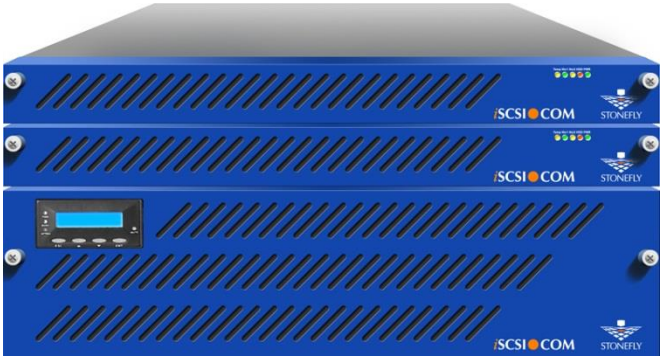
STONEFLY

Storage Concentrator User's Guide



StoneFly IP SAN, Scale Out NAS, Backup & Disaster Recovery, Hyper-Converged Storage & Server Appliances, Enterprise Cloud and Virtual Storage Appliances

This page is intentionally left blank.

Last update 11/30/2017.

This page is intentionally left blank.

TABLE OF CONTENTS

Change List			
Revision	Change Section/Description	Date	By
6.4.2.9 & 6.4.2.13	Added the Change List table	01-21-2011	B.J.
	Section 2.7.14 was updated to include description for Max Provisioned Space license.	01-25-2011	V.N.
	Fixed the name of section 4.3.	01-25-2011	V.N.
	Updated the contents of section 6.4.2; replaced contents of section 6.4.3; created new section 6.4.4.	01-25-2011	V.N.
	Updated section 1.3.5 – Active-Active operation section.	02-08-2011	B.J.
	Implemented check box "Start new replication at resume time" for replication schedule. Updated section 6.4.3.	02-16-2011	V.N
	Section 2.7.14 was updated to include description for Max Hosts license.	03-11-2011	V.N
	Customer Service was replaced with Customer Support.	03-11-2011	V.N
	Updated Appendix 8: StoneFly Storage Concentrator System Event Messages	03-11-2011	J.B.
	Updated the document date, TOC and version.	03-15-2011	B.J.
	Modified section 16.1.2 (Modifying Windows IO Timers), and removed the last check for the value.	04-11-2011	B.J.
	Modified section 2.7.14 to support Evaluation licenses for Max Hosts and Max Provisioned Space.	04-15-2011	V.N.
	Updated Software Upgrade section and moved it as an appendix.	05-13-2011	D.M., B.J.
	Limits section: Changed max # of "Snap enabled volumes" to "Snap enabled/Replicated volumes"	05-18-2011	B.J.
7.0.0.18	General update to include new NAS functionality.	01-04-2012	B.J., V.N., J.B.
7.0.1.5	NAS volume license is implemented.	02-17-2012	V.N.
	Updated Appendix 8: StoneFly Storage Concentrator System Event Messages.	02-29-2012	J.B.
7.0.1.9	Changes to Admin General page for new configurable options for NAS Volume CIFS User Policies. Changes to Resource Detail page for new Target Port Group field. Changes to the NAS Volume Configuration page for CIFS Account Unlock.	04-27-2012	J.B.

Revision	Change Section/Description	Date	By
7.0.2.4	Implement Scale Out NAS functionality.	08-02-2012	V.N.
	Change max number of host sessions from 200 to 1022.	08-02-2012	V.N.
	NAS volume renaming.	08-02-2012	V.N.
	Added DNS Server.	08-10-2012	J.B.
7.0.2.11	Update screen shots for NAS chapters 8 and 9.	29-01-2013	V.N.
	Set Copyright 2013.	29-01-2013	V.N.
	Added 2.7.24: NAS Server and Active Directory domain.	01-02-2013	V.N.
	Added 8.4: CIFS Users access.	01-02-2013	V.N.
	Updated Appendix 8: StoneFly Storage Concentrator System Event Messages.	02-02-2013	J.B.
8.0.0.x	Implement deduplication pools.	08-13-2013	V.N.
	Updated Appendix 8: StoneFly Storage Concentrator System Event Messages.	09-06-2013	J.B.
	Update 8.1: Deduplicated Pool Management, 8.2: Deduplicated Volume Status	09-13-2013	V.N.
	Set Copyright 2014.	01-21-2014	V.N.
	Update 2.4.1: Deduplicated Volumes and Pools.	01-21-2014	V.N.
	Update Chapter 8: Deduplicated Volumes Management.	01-21-2014	V.N.
	Added 2.4.1.7: Memory Limits for Deduplicated Volumes	01-23-2014	V.N.
	Update A4.2.7: Storage Concentrator Configuration Limits	01-23-2014	V.N.
	Updated Appendix 8: StoneFly Storage Concentrator System Event Messages.	01-28-2014	J.B.
	Update screen snapshots for NAS Volumes and NAS Segments Summary screens.	03-18-2014	V.N.
8.0.1.x	Set Copyright 2015, set version 8.0.1.x, and replace GUI screen snapshots. Added 9.10: NAS Volume Snapshots. Added Chapter 13: Flash Cache.	07-20-2015	V.N.
	Updated Appendix 8: StoneFly Storage Concentrator System Event Messages.	07-20-2015	J.B.
	Added 9.11: Directory Quotas	08-04-2015	V.N.
8.0.2.x	Updated GUI screenshots. Minor text corrections. Set copyright 2017, version 8.0.2x.	11-16-2017	A.M.
	Updated GUI screenshots. Added Appendix 1: Object Storage Resources. Updated section 10.3: Segment Migration. Added Section 10.4: Segment Replacement.	11-30-2017	A.M.

Chapter 1

Overview

1.1 Introduction

The StoneFly *Storage Concentrator*™ appliance software provides the interface through which a system administrator can configure and monitor the following:

- Logical volume management of all connected storage
- iSCSI interfaces that interconnect the host systems with the disk subsystems
- Upper level applications that provide an extensive suite of storage management functionality
- System operation and status
- FailOver protection
- Synchronous mirroring
- Snapshots
- Asynchronous replication
- Volume encryption
- Thin provisioning and deduplication
- Creating and managing NAS volumes that can be exported to NFS and CIFS clients

Most of these features are visible and can be operated only if the system has the appropriate license. For details see "Feature Licensing".

The following sections provide an overview of the StoneFly *Storage Concentrator* operating system that is offered in StoneFly's IP Storage Appliances, as well as StoneFly's Storage Concentrator Virtual Machine (SCVM™) a standalone virtual version that is compatible with VMware and Windows virtual and cloud environments. SCVM™ provides the same exact capabilities and functionalities available on the StoneFly's IP Storage Appliances, except that it would be running as a Virtual Machine on the user's environment.

1.2 Conventions

The tables that follow list the conventions used throughout this User's Guide.

1.2.1 Icons

Icon	Type	Description
	Note	Special instructions or information
	Warning	Risk of system damage or a loss of data

1.2.2 Text Conventions

Convention	Description
Boldface word	An action is required.
"Type" or "Enter"	Input the requested information

1.3 Storage Technologies

1.3.1 Storage Area Network (SAN)

Storage Area Network (SAN) is a separate and specialized network whose primary purpose is to transfer data between computer systems and storage elements. A SAN consists primarily of a communication infrastructure and a management layer. The communication infrastructure provides physical connections and the management layer organizes the connections, storage elements, and computer systems so that data transfer is secure and robust.

1.3.2 iSCSI

iSCSI is a protocol that enables the transmission of block-level SCSI data between storage resources and computers over a standard IP network. iSCSI combines Ethernet-based IP networking with the SCSI command set.

1.3.3 Network Attached Storage (NAS)

Network Attached Storage (NAS) is a storage technology that allows users to create shared volumes that can be accessed by one or more users. The difference between iSCSI volumes and NAS volumes are that NAS volumes (NAS shares) can be accessed by more than one user/server, but iSCSI volumes in general can only be accessed by one server.

1.3.4 Storage Provisioning

Storage Provisioning is the process of presenting a transparent, uniform, and logical representation of physical storage resources to storage clients (applications and users). Storage provisioning is not restricted by the type of storage device or medium, server platform, or connection methodology. Storage provisioning dynamically maps data from the logical storage space required by applications to the actual physical storage space.

1.3.5 Continuous Data Protection

Continuous Data Protection is an integrated data recovery solution that replicates not only files, but also databases and entire applications. An advanced Rewind technology provides a comprehensive undo capability that allows system administrators to instantly roll back damaged data resources to a previous, valid state. Because the Rewind is application aware, it can roll back one transaction at a time until the exact point of data loss.

1.3.6 Active-Active Operations

The StoneFly IP SANs that are built as a cluster such as Voyager DX, Voyager FC Plus, USO-HA, USO-FC, USS-HA, Voyager WX, USC-HA, UES-HA, UDS-HA, DR365-HA and DR365 Fusion

operate as an Active-Active cluster. Active-Active clustering in not available for standalone ISC, USO, USS, USC, DR365, Z-Series DR and M-Series DR class of products. In clustered configurations, two Storage Concentrators cooperate to achieve twice the possible bandwidth and system redundancy. Each system is capable of accepting the data traffic intended for one or more volumes. The activity level for each Storage Concentrator is set by the number of volumes assigned to each system. The Administrator balances this load by observing statistics related to IO traffic and then moving volumes between the two systems. A volume cannot be assigned to both Storage Concentrators at the same time.

1.3.7 FailOver

FailOver is an important fault tolerance function of mission-critical systems that require constant accessibility. FailOver adds a layer of redundancy to a storage network. FailOver is a feature of an Active-Active cluster (see above).

If a component of either cluster member fails, FailOver automatically reassigns all data traffic to the healthy Storage Concentrator. The system with failed components is then re-started to attempt a return to a healthy status. If the failure occurs in a hardware component it may not be possible to restore health by a restart. It may be necessary to replace the failed hardware component.

If the Secondary system returns to full health, the volumes assigned to the Secondary Storage Concentrator are redirected back to it from the Primary Storage Concentrator.

1.3.8 Mirroring

The StoneFly Synchronous Mirroring feature supplies host-independent, mirrored data storage that duplicates production data onto physically separate mirrored target images. This duplication occurs transparently to users, applications, databases, and host processors. The software duplicates block-level changes as they occur to one or more volumes: either to another local Storage Concentrator volume image or to another Storage Concentrator volume image at a campus location over standard IP connections.

StoneFly mirroring offers the benefit of protecting a critical volume from being a single point of failure, and providing continuous access to a volume without interruption to data availability when loss of access occurs to one of the images.

The StoneFly Asynchronous Mirroring feature allows data to be kept at two different locations. The local data volumes are synchronized with mirror volumes at the remote site at irregular intervals. The synchronization operations may be scheduled or allowed to happen at any time the local data volume falls quiet for a prescribed period of time.

1.3.9 Snapshot

StoneFly Snapshot creates virtual, temporary, and perishable point-in-time images of an active Live Volume. Snapshots contain a view of the volume at the exact point in time that the snapshot was taken. Snapshots can be created nearly instantaneously. The snapshot volume appears to the host as if it was a regular logical volume. Even after changes are

made to the original volume, Snapshots preserve a copy exactly as it existed when it was taken. Snapshots persist across reboots and can be mounted and accessed just like any other volume.

1.4 StoneFly Storage Appliances

1.4.1 Current Families of StoneFly Appliances:

Hardware Appliances – Cluster Solutions:

- Voyager DX IP SAN Appliance
- Voyager FC IP SAN Appliance
- USO-HA Unified Scale Out IP SAN + NAS Appliance
- USO-FC Unified Scale Out IP SAN + NAS Appliance
- USS-HA Unified Storage & Server Appliance
- USC-HA Unified Storage Concentrator SAN Gateway
- UES-HA Unified Encryption SAN Gateway
- UDS-HA Unified Deduplication SAN Gateway
- DR365-HA Backup & Disaster Recovery Appliance

Hardware Appliances – Standalone Solutions:

- ISC Integrated Storage Concentrator IP SAN Appliance
- USO Unified Scale Out IP SAN + NAS Appliance
- SSO Super Scale Out NAS Appliance
- TSO Twin Scale Out NAS Appliances
- USS Unified Storage & Server Hyper-Converged Appliance
- USC Unified Storage Concentrator SAN Gateway Appliance
- DR365 Backup & Disaster Recovery Appliance

SDUS - Software-Defined Unified Storage Solutions:

- SCVM – Storage Concentrator Virtual Machine (Virtual Storage Appliance)
- StoneFly Cloud Drive

Note: There are older generations of StoneFly appliances that are still supported with the latest StoneFusion software.

For description of each family and related datasheets, please refer to www.stonefly.com

All StoneFly appliances use a common Operating System called StoneFusion which is StoneFly's state of the art software that powers all StoneFly appliances. SCVM is a Software-Defined Unified Storage appliance, but other appliances are total IP Storage solutions combined with storage resources.

1.4.2 Storage Concentrator

The term Storage Concentrator is used throughout this document and is mainly the engine that delivers iSCSI, Fibre Channel and/or NAS storage volumes. Configuring and managing these volumes is accomplished using a browser-based graphical user interface (GUI) resident in the Storage Concentrator. The system administrator uses the graphical user interface to allocate storage to create iSCSI, Fibre Channel and/or NAS volumes and authorizes their use by individual host systems.

1.4.3 StoneFusion

The *Storage Concentrator* uses the proprietary StoneFusion™ Network Storage Platform that intelligently optimizes storage assets, offering the functionality traditionally associated with expensive midrange and high-end storage systems, and host-based volume management software. The StoneFusion architecture includes:

- An extensible in-band metadata storage-mapping layer
- Intelligent iSCSI storage packet routing software, providing aggregation and bi-directional data transfer for increased throughput
- A metadata database that tracks physical data locations to ensure data integrity
- Online storage management to easily consolidate free storage space maximizing storage resources

1.4.4 Administrative Interface

The administrative interface is accessed from a computer on the network via a web browser. The following management functions are available through the interface:

- **Volumes:** Create, manage, and delete storage volumes from any available resource.
- **Hosts:** Create, edit, and delete hosts.
- **Sessions:** Monitor the number of connections between hosts and volumes.
- **Resources:** Discover resources (devices) on the iSCSI bus that provide the space for storage volumes.
- **NAS:** Create, manage, and delete NAS volumes.
- **System:** The System Management screen provides information regarding the system setup, the configuration of the *Storage Concentrator* network settings, and FailOver settings.
- **Users:** Add, edit, delete, or view user information.
- **Reports:** The reporting function provides statistical information on the *Storage Concentrator* device, resources, volumes, and sessions. System logs provide a system-level sequence of events.

Figure 1-1 Storage Concentrator Administrative Interface Overview

1.5 Typical Configurations

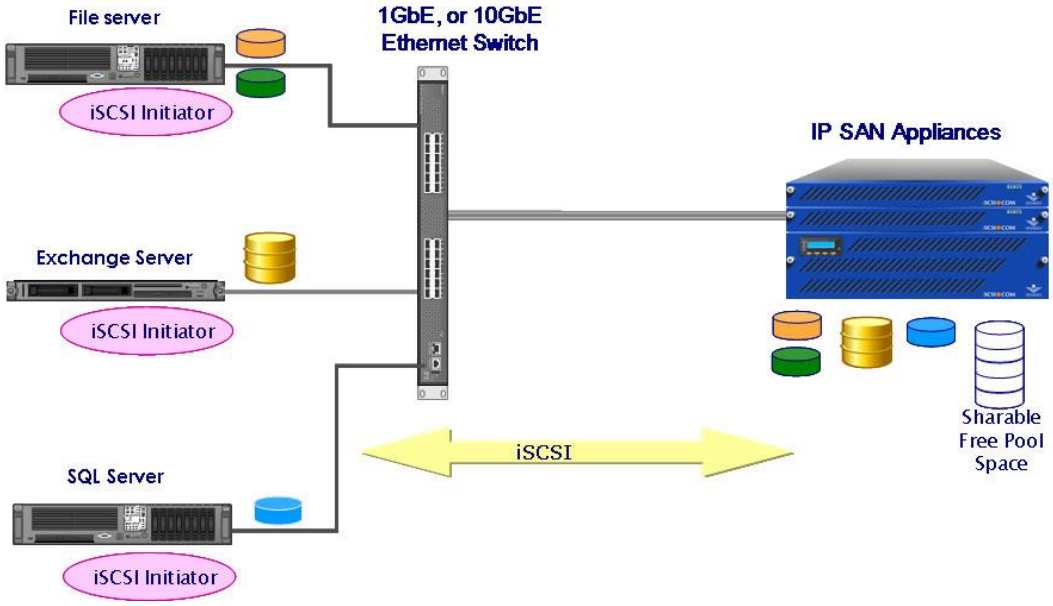


Figure 1-2 Typical iSCSI environment

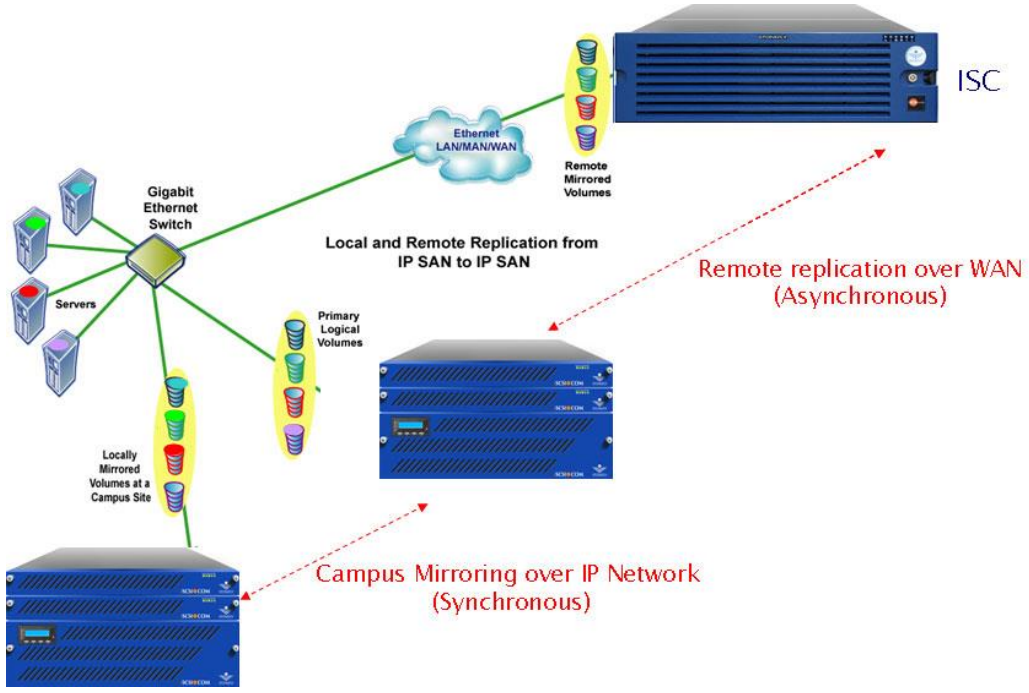


Figure 1-3 Example of a Campus (Synchronous Mirror), and Remote Replication (Asynchronous Mirror).

Chapter 2

Administrative Interface

2.1 Launching the Administrative Interface

The administrative interface resides on the *Storage Concentrator*. It is run from a network computer via a browser which can access the network where the *Storage Concentrator* is located.

Supported browsers include:

- Netscape 4.7 or later (Windows PC and Linux)
- Internet Explorer 5.0 or later (Windows PC only)
- Mozilla Firefox

To access the administrative interface, use the steps that follow:

1 Launch your web browser.

2 Type the IP address for the Management Port of the *Storage Concentrator* in the address field of the browser. For more information, refer to the Setup Guide, "Configuring the Network Settings."

The address field in the browser must include **https://** to access the administrative interface.

The following login screen appears.

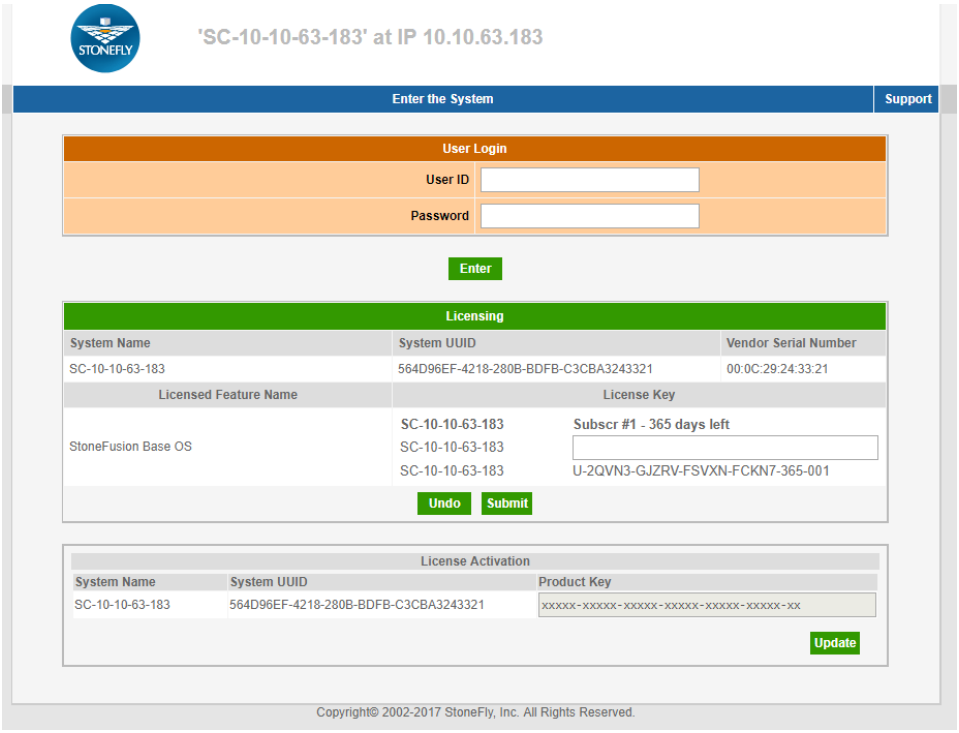


Figure 2-1 Storage Concentrator Login Screen

2.1.1 Accessing System Management Functions

To perform system management functions, you must log in with a user ID and password that is configured with administrative privileges.

Passwords must be between 6 to 15 alphanumeric characters with one character being a numeric character. Passwords cannot be the same as the corresponding username. StoneFly validates only the first eight characters during login. Characters after 8 are ignored, but supported for user convenience. For more information on creating users, see "Adding Users".

When logging in for the first time, the administrator can use the following: User ID: **stonefly**. Password: **stonefly**. It is strongly recommended that the stonefly password be changed at the initial configuration. It is also recommended that each system administrator have an individual user ID and password. For more information on creating users, see "Adding Users".

1 Enter your **User ID.**

2 Enter your **Password**.

3 Press **Enter** or click **Submit**. The *Storage Concentrator* home screen appears.

If you attempt to log in with an administrator user ID and password that are in use, a warning message dialog box appears.

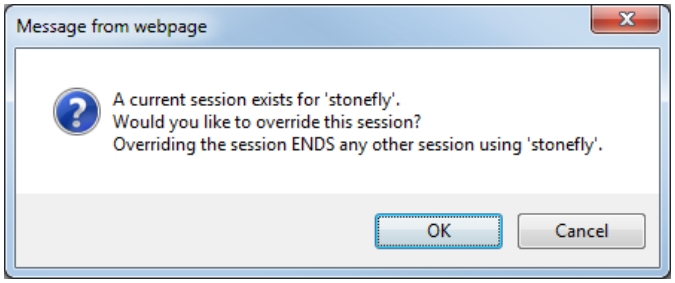


Figure 2-2 The "current session exists" Warning Dialog Box

If this screen appears, choose one of the following options:

1 Click OK - This logs you in and logs off the user who is logged in with the user ID and password you entered.

2 Click Cancel - The initial login screen appears. Repeat steps 1-3 using a different user ID and password.

2.2 Home Page

After successfully logging into the Storage Concentrator, the home page will appear.

Figure 2-3 The Storage Concentrator Home Page (Failover Cluster is pictured here)

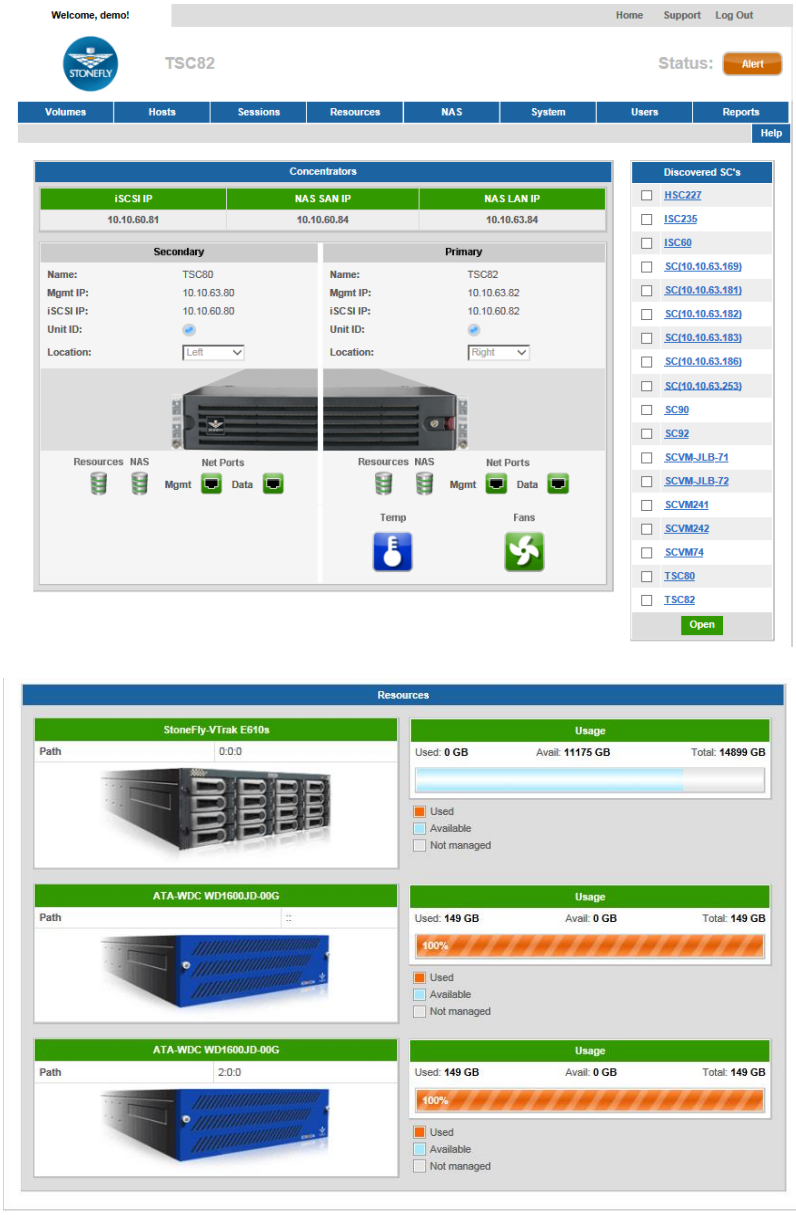


Figure 2-4 The Voyager-TSC Storage Concentrator Home Page has a different layout, and is pictured here.

The bar across the top of the administrative interface allows the system administrator to access information regarding system information, support, and system status.

- **Home**: Displays this home page with a summary of all system information
- **Support**: Displays StoneFly support information
- **Status**: Displays the status of the *Storage Concentrator*:
  - **Good**: Indicates normal operation
  - **Down**: Indicates that the *Storage Concentrator* is not running correctly and must have the power cycled to recover
  - **Alert**: Indicates that unacknowledged critical messages have been detected and

have been written to the system log

- **Logout**: logs out the current session.

The *Storage Concentrator* can be configured to notify the system administrator when Alerts occur. For more information on acknowledging Alerts in the system log, see "Logs". For more information on Alert notification, see "Notifications".

2.2.1 Home Page Images

Primary — The Primary Storage Concentrator is currently managing the storage volumes. Clicking on the photo image of the unit takes you to the Diagnostics Page. The Management IP is the address of the management port. The iSCSI IP is the address of the IPSAN port(s).

If a Storage Concentrator FailOver Cluster is configured for this Storage Concentrator, there is also a Cluster IP address displayed which is the master address of the FailOver Cluster, including both the Primary and Secondary units. For more information on FailOver Clusters, see "Setting up FailOver".

The **Unit ID** (Unit Identification) button can be used to cause a blue LED to visually identify the system, usually by blinking. The **Unit ID** button will not be displayed on systems that do not support the **Unit ID** function, or are currently inaccessible.

The Voyager-TSC chassis supports two systems in the same chassis. The right/left location is configured using the **Location** control. The **Unit ID** button should be used to determine or confirm the chassis location setting.

Secondary — Displays an image of the Secondary Storage Concentrator if the unit is in a cluster. If no unit is shown, there currently is no Secondary unit. Clicking on the Secondary unit takes you to the Diagnostics page. The Management IP is the address of the management port. The iSCSI IP is the address of the Gigabit Ethernet port(s)

Resource Status – The Resource icon indicates the status of the items listed on the Resource screen. If any Resource is not accessible the icon will appear Red instead of Green. Failovers are performed if the icon is Red for the Primary SC.

Temperature Indicator— The thermometer indicator displays the Primary Storage Concentrator's temperature status. If the thermometer is red, there is an alert regarding temperature on the diagnostics page. If it is green, the unit is OK. Clicking on the thermometer icon takes you to the Diagnostics page.

Fan Status— The fan indicator displays the Primary Storage Concentrator's fan status. If the fan is red and stationary, there is an alert regarding one or more of the fans on the diagnostics page. If it is green all the fans are operating within normal limits. Clicking on the fan icon takes you to the Diagnostics page.

Power Supply Status— The power supply indicator displays the Primary Storage Concentrator's power supplies and power related sensors.

If the power icon is red, there is an error regarding one or more of the power sensors on the

diagnostics page. If it is green, all are operating normally.

Clicking on the power icon takes you to the Diagnostics page. This icon is not shown on systems where there is no HW sensor support.

When UPS Management is enabled, the overall UPS status is also indicated here, with detailed status on the UPS Management page.

Port Status— For FailOver Clusters ONLY. The port icons display status of either the Primary or Secondary Storage Concentrator ports. If any of the port indicators are red, there is an Alert regarding the port listed on the diagnostics page. If all the ports are green, then all ports are OK. Clicking on any port icon will take you to the Diagnostics page.

Resources— For FailOver Clusters ONLY. Each Storage Concentrator resource that has been discovered is displayed with an icon on the home page. Clicking on the resource takes you to the Resource Management Summary page. The resource icon is shown with its name, including Manufacturer and Model number, and path, which is the SCSI path to the device.

Used and available space for the resource is displayed in the display bar to the right of the resource. If the display is blue, the space has been used and displays the amount in GB used. Space that has been allocated, but not used is displayed as green and is shown as available. Any resource that is not managed by the Storage Concentrator (pass through or none) shows as yellow.

Discovered SC's— Displays a list of other Storage Concentrators that have been discovered by this Storage Concentrator.

Clicking on the name of the SC will open a new browser window for that system.

Hovering over the link will display information about that system.

If windows for more than one SC are needed, select their check-boxes and click Open.

SC Discovery is achieved through the use of the Service Location Protocol (SLP) which uses broadcast and multicast IP frames over the management network.
SC Discovery is enabled by default. However, should the installation or network environment be such that this feature is unnecessary, or undesired, it may be disabled on the System Admin GUI page.

Note: Only Storage Concentrators that are operational are discovered and displayed. SC's running a SW version that does not support SC Discovery will not be detected. SC's that have the SC Discovery feature disabled will not be discovered.

When the SC is operating in a cluster, a similar "**Discovered SC's**" table will appear on the secondary system's management screen.

If the "Advanced Features" not licensed, or the SC discovery is disabled, or if there are no other SC's discovered, the "**Discovered SC's**" table does not appear.

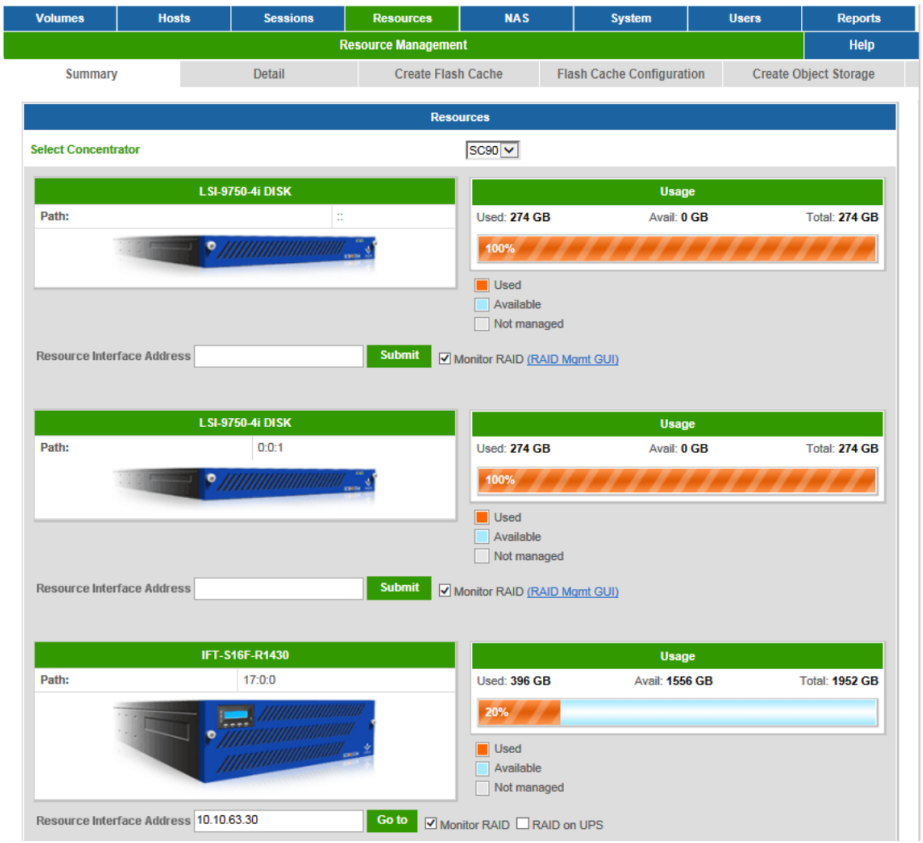
2.2.2 Master Storage Concentrator Menu

The main menu for the Storage Concentrator is displayed on the left hand side of the screen. The picture above shows the menu in its Nested Menus format. View the detail screen for each user to set this feature to "On". The following buttons are active:

- **Volumes** — Click to access volume management functions.
- **Hosts** — Click to access host management functions.
- **Sessions** — Click to access session management functions.
- **Resources** — Click to access resource management functions.
- **NAS** – Click to access NAS volume management functions.
- **System** — Click to access system management functions.
- **Users** — Click to access user management functions.
- **Reports** — Click to access reporting functions.

2.3 Resources

After initial configuration is completed, the next step is to discover resources of the IP Storage. A resource device is a data storage subsystem. Resource devices may include RAID drives and JBODs (just a bunch of disks). A resource device is typically attached directly to the Storage Concentrator.



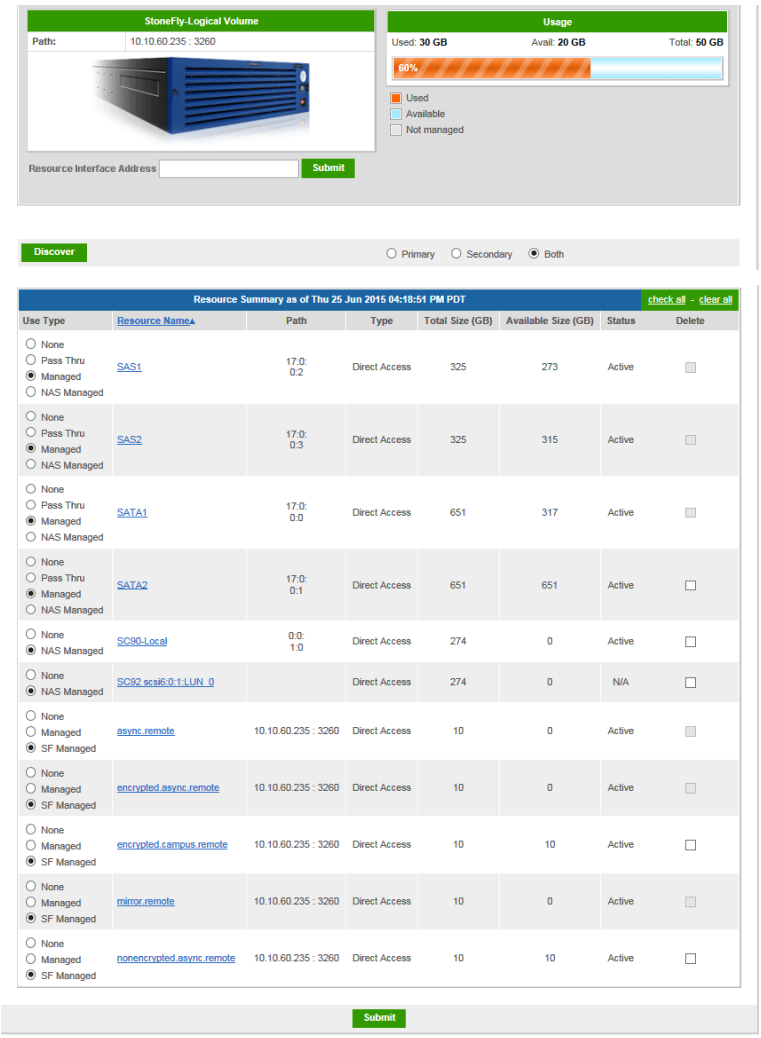


Figure 2-5 Resource Management Summary Screen

If a Failover cluster is configured, a summary page for the Secondary unit can be displayed by selecting it from the pull down list on the Resource Management Summary screen. Resources that cannot be seen by both members of a cluster are marked with a red asterisk.

If possible, have all storage resources configured prior to allocating volumes.

2.3.1 Discovering Resources

Discovering resources is the process the Storage Concentrator uses to query the resource device for information. At start up, the Storage Concentrator automatically discovers any resources that are attached.

To add resources using the Discover button on the Resource Management Summary Screen, use the steps that follow:

1. Click **Resources**. Any previously discovered resources display on the Resource

Summary screen.

2. Select the radio button to discover resources for the Primary unit, Secondary unit, or both (the default):

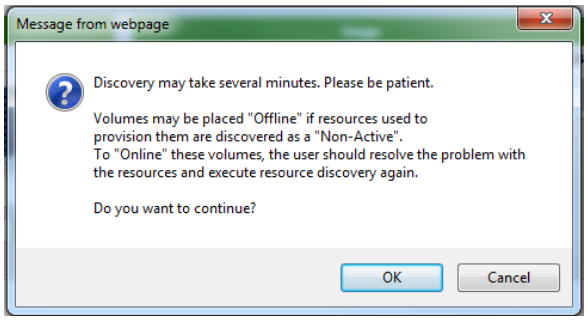
   Primary Unit: Selecting the Primary Unit will only discover resources for the Primary Unit. Since both the Primary Unit and the Secondary Unit are resource aware, this will reduce the time required to discover resources.

   Secondary Unit: Selecting the Secondary Unit will only discover resources for the Secondary Unit.

   Both: (default) Selecting both will discover resources for the Primary Unit and the Secondary Unit.

   Select either the Secondary Unit or the Primary Unit if you have a problem (such as a cable issue) on just one unit. Or select one or the other when there are a lot of resources to discover and you want to prevent the discovery process from timing out by discovering on the Primary Unit first and then the Secondary Unit. Generally, using both will be the preferred discovery method in most cases.

3. Click **Discover**. The following dialog box appears.



4. Click **OK**.

2.3.2 "Use" Types for Resources

When discovery is complete, all available resources will appear for the Primary Unit. The display will include direct-connected devices and any iSCSI resources.

Once discovered, each resource must be assigned a use type. The *Storage Concentrator* supports four use types:

None: This resource is not currently managed by the *Storage Concentrator.*

Pass Thru: The resource will accept SCSI commands from a host without any intervention from the *Storage Concentrator* except Reservation and Persistent Reservation SCSI commands. For Pass Thru volumes these commands are handled by the *Storage Concentrator.* This resource is not added to the storage pool.

Managed: The resource will be managed by the *Storage Concentrator*

Flash Cache: The resource is for use as a caching device for a resource managed by the

Storage Concentrator

NAS Managed: The resource is for use in a direct NAS Segment for NAS Volumes

SF Managed: The resource is available as a Campus Mirror resource through a secondary Storage Concentrator or as an Asynchronous Mirror through a remote Storage Concentrator.

The only available use types for iSCSI resources are "None", "SF Managed" and "Managed". "Pass Thru" is not displayed for iSCSI resources.

The pool of "Managed" resources can include local resources and remote resources at the same time. Local resources are attached to the Storage Concentrator directly through SAS, SCSI or FC HBA's. Access to remote resources is provided through iSCSI target portals. It's not recommended to split synchronous image provisioning between local and remote resources. The same precaution applies when a spanned volume is provisioned or snap space is allocated.

The Storage Concentrators use a "Managed" resource space to keep metadata so that the resource size available for volume provision is less than the total size of the resource.

Information stored on remote volumes will be lost after the volume is set up as a "Managed" resource. Don't let iSCSI hosts access the volume directly. The volume has to be used exclusively by the system that set up it as a "Managed" resource. When the use type for a remote "Managed" resource is changed to "None" or to "SF_Managed", then information on the resource becomes invalid.

To select the use type for a resource, use the steps that follow:

1 Click one of the following buttons:

- **None**
- **Pass Thru**
- **Managed**
- **Flash Cache**
- **NAS Managed**
- **SF Managed**

Note that only the Use Type choices valid for each resource will be shown.

2 Click **Submit**. The pass thru and managed resources are now available and storage volumes can now be created.

If you select SF Managed for a Campus resource or for a Remote resource that has not been discovered by both the Primary and Secondary Units, you will receive an Alert message. To correct the Alert, discover the resource for both the Primary Unit and the Secondary Unit. See Expanding a Mirror Volume.

Resource summary information in a clustered system will vary depending on which *Storage Concentrator* you are viewing. Resource summary information includes:

- **Use Type**: Indicates how the resource is being managed
- **Resource Interface Address** — This field holds the URL used to connect to the external RAID device for configuration and maintenance.

  When the resource is first discovered, this field is blank. It can be populated with an IP address, or a URL such as http://ip_address, http://ip_address:port, https://ip_address, or telnet://ip_address.

  When this field is changed, either hit Enter, or use the Submit button to the right. A new browser window will open using the URL provided.

  When the field is not blank, the button label changes to Go to which can be used to open a browser window to manage the RAID.

  Internal RAIDs do not require the Resource Interface Address to be set to manage them -- the RAID Mgmt GUI link to the right may be used instead.

  On RAID systems that are recognized by the SC, the RAID's published default login user name and password is shown when the mouse hovers over the Go to button or the RAID Mgmt GUI link. The RAID default login user name and password are not shown to non-administrative SC GUI users, nor when the SC has been configured with a specific RAID user name or password.
- **Monitor RAID** — The Monitor RAID check-box shows the current state of RAID Monitoring by the Storage Concentrator. If the check-box is checked, the RAID is being monitored.

  When a RAID is being monitored, the Storage Concentrator collects RAID event log information and presents this in its own logs. It polls the RAID periodically recording any new events.

  A * indicator and footnote indicates when one or more RAIDs have logged critical errors. The SC Event Log, and RAID Management GUI should be used to resolve the problem. The * indicator will only be cleared when the RAID events are flagged as acknowledged in the SC Log.

  Whether or not a RAID should be monitored is controlled by changing the check-box. A communication and interaction test is made at the time the check-box is checked, and the result reported to the user.

  Note that external RAIDs require that the Resource Interface Address be set before monitoring is possible; otherwise the check-box is disabled.

  Internal RAIDs can also be monitored, and do not require that the Resource Interface Address be set.

Also note that not all RAID vendors and models are supported for monitoring. The Monitor RAID check-box and RAID Management Password fields are hidden on devices that are known to not be supported.

Whether or not the RAID that is providing the resource storage is actually supported will be evident through the results of the communication and interaction test.

If the RAID password has been changed from the default, it may be set on the Resource Detail page. Otherwise the factory default RAID password is used.

- **RAID on UPS** — The RAID on UPS check-box indicates whether or not the RAID is connected to the UPS that is managed by the Storage Concentrator.

  When the SC is managing a UPS, and this check-box is checked, it is assumed that the RAID is also powered by the same UPS.

  Should a UPS low battery shutdown occur, the RAID will be shut down at the same time that the SC is. This causes any I/Os cached in the RAID to be flushed to disk before power is lost. It also prevents any new I/O's until the RAID is power cycled by the UPS.

  Only RAIDs that are being monitored as indicated by the Monitor RAID check-box are able to be shutdown. As not all RAID vendors and models are supported for monitoring, such unsupported RAIDs would have their RAID on UPS check-box hidden. Otherwise, the RAID on UPS check-box is disabled until Monitor RAID is enabled.

  If the SC is not configured to manage its UPS, or if the RAID is not being monitored, the RAID on UPS has no effect and will hidden.

  Internal RAIDs are always shutdown with the SC when a UPS low battery shutdown occurs. There is no configuration needed for internal RAIDs.

  When the RAID is not powered by any UPS, or by a different UPS than the one that the SC is managing, the RAID on UPS check-box should not be checked. Otherwise, the RAID would be shut down when it might still be in use, and it would require manual intervention for it to be restarted. This is especially important when the RAID is being shared with other systems.
- **Resource Name**: A default name assigned to the resource by the *Storage Concentrator*. The resource name can be edited later. For more information, see "Editing Resources".
- **HBA: bus: Target ID: LUN**: The SCSI address for the resource
- **Type**: Type of SCSI storage device. This information is provided by the resource
- **Block Size**: Number of bytes in each block
- **Total Size (GB)**: Total number of gigabytes on the resource
- **Available Size (GB)**: Number of gigabytes of space available for new volumes
- **Status**: A resource can have one of the following status:
  - **Active**: Discovered and selected to be managed
  - **Not managed**: Discovered but not selected to be managed by the *Storage Concentrator*
  - **Non-active**: Has been discovered in the past, but was not found with the most recent

discovery attempt

- **Off-line**: The *Storage Concentrator* is not able to successfully initiate a session with this resource

A resource can be removed from the Storage Concentrator at any time by clicking Delete next to the resource to be removed. For more information, see "Removing Resources".

A red asterisk indicates that the resource status is in question and needs to be investigated by the system administrator.

2.3.3 Adding Resources

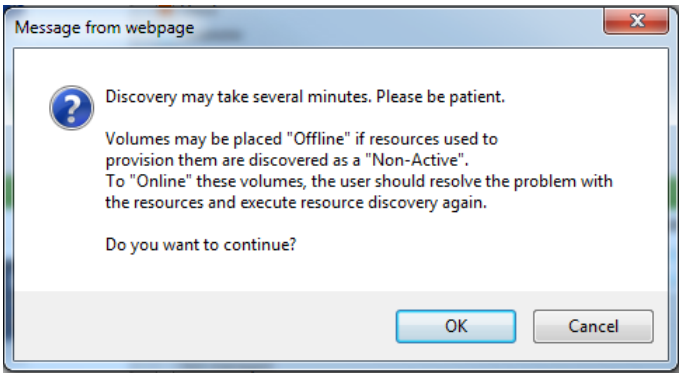
As storage needs change and grow, additional storage resources may be added to accommodate your storage requirements. Discovering resources is the process the Storage Concentrator uses to query the resource device for information.

You can use the Discover button on the Resource Management Summary screen to ADD storage resources without rebooting the Storage Concentrator. Existing resources may not be modified.

Storage resources can be added without rebooting the *Storage Concentrator* when the *Storage Concentrator* is connected to a disk enclosure that has empty slots available for additional hot--swappable drives, and additional drive(s) are inserted into the enclosure. When the **Discover** button on the Resource Management summary screen is selected, the disk drives are added as storage resources and are available for use by the *Storage Concentrator*.

To add resources using the Discover button on the Resource Management Summary Screen, use the steps that follow:

1. Select the radio button to discover resources for the Primary unit, Secondary unit, or both (the default).
2. Click **Discover**. The following dialog box appears.



3. Click **OK**.

When discovery is complete, all available resources will appear for the Primary Unit regardless of which radio button has been selected.

4. Click the **Managed** button for each resource you want the *Storage Concentrator* to manage.
5. When all resources to be managed by the *Storage Concentrator* have been selected, click **Submit**.

These resources are now available and storage volumes can be created.

To add storage that requires rebooting the *Storage Concentrator*, use the steps that follow:

1. Shut down the Storage Concentrator. For more information, see "Shutting Down". If hosts are currently online, shut them down.
2. Configure the storage resource device.
3. Power on the Storage Concentrator.
4. After the resources are configured on the Storage Concentrator, restart the host computers.

The *Storage Concentrator* automatically discovers any newly configured resources and rediscovers any existing resources.

2.3.4 Editing Resources

The resource name is the only setting that can be modified after initial discovery.

2.3.4.1 To edit the resource name, use the steps that follow:

1 Click **Resources**

2 Click on the **Resource Name** to be edited or click on **Details** on the Resource Management Summary screen. The Resource Details screen appears.

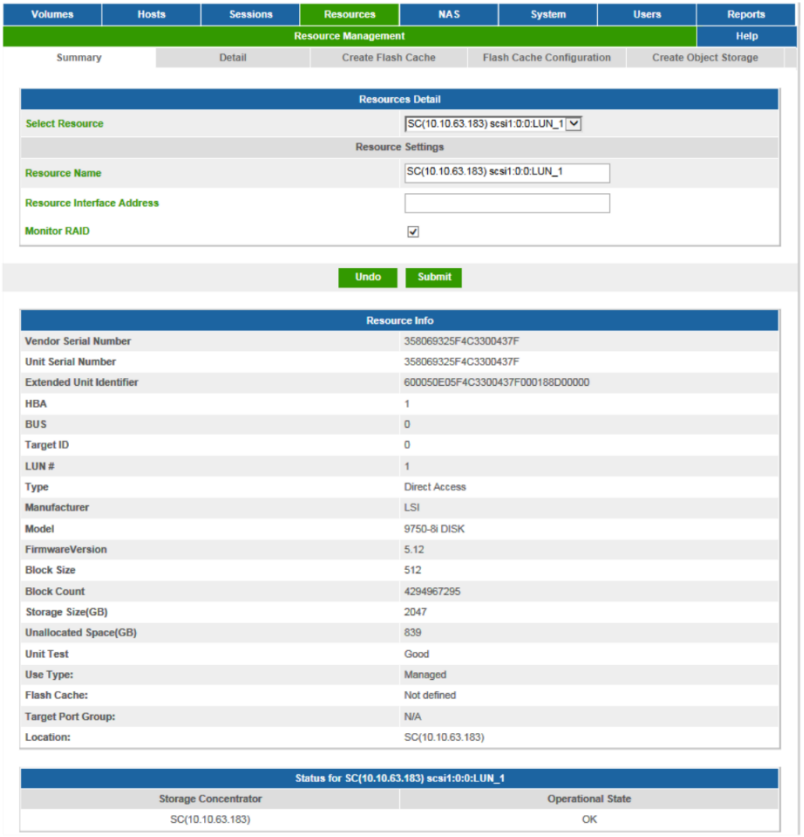
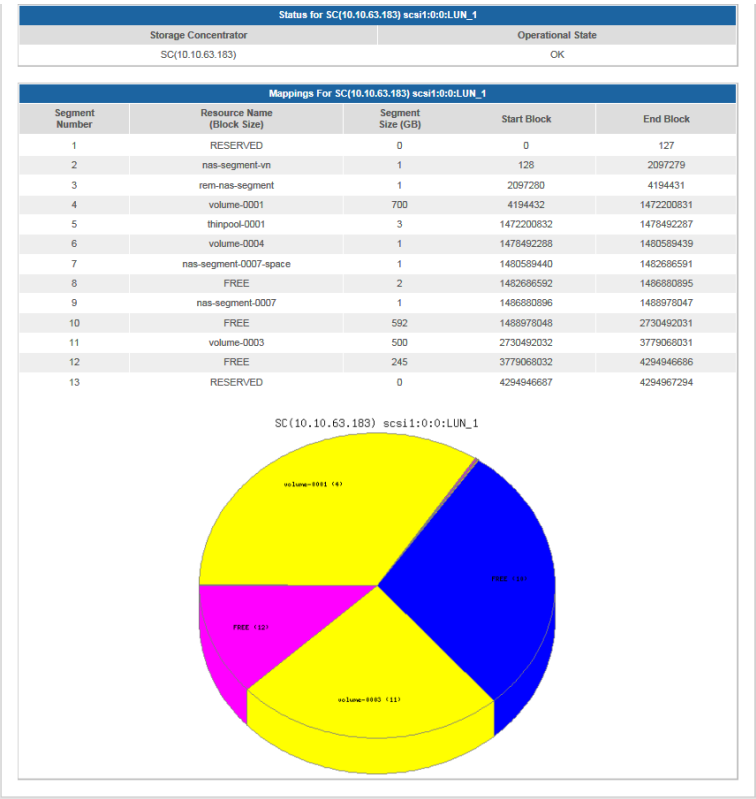

Volumes
Hosts
Sessions
Resources
NAS
System
Users
Reports
Resource Management
Help
Summary
Detail
Create Flash Cache
Flash Cache Configuration
Create Object Storage
Resources Detail
Select Resource
SC(10.10.63.183) scsi1:0:0:LUN_1
Resource Settings
Resource Name
SC(10.10.63.183) scsi1:0:0:LUN_1
Resource Interface Address
Monitor RAID
Undo
Submit
Resource Info
Vendor Serial Number 358069325F4C3300437F
Unit Serial Number 358069325F4C3300437F
Extended Unit Identifier 600050E05F4C3300437F000188D00000
HBA 1
BUS 0
Target ID 0
LUN # 1
Type Direct Access
Manufacturer LSI
Model 9750-8i DISK
FirmwareVersion 5.12
Block Size 512
Block Count 4294967295
Storage Size(GB) 2047
Unallocated Space(GB) 839
Unit Test Good
Use Type: Managed
Flash Cache: Not defined
Target Port Group: N/A
Location: SC(10.10.63.183)
Status for SC(10.10.63.183) scsi1:0:0:LUN_1
Storage Concentrator
Operational State
SC(10.10.63.183)
OK

Figure 2-6



Status for SC(10.10.63.183) scsi1:0:0:LUN_1

Storage Concentrator	Operational State
SC(10.10.63.183)	OK

Mappings For SC(10.10.63.183) scsi1:0:0:LUN_1

Segment Number	Resource Name (Block Size)	Segment Size (GB)	Start Block	End Block
1	RESERVED	0	0	127
2	nas-segment-vn	1	128	2097279
3	rem-nas-segment	1	2097280	4194431
4	volume-0001	700	4194432	1472200831
5	thinpool-0001	3	1472200832	1478492287
6	volume-0004	1	1478492288	1480589439
7	nas-segment-0007-space	1	1480589440	1482686591
8	FREE	2	1482686592	1486880895
9	nas-segment-0007	1	1486880896	1488978047
10	FREE	592	1488978048	2730492031
11	volume-0003	500	2730492032	3779068031
12	FREE	245	3779068032	4294946686
13	RESERVED	0	4294946687	4294967294

Figure 2-7

Resource Details Screen

3 Edit the resource name. Click **Undo** to revert to the saved settings.

4 Click Submit to save the changes.

The database tables in the Storage Concentrator are updated with the new resource name.

Resource detail information includes:

Serial Number: Serial number assigned by the manufacturer

Extended Unit Identifier: A unique ID number provided by the manufacturer. There are many ways this number is generated, including the MAC address, serial number, model number, etc. StoneFly merely displays the EUI number.

Bus: SCSI information

HBA: SCSI information

Target ID: SCSI information

Type: Type of SCSI storage device

LUN #: Logical unit number assigned to the device

Target Portal: Target portal address (if applicable)
If the resource has a target portal address, it does not display HBA, BUS, or LUN and instead displays the target portal address.

Manufacturer: Name of the manufacturer of the resource device

Model: Model number assigned by the manufacturer

Firmware Version: Current version of the firmware running on the resource

Block Size: Number of bytes in each block

Block Count: Number of blocks on the resource

Storage Size (GB): Total number of gigabytes of storage space on the resource

Unallocated Space (GB): Number of gigabytes of space available for new volumes

Unit Test: Reflects the manufacturer specific, SCSI test unit ready information

Use Type: Managed, SF Managed, Pass Thru, NAS Managed, Flash Cache or None

Flash Cache: Status of flash cache device if assigned to the resource

Target Port Group*:* When available, this number indicates which controller of a dual controller RAID the SC is attached to. When not available, N/A is displayed. This field is suppressed for resources provided by an iSCSI target portal.

Operational State: A resource can have one of the following states:

OK: The Storage Concentrator and the resource have initiated a session

Off-Line: The Storage Concentrator is not able to successfully initiate a session with this resource

Ready: On pass thru resources only, this state displays if there is currently no active front-end session between this resource and a host

Mappings for the selected resource: Refers to the physical locations (blocks) of the volumes on the resource. Resources that have a pass thru use type, display the volume name only when a volume is associated with the pass thru resource.

Segment number: A location on the resource

Resource name (block size): Name assigned to the resource

Segment size (GB): Total number of gigabytes allocated to this segment from this particular resource

Start Block: Block number where the volume segment starts on the resource.

End Block: Block number where the volume segment ends on the resource.

Chart of allocated space and free space within the resource: A pie chart that identifies the name of the volume, free space and the resource's segment number. Each segment is assigned a different color.

2.3.5 RAID Monitoring

The Storage Concentrator can monitor certain supported internal/external RAID subsystem's event logs, and report their events using the SC logging and event notification system. This provides a single place to monitor the systems.

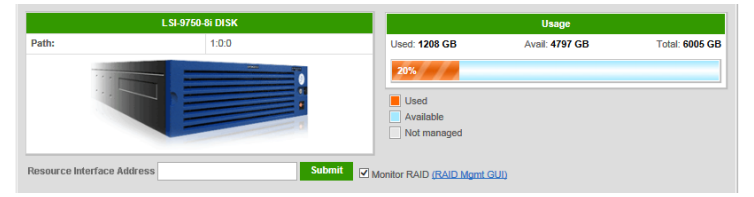


Figure 2-8 Resource Summary Page -- RAID Supporting RAID Monitoring.

Monitor RAID —The *Monitor RAID* check-box shows the current state of RAID Monitoring by the Storage Concentrator. If the check-box is checked, the RAID is being monitored. When a RAID is being monitored, the Storage Concentrator collects RAID event log information and presents this in its own logs. It polls the RAID periodically recording any new events. A * indicator and footnote indicates when one or more RAIDs have logged critical errors. The SC Event Log and RAID Management GUI should be used to resolve the problem. The * indicator will only be cleared when the RAID events are flagged as acknowledged in the SC Log. Whether or not a RAID should be monitored is controlled by changing the check-box. A communication and interaction test is made at the time the check-box is checked, and the result reported to the user.

Note that external RAID subsystems require that the *Resource Interface Address* be set before monitoring will be possible, otherwise the check-box is disabled. Internal RAID arrays can also be monitored, and do not require that the *Resource Interface Address* be set. Also note that not all RAID vendors and models are supported for monitoring. The *Monitor RAID* check-box and *RAID Management Password* fields are hidden on devices that are known to not be supported.

Whether or not the RAID that is providing the resource storage is actually supported will be evident through the results of the communication and interaction test.

If the RAID password has been changed from the default, it may be set on the *Resource Detail* page. Otherwise the factory default RAID password is used.

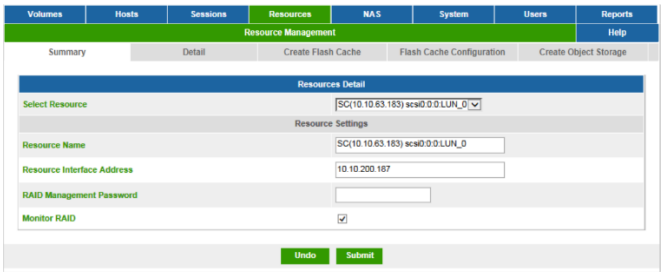


Figure 2-9 Resource Detail Page -- RAID Supporting RAID Monitoring.

Resource Interface Address —This field holds the URL used to connect to the external RAID device for configuration and maintenance.
When the resource is first discovered, this field is blank. It can be populated with an IP

address, or a URL such as ***http://ip_address***, ***http://ip_address:port***, ***https://ip_address***, or ***telnet://ip_address***.

Internal RAID arrays do not require the ***Resource Interface Address*** to be set in order to manage them.

RAID Management Password — When ***RAID Monitoring*** is enabled, this field holds the RAID management password that is used to login to the RAID.
Note: If the password field is left blank, the factory default RAID password is used.

2.3.6 Removing Resources

A resource device will need to be removed when the device is no longer available as a source for storage volume allocation.

To remove a resource, use the steps that follow:

1. Click **Resources**. The Resource Summary screen appears.
2. Click the check box under **Delete** for the resource to be removed.
3. Click Submit to remove the resource. The database tables in the Storage Concentrator are updated with the new status.

A resource that has a volume associated with it cannot be removed. If a volume is associated with the resource selected for removal, a warning dialog box appears and will prevent the resource from being deleted.

Click **Select All** if you wish to select all the resources that are listed. Click **Clear All** to deselect any selected resources.

2.4 Volumes and Security

2.4.1 Volumes

A volume is an arbitrarily sized space on one physical resource or one that spans multiple physical resources. Volumes behave like a disk block device. Volumes can be used as an address space for a file system or as a device swap space.

2.4.1.1 Access Control Lists (ACLs) and CHAP

An access control list (ACL) is a list that controls which hosts have access to which volumes. When a host attempts to access a volume for the first time, the *Storage Concentrator* allows access based on the current ACL. CHAP specifies a secret known only to the proper host and the *Storage Concentrator* to prevent unauthorized access to volumes.

There are two parts to volume management: creating the volume and setting security for the volume.

2.4.1.2 Overview of Encrypted Volumes

A new volume may be created as a non-encrypted or encrypted volume depending on its intended use. The decision to create an encrypted volume can be done only at the time it is created. The decision cannot be made after the volume is in use. An encrypted volume is created as a simple volume. An encrypted volume may be snapshot-enabled or mirrored later.

Managed and SF_Managed resources can be used to mirror encrypted volume. The encryption feature can be used on standalone or clustered Storage Concentrators.

When the volume is created the Administrator must enter a password. This password is used to determine the encryption key for the volume. Each volume should have a unique password that will generate a unique encryption key. The information about the password is not kept inside the Storage Concentrator. The passwords for all the volumes on a specific Storage Concentrator are stored on a USB memory drive. The USB drive is uniquely identified for its specific SC and volumes. There are two times when the USB drive must be inserted into the SC:

1. When system is standalone and the Storage Concentrator is rebooted or when the system is clustered and one Storage Concentrator is rebooted and the other one is still down.
2. When a volume will be created or deleted.

Except for these times, the USB drive should be removed and stored in a safe place. The Administrator should keep the information regarding volume passwords also in a safe place, preferably at a different location than the USB drive. The contents of the USB drive may be duplicated or saved on other types of media. The physical security of volume data is compromised while the USB drive is plugged into the SC. The end-user can set up the scheduler to check for the presence of the USB drive. If a drive is found, the scheduler will generate a log message with a critical severity level. The message can be retransmitted thru E-mail or SNMP notification.

Keep the USB drive safe but accessible in the case there are any problems that will generate an

automatic reboot of the standalone Storage Concentrator. If the standalone SC reboots without the USB drive inserted all encrypted volumes will be set Offline and will not be available to the servers in your IPSAN. The USB drive must be inserted and then the volumes must be set online. It can be done for each encrypted volume separately by using Volume Configuration or for all encrypted volumes simultaneously by using the "Online" button on the Volume Summary page. Refer to the section on Volume Configuration for instructions on bringing volumes online. The same procedure has to be executed to make encrypted volumes online in case of clustered Storage Concentrators.

The status of the USB drive is reported on the System->Admin screen. Each report indicates either a good situation or a problem. A good report is that the USB drive is not present. A USB drive is also allowed as the storage for the SC database files. If this database-restore USB drive is the only USB drive at boot time the system will not use it for encryption keys and the USB status will indicate this fact. The Volume->Summary screen tracks the encryption status of each volume.

Should the USB encryption drive be lost or damaged it may be restored if the passwords for each volume are known. Refer to the section on "Restoring the Encryption Keys to your USB Drive" at the end of the section on Volume Security.

Encrypted volumes created with version prior 6.3.2 don't support campus and asynchronous images. Starting 6.3.2 the newly created encrypted volumes can be mirrored by using appropriate SF_Managed resources. The remote resource has to be provisioned on the remote system as an encrypted volume with the same password that the local system uses for the local image.

2.4.1.3 System Metadata Volume

The "system-metadata" volume is a special volume that keeps internal data that is generated at run time during some operations. Currently metadata is generated for synchronously mirrored volumes during the period of time in which at least one of the images loses its synchronization. The collected metadata represents information about the volume blocks that are overwritten during this period. The presence of the metadata volume enables the system to significantly increase its performance during a rebuild. If the "system-metadata" volume is not created for some reason or is not available, the standard rebuild procedure is executed.

2.4.1.4 Thin Volumes

Thin volumes have most of the same features as any regular volume. The main difference is that thin volumes allocate data blocks only when IOs written to the volume have to be executed outside of the allocated range. Special space provisioned on the local resources is used to make this new allocation. The space is defined as a "thin pool". It provides storage for thin volume data segments and thin volume configuration information needed to handle read and write commands. Up to 10 thin pools can be created per system (see "Storage Concentrator Configuration Limits"). Thin pools can be used to create more than one thin volume. The maximum limit is no more than 128 thin volumes per pool, but additional restrictions are applied each time a new thin volume is created or an existing thin volume is expanded.

Thin pool space is split between two segments. The metadata segment stores configuration information about volumes provisioned in the pool. The data segment provides space to keep volume data. The data segment usage increases each time the system writes to unallocated blocks. The size of the metadata also increases because the metadata segment keeps all of the mapping information between thin volume blocks and the new blocks assigned to the volume at the data segment.

The newly created thin volume does not allocate data and metadata segments, but the pool does reserve some number of metadata blocks to handle future writes into the volume. These blocks are not assigned to any specific thin volume and are used only to check how many volumes the pool can handle in addition to the existing volumes.

The total size of the allocated metadata from all thin pools and all snap enabled volumes is limited by the system RAM ("Storage Concentrator Configuration Limits"). The metadata is stored on local resources, but at run time a copy of this information must be located in the system RAM so that the Storage Concentrator can use it to handle IOs. For performance reasons, metadata should use only physically available RAM and avoid the usage of virtual memory. Another limitation on metadata size is the requirement to be consistent with cluster failover. The new Primary Storage Concentrator has to read all of the metadata from the resources before it can start to handle host IOs. Delays above some limit are treated by iSCSI hosts as an IO failure and are reported to applications.

The available system RAM is checked each time a new thin pool or new thin volume is created or expanded. If the system can handle the request it proceeds with the execution. If the RAM usage is close to the limit or will exceed the limit, then the appropriate recommendations are made so that the user can adjust the request or make some other system configuration changes.

For cluster systems, thin pool active/active load balancing is inherited by the volumes provisioned from the pool. Thin volumes do not have individual load balancing settings.

Thin volumes provide benefits but require additional management and controls that the regular volumes do not have. For details on thin volume management, see "Thin Volumes Management".

2.4.1.5 Deduplicated Volumes

Volume deduplication represents the more advanced feature than the thin provisioning. The main difference is that a deduplicated volume allocates blocks of storage only on writes where the contents of the block are not identical to contents of other blocks allocated in the volume's pool. Deduplication block equivalence is determined through the use of cryptographically strong hashing.

Space is provisioned from local storage resources for the pool and is used for deduplicated volume allocations. The space is defined as a **Deduplicated Pool**. It provides storage all of for deduplicated volume data and configuration info needed to handle read and write commands. Up to 10 Deduplicated Pools can be created per system. A single pool can be used to create multiple deduplicated volumes, with duplications within each volume, and between the volumes in the pool being merged. The maximum limit is no more than 128 deduplicated volumes per pool, but additional memory and metadata based restrictions can occur, and are applied when a new deduplicated volume is created, or existing one is expanded.

Deduplicated volumes are thinly provisioned volumes in that they consume no space until they are written to, and only the blocks written consume space. Data read but never written will return all zeros. By default, deduplicated volumes will not store data blocks with all zeros, and when all zeros are written to an allocated block, that volumes allocation will be freed. If it was the last volume using the block before the write of zeros, the block would also be freed. The policy on whether or not to store all zero blocks can be changed on a per-volume basis with the **Refuse Space Allocation for Zero Writes** setting in the **Volume Configuration -**

Deduplicated Volumes configuration page.

Deduplicated Pool space is split between two sections. The metadata section stores information about volumes provisioned in the pool. The data section provides space for volume data. The data section usage goes up when write happens for block that is not identical to any previously allocated blocks in the pool. The metadata usage also goes up when a volume block address is written to, but had never been written to before, because the metadata section maintains the address mapping for the deduplicated volume blocks and the host's block address.

A newly created deduplicated volume does not have any allocated data and metadata sections, but the pool reserves a number of metadata blocks to handle future writes into the volume. These blocks are not assigned to any specific deduplicated volume and are held in reserve. This is why a deduplication pool with no volumes, or only has volumes that have never been written to consumes metadata.

Total size of allocated metadata from all thin pools, deduplicated pools and snap enabled volumes is limited by system RAM. The metadata are persistently stored on storage resources, but at run time copy of it is maintained in system RAM so the Storage Concentrator efficiently access it to handle IO's. Another limitation on total metadata size is a requirement for SC cluster failover. The new primary SC must read all metadata from storage resources before it can handle host IO's to the volumes dependent upon that metadata, which limits the overall size of the metadata.

Each time when a deduplicated pool or volume is provisioned or expanded, the available system RAM is checked. If the system cannot handle the request without exceeding the RAM usage limit, recommendations are made so user can adjust their request or make other changes to the system configuration.

The user should avoid host operations that attempt to write data into most or all of the deduplicated volume blocks. For example, a full format command, or destructive bad disk block check can update all blocks of the volume. While not an issue with a regular volume, a deduplicated volume has to allocate data blocks during such procedures, thickening it to maximum size, and potentially significantly reducing the deduplication efficiency. In such cases, the storage utilization benefit of deduplication disappears, but user continues to pay a performance penalty that is associated with deduplicated volumes.

For details on deduplicated volume management, see “
Deduplicated Volumes Management”.

2.4.1.6 Operational Considerations for Use of Deduplicated Volumes

There are some considerations that have to be taken into account when creating a deduplicated volume:

1. Not every type of application data is a good fit for deduplication.
2. Volumes that are host or SC encrypted will have a greatly reduced deduplication ratio as the encryption results in a different data pattern for the same data in different parts of the same or other volumes.

3 Compressed media formats for images, music, and video, like JPEG, mp3, .MOV, etc. may not deduplicate well because most of data redundancy is eliminated by the format.
4 The user should avoid using deduplicated volumes in performance critical applications where the deduplication overhead would impact operations.
5 The **Information** screen from SC GUI **System** menu can be used to check free memory and CPU resources available on the system during peak load periods.
6 The Storage Concentrator provides the ability to on-line migrate a regular volume contents to a deduplicated volume, and a deduplicated volume to a regular volume. This is done by creating a mirrored volume with one deduplicated and one regular image, with the subsequent deletion of the source image after the mirror becomes in sync. The procedure can be executed without setting volume offline, but does require space for both types of images during the migration.
7 A good candidate for deduplication is a volume that hosts archivals, backups, user documents, virtual files, or software deployment files that contain data that is modified infrequently and read frequently.
8 The deduplication pool block size is very important, and should match what the file-system or other storage application uses for its allocation block. For many file- system types, this is 4 KB, but the size can be overridden at creation, or vary based on the file-system size. If the deduplication pool block size is larger than the host applications block size, poor deduplication ratios would be expected, even when the data was highly duplicated. If the deduplication block size is smaller than the host applications block size, the deduplication ratio would not suffer, but the SC metadata and performance loading would be higher than it need be.
9 The other important parameter that has to be setup when deduplication pool is created is "Requested Pool Deduplication Ratio Limit (n:1)". This is the requested limit for the best case maximum deduplication ratio for all volumes in the pool when it is completely full. This parameter is used to size the amount of meta-data used to map the volume disk addresses to deduplication blocks. Too low of a value would exhaust the pools meta-data blocks before all of the pool data blocks are consumed. Too high of a value would waste pool and system resources and significantly limit the overall system configuration for an overly optimistic deduplication ratio unachievable with the given volume(s) data. Since the Pool Deduplication Ratio Limit can be adjusted upwards during deduplication pool expansion, it is recommended to start with a lower value until enough representative volume data is stored in the pool to guide an adjustment upwards if warranted.
10 The hosts file-system or application blocks must be aligned with the start of the SC deduplicated volume block. Otherwise, very poor deduplication ratios would occur even for highly duplicated data. The first block that the host file-system or application uses need not be the first block on the deduplicated volume, but it must be on an even deduplication block boundary. Such alignment is often handled automatically, or through controls used when the host creates its partitions and file-systems.
11 A deduplicated pool can, and should be oversubscribed; otherwise, there is no reason to use the deduplication feature. This means that the sum of the configured volume sizes within the pool is larger than the pool size, subtracting metadata.
12 When a deduplication pool is oversubscribed, there is the possibility that it runs out of space for data blocks, or space for metadata. This can occur due to writes to areas of the volume that had never been written to before, or re-writes with data patterns that reduce the deduplication ratio requiring more storage than is available. When that occurs, the deduplicated volume that was doing the write will automatically be placed offline for both reads and writes until manual intervention. The same would happen on any attempt to write any other volume in the same pool. Reads are allowed, as long as no writes are attempted while space is exhausted.
13 Deduplication volumes that are off-lined due to pool data or metadata exhaustion can be brought back on-line in the following ways:

a. The deduplication pool can be expanded, giving it more space.
b. Thin provisioned space reclamation can be done on one or more of the volumes in the pool. This causes no longer used space in the file-system to be released from the volume. A utility is run on the host where the volumes are mounted that writes zeros into blocks no longer unused by the file-system.
c. One or more of the deduplicated volumes can be migrated out of the pool by using each as the first image in a new pool, adding a second image, syncing the mirror, deleting the first image, and reverting the mirror to back to a volume.
d. Delete any volumes from the deduplication pool that are no longer needed.
e. Once deduplication pool space is added or freed, the each volume would be **Activated** to make it available again.

Deduplicated volumes have storage and performance implications. The deduplication process requires additional computing power and memory resources from system where it runs. I/O performance for deduplication volumes is lower when compare to regular volumes. This is because:

1 Write I/O's that are larger than the pool's deduplication block size must be broken into smaller I/O's, less than or equal to the block size.
2 Write I/O's that are not evenly aligned on the pool's deduplication block size boundary require the neighboring data to be read. Small write I/O's necessarily become read/modify/hash/write operations.
3 Read I/O's that are larger than the pool's deduplication block size must be broken into smaller I/O's, less than or equal to the block size.
4 Deduplication is done in parallel with the host I/O's. This means that the hosts write request is completed without the block being deduplicated, with the more time consuming search for a duplicate occurring in parallel. This reduces the write I/O latency for the hosts I/O's, but increases the overall SC and storage system loading, as an unnecessary block allocation and multiple metadata writes occurred when a duplicate is later found.
5 The deduplication data block cryptographic hash calculation requires significant CPU resources.
6 With a significantly high deduplication ratio, there are fewer write I/O's and fewer disk blocks allocated, improving overall disk utilization and cache efficiency.
7 Storage systems tend to optimize sequential I/O accesses, performing read-ahead, and streaming I/O's. With deduplicated volumes, because the deduplication blocks are dynamically allocated and freed, and often shared, their physical location has no correlation to the virtual address the host uses; they would not be sequential to the storage system. On deduplicated volumes, all I/O is random, even when the incoming requests are sequential.

2.4.1.7 Memory Limits for Deduplicated Volumes

Size of system RAM and size of Boot disk have significant effect on system ability to create and handle deduplicated pools and volumes. For regular volumes the main factor is size of available resources. User could allocate new regular volume up to size of available space without restrictions. Pools of deduplicated volumes have to manage not only blocks of volume's data but also maps that let Storage Concentrator find data location in deduplication space for any volume's LBA. These maps represent the most significant part of additional information that each deduplication pool has to handle. All together this information is named as "metadata". The metadata portion of deduplicated pool is stored on disks where the pool space is allocated. The main complication for deduplication is fact that the metadata has to be used at run time to handle IO's for deduplicated volumes. This is the main reason why system has to have enough RAM to store metadata. Metadata for snap enabled volumes and

thin pools have to be counted for the same reason too. Each time when new deduplicated pool has to be created check is done to verify how much of metadata the system RAM could handle in addition to metadata assigned to existing snap enabled volumes, thin pools and deduplicated pools. System let create new deduplicated pool of size that could be handled by the system RAM. At some point no new deduplicated pools could be added in spite fact that the system has available storage.

The other constrain is intention to have balance between metadata and data portions of deduplicated space. In the ideal case the metadata has to have ability to map LBA's from all volumes to all potential data blocks at the deduplicated space. How it can be accomplished? The answer on this question depends of how much deduplication is expected. If the expected deduplication ratio is 4:1, the data portion maybe 4 times less than total size of all volumes in the pool. With expected deduplication ratio 8:1, the data portion maybe 8 times less. This is the reason why one of parameters that has to be selected during deduplicated pool creation is "Requested Pool Deduplication Ratio Limit". See 2.4.1.10: Creating a Deduplicated Pool. It has range from 1:1 to 32:1. The parameter could be reset later only during pool expansion. See 2.4.3.2: Expanding a Volume.

Size of deduplicated block is the other parameter that is critical to define seizes of new deduplicated pools and volumes. The available values are 4K, 8K, 16K, 32K and 64K. Pools with 8K deduplication blocks require 2 times less mappings for volume's LBA's if compare to pools with 4K deduplication blocks. But it may be less deduplication when 8K deduplication blocks are used. It has no sense to use deduplication when deduplication ratio drops to be close to 1:1. This is the reason why recommended size is 4K. With 4K deduplication block selection it gives the next rough estimation for size metadata portion: it requires 1GB of metadata per each 1TB of volume space approximately. User should not be licensed to run deduplication on systems with less than 7GB of RAM. See A5.2.7.

During deduplicated pool and volume creation system uses precise formulas not an estimation that was mentioned above. These formulas could be modified during updates from one release of the software to another.

Recommendations:

- Storage Concentrator installations that do not support deduplicated pools should not get license to enable Deduplicated Volumes after they are upgraded to version that supports deduplicated pools. The main obstacle is that these systems usually do not have enough RAM and size of Boot disk is not enough to handle deduplication. The first problem can be fixed by adding more RAM to the system. The second one requires new installation not an upgrade. User's data have to be preserved in some way and put back after new installation is done.

- Field "Usable Space" on Create Deduplicated Pool screen can be used as an indicator of what is the largest deduplication pool could be created. Select different deduplication "Block Size" or "Requested Pool Deduplication Ration Limit" to see what is the "Usable Space" is in this case.

- The "Usable Space" is reduced after new deduplication pool is created. By doing this system makes reservation to support future IO traffic to the pool. Size of already allocated volume's data blocks at pools does not affect the "Usable Space" value.

- Size of new deduplicated volume is limited by how much of data blocks the

existing deduplicated volumes in the pool are already support without acquiring additional resources from the pool. The more pool resources are already acquired by its volumes the smaller new deduplicated volume could be created. The volumes current ability is counted here not the potential that has to handle full range of all volume's LBA's. Deduplicated volume gets resources from the pool dynamically when there is no space in already acquired resources to handle new writes. The implementation creates situation when volumes are competing at run time for metadata to handle its own LBA's, but sharing data blocks with other volumes from the pool.

- To find what is the current limit for a new deduplicated volume user has to select appropriate pool, put very large number into field "Desired Volume Size" and hit button "Submit". The pop-up screen will show message that the request is rejected but the maximum recommended volume size is suggested.

2.4.1.8 Creating a Volume

There are five methods for creating volumes that are exposed by Storage Concentrator as an iSCSI targets:

The Auto Create option automatically creates the volume from the available space on any available resource.

The Manual Create option allows volumes to be created by manually selecting space on specific resources.

The Pass Thru option allows volumes to be created from resources that have been assigned the Pass Thru use type.

The Thin Volume Create option lets the user select the pool to provision the volume from.

New thin pools can be created manually or automatically when the Thin Pool Create option is selected on the Thin Volume Create page.

The Create Deduplicated Volume option lets the user select the pool to provision the volume from.

New deduplicated pools can be created manually or automatically when the Deduplicated Pool Create option is selected on the Create Deduplicated Volume page.

A Pass Thru resource will accept SCSI commands from a host without any intervention from the *Storage Concentrator* except Reservation and Persistent Reservation commands.

To create a volume automatically using the Auto Create option, use the steps that follow:

1 Click **Volumes**. The list of existing volumes displays on the screen.

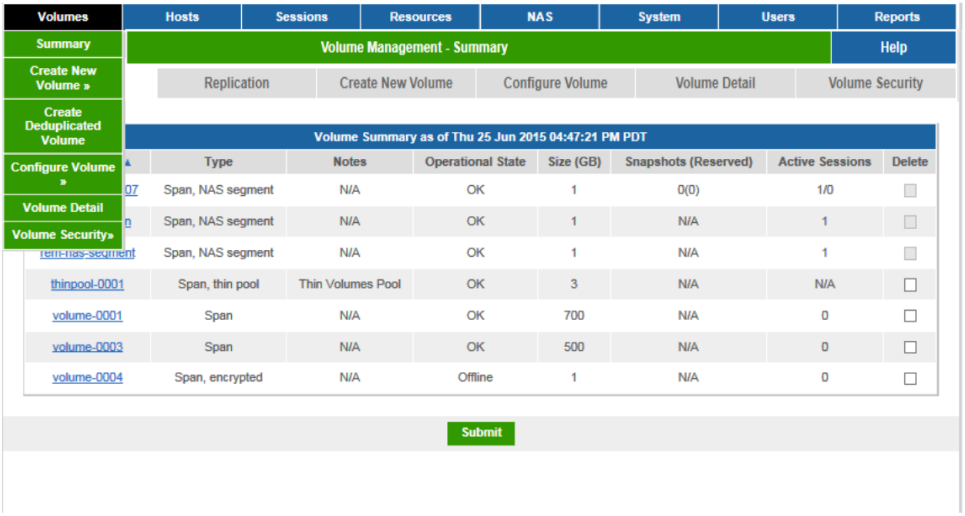


Figure 2-10 Volume Management Summary Screen

2 Click **Create New Volume**. The Volume Management Create New Volume screen appears.

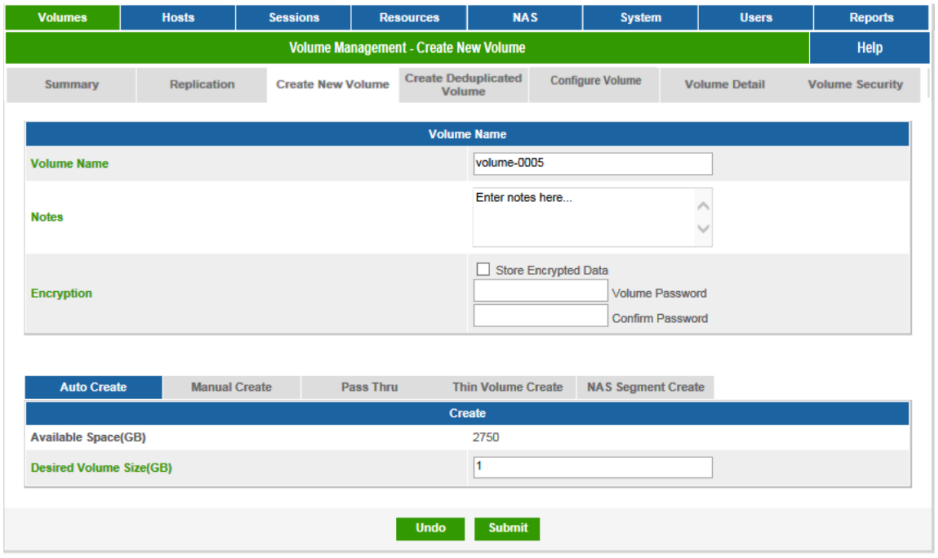


Figure 2-11 Volume Management Create New Volume Screen (No Cluster, Volume Encryption is Licensed)

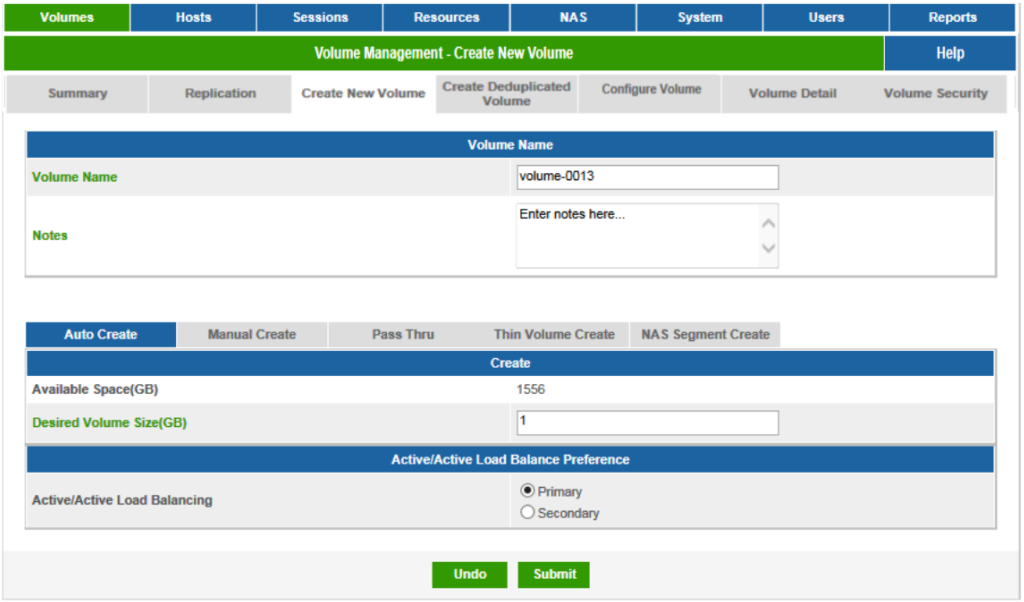


Figure 2-12 Volume Management Create New Volume Screen (in a Cluster, without License for Volume Encryption)

3 Enter a name for the volume in the **Volume Name** field.

4 Enter any descriptive notes regarding this volume in the **Notes** field. These notes appear on the Volume Management Summary screen.

5 If the volume is being created on a system with a Volume Encryption License, the encryption password fields display on the screen. Use these fields to create an encrypted volume. Type the same password into each password field. DO NOT LOSE THIS PASSWORD. As the volume is being created the new password and encryption keys are created and written to the USB device. See the Encryption Overview above.

6 Enter a size for the volume in the **Desired Volume Size** field. The size must be in whole GigaBytes and must be a minimum of one (1) GigaByte.

7 If this volume is being created in a Failover Cluster, the screen includes the selection of the Primary or Secondary *Storage Concentrator* to service the IO's for the volume. Click on the radio button associated with the desired system. (See the information on Load Balancing in Section 3: Failover.)

Click **Undo** to revert to the saved settings

8 Click **Submit**. The following dialog box appears.

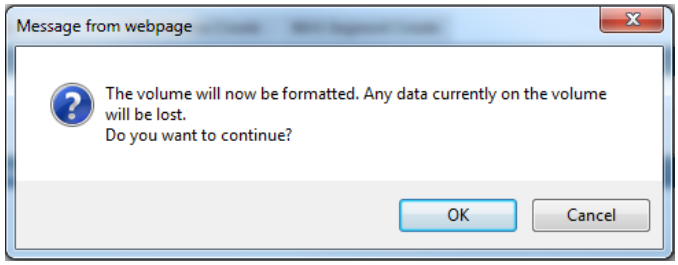


Figure 2-13 Create Volume Dialog Box

9 Click **OK** to continue or **Cancel** to end. A volume with the specified name and size is created automatically. The Volume Security screen opens to allow you to create an access control list for this volume from any known hosts. For information on creating access control lists, see "Volume Security".

To create a volume manually using the Manual Create option, use the steps that follow:

1 Click **Volumes**. The list of existing volumes displays on the screen.

2 Click **Create New Volume**. The Volume Management Create New Volume screen appears.

3 Enter a name for the volume in the **Volume Name** field.

4 Enter any descriptive notes regarding this volume in the **Notes** field.

5 Click Manual Create. A list of available resources appears.

6 If the volume is being created on a system with a Volume Encryption License, the encryption password fields display on the screen. Use these fields to create an encrypted volume. Type the same password into each password field. DO NOT LOSE THIS PASSWORD. As the volume is being created the new password and encryption keys are created and written to the USB device. See the Encryption Overview above.

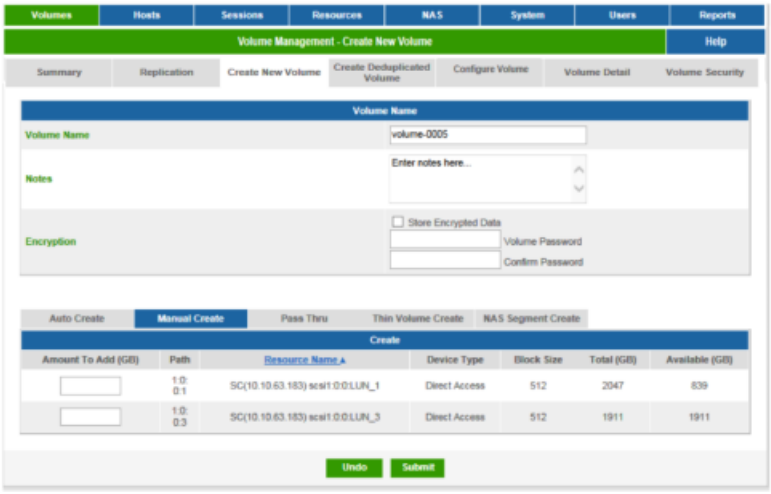


Figure 2-14 Volume Management Manual Create Screen (No Cluster, Volume Encryption is Licensed)

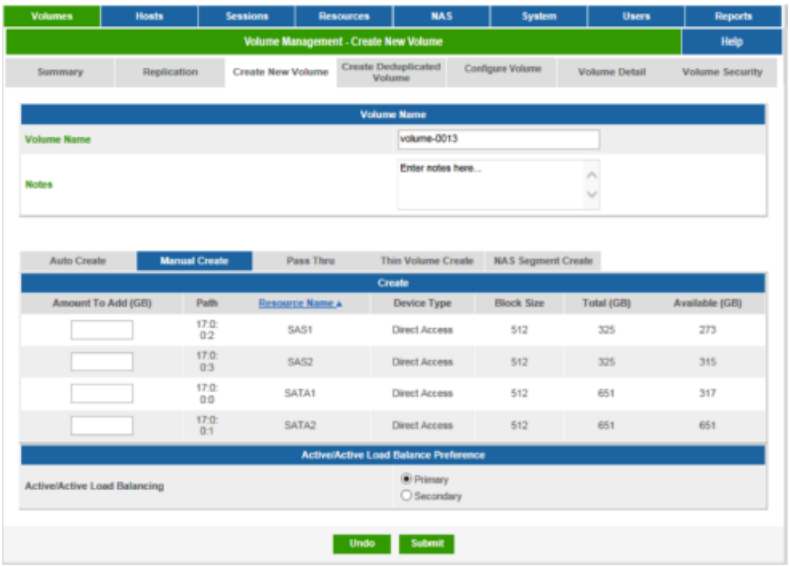


Figure 2-15 Volume Management Manual Create Screen (in a Cluster, without Volume Encryption License)

7 For managed resource types, in the **Amount to Add** field, enter the amount of space from this resource that is to be used to create this volume.

8 Repeat step 6 for each managed resource type that is being used to create the volume until the total amount of space desired for this volume has been allocated.

9 If this volume is being created in a Failover Cluster the screen includes the selection of the Primary or Secondary *Storage Concentrator* to service the IO's for the volume. Click on the radio button associated with the desired system. (See the information on Load Balancing in Section 3: Failover.)

Click **Undo** to revert to the saved settings.

10 Click **Submit**. A volume with the specified name and size is created. The Volume Security screen opens to allow you to create an access control list and a CHAP secret for this volume from any known hosts. For information on creating ACLs and setting a CHAP secret, see "Volume Security".

To create a pass thru volume using the Pass Thru option, use the steps that follow:

1 Click **Volumes**. A list of existing volumes appears.

2 Click **Create New Volume**. The Volume Management Create New Volume screen appears.

3 Enter a name for the volume in the **Volume Name** field.

4 Enter any descriptive notes regarding this volume in the **Notes** field.

5 Click the **Pass Thru** button. A list of available pass thru resources appears. No encryption is allowed on Pass-Thru volumes.

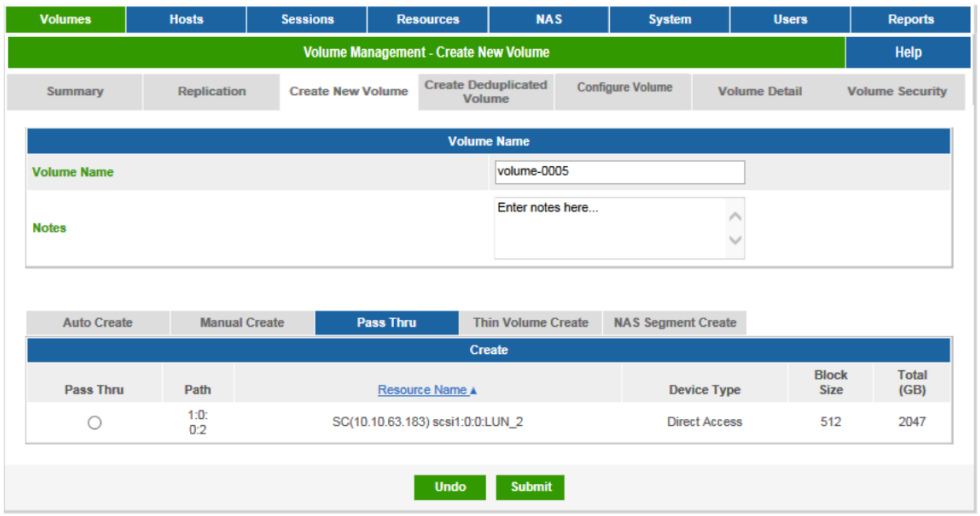


Figure 2-16 Volume Management Pass Thru Screen (No Cluster)

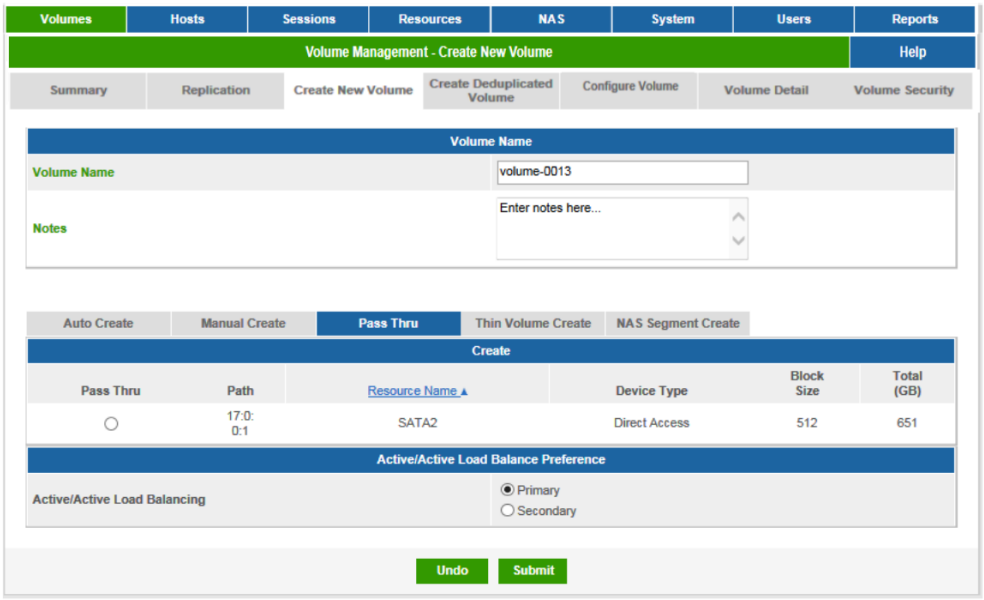


Figure 2-17 Volume Management Pass Thru screen (in a Cluster)

6 Click the **Pass Thru** check box on each resource to add to the volume.

All Pass Thru resources selected for a single volume must be in the same device. (The TargetID must match)

Click **Undo** to revert to the saved settings

7 If this Pass-Thru volume is being created in a Failover Cluster, the screen includes the selection of the Primary or Secondary *Storage Concentrator* to service the IO's for the volume. Click on the radio button associated with the desired system. (See the information on Load Balancing in Section 3.2.4)

8 Click **Submit**. A pass through volume is created.

To create a thin volume using the Thin Volume Create option, use the steps that follow:

1 Click **Volumes**. A list of existing volumes appears.

2 Click **Create New Volume**. The Volume Management Create New Volume screen appears.

3 Enter a name for the volume in the **Volume Name** field.

4 Enter any descriptive notes regarding this volume in the **Notes** field.

5 Enter a size for the volume in the **Desired Volume Size** field. The size must be in whole GigaBytes and must be a minimum of one GigaByte.

6 Click the **Thin Volume Create** button. A list of available thin pools appears. No encryption is allowed on thin volumes.

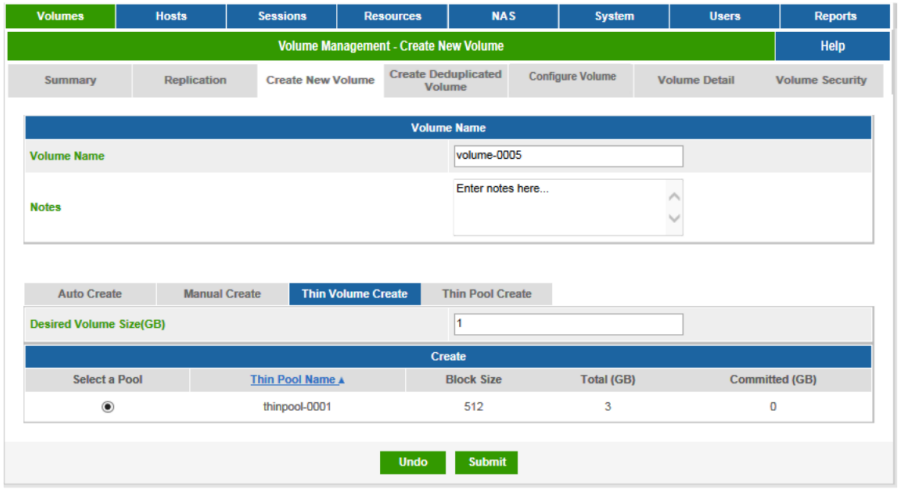


Figure 2-18 Create Thin Volume Screen (No Cluster).

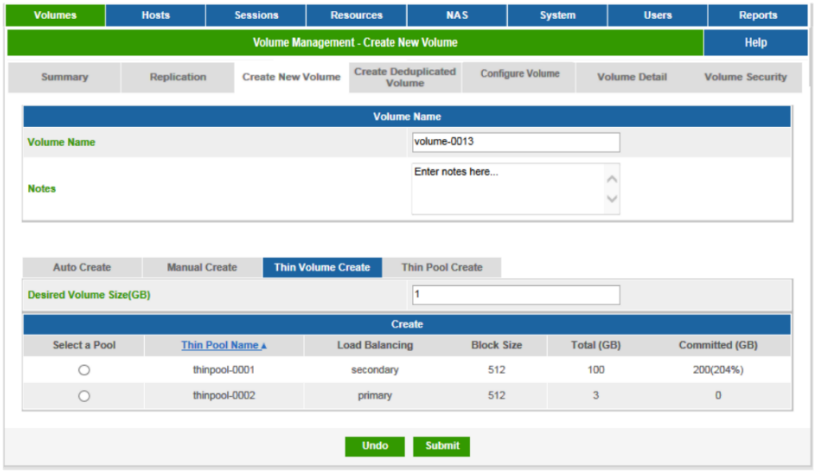


Figure 2-19 Create Thin Volume Screen (in a Cluster).

7 Click the **Select Pool** radio button on the pool that is to be used to provision the volume.

If a thin volume is being created in a Failover Cluster, the screen includes information about the Primary or Secondary *Storage Concentrator* to be used to service the IO's for the volume. The volume inherits load balancing from the selected pool.

8 Click **Undo** to revert to the saved settings

9 Click **Submit**. A thin volume is created. The Volume Security screen opens to allow you to create an access control list and a CHAP secret for this volume from any known hosts. For

information on creating ACLs and setting a CHAP secret, see "Volume Security".

To create a deduplicated volume using the Create Deduplicated Volume option, use the steps that follow:

1. Click **Volumes**. A list of existing volumes appears.
2. Click **Create Deduplicated Volume**. The Volume Management Create Deduplicated Volume screen appears.
3. Enter a name for the volume in the **Volume Name** field.
4. Enter any descriptive notes regarding this volume in the **Notes** field.
5. If the volume is being created on a system with a Volume Encryption License, the encryption password fields display on the screen. Use these fields to create an encrypted volume. Type the same password into each password field. DO NOT LOSE THIS PASSWORD. As the volume is being created the new password and encryption keys are created and written to the USB device. See the Encryption Overview above.
6. Enter a size for the volume in the **Desired Volume Size** field. The size must be in whole GigaBytes and must be a minimum of one GigaByte.

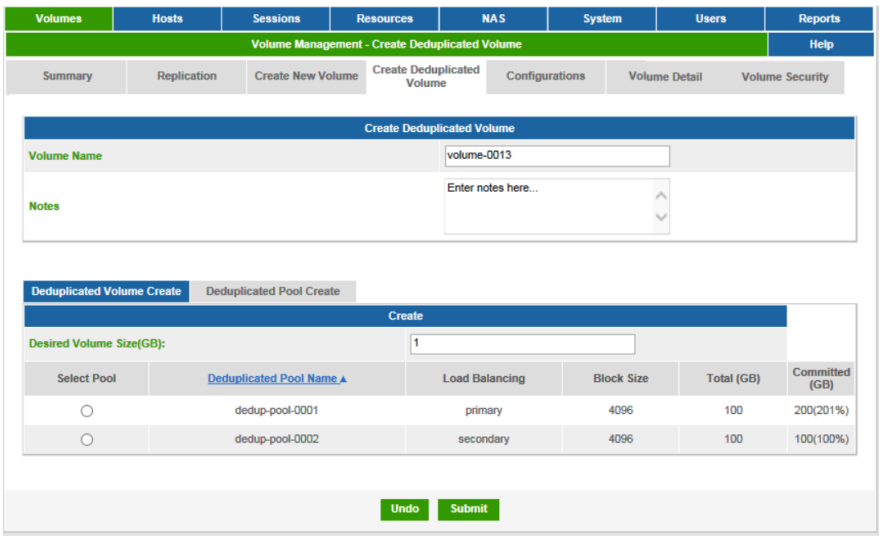


Figure 2-20 Create Deduplicated Volume Screen (in a Cluster).

7. Click the **Select Pool** radio button on the pool that is to be used to provision the volume.

If a deduplicated volume is being created in a Failover Cluster, the screen includes information about the Primary or Secondary *Storage Concentrator* to be used to service the IO's for the volume. The volume inherits load balancing from the selected pool.

8 Click **Undo** to revert to the saved settings

9 Click **Submit**. A deduplicated volume is created. The Volume Security screen opens to allow you to create an access control list and a CHAP secret for this volume from any known hosts. For information on creating ACLs and setting a CHAP secret, see "Volume Security".

2.4.1.9 Creating a Thin Pool

The first thin pool has to be created before the system can execute the first request to create a thin volume. The user will be redirected to the Thin Pool Create screen automatically. Subsequent thin pools can be created by accessing the Thin Pool Create screen from the Thin Volume Create GUI page. There are manual and auto Thin Pool Create options that are similar to the same options for creating a regular volume, see "Creating a Volume" for more details.

Thin pools cannot be on an encrypted volume. If the system is a failover cluster, the pool's load balancing is inherited by all thin volumes that are provisioned in the pool. Thin pools do not have Volume Security attributes because they are used internally as space is designated for thin volume data segment allocation.

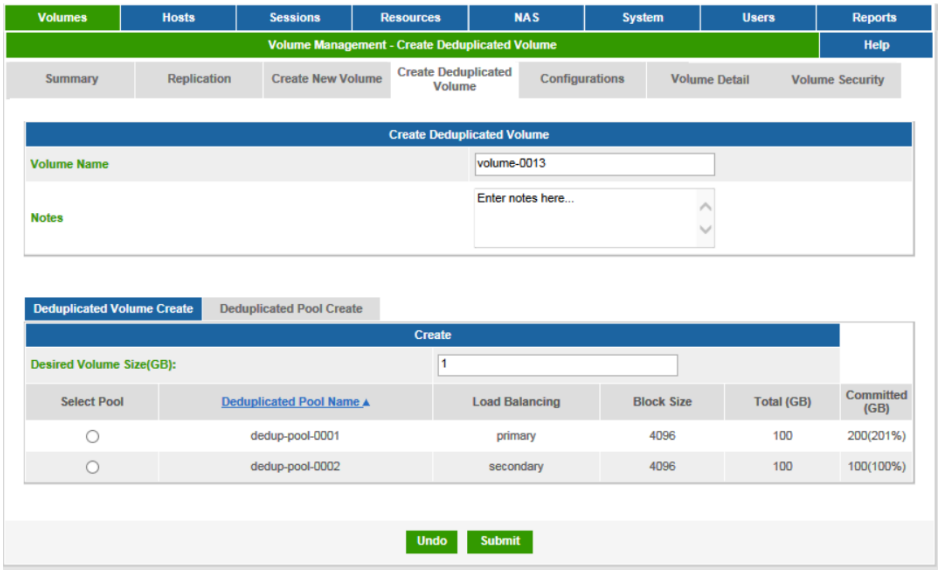


Figure 2-21 Volume Management Thin Pool Manual Create Screen (No Cluster)

2.4.1.10 Creating a Deduplicated Pool

The first deduplicated pool has to be created before the system can execute the first request to create a deduplicated volume. The user will be redirected to the Deduplicated Pool Create screen automatically. Subsequent deduplicated pools can be created by accessing the Deduplicated Pool Create screen from the Create Deduplicated Volume GUI page. There are manual and auto Deduplicated Pool Create options that are similar to the same options for creating a regular volume, see "Creating a Volume" for more details.

Deduplicated pools cannot be on an encrypted volume. If the system is a failover cluster, the

pool's load balancing is inherited by all deduplicated volumes that are provisioned in the pool. Deduplicated pools do not have Volume Security attributes because they are used internally as space is designated for deduplicated volume data segment allocation.

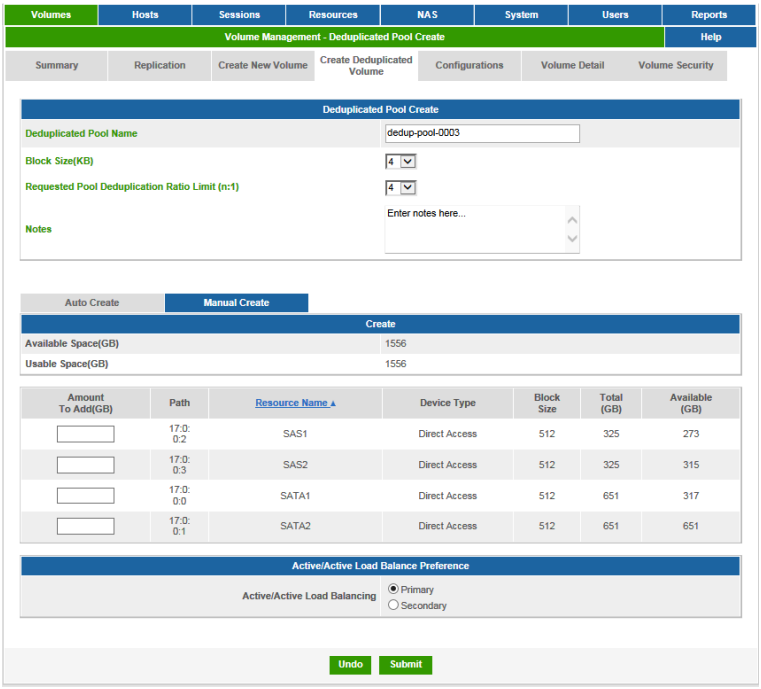


Figure 2-22 Volume Management Deduplicated Pool Manual Create Screen (in a Cluster)

2.4.2 Volume Security

After a volume is created, security can be set using CHAP and an access control list. The access control list allows hosts to access the storage on the volumes. (A host is a computer connected to storage. Typically a host is a server running applications or providing services that access and consumes storage.)

Before attempting to create an access control list or setting a CHAP secret, at least one host must be recognized by the *Storage Concentrator.* For more information on hosts, see "Adding a Host".

2.4.2.1 About CHAP

CHAP (Challenge Handshake Authentication Protocol) allows you to set a Password or "Secret" as a gatekeeper for communication between a host initiator and a volume. Combined, access control lists and CHAP provide a high degree of security to ensure that only specified hosts have access to Storage Concentrator volumes.

CHAP is supported at the Volume Level and at the Host level in the Storage Concentrator. Depending on your host initiator, you may want to specify host CHAP, volume CHAP, both host and volume with the same or different secrets, or use neither. (Please check with the initiator manufacturer to determine what level of CHAP is supported at the host level.)

The following are some sample CHAP scenarios:

You may first set up CHAP on the Storage Concentrator or on the Host. For example, to set up Volume CHAP using the Microsoft software initiator:

1. When you are first setting up a Storage Concentrator, logout the host sessions and make sure no ACLs apply to the volumes you will use with CHAP.
2. Log into the Storage Concentrator using the Microsoft Initiator without CHAP. Add a target portal to the Microsoft Initiator, which causes a discovery login. You will now have one host listed on the host page of the SC.
3. On the Storage Concentrator, create volumes and assign Read/Write access and a CHAP secret per volume and per host.
4. Use the Microsoft Initiator GUI to login to an available target using the advanced button and specify the CHAP secret that matches the CHAP secret on the Storage Concentrator. If your CHAP secret does not match, you will be unable to get a response from the *Storage Concentrator*.

If you wanted to set up Host Chap with the Microsoft Initiator:

1. When you are first setting up a Storage Concentrator, logout the host sessions and make sure no ACLs apply to the volumes you will use with CHAP.
2. Add hosts on the *Storage Concentrator* using the Host Access screen and assign a CHAP Username and Password.
3. Use the Microsoft Initiator GUI to add a Target Portal specifying the CHAP secret that matches the CHAP secret on the Storage Concentrator. If your CHAP secret does not match, you will be unable to get a response from the Storage Concentrator.

To configure security settings for a volume, use the steps that follow:

The volume creation process brings you to this screen automatically. If you are creating volumes at a different time, use the steps that follow.

1. On the **Volume Management** screen, click **Volume Security**. The Volume Management Security screen appears.

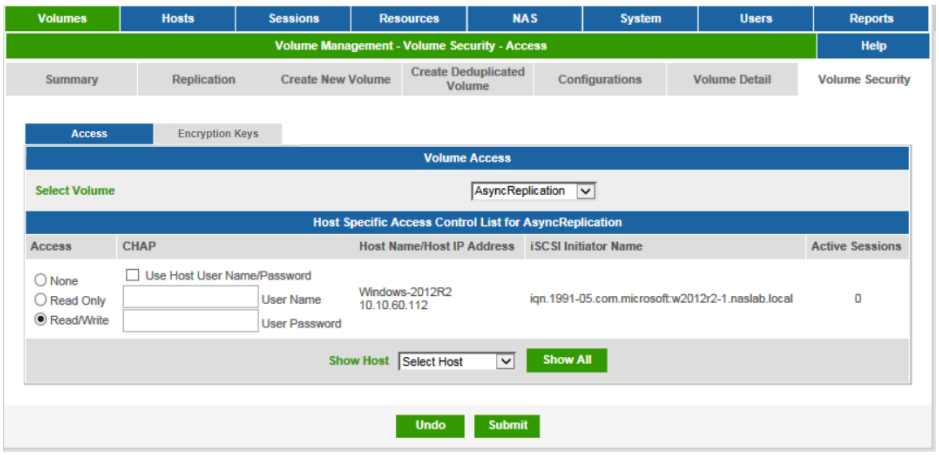


Figure 2-23 Volume Management Security Screen

A list of hosts appears with the following information:

The list only contains hosts that currently have access to the selected volume.

Host IP Address: The IP address of the host is entered into the system in one of two ways. It can be added automatically whenever a new host makes a request to the *Storage Concentrator,* or it can be added manually in the Hosts screen. For information on adding a host manually, see "Adding a Host".

Host Name: The name of the host may be entered into the system in one of two ways. It can be added automatically whenever a new host makes a request to the *Storage Concentrator*, or it can be added manually in the *Storage Concentrator* administrative interface. If the name of the host is added through the host request process, the name of the host will be the IP address of the host. If the name of the host is added in the *Storage Concentrator*, a user-defined name may be assigned to the host.

iSCSI Initiator Name: The name given to the initiator when it was created with its vendor-specific software.

Active Sessions: The number of connections between the host and the currently selected volume.

Show Hosts: The available hosts that can be given access to this volume.

2 From the volume list, select a volume to assign to the host.

3 Select the desired access permission:

None

Read Only: This permission is available only on volumes with the use type of Managed and Pass Thru

Read/Write: This permission is available only on volumes with the use type of Managed

4 Select **CHAP** by clicking on the Checkbox "Use Host User Name/Password if you wish to associate the name of the iSCSI Initiator to the IP address shown.

To have access to this volume when CHAP is selected, the host initiator must know the Specified CHAP Secret (password). Select a CHAP password that is the same as that provided to the host initiator. For example, the Microsoft initiator requires secret with 12 or more alphanumeric characters.

Most initiators support CHAP on the volume level, but some only support CHAP on the host level. Some initiators support CHAP at both the Host level and the Volume level. To specify a CHAP secret for a host, see "Adding a Host".

5 To add hosts that are allowed access to the volume, do one of the following:

a. Select a host from the **Show Host** list to allow a new host to have access to this volume. Once selected, the host is added to the table. Assign the type of access permission, and click **Submit**.
b. Click **Show All** to select all the available hosts, then assign the type of access permission, and click **Submit**.

Click **Undo** to revert to the saved settings.

6 Click **Submit**

If there is a mismatch between the CHAP secret on the initiator and the *Storage Concentrator*, you will get a target Alert message on the host.

2.4.2.2 Restoring the Encryption Keys to your USB drive

The USB drive holding encryption Keys may be lost or damaged. The Storage Concentrator does not function properly if the Encryption Keys are missing at boot times and other volume configuration points. The Storage Concentrator User Interface allows a non-encrypted USB drive to be formatted and rebuilt as the official encryption USB drive. The non-encrypted USB drive is identified on the System->Admin screen. The USB drive status states that the current USB drive does not contain the encryption information for this Storage Concentrator's volumes. Once that status is encountered the USB drive may be rebuilt to contain the encryption information for this Storage Concentrator. The rebuild/repair process requires the Passwords that were used to create the individual volumes. Without these passwords the USB drive cannot be restored.

Steps to restore the Encryption USB drive:

1 Navigate to the Volumes->Volume Security screen. The default screen contents refer to assigning the volume to a specific host. The screen includes a button for the "USB Key". This button initiates the USB drive rebuild/repair process, or checks for USB drive presence in the system. The "Rebuild" button has to be used to restore an USB encryption drive by using a new USB device if the old one is not available for some reason. The "Repair" button repairs an existing USB encryption drive if the drive does not have valid encryption information for some of the encrypted volumes.

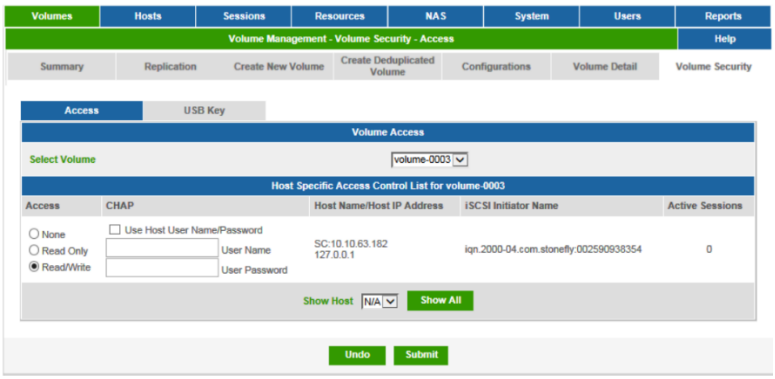


Figure 2-24

2 Click on the "USB Key" button to view the USB drive status. It should appear as shown below. There should not be any encryption information on the drive. If the drive contains information for a different Storage Concentrator, then the status information will notify the user of such.

Figure 2-25

3 Click on the "Rebuild" button to format and empty the USB drive. This prepares the USB drive to hold the new encryption information. Click on the "Format" button to start the format process.

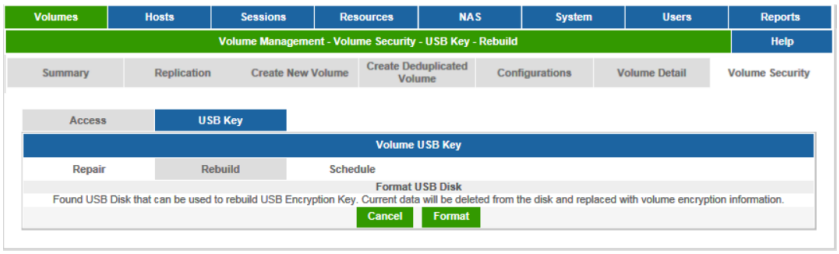
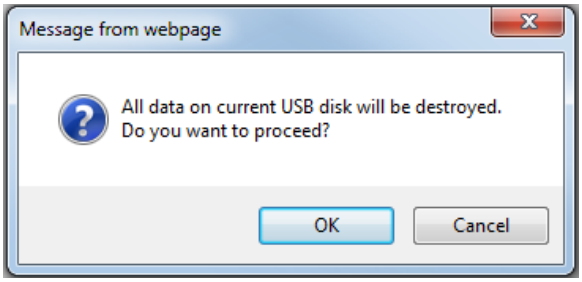


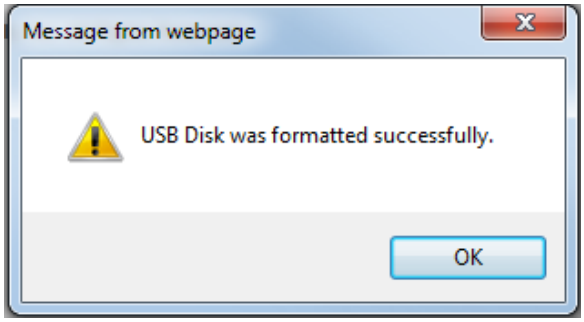
Figure 2-26

A pop-up message confirms the format operation prior to deleting any information

from the device. Click on OK to start the format process.



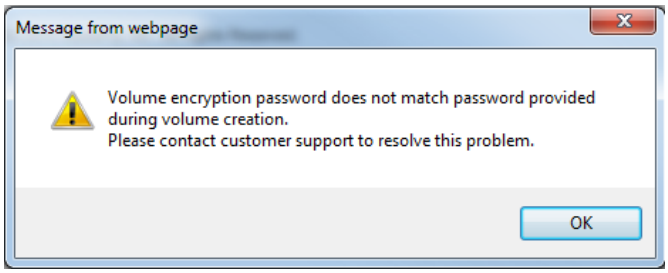
4 At the end of the format process the Storage Concentrator displays a pop-up message to indicate success or failure. The following message appears after a successful format.



5 After the format process is complete, the screen will be updated to show the first encrypted volume in the Storage Concentrator. Type the correct password in the Password edit field and click on Submit to rebuild the first encryption key. Each different encrypted volume is displayed in a similar screen. Enter the proper password for each volume as they appear in the screen as shown below.

Figure 2-27

If the entered password matches the original password, the correct data is placed on the USB drive. If the entered password does not match the original the following message is displayed. Click OK and try again.



If the original password is not known or found by the time of the next reboot of the Storage Concentrator (such as a software upgrade), the volume will become unusable. **Do not lose the list of Passwords! Make copies of the USB drive!**

2.4.2.3 Encryption USB Drive Check Scheduler

The USB drive holding the encryption Keys should be present in the system only when encryption volume has to be created or deleted or in case a standalone Storage Concentrator has to be rebooted. Most of time the drive should to be stored in a safe place away from the Storage Concentrator. Users can set the scheduler to check for drive presence. If the USB drive is present the scheduler has to generate log message with critical severity level. Users can establish notifications that retransmit all messages with this severity through E-mail or SNMP.

Steps to manage the Scheduler for Checking for the USB Encryption USB Drive:

- Navigate to the Volumes->Volume Security->USB Key->Scheduler screen.

  Select the hour and minutes from the pull down menus and click "Submit" button to set the time schedule. To delete the schedule, click on the check box under "Remove From Scheduler" and then click on the "Submit" button.

