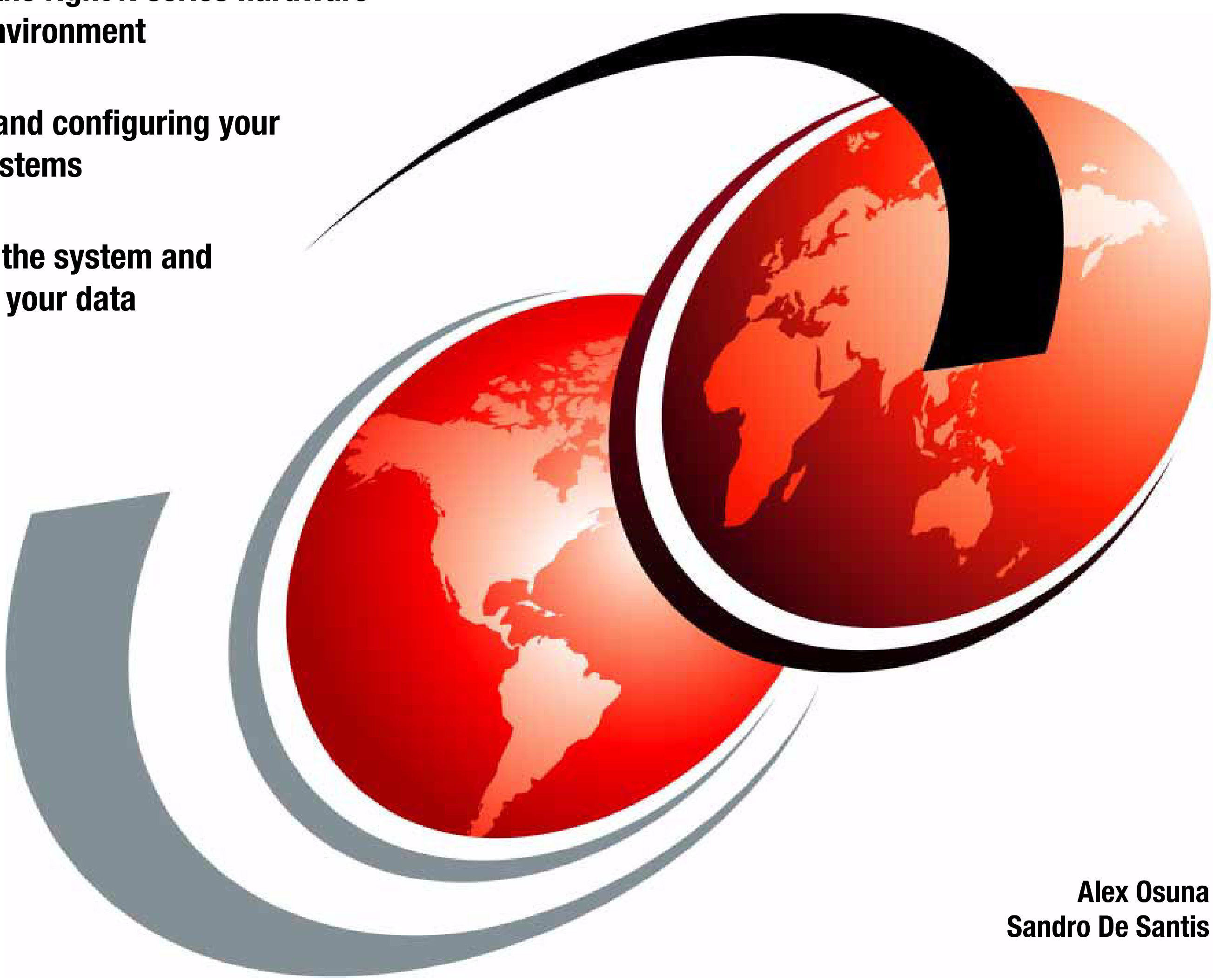
IBM System Storage N Series Hardware Guide

Selecting the right N series hardware for your environment

Installing and configuring your storage systems

Managing the system and protecting your data



Alex Osuna
Sandro De Santis

Redbooks

ibm.com/redbooks

International Technical Support Organization

IBM System Storage N Series Hardware Guide

October 2010

SG24-7840-00

Note: Before using this information and the product it supports, read the information in “Notices” on page ix.

First Edition (October 2010)

This edition applies to Data ONTAP Version 7.3.3 and above.

Contents

Notices ix
Trademarks x

Preface xi
The team who wrote this book xi
Now you can become a published author, too! xii
Comments welcome. xii
Stay connected to IBM Redbooks xii

Chapter 1. Introduction to IBM System Storage N series 1
1.1 IBM System Storage N series hardware 3
1.1.1 Maximum number of network interfaces 3
1.1.2 IBM System Storage N series A models hardware reference 4
1.1.3 IBM System Storage N series A and G models hardware quick reference 5
1.2 IBM System Storage N series standard software features 6
1.3 Optional software 7
1.4 Performance Accelerator Module (PAM) 9
1.5 IBM System Storage N series unified storage systems 9
1.5.1 IBM System Storage N3000 introduction 11
1.5.2 IBM System Storage N5000 introduction 18
1.5.3 IBM System Storage N6000 introduction 19
1.5.4 IBM System Storage N7000 introduction 32
1.6 IBM System Storage N series Gateways 38
1.6.1 IBM System Storage N series Gateway highlights 40
1.6.2 Gateway RAID 41
1.6.3 IBM System Storage N6040, N6060, and N6070 Gateway models 42
1.6.4 IBM System Storage N7700 and 7900 Gateway models 42
1.6.5 LUN sizing. 43
1.6.6 LUN mapping 43
1.7 N series expansion units 44
1.7.1 Intermixing EXN units with N series A models 44
1.7.2 EXN1000. 45
1.7.3 EXN4000. 46
1.7.4 EXN3000. 47

Chapter 2. Active/active configuration and management 49
2.1 Active/active overview 50
2.1.1 Benefits of active/active configurations 53
2.1.2 Standard active/active requirements and restrictions 54
2.1.3 Data ONTAP and active/active configuration 55
2.1.4 Active/active configuration node interaction 55
2.1.5 Required licenses 56
2.2 Configuring an active/active configuration 56
2.2.1 Configuration variations for standard active/active configurations 56
2.2.2 Best practices for active/active configurations 56
2.2.3 Enabling licenses on the active/active configuration. 57
2.2.4 Configuring network interfaces for the active/active configuration 58
2.2.5 Setting options and parameters 61
2.2.6 Testing takeover and giveback 62

2.2.7 Eliminating single points of failure with active/active configurations 63
2.2.8 Fibre Channel topologies for active/active configuration. 64
2.3 Managing an active/active configuration 85
2.3.1 Managing an active/active configuration in normal mode 85
2.3.2 Halting a node without takeover 85
2.3.3 Basic active/active configuration management 85
2.3.4 Active/active configuration failover basic operations. 91
2.3.5 Connectivity during failover 91
2.3.6 Reasons for active/active configuration failover 91
2.3.7 Failover due to disk mismatch. 92
2.3.8 Checking active/active configuration status 92
2.4 Active/active configuration failover mode (cfmode) for Fibre Channel 93
2.4.1 Summary of cfmode settings and supported systems 94
2.4.2 Cfmode restrictions for IBM System Storage N series Data ONTAP 7.3 Filer and Gateway 94
2.4.3 Setting cfmode 95
2.4.4 Single_image cfmode 96
2.4.5 Partner mode 102
2.4.6 Standby mode. 106
2.4.7 Dual fabric mode. 109
2.4.8 Mixed mode 112

Chapter 3. MetroCluster 117
3.1 Overview of MetroCluster 118
3.2 Benefits of using MetroCluster 119
3.3 Synchronous mirroring with SyncMirror 120
3.4 Business continuity with IBM System Storage N series 123
3.5 Implementing MetroCluster 123
3.5.1 Stretch MetroCluster 123
3.5.2 Fabric MetroCluster. 124
3.5.3 Fabric MetroCluster HBAs and cables 126
3.5.4 Using DWDM switches 126
3.5.5 Fabric-attached MetroCluster switches and drives 126
3.5.6 Tips for using MetroCluster. 127
3.6 MetroCluster: Cluster_Remote license 128
3.7 MetroCluster configurations 128
3.8 Cable selection 128
3.9 SyncMirror setup 130
3.10 Failure scenarios. 130
3.10.1 MetroCluster host failure. 131
3.10.2 MetroCluster IBM System Storage N series failure. 131
3.10.3 IBM System Storage N series and expansion unit failure. 132
3.10.4 MetroCluster interconnect failure 133
3.10.5 MetroCluster site failure 134

Chapter 4. FlexScale and the Performance Accelerator Module (PAM) 137
4.1 Overview of FlexScale 138
4.2 Performance Accelerator Module 138
4.2.1 How the PAM module accelerates performance. 139
4.2.2 Prerequisites for using the PAM II module 141
4.2.3 Supported N series configuration with PAM II cards. 141
4.3 FlexScale functionality and configuration 142
4.4 Checking FlexScale configuration. 143

4.4.1 Display FlexScale configuration 143
4.4.2 Display FlexScale usage and access information. 143

Chapter 5. Multipathing with the IBM System Storage N series 145
5.1 Overview 146
5.2 Data ONTAP for Windows MPIO 146
5.2.1 New load-balancing properties 147
5.2.2 Support for claiming iSCSI LUNS 147
5.2.3 Support for enhanced co-existence with other DSMs. 147
5.2.4 Multi-protocol LUN support (iSCSI and FCP paths to the same LUN) 147
5.2.5 Failover and load balance policy concepts 148
5.2.6 I_T and I_T_L nexus 149
5.2.7 Supported cluster mode settings for FCP 150
5.2.8 Performance factors 150
5.3 Installation of IBM Data ONTAP DSM. 151
5.3.1 Installation of IBM Windows Host Utilities Version 5.x 152
5.3.2 Data ONTAP DSM for Windows Version 3.2 R1. 154
5.3.3 Installation steps 155
5.3.4 Data ONTAP DSM repair option 157
5.4 Managing DSM using the GUI 158
5.5 Managing DSM using the CLI 161
5.5.1 Discovering new virtual disks (LUNs) 161
5.5.2 The main CLI commands 161
5.5.3 Using dcmcli dsm 161
5.5.4 Using dcmcli protocol 162
5.5.5 Using dcmcli lun 163
5.5.6 Using dcmcli path 164
5.6 Multiple path I/O support for Red Hat Linux 167
5.6.1 HBA installation and configuration 167
5.6.2 Uninstall previous HBA drivers 169
5.6.3 Verify native Linux multipath packages. 169
5.6.4 Download and install external HBA drivers. 170
5.6.5 Setting HBA and driver parameters 170
5.6.6 Bundled HBA drivers and applications 171
5.6.7 WWPN of your host HBAs 172
5.6.8 Configuring the N Series storage system for FC. 173
5.6.9 Loading the external HBA driver 173
5.6.10 Loading the bundled HBA driver 174
5.6.11 Configuring dm-multipath 174
5.6.12 Starting the multipath service 175
5.6.13 Accessing LUNs with dm-multipath support 175
5.6.14 Stopping dm-multipath support 178
5.7 Multiple path I/O support for Native AIX O/S. 179
5.7.1 Uninstalling Host Utilities kit 180
5.7.2 Installing the Host Utilities kit 181
5.7.3 N series FCP cluster mode 183
5.7.4 Creating, configuring, and managing LUNs without VIO. 183
5.7.5 Creating volume groups and file systems on AIX hosts 186
5.7.6 Using sanlun to display host LUN information 189
5.7.7 Displaying host HBA information. 191

Chapter 6. Disk sanitization 193
6.1 Data ONTAP disk sanitization. 194

6.2 Data confidentiality 195
6.2.1 Background 195
6.2.2 Technology drivers 195
6.2.3 Costs and risks 196
6.3 Data ONTAP sanitization operation 196

Chapter 7. System Manager 199
7.1 Introduction to System Manager 200
7.2 System Manager installation 201
7.3 System Manager use 201
7.3.1 Initial setup 212
7.3.2 System Manager and advanced features 212

Chapter 8. IBM System Storage N series administration 219
8.1 Administration methods. 220
8.2 Starting, stopping, and rebooting the storage system 221
8.2.1 Starting the IBM System Storage N series storage system 221
8.2.2 Stopping the IBM System Storage N series storage system 222
8.2.3 Rebooting the system 227
8.3 Checking the Data ONTAP software version 228
8.3.1 Data ONTAP version 8.0-7 mode 229
8.4 Storage management 229
8.4.1 Locating a failed disk 229
8.4.2 Adding new disks 236
8.5 AutoSupport Service 236

Chapter 9. Architecting an N series solution 239
9.1 Primary issues affecting planning 240
9.2 Performance and throughput 240
9.2.1 Effects of optional features 244
9.2.2 Future expansion 244
9.2.3 Application considerations 245
9.2.4 Backup servers 247
9.2.5 Backup and recovery 248
9.2.6 Resiliency to failure 249
9.3 Summary 250

Chapter 10. Boot from SAN with IBM System Storage N series 251
10.1 Overview 252
10.2 Configure SAN boot for IBM System x servers 253
10.2.1 Configuration limits and recommendations 253
10.2.2 Best practices 254
10.2.3 Basics of the boot process 256
10.2.4 Configuring SAN booting before installing Windows or Linux systems 257
10.2.5 Windows 2003 Enterprise SP2 installation 275
10.2.6 Windows 2008 Enterprise installation 277
10.2.7 Red Hat Enterprise Linux 5.2 installation 283

Appendix A. Getting started 287
Preinstallation planning 288
Collecting documents 288
Initial worksheet for setting up the nodes 288
Start with the hardware 303
Connecting your system to an ASCII terminal console 304

Power on N series 305
Data ONTAP update 309
Obtain Data ONTAP software from IBM NAS website 310
Install Data ONTAP system files 312
Download Data ONTAP to the storage system 317
Network setup using console 319
Change IP address 320
DNS setup 321

Appendix B. Power checklists 331
Power checklist for using an IBM cabinet 332
  Determining type of rack, power connector, and receptacle needed 332
  Power (line) cords 332
Power checklists for using a customer cabinet 333

Appendix C. N series hardware universe reference tables 335

Related publications 343
IBM Redbooks 343
Other publications 343
Online resources 343
How to get Redbooks 343
Help from IBM 344

Index 345

Notices

This information was developed for products and services offered in the U.S.A.

IBM may not offer the products, services, or features discussed in this document in other countries. Consult your local IBM representative for information on the products and services currently available in your area. Any reference to an IBM product, program, or service is not intended to state or imply that only that IBM product, program, or service may be used. Any functionally equivalent product, program, or service that does not infringe any IBM intellectual property right may be used instead. However, it is the user's responsibility to evaluate and verify the operation of any non-IBM product, program, or service.

IBM may have patents or pending patent applications covering subject matter described in this document. The furnishing of this document does not give you any license to these patents. You can send license inquiries, in writing, to:
IBM Director of Licensing, IBM Corporation, North Castle Drive, Armonk, NY 10504-1785 U.S.A.

The following paragraph does not apply to the United Kingdom or any other country where such provisions are inconsistent with local law: INTERNATIONAL BUSINESS MACHINES CORPORATION PROVIDES THIS PUBLICATION "AS IS" WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESS OR IMPLIED, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF NON-INFRINGEMENT, MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. Some states do not allow disclaimer of express or implied warranties in certain transactions, therefore, this statement may not apply to you.

This information could include technical inaccuracies or typographical errors. Changes are periodically made to the information herein; these changes will be incorporated in new editions of the publication. IBM may make improvements and/or changes in the product(s) and/or the program(s) described in this publication at any time without notice.

Any references in this information to non-IBM websites are provided for convenience only and do not in any manner serve as an endorsement of those websites. The materials at those websites are not part of the materials for this IBM product and use of those websites is at your own risk.

IBM may use or distribute any of the information you supply in any way it believes appropriate without incurring any obligation to you.

Information concerning non-IBM products was obtained from the suppliers of those products, their published announcements or other publicly available sources. IBM has not tested those products and cannot confirm the accuracy of performance, compatibility or any other claims related to non-IBM products. Questions on the capabilities of non-IBM products should be addressed to the suppliers of those products.

This information contains examples of data and reports used in daily business operations. To illustrate them as completely as possible, the examples include the names of individuals, companies, brands, and products. All of these names are fictitious and any similarity to the names and addresses used by an actual business enterprise is entirely coincidental.

COPYRIGHT LICENSE:

This information contains sample application programs in source language, which illustrate programming techniques on various operating platforms. You may copy, modify, and distribute these sample programs in any form without payment to IBM, for the purposes of developing, using, marketing or distributing application programs conforming to the application programming interface for the operating platform for which the sample programs are written. These examples have not been thoroughly tested under all conditions. IBM, therefore, cannot guarantee or imply reliability, serviceability, or function of these programs.

Trademarks

IBM, the IBM logo, and ibm.com are trademarks or registered trademarks of International Business Machines Corporation in the United States, other countries, or both. These and other IBM trademarked terms are marked on their first occurrence in this information with the appropriate symbol (® or ™), indicating US registered or common law trademarks owned by IBM at the time this information was published. Such trademarks may also be registered or common law trademarks in other countries. A current list of IBM trademarks is available on the Web at http://www.ibm.com/legal/copytrade.shtml

The following terms are trademarks of the International Business Machines Corporation in the United States, other countries, or both:

AIX®
DB2®
ESCON®
FICON®
IBM®
Redbooks®
Redbooks (logo) ®
System Storage®
System x®
Tivoli®
xSeries®

The following terms are trademarks of other companies:

Snapshot, SecureAdmin, RAID-DP, FlexShare, FlexCache, WAFL, VFM, SyncMirror, SnapVault, SnapValidator, SnapRestore, SnapMover, SnapMirror, SnapManager, SnapLock, SnapDrive, NearStore, FlexVol, FlexClone, FilerView, Data ONTAP, NetApp, and the NetApp logo are trademarks or registered trademarks of NetApp, Inc. in the U.S. and other countries.

Microsoft, Windows NT, Windows, and the Windows logo are trademarks of Microsoft Corporation in the United States, other countries, or both.

Celeron, Intel Xeon, Intel, Intel logo, Intel Inside logo, and Intel Centrino logo are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries.

UNIX is a registered trademark of The Open Group in the United States and other countries.

Linux is a trademark of Linus Torvalds in the United States, other countries, or both.

Other company, product, or service names may be trademarks or service marks of others.

Preface

This IBM® Redbooks® publication presents a detailed look at the features, benefits, and capabilities of the IBM System Storage® N series hardware offerings.

In an increasingly demanding and competitive business landscape, effective data management is essential to the success of your business. Employees, partners, and customers need access to up-to-date information to work productively, make timely decisions, and meet business goals. You are under pressure to deliver more storage capacity and at higher levels of service, yet critical resources—staffing, budget, power, cooling, and floor space—are often constrained.

These demands call for enterprise-class storage systems with the flexibility to support changing IT requirements, accommodate continuous data growth, satisfy application-level service requirements, and unify SAN and NAS infrastructures at a low overall cost.

The IBM System Storage N series systems are designed to help you tackle the challenge of effective data management using virtualization technology and a unified storage architecture. The N series delivers low to high-end enterprise storage and data management capabilities with midrange affordability. Built-in serviceability and manageability features help support your efforts to increase reliability; simplify and unify storage infrastructure and maintenance; and deliver exceptional economy.

The team who wrote this book

This book was produced by a team of specialists from around the world working at the International Technical Support Organization, San Jose Center.

Alex Osuna is a project leader at the International Technical Support Organization, Tucson Center. He writes extensively on IBM System Storage N series and Tape. Before joining the ITSO 5 years ago, Alex worked in the Tivoli® Western Region as a Systems Engineer. alex has over 32 years in the I/T industry focused mainly on storage. He has 29 years with IBM and holds certifications from IBM, Red Hat, Microsoft® and the Open Group.

Sandro De Santis Sandro De Santis is a Certified IT Specialist in Italy. He has ten years of experience in the storage field. He holds a degree in Nuclear Engineering. His areas of expertise include enterprise disk, virtualization, high performance computing. He has written extensively several papers on parallel sysplex, storage management, business continuity.

Thanks to the following people for their contributions to this project:

Todd Gray
Netapp

Kyle Burrell
IBM Integrated Storage Management - ITSM, ESS, SAN, Netapp, N series

Rucell Javier
IBM Client Technical Specialist Field Technical Sales

Roland Tretau
IBM Consulting IT Specialist - NAS & N Series Solutions

Paola Pescione
IBM FTSS manager, Storage Italy

Emma Jacobs
IBM Redbooks graphics specialist

Now you can become a published author, too!

Here's an opportunity to spotlight your skills, grow your career, and become a published author - all at the same time! Join an ITSO residency project and help write a book in your area of expertise, while honing your experience using leading-edge technologies. Your efforts will help to increase product acceptance and customer satisfaction, as you expand your network of technical contacts and relationships. Residencies run from two to six weeks in length, and you can participate either in person or as a remote resident working from your home base.

Find out more about the residency program, browse the residency index, and apply online at:

ibm.com/redbooks/residencies.html

Comments welcome

Your comments are important to us!

We want our books to be as helpful as possible. Send us your comments about this book or other IBM Redbooks publications in one of the following ways:

- Use the online **Contact us** review Redbooks form found at:

  ibm.com/redbooks

- Send your comments in an e-mail to:

  redbooks@us.ibm.com

- Mail your comments to:

  IBM Corporation, International Technical Support Organization
  Dept. HYTD Mail Station P099
  2455 South Road
  Poughkeepsie, NY 12601-5400

Stay connected to IBM Redbooks

- Find us on Facebook:

  http://www.facebook.com/IBMRedbooks

- Follow us on twitter:

  http://twitter.com/ibmredbooks

- Look for us on LinkedIn:

  http://www.linkedin.com/groups?home=&gid=2130806

- Explore new Redbooks publications, residencies, and workshops with the IBM Redbooks weekly newsletter:

  https://www.redbooks.ibm.com/Redbooks.nsf/subscribe?OpenForm

- Stay current on recent Redbooks publications with RSS Feeds:

  http://www.redbooks.ibm.com/rss.html

1

Introduction to IBM System Storage N series

The IBM System Storage N series offers additional choice to organizations facing the challenges of enterprise data management. The IBM System Storage N series is designed to deliver high-end value with midrange affordability. Built-in enterprise serviceability and manageability features help to support customer efforts to increase reliability, simplify and unify storage infrastructure and maintenance, and deliver exceptional economy.

In this chapter we introduce the IBM System Storage N series and describe its hardware features. The IBM System Storage N series provides a range of reliable, scalable storage solutions for a variety of storage requirements. These capabilities are achieved by using network access protocols such as Network File System (NFS), Common Internet File System (CIFS), HTTP, and iSCSI, as well as storage area network technologies such as Fibre Channel (FC). Utilizing built-in Redundant Array of Independent Disks (RAID) technologies, all data is well protected, with options to enhance protection through mirroring, replication, Snapshots, and backup. These storage systems are also characterized by simple management interfaces that make installation, administration, and troubleshooting straightforward.

This type of flexible storage solution enables you to:

- Tune the storage environment to a specific application while maintaining flexibility to increase, decrease, or change access methods with minimal disruption.
- React easily and quickly to changing storage requirements. If additional storage is required you can expand it quickly and non-disruptively. If existing storage is deployed incorrectly you have the capability to reallocate available storage from one application to another quickly and simply.
- Maintain availability and productivity during upgrades. If outages are necessary, they can be kept to the shortest time possible.
- Create effortless backup and recovery solutions that operate in a common manner across all data access methods.
- Simplify your infrastructure with file- and block-level services in a single system.

- Tune the storage environment to a specific application while maintaining its availability and flexibility.
- Change the deployment of storage resources non-disruptively, easily, and quickly. Online storage resource redeployment is possible.
- Easily and quickly implement the upgrade process. Non-disruptive upgrade is possible.
- Achieve strong data protection solutions with support for online backup and recovery.
- Include added value features such as deduplication to optimize space management.

All N series storage systems utilize a single operating system across the entire platform and offer a combination of multiple advanced function software features that provide one of the industry's most multifaceted storage platforms, including comprehensive system management, storage management, onboard copy services, virtualization technologies, and disaster recovery and backup solutions.

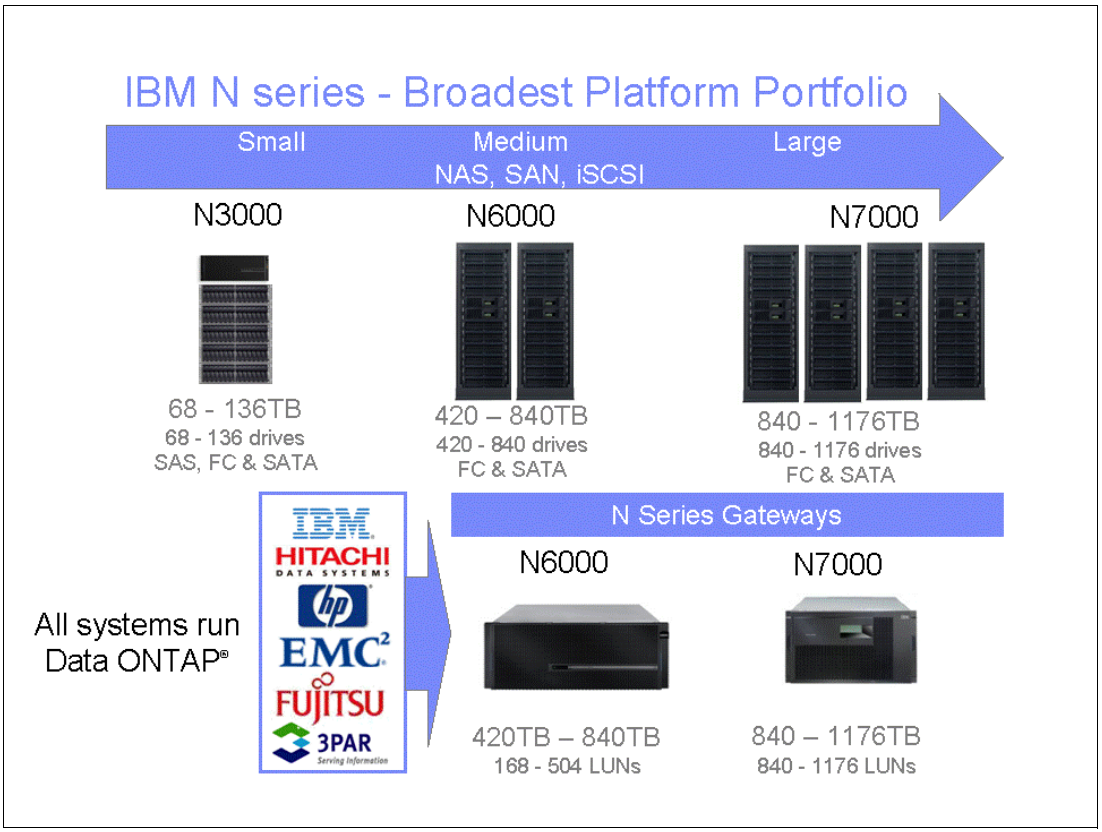


Figure 1-1 N series portfolio

1.1 IBM System Storage N series hardware

In the following sections we discuss the N series models available at the time of this writing. Table 1-1 identifies all the N series models released by IBM to date that belong to the N3000, N6000, and N7000 series.

Table 1-1 IBM System Storage N series storage systems

A1X and A2X models	Maximum drives	Maximum raw capacity in TB[1]
N3300	68	68
N3400	136	136
N3600[2]	104	104
N3700[3]	56	56
N5200[3]	168	168
N5300[3]	336	336
N5500[3]	336	336
N5600[3]	504	504
N6040[4]	420	420
N6060[2,4]	672	672
N6070[2,4]	840	840
N7600[3]	840	840
N7700[2,3]	840	840
N7800[3]	1008	1008
N7900[2]	1176	1176

1. The more dense disk drive available is the 2 TB 7200 RPM SATA Systems running Data ONTAP® 7.3.x can assign no more than seven 2 TB data disk drives to an aggregate due to maximum aggregate size limitations. These limitations do not include disks used for aggregate RAID parity or hot spares. Also the though you can populate the N series with the maximum disks allowed with 2 TB disks the maximum capacity allowed does not increase. 2TB SATA drives will have lower maximum number of spindle count.
2. Single node model no more available.
3. Withdrawn model.
4. The N6000 models series also acts as a gateway with Data ONTAP 7.3 and later.

1.1.1 Maximum number of network interfaces

Beginning with Data ONTAP 7.3, storage systems, except for the smallest models, are no longer limited to 128 interfaces per storage system. Larger storage systems can accommodate from 256 to 1,024 network interfaces per system depending on the system's memory and whether they are in a clustered, active/active configuration.

Run the `sysconfig` command and refer to the memory size field displayed for your storage system's slot 0 system board to determine your storage system memory.

Table 1-2 on page 4 shows the new maximum number of network interfaces allowed for each system. The total number of interfaces can include physical, Virtual Interface (VIF), virtual local area network (VLAN), vh, and loopback interfaces.

Table 1-2 N series Gateway models

G1X and G2X models	Maximum capacity
N6040	420 TB
N6060	672 TB
N6070	840 TB
N7700	840 TB
N7900	1176 TB

Using 2 TB SATA drives will lower the maximum spindle count. The number of physical interfaces depends on the storage system model. Up to 16 VIFs can be supported by each storage system. The maximum number of VLANs that can be supported equals the maximum number of network interfaces shown in Table 1-3 minus the total number of physical interfaces, VIFs, vh, and loopback interfaces supported by the storage system.

Table 1-3 Storage system memory and maximum number of interfaces

Storage system memory	Maximum number of interfaces
2 GB or less	128
2 GB or less in an active/active configuration	256
6 GB or less	256
6 GB or less in an active/active configuration	512
More than 6 GB	512
More than 6 GB in an active/active configuration	1024

1.1.2 IBM System Storage N series A models hardware reference

Table 1-4 provides a quick overview of the features of A model hardware.

Table 1-4 A model hardware quick reference

Feature	N3300	N3400	N3600	N6040	N6060	N6070	N7700	N7900
Maximum raw capacity in TB A1X models	68	136	104	420	672	840	840	1176
Maximum raw capacity in TB A2X models	68	136	104	420	672	840	840	1176
Fibre Channel disk drives	10 K RPM - 300 GB 15 K RPM - 144 GB, 300 GB, 450 GB, 600 GB							
Serial Advanced Technology Attachment (SATA) disk drives	7.2 K RPM - 500 GB, 1 TB, 2 TB							
Serial Attached SCSI (SAS) disk drives	15K RPM - 300 GB, 450 GB, 600 GB (SAS drives) 7,2K RPM 500 GB, 1 TB, 2 TB (SATA drives)							
Maximum number of disks	68	136	104	420	672	840	840	1176
Expansion units supported	EXN1000 (SATA), EXN2000 (FC 2 Gbps), EXN4000 (FC 4 Gbps), EXN3000 (SAS 3 Gbps)[1]							

1. EXN3000 supports only SAS drives and 500 GB, 1 TB, 2 TB SATA drives

Table 1-5 provides a quick overview of the features of N series G models.

Table 1-5 N series G models quick reference

Feature	N6040	N6060	N6070	N7700	N7900
Maximum raw capacity in TB G10/G20 models	420	672	840	840	1176
Max. number of logical units (LUNs) on back-end disk storage array	420	672	840	840	1176
Max LUN number	2048	2048	2048	1024	2048
Maximum volume size in TB[1]	16	16	16	16	16

1. Check also possible operating system LUN size dependencies.

1.1.3 IBM System Storage N series A and G models hardware quick reference

Table 1-6 provides a quick reference to IBM System Storage N series A and G model hardware.

Table 1-6 Storage system reference

Function	N3300	N3400	N3600	N6040	N6060	N6070	N7700	N7900
Network protocol support	NFS V2/V3/V4 over UDP or TCP, PCNFSD V1/V2 for (PC) NFS client authentication, Microsoft CIFS, iSCSI, FCP, VLD, HTTP 1.0, HTTP1.1 virtual host							
Other protocol support	SNMP, NDMP, LDAP, NIS, DNS,							
Onboard I/O ports per node	4 x GbE 4 x FC	8 x GbE 4 x FC 2 x SAS	4 x GbE 4 x cal FC	4 x GbE 8 x FC		4 x Gbe 8 x FC	6 x GbE 8 x FC	6 x GbE 8 x FC
PCI expansion slots per node	N/A	N/A	1 x PCI-E	4 x PCI-E	4 x PCI-E	4 x PCI-E	3 x PCI-x5 10 x PCI-E	3 x PCI-E 10 x PCI-E
Non-volatile RAM (NVRAM) in MB per node	128	256	256	512	2048	2048	2048	4096
Memory in GB per node	1	4	2	4	8	16	16	32
Rack space per node	2U for two nodes	2U for two nodes	4U for two nodes	6U for two nodes	6U for two nodes	6U for two nodes	6U	6U
Processors (A1X)	One 2.2 GHz Mobile Celeron®	One 1.66 GHz Intel® DUal Core Xeon	One 2.2 GHz Mobile Celeron	2.4 GHz AMD dual-core 64-bit Opteron	Two 2.6 GHz AMD dual-core 64-bit Opteron	Two 2.6 GHz AMD dual-core 64-bit Opteron	Two 2.6 GHz AMD dual-core Opteron	Four 2.6 GHz AMD dual-core Opteron
Processors (A2X)	Two 2.2 GHz Mobile Celeron	Two 1.66 GHz Intel Dual-core Xeon	Two 2.2 GHz Mobile Celeron	Two 2.4 GHz AMD dual-core 64-bit Opteron	Four 2.6 GHz AMD dual-core 64-bit Opteron	Four 2.6 GHz AMD dual-core 64-bit Opteron	Four 2.6 GHz AMD dual-core Opteron	Eight 2.6 GHz AMD dual-core Opteron

1.2 IBM System Storage N series standard software features

The following standard software features are available at no charge with IBM System Storage N series storage systems:

- Data ONTAP: Operating system software that optimizes data serving and allows multiple protocol data access.
- File Transfer Protocol (FTP): A standard Internet protocol, FTP is a simple way to exchange files between computers on the Internet.
- Telnet: Protocol that provides a general, bidirectional communications facility. It provides user-oriented command line login sessions between hosts.
- Snapshot™: Enables online backups, providing near instantaneous access to previous versions of data without requiring complete, separate copies.
- FlexVol®: Creates multiple flexible volumes on a large pool of disks and provides dynamic, non disruptive (thin) storage provisioning, thereby achieving space and time efficiency. These flexible volumes can span multiple physical volumes without regard to size.
- FlexShare™: Enables administrators to take advantage of existing infrastructure and increase processing utilization without sacrificing the performance of critical business needs. With FlexShare, administrators can confidently consolidate different applications and data sets on a single storage system and can prioritize applications based on how critical they are to the business.
- Disk sanitization: The process of physically obliterating data by overwriting disks with specified byte patterns or random data in a manner that prevents recovery of current data by any known recovery methods. This feature enables you to carry out disk sanitization by using three successive byte overwrite patterns per cycle. By default, six cycles are performed.
- FilerView®: Allows IT administrators to fully manage N series storage systems from remote locations. It provides simple and intuitive web-based single-appliance administration.
- SnapMover®: Migrates data among N series clusters with no impact on data availability and no disruption to users.
- AutoSupport: A sophisticated, event-driven logging agent featured in the Data ONTAP operating software and inside each N series system. It continuously monitors the health of your system and issues alerts if a problem is detected. These alerts can also be in the form of email.
- SecureAdmin™: This Data ONTAP module enables authenticated, command-based administrative sessions between an administrative user and Data ONTAP over an intranet or the Internet.
- Domain Name System (DNS): The N series supports using a host naming file or a specified DNS server and domain.
- Network Information Service (NIS): The N series provides NIS client support and can participate in NIS domain authentication.
- Integrated automatic Redundant Array of Independent Disks (RAID) manager: The IBM System Storage N series and Data ONTAP provide integrated RAID management with RAID-Double Parity (default) and RAID 4.
- iSCSI Host Attach Kit for AIX®, Windows®, Linux®: Includes support software and documentation for connecting a supported host to an iSCSI network. The support software includes programs that display information about storage and programs to collect information needed by customer support to diagnose problems.

- Systems Manager: Provides comprehensive management of one or more N series storage systems by way of a simple, easy-to-use, intuitive GUI.
- SyncMirror®: The synchronous mirror of a volume, SyncMirror maintains a strict physical separation between the two copies of your mirrored data. In case of an error in one copy, the data is still accessible without any manual intervention.
- Open System SnapVault® (OSSV) software: Enables Windows, UNIX®, and Linux servers to take advantage of the N series Snapshot processes (for example, block-level incremental backups) to reduce the amount of backup data flowing over the WAN while maintaining data integrity and recoverability.
- NearStore option: Provides enhanced performance in a disk-based, secondary storage device used for enterprise applications.
- Advanced Single Instance Storage (ASIS): Significantly improves physical storage efficiency and network efficiency by enabling the sharing of duplicate data blocks. ASIS provides a data-deduplication solution native to N series.

1.3 Optional software

The following optional, chargeable software features are available for the IBM System Storage N series:

- Common Internet File System (CIFS): Provides file system access for Microsoft Windows environments.
- NFS: Provides file system access for UNIX and Linux environments.
- Hypertext Transfer Protocol (HTTP): Allows a user to transfer displayable web pages and related files.
- FlexCache for NFS: Provides efficient caching of files and volumes in a local N series storage system when the source volume resides in a remote location N series storage system, thus avoiding inefficient use of bandwidth resources.
- FlexClone: Provides instant replication of data volumes and sets without requiring additional storage space at the time of creation.
- FlexScale: Used exclusively to improve performance by managing the additional cache provided by Performance Accelerator Modules. These modules can be added as an option to an N series storage system.
- Multistore: Permits an enterprise to consolidate a large number of Windows, Linux, or UNIX file servers onto a single storage system. Placing many virtual N series storage systems on one physical appliance eases migration and multi-domain failover scenarios.
- SnapLock: Provides non-erasable and non-rewritable data protection that helps enable compliance with government and industry records retention regulations. SnapLock is not available in Data ONTAP 7.3.0; it is available starting in Data ONTAP 7.3.1.
- SnapMirror: Remote mirroring software that provides automatic block-level incremental file system replication between sites. It is available in synchronous, asynchronous, and semi-synchronous modes of operation.
- SnapRestore: Allows rapid restoration of the file system to an earlier point in time, typically within a few seconds.
- SnapVault: Provides disk-based backup for N series systems by periodically backing up a Snapshot copy to another system.
- SnapDrive: Enables Windows and UNIX applications to access storage resources on N series storage systems, which are presented to the Windows 2000 or later operation

system as locally attached disks. For UNIX it allows you to create storage on N series storage systems in the form of logical unit numbers (LUNs), file systems, logical volumes, or disk groups.

- SnapManager: Host software for managing Snapshots for backup and restore operations. Versions are available to that integrate easily with critical applications, in particular:
  - SnapManager for MS Exchange
  - SnapManager for SQLServer
  - SnapManager for MS SharePoint
  - SnapManager for Oracle
  - SnapManager for SAP
  - SnapManager for Virtual Infrastructures, which automates and simplifies backup and recovery of primary storage used by VMWare Virtual Infrastructure
  - SnapManager for Hyper-V
- SnapValidator: For Oracle deployments, provides an additional layer of integrity checking between the application and N series storage. SnapValidator allows Oracle to create checksums on data transmitted to N series storage for writes to disk and include the checksum as part of the transmission.
- Single Mailbox Recovery for Exchange (SMBR): A software option from SnapManager that takes near-instantaneous online backups of Exchange databases, verifies that the backups are consistent, and rapidly recovers Exchange within levels (storage group, database, folder, single mailbox, or single message). The potential results are improved service to internal clients, reduced infrastructure expenses, and significant time savings for Exchange administrators.
- Operations Manager, File Storage Resource Manager (FSRM) feature: Provides monitoring and management of storage resources, including applications, files, file systems, and networks.
- Operations Manager Core: Provides remote, centralized management of IBM System Storage N series data storage infrastructure, including global enterprise, storage network, and so on.
- Provisioning Manager: Allows IT administrators to enter set provisioning policies, automate complex provisioning processes, check policy conformance, and pool storage resources for higher utilization. Provisioning Manager also enables server administrators to provision storage within the confines set by the storage administrator.
- MetroCluster: Provides an enterprise solution for high availability over wide area networks between two clustered nodes of a single N series storage system.
- Virtual File Manager (VFM): A comprehensive solution for managing unstructured file data. It is designed to provide data management functionality for server and storage consolidation, migration, remote office data management, and disaster recovery features, all while avoiding disruption to users. It provides all of this functionality through automated policy-based data management using a global namespace.
- Cluster failover (CFO): Ensures high data availability for business-critical requirements by eliminating a single point of failure. It must be ordered for A2X clustered configurations or upgrades from A1X to A2X. Its active/active pairing delivers even more “nines to the right of the decimal point.”

1.4 Performance Accelerator Module (PAM)

The performance of a disk system is heavily dependent on disk count as well as on the RPM of each disk. While a greater disk count and higher RPMs provide better performance, this also increases the power consumption and space usage in the data center. The Performance Accelerator Module has been introduced for the N6000, and N7000 series as a means of improving performance by adding modules with additional cache memory.

Each Performance Accelerator Module currently consists of:

- 256 GB of SLC flash memory per module (PAM II)
- 512 GB of SLC flash memory per module (PAM II)

It is possible to add more than one PAM to an N series storage system. PAMs can be configured to operate in three modes:

- Mode 1: Behaves more like main memory
- Mode 2: Preferentially caches meta data
- Mode 3: Caches what otherwise would have been flushed

Using tools based on predictive cache statistics it is possible to determine the optimal settings and the number of modules to deploy.

Performance Accelerator Modules are available as optional adapters in N5000, N6000, and N7000 N series systems. When including Performance Accelerator Modules it is necessary to add the FlexScale licensed software.

PAM II cards are available for N6070 and N7000 systems.

1.5 IBM System Storage N series unified storage systems

The IBM System Storage N series storage systems offer multiprotocol connectivity using internal storage or storage provided by expansion units (Figure 1-2 on page 10). The IBM System Storage N series systems are designed to provide integrated block-level and file-level data access, allowing concurrent operation in IP SAN (iSCSI), FC SAN, NFS, and CIFS environments. Other storage vendors might require the operation of multiple systems to provide this functionality. IBM System Storage N series storage systems are designed to avoid costly downtime, both planned and unplanned, and improve your access to mission-critical data, thereby helping you gain a competitive advantage.

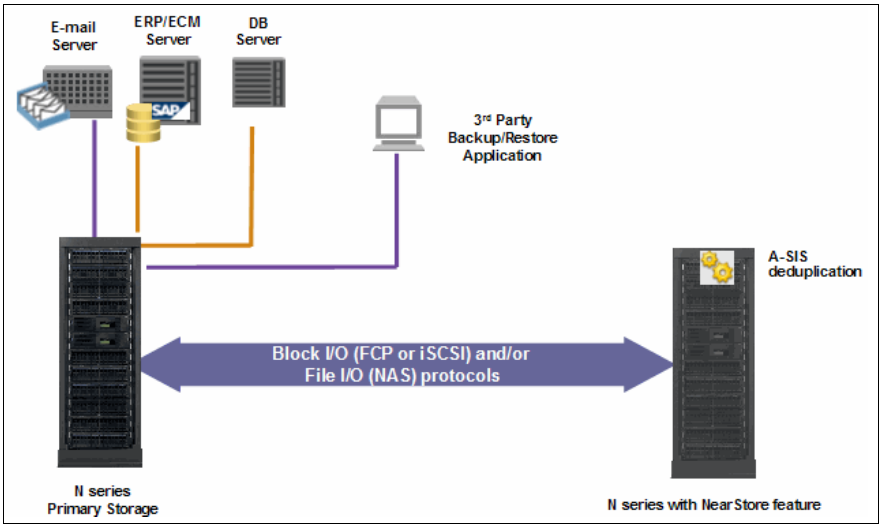


Figure 1-2 IBM System Storage N series A models

The N series is a specialized, *thin server* storage system with a customized operating system, similar to a stripped-down UNIX kernel, hereafter referred to as *Data ONTAP*. With this customized operating system, many of the server operating system functions that you are familiar with are not supported. Data ONTAP improves performance and reduces costs by eliminating unnecessary functions that do not pertain to a storage system.

N series units come with preconfigured software and hardware, and with no monitor or keyboard for user access. This is commonly called a *headless* system. A storage administrator accesses the systems and manages the disk resources from a remote console using a web browser or command line.

One of the typical characteristics of an N series storage systems product is its ability to be installed rapidly, using minimal time and effort to configure the system. The N series product is integrated seamlessly into the network, making it especially attractive when time and skills are limited in the customer's organization.

Drive flexibility

IBM System Storage N series products are designed to provide network-attached storage for environments where customers need to utilize their storage investment in a multifaceted environment. N series storage systems provide customers with a tremendous amount of versatility because they allow the solution to be populated with Fibre Channel, SAS disk drives, and SATA disk drives. An N series populated with Fibre Channel or SAS disk drives, or both, might be suitable for mission-critical high performance data transaction environments, whereas an N series populated with SATA disk drives might be attractive to customers that wish to use the platform for disk-to-disk backup scenarios, disaster recovery scenarios, archive data, or data like home directories, which do not have the demands of high performance transactional environments. Table 1-7 on page 11 identifies the appropriate drive types to satisfy various requirements.

Table 1-7 Drive positioning

Requirement	Fibre Channel drives	SAS drives	SATA drives
Online, high-performance, mission-critical production data repository	X	X	
Near-line storage used for tiered storage or infrequently accessed data	X	X	X
Data retention to help meet the needs of customers required to store data in non-erasable and non-rewritable (WORM) formats		X	X

Near-line storage

IBM System Storage N series with SATA drives offers near-line storage. Figure 1-3 shows an example of traditional disk-based backup and recovery. The left side shows the primary storage characterized by a higher cost and very fast performing system. On the far right the archive targets are represented traditionally by tape or optical jukeboxes. The N series secondary with SATA drives reduces access times to read and write data. A few years ago the concept of near-line storage in the middle for disk staging was introduced, enabling organizations to do daily backups to disk. Additionally, backups to tape could be done weekly or bi-weekly, thereby reducing the amount of data that must be written to tape. Moreover, data that is online is available for faster recovery. Another advantage is that this architecture makes use of your existing investment in primary storage, your backup application, and tape libraries.

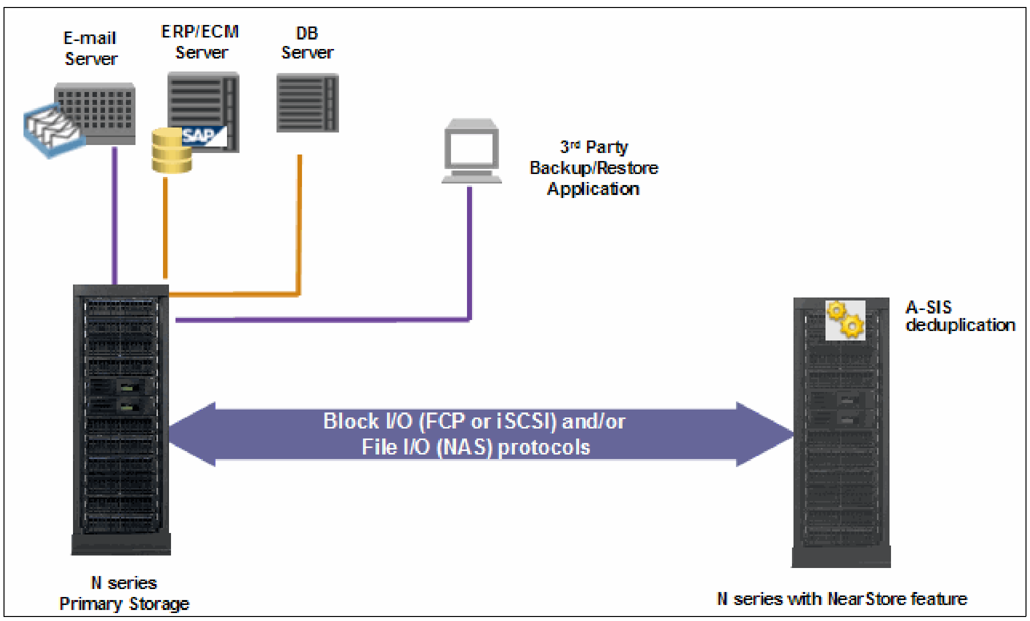


Figure 1-3 Near-line storage

1.5.1 IBM System Storage N3000 introduction

The N3000 systems are designed to provide primary and secondary storage for midsize enterprises. IT administrators can consolidate fragmented application-based storage and unstructured data into one unified, easily managed, expandable platform. N3000 systems offer integrated block-level and file-level data access, intelligent management software, and data protection capabilities (similar to those offered in higher end N series systems) in a cost-effective package. N series innovations include Serial-Attached SCSI (SAS) disk drive support, expandable I/O connectivity, and onboard remote management.

The N3000 systems are designed as the entry point to the entire N series family. These systems can provide the following key advantages:

- High availability: Takes advantage of proven features, including a high performing and scalable operating system, data management software, and redundancy features
- Backup and recovery features: Designed to support disk-based backup, with file- or application-level recovery with Snapshot and SnapRestore software features
- Simple replication and disaster recovery: Designed to provide an easy-to-deploy mirroring solution that is highly tolerant of wide area network (WAN) interruptions
- Management simplicity: Self-diagnosing systems designed to enable on-the-fly provisioning
- Versatile, single, integrated architecture: Designed to support concurrent block I/O and file serving over Ethernet and Fibre Channel SAN infrastructures

The N3000 series is compatible with the entire family of N series unified storage systems, which feature a comprehensive lineup of hardware and software designed to address a variety of possible deployment environments:

- N3300
  - 2859-A10 Single Node
  - 2859-A20 Clustered
- N3400
  - 2859-A11 Single Node
  - 2859-A21 Clustered
- N3600
  - 2862-A20 Clustered

The N3000 series supports Ethernet and Fibre Channel environments, enabling economical NAS, FC, and iSCSI deployments. The N3000 system functions as a *unification engine*, which is designed to allow you to simultaneously serve both file-level and block-level data across a single or multiple networks. Some solutions require multiple separately managed systems.

N3000 storage systems can offer significant advantages for distributed enterprises with remote and branch office sites. These organizations and others can use the SnapVault and SnapMirror software functions to implement a cost-effective data protection strategy by mirroring data to a central corporate data center.

There are no additional PCI adapter slots in the N3300 storage system. On the other hand, the N3600 storage system has one available PCIe adapter slot per node. For an A20 model, identical adapters must be added in pairs, one to each node, so that both nodes are populated with identical types of PCIe adapters.

N3300 and N3600

The N3300 and N3600 systems (Figure 1-4 and Figure 1-5 on page 13) provide multiple I/O connectivity options, a small footprint to hold high density Serial-Attached SCSI (SAS) drives, and external expansion using either low cost SATA drives or Fibre Channel disks for production applications while using Data ONTAP's Snapshot technology. SAS is the next generation of SCSI, which combines the advantages of parallel SCSI and serial FC. For further systems administration time and cost advantages, the systems come standard with

Remote Onboard Management capabilities to simplify remote system monitoring: managing cycle power, executing firmware upgrades, entering console commands, and running diagnostics to maintain the reliability of the system as well as business-critical data.

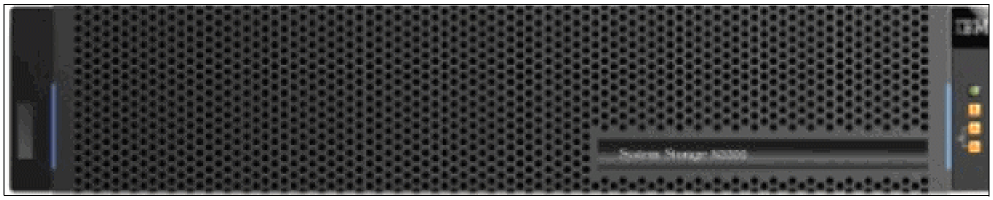


Figure 1-4 N3300

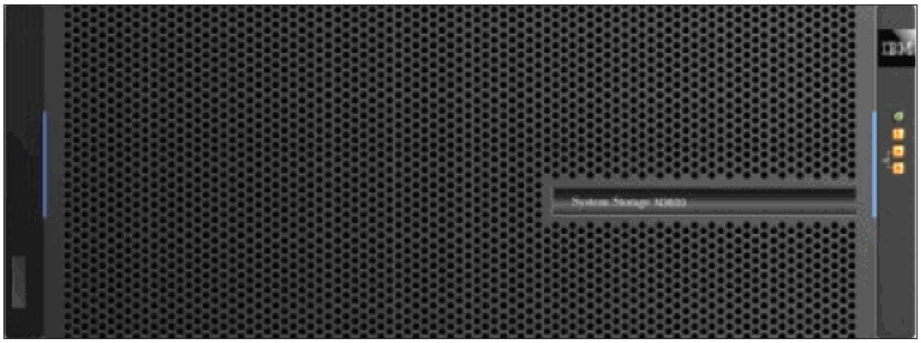


Figure 1-5 N3600

Figure 1-6 and Figure 1-7 show the rear views of N3300 and N3600, respectively. The single node models (A10) use a single control unit, while the dual-node clustered models (A20) use two control units.

Figure 1-6 N3300 rear view

Figure 1-7 N3600 rear view

N3300 is a 2U-high device with capacity for 12 internal SAS drive bays; it supports up to two external disk expansion units. Each controller has dual gigabit Ethernet ports, dual 4 Gbps Fibre Channel ports one console port, and one remote management port. These features are identified in Figure 1-8.

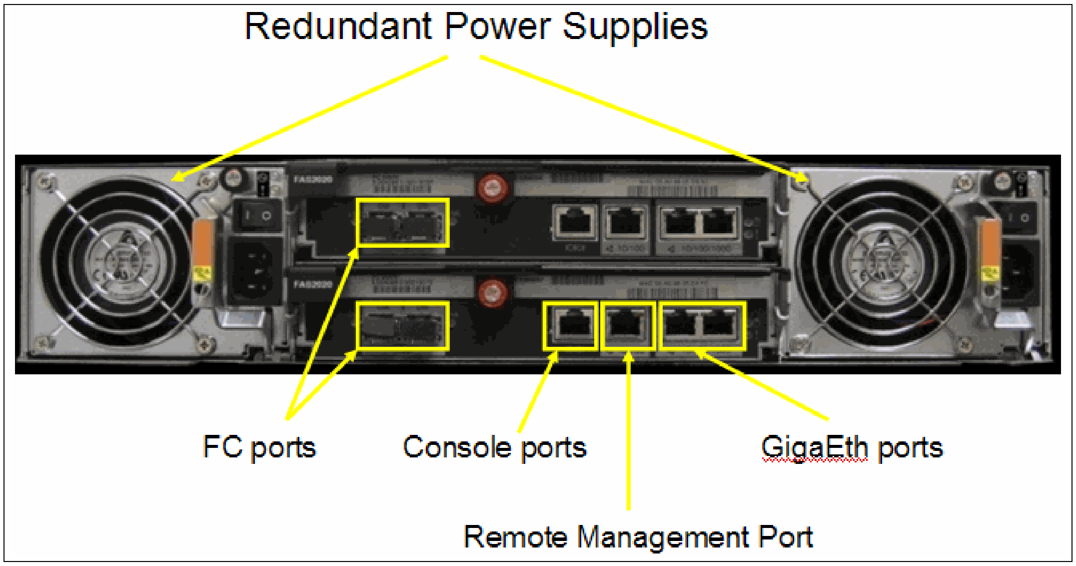


Figure 1-8 External ports on N3300

Tip: The N3300 series supports SAS, FC, and SATA disk technologies. The N3300 controller chassis supports from 6 to12 SAS or SATA disk drives. All disks must be of the same type. The N3300 controller can be configured with 0 disk drives and storage in disk expansion units like EXN1000/SATA or EXN4000/FC.

The N3600 has redundant power supplies (Figure 1-9).

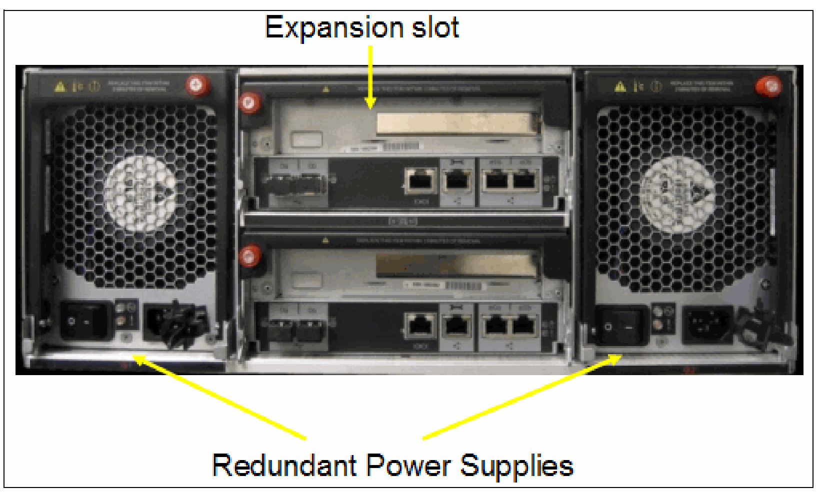


Figure 1-9 N3600 power supplies and expansion slot

N3600 is a 4U-high device with capacity for 20 internal SAS drive bays. N3600 can support up to six external disk expansion units. Each controller has dual gigabit Ethernet ports and dual 4 Gbps Fibre Channel ports (Figure 1-10). It also has one console port and one remote management port, as well as one PCIe slot on each controller for an expansion card.

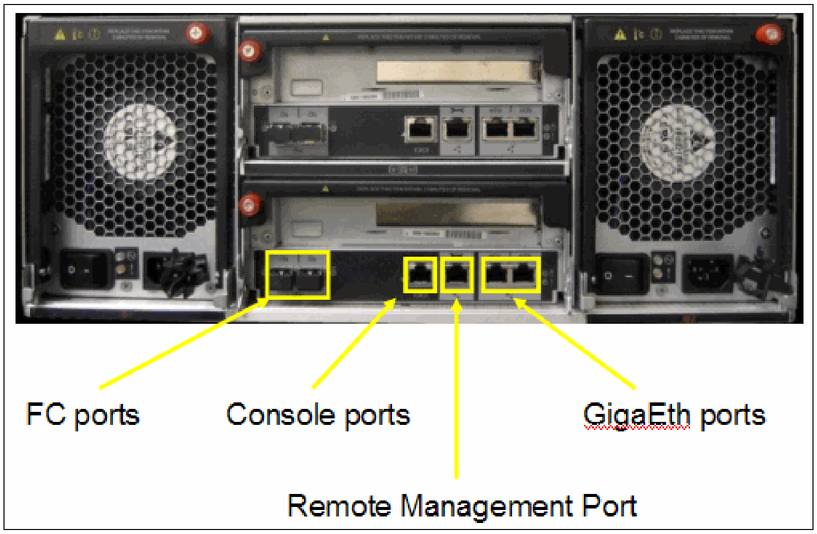


Figure 1-10 External ports on N3600

Tip: The N3600 series supports SAS, FC, and SATA disk technologies. The controller chassis supports either 20 SAS or 20 SATA disk drives. The N3600 requires a minimum of six SAS drives in the controller chassis.

The N3300/N3600 key specifications are:

- 2U high (N3300) or 4U high (N3600)
- Up to four external disk expansion units for N3300 and up to six external disk expansion units for N3600
- High-performance SAS infrastructure
- Single controller or dual controller (for HA)
- Unified storage: iSCSI, NAS, Fibre Channel
- Each controller: Dual gigabit Ethernet ports and dual 4 Gbps Fibre Channel ports
- Onboard remote platform management
- Internal SAS drive bays

The N3000 series is a small form-factor appliance that conserves valuable space in data centers or remote office locations and it is engineered for small to medium-sized enterprises.

N3400

N3400 is the newcomer of the N3000 family.

The IBM System Storage N3400 can provide primary and secondary storage for the midsize enterprise. It enables the IT department of such an organization to consolidate all of their distributed application-based storage and unstructured data into one unified, easily managed and expandable platform, thereby increasing their effectiveness. N3400 offers integrated block-level and file-level data access, intelligent management software, and data protection capabilities in a cost-effective package. The IBM System Storage N3400 and the other N3000 models provide innovation with internal controller support for Serial-Attached SCSI (SAS) or SATA drives, expandable I/O connectivity, and onboard remote management.

The new N3400 series can scale up to 24 TB of internal raw capacity and increase total raw capacity to 104 TB. Using 2 TB SATA drives will the lower maximum spindle count in the system.

Figure 1-11 and Figure 1-12 show front and back views of the N3400 controller module. In the rear panel both clustered controllers and stand-alone controller options are available.

.

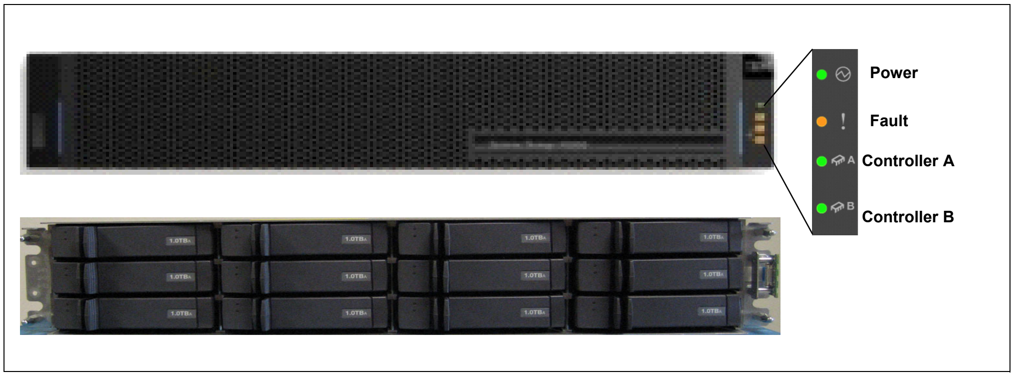


Figure 1-11 Front views of 3400 controller modules

N3400 is a 2U box with the capability to host up to 12 drives in the controller enclosure. If more capacity is needed the N3400 can be attached to external EXN1000 and EXN4000 units with SATA and FC drives. It is also possible to connect the controllers to the new EXN3000 shelf.

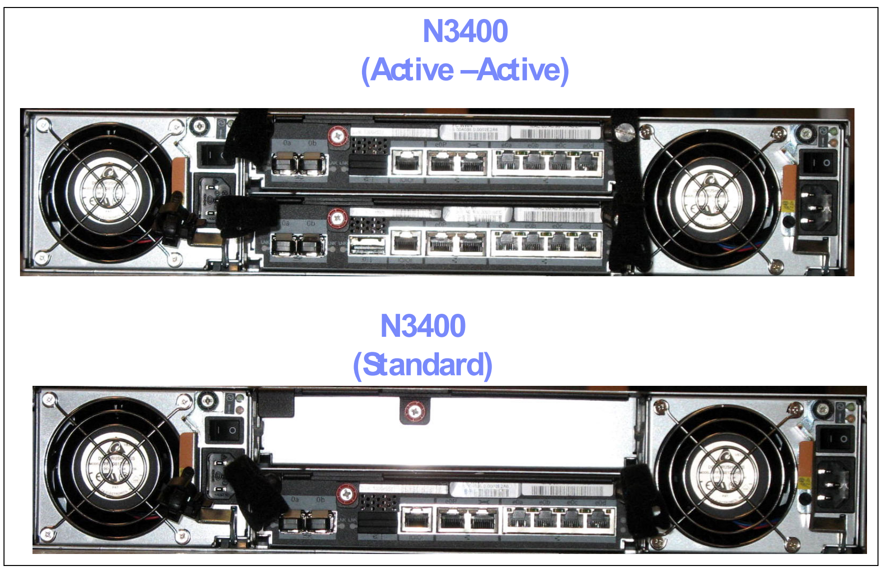


Figure 1-12 Comparison of N3400 configurations

N3400 has one SAS expansion port per controller with one Alternate Control Path (ACP). If you need to attach the EXN3000 shelf to the controller you can configure the shelf Alternate Control Path during the setup process. This will enable Data ONTAP to manage the EXN3000 on a separate network in order to increase availability and stability. The ACP is shown in Figure 1-13.

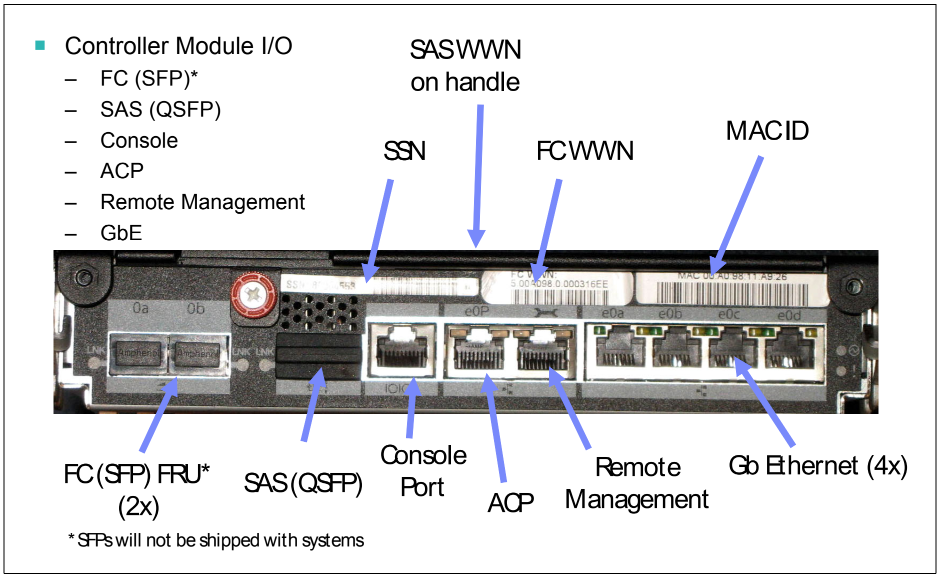


Figure 1-13 N 3400 communication ports

The N3400 has the following key specifications:

- 2U high
- Up to eight external EXN1000, EXN4000 expansion units (up to 6 in stack)
- High-performance SAS infrastructure
- Single controller or dual controller (for HA)
- Unified storage: iSCSI, NAS, Fibre Channel
- Each controller: Up to 8 gigabit Ethernet ports and two dual 4 Gbps Fibre Channel ports
- Onboard remote platform management
- Internal SAS drive bays

Starting from SAS firmware 0500 it is possible to perform a Non Disruptive Update (NDU) so disk I/Os are uninterrupted while SAS firmware is being updated.

Table 1-8 presents a summary of the specifications for the N3000 family.

Table 1-8 N3000 family specifications.

	N3300 series	**N3300 series**	**N3400 series**	**N3400 series**	**N3600 series**
Machine type Model	2859-A10	2859-A20	2859-A11	2859-A21	2862-A20
Controller configuration	Single	Dual Active/Active	Single	Dual Active/Active	Dual
GB of RAM	1	2	4	8	4
Fibre Channel ports[1]	2	4	2	4	2
Ethernet ports[2]	2	4	4	8	4
Maximum raw capacity in TB	68	68	136	136	104

	N3300 series	N3300 series	N3400 series	N3400 series	N3600 series
Number of disk drives	68	68	136	136	104
Maximum volume size in TB	8	8	16	16	16
Disk drives supported in controller (size, type, speed	SAS: 300 GB, 450 GB, 600 GB, 15,000 rpm SATA: 500 GB, 7,200 rpm; 1 TB, 2 TB		SAS: 300 GB, 450 GB, 600 GB 15,000 rpm SATA: 500 GB, 7,200 rpm; 1 TB, 2 TB		SAS: 300 GB, 450 GB, 600 GB, 15,000 rpm SATA: 500 GB, 7,200 rpm; 1 TB, 2 TB
Disk expansion units supported	EXN4000 - Fibre Channel Disk Storage Expansion Unit: • 4-Gbps Fibre Channel: 300 GB, 450 GB, 600 GB 15,000 rpm • 2-Gbps Fibre Channel: 300 GB, 450 GB, 600 GB 15,000 rpm EXN1000 - SATA Disk Storage Expansion Unit: • SATA: 500 GB, 7,200 rpm; 1 TB, 2 TB		EXN4000 - Fibre Channel Disk Storage Expansion Unit: • 4-Gbps Fibre Channel: 300 GB, 450 GB, 600 GB, 15,000 rpm • 2-Gbps Fibre Channel: 300 GB, 450 GB, 600 GB, 15,000 rpm EXN3000 - SAS Disk Storage Expansion Unit: • SAS: 300 GB, 450 GB, 600 GB, 15,000 rpm • SATA: 500 GB, 7,200 rpm; 1 TB, 2 TB EXN1000 - SATA Disk Storage Expansion Unit: • SATA: 500 GB, 7,200 rpm; 1 TB, 2 TB		EXN4000: • 4-Gbps Fibre Channel: 300 GB, 450 GB, 600 GB, 15,000 rpm • 2-Gbps Fibre Channel: 300 GB, 450 GB, 600 GB, 15,000 rpm EXN3000: • SAS: 300 GB, 450 GB, 600 GB, 15,000 rpm • SATA: 500 GB, 7,200 rpm; 1 TB EXN1000: • SATA: 500 GB, 7,200 rpm; 1 TB, 2 TB

1. 4 Gbps
2. 1 Gbps

1.5.2 IBM System Storage N5000 introduction

The N5300 and N5600 are suitable for environments that demand data in high availability, high capacity, and highly secure data storage solutions. The IBM System Storage N5000 series offers additional choices to organizations for enterprise data management. The IBM System Storage N5000 series is designed to deliver midrange to high end enterprise storage and data management value with midrange affordability. Built-in enterprise serviceability and manageability features help support your efforts to increase reliability, simplify and unify storage infrastructure and maintenance, and deliver exceptional economy.

The IBM N5000 A series comes in two models:

- N5300
  - 2869-A10 Single Node
  - 2869-A20 Clustered
- N5600
  - 2868 -A10 Single Node
  - 2868 -A20 Clustered

N5000 models can be configured for FC or SATA, or both.

N5000 models are no longer orderable.

The N5000 A10 models come in a compact 3U rack-mountable unit that can coexist in the same rack as an EXN1000, EXN2000, or EXN4000 storage expansion unit (Figure 1-14). The A20 models require 6U of space.

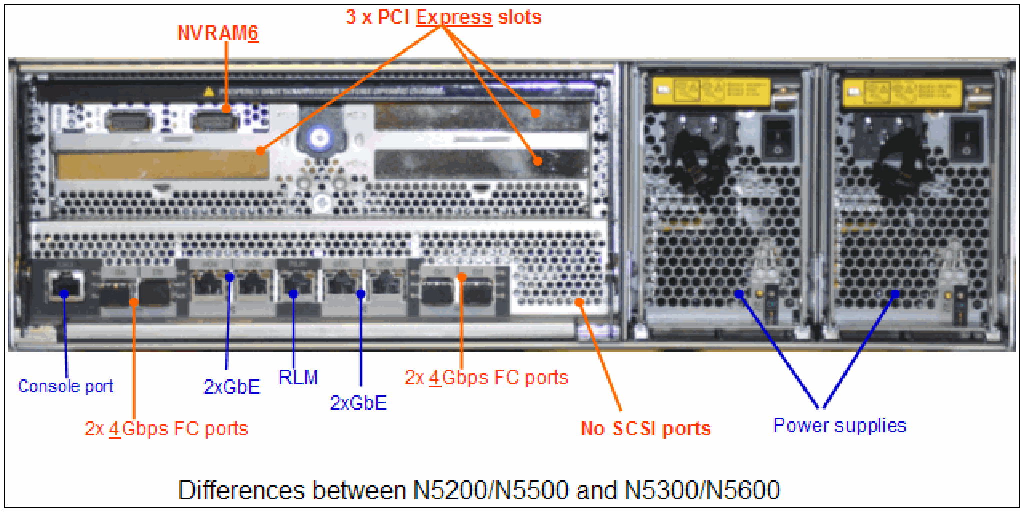


Figure 1-14 Rear view of N5600 and N5300

The N5600 and N5300 also use a BIOS prompt upon boot rather than a Common Firmware Environment (CFE) prompt.

The easily accessible rear of the N5000 series provides I/O connectivity and power supply access and status indications (Figure 1-14).

RAID group sizes

Table 1-9 provides information about RAID group sizes.

Table 1-9 RAID group size in drive type

Model	FC-AL drives default	FC-AL drives maximum	ATA drives default	ATA drives maximum
RAID 4	8	14	7	7
RAID DP	16	28	14	16

1.5.3 IBM System Storage N6000 introduction

IBM N6000 series systems offer a versatile storage platform for handling the large amounts of diverse data moving through large organizations. With an N6000 series system you can consolidate these varied data sets onto a unified storage platform supporting simultaneous block and file services for business and technical applications. The N6000 systems enable you to connect heterogeneous server environments (including Windows, UNIX, and Linux servers) and clients to one storage system using standard storage protocols and interfaces.

The IBM N6000 A series comes in two models:

- N6040
  - 2858-A10 Single Node
  - 2858-A20 Clustered
- N6070
  - 2858-A21 Clustered

Models in the N6000 family, as well as the N7000 family, support 8 GBit/sec FC connectivity on the SAN and 10 GBit/E IP connectivity (as a unified target for NIC and FCoE) on the LAN with proper adapters.

Unified model numbers

Previous models of the N series had separate model designations for the gateway (model G, for example, 2866-G11) and storage system (model A) versions (for example, 2866-A11). Starting with the N6000 there is only one model designation, model AXX (Figure 1-15 and Figure 1-16). Those models that are ordered with the gateway option can be identified with the letter G on the model and serial number label (Figure 1-15) located on the bezel. This helps reduce complexity for customers and sales personnel. In early releases of Data ONTAP 7.3 you might still see the G model designation in the **`sysconfig`** command (Example 1-1) on those N6000 series systems with the gateway option and in Autosupport problem reporting. Our support structure is set up to handle this. For manual problem reporting use the machine type and serial number on the label.

Figure 1-15 N6000 series with gateway option

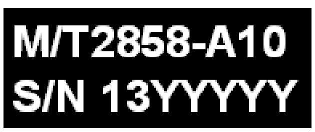


Figure 1-16 N6000 without gateway option

Example 1-1 N6000 with gateway option

```
slot 0: System Board 2.6 GHz (NetApp System Board XV A1)
         Model Name:         N6070
         Machine Type:       IBM-2858-G21
         Part Number:        110-00119
```

Figure 1-17 shows the front of an N6000 series unit, including the LED indicator lights; Figure 1-18 shows a unit with the front cover removed.

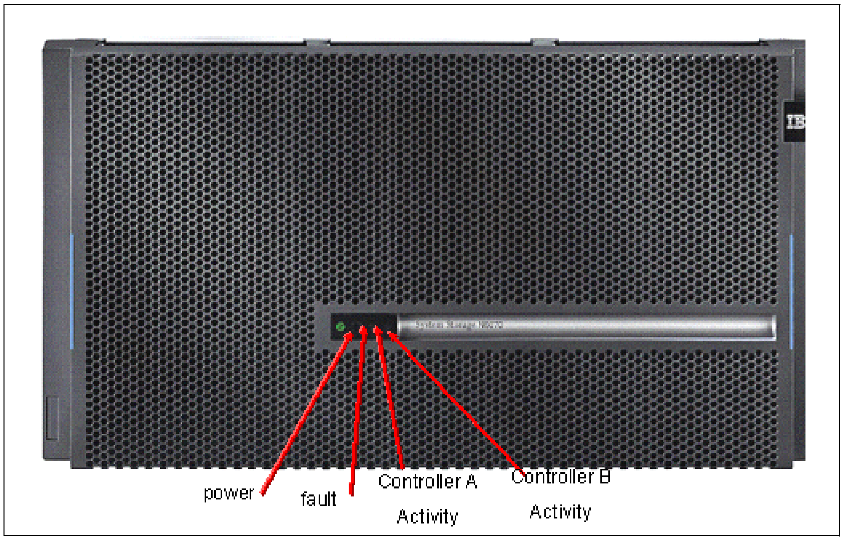


Figure 1-17 Front of an N6000 with LEDs

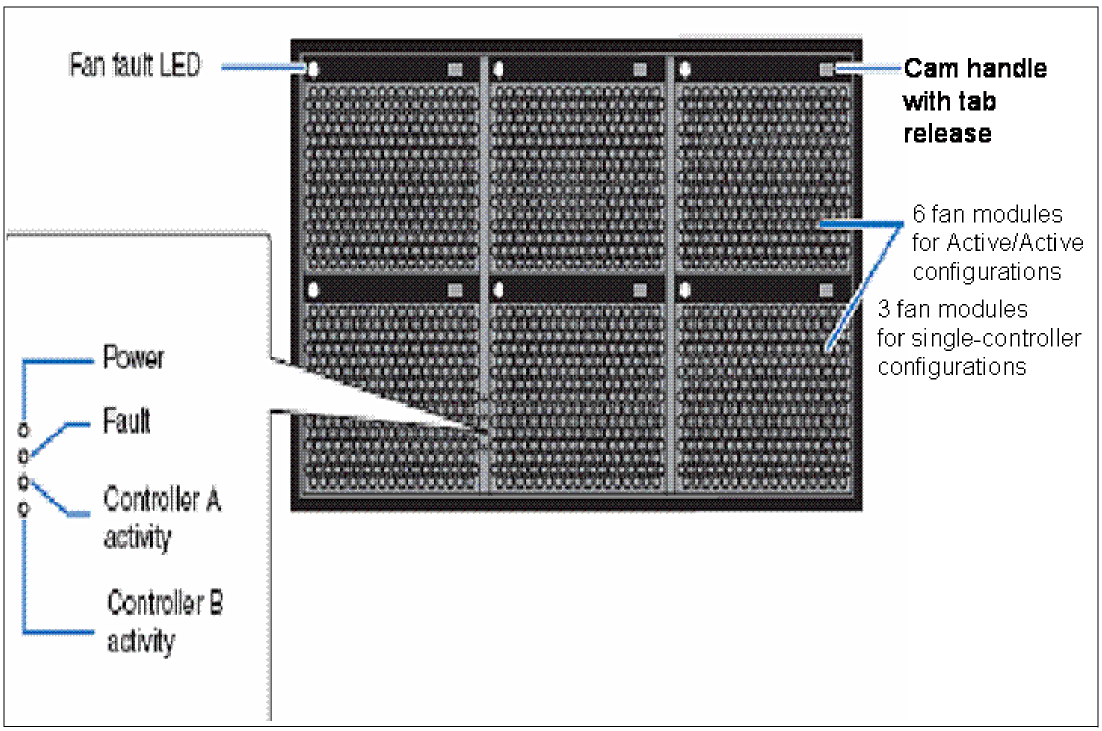


Figure 1-18 Front of an N6000 with front cover removed

Front LED operations

The four light emitting diodes (LEDs) change colors depending on status:

- Power LED
  - On (green) when main (12V) power is on.
  - Off when main power is removed.
    If power cords are installed and the rocker switches to on, standby power is still present.
- Fault
  - On (amber) when either controller indicates a fault.
  - There is no way to tell from the front (by LEDs alone) which controller is in fault. You can use the rear fault LED or console logs and EMS messages to determine where the fault is.
- Activity
  - One for A (top) controller, one for B (bottom) controller.
  - Driven on (green) or off by Data ONTAP.
- Fan field-replaceable units (FRUs)
  - These are only visible with the bezel off.
  - Controlled by Data ONTAP.
  - Turned on (amber) to identify which fan FRU to replace.

Figure 1-19 shows a rear view of a dual-controller configuration of the N6000 series model A20 or A21. The second controller would be absent in a single-controller configuration and there would be a empty bay in the chassis. The unified chassis allows both controllers to fit in a single chassis in an active/active configuration. This unified chassis design only requires 6U and eliminates the need for external communication cables between controllers.

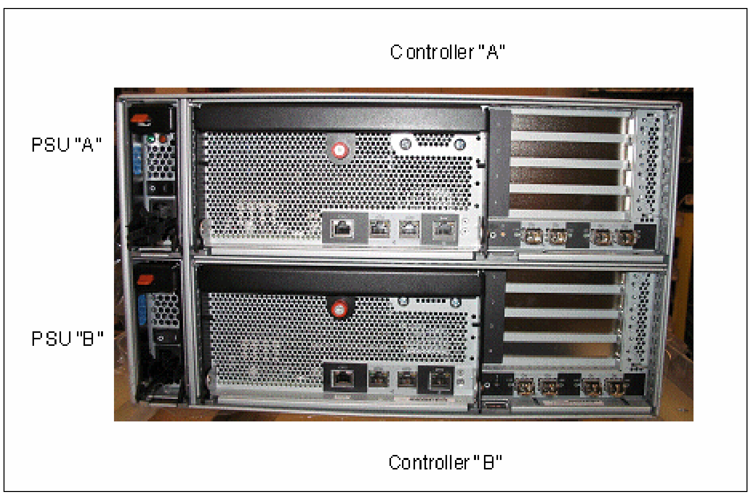


Figure 1-19 N6000 rear view

Figure 1-20 shows the available I/O ports on the base models of the N6000 series. More ports can be added through the PCI slots and appropriate feature cards.

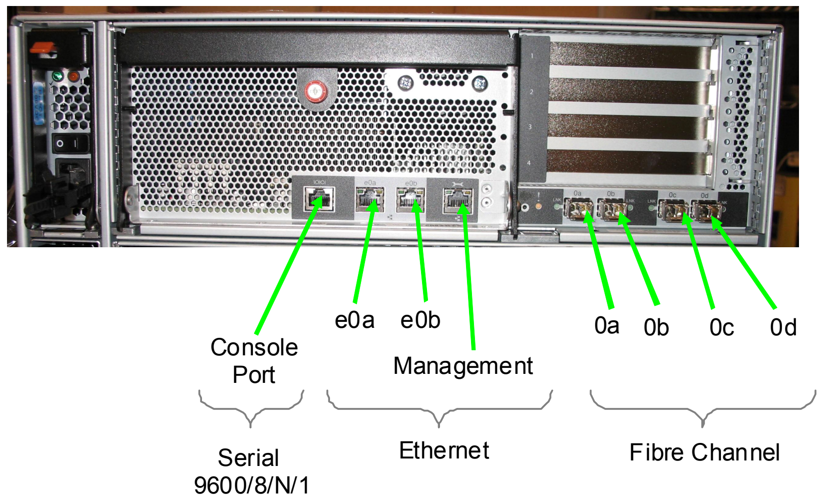


Figure 1-20 Rear view of a N6000 controller and its I/O ports

Controller rear LEDs

There are more LEDs in the rear of the controller that can provide status, too.

Power Supply Unit (PSU)

The PSU has two LEDs – green and amber (Table 1-10). Green and amber LEDs are mutually exclusive. There is no condition in which both are on at the same time.

Table 1-10 LED status

Green LED	Amber LED	Meaning
Off	Off	No external AC power or switch is off.
On	Off	PSU is on, operating normally.
Off	Flashing	PSU has faulted.

Ethernet

Each Ethernet interface has two link status LEDS:

- Amber indicates activity on the Fibre Channel
- Green indicates that the link is active

Controller LED

LEDs for the controllers are located on front and back and have the following display conditions:

- The rear LED mirrors the front fault LED, but is not OR’d with other controller.
- If the front fault LED is on, looking at the fault LEDs on the rear will tell you which controller is in fault. This is on by default after the power cycle. Data ONTAP turns it off only after it boots and determines that all is well.

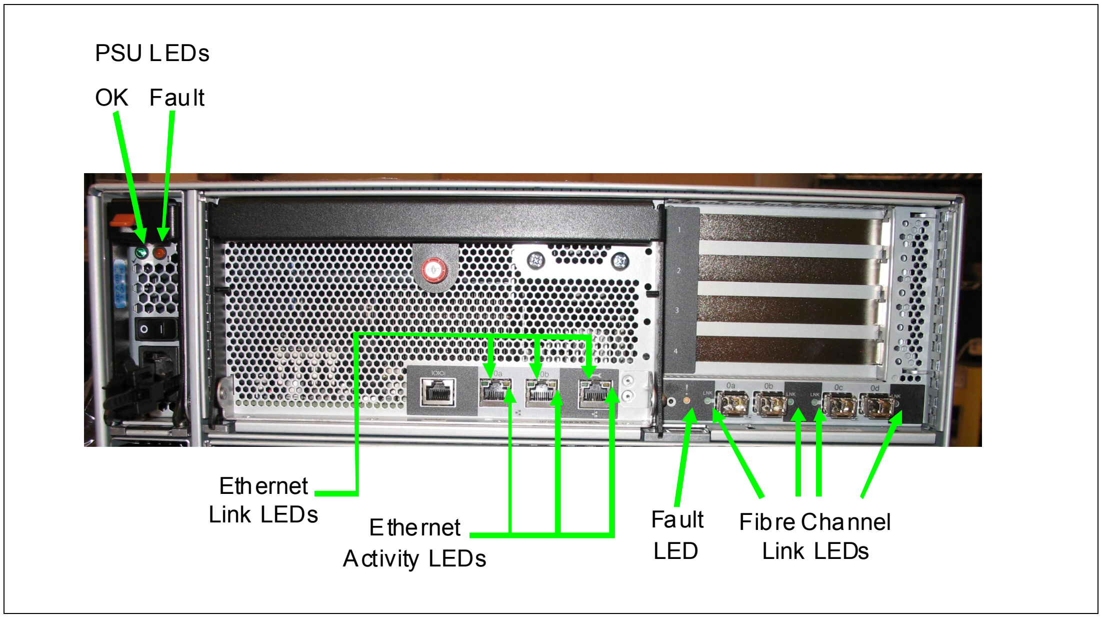


Figure 1-21 Rear of unit LEDs

System port 80 LEDs

The system port 80 LEDs indicate the status of the Data ONTAP operating system. As shown in Figure 1-22 there are eight green LEDs. These lights come on sequentially to indicate BIOS progress while the system is booting, then are off when Data ONTAP is running. They flash in a sweeping pattern if the system sitting at loader prompt.

There is one red NVRAM7 LED, which blinks when the NVRAM7 is holding data.

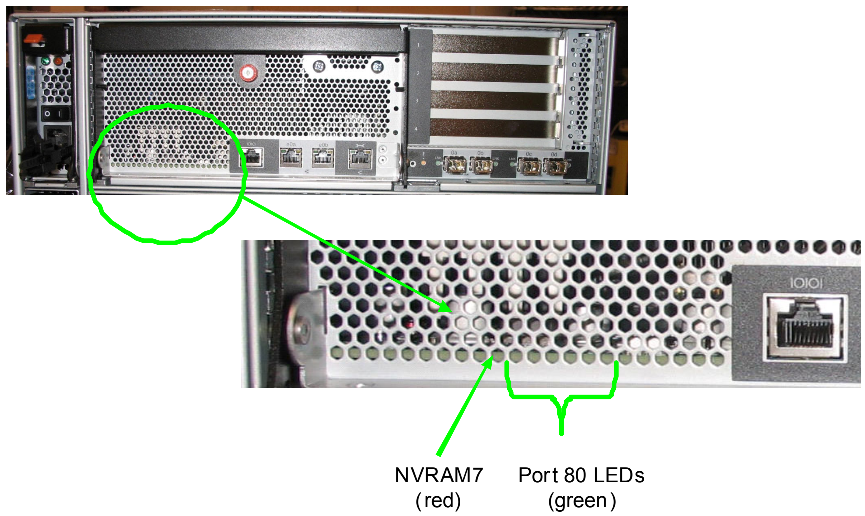


Figure 1-22 NVRAM and port 80 LEDs

Internal LEDs

There are several internal FRU LEDs (Figure 1-23 on page 26) backed by a *super-cap*, which remains lit for 5 to 10 minutes after power is removed (that is, to remove the controller from the chassis).

- Red NVRAM LED

  This blinks when NVRAM is holding data.

- As identified in Figure 1-23, the main controller board has:
  - Eight DIMM LEDs
  - NVRAM DIMM LED
  - RTC, remote LAN module (RLM), and compact flash LEDs

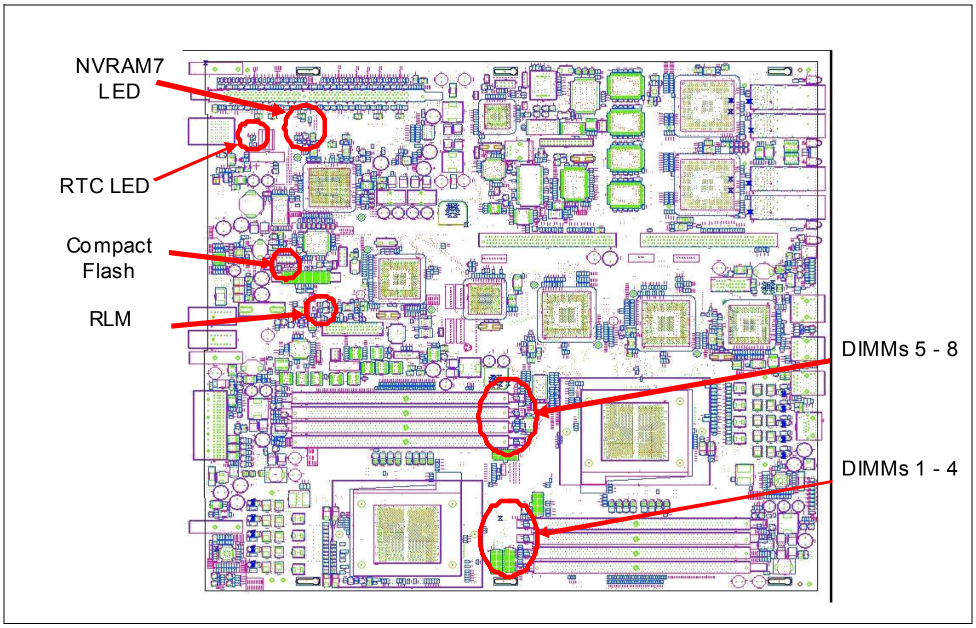


Figure 1-23 Internal LEDs

The riser has four PCI slot LEDs (Figure 1-24).

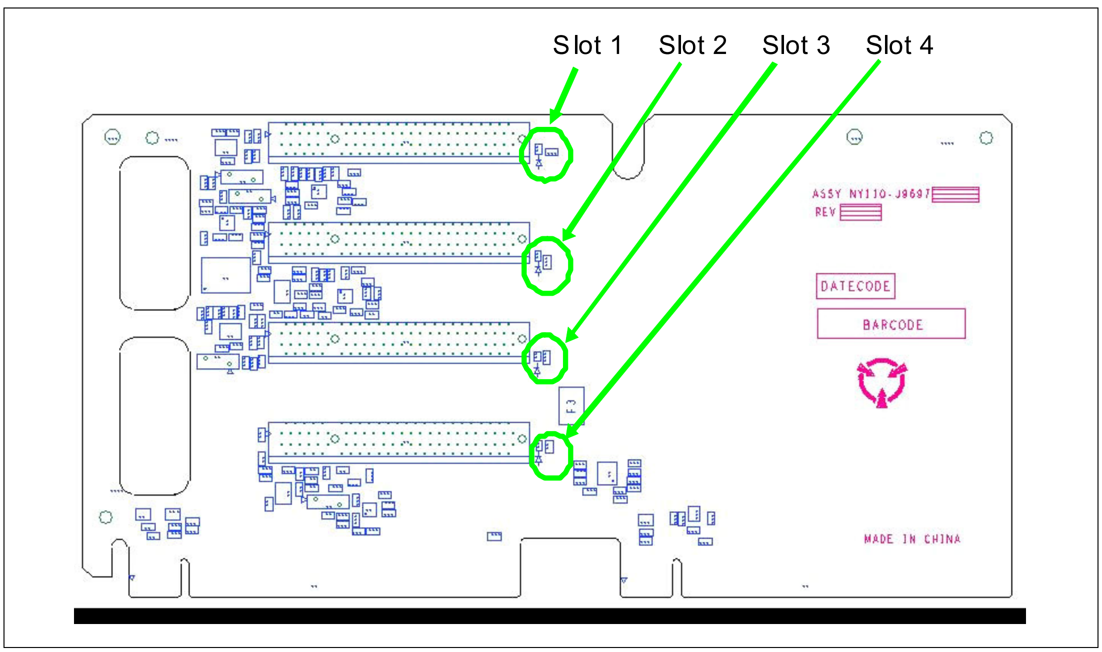


Figure 1-24 Riser board

Controller board components

Figure 1-25 is a view of the controller board for one node. In an active/active or A2X configuration there would be two controller boards for each node. The N6000 series uses NVRAM7. The NVRAM is embedded in the controller board.

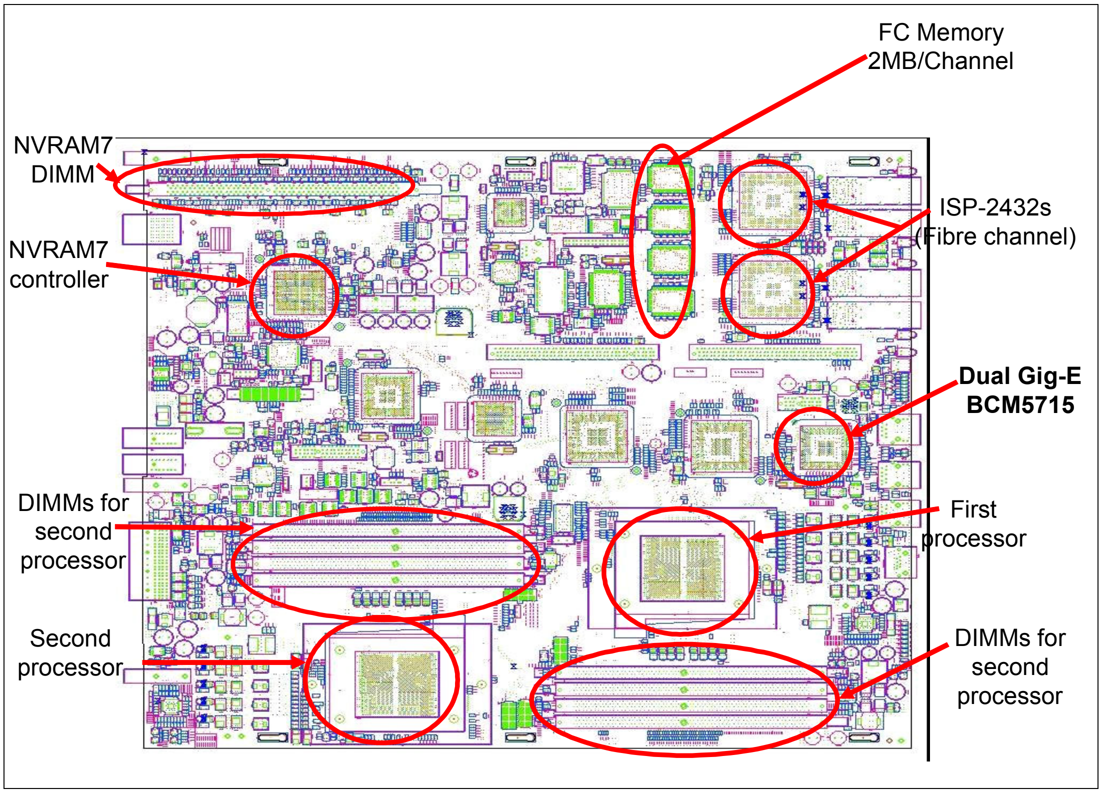


Figure 1-25 Controller board

Figure 1-26 shows a riser board, of which there is one per node. The riser board is where the PCI-E cards are installed.

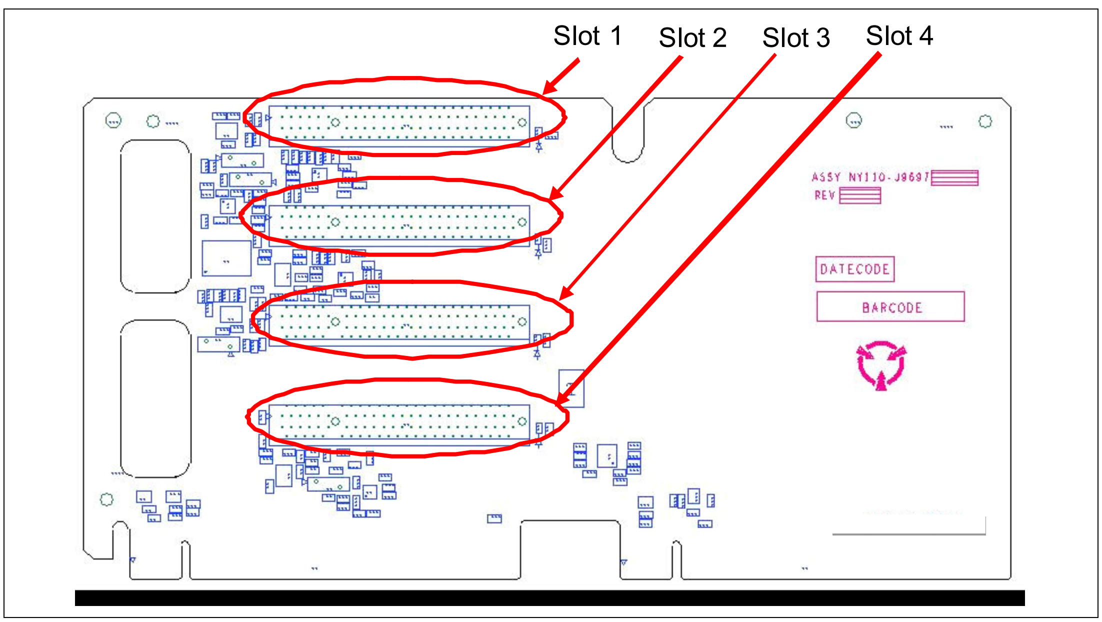


Figure 1-26 Riser board

Figure 1-27 presents a schematic view of the cooling configuration for the N6000 series.

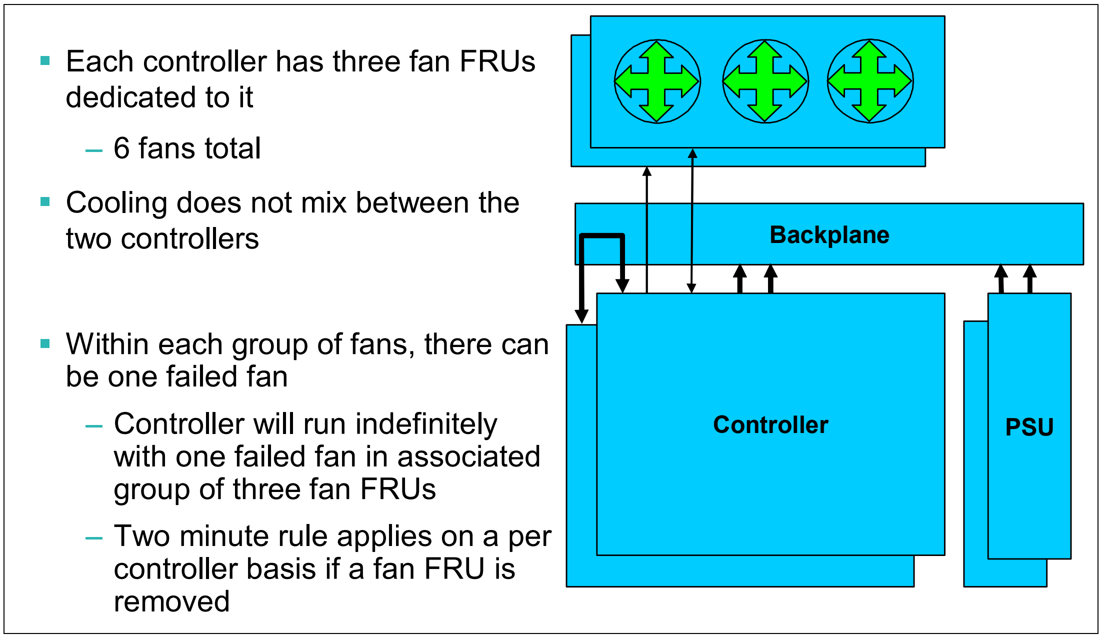


Figure 1-27 Cooling configuration

Dynamic removal and insertion of the controller

The N6000 controllers are hot pluggable. You do not have to turn off PSUs to remove a controller in a dual-controller configuration.

PSUs are independent components. One PSU can run an entire system indefinitely. There is no “two-minute rule” if you remove one PSU. PSUs have internal fans for self-cooling only.

New RLM design and internal Ethernet switch on the controller

The Data ONTAP management interface, known as e0M, provides a robust and cost-effective way to segregate management subnets from data subnets without incurring a port penalty. On the N6000 series, the traditional RLM port on the rear of the chassis (now identified by a wrench symbol) connects first to an internal Ethernet switch that provides connectivity to the RLM and e0M interfaces. Because the RLM and e0M each have unique TCP/IP addresses, the switch can discretely route traffic to either interface. You do not need to use a data port to connect to an external Ethernet switch. Set up of VLANs and VIFs is not required and not supported because e0M allows customers to have dedicated management networks without VLANs.

e0M highlights

Key features and capabilities of the e0M are:

- 100baseT link speed
- Auto-negotiate only
- Supports IP failover
- Shares the same external connector as the RLM (See Figure 1-28 on page 30.)
- Is a fully integrated Ethernet Media Access Control and Physical Layer Transceiver solution
- Use if switch is in advanced mode to read link status of external port (Example 1-2 on page 29)

Example 1-2 Switch status

```
itsotuc1*> ifswitch stat

--interface sw-RJ45--

Rx Good Frames  54661
Tx Frames       37887
Rx Discards     0
link            up
media           auto-100tx-fd

--interface sw-RLM--

Rx Good Frames  37887
Tx Frames       20788
Rx Discards     0
link            up
media           auto-100tx-fd

--interface sw-PartnerSwitch--

Rx Good Frames  8570
Tx Frames       8407
Rx Discards     0
link            up
media           auto-100tx-fd

--interface sw-e0M--

Rx Good Frames  0
Tx Frames       54242
Rx Discards     0
link            up
media           auto-100tx-fd
itsotuc1*>
```

The e0M interface can be thought of as another way to remotely access and manage the storage controller, much like the serial console, RLM, or standard network interface. The e0M interface should be used for network-based storage controller administration, monitoring activities, and ASUP reporting. The RLM is used when you require its higher level of support features. Host-side application data should connect to the appliance on a separate subnet from the management interfaces.

On the N6000 series there is an Ethernet switch within the controller. The management port (Figure 1-28 on page 30) provides access to both the RLM and the e0M port. Management traffic is all 10/100. The link light on the management port indicates the link to switch, not to the RLM or e0M. The switch configures itself to allow traffic to flow in appropriate directions at power-up.

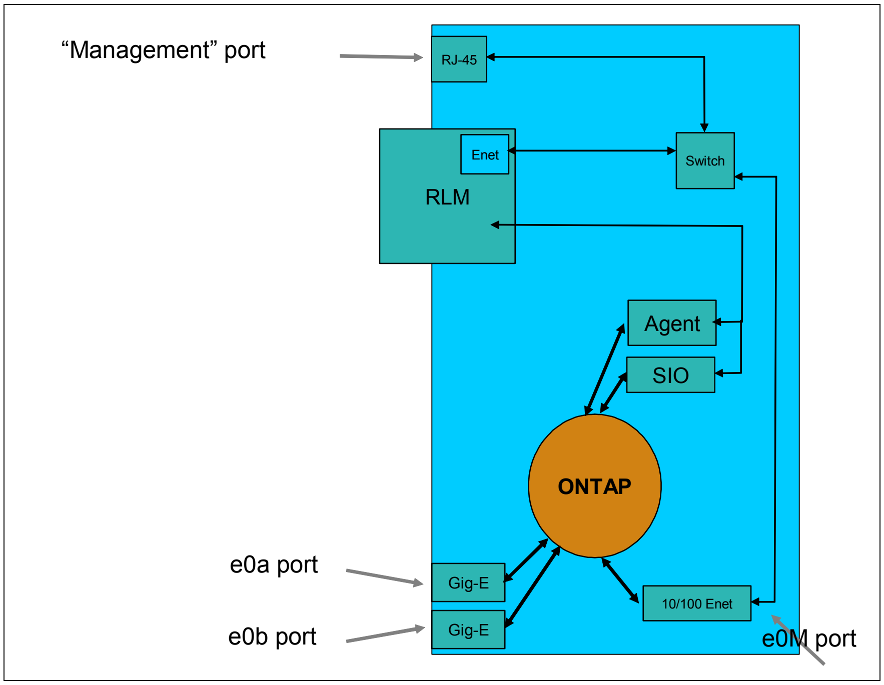


Figure 1-28 Ethernet internals

The possibility of loop storms

If the management ports of both controllers are connected to the same subnet you could experience loop storms (Figure 1-29).

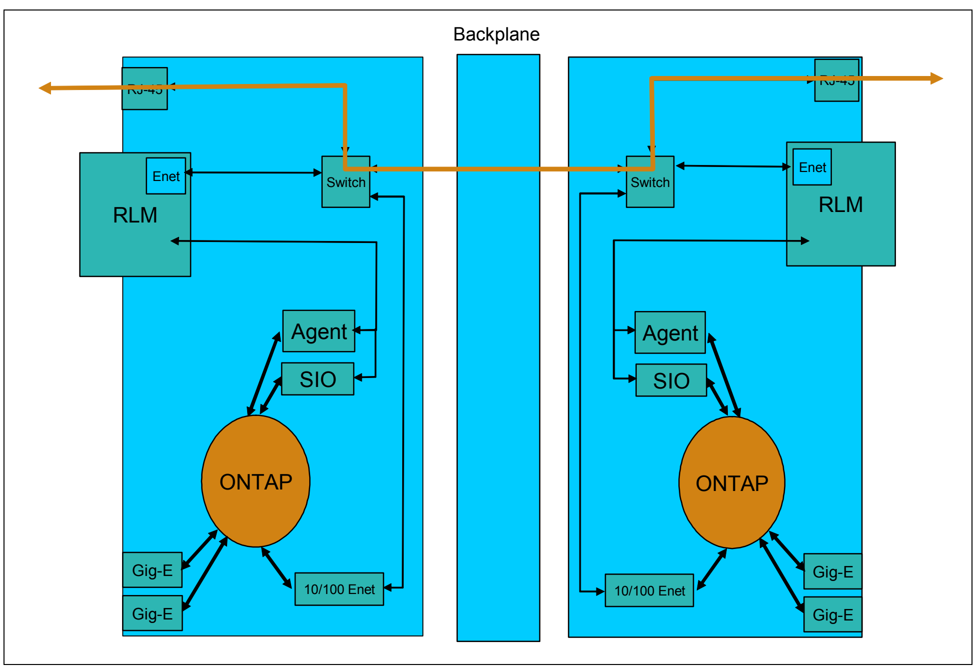


Figure 1-29 Loop storm configuration

Preventing the storm

A loop storm can be prevented by implementing *port blocking* in the Ethernet switch (Figure 1-30). The switch reads the configuration information at power-on. The switch is configured to disallow traffic to pass from the management port through to the backplane. All other paths through the switch are open for traffic.

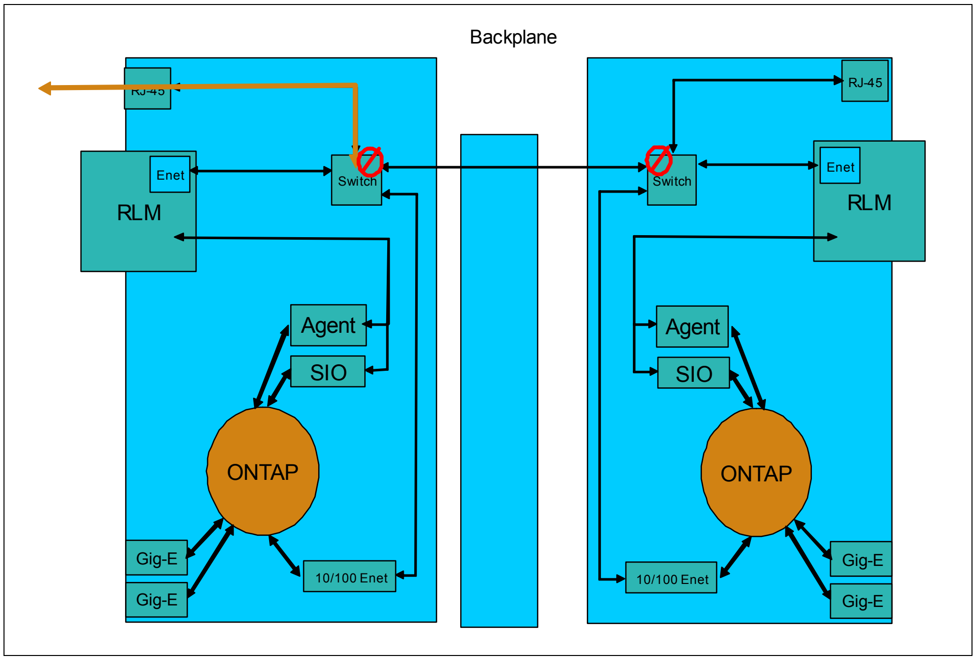


Figure 1-30 No loop storm configuration

RLM-assisted cluster failover (CFO)

In order to decrease the time required for cluster failover to occur when there is an event that the RLM is aware of, the RLM can communicate with the partner node instance of Data ONTAP. This capability was available in other N series models prior to the N6000 series, but the internal Ethernet switch makes the configuration much easier and facilitates quicker cluster failover, with some failovers occurring within 15 seconds.

Reliability improvements

The N6000 series improves reliability as compared with its predecessors. Some of the highlights are:

- It uses fewer cables because no external IB cables for clustering are required.
- Fewer connectors are needed: embedded NVRAM eliminates the PCIe connector.
- There are fewer components, specifically, two fewer power supplies.
- Improved component de-rating guidelines result in less stress on components.

