Quantum

Installation Guide Installation Guide Installation Guide

StorNext® 4.0.1

StorNext

6-00360-22 Rev A

StorNext 4.0.1 Installation Guide, 6-00360-22, Rev A, August 2010, Product of USA.

Contents

Figures

Tables

Preface

StorNext provides high performance data sharing and intelligent archiving that is compatible with an industry-leading range of operating systems, server platforms, and storage devices.

StorNext includes two main components:

- StorNext File System (SNFS) is high performance data sharing software that lets applications on multiple operating systems access a single data set.
- StorNext Storage Manager (SNSM) is an intelligent, policy-based data mover that automatically moves and locates data across multiple storage tiers.

In addition, StorNext includes client software that allows client systems to mount and work with StorNext file systems.

Audience

This document is intended for service professionals and system administrators who install StorNext software. The installer should have strong familiarity with the operating system on which StorNext is installed.

Only service professionals, experienced StorNext users, or individuals who have successfully completed StorNext training should configure StorNext. Attempting to configure StorNext without proper experience or training can affect system performance and may result in data loss.

Purpose

This document provides information about installing and configuring StorNext. The following topics are covered:

- System requirements
- Pre-installation tasks
- Installation
- Configuration
- Uninstalling StorNext

Document Organization

This document contains the following chapters:

- Chapter 1, Installing StorNext provides system requirements, pre-installation information, and instructions for installing StorNext File System and Storage Manger on a metadata controller (MDC).
- Chapter 2, Installing the StorNext Client provides system requirements, pre-installation information, and instructions for installing the StorNext client software.
- Chapter 3, Configuring StorNext describes how to use the StorNext configuration wizard and the Windows configuration utilities to create a basic StorNext configuration, and how to configure a distributed LAN client or server.
- Appendix A, "Uninstalling StorNext," describes how to remove StorNext software from a server or client system.
- Appendix B, "StorNext Administration Tasks,"describes how to perform file system tasks that do not apply to a typical StorNext setup.
- Appendix C, "StorNext Files,"describes the files used and modified to configure a file system, the cluster-wide control feature, and the file used to configure the cluster-wide security controls.
- Appendix D, "Quality of Service Guide,"describes how to configure and use the StorNext File System (SNFS) Quality of Service (QOS) feature.

Notational Conventions

This manual uses the following conventions:

Convention	Example
User input is shown in bold font.	**./DARTinstall**
Computer output and command line examples are shown in monospace font.	`./DARTinstall`
User input variables are enclosed in angle brackets.	**http://<ip_address>/cgi-bin/stats**

For UNIX and Linux commands, the command prompt is implied.	`./DARTinstall`
	`is the same as`
	`# ./DARTinstall`

File and directory names, menu commands, button names, and window names are shown in bold font.	**/data/upload**
Menu names separated by arrows indicate a sequence of menus to be navigated.	**Utilities > Firmware**

The following formats indicate important information:

Note: Note emphasizes important information related to the main topic.

Caution: Caution indicates potential hazards to equipment or data.

WARNING: Warning indicates potential hazards to personal safety.

Related Documents

Documents related to StorNext are shown below:

Quantum StorNext Documentation

Document No.	Document Title	Document Description
6-01658-09	StorNext User's Guide	Provides in-depth information about using StorNext to create and manage StorNext File Systems.
6-01376-13	StorNext File System Tuning Guide	Provides information for maximizing file system performance.
6-01620-12	StorNext Upgrade Guide	Provides procedures for upgrading from earlier versions of StorNext.
6-01688-09	StorNext CLI Reference Guide	Provides procedures for performing common tasks using the command line interface.
6-00431-27	StorNext 4.0 Release Notes	Includes the latest product information, fixes, and issues.

Contacting Quantum

More information about this product is available on the Quantum Service and Support website at www.quantum.com/ServiceandSupport. The Quantum Service and Support website contains a collection of information, including answers to frequently asked questions (FAQs). You can also access software, firmware, and drivers through this site.

To request a software upgrade, visit www.quantum.com/ServiceandSupport/Upgrade/Index.aspx. For further assistance, or if training is desired, contact Quantum Global Services:

United States	800-284-5101 (toll free) 949-725-2100
EMEA	00800-4-782-6886 (toll free) +49 6131 3241 1164
APAC	+800 7826 8887 (toll free) +603 7953 3010
For worldwide support: http://www.quantum.com/ServiceandSupport/Index.aspx	

(Local numbers for specific countries are listed on the Quantum Service and Support Website.)

Chapter 1
Installing StorNext

This chapter describes how to install StorNext File System (SNFS) and StorNext Storage Manager (SNSM) on a metadata controller (MDC). Install both SNFS and SNSM for storage systems that require policy-based data movement (for example, systems that include tape drives or libraries).

To ensure successful operation, do the following tasks before installing StorNext:

- Make sure the MDC meets all operating system and hardware requirements (see Storage Manager System Requirements on page 2).
- Make sure all storage devices are correctly configured and are visible to the MDC (see Getting Ready to Install on page 6)
- (Optional) Run the pre-installation script to check for available disk space and view recommended locations for support directories (see Pre-Installation Script on page 11).

When you are ready to install StorNext File System and Storage Manager on the MDC, run the installation script (see StorNext Installation Script on page 14).

Storage Manager System Requirements

To run StorNext File System and Storage Manager, the metadata controller must meet all operating system, hardware, and LAN requirements.

Operating System Requirements

The operating systems, kernel versions, and hardware platforms supported by StorNext SNFS and SNSM are presented in Table 1. Make sure the MDC uses a supported operating system and platform, and if necessary update to a supported kernel version before installing StorNext.

Table 1 Storage Manager Supported Platforms

Operating System	Kernel	Platform
Red Hat® Enterprise Linux® 5	2.6.18-53.EL (Update 1) 2.6.18-92.EL (Update 2) 2.6.18-128.1.1.EL (Update 3 2.6.18-164.EL (Update 4)	x86 64-bit
SUSE™ Linux Enterprise Server 10	2.6.16-46-0.12 (SP1) 2.6.16-60-0.27 (SP2) 2.6.16.60-0.54.5 (SP3)	x86 64-bit

* MDC support only. Storage Manager is not supported on Red Hat Enterprise Linux 4.

Hardware Requirements

The minimum amount of RAM and available hard disk space required to run StorNext SNFS and SNSM are presented in Table 2. StorNext utilizes database and journal files, and these are stored on the MDC. Consequently, the amount of local disk space that is required increases with the number of data files stored on StorNext file systems.

If necessary, upgrade the RAM and local disk storage in the MDC to meet the minimum requirements before installing StorNext.

Note: The RAM requirements in Table 2 are for running StorNext File System and Storage Manager only. Running additional software (including the StorNext client software) requires additional RAM.

StorNext Media on DVD

Beginning with StorNext 4.0, installation and upgrade media is shipped on DVDs. (Media was previously shipped on CDs.) Both software and documentation are now stored on a single DVD, eliminating multiple discs for different operating systems.

If you plan to install from media, you must have a DVD ROM drive to perform the installation or upgrade.

StorNext File System and Storage Manager Requirements

The hardware requirements for StorNext File System and Storage Manager are presented in Table 2.

Table 2 File System and Storage Manager Hardware Requirements

No. of File Systems	RAM	File System Disk Space	Storage Manager Disk Space
1–4*	4 GB	2 GB	• For application binaries, log files, and documentation: up to 30GB (depending on system activity) • For support directories: 3 GB per million files stored
5–8**	8 GB	4 GB	

* Two or more CPU cores are recommended for best performance.
** Two or more CPU cores are required for best performance.

Partitioning Local Hard Disks

StorNext can be installed on any local file system (including the root file system) on the MDC. However, for optimal performance, as well as to aid disaster recovery, follow these recommendations:

- Avoid installing StorNext on the root file system.
- Partition local hard disks so that the MDC has four available local file systems (other than the root file system) located on four separate hard drives.

Note: You can run the pre-installation script to help determine the estimated size of and optimal location for StorNext support directories. For more information, see Pre-Installation Script on page 11.

Additional Memory and Disk Requirements for Deduplication and Replication

In order to use the data deduplication and replication features in StorNext 4.0, your system must have the following memory and disk capacity **in addition to** the base memory and disk capacity required to run StorNext File System and Storage Manager.

Note: Deduplication is supported only on 64-bit operating systems.

Minimum Additional Memory and Disk Capacity Required for the Deduplication Repository

- 50 MB available hard disk space

Minimum Additional Memory and Disk Capacity Required for Systems Licensed for 0 - 1 TB of Data

- 1 GB additional RAM
- 1 TB available hard disk space

Minimum Additional Memory and Disk Capacity Required for Systems Licensed for 1 - 10 TB of Data

- 6 GB additional RAM
- 10 TB available hard disk space

Minimum Additional Memory and Disk Capacity Required for Systems Licensed for 10 - 50 TB of Data

- 13 GB additional RAM
- 50 TB available hard disk space

Minimum Additional Memory and Disk Capacity Required for Systems Licensed for 50 - 150 TB of Data

- 28 GB additional RAM
- 150 TB available hard disk space

LAN Requirements

The following LAN requirements must be met before installing StorNext on the MDC:

- In cases where gigabit networking hardware is used and maximum StorNext performance is required, a separate, dedicated switched Ethernet LAN is recommended for the StorNext metadata network. If maximum StorNext performance is not required, shared gigabit networking is acceptable.
- A separate, dedicated switched Ethernet LAN is mandatory for the metadata network if 100 Mbit/s or slower networking hardware is used.
- The MDC and all clients must have static IP addresses.

  Verify network connectivity with pings, and also verify entries in the **/etc/hosts** file. Alternatively, telnet or ssh between machines to verify connectivity.
- If using Gigabit Ethernet, disable jumbo frames and TOE (TCP offload engine).
- The hostname **localhost** is resolvable on the MDC.

- The hostname **localhost** on the MDC resolves to a an IPv4 address on the loopback device.
- The hostname **localhost** must not resolve to an IPv6 address.

Note: StorNext does not support file system metadata on the same network as iSCSI, NFS, CIFS, or VLAN data when 100 Mbit/s or slower networking hardware is used.

Other Requirements

The following requirements must be met before installing StorNext on the MDC:

- The MDC does not have SELinux enabled.
- Quantum recommends that system clocks are synchronized using NTP for easier debugging, particularly in an HA environment.

Getting Ready to Install

Before installing StorNext SNFS and SNSM, complete the following tasks to ensure successful installation:

- Make sure you understand changes in licensing procedures (see StorNext Licensing).
- Correctly configure all storage devices (see Configuring Storage Devices).
- If using LUNs larger than 1 TB, decide on a label type and install any necessary operating system patches (see Planning for LUNs Larger than 1 TB).
- (Linux only) Install the kernel source code (see Installing the Linux Kernel Source Code on page 10).

StorNext Licensing

Beginning with StorNext 4.0, licensing has changed significantly. Separate licenses are now required for different purchased StorNext

components (such as File System and Storage Manager) and features (such as Replication and Distributed Data Mover.)

Here is a list of all the StorNext licenses:

- **File System**: A File System license enables you to create and modify StorNext-supported file systems.
- **LAN Client**: You must have a Distributed LAN Client license for each LAN client you use with StorNext (in addition to any SAN clients).
- **Storage Manager**: A Storage Manager license provides full access to StorNext's Storage Manager features that are not licensed separately.
- **Replication**: A replication license is required if you want to use StorNext's Data Replication feature.
- **Deduplication**: A deduplication license is required if you want to use StorNext's Data Deduplication (blockpool) feature.
- **Vaulting**: A Vaulting license provides the ability to move seldom-used media to a manual archive vault, freeing room for media in the managed archives.
- **Storage Disk**: You must have a Storage Disk license to be able to configure and use StorNext storage disks.
- **Checksum**: A Checksum license enables you to verify data integrity by ensuring that the checksum created when data was stored matches the checksum upon data retrieval.
- **Distributed Data Mover (DDM)**: The number displayed is the maximum number of clients that can be used to run mover processes.
- **Failover (HA)**: A Failover (High Availability) license is required if you plan to use StorNext's HA failover features.
- **Maintenance**: A Maintenance license verifies that your system has an active maintenance contract and is entitled to install software upgrades. A maintenance license has an expiration date equal to the expiration date of your service contract. A new maintenance license with a new expiration date will be provided with each contract renewal. Once a maintenance license expires, the installed StorNext version continues to function indefinitely without any limitations, but the system cannot be upgraded to any StorNext version that has a release date after the expiration date of the maintenance license.

Installation Steps

In light of the new licensing implementation, here is a summary of the steps required to install and license StorNext:

1. Run the pre-installation script. (See Pre-Installation Script on page 11).
2. Run the installation script. (See StorNext Installation Script on page 14).
3. Log into the StorNext GUI.
4. Stop services on the System Control screen (File System and Storage Manager).
5. Obtain licenses for all purchased StorNext components and features from Quantum Technical Support. (See the Configuration Wizard chapter in the StorNext User's Guide.)
6. Enter license information through the StorNext Configuration Wizard. (See the Configuration Wizard chapter in the StorNext User's Guide.)
7. Restart services on the System Control screen.

Configuring Storage Devices

Before installing StorNext SNFS and SNSM, make sure that all LUNs are visible to the MDC. (A LUN, or logical unit number, is a logical device that corresponds to one or more disks, drives, or storage devices.)

If there are any connection issues, resolve them before installing StorNext. For assistance in configuring storage devices, refer to the documentation that came with the storage device, or contact the manufacturer.

Note: LUNs that you plan to use in the same stripe group must be the same size. Consider this when configuring storage devices. (For more information about stripe groups, see the *StorNext User's Guide*.)

Caution: StorNext does not support the connection of multiple devices through fibre channel hubs. Multiple devices must be connected using fibre channel switches.

Planning for LUNs Larger than 1 TB

StorNext supports LUNs greater than 1 TB in size if they are allowed by the operating system. To enable support for 1 TB or larger LUNs, all StorNext LUNs must be correctly labeled according to the requirements of the operating system running on the MDC as well as the operating system running on all connected clients. Disk LUNs can have one of three labels: VTOC, EFI, or sVTOC (short VTOC).

Required disk LUN label settings based on operating system and LUN size are presented in Table 3. Before installing StorNext, decide what label type will be used, and then install any required operating system patches or updates (for both MDC and client) as indicated in the notes for Table 3.

Note: After labeling a disk LUN, you must reboot systems running Solaris before they can access the disk LUN.

Table 3 Required Disk LUN Labels

Operating System	Label for LUNs < 1 TB	Label for LUNs 1–2 TB	Label for LUNs > 2 TB
Windows[a]	VTOC, EFI	VTOC, EFI	EFI, sVTOC
Linux 2.6 Kernel	VTOC, EFI	VTOC, EFI	EFI, sVTOC
Linux 2.4 Kernel	VTOC, EFI	VTOC, EFI	Not supported
Solaris 9[b]	VTOC, EFI	EFI	Not supported
Solaris 10[c]	VTOC, EFI	EFI	EFI
HP-UX[d]	VTOC, EFI	VTOC, EFI	EFI, sVTOC
IBM AIX[e]	VTOC, EFI	VTOC, EFI	EFI, sVTOC
SGI IRIX	VTOC	VTOC	Not supported
Apple Xsan 2.1.1 (MacOS X 10.5 Leopard)	VTOC, EFI	VTOC, EFI	EFI, sVTOC

Operating System	Label for LUNs < 1 TB	Label for LUNs 1–2 TB	Label for LUNs > 2 TB
Apple Xsan 2.2 (MacOS X 10.6 Snow Leopard)	VTOC, EFI	VTOC, EFI	EFI, sVTOC

a. For Windows XP 32-bit systems only, support for LUNs larger than 2 TB is limited to the first 2 TB. All other supported Windows platforms (including Windows XP 64-bit) support LUNs greater than 2 TB.
b. Support for EFI labels requires the Multi terabyte Disk Support patch.
c. Support for 2 TB or larger LUNs requires Solaris 10 Update 2 (06/06).
d. Supports only the first 2 TB of LUNs larger than 2 TB.
e. Supports only LUNs of 2.2 TB or less.

Note: For best results, Quantum recommends using persistent binding for disk LUNs. For more information, contact the vendor of your HBA (host bus adapter).

Installing the Linux Kernel Source Code

For management servers running Red Hat Enterprise Linux version 5, before installing SNFS and SNSM you must first install the kernel header files (shipped as the kernel-devel or kernel-devel-smp RPM, depending on your Linux distribution).

For servers running SUSE Linux Enterprise Server, you must install the first kernel source code (shipped as the kernel-source RPM). StorNext will not operate correctly if these packages are not installed. You can install the kernel header files or kernel source RPMs by using the installation disks for your operating system.

Verifying Hostname Length

The maximum hostname length for a StorNext server is limited to 255 characters. Before you begin the installation, verify that the destination hostname is not longer than 255 characters.

DDisk Support Discontinued

Beginning with StorNext 4.0, with the introduction of the deduplication feature DDisk is no longer supported due to conflicting technology. *Any existing DDisks will not be transferred to 4.0 when you upgrade.* For

assistance transitioning from DDisk to deduplication, contact Quantum Technical Support.

Pre-Installation Script

The StorNext pre-installation script (**snPreInstall**) is included on the StorNext installation DVD. When you run **snPreInstall**, you are prompted for information about your system. The pre-installation script uses this information to estimate the amount of local disk space required for SNFS and SNSM support directories. In addition, the script recommends the optimal locations for support directories.

StorNext uses five directories to store application support information. These directories are stored locally on the metadata controller, except for the **Backup** directory, which is stored on the managed file system.

The StorNext support directories are described in Table 4.

Table 4 StorNext Support Directories

Support Directory*	Description
Database `/usr/adic/database`	Records information about where and how data files are stored.
Journal `/usr/adic/database/journal`	Records changes made to the database.
Mapping `/usr/adic/TSM/internal/mapping_dir`	Contains index information that enables quick searches on the file system.
Metadata `/usr/adic/database/metadumps`	Stores metadata dumps (backups of file metadata).

* Depending on file system configuration, the support directory locations may be different. However, symbolic links in **/usr/adic/** will always point to the support directories.

Before You Begin

Before running the pre-installation script, be prepared to answer the following questions:

- Is this an upgrade installation?
- What local file systems can be used to store support information?
- Which version of StorNext will be installed?
- What is the maximum number of directories expected (in millions)?
- What is the maximum number of files expected (in millions)?
- How many copies will be stored for each file?
- How many versions will be retained for each file?

Note: Keep in mind that storage needs typically grow rapidly. Consider increasing the maximum number of expected directories and files by a factor of 2.5x to ensure room for future growth.

Note: The pre-installation script ignores unmounted file systems. Before running **snPreInstall**, be sure to mount all local file systems that will hold StorNext support information.

Running snPreInstall

To run the pre-installation script, use the StorNext installation DVD.

1. Log on to the MDC as **root**.
2. Mount the StorNext installation DVD and change to the DVD root directory.
3. List the installation directories on the DVD. At the command prompt, type:

   ls -l
4. Identify the correct installation directory for your operating system and hardware platform, and then change to that directory.

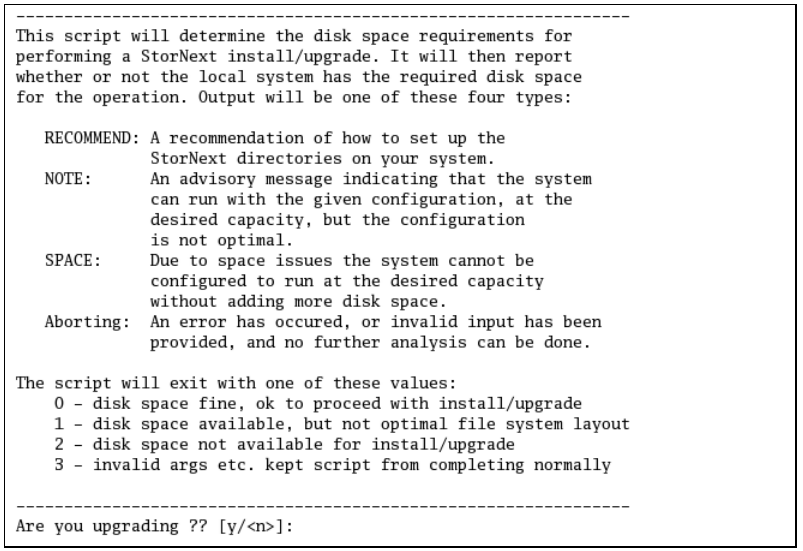
   For example, for Red Hat Linux 5 running on an x86 64-bit platform, change to the **RedHat50AS_26x86_64** directory.
5. Run the script. At the command prompt, type the following:

   . /usr/adic/.profile

./snPreInstall

The pre-installation script runs (Figure 1).

Figure 1 Pre-Installation Script



```
-----------------------------------------------------------------
This script will determine the disk space requirements for
performing a StorNext install/upgrade. It will then report
whether or not the local system has the required disk space
for the operation. Output will be one of these four types:

   RECOMMEND: A recommendation of how to set up the
              StorNext directories on your system.
   NOTE:      An advisory message indicating that the system
              can run with the given configuration, at the
              desired capacity, but the configuration
              is not optimal.
   SPACE:     Due to space issues the system cannot be
              configured to run at the desired capacity
              without adding more disk space.
   Aborting:  An error has occured, or invalid input has been
              provided, and no further analysis can be done.

The script will exit with one of these values:
    0 - disk space fine, ok to proceed with install/upgrade
    1 - disk space available, but not optimal file system layout
    2 - disk space not available for install/upgrade
    3 - invalid args etc. kept script from completing normally

-----------------------------------------------------------------
Are you upgrading ?? [y/<n>]:
```

6 Follow the onscreen instructions, and answer each question when prompted.

The script outputs results based on the information you entered.

Interpreting snPreInstall Output

After you enter all requested information, the pre-installation script outputs the following results:

- Estimated disk space required for each support directory.
- Recommended file system location for each support directory.

Note: For optimal performance, each support directory (other than the **Backup** directory) should be located on its own local file system, and each local file system should reside on a separate physical hard disk in the MDC.

There are three types of directory recommendations, as indicated by the text preceding the recommendation. The three types of recommendations are described in Table 5.

Table 5 Pre-Installation Script Recommendation Types

Type	Description
RECOMMEND	The suggested directory locations will result in optimal system performance.
NOTE	The suggested directory locations are valid but system performance will not be optimal.
SPACE	Additional hard disk space is required to store the required support directories.

The pre-installation script bases directory location recommendations on the following criteria:

- To aid disaster recovery, the **Database** and **Journal** directories should be located on different file systems.
- For optimal performance, the **Metadata** directory should not be located on the same file system as (in order of priority) the **Journal**, **Database**, or **Mapping** directory.

Note: These recommendations apply to non-HA systems only.

Note: Do not change the location of support directories manually. Instead, use the installation script to specify the location for support directories. For more information, see StorNext Installation Script.

StorNext Installation Script

When you are ready to install StorNext, use the StorNext installation script to install StorNext File System and Storage Manager on the metadata controller. The installation script also installs the client software on the MDC.

Upgrading From a StorNext File System-Only Configuration

If you have StorNext File System only and are adding StorNext Storage Manager, some components remain from the File System installation and the new installation could fail. To prevent installation failure, you should first remove the old components.

The following procedure describes how to remove the old components before starting the new StorNext installation. Follow the steps below *only* if you are upgrading from a StorNext File System-only configuration to StorNext File System and Storage Manger.

1. Log on to the MDC as **root**.
2. Save **/etc/fstab** and **/usr/cvfs/config**
3. Remove the existing StorNext File System rpms.
4. Install StorNext as described in Launching the StorNext Installation Script on page 15.
5. Restore **/etc/fstab** and **/usr/cvfs/config**

Launching the StorNext Installation Script

The StorNext installation script lets you choose installation configuration options and install StorNext. To launch the script, use the correct StorNext installation DVD for your operating system.

1. Log on to the MDC as **root**.
2. Mount the StorNext installation DVD and change to the DVD root directory.

   Note: When you mount a DVD in a Red Hat 5 system, DVDs are mounted by default with a noexec (non-executable) option which prevents you from proceeding with the installation.

   For Red Hat users only, before proceeding you must remount the DVD by typing **mount -o remount, exec ...**

3. List the installation directories on the DVD. At the command prompt, type:

   ls -l

4 Identify the correct installation directory for your operating system and hardware platform, and then change to that directory.

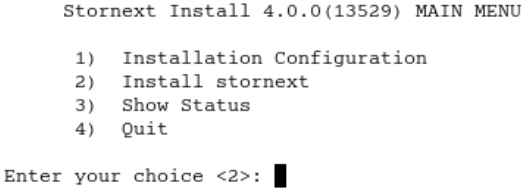
For example, for Red Hat Linux 5 running on an x86 64-bit platform, change to the **RedHat50AS_26x86_64** directory.

5 Run the script. At the command prompt, type:

./install.stornext

The installation script runs (Figure 2).

Figure 2 StorNext Installation Script



6 Type the number corresponding to a menu option and press **<Enter>**. Installation menu options are described in Table 6.

Table 6 StorNext Installation Main Menu Options

Menu Option	Description
1	Change installation options, including the location of application directories and the default media type (see Changing Installation Configuration Options on page 17).
2	Install StorNext using the installation options specified on the Configuration Menu (see Performing an Installation Operation on page 18).
3	View the status of an installation operation that is currently in progress.
4	Exit the StorNext installation script.

Changing Installation Configuration Options

Use the Configuration Menu to change StorNext installation options. You can specify the location of application and support directories and change the default media type for storage devices.

On the Main Menu, type **1** and press **<Enter>**. The Configuration Menu appears (Figure 3).

Figure 3 Configuration Menu

```
      Stornext 4.0.0(13529) Configuration Menu

      __VALUE_____________________          __DESCRIPTION___________
  1)  /usr/adic/perl                        Perl installation directory
  2)  /usr/adic/docs                        Docs installation directory
  3)  /usr/adic/PCL                         Pcl installation directory
  4)  /adic/database                        Dbm installation directory
      /usr/adic/SRVCLOG                     Srvclog installation directory
  6)  /usr/adic/PSE                         Pse installation directory
      /usr/adic/MSM                         Media Manager installation directory
      /usr/adic/TSM                         Tertiary Manager installation directory
  9)  /usr/adic/java                        Java5 installation directory
      /usr/adic/tomcat                      Tomcat installation directory
      /usr/adic/gui                         Sngui installation directory
      /usr/cvfs                             Disk Manager installation directory
 13)  /adic/database_meta                   Metadumps directory
 14)  /adic/mapping_dir                     TSM mapping directory
 15)  /adic/database_jnl                    Database journal directory
      LTO                                   Default media type

Enter number to modify or hit <RETURN> to continue:
```

Changing Application Directory Locations

When you first enter the Configuration Menu, it displays recommended locations for StorNext application directories based on the available file systems detected by the installation script. To accept the recommended directory locations and return to the Main Menu, press **<Enter>**.

Alternately, customize the installation by specifying the file systems where application and support files are stored.

1. On the Configuration Menu, type the number that corresponds to the directory location you want to customize (**1**–**16**), and then press **<Enter>**.

   A list of available local file systems is shown.

2. To change the file system where the directory is located, type the name of the file system and press **<Enter>**.

3 To confirm the change, type **yes** and press **<Enter>**.

To customize additional directory locations, repeat steps 1–3.