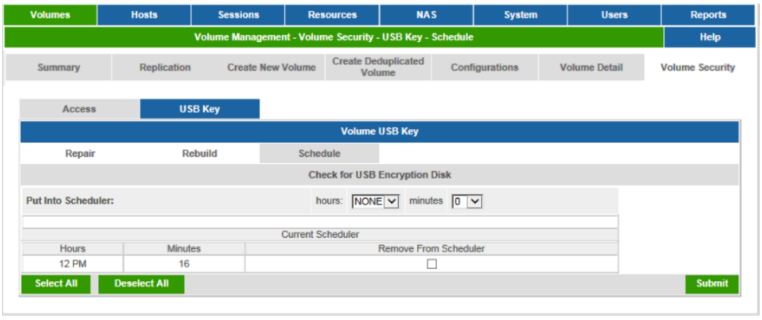


Figure 2-28

2.4.3 Volume Configuration

The General Configuration screen allows you to view and edit the following settings:

Logical Volume Name: System name of the volume

Volume Notes: Descriptive notes about the volume

Volume Control: Settings that allow you to put the volume online or offline

iSCSI Target Name: To change the iSCSI name of the volume.

To edit Volume Configuration, use the steps that follow:

1 Select Volume Management

2 Select Volume Configuration

3 Select General Configuration

4 Select the Volume you wish to edit from the drop down menu.

5 Edit the settings as desired.

Click **Undo** to revert to the saved settings.

6 Click **Submit**

The Volume Configuration screen will change depending on the volumes that have already been created. The drop down menus will offer the volumes that are available

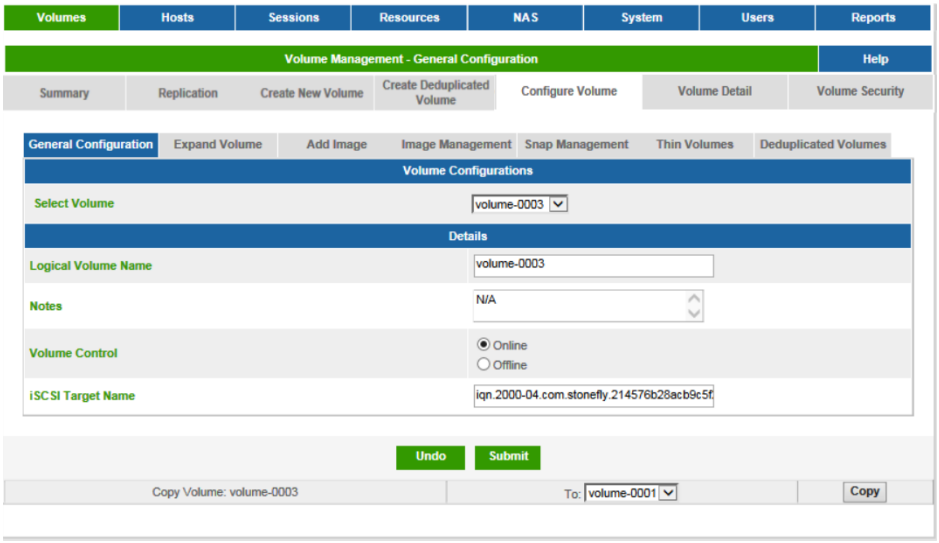


Figure 2-29 Volume Configuration Screen

2.4.3.1 Copy Volume

The Copy Volume function makes an exact copy of a spanned volume, a mirror volume, a pass-thru volume, a thin volume, a deduplicated volume or a Snapshot Live Volume on Stand Alone or Cluster of One Storage Concentrators. For Clusters of Storage concentrators the copy can be executed for spanned and pass-thru volumes only. The copy destination must be the same size or larger than the original. This function is primarily designed to create duplicate volumes for disaster recovery or data distribution. You may run up to five (5) copies of any volumes simultaneously.

To use the copy volume function, you must first create a destination volume of the same size or larger using the Create Volume Function (see section: "Creating a Volume"). Once the volume is created, you will need to copy from the original volume to the destination.

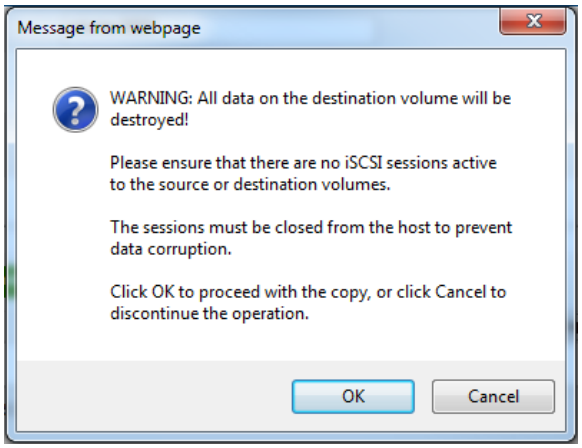
If a destination volume of an appropriate size has not been created, the Copy Volume function will not display.

1. Select **Volume**.
2. Click on **Volume Configuration** screen (as shown above).
3. Select the Volume you want to copy **from** the pull down menus at the top of the screen.
4. Select the Volume you want to copy **to** from the pull down menu next to the Copy Button. Only eligible volume selections will appear in this menu.

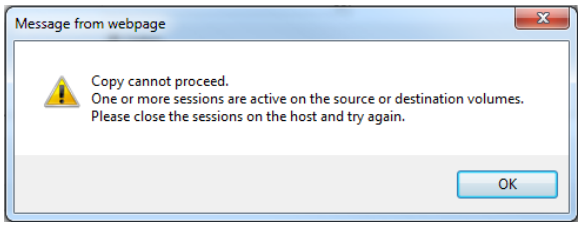
The volume you are copying from must be "Online" to be able to copy from it. Offline volumes are not available to copy. There must be no hosts logged into the source or destination volumes, and they remain inaccessible to hosts during the copy.

Volumes enabled with encryption are not eligible for the Volume Copy feature. A volume must be copied by performing the copy at the host level to insure the encryption is performed properly. The process includes:

1. Identify the volume to be copied.
- Create an encrypted volume to receive the data. The volume must be at least as big as the original volume. If encryption is desired on the copy it must be chosen at the time the copy is created. The system cannot guarantee the use of the same encryption key on both the original and the copy.
- Mount the copy volume on the server and use a function of the host operating system to copy from the original volume to the copy volume.
- Click **Copy.** The system will display the following message:



- Click **OK** if you have already logged out all host sessions. If you have not logged out all host sessions, click **Cancel.** If you click **OK** and all host sessions have not been logged out, the system displays the following Alert message. Click **OK.** Complete logging out all hosts and then proceed with the copy.



- Once the copy has commenced, wait for it to finish before continuing.

Depending on the size of the volume, the copy function will take some time to complete since it is creating a block-level copy of the entire volume. You may make only one copy at a time. Once you have completed the first copy, you may make additional copies of the volume.

2.4.3.2 Expanding a Volume

To increase the size of a volume after it is created,

1 Click Volume

2 Click Volume Configuration

3 Click Expand Volume. The Expand Volume screen appears.

With StoneFusion Version 3.0 or above you do not need to log out of sessions to a volume on the host initiator before expanding a volume. You may expand the volume while it is online. If you are running StoneFusion Version 2.2 or earlier, you will need to log out all sessions to the volume on the initiators before expanding the volume.

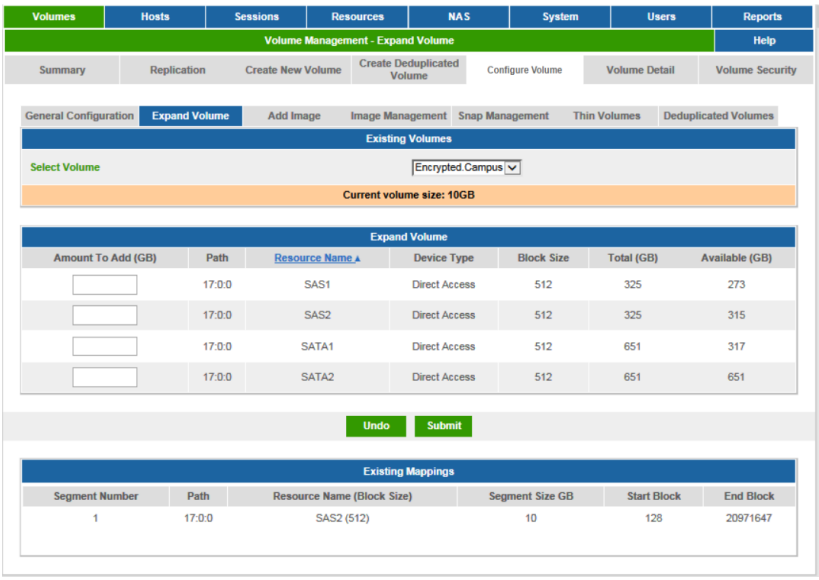


Figure 2-30 The Expand Volume Screen

1 Select the volume you wish to expand from the drop down menu. All available resources are shown. Encrypted volumes appear on the dropdown list along with the non-encrypted volumes.

2 Enter the desired expansion amount in GigaBytes. The minimum expansion amount is 1 GB. The maximum size is equal to the Available GB. You do not need to expand volumes symmetrically, except you must expand all images simultaneously in mirror volumes. You may select space from one resource or expand multiple resources to reach the desired expansion amount.

On the Host Server (Microsoft Windows):

1 Go to Disk Management.

2 From the top menu, select Rescan Disks, and you should see the extra blank unallocated space beside your partition.

3 If you have Windows Server 2003, use the diskpart command to expand your partition to include the extra disk space

4 If you have Windows Server 2008 or later, convert the disk to "Dynamic disk". The Microsoft initiator did not support dynamic disks in Server 2003, but it does for Server 2008. Check the Microsoft initiator release notes for more information.

5 Right click on the original partition and expand/extend it to include the extra disk space.

All images in a Mirror Volume must be expanded symmetrically and simultaneously. For more information on Mirror Volumes, see "StoneFly Mirroring

Overview".

Do not attempt to expand a Campus Mirror image or a Snapshot. Follow the procedure below for expanding a Campus Mirror image. Expanding a Snapshot is not possible but you may expand the Snapshot Live Volume and the Snapspace.

4 Click **Submit** to expand the volume.

If you previously logged out of the host sessions before expanding the volume, remember to log the hosts back in to the volume after the expansion is complete.

2.4.3.3 Expanding a Volume with Campus Mirror Image

To expand a volume with a Campus Mirror image, you must follow the following steps:

1 Expand the volume used as the Campus Mirror image on the Secondary *Storage Concentrator.*

2 Rediscover the volume on the Primary *Storage Concentrator*. Verify its size.

3 Go to the "Expand Volume" GUI page on the Primary *Storage Concentrator* and execute the expansion.

2.4.3.4 Expanding a Volume with Asynchronous Mirror Image

To expand a volume with an Asynchronous Mirror image, you must follow the following steps:

1 Expand the volumes used as the Asynchronous Mirror image on the Remote *Storage Concentrator.*

2 Rediscover the volume on the Primary *Storage Concentrator*. Verify its size.

3 Go to the "Expand Volume" GUI page on the Primary *Storage Concentrator* and execute the expansion.

2.4.3.5 Expanding a Snap-enabled Live Volume

1 Select the volume you wish to expand from the drop down menu. All available resources are shown.

2 Enter the desired expansion amount in GigaBytes. The minimum expansion amount is 1 GB. The maximum size is equal to the available GB.

3 If after expanding the volume you wish to expand the Snapspace as well, repeat the procedure for the Snapspace.

4 After the expansion is complete, you may resume taking Snapshots.

2.4.3.6 Expanding the Snapspace for a Snap-enabled Live Volume

1 An additional feature has been added that allows a user to expand the amount of the Snapspace assigned to a Snap-enabled Live Volume through the Volume Expansion menu. This feature is most useful for adjusting the space needed to hold changes in Asynchronous Mirrors.

2 Navigate to the Expansion page. Select the Snap-enabled volume from the drop down menu and select the Snapspace volume you wish to expand. Note that only the Snapspace appears on the list. The size indicated is the size of the Snapspace. (See image below)

3 The menu allows the amount of expansion to be entered for a specific resource. Enter the amount in the text box to the left of the desired Resource.

4 Click **Submit.** The size of the Snapspace is automatically increased.

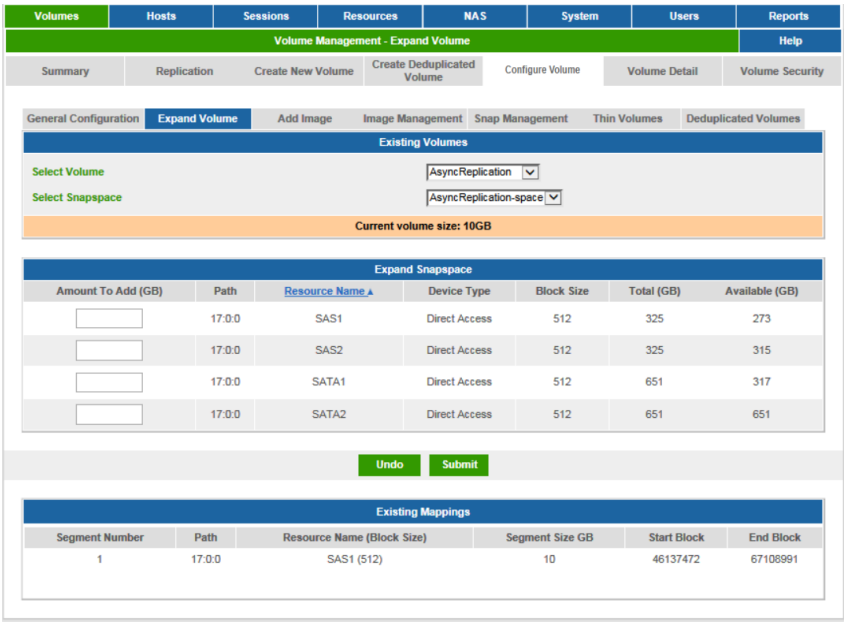


Figure 2-31 Selecting the Snapspace to be Expanded

2.4.3.7 Removing Volumes

Do not remove a volume until all storage on that volume is moved to another location, either to another volume or to backup. If a volume is selected for removal, a warning dialog box appears to remind you that all data will be lost.

To remove volumes, use the steps that follow:

1 Click Volumes. A list of all known volumes appears.

2 Click the check box under **Delete** for the volume to be removed.

3 Click **Submit** to remove the volume.

4 The system asks if it is OK to proceed with the deletion. Click yes to continue. The database tables in the *Storage Concentrator* are updated with the new status.

2.4.3.8 Removing Existing Hosts from the ACL

When a host no longer requires access to a storage volume, the host may be removed from the ACL.

To remove a host from the ACL, use the steps that follow:

1 On the **Volumes Summary** screen, click **Volume Security.** The Host Specific Access Control List appears. The list displays a line for each host that has been given access to the volume. Typically there is only one line for one host. Other types of host configurations, such as host Clusters, allow multiple hosts to be given access to the same volume.

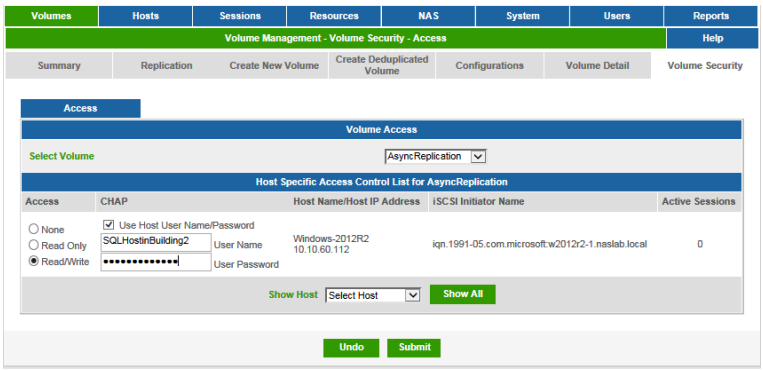


Figure 2-32 Volume Security Screen

2 Select a **Volume** from the volume list. The list of all hosts with access permission appears. Each host is marked with their access permissions.

3 In the **Access** column, click **None** for the host that should no longer have access to this volume.

Click **Undo** to revert to the saved settings.

4 Click **Submit** to remove access to this volume for this host.

2.4.4 Volume Detail

All essential information regarding the volume appears on the Volume Detail screen.

Active Sessions: The number of connections between the host and the currently selected volume

Volume Block Size: The size of the blocks on the volume in bytes

Volume Size (GB): The size of the volume in GigaBytes

Volume Type:

Span is a volume type for a managed volume.

Pass Thru is the volume type for volumes on resources designated as pass thru.

Mirror is the volume type for volumes with any mirror images.

Dedup or **Thin Volume** for volumes provisioned at deduplicated or thin pools.

Encryption: "Yes" or "No" based on the way the volume was created.

Image Type: Either Synchronous or Asynchronous.

iSCSI Target Name: Name assigned to the resource during configuration.

SnapShots (Reserved): The number of snapshots used in Asynchronous Mirrors is indicated inside the parentheses. The field does not appear for non-snapshot volumes.

Allowed Hosts: A numbered list of hosts who are allowed access to this volume. The Host is identified by its SAN IP Address.

MPIO enabled: Displays either Yes or No to indicate if MPIO is possible on this volume. This shows only for legacy volumes created prior to Release 4.2.

Operational State: A resource can have one of the following states:

OK: The *Storage Concentrator* and the resource have initiated a session.

Off-Line: The *Storage Concentrator* is not able to successfully initiate a session with this resource.

Ready: On pass thru resources only, this state displays if there is currently no active front-end session between this resource and a host.

If the volume is a mirror, additional information on the images in the mirror volume is displayed

- **Unit Test**: This field displays the results of the SCSI test unit ready command.

Each volume is described as it exists on the storage devices in the storage pool:

- **Segment Number**: Ordinal number indicating a portion of the volume's allocated space. The numbers are most often associated with volume expansions or using several resources in a Manual volume creation process.
- **Path**: The SCSI address of the Resource in the storage pool.
- **Resource Name**: Name assigned to the resource.
- **Segment Size (GB)**: Total number of GigaBytes allocated to this segment

from this particular resource.

- **Start Block**: Block number where the volume segment starts on the resource.
- **End Block**: Block number where the volume segment ends on the resource.

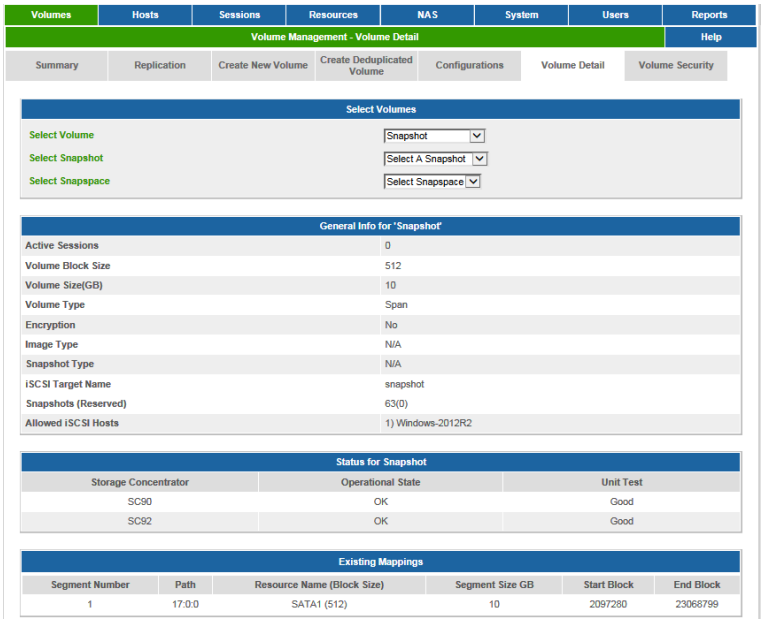


Figure 2-33 Volume Detail Screen

2.5 Hosts

In order for the *Storage Concentrator* to access the volumes that provide storage on the network, the *Storage Concentrator* must recognize a host. There are two ways hosts can be added so that the *Storage Concentrator* recognizes them.

One, the host may initiate the discovery of the *Storage Concentrator*. Refer to the vendor-specific documentation for information on initiating discovery of external devices. Two, hosts can be added manually in the *Storage Concentrator's* administrative interface.

2.5.1 Adding a Host

A host is added to the *Storage Concentrator* so that it can have access to the volumes which provide storage on the network.

To add a host using the *Storage Concentrator* administrative interface, use the steps that follow:

1 Click Hosts. The Host Management Summary screen appears, listing any previously recognized hosts.

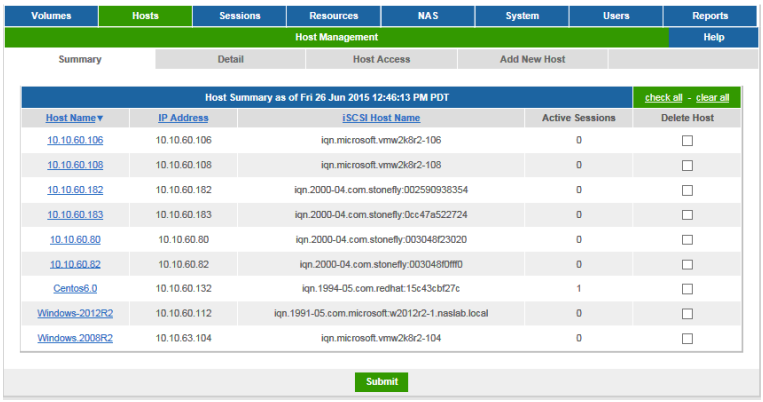


Figure 2-34 Host Management Summary Screen

Click **Host Name**, **IP Address**, or **iSCSI Host Name** in the column heading to change the sort order of the hosts.

2 Click **Add New Host**.

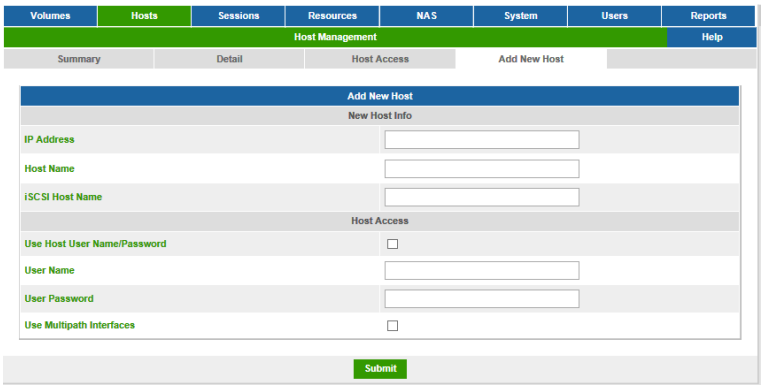


Figure 2-35 Host Management Add New Host Screen

3 Enter the following information for the host:

IP Address

If the security option is going to be selected in the Volume ACL's screen, the host's IP address must be the one that is associated with the iSCSI initiator's name.

Host Name (may be the same as the IP Address)

iSCSI Host Name (the host name specified in the initiator for the host)

If you want to specify a User Name and Password for this host click on Use Host User Name/ Password under Host Security. If you select this checkbox, you will need to type in the appropriate CHAP User Name and Password in the fields provided.

Some initiators only support CHAP at the Host level, and not at the Volume level. Some initiators support both. If you specify a CHAP secret at the host level, you may also specify a CHAP secret at the volume level only if your initiator supports this feature. For more information on Volume CHAP, see "Volume Security".

If there is a mismatch between the CHAP secret on the initiator and the *Storage Concentrator*, you will get a target error message on the host.

4 Click **Submit**.

You can also add future hosts in the Host Management screen. Future hosts are those that are not powered on, those that cannot initiate contact with the *Storage Concentrator*, or those that are not installed yet.

2.5.2 Editing a Host

The name and IP address of a host and the name of an iSCSI initiator can be edited from the Host Details screen.

To edit Host information, use the steps that follow:

1 On the Host Management screen, click **Detail**. The Host Management Detail screen appears.

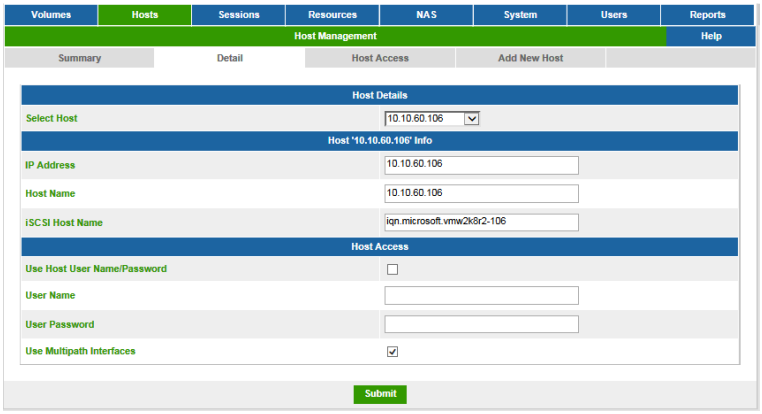


Figure 2-36 Host Management Detail screen

2 Select a **Host** from the host list.

3 Edit the fields as desired. They are:

IP Address — IP address assigned to the host. Edit as needed.

Host Name — Name assigned to the host appears here. Edit as needed.

iSCSI Host Name — Name assigned to the iSCSI host appears here. Edit as needed.

Use Host User Name/Password — If you have specified a CHAP secret on the Host Add page, you can modify it here. Or, if you did not specify a CHAP secret when you added the host, then you can do so here.

Select this option by selecting the checkbox if you are using Host CHAP. Under Host Security, click on Use Host User Name/Password if you want to specify a User Name and Password for this host. If you select this checkbox, you will need to type in the appropriate CHAP User Name and CHAP Secret in the fields provided.

Some initiators only support CHAP at the Host level, and not at the Volume level. Some initiators support both. If you specify a CHAP secret at the host level, you may also specify a CHAP secret at the volume level only if your initiator supports this feature.

If there is a mismatch between the CHAP secret on the initiator and the Storage

Concentrator, you will get a target error message on the host.

Do not select this option and do not enter a username or password for the Host if you will not be enabling CHAP Host Security.

User Name — The host name appears here. Edit as needed.

Password — Enter your CHAP Secret here. It must match the CHAP secret specified on your initiator. Edit as needed.

Use Multipath Interfaces — If you have configured your Storage Concentrator (SC) and this host to use multipath interfaces, you should set this check-box so that the multipath interfaces will be exposed to this host during iSCSI discovery.

The default is to not expose multipath interfaces to a host.

This field is not shown when there are no multipath interfaces configured on the SC.

Multipath interfaces are managed on the "System -> Network -> Local iSCSI Data Port" GUI page.

SC hosts that use this SC as an iSCSI target for campus mirror images, replication, etc. do not support multipath interfaces. This check-box will be disabled for SC hosts.

4 Click **Submit**.

2.5.3 Removing a Host

A host can be removed from the *Storage Concentrator* when it no longer requires access to storage volumes.

To remove a host, use the steps that follow:

1 Click **Hosts**.

2 On the Host Management Summary screen, click the check box under **Delete** for the host to be removed.

Click **Select All** if you wish to select all the hosts that are listed. Click **Clear All** to deselect all selected hosts.

Deleting a host will prevent it from having access to volumes managed by the *Storage Concentrator*.

3 Click **Submit** to remove the host.

2.5.4 Host Access

The Host Access Screen is designed to allow easy management of hosts and their access to volumes. You can change host security settings, access, and add new volumes to hosts.

To use host access, Click on Hosts, and then Host Access. The following screen appears:

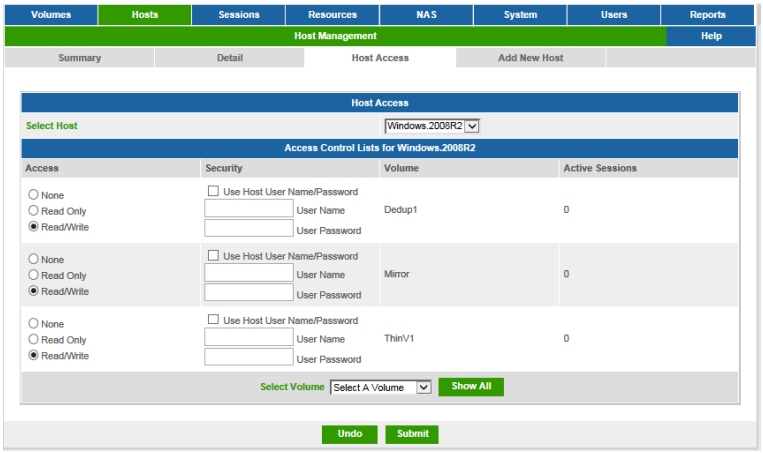


Figure 2-37 Host Access Screen

1 Select the Host desired using the drop down menu. If you select a Snapshot Live Volume, Snapshots will be displayed.

2 Select Volumes from the drop down menu, or Click the Show All button to display all volumes.

3 Change Security settings as desired.

4 Click **Submit** to save your changes.

2.6 Sessions

The Sessions screen displays all active sessions between hosts and the *Storage Concentrator* and allows a system administrator to arbitrarily end any session.

To display active sessions, use the steps that follow:

1 Click **Sessions**. The Sessions screen appears.

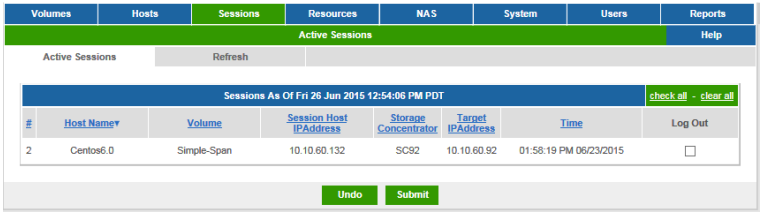


Figure 2-38 The following information is listed for each active session:

Host Name — The Storage Concentrator's (SC's) name for the host that initiated the session. Often, this name is based on the IP address that the host logs in from.

Volume — The volume name the iSCSI session is logged in to.

Session Host IP Address — The host's IP address that the host initiated the session from. Since hosts may have multiple network interfaces, the IP address may vary for sessions from the same host.

Storage Concentrator — The Storage Concentrator that the session is logged in to. In an SC cluster, this could be either SC due to the Active/Active feature. This column is only shown when there are two SC's in a cluster configuration. (Please consult the sections on Active-Active Load Balancing contained in Section 3: Failover.)

Target IP Address — The SC's target IP address that the session terminates on. In an SC cluster, or when there are SC multipath network interfaces defined, this address may vary. This column is only shown when there are two SC's in a cluster configuration, or there are SC multipath network interfaces configured.

Time — The date and time that the iSCSI session was established.

Logout — Click logout next to the session to be ended.

To end an active session, use the steps that follow:

1 Click L**ogout** next to the session to be ended.

2 Click **Submit** to end the session. The session ends.

As long as the host has not been removed, the host can immediately log in again if desired.

2.7 System

The System Management screen provides information regarding the system setup, routing network settings, and the configuration of the *Storage Concentrator* network settings for the iSCSI Host GbE port, Management GbE port, target portals, and default gateway. System configuration typically occurs at installation; the System Management screen allows these settings to be edited at a later time.

2.7.1 System Information

The System Management screen also facilitates setting up the FailOver function. FailOver is a process that automatically redirects user requests from a failed *Storage Concentrator* to the other *Storage Concentrator.* That *Storage Concentrator* takes over the operations of the failed system. For details on implementing the FailOver function, see "Setting up FailOver".

The *Storage Concentrator* can be rebooted or shut down from the System Management screen.

Diagnostic information that may be useful when troubleshooting hardware problems is also available from the System Management screen.

Software upgrades are a system management function and are covered in a separate section of this manual. For more information, see "Upgrading the Storage Concentrator Software".

The System Information screen is a view-only screen that displays the current settings of the *Storage Concentrator*.

To access the System Information screen, click **System**. The System Management Information screen appears.

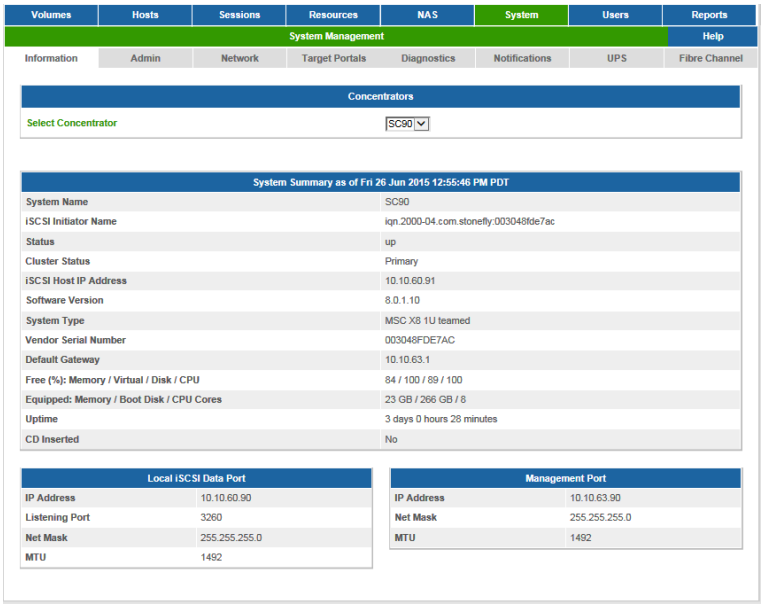


Figure 2-39 System Management Information screen

System Management information includes:

- System Name
- Status
- Cluster Status
- Software Version
- Serial Number
- Default Gateway
- iSCSI Host GbE Port settings
- Management Port settings
- Free Memory and CPU Utilization Percentages
- Uptime (in days, hours, and minutes)
- CD Contained in System Tray (Yes/No)

These settings can be modified after initial set up if there are changes to the network configuration.

2.7.2 Local iSCSI Data Port Settings

To edit the Local iSCSI Data Port settings, use the steps that follow:

1 Click System.

2 Click Network.

3 Click **Local iSCSI Data Port Settings**. The Local iSCSI Data Port Settings screen appears.

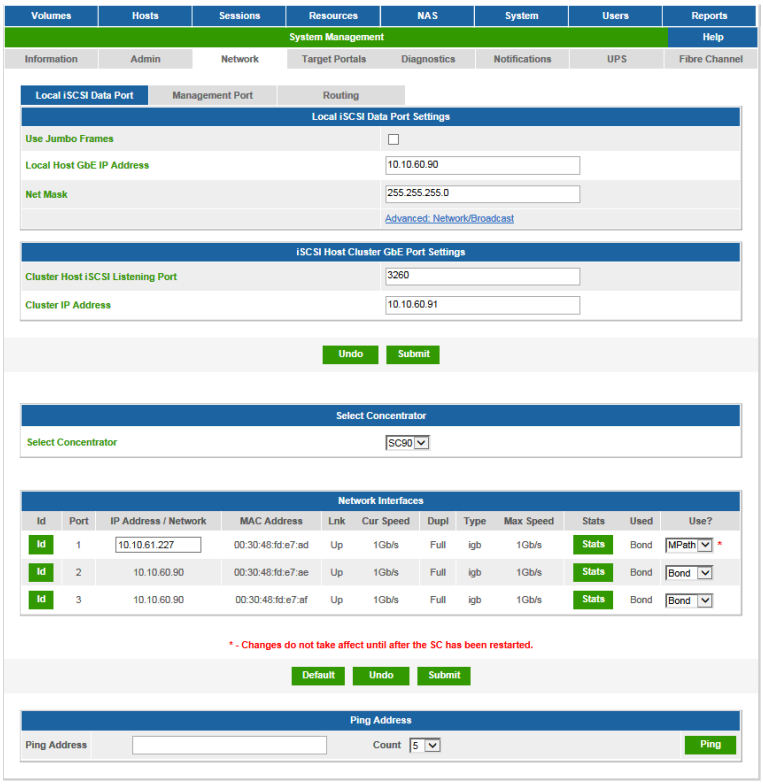


Figure 2-40 Network Local iSCSI Data Port Settings screen

2.7.2.1 Use Jumbo Frames

Jumbo Frames are a standard TCP/IP method of packing more data into each transmission unit. Ethernet has used 1514 byte frame sizes since it was created. The CRC adds another 4 bytes for a total of 1518. "Jumbo frames" extends Ethernet frame size to 8014 bytes (plus 4 CRC). The increase in frame size increases transmission performance, but networks must be capable of carrying the larger frame size. Many external networks are not able to process jumbo frames.

To use jumbo frames effectively, all devices on the network must be capable of processing them—in an iSCSI network, that includes initiators, switches, storage, etc. Some initiators are capable of supporting Jumbo Frames if they are set up with that option. If you intend to use Jumbo Frames on the Storage Concentrator, set all other devices such as initiators and switches to Jumbo Frames before selecting the Use Jumbo Frames option.

When using Jumbo Frames on the network, all equipment must be configured with a Frame size of 9014 excluding header size. If other equipment being used includes the header size in its setting that amount needs to be added to the frame size. For switches and V-LANs, 4 bytes must be added for CRC.

The Storage Concentrator can support Jumbo Frames by selecting the Use Jumbo Frames box. Only you can determine if your network can support Jumbo Frames.

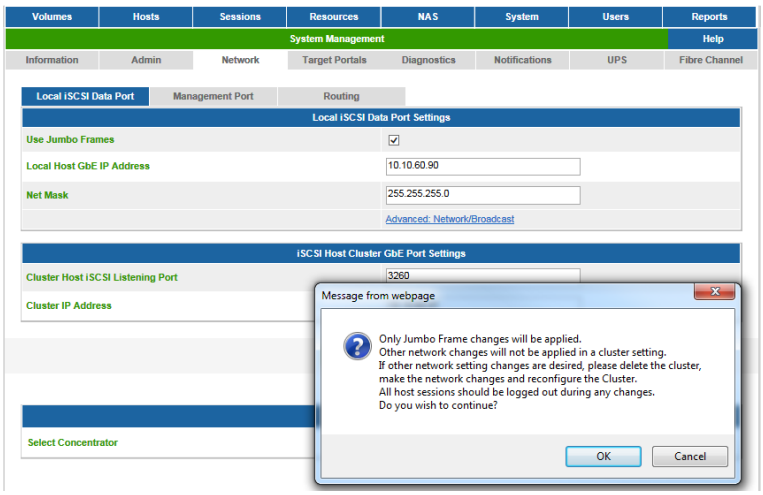


Figure 2-41 The display when Jumbo Frame is selected

Press **OK** to use Jumbo Frames.

2.7.2.2 IP Address and NetMask

If you are using the dual or quad GbE port option, there is only one IP Address required for the multiple ports. Plugging both ports into a GbE switch will provide the Adaptive Load Balancing feature across the two ports. No features or configurations are required in the GbE switch for this feature. See "Typical Configurations for using Dual iSCSI GbE Ports".

2.7.2.3 iSCSI Listening Port or Cluster Host iSCSI Listening Port

The iSCSI Listening Port will show up either under Local iSCSI Data Port Settings or under the iSCSI Host Cluster GbE Port Settings, depending on whether the Storage Concentrator is in a stand-alone or FailOver Cluster configuration. The standard setting is 3260.

The iSCSI Listening Port is set to the industry standard of 3260. If an iSCSI initiator is configured with the incorrect port number in the initiator's configuration software, it will not log onto the *Storage Concentrator*.

2.7.2.4 Network Interfaces

The number of SAN ports is adjustable. The SC may contain 2 or more possible SAN ports. A table of the Network Interfaces is displayed on this screen. The right-hand '**Use?**' column allows the Administrator to select the number of ports to be used in the SAN. The ports set to '**None**' are idle and do not carry IOs. The left-hand column includes an "**ID**" button for each Network Interface. The "**ID**" button flashes the lights on the selected port to allow cables to be inserted into the correct ports on the equipment.

Use? — The 'Use?' column shows and controls the current configuration state for each physical network interface. This column only appears when there is more than one physical network interface available. The choices are:

Bond - This interface will be used as one of the main SAN data network interfaces. These are the interfaces that are configured in 'Local iSCSI Data Port Settings' section of the page.

There can be multiple bonded interfaces, but there always must be at least one.

When there are multiple bonded interfaces, traffic is dynamically distributed amongst the interfaces to balance the load.

Note however that network bonding never uses more than one network interface to a single iSCSI host port at a time, although all bonded interfaces can be in use when there are multiple hosts.

Note that some SC platforms do not support network bonding -- for those, selection of only a single 'Bond' interface will be allowed.

None - This interface will not be used.

MPath - Short for multipath, this interface will be used as an extra interface for hosts that support iSCSI storage multipathing.

Multipath interfaces are added when there is a need for more than a single network interface bandwidth between the SC and a single iSCSI host.

Use of Multipath interfaces is never required. iSCSI hosts often need additional configuration to use multipath interfaces.

Each multipath interface is configured with its own unique IP network address (e.g. 192.168.1.0, not a full IP address). The full multipath interface address is formed from the host portion of the 'Local iSCSI Data IP Address ' setting.

When an interface is first changed to 'MPath', the network field will be blank. This must be changed to a valid network address before the interface can be used.

Multipath interfaces are only used by hosts that are configured for, and enabled to use them. The setting to enable their use is found on the "Hosts -> Detail" GUI page.

7 Click **Submit** to change the iSCSI Listening Port network settings.

2.7.2.5 Ping

Ping is available on the iSCSI Host GbE Port Network Settings, the Management Port Settings Screen and the Routing Screen. The feature works the same on all three pages. To ping a network IP address:

1 Enter the IP address in the field provided.

2 Enter the number of times you would like the ping to attempt.

If you are using "Jumbo Frames" (see "Local iSCSI Data Port Settings"), then you may Ping in Jumbo Frames. This option only shows up if you have already selected the Jumbo Frames option for the Local iSCSI Data Port setting. Click the checkbox if you would like to Ping using Jumbo Frames.

3 Click on the Ping button to execute the ping. The *Storage Concentrator* will report back whether the ping was successful.

2.7.3 Network and Broadcast IP Settings

The *Storage Concentrator* automatically configures the network and broadcast IP settings based on the IP address and netmask settings. These settings can be manually changed through the Advanced: Network/Broadcast screen.

To edit the network or broadcast IP settings, use the steps that follow:

1 Click **System**.

2 Click **Network**.

3 Click **iSCSI Data Port** or **Management Port**.

4 On the iSCSI Data Port screen or the Management port screen, click **Advanced: Network/Broadcast**.

The following screen appears when Advanced: Network/ Broadcast is selected from the Local iSCSI Data Port Settings screen.

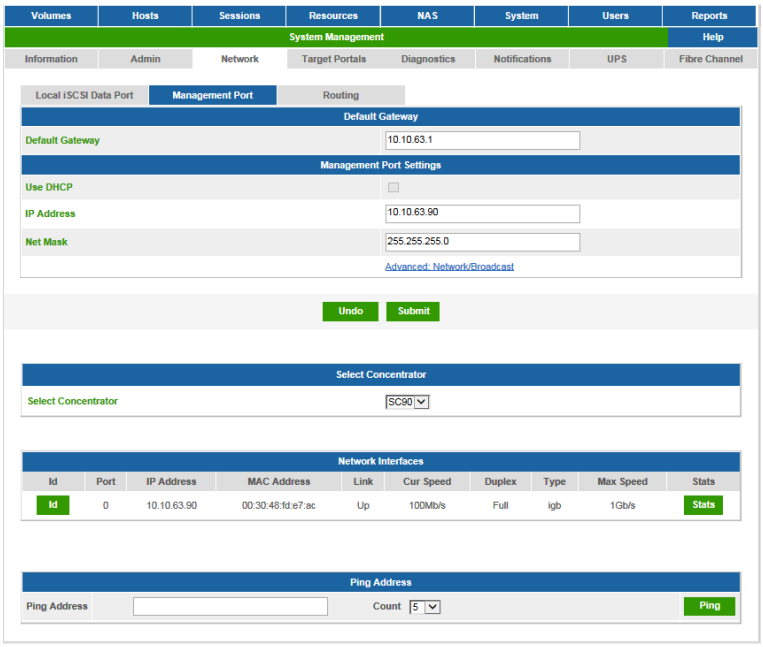


Figure 2-42 System Management, Network and Broadcast IP settings screen

The *Storage Concentrator* must be rebooted after changing network settings. If you have completed your changes, reboot now. Otherwise, continue with your changes and reboot when finished. For more information on rebooting, see "Rebooting the Storage Concentrator".

2.7.3.1 Management Port and Default Gateway Settings

Network settings can be modified after initial set up if there are changes to the network configuration.

Enter the appropriate values in these fields:

Default Gateway

IP Address

NetMask

Click **Undo** to revert to the saved settings.

Click **Submit** to send the configuration information to the database.

The following screen might appear.

Depending on the speed of your browser connection, you may not see this screen. Your changes will still take effect, however you will need to manually set your browser's URL to point to the new IP address.

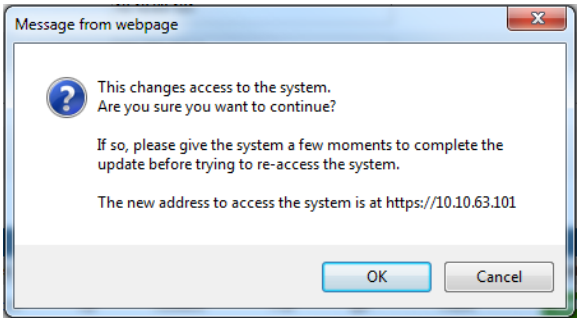


Figure 2-43 System Management Port Change screen

Click the new IP address to confirm the change to the Management GbE Port setting.

2.7.3.2 Network Interfaces

The Network Interfaces table is to provide information about the Management Port. The "ID" button may be used to flash the lights of the correct port on the equipment. In addition, some basic network statistics are available by clicking on "Stats". No selections are required for the Management port. (See the description of the iSCSI Data Port)

The *Storage Concentrator* automatically configures the network and broadcast settings based on the IP address and NetMask settings. The network and broadcast settings can be manually changed through the **Advanced: Network/ Broadcast** screen. For more information on editing the network and broadcast settings, see "Network and Broadcast IP Settings".

2.7.4 Routing

To access hosts on other networks, routing information to those networks must be configured in the System Management Routing screen. If the host to be communicated with has a network setting different from the one listed in the iSCSI Host GbE screen a route must be added. For example, if the *Storage Concentrator* network setting is 26.34.128.50 and the host network setting is 106.39.212.6, a route to the host must be configured.

A route to a host that is no longer needed can be deleted from the routing screen.

To configure the routing information, use the steps that follow:

1 Click **System**.

2 Click **Network**.

3 Click **Routing**. The Routing screen appears.

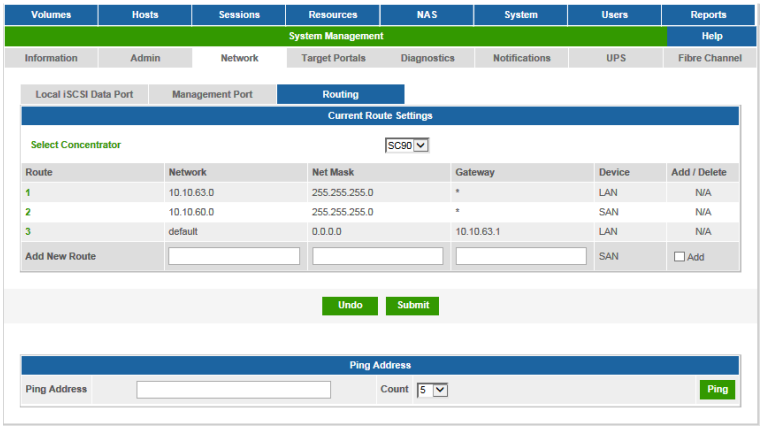


Figure 2-44 System Management, Routing screen

4 For Storage Concentrator FailOver Clusters, select the Concentrator IP address from the drop down menu. If changing the routing information for a FailOver Cluster, you must change the Routing information for both the Primary unit and the Secondary unit.

5 In the **Add New Route** fields, enter the **Network**, **Netmask** and **Gateway** settings for the new route.

6 Click the **Add** check box.

7 Click **Submit** to configure the routing information.

To remove a route from the Storage Concentrator, use the steps that follow:

1. Click the delete radio button next to the host whose routing information is to be removed.

Only host routes that do not specify the network itself can be deleted, and only one host route can be selected for deletion at a time.

5 Click **Submit** to remove the route. The *Storage Concentrator* is updated to remove the host route from the database.

2.7.5 Administrative Functions

Administrative functions that are available from the System Management screen include assigning the *Storage Concentrator* a new system name, changing the default number of log

records for the database, rebooting and shutting down the system, and setting the date and time. Performing software upgrades and saving and restoring user configuration information are done through the Admin screen.

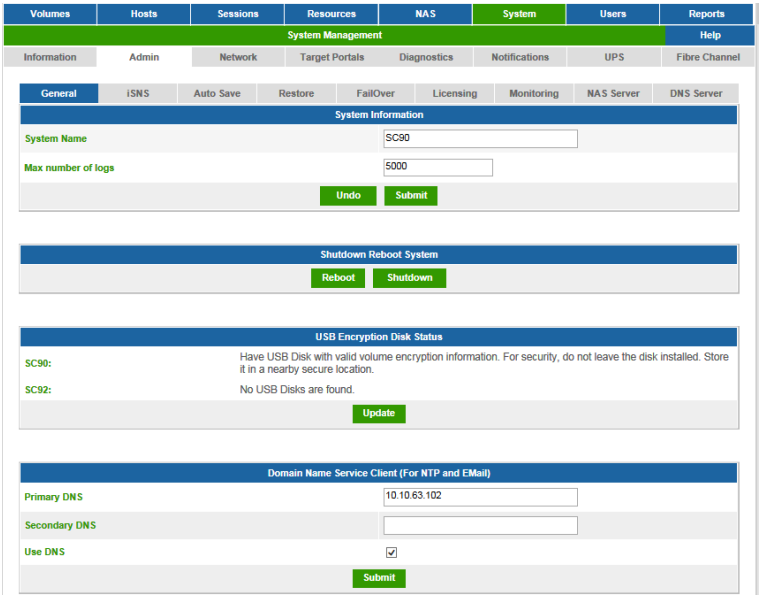


Figure 2-45 System Management, Admin screen – part 1

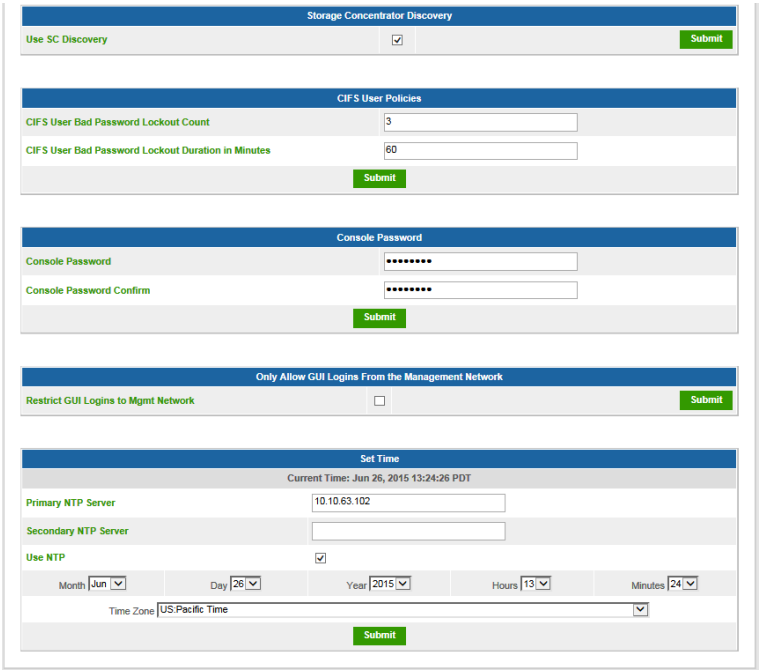


Figure 2-46 System Management Admin screen – part 2

To change the system name and the number of log records, use the steps that follow:

1 Click **System**.

2 Click **Admin**. The System Management Admin screen appears.

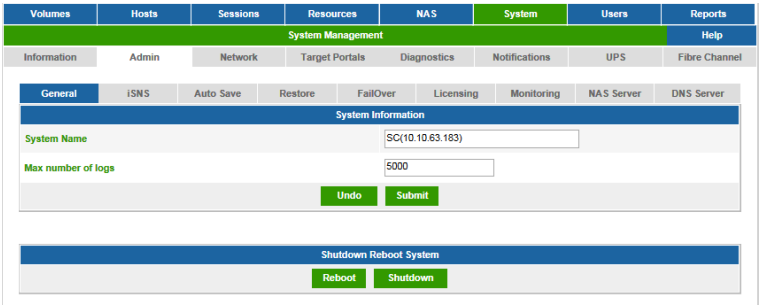


Figure 2-47 System Management Admin screen

3 Edit the **System Name** for the *Storage Concentrator*, if desired.

4 Change the number of records for the database in the **Max # of logs** field, if desired.

The maximum number of logs is the number of log records kept in the database. When this number is reached, the oldest record is overwritten when a new record is added. This prevents the log table from using up too much disk space.

5 Click **Reboot**.

The *Storage Concentrator* must be rebooted for the new log settings to be recognized.

2.7.6 Rebooting the Storage Concentrator

The system will require rebooting after network configuration changes, and following restoring user configuration data.

To reboot the system, use the steps that follow:

1 Click **System**.

2 Click **Admin**. The System Management Admin screen appears.

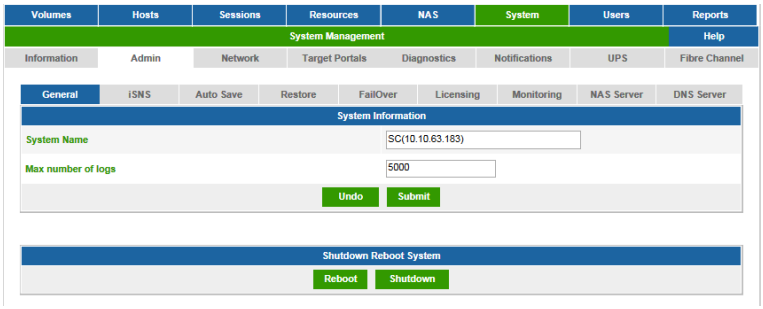


Figure 2-48 System Management Admin screen

3 Click **Reboot**.

2.7.7 Shutting Down the Storage Concentrator

The *Storage Concentrator* must be shut down when adding or reconfiguring resources.

To shut down the system, use the steps that follow:

1 Click System

2 Click Admin.
The System Management Admin screen appears as above.

3 Click Shutdown. (Unlike other *Storage Concentrators*, the High Availability Storage Concentrator, or HSC, does not power off during a shutdown. Wait until all activity lights are quiet and use the power switch to turn it off.)

Unlike a reboot, a shutdown of the Storage Concentrator will require physical access to power it back on.

2.7.8 USB Device Status

USB port statuses are reported on the System->Admin GUI page. Most of the reports are self-explanatory:

USB port statuses for primary SC:

- Have Encryption Keys Disk with valid volume encryption information.
- Have Encryption Keys Disk with valid volume encryption information. For security, do not leave the disk installed. Store it in a nearby secure location.
- Have Encryption Keys Disk that can be used to store volume encryption information.
- Have Encryption Keys Disk that is dedicated to another StoneFly service. Please insert

another disk.
- Access to Encryption Keys disk is blocked by another StoneFly service. Please repeat request later.
- No Encryption Keys Disks are found. Insert disk with encryption information into this system when needed.
- No Encryption Keys Disks are found. Please insert disk with encryption information.
- More than one Encryption Keys Disk with volume encryption information is detected. Please use no more than one encryption disk.
- Have only a non-DOS formatted USB Disks. Please use DOS formatted disks.
- Have Encryption Keys Disk with invalid volume encryption information. Please use valid disk.
- Have Encryption Keys Disk without any volume encryption information. Please use valid disk.
- Have Encryption Keys Disk with encryption information for invalid volumes. Please use appropriate disk.
- Encryption Keys Disk lost some of the encryption information but can be repaired.
- Have valid Encryption Keys Disks on different clustered Storage Concentrators. Please remove Encryption Keys Disk from the Secondary Storage Concentrator {<SC_Name>}.

USB port statuses for secondary SC:

- Have Encryption Keys Disk with valid volume encryption information. For security, do not leave the disk installed. Store it in a nearby secure location.
- Have Encryption Keys Disk that can be used by any StoneFly service.
- Have Encryption Keys Disk that is dedicated to another StoneFly service.
- Access to Encryption Keys disk is blocked by another StoneFly service.
- No USB Disks are found.
- More than one Encryption Keys Disk with volume encryption information is detected. Please use no more than one encryption disk.
- Have only a non-DOS formatted USB Disks. Please use DOS formatted disks.
- Have Encryption Keys Disk with invalid volume encryption information.
- Have Encryption Keys Disk without any volume encryption information.
- Have Encryption Keys Disk with encryption information for invalid volumes.
- Encryption Keys Disk lost some of the encryption information but can be repaired.
- Have valid Encryption Keys Disks on different clustered Storage Concentrators. Please remove Encryption Keys Disk from the Secondary Storage Concentrator {<SC_Name>}.
- Have Encryption Keys Disk with valid volume encryption information. Please remove this disk.

If these reports are displayed please consult StoneFly Support:

- Unknown problem with USB device.
- Unknown problem with database.
- Failure to assign USB device to encryption service.
- Invalid argument.
- Failure to copy encryption information to USB Disk.
- Failure to delete volume encryption information from USB Disk.
- Invalid command.
- Failure to create digest for encryption password. Please contact customer support to resolve this problem.
- Volume encryption password does not match password provided during volume

creation. Please contact customer support to resolve this problem.
- Failure to format USB Disk.
- Failure to access secondary SC.

GUI Log messages also report the status of the USB port:

- SC {<SC_NAME>}: have Encryption Keys Disk with valid volume encryption information.
- SC {<SC_NAME>}: have Encryption Keys Disk that can be used to store volume encryption information.
- SC {<SC_NAME>}: have Encryption Keys Disk that is dedicated to another StoneFly service. Please insert another disk.
- SC {<SC_NAME>}: access to USB port is taken by another StoneFly service. Please repeat request later.
- SC {<SC_NAME>}: no USB Disks are found. Please insert disk with encryption information.
- SC {<SC_NAME>}: more than one Encryption Keys Disk with volume encryption information are detected. Please use no more than one encryption disk.
- SC {<SC_NAME>}: have only non-DOS formatted USB Disks. Please use DOS formatted disks.
- SC {<SC_NAME>}: have Encryption Keys Disk with invalid volume encryption information. Please use valid disk.
- SC {<SC_NAME>}: unknown problem with USB device.
- SC {<SC_NAME>}: unknown problem with database.
- SC {<SC_NAME>}: failure to assign USB device to encryption service.
- SC {<SC_NAME>}: invalid argument.
- SC {<SC_NAME>}: failure to copy encryption information to Encryption Keys Disk.
- SC {<SC_NAME>}: failure to delete volume encryption information from Encryption Keys Disk.
- SC {<SC_NAME>}: invalid command.
- SC {<SC_NAME>}: have USB Disk without any volume encryption information. Please use valid disk.
- SC {<SC_NAME>}: have Encryption Keys Disk with encryption information for invalid volumes. Please use appropriate disk.
- SC {<SC_NAME>}: failure to create digest for encryption password. Please contact customer support to resolve this problem.
- SC {<SC_NAME>}: volume encryption password does not match password provided during volume creation. Please contact customer support to resolve this problem.
- SC {<SC_NAME>}: failure to format USB Disk.
- SC {<SC_NAME>}: Encryption Keys Disk lost some of the encryption information but can be repaired.
- SC {<SC_NAME>}: failure to access secondary SC.
- SC {<SC_NAME>}: have valid Encryption Keys Disks on different clustered Storage Concentrators. Please remove Encryption Keys Disk from the Secondary Storage Concentrator.
- SC {<SC_NAME>}: have Encryption Keys Disk with valid volume encryption information. It's insecure to keep disk in the USB port permanently.
- SC {<SC_NAME>}: more than one Encryption Keys Disk with volume encryption information are detected. It's insecure to keep disk in the USB port permanently.
- SC {<SC_NAME>}: have Encryption Keys Disk with encryption information for invalid volumes. It's insecure to keep disk in the USB port permanently.
- SC {<SC_NAME>}: Encryption Keys Disk lost some of the encryption information but can be repaired. It's insecure to keep disk in the USB port permanently.
- SC {<SC_NAME>}: have valid Encryption Keys Disks on different clustered Storage

Concentrators. It's insecure to keep disk in the USB port permanently.

2.7.9 Using DNS

The *Storage Concentrator* supports DNS (Short for Domain Name System or Service or Server), an Internet service that translates domain names into IP addresses. Because domain names are alphabetic, they're easier to remember. The Internet however, is really based on IP addresses. Every time you use a domain name, therefore, a DNS service must translate the name into the corresponding IP address.

The Storage Concentrator supports DNS only for NTP and SMTP services. DNS for data or management traffic is not supported.

To use DNS, you must provide the IP address of the DNS server.

To Use DNS, use the steps that follow:

1. Enter a primary IP Address in the field provided.
2. Entering a secondary IP address for an additional DNS server is optional.
3. Select the Checkbox "Use DNS."
4. Click Submit

2.7.10 Storage Concentrator Discovery

The Storage Concentrator supports the ability to discover neighboring SC's for more convenient GUI access to them. Discovered SC's appear on the GUI home page.

SC Discovery is achieved through the use of the Service Location Protocol (SLP) which uses broadcast and multicast IP frames over the management network.

SC Discovery is enabled by default. However, should the installation or network environment be such that this feature is unnecessary, or undesired, it may be disabled by clearing the **Use SC Discovery** check-box and clicking Submit.

Note: Only Storage concentrators that are operational are discovered and displayed. SC's running a SW version that does not support SC Discovery will not be detected. SC's that have **Use SC Discovery** disabled will not be discovered by other SC's.

2.7.11 CIFS User Policies

2.7.11.1 CIFS User Bad Password Lockout Count

For NAS Volumes using the CIFS protocol, the CIFS account can be automatically locked out after the specified number of bad password login attempts. Once the account is locked out, further attempts to login will fail until the lockout duration period has expired, or the administrator manually unlocks the account. Use '0' to never lockout the account.

Changing this setting takes immediate effect, an SC reboot is not required. This field only appears when the NAS feature is licensed.

A locked out NAS Volume account is indicated on the NAS Volume Management Configure Volume GUI page, and can be unlocked by the administrator there.

2.7.11.2 CIFS User Bad Password Lockout Duration in Minutes

For NAS Volumes using the CIFS protocol, the CIFS account can be automatically locked out for the specified number minutes after a number of bad password login attempts.

Once the account is locked out, further attempts to login will fail until this lockout duration period has expired, or the administrator manually unlocks the account. Use '-1' to never reset the lockout, and require that the administrator reset the account.

Changing this setting takes immediate effect, an SC reboot is not required. This field only appears when the NAS feature is licensed.

A locked out NAS Volume account is indicated on the NAS Volume Management Configure Volume GUI page, and can be unlocked by the administrator there.

2.7.12 Console Password

Allows for the Storage Concentrator system console and serial port CLI menu login 'console' password to be changed to value different from the published default.

If the password is not changed here, it remains the default value.

In an SC cluster, this setting applies to both concentrators -- it is a global setting.

Changing this setting takes immediate effect; an SC reboot is not required.

2.7.13 Only Allow GUI Logins From the Management Network

Restrict GUI Logins to Mgmt Network-- By default, the Storage Concentrator allows GUI login sessions over both the Management and the SAN networks.

There are cases where an installation may need to restrict the GUI logins to only be allowed from the Management network.

In an SC cluster, this setting applies to both concentrators -- it is a global setting.

Changing this setting takes immediate effect; an SC reboot is not required. However, the browser session will be reset, and a GUI page refresh or a login may be required.

2.7.14 Setting the Time and Using NTP Services

To change the date and time settings, use the steps that follow:

1 Click **System**.

2 Click **Admin**. The System Management Admin screen appears.

3 In the **Set Time** area of the screen, change the month, day, year, hour, minutes, and time zone settings as needed.

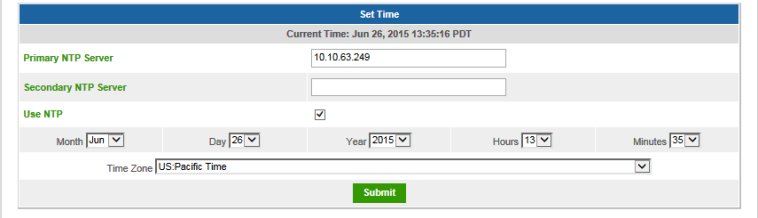


Figure 2-49

4 Click **Submit**.

Be sure to use the 24 hour clock when setting the hour. For example, 12:00 p.m. is hour 12, 1:00 p.m. is hour 13, and 2:00 p.m. is hour 14, and so on.

2.7.14.1 Using an NTP Server:

The Storage Concentrator can utilize an NTP (Network Time Protocol) server to set its internal clock. NTP is an Internet standard protocol that assures accurate synchronization to the millisecond of computer clock times in a network of computers. An NTP Server synchronizes client workstation clocks to the U.S. Naval Observatory Master Clocks. Running as a continuous background on the Storage Concentrator, NTP sends periodic time requests to the NTP server, obtaining time stamps and using them to adjust its clock.

The Storage Concentrator can't be used as an NTP server. It accesses the specified NTP server on the network to set its internal clock.

1 Enter the address of the NTP server. If you have specified a DNS server, you may enter the name of the DNS server rather than an IP address.

2.7.15 iSNS

Internet Storage Name Service (iSNS) facilitates automated discovery, management and configuration of iSCSI devices on a TCP/IP network. In any storage network, hosts (initiators) need to know which storage resources (or targets) they can access.

An Internet storage name server lets servers automatically identify and connect to authorized storage resources.

To use iSNS:

1 Click on the checkbox on the iSNS page to select it.

2 Enter the IP address of the iSNS server.

3 Click **Submit**.

To turn iSNS off, uncheck the box and click **Submit**.

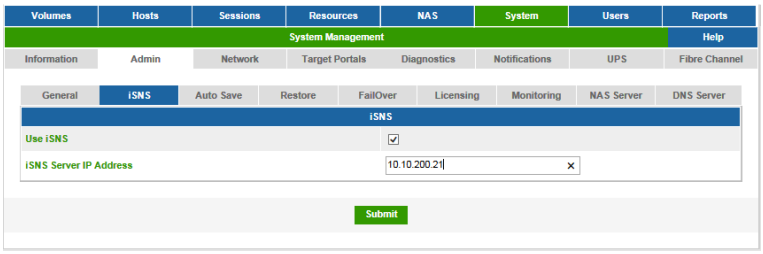


Figure 2-50 iSNS Screen

2.7.16 Auto Save

The Auto Save feature provides backup of user configuration data in the unlikely event that the *Storage Concentrator* goes down. Once Auto Save is set up, any configuration changes will automatically copy the information to an FTP site or to a local device such as the floppy drive or a USB Flash drive.

To Auto Save to an FTP site, use the steps that follow:

1 From the **System** screen, click **Admin**.

2 Click **Auto Save.** The System Management Auto Save screen appears.

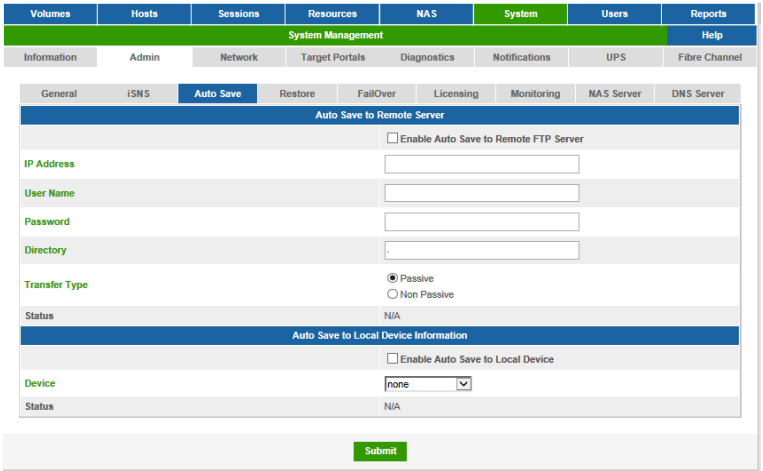


Figure 2-51 System Management Auto Save screen

3 Click the check box for **Enable Auto Save to Remote FTP Server** and then enter the appropriate information into these fields:

IP Address: IP address of the remote FTP server for backup

User Name: The FTP user name

Password: The FTP password

Directory: The FTP directory path to the location where the configuration data is going to be saved. The data is saved to the filename StoneFlyDB.tar.gz. The data is encrypted during backup and decrypted during restore.

This file should be saved to a secure location, and made accessible only by the user dedicated for backup.

In an environment where there is more than one installed Storage Concentrator, it is imperative that each *Storage Concentrator* has its own directory path to store the FTP information.

4 Click **Submit** to initiate the Auto Save process to an FTP site.

The Auto Save process is initiated immediately upon clicking submit.

The status field shows the current status of the FTP server. It will display all FTP messages, including Alert messages in case of failure.

To Auto Save locally to a floppy disk or USB Flash drive, use the steps that follow:

1 From the **System** screen, click **Admin**.

2 Click **Auto Save**. The System Management Auto Save screen appears.

3 Remove the front bezel and insert a floppy disk into the floppy drive on the front of the *Storage Concentrator* or insert a USB Flash disk into any available USB port. Do not use both the USB Flash disk and the floppy.

If you are using a *Storage Concentrator* FailOver Cluster, the same type of storage device must be inserted into BOTH the Primary and Secondary *Storage Concentrators*. Otherwise the database backup will fail depending on which unit is Primary.

4 Click the check box for Enable Auto Save to Local Device.

5 Choose the appropriate device from the device list.

6 Click **Submit** to initiate the Auto Save process to a local device.

The Auto Save process is initiated immediately upon clicking submit.

7 The status field shows the current status of the device.

2.7.17 Restore

Restore allows the user configuration data that was saved with the Auto Save function to be retrieved.

To retrieve user configuration data from the FTP server, use the steps that follow:

1 From the **System** screen, click **Admin**.

2 Click **Restore.** The System Management Restore screen appears.

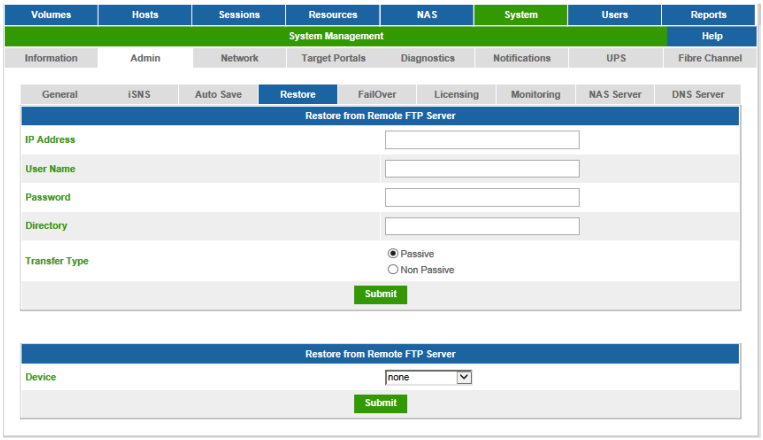


Figure 2-52 System Management, Restore screen

3 Enter the appropriate information into the **Restore from Remote FTP Server** fields:

IP Address: IP address of the remote FTP server for backup

User Name: The FTP user name

Password: The FTP password

Directory: The FTP directory path to the location from where the data will be retrieved. The filename is StoneFlyDB.tar.gz

The status field shows the current status of the FTP server.

3 Click **Submit** to initiate the restore process.

Following the restore process, the *Storage Concentrator* automatically reboots.

To retrieve user configuration data from a local device, use the steps that follow:

1 From the **System** screen, click **Admin**.

2 Click **Restore**. The System Management Restore screen appears.

3 Remove the front bezel of the *Storage Concentrator*. Insert the floppy disk that was used in the Auto Save process into the floppy drive on the front of the *Storage Concentrator* or insert a USB Flash disk into any available USB port.

4 Choose the appropriate device from the device list.

The status field shows the current status of the device.

5 Click **Submit** to initiate the restore process. Following the restore process, the *Storage Concentrator* automatically reboots.

2.7.18 Feature Licensing

After StoneFusion release 6.0.1.x some features are individually licensed. The process requires a System Serial Number to create the keys. A license key is specific to only one feature on one Storage Concentrator. The keys are not transferable to any other system. Any change to the Serial Number requires a new license key.

All SC systems now require a "StoneFusion Base OS" license. Otherwise they are considered as time limited evaluation systems. New systems are pre-licensed during production installation. Upgrades from older releases are implicitly "StoneFusion Base OS" licensed.

Most SC systems now require an "Advanced Features" license. This license is needed to enable support for SC Clustering, Synchronous Mirrors, Campus Mirrors, Remote Managed Storage, etc. The "Advanced Features" license is also a prerequisite for other licensed features such as "Asynchronous Replication", "Thin Volumes", "Deduplicated Volumes" and "Volume Encryption". Storage Concentrator Systems without "Advanced Features" licenses are generally simple, application specific systems that have no need these features. New systems are pre-licensed for "Advanced Features" during production installation. Upgrades from older releases are implicitly "Advanced Feature" licensed.

Features can now be licensed for Evaluation. Evaluation licenses allow a feature to be used for a limited time before a purchase decision is made. Contact StoneFly Customer Support to obtain Full or Evaluation Feature Licenses for your Storage Concentrator.

The NAS Volumes license couldn't be enabled on cluster of one or clustered systems. To do this, have to delete cluster, enable NAS volumes license on both systems and re-create the cluster.

The Max Hosts license limits number of hosts that can be added to the list of available hosts and number of hosts that can perform iSCSI discovery on the system. Hosts that are resident on other Storage Concentrators are not included into this limit.

The Max Provisioned Space license limits total size of resource space that is used to allocate:

- volumes, deduplicated and thin pools that are provisioned as a spans;
- volume's synchronous images, deduplicated and thin images are excluded;
- images of thin and deduplicated pools;
- pass thru volumes;
- NAS volume segments;
- NAS metadata segments.

The Max Provisioned Space license does not limit space allocated by:

- snap spaces;
- system metadata volume

and it does not include the size of:

- thin and deduplicated volumes;
- thin and deduplicated images;
- asynchronous images;
- snapshots;
- NAS volumes;
- NAS metadata volume;
- NAS volume snapshots;
- Flash Cache Caching devices.

Upgrades from older releases are implicitly "Max Provisioned Space" and "Max Hosts" unlimited licensed. Host access to system with expired "Max Provisioned Space" or "Max Hosts" Evaluation license is disabled.

1 The licensing feature is accessed on the System->Admin screen as shown in the partial screen below. Click on the Licensing link to display the screen on which the license keys are entered.

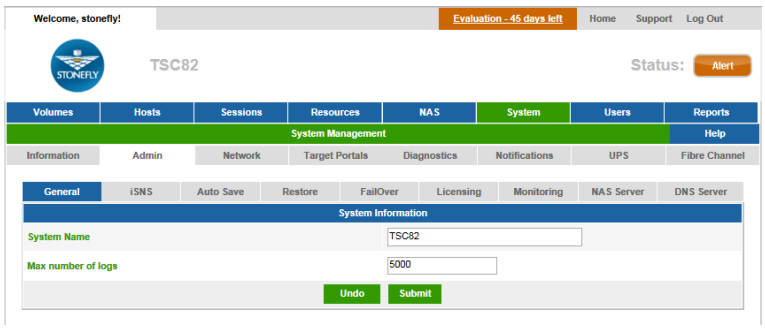


Figure 2-53 System Management, System Information Screen

2 The Licensing screen displays the System Serial Number that must be submitted to StoneFly's Support Team to generate the keys. A newly installed Storage Concentrator may have license keys pre-installed. In that case the word "Licensed" will appear, along with the key assigned. If the key is for an evaluation, the word "Eval" will appear instead of "Licensed", along with the number of days remaining in the evaluation

period.

3 The "NAS Volumes" and "Deduplicated Volumes" licenses can be activated only on systems that have specific size of RAM and boot disk space. Check "Storage Concentrator Configuration Limits".

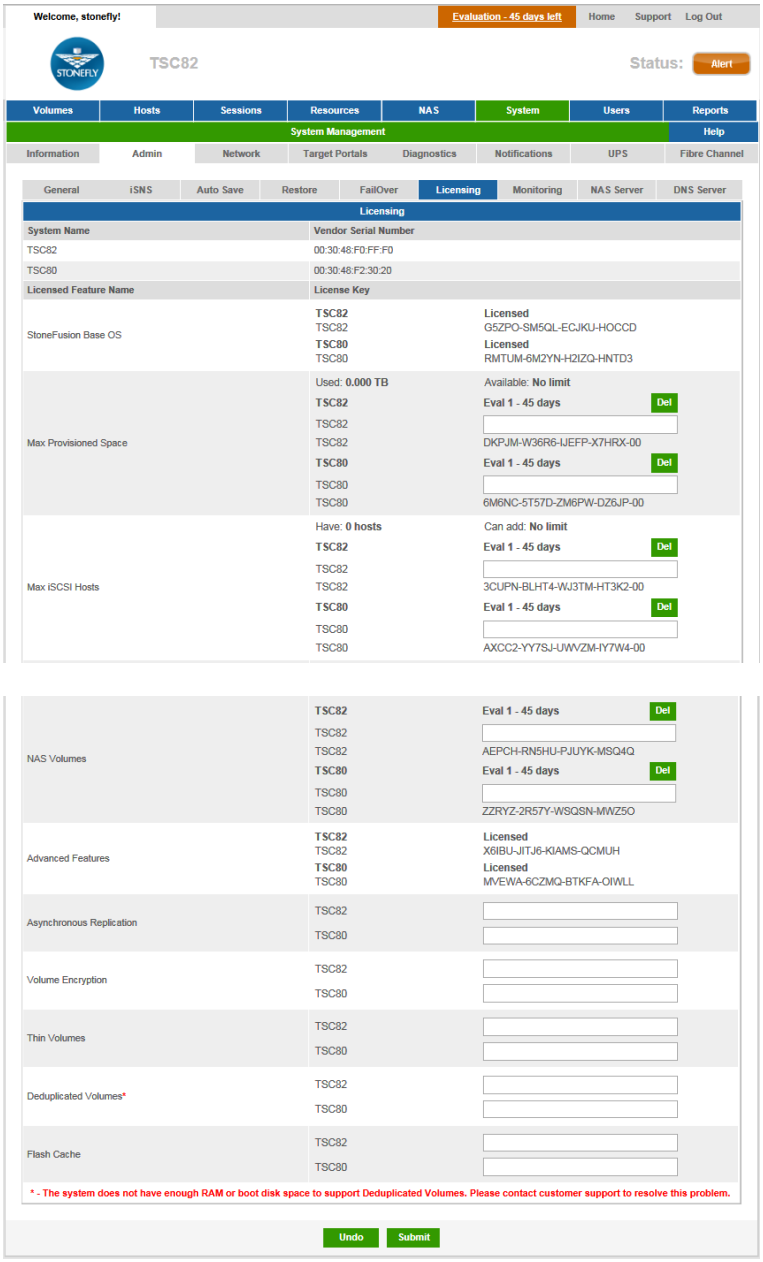


Figure 2-54 System Management, Licensing Screen

4 Contact StoneFly Support to give them the information displayed on this screen. Support returns a key for each authorized feature.

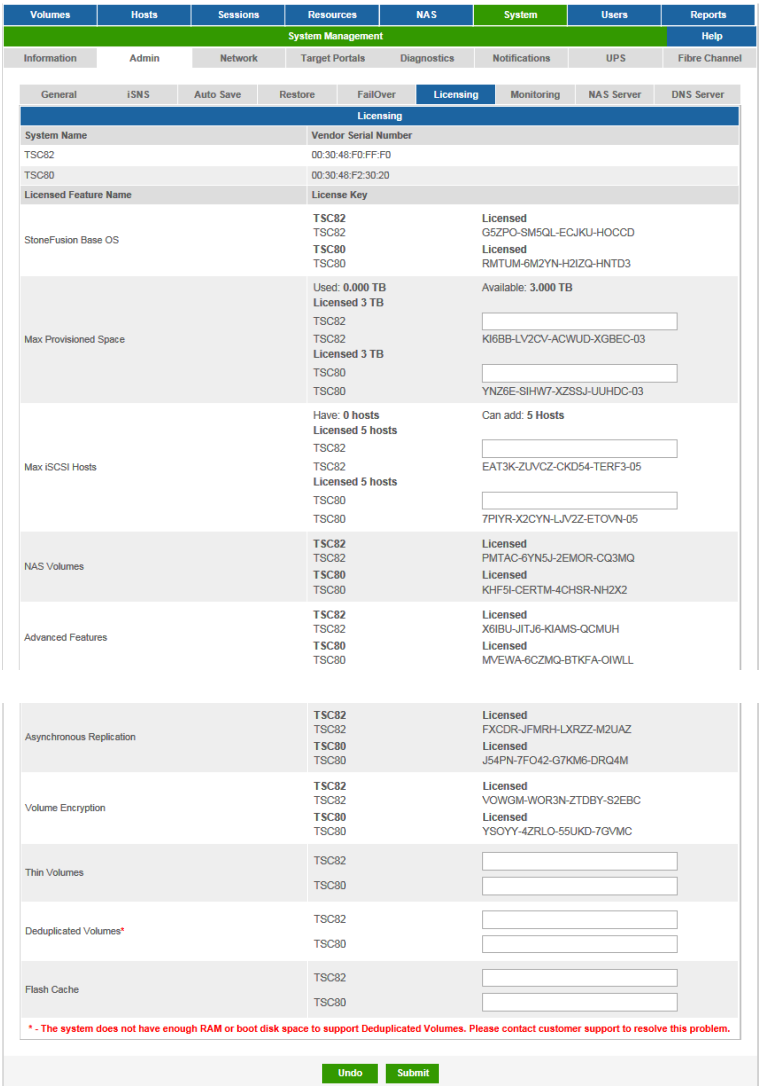


Figure 2-55 System Management, Licensing Screen-Filled

5 As shown above, enter the complete key into the appropriate field. It is not necessary to enter all the keys at the same time. Some features may not be authorized. Leave unauthorized key fields blank. Click **Submit**.

6 If the key is correctly entered, the word "Licensed" will appear, along with the key assigned. If the key is for an evaluation, the word "Eval" will appear instead of "Licensed", along with the number of days remaining in the evaluation period. See the following screen image as an example. A key for other features may be entered later. Please save a copy of the key values in a safe place.

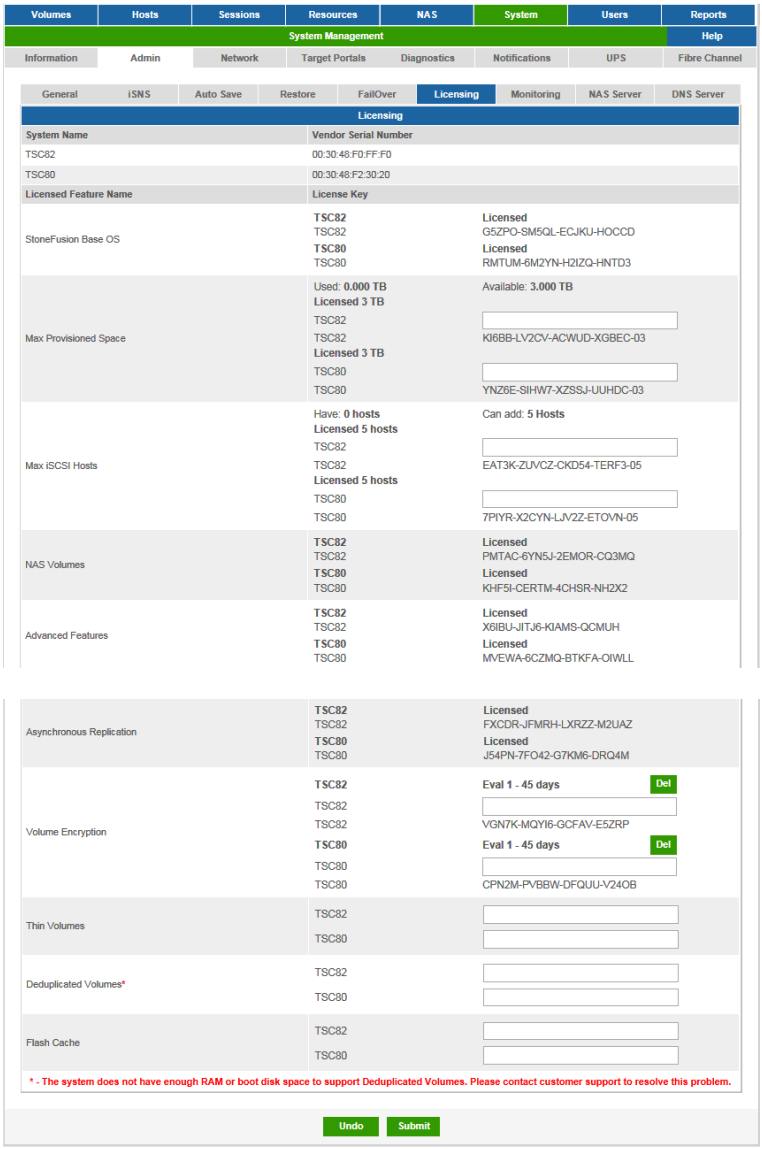


Figure 2-56 System Management, Licensing Screen-Eval

7 When the system is operating **any** Evaluation Licenses, the status line on each GUI page will indicate this along with the number of days left on the soonest to expire evaluation license. Clicking on the status will take you directly to the Licensing page. Note that a when an evaluation license expires, that features operation becomes inhibited.

There are some considerations regarding Licensed Features for Storage Concentrator systems working together in a clustered configuration:

A Feature License Key must be obtained for each Storage Concentrator system in the cluster.

In a cluster, both Storage Concentrator "Vendor Serial Numbers" should be provided when obtaining Feature License Keys. The two generated license keys are different and must be entered into the correct field.

The Feature License Keys needs to be manually entered for both of the two Storage Concentrator systems. This may be done before the two systems are clustered, or afterwards.

For two Storage Concentrators to be allowed to be clustered, they must both have the same feature licenses. An attempt to form the cluster will be blocked when this is not the case. Max Provisioned Space and Max Hosts licenses require that both Storage Concentrators have the same licensed amount, or that both systems should not be licensed.

It is not required that the Storage Concentrator system that was issued the license key be the primary storage concentrator when the key is entered. A failover is not required to enter a License Key.

Until **both** SC's have valid, non-expired, license keys entered, the feature is disabled.
The GUI title bar indicates "**LICENSE PROBLEM**" when one or more features are disabled due to asymmetric licensing.
A footnote appears on the GUI Licensing page also indicating this situation.

When Evaluation licenses are used, the first one to expire for a feature on either Storage Concentrator disables the feature.

When a licensed Storage Concentrator is removed from a cluster, its database is cleared, but attempts are made to retain the feature license records.
There are cases where the loss of Feature License settings can occur, and the license keys must then be re-entered.

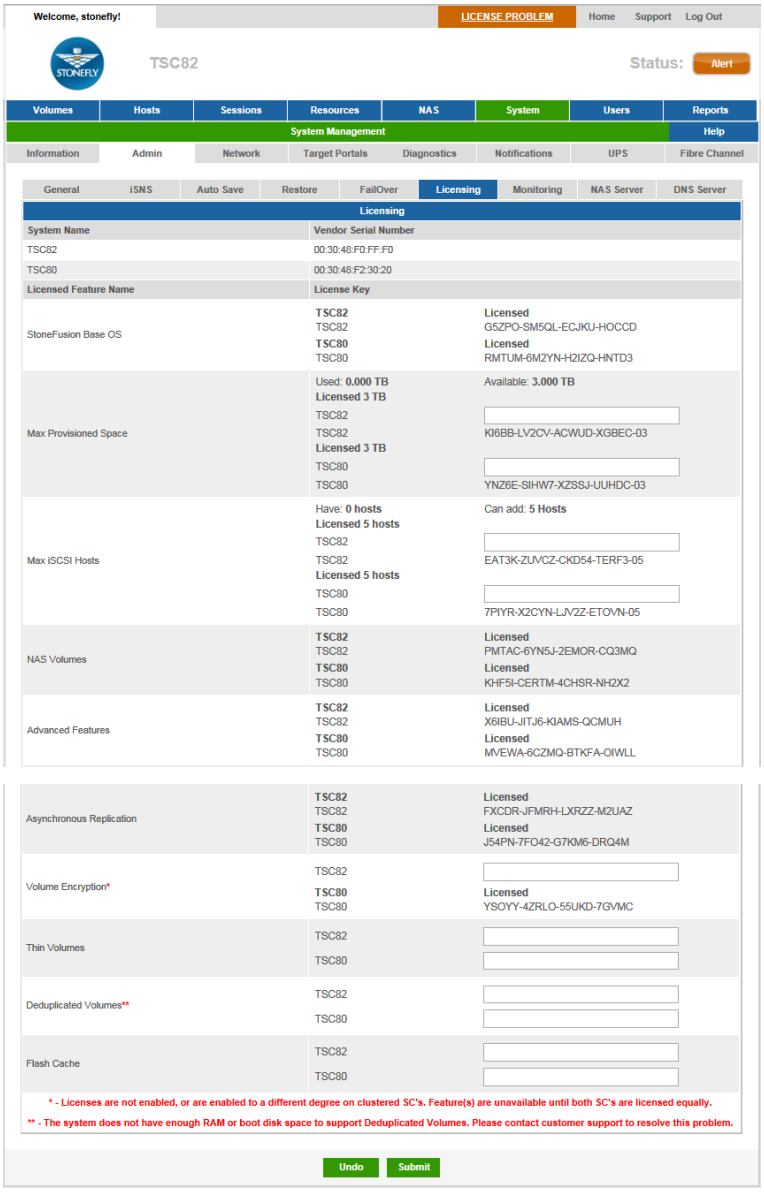


Figure 2-57 System Management, Licensing Screen-Problem

The GUI title bar indicates "**LICENSE PROBLEM**" when system does not have any more licensed space for new volume allocation but free resource space is still available.

The GUI title bar indicates "**LICENSE PROBLEM**" or "**NOT LICENSED**" when "Max Provisioned Space" or "Max Hosts" evaluation license expires.

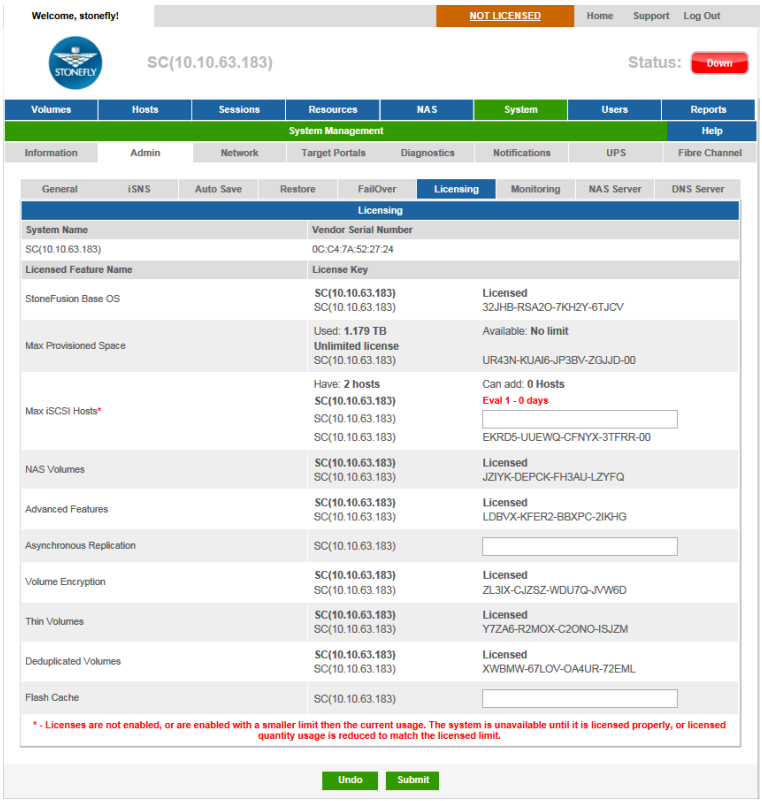


Figure 2-58 System Management, Licensing Screen-Expired

Any operations that let Storage Concentrator handle iSCSI sessions with hosts are disabled on system with expired "Max Provisioned Space" or "Max Hosts" license.

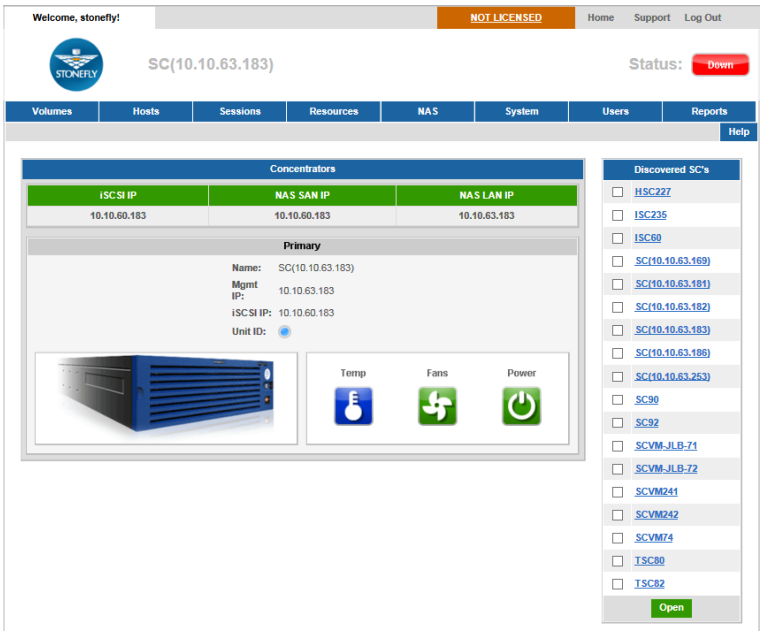


Figure 2-59 License Problem indicated in the Main System Configuration GUI

The appropriate GUI log message is also generated in this case.

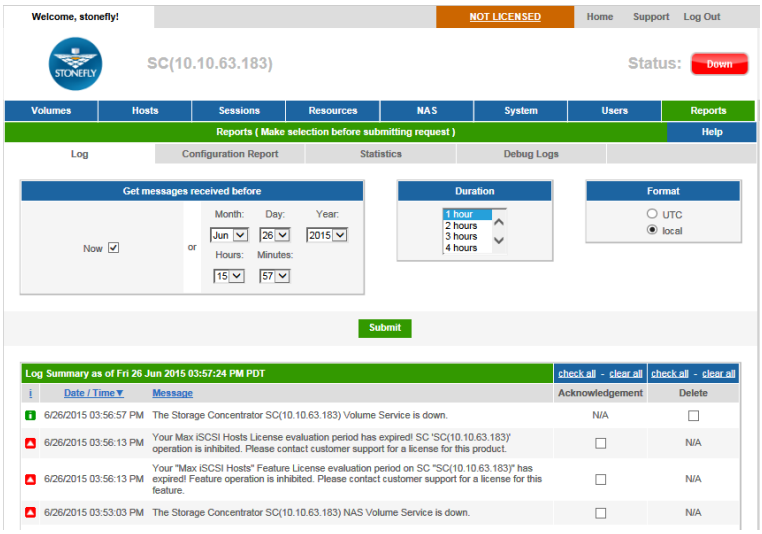


Figure 2-60 License Problem indicated in the System Log GUI

Additional information on Feature Licenses can be found in the online help on the licensing page.

2.7.19 Target Portals

Target Portals are an iSCSI term for a communications portal between two iSCSI devices. The StoneFusion operating system features an iSCSI initiator that enables the *Storage Concentrator* to communicate with other iSCSI devices. Currently, the only iSCSI device supported using a target portal is another *Storage Concentrator*.

The primary use for target portals is to set up a link between a primary *Storage Concentrator* and a secondary *Storage Concentrator* at a campus mirror location (see the chapter "StoneFly Mirroring Synchronous Mirroring") or a link between a local *Storage Concentrator* and a remote *Storage Concentrator* at an asynchronous mirror location (see the chapter "Asynchronous Mirrors").

To set up a target portal, use the steps that follow:

1 Navigate to System>Target Portals>. The Target Portal screen appears.

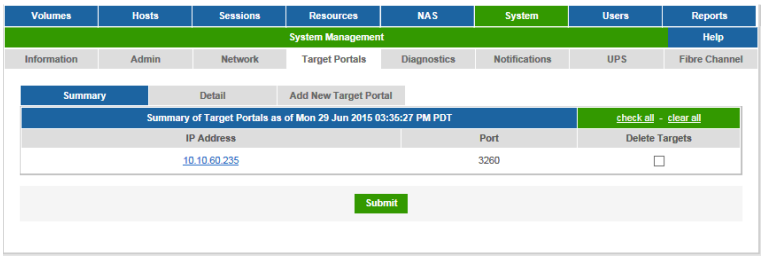


Figure 2-61 The Target Portal screen

2 Select **Add New Target Portal.** The Add New Target Portal screen appears.

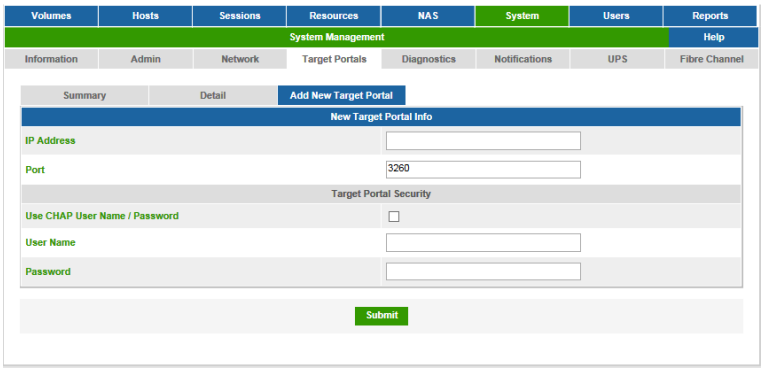


Figure 2-62 Add New Target Portal screen

3 Enter the IP Address of the Storage Concentrator to be used as a Secondary Storage Concentrator

4 The default setting for Port is 3260, the iSCSI listening port.

5 Enter the security settings for the Target Portal, including specifying a CHAP secret.

6 Click on **Submit**

To Check the Status of a Target Portal, click on the Detail button. The Target Portal Detail screen appears.

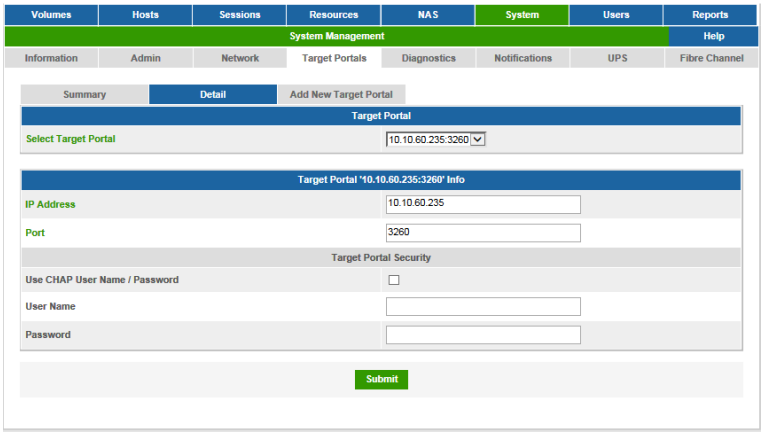


Figure 2-63 Target Portal Detail Screen

To delete a target portal:

1 Click **System**>**Target Portal**
The Target Portal summary screen appears:

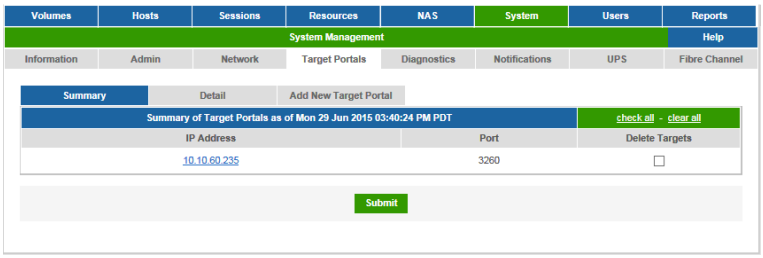


Figure 2-64 Target Portal Summary Screen

2 Select the Target Portal you wish to delete under Delete Targets.

3 Click **Submit**.

2.7.20 Diagnostics

The *Storage Concentrator* provides system information that may be useful when troubleshooting hardware problems.

To view diagnostic information, use the steps that follow:

1 Click System.

2 Click Diagnostics.

The diagnostic report includes the following information:

Monitors: in a FailOver cluster the following status information is provided for the cluster devices listed below. Possible statuses include: Healthy, Failed, Disabled, And Critical.

Cluster Monitors:

iSCSI Host IP Address: IP address resources

Monitor for *Storage Concentrator* name (IP address):

Web Service: https service resource

Management Port Link: 10/100/1000 link status resource

Data Port Link: The Local iSCSI Data Port link status resource

Volume Services: virtual device resource

Monitor for *Storage Concentrator* name (IP address):

Web Service: https service resource

MS Port Link: Management Port link status resource

Data Port Link: The Local iSCSI Data Port link status resource

Volume Services: virtual device resource

Figure 2-65 The Storage Concentrator Diagnostics Screen

Temperatures

CPU or Processor: The status and temperature of the main processor and motherboard

Ambient: The status and temperature of the environment

BP Bottom Temp: The status and temperature of the air in the chassis and backplane

Fans

The status and speed of the fan in the power supply bay, and in the main chassis of the unit

Power Supplies

The status of the power supplies, and power related sensors.

When UPS Management is enabled, the overall UPS status is shown here, with detailed status on the UPS Management page.

If you suspect a hardware problem, contact StoneFly Technical Support at 510-265-1616, 24 hours a day, 7 days a week.

2.7.21 Notifications

The notifications feature provides the system manager immediate notification of critical Alerts and warnings via e-mail.

To configure notifications, use the steps that follow:

1 Click System.

2 Click Notifications. The System Management Notifications screen appears.

3 Enter the **SMTP Server IP Address** (or Domain Name if using DNS).

The SMTP Server must be able to accept requests from the Storage Concentrator.

4 Enter the **SMTP Server Port**.

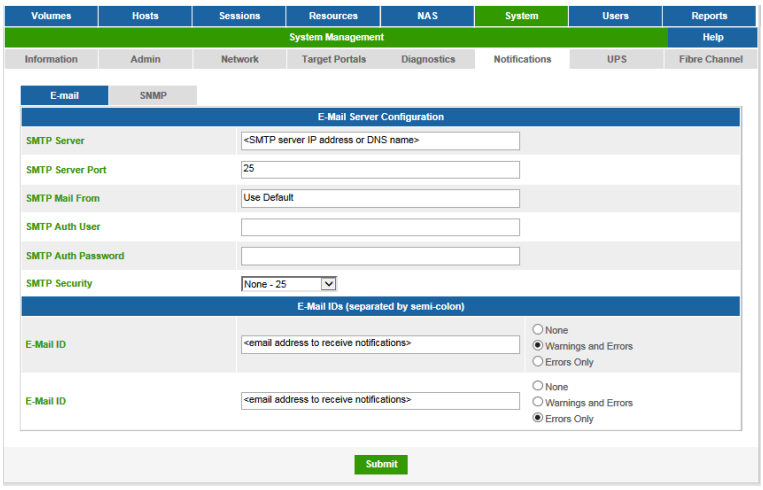


Figure 2-66 System Management, Notifications screen

5 Enter the email address for the **SMTP Mail From** portion of the email notifications. The *Storage Concentrator* supplies a default based on its own IP Address. Some email systems are more restrictive with the "MAIL FROM" field for SMTP email to be accepted. You may enter a valid email address value that will be used instead of the default value if desired.

6 **SMTP Auth User** — Many SMTP email servers require user name and password authentication to accept messages for transmission. If your server requires authentication, enter the user login name here. Note that if authentication is not required, both the SMTP Auth User and SMTP Auth Password fields should be left blank.

7 **SMTP Auth Password** — Many SMTP email servers require user name and password authentication to accept messages for transmission. If your server requires authentication, enter the password here. Note that if authentication is not required, both the SMTP Auth User and SMTP Auth Password fields should be blank.

8 Enter the e-mail address that will be sent notifications. Enter multiple addresses by using a semicolon between email addresses.

9 Select the notification type:
None: Select to temporarily turn off notifications
Warnings and Alerts: Select to receive both warning level Alerts and critical level Alerts
Alerts Only: Select to receive only critical level errors

10 Click **Submit**.

A list of the warning and critical level notification messages can be found in the "*StoneFly Storage Concentrator System Event Messages*" document.

2.7.22 UPS Management

The UPS Management page is where the use of a connected UPS system can be configured, monitored, and controlled.

A UPS system can provide backup power for a Storage Concentrator (SC) in the event of a power outage. This allows for the SAN to survive a short power outage without the equipment needing to be restarted, and can provide for a clean shutdown should the power outage be prolonged. When AC power is restored and the UPS batteries are recharged to a sufficient level, the SC would be powered back up automatically.

For a proper controlled shutdown in the case of a prolonged outage, the SC must be configured to monitor and manage the UPS system.

Note: Only one system can control or manage the UPS system. Because there is a dialog between the UPS and the server determining when a shutdown is required, a single system must be delegated to be in charge of the UPS that is powering it. Although it might seem possible to connect the UPS control/management interface to multiple servers, this should not be done.

UPS Management is per SC. When there are two clustered SC's, each must be configured for the management of their UPS when there are two UPS's. When there is a single UPS that is shared, the directly connect SC should be configured as the **Master**, and the other SC should be a **Slave** to the first SC.

When an external RAID is also powered by a UPS being managed by the SC, the ability to shut-down the RAID at the same time that the SC is shutdown should be enabled. This setting is found on the **Resource Detail** page for the RAID's resources.

An SC should only be cabled for communications to a single UPS. If connected to multiple UPS systems, only one UPS would be used, and it would not be predictable which one that would be.

Note: All equipment necessary for the SC to communicate with the UPS, and other SAN systems such as clustered SC's and external RAID's should be also powered by the UPS. For example, should a powered USB hub be used for the UPS connection, it should be powered by the UPS as well. The same applies for any network switches that connect clustered SC's.

To configure UPS Management, use the steps that follow:

1 Click System.

2 Click UPS. The UPS Management screen appears.

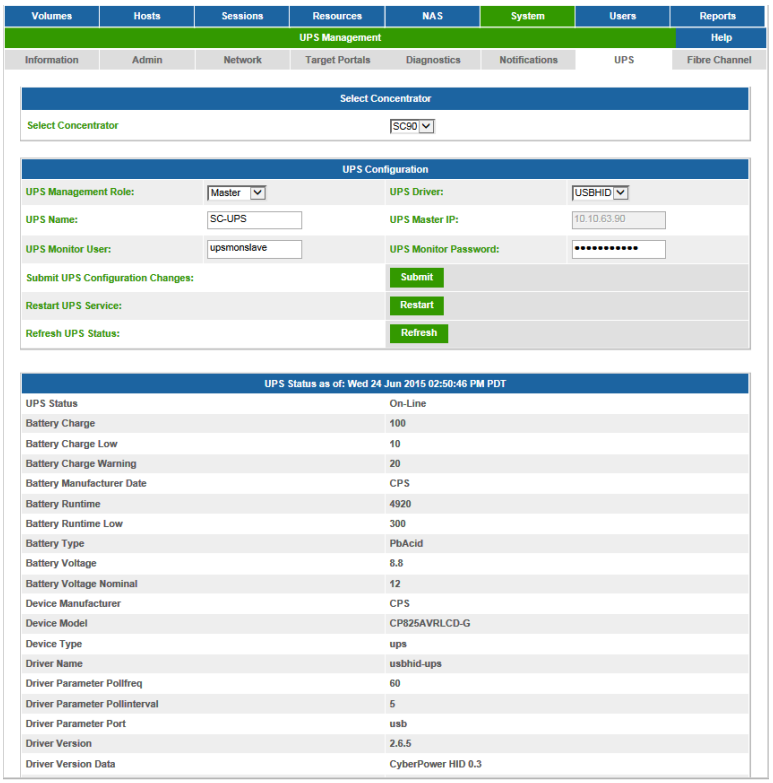


Figure 2-67 System Management, UPS Management Configuration screen

4 Set the **UPS Management Role**.
This controls whether there is a UPS system connected and being managed, and what management role is being used.
When set to **Disabled**, no UPS Management is performed. This would be the case for no UPS system present. This is the default setting.
When set to **Master**, the selected SC is directly connected to communicate with its UPS, and continuously polls it for status, responds to changes in status, and submits commands to the UPS on user request. The SC is assumed to be powered by the same UPS.
When set to **Slave**, the selected SC is **not** connected to communicate with the UPS, but

it is still powered by it. Instead of communicating with the UPS, the SC communicates with the **UPS Master**, and continuously polls it for status, and responds to changes in status, and commands from it. The SC is assumed to be powered by the UPS controlled by the **UPS Master**. A common configuration would be for a single UPS to be shared by both SC's in a cluster, and be controlled by the SC acting as the **Master**, with the other SC being a **Slave** to the first.
Note: The role of UPS **Master** or **Slave** is completely independent from the cluster SC status of Primary or Secondary. For example, half of the time, the cluster secondary may be the UPS master.
A configuration change here does not take effect until the **Submit** button is pushed.
Only administrative level users are allowed to make configuration changes.

5 Set the **UPS Driver**.
This sets the driver that the SC uses to interact with the UPS system. This controls both the interface type, and protocol that is used over that interface.
When set to **USBHID**, the USB interface is used to connect to the UPS, with the HID (Human Interface Device) Power Device class protocol.
The **USBHID** driver supports at least the following UPS systems: all MGE UPS SYSTEMS USB models, some APC models, some Belkin models, some Cyber Power Systems models, some TrippLite models, and some Liebert models.
The UPS Driver setting has no significance when the **UPS Management Role** is not set to **Master**.
A configuration change here does not take effect until the **Submit** button is pushed.
Only administrative level users are allowed to make configuration changes.

6 Set the **UPS Name**. This is the name that is used to refer to the UPS. For a simple single **UPS Master** configuration, the default name is sufficient.
However, when there are both **UPS Master** and **Slave** SC's (or other systems) configured, the ability to use a different **UPS Name** may be required.
A configuration change here does not take effect until the **Submit** button is pushed.
Only administrative level users are allowed to make configuration changes.

7 Set the **UPS Master IP**, which is the IP address of the UPS **Master** system.
When the **UPS Management Role** is **Master**, this is a display field only, and shows the SC's management IP address.
When the **UPS Management Role** is **Slave**, the IP address of the UPS **Master** SC (or other system) should be entered.
A configuration change here does not take effect until the Submit button is pushed.
Only administrative level users are allowed to make configuration changes.

8 Set the **UPS Monitor User,** which is the user name that is accepted by the UPS **Master** from the **Slave** during UPS monitoring communications.
Generally, for a single UPS shared between two SC's in a cluster, the default **UPS Monitor User** name is sufficient. When there are other systems involved though, the ability to configure the **UPS Monitor User** may be required.
A configuration change here does not take effect until the **Submit** button is pushed.
Only administrative level users are allowed to make configuration changes.

9 Set the **UPS Monitor Password**, which is the password that is accepted by the UPS **Master** from the **Slave** during UPS monitoring communications.
Generally, for a single UPS shared between two SC's in a cluster, the default **UPS Monitor Password** is sufficient. When there are other systems involved though, the ability to configure the **UPS Monitor Password** may be required.
A configuration change here does not take effect until the Submit button is pushed.

10 Only administrative level users are allowed to make configuration changes.

11 Click **Submit**. If the UPS is connected, supported, and working properly, **UPS Status** fields should be displayed. The UPS Status section displays status values obtained from the UPS system.
The specific values shown here depend upon the make and model of the UPS system. The meaning of the values is usually self-explanatory. More information should be available from the UPS vendor's documentation.
The **UPS Status** value is always present, and is listed first. If this value indicates an **Error**, it means that there was a problem interacting with the UPS system. There may be other values giving more detail on the error. Should an Error be indicated, verify that the UPS is properly connected to the SC, and perform a **Restart UPS Service** to attempt to restore access to it.
The **UPS Status** error status of **Data stale** indicates that the UPS data was available at one point, but currently data updates are not functioning. One example where this would occur is on the SC that is a UPS **Slave** to the other UPS **Master** SC in the cluster when that SC is not running. A Data stale condition can sometimes resolve itself over time.
The **UPS Status** error status of **Driver not connected** or **Connection failure** are indications of a more persistent problem, such as the UPS not being connected to the UPS **Master** SC, or there being no network path from the **Slave** to the **Master**. For these cases, normal **UPS Status** would not resume until the condition is cleared and a successful **Restart UPS Service** is performed.
Note: Connecting an unsupported UPS system would also result in an **Error** status.

12 At any point, you may click the **Refresh** button to update the **UPS Status** values.

13 At any point, you may click the **Restart** button to restart the UPS service that monitors the UPS.
The Restart button can be used to restart the SC UPS management service when the UPS Status or SC event log messages indicate problems communicating with the UPS system.
For example, if the UPS was disconnected when the SC was rebooted, it would not be able to communicate with the UPS. Clicking **Restart** would restart the service, and allow the UPS to then be discovered.
Only administrative level users are allowed to use the **Restart** button.

14 If the SC system(s) are connected to an external RAID, and that RAID is powered by a UPS managed by the SC's, the SC needs to be configured to shut-down the RAID. Go to the **Resource Management Detail** page for a managed resource provided by the RAID.

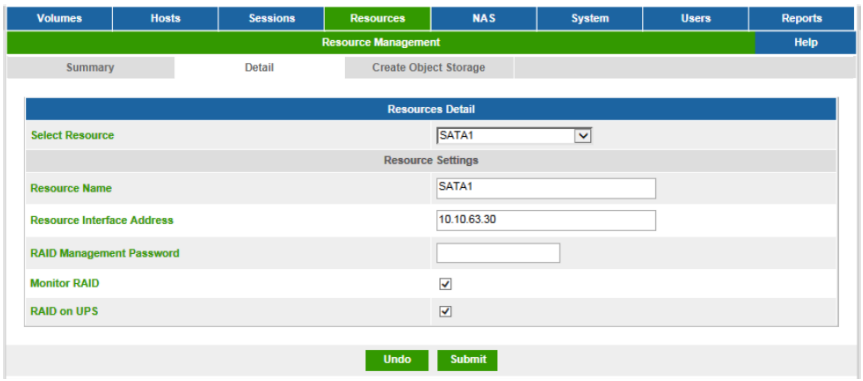


Figure 2-68 Resource Management, Details Screen

15 The **RAID on UPS** check-box indicates whether or not the RAID is connected to the UPS that is managed by the Storage Concentrator.
When the SC is managing a UPS, and this check-box is checked, it is assumed that the RAID is also powered by the same UPS.
Should a UPS low battery shut-down occur, the RAID will be shut-down at the same time that the SC is. This causes any I/Os cached in the RAID to be flushed to disk before power is lost. It also prevents any new I/O's until the RAID is power cycled by the UPS.
Only RAIDs that are being monitored as indicated by the **Monitor RAID** check-box are able to be shut-down. As not all RAID vendors and models are supported for monitoring, such unsupported RAIDs would have their RAID on UPS check-box hidden.
If the SC is not configured to manage its UPS, or if the RAID is not being monitored, the **RAID on UPS** has no effect.
Internal RAIDs are always shutdown with the SC when a UPS low battery shut-down occurs. There is no configuration needed for internal RAIDs.
When the RAID is not powered by any UPS, or by a different UPS than the one that the SC is managing, the **RAID on UPS** check-box should **not** be checked. Otherwise, the RAID would be shut down when it might still be in use, and it would require manual intervention for it to be restarted. This is especially important when the RAID is being shared with other systems.

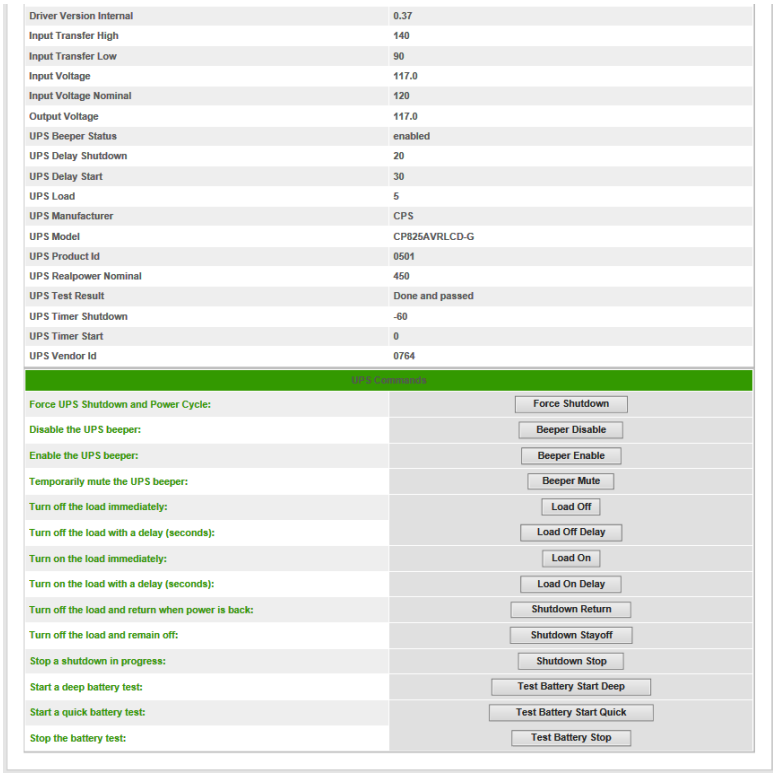


Figure 2-69 System Management, UPS Management Commands screen

2.7.22.1 UPS Commands

There are a number of test and control functions available in the **UPS Commands** section of

the page.

The specific tests and commands available here depend upon the make and model of the UPS system.

Warning: Many UPS commands result in a power down, or a power cycle of all equipment powered by the UPS. In some cases, power may not be restored without manual intervention at the UPS.

The commands indicate success when they were submitted to the UPS. The actual success of the test, or action commanded is indicated by changes in the **UPS Status** variables, and the UPS system state. Note that there may be a short delay before the UPS takes the action.

The **Force UPS Shutdown and Power Cycle** command will usually be present. This function is useful in that it can verify that everything will work properly in the event of a long AC power failure. An alternative test is to disconnect the UPS from the AC power, and then wait for the batteries to drain fully, which may take hours.

The **Force UPS Shutdown and Power Cycle** command simulates in the UPS a power failure followed by a Low Battery condition. This triggers the SC to shut down, and the UPS to then remove power from it and all other UPS powered equipment. After a short delay, the UPS should power the SC back up, the SC then reboot, and everything goes back to normal.

Only administrative level users are allowed to use **UPS Commands**.

2.7.23 System Monitoring

The Storage Concentrator can be configured for centralized monitoring by the Nagios Monitoring System (http://www.nagios.org/). The System Management – System Monitoring page is where this is done.

StoneFly can provide a service to remotely monitor Storage Concentrator (SC) systems to assist in managing them, and proactively respond to any issues that may arise.

Contact StoneFly for details.

Alternatively, the customer may setup and operate their own Nagios Monitoring System internally.

For System Monitoring to work, the Nagios NRPE service running on the Storage Concentrator must be configured, and enabled -- it is disabled by default.

The only information offered to the monitoring system is system status and version information such as disk and memory utilization, uptime, SC software version, etc. No customer sensitive information is exchanged with the monitoring system. Also, there is no ability to remotely control or reconfigure the SC system -- only monitoring functions are provided.

In order for StoneFly to be able to remotely monitor a Storage Concentrator, Internet access to the Nagios NRPE service must be provided by adding firewall access and forwarding rules to allow it.

Note: The Storage Concentrator must never be placed directly on the Internet -- a network firewall must always be present. The firewall rules should only allow access to the Nagios NRPE service listening port.

When configuring the Nagios NRPE network service on the Storage Concentrator, StoneFly customer support must be contacted so they can add the necessary configuration settings to their Nagios Monitoring Systems so that the monitoring can occur.

For each SC being monitored, this will include the firewall's external Internet facing IP address for the SC, and the TCP port that the Nagios NRPE service is listening on. Note that the external IP address provided to StoneFly may not be the SC Management IP address when the firewall is using Network Address Translation (NAT). Likewise, the TCP port provided may not be the Nagios NRPE listening port configured on the SC when the firewall or other network systems are performing port forwarding.

When StoneFly centralized monitoring is arranged, the Storage Concentrator E-Mail Notifications should also be configured so that these will also be received by StoneFly. StoneFly customer support will provide the correct email address to use in the notifications.

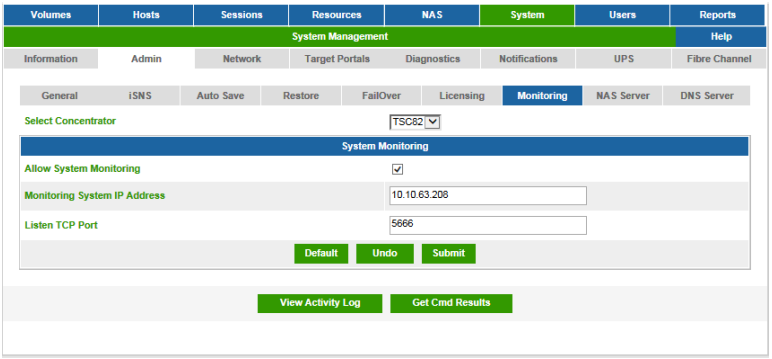


Figure 2-70 System Management, System Monitoring screen

2.7.23.1 System Monitoring Fields

Select Concentrator — When in a Storage Concentrator cluster, this control selects which SC the System Monitoring settings are being viewed or changed on.