Availability improvements

The N6000 has added the following new features that improve availability:

- Better cooling:
  - Three fan FRUs per controller versus two in N5000 systems.
  - Overall airflow much better with more thermal margin.
- One fewer two-minute rule to worry about; a PSU can be removed indefinitely.

Serviceability improvements

The N6000 series also includes serviceability improvements to make it easier and faster to service, including:

- Controller and NVRAM7 are combined in a single FRU
  - Fewer FRUs to diagnose
  - Fewer FRUs to stock
- Improved RLM mechanical guides
  - Easier to install
  - Easier to remove

> **Note:** The N6000 series only runs with Data ONTAP 7.2.5 or 7.3 and later.

NVRAM 7

One of the changes with the N6000 and its NVRAM 7 is that the serial number matches that of the system board (Example 1-3).

Example 1-3 Matching serial numbers

```
slot 0: NVRAM (NetApp NVRAM VII)
   Serial Number:      700761
slot 0: System Board 2.6 GHz (NetApp System Board XV B7
               Model Name:         FAS3170
               Part Number:        110-J9693
               Revision:           B7
               Serial Number:      700761
```

1.5.4 IBM System Storage N7000 introduction

The IBM System Storage N7000 series offers additional choices to organizations facing the challenges of enterprise data management. The IBM System Storage N7000 series is designed to deliver high end enterprise storage and data management value with midrange affordability. Built-in enterprise serviceability and manageability features help support your efforts to increase reliability, simplify and unify storage infrastructure and maintenance, and deliver exceptional economy.

The IBM N7000 A series comes in two models:

- N7700
  - 2866-A11
  - 2866-A21 Clustered
- N7900
  - 2867-A11
  - 2867-A21 Clustered

N7000 units support FC, SATA, or both. Both can be used behind a single controller, but not in the same drawer.

N7700 models are no longer orderable.

Like its N5000 predecessor, the front of the N7000 series unit has an LCD display with three standard LEDs indicating system activity, status, and power (Figure 1-31).

Figure 1-31 Front view of the N7000

From the rear of the N7000 you can see the redundant power supplies, the NVRAM card, the gigabit Ethernet interfaces, and the Fibre Channel interfaces. The console port and RLM port are also located on the rear (Figure 1-32 on page 33).

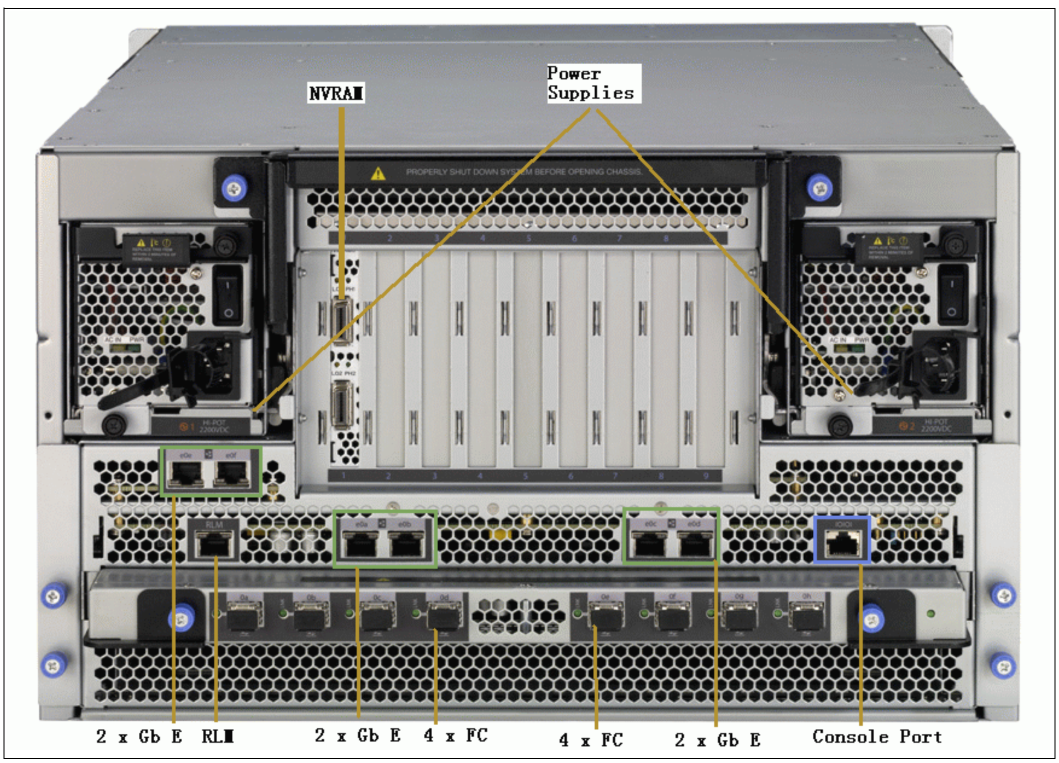


Figure 1-32 Rear view of the N7000

Each N7000 node requires 6U of rack space. Each EXN1000/4000 expansion unit requires 3U of rack space. Each N7000 node requires at least one expansion unit (Figure 1-33).

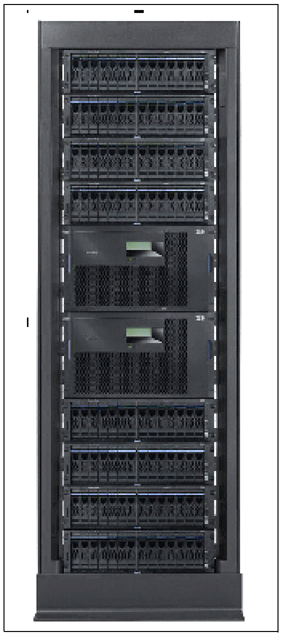


Figure 1-33 N7000 racked

A dual-node N7900 supports a maximum of 84 storage expansion units. Each rack holds a maximum of 12 expansion units (Figure 1-34). The N7000 products are installed by IBM Service or by a qualified IBM Business Partner. These systems are not appropriate for customer setup.

Figure 1-34 Clustered N7000 with multiple expansion units

Removing the bezel reveals the CompactFlash card reader, and directly below it, the remote LAN module. The RLM is required in all N 7000 series systems. The systems cannot boot unless the card is present. Also, you will see five fan units. The fans are hot swappable and are numbered in Figure 1-35 for your reference.

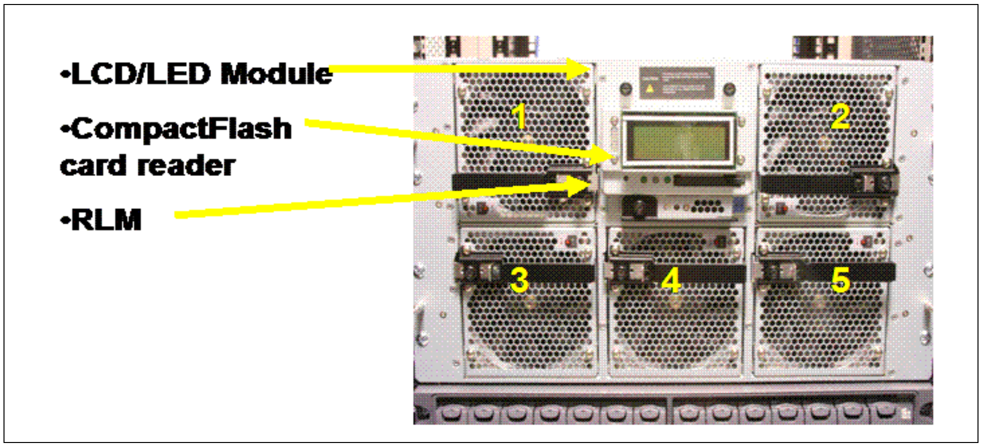


Figure 1-35 Front of N7000 with bezel removed

There are two handles on each side of the system to help when lifting the system (Figure 1-36). The system is very heavy. Fully loaded, it weighs 120 pounds. We recommend that before lifting the system you remove the fan units and the two power supplies. This reduces the weight to slightly over 90 pounds. We recommend that the lifting be done by three people.

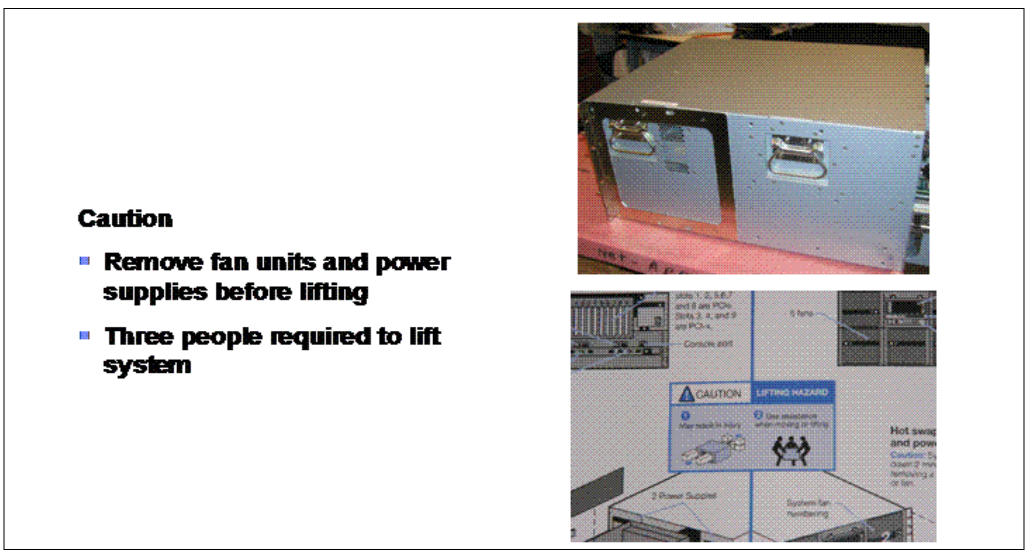


Figure 1-36 Lifting N7000

Figure 1-37 shows the rear view of the N7000 with the following areas identified:

- Two hot swappable power supplies.
- Nine PCI slots, numbered 1 through 9, from left to right.
- Ethernet ports, and specifically, a console port and an RLM port.
- The Fibre Channel tray, referred to as the FC tray, with eight onboard Fibre Channel ports. This tray is a field-replaceable unit.

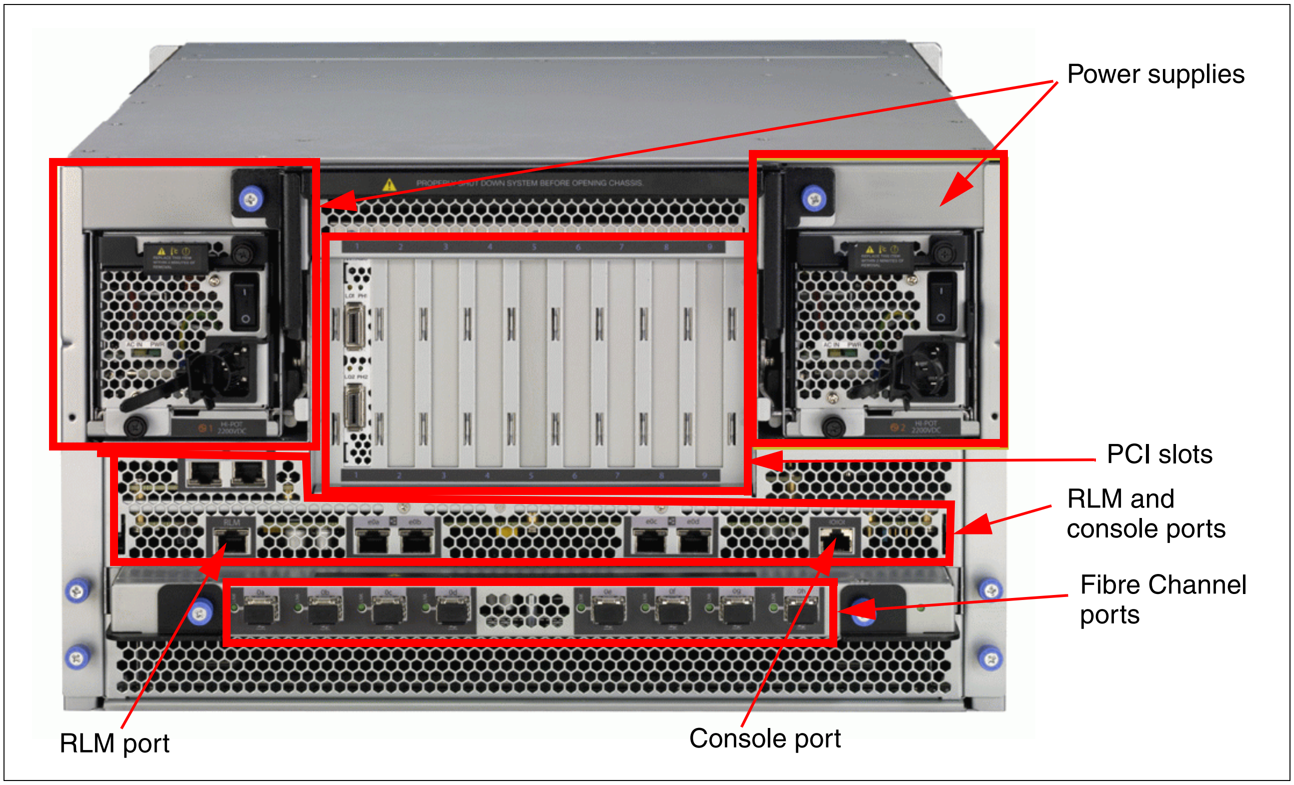


Figure 1-37 Rear view of an N7000

Inside the N7000, looking from the top (Figure 1-38), the PCI slots and system memory are visible, but not the processors. They are on the other side of the motherboard tray. Recall that the N7900 has eight CPUs.

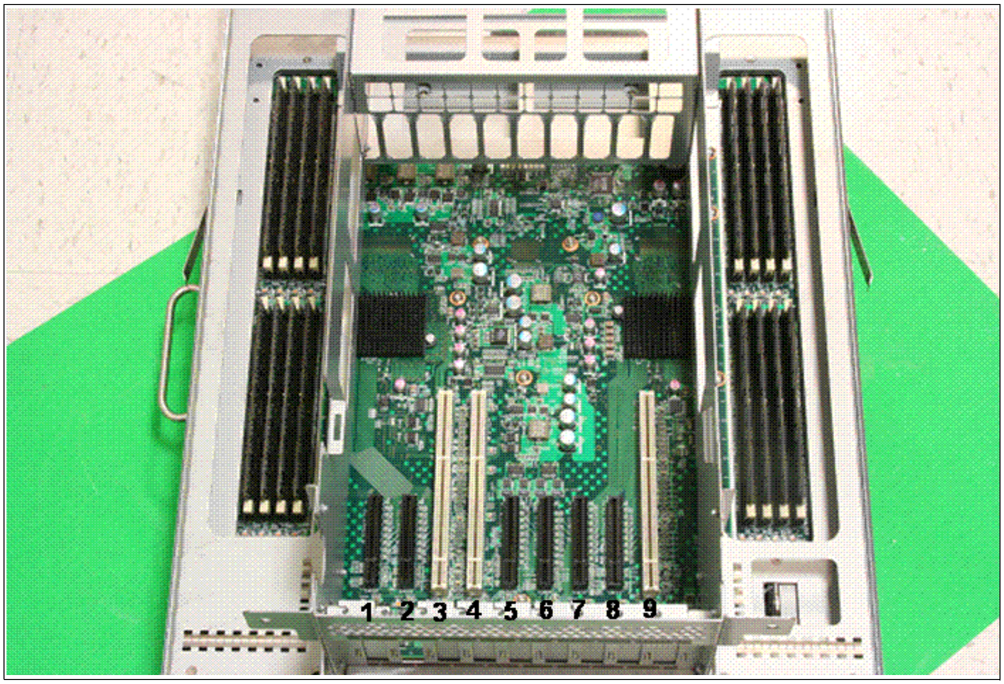


Figure 1-38 Top of N7000

From this perspective, you can see the nine PCI slots. Slots 3, 4, and 9 are white and represent PCI-X. Slots 1, 2, 5, 6, 7, and 8 are PCI-Express.

Notice that the NVRAM6 adapter resides in slot 2 on this standalone system. If this were an active/active configuration, the NVRAM6 adapter would reside in slot 1 and would also be used as the cluster interconnect card.

Keep in mind that the N7800 and N7900 use an NVRAM6 adapter with 4 GB of memory, and the N7600 and N7700 use an NVRAM6 adapter with 1 GB of memory (Figure 1-39).

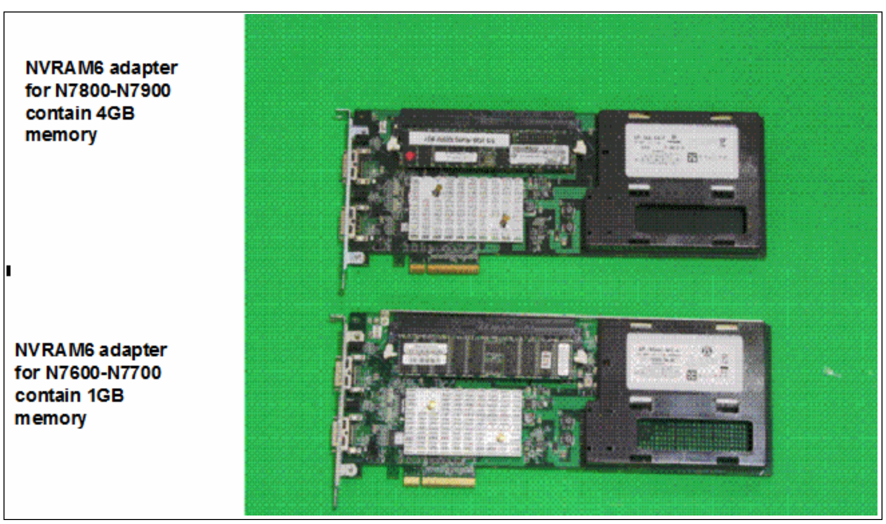


Figure 1-39 NVRAM

The fan units, PCI slots, and memory DIMMs have LEDs to indicate a failed component (Figure 1-40).

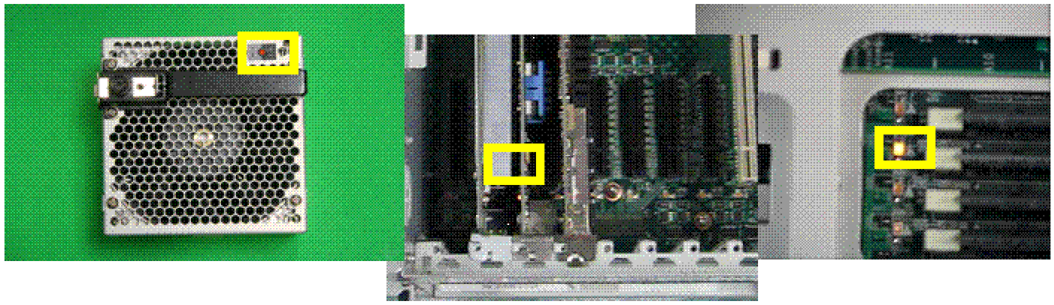


Figure 1-40 New LEDs (left to right): Fan unit, PCI adapter, DIMM

RAID group size

Table 1-11 lists RAID group sizes by drive type.

Table 1-11 RAID group size in drive type

Model	FC-AL drives default	FC-AL drives maximum	ATA drives default	ATA drives maximum
N7700 RAID 4	8	14	7	7
N7700 RAID DP	16	28	14	16
N7900 RAID 4	8	14	7	7
N7900 RAID DP	16	28	14	16

1.6 IBM System Storage N series Gateways

The IBM System Storage N series Gateway, an evolution of the N5000 series product line, is a network-based virtualization solution that virtualizes tiered, heterogeneous storage arrays, allowing customers to leverage the dynamic virtualization capabilities available in Data ONTAP across multiple tiers of IBM and vendor-acquired storage (Figure 1-41). Like all IBM System Storage N series storage systems, the IBM System Storage N series Gateway family is based on the industry-hardened Data ONTAP microkernel operating system, which unifies block and file storage networking paradigms under a common architecture and provides a complete suite of IBM System Storage N series advanced data management capabilities for consolidating, protecting, and recovering mission-critical data for enterprise applications and users.

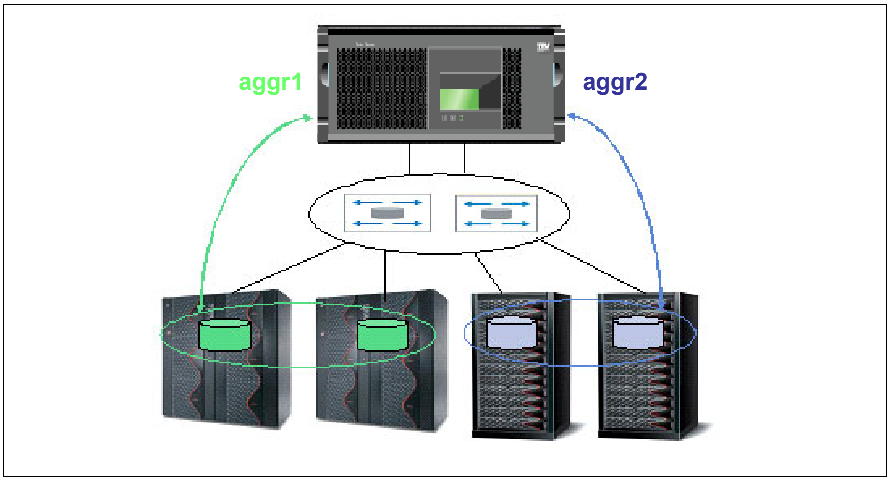


Figure 1-41 Heterogeneous storage

The industry's most comprehensive virtualization solution, the N series Gateway provides proven and innovative data management capabilities for sharing, consolidating, protecting, and recovering mission-critical data for enterprise applications and users and seamlessly integrates into mission-critical enterprise-class SAN infrastructures. These data management capabilities, when deployed with disparate storage systems, simplify heterogeneous storage management.

The IBM System Storage N series Gateway presents shares, exports, or LUNs that are built on flexible volumes that reside on aggregates. The N series Gateway is also a host on the storage array SAN. N series Gateways can take storage array LUNs (which are treated as disks) and virtualize them through Data ONTAP, presenting a unified management interface.

This simple, elegant data management solution can decrease management complexity, improve asset utilization, streamline operations to increase business agility and reduce total cost of ownership enhance data protection, enable rapid recovery, and broaden centralized storage usage by provisioning SAN capacity for business solutions requiring NAS, SAN, or IP SAN data access (Figure 1-42).

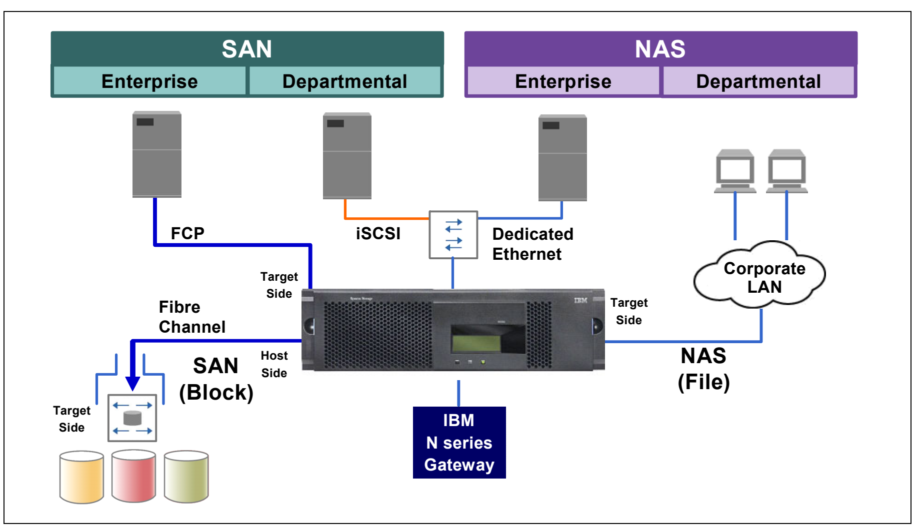


Figure 1-42 Gateway topology

With Data ONTAP, the N series Gateway now supports attachment of heterogeneous storage systems as well as IBM expansion units of the type used with N series storage systems (Figure 1-43).

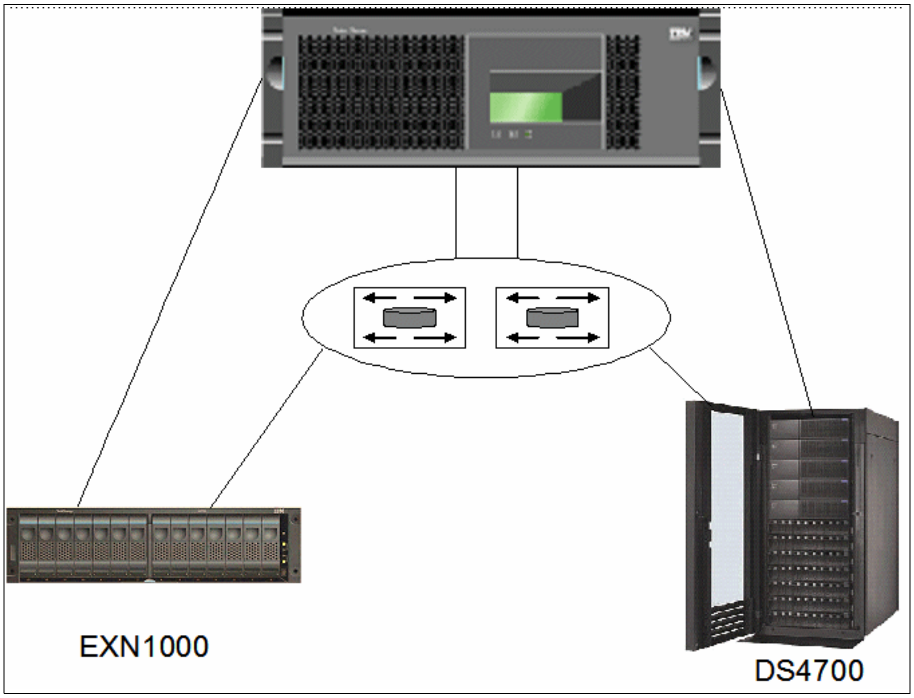


Figure 1-43 Storage systems with expansion units

Two halves to set up

An N series Gateway implementation can be thought of as a front-end implementation and a back-end implementation. Front-end set up includes configuring the N series Gateway for all protocols (NAS or FCP), implementing any snap features (Snapshot, SnapMirror, SnapVault, and so on), and setting up backup including NDMP dumps to tapes. The back-end implementation includes all tasks required to set up the N series Gateway system up to the point where it is ready for Data ONTAP installation. These tasks include array LUN formatting, port assignment, cabling, switch zoning, assigning LUNs to the N series Gateway system, creating aggregates, and loading Data ONTAP.

1.6.1 IBM System Storage N series Gateway highlights

IBM System Storage N series Gateway provides a number of key features that enhance the value and reduce the management costs of utilizing a storage area network. An N series Gateway:

- Simplifies storage provisioning and management
- Lowers storage management and operating costs
- Increases storage utilization
- Provides comprehensive, simple-to-use data protection solutions

- Improves business practices and operational efficiency
- Transforms conventional storage systems into a better managed storage pool (Figure 1-44)

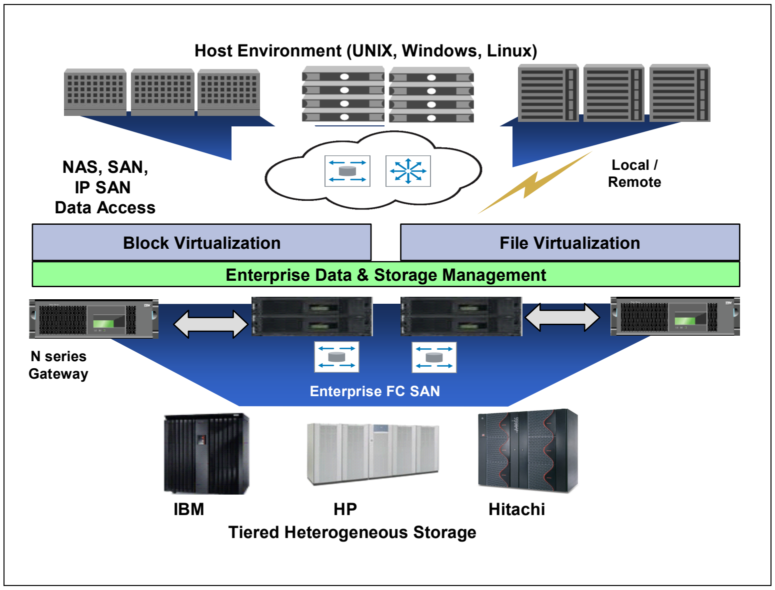


Figure 1-44 Tiered heterogeneous storage

1.6.2 Gateway RAID

Gateways use RAID 0 and rely on the RAID protection provided by the underlying storage subsystems (Figure 1-45). Physical disk operations such as scrubbing are disabled.

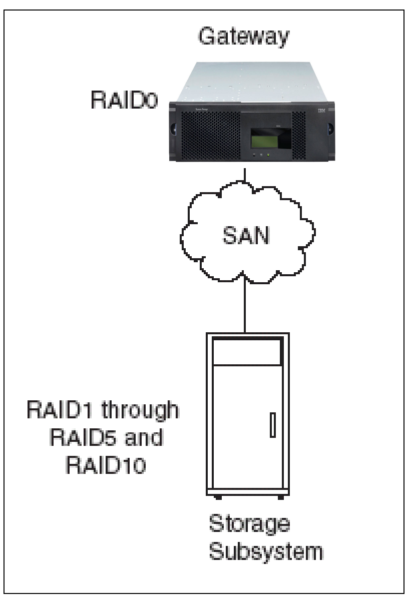


Figure 1-45 RAID configuration

When external storage is attached to the Gateway the volume status shows `raid0`, as seen in Example 1-4. Storage from attached shelves appear as `RAID 4` or `raid_dp`.

Example 1-4 vol status command for Gateway volumes

```
itsotuc2*> vol status -v vol3
         Volume State      Status            Options
           vol3 online     raid0, flex     nosnap=on, nosnapdir=off,
                                             minra=off,
                                             no_atime_update=off,
                                             nvfail=off,
                                             snapmirrored=off,
                                             create_ucode=on,
                                             convert_ucode=on,
                                             maxdirsize=31457,
                                             fs_size_fixed=off,
                                             guarantee=volume,
                                             svo_enable=off,
                                             svo_checksum=off,
                                             svo_allow_rman=off,
                                             svo_reject_errors=off,
                                             fractional_reserve=100,
                Containing aggregate: 'aggr0'

                Plex /aggr0/plex0: online, normal, active
                    RAID group /aggr0/plex0/rg0: normal
```

1.6.3 IBM System Storage N6040, N6060, and N6070 Gateway models

The N6000 Appliance models when provided with *feature code 9551* are Gateway models. For more details check:

http://www-03.ibm.com/systems/storage/network/

This means that the system can be attached to both EXNx000 shelves and other disk systems (if qualified). To determine whether your disk system (IBM or not) is certified check the compatibility matrix in:

ftp://public.dhe.ibm.com/storage/nas/nseries/nseries_gateway_interoperability.pdf

The Gateway models of the N6000 family are:

- N6040
  - 2858-A10 Single Node
  - 2858-A20 Clustered
- N6060
  - 2858-A22 Clustered
- N6070
  - 2858-A21 Clustered

1.6.4 IBM System Storage N7700 and 7900 Gateway models

IBM System Storage N7000 series Gateway models offer additional choices to organizations facing the challenges of enterprise data management. The IBM System Storage N7000 series is designed to deliver high-end enterprise storage and data management value with midrange affordability. Built-in enterprise serviceability and manageability features help support your efforts to increase reliability, simplify and unify storage infrastructure and maintenance, and deliver exceptional economy. The IBM System Storage N series Gateway models N7700 and N7900 deliver all the features that the N5000/N6000 series does, but with increased processing, memory, NVRAM, and total storage capacity. The N7000 series Gateway hardware is identical to the A1x and A2x models, the difference being in the enabled features and disk attachment by Data ONTAP.

The IBM System Storage N7000 series in Gateway mode comes in two models:

- N7700
  – 2866-A21 Clustered
- N7900
  – 2867-A21 Clustered

On either model feature code 9551 must be enabled to support external disk system storage.

1.6.5 LUN sizing

Gateway support for LUN sizes is:

- Maximum LUN size: 1024 GB
- Minimum LUN size: 100 MB

> **Note:** The Data ONTAP definition of GB is one GB is equal to 1000 x 1024 x 1024 bytes.
>
> Therefore, the maximum LUN size that Data ONTAP supports is 1024* 1000 * 1024 * 1024 = 1,048,576,000,000 bytes.

1.6.6 LUN mapping

Storage subsystem LUNs are converted to disks for the IBM System Storage N series Gateway (Figure 1-46). When external storage is used the logical disk count equals the physical LUN count associated with the N series. Figure 1-46 shows an example of an array LUN mapped to a gateway disk.

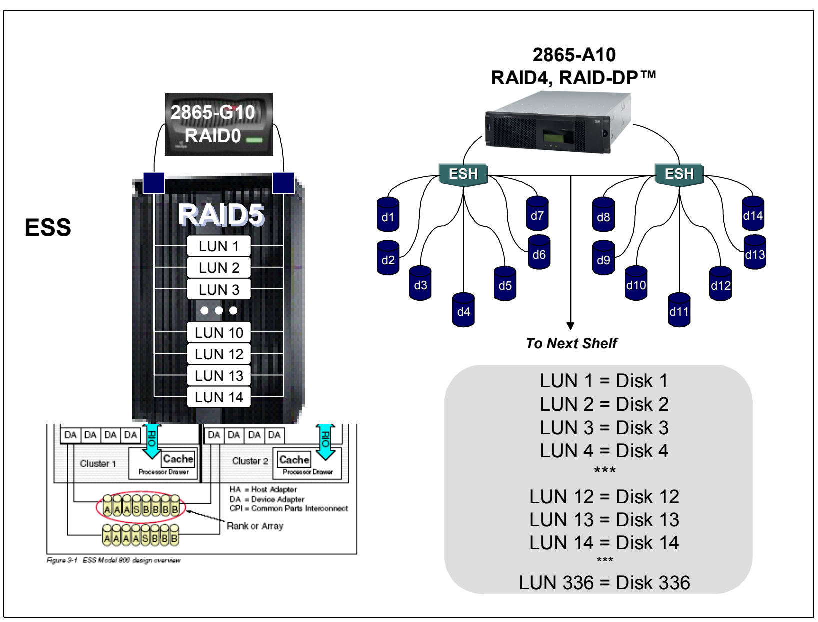


Figure 1-46 LUN to N series Gateway disk relationship

LUNs are added to the Gateway through the same volume wizard that we use on the N series A models (Figure 1-47).

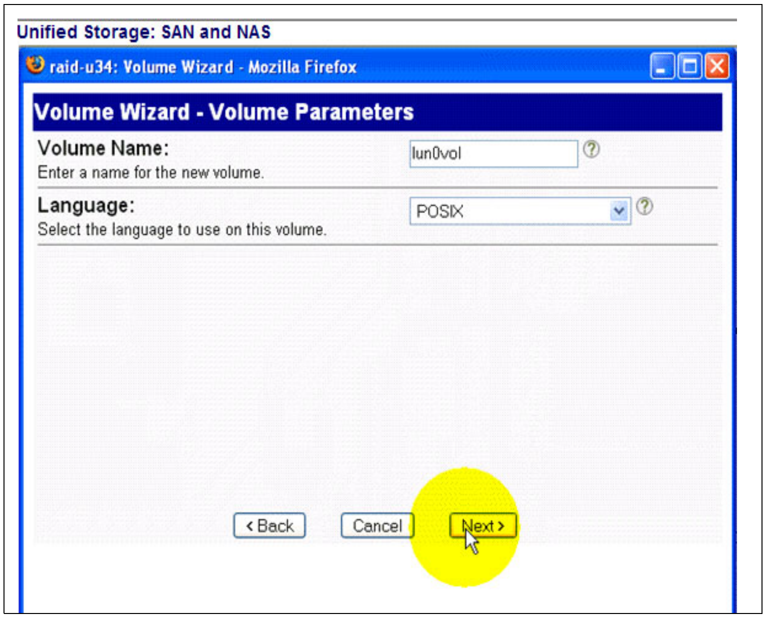


Figure 1-47 Volume wizard

Note: Do not map LUN 0 to gateway systems, even if LUN 0 is a storage LUN.

1.7 N series expansion units

There are currently three disk storage expansion units specifically designed for the IBM System Storage N series storage systems:

- IBM EXN4000 Fibre Channel disk storage expansion unit with 4 Gbps interfaces
- IBM EXN1000 Serial Advanced Technology Attachment storage expansion unit
- IBM EXN3000 Serial Attached SCSI expansion unit

Note: EXN expansion units can be used for attachment to a gateway with Data ONTAP 7.3 and later.

Multiple EXN1000s, each having different SATA disk drive feature codes, can be attached to the same N series storage system on the same Fibre Channel loop. Multiple EXN4000s, each having different Fibre Channel disk drive feature codes, can be attached to the same N series storage system on the same Fibre Channel loop. Multiple EXN3000s, each having SAS or SATA disk drives, can be attached to the same N series storage system on the same SAS loop. For the latest storage expansion unit support information, visit the following website:

http://www.ibm.com/storage/support/nas/

1.7.1 Intermixing EXN units with N series A models

EXN4000s are Fibre Channel disk storage expansion units. Intermixing Fibre Channel and SATA disk drives in a supported N series storage system configuration is supported as follows:

- Intermixing Fibre Channel disk expansion units with SATA disk expansion units on the same loop is not supported.

- EXN4000s (Fibre Channel disk drives) and EXN1000s (SATA disk drives) can be attached to the same N series storage system only if the Fibre Channel disk expansion units (EXN4000s) are on separate loops from the SATA disk expansion units (EXN1000s).
- Intermixing SAS disk expansion units with SATA disk expansion units on the same loop is not supported.

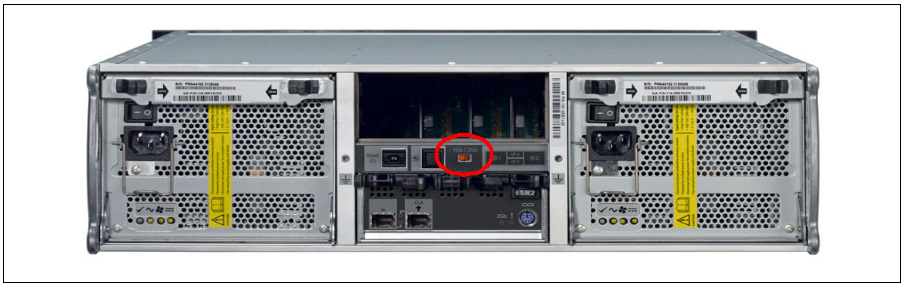


Figure 1-48 Speed switches

1.7.2 EXN1000

The EXN1000 uses the same shelf and hardware as the EXN2000 and EXN4000, so it has the same dimensions. As shown in Figure 1-49, it also supports the same number of disks per shelf (14). The main differences are:

- Drive type supported: SATA versus Fibre Channel
- Interface module: AT-FCX versus ESH2

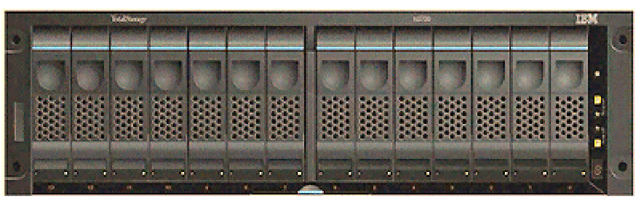


Figure 1-49 EXN1000 expansion unit

AT-FCX refers to the controller module (Figure 1-50) of the Serial Advanced Technology Attachment storage expansion unit.

Figure 1-50 AT-FCX module

Data ONTAP supports up to 400 RAID groups per storage system or cluster. When configuring your aggregates, keep in mind that each aggregate requires at least one RAID group and that the total of all RAID groups in a storage system cannot exceed 400.

1.7.3 EXN4000

The EXN4000 (Figure 1-51) uses the same shelf and hardware as the EXN2000 so it has the same dimensions. EXN4000 also supports the same number of disks per shelf: 14. EXN4000 uses ESH4 as its controller module. ESH4 refers to the third-generation, multiloop speed ESH module. ESH4 can function at 1 Gb, 2 Gb, or 4 Gb loop speed when it works with EXN4000. The ESH4 has LEDs that indicate whether the module is functioning normally (Figure 1-53 on page 47), whether there are any problems with the hardware, and the loop speed operation of the EXN4000. The main differences are:

- Higher bandwidth for heavy sequential workload
- Fewer HBAs or slots used to achieve higher bandwidth needs

The EXN4000 FC storage expansion unit will run at 2 Gbps FC when attached to systems that do not have 4 Gbps capability. ENX4000 can also be added to loops with existing EXN2000 loops.

Figure 1-51 EXN4000 expansion unit

Figure 1-52 on page 47 shows the rear view and the fans.

Figure 1-52 2xESH4 and 2xPSU/fans

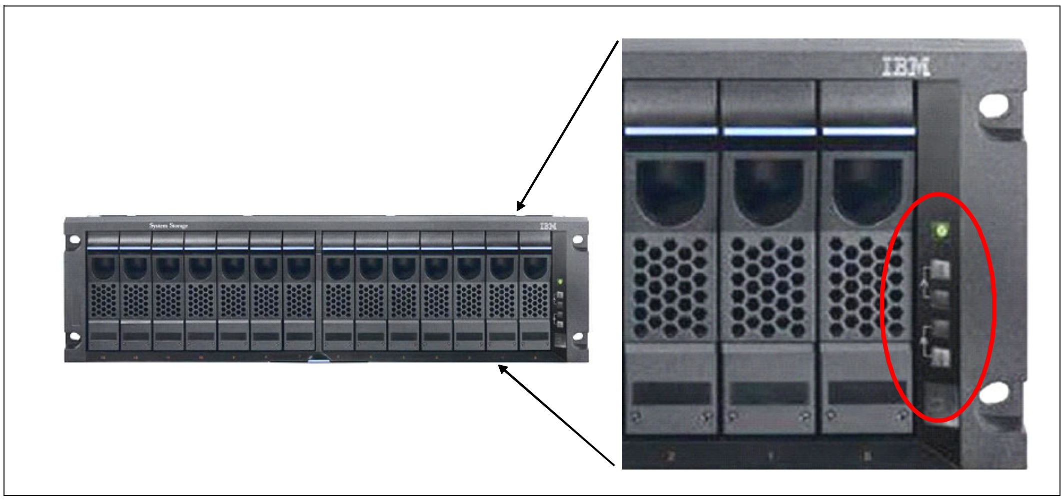


Figure 1-53 Location of LEDs for an ESH4

EXN4000 is the replacement for the EXN2000 FC storage expansion unit.

1.7.4 EXN3000

The IBM N series EXN3000 SAS Expansion Disk Shelf (Figure 1-54 on page 48) can host up to 48 TB of raw capacity in only 4U and 24 drives, and it is available with performance-optimized (SAS) drives or capacity-optimized (SATA) drives. N3400 and N3600 can be connected to up to four EXN3000.

The EXN3000:

- Offers a greater capacity density and new resiliency features such as the ACM and point-to-point SAS technology
- Can offer greater bandwidth with up to 12 Gbps SAS (with 4 wide SAS ports x 3Gbps) versus the current 4Gbps of FC.
- Reduces power consumption with more than 10% Watt consumed per TB of storage.
- Can be attached to new or to already installed N6000 and N7000 series systems and can be attached to N3400 and N3600 systems.

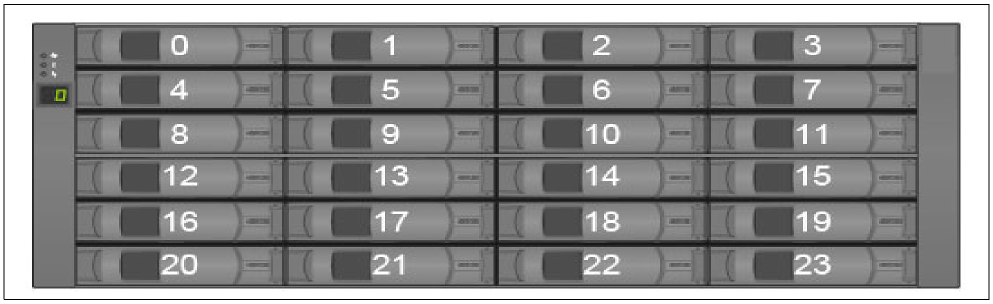


Figure 1-54 EXN3000

2

Active/active configuration and management

IBM System Storage N series active/active configuration consists of two nodes that are able to take over and fail over their resources or services to the associated counterpart nodes. You can configure the two nodes so that they share access to a single set of disks, subnets, and tape drives, or you can configure them to have two distinct sets of storage, each owned by one of the nodes. This functionality assumes that all resources can be accessed by each node. This chapter discusses aspects of determining Active/Active status, as well as Active/Active management.

Topics covered are:

- Active/active overview
- Configuring an active/active configuration
- Managing an active/active configuration
- Active/active configuration failover mode (cfmode) for Fibre Channel

2.1 Active/active overview

An IBM System Storage N series active/active configuration consists of two nodes that can take over and fail over their resources or services to associated counterpart nodes, using the active/active failover (CFO) software feature. Each node continually monitors its partner, mirroring the data for the other's NVRAM. This means that both nodes must have access to all disks physically (cabling) and logically. Figure 2-1 illustrates a basic active/active configuration.

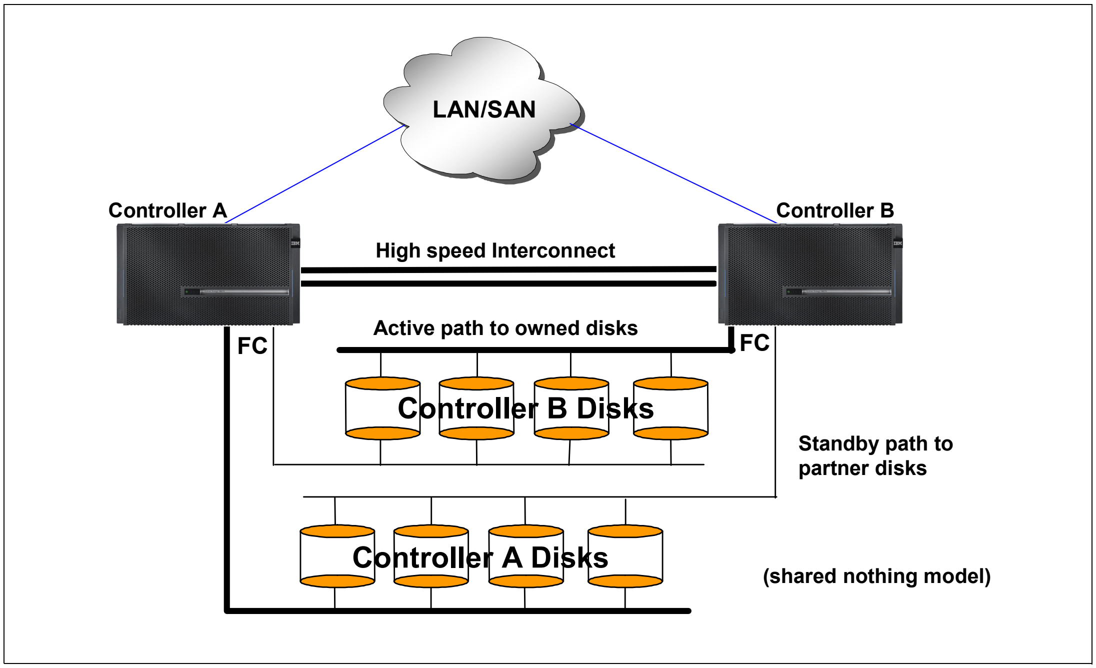


Figure 2-1 Basic active/active configuration

IBM N3000 and the N6000 series systems contain both active/active nodes in the same enclosure, and thus do not require interconnect cables.

In all IBM N5000 and N7000 series models, a standard active/active configuration contains two nodes, with each node contained in a different enclosure. Both nodes must be the same N series model. The two nodes are connected with an Infiniband (IB) active/active cable (interconnect cable) that is attached to an NVRAM5 adapter (for N5200 and N5500 modules) or NVRAM6 adapter (for N5300, N5600, and N7000 series models), which allows one node to serve data to the disks of its failed partner node.

The N6000 series has improved reliability using fewer cables: no external IB cables for clustering. In order to decrease the time required for cluster failover to occur when there is an event that the RLM is aware of, the RLM can communicate with the partner node instance of Data ONTAP. The internal Ethernet switch makes the configuration much easier and facilitates quicker cluster failover.

Active/active configuration terminology

Table 2-1 on page 51 identifies the types of N series active/active configurations (or High Availability pairs) and where each might be applied.

Table 2-1 Configuration types

Active/active configuration type	If A-SIS active	Distance between nodes	Failover possible after loss of entire node (including storage)	Notes
Standard active/active configuration	No	Up to 500 meters[a]	No	Use this configuration to provide higher availability by protecting against many hardware single-point-of-failure.
Mirrored active/active configuration	Yes	Up to 500 meters[a]	No	Use this configuration to add increased data protection to the benefits of a standard active/active configuration.
Stretch MetroCluster	Yes	Up to 500 meters (270 meters if FC 4 Gbps)	Yes	Use this configuration to provide data and hardware duplication to protect against a local disaster.
Fabric-attached MetroCluster	Yes	Up to 100 Km depending on switch configuration. For gateway systems up to 30 Km	Yes	Use this configuration to provide data and hardware duplication to protect against a larger scale disaster.

a. SAS configurations are limited to 5 meters between nodes

Certain terms have very particular meanings when used to refer to active/active configuration. The specialized meanings of these terms follow.

- An *active/active configuration* is a pair of storage systems configured to serve data for each other if one of the two systems becomes impaired. In Data ONTAP documentation and other information resources, active/active configurations are sometimes also referred to as *active/active pairs*.
- When in an active/active configuration, systems are often called *nodes*. One node is sometimes called the *local node*, and the other node is called the *partner node* or *remote node*.
- *Controller failover*, also referred to as *cluster failover* (CFO), refers to the technology that enables two storage systems to take over each other’s data, thus improving data availability.
- *FC direct-attached topologies* are topologies in which the hosts are directly attached to the storage system. Direct-attached systems do not use a fabric or FC switches.
- *FC dual fabric topologies* are topologies in which each host is attached to two physically independent fabrics that are connected to storage systems. Each independent fabric can consist of multiple FC switches. A fabric that is zoned into two logically independent fabrics is not a dual fabric connection.
- *FC single fabric topologies* are topologies in which the hosts are attached to the storage systems through a single FC fabric. The fabric can consist of multiple FC switches.
- *iSCSI direct-attached topologies* are topologies in which the hosts are directly attached to the storage controller. Direct-attached systems do not use networks or Ethernet switches.
- *iSCSI network-attached topologies* are topologies in which the hosts are attached to storage controllers through Ethernet switches. Networks can contain multiple Ethernet switches in any configuration.

- *Mirrored active/active configuration* is similar to the standard active/active configuration, except that there are two copies, or *plexes*, of the data. This is also called *data mirroring*.
- *Remote storage* refers to the storage that is accessible to the local node, but is at the location of the remote node.
- *Single storage controller configurations* are topologies in which there is only one storage controller used. Single storage controller configurations have a single point of failure and do not support cfmodes in FC SAN configurations.
- *Standard active/active configuration* refers to a configuration set up so that one node automatically takes over for its partner when the partner node becomes impaired.

MetroCluster configurations are discussed in the next chapter.

How the interconnect works

The interconnect adapters are among the most critical components in active/active controllers. Data ONTAP uses these adapters to transfer system data between the partner nodes, thereby maintaining data synchronization within the NVRAM on both controllers. Other critical information is also exchanged across the interconnect adapters, including the heartbeat signal, system time, and details concerning temporary disk unavailability due to pending disk firmware updates. The following section explains why NVRAM must be identical on both nodes.

Note: Because NVRAM5 and NVRAM6 cards provide integrated interconnect hardware functionality, standalone interconnect cards are not used (or necessary) when NVRAM5 or NVRAM6 cards are present, except when using the Fabric MetroCluster configuration.

How active/active controllers handle NVRAM

Data ONTAP uses the WAFL® file system to manage data processing and utilizes NVRAM to guarantee data consistency before committing writes to disks. Data within the NVRAM is copied to system memory through Direct Memory Access (DMA). If the storage controller encounters a power failure, the most current data is protected by the NVRAM, and file system integrity is maintained.

Within the active/active controller environment, each node reserves half of the total NVRAM size for the partner node's data to ensure that exactly the same data exists in NVRAM on both storage controllers (Figure 2-2 on page 53). Therefore, only half of the NVRAM in the active/active controller is dedicated to the local node. Dividing the NVRAM in half to guarantee data consistency incurs approximately a 2-3% performance penalty. If failover occurs, when the surviving node takes over the failed node, all WAFL checkpoints stored in NVRAM are flushed to disk. The surviving node then combines the split NVRAM and recovers the lost performance. Once the surviving node restores disk control and data processing to the recovered failed node, all NVRAM data belonging to the partner node is flushed to disk during the course of the giveback operation.

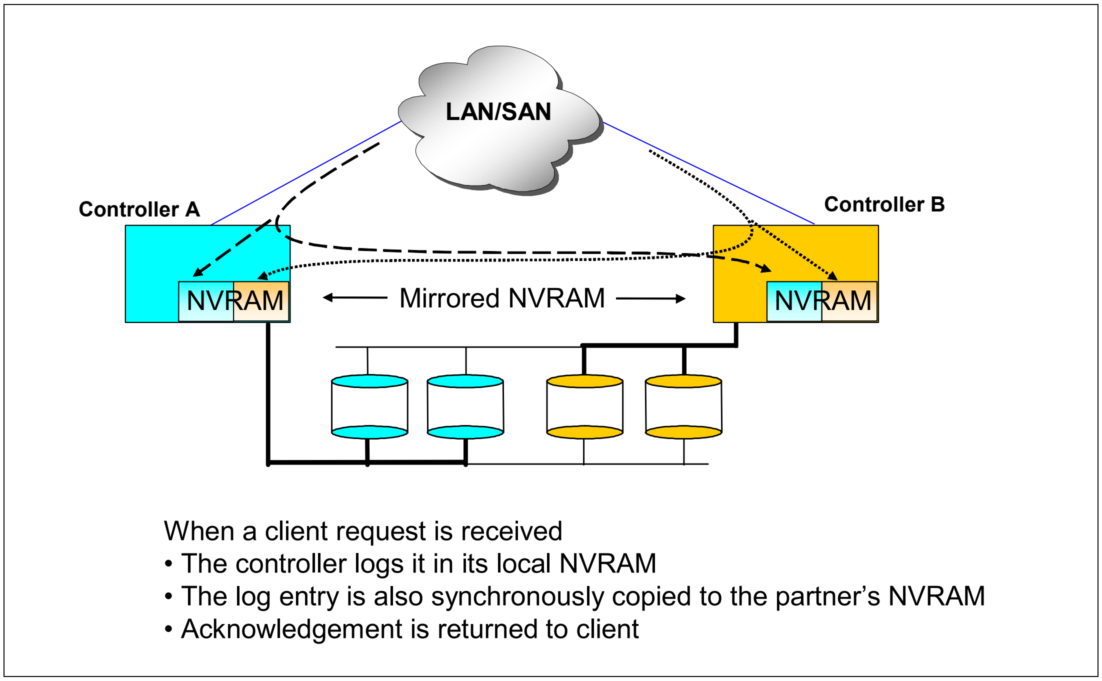


Figure 2-2 Mirrored NVRAM

Mailbox disks store active/active controller synchronization

To ensure that both nodes within the active/active controller configuration maintain the correct and current status of one another, node status and heartbeat information is stored on each node in the mailbox disks (a redundant set of disks used in coordinating takeover or giveback operations). If one node stops functioning, the surviving partner node uses the information on the mailbox disks to perform takeover processing, which creates a virtual storage system. The mailbox heartbeat information prevents an unnecessary failover from occurring in the event of interconnect failure. Moreover, if active/active configuration information stored on the mailbox disks is out of sync at boot time, active/active controller nodes automatically resolve the situation. The N series storage system failover process is extremely robust, preventing split-brain issues from occurring.

2.1.1 Benefits of active/active configurations

Configuring storage systems in an active/active configuration provides the following benefits:

- Fault tolerance

  When one node fails or becomes impaired a takeover occurs, and the partner node continues to serve the failed node's data.

- Nondisruptive software upgrades

  When you halt one node and allow takeover, the partner node continues to serve data for the halted node while you upgrade the node that you halted. Or, in the case of a failover event, the partner node automatically takes over. (Figure 2-3 on page 54 illustrates a failover configuration.) For more information about nondisruptive upgrades, see the *IBM System Storage N series Data ONTAP Upgrade Guide* for your version of Data ONTAP at:

  http://www-03.ibm.com/systems/storage/network/

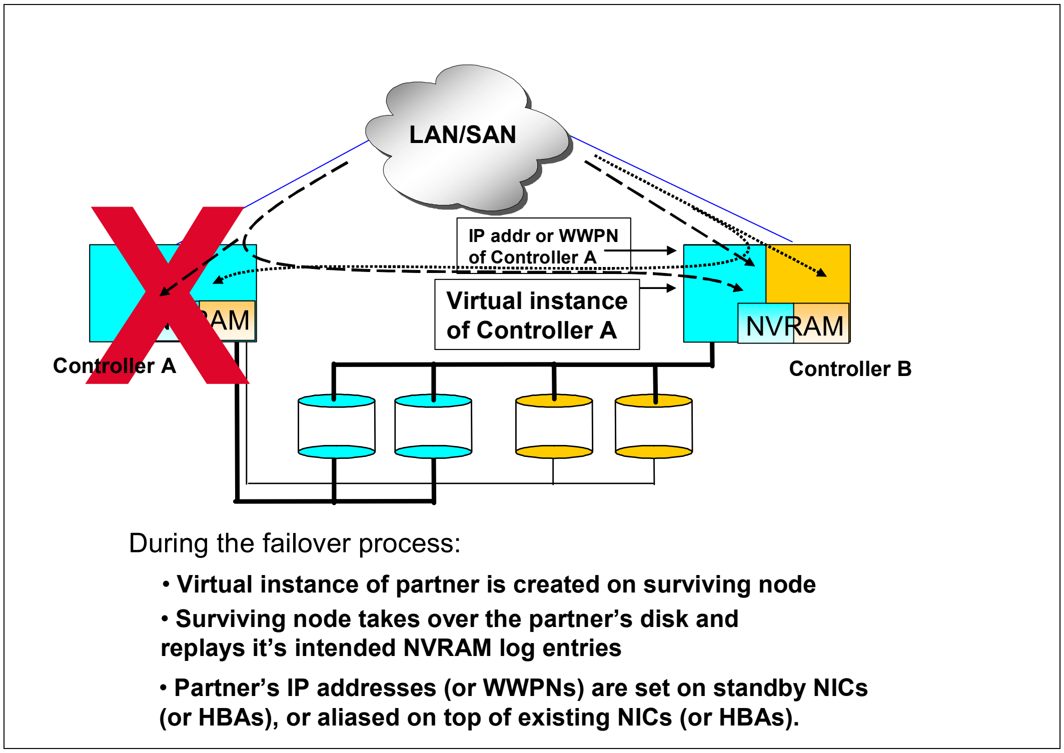


Figure 2-3 Failover configuration

- Non-disruptive storage system and disk maintenance

  When you halt one node and allow takeover, the partner node continues to serve data for the halted node while you replace or repair hardware in the node that you halted (or that was halted due to failure).

2.1.2 Standard active/active requirements and restrictions

You must follow certain requirements and restrictions when setting up a new standard active/active configuration. These restrictions are as follows:

- Architecture compatibility

  Both nodes must have the same system model. See the Data ONTAP Release Notes for the list of supported systems at:

  `www.ibm.com/storage/support/nas`

  > **Note:** In the case of systems with two controller modules in a single chassis, both nodes of the active/active configuration are located in the same chassis and have an internal interconnect.

- Storage capacity

  The number of disks must not exceed the maximum configuration capacity. In addition, the total storage attached to each node must not exceed the capacity for a single node.

Note: After a failover, the takeover node temporarily serves data from all the storage in the active/active configuration. When the single-node capacity limit is less than the total active/active configuration capacity limit, the total disk space in a active/active configuration can be greater than the single-node capacity limit. It is acceptable for the takeover node to temporarily serve more than the single-node capacity would normally allow, as long as it does not own more than the single-node capacity.

- Disks and disk shelf compatibility
  - Both Fibre Channel and SATA storage are supported in standard active/active configurations, as long as the two storage types are not mixed on the same loop.
  - One node can have only Fibre Channel storage and the partner node can have only SATA storage if needed.
- Active/active interconnect adapters and cables must be installed.
- Nodes must be attached to the same network and the network interface cards (NICs) must be configured correctly.
- The same system software, such as Common Internet File System (CIFS), Network File System (NFS), or SyncMirror, must be licensed and enabled on both nodes.

Note: If a takeover occurs, the takeover node can provide only the functionality for the licenses installed on it. If the takeover node does not have a license that was being used by the partner node to serve data, your active/active configuration loses functionality after a takeover.

2.1.3 Data ONTAP and active/active configuration

In a standard active/active configuration, one Data ONTAP function is to monitor each other's partner node through a heartbeat signal sent between the nodes. Data from the NVRAM of one node is mirrored by its partner node, and each node can take over the partner's disks and volumes if the partner fails.

Keep in mind that this mirroring reduces the amount of NVRAM available to each node, so you should factor this into your initial sizing. Also, the nodes maintain synchronization of time.

2.1.4 Active/active configuration node interaction

For basic operations the nodes operate independently, performing operations on a separate basis. Interactions under normal conditions have the following characteristics:

- Each node owns its spare disks and does not share them with the other node.
- Each node has two mailbox disks on the root volume (four, if the root volume is mirrored using the SyncMirror feature). The mailbox disks are used to perform the following tasks:
  - Maintain consistency between the pair.
  - Continually check whether the other node is running or whether it has performed a takeover.
  - Store active/active information that is not specific to any particular node.
- The nodes can reside on the same Microsoft Windows domain or on different domains.

2.1.5 Required licenses

Table 2-2 lists the required licenses for active/active configuration.

Table 2-2 Active/active license

Active/active configuration	Required licenses
Standard (CFO) cluster	Cluster

The Cluster Failover Option can be purchased as a stand-alone feature (CFO) or in the SAN Bundle feature (CFO + FCP) depending on your specific needs.

2.2 Configuring an active/active configuration

This section describes how to bring up a new standard active/active configuration for the first time and how to enable licenses, set options, configure networking, test the configuration, and discuss the different modes of active/active configurations.

2.2.1 Configuration variations for standard active/active configurations

The following list describes configuration variations that are supported for standard active/active configurations:

- Asymmetrical configurations

  In an asymmetrical standard active/active configuration, one node has more storage than the other. This is supported as long as neither node exceeds the maximum capacity limit for the node.

- Active/passive configurations

  In this configuration, the passive node has only a root volume, and the active node has all the remaining storage and services all data requests during normal operation. The passive node responds to data requests only if it has taken over the active node.

- Shared loops or stacks

  If your standard active/active configuration is using software-based disk ownership, you can share a loop or stack between the two nodes. This is particularly useful for active/passive configurations.

- Multipath storage

  Multipath storage for active/active configurations provides a redundant connection from each node to every disk. It can prevent some types of failovers.

2.2.2 Best practices for active/active configurations

Follow these best practices to ensure that active/active storage systems achieve maximum uptime:

- Make sure that the active/active controller units and shelves are on different power supplies or grids so that a single power outage does not affect both controller units and shelves.
- Use virtual interfaces (VIFs) to provide redundancy and improve the availability of network communication. The virtual interfaces are set up during initial installation or the subsequent initiation of setup.

- Maintain a consistent configuration between active/active nodes (such as Data ONTAP versions). An inconsistent active/active storage system configuration is often related to failover problems.
- Test the failover capability periodically (for example, during planned maintenance) to ensure a proper active/active storage system configuration.
- Follow the documented procedures in the upgrade guide when upgrading active/active storage systems.
- Make sure that active/active nodes have sufficient resources to adequately support workload during takeover mode.
- Periodically use the active/active configurations checker to help ensure that failovers are successful.
- Make sure the /etc/rc file is correctly configured. See Example 2-1.

Example 2-1 Example of /etc/rc files

```
/etc/rc on itsotuc1:
hostname itsotuc1
ifconfig e0 `hostname`-e0 mediatype 100tx-fd netmask 255.255.255.0
vif create multi vif1 e3a e3b e3c e3d
ifconfig vif1 `hostname`-vif1 mediatype 100tx-fd netmask 255.255.255.0 partner vif2
route add default 10.10.10.1 1
routed on
savecore
exportfs -a
nfs on

/etc/rc on itsotuc2:
hostname itsotuc2
ifconfig e0 `hostname`-e0 mediatype 100tx-fd netmask 255.255.255.0
vif create multi vif2 e3a e3b e3c e3d
ifconfig vif2 `hostname`-vif2 mediatype 100tx-fd netmask 255.255.255.0 partner vif1
route add default 10.10.10.1 1
routed onsavecore
exportfs -a
nfs on
```

2.2.3 Enabling licenses on the active/active configuration

To enable a license on the active/active configuration:

1. For each required license, enter the license and code on both node consoles, as shown in the Example 2-2.

   Example 2-2 Enabling license

   ```
   license add xxxxx
   where xxxxxis the license code you received for the feature
   ```

2. Reboot both nodes using the command **reboot**.
3. Enable active/active capability on each node by entering the command **cf enable** on the local node console.
4. Verify that active/active capability is enabled by entering the command **cf status** on each node console, as shown in the Example 2-3.

Example 2-3 Confirming if active/active configuration is enabled

```
cf status
Cluster enabled, nas2 is up
```

5. Repeat for any other licenses that you need to enable using the license type and code for each licensed product installed on the active/active configuration.

2.2.4 Configuring network interfaces for the active/active configuration

You must configure network interfaces so that if takeover occurs, interfaces on the operating node take over interfaces on the failed-over node and hosts can still reach data over the network.

Important considerations:

- Both nodes in the active/active configuration must have interfaces that access the same collection of networks and subnetworks.
- You must gather the following information before configuring the interfaces:
  - The IP addresses for both the local node and partner node
  - The netmasks for both the local node and partner node
- You should always use multiple NICs with VIFs to improve networking availability for both stand-alone storage systems and systems in an active/active configuration.

There are three types of interface configurations that you can enable in a active/active configuration:

- Shared

  This type of interface supports both the local and partner nodes. It contains both the local node and partner node IP addresses. During takeover, it supports the identity of both nodes.

- Dedicated

  This type of interface only supports the node in which it is installed. It contains the local node IP address only and does not participate in network communication beyond local node support during takeover. It is paired with a standby interface.

- Standby

  This type of interface is on the local node, but only contains the IP address of the partner node. It is paired with a dedicated interface.

Table 2-3 shows the interface roles in normal and takeover modes.

Table 2-3 Interface roles

Interface type	Normal mode	Takeover mode
Shared	Supports the identity of the local node	Supports the identity of both the local node and the failed node
Dedicated	Supports the identity of the local node	Supports the identity of the local node
Standby	Idle	Supports the identity of the failed node

Takeover configuration with shared interfaces

Figure 2-4 shows the use of two NICs to provide the two interfaces.

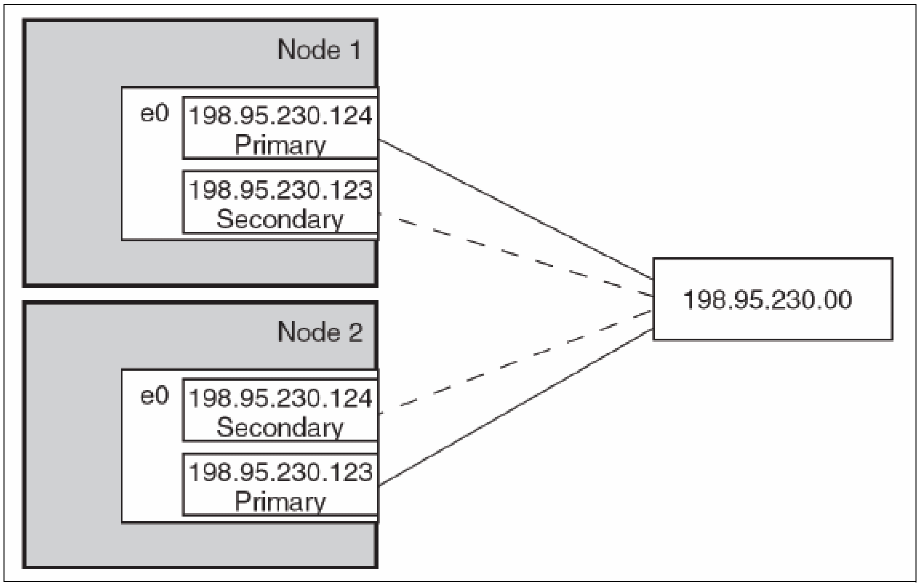


Figure 2-4 Takeover configuration with shared interfaces

If node 1 fails, interface e0 on node 1 stops functioning, but the secondary address on e0 on node 2 handles the node 1 network connection with the 230 network.

If node 2 fails, e0 on node 2 stops functioning, but e0 on node 1 substitutes for the failed interface and handles the node 2 network connection with the 230 network.

Configuring shared interfaces

Enter the command **`ifconfig`** on the command line of both nodes in a active/active configuration, as shown in Example 2-4, and in the /etc/rc file of both nodes so that the command is permanent.

Example 2-4 Configuring shared interfaces

```
ifconfig e0 199.9.204.255 partner 199.9.204.254
```

The following synopsis helps to determine the correct parameters of the **`ifconfig`** command:

`ifconfig` *`interface local_address partner partner_address`*`{`*`other_options`*`}`

Parameters have the following meanings:

Interface	Name of the local interface
local_address	IP address of the local interface
partner_address	IP address or interface name that the local interface services when the local node takes over the other node in the active/active configuration
other_options	Any other options needed to correctly configure the interface in your network environment

Takeover configuration with dedicated/standby interfaces

Figure 2-5 shows the use of two NICs for each interface, one on each storage system. One NIC acts as a dedicated interface and the other acts as a standby interface.

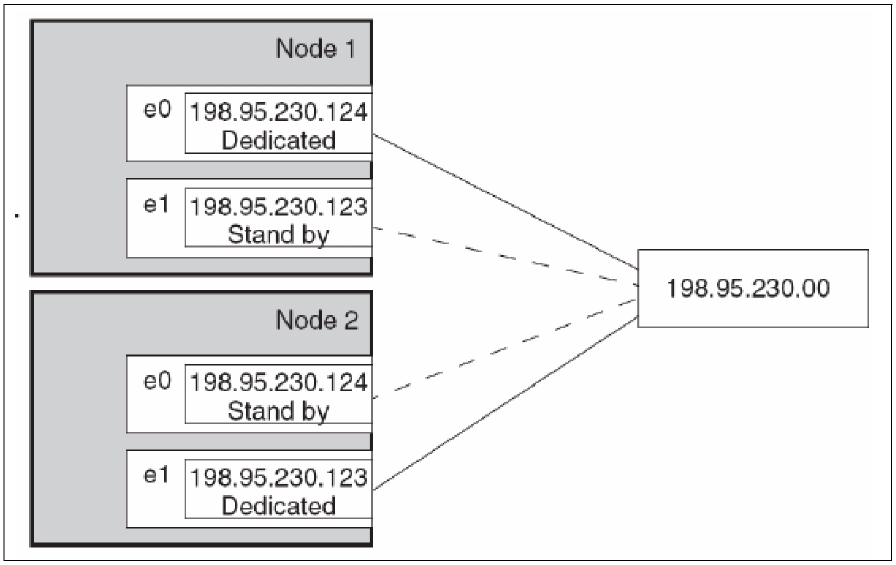


Figure 2-5 Takeover configuration with dedicated/standby Interface

If node 1 fails, interface e0 on node 1 stops functioning, but e0 on node 2 substitutes for the failed interface and handles the node 1 network connection with the 230 network.

If node 2 fails, e1 on node 2 stops functioning, but e1 on node 1 substitutes for the failed interface and handles the node 2 network connection with the 230 network.

Configuring dedicated and standby interfaces

On nodeA, enter the following commands on the command line and also enter them in the /etc/rc file so that the command is permanent. To configure the dedicated local interface for nodeA:

```
ifconfig interfaceA1 addressA1 {other_options}
```

Parameters have the following meanings:

interfaceA1	Name of the dedicated local interface for nodeA
addressA1	IP address of the dedicated local interface for nodeA
other_options	Any other options needed to correctly configure the interface in your network environment

To configure the standby interface for nodeA to take over the dedicated interface of nodeB on takeover:

```
ifconfig interfaceA2 partner addressB1
```

In this command:

interfaceA2	Name of the standby interface for nodeA
addressB1	IP address or interface name of the dedicated interface of nodeB

On nodeB, enter the following commands on the command line and in the /etc/rcfile. To configure the dedicated local interface for nodeB:

```
ifconfig interfaceB1 addressB1 {other_options}
```

Parameters have the following meanings:

interfaceB1	Name of the dedicated local interface for nodeB
addressB1	IP address of the dedicated local interface for nodeB
other_options	Any other options needed to correctly configure the interface in your network environment

To configure the standby interface on nodeB to take over the dedicated interface of nodeA on takeover:

```
ifconfig interfaceB2 partner addressA1
```

In this command:

interfaceB2	Name of the standby interface on nodeB
addressA1	IP address or interface name of the dedicated interface for nodeA

Interface types and configurations

Table 2-4 lists the configurations supported by each type of interface in an active/active configuration.

Table 2-4 Configurations supported in an active/active configuration

Interface	Shared	Dedicated	Standby	Partner parameter
Ethernet	X	X	X	IP address or interface name
Gigabit Ethernet	X	X	X	IP address or interface name
Virtual interface	X	X	X	Virtual interface name
VLAN interface	X	X	X	IP address or interface name

Note: Some storage systems, such as the N6040 and N6070 systems, include an e0M interface that is dedicated to management traffic. This port can be partnered in an active/active configuration in the same way as a regular Ethernet interface.

2.2.5 Setting options and parameters

Some options must be the same on both nodes in the active/active configuration, while some can be different, and some are affected by failover events.

In an active/active configuration, options are one of the following types:

- Options that must be the same on both nodes for the active/active configuration to function correctly.
- Options that might be overwritten on the node that is failing over. These options must be the same on both nodes to avoid losing system state after a failover.
- Options that should be the same on both nodes so that system behavior does not change during failover.
- Options that can be different on each node.

Note: You can determine whether an option must be the same on both nodes of an active/active configuration from the comments that accompany the option value when you enter the `option` command. If there are no comments, the option can be different on each node.

Setting matching node options

Because some Data ONTAP options must be the same on both the local and partner node, you must check these options with the **`options`** command on each node and change them as necessary.

Follow these steps to check the options:

1. View and note the values of the options on the local and partner nodes, using the following command on each console:

   ```
   options
   ```

   The current option settings for the node are displayed on the console. Output similar to the following is displayed:

   ```
   autosupport.doit DONT
   autosupport.enable on
   ```

2. Verify that the options with comments in parentheses are set to the same value for both nodes. The comments are as follows:

   ```
   Value might be overwritten in takeover
   Same value required in local+partner
   Same value in local+partner recommended
   ```

3. Correct any mismatched options using the following command:

   ```
   options option_name option_value
   ```

Note: See the na_options man page for more information about the options.

Parameters that must be the same on each node

The parameters listed in Table 2-5 must be the same so that takeover is smooth and data is transferred between the nodes correctly.

Table 2-5 Parameters that must be the same in both nodes

Parameter	Setting for...
date	date, rdate
NDMP (on or off)	ndmp (on or off)
route table published	route
route enabled	routed (on or off)
Time zone	timezone

2.2.6 Testing takeover and giveback

After you configure all aspects of your active/active configuration, execute the following steps to verify that it operates as expected:

1. Check the cabling on the active/active configuration interconnect cables to make sure that they are secure.
2. Verify that you can create and retrieve files on both nodes for each licensed protocol.
3. Enter the following command from the local node console:

   ```
   cf takeover
   ```

   The local node takes over the partner node and the following message is displayed:

   ```
   takeover completed
   ```

4. Test communication between the local node and partner node. For example, you can use the **`fcstat device_map`** command to ensure that one node can access the other node's disks.
5. Give back the partner node by entering the following command:

   ```
   cf giveback
   ```

   The local node releases the partner node, which reboots and resumes normal operation. The following message is displayed on the console when the process is complete:

   ```
   giveback completed
   ```

6. Proceed as shown in Table 2-6, depending on whether you got the message that giveback was completed successfully.

Table 2-6 Takeover and giveback messages

If takeover and giveback...	Then...
Is completed successfully	Repeat steps 2 through 5 on the partner node.
Fails	Attempt to correct the takeover or giveback failure.

2.2.7 Eliminating single points of failure with active/active configurations

Table 2-7 lists the ways that using active/active configurations help you to avoid single points of failure (SPOFs) in various hardware components.

Table 2-7 Avoiding single points of failure by using active/active configurations

Hardware component	SPOF		SPOF eliminated
	Non-active/ active	Active/active	
IBM System Storage N series storage system	Yes	No	If a storage system fails, cluster failover automatically fails over to its partner storage system and serves data from the takeover system.
NVRAM	Yes	No	If an NVRAM adapter fails, CF automatically fails over to its partner storage system and serves data from the takeover storage system.
CPU fan	Yes	No	If the CPU fan fails, the node gracefully shuts down. CF automatically fails over to its partner storage system and serves data from the takeover storage system.
Multiple NICs with VIFS (virtual interfaces)	No	No	If one of the networking links fails, the networking traffic is automatically sent over the remaining networking links on the storage system. No failover is needed in this situation. If all NICs fail, you can initiate failover to a partner storage system and serve data from the takeover storage system. **Note:** Always use multiple NICs with VIFS to improve networking availability for both single storage systems and active/active storage systems.
Single NIC	Yes	No	If a NIC fails, you can initiate a failover to its partner storage system and serve data from the takeover storage system.

Hardware component	SPOF		SPOF eliminated
	Non-active/ active	Active/active	
FC-AL card	Yes	No	If an FC-AL card for the primary loop fails, the partner node attempts a failover at the time of failure. If the FC-AL card for the secondary loop fails, the failover capability is disabled, but both storage systems continue to serve data to their respective applications and users, with no impact or delay.
Disk drive	No	No	If a disk fails, the storage system can reconstruct data from the RAID 4 or RAID DP. No failover is needed in this situation.
Disk shelf (including backplane)	No	No	A disk shelf is a passive backplane with dual power supplies, dual fans, dual ESH2s, and dual FC-AL loops. It is the most reliable component in a storage system.
Power supply	No	No	Both the storage system and the disk shelf have dual power supplies. If one power supply fails, the second power supply automatically kicks in. No failover is needed in this situation.
Fan (storage system or disk shelf)	No	No	Both the storage system head and disk shelf have multiple fans. If one fan fails, the second fan automatically provides cooling. No failover is needed in this situation.
Cluster adapter	N/A	No	If a cluster adapter fails, the failover capability is disabled but both storage systems continue to serve data to their respective applications and users.
Active/active configuration interconnect cable	N/A	No	The cluster adapter supports dual cluster interconnect cables. If one cable fails, the active/active traffic (heartbeat and NVRAM data) is automatically sent over the second cable with no delay or interruption. If both cables fail, the failover capability is disabled, but both storage systems continue to serve data to their respective applications and users.

2.2.8 Fibre Channel topologies for active/active configuration

Supported FC configurations for active/active configuration include direct-attached, single-fabric, and dual-fabric topologies.

Active/active configurations with dual, physically independent storage fabrics are recommended for SAN solutions. This provides redundancy at the fabric and storage system layers, which is particularly important because these layers typically support many hosts.

Note: The use of a heterogeneous SAN, or a fabric with storage from multiple vendors, is supported as long as the storage systems are contained in exclusive zones, or VSANs. A tape SAN should also be in a separate zone or VSAN, apart from the disk SAN.

Cascade, mesh, and core-edge fabrics are all industry-accepted methods of connecting FC switches into a fabric, and all are supported.

Fibre Channel supported hop count

The maximum supported FC hop count, or the number of inter-switch links (ISLs) crossed between a particular host and storage system, depends on the hop count that the switch supplier and storage system support for FC configurations. Table 2-8 shows the supported hop count for each switch supplier.

Table 2-8 Supported hop count

Switch supplier	Supported hop count
Brocade	5
Cisco	5
McData	3
Qlogic	4

Fibre Channel switch configuration best practices

A fixed link speed setting is highly recommended, especially for large fabrics, because it provides the best performance for fabric rebuild times. In large fabrics, this can create significant time savings.

Although auto-negotiation provides the greatest flexibility, it does not always perform as expected. Also, it adds time to the overall fabric-build sequence because the FC port must auto-negotiate.

The following sections show detailed SAN configuration diagrams for the N series storage system in active/active configuration. However, it is supported for single-controller configurations as well. For simplicity, the diagrams show only a single fabric or, in the case of the dual-fabric configurations, two fabrics. For complete information about SAN configuration for N series Storage, see the *iSCSI and Fibre Channel Configuration Guide* at:

http://www-01.ibm.com/support/docview.wss?uid=ssg1S7002547&rs=573

N7000 series supported topologies

Each N7000 series storage system is capable of supporting 2 Gb, 4 Gb, or 8 Gb FC target connections, but you cannot use all on the same storage system or on two different storage systems in an active/active configuration.

N7600 and N7800 storage systems are supported by single_image and, in some configurations, standby cfmode. N7700 and N7900 storage systems are only supported by single_image cfmode.

N7000 series target port configuration recommendations

Table 2-9 shows recommended target ports using single_image cfmode. The port pairs on a N7000 series storage system that share an ASIC are 0a+0b, 0c+0d, 0e+0f, 0g+0h.

Table 2-9 Target port configuration

Number of target ports	Ports
1	0h
2	0h, 0d
3	0h, 0d, 0f
4	0h, 0d, 0f, 0b
5	0h, 0d, 0f, 0b, 0g
6	0h, 0d, 0f, 0b, 0g, 0c

Number of target ports	Ports
7	0h, 0d, 0f, 0b, 0g, 0c, 0e,
8	0h, 0d, 0f, 0b, 0g, 0c, 0e, 0a

N7000 series: Dual-fabric active/active configuration, two onboard FC ports

Figure 2-6 and Table 2-10 show the supported topology and attributes specific to this configuration.

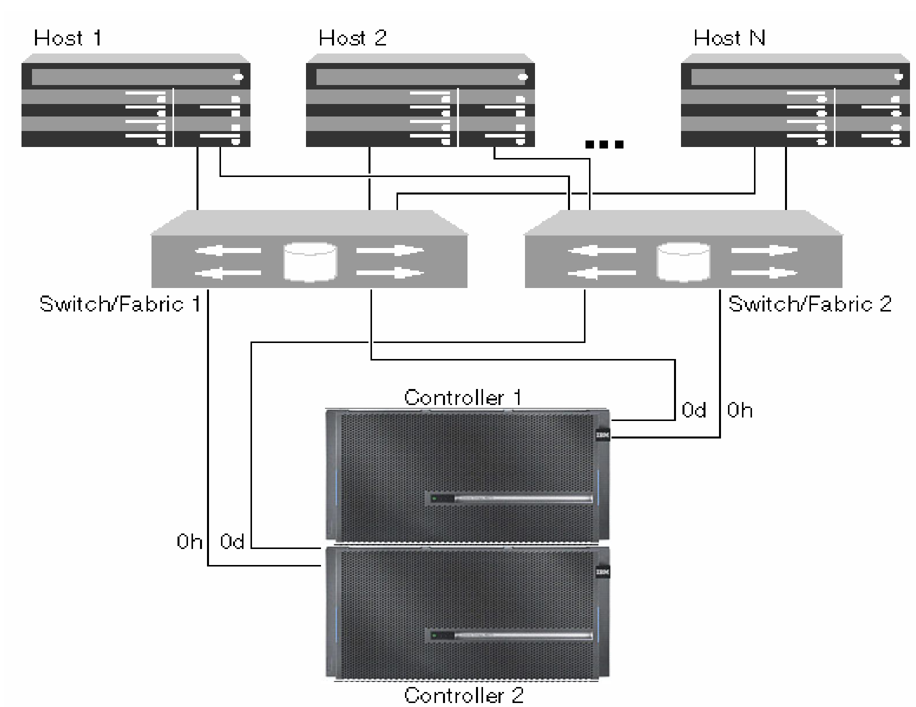


Figure 2-6 Dual-fabric active/active configuration, two onboard FC ports

Table 2-10 Dual-fabric active/active configuration, two onboard FC ports

Attribute	Value
Fully redundant	Yes
Type of fabric	Dual fabric
Different host operating systems	Yes, with multiple host configurations
FC ports or adapters	Two onboard FC ports per controller
Supported cfmodes	Single_image cfmode
Multipathing required	Yes, for full redundancy
Type of configuration	Active/active configuration

N7000 series: Dual-fabric active/active configuration, two ports of a 4 Gb FC target HBA

Figure 2-7 and Table 2-11 show the supported topology and attributes specific to this configuration.

> **Note:** The 4 Gb FC target HBA port numbers (5a and 5b) are examples. The actual port numbers might vary depending on the expansion slot in which the FC target HBAs are installed. Also 8 Gb target ports can now be configured on PCIe slots.

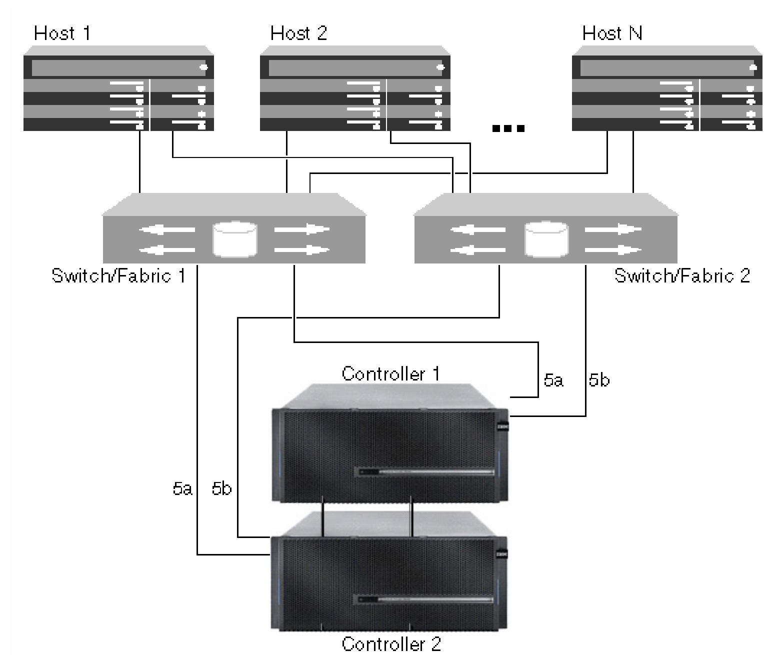


Figure 2-7 Dual-fabric active/active configuration, two ports of a 4 Gb FC target HBA

Table 2-11 Dual-fabric active/active configuration, two ports of a 4 Gb FC target HBA

Attribute	Value
Fully redundant	Yes
Type of fabric	Dual fabric
Different host operating systems	Yes, with multiple host configurations
FC ports or adapters	Two ports of a 4-Gb FC target HBA per controller
Supported cfmodes	Single_image cfmode
Multipathing required	Yes, for full redundancy
Type of configuration	Active/Active configuration

N7000 series: Direct-attached active/active configuration, one onboard FC port

Figure 2-8 and Table 2-12 show the supported topology and attributes specific to this configuration.

Additional hosts can be attached directly to the onboard FC ports on each controller.

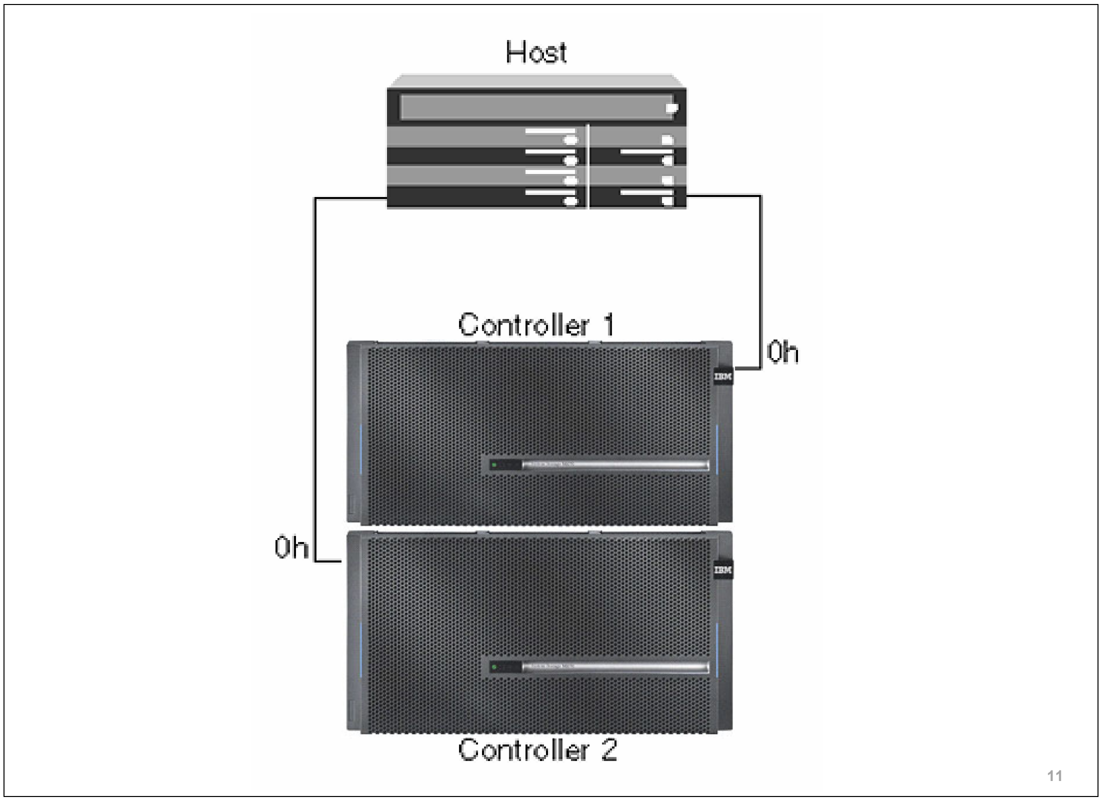


Figure 2-8 Direct-attached active/active configuration, one onboard FC port

Table 2-12 Direct-attached active/active configuration, one onboard FC port

Attribute	Value
Fully redundant	Yes
Type of fabric	None
Different host operating systems	Yes, with multiple host configurations
FC ports or adapters	One onboard FC ports
Supported cfmodes	Single_image cfmode
Multipathing required	Yes, for full redundancy
Type of configuration	Active/active configuration

N6000 series supported topologies

N6040 and N6070 systems are supported by single_image cfmode.

Each N6000 series storage systems supports 4 Gb FC target HBAs. 8 Gb FC target HBAs are supported on PCIe slots.

N6000 series target port configuration recommendations

Table 2-13 shows recommended target ports using single_image cfmode.

The port pairs on an N6000 series storage system that share an ASIC are target ports 0a+0b, 0c+0d.

Table 2-13 Recommended target ports using single_image cfmode

Number of target ports	Ports
1	0d
2	0d, 0b
3	0d, 0b, 0c
4	0d, 0b, 0c, 0a

N6040 and N6070 supported topologies

N6040 and N6070 storage systems are supported by single_image cfmode.

The N6040 and N6070 storage systems have four onboard 4 Gb FC ports per controller and each port can be configured as either an FC target port or an initiator port. For example, you can configure two ports as SAN targets and two ports as initiators for disk shelves.

N6040 and N6070: Dual-fabric active/active configuration, four onboard FC ports

Figure 2-9 and Table 2-14 show the supported topology and attributes specific to this configuration.

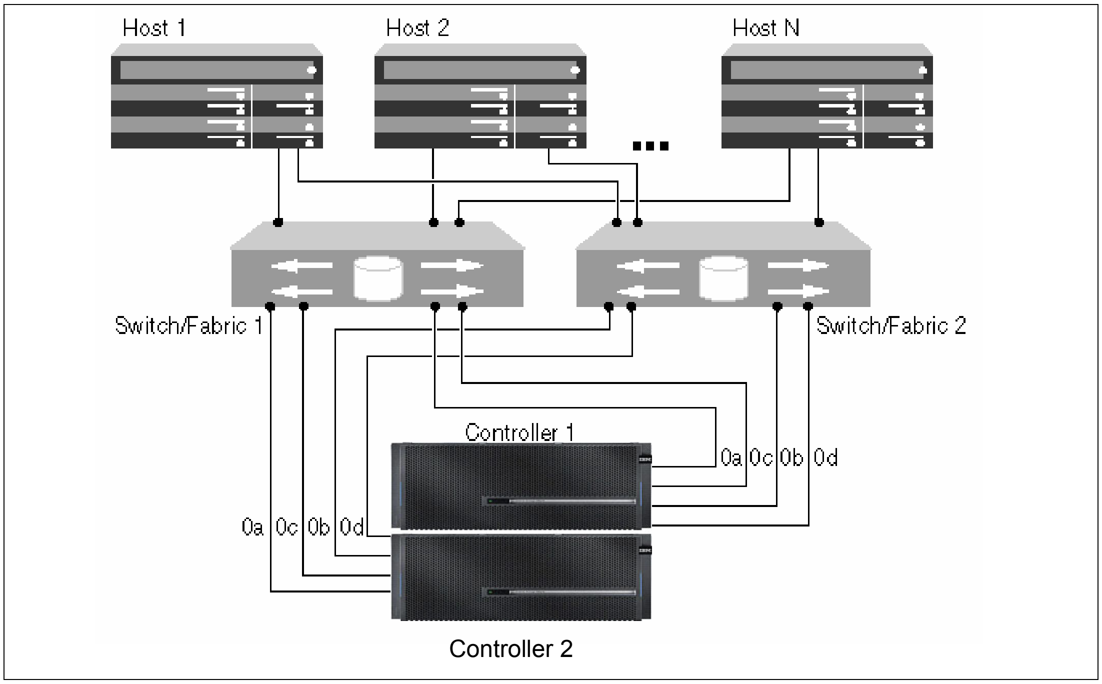


Figure 2-9 N6040 and N6070: Dual-fabric active/active configuration, four onboard FC ports

Table 2-14 N6040 and N6070: Dual-fabric active/active configuration, four onboard FC ports

Attribute	Value
Fully redundant	Yes
Type of fabric	Dual fabric
Different host operating systems	Yes, with multiple host configurations
FC ports or adapters	Four onboard FC ports per controller
Supported cfmodes	Single_image cfmode
Multipathing required	Yes, for full redundancy
Type of configuration	Active/active configuration

N6040 and N6070: Single-fabric active/active configuration, two onboard FC ports

Figure 2-10 and Table 2-15 show the supported topology and attributes specific to this configuration.

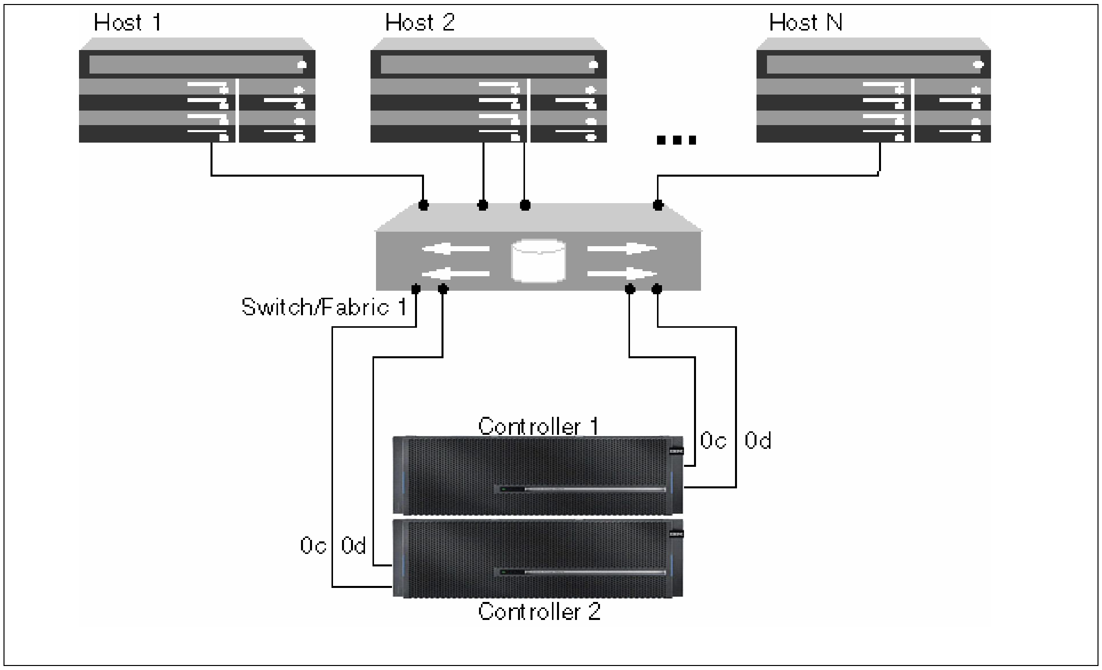


Figure 2-10 N6040 and N6070: Single-fabric active/active configuration, two onboard FC ports

Each controller can have one or up to the maximum number of target ports supported per controller connecting into the fabrics; an active/active configuration can have two or up to the maximum number of target ports supported per configuration connecting into the fabric. See the system configuration guide for the version of Data ONTAP being used by the controllers for these values.

Table 2-15 N6040 and N6070: Single-fabric active/active configuration, two onboard FC ports

Attribute	Value
Fully redundant	No, due to the single fabric
Type of fabric	Single fabric
Different host operating systems	Yes, with multiple host configurations
FC ports or adapters	Two onboard FC ports per controller
Supported cfmodes	Single_image cfmode
Multipathing required	Yes, if a host has multiple paths to a single LUN
Type of configuration	Active/active configuration

N6040 and N6070: Dual-fabric active/active configuration, two ports of multiple 4-Gb FC target HBAs

Figure 2-11 and Table 2-16 show the supported topology and attributes specific to this configuration.

Note: The 4 Gb FC target HBA port numbers (2a, 2b, 4a, and 4b) are examples. The actual port numbers might vary, depending on the expansion slot in which the FC target HBAs are installed.

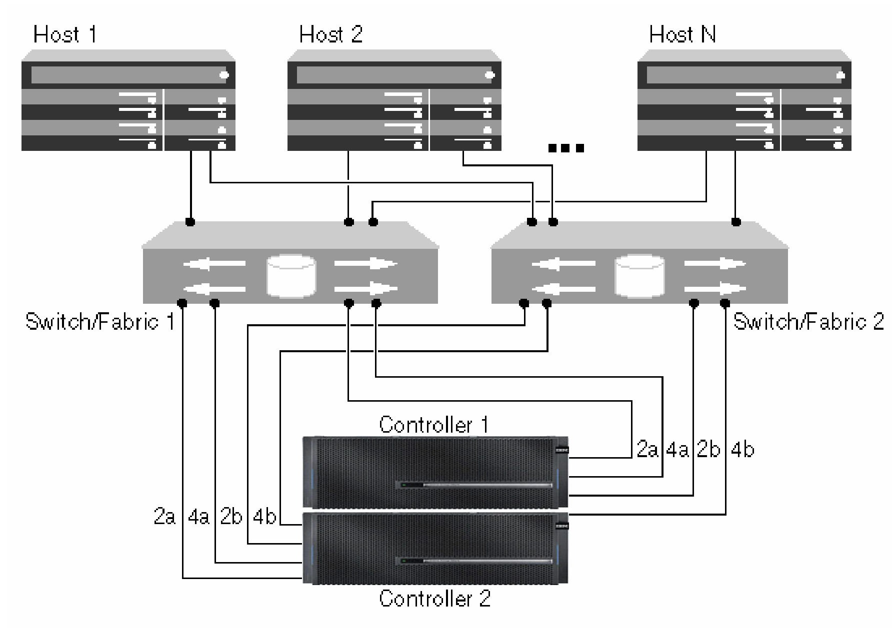


Figure 2-11 N6040 and N6070: Dual-fabric active/active configuration, two ports of multiple 4 Gb FC target HBAs

Table 2-16 N6040 and N6070: Dual-fabric active/active configuration, two ports of multiple 4 Gb FC target HBAs

Attribute	**Value**
Fully redundant	Yes
Type of fabric	Dual fabric
Different host operating systems	Yes, with multiple host configurations
FC ports or adapters	Two ports of multiple 4 Gb FC target HBAs per controller
Supported cfmodes	Single_image cfmode
Multipathing required	Yes, for full redundancy
Type of configuration	Active/Active configuration

N6040 and N6070: Direct-attached active/active configuration, one port of a 4 Gb FC target

Figure 2-12 and Table 2-17 show the supported topology and attributes specific to this configuration.

You might have to set the Data ONTAP FCP adapter media type to loop on the 4 Gb FC target HBAs for some direct-attached topologies.

Note: The 4 Gb FC target HBA port number (2a) is an example. The actual port number might vary, depending on the expansion slot in which the FC target HBA is installed.

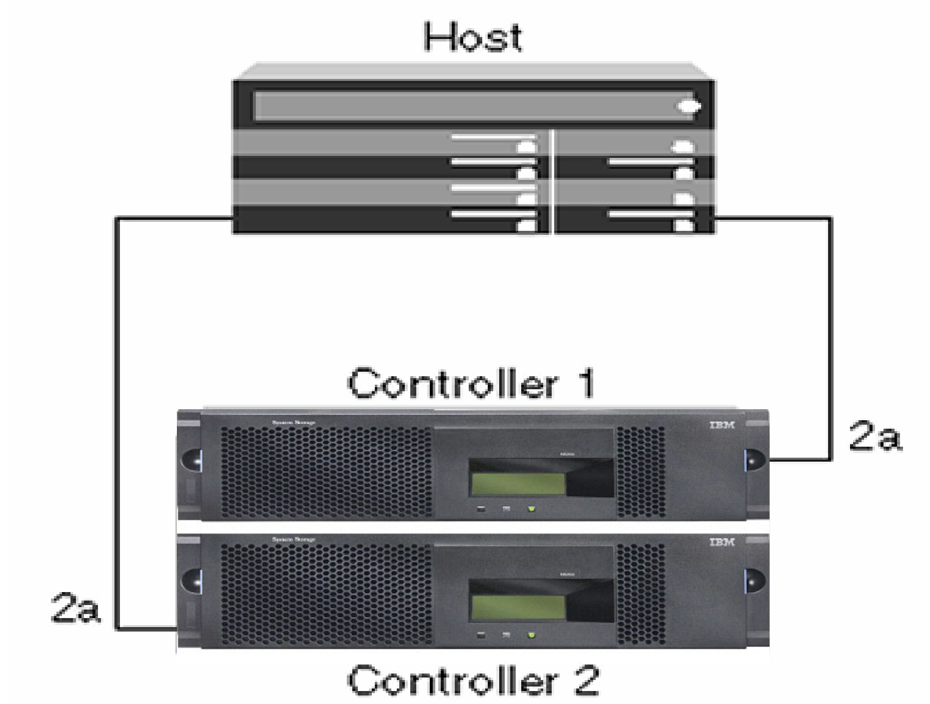


Figure 2-12 N6040 and N6070: Direct-attached active/active configuration, one port of 4 Gb FC target

Table 2-17 N6040 and N6070: Direct-attached active/active configuration, one port of 4 Gb FC target

Attribute	**Value**
Fully redundant	Yes
Type of fabric	None
Different host operating systems	Yes, with multiple host configurations
FC ports or adapters	One port of a 4 Gb FC target HBA per controller
Supported cfmodes	Single_image cfmode
Multipathing required	Yes, for full redundancy
Type of configuration	Active/active configuration

N5000 series supported topologies

N5200 and N5500 systems are supported by single_image, standby, and partner cfmodes. N5300 and N5600 systems are supported by single_image and standby cfmodes.

Each N5200 and N5500 series storage system supports 2 Gb or 4 Gb FC target connections, but you cannot use both on the same storage system or on two different storage systems in an active/active configuration. If you use target expansion adapters, then you can only use onboard adapters as initiators.

Each N5300 and N5600 series storage system supports 2 Gb, 4 Gb, or 8 Gb FC target connections, but you cannot use all on the same storage system or on two different storage systems in an active/active configuration. If you use target expansion adapters, then you can only use onboard adapters as initiators.

N5000 series target port configuration recommendations

Table 2-18 shows recommended target ports using single_image cfmode.

The port pairs on a N5000 series storage system that share an ASIC are 0a+0b and 0c+0d.

Table 2-18 Recommended target ports using single_image cfmode

Number of target ports	Ports
1	0d
2	0d, 0b
3	0d, 0b, 0c
4	0d,0 b, 0c, 0a

N5300 and N5600 supported topologies

N5300 and N5600 storage systems are supported by single_image and standby cfmodes.

The N5300 and N5600 storage systems have four onboard 4 Gb FC ports per controller and each port can be configured as either an FC target port or an initiator port. For example, you can configure two ports as SAN targets and two ports as initiators for disk shelves.

N5300 and N5600: Dual-fabric active/active configuration, four onboard FC ports (single_image cfmode)

Figure 2-13 and Table 2-19 show the supported topology and attributes specific to this configuration.

> **Note:** Although this topology appears similar to a topology with standby cfmode support, the following diagram shows the difference in cabling. Ports 0a and 0c connect to switch or fabric 1 and ports 0b and 0d connect to switch or fabric 2.

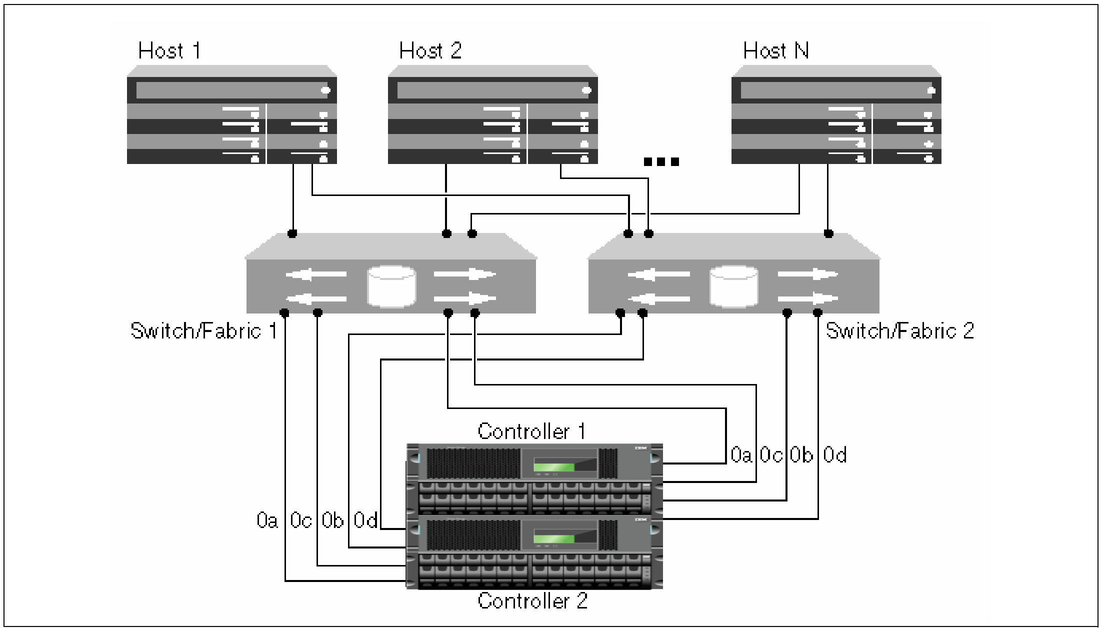


Figure 2-13 N5300 and N5600: Dual-fabric active/active configuration, four onboard FC ports (single_image cfmode)

Table 2-19 N5300 and N5600: Dual-fabric active/active configuration, four onboard FC ports (single_image cfmode)

Attribute	Value
Fully redundant	Yes
Type of fabric	Dual fabric
Different host operating systems	Yes, with multiple host configurations
FC ports or adapters	Four onboard FC ports per controller
Supported cfmodes	Single_image cfmode
Multipathing required	Yes, for full redundancy
Type of configuration	Active/Active configuration

N5300 and N5600: Dual-fabric active/active configuration, four onboard FC ports (single_image and standby cfmodes)

Figure 2-14 and Table 2-20 show the supported topology and attributes specific to this configuration.

Note: Although this topology appears similar to the topology with single_image only cfmode support, there are significant differences in the cabling. Ports 0a and 0b connect to switch or fabric 1 and ports 0c and 0d connect to switch or fabric 2.