4 When you are done customizing the installation, press **<Enter>** to return to the Main Menu.

Note: To help determine the optimal location for StorNext support directories, use the StorNext pre-installation script. For more information, see Pre-Installation Script on page 11.

Changing the Default Media Type

If you do not specify a different media type, the StorNext installation script selects LTO as the default media type for storage devices. If storage devices in your system use a different media type, change the default media type before installing StorNext.

1 On the Configuration Menu, type **16** and press **<Enter>**.

A list of valid default media types is shown. The valid media types are: SDISK, LTO, LTOW, 3590, 3592, 9840, 9940, AITW, AIT, DLT4, and T10K.

2 Type the correct default media type for storage devices in your system and press **<Enter>**.

3 To confirm the change, type **yes** and press **<Enter>**.

4 When you are done customizing the installation, press **<Enter>** to return to the Main Menu.

Performing an Installation Operation

To install StorNext using the options specified on the Configuration Menu, perform an installation operation.

1 On the Main Menu, type **2** and press **<Enter>**.

The StorNext installation script runs. The script displays the current progress of the installation.

2 When the installation is complete, press **<Enter>** to return to the Main Menu.

3 To exit the installation script, type **4** and press **<Enter>**.

Note: The script displays the URL at which the MDC can be accessed. Make a note of this information.

You can now access StorNext File System and Storage Manager, and run the configuration wizard. For more information, see Chapter 3, Configuring StorNext.

Post-Installation Steps

After StorNext is installed, you are ready to configure your system. Quantum recommends using the StorNext Configuration Wizard, which walks you through the steps of configuring your system for the first time. For more information about the Configuration Wizard, see Using the Configuration Wizard on page 53.

Clear Browser Cache

If you are upgrading to StorNext 4.0 from a previous 3.x release, before logging in to the StorNext GUI for the first time you must first clear your Internet browser's cache. Doing this ensures that any settings from an earlier StorNext release are cleared and will not affect StorNext 4.0 operation.

SSL Certificate Exception

When you log into StorNext for the first time, you might see a message warning you about a security certificate. Refer to the Quantum Knowledge Base for a permanent workaround to this issue.

For a temporary solution, create a certificate exception that will allow you to log into StorNext without seeing the warning message during subsequent logins.

Starting StorNext Components

When you log in to StorNext for the first time, the **System Control** screen appears. On this screen you can start StorNext File System and Storage Manager by clicking **Start**.

If you would like to bypass the System Control screen and display the Configuration Wizard Welcome Screen, you can run the command `service cvfs start`.

Obtaining StorNext Permanent Licenses

StorNext comes with temporary 30-day licenses, but you should enter permanent licenses through the StorNext GUI as soon as it is convenient.

The StorNext User's Guide describes how to obtain permanent licenses from Quantum Technical Support. After you have obtained the licenses, you should stop StorNext services and then follow the directions in adding permanent licenses described in the Configuration Wizard chapter in the StorNext User's Guide.

After entering license information, you should start services on the System Control screen.

Installing a StorNext 4.x .RPM HA System

This chapter describes how to set up a StorNext 4.x HA system using an RPM-only installation. Steps are run on both the primary and secondary nodes, and are indicated accordingly.

On the Primary

Run the following commands on the **Primary** node:

1 Install the three StorNext RPMs by running this command:

```
# rpm -ivh snfs-client-SuSE100ES_261621-
4.0.0.12838.x86_64.rpm snfs-server-SuSE100ES-
4.0.0.12838.x86_64.rpm snfs-SuSE100ES-
4.0.0.12838.x86_64.rpm
```

Note: The actual build numbers may be different than those shown in the example. For example, the build number for `snfs-SuSE100ES-4.0.0.12838.x86_64.rpm` may not be 12838.

2 Populate the `fsnameservers` file with the IP addresses of the fsnameserver systems by running this command:

```
# vi /usr/cvfs/config/fsnameservers
```

3 Create the `ha_peer` file by running the command below. This file should contain the IP address of the **Secondary** in the HA pair.

```
# vi /usr/cvfs/config/ha_peer
```

4 Start up the fsmpm by running this command:

```
# /etc/init.d/cvfs start
```

5 Create the `.cfgx` file by running this command:

```
# /usr/cvfs/bin/sncfgedit -n snfs1
```

(For more information, see the `snfs.cfgx` man page.)

After creating the file, do the following:

a Delete first two lines at top that start with `****`

b Change the `haFsType` entry to `HaUnmanaged`. That is, change this:

<haFsType>**HaUnmonitored**</haFsType>

to this:

<haFsType>**HaUnmanaged**</haFsType>

c Fix the <diskTypes> and <stripeGroups> sections at the bottom.

6 Make the file system by running this command:

```
# /usr/cvfs/bin/cvmkfs snfs1
```

7 Create the mount point by running this command:

```
# mkdir /stornext/snfs1
```

8 Add the file system to the fsmlist by running this command:

```
# vi /usr/cvfs/config/fsmlist
```

9 Configure the file system to automatically mount after reboot. To do this, edit the `/etc/fstab` file so that it contains the following line:

```
<file_system> <mount_point> cvfs rw 0 0
```

where <file_system> is the name of the StorNext file system and <mount_point> is the directory path created in Step 8.

10 Start and mount the file system by running this command:

```
# /usr/cvfs/bin/cvadmin -e 'start snfs1'
# mount /stornext/snfs1
```

On the Secondary

Run the following commands on the **Secondary** node:

1 Install RPMs by running this command:

```
# rpm -ivh snfs-client-SuSE100ES_261621-
4.0.0.12838.x86_64.rpm snfs-server-SuSE100ES-
4.0.0.12838.x86_64.rpm snfs-SuSE100ES-
4.0.0.12838.x86_64.rpm
```

Note: The actual build numbers may be different than those shown in the example. For example, the build number for `snfs-SuSE100ES-4.0.0.12838.x86_64.rpm` may not be 12838.

2 Create the `fsnameservers` file by running this command:

```
# vi /usr/cvfs/config/fsnameservers
```

Note: The `fsnameservers` file on the secondary must be the same as the `fsnameservers` file on the primary.

3 Create the `ha_peer` file by running the command below. This file should contain the IP address of the **Primary** in the HA pair.

```
# vi /usr/cvfs/config/ha_peer
```

4 Start up the fsmpm by running this command:

```
# /etc/init.d/cvfs start
```

5 Copy the .cfgx file from the Primary.

6 Add the file system to the fsmlist by running this command:

```
# vi /usr/cvfs/config/fsmlist
```

7 Add the file system to `/etc/fstab` by running this command:

```
# vi /etc/fstab
```

8 Start and mount the file system by running this command:

```
# /usr/cvfs/bin/cvadmin -e 'start snfs1'
# mount /stornext/snfs1
```

Installation Notes

Following are some considerations pertaining to RPM installations:

- For RPM installations (specifically, when there is no “HA shared” file system), the ha manager program `/usr/cvfs/bin/snhamgr` is not functional.
- To turn off HA functionality, change `haFsType` back to `HaUnmonitored` and bounce the FSM. Before doing this, ensure that the Secondary is turned off and unplugged.

Edit /etc/hosts File

SuSe Linux distributions automatically associate the FQDN of the local machine with the address 127.0.0.2 in the `/etc/hosts` file. There is no benefit from doing this when the machine is connected to a network that can resolve its name to an IP address.

However, the existence of this entry can sometimes cause a failure of configuration synchronization within and between the server computers in an HA configuration. For this reason, the 127.0.0.2 entry should be deleted from the /etc/hosts file.

Chapter 2
Installing the StorNext Client

This chapter describes how to install the StorNext client software. The StorNext client software lets you mount and work with StorNext file systems.

To ensure successful operation, make sure the client system meets all operating system and hardware requirements (see Client System Requirements on page 26).

To install the StorNext client software, first download the client software from the metadata controller (MDC) (see Downloading the StorNext Client Software on page 29).

After downloading the client software, install and configure it using the appropriate method for your operating system (see Installing the StorNext Client on Linux or Unix on page 31 or Installing the StorNext Client on Windows on page 39).

Note: Before installing the StorNext client software, you must install and configure StorNext on an MDC. For more information, see Chapter 1, Installing StorNext.

Client System Requirements

To run the StorNext client software, the client system must meet all operating system and hardware requirements.

Operating System Requirements

The operating systems, releases and kernels, and hardware platforms supported by the StorNext client software are presented in Table 7. Make sure the client system uses a supported operating system and platform, and if necessary update to a supported release or kernel version before installing StorNext.

Table 7 StorNext Client Supported Platforms

Operating System	Release or Kernel	Platform
Windows XP	SP2 and SP3	x86 32-bit x86 64-bit
Windows 2003 Server	R2 SP2	x86 32-bit x86 64-bit
Windows Vista	SP1 and SP2	x86 32-bit x86 64-bit
Windows 2008 Server	SP1, SP2 and R2	x86 32-bit x86 64-bit
Windows 7	n/a	x86 32-bit x86 64-bit
Red Hat Enterprise Linux 4	2.6.9-67.EL (Update 6) 2.6.9-78.EL (Update 7) 2.6.9-89 EL (Update 8)	x86 32-bit x86 64-bit
Red Hat Enterprise Linux 5	2.6.18-53.EL (Update 1) 2.6.18-92.EL (Update 2) 2.6.18-128.1.1.EL (Update 3) 2.6.18-164.EL (Update 4)	x86 64-bit
SUSE Linux Enterprise Server 10	2.6.16-46-0.12 (SP1) 2.6.16-60-0.27 (SP2) 2.6.16.60-0.54.5 (SP3)	x86 32 bit x86 64-bit

Operating System	Release or Kernel	Platform
SUSE Linux Enterprise Server 11	2.6.27.19-5	x86 64-bit
Sun Solaris 10	Generic 127128-11	Opteron x86 64-bit and Intel x86 64-bit
	Generic 120011-14	Sparc 64-bit
HP-UX®	11i v3 (Requires the "0909 Patch set")*	Itanium-64
IBM AIX®	6.1	64-bit Power Architecture
Apple Mac OS® X	10.5.8 w/Xsan 2.1.1, StorNext MDC	x86 32-bit PowerPC
	10.6.1 w/Xsan 2.2, StorNext MDC	x86 32-bit x86 64-bit
	10.6.1 w/Xsan 2.2.1, StorNext MDC	x86 32-bit x86 64-bit

* StorNext 4.0 supports HP-UX release HP11iv3. If you do not want to upgrade to HP11iv3 and want to keep clients at the HP11iv2 level, those clients must remain on StorNext 3.5.

The other option is to upgrade both HP-UX clients and StorNext in the following order:

1. Uninstall StorNext 3.5
2. Upgrade to HP11iv3
3. Install StorNext 4.0.

Hardware Requirements

To install and run the StorNext client software, the client system must meet the following minimum hardware requirements.

For SAN (FC-attached) clients or for distributed LAN clients:

- 1 GB RAM

- 500 MB available hard disk space

For SAN clients acting as a distributed LAN server:

- 2 GB RAM
- 500 MB available hard disk space

Note: Distributed LAN servers may require additional RAM depending on the number of file systems, distributed LAN clients, and NICs used. See "Distributed LAN Server Memory Tuning" in the *StorNext User's Guide* for Distributed LAN Server memory tuning guidelines.

StorNext File System Client Buffer Cache

As of StorNext 3.0, multiple mounted file systems typically share a single buffer cache. A StorNext client buffer cache is created for each different **cachebufsize**. By default, all file systems have the same **cachebufsize** of 64K, so they all share the same buffer cache.

Note: These settings do not apply to Apple Xsan clients, which do not use the StorNext buffer cache.

The amount of memory consumed by default for each **cachebufsize** depends on the platform type and the amount of memory in the system. Table 8 shows the default amount of memory consumed by **cachebufsize**.

Table 8 Default Memory Consumed by cachebufsize

Platform Type	<=2GB Memory	>2GB Memory
32-bit Windows	32MB	64MB
32-bit Linux	64MB	128MB
All other platforms	64MB	256MB

To see information about the buffer cache after mounting file systems, use the **cvdb(1)** command with the **-b** option. To change the amount of memory used by the buffer cache at mount time, use the **buffercachecap** parameter.

On Windows, the non-paged pool is used for buffer cache memory until it consumes up to 64 megabytes (32-bit systems) or 64 gigabytes (64-bit systems). Any additional buffer cache memory comes from the paged pool.

Downloading the StorNext Client Software

The StorNext client software can be downloaded from a metadata controller. (To download the client software, the client system must have network access to the MDC.)

Note: You can download the client software only from MDCs running Linux.

1. On the client system, point a web browser to the URL (host name and port number) of the MDC. For example, **http://servername:81**

   Use one of the following web browsers to access the MDC (make sure pop-up blockers are turned off):

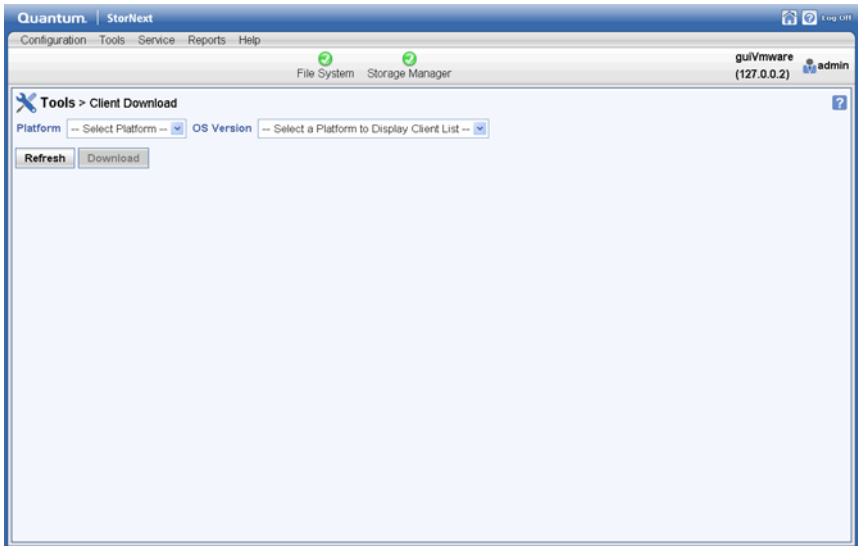
   - Internet Explorer 6.0 or later (including 7.0)
   - Mozilla Firefox 1.5 or later (including 2.0 or later)

2. When prompted, type the username and password for the MDC, and then click **OK**. (The default username is **admin**, and the default password is **password**.)

   The StorNext home page appears.

1. Choose **Client Download** from the **Tools** menu. The **Tools > Client Download** screen appears.

Figure 4 Client Download Screen

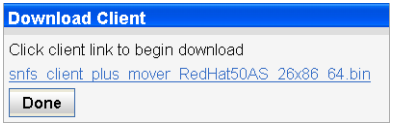


2 Select from the **Platform** list the desired operating system.

3 Select from the **OS Version** list the desired operating system version corresponding to the platform you selected.

4 When a window appears containing a link to the client software download location, click the link to begin downloading.

Figure 5 Client Download Link



5 Click **Download** to begin the process.

6 When prompted, choose the **Save to Disk** option, and then click **OK**.

7 Browse to the location where you want to save the file, and then click **Save**.

8 After the client package has been saved, click **Done**.

9 Continue with the installation procedure for your operating system as described in the StorNext Installation Guide or the online help.

10 After the download is complete, click **Done** to close the **Download Client Software** window.

Continue with the correct procedure for your operating system:

- Installing the StorNext Client on Linux or Unix on page 31
- Installing the StorNext Client on Windows on page 39.

Installing the StorNext Client on Linux or Unix

After downloading the StorNext client software from the metadata controller, install it on the client system using the correct procedure for your operating system:

- Installing the StorNext Client on Linux on page 31
- Installing the StorNext Client on Sun Solaris on page 34
- Installing the StorNext Client on HP-UX on page 35
- Installing the StorNext Client on IBM AIX on page 37

Note: Regardless of the operating system platform onto which you are installing, you must view and accept the StorNext end user license agreement (EULA) before you can install StorNext.

Installing the StorNext Client on Linux

To run the StorNext client software on Red Hat Linux or SUSE Linux Enterprise, first install the client software package, and then configure the client.

Caution: Before installing the StorNext client software, you must install the kernel source code. You can install the kernel source code by using the installation disks for your operating system.

1 Log on to the client system as **root**.

2 Change to the directory where the client software archive file you downloaded from the MDC is located.

3 Run the client software binary. At the command prompt, type:

```
./<archive name>
```

where `<archive name>` is the name of the software archive file you downloaded from the MDC. (For example, the file for RH5 is `sn_dsm_linuxRedHat50AS_x86_64_client.bin`)

4 When you are presented with the StorNext End User License Agreement (EULA), press **<ENTER>** to read the EULA. After reading the EULA enter **y** (lower case only) to accept the EULA, or enter any other key to decline.

After you accept the EULA, the client software is extracted to `/tmp/stornext`.

5 Change directories to `/tmp/stornext.`

6 List the packages extracted from the software archive file. At the command prompt, type:

ls -l

Identify the correct package to install. The correct package begins with **snfs-client** and ends with the **.rpm** file name extension.

Note: The file that ends with **.rpm.md5sum** is a checksum file, not the client software package.

---

7 Install the client software package and the snfs rpms. At the command prompt, type:

rpm -ivh <package_name> <snfs-client package>

where **<package_name>** is the name of the client software package you identified in Step 6.

8 Configure the boot order for system services. Do one of the following:

- For Red Hat Linux, at the command prompt, type:

  ```
  chkconfig --level 3456 cvfs on
  ```

- For SUSE Linux Enterprise, at the command prompt, type:

  ```
  chkconfig -s raw 235
  chkconfig -s cvfs 345
  ```

9 Edit the **/usr/cvfs/config/fsnameservers** text file to contain the IP address of the MDC the client will connect to.

The **fsnameservers** file on the client must be exactly the same as on the MDC. If the **fsnameservers** file does not exist, use a text editor to create it.

Note: The **fsnameservers** file only needs to contain the name of the MDC. For example, if the MDC is named snserver, then the **fsnameservers** file should contain a single line with a single word: snserver

---

10 Create a mount point for the file system. At the command prompt, type:

```
mkdir -p <mount point>
```

```
chmod 777 <mount point>
```

where *<mount point>* is the directory path where you want the file system to be mounted. For example: **/stornext/snfs1**

11 Configure the file system to automatically mount after reboot. To do this, edit the **/etc/fstab** file so that it contains the following line:

```
<file system> <mount point> cvfs verbose=yes 0 0
```

where *<file system>* is the name of the StorNext file system and *<mount point>* is the directory path created in Step 10.

12 Reboot the client system.

After reboot, the StorNext file system is mounted at the mount point you specified.

Note: To manually mount a file system, at the command prompt, type:

```
mount -t cvfs <file system> <mount point>
```

where *<file system>* is the name of the StorNext file system and *<mount point>* is the directory path where you want the file system to be mounted.

Installing the StorNext Client on Sun Solaris

To run the StorNext client software on Sun Solaris, first install the client software package, and then configure the client.

1. Log on to the client system as **root**.
2. Change to the directory where the client software archive file you downloaded from the MDC is located.
3. Run the client software binary. At the command prompt, type:

   `./<archive name>`

   where `<archive name>` is the name of the software archive file you downloaded from the MDC. (For example, the file for RH5 is `sn_dsm_linuxRedHat50AS_x86_64_client.bin`)
4. When you are presented with the StorNext End User License Agreement (EULA), press **<ENTER>** to read the EULA. After reading the EULA enter **y** (lower case only) to accept the EULA, or enter any other key to decline.

   After you accept the EULA, the client software is extracted to `/tmp/stornext`.
5. Change directories to `/tmp/stornext.`
6. Install the client software package. At the command prompt, type:

   `pkgadd -d .`
7. Type `1` to select the **ADICsnfs** package.
8. Type `y` to confirm installation of the **ADICsnfs** package. When installation is complete, type `q` to quit the installation program.
9. Edit the **/usr/cvfs/config/fsnameservers** text file to contain the IP address of the MDC the client will connect to.

   The **fsnameservers** file on the client must be exactly the same as on the MDC. If the **fsnameservers** file does not exist, use a text editor to create it.

   Note: The **fsnameservers** file only needs to contain the IP address of the MDC. For example, if the IP address for the MDC is 192.168.143.1, then the **fsnameservers** file should contain a single line with 192.168.143.1

10 Create a mount point for the file system. At the command prompt, type:

```
mkdir -p <mount point>
```

```
chmod 777 <mount point>
```

where <*mount point*> is the directory path where you want the file system to be mounted. For example: **/stornext/snfs1**

11 Configure the file system to automatically mount after reboot. To do this, edit the **/etc/vfstab** file so that it contains the following line:

```
<file system> - <mount point> cvfs 0 auto rw
```

where <*file system*> is the name of the StorNext file system and <*mount point*> is the directory path created in Step 10.

12 Reboot the client system.

After reboot, the StorNext file system is mounted at the mount point you specified.

Note: To manually mount a file system, at the command prompt type:

```
mount -F cvfs <file system> <mount point>
```

where <*file system*> is the name of the StorNext file system and <*mount point*> is the directory path where you want the file system to be mounted.

Installing the StorNext Client on HP-UX

To run the StorNext client software on HP-UX, first install the client software package, and then configure the client.

1 Log on to the client system as **root**.

2 Change to the directory where the client software archive file you downloaded from the MDC is located.

3 Run the client software binary. At the command prompt, type:

```
./<archive name>
```

where <archive name> is the name of the software archive file you downloaded from the MDC. (For example, the file for RH5 is sn_dsm_linuxRedHat50AS_x86_64_client.bin)

4 When you are presented with the StorNext End User License Agreement (EULA), press **<ENTER>** to read the EULA. After reading the EULA enter **y** (lower case only) to accept the EULA, or enter any other key to decline.

   After you accept the EULA, the client software is extracted to `/tmp/stornext`.

5 Change directories to `/tmp/stornext.`

6 List the packages extracted from the software archive file. At the command prompt, type:

   ```
   ls -l
   ```

   Identify the correct package to install. The correct package begins with **snfs-client** and ends with the **.depot** file name extension.

7 Install the client software package. At the command prompt, type:

   ```
   swinstall -s <package path and name> -x
   mount_all_filesystems=false \*
   ```

   where <*package path and name*> is the full path and name of the client software package you identified in Step 6.

8 Edit the **/usr/cvfs/config/fsnameservers** text file to contain the IP address of the MDC the client will connect to.

   The **fsnameservers** file on the client must be exactly the same as on the MDC. If the **fsnameservers** file does not exist, use a text editor to create it.

   Note: The **fsnameservers** file only needs to contain the name of the MDC. For example, if the MDC is named snserver, then the **fsnameservers** file should contain a single line with a single word: snserver

9 Create a mount point for the file system. At the command prompt, type:

   ```
   mkdir -p <mount point>
   ```

   ```
   chmod 777 <mount point>
   ```

   where <*mount point*> is the directory path where you want the file system to be mounted. For example: **/stornext/snfs1**

10 Configure the file system to automatically mount after reboot. To do this, edit the **/etc/fstab** file so that it contains the following line:

```
<mount point> <mount point> cvfs rw,fsname=<file system>
0 0
```

where *<mount point>* is the directory path created in Step 9 and *<file system>* is the name of the StorNext file system.

11 Reboot the client system.

After reboot, the StorNext file system is mounted at the mount point you specified.

Note: To manually mount a file system, at the command prompt, type:

```
mount -F cvfs <mount point> <mount point>
```

where *<mount point>* is the directory path where you want the file system to be mounted.

Installing the StorNext Client on IBM AIX

To run the StorNext client software on IBM AIX, first install the client software package, and then configure the client.

1 Log on to the client system as **root**.

2 Change to the directory where the client software archive file you downloaded from the MDC is located.

3 Run the client software binary. At the command prompt, type:

```
./<archive name>
```

where `<archive name>` is the name of the software archive file you downloaded from the MDC. (For example, the file for RH5 is `sn_dsm_linuxRedHat50AS_x86_64_client.bin`)

4 When you are presented with the StorNext End User License Agreement (EULA), press **<ENTER>** to read the EULA. After reading the EULA enter **y** (lower case only) to accept the EULA, or enter any other key to decline.

After you accept the EULA, the client software is extracted to `/tmp/stornext`.

5 Change directories to `/tmp/stornext.`

6 List the packages extracted from the software archive file. At the command prompt, type:

```
ls -l
```

Identify the correct package to install. The correct package begins with **snfs** and ends with the **.bff** file name extension.

7 Install the client software package. At the command prompt, type:

```
installp -ac -d <package name> all .
```

where *<package name>* is the name of the client software package you identified in Step 6.

8 Edit the **/usr/cvfs/config/fsnameservers** text file to contain the IP address of the MDC the client will connect to.

The **fsnameservers** file on the client must be exactly the same as on the MDC. If the **fsnameservers** file does not exist, use a text editor to create it.

Note: The **fsnameservers** file only needs to contain the name of the MDC. For example, if the MDC is named snserver, then the **fsnameservers** file should contain a single line with a single word: snserver

---

9 Create a mount point for the file system. At the command prompt, type:

```
mkdir -p <mount point>
```

```
chmod 777 <mount point>
```

where *<mount point>* is the directory path where you want the file system to be mounted. For example: **/stornext/snfs1**

10 Configure the file system to automatically mount. At the command prompt, type:

```
crfs -v cvfs -d <file system> -a verbose=yes -a type=cvfs
-A yes -m <mount point>
```

where *<file system>* is the name of the StorNext file system and *<mount point>* is the directory path created in Step 9.

The StorNext file system is mounted at the mount point you specified.

Note: To manually mount a file system, at the command prompt, type:

```
mount <mount point>
```

where *<mount point>* is the directory path where you want the file system to be mounted.

Installing the StorNext Client on Windows

The StorNext setup wizard guides you through the process of installing the StorNext client software on Windows machines.

Before installing StorNext, remove any previously installed versions (see Removing a Previous Version of StorNext on page 40).

When you are ready, use the setup wizard to install StorNext (see Running the Setup Wizard on page 41).

(Optional) After installation, restore the previous client configuration (see Restoring a Previous Client Configuration on page 45).

Note: You must log on as an Administrator to install StorNext.

If you are installing on Windows Vista, answer Yes to any messages asking if you want to run the installation process with administrative privileges.

Removing a Previous Version of StorNext

If a previous version of StorNext exists on the system, you must remove it before installing the new version.

1. Unzip the client software archive file you downloaded from the MDC.
2. Open the unzipped folder and double-click the client software installer file. This file is named **SnfsSetup32.exe** (32-bit systems) or **SnfsSetup64.exe** (64-bit systems).

   The **StorNext Installation** window appears (Figure 6).

Figure 6 StorNext Installation Window: Remove StorNext

3. Click **Remove StorNext**.

   A dialog box appears informing you that the current client configuration has been saved.
4. Note the name and location of the saved configuration file, and then click **OK**.
5. If prompted, click **Yes** to continue, and then click **OK** to finish the removal.
6. When the removal is complete, click **Yes** to reboot the system.

Note: After installing the new version of StorNext, you can restore the saved client configuration (see Restoring a Previous Client Configuration on page 45).

Running the Setup Wizard

To launch the setup wizard, access the folder for your operating system on the StorNext installation DVD, or use the client software installer you downloaded from the MDC.