With the exception of the 'Listen TCP Port' field, the settings are common to both concentrators in the cluster.

Allow System Monitoring — Select this check-box to enable the Nagios NRPE network service.

To change this setting, click the check-box and click submit.

Monitoring System IP Address — This is the IP Address that the Nagios Monitoring System accesses the Storage Concentrator from.

For security, only this IP Address is allowed by the SC Nagios NRPE network service. All other source addresses are rejected.

The default IP Address is the StoneFly central monitoring system is 204.128.202.129 (monitor.dnfcorp.com).

Note that in some cases, when port forwarding is used, the IP Address would be that of the network device performing the forwarding and is not the external Internet IP address.

Listen TCP Port — This is the TCP Port that the Nagios monitoring network NRPE service on the SC listens on.

The default TCP Port is the Nagios NRPE default 5666.

Note that in some cases, based upon the firewall and port forwarding systems being used, the TCP Port would need to be changed.

In some cases, each SC would need to have a unique TCP Port number. This is the case when the firewall is configured such that there is a single external IP address for all of the SC's. For this reason, when there are two SC's operating in a cluster, each can be configured with a different TCP Port.

The other settings on this page are common for both SC's in the cluster, but the TCP Port is not.

When changing the TCP Port setting in an SC cluster, make sure to make the necessary changes for both SC's.

Note that when port forwarding is used, the TCP Port number used here is not the same port number that would be provided to StoneFly for external access.

2.7.23.2 System Monitoring Buttons

View Activity Log — The **'View Activity Log'** button will report recent access attempts and failures by the Nagios Monitoring System to communicate with the NRPE service on the SC.

Attempts are indicated by 'START' lines, and failed attempts are indicated by 'FAIL' lines.

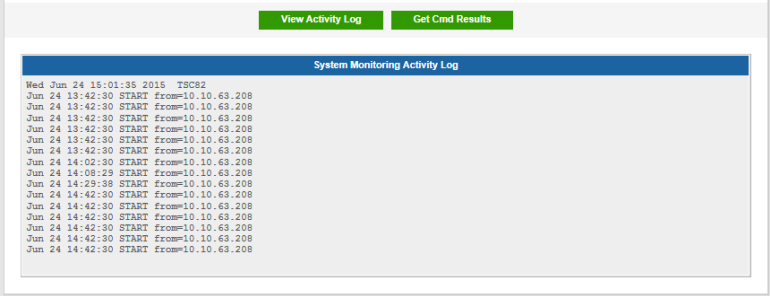


Figure 2-71 System Monitoring Activity Log

Get Cmd Results — The **'Get Cmd Results'** button will list all of the Nagios NRPE commands supported by the Storage Concentrator, the command options used, and the output that is returned when the command is run.

This can be used to show the information that the Nagios Monitoring System will receive. The command names would be needed should an internal Nagios system be implemented.

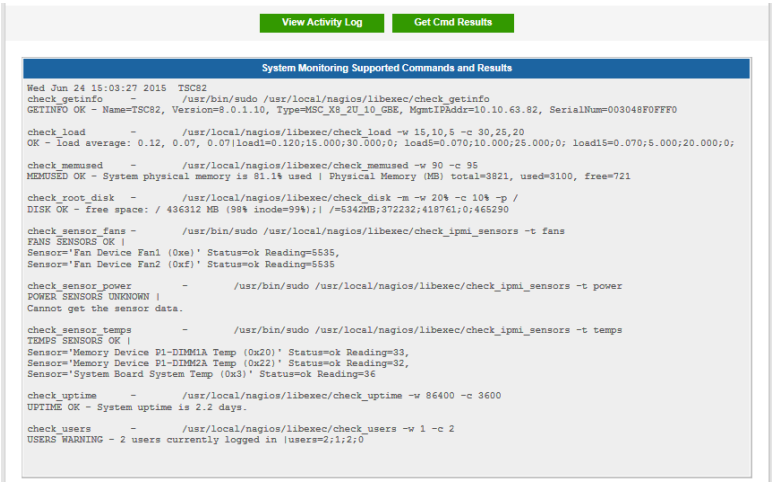


Figure 2-72 System Monitoring Supported Commands and Results

Default — The '**Default**' button resets the configuration to the default settings.

The default changes are submitted and take immediate effect.

Undo — Click to revert to the last saved settings.

Submit — Click to commit changes make to the input fields.

Note that changes here made do not require an SC reboot to take effect -- they are in effect immediately.

2.7.24 NAS Server

NAS Volume CIFS user name and password authentication can operate in two modes.
The default is the "Workgroup" mode where the SC system(s) manage the CIFS user accounts and passwords.
The other CIFS authentication mode is the use of a centralized Windows Active Directory Server (ADS) (also known as a Windows Domain Controller). Note that Workgroup users may continue to be used even when the ADS authentication mode is configured.

In an SC cluster, the NAS Server CIFS User Authentication configuration applies to both cluster nodes. The changes are made to both cluster nodes, and no reboot or fail-over is required to affect the changes on the secondary.

In an SC Scale Out (SO) configuration, the NAS Server CIFS User Authentication configuration applies to all of the SO nodes; they all must have the same configuration. Changes made are applied to all SO nodes. Status shown is from all SC SO nodes.

The management of the SC NAS Server CIFS User Authentication GUI page has the following sections:

1. NAS Server User Authentication Configuration

This section configures the CIFS user authentication mode and settings. This section is where the configuration settings are initially made or changed, shows the current settings. The configuration generally only needs to be set up once.
The fields displayed in this section vary based on the **CIFS Authentication Mode** in use, either **Workgroup** or **Active Directory**.

2. Active Directory Status
   This section shows the current status accessing the Windows ADS server. All clustered and SO SC nodes are shown. The status is updated each time the GUI page is refreshed. This section is only shown when the **CIFS Authentication Mode** is set to **Active Directory**.
3. NAS Server Active Directory Command
   This section allows Active Directory management commands to be submitted to the Windows ADS Server. These commands generally only need be performed once, during initial setup. The Active Directory command is performed on all SC nodes in an SC cluster. The command is performed on all SC nodes in a SO configuration. This section is only shown when the **CIFS Authentication Mode** is set to **Active Directory**.
4. Active Directory Domain Command Results
   This section shows the results of the submitted Active Directory command. This section is only shown when the **CIFS Authentication Mode** is set to **Active Directory**, and only immediately after a command was submitted.

The requested action requires that the CIFS services be restarted. This will cause all client sessions to be disconnected, and need to reconnect. If user authentication settings were changed, manual intervention may be needed for the clients to reconnect.
Note that each section with submission buttons, etc. operates independently. Work in one section at a time, otherwise changes made in the other sections may be lost and have to be re-entered.

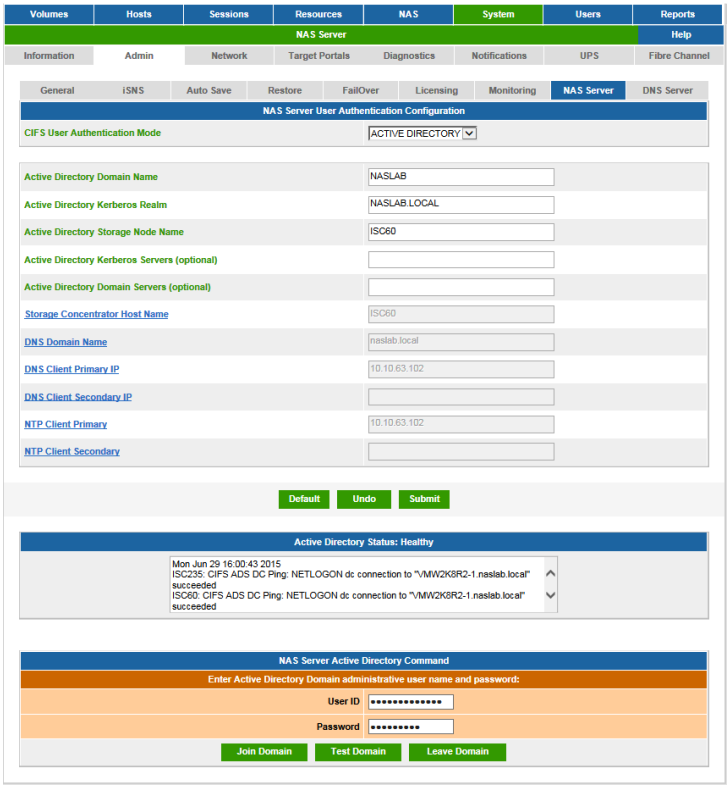


Figure 2-73 NAS Server screen

2.7.24.1 NAS Server Fields

Active Directory Domain Name— The **Active Directory Domain Name** field is the name of the Windows Active Directory domain that the SC is to be a member of.
This should not be a fully qualified domain name, but only the short name; the DNS domain that the SC belongs to is appended to this value.
This field is only shown when the **CIFS User Authentication Mode** is set to **ACTIVE DIRECTORY**.

Active Directory Domain Servers— The **Active Directory Domain Servers** is an optional list of one or more space separated IP addresses or DNS names for the Windows Active Directory servers.
This field is optional, and is seldom needed.
This field is only shown when the **CIFS User Authentication Mode** is set to **ACTIVE DIRECTORY**.

Active Directory Kerberos Realm— The **Active Directory Kerberos Realm** is the fully qualified Kerberos Realm name. This value is similar, but not necessarily identical to the Active Directory domain controller fully qualified DNS name. It should always be entirely upper case. For example: MYDC.MYDOMAIN.COM.
This field is only shown when the **CIFS User Authentication Mode** is set to **ACTIVE DIRECTORY.**

Active Directory Kerberos Servers— The **Active Directory Kerberos Servers** is an optional list of one or more space separated IP addresses or DNS names for the Windows Active Directory Kerberos servers.
This field is optional, and is seldom needed.
This field is only shown when the **CIFS User Authentication Mode** is set to **ACTIVE DIRECTORY**.

Active Directory Storage Node Name— The **Active Directory Storage Node Name** is the machine name that the SC, when clustered, both SC's will be known as in the Windows Active Directory domain.
This field is only shown when the **CIFS User Authentication Mode** is set to **ACTIVE DIRECTORY**.

CIFS User Authentication Mode— The **CIFS User Authentication Mode** can be set to either **WORKGROUP** or **ACTIVE DIRECTORY** mode.
In the **WORKGROUP** mode, the SC(s) perform CIFS user authentication themselves, using the CIFS Users and passwords configured.
In **ACTIVE DIRECTORY** mode, the SC(s) use the configured Windows Active Directory server to authenticate CIFS users.
CIFS users will only successfully authenticate after the SC(s) have *joined* the domain, and when the communication with the Active Directory server is successful.
The setting of this field governs the visibility of many other fields on this GUI page.
While in **ACTIVE DIRECTORY** mode, **WORKGROUP** CIFS users defined in the SC(s) may also be used.

DNS Client Primary IP— The **DNS Client Primary IP** field shows the current setting of the primary IP address DNS server that the SC(s) are using to resolve DNS names into IP addresses.
When in **ACTIVE DIRECTORY** mode, the Windows Active Directory domain controller should generally be used as the DNS server.

Either one or both of primary and secondary values must be configured, and functioning.
Changes to this value are made on another SC GUI page, the field label is a link to go to that page.
This field is only shown when the **CIFS User Authentication Mode** is set to **ACTIVE DIRECTORY**.

DNS Client Secondary IP— The **DNS Client Secondary IP** field shows the current setting of the secondary IP address DNS server that the SC(s) are using to resolve DNS names into IP addresses.
When in **ACTIVE DIRECTORY** mode, the Windows Active Directory domain controller should generally be used as the DNS server.
Either one or both of primary and secondary values must be configured, and functioning.
Changes to this value are made on another SC GUI page, the field label is a link to go to that page.
This field is only shown when the **CIFS User Authentication Mode** is set to **ACTIVE DIRECTORY**.

DNS Domain Name— The **DNS Domain Name** field shows the current setting of the SC's DNS domain name, for example *stonefly.com*. This value is appended to the **Active Directory Domain Name** to form a fully qualified DNS name for the Active Directory server.
When in **ACTIVE DIRECTORY** mode, the Windows Active Directory domain should generally be used as the **DNS Domain Name**.
Changes to this value are made on another SC GUI page, the field label is a link to go to that page.
This field is only shown when the **CIFS User Authentication Mode** is set to **ACTIVE DIRECTORY**.

NTP Client Primary— The **NTP Client Primary** field shows the current setting of the primary DNS name or IP address for the NTP server that the SC(s) use to maintain time of day clock synchronization using the Network Time Protocol.
When in **ACTIVE DIRECTORY** mode, the Windows Active Directory domain controller should generally be used as the NTP server.
Either one or both of primary and secondary values must be configured, and functioning.
Changes to this value are made on another SC GUI page, the field label is a link to go to that page.
This field is only shown when the **CIFS User Authentication Mode** is set to **ACTIVE DIRECTORY**.

NTP Client Secondary— The **NTP Client Secondary** field shows the current setting of the secondary DNS name or IP address for the NTP server that the SC(s) use to maintain time of day clock synchronization using the Network Time Protocol.
When in **ACTIVE DIRECTORY** mode, the Windows Active Directory domain controller should generally be used as the NTP server.
Either one or both of primary and secondary values must be configured, and functioning.
Changes to this value are made on another SC GUI page, the field label is a link to go to that page.
This field is only shown when the **CIFS User Authentication Mode** is set to **ACTIVE DIRECTORY**.

Password— The **Password** field is for the Windows Active Directory administrative user account used when one of the Active Directory Commands are used.
This account is not used during CIFS user authentication, only when the SC(s) machine account is being configured in the domain.

By default, the Active Directory administrative user account is stored in the SC database, which is indicated by the field being pre-populated. This information can be removed from the SC if desired by deleting the values from the field and then using one of the Active Directory Commands.
This field is only shown when the **CIFS User Authentication Mode** is set to **ACTIVE DIRECTORY**.

Storage Concentrator Host Name— The **Storage Concentrator Host Name** is the host name of the SC. Two SC's in a cluster should have different host names.
Changes to this value are made on another SC GUI page, the field label is a link to go to that page.
This field is only shown when the **CIFS User Authentication Mode** is set to **ACTIVE DIRECTORY**.

User ID— The **User ID** field is for the Windows Active Directory administrative user account used when one of the Active Directory Commands are used.
This account is not used during CIFS user authentication, only when the SC(s) machine account is being configured in the domain.
By default, the Active Directory administrative user account is stored in the SC database, which is indicated by the field being pre-populated. This information can be removed from the SC if desired by deleting the values from the field and then using one of the Active Directory Commands.
This field is only shown when the **CIFS User Authentication Mode** is set to **ACTIVE DIRECTORY**.

2.7.24.2 NAS Server Buttons

Default— The **Default** button is used to reset all of the *NAS Server User Authentication Configuration* fields to their default values, and submits them.
The changes will apply to both SC's in a cluster configuration and to all Scale Out nodes in a SO configuration.

Join Domain— The **Join Domain** button is used to (re-)register the SC(s) with the Windows Active Directory domain controller.
The correct **User ID** and **Password** for an administrative account for the ADS server must be provided, or have been provided previously.
The command will be performed on both SC's in a cluster configuration, and to all Scale Out nodes in a SO configuration.
The results of the command are shown in the *Active Directory Domain Command Results* section of the page.
This field is only shown when the **CIFS User Authentication Mode** is set to **ACTIVE DIRECTORY**.

Leave Domain— The **Leave Domain** button is used to deregister the SC(s) from the Windows Active Directory domain controller.
The correct **User ID** and **Password** for an administrative account for the ADS server must be provided, or have been provided previously.
The command will be performed on both SC's in a cluster configuration, and to all Scale Out nodes in a SO configuration.
The results of the command are shown in the *Active Directory Domain Command Results* section of the page.
Once the SC(s) leave the domain, CIFS User authentications through the Active Directory

server will fail.
This field is only shown when the **CIFS User Authentication Mode** is set to **ACTIVE DIRECTORY**.

Submit— The **Submit** button submits all of the settings in the *NAS Server User Authentication Configuration* fields to their configured values. If the values are accepted, they are committed to the SC DB.
The changes will apply to both SC's in a cluster configuration and to all Scale Out nodes in a SO configuration.

Test Domain— The **Test Domain** button is used to test the ability of the SC(s) to command the Windows Active Directory domain controller.
The correct **User ID** and **Password** for an administrative account for the ADS server must be provided, or have been provided previously.
The command will be performed on both SC's in a cluster configuration, and to all Scale Out nodes in a SO configuration.
The results of the command are shown in the *Active Directory Domain Command Results* section of the page.
This field is only shown when the **CIFS User Authentication Mode** is set to **ACTIVE DIRECTORY**.

Undo— The **Undo** button is used to undo any changes to the *NAS Server User Authentication Configuration* fields, restoring them to their previously committed values.

2.7.25 DNS Server

The Storage Concentrator (SC) provides a Domain Name Service (DNS) Server that can be used by clients to access the SC by a DNS name instead of directly by an IP address.

For iSCSI clients, using DNS names instead of IP addresses can be convenient. For DNS clients, DNS name use is more important, as they tend to be greater in number, and are often general purpose systems.

For Scale Out NAS, the client use of DNS instead of direct IP address is recommended as it provides a way for the NAS client load to be distributed amongst the many Scale Out SC's.

In both clustered SC's, and Scale Out SC configurations, the DNS Server name to address records are synchronized so that all of the SC's return the same name look-up responses.

There are two ways that the SC DNS Server can be used.

It can be configured to be a sub-domain of an existing corporate DNS server. The corporate DNS would then refer or forward requests to the SC DNS Server, serving all other requests itself. The advantages here are that the existing clients to not need to update their DNS server IP addresses, and the DNS look-up load is kept off the SC systems.

To configure the use of the SC DNS Server to be downstream from the corporate DNS server, the corporate DNS server should be configured to forward to an SC an SC specific DNS sub-domain, or a defined set of names that are also known to the SC DNS Server. In this case, the SC DNS Server configuration settings Primary and Secondary Upstream DNS Server IP settings should be left blank. Always configure both a primary and secondary DNS IP

addresses on the corporate DNS server pointing to two different SC's in the deployment, so that if one is down, the other can still provide DNS service.

The second way to use the SC DNS Server is for it to be the main DNS server that clients point to for their DNS services, and the SC DNS Server itself forward non-SC related name look-up requests to a corporate DNS Server. In this case, both the Primary and Secondary Upstream DNS Server IP address settings must be set, and point to the primary and secondary corporate DNS Servers. The SC DNS Server caches results from the upstream server to lessen the load and improve response.

The SC client's use of DNS instead of direct IP addresses is entirely optional, even in Scale Out NAS configurations. A configuration may not have enough clients to bother with DNS, or may choose to statically assign clients to Scale Out SC's instead of relying on their dynamic distribution. Alternatively, an existing corporate DNS could be populated with SC names and IP addresses by its administrator.

Note: Currently, SC SAN MPath network interface addresses are not supported; there is no way to assign unique DNS names for them.

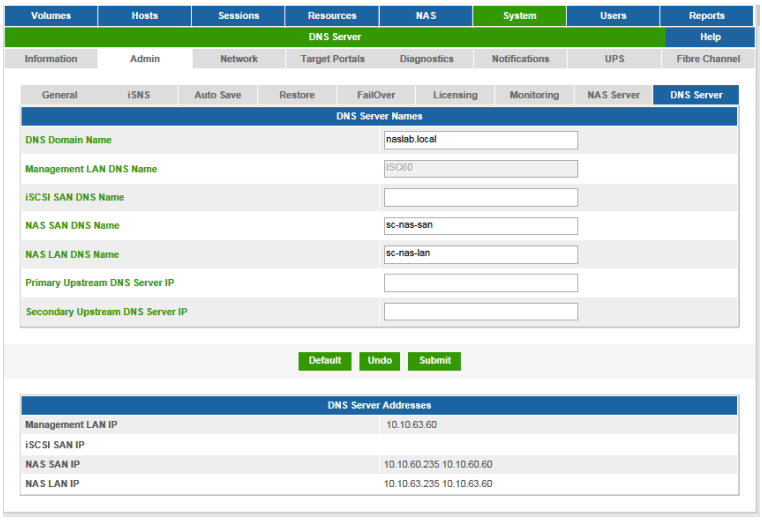


Figure 2-74 DNS Server screen

2.7.25.1 **DNS Server Fields**

DNS Domain Name — The DNS Domain Name is the suffix that will be added to all DNS host short names to form a fully qualified DNS name, without the leading '.', e.g. "stonefly.com".

Management LAN DNS Name — The Management LAN DNS Name is the DNS form of the System Name and maps to the SC Management IP address on the LAN network. This DNS name is used to access the SC GUI.

The Management LAN DNS Name always uniquely addresses a single SC, and will map to a single constant IP address.

The Management LAN DNS Name name cannot be changed on the DNS Server GUI page, but

it can be set on the System -> Admin -> General GUI page.

iSCSI SAN DNS Name — The iSCSI SAN DNS Name is the DNS name that iSCSI clients should use to configure their iSCSI initiator client.

The iSCSI SAN DNS Name uniquely addresses a single SC, or a pair of SC's when they are clustered.

NAS SAN DNS Name — The NAS SAN DNS Name is the DNS name that CIFS and NFS NAS clients should use to access the SC's over the SAN network.

The NAS SAN DNS Name addresses all NAS Scale Out SCs, and would typically resolve to multiple IP addresses. Note that clustered SC's that do not have a Cluster NAS Data IP Alias defined on the System -> Admin -> FailOver -> Setup Cluster GUI page are not addressed by this DNS name as there is no IP address assigned.

The NAS SAN DNS Name only appears when the NAS Volumes feature is licensed.

NAS LAN DNS Name — The NAS LAN DNS Name is the DNS name that CIFS and NFS NAS clients should use to access the SC's over the LAN network.

The NAS SAN DNS Name addresses all NAS Scale Out SCs, and would typically resolve to multiple IP addresses. Note that clustered SC's that do not have a Cluster NAS Management IP Alias defined on the System -> Admin -> FailOver -> Setup Cluster GUI page are not addressed by this DNS name as there is no IP address assigned.

The NAS LAN DNS Name only appears when the NAS Volumes feature is licensed.

Primary Upstream DNS Server IP — The Primary Upstream DNS Server IP address is the first IP address that the SC DNS Server will use to forward a name look-up request that it does not have knowledge of, and that is not already in its cache, to the upstream DNS server.

If the Upstream DNS Server IP addresses are not configured, name look-ups that cannot be resolved will fail.

Secondary Upstream DNS Server IP — The Secondary Upstream DNS Server IP address is the second IP address that the SC DNS Server will use to forward a name look-up request that it does not have knowledge of, and that is not already in its cache, to the upstream DNS server.

If the Upstream DNS Server IP addresses are not configured, name look-ups that cannot be resolved will fail.

Management LAN IP — The Management LAN IP field shows the IP address that the Management LAN DNS Name resolves to.

In SC cluster configurations, the Management LAN IP address should not be configured by clients as their DNS server IP address as one of the SC cluster nodes may be down. Instead, use one of the other IP addresses listed here.

iSCSI SAN IP — The iSCSI SAN IP field shows the IP address that the iSCSI SAN DNS Name resolves to.

NAS SAN IP — The NAS SAN IP field shows the IP address that the NAS SAN DNS Name resolves to. In a NAS Scale Out configuration, there will be multiple IP addresses listed here.

The NAS SAN IP field only appears when the NAS Volumes feature is licensed.

NAS LAN IP — The NAS LAN IP field shows the IP address that the NAS LAN DNS Name resolves to. In a NAS Scale Out configuration, there will be multiple IP addresses listed here.

The NAS LAN IP field only appears when the NAS Volumes feature is licensed.

2.7.25.2 DNS Server Buttons

Default — Click to reset all of the DNS Server settings to their default values.

Undo — Click to revert to the last saved settings.

Submit — Click to commit changes made to the DNS Server configuration input fields.

Note that changes here made do not require an SC reboot to take effect -- they are in effect immediately.

2.8 Users

Adding, editing, deleting, and viewing user information is done through the users function. There are two types of users on the *Storage Concentrator* system: an administrator and an observer. The administrator of the *Storage Concentrator* system manages the setup of hardware, the configuration of the storage management system, and the setup of users.

When logging in for the first time, the administrator can use the following: User ID: **stonefly**. Password: **stonefly**. It is strongly recommended that the stonefly password be changed at the initial configuration. It is also recommended that each system administrator have an individual user ID and password. For more information on creating users, see "Adding Users".

An observer can view system information only; he or she cannot make any changes.

There can be only one administrator logged into the interface at any time. There is no limit to the number of simultaneous observer logins.

2.8.1 Adding Users

Each user should have a unique user ID and password.

To add a new user, follow these steps:

1 Log in with your user ID and password.

You must have the **Administrative (All)** privilege to be able to add users.

2 Click Users. The User Management Summary screen displays the current settings for any existing users.

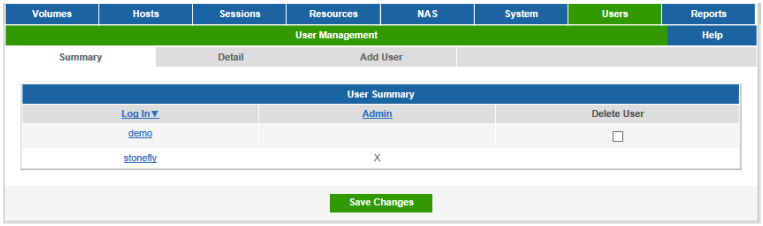


Figure 2-75 User Management, Summary screen

3 Click **Add User**. The Add New User screen appears.

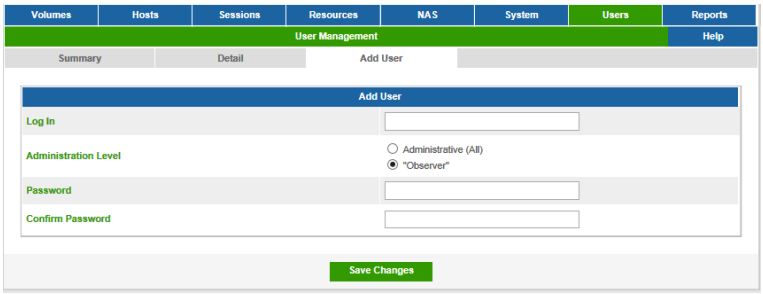


Figure 2-76 Add New User screen

4 In the **Log In** field, enter the user ID for the new user.

5 In the **Administrative Level** area, choose a user privilege for the new user:

Select **Administrative (All)** to give this user administrative privilege

Select **Observer** to give this user view-only privilege

6 In the **Password** field, enter a password for the new user.

The password cannot include the login name. The password must be between 6 and 15 characters long and one character must be a non–alpha character. However, StoneFly validates only the first eight characters. Characters after 8 are ignored but supported for user convenience.

7 In the **Confirm Password** field, retype the password for the new user.

8 **Use Nested Menus**: Select the radio button for **Yes** or **No.** The default is no. Nested menus are only available on certain browsers that support style sheets, such as Internet Explorer and Firefox. Nested menus allow faster menu-driven navigation.

If your browser does not support nested menus, use the default setting of **No**.

9 Click **Submit** to add the new user information to the *Storage Concentrator* database.

2.8.2 Editing Users

A user's privilege level or password may need to be changed. Modifications to user information can be made in the User Management Detail screen.

To edit a user, follow these steps:

You must have administrative privileges to edit users**.**

1 Log in with your user ID and password.

2 Click **Users**. The User Management Summary screen appears and the current settings for any existing users display.

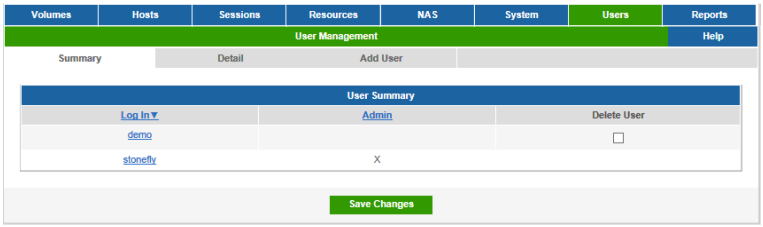


Figure 2-77 User Management, Summary screen

3 Click **Detail**.

4 Select the name of the user you want to edit from the **Select User** list. The User Management Detail screen appears.

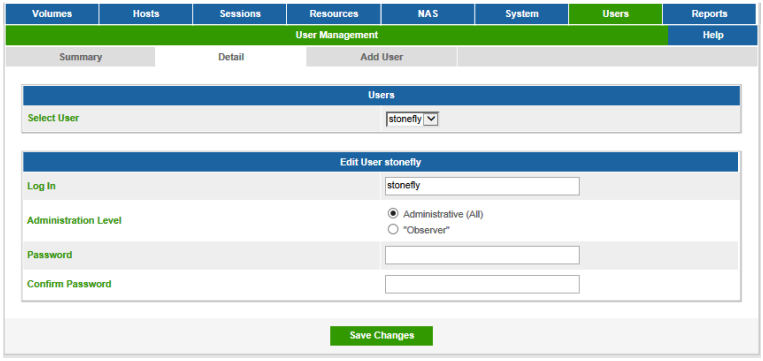


Figure 2-78 User Management, Detail screen

5 Change any of the following information:

Log In: Changes the login name of this user

Administration Level: Changes this user's privileges

Password: Changes this user's password

Confirm Password: Confirms the password change

6 Click **Submit** to save your changes.

2.8.3 Removing Users

To maintain system security, it is recommended that any user who should no longer have access to the *Storage Concentrator* be removed from the system.

To remove a user, use the steps that follow:

You must have administrative privileges to delete users**.**

1 Log in with your user ID and password.

2 Click **Users**. The User Management Summary screen appears and the current settings for any existing users display.

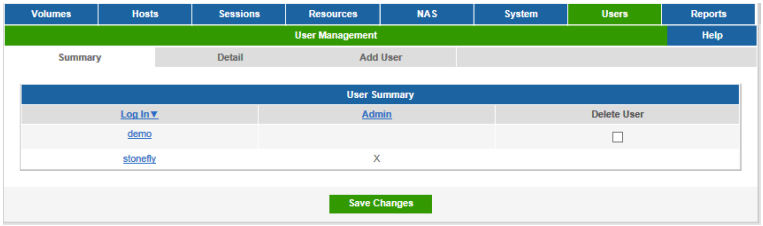


Figure 2-79 User Management, Summary screen

3 Click **Delete User** next to the name of the user you want to remove.

4 Click **Submit** to delete the user.

2.8.4 Viewing User Information

The *Storage Concentrator* provides user information for viewing only purposes.

To view a user's settings without making changes, use the steps that follow:

1 Log in with any user ID and password.

2 Click **Users.** The User Summary screen appears and the current settings for any existing users display.

2.9 Reports

The reporting function provides session information, logs of system events, system configuration detail and IO statistics. The session information reports statistical data about the *Storage Concentrator* device, resources, volumes, and sessions. The system event log tracks system-level sequences of events. This information can be used for troubleshooting. The configuration report provides detailed information regarding the system configuration. The Statistics reports display graphical representations of the IO flow to and from each volume in the system. These graphs are best used to make Active-Active Load Balancing adjustments.

2.9.1 Logs

The Log screen provides a sequential display of all system events. This information can be used when troubleshooting system problems. The number of log records kept in the database is set up in the System Management Admin screen. The maximum number of logs is the number of log records kept in the database. When this number is reached, the oldest record is overwritten when a new record is added. This prevents the log table from using up too much disk space. For more information on setting the number of records in the log file, see "Administrative Functions".

To view log files, use the steps that follow:

1 Click **Reports**. The Reports Log screen appears.

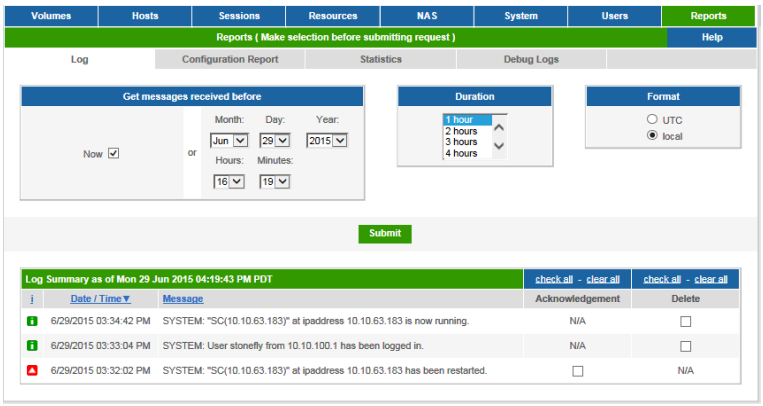


Figure 2-80 Log screen

3 In the duration area, indicate how many hours of historical log information you want to display.

4 Select **Now** (the current time) or a specific date and time. This is the starting point for the system events log that will be displayed.

5 Click **Submit** to display the historical system events log information. All logs with data gathered within the selected parameters displays on screen.

The Log information includes:

Time: Time that the log record was recorded in UTC or local time

Message: The text of the log record

Acknowledge: Indicates that a system-critical message has been generated and requires an acknowledgement by the system administrator

System critical warnings generate the status "Alert" on each page. These warnings must be acknowledged before the Alert status can be cleared.

Multiple copies of identical log entries may be suppressed to conserve log space.

Delete: Click the checkbox to delete the message from the Log report

2.9.2 Managing Logs

Logs should be periodically viewed to monitor the system status. If a system-critical message is generated, a check box will appear in the Acknowledge column of the Log Report screen. You may wish to manually delete superfluous log records to view messages more easily.

To acknowledge system-critical messages, use the steps that follow:

1 Click the check box under **Acknowle**dge on the Log Report screen.

2 Click **Submit**. The system is updated with the acknowledgement.

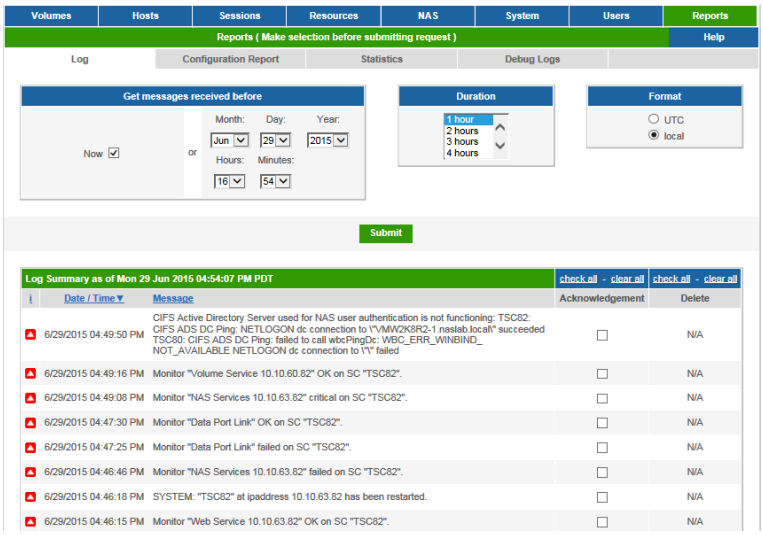


Figure 2-81 Log Management Screen

To manually delete log records, use the steps that follow:

1. Click the check box under **Delete** on the Log Report screen.
2. Click **Submit**. The selected logs will be deleted and the screen will refresh to show the updated information.

Log entries can be sorted using the column indicator at the top.

2.9.3 Configuration Report

The configuration report provides an onscreen and printable version of all system configuration information.

To print a configuration report, use the steps that follow:

1. Click **Reports**.
2. Click **Configuration Report**. Complete system configuration information appears on the screen.

Figure 2-82 Configuration Report screen

3 Click **Printable Version**. A printable copy of the system configuration information is generated. The configuration report information includes:

Storage Concentrator information:

System Name: Name assigned to the Storage Concentrator

iSCSI Listening Port: Network setting for the iSCSI Listening port

Local iSCSI Data Port IP Address: Network settings for the Data port

Management port IP Address: Network settings for the Management port

Default Gateway: Network settings for the default gateway

Volume configuration information for all volumes:

Volume Name: Name assigned to the volume

Notes: Descriptive information about the volume

Block: Block size in bytes

iSCSI Target Name: Name assigned to the resource during configuration

Type: Span is a volume type for a managed volume. **Pass Thru** is the volume type for volumes on resources designated as pass thru. **Mirror** is a volume type for a volume with mirror images.

Segment Number: Location on the resource where the volume resides

LUN Name (Block Size): Name assigned to the resource where the volume resides and the block size in bytes

Segment Size (GB): Total number of gigabytes allocated to this segment from this particular resource

Start Block: Block number where the volume segment starts on the resource.

End Block: Block number where the volume segment ends on the resource.

Resource information for all resources:

Serial Number: Serial number assigned by the manufacturer

BUS: SCSI information provided by the resource

HBA: SCSI information provided by the resource

Target ID: SCSI information provided by the resource

LUN #: SCSI information provided by the resource

Target Portal: Address for the target portal.

If the resource has a target portal address, it does not display HBA, BUS, or LUN and instead displays the target portal address.

Type: Type of SCSI storage device. This information is provided by the resource

Manufacturer: Name of the manufacturer of the resource device

Model: Model number assigned by the manufacturer

Firmware Version: Current version of firmware running on the resource

Block Size: Number of bytes in each block

Block Count: Number of blocks on the resource

Storage Size (GB): Total number of gigabytes of storage space on the resource

Unallocated Space (GB): Number of gigabytes of space available for new volumes

Type: Managed, Pass Thru, or None

2.9.4 Statistics

The Storage Concentrator monitors the number of reads and writes to all volumes. The counts are reported to the Administrative Interface periodically. The Statistics feature of the Administrative Interface analyzes the data between two of the sample periods and produces a graphical representation of the data.

For example, a chart of the number of Read IO's for a particular volume between 8:00 AM and 12:00 PM (noon) can be graphed. The graph can be saved as an HTML image or the requestor may choose to have a tabularized version of the data points in a comma separated value file. The file may be examined using a program such as Microsoft Excel to analyze the data further.

The statistics on IO flow are useful in many different ways. One of the best uses for these statistics is in setting the Active-Active Load Balancing. The use of the feature and its application to the Active-Active features are explained in Section 3: Failover.

2.9.5 Debug Logs

The Reports Debug Log screen allows a user with the Storage Concentrator administrative privileges collect system debug logs for submission to StoneFly Customer Support for analysis. It also provides debug log data generation, and maintenance functions.

To collect debug logs, use the steps that follow:

1 Click **Reports**.

2 Click **Debug Logs**.

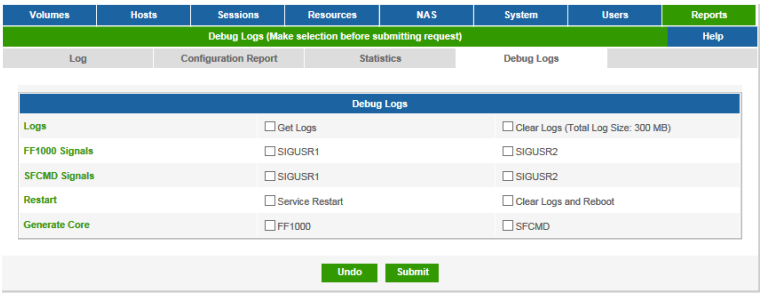


Figure 2-83 Debug Logs

3 Select the appropriate options, and then click **Submit**.

4 The fields are:

User ID, Password — The "User ID" and "Password" fields appear when the "Debug" page is used and the user has not already logged into the Storage Concentrator. This will always be the case when the "Debug" GUI page is used on a secondary SC in a cluster.

An administrator level user id must be used to perform any of the operations on this page.

When the "User ID" and "Password" fields are displayed, and are empty, they need to be filled in before the "Submit" button is pressed. To logout, click the "Logout" link on the top toolbar.

Get Logs — Check the "Get Logs" check-box to collect Storage Concentrator debug logs. These logs are often requested by StoneFly Customer Support to help resolve technical issues.

A single password protected compressed archive file is generated and saved on the host running the browser. Usually, this file would be uploaded to the StoneFly support ftp site.

When the "Get Logs" check-box is checked, default settings for other related check-boxes are also checked.

Note: When the Storage Concentrator is operating in a cluster, it is usually necessary to collect logs on both the primary and secondary SC. To collect logs on the cluster secondary SC, browse to its management page, push the "Get Logs" button, and enter an administrative user name and password in the fields provided.

Because the debug logs can be large, it may take some time to collect, archive, and compress them. An estimate is made on the size of the compressed archive before it generated.

The logs are not collected until the "Submit" button is pressed.

Clear Logs — Check the "Clear Logs" check-box to delete Storage Concentrator debug logs. This is sometimes done after a set of logs have been collected to reduce the size the next time logs are collected.

An estimate of the total disk space consumed by debug logs is displayed.

Because the system is not restarted as part of clearing the logs, some of the opened log files cannot be deleted. The total disk space consumed should not be expected

to reach 0.

Note: When the Storage Concentrator is operating in a cluster, one should consider performing the Clear Logs on both the primary and secondary systems.

The logs are not cleared until the "Submit" button is pressed.

FF1000 or SFCMD, SIGUSR1 or SIGUSR2 — The "SIGUSR1" and "SIGUSR2" check-boxes direct the system to send a signal to the "FF1000" or "SFCMD" Storage Concentrator process to flush to disk memory stored debug log information prior to collecting the debug logs.

Unless otherwise instructed by StoneFly Customer Support, the default settings that are made when "Get Logs" is checked are adequate.

The enabled signals are not sent until the "Submit" button is pressed.

Service Restart — The "Service Restart" check-box directs the system to perform a Storage Concentrator software restart instead of the hardware reset that is normally performed. Functionally, the system will appear as if it had been rebooted.

The "Service Restart" option should not be used, unless directed to do so by StoneFly Customer Support.

The "Service Restart" does not occur until the "Submit" button is pressed.

Clear Logs and Reboot — Check the "Clear Logs and Reboot" check-box to delete all of the Storage Concentrator debug logs, and then automatically reboot it.

The system reboots to allow the deletion of currently opened files.

If an interruption in SC operation is not desired, consider using the "Clear Logs" function instead.

The logs are not cleared and the system rebooted until the "Submit" button is pressed.

Chapter 3

FailOver

3.1 Introduction

FailOver is an important fault tolerance function of mission-critical systems that require constant accessibility. FailOver adds another layer of redundancy to network storage. In software releases after 5.0 the clustered Storage Concentrators operate in a mutual, Active-Active configuration. Both cluster members may receive IO requests from the hosts.

Before setting up a FailOver cluster, each Storage Concentrator must be configured using the procedures in the Storage Concentrator Setup Guide, Section 4, "Initial Installation." Failure to do so can result in damage to the Storage Concentrator and interruption of network services.

3.1.1 FailOver Concepts

FailOver

If a component in either cluster member fails, FailOver automatically reassigns the operations of the failed system to the surviving system. FailOver automatically redirects user requests from the failed system to the good system.

Active-Active Clusters

A cluster consists of two *Storage Concentrators.* In a clustered system there is one Primary and one Secondary *Storage Concentrator*. All hosts access the cluster through a single IP Address. This allows all iSCSI target discovery to be accomplished with a single entry in the iSCSI initiator interface on the host. The session between the host and its discovered targets may be directed to either of the Storage Concentrators by the Administrator. The balance of IO traffic between the two Storage Concentrators may be adjusted at any time using the Administrative Interface on the Primary Storage Concentrator.

Primary Storage Concentrator

The Primary *Storage Concentrator* is the *Storage Concentrator* in a cluster that displays the User Interface and processes all management requests. In the default configuration all IO from the hosts for any volume is processed by the Primary *Storage Concentrator*.

Secondary Storage Concentrator

The Secondary *Storage Concentrator* is the other member of an Active-Active cluster. It may be configured to process the IO requests for volumes created for the IPSAN. The Secondary *Storage Concentrator* does not display the User Interface. It monitors all FailOver conditions and in the event of a failure at the Primary *Storage Concentrator*, takes over the all active sessions. As a result of the Failover, the Secondary *Storage Concentrator* in the cluster transparently becomes the Primary *Storage Concentrator*.

The original Primary Storage Concentrator's volumes are now managed by the new Primary Storage Concentrator. There can be only one Secondary Storage Concentrator in the Active-Active cluster.

Active-Active Load Balancing

Active-Active Load Balancing is the process of assigning and re-assigning volumes to the Primary or Secondary Storage Concentrators. There is a screen to control the location of all volumes, allowing adjustment of the assignments. The decision to re-balance is supported by statistics screens in the Administrative Interface that periodically report the utilization of each volume. In addition, the System->Information screen reports the amount of free memory and CPU for each system.

Network-based Adaptive Load Balancing

As part of highly redundant system design, a Storage Concentrator supports more than one path from servers to storage. This ensures there is no single point of failure in the network. Multiple Network Interface Cards (NIC) adapter ports with Adaptive Load Balancing in the host servers provide emergency backup protection, as well as increased bandwidth. If one server link goes down, the other port automatically accepts the additional load with no interruption in server operation.

In addition, a dual-port gigabit Ethernet card option in the *Storage Concentrator* Cluster supports Adaptive Load Balancing across the two ports, providing increased bandwidth and port FailOver. For information on using the dual-port Adaptive Load Balancing feature, see "Typical Configurations for using Dual iSCSI GbE Ports".

FailOver Events

The following events put the system into FailOver mode:

- Manual FailOver from the Administrative Interface
- Power failure on the Primary or Secondary *Storage Concentrator*
- Network connection failure on the Data Network port if using a single Data port. If you are using a dual Data port card with ALB in the Storage Concentrator Cluster then a FailOver event will be triggered only if both ports go down.
- Back-end resource failure to the Primary or Secondary Storage *Concentrator*
- Network connection failure from the management port to the Primary or Secondary *Storage Concentrator*

3.2 Setting up FailOver

Setting up FailOver is a three-step process: First, at the **Primary** *Storage Concentrator*, you must enter the IP address of the *Storage Concentrator* that is applying to become the **Secondary** *Storage Concentrator* in this cluster and setup a cluster IP address against which the hosts will perform iSCSI target discovery. Second, at the *Storage Concentrator* that is applying to become the **Secondary** *Storage Concentrator*, you must enter the IP address of the **Primary** *Storage Concentrator* in this cluster. Third, the hosts must be assigned to the cluster.

Specific details how to set FailOver Configuration for system that supports NAS volumes can be found at section 9.7.

3.2.1 Designating a Primary Storage Concentrator

To designate a Primary *Storage Concentrator*, use the steps that follow:

1. Launch the administrative interface using the Management port IP address of the *Storage Concentrator* that will be the **Primary** *Storage Concentrator* in the cluster.
2. Log into the administrative interface with your user ID and password.
3. Click **System**.
4. Click **Admin** and then **FailOver**. The **FailOver Setup Cluster** screen appears.

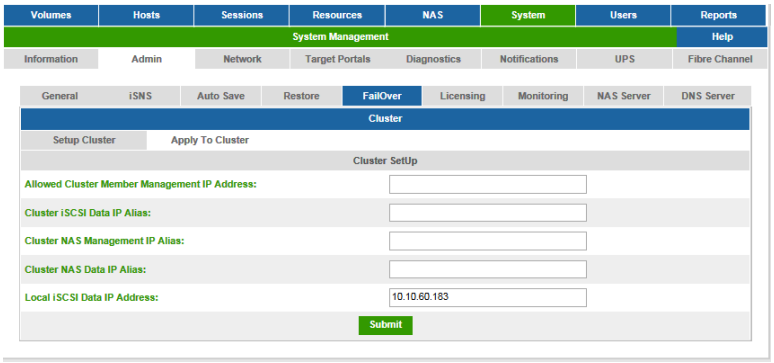


Figure 3-1 System Management, FailOver Setup Cluster screen

The information on the screen includes:

Allowed Cluster Member Management IP Address: The Management port IP address for the Secondary *Storage Concentrator* in the cluster is entered here.

Cluster iSCSI Data IP Alias: The IP address that allows hosts to connect to the storage volumes that are part of the cluster is entered here.

Cluster NAS Management IP Alias: The Management Network IP Address of the FailOver Cluster to use by NAS clients is entered here.

Cluster NAS Data IP Alias: The Data Network IP Address of the FailOver Cluster to use by NAS clients is entered here.

A host must be part of the cluster to be protected in a failover event. If a host is not addressing the cluster through its **Cluster iSCSI Data IP Alias** and one of the *Storage Concentrator* fails, the host will no longer have access to the back-end storage managed by the other *Storage Concentrator*.

If the cluster iSCSI Data IP alias is changed at the Primary *Storage Concentrator*, all hosts in the cluster must be reconfigured with the new cluster iSCSI Data IP alias.

Local iSCSI Data IP Address: By default, the IP address for the Local iSCSI Data Network port appears here. Change this setting only if this IP address is used for the Cluster iSCSI Data IP alias.

5 In the **Allowed Cluster Member Management IP Address** field, enter the Management Network IP address for the *Storage Concentrator* that is applying to be the **Secondary** *Storage Concentrator* in this cluster.

6 In the **Cluster iSCSI Data IP Alias** field, enter the IP address that allows hosts to connect to the storage volumes that are part of the cluster.

The Cluster iSCSI Data IP alias must be different from the Local iSCSI Data IP address. If you use the IP address of the Local iSCSI Data IP address for the Cluster iSCSI Data IP alias, change the IP address of the Local iSCSI Data IP Address before submitting the settings.

7 The IP address that was entered for the iSCSI Data IP Address port during initial setup appears in the **Local iSCSI Data IP Address** field. Change this setting only if this IP address is used for the Cluster iSCSI Data IP Alias. For more information on setting up the iSCSI Data port, see "Local iSCSI Data Port Settings".

8 Setup for **Cluster NAS Management** and **Data IP Aliases** is exposed on screen only in case when system has NAS Volume license. In this case both aliases have to be provided.

9 Click **Submit**. The following screen appears.

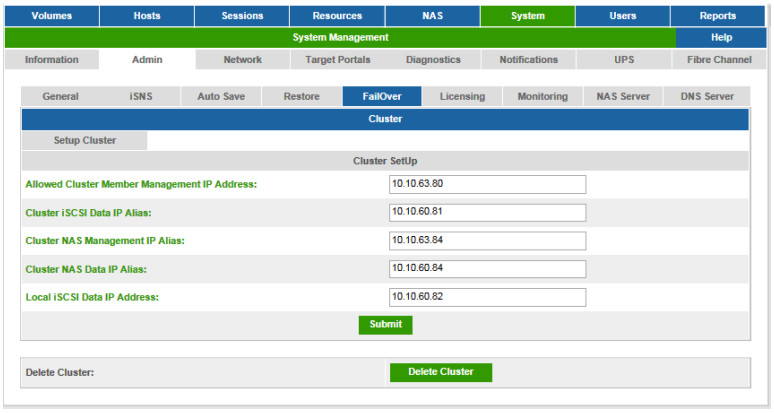


Figure 3-2 System Management, FailOver Setup Cluster screen – Waiting to accept Secondary member

This screen displays with the same contents prior to accepting the Secondary member into the cluster.

If the Allowed Cluster Member Management IP address of the *Storage Concentrator* that is applying to be the Secondary *Storage Concentrator* is changed before this *Storage Concentrator* has applied and been accepted, the *Storage Concentrator* will not be allowed to join the cluster. For more information on applying the Secondary *Storage Concentrator* to the cluster, see "Applying the Secondary Storage Concentrator to the Cluster".

The **Delete Cluster** button allows you to remove the *Storage Concentrator* from the cluster. For more information, see "Making a Storage Concentrator a Stand-Alone System".

Once the Secondary *Storage Concentrator* is accepted into the cluster, the following screen appears.

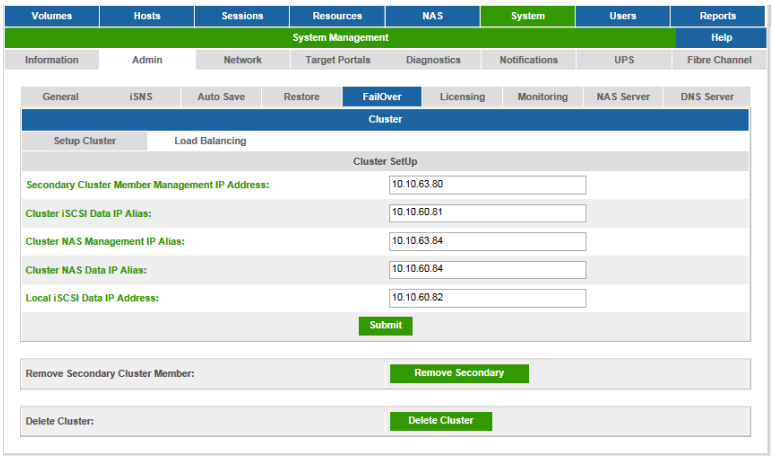


Figure 3-3 System Management, FailOver Setup Cluster screen - Secondary accepted to cluster

The new information on the screen includes:

The **Remove Secondary** button. This button is used to eliminate the Secondary *Storage Concentrator* from the cluster. In the event of a major equipment failure where the entire unit must be replaced it is necessary to remove the original Secondary *Storage Concentrator* and re-apply from the new Secondary *Storage Concentrator.*

3.2.2 Applying the Secondary *Storage Concentrator* to the Cluster

To apply to the cluster from the Secondary *Storage Concentrator*, use the steps that follow:

1. Launch the administrative interface using the Management GbE port IP address of the *Storage Concentrator* that will be the **Secondary** *Storage Concentrator* in the cluster.
2. Log into the administrative interface with your user ID and password.
3. Click **System**.
4. Click **Admin** and then **FailOver**.

   The **FailOver Setup Cluster** screen appears. (see "System Management FailOver Setup Cluster screen - submitting to cluster" above)

5. Click **Apply to Cluster**. The following screen appears.

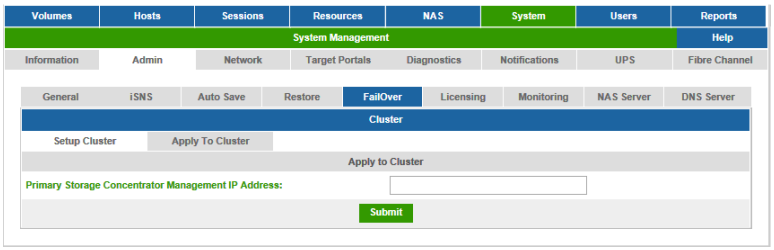
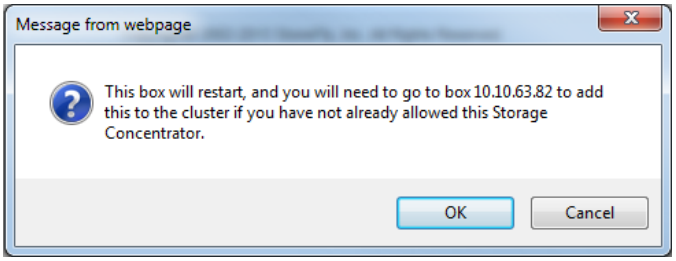


Figure 3-4 System Management, FailOver Apply to Cluster screen

6 In the **Primary Storage Concentrator GbE IP Address** field enter the Management GbE port IP address of the **Primary** *Storage Concentrator* in the cluster.

7 Click **Submit**. The following message appears.



8 Click **OK** to continue.

The following screen appears and remains displayed until the Secondary *Storage Concentrator* is accepted as a member of the cluster.

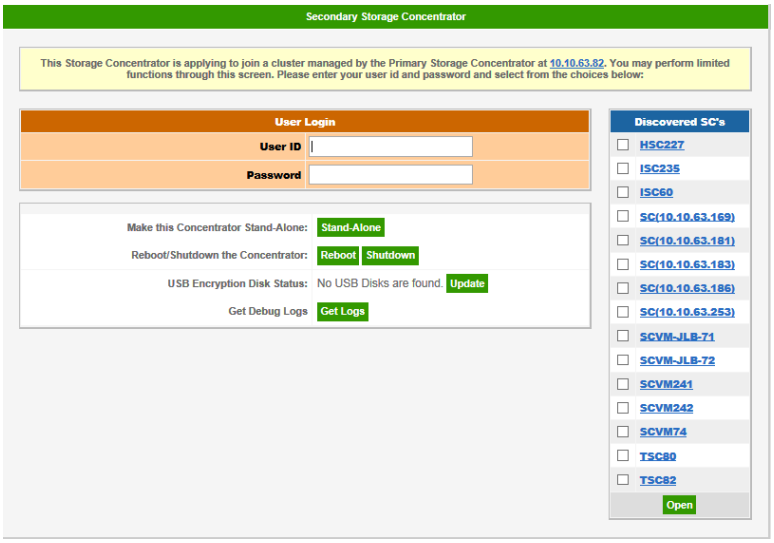


Figure 3-5 System Management, Apply to Cluster screen -Joining

The message at the top of the screen indicates that this Secondary *Storage Concentrator* is applying to join the cluster. When this Secondary *Storage Concentrator* is accepted as a member of the cluster, the following screen appears.

To see the following screen, you must refresh your browser. The Secondary unit will join the cluster. The screen will remain unchanged until the browser is refreshed.

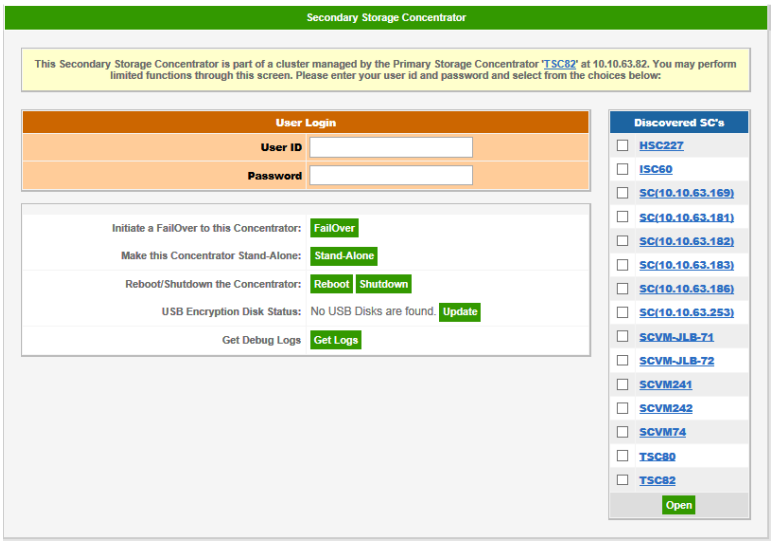


Figure 3-6 System Management, Apply to Cluster screen - accepted to cluster (with licensed Volume Encryption)

The information at the top of the screen confirms that this Secondary *Storage*

Concentrator has joined the cluster that is being controlled by the Primary *Storage Concentrator* named on the screen.

When a *Storage Concentrator* has applied to or been accepted to a cluster as the Secondary *Storage Concentrator*, only the following functions are available from that *Storage Concentrator's* Administrative Interface:

Making the Secondary *Storage Concentrator* a stand-alone system

Rebooting the Secondary *Storage Concentrator*

Shutting down the Secondary *Storage Concentrator*

In order for a host to be protected by the Secondary *Storage Concentrator* in a FailOver event, the host initiator must be configured to point to the cluster IP address that is designated at the Primary *Storage Concentrator* Cluster Setup screen.

If the cluster iSCSI Data IP Alias is changed at the Primary *Storage Concentrator*, all hosts in the cluster must be reconfigured with the new cluster iSCSI Data IP alias.

3.2.3 Assigning Hosts to the Cluster

To configure the host initiator to point to the Cluster iSCSI Data IP alias, do one of the following:

At the host initiator, configure the host to point to the target at the Cluster iSCSI Data IP alias designated on the Primary *Storage Concentrator* Cluster Setup screen. For more information on configuring the host initiator, see the documentation that is provided with the host system.

The current Local iSCSI Data IP Address may be entered in the Cluster iSCSI Data IP alias field. If you use the IP address of the Local iSCSI Data IP Address for the Cluster iSCSI Data IP alias, change the IP address of the Local iSCSI Data IP Address before submitting the settings. For more information on setting up the iSCSI Data port, see "Local iSCSI Data Port Settings".

The Cluster iSCSI Data IP alias must be different from the Local iSCSI Data IP address. If you use the IP address of the Local iSCSI Data port for the Cluster iSCSI Data IP alias, change the IP address of the Local iSCSI Data port before submitting the settings.

3.2.4 Active-Active Load Balancing

At any time after volumes have been created and a Failover cluster is configured, the Failover screen allows access to the **Load Balancing** button.

To adjust the Load Balancing within a Failover Cluster do the following steps:

1 Navigate to the Failover screen by clicking **System**

2 Click **Admin** and then **Failover**

3 Click on **Load Balancing** and the following screen displays

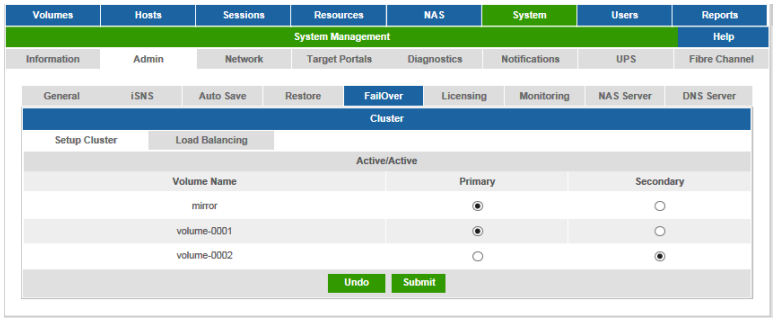


Figure 3-7 System Management, Load Balancing screen

Active-Active Load Balancing

4 Select the Storage Concentrator to service the connection to the host system, either Primary or Secondary.

5 Click the appropriate radio button for each volume that needs to be re-assigned.

6 Click **Submit**. The screen re-displays.

7 Click on the **Sessions** button to display the Sessions screen as below:

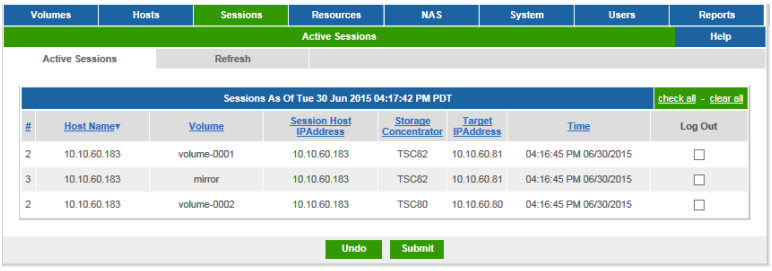


Figure 3-8 System Management, List of Active Sessions

Observe the column titled "Storage Concentrator". The name of the Storage Concentrator to which the volume has been assigned appears in the column.

3.2.5 Using Statistics to accomplish Load Balancing

The process of deciding to re-assign a volume from one cluster member to the other can be based on statistical analysis of the IO flow to the volume and utilization of the clustered Storage Concentrators. A statistics screen is provided that periodically report the IO flow for each volume according to each member of the cluster.

To display statistical information about volumes, perform the following steps:

1 Click on the **Reports** navigation button.

2 Click on **Statistics** button to display the following screen:

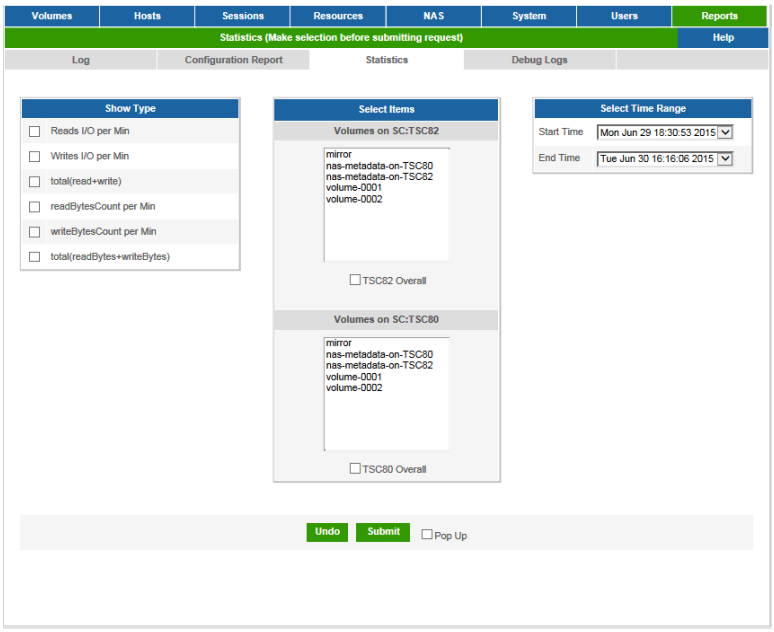


Figure 3-9 Statistics Screen

The **Statistics** screen contains the following information and buttons:

Show Type: This area of the screen is used to select the type of information to be display for each volume:

Reads IO per Min: total number of read IO's performed on the volume, expressed as IO's per minute.

Writes IO per Min: Total number of write IO performed on the volume, expressed as IO's per minute

Total (read+write): Show the sum of the two items above.

readBytesCount per Minute: Show the number of bytes read per minute for a specified volume

writeBytesCount per Minute: Show the number of bytes written per minute for a specified volume

totalBytesCount: Show the total number of bytes read or written per minutes for the specified volume.

Select Items: A list of the volumes on each Storage Concentrator. Select the volume you want to evaluate on each system.

Tables of Volumes: Select one or more Volume on each Storage Concentrator or use the **Overall** button.

Overall: The **Overall** button is used to select a summary for all IO traffic on a Storage Concentrator. The name of each Storage Concentrator is displayed at the side of the **Overall** button.

Select Time Range: Specify the start time and end time of the period to be analyzed and graphed.

Start Time: Select a time from the list of times for which statistics were taken.

Stop Time: Select a time from the list of times for which statistics were taken.

3 Make all selections and click on **Submit.** The resulting graph is the analysis of the data captured between the **Start Time** and the **Stop Time**.

An example of a graph is shown below. In this graph the time period selected is a 12 hour period. The **total IO per Min** type of data is selected for the **Overall** results of each Storage Concentrator. In this case the resulting graph indicates that during the early portion of the period all volumes were assigned to the Primary SC. Approximately nine hours into the period one of the volumes was re-assigned to the Secondary SC. The two different **total IO** lines after the ninth hour indicate activity on both Storage Concentrators.

The **Get CSV** button on the graph provides a Comma Separated Value (CSV) file containing the data points used in the graph. The CSV file can be loaded into a program like Microsoft Excel for additional analysis.

4 Using the data displayed in the Statistics graph, volumes may be re-assigned to balance the load. Graphs might be generated for each volume on each Storage Concentrator and then compared to the **Overall** charts to equalize the activity levels. Be sure to consult the **Free CPU** and **Free Memory** percentages on the System->Information

screen as another indicator of the activity on each system.

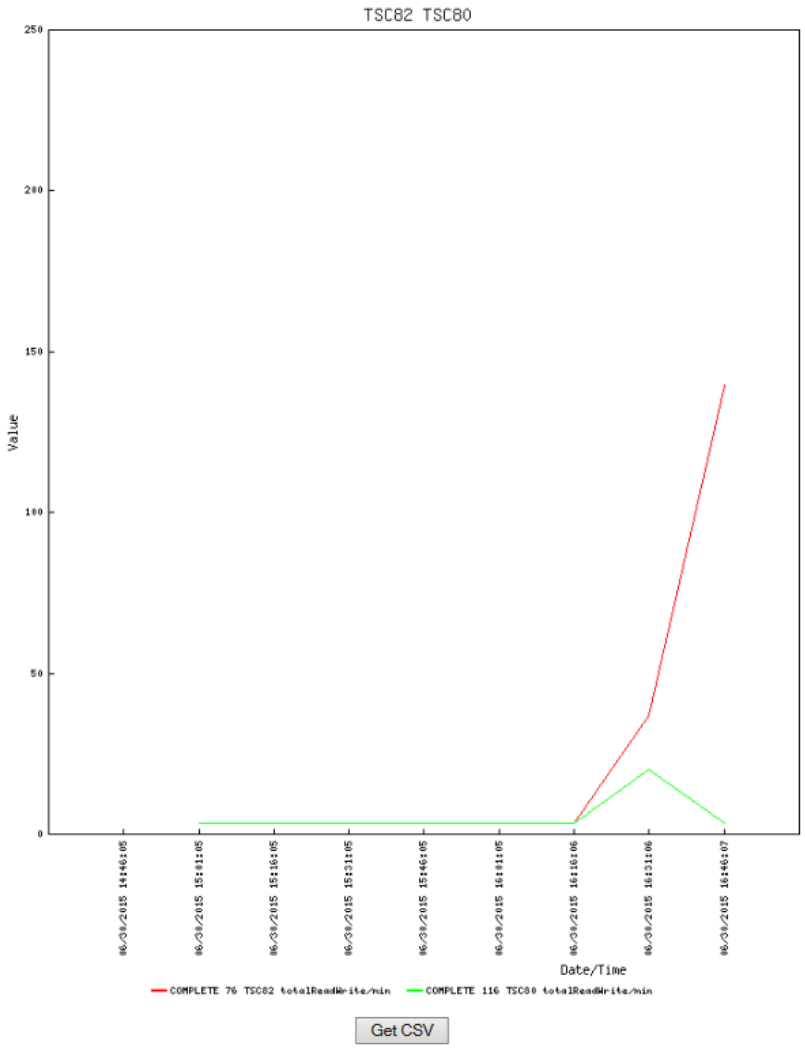


Figure 3-10 An Example of Statistical Graph

3.2.6 Making a *Storage Concentrator* a Stand-Alone System

Making a *Storage Concentrator* a stand-alone system, removes the *Storage Concentrator* from the cluster. You can make the Primary *Storage Concentrator* or the Secondary *Storage Concentrator* a stand-alone system. Make the Secondary *Storage Concentrator* a stand-alone system before making the Primary *Storage Concentrator* a stand-alone system. Typically, a Secondary *Storage Concentrator* is made a stand-alone system if it no longer needs to be part of a cluster or was added to the wrong cluster.

If a Secondary *Storage Concentrator* was added to the wrong cluster, it must first be deleted from the incorrect cluster and then added to the correct cluster.

When a Secondary *Storage Concentrator* is made a stand-alone system, it reinitializes the system files and loses all the information regarding the volumes that are currently managed by the **Primary** *Storage Concentrator*. When a **Primary** *Storage Concentrator* is made a stand-alone system, all system information remains intact.

To make the Primary *Storage Concentrator* a stand-alone system, use the steps that follow:

1 At the Primary *Storage Concentrator*, log in using your administrative user ID and password.

2 Click **System**,

3 Click **Admin** and then **FailOver**. The **FailOver Setup Cluster** screen appears.

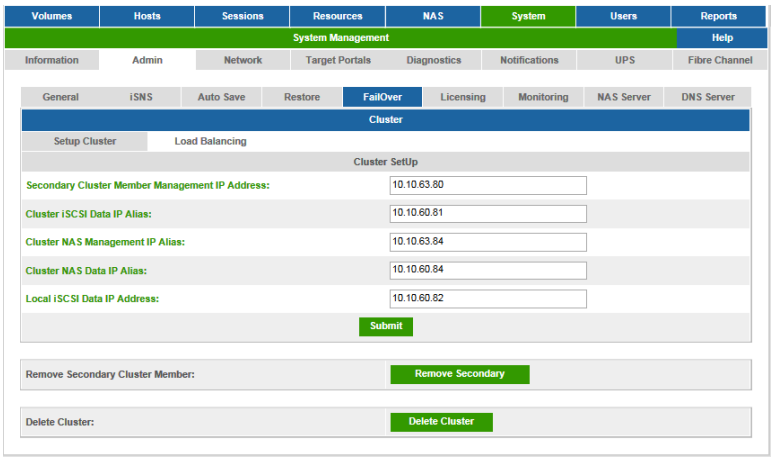


Figure 3-11 System Management, FailOver Setup Cluster screen

4 Click **Delete Cluster**. The Secondary Cluster Member Management IP Address field, the Cluster NAS iSCSI Data and Management IP Alias fields and the Cluster iSCSI Data IP alias field on the FailOver Cluster Setup screen are cleared. The Secondary *Storage Concentrator* is put into Stand-Alone mode. The Secondary *Storage Concentrator* loses its knowledge of the resources and volumes. Only the Primary *Storage Concentrator* retains the volumes and resources.

To make a Secondary *Storage Concentrator* a stand-alone system, use the steps that follow:

1 At the Secondary *Storage Concentrator*, enter your user ID and password.

2 Click **Stand-Alone**.

Both systems will still be connected to the same Storage array(s). If you intend to cluster the two systems together again, leave the storage connections unchanged. You are able to use the stand-alone Storage Concentrator to manage different resources on the same array. You may use the new stand-alone system to manage a different storage array by changing the storage connections.

3.2.7 Rebooting the Secondary Storage Concentrator

To reboot the Secondary *Storage Concentrator*, use the steps that follow:

1 Enter your **user ID** in the user ID field.

2 Enter your **password** in the password field.

3 Click **Reboot**.

3.2.8 Shutting Down the Secondary Storage Concentrator

When adding or reconfiguring resources, the *Storage Concentrator* must be shut down.

To shut down the Secondary *Storage Concentrator*, use the steps that follow:

1 Enter your **user ID** in the user ID field.

2 Enter your **password** in the password field.

3 Click **Shutdown**.

3.2.9 Changing the Secondary *Storage Concentrator*

To change the Secondary *Storage Concentrator* in a cluster, use the steps that follow:

1. At the Primary *Storage Concentrator*, login with your user ID and password, and then click **System->Admin->Failover-Setup Cluster**.
2. Then click **Remove Secondary**.
3. In the **Secondary Cluster Member Management IP Address** field, enter the Management IP address for the *Storage Concentrator* that will become the new Secondary *Storage Concentrator*.
4. Click **Submit**. The new Secondary *Storage Concentrator* must now apply to join the cluster. Please review the instruction above on how to join a cluster.

3.2.10 No Maintenance Window Upgrades Using FailOver

Please review the information in Section 6, Software Upgrades, to complete an upgrade without taking the cluster offline.

3.3 Adaptive Load Balancing (ALB)

A dual-port gigabit Ethernet card option in the *Storage Concentrator* Cluster supports Adaptive Load Balancing across the two ports, providing increased bandwidth and port FailOver. If both ports are connected, ALB operates automatically with no user setup or intervention. For more information on this feature, see "Typical Configurations for using Dual iSCSI GbE Ports".

3.4 Replicating Storage Concentrator Server Volumes

StoneFly Replicator™ uses existing infrastructure to keep critical data protected and highly available. Replicator provides storage-independent replication that operates at the block level over an iSCSI network.

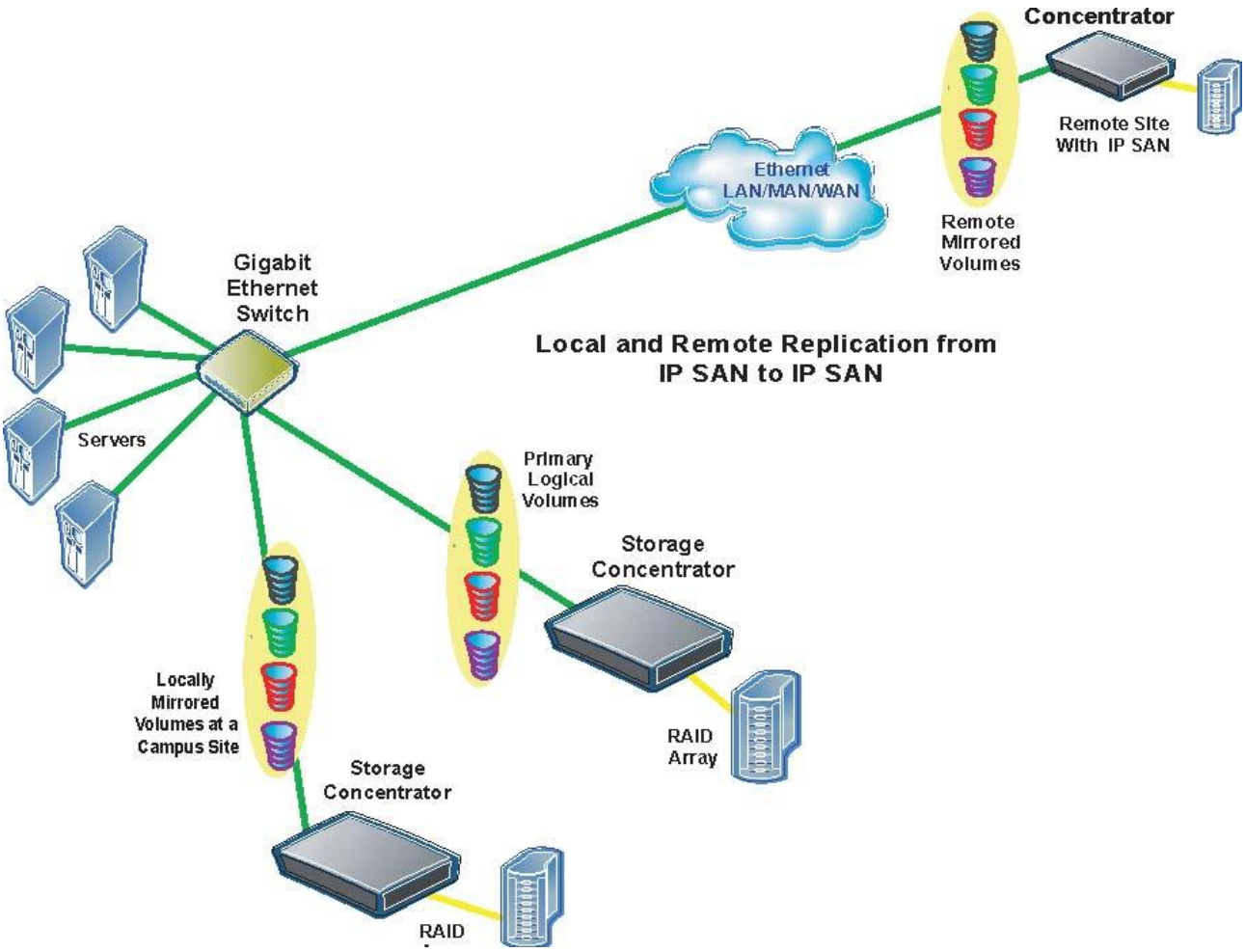


Figure 3-12 Replicating Server Volumes both locally and remotely

For more information about Replicator operation, see the StoneFly *Replicator Installation and User's Guide*.

Chapter 4

Synchronous Mirroring

This chapter provides an overview of StoneFly Mirroring™, including the steps necessary to add and manage synchronous mirror images of available storage volumes. For information on how to create a volume, see "Creating a Volume".

4.1 StoneFly Mirroring Overview

The StoneFly synchronous mirroring feature supplies host-independent mirrored data storage that duplicates production data onto physically separate mirrored target images transparently to users, applications, databases, and host processors. Synchronous mirroring implies that the *Storage Concentrator* waits for a write-complete acknowledgement from all volumes before presenting a write completion status to the host.

The software duplicates block-level changes as they occur to one or more volumes either to another local *Storage Concentrator* volume image or to another *Storage Concentrator* volume image at a campus location over standard IP connections.

Mirroring offers the benefit of protecting a critical volume from being a single point of failure, and providing continuous access to a volume without interruption to data availability when loss of access occurs to one of the images.

Synchronous mirroring is useful for applications such as:

- Backup/restore
- Business continuance, including high availability and no single point of failure configurations
- Archiving
- Data migration
- Disaster recovery
- Content distribution
- Business intelligence---data mining

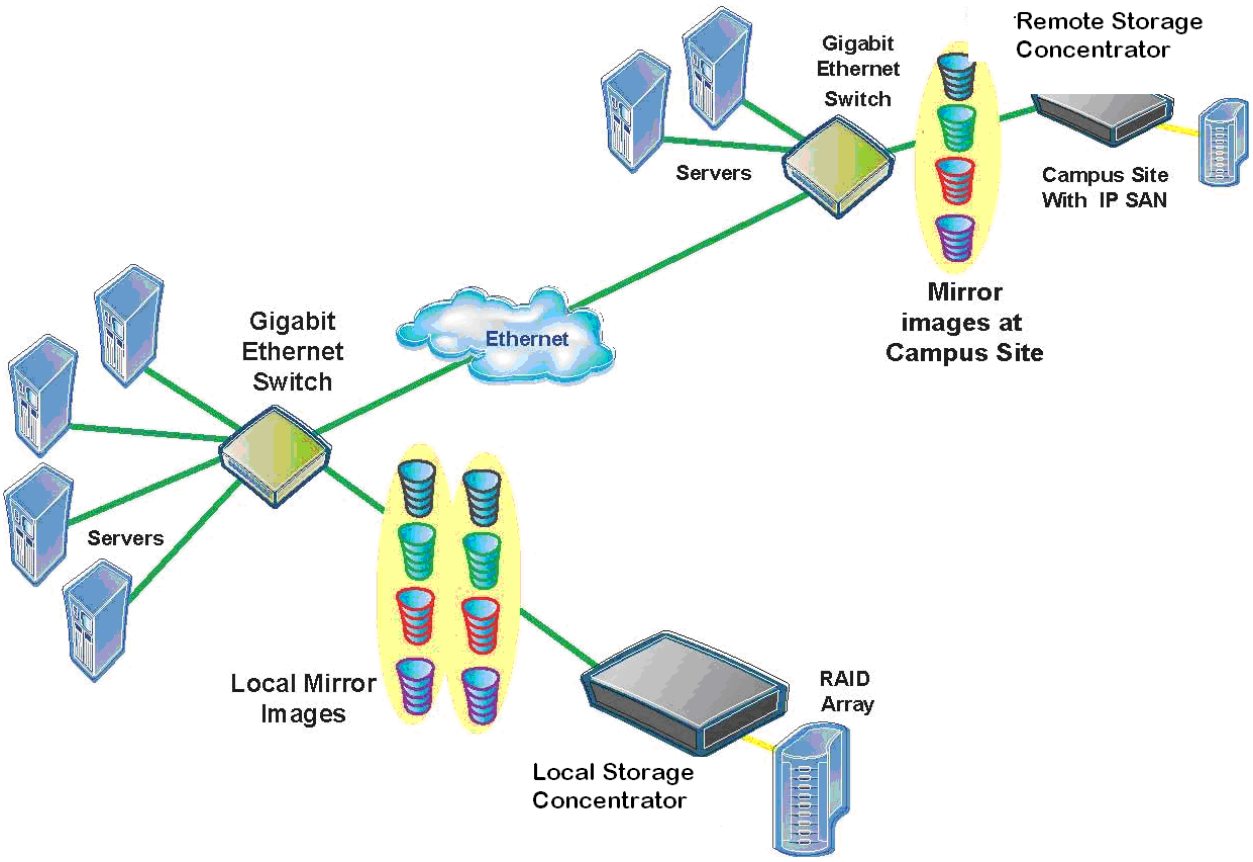


Figure 4-1 A Typical StoneFly Mirroring Configuration (Volumes Mirrored Locally and to a Campus Location)

4.1.1 Optimizing Synchronous Mirroring Environments

The StoneFly Mirroring campus mirroring feature is designed for use in networking infrastructure with sufficient bandwidth to handle data transfers. A dedicated gigabit Ethernet line provides an optimal connection between a local and campus mirror site. A gigabit Ethernet line that is already heavily used for I/O may not provide sufficient bandwidth.

Mirroring waits for a write-complete acknowledgement from all images in a mirror volume before presenting a write completion status to the host. To maximize performance, each mirror volume can have no more than four images simultaneously, in any combination of local and campus mirrors. However, the system can maintain many mirror volumes, each with no more than four images.

Distance increases latency, so the campus sites should be selected carefully to maximize both business continuity requirements and performance.

4.1.2 StoneFly Mirroring Terminology

Storage resources should also be selected with a degree of fault tolerance. Logically, every attempt should be made to select the resource that ensures that mirror images are on different physical resources or different drives than the original volume. In best storage practices, the original volume and the images will always reside on different physical storage resources.

Volume

The volume is storage presented to a host through the Storage Concentrator. The logical volume is the starting point before creating any mirror images.

Mirror Volume

A Mirror Volume is composed of multiple mirror images. Mirroring supports up to four mirror images in a mirror volume. The Mirror Volume is what the applications and hosts see as the storage device.

Mirror Image

A mirror image contains an exact duplicate of all other images in a mirror volume. A mirror image is grouped with other mirror images to comprise a mirror volume (each mirror image contains a duplicate copy of the data). Mirror images can be either local mirrors or they can be campus mirrors. The size of the mirror image must be equal to the size of all other images in the mirror volume.

Local Storage Concentrator

The Storage Concentrator that manages the mirror volume.

Local Mirrors

Mirrored images that are located behind the Local Storage Concentrator.

Campus Remote Storage Concentrator

Another Storage Concentrator in a campus configuration that presents mirror images to the Local Storage Concentrator. Either the Local Storage Concentrator or the Campus Remote Storage Concentrator or both can be single units or FailOver clusters.

Campus Mirrors

Mirror images that are behind the Campus Remote Storage Concentrator. Connection between the Local and Campus Remote Storage Concentrators should be a high speed, low latency connection.

Target Portal

An iSCSI communications gateway between a Local Storage Concentrator and a Campus Remote Storage Concentrator at a campus site. Once the target portal has been created, the system will recognize storage at the campus site as an available resource for creating Campus mirror images.

4.1.3 Mirroring Function Definitions

Secondary Servers

A server attached to the Campus Remote Storage Concentrator has access to campus mirror images; however it views them as a local virtual volume. A server attached to the Campus Remote Storage Concentrator can also have its own local mirror volumes. These volumes are accessible as campus mirrors from the Local Storage Concentrator.

Rebuild

The process of synchronizing a mirror image to an existing mirror volume. The volume is duplicated onto the mirror image at the block level in its entirety.

Detach Image

Detaching an image allows it to be mounted and used by other software applications. The most common usage is to make a backup copy of the detached image. The detached image retains the mirror volume information and can be reattached at any time.

Reattach Image

A previously detached mirror image can be reattached to the original mirror volume. Reattaching a mirror image removes it as a stand-alone image volume and initiates a rebuild operation on the reattached image.

Promote Image

Promoting a mirror image makes it a stand-alone volume. The promoted volume does not retain any mirrored volume information but contains an exact copy of the data at the time of promotion. The new stand-alone volume is accessible to the hosts and can retain security information from the mirror volume.

Once a volume has been promoted and exists as a stand-alone volume, it can then be used as a mirror volume. For example, a campus mirror image can be promoted and then mirrored to servers at the campus site for a safe and secure method of data migration.

Revert Mirror to Span

Reverting the Mirror to a Span volume converts the mirror volume to a regular spanned volume. This eliminates the critical warnings from having a mirror volume with only one good image. Users may want to Revert Mirrors to reduce overhead and increase system performance. Once a mirror volume has been reverted to a spanned volume, it can become a mirror volume by adding images to it.

4.2 Adding Mirror Images

This function allows the user to create a mirror volume using a previously defined Storage Concentrator volume. Once a mirror image is added to an existing non-mirrored volume, a mirror volume is created. Adding a mirror image initiates the rebuild operation on the newly added mirror image. Multiple mirror images can be added to a mirrored volume. They can be local mirror images or campus mirror images through a target portal.

4.2.1 To add a Mirror Image:

1 From the *Volumes* screen, click **Add Image**.

2 From the *Volume Select* drop down menu, select the correct volume. StoneFly Mirroring can mirror Snapshot Live Volumes or Snapspace, but not Snapshot Images.

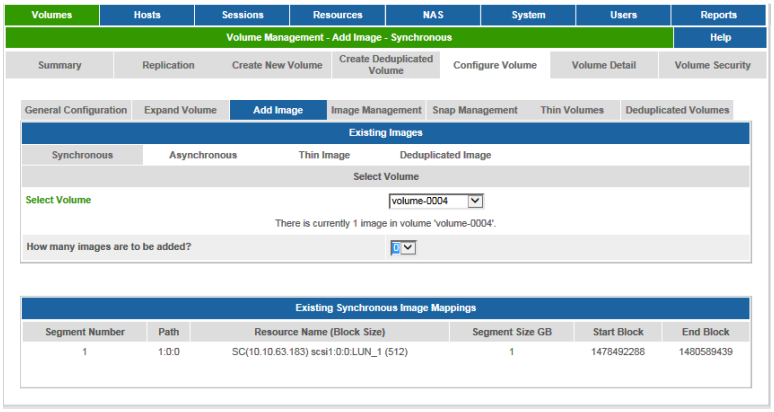


Figure 4-1 Volume Management Screen

3 From the drop down menu, select the number of images you wish to add. You can select any number less than the maximum of four images per Mirror Volume.

The screen displays an automatically generated Mirror Image name consisting of the name of the original volume and an image number in the *Image Name* field and displays available resources in the *Select Resources* table. In addition, images that were previously generated are listed in the *Existing Image Mappings* table.

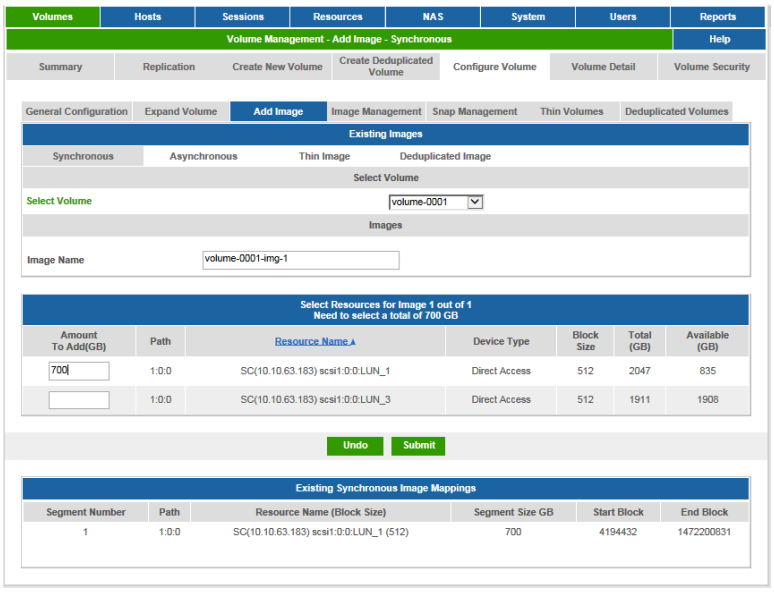


Figure 4-2 Volume Management with Image Mapping Tables

4 In the *Amount to Add (GB)* column of the *Select Resources* table, enter the total number of gigabytes in the row containing one of the available resources.

You may select the amount from one or more resources, but the total amount for the image must equal the size of the original volume.

To achieve fault tolerance, select the resource that ensures images are on different resources or different drives than the original volume. In best storage practices, the original volume and the images will always reside on different physical storage resources.

If you added more than one image, repeat this step until you have allocated space for each new image.

5 Click **Submit** to build the Mirror Image.

6 The image will automatically start building at the low priority level. You can change the Rebuilding Priority level from low to medium or high in the General Settings, but the system will use more operational resources to rebuild the image(s) at the higher priority setting.

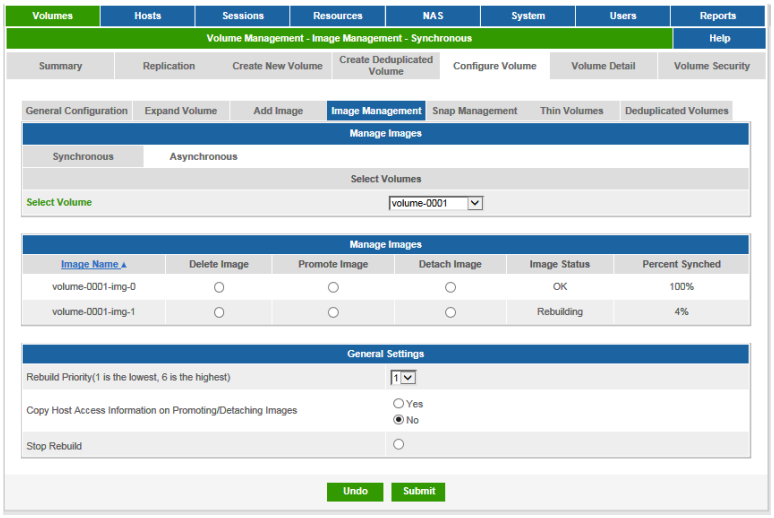


Figure 4-3 Volume Management General Settings

Changing the Rebuild Priority Setting from low to medium or high will consume more internal *Storage Concentrator* resources and should not be changed without considering the overhead on system resources and system performance. Medium and High settings should only be used if host I/O is not a priority.

From the Volumes screen, click **Summary**. The Operational State appears *Critical* if there is only one functioning image available. *Critical* will also appear if there is only one image available while one or more are rebuilding. The Operational State appears *Degraded* if an image fails and two or more images are functioning.

4.3 Adding Thin Mirror Image

4.3.1 To add a Thin Mirror Image:

1 From the *Volumes->Configure Volumes* screen, click **Add Image->Thin Image**.

2 From the *Volume Select* drop down menu, select the correct volume. The screen can be used to mirror Snapshot Enabled Live Volume, but not Snapshots or Deduplicated Volumes.

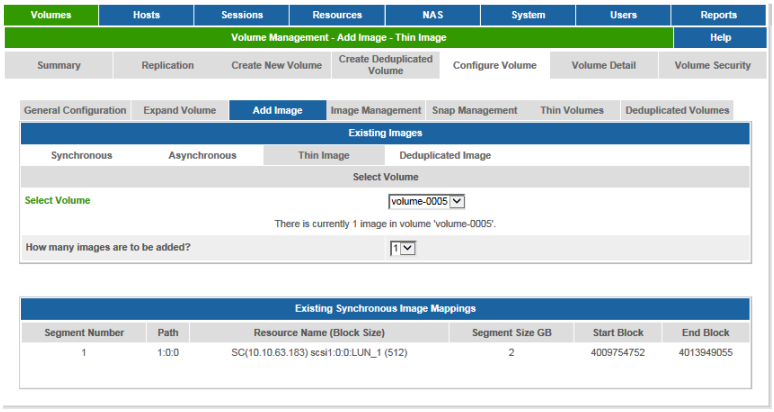


Figure 4-4 Volume Management, Add Thin Image

3 From the drop down menu, select the number of thin images you wish to add. You can select any number less than or equal to the maximum of allowed thin images per Mirror Volume. See Storage Concentrator Configuration Limits.

The *Image Name* field displays an automatically generated name for the Mirror Image consisting of the name of the original volume and an image number. Available thin pools are listed in the *Select Pool* table. Images that were previously generated are listed in the *Existing Image Mappings* table.

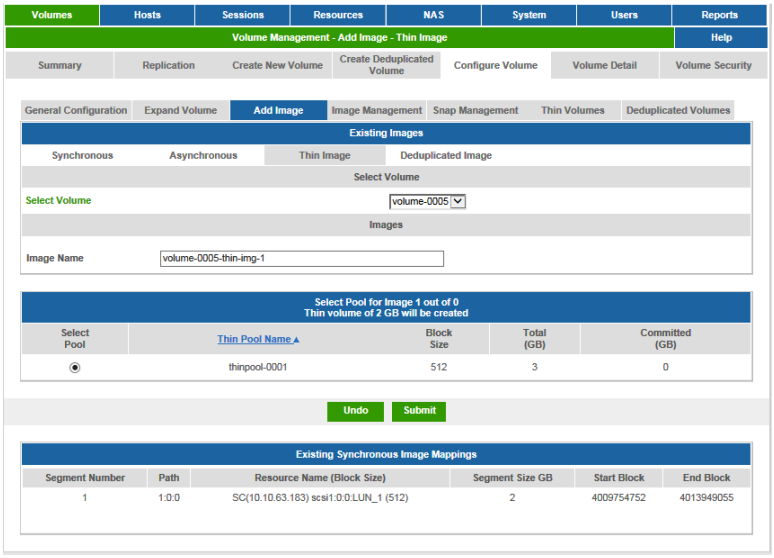


Figure 4-5 Volume Management, Existing Image mapping table

4 Select the Thin Pool as a space where the Thin Image has to be allocated.

5 Click **Submit** to build the Thin Mirror Image.

6 The image will automatically start building at the low priority level. You can change the Rebuilding Priority level from low to medium or high in the General Settings, but the system will use more operational resources to rebuild the image(s) at the higher priority setting.

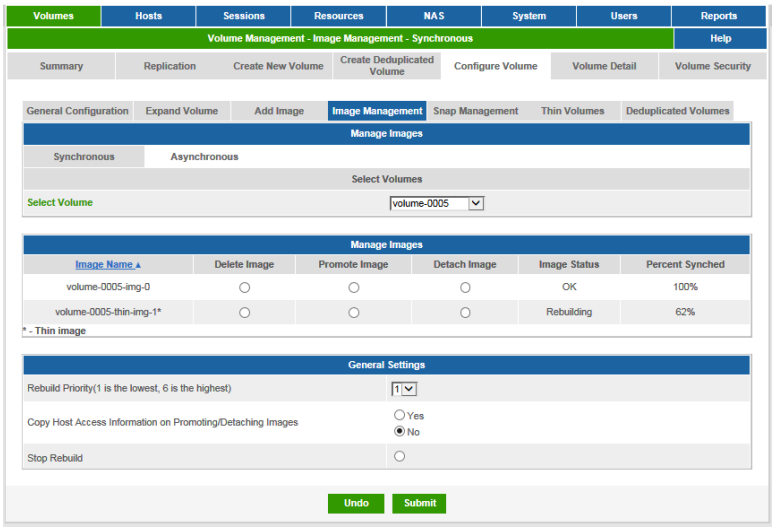


Figure 4-6 Volume Management, Existing Synchronous Image

Changing the Rebuild Priority Setting from low to medium or high will consume more internal *Storage Concentrator* resources and should not be changed without considering the overhead on system resources and system performance. Medium and High rebuild settings should only be used if host I/O is not a priority.

From the Volumes screen, click **Summary**. The Operational State appears *Critical* if there is only one functioning image available. *Critical* will also appear if there is only one image available while one or more images are rebuilding. The Operational State appears *Degraded* if an image fails and two or more images are functioning. See Mirror Volume States.

All of the data block contents from the selected volume will be copied to the newly created thin image during the rebuild. Checks for zeroes will be performed automatically to prevent unneeded allocations for the image. Thin volume blocks will be allocated for volume data. At the same time, thin blocks will be allocated for volume blocks that were never initialized with zeroes but have non-zero contents, and volume blocks that still have data for deleted files. This can make thin images to be 100% allocated during the rebuild and all benefits of thin provisioning will be eliminated.

Use these type of synchronous images only when the original volume is cleaned up and initialized or there are plans to do a migration from regular to thin volume and execute space reclamation later. See Thin Volume Space Reclamation for more details. Host utilities from the Thin Space Reclamation procedure can be used to initialize the original volume of unused space with zeroes.

4.4 Adding Deduplicated Mirror Image

4.4.1 To add a Deduplicated Mirror Image:

1 From the *Volumes->Configure Volumes* screen, click **Add Image->Deduplicated Image**.

2 From the *Volume Select* drop down menu, select the correct volume. The screen can be used to mirror Snap Enabled Live Volume, but not Snapshots or Thin Volumes.

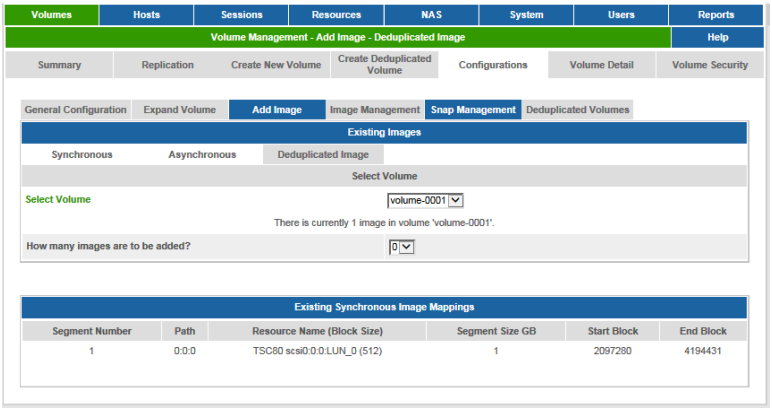


Figure 4-7 Volume Management, Add Deduplicated Image

3 From the drop down menu, select the number of deduplicated images you wish to add. You can select any number less than or equal to the maximum of allowed deduplicated images per Mirror Volume. See Storage Concentrator Configuration Limits.

The *Image Name* field displays an automatically generated name for the Mirror Image consisting of the name of the original volume and an image number. Available deduplicated pools are listed in the *Select Pool* table. Images that were previously generated are listed in the *Existing Image Mappings* table.

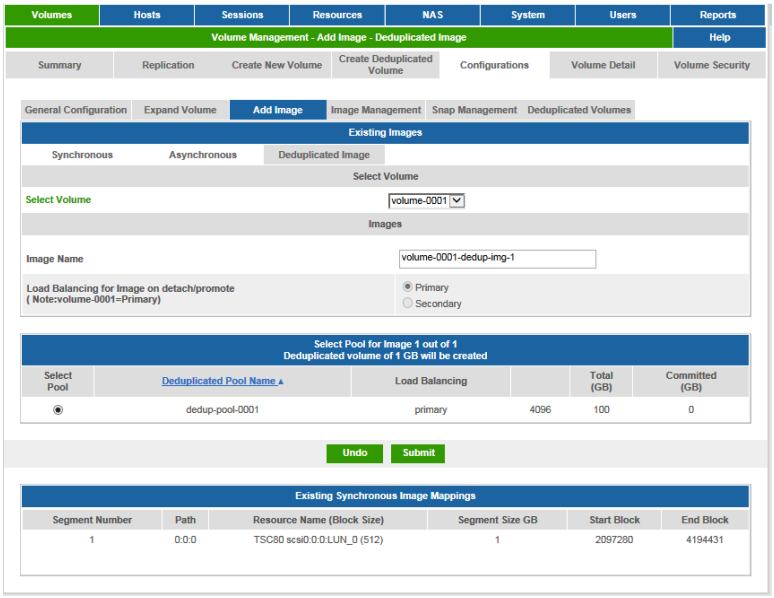


Figure 4-8 Volume Management, Existing Image mapping table

4 Select the Deduplicated Pool as a space where the Deduplicated Image has to be allocated.

5 Click **Submit** to build the Deduplicated Mirror Image.

6 The image will automatically start building at the low priority level. You can change the Rebuilding Priority level from low to medium or high in the General Settings, but the system will use more operational resources to rebuild the image(s) at the higher priority setting.

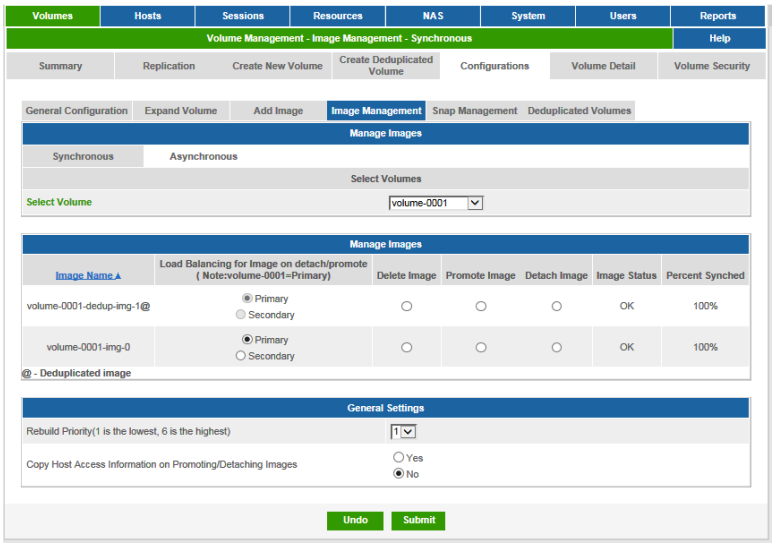


Figure 4-9 Volume Management, Existing Synchronous Image

Changing the Rebuild Priority Setting from low to medium or high will consume more internal *Storage Concentrator* resources and should not be changed without considering the overhead on system resources and system performance. Medium and High rebuild settings should only be used if host I/O is not a priority.

From the Volumes screen, click **Summary**. The Operational State appears *Critical* if there is only one functioning image available. *Critical* will also appear if there is only one image available while one or more images are rebuilding. The Operational State appears *Degraded* if an image fails and two or more images are functioning. See Mirror Volume States.

All of the data block contents from the selected volume will be copied to the newly created deduplicated image during the rebuild. Checks for zeroes will be performed automatically to prevent unneeded allocations for the image. Deduplicated pool blocks will be allocated for volume data that are not present in the pool already. During rebuild, deduplicated blocks may be allocated for volume blocks that were never initialized with zeroes but have non-zero contents and volume blocks that still have data for deleted files. This can make deduplicated images to be 100% allocated during the rebuild and all benefits of deduplicated provisioning will be eliminated.

Use these types of synchronous images only when the original volume is cleaned up and initialized or there are plans to do a migration from regular to deduplicated volume and execute space reclamation later. See Thin Volume Space Reclamation for more details. Host utilities from the Thin Space Reclamation procedure can be used to initialize the original volume of unused space with zeroes.

4.5 Managing Mirror Images and Mirror Volumes

4.5.1 Mirror Image States

Mirror Image State	Description
OK	Mirror image is synchronized with the mirror volume and I/Os are being completed successfully on mirror image.
NOT SYNCED	Mirror image is not synchronized with mirror volume (a rebuild operation is required). No I/Os are occurring on the mirror image.
REBUILDING	Mirror image is being synchronized with mirror volume. I/Os are completing successfully on the mirror image.
DETACHED	Mirror image has been detached from the mirror volume and is accessible as a stand-alone volume. No mirror volume I/Os are occurring on the mirror image.
CRITICAL I/O FAILURE	There has been an I/O failure to the last good image. The I/O failure is reported to the host. The host may retry and subsequent I/Os may succeed, but the image stays in Critical I/O Failure status until there is another image that is rebuilt. Then the Mirror image status goes to either OK or Failed.
FAILED	Mirror image has failed due to an I/O error.
OFFLINE	Mirror image is not available on the mirror volume (user has forced mirror image offline), volume has failed to provision mirror image, or the mirror image status cannot be determined).

4.5.2 Mirror Volume States

Mirror Volume States	Description
OK	All mirror images on the mirror volume are OK (does not consider DETACHED mirror images).
DEGRADED	Two or more mirror images on the mirror volume are OK and at least one mirror image is REBUILDING, OFFLINE, FAILED, or NOT SYNCED.
CRITICAL	Only one mirror image on the mirror volume remains OK and all other images are REBUILDING, OFFLINE, FAILED, NOT SYNCED, or DETACHED.
FAILED	All mirror images on the mirror volume are REBUILDING, OFFLINE, FAILED, NOT SYNCED, or DETACHED. Mirror volume remains accessible to hosts.
OFFLINE	Volume is no longer accessible (user has forced volume offline, volume has failed to provision, or status cannot be determined of mirror volume).

4.5.3 Delete Mirror Images

This function allows the user to remove a previously defined mirror image from a mirrored volume and free the resources that previously stored the image. The mirror image will not be accessible to a host through the mirrored volume after it is removed.

4.5.3.1 To delete an image:

1. From the *Volumes* screen, click **Image Management**.
2. From the Delete Image column of the Images table, select the image that you want to delete.
3. Click **Submit** to delete the Mirror Image.

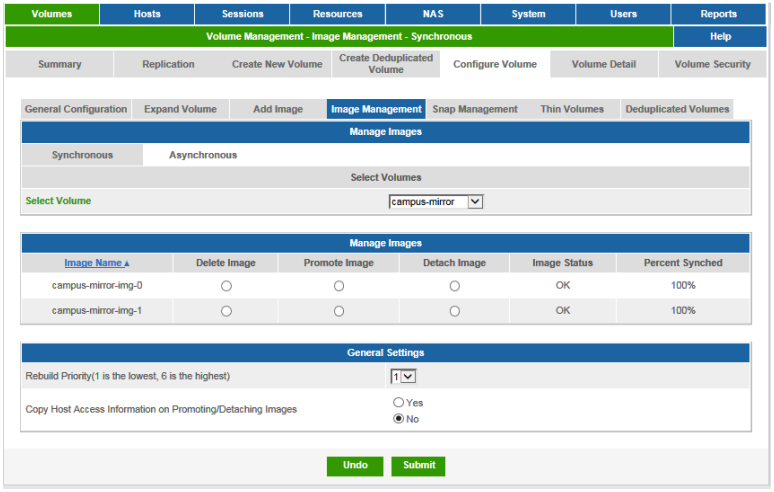


Figure 4-10 Deleting a Campus Mirror Image

When deleting a campus mirror image, it is critical that you delete the image from the Local *Storage Concentrator* before deleting the volume that supports the image from the Campus Remote *Storage Concentrator*.

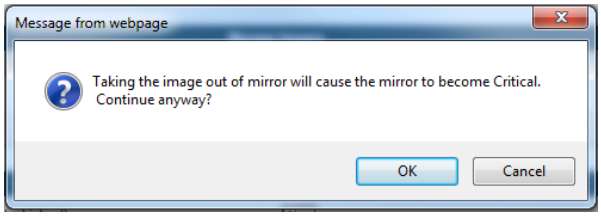
If you delete the volume that contains a campus mirror image from a Campus Remote *Storage Concentrator* BEFORE deleting the image on the Local *Storage Concentrator*, it will cause a FailOver in clustered configurations.

4.5.4 Promote Mirror Images

This function allows the user to promote a previously defined mirror image to a stand-alone volume (used to create a long-lived stand-alone volume). This volume does not retain any mirrored set information. The volume will remain accessible to the hosts but will no longer be mirrored.

To promote an image:

1. From the **Volumes** screen, click **Image Management**.
2. From the **Promote Image** column of the Images table, select the image that you want to promote.
3. To copy the security settings for the mirror volume to the promoted mirror image, select the **Yes/No** option for Copying the Security Information, including access control lists and CHAP secrets. The default setting is **Yes.**
4. Click **Submit** to promote the Mirror Image.
5. If you have only two mirror images and promoting the image will cause there to be only a single mirror image, the system will generate the following warning message:



Click **OK** to continue or **Cancel**.

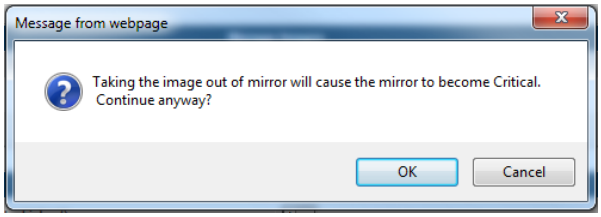
Simultaneously selecting both promote and detach image is not permitted. To reset selections, use UNDO at any time.

4.5.5 Detach Mirror Images

This function allows the user to detach a mirror image from a mirrored volume and make it available as a stand-alone volume. This image retains the mirror volume information to allow for reattachment to the mirrored volume at a later time. The detached mirror image will remain accessible to a host until it is reattached.

To detach a mirror image:

1. From the *Volumes* screen, click **Image Management**.
2. From the **Detach Image** column of the Images table, select the image that you want to detach.
3. To copy the security settings for the mirror volume to the promoted mirror image, select the **Yes/No** option for Copying the Security Information, including access control lists and CHAP secrets. The default setting is **Yes.**
4. Click **Submit** to detach the Mirror Image.
5. If you have only two mirror images and detaching the image will cause there to be only a single mirror image, the system will generate the following warning message:



This function allows the user to reattach a previously detached mirror image. The mirror image is reattached to the original mirror volume. Reattaching a mirror image removes it as a stand-alone image volume and initiates the rebuild operation on the reattached image. The rebuild operation will recreate the entire mirror image at the block level from scratch.

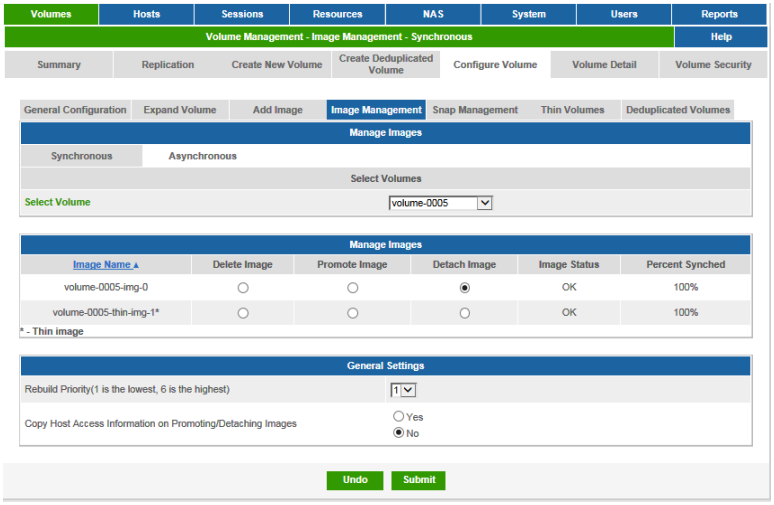


Figure 4-11 Detaching a Mirror Image

4.5.6 Reattaching a Detached Image

To reattach an image, click on the radio button of the image below **Reattach image** column, and then click on **Submit**.

Security settings cannot be retained on a Reattach operation. The security settings will be reset to those of the mirror volume upon rebuild.

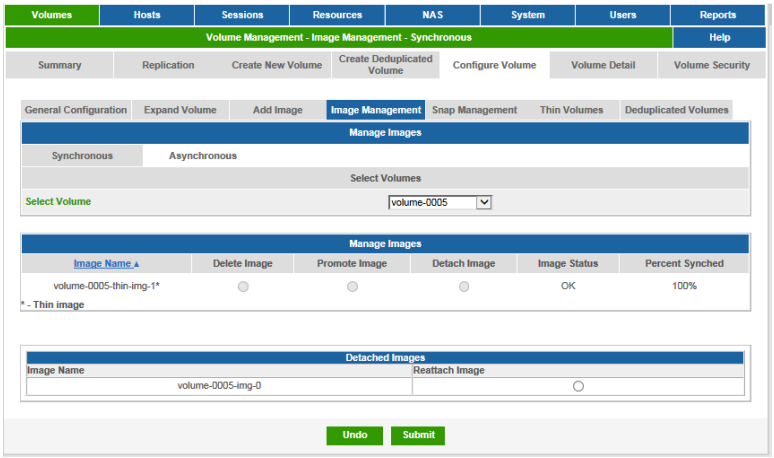


Figure 4-12 Reattaching a Mirror Image

4.5.7 Reverting Mirror Volumes to Spanned Volumes

Reverting a Mirror volume to a Span volume converts the mirror to a regular spanned volume. This eliminates the critical warnings from having a mirror volume with only one good image. Users may want to Revert Mirrors to reduce overhead and increase system performance. Once a mirror volume has been reverted to a spanned volume, it can become a mirror volume by adding images to it.

Mirror volumes can only be converted to spanned volumes if there is only one image remaining. The User Interface will not display the Revert to Span option until all other mirror images are **promoted or deleted**. Detaching the image will not create the conditions required to Revert a Mirror to Span.

To Revert a Mirror to a Span:

1. Promote or delete all images until a single image remains
2. **Select** the Correct Volume from the Pull down menu
3. Click on the **Revert** button under Revert the Mirror to a Span
4. After reverting the volume, the screen will say that there are no images to display under the Images menu and the volume will appear on the volume detail screen as a span volume

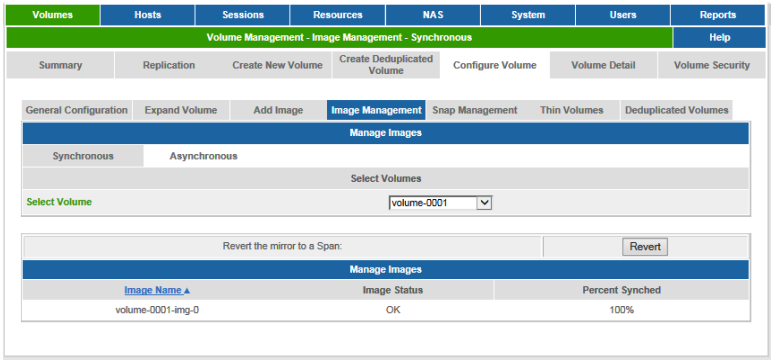


Figure 4-13 Image Management Screen with Revert Mirror to a Span option

4.5.8 Managing General Settings

The features in the General Settings table allow you to change the rate at which an image rebuilds or to retain or discard the Security Settings upon detaching or promoting an image.

To Change Rebuild general settings:

1 From the *Volumes* screen, click **Image Management**.

2 Select one of the following settings:

To change the speed of the rebuild, select a priority from the menu in the "*Rebuild Priority"* row. Changing the rebuild priority while an image is rebuilding will restart the rebuild process from the beginning.

To cancel the rebuild, select the radio button in the *Stop Rebuild* row.

3 Click **Submit** to initiate the setting.

Changing the Rebuild Priority Setting from low to medium or high will consume more internal Storage Concentrator resources and should not be changed without considering the overhead on storage and system resources and system performance.

If you have multiple rebuild operations running simultaneously, system performance may be significantly degraded.

4.5.9 Copying Mirror Security Settings on Promoting or Detaching Images

The security settings, including access control and CHAP secrets, are the same for each mirror image in a mirror volume. When detaching or promoting images, the administrator has the option of retaining the security settings for these images.

1 Select **Yes** to keep the settings of the mirror volume.

2 Select **No** to discard the security settings of the mirror volume. This is the default.

3 Click **Submit** to Change the Security Setting

Security settings cannot be retained on a Reattach operation. The security settings will be reset to those of the mirror volume upon rebuild.

Detached and promoted images are retaining volume encryption passwords.

4.5.10 Expanding a Mirror Volume

To increase the size of a mirror volume after it is created, **use the Expand Volume** function. You must expand all images in a mirror volume symmetrically and simultaneously. See "Expanding a Volume". To expand a Campus Mirror Image, See "Expanding a Campus Mirror Image".

4.5.11 Setting Up Campus Mirror Sites with Mirroring

The ability to mirror images to another *Storage Concentrator* at a campus location provides a robust business continuance option, and supports applications such as content distribution, data migration, data mining, etc.

Before setting up a campus mirror location, you will first need to set up the Campus Remote *Storage Concentrator* and its associated Resources and Volumes. For instructions on setting up a *Storage Concentrator*, see the *Storage Concentrator Setup Guide* and *Storage Concentrator User's Guide* "Using the Administrative Interface", including "Resources" and "Volumes and Security".

First you must set up a Target Portal for the local *Storage Concentrator* to be used to access data port on *Campus Remote Storage Concentrator*. For more information on Target Portals, see "Target Portals".

4.5.12 Setting Up a Target Portal

1 Navigate to System>Target Portals>Add New Target Portal. The Add New Target Portal screen appears

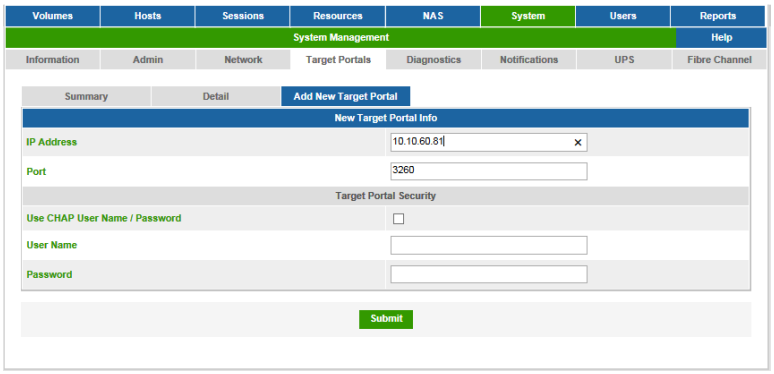


Figure 4-14 System Management, Adding Target Portal

2 Enter the IP Address of the *Storage Concentrator* to be used as a Campus Remote *Storage Concentrator*

3 The default setting for Port is 3260, the iSCSI listening port

4 Enter the security settings for the Target Portal, including specifying a CHAP secret if desired

5 Click on **Submit**.

Next the volume of appropriate size has to be created on the Remote *Storage Concentrator*. For more information, see "Volumes and Security". You must grant local *Storage Concentrator* access to the volume on Campus Remote *Storage Concentrator*. To do this, navigate to the Volumes/Volume Security/Access Page and set read/write access to the local *Storage Concentrator*. For more information, see "Volume Security".

To manage the resources behind the Target Portal, navigate to the Resources Page on local *Storage Concentrator*. Click on "Discover" button. The discovery may take a few minutes.

Storage resources associated with the Remote *Storage Concentrator* linked by the Target Portal will show up on the resources page. However, to make this storage accessible for mirror images, configure the use type as SF Managed.