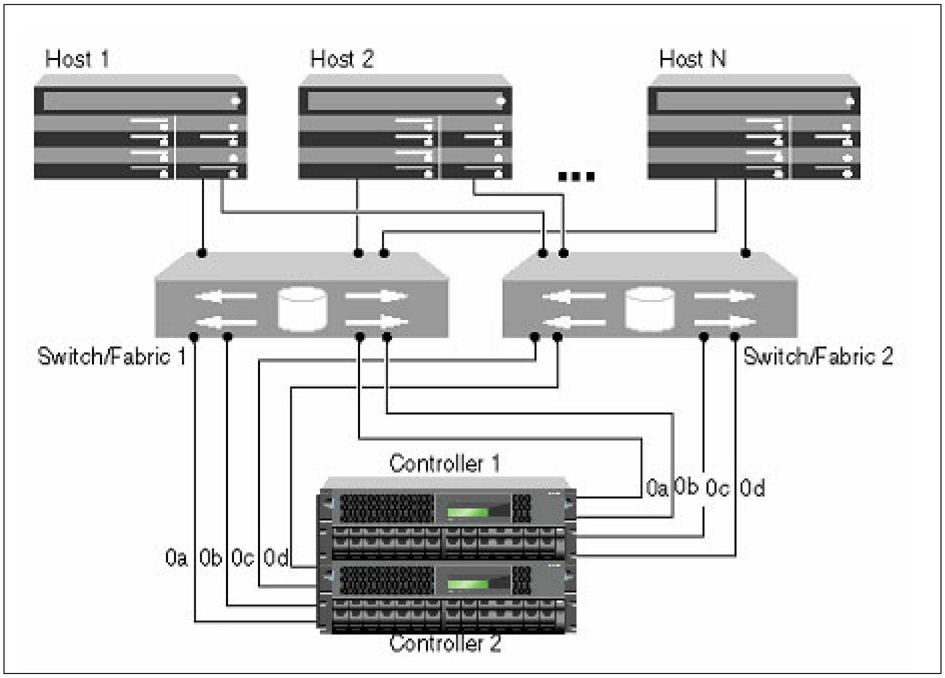


Figure 2-14 N5300 and N5600: Dual-fabric active/active configuration, four onboard FC ports (single_image and standby cfmodes

Table 2-20 N5300 and N5600: Dual-fabric active/active configuration, four onboard FC ports (single_image and standby cfmodes)

Attribute	**Value**
Fully redundant	Yes
Type of fabric	Dual fabric
Different host operating systems	Yes, with multiple host configurations
FC ports or adapters	Four onboard FC ports per controller
Supported cfmodes	Single_image and standby cfmodes
Multipathing required	Yes, for full redundancy
Type of configuration	Active/Active configuration

N5300 and N5600: Dual-fabric active/active configuration, two ports of multiple 4 Gb FC target HBAs (single_image and standby cfmodes)

Figure 2-15 and Table 2-21 show the supported topology and attributes specific to this configuration.

> **Note:** Although this topology is similar to the topology supported only by single_image only cfmode, there are significant differences in the cabling. Ports 2a and 2b connect to switch or fabric 1 and ports 4a and 4b connect to switch or fabric 2.
>
> The 4 Gb FC target HBA port numbers (2a, 2b, 4a, and 4b) are examples. The actual port numbers might vary, depending on the expansion slot in which the FC target HBAs are installed.

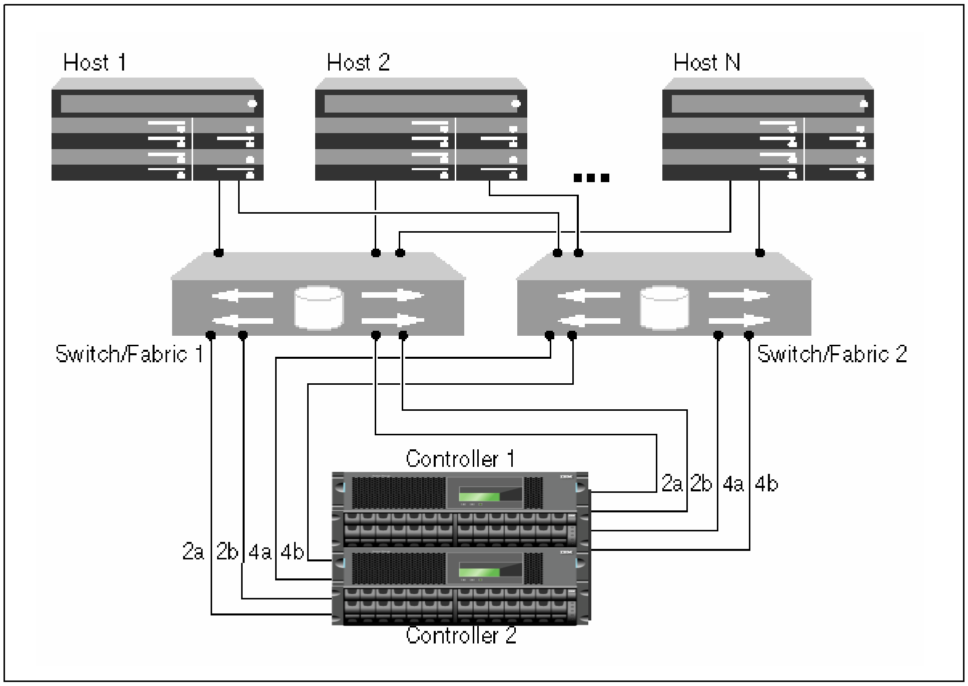


Figure 2-15 N5300 and N5600: Dual-fabric active/active configuration, two ports of multiple 4 Gb FC target HBAs (single_image and standby cfmodes)

Table 2-21 N5300 and N5600: Dual-fabric active/active configuration, two ports of multiple 4 Gb FC target HBAs (single_image and standby cfmodes)

Attribute	Value
Fully redundant	Yes
Type of fabric	Dual fabric
Different host operating systems	Yes, with multiple host configurations
FC ports or adapters	Two ports of multiple 4 Gb FC target HBAs per controller
Supported cfmodes	Single_image and standby cfmodes
Multipathing required	Yes, for full redundancy
Type of configuration	Active/active configuration

N5300 and N5600: Direct-attached active/active configuration, one port of a 4 Gb FC target HBA

Figure 2-16 and Table 2-22 show the supported topology and attributes specific to this configuration.

You might have to set the Data ONTAP FCP adapter media type to loop on the 4 Gb FC target HBAs for some direct-attached topologies.

> **Note:** The 4 Gb FC target HBA port number (2a) is an example. The actual port number might vary depending on the expansion slot in which the FC target HBA is installed.

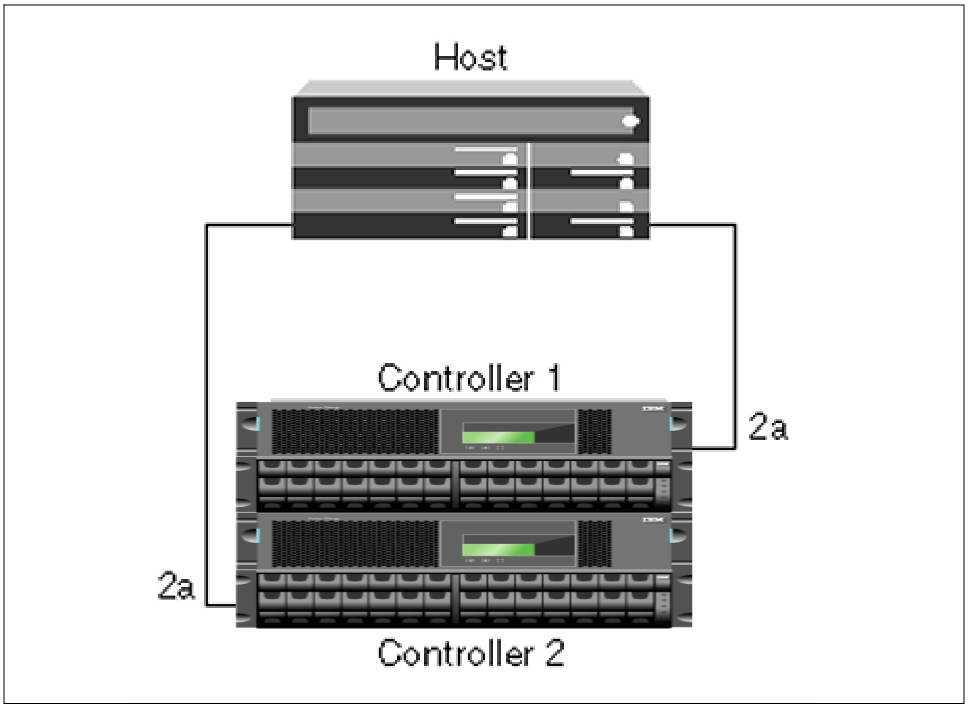


Figure 2-16 N5300 and N5600: Direct-attached active/active configuration, one port of a 4 Gb FC target HBA

Table 2-22 N5300 and N5600: Direct-attached active/active configuration, one port of a 4 Gb FC target HBA

Attribute	Value
Fully redundant	Yes
Type of fabric	None
Different host operating systems	Yes, with multiple host configurations
FC ports or adapters	One port of a 4 Gb FC target HBA per controller
Supported cfmodes	Single_image cfmodes
Multipathing required	Yes, for full redundancy
Type of configuration	Active/active configuration

N5300 and N5600: Single-fabric active/active configuration, two onboard FC ports

Figure 2-17 and Table 2-23 show the supported topology and attributes specific to this configuration.

Each controller can have one or up to the maximum number of target ports supported per controller connecting into the fabrics, while an active/active configuration can have two or up to the maximum number of target ports supported per configuration connecting into the fabric. See the system configuration guide for the version of Data ONTAP being used by the controllers for these values.

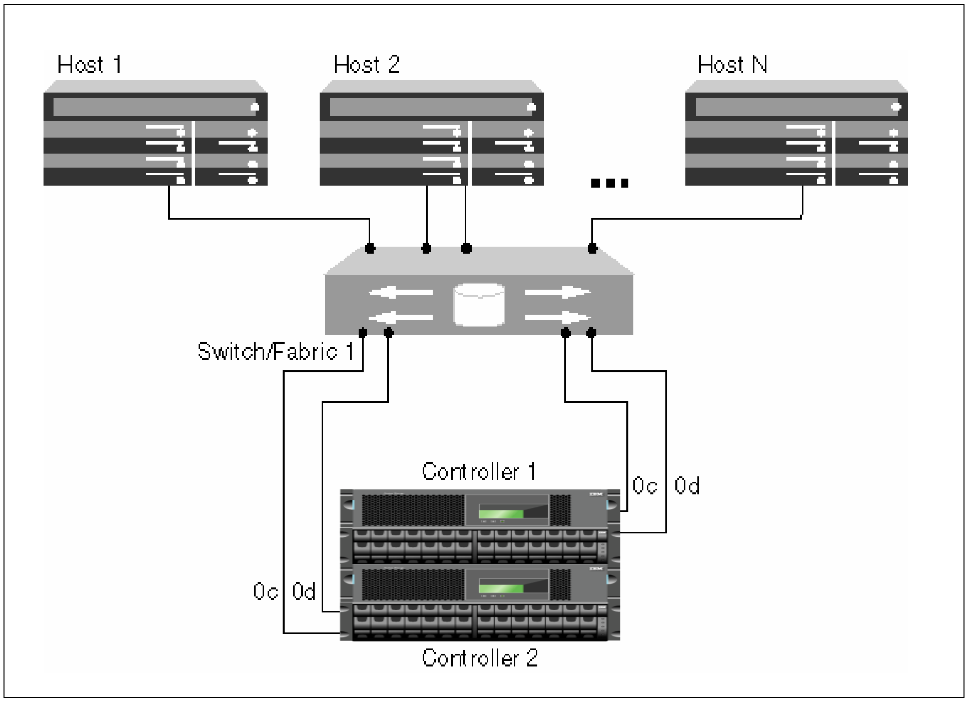


Figure 2-17 N5300 and N5600: Single-fabric active/active configuration, two onboard FC ports

Table 2-23 N5300 and N5600: Single-fabric active/active configuration, two onboard FC ports

Attribute	**Value**
Fully redundant	No, due to the single fabric
Type of fabric	Single fabric
Different host operating systems	Yes, with multiple host configurations
FC ports or adapters	Two onboard FC ports per controller
Supported cfmodes	Single_image and standby cfmodes
Multipathing required	Yes, if a host has multiple paths to a single LUN No, if in standby cfmode and with single-attached hosts
Type of configuration	Active/active configuration

N5200 and N5500 supported topologies

N5200 and N5500 storage systems are supported by single_image, partner, and standby cfmodes.

The N5200 and N5500 storage systems have four onboard 2 Gb FC ports per controller and each port can be configured as either an FC target port or an initiator port.

2 Gb FC target ports are supported with the onboard 2 Gb FC ports on the N5200 and N5500 storage systems. 4 Gb FC target connections are supported with 4 Gb FC target HBAs.

Each N5000 series storage system supports 2 Gb or 4 Gb FC target HBAs, but you cannot use both on the same storage system or on two different storage systems in an active/active configuration. If you use target expansion HBAs, then you can only use onboard ports as initiators.

N5200 and N5500: Dual-fabric active/active configuration, four onboard FC ports (single_image and partner cfmodes)

Figure 2-18 and Table 2-24 show the supported topology and attributes specific to this configuration.

Note: Although this topology is similar to the topology supported by single_image, standby, and partner cfmodes, there are significant differences in the cabling. Ports 0a and 0c connect to switch or fabric 1 and ports 0b and 0d connect to switch or fabric 2.

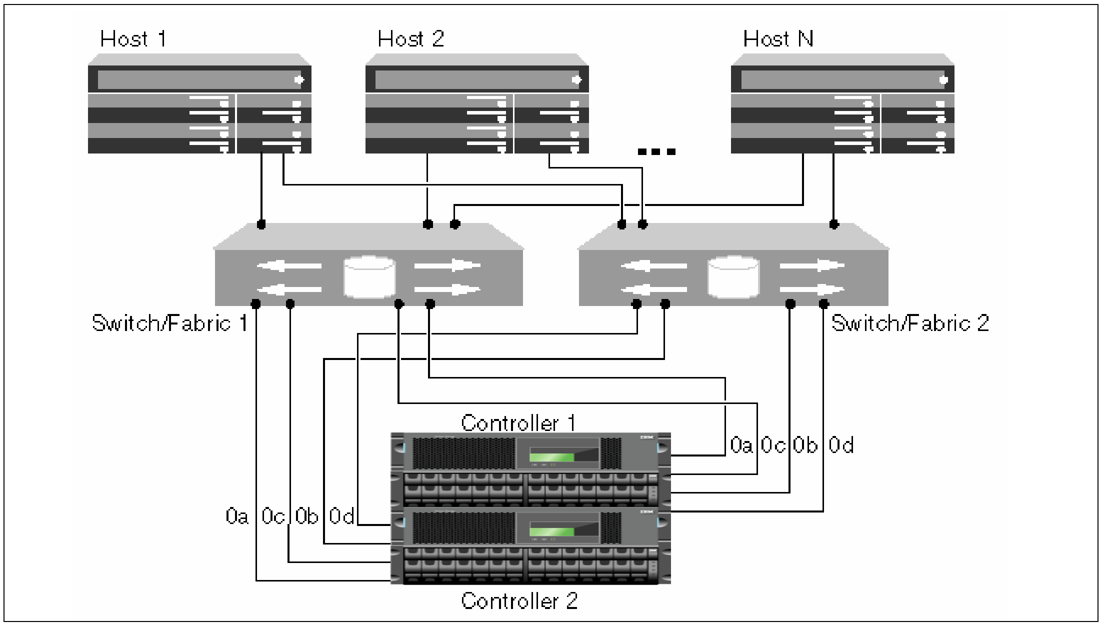


Figure 2-18 N5200 and N5500: Dual-fabric active/active configuration, four onboard FC ports (single_image and partner cfmodes)

Table 2-24 N5200 and N5500: Dual-fabric active/active configuration, four onboard FC ports (single_image and partner cfmodes)

Attribute	Value
Fully redundant	Yes
Type of fabric	Dual fabric
Different host operating systems	Yes, with multiple host configurations
FC ports or adapters	Four onboard FC ports per controller
Supported cfmodes	Single_image and partner cfmodes
Multipathing required	Yes, for full redundancy
Type of configuration	Active/active configuration

N5200 and N5500: Dual-fabric active/active configuration, two 4 Gb FC target HBAs (single_image cfmode)

Figure 2-19 and Table 2-25 show the supported topology and attributes specific to this configuration.

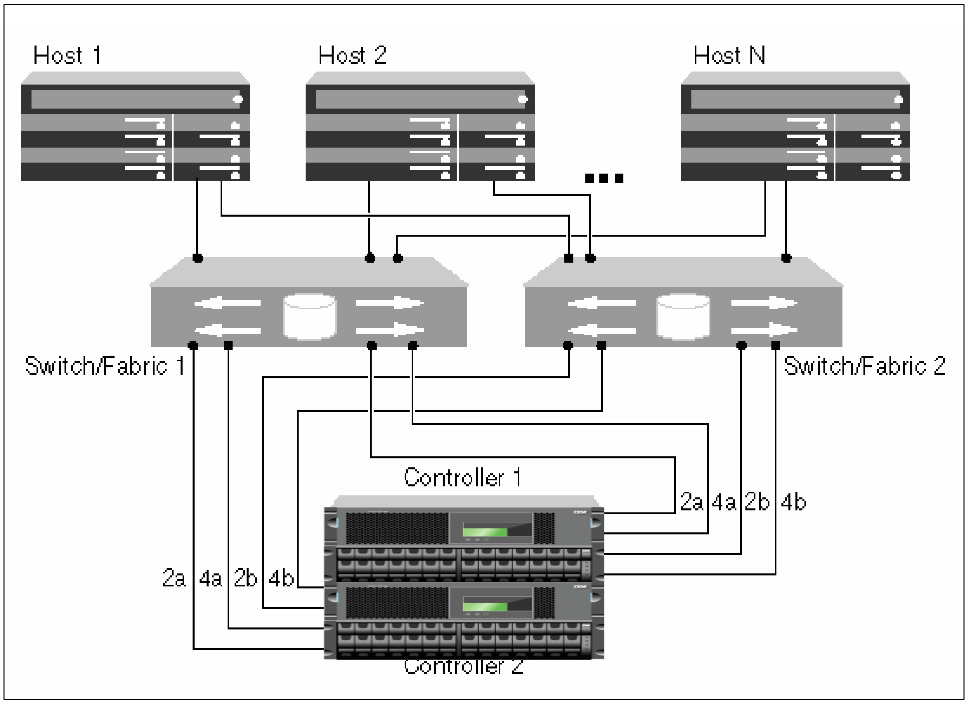


Figure 2-19 N5200 and N5500: Dual-fabric active/active configuration, two 4 Gb FC target HBAs (single_image cfmode)

> **Note:** Although this topology is similar to the target topology supported by standby cfmode, there are significant differences in the cabling of this configuration. Ports 2a and 4a connect to switch or fabric 1 and ports 2b and 4b connect to switch or fabric 2.
>
> The 4 Gb FC target HBA port numbers (2a, 2b, 4a, and 4b) are examples. The actual port numbers might vary, depending on the expansion slot in which the FC target HBAs are installed.

Table 2-25 N5200 and N5500: Dual-fabric active/active configuration, two 4 Gb FC target HBAs (single_image cfmode)

Attribute	Value
Fully redundant	Yes
Type of fabric	Dual fabric
Different host operating systems	Yes, with multiple host configurations
FC ports or adapters	Two 4 Gb FC target HBAs per controller
Supported cfmodes	Single_image cfmodes
Multipathing required	Yes, for full redundancy
Type of configuration	Active/active configuration

N5200 and N5500: Dual-fabric active/active configuration, two 4-Gb FC target HBAs (standby cfmode)

Figure 2-20 and Table 2-26 show the supported topology and attributes specific to this configuration.

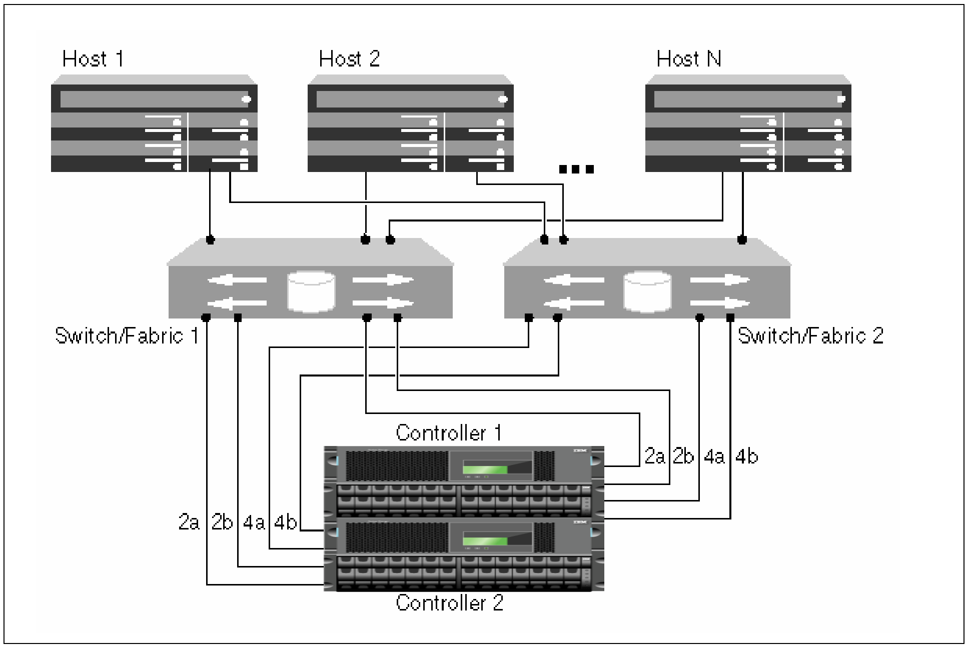


Figure 2-20 N5200 and N5500: Dual-fabric active/active configuration, two 4 Gb FC target HBAs (standby cfmode)

Note: Although this topology is similar to the target topology supported by single_image only cfmode, there are significant differences in the cabling of this configuration. Ports 2a and 2b connect to switch or fabric 1 and ports 4a and 4b connect to switch or fabric 2.

The 4 Gb FC target HBA port numbers (2a, 2b, 4a, and 4b) are examples. The actual port numbers might vary, depending on the expansion slot in which the FC target HBAs are installed.

Table 2-26 N5200 and N5500: Dual-fabric active/active configuration, two 4 Gb FC target HBAs (standby cfmode)

Attribute	Value
Fully redundant	Yes
Type of fabric	Dual fabric
Different host operating systems	Yes, with multiple host configurations
FC ports or adapters	Two 4 Gb FC target HBAs per controller
Supported cfmodes	Single_image and standby cfmodes
Multipathing required	Yes, for full redundancy
Type of configuration	Active/active configuration

N5200 and N5500: Direct-attached active/active configuration, one onboard FC port

Figure 2-21 and Table 2-27 show the supported topology and attributes specific to this configuration.

You might have to set the Data ONTAP FCP adapter media type to loop on the 4 Gb FC target HBAs for some direct-attached topologies.

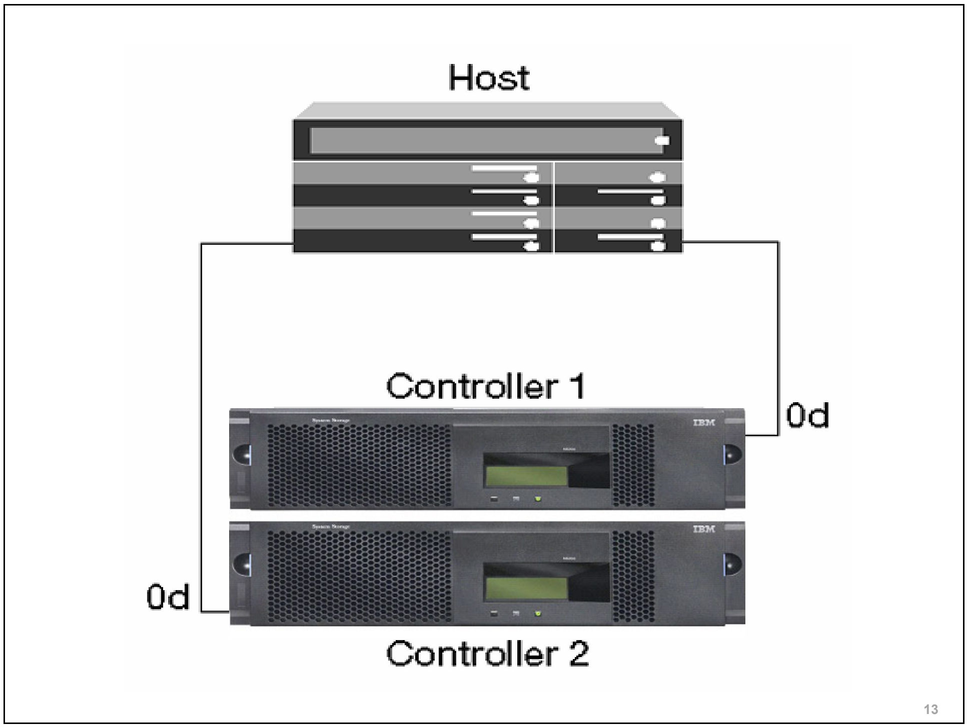


Figure 2-21 N5200 and N5500: Direct-attached active/active configuration, one onboard FC port

Table 2-27 N5200 and N5500: Direct-attached active/active configuration, one onboard FC port

Attribute	Value
Fully redundant	Yes
Type of fabric	None
Different host operating systems	Yes, with multiple host configurations
FC ports or adapters	One onboard FC port per controller
Supported cfmodes	Single_image cfmodes
Multipathing required	Yes, if a host has multiple paths to a LUN
Type of configuration	Active/active configuration

2.3 Managing an active/active configuration

This section describes the considerations and activities related to managing an active/active configuration.

2.3.1 Managing an active/active configuration in normal mode

At a high level, the tasks involved in managing an active/active configuration in normal mode are:

- Monitoring active/active configuration status
- Viewing information about the active/active configuration
  - Displaying the partner's name
  - Displaying disk information
- Enabling and disabling takeover
- Enabling and disabling immediate takeover of a panicked partner
- Halting a node without takeover
- Performing a takeover

For detailed information about managing an active/active configuration in normal mode refer to the *IBM System Storage N series Data ONTAP 7.3 Active/Active Configuration Guide* at:

http://www-01.ibm.com/support/docview.wss?uid=ssg1S7002467

2.3.2 Halting a node without takeover

You can halt the node and prevent its partner from taking over. For example, you might need to perform maintenance on both the storage system and its disks and want to avoid an attempt by the partner node to write to those disks.

To do this, enter the following command:

```
halt -f
```

The synopsis for the **`halt`** command is:

```
halt [-d] [-t interval] [-f]
```

`-d`	The storage system performs a core dump before halting.
`-t interval`	The storage system halts after the number of minutes specified by interval.
`-f`	Prevents one partner in a clustered storage system pair from taking over the other after the storage system halts.

2.3.3 Basic active/active configuration management

This section discusses simple active/active configuration management, including forced active/active configuration takeover.

Takeover

There are two ways to take over resources from one node to another node: using the command line or the GUI.

> **Important:** Taking over resources affects the client environment. In particular, Windows users and shares (CIFS services) are affected by this procedure.

Note: Issue the **cf takeover** command on the node that will remain operating and will take over resources of the other node. In our example, we took node itsosj_n2 offline by issuing cf takeover on node itsosj_n1.

Use the following steps if working from a command line:

1. Check the active/active status with the **cf status** command (Example 2-5).

Example 2-5 cf status: Check status

```
itsosj-n1> cf status
Cluster enabled, itsosj-n2 is up.
itsosj-n1>
```

2. Issue the **cf takeover** command as shown in Example 2-6.

Example 2-6 cf takeover command

```
itsosj-n1> cf takeover
cf: takeover initiated by operator
itsosj-n1> Wed Sep 14 00:07:11 CEST [itsosj-n1: cf.misc.operatorTakeover:warning]:
Cluster monitor: takeover initiated by operator
Wed Sep 14 00:07:11 CEST [itsosj-n1: cf.fsm.nfo.acceptTakeoverReq:warning]: Negotiated
failover: accepting takeover request by partner, reason: operator initiated cf takeover.
Asking partner to shutdown gracefully; will takeover in at most 180 seconds.
Wed Sep 14 00:07:14 CEST [itsosj-n1: cf.fsm.firmwareStatus:info]: Cluster monitor:
partner rebooting
Wed Sep 14 00:07:14 CEST [itsosj-n1: cf.fsm.nfo.partnerShutdown:warning]: Negotiated
failover: partner has shutdown
Wed Sep 14 00:07:14 CEST [itsosj-n1: cf.fsm.takeover.nfo:info]: Cluster monitor:
takeover attempted after cf takeover command
Wed Sep 14 00:07:14 CEST [itsosj-n1: cf.fsm.stateTransit:warning]: Cluster monitor: UP
--> TAKEOVER
Wed Sep 14 00:07:14 CEST [itsosj-n1: cf.fm.takeoverStarted:warning]: Cluster monitor:
takeover started
Wed Sep 14 00:07:16 CEST [itsosj-n1: cf_takeover:info]: NVRAM takeover: partner nvram is
disabled
Wed Sep 14 00:07:17 CEST [itsosj-n2/itsosj-n1: wafl.vol.loading:debug]: Loading Volume
partner:vol_itsosj01
Wed Sep 14 00:07:17 CEST [itsosj-n2/itsosj-n1: wafl.vol.loading:debug]: Loading Volume
partner:vol0
Wed Sep 14 00:07:18 CEST [itsosj-n2/itsosj-n1: wafl.maxdirsize.boot.notice:warning]:
partner:vol0: This volume's maxdirsize (2621KB) is higher than the default (1310KB).
There may be a performance penalty when doing operations on large directories.
Replaying takeover WAFL log
Wed Sep 14 00:07:18 CEST [itsosj-n2/itsosj-n1: wafl.vol.guarantee.fail:error]: Space for
volume vol_itsosj01 is NOT guaranteed
Wed Sep 14 00:07:18 CEST [itsosj-n2/itsosj-n1: wafl.takeover.nvram.missing:error]: WAFL
takeover: no partner area found during wafl replay
Wed Sep 14 00:07:18 CEST [itsosj-n2/itsosj-n1: wafl.replay.done:info]: WAFL log replay
completed, 0 seconds
Wed Step 14 00:07:20 CEST [itsosj-n2/itsosj-n1: cf_takeover:ALERT]: Language not set on
volume vol0. Using lang config "C". Use vol lang to set language.
ifconfig: 'ns0' cannot be configured: Address does not match any partner interface.
ifconfig: ns0: no such interface
add net default: gateway 9.1.37.65: network unreachable
Wed Sep 14 00:07:20 CEST [itsosj-n2/itsosj-n1: net.ifconfig.noPartner:error]: ifconfig:
'ns0' cannot be configured: Address does not match any partner interface.
```

```
Wed Sep 14 00:07:23 CEST [itsosj-n1: net.ifconfig.takeoverError:warning]: WARNING: 1
error detected during network takeover processing WARNING: Some network clients may not
be able to access the cluster during takeover
Wed Sep 14 00:07:23 CEST [itsosj-n1: cf.rsrc.takeoverOpFail:error]: Cluster monitor:
takeover during ifconfig_2 failed; takeover continuing...
CIFS partner server is running.
Wed Sep 14 00:07:23 CEST [itsosj-n1 (takeover): cf.rsrc.transitTime:notice]: Top
Takeover transit times wafl=2240, registry_postrc_phase1=1630, rc=990 {options=570,
hostname=160, options=10}, raid=720, wafl_sync=500, ifconfig=240, registry_prerc=220,
raid_replay=180, sshd=160, syslog=120
Wed Sep 14 00:07:23 CEST [itsosj-n1 (takeover): cf.fm.takeoverComplete:warning]: Cluster
monitor: takeover completed
Wed Sep 14 00:07:33 CEST [itsosj-n2/itsosj-n1: nbt.nbns.registrationComplete:info]: NBT:
All CIFS name registrations have completed for the partner server.

itsosj-n1(takeover)> Wed Sep 14 00:08:01 CEST [itsosj-n1 (takeover):
monitor.globalStatus.critical:CRITICAL]: This node has taken over itsosj-n2. Disk on
adapter v0, shelf 2, bay 6, failed by administrator. Disk on adapter v0, shelf 2, bay 4,
failed by administrator. Disk on adapter v0, shelf 1, bay 3, failed by administrator.
Wed Sep 14 00:08:01 CEST [itsosj-n2/itsosj-n1: monitor.globalStatus.critical:CRITICAL]:
itsosj-n1 has taken over this node.

itsosj-n1(takeover)>
```

3. Check the status, as shown in Example 2-7.

Example 2-7 cf status: verification if takeover completed

```
itsosj-n1 has taken over itsosj-n2.
Takeover due to negotiated failover, reason: operator initiated cf takeover
itsosj-n1(takeover)>
```

As illustrated in Figure 2-22, you can perform the takeover using FilerView and selecting **Cluster** → **Manage** → **Initiate Takeove**r. Click **OK** when the FilerView Alert! appears.

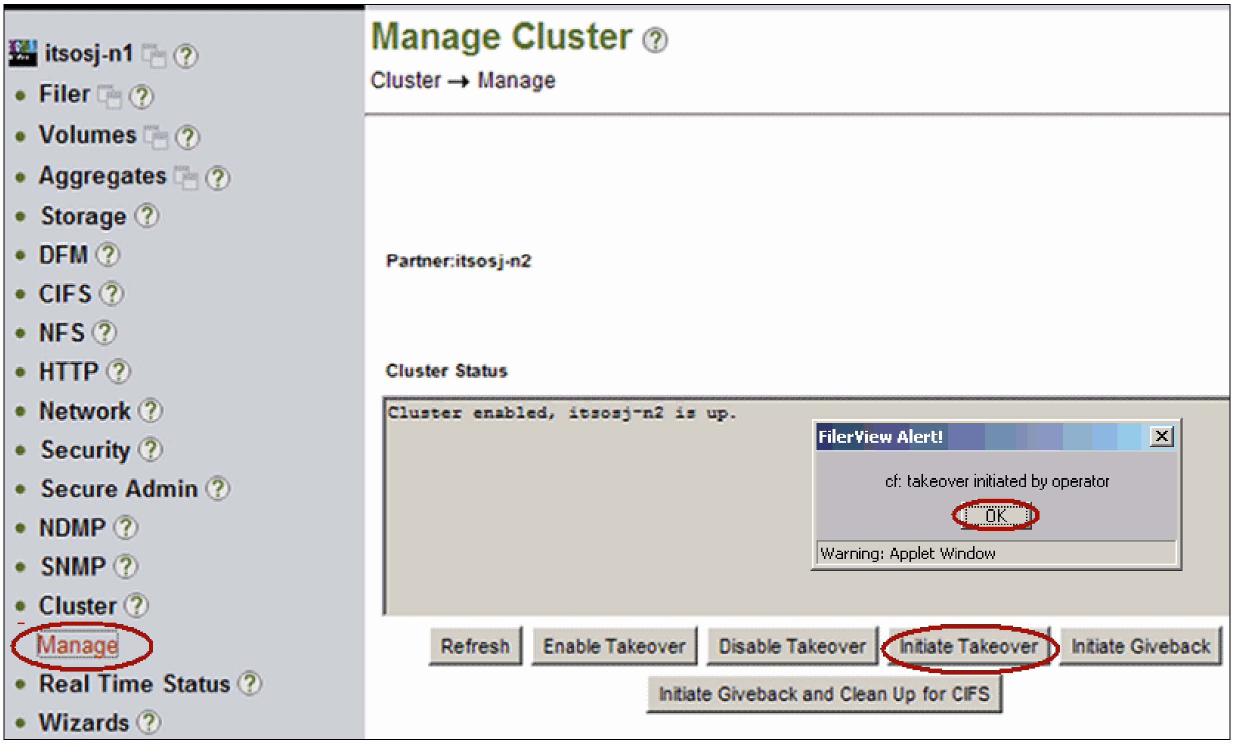


Figure 2-22 Active/active configuration takeover initiated by FilerView

Click **Refresh** to see the most current cluster status. The result is shown in Figure 2-23.

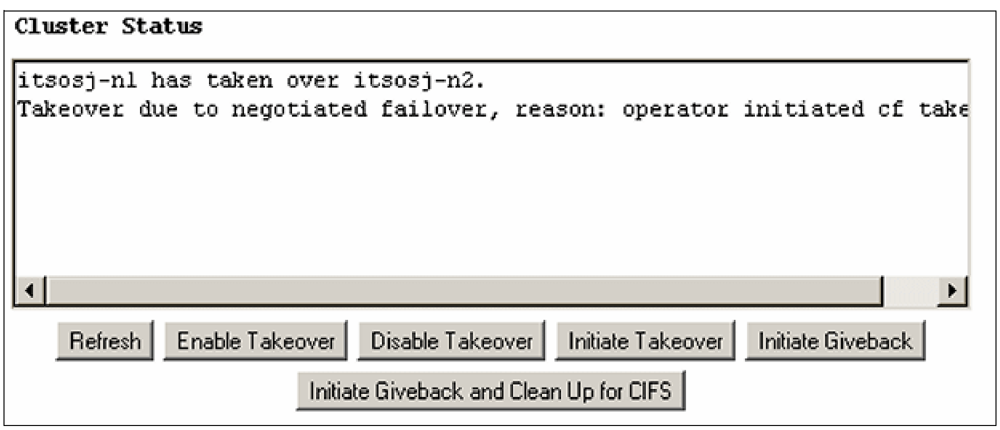


Figure 2-23 FilerView: Cluster status

Giveback

To give back resources, issue the **cf giveback** command as shown in Example 2-8.

Example 2-8 cf giveback

```
itsosj-n1(takeover)> cf giveback
Wed Sep 14 00:10:15 CEST [itsosj-n1 (takeover): cf.misc.operatorGiveback:info]: Cluster
monitor: giveback initiated by operator
Wed Sep 14 00:10:15 CEST [itsosj-n1: cf.fm.givebackStarted:warning]: Cluster monitor:
giveback started
itsosj-n1(takeover)>
CIFS partner server is shutting down...

CIFS partner server has shut down...
Wed Sep 14 00:10:17 CEST [itsosj-n1: cf.rsrc.transitTime:notice]: Top Giveback transit
times wafl=1580, ndmpd=560, raid=80, registry_giveback=20, java=10, snapmirror=10, nfsd=10,
fmfsm_reserve=0, fmdisk_inventory=0, raid_disaster_early=0
Wed Sep 14 00:10:17 CEST [itsosj-n1: cf.fm.givebackComplete:warning]: Cluster monitor:
giveback completed
Wed Sep 14 00:10:17 CEST [itsosj-n1: cf.fsm.stateTransit:warning]: Cluster monitor:
TAKEOVER --> UP
Wed Sep 14 00:10:18 CEST [itsosj-n1: cf.fsm.takeoverByPartnerDisabled:notice]: Cluster
monitor: takeover of itsosj-n1 by itsosj-n2 disabled (unsynchronized log)
Wed Sep 14 00:10:18 CEST [itsosj-n1: cf.fsm.firmwareStatus:info]: Cluster monitor: partner
rebooting
Wed Sep 14 00:10:26 CEST [itsosj-n1: cf.fsm.partnerNotResponding:notice]: Cluster monitor:
partner not responding
Wed Sep 14 00:10:29 CEST [itsosj-n1: cf.fm.timeMasterStatus:info]: Acting as cluster time
slave
Wed Sep 14 00:10:30 CEST [itsosj-n1: cf.fsm.partnerOk:notice]: Cluster monitor: partner ok
Wed Sep 14 00:10:30 CEST [itsosj-n1: cf.fsm.takeoverOfPartnerDisabled:notice]: Cluster
monitor: takeover of itsosj-n2 disabled (partner booting)
Wed Sep 14 00:10:33 CEST [itsosj-n1: cf.fsm.takeoverOfPartnerDisabled:notice]: Cluster
monitor: takeover of itsosj-n2 disabled (unsynchronized log)
Wed Sep 14 00:10:35 CEST [itsosj-n1: cf.fsm.takeoverOfPartnerEnabled:notice]: Cluster
monitor: takeover of itsosj-n2 enabled
Wed Sep 14 00:10:35 CEST [itsosj-n1: cf.fsm.takeoverByPartnerEnabled:notice]: Cluster
monitor: takeover of itsosj-n1 by itsosj-n2 enabled

itsosj-n1>
```

You can check the cluster status by issuing the **cf status** command, as shown in Example 2-9.

Example 2-9 cf status: check for successful giveback

```
itsosj-n1> cf status
Cluster enabled, itsosj-n2 is up.
itsosj-n1>
```

Figure 2-24 illustrates how to perform the active/active configuration giveback using FilerView.

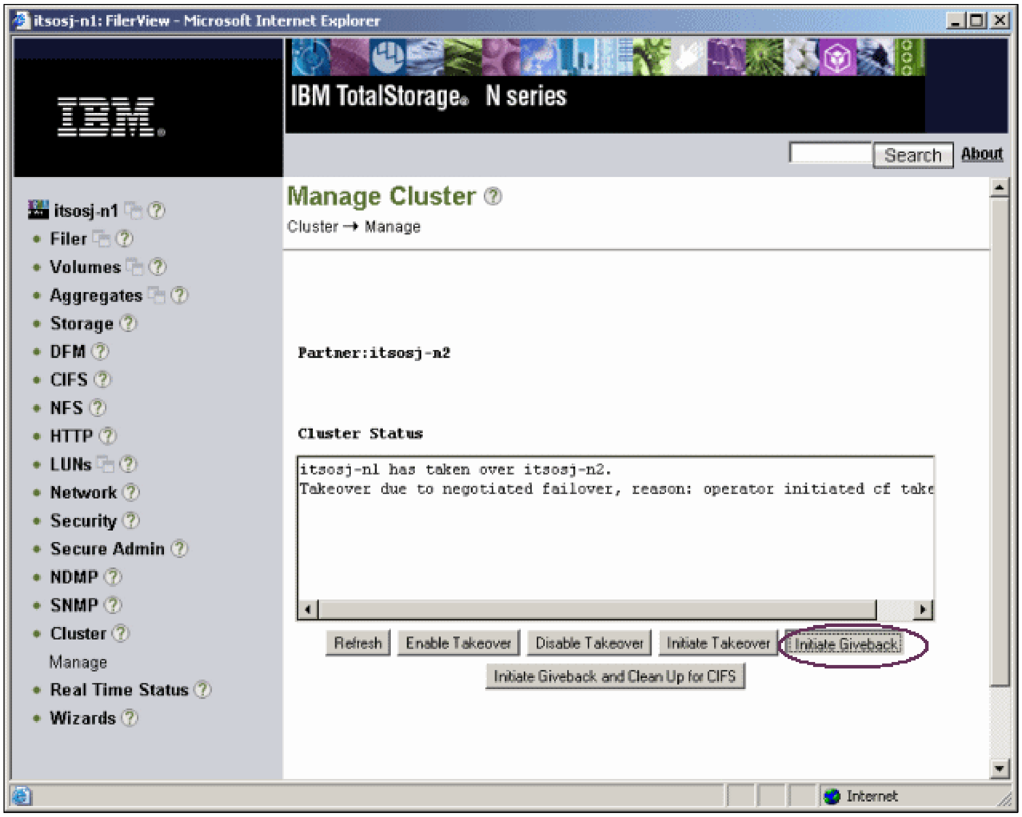


Figure 2-24 FilerView: initiate giveback

After refreshing the view, the active/active configuration status should be `Cluster enabled`. As shown in Figure 2-25, itsosj-n2 is up.

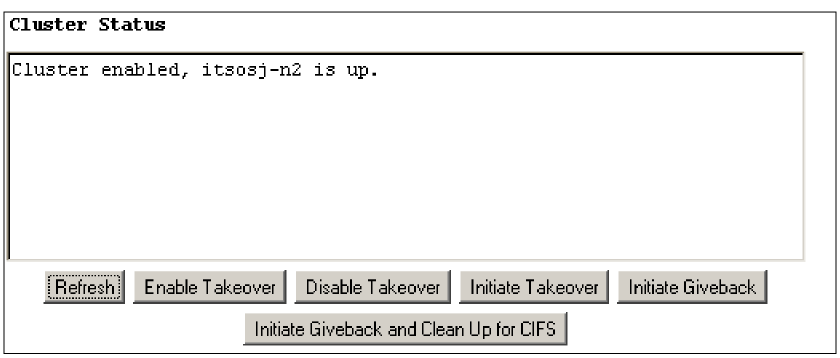


Figure 2-25 FilerView: status after refresh

2.3.4 Active/active configuration failover basic operations

When a failover occurs, the running partner node in the active/active configuration takes over the functions and disk drives of the failed node by creating an emulated storage system that performs the following tasks:

- Assumes the identity of the failed node.
- Accesses the failed nodefs disks and serves its data to clients.
- The partner node maintains its own identity and its own primary functions, but also handles the added functionality of the failed node through the emulated node.

> **Note:** When a failover occurs, existing CIFS sessions are terminated. A graceful shutdown of the CIFS sessions is not possible, and some data transfers could be interrupted.

2.3.5 Connectivity during failover

Both front-end and back-end operations are affected during a failover. On the front end are the IP addresses, used as well as the host name. On the back end there is the connectivity and addressing to the disk subsystem. Both the back-end and front-end interfaces must be configured correctly for a successful failover.

2.3.6 Reasons for active/active configuration failover

The conditions under which takeovers occur depend on how you configure the active/active configuration. Takeovers can be initiated when one of the following conditions occurs:

- An active/active node that is configured for immediate takeover on panic undergoes a software or system failure that leads to a panic.
- A node that is in an active/active configuration undergoes a system failure (for example, NVRAM failure) and cannot reboot.

  > **Note:** If the storage for a node also loses power at the same time, a standard takeover is not possible.

- There is a mismatch between the disks that one node can see and the disks that the other node can see.
- One or more network interfaces that are configured to support failover becomes unavailable.
- A node cannot send heartbeat messages to its partner. This could happen if the node experienced a hardware failure or software failure that did not result in a panic, but still prevented it from functioning correctly (such as a failure in the interconnect cable).
- You halt one of the active/active nodes without using the -f flag. The -f flag applies only to storage systems in an active/active configuration. If you enter the `halt -f` command on an N series, its partner does not take over.
- You initiate a takeover manually.

2.3.7 Failover due to disk mismatch

When communication between active/active nodes is first established through the active/active configuration interconnect adapters, the nodes exchange a list of disk shelves that are visible on the A loop and B loop of each node. If a system later sees that the B loop shelf count on its partner is greater than its local A loop shelf count, the system concludes that it is impaired and prompts its partner to initiate a takeover.

2.3.8 Checking active/active configuration status

You can check active/active configuration status by using the **`cf status`** and **`cf monitor`** commands, as illustrated in Example 2-10.

Example 2-10 The cf status and cf monitor commands

```
itsosj-n1> cf status
Cluster enabled, itsosj-n2 is up.
itsosj-n1>
itsosj-n1> cf monitor
  current time: 14Sep2005 00:04:05
  UP 02:49:33, partner 'itsosj-n2', cluster monitor enabled
  Interconnect is up, takeover capability on-line
  partner update TAKEOVER_ENABLED (14Sep2005 00:04:04)
itsosj-n1>
```

You can also check cluster status using FilerView **Cluster** → **Manage** (Figure 2-26).

Figure 2-26 FilerView: Manage Cluster window

2.4 Active/active configuration failover mode (cfmode) for Fibre Channel

Data ONTAP provides multiple modes of operation, called cfmodes, which are required to support homogeneous and heterogeneous host operating systems. This section provides an overview of each cfmode setting and describes how to change the default cfmode to satisfy the requirements of your configuration.

Cluster failover mode, or cfmode, is functionality within Data ONTAP that defines how Fibre Channel ports behave during failover in an active/active configuration. Selecting the correct cfmode is critical to ensuring that your LUNs are accessible and optimizing your storage system's performance in the event of a failover.

The FCP cfmode setting controls how the target ports perform the following tasks:

- Log into the fabric.
- Handle local and partner traffic for an active/active configuration in normal operation and in takeover.
- Provide access to local and partner LUNs in an active/active configuration.

> **Important:** The cfmode setting of your storage system and the number of paths available to the storage system must align with cabling, configuration limits, and zoning requirements.
>
> A cfmode setting is not available on single-controller configurations. You can change the cfmode setting from the storage system console by setting privileges to “Advanced” and then using the **`fcp set`** command.

Detailed descriptions of port behavior with each cfmode are available in the *IBM System Storage N series Data ONTAP 7.3 Block Access Management Guide for iSCSI and FCP* at:

http://www-01.ibm.com/support/docview.wss?uid=ssg1S7002547

There are five cfmodes:

- Partner mode: Both ports A and B are active.
- Standby mode: Port A is active and port B is in standby.
- Mixed mode: Supported, not recommended (virtual ports).
- Dual_fabric mode: N3700 only (virtual ports).
- Single_image cfmode: Allows all ports to host LUNs from both storage systems, and it makes the filer act as a single FCP device server with only one FCP nodename.

2.4.1 Summary of cfmode settings and supported systems

Table 2-28 summarizes the cfmodes, supported systems, benefits, and limitations.

Table 2-28 Cfmodes characteristics

cfmode	Supported systems	Benefits and limitations
Partner	All systems except N3700, N7000 series, and systems with 4 Gb adapters	Supports all host OS types. Supports all switches.
Single-image	All systems	Supports all host OS types. Supports all switches. Makes all LUNs available on all target ports.
Dual_fabric	N3700 only	Supports all host OS types. Requires fewer switch ports. Does not support all switches. Requires switches that support public loop.
Standby	All systems except N3700	Requires more switch ports. Supports only Windows and Solaris hosts.
Mixed	All systems except IBM System Storage N series N3700 (2863-A20) storage system and N7000 series	Supports all host OS types. Does not support all switches. Requires switches that support public loop.

2.4.2 Cfmode restrictions for IBM System Storage N series Data ONTAP 7.3 Filer and Gateway

There are a number of restrictions to consider when deciding which cfmode to implement. They are identified in the following list.

- The cfmode settings must be set to the same value for both nodes in an active/active configuration. If the cfmode settings are not identical, your hosts might not be able to access data stored on the system.
- Upgrading to a 4 Gb HBA in a SAN environment: If you attempt to run a N5000, N6000, or N7000 series storage system with a 4 Gb HBA in an unsupported cfmode, the 4 Gb HBA is set to offline and an error message is displayed.

  You must change the cfmode to single_image or standby cfmode before upgrading. Only single_image and standby cfmodes are supported with the new 4 Gb Fibre Channel host bus adapters (HBAs) on N5000, N6000, and N7000 series storage systems.

  In addition, Data ONTAP does not allow changing from a supported cfmode to an unsupported cfmode with the 4 Gb HBA installed on these systems.

- Single_image is the only supported cfmode for new installations starting with the Data ONTAP 7.3 release.

  If you are upgrading to Data ONTAP 7.3 from a prior release, follow these guidelines:

  – On existing systems, including N3700 storage systems, you can continue to use other cfmodes that are supported on your systems. You can freely change from one supported cfmode to any other supported cfmode on these systems. See "How to manage FCP with Active/Active configurations" in the *IBM System Storage N series Data ONTAP 7.3 Block Access Management Guide for iSCSI and FCP* for the list of supported cfmodes by storage system model, at:

    `http://www-01.ibm.com/support/docview.wss?uid=ssg1S7002473&rs=555`

  – On N3300 and N3600 storage systems, and N5000, N6000, and N7000 series systems, you can continue to run the existing cfmode after upgrading. If you change to single_image cfmode, you cannot revert to other cfmodes. See "IBM System Storage N series Changing the Cluster cfmode Setting in Fibre Channel SAN Configurations" for instructions on upgrading to single_image cfmode, at:

    `http://www-01.ibm.com/support/docview.wss?uid=ssg1S7001382`

Note: Data ONTAP 8.0 supports only single_image cfmode.

Data ONTAP Version 7.2 and earlier supports the following cfmodes:

- Standby
- Partner
- Mixed
- dual_fabric

2.4.3 Setting cfmode

The cfmode setting applies to the entire storage system and requires that the FCP service be stopped to change it. When an igroup is created, the default settings are determined by the cfmode of the storage system.

The cfmode selected determines the multipathing host software requirements for all supported host operating systems. It is subject to the following conditions:

- The cfmode setting sets the target driver mode for the entire storage system.
- The FCP service must be stopped first.
- Each active/active head must have the same setting. Different settings on each head prevents transparent takeover.
- When an igroup is created, the multipath requirement is determined by the cfmode.

Executing fcp setup

This script helps determine the correct cfmode setting for the storage system:

```
Executing fcp set cfmode [ dual_fabric | mixed | partner | standby ]
```

The command shown in Example 2-11 on page 96 requires priv set advanced mode and allows you to set the cfmode manually.

Example 2-11 The priv set advanced command

```
itsotuc2> priv set advanced
Warning: These advanced commands are potentially dangerous; use
         them only when directed to do so by IBM personnel.
itsotuc2*>
```

2.4.4 Single_image cfmode

The single_image cfmode setting is available starting with Data ONTAP 7.1 and is the default setting starting with Data ONTAP Version 7.2. In single_image mode, an active/active configuration has a single global WWNN, and both systems in the active/active configuration function as a single Fibre Channel node. Each node in the active/active configuration shares the partner node's LUN map information.

All LUNs in the active/active configuration are available on all ports in the active/active configuration by default. As a result, there are more paths to LUNs stored on the active/active configuration, because any port on each node can provide access to both local and partner LUNs. You can specify the LUNs available on a subset of ports by defining port sets and binding them to an igroup. Any host in the igroup can access the LUNs only by connecting to the target ports in the port set.

Figure 2-27 shows an example configuration with a multi-attached host. If the host accesses lun_1 through ports 4a, 4b, 5a, or 5b on NodeA, then NodeA recognizes that lun_1 is a local LUN. If the host accesses lun_1 through any of the ports on NodeB, lun_1 is recognized as a partner LUN and NodeB sends the SCSI requests to NodeA over the active/active configuration interconnect.

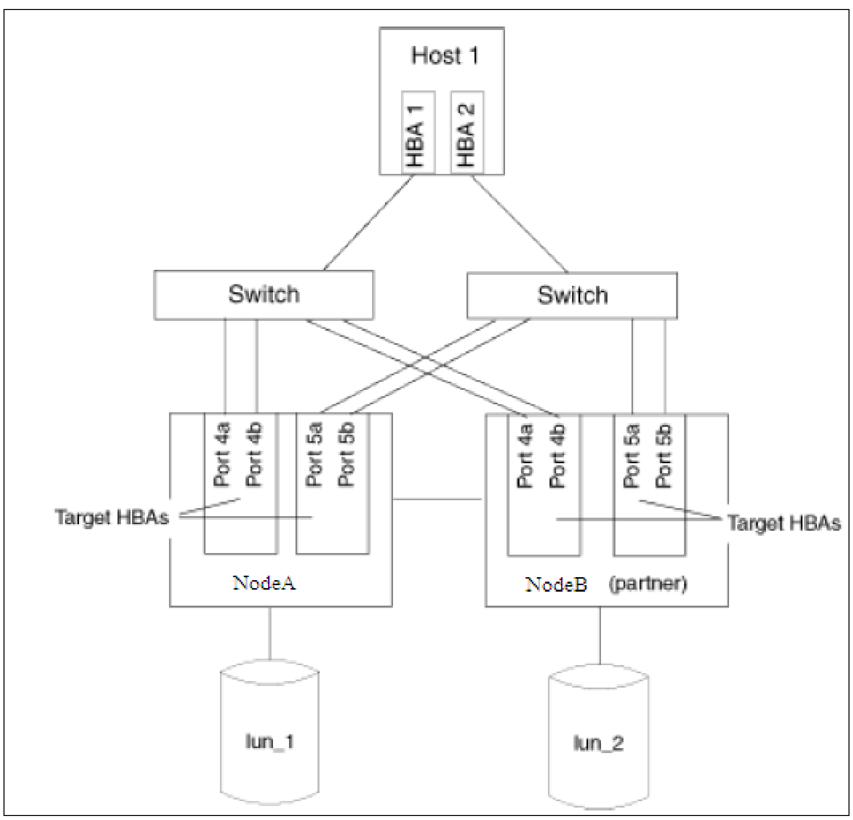


Figure 2-27 Single_image cfmode in an active/active configuration

This configuration is possible only with an active/active storage controller configuration. When properly configured with multipathing software, dual-attached hosts are fully redundant from the storage perspective because the HBA, wiring, fabrics, storage controller, and disks are all redundantly configured.

For multiple-host configurations, the hosts can be heterogeneous (that is, Windows and UNIX).

Features

Note the following points regarding Single_image cfmode:

- It supports all storage systems and switches.
- It makes efficient use of FC target ports.
- All LUNS are visible on all ports.

Figure 2-28 illustrates SSI with node failure. One node of an active/active N6000 configuration failed. Now only two paths are available to the shelves (expansion units). Controller 2 has taken over the drives from controller 1. As a result, only two paths remain to the remaining controller and shelves.

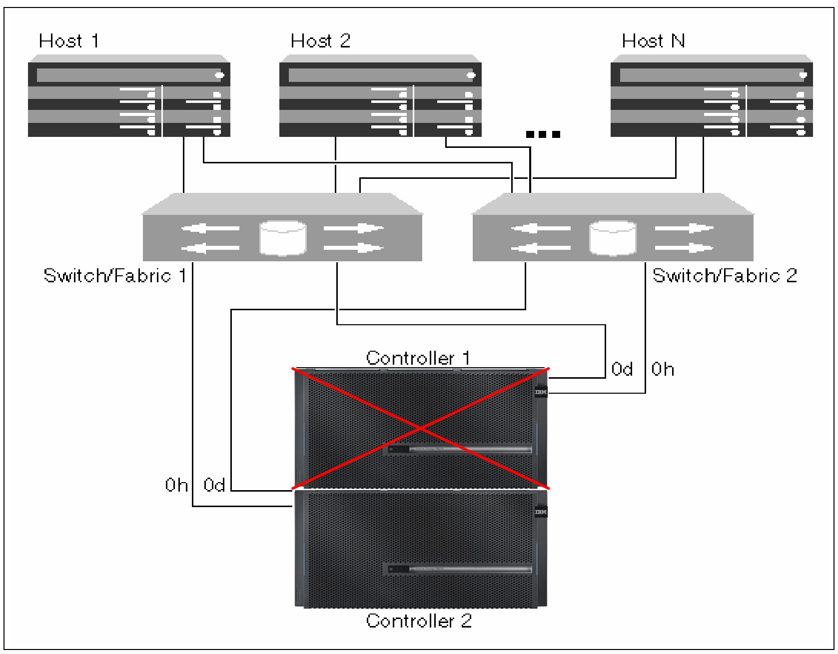


Figure 2-28 SSI with node failure

Access after failover

After a takeover occurs, the surviving partner node can be viewed as having two identities (its own identity and its partner's identity) that exist simultaneously on the same storage system unit. Each identity can access only the appropriate volumes and networks.

You can send commands or log in to either storage system unit by using the **rsh** or **telnet** command, allowing remote scripts that invoke storage system commands through a remote shell connection to continue to operate normally.

- Access with **rsh**

  Using **rsh**, commands sent to the failed storage system unit through a remote shell connection are serviced by the partner storage system, as are **rsh** command login requests.

- Access with **telnet**

  Logging in to a failed storage system unit using a **telnet** session, you see a message alerting you that your storage system failed and to log into the partner node instead. See Example 2-12. If you are logged into the partner node, you can access the failed storage system or its resources from the partner node by using the **partner** command.

Example 2-12 Telnet session to failed node

```
[yepper@yepper ~]$ telnet 192.168.2.112
Trying 192.168.2.112...
Connected to 192.168.2.112 (192.168.2.112).
Escape character is '^]'.

Running in failover mode
Connection closed by foreign host.
```

Multipathing requirements for single-image cfmode

Multipathing software is required on the host so that SCSI commands fail over to alternate paths when links go down due to switch failures or cluster failovers. In the event of a failover, none of the adapters on the takeover storage system assume the WWPNs of the failed storage system.

Guidelines for migrating to single_image cfmode

The default cfmode for new installations is single_image. If you are migrating to single_image cfmode from an existing system, follow the important guidelines in this section.

For additional information, refer to “IBM System Storage N series Changing the Cluster cfmode Setting in Fibre Channel SAN Configurations” at:

http://www-01.ibm.com/support/docview.wss?rs=1147&uid=ssg1S7001382

Reasons for changing to single_image cfmode

The single_image cfmode provides a number of advantages over the other cfmodes. Some of them are:

- The host can access all LUNs through any target port on the active/active configuration.
- The single_image mode is supported on all storage systems.
- The single_image mode is compatible with all supported FCP hosts and active/active configuration storage systems that support the FCP protocol. There are no switch limitations. You can connect a storage system in the active/active configuration in single_image mode to any FCP switch supported by Data ONTAP.

You might also want to change your cfmode setting for the following reasons:

- To increase the number of paths to the active/active configuration. For example, you might want to use the single_image setting to increase the number of available paths in your configuration. All ports on each storage system are available to access local LUNs. In addition, single_image node is available for all supported SAN hosts.

- To support heterogeneous configurations. For example, you upgraded an existing SAN configuration with Solaris or Windows hosts and you want to add HP-UX or AIX hosts to the configuration. HP-UX and AIX hosts do not support failover when the storage systems in an active/active configuration are in standby mode. If you want to add these hosts to an existing homogenous configuration that has Solaris-only or Windows-only hosts, you must change the cfmode setting on the storage system to either single_image or partner.

Solaris or Windows hosts in the same environment as HP-UX or AIX hosts must also have multipathing software installed when the cfmode setting is single_image or partner.

Impact of changing to single_image cfmode

When you change the cfmode setting of the active/active configuration to single_image, several configuration components are affected.

Carefully consider the following impacts before changing to single_image cfmode:

- Host access to LUNs

  Hosts cannot access data on mapped LUNs. When you change the cfmode setting, you change the available paths between the host and the storage systems in the active/active configuration. Some previously available paths are no longer available and some new paths become available. The LUNs might be accessible but cannot be used until you reconfigure the host to discover the new paths. You must reconfigure every host that is connected to the active/active configuration to discover the new paths. The LUNs are not accessible until you reconfigure the host. The procedure depends on your host operating system.

- Multipathing software

  If you have multipathing software in your configuration, changing the cfmode setting might also affect the multipathing policy.

- Switch zoning

  When you change the active/active configuration's cfmode setting to single_image cfmode, both nodes in the active/active configuration use the same WWNN. One node assumes the WWNN of its partner. If your storage system connects to a switch by using soft zoning (zoning by WWPN), you must update your zones to accommodate the WWNN change. Systems that connect to switches using hard (port) zoning are not affected.

- Cabling

  The single_image setting makes more target ports available to the host. This means that you might have to change your cabling configuration.

Downtime planning

Changing the cfmode setting on the storage system requires host reconfiguration and, in some cases, you might have to reboot the host.

The procedures also require you temporarily to stop host I/O and take host applications offline.

You should schedule downtime for your configuration before you change cfmode settings.

How Data ONTAP displays information about target ports in single_image cfmode

Example 2-13 shows how Data ONTAP displays target ports when the active/active configuration is in single_image mode and in normal operation.

Each system has two adapters. Note that all ports show the same WWNN (node name), and the mediatype of all adapter ports is set to auto. This means that the ports log into the fabric using point-to-point mode. If point-to-point mode fails, then the ports try to log into the fabric in loop mode. You can use the **fcp config mediatype** command to change the default mediatype of the ports from auto to another mode according to the requirements of your configuration.

Example 2-13 fcp config output

```
storage_system1> fcp config
4a: ONLINE [ADAPTER UP] PTP Fabric
      host address 011f00
      portname 50:0a:09:81:82:00:96:d5 nodename 50:0a:09:80:82:00:96:d5
      mediatype auto
4b: ONLINE [ADAPTER UP] PTP Fabric
      host address 011700
      portname 50:0a:09:82:82:00:96:d5 nodename 50:0a:09:80:82:00:96:d5
      mediatype auto
5a: ONLINE [ADAPTER UP] PTP Fabric
      host address 011e00
      portname 50:0a:09:83:82:00:96:d5 nodename 50:0a:09:80:82:00:96:d5
      mediatype auto
5b: ONLINE [ADAPTER UP] PTP Fabric
      host address 011400
      portname 50:0a:09:84:82:00:96:d5 nodename 50:0a:09:80:82:00:96:d5
      mediatype auto

storage_system2> fcp config
4a: ONLINE [ADAPTER UP] PTP Fabric
      host address 011e00
      portname 50:0a:09:81:92:00:96:d5 nodename 50:0a:09:80:82:00:96:d5
      mediatype auto
4b: ONLINE [ADAPTER UP] PTP Fabric
      host address 011400
      portname 50:0a:09:82:92:00:96:d5 nodename 50:0a:09:80:82:00:96:d5
      mediatype auto
5a: ONLINE [ADAPTER UP] PTP Fabric
      host address 011f00
      portname 50:0a:09:83:92:00:96:d5 nodename 50:0a:09:80:82:00:96:d5
      mediatype auto
5b: ONLINE [ADAPTER UP] Loop Fabric
      host address 0117da
      portname 50:0a:09:84:92:00:96:d5 nodename 50:0a:09:80:82:00:96:d5
      mediatype auto
```

How Data ONTAP avoids igroup mapping conflicts with single_image cfmode

Each node in the active/active configuration shares its partner's igroup and LUN mapping information.

Data ONTAP uses the active/active configuration interconnect to check igroup and LUN mapping information and also provides the mechanisms for avoiding mapping conflicts.

igroup ostype conflicts

When you add an initiator WWPN to an igroup, Data ONTAP verifies that there are no igroup ostype conflicts.

An ostype conflict occurs, for example, when an initiator with the WWPN 10:00:00:00:c9:2b:cc:39 is a member of an AIX igroup on one node in the active/active configuration and the same WWPN is also a member of a group with the default ostype on the partner.

Reserved LUN ID ranges

By reserving LUN ID ranges on each storage system, Data ONTAP provides a mechanism for avoiding mapping conflicts.

If the active/active configuration interconnect is down, and you try to map a LUN to a specific ID, the **`lun map`** command fails. If you do not specify an ID in the **`lun map`** command, Data ONTAP automatically assigns one from a reserved range.

The LUN ID range on each storage system is divided into three areas:

- IDs 0 to 192 are shared between the nodes. You can map a LUN to an ID in this range on either node in the active/active configuration.
- IDs 193 to 224 are reserved for one storage system.
- IDs 225 to 255 are reserved for the other storage system in the active/active configuration.

Bringing LUNs online

The **`lun online`** command fails when the active/active configuration interconnect is down to avoid possible LUN mapping conflicts.

When to override possible mapping conflicts

When the active/active configuration interconnect is down, Data ONTAP cannot check for LUN mapping or igroup ostype conflicts.

The following commands fail unless you use the -f option to force them. The -f option is only available with these commands when the active/active configuration interconnect is down and the cfmode is single_image.

- **`lun map`**
- **`lun online`**
- **`igroup add`**
- **`igroup set`**

You might want to override possible mapping conflicts in disaster recovery situations or situations in which the partner in the active/active configuration cannot be reached and you want to regain access to LUNs. For example, the following command maps a LUN to an AIX igroup and assigns a LUN ID of 5, regardless of any possible mapping conflicts:

```
lun map -f /vol/vol2/qtree1/lun3 aix_host5_group2 5
```

2.4.5 Partner mode

The partner cfmode is supported on all FCP-licensed systems except the IBM System Storage N series N3700 (2863-A20) storage system, N7000 series, and systems with 4 Gb adapters. It is also supported for all host OS types.

For systems with HBAs, port A and port B are both active. Port A on each HBA provides access to local LUNs, and port B provides access to LUNs on the partner system. The target ports log into the fabric using a point-to-point topology.

If you have the IBM N5000 series storage system with a new installation of Data ONTAP, the state of the onboard Fibre Channel port depends on your configuration. In the default two-port configuration, ports 0c and 0d connect to the SAN. Port 0c provides access to local LUNs, and port 0d provides access to LUNs on the partner. In a four-port configuration in which all onboard ports connect to the SAN, ports 0a and 0c on each node in the active/active configuration provide access to local LUNs, and ports 0b and 0d provide access to LUNs on the partner.

Figure 2-29 shows a sample configuration with a multi-attached host connecting to an active/active configuration with target HBAs. The solid lines represent paths to LUNs on the local storage system and the dotted lines represent paths to partner LUNs.

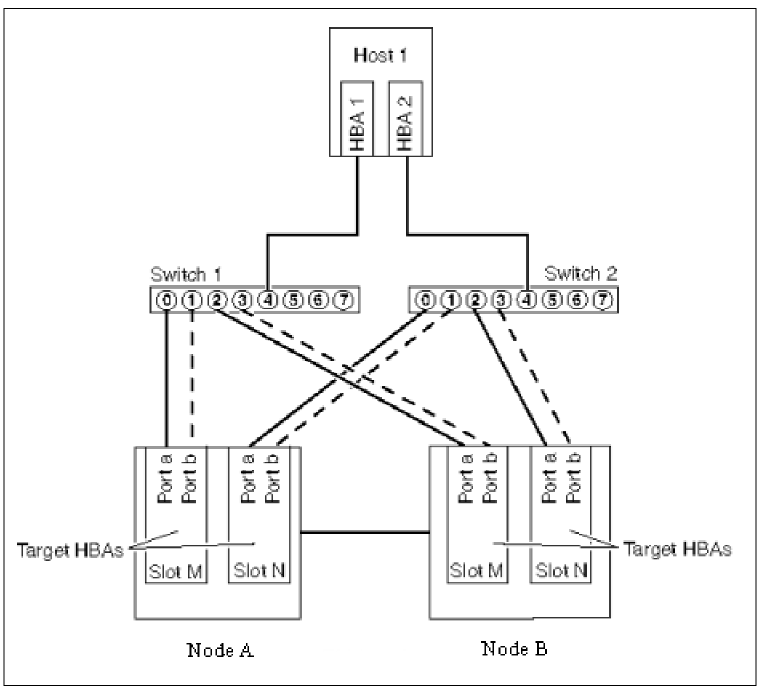


Figure 2-29 Partner cfmode in an active/active configuration

Multipathing requirements for partner cfmode

Partner mode requires host-side multipathing software. The multipathing policy is active/passive. The primary paths to the LUNs are always through the A ports. The B ports are secondary paths.

Table 2-29 on page 103 shows the available paths between the host and the storage systems in the active/active configuration shown in Figure 2-29.

Table 2-29 Available paths between the host and the nodes in the active/active configuration

	Type of path	Target FCP ports
Node A LUN	Local/primary	Port a, slot M on node A Port a, slot N on node A
	Partner/secondary	Port b, slot M on node B Port b, slot N on node B
Node B LUN	Local/primary	Port a, slot M on node B Port a, slot N on node B
	Partner/secondary	Port b, slot M on node A Port b, slot N on node A

How Data ONTAP displays target ports in partner cfmode

When the FCP cfmode setting is partner, the local and partner addresses of the WWNN and WWPN have a pattern of 50:a9:80:nn:nn:nn:nn.

The WWPN and WWNN of the B ports are based on the WWNN of the partner storage system in the active/active configuration. For example, port B on node A presents the WWNN of node B. This means that the WWNNs and WWPNs do not change during a takeover. The static address requirements for the N_Port ID, S_ID, and D_ID are maintained.

Example 2-14 and Figure 2-30 on page 104 show how Data ONTAP displays WWNN and WWPN when the storage system's cfmode is set to partner and the active/active configuration is in normal operation.

Example 2-14 Adapter settings with cfmode of partner

```
storage_system> fcp config
   9a: ONLINE [ADAPTER UP] PTP Fabric
         host address 021b00
         portname 50:a9:80:01:03:00:e0:73 nodename 50:a9:80:00:03:00:e0:73
         mediatype ptp partner adapter 9a
   9b: ONLINE [ADAPTER UP] PTP Fabric
         host address 021a00
         portname 50:a9:80:0a:03:00:e0:5f nodename 50:a9:80:00:03:00:e0:5f
         mediatype ptp partner adapter 9b
   11a: ONLINE [ADAPTER UP] PTP Fabric
         host address 021500
         portname 50:a9:80:03:03:00:e0:73 nodename 50:a9:80:00:03:00:e0:73
         mediatype ptp partner adapter 11a
   11b: ONLINE [ADAPTER UP] PTP Fabric
         host address 021600
         portname 50:a9:80:0c:03:00:e0:5f nodename 50:a9:80:00:03:00:e0:5f
         mediatype ptp partner adapter 11b
```

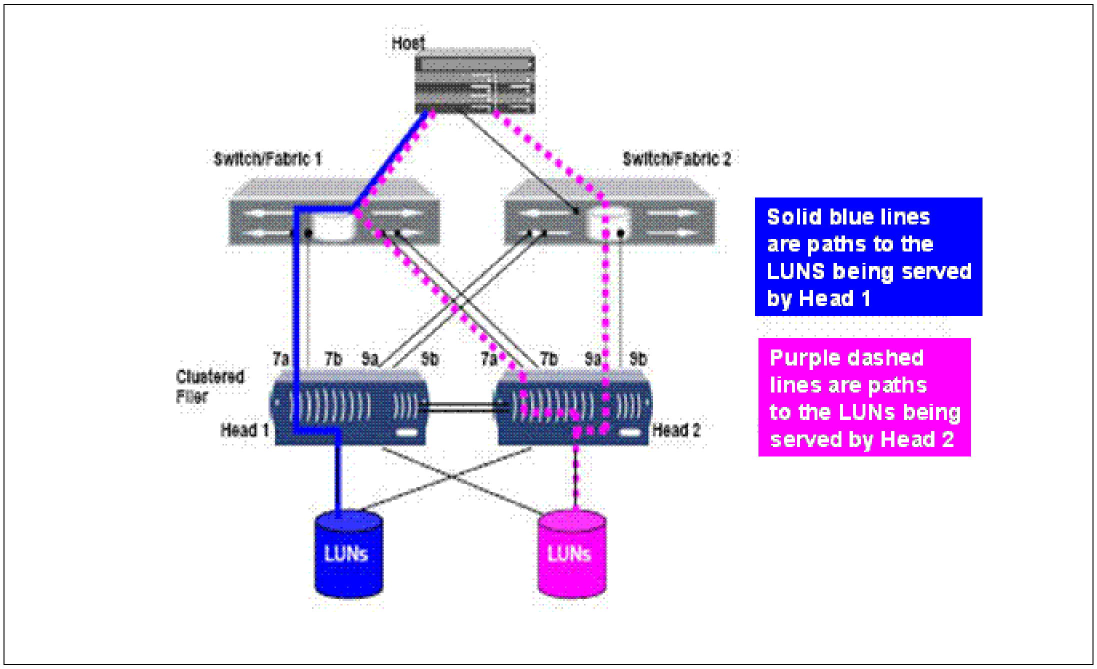


Figure 2-30 Partner mode active/passive, no load balancing

Dual Fibre Channel target cards are supported with partner, standby, and mixed cfmodes.

The key points to remember are:

- Each N series serves its own LUNs on its A ports, and its partner's LUNs on its B ports.
- Each access through a B port is normally a proxy path that requires passing the block over the active/active configuration interconnect.
- If an N series head is down, the partner head serves the LUN data on behalf of the downed head directly over the B port.

What happens when a link fails with partner cfmode

When a link fails, the host loses a primary path and fails over to the other primary paths. If there are no other primary paths available, the host can access LUNs through the secondary paths. For example, see Figure 2-31 on page 105. If the host loses the path through the A port of the HBA in slot M to NodeA, the host can access LUNs through the B ports on NodeB.

The failover method depends on the host and multipathing software. For example, if you have VERITAS Volume Manager (VxVM) with Dynamic MultiPathing software and the Array Support Library (ASL) on a Solaris host, all LUNs that share active paths form a group. If all active paths fail for a LUN in a group, all LUNS in the group fail over to the secondary paths.

For detailed information about how each host handles failover, see "Installing the ASL and APM" in FCP Host Utilities documentation for Solaris at:

http://www-01.ibm.com/support/docview.wss?rs=1196&uid=ssg1S7002489

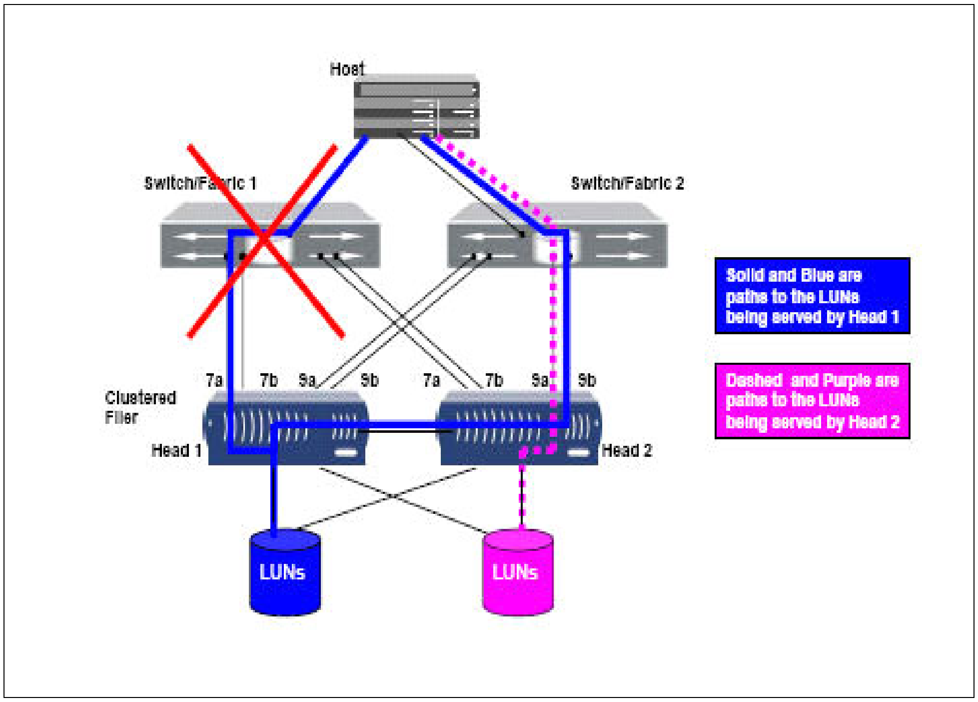


Figure 2-31 Switch failure: partner mode

What happens during a takeover with partner cfmode

If NodeB takes over for NodeA, the host continues to access LUNs on NodeA through the B ports on NodeB. The WWNN and WWPN of the B ports on NodeB do not change.

This enables HP-UX and AIX hosts, which track target devices based on WWPN/WWNN and N_Port ID (the switch-assigned addresses), to maintain correct information about available paths.

Active/active configuration failover event

If an active/active failover occurs, the takeover N series serves its own LUNs on its A ports, as is normal. The partner's LUNs are served through the B port and the N series direct connection to disk shelves, as illustrated in Figure 2-32.

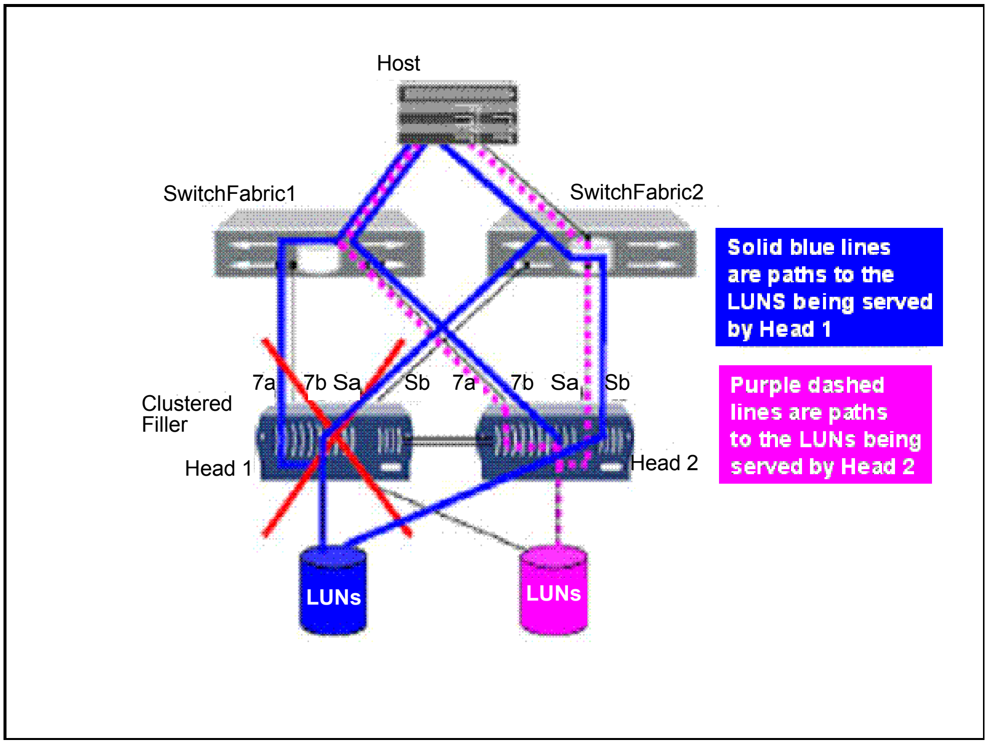


Figure 2-32 Head 1 failure in partner mode

2.4.6 Standby mode

Note the following points regarding standby mode:

- Under normal operations, FC ports for LUNs being served by a head are only active on that head.
- An FC port's LUNs on its partner head are only activated at N series cluster failover.
- Ports on the partner head take over the identity (WWPN and WWNN) of the ports from the original head. This is analogous to IP failover where the IP address fails over to the other head at cluster failover (with the MAX address).
- Standby mode does not work for HP-UX or AIX.
- N3000 models are not supported.
- There is no problem of a host accessing through a secondary (proxy path, IC) path.
- The standby cfmode setting requires more switch ports because port A and port B on each HBA must connect to the switch, even though B becomes active only in the event of a takeover.
- Virtual standby port is not used in dual_fabric cfmode.
- Each N series serves its own LUNs on its A ports.
- With active/passive multipathing, no load balancing occurs; all traffic travels through one host port.

- Multi-active/passive multipathing is normally used for standby mode.
- Active/active multipathing is used for MPIO Windows systems, allowing multipathing from a single host. However, on a LUN basis active/passive multipathing is used (that is, a single LUN only accesses through one path unless the LUNs are round-robined across the ports in standby mode).

Port A on each target HBA operates as the active port, and port B operates as a standby port. When the active/active configuration is in normal operation, port A provides access to local LUNs, and port B is not available to the initiator. When one system in the active/active configuration fails, port B on the partner system becomes active and provides access to the LUNs on the failed system. Then port B assumes the WWPN of the port A on the failed partner. The ports log in to the fabric in point-to-point mode.

Figure 2-33 shows a sample configuration with a multi-attached host connecting to an active/active configuration with target HBAs. The solid lines represent paths to LUNs on the local system and the dotted lines represent paths to partner LUNs.

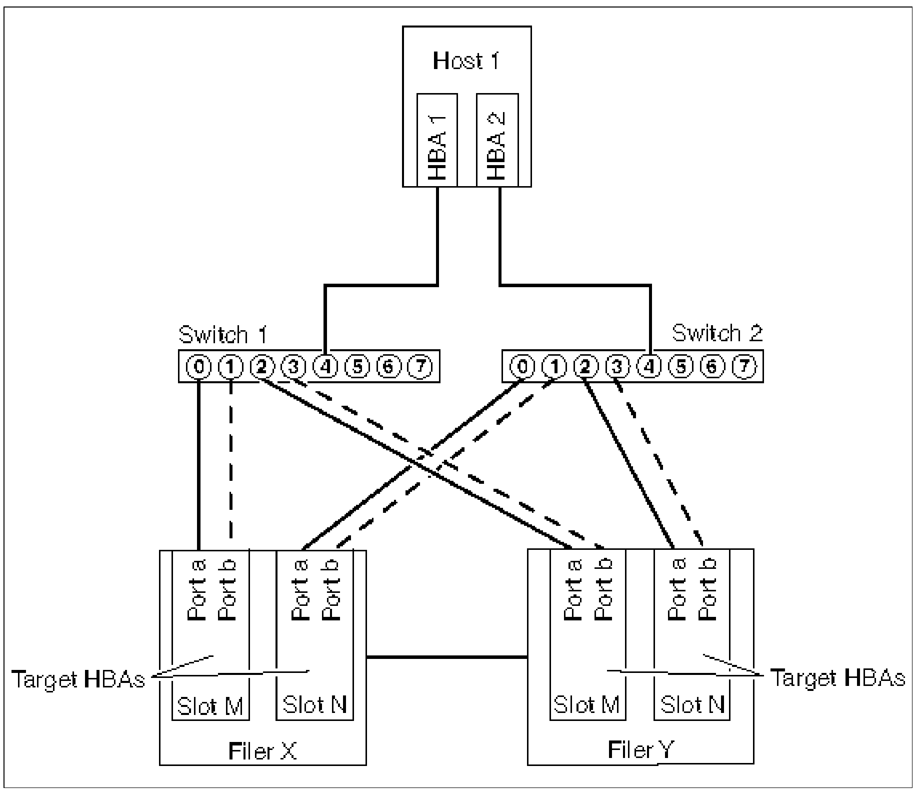


Figure 2-33 Standby cfmode in an active/active configuration

The points to remember are:

- Each N series serves its own LUNs on its A ports.
- B ports are online but have different WWNNs and no LUNs available.
- Standby mode is the same as IP card failover.

Status

Example 2-15 on page 108 illustrates the adapter settings after changing the cfmode to standby. Note how the nodename address on 5a is the nodename of the local host. The nodename for port 5b is the nodename derived from the Qlogic card.

In standby mode, if there were a cluster failover, the nodename would become that of the partner.

Example 2-15 cfmode status standby mode

```
itsotuc>fcp nodename
Fibre Channel nodename: 50:0a:09:80:82:80:c4:d1 (500a09808280c4d1)
itsotuc>fcp show cfmode
fcp show cfmode:standby
itsotuc>fcp config
5a:     LINK NOT CONNECTED <ADAPTER UP>
          host addres 00000
          portname 50:0a:09:83:82:80:c4:d1 nodename 50:0a:09:80:82:80:c4:d1
          mediatype ptp partner adapter None

5b      ONLINE <ADAPTER UP> PTP Fabric Standby
          host address 031100
          portname 20:01:00:e0:8b:27:ae:6b nodename 20:01:00:e0:8b:27:ae:6b
      mediatype ptp partner adapter 5b
```

Switch failure

If the switch in fabric 1 experiences a failure, the multipath layer selects an alternate path through switch/fabric 2.

Active/active configuration failover event in standby mode

Figure 2-34 and Example 2-16 show a configuration in which port B operates as a standby port.

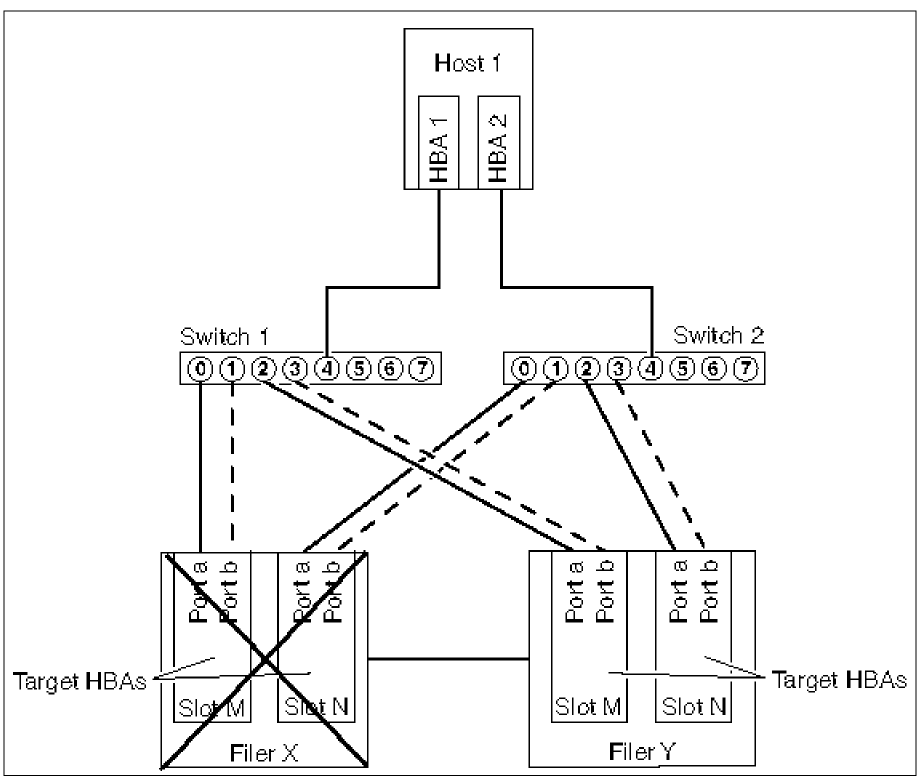


Figure 2-34 Cluster failover event in standby mode

Both storage systems in the active/active configuration show each storage system with two target HBAs, in slots M and N. For each storage system, the slot-M HBA connects to switch 1 and slot-N HBA connects to switch 2. The solid lines indicate active connections. The dotted lines indicate standby connections.

Example 2-16 Cluster failover event standby mode

```
If Filer X fails, then Filer Y takes over and the following occurs:
   Slot-M-port b on Filer Y takes over for slot-M-port a in Filer X
   Slot-N-port b on Filer Y takes over for slot-N-port a in Filer X
```

Port B on each HBA in Filer Y becomes active and enables the host to access the storage until Filer X is repaired and running. Each B port assumes the WWNN and WWPN of the corresponding A port on the failed storage system.

2.4.7 Dual fabric mode

Note the following points regarding dual fabric cfmode:

- The dual_fabric cfmode is only supported on IBM System Storage N series N3700 (2863-A20) storage system active/active configurations.
- The Fibre Channel target port of each system in the configuration supports three virtual ports:
  - Virtual local port, which provides access to LUNs on the local system
  - Virtual standby port, which is not used
  - Virtual partner port, which provides access to LUNs on the partner node

  > **Note:** For switched configurations, dual_fabric mode requires switches that support public loop.

- A host cannot distinguish between a virtual port and a physical port.
- It is the same as partner cfmode, but with virtual ports (local and partner).
- All hosts require multipathing software in order to function in an active/active configuration failover environment.
- This functionality was primarily implemented to support the single port in the N3700 for FCP SAN. However, a benefit for a modular system environment is that it allows both A and B ports to be used simultaneously. It effectively doubles the number of available paths. A host reconfiguration will be required to handle the extra paths.
- The cfmode setting for this scenario is mixed mode. A cfmode of dual_fabric also displays virtual ports, but the standby port is not used. Therefore, host multipathing software is required for hosts running in this configuration. This enables the host to use the partner port in the event of a path failure.

Figure 2-35 on page 110 illustrates normal operation for Dual Fabric mode.

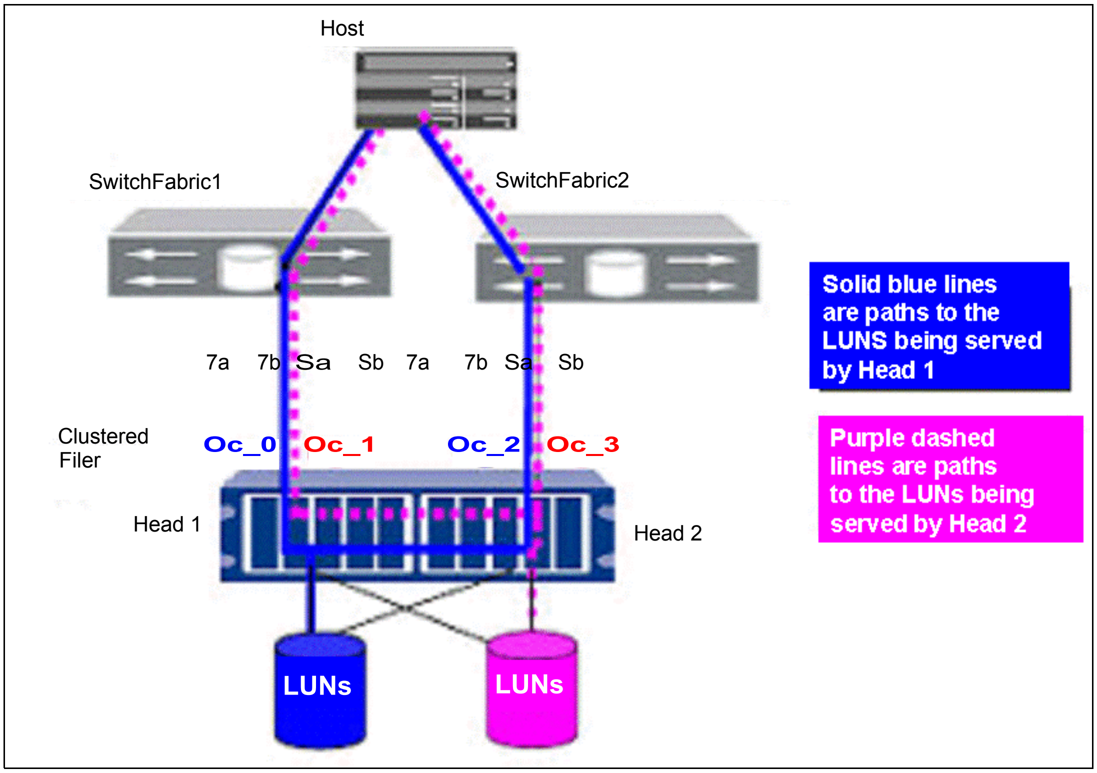


Figure 2-35 Dual fabric mode: Normal operations

Figure 2-36 illustrates switch failure for dual fabric mode.

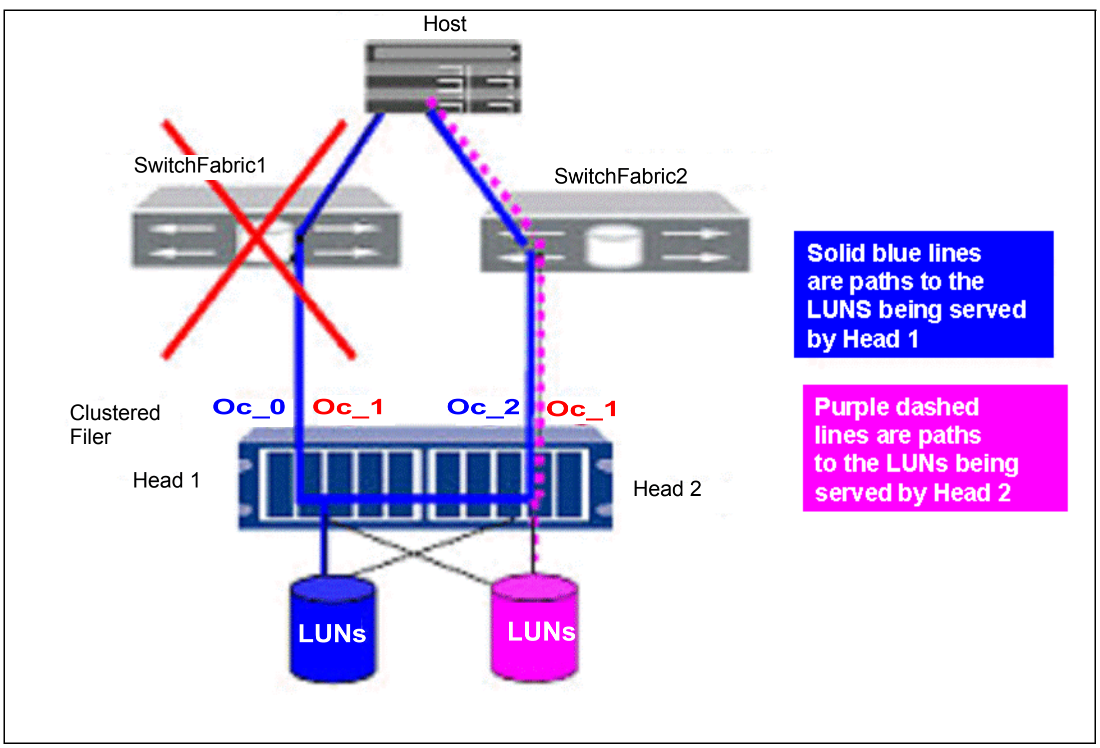


Figure 2-36 Dual fabric mode: Switch failure

Figure 2-37 illustrates a CFO event for dual fabric.

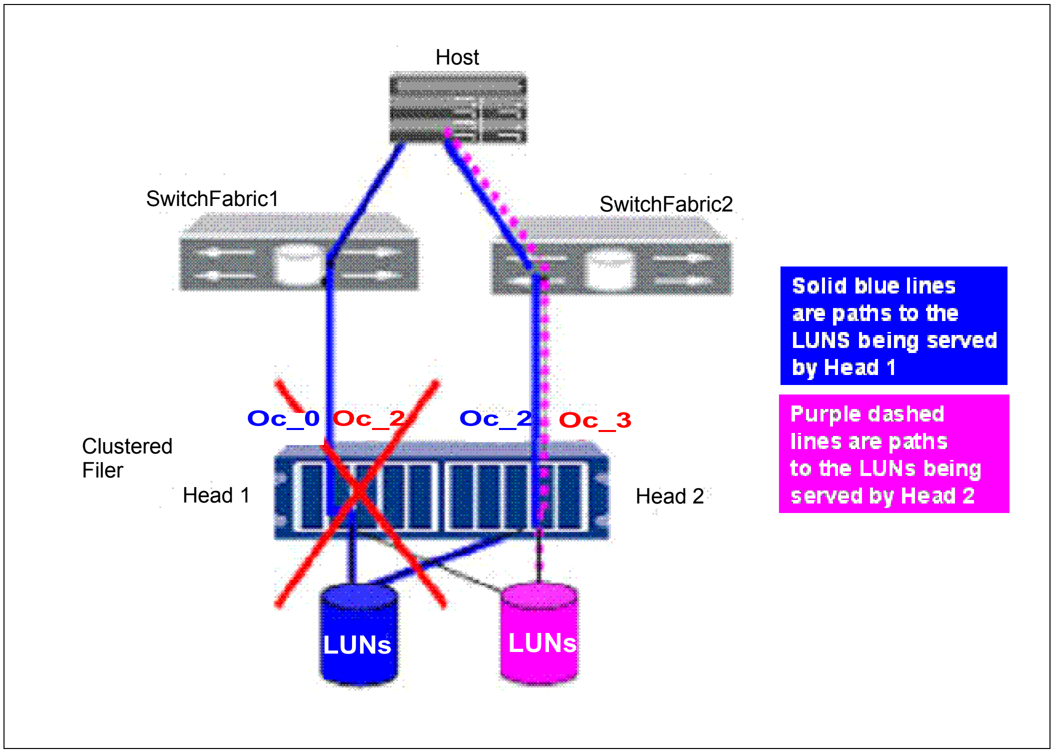


Figure 2-37 CFO event: Dual fabric

Example 2-17 displays the adapter settings with a cfmode of “dual_fabric” on an active/active N3700. Note how the adapter has three virtual ports for local, standby, and partner. The nodename for 0c_0 is for the local storage system. The nodename for port 0c _1 and 0c _2 is the nodename of the partner storage system for active/active configuration failover.

The listing for a cfmode of dual_fabric is very similar. The built-in FCP card is in slot 0c. Therefore, the portnames are 0c_0, 0c_1, and 0c_2.

Example 2-17 The cfmode dual fabric

```
itsotuc>fcp nodename
Fibre Channel nodename: 50:0a:09:84:82:80:c4:d1 (500a09848280c4d1)
itsotuc>fcp show cfmode
fcp show cfmode: dual_fabric
itsotuc>fcp config
0c:      ONLINE <ADAPTER UP> Loop Fabric
          host address 0311da
          portname 50:0a:09:84:82:80:c4:d1 nodename 50:0a:09:80:82:80:c4:d1
          mediatype loop partner adapter 0c
0c_0:     ONLINE Local
          portname 50:0a:09:84:82:80:c4:d1 nodename 50:0a:09:80:82:80:c4:d1
          loopid 0x7 portid 0x0311da
0c_1:     ONLINEd BY USER/SYSTEM Standby
          portname 50:0a:09:84:82:00:c4:d0 nodename 50:0a:09:80:82:00:c4:d0
          loopid 0x0 portid 0x000000
0c_2:     ONLINE Partner
          portname 50:0a:09:8c:82:00:c4:d0 nodename 50:0a:09:80:82:00:c4:d0
          loopid 0x9 portid 0x0311d6
```

2.4.8 Mixed mode

The characteristics of mixed mode are presented in the following list:

- The mixed cfmode is supported on all systems except the N3000 and the N7000 series.
- Mixed mode is supported, but not recommended. Use it only when needed.
- Each FCP target port supports three virtual ports:
  - Virtual local port, which provides access to LUNs on the local system.
  - Virtual standby port, which provides access to LUNs on the failed system when a takeover occurs. The standby virtual port assumes the WWPN of the corresponding port on the failed partner.
  - Virtual partner port, which provides access to LUNs on the partner system. This port enables hosts to bind the physical switch port address to the target device and allows hosts to use active/passive multipathing software.
- In mixed mode, the target ports connect to the fabric in loop mode. This means that you cannot use mixed mode with switches that do not support public loop.
- A host cannot distinguish between a virtual port or a physical port.
- It allows all ports to be simultaneously active.
- AIX or HP-UX hosts connected to an active/active configuration in mixed mode must have multipathing software installed.
- Under normal operations:
  - Each N series serves its own LUNs on its A ports (Figure 2-38 on page 113).
  - Active/passive multipathing, no load balancing, all traffic through one host port.
  - Multi-active/passive, normally for standby mode.
  - Active/active is used for MPIO Windows. However, it is active/passive per LUN (that is, a single LUN only accesses through one path unless that path fails, but the LUNs are round-robined across the ports in standby mode).

Status

Example 2-18 shows the adapter settings with a cfmode of mixed. Note that the adapter now has three virtual ports for local, standby, and partner. The nodename for 5b_0 is for the local storage system. The nodename for port 5b_1 and 5b_2 is the nodename of the partner storage system for active/active configuration failover.

Example 2-18 Status mixed mode

```
itsotuc>fcp config 5b
5b:      ONLINE <ADAPTER UP> Loop Fabric
          host address 0311da
          portname 50:0a:09:84:82:80:c4:d1 nodename 50:0a:09:80:82:80:c4:d1
          mediatype loop partner adapter 5b
5b_0:     ONLINE Local
          portname 50:0a:09:84:82:80:c4:d1 nodename 50:0a:09:80:82:80:c4:d1
          loopid 0x7 portid 0x0311da
5b_1:     ONLINEd BY USER/SYSTEM Standby
          portname 50:0a:09:84:82:00:c4:d0 nodename 50:0a:09:80:82:00:c4:d0
          loopid 0x0 portid 0x000000
5b_2:     ONLINE Partner
          portname 50:0a:09:8c:82:00:c4:d0 nodename 50:0a:09:80:82:00:c4:d0
          loopid 0x9 portid 0x0311d6
```

Figure 2-38 illustrates mixed mode path access.

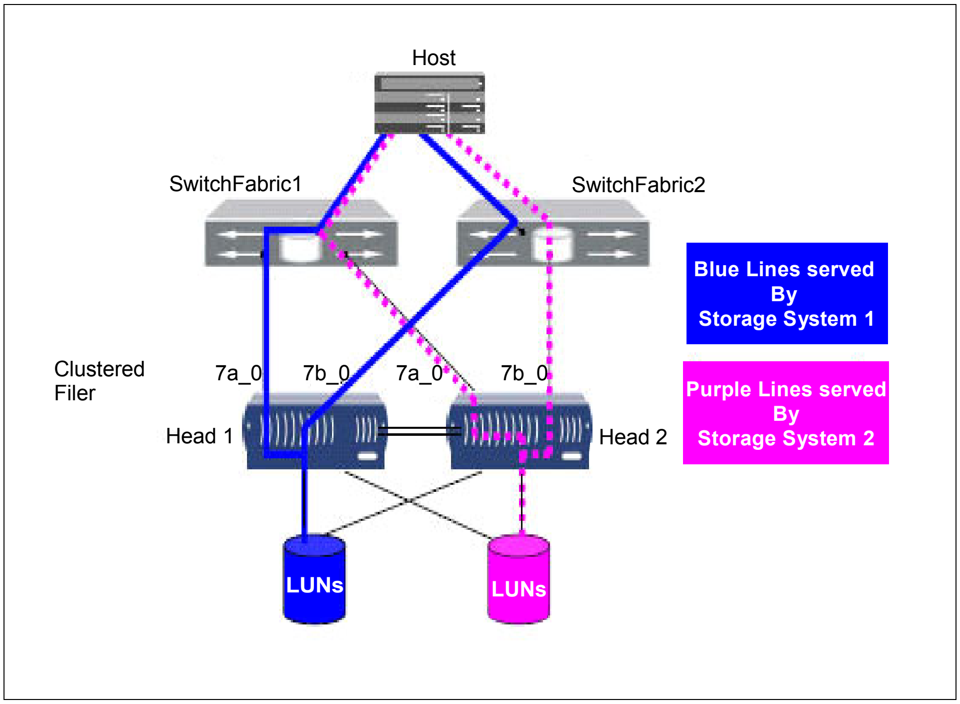


Figure 2-38 Mixed mode path access

Switch failure

If the switch in fabric 1 experiences a failure, the multipath layer selects an alternate path through switch/fabric 2. Figure 2-39 illustrates mixed mode switch failure for Linux, Solaris, Windows, AIX and HP-UX.

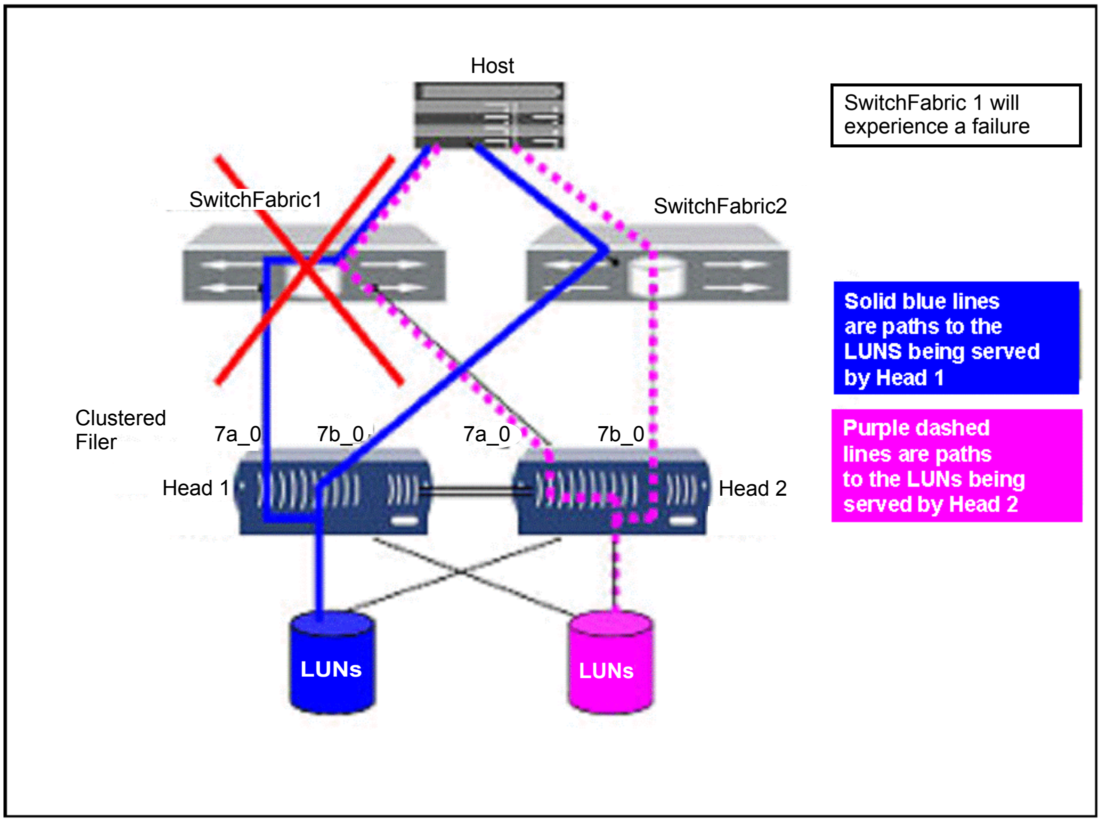


Figure 2-39 Mixed mode switch failure for Linux, Solaris, Windows

Active/active configuration failover event

If an active/active configuration failover occurs, the takeover N series serves its own LUNs on its A ports, as is normal. The partner's LUNs are served through the B port and the N series direct connection to disk shelves. Figure 2-40 on page 115 illustrates a CFO event for Linux, Solaris, and Windows.

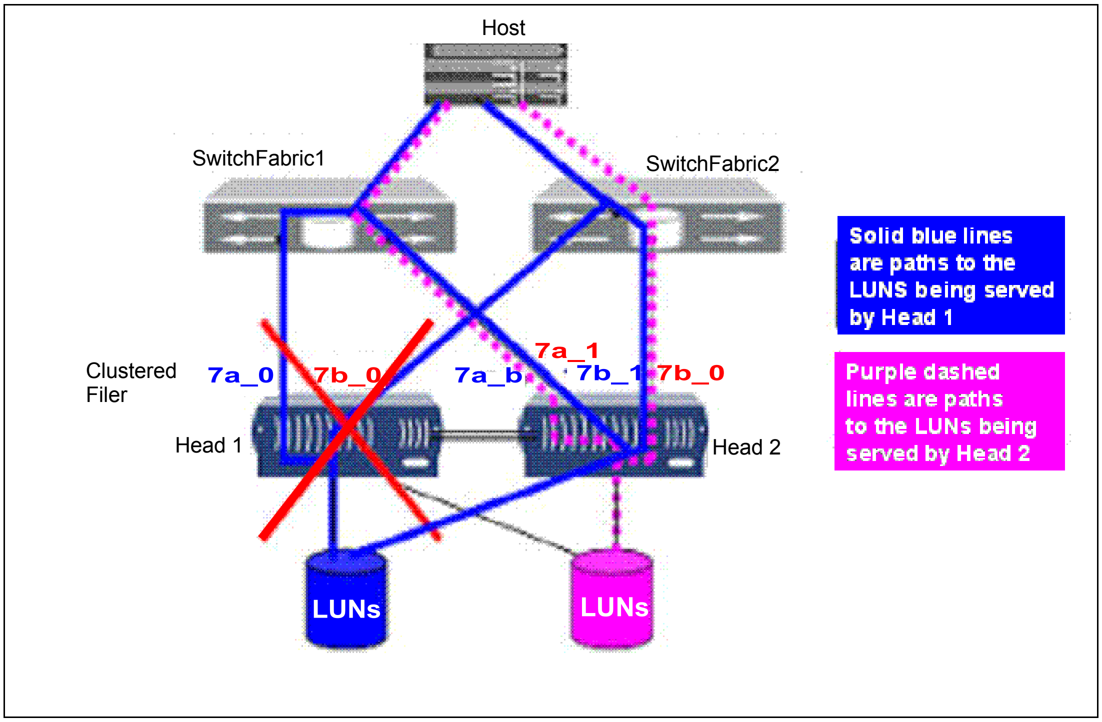


Figure 2-40 CFO event for Linux, Solaris, and Windows

Figure 2-41 illustrates a CFO event for AIX and HP-UX.