1. If necessary, unzip the client software archive file you downloaded from the MDC.
2. Open the unzipped folder and double-click the client software installer file. This file is named **SnfsSetup32.exe** (32-bit systems) or **SnfsSetup64.exe** (64-bit systems).

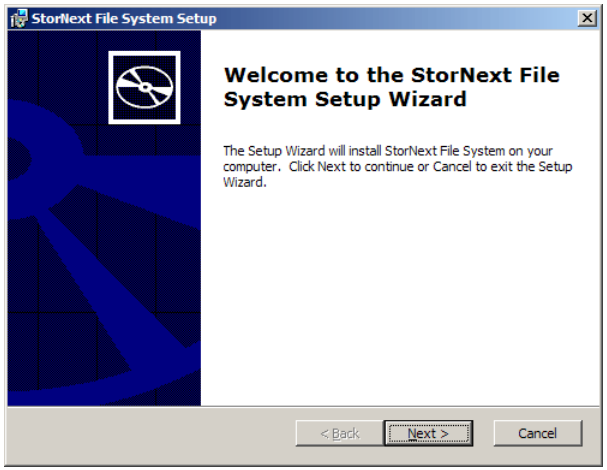
   The **StorNext Installation** window appears (Figure 7).

Figure 7 StorNext Installation Window: Install StorNext

3. Click **Install StorNext** to begin installation.

   The StorNext setup wizard appears (Figure 8).

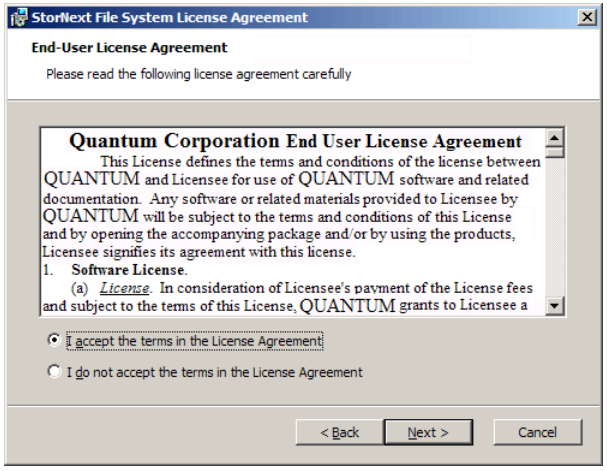
Figure 8 SNFS Setup: Welcome



4 Click **Next** to continue.

The **License Agreement** window appears (Figure 9).

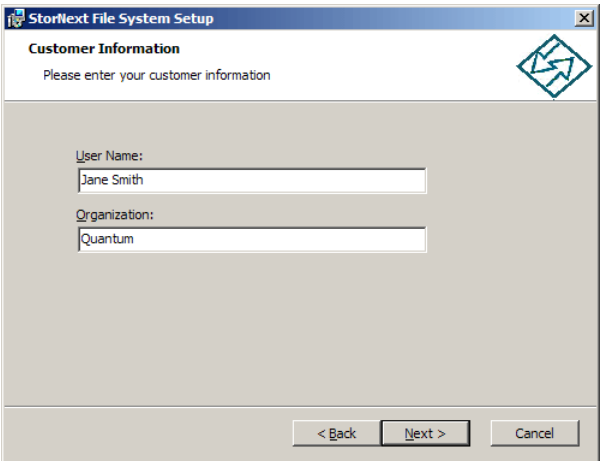
Figure 9 SNFS Setup: End-User License Agreement



5 Click the option to accept the license agreement, and then click **Next** to continue.

The **Customer Information** window appears (Figure 10).

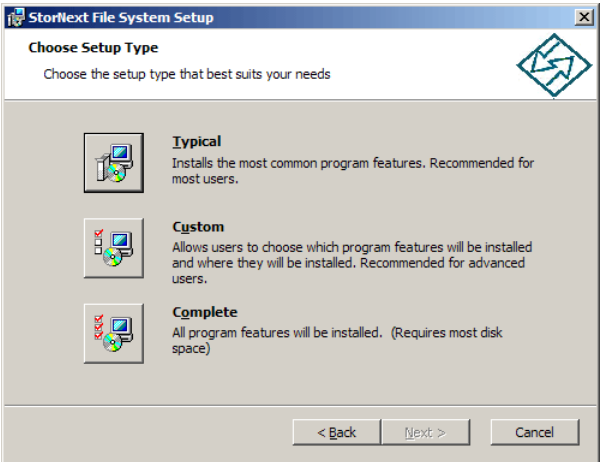
Figure 10 SNFS Setup: Customer Information



6 Type your name and the name of your company or organization in the boxes, and then click **Next** to continue.

The **Choose Setup Type** window appears, (Figure 11).

Figure 11 SNFS Setup: Choose Setup Type



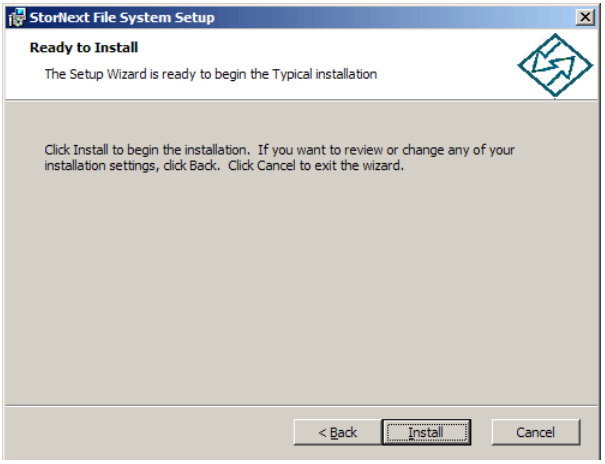
7 Click an installation option:

- **Typical:** (Recommended) Installs the StorNext client software and help files in the default location (**C:\Program Files\StorNext**).
- **Custom:** Lets you choose which components to install and specify an installation location. When ready, click **Next**.

- **Complete:** Same as a typical installation.

The **Ready to Install** window appears (Figure 12).

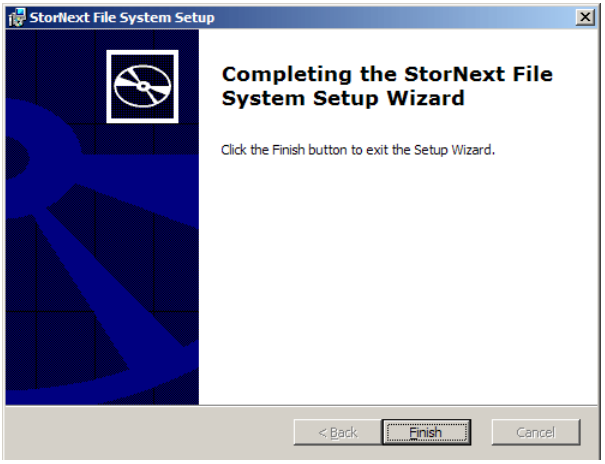
Figure 12 SNFS Setup: Ready to Install



8 Click **Install** to continue.

Wait while the setup wizard installs StorNext. When installation is complete, the **Finish** window appears (Figure 13).

Figure 13 SNFS Setup: Finish



9 Click **Finish** to close the setup wizard.

10 On the **StorNext Installation** window, click **Exit**.

You are prompted to reboot the system.

11 Click **Yes** to reboot.

You can now configure StorNext File System. For more information, see Chapter 3, Configuring StorNext.

Restoring a Previous Client Configuration

If you saved a client configuration file (for example, when removing a previous version of StorNext), you can import it after installing StorNext. This configures StorNext using the same settings as the previous installation.

1 If necessary, unzip the client software archive file you downloaded from the MDC.

2 Open the unzipped folder and double-click the client software installer file. This file is named **SnfsSetup32.exe** (32-bit systems) or **SnfsSetup64.exe** (64-bit systems).

The **StorNext Installation** window appears (Figure 14).

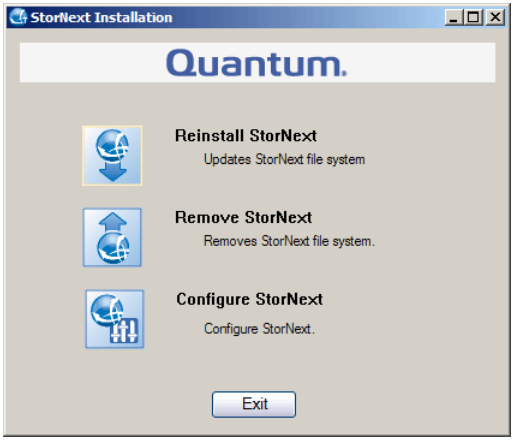


Figure 14 StorNext Installation Window: Configure StorNext

3 Click **Configure StorNext**.

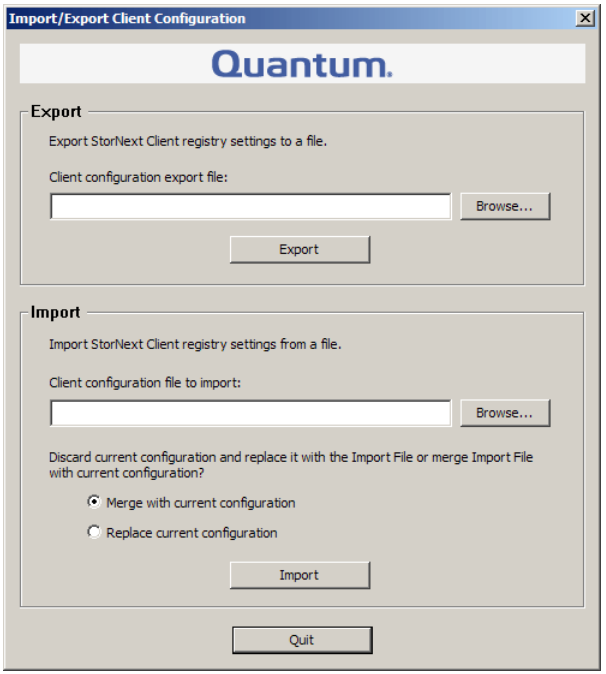
The **StorNext Configuration** window appears (Figure 15).

Figure 15 StorNext Configuration Window

4 Click **Import/Export Client Settings**.

The **Import/Export Client Configuration** window appears (Figure 16).

Figure 16 Import/Export Client Configuration Window



5 Under **Import**, click **Browse**. Locate the client configuration (***.reg**) file to import, and then click **Open**.

Client configuration files saved during removal of a previous version of StorNext are located in one the following directories:

- **C:\SNFS\config**
- **C:\Program Files\StorNext\config**

6 Click an option for handling current configuration settings:

- **Merge with current configuration:** The imported configuration is merged with the current configuration settings.
- **Replace current configuration:** The imported configuration overwrites the current configuration settings.

7 Click **Import**, and then click **Yes** to confirm.

A message appears informing you the configuration settings were successfully added to the registry.

8 Click **OK**, and then click **Quit** to close the **Import/Export Client Configuration** window.

Chapter 3
Configuring StorNext

This chapter describes how to configure StorNext after installation. To configure StorNext, enter license information and create one or more file systems. In addition, on metadata controllers (MDCs) running StorNext Storage Manager, you can add storage devices and media, create storage policies, and set up e-mail notifications. In addition, Linux or Windows SAN clients can be configured as a distributed LAN client or a distributed LAN server.

For systems running Unix or Linux, use the web-based GUI (graphical user interface) to access StorNext running on an MDC (see The StorNext GUI on page 50).

The StorNext Configuration Wizard guides you through the process of setting up StorNext (see The StorNext Configuration Wizard on page 52).

To configure a StorNext client as a distributed LAN client or server, edit the distributed LAN configuration file (see Configuring a Distributed LAN Client Server on page 90).

For systems running Windows, use the Windows-based configuration utilities to set up server and client features (see Windows Configuration Utilities on page 54).

The StorNext GUI

The StorNext GUI provides access to StorNext features and displays system status information. Use the StorNext GUI to create and work with file systems. In addition, on metadata controllers running Storage Manager, you can use the StorNext GUI to configure storage devices and media, and to set up storage policies.

Note: The StorNext GUI is available only for MDCs running Linux. For MDCs running on Windows, use the configuration utilities to configure StorNext (see Windows Configuration Utilities on page 54).

Accessing the StorNext GUI

The StorNext GUI is browser-based and can be remotely accessed from any machine with access to the StorNext server.

Use this procedure to access the StorNext GUI.

1 Open a Web browser.

Note: The following browsers have been tested to work with StorNext. Browsers not listed may work but are not recommended.

- Internet Explorer 7.x and 8.x
- FireFox 3.x

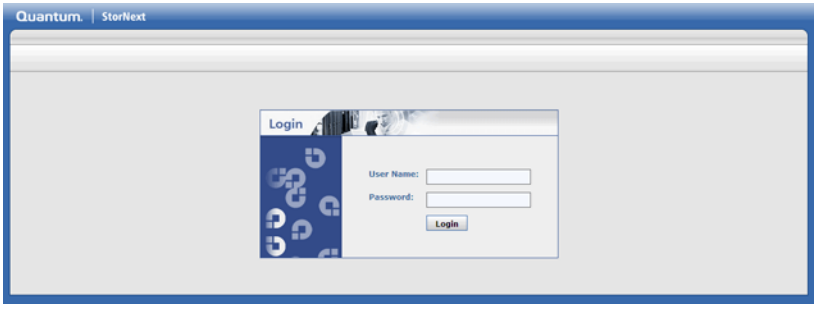
To ensure proper browser operation, all monitors must be set to display at a minimum resolution of 1024 x 768. If you use a popup blocker, be sure to enable pop-ups in order to ensure that StorNext displays properly.

2 In the browser's **Address** field, type the full address of the machine and its port number, and then press **Enter**. For example: http://<*machine name*>:<port number>. Use the name of the machine and port number you copied when you installed the StorNext software.

Note: Typically, the port number is 81. If port 81 is in use, use the next unused port number. (I.e., 82, 83, etc.)

After you enter the machine name and port number, the following window appears:

Figure 17 StorNext Login Window

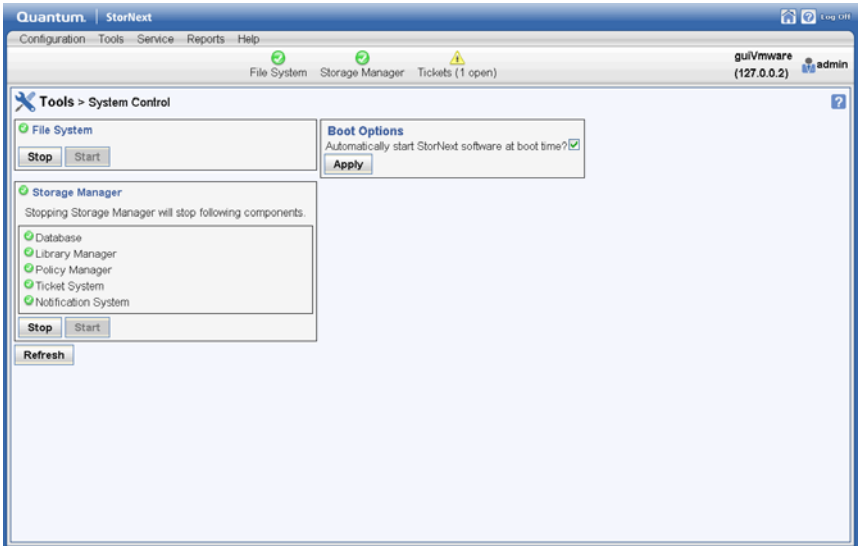


3 In the **User ID** field, type admin.

4 In the **Password** field, type password.

5 Click **Login**. The initial StorNext **System Control** screen appears.

Figure 18 System Control



6 On this screen you can determine if the StorNext File System and Storage Manager components are currently started. If not, click **Start** for each component to start them.

7 Click the home (house) icon in the upper right corner to go to the StorNext Home Page.

Note: When you log into StorNext for the first time, you might see a message warning you about a security certificate. Refer to the Quantum Knowledge Base for a permanent workaround to this issue. For a temporary solution, create a certificate exception that will allow you to log into StorNext without seeing the warning message during subsequent logins.

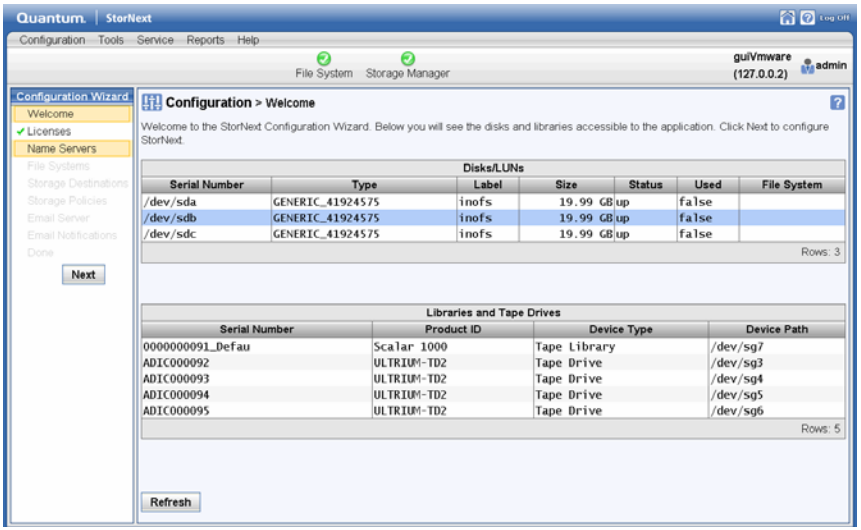
The StorNext Configuration Wizard

The first time you log onto the StorNext GUI, the Configuration Wizard appears. The wizard guides you step-by-step through the process of configuring StorNext.

Note: Using the Configuration Wizard to configure StorNext is optional. All configuration tasks can be accessed at any time using the StorNext Setup menu.

The Configuration Wizard is shown in Figure 19.

Figure 19 StorNext Configuration Wizard: Storage Manager



Using the Configuration Wizard

The Configuration Wizard consists of nine steps. The wizard lets you navigate between steps and tracks your progress as you complete each step.

- To go to the next step, click **Next**.
- To return to a previous step, click **Back** (or click the step in the list).

These are the steps you'll complete in the Configuration Wizard:

- **Step 1: Welcome**: View disks and libraries currently available for StorNext usage
- **Step 2: Licenses**: Enter StorNext License Information

- **Step 3: Name Servers**: Specify and order the machines acting as StorNext name servers
- **Step 4: File Systems**: Add a StorNext file system
- **Step 5: Storage Destinations:** Add a library, storage disks, and other storage destinations
- **Step 6**: **Storage Policies:** Add a Storage Manager or replication storage policy
- **Step 7: Email Server:** Specify an email server to handle StorNext notifications
- **Step 8: Email Notification:** Add email notifications recipients
- **Step 9: Done:** Signify that you are finished using the Configuration Wizard. You can also convert to a high availability (HA) system.

To see more detailed instructions for completing Configuration Wizard tasks, refer to "Chapter 3: The Configuration Wizard" in the *StorNext User's Guide*.

Windows Configuration Utilities

To configure StorNext File System or the StorNext client software on Windows operating systems, use the Windows configuration utilities. The configuration utilities let you set up a Windows-based metadata controller, configure a StorNext client, and work with StorNext file systems.

To run a configuration utility, on the Windows **Start** menu, click **All Programs > StorNext File System**, and then click the utility.

The StorNext client software includes the following configuration utilities:

- Client Configuration on page 55
- Disk Device Labeler on page 67
- License Identifier on page 69
- Simple File System Configuration on page 69
- Start File System Services on page 71

StorNext File System server includes all the configuration utilities installed by the client software as well as the following additional utilities:

Client Configuration

The Client Configuration utility lets you view and modify properties for the StorNext client software. Use the utility to specify mount points and mount options, set various cache settings, and configure a distributed LAN server or a distributed LAN client.

To run the Client Configuration utility, on the Windows **Start** menu, click **All Programs > StorNext File System > Client Configuration**.

The Client Configuration utility includes four tabs:

Click a tab to view or modify StorNext client properties. After making changes on one or more tabs, click **OK** to save the changes. A message appears prompting you to restart the system. Click **Yes**. Most changes do not take affect until the system is restarted.

Mount Point

StorNext file systems can be mapped to a drive letter or to a folder. When mapping to a folder, the folder must be empty or non-existent, and must exist within an NTFS file system (e.g., **C:\Mount\MySnfs**).

Use the **Mount Point** tab (Figure 20) to specify drive mappings. Table 9 describes the fields on the **Mount Point** tab.

Note: Make sure the host list is the same for all clients on the SAN. An incorrect or incomplete host list may prevent the client from connecting to the file system.

Figure 20 Client Configuration: Mount Point

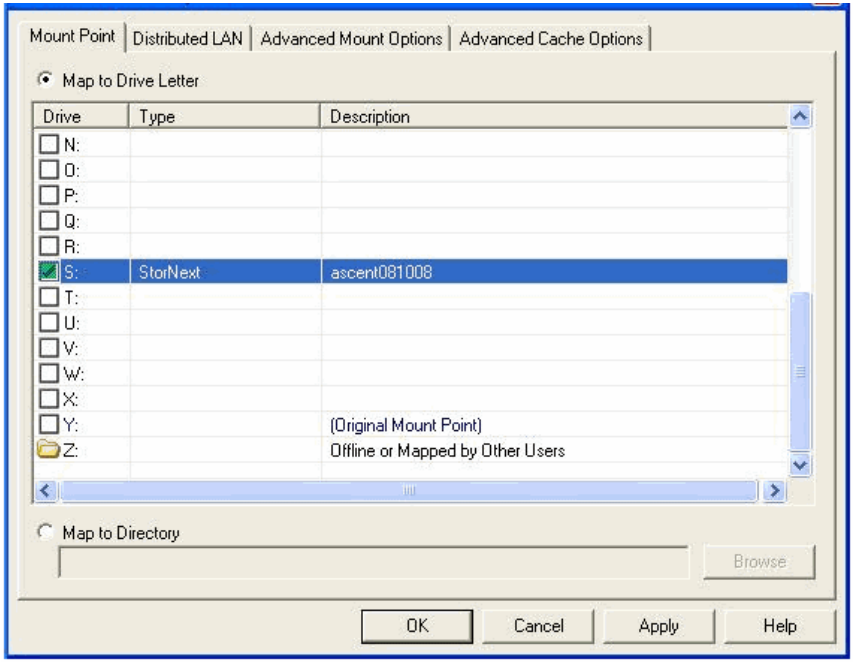


Table 9 Client Configuration: Mount Point

Field / Button	Description
Map to Drive Letter	Select Map to Drive Letter to map the drive to a specific drive letter.

Field / Button	Description
Drive (Letter)	Select the desired drive letter.
Map to Directory	Instead of selecting Map to Drive Letter, select Map to Directory. Enter a directory path, or click **Browse** to navigate to the directory path.

Note: If this is a Distributed LAN client or server, be sure to enter the parameters on the Distributed LAN tab.

Distributed LAN

The Distributed LAN tab is used to enable and configure StorNext Distributed LAN (DLAN) Clients and Servers for a file system. The Enable Client and Enable Server radio buttons determine whether the system is a Client or Server. Note that separate licensing is required for DLAN.

StorNext Distributed LAN Clients provide the same functionality as normal StorNext Clients but do not need to be connected directly to the SAN. They instead send I/O via the TCP/IP network to a Distributed LAN Server, which is a standard StorNext SAN client; a normal SAN client with DLAN Server Selected in its configuration. Metadata is sent over the TCP/IP network directly to the File System Metadata Controller (Server/MDC).

Distributed LAN Clients request the location of normal DLAN Servers via the Name Server.

- **Distributed LAN Server:** A distributed LAN server has direct access to data disks (using fibre channel) and makes the shared storage pool available to distributed LAN clients over a network.
- **Distributed LAN Client:** A distributed LAN client does not have direct access to data disks and instead accesses the shared storage pool by connecting to a distributed LAN server over a network. (Distributed LAN clients have full StorNext client functionality.)

Note: A StorNext SAN client configured as a distributed LAN server requires a dedicated system. Do not run other applications or services on a system configured as a distributed LAN server.

Use the **Distributed LAN** tab (Figure 21) to configure a distributed LAN server or a distributed LAN client. Table 10 describes the fields on the **Distributed LAN** tab.

Figure 21 Client Configuration: Distributed LAN

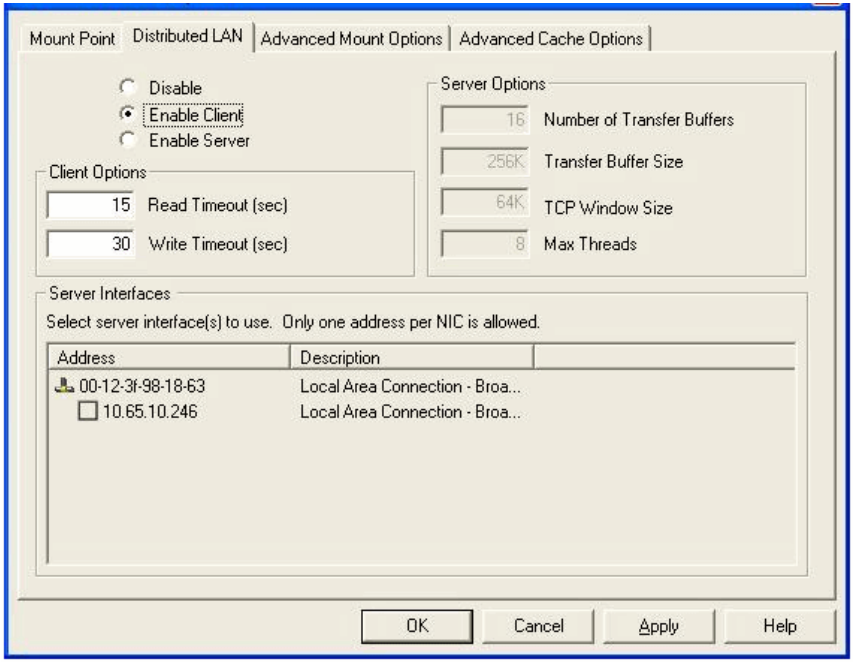


Table 10 Client Configuration: Distributed LAN

Field / Button	Description
Disable	Select this option to disable the StorNext client as a distributed LAN client or distributed LANserver.
Enable Client	Select to configure the StorNext client as a distributed LAN client. The StorNext client will connect to all available distributed LAN servers. (If you select this check box, all other fields on the tab become unavailable.)

Field / Button	Description
Enable Server	Select to configure the StorNext client as a distributed LAN server. The StorNext client will make the disks for this file system available to distributed LAN clients over the specified LAN.
Number of Transfer Buffers	Specify the number of socket transfer buffers used for Distributed LAN Client I/O.
Read Timeout (sec)	Specify how long the Distributed LAN Client will wait for a read operation to complete before retrying it on a different Distributed LAN Server.
Write Timeout (sec)	Specify how long the Distributed LAN Client will wait for a write operation to complete before retrying it on a different Distributed LAN Server.
Transfer Buffer Size (kb)	Specify the size in kilobytes of the socket transfer buffers used for Distributed LAN Client I/O.
TCP Window Size (kb)	Specify the size in kilobytes of the TCP window used for Distributed LAN Client I/O connections.
Max Threads	Specify the number of kernel threads that will be used to service Distributed LAN Client I/O requests.
Server Interfaces	To complete Distributed LAN Server configuration, select one or more network interfaces to be used for Distributed LAN Client traffic.