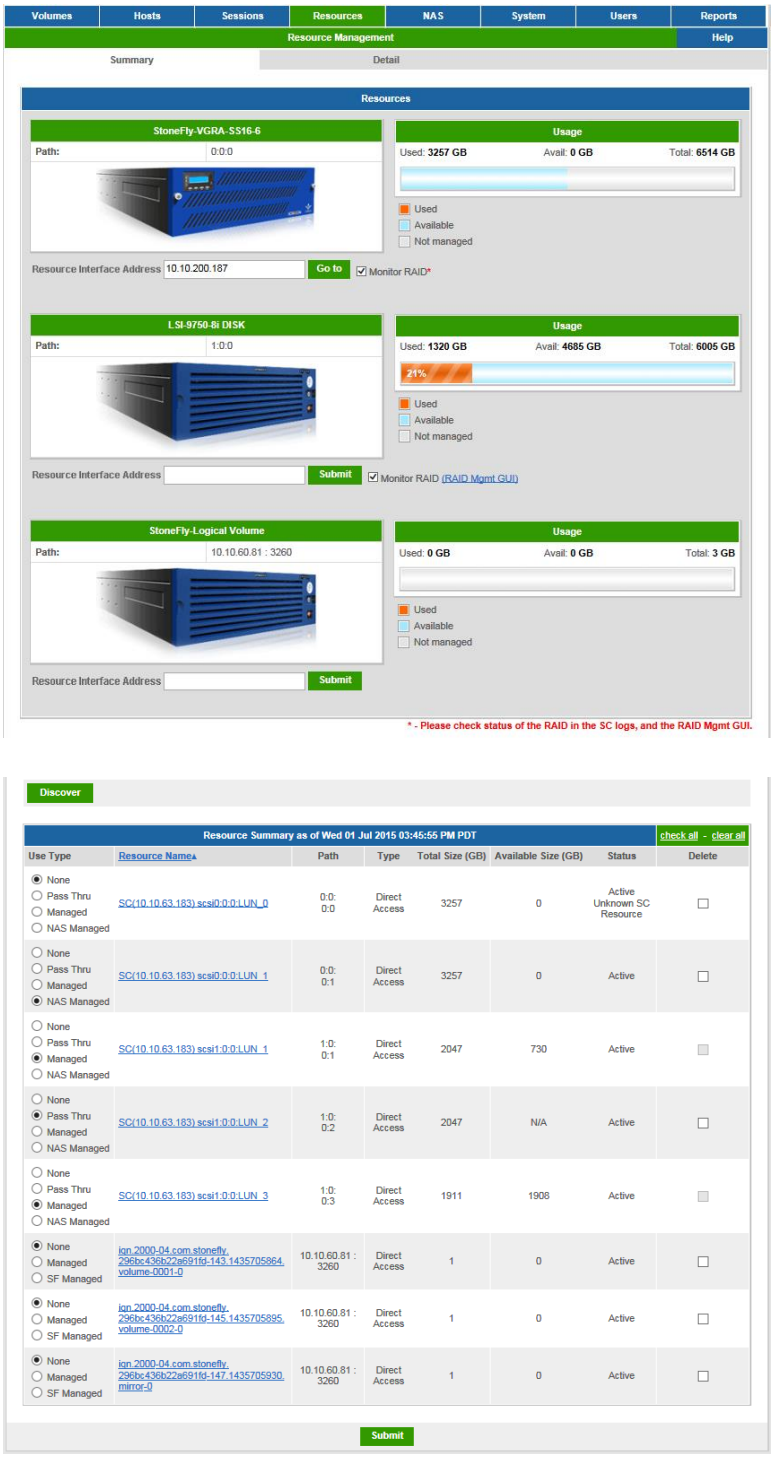


Figure 4-15 System Management, Target portal Resource Table

The Target Portal resources managed by a *Storage Concentrator* will show up with a *Storage Concentrator* icon in the Resources Summary section and shows up in the Resource Summary table with a Path using the IP Address of the Target Portal.

Campus Mirrors use volumes from other Storage Concentrators as resources. The configuration of the remote volume is important. The remote volume can be assigned to any SC of a clustered pair of remote SC's. A Pass Thru volume cannot be used as a campus resource. In addition, StoneFly does not recommended setting up a campus mirror volume as read-only.

There can be no more than 40 total mirrored volumes. The system will generate an error message if attempting to mirror another volume when the maximum number has been reached.

To make the Storage accessible for Campus Mirrors:

1. Select the Radio Button for **SF Managed**
2. Click **Submit**. The available storage will turn from yellow to Green as shown on the resources page to show that it is now available for use with Campus mirrors.

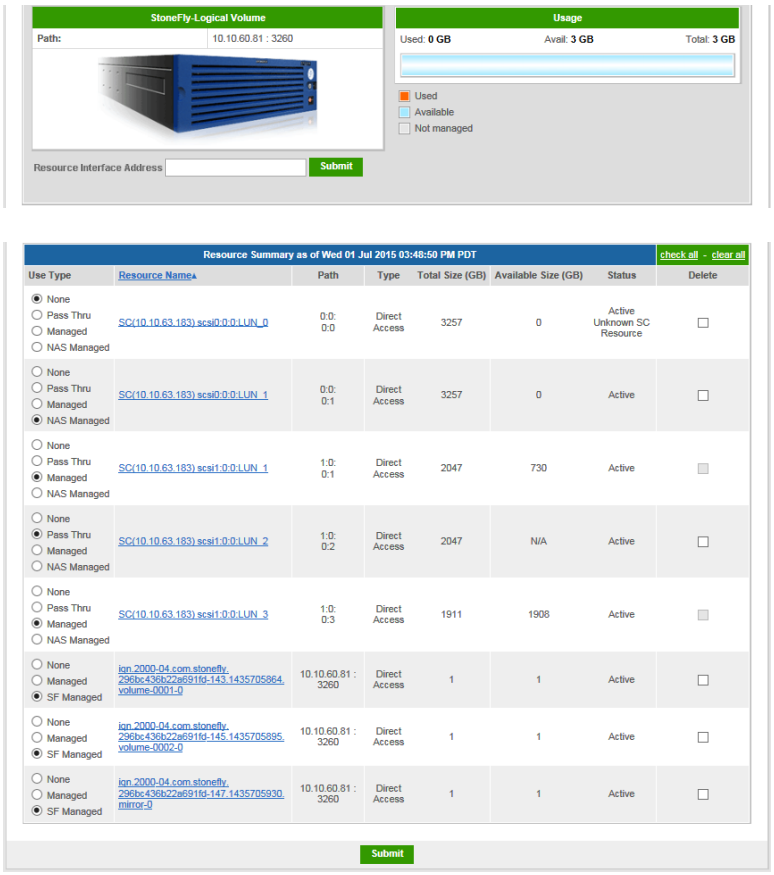


Figure 4-16 Using Campus Remote Storage for Campus Mirrors

Next, navigate to the Volume Page. Select add mirror images according to the instructions on "Adding Mirror Images". Select number of images to add. Follow instructions on screen.

Campus Remote *Storage Concentrator* resources can be identified by the IP address in the Path and by the Resource Name.

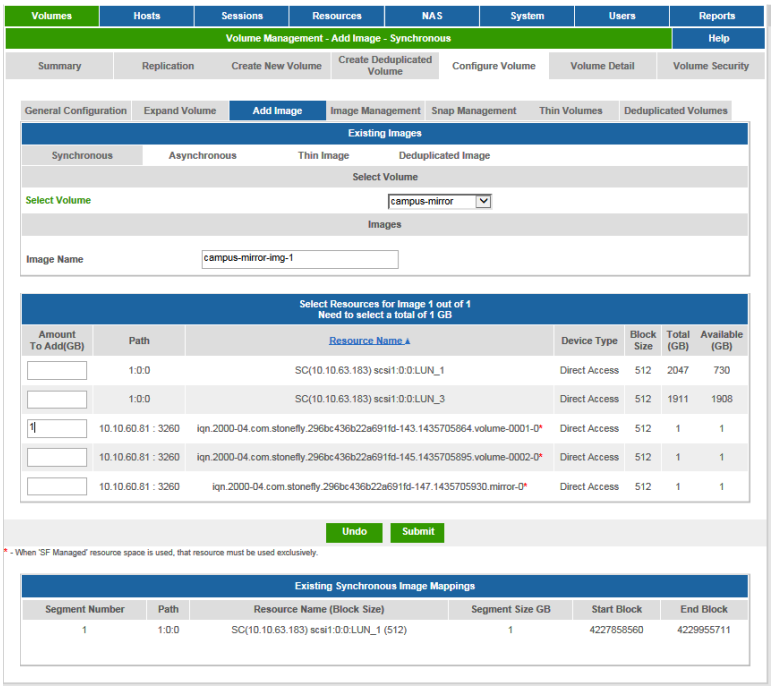


Figure 4-17 Campus Remote Resources

Select amount of storage to add from the appropriate remote resource and click button "Submit". Duplication from local image to the campus image will start immediately after the last requested image is added. User can change rebuild priority any time.

4.6 SFSANCLI & VSS

Many host-based applications, such as backup, can operate in an automated fashion using batch files. There are two methods to accomplish automation in backups. Using VSS and using SFSANCLI. Using VSS is the preferred method.

StoneFly has created a command line interface for common functions required by backup applications.

SFSANCLI is a command line interface for StoneFusion that works with Windows 2000 and 2003. Functions supporting Mirroring include getting the status of a mirror image, detaching an image and reattaching an image.

Quiescing volumes must take place in the application software. Backup applications, have the ability to run scripts and batch files for applications such as SQL and Microsoft Exchange that quiesce a host volume. Regardless of the method used, quiescing volumes is a critical component of any zero backup window implementation. StoneFly provides a Command Line Interface called SFSANCLI.exe to help in developing scripts or batch files for automatic backup of Snapshots.

SFSANCLI was created before Microsoft VSS was invented. StoneFly provides a VSS agent that can be installed on windows systems up to windows 2008. Use of VSS along with the StoneFly snapshot is now the preferred method for use with backup applications to enable a clean backup.

Additional information on what VSS is can be found by searching Microsoft webpage for VSS. StoneFly VSS agent (Hardware provider) or SFSANCLI can be obtained by contacting your StoneFly sales representative.

4.7 Disaster Recovery Scenarios

This section covers a few sample scenarios for recovering from a disaster using a campus mirror site. There are many other possible ways to recover but the following provides some basic guidelines.

To start, ***Figure 4-18*** is a simple IP Storage with a campus mirror site RAID array.

Figure 4-18 Simple Campus Mirroring Configuration

In this scenario, Host X has access to the mirror volume (MV1), comprised of one local image behind SCA and one campus image behind SCB There is a target portal from SCA to SCB. SCB has a managed volume that is SF managed by SCA that stores the campus mirror image of MV1.

Recovery Scenario 1: Temporary Access to Campus Site

If the local site is completely destroyed (fire, flood, etc.), setting up a new host at another site on the network (Host Y) will provide users access to their data using the image of MV1 at the Campus Site as shown in the following illustration.

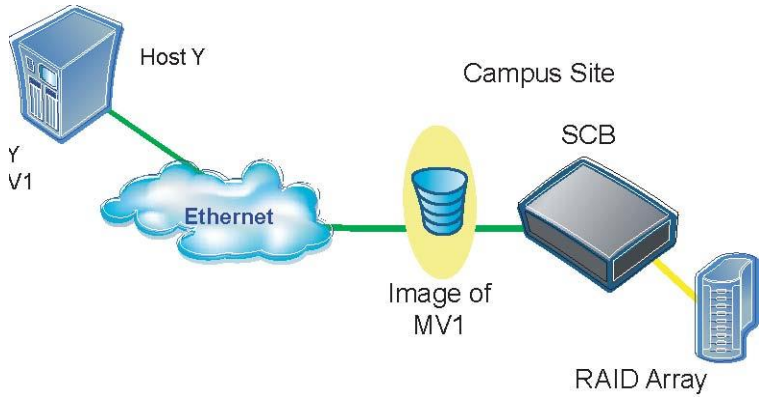


Figure 4-19 Using SCB Provide Host Y with ACLs to MV1

To provide Host Y access to the Image of MV1, the administrator provides access control to Host Y using the volume security settings on SCB. This allows applications to temporarily access the data from the mirror volume.

Restore Scenario 2: Restore Original Campus Mirror Configuration

To restore the original Local/Campus mirror configuration to pre-disaster status requires the following steps.

The following procedure assumes that all applications have been shut down and there is currently no data I/Os to the volume.

Section One: Restore the Data at the New Local Site

1 Install new equipment at local site. The configuration will appear as shown in Figure 4-20:

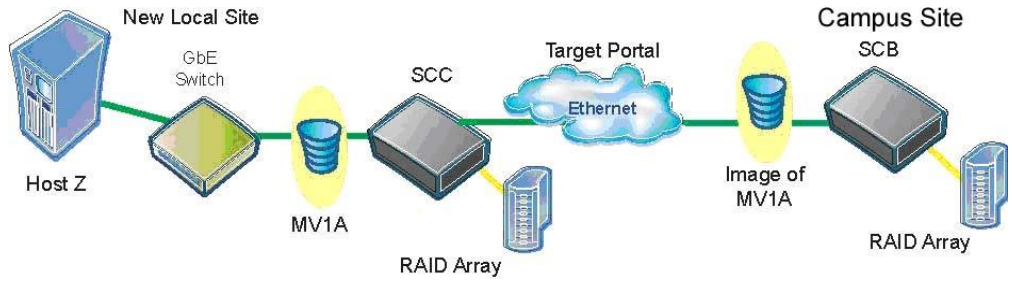


Figure 4-20 Campus Mirroring Configuration Example

2 Create a managed span volume on SCC at the new local site that is the same size as the Image of MV1. This volume initially will be used as a remote image for SCB.

3 Add a Target Portal on SCB to SCC. On SCB, discover the storage resources and set as SF managed the volume created in step 2 on SCC.

4 Create a campus image of MV1 on SCC called MV1A. Allow this image to fully rebuild.

There is now a copy of the data on the new local site. There are other ways to recreate the data on the new local site. One would be to physically move the RAID Array from behind SCB to behind SCC copy the data and then move it back.

Section Two: Restores the Original Campus Mirror

1 Add an ACL for Host Z to the Span Volume MV1A on SCC. Host Z can now service end users on the new local site.

2 Delete (or promote) MV1 on Campus Site.

3 Create a campus image of MV1A on SCB.

4 Add a target Portal from SCC to SCB. Add an ACL for SCC on SCB

5 From SCC, discover the volume on SCB. Add an image of MV1A to campus site.

The Local/Campus mirror configuration is now back to its pre-disaster state.

This page is intentionally left blank.

Chapter 5

Asynchronous Mirroring

5.1 Introduction to Asynchronous Mirroring

Asynchronous Mirroring can be used to replicate data between two or more geographically separate sites. For the purposes of this document, the source data location will be referred to as the local, or production, site and the other as the target or remote site. The sites are represented by two different Storage Concentrator systems, one at each site, regardless of the distance between them. Asynchronous mirroring is associated with a desire to have a level of data recovery in a remote location in the face of a natural or un-natural disaster. Recently, more companies have been required to keep data at certain distances (in some cases hundreds of miles) away from the production site to insure protection of the data in the event a disaster at the local site. The StoneFly method for providing asynchronous mirrors includes the ability for the local Storage Concentrator to identify changes in a volume and periodically move those changes to a remote site. This process is done by coupling StoneFly Mirroring and Snapshot technology. The StoneFly Snapshot technology tracks the changes in the local (production) volume while the Refection technology maintains the movement of the data to the mirror image at the remote site.

The Storage Concentrator uses two methods to identify the changes which must be sent to the remote site: measuring a period of inactivity, or, using a predefined data replication schedule. The task is the same regardless of the amount of change. Using one of the two trigger methods, the local, or production Storage Concentrator triggers a replication event that sends the changes to the remote site. In many cases the communications link between the two sites described above is a slower link than would be used to process that data at the local site. The initial movement of the data to the remote site should be planned in advance of the deployment of asynchronous replication. For instance, the link's throughput capability is critical and must be determined in advance to be able handle the flow of replicated data to the remote site. The desired level of synchronization between the two sites is partly expressed in the size and shape of the communications link and the space allocated to hold the changes before they are sent to the remote site. A slow link with a high rate of change in the data will dictate a gap (perhaps large) in the synchronization between the sites. A high speed link and smaller rates of change in the data can keep the two sites closely synchronized.

Consideration should be taken to understand your organization's Recovery Point Objective and Recovery Time Objective (RPO and RTO respectively). If you require a highly-synchronized RPO and short RTO, bandwidth and replication scheduling should be planned carefully. See also Appendix 3: Preparing for StoneFly Asynchronous Mirrors.

It is also good to think about how the replicated data will be used in a Recovery Plan. The communications discussion above can also affect the way the data is used at the remote site. Bringing data back to the local IT site through a slow link may be prohibitive in the desire to get an organization up and running again. Even a single large file needed at the local IT site could take days of transmission time over something like a T1 link.

5.2 Operational Overview

StoneFly Asynchronous Replication is implemented using StoneFly Mirroring and Snapshot technology. The asynchronous replication processes are coupled with snapshots to track all the changed blocks in the volume. **It is important to review the sections of this User Guide that describe Campus Mirrors and Snapshots before implementing Asynchronous Mirror features on any volume**.

A simple view of the way StoneFly Asynchronous Mirroring work is to describe the two volumes that must be synchronized during the mirroring or replication process:

1. The Remote(Replication) Volume
2. The Production Volume

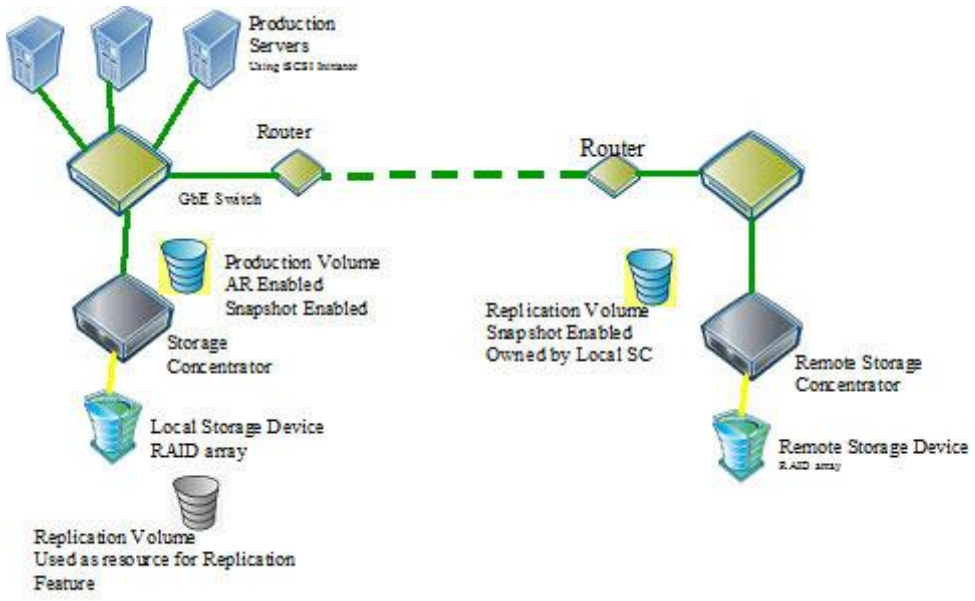


Figure 5-1 An example of Asynchronous Mirroring Configuration

5.2.1 Overview of the Remote Volume

The basis of the replication process is the existence of an additional copy of the data at one or more remote sites. The first step toward Asynchronous Mirroring is to provide a volume at the remote site. The remote volume must be at least the size of the volume being mirrored from the production site. This remote volume is expected to contain the exact same data as the volume at the production site. Extra space on remote volume let user execute expansion for production volume later if the expansion is needed. Data is being added to the remote volume through time at an unknown rate and may be in an unstable state. Any disruption to the communications link between the production

site and the remote site may leave IO's in a set of changes on the link and not completed. The use of the volume in a disaster recovery situation may require the use of tools that will repair any problems in the file system or application data.

Alternately, the remote volume may be enabled for snapshots. This is highly recommended to improve the accuracy and stability of the replicated data at the remote site. Each time a set of changes is sent to the remote volume, a snapshot is requested by the Storage Concentrator at the production site. These snapshots become recovery points at the remote site. The amount of Snapspace allocated at the remote site determines how many recovery points are possible. Allocating a larger Snapspace provides more recovery points up to limit defined by the software implementation.

5.2.2 Overview of the Production Volume

The Production Volume is the source of all changes in the replication process. Changes are captured through the use of snapshot technology that tracks all changed blocks in a volume. Once the time for a replication has been identified, a differencing process takes place and all changed blocks are transferred to the remote site. A special set of snapshots are established on the production volume to allow for the differencing process to take place. Once the snapshots have been established on the volume a second step is used to connect the volume to one or more of the available remote volumes.

The Storage Concentrator's snapshot technology uses a separately allocated space in the Resource pool for tracking changes. The allocation of this space for the Production Volume is a buffer against problems with communications to the remote site. A larger snapshot allocation provides more space to hold changes while the communications link may be down. The size of the space directly represents the available time to recover any communications problems before the synchronization cannot be done between the two sites. When this space fills up a full re-synchronization of the two sites must be performed. However, the Snapspace may be increased manually as more is learned about the replication environment. See the section on Volume Configuration, subsection "Expanding the Snapspace for a Snap-enabled Live Volume".

5.2.3 Synchronizing the Data in the Volumes

In order to start the replication process on any volume it is necessary to decide on the initial synchronization method. There are three approaches to synchronization provided by the Storage Concentrator:

Option #1: Starting when there are no data on Production and Remote Volumes. This means that either the volume is completely empty; every byte of both

volumes can be overwritten later. There is no need for any additional synchronization of the Remote Volume. The replication process can simply start at the earliest possible time.

Option #2: Sending a copy of the production volume to the remote site and copying it into the Remote Volume. In this scenario the data in the Production Volume is expected to be different than the data in the Remote Volume due to the changes taking place while the data is in transit to the remote location The process starts when the Administrator at the Production Site sets a snapshot on the Production Volume. An exact copy of the volume is made using this snapshot. Then the copy is sent to the remote site. When the replication process is started at the production site, this specific snapshot is used to determine all the changes that were made since the copy was sent to the remote site. The first replication will move the bulk of the changes to the remote site. The next Asynchronous Replication snapshot will track the changes that occurred while the re-synchronization process was going on. There is no need to take the Production Volume offline while the re-synchronization process is going on.

Option #3: When the Storage Concentrator equipment is present at the Production Site for initialization, Option #3 can move the initial volume copy to the remote equipment before the equipment is sent to the Remote Site. Once the initialization operations are complete the Remote Storage Concentrator is simply shutdown and moved to the Remote Site. The Snapspace must be large enough to cover the time for the move. After the remote equipment is installed and connected to the equipment in the Production Site, the Asynchronous Mirroring process should automatically restart. This is similar to the situation in which the link between the two sites is disrupted. The replications begin again when the link is restored.

Alternate Option #3: Alternate use: If the remote equipment is installed and working at the Remote Site, the initialization process may be accomplished by copying the entire volume to the remote site through the network link. Depending on the size of the mirrored volume and the amount of available bandwidth, this may be the best approach to start your replication. If this option is selected, the replication process begins immediately to copy the volume to the Remote Site and to track any changes to the Production Volume while the initial synchronization is being done. The new changes are tracked in the replication snapshots and then moved to the remote site after the original copy is complete. The size of the Snapspace allocated for the volume must be large enough to contain the changes while the copy is made. A rough estimate of the copy time is possible using the link speed and the size of the volume. All blocks in the volume will be copied. The Storage Concentrator has no knowledge of whether any particular block is used or not used. This option does not require any downtime for the Production Volume.

These three options appear in the implementation section several times. Each time the replication process is configured and enabled on a new or existing volume one of these options must be selected.

5.2.4 Keeping the Remote Volume Synchronized

There are several methods for insuring that replication data is passed from the production site to the remote site. The most straightforward methods include:

1. Using the IO Idle Interval setting to recognize a quiet time and triggering a replication event.
2. Setting a schedule of exact times a replication event is desired.

Both of these methods can result in replication data being passed to the remote site. The use of the IO Idle Interval recognizes a period of time during which there are not write IO's for the Production Volume. The purpose of this time interval is to establish points in time when a file system or data base is considered quiesce.

The use of the schedule can also involve the IO Idle Interval but attempts to control the flow of data to the remote site. The scheduler uses a list of times to open up the volume for replication events. The system may be configured to start searching for an IO Idle Interval period at 2:15 AM and again at 2:15 PM. On the next occurrence of the IO Idle Interval after the 2:15 AM time, the system will trigger a replication event and pass all the data that has been changed in the volume since the last replication event.

5.3 Implementing Asynchronous Mirrors

The instructions for successfully implementing Asynchronous Mirrors start with the assignment of Snapshots to the volume. The second step is adding the Asynchronous Mirror(s).

5.3.1 Configuring Asynchronous Mirrors for Newly Created Volumes

Many of the steps for implementing Asynchronous Mirrors on newly created volumes can be applied to any volume. Read these steps carefully before starting replication features on any volume.

To configure a new volume for Asynchronous Mirrors:

1. Define the size and location of each volume to be remotely mirrored.
2. Create the local volumes to hold the production data.
3. It is possible to add a Synchronous Mirror image to the new volume along with the Asynchronous Mirror image. Add the Synchronous Mirror images before adding any Asynchronous Mirror images. See the earlier sections of the User Guide for instructions on how these Mirroring features are applied to any volume. (see also: "To add a Mirror Image")
4. Allow all the Synchronous Mirror images to completely rebuild before proceeding to add Asynchronous Images.
5. Create volumes on the remote site *Storage Concentrator* to prepare for the Asynchronous Mirroring process. These volumes at the remote site must be at least the size of the Production volumes at the Local Site. If the remote site is an SC Cluster the volumes must be assigned to the Primary remote SC.
6. Use the remote site SC's Snapshot Management screens to enable the remote volumes for Snapshots. Do **not** select the **Enable Volume Asynchronous Replication** at the remote site.
7. Use information for "Local iSCSI Data Port" from the Remote System Information Page (see "System Information") to create Target Portal on Local Site to iSCSI Data Port on Remote Site (see "Target Portals"). If the Target Portal is created successfully, the Local Site SC's have to be placed into list of available iSCSI hosts on the Remote Site automatically (see "Hosts").
8. Set the Volume Security of the Remote Site volumes to allow the Local Site SC's access to them (see "Volume Security"). If the Local Site is an SC Failover Cluster the remote volumes must be set to allow access by both of the local SC's.
9. Using the Local Site SC's Resources Summary screen Discover the new Remote Site volumes as Resources. Set them for "SF Managed" use.
10. Enable Snapshots on the volumes for which Asynchronous mirrors will be requested.

11. Be sure to click **Yes** on the "Enable Volume Asynchronous Replication" line. Then click **Enable Replication**. (If you forget to click on the Yes, you can do it anytime later or right away on another snapshot.)

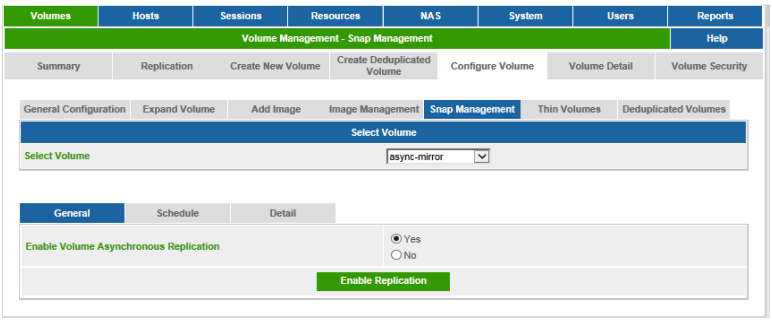


Figure 5-2 Enabling Volumes for Snapshots with Asynchronous Mirrors

12. Request a SnapSpace large enough to hold the changes made in a typical day plus extra space to hold changes while the remote link may be down. This size is balanced against the frequency of replication and the throughput of the communications link.

13. Finally, use the Local Site SC's Add Image –> Asynchronous screen to request one to three Asynchronous Mirrors:

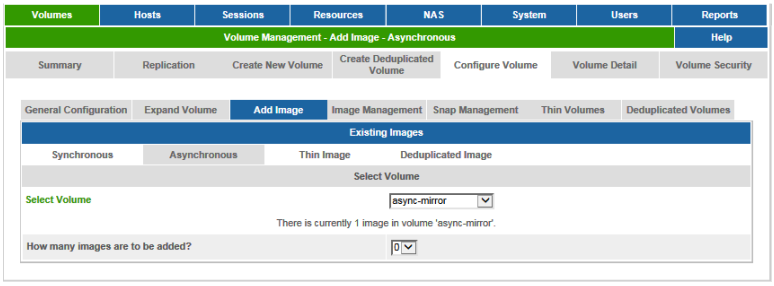


Figure 5-3 Asynchronous Mirroring Screen

14. Activate the replication process on one of the resources from the Remote Site SC as shown below. If you have requested more than one Asynchronous Mirror the screen will be repeated to allow initialization of all mirror images.

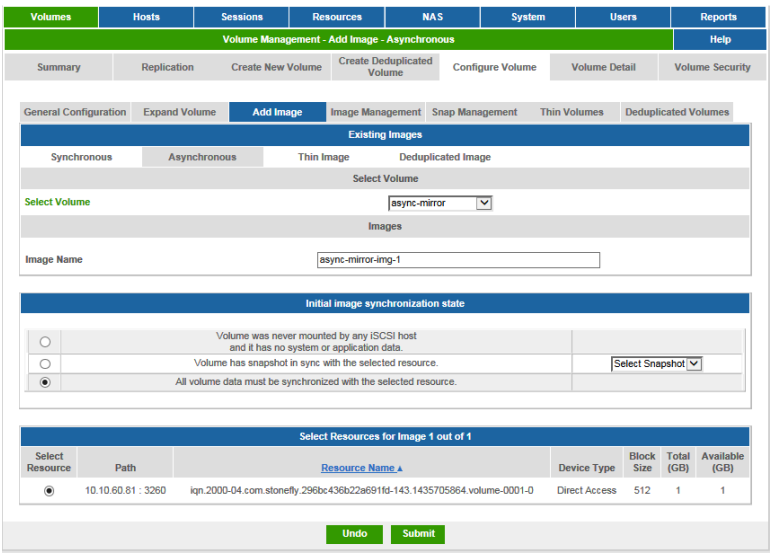


Figure 5-4 Add Asynchronous Mirror Image to volume "async-mirror" on "local" Storage Concentrator. Volume "volume-0001" from the "remote" System is selected.

Option #1 is for volumes that are built up through time. The host has not formatted the volume yet so it is empty. Select this Option to start replicating on the next change.

Option #2 is for volumes that are built up from a known data set. If the volume will be pre-loaded with data for use by the applications Option #2 allows a snapshot to signify the data set is loaded in the remote site volume. The remote volume must be an exact image, block for block of the selected snapshot. Usually this means it is created as a local mirror image and then moved to the Remote Site. The snapshot must be taken at the point of separation.

Option #3 is for most volumes that have been in use for a while. If the remote site equipment is present at the Local Site the volume may be established on the remote equipment before the equipment is moved to the remote site. Remember to suspend Asynchronous replications during the move. Resume operations when the remote site is working.

Alternate Option #3 Alternate use: If the volume cannot be mirrored at the local site prior to setting up the remote site's Storage Concentrator(s) use this setting to copy the entire volume to the remote site through the communications link. This is hard on the communications link but can be done while the local volume is in normal operations. Achieving synchronization may take a long time.

5.3.2 Configuring Asynchronous Replication for Existing Volumes

1. Identify the volume(s) to be replicated. Note the size of the Volume.
2. Decide if the volume should be expanded now to accommodate future needs. If yes, do it now. Replication must be stopped to allow Volume Expansion.
3. At each DR Site: Allocate a volume the same or larger size as the production Volume.
4. At each DR Site: Enable the new Volume for Snapshots. Select a Snapspace size large enough to hold several periodic updates.
5. At the local Site: Create Target Portals to access iSCSI Data Port at each DR Site (see "Target Portals").
6. At each DR Site: Set Volume Security to allow access for the SC(s) at the Local Site.
7. At the Local Site: Discover the new resource from the DR Site. Set the use as SF Managed.
8. At the Local Site: Enable the existing volume for Snapshots. Use the "Enable for Asynchronous Replication" button. Select a Snapspace large enough for 20% to 30% of the original volume size.
9. Select one of the two following algorithms:
   a. Option #2 is for volumes that are built up from a known data set. If the volume is loaded with data for use by the applications Option #2 allows a snapshot to signify the data set as it is being loaded in the DR Site volume. The DR Site volume must be an exact image, block for block. Usually this means it is created as a local mirror image and then moved to the DR Site. The snapshot must be taken at the point of separation.
   b. Option #3 is for volumes that have been in use for a while and cannot be mirrored at the local site prior to setting up the DR Site SC's. In this case the volume contents are built up by moving the data across the link to the DR Site volume. This is hard on the communications link but can be done while the local volume is in normal operations. Achieving synchronization may take a long time.
10. At the Local Site: Enable the Asynchronous Replication using the Add Image -> Asynchronous screen. Refer to the process as described for New volume above.

5.3.3 Expanding Asynchronous Replication Volumes

To expand the size of a volume that is enabled for Asynchronous Replication see "Expanding a Volume with Asynchronous Mirror Image".

5.4 Managing Asynchronous Replication through the User Interface

5.4.1 The Replication Summary Screen

From the Volumes navigation bar, access the Replication Summary screen to see all Asynchronous Images in one screen.

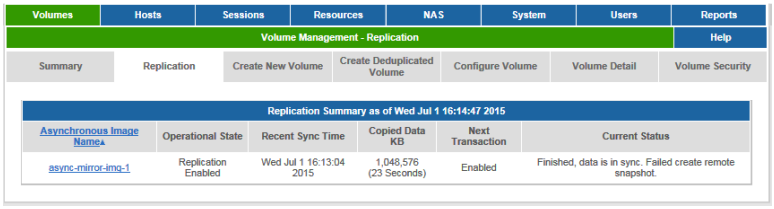


Figure 5-5 Asynchronous Images Screen on "local" Storage Concentrator

The following fields are presented for each Asynchronous Image:

Asynchronous Image name – Each image is listed with the name selected at the time the asynchronous image was created. Note that images start numbering at one. The production volume is numbered "zero". Only the remote images are listed in this table.

Operational State – The status of the images is indicated here. The term "Replication Enabled" indicates a good image.

Recent Sync Time – The most recently completed replication is noted in this field. The time indicates the start time of the replication.

Copied Data KB – The amount of data copied on the most recently completed replication is listed in kilobytes.

Next Transaction – The ability to continue the replication process is shown in this filed. An asynchronous image that is suspended would be indicated in this field. The status "Enabled" indicates that the next replication can begin at its appointed time whether scheduled or free form.

Current Status – The current Status indicates if a replication is in progress at the time the screen was displayed. A status of "Finished, data is in sync" indicates the last replication was completed but the next replication has not yet begun. A status of "Standby" indicates that the first replication has not occurred yet.

To see all the details about a specific image click on the name of the image to display the Image Management screen for that image. See the next section for details on the Image Management screen.

5.4.2 The Asynchronous Image Management Screen

Several new screens provide information and adjustments to the Asynchronous Replication Features. Accessing the Image Management->Asynchronous->General screen provides all the information about how the replication process is being carried out. The screen includes information about the last synchronization point.

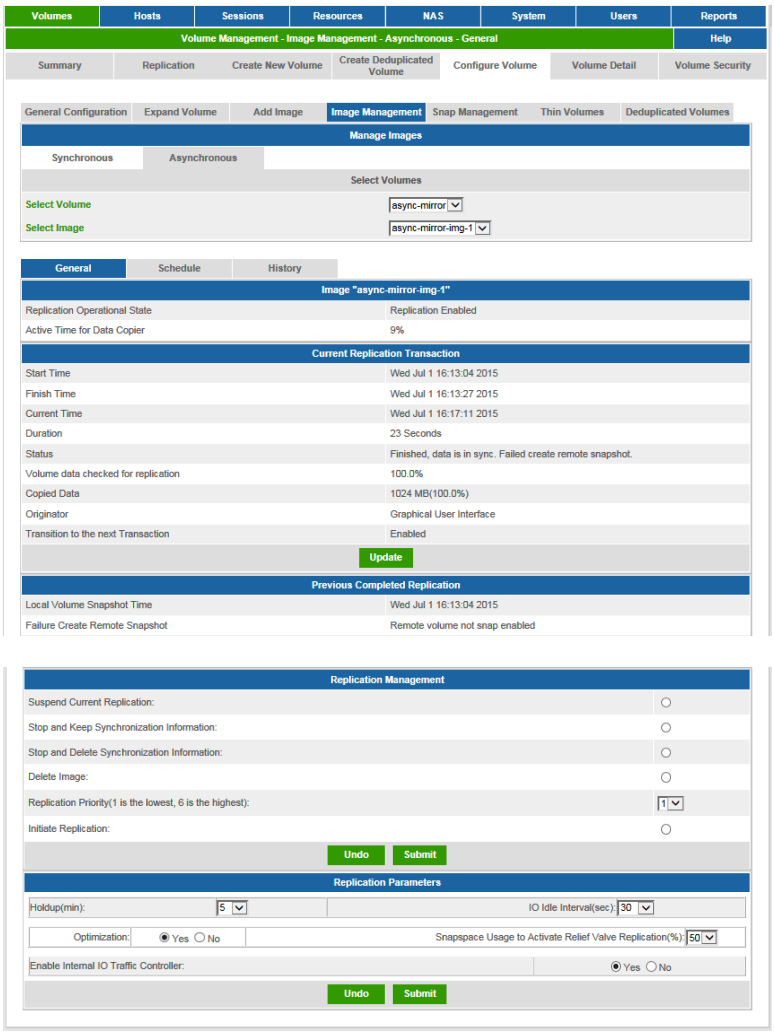


Figure 5-6 Asynchronous Image Management on "local" Storage Concentrator

Select Volume pull-down: Select which volume to display

Select Image (pull-down): Select which of the remote images to display. There may be more than one remote site.

Replication Operational State: Enabled or Disabled

Active Time for Data Copier: This number is reported as a percentage of the time that the last replication took before the next replication began. Determine the number of seconds used to copy the data in the last replication and divide that number by the difference of the time from the start of the last replication to the start of the next replication.

$$Replication\ bandwidth = \frac{Transfer\ time\ (in\ seconds)}{Recent\ start\ time - last\ start\ time}$$

As this number approaches 100% it indicates the Asynchronous Mirroring process is not able to keep up with the rate of change in the volume.

Current Replication Section

Start Time: The time the current or last replication event was recognized.

Finish Time: The time the most recent replication completed. If the replication process is still moving data to the remote site this field will report "Not Finished"

Current Time: The time when this screen was last updated. Update the screen to match the actual time. This data can be compared to the Finish Time to know the period since the last synchronization point.

Duration: The time period in minutes and seconds that it took to move the most recent replication data to the remote site. Compare this data to the amount of data moved to understand the utilization and efficiency of the communications link to the remote site.

Current Status: This data reports on the activities and results of the most recent replication event. The data is represented in two pieces: Completion status and Synchronization status.

Completion Status

Finished: No data is currently being passed to the remote site. The system is awaiting the next replication event.

Not Finished: Replication data is currently being moved to the remote site.

Synchronization Status

Data are in sync: The most recent replication operation was successful. The data at the remote site matches the data at the production site.

Data are out of sync: Some error has occurred that prevented the most recent replication operation from completing successfully.

Failed to create Remote Snapshot: The remote site is not able to create another snapshot for this volume. However, the replication volume continues to be updated at the remote site.

Volume Data Checked for Replication:

Copied Data: (in Megabytes)

Originator: This data indicates the trigger event for the most recent replication. There are several possible triggers:

Internal IO Traffic Controller: This trigger is the normal operation for the Asynchronous Mirroring Features. The **IO Idle Interval** has been detected and started the most recent replication operation.

Relief Valve for Internal Snapspace Controller: This trigger is enabled to move data based on the amount of Snapspace used to prepare for the next replication. The trigger is designed to move data before the next **IO Idle Interval** is detected.

Graphical User Interface: This trigger indicates that the button below, **Initiate Replication** was selected by the Administrator.

Transition to the Next Replication:

Enabled means the next replication activity can take place as scheduled.

Disabled means no more replications will occur

Suspended means the replication process will not continue without Administrator intervention

Current Synchronization Section

Local Volume Snapshot Time: The time the most recent replication operation completed.

Remote Volume Snapshot Name: The name applied to the snapshot at the remote site. Each time a replication operation is successful a request is sent to the remote site to take a snapshot on the remote volume. The requests sent to the remote site include the specific name to be given this snapshot.

Replication Management Section

Delete Image: Selecting this button removes this asynchronous image for the volume. The volume may have more than one asynchronous image. The other images continue to operate after this image is deleted.

Suspend Current Replication: By selecting this button the movement of data to the remote site is stopped until restarted. Changes to the volume are tracked in the volume's local Snapspace.

Stop and Keep Synchronization Information: Select this function to cancel the current replication operation. The previously completed replication to the remote is retained, and is reported above.

Stop and Delete Synchronization Information: Select this function to cancel the current replication operation, and to delete the replication synchronization state. The next replication shall then require a full data transfer of all of the volume's data to the remote.

Replication Priority: Select '1','2','3','4','5' or '6' to control the amount of resources used to complete the replication operations on this image. The resources are allocated as the number of IO tasks invoked to update the remote volume. A LOW ('1') setting has fewer IO tasks than a HIGH ('6') setting.

Initiate Replication: Select this function to start a replication operation immediately. The function does not wait for an IO Idle Interval.

Replication Parameters Section

Holdup: Set the amount of holdup time before the next replication operation can be started. The value is set in minutes with a default of 5 minutes. The setting has effect when previous replication was finished successfully and all data were synchronized between local and remote snapshots.

IO Idle Interval: Set the amount of time the volume must have no modifications before a replication operation is started. The value is set in seconds with a default of 30 seconds. The same time interval is used by Relief Valve to check snapspace usage.

Optimization: When enabled invokes procedure to reduce size of data that have to be copied to the remote volume during replication. The optimization procedure consumes additional processor time on local system, but in general it can get additional performance by reducing amount of data that have to be copied through network. Use the **Submit** button to enable/disable replication IO optimization at any time.

Snapspace Usage to Activate Relief Valve Replication: Set the amount of snapspace usage before pushing the existing changes to the remote volume. The Relief Valve function is used when a volume is too busy to encounter an IO Idle Interval. Consider setting the IO Idle Interval to a shorter period if the only replications come as a result of Relief Valve operations. Use the **Submit** button to adjust these Replication Parameters at any time.

Enable Internal IO Traffic Controller: When the Controller is disabled no attempt is made to find an Idle IO Interval period to trigger the replication operation. In this case replication can be initiated manually by using GUI replication management screen, externally from hosts or can be requested by Relief Valve if snapspace usage reached some predefined limit. If the Controller is enabled replication is initiated each time when specific idle IO interval is detected. The default state for Controller is enabled. User should not change it without providing other tools to trigger replication automatically.

5.4.3 Suspending the Mirroring Process with the Internal Scheduler

This function allows the user to create a list of time intervals during which processing of a replication operations can be put on hold. The list composed of separate weekday and weekend settings. Initially the list is empty with all hour fields set as 'NONE'. In this case the schedule does not prevent replication from running. Minute settings have no effect if hour is set to 'NONE'. The list defines times when replication operation has to be resumed or suspended. When time comes to resume replication the schedule does not trigger start of the new replication because there is no guarantee that the volume data are at quiesced state at the moment defined by the schedule. The previously suspended replication will be resumed only.

Changes at the schedule have effect on replication during the next 10 minutes interval. For each moment the replication state is defined by the schedule using the most recent setting. For example if the last settings for weekday is 'resume at 8PM' and all weekend settings are 'NONE', replication will be enabled at 8PM on Friday and has to stay at this state at least till 12AM Monday. The last weekend setting is in effect from its time on Sunday till time of the first weekday's setting on Monday. If system time is changed, the schedule is checked and replication state is adjusted regarding to the new system time during the next 10 minutes.

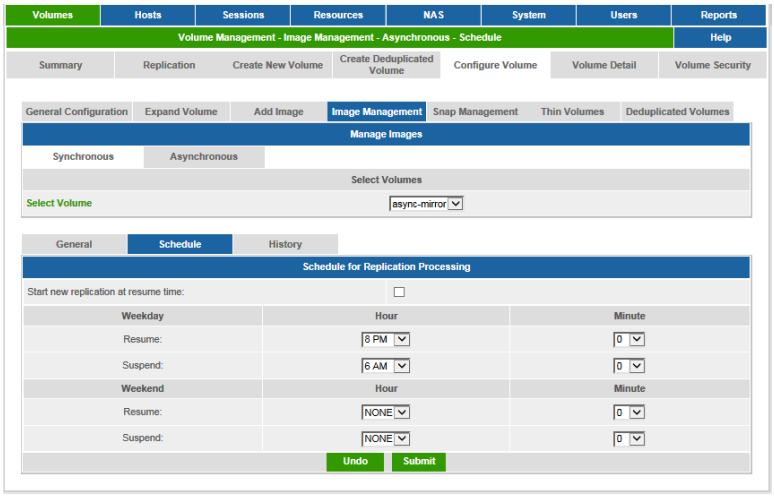


Figure 5-7 Asynchronous Images Schedule Screen on "local" Storage Concentrator

Select Volume (pull down list): The pull down list contains all the volumes that have been enabled for Asynchronous Replication. Use the mouse to highlight the name of a particular volume on the list to display its specific settings.

Start new replication at resume time: The default schedule's behavior is to resume replications that were suspended early. If there is no suspended replication the schedule let other facilities to trigger replication, but it does not start the new replication itself. User can change this default behavior by checking the "Start new replication at resume time" check box. If the check box is checked, the schedule has to start new replication at the "resume" time.

5.4.4 Replication History

This screen allows the user to retrieve all successful replications that are still presented in database. The database can keep no more than specific number of replication records. The oldest records that are beyond this limit are deleted automatically.

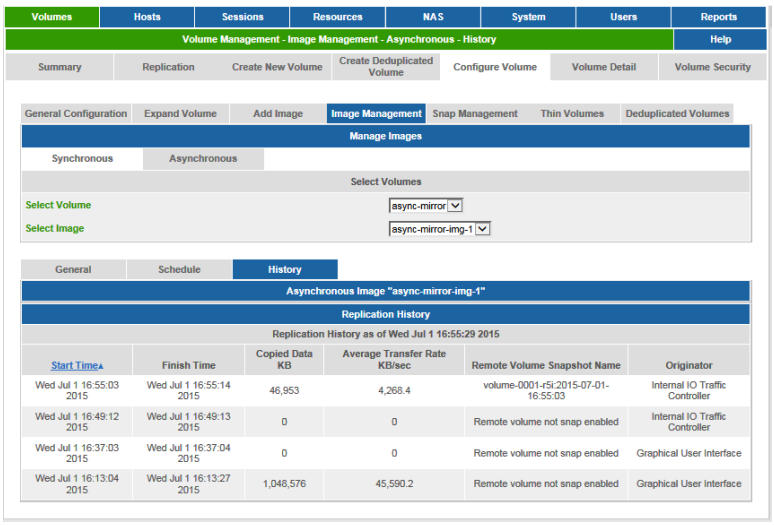


Figure 5-8 Replication History Screen on "local" Storage Concentrator

Select Volume: Select a volume from the drop down volume list. All applicable volumes will be listed.

Select Image: Select an image from the drop down image list. All asynchronous images for the current volume are listed.

Start Time: Reports the time the successful replication operation started.

Finish Time: Reports the time the successful replication operation completed.

Copied Data: Displays the amount of data in KB transferred to the remote site during the replication operation.

Average Transfer Rate: Displays the average amount of data in KB transferred to the remote site per second during the replication operation. The rate usually is less than network transfer rate because time to search for replication data and time to execute replication management commands are taken into account. Two or more asynchronous replications can share the same network bandwidth. In this case the network transfer rate is split between all involved replications.

Remote Volume Snapshot Name: This field reports the name of the snapshot created for completed replication operation.

Originator: Displays the function within the Asynchronous Replication features that caused the replication operation to take place.

5.5 Disaster Recovery with Asynchronous Replication

When Asynchronous Replication features are applied to any volume the copy of the data at the remote site may be used to re-establish servers and applications following an outage or disaster. Each user of these features should carefully consider and plan for this recovery operation. Some guidelines and suggestions for the recovery process are included here. The exact plan for each recovery operation may be different.

Details of the next cases are described below:

Planned Failover and Recovery. User has time to prepare the local system and servers for shutdown. After that all applications has to be supported by the remote system. When it became possible, the local system has to be put on line and all services have to be redirected to use the local system as a primary. Original status of the remote system has to be restored.

Unplanned Failover and Managed Recovery. User does not have time to prepare the local system and services. When failure of the local system is detected, the remote system has to be activated to support applications. When it became possible, the local system has to be put on line and all services have to be redirected to use the local system as a primary. Original status of the remote system has to be restored.

Primary Site Hardware Replacement. This is the case when hardware on the local site has to be replaced before the original status of the local and remote systems can be restored.

Planned Shutdown and Recovery. User has time to prepare the local system and servers for shutdown. The local system and servers are staying down until the system can be put online again. The remote system is not used to support servers.

Unplanned Shutdown and Managed Recovery. User does not have time to prepare the local system and servers for shutdown. The local system and servers are staying down until the system can be put online again. The remote system is not used to support servers.

The real situation can be different. For example, user is making preparation for disaster by using the "**Planned Shutdown and Recovery**" procedure, but has to replace hardware during recovery later. Or the remote system has to support servers for some period of time before the local system can be put on line again.

5.5.1 Planned Failover and Recovery

This is the case when the local system has to be shut down and all hosts that are using targets provided by the local system have to be redirected to use targets from the remote site. After some period of time the local system has to be activated and restored as a primary system. There are 4 different periods in procedure to handle the disaster.

The first one is when the planned failover is executed.

The second is for time when local system is down and the remote system is handling all applications.

The third period is when local system is put online, remote system is still handling all activity, but it is replicating all changes to the local system.

The final period is time when planned failover is executed to restore primary status for the local system.

The local system and remote system can be standalone Storage Concentrator or cluster. For simplicity, we are going to address the remote system as a standalone Storage Concentrator. The scope of the current case includes a single volume on the local system that is replicating to the single remote location and iSCSI hosts that are using the selected volume. All other volumes that are provisioned on system and all other hosts that are accessing these volumes are out of the consideration. Each individual volume can be handled in the same way.

5.5.1.1 Local System Shutdown

STOP APPLICATIONS.

DISCONNECT ALL HOSTS FROM THE LOCAL SYSTEM. TO DO THIS, YOU HAVE TO OPEN THE MANAGEMENT SOFTWARE FOR ISCSI INITIATOR ON EACH HOST AND LOGOUT SESSIONS.

ON EACH HOST, DELETE THE TARGET PORTAL THAT IT IS USING TO ACCESS THE VOLUME ON THE LOCAL SYSTEM. YOU MAY HAVE TO USE ISCSI INITIATOR MANAGEMENT SOFTWARE TO DO THIS.

OPEN THE MANAGEMENT GUI FOR LOCAL STORAGE CONCENTRATOR AND INITIATE VOLUME REPLICATION TO THE REMOTE SYSTEM. SEE IMAGE BELOW. WAIT UNTIL THE REPLICATION HAS FINISHED SUCCESSFULLY. USE THE SAME MANAGEMENT PAGE TO SUSPEND VOLUME REPLICATION. AT THIS POINT, THE CONTENT OF LOCAL VOLUME AND REMOTE VOLUME IS NOW IN SYNC. FOR MORE INFORMATION, SEE "THE ASYNCHRONOUS IMAGE MANAGEMENT SCREEN".

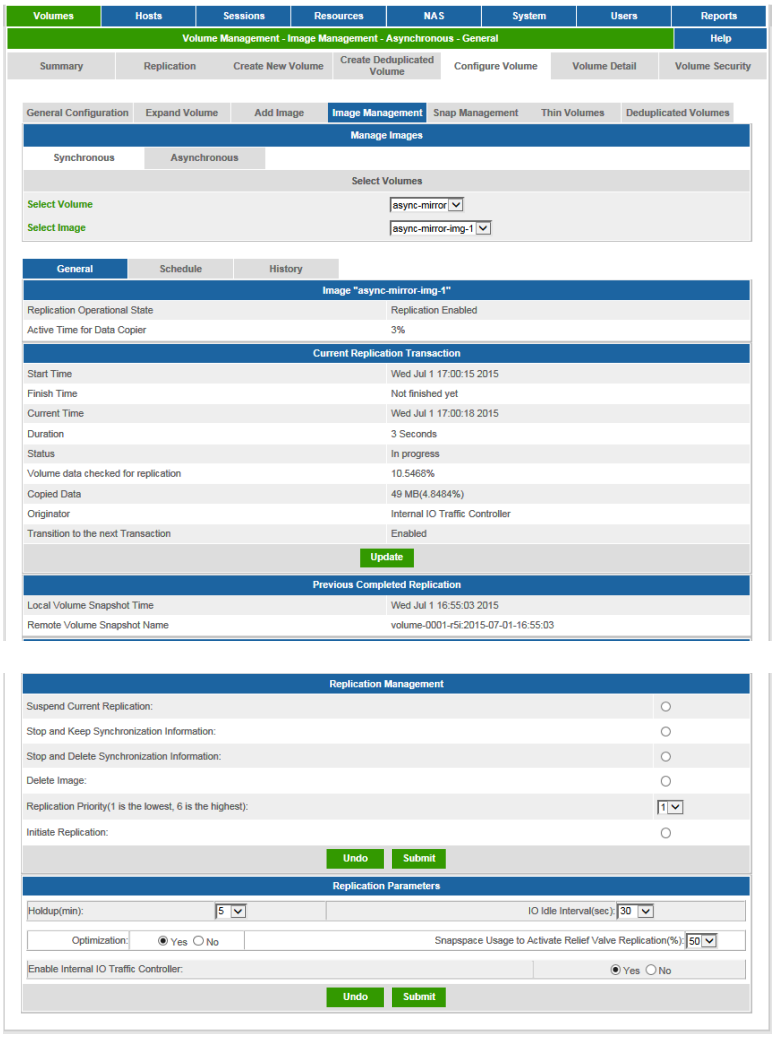


Figure 5-9 Asynchronous Management Screen for "async-mirror" on "local" Storage Concentrator

Check optimization status for volume replication. If optimization is enabled now ("Yes") it can be set "enabled" during the recovery later. See image above.

Shutdown local system. If the local system is a cluster, shut down the secondary Storage Concentrator first and then the primary Storage Concentrator. For more information, see "Shutting Down".

5.5.1.2 Remote System Management

Open the Management GUI for remote Storage Concentrator. Go to Volume Security page and delete "local" Storage Concentrator from the remote volume access list. For more information, see "Removing Existing Hosts from the ACL".

Enable asynchronous replication for the volume on the remote system. See image below. For more information, see "Configuring Asynchronous Replication for Existing

Volumes".

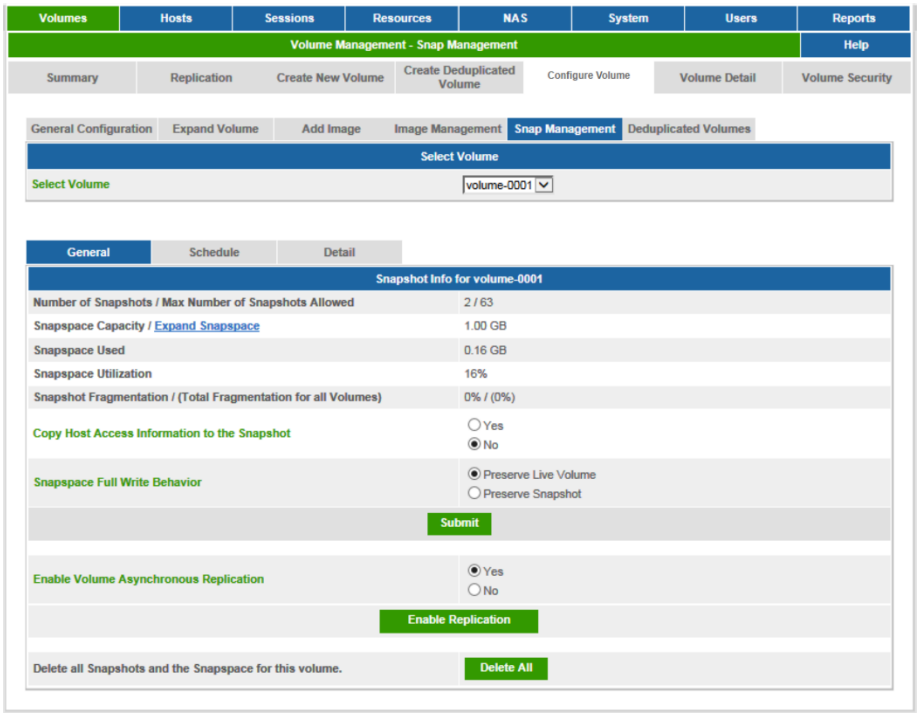


Figure 5-10 Snap Manangement Screen for "volume-0001" on "remote" System

Create "Snapshot Hold" for volume on remote system. The contents of the "Snapshot Hold" will be frozen and will be in sync with volume contents on local system while the remote volume itself will accumulate all future modifications. See image below.

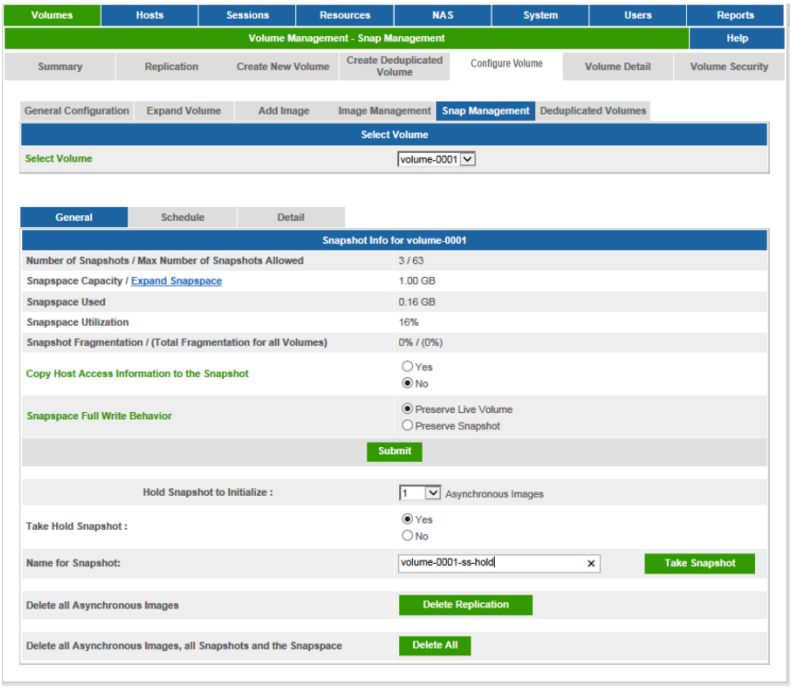


Figure 5-11 Creating Snapshot Hold for "volume-0001" on "remote" System

On each host that has to access volume on remote system open iSCSI initiator Management software and create target portal to access data port on remote system.

Verify that all needed iSCSI hosts are at the remote system host list. See image below. For more information, see "Hosts".

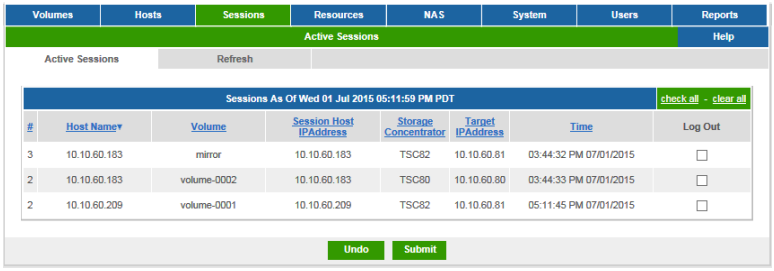


Figure 5-12 Host Management, Summary of Hosts for "remote" System

Go to Volume Security page and enable read/write access for each appropriate host. For more information, see "Volume Security".

On each host, use iSCSI initiator Management Software to create session to the volume on remote system.

Restart applications.

5.5.1.3 Local System Activation

Continue to run applications while assessing volume on remote system.

When possible, power up the local system. Wait when status for all services is healthy. For more information, see "Home Page". Don't let the hosts access the local system. **Do not restart volume replication**.

Use the Storage Concentrator Management GUI for local system to delete the volume asynchronous images. For more information, see "The Asynchronous Image Management Screen".

On the local system, delete the remote volume from the resource list. For more information, see "Removing Resources".

On the local system, delete the target portal used to access data port on the remote system. For more information, see "Target Portals".

On the remote system, create a target portal to access data port on the local system. For more information, see "Target Portals".

On the local system, verify that the remote Storage Concentrator is at the local system host list. For more information, see "Hosts".

On the local system, go to the Volume Security page and enable read/write access to the volume from the remote Storage Concentrator. For more information, see "Volume Security".

On the remote system, go to the GUI Resources Summary page and click "Discover" button. Wait until discovery is finished before you continue. For more information, see "Discovering Resources".

Verify that the volume from the local system shows as a resource. Set "Use Type" for this resource as an "SF_Managed". For more information, see ""Use" Types for Resources".

On the remote system, use the discovered resource to create a volume asynchronous image. Apply the second option to set initial synchronization state. Use the "Snapshot Hold" that was created earlier.

For more information, see "Configuring Asynchronous Replication for Existing Volumes".

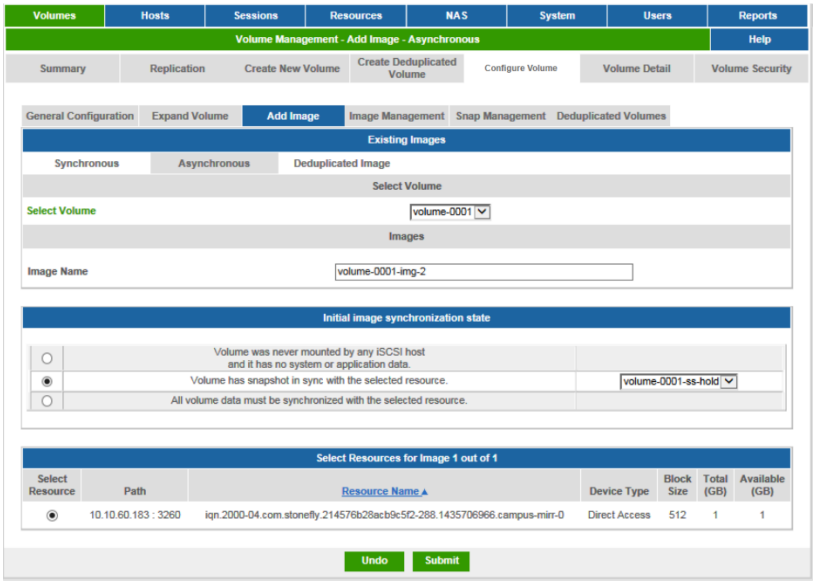
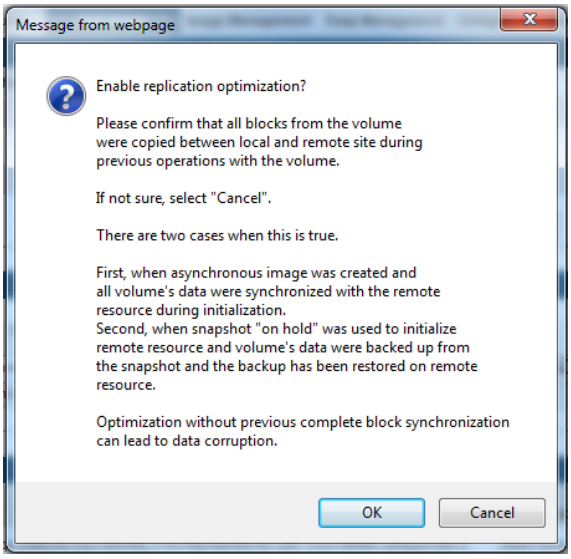


Figure 5-13 Volume Manangment, setup Asynchronous Replication for "volume-0001" on "remote" System

If optimization was enabled earlier, make sure to enable it now (Click [OK] when the following prompt appears.) Do not enable replication optimization if optimization was not enabled before.



On the remote system, start replication manually. For more information, see "The Asynchronous Image Management Screen".

Let the remote system replace the "Snapshot Hold" as a synchronization point before going to the next step. For more information, see "Managing Snapshots and Snapspace" and "Information Displayed about. Snapshots".

5.5.1.4 Restore Local System as a Primary Site

Stop applications.

DISCONNECT ALL HOSTS FROM THE REMOTE SYSTEM. TO DO THIS, YOU HAVE TO OPEN ISCSI INITIATOR MANAGEMENT SOFTWARE ON EACH HOST AND LOGOUT SESSIONS.

On each host, delete the target portal that it is using to access the volume on the remote system. You may have to use iSCSI initiator management software to do this.

Open the Management GUI for the remote Storage Concentrator and initiate volume replication to the local system. Wait until replication is finished successfully before you continue. Delete volume replication. Keep volume snap enabled. The content of the remote volume and local volume is now in sync. For more information, see "The Asynchronous Image Management Screen".

On the remote system, delete the local volume from the resource list. For more information, see "Removing Resources".

On the remote system, delete the target portal to access the data port on the local system. For more information, see "Target Portals".

On the local system, delete the remote system from the local volume access list. For more information, see "Removing Existing Hosts from the ACL".

On the local system, delete the remote system from the host list. For more information, see "Removing a Host".

On the local system, create a target portal to access the data port on the remote system. For more information, see "Target Portals".

Verify that the local system is on the remote system host list. For more information, see "Hosts". This step is similar to the step 5 from the "Remote System Management".

On the remote system, go to the Volume Security page and enable read/write access to the volume for the host that represents the local system. For more information, see "Volume Security".

On the local system, go to the GUI Resources Summary page and click the "Discover" button. Wait until discovery is finished before you continue. For more information, see "Discovering Resources".

Verify that the volume from the remote system shows up as a resource. Set "Use Type" for this resource as a "SF_Managed". For more information, see ""Use" Types for Resources".

Enable asynchronous replication for the volume on the local system. This step is similar to the step 2 from the "Remote System Management".

Create a "Snapshot Hold" for the volume on the local system. The contents of the "Snapshot Hold" will be frozen and will be in sync with the contents of local and remote volumes. This step is similar to the step 3 from the "Remote System Management".

On the local system, use the discovered resource to create the volume asynchronous image. Apply the second option to set the initial synchronization state. Use the "Snapshot Hold" that was created early. If optimization was enabled earlier, make sure to enable it now. **Do not enable replication optimization if optimization was not enabled before**. This step is similar to the steps 11 and 12 from the "Local System Activation".

On the local system, start replication manually. This step is similar to the step 13 from the "Local System Activation".

On each host that has to access volume on local system, open the iSCSI initiator Management software and create target portal to access the data port on the local system.

Verify that all of the needed iSCSI hosts are on the local system host list. For more information, see "Hosts". This step is similar to the step 5 from the "Remote System Management".

Go to the Volume Security page and enable read/write access for each appropriate host. For more information, see "Volume Security".

On each host, use iSCSI initiator Management Software to create sessions to the volume on the local system.

Restart applications.

5.5.1.5 Restrictions

The "Snapshot Hold" accumulates all changes to the remote volume that were made during period of time when the local system is down. The remote system can keep the snapshot alive until unused snap space is available to provide extension of the snapshot data. When the snap space is exhausted, the "Snapshot Hold" is deleted automatically. In this case the second option can't be used again to set initial synchronization state. User has to employ option 3 to set replication during the second attempt. It can be more time consuming and there is no guarantee that there are enough snap space to finish the second attempt successfully. If this is the case user can relocate local system to be close to the remote system and repeat the synchronization procedure.

5.5.2 Unplanned Failover and Managed Recovery

This is the case when local system is shut down without any preliminary warning so no management steps were executed on local system before it went down. If hosts are still on line, user has to do the next:

STOP APPLICATIONS.

DISCONNECT ALL HOSTS FROM THE LOCAL SYSTEM. TO DO THIS, YOU HAVE TO OPEN THE ISCSI INITIATOR MANAGEMENT SOFTWARE ON EACH HOST AND LOGOUT SESSIONS.

On each host, delete the target portal that it is using to access the volume on the local system. You may have to use iSCSI initiator management software to do this.

User will have to know the optimization status for volume replication. If optimization is enabled it can be set "enabled" during the recovery later. This information has to be checked before disaster at time when replication between local and remote system is provisioned initially.

There is no synchronization between contents of local and remote volumes, but in case of established and successful replication, both sites should have snapshots with contents in sync. These snapshots have to be used during the recovery.

5.5.2.1 Remote System Management

Open the Management GUI for the remote Storage Concentrator. Go to the Volume Security page and delete the "local" Storage Concentrator from the remote volume access list. For more information, see "Removing Existing Hosts from the ACL".

Find the most recent snapshot created for the remote volume during replication. Rollback the volume to the snapshot. Remember name of the snapshot. See picture below.

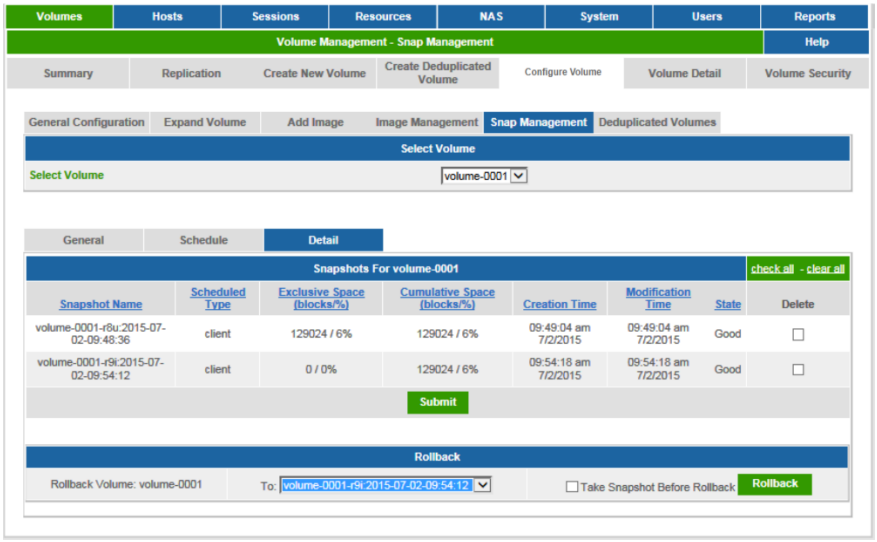


Figure 5-14 List of Snapshots for "volume-0001" on "remote" System

Enable asynchronous replication for the volume on the remote system. This step is similar to the step 2 from the "Remote System Management".

Create a "Snapshot Hold" for the volume on the remote system. The contents of the "Snapshot Hold" will be frozen and will be in sync with the contents of the snapshot used to rollback. The volume itself will accumulate all future modifications. This step is similar to the step 3 from the "Remote System Management".

On each host that has to access the volume on the remote system, open the iSCSI initiator Management software and create a target portal to access the data port on the remote system.

Verify that all of the needed iSCSI hosts are on the remote system host list. For more information, see "Hosts". This step is similar to the step 5 from the "Remote System Management".

Go to the Volume Security page and enable read/write access for each appropriate host. For more information, see "Volume Security".

On each host use the iSCSI initiator Management Software to create session to the volume on the remote system.

Restart applications.

5.5.2.2 Local System Activation

Continue to run applications by assessing the volume on the remote system.

When possible, power up the local system. Wait when status for all services is healthy. Do not let the hosts access the local system.

On the local system, suspend the volume replication for all asynchronous images. See picture below.

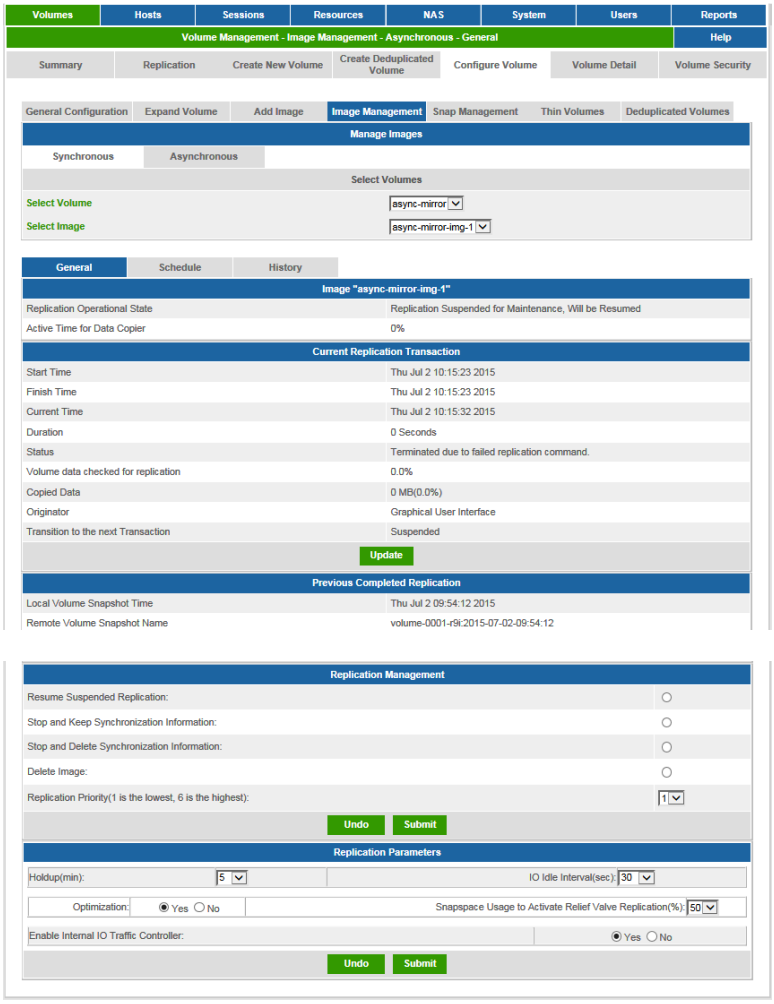


Figure 5-15 Volume Manangment on "local" System, Asynchronous Images Screen with suspended replication for "async-mirror"

On the local system, check the volume replication snapshots. Find an active snapshot with the same modification and creation time that was used to generate name for the remote volume rollback snapshot. See picture below.

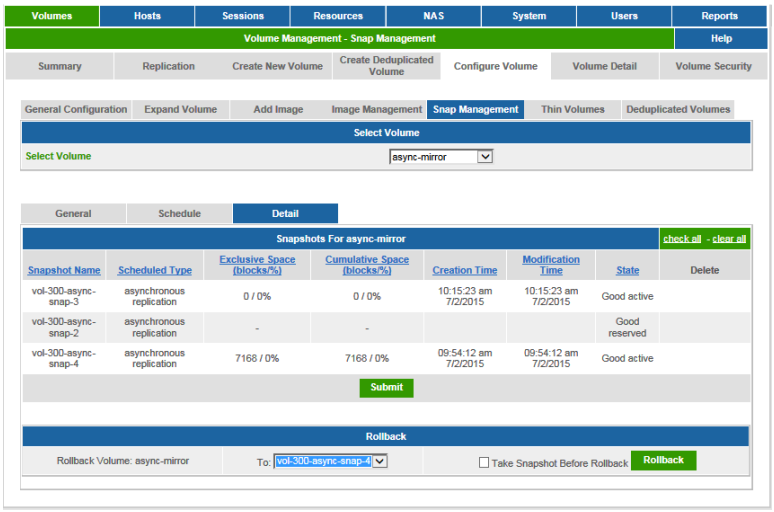


Figure 5-16 Volume Management on "local" Storage Concentrator, Snapshot images for volume "async-mirror"

Rollback the local volume to the snapshot. The content of the local volume is now in sync with the contents of the "Snapshot Hold" on the remote system. See picture above.

Use the Storage Concentrator Management GUI for the local system to delete the volume asynchronous images. For more information, see "The Asynchronous Image Management Screen".

On the local system, delete the remote volume from the resource list. For more information, see "Removing Resources".

On the local system, delete the target portal used access the data port on the remote system. For more information, see "Target Portals".

On the remote system, create a target portal to access the data port from the local system. For more information, see "Target Portals".

On the local system, verify that the remote Storage Concentrator is on the local system host list. For more information, see "Hosts". This step is similar to the step 5 from the "Remote System Management".

On the local system, go to the Volume Security page and enable read/write access to the volume from the remote Storage Concentrator. For more information, see "Volume Security".

On the remote system, go to the GUI Resources Summary page and click the "Discover" button. Wait until discovery is finished before you continue. For more information, see "Discovering Resources".

Verify that the volume from the local system shows as a resource. Set "Use Type" for this resource as an "SF_Managed". For more information, see ""Use" Types for Resources".

On the remote system, use the discovered resource to create the volume asynchronous image. Apply the second option to set initial synchronization state. Use the "Snapshot Hold" that was created earlier. If optimization was enabled earlier, make sure to enable it now. **Do not enable replication optimization if optimization was not enabled before**. This step is similar to the steps 11 and 12 from the "Local System Activation".

On the remote system, start the volume replication manually. For more information, see "The Asynchronous Image Management Screen".

Let the remote system replace the "Snapshot Hold" as a synchronization point before going to the next step. For more information, see "Managing Snapshots and Snapspace" and "Information Displayed about. Snapshots".

5.5.2.3 Restore Local System as a Primary Site

Stop applications.

DISCONNECT ALL HOSTS FROM THE REMOTE SYSTEM. TO DO THIS, YOU HAVE TO OPEN THE ISCSI INITIATOR MANAGEMENT SOFTWARE ON EACH HOST AND LOGOUT SESSIONS.

On each host, delete the target portal that the host is using to access the volume on the remote system. You may have to use the iSCSI initiator management software to do this.

Open the Management GUI for the remote Storage Concentrator and initiate the volume replication to the local system. Wait until the replication is finished successfully before you continue. Delete the volume replication. Keep the volume snap enabled. The content of the remote volume and local volume is now in sync. For more information, see "The Asynchronous Image Management Screen".

On the remote system, delete the local volume from the resource list. For more information, see "Removing Resources".

On the remote system, delete the target portal used to access the data port on the local system. For more information, see "Target Portals".

On the local system, delete the remote system from the local volume access list. For more information, see "Removing Existing Hosts from the ACL".

On the local system, delete the remote system from the host list. For more information, see "Removing a Host".

On the local system, create a target portal to access the data port on the remote system. For more information, see "Target Portals".

Verify that the local system is on the remote system host list. For more information, see "Hosts". This step is similar to the step 5 from the "Remote System Management".

On the remote system, go to the Volume Security page and enable read/write access to the volume for the host that represents the local system. For more information, see "Volume Security".

On the local system, go to the GUI Resources Summary page and click the "Discover" button. Wait until discovery is finished before you continue. For more information, see "Discovering Resources".

Verify that the volume from the remote system shows as a resource. Set "Use Type" for this resource as a "SF_Managed". For more information, see ""Use" Types for Resources".

Enable asynchronous replication for the volume on the local system. This step is similar to the step 2 from the "Remote System Management".

Create a "Snapshot Hold" for the volume on the local system. The contents of the "Snapshot Hold" will be frozen and will be in sync with the contents of the local and the remote volumes. This step is similar to the step 3 from the "Remote System Management".

On the local system, use the discovered resource to create the volume asynchronous image. Apply the second option to set initial synchronization state. Use the "Snapshot Hold" that was created earlier. If optimization was enabled earlier, make sure to enable it now. **Do not enable replication optimization if optimization was not enabled before.** This step is similar to the steps 11 and 12 from the "Local System Activation".

On the local system, start the volume replication manually. For more information, see "The Asynchronous Image Management Screen".

On each host that has to access the volume on the local system, open the iSCSI initiator Management software and create a target portal to access the data port on the local system.

Verify that all of the needed iSCSI hosts are on the local system host list. For more information, see "Hosts". This step is similar to the step 5 from the "Remote System Management".

Go to the Volume Security page and enable read/write access for each appropriate host. For more information, see "Volume Security".

On each host use the iSCSI initiator Management Software to create session to the volume on the local system.

Restart applications.

5.5.2.4 Restrictions

The "Snapshot Hold" accumulates all changes to the remote volume that were made during period of time when the local system is down. The remote system can keep the snapshot alive until unused snap space is available. When the snap space is exhausted, the "Snapshot Hold" is deleted automatically. In this case the second option can't be used again to set initial synchronization state. User has to employ option 3 to set replication the second time. It can be more time consuming and there is no guarantee that there are enough snap space to finish the second attempt successfully. If this is

the case user can relocate local system to be close to the remote system and repeat the synchronization procedure.

5.5.3 Primary Site Hardware Replacement

This is the case when failover from the local system was executed in planned or unplanned way, but hardware on the primary site has to be replaced before activity can be restored there. User has to follow guidelines for each case described early except the "Local System Activation". This step has to be executed in different way.

5.5.3.1 Local System Activation

Continue to run applications by assessing the volume on the remote system.

When possible, power up the replacement unit on the primary site.

Follow standard recommendations to set up the replacement system.

Wait when status for all services is healthy before you continue.

Provision local volume of the same size as it was before.

Don't let hosts access the volume.

Enable snapshots on the volume.

On the remote system, create a target portal to access the data port on the local system. For more information, see "Target Portals".

On the local system, verify that the remote Storage Concentrator is on the local system host list. For more information, see "Hosts". This step is similar to the step 5 from the "Remote System Management".

On the local system, go to the Volume Security page and enable read/write access to the volume from the remote Storage Concentrator. For more information, see "Volume Security".

On the remote system, go to the GUI Resources Summary page and click the "Discover" button. Wait until discovery is finished before you continue. For more information, see "Discovering Resources".

Verify that the volume from the local system shows as a resource. Set "Use Type" for this resource as a "SF_Managed". For more information, see ""Use" Types for Resources".

On the remote system, use the discovered resource to create the volume asynchronous image. Apply the third option to set initial synchronization state. For more information, see "Synchronizing the Data in the Volumes". If the "Snapshot Hold" was created earlier, delete it. For more information, see "Delete Snapshots".

Let the remote system finish initial synchronization before going to the next step. For

more information, see "The Asynchronous Image Management Screen".

5.5.3.2 Restrictions

All changes to the remote volume made during period of time when initial synchronization is in progress are accumulated by special replication snapshot. The remote system can keep this snapshot alive until unused snap space is available. When the snap space is exhausted, the snapshot is deleted automatically. It means that if the remote volume does not have enough snap space, the initial synchronization cannot be finished successfully. User can relocate the local system to be close to the remote system and repeat the procedure. When synchronization is done, the local system can be sent to the primary site for activation.

5.5.4 Planned Shutdown and Recovery

This is the case when the local system and all services that are relaying on data provided by the local system have to be shutdown. After some period of time the local system and the services have to be activated. There are 2 different periods in procedure to handle the disaster.

The first one is when the planned shutdown is executed.

The second one is the period to put local site and services on line.

5.5.4.1 Local System Shutdown

STOP APPLICATIONS.

DISCONNECT ALL HOSTS FROM THE LOCAL SYSTEM. TO DO THIS, YOU HAVE TO OPEN THE MANAGEMENT SOFTWARE FOR ISCSI INITIATOR ON EACH HOST AND LOGOUT SESSIONS.

OPEN THE MANAGEMENT GUI FOR THE LOCAL STORAGE CONCENTRATOR AND INITIATE THE VOLUME REPLICATION TO THE REMOTE SYSTEM. WAIT UNTIL REPLICATION IS FINISHED SUCCESSFULLY BEFORE YOU CONTINUE. SUSPEND THE VOLUME REPLICATION. THE CONTENT OF THE LOCAL VOLUME AND REMOTE VOLUME IS NOW IN SYNC. FOR MORE INFORMATION, SEE "THE ASYNCHRONOUS IMAGE MANAGEMENT SCREEN".

SHUTDOWN THE LOCAL SYSTEM. IF THE LOCAL SYSTEM IS A CLUSTER, SHUTDOWN THE SECONDARY STORAGE CONCENTRATOR FIRST AND THE PRIMARY STORAGE CONCENTRATOR THE SECOND. FOR MORE INFORMATION, SEE "SHUTTING DOWN".

5.5.4.2 Restore Local System as a Primary Site

When possible, power up the local system and servers. Wait when status for all services is healthy.

Verify that the volume from the remote system shows as a resource. For more information, see "Resources". Resume asynchronous replication for the volume on the local system. For more information, see "The Asynchronous Image Management Screen".

On each host use the iSCSI initiator Management Software to create session to the

volume on the local system.

Restart applications.

5.5.5 Unplanned Shutdown and Managed Recovery

This is the case when the local system is shut down without any preliminary warning so no management steps were executed on the local system before it went down. There is no server access to the volume on the remote site in this case.

If hosts are still on line, user has to do the next:

STOP APPLICATIONS.

DISCONNECT ALL HOSTS FROM THE LOCAL SYSTEM. TO DO THIS, YOU HAVE TO OPEN THE ISCSI INITIATOR MANAGEMENT SOFTWARE ON EACH HOST AND LOGOUT SESSIONS.

5.5.5.1 Restore Local System as a Primary Site

When possible, power up the local system and servers. Wait when status for all services is healthy.

Verify that the volume from the remote system shows as a resource. For more information, see "Resources".

On each host use the iSCSI initiator Management Software to create session to the volume on the local system.

If it is needed, execute server's recovery procedure for the volume.

Restart applications.

5.5.6 Test Replication

To verify that data on the remote system has the appropriate context, the next procedure can be employed:

On the remote system open the GUI Management and find the replication snapshot previous to the most recent one. The most recent replication snapshot should not be touched because it can be used to handle the possible recovery. The volume itself can be in state of updating its contexts. So the volume is not an appropriate target to open session to. The remote system is rotating the replication snapshots each time when the number of the snapshots reaches 63 or snap space usage reaches threshold. It means that the selected snapshot will be deleted early or later automatically. User has some time to test the replication before the selected snapshot is deleted. See picture below.

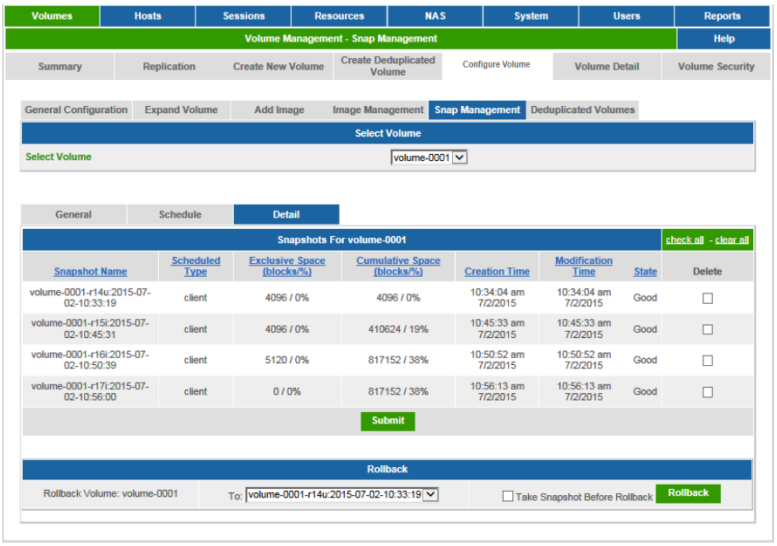


Figure 5-17 Volume Management on "remote" System, Snapshot images for "volume-0001"

Select iSCSI host and create a target portal to access the data port on the remote system.

On the remote system, verify that the selected host is on the remote system host list. For more information, see "Hosts". This step is similar to the step 5 from the "Remote System Management".

On the remote system, go to the Volume Security page and enable read/write access to the selected snapshot from the selected host. See picture below. For more information, see "Volume Security".

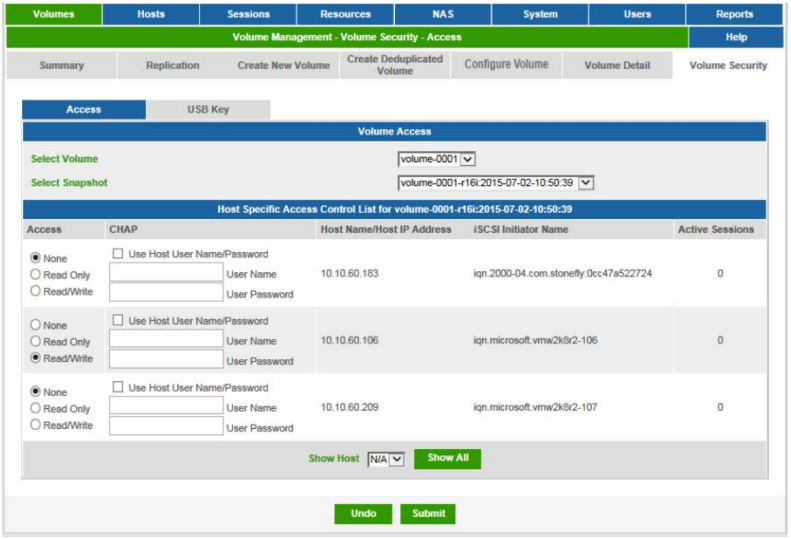


Figure 5-18 Volume Management on "remote" System, Volume Access Security for Replication Snapshot

On selected host, use the iSCSI initiator Management Software to create session to the selected snapshot.

On host system, use any available tools to verify context of the snapshot.

After the test is done, close session to the snapshot. Delete the created target portal from the host.

Remove the host from the snapshot access list. For more information, see "Removing Existing Hosts from the ACL".

5.5.6.1 Restrictions

Test has to be executed before the remote system deletes the used snapshot.

This page is intentionally left blank.

Chapter 6

Snapshot

This chapter provides an overview of StoneFly Snapshot, including the steps necessary to automatically take scheduled Snapshots and manage Snapshots of storage volumes. (For information on how to create a volume, see "Creating a Volume".)

6.1 StoneFly Snapshot Overview

StoneFly Snapshot creates virtual, temporary, and perishable point-in-time copies of an active Live Volume. A Snapshot contains an image of the volume at the exact point-in-time that the Snapshot was taken. Snapshots can be created nearly instantaneously. The snapshot volume appears to the host as if it was a regular logical volume. Even after changes are made to the original volume, a Snapshot preserves a copy exactly as it existed when it was taken. Snapshots persist across reboots and can be mounted and accessed and written to like any other volume. Snapshot data are encrypted if the original volume was encrypted.

Snapshots are not replication, like StoneFly Mirroring. An exact copy would mean that five Snapshots require five times the storage space of the original. Snapshots use Copy-On-First-Write (COFW) mechanisms that only copy data blocks if and when an application changes them. Often, this "copy" is of a pointer followed by an allocation of a data block.

The primary benefits of Snapshots are their smaller storage requirements and their comparatively low system overhead.

Snapshots impose read and write overhead in exchange for their convenience. Write performance will be slightly degraded as each write I/O must be copied to the Snapspace and metadata saved to maintain the integrity of the live volume and all the snapshots. Keep in mind that performance with increasing numbers of Snapshots can be highly variable depending on the amount of data changed between each.

Snapshots are not meant to be kept for an indefinite time; rather they are created, used and later overwritten to make space for more current snapshots. Due to the nature of the Snapshot implementation, it may be necessary to overwrite or delete Snapshots to free up space to make new ones or to preserve the availability of the Live Volume. Volumes that change significantly between Snapshots may not be able to support as many as a volume with less change.

6.1.1 Typical StoneFly Snapshot Configuration

Snapshots are useful for applications such as:

- Backup/restore---Reduces the backup window to the time required to quiesce the volume and take a snapshot
- Data migration
- Disaster recovery---a rollback feature provides the ability to return to an earlier point in time. This is useful to return to a known good state and allows recover in the event of data corruption or a virus
- Temporary copies of a volume that can be used and discarded

The Snapshot volume comprises the Live Volume, the Snapspace where the changes are written and all the snapshots.

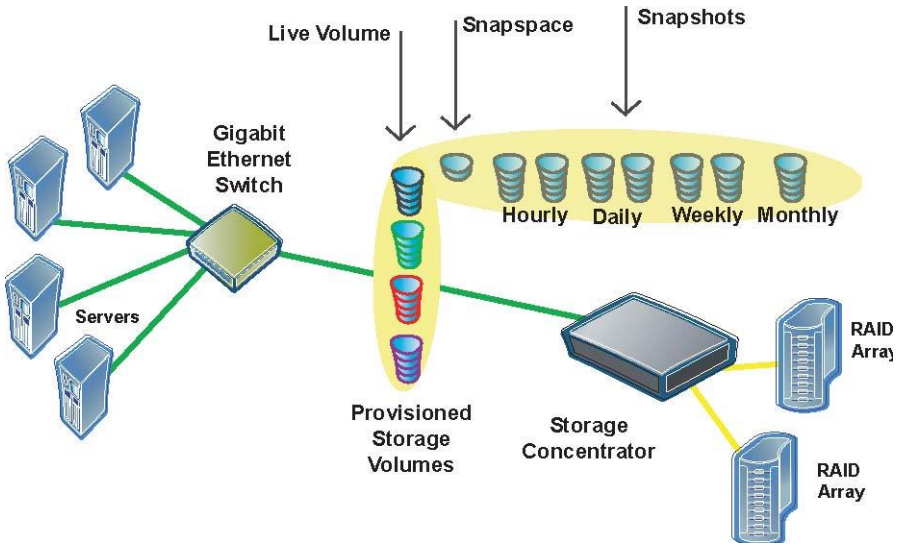


Figure 6-1 Typical StoneFly Snapshots

6.1.2 StoneFly Snapshot Terminology

Volume

The volume is storage presented to a host through the Storage Concentrator. The logical volume is the starting point before creating a Snapshot Live Volume.

Live Volume

The volume that is being accessed by the host for normal operations. The data on this volume is complete and not accessed via pointers and/or data structures. The difference between a regular volume and a Live Volume is that the Live Volume has been Snapshot enabled. The Live Volume can also be mirrored, depending on performance requirements since there will be overhead associated with both Snapshots and mirrors on the same volume.

6.1.3 Snapshot Function Definitions

Snapshot

A Snapshot is a virtual volume that represents a point-in-time image of a Live Volume. The Snapshot data is a combination of data from the Live Volume and data from the Snapspace. A Snapshot cannot be mirrored.

Snapspace

Snapspace is storage space configured by the Storage Concentrator used to store the changed data for maintaining snapshots. This data space is not directly readable by hosts. For fault tolerance, the Snapspace can be mirrored. (The Snapspace must be mirrored on HSC installations.) The size of the Snapspace may be increased by the Administrator. See the section on Volume Configuration.

Take Snapshot

The process of creating a new Snapshot of the Live Volume. The user can select whether the Snapshot inherits the security features of the Live Volume, including ACLs, CHAP secrets, etc.

Rollback

Rollback is a disaster recovery feature that allows users to return to the last known "good" data point in the event of corruption, viruses, or lost files. Essentially, performing a rollback returns the Live Volume to the known state of one of the Snapshots.

Preserve Snapshots or Preserve Live Volume

When the Snapspace is at 100% of capacity or filled to the point that taking another Snapshot will exceed this threshold, the system can be configured to either delete some older Snapshots or refuse writes to the Live Volume. The recommended alternative is to preserve the Live Volume by deleting Snapshots. Some installations that require Snapshots for backup purposes may not want to select this option if doing so means their system will not be completely backed up. In this case, the user may elect to Preserve Snapshots. For most installations, it is preferable to preserve the Live Volume (this is the default option). The policy is applied to the Live Volume, not to individual Snapshots.

The Storage Concentrator will display a warning message when the Snapspace is 80% full. It will display another when the Snapspace has reached 95% full. It is recommended that you delete enough Snapshots to bring the Snapspace below the 80% full threshold level. Or the size of the Snapspace may be increased.

Snapshot Fragmentation/Total Fragmentation for All Volumes:

Because the system must maintain a record of all transactions in the Snapspace, the index tables that keep track can sometimes become full. This occurs when there are a lot of random small changes, such as that introduced by backup software or virus checkers. This can cause depletion of available space to keep track. Snapshot Fragmentation/Total Fragmentation for All Volumes values are displayed as percentages that represent the amount of the index table fragmentation on the volume selected as well as the total amount of fragmentation for all Snapshot volumes. If these values are high, then steps should be taken to reduce the total percentages by deleting snapshots.

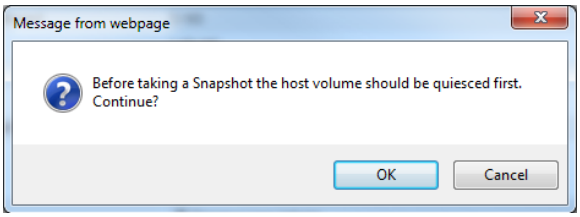
These percentages represent the amount of the index table fragmentation on the volume selected as well as the total amount of fragmentation for all Snapshot volumes. When the Total Fragmentation for All Volumes approaches 95%, a warning message is

sent to the Logs and notifications. When the value approaches 100%, the system will delete the oldest snapshot in the volume that has the most fragmentation. If a particular volume represents a high level of fragmentation compared to all other volumes, deleting the snapshots in this volume before the values reach maximums will often cause a significant reduction in the total fragmentation level.

Do NOT under any circumstances attempt to defragment a disk array due to the numbers shown by the Storage Concentrator Snapshot Fragmentation values. These figures do not have any relationship to disk fragmentation.

Quiescing Volumes:

All Snapshots need to be taken with the Live Volume quiesced. This is accomplished primarily by logging off the hosts connected to the volume. If taking a Snapshot manually, the system will automatically query the user if host volumes are quiesced as shown below.



When using the automatic Snapshots feature, it is important that software processes are quiesced. Applications that work in the Microsoft Volume Shadow Copy Services (VSS) features have the ability to run scripts and batch files for applications such as SQL and Microsoft Exchange that quiesce a host volume. There are also software utilities, such as St. Bernard's Open File Manager, which allow the system to take snapshots of open files on application servers. Regardless of the method used, quiescing volumes is a critical component of a sound Snapshot implementation.

Total Number of Snapshots:

StoneFly Snapshots are limited to 63 Snapshots of each live volume. This includes all Snapshots taken automatically and all taken manually. The system will only allow eight daily Snapshots until a weekly is taken, six weekly Snapshots before a monthly is taken, and so on. Scheduling Snapshots in an orderly fashion conserves Snapspace and provides the widest range of data protection. The allotted Snapspace will determine how many actual Snapshots the system can manage. The more data changes on the Live Volume, or the more data that is added to the volume, the more Snapspace will be required to maintain the system. The recommended size of the Snapspace is 20-30% the size of the live volume. The size may be increased if necessary later.

6.1.4 Optimizing Snapshot Environments

Enhancing Availability:

The Snapspace for any Snapshot enabled volume should reside on the same Resource as the Live Volume. In situations where a Resource fails the Live Volume and Snapspace must both be available to process IO's against the Live Volume. If the Snapspace is on a different Resource the failure of the that resource will cause IO's to fail against the Live Volume even if the Live Volume is on a surviving Resource.

Mirroring Live Volumes and Snapspace:

Using StoneFly Mirroring, the Storage Concentrator allows users to mirror the Live Volume and the Snapspace, but not the Snapshots. Care should be taken to do this only with extremely critical volumes where system performance is not a problem. Mirroring a Snapshot Live Volume requires system overhead for maintaining both the mirror images and the Snapshots. The requirement for Cluster Mirroring in the High Availability Storage Concentrators (HSC) applies to Snapspace also. Each Snapshot enabled volume on the HSC internal storage must have mirrored Live Volume and Snapspace.

6.2 Creating Snapshots

To add a Snapshot:

1 From the *Volume Configuration* screen, click **Snap Management**.

2 From the *Volume Select* drop down menu, select the correct volume.

3 Select radio button for **Copy the Volume Security Information to the Snapshot**. Selecting **Yes** confers the Snapshot with the Host ACLs, CHAP Secret, and all other volume security information. Selecting **No** (the default) will apply no security information to the Snapshot. Security parameters can be added later

4 Select the **Snapspace Full Write Behavior** you desire.

Preserve live volume: The system will delete Snapshots in FIFO (First In First Out) order to make room in the Snapspace to take new snapshots. When the Snapspace reaches 95% utilization, the system deletes Snapshots in FIFO order until utilized Snapspace is 80% or less. However, it does not delete the last remaining Snapshot even if capacity used is over 95%. Users will receive warnings when the Snapspace reaches 80% capacity and then again just before deleting the first Snapshot. ***This is the default, recommended option.***

Preserve Snapshots: The system will stop taking new Snapshots and prevent new I/O writes until Snapshots are deleted to free up space. This assures completion of all required data and metadata writes to the Live Volume and Snapspace and preserves the integrity of both the Live volume and all previous Snapshot data. However, users may not want to use this advanced option if I/O to the volume is critical because reaching capacity will cause the hosts to unable to write to the volumes until Snapshots are deleted.

5 Click **Take Snapshot**.

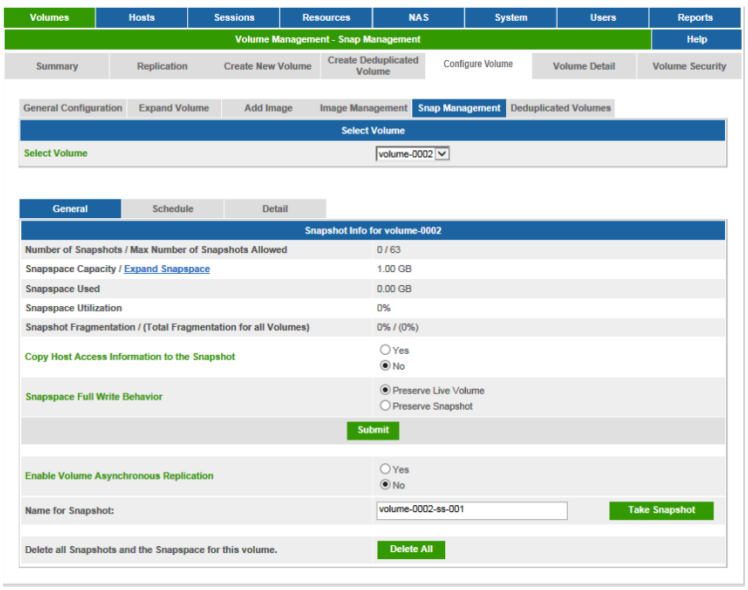
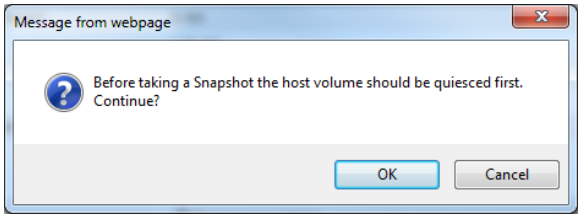


Figure 6-2 Snap Management General Screen

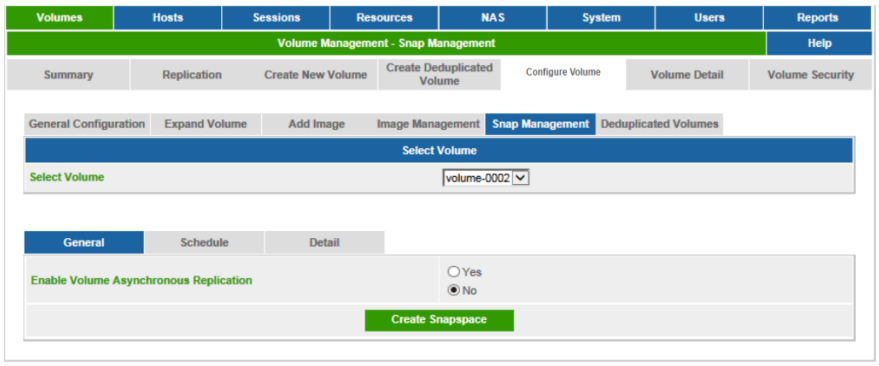
6 The *Storage Concentrator* displays a reminder message that the host volume should be quiesced before taking the Snapshot.

Click **OK** to continue or Cancel to abort the Snapshot



7 Select a name for the Snapshot. The system will automatically suggest a name, but you can enter a different name in the field.

8 If this is the first Snapshot of a volume, the system will ask you to set up Snapspace for the Snapshots. Click **Create Snapspace**.



9 Next, select the resource for the Snapspace. The system will display all the available resources. In best storage practices, the original volume and the Snapspace should reside on different physical storage resources for performance reasons.

The Storage Concentrator will recommend the size of the Snapspace. The recommended size is approximately 10-20% of the size of the Live Volume with a maximum size of 1TB. The amount to add is represented as a whole number in GB. You may use more than one resource to create the Snapspace, selecting from many available resources. The system displays current mappings for the base volume to help you determine the appropriate resource.

If possible, place the Snapspace on the same resource as the Live Volume. Any problems that occur on the resources will guarantee that a Snapshot Volume will continue to work if it's resource is good.

Enter the number(s) for the total amount under the best choice of Resources displayed. Click **Submit.** Remember that the size may be increased later if desired.

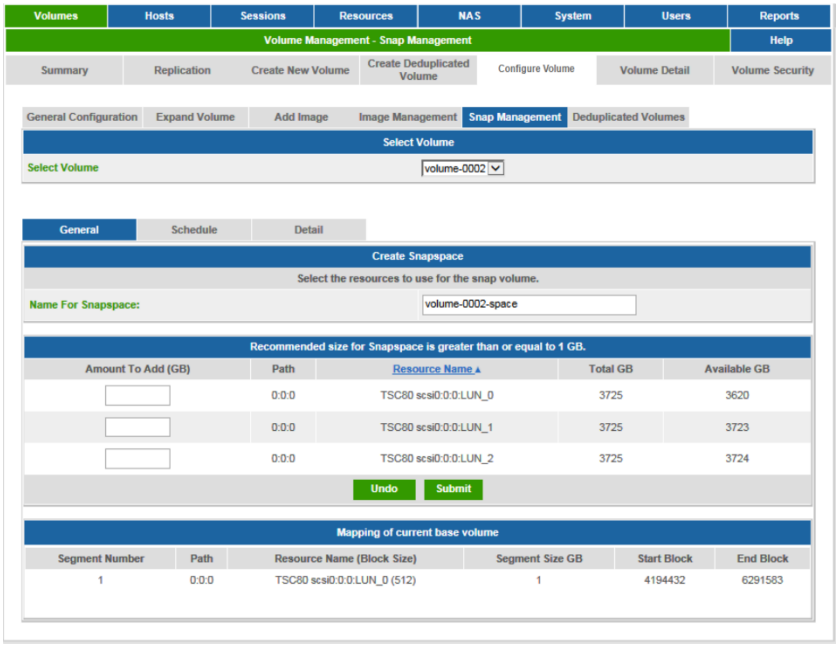
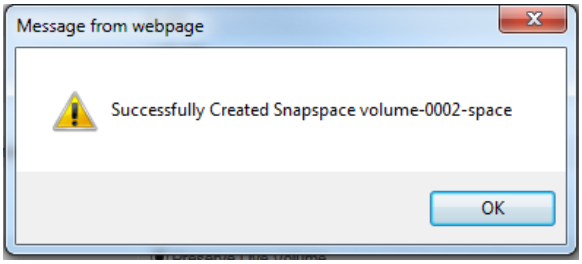


Figure 6-3 Selecting the Resources to Use for Snapspace

The Storage Concentrator will configure the Snapspace. The following message will display:



After the Snapshot is taken, the Snap Management General screen will appear like the following example.

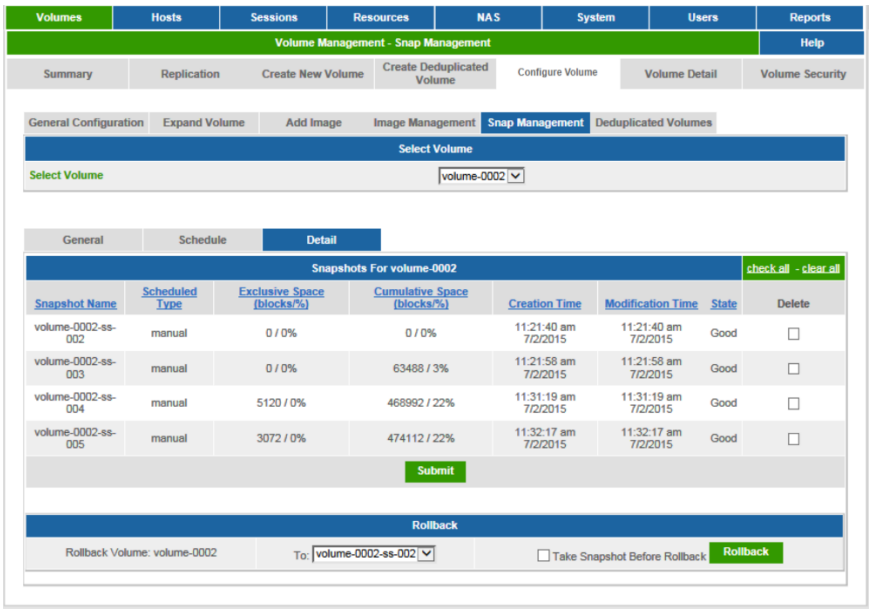


Figure 6-4 Volume Management, Snap Management Screen

6.2.1 Information Displayed about Snapshots

The *Snap Management Detail* Screen provides information about Snapshots

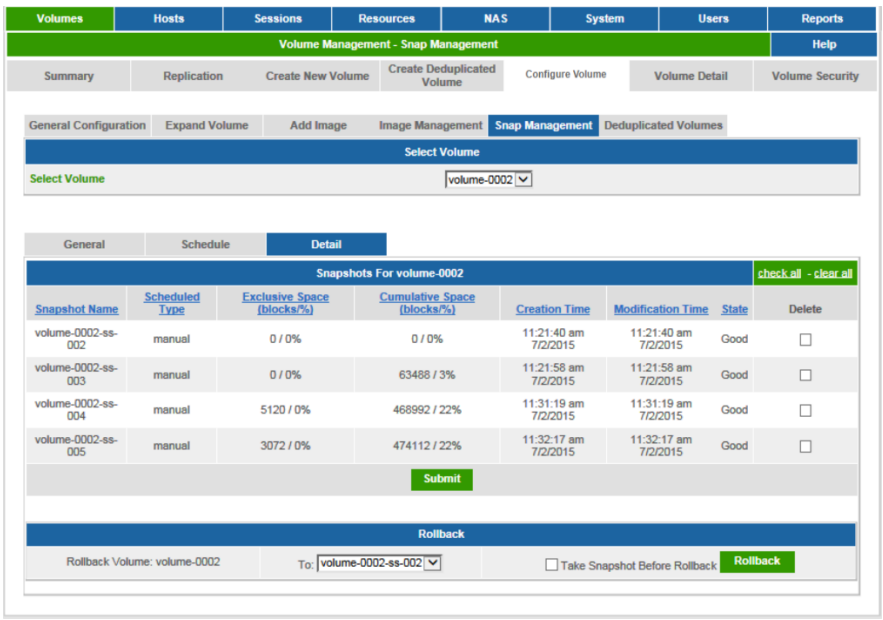


Figure 6-5 Snap Management Detail Screen

The following information is displayed (Snapshots are displayed in chronological order):

Name: The System Name for the Snapshot

Scheduled Type: Manual, Hourly, Daily, Weekly, Monthly

Exclusive Space: (blocks/Percentage Used): The space in blocks that are required in the Snapspace to maintain each individual Snapshot, as well as the percentage of the Snapspace.

Cumulative Space: (blocks/Percentage Used): The cumulative space in blocks that have been used in the Snapspace for a particular Snapshot as well as all the previous Snapshots. The most recent Snapshot will require the least amount of cumulative space since it will have the fewest number of changes from the Live Volume. The percentage displayed is the total percentage of the Snapspace required to maintain a particular Snapshot as well as all previous Snapshots. The last Snapshot taken will display the total cumulative blocks and percentage used.

Creation Time: The time of day and date the Snapshot was created.

Last Modification Time: The last time the Snapshot was written to. Changes to the Live Volume will ripple through the Snapshots so that many will display similar last modification times.

State: Good, Invalid, or Offline. (Invalid occurs if there is Snapshot data loss or corruption.)

6.2.2 Scheduling Snapshots

The *Storage Concentrator* can automatically take Snapshots at predetermined intervals using the Snapshot scheduler. Administrators can configure the system to automatically take Snapshots hourly, daily, weekly and monthly. StoneFly has configured the total number of Snapshots to take advantage of n+1 Snapshots in each category (e.g. Daily Snapshots has a maximum of eight retained Snapshots---one for each day of the week plus one).

Type of Snapshot	Max. Number of Snapshots
Hourly	25
Daily	8
Weekly	6
Monthly	13
Manually Taken Snapshots (If maximum number in all other categories are taken AND if there is sufficient Snapspace.)	11

The scheduler provides the widest range of coverage possible. Administrators will want to set up their systems to take scheduled daily Snapshots after the hourly capacity is reached, weekly Snapshots after daily capacity is reached, etc.

The scheduler does not quiesce volumes. Quiescing volumes must take place in the application software.

Hourly Snapshots

To schedule an Hourly Snapshot:

1 From the *Volume Configuration* screen, click **Snap Management**.

2 From the *Snap Management* screen Select **Schedule**.

3 From the Snapshot Schedule, select **Hourly**.

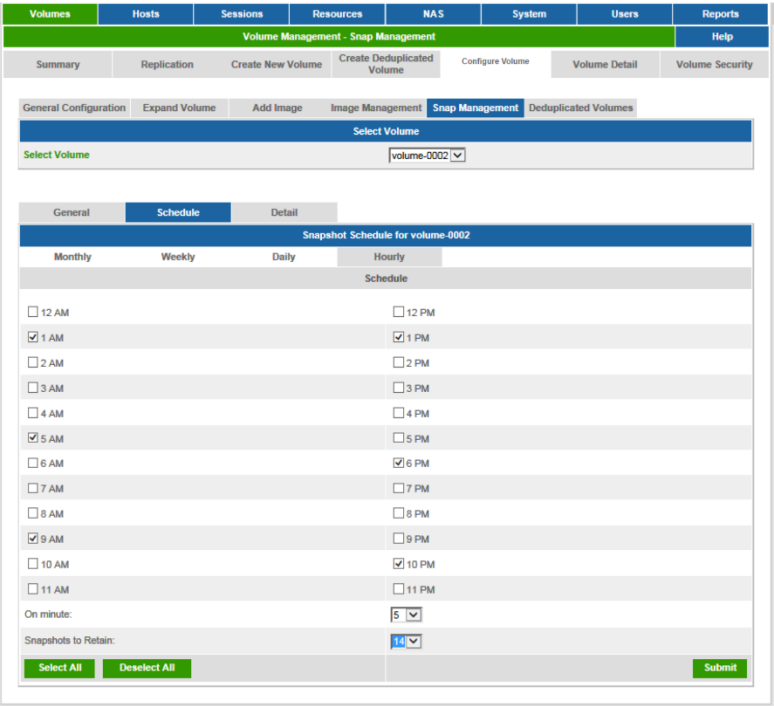


Figure 6-6 Schedule Hourly Snapshots Screen

4 Select the Hours desired by clicking the Checkboxes.

5 Select the minute desired from the drop down menu. All hourly Snapshots will all be taken at the same number of minutes after the hour.

6 Select the number of hourly Snapshots to retain. The maximum number of retained hourly Snapshots is 25. Once the number of hourly retained Snapshots has been reached, the system will automatically start overwriting

snapshots to make room for new Snapshots.

7 Click **Submit.** The Snapshots will then be taken at the hours specified.

Daily Snapshots

To schedule a Daily Snapshot:

1 From the *Volume Configuration* screen, click **Snap Management**.

2 From the *Snap Management* screen Select **Schedule**.

3 From the Snapshot Schedule, select **Daily**.

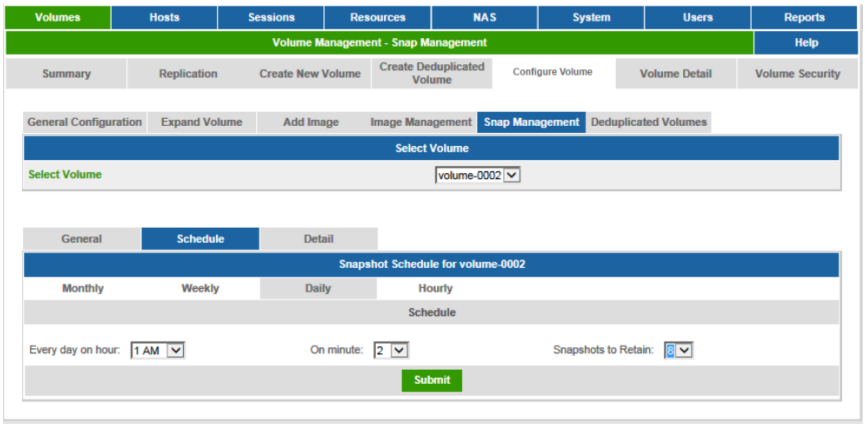


Figure 6-7 Schedule Daily Snapshots Screen

4 Select the Hour desired for the daily Snapshot using the dropdown menu.

5 Select the minute desired from the drop down menu. It is recommended that you do not select the same time as any other Scheduled Snapshot.

6 Select the number of Daily Snapshots to retain. The maximum number of retained Daily Snapshots is 8. Once the number of daily retained Snapshots has been reached, the system will automatically start overwriting Snapshots to create room for new ones.

7 Click **Submit**. The daily Snapshots will then be taken at the scheduled hour.

Weekly Snapshots

To schedule a Weekly Snapshot:

1 From the *Volume Configuration* screen, click **Snap Management**.

2 From the *Snap Management* screen Select **Schedule**.

3 From the Snapshot Schedule, select **Weekly**.

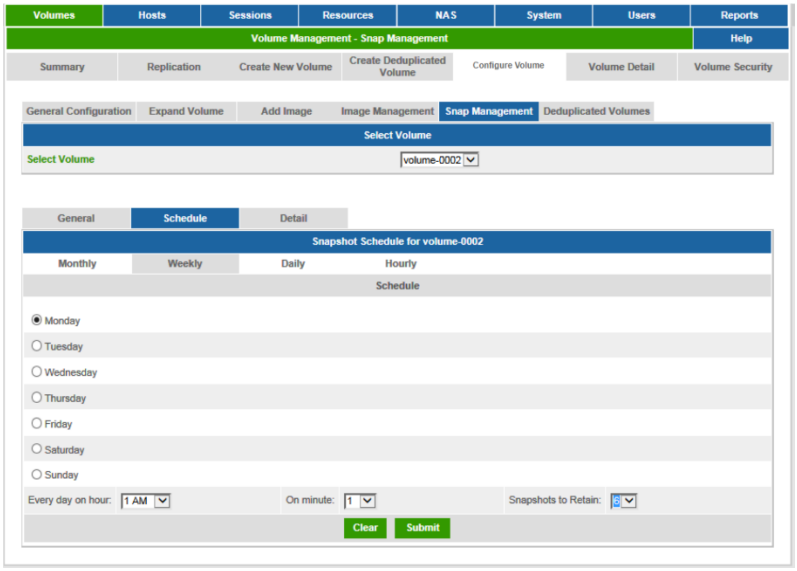


Figure 6-8 Schedule Weekly Snapshots

4 Select the checkboxes for the day desired for the weekly Snapshot. Then select the hour and minute desired using the drop down menus. It is recommended that you do not select the same time as any other scheduled Snapshot.

5 Select the number of weekly Snapshots to retain. The maximum number of retained weekly Snapshots is 6. Once the number of weekly retained Snapshots has been reached, the system will automatically start deleting Snapshots to make room for new Snapshots.

6 Click **Submit.** The Weekly Snapshots will then be taken at the days and hours specified.

Monthly Snapshots

To schedule a Monthly Snapshot:

1 From the Volume Configuration screen, click **Snap Management**.

2 From the Snap Management screen Select **Schedule**.

3 From the Snapshot Schedule, select **Monthly**.

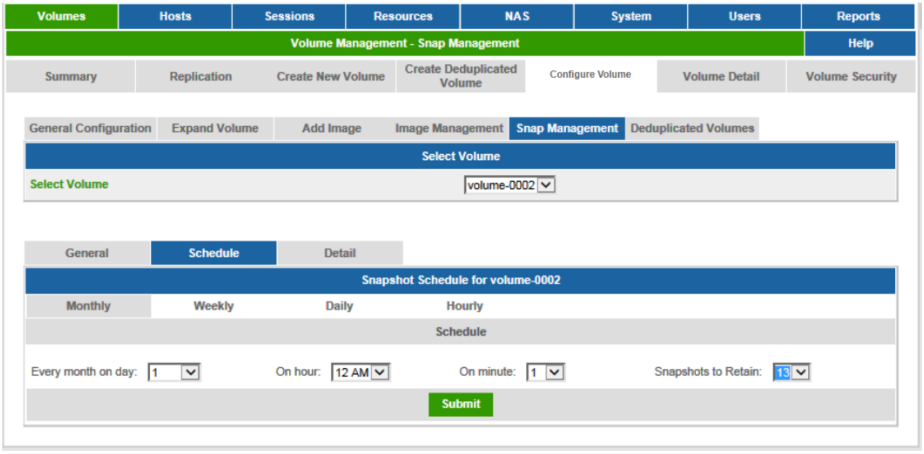


Figure 6-9 Schedule Monthly Snapshots

4 Select the day of the month desired from the dropdown menu. Then select the hour and minute desired using the drop down menus. It is recommended that you do not select the same time as any other scheduled Snapshot.

When you select a day for scheduled monthly snapshots, be sure to pick a day that is available all months. Not all months have all days (e.g. February normally has only 28).

5 Select the number of monthly Snapshots to retain. The maximum number of retained monthly Snapshots is 13. Once the number of monthly retained Snapshots has been reached, the system will automatically start deleting Snapshots to make room for new Snapshots, if you set up your live volume to **Preserve Live Volume** (recommended).

6 Click **Submit.** The monthly Snapshots will then be taken at the days and hours specified.

6.3 Managing Snapshots and Snapspace

6.3.1 5.3.1 Increasing the Size of the Snapspace

The process to increase the size of the Snapspace is one method to allow more snapshots to be preserved in the system. There must be available space on one of the resources in the system before using this function. In order to enhance the availability of the snapshot volume, place the new allocation for the Snapspace on the same resource as the originally allocated space. Please refer to the section titled ”Expanding the Snapspace of a Snapshot Live Volume” under Volume Configuration.

6.3.2 Delete Snapshots

This function allows the user to remove a Snapshot image from a Snap Volume and free the resources that previously stored the Snapshot. The Snapshot will not be accessible to a host after it is removed.