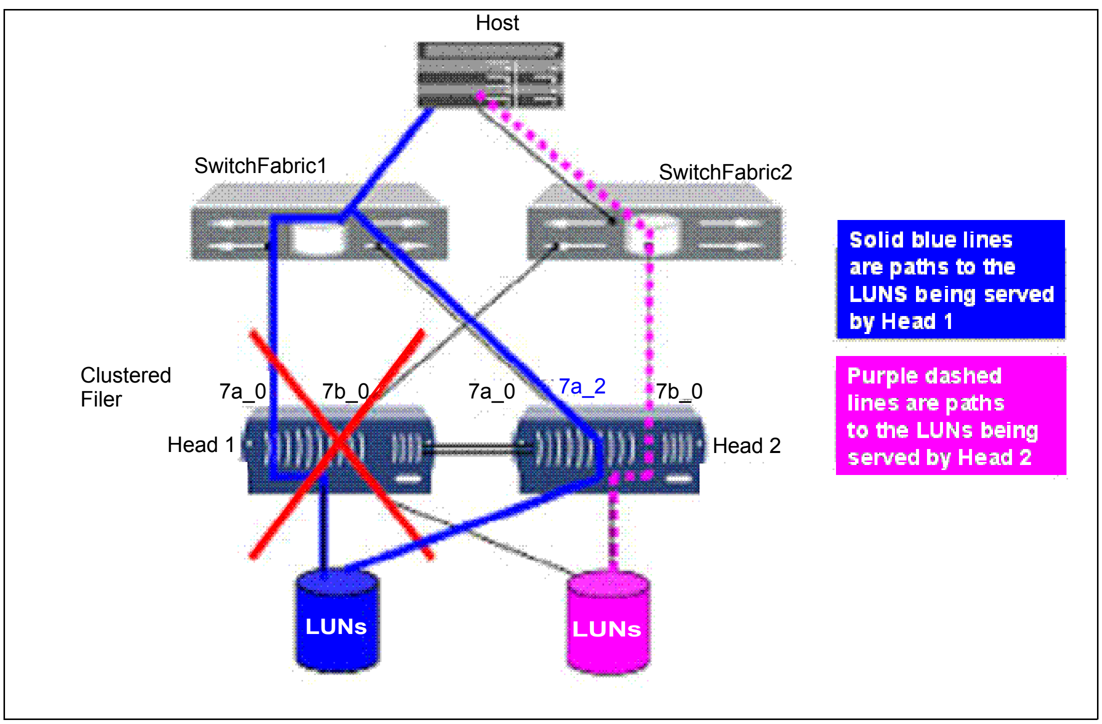


Figure 2-41 CFO event for AIX and HP-UX

Multi-ID support

Multi-ID functionality enables multiple virtual ports per physical port. It utilizes loop topology and requires the switch to support public loop. As illustrated in Figure 2-42, one physical port can support:

- One virtual port for local access
- One virtual port for partner access
- One virtual port for standby access

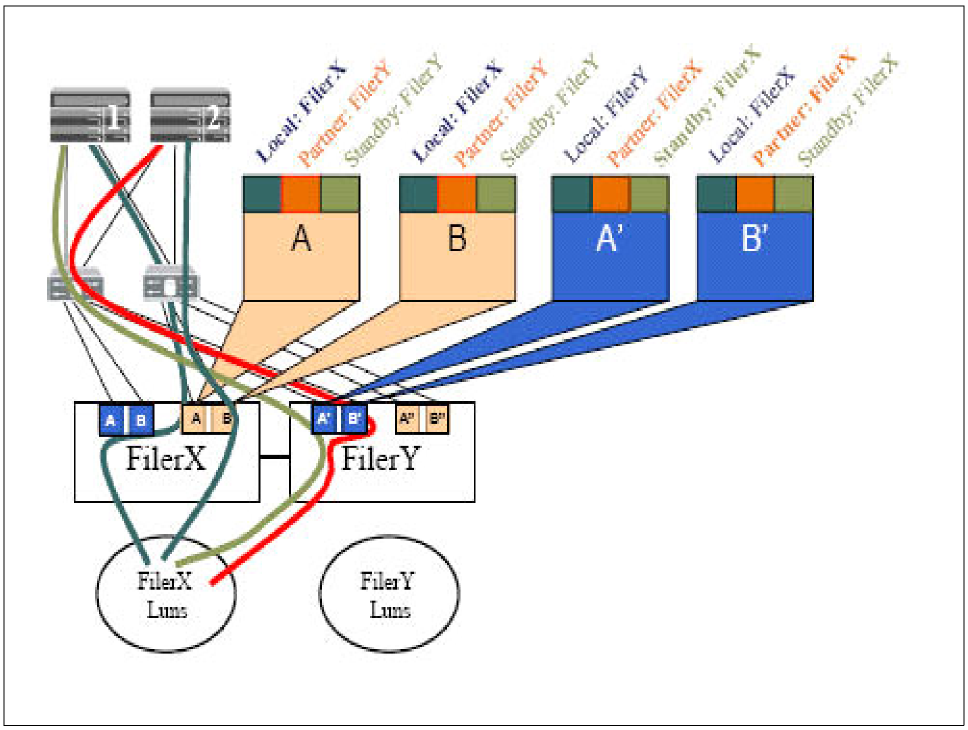


Figure 2-42 Multi-ID support

On the host side, it requires multipathing software to use partner access. This functionality was primarily implemented to support the single port in the N3700 for FCP SAN, but a benefit for a modular systems environment like the N5000 series is that it allows both A and B ports to be used simultaneously. It effectively doubles the number of available paths. A host reconfiguration is required to handle the extra paths.

The cfmode setting for this scenario is mixed mode. A cfmode of dual_fabric also displays three virtual ports, but the standby port is not used. Therefore, host multipathing software is required for hosts running in this configuration. This enables the host to use the partner port in the event of a path failure.

3

MetroCluster

This chapter discusses the MetroCluster feature, which is an integrated, high-availability, business continuance solution that allows clustering of two N5000, N6000, or N7000 storage storage systems at distances up to 100 kilometers.

The primary goal of MetroCluster is to provide mission-critical applications with redundant storage services in case of site-specific disasters. By synchronously mirroring data between two sites, it tolerates site-specific disasters with minimal interruption to applications and zero data loss.

The following topics are covered:

- Benefits of using MetroCluster
- Synchronous mirroring with SyncMirror
- Business continuity with IBM System Storage N series
- Implementing MetroCluster
- MetroCluster configurations
- Prerequisites for MetroCluster usage
- SyncMirror setup
- Failure scenarios

3.1 Overview of MetroCluster

MetroCluster, as illustrated in Figure 3-1, expands the capabilities of the N series portfolio of high-availability and disaster recovery solutions—a portfolio that includes failover, data replication, and backup solutions.

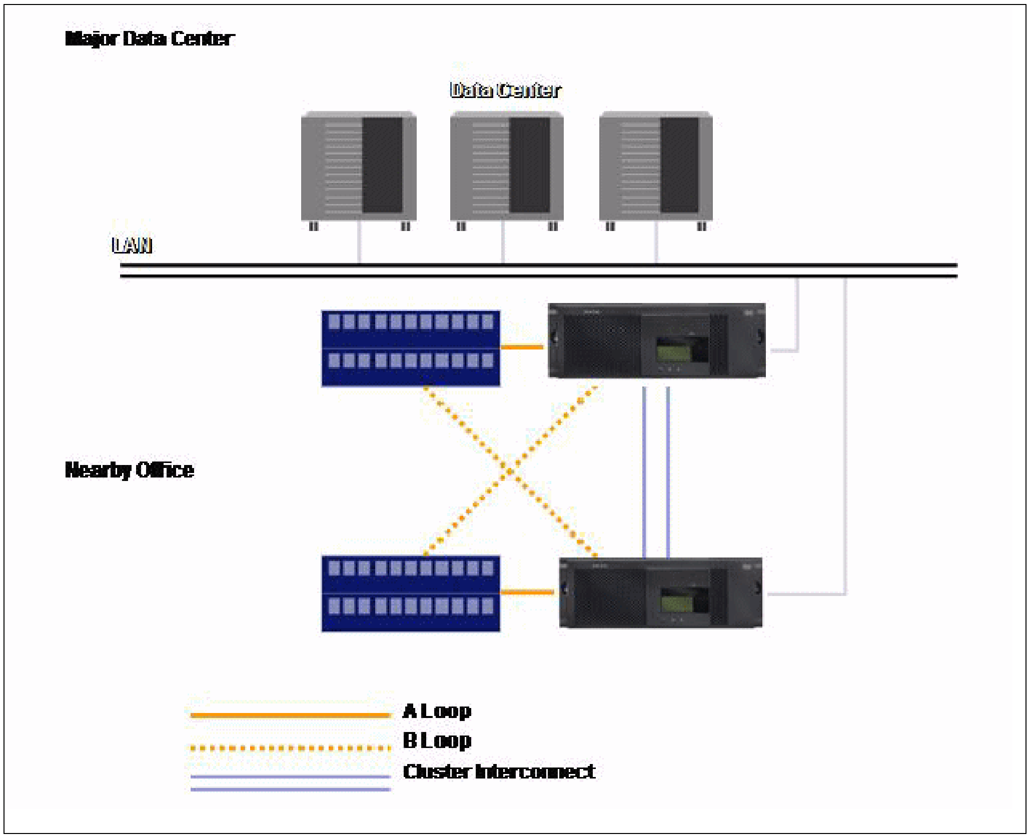


Figure 3-1 MetroCluster

MetroCluster software provides an enterprise solution for high availability over wide area networks (WANs). MetroCluster deployments of N series storage systems are used for:

- Business continuance
- Disaster recovery
- Achieving recovery point and recovery time objectives

MetroCluster technology is an important component of enterprise data protection strategies. Using MetroCluster, if a disaster occurs at a source site, businesses can continue to run and access data from a clustered node in a remote site.

A MetroCluster system is made up of the following components:

- Multiple storage controllers, HA configuration: Provides automatic failover capability between sites in case of hardware failures.
- SyncMirror: Provides an up-to-date copy of data at the remote site. Data is ready for access after failover without administrator intervention.
- Cluster remote: Provides a mechanism for an administrator to declare a site disaster and initiate a site failover through a single command for ease of use.
- FC switches: Provide storage system connectivity between sites that are greater than 500 meters apart.

MetroCluster allows the active/active configuration to be spread across data centers up to 100 kilometers apart. In the event of an outage at one data center, the second data center can assume all affected storage operations lost with the original data center. SyncMirror is required as part of MetroCluster to ensure that an identical copy of the data exists in the second data center should the original data center be lost.

- MetroCluster along with SyncMirror extends active/active clustering across data centers up to 100 kilometers apart. Either dark-fiber or DWDM between the switches is required in this configuration.
- MetroCluster and SyncMirror provide the highest level of storage resiliency across a local region.
- The highest levels of regional storage resiliency ensure continuous data availability in a particular geography.

3.2 Benefits of using MetroCluster

Using MetroCluster in your enterprise provides the following benefits:

- MetroCluster is designed to be a simple-to-administer solution that extends failover capability from within a data center to a remote site.
- It is designed to provide replication of data from the primary site to a remote site, helping keep data at the remote site current.
- The combination of failover and data replication aids in the recovery from disaster by helping to prevent the loss of data in less time than is otherwise possible.
- MetroCluster extends clustered failover capabilities from a primary site to a remote site.
- MetroCluster replicates data from the primary site to the remote site to ensure that data there is completely up-to-date and available.
- If site A goes down, MetroCluster allows you to rapidly resume operations at a remote site minutes after a disaster, as illustrated in Figure 3-2.

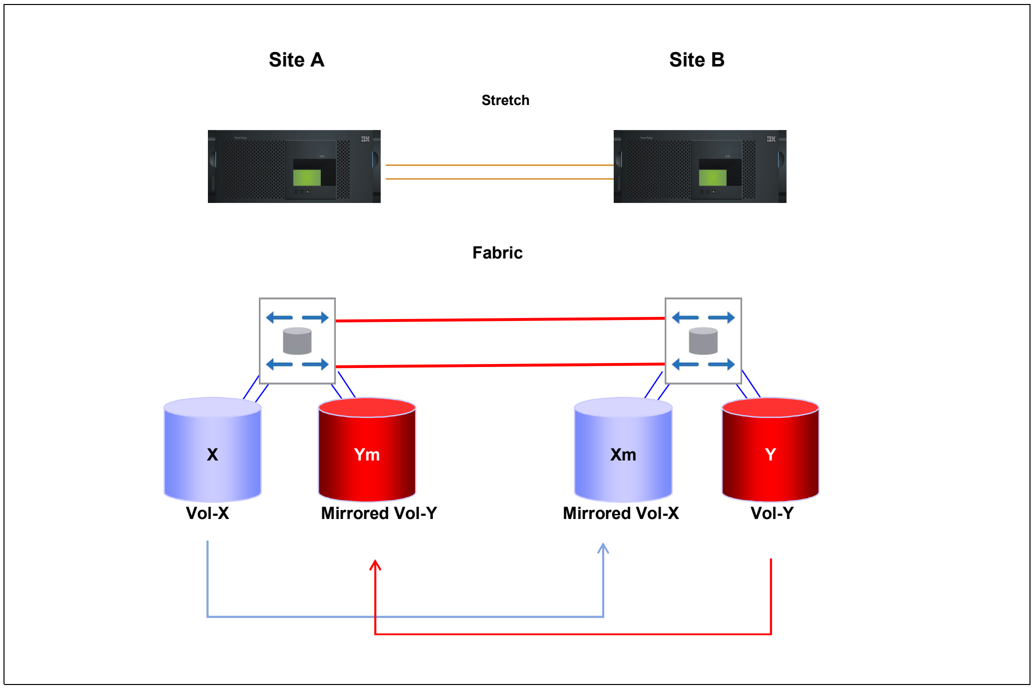


Figure 3-2 MetroCluster disaster recovery

For the highest level of storage resiliency, we recommend that you use MetroCluster with SyncMirror.

3.3 Synchronous mirroring with SyncMirror

Synchronous mirroring for MetroCluster mirrors the WAFL volumes (aggregates). Both copies or plexes are updated synchronously on writes, thus ensuring consistency.

The design of IBM System Storage N series and MetroCluster provides data availability even in the event of an outage, whether it is due to a disk problem, cable break, or host bus adapter (HBA) failure. SyncMirror can instantly access the mirrored data without any operator intervention or disruption to client applications. Read performance is optimized by performing application reads from both plexes (Figure 3-3 on page 121).

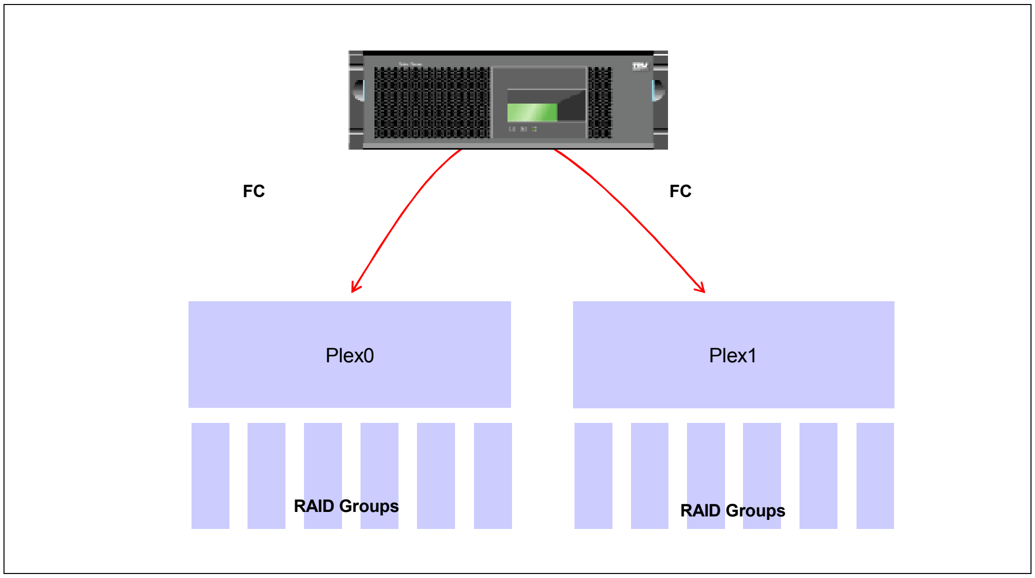


Figure 3-3 Synchronous mirroring

Performance is directly affected by the distance between plexes. Additional performance benefits of synchronous mirroring include low write impact and latencies.

SyncMirror local (without MetroCluster) is basically a standard cluster with one or both controllers mirroring their RAID to two separate shelves. The caveat in failover is that if you lose a controller and one of its RAID sets (plexes), the partner does not take over the other RAID set (plex). Therefore, without MetroCluster, all of the same rules apply as for a normal cluster:

- If controller A fails, partner B takes over.
- If loop A (Plex0) on controller A fails, controller A continues operation by running from loop B (Plex1).
- If controller A fails and either loop A or loop B (Plex0/Plex1) fails, you will not be able to continue.

MetroCluster protects against the following scenario: If controller A fails and its SyncMirrored shelves attached to loop A (Plex0) or loop B (Plex1) fail simultaneously, partner B takes over operation for partner A and its SyncMirrored plex on either loop A (Plex0) or loop B (Plex1). See Figure 3-4.

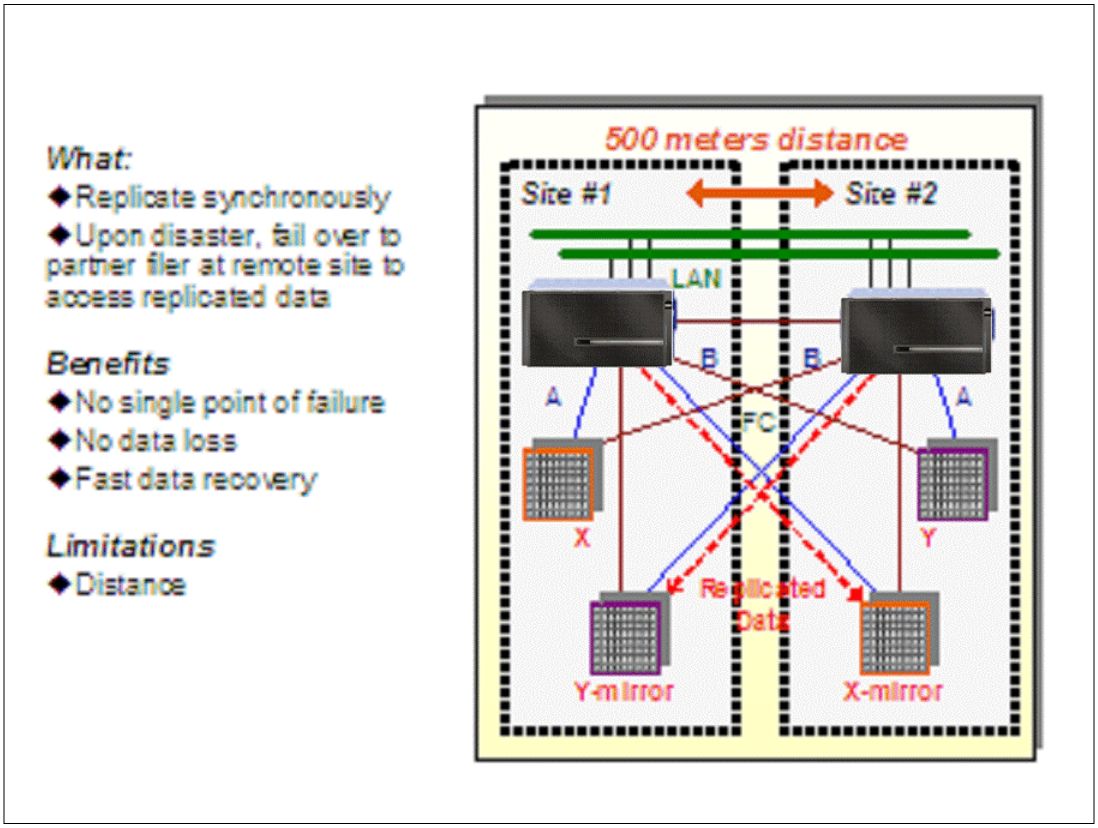


Figure 3-4 MetroCluster protection

RAID SyncMirror, which is the replication used in MetroCluster, requires that a disk at a particular location in a RAID group in one plex be the same size as the corresponding disk in the other plex. So while you can use larger disks for the new plex at the new remote site, those disks will be *downsized* to the lower size on the other side when you mirror the volumes. You can also use disks of different speeds (for example, 10 K and 15 K), but make sure that they are in separate aggregates.

3.4 Business continuity with IBM System Storage N series

The N series offers several different levels of protection and several different options. MetroCluster is just one of the options offered by the N series. MetroCluster fits into the campus-level distance requirement of business continuity. See Figure 3-5.

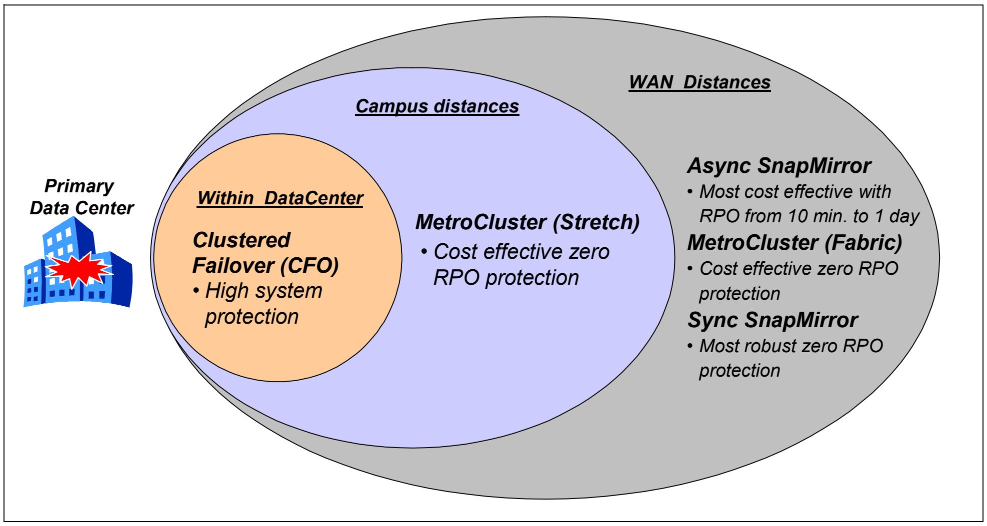


Figure 3-5 Business continuity with IBM System Storage N series

3.5 Implementing MetroCluster

There are two ways to implement MetroCluster: Stretch MetroCluster or Fabric MetroCluster. The appropriate choice is based on your business needs and the distances required. This section describes the two implementations.

3.5.1 Stretch MetroCluster

In a Stretch MetroCluster (sometimes referred to as *non-switched*), the controllers and storage are attached to Fibre Channel switches and the switches have GBICs to communicate across the WAN to one another. SyncMirror is built into MetroCluster so that every write is written to two separate expansion units in two separate aggregate groups.

The advantage of MetroCluster is the ability to take a high-availability solution and *stretch* it outside of the frame. This gives you the ability to do a site failover with a single command. Think about taking a disk subsystem like the N series and separating the storage systems miles apart and maintaining two separate disk groups to ensure failover, instead of building two completely separate N series systems and synchronously mirroring between the two.

Stretch MetroCluster has the following characteristics:

- It provides a disaster recovery option at distances up to 500 meters (depending on speed and cable type) between each N series system. See Figure 3-6.
- It only requires the MetroCluster license and the SyncMirror license.
- It is available on N5000, N6000, and N7000 storage systems.

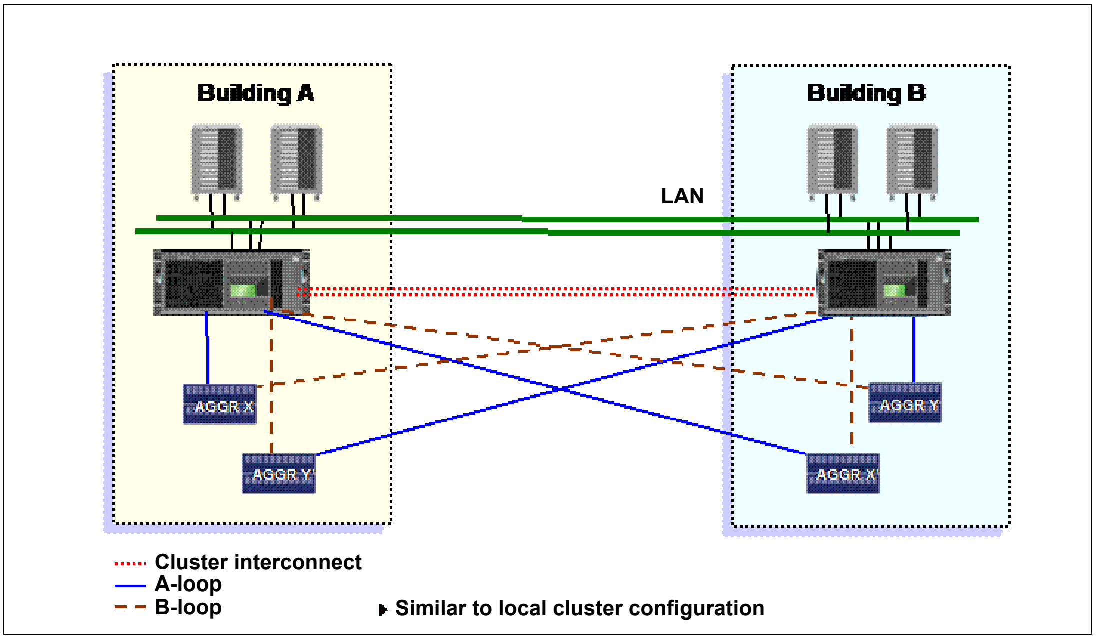


Figure 3-6 MetroCluster less than 500 meters

You can use SATA disks, either in a mirrored configuration or a non-mirrored configuration (with one of the nodes, only) on a Stretch MetroCluster. Mixed SATA and FC configurations are allowed, provided that the following requirements are met:

- There is no intermixing of FC and SATA shelves on the same loop.
- Mirror shelves must be of the same type as their parents.

3.5.2 Fabric MetroCluster

Fabric MetroCluster (also referred to as *switched*) provides a disaster recovery option at distances up to 100 km using a Fibre Channel switched network. Fabric MetroCluster uses switches for longer distance disaster recovery solutions. It is available on N5000, N6000, and N7000 series.

Figure 3-7 illustrates MetroCluster at more than 500 meters.

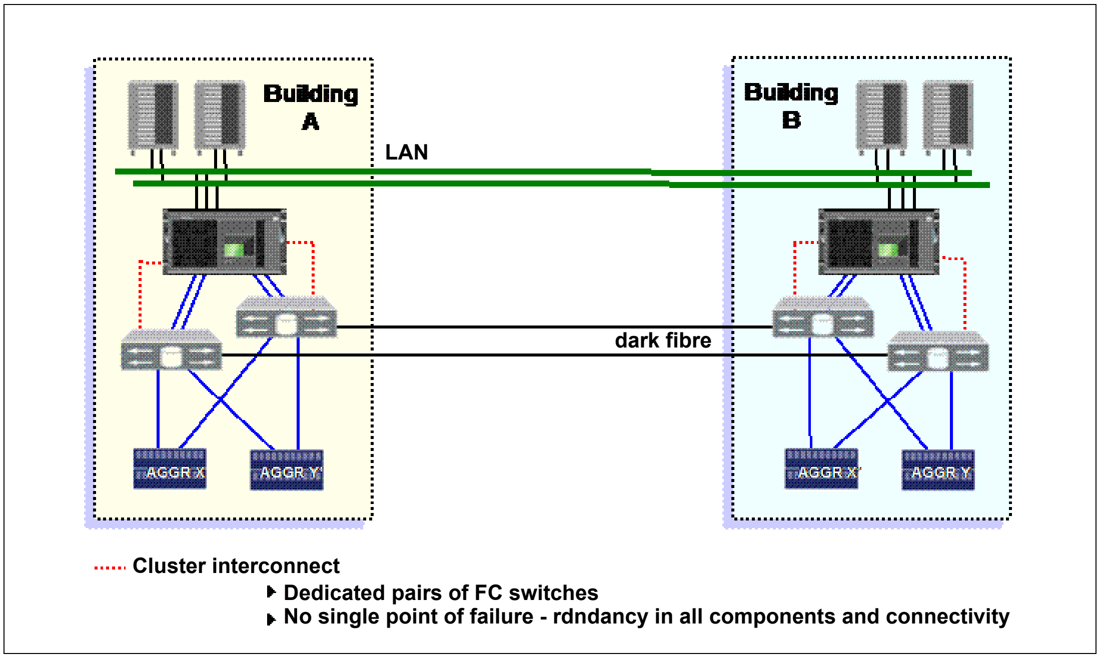


Figure 3-7 MetroCluster greater than 500 meters

Fabric MetroCluster configuration details

Make sure that you have the correct storage initiator HBAs:

- Feature code 1006 with N5200 and N5500
- Feature code 1014 or 1029 with N5300, N5600, N6000, N7000

More detailed information can be found in the interoperability guide at:

ftp://service.boulder.ibm.com/storage/nas/nseries/nseries_metrocluster.pdf

The heartbeat and NVRAM use a separate dual-port MetroCluster HBA. This HBA is dedicated and not used for the I/O traffic of SyncMirror. Table 3-1identifies the appropriate feature codes.

Table 3-1 Fabric MetroCluster Interconnect HBAs

<table>
<tr><th>Platform</th><th>N5200 and N5500</th><th>N5300 and N5600</th><th>N6000 and N7000</th><th>Minimum level of Data ONTAP</th></tr>
<tr><td>Feature Code 1018</td><td>Yes</td><td colspan="2">No</td><td>7.1.H12</td></tr>
<tr><td>Feature Code 1032</td><td>No</td><td colspan="2">Yes</td><td>7.2.3</td></tr>
</table>

Additional requirements to implement Fabric MetroCluster are that you must use clustered models, and you must have a no-charge SyncMirror license on each storage controller.

3.5.3 Fabric MetroCluster HBAs and cables

In this section we discuss feature 1042 - Copper-Fiber Converter. It is used to convert IB-copper to fiber and is required to connect the N series heads over a distance greater than 30 m. This adaptor has MPO connectors and needs an MTP plug (NVRAM5).

For details see the MetroCluster interoperability matrix for Storage initiator HBAs and interconnect HBAs:

ftp://service.boulder.ibm.com/storage/nas/nseries/nseries_metrocluster.pdf

3.5.4 Using DWDM switches

Dense Wave Division Multiplexing (DWDM) is a method of multiplexing multiple channels of fiber-optic-based protocols, such as ESCON®, Fibre Channel, FICON®, and Gbit Ethernet, onto physical cables by assigning different wavelengths of light (that is, colors) to each channel and then fanning it back out at the receiving end. The major players in the enterprise class DWDM marketplace are Nortel Networks, Cisco (ONS 15540), and Lucent.

Dense Wave Division Multiplexors are data link Layer 2 tools. Thus, the typical DWDM machine does not perform any switching, routing, or protocol conversion.

Figure 3-8 depicts a Fabric MetroCluster installation using DWDM.

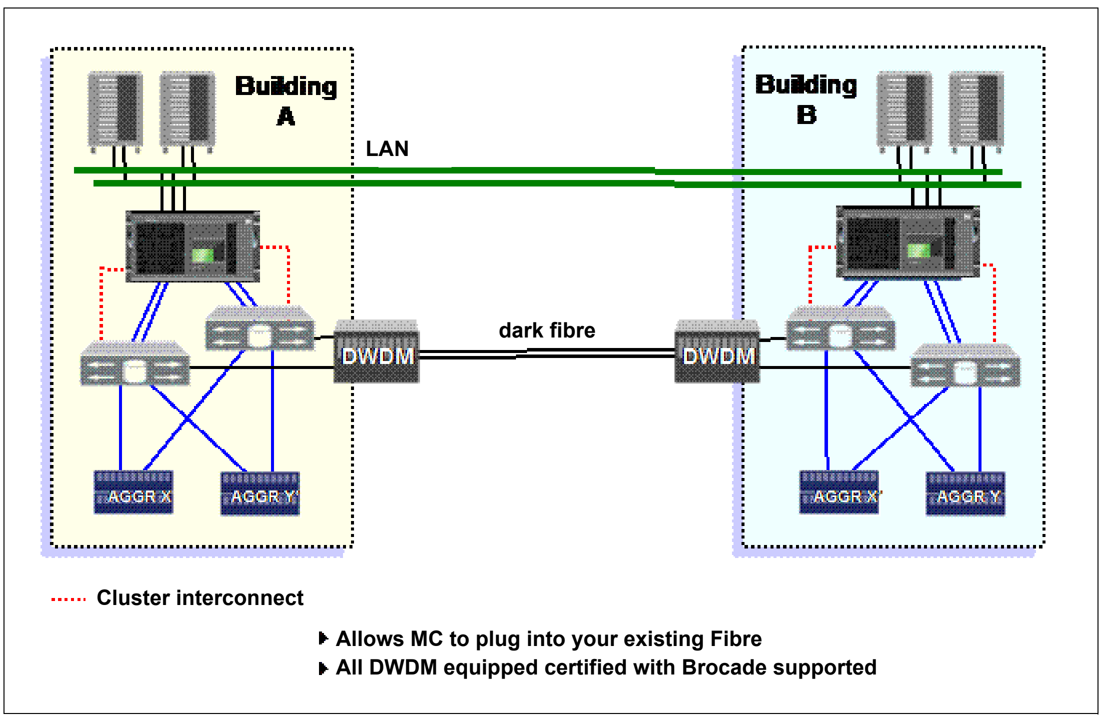


Figure 3-8 DWDM environment

3.5.5 Fabric-attached MetroCluster switches and drives

For complete interoperability information about Fabric MetroCluster with SAN switches, refer to the N series MetroCluster interoperability matrix.

FC drives are supported. SATA drives are not available for use in Fabric MetroCluster configurations. SAS drives are not yet supported.

For installation details refer to *IBM Data ONTAP Cluster Installation and Administration Guide:*

http://www-1.ibm.com/support/docview.wss?uid=ssg1S7001533&aid=1

3.5.6 Tips for using MetroCluster

This section presents additional recommendations for implementing MetroCluster.

Supported transport

MetroCluster runs over Fibre Channel (FC) and requires dark fiber links between the primary site and the remote site. MetroCluster does not run over an IP network. However, the clients accessing volumes on an N series MetroCluster configuration can use the IP network to access the N series.

Connecting at distances greater than 30 m

To connect the N series heads over a distance greater than 30 m, you must convert the IB-copper to fiber using feature 1042 - Copper-Fiber Converter. This adaptor has MPO connectors and requires an MTP plug (NVRAM5).

What is MTP/MPO? Optical links specify the LC duplex connector for 1X links, the MTP/MPO for 4X links, and dual MTP/MPO for 12X links. All of these optical connectors have been accepted as ad hoc industry standards for other optical data (communication protocols as well).

Fibre Channel SAN in a MetroCluster world

For those who are experienced in Fibre Channel technology, and storage area networks (SANs) in particular, there are differences and restrictions worth discussing relative to how Fabric MetroCluster utilizes this technology.

Fabric MetroCluster configurations use Fibre Channel switches as the means to separate the controllers by a greater distance. The switches are connected between the controller heads and the disk shelves, and to each other. Each disk drive or LUN individually logs into a Fibre Channel fabric. The nature of this architecture requires, for performance reasons, that the two fabrics be completely dedicated to Fabric MetroCluster. Extensive testing was done to ensure adequate performance with switches included in a Fabric MetroCluster configuration. For this reason, Fabric MetroCluster requirements prohibit the use of any other model or vendor of Fibre Channel switch than the Brocade included with the Fabric MetroCluster.

Also for performance reasons there is a current maximum number of drives. Higher drive count solutions will be available in the future. Keep in mind that this is for Fabric MetroCluster only. The maximum number of disks in a Stretch MetroCluster depends solely on the model:

- Up to 672 drives (48 shelves) on N7000, N6060, N6070
- Up to 504 drives (36 shelves) on N5600 with 16-port switches
- Up to 336 drives (24 shelves) on the other systems with 24-port switches

In a traditional SAN there is great flexibility in connecting devices to ports as long as the ports are configured correctly and any zoning requirements are met. A MetroCluster, however, expects certain devices to be connected to specific ports or ranges of ports. It is therefore critical that cabling be exactly as described in the installation procedures. Also, no switch-specific functions such as trunking or zoning are currently used in a Fabric MetroCluster, making switch management minimal.

3.6 MetroCluster: Cluster_Remote license

The Cluster_Remote license provides features that enable the administrator to declare a site disaster and initiate a site failover using a single command.

It enables the **`cf forcetakeover -d`** command, which initiates a takeover of the local partner even in the absence of a quorum of partner mailbox disks. This gives the administrator the ability to declare a site-specific disaster and have one node take over its partner's identity without a quorum of disks.

Several requirements must be in place to enable takeover in a site disaster:

- Root volumes of both storage systems *must* be synchronously mirrored.
- Only synchronously mirrored aggregates are available during a site disaster.

Administrator intervention, that is, issuing the forcetakeover command, is required as a safety precaution against a *split brain* scenario.

Important: Site-specific disasters are not the same as a normal cluster failover.

3.7 MetroCluster configurations

For supported configurations see the MetroCluster configuration guide at:

ftp://service.boulder.ibm.com/storage/nas/nseries/nseries_metrocluster.pdf

3.8 Cable selection

Table 3-2 is a cable selection chart that can help you to determine whether a desired cable length is within the supported maximum specification. However, because there are many variables (such as types of cables, panels, and so forth), the best way to determine the optimum cable type is by testing the cable length in the environment.

Table 3-2 Cable selection chart

Cable type	Fiber core type	Mode	Wave length	Maximum distance (m)	Attenuation (Db/Km)	Maximum channel attenuation	Splice loss	Connector pair loss
1 Gbps								
OM2	50/ 125 um	Multi	850	550	3.00	3.25	0.3	0.75
OM3	50/ 125 um	Multi	850	550	3.00	3.50	0.3	0.75
OS1	9/125 um	Single	1310	2000	0.40	7.80	0.3	0.75
2 Gbps								
OM2	50/ 125 um	Multi	850	550[a]	3.00	2.62	0.3	0.75
OM3	50/ 125 um	Multi	850	550	3.00	3.25	0.3	0.75
OS1	9/125 um	Single	1310	2000	0.40	7.80	0.3	0.75

Cable type	Fiber core type	Mode	Wave length	Maximum distance (m)	Attenuation (Db/Km)	Maximum channel attenuation	Splice loss	Connector pair loss
4 Gbps								
OM2	50/ 125 um	Multi	850	150	3.00	2.06	0.3	0.75
OM3	50/ 125 um	Multi	850	150	3.00	3.00	0.3	0.75
OS1	9/125 um	Single	1310	500	0.40	7.80	0.3	0.75
IB 1X 250 MBps								
OM2	50/ 125 um	Multi	850	250	3.50	2.38	0.3	0.75
OM3	50/ 125 um	Multi	850	500	3.50	3.25	0.3	0.75

a. According to the maximum channel attenuation (2.62 dB) and an attenuation of 3.00 dB/km, the maximum distance for this cable type is 406 m, so be careful with longer distances (up to 550 m) with OM2 cable at 2 Gbps.

Table 3-2 summarizes data related to optical cabling for data communications that is available in documents published by various standards organizations. We focus on data that is relevant to fiber deployments, which are supported on N series systems.

To determine whether the desired cable run length is within the supported maximum specification:

- Determine the needed transfer rate based on the type of shelf that you use.
- Find out what fiber type is installed for the system.
- Determine the number of connectors in the path between the nodes.

Consider a client who wants to run an FC cable over approximately 260 m (850 ft) and wants to run it at 2Gbps.

In this example, also assume:

- The desired transfer rate is 2 Gbps.
- Fiber core type is 50/125 OM2 Multimode cabling
- There are two connector pairs.

The table shows that 260 m is within the operating maximum distance for all Fibre Channel transfer rates. This also assumes that the fiber connection is point-to-point with only source and destination connections and no patch panels or splices.

3.9 SyncMirror setup

To see the volume /plex/raidgroup relationship, use the **`sysconfig -r`** command, as shown in Figure 3-9. Use the **`aggr mirror`** command to start mirroring the plexes.

```
Volume spiel (online, normal, mirrored) (zoned checksums)
  Plex /spiel/plex0 (online, normal, active)
    RAID group /spiel/plex0/rg0

      RAID Disk Device          HA   SHELF BAY CHAN   Used (MB/blks)   Phys (MB/blks)
      --------- ------          ---------------- ----  --------------   --------------
      parity    3.2             3      0    2  FC:A   8579/17570816    8683/17783112
      data      3.3             3      0    3  FC:A   8579/17570816    8683/17783112

  Plex /spiel/plex2 (online, normal, active)
    RAID group /spiel/plex2/rg0

      RAID Disk Device          HA   SHELF BAY CHAN   Used (MB/blks)   Phys (MB/blks)
      --------- ------          ---------------- ----  --------------   --------------
      parity    6.8             6      1    0  FC:A   16979/34774016   17560/35964296
      data      6.9             6      1    1  FC:A   8579/17570816    17560/35964296
```

Figure 3-9 Viewing the volume status

3.10 Failure scenarios

It is difficult to predict which component of a MetroCluster configuration might fail. The following examples illustrate possible failure scenarios and the resulting configurations when using MetroCluster.

3.10.1 MetroCluster host failure

In this scenario, host 1 is lost but access to data continues uninterrupted. Host 2 continues accessing data (Figure 3-10).

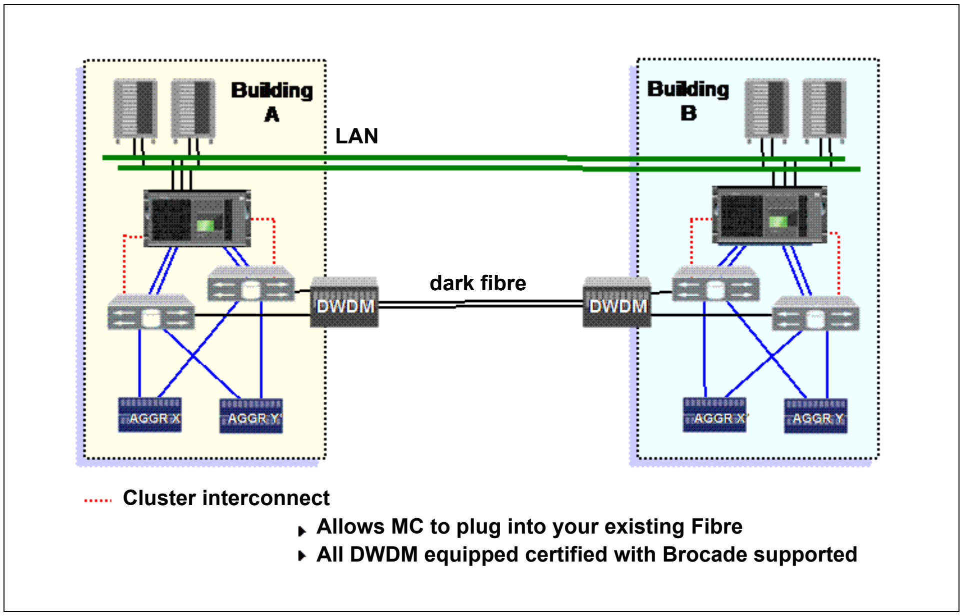


Figure 3-10 Host failure

3.10.2 MetroCluster IBM System Storage N series failure

In this scenario, N series N1 has been destroyed by environmental factors. N series takes over access to its disks (Figure 3-11). The fabric switches provide the connectivity for the N series N2 and the hosts to continue to access data without interruption.

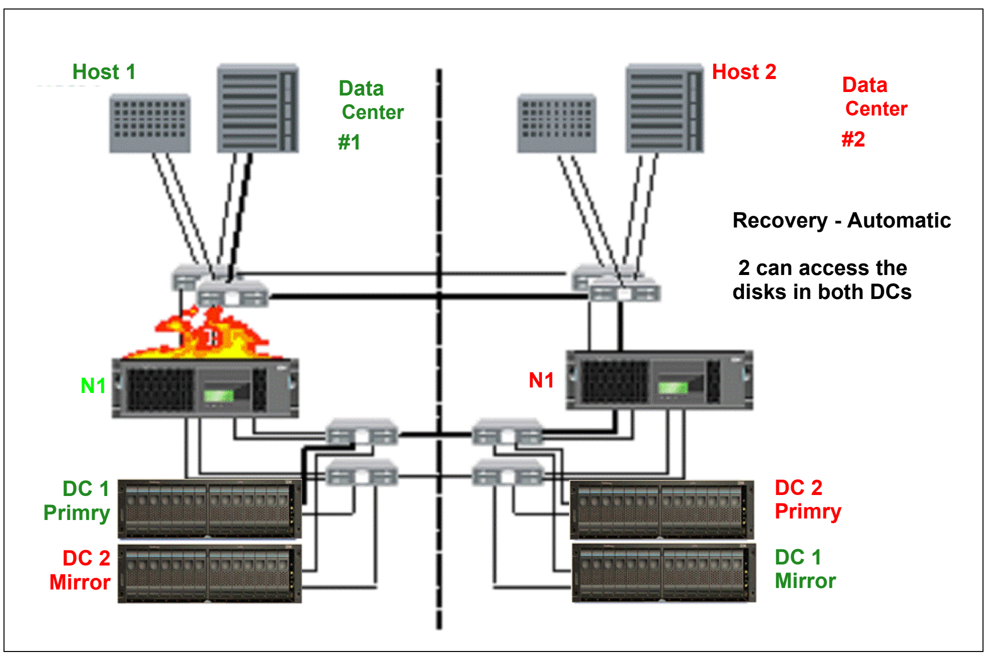


Figure 3-11 IBM System Storage N series failure

3.10.3 IBM System Storage N series and expansion unit failure

This scenario is more likely when a catastrophic action has occurred to a rack with an N series and its expansion units. Here both the N series and the expansion units have become unavailable (Figure 3-12).

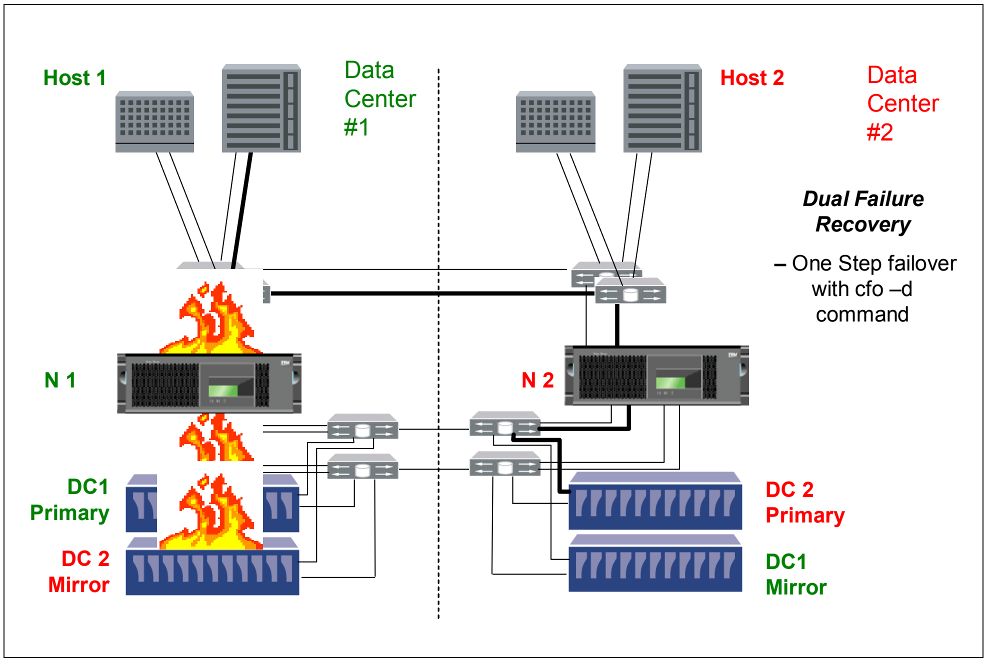


Figure 3-12 Controller and expansion unit failure

In order to continue access, a cluster failover must be performed by using the **`cfo -d`** command. Data access is restored because DC1 mirror was synchronized with DC1 primary. Through connectivity provided by the fabric switches, all hosts will be able to once again have access to required data.

3.10.4 MetroCluster interconnect failure

In this scenario, the fabric switch interconnects have failed (Figure 3-13). Although this is not a critical failure, resolution should occur promptly in case of a more critical failure.

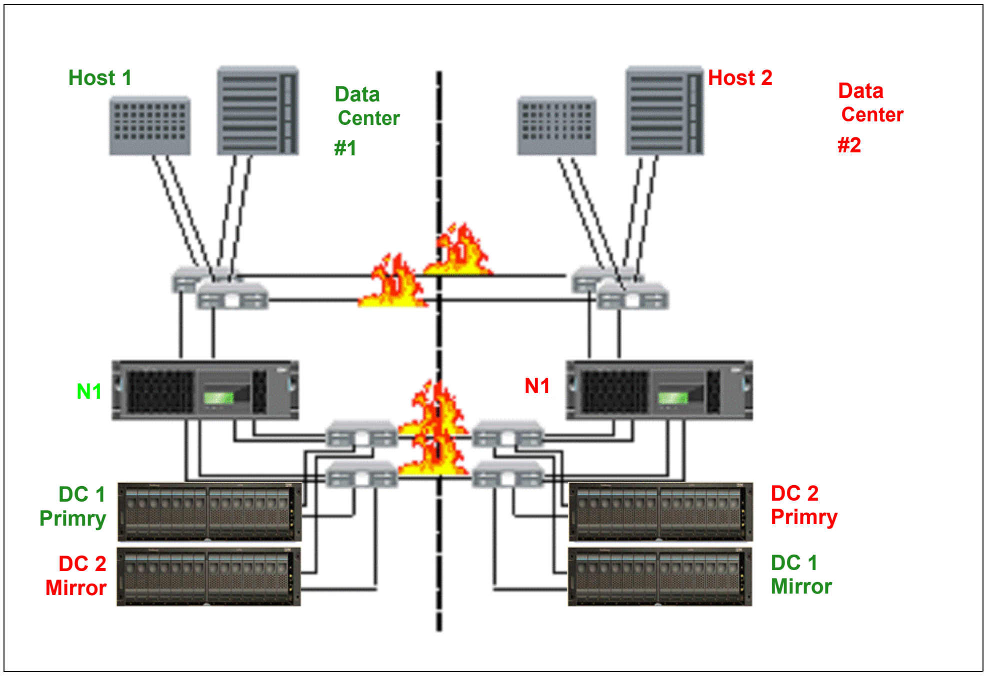


Figure 3-13 Interconnect failure

During this period, data access is uninterrupted to all hosts. However, mirroring and failover are disabled, thus reducing data protection. When the interconnect failure is resolved, re-syncing of mirrors occurs.

3.10.5 MetroCluster site failure

In this scenario, a site disaster has occurred and all switches, storage systems, and hosts have been lost (Figure 3-14). To continue data access, a cluster failover must be initiated by using the **`cfo -d`** command. Both primaries now exist at data center 2, and hosting of Host1 is also done at data center 2.

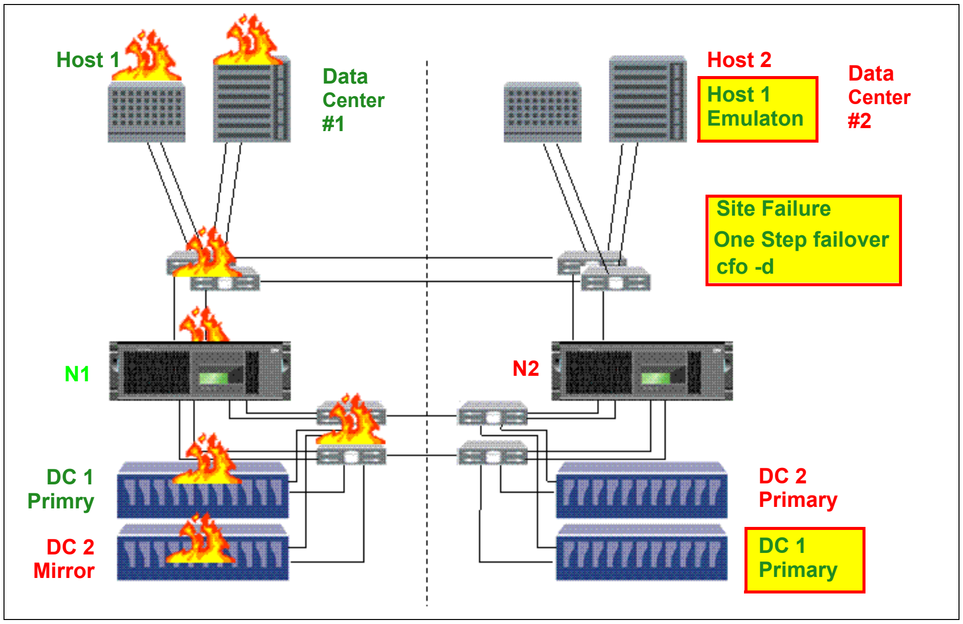


Figure 3-14 Site failure

Note: If the site failure is staggered in nature and the interconnect fails before the rest of the site is destroyed, there is a chance of data loss. This occurs because processing has continued after the interconnect has failed. Typically, site failures occur pervasively and at the same time.

MetroCluster site recovery

After the hosts, switches, and storage systems have been recovered at data center 1, a recovery can be performed. A **cf giveback** command is issued to resume normal operations (Figure 3-15). Mirrors are resynchronized and primaries and mirrors are reversed to their previous status.

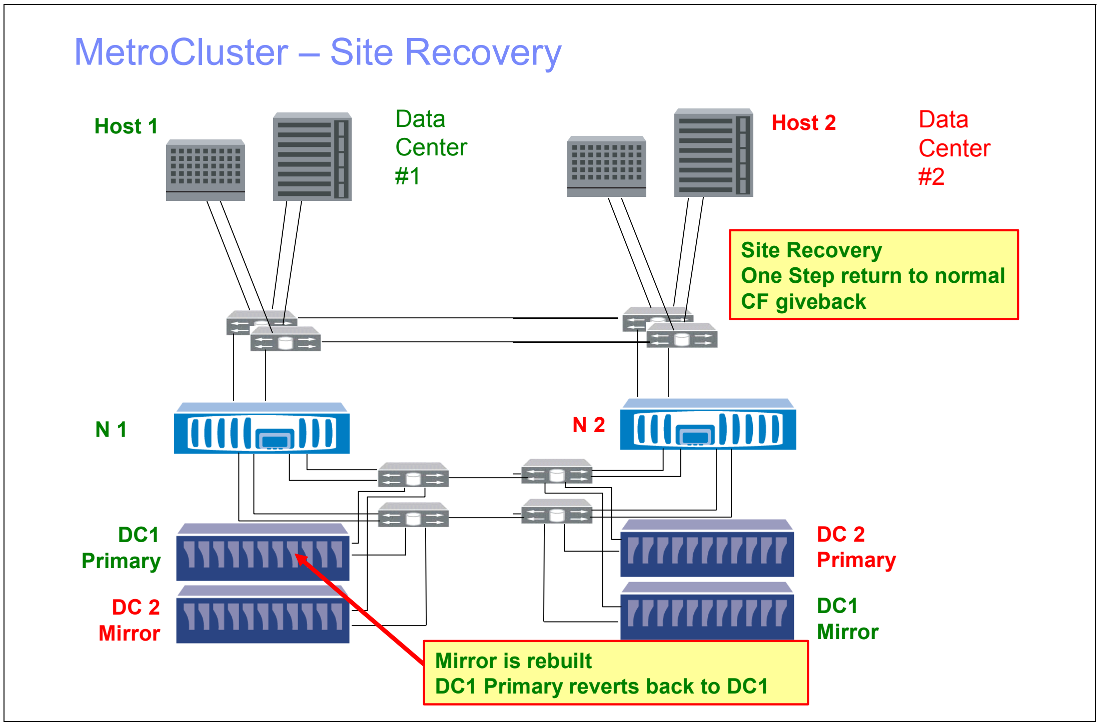


Figure 3-15 MetroCluster recovery

4

FlexScale and the Performance Accelerator Module (PAM)

This chapter provides an overview of FlexScale, which includes the Performance Accelerator Module (a hardware component) and Write Anywhere File Layout (WAFL, an extended cache software component).

The following topics are covered:

- Overview of FlexScale
- Performance Accelerator Module (PAM II cards)
- FlexScale functionality and configuration
- Checking FlexScale configuration

4.1 Overview of FlexScale

FlexScale provides a refined and efficient design with additional WAFL cache memory to improve the performance of IBM System Storage N series storage subsystems by reducing disk reads.

This solution consists of a hardware and software component, namely the Performance Accelerator Module and the extended WAFL cache memory, which is configured by using the command `options flexscale`.

The traditional technique of improving performance by increasing the spindle count of an aggregate (or traditional volume) has proven to be crude and inefficient because drives' capacity for growth has outpaced their performance.

> **Note:** This solution is suitable for all types of workloads but would provide the greatest benefit from IBM System Storage N series storage subsystems serving intensive random read transactions.

4.2 Performance Accelerator Module

A PAM module provides an additional 16 GB (PAM I), 256 GB, or 512 GB (PAM II) of extended cache for your IBM System Storage N series storage subsystem, depending on the model. Up to five modules can be installed. Each module must be installed on a PCI express slot and only consumes an additional 18 W of power (per module). Extra rack space and ventilation is not required, thus making it an environmentally friendly option (Figure 4-1).

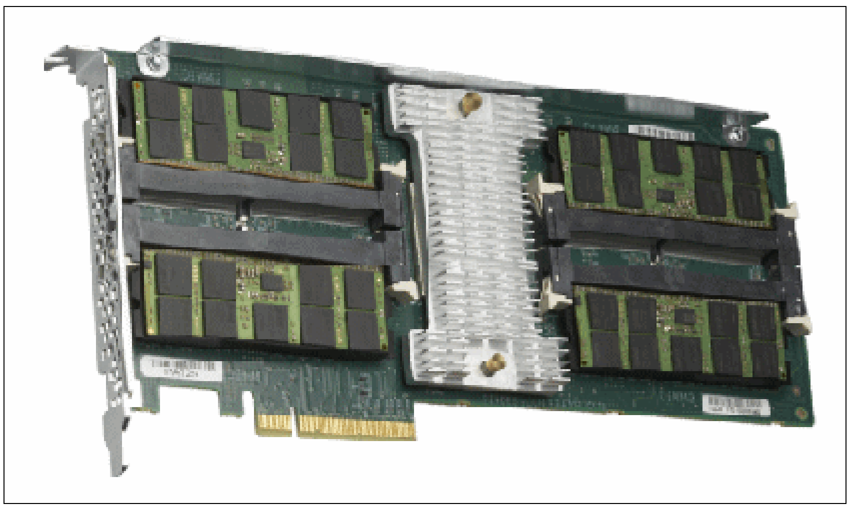


Figure 4-1 Performance Accelerator Module (PAM I)

4.2.1 How the PAM module accelerates performance

The IBM System Storage N series storage subsystem fulfills any read requests from memory before seeking the data from disks. On a read-intensive system, the read cache (system memory) can be rapidly depleted from multiple read requests (see Figure 4-2). FlexScale addresses this issue by extending the available WAFL cache.

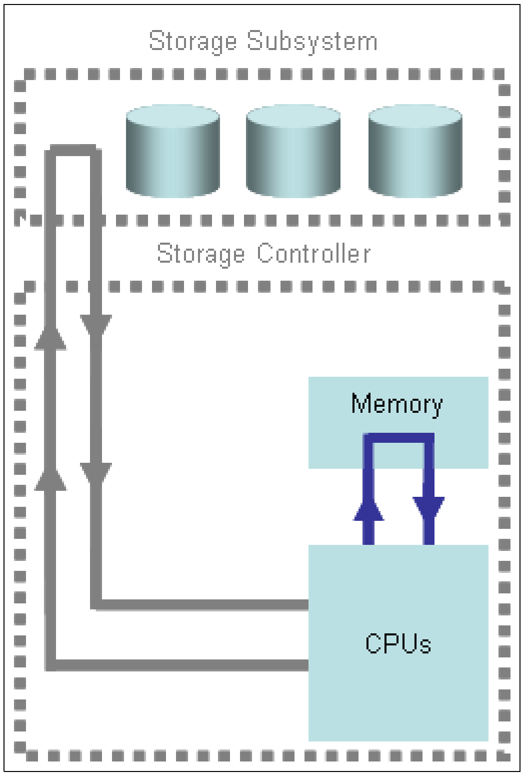


Figure 4-2 Read request without PAM module installed

With the PAM module installed, any read requests are served from the module if the data is not available from memory. The disk is read only when the request cannot be fulfilled by either cache. Seeking data from the PAM module is slower than system memory but it is still significantly faster than disk.

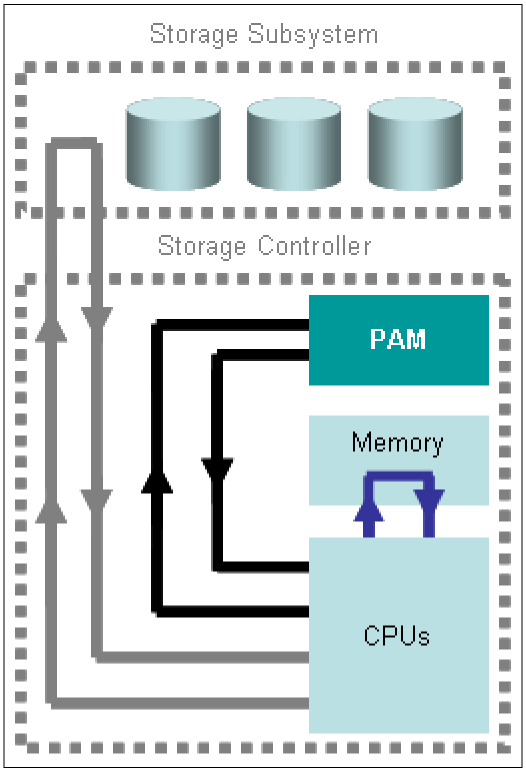


Figure 4-3   Read request with PAM module installed

Typically, cached data is purged when additional storage is required on the system memory. With the PAM module installed, cached data is relocated there. Any future read request will be fulfilled by the module (for the same data), thereby improving the storage subsystem response time (Figure 4-4).

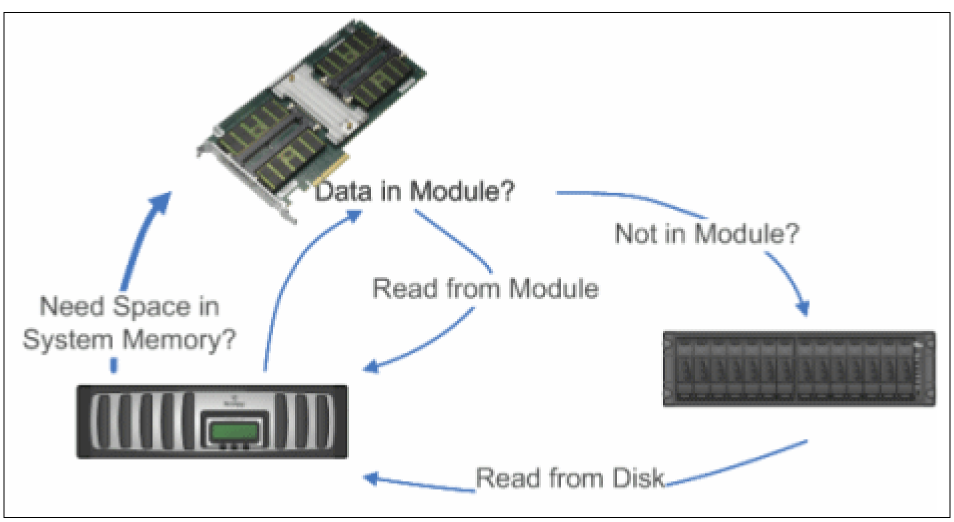


Figure 4-4   Additional storage for WAFL extended cache

4.2.2 Prerequisites for using the PAM II module

The PAM II module is only supported on:

- Systems running Data ONTAP Version 7.3 or later.
- The following N series and gateway models:
  - N5600
  - N6040
  - N6060
  - N6070
  - N7600
  - N7700
  - N7800
  - N7900

  It must be installed into the PCI-e slots.
- An active/active configuration. An equal number of the PAM II modules must be installed on both nodes.

The PAM II module is suitable for all file and SAN environments, and is supported on all Data ONTAP software 7.3 and later.

	Original PAM	PAM II
Size per Card	16GB	256GB, 512GB
Memory Technology	DRAM	SLC NAND Flash
Max. Cache in N7000	160GB	2TB

Figure 4-5 PAM capacities

4.2.3 Supported N series configuration with PAM II cards

Table 4-1 provides a list of supported systems and configurations of PAM II modules.

Table 4-1 Supported systems

Filer/gateway	PAM II modules supported	Extended read cache
N7800 and N7900	4 x 512 GB x node	2 TB
N7700 and N6070	2 x 512 GB x node	1 TB
N6060 and N5600	2 x 256 GB x node	512 GB
N6040	1 x 256 GB x node	256 GB

Cache extension with PAM II

The SLC NAND flash-based PAM II joins the DRAM-based PAM I card, which was already introduced with the N5200 and N5500. Each PAM II card is up to 512 GB, with up to 4 cards in PCI Express slots in the controller (meaning no additional rack space is consumed). PAM family modules are intelligent, level 2 read caches for the storage controller. They cache metadata and user data with the aim of avoiding disk reads and the added latency and limited I/O throughput that result from such reads.

Some potential for confusion exists in comparisons between the PAM implementation and other PCIe-based solid state cards. Whether solid state technology is used as persistent storage or as cache, the architectural context matters: PAM is used in SAN and NAS environments, while typically other PCIe cards are used in a DAS (that is, server-attached) context. Another key distinction is functionality. PAM is a dedicated, intelligent read cache that is tightly integrated with the Data ONTAP operating system, while server-based flash cards are generally used for persistent storage.

4.3 FlexScale functionality and configuration

The WAFL extended cache software component is a licensed feature that enables user data blocks to be cached. FlexScale is available on Data ONTAP version 7.3 and later and supports a maximum of 127 aggregates (or traditional volumes). If there are more than 127 aggregates on the IBM System Storage N series storage subsystem, only data from the first 127 aggregates is cached.

> **Note:** if you enable the FlexScale option without a license or PAM module installed, it runs in *predictive cache statistics* (PCS) mode. This mode provides valuable statistics to assist with planning and configuration of a FlexScale deployment.

FlexScale modes and their functionality

There are three modes in which the WAFL extended cache handles user data blocks on the IBM System Storage N series storage subsystem:

- Caching normal user data

  In this mode, the WAFL extended cache saves normal user data blocks. The following option should be enabled:

  ```
  options flexscale.normal_data_blocks on
  ```

  If the option is set to off, only metadata blocks are cached.[1]

- Caching only system metadata

  If the working set of the storage subsystem is very large, such as a large email server, you can cache only the system metadata in WAFL extended cache by turning off both lopri_blocks and normal_data_blocks[1]:

  ```
  options flexscale.lopri_blocks       off
  options flexscale.normal_data_blocks off
  ```

- Caching low priority user data

  Low priority user data blocks that are read in large sequential scans and blocks written to the storage subsystem through a network-attached storage (NAS) protocol such as Network File System (NFS) are normally not cached by FlexScale. To enable caching of

[1] With the exception of volumes with the FlexShare cache setting set to keep (priority set volume myvol cache=keep).

the low priority user data blocks, both options flexscale.normal_data_blocks and flexscale.lopri_blocks must be enabled:

```
options flexscale.lopri_blocks       on
options flexscale.normal_data_blocks on
```

4.4 Checking FlexScale configuration

Data ONTAP enables you to display information about FlexScale configuration, usage, and access information.

4.4.1 Display FlexScale configuration

To display FlexScale information, enter the following in the CLI:

```
stats show -p flexscale
```

The output is as follows (PCS mode shown):

```
FlexScale Configuration
ext_cache:ext_cache:state-string:Active
ext_cache:ext_cache:cache_objects:3
ext_cache:ext_cache:block_checksums:0
ext_cache_obj:ec0:type:PCS
ext_cache_obj:ec0:blocks:1572864
ext_cache_obj:ec0:usage:33%
ext_cache_obj:ec1:type:PCS
ext_cache_obj:ec1:blocks:1572864
ext_cache_obj:ec1:usage:11%
ext_cache_obj:ec2:type:PCS
ext_cache_obj:ec2:blocks:3145728
ext_cache_obj:ec2:usage:0%
---
```

4.4.2 Display FlexScale usage and access information

Information regarding FlexScale usage and access is useful if produced and collated periodically, which can be achieved by entering the following command in the CLI:

```
stats show -p flexscale-access [-i interval] [-n num]
```

- If no options are used, a single 1-second snapshot of statistics is used.
- -i *interval* specifies that output is to be produced periodically, with an interval of *interval* seconds between each set of output.
- -n *num* terminates the output after *num* number of iterations, when the -i option is also used. If no *num* value is specified, the output runs forever until a user issues a break in pressing by pressing Ctrl+C.

Example 4-1 shows output from the **stats show -p flexscale-access -i 1 -n 10** command.

Example 4-1 Output from -p flexscale-access -i -n 1- command

```
Cache                                                Reads        Writes
Usage    Hit    Miss    Hit Evict   Inval Insert Chain Blocks Chain Blocks
    %     /s      /s      %    /s      /s     /s    /s     /s    /s     /s
   37      0       0      0     0       0      0     0      0     0      0
   37      0       0      0     0       0      0     0      0     0      0
   37      1       0    100     0       0      0     0      0     0      0
   25    122      12     91     0  187158      0     0      0     0      0
   12     23       1     95     0  215923      0     0      0     0      0
   12      0       0      0     0       0      0     0      0     0      0
   12      0       0      0     0       0      0     0      0     0      0
   12      0       0      0     0       0      0     0      0     0      0
   12      0       0      0     0       0      0     0      0     0      0
   12      0       0      0     0       1      0     0      0     0      0
```

5

Multipathing with the IBM System Storage N series

In this chapter, we introduce the concepts of multipathing. We briefly outline the installation steps and describe the management interface for the Windows, Linux, and AIX operating systems (OS).

The following topics are covered:

- Overview
- Data ONTAP for Windows MPIO
- Installation of IBM Data ONTAP DSM
- Managing DSM using the GUI
- Managing DSM using the CLI
- Multiple path I/O support for Red Hat Linux
- Multiple path I/O support for Native AIX O/S

5.1 Overview

Multipathing I/O (MPIO) provides multiple storage paths from hosts (initiators) to their IBM System Storage N series targets. The multiple paths provide redundancy against failures of hardware like cabling, switches, and adapters, and also provide higher performance thresholds by aggregation or optimum path selection.

Active/active, and active-passive, are usually used to categorize multipathing. However, current midrange storage solutions support multiple active and multiple passive paths simultaneously. A multipathing solution is generally considered to be active/active when an I/O for a single LUN travels multiple paths simultaneously to its destination.

The following settings are used to decide how to spread the I/O across the available paths:

- Active-passive: A single path is used, and the passive paths are only used when the active path is unavailable.
- Round robin: The I/O is spread evenly across the available paths using a round-robin algorithm.
- Least queue depth: The path chosen for an I/O is the available path that has the smallest number of outstanding I/Os.
- Fewest bytes: The path chosen for an I/O is the path with the lowest number of outstanding bytes.
- Weighted path: Each path is given a weight, which is used to determine which path should be used for an I/O.

5.2 Data ONTAP for Windows MPIO

Data ONTAP device-specific module (DSM) works with Microsoft Windows MPIO drivers to manage multiple paths between the N series storage subsystems and Windows hosts, providing storage-system-specific intelligence to correctly identify paths and managing path failure and recovery.

DSM Version 3.2 is still available. Version 3.3.1 is the latest version, designed for Windows 2003 and Windows 2008 (R2 with DSM v3.3.1) and requiring ONTAP 7.2.2 or later to be running on the storage subsystem. The latest Fibre Channel (FC) and iSCSI Host Utilities kit is recommended and available for registered users at:

http://www-304.ibm.com/systems/support/supportsite.wss/supportresources?brandind=5000029&familyind=5364556&taskind=1

To upgrade a host to Windows 2008 R2, you must upgrade to Windows Host Utilities 5.2 and Data ONTAP DSM 3.3.1 for Windows MPIO before upgrading to Server 2008 R2.

> **Note:** MPIO and iSCSI initiator are built into Windows 2008. Therefore, it is not required to be installed separately.

The following new features are offered in Version 3.2:

- New load-balancing policies
- Support for claiming iSCSI LUNS
- Support for enhanced co-existence with other DSMs
- Multi-Protocol LUN support (iSCSI and Fibre Channel Protocol (FCP) paths to the same LUN)

5.2.1 New load-balancing properties

The two new policies introduced in DSM 3.2 for enhancing support of load balancing for iSCSI and FC LUNS are:

- Least queue depth
- Least weighted paths

DSM 3.2 and 3.3.1 continue to support existing policies from the earlier versions, such as round-robin, round-robin with subset, failover only, and auto-assigned.

5.2.2 Support for claiming iSCSI LUNS

The Data ONTAP DSM 3.2 and 3.3.1 now support claiming LUNS with iSCSI paths. You can view all the iSCSI LUNS along with the following characteristics:

- State of the path (active or passive)
- DSM ID
- Target adapter IP address
- Preferred path elected by the policy

5.2.3 Support for enhanced co-existence with other DSMs

If more than one DSM co-exists within a host, it causes constraints in setting some of the MPIO parameter's values because all these settings are tested and tuned for a specific storage vendor and differ from other DSM providers. Since all these settings were previously stored in the registry on a global basis, all DSM providers had to share a common registry location, which led to co-existence issues.

DSM 3.2 now integrates with MPIO DDK 1.20, which supports a feature that allows a vendor DSM to have its own set of parameters on a non-global basis. The installer sets the DSM-specific parameter in the registry and ensures that N series defined parameters are used.

The Data ONTAP DSM 3.3.1 for Windows MPIO uses MPIO version 1.22 for Windows Server 2003.

> **Note:** The native Microsoft msdsm will not be used when ONTAP DSM is installed on Windows Server 2008. Subsequently, any N series LUNs will be claimed by ONTAP DSM.

5.2.4 Multi-protocol LUN support (iSCSI and FCP paths to the same LUN)

Multi-protocol LUN support is the ability to utilize both FCP and iSCSI to communicate with a LUN. The ONTAP DSM can claim devices from both protocols and associate them with the same MPIO pseudo LUNS. Load balancing then occurs between the three different path types. The DSM maintains internal weights for the paths based on these attributes' preference: Non-proxy FCP, proxy FCP, and iSCSI.

> **Note:** The default cfmode on a new N series system is single_image from Data ONTAP Version 7.2 or later. Ensure that the cfmode is set to single_image if you are upgrading from previous versions. To support both FCP and iSCSI paths to the same LUN, the N series storage subsystem must be running single_image cfmode.
>
> **Note for N3700:** dual_fabric mode is the only available cfmode for this N series model.

Path limits

A maximum of 32 paths are permitted to a LUN. This applies to any mix of FCP and iSCSI protocols, which is a Windows MPIO layer limitation. Any path exceeding the limit will not be claimed by DSM, and could lead to unpredictable behavior and possible data loss.

Ease of DSM management

The Data ONTAP DSM 3.2 can be invoked using the CLI or GUI and is integrated with the Windows 2003 and 2008 computer management Microsoft Management Console (MMC). On Windows 2008, it can be invoked from the storage management node's server manager console, which is the default interface available at startup to manage domain-wide roles and features.

5.2.5 Failover and load balance policy concepts

Active I_T_L nexuses (refer to 5.2.6, "I_T and I_T_L nexus" on page 149, for a working definition of a nexus) between the LUN and Windows hosts are chosen by ONTAP DSM based on the following:

- Load balance policy of the LUN
- Whether the path is local or proxy
- Whether the path is FCP or iSCSI
- State of all possible paths
- Load on each path

Non-proxy or optimized FCP paths are chosen by default whenever available. When no optimized FCP paths are available, proxy or non-optimized FCP paths are chosen. Only when no FCP paths are available will iSCSI paths be chosen. This default DSM algorithm can be altered by selecting an alternative load balance policy.

The six load balance policies that can be selected are:

- Least queue depth: I/O to the LUN is sent to an available path with the smallest outstanding queue length. The queue is determined at the I_T nexus level, not the LUN path level. Administrator intervention is not require to maximize the bandwidth utilization with this policy. The default DSM path selection is made obvious in this policy.
- Least weighted paths: User-configurable available path weight is applied to the LUN, with zero being the highest priority (lowest weight). A maximum value of 2146999999 (lowest priority) can be set.
- Round robin: The round robin policy is an active/active policy. All optimized (non-proxy) paths to the storage system are used when available.
- Round robin with subset: Like the round robin policy, this policy uses multiple paths. However, optimized standby paths can be specified and paths can be chosen when they are available.
- Although you can specify non-optimized (proxy) paths as part of the active subset, this is not recommended. By default, all optimized paths are initially selected. To specify the subset, make individual paths active or passive as desired.

- Failover only: The failover only policy enables you to manually select a single preferred I_T_L nexus. This I_T_L nexus is used whenever it is available.
- Auto assigned: The auto assigned policy attempts to spread the load evenly across all available local paths. For each LUN, only one path is used at a time. Whenever a path changes state, the DSM automatically rebalances the load.

5.2.6 I_T and I_T_L nexus

An initiator-target (I_T) nexus represents the path from the host's host bus adapter (HBA) (initiator) to the storage systems's HBA (target). An initiator-target-LUN (I_T_L) nexus represents one virtual disk (LUN) as seen by the DSM. The DSM groups all I_T_L nexuses to the same LUN together, and presents a single virtual disk to the Windows disk manager.

The I_T_L nexus is assigned an 8-character DSM identifier. The identifier is made up of four fields:

- Port
- Bus
- Target
- LUN

Table 5-1 shows the Initiator_Target_Lun (I_T_L) nexus composition.

Table 5-1 Initiator_Target_Lun (I_T_L) nexus constitution

DSM ID	Port	Bus	Target	LUN
03000101	03	00	01	01

Each path (I_T nexus) also has an 8-character identifier made up of four fields. The first three fields are identical to the DSM ID. The fourth field indicates the path type:

- 01 for optimized FCP
- 02 for non-optimized FCP
- 03 for iSCSI path

Table 5-2 shows the Initiator-Target (I_T) nexus composition.

Table 5-2 Initiator-Target (I_T) nexus constitution

Nexus ID	Port	Bus	Target	Path Type
03000101	03	00	01	01 - non-proxy FC
03000202	03	00	02	02 - proxy FC
05000103	05	00	01	03 - iSCSI

5.2.7 Supported cluster mode settings for FCP

Table 5-3 lists the FCP cfmode setting required on the storage system for each load balance policy. Data ONTAP 7.2.2 or later is required.

Note: Mixed FCP and iSCSI access to the same LUN requires the single_image cfmode setting. For iSCSI-only access, the cfmode setting does not matter.

Table 5-3 Supported FCP cluster modes and load balance policies

Policy	Supported cfmode values
Auto assigned	single_image, standby, partner, dual_fabric
FailOver only	single_image, standby, partner, dual_fabric
Least weighted paths	single_image, standby, partner, dual_fabric
Round robin	single_image, standby
Round robin with subset	single_image, standby
Least queue depth	single_image, standby

The FCP cfmode setting required on the storage system for each load balance policy with Data ONTAP 7.2.2 or later and DSM 3.3.1 is unique: single_image.

Note: Although the standby, partner, and dual-fabric cfmode values are supported, we strongly recommend the single_image cfmode setting.

5.2.8 Performance factors

The following performance considerations (depending on your environment) can help you achieve improved operation of your N series:

- About optimized paths: The DSM differentiates between non-proxy and proxy FCP paths. Proxy FCP paths use the cluster inter-connectivity between the active/active configuration in the controller of the N series storage subsystems, which is less efficient than local non-proxy FCP paths.

  By default, optimized paths are used unless they are unavailable or it is explicitly set in load balance policies like failover only or round robin with the subset policy, or the proxy paths have a lower weight set in the least weighted paths policy.

- Windows clusters: Special rules are applied in the case of Windows cluster configurations, If the state of any path changes (available becomes unavailable or vice versa), a 2-minute wait is given for the I/O to stabilize before moving LUNs owned by a clustered Windows host to balance the load among paths.
- Disabled paths: Any manually disabled I_T_L nexus will not be utilized (failed over upon) unless all enabled I_T_L nexus are unavailable. But once the active I_T_L nexus is restored, DSM fails back to its original nexus status quo.

  DSM 3.3.1 can use a disabled path.

5.3 Installation of IBM Data ONTAP DSM

In this section we briefly outline the steps to prepare for and perform the installation of ONTAP DSM. We assume that your existing host has a working and supported configuration with multiple physical paths to your N series storage. Consult the IBM website to qualify your configuration before commencing any software installation.

There are four Fibre Channel host bus adapters (FC HBAs) installed on our Windows 2003 host, as shown in Figure 5-1.

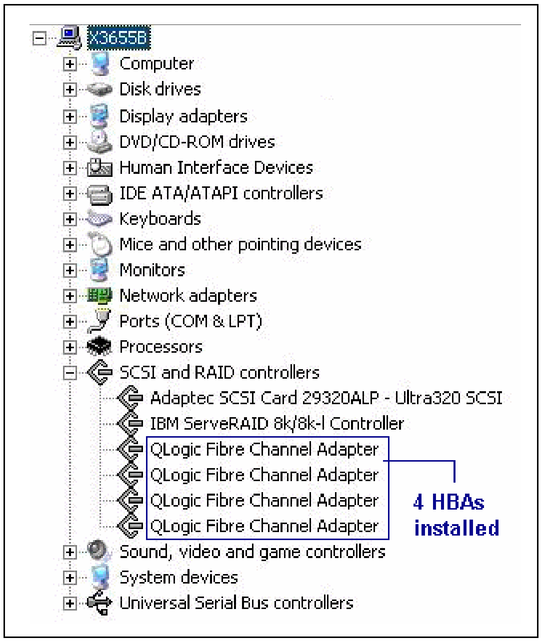


Figure 5-1 Four HBA adapters installed on Windows host

The four HBAs would cause the same LUN to appear as four distinct disks in the MMC Disk Management, as shown in Figure 5-2. Data corruption would occur if each of the four disks were formatted and used. Data ONTAP DSM will prevent that from happening.

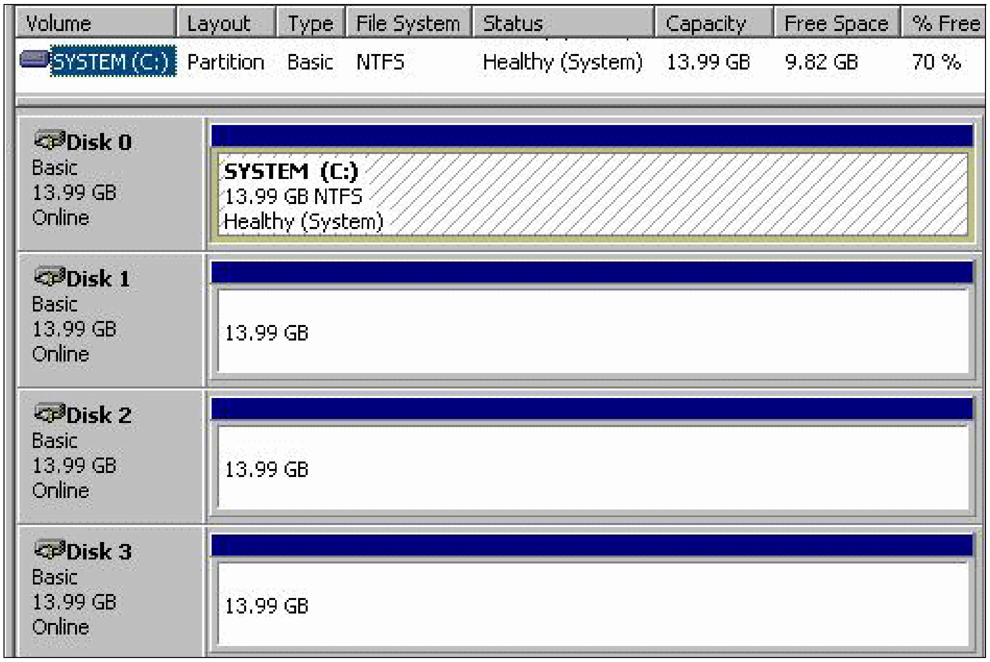


Figure 5-2 Identical LUN presented four times on Windows host

5.3.1 Installation of IBM Windows Host Utilities Version 5.x

Download the Windows Host Utilities software from the IBM website and follow the prompts to run the file. This installs the software and updates the Windows registry and HBA parameters. A reboot is required to complete the installation.[1] See Figure 5-3.

Note: Consult the IBM website for the latest hotfixes for your host.

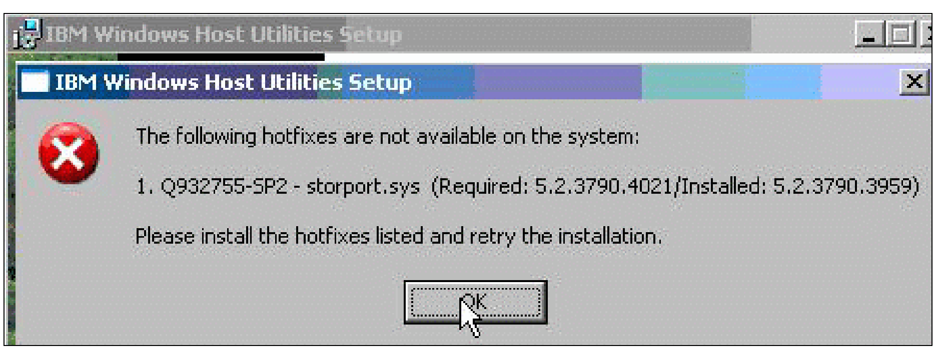


Figure 5-3 Prerequisite hotfix required for Windows Host Utilities software

Example 5-1 shows an unattended installation of Windows Host Utilities.

[1] Verify that your entire configuration is supported, including Windows operating system version, service pack level, and required hotfixes; HBA model and firmware version; Fibre Channel switch model and firmware version; iSCSI initiator; multipathing software; Data ONTAP version and cfmode setting; and option software such as SnapDrive® for Windows.

Example 5-1 Unattended installation of Windows Host Utilities

```
You may choose to install the software unattended; instructions as follows:
   msiexec /i installer.msi /quiet
   PROTOCOLS={0 | 1 | 2 | 3}
   MULTIPATHING={0 | 1}
   [INSTALLDIR=inst_path]
installer is the name of the .msi file for your CPU architecture
PROTOCOLS specifies which protocols are supported. Allowed values are 0 for none,
1 for FCP only, 2 for iSCSI only, and 3 for both.
MULTIPATHING specifies whether MPIO support is installed. Allowed values are 0 for
no, for yes.
inst_path is the path where the Host Utilities files are installed. The default
path is C:\Program Files\NetApp\Windows Host Utilities\.
```

FCP and iSCSI protocols are both supported in the latest version of the Host Utilities software, as shown in Figure 5-4. You can reinitiate the installation from the Windows Add or Remove programs applet to add or remove any protocols in the future by choosing the **Repair** option.[2]

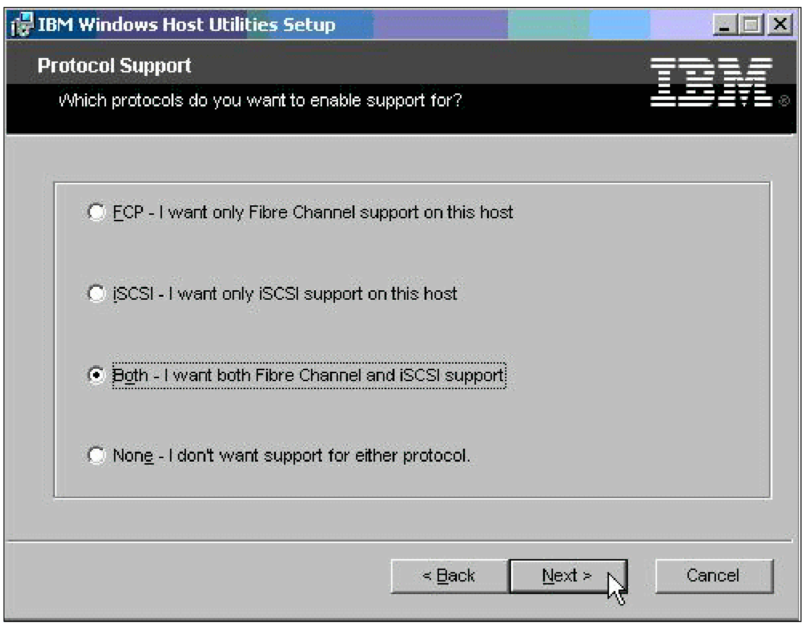


Figure 5-4 Multi protocol supported by Windows Host Utilities software

We selected support for Multpath I/O in our installation, as shown in Figure 5-5. Select the single path option if it is the only valid option for your existing host. Repeat the installation of DSM by selecting **Repair** when the configuration of your systems changes.[2]

[2] You might need to repair settings if you install a new HBA or configure Windows clustering after installing Windows Host Utilities. To repair Windows Host Utilities from your hosts, use the Windows Add or Remove Programs applet and follow the prompts. The repair option updates the Windows registry and Fibre Channel HBAs with the required settings.

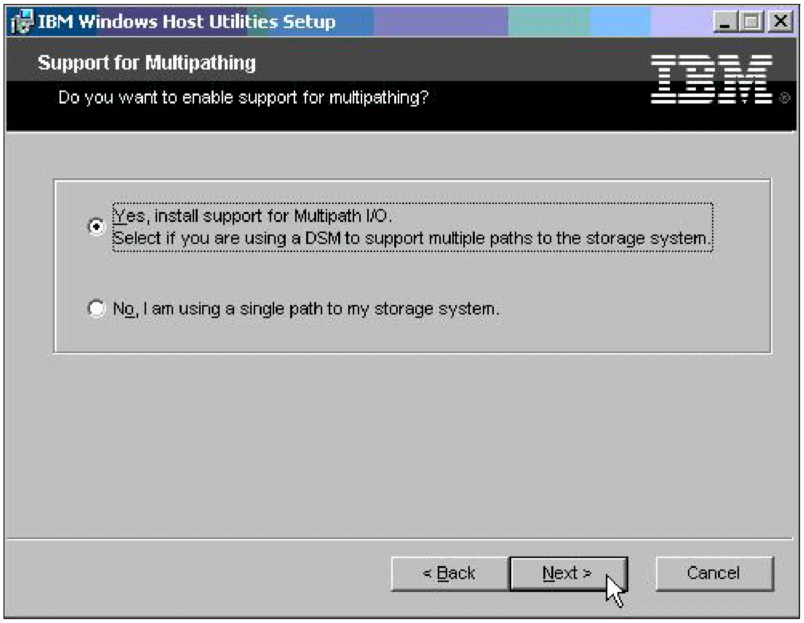


Figure 5-5 Multipath I/O supported by Windows Host Utilities software

> **Note:** The optional Windows multipathing feature on Windows Server 2008 must be installed before installing DSM.

Reboot your host to finalize the installation. Confirm the installation of the Windows Host Utilities by reviewing the disk devices on the MMC Device Manager (Figure 5-6).

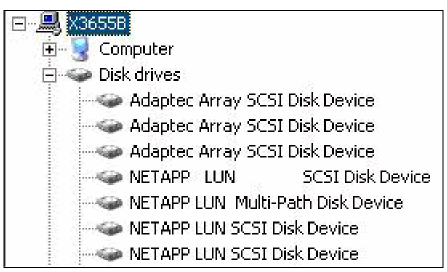


Figure 5-6 Multipath and iSCSI device drivers installed

5.3.2 Data ONTAP DSM for Windows Version 3.2 R1

If you are running an earlier version of the Data ONTAP DSM for Windows MPIO with Windows Server 2003, you must upgrade to DSM 3.2 before upgrading your host from Windows Server 2003 to Windows Server 2008.[3]

Data ONTAP DSM for Windows MPIO works with SnapDrive 5.0 for Windows and later. Check the IBM website for the latest information regarding interoperability between SnapDrive and Data ONTAP:

http://www-03.ibm.com/systems/storage/network/interophome.html

[3] Confirm that you have a supported configuration of the following components from their respective vendor websites: Data ONTAP version on N series storage subsystem; Windows OS version on your host; SnapDrive for Windows software version on your host; FCP Host Utilities software versions on your host; FC HBA model, driver, and firmware versions on your host computer; FC switch model and firmware version; and optional iSCSI software initiator and iSCSI Host Utilities.

If you have an earlier version of SnapDrive on your Windows host, remove it or upgrade it before the DSM is installed. Record the SnapDrive for Windows license before removing SnapDrive 4.1 or earlier from your host. Refer to the installation guide for your specific version of SnapDrive regarding any upgrade or removal procedures.

Note: There are two license keys for SnapDrive:
- LUN management component key
- MPIO component key

5.3.3 Installation steps

Follow these steps to install:

1. Download or copy the appropriate installation file for your Windows host's processor architecture from the IBM website.

   http://www-947.ibm.com/systems/support/supportsite.wss/allproducts?brandind=5000029&taskind=1

2. Stop any I/O to the host, including any I/O using the iSCSI protocol.
3. For hosts running MSCS, stop the cluster service.
4. Launch the installer program and follow the instructions in the window.
5. Enter the 14-character MPIO license key when prompted.
6. Enter the user name and password of the account on the host under which the DSM management service will be logged on. This account must be in the Windows administrators group. Press **OK** in the dialog box that appears, as shown in Figure 5-7.

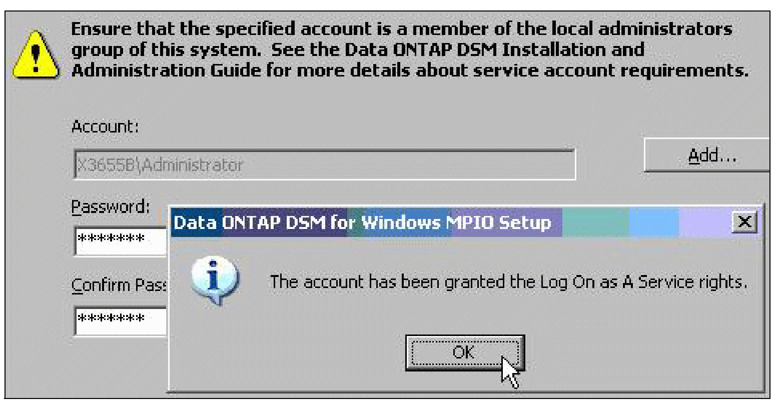


Figure 5-7 Windows account granted for log-on service rights

7. Select the type of LUNs to claim: FCP only, iSCSI only, or both, as shown in Figure 5-8.

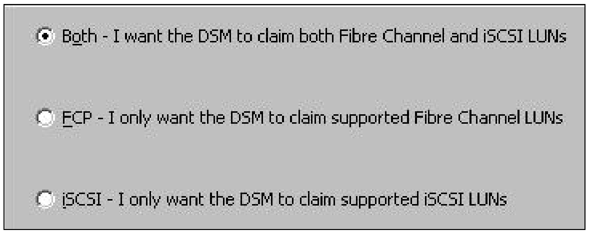


Figure 5-8 Supported protocols on ONTAP DSM

8. When prompted, click **Yes** to reboot the Windows host. The Data ONTAP DSM service appears in the MMC Services applet after the reboot, as shown in Figure 5-9.

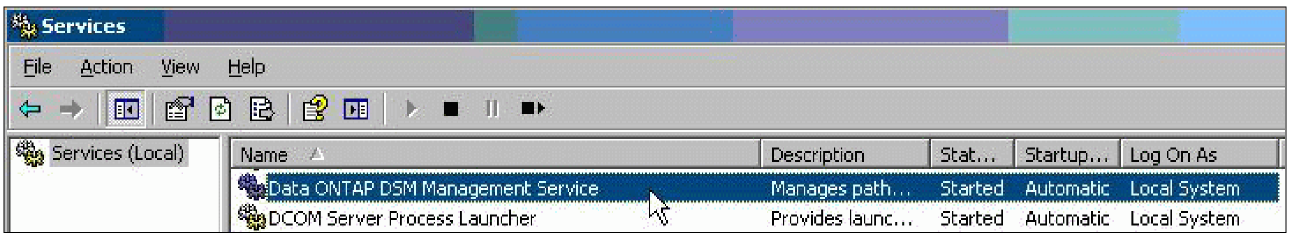


Figure 5-9 ONTAP DSM service running on successful installation

> **Note:** If the installer reports a problem, such as a required hotfix not found, correct the issue and run the installer again.

Example 5-2 shows an unattended installation of Data ONTAP DSM.

Example 5-2 Unattended installation of ONTAP DSM

```
You may choose to install the software unattended; instructions as follows:
   msiexec /package installer.msi /quiet /l*v log_file_name LICENSECODE=key
   SVCUSERNAME=domain\user SVCUSERPASSWORD=password SVCCONFIRMUSERPASSWORD=password
   PROTOCOLS={ 1 | 2 | 3 } [INSTALLDIR=inst_path] [TEMP_FOLDER=temp_path]
installer.msi is the DSM installation program for your Windows host's processor
architecture.
log_file_name is the file path and name for the MSI installer log. Note the first character
of the l*v option is a lower case L.
key is the MPIO license code for the DSM.
domain\user is the Windows domain and user name of an account in the Administrators group
on the Windows host under which the DSM management service will be logged on.
password is the password for the account above.
inst_path is the path where the DSM files are installed. The default path is C:\Program
Files\NetApp\MPIO\.
temp_path is the path where log files are written (except the MSI installer log). The
default path is C:\temp\netapp\.
PROTOCOLS={ 1 | 2 | 3 } specifies which LUN types are claimed by the DSM. Allowed values
are 1 for FCP only, 2 for iSCSI only, and 3 for both.
```

Note: To include the silent install command in a script, use **`start /b /wait`** before the installer.exe command:

```
start /b /wait msiexec /package installer.msi ...
```

The wait option is needed to get the correct installation return value. If you just run installer.msi, it returns `success` if the Windows installer is successfully launched. However, the installation itself may still fail. By using the wait option, the return code describes the success or failure of the actual installation.

The disk referred to in Figure 5-2 on page 152 appears only once after the successful installation of Data ONTAP DSM, as shown in Figure 5-10.

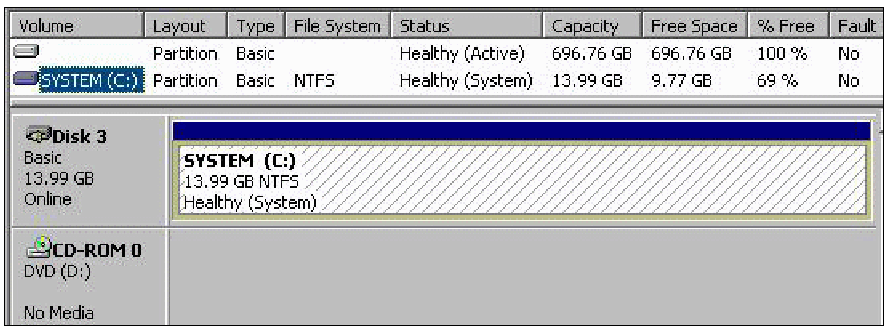


Figure 5-10 One LUN presented with DSM installed (compare with Figure 5-2 on page 152)

5.3.4 Data ONTAP DSM repair option

The repair option is available from the Windows Add or Remove programs applet, which updates the Windows registry settings and copies the DSM and MPIO driver files to their respective location. The repair option can also change the types of LUNs claimed by DSM.

Run the repair options by following these steps:

1. Open the Windows Add or Remove programs applet.
2. Select **Data ONTAP DSM for Windows MPIO**.
3. Select the **Repair** option.
4. Enter the user name used in step 6 on page 155, which DSM used to log on.
5. Verify that the selection for the type of LUNs claimed is correct.
6. Follow the instructions and reboot your host when prompted.

Changing the LUN types claimed by DSM

During your initial installation of DSM, specify the LUN types claimed: iSCSI only, FCP only, or both (step 7 on page 156). To change the types of LUNs claimed after your initial installation requires a reboot of your host, you can either use the repair option or the CLI (Example 5-10 on page 162).

5.4 Managing DSM using the GUI

Data ONTAP DSM can be managed using either the graphical user interface (GUI) or the command-line interface (CLI). This section describes management using the GUI, the use of the CLI is covered in 5.5, “Managing DSM using the CLI” on page 161.

Note: You must manually initiate a rescan to discover any new LUNs added to your storage system. To discover new virtual disks:

1. Click **Start** → **Administrative Tools** → **Computer Management**.
2. Select **Storage** → **Disk Management** in the left pane.
3. The Initialize and Convert Disk wizard starts. Follow the prompts to initialize the disks, but do not convert them to dynamic disks. Partition and format the disks if desired. If the wizard does not start, select **Action** → **Rescan Disks**.

Refresh the GUI to display the new virtual disks DSM of the MMC.

Open the MMC to manage ONTAP DSM. A new section appears when you click **Storage** on the left pane. When the individual disk is selected (Figure 5-11) the multipath information is displayed in the right pane. The columns shown in Table 5-4 will appear on the right pane.

Table 5-4 Multipath I/O columns and description

Column name	Description
State	All paths can be enabled or disabled. Only enabled paths can be active or passive.
DSM ID	Unique ID of the virtual disk.
Nexus ID	Unique ID of the path type.
Initiator HBA Name	Initiator FCP HBA name or iSCSI adapter.
Initiator HBA Address	Initiator FCP HBA WWPN or iSCSI IP address.
Target Adapter/Portal IP	Target FCP WWPN or iSCSI IP address.
Optimized Path	Yes - non-proxy FCP. No - proxy FCP. Not Applicable - iSCSI.
Preferred Path	Yes - non-proxy FCP. No - proxy FCP and iSCSI.
Path Weight	Only applicable when least weight path policy is used.

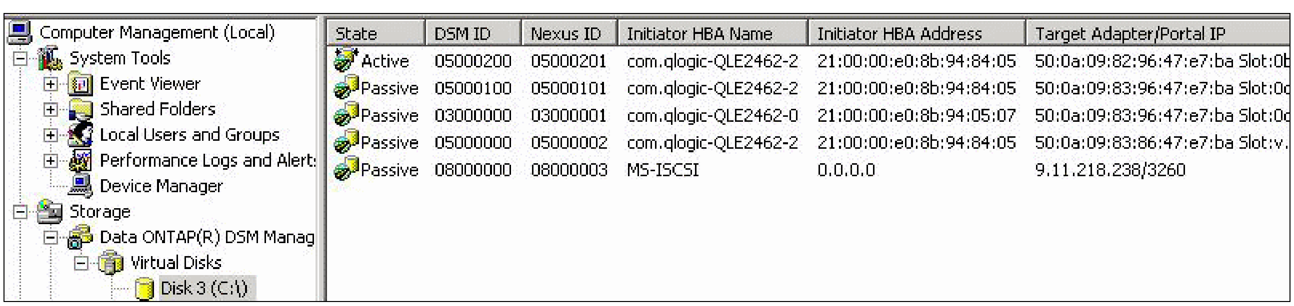


Figure 5-11 Multi-protocol paths to disk 3 on Windows host

Note: You can make a passive path active only when the load balance policy is failover only, round robin with subset, or least queue depth.

With policy auto assigned, you can disable the active path, forcing a passive path to become active.

With policy round robin, you can disable all active paths, forcing all passive paths to become active.

With policy least weighted paths, you can adjust the path weight to specify which path is active.

When you click the virtual disks on the left pane of the MMC, all virtual disks claimed by Data ONTAP DSM are displayed, as shown in Figure 5-12. Table 5-5 provides a brief description of the columns.

Table 5-5 Virtual disks columns and description

Column name	Description
Physical Device ID	Physical device recognized by Windows OS
Storage System	Target N series storage subsystem
Storage System Path	Target N series storage subsystem volume or qtree location
LUN	LUN ID
Serial Number	Unique LUN serial number
Disk Size	Disk size
Load Balance Policy	LBP used by LUN
Cluster Mode	FCP cluster mode used by LUN

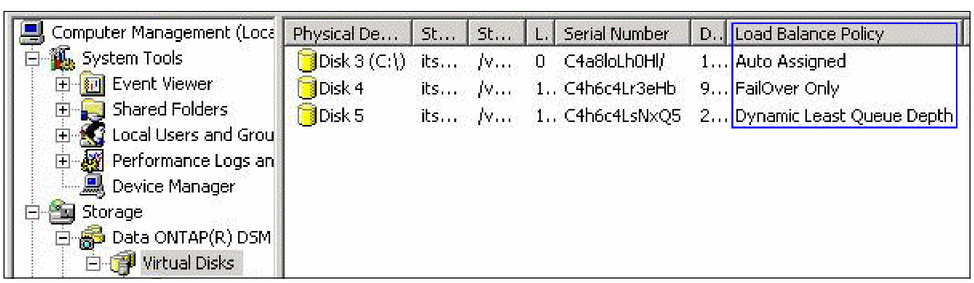


Figure 5-12 Virtual disks and their load balance polices

The same output shown in Figure 5-12 on page 159 can be achieved from the CLI with command **`dsmcli path list -v`**. Refer to Example 5-20 on page 164.

Right-click **Data ONTAP DSM Management** in the left pane to set the default load balance policy for any new virtual disks, as shown in Figure 5-13. Any existing virtual disk is not affected.

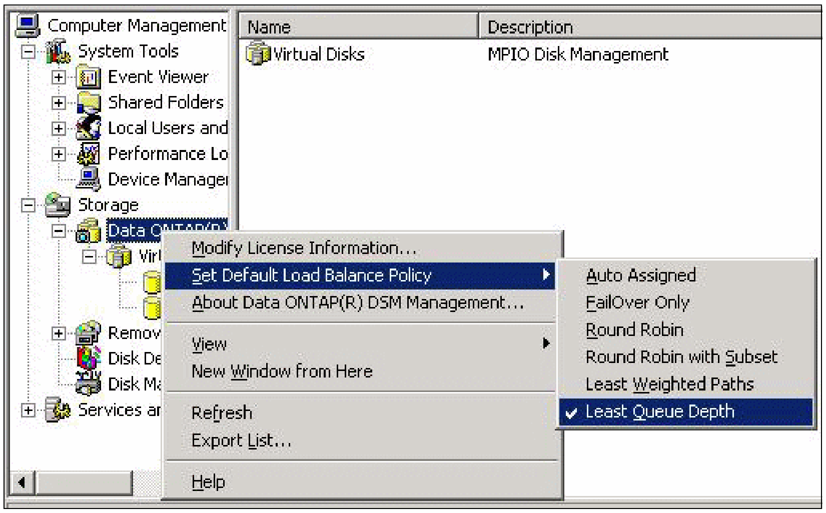


Figure 5-13 Setting default load balance policy for new virtual disks on DSM

To set the LBP of an existing LUN, right-click the chosen disk in the left pane. The existing policy is displayed with a check mark. Choose and highlight the desired policy, as shown in Figure 5-14.

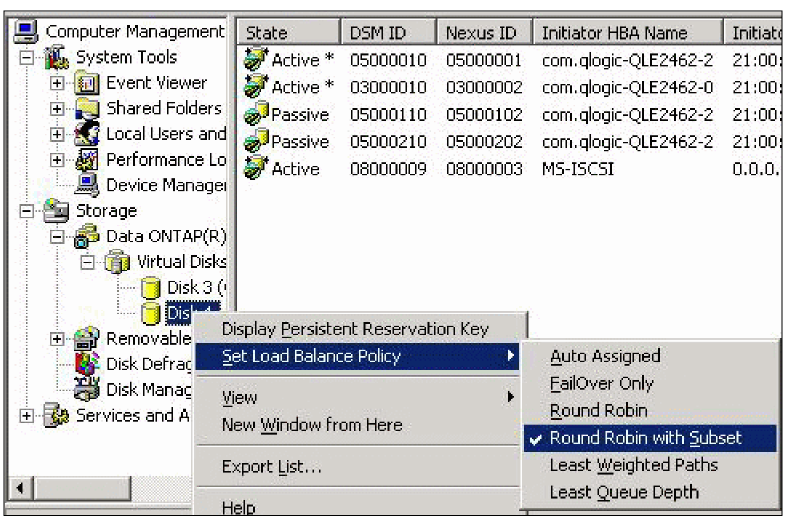


Figure 5-14 Setting load balance policy for an individual virtual disk on DSM

Note: For the auto assigned and failover only policies, you cannot disable the active path. You must first activate another path so that the path you want to disable becomes a passive path.

You should avoid making non-optimized (proxy) paths preferred.

5.5 Managing DSM using the CLI

This section describes how to manage the Data ONTAP DSM using the command line interface.