Advanced Mount Options

The **Mount Options** tab (Figure 22) displays the mount options for the drive currently selected on the **Drive Mappings** tab. Table 10 describes the fields on the **Distributed LAN** tab.

Caution: Changing the values on the **Mount Options** tab can affect system performance and stability. Do not change mount options unless instructed to do so by the Quantum Technical Assistance Center.

Figure 22 Client Configuration: Advanced Mount Options

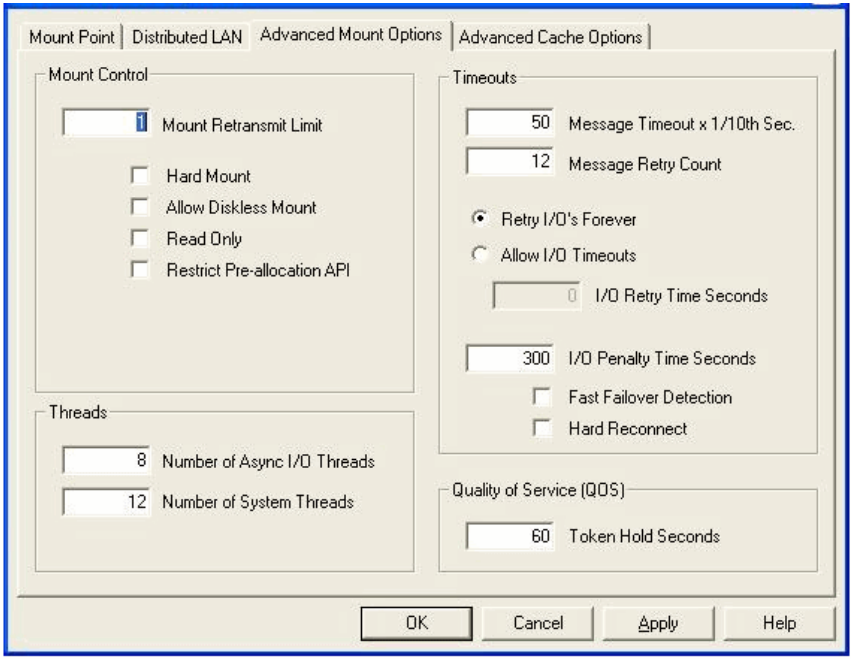


Table 11 Client Configuration: Advanced Mount Options

Field / Button	Description
Mount Retransmit Limit	Specify the number of times the driver will re-transmit a message to the FSM. After a successful connection, this value is no longer used. The default is 1.

Field / Button	Description
Hard Mount	When this box is checked, the driver attempts to mount the file system forever. The default is off (Soft Mount). Quantum recommends that Hard Mount NOT be used on Windows systems.
Allow Diskless Mount	When this box is checked, the file system can be mounted and accessed without all the disks being accessible in the file system stripe groups. In this mode, file metadata can be viewed and changed, but no data can be accessed. The default is to not allow a file system to be mounted unless all disks are accessible.
Read Only	When this box is checked, the file system is mounted in read-only mode. The default is off or not checked.
Restrict Pre-allocation API	When set to yes, non-administrator users are unable to use the preallocation ioctl.
Number of Async I/O Threads	Specify the number of threads created to be used as asynchronous I/O threads for user applications. The default is 8. The allowed range is 4 to 32.
Number of System Threads	Specify the number of threads created for use by the file system. The default value is 16.
Message TImeout x 1/10th Sec.	Specify (in tenths of a second) the amount of time before a message to the FSM is timed out. The default is 50 (five seconds).
Message Retry Count	Specify the number of attempts to be made for sending a message to the File System Service. The default is 12.
Retry I/Os Forever or Allow I/O Timeouts	Choose either Retry I/Os Forever or Allow I/O Timeouts, depending on whether you want to allow timeouts or keep trying I/O. The default value is to Retry I/Os Forever.

Field / Button	Description
I/O Retry Time Seconds	When you choose Allow I/O Timouts, this option controls the I/O retry behavior. This parameter establishes the amount of time that may elapse during an I/O retry sequence. An I/O retry sequence consists of the following: • Retry an I/O request across all available paths that are currently present. • Compare the current time to the Instantiation time of the I/O request, if at least <n> seconds have elapsed, return the I/O request in error, otherwise reset the paths used, and retry again.
I/O Penalty TIme Seconds	This option controls the Multi-Path I/O penalty value, where <n> is expressed in seconds with a minimum value of 1 and a maximum value of 15. This parameter establishes the amount of time a **Path_In_Error** will be bypassed in favor of an Operational Path during a Multi-Path selection sequence. If all paths are currently in the Path_In_Error state, the first available path will be selected, independent of the Path_In_Error state. The default value is 5 seconds.
Fast Failover Detection	When this box is checked, the client uses a more aggressive approach to detecting if the FSM has failed. With this option on, the client monitors the FSM heartbeat. If no heartbeat is detected within three (3) seconds, a failover request is initiated. This option is desirable for near realtime activities that cannot sustain more than 10 seconds of access loss to the server. Quantum does not recommend enabling this option for normal file system operations.

Field / Button	Description
Hard Reconnect	When this box is checked, the file system attempts to reconnect to the FSM forever. If this option is not enabled, the file system attempts to reconnect to the FSM for the number of times specified at the Mount Retransmit Limit field before failing the request. The default value is off.
Token Hold Seconds	The QOS Token Hold Time parameter is applicable only when using the StorNext Quality of Service (QOS) feature for real-time IO. This parameter determines the number of seconds a client stripe group holds on to a non-realtime I/O token during periods of inactivity. If no I/O is performed on a stripe group within the specified number of seconds, the token is released back to the FSM. The default value is 60 seconds. The parameter should be specified in five-second increments; if the parameter is not a multiple of five, it will be rounded up automatically.

Advanced Cache Options

The Data Buffer Cache keeps the file system data in memory on the local computer to speed up performance for small I/O operations.

The **Advanced Cache Options** tab (Figure 23) displays performance values that control how many file system lookup names are kept in memory.

Caution: Changing the values on the **Advanced Cache Options** tab can affect system performance and stability. Do not change cache parameters unless instructed to do so by the Quantum Technical Assistance Center.

Figure 23 Client Configuration: Advanced Cache Options

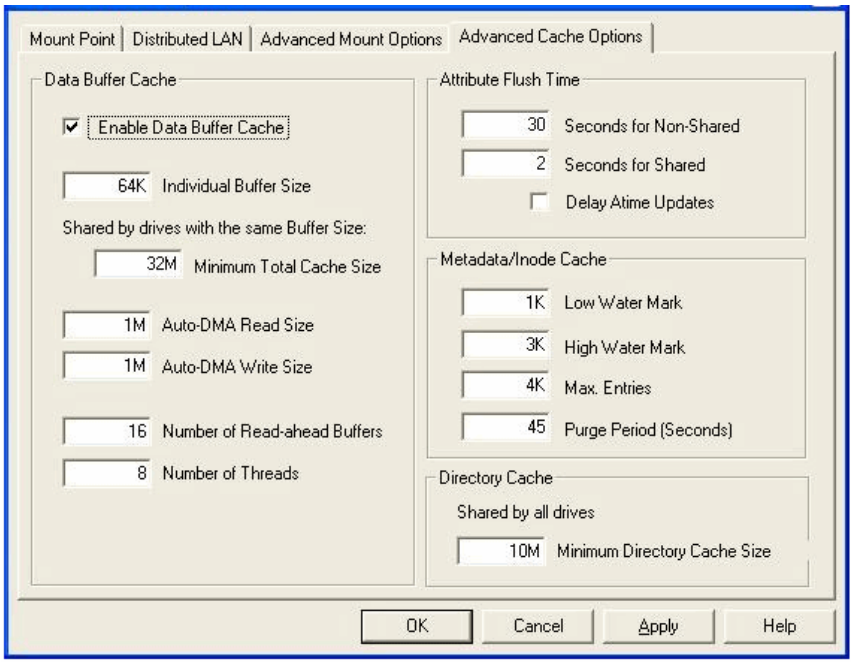


Table 12 Client Configuration: Advanced Cache Options

Field / Button	Description
Enable Data Buffer Cache	When this box is not checked, the file system will not buffer any data. All files will be read directly into the application's memory using DMA. Requests that are not properly aligned will be read into temporary buffers, then copied to the application's buffer. If Buffer Cache is disabled, the remaining Data Buffer Cache options are ignored.

Field / Button	Description
Individual Buffer Size	This option sets the size of each cache buffer. This determines the I/O transfer size used for the buffer cache. To determine optimal performance, try different sizes or contact your RAID vendor. This size must be divisible by the file system block size. The maximum value allowed is 512K and the minimum value allowed is the file system block size. The value default is 64K.
Minimum Total Cache Size	This value controls the amount of memory used to cache data from the file system. This parameter is shared by all file systems with the same block size.
Auto-DMA Read/ Write Sizes	Buffer Cache Auto DMA Read Length and Buffer Cache Auto DMA Write Length control the minimum I/O size for DMA transfers. That is, any smaller transfer will always go through the buffer cache. The default value is 1MB.
Number of Read-ahead Buffers	This option controls the size of the read-ahead window used by the buffer cache. The default value is 16 buffers with a maximum of 100. Using a value of 0 disables read-ahead.
Number of Threads	This option controls the number of dedicated background threads used by the buffer cache for performing I/O. The default value is 8, with a minimum of 1 and a maximum of 100. Using a larger value can increase I/O concurrency in the buffer cache and therefore improve performance.

Field / Button	Description
Seconds for Shared/ Non-Shared	The Attribute Flush Time parameters control the length of time attributes are held on the client before they are flushed back to the FSM (unless there is an explicit attribute request from the FSM). The default values are 30 seconds for non-shared files (Attribute Flush Time, non shared), and 2 seconds for shared files (Attribute Flush Time, shared). Setting these parameters lower will greatly increase the amount of metadata traffic between clients and the FSM. However, in circumstances where files are frequently shared on the SAN, setting the Attribute Flush Time Shared can result in other clients seeing size updates more frequently if a file is being written on one client and read on another. A value of zero is invalid, and will result in using the default setting.
Delay Atime Updates	When this box is checked, the file system delays Atime (access time) updates when reading a file to the time when the file is closed. This cuts down on FSM metadata updates at the expense of coherency.
Low Water Mark	Metadata Caches contain non-data information such as file names, sizes, update times, etc. Tuning the low and high water marks and the frequency of purge passes can help certain large mix applications.
High Water Mark	
Max. Entries	
Purge Period (Seconds)	
Minimum Directory Cache Size	This option sets the size of the directory cache. Directory entries are cached on the client to reduce client-FSM communications during directory reads. The default value is 10 MB.

Disk Device Labeler

The Disk Device Labeler utility lets you configure StorNext File System storage area network (SAN) disks. Use the Disk Device Labeler to create a list of disk labels, associated device names, and (optional) the sectors to use.

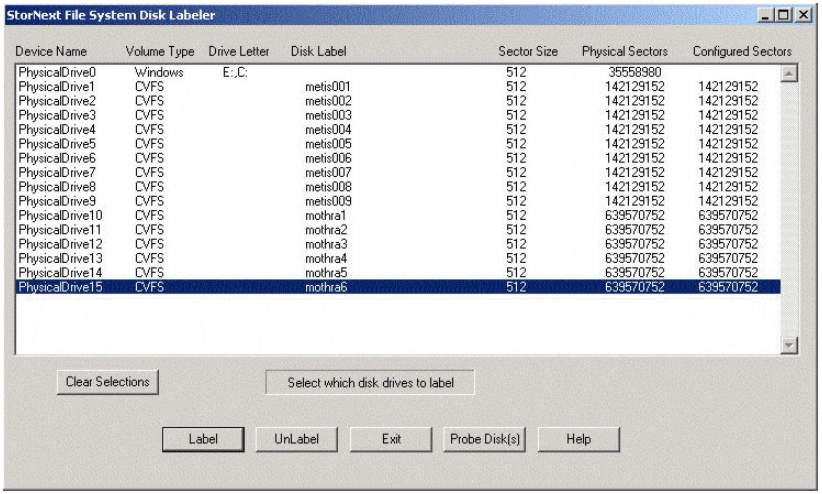
> **Note:** Run the Disk Device Labeler on a system that has visibility to all disk devices on the SAN.

The file system uses the volume labels to determine which disk drives to use. The label name written to a disk device must match the disk name specified in the Server Configuration utility. For more information, see Server Configuration on page 81.

> **Caution:** Modifying the label of a system disk may render the system inoperable and require you to repair the volume. Only label disk devices that are intended for use with the StorNext File System SAN.

To run the Disk Device Labeler utility, on the Windows **Start** menu, click **All Programs > StorNext File System > Disk Device Labeler**. The **Disk Labeler** window (Figure 24) appears.

Figure 24 Disk Device Labeler



On the **Disk Labeler** window, do one of the following:

- Click **Label** to write a volume label to the selected device(s) (see Labeling Disks below).
- Click **UnLabel** to remove the existing volume label from the selected device(s).
- Click **Exit** to quit the Disk Device Labeler utility.
- Click **Probe Disk(s)** to pulse the activity light of the selected device(s). Use this feature to correctly identify disks before labeling them.
- Click **Clear Selections** to deselect all devices in the list.

Labeling Disks

When you select one or more disks and click **Label**, a confirmation screen appears asking if you are sure you want to proceed. Click **OK** to continue. The **Disk Labeler** dialog box appears (Figure 25). Table 13 describes the fields on the on the **Disk Labeler** dialog box.

Figure 25 Disk Labeler Dialog Box

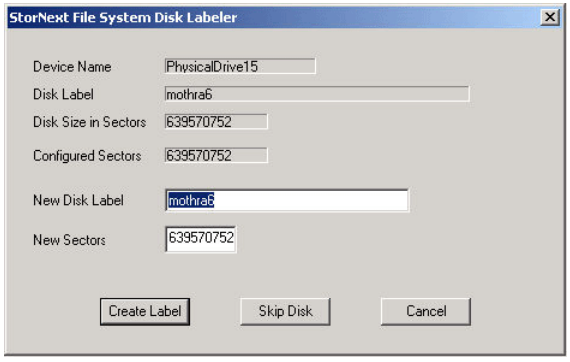


Table 13 Disk Labeler Dialog Box

Field / Button	Description
New Disk Label	Type the label for the disk.
New Sectors	(Optional) Type the number of sectors on the disk.
Create Label	Write the new label to the disk and proceed to the next selected disk.

Field / Button	Description
Skip Disk	Do not write a label to the disk and proceed to the next selected disk.
Cancel	Close the **Disk Labeler** dialog box.

License Identifier

Use the License Identifier utility to display the host license identifier. The host license identifier is required to obtain a permanent license for StorNext.

To run the License Identifier utility, on the Windows **Start** menu, click **All Programs > StorNext File System > License Identifier**. A dialog box displays the host license identifier. Record this information.

To obtain a permanent license, contact the Quantum Technical Assistance center at licenses@quantum.com and provide the following information:

- The product serial number from the StorNext box or DVD.
- The number of client machines connecting to the MDC.
- The host license identifier you recorded.

A Quantum support representative will send you a **license.dat** file. Copy the file to the **C:\Program Files\StorNext\config** directory. (If there is a temporary license file, rename the file or move it to a backup location.)

Note: To prevent accidentally losing a valid license, be sure to back up or rename any existing **license.dat** files.

Simple File System Configuration

The Simple File System Configuration utility can be used instead of the Server Configuration utility to configure a basic StorNext file system with a single stripe group.

Note: Before configuring a file system, you should label disk devices. For more information, see Disk Device Labeler on page 67.

To run the Simple File System Configuration utility, on the Windows **Start** menu, click **All Programs > StorNext File System > Simple File**

System Configuration. The **Simple Configuration Setup** window appears (Figure 26). Table 14 describes the fields on the **Simple Configuration Setup** window.

To configure a simple file system, select the disks to use in the configuration. Specify settings (file system name, block size, stripe size, and maximum connections), and then click **Configure**.

Figure 26 Simple File System Configuration

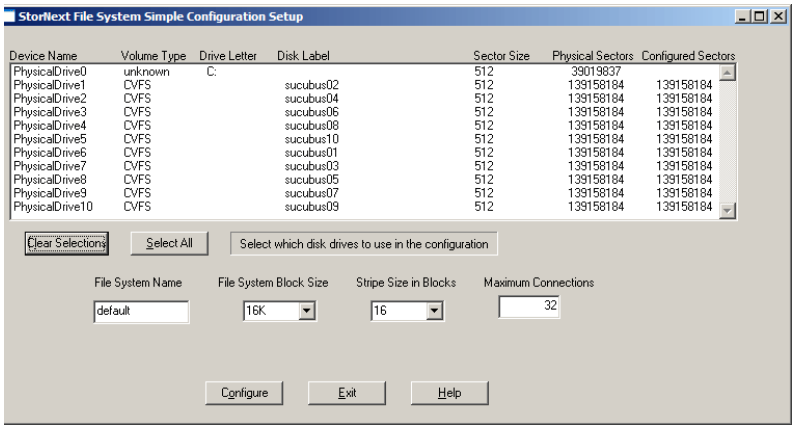


Table 14 Simple File System Configuration

Field / Button	Description
Clear Selections	Click to deselect all devices in the list.
Select All	Click to select all devices in the list.
File System Name	Type the name of the file system. This is the name used by clients when establishing a mount point for the file system.
File System Block Size	Select the file system block size (in bytes). This is the minimum allocation size used by the file system.
Stripe Size in Blocks	Select the stripe size (in blocks). This is the number of file system blocks to write before switching to the next disk in the stripe group.

Field / Button	Description
Maximum Connections	Type the maximum number of clients that can simultaneously mount the file system. (This value may be overridden by values in your product license code.)
Configure	Click to save the configuration using the current settings. The configuration file is saved in the StorNext configuration directory.
Exit	Click to quit the Simple File System Configuration utility.

Start File System Services

The Start File System Services utility starts all StorNext services on an MDC or StorNext client.

The StorNext services must be running on the MDC for file systems to be active and available. In addition, the StorNext services must be running to use the StorNext configuration utilities and to mount file systems using the client software.

To start StorNext File System services, on the Windows **Start** menu, click **All Programs > StorNext File System > Start File System Services**.

Stop and Remove File System Services

The Stop and Remove File System Services utility stops all StorNext services on an MDC or StorNext client, and also removes registry entries that automatically start the services on bootup.

To stop and remove StorNext File System services, on the Windows **Start** menu, click **All Programs > StorNext File System > Stop and Remove File System Services**.

To start the StorNext services again, you must use the Start File System Services utility. Rebooting the system will not restart services. For more information, see Start File System Services.

Stop File System Services

The Stop File System Services utility stops all StorNext services on an MDC or StorNext client.

To stop StorNext File System services, on the Windows **Start** menu, click **All Programs > StorNext File System > Stop System Services**.

To start the StorNext services again, reboot the system or use the Start File System Services utility. For more information, see Start File System Services.

Version Information

The Version Information utility displays information about the currently installed version of the StorNext server and/or client software, such as the version number, build number, and platform.

To run the Version Information utility, on the Windows **Start** menu, click **All Programs > StorNext File System > Version Information**. A dialog box displays version information for the StorNext server and/or client software installed on your system.

Check (Read-Only) a File System

The Check (Read-Only) a File System utility allows you to check a StorNext file system for metadata corruption (due to a system crash, bad disk, or other failure). Run the utility on an MDC that contains the file system you want to check.

To check a file system, on the Windows **Start** menu, click **All Programs > StorNext File System > Check (Read-Only) a File System**. Type the number that corresponds to the file system you want to check, and then press **<Enter>**.

Because the file system check is run in read-only mode, any problems that exist are not repaired. If the utility identifies errors or corruption in metadata, you must repair the file system (see Repair a File System on page 74).

File System Startup List

The File System Startup List utility lets you modify the File System Manager (FSM) services list and set file system priority.

The File System Manager is a process that manages the name space, allocations, and metadata coherency for a file system. Each file system uses its own FSM process. When there are multiple file systems (and therefore multiple FSM processes), the FSM services list controls which FSM processes are run when the server starts up, and also sets the priority for each file system (for failover configurations).

To run the File System Startup List utility, on the Windows **Start** menu, click **All Programs > StorNext File System > File System Startup List**. The **FSM Services List** window appears (Figure 27). Table 15 describes the fields on the **FSM Services List** window.

Figure 27 File System Startup List

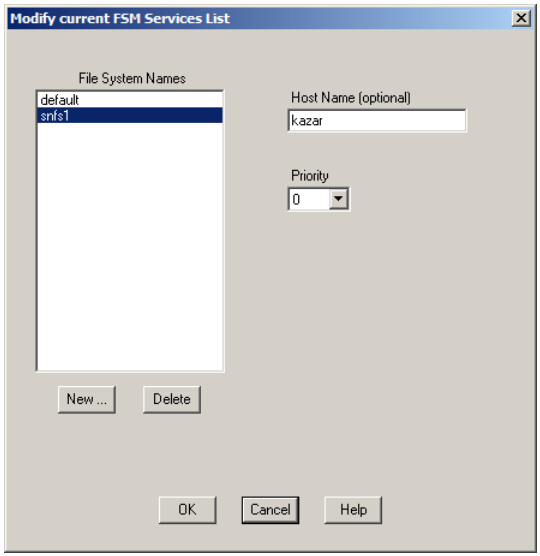


Table 15 File System Startup List

Field / Button	Description
New	Click to add a file system to the FSM services list. Type the name of the file system and click **OK**.
Delete	Click to remove the selected file system from the FSM services list.
Host Name	(Optional) Type the name of the host on which the file system is running.
Priority	(Optional) Select the priority for the selected file system. This priority is used for failover configurations. **0** is highest priority and **9** is lowest priority.
OK	Click to save changes to the FSM services list.

Field / Button	Description
Cancel	Click to quit the File System Startup List utility without saving changes.

Re-initialize a File System

The Re-initialize a File System utility allows you to initialize an existing file system. Initializing a file system prepares it for use.

Caution: Re-initializing a file system will destroy all data on the file system.

To initialize a file system, on the Windows **Start** menu, click **All Programs > StorNext File System > Re-initialize a File System**. Type the number that corresponds to the file system you want to re-initialize, and then press **<Enter>**.

Repair a File System

The Repair a File System utility lets you repair corrupted metadata on a file system. Repair a file system if errors were identified when checking the file system (see Check (Read-Only) a File System on page 72).

The file system must be inactive in order to be repaired. To stop a file system, use the Server Administration utility (see Server Administration).

To repair a file system, on the Windows **Start** menu, click **All Programs > StorNext File System > Repair a File System**. Type the number that corresponds to the file system you want to repair, and then press **<Enter>**.

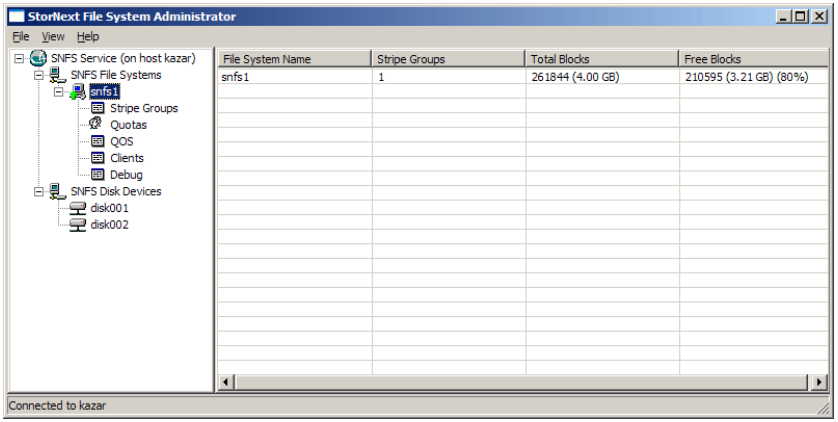
Note: Multiple repair passes may be needed to fully fix a damaged file system. Run the repair utility repeatedly until no problems are reported by the utility.

Server Administration

The Server Administration utility lets you view and modify stripe group properties and set quotas. A stripe group is a logical storage unit made up of one or more disks. A quota is a space limit that is set for specified users or groups.

To run the Server Administration utility, on the Windows **Start** menu, click **All Programs > StorNext File System > Server Administration**. The **Administrator** window appears (Figure 28). The left pane shows file systems running on the currently connected MDC. Expand a file system to see stripe groups, quotas, and other properties.

Figure 28 Server Administration



On the **Administrator** window, do one of the following:

- To connect to an MDC, click **File > Connect to**. Type the host name and click **OK**. File systems on the server appear in the left pane.
- To start a file system, click it in the left pane. Click **File > Start File System**, and then click **Start**.
- To activate a file system, click it in the left pane. Click **File > Activate File System**, and then click **Activate**.
- To stop a file system, click it in the left pane. Click **File > Stop File System**, and then click **Stop**.
- To update the list of file systems in the left pane, click **View > Refresh**.

For more information about viewing and modifying file system properties and quotas, see the following sections:

- File System Properties on page 76
- Stripe Group Properties on page 77
- Quota Properties on page 79
- Quality of Service Information on page 80

- Clients Information on page 80

Note: Do not change settings on the **Debug Settings** dialog box unless directed to do so by the Quantum Technical Assistance Center.

File System Properties

To view or change file system properties, click a file system in the left pane, and then click the file system name in the right pane. The **File System Properties** dialog box appears (Figure 29). Table 16 describes the fields on the **File System Properties** dialog box.

After making changes, click **OK**. (Not all fields can be modified on this dialog box.)

Figure 29 Server Administration: File System Properties

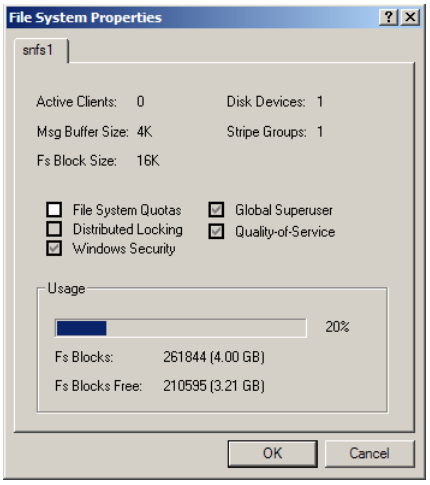


Table 16 Server Administration: File System Properties

Field / Button	Description
Active Clients	The number of active clients on the file system.
Msg Buffer Size	The size of the message buffer.
Fs Block Size	The file system block size.

Field / Button	Description
Disk Devices	The number of disk devices in the file system.
Stripe Groups	The number of stripe groups in the file system.
File System Quotas	Select to enable file system quotas.
Distributed Locking	Indicates if distributed locking is enabled or disabled.
Windows Security	Indicates if Windows security is enabled or disabled.
Global Superuser	Indicates if the global superuse is enabled or disabled.
Quality-of-Service	Indicates if Quality of Service is enabled or disabled.
Usage	Displays the amount of used and free storage space in the file system.

Stripe Group Properties

To view or change stripe group properties, expand a file system in the left pane, click **Stripe Groups**, and then click the stripe group name in the right pane. The **Stripe Group Properties** dialog box appears (Figure 30). Table 17 describes the fields on the **Stripe Group Properties** dialog box.

After making changes, click **OK**. (Not all fields can be modified on this dialog box.)