To delete a Snapshot:

1. From the *Volume Configuration* screen, click **Snap Management**.
2. Click on **Detail.**
3. From the Delete Snapshot column of the Snapshot detail table, select the image that you want to delete.
4. Click **Submit** to delete the Snapshot.

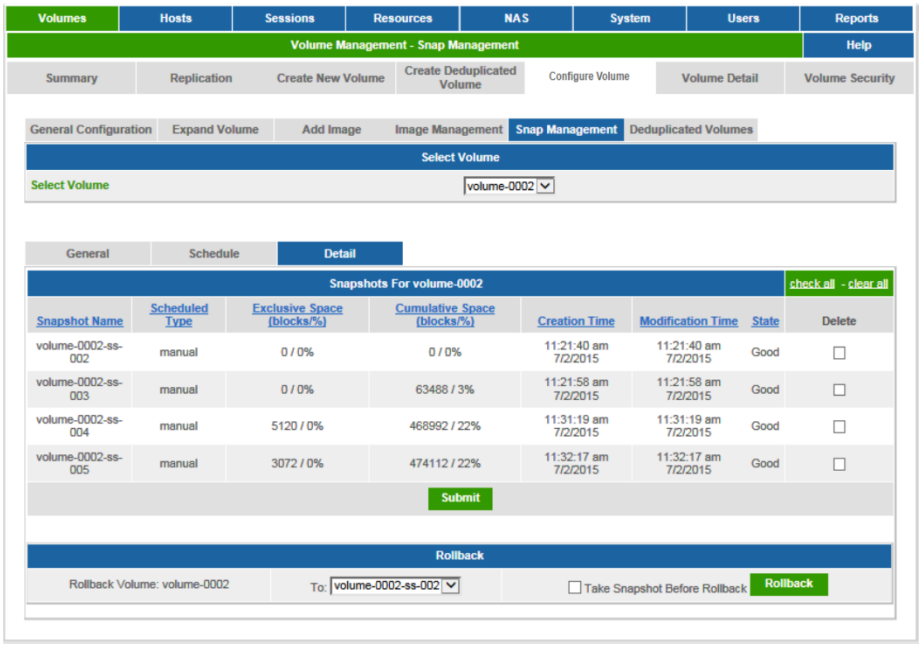


Figure 6-10 Snap Management Detail Screen

6.3.3 Deleting All Snapshots and Snapspace

Deleting all Snapshots and the Snapspace will revert the volume to a normal span volume. This may be necessary to expand the volume or because Snapshots are no longer necessary for the volume.

To delete all Snapshots and the Snapspace:

1 From the *Volume Configuration* screen, click **Snap Management**. The General screen is displayed:

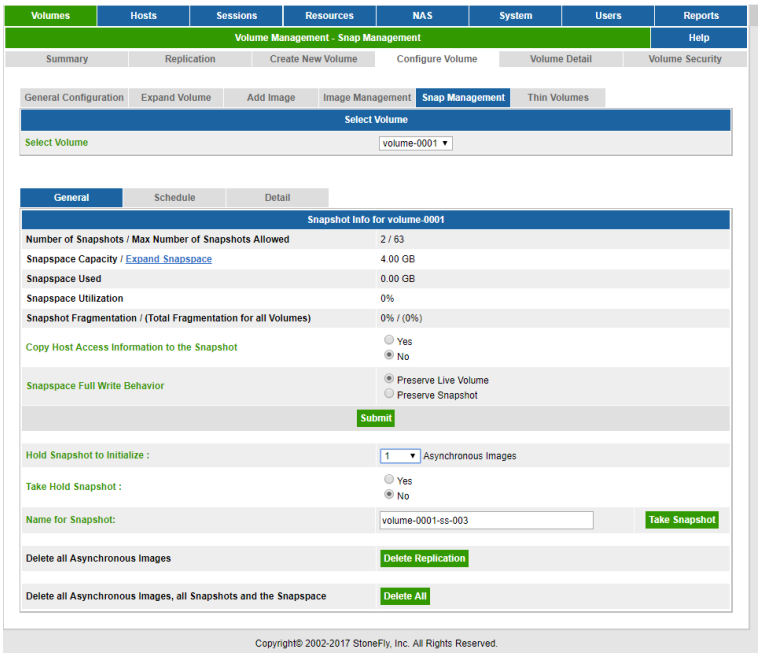
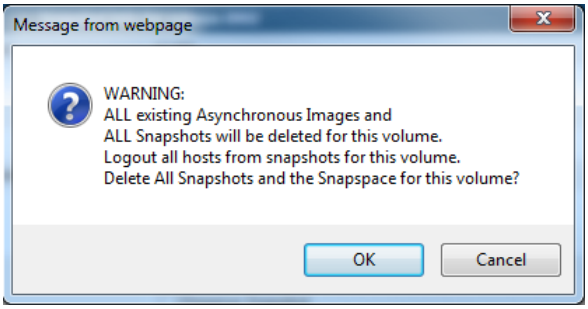


Figure 6-11 Snap Management General Screen

2 Select the desired volume in the pull-down list **Select Volume**. The **General** screen re-displays with the proper contents.

3 Click **Delete All** next to "Delete all asynchronous Images, all Snapshots and the Snapspace".

4 The Storage Concentrator will display the following warning.



5 Click **OK** to delete all the asynchronous Images, Snapshots and Snapspace or **Cancel** if you do not want to revert the volume to a span volume

If the Snapshot Live Volume was also mirrored, selecting this button deletes all the Snapshots and Snapspace and reverts the volume to a mirror volume rather than a span volume.

6.3.4 Rollback

Rollback is a disaster recovery feature that allows users to return to the last known "good" data point in the event of corruption or viruses. Essentially, performing a rollback returns the Live Volume to a previous known state. All writes that have been completed since the Snapshot was taken are overwritten so that the Live Volume appears just as it was when the Snapshot was taken.

The primary application for Rollback is to return to an earlier point-in-time to try and restore the volume in the event of corruption or a virus.

Special Considerations When Using Rollback:

1 If you write to a Snapshot and then use that Snapshot for Rollback, the Live Volume will also have the changes made to the Snapshot.

2 You can rollback multiple times to get the desired result. For example, if you rollback to find a lost file and the file isn't there on the first Snapshot used, you can rollback again using an older Snapshot to locate the item required. This only works if you rollback in reverse time order.

3 You will lose new data by rolling back to an earlier point-in-time. Therefore, unless a Live Volume is really corrupted or there is a virus that would make containment more difficult, it is recommended to make sure there is a backup of the Live Volume before rollback. One way to accomplish this is to mirror the live volume and then promote the mirror before rollback.

4 You will need to log out of all host sessions to the Live Volume before performing the rollback. The system will automatically log out all sessions if you do not do so beforehand.

It is highly advisable to take a snapshot of the Live Volume immediately prior to starting the rollback process. That way, if the result from the rollback is not what you expected, you can revert to the prior state using the snapshot taken immediately before the rollback.

To rollback the Live Volume using a Snapshot:

1 It is recommended that you log out all host sessions before starting the rollback process.

The system will automatically log out all sessions if you do not do so beforehand.

2 From the Volume Configuration screen, click **Snap Management**.

3 Click **Detail**

4 The Snap Management Detail screen appears

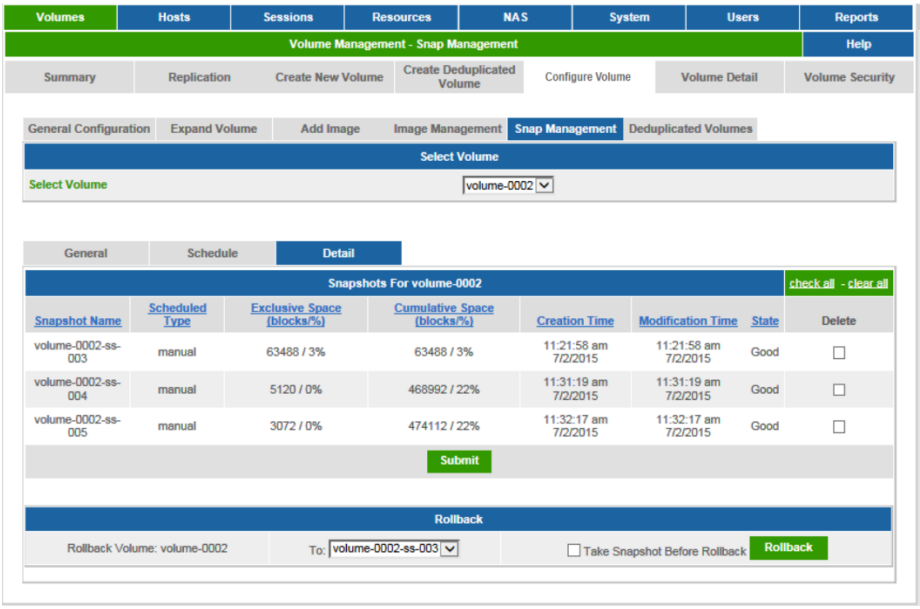
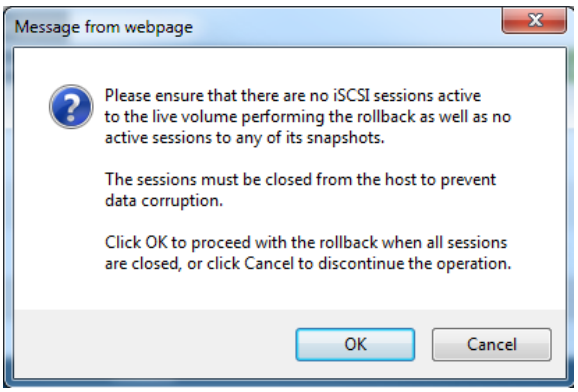


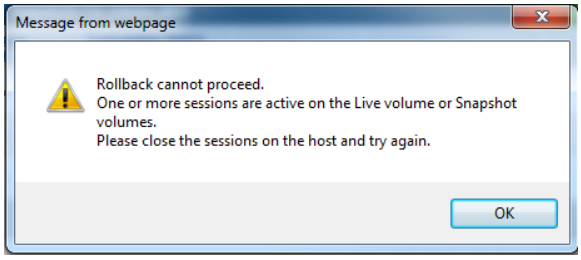
Figure 6-12 Snap Management Detail Screen

5 To rollback the Live Volume, select the Snapshot of the point-in-time that you want using the pull down menu. If you would like a Snapshot taken immediately before the rollback, select the checkbox for "Take Snapshot Before Rollback". If you do not wish to take a Snapshot before rollback, then deselect this option.

6 Click **Rollback.** The system displays the following warning message:



7 If you have already logged out all host sessions. Click **OK.** If you have not, Click **Cancel** to exit and log out all sessions. If you click **OK** and there are still sessions logged in, the system will display the following error message.



8 Once all sessions are logged out, the system will start the rollback process. Depending on how large the Live Volume is and how many changes there were since the time the Snapshot was taken, it may take some time to complete the rollback process.

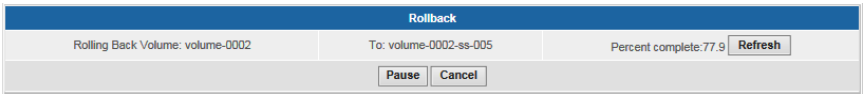


Figure 6-13 Rollback Progress Message

9 To view the percentage complete, click on the **Refresh** button.

10 If you need to cancel or pause the rollback process, click the **Cancel** or **Pause** button.

Cancel will leave the Live Volume in an unknown state and is inadvisable unless there is a serious hardware issue and the rollback cannot proceed. Use Cancel to abort the process if some problem occurs or want to restart the rollback. Pause will merely halt the rollback process.

If the result from the rollback is not what you expected, you can revert to the prior state using the Snapshot taken immediately before the rollback.

6.3.5 Mirroring the Live Volume and Snapspace

StoneFly Mirroring can mirror Snap-enabled Volumes and Snapspaces, but not Snapshot Images. (See "To add a Mirror Image")

6.3.6 Expanding a Snap-enabled Volume

Expansion of a Snap-enabled Volume is possible. See "Expanding a Volume".

If expanding a Snap-enabled Volume makes the Snapspace too small to be efficient, expand the Snapspace to maintain its relative size of 20-30% of the Live Volume.

This page is intentionally left blank.

Chapter 7
Thin Volumes Management

There are two aspects of thin volume management. The first one is thin pool management for the specific thin volume it belongs to. The second aspect is management of the thin volume itself. In both cases, the main focus is how much of the thin pool space is already allocated by volume data and by supporting metadata. Thin volumes have to be put offline if the pools are not able to provide space to handle the new data. This is the reason why this page is oriented to report current thin pool space usage and lets the user set up different alarms when the usage reaches the specific limit.

Thin pool space is split between two segments. The metadata segment stores configuration information about volumes provisioned in the pool. The data segment provides space to keep volume data. The data segment usage increases each time the system writes to unallocated blocks. Size of the metadata also increases because the metadata segment keeps all of the mapping information between thin volume blocks and the new blocks assigned to the volume at the data segment.

The newly created thin volume does not allocate data and metadata segments, but the pool does reserve some number of metadata blocks to handle future writes into the volume. These blocks are not assigned to any specific thin volume and are used only to check how many volumes the pool can handle in addition to the existing volumes.

The total size of the allocated metadata from all thin pools and all snap enabled volumes is limited by the system RAM. The metadata is stored on local resources but at run time a copy of this information must be located in the system RAM so that the Storage Concentrator can use it to handle IOs. For performance reasons, metadata should use only physically available RAM and avoid the usage of virtual memory. Another limitation on metadata size is the requirement to be consistent with cluster failover. The new Primary Storage Concentrator has to read all of the metadata from the resources before it can start to handle host IOs. Delays above some limit are treated by iSCSI hosts as an IO failure and are reported to applications.

The available system RAM is checked each time a new thin pool or new thin volume is created or expanded. If the system can handle the request it proceeds with the execution. If the RAM usage is close to the limit or will exceed the limit, then the appropriate recommendations are made so that the user can adjust the request or make some other system configuration changes.

System memory usage checks are performed on a periodical basis. If the memory usage is close to the limit, the Storage Concentrator can delete some snapshots automatically. If all snapshots are already deleted but the memory usage is still above the threshold, then a GUI message is generated. The message advises the user to handle this problem by checking thin volumes for potential deletion or for modification back to thick volumes.

7.1 Thin Pool Management

Select the pool from the pull-down menu to manage the specific pool (No cluster):

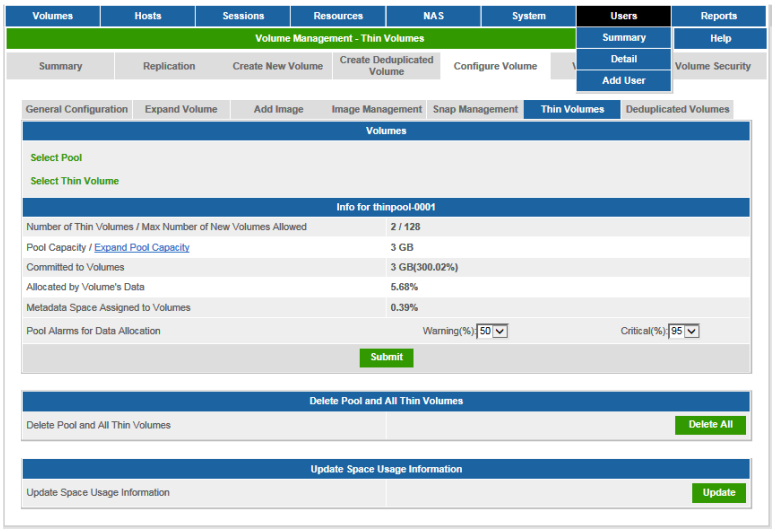


Figure 7-1 Thin Volume Management Screen

The following fields are presented for each thin pool:

Number of Thin Volumes – Thin volumes have been provisioned in the pool.

Number of New Volumes Allowed – Maximum number of thin volumes that can be created in the pool in addition to the existing ones.

Pool Capacity – Size of the pool space in GigaBytes.

Committed to Volumes – Amount of space that all thin volumes are subscribed for in GigaBytes and in percent of pool space data segment.

Allocated by Volume's Data – Percent of pool space data segment allocated by thin volume data.

Metadata Space Assigned to Volumes – Percent of pool metadata blocks assigned to volumes with data.

Assigned and Reserved – Percent of pool metadata blocks assigned to volumes with data or reserved for volumes without data.

Pool Alarms for Data Allocation – Threshold for data segment usage to generate GUI log message with warning or critical severity levels.

Load Balancing – Load Balancing for the pool (in a cluster). Thin volumes provisioned in the pool inherit the Load Balancing settings from the pool. To change Load balancing see "Adaptive Load Balancing (ALB)".

Use the "Update" button to update the screen with the most recent information about thin pool space usage. Change alarm levels by using pull-down menus and hit the

"Submit" button to enforce the changes. Use the "Delete All" button to delete the pool and all allocated thin volumes. The "Delete All" option erases data from all thin volumes. There is no procedure to restore these volumes. Be very cautious with this option. The delete procedure from the Volume Summary screen has to be used in most cases (see "Removing Volumes").

7.2 Thin Volume Status

Select the volume from the pull-down menu to manage the specific volume from the selected pool (No cluster):

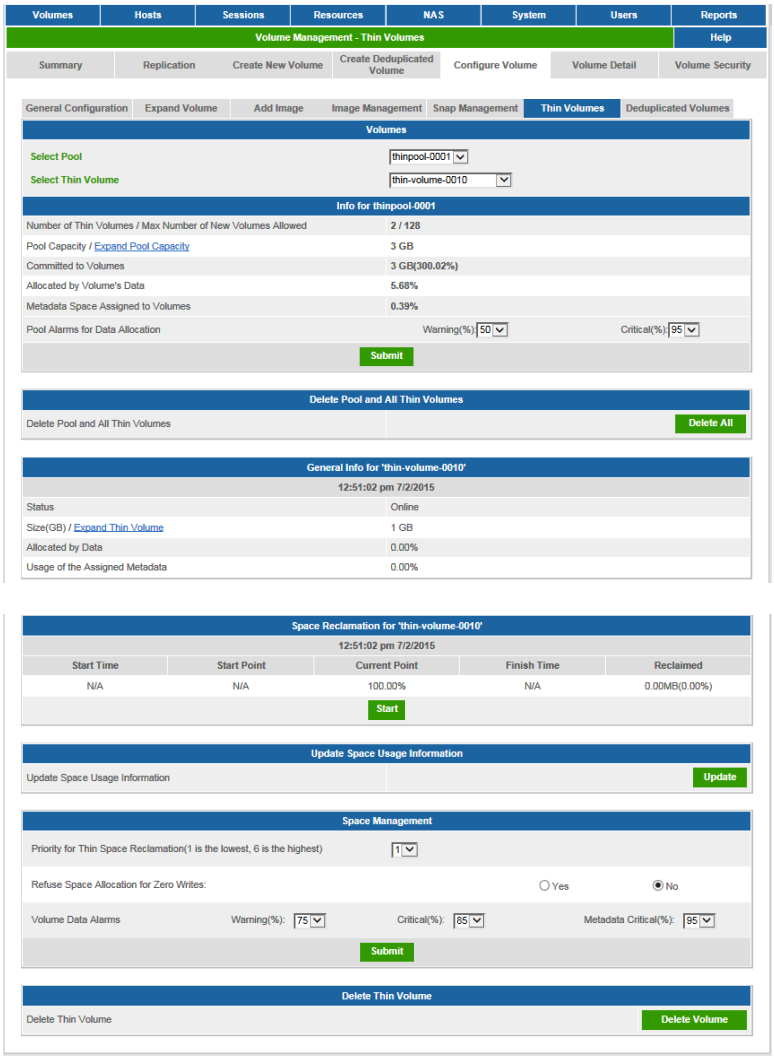


Figure 7-2 Volume Management, Thin Volumes Screen

The status options for thin volumes are "Online", "Offline", "Non Active, Data Full", "Non Active, Meta Full", or "Non Active, Meta Max".

Use the Volume Configuration General Configuration screen to handle "Offline" and "Online" states of the volume, see "Volume Configuration".

The volume is at the "Non Active, Data Full" status, if it failed to allocate new data segments from the pool. When the volume failed to get additional segments to store metadata, it has to be put into "Non Active, Meta Full" state. If the volume reached the metadata allocation limit, the volume state has to be "Non Active, Meta Max". In all these "Non Active" cases, there are no active iSCSI sessions to the volume and new iSCSI logins are rejected. The user has to find a way to provide the pool with additional

resources. Then the "Activate" button can be used to put volume online.

The “Activate” button is only exposed for thin volumes that are not able to handle write IOs because of the exhaustion of data or metadata segments needed to get new allocation. The button has to be used after some measures are taken. These measures include thin pool expansion, deletion of some thin volumes, transferring some thin volumes to thick volumes, or execution of the thin space reclamation procedure.

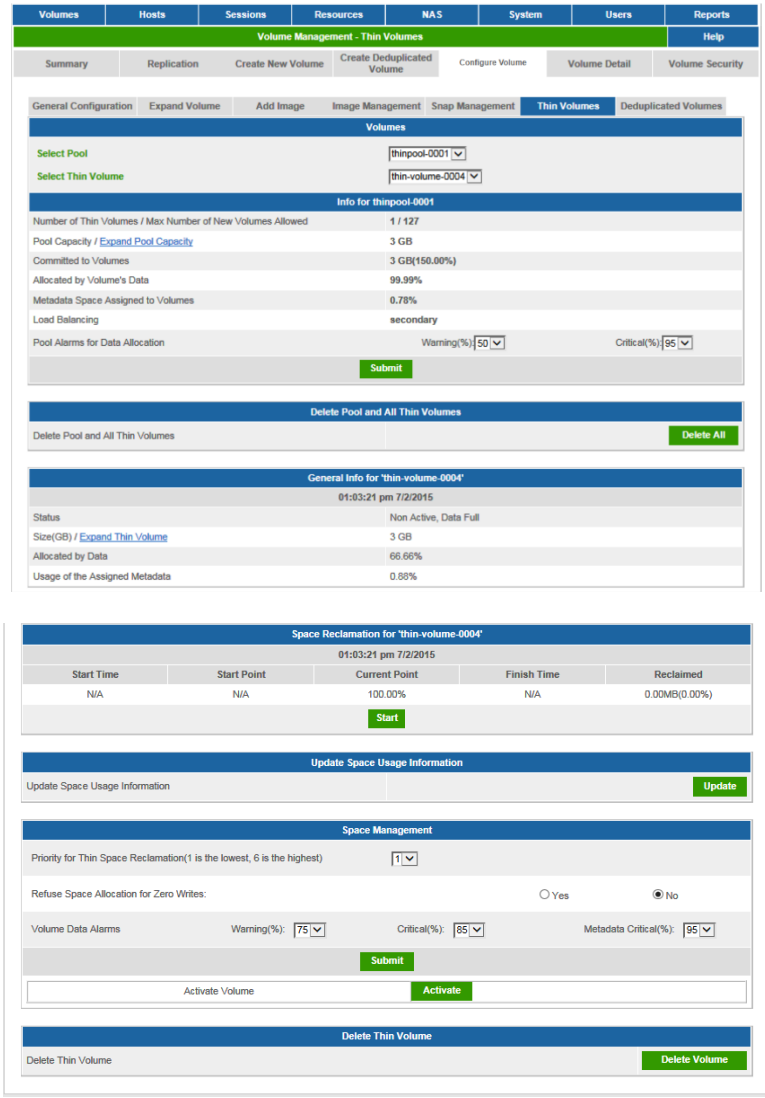


Figure 7-3 Thin Volume Management screen for a thin pool that is out of free data segments (in a cluster).

7.3 Thin Volume Space Reclamation

The space allocated to a Thin Volume can only increase, and will never decrease unless the special reclamation procedure is applied to the volume. The reclamation procedure consists of steps that have to be executed on the host that mounted the Thin Volume and steps that have to be executed on the Storage Concentrator where the volume is provisioned. Reclamation has to be initiated manually. The procedure is only effective in cases where the Thin Volume has allocated space that is not used by the host's file system at the time when the reclamation is executed. The reclamation has to release this space to the Pool and provides available space to support new writes. Users should not rely on space reclamation as an ultimate tool and should become accustomed with the other options available that allow the thin volume to be kept online.

Special attention must be paid to Thin Volume and Pool space utilizations. Thin Pool over-subscription should be avoided as once the Pool capacity is reached, new writes cannot occur until one or more Thin Volumes that are off-line are copied out of the Pool and then deleted from it.

Host applications and procedures that tend to fully thicken a Thin Volume should be avoided since the Thin Volume's unused space cannot later be reclaimed.

The next options are available only in cases where the pool is close to free space exhaustion or is already out of free space:

The pool can be expanded by using available (not used) storage resources. See "Expanding a Volume". New storage resources can be added to the system to provide this expansion. See "Adding Resources";

Pool space used by some thin volumes can be released by deleting these volumes from the pool. Thin volumes left in the pool will reclaim the storage at run time.

Thin volumes with allocation sizes above some limit can be transferred into regular volumes. The Storage Concentrator provides an offline procedure for this transformation. See "Copy Volume". The original thin volume has to be deleted after the copy is executed successfully.

Thin volumes can be copied to regular volumes by using any host-based disk copy utility. The original thin volume has to be deleted after the copy is executed successfully.

Thin volumes can be copied to a newly created thin volume from the same pool by using host-based file-to-file partition copy utilities. The original thin volume has to be deleted after the copy is executed successfully. Thin volumes used for the copy should not have any allocated blocks with unused information because the copy utility transfers only the currently stored files. If the size of the reclaimed space is not large enough then other options will have to be considered.

7.3.1 Space Reclamation for Thin Volumes Mounted on Microsoft Windows

1 On the SC, go to the "Volume Configuration->Thin Volumes" management screen for the specific Thin Volume, and set the "Refuse Space Allocation for Zero Writes" to "Yes", and hit the "Submit" button.

2 If the Thin Volume is already at 100% utilization and is offline for host access, then run an initial "Space Reclamation" by hitting the "Start" button to free up enough space so that the volume can be changed to an online status and the host can perform some overhead writes.

The reclaim process does not need to complete, but verify that some space has been reclaimed using the "Update" button, then stop the process since it will be repeated again below.

The host iSCSI session will need to be re-established, and a utility 'chkdsk' may be needed.

Note that in this total space exhaustion case, all other host applications performing I/O to the same volume should be stopped until the unused space is zeroed and reclaimed.

3 Use utility 'sdelete.exe' on the host from the Microsoft System Internals with the "-p 1 -c" options on the drive letter for the Thin Volume:

sdelete.exe -p 1 -c f:

The 'sdelete.exe' utility can be downloaded from the www.sysinternals.com.

4 On the SC GUI "Volume Configuration->Thin Volumes" management screen for the specific Thin Volume, start a reclamation with the "Start" button, and wait for it to finish. Check the status with the "Update" button.

5 On the SC, go to the "Volume Configuration -> Thin Volumes" management screen for the specific Thin Volume and restore the "Refuse Space Allocation for Zero Writes" back to its normal setting of "No", and then hit the "Submit" button.

Note that the SC and host will operate normally with the setting left at "Yes", but there is a performance penalty in doing so.

7.3.2 Space Reclamation for Thin Volumes Mounted on Linux

1 On the SC, go to the "Volume Configuration->Thin Volumes" management screen for the specific Thin Volume and set the "Refuse Space Allocation for Zero Writes" to "Yes", and hit the "Submit" button.

2 If the Thin Volume is already at 100% utilization and is offline for host access, then run an initial "Space Reclamation" by hitting the "Start" button to free up enough space so that the volume can be changed to an online status and the host can perform some overhead writes.

The reclaim process does not need to complete, but verify that some space has been reclaimed using the "Update" button, then stop the process since it will be repeated again below.

The host iSCSI session will need to be re-established, and a utility 'fsck' may be needed before the file-system can be mounted.

Note that in this total space exhaustion case, all other host applications performing I/O to the same volume should be stopped until the unused space is zeroed and reclaimed.

3 On the host, we will use 'dd' to create a new large file that is entirely zeroed until the entire file-system becomes full. In this example, '/thin-test/' is the mount point for the Linux file-system on the SC Thin Volume. As the root user, create a temporary "all zeros" file on the r/w mounted file-system that is being reclaimed:

```
# dd bs=1M if=/dev/zero of=/thin-test/zeroing-file dd:

writing `/thin-test/zeroing-file': No space left on

device 934+0 records in 933+0 records out
```

Perform the 'sync' command at the end:

```
# sync
```

The "No space left on device" error is expected. If it does not appear, it may be that the maximum file size for a single file is less than the amount of free space on the file system. In that case, multiple "zeroing files" will need to be created until the "No space left on device" error is seen.

4 Delete all of the temporary "zeroing files" that were just created.

```
# rm /thin-test/zeroing-file rm: remove regular file

`/thin-test/zeroing-file'? y
```

5 On the SC GUI "Volume Configuration->Thin Volumes" management screen for the specific Thin Volume, start a reclamation with the "Start" button, and wait for it to finish. Check the status with the "Update" button.

6 On the SC, go to the "Volume Configuration -> Thin Volumes"

management screen for the specific Thin Volume and restore the "Refuse Space Allocation for Zero Writes" back to its normal setting of "No", and then hit the "Submit" button.

Note that the SC and host will operate normally with the setting left at "Yes", but there is a performance penalty in doing so.

7.4 Thin Friendly Hosts and Applications

The user should avoid doing host operations that cause writes to most or all of the thin volume's blocks because this would cause allocation of storage blocks and the volume would no longer be thin.

For example, a full file-system format, defragmentation, or security erase process may update most of the blocks of the volume. While this is fine with a regular volume, this would require a thin volume to allocate all allocated space. The benefits of thin provisioning disappear but the user has to continue to pay a penalty of slightly lower performance than what is usually associated with thin volumes.

Another example is system-restore, recycle bin, and host volume shadow copies where previously deleted file versions are retained for later recovery. This is fine if it is configured so that the maximum amount spaced used for the old contents is limited. However, if it is allowed to use the entire size of the thin volume, use of these features might be considered inconsistent with thin volumes.

Thin volumes should not be used as a boot or swap disk because the thin volume can then be placed offline if the thin pool is out of data segments needed to write into an unallocated range of volume blocks. The thin pool exhaustion may be caused by any of the volumes in the pool if over-subscription is used.

Operating Systems and applications can be considered "thin friendly" when:

The file systems efficiently reuse already allocated space (from previously deleted files) instead of using new storage blocks.

Applications reserve disk space but do not initialize it.

Data is not moved from one disk location to another previously unused location.

Thin friendly operations tend to avoid unnecessary new writes, and reuse blocks that are already allocated. Systems that use new blocks can thicken the volume beyond available pool space. When over-subscription is used, a volume will be unavailable for new writes in spite of the fact that from OS perspective it has available space.

Test thin volumes can be created to test OS and application handling of the volume. Monitor thin volume usage with the Storage Concentrator Volume Configuration Thin Volume Status page. If the test keeps volume allocation size within limits, the appropriate thin volume can be created to accommodate production.

7.4.1 Microsoft Windows Thin Volume Usage Guidelines

Please consult the following recommendations:

Guidelines for thinly provisioned LUNs in Windows Server 2008 R2, Windows 7, Windows Server 2008, Windows Vista, and Windows Server 2003
http://support.microsoft.com/kb/959613

7.4.2 VMware Thin Volume Usage Guidelines

Note that VMware systems provide thin provisioning features at the host level as well as operating with Storage Concentrator storage system thin volumes. Both features can be used independently, or together. It is important to be clear about which feature location is being discussed in the following references.

Please consult the following recommendations:

Dynamic Storage Provisioning - Considerations and Best Practices for Using Virtual Disk Thin Provisioning
http://www.vmware.com/files/pdf/VMware-DynamicStorageProv-WP-EN.pdf

Using thin provisioned disks with virtual machines
http://kb.vmware.com/selfservice/microsites/search.do?language=en_US&cmd=displayKC&externalId=1005418

VMware KB: Issues with storage related commands when using thin provisioned LUNs with your storage array
http://kb.vmware.com/selfservice/microsites/search.do?language=en_US&cmd=displayKC&externalId=1003826

7.4.3 Linux Thin Volume Usage Guidelines

To be determined.

Chapter 8
Deduplicated Volumes Management

The main focus of management for deduplicated volumes has to be monitoring of how much of the deduplicated pool space is already allocated by volume data and by supporting metadata. Deduplicated volumes have to be put offline if the pools are not able to provide space to handle the new data. This is the reason why this page is oriented to report current deduplicated pool space usage and lets the user set up different alarms when the usage reaches the specific limit.

8.1 Deduplicated Pool Management

Select the pool from the pull-down menu to manage the specific pool (in a cluster):

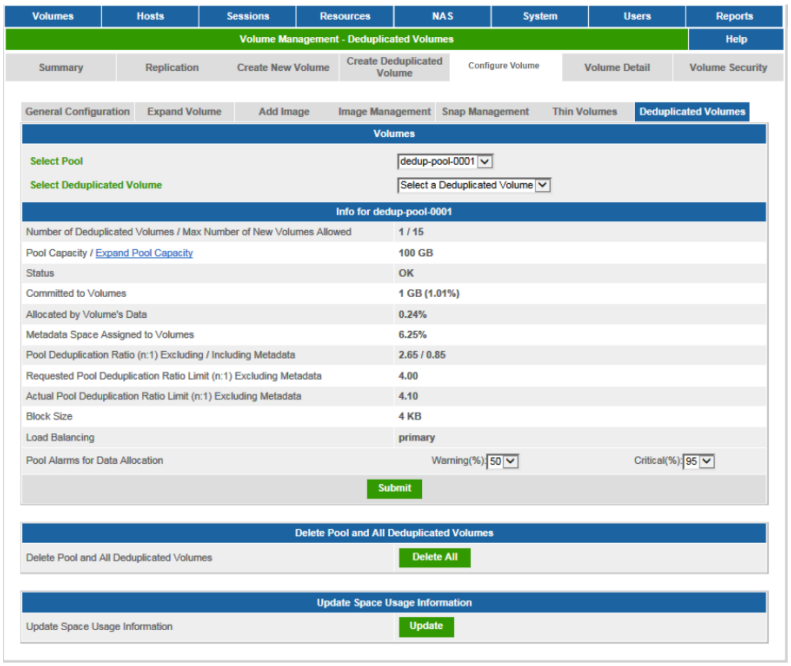


Figure 8-1 Deduplicated Volumes Management Screen

The following fields are presented for each deduplicated pool:

Number of Deduplicated Volumes – Number of deduplicated volumes provisioned in the pool.

Number of New Volumes Allowed – Maximum number of deduplicated volumes that can be created in the pool in addition to the existing ones. The number is defined by how many metadata blocks are available in the pool to support creation of new deduplicated volumes.

Pool Capacity – The size of the pool space in gigabytes, used for both volume data blocks, and pool and per-volume metadata.

Status - OK indicates that the pool is fully operational. Offline indicates the pool is not operational due to failure or because it set "Offline" by user. Additional indicators show that the pool is in process of execution of specific asynchronous commands. If one of these indicators is present the deduplication pool is blocked from accepting additional commands from GUI until the current asynchronous command is finished. Due to significant size of deduplication metadata that asynchronous command has to handle it could take tens of minutes. The next Deduplicated Pool commands are executed in asynchronous mode: provision deduplicated pool, format deduplicated space, delete deduplicated volume and calculate deduplication ratio for volume. Click on button "Update" to refresh the Status field.

Committed to Volumes –Amount of space that all deduplicated volumes are subscribed to in gigabytes and in percent of the pool space data blocks. When the pool is over-

committed, the size will be greater than the size of the pool, and the percentage greater than 100%.

Allocated by Volume's Data – Percent of pool space data blocks allocated by data that belongs to all deduplicated volumes. This is the pool space consumed and approaches, but does not exceed 100%, with warnings issued as it does so.

Metadata Space Assigned to Volumes – Percent of pool metadata blocks assigned to volumes with data. This value and approaches, but does not exceed 100%, with warnings issued as it does so.

Pool Deduplication Ratio - Indication of how effective deduplication is for the pool in total. A higher number represents the better deduplication.

The ratio between the number of data blocks in all deduplicated volumes that have reference to data blocks allocated in the pool, shared both internally within each volume, and between different volumes in the pool, and the total number of allocated data blocks in the pool, excluding pool blocks used for metadata.

The ratio between the number of data blocks in all deduplicated volumes that have reference to data blocks allocated in the pool, shared both internally within each volume, and externally between different volumes in the pool, and the total number of allocated data blocks in the pool, including pool blocks used for metadata.

Requested Pool Deduplication Ratio Limit - This shows the user requested limit for the best case maximum deduplication ratio for all volumes in the pool when it is completely full. This is a limit placed on the deduplication ratio, not a measured ratio. This parameter was used to size the amount of meta-data used to map the volume disk addresses to deduplication blocks. Too low of a value would exhaust the pools meta-data blocks before all of the pool data blocks are consumed. Too high of a value would waste pool and system resources and significantly limit the overall system configuration for an overly optimistic deduplication ratio unachievable with the given volume(s) data.

Actual Pool Deduplication Ratio Limit - This shows the actual limit for the best case maximum deduplication ratio for all volumes in the pool when it is completely full. This is a limit placed on the deduplication ratio, not a measured ratio. This parameter was used to size the amount of meta-data used to map the volume disk addresses to deduplication blocks. Since the actual limit may not be the same as the user requested limit, both are shown.

Block Size – The size in bytes of deduplication blocks that are allocated from pool. This is not the SCSI I/O block size, which is defined by the of storage resources where pool is provisioned. The deduplication block size represents the smallest amount of data that can be allocated or shared in the pool.

Pool Alarms for Data Allocation – Threshold for pool data blocks usage to generate GUI log message with warning or critical severity level. Used to monitor the pool's data blocks usage to avoid situation when pool is out of free data blocks. When pool's data blocks are exhausted, newly allocating writes will be blocked, and cause the volume to go offline until the pool is expanded, or data blocks freed.

Load Balancing – Load Balancing for the pool (in a cluster). The setting is inherited by any deduplicated volume provisioned in the pool. For more details or to modify the settings see "Adaptive Load Balancing (ALB)".

Use the "Update" button to update the screen with the most recent information about deduplicated pool space usage. Change alarm levels by using pull-down menus and hit the "Submit" button to enforce the changes. Use the "Delete All" button to delete the pool and all allocated deduplicated volumes. The "Delete All" option erases data from all deduplicated volumes. There is no procedure to restore these volumes. Be very cautious with this option. The delete procedure from the Volume Summary screen has to be used in most cases (see "Removing Volumes").

8.2 Deduplicated Volume Status

Select the volume from the pull-down menu to manage the specific volume from the selected pool (in a cluster):

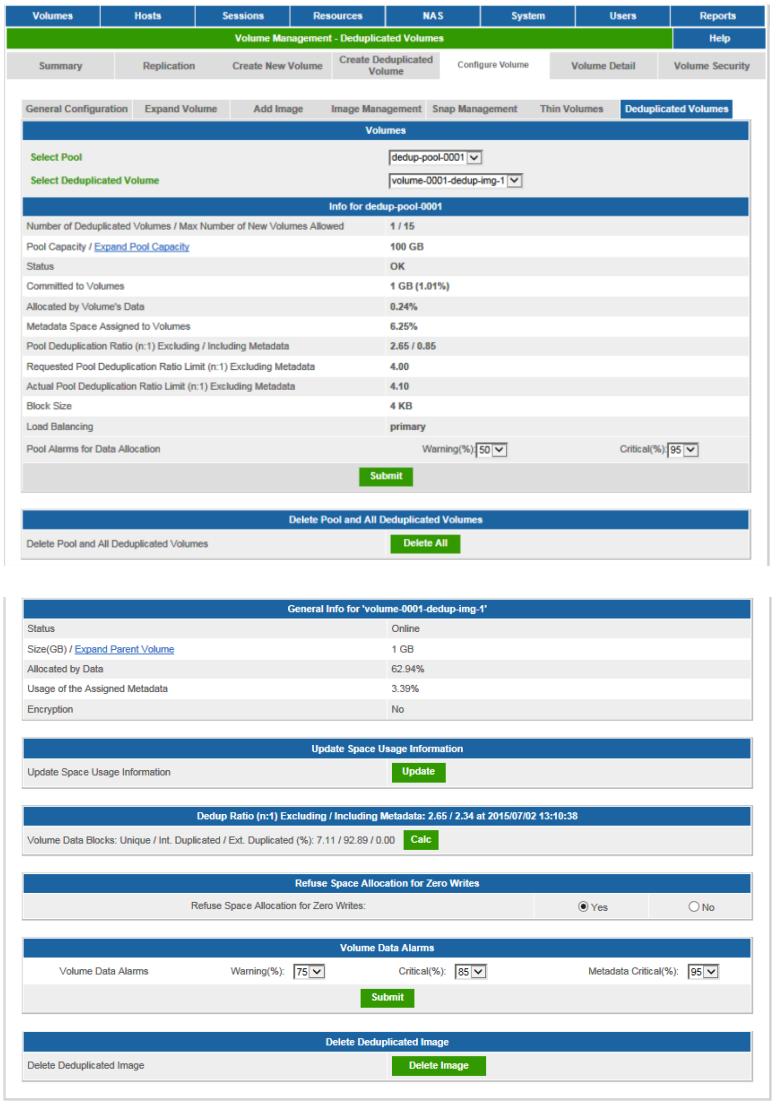


Figure 8-2 Volume Management, Deduplicated Volumes Screen

Status - Current status of the volume. The status options are "Online", "Offline", "Non Active, Data Full", "Non Active, Meta Full", "Non Active, Meta Max", "Provisioning", "Calc.Dedup.Ratio" or "Deleting Volume".

The volume is in the "Non Active, Data Full" status if it has failed to allocate new data blocks from the pool. When a volume fails to get additional deduplication blocks to store metadata, it is in the "Non Active, Meta Full" state. If the volume reaches its metadata allocation limit the volume state is "Non Active, Meta Max". In both cases these "Non Active" cases there are no active iSCSI sessions to the volume and new iSCSI logins are rejected. The user must provide pool with additional resources, or free existing ones. The volume can be placed online with the "Activate" button.

Use the Volume Configuration General Configuration page to handle "Offline" state of the volume., see "Volume Configuration".

The volume states "Provisioning", "Calc.Dedup.Ratio" or "Deleting Volume" are indicators that the deduplicated pool is still in process of executing the specific asynchronous command. Use the update button to refresh status and other fields.

Size - The advertised size of the volume in gigabytes. The amount of space used from the pool will be generally be less than this size due to thin provisioning, and internal and external deduplication block sharing with duplicates in the volume, and other volumes in the pool.

Allocated by Data - Percent of the advertised volume space that are used, having reference to data blocks allocated at pool. This is indicator of how much volume space is allocated by data. It includes data blocks for files that are used now and for files that are deleted but not cleared by the host system. This indicator is equivalent of the same indicator used by thin volumes. Volumes with high percent of allocated data and low deduplication rate should be first candidates for transformation to regular volumes.

Usage of the Assigned Metadata - Percent of metadata blocks assigned to the volume that are in use. After volume reaches 100% of metadata block usage it can get additional free metadata blocks from the pool so the volume can continue handle IO traffic without interruption, up to a limit. The critical value for this indicator occurs when pool does not have any more free metadata blocks to give to the volume, and the usage is high.

Encryption - "Yes" if the volume is encrypted. Note that Volume encryption makes unlikely to find deduplicated blocks between encrypted volume and other volumes in the pool. Deduplication for encrypted volumes would be mostly internal to the volume.

Dedup Ratio - The ratio between the number of volume data blocks that have reference to data blocks allocated in the pool and the total number of allocated data blocks to the volume in the pool, excluding blocks assigned to volume and used for metadata.

The ratio between the number of volume data blocks that have reference to data blocks allocated in the pool and the total number of allocated data blocks to the volume in the pool, including volume blocks assigned to volume and used for metadata.

This is an indication of internal deduplication for the volume. The ratio does not show information how data blocks of the volume are shared with other volumes in the pool.

Volume Data Blocks - Displays three indicators.

The first one is indicator of number of unique blocks in the volume. It shows how much of volume's allocated data are specific to the volume and are not shared internally and externally with other volumes in the pool.

The second one has name "Int. Duplicated". It shows how much of volume's allocated data are duplicated internally and do not shared with other volumes in the pool.

The last one is "Ext. Duplicated" and it shows how much of volume's allocated data are shared with other volumes in the pool.

The indicators let user understand what is the main contributer to the volume deduplication. The volume deduplication ratio is general indicator and has to be used in combination with these three indicators.

Click button “Calc” to request the calculation of the volumes Dedup Ratio. Since it may take time for the calculation to complete, the new values may not appear until the page is refreshed later with the update button.

Refuse Space Allocation for Zero Writes - When "Yes", all zero deduplication blocks are not stored, and do not consume pool data or metadata resources. When "No", all zero blocks are stored. Blocks that have never been written are assigned the value of all zeros. Must be set to "Yes" when executing space reclamation on the volume. See user's guide for details. The recommended setting is "Yes".

There is no sense to check for zero contents in blocks that have to be written into encrypted deduplicated volumes because the blocks are already encrypted so probability to find zero contents is very low. For this reason the feature to refuse space allocation is disabled for encrypted volumes.

Volume Data Alarms - Threshold to monitor volume data space allocation and generate GUI log messages with warning or critical severity levels. This indication is used in conjunction with other volume deduplication indicators because it shows the volume thinness only, not the pool over all. These alarms would be a trigger to perform volume space reclamation at the host to release blocks from deleted files.

Threshold to monitor volume metadata space utilization and generate GUI log messages with warning or critical severity levels. This alarm can go "on" and back "off" as the volume gets new portion of free metadata blocks from the pool. The user should check the metadata blocks availability at the pool level when receiving this alarm.

The “Activate” button is only exposed for deduplicated volumes that were not able to handle a write IO because of the exhaustion of data or metadata blocks needed to handle a new allocation. The button is used after restorative measures are taken. These measures include deduplicated pool expansion, transferring some deduplicated volumes to thick volumes, execution of deduplicated space reclamation procedure, or the deletion of some unneeded deduplicated volumes.

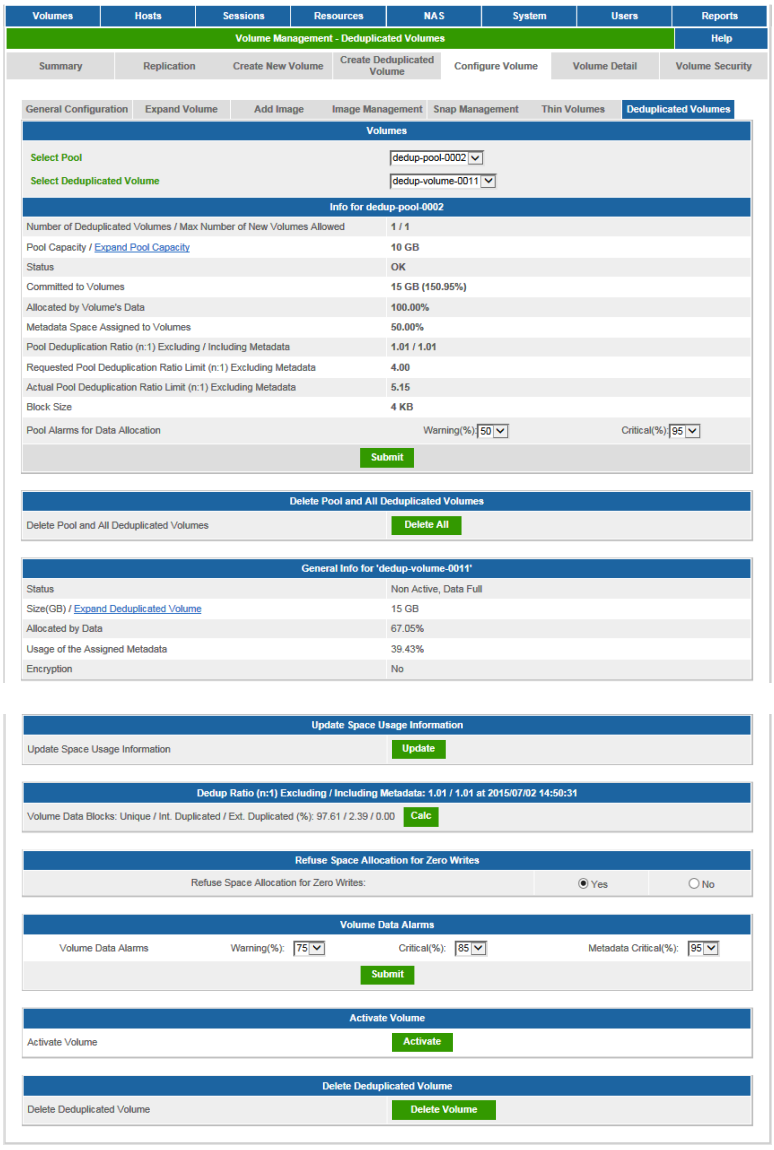


Figure 8-3 Deduplicated Volume Management screen for a pool that is out of free data segments (in a standalone system).

8.3 Deduplicated Volume Space Reclamation

The space allocated to a Deduplicated Volume can only increase, and will never decrease unless the special reclamation procedure is applied to the volume. The reclamation procedure consists of steps that have to be executed on the host that mounted the Deduplicated Volume and steps that have to be executed on the Storage Concentrator where the volume is provisioned. Reclamation has to be initiated manually. The procedure is only effective in cases where the Deduplicated Volume has allocated space that is not used by the host's file system at the time when the reclamation is executed. Because the allocated blocks are shared between all Deduplicated Volumes in the pool the reclamation has to be done for all Pool's volumes. The reclamation has to release this space to the Pool and provides available space to support new writes. Users should not rely on space reclamation as an ultimate tool and should become accustomed with the other options available that allow the deduplicated volume to be kept online.

Special attention must be paid to Deduplicated Volume and Pool space utilizations. Deduplicated Pool over-subscription should be avoided as once the Pool capacity is reached, new writes cannot occur until one or more Deduplicated Volumes that are off-line are copied out of the Pool and then deleted from it.

Host applications and procedures that tend to fully thicken a Deduplicated Volume should be avoided since the Deduplicated Volume's unused space cannot later be reclaimed.

The next options are available only in cases where the pool is close to free space exhaustion or is already out of free space:

The pool can be expanded by using available (not used) storage resources. See "Expanding a Volume". New storage resources can be added to the system to provide this expansion. See "Adding Resources";

Pool space used by some deduplicated volumes can be released by deleting these volumes from the pool. Deduplicated volumes left in the pool will reclaim the storage at run time.

Deduplicated volumes with allocation sizes above some limit can be transferred into regular volumes. The Storage Concentrator provides an offline procedure for this transformation. See "Copy Volume". The original deduplicated volume has to be deleted after the copy is executed successfully.

Deduplicated volume can be copied online by mirroring and executing rebuild procedure for the mirror. After the rebuild is finished successfully the deduplicated image has to be deleted.

Deduplicated volumes can be copied to regular volumes by using any host-based disk copy utility. The original deduplicated volume has to be deleted after the copy is executed successfully.

Deduplicated volume can be copied to a newly created deduplicated volume from the same pool by using host-based file-to-file partition copy utilities. The original deduplicated volume has to be deleted after the copy is executed successfully.

Deduplicated volumes used for the copy should not have any allocated blocks with unused information because the copy utility transfers only the currently stored files. If the size of the reclaimed space is not large enough then other options will have to be considered.

8.3.1 Space Reclamation for Deduplicated Volumes Mounted on Microsoft Windows

1 On the SC, go to the "Volume Configuration->Deduplicated Volumes" management screen for the specific Deduplicated Volume, and set the "Refuse Space Allocation for Zero Writes" to "Yes", and hit the "Submit" button.

2 To run the Space Reclamation procedure Deduplication Pool should not hit 100% usage of total data segments. If it happens, other methods to free space has to be used first. After that if the Deduplicated Volume is offline for host access, the volume can be changed to an online status so the host can perform some overhead writes. The host iSCSI session will need to be re-established, and a utility 'chkdsk' may be needed. Note that in this total space exhaustion case, all other host applications performing I/O to the same volume should be stopped until the unused space is zeroed and reclaimed.

3 Use utility 'sdelete.exe' on the host from the Microsoft System Internals with the "-p 1 -c" options on the drive letter for the Deduplicated Volume:

sdelete.exe -p 1 -c f:

The 'sdelete.exe' utility can be downloaded from the www.sysinternals.com.

4 On the SC GUI "Volume Configuration->Deduplicated Volumes" management screen for the specific Deduplicated Volume, check the status with the "Update" button.

8.3.2 Space Reclamation for Deduplicated Volumes Mounted on Linux

1 On the SC, go to the "Volume Configuration->Deduplicated Volumes" management screen for the specific Deduplicated Volume and set the "Refuse Space Allocation for Zero Writes" to "Yes", and hit the "Submit" button.

2 To run the Space Reclamation procedure Deduplication Pool should not hit 100% usage of total data segments. If it happens, other methods to free space has to be used first. After that if the Deduplicated Volume is offline for host access, the volume can be changed to an online status so the host can perform some overhead writes. The host iSCSI session will need to be re-established, and a utility 'fsck' may be needed before the file-system can be mounted. Note that in this total space exhaustion case, all other host applications performing I/O to the same volume should be stopped until the unused space is zeroed and reclaimed.

3 On the host, we will use 'dd' to create a new large file that is entirely zeroed until the entire file-system becomes full. In this example, '/dedup-test/' is the mount point for the Linux file-system on the SC Deduplicated Volume. As the root user, create a temporary "all zeros" file on the r/w mounted file-system that is being reclaimed:

```
# dd bs=1M if=/dev/zero of=/dedup-test/zeroing-file dd:

writing `/dedup-test/zeroing-file': No space left on

device 934+0 records in 933+0 records out
```

Perform the 'sync' command at the end:

```
# sync
```

The "No space left on device" error is expected. If it does not appear, it may be that the maximum file size for a single file is less than the amount of free space on the file system. In that case, multiple "zeroing files" will need to be created until the "No space left on device" error is seen.

4 Delete all of the temporary "zeroing files" that were just created.

```
# rm /dedup-test/zeroing-file rm: remove regular file

`/dedup-test/zeroing-file'? y
```

5 On the SC GUI "Volume Configuration->Deduplicated Volumes" management screen for the specific Deduplicated Volume check the status with the "Update" button.

8.4 Deduplication Friendly Hosts and Applications

The user should avoid doing host operations that cause writes to most or all of the deduplicated volume's blocks because this would cause allocation of storage blocks and the volume would no longer be thin. The recommendations in most cases are the same as for thin volumes. See "Thin Friendly Hosts and Applications".

Chapter 9

NAS Functions

NAS which stands for "Network Attached Storage" is a storage technology that allows users to create shares/volumes that can be accessed by one or more users. Difference between iSCSI volumes and NAS volumes are that NAS volumes can be accessed by more than one user/server, and iSCSI volumes in general can only be accessed by one server.

9.1 Introduction to StoneFly NAS implementation

StoneFly NAS is an integrated part of StoneFusion which is the operating system of all Stonefly's IP Storage appliances. All NAS functionalities are controlled from NAS menu of StoneFusion GUI as shown below.

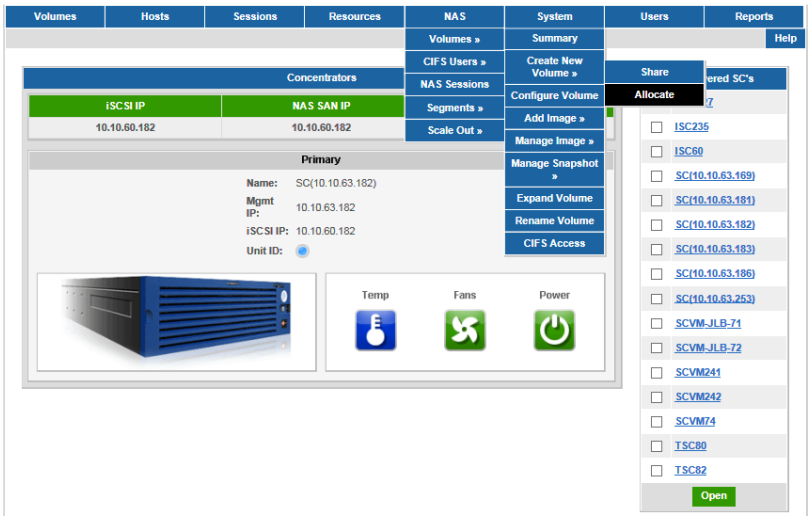


Figure 9-1 NAS Initial Menu GUI

NAS Segment: NAS Segments are a chuck of storage that will be used for NAS functionality. In other words, a NAS segment is part of the total storage resource that can be used for NAS. System can have multiple segments and each node of FailOver cluster has at least more than 1 NAS segment. NAS Volumes are created under segments. Volume's files can be distributed between different segments allocated on the same node or can be copied for redundancy between segments allocated on different nodes. NAS Segments themselves are not visible to NAS clients, only NAS Volumes are. Each NAS Segment is defined internally as a regular StoneFusion iSCSI volume or as a StoneFusion resource designated for usage as a NAS Segment . Unlike iSCSI volumes that are used by iSCSI host initiators, NAS Segment "volumes" are not exposed as iSCSI targets to any external host.

NAS Volumes: NAS volume is a piece like a folder or a subdirectory of a segment, which is created by NAS administrator and is given to different clients. When creating NAS volumes, NAS administrator decides what kind of protocol should the NAS volume support. There are different NAS protocols with CIFS and NFS being the two predominant protocols used by majority of the users. CIFS protocol is used mainly in Windows environment and NFS is used in Linux environment. When creating a NAS volume, administrator need to decide which protocol should be supported on that particular NAS volume. One NAS segment can have multiple volumes, each can support a different protocol. There are two ways to create a new NAS volume. One is to allocate new NAS segment and set new NAS volume based on this segment. The other way is to share ownership of the existing segments with other NAS volume or group of NAS volumes. From client's perspective there is no difference on how the NAS volume is created. After the NAS volume is created, NAS administrator decides which users can access that volume.

9.2 Create a NAS Volumes

9.2.1 Create a NAS Segment

Space for new NAS segment can be allocated from the available pool and has to be selected automatically or manually on managed local storage resources. Alternatively, the "Create Direct" tab can be used to create a direct NAS Segment using a NAS Managed resource.

There is a suggested default name for the segment being created. User can replace it with its own custom name. System does not provide option to rename the segment later.

The desired size of the segment is set after considering how many NAS volumes will be sharing ownership of the segment and what can be total size of used space on these volumes. Node for the segment has to be selected for FailOver Clusters. For standalone system or Cluster of One there are no reason to make this selection because the system has only one Storage Concentrator. Each segment provides service for dependent NAS volumes only on its node. The segment does not get provisioned on other node.

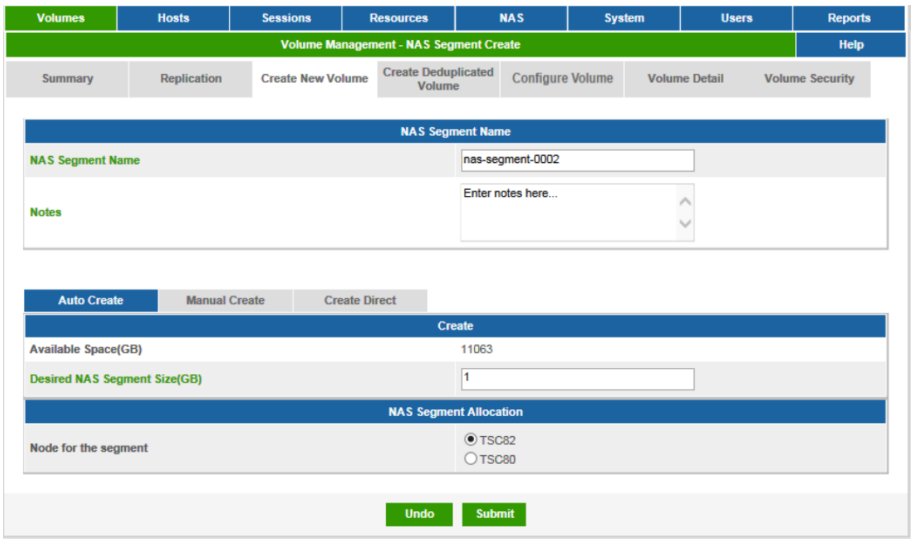


Figure 9-2 NAS Volume Manangment, Segment Creation Screen

After the segment is created its status can be checked on segments summary screen:

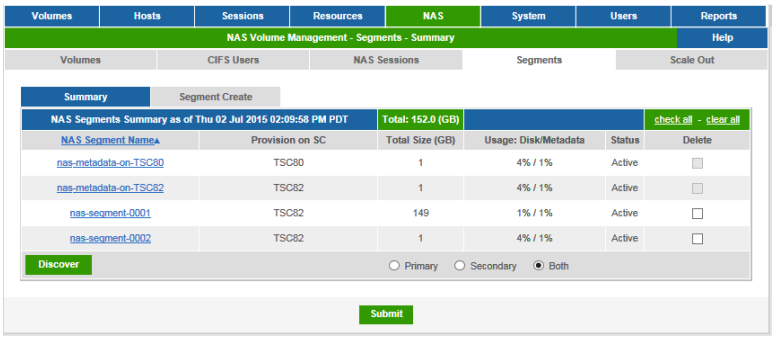


Figure 9-3 NAS Volume Management, Segment Summary Screen

Click the check box to select this segment to be deleted. A segment cannot be deleted if one or more NAS volumes use it; the check box is disabled in this case. A segment can have one of the following statuses: active, non-active, and off-line (see section 2.3.2). The Discover button provides the same functionality as the Discover button on Resource Summary screen (see section 2.3.1). The "Usage: Disk / Metadata" is the segment utilization in percentage for both file storage and file-system meta-data. When either of these sizes approaches 100%, delete unneeded files, or NAS volume expansion should be considered. When this information is Not Available, 'N/A' is displayed.

9.2.2 Create a Volume at Unused Segments

At least the first NAS volume has to be created by using this option. Later user has option to create new NAS volume by sharing NAS segments with the existing volumes.

Select item "Allocate" from "Create New Volume" menu. The system checks for available not used NAS segments. If some of the nodes do not have the required segment, user will be redirected to "Segment Create" screen. User has to follow recommendation of section 9.2.1 to provision the needed segment. Size of the new segment has to be equal to size of the already existing segment that will be used to provision new volume on the other node if this is the case.

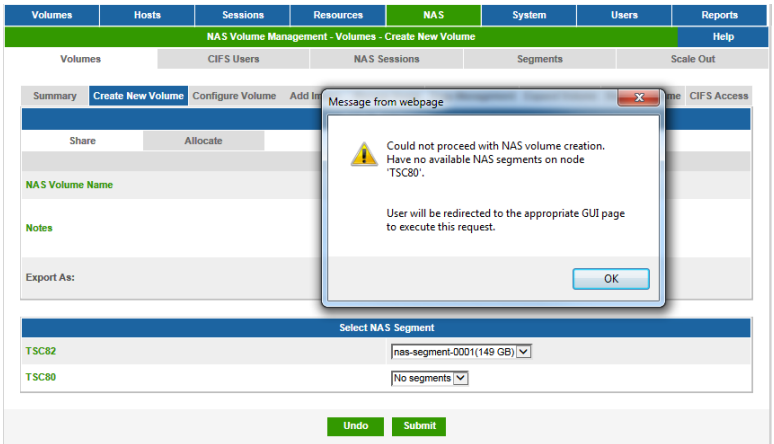


Figure 9-4 NAS Volume Management, Failure Warning Message

When the required segment is created successfully, user will be returned to the "Create New Volume" screen automatically:

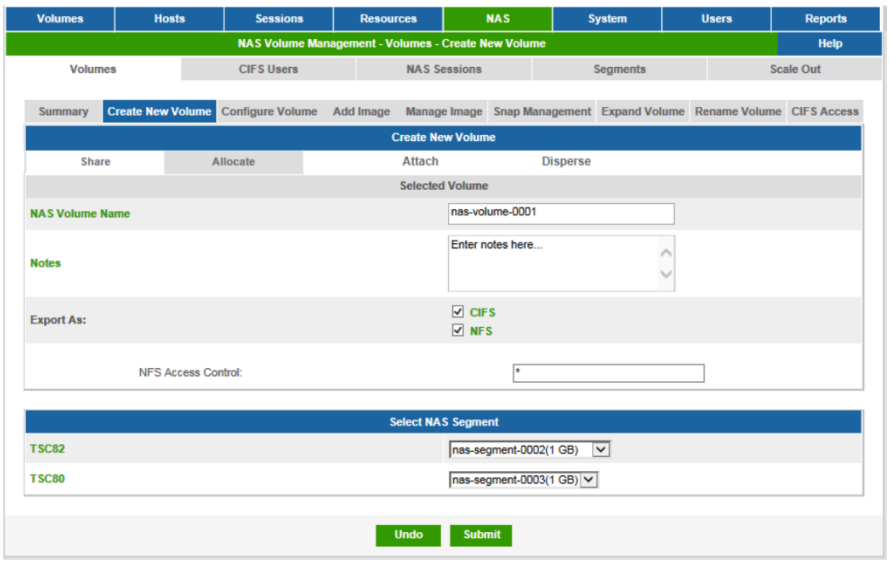


Figure 9-5 NAS Volume Management, Volume Creation Screen

Fill up all required fields on the screen. Currently, the NAS Volume name cannot be changed once created so it should be chosen carefully. To effect a change in the name of a NAS Volume, the content would need to be backed up externally and restored to a second NAS Volume, the clients aimed at the new volume, and then the first is deleted.

The NAS network protocols supported are NFS and CIFS (sometimes called SMB). Either, neither or both type of exports can be selected for a volume. Exporting using both protocols on the same NAS Volume is not usually recommended due to client file format differences, and different NAS file locking protocol semantics.

Access to CIFS exports is controlled through a CIFS user name and password. Multiple NAS Volumes may use the same user name, but in that case they must also share the same password. In such a case, changing the CIFS export password for one volume will apply to the other volumes as well.

Access to NFS exports is controlled through a list of NFS client IP addresses. A comma separated list of IP addresses of the NFS clients allowed to access the NAS volume. By default, all clients are allowed, with the setting of '*'. The wild-card character '*' can be used in fields of the IP address (e.g. "192.168.0.*,192.168.1.*").

Use the NAS Segment drop down list(s) of available NAS Segments to select the appropriate one to use for the storage for the NAS Volume. In Storage Concentrator Clusters, the same sized segments must be used for both SC's for mirroring purposes.

Status of the newly created NAS volume can be checked on "NAS Volumes Summary" screen:

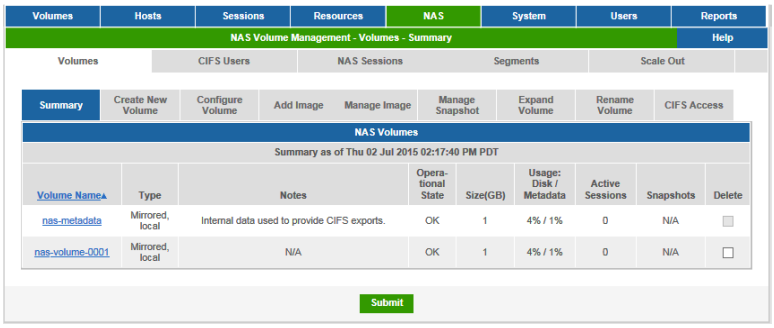


Figure 9-6 NAS Volume Management, Volume Summary

Operational State OK indicates that a session with the volume can be initiated. Offline indicates a session could not be initiated. Disabled indicates that user disabled NAS exports. REBUILDING indicates that volume is mirrored and process of image rebuild is in progress. NOT SYNCED indicates that volume is mirrored and images have to be rebuild.

The "Usage: Disk / Metadata" is the volume utilization in percentage for both file storage and file-system meta-data. When either of these sizes approaches 100%, delete unneeded files, or NAS volume expansion should be considered. When this information is Not Available, 'N/A' is displayed.

Volume size represents number of gigabytes in all volume's segments adjusted on mirroring. The each volume counts the shared NAS segment as many times as the segment is shared.

Active sessions field includes number of connections between CIFS clients and the volume. Number of connections between NFS clients and the volume is not included.

9.2.3 Create a Volume by Sharing NAS Segments

Select item "Share" from "Create NAS Volume" menu. Select volume from pull down list to share its segments with the volume to be created. Fill up other fields following the same recommendations as at the section 9.2.2.

Although the two NAS Volumes would share the same storage, they would not share any files. Different NAS export protocols and access settings can be used on each shared NAS Volume. When all disk space is consumed on one of the shared volumes, it is exhausted on all of them.

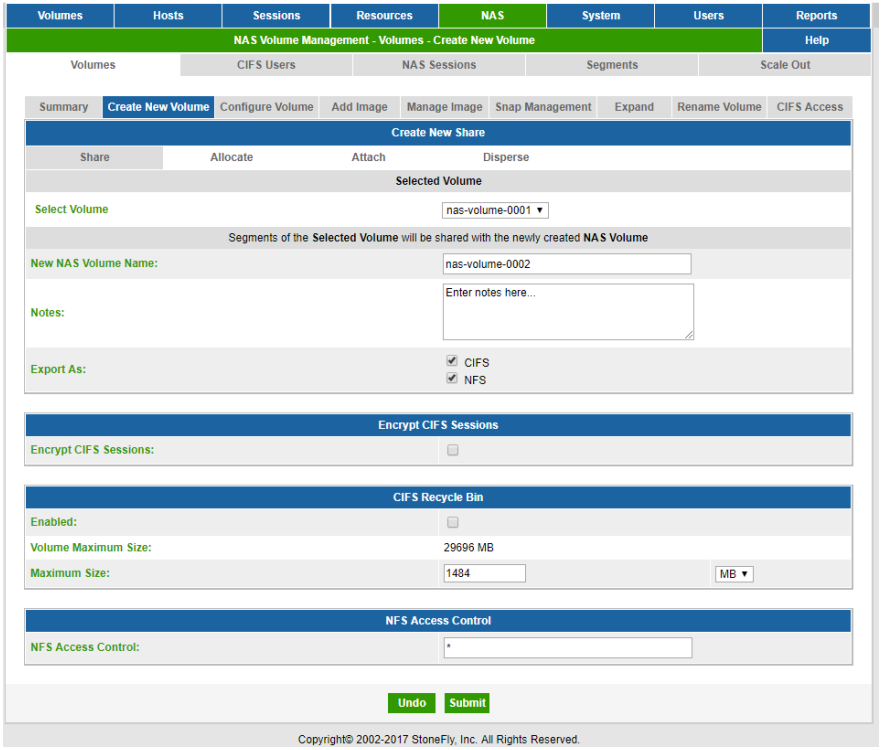


Figure 9-7

After the volume is created successfully, user is redirected to "NAS Volumes Summary" screen:

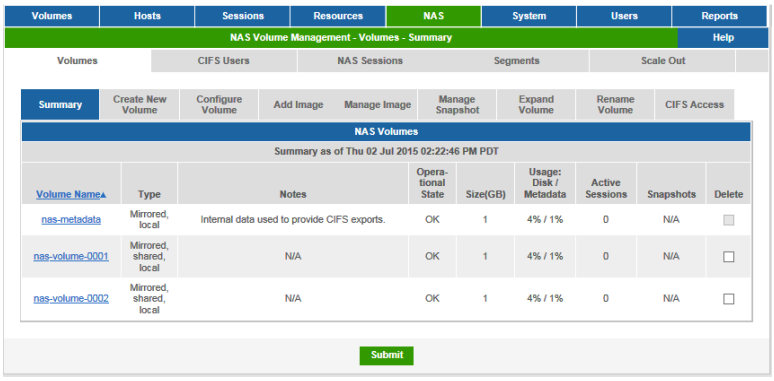


Figure 9-8

The newly created volume and volume used to share space have attribute "shared" at the "Type" field.

9.3 Volume Configuration

Select NAS volume from the pull-down list on "Configure Volume" screen. Current configuration of the selected volume will be displayed.

User can use this screen to modify volume's notes, modify list of IP addresses of allowed NFS clients or disable/enable NAS exports completely. See section 9.2.2 for detailed information about each field.

The image rebuild section only appears on Clustered Storage Concentrator for volumes that have two images defined.

Click button **"Start Rebuild**" Click to start a rebuild of the NAS Volume images. A NAS Volume rebuild performs an image consistency check between the two images, causing them to synchronize their contents when necessary. Note that it is valid to request a rebuild even when both images are already "OK".

Click button "**Refresh**" to update displayed information about the volume images rebuild.

Click button "**Stop**" to stop a NAS Volume image rebuild.

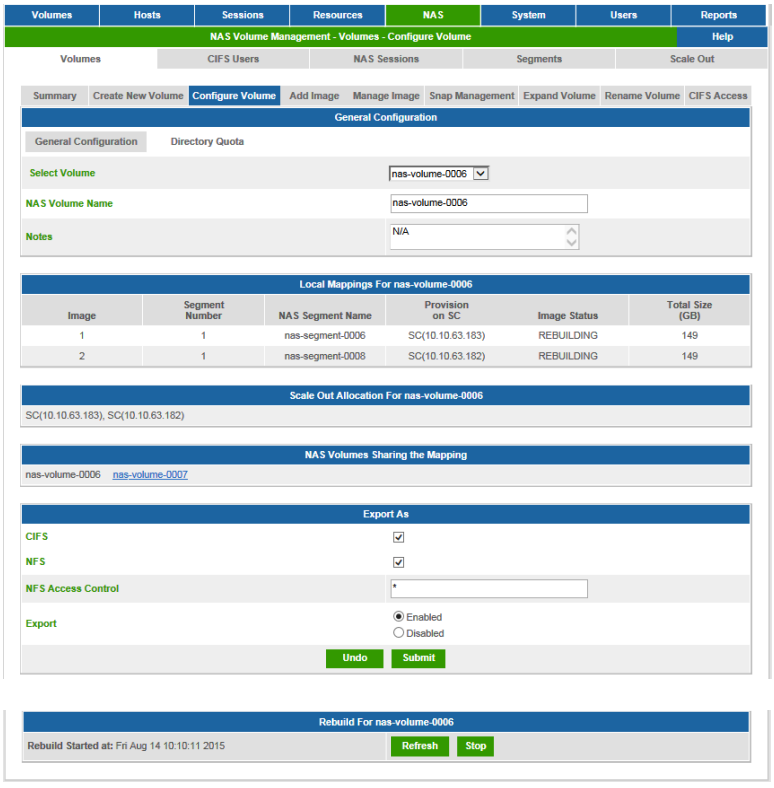


Figure 9-9

Information about the selected volume's segments presented in the mapping section.

The "Segment Number" is the segment slice number. The first slice of the volume is created when the volume is allocated initially. Each volume expansion adds new slice to the volume and new segments get appropriate numbers.

The "Image" is the image number for the segment. Mirrored volumes have two images and image numbers are 1 and 2. Distributed volumes have single image and the image number is 1.

Status of the segment can have the next values:

- DOWN when the Storage Concentrator where the segment is provisioned is down;
- NON ACTIVE when a segment's resources are not provisioned;
- FAILED when test command issued to the segment failed;
- NOT MOUNTED when the segment can't be exposed to the NAS subsystem;
- REBUILDING when the segment belongs to a mirror and the process of mirror rebuilding is in progress;
- NOT SYNCED when the segment belongs to a mirror and the image may be out of synchronization with the other image and needs to be rebuilt;
- OK none of the above and the segment is operating properly.

Note that for NAS Volumes, when an image is either REBUILDING or NOT SYNCED, most of the files are often present on both images, but because the volume was known to be offline at some point, some of the changes made to the other image may be missing.

Also note that it is possible in some cases for both images to be either REBUILDING or be NOT SYNCED. This means that both images had been offline at some point without their content having been brought back into synchronization before the other image became offline. When both images are online, even if not "OK", the SC will still deliver the correct, most updated copy of the files. However, if one image is offline, and the other is not "OK", there is the possibility of old content being delivered.

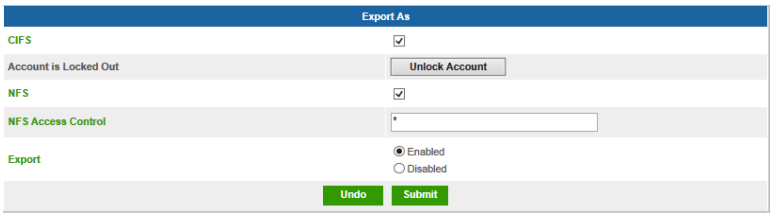
List of NAS volumes that share the same segments with the selected volume is present only for volumes that have shares.

9.3.1 CIFS Account Unlock

The **Unlock Account** button is only shown when a CIFS User Account has become locked out due to too many bad password login attempts.

Depending upon the system configuration, the account may automatic unlock after a period of time, or the administrator may unlock the locked account manually at any time with the **Unlock Account** button.

The CIFS User Account lockout policies can be set by the administrator in the CIFS User Policies section of the System Management Admin General GUI page.



9.3.2 Image Manual Repair

There are cases when the files on the two images of a NAS Volume differ, and the correct version to use cannot be automatically determined. In such cases, it will not be possible to rebuild the images and manual intervention is required to select the correct version of the files to use.
Only NAS Volumes with two images need rebuilding, so only clustered SC's with mirrored NAS Volumes would ever need manual repair.
The **Image Repair** section of the GUI page is only shown when a repair is needed, or is in progress.
The **Repair Log** is the result of the last image rebuild attempt, and is used as the list of files that require manual selection. These conflicting file versions are sometimes referred to as 'split-brain' files.
During image manual repair, normal access to the NAS Volume can continue, but the files that cannot be automatically reconciled may not be accessible.
To affect a NAS Volume Image Repair:

When a manual repair is needed, the repair log should be opened using the **Repair Log** button, and be saved to a local file since it will be deleted and regenerated on the next manual or automatic rebuild attempt.

Enter a temporary CIFS export administrative username and password to use for the image repair shares.
NOTE: Normal users and clients should not use the repair shares, only an administrator affecting the manual repair should be allowed access.
Do not use the repair exports for normal storage access as this would interfere with normal NAS Volume image operation.

Press the **Enable Repair** button to cause all of the NAS Volumes images NAS Segments to be exported as CIFS Shares using the username and password provided.
The names of the exported CIFS Shares are in the form:
"*NAS Volume Name_NAS Segment Name*".
Note that for NFS NAS Volumes, the CIFS repair shares are used for repair, there is not an NFS repair share.

From a client system, mount the CIFS repair shares for all of the NAS Volume's images segments.
Each NAS Volume's images segments are only accessible from the SC that manages it, so the client must use different SC IP addresses to access each images segments share.
If both the NAS SAN IP and NAS LAN IP addresses are configured, each may be used to access the images. Otherwise, each SC's iSCSI SAN IP address should be used instead of the Cluster SAN IP.
Because each image segments are only accessible from the SC that manages it, the SC cluster must be operating duplex to reach both images segments.

Using the **Repair Log**, for each file listed, locate both copies in the mounted repair image segments. Compare the two file versions, and using the file size, last modified time, and the files contents, select the file that is most likely to be the correct version. After first making a backup of the other file to local storage, delete the incorrect file version.

Repeat the above step for all of the files listed in the Repair Log.

Once all of the files have been reconciled, disconnect the client from the image repair shares, and press the **Disable Repair** button to remove the temporary repair exports. Disabling the repair exports causes both images to be marked as needing a rebuild, and will start a rebuild of the volume.

Once the rebuild completes, if it was successful, the repair is complete.
If the repair was not successful, repeat the above steps until it is.

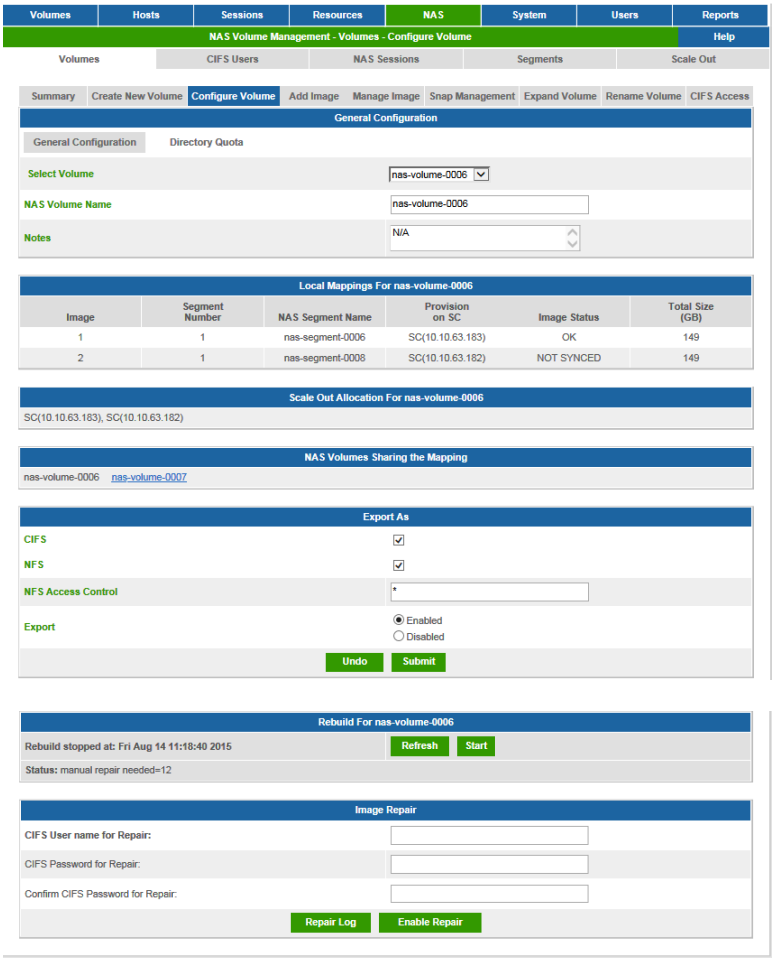


Figure 9-10

9.4 Volume CIFS Access

NAS Volume CIFS access provides volume security by creating or modifying the access, which restricts user access to the selected volume.

The CIFS access applies to the entire NAS volume, not specific files or directories. For example, a user may have R/W access to the NAS Volume, but files within that share may be hidden, or read-only to them based on the file permissions on the file/directory. The CIFS permissions on files and directories are viewed and set by the Windows client.

Each CIFS user is defined in either the local Storage Concentrator (SC) Workgroup, or in a Windows Active Directory Server.

The "Workgroup" represents users that have user accounts defined by the SC itself. The SC is responsible to keep these accounts in sync with other clustered or Scaled Out SC's.

The Active Directory Server domain is defined by a Windows Active Directory Server (ADS). Users from the Active Directory domain are defined externally. Each Storage Concentrator gets list of users automatically from the ADS domain controller.

The Active Directory domain also defines groups. In this case Storage Concentrator can offer access to these groups as a whole and uses the group name to manage the group access to NAS volumes. In this case, access is granted to the ADS defined group, so the SC does not know the users that are members of the group, which may change over time.

There are no groups for the "Workgroup" mode of authentication.

9.4.1 "Workgroup" Users

To create new "Workgroup" user use "Add/Update User" screen from the "CIFS Users" menu:

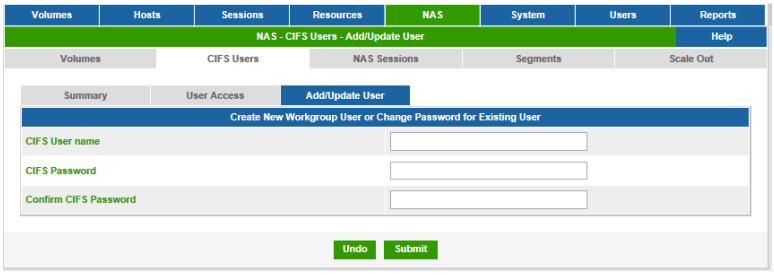


Figure 9-11

The page can be used to modify password for existing user.

Any existing CIFS user name can be used including names for users and groups from the Active Directory domain. In this case the newly created Workgroup user inherits NAS volume access list that the Active Directory user or group with the same name had before.

9.4.2 Active Directory Users

Before the system can manage Active Directory users it has to join the Windows Active Directory Server. For details see "NAS Server".

After the system joined the Active Directory Server successfully list of all available users and groups of users has to be managed on "CIFS Access" screen from NAS Volume menu.

Active Directory users and groups with assigned NAS volumes can be handled on "Summary" and "User Access" screens from the "CIFS Users" menu. The screens do not show Active Directory users that have no access to the Storage Concentrator NAS volumes.

9.4.3 Access for Active Directory Users and Groups

Use "CIFS Access" screen from NAS Volumes configuration menu. The **CIFS User Authentication Mode** has to be set to **ACTIVE DIRECTORY** on the *NAS Server Configuration* GUI page.

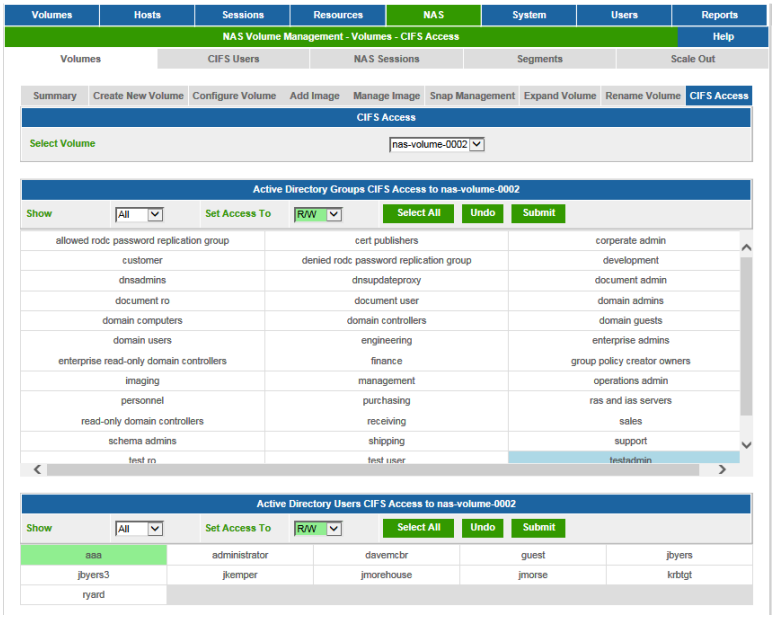


Figure 9-12 AD Users and Groups Access

9.4.3.1 CIFS Access Screen Fields

Show — Select one of the following options to filter the displayed users by the kind of access they currently have:

1. 'Any' to show users with any access to the volume except the 'None'.

2. 'All' to show all Active Directory users.
3. 'None' for users that don't have access to the volume.
4. 'R/O' for users with Read Only access to the volume.
5. 'R/W' for users with Read Write access to the volume.
6. 'Admin' for users with administrative privilege. Administrative users could perform any of file operations as a super-user(root). In addition they have Read Write access to the volume.

The level of access for a user is shown by different background colors based on access type: The color "white" is used to show users without any access, the color "light blue" shows users with Read Only access, the color "green" shows users with Read Write access and the color "pink" shows users with Administrative access.

The level of access for a user is also shown in "hover" text when the mouse pointer is held over the user's name.

Set Access To — Used to assign new access to users at the table. After selection at the field is made, the appropriate set of users has to be selected at the table.

Toggle changes to a user's selection by clicking the cell that represents the user. The user's access has to be different from selection at the "Set Access To" field. If selection is appropriate the user's name will be displayed in **bold** and it will be underlined. The background color will be changed to color representing the desired access.

When there are a large number of users, the browser's find in page function (often Ctrl-F) can be used to locate a specific user.

Additional users can be selected in the same way.

To unselect the previously selected user have to click on the user's cell. The background color and font have to be restored to the original values.

Button **Select All** can be used to select all available users simultaneously. Some users may then be unselected if desired.

To restore the current access list use button **Undo**.

After the appropriate selections are done for the selected "Set Access To" value, click the **Submit** button to update Storage Concentrator Configuration with new information.

9.4.4 Access for Workgroup Users

This section shows how to manage users defined by the Storage Concentrator itself. The management screen is shared with the Active Directory Users management:

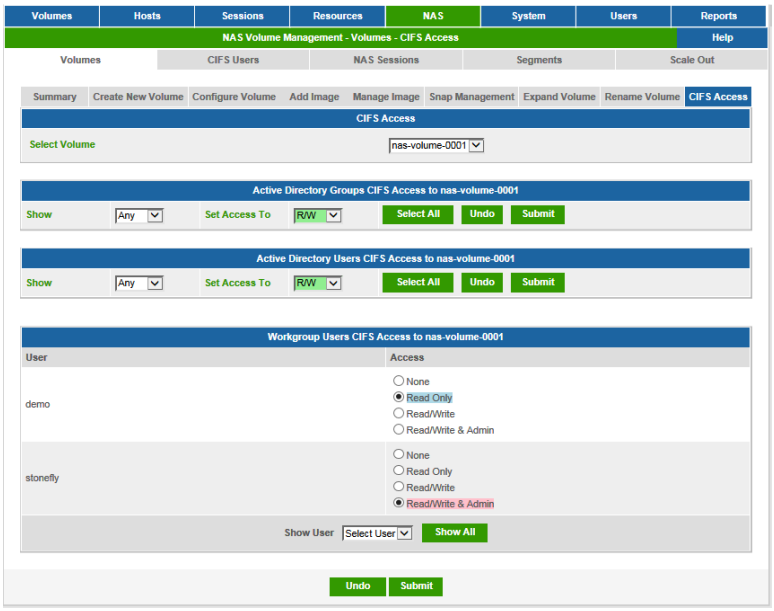


Figure 9-13 Workgroup Users Access

9.4.4.1 CIFS Access Screen Fields

Access — Set the access permissions for the user. Read Only, Read/Write and Administrator are available only for NAS volumes that are managed by the Storage Concentrator Configuration. The Administrative permission includes the Read/Write permission. In addition the administrator can perform any of file operations as a super-user(root). The options for this field are None, Read Only, Read/Write and Read/Write&Admin.

User — The Workgroup user name.

Show User — Select from the available list of users using the drop down menu. The easiest way to use this feature is to select users from the drop down menu one by one. Set appropriate permissions. Then click on **Submit**. The **Show All** button can be used instead to display all available Workgroup users.

9.4.5 Deleting User's Access

The CIFS Users Summary screen provides high level information on all configured CIFS Users across current Storage Concentrator configuration. The screen shows only Active Directory domain users that have any access to the Storage Concentrator NAS volumes. Active Directory users without access can be managed on "CIFS Access" screen only.

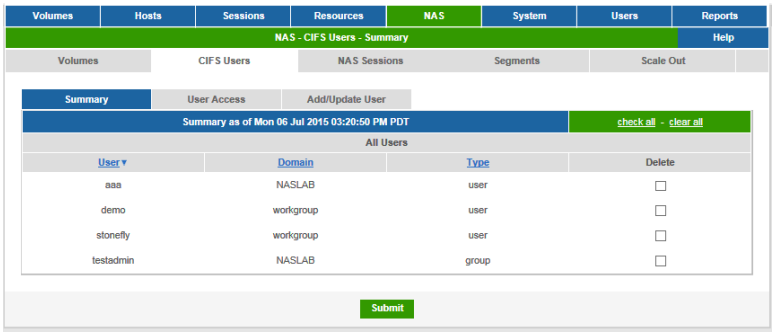


Figure 9-14 CIFS Users Summary

To delete user's access to all NAS volumes click the check box under delete to remove user or group. Apply "**Submit**" button after all selections for deletion are done. After the deletion the user or the group can't be used to access NAS volumes provisioned at the Storage Concentrator configuration. The deleted Active Directory user or group is still listed on the Active Directory Server. The deleted "Workgroup" user is not listed at the Storage Concentrator configuration any more.

9.4.6 Modifying User's Access

"Workgroup" users and Active Directory domain users and groups can be managed on volume oriented "CIFS Access" screen. The "User Access" screen from CIFS Users menu provides the same functionality but on per user basis. The screen can manage only Active Directory users and groups that have at least some access to the Storage Concentrator NAS volumes. The restriction does not apply to "Workgroup" users.

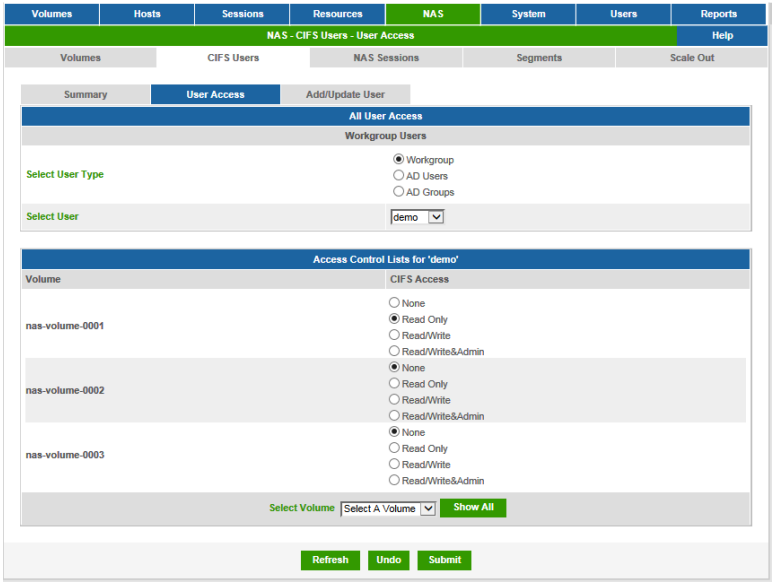


Figure 9-15 CIFS Users Access

9.4.6.1 User Access Screen Fields

Select User Type — Select one of three possible options. The 'Workgroup' option lets handle users defined by local systems. The 'AD Users' selection lets select any of CIFS users that belong to Active Directory domain. The last selection 'AD Groups' lets handle groups of CIFS users defined by Active Directory domain.

Select User — Select a user or group from the drop down list. When the list is long, a specific name can be selected by quickly typing the first few letters of the name.

CIFS Access — Set the access permissions for the user. Read Only, Read/Write and Administrator are available only for NAS volumes that are managed by the Storage Concentrator Configuration. The Administrative permission includes the Read/Write permission. In addition the administrator can perform any of file operations as a super-user(root). The options for this field are None, Read Only, Read/Write and Read/Write&Admin.

Volume — Lists of NAS volumes that this user currently has access to.

Select Volume — Select from the available list of volumes using the drop down menu. The easiest way to use this feature is to select volumes from the drop down menu one by one. Set appropriate permissions. Then click on **Submit**. The **Show All** button can be used instead to display list of all available NAS volumes. Use button **Refresh** to show the current ACL's for the selected user.

9.5 NAS Sessions

The NAS Sessions screen displays active NAS sessions between clients and the Storage Concentrator.

On table column headings that are highlighted, click on the column heading to sort by that field. Click again to reverse the sort order.

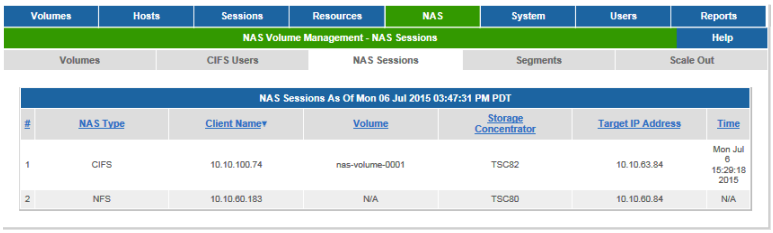


Figure 9-16

The NAS client name is shown when it is available, along with its IP address. This is the IP address that the client established the session from.

The NAS session type can be either CIFS (also known as SMB), or NFS.

The "Volume" is volumes name the NAS session is logged in to. Note that for NFS sessions, the volume name currently cannot be determined, so N/A (Not Available) is displayed.

The Storage Concentrator is the node name that the session is logged in to. In an SC cluster, this could be either SC. This column is only shown when there are two SC's in a cluster configuration.

The "Target IPAddress" is SC's target IP address that the session terminates on. In an SC cluster, or when there are SC Multipath network interfaces defined, or the LAN network is also used for NAS, this address may vary and indicates which interface/network is being used.

The date and time that the NAS session was established is displayed at the last column. Note that NAS sessions drop due to inactivity, and reestablish automatically based on need. For NFS sessions, the time the session was established currently cannot be determined, so N/A (Not Available) is displayed.

9.6 Delete NAS Volume

Do delete NAS volume use "NAS Volumes Summary" screen (see picture from section 9.2.2). After the selected volume is deleted, user has to check NAS segments "Summary" screen and make decision to keep ones that became unused after the NAS volume deletion or to delete them.

9.7 NAS Volume Expansion

The expansion of NAS Volume is procedure that consists of the next steps. During the first Step new NAS segments of the same size are created on each Storage Concentrator where the volume is allocated. Segments are added to the volume during the second Step. Final the third Step has to create layout of the volume directories on the newly added segments and move volume files between old segments and new segments to balance access. Completion of the volume rebalance could be checked on the NAS Volume Configuration GUI screen.

Pre-existing NAS Segments that do not belong to other volumes can be used for expansion.

All NAS volumes that are sharing the same segments will be expanded simultaneously. It's enough to pick up one volume from the group and initiate expansion for the selected one.

Volumes could be online during expansion.

Mirrored volumes have to be in sync before expansion can be initiated.

9.7.1 Fields

Select Volume — Select volume from the drop down list for expansion.

Select NAS Segment— Use the NAS Segment drop down list(s) of available NAS Segments to select the appropriate one to use as an expansion for the NAS Volume. If the selection list is empty, the user will be redirected to "NAS Segment Create" GUI screen automatically. Adjust the size, name and set notes for the new NAS Segment during segment creation procedure. After the segment(s) are created, the GUI will return to the "Expand NAS Volume" screen automatically and the user may proceed with the expansion. In Storage Concentrator Clusters, the same sized segments must be used for both SC's for mirroring purposes.

9.7.2 Buttons

Start — Click to start procedure of the selected volume expansion.

Undo — Click to revert to the last saved settings.

Submit — Click to add selected segments to the volume and to start the volume rebalancing.

9.8 Rename Volume

NAS volume renaming is very expensive procedure because volume name is used internally as a tag for objects created to provide volume functionality and is used externally by NFS and CIFS clients. This is the reason why scope of this procedure is limited to stand-alone Storage Concentrators only. The renaming maybe needed during NAS Scale Out configuration creation. Storage Concentrator can join Scale Out NAS configuration if names of its NAS volumes are different from names of NAS volumes already provisioned at the configuration. In case of conflict the simplest solution is to rename NAS volumes at system that has to join the existing Scale Out configuration.

All client sessions to volume will be closed during renaming automatically. When renaming is done user has to reconnect NAS clients by using new volume name.

9.9 FailOver Clusters

In general a NAS FailOver Cluster has the same features and based on the same 2-node model as a StoneFly iSCSI FailOver Cluster (see section **Error! Reference source not found.**). But ability to provide support for NAS volumes requires additional considerations.

First, NAS volumes are synchronously mirrored on each Storage Concentrator at the cluster, so user has to specify separate storage segments to allocate images on both nodes.

Second, two additional IP address aliases are defined, one is for Data Network, the other is for Management Network. These aliases are used exclusively by NAS clients to access NAS volumes. When both Storage Concentrators are online, one of the NAS IP aliases is assigned to one node and the other is assigned to other node, so the incoming and outgoing IO traffic is going thru one Storage Concentrator for Data Network based clients, and thru other Storage Concentrator for Management Network based clients. When one of the nodes is down, the other node gets both NAS aliases and is responsible to handle external IO traffic for both networks.

When the failed node is returning to the cluster the system is responsible to rebuild images and make them synchronous again. The system scans NAS volume files on both images and replaces the old copy with the found newest version. The NAS volume rebuild procedure is executed in background so client's access to the volume does not interrupt.

9.9.1 NAS Specific to Set FailOver Cluster

Before proceeding with the setup user has to verify that NTP service is enabled. For details of setting the NTP service see section 2.7.10.

Storage Concentrator has to have enough available space on managed local resources (see section 2.3) to mirror all existing NAS volumes. The extra 2 Gig of the space will be used to create internal NAS volume "nas-metadata".

Storage Concentrator that is designated to be the other member of the cluster has to be online at this moment and has to have licenses that match licenses that are set on system that user is configuring to be the Primary.

Sessions from all NAS clients have to be closed until cluster is set up and all monitors are in good state.

9.9.1.1 NAS IP Aliases

To set FailOver Cluster user has to define two NAS IP Aliases in addition to IP addresses used before support for NAS volumes was implemented:

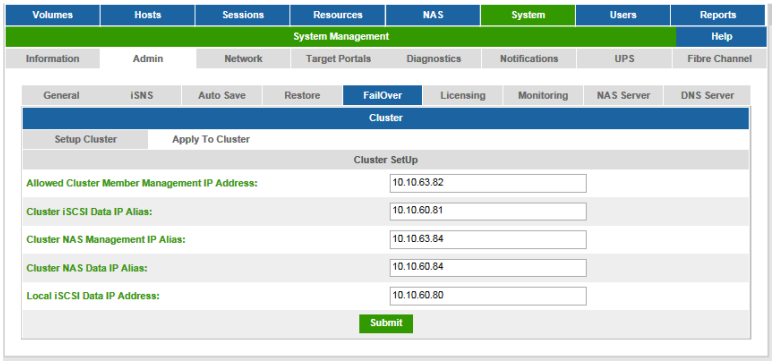


Figure 9-17

HIT "SUBMIT" BUTTON TO INITIATE CLUSTER CREATION.

The next screen is indication that the process of setting primary Storage Concentrator is successful. The system represents Cluster of One and has to handle NAS and SAN clients requests by using IP aliases only.

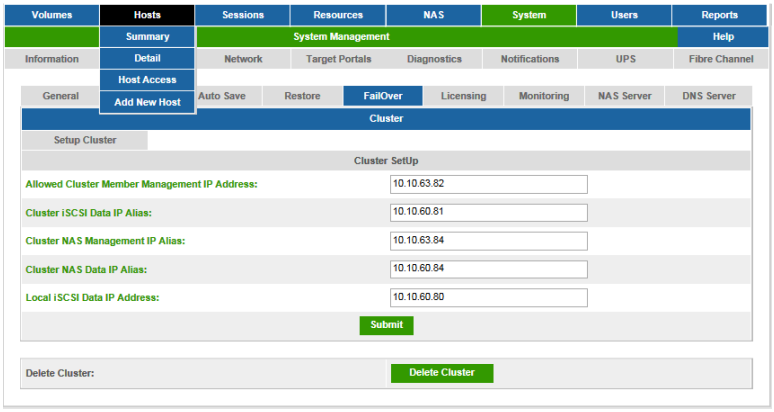


Figure 9-18

User has to change configuration of all iSCSI hosts to start use Cluster iSCSI Data IP Alias. NAS clients have to be modified too and start to use one of Cluster NAS IP Aliases.

9.9.1.2 State of Cluster Monitors

When status all monitors of the Primary Storage Concentrator is OK:

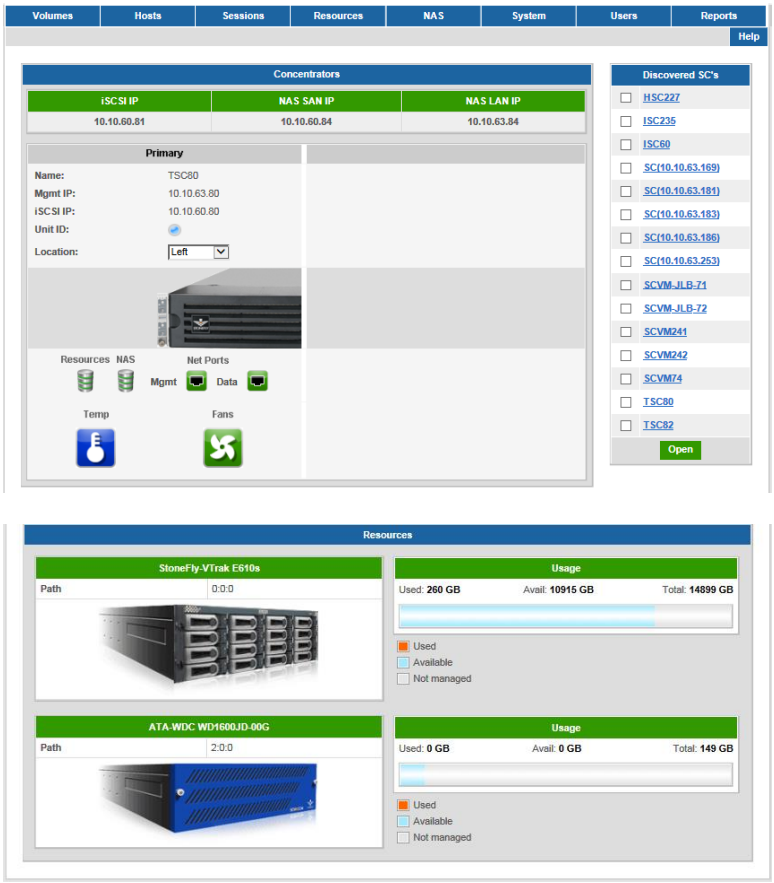


Figure 9-19

It's appropriate for the Secondary Storage Concentrator to join the cluster (for more details see 3.2.2).

It takes time to provision cluster and to create the "nas-metadata" volume. When this step id done the "Home" screen has to look like the next:

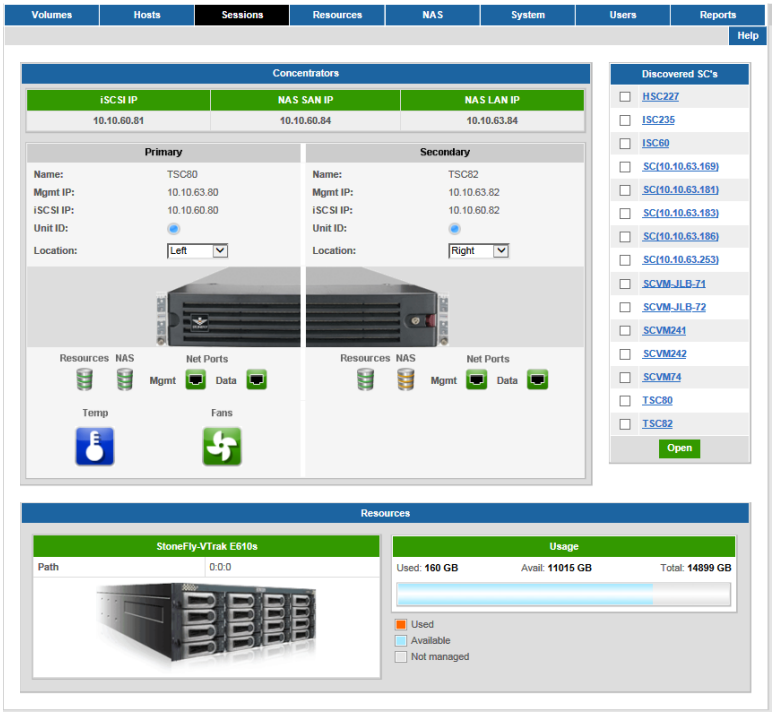


Figure 9-20

NAS monitor on the Secondary Storage Concentrator is in "critical" state because the system used to build cluster from had NAS volumes initially. These volumes are preserved and are defined as distributed on the Primary Storage Concentrator only. If the Primary fails, NAS clients lost connection to the volumes. This is the reason why the NAS monitor on the Secondary is in critical state. Distributed volumes have to be mirrored to handle this problem. For system without original NAS volumes the monitors have to be at good state on both nodes.

9.9.1.3 The Metadata Volume

For systems with pre-existing NAS volumes the NAS Volumes summary screen has to look like the next one:

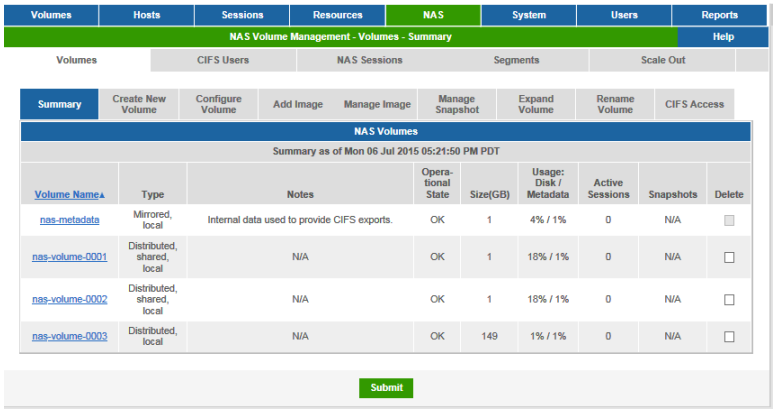


Figure 9-21

The special system volume "nas-metadata" is defined for FailOver Clusters only. The volume is used by the system internally.

Clusters without the "nas-metadata" volume or with "nas-metadata" volume at "offline" state are unable to provide NFS and CIFS exports to external NAS clients.

The "nas-metadata" volume can't be deleted from the FailOver Cluster explicitly.

The "nas-metadata" volume is created by system automatically when the Secondary Storage Concentrator joins the cluster. It is deleted by the system automatically when the Secondary Storage Concentrator is deleted from the cluster or when the cluster is deleted itself.

Should the SC fail to create "nas-metadata" volume automatically, the user will be asked to do this manually next time when they use the NAS Volume Management GUI screens.

9.9.1.4 Add Image

Distributed volumes can't provide non-interrupted service in case of fail over because the volumes are allocated on one of the cluster's nodes. When the node is down the volume is not available. To fix this problem the procedure "Add Image" has to be invoked for all distributed NAS volumes. For each single-node volume it let user create synchronous image allocated on other node. To do this select volume from the pull-down list and hit button "Start Image Create":

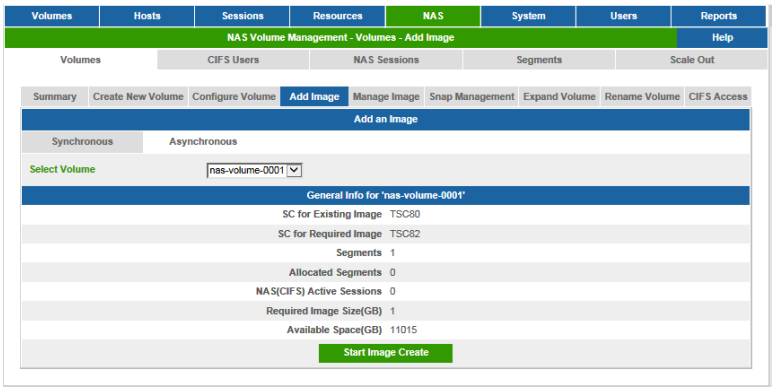


Figure 9-22

All NAS volumes that are sharing the same NAS segments will be mirrored simultaneously. It's enough to initiate mirroring for any one volume from the group. The procedure consists of some number of steps managed by the user. New NAS segments have to be created during first steps of the procedure. To do this user will redirect to the "Create NAS Segment" screen and returned back to the "Add Image" screen automatically. The procedure can be stopped any time before the user made the final decision to proceed with image creation:

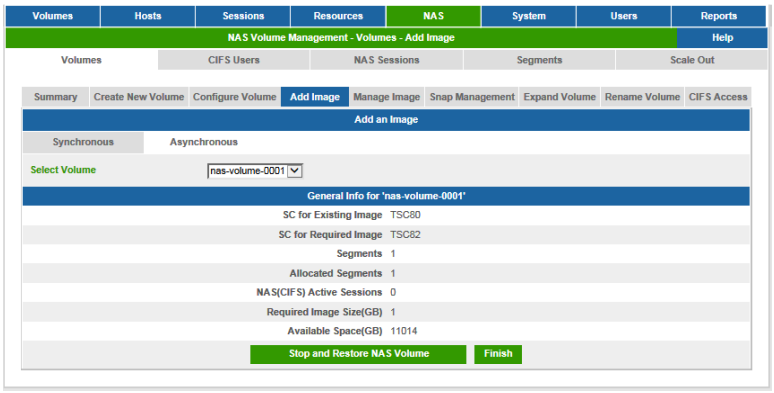


Figure 9-23

After all distributed images are mirrored user can open sessions from NAS clients. NAS IP aliases have to be used by the clients only as it was specified at the section 9.9.1.1.

9.9.1.5 Image Rebuild

The newly created image has to be synchronized with the original image. Invoking the rebuild process does this. It starts automatically and runs in background.

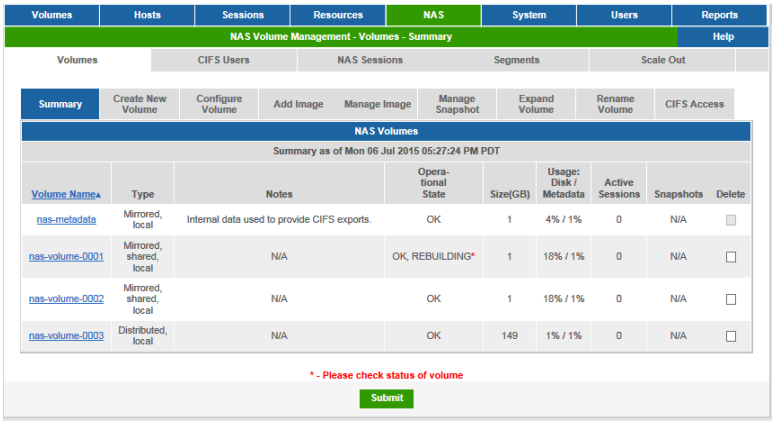


Figure 9-24

9.9.2 NAS Specific to Delete FailOver Cluster

In some cases user has to take one of cluster's node out of the system due to hardware failure or some specific upgrades. Deleting the FailOver Cluster can do it. It will transform the Primary Storage Concentrator node into regular Storage Concentrator. If user wants to preserve cluster's settings the Delete Secondary Storage Concentrator can be invoked. Before proceeding user has to close sessions from NAS clients and has to verify that all images are in sync. It's recommended to initiate rebuild manually and wait until the rebuild is finished. In both procedures mirrored NAS volumes will be transformed into distributed automatically. The "NAS-metadata" volume will be deleted automatically. NAS volumes distributed on the Secondary node will be deleted. To preserve these volumes user has to mirror the volume before the Secondary Node deletion.

9.10 NAS Volume Snapshots

This function allows the user to snap enable previously defined NAS volume. All segments of the volume have to be provisioned as a snap enabled StoneFly iSCSI volumes. The segments have to be created by using only "Managed" resources. NAS volumes with segments created from "NAS Managed" resources can't be made snap enabled. NAS volumes that are sharing segments with other NAS volumes can't be made snap enabled too. NAS Volume Snapshot creates virtual, temporary, and perishable point-in-time images of the Volume. A Snapshot contains an image of the volume at the exact point in time that the snapshot was taken. NAS volume Snapshot is not a NAS volume. The snapshot appears to the NAS volume clients in read-only mode. Even after changes are made to the original volume, Snapshot preserves a copy exactly as it existed when it was taken. Snapshots persist across reboots and can be mounted and accessed for reading just like any other volume. NAS volume snapshot implementation is based on creating iSCSI volume snapshots on segments that the NAS volume is provisioned on. For details of iSCSI volume snapshots see Chapter 6 (Snapshots).

9.10.1 Snap Enable NAS Volume

Before to create NAS volume snapshot the volume has to be made snap enabled:

1 From the *NAS Volumes* screen, click **Snap Management.**
2 From the *Select Volume* drop down menu, select the correct volume.

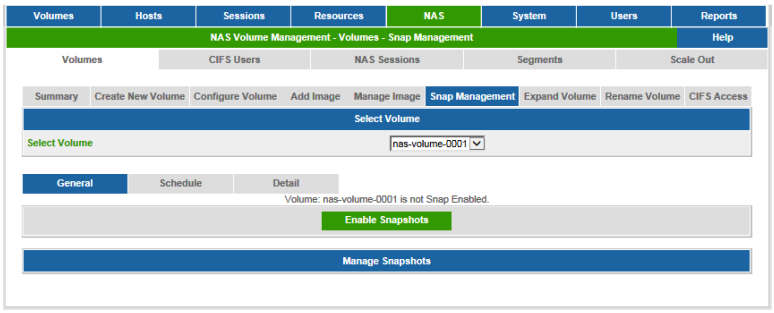
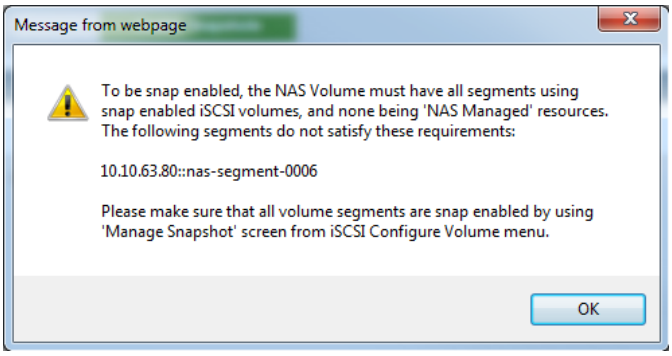


Figure 9-25 Enable NAS Snapshots

3 Click **Enable Snapshots.**

4 If not all segments for selected NAS volume are snap enabled iSCSI volumes the next message is generated:



Make all listed snapshots snap enabled. See 6.2 for details.

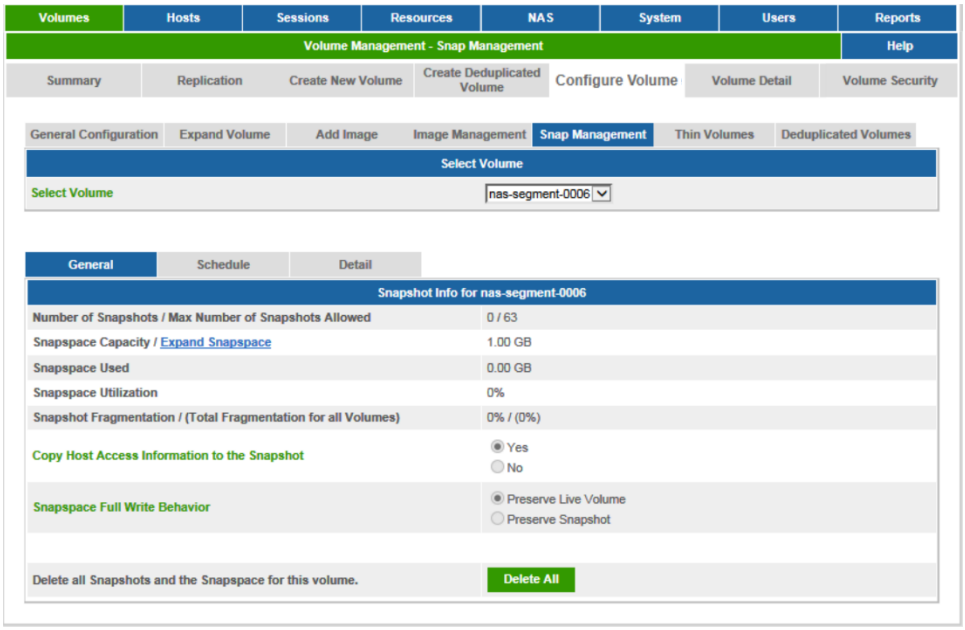


Figure 9-26 Snap Enable NAS Segment

After all requested segments are snap enabled go to step 1 and try to enable snapshots for the selected NAS volume again.

5 In case of successful completion the *Snap Management* screen for the selected NAS volume has to have the next view:

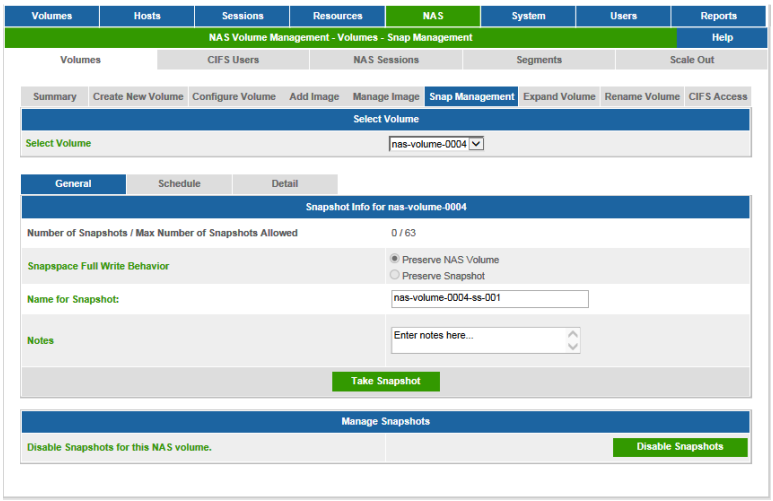


Figure 9-27 Snap Enabled NAS Volume Screen

For now each snap enabled NAS volume could support up to 63 snapshots. In case when at least one of the volume's segments becomes out of snap space or segment's snap space utilization reaches 95%, the system deletes NAS snapshots in FIFO order until segment's utilized snap space is 80% or less.

6 The **Disable Snapshot** button has to be used to delete all created snapshots and to remove snap enabled attribute from the selected NAS volume.

9.10.2 Creating Snapshots

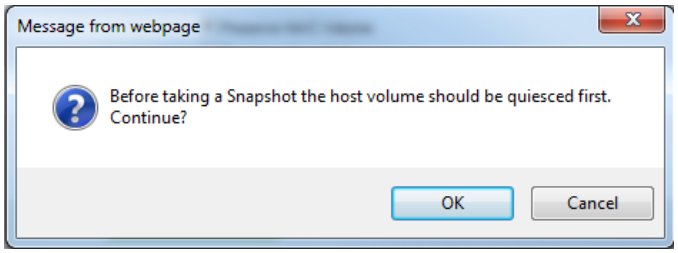
1 From the *NAS Volumes* screen, click **Snap Management.**

2 From the *Select Volume* drop down menu, select the correct volume. Selected NAS volume has to be snap enabled already.