5.5.1 Discovering new virtual disks (LUNs)

Manually rescan for any new virtual disks when a new LUN is added to your N series storage. To discover any new virtual disks:

1. Select **Administrative Tools** → **Computer Management**.
2. Select **Storage Disk Management** in the left pane.
3. The Initialize and Convert Disk wizard starts. Follow the prompts to initialize the disks, but do not convert them to dynamic disks. Partition and format the disks if desired.

The new virtual disks can be presented from the CLI with the **`dsmcli lun list`** command.

> **Note:** The DSM limits the LUN path on the storage system to 128 characters. The DSM fails to claim disks with longer path names. The LUN path (/vol/vol2/server3_luns/lun4, for example) is specified when you create the LUN on the storage system. If you are using SnapDrive for Windows, the 128-character limit applies to the sum of the LUN path and name fields in the Create Disk wizard.

5.5.2 The main CLI commands

The CLI uses the **`dsmcli`** command. It has four main options:

- **`dsmcli dsm`** displays information about installed device-specific modules and displays or sets the default load balance policy and persistent reservation key.
- **`dsmcli lun`** displays information about LUNs managed by the DSM and sets the load balance policy for individual LUNs.
- **`dsmcli path`** lists the failover status of each LUN, enables and disables paths, and sets a passive path to active.
- **`dsmcli protocol`** controls whether the DSM claims iSCSI LUNs.

5.5.3 Using dcmcli dsm

The commands we discuss in this section list information about the installed DSM. See Example 5-3 and Example 5-4.

Example 5-3 dsmcli dsm syntax

```
dsmcli dsm [ list | version ]
```

Example 5-4 dsmcli dsm examples

```
c:\dsmcli dsm list
Number of DSM registered with MPIO is 1
List of DSMs registered with MPIO:
        "Data ONTAP Multi-Path Device Specific Module"
```

```
c:\dsmcli dsm version
DSM MGMT Dll version 3.2.14889.2055 built on 04/25/2008 21:01:40
ONTAPDSM Driver   version 3.2.14879.1245 built on 04/15/2008 12:48:26
MSFT MPIO Driver version 1.21.3790.2191 built on 05/01/2008 22:07:58
MSISCDSM Driver version 5.2.3790.3640 built on 05/01/2008 22:07:42
```

Setting the default load balance policy

The default load balance policy only applies to newly created virtual disks. Use **dsmcli dsm setdefaultlbp** to change the policy of an existing virtual disk (Example 5-5 and Example 5-6).

Example 5-5 dsmcli dsm getdefaultlbp and setdefaultlbp syntax

```
dsmcli dsm getdefault
dsmcli dsm setdefault -l [ FO | AA | RR | RS | WP | LQD ]
```

Example 5-6 dsmcli dsm getdefaultlbp and setdefaultlbp example

```
C:\>dsmcli dsm getdefaultlbp
dsmcli:The default load balance policy for the dsm is   LEAST QUEUE DEPTH

C:\>dsmcli dsm setdefaultlbp -l AA
dsmcli:The default load balance policy for the dsm is   AUTO ASSIGNED.
```

5.5.4 Using dcmcli protocol

This command displays whether Data ONTAP DSM claims iSCSI LUNs, which is determined during installation if you specified to claim FCP, iSCSI, or both protocols (Example 5-7 and Example 5-8).

Example 5-7 dsmcli protocol syntax

```
dsmcli protocol iscsi -l
```

Example 5-8 dsmcli protocol example

```
C:\>dsmcli protocol iscsi -l
dsmcli:Ontap dsm manages both the FCP and iSCSI luns. iSCSINodeName for this host
is iqn.1991-05.com.microsoft:x3655b
```

Changing whether DSM claims iSCSI LUNs

This command changes whether the Data ONTAP DSM claims iSCSI LUNs (with the **-c** option) or not (with the **-d** option). You must reboot the Windows host for this change to take effect (Example 5-9 and Example 5-10).

Example 5-9 Syntax of dsmcli protocol to claim iSCSI LUN

```
dsmcli protocol iscsi [ -c | -d ]
```

Example 5-10 Example of claiming iSCSI with dsmcli protocol command

```
C:\>dsmcli protocol iscsi -c
dsmcli:Ontapdsm will manage the iscsi luns also. Please reboot the system for the
changes be effective.
```

5.5.5 Using dcmcli lun

These commands display information about the LUNs mapped to the Windows host. A new LUN does not present itself on the host's CLI until a manual rescan is performed from the Windows Disk Manager (Example 5-11 and Example 5-12).

Example 5-11 dsmcli lun syntax

```
dsmcli lun [ list | attributes ]
```

Example 5-12 dsmcli lun example

```
C:\>dsmcli lun list
Disks managed by  ONTAPDSM
C4a8loLhOHl/    Data ONTAP Multi-Path Device Specific Module
C4h6c4Lr3eHb    Data ONTAP Multi-Path Device Specific Module

C:\>dsmcli lun attributes
Disks managed by  ONTAPDSM
SerialNumber     Storage System  Storage System Path            MountPath
************     **************  ********************           *********
C4a8loLhOHl/    itsotuc4        /vol/x3655b/LUN_SAN_BOOT_3655B     C:\
C4h6c4Lr3eHb    itsotuc3        /vol/x3655b_iSCSI/mpio_win     E:\
```

Changing the load balance policy

The command shown in Example 5-13 displays the load balance policy for a single LUN (Example 5-14). Serial number can be displayed using the commands described previously.

Example 5-13 dsmcli lun set lbp syntax

```
dsmcli lun setlbp -l < FO | AA | RR | RS | WP | LQD> [-s serialNumber | -d
DriveLetter]
```

Example 5-14 dsmcli lun set lbp example

```
C:\>dsmcli lun setlbp -l RR -s C4a8loLhOHl/
dsmcli:The load balance policy for the lun is set to ROUND ROBIN

C:\>dsmcli lun setlbp -l AA -d C:\
dsmcli:The load balance policy for the lun is set to AUTO ASSIGNED
```

Listing available load balance policies

The command shown in Example 5-15 displays the load balance policies that can be used with the specified LUN (Example 5-16). Different policies might be available, depending on your configuration.

Example 5-15 dsmcli lun getlbp syntax

```
dsmcli lun getlbp  [-s serialNumber | -d DriveLetter]
```

Example 5-16 dsmcli lun getlbp example

```
C:\>dsmcli lun getlbp -s C4a8loLhOHl/
dsmcli: The supported Load Balance policies for this LUN are
 FAILOVER
 WEIGHTED PATHS
```

```
AUTO ASSIGNED
ROUND ROBIN
ROUND ROBIN WITH SUBSET
LEAST QUEUE DEPTH
```

Viewing the persistent reservation key

The Data ONTAP DSM for Windows MPIO uses a persistent reservation key to track which host in a Microsoft Windows 2003 cluster (MSCS) or Windows 2008 Failover Cluster is currently allowed to write to a virtual disk (LUN). You can view the key for each virtual disk (Example 5-17 and Example 5-18).

> **Note:** In Windows Server 2003, all virtual disks have the same value on the host. In Windows Server 2008, each virtual disk has a unique value.

Example 5-17 dsmcli lun getprkey syntax

```
dsmcli lun getprkey [-s serialNumber | -d DriveLetter]
```

Example 5-18 dsmcli lun getprkey example

```
C:\>dsmcli lun getprkey -d C:\
The Persistent Reservation Key for lun C4a8loLhOHl/ is: 53 56 50 50 10 21 2 49
```

5.5.6 Using dcmcli path

The command shown in Example 5-19 displays all of the paths for one or all LUNs mapped to the Windows host (Example 5-20). To display the paths for an individual LUN, specify the serial number or drive letter.

Example 5-19 dsmcli path list syntax

```
dsmcli path list [ -v ] [-s SerialNumber | -d DriveLetter]
```

Example 5-20 dsmcli path list example

```
C:\>dsmcli path list -v -d C:\
Serial Number: C4a8loLhOHl/
MPIO Paths: 5
Load Balance Policy:  AUTO ASSIGNED

Dsm Id:         0x5000200
SCSI Address:
 Scsiport : 5
 HostPathId : 0
 Targetid : 2
 lun : 0
Path State:     ACTIVE
Protocol : FCP
Host Adapter Name : com.qlogic-QLE2462-2
Host Adapter WWN Port : 21:00:00:e0:8b:94:84:05
Proxy Path : NO
CFMode: SINGLE IMAGE
Target Adapter:50:0a:09:82:96:47:e7:ba
Target Node Name:50:0a:09:80:86:47:e7:ba
```

```
Dsm Id:         0x5000100
SCSI Address:
 Scsiport : 5
 HostPathId : 0
 Targetid : 1
 lun : 0
Path State:     PASSIVE
Protocol : FCP
Host Adapter Name : com.qlogic-QLE2462-2
Host Adapter WWN Port : 21:00:00:e0:8b:94:84:05
Proxy Path : NO
CFMode: SINGLE IMAGE
Target Adapter:50:0a:09:83:96:47:e7:ba
Target Node Name:50:0a:09:80:86:47:e7:ba


Dsm Id:         0x3000000
SCSI Address:
 Scsiport : 3
 HostPathId : 0
 Targetid : 0
 lun : 0
Path State:     PASSIVE
Protocol : FCP
Host Adapter Name : com.qlogic-QLE2462-0
Host Adapter WWN Port : 21:00:00:e0:8b:94:05:07
Proxy Path : NO
CFMode: SINGLE IMAGE
Target Adapter:50:0a:09:83:96:47:e7:ba
Target Node Name:50:0a:09:80:86:47:e7:ba


Dsm Id:         0x5000000
SCSI Address:
 Scsiport : 5
 HostPathId : 0
 Targetid : 0
 lun : 0
Path State:     PASSIVE
Protocol : FCP
Host Adapter Name : com.qlogic-QLE2462-2
Host Adapter WWN Port : 21:00:00:e0:8b:94:84:05
Proxy Path : YES
CFMode: SINGLE IMAGE
Target Adapter:50:0a:09:83:86:47:e7:ba
Target Node Name:50:0a:09:80:86:47:e7:ba


Dsm Id:         0x8000000
SCSI Address:
 Scsiport : 8
 HostPathId : 0
 Targetid : 0
 lun : 0
```

```
Path State:     PASSIVE
Protocol : iSCSI
Initiator Name : MS-iSCSI
Initiator Portal : 0.0.0.0/1033
Target Portal IP : 9.11.218.238
```

Changing path status

The commands shown in Example 5-21 change the status of the specified path. The actions you can take depend on the load balance policy of the virtual disk and the status of the other paths. See Example 5-22.

Specify the path that you want to change using the DSM ID. You can get the DSM ID using the **dsmcli path list** command. The 0x prefix for the DSM ID is optional.

> **Note:** You can make a passive path active for any load balance policy except auto assigned. With the auto assigned policy you can disable the active path, forcing a passive path to become active. You cannot disable a path if no other path is available to take over. There must always be an active path.
>
> The passive option applies only to the round robin with subset and least queue depth load balance policies.
>
> Although you can make a non-optimized (proxy) path active, you should avoid doing so if any optimized paths are available.

Example 5-21 dsmcli path status syntax

```
dsmcli path < enable | disable | active | passive | weight | list > < -p DsmId >
<-w weight>
```

Example 5-22 dsmcli path status example

```
C:\>dsmcli path active -p 0x5000100
dsmcli:Path 0x05000100 is set to ACTIVE state

C:\>dsmcli path disable -p 0x5000100
dsmcli:Path 0x05000100 is DISABLED

C:\>dsmcli path enable -p 0x5000100
dsmcli:Path 0x05000100 is ENABLED
```

Changing path weight

The command shown in Example 5-23 sets the weight assigned to each path for virtual disks with the least weighted path load balance policy. See Example 5-24.

The available path with the lowest weight is used to access the virtual disk. If multiple paths with the same weight value are available, the DSM selects the path shared with the fewest other LUNs. The weight value can be set from 0 (highest priority) to 2146999999 (lowest priority). The default is 255.

Example 5-23 dsmcli path weight syntax

```
dsmcli dsmcli path weight -p Dsm_Id -w weight
```

Example 5-24 dsmcli path weight example

```
C:\>dsmcli path weight -p 0x8000000 -w 200
dsmcli:Path 0x08000000 weight is set to  200
```

5.6 Multiple path I/O support for Red Hat Linux

In this section we discuss the software component for the installation of FC Linux Host Utilities kit (5.0 and 5.1). Physical cabling and HBA configuration and installation, and MPIO support for the boot device (local or from SAN) are beyond the scope of this book. This section covers the items pertaining to the host utility kits and enabling MPIO support on the host.

Note: For details about the Linux install on a multipath boot device refer to:

http://publib.boulder.ibm.com/infocenter/systems/index.jsp?topic=/liaai/multipath/liaaiiscsimultirhel.htm

Ensure that you have a supported combination of the following components before beginning the installation:

- Linux OS version
- HBA model (firmware and driver version)
- FC switch model (firmware version)
- N series storage (Data ONTAP version)

Visit the following website to view the IBM System Storage N series interoperability matrix for FC and iSCSI SAN, Antivirus and UPS link:

http://www-03.ibm.com/systems/storage/network/interophome.html

The FC Linux Host Utilities kit supports both QLogic and Emulex brand HBAs with bundled and external HBA drivers.

Note: Although both QLogic and Emulex HBAs are supported, we do not recommend installing different HBA brands on the same host.

5.6.1 HBA installation and configuration

This section explains how to install the Host Utilities kit and HBA drivers and complete the initial configuration of the storage system. Linux native dm-multipath support is required to create a highly available connection between the host and N series storage. Device mapper and multipath tools packages are also needed.

> **Note:** Log in with root privileges before you configure the initiator. We used Red Hat Enterprise Linux Server release 5.2 for the examples shown here.

Ensure that the correct version of the Host Utilities kit (Linux Host Utilities 5.1 is now supported) is downloaded for your version of OS from this IBM website:

http://www-304.ibm.com/systems/support/supportsite.wss/supportresources?brandind=5000029&familyind=5364556&taskind=1

Complete the following steps to install the Host Utilities kit:

1. Remove any previous version of the Host Utilities kit. Run the **`./uninstall`** command in the directory where the previous version was installed (default directory /opt/ontap/santools).
2. Go to the directory in which you have placed the downloaded Host Utilities kit file.
3. To uncompress the file, enter the following command:

   ```
   gunzip ibm_linux_host_utilities_5_1.tar.gz
   ```

4. To extract the files, enter the following command:

   ```
   tar -xvf ibm_linux_host_utilities_5_1.tar
   ```

5. Change to the ibm_linux_host_utilities_5_1 directory. By default, this is a sub-directory of the working directory from which you extracted in the previous step.
6. Enter the following command:

   ```
   ./install
   ```

   The following directories and files are installed.

   ```
   opt/ontap/
   opt/ontap/santools/
   opt/ontap/santools/sanlun
   opt/ontap/santools/brocade_info
   opt/ontap/santools/cisco_info
   opt/ontap/santools/filer_info
   opt/ontap/santools/linux_info
   opt/ontap/santools/mcdata_info
   opt/ontap/santools/qlogic_info
   opt/ontap/santools/uninstall
   opt/ontap/santools/FTP.pm
   opt/ontap/santools/SHsupport.pm
   opt/ontap/santools/Telnet.pm
   opt/ontap/santools/san_version
   opt/ontap/santools/mpath_prio_ontap
   opt/ontap/santools/mpath_prio_ontap.static
   opt/ontap/man/
   opt/ontap/man/man1/
   opt/ontap/man/man1/brocade_info.1
   opt/ontap/man/man1/cisco_info.1
   opt/ontap/man/man1/filer_info.1
   opt/ontap/man/man1/linux_info.1
   opt/ontap/man/man1/mcdata_info.1
   opt/ontap/man/man1/qlogic_info.1
   opt/ontap/man/man1/sanlun.1
   ```

5.6.2 Uninstall previous HBA drivers

You must remove any previous external HBA driver modules.

QLogic HBA drivers

Follow these steps to remove the QLogic HBA driver modules:

1. Log into the Linux host as root.
2. Stop the qlremote process if it is running by entering the following command:

   ```
   pkill qlremote
   ```

3. Change to the directory where the driver installer files are located.
4. Uninstall the previous QLogic driver on the host. Ensure that the driver is not being used (usage count of qla2xxx is 1):
   - From the SANsurfer package that is installed on the host, run the following command:

     ```
     ./qinstall --uninstall
     ```

   - For a newly installed host with the bundled drivers loaded, run the following command:

     ```
     modprobe -r qla2300/qla2400/qla2500
     ```

Emulex HBA drivers

Follow these steps to remove the Emulex HBA driver modules:

1. Log into the Linux host as root.
2. Go to the directory where the driver installer files are located.
3. Uninstall the previous Emulex driver on the host. Ensure that the driver is not being used (usage count of lpfc is 0):
   - For a freshly installed host with the bundled drivers loaded, run the following command:

     ```
     modprobe -r lpfc
     ```

   - Run the following command to remove the drivers:

     ```
     ./lpfc-install --uninstall
     ```

5.6.3 Verify native Linux multipath packages

We recommend using dm-multipath to create a highly available connection between your host and the N series storage subsystem. Complete these steps to verify that the correct multipathing packages have been installed:

1. If you plan to use dm-multipath support, enter the following:

   ```
   rpm -q device-mapper
   rpm -q device-mapper-multipath
   ```

   The **`rpm -q`** command returns a list of the multipathing packages installed.

2. Install the required packages if desired.

5.6.4 Download and install external HBA drivers

The HBA requires a specific driver version customized for use with the IBM System Storage N series storage subsystem. Confirm your HBA interoperability from the following IBM website (find the IBM System Storage N series interoperability matrix for FC and iSCSI SAN, Antivirus and UPS link):

http://www-03.ibm.com/systems/storage/network/interophome.html

The appropriate management application for the HBA is also required:

- QLogic
  a. Go to:
     http://support.qlogic.com/support/drivers_software.asp
  b. Under OEM Models, select **NetApp®**
  c. Find and download the driver version listed in the support matrix.
- Emulex
  a. Go to:
     http://www.emulex.com/ts/index.html
  b. Under Storage and System Supplier Qualified and Supported HBAs, select **NetApp**.
  c. Find the driver version listed in the support matrix and download it. Also download the Emulex applications package from the same location.

Follow the installation instructions. A reboot might be required.

5.6.5 Setting HBA and driver parameters

HBA and driver parameters must be set for multipathing and storage system failover to work optimally. To modify the QLogic HBA parameters, complete these steps:

1. Run the `scli` command.
2. From the main menu, select the **Configure HBA Settings** option. The navigation of the menu may vary depending on the version of SANsurfer FC HBA CLI installed. The following steps lead to the desired menu on SANsurfer Version 1.7.1 Build 23:

   Main Menu → 3: HBA Parameters → 1: Port 1: WWPN: <FC HBA → → 2: Configure HBA Parameters
3. Select the first of the two HBA port listed.
4. Configure the HBA parameter ConnectionOptions according to the connection to the N series storage subsystem:
   – Point to Point Only: SSI, Standby, or Partner FC cfmodes
   – Loop Only: Dual Fabric or Mixed FC cfmodes
5. Configure the Data Rate option of the QLogic port speed to the maximum supported, depending on that of the switch or target port to which it is connected.
6. Modify the Port Down Retry Count option to 30.
7. Modify the Link Down Timeout option to 20.
8. Select Commit Changes for the select HBA.
9. Repeat these steps for every other HBA installed on your host.
10. Return to the main menu and select **Quit**.

Edit /etc/modprobe.conf to modify the QLogic driver parameters by performing the following steps:

1. Unload the external QLogic HBA driver by entering the appropriate command for your HBA model:

```
./qlinstall -ul qla2500
./qlinstall -ul qla2400
./qlinstall -ul qla2300
```

2. Append the options qla2xxx in /etc/modprobe.conf on a single line. Only one option, qla2xxx, is allowed:

```
ql2xsuspendcount=6 MaxRetriesPerPath=10 MaxRetriesPerIo=10 ql2xfailover=0
```

3. Save the file and exit /etc/modprobe.conf.
4. Reload the external QLogic HBA driver using the appropriate command for your HBA:

```
./qlinstall -l qla2500
./qlinstall -l qla2400
./qlinstall -l qla2300
```

5. Enter the appropriate command to rebuild the initial ramdisk (initrd) image with the new parameter:

```
./qlinstall -br -in qla2500
./qlinstall -br -in qla2400
./qlinstall -br -in qla2300
```

6. Reboot your host with the updated image.

Note: No changes are required for Emulex HBAs.

5.6.6 Bundled HBA drivers and applications

HBA and driver parameters must be set for multipathing and storage system failover to work optimally.

Note: Individual firmware modules for qla2300, qla2400, and qla2500 have been integrated into the base qla2xxx module in the Red Hat Linux 5 series.

To modify the QLogic HBA parameters, perform these steps:

1. Run the **scli** command.
2. Set the following HBA parameters:
   - Port Down Retry Count to 30
   - Link Down Timeout to 20
3. Configure the HBA parameter ConnectionOptions according to the connection to the N series storage subsystem:
   - Point to Point Only: SSI, Standby, or Partner FC cfmodes
   - Loop Only: Dual Fabric or Mixed FC cfmodes
4. Configure the Data Rate option of the QLogic port speed to the maximum supported, depending on which switch or target port it is connected to.
5. Select **Commit Changes** for the select HBA.
6. Repeat these steps for every other HBA installed.

The default values work fine in the /etc/modprobe.conf of the QLogic bundled drivers.

Note: No changes are required for bundled Emulex HBAs.

5.6.7 WWPN of your host HBAs

Each HBA has a unique WWPN used to identify it, which is added to an igroup to enable the initiator to access a particular LUN on the N series storage subsystem.

WWPN of QLogic HBAs

Use the SANsurfer FC HBA CLI to obtain the WWPN of QLogic HBAs by entering the following command (the output is shown in Example 5-25):

```
scli -l
```

Example 5-25 Output of command: scli -l; WWPN in bold

```
--------------------------------------------------------------------------------
Host Name                         : x366
HBA Instance                      : 0
HBA Model                         : QLA2340
HBA Description                   : QLA2340/QLA2340L (PCI to FC single channel)
HBA ID                            : 0-QLA2340
HBA Alias                         :
HBA Port                          : 1
Port Alias                        :
Node Name                         : 20-00-00-E0-8B-0A-18-27
Port Name                         : 21-00-00-E0-8B-0A-18-27
Port ID                           : 6F-12-00
Serial Number                     : G65368
Driver Version                    : 8.02.00-k5-rhel5.2-04
BIOS Version                      : 1.25
Firmware Version                  : 3.03.26 IPX
Actual Connection Mode            : Point to Point
Actual Data Rate                  : 2 Gbps
PortType (Topology)               : NPort
Target Count                      : 1
PCI Bus Number                    : 2
PCI Device Number                 : 1
HBA Status                        : Online
--------------------------------------------------------------------------------
Host Name                         : x366
HBA Instance                      : 1
HBA Model                         : QLA2340
HBA Description                   : QLA2340/QLA2340L (PCI to FC single channel)
HBA ID                            : 1-QLA2340
HBA Alias                         :
HBA Port                          : 1
Port Alias                        :
Node Name                         : 20-00-00-E0-8B-12-0E-54
Port Name                         : 21-00-00-E0-8B-12-0E-54
Port ID                           : 6F-13-00
Serial Number                     : M01166
Driver Version                    : 8.02.00-k5-rhel5.2-04
BIOS Version                      : 1.52
```

```
Firmware Version                 : 3.03.26 IPX
Actual Connection Mode           : Point to Point
Actual Data Rate                 : 2 Gbps
PortType (Topology)              : NPort
Target Count                     : 1
PCI Bus Number                   : 4
PCI Device Number                : 1
HBA Status                       : Online
--------------------------------------------------------------------------------
```

WWPN of Emulex HBAs

Use the Emulex HBAnywhere for Linux CLI to obtain the WWPN of Emulex HBAs:

```
/usr/sbin/hbanywhere/hbacmd ListHBAs
```

5.6.8 Configuring the N Series storage system for FC

Refer to the *Data ONTAP 7.3 Block Access Management Guide for iSCSI and FCP*, GC26-7973 to configure the N series storage subsystem for FC. To query your N series FC cfmode, enter the following command in the N series CLI:

```
fcp show cfmode
```

Example 5-26 shows the results of the command.

Example 5-26 Output of FCP show cfmode on N series storage subsystem

```
fcp show cfmode: single_image
```

Note: We assume that your N series FC cfmode is single_image. Consult the following website to verify the supported cfmode on your N series and host configuration:

http://www-03.ibm.com/systems/storage/network/interophome.html

Go to the IBM System Storage N series interoperability matrix for FC and iSCSI SAN, Antivirus and UPS link.

5.6.9 Loading the external HBA driver

Load the external HBA driver once you have a LUN mapped to your host. All mapped LUNs will be discovered from the N series storage subsystem when the external HBA driver is started. Multipath devices will be created if **dm-multipath** is configured.

Note: Be sure that a LUN is mapped to the host as LUN 0 before loading the external HBA driver. Also make sure the multipath is configured before starting FC.

To load the external QLogic HBA driver enter the following at the CLI of you host:

```
modprobe -v qla2300
modprobe -v qla2400
modprobe -v qla2500
```

To load an external Emulex HBA driver enter:

```
modprobe -v lpfc
```

5.6.10 Loading the bundled HBA driver

Load the bundled HBA driver once you have a LUN mapped to your host. All mapped LUNs will be discovered from the N series storage subsystem when the external HBA driver is started. Multipath devices will be created if **`dm-multipath`** is configured.

> **Note:** Be sure that a LUN is mapped to the host as LUN 0 before loading the external HBA driver. Also make sure that multipath is configured before starting FC.

To loading an external QLogic HBA driver enter the following at the CLI of you host (for Red Hat Enterprise Linux 5):

```
modprobe -v qla2xxx
```

To load an external Emulex HBA driver enter:

```
modprobe -v lpfc
```

5.6.11 Configuring dm-multipath

Perform these steps to configure dm-multipath:

1. Rename the existing /etc/multipath.conf to /etc/multipath.conf.old if it exists. Create a new /etc/multipath.conf file and insert the code provided in Example 5-27 into the file.
2. Save the changes.
3. Edit the /etc/multipath.conf file to exclude (blacklist) all devices that are not displayed from the **`sanlun lun show`** command.
   a. Blacklist the non N series SCSI devices installed using the WWID method. Run the **`scsi_id`** command to obtain the WWID:
   ```
   scsi_id -gus /block/sda
   ```
   b. Blacklist other devices using the devnode method:
   ```
   devnode “^hd[a-z]”
   devnode "^(ram|raw|loop|fd|md|dm-|sr|scd|st)[0-9]*"
   devnode "^cciss!c[0-9]d[0-9]*"
   ```
4. Save the changes and exit.

Example 5-27 Changes required to /etc/multipath.conf on Red Hat Enterprise Linux 5.2

```
defaults
{
   user_friendly_names yes
   max_fds 4096
}
blacklist
{
   wwid <DevId>
   devnode "^(ram|raw|loop|fd|md|dm-|sr|scd|st)[0-9]*"
   devnode "^hd[a-z]"
   devnode "^cciss!c[0-9]d[0-9]*[p[0-9]*]"
}
devices
{
   device
```

```
    {
        vendor "NETAPP"
        product "LUN"
        getuid_callout "/sbin/scsi_id -g -u -s /block/%n"
        prio_callout "/sbin/mpath_prio_ontap /dev/%n"
        features "1 queue_if_no_path"
        hardware_handler "0"
        path_grouping_policy group_by_prio
        failback immediate
        rr_weight uniform
        rr_min_io 128
        path_checker directio
    }
}
```

Note: max_fds is available with Red Hat Enterprise Linux 5 Update 2 and later. It is not supported on Red Hat Enterprise Linux 5.0 and Red Hat Enterprise Linux 5 Update 1.

5.6.12 Starting the multipath service

Start the multipath service on your host when you have successfully loaded the external HBA drivers. Perform these steps:

1. To start the multipath daemon, enter the following command on your host CLI:

   ```
   /etc/init.d/multipathd start
   Starting multipathd daemon:                              [  OK  ]
   ```

2. Run the following command to configure the dm-multipath devices:

   ```
   multipath
   ```

You can configure the multipath service to start automatically upon reboot with the following commands on your host:

```
chkconfig --add multipathd
chkconfig multipathd on
```

5.6.13 Accessing LUNs with dm-multipath support

Perform the steps in this section to access LUNs with dm-multipath support.

Verify HBA setting and dm-multipath configuration on QLogic HBA

Once dm-multipath is started, verify that the HBA settings are correct, the desired multipathing is correctly configured, and the multipath maps were created.

- HBA settings

  a. Verify the ql2xsuspendcount, MaxRetriesPerIo, and MaxRetriesPerPath parameters by checking the values for these parameters in the corresponding qla2xxx directory in the sysfs, /sys/module/qla2xxx/, or /sys/module/qla2xxx/parameters/.

     ```
     cat /sys/module/qla2xxx/ql2xsuspendcount
     ```

  b. PortDownRetryCount and LinkDownTimeout, along with Connection Options and Data Rate (which must be manually set), can be verified as follows:

     ```
     /opt/QLogic_Corporation/SANsurferCLI/scli -c all
     ```

Example 5-28   Output of scli -c all

```
--------------------------------------------------------------------------------
HBA Instance 0: QLA2340 Port 1 WWPN 21-00-00-E0-8B-0A-18-27 PortID 6F-12-00
--------------------------------------------------------------------------------
Connection Options              : 1 - Point-to-Point Only.
Data Rate                       : 2 Gbps
Frame Size                      : 2048
Hard Loop ID                    : 125
Loop Reset Delay (seconds)      : 5
Enable Host HBA BIOS            : Enabled
Enable Hard Loop ID             : Enabled
Enable FC Tape Support          : Enabled
Operation Mode                  : 0 - Interrupt for every I/O completion
Interrupt Delay Timer (100ms)   : 0
Execution Throttle              : 256
Enable Extended Error Logging   : Disabled
Login Retry Count               : 30
Enable LIP Reset                : Disabled
Port Down Retry Count           : 30
Enable LIP Full Login           : Enabled
Link Down Timeout (seconds)     : 20
Enable Target Reset             : Enabled
LUNs Per Target                 : 0
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
HBA Instance 1: QLA2340 Port 1 WWPN 21-00-00-E0-8B-12-0E-54 PortID 6F-13-00
--------------------------------------------------------------------------------
Connection Options              : 1 - Point-to-Point Only.
Data Rate                       : 2 Gbps
Frame Size                      : 2048
Hard Loop ID                    : 125
Loop Reset Delay (seconds)      : 5
Enable Host HBA BIOS            : Disabled
Enable Hard Loop ID             : Enabled
Enable FC Tape Support          : Enabled
Operation Mode                  : 0 - Interrupt for every I/O completion
Interrupt Delay Timer (100ms)   : 0
Execution Throttle              : 256
Enable Extended Error Logging   : Disabled
Login Retry Count               : 30
Enable LIP Reset                : Disabled
Port Down Retry Count           : 30
Enable LIP Full Login           : Enabled
Link Down Timeout (seconds)     : 20
Enable Target Reset             : Enabled
LUNs Per Target                 : 0
--------------------------------------------------------------------------------
```

Note: Default settings work fine for Emulex HBAs.

- Multipath configuration

  Use the **`multpath`** command to view the dm-multipath configuration. Use the **`-v`** option for a detailed display. For a list of available options, use the **`-h`** option.

```
multipath -v3 -d -ll
```

This command will display all configuration details, but prevents updating the multipath maps with the **-d** option used.

- Multipath status

  Verify that the multipath daemon is running by issuing the following command:

  ```
  /etc/init.d/multipathd status
  ```

- Multipath devices

  To view a list of multipath devices, run the following command:

  ```
  multipath -ll
  ```

  Example 5-29 shows the output of this command.

Example 5-29 Output of multipath -ll

```
mpath2 (360a98000433468366334 4d33634b6842) dm-3 NETAPP,LUN
[size=1.0G][features=1 queue_if_no_path][hwhandler=0]
\_ round-robin 0 [prio=8][active]
 \_ 0:0:0:2 sdc 8:32  [active][ready]
 \_ 1:0:0:2 sdf 8:80  [active][ready]
mpath1 (360a98000433468366334 4c6945315875) dm-2 NETAPP,LUN
[size=1.0G][features=1 queue_if_no_path][hwhandler=0]
\_ round-robin 0 [prio=8][active]
 \_ 0:0:0:1 sdb 8:16  [active][ready]
 \_ 1:0:0:1 sde 8:64  [active][ready]
```

Identify and view a list of LUNs

The dm-multipath code creates persistent devices for each LUN within the /etc/mapper/ directory on your host. Use these entries for multipath LUNs, which can be used to create a file system directly on a multipath device in /etc/mapper/. The entries in /etc/mapper/ can also be used in /etc/fstab file.

Use the **sanlun lun show all** command to review all LUNs mapped to your host (Example 5-30).

Example 5-30 Output of sanlun lun show all

```
filer:            lun-pathname          device filename  adapter  protocol         lun
size         lun state
itsotuc3:  /vol/rh_boot_x366/data_vg_SdLun  /dev/sdb        host0    FCP         1.0g
(1082130432)     GOOD
itsotuc3:  /vol/rh_data_x366/fcp_mpio      /dev/sdc        host0    FCP           1g
(1073741824)     GOOD
itsotuc3:  /vol/rh_boot_x366/data_vg_SdLun  /dev/sde        host1    FCP         1.0g
(1082130432)     GOOD
itsotuc3:  /vol/rh_data_x366/fcp_mpio      /dev/sdf        host1    FCP           1g
(1073741824)     GOOD
```

Use the **sanlun lun show -p** command to review multipath information (Example 5-31).

Example 5-31 Output of sanlun lun show -p

```
itsotuc3:/vol/rh_boot_x366/data_vg_SdLun (LUN 1)               Lun state: GOOD
Lun Size:    1.0g (1082130432)   Filer_CF_State: Cluster Enabled
Protocol: FCP           Filer Partner: itsotuc4
DM-MP DevName: mpath1   (360a98000433468366334 4c6945315875)     dm-2
```

```
Multipath-provider: NATIVE
------ ---------- ------- ------------ --------------- ---------------
sanlun      Filer                                Primary        Partner
  path       Path   /dev/         Host             Filer          Filer
 state       type    node          HBA              port           port
------ ---------- ------- ------------ --------------- ---------------
    UP  primary       sdb        host0              0b             --
    UP  primary       sde        host1              0b             --

itsotuc3:/vol/rh_data_x366/fcp_mpio (LUN 2)            Lun state: GOOD
Lun Size:      1g (1073741824)   Filer_CF_State: Cluster Enabled
Protocol: FCP           Filer Partner: itsotuc4
DM-MP DevName: mpath2   (360a98000433468366334 4d33634b6842)     dm-3
Multipath-provider: NATIVE
------ ---------- ------- ------------ --------------- ---------------
sanlun      Filer                                Primary        Partner
  path       Path   /dev/         Host             Filer          Filer
 state       type    node          HBA              port           port
------ ---------- ------- ------------ --------------- ---------------
    UP  primary       sdc        host0              0b             --
    UP  primary       sdf        host1              0b             --
```

Tuning dm-multipath configuration

Two Linux parameters are available to advanced administrators to extract additional performance from the LUNs:

- queue_depth is one of the critical settings to obtain best performance from the SCSI device queue. The setting is found in sysfs and can be altered by echoing a new value for the queue_depth.

  ```
  echo "64" > /sys/block/sdb/device/queue_depth
  ```

- The rr_min_io setting for dm-multpath specifies the number of I/O operations sent through a path before switching to the next path. Lowering this value from the default value of 1000 dramatically improves performance, especially for workload with large I/O of 64 KB or more in size.

  Other performance tests have shown that the value for rr_min_io should be less than or equal to the number of threads in the system that issue a single I/O and block, waiting for it to complete.

  Note: The rr_min_io value is recommended for optimal performance tuning and is dependent on the workload, host memory, and CPU speed.

  Set the rr_min_io value as follows in the /etc/multipath.conf.

  ```
  device
  {
     rr_min_io 128
  }
  ```

5.6.14 Stopping dm-multipath support

If you decide to stop the multipath support on your host, stop the affected services in the following order:

1. Unmount all file systems on the FC LUNs.

2. Flush the multipath maps by entering the command:

   `multipath -F`

3. Stop the multipath daemon by entering the command:

   `/etc/init.d/multipathd stop`

5.7 Multiple path I/O support for Native AIX O/S

This section provides a brief overview of the FCP AIX Host Utilities kit installation for Native OS, which also includes creating, configuring, and managing LUNs on non-VIO hosts. The Host Utilities kit provides the necessary tools for your AIX host to connect to the N series storage subsystem. IBM Path Control Module (PCM) will be as used as the MPIO solution.

The AIX Host Utilities 5.0 also supports Power VM Live Partition Mobility on IBM system p.

> **Note for AIX systems with VIO:** The FCP AIX Host Utilities kit for Native OS supports hosts running an AIX Virtual I/O (VIO) server. If you intend to install the Host Utilities kit with your host running VIO, ensure that your VIO environment is set up first.
>
> Different commands are used to set up your environment if you are running VIO on your hosts. For example, you are required to run as padmin, not root, on a VIO host. To discover LUNs on VIO servers, the **`cfgdev`** command is used instead of **`cfgmgr`**.
>
> - Installing Host Utilities kit in a VIO environment: You are logged in as padmin in a VIO environment. Use the **`oem_setup_env`** command to change to root. (The Host Utilities kit requires you to be logged in as root.)
> - Installing SAN Toolkit package on VIO server: You must install this package to each VIO client in addition to the VIO server. Although you are unable to run the sanlun utility on the client, you can run the diagnostic scripts from the SAN Toolkit.
> - Troubleshooting on VIO servers:
>   - To collect information about the host, run the ontap_collect (instead of aix_info) file.
>   - If you have virtual SCSI LUNs set up, the sanlun utility does not recognize them because LUNs are identified as VID ‘IBM’ and PID ‘VDASD’.
>
> Check the following IBM websites for current support and interoperability information:
>
> http://www.ibm.com/storage/support/nas
> http://www-03.ibm.com/systems/storage/product/interop.html

With the native AIX MPIO, you can configure multiple network paths between the host and the N series storage system. If one path fails, FCP traffic continues on the alternate paths.

> **Note:** Our AIX host is running the following OS level:
>
> ```
> # oslevel -g
> Fileset          Actual Level      Maintenance Level
> bos.rte          6.1.1.0           6.1.0.0
> ```

Multiple HBAs will be installed to provide multipathing to your storage target. Do not install different HBA brands on a single host.

5.7.1 Uninstalling Host Utilities kit

The older Host Utilities kit must be removed because it uses a different stack from the current 4.2 version. You should not have both the earlier and the current version installed on your host. Complete these prerequisites before you uninstall the previous version of the software:

1. Unmount the file systems that contain the volume groups mapped to the N series LUNs.
2. Quiesce and export the volume groups.
3. Stop multipathing software.
4. Remove the existing Data ONTAP LUN handles, which disconnects ONTAP LUNs from the host before you uninstall the existing version of the Host Utilities kit.

To remove the previous version of the Host Utilities kit with AIX SMIT, perform these steps:

1. Log into your host as root.
2. Invoke SMIT by running the following command:

   `# smit`

3. Select the Software Installation and Maintenance menu option:

Example 5-32 SMIT menu

```
root> # smit
System Management
Move cursor to desired item and press Enter.

Software Installation and Maintenance
Software License Management
Devices
System Storage Management (Physical & Logical Storage)
Security & Users
Communications Applications and Services
Print Spooling
Problem Determination
Performance & Resource Scheduling
System Environments
Processes & Subsystems
Applications
Cluster System Management
Using SMIT (information only)

F1=Help F2=Refresh F3=Cancel
Esc+8=Image
Esc+9=Shell Esc+0=Exit Enter=Do
```

4. Select the menu option Software Maintenance and Utilities on the current screen to proceed:

Example 5-33 Software Maintenance and Utilities option

```
Software Installation and Maintenance

Move cursor to desired item and press Enter.

Install and Update Software
List Software and Related Information
Software Maintenance and Utilities
Network Installation Management
System Backup Manager
```

```
F1=Help F2=Refresh F3=Cancel
Esc+8=Image
Esc+9=Shell Esc+0=Exit Enter=Do
```

5. Select the menu option Remove Installed Software on the current screen to proceed:

Example 5-34 Remove Installed Software option

```
Software Maintenance and Utilities

Move cursor to desired item and press Enter.

Commit Applied Software Updates (Remove Saved Files)
Reject Applied Software Updates (Use Previous Version)
Remove Installed Software
Copy Software to Hard Disk for Future Installation
Check Software File Sizes After Installation
Verify Software Installation and Requisites
Clean Up After Failed or Interrupted Installation

F1=Help F2=Refresh F3=Cancel
Esc+8=Image
Esc+9=Shell Esc+0=Exit Enter=Do
```

6. There are two ways in which you can remove the software. Choose either method a or b.
   a. Enter the package name `Ontap.mpio_attach_kit` and make sure the *`preview only`* is set to *`no`*.
   b. Follow these steps:
      i. Press F4 to display a list of names.
      ii. Scroll down the list of names.
      iii. Select Ontap.mpio_attach_kit and press Enter.
      iv. Press Enter to start the uninstall when prompted for confirmation.

An example of a successful removal is shown in Example 5-35.

Example 5-35 Success response

```
installp: bosboot process completed.
+-----------------------------------------------------------------------------+
Summaries:
+-----------------------------------------------------------------------------+
Installation Summary
--------------------
Name                         Level    Part     Event     Result
-----------------------------------------------------------------------------
Ontap.mpio_attach_kit.confi  4.1.0.0  USR     DEINSTALL  SUCCESS
Ontap.mpio_attach_kit.fcp    4.1.0.0  USR     DEINSTALL  SUCCESS
Ontap.mpio_attach_kit.iscsi  4.1.0.0  USR     DEINSTALL  SUCCESS
Ontap.mpio_attach_kit.pcmod  4.1.0.0  USR     DEINSTALL  SUCCESS
```

5.7.2 Installing the Host Utilities kit

The Host Utilities kit for Native OS (AIX) provides support for the AIX native stack and for VIO servers. Ensure that you are running the supported HBAs and MPIO. If you are running an environment like VIO, you must set up that environment first.

Before you install the Host Utilities and SAN Toolkit package:

- Verify that your host meets the requirements identified on the following IBM website:

  http://www.ibm.com/storage/support/nas

- Remove the earlier version of the Host Utilities kit using the instructions in 5.6.14, “Stopping dm-multipath support” on page 178.

The FCP AIX Host Utilities kit includes:

- The host settings file, which is the software package for Native OS using MPIO.
- The Data ONTAP SAN Toolkit file, which includes the sanlun utility and diagnostic scripts.

Both host settings and SAN Toolkit files can be obtained from the following website and can be used for hosts running AIX and VIO:

http://www-304.ibm.com/systems/support/supportsite.wss/selectproduct?brandind=5000029&familyind=5364546&taskind=1

The following steps are used to install both the Host Utilities kit and the SAN Toolkit from the AIX System Management Interface Tool (SMIT). Alternatively, you can use the CLI command **installp** to install packages. Perform these steps:

1. Change to the directory where you have saved the downloaded software packages. Enter the following command to uncompress them:

   ```
   # zcat ibm_aix_mpio_5.0.tar.Z | tar -xvf -
   ```

   The extracted software is placed in three directories: MPIO, SAN, and SAN_Tool_Kit.

2. Log into your host as root. Execute the **oem_setup_env** if you are logged into the host as padmin for VIO environment.
3. Start SMIT by entering the following command:

   ```
   # smit install
   ```

4. At the main menu screen of SMIT select the option **Install and Update Software**.
5. At the next screen, select the **Install Software** option.
6. At the Install Software screen, specify the location of the downloaded packages using one of the following options:
   - Manually, typing the location of the files.
   - By pressing F4, which displays a list of options. If you choose this option:
     i. At the prompt for software to install enter `ibm_aix_mpio_5.0`.
     ii. Press Enter to continue the installation.
7. Press Enter to confirm the installation of the software package when challenged by SMIT.
8. SMIT displays the information about the installation (see Example 5-36) Do not be alarmed if the installation includes both FCP and iSCSI file sets; this is the default behavior.

Example 5-36 installation information

```
+-----------------------------------------------------------------------------+
                                Summaries:
+-----------------------------------------------------------------------------+

Installation Summary
--------------------
Name                        Level           Part        Event       Result
```

```
---------------------------------------------------------------------
Ontap.mpio_attach_kit.pcmod 5.0.0.0        USR         APPLY       SUCCESS
Ontap.mpio_attach_kit.iscsi 5.0.0.0        USR         APPLY       SUCCESS
Ontap.mpio_attach_kit.fcp   5.0.0.0        USR         APPLY       SUCCESS
Ontap.mpio_attach_kit.confi 5.0.0.0        USR         APPLY       SUCCESS

The following is displayed upon a successful install of the SAN Toolkit:
+--------------------------------------------------------------------+
                                Summaries:
+--------------------------------------------------------------------+

Installation Summary
--------------------
Name                        Level           Part        Event       Result
--------------------------------------------------------------------
Ontap.SAN_toolkit.scripts   5.0.0.0         USR         APPLY       SUCCESS
Ontap.SAN_toolkit.sanlun    5.0.0.0         USR         APPLY       SUCCESS
```

9. This completes the installation of the Host Utilities kit software package. Check smit.log and smit.script for installation log.

Note: These installation steps install the Host Utilities kit. Repeat steps 3 to 8 to install the SAN Toolkit package, santoolkit_aix_ibm_5.0.tar.Z. The SAN Toolkit must be installed on each VIO client in addition to the server.

This book uses names for the packages in the examples to make the examples easier to read. The final names for these packages and sequence of operations may differ from those used in examples. Do not be concerned about minor differences. Refer to the appropriate documentation on the IBM Support website for current details.

Check the HBA initiator queue depth after you install both software packages. Enter the following command to query the queue depth:

```
# lsattr -El fcs2
num_cmd_elems 200       Maximum number of COMMANDS to queue to the adapter
True

# lsattr -El fcs5
num_cmd_elems 200       Maximum number of COMMANDS to queue to the adapter
True
```

Note: The default values for num_cmd_elems of 200 for the HBAs and queue_depth of 12 on hdisk handles are a good starting point.

5.7.3 N series FCP cluster mode

We assume that your N series cluster is configured to FCP cfmode single_image. Refer to Chapter 2, “Active/active configuration and management” on page 49, for a discussion of the merits of this FCP cfmode.

5.7.4 Creating, configuring, and managing LUNs without VIO

This section describes how to prepare the N series storage subsystem for LUN configuration, and also how to create and configure LUNs.

Creating and mapping igroups and LUNs

Refer to the *IBM Data ONTAP Block Management Access Management Guide for FCP* for information about creating and mapping igroup and LUNs on N series storage subsystems at:

```
www.ibm.com/storage/support/nas
```

LUN configuration overview

Use the native AIX MPIO to provide multiple path access to LUNs on non-VIO hosts by performing these steps:

1. Log into the host as root.
2. Run the following command to discover new LUNs created:

   ```
   # cfgmgr
   ```

3. Run the following command to verify that the hdisk MPIO devices have been created. Note the hdisk instance numbers for path configuration:

   ```
   # lsdev -Ccdisk

   hdisk0  Available 1S-08-00-8,0 16 Bit LVD SCSI Disk Drive
   hdisk1  Available 1S-08-00-9,0 16 Bit LVD SCSI Disk Drive
   hdisk39 Available 1n-08-02     MPIO Ontap FCP Default PCM Disk
   hdisk40 Available 1D-08-02     MPIO Ontap FCP Default PCM Disk
   ```

4. Run the **lsattr -El hdisk39** command to acquire information regarding your setup (Example 5-37).

Example 5-37 Setup information

```
# lsattr -El hdisk39
PCM             PCM/friend/OntapDefaultPCM Path Control Module            False
algorithm       round_robin                Algorithm                      True
clr_q           no                         Device CLEARS its Queue on error True
dist_err_pcnt   0                          Distributed Error Sample Time    True
dist_tw_width   50                         Distributed Error Sample Time    True
hcheck_cmd      inquiry                    Health Check Command           True
hcheck_interval 30                         Health Check Interval          True
hcheck_mode     nonactive                  Health Check Mode              True
location                                   Location Label                 True
lun_id          0x0                        Logical Unit Number ID         False
lun_reset_spt   yes                        LUN Level Reset                True
max_transfer    0x40000                    Maximum TRANSFER Size          True
node_name       0x500a09808647e7ba         FC Node Name                   False
pvid            none                       Physical volume identifier     False
q_err           yes                        Use QERR bit                   True
q_type          simple                     Queuing TYPE                   True
queue_depth     12                         Queue DEPTH                    True
reassign_to     120                        REASSIGN time out value        True
reserve_policy  no_reserve                 Reserve Policy                 True
rw_timeout      30                         READ/WRITE time out value      True
scsi_id         0x6f0800                   SCSI ID                        False
start_timeout   60                         START unit time out value      True
ww_name         0x500a09828647e7ba         FC World Wide Name             False

# lsattr -El hdisk40
PCM             PCM/friend/OntapDefaultPCM Path Control Module            False
algorithm       round_robin                Algorithm                      True
clr_q           no                         Device CLEARS its Queue on error True
dist_err_pcnt   0                          Distributed Error Sample Time    True
dist_tw_width   50                         Distributed Error Sample Time    True
```

```
hcheck_cmd      inquiry                     Health Check Command            True
hcheck_interval 30                          Health Check Interval           True
hcheck_mode     nonactive                   Health Check Mode               True
location                                    Location Label                  True
lun_id          0x0                         Logical Unit Number ID          False
lun_reset_spt   yes                         LUN Level Reset                 True
max_transfer    0x40000                     Maximum TRANSFER Size           True
node_name       0x500a09808647e7ba          FC Node Name                    False
pvid            none                        Physical volume identifier      False
q_err           yes                         Use QERR bit                    True
q_type          simple                      Queuing TYPE                    True
queue_depth     12                          Queue DEPTH                     True
reassign_to     120                         REASSIGN time out value         True
reserve_policy  no_reserve                  Reserve Policy                  True
rw_timeout      30                          READ/WRITE timeout value       True
scsi_id         0x6f0800                    SCSI ID                         False
start_timeout   60                          START unit timeout value       True
ww_name         0x500a09828647e7ba          FC World Wide Name                False
```

5. You might have to order and prioritize the path to the N series storage system dependent on the version of Data ONTAP that supports Asymmetric Logical Unit Access (ALUA).
   – Refer to *Data ONTAP Block Access Management Guide for FCP* if the version of Data ONTAP supports ALUA.
   – Otherwise, you must execute the dotpaths utility (Example 5-38) to set the path priority, which is installed during the Host Utilities kit installation.

Example 5-38 dotpaths utility

```
# dotpaths -v
hdisk39 (path 0): Optimized path - Current priority of 255 is correct
hdisk40 (path 0): Optimized path - Current priority of 255 is correct
Path priority set and/or verified for 2 disks, 2 total paths.
```

6. Run the **sanlun lun show -p** command to display information about your MPIO setup (Example 5-39).

Example 5-39 sanlun lun show command

```
# sanlun lun show -p
ONTAP_PATH: itsotuc4:/vol/aix_mpio/mpio1
LUN: 0
LUN Size:     1g (1073741824)
Host Device: hdisk40
LUN State: GOOD
Controller_CF_State: Cluster Enabled
Controller Partner:
Multipath Provider: AIX Native
Multipathing Algorithm: round_robin

--------- ----------- ------ ------ ----------- ----------
MPIO      Controller  AIX           Controller  AIX MPIO
path      path        MPIO   host   target HBA  path
status    type        path   HBA    port        priority
--------- ----------- ------ ------ ----------- ----------
Enabled       primary path0  fcs5       0b         255
Enabled       primary path1  fcs2       0b         1
```

```
ONTAP_PATH: itsotuc4:/vol/aix_mpio/mpio0
LUN: 0
LUN Size:     1g (1073741824)
Host Device: hdisk39
LUN State: GOOD
Controller_CF_State: Cluster Enabled
Controller Partner:
Multipath Provider: AIX Native
Multipathing Algorithm: round_robin
```

Note: You must determine which device to use as the physical volume for the LVM before you create a volume group. You can use the **lsdev -Cc disk** command to determine which host disk device MPIO uses as the I/O access device for the LUN.

5.7.5 Creating volume groups and file systems on AIX hosts

This section describes one of the methods for setting up volume groups and file sysytems on AIX hosts.

Creating volume groups

You can create volume groups using numerous tools on an AIX host. We used SMIT for this example. Refer to AIX documentation for additional information about managing volume groups, logical volumes, and file systems.

Follow these steps to use the AIX SMIT tool to create a new volume group once you have determined which device to use as the physical volume.

1. Launch the tool by running the following:

   ```
   # smit vg
   ```

2. Provide a name for the new volume group and the desired physical devices in SMIT. Select **Add a Volume Group** to create a new volume group (Example 5-40).

   Example 5-40 Adding a Volume grouop

   ```
   Add an Original Volume Group

   Type or select values in entry fields.
   Press Enter AFTER making all desired changes.

                                             [Entry Fields]
     VOLUME GROUP name                      [nseries_mpio]    +
     Physical partition SIZE in megabytes                     +
   * PHYSICAL VOLUME names                  [hdisk39 hdisk40] +
     Force the creation of a volume group? no                 +
     Activate volume group AUTOMATICALLY   yes                +
       at system restart?
     Volume Group MAJOR NUMBER              []                +#
     Create VG Concurrent Capable?          no                +
   ```

3. Press Enter to continue creating the volume group.

   ```
   COMMAND STATUS

   Command: OK  stdout: yes  stderr: no
   ```

```
Before command completion, additional instructions may appear below.

0516-1254 /usr/sbin/mkvg: Changing the PVID in the ODM.
0516-1254 /usr/sbin/mkvg: Changing the PVID in the ODM.
nseries_mpio
```

4. Press F10 to exit the SMIT tool.
5. To verify the new volume group information run the following command:

   ```
   # lsvg nseries_mpio
   ```

   Our results are shown in Example 5-41.

Example 5-41 Verifying the new volume group

```
VOLUME GROUP:       nseries_mpio    VG IDENTIFIER:
0007041a00004c000000011df8d2d883
VG STATE:           active          PP SIZE:        4 megabyte(s)
VG PERMISSION:      read/write      TOTAL PPs:      510 (2040 megabytes)
MAX LVs:            256             FREE PPs:       510 (2040 megabytes)
LVs:                0               USED PPs:       0 (0 megabytes)
OPEN LVs:           0               QUORUM:         2 (Enabled)
TOTAL PVs:          2               VG DESCRIPTORS: 3
STALE PVs:          0               STALE PPs:      0
ACTIVE PVs:         2               AUTO ON:        yes
MAX PPs per VG:     32512
MAX PPs per PV:     1016            MAX PVs:        32
LTG size (Dynamic): 256 kilobyte(s) AUTO SYNC:      no
HOT SPARE:          no              BB POLICY:      relocatable
```

Setting up a new file system

You can access storage space on a volume group using numerous methods. We used SMIT in this example. Refer to IBM documentation for additional information about managing file systems.

Use these steps to set up a new file system:

1. Run the following command:

   ```
   # smit fs
   ```

2. Select the appropriate menu options (Example 5-42).

Example 5-42 Accessing storage space

```
 → Add / Change / Show / Delete File Systems
 → Journaled File Systems
 → Add a Journaled File System
 → Add a Standard Journaled File System
Move cursor to desired item and press Enter.
Add a Standard Journaled File System
  Add a Compressed Journaled File System
  Add a Large File Enabled Journaled File System

   +-------------------------------------------------+
   |  Volume Group Name                              |
   |                                                 |
```

```
    | Move cursor to desired item and press Enter.   |
    |                                                |
    |  rootvg                                        |
    |  ermtstvg                                      |
    |  nseries_mpio                                  |
    |                                                |
    | F1=Help   F2=Refresh  F3=Cancel                |
    | F8=Image  F10=Exit    Enter=Do                 |
F1=| /=Find    n=Find Next                          |    F8=Image
F9=+------------------------------------------------+
```

3. Select the volume group that you created previously.

Example 5-43 Volume group selection

```
Add a Standard Journaled File System

Type or select values in entry fields.
Press Enter AFTER making all desired changes.

                                          [Entry Fields]
  Volume group name                       nseries_mpio
  SIZE of file system
          Unit Size                       Megabytes     +
*         Number of units                 [1024]        #
* MOUNT POINT                             [/mpio_fs]
  Mount AUTOMATICALLY at system restart?   no           +
  PERMISSIONS                              read/write   +
  Mount OPTIONS                           []            +
  Start Disk Accounting?                   no           +
  Fragment Size (bytes)                    4096         +
  Number of bytes per inode                4096         +
  Allocation Group Size (MBytes)           8            +
  Logical Volume for Log                                +

F1=Help   F2=Refresh  F3=Cancel  F4=List
F5=Reset  F6=Command  F7=Edit    F8=Image
F9=Shell  F10=Exit    Enter=Do
```

4. Press Enter to continue the creation of the new file system (Example 5-44).

Example 5-44 Creating new filesystem

```
COMMAND STATUS

Command: OK           stdout: yes          stderr: no

Before command completion, additional instructions may appear below.

Based on the parameters chosen, the new /mpio_fs JFS file system
is limited to a maximum size of 134217728 (512 byte blocks)

New File System size is 2097152

F1=Help      F2=Refresh  F3=Cancel  F6=Command
F8=Image     F9=Shell    F10=Exit   /=Find
```

```
n=Find Next
```

5. Press F10 to exit SMIT.
6. Mount and verify the new file system

```
# mount /mpio_fs
# df -k /mpio_fs
Filesystem    1024-blocks      Free %Used    Iused %Iused Mounted on
/dev/lv00         1048576   1015612    4%       17     1% /mpio_fs
```

5.7.6 Using sanlun to display host LUN information

Perform these steps to display LUN information:

1. Log into your host as root.
2. Type the following into your CLI to display the LUN information:

   `sanlun lun show [-v] [-d host device filename | all | filername | filername:filer_pathname]`

   The parameters are:

   -v Verbose output (cannot be used with -d).
   -d Device option. Choose from the following
   host device filename: The special device file on the host.
   all: All attached LUNs on your host.
   filername: Host name of the N series storage system.
   filer_pathname: Path name of the LUN on the N series storage subsystem.

The following examples show the output of the **sanlun** command with various options specified.

Example 5-45 shows the display listing of all paths associated with the LUN (cannot be used with -d option).

Example 5-45 Display LUN paths

```
# sanlun lun show -p
ONTAP_PATH: itsotuc4:/vol/aix_mpio/mpio1
LUN: 0
LUN Size:     1g (1073741824)
Host Device: hdisk40
LUN State: GOOD
Controller_CF_State: Cluster Enabled
Controller Partner:
Multipath Provider: AIX Native
Multipathing Algorithm: round_robin
--------- ----------- ------ ------ ----------- ----------
MPIO      Controller  AIX           Controller  AIX MPIO
path      path        MPIO   host   target HBA  path
status    type        path   HBA    port        priority
--------- ----------- ------ ------ ----------- ----------
Enabled       primary path0  fcs5       0b         255
Enabled       primary path1  fcs2       0b         1

ONTAP_PATH: itsotuc4:/vol/aix_mpio/mpio0
```

```
LUN: 0
LUN Size:     1g (1073741824)
Host Device: hdisk39
LUN State: GOOD
Controller_CF_State: Cluster Enabled
Controller Partner:
Multipath Provider: AIX Native
Multipathing Algorithm: round_robin
--------- ----------- ------ ------ ----------- ----------
MPIO      Controller  AIX           Controller  AIX MPIO
path      path        MPIO   host   target HBA  path
status    type        path   HBA    port        priority
--------- ----------- ------ ------ ----------- ----------
Enabled       primary path0  fcs2       0b         255
Enabled       primary path1  fcs5       0b         1
```

Example 5-46 is the summary listing of the LUNs associated with host device /dev/hdisk40.

Example 5-46 Display LUN served by device

```
# sanlun lun show -d /dev/hdisk40
filer:  lun-pathname  device filename  adapter  protocol   lun size      lun state
itsotuc4: /vol/aix_mpio/mpio1 hdisk40   fcs5    FCP       1g (1073741824)     GOOD
```

Example 5-47 displays verbose output of LUNs available on host.

Example 5-47 Display verbose output of LUN

```
#sanlun lun show -v all
filer:   lun-pathname    device  filename adapter protocol  lun size     lun state
 itsotuc4:  /vol/aix_mpio/mpio0  hdisk39    fcs2    FCP     1g (1073741824)   GOOD
        Serial number: C4a8loMF2C5X
        Filer FCP nodename:500a09808647e7ba  Filer FCP portname:500a09828647e7ba
        Filer adapter name: 0b
        Filer IP address:       09.11.218.238
        Filer volume name:aix_mpio   FSID:0x273bd977
        Filer qtree name:/vol/aix_mpio   ID:0x0
        Filer snapshot name:   ID:0x0
 itsotuc4:  /vol/aix_mpio/mpio1  hdisk40     fcs5    FCP   1g (1073741824)   GOOD
        Serial number: C4a8loMF2Lhk
        Filer FCP nodename:500a09808647e7ba  Filer FCP portname:500a09828647e7ba
        Filer adapter name: 0b
        Filer IP address:       09.11.218.238
        Filer volume name:aix_mpio   FSID:0x273bd977
        Filer qtree name:/vol/aix_mpio   ID:0x0
        Filer snapshot name:   ID:0x0
```

Example 5-48 shows a summary listing of all the LUNs available on the host served by N series storage subsystem `itsotuc4`.

Example 5-48 Display LUN served by storage subsystem

```
# sanlun lun show itsotuc4
filer:   lun-pathname   device filename  adapter  protocol  lun size     lun state
 itsotuc4:  /vol/aix_mpio/mpio0  hdisk39   fcs2    FCP    1g (1073741824)     GOOD
```

```
itsotuc4:  /vol/aix_mpio/mpio1  hdisk40   fcs5    FCP    1g (1073741824)     GOOD
```

5.7.7 Displaying host HBA information

Perform these steps to display LUN information:

1. Log into your host as root.
2. Type the following into your CLI to display the LUN information:

   `sanlun fcp show adapter [ -c | [-v] [adapter_name | all ]]`

   The parameters are:

-v	verbose output
-c	configuration information to create igroups
all	information for all FCP adapters
adapter_name	information for specified adapter

To display detailed information about all FCP adapters, run the command in Example 5-49.

Example 5-49 Display FCP adapter information

```
# sanlun fcp show adapter -v
adapter name:      fcs2
WWPN:              10000000c955e566
WWNN:              20000000c955e566
driver name:       /usr/lib/drivers/pci/efcdd
model:             df1000f9
model description: FC Adapter
serial number:     1F6240C38A
hardware version:  Not Available
driver version:    6.1.1.0
firmware version:  391304
Number of ports:   1
port type:         Fabric
port state:        Operational
supported speed:   2 GBit/sec
negotiated speed:  2 GBit/sec
OS device name:    fcs2

adapter name:      fcs5
WWPN:              10000000c955e581
WWNN:              20000000c955e581
driver name:       /usr/lib/drivers/pci/efcdd
model:             df1000f9
model description: FC Adapter
serial number:     1F6240C135
hardware version:  Not Available
driver version:    6.1.1.0
firmware version:  391304
Number of ports:   1
port type:         Fabric
port state:        Operational
supported speed:   2 GBit/sec
negotiated speed:  2 GBit/sec
OS device name:    fcs5
```

If you omit the **-v** option while running this command, only the WWPN associated with the adapter will be displayed:

```
# sanlun fcp show adapter
fcs2            WWPN:10000000c955e566
fcs5            WWPN:10000000c955e581
```

6

Disk sanitization

This chapter explains disk sanitization and the process of physically removing data from a disk by overwriting patterns on the disk in a manner that precludes the recovery of that data by any known recovery methods.

It also presents the Data ONTAP disk sanitization feature and briefly addresses data confidentiality, technology drivers, costs and risks, and the sanitizing operation.

The following topics are covered:

- Data ONTAP disk sanitization
- Data confidentiality
- Data ONTAP sanitization operation

6.1 Data ONTAP disk sanitization

The Data ONTAP disk sanitization feature enables you to carry out disk sanitization by using three successive byte overwrite patterns per cycle and a default six cycles per operation (Example 6-1). Previously, compliance was done with crude and inefficient methods of degaussing, demagnetizing, disk drive destruction, or retaining and storage of disk drives.

Example 6-1 Disk sanitization progress

```
Tue Jun 24 02:40:10 Disk sanitization initiated on drive 8a.43 [S/N
3FP20XX400007313LSA8]

Tue Jun 24 02:40:10 Disk sanitization initiated on drive 8a.44 [S/N
3FP0RFAZ00002218446B]

Tue Jun 24 02:40:10 Disk sanitization initiated on drive 8a.45 [S/N
3FP0RJMR0000221844GP]

Tue Jun 24 02:53:55 Disk 8a.44 [S/N 3FP0RFAZ00002218446B] format
completed in 00:13:45.

Tue Jun 24 02:53:59 Disk 8a.43 [S/N 3FP20XX400007313LSA8] format
completed in 00:13:49.

Tue Jun 24 02:54:04 Disk 8a.45 [S/N 3FP0RJMR0000221844GP] format
completed in 00:13:54.

Tue Jun 24 02:54:11 Disk 8a.44 [S/N 3FP0RFAZ00002218446B] cycle 1
pattern write of 0x47 completed in 00:00:16.

Tue Jun 24 02:54:11 Disk sanitization on drive 8a.44 [S/N
3FP0RFAZ00002218446B] completed.

Tue Jun 24 02:54:15 Disk 8a.43 [S/N 3FP20XX400007313LSA8] cycle 1
pattern write of 0x47 completed in 00:00:16.

Tue Jun 24 02:54:15 Disk sanitization on drive 8a.43 [S/N
3FP20XX400007313LSA8] completed.

Tue Jun 24 02:54:20 Disk 8a.45 [S/N 3FP0RJMR0000221844GP] cycle 1
pattern write of 0x47 completed in 00:00:16.

Tue Jun 24 02:54:20 Disk sanitization on drive 8a.45 [S/N
3FP0RJMR0000221844GP] completed.

Tue Jun 24 02:58:42 Disk sanitization initiated on drive 8a.43 [S/N
3FP20XX400007313LSA8]

Tue Jun 24 03:00:09 Disk sanitization initiated on drive 8a.32 [S/N
43208987]

Tue Jun 24 03:11:25 Disk 8a.32 [S/N 43208987] cycle 1 pattern write
of 0x47 completed in 00:11:16.

Tue Jun 24 03:12:32 Disk 8a.43 [S/N 3FP20XX400007313LSA8]
sanitization aborted by user.

Tue Jun 24 03:22:41 Disk 8a.32 [S/N 43208987] cycle 2 pattern write
of 0x47 completed in 00:11:16.
```

```
Tue Jun 24 03:22:41 Disk sanitization on drive 8a.32 [S/N 43208987]
completed.
```

6.2 Data confidentiality

In every industry, IT managers face increasing pressure to ensure the confidentiality of corporate, client, or patient data. In addition, companies and managers in certain industries must comply with laws that specify strict standards for handling, distributing, and using confidential client, corporate, or patient information.

Although there are methods and products to aid in data storage and transmission security as the data moves through the system, assuring confidentiality of data on desktop or notebook computers when they leave the premises for disposal presents a different set of challenges and exposures. In the following sections we lay out some of those challenges and attempt to demonstrate the value of third-party disposal.

6.2.1 Background

Data confidentiality has always been an issue of ethical concern. But with the enactment of laws to protect the privacy of individual health and financial records, it has become a legal concern as well.

Most IT managers have some kind of strategy in place for securing customer information within their networks and, especially in the healthcare industry, controlling data interchange with vendors to ensure patient privacy.

The market offers various products and services to assist managers with these challenges. Many offer ways to integrate confidentiality and compliance into daily operations.

6.2.2 Technology drivers

As technology advances, upgrades, disk subsystem replacements, and data life-cycle management will require the movement or migration of data. To ensure that the data movement does not create a security risk by leaving data patterns behind, IBM System Storage N series offers the disk sanitization feature.

You might also sanitize disks if you want to help ensure that data currently on those disks is physically unrecoverable. For example, you might have some disks that you intend to remove from one storage system and you want to re-use those disks in another appliance or simply dispose of the disks (Figure 6-1).

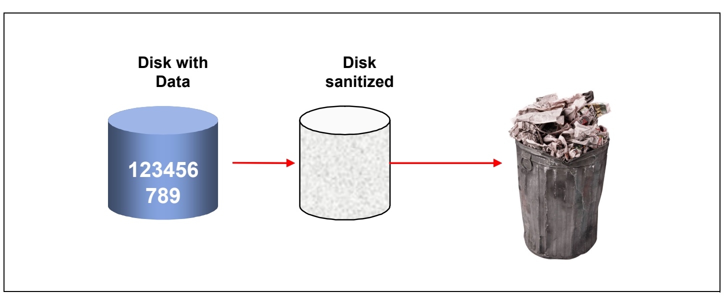


Figure 6-1 Disposing of disks

6.2.3 Costs and risks

There are two critical factors that all enterprises must consider when making a decision involving the cost and risk of hard drive sanitization practices.

- The cost of running sanitization programs on a fleet of computers can be prohibitive. Even in smaller organizations, the number of hard drives that must be cleansed can be unmanageable. Most IT managers do not have the time or resources to accomplish such a task without impacting other core business responsibilities. Should a company choose to circumvent these costs and simply destroy their hard drives (many of which could be reused), they dispose of equipment that still has market value.
- At the same time, companies must recognize the significant risk associated with breaches of private information. When companies do not properly sanitize exiting storage devices, they expose themselves to a myriad of public relations, legal, and business repercussions should any confidential data be leaked. Because governments around the world continue to pass and enforce regulations for electronic data security, IT managers must act quickly to adopt and implement appropriate hard drive sanitization practices.

Using Data ONTAP, IBM System Storage N series offers an effective sanitization method, while reducing both costs and risks. The disk sanitization algorithims are built into Data ONTAP and require only licensing and no additional software installation.

6.3 Data ONTAP sanitization operation

With the **disk sanitize start** command, Data ONTAP begins the sanitization process on each of the specified disks. The process consists of a disk format operation, followed by the specified overwrite patterns repeated for the specified number of cycles.

The time to complete the sanitization process for each disk depends on the size of the disk, the number of patterns specified, and the number of cycles specified.

The following command starts the sanitization process on the disks listed:

```
disk sanitize start
[-p <pattern>|-r [-p <pattern>|-r [-p <pat_tern>|-r]]] [-c <cycles>] <disk_list>
```

- The -p option defines the byte patterns and the number of write passes in each cycle.
- The -r option can be used to generate a write of random data, instead of a defined byte pattern.
- If no patterns are specified, the default is 3 using pattern 0x55 on the first pass, 0xaa on the second, and 0x3c on the third.
- The -c option specifies the number of cycles of pattern writes. The default is 1 cycle.

All sanitization process information is written to the log file at /etc/sanitization.log. The serial numbers of all sanitized disks are written to /etc/sanitized_disks.

The command shown in Example 6-2 invokes one format overwrite pass and 18 pattern overwrite passes of disk 7.3.

Example 6-2 disk sanitize start command

```
disk sanitize start -p 0x55 -p 0xAA -p 0x37 -c 6 7.3
```

The sanitization process can take a long time. To view the progress, use the **disk sanitize status** command, as shown in Example 6-3.

Example 6-3 disk sanitize status command

```
itsotuc4*> disk sanitize status
sanitization for 0c.24 is 10 % complete
```

The **disk sanitize release** command allows the user to return a sanitized disk to the spare pool.

The **disk sanitize abort** command is used to terminate the sanitization process for the specified disks:

```
disk sanitize abort <disk_list>
```

If the disk is in the format stage, the process will be aborted when the format is complete. A message will be displayed when the format is complete and when an abort is complete.

7

System Manager

This chapter discusses the N series System Manager tool, which allows the user to manage the N series storage system with limited experience and knowledge of the N series hardware and software features. System Manager helps with both setup and basic administration tasks and can help the user to manage the box without accessing the Filer View, which requires more technical skill.

The following topics are covered:

- System Manager requirements
- Installation steps
- N series systems discovery
- N series system setup
- N series system configuration.

7.1 Introduction to System Manager

System Manager is a new and easy way to administer N series systems. It is a stand-alone application and is executed as a Microsoft Management Console (MMC) snap-in. It does not require any knowledge of N series.

System Manager key features include:

- System setup and configuration management
- Protocol management (NFS, CIFS, iSCSI and FCP)
- Shares/exports management
- Storage management (volumes, aggregates, disks, qtrees)

Currently the following Windows desktop platforms are supported:

- Windows XP, Vista, Server 2003, and 2008

System Manager release 1.0.1 supports Data ONTAP 7.2.3 -7.3.x. For more details and updated information check the IBM support website at:

http://www.ibm.com/storage/support/nas/

The three N series management tools are:

- System Manager (no charge)
  - A solution for managing a single N series storage system.
- FilerView (no charge)
  - A device manager for systems not currently qualified for use with System Manager (gateways and Data ONTAP releases prior to 7.2.3 and starting from 8.0)
  - Also recommended for features not yet supported in System Manager (for example, basic SnapMirror® management, quotas, MetroCluster platforms)
- Operations Manager (cost depending on product purchased)
  - An N series solution for managing multiple N series storage systems
  - Scalable management, monitoring, and reporting software for enterprise-class environments.

System Manager is shipped on CD with N series hardware or it can be downloaded from the IBM website at:

http://www.ibm.com/storage/support/nas/

7.2 System Manager installation

Before you begin installing System Manager, check the compatability matrix and prerequisites to ensure you have an appropriate environment set up. Perform the following steps to install System Manager:

1. Download the System Manager to your desktop and run the .exe file.
2. When prompted by the setup wizard, accept the licence agreement and select your installation folder. Specify whether others can use the application, or just you.
3. Click **Next** for the confirm installation page, and click **Next** again.
4. When the installation is complete, click the **System Manager** icon on the desktop to launch the application.

7.3 System Manager use

When you first launch the IBM System Manager application a Welcome panel is displayed. You have two options here:

- Add storage system: Select this option to add a storage system or an active/active cluster to your configuration if you know its N series host name or IP address. Click **Add**.
- Discover storage systems: This option will find the storage systems on a network subnet. Click **Discover**.

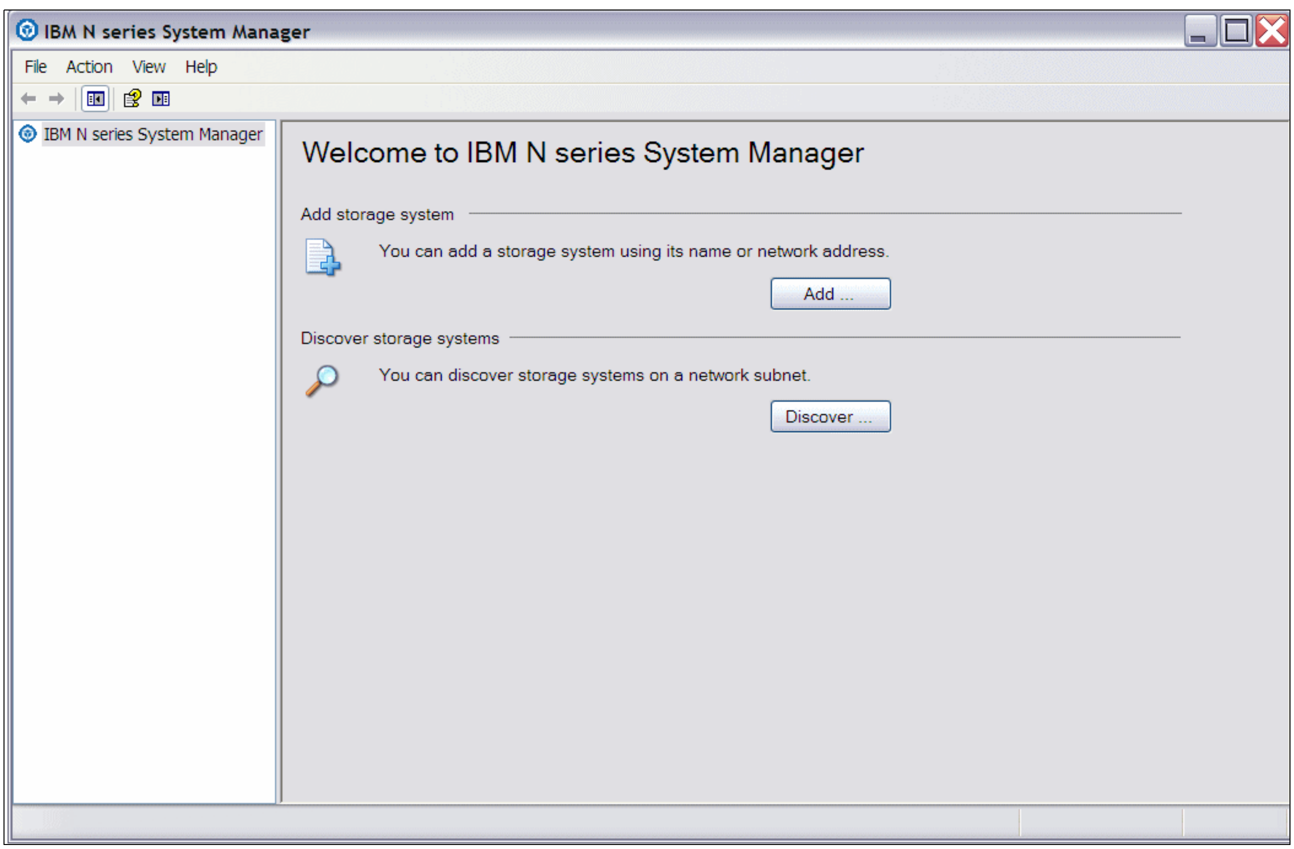


Figure 7-1 Option screen

We selected the first option and provided the IP address of our N series system in the **Add** box returned[1]. The resulting screen is shown in Figure 7-2.

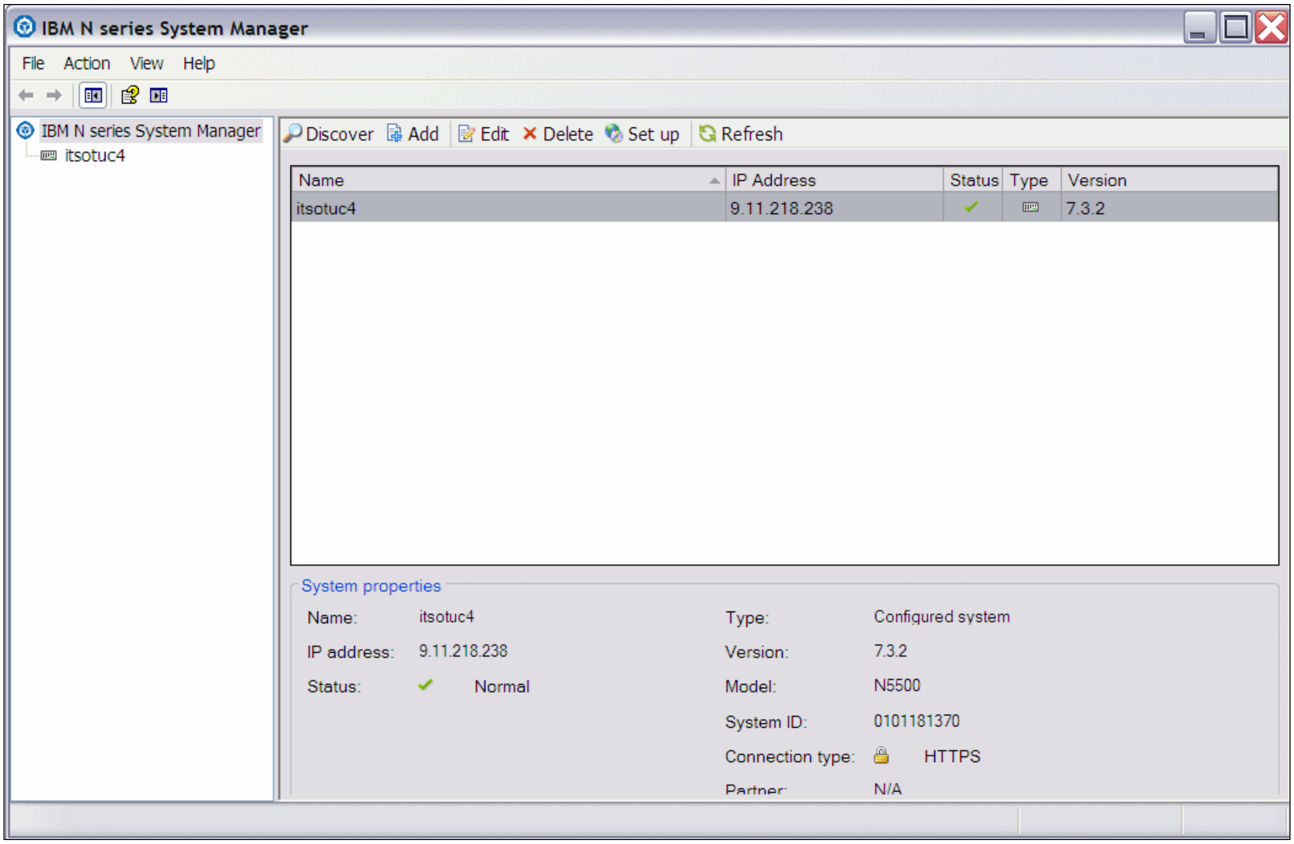


Figure 7-2 Selection result

Check that your system is included in the managed system list. The host Name (itsotuc4 is the only N series in the list), IP Address, Status, Type, and Data ONTAP level are identified in the upper section of the screen. The lower section provides some details about the box. Double-click the host name in the left panel to access the next data screen, which provides additional details about the properties and performance of your storage system (Figure 7-3 on page 203).

[1] Here you can also specify the SNMP community used to discover your storage systems if you are using a non-public SNMP community

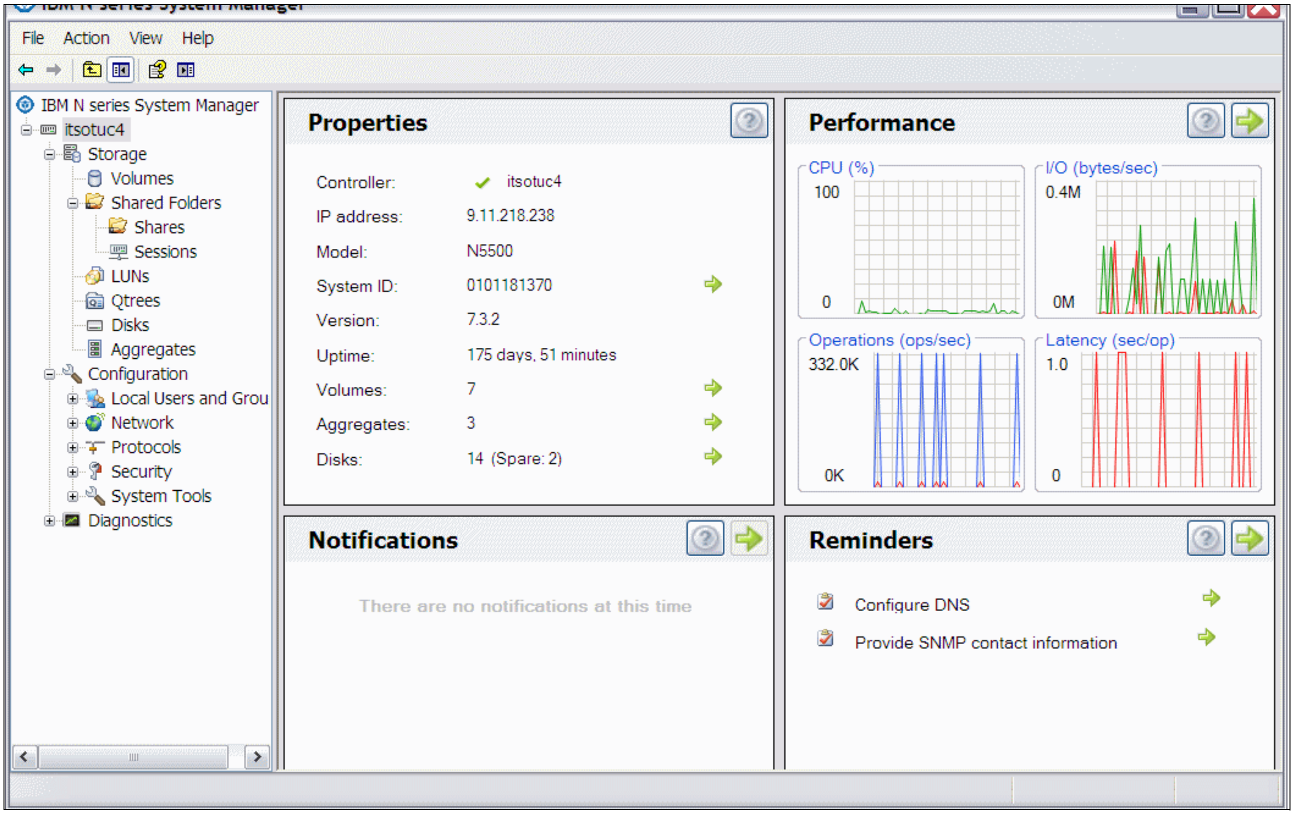


Figure 7-3 N series properties

This page displays details about your current configuration in an easy to use format. The panes of the display are:

- Properties: The top left pane shows the properties of the file system. The green arrows in the display indicate fields for which you can access panels to perform management tasks. Click the appropriate arrow to manage, for example, Volumes or Aggregates.
- Performance: The top right pane presents key performance data in a graphical format. The metrics presented are CPU usage (%), I/O rate (byte/sec), operations (ops/sec) and latency (sec/ops). To view this data numerically rather than graphically click the arrow at the top right of the pane.
- Notifications: The bottom left pane lists the five highest severity notifications for the individual storage system.
- Reminders: The bottom right pane shows reminders about any optional but important configuration tasks that were not performed during application setup and provides links to access the pertinent setup panels.

Figure 7-4 shows the Aggregate details panel that was returned when we clicked the arrow in the Aggregates field of the Properties pane. The same screen would be accessed by clicking the Aggregates line in the left side navigation pane.

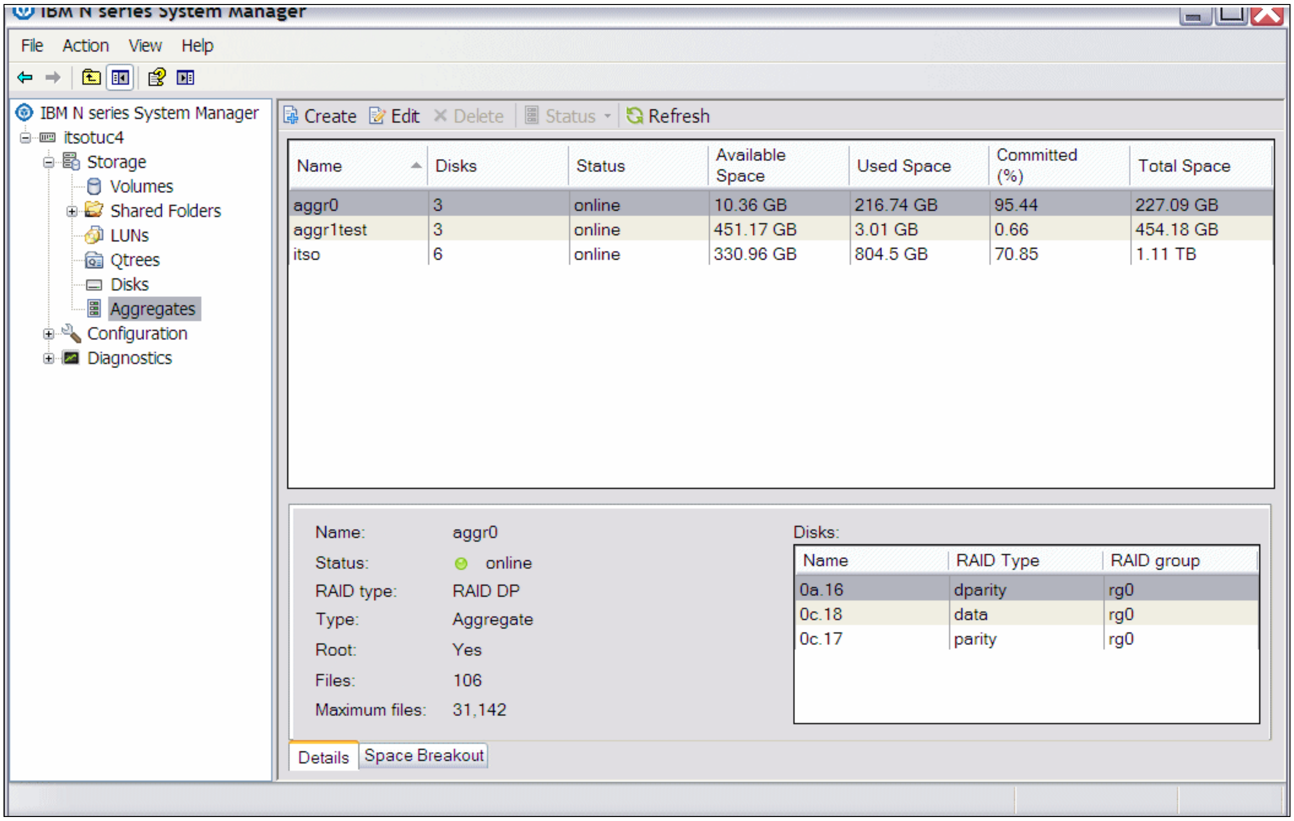


Figure 7-4 Aggregate configuration

This panel includes the navigation pane and the aggregates list. You can view Details or the Space Breakout for each aggregate listed by highlighting the aggregate name and clicking the tab of interest at the bottom of the detail pane.

Creating a new aggregate

At the top of the list pane are action buttons used to manage the storage system. Management is easy because wizards guide you through each configuration task. For example, use the following steps to add a new aggregate to your N series configuration:

1. Click **Create.**

2. In the Create Aggregate Wizard returned, specify details for the new aggregate (Figure 7-5). Type the aggregate name (we entered `aggr1`), the RAID type (we chose RAID-4). You can let the application select the disk drives automatically or you can manually select which drives will be used to define the new aggregate. Also on this screen you select the drive type, where applicable.

   When your selections are made click **Next**.

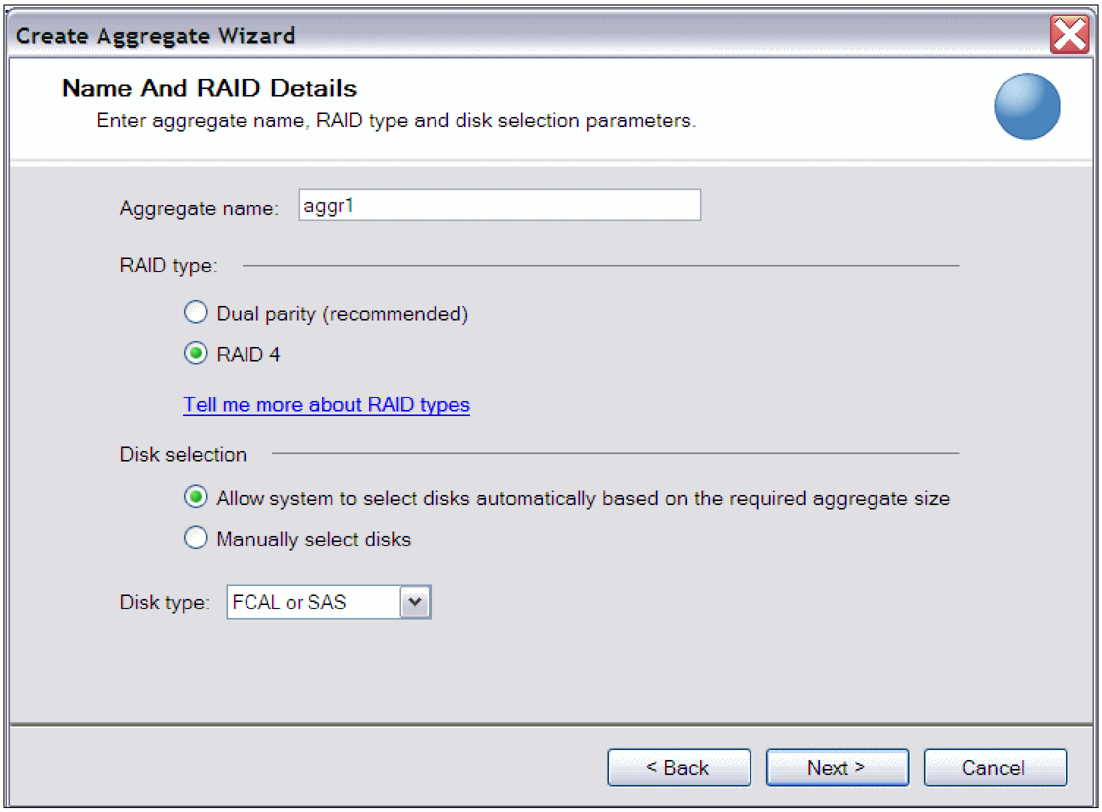


Figure 7-5 Aggregate naming

3. Specify the aggregate size on the next panel (Figure 7-6). Review the available options displayed by the wizard, then make your selection by sliding the pointer on the capacity scale and check that the number of drives is correct. Click **Next**.

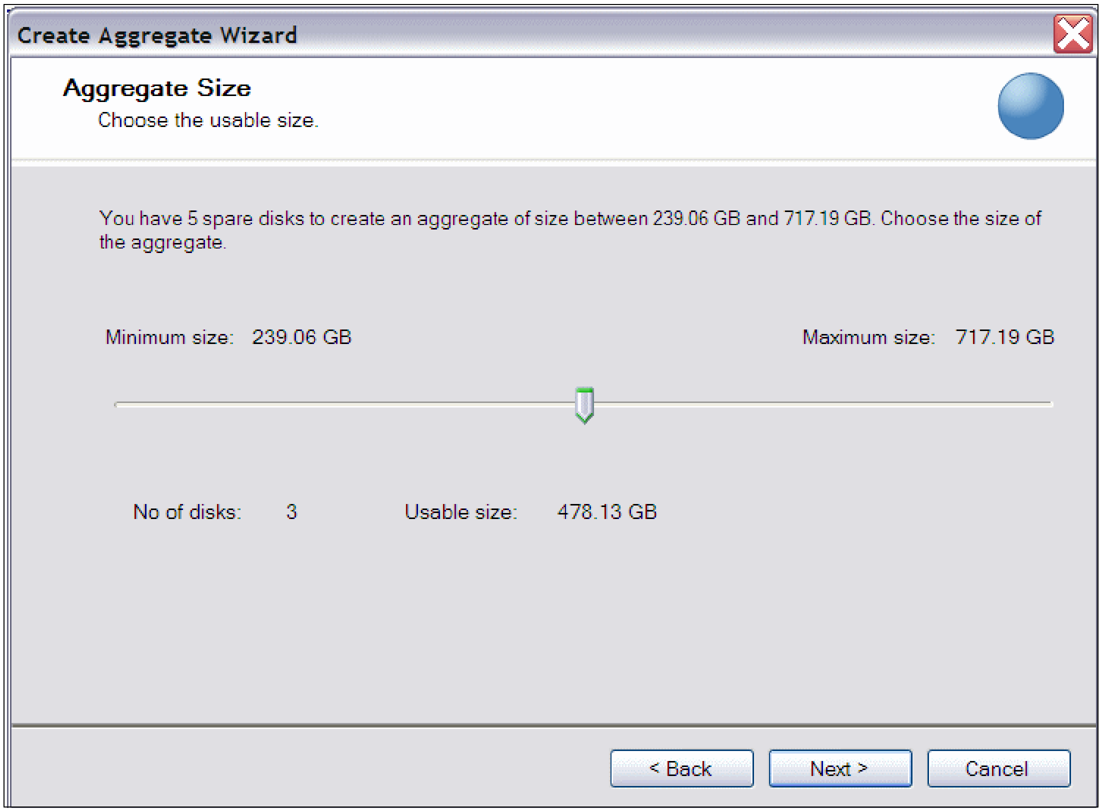


Figure 7-6 Aggregate size

4. The Aggregate Summary panel with all technical information is displayed (Figure 7-7 on page 207). Check that the values shown match your requirements, then click **Next** to begin the actual creation of the aggregate. Progress is indicated on the panel shown in Figure 7-8 on page 207.

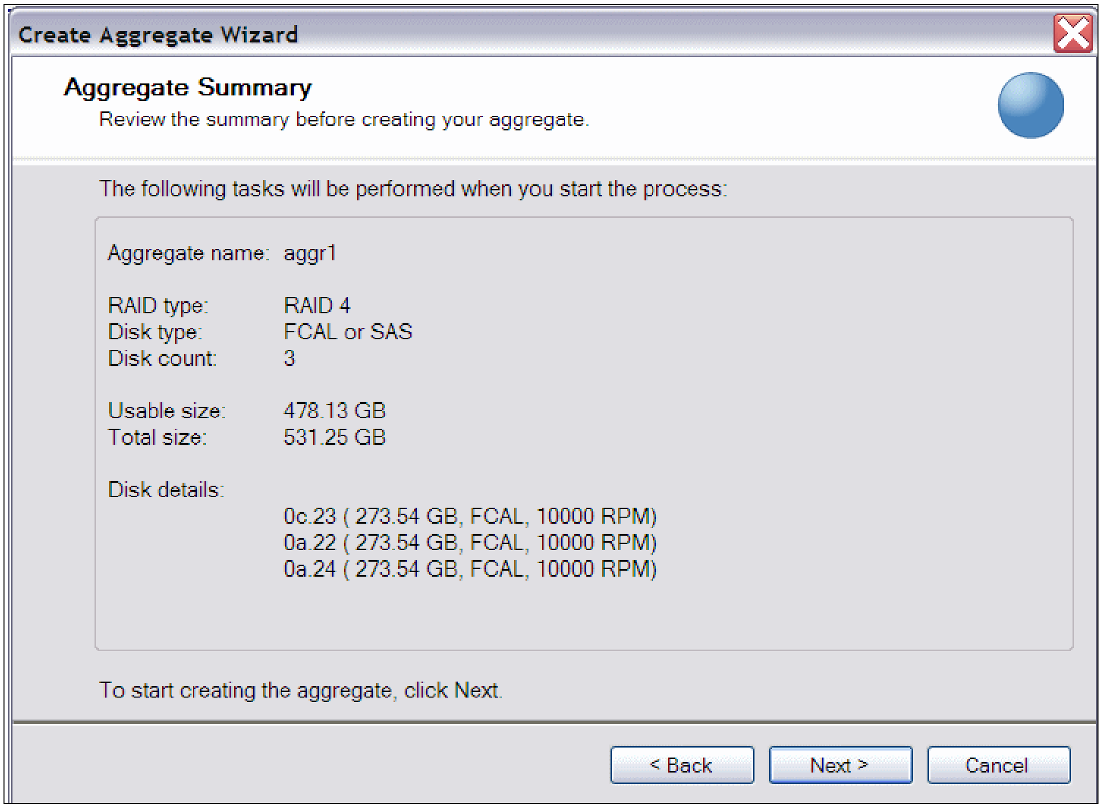


Figure 7-7 Aggregate Summary panel

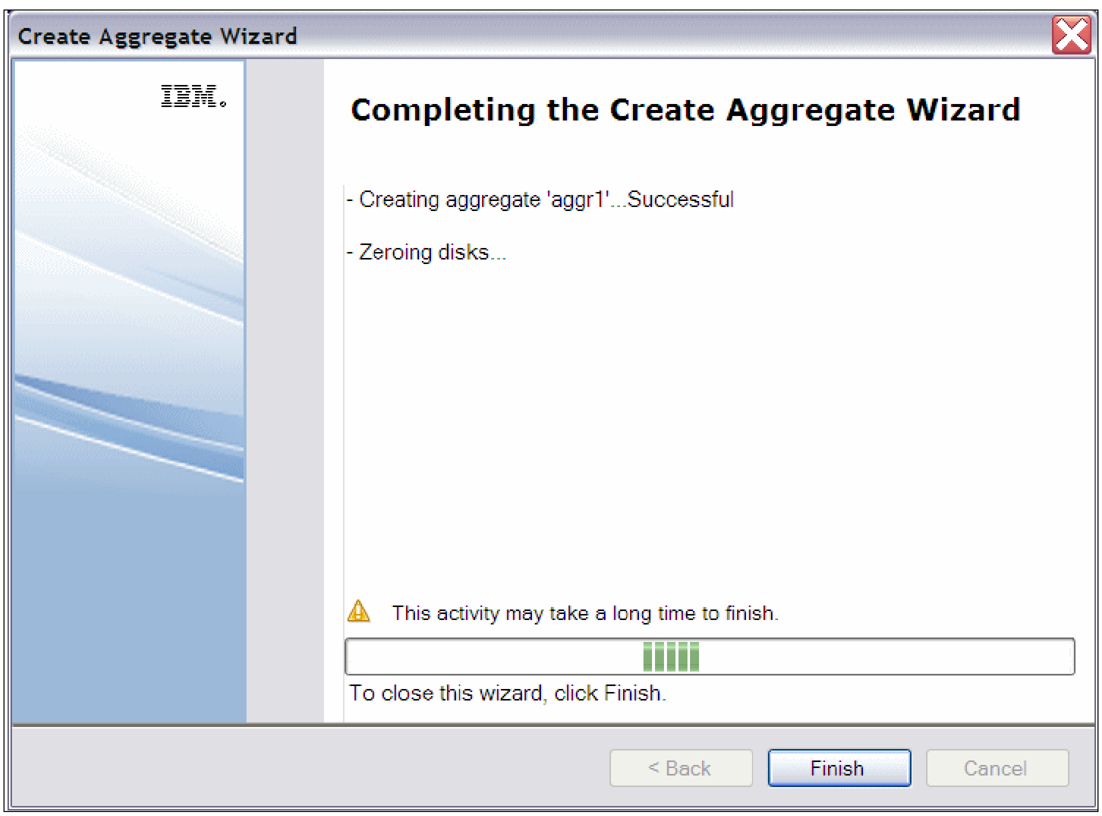


Figure 7-8 Aggregate creation

5. To verify that the System Manager has sent the commands to the filer use the FilerView. From a browser window, click **Aggregate** → **Manage**. There will be a new status reflecting the modifications (Figure 7-9). As you can see in this window, the new aggr1 is added to the current list even though it is not online but in initializing status.

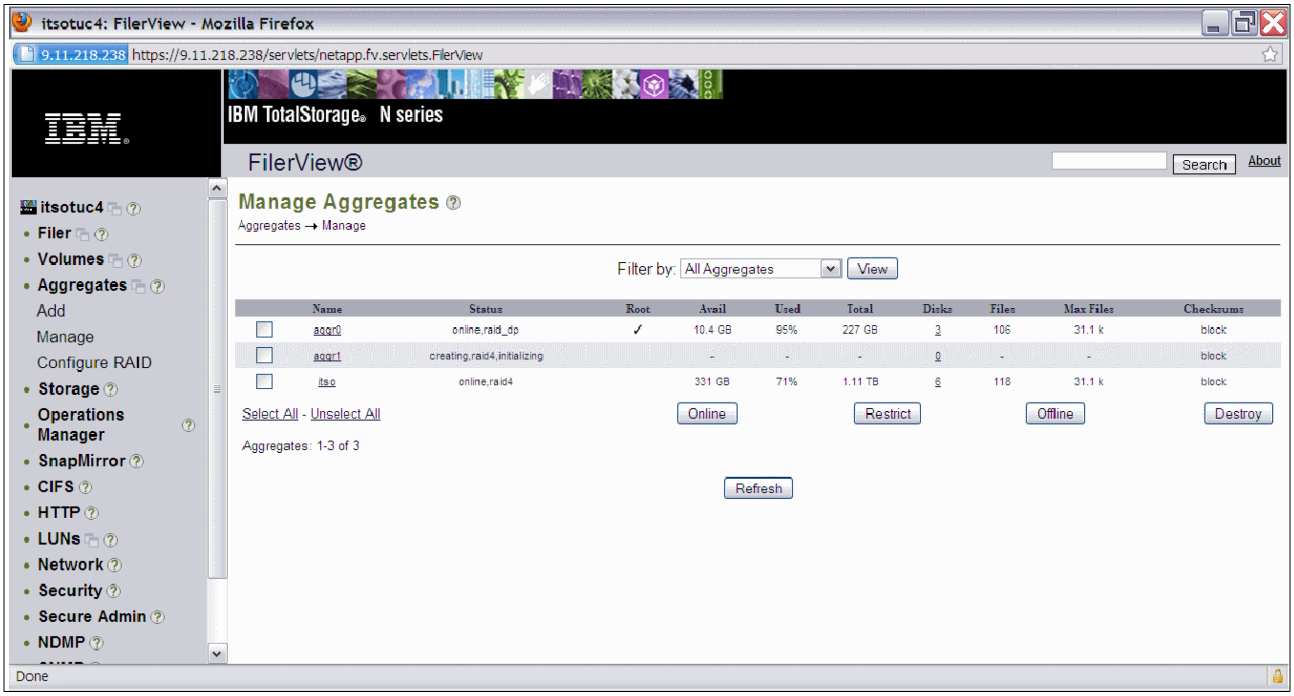


Figure 7-9 Aggregate verification

6. You can close both the wizard menu in the System Manager and the FilerView at this point. The task of creating the aggregate, which might take a long time, will continue in the background. Later you can open the System Manager and check the updated configuration; the new resource will eventually show as “online” in the aggregate panel (Figure 7-10). Click the **Details** button at the bottom of the screen to check the aggregate properties and the disks that the system has selected to define the aggregate. Click the **Space Breakout** button to view details about the resources belonging to the aggregate.

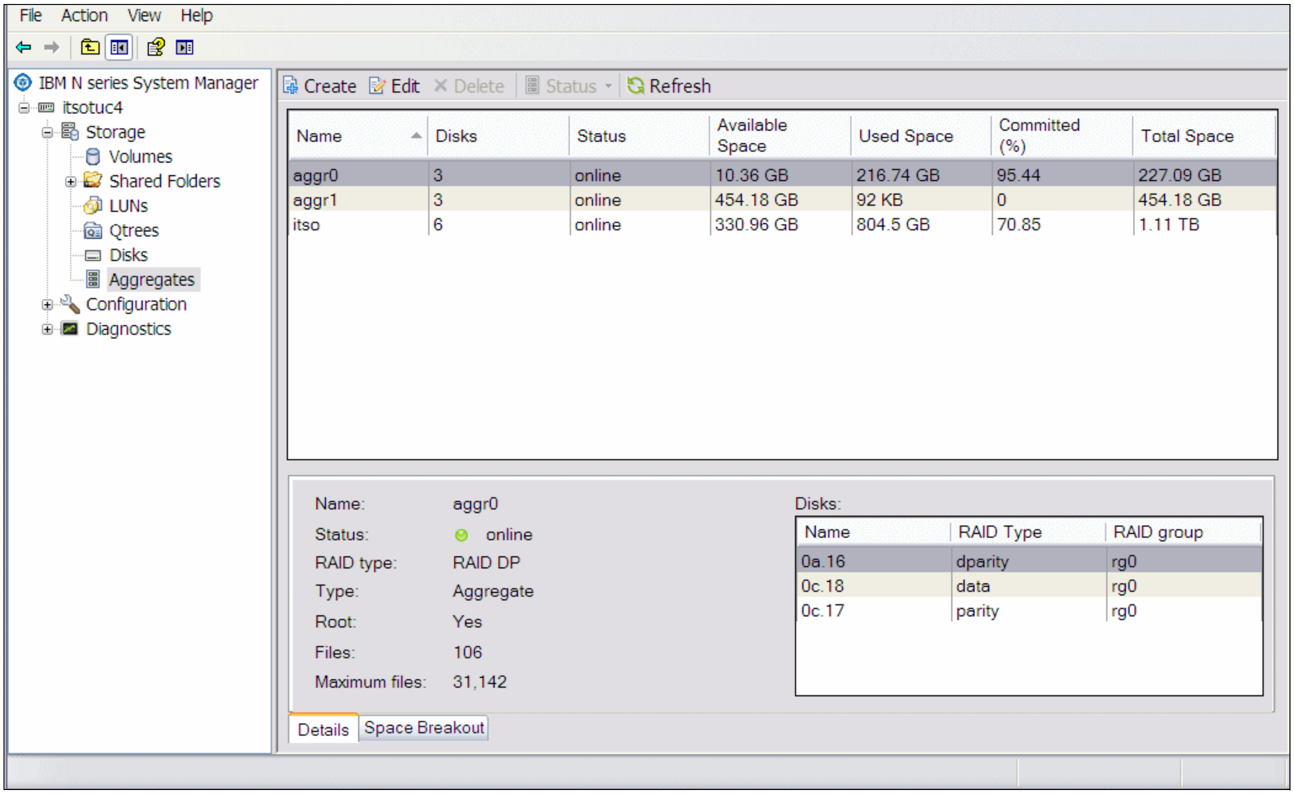


Figure 7-10 Aggregate properties

Creating a new volume in the aggregate

You can define and edit volumes and shares, again using the System Manager to simplify the process. Use the following steps to create a new volume:

1. Navigate to the volumes panel. At the top of the main pane click **Create**. This starts the wizard for creating a new volume in the aggregate.
2. Enter the name (we typed `ASIS_vol1`), select the storage type (`NAS`), the aggregate (`aggr1`), the size (`10 GB`), and the space guarantee settings (none, in our example) as shown in Figure 7-11 on page 210.
3. Click **Create** to complete the procedure.