Figure 30 Server Administration: Stripe Group Properties

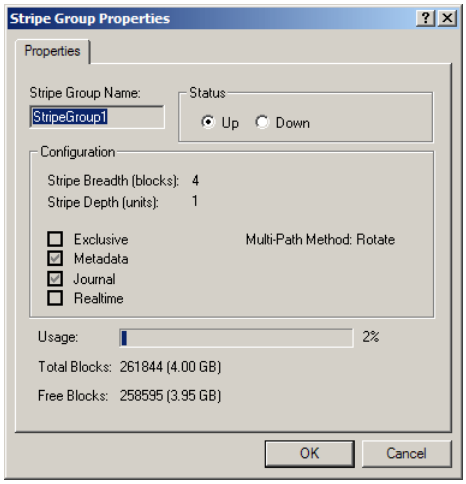


Table 17 Server Administration: Stripe Group Properties

Field / Button	Description
Stripe Group Name	The name of the stripe group.
Status	Shows the current status of the stripe group. Click **Up** to make the stripe group active or click **Down** to make the strip group inactive.
Stripe Breadth	The number of file system blocks to write before switching to the next disk in the stripe group.
Stripe Depth	The number of disks in the stripe group.
Exclusive	Indicates if only specified file types (associated with the stripe group affinities) can be stored on the stripe group.
Metadata	Indicates if file system metadata can be stored on the stripe group.
Journal	Indicates if the file system logging journal can be stored on the stripe group.
Realtime	(no longer supported)

Field / Button	Description
Multi-Path Method	Indicates the method the file system uses to access the disk: round, static, or sticky.
Usage	Displays the amount of used and free storage space in the stripe group.

Quota Properties

Quotas define the amount of storage space a specified user or group is allowed to own. You can set hard or soft quota limits.

To view or change quota properties, expand a file system in the left pane, and then click **Quotas**. The **Quota Properties** dialog box appears (Figure 31). There are two tabs:

- Set quotas for users on the **User Quotas** tab.
- Set quotas for groups on the **Group Quotas** tab.

Table 18 describes the fields on the **User Quotas** and **Group Quotas** tabs. After making changes, click **OK**. (Not all fields can be modified on this dialog box.)

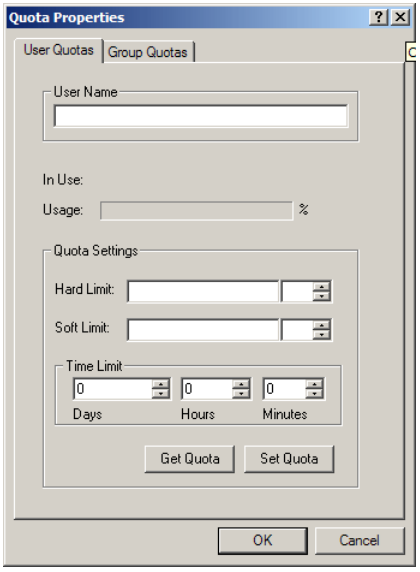


Figure 31 Server Administration: Quota Properties

Table 18 Server Administration: Quota Properties

Field / Button	Description
User Name / Group Name	Type the name of the user or group to set a quota for.
Usage	Displays the percentage of the quota that has been used.
Hard Limit	Specify an amount in B, KB, MB, GB, or TB, or EB. This is the maximum amount of space the specified user or group can use.
Soft Limit	Specify an amount in B, KB, MB, GB, or TB, or EB. Once the user or group uses this amount of space, a warning is sent. (Typically this is 80% of the hard limit.)
Time Limit	Specify the amount of time it takes for the soft limit to turn into a hard limit.
Get Quota	Click to get quota settings for the specified user or group.
Set Quota	Click to set a quota for the specified user or group using the current settings.

Quality of Service Information

To view Quality of Service (QOS) information for a file system, expand a file system in the left pane, and then click **QOS**. Quality of Service information appears in the right pane, including stripe group, real-time limit, non-real-time reserve, committed RTIO, and the number of non-real-time clients.

Clients Information

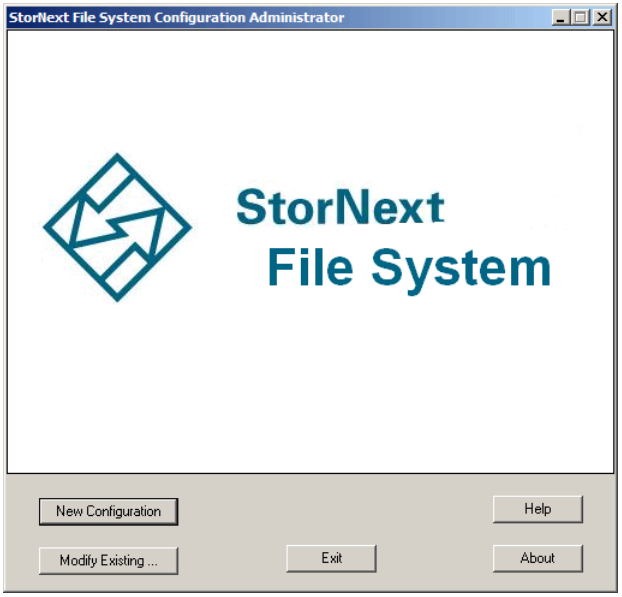
To view information about clients connected to the file system, expand a file system in the left pane, and then click **Clients**. Information about the currently connected clients appears in the right pane, including client ID, type, location, up time, and when the client license expires.

Server Configuration

The Server Configuration utility lets you view and modify properties for an MDC. Use this utility to create a new server configuration or modify an existing configuration.

To run the Server Configuration utility, on the Windows **Start** menu, click **All Programs > StorNext File System > Server Configuration**. The **Configuration Administrator** window appears (Figure 32).

Figure 32 Configuration Administrator



On the **Configuration Administrator** window, do one of the following:

- Click **New** to create a new configuration file.
- Click **Modify** to edit an existing configuration file. Browse for the configuration (***.cfg**) file and click **Open**.
- Click **Exit** to close the Configuration Administrator window.

When you create a new configuration file or open an existing configuration file for editing, the StorNext File System Configuration window appears. This window includes four tabs:

- Global Settings on page 82
- Disk Types on page 84
- Disk Definitions on page 86

- Stripe Groups on page 88

Click a tab to view or modify StorNext File System properties. After making changes on one or more tabs, click **OK** to save the changes. Specify a name and a location for the ***.cfg** file, and then click **Save**.

Note: Before configuring a file system, you should label disk devices. For more information, see Disk Device Labeler on page 67.

Global Settings

Use the **Global Settings** tab (Figure 33) to specify general file system properties. Table 19 describes the fields on the **Global Settings** tab.

Figure 33 Server Configuration: Global Settings

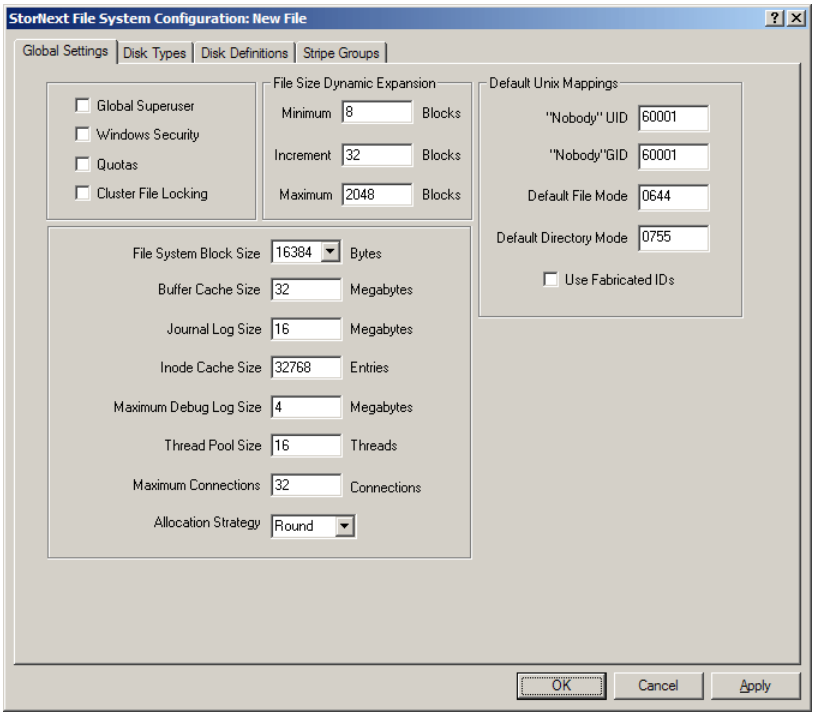


Table 19 Server Configuration: Global Settings

Field / Button	Description
Global Superuser	(Optional) Select to allow a user with superuser privileges to have global access rights on the file system.
Windows Security	(Optional) Select to use Windows Security Reference Monitor (ACLs) on Windows clients.
Quotas	(Optional) Select to enforce file system quotas and track storage usage for individual users and groups.
Cluster File Locking	(Optional) Select to send a warning when a user attempts to access a file that is locked or already in use.
File Size Dynamic Expansion - Minimum*	(Deprecated) Type the minimum number (in blocks) to allocate when a file is created.
File Size Dynamic Expansion - Increment*	(Deprecated) Type the initial increment (in blocks) by which a file is expanded.
File Size Dynamic Expansion - Maximum*	(Deprecated) Type the maximum increment (in blocks) by which a file is expanded.
File System Block Size	Select the file system block size (in bytes).
Buffer Cache Size	Type the amount of memory (in MB) used for general metadata information caching.
Journal Log Size	Type the maximum size (in MB) for the journal log.
Inode Cache Size	Type the number of entries in the inode cache.
Maximum Debug Log Size	Type the maximum size (in MB) for the debug log.

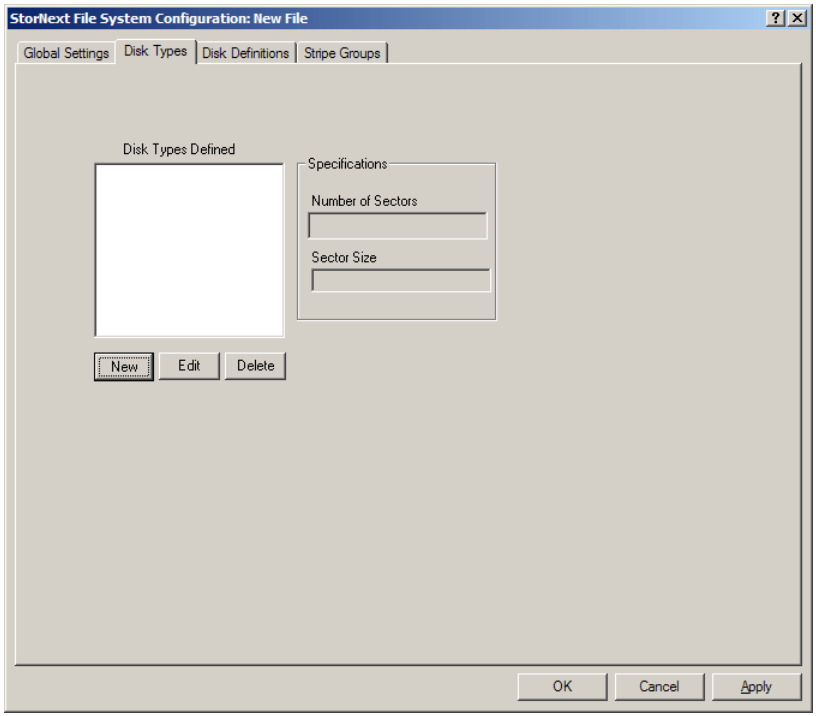
Field / Button	Description
Thread Pool Size	Type the number of threads the FSS uses when reading and storing files.
Maximum Connections	Type the maximum number of simultaneous connections (SNFS clients and Administrative Tap clients) allowed by the FSS.
Allocation Strategy	Select the method of allocation for new disk blocks: • **Round:** Alternate between similar stripe groups. (Default) • **Fill:** Use the stripe group with the least amount of free space (that can accommodate the file) before using the stripe group with the next smallest amount of free space. • **Balance:** Use the stripe group with the most free space.
"Nobody" UID	Type the Windows user ID to use when no user ID can be found using Active Directory.
"Nobody" GID	Type the Windows group ID to use when no group ID can be found using Active Directory.
Default File Mode	Type the permission mode to use when creating files.
Default Directory Mode	Type the permission mode to use when creating directories.
Use Fabricated IDs	Select if the metadata controller is running on Apple MacOS X.

* File size dynamic expansion is calculated on a file-by-file basis for better performance. If this is not working optimally, you can use these settings to override the file-by-file calculations.

Disk Types

Use the Disk Types tab (Figure 34) to define disk types used in the file system.

Figure 34 Server Configuration: Disk Types



On the **Disk Types** tab, do one of the following:

- To add a new disk type, click **New**.
- To modify a disk type, select the disk type in the list and click **Edit**.
- To delete a disk type, select it in the list and click **Delete**.

When you add or modify a disk type, the **Enter New Disk Type** dialog box appears (Figure 35). Table 20 describes the fields on the **Enter New Disk Type** dialog box.

Note: The **Sector** and **Sector Size** fields are populated with values from the Disk Device Labeler utility. For more information, see Disk Device Labeler on page 67.

Figure 35 Server Configuration: Enter New Disk Type

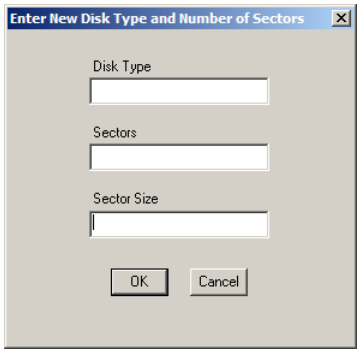


Table 20 Server Configuration: Enter New Disk Type

Field / Button	Description
Disk Type	Type a descriptive name for the disk.
Sectors	Type the number of physical sectors on the disk.
Sector Size	Type the size of sectors on the disk.
OK	Click to create the new disk type with the specified properties.
Cancel	Click to close the **Enter New Disk Type** dialog box without saving changes.

Disk Definitions

Use the **Disk Definitions** tab (Figure 36) to create disk definitions and modify disk specifications. Table 21 describes the fields on the **Disk Definitions** tab.

Figure 36 Server Configuration: Disk Definitions

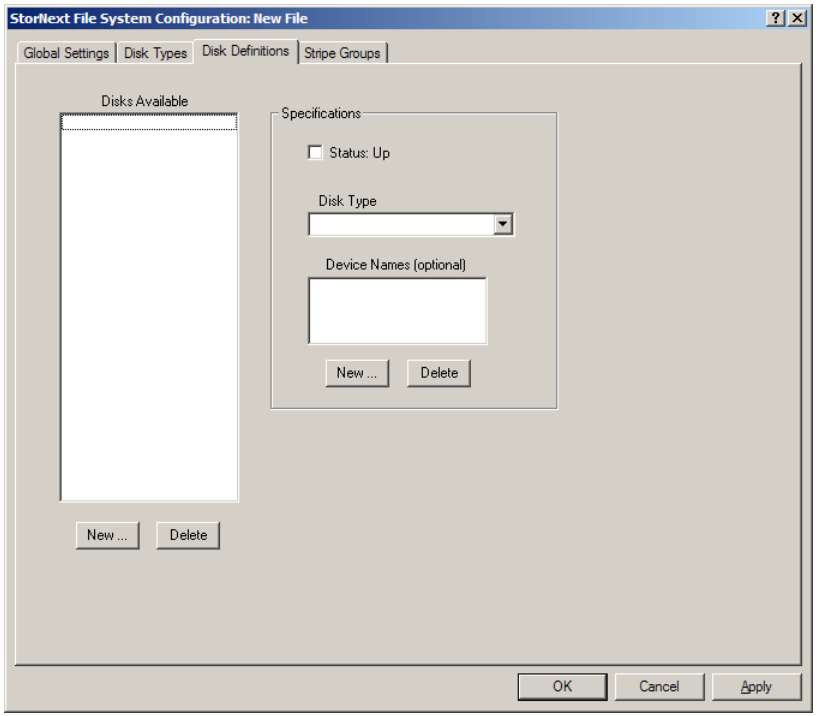


Table 21 Server Configuration: Disk Definitions

Field / Button	Description
Disks Available - New	Click to create a new disk definition. Type a name for the disk and click **OK**.
Disks Available - Delete	Click to delete the selected disk definition.
Status: Up	Select to enable the selected disk definition. When enabled, the defined disk is available to the configured file system for I/O.
Disk Type	Select a disk type for the selected disk definition. (Disk types are defined on the **Disk Types** tab.)
Device Names - New	Click to add a device name to the disk definition. Type a device name and click **OK**.

Field / Button	Description
Device Names - Delete	Click to delete the selected device name.

Stripe Groups

Use the Stripe Groups tab (Figure 37) to define stripe groups. (A stripe group is a logical storage unit consisting of one or more disk drives.) Table 22 describes the fields on the Stripe Groups tab.

Figure 37 Server Configuration: Stripe Groups

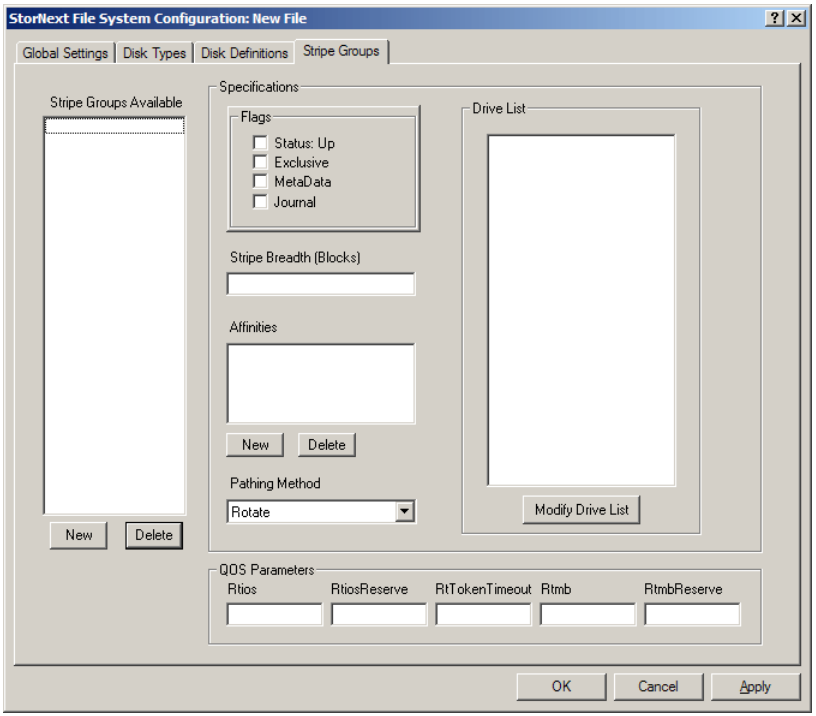


Table 22 Server Configuration: Stripe Groups

Field / Button	Description
Stripe Groups Available - New	Click to create a new stripe group. Type a name for the stripe group and click **OK**.

Field / Button	Description
Stripe Groups Available - Delete	Click to delete the selected stripe group.
Status: Up	Select to enable the selected stripe group. When enabled, the stripe group is available to the configured file system for I/O.
Exclusive	Select to allow only file types associated with the affinities specified under **Affinities** to be stored on the stripe group.
MetaData	Select to allow file system metadata to be stored on the stripe group.
Journal	Select to allow the file system logging journal to be stored on the stripe group.
Affinities - New	Click to add a new affinity to the stripe group. Type the name of the new affinity and click **OK**. (Affinity names cannot be longer than eight characters.)
Affinities - Delete	Click to the delete the selected affinity.
Pathing Method	Select a method for how the file system accesses the disk: • **Rotate:** Chooses an I/O path based on load. • **Static:** Always uses the same I/O path. • **Sticky:** Chooses an I/O path based on the identity of the target file.
Rtios	Type the maximum number of disk I/O operations per second available to real-time applications using the Quality of Service (QOS) API.
RtiosReserve	Type the minimum number of disk I/O operations per second reserved for non-real-time applications.
RtTokenTimeout	Type the number of seconds to wait for clients to respond to a token callback.

Field / Button	Description
Rtmb	Type the maximum number of MBs per second available to real-time applications using the Quality of Service (QOS) API.
RtmbReserve	Type the minimum number of MBs per second reserved for non-real-time applications.
Modify Drive List	Click to add or remove disks for the selected stripe group. Use the left and right arrows to add or remove disks in the **Primary Disks** list, and then click **OK**.

Caution: When adding a disk to an existing stripe group, make sure the disk is added to the bottom of the **Drive List**. If the disk is inserted in the middle of the list, the file system behaves as if files on the file system are corrupted.

Configuring a Distributed LAN Client Server

Traditional StorNext SAN clients access data storage (hard disk drives) in the SAN using a high throughput, low-latency fibre channel connection. In addition to this fibre channel connection, StorNext also provides a Distributed LAN client and server feature. StorNext clients can access data disks over a network instead of directly accessing disks.

A StorNext client running on Linux or Windows can be configured as a distributed LAN client or a distributed LAN server. (This is a licensed feature.)

- **Distributed LAN Server:** A distributed LAN server has direct access to data disks (using fibre channel) and makes the shared storage pool available to distributed LAN clients over a network.
- **Distributed LAN Client:** A distributed LAN client does not have direct access to data disks and instead accesses the shared storage

pool by connecting to a distributed LAN server over a network. (Distributed LAN clients have full StorNext client functionality.)

Note: A StorNext client configured as a distributed LAN server requires a dedicated system. Do not run other applications or services on a system configured as a distributed LAN server.

To configure a Linux client as a distributed LAN client, see Configuring a Distributed LAN Client on Linux on page 91. To configure a Linux client as a distributed LAN server, see Configuring a Distributed LAN Server on Linux on page 91.

To configure a Windows client as a distributed LAN server or a distributed LAN client, use the Windows Client Configuration utility. For more information, see Client Configuration on page 55 and Distributed LAN on page 57.

Configuring a Distributed LAN Client on Linux

To configure a StorNext client as a distributed LAN client, edit mount options in the **/etc/fstab** file.

1. Stop the StorNext client. At the command prompt, type:

   /etc/init.d/cvfs stop

2. Configure the client to mount a file system as a distributed LAN client. To do this, edit the **/etc/fstab** file so that it contains the following line:

   <file_system> <mount_point> cvfs rw,diskproxy=client 0 0

   where **<file_system>** is the name of the StorNext file system and **<mount_point>** is the directory path where the file system is mounted.

3. Start the StorNext client. At the command prompt, type:

   /etc/init.d/cvfs start

Configuring a Distributed LAN Server on Linux

To configure a StorNext client as a distributed LAN server, edit mount options in the **/etc/fstab** file, and then configure distributed LAN server options.

1. Stop the StorNext client. At the command prompt, type:

/etc/init.d/cvfs stop

2 Configure the client to mount a file system as a distributed LAN server. To do this, edit the **/etc/fstab** file so that it contains the following line:

<file_system> <mount_point> cvfs rw,diskproxy=server 0 0

where **<file_system>** is the name of the StorNext file system and **<mount point>** is the directory path where the file system is mounted.

3 Change to the **/usr/cvfs/bin/** directory.

4 Launch the StorNext distributed LAN server configuration utility. At the command prompt, type:

./sndpscfg -e

Note: To get help for the **sndpscfg** utility, at the command prompt, type: **./sndpscfig -h**

The distributed LAN server configuration utility appears (Figure 38).

Figure 38 sndpscfg Configuration Utility

```
# Disk Proxy Server Configuration
# -------------------------------

# ==================
# Tunable Parameters
# ==================

# Un-comment and modify any of the following lines to override the default
#  values for tunable parameters.
#tcp_window_size_kb 64
#transfer_buffer_size_kb 256
#transfer_buffer_count 16
#server_buffer_count 8
#daemon_threads 8

# ======================
# Interface Configuration
# ======================

# One or more interfaces must be configured for Disk Proxy Server I/O.

# Un-comment one (and only one) of the 2 lines below to allow Disk Proxy
#  Server I/O on interface 'eth0' and to select the address to listen for
#  connections on.
#interface eth0
interface eth0 address fec0::215:f2ff:fe68:afa2
```

5 Edit the default distributed LAN server configuration:

- (Required) Uncomment (delete the **#** next to) the line that corresponds to the network interface used by the distributed LAN server when connecting to distributed LAN clients. (For best performance, do not choose the network used for metadata.)
- (Optional) To change a tunable parameter from the default value, uncomment the parameter, and then specify a new value. Table 23 describes the tunable parameters.

Note: The distributed LAN server configuration utility uses the default text editor. If a default text editor has not been specified, the utility uses the **vi** editor.

Table 23 Distributed LAN Server Parameters

Field / Button	Description
tcp_window_size_kb	The TCP window size (in KB) used by the distributed LAN server. (Default: 64)
transfer_buffer_size_kb	The transfer buffer size (in KB) used by the distributed LAN server. A larger buffer may increase performance for larger files. (Default: 256)
transfer_buffer_count	The number of transfer buffers used by the distributed LAN server. This parameter is used only by Windows servers and clients. Linux servers pass the value of this parameter to Windows clients. (Default: 16)
server_buffer_count	The number of I/O buffers allocated for each network interface on the distributed LAN server. This parameter is used only by Linux servers. (Default: 8)
daemon_threads	The maximum number of daemon threads used by the distributed LAN server. (Default: 8)

6 When you are done making edits, type **:x** to save and close the configuration file.

The configuration is saved in the file **dpserver** in the **/user/cvfs/config/** directory.

Note: To edit this file at a later date, you can run the **sndpscfg** utility again or edit the **dpserver** file directly.

7 Start the StorNext client. At the command prompt, type:

/etc/init.d/cvfs start

Creating an Override Configuration

By default, the settings specified in the **dpserver** configuration file are used for all file systems mounted by the distributed LAN server. To use different settings for a file system, create an override configuration file for that file system. At the command prompt, type:

./sndpscfg -E <file_system>

where **<file_system>** is the name of the file system to which the override settings apply. Edit the configuration file, and then save and exit. A new configuration file with the name **dpserver.<file_system>** is created in the **/user/cvfs/config/** directory.

Using Distributed LAN Client and Server Commands

Use the **cvadmin** commands described in Table 24 to see information about distributed LAN servers and clients. (For more information about using **cvadmin** commands, see the *CLI Reference Guide*.)

Table 24 Distributed LAN Client and Server Commands

Command	Description
proxy	Lists all distributed LAN servers for the currently selected file system.
proxy long	Lists all distributed LAN servers for the currently selected file system, as well as tunable parameters and attached disks.
proxy who <distributed_LAN_client>	Lists distributed LAN servers connected to the specified distributed LAN client, as well as current I/O statistics.

Command	Description
proxy who <distributed_LAN_server>	Lists distributed LAN clients connected to the specified distributed LAN server, as well as current I/O statistics.

Distributed LAN Clients in HA Environments:

When a StorNext High Availability (HA) configuration is used and the StorNext Distributed LAN Client/Server (DLC) feature is configured on the HA server nodes, each node in the HA pair should be configured separately and independently. This allows for the best and highest availability of the DLC capability.

Ideally, each node in the HA pair should have the same number of network interfaces and be on the same networks. The dpserver configuration should be done before converting an MDC pair into an HA pair.