3 Set name for snapshot or keep the default name that is offered by the system. Put comments into the *Notes* field if needed.

4 Click **Take Snapshot.**

5 The *Storage Concentrator* displays a reminder message that the host volume should be quiesced before taking the Snapshot. Click **OK** to continue or **Cancel** to abort snapshot.



6 Snap Management screen will be updated to take the next snapshot if the previous snapshot is created successfully:

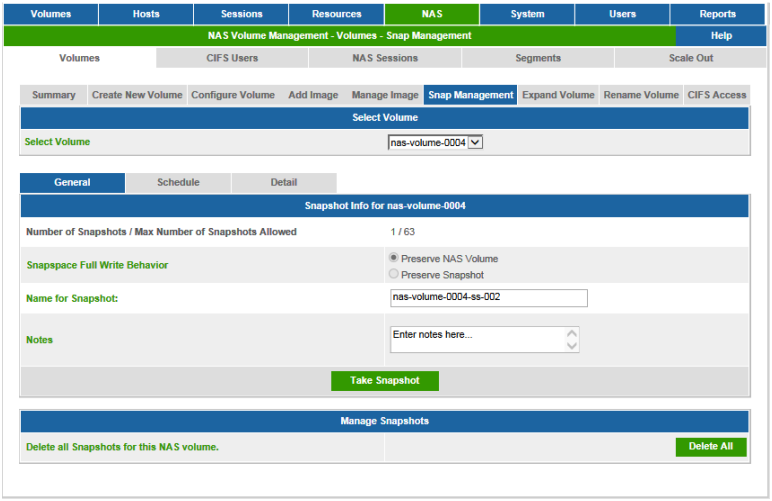


Figure 9-28 Taking NAS Snapshot Screen

9.10.3 Scheduling Snapshots

The *Storage Concentrator* can automatically take Snapshots at predetermined intervals using the Snapshot scheduler. Administrators can configure the system to automatically take Snapshots hourly, daily, weekly and monthly. StoneFly has configured the total number of Snapshots to take advantage of n+1 Snapshots in each category (e.g. Daily Snapshots has a maximum of eight retained Snapshots---one for each day of the week plus one).

Type of Snapshot	Max. Number of Snapshots
Hourly	25
Daily	8
Weekly	6
Monthly	13
Manually Taken Snapshots (If maximum number in all other categories are taken AND if there is sufficient	11

Snapspace.)	

The scheduler provides the widest range of coverage possible. Administrators will want to set up their systems to take scheduled daily Snapshots after the hourly capacity is reached, weekly Snapshots after daily capacity is reached, etc.

The scheduler does not quiesce volumes. Quiescing volumes must take place in the application software.

Hourly Snapshots

To schedule an Hourly Snapshot:

1 From the *NAS Volumes* screen, click **Snap Management**.

2 From the *Snap Management* screen Select **Schedule**.

3 From the Snapshot Schedule, select **Hourly**.

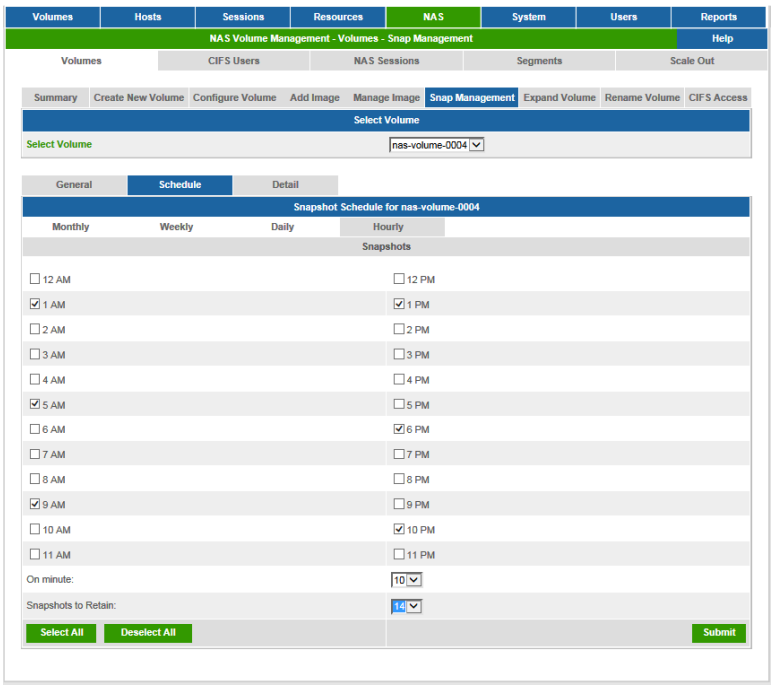


Figure 9-29 Schedule Hourly Snapshots Screen

4 Select the Hours desired by clicking the Checkboxes.

5 Select the minute desired from the drop down menu. All hourly Snapshots will all be taken at the same number of minutes after the hour.

6 Select the number of hourly Snapshots to retain. The maximum number of retained hourly Snapshots is 25. Once the number of hourly retained Snapshots has been reached, the system will automatically start overwriting snapshots to make room for new Snapshots.

7 Click **Submit.** The Snapshots will then be taken at the hours specified.

Daily Snapshots

To schedule a Daily Snapshot:

1 From the *NAS Volumes* screen, click **Snap Management**.

2 From the *Snap Management* screen Select **Schedule**.

3 From the Snapshot Schedule, select **Daily**.

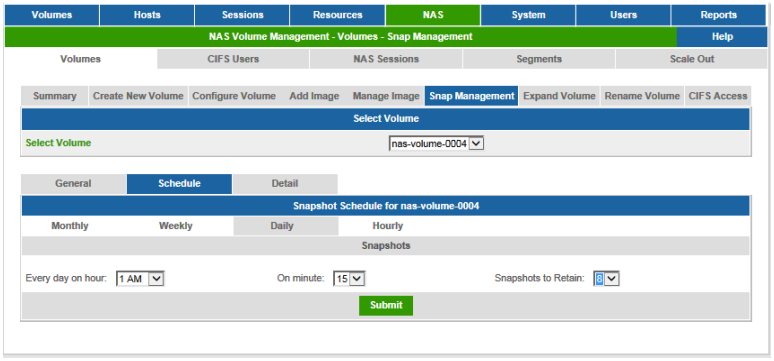


Figure 9-30 Schedule Daily Snapshots Screen

4 Select the Hour desired for the daily Snapshot using the dropdown menu.

5 Select the minute desired from the drop down menu. It is recommended that you do not select the same time as any other Scheduled Snapshot.

6 Select the number of Daily Snapshots to retain. The maximum number of retained Daily Snapshots is 8. Once the number of daily retained Snapshots has been reached, the system will automatically start overwriting Snapshots to create room for new ones.

7 Click Submit. The daily Snapshots will then be taken at the scheduled hour.

Weekly Snapshots

To schedule a Weekly Snapshot:

1 From the *NAS Volumes* screen, click **Snap Management**.

2 From the *Snap Management* screen Select **Schedule**.

3 From the Snapshot Schedule, select **Weekly.**

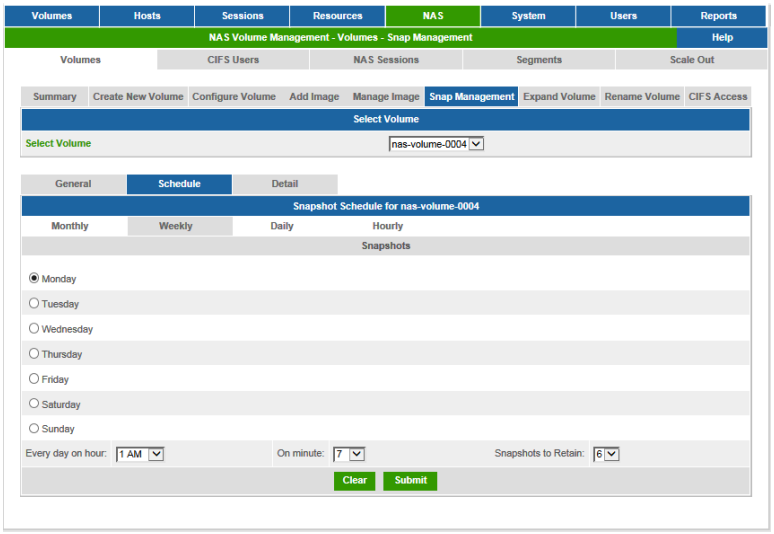


Figure 9-31 Schedule Weekly Snapshots

4 Select the checkboxes for the day desired for the weekly Snapshot. Then select the hour and minute desired using the drop down menus. It is recommended that you do not select the same time as any other scheduled Snapshot.

5 Select the number of weekly Snapshots to retain. The maximum number of retained weekly Snapshots is 6. Once the number of weekly retained Snapshots has been reached, the system will automatically start deleting Snapshots to make room for new Snapshots.

6 Click **Submit.** The Weekly Snapshots will then be taken at the days and hours specified.

Monthly Snapshots

To schedule a Monthly Snapshot:

1 From the *NAS Volumes* screen, click **Snap Management**.

2 From the Snap Management screen Select **Schedule**.

3 From the Snapshot Schedule, select **Monthly**.

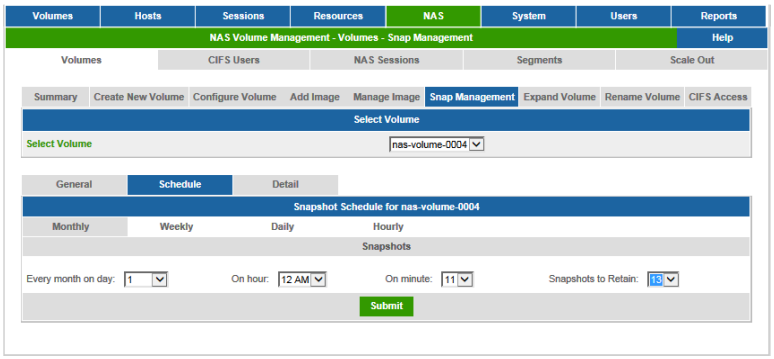


Figure 9-32 Schedule Monthly Snapshots

4 Select the day of the month desired from the dropdown menu. Then select the hour and minute desired using the drop down menus. It is recommended that you do not select the same time as any other scheduled Snapshot.

When you select a day for scheduled monthly snapshots, be sure to pick a day that is available all months. Not all months have all days (e.g. February normally has only 28).

5 Select the number of monthly Snapshots to retain. The maximum number of retained monthly Snapshots is 13. Once the number of monthly retained Snapshots has been reached, the system will automatically start deleting Snapshots to make room for new Snapshots.

6 Click **Submit.** The monthly Snapshots will then be taken at the days and hours specified.

9.10.4 Managing Snapshots

This function allows the user to remove a Snapshot image from a NAS Volume and free the resources that previously stored the Snapshot. The Snapshot will not be accessible to a host after it is removed.

9.10.4.1 Delete Snapshot

To delete a Snapshot:

1 From the *NAS Volumes* screen, click **Snap Management**.

2 Select NAS Volume from the drop-down menu.

3 Click on **Detail**.

4 From the Delete Snapshot column of the Snapshot detail table, select the image that you want to delete.

5 Click **Submit** to delete the Snapshot.

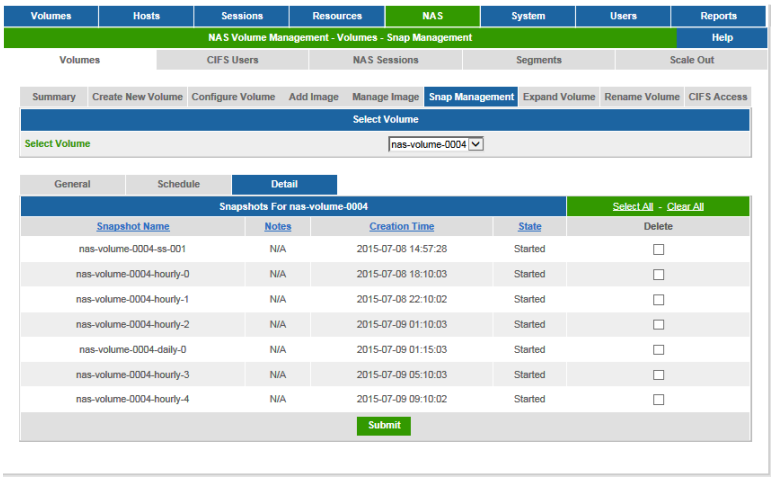


Figure 9-33 Snap Management Detail Screen

9.10.4.2 Snapshot Segments

Similar to NAS volume each NAS snapshot has segments on each node where the volume has its own snapshots. Snapshot segments are created and deleted by Storage Concentrator automatically when snapshot is created or deleted. User has ability to delete NAS snapshot segments that for some reason are not used by any NAS snapshot anymore. The segments are listed with all other NAS segments on *NAS Segments Summary* screen:

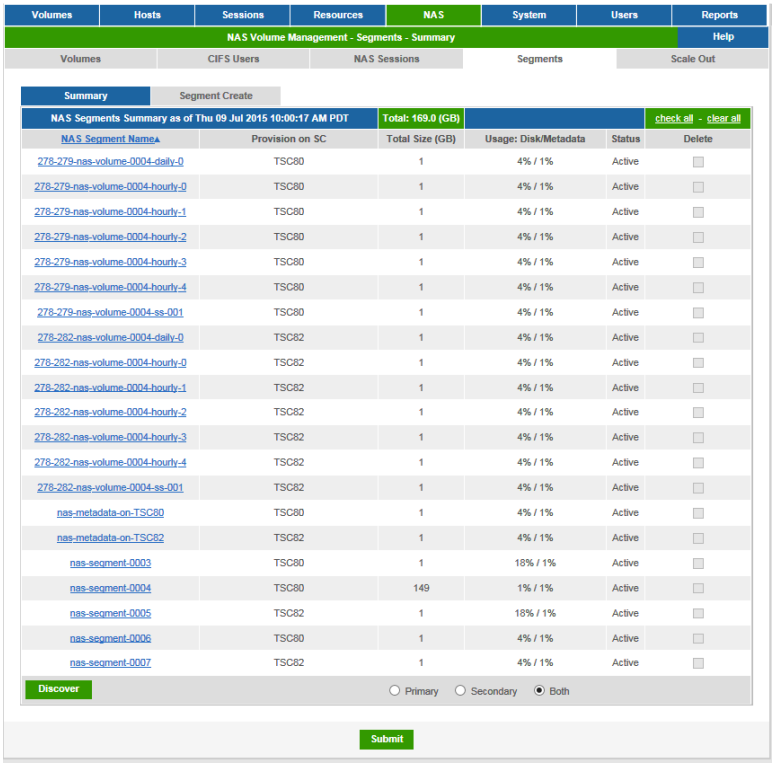


Figure 9-34 NAS Segments Summary Screen

Snapshot segments are based on iSCSI volume snapshots that are taken on each segment that NAS volume is provisioned on. As with snapshot segments these snapshots are managed by Storage Concentrator automatically. The snapshots are listed on *Volumes Configurations Snap Management Detail* screen:

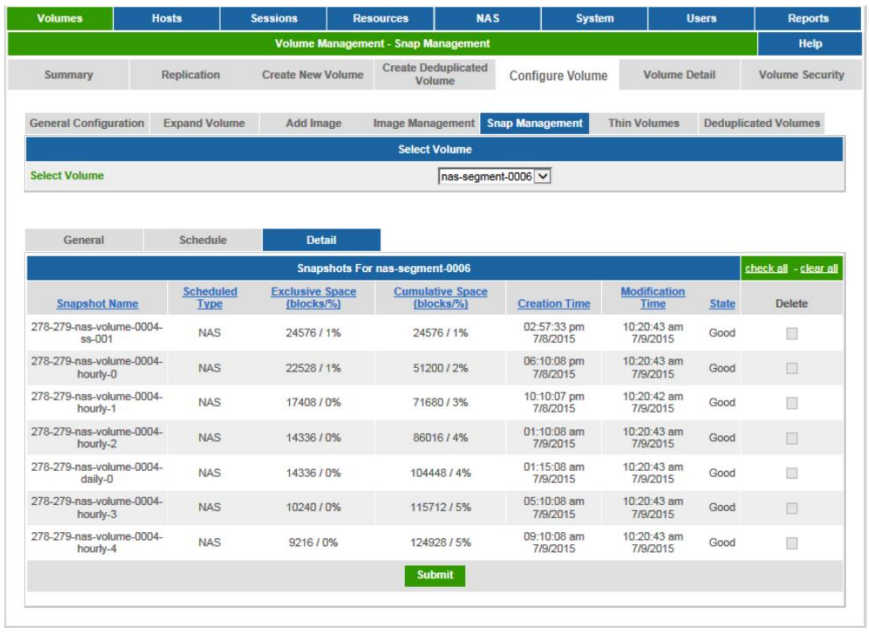


Figure 9-35 Segment Snapshots Detail Screen

9.10.5 Accessing Snapshots

To access contents of snapshot the snapshot has to be mounted on the same system where volume is mounted. User has to know name of the selected snapshot. Name can be found in text file "snapshots.txt" located at the volume's directory "\.volume-metadata". After that the snapshot should be mounted as "<Volume Name>\.snaps\<Snapshot Name>".

For example, if on Windows Network Drive(Y:) is mapped to volume by using folder name "\\10.10.63.84\nas-volume-0004":

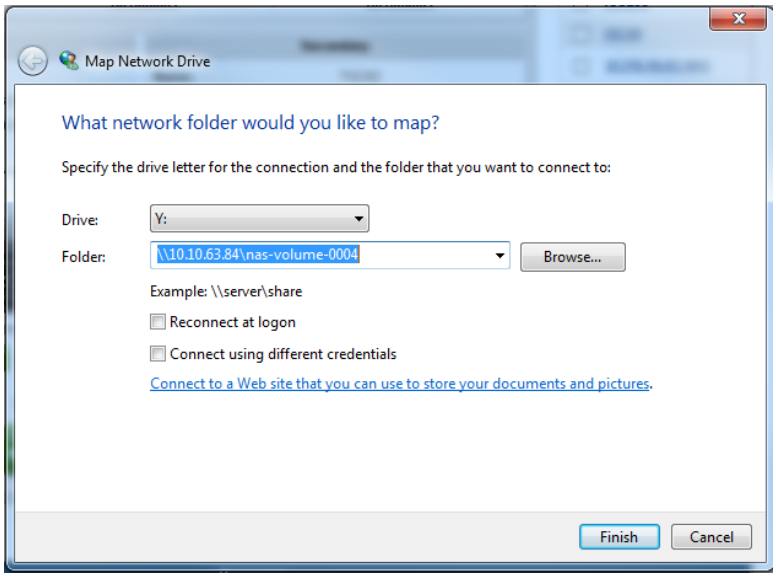
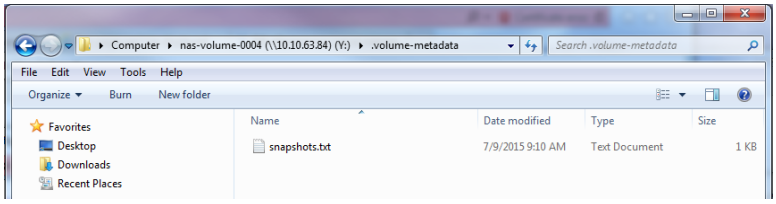


Figure 9-36 Mapping NAS Volume

Contents of the file "Y:.volume-metadata\snapshots.txt" are:



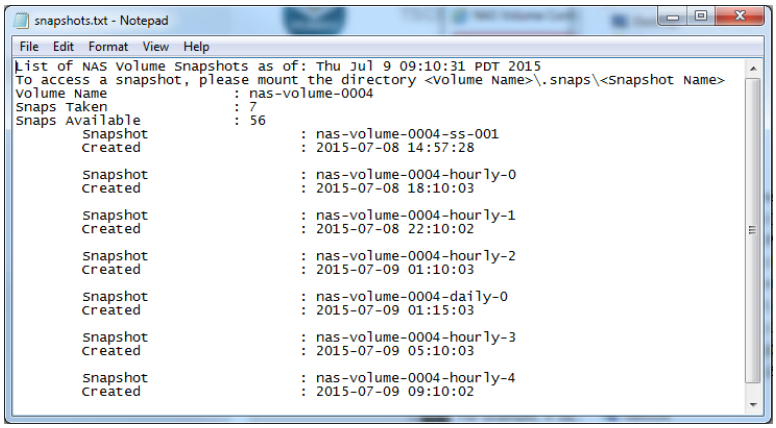


Figure 9-37 Contents of the "snapshots.txt"

Snapshot "nas-volume-0004-hourly-2" can mapped to Network Drive by using folder name "\\10.10.63.84\nas-volume-0004\.snaps\nas-volume-0004-hourly-2".

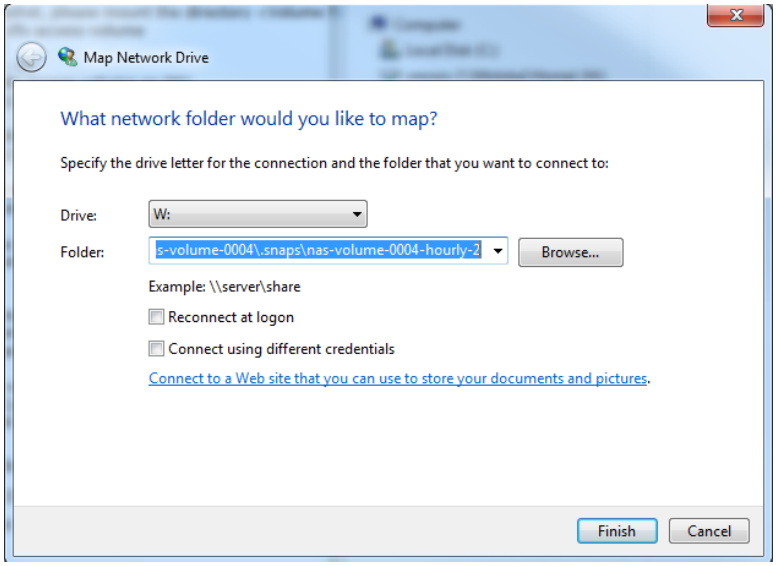


Figure 9-38 Mapping Snapshot

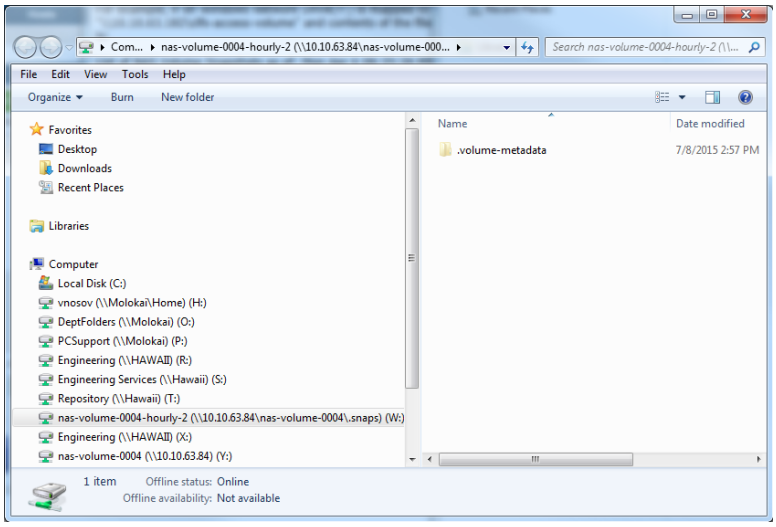


Figure 9-39 Browsing Volume and Snapshot

NAS snapshot is mounted as a read only volume.

There is no functionality to rollback NAS volume to specific point-in-time snapshot copy. User has only ability to restore files and folders by mounting snapshots and coping files to the volume.

It's important to remember that NAS volume snapshots are not preserved when standalone system is upgraded or its configuration is restored. NAS volume continues to be snap enabled and it generates new snapshots after upgrade regarding to snapshot schedule.

NAS volume snapshots are deleted when synchronous image is added to the volume as part of Storage Concentrator Cluster creation. Snapshots are deleted when Storage Concentrator Cluster is deleted or Secondary system is removed from Cluster

9.11 Directory Quotas

The Directory Quota functionality allows the user to set and manage limits on disk space used by directories or the volume. The feature has to be enabled first. After that the user could set quota configuration on specific directory. Absolute path of the directory has to be specified. Use path "/" to set limit on the volume. It will limit total disk space used by NFS and CIFS volume clients. Use path "/cifs_share" to limit disk space used by CIFS clients. Absolute path for directory "dir" accessible by CIFS client should be "/cifs_share/dir".

Two definitions are associated with quotas. The "Hard Limit" is the quota limit itself; the "Soft Limit" is a percentage of the hard limit. By default, the soft limit is 80% of the hard limit. The Directory Quota facility checks size of used space of volume directories with quota configuration and updates volume servers so the servers could enforce the limits. The default frequency for checks when the usage is below the soft limit is 60 seconds. When the usage is above the soft limit the default frequency for checks became 5 seconds. User has to be aware that servers let clients cross hard limit before size of the used space is updated on server side. For example, the used space is 95% of hard limit and client is writing to the directory 10% more data. If client able finish writes in less than 5 seconds and before the used space is updated, the final used space will be 5% above the hard limit. But the next writes made after update is executed will be rejected because the quota is reached already.

Quota limits set on volumes and directories are reported as volume or directory sizes by utilities that display the disk usage. For volumes and directories without quota configuration the physical size of volume or directory is reported as usually.

In Scale Out Configuration it could take up to minute update NAS volume quotas on all nodes after segment replacement. Do not be confused to see different list of quotas initially.

For each NAS volume the system generates messages stored at the Report Log with severity level "Warning" when there are changes in number of quota configurations with used size exceeding soft or hard limit. For detail information user has to open Directory Quota Management screen. These messages could be re-transmitted as E-mail or SNMP notifications. System monitors quotas periodically so the appropriate notification is generated post-factum with delay. The same message could be repeated not early than 24 hours after it is generated the previous time. In Scale Out Configuration NAS volume quotas are monitored for notification by Storage Concentrator where first segment of the volume is allocated. To avoid problem of checking different Report Logs user could set E-mail or SNMP notification on all Scale Out Nodes so all notification messages will be collected at the single destination.

If used space is approaching quota limit the next actions could be taken. System Administrator has option to delete quota that exceeded limit or to increase the quota's limit. Directory user could delete unneeded files from the quota's directory.

9.11.1 Enabling Quotas

From the NAS Volume Management Configure Volume menu select *Directory Quota* item. From the Select Volume drop down menu select the correct volume. Check the *Enabled* radio button. Click *Submit* button.

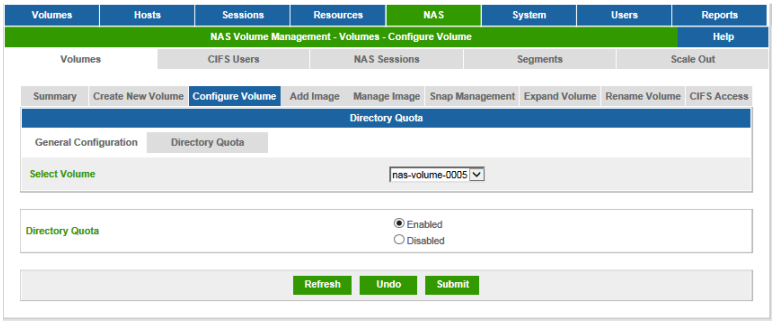


Figure 9-40 Enabling Directory Quotas

9.11.2 Setting Limits

For volume with enabled Directory Quota functionality the Directory Quota screen looks like the next one before the first quota configuration is set on the volume:

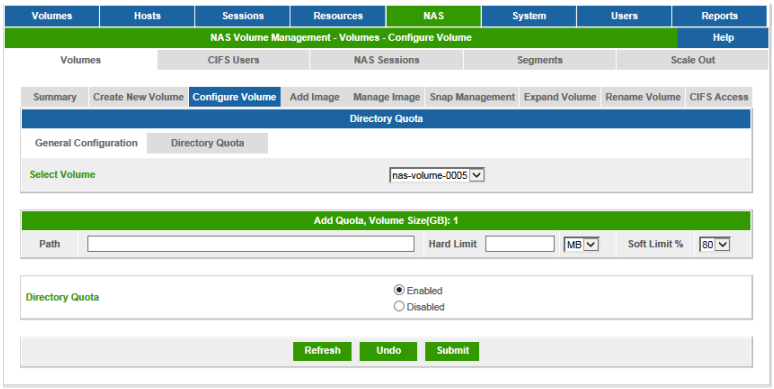


Figure 9-41 Setting Quota Limits

The next fields have to be set:

Path - Absolute path of the directory that is going to get quota configuration. The directory has to exist already except two cases. Use "/" to limit total size of the volume used space. Use "/cifs_share" to limit total size used by CIFS clients if the volume's CIFS export is enabled. The directory "/cifs_share" is created automatically when CIFS export is enabled. In all other cases request to set usage limit on nonexistent directory

fails. For existing directories use path "/Dir1" to limit size available to NFS clients at the directory "Dir1". Use path "/cifs_share/Dir2" to limit size available to CIFS clients at the directory "Dir2".

Hard Limit – Set used size limit on specific directory of the volume. It makes sense to set hard limit that is no more than total volume size. But any value could be used as well. Select units from the list. Possible options are KB, MB, GB, TB and PB. The hard limit could be set less than already used space. If this is the case it prevents clients from writing more data to the directory.

Soft Limit – Is percentage of the Hard Limit. It is used to adjust quota functionality when the used space size exceeds this value.

9.11.3 Modify Quotas

The Directory Quota screen can be used to modify or delete specific quotas. From the Select Volume drop down menu select volume. At the table of existing quotas find the correct quota. Use *Add Quota* area to modify the quota. Put the quota's directory name into the *Path* field. Set new *Hard Limit* and/or new *Soft Limit.* Click on button *Submit* button.

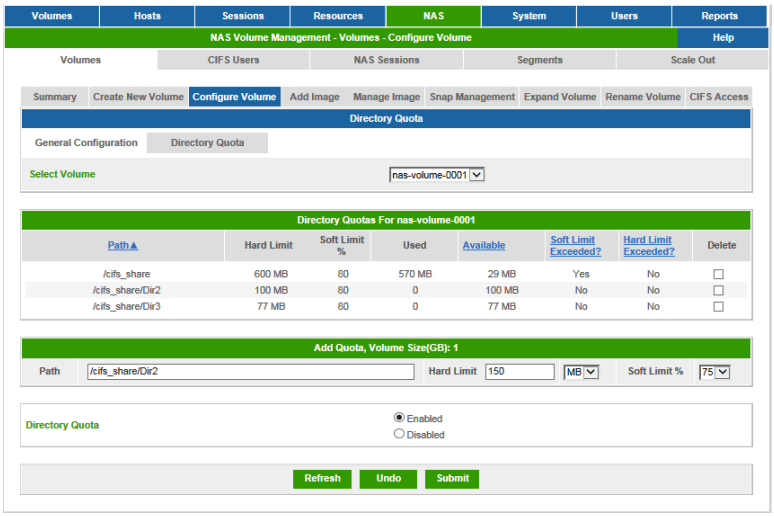


Figure 9-42 Modify Quota Limits

To delete the specific quotas select quotas at the *Delete* column and click on Submit button.

9.11.4 Disabling

Quota functionality for NAS volume can be disabled and all defined quota configurations will be deleted at the same time by selecting radio button *Disabled* and clicking button *Submit*.

Chapter 10

Scale Out NAS

Deployment of Scale Out NAS configuration provides user with ability to expand NAS volume capacity beyond single Storage Concentrator. Any node at Scale Out NAS configuration can be used by a NFS or CIFS client as a access point to volumes provisioned at the configuration.

Client session can be established manually by IP address or by using the Storage Concentrator DNS Server.

10.1 Create Scale Out Configuration

NAS Scale Out configuration can be created by using Storage Concentrators with valid NAS Volumes license. The systems have to be on the same data and management networks. NTP server has to be setup and NTP has to be enabled. System names have to be unique for Storage Concentrators selected to be a nodes at the configuration. The same requirement is applied to names of NAS volumes that are already provisioned on each individual Storage Concentrator before it can join configuration. The same accounts for CIFS export have to have the same password across all Storage Concentrators at the NAS Scale Out configuration.

Scale Out NAS configuration can be created by two ways. The first one is when one standalone Storage Concentrator is joining the other standalone Storage Concentrator. The second way is to use two standalone Storage Concentrators with enabled NAS Volumes license and create FailOver Cluster of two SC's. In the first case the result will be Scale Out Configuration that supports distributed NAS volumes only. In the second case configuration supports mirrored NAS volumes.

Any standalone Storage concentrator can join the initially created Scale Out configuration later. It provides user with ability to create configuration with required number of nodes.

Additional FailOver Clusters can be created by using available standalone nodes inside configuration. There is no requirement that all nodes have to be clustered, but clustering provides additional protection for Scale Out configuration in case of failure on one of the clustered nodes.

Two or more Scale Out configurations can't be merged together to create single configuration.

10.1.1 Join Scale Out Configuration

User has to open NAS Scale Out Summary screen on system that has to be a member of Scale Out configuration:

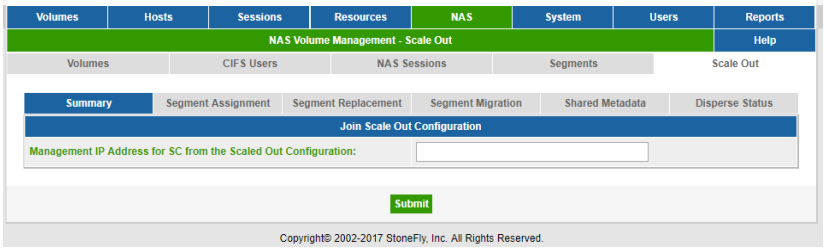


Figure 10-1

Screen on Figure 10-1 is specific for system that can join Scale Out configuration. If the system is already a member of Scale Out configuration the screen has different layout. See section 10.2.

To initiate request user has to provide Management IP Address for any Storage Concentrator that is already a member of the Scale Out Configuration or for standalone

Storage Concentrator that does not belong to any Scale Out configuration and hit **Submit** button. The selected member of the Scale Out Configuration has to be an active Storage Concentrator if it belongs to FailOver Cluster.

When system joined configuration successfully, the NAS Scale Out Summary screen has to get different layout and represents NAS Scale Out configuration monitoring screen.

10.1.2 FailOver Cluster as a Scale Out Configuration

FailOver Cluster of two Storage Concentrators with enabled NAS Volumes license represents case of Scale Out NAS configuration. Use standard procedure to create FailOver Cluster. See section 3.2. After the FailOver Cluster is created the NAS Scale Out Summary screen has to look like the next one:

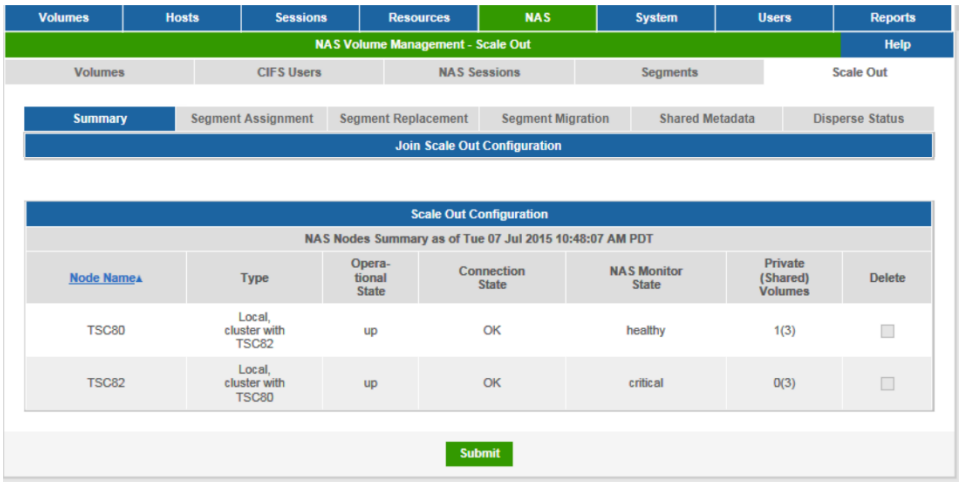


Figure 10-2

10.2 Node Monitoring and Management

The NAS Scale Out Configuration Summary screen provides status of configuration that the local Storage Concentrator belongs to.

If the local system does not belong to any Scale Out configuration the Summary screen provides ability for local system join the existing configuration. See section 10.1.1 for example.

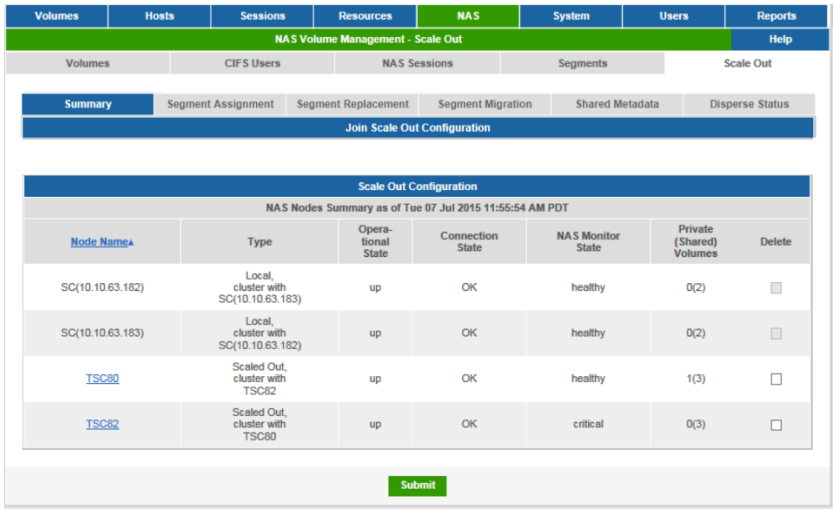


Figure 10-3

- **Node Name**: Name of Storage Concentrators that belong to the same Scale Out configuration as the local one.
- **Type**: Node type can be "local" for nodes that the current browser window represents the management interface to or "Scaled Out" for all other nodes. If node is cluster, cluster of one or is selected to be a secondary by other node, the node type has the appropriate note.
- **Operational State**: Node operational state. Can have one of the following states: "up", "down" and "no response".
- **Connection State**: A network connection state between local active node and the selected node. Can have one of the following states: "OK" and "failed".
- **NAS Monitor State**: Information is available only for nodes that are members of FailOver Clusters. The value can be "n/a", "unknown", "healthy", "critical", "failed" and "disabled" and represents the NAS Monitor state from the specific node.
- **Private (Shared) Volumes**: Private - total number of NAS volumes with segments allocated on the selected node only. Shared - total number of NAS volumes allocated on the selected node and on one or more other nodes from the Scale Out configuration.
- **Delete**: Click the check box to select this node to be deleted from the Scale Out configuration and hit "Submit" button when all selections for node deletion are done.

  Member of FailOver cluster can be deleted only after the FailOver cluster is deleted first.

A node cannot delete itself or member of its FailOver cluster; the check box is disabled in this case.

A node cannot be deleted if one or more NAS volumes have segments on the selected for deletion node and on some other nodes from the Scale Out configuration.

10.3 Segment Migration

The NAS volume segment migration procedure is effective way to move volume contents allocated on one node to different node at the same Scale Out configuration. The destination node has to have available space to handle the selected portion of volume's data. The procedure is executed without interrupting NAS service. Regarding to size of the volume data the reallocation can take significant time. The procedure is initiated by using one GUI screen, but progress of all currently active migration procedures has to be checked on the other GUI screen. To initiate migration user has to open management GUI on system where the destination segments have to be allocated.

If NAS volume is mirrored, the migration is done for each image separately and one after another.

If the destination system is rebooted during segment migration, all migrations have to be started again. User intervention is not required to do this. System will do it itself.

For mirrored volumes, segment migration can be done only with destination segments on Scale Out nodes that are members of FailOver Clusters.

Mirrored volumes have to be in sync before segment migration can be initiated.

For distributed volumes segment migration can be done only with destination segment on Scale Out nodes that are Standalone Storage Concentrators.

Segment migration is executed simultaneously for all NAS volumes that are sharing the source segment. It can take different time to finish the migration for different shares. It's enough to pick up one volume from the group and initiate migration for the selected one.

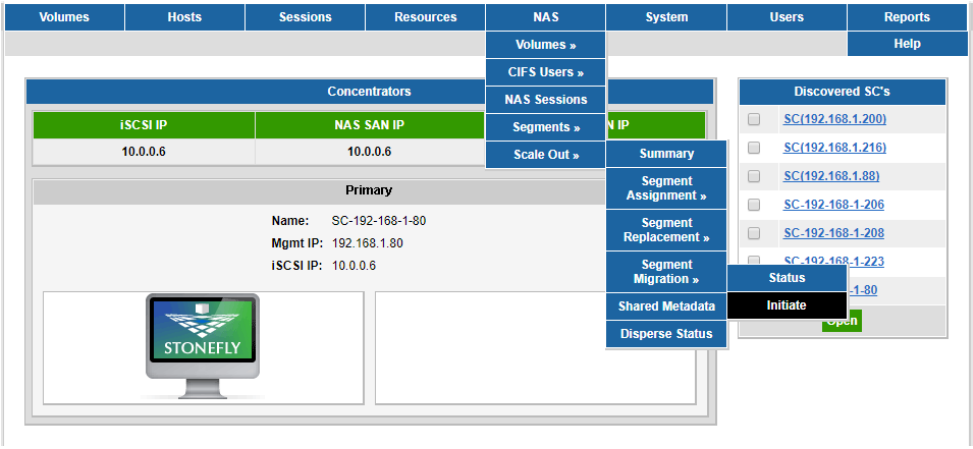
After migration is done, the source segment can be deleted.

Pre-existing NAS Segments that do not belong to other volumes can be used for migration.

No more than one migration process can be active for each NAS volume.

10.3.1 Initiate Segment Migration

User has to open "Segment Migration Initiate" screen on node that has to host destination segment for the migration. Navigate to NAS -> Scale Out -> Segment Migration -> **Initiate** as shown in the figure below.



"Select volume" from the drop down list. All appropriate NAS volumes that are configured in the Scale Out configuration have to be presented at the list.

"Select Scale Out node" from the drop down list. All nodes from the Scale Out configuration where the selected volume has allocated segments have to be presented at the list.

"Select Segment to Migrate" from the drop down list. For mirrored volumes both segment replicas have to be selected. For distributed volumes single segment has to be selected.

Click **Start** to proceed with segment selection to be used as the migration segments.

Use the NAS Segment drop down list(s) of available NAS Segments to select the appropriate one to use as a new allocation for the NAS Volume.
If the selection list is empty, the user will be redirected to "NAS Segment Create" GUI screen automatically. Adjust the name and set notes for the new NAS Segment during segment creation procedure. After the segment(s) are created, the GUI will return to the "Segment Migration Initiate" screen automatically and user may proceed with the migration.
In Storage Concentrator Clusters, the same sized segments must be used for both SC's for mirroring purposes.

Click **Submit** to start segment migration for the volume.

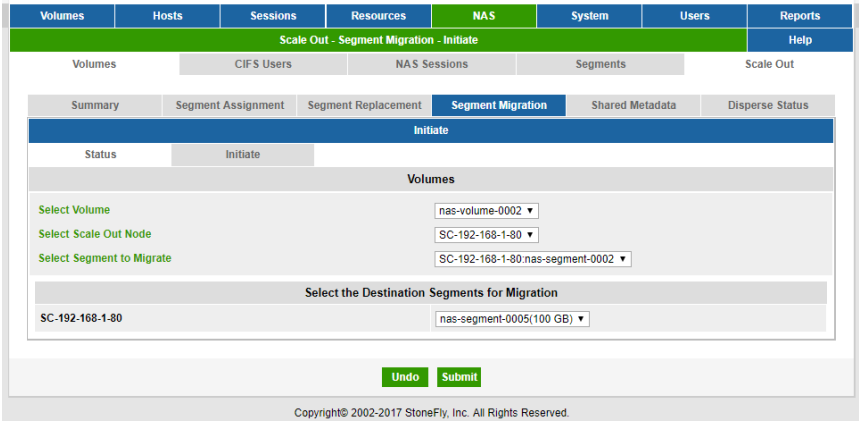


Figure 10-4

If segment migration is initiated successfully user will be redirected to "Segment Migration Status" GUI screen. See the next section 10.3.2.

10.3.2 Manage Segment Migration

Management for NAS segment Migration can be done from GUI opened on any active Storage Concentrator in the Scale Out configuration. The migration process can be aborted any time before user executed the final "Commit" request. After migration is committed volume stops to use source segments and starts to use destination segments. User can delete the source segments any time now or reuse them for different volumes.

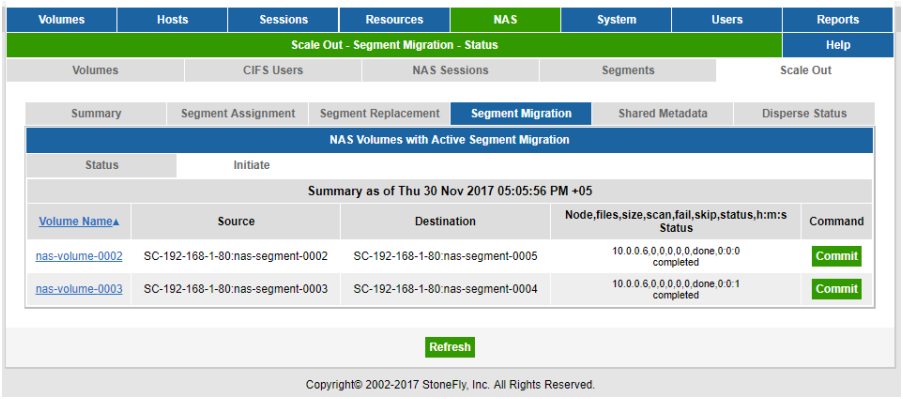


Figure 10-5

- **Volume Name**: NAS volumes with segment migration processes.
- **Source**: Node name and name of the segment selected for migration.
- **Destination**: Node name and name of the segment where data from the source segment has to be migrated.
- **Status**: Current status of the migration. The first lines are the detailed status of each segment, with the last line being the summary status of the whole migration process.
- **Command**: Only command that is appropriate for current state of segment migration is shown at the field. The command can be one of the following:

  - "Stop" for migrations that are in process of migrating files from the source segment to the destination segment. The "Status" field has to show how many files are migrated;
  - "Start" to continue the migration that had previously been stopped;

  - "Commit" to modify volume configuration and to let volume use the destination segment and to stop use the source segment. This command is available when migration of all files from the source segment to the destination segment is completed;

  - "Rollback" is used to restore original segment assignment for the mirrored volume. Segment migration can be aborted any time. If segment migration for mirrored volume is aborted at time of migrating files for the second image, the command "Rollback" is used to restore segment assignment for the first image. The first image segment replacement is committed at this moment already. This is the reason why there are needs for special command;

  - "n/a" for case of failure to detect current status of the segment migration. User should verify the state of the Scale Out nodes, and then try again.

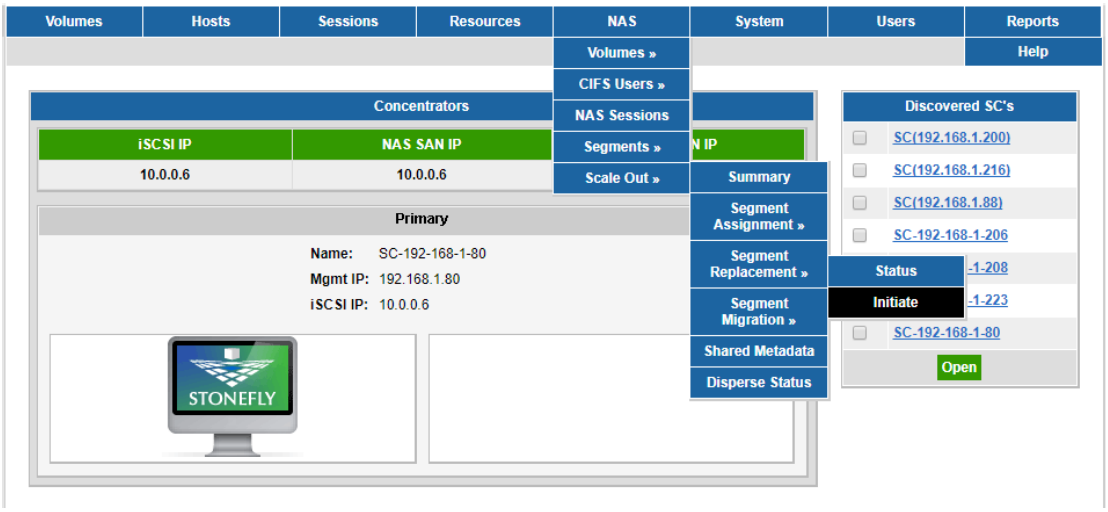
10.4 Segment Replacement

Difference between "Segment Migration" and "Segment Replacement" is that the "Segment Replacement" is not copying data from old segment to new segment. This means that the "Segment Replacement" can be used when old segment is not available. At the same time if "Segment Replacement" is used for Disperse or Replicated volumes, the volume is able to restore data on new segment without copying them from the old one - by invoking the volume healing procedure automatically.

The "Segment Migration" procedure is copying data from old segment to new segment.

10.4.1 Initiate Segment Replacement

User has to open "Segment Replacement Initiate" screen on node that has to host destination segment for the replacement. Navigate to NAS -> Scale Out -> Segment Replacement -> **Initiate** as shown in the figure below.



"Select volume" from the drop down list. All appropriate NAS volumes that are configured in the Scale Out configuration have to be presented at the list.

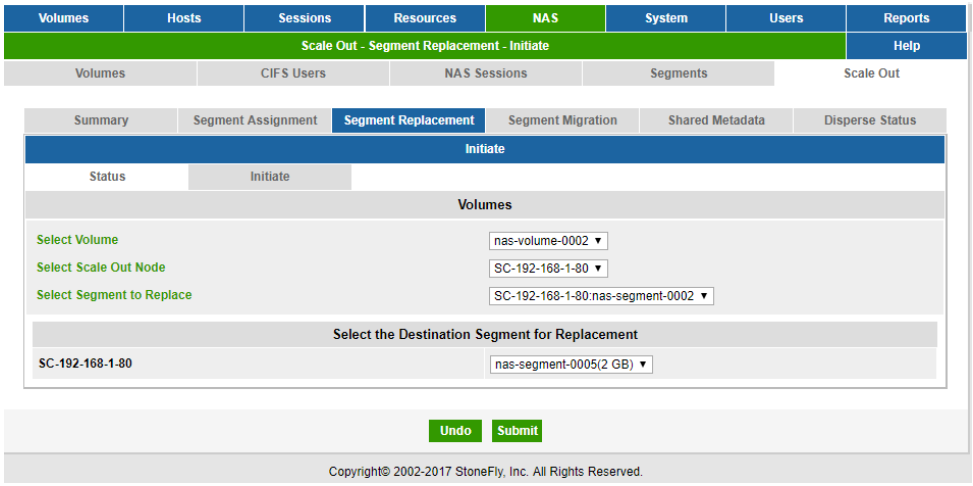
"Select Scale Out node" from the drop down list. All nodes from the Scale Out configuration where the selected volume has allocated segments have to be presented at the list.

"Select Segment to Replace" from the drop down list. For mirrored volumes both segment replicas have to be selected. For distributed volumes single segment has to be selected.

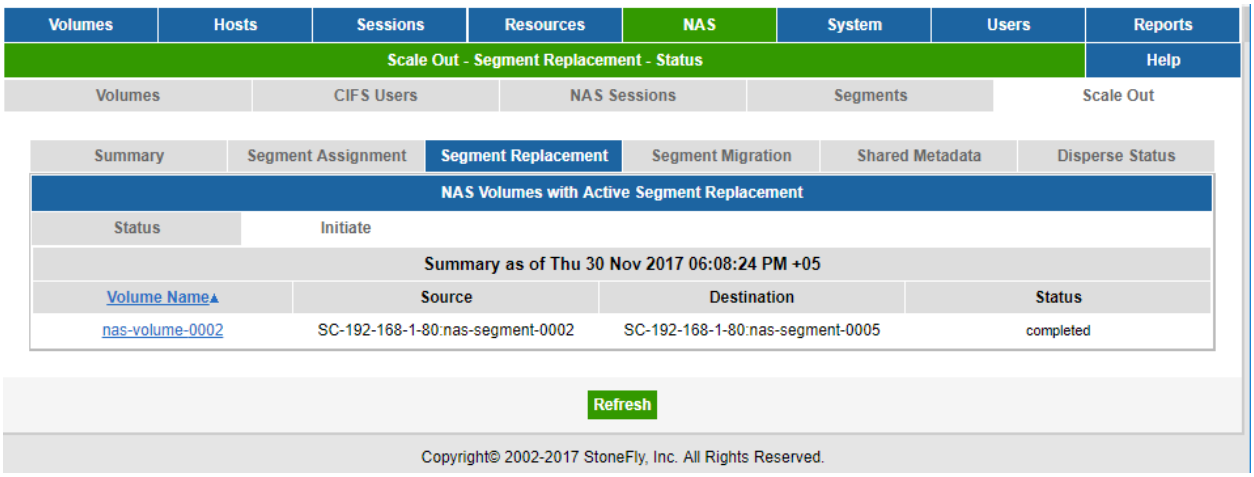
Click **Start** to proceed with segment selection to be used as the replacement segments.

Use the NAS Segment drop down list(s) of available NAS Segments to select the appropriate one to use as a new allocation for the NAS Volume.
If the selection list is empty, the user will be redirected to "NAS Segment Create" GUI screen automatically. Adjust the name and set notes for the new NAS Segment during segment creation procedure. After the segment(s) are created, the GUI will return to the "Segment Replacement Initiate" screen automatically and user may proceed with the replacement.
In Storage Concentrator Clusters, the same sized segments must be used for both SC's for mirroring purposes.

Click **Submit** to start segment replacement for the volume.



If segment replacement is initiated successfully user will be redirected to "Segment replacement Status" GUI screen. See the figure below.



10.5 NAS Volume Management

GUI screens used to manage and monitor NAS volumes on local systems are used to provide the same functionality for volumes in Scale Out configuration. See Chapter 9. Some obvious adjustments are done to show specific if this is the case for the volume. It's user responsibility to open GUI screen with administrative privilege and to use only one of them to manage the Scale Out Configuration. The other GUI screens can be used to manage SAN components of the each Scale Out node.

The "NAS Volumes Summary" GUI screen shows current status of all volumes provisioned in the Scale Out configuration. The next volume types can be presented on the screen. Type '**Distributed**' is used for volumes with segments provisioned on single Storage Concentrator or volumes expanded from single Storage Concentrator to any set of single Storage Concentrators at Scale Out configuration. Type '**Mirrored**' is used for volumes from FailOver Cluster or from set of FailOver Clusters at Scale Out configuration with NAS segments on both Storage Concentrators from the selected set of nodes. Type '**shared**' is used for volumes sharing all NAS segments with at least one other NAS volume. Type '**guest**' is used for volumes at Scale Out configuration without segments on local Scale Out node. Type '**local**' is used for volumes with segments allocated on local Scale Out node only. Type '**scale out local**' is used for volumes allocated on local Scale Out nodes and on some other nodes from the Scale Out configuration.

The "NAS Volume Expand" screen can be used to expand NAS volumes to other nodes in the Scale Out configuration. The expansion request has to be initiated on system where the new volume segments have to be allocated.

This page is intentionally left blank.

Chapter 11

Troubleshooting

11.1 Common Problems

Refer to the following table for helpful hints on solving some of the most common problems. If you cannot find a resolution to your problem below, contact StoneFly technical support. To contact StoneFly, call 510-265-1616 (Select Support), 24 hours a day, 7 days a week.

For FAQs and videos, go to the support section of the StoneFly web site at: www.stonefly.com (Requires a login account)

Please have the following information available when contacting technical support for assistance:

Model Number:_____

Serial Number:_____

Software Version:_____

Operating System:_____

Initiator Type and Version_____

Maintenance Contract:_____

Problem	Possible cause	Solution
The *Storage Concentrator* administrative interface is not accessible from the computer	The computer is not on the network.	Ensure that the computer is running a supported browser and is on the same network or subnet as the *Storage Concentrator.*
	An incorrect IP address was typed in the address line of the browser.	Try the following: • Verify and retype the IP address in the browser. • Make sure the IP address begins with **https** and not http.
	Incorrect network settings were entered during initial installation.	Refer to your *Storage Concentrator* Setup Guide, "Configuring the Network Settings" for more information.
	The network cable is not properly connected to the computer or the *Storage Concentrator.*	Try the following: • Check all network cable connections. • Make sure the network cable is not damaged
The *Storage Concentrator* cannot be configured using the factory supplied IP address.	The cable is not a network crossover cable.	Use a network crossover cable.
	The network crossover cable is plugged into the incorrect port on the back of the *Storage Concentrator*.	Plug the network crossover cable into the Management GbE port on the back of the *Storage Concentrator*. For more information, see the Setup Guide, Section 4, "Initial Installation".
	The computer is not configured to communicate in the 192.168.0.254 network.	Configure the computer to communicate in the 192.168.0.254 network. For more information, see the Setup Guide, Section 4, "Initial Installation"
During a FailOver, the hosts report "Delayed Write Errors"	The Windows 2000 disk "TimeOutValue" in the registry needs to be changed	Change the Disk "TimeOutValue" in the registry to 300 seconds

11.2 System Recovery

In the unlikely event that the system experiences a catastrophic failure, StoneFly has included a recovery CD with your system.

System information must have been saved with the Auto Save process prior to attempting a restore. For more information on the Auto Save procedure, see "Auto Save".

To recover the system, use the steps that follow:

1. If the *Storage Concentrator* is running, perform a shutdown of the *Storage Concentrator*.
2. Remove the front panel bezel from the *Storage Concentrator*.
3. Insert the recovery CD into the CD-ROM drive.
4. Power up the *Storage Concentrator*.

The *Storage Concentrator* will boot from the CD ROM and will perform a complete install of the software. Any existing user configuration data will be overwritten. The CD ROM is ejected when the load of the software is complete. Remove the CD ROM and restart the *Storage Concentrator*.

6. To update the *Storage Concentrator* with the user configuration data, complete the restore process, using the information under "Restore".

Chapter 12

Service Menu

12.1 Introduction

The Service menu command line interface configures the system parameters of the *Storage Concentrator*. The purpose of the Service menu is to recover *Storage Concentrators* that have unknown IP addressing. The Service menu should only be used when troubleshooting configuration problems with the *Storage Concentrator*.

Ensure that all users are logged out of the *Storage Concentrator* Administrative Interface before accessing the Service menu. When the Service menu is accessed any users in the system will be logged off without warning. Use of the GUI *Storage Concentrator* Administrative Interface is not possible until one is logged out of the command line Service Menu.

The Service menu has three modes. The mode you see and the options that are available are determined by whether a *Storage Concentrator* is Primary, Secondary, or in FailOver mode. The three Service menu modes are:

Storage Concentrator Stand-alone Mode: Appears when the *Storage Concentrator* is the Primary *Storage Concentrator.*

Storage Concentrator Secondary Mode: Appears when the *Storage Concentrator* is the Secondary *Storage Concentrator.*

Storage Concentrator FailOver Mode: Appears when the Primary *Storage Concentrator* is in FailOver mode.

12.2 Accessing the Service Menu

To access the Service menu, use the steps that follow:

1 Power on the Storage Concentrator.

- Use CRT: Connect a CRT monitor and a keyboard to the Storage Concentrator.

- or -

- Use Serial Port: Connect a computer with a terminal emulation program, such as Hyper Terminal, to the serial port using a null modem cable.

  The terminal emulation program must be set up as follows: Terminal Mode = VT - 100, Bits/Sec = 9600, Data Bits = 8, Parity = no parity, Stop Bits = 1 stop bit.

3 Press **Enter** to display the login screen. The *Storage Concentrator* Login screen appears.

4 At the user ID prompt type **console**.

5 At the password prompt type **coni100o**.

The user ID and password are case sensitive.

12.3 Using Storage Concentrator Stand-Alone Mode

When the *Storage Concentrator* is in stand-alone mode, the Stand-alone Mode Service menu appears with the following options:

1 **Admin**: Provides administrative functions while in the Service menu. The system administrator can reboot, shutdown, save configuration information, enable or disable Telnet, and restore configuration information.

2 **Network**: Allows the system administrator to configure the default gateway setting and the Management GbE port settings for IP address, Netmask, Network, and Broadcast settings for the *Storage Concentrator*.

q **Logout**: Logs the user out of the Stand-alone Mode Service menu.

Ensure that all Telnet sessions are ended before logging out. The logout function will terminate any active Telnet sessions.

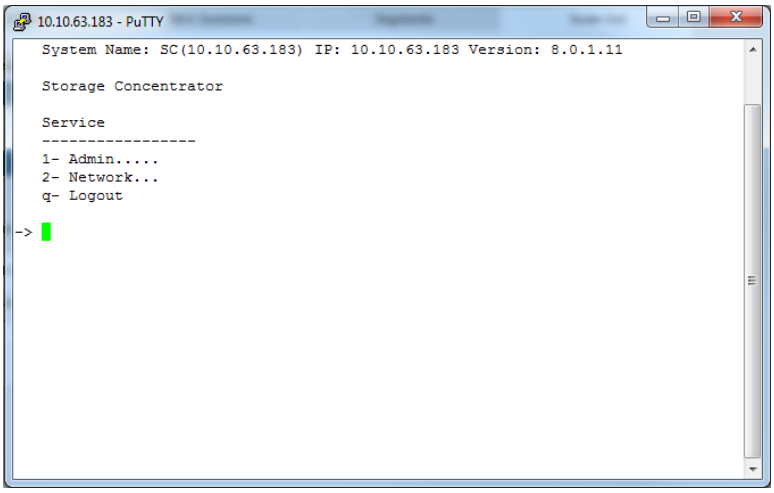


Figure 12-1

12.3.1 Using the Stand-Alone Mode Admin Menu

The Stand-alone Mode Admin menu allows the system administrator to reboot, shut down, save configuration information, enable or disable Telnet, and restore configuration information.

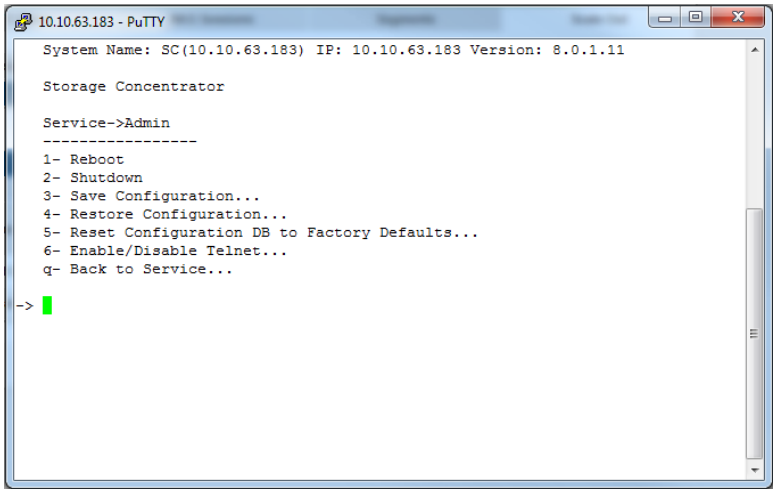


Figure 12-2

To reboot the *Storage Concentrator* from the Stand-alone Mode Admin menu, use the steps that follow:

1. From the Stand-alone Mode Service menu, type **1** and press **Enter**. The Stand-alone Mode Admin menu appears.
2. To select reboot, type **1** and press **Enter**. The following prompt appears:

   About to reboot, are you sure (y/n)?

3. Type **Y** and press **Enter** to reboot. The *Storage Concentrator* reboots.

To shut down the *Storage Concentrator* from the Stand-alone Mode Admin menu, use the steps that follow:

1. From the Stand-alone Mode Service menu, type **1** and press **Enter**. The Stand-alone Mode Admin menu appears.
2. To select shutdown, type **2** and press **Enter**. The following prompt appears: About to shutdown, are you sure (y/n)?
3. Type **Y** and press **Enter** to shut down. The *Storage Concentrator* will shut down. (The High Availability Storage Concentrator, or HSC, does not power off during a shutdown. Wait until all activity lights are quiet and use the power switch to turn it off.)

To save the *Storage Concentrator* configuration information from the Stand-alone Mode Admin menu, use the steps that follow:

1 From the Stand-alone Mode Service menu, type **1** and press **Enter**. The Stand-alone Mode Admin menu appears.

2 To select Save Configuration, type **3** and press **Enter**.

3 Select one of the following options:

2 - To Floppy Drive: Saves to a floppy disk.
q - Back to System: Exits the Stand-alone Mode Admin menu and returns to the Stand-alone Mode Service menu.

If 2 is selected, the following message appears when the save is complete:

Done (Type Enter)

4 Press **Enter**.

To restore the *Storage Concentrator* configuration information from the Stand-alone Mode Admin menu, use the steps that follow:

1 From the Stand-alone Mode Service menu, type **1** and press **Enter**. The Stand-alone Mode Admin menu appears.

2 To select Restore Configuration, type **4** and press **Enter**.

3 Select one of the following options and press **Enter**:

1 **From Floppy Drive**: Restores from a floppy disk.

q **Back to System**: Exits the Stand-alone Mode Admin menu and returns to the Stand-alone Mode Service menu.

If 1 is selected, the following message appears:

This will reboot the system, continue (y/n)?

4 Type **Y** to start the restore. The configuration information is restored and the system is automatically rebooted when the restore is complete.

To reset the *Storage Concentrator* configuration database to the Factory Default settings, select the "Reset Configuration DB to Factory Defaults..." option:

About to reset this SC's configuration database back to factory defaults. All volumes and resources will be deleted. If this SC is a member of a cluster, the other SC's configuration is not changed. The system is restarted.
Resetting to Factory Defaults -- are you sure? (y/n)

To enable or disable a Telnet session from the Stand-alone Mode Admin menu, use the steps that follow:

1 From the Stand-alone Mode Service menu, type **1** and press **Enter**. The Stand-alone Mode Admin menu appears.

2 To select Enable/Disable Telnet, type **5** and press **Enter**.

3 Select one of the following options and then press **Enter**:

- **Enable Telnet temporarily**: Allows anyone to Telnet into the *Storage Concentrator*. The temporary access will be disabled as soon as the user quits out of the Service Menu.
- **Disable Telnet**: Prevents the establishment of any new Telnet sessions.
- **Back to Admin**: Exits this menu and returns to the Stand-alone Mode Admin menu.

12.3.2 Using the Stand-Alone Mode Network Menu

The Stand-alone Mode Network menu allows the system administrator to configure the default gateway setting and the Management GbE port settings for IP address, Netmask, Network, and Broadcast settings for the *Storage Concentrator.*

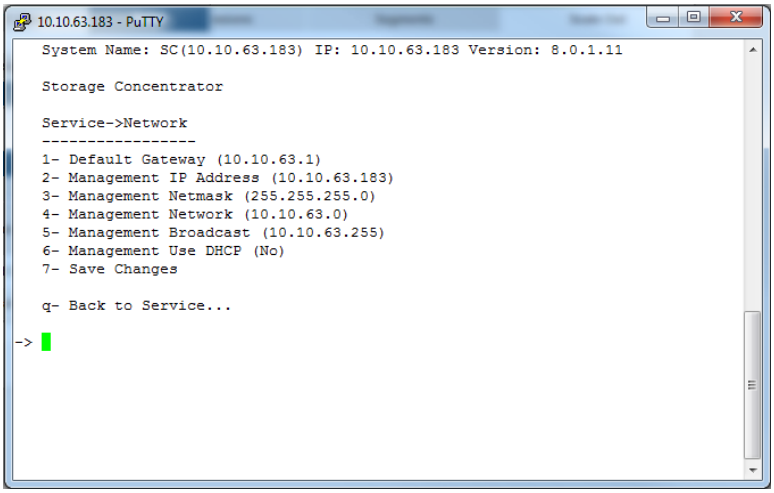


Figure 12-3

To configure the network settings, use the steps that follow:

1 From the Stand-alone Mode Service menu, type **2**. The Stand-alone Mode Network menu appears with the current network settings displayed in parentheses.

2 Select one of the following options:

1 - **Default Gateway**

2 - **Management GbE IP Address**

3 - **Management GbE Netmask**

4 - **Management GbE Network**

5 - **Management GbE Broadcast**

q - Logout: Type **q** to log out of the Stand-alone Mode

Service menu

3 If 1 - 5 are selected, type the appropriate network setting and press **Enter**. The new network settings are displayed on the Stand-alone Mode Network menu.

If an invalid entry is made, a warning message appears and prompts you to enter a valid setting.

4 When finished, press **q** to logout.

Ensure that all Telnet sessions are ended before logging out. The logout function ends any active Telnet sessions without warning.

12.4 Using Storage Concentrator Secondary Mode

When the *Storage Concentrator* is in Secondary mode, the Secondary Mode Service menu appears with the following options:

1 - Admin: Provides administrative functions while in the Service menu. The system administrator can reboot, shut down, save configuration information, enable or disable Telnet, and restore configuration information.

3 - Cluster: Allows the system administrator to make this *Storage Concentrator* a stand-alone system or to force this *Storage Concentrator* into FailOver mode.

q - Back to Service: Returns to the Secondary Mode Service Menu.

Ensure that all Telnet sessions are ended before logging out. The logout function will end any active Telnet sessions without warning.

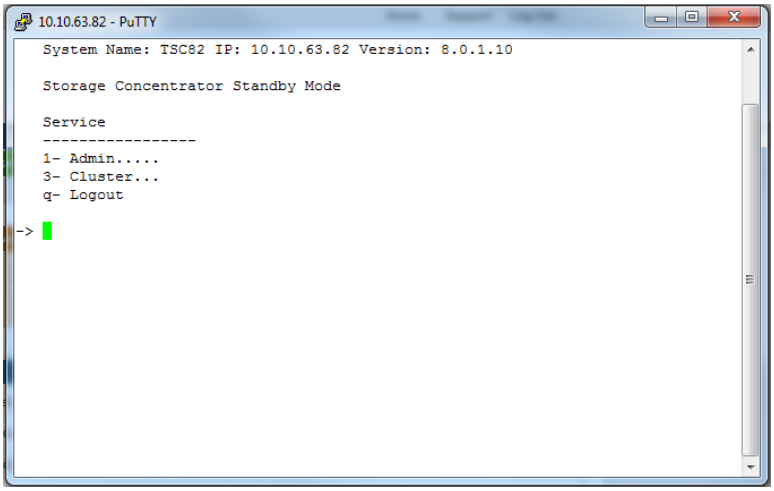


Figure 12-4

12.4.1 Using the Admin Menu Secondary Mode

The Secondary Mode Admin menu allows the system administrator to reboot, shutdown, and enable or disable Telnet.

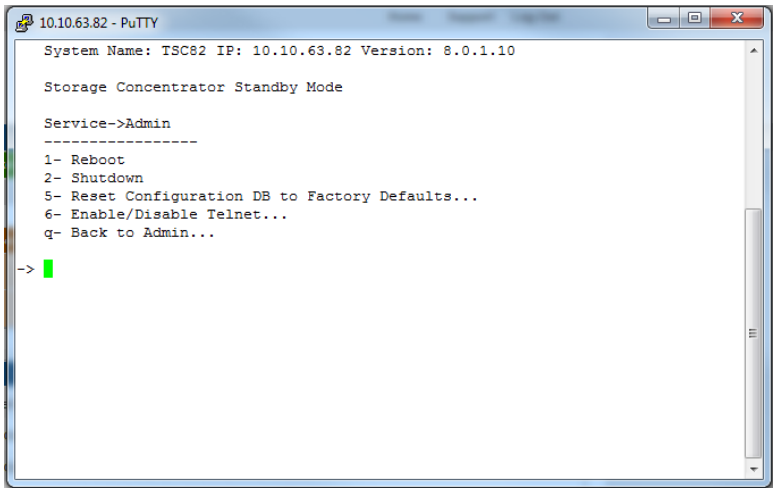


Figure 12-5

To reboot the *Storage Concentrator* from the Secondary Mode Admin menu, use the steps that follow:

1 From the Secondary Mode Service menu, type **1** and press **Enter**. The Secondary Mode Admin menu appears.

2 To select reboot, type **1** and press **Enter**. The following prompt appears: About to reboot, are you sure (y/n)?

3 Type **Y** and press **Enter** to reboot. The *Storage Concentrator* reboots.

To shut down the *Storage Concentrator* from the Secondary Mode Admin menu, use the steps that follow:

1 From the Secondary Mode Service menu, type **1** and press **Enter**. The Secondary Mode Admin menu appears.

2 To select shutdown, type **2** and press **Enter**. The following prompt appears: About to shut down, are you sure (y/n)?

3 Type **Y** and press **Enter** to shut down. The *Storage Concentrator* will shut down. (The High Availability Storage Concentrator, or HSC, does not poweroff during a shutdown. Wait until all activity lights are quiet and use the power switch to turn it off.)

To enable or disable a Telnet session from the Secondary Mode Admin menu, use the steps that follow:

1 From the Secondary Mode Service menu, type **1** and press **Enter**. The Secondary Mode Admin menu appears.

2 To select Enable/Disable Telnet, type **4** and press **Enter**.

3 Select one of the following options and then press **Enter**:

1 - Enable Telnet temporarily: Allows anyone to Telnet into the *Storage Concentrator*.

2 - Disable Telnet: Ends all Telnet sessions.

q - Back to Admin: Exits this menu and returns to the Secondary Mode Admin menu.

12.4.2 Using the Secondary Mode Cluster Menu

The Secondary Mode Cluster menu allows the system administrator to remove this *Storage Concentrator* from the cluster and make it a stand-alone system or to force this *Storage Concentrator* into FailOver mode.

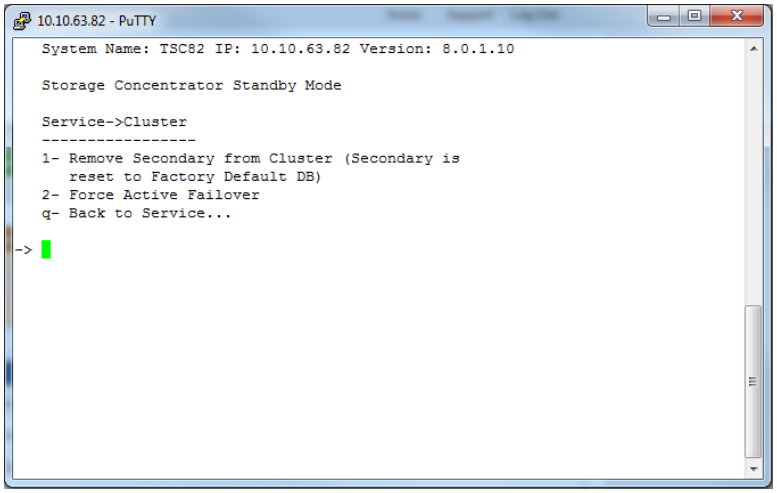


Figure 12-6

To remove this *Storage Concentrator* from the cluster and make it a stand-alone system, from the Cluster menu, use the steps that follow:

1 From the Secondary Mode Service menu, type **3** and press **Enter**. The Secondary Mode Cluster menu appears.

2 To remove the *Storage Concentrator* from the cluster and make it a stand-alone system, type **1** and press **Enter**. The following prompt appears:

About to switch to stand-alone, are you sure (y/n)?

3 Type **Y** and press **Enter** to make the *Storage Concentrator* a stand-alone system. The *Storage Concentrator* is now a stand-alone system.

To force this *Storage Concentrator* into FailOver mode, use the steps that follow:

1. From the Secondary Mode Service Menu, type **3** and press **Enter**. The Secondary Mode Cluster menu appears.
2. To force this *Storage Concentrator* into FailOver mode, type **2** and press **Enter**. The following prompt appears:

   About to force a FailOver, are you sure (y/n)?

3. Type **Y** and press **Enter** to force this *Storage Concentrator* into FailOver mode. The *Storage Concentrator* is now in FailOver mode.

To return to the Secondary Mode Service menu from the Secondary Mode Cluster menu, use the steps that follow:

Press **q - Back to Service**: Exits the Secondary Mode Cluster menu and returns to the Secondary Mode Service menu.

12.5 Using Storage Concentrator FailOver Mode

When the *Storage Concentrator* is in FailOver Mode, the FailOver Mode Service menu appears with the following options:

1 - Admin: Provides administrative functions while in the Service menu. The system administrator can reboot, shut down, save configuration information, enable or disable Telnet, and restore configuration information.

2 - Network: Allows the system administrator to configure the default gateway setting and the Management GbE port settings for IP address, Netmask, Network, and Broadcast settings for the *Storage Concentrator*.

3 - Cluster: Allows the system administrator to make this *Storage Concentrator* a stand-alone system.

q - Logout: Logs the user out of the FailOver Mode Service menu.

Ensure that all Telnet sessions are ended before logging out. The logout function will end any active Telnet sessions without warning.

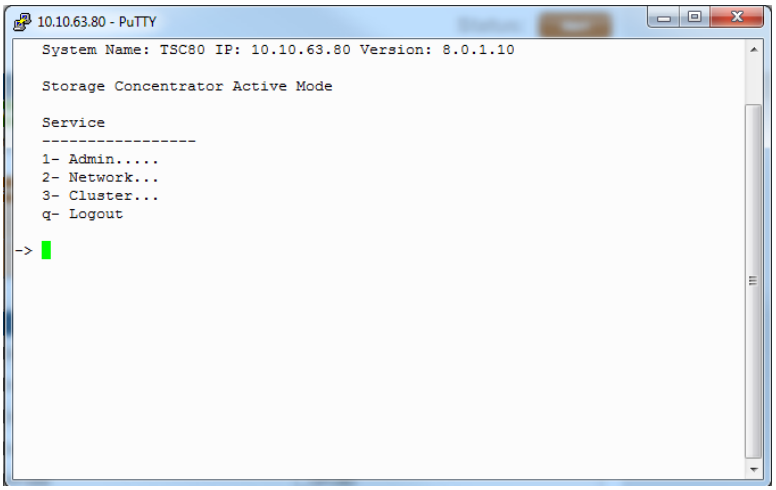


Figure 12-7

12.5.1 Using the FailOver Mode Admin Menu

The FailOver Mode Admin menu allows the system administrator to reboot, shut down, save configuration information, enable or disable Telnet, and restore configuration information.

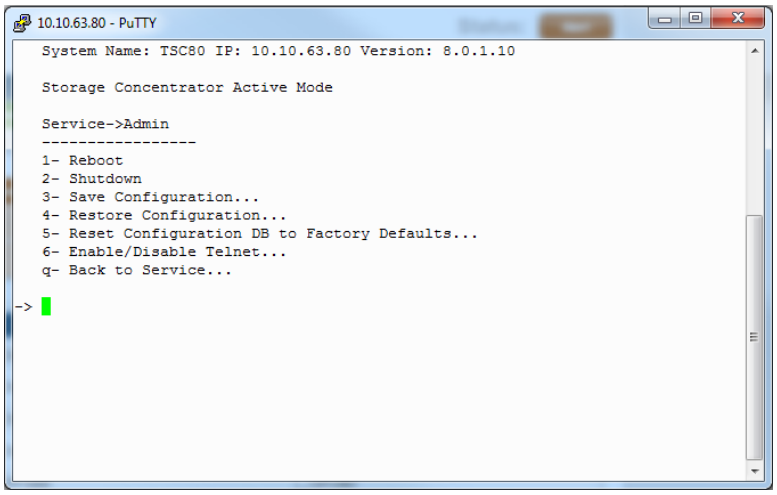


Figure 12-8

To reboot the *Storage Concentrator* from the FailOver Mode Admin menu, use the steps that follow:

1. From the FailOver Mode Service menu, type **1** and press **Enter**. The FailOver Mode Admin menu appears.
2. To select reboot, type **1** and press **Enter**. The following prompt appears: About to reboot, are you sure (y/n)?
3. Type **Y** and press **Enter** to reboot. The *Storage Concentrator* reboots.

To shut down the *Storage Concentrator* from the FailOver Mode Admin menu, use the steps that follow:

1. From the FailOver Mode Service menu, type **1** and press **Enter**. The FailOver Mode Admin menu appears.
2. To select shut down, type **2** and press **Enter**. The following prompt appears:

   About to shut down, are you sure (y/n)?
3. Type **Y** and press **Enter** to shut down. The *Storage Concentrator* will shut down.

To save the *Storage Concentrator* configuration information from the FailOver Mode Admin menu, use the steps that follow:

1. From the FailOver Mode Service menu, type **1** and press **Enter**. The FailOver Mode Admin menu appears.
2. To select Save Configuration, type **3** and press **Enter**.

3 Select one of the following options:

2 - To Floppy Drive: Saves to a floppy disk.

q - Back to Service: Exits the FailOver Mode Admin menu and returns to the FailOver Mode Service menu. If option 2 is selected, the following message appears when the save is complete: Done (Type Enter)

4 Press **Enter**.

To enable or disable a Telnet session from the FailOver Mode Admin menu, use the steps that follow:

1 From the FailOver Mode Service menu, type **1** and press **Enter**. The FailOver Mode Admin menu appears.

2 To select Enable/Disable Telnet, type **4** and press **Enter**. 3 Select one of the following options and then press **Enter**:

1 - Enable Telnet temporarily: Allows anyone to Telnet into the *Storage Concentrator*. The Telnet session will end when the user quits the service menu.

2 - Disable Telnet: Ends all Telnet sessions.
q - Back to Service: Exits the FailOver Mode Admin menu and returns to the FailOver Mode Service menu.

To restore the *Storage Concentrator* configuration information from the FailOver Mode Admin menu, use the steps that follow:

1 From the FailOver Mode Service menu, type **1** and press **Enter**. The FailOver Mode Admin menu appears.

2 To select Restore Configuration, type **5** and press **Enter**.

3 Select one of the following options and press **Enter**:

2 - From Floppy Drive: Restores from a floppy disk.

q - Back to Service: Exits the FailOver Mode Admin menu and returns to the FailOver Mode Service menu.

If option 2 is selected, the following message appears:

This will reboot the system, continue (y/n)?

3 Type **Y** to start the restore. The configuration information is restored and the system is automatically rebooted when the restore is complete.

12.5.2 Using the FailOver Mode Network Menu

The FailOver Mode Network menu does not allow the system administrator to make any changes to the default gateway setting or the Management GbE port settings for IP address, Netmask, Network, and Broadcast settings for the *Storage Concentrator.* The following message appears:

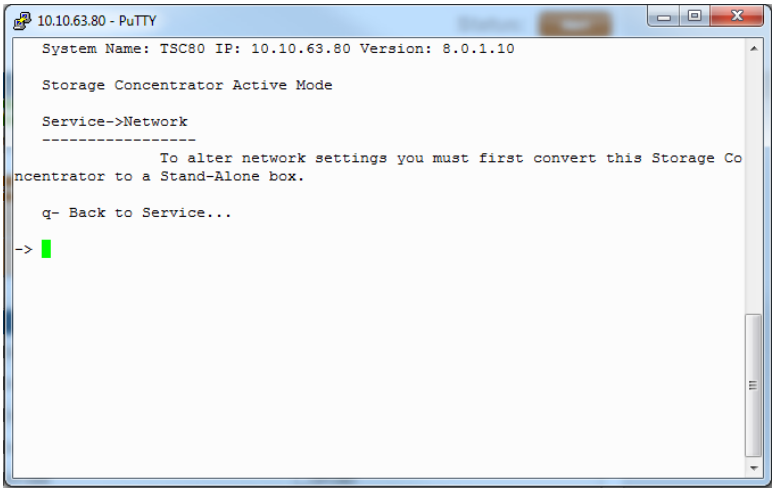


Figure 12-9

12.6 Using the FailOver Mode Cluster Menu

The FailOver Mode Cluster menu allows the system administrator to make this *Storage Concentrator* a stand-alone system.

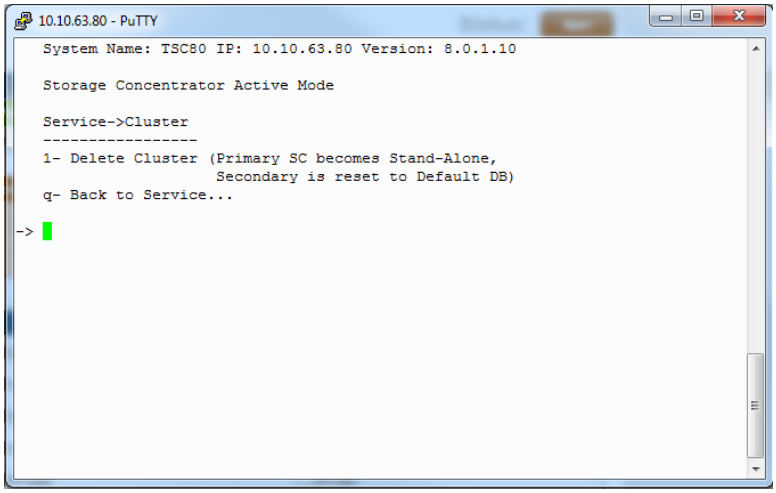


Figure 12-10

To delete the cluster and make both cluster members into stand-alone systems, use the steps that follow:

1. From the FailOver Mode Service Menu, type **3** and press **Enter**. The FailOver Mode Cluster menu appears.
2. To delete the cluster and make it a stand-alone system, type **1** and press **Enter**. The following prompt appears:

   About to switch to stand-alone, are you sure (y/n)?

3. Type **Y** and press **Enter** to make the *Storage Concentrator* a stand-alone system. The *Storage Concentrator* is now a stand-alone system.

To return to the FailOver Mode Service menu from the FailOver Mode Cluster menu, use the step that follows

Press **q - Back to Service**: Exits the FailOver Mode Cluster menu and returns to the FailOver Mode Service menu.

This page is intentionally left blank.

Chapter 13

Flash Cache

The Flash Cache function allows create cache device for the selected Managed or SF Managed resource. Depending on caching mode any read and write IO's to the resource are handled by the caching device first. The functionality is available only on standalone Storage Concentrators with valid Flash Cache license.

13.1 Selecting Resource

To create Flash Cache Device one of Storage Concentrator resources has to be designated as a device to serve as a buffer for IO's handled by the Flash Cache. Use Resource Summary screen to make this assignment.

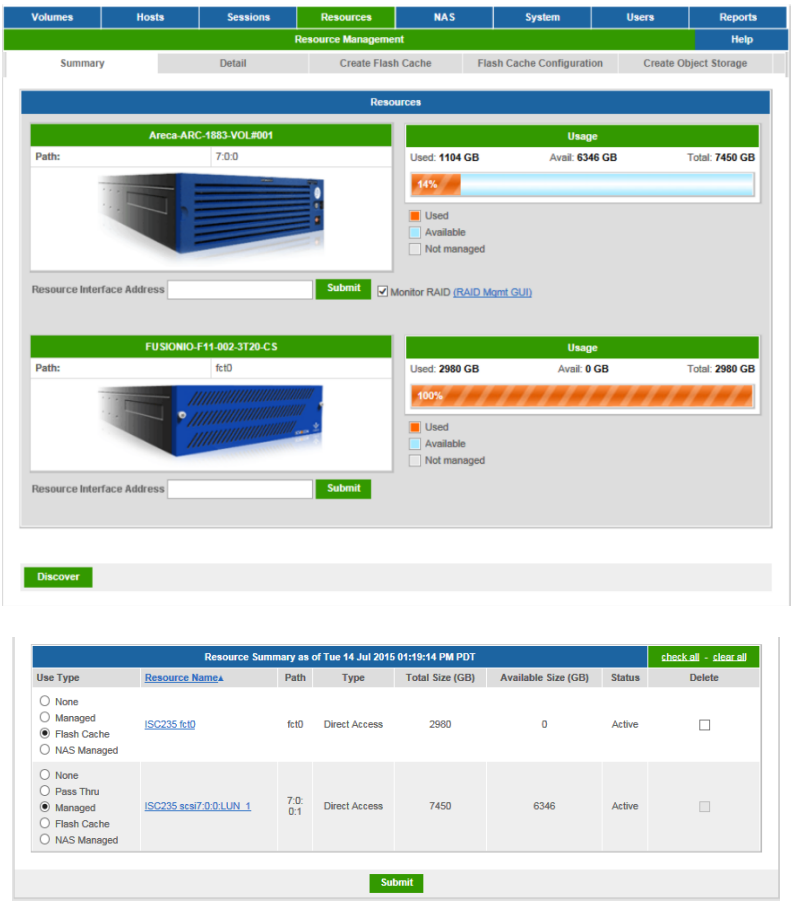


Figure 13-1 Selecting Resource for Flash Cache device

Use type **Flash Cache** has to be assigned to appropriate resource that provides higher performance compare to other **Managed** resources. Size the Flash Cache resource has to be at least 10% of the resource size that the Flash Cache device is going to serv.

13.2 Create Flash Cache Device

Open **Create Flash Cache** screen from the Resource Management menu:

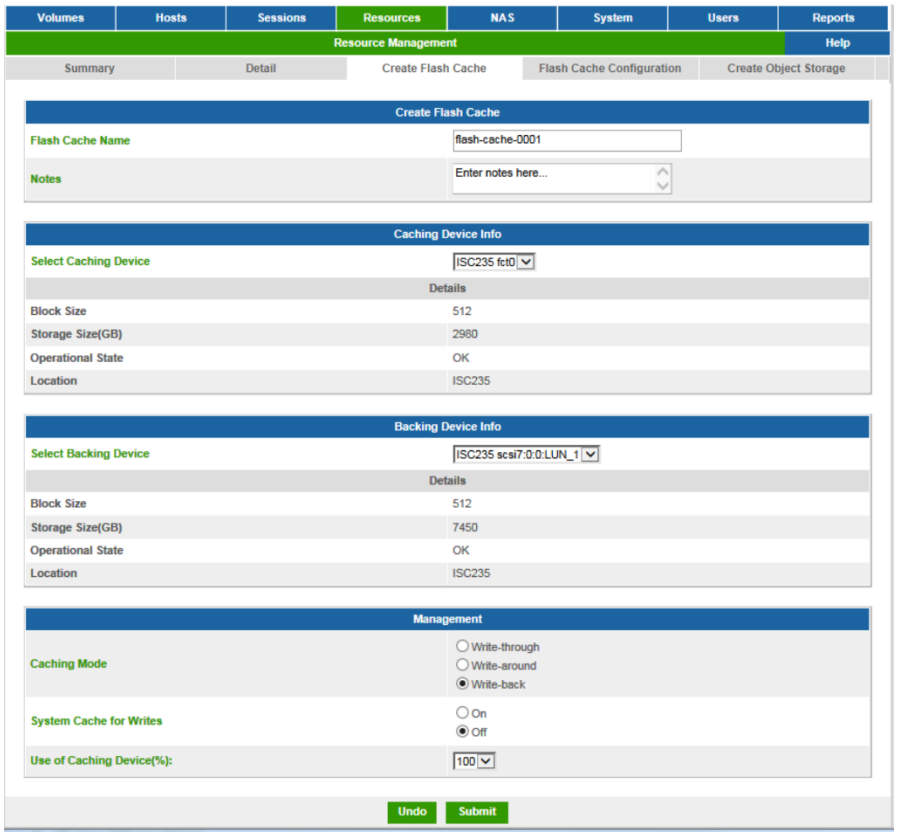


Figure 13-2 Creating Flash Cache Device

1 Enter a name of the flash cache device in **Flash Cache Name** field.

2 Enter descriptive **Notes** regarding this flash cache if needed.

3 Select **Caching Device**. Choose a resource from the list of available flash cache type resources. The resource will be used as a block cache for write and read commands to and from the selected backing device.

4 Select **Backing Device**. Choose a resource from the list of available Managed SF Managed resources to enable block caching for.

5 **Caching Mode**. The following write caching modes are available: write-through, write-around and write-back.

The **write-through** mode is the safest. All writes are cached to the caching device and are also written to the backing device before response is reported to I/O submitter. Blocks cached due to writes are also used to cache future reads.

The **write-around** mode is also very safe. Writes are not written to the caching device at all, but are directly written to the backing device.

The **write-back** caching mode is the fastest, but is the least safe. Writes go to the caching device first, and are written to the backing device at some time later. The I/O is completed to the I/O submitter long before it is written to the backing device. Blocks cached due to writes are also used to cache future reads.

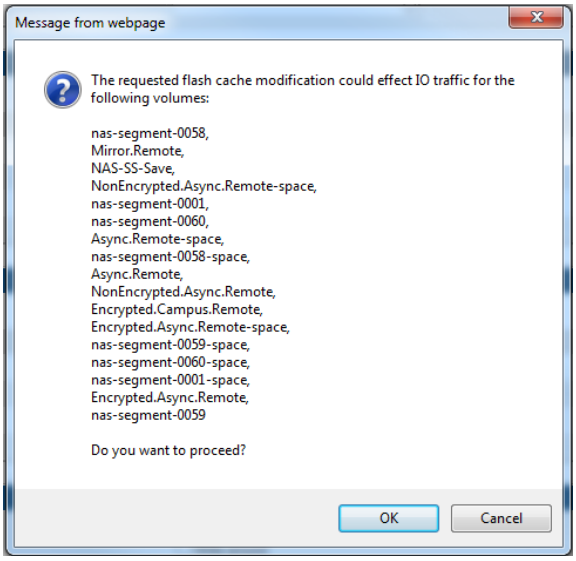
The caching mode only affects write caching. The read of a block not already in cache will always cause it to then be cached as part of the read.

6 **System Cache for Writes**. The system buffer cache can be enabled or disabled for writes that have to be submitted to the flash cache device. When the buffer cache is enabled writes are cached by the buffer cache and response is reported immediately. The cached blocks are submitted to the flash cache device in the background later. Cached blocks are used for reads and not cached reads are cached. The buffer cache is resides in system memory. It is fast but has limited size. The buffer cache is prone to errors: if something happens to machine before the data in the cache waiting to be written gets written, the changes in the cache are usually lost. This option has to be set "On" on systems that have at least UPS to protect from unexpected power outages.

7 **Use of Caching Device**. Select what percent of the caching device use to keep the device metadata and cached blocks. It is recommended that size of the caching device has to be between 10 and 20 percent of size of the backing device. The cached device size above these limits does not provide additional benefits but it increases time to maintain the device.

8 Click **Submit** button to create the selected Flash Cache device.

9 If Backing Device is used to provision volumes the next pop up message is generated:



After the cache device is created the backing resource is exposed on GUI screens with special footnote:

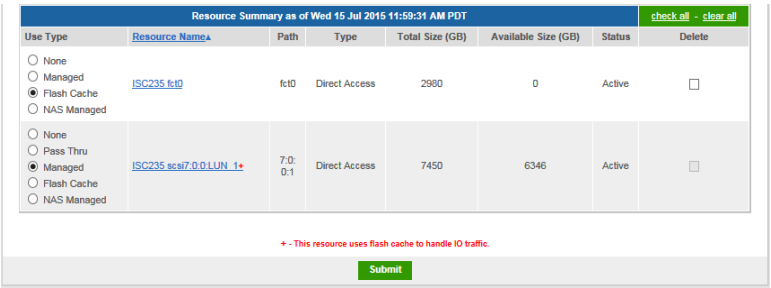


Figure 13-3 Footnote for Resources with Enabled Flash Cache

13.3 Manage Flash Cache Device

13.3.1 Device Status

Open **Flash Cache Configuration** screen from Resource Management menu to see current status of the selected Flash Cache device:

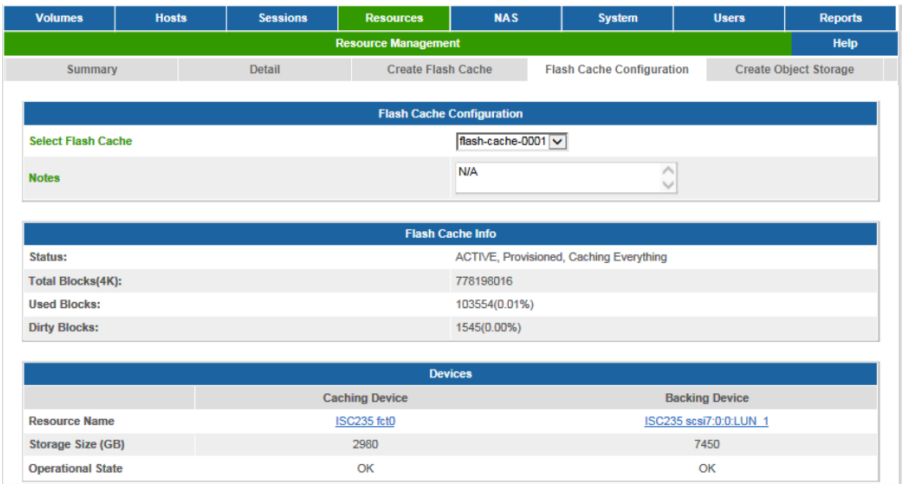


Figure 13-4 Status of Flash Cache Device

The appropriate Flash Cache device has to be selected from the drop down list first. The notes for the device can be modified on this screen. Value of **Total Blocks** is maximum number of 4K data blocks that can be cached by the cache device. The **Used Blocks** value is number and percent of Total Blocks that were used for caching after the cache device is created. The **Dirty Blocks** value is number and percent of Total Blocks that are cached and have to be written to the backing device.

The **Status** field could have next values:

- **No access** – Have failure to access the caching device. IO's are going to the backing device directly;
- **Not loaded** – The cache device is not loaded. IO's are going to the backing device directly;
- **Not provisioned** – The cache device is not provisioned. IO's are going to the backing device directly;
- **ACTIVE** – The device operational state is active.
- **SUSPENDED** – The device is suspended by the operational system. Any IO that has already been mapped by the device but has not yet completed will be flushed. Any further IO to that device will be postponed for as long as the device is suspended;
- **Provisioned** – The device is in ACTIVE operational state and is managed by Storage Concentrator;
- **Bypass Cache** - The read and write IO's are bypassing system buffer cache and flash cache device and are going directly to the backing device. IO's do not split on 4K or smaller block sizes. This mode is equivalent to case when the backing resource does not have block cache completely. The flash cache device provisioned but it is not used;

- **Caching Nothing** - IO's are going through the flash cache device without caching. The option does not have effect on system buffer cache;
- **Caching Everything** - IO's are split on 4K or smaller block sizes and are handled by the cache device. The option does not have effect on system buffer cache.

13.3.2 Change Configuration

Most of the caching device settings has to be selected during creation and could not be changed later. The list includes Caching Mode and use of System Cache for Writes. User has ability to turn off and on caching and to modify parameters used during flashing of dirty blocks.

To change cache device settings open **Flash Cache Configuration** screen from Resource Management menu:

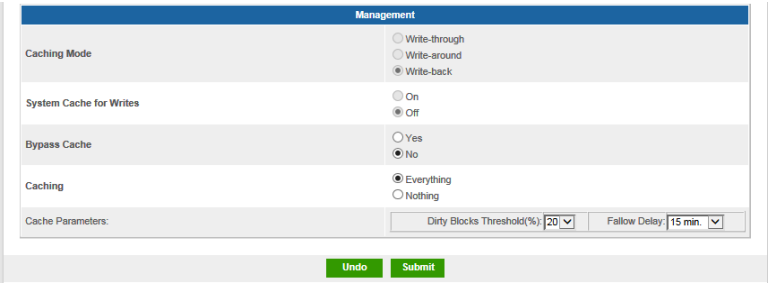


Figure 13-5 Modify Flash Cache Device

Caching Mode - Caching mode of the selected device. Can't be changed without deleting and re-creating the flash cache device.

System Cache for Writes - State of the system buffer cache. Can't be changed without deleting and re-creating the flash cache device.

Bypass Cache - If "Yes", the read and write IO's are bypassing system buffer cache and flash cache device and are going directly to the backing device. IO's do not split on 4K or smaller block sizes. This mode is equivalent to case when the backing resource does not have block cache completely. The flash cache device provisioned but it is not used. Click Submit if the setting has to be changed or undo to clear your changes on the screen.

Caching - When "Nothing" is set, IO's are going through the flash cache device without caching. The option does not have effect on system buffer cache. IO's are split on 4K or smaller block sizes as when the option "Everything" is set. Click Submit if the setting has to be changed or undo to clear your changes on the screen.

Dirty Blocks Threshold - Available only for Write-back cache devices. The flash cache will attempt to keep the dirty blocks in each set under this percent. A lower dirty threshold increases backing device writes, and reduces block overwrites, but increases the blocks available for read caching. Click Submit if the setting has to be changed or undo to clear your changes on the screen.

Fallow Delay - Available only for Write-back cache devices. The device cleans dirty blocks that have been "idle" (not read or written) for "Fallow Delay". Default is 15 minutes. Setting

this to "disabled" disables idle cleaning completely. Click Submit if the setting has to be changed or undo to clear your changes on the screen.

13.3.3 Flash Cache Effect on System Performance

Creating cache device for resource makes sense only it improves performance of the resource. User has ability to turn caching off and on temporally (see 13.3.2 Change Configuration) and check Storage Concentrator performance to execute IO's for specific hosts.

Storage concentrator provides ability to show how many IO's was not handled by cache device because the IO size. IO's less than 4K are not cached by the device. The original IO could be larger enough but if it not aligned on 4K boundary it submits to the cache device blocks that have to go without caching directly to backing device. In successful cases of Flash Cache implementation number of these IO's has to be reduced to minimum.

The other option is to analyze number of read and write IO's handled by caching and backing devices. The more number of hits for cached data blocks the better performance of the cache device has to be.

The Flash Cache Configuration screen from the Resource Management menu provides functionality to collect statistics for number of blocks handled by the caching and backing devices. The sampling interval could be 1 second or user could start and stop sampling in any time.

To see IO's current statistics for 1 second sampling interval click button **Refresh**:

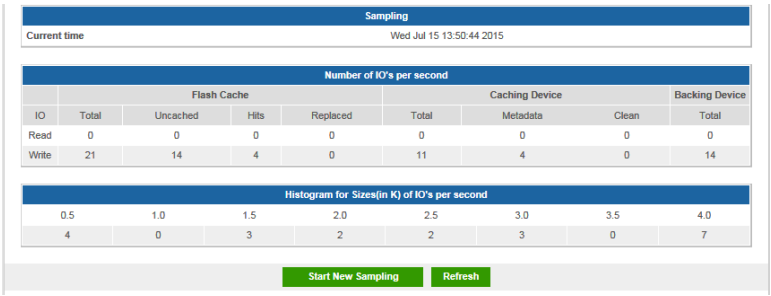


Figure 13-6 One second IO Statistics

The screen show pattern of IO's handled by different components of the flash cache device at moment displayed in the **Current time** field:

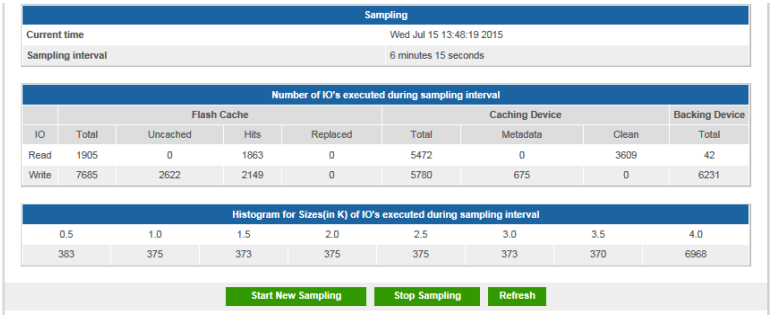
- **Total** - The total number of IO's.
- **Uncached** - The number of IO's that bypassed the cache. Some common cases are IO's that are smaller than the Flash Cache block size, or are not aligned on that size.
- **Hits** - The number of IO's that are served from the caching device instead of the backing device.
- **Replaced** - The number of cache blocks in use that needed to be reclaimed for use for new reads or writes.
- **Metadata** - The number of writes to update Flash Cache metadata when IO's are

write cached.

- **Clean** - The number of reads from the cache done when cached write IO's are committed to the backing device.

Sizes of read and write IO's handled by the flash cache device at moment shown in the **Current time** field are displayed at **Histogram for Sizes of IO's per second**.

To start new sampling interval click button **Start New Sampling**. Use button **Refresh** to update screen with statistics accumulated from start of the sampling interval, use button **Stop Sampling** to stop sampling and update screen with statistics collected between start and stop moments:



Sampling

Current time	Wed Jul 15 13:48:19 2015
Sampling interval	6 minutes 15 seconds

Number of IO's executed during sampling interval

	Flash Cache				Caching Device			Backing Device
IO	Total	Uncached	Hits	Replaced	Total	Metadata	Clean	Total
Read	1905	0	1863	0	5472	0	3609	42
Write	7685	2622	2149	0	5780	675	0	6231

Histogram for Sizes(in K) of IO's executed during sampling interval

0.5	1.0	1.5	2.0	2.5	3.0	3.5	4.0
383	375	373	375	375	373	370	6968

Start New Sampling | Stop Sampling | Refresh

Figure 13-7 IO Statistics for User Defined Sampling Interval

13.4 Delete Flash Cache

Open Flash Cache Configuration screen from Resource Management menu and select appropriate caching device from the drop down menu:

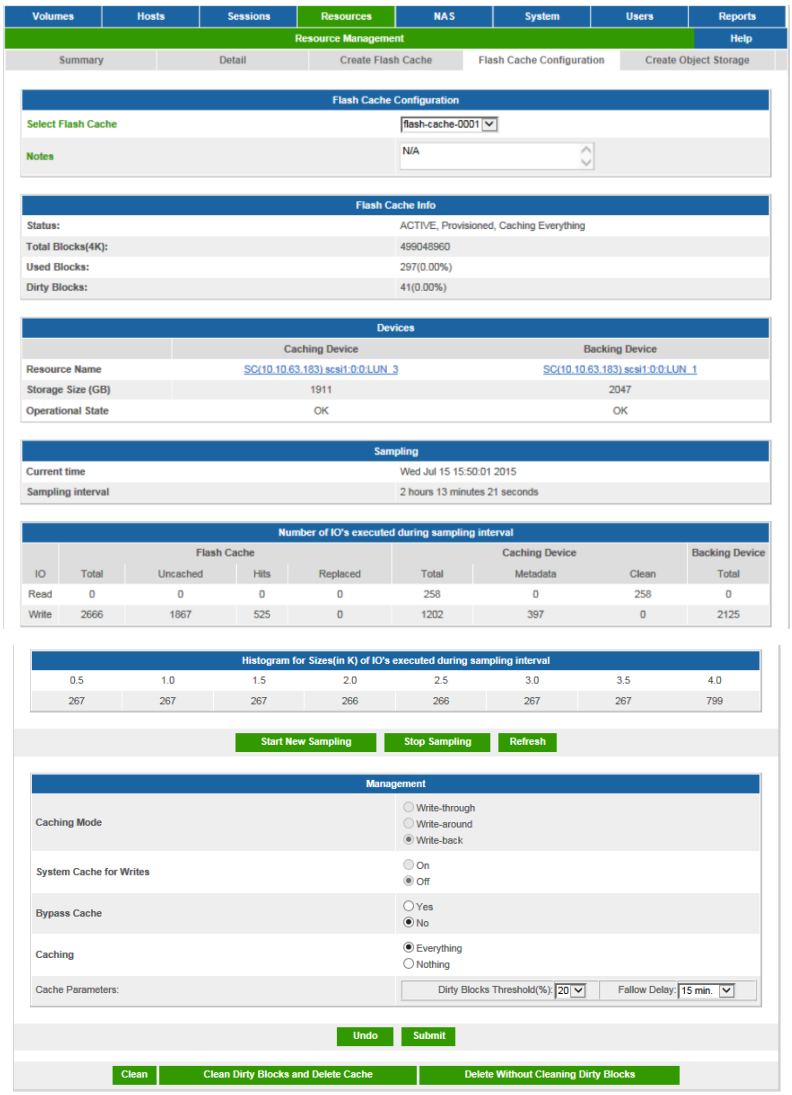


Figure 13-8 Delete Flash Cache Device

Click button **Clean** to initiate process of dumping dirty blocks from caching to backing device. The functionality is available only for devices with Write-back mode.

Use button **Clean Dirty Blocks and Delete Cache** to delete the selected flash cache device and restore IO traffic through the backing device without using of block caching.

Click button **Delete Without Cleaning Dirty Blocks** to delete the selected flash cache device and restore IO traffic through the backing device without using of block caching. The functionality is available only for devices with Write-back mode. It does not dump dirty blocks to the backing device. Some data could be lost.

Appendix 1

Object Storage Resources

A1.1 Introduction

Fragmented storage adds complexity and slows down innovation for business applications. Object storage helps you break down these silos by providing massively scalable, cost-effective storage to store any type of data in its native format.

With StoneFly Storage Concentrator object storage service, you manage your storage in one place with an easy-to-use GUI interface. You can use policies to optimize storage costs and tiering between different storage classes. Storage Concentrator Object Storage service makes storage easier to use to gain insights, perform analysis and make better decisions faster.

A1.2 Attaching AWS / Compatible AWS S3 Storage to Storage Concentrator

Please refer to the "Attaching Amazon AWS Storage Resources to Storage Concentrator" manual.

A1.3 Attaching Azure Blob Storage to Storage Concentrator

Please refer to the "Attaching Azure Blob Storage Resources to Storage Concentrator" manual.

Appendix 2

iSCSI Initiators

A2.1 Introduction

StoneFly' *Storage Concentrator* is the interface between hosts and servers and storage devices in an IP network. iSCSI is the protocol that makes this possible. Each host that wants to use the storage devices presented by the *Storage Concentrator* must implement the iSCSI protocol and link to the Ethernet connection to the storage network. The iSCSI Initiator (also called client) is the component that makes the network connection between a host and the *Storage Concentrator*. The initiator is responsible for making the *Storage Concentrator* that is attached to the network appear to the host system as a locally attached SCSI device. The initiator can be a software driver or a hardware card. Hardware initiators are referred to as iSCSI HBAs, Storage Adapters, or as TOEs (TCP offload engine) and are usually installed in a PCI/PCI-X slot in the host computer. Software initiators are loaded onto the host or server as a driver and use the main CPU of the host computer. The most widely used software initiator is the Microsoft iSCSI initiator, but others are available for Linux, Unix, Novell, and so on.

The following information describes how to successfully implement and use iSCSI initiators in IP networks.

Network Overview

A standard IP network is used to connect the iSCSI initiators and the *Storage Concentrator*. A single host may be directly connected to the *Storage Concentrator* with a CAT5 cable or multiple hosts may be connected to the *Storage Concentrator* through an IP network, which includes switches and routers. No restrictions are placed on the use of iSCSI initiators in IP networks. The network design for the initiators determines the amount of configuration that must be done to establish the connection between the iSCSI initiator on the host(s) and the *Storage Concentrator*.

The *Storage Concentrator* Cluster comes standard with dual Gigabit Ethernet (GbE) ports with load balancing across the ports.

Software Initiators: The primary method of connecting hosts to is to use a host software-based initiator, such as the Microsoft iSCSI initiator. Other initiators, including hardware-based initiators will use either the Microsoft Initiator or proprietary software drivers to support iSCSI communication.

Hardware components for host systems: If a software initiator is used as the iSCSI initiator, the host may need a separate standard Ethernet network adapter card dedicated to iSCSI use. This adapter card can be any speed including 10baseT, 100baseT or 1000baseT (GbE), or a combination 10/ 100/1000. It is recommended that a 1000baseT (GbE) card be used for iSCSI storage applications. If a hardware initiator (such as an iSCSI HBA) is used as the iSCSI initiator, the initiator may not support regular TCP/IP communication with other computers on the network. In this case, the host needs a separate network adapter card to provide Ethernet connections to other machines.

Routing: Because the hosts are not required to be located on the same network as the *Storage Concentrator*, the *Storage Concentrator's* GUI allows routing and gateway information to be entered during configuration. All hosts that appear on the same network as the *Storage Concentrator's* Host iSCSI GbE port are immediately available to the *Storage Concentrator* for connection. If a host is located on a different network,

the routing and gateway information must be entered into the *Storage Concentrator* configuration to make the connection possible.

A2.2 Usage Models

The features of each iSCSI initiator may dictate a different usage model. However, several models can be applied to making the best connection between the initiator and the *Storage Concentrator*. Typically there is not a distinction between software-based or hardware-based iSCSI initiators. The distinction between different iSCSI initiators comes from the way each initiator is managed.

To successfully connect the iSCSI initiator and the *Storage Concentrator*, the IP address of the initiator must be in the access control list (ACL) assigned to the volumes.

1 **Hosts must know the IP address of the *Storage Concentrator*.** The connection between the iSCSI initiator and the *Storage Concentrator* is a TCP/IP connection. The connection begins when the IP address of the iSCSI GbE Port on the *Storage Concentrator* is provided to the iSCSI initiator. Be sure to use the IP address the administrator has assigned to the Host iSCSI GbE port and not the IP address assigned to the Management GbE port. The *Storage Concentrator* may be referred to as the "iSCSI Target" in the management interface or in the configuration file of the iSCSI initiators.

2 **Hosts must appear on the Access Control List (ACL).** The Volume Security page in the *Storage Concentrator* GUI must contain the IP address of a host before it can grant access to any of the volumes managed by the *Storage Concentrator*. The IP address of the host may be manually entered in the *Storage Concentrator* GUI or automatically discovered by receiving I/Os from the host. If the proper IP address of the *Storage Concentrator* has been entered through the initiator management interface, the first I/O will be delivered to the *Storage Concentrator* automatically. If the *Storage Concentrator* is operational and on the network when the first I/O is issued, the IP address of that initiator is captured and retained for use in configuring ACLs within the *Storage Concentrator*. Some initiators may send more than one I/O to the *Storage Concentrator* before the final connection is made. If the *Storage Concentrator* does not capture the initiator's IP address on the first I/O, the initiator may need to be restarted or reset to send additional I/Os. Repeat the process until the IP address of the host appears on the ACL configuration page of the *Storage Concentrator* GUI.

3 **Hosts may or may not perform discovery.** Hardware and software iSCSI initiators fall into two categories: those that perform iSCSI discovery and those that do not. Discovery is the process in which the initiator searches the network for iSCSI target devices and then collects information from them.

a. Initiators that perform iSCSI discovery do not require user intervention on the host side. iSCSI discovery allows the *Storage Concentrator* administrator to configure volumes and assign them to hosts. This information is automatically captured by the initiator during discovery.

b. Initiators that do not perform iSCSI discovery require that the iSCSI target name of the *Storage Concentrator* be entered at the host before connecting to the *Storage Concentrator*. The iSCSI target name is either entered into a file or through the initiator management GUI. The required iSCSI target name can be found on the Volume>Details page of the *Storage Concentrator* GUI.

c. Some initiators implement discovery as an interface to services on the network that provide information about storage devices. One such method is iSNS. Refer to "iSNS" for information on using iSNS for discovery.

This page is intentionally left blank.

Appendix 3

Preparing for Asynchronous Mirroring

This Section provides information and checklists to prepare for Asynchronous Mirroring using StoneFly Storage Concentrators connected through a WAN connection. The information is specific to the StoneFly products but discusses issues that affect any Asynchronous Mirror or Replication scenario.

Asynchronous Mirrors present several issues that require prior planning and estimating before implementing the system in any environment. The only way to be successful is to make high-quality estimates. This document leads you through some discussions of the reasons for these estimates and suggests ways that these estimates can be made.

A3.1 Introduction to Asynchronous Mirrors

Asynchronous Mirrors can be used to replicate data between two geographically separate sites. For the purposes of this document, the source data location will be referred to as the local, or production, site and the other as the target or remote site. The sites are represented by two different Storage Concentrator systems, one at each site, regardless of the distance between them. Asynchronous mirroring is associated with a desire to have a level of data recovery in a remote location in the face of any unplanned disaster. In recent times more companies have been required to keep data at certain distances (in some cases hundreds of miles) away from the production site to insure protection of the data in the event of a disaster at the local site. The StoneFly asynchronous mirror process includes the ability for the local Storage Concentrator to identify changes in a volume and periodically move those changes to a remote site. This process is done by coupling StoneFly Mirroring and Snapshot technology. The StoneFly Snapshot technology tracks the changes in the local (production) volume while the Mirroring technology (mirroring) maintains the movement of the data to the mirror image at the remote site.

The Storage Concentrator uses two methods to identify when the changes must be sent to the remote site: (1) measuring a period of inactivity, or, (2) using a predefined data replication schedule. The task is the same regardless of the amount of change. Using one of the two trigger methods, the local, or production Storage Concentrator triggers a replication event that sends the changes to the remote site. The link's throughput capability is critical and must be qualified in advance to be able handle the flow of replicated data to the remote site. The desired level of synchronization between the two sites is expressed in: (1) the size and shape of the communications link and (2) the space allocated to hold the changes before they are sent to the remote site. A slow link with a high rate of change in the data will dictate a gap (perhaps large) in the synchronization between the sites. A high speed link and smaller rates of change in the data can keep the two sites closely synchronized.

The simple concepts above are formalized in terms such as Recovery Point Objective and Recovery Time Objective. Recover Point Objective describes the level of synchronization between the two data sets. As the synchronization gets better the Recovery Point Objective is closer to 100%. The Recovery Time Objective describes the amount of time needed to convert from using the data at the Production Site to using the data at the Remote Site. In many cases higher level of RTO (shorter time periods) is assisted by the use of Hot-standby servers at the Remote Site. Consideration should be taken to understand your organization's Recovery Point Objective and Recovery Time Objective (RPO and RTO respectively). If you require a highly-synchronized RPO and short RTO, bandwidth and replication scheduling should be planned carefully.

It is also good to think about how the replicated data will be used in a Recovery

Plan. The communications discussion above can also affect the way the data is used at the remote site. Bringing data back to the local IT site through a slow link may be prohibitive in the desire to get an organization up and running again. Even a single large file needed at the local IT site could take days of transmission time over something like a T1 link.

A3.2 Operational Overview

StoneFly Asynchronous Mirroring is implemented using StoneFly Mirroring (mirroring) and Snapshot technology. The asynchronous replication processes are coupled with snapshots to track all the changed blocks in the volume.

A simple view of the way StoneFly Asynchronous Mirrors work is to describe the two volumes that must be synchronized during the mirroring or replication process:

1. The Remote (Replication) Volume
2. The Production Volume

Many components outside the Storage Concentrators pay a role in the configuration. Each component should be thought out and pre-qualified for its intended role.

1. The Production Site includes the production servers and the SC Volumes used by them.
2. The Replication Volume is part of the Remote Site but appears in the diagram on the Production side as well. It is discovered and used as a storage Resource. It is "owned" by the Storage Concentrator at the Production Site.
3. Note the two routers in the diagram. Each is one end of the WAN link between the two sites. The distance between the routers and the health of the link are major contributors to the success of the configuration.

A3.2.1 Overview of the Remote Volume

The basis of the replication process is the existence of a second copy of data at a remote site. The first step toward Asynchronous Mirroring is to provide a volume at the remote site that is the same size as the volume being mirrored from the production site. This remote volume is intended to contain the exact same data as the volume at the production site. Data is being added to the remote volume through time at an unknown rate and may be in an unstable state if stopped unexpectedly. Any disruption to the communications link between the production site and the remote site may leave IO's on the link and not completed. The use of the volume in a disaster recovery situation may require the use of tools that will repair any problems in the file system or application data.

Alternately, the remote volume may be enabled for snapshots. This is highly recommended to improve the accuracy and stability of the replicated data at the remote site. Each time a complete set of changes is sent to the remote volume, a snapshot is requested by the production Storage Concentrator. These snapshots become recovery points at the remote site. The amount of Snapspace allocated at the remote site determines how many recovery points are possible. Allocating a larger

Snapspace provides more recovery points.

These remote volumes are specified to be discovered by the Production Site SC. The volumes are used as storage resources and not assigned to any server at the Production Site. The volumes may be used at the Remote Site by the hot-standby servers as part of the Recovery Plan.

A3.2.2 *Overview of the Production Volume*

The Production Volume is the source of all changes in the replication process. Changes are captured through the use of snapshot technology that tracks all changed blocks in a volume. Once the time for a replication has been triggered, a differencing process takes place and all changed blocks are transferred to the remote site. A special set of snapshots are established on the production volume to allow for the differencing process to take place. After the snapshots are established the volume is connected to one or more of the available remote volumes.

The Storage Concentrator's snapshot technology uses a separately allocated space in the Resource pool for tracking changes. The allocation of this space for the Production Volume is a buffer against problems with communications to the remote site. A larger snapshot allocation provides more space to hold changes while the communications link may be down. The size of the space directly represents the available time to recover any communications problems before the synchronization cannot be done between the two sites. When this space fills up a full re-synchronization of the two sites must be performed.

A3.3 Preparations and Checklists

The successful Asynchronous Mirror operation is based on the balance between the amount of change in the Data and the size of the communications link provided to carry this data to the remote site. The process begins by estimating data changes.

A3.3.1 *Estimating Data Changes*

There are several suggestions for estimating the amount of change in the data volumes. These techniques rely on the selection of drives that will be replicated to the remote site. The list should be sure to include all the current drives on all servers which must be processed. Any information about future needs should be factored into the process. One of the ways the Asynchronous Mirror implementation can be rendered insufficient is by not estimating the current data and how that might change in the near future.

Step 1 is to complete the list of drives to be replicated. List the drives with their current sizes. Then estimate any changes that might take place in the next 6 to 12 months. A chart is provided at the end of this document for listing the volumes.

Step 2 is to actually make some measurements that allow estimates of the data to be moved to the remote site. These estimates will be expressed in terms of bytes of data per second, minute, hour, or day being written to the file systems in the drive list. The timing of these measurements is critical. A two hour measurement in the middle of the week is insufficient to provide accurate estimates. The knowledge of the data administrator is important to factor into the measurement. Factor in usage cycles that may force the measurements higher on certain days of the month.