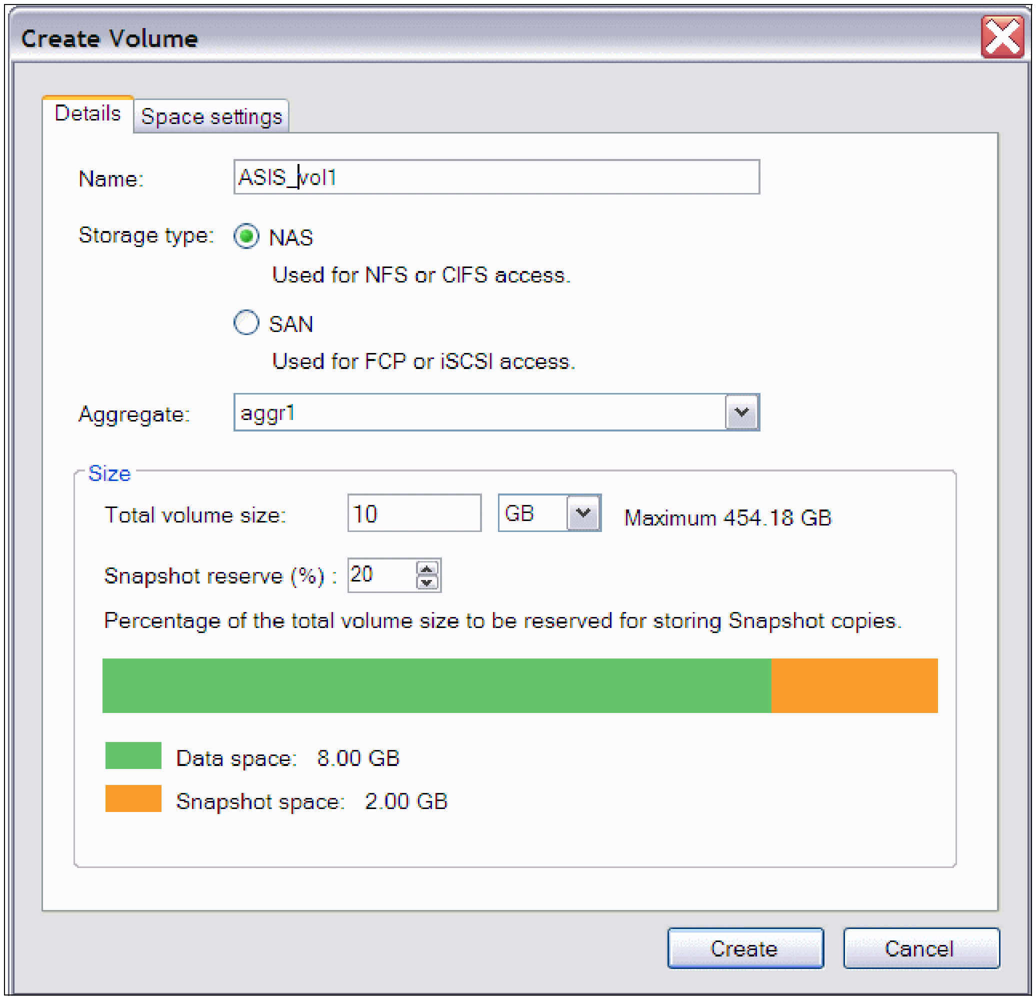


Figure 7-11 Volume creation

Defining a new share

The procedure to share folders is even easier using the System Manager, and consists of the following steps:

1. In the navigation pane on the left side of the System Manager screen, under the N series system (itsotuc4) click **Storage** → **Volumes** → **Shared Folders** → **Shares**.
2. Click **Create** to start the wizard and define the new share. Click **Next** from the initial wizard panel.
3. Select the folder to share and click **Next**, as shown in Figure 7-12 on page 211.

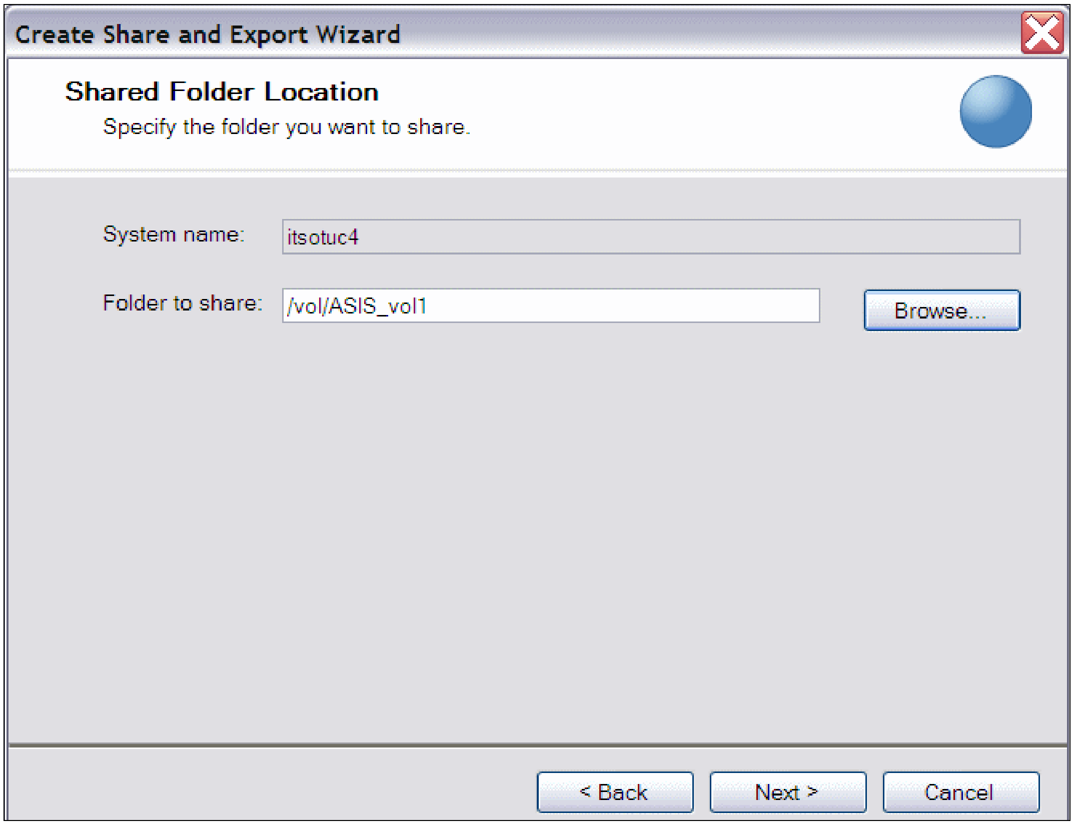


Figure 7-12 Creating shared folders

4. Select the CIFS share name and its description. Click **Next**.
5. Verify the settings in the Share Summary panel and click **Next**.
6. Click **Finish** to complete the procedure.

Display the Share list by clicking **Shares** in the navigation pane. The new dedup_test share will appear in the list, as shown in Figure 7-13.

You can also navigate to the Storage Manager application to check other resources (such as SW licences, protocols, and adapters) and perform basic management tasks on your N series system.

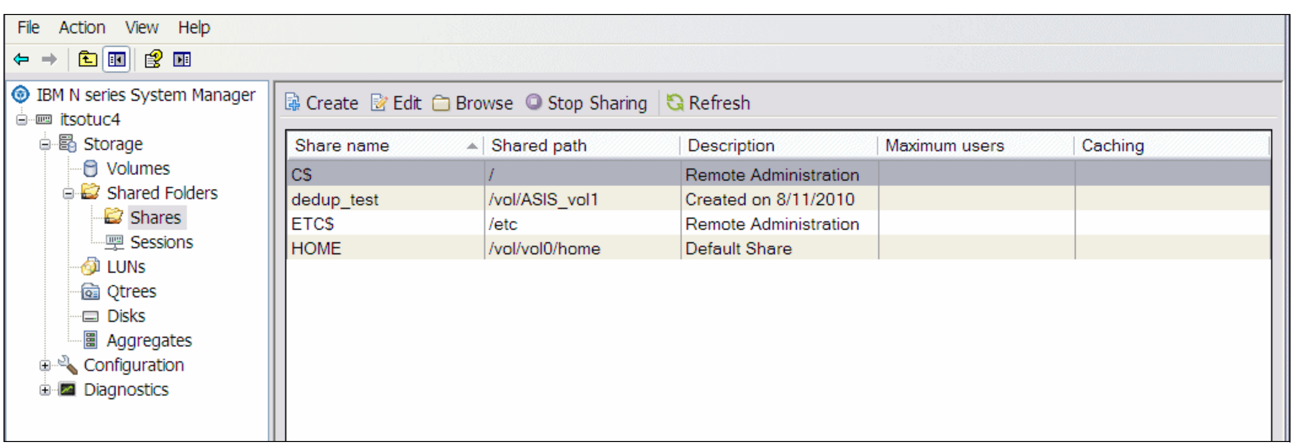


Figure 7-13 Displaying CIFS share

7.3.1 Initial setup

When a new storage system is discovered or added to the the list of managed systems it must be set up and configured for storage before it can be used. A “New system (requires setup)” message is displayed on the main systems page if the system has not been set up. The system name is displayed if the system has been set up.

You can use the System Management Setup wizard to set up your N series system (standalone or cluster).

You need to have the following information available for the storage system that you want to set up:

- Host name or IP address
- Administrator’s password
- Network address, gateway, and subnet
  If it is a clustered system you also need to know the partner IP address

To set up a storage system using the System Manager Setup wizard, follow these steps:

1. If you are not already in the main application window, click **System Manager** in the navigation pane.
2. Select the storage system that you want to set up.
3. Click **Set up** to start the Storage System Setup wizard. (If you have not set up the storage system, you can click the discovered IP address of the system and click the link for the Storage System Setup wizard.)
4. Type or select information as prompted by the wizard.
5. Click **Finish** to save the information.

After you discover or add to the list of managed systems a storage system that has not been configured, you must perform some initial configuration on the system to start provisioning the storage space. You can use the wizard to configure the system. You must separately configure each storage system node when you configure a clustered appliance.

To configure one or more storage systems using the Configuration wizard, follow these steps:

1. From the console tree, click the appropriate storage system to display a subtree of management features.
2. Click **Storage**.
3. Click **Storage Configuration wizard**. Type or select information as prompted by the wizard.
4. Click **Finish** to close the wizard.

After you set up a storage system using the wizard, the main systems page is displayed if you set up a system from that page. If you set up a system from the System Setup Dashboard window, the system dashboard is displayed.

7.3.2 System Manager and advanced features

System Manager includes support for many advanced features to simplify your storage system administration tasks. Complete details about supported features is beyond the scope of this book. In this section we demonstrate the use of one of the advanced features. The example is related to the FlexClone® capability to create and manage active copies of a source volume and then decide to terminate the relationship between parent and child resources.

We assume that the FlexClone licence is already active in your system. The following steps describe how we used System Manger to create a clone volume, make changes, and test the copy:

1. Define a clone from the ASIS_vol1 resource already defined. Go to **Storage** → **Volumes,** select the parent volume and click **Snapshot**.
2. Choose the name of the clone volume and decide which copy will be the base snapshot for the relationship. We chose to create a new snapshot (so the relationship is active and valid from that moment); the space guarantee is set to None because we are managing flexible volumes (Figure 7-14).

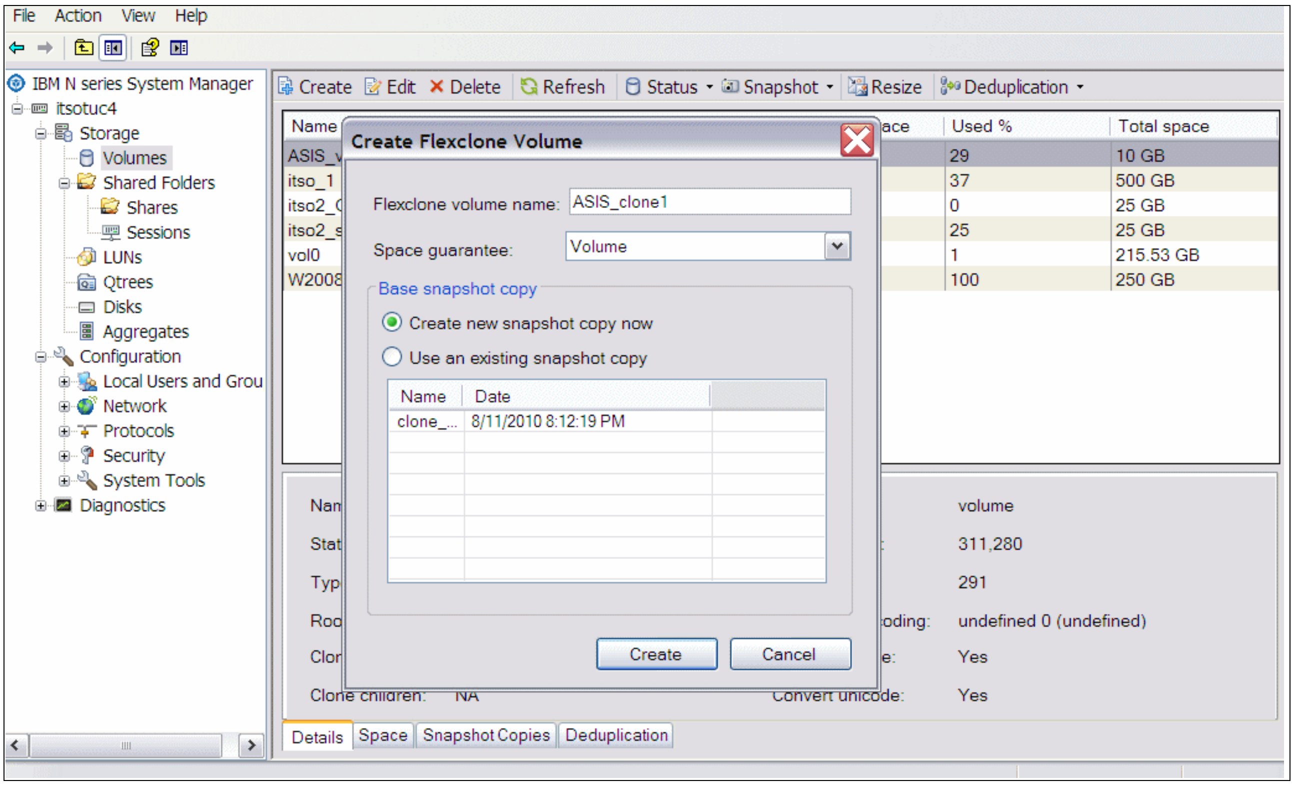


Figure 7-14 FlexClone creation

3. Click **Create**. The result is shown in Figure 7-15 on page 214, where a new volume is already available: the ASIS_clone1. This resource has the same characteristics as its parent, as shown in the figure.
4. Define a new share following the same procedure already described in order to prepare the test environment.

We plan to update the older volume with new data and check the differences between the source volume and its FlexClone volume. On the older volume we have already activated the deduplication feature (ASIS). After the initial ASIS run we decided to stop new deduplication sessions since new data was not added.

Now the situation has changed because we have decided to create the FlexClone and then modify the older volume adding new data.

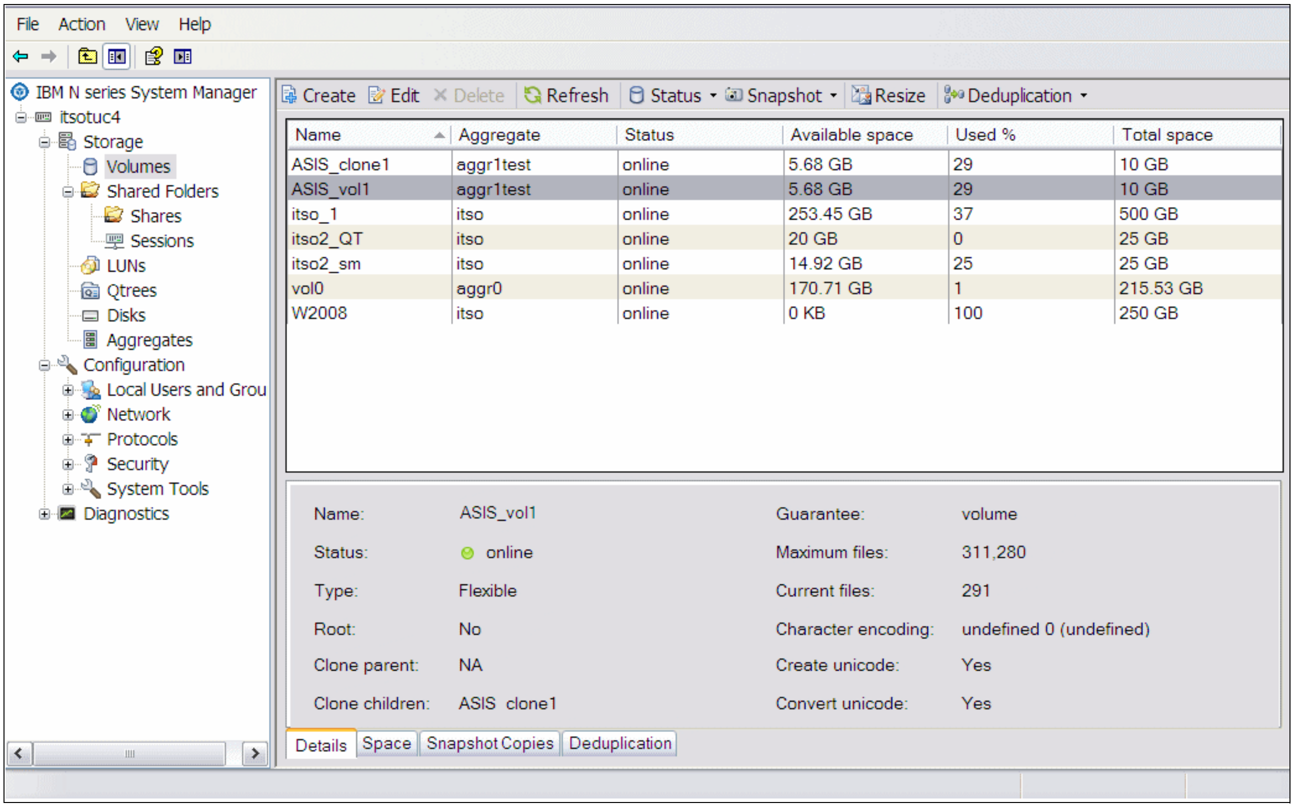


Figure 7-15 FlexClone display

5. Start a new ASIS session on the parent volume ASIS_vol1, checking that only the new data will be deduplicated, not the complete volume because this was already deduplicated and we don't need to scan the entire volume (Figure 7-16 on page 215).

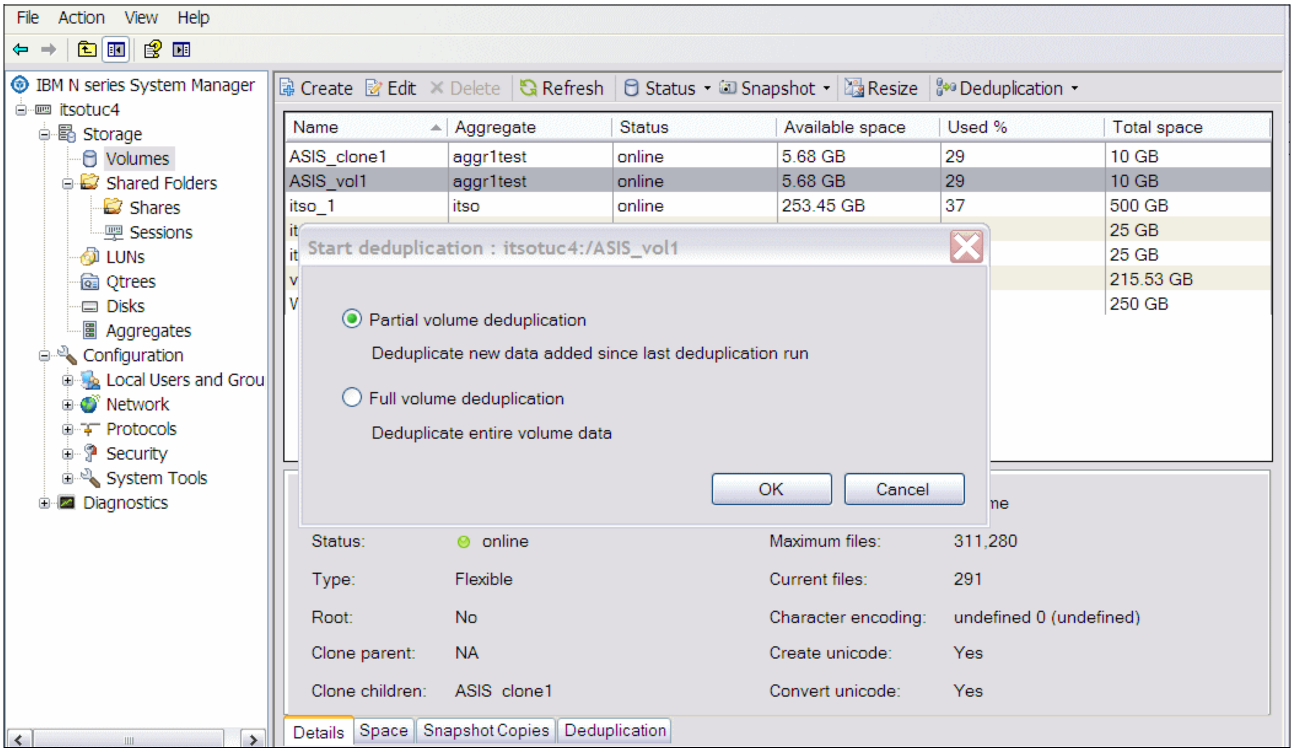


Figure 7-16 Starting deduplication

6. When ASIS has completed its run the volume has reduced its size. The differences are shown in Figure 7-17. The FlexClone volume ASIS_clone1 is not affected by these operations since it is still tied to the snapshot consistency point defined at its creation.

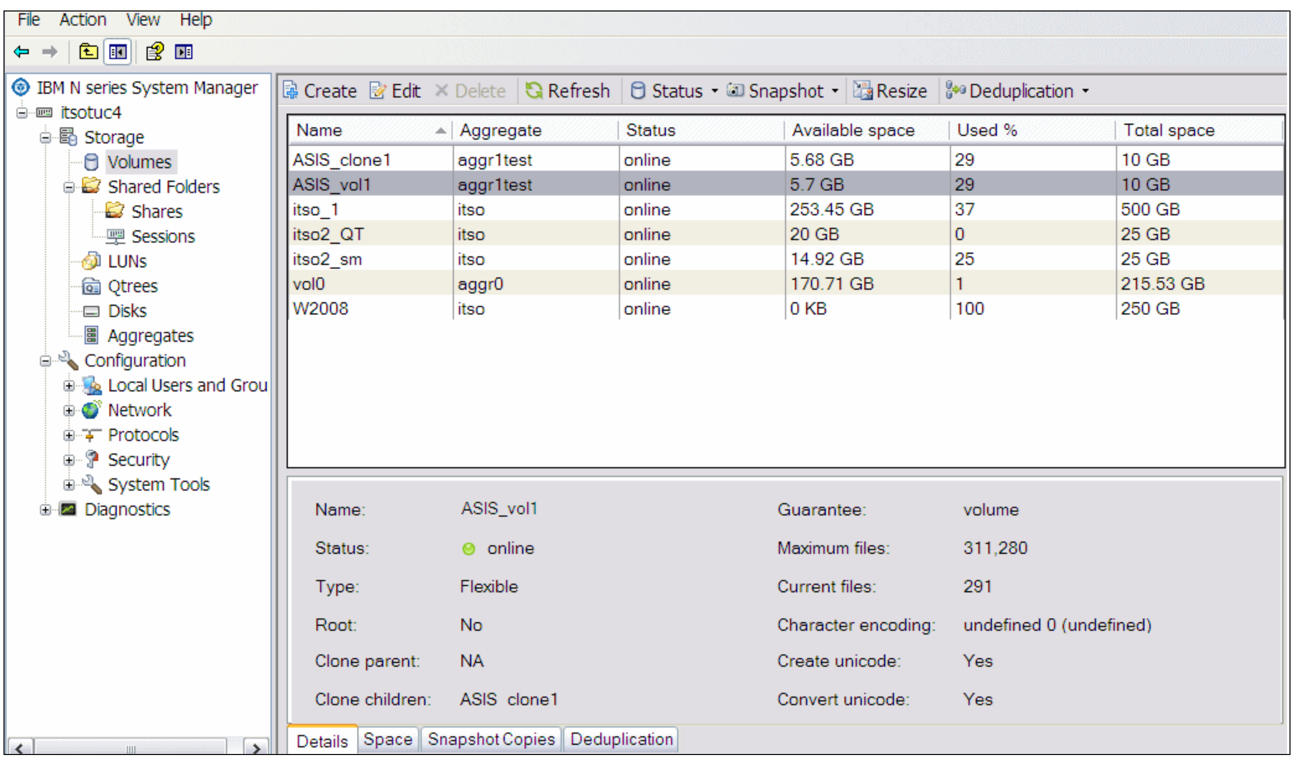


Figure 7-17 Deduplication complete

7. The ASIS_vol1 has several related active snapshots as defined previously. One of these snapshots is the one needed to define its clone. It is possible to modify the snapshot list and delete some of the available snapshots by simply clicking the selected snapshot on the list and clicking **Delete**, as shown in Figure 7-18. A dialog box will open and prompt you to confirm the snapshot deletion.
8. Continue to work with your source volume, modifying its content and adding new data. If no more consistency points are activated the flexClone volume will maintain the content and size at the time of its creation.

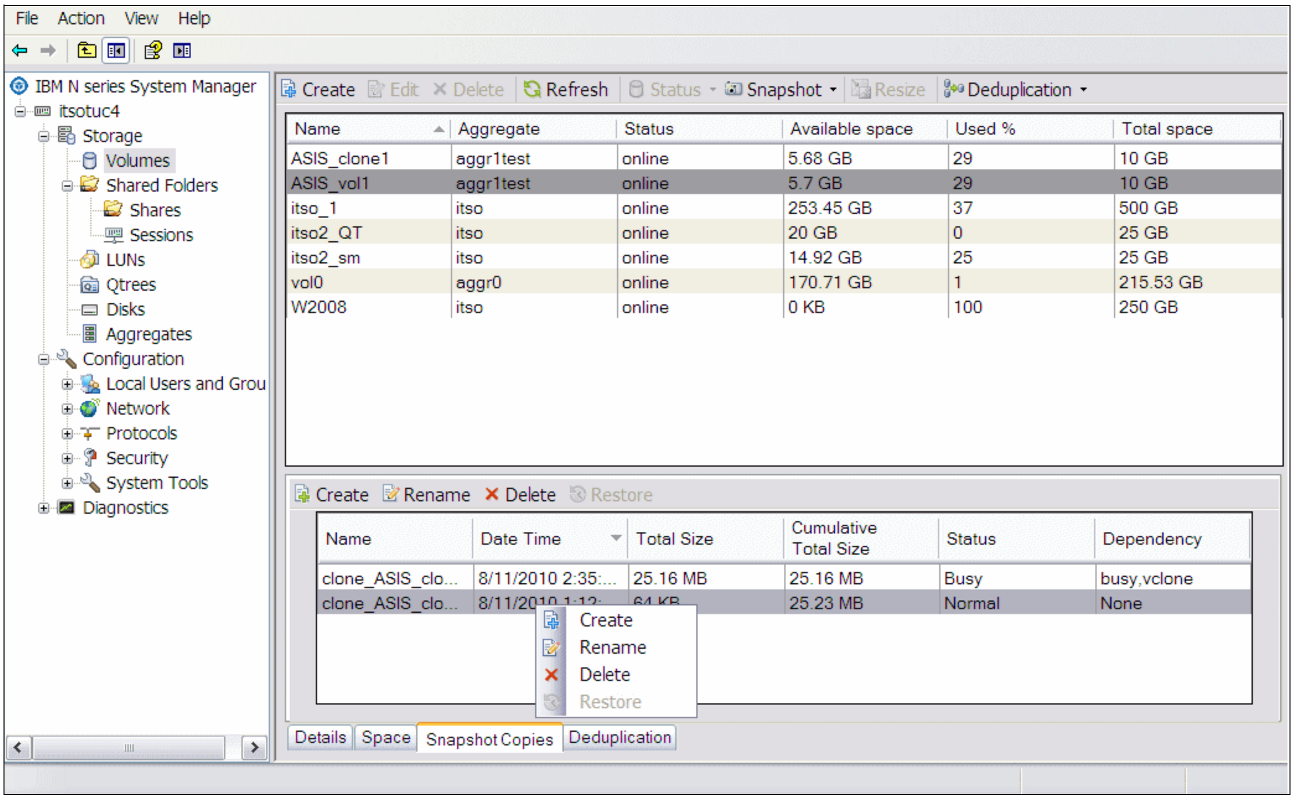


Figure 7-18 Snapshot list

9. We want to split the two volumes, so we decided to delete all snapshots (or at least that used to define the clone). Currently the FlexClone split action is not supported by System Manager, so we used the CLI and issued the relevant commands. Example 7-1 shows the commands for checking the status of the parent volume and defining the split. The status output indicates that ASIS_vol1 has a clone. After the **`start split`** command we waited for the split complete message and then issued the **`vol status`** command to verify that no relationships are active between ASIS_vol1 and ASIS_clone1.

Example 7-1 FlexClone split operation

```
Data ONTAP (itsotuc4.)
login:
Data ONTAP (itsotuc4.)
login:
Data ONTAP (itsotuc4.)
login: root
```

```
Password:
itsotuc4>
itsotuc4> vol status ASIS_vol1
       Volume State        Status          Options
       ASIS_vol1 online    raid4, flex    create_ucode=on,
                 sis       convert_ucode=on, guarantee=none,
                Volume has clones: ASIS_clone1
              Containing aggregate: 'aggr1test'
itsotuc4>
itsotuc4> vol clone split start ASIS_clone1
Fri Mar 5 14:50:57 MST [wafl.volume.clone.split.started:info]: Clone split was
started for volume ASIS_clone1
Clone volume 'ASIS_clone1' will be split from its parent.
Monitor system log or use 'vol clone split status' for progress.
itsotuc4> Fri Mar  5 14:50:57 MST [wafl.scan.sta
split on volume ASIS_clone1.
vol clone split staFri Mar 5 14:52:16 MST [wafl.scan.clone.split.complete:info]
: Clone split complete for Volume ASIS_clone1
tus ASIS_clone1
vol clone split status: The volume is not a clone
itsotuc4> vol status
         Volume State          Status            Options
         itso_1 online          raid4, flex       snapmirrored=on,
                                snapmirrored      maxdirsize=41861,
                                read-only         fs_size_fixed=on
           vol0 online          raid_dp, flex     root
       itso2_QT online          raid4, flex
  sangam_mirror online          raid4, flex       snapmirrored=on,
                                snapmirrored      maxdirsize=41861,
                                read-only         fs_size_fixed=on
         sangam online          raid4, flex       snapmirrored=on,
                                snapmirrored      maxdirsize=41861,
                                read-only         fs_size_fixed=on
    ASIS_clone1 online          raid4, flex       create_ucode=on,
                                sis               convert_ucode=on, guarantee=none,
      ASIS_vol1 online          raid4, flex       create_ucode=on,
                                sis               convert_ucode=on, guarantee=none,
       itso2_sm online          raid4, flex       snapmirrored=on,
                                snapmirrored      maxdirsize=41861,
                                read-only         fs_size_fixed=on
itsotuc4>
itsotuc4>
```

At this point you can modify both volumes, adding new data and starting deduplication if needed. No dependencies are active between the clone and its parent volume.

8

IBM System Storage N series administration

This chapter describes how to accomplish basic administration tasks on IBM System Storage N series storage systems.

This chapter covers the following topics:

- Administration methods
- Starting, stopping, and rebooting the IBM System Storage N series storage system
- Checking the Data ONTAP software version
- Storage management

8.1 Administration methods

The following methods can be used to administer an N series storage system:

- FilerView interface
- Operations Manager software
- Command-line interface (CLI)
- Mounting the root volume/mapping administration share

Details are provided in this chapter for using the FilerView web interface and the CLI to perform basic adminitration tasks.

Web interface

To access the N series through FilerView, open your browser and point to the following URL:

```
http://<filername or ip-address>/na_admin
```

To proceed, specify a valid user name and password.

Command-line interface

The CLI can be accessed through Telnet or a secure shell interface (SSH). Use the **help** command or simply enter a question mark (**?**) to obtain an overview of available commands.

Enter **help <*command*>** for a brief description of what the command does.

Enter **<*command*> help** for a list of the available options of the specified command, as shown in Figure 8-1.

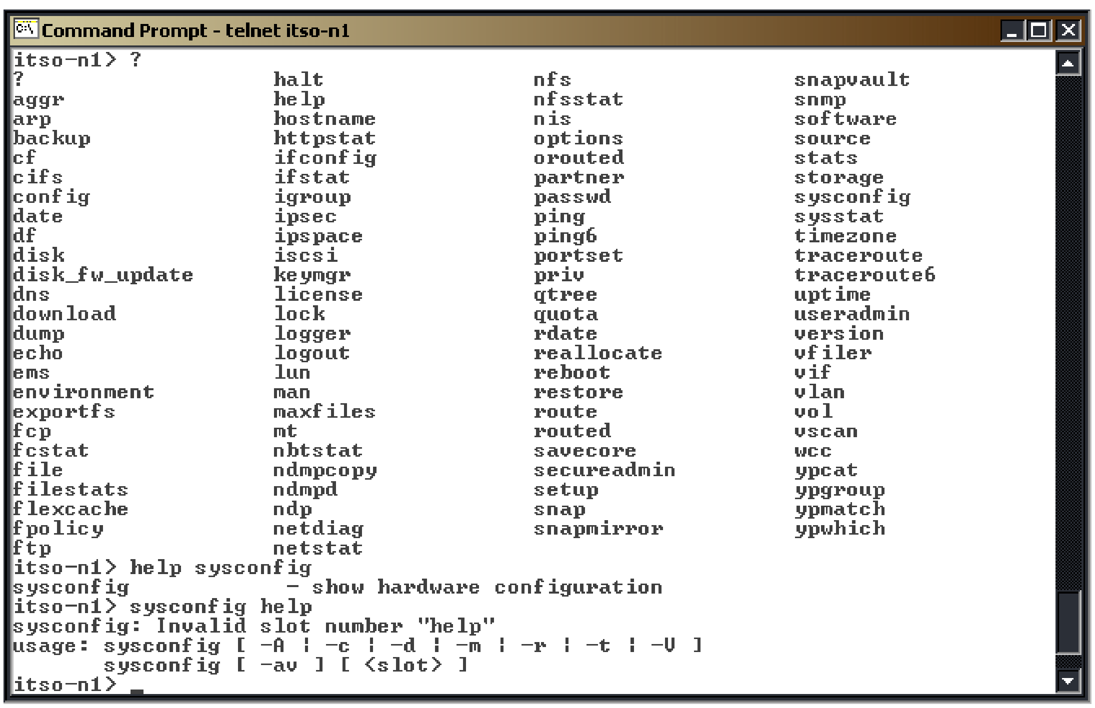
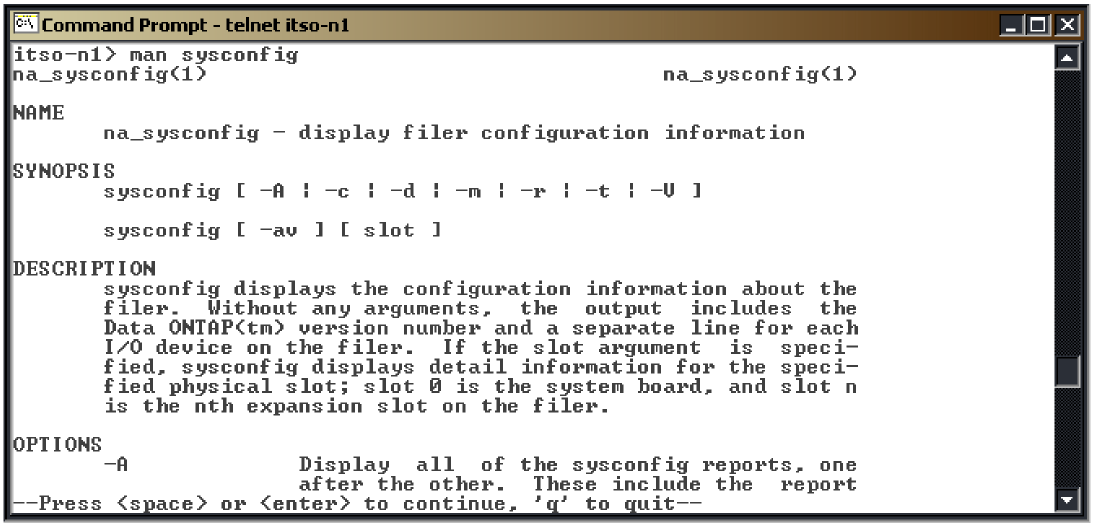


Figure 8-1 The help and ? commands

The manual pages can be accessed by entering the **man** command. Figure 8-2 provides a detailed description of a command and lists options (**man *<command>***).




```
itso-n1> man sysconfig
na_sysconfig(1)                                          na_sysconfig(1)

NAME
       na_sysconfig - display filer configuration information

SYNOPSIS
       sysconfig [ -A | -c | -d | -m | -r | -t | -V ]

       sysconfig [ -av ] [ slot ]

DESCRIPTION
       sysconfig displays the configuration information about the
       filer.  Without any arguments, the  output  includes  the
       Data ONTAP(tm) version number and a separate line for each
       I/O device on the filer.  If the slot argument  is  speci-
       fied, sysconfig displays detail information for the speci-
       fied physical slot; slot 0 is the system board, and slot n
       is the nth expansion slot on the filer.

OPTIONS
       -A              Display  all  of the sysconfig reports, one
                       after the other.  These include the  report
--Press <space> or <enter> to continue, 'q' to quit--
```

Figure 8-2 Results of a man command

8.2 Starting, stopping, and rebooting the storage system

This section describes boot, shutdown, and halt procedures.

Important: Reboot or halt should always be planned procedures. Users should be informed about these tasks in advance and give users enough time to save their changes to avoid loss of data.

8.2.1 Starting the IBM System Storage N series storage system

The IBM System Storage N series boot code resides on a Compact Flash card. After turning on the system, IBM System Storage N series will boot automatically from this card. You can enter an alternative boot mode by pressing Ctrl+C and choosing the **boot** option.

Attention: Power on the IBM System Storage N series storage system in the following order:

1. Expansion disk shelves
2. IBM System Storage N series (base unit)

Example 8-1 shows the boot procedure; Example 8-2 shows the boot options. The typical case is to boot in normal boot mode.

Example 8-1 Boot screen: press Ctrl+c for special boot menu

```
CFE version 1.2.0 based on Broadcom CFE: 1.0.35
Copyright (C) 2000,2001,2002,2003 Broadcom Corporation.
Portions Copyright (C) 2002,2003 Network Appliance Corporation.

CPU type 0x1040102: 650MHz
Total memory: 0x40000000 bytes (1024MB)

Starting AUTOBOOT press any key to abort...
Loading: 0xffffffff80001000/21792 0xffffffff80006520/10431377 Entry at 0xffffffff80001000
Starting program at 0xffffffff80001000
Press CTRL-C for special boot menu
```

Example 8-2 Boot menu

```
1) Normal Boot
2) Boot without /etc/rc
3) Change Password
4) Initialize all disks
4a) Same as option 4 but create a flexible root volume
5) Maintenance boot
Selection (1-5)?
```

8.2.2 Stopping the IBM System Storage N series storage system

Stopping and rebooting the IBM System Storage N series storage system prevents all users from accessing the N series. Before stopping or rebooting the system, ensure that maintenance is possible and that all users (file access, database user, and others) have been informed about the upcoming action.

Important: For a graceful shutdown of IBM System Storage N series storage systems, use the **halt** command or FilerView. This avoids unpredictable problems. Remember to shut down both nodes if an IBM System Storage N series A2x model has to be shut down.

Common Internet File System (CIFS) services

The **cifs sessions** command reports open sessions to the IBM System Storage N series storage system (Example 8-3).

Example 8-3 List open CIFS sessions

```
itsosj-n1> cifs sessions
Server Registers as 'ITSO-N1' in workgroup 'WORKGROUP'
Root volume language is not set. Use vol lang.
WINS Server: 9.1.38.12
Using Local Users authentication
====================================================
PC IP(PC Name) (user)           #shares   #files
9.1.57.45() (ITSO-N1\administrator - root) (using security signatures)
                                1         0
9.1.39.107() (ITSO-N1\administrator - root) (using security signatures)
                                3         0
itsosj-n1>
```

With the IBM System Storage N series storage systems, you can specify which users receive CIFS shutdown messages. By issuing the **cifs terminate** command, Data ONTAP, by default, sends a message to all open client connections. This setting can be changed by issuing the following command:

```
options cifs.shutdown_msg_level 0 | 1 | 2
```

The options are:

0: Never send CIFS shutdown messages.
1: Send CIFS messages to clients connected and with open files only.
2: Send CIFS messages to all open connections (default).

The **cifs terminate** command shuts down CIFS, ends CIFS service for a volume, or logs off a single station. The **-t** option can be used to specify a delay interval in minutes before cifs stops (Example 8-4).

Example 8-4 The cifs terminate -t command

```
itsosj-n1> cifs terminate -t 3
Total number of connected CIFS users: 1
     Total number of open CIFS files: 0
Warning: Terminating CIFS service while files are open may cause data loss!!
3 minutes left until termination (^C to abort)...
2 minutes left until termination (^C to abort)...
1 minute left until termination (^C to abort)...

CIFS local server is shutting down...

CIFS local server has shut down...
itsosj-n1>
```

You can even choose single workstations for which the cifs service should stop (Example 8-5).

Example 8-5 The cifs terminate command for a single workstation

```
itsosj-n1> cifs terminate -t 3 workstation_01
3 minutes left until termination (^C to abort)...
2 minutes left until termination (^C to abort)...
1 minute left until termination (^C to abort)...
itsosj-n1> Thu Sep  8 09:41:43 PDT [itsosj-n1: cifs.terminationNotice:warning]: CIFS: shut
down completed: disconnected workstation workstation_01.

itsosj-n1>
```

When you shut down an N series there is no need to specify the **cifs terminate** command because it will be run by the operating system automatically.

Note: Workstations running Windows 95/98 or Windows for Workgroups will not see the notification unless they are running WinPopup.

Depending on the CIFS message settings, pop-ups or similar messages should appear on the affected workstations (Figure 8-3).

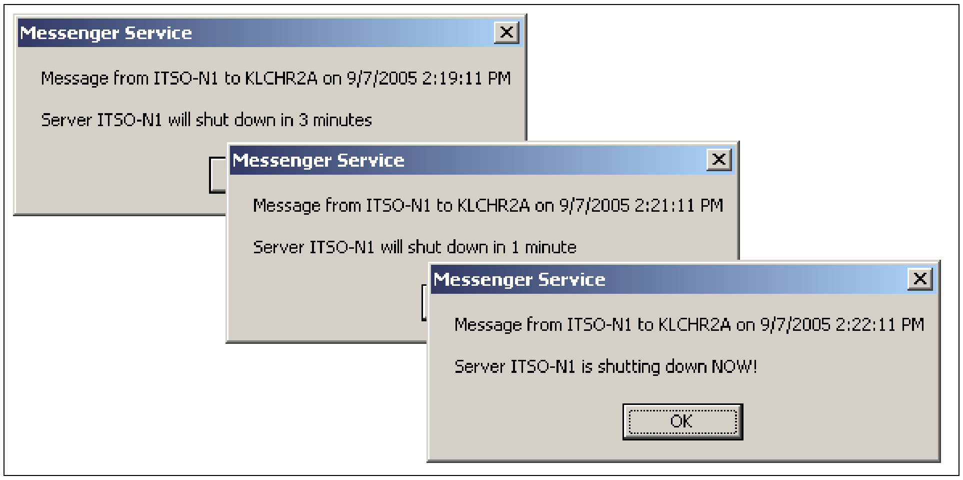


Figure 8-3 Shutdown messages on CIFS clients

To restart CIFS, issue the **`cifs restart`** command, as shown in Example 8-6. (The N series startup procedure starts the CIFS services automatically.)

Example 8-6 The cifs restart command

```
itsosj-n1> cifs restart
CIFS local server is running.
itsosj-n1>
```

You can verify whether CIFS is running by using the **`cifs sessions`** command. If CIFS is not running a message appears, as shown in Example 8-7.

Example 8-7 Checking whether CIFS is running on the N series

```
itsosj-n1> cifs sessions
CIFS not running.  Use "cifs restart" to restart
                   Use "cifs prefdc" to set preferred DCs
                   Use "cifs testdc" to test WINS and DCs
                   Use "cifs setup" to configure
itsosj-n1>
```

Halting the N series

You can use the command line or FilerView interface to stop the N series. You can use the **`halt`** command on the CLI to perform a graceful shutdown. The -t option causes the system to stop after the number of minutes that you specify (for example, **`halt -t 5`**). The **`halt`** command stops all services and shuts down the system gracefully to the Common Firmware Environment (CFE) prompt.

File system changes will be written to disk and the non-volatile Random Access Memory (NVRAM) content will be vacated.

We used the serial console because the IP connection will be lost after halting the N series (Example 8-8).

Example 8-8 Halt with command-line interface (serial console)

```
CFE version 1.2.0 based on Broadcom CFE: 1.0.35
Copyright (C) 2000,2001,2002,2003 Broadcom Corporation.
Portions Copyright (C) 2002,2003 Network Appliance Corporation.

CPU type 0x1040102: 650MHz
Total memory: 0x40000000 bytes (1024MB)
CFE>
```

Booting the N series

As described in 8.2.1, “Starting the IBM System Storage N series storage system” on page 221, the IBM System Storage N series storage systems automatically boot Data ONTAP from a PC Compact Flash card, which ships with the most current Data ONTAP release. The Compact Flash card contains sufficient space for an upgrade kernel. Use the **download** command to copy a boot kernel to the Compact Flash card.

The Common Firmware Environment (CFE) prompt provides several boot options:

- **boot_ontap**

  This boots the current version of Data ONTAP from the Compact Flash card.

- **boot_primary**

  This boots the current version of Data ONTAP from the Compact Flash card as the primary kernel (the same kernel as boot_ontap).

- **boot_backup**

  This boots the backup version of Data ONTAP from the Compact Flash card. The backup release is created during the first software upgrade to preserve the kernel that shipped with the system. It provides a known good release from which you can boot the system if it fails to automatically boot the primary image.

- **netboot**

  This boots from a Data ONTAP version stored on a remote HTTP or TFTP server. The netboot option enables you to boot an alternative kernel if the Compact Flash card becomes damaged. Upgrade the boot kernel for several devices from a single server.

To enable netboot, you must configure networking for the IBM System Storage N series storage system (using DHCP or static IP address) and place the boot image on a configured server.

> **Tip:** We recommend that you store a boot image on an http or TFTP server to protect against Compact Flash card corruption.

For more information about setting up netboot, refer to the following website:

http://www.ibm.com/storage/support/nas/

Usually you boot the N series after you issue the **halt** command with the **boot_ontap** or **bye** command. These commands end the CFE prompt and restart the N series, as shown in Example 8-9 on page 226.

Example 8-9 Start up the N series at the CFE prompt

```
CFE>bye
CFE version 1.2.0 based on Broadcom CFE: 1.0.35
Copyright (C) 2000,2001,2002,2003 Broadcom Corporation.
Portions Copyright (C) 2002,2003 Network Appliance Corporation.

CPU type 0x1040102: 650MHz
Total memory: 0x40000000 bytes (1024MB)

Starting AUTOBOOT press any key to abort...
Loading: 0xffffffff80001000/21792 0xffffffff80006520/10431377 Entry at 0xffffffff80001000
Starting program at 0xffffffff80001000
Press CTRL-C for special boot menu
..................................................................................
..................................................................................
.........................................Interconnect based upon M-VIA ERing Support
    Copyright (c) 1998-2001 Berkeley Lab
    http://www.nersc.gov/research/FTG/via
Wed Aug 31 19:00:46 GMT [cf.nm.nicTransitionUp:info]: Interconnect link 0 is UP
Wed Aug 31 19:00:46 GMT [cf.nm.nicTransitionDown:warning]: Interconnect link 0 is DOWN
Data ONTAP Release 7.1H1: Mon Aug 15 16:02:45 PDT 2005 (IBM)Copyright (c) 1992-2005 Network
Appliance, Inc.
Starting boot on Wed Aug 31 19:00:45 GMT 2005
Wed Aug 31 19:00:51 GMT [diskown.isEnabled:info]: software ownership has been enabled for
this system
Wed Aug 31 19:00:56 GMT [raid.cksum.replay.summary:info]: Replayed 0 checksum blocks.
Wed Aug 31 19:00:56 GMT [raid.stripe.replay.summary:info]: Replayed 0 stripes.
Wed Aug 31 19:00:57 GMT [localhost: cf.fm.launch:info]: Launching cluster monitor
Wed Aug 31 19:00:57 GMT [localhost: cf.fm.notkoverClusterDisable:warning]: Cluster monitor:
cluster takeover disabled (restart)
add net 127.0.0.0: gateway 127.0.0.1
DBG: Failed to get partner serial number from VTIC
DBG: Set filer.serialnum to: 310070722
Wed Aug 31 19:00:58 GMT [rc:notice]: The system was down for 71 seconds
Wed Aug 31 12:01:00 PDT [itsosj-n1: dfu.firmwareUpToDate:info]: Firmware is up-to-date on
all disk drives
Wed Aug 31 12:01:00 PDT [ltm_services:info]: Ethernet e0a: Link up
add net default: gateway 192.186.101.57: network unreachable
Wed Aug 31 12:01:02 PDT [rc:ALERT]: timed: time daemon started
Wed Aug 31 12:01:03 PDT [itsosj-n1: mgr.boot.disk_done:info]: Data ONTAP Release 7.1H1 boot
complete. Last disk update written at Wed Aug 31 11:59:46 PDT 2005
Wed Aug 31 12:01:03 PDT [itsosj-n1: mgr.boot.reason_ok:notice]: System rebooted.

Password:

itsosj-n1> Wed Aug 31 12:01:20 PDT [console_login_mgr:info]: root logged in from console
itsosj-n1>
```

Alternatively, you can use FilerView to shut down the N series (Figure 8-4). Choose **Filer → Shut Down and Reboot** and specify any appropriate options.

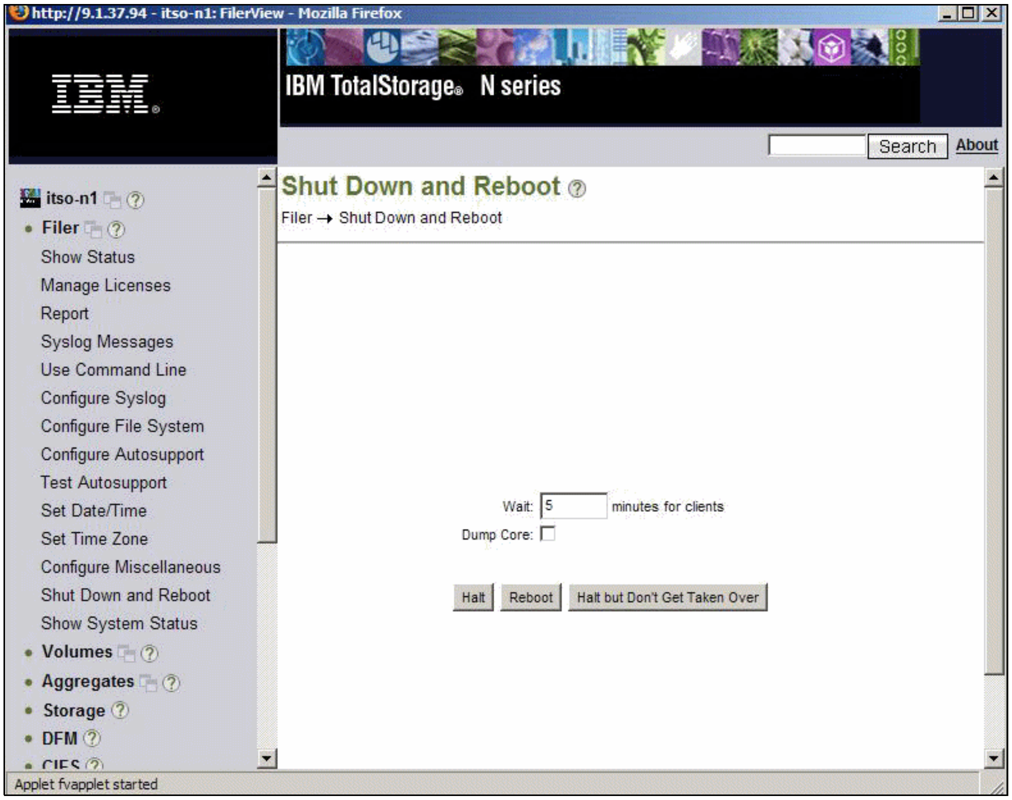


Figure 8-4 Halt with FilerView GUI

Additional confirmation and alert pop-up windows appear, as shown in Figure 8-5.

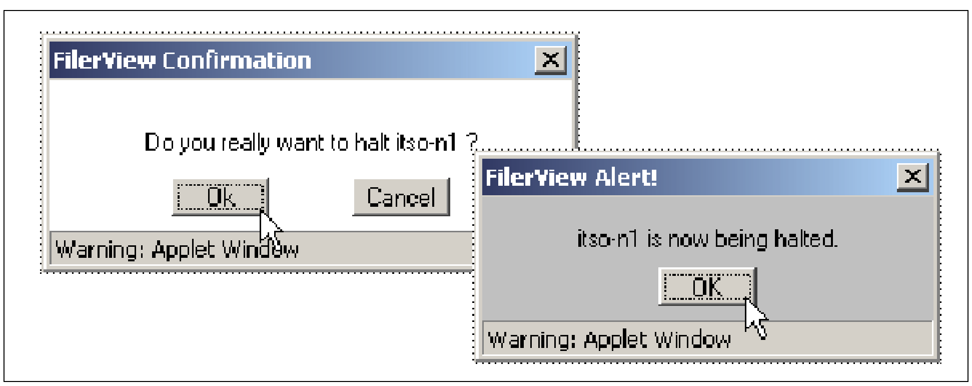


Figure 8-5 Confirmation on halting the N series with FilerView

Depending on the CIFS Message settings and Microsoft Windows Client settings, you may receive several messages on your CIFS client concerning the shutdown of the N series as shown previously in Figure 8-3.

8.2.3 Rebooting the system

The System Storage N series systems can be rebooted from the command line or from the FilerView interface.

Rebooting from the CLI halts the N series and then restarts it (Example 8-10).

Example 8-10 Rebooting from the command-line interface

```
[root@itso3775 node1]# reboot

Broadcast message from root (pts/2) (Thu Sep  8 13:23:47 2005):

The system is going down for reboot NOW!
```

Network File System (NFS) clients can maintain use of a file over a halt or reboot because NFS is a *stateless* protocol. CIFS, FCP, and iSCSI clients behave differently, and you can use the `-t` option to specify the time before shutdown.

If you choose FilerView to reboot the system, then proceed as shown in Figure 8-6.

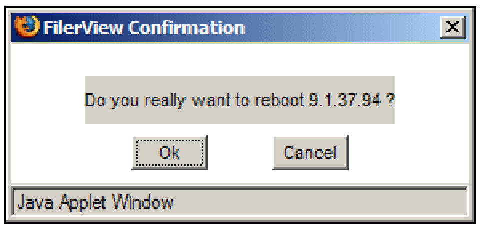


Figure 8-6 Reboot with FilerView

A confirmation and alert window pops up. Click **OK** to proceed if you definitely want to shut down the N series. See Figure 8-7.

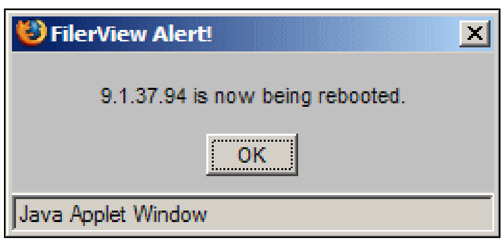


Figure 8-7 FilerView reboot confirmation and alert messages

Depending on the shutdown message settings, CIFS clients will get messages.

8.3 Checking the Data ONTAP software version

The Data ONTAP software level can be listed by using the **`version`** command from the command line, as shown in Example 8-11.

Example 8-11 Check Data ONTAP version using the command line

```
itsosj-n1> version
Data ONTAP Release 7.1X17: Mon Aug  8 02:50:45 PDT 2005 (IBM)
itsosj-n1>
```

Alternatively, you can use FilerView, as shown in Figure 8-8.

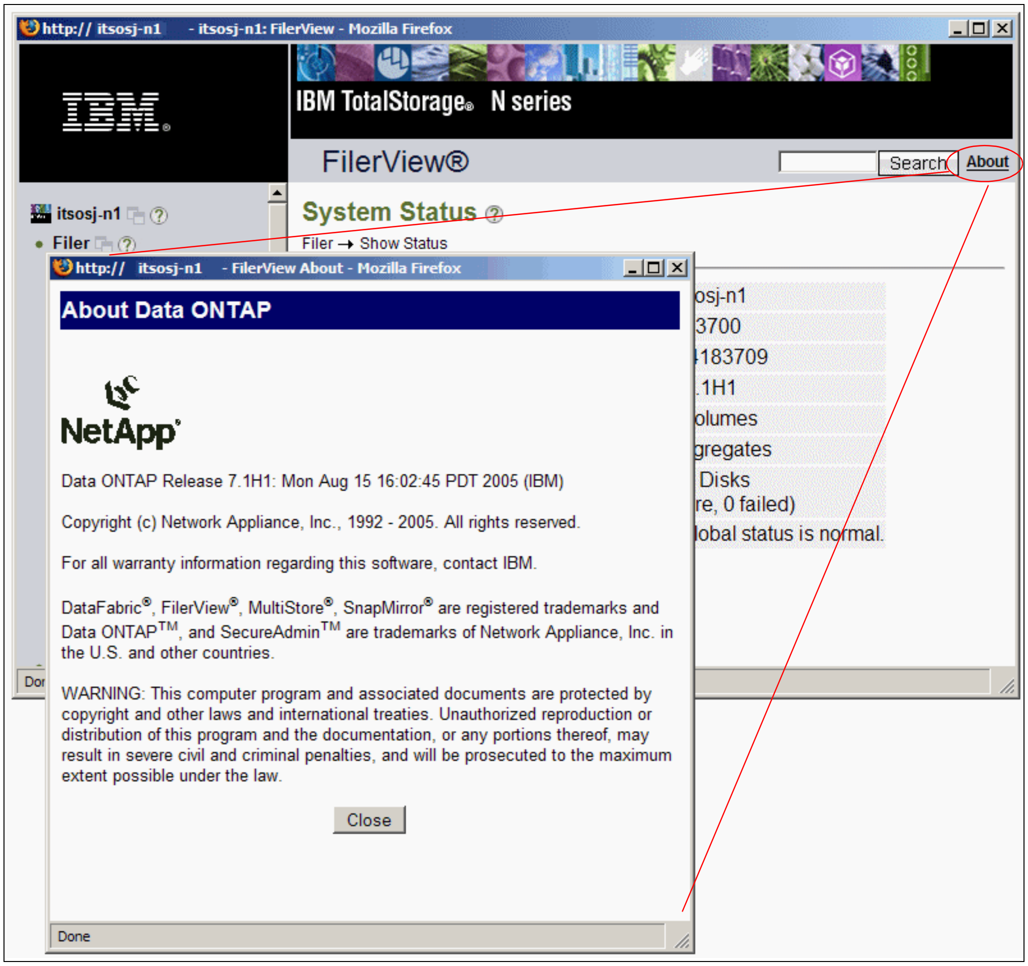


Figure 8-8 The FilerView About Data ONTAP window

8.3.1 Data ONTAP version 8.0-7 mode

Before installing or upgrading to ONTAP 8.0-7 mode check the information provided in Chapter 5, “Multipathing with the IBM System Storage N series” on page 145 and in Appendix A, “Getting started” on page 287.

8.4 Storage management

Storage management of the IBM System Storage N series storage system represents a large part of the administrator’s responsibilities. This section discusses some of the more common storage management tasks.

8.4.1 Locating a failed disk

The administrator can locate failed disks in an aggregate by issuing the **`aggr status -f`** or **`vol status -f`** commands on the command line. Example 8-12 on page 230 displays a failed disk that was located using the **`aggr status - f`** command.

Note: In a clustered environment, check both nodes for failed disks because disks assigned to the other node appear as *partner* disks. Failed disks from the other node may not show up as failed on both nodes.

Example 8-12 Finding failed disks

```
itsosj-n1> aggr status -f

Broken disks

RAID Disk       Device  HA  SHELF BAY CHAN Pool Type  RPM  Used (MB/blks)    Phys (MB/blks)
---------       ------  ------------- ---- ---- ---- ----- --------------    --------------
admin failed    v0.19   v0    1   3   FC:A   -  FCAL  N/A  36/74752          42/87168
itsosj-n1>
```

When a disk failure is detected, a hot spare automatically comes on line as a replacement. The IBM System Storage N series storage system enters a degraded mode and rebuild takes place if any spare is available. If the IBM System Storage N series is shut down while it is in degraded mode, reconstruction will be stopped. If the IBM System Storage N series will be powered on again, the reconstruction process starts from the beginning.

If no replacement is available, the IBM System Storage N series operates in degraded mode and shuts down after 24 hours. This value can be increased to 72 hours with the raid.timeout option. If you are running a clustered system, then change the setting on both nodes. Example 8-13 illustrates how to change the setting.

Example 8-13 The raid.timeout setting

```
itsosj-n1> options raid.timeout
raid.timeout                 24         (value might be overwritten in takeover)
itsosj-n1>
itsosj-n1> options raid.timeout 72
You are changing option raid.timeout which applies to both members of
the cluster in takeover mode.
This value must be the same in both cluster members prior to any takeover
or giveback, or that next takeover/giveback may not work correctly.
itsosj-n1> Fri Sep  9 22:05:19 CEST [itsosj-n1: reg.options.cf.change:warning]: Option
raid.timeout changed on one cluster node

itsosj-n1>
```

Attention: In general, we do not recommend that you change raid.timeout settings. If the IBM System Storage N series operates in degraded mode and no spare disk is available, the RAID array cannot be rebuilt and the N series will probably run without RAID protection. A second failure may cause data loss.

Make sure that you have enough spare disks available in your IBM System Storage N series, and replace failed disks after the N series has reconstructed data.

Removing a failed disk

To physically remove from the shelf the disk that you identified as failed:

1. Press down the release mechanism with one hand while grasping the top flange of the shelf with your other hand. See Figure 8-10 on page 235.
2. Pull the disk until it disengages and wait a few seconds until it stops spinning.

3. Remove the disk gently from the bay. See Figure 8-11 on page 235.

Removing a data disk

To remove a disk that reported errors but did not fail, note the disk number from the log messages that report errors (look at the numbers that follow the word *disk*).

Enter the **aggr status -r** command or the **vol status -r** command.

Note the device column of the output of the **sysconfig -r** command, which shows the disk ID of each disk. The location of the disk appears on the right of the disk ID, in the HA SHELF BAY column.

Use **disk fail [-i] disk** to fail the appropriate disk, as shown in Example 8-15 on page 232. Data ONTAP asks for confirmation on failing the specific disk. The -i option fails the disk immediately. Specify the disk that you identified in the log messages.

If no -i option has been specified, the specified disk will be pre-failed. Data will be copied to a replacement disk. If copying is successful, the disk is marked as failed. (This may take a while, depending on the size of the disk and the load of the IBM System Storage N series storage system.) When a copy operation fails, the system turns to degraded mode.

Important: Wait until the data has been copied before using the -i option (pre-failed disk).

If the -i option has been specified, the disk fails immediately and the system runs in degraded mode until the RAID system has been rebuilt. In our example we failed data disk v0.38 (aggregate: aggr_test, which contains two RAID groups, rg0 and rg1). v0.38 is in raid group: rg0; HA: 2 and SHELF BAY: 6.

First we located the disk (v0.38) to fail with the **sysconfig -r** command. See Example 8-14.

Example 8-14 Using the sysconfig -r command to locate the disk

```
itsosj-n1> sysconfig -r
......
Aggregate aggr_test (online, raid_dp) (block checksums)
  Plex /aggr_test/plex0 (online, normal, active)
    RAID group /aggr_test/plex0/rg0 (normal)

      RAID Disk Device  HA  SHELF BAY CHAN Pool Type  RPM  Used (MB/blks)    Phys (MB/blks)
      --------- ------  ------------- ---- ---- ---- ----- --------------    --------------
      dparity   v0.40   v0    2   8   FC:A   -  FCAL  N/A  36/74752          42/87168
      parity    v0.37   v0    2   5   FC:A   -  FCAL  N/A  36/74752          42/87168
      data      v0.38   v0    2   6   FC:A   -  FCAL  N/A  36/74752          42/87168
      data      v0.35   v0    2   3   FC:A   -  FCAL  N/A  36/74752          42/87168

    RAID group /aggr_test/plex0/rg1 (normal)

      RAID Disk Device  HA  SHELF BAY CHAN Pool Type  RPM  Used (MB/blks)    Phys (MB/blks)
      --------- ------  ------------- ---- ---- ---- ----- --------------    --------------
      dparity   v0.34   v0    2   2   FC:A   -  FCAL  N/A  36/74752          42/87168
      parity    v0.33   v0    2   1   FC:A   -  FCAL  N/A  36/74752          42/87168
      data      v0.32   v0    2   0   FC:A   -  FCAL  N/A  36/74752          42/87168
      data      v0.39   v0    2   7   FC:A   -  FCAL  N/A  36/74752          42/87168
......

itsosj-n1>
```

Then we failed the disk *without* the -i option. See Example 8-15. The system asked for confirmation (yes/no). We entered **y** for yes.

Example 8-15 Disk fail without -i option

```
itsosj-n1> disk fail v0.38
*** You are about to prefail the following file system disk, ***
*** which will eventually result in it being failed ***
  Disk /aggr_test/plex0/rg0/v0.38

      RAID Disk Device  HA  SHELF BAY CHAN Pool Type  RPM  Used (MB/blks)    Phys (MB/blks)
      --------- ------  ------------- ---- ---- ---- ----- --------------    --------------
      data      v0.38   v0    2   6   FC:A   -  FCAL  N/A  36/74752          42/87168
***
Really prefail disk v0.38? y
disk fail: The following disk was prefailed: v0.38
Disk v0.38 has been prefailed.  Its contents will be copied to a
replacement disk, and the prefailed disk will be failed out.
itsosj-n1>
```

The disk failure, which was triggered by the **disk fail** command, generated the messages shown in Example 8-16.

Example 8-16 Disk failure messages

```
itsosj-n1> Mon Sep 12 19:25:44 CEST [itsosj-n1: raid.rg.diskcopy.start:notice]:
/aggr_test/plex0/rg0: starting disk copy from v0.38 to v0.41
Mon Sep 12 19:25:55 CEST [itsosj-n1: raid.rg.diskcopy.done:notice]: /aggr_test/plex0/rg0:
disk copy from v0.38 to v0.41 completed in 0:10.57
Mon Sep 12 19:25:55 CEST [itsosj-n1:
raid.config.filesystem.disk.admin.failed.after.copy:info]: File system Disk v0.38 Shelf 2
Bay 6 [NETAPP   VD-16MB-FZ-520   0042] S/N [94283320] is being failed after it was
successfully copied to a replacement.
Mon Sep 12 19:25:55 CEST [itsosj-n1: raid.disk.unload.done:info]: Unload of Disk v0.38
Shelf 2 Bay 6 [NETAPP   VD-16MB-FZ-520   0042] S/N [94283320] has completed successfully

itsosj-n1>
```

Entering the **sysconfig -r** command displays the new data disk v0.41. See Example 8-17.

Example 8-17 The sysconfig -r command shows disk v0.41 as new data disk aggregate

```
itsosj-n1> sysconfig -r
.....
Aggregate aggr_test (online, raid_dp) (block checksums)
  Plex /aggr_test/plex0 (online, normal, active)
    RAID group /aggr_test/plex0/rg0 (normal)

      RAID Disk Device  HA  SHELF BAY CHAN Pool Type  RPM  Used (MB/blks)    Phys (MB/blks)
      --------- ------  ------------- ---- ---- ---- ----- --------------    --------------
      dparity   v0.40   v0    2   8   FC:A   -  FCAL  N/A  36/74752          42/87168
      parity    v0.37   v0    2   5   FC:A   -  FCAL  N/A  36/74752          42/87168
      data      v0.41   v0    2   9   FC:A   -  FCAL  N/A  36/74752          42/87168
      data      v0.35   v0    2   3   FC:A   -  FCAL  N/A  36/74752          42/87168

    RAID group /aggr_test/plex0/rg1 (normal)
```

```
      RAID Disk Device  HA  SHELF BAY CHAN Pool Type  RPM  Used (MB/blks)    Phys (MB/blks)
      --------- ------  ------------- ---- ---- ---- ----- --------------    --------------
     dparity   v0.34   v0    2   2   FC:A   -  FCAL  N/A  36/74752          42/87168
      parity    v0.33   v0    2   1   FC:A   -  FCAL  N/A  36/74752          42/87168
      data      v0.32   v0    2   0   FC:A   -  FCAL  N/A  36/74752          42/87168
      data      v0.39   v0    2   7   FC:A   -  FCAL  N/A  36/74752          42/87168
.....
itsosj-n1>
```

The failed disk v0.38 shows up in the **sysconfig -r** section Broken disks, as shown in Example 8-18.

Example 8-18 The sysconfig -r command shows broken disks (failed disks)

```
itsosj-n1> sysconfig -r
.....
Broken disks

RAID Disk       Device  HA  SHELF BAY CHAN Pool Type  RPM  Used (MB/blks)    Phys (MB/blks)
---------       ------  ------------- ---- ---- ---- ----- --------------    --------------
admin failed    v0.19   v0    1   3   FC:A   -  FCAL  N/A  36/74752          42/87168
admin failed    v0.36   v0    2   4   FC:A   -  FCAL  N/A  36/74752          42/87168
admin failed    v0.38   v0    2   6   FC:A   -  FCAL  N/A  36/74752          42/87168
.....
itsosj-n1>
```

Removing hot spare disks

Use the **sysconfig - r** or **aggr status -s** command to locate spare disks (Example 8-19). Spare disks are listed with their IDs and locations in the shelf and bay.

Example 8-19 Locating spare disks with the sysconfig -r command

```
itsosj-n1> sysconfig -r
Volume vol0 (online, RAID 4) (block checksums)
  Plex /vol0/plex0 (online, normal, active)
    RAID group /vol0/plex0/rg0 (normal)

      RAID Disk Device  HA  SHELF BAY CHAN Pool Type  RPM  Used (MB/blks)    Phys (MB/blks)
      --------- ------  ------------- ---- ---- ---- ----- --------------    --------------
      parity    0b.17   0b    1   1   FC:A   -  FCAL 10000 136000/278528000  137104/280790184
      data      0b.38   0b    2   6   FC:A   -  FCAL 10000 136000/278528000  137104/280790184


Spare disks

RAID Disk       Device  HA  SHELF BAY CHAN Pool Type  RPM  Used (MB/blks)    Phys (MB/blks)
---------       ------  ------------- ---- ---- ---- ----- --------------    --------------
Spare disks for block or zoned checksum traditional volumes or aggregates
spare           0b.19   0b    1   3   FC:A   -  FCAL 10000 136000/278528000  137104/280790184
spare           0b.21   0b    1   5   FC:A   -  FCAL 10000 136000/278528000  137104/280790184
spare           0b.25   0b    1   9   FC:A   -  FCAL 10000 136000/278528000  137104/280790184
spare           0b.27   0b    1   11  FC:A   -  FCAL 10000 136000/278528000  137104/280790184
spare           0b.28   0b    1   12  FC:A   -  FCAL 10000 136000/278528000  137104/280790184
spare           0b.39   0b    2   7   FC:A   -  FCAL 10000 136000/278528000  137104/280790184
spare           0b.40   0b    2   8   FC:A   -  FCAL 10000 136000/278528000  137104/280790184
spare           0b.41   0b    2   9   FC:A   -  FCAL 10000 136000/278528000  137104/280790184
spare           0b.42   0b    2   10  FC:A   -  FCAL 10000 136000/278528000  137104/280790184
spare           0b.43   0b    2   11  FC:A   -  FCAL 10000 136000/278528000  137104/280790184
spare           0b.44   0b    2   12  FC:A   -  FCAL 10000 136000/278528000  137104/280790184
itsosj-n1>
```

More conveniently, you can use FilerView to list spare disks. In the right frame of FilerView, click **Storage** → **Disk** → **Manage**. In the View Type menu, select **Spare Disks** and click **View** (Figure 8-9).

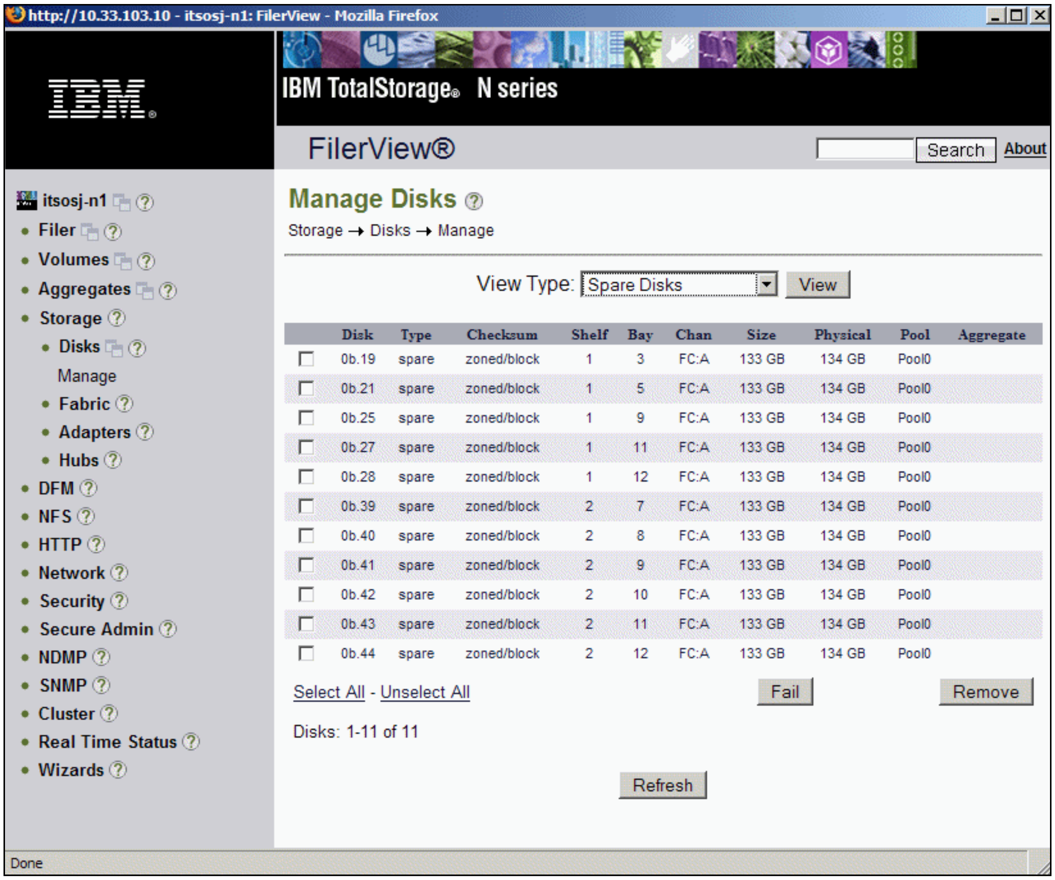


Figure 8-9 FilerView: List spare disk

To remove a Fibre Channel spare disk from the N series, issue the **`disk remove disk_name`** command. As Example 8-20 shows, we removed disk v0.41.

Example 8-20 Removing a spare disk

```
itsosj-n1>
itsosj-n1> disk remove v0.41
Fri Sep  9 22:27:50 CEST [itsosj-n1: raid.config.spare.disk.admin.removed:info]: Spare Disk
v0.41 Shelf 2 Bay 9 [NETAPP   VD-16MB-FZ-520   0042] S/N [94283323] is being removed by
administrator.
Fri Sep  9 22:27:50 CEST [itsosj-n1: raid.disk.unload.done:info]: Unload of Disk v0.41
Shelf 2 Bay 9 [NETAPP   VD-16MB-FZ-520   0042] S/N [94283323] has completed successfully
disk remove: The following disk was removed: v0.41
Removal and unload of disk v0.41 has been initiated.
You will be notified via the system log when unload is complete
itsosj-n1>
```

After removing the disk logically from the N series, wait until the disk stops spinning and then put on an antistatic wrist strap and ground leash and remove the disk physically from the shelf by following these steps:

1. Press the release mechanism down with one hand while grasping the top flange of the shelf with your other hand. See Figure 8-10 on page 235.

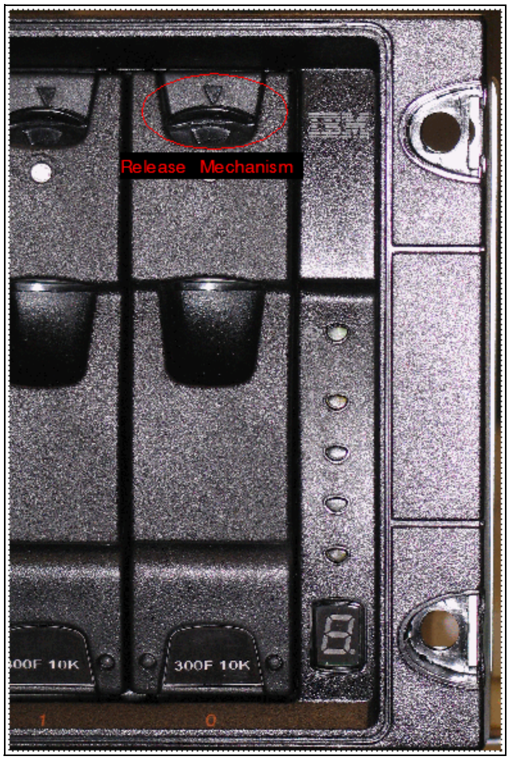


Figure 8-10 Disk drive release mechanism

2. Pull the disk until it disengages and wait a few seconds until it stops spinning.
3. Remove the disk gently from the bay (Figure 8-11).

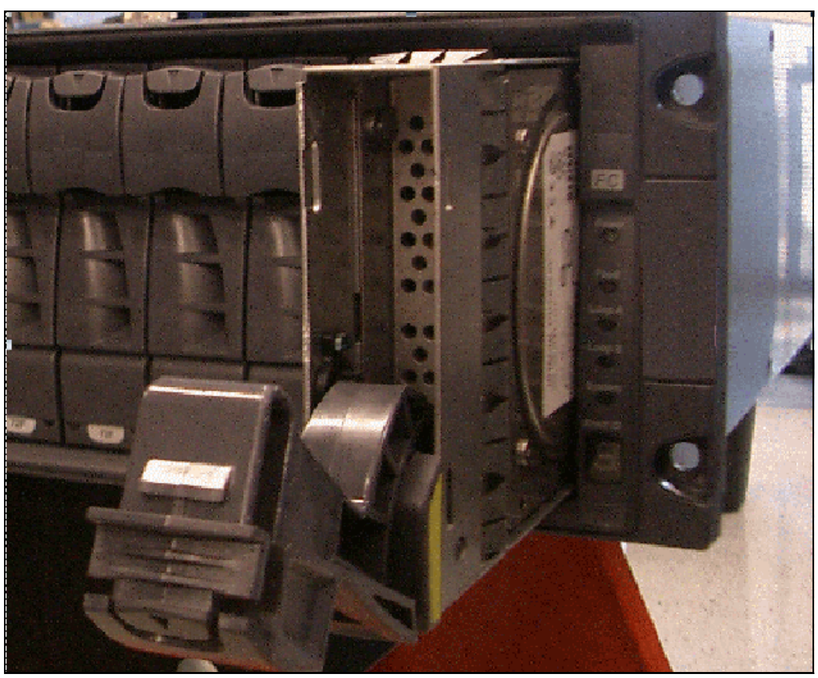


Figure 8-11 Remove disk drive from bay

8.4.2 Adding new disks

After a failed disk is removed physically from the shelf, replace the disk as soon as possible with a new one.

Note: New added disks should always have the same specifications as the existing disks in your shelf. Never use a unsupported disk in the IBM System Storage N series storage system.

To insert a disk:

1. Put on an antistatic wrist strap and ground leash.
2. Insert the disk into the bay with the release mechanism at the top (Figure 8-12).
3. Gently slide the disk in the bay until it engages the backplane. The release mechanism will click into place.

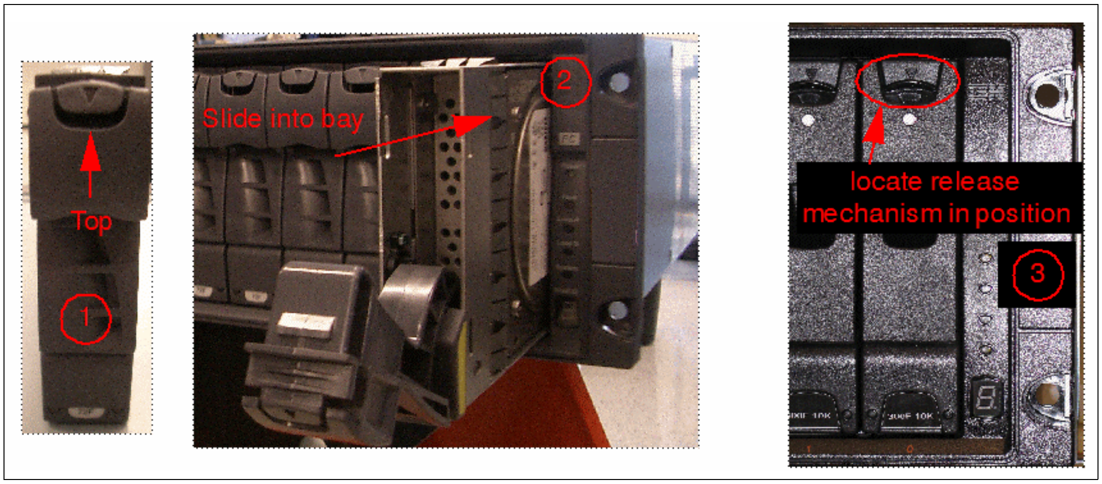


Figure 8-12 Insert disk

After a disk is added physically, the disk becomes the first spare disk and can be added to a RAID group. After a few seconds a message is displayed that a disk was installed.

4. You can verify that the disk was added by entering the **`aggr status -r`** command. The new inserted disk appears in the spare disk section.

When adding multiple disks, it can take up to a minute for them come up to speed and verify device addresses. SES bays must be populated with disks.

8.5 AutoSupport Service

AutoSupport Service for IBM System Storage N series is designed to continuously monitor the health of your storage system using sophisticated, event-driven logging. It allows your IT staff and IBM Support to keep a watchful eye on a multitude of preset conditions. It is designed to automatically send a status message to a centralized IBM support center (and your system administrator, if desired) so that you can take immediate action to correct potential problems. AutoSupport helps your staff monitor your storage environment to keep it working at top efficiency. It helps you anticipate needs before they become issues. Continuous updates to the rules-based AutoSupport processing platform allow you to take advantage of the IBM solution database and proactive notification.

Helping to keep your storage systems up and running

AutoSupport provides a number of powerful capabilities that can help you get the most from your N series storage investment:

- Alerts that identify critical system events such as:
  - Disk, power supply, or fan failure
  - System reboot
  - Cluster takeover
  - And so on
- Included within your warranty or based on your maintenance support entitlements, notification by IBM Support if replacement parts are required
- Constant monitoring of your storage system status
- If desired, automatic e-mail/pager notification of events to your system administrator
- Comprehensive case tracking capability on the web from IBM Support

Attention: Starting with Data ONTAP 7.2.5 AutoSupport Release 4.0 we capture the customer contact information and machine location information to create reliable (IBM) records containing user information. This information is utilized when opening a PMR, so it is import to fill in AutoSupport information correctly for prompt service.

AutoSupport configuration

We can configure ASUP using FilerView or CLI, or using the System Manager. For the detailed procedure consult the updated documentation available in the IBM support website. AutoSupport is enabled by default with Data ONTAP 7.1H2 and later when you configure the system for the first time.

Example 8-21 AutoSupport options

```
itsotuc4> options autosupport
autosupport.cifs.verbose     off
autosupport.content          complete
autosupport.doit             DONT
autosupport.enable           on
autosupport.from             postmaster@storage.tucson.ibm.com
autosupport.local.nht_data.enable off
autosupport.local.performance_data.enable off
autosupport.mailhost         mailhost
autosupport.minimal.subject.id hostname
autosupport.nht_data.enable  on
autosupport.noteto
autosupport.partner.to
autosupport.performance_data.enable on
autosupport.retry.count      15
autosupport.retry.interval   4m
autosupport.support.enable   on
autosupport.support.proxy
autosupport.support.to       callhome@de.ibm.com
autosupport.support.transport smtp
autosupport.support.url      eccgw01.boulder.ibm.com/support/electronic/nas
autosupport.throttle         on
autosupport.to               sandro1_desantis@it.ibm.com
itsotuc4>
```

Example 8-21 shows the output after we issued the related commands in options mode. You must customize the settings for your particular environment.

9

Architecting an N series solution

This chapter discusses the issues to consider when sizing an IBM System Storage N series storage system to a customer's environment. The following topics are covered:

- Performance and throughput
- Capacity requirements
- Effects of optional features
- Future expansion
- Application considerations
- Backup and recovery
- Resiliency to failure
- Configuration limits

A complete explanation is beyond the scope of this book, so only a high-level view planning considerations is presented.

9.1 Primary issues affecting planning

We start with the assumption that a decision has been made to utilize an IBM System Storage N series for storage. The topics of which model IBM System Storage N series to use, what amount of storage would be required on the IBM System Storage N series, which optional features are desired, and future expansion requirements were most likely discussed during the decision-making process.

9.2 Performance and throughput

The performance required from the storage subsystem is usually driven by the number of client systems relying on the IBM System Storage N series for storage service, as well as the demands of the applications running on those client systems. Keep in mind that performance involves a balance of all of the following:

- Performance of a particular IBM System Storage N series model
- Number of disks used for a particular workload
- Type of disks used
- How close to capacity the disks being run are
- Number of network interfaces in use
- Protocols used for storage access
- Workload mix (reads versus writes versus lookups, and so on)
  - Protocol choice
  - Percentage mix of read and write operations
  - Percentage mix of random and sequential operations
  - I/O sizes
  - Working set sizes for random I/O
  - Latency requirements
  - Background tasks running on the storage system (for example, SnapMirror)

> **Note:** Regarding background tasks, it is good practice to always size a storage system to have reserve capacity beyond what is expected to be its normal workload.

Capacity requirements

A key measurement of a storage system is the amount of storage that it provides. Vendors and installers of storage systems generally deal with raw storage capacities. Users, however, are generally only concerned with available capacity. Ensuring that the gap is bridged between raw capacity and usable capacity minimizes surprises both at installation time and in the future.

Raw capacity

Raw capacity is determined by taking the number of disks connected and multiplying by their capacity. So, for example, 14 disks (the maximum number provided in the IBM System Storage N series disk shelves) times 72 GB per drive results in a raw capacity of approximately 1000 GB or 1 TB.

Usable capacity

Usable capacity is determined by factoring out the portion of the raw capacity that goes to support the infrastructure of the storage system. This includes space used for operating system information, disk drive formatting, file system formatting, RAID protection impact, spare disk allocation, mirroring impact, and space used by the Snapshot protection mechanism.

Next we provide an example of where the storage would go in the example 14 x 72 GB drive system. Capacity usually gets utilized as follows:

- Spare disk for protection against a disk failure: It is good practice to allocate spare disk drives to every system. These are utilized if a disk drive fails so that the data on the failed drive can automatically be rebuilt without any operator intervention or downtime.

  For a 14-drive system, a minimum acceptable practice would be to allocate one spare drive.

- RAID-DP™: When a drive fails, it is the RAID information that allows the lost data to be recovered. With the IBM System Storage N series, the maximum protection against loss is provided by using the RAID-DP facility. If capacity is at a premium and reliability is not critical, then RAID 4 can be utilized instead.

  For a 14-drive system, a minimum acceptable practice is to allocate one drive to RAID 4. A better practice is to allocate two drives to RAID-DP.

- When setting up the disks, RAID groups must be defined. A RAID group is a set of data disks protected by one RAID 4 or two RAID-DP parity disks. RAID groups are then combined to create storage aggregates that subsequently have volumes (also referred to as *file systems*) or LUNs allocated on them.

  Normal practice would be to treat the 11 remaining disks as data disks, thus creating a single large aggregate.

At this point in our example we have allocated all 14 available disks. This allocation consists of the following:

  – Spare disk drive: 1
  – RAID parity disks: 2 RAID-DP
  – Data disks: 11

As you can see, we have now lost about 25% of our raw capacity to hardware protection issues. The remaining usable capacity becomes less deterministic from this point because of ever-increasing numbers of variables, but a few firm guidelines are still available.

One near-constant that remains is that you will lose approximately 5% of a disk's raw capacity when it is formatted. This results, in our example, in a 72 GB disk holding only about 68 GB of data (Figure 9-1).

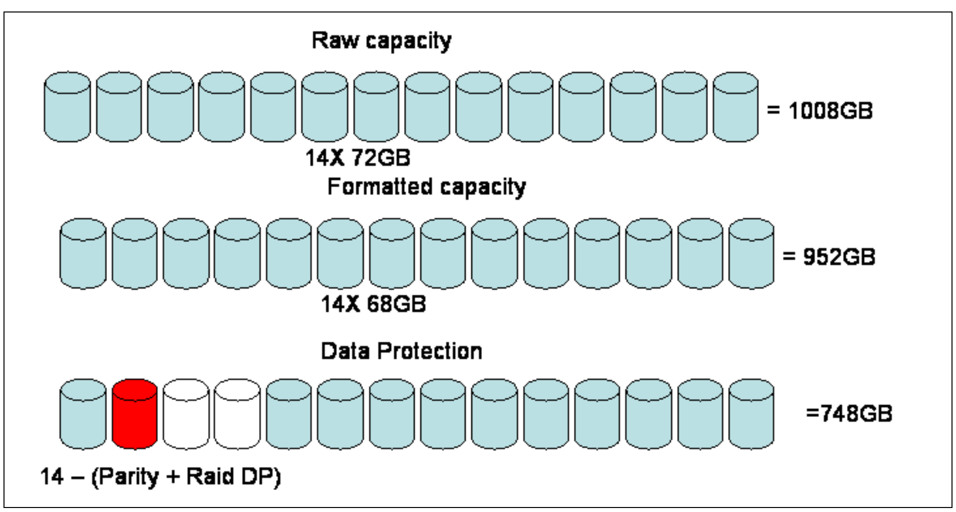


Figure 9-1 Capacity example

Impact of the WAFL file system

Another factor that affects capacity is imposed by the file system. The WAFL file system used by the IBM System Storage N series has less impact than many file systems, but the impact still exists. Generally, WAFL uses an additional 10% of the formatted capacity of a drive.

As a result, our example of 72 GB disk drives are now down to only a little over 60 GB before we put any user data on them at all. If, at this point, we take our 11 data drives and allocate them to a single large volume, we find that the resulting capacity is approximately 600 GB (Figure 9-2).

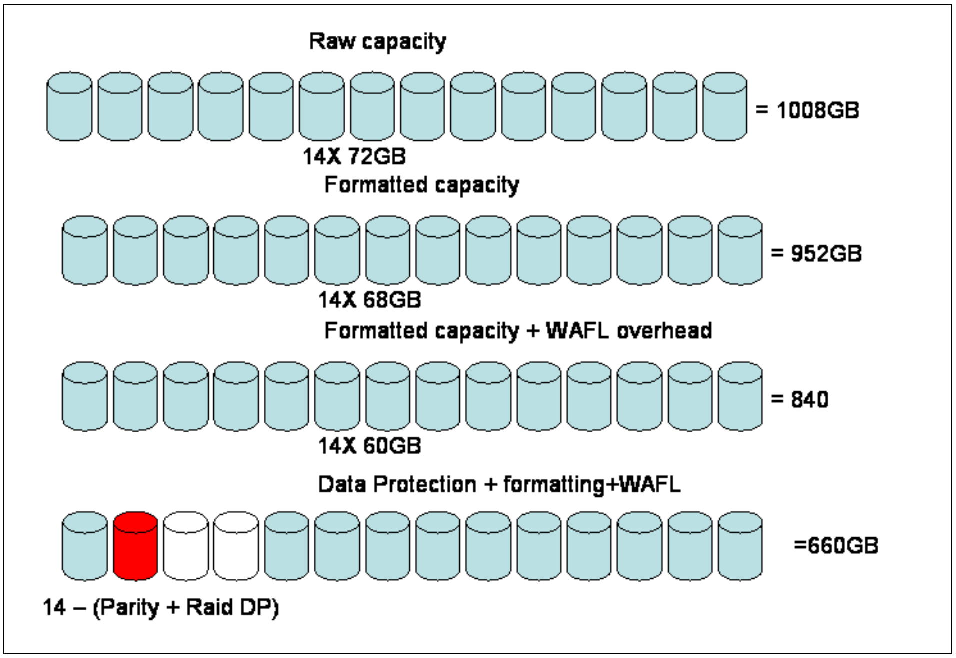


Figure 9-2 Total capacity

Impact of Snapshot protection

Finally, we must consider the impact of Snapshot protection on capacity. Snapshot is a built-in capability that does not utilize any space until it is actually used, but its use affects the apparent usable capacity of the storage system. It is common to run a storage system with 30% of space reserved for Snapshot use. This is space that appears to be unavailable to user storage. (The amount allocated for this purpose can be easily adjusted when necessary to a lower value or a higher value.)

Running with this 30% impact further reduces our 600 GB usable storage to a more realistic 400 GB. See Figure 9-3 on page 243.

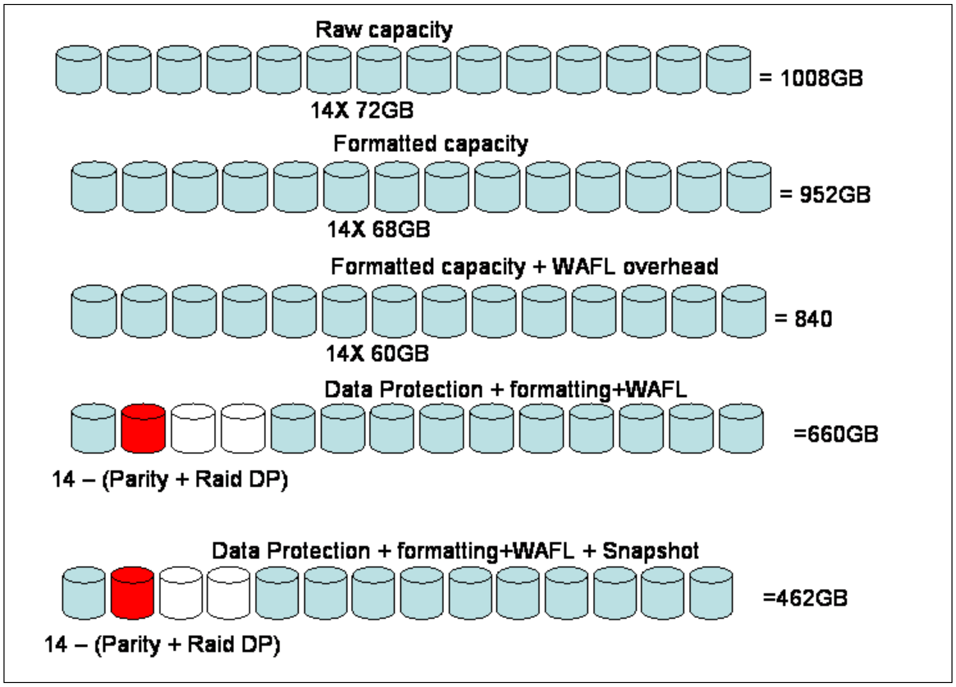


Figure 9-3 True capacity

If your storage system is installed and configured according to this example, running the `df` command on your server will yield a usable capacity in the 400 GB range. Returning to our original intent to focus on reconciling usable storage to raw storage, this example suggests early planning that provides for just under 50% of raw capacity to be ultimately available for storing user data.

Note: This is extremely important when discussing introducing the N series gateway in a pre-existing environment. The user must be aware that the final usable capacity will be different from that available on the external disk system before being virtualized by the gateway.

Other effects of Snapshot

In addition to reducing usable storage capacity, Snapshot protection impacts CPU utilization, memory, and NVRAM. The magnitude of this impact is based on the number of Snapshots done per day and the block change rate.

The factors that determine the impact of Snapshots are:

- How often Snapshots are done

  For example, if Snapshots are done every hour, it may not generally have any impact. However, during peak periods, it may make a big difference.

- What the change rate is (that is, how much of a volume or aggregate is changing)

  Snapshot technology optimizes storage because only changed blocks are copied. So for file access, the change rate is typically in the 3–5% range. For database applications, however, the change rate could be as high as 100%.

- How many Snapshots are kept

  The *frequency* of Snapshots affects resources like CPU and memory. The *number* of Snapshots kept affects storage capacity sizing.

CPU utilization

As a general rule, when your processors approach 90% utilization, it is time to consider increasing the amount of CPU processing power. At 90% average CPU utilization, you are probably already hitting peak periods of 100% or periods of non-responsiveness.

The optimal initial plan would be for 50% average utilization, with peak periods of 70% CPU utilization. While it is easy to give these guidelines after the fact when monitoring is available, for sizing purposes it is best to consult the information posted by the Standard Performance Evaluation Corporation and review workloads and results for initial CPU sizing estimates.

One of the additional things to consider is the configuration limits for iSCSI and Fibre Channel (Table 9-1). Each operating system has host-based configuration limits for FC, FCoE, and iSCSI. For best performance do not configure the system at the maximum values.

Table 9-1 Host operating system configuration limits for iSCSI, FCoE, and Fibre Channel

	Operating system					
Parameter	**Windows**	**Linux**	**HP-UX**	**Solaris**	**AIX**	**VMware**
Visible target ports per host	28	16	16	16	16	16
LUNs per host	64 Windows 2000 128 Windows 2003	128	512	512	128	2.x=128 3.x=256
Paths per LUN	8	4	8 more possible, but pvlinks will only utilize 8	16	16	2.x=4 3.x=8
Max LUN size	2 TB 12 TB Windows 2003 or later	2 TB	2 TB	1023 GB 12 TB with Solaris 9, VxVM, EFI, and appropriate patches	1 TB 12 TB with AIX 5.2ML7 or later and AIX 5.3ML3 or later	2 TB

9.2.1 Effects of optional features

A few optional features affect the early planning required. Most notably, heavy use of the SnapMirror option consumes large amounts of CPU resources. These resources are directly removed from the pool available for serving user and application data, thus resulting in what appears to be an overall reduction in performance. SnapMirror can impact available disk I/O bandwidth and network bandwidth, as well. So if heavy, constant use of SnapMirror is planned, these factors should be adjusted accordingly.

9.2.2 Future expansion

Many of the resources of the storage system can be expanded dynamically. However, early planning can make this expansion even easier and less disruptive if possible future requirements are considered from the start.

Adding disk drives is one simple example. The disk drives and shelves themselves are all hot-pluggable and can be added or replaced without service disruption. But if, for example, all available space in a rack is used by completely full disk shelves, how does a disk drive get added?

Where possible, a good practice from the very beginning is to try to avoid fully populating disk shelves. It is much more flexible to install a new storage system with two half-full disk shelves attached to it rather than a single full shelf. The added cost is generally minimal and is quickly recovered the first time additional disks are added.

Similar consideration can be given to allocating network resources. For instance, if a storage system has two available Gigabit Ethernet interfaces, it is good practice to install and configure both interfaces from the beginning. Commonly, one interface will be configured for actual production use and one configured as a standby in case of failure, although it is also possible (given a network environment that supports this) to configure both interfaces to be in use and providing mutual failover protection to each other. This arrangement provides additional insurance because both interfaces are constantly in use, so you will not find that the standby interface is broken when you need it most – at the time of failure.

Overall, it is valuable during the planning and deployment phase to consider how the environment might change in the future and to engineer in some flexibility from the very beginning.

9.2.3 Application considerations

Different applications and environments put different workloads on the storage system. In this section we discuss a few considerations that are best addressed early in the planning and installation phases.

Home directories and desktop serving

This is a traditional application for network-attached storage solutions. Since many different clients are attached to one or more servers, there is little possibility to effectively plan and model in advance of actual deployment. But a few common sense considerations can help:

- This environment is generally characterized by the use of Network File System (NFS) or Common Internet File System (CIFS) protocols.
- It is generally accessed using Ethernet with TCP/IP as the primary access mechanism.
- The mix of reading and writing is heavily tipped towards the reading side. Uptime requirements are generally less than those for enterprise application situations, so scheduled downtime for maintenance is not too difficult.

In this environment, the requirements for redundancy and maximum uptime are reduced. The importance of data writing throughput is also lessened. More important is the protection offered by Snapshot facilities to protect user data and provide for rapid recovery in case of accidental deletion or corruption. For instance, e-mail viruses can disrupt this type of environment more readily than an environment serving applications like Oracle or SAP.

Load balancing in this environment often takes the form of moving specific home directories from one storage system to another, or moving client machines from one subnet to another. Effective prior planning is difficult. The best planning takes into account that the production environment will be dynamic, and therefore flexibility is key.

It is especially important in this environment to install with maximum flexibility in mind from the very beginning. This environment also tends to make use of a large number of Snapshot images to maximize the protection offered to the user.

Enterprise applications

Previously the domain of direct attached storage (DAS) architectures, it is becoming much more common to deploy enterprise applications utilizing storage systems. These environments have significantly different requirements from the home directory environment. It is common for the emphasis to be on performance, uptime, and backup rather than on flexibility and individual file recovery.

Commonly, these environments utilize a block protocol such as iSCSI or FCP because they mimic DAS more closely than the use of NAS technologies. However, increasingly the advantages and flexibility provided by NAS solutions have been drawing more attention. Rather than designing to serve individual files, the configuration focuses on LUNs or the use of files as though they were LUNs. (An example would be a database application that uses files for its storage instead of LUNs. At its most fundamental, the database application does not treat I/O to files any differently than it does to LUNs, thus allowing the customer to choose the deployment that provides the combination of flexibility and performance required.)

Enterprise environments are usually deployed with their storage systems clustered. This minimizes the possibility of a service outage caused by a failure of the storage appliance. In clustered environments there is always the opportunity to spread workload across at least two active storage systems, so getting good throughput for the enterprise application is generally not difficult.

Of course, this assumes that the application administrator has a good idea of where the workloads are concentrated in the environment so that beneficial balancing can be accomplished. Clustered environments always have multiple I/O paths available, so it is important to balance the workload across these I/O paths as well as across server heads.

Finally, for mission-critical environments, it is important to plan for the worst-case scenario, running the enterprise when one of the storage systems has failed and the remaining single unit has to provide the entire load. In most circumstances, the mere fact that the enterprise is running in spite of a significant failure is viewed as positive, but there are some situations in which the full performance expectation must be met even after a failure. In this case, the storage systems must be sized accordingly.

Block protocols with iSCSI or FCP are also common. The use of a small number of files or LUNs to support the enterprise application means that the distribution of the workload is relatively easy to install and predict.

Microsoft Exchange

MS Exchange has a number of parameters that affect the total storage required of N series. The following are some examples of those parameters:

- Number of instances

  With Microsoft Exchange, you can specify how many instances of an email or document are saved. The default is 1. If you elect to save multiple instances, this should be taken into consideration for storage sizing.

- Number of logs kept

  Microsoft Exchange uses a 5 MB log size. The data change rate determines the number of logs generated per day for recovery purposes. It is not impossible for a highly active Microsoft Exchange server to generate up to 100 logs per day.

- Number of users

  This number, along with mailbox limit, user load, and percentage concurrent access, has a significant impact on the sizing.

- Mailbox limit

  The mailbox limit usually represents the quota assigned to users for their mailboxes. If you have multiple quotas for different user groups, this limit represents the average. This average, multiplied by the number of users, determines the initial storage space required for the mailboxes.

- I/O load per user

  For a new installation, it is difficult to determine the I/O load per user, but you can estimate the load by grouping the users. Engineering and development tend to have a high workload because of drawings and technical documents. Legal may also have a high workload due to the size of legal documents. Normal staff usage, on the other hand, consists of smaller sized I/O transaction workloads and frequency.

  ```
  IOPS/Mailbox = (average disk transfers/sec) / (number of mailboxes)
  ```

- Concurrent users

  Typically, an enterprise's employees do not all work in the same time zone or location. You must estimate the number of concurrent users for a peak period, which is usually the time zone when most of an enterprise's employees are online and during daytime operations.

- Number of storage groups

  Because a storage group cannot span N series storage systems, the number of storage groups has an impact on sizing. There is no recommendation on number of storage groups per IBM System Storage N series storage system, but the number and type of users per storage group should help determine the number of storage groups per storage system.

- Volume type

  Are FlexVols or traditional volumes used? The type of volume used impacts both performance and capacity.

- Drive type

  Earlier in this chapter we discussed the storage capacity impact of drive type. For Microsoft Exchange, the drive type and performance characteristics are also significant, especially with a highly utilized Exchange server. In an active environment, we recommend that you use smaller drives and higher performance characteristics such as RPM and Fibre versus SATA.

- Read-to-write ratio

  The typical read-to-write ratio is 70% to 30%.

- Growth rate

  Industry estimates have placed data storage growth rates at 50% or higher, and you should size for at least two years into the future.

- Deleted mailbox cache space

  This is a feature of Microsoft Exchange that must also be sized for storage usage on the N series. Microsoft allows for a time-specified retention of documents even after deletion of a mailbox. You also must size the storage impact of this feature.

9.2.4 Backup servers

Protecting and archiving critical corporate data is increasingly important. Deploying servers for this purpose is becoming more common and these situations call for their own planning guidelines.

A backup server generally is not designed to deliver great performance. Data center managers rely more on the fact that the backup server is available to receive the backup streams when they are sent. Often the backup server is an intermediate repository for data before it goes to backup tape and ultimately offsite, but frequently the backup server takes the place of backup tapes.

The write throughput of a backup server is frequently the most important factor to consider in planning. Another important factor is the number of simultaneous backup streams that a single server can handle. The more effective the write throughput and the greater the number of simultaneous threads, the more rapidly backup processes complete and the sooner that production servers are taken out of backup mode and returned to full performance.

Different IBM System Storage N series platforms have different capabilities in each of these areas, and the planning process should take these characteristics into account to ensure that the backup server is capable of the workload expected.

9.2.5 Backup and recovery

In addition to backup servers, all storage systems must be backed up. Generally, the goal is to have the backup process occur at a time and in a way that minimizes the impact on overall production. That is why it is common to find large numbers of backup processes scheduled to run during off-hours. But because all of these backups run more or less at the same time, the greatest I/O load put on the storage environment frequently is during these backup activities, instead of during normal production.

IBM System Storage N series storage systems have a number of backup mechanisms available. By using prior planning, the proper use (or usually, a combination of uses) allows you to deploy an environment that provides maximum protection against failure while also optimizing the storage and performance capabilities.

Issues to keep in mind include:

- Storage capacity used by Snapshots

  How much extra storage must be available for Snapshots to consume?

- Networking bandwidth consumed by SnapMirror

  In addition to the production storage I/O paths, SnapMirror needs bandwidth to duplicate data to the remote server.

- Number of possible simultaneous SnapMirror threads

  How many parallel backup operations can be run at once before some resource runs out? Resources to consider include CPU cycles, network throughput, maximum parallel threads (which is platform-dependent), and the amount of data requiring transfer.

- Frequency of SnapMirror operations

  The more frequently data is synchronized, the fewer the number of changes each time. More frequent operations result in background operations running almost all the time.

- Rate at which stored data is modified

  Data that does not change much (for example, archive repositories) does not need to be synchronized as often, and each operation takes less time.

- Use and effect of third-party backup facilities (for example, IBM Tivoli Storage Manager)

  Each third-party backup tool has its unique I/O impacts that must be accounted for.

- Data synchronization requirements of enterprise applications

  Certain applications (for example, DB2®, Oracle, Microsoft Exchange) must be quiesced and flushed prior to performing backup operations to ensure data consistency of backed-up data images.

9.2.6 Resiliency to failure

Like all data processing equipment, storage devices sometimes fail. Most often the failure is of small, uncritical pieces that have redundancy (disks, networks, fans, power supplies, and so on), and that generally have only a small impact (usually none at all) on the production environment. But unforeseen problems can cause rare and infrequent outages of entire storage systems. The most common issues are software problems that occur inside the storage system or infrastructure errors (DNS, routing tables, and so on) that prevent access to the storage system. If a storage system is running but cannot be accessed, the effect on the enterprise is effectively the same as it being completely out of service.

Designing 100% reliable configurations is difficult, time-consuming, and costly. It is more effective to strike a compromise that minimizes the likelihood of error while providing a mechanism to get the server back into service as quickly as possible. In other words, accept the fact that failures will occur, but have a plan ready (and practiced) ahead of time to recover when they do occur.

Spare servers

Some enterprises keep spare equipment around in case of failure. Generally, this is the most expensive solution and is only practical for the largest enterprises.

An often overlooked similar situation is the installation of new servers. Additional or replacement equipment is always being brought into most data environments. Bringing this equipment in a bit early and using it as spare or test equipment is good practice wherever possible. Storage administrators are given an opportunity to practice new procedures and configurations, as well as to test new software without having to do so on production equipment.

Local clustering

The decision to use the clustering features of IBM System Storage N series is determined by the availability and service level agreements affecting the data and applications that run on the IBM System Storage N series storage systems. If it is determined that clustering is needed, it will have an impact on sizing. Rather than sizing for all data, applications, and clients being serviced by one IBM System Storage N series node, the workload is instead divided over two or more nodes.