When deduplication/replication is enabled, Virtual IPs (VIPs) are configured and applied to the active blockpool server. In such a setting, the network interface(s) with VIP configured can potentially have more than one IP addresses (physical and virtual IPs). If such an interface is also used for Distributed LAN Clint, use the interface with the physical IP address to specify the network interface.

Installing Promise RAID Controller Software

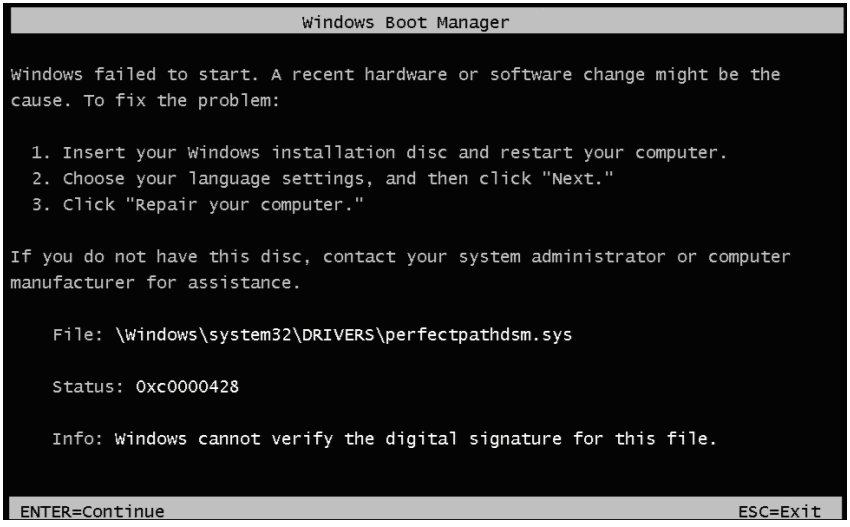
If you are using Promise RAID controllers on a Windows Server 2008 64-bit system, you must install Promise's PerfectPath software. If you do not install this software you can still access (read from and write to) your RAIDs, but you will be unable to failover your controller.

Promise is working on a solution to this problem, but in the meantime they have provided the following workaround.

1. Install the PerfectPath software on your Windows Server 2008 64-bit system.
2. Restart your system.

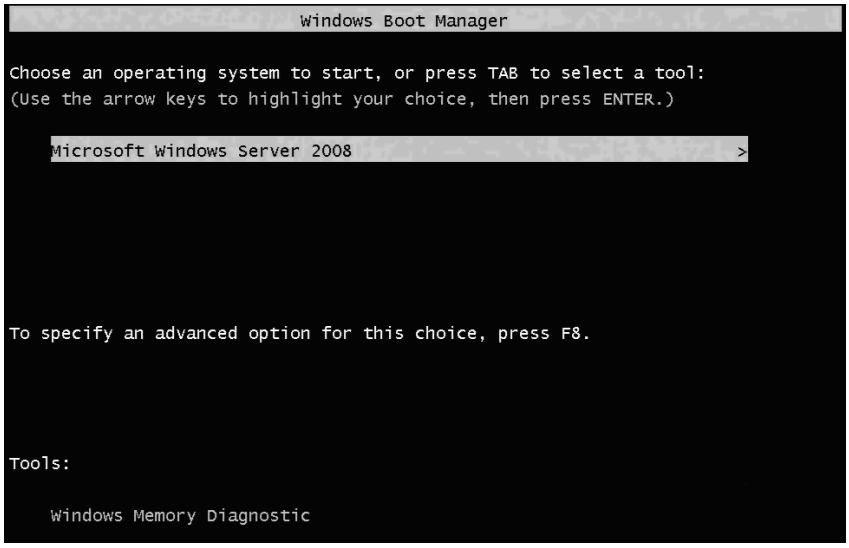
The login prompt will *not* appear after you restart. Instead, the **Windows Boot Manager** screen appears showing an error message: “Windows cannot verify the digital signature for this file” (**\Windows\system32\DRIVERS\ perfectpathdsm.sys)**

Figure 39 Windows Boot Manager Screen 1



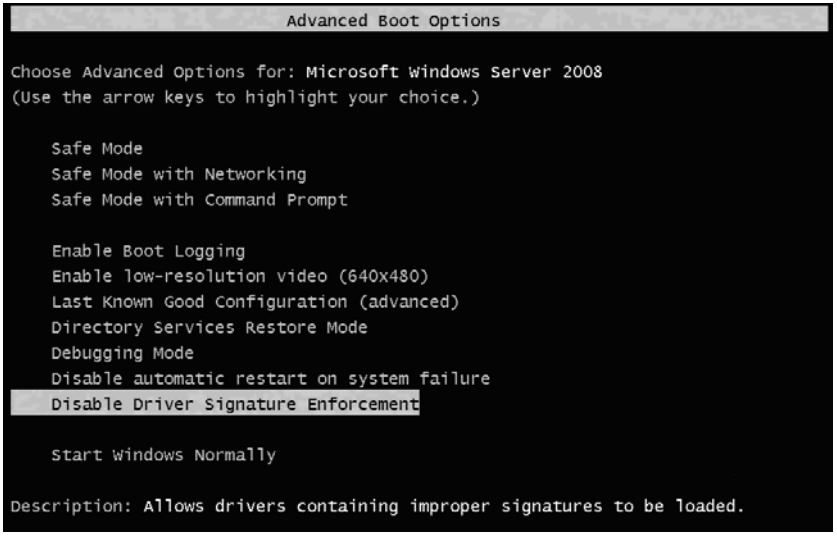
3 From the **Windows Boot Manager** screen, press **Enter** to continue. A second **Windows Boot Manager** screen appears, asking you to choose an operating system or specify an advanced option.

Figure 40 Windows Boot Manager Screen 2



4 On the second **Windows Boot Manager** screen, press **F8** to specify advanced options. The **Advanced Boot Options** screen appears.

Figure 41 Advanced Boot Options Screen



5 On the **Advanced Boot Options** screen, use the arrow keys to choose the option **Disable Driver Signature Enforcement**.

Choosing this option will cause the system to display the login prompt normally after you reboot.

6 Restart your system.

Appendix A
Uninstalling StorNext

If you no longer need to use StorNext, uninstall the software to remove it from your system. To uninstall StorNext, use the appropriate method for your operating system:

- Uninstalling StorNext on Linux on page 99
- Uninstalling StorNext on Windows on page 102

Uninstalling StorNext on Linux

To uninstall StorNext on a metadata controller (MDC), use the installation script (see Uninstalling SNSM and SNFS on page 99).

To uninstall the StorNext client software, use the appropriate method for your operating system (see Uninstalling the StorNext Client Software on page 101).

Uninstalling SNSM and SNFS

To uninstall StorNext File System and Storage Manager on an MDC running Unix or Linux, run the installation script with the **-remove** option. To launch the script, access the folder for your operating system on the StorNext installation DVD.

1 Log on to the MDC as **root**.

2 Mount the StorNext installation DVD and change to the DVD root directory.

Note: When you mount a DVD in a Red Hat 5 system, DVDs are mounted by default with a noexec (non-executable) option which prevents you from proceeding with the installation.

For Red Hat users only, before proceeding you must remount the DVD by typing **mount -o remount, exec ...**

3 List the installation directories on the DVD. At the command prompt, type:

ls -l

4 Identify the correct installation directory for your operating system and hardware platform, and then change to that directory.

For example, for Red Hat Linux 5 running on an x86 64-bit platform, change to the **RedHat50AS_26x86_64** directory.

5 Do one of the following steps:

- (Recommended) Completely uninstall StorNext and return the system to a pristine state. At the command prompt, type:

  ./install.stornext -removeall

- Uninstall StorNext binaries but leave files that are useful if you are going to perform an HA upgrade (such as license files, some configuration files, and log files). At the command prompt, type:

  ./install.stornext -remove

The installation script runs (Figure 42).

Figure 42 StorNext Remove Main Menu

```
Stornext Remove 4.0.0(13529) MAIN MENU

 1)  Installation Configuration
 2)  Remove stornext
 3)  Show Status
 4)  Quit

Enter your choice <4>:
```

6 Type **2** and press **<Enter>**.

Uninstalling the StorNext Client Software

To uninstall the StorNext client software, unmount all file systems and stop StorNext software. Then remove the client software package using the appropriate command for your operating system.

Note: This procedure removes the client software only. To uninstall StorNext File System and Storage Manager, see Uninstalling SNSM and SNFS on page 99.

1 Unmount all file systems. To unmount a file system, at the command prompt, type:

umount <mount -point>

where **<mount_point>** is the location where the file system is mounted.

2 Stop the StorNext client software. At the command prompt, type:

/etc/init.d/cvfs fullstop

3 Remove the StorNext software package using the correct commands for your operating system.

- Linux:

  rpm -e <package_name>

  where **<package_name>** is the name of the client software package that was installed.

- Solaris:

  pkgrm ADICsnfs

- HP-UX:

  swremove –x mount_all_filesystems=false snfsclient

- IBM AIX:

  installp –u <package_name>

  where **<package_name>** is the name of the client software package that was installed.

- SGI IRIX:

  Inst

remove cv_base

remove cv_clnt

go

Note: If you do not know the package name, you can download the client software for your operating system and check the package name (see Installing the StorNext Client on page 25).

Uninstalling StorNext on Windows

To uninstall StorNext (server or client) on a Windows system, use the Add or Remove Programs control panel.

1. Open the Windows Control Panel, and then double-click **Add or Remove Programs**.

   The **Add or Remove Programs** window appears.

2. In the list of currently installed programs, click **StorNext File System**, and then click **Remove**.
3. Click **Yes** to confirm the removal.
4. After the uninstall process is complete, reboot the system.

Appendix B

StorNext Administration Tasks

Use the following information to perform specific file system configuration tasks that do not apply to a typical StorNext set up.

Sharing a Library Between Applications

If two or more software applications (on one or multiple machines), are configured to use the same library, then you must use the following procedure to enable the library to be shared and to prevent media being overwritten when these applications are used. This task must be performed when the Configuration Wizard (CW) is run during initial StorNext setup and configuration. For more information on using the CW, see The StorNext Configuration Wizard on page 52.

Note: For library sharing to be effective, the steps in the following procedure must be performed in the order in which they are presented.

You must have root privileges to perform this task.

1 When you match hardware devices to their correct slots on the **Matched Devices With Slots** screen (you will see this screen during the execution of Step 4 of the Configuration Wizard, **Add Tape Drive**), only configure the tape drives that your software application

will use. You do not need to configure all tape drives listed as part of the library.

2 Once the tape drives are added and configured in Step 4 of the CW (**Add Tape Drive**), but before continuing to Step 5 of the CW (**Add Media**), go to a metadata controller and perform these steps:

   a Change directory to the **config** directory. Type:

   cd /usr/adic/MSM/internal/config

   b List all files in the config directory. Type:

   ls -l

   c Using a text editor, edit the **media_file_<library_name>** file. For example:

   vi media_file_<library_name>

   The **media_file** contains a filter for barcode labels that are read by StorNext. This action is performed by using regular expressions. (A regular expression provides a mechanism to select strings from a set of character strings. In the above example, the character strings refer to the barcode labels.) In the example above, **<library_name>** is the library that was configured in Step 3, **Add Library**, of the CW.

   d Set up the pattern matching of the media file.

   Multiple entries may exist in the media file. Read the examples in the file to learn how to set up pattern matching.

   These are examples of pattern matching:

   scsi_archive 1111.. (brings in all media with the 1111 prefix)

   scsi_archive AB9... (brings in all six-character media labels with the AB9 prefix)

   scsi_archive CLN1.. (brings in all media that have the CLN1 prefix)

3 Continue to the **Add Media section of the Configuration Wizard**.

Manually Modifying Storage Manager Configuration

StorNext employs global parameters that define the Storage Manager configuration. Although the default values in these files are sufficient for most installations, you can edit the files to customize the StorNext software for your specific installation. This section provides a procedure that can be used to modify the following StorNext configuration files:

- **/usr/adic/TSM/config/fs_sysparm_override**
- **/usr/adic/TSM/config/filesize.config**
- **/usr/adic/TSM/logs/log_params**

fs_sysparm_override Configuration File

The **fs_sysparm_override** configuration file is used to tune items related to system administration, drives and media, media manager interface, and other miscellaneous items. For more information about the parameters that can be placed into the **fs_sysparm_override** configuration file, refer to the **fs_sysparm.README** file.

Caution: Incorrect modification of the **fs_sysparm_override** configuration file can cause a serious, adverse effect on StorNext functionality. Before modifying this file, Quantum recommends that you contact Quantum Global Services.

Caution: If you have files larger than 100GB and are using LTO2 media, the MED_SEG_OVER_LTO parameter may be modified to a value less than or equal to 190G to reduce file fragmentation. This has the adverse effect of reducing the potential for parallel I/O for multiple file segments. Setting the MED_SEG_OVER_LTO parameter to a value larger than 190GB may result in allocation failures that prevent file movement to tape.

filesize.config Configuration File

The **filesize.config** configuration file is used to control the file steering feature and has these characteristics:

- Allows the placement of files on different media types, based on the size of the files
- Specifies which drive pool a policy class should use when storing data
- Contains specific syntax and usage information
- Enables the system administrator to make changes without recycling the Tertiary Manager software

log_params Configuration File

The **log_params** configuration file is used to control various levels of trace logging. The file specifies each log level and how to enable and/or disable it.

Use the following procedure to modify the **fs_sysparm_override**, **filesize.config**, or **log_params** configuration files.

1. Use a text editor to open one of the following configuration files:
   - **/usr/adic/TSM/config/fs_sysparm_override**
   - **/usr/adic/TSM/config/filesize.config**
   - **/usr/adic/TSM/logs/log_params**
2. Locate the parameter you want to modify and replace the setting with a new, valid value.

   When editing a file, be sure to follow the format used by entries in the file. For example, in the **fs_sysparm_override** and **filesize.config** files, all entries must be in the format: **name=value;**
3. Recycle the Tertiary Manager software.
   a. Stop the software by typing **TSM_control stop**
   b. Restart the software by typing **TSM_control start**

Connecting to a StorNext Network Using Xsan 2

If you are using Apple Xsan 2.1.1 or later, follow the procedure below to connect to a StorNext network. The procedure consists of mounting the StorNext file system onto the Mac OS X computer, and then creating an

automount.plist to enable mounting the StorNext file system whenever the Macintosh is rebooted.

Mounting SNFS on the Mac OS X Computer

Follow this procedure to mount the StorNext file system.

1. Connect the Apple computer to the SAN's metadata Ethernet and Fibre Channel networks.
2. Install Xsan 2 software on the Apple computer. (Xsan 2 is supported only by the Leopard operating system.)
3. Create the file **/etc/systemserialnumbers/xsan** using that Macintosh's Xsan serial number.

   You must create the directory **/etc/systemserialnumbers** if it doesn't already exist. The format of the serial number file is a two-line file:

   ```
   XSAN-020-XXX-XXX-X-XXX-XXX-XXX-XXX-XXX-XXX-X
   registered to|organization
   ```

   Note: The file does not have a trailing return on the last line. Use the following:
   `cat > /etc/systemserialnumbers/xsan and end with ^D^D` (where **^D^D** is control-D control-D)

   If you are unable to browse to the systemserialnumbers directory, use the cat or echo commands to insert information into the Xsan file under that directory.
4. Copy from the following template and paste into the file to create the file **config.plist** in **/Library/Filesystems/Xsan/config/**.

   Note: Note: Items in red indicate data you should enter which applies to your configuration. Do not enter the literal red text shown in the template.
   For example, for <string>My IP Address</string> you would enter your system metadata IP address or the regular system IP address, not the words "My IP Address."

   ```
   <?xml version="1.0" encoding="UTF-8"?>
   <!DOCTYPE plist PUBLIC "-//Apple//DTD PLIST 1.0//EN"
   "http://www.apple.com/DTDs/PropertyList-1.0.dtd">
   <plist version="1.0">
   ```

```
<dict>
    <key>computers</key>
    <array/>
    <key>metadataNetwork</key>
    <string>My IP Address</string>
    <key>ownerEmail</key>
    <string>me@mycompany.com</string>
    <key>ownerName</key>
    <string>My Name</string>
    <key>role</key>
    <string>CLIENT</string>
    <key>sanName</key>
    <string>My SAN name</string>
    <key>serialNumbers</key>
    <array>
        <dict>
            <key>license</key>
            <string>xsan client license number</string>
            <key>organization</key>
            <string>organization name</string>
            <key>registeredTo</key>
            <string>registered to name</string>
        </dict>
    </array>
</dict>
</plist>
```

5 Navigate to an existing StorNext server on the SAN and copy the **fsnameservers** file to your Macintosh client in **/Library/Filesystems/Xsan/config/**. This file should contain the IP addresses of the name server machines.

On Unix: the **fsnamerservers** file is located at **/usr/cvfs/config/**.

On Windows: the **fsnamerservers** file is located at **\%cvfsroot%\config** (where **%cvfsroot%** is the directory where the StorNext software is installed). This directory is usually located in **C:\Program Files\Stornext\config**.

Note: If the content of the **fsnameservers** file specifies a fully qualified domain name (FQDN), you must convert this to the metadata controller's IP address by using the command **host <FQDN>**

---

6 Run **RegisterSeRV** in Terminal to ensure that the serial number is recognized.

7 Run the following in Terminal to update Xsan with the new file system information:

launchctl unload /System/Library/LaunchDaemons/com.apple.xsan.plist

launchctl load /System/Library/LaunchDaemons/com.apple.xsan.plist

8 Run **cvadmin** to verify the file systems are present.

9 Run **cvlabel -l** to verify the all the luns used for file systems are displayed.

10 In the Terminal type **Xsanctl mount <fsname>**.

Where **<fsname>** is the file system name displayed in cvadmin.

11 Verify the automount.plist file is created in the **/Library/Filesystems/Xsan/config/** directory.

12 Verify the StorNext file system is now mounted on the Mac OS X computer.

XSan Upgrade Considerations

StorNext users migrating their metadata controllers from Apple Xsan to Linux should be aware of the following upgrade considerations:

- If the file system is running Xsan 2.1.1 or earlier, it should be a simple upgrade: just replace the MDC.
- If the file system is running Xsan 2.2 or later with “NamedStreams No” (which is NOT the default for Xsan 2.2), it should also be a simple upgrade: just replace the MDC.

If the file system is running Xsan 2.2 or later with “NamedStreams Yes,” you must completely remake (reformat) the file system. Before reformatting it is highly recommended that a complete backup be performed before migrating.

Xsan 2.2 defaults to making new file systems with “NamedStreams Yes” (actually called Extended Attributes in Xsan Admin's create FS dialog), which means that these file systems can't be migrated to StorNext File System without using the `cvmkfs` command.

Both the fsm and cvfsck will fail to start, panicking with a message about the inode version 0x205 being unknown and unsupported (see bug 29808 for the complete text from cvfsck).

If you have “NamedStreams No” you can migrate by commenting out the “NamedStreams No” line.

Caution: It is possible to take an Xsan 2.1 file system and update it to NamedStreams Yes with Xsan 2.2. Once that happens, migration is no longer possible.

Viewing a Fibre Channel Device

Use this procedure to add a Fibre Channel (FC) device. Before adding a FC device, first configure the Host Bus Adapter (HBA) card so you can view the device. Use the **fs_scsi -p** command to make sure you can view

your devices over FC. FC devices include tape libraries, individual drives, or RAID disk.

1. Type the **/usr/cvfs/bin/fs_scsi -p** command to make sure that the necessary devices can be viewed over FC.

- To view tape libraries, access the StorNext GUI to view your tape libraries.
- To view individual drives or RAID LUNs, access either the StorNext or SNFS GUI.

To add a disk, see Chapter 6, "Managing the File System," in the *StorNext User's Guide*. To add a tape drive to a configured library, see Chapter 8, "Managing Drives and Disks," in the *StorNext User's Guide*.

StorNext Log Rolling Administration

Use the following procedures to configure StorNext log rolling options.

Changing Log Rolling Times

Use this procedure to change the frequency of rolling the StorNext logs. This process requires that you edit the **tldm crontab** to set log times. Use this procedure to edit **tdlm crontab**.

1. Log on as **root**.
2. Edit the **tdlm crontab** and update the **sn_log_update script**.

   Below is an example crontab:

   0 1,7,13,19 * * * /usr/adic/guib/bin/cmdwrap /usr/adic/util/sn_log_update /usr/adic

   where **0 1,7,13,19 * * *** designates the times when logs run.

Log Rolling Options

You can change these options to optimize log rolling.

- **-s**: This option sets the directory where logs are saved (copied) to as they are rolled. This directory is typically a managed directory. For example:

  sn_log_update [-s <dir>]

  where **<dir>** is the directory where you want the logs to reside.
- **$DEF_SIZE = 2000000**: This is the default size at which logs are rolled. Edit this entry in the **/usr/adic/util/sn_log_update.cfg** file if you want the log sizes to be larger or smaller.
- **$DEF_LOGS = 28**: This is the default number of logs that are saved before they are deleted. Edit this entry in the **/usr/adic/util/sn_log_update.cfg** file if you want to save less than 28 logs or are saving the logs to a managed directory.

Avoiding Poor System Performance

Most systems run administrator scripts to rebuild namespace databases and/or to remove certain files. Unfortunately, running these scripts can degrade StorNext performance. It is recommended that you review administrator scripts to determine if they may adversely affect your StorNext installation and its performance.

For example, many UNIX systems have a find command in the **root crontab** that traverses the entire directory hierarchy while searching for old temporary and core files. Typically, the **find** commands are configured only to run on the local file system. However by default, these commands do not skip StorNext file systems because they are considered to be local, even though the file system manager may be running remotely. StorNext performance can be dramatically degraded when these commands visit a StorNext file system if the system contains a large number of files.

If StorNext performance is degraded to an unacceptable level due to the operation of administrator scripts, you can modify them to skip StorNext file systems. A sample command (original and modified is shown below:

On IRIX, by default, the **root crontab** contains the following find command:

find / -local -type f '(' -name core -o -name dead.letter ')' -atime +7 -mtime + 7 -exec rm -f '{}' ';'

To prevent the find command from searching in StorNext file systems, modify it as shown:

'(' -type d -fstype cvfs -prune ')' -o

The modified find command looks like this:

find / -local '(' -type d -fstype cvfs ')' -o -type f '(' -name core -o -name dead.letter ')' -atime +7 –mtime +7 -exec rm -f '{}' ';'

Appendix C

StorNext Files

StorNext creates a series of files that are used and modified to configure a file system. This section includes an expanded configuration file, a cluster-wide central control file, and a listing of the most commonly used StorNext files with descriptions and locations for each.

Example FSM Configuration File

On Linux, the StorNext configuration file uses an XML format (.cfgx). On Windows, the configuration file uses a text format (.cfg). However, the values contained in both files are similar.

You can locate an example StorNext configuration file in the following directory:

- Linux - **/usr/cvfs/examples/example.cfgx**
- Windows - **C:\Program Files\Stornext\config\example.cfg**

  If you installed StorNext in a location other than the default installation directory, the example configuration file is located in **C:\<install_directory>\config\example.cfg**

Linux Example Configuration File

Below are the contents of the StorNext example configuration file for Linux (**example.cfgx**):

```
<?xml version="1.0" encoding="UTF-8"?>
<configDoc xmlns="http://www.quantum.com/snfs" version="1.0">
   <config configVersion="0" name="example" fsBlockSize="16384"
   journalSize="16777216">
      <globals>
         <abmFreeLimit>false</abmFreeLimit>
         <allocationStrategy>round</allocationStrategy>
         <haFsType>HaUnmonitored</haFsType>
         <bufferCacheSize>33554432</bufferCacheSize>
         <cvRootDir>/</cvRootDir>
         <storageManager>false</storageManager>
         <dataMigrationThreadPoolSize>128</dataMigrationThreadPoolSize>
         <debug>00000000</debug>
         <dirWarp>true</dirWarp>
         <extentCountThreshold>49152</extentCountThreshold>
         <enableSpotlight>false</enableSpotlight>
         <enforceAcls>false</enforceAcls>
         <fileLocks>false</fileLocks>
         <fileLockResyncTimeOut>20</fileLockResyncTimeOut>
         <forcePerfectFit>false</forcePerfectFit>
         <fsCapacityThreshold>0</fsCapacityThreshold>
         <globalSuperUser>true</globalSuperUser>
         <inodeCacheSize>32768</inodeCacheSize>
         <inodeExpandMin>0</inodeExpandMin>
         <inodeExpandInc>0</inodeExpandInc>
         <inodeExpandMax>0</inodeExpandMax>
         <inodeDeleteMax>0</inodeDeleteMax>
         <inodeStripeWidth>0</inodeStripeWidth>
         <maxConnections>32</maxConnections>
         <maxLogs>4</maxLogs>
         <remoteNotification>false</remoteNotification>
         <reservedSpace>true</reservedSpace>
         <fsmRealTime>false</fsmRealTime>
         <fsmMemLocked>false</fsmMemLocked>
         <opHangLimitSecs>180</opHangLimitSecs>
         <perfectFitSize>131072</perfectFitSize>
         <quotas>false</quotas>
```

```
    <restoreJournal>false</restoreJournal>
    <restoreJournalDir/>
    <restoreJournalMaxHours>0</restoreJournalMaxHours>
    <restoreJournalMaxMb>0</restoreJournalMaxMb>
    <stripeAlignSize>0</stripeAlignSize>
    <trimOnClose>0</trimOnClose>
    <threadPoolSize>32</threadPoolSize>
    <unixDirectoryCreationModeOnWindows>644</
    unixDirectoryCreationModeOnWindows>
    <unixIdFabricationOnWindows>false</unixIdFabricationOnWindows>
    <unixFileCreationModeOnWindows>755</unixFileCreationModeOnWindows>
    <unixNobodyUidOnWindows>60001</unixNobodyUidOnWindows>
    <unixNobodyGidOnWindows>60001</unixNobodyGidOnWindows>
    <windowsSecurity>true</windowsSecurity>
    <eventFiles>true</eventFiles>
    <eventFileDir/>
    <allocSessionReservation>false</allocSessionReservation>
</globals>
<diskTypes>
    <diskType typeName="MetaDrive" sectors="99999999" sectorSize="512"/>
    <diskType typeName="JournalDrive" sectors="99999999"
    sectorSize="512"/>
    <diskType typeName="VideoDrive" sectors="99999999" sectorSize="512"/>
    <diskType typeName="AudioDrive" sectors="99999999" sectorSize="512"/>
    <diskType typeName="DataDrive" sectors="99999999" sectorSize="512"/>
</diskTypes>
<stripeGroups>
    <stripeGroup index="0" name="MetaFiles" status="up"
    stripeBreadth="262144" read="true" write="true" metadata="true"
    journal="false" userdata="false" realTimeIOs="200"
    realTimeIOsReserve="1" realTimeMB="200" realTimeMBReserve="1"
    realTimeTokenTimeout="0" multipathMethod="rotate">
        <disk index="0" diskLabel="CvfsDisk0" diskType="MetaDrive"/>
    </stripeGroup>
    <stripeGroup index="1" name="JournFiles" status="up"
    stripeBreadth="262144" read="true" write="true" metadata="false"
    journal="true" userdata="false" realTimeIOs="0"
    realTimeIOsReserve="0" realTimeMB="0" realTimeMBReserve="0"
    realTimeTokenTimeout="0" multipathMethod="rotate">
        <disk index="0" diskLabel="CvfsDisk1" diskType="JournalDrive"/>
    </stripeGroup>
```

```
    <stripeGroup index="2" name="VideoFiles" status="up"
    stripeBreadth="4194304" read="true" write="true" metadata="false"
    journal="false" userdata="true" realTimeIOs="0"
    realTimeIOsReserve="0" realTimeMB="0" realTimeMBReserve="0"
    realTimeTokenTimeout="0" multipathMethod="rotate">
       <affinities exclusive="true">
          <affinity>Video</affinity>
       </affinities>
       <disk index="0" diskLabel="CvfsDisk2" diskType="VideoDrive"/>
       <disk index="1" diskLabel="CvfsDisk3" diskType="VideoDrive"/>
       <disk index="2" diskLabel="CvfsDisk4" diskType="VideoDrive"/>
       <disk index="3" diskLabel="CvfsDisk5" diskType="VideoDrive"/>
       <disk index="4" diskLabel="CvfsDisk6" diskType="VideoDrive"/>
       <disk index="5" diskLabel="CvfsDisk7" diskType="VideoDrive"/>
       <disk index="6" diskLabel="CvfsDisk8" diskType="VideoDrive"/>
       <disk index="7" diskLabel="CvfsDisk9" diskType="VideoDrive"/>
    </stripeGroup>
    <stripeGroup index="3" name="AudioFiles" status="up"
    stripeBreadth="1048576" read="true" write="true" metadata="false"
    journal="false" userdata="true" realTimeIOs="0"
    realTimeIOsReserve="0" realTimeMB="0" realTimeMBReserve="0"
    realTimeTokenTimeout="0" multipathMethod="rotate">
       <affinities exclusive="true">
          <affinity>Audio</affinity>
       </affinities>
       <disk index="0" diskLabel="CvfsDisk10" diskType="AudioDrive"/>
       <disk index="1" diskLabel="CvfsDisk11" diskType="AudioDrive"/>
       <disk index="2" diskLabel="CvfsDisk12" diskType="AudioDrive"/>
       <disk index="3" diskLabel="CvfsDisk13" diskType="AudioDrive"/>
    </stripeGroup>
    <stripeGroup index="4" name="RegularFiles" status="up"
    stripeBreadth="262144" read="true" write="true" metadata="false"
    journal="false" userdata="true" realTimeIOs="0"
    realTimeIOsReserve="0" realTimeMB="0" realTimeMBReserve="0"
    realTimeTokenTimeout="0" multipathMethod="rotate">
       <disk index="0" diskLabel="CvfsDisk14" diskType="DataDrive"/>
       <disk index="1" diskLabel="CvfsDisk15" diskType="DataDrive"/>
       <disk index="2" diskLabel="CvfsDisk16" diskType="DataDrive"/>
       <disk index="3" diskLabel="CvfsDisk17" diskType="DataDrive"/>
    </stripeGroup>
</stripeGroups>
```

```
   </config>
</configDoc>
```