A3.3.2 *Estimating at Sites with No Storage Concentrator*

If no StoneFly Storage Concentrator exists at your site the estimates may be made through at least two different methods.

1. By using counters on each server. The purpose of these counters is to produce a record of the bytes per second or minute written to the drives over a reasonable period of time.
2. By gathering statistics form differential backups. The nightly or weekly backups may produce data on how big the changes were for the day or week.

The estimates that come out of these measurements are not exactly applicable to the snapshot technology used in the Storage Concentrator. The pattern of changes in the volume or drive is important. The snapshot technology works in multiple blocks of storage rather than a block by block tracking mechanism. Spacing small changes throughout a volume or drive commits more data blocks to be replicated than are actually changed. Tightly grouped changes tend to reduce the number of bytes to be transferred to the remote site.

A3.3.3 *Estimating at Sites with Existing Storage Concentrators*

The highest quality measurements are achieved when a Storage Concentrator is already available for the drives to be replicated. As described above, the replication process uses the Snapshot technology in the Storage Concentrator to track the amount of changes between each replication. By establishing snapshots on each volume to be replicated the administrator is able to see the exact size of the Snapspace used on a daily, weekly, monthly basis.

The snapshot features in the Storage Concentrator require an allocation of storage from the Resource pool to hold the changes for the volume. A large allocation should be made for the initial measurements. In this way an accurate measurement is available for a longer period of time. For instance, setting a Snapspace equal to the size of the volume allows measurements for a week or month at a time. It is not unreasonable to measure a week long amount of changes. The objective is to see what would happen if the WAN link went down for several days.

A3.3.4 *What to Do With the Measurements*

After the measurements are taken and recorded they must be used to estimate the data movement between the sites.

1. The Daily amount of change on each drive is measured and averaged. The assumption is that whatever the size of the WAN link, this amount of change can be moved to the Remote Site in a 24 hour period. If not then the site will

fall further and further behind each day and ultimately the Asynchronous Mirror feature will not be able to continue. The whole process must be restarted after fixing the problem.

2. An estimate of the total downtime on the WAN link must be estimated. This number might come from a Service Level Agreement with the WAN supplier or from knowledge of the supplier. One estimate that must not be overlooked is the amount of time to move the Remote Site equipment from the Production Site to its remote home. The normal initialization technique is to start Asynchronous Mirroring with all the equipment at the Production Site. This allows a very fast link for moving the data to the remote equipment. Then the remote equipment is disconnected and shipped to the remote site. If it takes a week to set up equipment at the Remote Site the Production SC must have sufficient space to hold the week's changes. After the installation is complete at the Remote Site the WAN link must be able to sustain the movement of all these changes in a reasonable amount of time. New changes will be tracked while the two sites re-synchronize themselves.

3. Now that the possible maximum amount of change is known some calculations can be made to see the length of time required to move those changes across a WAN link. For instance, a 10 Mbps link is frequently used for DR sites. This link works at about 1 MByte per second and approximately 3.6 GBytes per hour. This number recognizes the Ethernet packet headers and such that reduce the amount of actual data being transferred.

4. Is the proposed WAN link speed enough to move this data to the Remote Site in a reasonable time? Will the daily change rate move across the link properly? Increasing the speed of the link dramatically changes the efficiency of the Asynchronous Mirroring process.

A3.3.5 *Checklists*

Determining the space to hold changes and the size of the WAN link can be arrived at using the following chart. The method is designed to determine the greatest amount of data that must be moved to the remote site and the time that it must be done.

Step	Activity	Results
1	Create list of Drives to be Mirrored (use chart below)	
2	Establish the measurement method	
3	Run measurements for days or weeks.	
4	Record data changes for each drive and summarize the data changes per time period across all drives	

5	Estimate growth and activity increases per drive over the next 6 months	
6	Establish the daily quota of data to be mirrored to the remote site for each drive and as a total	
7	Estimate the possible downtime on the link in days, including initial set up – Ask for a Service Level Agreement from suppliers of WAN links	
8	Set your work space per drive to accommodate this amount of data change (based on #6 and #7)	
9	Establish the desired time to recover after the longest estimated link downtime (moving all the stored changes in #8 to the remote site)	
10	Order the WAN link to allow recovery in the desired time	

A3.3.6 *Recording Volume Information*

Volume Name	Volume Size	Average Daily Changes in Bytes	Size of Snapspace in Giga-Bytes

Size of Snapspace in Giga-Bytes = Average Daily Changes x number of days of downtime allowed or expected

This page is intentionally left blank.

Appendix 4

Safety Information

Please review the following safety information completely before installing your *Storage Concentrator*

Environmental	Operating temperature: +50F to +95F (10 C to 35 C)
	Power consumption: 200.1 Watts (687 BTU/hour) minimum, 217 Watts (743 BTU/hour) maximum
	System Cooling: 4x8 cm fan assemblies
	Altitude: -50 to 10,000 feet (-16 to 3048 meters)
	Relative humidity: 8% to 80% (non-condensing)
Compliance	FCC 47 CFR Part 15, Class A
	CSA C1088, Class A
	UL 60950; CUL 60950
	CE Mark (EN 55022 Class A/EN 55024)

Appendix 5

Technical Specs

A5.1 Storage Concentrator Technical Specifications

A5.1.1 Protocols and Standards

- iSCSI (IETF Version 1.0 Draft 20)
- IP (RFC 791, 894, 1092)
- TCP (RFC 793)
- IMCP (RFC 792, 950,1256)
- SCSI-2 and SCSI-3
- Fibre Channel
- Serial Attached SCSI (SAS)

A5.1.2 Security

CHAP and iSNS

A5.2 Initiators

- Windows NT/2000/2003 Server/2003 Storage Server,
- Linux
- Solaris
- AIX
- Novell

A5.2.1 Targets

- Parallel SCSI (Per Port): Up to 15 storage targets
- Fibre Channel (Per Port): Up to 127 storage targets per Fibre Channel Loop or any number in a switched fabric

A5.2.2 Logical Volume Management

- Volume Size: Minimum 1 GB, Maximum is based on the maximum of the Operating System using the volume.
- Maximum Number of sessions: 1022*

*Some iSCSI initiators may require more than one session to access storage — reducing the total number of storage sessions available.

A5.2.3 Management

Telnet (RFC 854) HTML HTTPS

A5.2.4 IEEE

IEEE 802.3 (10/100 Ethernet) IEEE 802.3z (Gigabit Ethernet)

A5.2.5 Regulatory

- FCC 47 CFR Part 15, Class A
- CSA C1088, Class A
- UL 60950, CUL 60950
- CE Mark (EN455022 Class A/EN455024

A5.2.6 Warranty

One year web-based and telephone technical support and advanced parts replacement. Extended and on-site warranties are available.

A5.2.7 Storage Concentrator Configuration Limits

Limits Per Storage Concentrator	Number
Hosts Sessions (total per cluster or per standalone SC)	1022
Total Number of Volumes	200
Total Number of Mirror Volumes	40
Mirror Images Per Volumes	4 (Sync) 3 (Async)
Total Number of Snapshots	40*63 = 2,520
Maximum Snapshots Per Volume	63
Snap Space Limit Per Volume	8,191GB
Total Snap Space Limit Per System with 512MB of RAM or less	4,095GB
Total Snap Space Limit Per System with 1024MB of RAM	6,143GB
Total Snap Space Limit Per System with 2048MB of RAM or more	24,575GB
Max Volumes per Thin Pool for System with 512MB of RAM or less	16
Max Volumes per Thin Pool for System with 1024MB of RAM	64
Max Volumes per Thin Pool for System with 2048MB of RAM or more	128
Max Total Size of Thin Pools for System with 512MB of RAM or less	4,072GB
Max Total Size of Thin Pools for System with 1024MB of RAM	16,288GB
Max Total Size of Thin Pools for System with 2048MB of RAM	32,577GB
Max Total Number of Thin Pools	10
Total Number of Async Images (40 volumes max, 3 per volume)	120
Total Number of iSCSI Hosts	200
Total Number of Snap Enabled or Replicated Volumes	40
Max Number of Snap Enabled or Replicated Volumes Per Thin Pool or Deduplicated Pool	5
Number of Hosts with access to a volume when SCSI Persistent Reservation is used.	32
Number of Hosts with access to a volume when SCSI Persistent Reservation is not used.	200
Minimum recommended System memory size to support NAS volumes	4GB
Minimum recommended System memory size to support Deduplicated Volumes	7GB
Minimum recommended System memory size to support maximum Deduplication Configuration	30GB

Limits Per Storage Concentrator	Number
Minimum Boot Disk space to support Deduplicated Volumes	72GB
Minimum Boot Disk space to support maximum Deduplication Configuration	288GB
Max Total Number of Deduplicated Pools	10

A5.2.8 *Typical Configurations for using Dual iSCSI GbE Ports*

A single *Storage Concentrator* provides the ability to share its storage resources with several server machines. Each server in the storage network provides a certain amount of the total bandwidth being presented to the *Storage Concentrator*. Typically each server is connected to the *Storage Concentrator* through a Gigabit Ethernet network interface card (NIC) or an iSCSI HBA. A single server with a Gigabit Ethernet connection has the ability to provide 100% of the available bandwidth presented to a single Gigabit Ethernet port on the *Storage Concentrator*. The most efficient method to expand the bandwidth is to connect a second gigabit Ethernet port in a teamed, or port aggregated, scenario.

The *Storage Concentrator* i5000 standard configuration includes two "teamed" Gigabit Ethernet ports. When this NIC is present in the *Storage Concentrator*, the unit automatically configures the two ports as a single iSCSI Data port with the two ports teamed as one entity within the *Storage Concentrator*. The pair of ports requires only one IP Address. The special purpose Intel driver code automatically detects the traffic flow and uses the two ports to achieve load balancing between them.

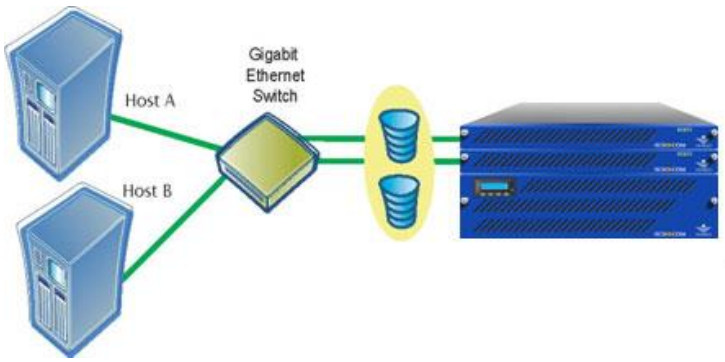


Figure App.-1

In the diagram above a single *IPSAN Appliance* is connected to two hosts/servers. The connection from HostA and HostB to the switch is one Gigabit. Even though most servers are not expected to run their network connections at 100%, the theoretically possible aggregate network load is two Gigabits/sec. By connecting both the teamed Gigabit Ethernet connections from the *Storage Concentrator* to the same switch it is

possible to insure that any load greater than one Gigabit/ sec is handled by the *Storage Concentrator*.

In addition, the Intel driver provides for backup between the two ports on the Storage Concentrator. In a single *Storage Concentrator* configuration, the failure of one of the ports will cause all traffic to pass through the other port.

A5.2.9 *Configuring Teamed NICs in a FailOver Cluster*

Storage Concentrator FailOver clusters have additional requirements to respond to failures in network components, including hosts, switches, cables and storage targets. The *Storage Concentrator* responds to failures in network components that directly connect to it (link status as detected by the ports in the *Storage Concentrator*). Other network components must be configured to provide redundancy on their own.

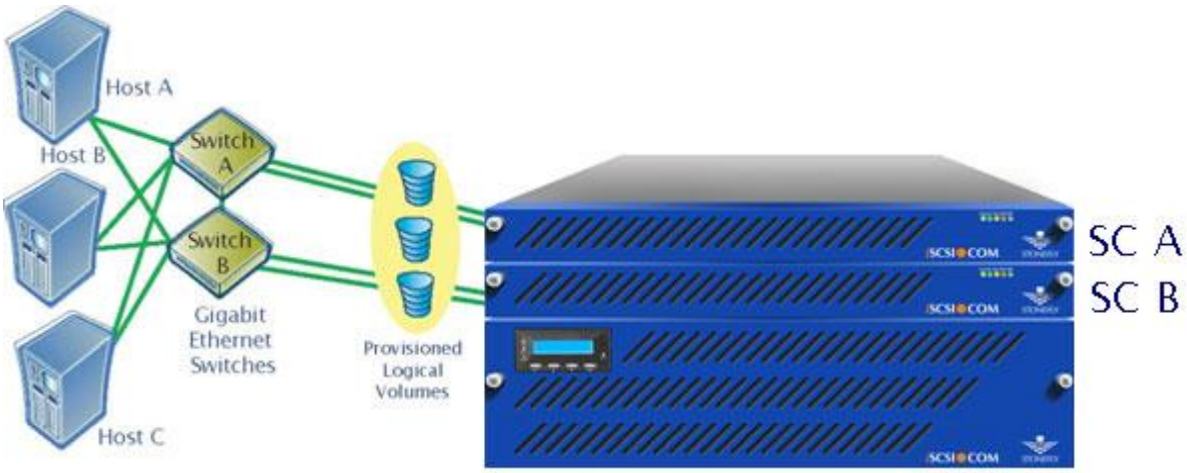


Figure App.-2

In the diagram above, traffic in the network between Hosts and Switches A and B is fully redundant due to the connections between the two switches. If Host A has traffic to the *Storage Concentrator* FailOver Cluster on both links the traffic going to Switch B, it must travel through the switch interconnect to get to the Primary *Storage Concentrator* (assumed to be SC A in this diagram). In the event that Host A's connection to Switch A dies all traffic will be applied to the Switch B connection. (Note: the connections on the host must also be configured for redundancy in some manner.) No FailOver is required by the *Storage Concentrators* since Switch A is still active. The traffic will pay a slight performance penalty while the host connection is down. Restoring the connection should cause the traffic to resume on both connections.

If the host connections are configured with only one active link and a backup link and the active port fails, the traffic can switch from Switch A to Switch B without requiring any action by the *Storage Concentrator* FailOver Cluster. The switch interconnect handles the delivery of the traffic to the Active-Active SC's.

In the diagram above Switches A and B are directly connected to the *Storage Concentrator*. A general failure (i.e., power loss) in a switch or the failure of a port connected to the Storage Concentrator will cause a FailOver event. The *Storage Concentrator* FailOver event does not require any action by the hosts other than the

expected re-login. The hosts are not required to shift their traffic patterns.

In a similar manner, the loss of connectivity to the GbE switch on the Primary or Secondary *Storage Concentrator* will cause a FailOver event. However, this FailOver does not require any shift in the traffic pattern.

This page is intentionally left blank.

Appendix 6

Network Configuration

This section describes some important Storage Concentrator (SC) SAN (Storage Area Network) Networking setup and configuration guidelines.

A6.1 Network Configuration and Wiring

The following issues related to the network switch configuration and wiring:

1. If the SAN network switch in use supports the Spanning Tree Protocol (STP), in single switch configurations, disable STP as it is seldom necessary in isolated SAN networks, and can result in long delays, disrupting the iSCSI traffic.

   In configurations where the SAN switches are not isolated and are connected to other switches, ensure that the STP is enabled only on switch ports that connect between switches, and disabled on all other switch ports. In HP ProCurve switches, this is done by configuring the non-switch-ports as "edge ports". In Cisco switches, enable "PortFast" mode on the non-switch-ports.

2. In multiple SAN network switch configurations, **make sure** that there is sufficient inter-switch bandwidth to avoid performance bottlenecks. For example, if the Storage Concentrator is configured with bonded 1 GbE NIC ports, providing only a single 1 GbE connection between switches will limit performance. Consider the network paths and the bandwidth available to them when implementing the network.

   The procedure to configure and verify the inter-switch networking is switch vendor/model specific; consult the product documentation for this information. When the network switches are not specifically configured for inter-switch trunking, or when the switch(s) are unmanaged (e.g. "dumb), only a single 1 GbE network interface would be used to carry network traffic regardless of the number that are physically wired.

3. The iSCSI Hosts initiators and SC targets should all be on the same IP subnet -- otherwise SAN data traffic may migrate to the LAN which is often slower, and can also destabilize the LAN. This occurs because when the initiator or target IP address is not on the same network as the SAN interface, and there is no other route, the traffic must then flow via the default route which uses the LAN interface.

4. There are cases where all of the iSCSI systems cannot share the same SAN network. One example is when a remote SC used for campus mirroring or asynchronous replication requires the use of different IP subnet. When this occurs, make sure to configure the needed routes on the SC GUI "System -> Network -> Routing" page. Also make sure that the Hosts have the necessary routes as well.

   One can determine that there is not an excessive amount of network traffic occurring on the LAN network by observing that the SC LAN network interface statistics are not too high with the SC GUI "System -> Network -> Management Port" "Stats" button.

5. When there are multiple switches used in the SAN network, make sure that the systems and switches are fully meshed – that is, that there is a route through one or more switch from every Ethernet port on every host to every SC Ethernet port.

6. The SC requires the use of a SAN network switch; point-to-point wiring without a SAN network switch is not supported.

7. The minimum SAN Ethernet interface speed supported is 1 Gbps Ethernet. Performance problems and timeouts may occur when full or half duplex 100 Mbps Ethernet speeds are used for SAN network traffic.

A6.2 SC SAN Network Port Bonding

The Storage Concentrator supports the bonding of separate Ethernet ports into a single "bonded" SAN network interface. This allows the SC to appear as a single IP address to different iSCSI Hosts, thus distributing the traffic across multiple network interfaces.

A6.2.1 *Balance-ALB Bonding*

The SC uses the bonding method known as Balance-ALB, or Adaptive Load Balancing. This bonding method requires no special network switch configuration or support, and thus can be used with unmanaged "dumb" switches.

Balance-ALB dynamically distributes network traffic to/from the SAN Host Ethernet interfaces amongst the available SC Ethernet ports. Note that traffic from a single Host Ethernet interface is never spread across multiple SC ports, but may be moved dynamically based on the current traffic loading to another less loaded port.

1. SC network interface bonding is supported by both 1 GbE and 10 GbE interfaces, but only one or the other can be bonded together. It is not possible to have both 1 GbE and 10 GbE network interfaces in the bonded interface set.

2. All of the SC bonded SAN network interfaces must have the same network visibility, all ports must be connected to the same network switch, or if multiple switches, they must be fully meshed. It is important that all Host Ethernet ports be visible from all SC Ethernet ports through one or more switches as the traffic for any Host can be moved to any SC port at any time.

3. SC Balance-ALB bonding is incompatible with any other switch supported bonding methods such as EtherChannel or IEEE 802.3ad, also known as port trunking, link aggregation, Ethernet trunking, NIC teaming, port channel, port teaming, port trunking, link bundling, Multi-link trunking (MLT), etc.

   There should be no EtherChannel or IEEE 802.3ad protocol configuration set in the network switches that the SC bonded network ports are connecting to.

   These protocols can be used on the Host side of the switch for Hosts supporting multiple Ethernet interfaces if desired.

4. When Balance-ALB network bonding is in use, on the SC the host to NIC port assignments are rebalanced based on load every 10 seconds. Given this, when running benchmarks, you should provide a 30 second test ramp-up time to allow the final network ports assignments to settle based on the new work-load.

A6.3 Jumbo Frames

The Storage Concentrator supports the use of Jumbo Frames on the SAN network ports. Ethernet "Jumbo Frames" are frames with more than 1,500 bytes of payload, the Maximum Transmission Unit (MTU).

Ethernet has used 1514 byte frame sizes since it was created. The CRC adds another 4 bytes for a total of 1518. "Jumbo Frames" can extend the Ethernet frame size up to a maximum of 16124 bytes (plus 4 CRC), depending upon the network hardware in use.

Using "Jumbo Frames" significantly reduces protocol overhead and also latency. If an 8K SCSI write was being done, this would take about 5 1500 MTU Ethernet frames, but only 1 9000 MTU frame. However, when there are large I/O's on the wire, Jumbo frames actually increase latency for small I/O's because it takes longer for 9000 octet frames to clear the wire so that small block I/O's will suffer more latency.

1. Many network switches and Host network interface cards/drivers are not able to process "Jumbo Frames". Unmanaged "dumb" switches usually cannot since there is no management interface through which to enable them.
2. To use "Jumbo Frames" effectively, all devices on the SAN network must be capable of processing them. In an iSCSI network, that includes Host initiators, switches, etc. Most initiators and switches are only capable of supporting Jumbo Frames when they are configured with that option. If you intend to use Jumbo Frames on the Storage Concentrator, set all other devices such as initiators and switches to use Jumbo Frames before enabling the **Use Jumbo Frames** option on the SC.
3. When using Jumbo Frames on the network, all equipment must be configured with the same Frame size.

   Using different "Jumbo Frame" sizes can work, but performance is impaired when the path MTU size needs to be rediscovered.

   Typical Jumbo MTU Size values that are seen in SAN networks are 4074, 8000, 9000, 16110.
4. Note that the frame size values the SC used are in terms of the TCP MTU size, but other equipment may use the Ethernet frame size. If so, that equipment may need to add to the MTU value 14 bytes for the Ethernet header, and 4 bytes for the CRC.
5. The Storage Concentrator supports Jumbo Frames when enabled with the **Use Jumbo Frames** check-box.

   When **Use Jumbo Frames** is enabled, the **Jumbo MTU Size** field specifies the MTU size that should be used. This value must be within the valid range supported by the Storage Concentrator network hardware, and varies with the type of hardware installed.

   Of all of the ports that are in the SC bonding set, the type with the smallest maximum MTU size will limit the maximum "Jumbo Frame" size that can be used. The number of ports in the bonding set does not influence the maximum MTU size.

   If an invalid value is chosen, an error pop-up will occur indicating the accepted range for that system. A default value appears when **Use Jumbo Frames** is first enabled.

For ISC-1G the max MTU with 6 bonded ports will be 9216. The limitation is the MB Ethernet port, if only the Intel NIC card ports were bonded, the max MTU could go up to 16110.

For ISC-10G the max MTU with 4 bonded ports will be 16110 assuming an Intel 10 GbE NIC card used.

For Voyager-1G the max MTU with 6 bonded ports will be 9216.

For Voyager-10G the max MTU with 4 bonded ports will be 16110.

6. It is essential to verify end-to-end that Jumbo Frame use is configured and working properly. The SC GUI "System -> Network -> Local iSCSI Data Port" page "Ping" feature with the "Test Jumbo Frame" option selected can be useful to confirm proper end-to-end Jumbo frame function with each Host on the SAN.
7. "Dumb" network switches that are unmanaged do not support jumbo frames. An attempt to use them with such switches will result in strange intermittent results. Small frames such as those needed for iSCSI Login, etc. will work fine, but as soon as the SAN traffic I/O block sizes increase, the I/O's will fail, or will work with long delays. Do not enable "Jumbo Frame" use in these environments.
8. In a mixed I/O environment where there are both small and large heavy I/O applications, "Jumbo Frames" can sometimes increase the latency for the small I/O's application causing them to take longer. Using a different host Ethernet interface for the volumes that require small block I/O's can alleviate this effect.

A6.4 SC SAN Multipath Interfaces

The SC ships with, and is generally configured with multiple 1 GbE or 10 GbE network interfaces bonded at the network level using the 'bonding' driver in 'balance-alb' mode.

This 'balance-alb' bonding provides network traffic distribution between multiple iSCSI host IP interfaces and the SC in both the transmit and receive directions.

However, with network bonding, there is never more than 1 GbE network interface used between the SC and any single host IP interface.

For the general case with multiple physical hosts, and the number of host IP interfaces is greater than or equal to the number of bonded interfaces, and all of the hosts are actively performing I/O, from the SC perspective, all if the interfaces will be in use with 'balance-alb' bonding.

However in virtualization environments, there is often either a single physical host, or more than one 1/10 GbE of iSCSI bandwidth is needed to a single host.

There are guidelines and configuration steps specific to the different host platforms. Please refer to the host specific Multipath Interface sections elsewhere in this document.

This page is intentionally left blank.

Appendix 7

Systems Considerations

A7.1 Windows Systems Considerations

This section is intended to provide some guidelines for users installing StoneFly IPSANs on windows systems. These guidelines apply to all StoneFly IPSANs unless specifically noted.

A7.1.1 Known Issues

#	Description
1	Windows Server 2000 hosts must not access volumes if they contain a Windows Server 2003 NTFS file system. The NTFS files systems are not compatible. This can be done inadvertently by assigning a Snapshot from a volume in use by a Windows Server 2003 host to be accessed by a Windows Server 2000 host. The Windows Server 2000 host may crash.
2	Windows 2000 and 2003 Server-family hosts using Dynamic Disks will become "offline" following a reboot of the host due to the timing of activation of dynamic volumes during the boot. In order to fix this after a reboot, you need to re-activate the disks and the file shares using the Disk Management control panel or a boot-time script. This is a limitation of Microsoft iSCSI initiator for windows 2000 & 2003. Microsoft initiator for Windows Server 2008 does not have this limitation.

A7.1.2 Modifying Windows IO Timers

A7.1.2.1 General Timeout value change

When a Storage Concentrator in a Cluster is performing a FailOver operation, or if an USO reboots for any reason, enough time must be allowed for the underlying storage subsystems.

To allow sufficient time, the Windows disk timeout value must be changed. The default value is 10 seconds. If time out is left at 10 seconds, the Operating System would think that an error has occurred and potentially return an error to the application.

IMPORTANT NOTE: When using Mirroring or Campus Mirroring, the value in the registry must be set to 300 seconds.

Warning: Altering the Windows registry incorrectly can render your system inoperable.

Perform the following steps with caution.

Refer to your documentation CD and go to the directory called "Windows-Timeout-Value-Change" or contact StoneFly support and obtain a script that changes the timeout values within windows registry. The script name is: DiskTimeOutValue.reg.

Contact StoneFly support and obtain a script that changes the timeout values within windows registry. The script name is: DiskTimeOutValue.reg.

Copy the file to your Windows Host.

Double click on the file. This will modify your Disk driver TimeOutValue to 12C. 12C is the HEX value for 300 seconds.

Reboot your Host for the change to take effect.

A7.1.2.2 Timeout value change if you use MPIO

An additional timer is used by the MPIO features to look for failed links. It is a Windows registry entry called "LinkDownTime". It is located in registry hives that may not be easy to find at first. Run "regedit" and search the registry for the value "LinkDownTime". The name may appear in several places in CurrentControlSet and other ControlSet's. Only modify the values in CurrentControlSet. Set "LinkDownTime" to 90 seconds on all MPIO devices on all MPIO servers. This setting ensures that Windows MPIO hosts connect to their volumes after a Failover.

Non-MPIO Windows hosts can benefit from an additional setting as well. Repeat the search as you did for of LinkDownTime, but this time search for occurrences of "MaxRequestHoldTime". This value should be reset from the default of 60 seconds to 90 seconds. Once again this is only for devices in the CurrentControlSet. There may be multiple locations in the registry, use the F3 button to find the correct registry hive.

A7.1.3 *Turn off Windows Networking TCP Delayed ACKs*

From the Microsoft iSCSI Software Initiator 2.x Users Guide:

Slow performance may occur during network congestion when RFC 1122-delayed acknowledgements extend the error recovery process. In these situations, the default 200 millisecond delay on the acknowledgement can significantly impact read bandwidth. Use of multipathing solutions which load balancing read requests across ports, increases the likelihood that simultaneous read completions from multiple ports will result in network congestion. This increases the likelihood of experiencing the problem.

As specified in RFC 1122, Microsoft TCP uses delayed acknowledgments to reduce the number of packets that are sent on the media. Instead of sending an acknowledgment for each TCP segment received, TCP in Windows 2000 and later takes a common approach to implementing delayed acknowledgments. As data is received by TCP on a particular connection, it sends an acknowledgment back only if one of the following conditions is true:

No acknowledgment was sent for the previous segment received.

A segment is received, but no other segment arrives within 200 milliseconds for that connection.

Typically, an acknowledgment is sent for every other TCP segment that is received on a connection unless the delayed ACK timer (200 milliseconds) expires. You can adjust the delayed ACK timer by editing the registry as outlined in the workaround below.

Fix:

Modify the TCP/IP settings for the network interfaces carrying iSCSI traffic to immediately acknowledge incoming TCP segments. This workaround solves the read performance issue. The procedure to modify the TCP/IP settings is different for Windows 2000 servers and Windows 2003 servers. Follow directions appropriate for the version you are running on your servers.

Note: These TCP/IP settings should not be modified for network interfaces not carrying iSCSI traffic as the increased acknowledgement traffic may negatively affect other applications.

On a server that runs Windows 2000 SP3 or later, follow these steps:

Start Registry Editor (Regedit.exe).

Locate and then click the following registry subkey:

HKEY_LOCAL_MACHINE\SYSTEM\CurrentControlSet\Services\Tcpip\Parameters\Interfaces

the interfaces will be listed underneath by automatically generated GUIDs like {064A622F-850B-4C97-96B3-0F0E99162E56}

Click each of the interface GUIDs and perform the following steps:

a. Check the IPAddress or DhcpIPAddress parameters to determine whether the interface is used for iSCSI traffic. If not, skip to the next interface.
b. On the Edit menu, point to New and then click DWORD value.
c. Name the new value TcpDelAckTicks and assign it a value of 0.

Exit the Registry Editor.

Restart Windows for this change to take effect.

On a server that runs Windows Server 2003 SP1 or later, follow these steps:

Start Registry Editor (Regedit.exe).

Locate and then click the following registry subkey:

HKEY_LOCAL_MACHINE\SYSTEM\CurrentControlSet\Services\Tcpip\Parameters\Interfaces

the interfaces will be listed underneath by automatically generated GUIDs like {064A622F-850B-4C97-96B3-0F0E99162E56}

Click each of the interface GUIDs and perform the following steps:

a. Check the IPAddress or DhcpIPAddress parameters to determine whether the interface is used for iSCSI traffic. If not, skip to the next interface.
b. On the Edit menu, point to New, and then click DWORD value.
c. Name the new value TcpAckFrequency, and assign it a value of 1.

Exit the Registry Editor.

Restart Windows for this change to take effect.

A7.1.4 *Using Microsoft MPIO features*

Support for Microsoft MPIO was introduced in Release 4.2.0.10 and later releases continue to support for this feature. Please review all issues and suggestions contained in this section to insure correct operation of MPIO. See **notes** above concerning Windows registry settings.

A new installation with no volumes created does not need to do anything to have MPIO supported volumes. If the Storage Concentrators are upgraded to the new release, any legacy volumes created using SW prior to 4.2.0.10 may not be enabled for MPIO features. The discussion that follows describes the actions required to convert all legacy volumes (created prior to 4.2.0.10) into MPIO enabled volumes. **IMPORTANT**: All legacy volumes are converted at the same time. Schedule a maintenance period to do these steps.

A7.1.5 *MPIO GUI Related Items*

If there are any non-MPIO enabled volumes, then a checkbox to "MPIO Enable" legacy volumes will appear in the System->admin->general page. If all volumes are enabled for MPIO this checkbox does not appear on the System screen.

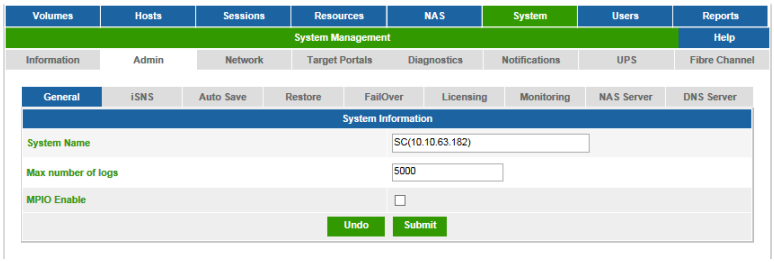


Figure App.-3

If there are non-MPIO enabled volumes, the volume detail page describes if the volume is not MPIO enabled the screen appears as:

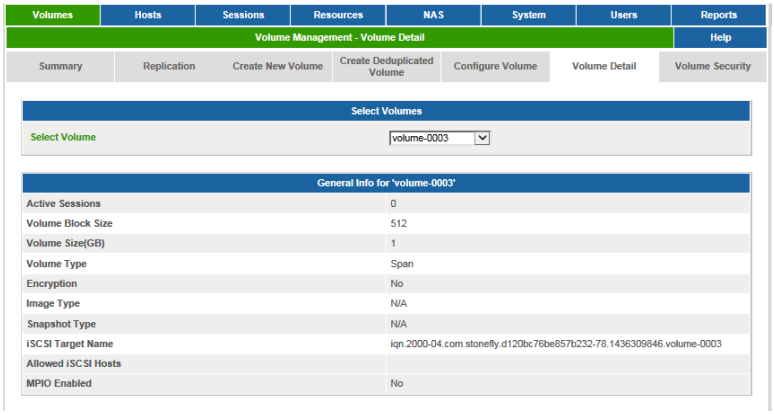


Figure App.-4

Or, if the volume is MPIO enabled the screen appears as:

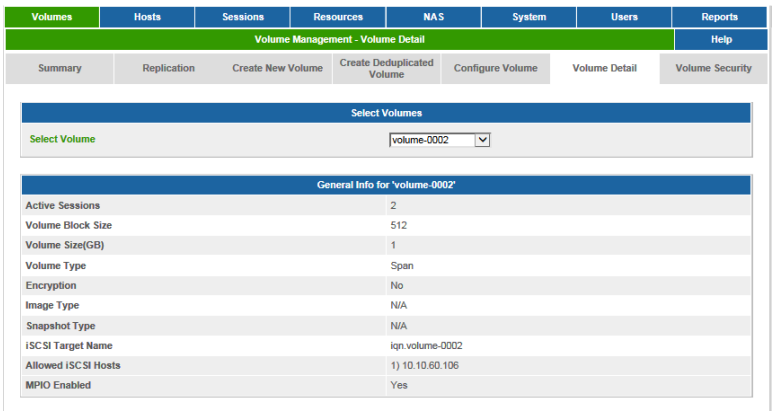


Figure App.-5

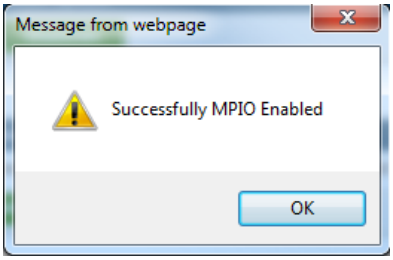
All volumes created using version 4.2 or higher will be automatically MPIO enabled, those created before 4.2 will not be MPIO capable until after the administrator performs a conversion using the MPIO Enable checkbox in the General Admin screen.

Doing the conversion:

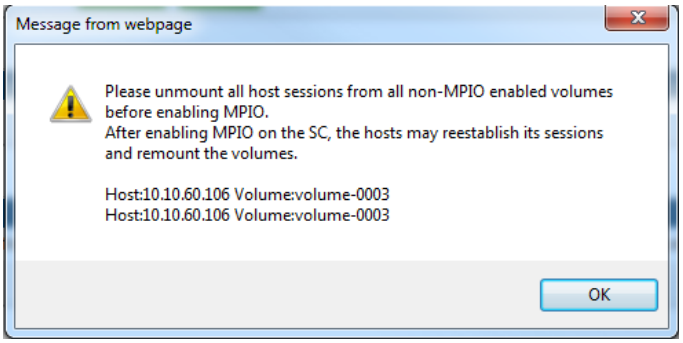
NOTE: All host sessions to non-MPIO Enabled volumes must first be closed. Do this from the Host system using the iSCSI Initiator management features.

Click MPIO Enable checkbox and submit.

On Success:



If there are sessions to non-MPIO enabled volumes an error dialog will appear listing sessions that should be closed at the host.



IMPORTANT: The Initiator should not be installed with MPIO unless the SC has been MPIO enabled.

After enabling MPIO on the SC, the initiator may be reinstalled to use MPIO.

NOTES on MPIO: Several new issues have been raised concerning MPIO functionality in Microsoft Windows 2003 Server.

The following Microsoft Hotfix is required when using MPIO on Windows 2003: An updated Storport storage driver (version 5.2.3790.2497) is available for Windows Server 2003
http://support.microsoft.com/kb/903081/

The following Microsoft Windows 2003 Hotfix is recommended for use with MPIO:

An updated Storport storage driver (version 5.2.3790.2630) is available for Windows Server 2003 systems at:http://support.microsoft.com/kb/912944/en-us#kb2

StoneFly has experienced some IO failures when the iSCSI connection is broken and re-established repeatedly. An example of this is if the Ethernet connection between the server and the Storage Concentrator is broken and fixed such as pulling one of the cables and re-inserting it over and over. After approximately 20 of these connection losses the Windows 2003 operating system will forget about the drive and it will not be managed by the Disk Management application. In all our test cases the iSCSI connection was re-established without any problems. Only the communications between the operating system and the Microsoft iSCSI initiator seems to be broken. Just logout of the volume at the Microsoft iSCSI initiator properties panel and then log back in. The volume will be established properly in the system.

To use Microsoft MPIO with SC Failover you need to extend the **'PDORemovePeriod'** MPIO timeout setting (per initiator host) using the following procedure:

- Run 'wbemtest.exe'
- Click on 'Connect'. Change 'Namespace' to 'root\wmi' and click on 'Connect' again (or click 'Login').
- Click on 'Enum Classes'. Select 'Recursive' and click OK
- Scroll down and select 'MPIO_TIMERS_COUNTERS' and double click that item
- Click on 'Instances' and double click the instance entry listed
- Scroll down and double click on 'PDORemovePeriod'
- Change the value from 2 seconds to 300 seconds (decimal). Then click on 'Save

Property'
- Now click on 'Save Object' and close all windows and panels
- Restart 'wbemtest.exe'
- Repeat steps 1 through 6 and verify that the PDORemoveperiod' is indeed set to 300 seconds
- Reboot the machine. Now the MPIO 'PDORemovePeriod' timer should be set to 300 seconds

A7.1.6 *Missing Shares to iSCSI storage after Windows reboot*

If you experience problems with Missing File Shares to iSCSI storage after a Windows OS reboot, please consult the Microsoft Knowledge Base article:

> File shares on iSCSI devices may not be re-created when you restart the computer
> http://support.microsoft.com/?kbid=870964

A7.1.7 *MPIO (Multi-path IO) in Microsoft Windows environments*

Microsoft Windows multi-pathing is performed through MPIO on the host. MPIO or Multi-path IO is used to increase performance by using more than one port of the SAN or more than one port of the Server to communicate with the SAN. There are two types of MPIO: 1- MPIO on the SAN, 2- MPIO from the server.

The easiest way is to use MPIO on the SAN is to have the SC SAN ports bonded together. By bonding the SC SAN ports together, StoneFly software handles MPIO automatically and assigns each host session IP to each SC port of the SAN. So SAN MPIO session distribution is accomplished automatically at the SC when the network ports are bonded together.

For MPIO to be available on the host using Windows 2003, the Microsoft iSCSI initiator should be installed with the MPIO option selected during installation. If the Microsoft iSCSI initiator was already installed without the MPIO option, it will be necessary to uninstall and reinstall it again to enable MPIO.

On Windows 2008, the iSCSI Initiator is pre-installed. To enable MPIO on Windows 2008 systems please follow the procedure:

> Configuring iSCSI MPIO on Windows Server 2008 R2 (full)
> http://blogs.technet.com/b/migreene/archive/2009/08/29/3277914.aspx

Depending on how many ports you have on your server, please review the notes below:

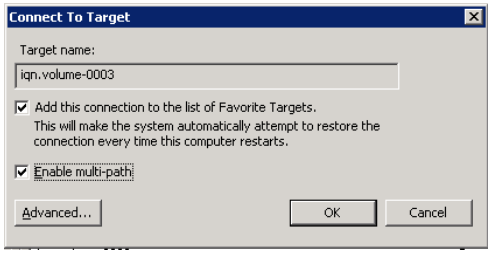
1 - MPIO from a server with one NIC to an SC that has two volumes: You do not need MPIO on the host, and you just allocate each volume to each SC in the case of a clustered IP Storage and login once for each volume.

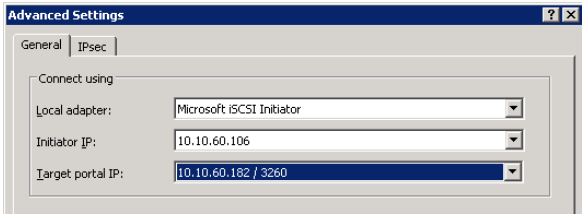
2 - If you have two iSCSI ports or more on your server host:

StoneFly IPSAN iSCSI ports should be bonded together from StoneFly GUI are assigned iSCSI IP address of the SC.

In a case of a clustered SAN, you need to know the SC cluster IP address.

For added bandwidth, an iSCSI session should be created for each server host IP network interface that has an IP address on the SAN network. All of these will connect to the same SC iSCSI IP address used by the SC bonded network interfaces – there is no other SC configuration required. The bonding logic will delegate the host iSCSI sessions to an appropriate SC network interface automatically.





For each host SAN IP address, use the Microsoft iSCSI Initiator GUI to create an iSCSI session using a different "Source IP:" for each session, making sure to check "Use multi-path" on the second and subsequent sessions.

You select the first host SAN "Source IP:" address, select the common/bonded SAN IP address for your SC for the "Target portal:", and then select ok and login.

Then repeat the same steps above, but this time you select the remaining host SAN IP addresses.

MPIO runs in different modes, failover, round robin, etc. Round-robin modes should be used for improved network bandwidth as multiple network interfaces can be used at the same time. Please refer to Microsoft iSCSI initiator documentation that explains how to enable MPIO and configure for round robin, etc. To confirm that you are in the round robin mode, go to Microsoft iSCSI initiator MPIO Tab, and make sure MPIO is set for Active-Active-Round Robin.

A7.1.8 *StoneFly Multi-path (MPath) and MPIO in Microsoft Initiator*

Microsoft Windows multi-pathing is performed through MPIO on the host.

Windows MPIO use does not require that any SC 'MPath' interfaces be defined; the SC SAN network ports can and should usually be left set to 'Bonded'. The use of SC 'MPath' interfaces with Windows MPIO is not necessary, and is not generally recommended.

However, should SC 'MPath' interfaces need to be defined (perhaps to support other iSCSI host types in the same SAN), the Windows iSCSI initiator and MPIO can also use the 'MPath' interfaces by configuring the Windows iSCSI initiator properly.

The following example shows how to configure Microsoft Windows Server 2008 with multiple session MPIO using SC MPath network interfaces in addition to SC Bonded interfaces. Although the Windows iSCSI screens are different in Windows 2000 and 2003 server, the concepts are similar.

1 - Assume the following configuration for a host with SAN IP addresses of 10.10.60.106, 10.10.61.106 and 10.10.62.106, and an SC with one or more bonded interfaces at IP address 10.10.60.182, and an 'MPath' interface on network 10.10.61.0 and 10.10.62.0:

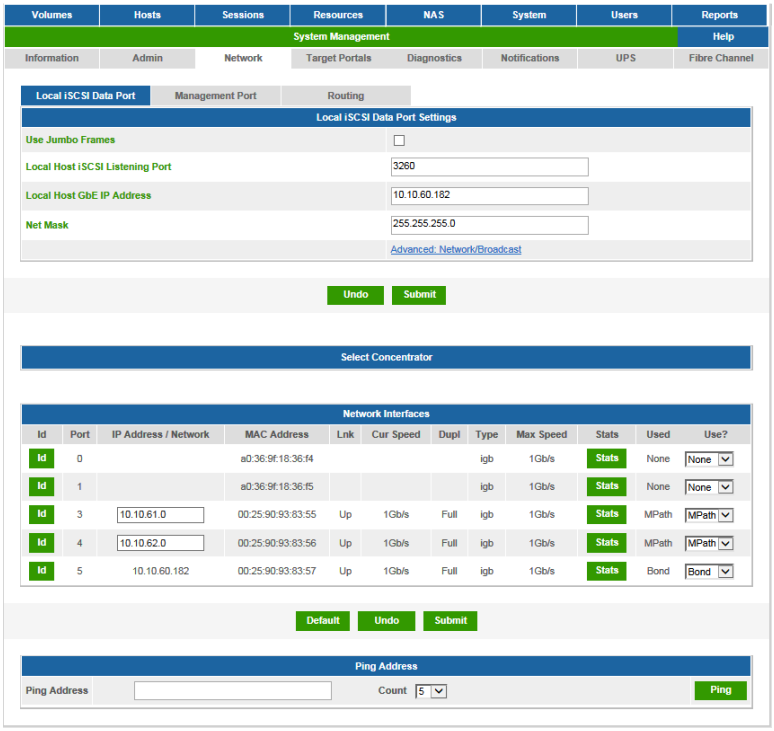
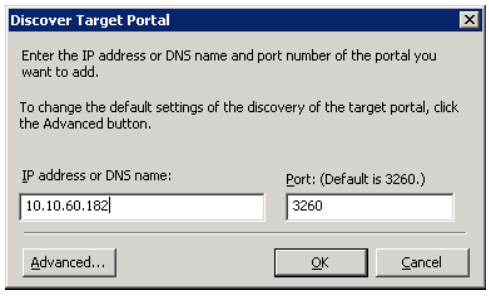
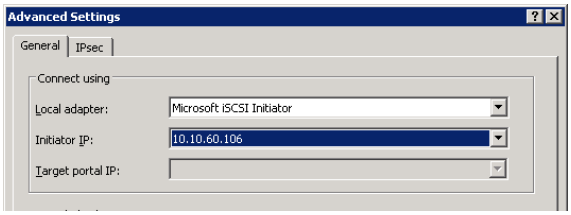


Figure App.-6

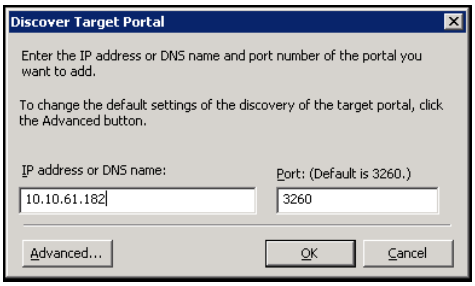
2 - Configure the Windows iSCSI initiator for discovery on both the SC Bonded IP address, and on each 'MPath' network interface, first to the main Bonded 10.10.60.0 network, then the 10.10.61 and 10.10.62 MPath networks. First create a Discovery Target Portal for the host and SC on the 10.10.60.0 network:

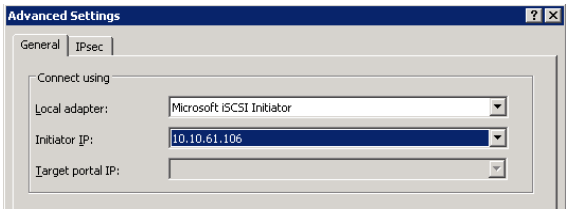


3 - Make sure to change the Advanced Settings "Local adapter" and "Initiator IP:" values appropriate for each network being configured:



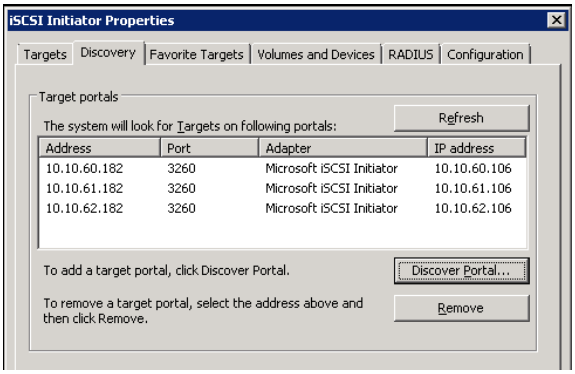
4 - Now create a Discovery Target Portal for the host and SC on the 10.10.61.0 network:



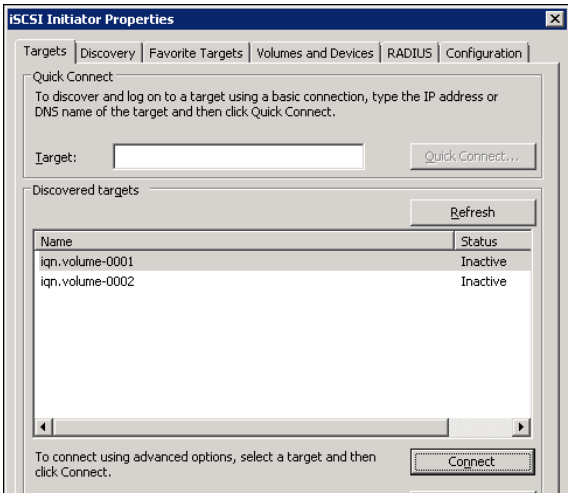


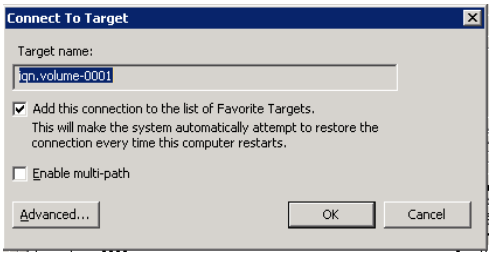
Create a Discovery Target Portal for the host and SC on the 10.10.62.0 network.

5 - Verify that the target portals are configured properly:

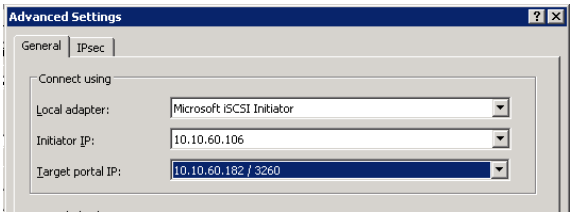


6 – Create persistent iSCSI sessions, one for each Windows host network IP address, to the appropriate SC SAN Bonded and MPath interfaces. Select a "Discovered target" and click Connect:

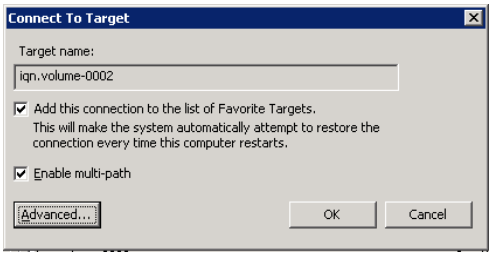




7 - For the SC Bonded port(s) network make sure to change the Advanced Settings "Local adapter", "Initiator IP:", and "Target portal IP:" values appropriate for each network being configured as below:



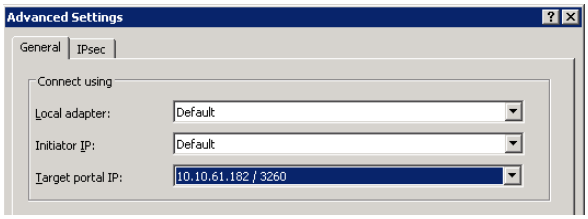
8 - Now create a iSCSI session through the 10.10.61.0 MPath network. Make sure to check "Enable multi-path" for the second and subsequent sessions:



9 - For the SC MPath network(s) make sure to change the Advanced Settings "Local adapter", "Initiator IP:", and "Target portal IP:" values appropriate for each network being configured as below.

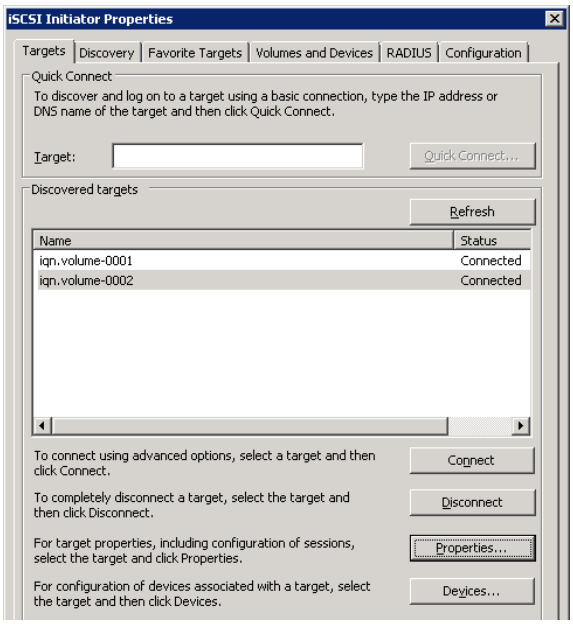
Note that the settings are different when the non-default "Target portal IP:" is being

used, the "Local adapter:" and "Initiator IP:" settings **must** be left set to "Default" in order to be able to select the desired MPath network portal IP:

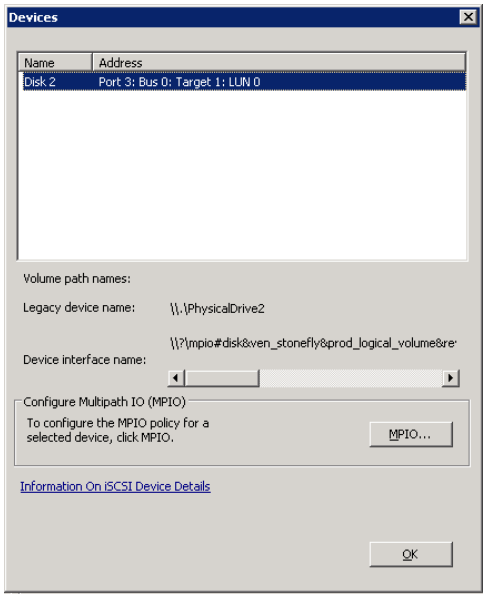


Now create an iSCSI session for the same target through the 10.10.62.0 MPath network. Make sure to check "Enable multi-path" and select the desired MPath network portal IP.

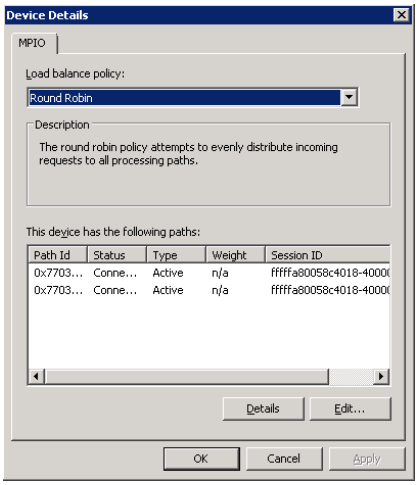
10 - To achieve a bandwidth advantage using multiple iSCSI network interfaces simultaneously, the MPIO policy must be set to "Round Robin". Start by selecting a target, clicking "Devices" button:



11 - Click on each Target Devices' MPIO button:



12 - Change to "Round Robin" if not already set so.



13 - On the SC Sessions page, verify that the iSCSI sessions are correctly connected on the main "Bonded" and MPath network(s):

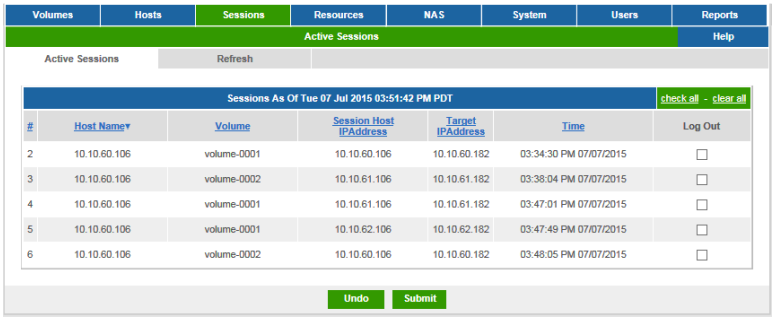


Figure App.-7

A7.2 Linux Systems Considerations

This section is intended to provide some guidelines for users installing StoneFly IPSANs on Linux systems. These guidelines apply to all StoneFly IPSANs unless specifically noted.

A7.2.1 *Modifying Linux IO Timers for Voyager and OptiSAN*

When a Storage Concentrator in a Cluster is performing a FailOver operation, enough time must be allowed for the underlying storage subsystems. When the Storage Concentrator determines that a failure has occurred, the FailOver process begins.

To allow sufficient time for the FailOver, the Linux disk I/O timeout value must be changed. The default value is 30 seconds. That value is not large enough to allow a full FailOver. The Operating System by default would think that an error has occurred and potentially return an error to the application. By extending the timeout value, the Storage Concentrator has the appropriate time needed to perform the error analysis, and then perform the FailOver.

The SCSI I/O timeout should be extended to be at least 90 seconds.

A common way to increase the I/O timeout is with a script that is added to startup. An example script is:

```
#!/bin/bash
# The default timeout value for each SCSI Device is 30.
# This value is too low to handle iSCSI cluster F/O's.
# To show the default timeout value for each SCSI Device:
for x in `ls /sys/class/scsi_device/`;
do
   cat /sys/class/scsi_device/$x/device/timeout;
done
# To modify the timeout value to 80 secs for each SCSI Device:
for x in `ls /sys/class/scsi_device/`;
do
```

```
    echo 90 > /sys/class/scsi_device/$x/device/timeout;
done
```

A7.3 VMware System Considerations

This section is intended to provide some guidelines for users connecting StoneFly IPSANs to VMware systems. These guidelines apply to all StoneFly IPSANs unless specifically noted.

A7.3.1 Modifying VMware ESX Default Timers

When a Storage Concentrator (SC) in a Cluster is performing a FailOver operation, enough time must be allowed for the underlying storage subsystems. When the Storage Concentrator determines that a failure has occurred, the FailOver process begins. To allow sufficient time for the fault detection and the FailOver, some VMware ESX default timeout values must be changed.

There are other cases where an iSCSI session must be temporarily disconnected, or where I/O's are temporarily suspended, even when not in an SC cluster. For example, when expanding snap-space, it is necessary to disconnect the hosts for a short period.

So the ESX timer adjustments the changes are required for an SC cluster environment for ESX I/O's to survive an SC F/O, but are recommended to better handle other infrequent situations in both clusters and non-clusters.

The following VMware ESX timer adjustments should be made:

1. Using the "VMware Client", increase the ESX "Configuration -> Advanced -> Settings -> SCSI" 'SCSI reaborttimeout' from 5000 to 50000 (decimal).
2. Using the "VMware Client", increase the ESX "Configuration -> Advanced -> Settings -> SCSI" 'SCSI scantimeout' from 1000 to 60000 (decimal).

NOTE: The ESX server must be rebooted for these setting changes to take effect.

A7.3.2 VMware Guest OS Considerations

It is not sufficient to only change the VMware timers without also considering the guest OS.

Otherwise, although VMware might wait longer, the hosts would timeout first and could still fail the I/Os during an SC fail-over.

Please refer to the related "Considerations …" documents for the recommended settings for the VMware guest OS types.

A7.3.3 VMware Multi-pathing.

In virtualization environments, there is often only a single physical host, and/or more than one 1 (or 10) GbE of iSCSI bandwidth is needed to the virtualization host.

A common problem use case for SC Multipath interfaces is where there is a single USO connected to a single VMware ESX server with many VM's.

Of course, if the ESX host only uses one network interface, no more than 1 (or 10) GbE of traffic can ever flow.

The direct approach is to add additional network interfaces as 'VMkernel' ports to the ESX networking configuration. This only adds network layer bonding on the ESX side, which will not result in the use of any more than one network interface at a time.

Instead, on the ESX side, storage 'multi-pathing' needs to be enabled. This was an experimental feature in ESX 3.5, but is now fully supported for ESX 4.0.

Before configuring the ESX host for multi-pathing, the SC should have some of its network interfaces defined as 'Multipath' interfaces. Unused network interfaces can be wired up and designated as SC 'Multipath' interfaces, or existing 'Bonded' interfaces can be reassigned. Each SC 'Multipath' interface must be on a unique SAN data subnet which is assigned during the configuration. Consult the SC User Guide, and SC GUI "Local iSCSI Data Port Settings" page and online help. Note that each iSCSI host must be enabled to be able to discover multipath interfaces; this is done on the SC GUI "Host Management Detail" page.

The instructions to properly configure "Host-based multipathing" in ESX 4.0 can be found here:

> iSCSI SAN Configuration Guide
> ESX 4.0, ESXi 4.0, vCenter Server 4.0
> EN-000110-00
> http://www.vmware.com/pdf/vsphere4/r40/vsp_40_iscsi_san_cfg.pdf
> Activate Multipathing for Software iSCSI Initiator
> Use this task only if your ESX/ESXi host has two or more physical network adapters that you designate for iSCSI traffic. This task explains how to activate host-based multipathing for your host by connecting the software iSCSI initiator to iSCSI VMkernel ports that you created for the network adapters.
> ...

The procedure involves disabling ESX bonding of these interfaces so that multipathing can be used.

Note that the two (or more) 'VMkernel' interfaces should be on different SAN sub-nets, otherwise network routing would send traffic through a single interface. Also, if both ESX VMkernel ports are on the same sub-net, ESX would create twice as many paths than were needed.

Note that the default for ESX 4.0 multipathing is not to use Round-robin, but Fixed, which only uses a single network interface at a time. To change the settings for each storage device to round-robin, refer to:

> iSCSI SAN Configuration Guide
> http://www.vmware.com/pdf/vsphere4/r40/vsp_40_iscsi_san_cfg.pdf
> Setting a Path Selection Policy
> For each storage device, the ESX/ESXi host sets the path selection policy based on the claim rules defined in the /etc/vmware/esx.conf file.
> ...

Even with ESX round-robin though, by default, the policy is to not switch paths until 1000 I/O's have been sent down one path -- only then, will it switch to the next path. To get more simultaneous use of the paths, the following 'esxcli' commands can be used (no GUI access for this):

```
# esxcli nmp device list
# esxcli nmp roundrobin getconfig -d
naa.600174d0010000000113003048318438
# esxcli nmp roundrobin setconfig -d
naa.600174d0010000000111003048318438
        type iops --iops 3
```

This changes the path rotation from every 1000 I/O's to every 3.

However, there are issues in ESX that cause this setting to be lost at reboot. Unfortunately, instead of the '--iops' value getting reset back to 1000, it instead gets set to 1449662136 which effectively disables round-robin.

Note also, the 'bytes' rotation policy does not appear to work at all, and round-robin is also disabled.

So the current recommendation is to either not attempt to change the round-robin policy settings at all, and use the default of 'iops' at 1000, or to write an ESX or 'vCenter' script that runs at ESX host boot to re-instantiate these settings.

A7.4 Citrix XenServer System Considerations

This section is intended to provide some guidelines for users connecting StoneFly IPSANs to Citrix XenServer systems. These guidelines apply to all StoneFly IPSANs unless specifically noted.

A7.4.1 *Modifying Citrix XenServer Default Timers*

When a Storage Concentrator (SC) in a Cluster is performing a FailOver operation, enough time must be allowed for the underlying storage subsystems. When the Storage Concentrator determines that a failure has occurred, the FailOver process begins. To allow sufficient time for the fault detection and the FailOver, some XenServer default timeout values must be changed.

There are other cases where an iSCSI session must be temporarily disconnected, or where I/O's are temporarily suspended, even when not in an SC cluster. For example, when expanding snap-space, it is necessary to disconnect the hosts for a short period.

So the XenServer timer adjustments the changes are required for an SC cluster environment for XenServer I/O's to survive an SC F/O, but is recommended to better handle other infrequent situations in both clusters and non-clusters.

The following Citrix XenServer timer adjustments should be made:

The SCSI I/O timeout should be extended to be at least 90 seconds.

A common way to increase the I/O timeout is with a script that is added to startup.

An example script is:

```
#!/bin/bash
# The default timeout value for each SCSI Device is 30.
# This value is too low to handle iSCSI cluster F/O's.
# To show the default timeout value for each SCSI Device:
for x in `ls /sys/class/scsi_device/`;
do
   cat /sys/class/scsi_device/$x/device/timeout;
done
# To modify the timeout value to 80 secs for each SCSI Device:
for x in `ls /sys/class/scsi_device/`;
do
   echo 90 > /sys/class/scsi_device/$x/device/timeout;
done
```

NOTE: The XenServer must be rebooted to test that these setting changes take effect

A7.4.2 *Citrix XenServer Guest OS Considerations*

It is not sufficient to only change the XenServer timers without also considering the guest OS.

Otherwise, although XenServer might wait longer, the hosts would timeout first and could still fail the I/Os during an SC fail-over.

Please refer to the related "Considerations …" sections for the recommended settings for the XenServer guest OS types.

Although the XenServer platform management and SAN network interfaces can be configured to use Jumbo frames, Ethernet Jumbo frames do not appear to be supported by the XenServer virtual switches. As such, using Jumbo frames on a Guest OS, especially when the Guest has its own iSCSI SW initiator configured for direct use of the SAN would not be useful.

A7.4.3 *Citrix XenServer Multi-pathing.*

In virtualization environments, there is often only a single physical host, and/or more than one 1 (or 10) GbE of iSCSI bandwidth is needed to the virtualization host.

A common problem use case for SC Multipath interfaces is where there is a single USO connected to a single Citrix XenServer server with many VM's.

Of course, if the XenServer host only uses one network interface, no more than 1 (or 10) GbE of traffic can ever flow.

The direct approach would be to also configure XenServer to use network bonding. This only adds network layer bonding on the XenServer side, which will **not** result in the use of any more than one network interface at a time.

Instead, on the XenServer side, storage 'multi-pathing' needs to be enabled.

Before configuring the XenServer host for multi-pathing, the SC should have some of its network interfaces defined as 'Multipath' interfaces. Unused network interfaces can be wired up and designated as SC 'Multipath' interfaces, or existing 'Bonded' interfaces can be reassigned. Each SC 'Multipath' interface must be on a unique SAN data subnet which is assigned during the configuration. Consult the SC User Guide, and SC GUI "Local iSCSI Data Port Settings" page and online help. Note that each iSCSI host must be enabled to be able to discover multipath interfaces; this is done on the SC GUI "Host Management Detail" page.

First, there must be multiple XenServer network interfaces configured for iSCSI storage. Each of these should be on a different SAN sub-net used for SC multipathing. This can be done using the XenCenter management GUI at "Server -> Management Interfaces -> New Interface", or using the procedure documented below:

> XenServer Administrator's Guide 5.5.0
> Published November 2009
> 1.1 Edition
> http://support.citrix.com/servlet/KbServlet/download/20636-102-641538/reference.pdf
> Configuring a dedicated storage NIC XenServer allows use of either XenCenter or the xe CLI to configure and dedicate a NIC to specific functions, such as storage traffic.
> …

Note that the XenServer iSCSI SAN network interfaces should be dedicated use, and not also be used for VM's. Consult the XenServer online help topic "Dedicating a NIC to a specific function" to configure the system properly.

XenServer storage multipath must then be enabled. Follow the procedure documented here:

> XenServer Administrator's Guide 5.5.0
> Storage Multipathing
> Dynamic multipathing support is available for Fibre Channel and iSCSI storage backends. By default, it uses round-robin mode load balancing, so both routes have active traffic on them during normal operation. You can enable multipathing in XenCenter or on the xe CLI.
> ...

Note that the 'host-param' 'other-config' multipath settings can be more easily enabled using the GUI at "Server -> Properties -> Multipathing".

By default, storage multipathing will be fail-over, not round-robin. This can be changed by, on the XenServer console, edit the '/etc/multipath.conf' file by adding a new 'device' to the 'devices {" section:

```
# vi /etc/multipath.conf
devices {
   device {
      vendor "StoneFly"
      product "Logical Volume"
      path_grouping_policy multibus
```

```
        getuid_callout "/sbin/scsi_id -g -u -s /block/%n"
        rr_weight uniform
        rr_min_io 100
        path_checker tur
        failback immediate
        no_path_retry 12
    }
    …
```

Reboot the XenServer for this edit to take effect.

Then confirm that there are multiple active paths to the iSCSI storage using the "multipath -ll" command:

```
[root@xenserver-223 ~]# multipath –ll
3600174d001000000012000304831843 8dm-0 StoneFly,Logical Volume
[size=100G][features=1 queue_if_no_path][hwhandler=0]
\_ round-robin 0 [prio=2][enabled]
\_ 13:0:0:0 sdd 8:48  [active][ready]
\_ 14:0:0:0 sde 8:64  [active][ready]
```

More information about the 'multipath' command output can be found at:

MultipathUsageGuide - Device-mapper and LVM2 Wiki
http://sources.redhat.com/lvm2/wiki/MultipathUsageGuide

For round-robin to be enabled there should be multiple active paths to each volume.

This page is intentionally left blank.

Appendix 8

Redundant SAN Network

In high availability SAN network, often a pair of network switches is used to improve availability. At times, one of these redundant switches needs to be replaced. Here, a redundant network switch is defined as one of a pair of switches, that have one or more inter-switch links, and where each which has one or more direct connection to both SC's, and all iSCSI hosts in the SAN. The goal is to perform the switch replacement without any SC failovers or reboots, and with little disruption to SAN data traffic. This is possible only when there is at least one functional network path from every host to both SC's, through the remaining switch, and only if the correct sequence is followed.

To remove a redundant network switch:

1) Log into the current primary SC GUI, and determine which is the cluster primary and which is the secondary.

2) Go to the SC GUI Sessions page and confirm that there the expected sessions from all of the iSCSI SAN hosts.

3) Go to the SC GUI Home and "System -> Diagnostics" pages and confirm that all primary and secondary monitors are "Healthy".

4) Locate the switch that is to be removed.

5) Trace the network cables from the switch to be removed back to the *Secondary* SC, and disconnect these cables from the switch.

6) On the SC GUI Home and "System -> Diagnostics" pages, the Secondary "Data Port Link" monitor should go Orange, and be status "Critical", but not be "Failed". The Primary should still be healthy. The Secondary should not reboot.

7) If there were any Secondary assigned iSCSI sessions, they will be temporarily moved to the Primary. This can be confirmed on the GUI "Sessions" page.

8) Now, trace the network cables from the switch being removed back to the *Primary* SC, and disconnect these cables from the switch.

9) On the SC GUI Home and "System -> Diagnostics" pages, the Primary and Secondary "Data Port Link" monitors should both be Orange, and be status "Critical", but not be "Failed". Neither SC should reboot.

10) Trace the network cables from the switch being removed back to each host, and disconnect them from the switch.

11) On the SC GUI "Sessions" page, confirm that all of the hosts still have at least one session to their volumes. Because all of the hosts are wired to both switches, and are
configured for either MPIO, network bonding, or Multipath, SAN data traffic should still occur on the remaining paths.

12) Disconnect the inter-switch network links to the other switch, and physically remove the switch.

To replace a redundant network switch:

1) Physically install the new switch, but do not connect any of the Ethernet cables yet. Power it on.

2) Connect to the switch's management interface and confirm its status and configuration settings.

3) Reconnect the inter-switch network links to the other switch, and confirm the inter-

switch networking is functional.

4) Reconnect the network cables from the *Primary* SC to the new switch.

5) On the SC GUI Home and "System -> Diagnostics" pages, the Primary "Data Port Link" monitor should Green, and the status now be "Healthy". There should not be a failover.

6) Reconnect the network cables from the *Secondary* SC to the new switch.

7) On the SC GUI Home and "System -> Diagnostics" pages, both the Primary and Secondary "Data Port Link" monitor should go Green, and the status now be "Healthy".

8) Reconnect the iSCSI SAN host network cables to the new switch.

9) On the SC GUI "Sessions" page, confirm that all of the hosts establish their redundant iSCSI sessions for the newly restored network paths. In some cases, it may require host
data I/O, or an iSCSI Initiator UI action to trigger this.

Notes:

1) In a cluster, the Secondary SC will not reboot on a "Critical" network fault -- only a "Failed" one. The distinction being that "Failed" is defined as all network interfaces being down.

2) In a cluster, upon a Primary SC network fault detection, a failover to the Secondary SC will not occur if the same network fault already exists on the Secondary.

3) When a network fault clears, if it clears first on the Primary SC, then on the Secondary SC, a failover will not occur. In the reverse order, a failover is likely to occur.

4) The network fault detection period is 5 to 10 seconds. If a cable can be moved or replaced within that short period of time, no SC failovers, reboots, or disruption to SAN data
traffic would occur. However, some network equipment can have long delays before establishing the link, so the full time period may not always be available.

5) Because there is a single SC management network port, a secondary reboot and/or a failover is unavoidable when changing the LAN network switch. The same is true when there are is only a single SAN network switch.

6) When SC GUI access is unavailable via the LAN network, often the GUI can reached through the SC SAN network IP address instead. This can help when resolving the issues
with the LAN network, or working a LAN network maintenance action. Note that you must use the SC SAN IP addresses and not the SC Cluster IP address.

This page is intentionally left blank.

Appendix 9

System Event Messages

The Storage Concentrator records system event messages in the GUI Reports Log page, and optionally can send them as notification events to an email address and/or SNMP trap.

The system events can be categorized as follows according to the severity levels:

Error – Events that require immediate attention. These are also referred to as 'Critical' and 'Serious'.

Warning – Events indicating an abnormal situation or supplemental events relating to an 'Error' event.

Informational – Events providing normal status or progress information.

Note that only 'Error' and 'Warning' events are subject to notifications. All events are stored in the GUI Log.

The following table lists the system event messages that the Storage Concentrator logs.

Many of the events are parameterized, with parameters represented by '{0}', '{1}', etc.

Some events can be reported at different severity levels based on context. The highest level used is shown.

Level	Event Message	Notes
Error	{0} at {1} has been booted into primary status. All mirrors/snaps are automatically taken offline. Both images of all NAS volumes are marked to need rebuilding, and client access is temporarily disabled. When the secondary SC comes up these mirrors/snaps will be placed back on line. Any failed images will be detached for later recovery if needed. If the secondary does not recover, the mirrors/snaps may be placed on line manually, but only do this if you are sure all images are valid (i.e. No failed images occurred while {0} was down).Note: It is best to bring up the second SC and let the system attempt to determine the correct status.	
Error	{0} at ipaddress {1} appears to be having startup problems. Please reboot and if problem continues contact Customer Support.	
Error	{0} to SC {1} failed with return code {2} for {3}. SenseData = {4}.	
Error	{0}: IO failed due to failure allocate data blocks. Thin volume "{0}" is set not available for iSCSI connections. Thin pool "{1}" can't be expanded to handle more data blocks because total size of allocated pools reached system limit. Volume "{0}" can be used for reading only. To do this, change volume "{0}" access attributes and activate volume by using "Thin Pools" management GUI page. As an alternative solution consider deletion or moving out some volumes from the pool "{1}" or deletion some thin pools completely. Then expand thin pool "{1}" if it possible and activate volume by using "Thin Pools" management GUI page.	
Error	{0}: IO failed due to failure allocate data blocks. Thin volume "{0}" is set not available for iSCSI connections. To activate this volume, expand thin pool "{1}" first and then activate volume by using "Thin Pools" management GUI page.	
Error	{0}: IO failed due to failure allocate metadata blocks. Thin volume "{0}" is set not available for iSCSI connections. To activate this volume, expand thin pool "{1}" by providing at least {2}GB of additional space. Then activate the volume by using "Thin Pools" management GUI page.	
Error	{0}: IO failed due to failure allocate metadata blocks. Thin volume "{0}" is set not available for iSCSI connections. Volume "{0}" can be used for reading only. To do this, change volume "{0}" access attributes and activate volume by using "Thin Pools" management GUI page. As an alternative solution consider deletion or moving out some volumes from the pool "{1}". Then activate the volume	

Level	Event Message	Notes
	by using "Thin Pools" management GUI page.	
Error	{0}: IO failed due to failure allocate metadata blocks. Thin volume "{0}" reached its maximum capacity to handle metadata blocks. Volume is set not available for iSCSI connections. Volume can be used for reading only. To do this, change volume "{0}" access attributes and activate volume by using "Thin Pools" management GUI page.	
Error	{0}: IO failed due to failure to allocate data blocks. Dedup volume "{0}" is set not available for iSCSI connections. To use this volume, first expand dedup pool "{1}", then re-activate the volume by using "Deduplicated Volumes" GUI page.	
Error	{0}: IO failed due to failure to allocate data blocks. Dedup volume "{0}" is set not available for iSCSI connections. Dedup pool "{1}" can't be expanded to handle more data blocks because total size of allocated pools has reached the system limit. Access to the volume is blocked until it is re-activated by using the "Deduplicated Volumes" GUI page. The volume can be used read-only by changing the volume access attributes. Alternatively, consider moving some, or all of pool volumes data out of the pool, or by completely deleting unneeded volumes from the pool. Note that a client utility that zeros unused file-system space may need to be run to release space held by deleted files.	
Error	{0}: IO failed due to failure to allocate data blocks. Thin volume "{0}" is set not available for iSCSI connections. Thin pool "{1}" can't be expanded to handle more data blocks because total size of allocated pools reached system limit. This volume can be used for reading only. To do this, change the volume access attributes, and activate the volume by using "Thin Volumes" management GUI page. As an alternative solution consider the deletion or moving out some volumes from the pool, or delete other thin pools completely. Then expansion of the thin pool should be possible and activate the volume by using "Thin Volumes" management GUI page.	
Error	{0}: IO failed due to failure to allocate data blocks. Thin volume "{0}" is set not available for iSCSI connections. To use this volume, first expand thin pool "{1}", then activate the volume by using "Thin Volumes" management GUI page.	
Error	{0}: IO failed due to failure to allocate metadata blocks. Dedup volume "{0}" reached its maximum capacity to handle metadata blocks. The volume is set not available for iSCSI connections. Access to the volume is blocked until it is re-activated by using the	

Level	Event Message	Notes
	"Deduplicated Volumes" GUI page. The pool must be expanded to prevent it from going offline again on write accesses. The volume can be used read-only by changing the volume access attributes. Alternatively, consider moving some, or all of pool volumes data out of the pool, or by completely deleting unneeded volumes from the pool. Note that a client utility that zeros unused file-system space may need to be run to release space held by deleted files. The volume can be used for reading only. To do this, change the volume access attributes, and activate volume by using "Deduplicated Volumes" GUI page.	
Error	{0}: IO failed due to failure to allocate metadata blocks. Dedup volume "{0}" is set not available for iSCSI connections. To use this volume, first expand dedup pool "{1}", then re-activate the volume by using "Deduplicated Volumes" GUI page.	
Error	{0}: IO failed due to failure to allocate metadata blocks. Dedup volume "{0}" is set not available for iSCSI connections. Access to the volume is blocked until it is re-activated by using the "Deduplicated Volumes" GUI page. The pool must be expanded to prevent it from going offline again on write accesses. The volume can be used read-only by changing the volume access attributes. Alternatively, consider moving some, or all of pool volumes data out of the pool, or by completely deleting unneeded volumes from the pool. Note that a client utility that zeros unused file-system space may need to be run to release space held by deleted files.	
Error	{0}: IO failed due to failure to allocate metadata blocks. Thin volume "{0}" is set not available for iSCSI connections. To activate this volume, expand thin pool "{1}" by providing at least {2}GB of additional space. Then activate the volume by using "Thin Volumes" management GUI page.	
Error	{0}: IO failed due to failure to allocate metadata blocks. Thin volume "{0}" is set not available for iSCSI connections. The volume can be used for reading only. To do this, change the volume access attributes, and activate the volume by using "Thin Volumes" management GUI page. As an alternative solution consider deletion or moving out some volumes from the pool. Then activate the volume by using "Thin Volumes" management GUI page.	
Error	{0}: IO failed due to failure to allocate metadata blocks. Thin volume "{0}" reached its maximum capacity to handle metadata blocks. The volume is set not available for iSCSI connections. The volume can be used for reading only. To do this, change the volume access attributes, and activate the volume by using "Thin	

Level	Event Message	Notes
	Volumes" management GUI page.	
Error	{0}: IO failed due to space full.	
Error	Access to NAS volume "{0}" on Storage Concentrator "{2}" is permanently disabled for CIFS user "{1}" during upgrade of the StoneFusion software. The disabled CIFS user has to be deleted and re-created.	
Error	An image failure/snap_failure occurred on {0} during upgrade. Abort the upgrade, and return this box to original version.	
Error	Attention System Administrator. SC {0}: encrypted image "{1}" was promoted. The image uses the same password as the volume "{2}". Please update records that can be used to restore or to rebuild the USB encryption key in future.	
Error	Attention System Administrator. SC {0}: image "{1}" was detached from the encrypted volume "{2}". The image uses the same password as the volume "{2}". Please update records that can be used to restore or to rebuild the USB encryption key in future.	
Error	Attention System Administrator. SC {0}: name of the encrypted volume was changed. Old name "{1}", new name "{2}". Please update records that can be used to restore or to rebuild the USB encryption key in future.	
Error	Can not use a raid resource as a passthru.	
Error	CIFS Active Directory Server used for NAS user authentication is not functioning: {0}	
Error	CIFS User "{0}" is disabled on Storage Concentrator "{1}". The disabled CIFS user has to be deleted and re-created again.	
Error	Copy of volume:{0} to volume:{1} Failed: Error:{2} Sense:{3}	
Error	Could not add the SC:{0}. Restart SC with IP address {1}.	
Error	Could not create new Resource {0}.	
Error	Could not ping host:{0} IP Address:{1} from SC:{2}. Please	

Level	Event Message	Notes
	verify iSCSI network and settings.	
Error	Could not ping SC:{0} IP Address:{1}. Please verify iSCSI network and settings.	
Error	Could not properly reapply to {0}.	SC had a problem applying to be a member of a cluster.
Error	Could not provision {0} on SC {1} from {2}. errorcode={3}.	
Error	Could not provision image:{0} for mirror:{1} on SC:{2}.	
Error	Could not provision node {0} on SC {1}. errorcode={2}.	
Error	Could not read label from {0} for resource {1} lun {2}.	
Error	Could not successfully provision Target Portal: {0} on SC:{1}.	
Error	Data allocation for deduplicated volume "{0}" has reached critical threshold. Consider expanding this deduplicated volume, or transferring it into a regular one.	
Error	Data allocation for thin volume "{0}" hit critical threshold {2}% of total volume size. Consider expanding this thin volume, or transferring it into a regular one.	
Error	Data allocation from deduplication pool "{0}" has reached critical threshold. Consider deleting some deduplicated volumes, transferring some deduplicated volumes into regular ones, changing access level for all deduplicated volumes to read only or expanding the deduplication pool.	
Error	Database has been restored. All mirrors have been be placed offline. Synchronization information for all asynchronous images has been deleted, and async replication is disabled. All regular Snapshots are deleted. Replication snapshots are preserved as reserved. FailOver configuration (if any) was deleted. SCSI Persistent Reservation meta-data has been cleared. Database Auto Save is disabled to protect the database backup that was restored.	

Level	Event Message	Notes
Error	Deduplicated volume:{0} could not be created. Problem:{1}	
Error	Deduplicated volume:{0} deleted.	
Error	Deduplication for the pool "{0}" has ceased due to problems with the deduplication service on Storage Concentrator "{1}". IO will continue without deduplication.	
Error	Deduplication pool "{0}" failed to restore a deduplicated volume not present in the system configuration. Space allocated by this volume can be released after the pool is deleted.	
Error	Deduplication pool "{0}" failed to restore a deduplicated volume not present in the system configuration because the pool is busy with deduplication space reclamation. Try to restore the volume after the reclamation is finished. For now space allocated by this volume is still help by the pool.	
Error	Deduplication pool "{0}" failed to restore deduplicated volume not present in the system configuration due to insufficient system memory. Space allocated by this volume can be released after the pool is deleted.	
Error	Deduplication pool "{0}" has a deduplicated volume not present in the system configuration. Was not able restore this volume because the maximum number of volumes have already been created. Delete unused volumes and try to restore deduplicated volume again. Space allocated by this volume can be released after the restored volume is deleted.	
Error	Deleted snaps from {0} as it marked as failed according to {1}.	
Error	Detected problem with iSCSI Host IP Address already in use, or the SAN network port is down. Can not activate {0}.	
Error	Detected problem with iSCSI Host IP Address already in use. Can not activate {0}.	
Error	Disabled asynchronous image {0} for {1} because the image snapshots are invalid.	
Error	Disabled remote replication for {0} because system is out of free snapspace.	

Level	Event Message	Notes
Error	Discovered a resource not usable by SC:{3} due to block size "{0}". Manufacturer={1} SerialNumber={2}.	
Error	Email Notification could not connect to address {0} for email user {1}. Error = {2}.	
Error	Email Notification to address {0} failed to email id {1}. Error = {2}.	
Error	Failed to login to target:{0} local initiator:{1} from SC:{2} sense data={3}	
Error	Failed to logout initiator:{1} from target: {2} on SC:{0}.	
Error	FailOver occurred: Storage Concentrator "{0}" at {1} is now Primary.	
Error	Failure to provision deduplicated pool '{0}' on SC {1}. The pool's metadata are using format un-supported by current version. Delete pool if SC has to continue to use the current software version. Save DB configuration and re-install software version that was used to create pool if the pool's data has to be preserved.	
Error	Feature License for "{0}" on SC "{1}" is invalid. This can occur on system hardware replacement, and on changes to virtual machines. A temporary evaluation license is being established. Please contact customer support for a replacement license key.	
Error	Hack attempt was made from ipaddress {0} by '{1}' user.	
Error	Have failure to expand deduplicated volume "{0}". Original volume configuration is restored. The volume is set offline. Use General Configuration Volume Management screen to put the volume online.	
Error	Have failure to provision cache device "{0}" for resource "{1}". Error code = {2}.	
Error	Have failure to remove NAS node "{0}" from the Scaled Out configuration. The failed system has to be restored to factory default configuration. Use console to do this. Don't restart system in the current configuration. It may corrupt nodes that are still present in the Scale Out configuration.	

Level	Event Message	Notes
Error	Host "{0}" from ipaddress {1} is using invalid target ipaddress {2} but should be using {3}.	
Error	Host {0} from {1} using invalid target ipaddress.	
Error	Host:{0} from ip:{1} is having problems logging into target:{2}. This target is set to be load balanced on the Secondary SC. To aid in recovery, the volume for the target has been assigned to the Primary SC. If the host can handle redirects you may set it back to the Secondary SC thru the Load Balancing page under System->Admin->FailOver->Load Balancing.	
Error	Host:{0} from ip:{1} is having problems logging into target:{2}. Please check the status of the volume and Access Control List(ACLS).	
Error	Host:{0} from ip:{1} is having problems logging into target:{2}. This target is set to be load balanced on the Secondary SC. To aid in recovery, the volume for the target has been assigned to the Primary SC. If the host can handle redirects you may set it back to the Secondary SC thru the Load Balancing page under System->Admin->FailOver->Load Balancing.	
Error	Host:{0} from ip:{1} is having problems logging into target:{2}. Please check the status of the volume and Access Control List(ACLS).	
Error	Host:{0} from ip:{1} is having problems logging into target:{2}. This target is set to be load balanced on the Secondary SC. To aid in recovery, the volume for the target has been assigned to the Primary SC. If the host can handle redirects you may set it back to the Secondary SC thru the Load Balancing page under System->Admin->FailOver->Load Balancing.	
Error	Host:{0} from ip:{1} is having problems logging into target:{2}. Please check the status of the volume and Access Control List(ACLS).	
Error	Host:{0} from ip:{1} is having problems logging into target:{2}. This target is set to be load balanced on the Secondary SC. To aid in recovery, the volume for the target has been assigned to the Primary SC. If the host can handle redirects you may set it back to the Secondary SC thru the Load Balancing page under System->Admin->FailOver->Load Balancing.	

Level	Event Message	Notes
Error	Host:{0} from ip:{1} is having problems logging into target:{2}. Please check the status of the volume and Access Control List(ACLS).	
Error	Host:{0} from ip:{1} is having problems logging into target:{2}. This target is set to be load balanced on the Secondary SC. To aid in recovery, the volume for the target has been assigned to the Primary SC. If the host can handle redirects you may set it back to the Secondary SC thru the Load Balancing page under System->Admin->FailOver->Load Balancing.	
Error	Host:{0} from ip:{1} is having problems logging into target:{2}. Please check the status of the volume and Access Control List(ACLS).	
Error	Host:{0} from ip:{1} is having problems logging into target:{2}. This target is set to be load balanced on the Secondary SC. To aid in recovery, the volume for the target has been assigned to the Primary SC. If the host can handle redirects you may set it back to the Secondary SC thru the Load Balancing page under System->Admin->FailOver->Load Balancing.	
Error	Host:{0} from ip:{1} is having problems logging into target:{2}. Please check the status of the volume and Access Control List(ACLS).	
Error	Host:{0} from ip:{1} is having problems logging into target:{2}. This target is set to be load balanced on the Secondary SC. To aid in recovery, the volume for the target has been assigned to the Primary SC. If the host can handle redirects you may set it back to the Secondary SC thru the Load Balancing page under System->Admin->FailOver->Load Balancing.	
Error	Host:{0} from ip:{1} is having problems logging into target:{2}. Please check the status of the volume and Access Control List(ACLS).	
Error	Host:{0} from ip:{1} is having problems logging into target:{2}. This target is set to be load balanced on the Secondary SC. To aid in recovery, the volume for the target has been assigned to the Primary SC. If the host can handle redirects you may set it back to the Secondary SC thru the Load Balancing page under System->Admin->FailOver->Load Balancing.	

Level	Event Message	Notes
Error	Host:{0} from ip:{1} is having problems logging into target:{2}. Please check the status of the volume and Access Control List(ACLS).	
Error	Image {0} has Failed. scsi_status={1} sense_data: {2} CDB: {3}. Please make sure that rebuild starts and completes successfully.	Synchro nous mirror image.
Error	Image: {0} failed and the secondary could not be communicated with appropriately.	Synchro nous mirror image.
Error	Invalid target reference.	
Error	IO failed due to failure to allocate data blocks. Dedup volume "{0}" is set not available for iSCSI connections. Dedup pool "{1}" can't be expanded to handle more data blocks because total size of allocated pools has reached the system limit. Access to the volume is blocked until it is re-activated by using the "Deduplicated Volumes" GUI page. The volume can be used read-only by changing the volume access attributes. Alternatively, consider moving some, or all of pool volumes data out of the pool, or by completely deleting unneeded volumes from the pool. Note that a client utility that zeros unused file-system space may need to be run to release space held by deleted files.	
Error	IO traffic has to bypass flash cache "{0}" and is going directly to cache backing device because the caching device "{1}" is in a failed state. The flash cache device is unprovisioned.	
Error	IO traffic thru the Writeback flash cache "{0}" and thru the cache backing device is stopped because the caching device "{1}" is in a failed state. The flash cache device is unprovisioned. It is possible that some recently written data to the cache will be lost if the caching device is not put online.	
Error	iSNS deregistration failed. Status is {0}.	
Error	iSNS registration failed. Status is {0}. Verify ipaddress in iSNS setup screen and review logs.	
Error	License "NAS Volumes" can't be enabled on clustered or cluster of	

Level	Event Message	Notes
	one systems. To do this, delete cluster, enable "NAS Volumes" license and recreate the cluster.	
Error	Licenses for the Features "{0}" are not enabled or enabled in different degree on both cluster nodes. Feature operation is inhibited until license keys are installed for both clustered SC's.	
Error	Metadata use for deduplicated volume "{0}" has reached critical threshold. Consider transferring this deduplicated volume into a regular one.	
Error	Metadata use for deduplicated volume "{0}" has reached critical threshold. Consider deleting some deduplicated volumes or transferring some deduplicated volumes into regular ones.	
Error	Mirror {0} is Critical.	
Error	Mirror {0} is Failed.	
Error	Monitor "{0}" failed on SC "{1}".	
Error	Name not unique.	
Error	Not all NAS volumes are mirrored.	
Error	Not enough space on resource.	
Error	Not enough system memory to expand volume "{0}" on SC "{1}".	
Error	Not enough system memory to provision the pool "{0}" on SC "{1}".	
Error	Not enough system memory to provision volume "{0}" on SC "{1}".	
Error	Problem provision encrypted volume {0}. Volume is "Offline". Fix problem with USB Encryption Keys Disk and make volume "Online" manually.	
Error	Problem provisioning volume {0} on SC {1}. errorcode={2}.	
Error	Problem setting up cluster: Concentrator "{0}" error code is {1}	

Level	Event Message	Notes
Error	Problem setting up cluster: return result code is {0}	
Error	Problem to provision encrypted volume {0} on SC {1}. Check USB Encryption Keys Disk.	
Error	Problem with Cluster Manager from SC "{1}": code={0}.	
Error	Process of segment replacement for NAS volumes "{0}" was terminated. Original segment assignment is restored.	
Error	RAID {0}: Event: {1}	Log event from a monitored RAID.
Error	Repair for asynchronous image "{0}" failed. Synchronization information for image "{0}" will be deleted automatically.	
Error	Repair for asynchronous image "{0}" failed. Synchronization information for image has been deleted. Restart remote replication manually by synchronizing all data blocks of the volume.	
Error	Resource:{0} has the same Extended Unit Identifier/Serial Number as another resource. No passthru volumes can be created in this situation.	
Error	Resource:{0}. No support for Flash Cache functionality in StoneFly clusters.	
Error	Rollback of {0} from {1} failed. status: {2} sense: {3} CDB: {4}	
Error	Rollback of {0} from {1} interrupted because system is out of free snapspace. Volume '{0}' lost data integrity and is unavailable as an iSCSI target. If pre-rollback snapshot was taken, rollback volume to this snapshot to restore the last state of the volume or consider deleting of the volume.	
Error	SC "{0}" has {1} route set up and SC "{2}" has {3} routes set up. In most cases, the network configurations of the Active and Standby boxes should be similar.	

Level	Event Message	Notes
Error	SC "{0}" has different multipath network configurations. SC "{0}" has networks "{1}", where SC "{2}" has networks "{3}". The multipath network configurations of the Active and Standby boxes should be similar.	
Error	SC "{0}" is available.	
Error	SC "{0}" is unavailable.	
Error	SC "{0}" Listening Port {1} does not match SC "{2}" Listening Port {3}	
Error	SC "{2}" Sensor "{0}" is at "{1}". Check Diagnostics.	
Error	SC {0} is waiting to handle sessions but the Cluster IPAddress is in use by another concentrator. The status of some resources or volumes do not appear to be OK for {1}, for example [{2}]. Please verify before continuing with upgrade.	
Error	SC {0} is waiting to handle sessions but the Cluster IPAddress is in use by another concentrator. If the status of all resources and volumes are OK, then start upgrade of other Concentrators.	
Error	SC {0}: access to USB port is taken by another StoneFly service. Please repeat request later.	
Error	SC {0}: failure to access secondary SC.	
Error	SC {0}: failure to assign USB device to encryption service.	
Error	SC {0}: failure to copy encryption information to USB Disk.	
Error	SC {0}: failure to create digest for encryption password. Please contact customer support to resolve this problem.	
Error	SC {0}: failure to delete volume encryption information from USB Disk.	
Error	SC {0}: failure to format USB Disk.	
Error	SC {0}: have Encryption Keys Disk that can be used to store	

Level	Event Message	Notes
	volume encryption information.	
Error	SC {0}: have Encryption Keys Disk that is dedicated to another StoneFly service. Please insert another disk.	
Error	SC {0}: have Encryption Keys Disk with encryption information for invalid volumes. Please use appropriate disk.	
Error	SC {0}: have Encryption Keys Disk with encryption information for invalid volumes. It's insecure to keep disk in the USB port permanently.	
Error	SC {0}: have Encryption Keys Disk with invalid volume encryption information. Please use valid disk.	
Error	SC {0}: have Encryption Keys Disk with valid volume encryption information.	
Error	SC {0}: have Encryption Keys Disk with valid volume encryption information. It's insecure to keep disk in the USB port permanently.	
Error	SC {0}: have Encryption Keys Disk without any volume encryption information. Please use valid disk.	
Error	SC {0}: have only non-DOS formatted USB Disks. Please use DOS formatted disks.	
Error	SC {0}: have valid Encryption Keys Disks on different clustered Storage Concentrators. It's insecure to keep disk in the USB port permanently.	
Error	SC {0}: have valid USB Encryption Keys Disks on different clustered Storage Concentrators. Please remove USB Encryption Keys Disk from the Secondary Storage Concentrator.	
Error	SC {0}: invalid argument.	
Error	SC {0}: invalid command.	
Error	SC {0}: more than one Encryption Keys Disk with volume encryption information are detected. Please use no more than one encryption disk.	

Level	Event Message	Notes
Error	SC {0}: more than one Encryption Keys Disk with volume encryption information are detected. It's insecure to keep disk in the USB port permanently.	
Error	SC {0}: no Encryption Keys Disks are found. Please insert disk with encryption information.	
Error	SC {0}: unknown problem with database.	
Error	SC {0}: unknown problem with USB device.	
Error	SC {0}: UPS: {1}	Event notification from a monitored UPS.
Error	SC {0}: USB Disk lost some of the encryption information but can be repaired.	
Error	SC {0}: USB Disk lost some of the encryption information but can be repaired. It's insecure to keep disk in the USB port permanently.	
Error	SC {0}: USB encryption key was rebuilt.	
Error	SC {0}: USB encryption key was repaired.	
Error	SC {0}: volume encryption password does not match password provided during volume creation. Please contact customer support to resolve this problem.	
Error	SC Secondary cluster node "{0}" at "{1}" was deleted from the cluster on the Primary, but it could not be accessed at the time. The old Secondary will still have an old copy of the SC database, and still think that it is in the cluster. This will interfere with the Primary should this Secondary ever be running again. Running it could cause damaging uncoordinated access to shared storage resources. The old Secondary should be removed from the network, and have its database cleared. This can be done by, on the Secondary, using either the SC System Console 'Default DB' function, or by reinstalling the SC software with the 'default.db' file	

Level	Event Message	Notes
	on the USB key disk in the root directory.	
Error	SC:{0} has been unable to synchronize the database with the cluster secondary so the secondary database does not reflect the current configuration of the cluster. Please correct the condition of the secondary. If the secondary cluster node cannot be made healthy, you may choose to remove it from the cluster. Do not allow the secondary to go active as primary without it first being a healthy secondary so that its DB copy can be updated to be current.	
Error	SC:{0} has monitors in the unknown/failed state. Your Storage Concentrator Failover pair is running without Redundancy. Please check the Diagnostics screen and verify that all Monitors are healthy. If any Monitors are not healthy, you may need to take corrective action.	
Error	SESSION: Host "{0}", from ipaddress {2}, failed to login in to thin target "{1}". Thin volume "{3}" is out of available data or metadata space.	
Error	Snapshot Fragmentation Threshold reached, highly fragmented snapshots will be deleted.	
Error	Snapshot: {0} failed and the secondary could not be communicated with appropriately.	
Error	Snapshot: {0} failed Snapspace could not be communicated with appropriately. All Snapshots have been deleted.	
Error	Snapshot:{0} {1} and will be deleted. CDB: {2} scsi_sense: {3} scsi_status: {4}.	
Error	Snapshot:{0} could not be created by the local node. Problem:{1}. It's possible that the snapshot is created already by other node from the Scale Out configuration. Check NAS Snap Management Detail screen to verify it.	
Error	Snapshot:{0} could not be created. Problem:{1}.	
Error	Snapshot:{0} Snapspace use threshold hit at {1}%.	
Error	Snapshots for {0} are invalid and will be deleted. CDB: {1} scsi_sense: {2} scsi_status: {3}.	

Level	Event Message	Notes
Error	Storage Concentrator "{0}" at {1} is now in the Primary(upgrade) state.	
Error	Storage Concentrator "{0}" from {1} failed to join the cluster: {2}.	
Error	System can't maintain IO traffic thru flash cache "{0}" and thru the cache backing device because the backing device "{1}" is in a failed state. The flash cache device is unprovisioned.	
Error	System failed to create the "nas-metadata" volume automatically. NFS and CIFS Exports are not available for any NAS volume until "nas-metadata" volume is recreated successfully. The system will attempt to create the "nas-metadata" volume on each attempt to create a new NAS volume, add an image to existing NAS volume, or add a new share. If the system has no shared "Managed" resources with space, "NAS Managed" resources will be used instead once available on both SC's. Use the resource summary screen to "Manage" or "NAS Manage" a storage resource on the new secondary.	
Error	System Metadata volume "system-metadata" failed while executing operation for volume "{0}" on SC:{1}. The volume operations that are based on this metadata will be suspended. Please check status of storage resources used to by the system metadata volume. Take corrective action if resources are not in a valid operational state.	
Error	SYSTEM: "{0}" at ipaddress {1} has been restarted.	
Error	SYSTEM: "{0}", total memory has fallen below the {1}GB minimum required to support NAS volumes functionality. Please check for failed system memory.	
Error	SYSTEM: "{0}", total memory has fallen below the {1}GB minimum required to support {2} volumes functionality. Please check for failed system memory.	
Error	SYSTEM: Could not provision Fibre Channel Target resource {0} on concentrator {1}. Provision error_code={2}.	
Error	SYSTEM: Could not provision NAS segment resource {0} on concentrator {1}. Provision error_code={2}.	

Level	Event Message	Notes
Error	SYSTEM: Could not provision NAS segments on concentrator {0}. Provision error_code={1}.	
Error	SYSTEM: Could not provision NAS segments on concentrator {0}. Try to resolve this by performing a NAS segment rediscovery on the NAS Segment Summary GUI page. Provision error_code={1}.	
Error	SYSTEM: Could not provision resource {0} on concentrator {1}. Provision error_code={2}.	
Error	SYSTEM: Could not retrieve serial number from {0}.	
Error	SYSTEM: Database Failure {0}.	
Error	SYSTEM: Failed to add route:{0} gateway:{1} network:{2} netmask:{3} on firefly {4}.	
Error	SYSTEM: update routes failed for network={0} netmask={1} gateway={2} device={3}.	
Error	The {0} network MTU setting has been reset to the default on SC "{1}" at {2}; was "{3}", now is "{4}".	
Error	The {0} network multipath port assignments have been changed on SC "{1}" at {2}; was "{3}", now is "{4}".	
Error	The {0} network port assignments have been changed on SC "{1}" at {2}; was "{3}", now is "{4}".	
Error	The {0} network port assignments have been reset to default on SC "{1}" at {2}; was "{3}", now is "{4}".	
Error	The Fibre Channel port "{0}" in SC "{1}" is configured, but has gone missing. FC targets through this port will be inaccessible. Resolve the HW problem, and/or remap the targets to another FC port.	
Error	The Maximum number of segments ({0}) in volume {1} have already been created.	
Error	The Maximum number of volumes have already been created.	

Level	Event Message	Notes
Error	The primary and secondary meta data backups do not match for resource {0}. Please call customer support for assistance.	The SC resource reserved space DB redundant copies are not equal.
Error	The resource "{0}" has changed in size from {1} GB to {2} GB. Old block size is {3}. New block size is {4}. It is strongly recommended to delete old "Non-Active" copy of the resource from the resource summary page.	
Error	The resource {0} has decreased in size. Shrinking of resources is not allowed.	
Error	The resource {0} has increased in size from {1} GB to {2} GB. It is strongly recommended that a reboot be done if a reboot has not been done since the resource increased in size.	
Error	The resource {0} used by the flash cache {1} has changed in size. Changing sizes of flash cache resources is not allowed. The flash cache device has to be deleted.	
Error	The resource "{0}" has increased in size from {1} GB to {2} GB.	
Error	The secondary was not synched appropriately. Please verify secondary status.	
Error	The secondary was not synched appropriately. Please verify secondary status.	The SC DB was not synchronized with the cluster secondary.
Error	The Storage Concentrator {0} CIFS Volume Service is down.	
Error	The Storage Concentrator {0} Management Service is down.	

Level	Event Message	Notes
Error	The Storage Concentrator {0} NAS Volume Service is down.	
Error	The Storage Concentrator {0} Volume Service is down.	
Error	There was a problem updating configuration at ipaddress {0}.	
Error	Thin pool "{0}" failed to restore a thin volume not present in the system configuration because the pool is busy with thin space reclamation. Try to restore the volume after the reclamation is finished. For now space allocated by this volume is still help by the pool.	
Error	Thin pool "{0}" failed to restore thin volume not presented in the system configuration. Space allocated by this volume can be released after the pool is deleted.	
Error	Thin pool "{0}" has a thin volume not present in the system configuration. Was not able restore this volume because the maximum number of volumes have already been created. Delete unused volumes and try to restore thin volume again. Space allocated by this volume can be released after the restored volume is deleted.	
Error	Thin volume:{0} could not be created. Problem:{1}	
Error	This is a test Message arg0 {0}, arg1 {1}, arg2 {2}, arg3 {3}, arg4 {4}.	
Error	To properly expand volume {0} it will need to be toggled offline/online.	
Error	Total memory for system has been increased. To provide more metadata space for existing deduplication pools consider deduplication pools expansion. For new deduplication pools the additional memory will be counted automatically.	
Error	Usage of Deduplication Metadata hit System Threshold {1}%, highly allocated deduplicated volumes should be deleted or transferred to regular volumes. Consider to do this for deduplicated volume:{0}.	
Error	Usage of Snap and Thin Metadata reached Threshold, highly fragmented snapshots will be deleted.	

Level	Event Message	Notes
Error	Usage of Snap, Thin, and Dedup Metadata reached Threshold, highly fragmented snapshots will be deleted.	
Error	Version of SC database is not correct. Have version {0}, was expected {1}. Try a reboot. If problem persists please call customer support.	
Error	Volume "{0}" could not be provisioned because cache device "{1}" used by the volume is in a failed state.	
Error	Volume "{0}" could not be updated because cache device "{1}" used by the volume is in a failed state.	
Error	Volume '{0}' has been unprovisioned from Storage Concentrator '{1}'.	
Error	VOLUME: Error zeroing volume "{0}".	
Error	Your "{0}" Feature License evaluation period on SC "{1}" has expired! Feature operation is inhibited. Please contact customer support for a license for this feature.	
Error	Your StoneFusion Base OS evaluation period has expired! SC operation is inhibited. Please contact customer support for a license for this product.	
Info	{0} is ok after retry.	
Info	{0} snapshot:{1} deleted.	
Info	{0} snapshot:{1} has to be deleted by user: "{2}".	
Info	{0} snapshot{1} deleted by user: "{2}".	
Info	{0} state change being ignored for failover as it recently had a changed state.	
Info	A resource from {0} {1} discovered by secondary SC:{2} has the same Extended Unit Identifier/Serial Number {3}/{4} as another discovered resource. Can not associate this resource to the secondary in this situation without first being managed by the primary.	

Level	Event Message	Notes
Info	Asynchronous image "{0}" has been repaired successfully, replication has resumed.	
Info	Asynchronous image {0} has been deleted.	
Info	Asynchronous image {0} inserted into mirror {1}. Initialization mode "{2}".	
Info	Automatic repair for asynchronous image "{0}" has been initiated.	
Info	Cache device "{0}" for resource "{1}" is provisioned.	
Info	Cannot perform {0} for {1} as Rollback in use. Rollback stage: {2}	
Info	Check for USB encryption Keys Disk could not be started. Problem:{0}	
Info	Communication to remote storage for asynchronous image "{0}" for volume "{1}" on SC:{2} is restored.	
Info	Copy of volume:{0} to volume:{1} completed.	
Info	Copy of volume:{0} to volume:{1} is {2} percent complete.	
Info	Copy of volume:{0} to volume:{1} was canceled before completion.	
Info	Could not create host for {0} as host name is not unique.	
Info	Could not create host for {0} as host name not valid.	
Info	Could not create host for {0} as iSCSI host name not valid.	
Info	Could not create host for {0} as iSCSI name not unique.	
Info	Could not create host for {0} as no alias.	
Info	Could not create host for {0} as the licensed number of iSCSI hosts are already created.	

Level	Event Message	Notes
Info	Could not create host for {0} if ipaddress not valid.	
Info	Could not create host for {0}.	
Info	Could not create new acl for host:{0}. Discovery string would be too big. Try shortening iSCSI Target Names of targets used by this host.	
Info	Could not do a autosave to {0}.	SC DB periodic backup failed.
Info	Could not do a full autosave recover following a resource resize for {0}.	The SC resource reserved space DB backup could not be moved during a resource resize.
Info	Could not retrieve DNS info on SC:{0} from DNS server {1} for {2}.	
Info	Could not retrieve DNS info on SC:{0} from DNS server {1}.	
Info	Database backup to {0} {1} {2}.	
Info	Database has been restored. All mirrors have been placed offline. Synchronization information for all asynchronous images has been deleted, and async replication is disabled. All regular Snapshots are deleted. All NAS volume Snapshots are deleted. Replication snapshots are preserved as reserved. FailOver configuration (if any) was deleted. SCSI Persistent Reservation meta-data has been cleared. Database Auto Save is disabled to protect the database backup that was restored.	
Info	Deduplicated volume "{0}" has been restored in pool "{1}" successfully. Delete this volume if it is no longer needed. Space allocated by this volume can be returned to the pool only after the	

Level	Event Message	Notes
	restored volume is deleted.	
Info	Deduplicated volume:{1} in pool:{0} is created, this makes {2} volumes existing for this pool.	
Info	Discovered a resource not usable by SC:{3} of type "{0}". Manufacturer={1} SerialNumber={2}.	
Info	Duplicate request for security for {0}.	
Info	Failed to Create device for FlashCache '{0}' by Device Mapper VD '{1}' on SC '{2}' sense data={3}	
Info	Failed to Create device for target:{0} lun:{1} local initiator:{2} from SC:{3} sense data={4}	
Info	Failed to discover targets target:{0} local initiator:{1} from SC:{2} sense data={3}	
Info	Failed to query device Manufacture:{0} target:{1} from SC:{2} sense data={3}	
Info	Feature License for "{0}" on SC "{1}" has been "{1}"{2}	
Info	Image {0} has been detached from {1}.	Synchro nous mirror image.
Info	Image {0} has been promoted from {1} and the mirror itself has been cleared.	Synchro nous mirror image.
Info	Image {0} has been promoted from {1}.	Synchro nous mirror image.
Info	Image {0} has been reattached to {1}.	Synchro nous mirror image.

Level	Event Message	Notes
Info	Image {0} inserted into mirror {1}.	
Info	Internally suspended replication for asynchronous image "{1}" from volume "{0}" has been automatically resumed.	
Info	Invalid access used to create security for {0}.	
Info	Invalid Host used to create security for {0}.	
Info	Invalid iSCSI target used to create security for {0}.	
Info	Invalid security used for {0}.	
Info	IO traffic for resource "{0}" is handled by the Flash Cache device "{1}".	
Info	IO traffic for resource "{0}" was redirected to the resource directly to bypass the Flash Cache device "{1}".	
Info	iSNS deregistration ok. Status is {0}.	
Info	iSNS registration succeeded. Status is {0}.	
Info	iSNS update succeeded. Status is {0}.	
Info	License generation login attempt failed from ipaddress "{0}" by user "{1}".	
Info	Local Storage Concentrator "{0}" joined the Scaled Out configuration.	
Info	Local Storage Concentrator "{0}" was removed from the Scaled Out configuration.	
Info	Maximum number of ACLs({1}) already created. Could not create new acl for {0}.	
Info	Maximum number of DedupSpace volumes already created. Could not create the new one for: {0}.	

Level	Event Message	Notes
Info	Maximum number of SnapEnabled volumes already created. Could not SnapEnable: {0}.	
Info	Maximum number of SnapEnabled volumes in the deduplication pool already created. Could not SnapEnable: {0}.	
Info	Maximum number of SnapEnabled volumes in the thin pool already created. Could not SnapEnable: {0}.	
Info	Maximum number of ThinSpace volumes already created. Could not create the new one for: {0}.	
Info	Mirror {0} is OK.	
Info	Monitor "{0}" critical on SC "{1}".	
Info	Monitor "{0}" OK on SC "{1}".	
Info	Monitor "{0}" unknown on SC "{1}".	
Info	NAS node "{0}" was removed from the Scaled Out configuration.	
Info	NAS snapshot:{1} of volume:{0} taken, this makes {2} snapshots existing for this volume.	
Info	NAS Volume "{0}" rebalance status: {1}	
Info	NAS volume "{0}" was expanded from {1} GB to {2} GB.	
Info	NAS volume {0} segments are no longer exported for manual repair.	
Info	NAS volume {0} segments have been exported for manual repair.	
Info	NAS volumes will be reassigned to use segment "{0}" from node "{1}" as a replacement for segment "{2}" from node "{3}".	
Info	Performed autosave following a resource resize for {0}.	The SC resource reserved space DB

Level	Event Message	Notes
		backup moved successfully during a resource resize.
Info	Performed autosave to {0}.	SC DB periodic backup worked.
Info	Problem creating ACL for {0}.	
Info	Process of segment replacement for NAS shares "{0}" was started. Check status of the replacement on GUI screen. When the replacement copied all data to the destination segment, use the status screen to commit the replacement. If shares are replicas, the replacement process starts for segment from the second image automatically. Continue to check status of the second replacement and commit it after the second replacement is done. The shares start to use new segments after the last commit command.	
Info	Process of segment replacement for NAS volume "{0}" was started. Check status of the replacement on GUI screen. When the replacement copied all data to the destination segment, use the status screen to commit the replacement. The source segments are removed after the commit command.	
Info	Process of segment replacement for NAS volumes "{0}" finished successfully.	
Info	Process of segment replacement for NAS volumes "{0}" was terminated. Original segment assignment is restored.	
Info	RAID '{0}' is successfully being monitored.	
Info	RAID '{0}' monitoring has been disabled.	
Info	Rebuild of {0} has been stopped by user.	Synchronous mirror rebuild.

Level	Event Message	Notes
Info	Rebuild of {0} has been stopped to perform internal maintenance. Rebuild will be restarted after the maintenance is finished.	Synchro nous mirror rebuild.
Info	Rebuild of {0} has completed.	Synchro nous mirror rebuild.
Info	Rebuild of NAS volume {0} has completed.	
Info	Rebuilding of {0} on {1} from block {2} with priority {3}.	Synchro nous mirror rebuild.
Info	Rebuilding of NAS volume {0} has started.	
Info	Repair for asynchronous image "{0}" finished. Replication snapshots have valid state now. Try to initiate replication manually. If problem is persistent call customer support.	
Info	Replication:{0} could not be started. Problem:{1}	
Info	Rollback of volume: {0} from snapshot: {1} is {2}% completed.	
Info	SC {0}: SCSI Persistent Reservation meta-data for host {1} access to target {2} has been deleted with an active SCSI PR registration present. If this host is re-allowed access to this target, you must manually clear SCSI Persistent Registration from the host first.	
Info	SC {0}: UPS management was successfully disabled.	
Info	SC:{0} has redundant paths to Resource: {1} and Resource: {2}, but is currently using the same path on both SCs. Improved path redundancy can be achieved by repairing any failed paths, and restarting the SC.	
Info	Schedule for Replication Processing could not be invoked. Problem:{0}	

Level	Event Message	Notes
Info	Scsi command {2} used by volume service on SC:{0} failed for {1}. Scsi status = {3}.	
Info	Scsi command {2} used by volume service on SC:{0} failed for {1}. Sense key = {3}, ASC:ASCQ = {4}.	
Info	Scsi command {2} used by volume service on SC:{0} timed out for {1}.	
Info	SESSION: Host "{0}", from ipaddress {1}, failed to login in to volume "{2}" host not using chap.	
Info	SESSION: Host "{0}", from ipaddress {2}, failed to login in to unavailable target "{1}".	
Info	SESSION: Host "{0}", successful login to volume {2}, from ipaddress {4}, to ipaddress {1}, given {3} access.	
Info	SESSION: Initiator "{0}", from ipaddress {3}, failed to login to target "{1}" using user name "{2}".	
Info	Shared space for NAS volumes "{0}" was expanded from {1} GB to {2} GB.	
Info	Snapshot:{1} of volume:{0} taken, this makes {2} snapshots existing for this volume.	
Info	Snapspace for "{0}" attribute "{1}" set to "{2}".	
Info	Storage Concentrator "{0}" from {1} has deleted from cluster.	
Info	Storage Concentrator "{0}" from {1} has joined cluster.	
Info	Storage Concentrator "{0}" from {1} has successfully joined the cluster. Storage Concentrator "{0}" from {1} has joined cluster.	
Info	Storage Concentrator "{0}" from {1} is joining the cluster.	
Info	Storage Concentrator "{0}" joined the Scaled Out configuration.	
Info	Successfully restarted replication after failure for asynchronous	

Level	Event Message	Notes
	image "{0}" for mirror "{1}" on SC {2}.	
Info	Synchronous image {0} inserted into mirror {1}.	
Info	SYSTEM: "{0}" at ipaddress {1} has started.	
Info	SYSTEM: "{0}" at ipaddress {1} is now running.	
Info	SYSTEM: The Debug utility has been navigated to by user {0}.	
Info	SYSTEM: The Debug utility has been used by user {0}. Parameter {1} is {2}.	
Info	SYSTEM: User {0} has been logged in.	
Info	Test unit ready not valid from {0} for {1} check_cond/sense_data={2}.	
Info	The {0} network multipath port assignments have been changed on SC "{1}" at {2}; was "{3}", now is "{4}".	
Info	The {0} oldest log messages were deleted because the max log limit had been reached.	
Info	There are multiple resources having the same StoneFly label. Label is {0}, Resource: {1} and Resource: {2}.	
Info	There was a problem to access reserved space on resource "{0}". Please check status of the resource.	
Info	Thin volume "{0}" has been restored in pool "{1}" successfully. Delete this volume if it is no longer needed Space allocated by this volume can be returned to the pool only after the restored volume is deleted.	
Info	Thin volume "{0}" was restored in the pool "{1}" successfully. Delete volume if it should not be used anymore or it's contents is outdated. Space allocated by this volume can be returned to the pool only after the restored volume is deleted.	
Info	Thin volume:{0} deleted.	

Level	Event Message	Notes
Info	Thin volume:{1} in pool:{0} is created, this makes {2} volumes existing for this pool.	
Info	Unit does not support inquiry. Device Query Manufacture:{0} target:{1} from SC:{2} sense data={3}.	
Info	update of mirror: {1} to resource: {0} state:{2}	
Info	Volume "{0}" restored management for asynchronous image "{1}". GUI is enabled for this image now.	
Info	Volume "{0}" temporarily suspended remote replication for asynchronous image "{1}" on "{2}". System resumes replication automatically. User may resume replication manually by selecting "suspend" replication for image "{1}", followed by "resume" for the same image.	
Info	Volume "{0}" temporarily suspended remote replication for asynchronous image "{1}" on "{2}". System resumes replication automatically. User may resume replication manually by selecting "resume" for image "{1}".	
Info	Volume '{0}' has been deleted.	
Info	Volume {0} is Offline.	
Info	Volume service failed for {1} on SC:{0}.	
Info	Volume service failed to get list of virtual volumes from SC:{0}.	
Info	Volume service failed to open iSCSI session with SC:{0}.	
Info	Volume: {0} expand was reverted back to {1} GB.	
Info	Volume: {0} expanded from {1} GB to {2} GB.	
Info	Volume: {0} expanded from {1} GB to {2} GB. Consider snapspace expansion for the volume too.	
Info	Your "{0}" Feature License {1} day evaluation period on SC "{2}" has only {2} days remaining. Please contact customer support for	

Level	Event Message	Notes
	a license for this feature.	
Warn	{0} at ipaddress {1} appears to be having startup problems. It is being automatically rebooted.	
Warn	{0} failed to provision snapshot "{1}" on "{2}" for use to replicate volume "{3}". When replication repair is in progress, wait for the repair. If there is no replication repair, or repair failed, it is strongly recommended to delete remote replication for volume "{3}" and recreate it.	
Warn	{0} snapshot:{1} auto deleted to free space.	
Warn	Access to NAS volume "{0}" on Storage Concentrator "{2}" is disabled for CIFS user "{1}" until all nodes in the configuration are functioning properly and are running the same version of the StoneFusion software.	
Warn	Access to the NAS volume "{0}" had been temporarily disabled, but access has now been restored.	
Warn	Access to the NAS volume "{0}" has been temporarily disabled due to the only online image not being in synch. Access will be restored when the other image comes online.	
Warn	Active boot check done on {0}.	
Warn	Asynchronous image "{0}" for volume "{1}" on SC:{2} has no response from remote storage. Possibly have a network problem. Check remote system. Replication will be restarted automatically when network and remote system becomes available again.	
Warn	Asynchronous image "{0}" has assignment miscompare for replication reserved snapshots. Remote replication has been disabled temporarily. Automatic repair will be started when the system is ready. In a SC cluster, the repair will be initiated when all nodes of the cluster reach the "healthy" state. If the secondary node is going to stay down, the user may remove the secondary node from the cluster to allow replication to resume.	
Warn	Asynchronous image "{0}" has invalid replication reserved snapshots. Automatic repair will be started when the system is ready. In a SC cluster, the repair will be initiated when all nodes of the cluster reach the "healthy" state. If the secondary node is going to stay down, the user may remove the secondary node	

Level	Event Message	Notes
	from the cluster to allow replication to resume. User intervention is not required.	
Warn	Could not add target:{0} from target portal:{1} as there are too many targets in this portals in this portal.	
Warn	Could not retrieve time on SC:{0} from NTP server {1}.	
Warn	Data allocation for deduplicated volume "{0}" has reached warning threshold.	
Warn	Data allocation from deduplication pool "{0}" has reached warning threshold.	
Warn	Data allocation from thin pool "{0}" hit critical threshold at {2}%. Consider deleting some thin volumes, transferring some thin volumes into regular ones, changing access level for all thin volumes to read only or expanding the thin pool.	
Warn	Data allocation from thin pool "{0}" hit threshold at {2}%.	
Warn	Deduplication for the pool "{0}" on Storage Concentrator "{1}" is restored. IO will continue with deduplication.	
Warn	Detached image {0} from {1} as it was marked as failed according to {2}. This image may be manually reattached if the mirror is ok and the detached image is no longer needed for testing.	
Warn	Feature License for "{0}" is invalid. This can occur on system hardware replacement. A valid license still exists in the cluster, so the invalid license is repaired and the feature remains licensed.	
Warn	Feature License for "{0}" on SC "{1}" is invalid, and has been re-keyed. This can occur on system hardware replacement, and on changes to virtual machines. The license remains valid, but with a new key.	
Warn	Fibre Channel target ports are not online: {0}	
Warn	Fibre Channel targets need to be mapped to hosts on ports on both SCs in a cluster. The following volumes are currently not mapped correctly: {0}	

Level	Event Message	Notes
Warn	Have internal problem with the System Metadata volume.	
Warn	Host "{0}" from ipaddress {1} is not on the SC data network {2}, nor are there routes to it. Please check your SAN network addressing and SC routes.	
Warn	Metadata use for deduplicated volume "{0}" has reached warning threshold.	
Warn	Metadata use for deduplication pool "{0}" has reached critical threshold. Consider deleting some deduplicated volumes, transferring some deduplicated volumes into regular ones, changing access level for all deduplicated volumes to read only. User can wait when critical threshold will be reported for specific deduplicated volume and then apply the recommended actions to this volume.	
Warn	Metadata use for deduplication pool "{0}" has reached warning threshold.	
Warn	Metadata use for thin pool "{0}" hit critical threshold at {1}%. Consider deleting some thin volumes, transferring some thin volumes into regular ones, changing access level for all thin volumes to read only. User can wait when critical threshold will be reported for specific thin volume and then apply the recommended actions to this volume.	
Warn	Metadata use for thin pool "{0}" hit threshold at {1}%.	
Warn	Metadata use for thin volume "{0}" hit critical threshold at {1}%. Consider transferring this thin volume into a regular one.	
Warn	Metadata use for thin volume "{0}" hit threshold at {1}%.	
Warn	Mirror {0} is Degraded.	
Warn	NAS client sessions(s) are not using SC NAS alias IP addresses, but should be. The following clients have sessions to invalid target IP addresses (Client Name/IP : Target IP): {0}	
Warn	NAS service critical on SC:{0} - {1}	

Level	Event Message	Notes
Warn	NAS service failed on SC:{0} - {1}	
Warn	Please create the System Metadata volume "system-metadata" of 1 GB size. Use "Create New Volume" GUI page.	
Warn	Please expand the System Metadata volume "system-metadata" by 1 GB of local resource space.	
Warn	Problem provisioning NAS volume {0} on SC {1}. errorcode={2}.	
Warn	Quota of used space for NAS volume '{0}' exceeded soft limit for {1} and hard limit for {2} directories. Open NAS volume Directory Quota management GUI screen for details.	
Warn	RAID '{0}' monitoring has failed: RAID is not configured for monitoring.	
Warn	RAID '{0}' monitoring has failed: RAID model is not supported for monitoring.	
Warn	RAID '{0}' monitoring has failed: RAID monitoring is disabled.	
Warn	RAID '{0}' monitoring has failed: Temporary inability to interact with the RAID.	
Warn	RAID '{0}' monitoring has failed: Timeout interacting with the RAID.	
Warn	RAID '{0}' monitoring has failed: Unable to contact the RAID.	
Warn	RAID '{0}' monitoring has failed: Unable to interact with the RAID.	
Warn	RAID '{0}' monitoring has failed: Unable to login to the RAID, check password.	
Warn	RAID '{0}' monitoring has failed: Unexpected internal error.	
Warn	Rebooting "{0}" by user "{1}" from {2}.	
Warn	Rebuild of NAS volume {0} has been stopped due to "{1}". Rebuild will be restarted automatically when possible.	