Failover performance

Another aspect of clustering is failover performance. Perhaps you have already determined that the data, application, or clients require constant availability of the IBM System Storage N series and have elected to use clustering. However, you might have sized for normal operations on each node and not failover. So what was once a normal workload for a single node has now doubled.

You also must consider the service level agreement for response time, data access, and application performance. How long can your customers work within a degraded performance environment? If the answer is not long at all, then the initial sizing of each node also must take failover workload into consideration. Because failover operation is infrequent and usually remedied very quickly, it is difficult to justify these additional standby resources unless maintaining optimum performance is critical. A example is a product ordering system with the data storage or application residing on an IBM System Storage N series storage system, and any impact on the ability to place an order affects sales.

Software upgrades

Frequently, the process of upgrading software causes unexpected outages. Good practice is that software on storage systems should not be upgraded if it provides acceptable and reliable service. Only when new functionality is provided in software or a known defect is expected to be encountered should server software be upgraded.

Upgrade recommendations for Data ONTAP, along with mechanisms for implementing the upgrade, are available on the web at:

`www.ibm.com/storage/support/nas`

Be sure you understand the recommendations from the vendor as well as the risks. Use all the available protection tools (Snapshots, mirrors, and so on) to provide a fallback in case the upgrade introduces more problems than it solves. And whenever possible, perform incremental unit tests on an upgrade before putting an upgrade into critical production.

Testing

Testing storage configurations is often a task not well performed by most enterprises because of cost and time requirements. However, as storage environments become ever more complex and critical, the need for customer-specific testing increases in importance. Customers should work with their storage vendors to determine an appropriate and cost-effective approach to testing various solutions to ensure that their storage configurations are running optimally.

Even more important is that testing of disaster recovery procedures become a regular and ingrained process for everyone involved with storage management.

9.3 Summary

While we provided only a high-level set of guidelines for planning here, we expect that consideration of the issues discussed will maximize the likelihood for a successful initial deployment of an IBM System Storage N series storage system. Other sources of specific planning templates exist or are under development and should be simple to locate using web search queries.

Deploying a network of storage systems is not greatly challenging, and most customers can successfully deploy it themselves by following these guidelines. IBM System Storage N series represents the best in storage appliances, and it is the appliance concept from the beginning that should provide a great deal of confidence. Because of the simplicity that appliances provide, if a mistake is made in the initial deployment, corrective actions are generally not difficult or overly disruptive. For many years customers have iterated their storage system environments into scalable, reliable, and smooth-running configurations. So getting it correct the first time is not nearly as critical as it was prior to the introduction of storage appliances.

If storage system planners and architects remember to keep things simple and flexible, success in deploying an IBM System Storage N series system can be expected.

10

Boot from SAN with IBM System Storage N series

Storage area network (SAN) boot is a technique that allows servers to utilize an operating system (OS) image installed on external SAN-based storage to boot, rather than booting off their own local disk or direct-attached storage. The term *SAN booting* means using a SAN-attached disk, such as a logical unit number (LUN), as a boot device for a SAN host.

Fibre Channel SAN booting does not require support for special SCSI operations. It is no different from any other SCSI disk operation. The host bus adapter (HBA) communicates with the system BIOS, which enables the host to boot from a LUN on the storage system.

This chapter describes in detail the process that you must go through to set up a Fibre Channel Protocol (FCP) SAN boot for your server using a LUN from an FCP SAN-attached N series storage system. We explain the concept of SAN boot and general prerequisites for using this technique, covering these implementations:

- Configure SAN boot with Windows 2003 Enterprise for System x® Servers.
- Configure SAN boot with Windows 2008 Enterprise Server for System x Servers.
- Configure SAN boot for System x Servers with Red Hat Enterprise Linux 5.2.

10.1 Overview

FCP SAN boot, remote boot, and *root boot* are the terms used when referring to a server configuration where the server's operating system is installed on a logical drive that is not resident locally to the server chassis. So why SAN boot? Why is that any better than booting the host OS from local storage? The primary benefits are:

- The ability to create a Snapshot copy of the host OS

  Create a Snapshot copy of the OS before installing a hot fix, service pack, or other risky change to the OS. If it goes bad, just restore the OS from the copy. Further details about Snapshot technology can be found at:

  http://www-03.ibm.com/systems/storage/network/software/snapshot/index.html

- Performance

  The host is likely to boot significantly faster in a SAN boot configuration because you can put several spindles under the boot volume.

- Fault tolerance

  There are multiple disks under the volume in a RAID 4 or RAID-DP configuration.

- The ability to clone flexvol volumes, creating FlexClone volumes

  This host OS cloned LUN can be used for testing purposes. Further information about FlexClone software can be found at:

  http://www-03.ibm.com/systems/storage/network/software/flexvol/index.html

- Interchangeable servers

  By allowing boot images to be stored on the SAN, servers are no longer physically bound to their startup configurations. Therefore, if a server were to fail, it would be very easy to replace it with another generic server and resume operations with the exact same boot image from the SAN (only some minor reconfiguration is required on the storage system). This quick interchange will help reduce downtime and increase host application availability.

- Provisioning for peak usage

  Since the boot image is available on the SAN, it is easy to deploy additional servers to temporarily cope with high workloads.

- Centralized administration

  SAN boot enables simpler management of the startup configurations of servers. Rather than needing to manage boot images at the distributed level at each individual server, SAN boot empowers administrators to manage and maintain the images at a central location in the SAN. This feature enhances storage personnel productivity and helps to streamline administration.

- Utilizes high availability features of SAN storage

  SANs and SAN-based storage are typically designed with high availability in mind. SANs can utilize redundant features in the storage network fabric and RAID controllers to ensure that users do not incur any downtime. Most boot images that are located on local disk or direct-attached storage do not share the same protection. Using SAN boot allows boot images to take advantage of the inherent availability built into most SANs. This helps to increase availability and reliability of the boot image and reduce downtime.

- Efficient disaster recovery process

  Assuming that data (boot image and application data) is mirrored over the SAN between a primary site and a recovery site, servers can take over at the secondary site if a disaster destroys servers at the primary site.

- Reduce overall cost of servers

  Locating server boot images on external SAN storage eliminates the need for local disk in the server. This helps lower costs and allows SAN boot users to purchase servers at a reduced cost but still maintain the same functionality. In addition, SAN boot minimizes the IT costs through consolidation, which reduces the use of electricity and floor space, and through cost savings achieved by centralized management.

10.2 Configure SAN boot for IBM System x servers

This section provides the configuration steps for System x series server SAN boot from N series.

10.2.1 Configuration limits and recommendations

The configuration limits and recommendations are:

- For Windows and Linux-based operating systems, the boot LUN must be assigned as LUN 0 (zero) when doing storage partitioning.
- Enable the BIOS on only one HBA. You must enable the BIOS on the second HBA only if you must reboot the server while the original HBA used for booting purposes, the cable, or the FC switch has failed. In this scenario, you would use QLogic Fast!UTIL or Emulex HBAnyware to select the active HBA, enable the BIOS, scan the BUS to discover the boot LUN, and assign the WWPN and LUN ID to the active HBA. However, when both HBA connections are functional, only one must have its BIOS enabled.
- During the installation of the operating system, there should only be one path active. The reason for this is that no multipathing software is available during the installation of the operating system. The second or alternate path can be activated after the installation of the operating system is complete. You must configure your SAN zoning or remove (disconnect) the HBA cables to leave only one path active.
- This implementation does not make any testing statements about supported configurations. You should always refer to the IBM System Storage N series interoperability matrix for FC and iSCSI SAN, available at:

  http://www-03.ibm.com/systems/storage/network/interophome.html

- In addition, you should review the supported configuration for your server and operating system.

The infrastructure and equipment used for our illustrations consists of the hardware and software listed in Table 10-1.

Table 10-1 Configuration

Server	Operating system	HBA model	N-Series IBM/NetApp	Data ONTAP version
IBM System x 3655 (7985)	Windows 2003 Enterprise SP2	QLOGIC QLE2462	N series 5500 (2865-A20)	7.3
	Windows 2008 Enterprise Server	QLOGIC QLE2462	N series 5500 (2865-A20)	7.3
IBM xSeries® 3850 (8863)	Red Hat Enterprise Linux 5.2	QLOGIC QLA2340	N series 5500 (2865-A20)	7.3
IBM xSeries 225 (8647)	Windows 2003 Enterprise SP2	Emulex LP9802	N series 5500 (2865-A20)	7.3

10.2.2 Best practices

Here are some best practice recommendations that will help you get the most satisfaction out of your N series:

- Fibre Channel queue depth: To avoid host queuing, the host queue depths should not exceed the target queue depths on a per-target basis. For more information about target queue depths and system storage controllers, refer to the FCP Configuration Guide at:

  `http://www-01.ibm.com/support/docview.wss?uid=ssg1S7002547`

- Check the appropriate interoperability matrix at the following website for the latest SAN booting requirements for your operating system:

  `http://www-03.ibm.com/systems/storage/network/interophome.html`

- Volume layout: Volumes containing boot LUNs should be separated from application data to preserve Snapshot data integrity and prevent Snapshot locking when using LUN clones. Even though volumes containing boot LUNs may not require much physical disk space, give the volume enough spindles so that performance is not bound by disk activity. With Data ONTAP Version 7 and later, volumes with boot LUNs can be created on the same aggregate in which the data volumes reside to maximize storage utilization without sacrificing performance.

- RHEL5 now has the capability to detect, create, and install to dm-multipath devices during install. To enable this feature, add the parameter `mpath` to the kernel boot line. At the initial Linux install screen type `Linux mpath` and press Enter to start the Red Hat installation.

- Windows operating system pagefile placement: For Windows 2003 and 2008 configurations, store the pagesys.sys file on the local disk if you suspect pagefile latency issues. Refer to this Microsoft article for more information about pagefiles:

  `http://support.microsoft.com/default.aspx?scid=kb;EN-US;q305547`

  The operating system pagefile is where Windows writes seldom-used blocks from memory to disk to free physical memory. This operation is called paging. There is often concern among administrators when placing the operating system pagefile on SAN storage systems.

  Potential issues with placing the pagefile on a SAN device are:

  – In instances where systems share common resources on the SAN, such as disk spindles, switch bandwidth, or controller CPU and cache, heavy paging operations of one system can impact storage system responsiveness for both operating system and application data for all connected systems.

 - Depending on the device configuration, paging to a SAN device might be slower than paging to local storage. This is unlikely because paging operations benefit from the write cache and multiple disk spindles available from enterprise-class SAN storage systems. These benefits far outweigh the latency induced by a storage networking transport unless the storage is oversubscribed.
 - Sensitivity to bus resets can cause systems to become unstable. Bus resets do not generally affect all systems connected to the SAN because Microsoft has implemented a hierarchical reset handling mechanism within its STORport drivers for Windows Server 2003 to address this behavior.
 - High latency during pagefile access can cause systems to crash with a STOP message (blue screen) or perform very poorly. The disk array should be monitored carefully to prevent oversubscription of the storage, which can result in high latency.
 - Some administrators concerned about paging performance may opt to keep the pagefile on a local disk while storing the operating system on an N series SAN. There are issues with this configuration as well.
 - If the pagefile is moved to a drive other than the boot drive, system crash dumps will not be written. This can be an issue when trying to debug operating system instability in the environment.
 - If the local disk fails and is not mirrored, the system will crash and not be able to boot until the problem is corrected.

 In addition, two pagefiles should not be created on devices with different performance profiles, such as a local disk and a SAN device. Attempting to distribute the pagefile in this manner may result in kernel inpage STOP errors.

 In general, if the system is paging heavily, performance will suffer whether the pagefile is on a SAN device or local disk. The best way to address this problem is to add more physical memory to the system or correct the condition that is causing severe paging. At the time of this publication, the costs of physical memory are such that a small investment can prevent paging and preserve the performance of the environment.

 It is also possible to limit the pagefile size or disable it completely to prevent SAN resource contention. If the pagefile is severely restricted or disabled to preserve performance, it is likely that application instability may result in cases where memory is fully utilized. This option is only recommended for servers that have enough physical memory to cover the anticipated maximum requirements of the application.

- Microsoft Cluster Services and SCSI port drivers: the Microsoft Cluster Service uses bus-level resets in its operation. Without the ability to isolate these resets from the boot device, installations using the SCSIport driver with Microsoft Windows 2000 or 2003 must utilize separate HBAs for the boot device and the shared cluster disks. In deployments where full redundancy is desired, a minimum of four HBAs will be required for MPIO. In Fibre Channel implementations, we strongly recommend that zoning be employed to separate the boot and shared cluster HBAs.

By deploying Microsoft Cluster Services on a Windows Server 2003 platform using STORport drivers, both the boot disks and shared cluster disks can be accessed through the same HBA (Figure 10-1). A registry entry is required to enable a single HBA to connect to both shared and non-shared disks in an MSCS environment. For details, refer to the “Server Clusters: Storage Area Networks - For Windows 2000 and Windows Server 2003” article at:

http://www.microsoft.com/technet/prodtechnol/windowsserver2003/technologies/clustering/starenet.mspx

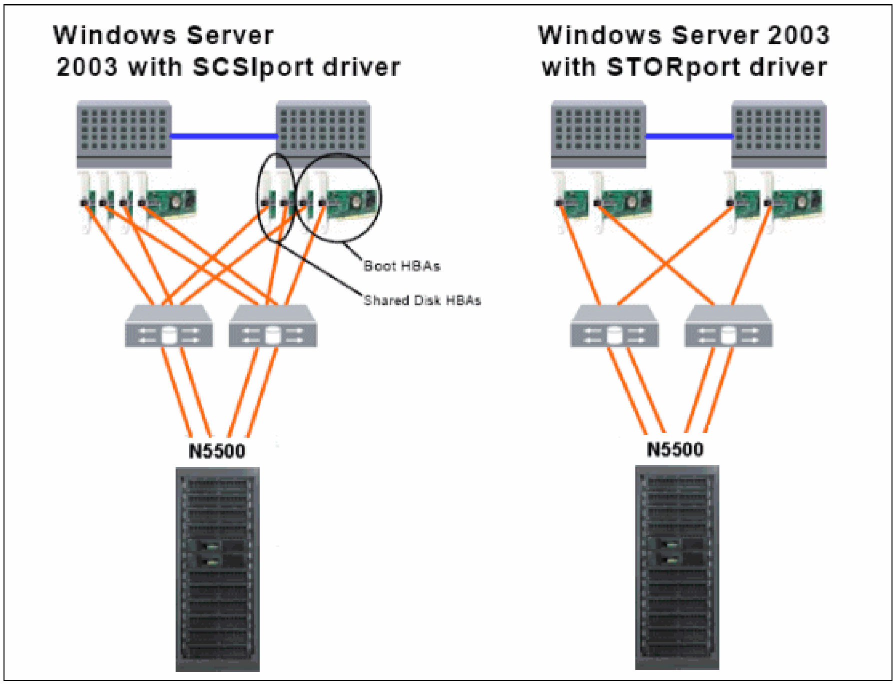


Figure 10-1 Windows Server 2003 platform using STORport drivers

10.2.3 Basics of the boot process

The boot process of the IA32 architecture has not changed significantly since the early days of the personal computer. Before the actual loading of the operating system from disk takes place, a pre-boot process is completed by the host BIOS routines:

1. Power on the self test.

   The BIOS initiates a diagnostic test of all hardware devices for which a routine exists. Devices for which the system BIOS does not have direct knowledge, such as installed HBAs, execute their own routines after the system tests have completed.

2. Initialize.

   The BIOS routines clear system memory, processor registers, and initialize all devices.

3. Set the boot device.

   Although multiple devices may be bootable (that is, the CD-ROM, a floppy disk drive, network adapter, storage HBA), only one can be the actual boot device. The BIOS determine the correct boot device order based on each device's *ready* status and the stored configuration.

4. Load the boot sector.

   The first sector of the boot device, which contains the MBR, is loaded. The MBR contains the address of the bootable partition on the disk, where the operating system resides.

10.2.4 Configuring SAN booting before installing Windows or Linux systems

Note: For the latest information about SAN booting, including restrictions and configuration recommendations, refer to Support for FCP/iSCSI Host Utilities on Windows at:

https://www-304.ibm.com/systems/support/myview/supportsite.wss/selectproduct?taskind=7&brandind=5000029&familyind=5364556&typeind=0&modelind=0&osind=0&psid=sr&continue.x=1

And for Linux Support for FCP/iSCSI Host Utilities see:

http://www-304.ibm.com/systems/support/myview/supportsite.wss/selectproduct?taskind=7&brandind=5000029&familyind=5364552&typeind=0&modelind=0&osind=0&psid=sr&continue.x=1

To use a LUN as a boot device, complete the following steps:

1. Obtain the WWPN of the initiator HBA installed in the host. The WWPN is required to configure the initiator group on the storage system and map the LUN to it.

   Note: After you obtain the WWPN for the HBA, you must create the LUN to use as a boot device, map it to an initiator group, and assign a LUN ID. You must assign a LUN ID of 0. Refer to IBM System Storage N series Data ONTAP 7.3 Block Access Management Guide for iSCSI and FCP at the following website for detailed information about how create the LUN that will be used as a boot device and map the LUN to an initiator group:

   http://www-01.ibm.com/support/docview.wss?uid=ssg1S7002547

2. Enable and configure BootBIOS on the HBA to use the LUN as a boot device.
3. Configure the PC BIOS boot order to make the LUN the first disk device.

Obtaining the WWPN of the initiator HBA

Before you create the LUN on the storage system and map it to an igroup, you must obtain the WWPN of the HBA installed in the host. The WWPN is required when you create the igroup. You can obtain the WWPN using one of the following tools:

- Emulex BIOS Utility
- Qlogic Fast!UTIL

Obtaining the WWPN using Emulex BIOS Utility

To obtain the WWPN using the Emulex BIOS Utility:

1. Reboot the host.
2. Press Alt+E to access the Emulex BIOS Utility.

3. Select the appropriate adapter and press Enter, as shown in Figure 10-2.

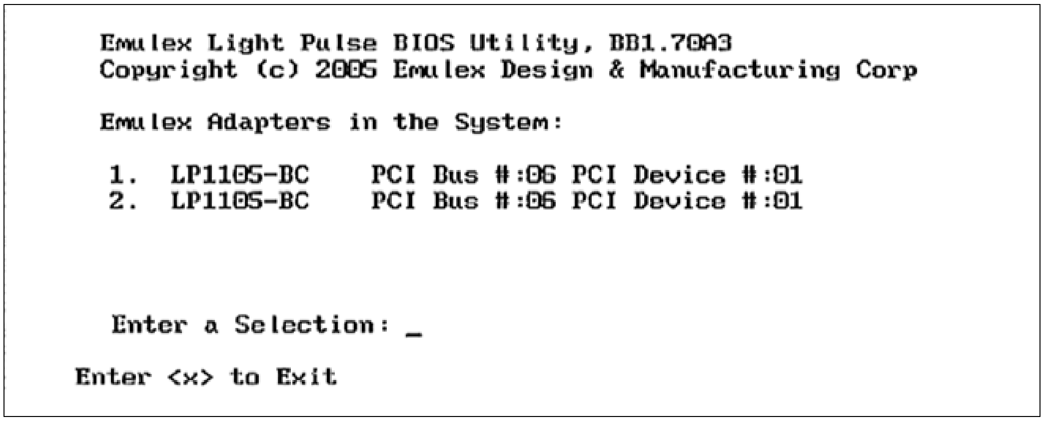


Figure 10-2 Emulex BIOS Utility

BootBIOS displays the configuration information for the HBA, including the WWPN, as shown in Figure 10-3.

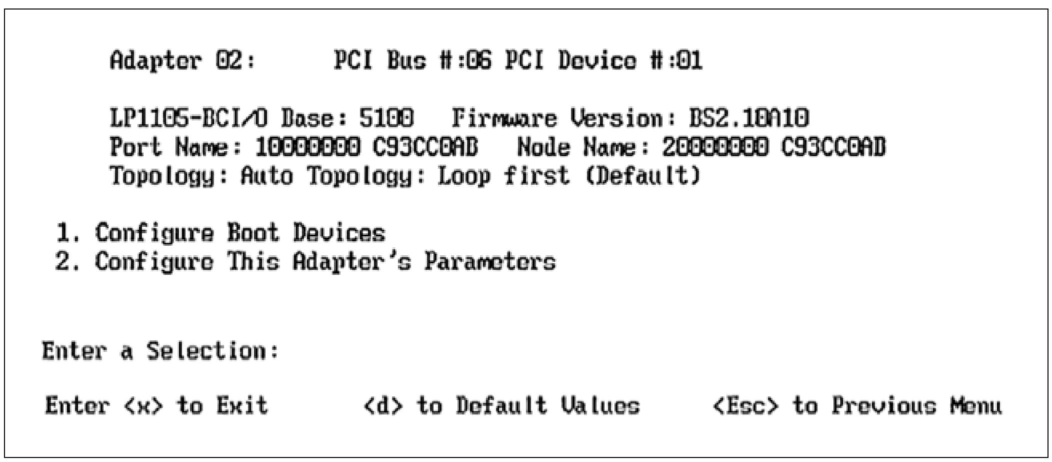


Figure 10-3 Adapter 02 screen

4. Record the WWPN for the HBA.

Obtaining the WWPN using Qlogic Fast!UTIL

To obtain the WWPN using Qlogic Fast!UTIL:

1. Reboot the host.
2. Press Ctrl+Q to access BootBIOS.

3. BootBIOS displays a menu of available adapters. Select the appropriate HBA and press Enter, as shown in Figure 10-4.

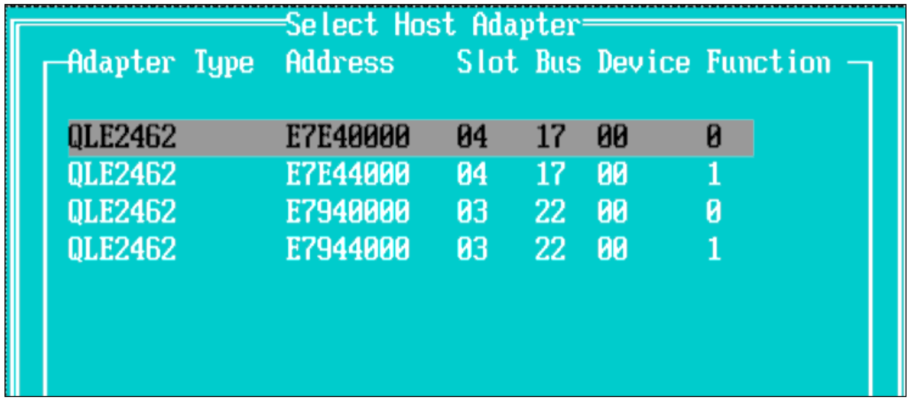


Figure 10-4 Selecting host adapter

4. The Fast!UTIL options appear. Select **Configuration Settings** and press Enter, as shown in Figure 10-5.

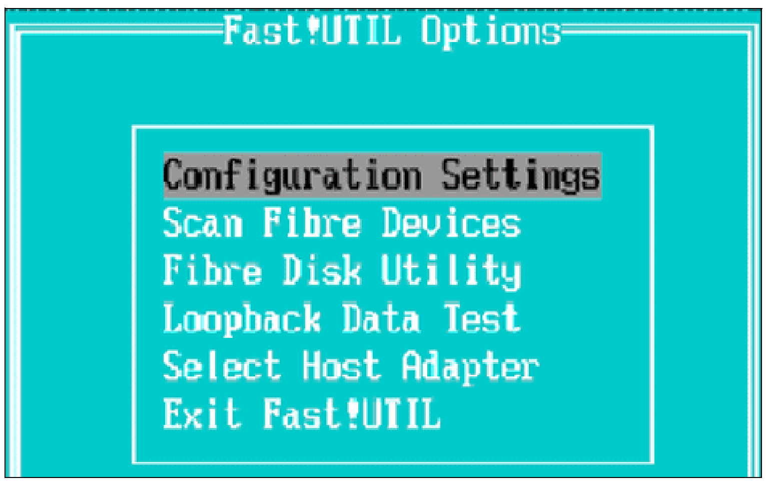


Figure 10-5 Fast!UTIL Options screen

5. Select **Adapter Settings** and press Enter (Figure 10-6).

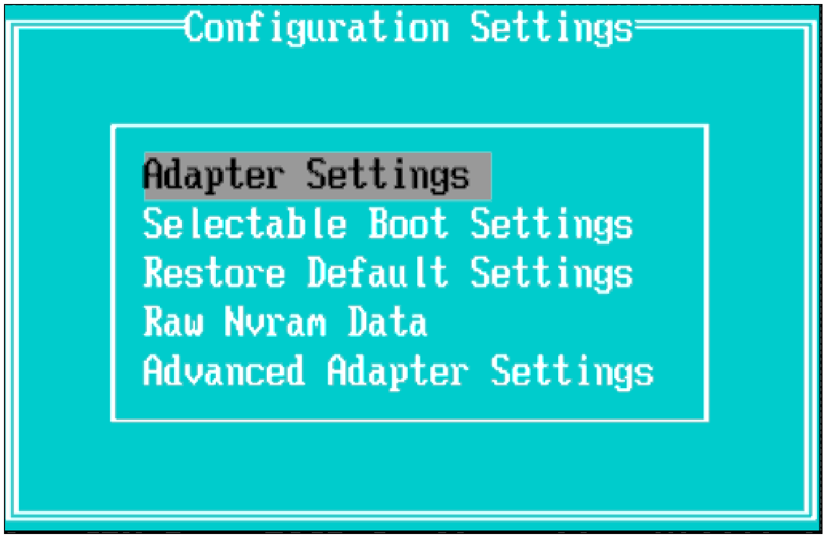


Figure 10-6 Configuration Settings screen

The adapter settings appear, including the WWPN, as shown in Figure 10-7.

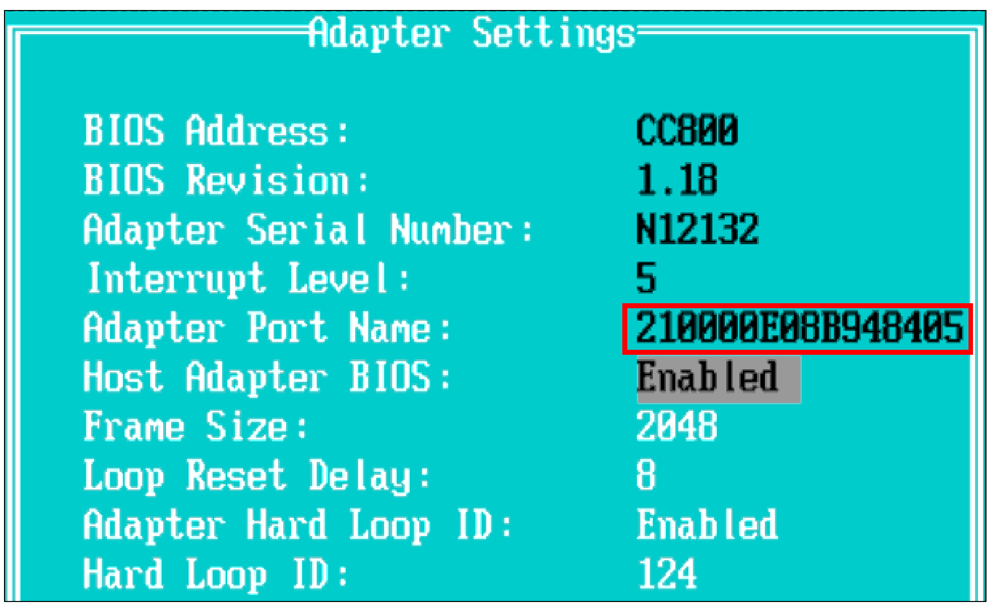


Figure 10-7 Enabling host adapter BIOS in Adapter Settings menu

6. Record the WWPN from the Adapter Port Name field.

Enable and configure BootBIOS on the HBA

BootBIOS enables the HBA to access the existing BIOS on Intel 32-bit, Intel Xeon® 64-bit, and AMD Opteron 64-bit systems. It also enables you to designate a Fibre Channel drive, such as a storage system LUN, as the host's boot device.

BootBIOS firmware is installed on the HBA that you purchased.

Note: Ensure that you are using the version of firmware required by this FCP Windows Host Utility.

BootBIOS firmware is disabled by default. To configure SAN booting, you must first enable BootBIOS firmware and then configure it to boot from a SAN disk.

You can enable and configure BootBIOS on the HBA using one of the following tools:

- Emulex LP6DUTIL.EXE: The default configuration for the Emulex expansion card for x86 BootBIOS in the Universal Boot Code image is not enabled at startup, disallowing access to the BIOS Utility on power up. Otherwise, press Alt+E. In Figure 10-8 the x86 BootBIOS is enabled at startup, so we press Alt+E to access the BIOS Utility.
- Qlogic Fast!UTIL: Enable BootBIOS for Qlogic HBAs using FastUTIL!.

Enabling and configuring Emulex BootBIOS

To enable BootBIOS:

1. Power on your server and press Alt+L to open the Emulex BIOS Utility.
2. Select the appropriate adapter and press Enter, as shown in Figure 10-8.

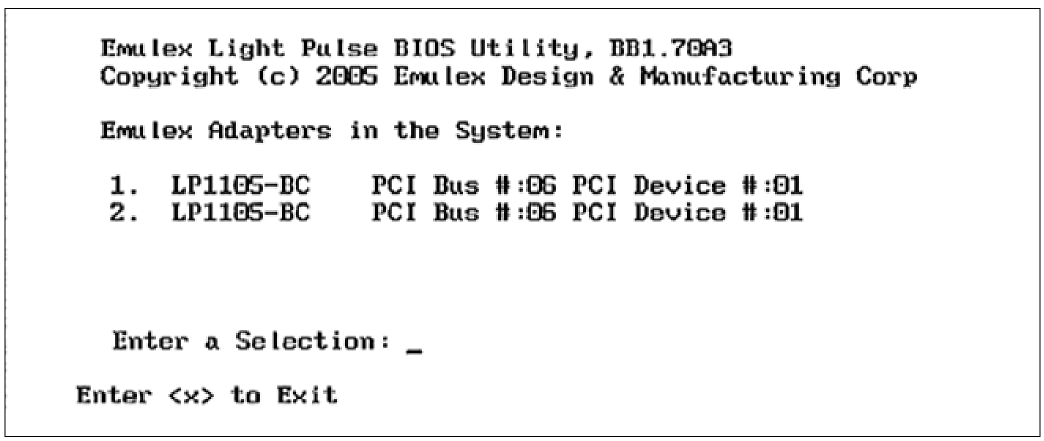


Figure 10-8 Emulex BIOS Utility

3. Select **2** to configure the adapter’s parameters and press Enter, as shown in Figure 10-9.

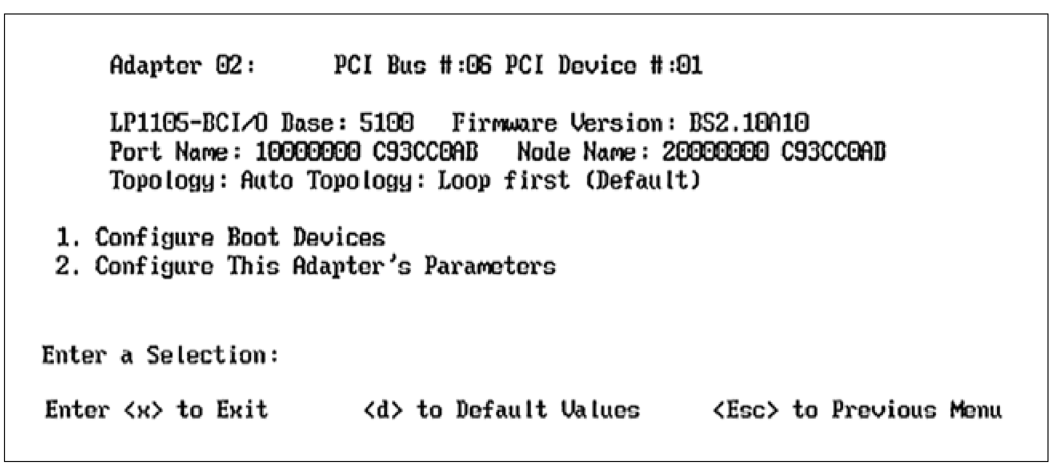


Figure 10-9 Adapter 02 screen

4. From the configure adapter’s parameters menu, select **1** to enable the BIOS, as shown in Figure 10-10.

```
Adapter 02:      PCI Bus #:06 PCI Device #:01

LP1105-BCI/O Base: 5100   Firmware Version: BS2.10A10
Port Name: 10000000 C93CC0AB   Node Name: 20000000 C93CC0AB
Topology: Auto Topology: Loop first (Default)

1. Enable or Disable BIOS
2. Change Default ALPA of this adapter
3. Change PLOGI Retry Timer (+Advanced Option+)
4. Topology Selection (+Advanced Option+)
5. Enable or Disable Spinup delay (+Advanced Option+)
6. Auto Scan Setting (+Advanced Option+)
7. Enable or Disable EDD 3.0 (+Advanced Option+)
8. Enable or Disable Start Unit Command (+Advanced Option+)
9. Enable or Disable Environment Variable (+Advanced Option+)
A. Auto Sector Format Select (+Advanced Option+)

 Enter a Selection: _

 Enter <x> to Exit          <Esc> to Previous Menu
```

Figure 10-10 Configure the adapter’s parameters screen

5. This screen shows the BIOS disabled. Select **1** to enable the BIOS, as shown in Figure 10-11.

```
Adapter 02:      PCI Bus #:06 PCI Device #:01


The BIOS is Disabled!!


Enable Press 1, Disable Press 2:_


Enter <x> to Exit          <Esc> to Previous Menu
```

Figure 10-11 Enable/disable BIOS screen

The BIOS is now enabled, as shown in Figure 10-12.

```
Adapter 01:              PCI  Bus:08 Device:01 Function:00

The BIOS is Enabled!!

Enable Press 1, Disable Press 2:

Enter <x> to Exit         <Esc> to Previous Menu
```

Figure 10-12 Enable BIOS success screen

6. Press Esc to return to the configure adapter's parameters menu, as shown in Figure 10-13.

```
    Adapter 02:      PCI Bus #:06 PCI Device #:01

    LP1105-BCI/O Base: 5100   Firmware Version: BS2.10A10
    Port Name: 10000000 C93CC0AB   Node Name: 20000000 C93CC0AB
    Topology: Auto Topology: Loop first (Default)

1. Enable or Disable BIOS
2. Change Default ALPA of this adapter
3. Change PLOGI Retry Timer (+Advanced Option+)
4. Topology Selection (+Advanced Option+)
5. Enable or Disable Spinup delay (+Advanced Option+)
6. Auto Scan Setting (+Advanced Option+)
7. Enable or Disable EDD 3.0 (+Advanced Option+)
8. Enable or Disable Start Unit Command (+Advanced Option+)
9. Enable or Disable Environment Variable (+Advanced Option+)
A. Auto Sector Format Select (+Advanced Option+)

  Enter a Selection: _

Enter <x> to Exit          <Esc> to Previous Menu
```

Figure 10-13 Configure adapter's parameters screen

7. Press Esc one more time to return to the main configuration menu. You are now ready to configure your boot devices. Select **1** to configure the boot devices, as shown in Figure 10-14.

> **Note:** The Emulex adapter supports FC_AL (public and private loop) and fabric point-to-point. During initialization, the adapter determines the appropriate network topology and scans for all possible target devices.

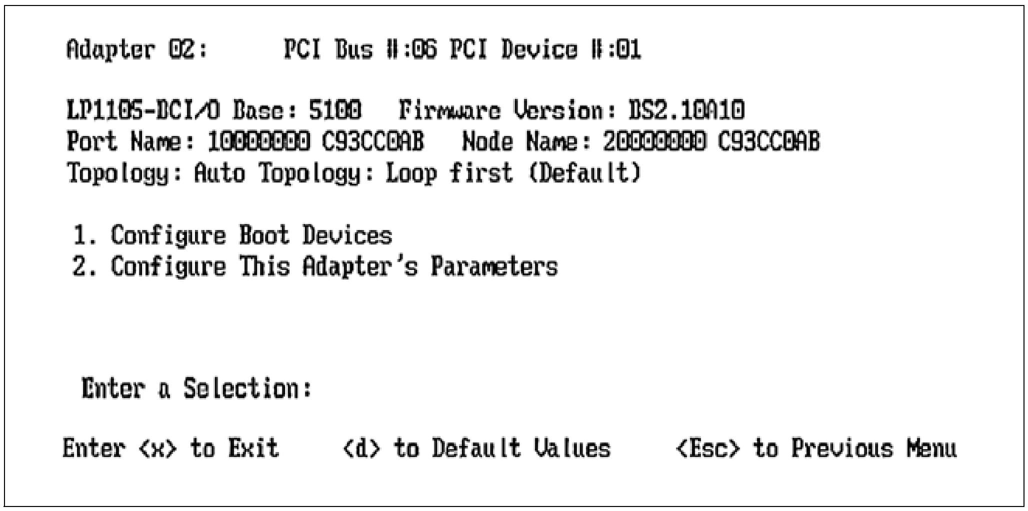


Figure 10-14   Adapter 02 screen

8. The eight boot entries are zero by default. The primary boot device is the first entry exhibit, and it is the first bootable device. Select a boot entry to configure. Select **1**, as shown in Figure 10-15.

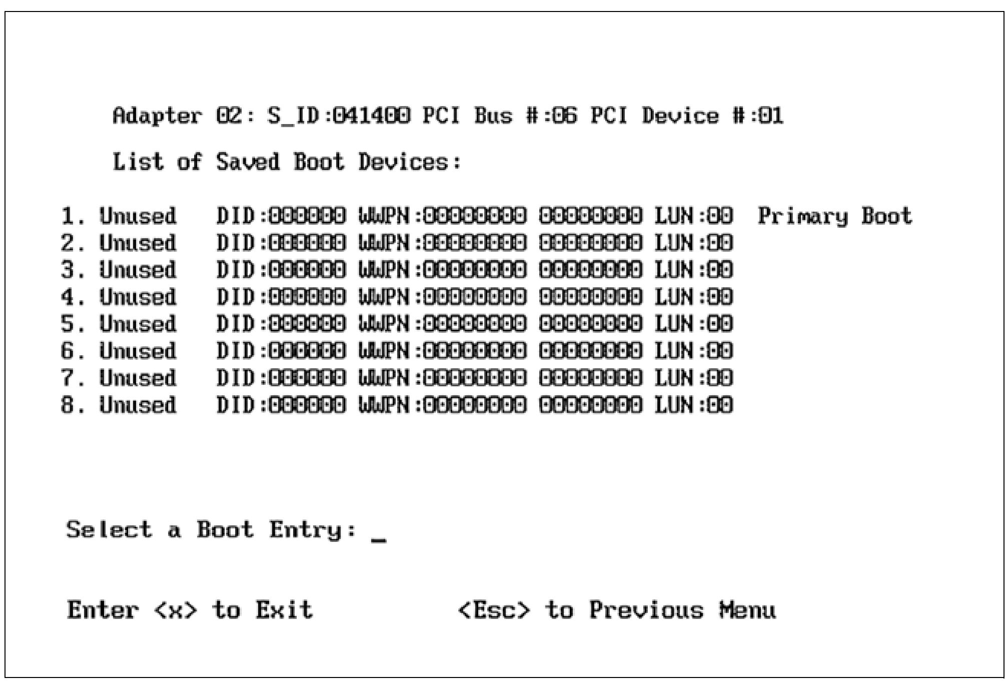


Figure 10-15   Configure boot device screen

Note: Target device failover: If the first boot entry fails due to a hardware error, the system can boot from the second bootable entry. If the second boot entry fails, the system boots from the third bootable entry and so on, up to eight distinct entries. This provides failover protection by automatically redirecting the boot device without user intervention.

9. At initialization, Emulex scans for all possible targets or boot devices. If the HBA is attached to a storage array, the storage device is visible. To view the LUNs, select the storage array controller. Figure 10-16 shows two arrays within the entry field. Select **01** and press Enter.

```
    Adapter 02: S_ID:041400 PCI Bus #:06 PCI Device #:01

00. Clear selected boot entry!!
01. DID:041300 WWPN:50050763 031840C6 LUN:00    IBM     2107900         .208
02. DID:041500 WWPN:50050763 030300C6 LUN:00    IBM     2107900         .208



  Select The Two Digit Number of The Desired Boot Device:

Enter <x> to Exit        <Esc> to Previous Menu       <PageDn> to Next Page
```

Figure 10-16 Boot device entry field screen

Note: Device scanning: The adapter scans the fabric for Fibre Channel devices and lists all the connected devices by DID and WWPN. Information about each device is listed, for example, starting LUN number, vendor ID, product ID, and product revision level.

10. A pop-up window requests entry of the starting LUN number to display. Enter `00` to display the first 16 LUNS, as shown in Figure 10-17.

```
    Adapter 02: S_ID:041400 PCI Bus #:06 PCI Device #:01

00. Clear selected boot entry!!
01. DID:041300 WWPN:50050763 031840C6 LUN:00    IBM     2107900         .208
02. DID:041500 WWPN:50050763 030300C6 LUN:00    IBM     2107900         .208

              DID:041300 WWPN:50050763 031840C6

           Enter two digits of starting LUN (Hex):

                   <Esc> to Previous Menu

  Select The Two Digit Number of The Desired Boot Device:01

Enter <x> to Exit        <Esc> to Previous Menu       <PageDn> to Next Page
```

Figure 10-17 Starting LUN number screen

11. BootBIOS displays a menu of bootable devices. The devices are listed in boot order. The primary boot device is the first device listed. If the primary boot device is unavailable, the host boots from the next available device in the list. In our example, as shown in Figure 10-18, only one LUN is available because we configured SAN zoning to one path only, as recommended in 10.2.1, “Configuration limits and recommendations” on page 253. Type `01` to select the primary boot entry, and press Enter.

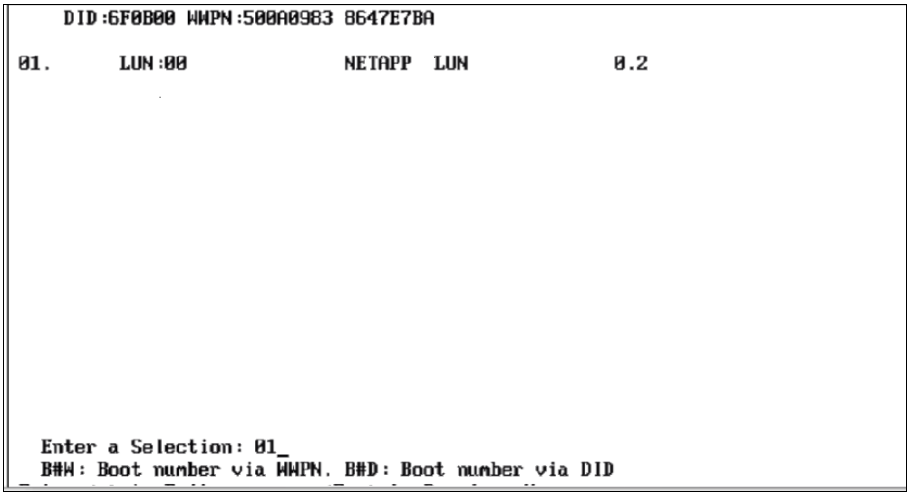


Figure 10-18 Bootable devices menu

12. After the LUN is selected another menu prompts you to specify how the boot device will be identified. We recommend using the Worldwide Port Name (WWPN) for all boot-from-SAN configurations. Select item **1** to boot this device using the WWPN, as shown in Figure 10-19.

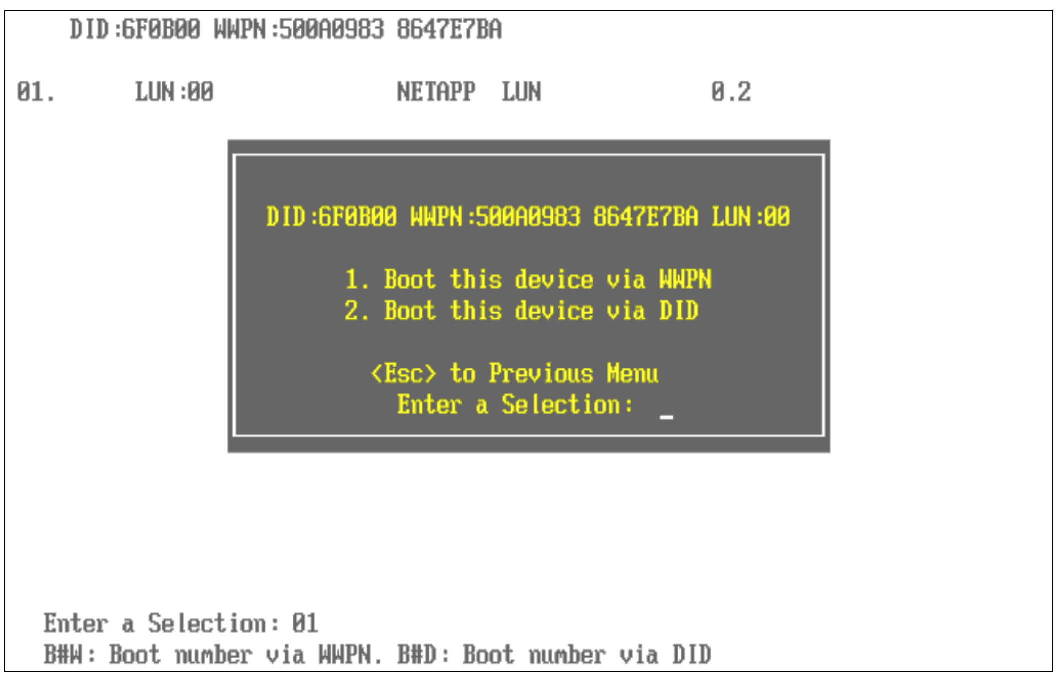


Figure 10-19 How the boot device will be identified

13. Once complete, press **x**, as shown in Figure 10-20, to exit and save your configuration. Your HBA’s BootBIOS should now be properly configured to boot from a SAN on the attached storage device.

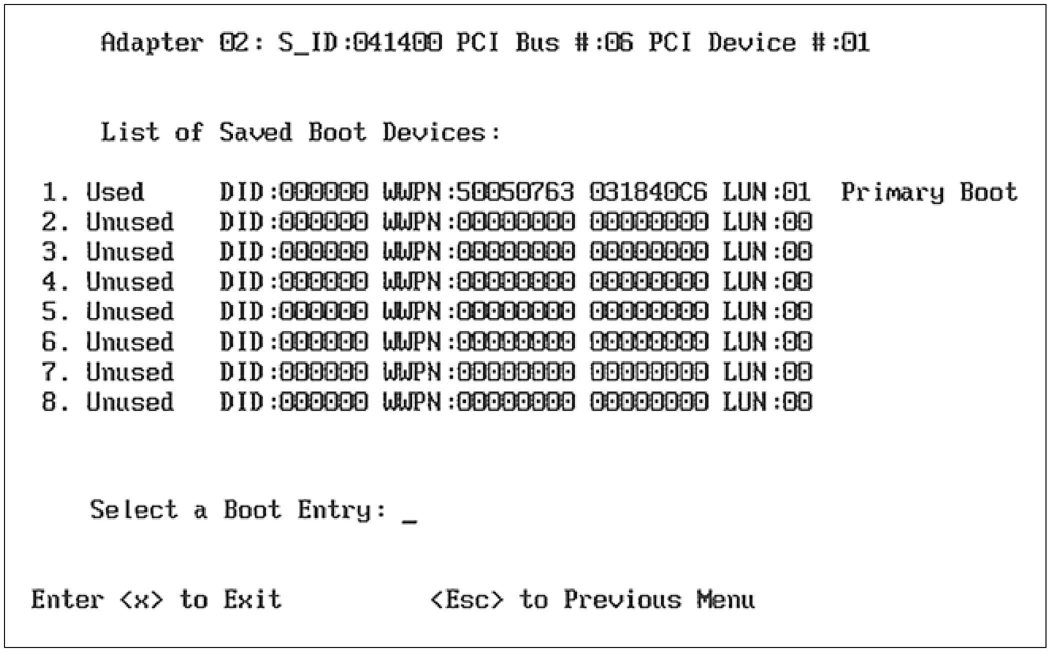


Figure 10-20 Exit Emulex Boot Utility and saved boot device screen

14. Press **Y** to reboot your system, as shown in Figure 10-21.

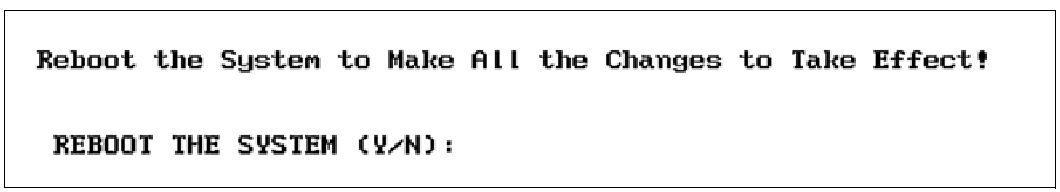


Figure 10-21 Reboot system confirmation screen

Enabling and configuring QLOGIC BootBIOS

Perform these steps to configure QLOGIC BootBIOS:

1. Power on or reboot your host.
2. Press Ctrl+Q or Alt+Q to enter the BIOS configuration utility, as shown in Figure 10-22.

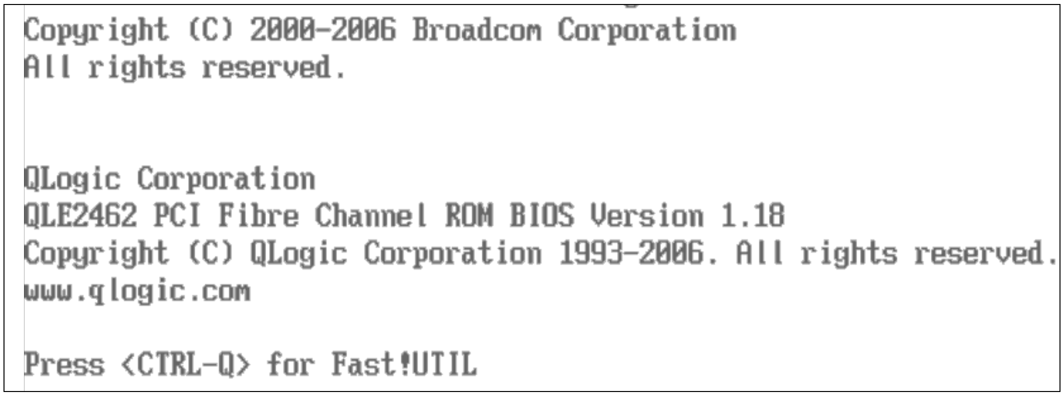


Figure 10-22 Pressing Ctrl+Q for Fast!UTIL screen

3. The Qlogic Fast!UTIL displays the available adapters, listed in boot order. The primary boot device is the first device listed. If the primary boot device is unavailable, the host boots from the next available device in the list. Select the first Fibre Channel adapter port and press Enter, as shown in Figure 10-23.

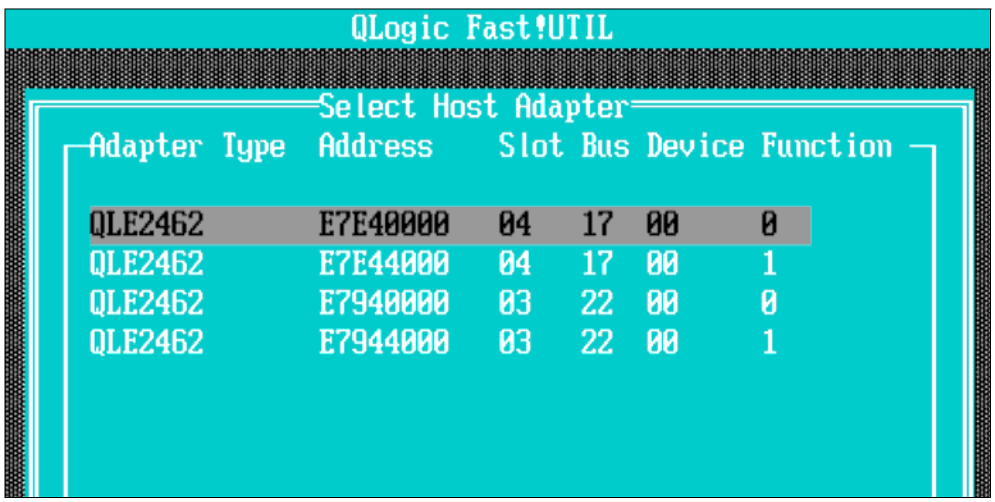


Figure 10-23 Qlogic Fast!UTIL menu

4. Select **Configuration Settings** and press Enter, as shown in Figure 10-24.

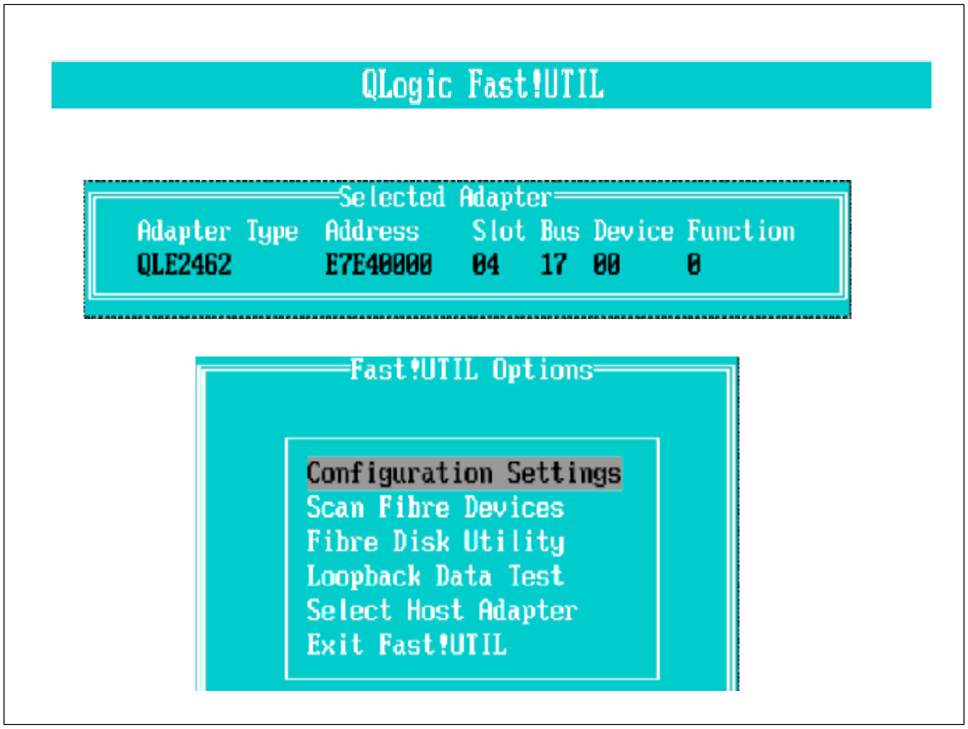


Figure 10-24 Configuration settings for QLE2462 adapter screen

5. Select **Adapter Settings** and press Enter, as shown in Figure 10-25.

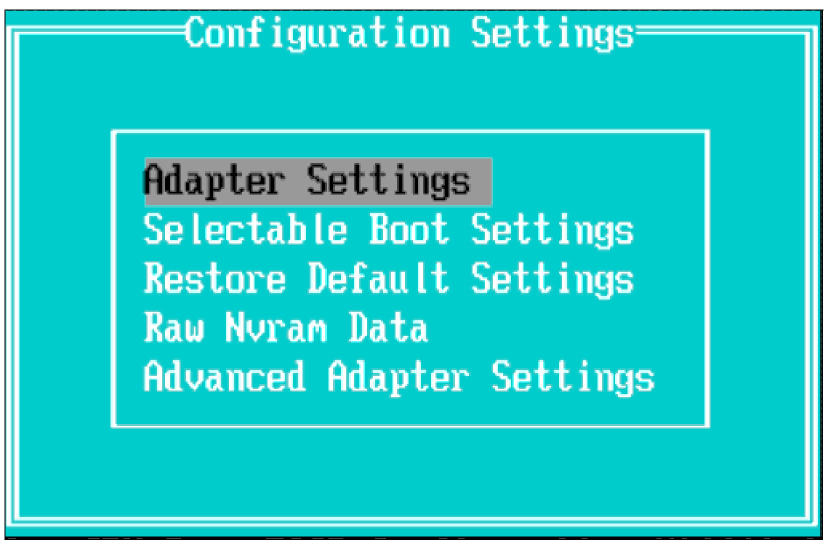


Figure 10-25 Adapter Settings screen

6. Scroll to Host Adapter BIOS, as shown in Figure 10-26.
   – If this option is disabled, press Enter to enable it.
   – If this option is enabled, go to the next step.

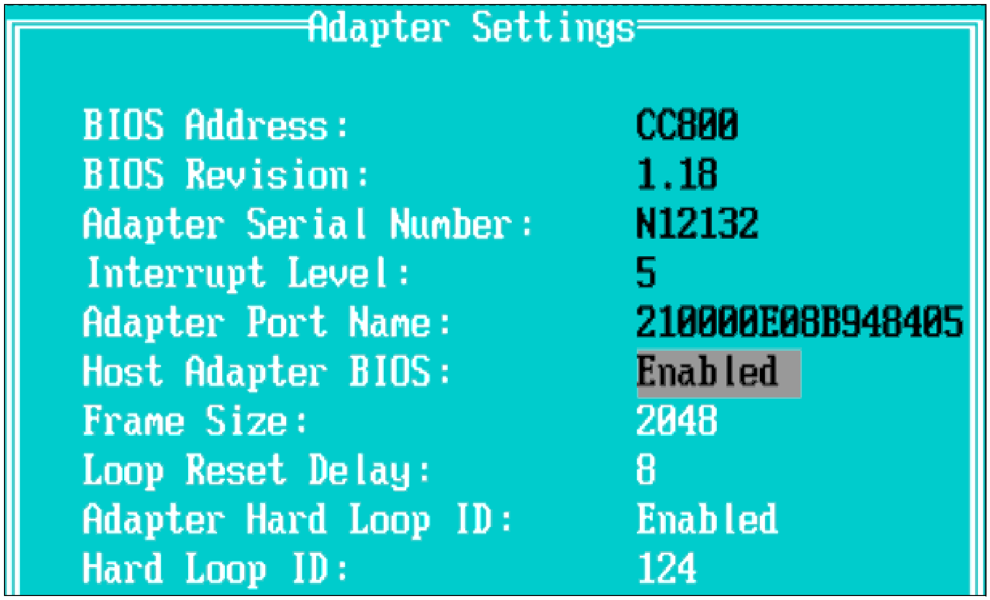


Figure 10-26 Enabling host adapter BIOS

7. Press Esc to return to the Configuration Settings screen. Scroll to **Selectable Boot Settings** and press Enter, as shown in Figure 10-27.

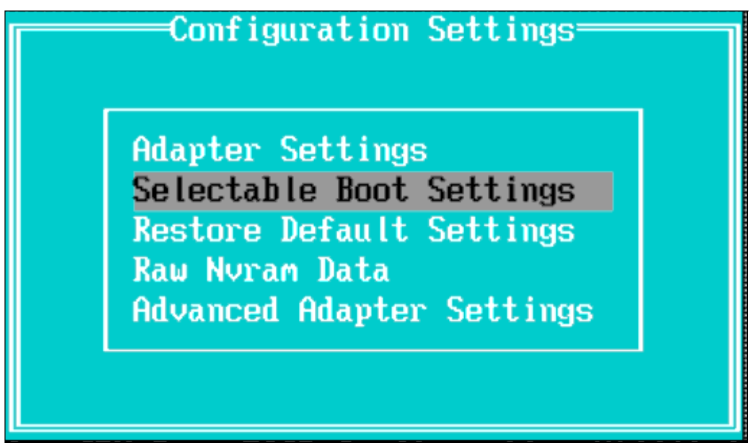


Figure 10-27 Accessing selectable boot settings

8. Scroll to Selectable Boot, as shown in Figure 10-28.
   - If this option is disabled, press Enter to enable it.
   - If this option is enabled, go to the next step.

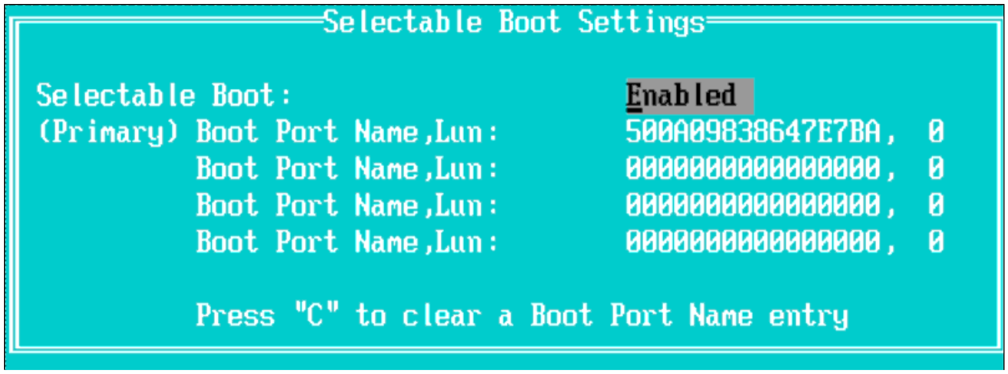


Figure 10-28 Enabling selectable boot in Selectable Boot Settings screen

9. Select the entry in the (Primary) Boot Port Name, LUN field, as shown in Figure 10-29, and press Enter.

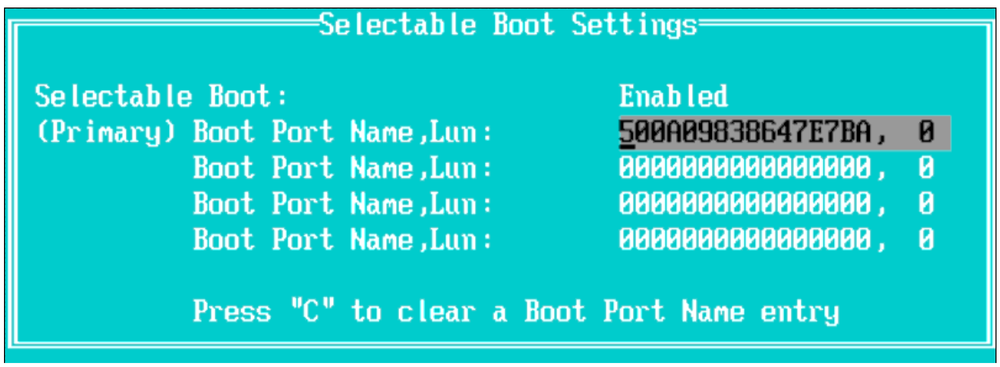


Figure 10-29 Selecting the (Primary) Boot Port Name

10. The available Fibre Channel devices appear, as shown in Figure 10-30. Select the boot LUN **0** from the list of devices and press Enter.

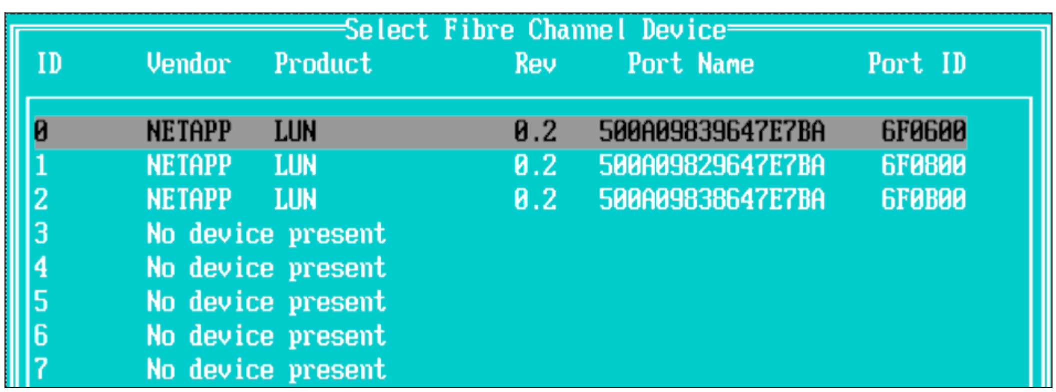


Figure 10-30 Select Fibre Channel Device screen

11. Press Esc to return to the previous screen. Press Esc again and you are prompted to save the configuration settings, as shown in Figure 10-31. Select **Save changes** and press Enter.

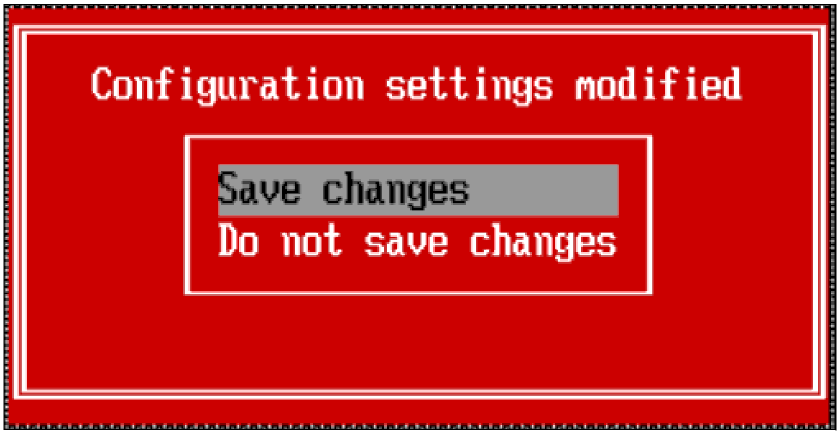


Figure 10-31 Saving the configuration settings

12. The changes are saved and you are returned to the configuration settings. Press Esc and you are prompted to reboot the system, as shown in Figure 10-32.

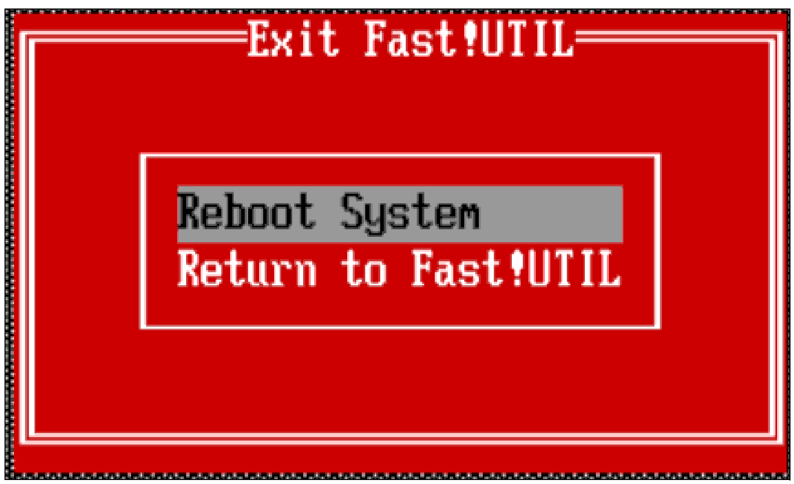


Figure 10-32 Exiting the Fast!UTIL

13. Select **Reboot system** and press Enter.

Configuring the PC BIOS boot order

If your host has an internal disk, you must enter BIOS setup to configure the host to boot from the LUN. You must ensure that the internal disk is not bootable through the BIOS boot order.

The BIOS setup program differs depending on the type of PC BIOS that your host is using. This section shows example procedures for the following BIOS setup programs:

- "IBM BIOS" on page 273
- "Phoenix BIOS 4 Release 6" on page 275

IBM BIOS

BIOS considerations: There may be slight differences within the System BIOS configuration and setup utility depending on the server model and BIOS version being used. Knowledge of BIOS and ROM memory space usage may be required in certain situations. Some older PC architecture limits ROM image memory space to 128 K maximum, and this limit becomes a concern if more devices that require ROM space are desired. If you have many HBAs in your server, you might receive a PCI error allocation message during the boot process. To avoid this, you must disable the boot options in the other HBAs that are not being used for SAN boot installation.

To configure the IBM BIOS setup program:

1. Reboot the host.
2. Press F1 to enter BIOS setup, as shown in Figure 10-33.

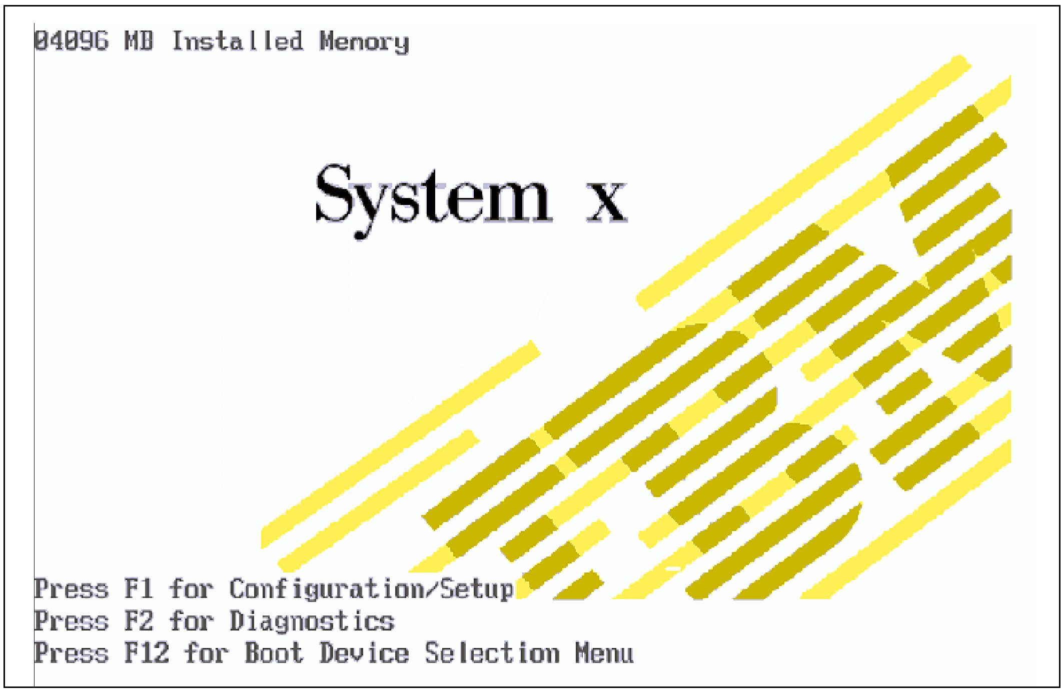


Figure 10-33 System x BIOS Setup screen

3. Select **Start Options**, as shown in Figure 10-34.

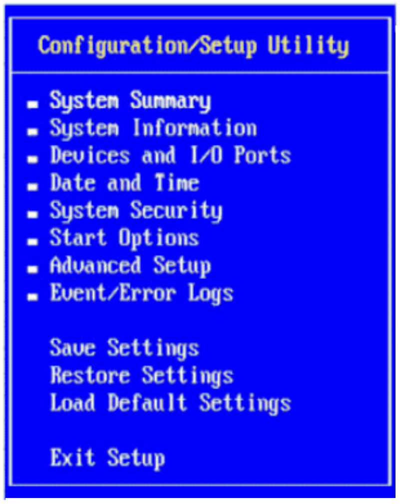


Figure 10-34 Selecting Start Options in Configuration/Setup Utility screen

4. Scroll to the PCI Device Boot Priority option and select the slot in which the HBA is installed, as shown in Figure 10-35.

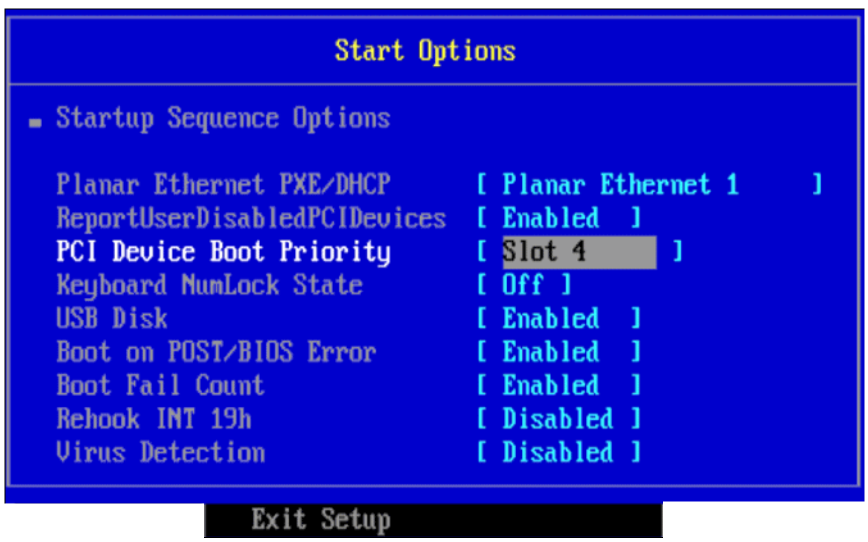


Figure 10-35 Selecting PCI Device Boot Priority in Start Options screen

5. Scroll up to Startup Sequence Options and press Enter. Make sure that the Startup Sequence Option is configured as shown in Figure 10-36.

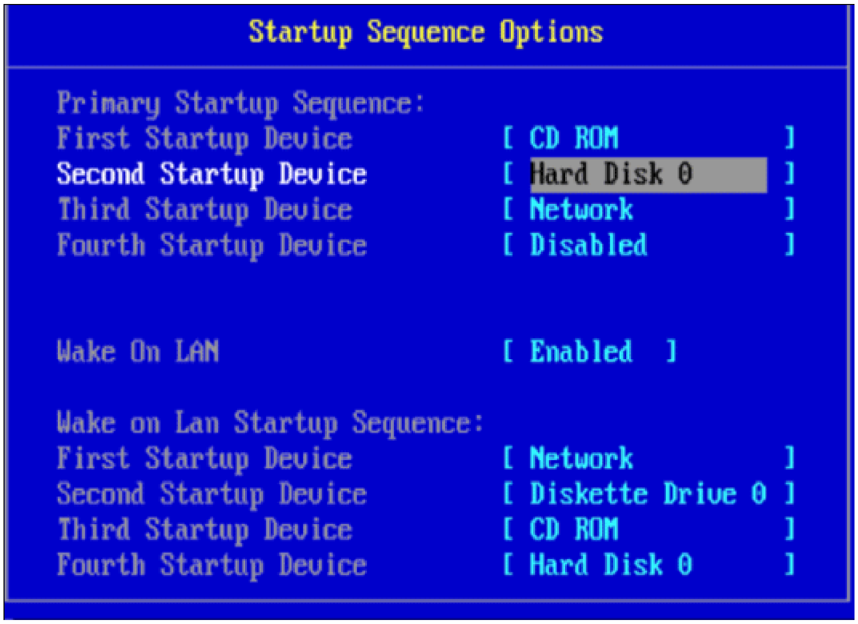


Figure 10-36 Selecting Hard Disk 0 in Startup Sequence Options screen

Phoenix BIOS 4 Release 6

To configure Phoenix BIOS to boot from the Emulex HBA:

1. Reboot the host.
2. Press F2 to enter BIOS setup.
3. Navigate to the **Boot** tab.
4. The Boot tab lists the boot device order. Ensure that the HBA is configured as the first boot device.
5. Select **Hard Drive**.
6. Configure the LUN as the first boot device.

10.2.5 Windows 2003 Enterprise SP2 installation

This section describes installation procedures for Windows 2003 Enterprise SP2.

Copying the SAN boot drivers

When you boot from a LUN, you must ensure that the operating system on the LUN has the required HBA driver for booting from a LUN. You must download these drivers from the Qlogic or Emulex website.

During the Windows 2003 installation, you must install the driver as a third-party SCSI array driver from a floppy disk.

Perform the following steps:

1. Download the Emulex or Qlogic driver for Windows 2003:
   - Emulex
     i. Go to:

        http://emulex.com/ts/downloads/windows/windows.htm

     ii. Download the Storport Miniport driver.
   - Qlogic
     i. Go to:

        http://driverdownloads.qlogic.com/QLogicDriverDownloads_UI/default.aspx

     ii. Click the appropriate HBA.

     iii. Click **Windows Server 2003** and download the STOR Miniport Microsoft Certified Boot from the SAN Driver Package.
2. Copy the driver files to a floppy disk.

Installing Windows 2003 Enterprise SP2

To install Windows 2003 on the LUN:

1. Insert the Windows 2003 CD-ROM and reboot the host.

   A message displays indicating the HBA BIOS is installed along with the boot LUN. See Example 17-1.

   Example 10-1 Message

   ```
   LUN: 00 NETAPP LUN
   BIOS is installed successfully!
   ```

   > **Note:** If the message does not display, do not continue installing Windows. Check to ensure that the LUN is created and mapped, and that the target HBA is in the correct mode for directly connected hosts. Also, ensure that the WWPN for the HBA is the same WWPN that you entered when creating the igroup.
   >
   > If the LUN appears but the message indicates that the BIOS is not installed, reboot and enable the BIOS.

2. When prompted, press any key to boot from the CD-ROM.
3. When prompted, press F6 to install a third-party SCSI array driver.
4. Insert the HBA driver floppy disk that you created previously when the following message is displayed:

   ```
   Setup could not determine the type of one or more mass storage devices
   installed in your system, or you have chosen to manually specify an adapter.
   ```

5. Press S to continue.
6. From the list of HBAs, select the supported HBA that you are using and press Enter. The driver for the selected HBA is configured in the Windows operating system.
7. Follow the prompts to set up the Windows operating system. When prompted, set up the Windows operating system in a partition formatted with NTFS.

8. The host system reboots and then prompts you to complete the server setup process as you normally would do. The rest of the Windows installation is the same as a normal installation.

Note: After you have successfully installed Windows 2003 you must add the remaining WWPN for all additional HBAs to the igroup and install the FCP Windows Host Utilities. See Chapter 5, "Multipathing with the IBM System Storage N series" on page 145, for detailed information about installing the FCP Host Utilities in a Windows System.

Current limitations to Windows boot from SAN

There are a number of advanced scenarios that are not currently possible in Windows boot from SAN environments, specifically:

- No shared boot images: Windows servers cannot currently share a boot image. Each server requires its own dedicated LUN to boot.
- Mass deployment of boot images requires Automated Deployment System (ADS): Windows does not currently support mass distribution of boot images. While cloning of boot images could help here, Windows does not have the tools for distribution of these images. In enterprise configurations, however, Windows ADS can help.
- Lack of standardized assignment of LUN 0 to controller: Some vendors' storage adapters automatically assign logical unit numbers (LUNs). Others require that the storage administrator explicitly define the numbers. With parallel SCSI, the boot LUN is LUN 0 by default.
- Fibre Channel configurations must adhere to SCSI-3 storage standards. In correctly configured arrays, LUN 0 is assigned to the controller (not to a disk device) and is accessible to all servers. This LUN 0 assignment is part of the SCSI-3 standard, since many operating systems do not boot unless the controller is correctly assigned as LUN 0. Correctly assigning LUN 0 to the controller allows it to assume the critical role in discovering and reporting a list of all other LUNs available through that adapter. In Windows, these LUNs are reported back to the kernel in response to the SCSI REPORT LUNS command.

Unfortunately, not all vendor storage arrays comply with the standard of assigning LUN 0 to the controller. Failure to comply with that standard means that the boot process may not proceed correctly. In some cases, even with LUN 0 correctly assigned, the boot LUN cannot be found, and the operating system fails to load. In the following cases (without HBA LUN remapping), the kernel finds LUN 0, but may not be successful in enumerating the LUNs correctly.

10.2.6 Windows 2008 Enterprise installation

The Windows 2008 server can be installed in two types:

- Full installation
- Core installation

Full installation supports GUI and no roles are installed by default (for example, print, file, DHCP, and so on). On the other hand, core installation does not support any GUI. It supports only command line and Windows power shell, which is why it does not require higher memory and disk.

There are few boot configuration changes introduced in the Windows 2008 server. The major change is that boot configuration data (BCD) stores contain boot configuration parameters and control how the operating system is started in Microsoft Windows Server 2008 operating systems. These parameters were previously in the Boot.ini file (in BIOS-based operating systems) or in the non-volatile RAM (NVRAM) entries (in Extensible Firmware Interface-based operating systems). You can use the Bcdedit.exe command line tool to modify the Windows code that runs in the pre-operating system environment by adding, deleting, editing, and appending entries in the BCD store. Bcdedit.exe is located in the \Windows\System32 directory of the Windows 2008 active partition. BCD was created to provide an improved mechanism for describing boot configuration data. With the development of new firmware models (for example, the Extensible Firmware Interface (EFI)), an extensible and interoperable interface was required to abstract the underlying firmware.

Windows Server 2008 R2 supports the ability to boot from a SAN, which eliminates the need for local hard disks in the individual server computers. In addition, performance for accessing storage on SANs has been greatly improved. Figure 10-37 shows how booting from a SAN can dramatically reduce the number of hard disks and decrease power consumption as a result.

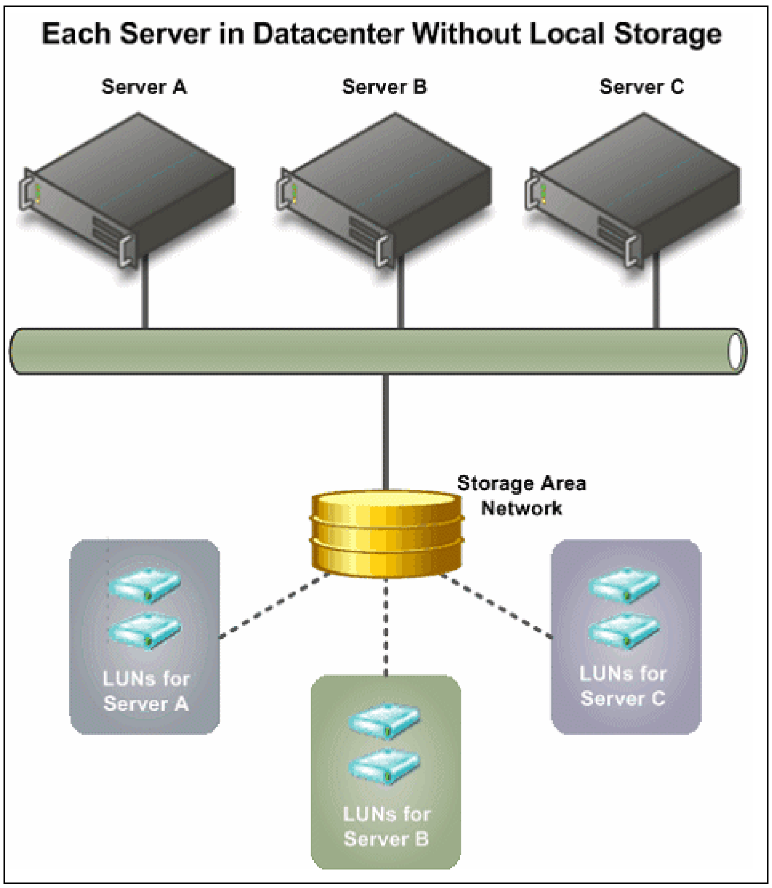


Figure 10-37 Centralizing storage to reduce power consumption

To install the Windows Server 2008 full installation option:

1. Insert the appropriate Windows Server 2008 installation media into your DVD drive. Reboot the server, as shown in Figure 10-38.

Figure 10-38 Rebooting the server

2. Select an installation language, regional options, and keyboard input, and click **Next**, as shown in Figure 3 on page 280.

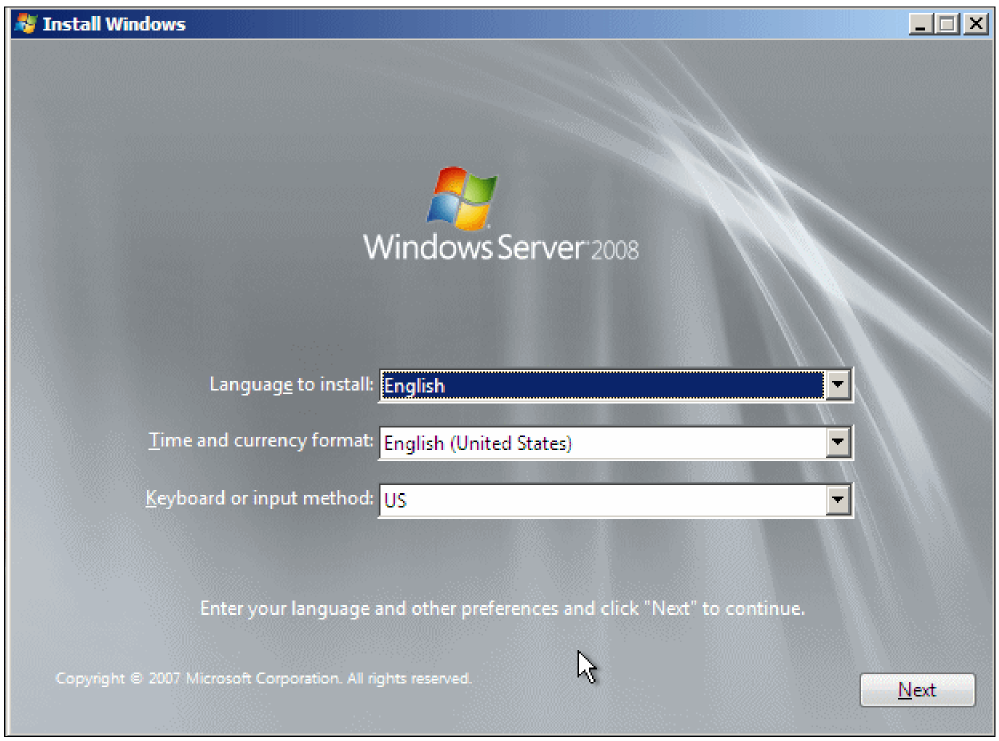


Figure 10-39 Selecting language to install, regional options, and keyboard input

3. Click **Install now** to begin the installation process, as shown in Figure 10-40.

Figure 10-40 Selecting Install now

4. Enter the product key and click **Next**, as shown in Figure 10-41.

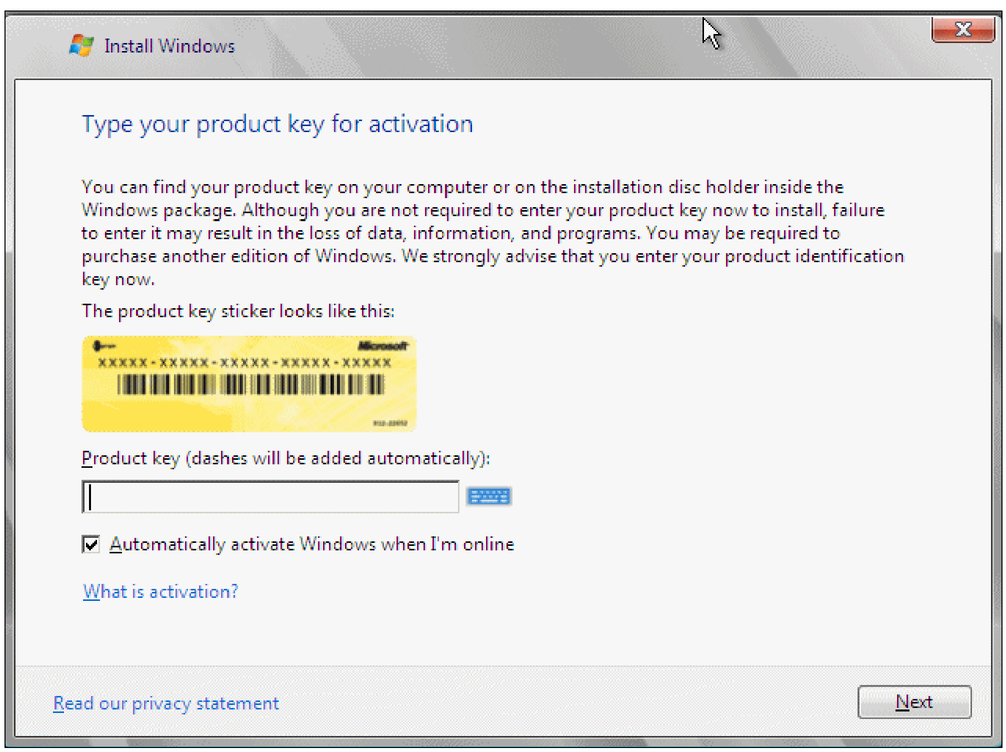


Figure 10-41 Entering the product key

5. Select the **I accept the license terms** check box and click **Next**, as shown in Figure 10-42.

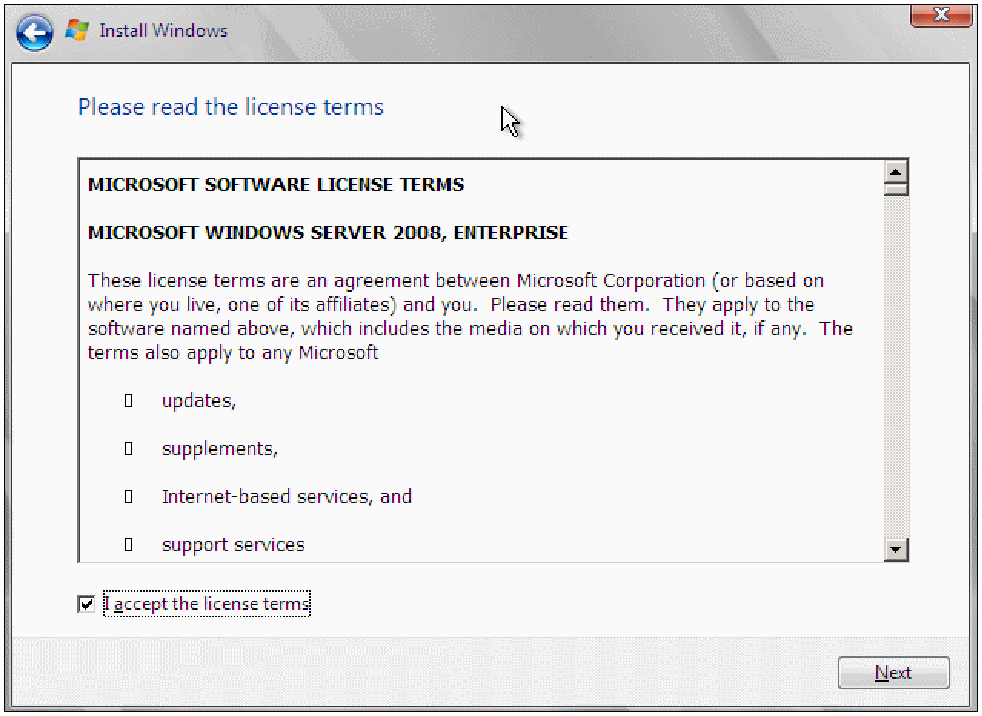


Figure 10-42 Accepting the license terms

6. Click **Custom (advanced)**, as shown in Figure 10-43.

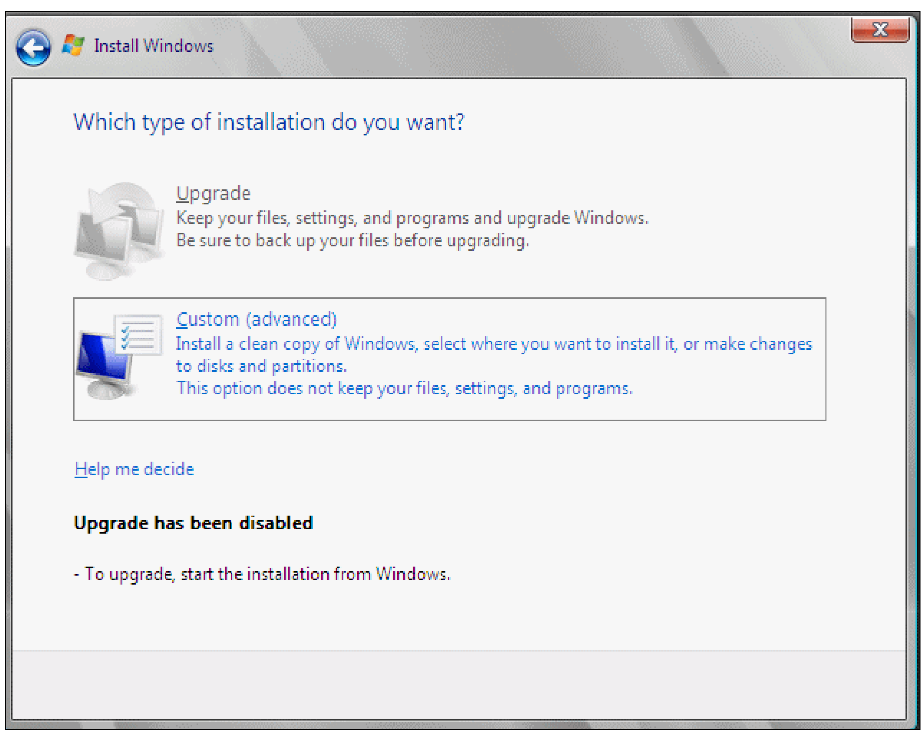


Figure 10-43 Selecting Custom installation option

7. If the window shown in Figure 10-44 does not show any hard drives, or if you prefer to install the HBA device driver now, click **Load Driver**.

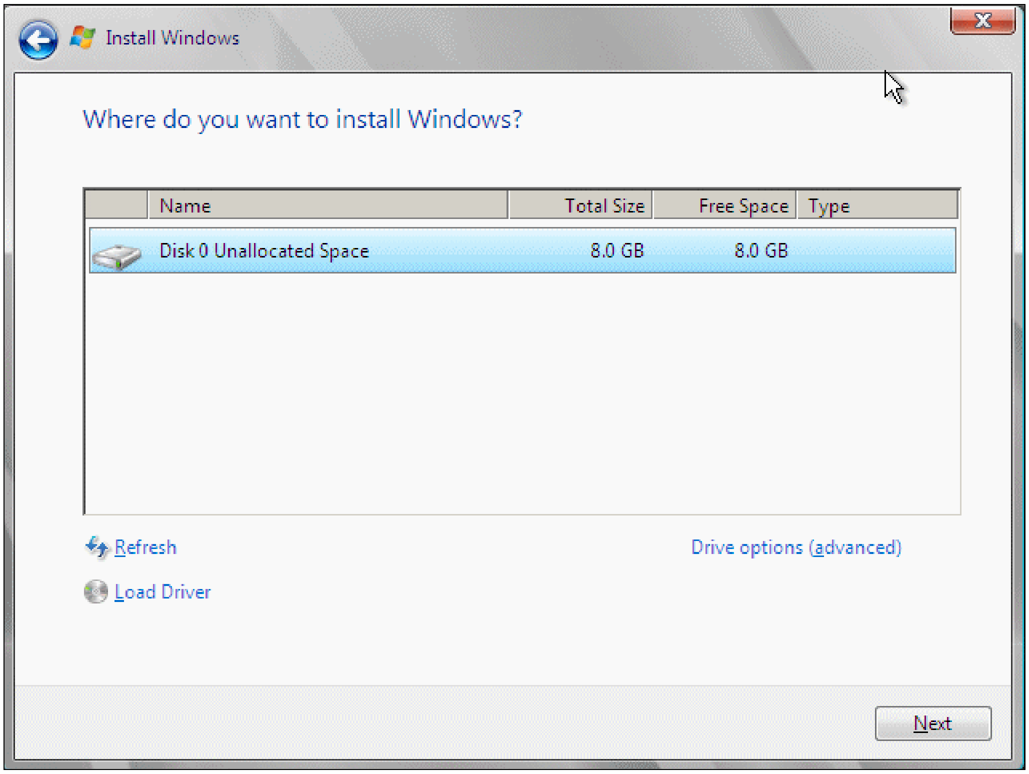


Figure 10-44 Where do you want to install Windows? screen

8. As shown in Figure 10-45, insert appropriate media containing the HBA device driver files and click **Browse**.

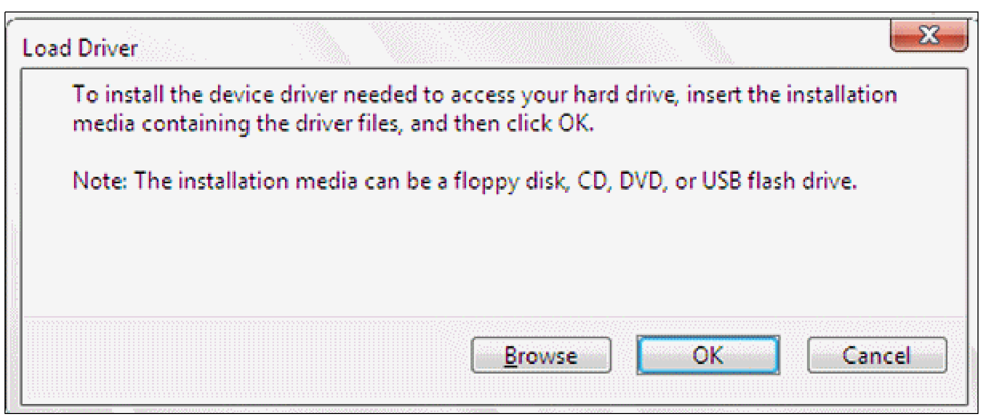


Figure 10-45 Load Driver screen

Click **OK** → **Next**.

9. Click **Next** again to leave the "Windows creates the partition automatically" window or click **Drive options (advanced)** to create the partition. Then click **Next** to start the installation process, as shown in Figure 10-46.

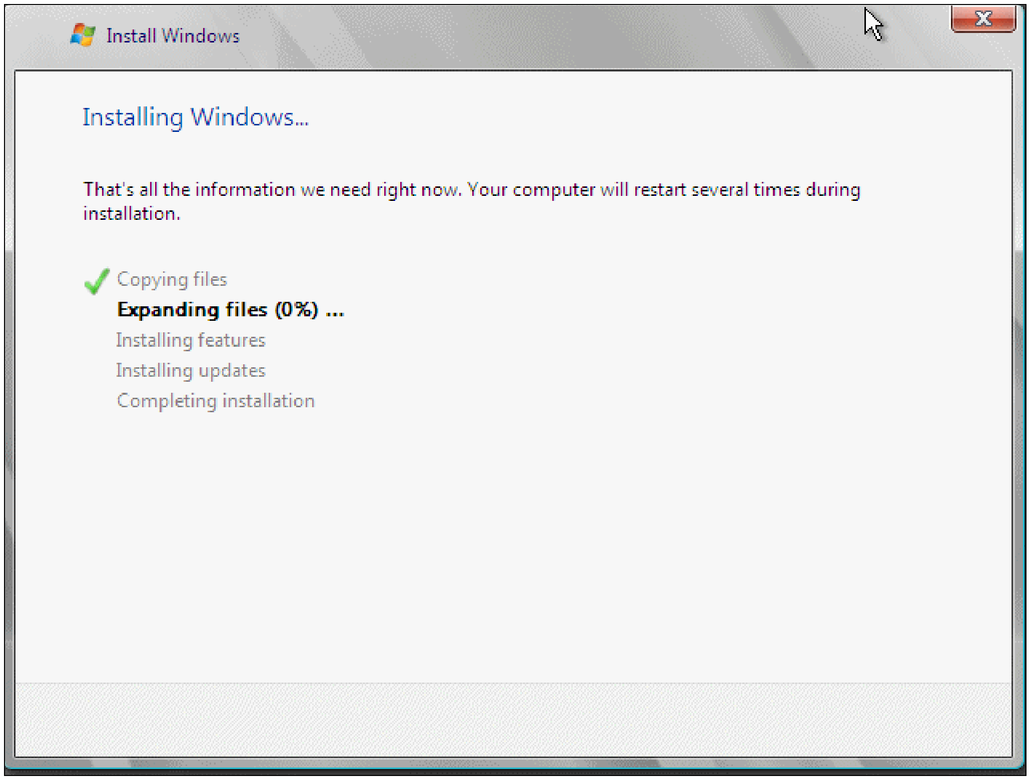


Figure 10-46 Installing Windows screen

10. When Windows Server 2008 Setup has completed installation, the server automatically restarts.
11. After Windows Server 2008 restarts, you will be prompted to change the administrator password before you can log on.
12. After you are logged on as the administrator, a configuration wizard window is displayed. Use the wizard for naming and basic networking setup.
13. Use the Microsoft Server 2008 "roles and features" functions to set up the server to your specific needs.

Note: After you have successfully installed Windows 2008 you must add the remaining WWPN for all additional HBAs to the igroup and install the FCP Windows Host Utilities. See Chapter 5, "Multipathing with the IBM System Storage N series" on page 145, for detailed information about installing the FCP Host Utilities in a Windows system.

10.2.7 Red Hat Enterprise Linux 5.2 installation

This section shows how to install Red Hat Enterprise Linux 5.2 boot from SAN with an IBM System x server.

Note: Always check hardware and software, including firmware and operating system compatibility, before you implement SAN boot in different hardware or software environments.

Linux boot process

This section is an overview of the Linux boot process in an x86 environment. In general, the boot process is as shown in Figure 10-47.

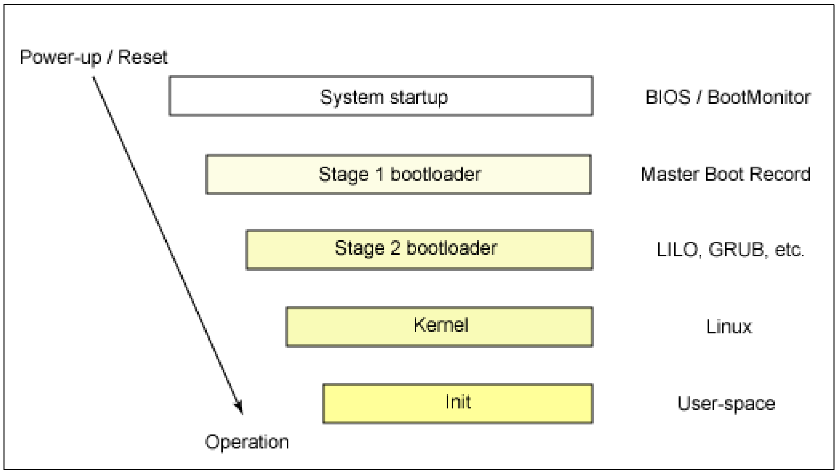


Figure 10-47 Linux boot process

System BIOS

The process starts when we power up or reset our System x machine. The processor executes the Basic Input/Output System (BIOS) code, which then does a power-on self test (POST) to check and initialize the hardware. It then locates a valid device to boot the system.

Boot loader

If a boot device is found, the BIOS loads the first stage boot loader stored in the master boot record (MBR) (the first 512 bytes of the bootable device) into memory. This first stage boot loader is then executed to locate and load into memory the second stage boot loader. (Boot loaders are in two stages because of the very limited size of the MBR.) In an x86 machine, the second stage boot loader can be the LInux Loader (LILO) or the GRand Unified Boot loader (GRUB). Once loaded, it presents a list of available kernels to boot.

OS kernel

Once a kernel has been selected, the second stage boot loader locates the kernel binary and loads into memory the initial RAM disk image. The kernel then checks and configures hardware and peripherals, and decompresses the initial RAM disk image into load drivers and modules necessary to completely boot the system. It also mounts the root device.

Continue system start

Once the kernel and its modules are loaded, a high-level system initialization is performed by the /sbin/init program, which will be the parent process of all other subsequent start-up processes. /sbin/init executes /etc/rc.d/rc.sysinit and its corresponding scripts. This is followed by executing /etc/inittab, /etc/rc.d/init.d/functions, and the appropriate rc directory as configured in /etc/inittab. For example, if the default runlevel in /etc/inittab is configured as runlevel 5, /sbin/init executes scripts under the /etc/rc.d/rc5.d/ directory.

Install Red Hat Enterprise Linux 5.2

The installation process explained here assumes that the server does not have any special hardware (SCSI card or HBA) that would require a specific Linux driver not included on the installer CD. If you have a device driver diskette for a specific device driver to be loaded during the installation process, you must type `Linux dd` at the installation boot prompt before the installation wizard is loaded.

> **Note:** RHEL5 now has the capability to detect, create, and install to dm-multipath devices during install. To enable this feature, add the parameter `mpath` to the kernel boot line. At the initial Linux install screen type `linux mpath` and press Enter to start the Red Hat installation.

The installation process is similar to local disk installation. In summary, the general steps required to set up a Linux SAN boot are:

1. Insert the Linux install CD and reboot the host. During the installation, you are able see the LUN and install the OS on it.
2. Click **Next** and follow the installation wizard as you normally would do with a local disk installation.

> **Important:** After you have successfully installed Red Hat Enterprise Linux 5.2, you must add the remaining WWPN for all additional HBAs to the igroup, and install the FCP Linux Host Utilities. See Chapter 5, “Multipathing with the IBM System Storage N series” on page 145, for detailed information about installing the FCP Host Utilities in a Linux system.

Align partitions on Linux with IBM LUNs for best performance

IBM LUNs connected by way of a block protocol (for example, iSCSI, FCP) to Linux hosts using partitions might require special partition alignment for best performance. See detailed information about this issue at:

http://www-01.ibm.com/support/docview.wss?uid=ssg1S1002716&rs=573

A

Getting started

This appendix provides information to help you document, install, and set up your IBM System Storage N series storage system.

It covers the following topics:

- Preinstallation planning
- Collecting documents
- Start with the hardware
- Power on N series
- Data ONTAP update (up to release 7.x)
- Obtain Data ONTAP software from IBM NAS website
- Install Data ONTAP system files
- Download Data ONTAP to the storage system
- Network setup using console
- Change IP address
- DNS setup
- Data ONTAP update (up to release 8.0-7 mode)

Preinstallation planning

Successful installation of the IBM System Storage N series storage system requires careful planning. This section provides information about this preparation.

Collecting documents

N series product documentation is available at:

http://www.ibm.com/systems/support/storage/nas

Collect all documents needed for installing new storage systems:

1. N series information requires unregistered users to complete the one-time registration and then log into the site using their registered IBM Identity with each visit. Detailed step-by-step instructions for N series registration can be downloaded from:

   http://www.ibm.com/systems/support/resources/jct01004c_systems_support_storage_register_NseriesStepByStep.pdf

2. Prepare the site and requirements of your system. Refer to *IBM System Storage N series Introduction and Planning Guide,* GA32-0543, to plan for the physical environment where the equipment will operate. This planning step includes the physical space, electrical, temperature, humidity, altitude, air flow, service clearance, and similar requirements. Also check this document for rack, power supplies, power requirements, and thermal considerations.

3. Use the hardware guide to install the N series storage system:

   – *Installation and Setup instructions for N series storage system,* GC26-7784
   – *Hardware and Service Guide for N series storage system*, GC26-7785

   There are different cabling instructions for single-node and dual-node (cluster) configuration.

   Note: Refer to the *Cluster Installation and Administration Guide or Active/Active Configuration Guide* GC26-7964, for your version of Data ONTAP for more information about clustering.

4. Refer to *IBM System Storage N series Data ONTAP Software Setup Guide,* GC27-2206, for information about how to set up the N series Data ONTAP. This document describes how to set up and configure new storage systems that run Data ONTAP software.

   To ensure interoperability of third-party hardware, software, and the N series storage system, refer to the appropriate Interoperability Matrix, which can be located at the following website:

   http://www-03.ibm.com/systems/storage/nas/interophome.html

Initial worksheet for setting up the nodes

For first-time installation on any of the N series models, Data ONTAP presents a series of questions regarding the storage system setup. The worksheets provided here make sure that you have the answers to these questions available before installation.

Initial worksheet

Table A-1 provides an initial worksheet for setting up the node.

Table A-1 Initial worksheet

Types of information		**Your values**
Storage system	Host name If the storage system is licensed for the Network File System (NFS) protocol, the name can be no longer than 32 characters. If the storage system is licensed for the Common Internet File System (CIFS) protocol, the name can be no longer than 15 characters.	
	Password.	
	Time zone.	See “Time zones” on page 292 for a list of valid time zones.
	Storage system location The text that you enter during the storage system setup process is recorded in the SNMP location information. Use a description that identifies where to find your storage system (for example, lab 5, row 7, rack B).	
	Language used for multiprotocol storage systems.	See “Supported languages” on page 299.
Administration host A client computer that is allowed to access the storage system through a Telnet client or through the Remote Shell protocol.	Host name.	
	IP address.	
Virtual interfaces The virtual network interface information must be identical on both storage systems in an active/active pair.	Link names (physical interface names such as e0, e0a, e5a, or e9b)	The default is set to no for most installations.
	Number of links (number of physical interfaces to include in the vif).	
	Name of virtual interface (name of vif, such as vif0).	

Ethernet interfaces	IP address.		
	Subnet mask.		
	Partner IP address If your storage system is licensed for controller takeover, record the interface name or IP address belonging to the partner that this interface should take over during an active/active configuration takeover.		The default is set to no for most installations.
	Media type (network type) (100tx-fd, tp-fd, 100tx, tp, auto (10/100/1000)).		The default is set to auto.
	Are jumbo frames supported?		The default is set to no.
	MTU size for jumbo frames.		
	Flow control (none, receive, send, full).		The default is set to full.
E0M Ethernet interface if available	IP address.		
	Subnet mask.		
	Partner IP address.		The default is set to no for most installations.
	Flow control (none, receive, send, full).		The default is set to full.
Router (if used)	Gateway name.		
	IP address.		
Would you like to continue setup through Web interface? You do this through the Setup Wizard.			The default is set to no.
DNS	Domain name.		
	Server address 1, 2, 3.		
NIS	Domain name.		
	Server address 1, 2, 3.		
Customer contact	Primary.	Name.	
		Phone.	
		Alternate phone.	
		E-mail address or IBM Web ID.	
	Secondary.	Name.	
		Phone.	
		Alternate phone.	
		E-mail address or IBM Web ID.	