Windows Example Configuration File

Below are the contents of the StorNext example configuration file for Windows (**example.cfg**):

```
# Globals

ABMFreeLimit no
AllocationStrategy Round
HaFsType HaUnmonitored
FileLocks No
BrlResyncTimeout 20
BufferCacheSize 32M
CvRootDir /
DataMigration No
DataMigrationThreadPoolSize 128
Debug 0x0
DirWarp Yes
ExtentCountThreshold 48K
EnableSpotlight No
ForcePerfectFit No
FsBlockSize 16K
GlobalSuperUser Yes
InodeCacheSize 32K
InodeExpandMin 0
InodeExpandInc 0
InodeExpandMax 0
InodeDeleteMax 0
InodeStripeWidth 0
JournalSize 16M
MaxConnections 32
MaxLogs 4
PerfectFitSize 128K
RemoteNotification No
ReservedSpace Yes
FSMRealtime No
FSMMemlock No
OpHangLimitSecs 180
```

```
Quotas No
RestoreJournal No
RestoreJournalMaxHours 0
RestoreJournalMaxMB 0
StripeAlignSize -1
TrimOnClose 0
ThreadPoolSize 32
UnixDirectoryCreationModeOnWindows 0755
UnixIdFabricationOnWindows No
UnixFileCreationModeOnWindows 0644
UnixNobodyUidOnWindows 60001
UnixNobodyGidOnWindows 60001
WindowsSecurity Yes
EventFiles Yes
AllocSessionReservation No
# Disk Types

[DiskType MetaDrive]
Sectors 99999999
SectorSize 512
[DiskType JournalDrive]
Sectors 99999999
SectorSize 512
[DiskType VideoDrive]
Sectors 99999999
SectorSize 512
[DiskType AudioDrive]
Sectors 99999999
SectorSize 512
[DiskType DataDrive]
Sectors 99999999
SectorSize 512
# Disks

[Disk CvfsDisk0]
Type MetaDrive
Status UP
[Disk CvfsDisk1]
Type JournalDrive
Status UP
```

```
[Disk CvfsDisk2]
Type VideoDrive
Status UP
[Disk CvfsDisk3]
Type VideoDrive
Status UP
[Disk CvfsDisk4]
Type VideoDrive
Status UP
[Disk CvfsDisk5]
Type VideoDrive
Status UP
[Disk CvfsDisk6]
Type VideoDrive
Status UP
[Disk CvfsDisk7]
Type VideoDrive
Status UP
[Disk CvfsDisk8]
Type VideoDrive
Status UP
[Disk CvfsDisk9]
Type VideoDrive
Status UP
[Disk CvfsDisk10]
Type AudioDrive
Status UP
[Disk CvfsDisk11]
Type AudioDrive
Status UP
[Disk CvfsDisk12]
Type AudioDrive
Status UP
[Disk CvfsDisk13]
Type AudioDrive
Status UP
[Disk CvfsDisk14]
Type DataDrive
Status UP
[Disk CvfsDisk15]
```

```
Type DataDrive
Status UP
[Disk CvfsDisk16]
Type DataDrive
Status UP
[Disk CvfsDisk17]
Type DataDrive
Status UP
# Stripe Groups

[StripeGroup MetaFiles]
Status Up
StripeBreadth 256K
Metadata Yes
Journal No
Exclusive Yes
Read Enabled
Write Enabled
Rtmb 0
Rtios 0
RtmbReserve 0
RtiosReserve 0
RtTokenTimeout 0
MultiPathMethod Rotate
Node CvfsDisk0 0

[StripeGroup JournFiles]
Status Up
StripeBreadth 256K
Metadata No
Journal Yes
Exclusive Yes
Read Enabled
Write Enabled
Rtmb 0
Rtios 0
RtmbReserve 0
RtiosReserve 0
RtTokenTimeout 0
MultiPathMethod Rotate
```

```
Node CvfsDisk1 0

[StripeGroup VideoFiles]
Status Up
StripeBreadth 4M
Metadata No
Journal No
Exclusive No
Read Enabled
Write Enabled
Rtmb 0
Rtios 0
RtmbReserve 0
RtiosReserve 0
RtTokenTimeout 0
MultiPathMethod Rotate
Node CvfsDisk2 0
Node CvfsDisk3 1
Node CvfsDisk4 2
Node CvfsDisk5 3
Node CvfsDisk6 4
Node CvfsDisk7 5
Node CvfsDisk8 6
Node CvfsDisk9 7
Affinity Video

[StripeGroup AudioFiles]
Status Up
StripeBreadth 1M
Metadata No
Journal No
Exclusive No
Read Enabled
Write Enabled
Rtmb 0
Rtios 0
RtmbReserve 0
RtiosReserve 0
RtTokenTimeout 0
MultiPathMethod Rotate
```

```
Node CvfsDisk10 0
Node CvfsDisk11 1
Node CvfsDisk12 2
Node CvfsDisk13 3
Affinity Audio

[StripeGroup RegularFiles]
Status Up
StripeBreadth 256K
Metadata No
Journal No
Exclusive No
Read Enabled
Write Enabled
Rtmb 0
Rtios 0
RtmbReserve 0
RtiosReserve 0
RtTokenTimeout 0
MultiPathMethod Rotate
Node CvfsDisk14 0
Node CvfsDisk15 1
Node CvfsDisk16 2
Node CvfsDisk17 3
```

Cluster-Wide Central Control

The purpose of this feature is to provide cluster-wide central control. (Currently the central control file is supported on the Linux platform only.) A central control file called **nss_cctl.xml** provides a way to restrict the behavior of SNFS cluster nodes (fsm, file system client, cvadmin client) from a central place: an NSS server.

This feature currently supports the following controls that allow you to specify:

1. Whether a client is allowed to mount as a proxy client.
2. Whether a client is allowed to mount as read/write or read-only.

3 Whether a user (especially a local administrator on Windows clients,) is allowed to take ownership of a file or directory on a StorNext file system.

4 Whether **cvadmin** running on a certain client is allowed to have super admin privilege to run destructive commands such as starting/stopping the file system, refreshing disks, changing quota settings, and so on.

5 Whether **cvadmin** running on certain client is allowed to connect to other fsms via the **-H** option.

6 Whether binary executable files on the StorNext file system are allowed to be executed.

7 Whether the **setuid** bit of a file is allowed to take effect.

The control file is in xml format and has a hierarchical structure. The top level element, **snfsControl**, contains control elements with the **securityControl** label for certain file systems. If you have different controls for different file systems, each file system should have its own control definition. A special virtual file system, **#SNFS_ALL#**, is used as the default control for file systems not defined in this control file. It is also used to define the **cvadmin**-related control on clients.

Note: You cannot have a file system named **#SNFS_ALL#**.

Each file system-related element (indicated by the label **securityControl**) has a list of **controlEntry** items. Each **controlEntry** item defines the client and the controls. The client type can be either **host** or **netgrp**. A host can be the IP address or the host name. (Both IP V4 and IP V6 are supported.) Netgrp specifies a group of consecutive IP addresses and has a network IP address (either IP V4 or V6,) and network mask bits. It is possible for there to be overlapping in IP address between an individual host and netgrp, but the individual host should be defined before the netgrp. If a client node has more than one IP address, then define the controls for each IP address.

The following controls are currently supported:

1 **mountReadOnly**: Control whether a client should mount as read-only. The default is read/write.

2 **mountDlanClient**: Control whether a client can mount as a proxy client. The default is not allowed.

3 **takeOwnership**: Control whether users on a Windows client are allowed to take ownership of a file or directory in a StorNext file system.

4 **snfsAdmin**: Controls whether **cvadmin** running on a host is allowed to have super admin privilege to run privileged commands such as start/stop fs.

5 **snfsAdminConnect**: Controls whether **cvadmin** running on a client is allowed to connect to other fsm via the **-H** option.

6 **exec**: Controls whether binary executable files on the file system are allowed to be executed. The default value is "true" (that is, the execution is allowed).

7 **suid**: Controls whether the **setuid** bit is allowed to take effect. (The default value is "true".)

If no match is found for a given client's IP address, the client has no privileges. If a file system has been defined but the client is not defined in that file system's control section (**securityControl**), the client has no access privileges to the specified file system.

Cluster-Wide Control File Example

The control file called **nss_cct1.xml** provides a feature overview and describes how to configure the central control file. The values entered in the following sample file are for example purposes only and should not necessarily be entered in your own control files.

```
<?xml version="1.0" encoding="UTF-8"?>

<!-- Cluster-Wide Central Control File -->

<!-- The nss_cctl.xml file provides a way to restrict the behavior of -->
<!-- SNFS cluster nodes (fsm, file system client, cvadmin client) from -->
<!-- a central place, i.e on nss server. As for SNFS 3.5, we support -->
<!-- the following controls: -->
<!-- 1. Whether a client is allowed to mount as a proxy client -->
<!-- 2. Whether a client is allowed to mount as read/write or read-only. -->
<!-- 3. Whether a user especially local admin on windows is allowed to -->
<!--    take ownership of file or directory on a Stornext file system. -->
<!-- 4. Whether cvadmin running on certain client is allowed to have -->
<!--    super admin privilege to run destructive commands such as -->
<!--    start/stop file system, refresh disks, and change quota setting, -->
<!--etc. -->
```

```
<!-- 5. whether cvadmin running on certain client is allowed to connect -->
<!--    to other fsm via "-H" option. -->
<!-- 6. whether an executable file on the file system can be executed.  -->
<!-- 7. whether to allow set-user-identifier bit to take effect. -->

<!-- The control file is in xml format and has hierarchical structure. -->
<!-- The top level element is "snfsControl", it contains control element -->
<!-- "securityControl" for certain file system. If you have different -->
<!-- controls for different file systems, then each file system should -->
<!-- has its own control definition. A special virtual file system  -->
<!-- "#SNFS_ALL#" is used as the default control for file systems not -->
<!-- defined in this control file. It is also used to define the cvadmin -->
<!-- related control on clients. Note: you cannot have a real file -->
<!-- system named as "#SNFS_ALL#". -->

<!-- Each file system related control element (securityControl) has a -->
<!-- list of "controlEntry", each "controlEntry" defines the client and -->
<!-- the controls. A client can be of type "host" or "netgrp". A "host" -->
<!-- can be ip address or host name. Both IP V4 and IP V6 are supported. -->
<!-- "netgrp" specifies a group of consecutive ip addresses. It has  -->
<!-- network ip address (either IP V4 or V6) and network mask bits. It -->
<!-- is possible that there is overlapping in ip addresses between  -->
<!-- individual host and netgrp, but the individual host should be  -->
<!-- defined before the netgrp. In this case, the netgrp control is -->
<!-- considered to be generic case, while the controls for individual is -->
<!-- considered to be special case. A special case takes precedence. -->

<!-- Currently there are seven controls supported: -->
<!-- 1. mountReadOnly: control whether a client should mount as -->
<!--    readonly. The default is read/write. -->
<!-- 2. mountDlanClient: control whether a client can mount as proxy  -->
<!--    client, the default is "mount not allowed". -->
<!-- 3. takeOwnership: control whether users on a windows client is  -->
<!--    allowed to take ownership of file or directory of a stornext -->
<!--    file system. The default is "take ownership not allowed". -->
<!-- 4. snfsAdmin: whether cvadmin running on a host is allowed to have -->
<!--    super admin privilege to run privileged commands such as  -->
<!--    start/stop fs. The default is that super admin privilege is not -->
<!--    honored. -->
<!-- 5. snfsAdminConnect: whether cvadmin running on a client is allowed -->
```

```
<!--    to connect to other fsm via "-H" option. The default is "-H" is -->
<!--    not allowed. -->
<!-- 6. exec: whether binary files on the file system is allowed to -->
<!--    be executed.  -->
<!-- 7. suid: whether set-user-identifier bit is allowed to take effect. -->

<!-- If no match is found for a given client's ip address, then the -->
<!-- client has no privilege to access a SNFS cluster. If a file system -->
<!-- has been defined but the client is not defined in that file  -->
<!-- system's control (securityControl), then the client has no access -->
<!-- privilege to the specified file system. -->

<!-- Currently only Linux platform is supported to be a nss server -->
<!-- capable of parsing this xml file. -->

<!-- The following is an example to define the nss_cctl.xml. It defines -->
<!-- the control of file system "snfs", and also the special virtual -->
<!-- file system "#SNFS_ALL#".   -->

<snfsControl xmlns="http://www.quantum.com/snfs/cctl/v1.0">
    <securityControl fileSystem="snfs">
<controlEntry>
    <client type="host">
<hostName value="192.168.230.132"/>
    </client>
    <controls>
<mountReadOnly value="true"/>
<mountDlanClient value="true"/>
<takeOwnership value="false"/>
<exec value="true"/>
<suid value="false"/>
    </controls>
</controlEntry>
<controlEntry>
    <client type="netgrp">
<network value="192.168.1.0"/>
<maskbits value="24"/>
    </client>
    <controls>
<takeOwnership value="true"/>
```

```
<mountReadOnly value="true"/>
    </controls>
</controlEntry>
    </securityControl>
    <securityControl fileSystem="#SNFS_ALL#">
<controlEntry>
    <client type="host">
<hostName value="linux_ludev"/>
    </client>
    <controls>
<snfsAdmin value="true"/>
<snfsAdminConnect value="true"/>
    </controls>
</controlEntry>
    </securityControl>
</snfsControl>
```

Commonly Used SNFS Files

The following is a list of commonly used SNFS files and provides the name and location of the files installed during a SNFS installation. Each entry also includes a brief description of the file's utility in SNFS functionality.

- **/usr/cvfs/bin/**

  cvadmin — Allows you to view and modify the active SNFS system(s).

  cvaffinity — Associates a stripe group with a media type.

  cvcp — A high-speed file copy utility.

  cvdb — A client debugging tool.

  cvdbset — A wrapper for cvdb.

  cvfsck — A utility to check and repair a SNFS file system.

  cvfsd — SNFS server daemon launched at mount time.

  cvfsdb — A metadata debugging tool.

cvfsid — Provides the SNFS site identifier

cvlabel — Lists and labels SNFS drives.

cvmkdir — Associates an affinity with a directory.

cvmkfile — Pre-allocates space and attributes for a file.

cvmkfs — Initializes a SNFS file system

cvupdatefs — Modifies the configuration of an existing SNFS file system.

cvversions — Lists SNFS version information.

fsm — A SNFS server daemon.

fsmpm — A SNFS portmapper daemon.

mount_cvfs — A mount helper for SNFS.

snfsdefrag — Defragments files and file systems.

- **/usr/cvfs/config/**

  fsmlist — A list of SNFS file systems that have been serviced.

  fsnameservers — Identifies the SNFS name service coordinators.

  license.dat — The SNFS license file.

- **/usr/cvfs/data/<FS Name>/log/**

  Cvlog — debug and runtime messages

- **/usr/cvfs/debug/**

  cvfsd.out — Displays debug and status messages from cvfsd.

  fsmpm.out — Displays debug and status messages from fsmpm.

  mountall.out — Displays debug and status messages from mount_cvfs.

  nssdbg.out — Displays debug and status messages from fsmpm.

- **/usr/cvfs/CLIENTS/** — The following are client installation files for various operating systems.

  sn_dsm_aix52_client.tar

  sn_dsm_irix65f_client.tar

  sn_dsm_irix65m_client.tar

  sn_dsm_linuxRH_80i386smp_client.tar

sn_dsm_linuxRH_80i386up_client.tar

sn_dsm_linuxRH_AS_3i386smp_client.tar

sn_dsm_linuxRH_AS_3ia64smp_client.tar

sn_dsm_linuxSuSE_81i386smp_client.tar

sn_dsm_linuxSuSE_81i386up_client.tar

sn_dsm_solaris58sparc64_client.tar

sn_dsm_solaris59sparc64_client.tar

sn_dsm_win2k_client.exe

sn_dsm_winnt_client.exe

- **/usr/cvfs/docs/**

  external_api.pdf — Documentation for the SNFS API.

- **/usr/cvfs/examples/** — The following are examples of files that are present in certain file system configurations. For example, the cvfail.example will be in the configuration file of a system running failover.

  cvfail.example

  cvlabels.example

  cvpaths.example

  example.cfg

  fsmlist.example

  fsnameservers.example

  fsports.example

  fsroutes.example

  nss_cctl.example

- **/usr/cvfs/lib/**

  cvextapi.a — A SNFS API library.

- **libcvfsdmapi.a** — A SNFS DMAPI library.

- **/usr/cvfs/src/**

  cvapi.c — A sample API application.

Appendix D
Quality of Service Guide

This appendix describes how to configure and use the StorNext File System (SNFS) Quality of Service (QOS) feature. QOS allows real-time applications to reserve some amount of bandwidth on the storage system. This is known as real-time I/O (RTIO). SNFS gates (that is, throttles) non-real-time applications so their I/O accesses do not interfere with the real-time application.

QOS is a *passive* implementation in that it does not actively monitor a process' activity and then schedule the process so that it receives the bandwidth it has requested. It is up to real-time applications to gate their own I/O requests to the desired rate. SNFS QOS provides a "get out of the way" gating for non-real-time I/O requests so they do not hinder the real-time requests.

The remainder of this document explains the client and server configuration settings for QOS; describes the use of tokens and callbacks for gating non-real-time I/O; describes setting real-time I/O on a file; and discusses the performance monitoring tools that are available for diagnosis.

Overview

QOS is stripe-group centric; all configuration and operations act on a specific stripe group. This allows stripe groups with differing access characteristics to use different real-time settings. When a client requests a certain amount of real-time I/O, the FSM informs all connected clients of the amount remaining and available for non-real-time I/O accesses.

This information is encapsulated in a token. Think of the token as a capability that enables the client to perform I/O without contacting the FSM. When the amount of real-time I/O on a stripe group changes, the FSM informs all connected clients about the change via a callback.

In both the client and the server, QOS is implemented as a state machine. Each state has a set of events and allowable actions. In this document, the term *state* thusly refers to the internal state machine.

Active vs. Passive

QOS is a passive, not active implementation of real-time I/O. In an active implementation (such as the SGI IRIX guaranteed rate I/O known as GRIO), the scheduler is tightly coupled with the I/O subsystem. The qualities of the disk subsystem are well known so the scheduler can guarantee that a process will be scheduled such that it will receive the required amount of bandwidth. Since SNFS is a cross-platform file system that does not have hooks in the operating system scheduler, it cannot provide such a guarantee.

In a passive implementation, a real-time process gates its I/O according to some outside metric (such as a frame rate for specific video formats). The file system then gates all other non-real-time I/O so they do not interfere.

It is very important that the differences between Active and Passive OoS operation are understood by the user. It is a misconception to think that QOS, despite its name, guarantees a specific amount of real-time I/O to a process.

Supported Platforms

QOS has been tested on Windows XP, Linux, IRIX, and Solaris. In Windows, an application gets a handle to a file to perform I/O, usually via the **Win32 CreateFile() API**. In UNIX, an application receives a file

descriptor (**fd**) via the **open(2)** system call. In this document, "handle" is synonymous with **fd**.

Non-Realtime Operation

Beginning with StorNext release 3.5, QOS has been enhanced to reserve bandwidth for non-realtime I/O clients. Differentiating between realtime and non-realtime I/O clients should result in better bandwidth utilization and overall performance.

During the mounting process, there is now provision to specify the amount of bandwidth to be reserved. Applications that do not have awareness of the external API can leverage the bandwidth amount. Also, clients with different load characteristics can be differentiated.

Non-realtime I/O support is provided through a central configuration file on the FSM server side, which defines the bandwidth reservation for non-realtime I/O requests on certain clients. This bandwidth reservation (RVIO) has higher priority than normal non-realtime clients, but lower priority over the realtime application (RTIO). FSM will always attempt to first satisfy realtime requests, and then follow the bandwidth reservation request. The remaining bandwidth is evenly distributed to all active non-realtime clients (excluding RVIO reserved clients).

To summarize, bandwidth priority is allocated in the following order:

- Realtime applications (RTIO)
- Bandwidth reservation (RVIO)
- Non-realtime I/O requests

Note: If there is not sufficient bandwidth to accommodate the bandwidth reservation request (RVIO), the client receives whatever bandwidth remains after RTIO requests have been granted. If there are multiple RVIO clients, then each client is allocated a RVIO bandwidth directly proportional to their requested bandwidth.

Configuration

The primary configuration for QOS is in the FSM configuration file. No client configuration is required, although there is a QOS tuning parameter that can be specified when the file system is mounted.

Unit of I/O

Real-time I/O is based on well-formed I/O. This means that for the purposes of determining bandwidth rates, well-formed I/O is characterized as being a stripe width in size. This makes the best utilization of the disks in the stripe group and maximizes the transfer rate. Internally, non-real-time I/O is tracked by number of I/O operations per second. An I/O operation is a minimum of a file system block size, and a maximum of the file system block size multiplied by the stripe breadth

(FsBlocksize * StripeBreadth).

FsBlockSize <= I/O <= (FsBlocksize * StripeBreadth)

Converting MB/sec to I/O/sec

Typically, it is easier to qualify an I/O subsystem in terms of MB/sec that can be sustained. However, internally the file system tracks everything on an I/O/sec basis. Note that the file system tracks only non-real-time I/O (that is, it gates only non-real-time I/O). An I/O is a minimum of the file system block size, and is typically the point at which the file system hands the request off to the disk driver (**IoCallDriver** in Windows, or a **strategy** call in UNIX).

The file system counts the number of I/Os that have taken place during a given second. If the number exceeds that which is allotted, the request is pended until I/O becomes available (typically in the next second). I/O is honored in FIFO fashion; no priority is assigned.

To convert between I/Os and MB/sec, SNFS uses a somewhat unique formula that quantifies I/O as well-formed. The rationale behind this is due to the way in which many video applications make real-time I/O requests. To optimize the disk subsystem, real-time I/Os are well-formed so they saturate the disks. In SNFS terminology, this would be an I/O that covers all of the disks in a stripe. This can be expressed as follows:

ios_sec = mb_sec /

(stripe_breadth * stripe_depth *fs_blocksize)

For example, with a file system blocksize of 4k, a **stripe_breadth** of 384, and a **stripe_depth** of four, the equivalent number of I/Os/sec for each well-formed I/O would be 216 Mb/sec / (384 * 4 * 4k). This is equivalent to 221184 k/sec / 6144k= 36 I/O/sec.

Server Configuration

All storage subsystems are different, so users must qualify the I/O subsystem and determine the maximum amount of I/O bandwidth available. SNFS relies on the correct setting in the configuration file; if the storage system changes (for example, because of a new disk array,) the user must re-qualify the I/O subsystem to determine the amount of bandwidth available. This amount will be specified in the FSM configuration file. The user can also specify the minimum amount of bandwidth to be provided to non-real-time applications.

There are five keywords controlling QOS that can be specified in the stripe group section of the FSM configuration file. Not all keywords need be present. Typically, the user specifies the RTIO bandwidth in terms of either number of I/O operations per second (**rtios**) or megabytes per second (**rtmb**). Keywords are not case sensitive.

For a minimum configuration, only the real-time limit (either **rtios** or **rtmb**) need be specified. All other configuration variables default to reasonable values.

Table 25 QOS Configuration Keywords

Name	Description	Default
Rtios	The maximum number of real-time I/Os allowed in a stripe group during any one-second period.	0 (no real-time)
Rtmb	Maximum amount of real-time MB/sec allowed on the stripe group during any one-second period.	0 (no real-time)

Name	Description	Default
RtiosReserve	Amount of reserve in I/Os/sec from the maximum allowed for non-real-time I/Os. Must be greater than the equivalent to 1MB/sec or the amount that can be transferred to a single stripe line.	Equivalent to 1MB/sec
RtmbReserve	Amount to reserve in MB/sec from the maximum allowed for non-real-time I/O.	Must be greater than 1. 1MB/sec
RtTokenTimeout	Time in seconds to wait for clients to respond to a token callback.	1.5 seconds

RTIO Upper Bound

The limit will be specified in terms of I/Os per second (parameter **Rtios**) or in terms of MB/sec (parameter **Rtmb**). Case is not sensitive. Note that I/Os per second are I/Os of any size to the disk subsystem. Either or both may be specified. If both are specified, the lower limit is used to throttle

I/O. If neither is specified, no real-time I/O is available on the stripe group. These parameters are applied to a stripe group definition.

Example (Linux)

```
<stripeGroup index="1" name="MyStripeGroup"
realTimeIOs="2048" realTimeMB="10">
</stripeGroup>
```

Example (Windows)

```
[StripeGroup MyStripeGroup]
   Rtios 2048
   Rtmb 10
```

The above example specifies that the storage system can support a maximum of 2048 I/Os per second at any instant, aggregate among all the clients, or 10 MB/sec, whichever is lower.