Level	Event Message	Notes
Warn	Replication failure for asynchronous image "{0}" from volume "{1}" on SC:{2}. Check that local and remote systems are OK. Replication will restart automatically.	
Warn	RESOURCE: Discovery Failed on "{1}" : {0} sense data = {2}	
Warn	SC {0}: SCSI Persistent Reservation meta-data has been deleted. This is expected during DB restore from backup, and some SC SW upgrades. All hosts will be notified. Please confirm normal operation on hosts that use SCSI Persistent Reservations.	
Warn	SC {0}: UPS command '{1}' was/was NOT issued successfully.	
Warn	SC {0}: UPS management service was/was NOT successfully restarted.	
Warn	SC {0}: UPS management was/was NOT successfully enabled with the role '{2}'.	
Warn	SESSION: Host "{0}", from ipaddress {1}, failed to login in to volume "{2}" not allowed in ACLS.	
Warn	Shutting down "{0}" by user "{1}" from {2}.	
Warn	Snapshot "{0}" used to replicate volume "{1}" has been repaired.	
Warn	Snapshot Fragmentation Threshold {1}%, highly fragmented snapshots should be deleted. Consider deleting snapshot:{0}.	
Warn	System failed to create image for volume "system-metadata" automatically.	
Warn	SYSTEM: "{0}" UUID has changed, old value={1}, new value={2}.	
Warn	SYSTEM: "{0}", total memory has been changed, old value={1}KB, new value={2}KB.	
Warn	The CPU in SC "{0}" at ipaddress {1} can support AES-NI HW encryption, but it is not enabled in the BIOS. Enable AES-NI for a significant performance improvement.	

Level	Event Message	Notes
Warn	The metadata mapping for volume "{0}" is damaged. Problem can be fixed by deleting images for volume "{0}" and reverting the volume to a span -- then re-mirror it.	
Warn	There are too many sessions({0}) logged in to the system, there should be max of {1}. If you are using loading, not all sessions may be able to login to the system during an actual failover.	
Warn	Thin data allocation from thin pool "{0}" hit critical threshold at {2}%. Consider deleting some thin volumes, transferring some thin volumes into regular ones, changing access level for all thin volumes to read only or expanding the thin pool.	
Warn	Thin data allocation from thin pool "{0}" hit threshold at {2}%.	
Warn	Total memory for system has been increased. To provide more metadata space for existing thin pools consider thin pools expansion. For new thin pools the additional memory will be counted automatically.	
Warn	Unable to mirror the System Metadata volume "system-metadata" automatically. The system will operate normally, and the mirroring will be retried later. You may also add a sync image manually.	
Warn	Unknown concentrator from {0} attempting to join cluster.	
Warn	Update of {0} failed to {1} from {2}.	There was a problem updating the mirror status.
Warn	Usage of Snap and Thin Metadata hit System Threshold {1}%, highly fragmented snapshots should be deleted. Consider deleting snapshot:{0}.	
Warn	Usage of Snap, Thin, and Dedup Metadata hit System Threshold {1}%, highly fragmented snapshots should be deleted. Consider deleting snapshot:{0}.	
Warn	Usage of Thin Metadata hit System Threshold {1}%, highly allocated thin volumes should be deleted or transferred to regular	

Level	Event Message	Notes
	volumes. Consider to do this for thin volume:{0}.	
Warn	Your StoneFusion Base OS {0} day evaluation period has only {1} days remaining. Please contact customer support for a license for this product.	

Appendix 10

Software Upgrades

A10.1 Upgrading the Storage Concentrator Software

The *Storage Concentrator* software should be upgraded when new features and functionality become available. Upgrade the software using the Storage Concentrator Recovery CD.

Important Note: No Asynchronous Mirroring operations should be done while an Upgrade is being performed. All Asynchronous Mirror operations should be suspended for the period of the upgrade. When the upgrade is complete normal Asynchronous Mirroring operations can resume without any loss of data at the remote site. Refer to the section on Asynchronous Mirroring for instructions on how to suspend operations.

Contact StoneFly's technical support to acquire the most current *Storage Concentrator* Recovery CD.

To contact StoneFly, call 510.265.1616, 24 hours a day, 7 days a week, or go to the StoneFly Web site at: www.stonefly.com.

There are typically two scenarios regarding software installation with the Recovery CD:

Upgrade the software and retain the current configuration data

Upgrade the software and restore the factory default configuration data

IMPORTANT NOTE: Only if you are upgrading to version 6.4.2.9 or above and you are using asynchronous replication, you need to pay close attention to "Asynchronous Mirroring Management Migration", few sections down.

IMPORTANT NOTE: This release cannot be installed in a network where the Storage Concentrator Management Port and the iSCSI Data Port are on the same subnet. If you have configured your network in this way you must re-implement your network to provide different subnets for the Management Port and the Data Port. After that, complete the installation as described below.

IMPORTANT NOTE: System running Releases earlier than 3.0.0.x requires System upgrades to install this release.

A10.1.1 Installation/Upgrade Prerequisites

6.4.2.9 may be installed as an upgrade to older StoneFly products. The new StoneFly IP Storage appliances will be preloaded with Release 6.4.2.9 at the factory.

A new installation should follow the instructions contained in the Storage Concentrator Setup Guide. An upgrade installation should follow the steps outlined below. Upgrades at existing Storage Concentrator sites should always prepare for the upgrade by completing a data backup cycle and saving the Storage Concentrator's database (see the User Guide section on "AutoSave".)

Remove all flexible diskettes and USB flash sticks from the Storage Concentrator before starting any upgrade procedure.

Important Note: Upgrades of existing Storage Concentrators with StoneFusion

Versions prior to 3.0.0.xxx require a maintenance period to complete the upgrade. This requires host connections to be logged out at the host initiator before upgrading the Storage Concentrator.

There should be no volume or volume-image modifications going on during the installation or upgrade process. Specifically, there should be no volume-image rebuilds going on. You should not be performing upgrades when a snapshot is scheduled.

In all Storage Concentrator i3000 installations, the attached storage should not be configured with bootable partitions. This may happen when older storage is reclaimed and attached to the i3000. Prior to attaching previously used storage it should be reviewed and the LUNS/drives formatted as blank drives. Storage devices previously used behind a Storage Concentrator i1000, i1500 or i2000 system that are moved behind an i3000 do not typically require this type of formatting.

Important Note! A change was made in the database for systems starting in 4.3.0.25. Installing 6.4.2.9 over a version of StoneFusion older than 4.3.0.25 requires special handling if the database must be restored in the future. You cannot restore from a database earlier than 4.3.0.25 when 6.4.2.9 is installed on your system. Save your database again as soon as the newly installed system is working correctly.

The iSCSI initiators used in the host systems should be upgraded to the latest revision of firmware for iSCSI Storage Adapters and HBAs, as well as software drivers for both hardware and software-based initiators.

The Windows Disk Registry "TimeOutValue" must be changed to a value of 300 seconds. Please refer to User's Guide System Considerations section for details of how to change this.

Customers configuring FailOver Clusters should not use the motherboard-based SCSI port on older i3000 systems. Instead, connect and/or move the storage devices to the PCI-based SCSI ports. Please refer to the ***QuickStart Guide*** to locate these ports. To relocate an existing array to a different port, power down both the Storage Concentrator and the Storage Array. Move the storage cable to the PCI SCSI ports. Power up both the Storage Array and the Storage Concentrator. Rediscover the Storage Array.

A10.1.2 *Upgrading a Stand-Alone Unit*

There are four steps required to upgrade a stand-alone unit with the StoneFusion software:

1 Save the Configuration Data.

2 Log out the host initiators that access the Storage Concentrator.

3 Install the new software using the Recovery CD.

4 Log back in the hosts that access the Storage Concentrator.

Save the Configuration Data

This step saves the *Storage Concentrator* configuration data to a floppy disk or USB flash disk. This step must be performed to ensure that you have a copy of the latest configuration information before upgrading the Storage Concentrator software.

To save the Configuration Data, use the steps that follow:

1 From the System screen, click Admin.

2 Click Auto Save. The System Management Auto Save screen appears.

3 Remove the bezel on the front of the Storage Concentrator and insert a floppy disk into the disk drive or insert a USB Flash disk into any available USB port.

4 Click the Enable Auto Save to Local Device check box.

5 Select the appropriate device from the device list (Floppy or USB Flash).

6 Click Submit to initiate the Auto Save process to a local device.

7 When the saving process is completed, remove the floppy disk or USB Flash disk from the Storage Concentrator.

Log Out the Hosts

Log Out all hosts that access the *Storage Concentrator*.

Failure to Log Out the Hosts may result in timeout or I/O errors.

Install the Software using the Recovery CD

Please **note** that some systems are equipped with an internal CD-ROM drive while others are delivered with an external, USB drive. Please use the appropriate device when these instructions request actions with the Recovery CD. The external, USB drive may be connected to any available USB port on the Storage Concentrator.

There are typically two scenarios regarding software installation with the Recovery CD:

Upgrade the software and retain the current configuration data

Upgrade the software and restore the factory default configuration data

Upgrade the Software and Retain the Current Configuration Data

In most cases you will upgrade the software and retain the current configuration data. The Recovery CD performs the following tasks when upgrading the *Storage Concentrator* software:

1 Automatically saves the configuration data.

2 Re-initializes the *Storage Concentrator*.

3 Installs the new software.

4 Automatically restores the configuration data.

To install the new software and retain the current database, use the steps that follow:

1 Remove the front panel bezel from the *Storage Concentrator*.

2 Do one of the following:

If the *Storage Concentrator* is powered on:

a. Insert the Recovery CD into the CD-ROM drive.

b. Reboot the *Storage Concentrator* from the **System** screen by clicking **Admin**, then **General**, then **Reboot**.

If the *Storage Concentrator* is powered off:

a. Power it on.

b. Insert the Recovery CD into the CD-ROM drive so that the *Storage Concentrator* boots off the CD.

The *Storage Concentrator* boots from the CD and installs the StoneFusion software. The existing configuration data is restored. When the software install is complete, the Recovery CD is ejected from the CD-ROM drive. The *Storage Concentrator* automatically restarts. The *Storage Concentrator* takes approximately 5-7 minutes to boot up.

A10.1.3 *Upgrade the Software and Restore Factory Default Configuration Data*

Note that in newer SC SW releases, the command line System Console has the menu option "Admin - Reset Configuration DB to Factory Defaults" to reset the system to the Factory Default Configuration. If the intent is to remain on the current SW version, this may be the preferred procedure.

To install the new software and restore the factory default configuration data, use the steps that follow:

1 On a blank diskette or a USB Flash disk, create a text file named **default.db**.


The default.db text file signals the Recovery CD to not save current configuration data and to restore the factory default configuration data.

2 Remove the front panel bezel from the *Storage Concentrator*.

3 Insert the diskette or USB Flash disk with the **default.db** text file into the appropriate location on the system.

Using this diskette or a USB Flash disk will delete the existing configuration data.

4 Do one of the following:

If the *Storage Concentrator* is powered on:

a. Insert the Recovery CD into the CD-ROM drive.

b. Reboot the *Storage Concentrator* from the **System** screen by clicking **Admin**, then **General**, then **Reboot**.

If the *Storage Concentrator* is powered off:

a. Power it on.

b. Insert the Recovery CD into the CD-ROM drive so that the *Storage Concentrator* boots off the CD.

The *Storage Concentrator* boots from the CD and installs the StoneFusion software. The factory default configuration data is restored.

If there is a monitor connected to the system, early messages indicating the detection of the '**default.db**' file and the resulting default DB installation should be expected.

When the software install is complete, the Recovery CD is ejected from the CD-ROM drive. The *Storage Concentrator* automatically restarts. The *Storage Concentrator* takes approximately 5-7 minutes to boot up.

5 Be sure to remove the diskette or USB Flash disk from the system.

Log the Hosts Back In

Log Back in the Initiators at the host systems.

Refer to the documentation that came with the initiator for step-by-step instructions for Log Ins/Log Outs.

A10.1.4 *Upgrades for FailOver Clusters Running 3.0.0.xxx and Above*

If upgrading existing Storage Concentrators in a FailOver Cluster with Software Versions 5.0.0.xxx, it is recommended that you follow the procedure outlined below to upgrade. First, you will upgrade the Secondary system. The Primary Unit will continue to serve up I/Os while the upgrade of the Secondary is taking place. The Secondary unit will come back after upgrade with a status of "Primary Upgrade." You will then perform a reboot on the Primary unit causing it to FailOver to the "Primary Upgrade" unit, causing it to become the Primary system and take over all I/Os from the hosts. You can then complete the upgrade on the formerly Primary unit, which will rejoin the cluster as a Secondary Unit. Please be patient as this process will take about one hour to fully complete. Following the upgrade the Primary *Storage Concentrator* will automatically re-assign volumes to the Secondary *Storage Concentrator*, as directed in the Load Balancing screen.

On the Secondary Unit

Insure that all mirroring activities such as rebuilds are finished prior to starting the upgrade process. The upgrade should be done during a period of low activity.

1 Backup your database as described above. Once complete, remove the media device on which the database was saved.

2 Insert the Upgrade/Recovery CD into its drive on the Secondary Unit

3 Reboot the Secondary Storage Concentrator from its **System->Admin** screen using the Reboot button. Wait for the CD to Eject from the Secondary Unit and remove it from the system. Do not log in until approximately 10 minutes AFTER the CD Ejects.

4 When the unit comes up, you will see that there is a Full Login Screen rather than the Secondary Screen. You may want to close and restart the browser to flush the old address from the system cache.

5 Log back into the Storage Concentrator GUI on the Secondary Unit. Wait for Diagnostics to be Healthy (the Primary Unit will show as Unknown). This will take approximately 10 minutes. Note that the Secondary unit will have a Cluster Status of "Primary (Upgrade)" on the System Information page. You can also tell if the software is at the updated software version as shown below.

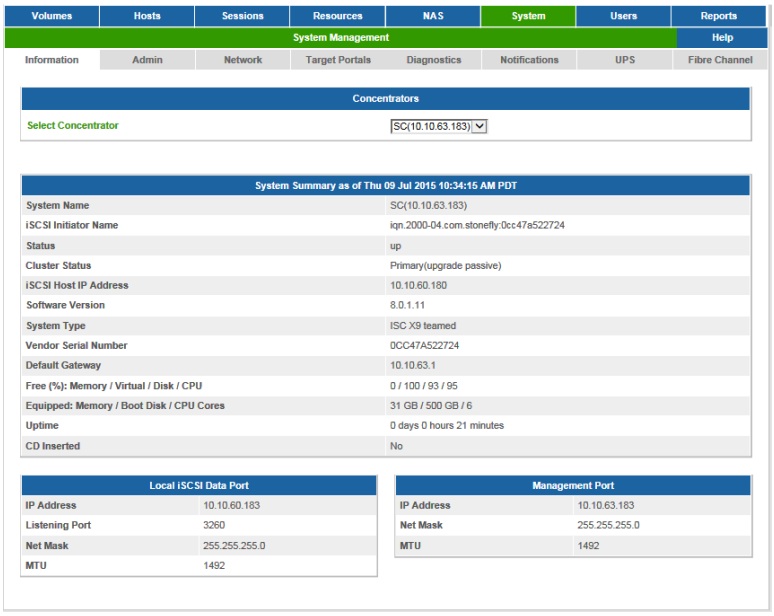
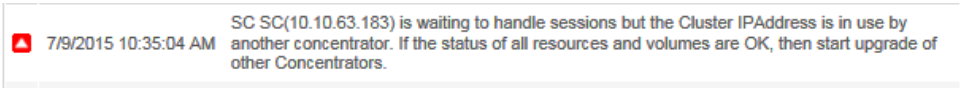


Figure App.-8 System Information Page Displaying Primary (upgrade) Status

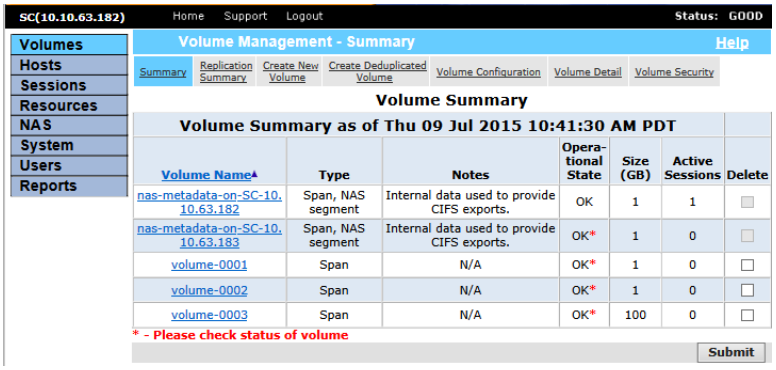
10 Verify the presence of a message in the Logs on the Secondary unit that says: "<system name> is waiting to handle sessions but the Cluster IP Address is in use by another Concentrator. If the status of all resources and volumes are ok, then start upgrade of other Concentrators." This indicates the former Secondary system has successfully updated and will become the Primary system when you are ready.

11



On the Primary Unit

1 Verify that Volumes and Resources are all OK (From Volume Summary and Resource Summary pages)



2 Insert the Upgrade/Recovery CD into its appropriate drive on the Primary Unit

3 Reboot the Primary Storage Concentrator from the GUI **<System, Admin, Reboot>**. Note that the Secondary Unit will become the new Primary Unit during this process. Wait for the CD to Eject and remove it from the formerly Primary Unit. Do not login until the system has completed rebooting (approximately 10 minutes after the CD ejects)

4 From the newly Primary Unit (previously the Secondary Unit), wait for the Diagnostics to be "Healthy" on the Diagnostics page for both Storage Concentrators. It will take **at least** 15 minutes for BOTH units to become Healthy

5 Check all volumes and make sure that the Status is OK

6 Check all initiators and make sure that all volumes are logged in and active

A10.1.5 Asynchronous Mirroring Management Migration

Only if you are upgrading to version 6.4.2.9 or above and you are using asynchronous replication, you need to pay close attention to this section.

The configuration for replication triggers and scheduling has been greatly simplified. A number of confusing controls have been removed, and there now is a single daily time to specify when to start a scheduled replication. NOTE: When using scheduled Async Replications and upgrading from a prior release, you should record the configuration settings before the upgrade, as they are not preserved during the upgrade. The desired Async Replication schedule settings can then be applied after the upgraded.

A10.1.6 Async Management General Page

Triggers

The Triggers link has been replaced by Schedule link. See Screen Shot 1 and Screen Shot 2.

Replication Priority

In previous releases the settings for replication priority where Low, Medium and High. New settings for replication priority are 1, 2, 3, 4, 5 and 6. The numeric value represents the number of threads used during replication operations. The correlation between the releases are Low = 1, Medium = 2, High = 3. The higher the replication priority the faster the replication will be completed. The time required to complete a Replication will depend on link speed, amount a data changed between replications and the Replication Priority.

Replication Point of no return (MB)

The section is obsolete. See Screen Shot 1.

Enable Internal IO Traffic Controller

This changed in 6.4.2.8 or above release. Select “Yes” to enable the Internal IO Traffic Controller. See Screen Shot 3. Select “No” to disable the Internal IO Traffic Controller. The default setting is “Yes”. In previous releases this was configured on the Triggers page as Disable Internal IO traffic Controller. See screen shot 3. With the Internal IO Traffic Controller enabled, replication will occur when the Volume in inactive for IO Idle Interval Seconds. With this setting replication will occur if there is fresh data to be replicated and the volume has not been written to for the set number of seconds. With this setting disabled replication will only occur for scheduled and Snapspace Usage to Activate Relief Valve Replication(%):

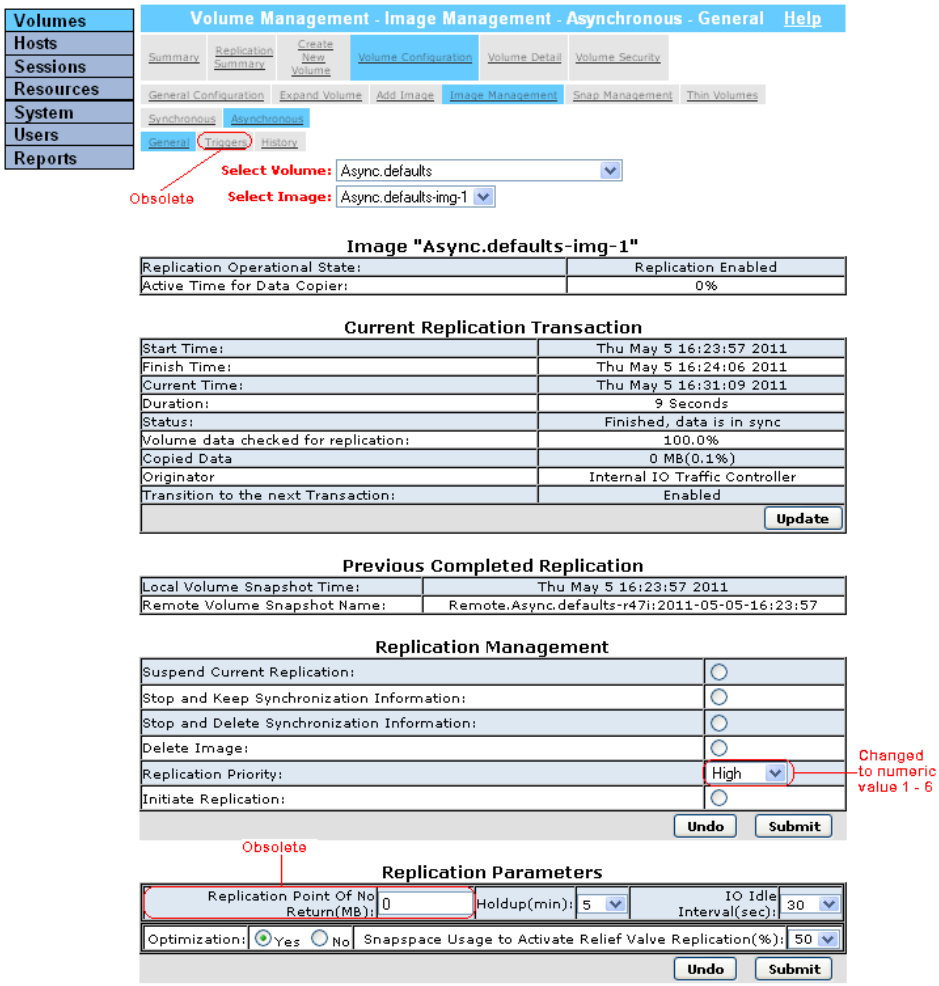


Figure App.-9 Screen Shot 1

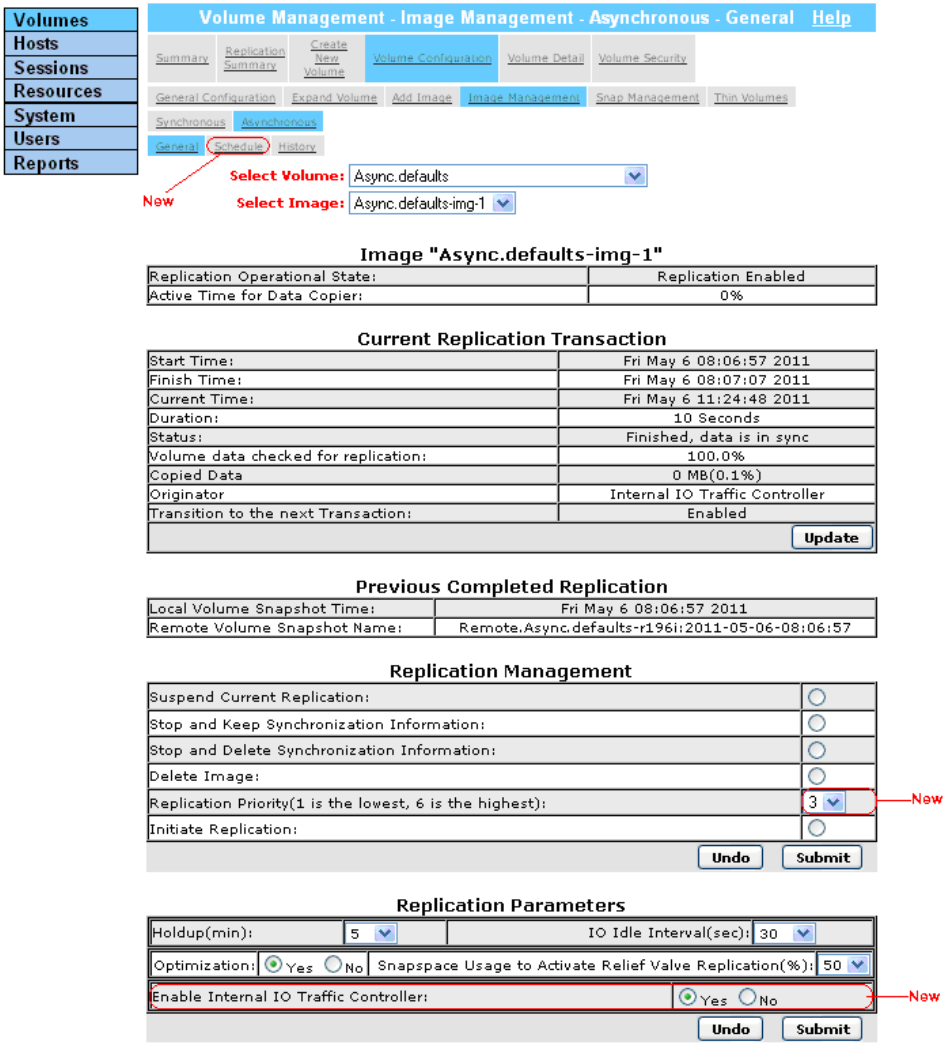


Figure App.-10 Screen Shot 2

A10.1.6.1 *Async Management Triggers (Obsolete)*

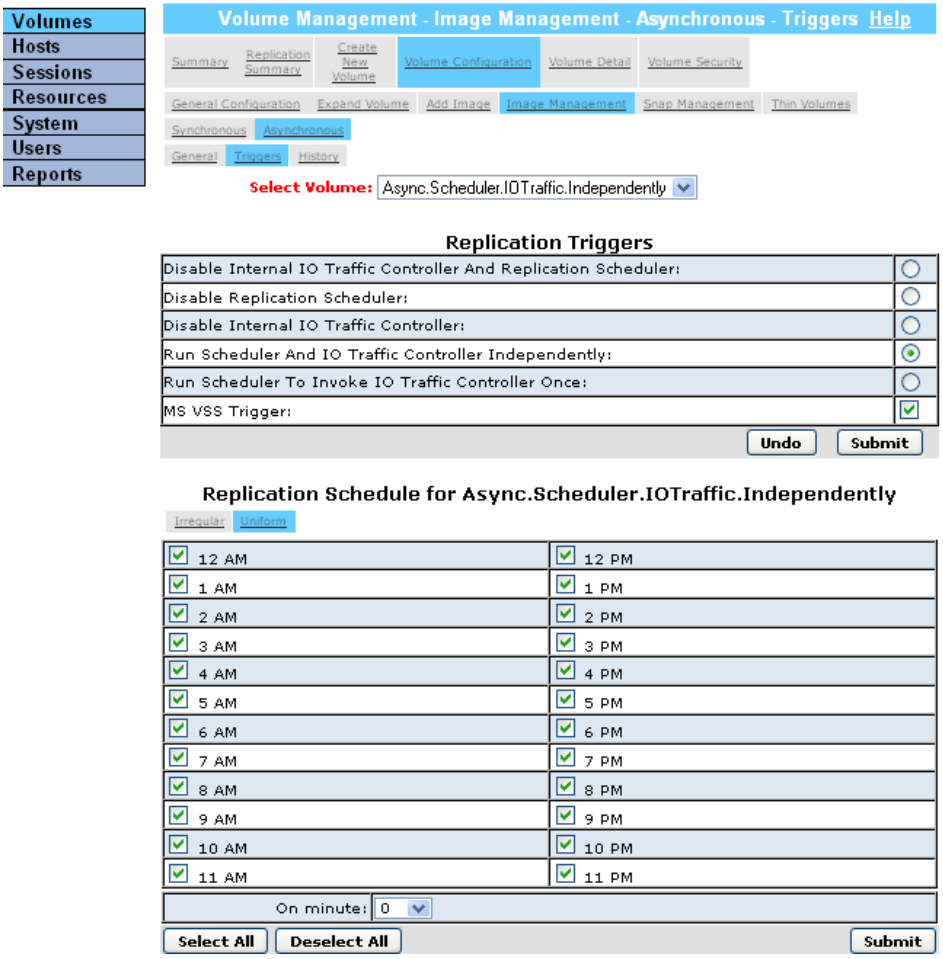


Figure App.-11 Screen Shot 3

Disable Internal IO Traffic Controller and Scheduler:

This is Obsolete. To configure the same ability Select "No" for Internal IO Traffic Controller (See Screen Shot 3). Also ensure Start new replication at resume time is unchecked (see Screen Shot 4). When configured in this manner Relief Valve SFVSSRM Replications will be the only Replications that occur. If async image has suspended replication, this replication has to be activated at the resume time regardless of settings for "Start new replication at resume time" check box.

Disable Replication Scheduler:

This is Obsolete.

Disable IO Internal IO Controller:

To disable the Internal IO Traffic Controller on 6.4.2.8 on newer releases Select "No" for Enable Internal IO Traffic Controller. The default setting is "Yes". See Screen Shot

Run Scheduler and IO Traffic Controller Independently:

To configure this setting use select "Yes" for Enable Internal IO Traffic Controller. See Screen Shot 2. Check Start new replication at resume time and configure resume and suspend times.

Run Scheduler to Invoke IO Traffic Controller Once:

This is Obsolete.

MS VSS Trigger:

This is no longer required. VSS or SFVSRM replication will occur when used.

A10.1.7 *Async Management Schedule Page*

Start new replication at resume time:

The default is unchecked. When uncheck replication will occur base on the IO Idle Interval (sec) value. When checked a replication will occur at the resume time. See Screen Shot 4.

Weekday Resume Suspend

This is new to 6.4.2.8 or above Release. These settings can be used to control when automatic replications can occur. The default setting of NONE does not limit the time period for replication to occur. Setting the Resume and Suspend times allows automatic replication to occur during the resume to suspend time window. During suspend times, manual and Relief Valve Replication will still occur.

Weekend Resume Suspend

This is new to 6.4.2. 8 or above Release. These settings can be used to control when automatic replications can occur. The default setting of NONE does not limit the time period for replication to occur. Setting the Resume and Suspend times allows automatic replication to occur during the resume to suspend time window. During suspend times, manual and Relief Valve Replication will still occur.

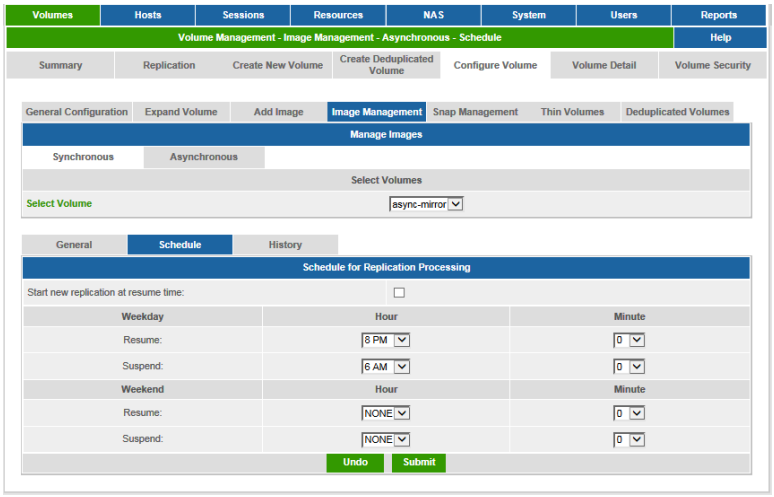


Figure App.-12 Screen Shot 4

This page is intentionally left blank.

Appendix 11

License Agreement and Warranty Information

A11.1 End-User License Agreement

IMPORTANT-READ CAREFULLY:

This StoneFly, Inc. (StoneFly) End-User License Agreement ("EULA") is a legal agreement between you (either an individual or a single entity) and StoneFly for the StoneFly software accompanying this EULA, which includes computer software and may include associated media, printed materials, and "online" or electronic documentation ("SOFTWARE PRODUCT" or "SOFTWARE"). By using the SOFTWARE PRODUCT, you agree to be bound by the terms of this EULA. If you do not agree to the terms of this EULA, you may not use the SOFTWARE PRODUCT.

1. Software PRODUCT LICENSE

The SOFTWARE PRODUCT is licensed, not sold. This EULA grants you the following non-exclusive rights:

1.1 **Use**. You may use the SOFTWARE PRODUCT as long as it is installed on the StoneFly hardware it was purchased with or on replacement StoneFly hardware in the event of a warranty replacement.

1.2 **Reproduction**. You may reproduce copies of the SOFTWARE PRODUCT for backup or archival purposes only.

2. DESCRIPTION OF OTHER RIGHTS AND LIMITATIONS

2.1 **Reverse Engineering, Decompilation, and Disassembly**. You may not, and you shall not allow third parties to, reverse engineer, decompile, or disassemble the SOFTWARE PRODUCT, except and only to the extent that such activity is expressly permitted by applicable law notwithstanding this limitation.

2.2 **Modification**. You may not, and you shall not allow third parties to, modify the SOFTWARE PRODUCT or incorporate the SOFTWARE PRODUCT into, or with, any other software.

2.3 **Separation of Components**. The SOFTWARE PRODUCT is licensed as a single product. The SOFTWARE PRODUCT's component parts may not be separated for use on more than one computer.

2.4 **Software Transfer**. You may permanently transfer all of your rights under this EULA, provided the recipient agrees to the terms of this EULA.

2.5 **Product Identification**. You may not, and you shall not allow third parties to, remove any SOFTWARE PRODUCT identification or other notices.

3. COPYRIGHT

The SOFTWARE PRODUCT is protected by copyright laws and international copyright treaties, as well as other intellectual property laws and treaties. All title, copyrights and all other intellectual property rights in and to the SOFTWARE PRODUCT (including but not limited to any images, photographs, animations, video, audio, music, text, and

"applets" incorporated into the SOFTWARE PRODUCT), the accompanying printed materials, and any copies of the SOFTWARE PRODUCT are owned by StoneFly or its suppliers. The SOFTWARE PRODUCT is protected by copyright laws and international treaty provisions. Therefore, you must treat the SOFTWARE PRODUCT like any other copyrighted material.

4. TRADEMARKS

"StoneFly" and any other registered and non-registered trademarks are trademarks of StoneFly, All Rights Reserved. StoneFly's failure to list a trademark in this Section shall not constitute a waiver of any trademark rights. All other trademarks in the SOFTWARE PRODUCT not owned by StoneFly are the property of their respective owners.

5. U.S. GOVERNMENT RESTRICTED RIGHTS

The SOFTWARE PRODUCT and documentation are provided with RESTRICTED RIGHTS. Use, duplication, or disclosure by the Government is subject to restrictions as set forth in subparagraph ©(1)(ii) of the Rights in Technical Data and Computer Software clause at DFARS 252.227-7013 or subparagraphs ©(1) and (2) of the Commercial Computer Software-Restricted Rights at 48 CFR 52.227-19, as applicable. Manufacturer is StoneFly, Inc., 26250 Eden Landing Road, Hayward, CA 94545.

6. LIMITED WARRANTY

6.1 **NO WARRANTIES**. STONEFLY EXPRESSLY DISCLAIMS ANY WARRANTY FOR THE SOFTWARE PRODUCT. THE SOFTWARE PRODUCT AND ANY RELATED DOCUMENTATION IS PROVIDED "AS IS" WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESS OR IMPLIED, INCLUDING, WITHOUT LIMITATION, THE IMPLIED WARRANTIES OR MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE, OR NONINFRINGEMENT. THE ENTIRE RISK ARISING OUT OF USE OR PERFORMANCE OF THE SOFTWARE PRODUCT REMAINS WITH YOU.

6.2 Some jurisdictions do not allow the exclusion of implied warranties, so the above exclusion may not apply to you. You may have other rights which vary from jurisdiction to jurisdiction.

6.3 **LIMITATION OF LIABILITY**. THE ENTIRE RISK AS TO THE RESULTS AND PERFORMANCE OF THE SOFTWARE PRODUCT IS ASSUMED BY YOU. STONEFLY SHALL NOT HAVE ANY LIABILITY TO YOU OR ANY OTHER PERSON OR ENTITY FOR ANY INDIRECT, INCIDENTAL, SPECIAL, OR CONSEQUENTIAL DAMAGES WHATSOEVER, INCLUDING, BUT NOT LIMITED TO, LOSS OF REVENUE OR PROFIT, LOST OR DAMAGED DATA OR OTHER COMMERCIAL OR ECONOMIC LOSS, EVEN IF STONEFLY HAS BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGES, OR THEY ARE FORESEEABLE. OUR MAXIMUM AGGREGATE LIABILITY TO YOU FOR DIRECT DAMAGES SHALL NOT EXCEED THE AMOUNT PAID BY YOU FOR THE SOFTWARE PRODUCT. THE LIMITATIONS IN THIS SECTION SHALL APPLY WHETHER OR NOT THE ALLEGED BREACH OR DEFAULT IS A BREACH OF A FUNDAMENTAL CONDITION OR TERM OR A FUNDAMENTAL BREACH.

Because some states/jurisdictions do not allow the exclusion or limitation of liability for consequential or incidental damages, the above limitation may not apply to you.

7. TERM

The EULA is effective until terminated. You may terminate the EULA at any time by returning or destroying all copies of the SOFTWARE PRODUCT and related documentation. The EULA will terminate automatically if you fail to comply with any term or condition of the EULA, including any attempt to transfer a copy of the SOFTWARE PRODUCT to another party except as provided in the EULA. You agree upon such termination, you will return or destroy all copies of the EULA and related documentation. The provisions of Sections 2 – 8 shall survive termination.

8. Miscellaneous

8.1 **Governing Law**. You agree that this EULA is governed by the laws of the State of California, without reference to conflicts of law principles or the United Nations Convention on Contracts for the International Sale of Goods. The sole jurisdiction and venue for actions related to the subject matter hereof shall be the state and U.S. federal courts having within their jurisdiction the location of the StoneFly's principal place of business. Both parties consent to the jurisdiction of such courts and agree that process may be served in the manner provided herein for giving notices or otherwise as allowed by California state or U.S. federal law. In any action to enforce this Agreement, the prevailing party shall be entitled to costs and attorneys' fees.

8.2 **Severability**. If any provision of these End-User Terms and Conditions is held to be invalid, illegal or unenforceable, it shall be severed and the remaining provisions shall continue in full force and effect.

8.3 **Entire Agreement and Waiver**. The EULA is the entire agreement regarding your use of the SOFTWARE PRODUCTS, superseding any other agreement or discussions, oral or written, and may not be changed except by a signed agreement. At no time shall a failure or delay in enforcing any provisions, exercising any option or requiring performance, be construed to be a waiver.

Should you have any questions concerning this EULA, or if you desire to contact StoneFly for any reason, please write to: StoneFly, Inc., 26250 Eden Landing Road, Hayward, CA 94545.

A11.2 One-Year Limited Warranty (U.S. Only)

StoneFly, Inc. ("StoneFly") manufactures its hardware products from parts and components that are new or equivalent to new in accordance with industry-standard practices. StoneFly warrants that the hardware products it manufactures will be free from defects in materials and workmanship during the term of the limited warranty. The limited warranty term is one year beginning from the date of shipment from StoneFly, as described in the following text.

THIS LIMITED WARRANTY GIVES YOU SPECIFIC LEGAL RIGHTS, AND YOU MAY ALSO HAVE OTHER RIGHTS, WHICH VARY FROM STATE TO STATE (OR JURISDICTION TO JURISDICTION). STONEFLY'S RESPONSIBILITY AND YOUR SOLE REMEDY UNDER THIS LIMITED WARRANTY IS LIMITED TO REPAIR AND REPLACEMENT AS SET FORTH IN THIS LIMITED WARRANTY STATEMENT. ALL EXPRESS AND IMPLIED WARRANTIES FOR THE PRODUCT, INCLUDING BUT NOT LIMITED TO ANY IMPLIED WARRANTIES OF AND CONDITIONS OF MERCHANTIBILITY AND FITNESS FOR A PARTICULAR PURPOSE, ARE LIMITED IN DURATION TO THE WARRANTY PERIOD SET FORTH ABOVE AND NO

WARRANTIES, WHETHER EXPRESS OR IMPLIED, WILL APPLY AFTER SUCH PERIOD.

SOME STATES (OR JURISDICTIONS) DO NOT ALLOW LIMITATIONS ON HOW LONG AN IMPLIED WARRANTY LASTS, SO THE PRECEDING LIMITATION MAY NOT APPLY TO YOU.

STONEFLY DOES NOT ACCEPT LIABILITY BEYOND THE REMEDIES SET FORTH IN THIS LIMITED WARRANTY STATEMENT OR LIABILITY FOR INCIDENTAL OR CONSEQUENTIAL DAMAGES, INCLUDING WITHOUT LIMITATION ANY LIABILITY FOR PRODUCTS NOT BEING AVAILABLE FOR USE OR FOR LOST DATA OR SOFTWARE.

SOME STATES (OR JURISDICTIONS) DO NOT ALLOW THE EXCLUSION OR LIMITATION OF INCIDENTAL OR CONSEQUENTIAL DAMAGES, SO THE PRECEDING EXCLUSION OR LIMITATION MAY NOT APPLY TO YOU. THIS LIMITATION IS INTENDED TO APPLY AND DOES APPLY WITHOUT REGARD TO WHETHER SUCH DAMAGES ARE CLAIMED, ASSERTED OR BROUGHT IN AN ACTION OR CLAIM SOUNDING IN TORT OR CONTRACT, OR ON THE WARRANTY, OR UNDER ANY OTHER LAW OR FORM OF ACTION.

These provisions apply to StoneFly's one-year limited warranty only. For provisions of any other service contract covering your system, refer to your invoice or the separate service contract that you will receive.

If StoneFly elects to exchange a system or component, the exchange will be made in accordance with StoneFly's Exchange Policy in effect on the date of the exchange. In any instance in which StoneFly issues a Return Material Authorization Number, StoneFly must receive the product(s) for repair prior to the expiration of the warranty period in order for the repair(s) to be covered by the limited warranty.

This limited warranty does not cover damage due to external causes, including accident, abuse, misuse, problems with electrical power, servicing not authorized by StoneFly, usage not in accordance with product instructions, failure to perform required preventive maintenance, and problems caused by use of parts and components not supplied by StoneFly.

This limited warranty does not cover any items that are in one or more of the following categories: software; external devices (except as specifically noted); accessories or parts added to a StoneFly system after the system is shipped from StoneFly; accessories or parts that are not installed in the StoneFly factory; or StoneFly Software & Peripherals products.

A11.3 Service Policy

Your StoneFly Storage Concentrator product is covered 7 days a week, 24 hours a day, and 365 days a year by StoneFly's web-based and telephone technical support for one year from the date of shipment from StoneFly, as described in the following text. Software maintenance for StoneFusion software for Storage Concentrators is provided for a period of one year from date of purchase.

In order to initiate StoneFly Customer Support coverage for your product you must first register your product via StoneFly's web site, www.stonefly.com, or by returning a completed Registration Card, which was included with your product documentation. Please have your StoneFly product serial number(s) available when you start the registration process. Please enter the serial numbers from each of the products purchased.

StoneFly offers optional on-site hardware service plans. For covered hardware products, in the event on-site service is required, a trained technician will arrive on-site, depending on location, within 4 hours after telephone trouble-shooting to help get systems back in operation as rapidly as possible (note: 4-hour response parts are sent directly to the customer via local courier).

If you purchase an extended service contract after purchasing an initial service contract, your system must be re-certified before a new service contract may be purchased. Any repair needs identified during re-certification are your responsibility, and repairs must be completed prior to purchase of the new service contract. Although reasonable effort is made to notify you before your service contract expires, StoneFly is not liable for any failure to do so.

A11.4 Legal Terms and Disclosures

The information in this document has been reviewed and is believed to be accurate. However, neither StoneFly nor its affiliates assume any responsibility for inaccuracies, errors, or emissions that may be contained herein. In no event will StoneFly or its affiliates be liable for direct, indirect, special, incidental, or consequential damages resulting from any defect or omission in this document, even if advised of the possibility of such damages. StoneFly reserves the right to make improvements or changes to this document and the products and services described at any time, without notice or obligation.

A11.5 Trade Names

StoneFly, the StoneFly logo, Storage Concentrator, Integrated Storage Concentrator, ISC, Modular Storage Concentrator, StoneFly Backup Advantage, StoneFusion, StoneFly Replicator CDP, ValueSAN, Unified Scale Out, USO, Super Scale Out, SSO, Twin Scale Out, TSO, Unified Storage & Server, USS, Unified Deduplicated Storage, UDS, Unified Encrypted Storage, UES, OptiSAN, StoneFly Voyager, DR365, DR365 Fusion, StoneFly Mirroring, Storage Concentrator Virtual Machine, SCVM, Software-Defined Unified Storage, SDUS, and StoneFly Cloud Drive are trademarks of StoneFly, Inc., a wholly owned subsidiary of Dynamic Network Factory, Inc. Other trademarks and trade names may be used in this document to refer to either the entities claiming the marks and names or their products. StoneFly disclaims any proprietary interest in trademarks and trade names other than its own. StoneFly cannot be responsible for errors in typography and photography. Service specifications are valid in the US only and subject to change without notice.

This page is intentionally left blank.

Appendix 12

Other Licenses

GNU General Public License

Standard versions of the following software modules are installed on your StoneFly product along with StoneFly-developed software, reference StoneFly End User License Agreement. StoneFly has not made modifications to these modules.

Listed below are the copyright notices for each of these software modules, as well a copy of the GNU General Public License, GNU Library General Public License, Source Code Agreement, Other Copyrights and the Artistic License that apply. Complete, machine-readable source code for each of these software modules is provided on your StoneFly product.

ANSIColor

Copyright 1996, 1997, 1998, 2000, 2001 Russ Allbery <rra@stanford.edu> and Zenin <zenin@bawdycaste.org>. This program is free software; you may redistribute it and/or modify it under the same terms as Perl itself.

CGI.pm

Copyright 1995-1997, Lincoln D. Stein. All rights reserved. It may be used and modified freely, but I do request that this copyright notice remain attached to the file. You may modify this module as you wish, but if you redistribute a modified version, please attach a note listing the modifications you have made.

libnet

© 1996-2000 Graham Barr. All rights reserved. This library is free software; you can redistribute it and/ or modify it under the same terms as Perl itself.

MIME-Base64

Copyright 1995-1999,2001 Gisle Aas. This library is free software; you can redistribute it and/or modify it under the same terms as Perl itself.

Time-HiRes

Copyright © 1996, 1997, 1998 Douglas E. Wegscheid. All rights reserved. This program is free software; you can redistribute it and/or modify it under the same terms as Perl itself.

ApacheDBI

© Edmund Mergl, November 20, 1999. You may distribute under the terms of either the GNU General Public License or the Artistic License, as specified in the Perl README file.

Data-Dumper

Copyright © 1996-98 Gurusamy Sarathy. All rights reserved. This program is free software; you can redistribute it and/or modify it under the same terms as Perl itself.

DBI.pm

Mail-Sendmail

SNMP

URI

Preamble

The licenses for most software are designed to take away your freedom to share and change it. By contrast, the GNU General Public License is intended to guarantee your freedom to share and change free software-to make sure the software is free for all its users. This General Public License applies to most of the Free Software Foundation's software and to any other program whose authors commit to using it. (Some other Free Software Foundation software is covered by the GNU Library General Public License instead.) You can apply it to your programs, too.

When we speak of free software, we are referring to freedom, not price. Our General Public Licenses are designed to make sure that you have the freedom to distribute copies of free software (and charge for this service if you wish), that you receive source code or can get it if you want it, that you can change the software or use pieces of it in new free programs; and that you know you can do these things.

To protect your rights, we need to make restrictions that forbid anyone to deny you these rights or to ask you to surrender the rights. These restrictions translate to certain responsibilities for you if you distribute copies of the software, or if you modify it.

For example, if you distribute copies of such a program, whether gratis or for a fee, you must give the recipients all the rights that you have. You must make sure that they, too, receive or can get the source code. And you must show them these terms so they know their rights.

We protect your rights with two steps: (1) copyright the software, and (2) offer you this license which gives you legal permission to copy, distribute and/or modify the software.

Also, for each author's protection and ours, we want to make certain that everyone understands that there is no warranty for this free software. If the software is modified by someone else and passed on, we want its recipients to know that what they have is not the original, so that any problems introduced by others will not reflect on the original authors' reputations.

Finally, any free program is threatened constantly by software patents. We wish to avoid the danger that re-distributors of a free program will individually obtain patent licenses, in effect making the program proprietary. To prevent this, we have made it clear that any patent must be licensed for everyone's free use or not licensed at all.

The precise terms and conditions for copying, distribution and modification follow.

A12.1 TERMS AND CONDITIONS FOR COPYING, DISTRIBUTING AND MODIFICATION

0 This License applies to any program or other work which contains a notice placed by the copyright holder saying it may be distributed under the terms of this General Public License. The "Program", below, refers to any such program or work, and a "work based on the Program" means either the Program or any derivative work under copyright law: that is to say, a work containing the Program or a portion of it, either verbatim or with modifications and/or translated into another language. (Hereinafter, translation is included without limitation in the term "modification".) Each licensee is addressed as "you".

Activities other than copying, distribution and modification are not covered by this License; they are outside its scope. The act of running the Program is not restricted, and the output from the Program is covered only if its contents constitute a work based on the Program (independent of having been made by running the Program). Whether that is true depends on what the Program does.

1 You may copy and distribute verbatim copies of the Program's source code as you receive it, in any medium, provided that you conspicuously and appropriately publish on each copy an appropriate copyright notice and disclaimer of warranty; keep intact all the notices that refer to this License and to the absence of any warranty; and give any other recipients of the Program a copy of this License along with the Program.

You may charge a fee for the physical act of transferring a copy, and you may at your option offer warranty protection in exchange for a fee.

2 You may modify your copy or copies of the Program or any portion of it, thus forming a work based on the Program, and copy and distribute such modifications or work under the terms of Section 1 above, provided that you also meet all of these conditions:

a) You must cause the modified files to carry prominent notices stating that you changed the files and the date of any change.

b) You must cause any work that you distribute or publish, that in whole or in part contains or is derived from the Program or any part thereof, to be licensed as a whole at no charge to all third parties under the terms of this License.

c) If the modified program normally reads commands interactively when run, you must cause it, when started running for such interactive use in the most ordinary way, to print or display an announcement including an appropriate copyright notice and a notice that there is no warranty (or else, saying that you provide a warranty) and that users may redistribute the program under these conditions, and telling the user how to view a copy of this License. (Exception: if the Program itself is interactive but does not normally print such an announcement, your work based on the Program is not required to print an announcement.)

These requirements apply to the modified work as a whole. If identifiable sections of that work are not derived from the Program, and can be reasonably considered independent and separate works in themselves, then this License, and its terms, do not apply to those sections when you distribute them as separate works. But when you distribute the same sections as part of a whole which is a work based on the Program, the distribution of the whole must be on the terms of this License, whose permissions for other licensees extend to the entire whole, and thus to each and every part regardless of who wrote it.

Thus, it is not the intent of this section to claim rights or contest your rights to work written entirely by you; rather, the intent is to exercise the right to control the distribution of derivative or collective works based on the Program.

In addition, mere aggregation of another work not based on the Program with the Program (or with a work based on the Program) on a volume of a storage or distribution medium does not bring the other work under the scope of this License.

3 You may copy and distribute the Program (or a work based on it, under Section 2) in object code or executable form under the terms of Sections 1 and 2 above provided that you also do one of the following:

a) Accompany it with the complete corresponding machine-readable source code, which must be distributed under the terms of Sections 1 and 2 above on a medium customarily used for software interchange; or,

b) Accompany it with a written offer, valid for at least three years, to give any third party, for a charge no more than your cost of physically performing source distribution, a complete machine-readable copy of the corresponding source code, to be distributed under the terms of Sections 1 and 2 above on a medium customarily used for software interchange; or,

c) Accompany it with the information you received as to the offer to distribute corresponding source code. (This alternative is allowed only for noncommercial distribution and only if you received the program in object code or executable form with such an offer, in accord with Subsection b above.)

The source code for a work means the preferred form of the work for making modifications to it. For an executable work, complete source code means all the source code for all modules it contains, plus any associated interface definition files, plus the scripts used to control compilation and installation of the executable. However, as a

special exception, the source code distributed need not include anything that is normally distributed (in either source or binary form) with the major components (compiler, kernel, and so on) of the operating system on which the executable runs, unless that component itself accompanies the executable.

If distribution of executable or object code is made by offering access to copy from a designated place, then offering equivalent access to copy the source code from the same place counts as distribution of the source code, even though third parties are not compelled to copy the source along with the object code.

4 You may not copy, modify, sublicense, or distribute the Program except as expressly provided under this License. Any attempt otherwise to copy, modify, sublicense or distribute the Program is void, and will automatically terminate your rights under this License. However, parties who have received copies, or rights, from you under this License will not have their licenses terminated so long as such parties remain in full compliance.

5 You are not required to accept this License, since you have not signed it. However, nothing else grants you permission to modify or distribute the Program or its derivative works. These actions are prohibited by law if you do not accept this License. Therefore, by modifying or distributing the Program (or any work based on the Program), you indicate your acceptance of this License to do so, and all its terms and conditions for copying, distributing or modifying the Program or works based on it.

6 Each time you redistribute the Program (or any work based on the Program), the recipient automatically receives a license from the original licensor to copy, distribute or modify the Program subject to these terms and conditions. You may not impose any further restrictions on the recipients' exercise of the rights granted herein. You are not responsible for enforcing compliance by third parties to this License.

7 If, as a consequence of a court judgment or allegation of patent infringement or for any other reason (not limited to patent issues), conditions are imposed on you (whether by court order, agreement or otherwise) that contradict the conditions of this License, they do not excuse you from the conditions of this License. If you cannot distribute so as to satisfy simultaneously your obligations under this License and any other pertinent obligations, then as a consequence you may not distribute the Program at all. For example, if a patent license would not permit royalty-free redistribution of the Program by all those who receive copies directly or indirectly through you, then the only way you could satisfy both it and this License would be to refrain entirely from distribution of the Program.

If any portion of this section is held invalid or unenforceable under any particular circumstance, the balance of the section is intended to apply and the section as a whole is intended to apply in other circumstances.

It is not the purpose of this section to induce you to infringe any patents or other property right claims or to contest validity of any such claims; this section has the sole purpose of protecting the integrity of the free software distribution system, which is implemented by public license practices. Many people have made generous contributions to the wide range of software distributed through that system in reliance on consistent application of that system; it is up to the author/donor to decide if he or she is willing to distribute software through any other system and a licensee cannot impose that choice.

This section is intended to make thoroughly clear what is believed to be a consequence of the rest of this License.

8 If the distribution and/or use of the Program is restricted in certain countries either by patents or by copyrighted interfaces, the original copyright holder who places the Program under this License may add an explicit geographical distribution limitation excluding those countries, so that distribution is permitted only in or among countries not thus excluded. In such case, this License incorporates the limitation as if written in the body of this License.

9 The Free Software Foundation may publish revised and/or new versions of the General Public License from time to time. Such new versions will be similar in spirit to the present version, but may differ in detail to address new problems or concerns.

Each version is given a distinguishing version number. If the Program specifies a version number of this License which applies to it and "any later version", you have the option of following the terms and conditions either of that version or of any later version published by the Free Software Foundation. If the Program does not specify a version number of this License, you may choose any version ever published by the Free Software Foundation.

10 If you wish to incorporate parts of the Program into other free programs whose distribution conditions are different, write to the author to ask for permission. For software which is copyrighted by the Free Software Foundation, write to the Free Software Foundation; we sometimes make exceptions for this. Our decision will be guided by the two goals of preserving the free status of all derivatives of our free software and of promoting the sharing and reuse of software generally.

NO WARRANTY

11 BECAUSE THE PROGRAM IS LICENSED FREE OF CHARGE, THERE IS NO WARRANTY FOR THE PROGRAM, TO THE EXTENT PERMITTED BY APPLICABLE LAW. EXCEPT WHEN OTHERWISE STATED IN WRITING THE COPYRIGHT HOLDERS AND/OR OTHER PARTIES PROVIDE THE PROGRAM "AS IS" WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESSED OR IMPLIED, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE. THE ENTIRE RISK AS TO THE QUALITY AND PERFORMANCE OF THE PROGRAM IS WITH YOU. SHOULD THE PROGRAM PROVE DEFECTIVE, YOU ASSUME THE COST OF ALL NECESSARY SERVICING, REPAIR OR CORRECTION.

12 IN NO EVENT UNLESS REQUIRED BY APPLICABLE LAW OR AGREED TO IN WRITING WILL ANY COPYRIGHT HOLDER, OR ANY OTHER PARTY WHO MAY MODIFY AND/OR REDISTRIBUTE THE PROGRAM AS PERMITTED ABOVE, BE LIABLE TO YOU FOR DAMAGES, INCLUDING ANY GENERAL, SPECIAL, INCIDENTAL OR CONSEQUENTIAL DAMAGES ARISING OUT OF THE USE OR INABILITY TO USE THE PROGRAM (INCLUDING BUT NOT LIMITED TO LOSS OF DATA OR DATA BEING RENDERED INACCURATE OR LOSSES SUSTAINED BY YOU OR THIRD PARTIES OR A FAILURE OF THE PROGRAM TO OPERATE WITH ANY OTHER PROGRAMS), EVEN IF SUCH HOLDER OR OTHER PARTY HAS BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGES.

GNU Library General Public License

Version 2, June 1991 Copyright © 1991 Free Software Foundation, Inc. 59 Temple

Place - Suite 330, Boston, MA 02111-1307, USA

Everyone is permitted to copy and distribute verbatim copies of this license document, but changing it is not allowed.

[This is the first released version of the library GPL. It is numbered 2 because it goes with version 2 of the ordinary GPL.]

Preamble

The licenses for most software are designed to take away your freedom to share and change it. By contrast, the GNU General Public Licenses are intended to guarantee your freedom to share and change free software—to make sure the software is free for all its users.

This license, the Library General Public License, applies to some specially designated Free Software Foundation software, and to any other libraries whose authors decide to use it. You can use it for your libraries, too.

When we speak of free software, we are referring to freedom, not price. Our General Public Licenses are designed to make sure that you have the freedom to distribute copies of free software (and charge for this service if you wish), that you receive source code or can get it if you want it, that you can change the software or use pieces of it in new free programs; and that you know you can do these things.

To protect your rights, we need to make restrictions that forbid anyone to deny you these rights or to ask you to surrender the rights. These restrictions translate to certain responsibilities for you if you distribute copies of the library, or if you modify it.

For example, if you distribute copies of the library, whether gratis or for a fee, you must give the recipients all the rights that we gave you. You must make sure that they, too, receive or can get the source code. If you link a program with the library, you must provide complete object files to the recipients so that they can relink them with the library, after making changes to the library and recompiling it. And you must show them these terms so they know their rights.

Our method of protecting your rights has two steps: (1) copyright the library, and (2) offer you this license which gives you legal permission to copy, distribute and/or modify the library.

Also, for each distributor’s protection, we want to make certain that everyone understands that there is no warranty for this free library. If the library is modified by someone else and passed on, we want its recipients to know that what they have is not the original version, so that any problems introduced by others will not reflect on the original authors’ reputations.

Finally, any free program is threatened constantly by software patents. We wish to avoid the danger that companies distributing free software will individually obtain patent licenses, thus in effect transforming the program into proprietary software. To prevent this, we have made it clear that any patent must be licensed for everyone’s free use or not licensed at all.

Most GNU software, including some libraries, is covered by the ordinary GNU General

Public License, which was designed for utility programs. This license, the GNU Library General Public License, applies to certain designated libraries. This license is quite different from the ordinary one; be sure to read it in full, and don't assume that anything in it is the same as in the ordinary license.

The reason we have a separate public license for some libraries is that they blur the distinction we usually make between modifying or adding to a program and simply using it. Linking a program with a library, without changing the library, is in some sense simply using the library, and is analogous to running a utility program or application program. However, in a textual and legal sense, the linked executable is a combined work, a derivative of the original library, and the ordinary General Public License treats it as such.

Because of this blurred distinction, using the ordinary General Public License for libraries did not effectively promote software sharing, because most developers did not use the libraries. We concluded that weaker conditions might promote sharing better.

However, unrestricted linking of non-free programs would deprive the users of those programs of all benefit from the free status of the libraries themselves. This Library General Public License is intended to permit developers of non-free programs to use free libraries, while preserving your freedom as a user of such programs to change the free libraries that are incorporated in them. (We have not seen how to achieve this as regards changes in header files, but we have achieved it as regards changes in the actual functions of the Library.) The hope is that this will lead to faster development of free libraries.

The precise terms and conditions for copying, distribution and modification follow. Pay close attention to the difference between a "work based on the library" and a "work that uses the library". The former contains code derived from the library, while the latter only works together with the library.

Note that it is possible for a library to be covered by the ordinary General Public License rather than by this special one.

A12.2 TERMS AND CONDITIONS FOR COPYING, DISTRIBUTION, AND MODIFICATION of Software Library

0 This License Agreement applies to any software library which contains a notice placed by the copyright holder or other authorized party saying it may be distributed under the terms of this Library General Public License (also called "this License"). Each licensee is addressed as "you".

A "library" means a collection of software functions and/or data prepared so as to be conveniently linked with application programs (which use some of those functions and data) to form executables.

The "Library", below, refers to any such software library or work which has been distributed under these terms. A "work based on the Library" means either the Library or any derivative work under copyright law: that is to say, a work containing the Library or a portion of it, either verbatim or with modifications and/or translated straightforwardly into another language. (Hereinafter, translation is included without

limitation in the term "modification".)

"Source code" for a work means the preferred form of the work for making modifications to it. For a library, complete source code means all the source code for all modules it contains, plus any associated interface definition files, plus the scripts used to control compilation and installation of the library.

Activities other than copying, distribution and modification are not covered by this License; they are outside its scope. The act of running a program using the Library is not restricted, and output from such a program is covered only if its contents constitute a work based on the Library (independent of the use of the Library in a tool for writing it). Whether that is true depends on what the Library does and what the program that uses the Library does.

1 You may copy and distribute verbatim copies of the Library's complete source code as you receive it, in any medium, provided that you conspicuously and appropriately publish on each copy an appropriate copyright notice and disclaimer of warranty; keep intact all the notices that refer to this License and to the absence of any warranty; and distribute a copy of this License along with the Library.

You may charge a fee for the physical act of transferring a copy, and you may at your option offer warranty protection in exchange for a fee.

2 You may modify your copy or copies of the Library or any portion of it, thus forming a work based on the Library, and copy and distribute such modifications or work under the terms of Section 1 above, provided that you also meet all of these conditions:

a) The modified work must itself be a software library.

b) You must cause the files modified to carry prominent notices stating that you changed the files and the date of any change.

c) You must cause the whole of the work to be licensed at no charge to all third parties under the terms of this License.

d) If a facility in the modified Library refers to a function or a table of data to be supplied by an application program that uses the facility, other than as an argument passed when the facility is invoked, then you must make a good faith effort to ensure that, in the event an application does not supply such function or table, the facility still operates, and performs whatever part of its purpose remains meaningful. (For example, a function in a library to compute square roots has a purpose that is entirely well-defined independent of the application. Therefore, Subsection 2d requires that any application-supplied function or table used by this function must be optional: if the application does not supply it, the square root function must still compute square roots.)

These requirements apply to the modified work as a whole. If identifiable sections of that work are not derived from the Library, and can be reasonably considered independent and separate works in themselves, then this License, and its terms, do not apply to those sections when you distribute them as separate works. But when you distribute the same sections as part of a whole which is a work based on the Library, the distribution of the whole must be on the terms of this License, whose permissions

for other licensees extend to the entire whole, and thus to each and every part regardless of who wrote it.

Thus, it is not the intent of this section to claim rights or contest your rights to work written entirely by you; rather, the intent is to exercise the right to control the distribution of derivative or collective works based on the Library.

In addition, mere aggregation of another work not based on the Library with the Library (or with a work based on the Library) on a volume of a storage or distribution medium does not bring the other work under the scope of this License.

3 You may opt to apply the terms of the ordinary GNU General Public License instead of this License to a given copy of the Library. To do this, you must alter all the notices that refer to this License, so that they refer to the ordinary GNU General Public License, version 2, instead of to this License. (If a newer version than version 2 of the ordinary GNU General Public License has appeared, then you can specify that version instead if you wish.) Do not make any other change in these notices.

Once this change is made in a given copy, it is irreversible for that copy, so the ordinary GNU

General Public License applies to all subsequent copies and derivative works made from that copy.

This option is useful when you wish to copy part of the code of the Library into a program that is not a library.

4 You may copy and distribute the Library (or a portion or derivative of it, under Section 2) in object code or executable form under the terms of Sections 1 and 2 above provided that you accompany it with the complete corresponding machine-readable source code, which must be distributed under the terms of Sections 1 and 2 above on a medium customarily used for software interchange.

If distribution of object code is made by offering access to copy from a designated place, then offering equivalent access to copy the source code from the same place satisfies the requirement to distribute the source code, even though third parties are not compelled to copy the source along with the object code.

5 A program that contains no derivative of any portion of the Library, but is designed to work with the Library by being compiled or linked with it, is called a "work that uses the Library". Such a work, in isolation, is not a derivative work of the Library, and therefore falls outside the scope of this License.

However, linking a "work that uses the Library" with the Library creates an executable that is a derivative of the Library (because it contains portions of the Library), rather than a "work that uses the library". The executable is therefore covered by this License. Section 6 states terms for distribution of such executables.

When a "work that uses the Library" uses material from a header file that is part of the Library, the object code for the work may be a derivative work of the Library even though the source code is not. Whether this is true is especially significant if the work can be linked without the Library, or if the work is itself a library. The threshold for this to be true is not precisely defined by law.

If such an object file uses only numerical parameters, data structure layouts and

accessors, and small macros and small inline functions (ten lines or less in length), then the use of the object file is unrestricted, regardless of whether it is legally a derivative work. (Executables containing this object code plus portions of the Library will still fall under Section 6.)

Otherwise, if the work is a derivative of the Library, you may distribute the object code for the work under the terms of Section 6. Any executables containing that work also fall under Section 6, whether or not they are linked directly with the Library itself.

6 As an exception to the Sections above, you may also compile or link a "work that uses the Library" with the Library to produce a work containing portions of the Library, and distribute that work under terms of your choice, provided that the terms permit modification of the work for the customer's own use and reverse engineering for debugging such modifications.

You must give prominent notice with each copy of the work that the Library is used in it and that the Library and its use are covered by this License. You must supply a copy of this License. If the work during execution displays copyright notices, you must include the copyright notice for the Library among them, as well as a reference directing the user to the copy of this License. Also, you must do one of these things:

a) Accompany the work with the complete corresponding machine-readable source code for the Library including whatever changes were used in the work (which must be distributed under Sections 1 and 2 above); and, if the work is an executable linked with the Library, with the complete machine-readable "work that uses the Library", as object code and/or source code, so that the user can modify the Library and then relink to produce a modified executable containing the modified Library. (It is understood that the user who changes the contents of definitions files in the Library will not necessarily be able to recompile the application to use the modified definitions.)

b) Accompany the work with a written offer, valid for at least three years, to give the same user the materials specified in Subsection 6a, above, for a charge no more than the cost of performing this distribution.

c) If distribution of the work is made by offering access to copy from a designated place, offer

equivalent access to copy the above specified materials from the same place.

d) Verify that the user has already received a copy of these materials or that you have already sent this user a copy.

For an executable, the required form of the "work that uses the Library" must include any data and utility programs needed for reproducing the executable from it. However, as a special exception, the source code distributed need not include anything that is normally distributed (in either source or binary form) with the major components (compiler, kernel, and so on) of the operating system on which the executable runs, unless that component itself accompanies the executable.

It may happen that this requirement contradicts the license restrictions of other proprietary libraries that do not normally accompany the operating system. Such a contradiction means you cannot use both them and the Library together in an executable that you distribute.

7 You may place library facilities that are a work based on the Library side-by-side in a single library together with other library facilities not covered by this License, and distribute such a combined library, provided that the separate distribution of the work based on the Library and of the other library facilities is otherwise permitted, and provided that you do these two things:

a) Accompany the combined library with a copy of the same work based on the Library, uncombined with any other library facilities. This must be distributed under the terms of the Sections above.

b) Give prominent notice with the combined library of the fact that part of it is a work based on the Library, and explaining where to find the accompanying uncombined form of the same work.

8 You may not copy, modify, sublicense, link with, or distribute the Library except as expressly provided under this License. Any attempt otherwise to copy, modify, sublicense, link with, or distribute the Library is void, and will automatically terminate your rights under this License. However, parties who have received copies, or rights, from you under this License will not have their licenses terminated so long as such parties remain in full compliance.

9 You are not required to accept this License, since you have not signed it. However, nothing else grants you permission to modify or distribute the Library or its derivative works. These actions are prohibited by law if you do not accept this License. Therefore, by modifying or distributing the Library (or any work based on the Library), you indicate your acceptance of this License to do so, and all its terms and conditions for copying, distributing or modifying the Library or works based on it.

10 Each time you redistribute the Library (or any work based on the Library), the recipient automatically receives a license from the original licensor to copy, distribute, link with or modify the Library subject to these terms and conditions. You may not impose any further restrictions on the recipients' exercise of the rights granted herein. You are not responsible for enforcing compliance by third parties to this License.

11 If, as a consequence of a court judgment or allegation of patent infringement or for any other reason (not limited to patent issues), conditions are imposed on you (whether by court order, agreement or otherwise) that contradict the conditions of this License, they do not excuse you from the conditions of this License. If you cannot distribute so as to satisfy simultaneously your obligations under this License and any other pertinent obligations, then as a consequence you may not distribute the Library at all. For example, if a patent license would not permit royalty-free redistribution of the Library by all those who receive copies directly or indirectly through you, then the only way you could satisfy both it and this License would be to refrain entirely from distribution of the Library.

If any portion of this section is held invalid or unenforceable under any particular circumstance, the balance of the section is intended to apply, and the section as a whole is intended to apply in other circumstances.

It is not the purpose of this section to induce you to infringe any patents or other property right claims or to contest validity of any such claims; this section has the sole purpose of protecting the integrity of the free software distribution system which is implemented by public license practices. Many people have made generous

contributions to the wide range of software distributed through that system in reliance on consistent application of that system; it is up to the author/donor to decide if he or she is willing to distribute software through any other system and a licensee cannot impose that choice.

This section is intended to make thoroughly clear what is believed to be a consequence of the rest of this License.

12 If the distribution and/or use of the Library is restricted in certain countries either by patents or by copyrighted interfaces, the original copyright holder who places the Library under this License may add an explicit geographical distribution limitation excluding those countries, so that distribution is permitted only in or among countries not thus excluded. In such case, this License incorporates the limitation as if written in the body of this License.

13 The Free Software Foundation may publish revised and/or new versions of the Library General Public License from time to time. Such new versions will be similar in spirit to the present version, but may differ in detail to address new problems or concerns.

Each version is given a distinguishing version number. If the Library specifies a version number of this License which applies to it and "any later version", you have the option of following the terms and conditions either of that version or of any later version published by the Free Software Foundation. If the Library does not specify a license version number, you may choose any version ever published by the Free Software Foundation.

14 If you wish to incorporate parts of the Library into other free programs whose distribution conditions are incompatible with these, write to the author to ask for permission. For software which is copyrighted by the Free Software Foundation, write to the Free Software Foundation; we sometimes make exceptions for this. Our decision will be guided by the two goals of preserving the free status of all derivatives of our free software and of promoting the sharing and reuse of software generally.

NO WARRANTY

15 BECAUSE THE LIBRARY IS LICENSED FREE OF CHARGE, THERE IS NO WARRANTY FOR THE LIBRARY, TO THE EXTENT PERMITTED BY APPLICABLE LAW. EXCEPT WHEN OTHERWISE STATED IN WRITING THE COPYRIGHT HOLDERS AND/OR OTHER PARTIES PROVIDE THE LIBRARY "AS IS" WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESSED OR IMPLIED, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE. THE ENTIRE RISK AS TO THE QUALITY AND PERFORMANCE OF THE LIBRARY IS WITH YOU. SHOULD THE LIBRARY PROVE DEFECTIVE, YOU ASSUME THE COST OF ALL NECESSARY SERVICING, REPAIR OR CORRECTION.

16 IN NO EVENT UNLESS REQUIRED BY APPLICABLE LAW OR AGREED TO IN WRITING WILL ANY COPYRIGHT HOLDER, OR ANY OTHER PARTY WHO MAY MODIFY AND/OR REDISTRIBUTE THE LIBRARY AS PERMITTED ABOVE, BE LIABLE TO YOU FOR DAMAGES, INCLUDING ANY GENERAL, SPECIAL, INCIDENTAL OR CONSEQUENTIAL DAMAGES ARISING OUT OF THE USE OR INABILITY TO USE THE LIBRARY (INCLUDING BUT NOT LIMITED TO LOSS OF DATA OR DATA BEING RENDERED INACCURATE OR LOSSES SUSTAINED BY YOU OR THIRD PARTIES OR A FAILURE OF THE LIBRARY TO

OPERATE WITH ANY OTHER SOFTWARE), EVEN IF SUCH HOLDER OR OTHER PARTY HAS BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGES.

How to Apply These Terms to Your New Libraries

If you develop a new library, and you want it to be of the greatest possible use to the public, we recommend

making it free software that everyone can redistribute and change. You can do so by permitting

redistribution under these terms (or, alternatively, under the terms of the ordinary General Public License).

To apply these terms, attach the following notices to the library. It is safest to attach them to the start of each source file to most effectively convey the exclusion of warranty; and each file should have at least the "copyright" line and a pointer to where the full notice is found.

One line to give the library's name and an idea of what it does.

Copyright © *year name of author*

This library is free software; you can redistribute it and/or modify it under the terms of the GNU Library General PublicLicense as published by the Free Software Foundation; either version 2 of the License, or (at your option) any later version.

This library is distributed in the hope that it will be useful, but WITHOUT ANY WARRANTY; without even the implied warranty of MERCHANTABILITY or FITNESS FOR A PARTICULAR PURPOSE. See the GNU Library General Public License for more details.

You should have received a copy of the GNU Library General PublicLicense along with this library; if not, write to the Free Software Foundation, Inc., 59 Temple Place - Suite 330, Boston, MA 02111-1307, USA.

Also add information on how to contact you by electronic and paper mail.

You should also get your employer (if you work as a programmer) or your school, if any, to sign a "copyright disclaimer" for the library, if necessary. Here is a sample; alter the names:

Yoyodyne, Inc., hereby disclaims all copyright interest in the library 'Frob' (a library for tweaking knobs) written by James Random Hacker.

signature of Ty Coon, 1 April 1990 Ty Coon, President of Vice

That's all there is to it!

SOURCE CODE Version 1.2D

AGREEMENT PLEASE READ THIS AGREEMENT CAREFULLY. By accessing and using the ***Source Code***, you accept this Agreement in its entirety and agree to only use the ***Source Code*** in accordance with the following terms and conditions. If you do not wish to be bound by these terms and conditions, do not access or use the ***Source Code***.

A12.3 YOUR REPRESENTATIONS

A12.3.1 You represent and warrant that:

a. If you are an entity, or an individual other than the person accepting this Agreement, the person accepting this Agreement on your behalf is your legally authorized representative, duly authorized to accept agreements of this type on your behalf and obligate you to comply with its provisions;

b. You have read and fully understand this Agreement in its entirety;

c. Your ***Build Materials*** are either original or do not include any ***Software*** obtained under a license that conflicts with the obligations contained in this Agreement;

d. To the best of your knowledge, your ***Build Materials*** do not infringe or misappropriate the rights of any person or entity; and,

e. You will regularly monitor the ***Website*** for any notices.

A12.3.2 DEFINITIONS AND INTERPRETATION

1 For purposes of this Agreement, certain terms have been defined below and elsewhere in this Agreement to encompass meanings that may differ from, or be in addition to, the normal connotation of the defined word.

a. "***Additional Code***" means ***Software*** in source code form which does not contain

i any of the ***Source Code***, or

ii derivative work (such term having the same meaning in this Agreement as under U.S. Copyright Law) of the ***Source Code***.

b. "***AT&T Patent Claims***" means those claims of patents (i) owned by AT&T and (ii) licensable without restriction or obligation, which, absent a license, are necessarily and unavoidably infringed by the use of the functionality of the ***Source Code***.

c. "***Build Materials***" means, with reference to a ***Derived Product***, the ***Patch*** and ***Additional Code***, if any, used in the preparation of such ***Derived Product***, together with written instructions that describe, in reasonable detail, such preparation.

d. "***Capsule***" means a computer file containing the exact same contents as a computer file downloaded from the ***Website***.

e. "***Derived Product***" means a ***Software Product*** which is a derivative work of the ***Source Code***.

f. "***IPR***" means all rights protectable under intellectual property law anywhere throughout the world, including rights protectable under patent, copyright and trade secret laws, but not trademark rights.

g. "***Patch***" means ***Software*** for ***changing*** all or any portion of the ***Source Code***.

h. "***Proprietary Notice***" means the following statement:

"This product contains certain software code or other information ("AT&T Software") proprietary to AT&T Corp. ("AT&T"). The AT&T Software is provided to you "AS IS". YOU ASSUME TOTAL RESPONSIBILITY AND RISK FOR USE OF THE AT&T SOFTWARE. AT&T DOES NOT MAKE, AND EXPRESSLY DISCLAIMS, ANY EXPRESS OR IMPLIED WARRANTIES OF ANY KIND WHATSOEVER, INCLUDING, WITHOUT LIMITATION, THE IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE, WARRANTIES OF TITLE OR NON-INFRINGEMENT OF ANY INTELLECTUAL PROPERTY RIGHTS, ANY WARRANTIES ARISING BY USAGE OF TRADE, COURSE OF DEALING OR COURSE OF PERFORMANCE, OR ANY WARRANTY THAT THE AT&T SOFTWARE IS "ERROR FREE" OR WILL MEET YOUR REQUIREMENTS.

Unless you accept a license to use the AT&T Software, you shall not reverse compile, disassemble or otherwise reverse engineer this product to ascertain the source code for any AT&T Software.

© AT&T Corp. All rights reserved. AT&T is a registered trademark of AT&T Corp."

i. "***Software***" means, as the context may require, source or object code instructions for controlling the operation of a central processing unit or computer, and computer files containing data or text.

j. "***Software Product***" means a collection of computer files containing ***Software*** in object code form only, which, taken together, reasonably comprise a product, regardless of whether such product is intended for internal use or commercial exploitation. A single computer file can comprise a ***Software Product***.

k. "***Source Code***" means the ***Software*** contained in compressed form in the ***Capsule***.

l. "***Website***" means the Internet website having the URL *http://www.research.att.com/sw/ download/*. AT&T may ***change*** the content or URL of the ***Website***, or remove it from the Internet altogether.

2 By way of clarification only, the terms ***Capsule, Proprietary Notice*** and ***Source Code*** when used in this Agreement shall mean the materials and information defined by such terms without any change, enhancement, amendment, alteration or modification (collectively, "***change***").

A12.3.3 *GRANT OF RIGHTS*

1 Subject to third party intellectual property claims, if any, and the terms and conditions of this Agreement, AT&T grants to you under:

a. the ***AT&T Patent Claims*** and AT&T's copyright rights in the ***Source Code***, a non-exclusive, fully paid-up license to:

i Reproduce and distribute the ***Capsule***;

ii Display, perform, use, and compile the ***Source Code*** and execute the resultant binary ***Software*** on a computer;

iii Prepare a ***Derived Product*** solely by compiling ***Additional Code***, if any, together with the code resulting from operating a ***Patch*** on the ***Source Code***; and,

iv Execute on a computer and distribute to others ***Derived Products***,

except that, with respect to the ***AT&T Patent Claims***, the license rights granted in clauses (iii) and (iv) above shall only extend, and be limited, to that portion of a ***Derived Product*** which is ***Software*** compiled from some portion of the ***Source Code***; and,

b. AT&T's copyright rights in the ***Source Code***, a non-exclusive, fully paid-up license to prepare and distribute ***Patches*** for the ***Source Code***.

2 Subject to the terms and conditions of this Agreement, you may create a hyperlink between an Internet website owned and controlled by you and the ***Website***, which hyperlink describes in a fair and good faith manner where the ***Capsule*** and ***Source Code*** may be obtained, provided that, you do not frame the ***Website*** or otherwise give the false impression that AT&T is somehow associated with, or otherwise endorses or sponsors your website. Any goodwill associated with such hyperlink shall inure to the sole benefit of AT&T. Other than the creation of such hyperlink, nothing in this Agreement shall be construed as conferring upon you any right to use any reference to AT&T, its trade names, trademarks, service marks or any other indicia of origin owned by AT&T, or to indicate that your products or services are in any way sponsored, approved or endorsed by, or affiliated with, AT&T.

3 Except as expressly set forth in Section 3.1 above, no other rights or licenses under any of AT&T's ***IPR*** are granted or, by implication, estoppels or otherwise, conferred. By way of example only, no rights or licenses under any of AT&T's patents are granted or, by implication, estoppels or otherwise, conferred with respect to any portion of a ***Derived Product*** which is not ***Software*** compiled from some portion, without ***change***, of the ***Source Code***.

A12.3.4 *YOUR OBLIGATIONS*

1 If you distribute ***Build Materials*** (including if you are required to do so pursuant to this Agreement), you shall ensure that the recipient enters into and duly accepts an agreement with you which includes the minimum terms set forth in Appendix A (completed to indicate you as the LICENSOR) and no other provisions which, in AT&T's opinion, conflict with your obligations under, or the intent of, this Agreement. The agreement required under this Section 4.1 may be in electronic form and may be distributed with the ***Build Materials*** in a form such that the recipient accepts the agreement by using or installing the ***Build Materials***. If any ***Additional Code*** contained in your ***Build Materials*** includes ***Software*** you obtained under license, the agreement shall also include complete details concerning the license and any restrictions or obligations associated with such ***Software***.

2 If you prepare a ***Patch*** which you distribute to anyone else you shall:

a. Contact AT&T, as may be provided on the ***Website*** or in a text file included with the ***Source Code***, and describe for AT&T such ***Patch*** and provide AT&T with a copy of such ***Patch*** as directed by AT&T; or,

b. Where you make your ***Patch*** generally available on your Internet

website, you shall provide AT&T with the URL of your website and hereby grant to AT&T a non-exclusive, fully-paid up right to create a hyperlink between your website and a page associated with the ***Website***.

3 If you prepare a ***Derived Product***, such product shall conspicuously display to users, and any corresponding documentation and license agreement shall include as a provision, the ***Proprietary Notice***.

A12.3.5 *YOUR GRANT OF RIGHTS TO AT&T*

1 You grant to AT&T under any ***IPR*** owned or licensable by you which in any way relates to your ***Patches***, a non-exclusive, perpetual, worldwide, fully paid-up, unrestricted, irrevocable license, along with the right to sublicense others, to (a) make, have made, use, offer to sell, sell and import any products, services or any combination of products or services, and (b) reproduce, distribute, prepare derivative works based on, perform, display and transmit your Patches in any media whether now known or in the future developed.

A12.3.6 *AS IS CLAUSE / LIMITATION OF LIABILITY*

1 The ***Source Code*** and ***Capsule*** are provided to you "AS IS". YOU ASSUME TOTAL RESPONSIBILITY AND RISK FOR YOUR USE OF THEM INCLUDING THE RISK OF ANY DEFECTS OR INACCURACIES THEREIN. AT&T DOES NOT MAKE, AND EXPRESSLY DISCLAIMS, ANY EXPRESS OR IMPLIED WARRANTIES OF ANY KIND WHATSOEVER, INCLUDING, WITHOUT LIMITATION, THE IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE, WARRANTIES OF TITLE OR NON-INFRINGEMENT OF ANY ***IPR*** OR TRADEMARK RIGHTS, ANY WARRANTIES ARISING BY USAGE OF TRADE, COURSE OF DEALING OR COURSE OF PERFORMANCE, OR ANY WARRANTY THAT THE ***SOURCE CODE*** OR ***CAPSULE*** ARE "ERROR FREE" OR WILL MEET YOUR REQUIREMENTS.

2 IN NO EVENT SHALL AT&T BE LIABLE FOR (a) ANY INCIDENTAL, CONSEQUENTIAL, OR INDIRECT DAMAGES (INCLUDING, WITHOUT LIMITATION, DAMAGES FOR LOSS OF PROFITS, BUSINESS INTERRUPTION, LOSS OF PROGRAMS OR INFORMATION, AND THE LIKE) ARISING OUT OF THE USE OF OR INABILITY TO USE THE ***SOURCE CODE*** OR ***CAPSULE***, EVEN IF AT&T OR ANY OF ITS AUTHORIZED REPRESENTATIVES HAS BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGES, (b) ANY CLAIM ATTRIBUTABLE TO ERRORS, OMISSIONS, OR OTHER INACCURACIES IN THE ***SOURCE CODE*** OR ***CAPSULE***, OR (c) ANY CLAIM BY ANY THIRD PARTY.

3 BECAUSE SOME STATES DO NOT ALLOW THE EXCLUSION OR LIMITATION OF LIABILITY FOR CONSEQUENTIAL OR INCIDENTAL DAMAGES, THE ABOVE LIMITATIONS MAY NOT APPLY TO YOU. IN THE EVENT THAT APPLICABLE LAW DOES NOT ALLOW THE COMPLETE EXCLUSION OR LIMITATION OF LIABILITY OF CLAIMS AND DAMAGES AS SET FORTH IN THIS AGREEMENT, AT&T'S LIABILITY IS LIMITED TO THE GREATEST EXTENT PERMITTED BY LAW.

A12.3.7 *INDEMNIFICATION*

1 You shall indemnify and hold harmless AT&T, its affiliates and authorized

representatives against any claims, suits or proceedings asserted or commenced by any third party and arising out of, or relating to, your use of the ***Source Code***. This obligation shall include indemnifying against all damages, losses, costs and expenses (including attorneys' fees) incurred by AT&T, its affiliates and authorized representatives as a result of any such claims, suits or proceedings, including any costs or expenses incurred in defending against any such claims, suits, or proceedings.

A12.3.8 *GENERAL*

1 You shall not assert against AT&T, its affiliates or authorized representatives any claim for infringement or misappropriation of any ***IPR*** or trademark rights in any way relating to the ***Source Code***, including any such claims relating to any ***Patches***.

2 In the event that any provision of this Agreement is deemed illegal or unenforceable, AT&T may, but is not obligated to, post on the ***Website*** a new version of this Agreement which, in AT&T's opinion, reasonably preserves the intent of this Agreement.

3 Your rights and license (but not any of your obligations) under this Agreement shall terminate automatically in the event that (a) notice of a non-frivolous claim by a third party relating to the ***Source Code*** or ***Capsule*** is posted on the ***Website***, (b) you have knowledge of any such claim, (c) any of your representations or warranties in Article 1.0 or Section 8.4 are false or inaccurate, (d) you exceed the rights and license granted to you or (e) you fail to fully comply with any provision of this Agreement. Nothing in this provision shall be construed to restrict you, at your option and subject to applicable law, from replacing the portion of the Source Code that is the subject of a claim by a third party with non-infringing code or from independently negotiating for necessary rights from the third party.

4 You acknowledge that the ***Source Code*** and ***Capsule*** may be subject to U.S. export laws and regulations, and, accordingly, you hereby assure AT&T that you will not, directly or indirectly, violate any applicable U.S. laws and regulations.

5 Without limiting any of AT&T's rights under this Agreement or at law or in equity, or otherwise expanding the scope of the license and rights granted hereunder, if you fail to perform any of your obligations under this Agreement with respect to any of your ***Patches*** or ***Derived Products***, or if you do any act which exceeds the scope of the license and rights granted herein, then such ***Patches***, ***Derived Products*** and acts are not licensed or otherwise authorized under this Agreement and such failure shall also be deemed a breach of this Agreement. In addition to all other relief available to it for any breach of your obligations under this Agreement, AT&T shall be entitled to an injunction requiring you to perform such obligations.

6 This Agreement shall be governed by and construed in accordance with the laws of the State of New York, USA, without regard to its conflicts of law rules. This Agreement shall be fairly interpreted in accordance with its terms and without any strict construction in favor of or against either AT&T or you. Any suit or proceeding you bring relating to this Agreement shall be brought and prosecuted only in New York, New York,USA

A12.4 Other Copyrights

Portions of code may be covered by the following copyrights:

A12.4.1 X Consortium

Copyright © 1996 X Consortium

Permission is hereby granted, free of charge, to any person obtaining a copy of this software and associated documentation files (the "Software"), to deal in the Software without restriction, including without limitation the rights to use, copy, modify, merge, publish, distribute, sublicense, and/or sell copies of the Software, and to permit persons to whom the Software is furnished to do so, subject to the following conditions:

The above copyright notice and this permission notice shall be included in all copies or substantial portions of the Software.

THE SOFTWARE IS PROVIDED "AS IS", WITHOUT WARRANTY OF ANY KIND, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO THE WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE AND NONINFRINGEMENT. IN NO EVENT SHALL THE X CONSORTIUM BE LIABLE FOR ANY CLAIM, DAMAGES OR OTHER LIABILITY, WHETHER IN AN ACTION OF CONTRACT, TORT OR OTHERWISE, ARISING FROM, OUT OF OR IN CONNECTION WITH THE SOFTWARE OR THE USE OR OTHER DEALINGS IN THE SOFTWARE.

Except as contained in this notice, the name of the X Consortium shall not be used in advertising or otherwise to promote the sale, use or other dealings in this Software without prior written authorization from the X Consortium.

X Window System is a trademark of X Consortium, Inc.

A12.4.2 Berkeley-based copyrights:

General

Redistribution and use in source and binary forms, with or without modification, are permitted provided that the following conditions are met:

1 Redistributions of source code must retain the above copyright notice, this list of conditions and the following disclaimer.

2 Redistributions in binary form must reproduce the above copyright notice, this list of conditions and the following disclaimer in the documentation and/or other materials provided with the distribution.

3 The name of the author may not be used to endorse or promote products derived from this software without specific prior written permission.

THIS SOFTWARE IS PROVIDED BY THE AUTHOR "AS IS" AND ANY EXPRESS OR IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE ARE DISCLAIMED.

IN NO EVENT SHALL THE AUTHOR BE LIABLE FOR ANY DIRECT, INDIRECT, INCIDENTAL, SPECIAL, EXEMPLARY, OR CONSEQUENTIAL DAMAGES (INCLUDING, BUT NOT LIMITED TO, PROCUREMENT OF SUBSTITUTE GOODS OR SERVICES; LOSS OF USE, DATA, OR PROFITS; OR BUSINESS INTERRUPTION) HOWEVER CAUSED AND ON ANY THEORY OF LIABILITY, WHETHER IN CONTRACT, STRICT LIABILITY, OR TORT (INCLUDING NEGLIGENCE OR OTHERWISE) ARISING IN ANY WAY OUT OF THE USE OF THIS SOFTWARE, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGE.

A12.4.3 *UCB/LBL*

Copyright © 1993 The Regents of the University of California. All rights reserved.

This software was developed by the Computer Systems Engineering group at Lawrence Berkeley Laboratory under DARPA contract BG 91-66 and contributed to Berkeley.

All advertising materials mentioning features or use of this software must display the following acknowledgement: This product includes software developed by the University of California, Lawrence Berkeley Laboratory.

Redistribution and use in source and binary forms, with or without modification, are permitted provided that the following conditions are met:

1 Redistributions of source code must retain the above copyright notice, this list of conditions and the following disclaimer.

2 Redistributions in binary form must reproduce the above copyright notice, this list of conditions and the following disclaimer in the documentation and/or other materials provided with the distribution.

3 All advertising materials mentioning features or use of this software must display the following

acknowledgement: This product includes software developed by the University of California, Berkeley and its contributors.

4 Neither the name of the University nor the names of its contributors may be used to endorse or promote products derived from this software without specific prior written permission.

THIS SOFTWARE IS PROVIDED BY THE REGENTS AND CONTRIBUTORS “AS IS” AND ANY EXPRESS OR IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE ARE DISCLAIMED. IN NO EVENT SHALL THE REGENTS OR CONTRIBUTORS BE LIABLE FOR ANY DIRECT, INDIRECT, INCIDENTAL, SPECIAL, EXEMPLARY, OR CONSEQUENTIAL DAMAGES (INCLUDING, BUT NOT LIMITED TO, PROCUREMENT OF SUBSTITUTE GOODS OR SERVICES; LOSS OF USE, DATA, OR PROFITS; OR BUSINESS INTERRUPTION) HOWEVER CAUSED AND ON ANY THEORY OF LIABILITY, WHETHER IN CONTRACT, STRICT LIABILITY, OR TORT (INCLUDING NEGLIGENCE OR OTHERWISE) ARISING IN ANY WAY OUT OF THE USE OF THIS SOFTWARE, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGE.

The “Artistic License” **Preamble**

The intent of this document is to state the conditions under which a Package may be copied, such that the Copyright Holder maintains some semblance of artistic control over the development of the package, while giving the users of the package the right to use and distribute the Package in a more-or-less customary fashion, plus the right to make reasonable modifications.

A12.5 Definitions:

"Package" refers to the collection of files distributed by the Copyright Holder, and derivatives of that collection of files created through textual modification.

"Standard Version" refers to such a Package if it has not been modified, or has been modified in accordance with the wishes of the Copyright Holder as specified below.

"Copyright Holder" is whoever is named in the copyright or copyrights for the package.

"You" is you, if you're thinking about copying or distributing this Package.

"Reasonable copying fee" is whatever you can justify on the basis of media cost, duplication charges, time of people involved, and so on. (You will not be required to justify it to the Copyright Holder, but only to the computing community at large as a market that must bear the fee.)

"Freely Available" means that no fee is charged for the item itself, though there may be fees involved in handling the item. It also means that recipients of the item may redistribute it under the same conditions they received it.

1 You may make and give away verbatim copies of the source form of the Standard Version of this Package without restriction, provided that you duplicate all of the original copyright notices and associated disclaimers.

2 You may apply bug fixes, portability fixes and other modifications derived from the Public Domain or from the Copyright Holder. A Package modified in such a way shall still be considered the Standard Version.

3 You may otherwise modify your copy of this Package in any way, provided that you insert a prominent notice in each changed file stating how and when you changed that file, and provided that you do at least ONE of the following:

a. place your modifications in the Public Domain or otherwise make them Freely Available, such as by posting said modifications to Usenet or an equivalent medium, or placing the modifications on a major archive site such as uunet.uu.net, or by allowing the Copyright Holder to include your modifications in the Standard Version of the Package.

b. use the modified Package only within your corporation or organization.

c. rename any non-standard executables so the names do not conflict with standard executables, which must also be provided, and provide a separate manual page for each non-standard executable that clearly documents how it differs from the Standard Version.

d. make other distribution arrangements with the Copyright Holder.

4 You may distribute the programs of this Package in object code or executable form, provided that you do at least ONE of the following:

a. distribute a Standard Version of the executables and library files, together with instructions (in the manual page or equivalent) on where to get the Standard Version.

b. accompany the distribution with the machine-readable source of the Package with your modifications.

c. give non-standard executables non-standard names, and clearly document the differences in manual pages (or equivalent), together with instructions on where to get the Standard Version.

d. make other distribution arrangements with the Copyright Holder.

5 You may charge a reasonable copying fee for any distribution of this Package. You may charge any fee you choose for support of this Package. You may not charge a fee for this Package itself. However, you may distribute this Package in aggregate with other (possibly commercial) programs as part of a larger (possibly commercial) software distribution provided that you do not advertise this Package as a product of your own. You may embed this Package's interpreter within an executable of yours (by linking); this shall be construed as a mere form of aggregation, provided that the complete Standard Version of the interpreter is so embedded.

6 The scripts and library files supplied as input to or produced as output from the programs of this Package do not automatically fall under the copyright of this Package, but belong to whoever generated them, and may be sold commercially, and may be aggregated with this Package. If such scripts or library files are aggregated with this Package via the so-called "undump" or "unexec" methods of producing a binary executable image, then distribution of such an image shall neither be construed as a distribution of this Package nor shall it fall under the restrictions of Paragraphs 3 and 4, provided that you do not represent such an executable image as a Standard Version of this Package.

7 Subroutines (or comparably compiled subroutines in other languages) supplied by you and linked into this Package in order to emulate subroutines and variables of the language defined by this Package shall not be considered part of this Package, but are the equivalent of input as in Paragraph 6, provided these subroutines do not change the language in any way that would cause it to fail the regression tests for the language.

8 Aggregation of this Package with a commercial distribution is always permitted provided that the use of this Package is embedded; that is, when no overt attempt is made to make this Package's interfaces visible to the end user of the commercial distribution. Such use shall not be construed as a distribution of this Package.

9 The name of the Copyright Holder may not be used to endorse or promote products derived from this software without specific prior written permission.

10 THIS PACKAGE IS PROVIDED "AS IS" AND WITHOUT ANY EXPRESS OR IMPLIED WARRANTIES, INCLUDING, WITHOUT LIMITATION, THE IMPLIED WARRANTIES OF MERCHANTIBILITY AND FITNESS FOR A PARTICULAR PURPOSE

Appendix 13

Glossary

A13.1 Glossary

Access Control List (ACL)

A list that controls which hosts have access to which volumes. When a host attempts to log in, the *Storage Concentrator* requests authorization and allows access based on the list.

Access Path

The route used by a computer to communicate with a storage device. The path accesses the host bus adapter, host storage interconnection controller and logical unit. Some configurations support multiple access paths to a single device.

Administrative Interface

The graphical administrative interface is accessed from a computer on the network via a web browser. The following functions are available through the interface: *Storage Concentrator* discovery, physical resource management, storage pool management, logical volume management, target management, Access Control List management (ACL), system management, session management.

Adaptive Load Balancing

A feature that allows for more than one networking path providing port FailOver protection, as well as increased bandwidth. If one port goes down, the other port automatically accepts the additional load. There is no interruption in server operation, and a network alert is generated to inform IT staff of the problem.

Application Programming Interface (API)

The interface used by an application program to request services; usually denotes interfaces between applications and the software components that comprise the operating environment (e.g., operating system, file system, volume manager, and device drivers).

Array or Array Configuration

A storage array, i.e., a disk array or RAID array.

1. Assignment of the disks and operating parameters for a disk array. Disk array configuration includes designating the member disks or extents of the array and the order in which they are to be used, as well as setting parameters such as stripe depth, RAID model, cache allowance, spare disk assignments.
2. The arrangement of disks and operating parameters that result from such an assignment.

Asynchronous I/O Operation

An I/O operation whose initiator does not await its completion before proceeding with other work. Asynchronous I/O operations enable an initiator to have multiple

concurrent I/O operations in progress.

Asynchronous Mirroring

Transactions that maintain synchronization for logical volumes occur in a batch rather than real time mode. An asynchronous mirroring is desired when mirrored logical volumes are separated by distance in order to reduce the effects of distance-induced latency. It is also desirable to use Asynchronous Mirroring or replication to reduce the bandwidth requirements of the network connection.

Backup

1 (noun) A collection of data stored on non-volatile (usually removable) storage media for purposes of recovery in case the original copy of data is lost or becomes inaccessible. The data is also called the backup copy. To be useful for recovery, a backup must be made by copying the source data image when it is in a consistent state.

2 (verb) The act of creating a backup.

Campus Mirrors

Mirror images that are behind the Secondary *Storage Concentrator*

CHAP

CHAP (Challenge Handshake Authentication Protocol) allows you to set a Password or "Secret" for as a gatekeeper for communication between a host initiator and a volume.

Clusters

You must have at least two *Storage Concentrators* to implement FailOver on an IP Storage. Multiple *Storage Concentrators* are known as a cluster. In a clustered pair there will be one Primary and one Secondary *Storage Concentrator*. A cluster appears as a single entity to hosts on the network.

Concatenation

A logical joining of two series of data, usually represented by the symbol "|". In data communications, two or more data are often concatenated to provide a unique name or reference (e.g., S_ID | X_ID). Volume managers concatenate disk address spaces to present a single larger address space.

Configuration

1 The process of installing or removing hardware or software components required for a system or subsystem to function.

2 Assignment of the operating parameters of a system, subsystem, or device. For example, disk array configuration includes designating the member disks or extents for the array, as well as setting parameters such as stripe depth, RAID model, and cache allowance.

3 The collection of hardware and software components and operating

parameters for an operating system.

Controller

The control logic in a storage subsystem that performs command transformation and routing, aggregation (RAID, mirroring, striping, or other aggregation), high-level error recovery, and performance optimization for multiple storage devices.

Controller-based Array

A disk array whose control software executes in a disk subsystem controller. The member disks of a controller-based array are necessarily part of the same disk subsystem that includes the controller.

Database Management System (DBMS)

A set of computer programs with a user and/or programming interface that supports defining the format of a database, and creating and accessing the data. A database management system removes the need for a user or program to manage low-level database storage. It also provides security and assures the integrity of the data it contains. Database management systems may be relational (table-oriented) or object oriented.

Data Transfer Rate

The amount of data per unit time moved across an I/O bus while executing an I/O load. For any I/O load, the data transfer capacity of an I/O subsystem is limited by its data transfer rate. For disk subsystem I/O, data transfer rate is usually expressed in MB/second (millions of bytes per second where 1 million is 10^6).

Detach Image

Detaching an image allows it to be mounted and used by other software applications. The most common usage is to make a backup copy of the detached image. The detached image retains the mirror volume information and can be rejoined at any time.

Disk Array

A set of disks from one or more commonly accessible disk subsystems, combined with a body of control software. The control software presents the storage capacity of the disk to hosts as one or more logical disks. When it runs on a disk controller, control software is often called firmware or microcode. Control software that runs in a host computer is usually called a volume manager.

A disk subsystem which includes control software with the capability to organize disks as disk arrays.

Disk Array Subsystem Disk Drive

A non-volatile, randomly addressable, writable data storage device. Subdivision of a disk drive, disk array, or RAID array.

Disk Partitions Distributed Lock Manager (DLM)

Software, hardware or a combination of hardware and software that prevents multiple writers from altering a data element simultaneously or in a fashion that would lead to data corruption.

Failed Over

A mode of operation for failure-tolerant systems in which a component has failed and its function has been assumed by a redundant component. A system operating in a failed-over mode that protects against single failures is not failure tolerant, since failure of the redundant component may render the system unable to function. Some systems (e.g., clusters) are able to tolerate more than one failure; these remain failure tolerant until no redundant component is available to protect against further failures.

FailOver

The automatic substitution of a functionally equivalent system component for a failed one. FailOver automatically redirects user requests from the failed or down system to the backup system that takes over the operations of the primary system.

File Server

A computer whose primary purpose is to serve files to clients. A file server is a general-purpose computer capable of hosting additional applications or capable of only serving files. The server is also called a host.

File System

Software that imposes structure on the address space of one or more physical or logical disks so that applications may deal more conveniently with abstract-named data objects of variable size (files). File systems are often supplied as operating system components, but are implemented and marketed as independent software components.

Graphical User Interface (GUI)

A user interface for intelligent devices that is characterized by pictorial displays and highly structured forms-oriented input.

Hardware-based

Functionality implemented in high-speed physical (digital) hardware components, such as logic gates, inside high-density field- programmable gate arrays (FPGAs) or application-specific integrated circuits (ASICs).

Host

A computer connected to storage; typically a server running applications or providing services that access and consume storage.

Host Bus Adapter (HBA)

An I/O adapter that connects a host I/O bus to the memory system of a computer.

Image

See Mirror Image.

Input/Output (I/O)

The process of moving data between the main memory of a computer system and an external device or interface such as a storage device, display, printer, or network connected to other computer systems. I/O is a collective term for reading, or moving data into a computer system's memory, and writing, or moving data from a computer system's memory to another location.

Initiator

SCSI device (usually a host system) that requests an operation to be performed by another SCSI device, the target.

IOPS

Input/Output Per Second. It is the number of inputs and outputs or read/writes per second.

iSCSI

iSCSI is a protocol that enables the transmission of block-level SCSI data between storage devices and computers over a standard IP network. iSCSI combines Ethernet based IP networking with the SCSI command set, the core command set used in all storage configurations.

iSCSI Client

A logical entity, typically a host, which includes at least one iSCSI Initiator.

iSCSI Initiator

A logical entity, typically within a host, that sends (iSCSI) SCSI commands to targets to be executed.

iSCSI Server

A logical entity, typically a storage controller or gateway, which includes at least one iSCSI Target.

iSCSI Target

A group of physical storage devices containing at least 1 Logical Unit Number (LUN).

iSNS

iSNS (Internet Storage Naming Service) is a discovery protocol that facilitates automated discovery, management and configuration of iSCSI devices on a TCP/IP network. In any storage network, hosts (initiators) need to know which storage resources (or targets) they can access. An Internet storage name server lets servers automatically identify and connect to authorized storage resources.

JBOD

Acronym for "Just a Bunch Of Disks." Originally used to mean a collection of disks without the coordinated control provided by control software; today the term JBOD most often refers to a cabinet of disks whether or not RAID functionality is present.

Live Volume

The volume that is being accessed by the host for normal operations. The data on this volume is complete and not accessed via pointers and/or data structures. The difference between a regular volume and a Live Volume is that the Live Volume has been Snapshot enabled.

Logical Device

A device presented to an operating environment by control software or by a volume manager. From an application standpoint, a logical device is equivalent to a physical one. In some implementations, logical devices may differ from physical ones at the operating system level (e.g., booting from a host based disk array may not be possible).

Logical Partition Logical Unit

A logical partition is a segmentation of a logical volume.

Logical Unit Number (LUN)

The entity within a SCSI target that executes I/O commands. SCSI I/O commands are sent to a target and executed by a logical unit within that target. A SCSI physical disk typically has a single logical unit. Tape drives and array controllers may incorporate multiple logical units to which I/O commands can be addressed. Each logical unit exported by an array controller corresponds to a logical disk. (Common practice uses the terms "Logical Unit" and "LUN" interchangeably, although this is not strictly correct).

A SCSI representation of a system drive on a given channel and target ID. An encoded three-bit identifier for the logical unit.

Logical Volume

An arbitrary-sized space in a volume group that can be used as an address space for a file system or as device swap space. Logical volumes behave like disk block devices, except that, unlike physical disk partitions, they can be dynamically grown, shrunk and moved about without rebooting an operating system or entering into a maintenance or stand-alone mode.

LUN Zoning

In Fibre Channel, several devices grouped by function or by location. All devices connected to a connectivity product may include configuration of one or more zones. Devices in the same zone can see each other; devices in different zones cannot. A fabric management service used to create logical device subsets within a Storage Area Network (SAN).

Mapping

Conversion between two data addressing spaces. For example, mapping refers to the conversion between physical disk block addresses and the block addresses of the logical disks presented to operating environments by control software.

Management Information Base (MIB)

In SNMP, a collection of data elements that define the device settings the *Storage Concentrator* can control and the information it can retrieve it from a storage device.

Local Mirrors

Mirrored images that are located behind the Primary Storage Concentrator.

Metadata

Data that describes data. In disk arrays, meta-data consists of items such as array membership, member extent sizes and locations, descriptions of logical disks and partitions, and array state information. In file systems, meta-data includes file names, file properties and security information, and lists of block addresses at which each file's data is stored.

Mirroring

StoneFly Replicator™ software provides storage-independent replication that operates at the block level over an iSCSI network. Replicator supports local and remote replication and can operate synchronously or asynchronously.

StoneFly Mirroring™ synchronous mirroring feature supplies host-independent mirrored data storage that duplicates production data onto physically separate mirrored target images transparently to users, applications, databases, and host processors. Synchronous mirroring implies that the *Storage Concentrator* waits for a write-complete acknowledgement from all volumes before presenting a write completion status to the host.

Mirror Image

A mirror image contains an exact duplicate of all other images in a mirror volume. A mirror image is grouped with other mirror images to comprise a mirror volume (each mirror image contains a duplicate copy of the data). Mirror images can be either local mirrors or they can be campus mirrors. The size of the mirror image must be equal to the size of all other images in the mirror volume.

Mirror Volume

A Mirror Volume is composed of multiple mirror images. StoneFly Mirroring supports up to four mirror images in a mirror volume. The Mirror Volume is what the applications and hosts see as the storage device.

Mount Directory

Directory in any file system where the top directory of a descendent file system is mounted. The contents of the mount directory are the contents of the top directory in the mounted file system. If the mount directory is not empty before the file system is mounted, any existing files and directories in that directory as well as any descendent directories become invisible (and inaccessible) until the file system is un-mounted.

Mount Point

Synonym for mount directory. The mount point is the location (directory) where a file system, known to the host system, is mounted. Usually defined in terms of which system the file system is mounted on and where on that system the file system is mounted.

Network Attached Storage (NAS)

A term used to refer to storage elements that connect to a network and provide file access services to computer systems. Abbreviated NAS. A NAS Storage Element consists of an engine, which implements the file services, and one or more devices, on which data is stored. NAS elements may be attached to any type of network. A class of systems that provide file services to host computers. A host system that uses network attached storage uses a file system device driver to access data using file access protocols such as NFS or CIFS. NAS systems interpret these commands and perform the internal file and device I/O operations necessary to execute them.

Network File System

In NAS, a distributed file system and its associated network protocol originally developed by Sun Microsystems Computer Corporation and commonly implemented in UNIX systems, although most other computer systems have implemented NFS clients and/or servers. Abbreviated NFS. The IETF is responsible for the NFS standard.

Partition

1) Subdivision of the capacity of a physical or logical disk. Partitions are consecutively numbered ranges of blocks that are recognized by MS-DOS, Windows, and most UNIX operating systems.

2) Synonym for the type of extent used to configure arrays.

3) A contiguously addressed range of logical blocks on a physical media that is identifiable by an operating system via the partition's type and subtype fields. A partition's type and subtype fields are recorded on the physical media and hence make the partition self-identifying.

Primary Storage Concentrator

The *Storage Concentrator* that manages the mirror volume.

Promote Image

Promoting a mirror image makes it a stand-alone volume. The promoted volume does not retain any mirrored volume information but contains an exact copy of the data at the time of promotion. The new stand-alone volume is accessible to the hosts and can retain security information from the mirror volume.

Partitioning

Presentation of the usable storage capacity of a disk or array to an operating environment in the form of several logical disks whose aggregate capacity approximates that of the underlying physical or logical disk. Partitioning is common in MS-DOS, Windows, and UNIX environments. Partitioning is useful with hosts that cannot support the full capacity of a large disk or array as one device. It can also be useful administratively, for example, to create hard subdivisions of a large logical disk.

Path

Path is the access path from a host computer to a storage device.

Physical Extent

A unit of storage space on a physical device containing a specific amount of storage space. A collection of physical extents is then managed by the StoneFly Volume Manager as a logical volume, and a group of logical volumes make up a volume group. Physical extents are the smallest manageable element in a logical volume that can be managed by the StoneFly Volume Manager. The kernel and file system then use standard disk or file-system blocks when writing or reading to and from.

Physical Volume

A physical device such as a disk drive or RAID sub-system that, usually together with other devices, is configured as a volume group for subsequent division into one or more logical volumes. Logical volumes can be of arbitrary size (in multiples of physical extent size), whereas the physical volume is defined by the physical storage boundaries of the actual recording media or recording surface area in the device.

Primary Storage Concentrator

A Primary *Storage Concentrator* is the *Storage Concentrator* in a cluster that has an active Webserver for the Administrative Interface. The Primary *Storage Concentrator* configures the volumes and back-end sessions on the Secondary *Storage Concentrator(s)* within the cluster.

Port

An I/O adapter used to connect an intelligent device (node) to an I/O bus or network storage subsystems. Port is the synonym for the head end of a device I/O bus containing the arbitration logic.

Provisioning

The logical volume management services provided by the *Storage Concentrator* system create a centrally administered SAN infrastructure ideally suited for provisioning enterprise or departmental SAN, over an IP network.

Storage provisioning is the process of presenting a uniform and logical representation of physical storage resources transparent to the consumers of the storage (applications and users). Storage provisioning is not restricted by the type of storage, server platform or connection methodology. Storage provisioning dynamically maps data from the logical storage space required by applications to the actual physical storage space.

RAID

RAID is an acronym for Redundant Array of Independent Disks, a Firmware/hardware implementation of a disk array controller in which data is stored on disks in such a manner as to improve performance and avoid data loss in the presence of a disk failure.

Raw Device / Raw Partition

A disk partition not managed by a volume manager. The term raw partition is frequently encountered when discussing database systems because some database system vendors recommend volumes or files for underlying database storage, while others recommend direct storage on raw partitions.

Reattach Image

A previously detached mirror image can be reattached to the original mirror volume. Reattaching a mirror image removes it as a stand-alone image volume and initiates a rebuild operation on the reattached image.

Rebuild

The process of creating a mirror image from an existing volume. The volume is duplicated onto the mirror image at the block level in its entirety.

Relational Database Management System (RDBMS)

An RDBMS is a type of database management system that stores data in the form of related tables. Relational databases are powerful because they require few assumptions about how data is related or how it will be extracted from the database. As a result, the same database can be viewed in many different ways.

An important feature of relational systems is that a single database can be spread across several tables

Resource (Physical Device)

A disk drive, RAID (Redundant Array of Independent Disks) subsystem or other mass storage device and the data–storing media it contains. Sometimes referred to as a physical volume.

Revert Mirror to Span

Reverting a Mirror to a Span converts the mirror volume to a regular spanned volume.

Rollback

In Snapshot, rollback is a disaster recovery feature that allows users to return to the last known "good" data point in the event of corruption, viruses, or lost files. Essentially, performing a rollback returns the Live Volume to the last good known state.

Scalability

Capable of being changed in size and configuration. It typically refers to a computer, product or system's ability to expand.

Secondary Servers

A server attached to the Secondary *Storage Concentrator* has access to campus mirror images; however it views them as a local virtual volume. A server attached to the Secondary *Storage Concentrator* can also have its own local mirror volumes. These volumes are accessible as campus mirrors from the Primary *Storage Concentrator*.

Secondary Storage Concentrator

Another *Storage Concentrator* in a campus configuration that presents mirror images to the primary *Storage Concentrator*. Either the Primary *Storage Concentrator* or the Secondary *Storage Concentrator* or both can be single units or FailOver clusters.

SES

SES is an acronym for SCSI-3 Enclosure Services. Provides a means of SCSI access to multiple devices within an enclosure.

Simple Network Management Protocol (SNMP)

SNMP is an IETF protocol for monitoring and managing systems and devices in a network. The data being monitored and managed is defined by a MIB. The functions supported by the protocol are the request and retrieval of data, the setting or writing of data, and traps that signal the occurrence of events.

Small Computer Storage Interface (SCSI)

SCSI is a collection of ANSI standards and proposed standards that define I/O buses primarily intended for connecting storage subsystems or devices to hosts through host bus adapters. Originally intended primarily for use with small (desktop workstation) computers, SCSI has been extended to serve most computing needs, and is arguably the most widely implemented I/O bus in use today.

Snapshot

A point-in-time copy of a Live Volume. Snapshots contain an image of the volume at the exact point in time that the snapshot was taken. Snapshots can be created nearly instantaneously. The snapshot volume appears to the host as if it was a regular logical volume.

Snapshot Fragmentation/ Total Fragmentation for All Volumes

These percentages represent the amount of the index table fragmentation on the volume selected as well as the total amount of fragmentation for all Snapshot volumes. These percentages have no relationship to disk fragmentation.

Snapshot Volume

A Snapshot Volume is a virtual volume that represents a point-in-time image of a Live Volume. The Snapshot Volume data is a combination of data from the Live Volume and data from the Snapspace.

Snapspace

Snapspace is storage space configured by the Storage Concentrator used to store the changed data for maintaining snapshots. This data space is not directly readable by hosts. For fault tolerance, the Snapspace can be mirrored.

Source

A place from which data is taken. The place from which the data is acquired is called the *source*, whereas the place it is sent to or moved to is called the destination or target.

Spanning

A volume that is created comprised of regions or sections of several physical devices.

Secondary Storage Concentrator

A Secondary Storage Concentrator is a redundant Storage Concentrator that can be assigned volumes and accept sessions from hosts for those volumes. The Secondary Storage Concentrator in the cluster will transparently (to the hosts) become the Primary Storage Concentrator in the event of a failure in the Primary Storage Concentrator. The original Primary Storage Concentrator's storage volumes are now managed by the new Primary Storage Concentrator.

Storage Area Network (SAN)

A storage area network (SAN) is a separate and specialized network whose primary purpose is the transfer of data among storage elements and between computer systems and storage elements. A SAN consists primarily of a communication infrastructure and a management layer. The communication infrastructure provides physical connections and the management layer organizes the connections, storage elements, and computer systems so that data transfer is secure and robust.

Storage Concentrator

The *Storage Concentrator* delivers iSCSI target volumes to hosts over TCP connections in an Ethernet network. Configuring and managing these iSCSI target volumes is accomplished using a browser-based graphical user interface (GUI) resident in the *Storage Concentrator*. Storage resources are connected to the *Storage Concentrator* through a parallel SCSI connection or other connections. The system administrator uses the graphical user interface to allocate blocks of storage to create the iSCSI target volumes and authorizes their use by individual host systems.

Target

Target is synonymous with destination, a target is a file, a device, or any type of location to which data is moved or copied. The target is the provider of storage.

Target Portal

An iSCSI communications gateway between a primary Storage Concentrator and a Secondary Storage Concentrator at a campus site. Once the target portal has been created, the system will recognize storage at the campus site as an available resource for creating secondary mirror images.

Volume Group

A volume group (also referred to as SPAN) is one or more separate physical devices (disk drives or RAID sub-systems) called physical volumes that are configured to form a single large storage area that is then divided into one or more separate storage areas of arbitrary size called logical volumes. Each logical volume can then be a used as a separate storage space for file system or for raw data storage by specific applications; without being constrained by physical disk boundaries. This feature is useful in various situations such as creating large file systems that exceed the size of a single disk or disk array, creating two or more separate file systems on a single disk device and/or creating a number of file systems having dissimilar geometries. Volume group is a management abstraction term. For example, management actions and policies are applied to volume groups.

Appendix 14

Index

A

B

C

D

E

F

H

I

L

M

N

O

P

Q

R

S

T

U

V

W

This page is intentionally left blank.

StoneFly, Inc.

26250 Eden Landing Road

Hayward, CA 94545

(510) 265-1616

support@stonefly.com

www.stonefly.com

www.iscsi.com