Most real-time I/O requests will be a stripe line at a time to maximize performance. Non-real-time I/Os will be a minimum of a file system block size.

Note: It is important to realize that the **rtios** and **rtmb** settings refer to the total amount of sustained bandwidth available on the disk subsystem. Any I/O, either real-time or non-real-time, will ultimately be deducted from this overall limit. The calculations of available real-time and non-real-time are discussed later.

Specifying **rtmb** in the FSM configuration file is only recommended if all I/Os are well formed (that is, a full stripe width). Otherwise, the conversion between MB/sec and I/Os/sec using the well-formed I/O calculation could lead to unexpected results.

Reserve

To prevent deadlock, the QOS implementation never allows zero I/O/sec for non-real-time I/O. Otherwise, a system could block with many critical file system resources held waiting for I/O to become available. This is especially true via flush-on-close I/O via the buffer cache. It becomes extremely difficult to diagnose system hangs because no I/O is available. For this reason, QOS always reserves some amount of I/O for non-real-time I/O.

The minimum amount of non-real-time I/O reserved for non-real-time applications is one MB/sec. This can be changed via the stripe group section parameters (again, case is not sensitive). If both are specified, the lower of the two amounts is chosen. This amount is shared by all non-real-time applications on each client.

Example (Linux)

```
<stripeGroup index="1" name="MyStripeGroup"
realTimeIOsReserve="256" realTimeMBReserve="2">
</stripeGroup>
```

Example (Windows)

```
[StripeGroup MyStripeGroup]
   RtiosReserve 256
```

```
RtmbReserve 2
```

Token Timeouts

The **RtTokenTimeout** parameter controls the amount of time the FSM waits for clients to respond to callbacks. In most normal SANs, the default two-second setting is sufficient. This value may need to be changed for a SAN that has a mixture of client machine types (Linux, NT, IRIX, etc.) that all have different TCP/IP characteristics. Also, large numbers of clients (greater than 32) may also require increasing the parameter.

For example, if the FSM should ever fail, the clients will attempt to reconnect. When the FSM comes back online, the amount of time the clients take to re-establish their TCP/IP connection to the FSM can differ wildly. To avoid unnecessary timeouts, the **RtTokenTimeout** parameter can be increased, meaning the FSM waits longer for callback responses.

If a client times out on a token retraction, the original requestor receives an error from the FSM that includes the IP address of the offending client. This error is logged to **syslog**, and alternatively to the desktop on Windows clients. This can help in diagnosing reconnect failures, and in determining if the token time value should be increased.

Client Configuration

When a client obtains a non-real-time I/O token from the FSM, the token allows the client a specific amount of non-real-time I/O. If the client is inactive for a period of time, the token is relinquished and the non-real-time I/O released back to the FSM for distribution to other clients. The timeout period is controlled by the **nrtiotokenhold** mount option on UNIX platforms, and the **QOS Token Hold** Time parameter in the mount options tab of the SNFS control panel on Windows platforms. The default is sixty (60) seconds.

This means that after sixty seconds without non-real-time I/O on a stripe group, the non-real-time token for that stripe group is released. The parameter should be specified in five (5) second increments. If it is not, it will be silently rounded up to the next five-second boundary. If the **syslog** level is set to **debug**, the file system dumps out its mount parameters so the value can be seen.

Real-time I/O

A process requests real-time (ungated) I/O by using the SNFS External API **SetRtio** call (**F_SETRIO ioctl**). A library function is included in the External API sample source code that provides all the required cross-platform handling.

As an example, assume that a video playback application requires a constant rate of 186 MB/sec to correctly display images without dropping any frames. The application gates itself; that is, it requests I/O at a rate to satisfy the requirements of correctly displaying an image. QOS provides a mechanism so other I/O requests do not perturb the real-time display.

In the following example, assume the I/O subsystem has been qualified at 216 MB/sec. The file system block size is 4k. The disk subsystem is actually a large RAID array that internally maps many drives to a single LUN. There are four LUNs in the stripe group; each LUN is optimized for a 1.5 MB transfer. This corresponds to the following in the fsm configuration file:

Example (Linux)

```
<stripeGroup index="1" name="MyStripeGroup"
stripeBreadth="384" realTimeMB="216">
   <disk index="0" diskLabel="CvfsDisk0"
   diskType="VideoDrive"/>
   <disk index="0" diskLabel="CvfsDisk1"
   diskType="VideoDrive"/>
   <disk index="0" diskLabel="CvfsDisk2"
   diskType="VideoDrive"/>
   <disk index="0" diskLabel="CvfsDisk3"
   diskType="VideoDrive"/>
</stripeGroup>
```

Example (Windows)

```
[StripeGroup MyStripeGroup]
     StripeBreadth 384
```

```
Node CvfsDisk0 0
Node CvfsDisk1 1
Node CvfsDisk2 2
Node CvfsDisk3 3
Rtmb 216
```

Also, assume there is only one stripe group for user data in the file system. As recommended by Quantum, there may be other stripe groups for metadata and journal that are not shown.

SetRtio

Initially, all stripe groups in the file system are in non-real-time mode. Clients make their requests directly to the I/O subsystem without any gating. In our example, the process requires 186 MB/sec and the system designers know there will never be a need to support more than one stream at 186 MB/sec.

The **SetRtio** request has a number of flags and parameters to control its operation. These are all documented in the **external_api.pdf** file that describes the external API in detail. For this example, set the handle for the indicated stripe group using the **RT_SET** parameter.

Oversubscription

In most cases, system designers ensure that the amount of **RTIO** is not oversubscribed. This means that processes will not ask for more **RTIO** than is specified in the configuration file. However, it is possible to request more **RTIO** than is configured. The API uses the **RT_MUST** flag to indicate that the call must succeed with the specified amount. If the flag is clear, the call allocates as much as it can. In both cases, the amount allocated is returned to the caller.

Handles

The **SetRtio** call accepts two different types of handles. The first is a handle to the **root** directory. In this mode the stripe group is put into real-time mode, but no specific file handle is tagged as being **ungated**. Real-time I/O continues on the stripe group until it is explicitly cleared with a **SetRtio** call on the **root** directory that specifies the **RT_CLEAR** flag; the file system is unmounted; or the system is rebooted. It is up to the application to make a subsequent call to **EnableRtio** (**F_ENABLERTIO**) on a specific handle.

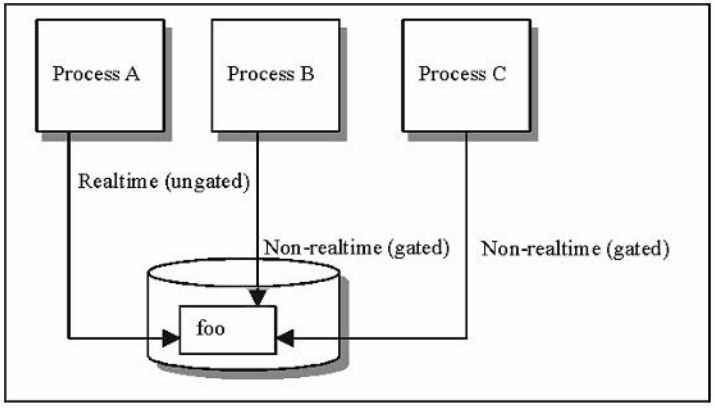
If the handle in the **SetRtio** call refers to a regular file, it is the equivalent of a **SetRtio** call on the root directory followed by an **EnableRtio** call. The file handle will be ungated until it is closed, cleared (**RT_CLEAR** in a **SetRtio** call), or disabled (**DisableRtio**). When the handle is closed, the amount of real-time I/O is released back to the system. This causes the FSM to readjust the amount of bandwidth available to all clients by issuing a series of callbacks.

The client automatically issues a call to the FSM with the **RT_CLEAR** flag specifying the amount of real-time I/O set on the file. If multiple handles are open on the file—each with a different amount of real-time I/O—only the last file close triggers the releasing action; all aggregate **RTIO** are released.

This automatic clearing of real-time I/O is carried out in the context of the process that is closing the file. If the FSM cannot be reached for some reason, the request is enqueued on a daemon and the process closing the file is allowed to continue. In the background, the daemon attempts to inform the FSM that the real-time I/O has been released.

Different processes can share the same file in real-time and non-real-time mode. This is because the level of gating is at the handle level, not the file level. This allows a real-time process to perform ingest of material (video data) at the same time as non-real-time processes are performing other operations on the file.

Figure 43 Sharing Access to Files



In Figure 43, Process A has ungated access to file **foo**. Processes B and C also are accessing file **foo**, but the client gates their I/O accesses. If multiple handles are open to the same file and all are in real-time mode, only the last close of the handle releases the real-time I/O back to the

system. This is because on most platforms the file system is informed only on the last close of a file.

Ungated files

It is also possible to denote using the **RT_NOGATE** flag that a handle should not be gated without specifying any amount of real-time I/O. This is useful for infrequently accessed files (such as index files) that should not be counted against the non-real-time I/O. System designers typically allow for some amount of overage in their I/O subsystem to account for non-gated files.

Calculating Available RTIO

When the FSM receives a request for RTIO, it takes the amount reserved into consideration. The reserve amount functions as a soft limit beyond which the FSM will not traipse. The calculation for **RTIO** is as follows:

avail_rtio = rtio_limit - rtio_current

avail_rtio -= rtio_reserve

In the above calculation, **rtio_limit** is the stripe group's maxmimum number of IOs (the value of parameter **Rtios**), **rtio_current** is the total number of currently reserved real-time IOs, and **rtio_reserve** is the minimum IOs reserved for non-realtime IOs (the value of **RtiosReserve**).

All internal calculations are done in terms of I/O/sec.

Non-Realtime I/O Configuration

In order to reserve bandwidth for non-realtime I/O (RVIO) operations on certain clients, you must specify the amount in the QOS central configuration file. This file is located under **$CVFSROOT/config/** (for example, **/usr/cvfs/config)**. The configuration file name consists of the file system name and the suffix **_rvio.opt**. For example, the central configuration file name for a file system called **fs1** would be **fs1_rvio.opt**.

The configuration doesn't affect the normal RTIO requests; it affects only the bandwidth allocation to non-RTIO requests. Note the following characteristics of RVIO:

1. RTIO has higher priority and is always satisfied first. RVIO uses whatever bandwidth is left.
2. RVIO reservation is not guaranteed. Instead, it is adjusted dynamically. When there are new RVIO or RTIO requests, the bandwidth reserved for clients with previous RVIO requests is readjusted. Each RVIO client will get the RVIO bandwidth reserved proportional to its request amount.
3. The RVIO bandwidth allocated for a client is the total bandwidth for non-RTIO requests on the client. The bandwidth (token) will not return like a normal non-RTIO client does if it's been idle for a hold time. Non-RTIO requests on clients with an RVIO reservation is subject to the same throttling mechanism as non-RTIO requests on normal clients.
4. Other active non-RTIO clients equally share the bandwidth left over after the possible RTIO and RVIO allocations. If no bandwidth is left, the minimal reserved bandwidth (defined in the stripe group section of the file system configuration file) is used.

A central configuration file has multiple entries, and each entry defines the RVIO reservation for a client. Following are the fields of an entry:

host bw-type sg=yy,[sg=yy]

- **host** is the host name, IP address (either V4 or V6), or FQDN, etc. of the client. The host name should be able to be resolved (converted to a valid IP address) by the FSM server. If there are multiple IPs on the host, use the IP that is used to communicate with the FSM.
- **bw-type** is the type of bandwidth to be specified. There are two bandwidth type options:
  - **qosmb** means the subsequent value for bandwidth reservation uses MB/s (megabytes per second) as the unit of measure.
  - **qosios** means the subsequent value for bandwidth reservation uses IOs/s (inputs/outputs per second) as the unit of measure.
- **sg** stands for stripe group. You may specify a stripe group by entering the stripe group name, or by stripe group number starting from 0 for the first stripe group. If you'd like to specify the same RVIO bandwidth reservation for all eligible stripe groups, you can use the asterisk wildcard in the format ***=yy**.

Any line in the configuration file starting with the pound sign (#) is considered a comment line and is not executed.

For additional information about entering the configuration file for QOS, see the man page called **qos_config.1**.

Configuration File Examples

Following are some example configuration file entry examples:

```
box1.xyz.com qosios 2=200,4=50,sg5=45
10.10.1.8     qosmb *=60
```

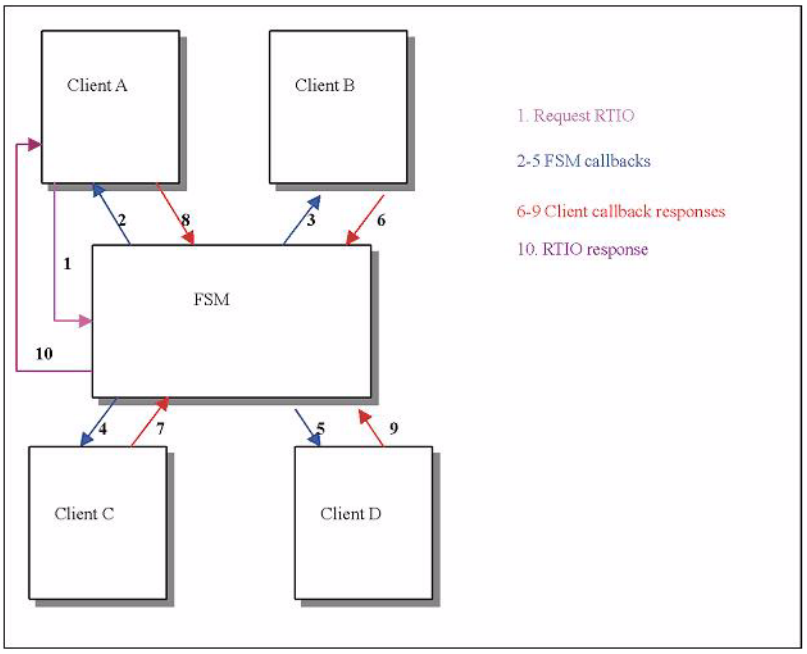
Callbacks

The cornerstones of the communications between the FSM and the client are callbacks and tokens. A callback is an unsolicited message from the FSM to the client requesting that the client adjust its real-time I/O parameters. The callback contains a token that specifies the amount of non-real-time I/O available on a stripe group.

Initially, all stripe groups in a file system are in non-real-time (ungated) mode. When the FSM receives the initial request for real-time I/O, it first issues callbacks to all clients informing them that the stripe group is now in real-time mode. The token accompanying the message specifies no I/O is available for non-real-time I/O. Clients must now obtain a non-real-time token before they can do any non-real-time I/O.

After sending out all callbacks, the FSM sets a timer based on the **RtTokenTimeout** value, which by default is set to 1.5 seconds. If all clients respond to the callbacks within the timeout value the **RTIO** request succeeds, and a response is set to the requesting client.

Figure 44 Callback Flow for Initial RTIO Request



In the above diagram, a process on client A requests some amount of **RTIO** in Step 1. Since this is the first request, the FSM issues callbacks to all connected clients (Steps 2-5) informing them that the stripe group is now in real-time mode. The clients respond to the FSM in Steps 6-9. After all the clients have responded, the FSM responds to the original requesting client in Step 10.

If the timer expires and one or more clients have not responded, the FSM must retract the callbacks. It issues a response to the requesting client with the IP number of the first client that did not respond to the callback. This allows the requesting client to log the error with the IP number so system administrators have a chance of diagnosing the failure. It then sends out callbacks to all the clients to which it first sent the callbacks, retracting them to the original state. In our example, it would set the stripe group back to non-real-time mode.

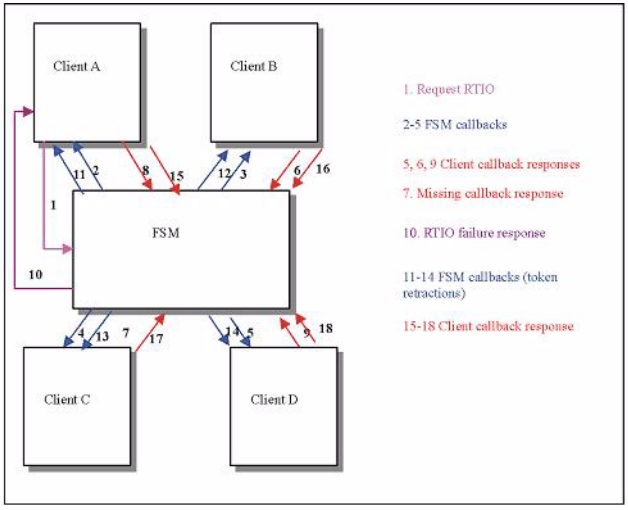
After sending out the callbacks, the FSM waits for a response using the **RtTokenTimeout** value as before. If a client again does not respond within the timeout value, the callbacks are retracted and sent out again. This repeats until all clients respond. During this time of token retractions, real-time requests cannot be honored and will only be enqueued.

Callback Failures

The FSM must handle a case where a client does not respond to a callback within the specified timeout period (**RtTokenTimeout**). If a client does not respond to a callback, the FSM must assume the worst: that it is a rogue that could wreak havoc on real-time I/O. It must retract the tokens it just issued and return to the previous state.

As mentioned earlier, the original requestor will receive an error (**EREMOTE**) and the IP address of the first client that did not respond to the callback. The FSM enters the token retraction state, and will not honor any real-time or token requests until it has received positive acknowledgement from all clients to which it originally sent the callbacks.

Figure 45 Callback Retraction Example



In Figure 45, Client A requests some amount of **RTIO** as in Figure 44. However, assume that Client C did not respond to the initial callback in time (step 7). The FSM will return a failure to Client A for the initial **RTIO** request, then send out callbacks to all clients indicating the stripe group is no longer real-time (steps 11-14). In the example, Client C responds to the second callback, so the FSM will not send out any more callbacks. The stripe group is back in non-real-time mode.

Note that this can have interesting repercussions with file systems that are soft mounted by default (such as Windows). When the caller times out because other clients are not responding and then gives up and returns an error to the application, if at some point the FSM is able to process the **RTIO** request it may result in the stripe group being put into

real-time mode after the original caller has received an error code. Both the FSM and clients log their actions extensively to **syslog**, so if this situation arises it can be detected.

In Figure 45, if the stripe group were already in real-time mode the FSM would only send out callbacks to those clients that already have tokens. Once all clients responded to the token callbacks, the stripe group would be back in its original state.

Tokens

A token grants a client some amount of non-real-time I/O for a stripe group. Tokens are encapsulated in callback messages from the FSM. Initially, no tokens are required to perform I/O. Once a stripe group is put into real-time mode, the FSM sends callbacks to all clients informing them that they will need a token to perform any non-real-time I/O. The first I/O after receiving the callback will then request a non-real-time I/O token from the FSM.

The FSM calculates the amount of non-real-time bandwidth using the following formula:

avail_nrtio = rtio_limit - rtio_current - rvio_current;

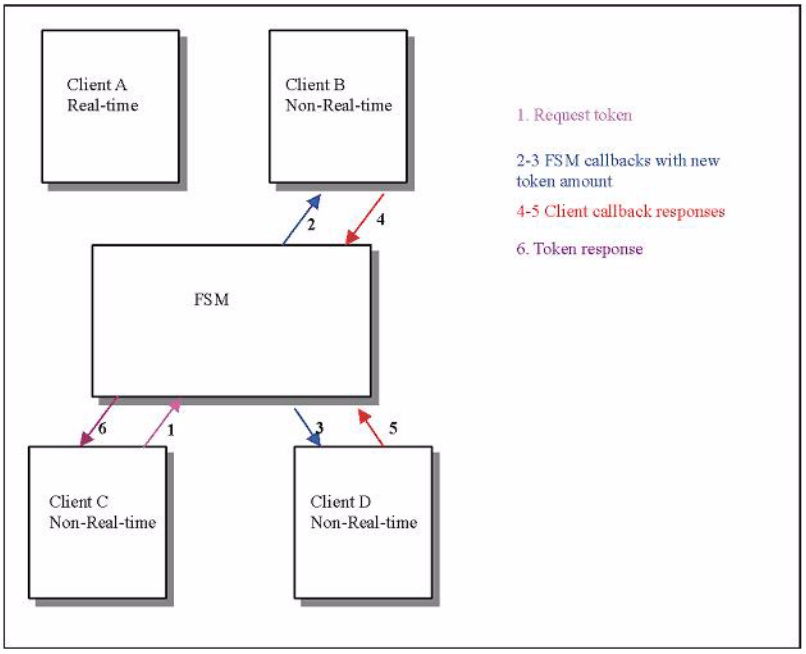
avail_nrtio /= current_num_nonrtio_clients + 1

In the above calculation, **rvio_current** is the total bandwidth reserved by current RVIO clents. The amount of existing real-time I/O (**rtio_current**) has already been adjusted with the reserve parameter. As each client requests a non-real-time I/O token, the number of clients increases (**current_num_nonrtio_clients** in the above formula) and the amount of available non-real-time I/O decreases.

Each time there is a change in the amount of non-real-time I/O available, the FSM sends callbacks to the clients with tokens. It is important to note that unlike the initial set of callbacks where the FSM sent callbacks to all connected clients, it is now only necessary to send callbacks to those clients that have an existing token.

Once a client has a token, it can perform as much I/O per second as is allowed by that token. It does not need to contact the FSM on every I/O request. The FSM will inform the client whenever the token changes value.

Figure 46 Non-Real-time Token Adjustments



In Figure 46, assume the stripe group is already in real-time mode as a result of an **RTIO** request from client A. Clients B and D are doing non-real-time I/O to the stripe group and have a token that specifies the amount of non-real-time I/O available. Client C then requests a non-real-time I/O token in Step 1. The FSM calls back to Clients B and D and specifies the new token amount in Steps 2-3. The clients respond in Steps 4-5, acknowledging the new token amount. The FSM then responds to Client C with the new token in Step 6.

Failure Semantics

There are two major failures that affect QOS: FSM crashes and client crashes. These can also be loss of communication (network outages). For client and server failures, the system attempts to readjust itself to the pre-failure state without any manual intervention.

FSM Failures

If the FSM crashes or is stopped, there is no immediate affect on real-time (ungated) I/O. As long as the I/O does not need to contact the FSM for some reason (attribute update, extent request, etc.), the I/O will

continue. From the standpoint of QOS, the FSM being unavailable has no affect.

Non-real-time I/O will be pended until the FSM is re-connected. The rationale for this is that since the stripe group is in real-time mode, there is no way to know if the parameters have changed while the FSM is disconnected. The conservative design approach was taken to hold off all non-real-time I/O until the FSM is reconnected.

Once the client reconnects to the FSM, the client must re-request any real-time I/O it had previously requested. The FSM does not keep track of QOS parameters across crashes; that is, the information is not logged and is not persistent. Therefore, it is up to the clients to inform the FSM of the amount of required **RTIO** and to put the FSM back into the same state as it was before the failure.

In most cases, this results in the amount of real-time and non-real-time I/O being exactly the same as it was before the crash. The only time this would be different is if the stripe group is oversubscribed. In this case, since more **RTIO** had been requested than was actually available, and the FSM had adjusted the request amounts, it is not deterministically possible to re-create the picture exactly as it was before. Therefore, if a deterministic picture is required across reboots, it is advisable to not over-subscribe the amount of real-time I/O.

The process of each client re-requesting **RTIO** is exactly the same as it was initially; once each client has reestablished its **RTIO** parameters, the non-real-time I/O is allowed to proceed to request a non-real-time token. It may take several seconds for the SAN to settle back to its previous state. It may be necessary to adjust the **RtTokenTimeout** parameter on the FSM to account for clients that are slow in reconnecting to the FSM.

Client Failures

When a client disconnects either abruptly (via a crash or a network partition,) or in a controlled manner (via an unmount), the FSM releases the client's resources back to the SAN. If the client had real-time I/O on the stripe group, that amount of real-time I/O is released back to the system. This causes a series of callbacks to the clients (all clients if the stripe group is transitioning from real-time to non-real-time,) informing them of the new amount of non-real-time I/O available.

If the client had a non-real-time I/O token, the token is released and the amount of non-real-time I/O available is recalculated. Callbacks are sent

to all clients that have tokens informing them of the new amount of non-real-time I/O available.

Client Token Releases

While it is not a failure case, the handling of a client token release is exactly the same as in the case where the client disconnected. All clients retain non-real-time tokens for a fixed amount of time. The default is 60 seconds. This can be controlled via the **nrtiotokentimeout** mount option. After the specified period of inactivity (i.e., no non-real-time I/O on the stripe group), the client will release the token back to the FSM. The FSM will re-calculate the amount of non-real-time bandwidth available, and send out callbacks to other clients.

Therefore, if a situation exists where a periodic I/O operation occurs every 70 seconds, it would be beneficial to set the **nrtiotokentime** mount option to something greater than or equal to 70 seconds to cut down on system and SAN overhead.

Monitoring

The current real-time statistics are available via the **cvadmin** utility. The **show long** command has been enhanced to provide information as to the current limit, the minimum amount reserved for non-real-time I/O, the number of active clients, the amount currently committed, and the amount a non-real-time application could hope to get when requesting I/O.

Whenever the stripe group status changes (such as from non-real-time to real-time mode), an event is logged to **syslog** (system event log on Windows platforms).

On the NT platform, real-time performance is also available via the **perfmon** utility. There are counters for both the client (SNFS Client) and FSM (SNFS File System Server [FSS]). In the client, a number of **rtio_xxx** counters are available to track the number of real-time I/Os/sec, number of non-real-time I/O requests, non-real-time I/O starvation, and other counters. A full explanation of each counter is provided with the **perfmon** utility by clicking Explain. In the FSM, information about the number of outstanding non-real-time clients, available **RTIO**, and other QOS information is available.

Performance counters are not enabled by default in the client. To enable them, you must execute **cvdb -P** ; this toggles the state of collection. To turn off the **perfmon** counters, do another **cvdb -P**. To re-enable, again do **cvdb -P**, and so on. To clear all client counters to zero, disable, then re-enable the counters with **cvdb -P**.

Sharing Disk Resources

As a final example, consider the case of a full-featured disk subsystem (Redundant Array of Very Expensive Disks, or RAVED) using a very high speed interconnect. Many clients can share these disk arrays, but it is sometimes desirable to limit a client's access to the array. QOS provides a mechanism for political bandwidth management so that no one client can consume all the bandwidth of the channel interconnects.

In such a scenario, the stripe group would always be in real-time mode. Each client would have a token specifying the number of I/Os/sec permissible. If there is need to assign different reserved bandwidth for non-real-time clients, specify the client’s bandwidth reservation in the RVIO config file.

The foundation of such an approach is a simple program that puts the stripe group into real-time mode as soon as the FSM is up and servicing requests. An example of such a program is included in the source code for the External API.

Once the stripe group is in real-time mode, the bandwidth as specified in the FSM configuration file is shared by all clients. In cases such as this, the real-time limit (**rtios** or **rtmb**) is calculated to be the total bandwidth desired to each client times the number of possible clients.

As each client attempted to access the disk subsystem, it would obtain a token. The FSM would send out callbacks adjusting down the amount of bandwidth available. No one client would be allowed to exceed the threshold specified in the non-realtime token. This assures fairness among all the clients.