Machine location	Business name.		
	Address.		
	City.		
	State.		
	Country code. (Value must be two upper case letters.)		See “Country codes” on page 299.
	Postal code.		
CIFS	Windows domain.		
	WINS servers.		
	Multiprotocol or NTFS only storage system		
	Should CIFS create default etc/passwd and etc/group files? Enter `y` here if you have a multiprotocol environment. Default UNIX accounts are created, which are used when performing user mapping. As an alternative to storing this information in a local file, the generic user accounts can be stored in the NIS or LDAP databases. However, when generic accounts are stored in the local passwd file, mapping of a Windows user to a generic UNIX user and mapping of a generic UNIX user to a Windows user work better than when generic accounts are stored in NIS or LDAP.		
	NIS group caching NIS group caching is used when access is requested to data with UNIX security style. UNIX file and directory style permissions of rwxrwxrwx are used to determine access for both Windows and UNIX clients. This security style uses UNIX group information.	Enable?	
		Hours to update the cache.	
	CIFS server name if different from default.		

<table>
<tr><td rowspan="8"></td><td colspan="2">User authentication style:
1. Active Directory authentication (Active Directory domains only).
2. Windows NT® 4 authentication (Windows NT or Active Directory domains only).
3. Windows workgroup authentication using Storage systems local user accounts.
4. etc/password and/or NIS/LDAP authentication.</td><td></td></tr>
<tr><td rowspan="6">Windows Active Directory Domain.</td><td>Windows Domain Name.</td><td></td></tr>
<tr><td>Time server names or IP addresses.</td><td></td></tr>
<tr><td>Windows user name.</td><td></td></tr>
<tr><td>Windows user password.</td><td></td></tr>
<tr><td>Local administrator name.</td><td></td></tr>
<tr><td>Local administrator password.</td><td></td></tr>
<tr><td colspan="2">CIFS administrator or group
You can specify an additional user or group to be added to the storage system's local BUILTIN\Administrators group, thus giving them administrative privileges as well.</td><td></td></tr>
</table>

Time zones

You must select a valid time zone or alias from the lists provided, record it in the configuration worksheet, and enter the value at the setup prompt.

You can enter a geographic region or you can use one of the following aliases (Table A-2 through Table A-18 on page 298) to represent its corresponding time zone descriptions.

Table A-2 Africa

Africa/Abidjan	Africa/Dar_es_Salaam	Africa/Malabo
Africa/Accra	Africa/Djibouti	Africa/Maputo
Africa/Addis_Ababa	Africa/Douala	Africa/Maseru
Africa/Algiers	Africa/El_Aaiun	Africa/Mbabane
Africa/Asmara	Africa/Freetown	Africa/Mogadishu
Africa/Asmera	Africa/Gaborone	Africa/Monrovia
Africa/Bamako	Africa/Harare	Africa/Nairobi
Africa/Bangui	Africa/Johannesburg	Africa/Ndjamena
Africa/Banjul	Africa/Kampala	Africa/Niamey
Africa/Bissau	Africa/Khartoum	Africa/Nouakchott
Africa/Blantyre	Africa/Kigali	Africa/Ouagadougou
Africa/Brazzaville	Africa/Kinshasa	Africa/Porto-Novo
Africa/Bujumbura	Africa/Lagos	Africa/Sao_Tome
Africa/Cairo	Africa/Libreville	Africa/Timbuktu

Africa/Casablanca	Africa/Lome	Africa/Tripoli
Africa/Ceuta	Africa/Luanda	Africa/Tunis
Africa/Conakry	Africa/Lubumbashi	Africa/Windhoek
Africa/Dakar	Africa/Lusaka	

Table A-3 America

America/Adak	America/Eirunepe	America/Montevideo
America/Anchorage	America/El_Salvador	America/Montreal
America/Anguilla	America/Ensenada	America/Montserrat
America/Antigua	America/Fort_Wayne	America/Nassau
America/Araguaina	America/Fortaleza	America/New_York
America/Argentina/Buenos_Aires	America/Glace_Bay	America/Nipigon
America/Argentina/Catamarca	America/Godthab	America/Nome
America/Argentina/ComodRivadavia	America/Goose_Bay	America/Noronha
America/Argentina/Cordoba	America/Grand_Turk	America/North_Dakota/Center
America/Argentina/Jujuy	America/Grenada	America/North_Dakota/New_Salem
America/Argentina/La_Rioja	America/Guadeloupe	America/Panama
America/Argentina/Mendoza	America/Guatemala	America/Pangnirtung
America/Argentina/Rio_Gallegos	America/Guayaquil	America/Paramaribo
America/Argentina/San_Juan	America/Guyana	America/Phoenix
America/Argentina/San_Luis	America/Halifax	America/Port-au-Prince
America/Argentina/Tucuman	America/Havana	America/Port_of_Spain
America/Argentina/Ushuaia	America/Hermosillo	America/Porto_Acre
America/Aruba	America/Indiana/Indianapolis	America/Porto_Velho
America/Asuncion	America/Indiana/Knox	America/Puerto_Rico
America/Atikokan	America/Indiana/Marengo	America/Rainy_River
America/Atka	America/Indiana/Petersburg	America/Rankin_Inlet
America/Bahia	America/Indiana/Tell_City	America/Recife
America/Barbados	America/Indiana/Vevay	America/Regina
America/Belem	America/Indiana/Vincennes	America/Resolute
America/Belize	America/Indiana/Winamac	America/Rio_Branco
America/Blanc-Sablon	America/Indianapolis	America/Rosario
America/Boa_Vista	America/Inuvik	America/Santiago
America/Bogota	America/Iqaluit	America/Santo_Domingo
America/Boise	America/Jamaica	America/Sao_Paulo

America/Buenos_Aires	America/Jujuy	America/Scoresbysund
America/Cambridge_Bay	America/Juneau	America/Shiprock
America/Campo_Grande	America/Kentucky/Louisville	America/St_Barthelemy
America/Cancun	America/Kentucky/Monticello	America/St_Johns
America/Caracas	America/Knox_IN	America/St_Kitts
America/Catamarca	America/La_Paz	America/St_Lucia
America/Cayenne	America/Lima	America/St_Thomas
America/Cayman	America/Los_Angeles	America/St_Vincent
America/Chicago	America/Louisville	America/Swift_Current
America/Chihuahua	America/Maceio	America/Tegucigalpa
America/Coral_Harbour	America/Managua	America/Thule
America/Cordoba	America/Manaus	America/Thunder_Bay
America/Costa_Rica	America/Marigot	America/Tijuana
America/Cuiaba	America/Martinique	America/Toronto
America/Curacao	America/Mazatlan	America/Tortola
America/Danmarkshavn	America/Mendoza	America/Vancouver
America/Dawson	America/Menominee	America/Virgin
America/Dawson_Creek	America/Merida	America/Whitehorse
America/Denver	America/Mexico_City	America/Winnipeg
America/Detroit	America/Miquelon	America/Yakutat
America/Dominica	America/Moncton	America/Yellowknife
America/Edmonton	America/Monterrey	

Table A-4 Antarctica

Antarctica/Casey	Antarctica/McMurdo	Antarctica/Syowa
Antarctica/Davis	Antarctica/Palmer	Antarctica/Vostok
Antarctica/DumontDUrville	Antarctica/Rothera	
Antarctica/Mawson	Antarctica/South_Pole	

Table A-5 Asia

Asia/Aden	Asia/Hong_Kong	Asia/Pyongyang
Asia/Almaty	Asia/Hovd	Asia/Qatar
Asia/Amman	Asia/Irkutsk	Asia/Qyzylorda
Asia/Anadyr	Asia/Istanbul	Asia/Rangoon
Asia/Aqtau	Asia/Jakarta	Asia/Riyadh
Asia/Aqtobe	Asia/Jayapura	Asia/Saigon

Asia/Ashgabat	Asia/Jerusalem	Asia/Sakhalin
Asia/Ashkhabad	Asia/Kabul	Asia/Samarkand
Asia/Baghdad	Asia/Kamchatka	Asia/Seoul
Asia/Bahrain	Asia/Karachi	Asia/Shanghai
Asia/Baku	Asia/Kashgar	Asia/Singapore
Asia/Bangkok	Asia/Katmandu	Asia/Taipei
Asia/Beirut	Asia/Kolkata	Asia/Tashkent
Asia/Bishkek	Asia/Krasnoyarsk	Asia/Tbilisi
Asia/Brunei	Asia/Kuala_Lumpur	Asia/Tehran
Asia/Calcutta	Asia/Kuching	Asia/Tel_Aviv
Asia/Choibalsan	Asia/Kuwait	Asia/Thimbu
Asia/Chongqing	Asia/Macao	Asia/Thimphu
Asia/Chungking	Asia/Macau	Asia/Tokyo
Asia/Colombo	Asia/Magadan	Asia/Ujung_Pandang
Asia/Dacca	Asia/Makassar	Asia/Ulaanbaatar
Asia/Damascus	Asia/Manila	Asia/Ulan_Bator
Asia/Dhaka	Asia/Muscat	Asia/Urumqi
Asia/Dili	Asia/Nicosia	Asia/Vientiane
Asia/Dubai	Asia/Novosibirsk	Asia/Vladivostok
Asia/Dushanbe	Asia/Omsk	Asia/Yakutsk
Asia/Gaza	Asia/Oral	Asia/Yekaterinburg
Asia/Harbin	Asia/Phnom_Penh	Asia/Yerevan
Asia/Ho_Chi_Minh	Asia/Pontianak	

Table A-6 Atlantic

Atlantic/Azores	Atlantic/Faeroe	Atlantic/Reykjavik
Atlantic/Bermuda	Atlantic/Faroe	Atlantic/South_Georgia
Atlantic/Canary	Atlantic/Jan_Mayen	Atlantic/St_Helena
Atlantic/Cape_Verde	Atlantic/Madeira	Atlantic/Stanley

Table A-7 Australia

Australia/ACT	Australia/Hobart	Australia/Queensland
Australia/Adelaide	Australia/LHI	Australia/South
Australia/Brisbane	Australia/Lindeman	Australia/Sydney
Australia/Broken_Hill	Australia/Lord_Howe	Australia/Tasmania
Australia/Canberra	Australia/Melbourne	Australia/Victoria
Australia/Currie	Australia/NSW	Australia/West
Australia/Darwin	Australia/North	Australia/Yancowinna
Australia/Eucla	Australia/Perth	

Table A-8 Brazil

Brazil/Acre	Brazil/East
Brazil/DeNoronha	Brazil/West

Table A-9 Canada

Canada/Atlantic	Canada/Eastern	Canada/Pacific
Canada/Central	Canada/Mountain	Canada/Saskatchewan
Canada/East-Saskatchewan	Canada/Newfoundland	Canada/Yukon

Table A-10 Chile

Chile/Continental	Chile/EasterIsland

Table A-11 Europe

Europe/Amsterdam	Europe/Kiev	Europe/Sarajevo
Europe/Andorra	Europe/Lisbon	Europe/Simferopol
Europe/Athens	Europe/Ljubljana	Europe/Skopje
Europe/Belfast	Europe/London	Europe/Sofia
Europe/Belgrade	Europe/Luxembourg	Europe/Stockholm
Europe/Berlin	Europe/Madrid	Europe/Tallinn
Europe/Bratislava	Europe/Malta	Europe/Tirane
Europe/Brussels	Europe/Mariehamn	Europe/Tiraspol
Europe/Bucharest	Europe/Minsk	Europe/Uzhgorod
Europe/Budapest	Europe/Monaco	Europe/Vaduz
Europe/Chisinau	Europe/Moscow	Europe/Vatican
Europe/Copenhagen	Europe/Nicosia	Europe/Vienna
Europe/Dublin	Europe/Oslo	Europe/Vilnius
Europe/Gibraltar	Europe/Paris	Europe/Volgograd
Europe/Guernsey	Europe/Podgorica	Europe/Warsaw

Europe/Helsinki	Europe/Prague	Europe/Zagreb
Europe/Isle_of_Man	Europe/Riga	Europe/Zaporozhye
Europe/Istanbul	Europe/Rome	Europe/Zurich
Europe/Jersey	Europe/Samara	
Europe/Kaliningrad	Europe/San_Marino	

Table A-12 Indian (Indian Ocean)

Indian/Antananarivo	Indian/Comoro	Indian/Mauritius
Indian/Chagos	Indian/Kerguelen	Indian/Mayotte
Indian/Christmas	Indian/Mahe	Indian/Reunion
Indian/Cocos	Indian/Maldives	

Table A-13 Mexico

Mexico/BajaNorte	Mexico/BajaSur	Mexico/General

Table A-14 Pasific

Pacific/Apia	Pacific/Johnston	Pacific/Ponape
Pacific/Auckland	Pacific/Kiritimati	Pacific/Port_Moresby
Pacific/Chatham	Pacific/Kosrae	Pacific/Rarotonga
Pacific/Easter	Pacific/Kwajalein	Pacific/Saipan
Pacific/Efate	Pacific/Majuro	Pacific/Samoa
Pacific/Enderbury	Pacific/Marquesas	Pacific/Tahiti
Pacific/Fakaofo	Pacific/Midway	Pacific/Tarawa
Pacific/Fiji	Pacific/Nauru	Pacific/Tongatapu
Pacific/Funafuti	Pacific/Niue	Pacific/Truk
Pacific/Galapagos	Pacific/Norfolk	Pacific/Wake
Pacific/Gambier	Pacific/Noumea	Pacific/Wallis
Pacific/Guadalcanal	Pacific/Pago_Pago	Pacific/Yap
Pacific/Guam	Pacific/Palau	
Pacific/Honolulu	Pacific/Pitcairn	

Table A-15 GMT

Greenwich	Etc/GMT+3	Etc/GMT-13
GMT	Etc/GMT+4	Etc/GMT-14
GMT+0	Etc/GMT+5	Etc/GMT-2
GMT-0	Etc/GMT+6	Etc/GMT-3
GMT0	Etc/GMT+7	Etc/GMT-4
Etc/GMT	Etc/GMT+8	Etc/GMT-5

Etc/GMT+0	Etc/GMT+9	Etc/GMT-6
Etc/GMT+1	Etc/GMT-0	Etc/GMT-7
Etc/GMT+10	Etc/GMT-1	Etc/GMT-8
Etc/GMT+11	Etc/GMT-10	Etc/GMT-9
Etc/GMT+12	Etc/GMT-11	Etc/GMT0
Etc/GMT+2	Etc/GMT-12	Etc/Greenwich

Table A-16 ETC

Etc/Greenwich	Etc/UTC	Etc/Zulu
Etc/UCT	Etc/Universal	

Table A-17 US

US/Alaska	US/Eastern	US/Pacific
US/Aleutian	US/Hawaii	US/Pacific-New
US/Arizona	US/Indiana-Starke	US/Samoa
US/Central	US/Michigan	
US/East-Indiana	US/Mountain	

Table A-18 Miscellaneous

Arctic/Longyearbyen	Iceland	PST8PDT
CET	Iran	Poland
CST6CDT	Israel	Portugal
Cuba	Jamaica	ROC
EET	Japan	ROK
EST	Kwajalein	Singapore
EST5EDT	Libya	Turkey
Egypt	MET	UCT
Eire	MST	UTC
Factory	MST7MDT	Universal
GB	NZ	W-SU
GB-Eire	NZ-CHAT	WET
HST	Navajo	Zulu
Hongkong	PRC	

Supported languages

To use UTF-8 as the NFS character set, append .UTF-8. See Table A-19.

Table A-19 Language abbreviations

Language	Abbreviation	Language	Abbreviation
Arabic	ar	Norwegian	no
Croatian	hr	Polish	pl
Czech	cs	Portuguese	pt
Danish	da	POSIX	C
Dutch	nl	Romanian	ro
English	en	Russian	ru
English (U.S.)	en_US	Simplified Chinese	zh
Finnish	fi	Simplified Chinese (GBK)	zh.GBK
French	fr	Slovak	sk
German	de	Slovenian	sl
Hebrew	he	Spanish	es
Hungarian	hu	Swedish	sv
Italian	it	Traditional Chinese euc-tw	zh_TW
Japanese euc-j	ja	Traditional Chinese Big 5	zh_TW.BIG5
Japanese PCK (sjis)	ja_JP.PCK	Turkish	tr
Korean	ko		

Country codes

Table A-20 provides country codes.

Table A-20 Country codes

Code	Country	Code	Country
AF	Afghanistan	LS	Lesotho
AL	Albania	LR	Liberia
DZ	Algeria	LY	Libyan Arab Jamahiriya (Lybia)
AS	American Samoa	LI	Liechtenstein
AD	Andorra	LT	Lithuania
AO	Angola	LU	Luxembourg
AI	Anguilla	MK	Macedonia
AQ	Antarctia	MG	Madagascar
AG	Antigua and Barbuda	MW	Malawi
AR	Argentina	MY	Malaysia
AM	Armenia	MV	Maldives

AW	Aruba	ML	Mali
AU	Australia	MT	Malta
AT	Austria	MH	Marshall Islands
AZ	Azerbaijan	MQ	Martinique
PT	Azores	MR	Mauritania
BS	Bahamas	MU	Mauritius
BH	Bahrain	YT	Mayotte
ES	Balearic Islands	MX	Mexico
ID	Bali	FM	Micronesia, Federated States of
BD	Bangladesh	US	Midway Island
BB	Barbados	MD	Moldavia
BY	Belarus	MC	Monaco
BE	Belgium	MN	Mongolia
BZ	Belize	MS	Montserrat
BJ	Benin	MA	Morocco
BM	Bermuda	MZ	Mozambique
BT	Bhutan	MM	Myanmar
BO	Bolivia	NA	Namibia
ID	Borneo	NR	Nauru
BA	Bosnia and Herzegovina	NP	Nepal
BW	Botswana	NL	Netherlands
BV	Bouvet Island	AN	Netherlands Antilles
BR	Brazil	NC	New Caledonia
IO	British Indian Ocean Territory	NZ	New Zealand
VG	British Virgin Islands	NI	Nicaragua
BN	Brunei Darussalem	NE	Niger
BG	Bulgaria	NG	Nigeria
BF	Burkina Faso	NU	Niue
BI	Burundi	NF	Norfolk Island
TC	Caicos Islands	MP	Northern Mariana Islands
CM	Cameroon	NO	Norway
CA	Canada	OM	Oman
CV	Cape Verde	PK	Pakistan
KY	Cayman Islands	PW	Palau
CF	Central African Republic	PA	Panama

TD	Chad	PG	Papua New Guinea
CL	Chile	PY	Paraguay
CN	China	PE	Peru
HK	China (Hong Kong S.A.R.)	PH	Philippines
MO	China (Macau S.A.R.)	PN	Pitcairn
CX	Christmas Island	PL	Poland
CC	Cocos (Keeling) Islands	PT	Portugal
CO	Colombia	PR	Puerto Rico
KM	Comoros	QA	Qatar
CG	Congo	RE	Reunion
CD	Congo, The Democratic Republic of the	RO	Romania
CK	Cook Islands	RU	Russia
CR	Costa Rica	RW	Rwanda
CI	Cote d'Ivoire	KN	Saint Kitts and Nevis
HR	Croatia	LC	Saint Lucia
CU	Cuba	VC	Saint Vincent and the Grenadines
CY	Cyprus	WS	Samoa (Western)
CZ	Czech Republic	SM	San Marino
DK	Denmark	ST	Sao Tome and Principe
DJ	Djibouti	SA	Saudi Arabia
DM	Dominica	SN	Senegal
DO	Dominican Republic	YU	Serbia
TL	East Timor	T9	Service Agent Test Systems
EC	Ecuador	SC	Seychelles
EG	Egypt	SL	Sierra Leone
SV	El Salvador	SG	Singapore
GQ	Equatorial Guinea	SK	Slovakia
ER	Eritrea	SI	Slovenia
EE	Estonia	SB	Solomon Islands
ET	Ethiopia	SO	Somalia
FK	Falkland Islands (Malvinas)	ZA	South Africa
FO	Faroe Islands	GS	South Georgia and the South Sandwich Islands
FJ	Fiji	ES	Spain

FI	Finland	LK	Sri Lanka
FR	France	US	St. Croix
FX	France Metropolitan	SH	St. Helena
GF	French Guiana	PM	St. Pierre and Miquelon
PF	French Polynesia	US	St. Thomas
TF	French Southern Territories	SD	Sudan
GA	Gabon	SR	Suriname
GM	Gambia	SJ	Svalbard and Jan Mayen Islands
GE	Georgia	SZ	Swaziland
DE	Germany	SE	Sweden
GH	Ghana	CH	Switzerland
GI	Gibraltar	SY	Syrian Arab Republic (Syria)
GR	Greece	TW	Taiwan
GL	Greenland	TJ	Tajikistan
GD	Grenada	TZ	Tanzania
GP	Guadaloupe	TH	Thailand
GU	Guam	TG	Togo
GT	Guatemla	TK	Tokelau
GN	Guinea	TO	Tonga
GW	Guinea-Bissau	TT	Trinidad And Tobago
GY	Guyana	TN	Tunisia
HT	Haiti	TR	Turkey
HM	Heard and McDonald Islands	TM	Turkmenistan
HN	Honduras	TC	Turks and Caicos Islands
HU	Hungary	TV	Tuvalu
IS	Iceland	UG	Uganda
IN	India	UA	Ukraine
ID	Indonesia	AE	United Arab Emirates (Abu Dhabi)
IR	Iran	AE	United Arab Emirates (Dubai)
IQ	Iraq	GB	United Kingdom
IE	Ireland	US	United States
IL	Israel	UM	United States Minor Outlying Islands
IT	Italy	UY	Uruguay
JM	Jamaica	UZ	Uzbekistan
JP	Japan	VU	Vanuatu

US	Johnston Atoll	VA	Vatican City State (Holy See)
JO	Jordan	VE	Venezuela
KZ	Kazakhstan	VN	Vietnam
KE	Kenya	VG	Virgin Islands (British)
KH	Khmer Rep Cambodia	VI	Virgin Islands (U.S.)
KI	Kiribati	US	Wake Island
KP	Korea, Democratic People's Republic of (North Korea)	WF	Wallis and Futuna Islands
KR	Korea, Republic of (South Korea)	EH	Western Sahara
KW	Kuwait	YE	Yemen
KG	Kyrgyzstan	YU	Yugoslavia
LA	Lao People's Democratic Republic (Laos)	ZM	Zambia
LV	Latvia	ZW	Zimbabwe
LB	Lebanon		

Remote LAN module (RLM) configuration worksheets

Table A-21 provides a worksheet for RLM configuration.

Table A-21 RLM worksheets

Node setup	
Would you like to enable DHCP on the RLM interface?	
IP address.	
Network mask (subnet mask).	
Gateway.	
Name or IP address of AutoSupport mailhost.	

Start with the hardware

Refer to your appropriate installation and setup instructions for your model at:

`http://www.ibm.com\storage\support\nas`

Check the instructions on the document about:

- Unpacking the N series
- Rack mounting
- Connecting to storage expansions
- Power and network cable installations

The IBM System Storage N series come with pre-configured software and hardware, and with no monitor or keyboard for user access. This is commonly termed a *headless* system. A storage administrator accesses the systems and manages the disk resources from a remote console using a Web browser or command line after initial setup. Otherwise, they use a serial port.

Connecting your system to an ASCII terminal console

The ASCII terminal console enables you to monitor the boot process, helps you configure your N series system after it boots, and enables you to perform system administration.

To connect an ASCII terminal console to the N series system, perform these steps:

1. Set the following communications parameters of your system (Table A-22). For example, you can use hyperterminal or putty for Windows users (see Figure A-1) and for Linux users you can use a terminal program like minicom or screen.

Table A-22 Communication parameters

Parameter	Setting
Baud	9600
Data bit	8
Parity	None
Stop bits	1
Flow control	None

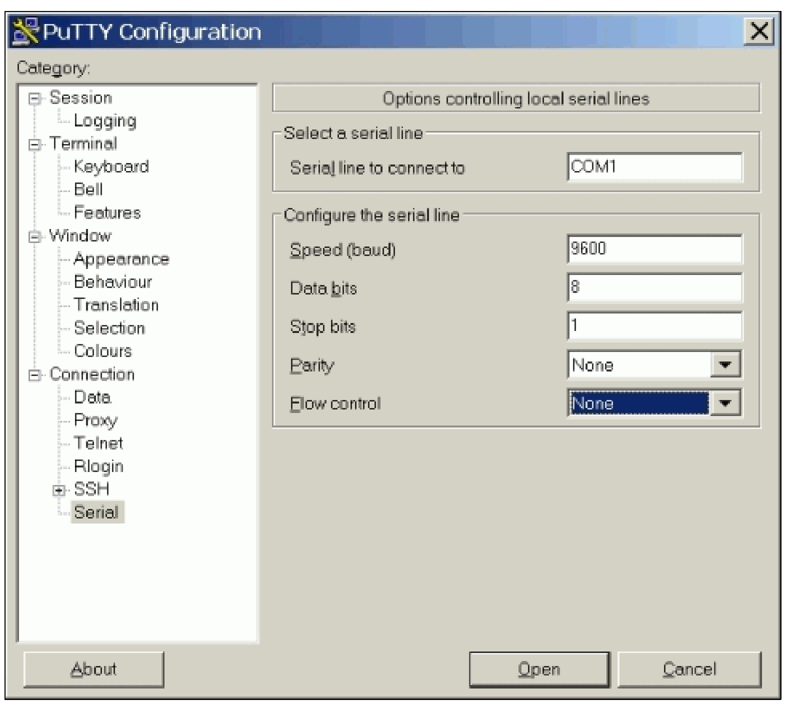


Figure A-1 PuTTY

Note: See your terminal documentation for information about changing your ASCII console terminal settings.

2. Connect the DB-9 null modem cable to the DB-9 to RJ-45 adapter cable, and then connect the RJ-45 end to the console port on the N series system and the other end to the ASCII terminal.
3. Connect to the ASCII terminal console.

Power on N series

After you connect all power cords to the power sources, continue with these steps:

1. Power on N series machines. Always do these steps sequentially to power on N series machines:
   a. Turn on the power to only the expansion units, making sure that you turn them on within 5 minutes of each other.
   b. Turn on the N series filers.
2. Initialize Data ONTAP. This steps provides information if you want to format all disks on a filer and reinstall Data ONTAP. This step can also be used as problem resolution instructions for a situation in which a newly purchased storage system cannot find a root volume (vol0) when trying to boot.

   Otherwise, you can skip this step and continue to step 3.

   Note: This procedure removes all data from all disks.

   a. Turn on the machine. The system begins to boot.

      If you are already at the storage system prompt, enter the following command:

      `halt`

      Then the storage system console displays the boot environment prompt. The boot environment prompt may be CFE> or LOADER>, depending on your storage system. See Example A-1.

Example A-1 N series halt

```
n3300a> halt

CIFS local server is shutting down...

CIFS local server has shut down...

Wed May  2 03:00:13 GMT [n3300a: kern.shutdown:notice]: System shut down because :
"halt".

AMI BIOS8 Modular BIOS
Copyright (C) 1985-2006,  American Megatrends, Inc. All Rights Reserved
Portions Copyright (C) 2006 Network Appliance, Inc. All Rights Reserved
BIOS Version 3.0X11
................

Boot Loader version 1.3
Copyright (C) 2000,2001,2002,2003 Broadcom Corporation.
```

```
Portions Copyright (C) 2002-2006 Network Appliance Inc.

CPU Type: Mobile Intel(R) Celeron(R) CPU 2.20GHz
LOADER>
```

b. When the message `Press CTRL C for special menu` appears, press Ctrl+C to access the special boot menu. See Example A-2.

Example A-2 Boot menu

```
LOADER> boot_ontap
Loading:...............0x200000/33342524 0x21cc43c/31409732 0x3fc0a80/2557763
0x42311c3/5 Entry at 0x00200000
Starting program at 0x00200000
cpuid 0x80000000: 0x80000004 0x0 0x0 0x0
Press CTRL-C for special boot menu
Special boot options menu will be available.
Wed May  2 03:01:27 GMT [fci.initialization.failed:error]: Initialization failed on
Fibre Channel adapter 0a.
Wed May  2 03:01:27 GMT [fci.initialization.failed:error]: Initialization failed on
Fibre Channel adapter 0b.

Data ONTAP Release 7.2.4L1: Wed Nov 21 06:07:37 PST 2007 (IBM)
Copyright (c) 1992-2007 Network Appliance, Inc.
Starting boot on Wed May  2 03:01:12 GMT 2007
Wed May  2 03:01:28 GMT [nvram.battery.turned.on:info]: The NVRAM battery is turned
ON. It is turned OFF during system shutdown.
Wed May  2 03:01:31 GMT [diskown.isEnabled:info]: software ownership has been
enabled for this system
```

c. At the 1-5 special boot menu, choose either option 4 or option 4a. Option 4 creates a RAID 4 traditional volume. Selecting option 4a creates a RAID-DP aggregate with a root FlexVol. The size of the root flexvol is dependant upon platform type. See Example A-3.

Example A-3 Special boot menu

```
(1)  Normal boot.
(2)  Boot without /etc/rc.
(3)  Change password.
(4)  Initialize owned disks (6 disks are owned by this filer).
(4a) Same as option 4, but create a flexible root volume.
(5)  Maintenance mode boot.

Selection (1-5)? 4
```

d. Answer `Y` to the next two displayed prompts to zero out your disks. See Example A-4.

Example A-4 Initialize disks

```
Zero disks and install a new file system? y
This will erase all the data on the disks, are you sure? y
Zeroing disks takes about 45 minutes.
Wed May  2 03:01:47 GMT [coredump.spare.none:info]: No sparecore disk was found.
.................................................................................
.................................................................................
........
```

Attention: Zeroing disks can take 40 minutes or more to complete. Do not turn off power to the system or interrupt the zeroing process.

e. After the disks have been zeroed, the system begins to boot, and then it stops at the first installation question, which is displayed on the console windows:

```
Please enter the new hostname [ ]:
```

See Example A-5.

Example A-5 Initialize complete

```
Wed May  2 03:32:00 GMT [raid.disk.zero.done:notice]: Disk 0c.00.7 Shelf ? Bay ?
[NETAPP   X286_S15K5146A15 NQ06] S/N [3LN11RGT0000974325E5] : disk zeroing complete
Wed May  2 03:32:01 GMT [raid.disk.zero.done:notice]: Disk 0c.00.8 Shelf ? Bay ?
[NETAPP   X286_S15K5146A15 NQ06] S/N [3LN1322S0000974208ZC] : disk zeroing complete
Wed May  2 03:32:02 GMT [raid.disk.zero.done:notice]: Disk 0c.00.1 Shelf ? Bay ?
[NETAPP   X286_S15K5146A15 NQ06] S/N [3LN11G4G00009742TXB2] : disk zeroing complete
Wed May  2 03:32:02 GMT [raid.disk.zero.done:notice]: Disk 0c.00.9 Shelf ? Bay ?
[NETAPP   X286_S15K5146A15 NQ06] S/N [3LN11RCB00009742TX02] : disk zeroing complete
Wed May  2 03:32:09 GMT [raid.disk.zero.done:notice]: Disk 0c.00.10 Shelf ? Bay ?
[NETAPP   X286_S15K5146A15 NQ06] S/N [3LN1321A0000974209ZM] : disk zeroing complete
Wed May  2 03:32:10 GMT [raid.disk.zero.done:notice]: Disk 0c.00.11 Shelf ? Bay ?
[NETAPP   X286_S15K5146A15 NQ06] S/N [3LN120QE00009742TT87] : disk zeroing complete
Wed May  2 03:32:11 GMT [raid.vol.disk.add.done:notice]: Addition of Disk
/vol0/plex0/rg0/0c.00.7 Shelf 0 Bay 7 [NETAPP   X286_S15K5146A15 NQ06] S/N
[3LN11RGT0000974325E5] to volume vol0 has completed successfully
Wed May  2 03:32:11 GMT [raid.vol.disk.add.done:notice]: Addition of Disk
/vol0/plex0/rg0/0c.00.1 Shelf 0 Bay 1 [NETAPP   X286_S15K5146A15 NQ06] S/N
[3LN11G4G00009742TXB2] to volume vol0 has completed successfully
Wed May  2 03:32:11 GMT [wafl.vol.add:notice]: Volume vol0 has been added to the
system.
.
.
.
Please enter the new hostname []:
```

f. Complete the initial setup. See Example A-6 for the initial setup.

g. Install the full operating system. FilerView can be used after the full operating system is installed.

   The full installation procedure is similar to the Data ONTAP update procedure. See “Data ONTAP update” on page 309.

3. The system begins to boot. Complete the initial setup by answering all the installation questions as you have prepared them in the initial worksheet. Refer to the *IBM System Storage Data ONTAP Software Setup Guide,* GA32-0530.

   See Example A-6 for N3300 setup.

Example A-6 Setup

```
Please enter the new hostname []: n3000a
Do you want to configure virtual network interfaces? [n]:
Please enter the IP address for Network Interface e0a []: 9.11.218.246
Please enter the netmask for Network Interface e0a [255.0.0.0]: 255.255.255.0
Should interface e0a take over a partner IP address during failover? [n]:
Please enter media type for e0a {100tx-fd, tp-fd, 100tx, tp, auto (10/100/1000)} [auto]:
Please enter flow control for e0a {none, receive, send, full} [full]:
Do you want e0a to support jumbo frames? [n]:
Please enter the IP address for Network Interface e0b []:
Should interface e0b take over a partner IP address during failover? [n]:
Would you like to continue setup through the web interface? [n]:
Please enter the name or IP address of the default gateway: 9.11.218.1
   The administration host is given root access to the filer's
```

```
   /etc files for system administration.  To allow /etc root access
   to all NFS clients enter RETURN below.
Please enter the name or IP address of the administration host:
Where is the filer located? []: Tucson
Do you want to run DNS resolver? [n]:
Do you want to run NIS client? [n]:
This system will send event messages and weekly reports to IBM Technical Support. To
disable this feature, enter "options autosupport.support.enable off" within 24 hours.
Enabling Autosupport can significantly speed problem determination and resolution should
a problem occur on your system. For further information on Autosupport, please see:
http://now.netapp.com/autosupport/
   Press the return key to continue.

   The Baseboard Management Controller (BMC) provides remote management capabilities
   including console redirection, logging and power control.
   It also extends autosupport by sending down filer event alerts.

Would you like to configure the BMC [y]: n
Name of primary contact (Required) []: administrator
Phone number of primary contact (Required) []: 1-800-426-4968
Alternate phone number of primary contact []: 1-888-7467-426
Primary Contact e-mail address or IBM WebID? []: admin@itso.tucson.ibm.com
Name of secondary contact []:
Phone number of secondary contact []:
Alternate phone number of secondary contact []:
Secondary Contact e-mail address or IBM WebID? []:
Business name (Required) []: itso
Business address (Required) []: Rita Road
City where business resides (Required) []: tucson
State where business resides []: arizona
2-character country code (Required) []: us
Postal code where business resides []:
   The root volume currently contains 2 disks;  you may add more
   disks to it later using the "vol add" or "aggr add" commands.
   Now apply the appropriate licenses to the system and install
   the system files (supplied on the Data ONTAP CD-ROM or downloaded
   from the NOW site) from a UNIX or Windows host.  When you are
   finished, type "download" to install the boot image and
   "reboot" to start using the system.
Thu May  3 05:33:10 GMT [n3300a: init_java:warning]: Java disabled:  Missing
/etc/java/rt131.jar.
Thu May  3 05:33:10 GMT [dfu.firmwareUpToDate:info]: Firmware is up-to-date on all disk
drives
Thu May  3 05:33:13 GMT [n3300a: 10/100/1000/e0a:info]: Ethernet e0a: Link up
add net default: gateway 9.11.218.1
Thu May  3 05:33:15 GMT [n3300a: httpd_servlet:warning]: Java Virtual Machine not
accessible
There are 4 spare disks; you may want to use the vol or aggr command
to create new volumes or aggregates or add disks to the existing volume.
Thu May  3 05:33:15 GMT [mgr.boot.disk_done:info]: Data ONTAP Release 7.2.5.1 boot
complete. Last disk update written at Thu Jan  1 00:00:00 GMT 1970
Clustered failover is not licensed.
Thu May  3 05:33:15 GMT [cf.fm.unexpectedAdapter:warning]: Warning: clustering is not
licensed yet an interconnect adapter was found. NVRAM will be divided into two parts
until adapter is removed
Thu May  3 05:33:15 GMT [cf.fm.unexpectedPartner:warning]: Warning: clustering is not
licensed yet the node once had a cluster partner
Thu May  3 05:33:16 GMT [mgr.boot.reason_ok:notice]: System rebooted.
Thu May  3 05:33:16 GMT [asup.config.minimal.unavailable:warning]: Minimal Autosupports
unavailable. Could not read /etc/asup_content.conf
```

```
n3300a> Thu May  3 05:33:18 GMT [n3300a: console_login_mgr:info]: root logged in from
console
```

4. Add software licenses by entering the command:

   `license add <license>`

   See Example A-7.

   Example A-7 Example NFS license

```
n3300a> license add XXXXXXX
n3300a> Wed May 3 23:19:30 GMT [rc:notice]: nfs licensed
```

5. Always consider updating firmware and Data ONTAP to the recommended version. See “Data ONTAP update” on page 309.
6. Do the these steps again on the second filer for N series with model A20 or A21.

Data ONTAP update

Refer to the IBM System Storage N series Data ONTAP Matrix on the IBM NAS website to identify the compatible IBM System Storage N series hardware for the currently supported releases of Data ONTAP:

`http:\\www.ibm.com/storage/support/nas`

We strongly recommend updating the installed N series storage controller to the latest recommended Data ONTAP release. This recommendation is based on metrics that demonstrate reliability over a large number of customer installations as well as completion of compatibility testing with other products.

Upgrading Data ONTAP software requires several prerequisites, installing system files, and downloading the software to the system CompactFlash. Required procedures might include the following:

- Update the motherboard firmware (system firmware).

  To determine whether your storage system needs a system firmware update, compare the version of installed system firmware with the latest version available.

- Update the disk firmware.

  When you update the storage system software, disk firmware is updated automatically as part of the storage system software update process. A manual update is not necessary unless the new firmware is not compatible with the storage system disks.

- Update the Data ONTAP.

  The latest system firmware is included with Data ONTAP update packages for CompactFlash-based storage systems. New disk firmware is sometimes included with Data ONTAP update packages. For more information see the *Data ONTAP Upgrade Guide at:*

`http:\\www.ibm.com/storage/support/nas`

There are two methods to upgrade storage systems in an active/active configuration:

- Nondistruptive

  The nondisruptive update method is appropriate when you need to maintain service availability during system updates. When you halt one node and allow takeover, the partner node continues to serve data for the halted node while you update the node that you halted.

- Standard

  The standard update method is appropriate when you can schedule downtime for system updates.

Upgrading Data ONTAP for a single system requires downtime.

Important: Review the *Data ONTAP Release Notes and IBM System Storage N series Data ONTAP Upgrade Guide* for your version of Data ONTAP at:

`http:\\www.ibm.com/storage/support/nas`

Obtain Data ONTAP software from IBM NAS website

To obtain Data ONTAP, perform these steps:

1. Log in using a registered N series user.
2. Click **Software Packages** for registered N series customers. See Figure A-2.

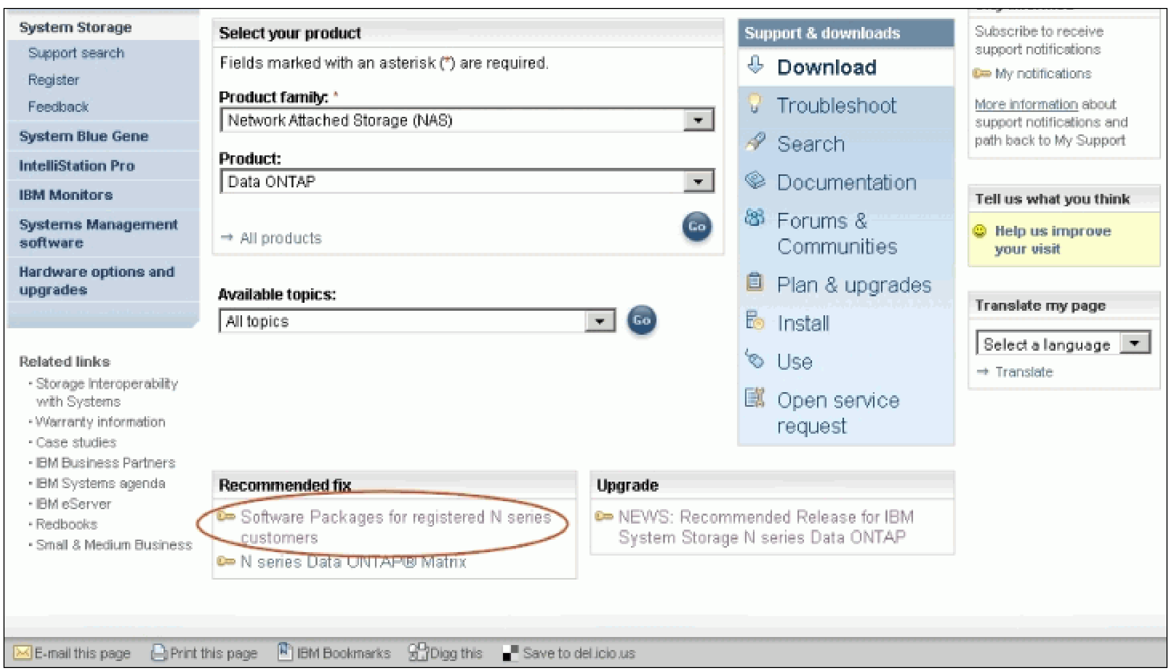


Figure A-2 Software packages

3. Once provided with the Web access keys on the N series and related host software downloads for the individual machines they own, all software updates for these machines will be visible. See Figure A-3.

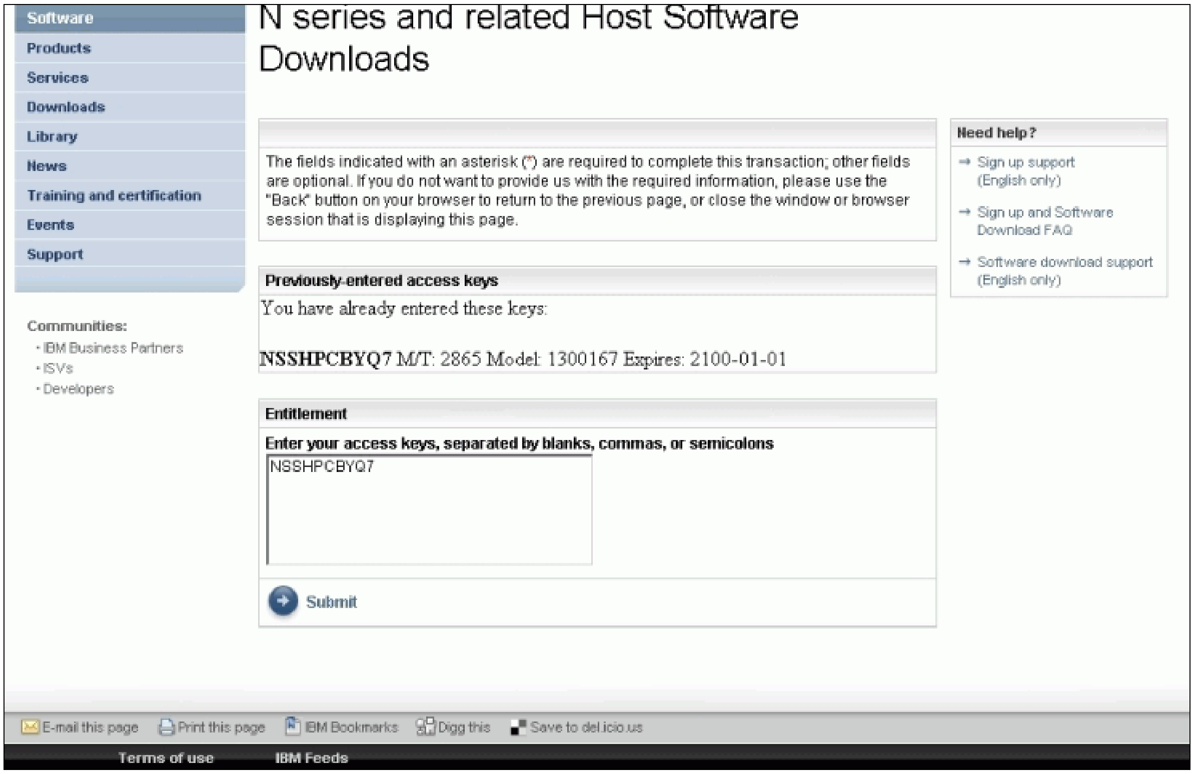


Figure A-3 Web access keys

4. Select the Data ONTAP version. See Figure A-4.

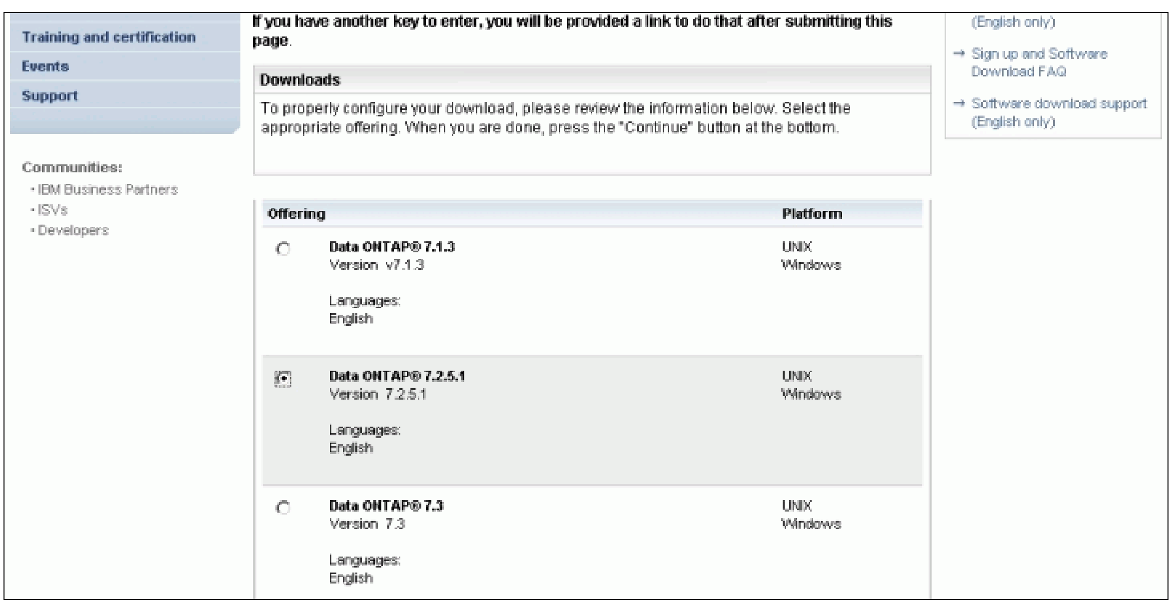


Figure A-4 Data ONTAP versions

5. Select the installation kit that you want to download. Check and confirm the license agreement to start downloading the software. Example on Figure A-5.

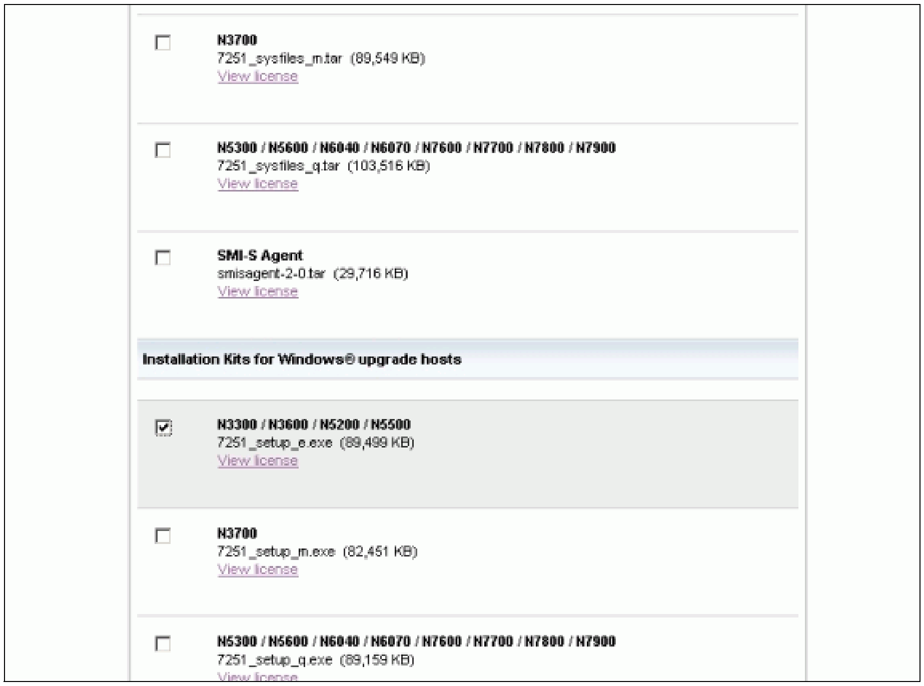


Figure A-5 Installation kits

Install Data ONTAP system files

You can install Data ONTAP system files from a UNIX client, Windows client, or HTTP server. To do install from a Windows client:

1. Set up CIFS on the filer:

   a. Add a CIFS license. See Example A-8.

Example A-8 CIFS license

```
n3300a*> license add XXXXXXX
      Run cifs setup to enable cifs.
```

   b. Set up the CIFS to install Data ONTAP. Enter:

   `cifs setup`

   See Example A-9 on page 313.

Example A-9 Basic CIFS setup

```
n3300a*> cifs setup
This process will enable CIFS access to the filer from a Windows(R) system.
Use "?" for help at any prompt and Ctrl-C to exit without committing changes.

      Your filer does not have WINS configured and is visible only to
      clients on the same subnet.
Do you want to make the system visible via WINS? [n]:
      A filer can be configured for multiprotocol access, or as an NTFS-only
      filer. Since NFS, DAFS, VLD, FCP, and iSCSI are not licensed on this
      filer, we recommend that you configure this filer as an NTFS-only
      filer
(1) NTFS-only filer
(2) Multiprotocol filer
Selection (1-2)? [1]: 1
      CIFS requires local /etc/passwd and /etc/group files and default files
      will be created.  The default passwd file contains entries for 'root',
      'pcuser', and 'nobody'.
Enter the password for the root user []:
Retype the password:
      The default name for this CIFS server is 'N3300A'.
Would you like to change this name? [n]:
      Data ONTAP CIFS services support four styles of user authentication.
      Choose the one from the list below that best suits your situation.

(1) Active Directory domain authentication (Active Directory domains only)
(2) Windows NT 4 domain authentication (Windows NT or Active Directory domains)
(3) Windows Workgroup authentication using the filer's local user accounts
(4) /etc/passwd and/or NIS/LDAP authentication
Selection (1-4)? [1]: 4
What is the name of the Workgroup? [WORKGROUP]:
CIFS - Starting SMB protocol...
Welcome to the WORKGROUP Windows(R) workgroup

CIFS local server is running.
n3300a*> cif        Wed May  2 04:25:30 GMT [nbt.nbns.registrationComplete:info]:
NBT: All CIFS name registrations have completed for the local server.
```

c. Give share access for C$. This access should be set again later for security purposes. Use this command:

```
cifs access <share> <user|group> <rights>
```

See Example A-10.

Example A-10 Share CIFS access

```
n3300a*> cifs access C$ root "Full Control"
1 share(s) have been successfully modified
n3300a*> cifs shares
Mount Point                       Description
----         -----------                     -----------
ETC$         /etc                            Remote Administration
         ** priv access only **
HOME         /vol/vol0/home                  Default Share
         everyone / Full Control
C$           /                               Remote Administration
         root / Full Control
```

2. Map the system storage to a drive. You must log in as administrator or log in using an account that has full control on the storage system C$ directory.
   a. Click **Tools** → **Map Network Drive**. See Figure A-6.

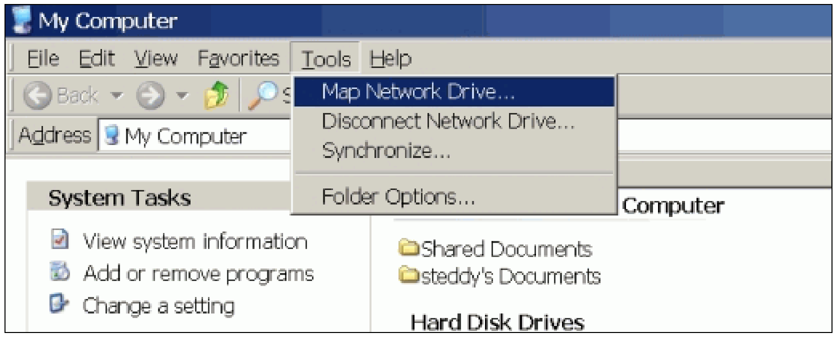


Figure A-6 Map Network Drive

   b. Enter the network mapping address. See Figure A-7.

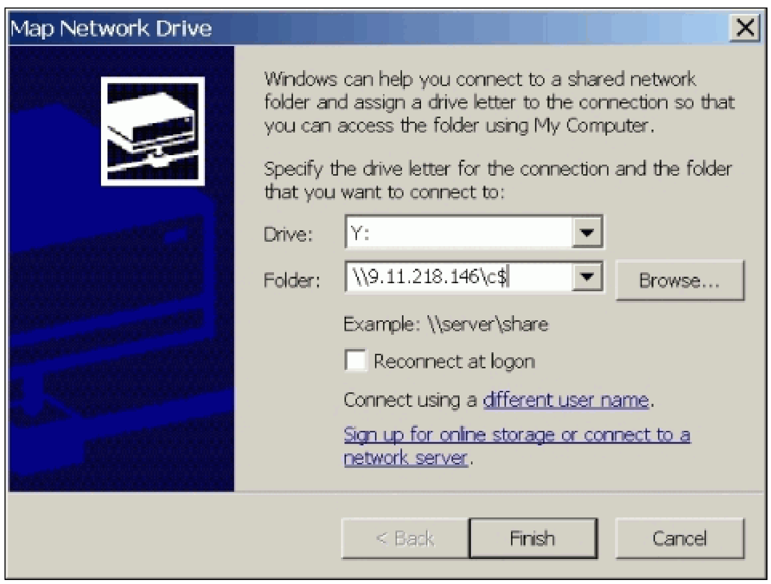


Figure A-7 Mapping address

   c. Enter a user name and password to access the storage system. See Figure A-8.

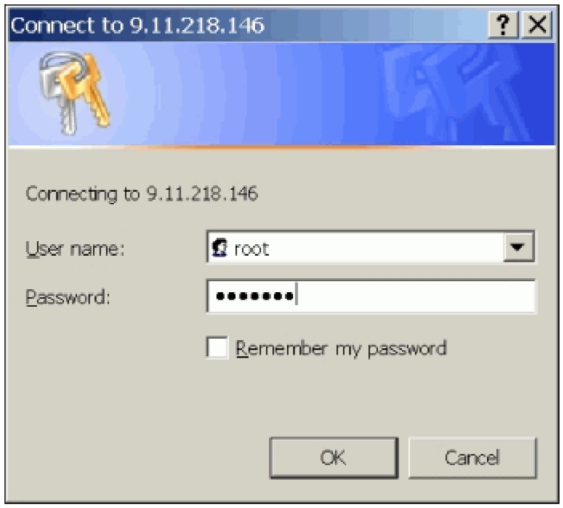


Figure A-8 Storage access

d. The drive now has been mapped. See Figure A-9.

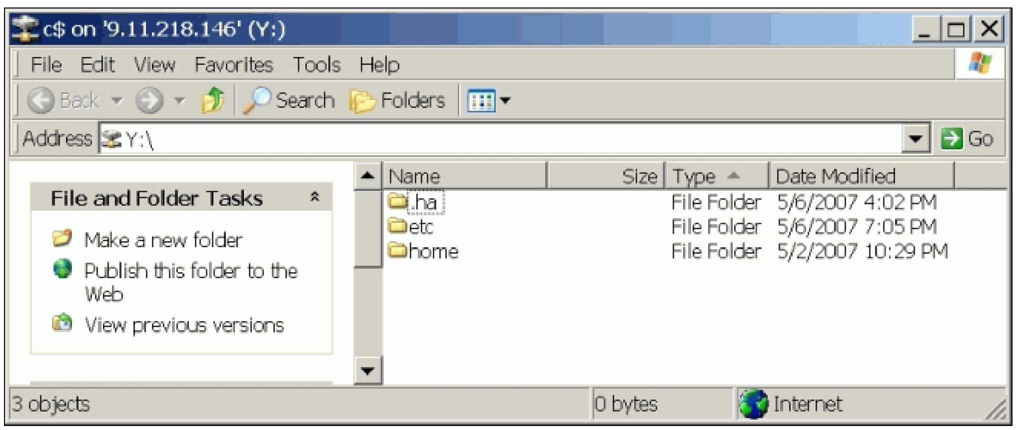


Figure A-9 Drive mapping example

3. Run the Data ONTAP installer:
   a. Go to the drive to which you previously downloaded the software ("Obtain Data ONTAP software from IBM NAS website" on page 310).
   b. Double-click the files that you downloaded. A dialog box appears. See Figure A-10.

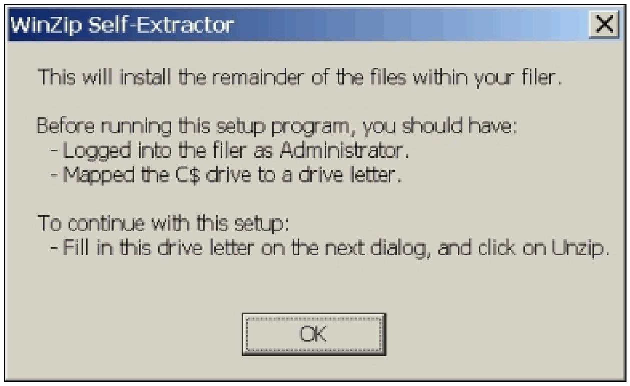


Figure A-10 Winzip self-extractor

c. In the WinZip dialog box, enter the letter of the drive to which you mapped the storage system. For example, if you chose drive Y, replace DRIVE:\ETC with the following path:

```
Y:\ETC
```

See Figure A-11.

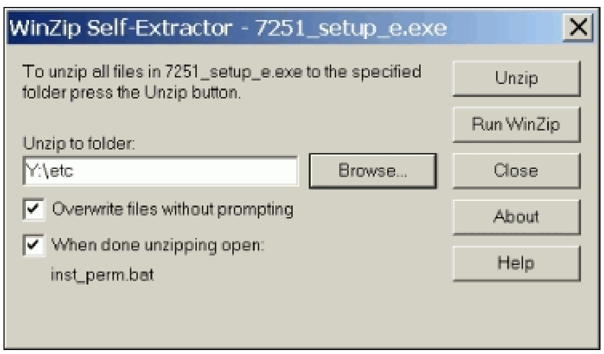


Figure A-11 Extract path

d. Ensure that the following check boxes are selected:

- Overwrite Files Without Prompting
- When Done Unzipping Open

Leave the options as they are.

e. Click the **Unzip** button. In the dialog box's lower pane, confirmation messages are displayed as files are decompressed. See Figure A-12.

Figure A-12 Extraction finished

f. Run the script installer. See Figure A-13.

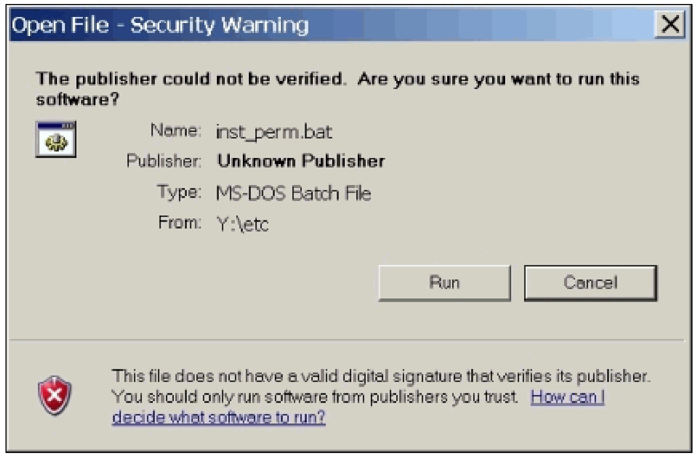


Figure A-13 Script installer

g. Check the script output for minimum requirements. See Figure A-14.

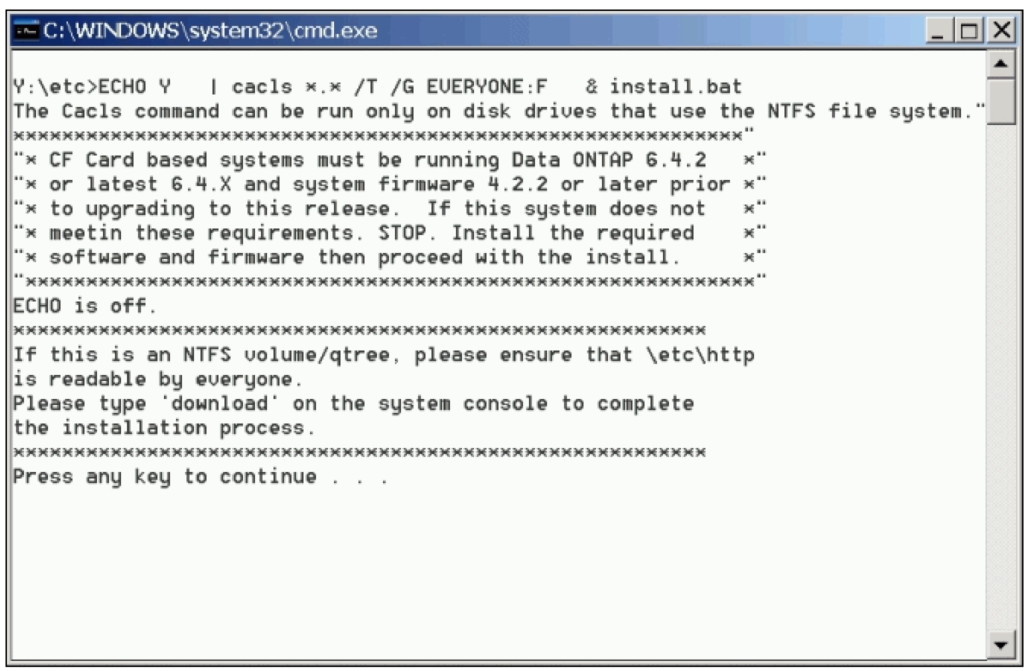


Figure A-14 Script output

Download Data ONTAP to the storage system

The following steps describe the standard update method. For the nondistruptive method on an active/active configuration see the *Data ONTAP Upgrade Guide*. Perform these steps:

1. Install Data ONTAP. Type the **download** command to copy the kernel and firmware data files to the CompactFlash card. The download command provides a status message similar to Example A-11.

Example A-11 Download process

```
n3300a*> download
download: You can cancel this operation by hitting Ctrl-C in the next 6 seconds.
```

```
download: Depending on system load, it may take many minutes
download: to complete this operation. Until it finishes, you will
download: not be able to use the console.
Thu May  3 05:43:50 GMT [download.request:notice]: Operator requested download initiated
download: Downloading boot device
Version 1 ELF86 kernel detected.
...........
download: Downloading boot device (Service Area)
......
n3300a*> Thu May  3 05:49:44 GMT [download.requestDone:notice]: Operator requested
download completed
```

2. Check whether your system requires a firmware update.

   At the console of each storage system, enter the following commands to compare the installed version of system firmware with the version on the CompactFlash card. Display the version of your current system firmware:

   `sysconfig -a`

   See Example A-12.

Example A-12   sysconfig -a

```
n3300a*> sysconfig -a
   Data ONTAP Release 7.2.5.1: Wed Jun 25 11:01:02 PDT 2008 (IBM)
   System ID: 0135018677 (n3300a); partner ID: 0135018673 (n3300b)
   System Serial Number: 2859138306700 (n3300a)
   System Rev: B0
   slot 0: System Board 2198 MHz (System Board XIV D0)
                Model Name:         N3300
                Machine Type:       IBM-2859-A20
                Part Number:        110-00049
                Revision:           D0
                Serial Number:      800949
                BIOS version:       3.0
                Processors:         1
                Processor ID:       0xf29
                Microcode Version:  0x2f
                Memory Size:        896 MB
                NVMEM Size:         128 MB of Main Memory Used
                CMOS RAM Status:    OK
                Controller:         B
...
```

Display the firmware version on the CompactFlash:

`version -b`

See Example A-13.

Example A-13   version -b

```
n3300a*> version -b
1:/x86_elf/kernel/primary.krn: OS 7.2.5.1
1:/backup/x86_elf/kernel/primary.krn: OS 7.2.4L1
1:/x86_elf/diag/diag.krn:  5.3
1:/x86_elf/firmware/deux/firmware.img: Firmware 3.1.0
1:/x86_elf/firmware/SB_XIV/firmware.img: BIOS/NABL Firmware 3.0
1:/x86_elf/firmware/SB_XIV/bmc.img: BMC Firmware 1.1
```

Table A-23 Firmware update requirement

If the version of the newly loaded firmware displayed by the version command is...	Then...
The same as the installed version displayed by sysconfig	Your storage system does not need a system firmware update.
Later than the installed version displayed by sysconfig	Your storage system needs a system firmware update.
Earlier than the installed version displayed by sysconfig	Do *not* update system firmware.

3. Shut down the machine using the **halt** command. After the storage system shuts down, the firmware boot environment prompt appears. See Example A-14.

Example A-14 Halting process

```
n3300a*> halt
CIFS local server is shutting down...
waiting for CIFS shut down (^C aborts)...
CIFS local server has shut down...
Thu May  3 05:51:54 GMT [kern.shutdown:notice]: System shut down because : "halt".
AMI BIOS8 Modular BIOS
Copyright (C) 1985-2006,  American Megatrends, Inc. All Rights Reserved
Portions Copyright (C) 2006 Network Appliance, Inc. All Rights Reserved
BIOS Version 3.0
................
Boot Loader version 1.3
Copyright (C) 2000,2001,2002,2003 Broadcom Corporation.
Portions Copyright (C) 2002-2006 Network Appliance Inc.
CPU Type: Mobile Intel(R) Celeron(R) CPU 2.20GHz
LOADER>
```

4. From the environmental prompt you can update your firmware by doing update_flash.
5. At the firmware environment boot prompt, enter **bye** to reboot the system using the new software and, if applicable, the new firmware. See Figure A-15 on page 320.

Example A-15 Reboot the system

```
LOADER> bye
AMI BIOS8 Modular BIOS
Copyright (C) 1985-2006,  American Megatrends, Inc. All Rights Reserved
Portions Copyright (C) 2006 Network Appliance, Inc. All Rights Reserved
BIOS Version 3.0
..................
```

In Data ONTAP 7.2 and later, disk firmware updates for RAID 4 aggregates must complete before the new Data ONTAP version can finish booting. Storage system services are not available until the disk firmware update completes.

6. Check the **/etc/messages** and **sysconfig -v** outputs to verify that the updates were successful.

Network setup using console

The easiest way to change network configuration is by using **setup** command. But the new contents do not take effect until the filer is rebooted. This section discusses how to change some network configuration without rebooting the filer.

Change IP address

To change the IP address of a filer, perform these steps:

1. List the contents of the /etc/hosts file to note the filer's name and associated IP address. For example, in the following listing, the filer's name is n3300a and its IP address is 9.11.218.146, associated with interface e0a. See Example A-16.

Example A-16 List host name

```
n3300a>  rdfile /etc/hosts
#Auto-generated by setup Sat May  5 23:06:14 GMT 2007
127.0.0.1 localhost
9.11.218.146    n3300a n3300a-e0a
# 0.0.0.0       n3300a-e0b
```

2. To change the network IP address:

```
ifconfig <interface_name> <new_IP_address> netmask <mask>
```

> **Note:** You must be connected to the console. If you are connected by telnet, the connection will be terminated after running the **ifconfig** command.

See Example A-17.

Example A-17 Change network IP

```
n3300a> ifconfig e0a 9.11.218.147 netmask 255.255.255.0
n3300a> netstat -in
Name    Mtu    Network      Address      Ipkts  Ierrs  Opkts  Oerrs  Collis  Queue
e0a     1500 9.11.218/24  9.11.218.147    33k      0    13k      0       0      0
e0b*    1500  none         none             0      0      0      0       0      0
lo      8160  127          127.0.0.1       52      0     52      0       0      0
```

3. If you want this IP address to be persistent after the filer rebooted, update the '/etc/hosts' for IP address changes in the associated interface. For netmask and other network parameters, update the '/etc/rc'. You can modify this file from the N series console, CIFS, or NFS. For this example, we use a CIFS connection to update these files. See Figure A-15 and Figure A-16 on page 321.

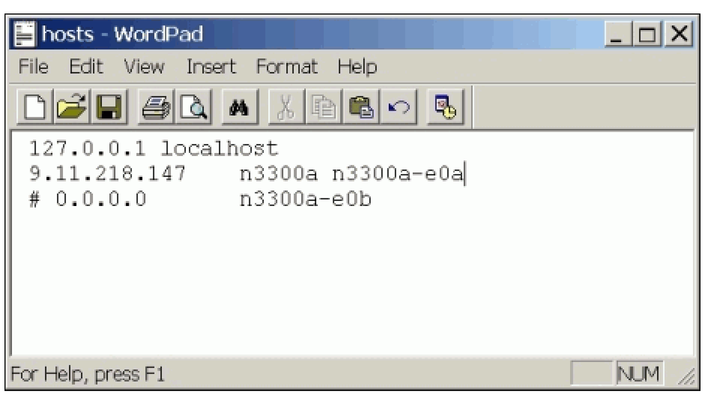


Figure A-15 List host name from Windows

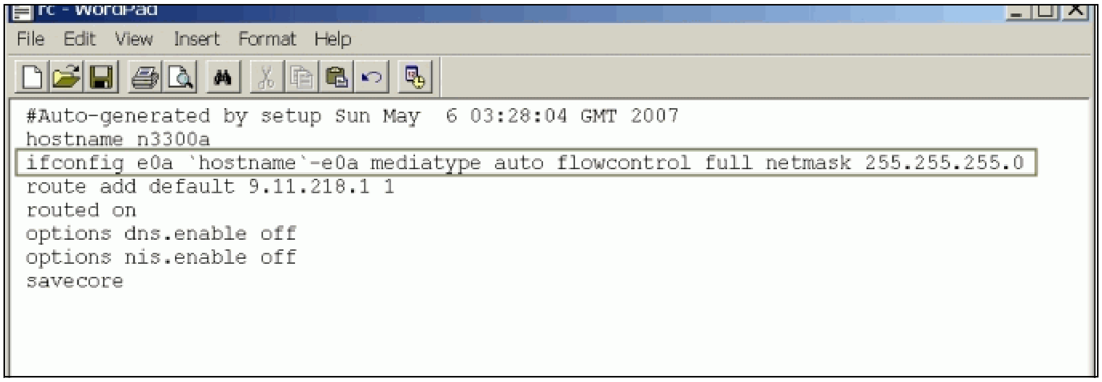


Figure A-16 /etc/rc file

DNS setup

To set up DNS:

1. Create/update the file '/etc/resolv.conf'. Then add/update these entries to the add name server:

   ```
   nameserver ip_address
   ```

   See Figure A-17.

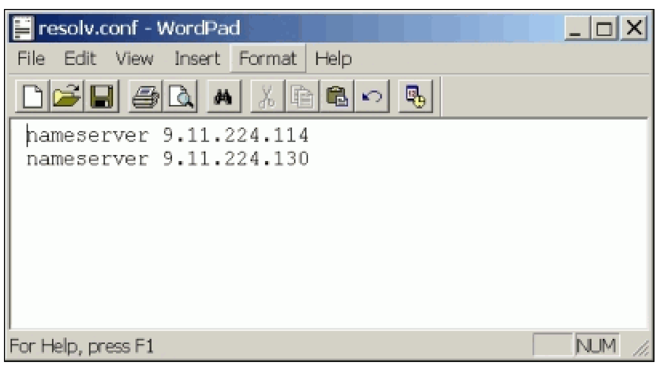


Figure A-17 Name server

2. Update/confirm the DNS domain name with the following commands:

   – To display the current DNS domain name:

   ```
   options dns.domainname
   ```

   – To update the DNS domain name:

   ```
   options dns.domainname <domain name>
   ```

See Example A-18.

Example A-18 Update DNS domain name

```
#---check the dns domainname---
n3300a> options dns.domainname
dns.domainname                          (value might be overwritten in takeover)
#---update
n3300a> options dns.domainname itso.tucson.ibm.com
You are changing option dns.domainname which applies to both members of the cluster in
takeover mode.
This value must be the same in both cluster members prior to any takeover or giveback,
or that next takeover/giveback may not work correctly.
Sun May  6 03:41:01 GMT [n3300a: reg.options.cf.change:warning]: Option dns.domainname
changed on one cluster node.
n3300a> options dns.domainname
dns.domainname               itso.tucson.ibm.com (value might be overwritten in
takeover)
```

3. Check that the DNS is already enable using the **dns info** command. To enable DNS:

```
options dns.enable on
```

See Example A-19.

Example A-19 Enabling DNS

```
n3300a> dns info
DNS is disabled
n3300a>
n3300a>
n3300a> options dns.enable on
Sun May  6 03:50:06 GMT [n3300a: reg.options.overrideRc:warning]: Setting option
dns.enable to 'on' conflicts with /etc/rc that sets it to 'off'.
** Option dns.enable is being set to "on", but this conflicts
** with a line in /etc/rc that sets it to "off".
** Options are automatically persistent, but the line in /etc/rc
** will override this persistence, so if you want to make this change
** persistent, you will need to change (or remove) the line in /etc/rc.
You are changing option dns.enable which applies to both members of
the cluster in takeover mode.
This value must be the same in both cluster members prior to any takeover
or giveback, or that next takeover/giveback may not work correctly.
Sun May  6 03:50:06 GMT [n3300a: reg.options.cf.change:warning]: Option dns.enable
changed on one cluster node.
n3300a>
n3300a>
n3300a> dns info
DNS is enabled

DNS caching is enabled

0 cache hits
0 cache misses
0 cache entries
0 expired entries
0 cache replacements

IP Address                                              State   Last Polled              Avg RTT Calls  Errs
------------------------------------------------------------------------------------------------------------------------------
9.11.224.114                                            NO INFO                                 0     0     0
9.11.224.130                                            NO INFO                                 0     0     0
```

```
Default domain: itso.tucson.ibm.com
Search domains: itso.tucson.ibm.com tucson.ibm.com ibm.com
```

4. To make this persistent after filer reboot, update the‘/etc/rc’ to ensure that the name server exists. See Figure A-18.

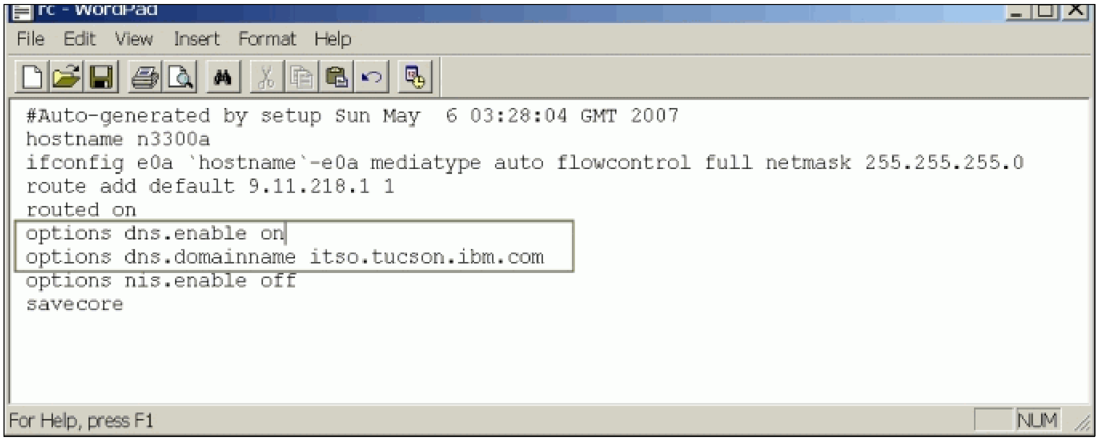


Figure A-18 /etc/rc file

Update to Data ONTAP 8

Before upgrading to DOT8 7-mode you should carefully inspect your system including installed hardware and software. Upgrade all software to the most current release.

Only migrations from 7.3.x to DOT 8 7-mode provide the possibility for an non disruptive upgrade (NDU). This upgrade path is the only one that can be reverted without data loss.

All other migration paths require a clean install, where all data must be backed up, because the systems are installed from scratch and existing data is erased.

To organize your upgrade process follow these high-level steps:

1. Review your current system hardware and licenses.
2. Review all necessary documentation.
3. Generate an AutoSupport email.
4. Obtain the Data ONTAP upgrade image.
5. Install the software and download the new version to the CompactFlash card.
6. Reboot the system.
7. Verify the install.

Before performing the storage controller non disruptive upgrade perform the following steps:

1. Validate the high-availability controller configuration.
2. Remove all failed disks to allow giveback operations to succeed.
3. Upgrade disk and shelf firmware.
4. Verify that system loads are within the acceptable range; load should be less than 50% on each system.

Table A-24 shows supported NDU upgrade paths.

Table A-24 Supported high-availability configuration upgrade paths

Source	Release	Upgrade	Revert	NDU
7.2.x	7-mode	yes	yes	no
7.3.x	7-mode	yes	yes	yes

Evaluate free space for LUNs

Before upgrading a storage system in a SAN environment, you must ensure that every volume containing LUNs has available at least 1 MB of free space. The space is needed to accommodate changes in the on-disk data structures used by the new version of Data ONTAP.

System requirements

Generally, DOT8 requires you to use 64-bit hardware. Older 32-bit hardware is not supported. Currently, the supported systems and hardware are:

- N series: N7900, N7700, N6070, N6060, N6040, N5600, N5300, N3040
- Performance acceleration cards (PAM)

Revert considerations

The N series does not support NDU for the revert process for DOT 8 7-mode. The following restrictions apply to the revert process:

- User data will be temporarily offline and unavailable during the revert.
- Administrators will need to plan when the data is offline to limit the unavailability window and make it fall within the timeout window for the SAN attach kits on hosts.
- Administrators must disable DOT 8.0 7-mode features before reverting.
- 64-bit aggregates and 64-bit volumes cannot be reverted, thus the data will need to be migrated.
- You cannot revert while an upgrade is in progress.
- The *revert_to* command will remind administrators of the features that need to be disabled to complete the reversion.
- FlexVol volumes must be online during the reversion.
- Space guarantees should be checked after the reversion.
- You must delete any Snapshot copies made on Data ONTAP 8.0.
- You must re initialize all SnapVault relationships after the revert because all snapshots associated with Data ONTAP 8.0 were deleted.
- SnapMirror sources must be reverted before SnapMirror destinations are reverted.

Example 10-2 shows details of the **revert_to** command.

Example 10-2 revert_to command

```
TUCSON1> revert_to
usage: revert_to [-f] 7.2 (for 7.2 and 7.2.x)
       revert_to [-f] 7.3 (for 7.3 and 7.3.x)

    -f   Attempt to force revert.
TUCSON1>
```

You cannot revert while the upgrade is still in progress. Use the command shown in Example A-20 to check for upgrade processes that are still running.

Example A-20 WAFL scan status

```
TUCSON1> priv set advanced
Warning: These advanced commands are potentially dangerous; use
         them only when directed to do so by IBM
         personnel.
TUCSON1*> wafl scan status
Volume vol0:
 Scan id                   Type of scan     progress
       1    active bitmap rearrangement     fbn 454 of 1494 w/ max_chain_len 7
...
```

Example A-21 shows how the revert process was performed. First, all 64-bit aggregates were removed, all snapshots were deleted for all volumes and aggregates (the command in Example A-21) and snapshot schedules were disabled. SnapMirror also had to be disabled. Then the **software upgrade** command was issued. Finally, the **revert_to** command was issued. The system rebooted to the firmware level prompt. You are now able to perform a netboot or use the **autoboot** command.

Example A-21 The revert process

```
TUCSON1> snapmirror off
...
TUCSON1> snap delete -A -a aggr0
...
TUCSON1> software list
727_setup_q.exe
732_setup_q.exe
8.0RC3_q_image.zip
TUCSON1> software update 732_setup_q.exe
...
TUCSON1> revert_to 7.3
...
autoboot
...
TUCSON1> version
Data ONTAP Release 7.3.2: Thu Oct 15 04:39:55 PDT 2009 (IBM)
TUCSON1>
```

You can use the netboot option for a fresh install of the storage system, booting from a Data ONTAP version stored on a remote HTTP or TFTP (Trivial File Transfer Protocol) server.

Prerequisites: For this procedure we are assuming that the hardware is fine, and includes a 1GB CompactFlash card, an RLM card, and a network interface card. Perform the following steps for a netboot install:

1. Upgrade BIOS if necessary.

   ```
   ifconfig e0c -addr=10.10.123.??? -mask=255.255.255.0 -gw=10.10.123.1
   ping 10.10.123.45
   flash tftp://10.10.123.45/folder.(system_type).flash
   ```

2. Enter one of the following commands at the boot environment prompt:

   – If you are configuring DHCP, enter:

```
ifconfig e0a -auto
```

– If you are configuring manual connections, enter:

```
ifconfig e0a -addr=filer_addr -mask=netmask -gw=gateway -dns=dns_addr
-domain=dns_domain
```

filer_addr is the IP address of the storage system.

netmask is the network mask of the storage system.

gateway is the gateway for the storage system.

dns_addr is the IP address of a name server on your network.

dns_domain is the Domain Name Service (DNS) domain name.

If you use this optional parameter, you do not need a fully qualified domain name in the netboot server URL; you need only the server's host name.

3. Set up Boot environment.

```
set-defaults
setenv ONTAP_NG true
setenv ntap.rlm.gdb 1
setenv ntap.init.usebootp false
setenv ntap.mgwd.autoconf.disable true
```

Depending on N6xxx or N7xxx, set it to e0c for now. You can set it back to e1a later.

```
setenv ntap.bsdportname e0f
setenv ntap.bsdportname e0c
"a New variable for BR nay be needed."
setenv ntap.givebsdmgmtport true #before installing build
setenv ntap.givebsdmgmtport false #after installing build
"FOR 10-MODE"
setenv ntap.init.boot_clustered true
ifconfig e0c -addr=10.10.123.??? -mask=255.255.255.0 -gw=10.10.123.1
ping 10.10.123.45
```

4. Netboot from loader prompt.

```
netboot http://10.10.123.45/home/bootimage/kernel
```

5. Enter the NFS root path.

```
10.10.123.45/vol/home/web/bootimage/rootfs.img
```

The NFS root path is the IP address of an NFS server followed by the export path.

6. Press Ctrl-C for Boot Menu.
7. Select Software Install (option 7).

Enter the URL to install the image.

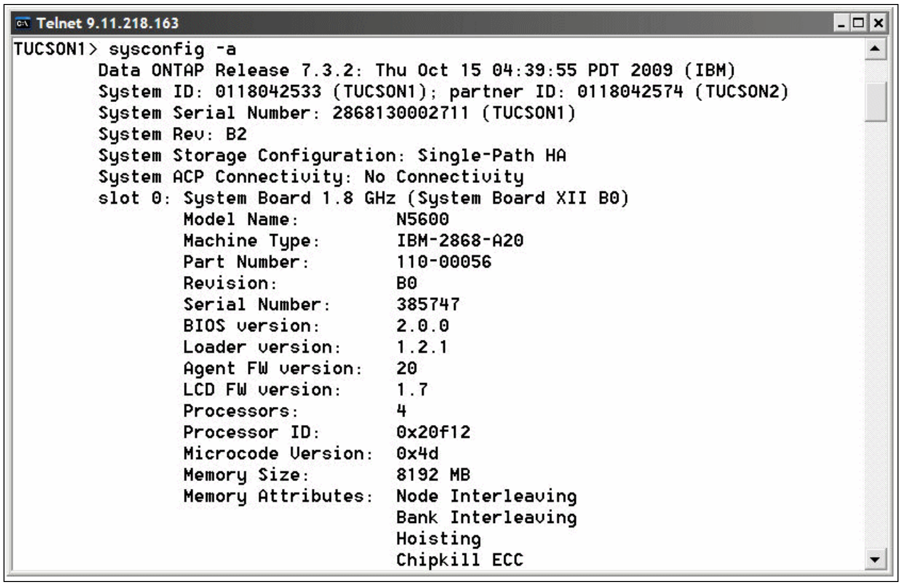
```
http://10.10.123.45/bootimage/image.tgz
```

Note: The provided URLs are examples only, and should be replaced with the URLs for your environment.

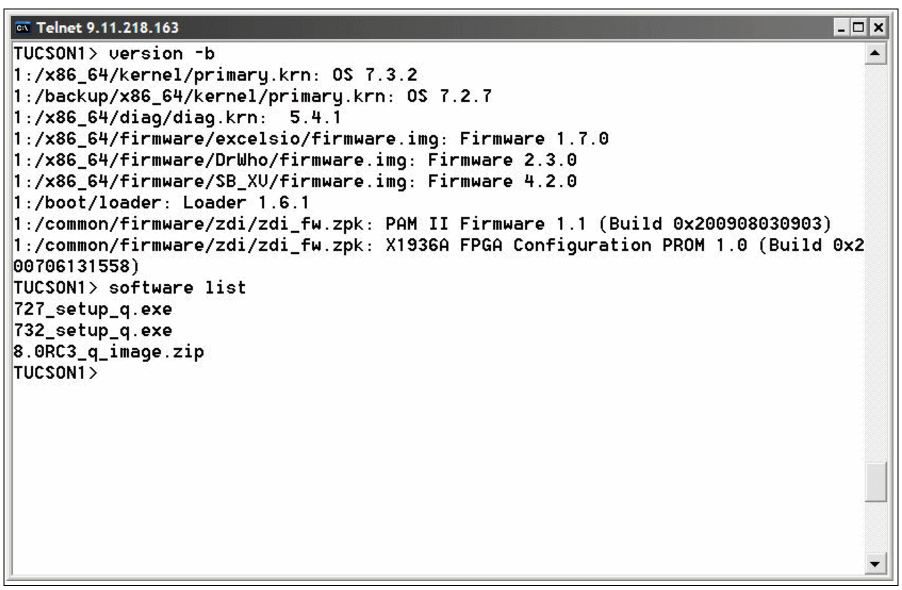
Update example

Our test environment in composed of two N5600 system, each with a designated EXN4000 shelf. An upgrade will be performed from DOT 7.3.2. In case a clean install is required, DOT 8 7-mode also supports the netboot process.

First the current system configuration has to be reviewed by using the **sysconfig -a** command. The output is displayed in Figure A-19 on page 327.


```
Telnet 9.11.218.163
TUCSON1> sysconfig -a
        Data ONTAP Release 7.3.2: Thu Oct 15 04:39:55 PDT 2009 (IBM)
        System ID: 0118042533 (TUCSON1); partner ID: 0118042574 (TUCSON2)
        System Serial Number: 2868130002711 (TUCSON1)
        System Rev: B2
        System Storage Configuration: Single-Path HA
        System ACP Connectivity: No Connectivity
        slot 0: System Board 1.8 GHz (System Board XII B0)
                Model Name:         N5600
                Machine Type:       IBM-2868-A20
                Part Number:        110-00056
                Revision:           B0
                Serial Number:      385747
                BIOS version:       2.0.0
                Loader version:     1.2.1
                Agent FW version:   20
                LCD FW version:     1.7
                Processors:         4
                Processor ID:       0x20f12
                Microcode Version:  0x4d
                Memory Size:        8192 MB
                Memory Attributes:  Node Interleaving
                                    Bank Interleaving
                                    Hoisting
                                    Chipkill ECC
```
Figure A-19 sysconfig -a command

To verify the existing firmware level use the **version -b** command as shown in Figure A-20.

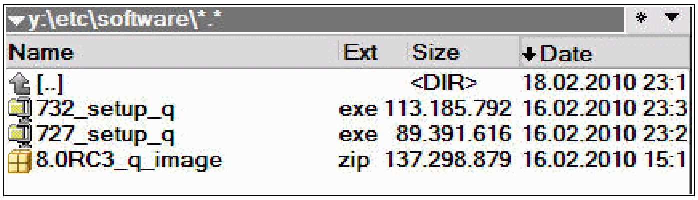

```
Telnet 9.11.218.163
TUCSON1> version -b
1:/x86_64/kernel/primary.krn: OS 7.3.2
1:/backup/x86_64/kernel/primary.krn: OS 7.2.7
1:/x86_64/diag/diag.krn:  5.4.1
1:/x86_64/firmware/excelsio/firmware.img: Firmware 1.7.0
1:/x86_64/firmware/DrWho/firmware.img: Firmware 2.3.0
1:/x86_64/firmware/SB_XV/firmware.img: Firmware 4.2.0
1:/boot/loader: Loader 1.6.1
1:/common/firmware/zdi/zdi_fw.zpk: PAM II Firmware 1.1 (Build 0x200908030903)
1:/common/firmware/zdi/zdi_fw.zpk: X1936A FPGA Configuration PROM 1.0 (Build 0x2
00706131558)
TUCSON1> software list
727_setup_q.exe
732_setup_q.exe
8.0RC3_q_image.zip
TUCSON1>
```
Figure A-20 Verify version

You can also use the **license** command to verify what software is licensed on the system. (Confidentiality requires us to omit this screen shot).

Next, review all necessary documentation including the *Data ONTAP Upgrade Guide* and *Data ONTAP Release Notes* for the destination version of Data ONTAP. Customers can obtain these document from IBM's support website at:

http://www.ibm.com/storage/support/nas

The directory */etc/software* hosts installable ONTAP releases (see Figure A-21 on page 328). The install images have been copied from a Windows client using the administrative share *\\filer_ip\c$* .

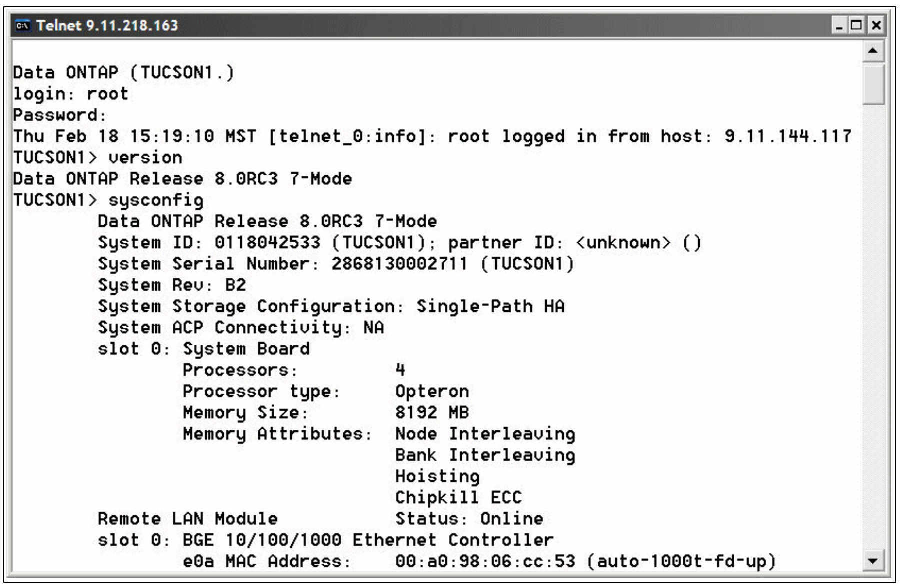


y:\etc\software*.*

Name	Ext	Size	Date
[..]		<DIR>	18.02.2010 23:1
732_setup_q	exe	113.185.792	16.02.2010 23:3
727_setup_q	exe	89.391.616	16.02.2010 23:2
8.0RC3_q_image	zip	137.298.879	16.02.2010 15:1

Figure A-21 Windows client share

Starting with DOT 8, software images end with *.zip* and are no longer *.exe* or *.tar* files. The **software** command must be used to install or upgrade DOT 8 versions. We used the **software update** command. At the time of this writing only DOT 8 7-mode Release Candidate 3 was available, therefore all tasks were performed using this software version. The update process is displayed in Figure A-22 and takes much longer compared to previous DOT version, so be patient.



```
Telnet 9.11.218.163

Data ONTAP (TUCSON1.)
login: root
Password:
Thu Feb 18 15:19:10 MST [telnet_0:info]: root logged in from host: 9.11.144.117
TUCSON1> version
Data ONTAP Release 8.0RC3 7-Mode
TUCSON1> sysconfig
        Data ONTAP Release 8.0RC3 7-Mode
        System ID: 0118042533 (TUCSON1); partner ID: <unknown> ()
        System Serial Number: 2868130002711 (TUCSON1)
        System Rev: B2
        System Storage Configuration: Single-Path HA
        System ACP Connectivity: NA
        slot 0: System Board
                Processors:         4
                Processor type:     Opteron
                Memory Size:        8192 MB
                Memory Attributes:  Node Interleaving
                                    Bank Interleaving
                                    Hoisting
                                    Chipkill ECC
        Remote LAN Module           Status: Online
        slot 0: BGE 10/100/1000 Ethernet Controller
                e0a MAC Address:    00:a0:98:06:cc:53 (auto-1000t-fd-up)
```

Figure A-22 sysconfig command

When the system reboots, press **CTRL-C** to access the first boot menu.

> **Important:** The boot loader must be upgraded, otherwise Data ONTAP 8 will not load. Instead the previously installed version will continue to boot.

The boot loader of the system will be upgraded using the **update_flash** command as shown in Figure A-23 on page 329.

> **Attention:** Ensure that all firmware is up to date! If you are experiencing long boot times, you can disable the auto update of disk firmware prior to downloading Data ONTAP:
> **options raid.background_disk_fw_update.enable off**

```
Boot Loader version 1.2.1
Copyright (C) 2000,2001,2002,2003 Broadcom Corporation.
Portions Copyright (C) 2002-2005 Network Appliance Inc.

CPU Type: Dual Core AMD Opteron(tm) Processor 265
> setenv USE_SECONDARY false


Starting AUTOBOOT press Ctrl-C to abort...

Autoboot of PRIMARY image aborted by user.

LOADER> update_flash
New BIOS Version: 2.4.0
New Loader Version: 1.7
Saving Primary Image to Secondary
Programming .+.+.+.+.+.+.+.+.+.+.+.+.+.+.+.+.+.+.+.+.+.+.+.+.+.+.+.+.+.+.+.+.+ don
e. 2097152 bytes written
Updating Primary Boot Flash
Programming .+.+.+.+.+.+.+.+.+.+.+.+.+.+.+ done. 917504 bytes written
LOADER>
```

Figure A-23 Boot Loader

Next, use the **autoboot** command and perform another **reboot** if DOT 8 did not load immediately after the flash update.

After the boot process has completed the correct version can be verified by using the **version** and **sysconfig** commands as shown in Figure A-24.

```
Telnet 9.11.218.163

Data ONTAP (TUCSON1.)
login: root
Password:
Thu Feb 18 15:19:10 MST [telnet_0:info]: root logged in from host: 9.11.144.117
TUCSON1> version
Data ONTAP Release 8.0RC3 7-Mode
TUCSON1> sysconfig
        Data ONTAP Release 8.0RC3 7-Mode
        System ID: 0118042533 (TUCSON1); partner ID: <unknown> ()
        System Serial Number: 2868130002711 (TUCSON1)
        System Rev: B2
        System Storage Configuration: Single-Path HA
        System ACP Connectivity: NA
        slot 0: System Board
                Processors:         4
                Processor type:     Opteron
                Memory Size:        8192 MB
                Memory Attributes:  Node Interleaving
                                    Bank Interleaving
                                    Hoisting
                                    Chipkill ECC
        Remote LAN Module           Status: Online
        slot 0: BGE 10/100/1000 Ethernet Controller
                e0a MAC Address:    00:a0:98:06:cc:53 (auto-1000t-fd-up)
```

Figure A-24 sysconfig command

B

Power checklists

This appendix provides guidance regarding power checklists when using IBM cabinets and customer cabinets. It also discusses how to determine which type of rack, power connector, and receptacle to use.

This appendix covers the following topics:

- Power checklist for using an IBM cabinet
- Power checklists for using a customer cabinet

Power checklist for using an IBM cabinet

Refer to Chapter 4 in the *IBM System Storage N series Introduction and Planning Guide*, GA32-0543, for information about power and site requirements when using an IBM cabinet. The publication is available at the following address:

http://www-1.ibm.com/support/docview.wss?rs=1144&context=STQTD9&context=STQTDP&context=STQTED&context=STQTFY&dc=DA400&q1=ssg1*&uid=ssg1S7001346&loc=en_US&cs=utf-8&lang=en

Determining type of rack, power connector, and receptacle needed

For rack and power information see the *N series Introduction and Planning Guide*, GA32-0543. Also refer to this IBM publication for detailed information about rack and power options:

ftp://ftp.software.ibm.com/pc/pccbbs/pc_servers_pdf/dmg_rsp_mar2005.pdf

Power (line) cords

These features provide two line cords (IEC 320-C14, which is shown in Figure B-1) to supply external power to the N series.

Figure B-1 IEC 320 C14 Plug Feature Code #9000

One of the following power cord features must be specified:

- The PDU power cord for IEC 320
- Sheet C13 outlets (shown in Figure B-2)
- One based on the country or region in which the system will be installed

Figure B-2 IEC 320 C13 receptacle

Note the following points:

- Minimum required: One #90xx feature

- Maximum allowed: One #90xx feature
- Corequisites: None
- Limitations: Only one power cord feature may be ordered per machine
- Customer setup: Yes
- Plant install or field install: Plant only

Power checklists for using a customer cabinet

Refer to Chapter 4 in the *IBM System Storage N series Introduction and Planning Guide*, GA32-0543, for information about power and site requirements when using a customer cabinet. The publication is available at the following address:

http://www-1.ibm.com/support/docview.wss?rs=1144&context=STQTD9&context=STQTDP&context=STQTED&context=STQTFY&dc=DA400&q1=ssg1*&uid=ssg1S7001346&loc=en_US&cs=utf-8&lang=en

C

N series hardware universe reference tables

In this appendix we include some reference tables that can be useful for selecting the needed HW features before installing and configuring the systems. In particular, we provide:

- N series appliance models with technical specification (Figure C-1 on page 336)
- N7000/6000/3000 series model rear views (Figure C-2 on page 337)
- N series Gateway models with technical specification (Figure C-3 on page 338)
- Fibre channel and Ethernet adapter cards summary (Figure C-4 on page 339)
- SAS and special purposes adapter cards summary (Figure C-5 on page 340)
- Disk drive summary table (Figure C-6 on page 341)
- Shelf enclosures table and N3300 rear view (Figure C-7 on page 342).

For the most up-to-date specifications, see:

http://www-03.ibm.com/systems/storage/network/

Select the specific model to determine whether updates are available.

IBM Hardware Universe – 1 – 2010/03/10

IBM

IBM N series appliance models (Axx)

Model	N7900	N7700	N6070	N6060	N6040	N3600	N3400	N3300
System Capacity Raw Max (2TB dr)	2352TB	1680TB	1680TB	1344TB	840TB	208TB	272TB	136TB
Max Aggregate[2] (32-bit / 64-bit)	16TB / 100TB	16TB / 70TB	16TB / 70TB	16TB / 50TB	16TB / 40TB	16TB / -	16TB / 30TB	16TB[5] / -
Max EXN3000 Shelves	49	35	35	28	17.5	4	4	-
Max EXN-4000/ 1000 Shelves	84	60	60	48	30	6	8	4
Max Drive Quantity	1176	840	840	672	420	104 (20 Int + 84 Ext)	136 (12 Int + 124 Ext)	68 (12 Int + 56 Ext)
Environmental (HA Pair/Single Controller)								
Height (HA/Single)	12U / 6U	12U / 6U	6U / 6U	6U / 6U	6U / 6U	4U / 4U	2U / 2U	2U / 2U
Weight (HA/Single)	242 lb. (109.6 kg) 121 lb. (54.8 kg)	242 lb. (109.6 kg) 121 lb. (54.8 kg)	122 lb. (55.3 kg) 95 lb. (43.1 kg)	122 lb. (55.3 kg) 95 lb. (43.1 kg)	122 lb. (55.3 kg) 95 lb. (43.1 kg)	110 lb. (50 kg) with drives	66 lb. (29.9 kg) with drives	66 lb. (29.9 kg) with drives
AC Power (HA only)	100-120V 11.2A 200-240V 5.8A	100-120V 10.8A 200-240V 5.6A	100-120V 8.1A 200-240V 4A	100-120V 7.7A 200-240V 4.3A	100-120V 5.9A 200-240V 2.9A	100-120V 5.7A 200-240V 2.9A	100-120V 4.6A 200-240V 2.3A	100-120V 4.1A 200-240V 2.2A
Thermal[3] (HA/Single)	3,740 BTU/hr 1,870 BTU/hr	3,624 BTU/hr 1,812 BTU/hr	2,761 BTU/hr 1,602 BTU/hr	2,594 BTU/hr 1,502 BTU/hr	2,026 BTU/hr 1,272 BTU/hr	2,247 BTU/hr 1,988 BTU/hr	1,518 BTU/hr 1,279 BTU/hr	1,587 BTU/hr 1,298 BTU/hr
Platform Specifications (HA Pair/Single Controller)								
Processor (HA/Single)	8 / 4 64-bit dual-core	4 / 2 64-bit	4 / 2 64-bit dual-core	4 / 2 64-bit dual-core	2 / 1 64-bit dual-core	2 / 1 32-bit	2 / 1 32-bit dual-core	2 / 1 32-bit
Memory (HA/Single)	64GB / 32GB	32GB / 16GB	32GB / 16GB	16GB / 8GB	8GB / 4GB	4GB / 2GB	8GB / 4GB	2GB / 1GB
NVRAM (HA/Single)	4GB / 2GB	1GB / 512MB	4GB / 2GB onboard	4GB / 2GB onboard	1GB / 512MB onboard	512MB / 256MB NVMEM	1GB / 512MB NVMEM	256MB / 128MB NVMEM
PCI Slots (HA/Single)	10 / 5 (PCIe) 6 / 3 (PCI-X)	10 / 5 (PCIe) 6 / 3 (PCI-X)	8 / 4 (PCIe)	8 / 4 (PCIe)	8 / 4 (PCIe)	2 / 1 (PCIe)	-	-
Ethernet (HA/Single)	12 / 6 GbE RJ45	12 / 6 GbE RJ45	4 / 2 GbE RJ45	4 / 2 GbE RJ45	4 / 2 GbE RJ45	4 / 2 GbE RJ45	8 / 4 GbE RJ45	4 / 2 GbE RJ45
FC Ports (HA/Single)	16 / 8 4Gb SFP[4]	16 / 8 4Gb SFP[4]	8 / 4 4Gb SFP[4]	8 / 4 4Gb SFP[4]	8 / 4 4Gb SFP[4]	4 / 2 4Gb SFP[4]	4 / 2 4Gb SFP[4]	4 / 2 4Gb SFP[4]
SAS Ports (HA/Single)	- / -	- / -	- / -	- / -	- / -	- / -	2 / 1 12Gb QSFP	- / -
Data ONTAP® (Min Release)	7.2.4 / 8.0 /	7.2.4 / 8.0	7.2.5 / 8.0	7.2.6 / 7.3.1 / 8.0	7.2.5 / 8.0	7.2.2L1	7.3.2 / 8.0	7.2.2L1

Notes

[1] System capacity is calculated using base 10 arithmetic (i.e. 1TB=1,000,000,000,000 bytes) and is derived based on the type, size, and number of drives.

[2] Maximum volume/aggregate size is calculated using base 2 arithmetic (1TB = 2^{40} bytes).

[3] The thermal dissipation values shown are based on typical system values at 100-120V input voltage. Please refer to the Site Requirements

[4] Autosensing ports: 1, 2, 4Gb.

[5] Beginning with Data ONTAP 7.3.1, N3300 systems support aggregates up to 16TB raw capacity, provided that the root volume is hosted in a dedicated aggregate (no user data) or by maintaining two spare disks per controller.

HA Pair Notes

HA pair configuration means two controllers - both actively serving data - are connected for automatic failover protection. The interconnect options are applicable to HA pair configurations using NVRAM5 and NVRAM6 adapters.

NOTE: In order to achieve the maximum 500 meter distance between controllers in a HA pair, the interconnect cable must be a direct point-to-point connection with no intermediate device between them (such as patch panel).

HA configurations that must go through a patch panel can use specific part numbers in conjunction with the Cu-to-Op converters.

Terms and Abbreviations

BTU – British Thermal Unit
Cu – Copper Connector
FC – Fibre Channel
GbE – Gibabit Ethernet
HA – High Availability (2 clustered nodes)
NVRAM – Non-Volatile RAM
Op – Optical Connector
PSU – Power Supply Unit
SAS – Serial-attached SCSI
SATA – Serial ATA
SFP – Small Form-Factor Pluggable
VAC – Volts Alternating Current

Figure C-1 N series models

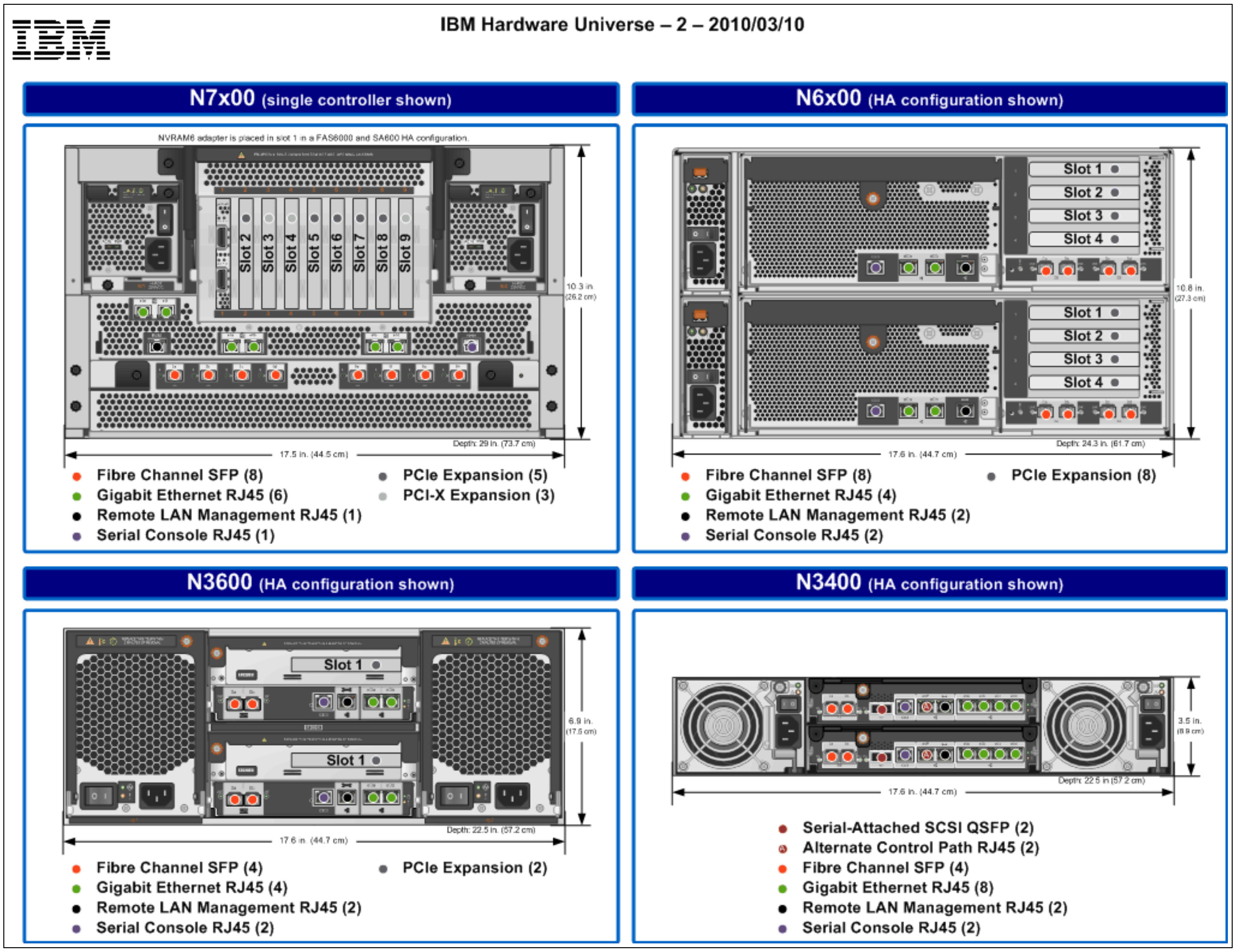


Figure C-2 Rear view of N series

IBM

IBM Hardware Universe – 3 – 2010/03/10

IBM N series Gateway

Model	N7900	N7700	N6070	N6060	N6040
System Capacity Raw Maximum[1]	1176TB	840TB	840TB	672TB	420TB
Max Number of LUNs	1176	840	840	672	420

V-Series Supported Arrays

EMC CLARiiON
EMC Symmetrix
Fujitsu ETERNUS
Hitachi TagmaStore USP
Hitachi Lightning
Hitachi Thunder
HP StorageWorks
IBM TotalStorage
3Par InServ
TMS RamSan

Note: Please check the IBM web site for details and the latest N series gateway support matrix.

Environmental (HA Pair/Single Controller)

	N7900	N7700	N6070	N6060	N6040
Height (HA/Single)	12U / 6U	12U / 6U	6U / 6U	6U / 6U	6U / 6U
Weight (HA/Single)	242 lb. (109.6 kg) 121 lb. (54.8 kg)	242 lb. (109.6 kg) 121 lb. (54.8 kg)	122 lb. (55.3 kg) 95 lb. (43.1 kg)	122 lb. (55.3 kg) 95 lb. (43.1 kg)	122 lb. (55.3 kg) 95 lb. (43.1 kg)
AC Power (HA only)	100-120V 11.1A 200-240V 5.8A	100-120V 22A 200-240V 10A	100-120V 8.1A 200-240V 4A	100-120V 7.7A 200-240V 4.3A	100-120V 5.9A 200-240V 2.9A
Thermal[2] (HA/Single)	3,740 BTU/hr 1,870 BTU/hr	3,624 BTU/hr 1,812 BTU/hr	2,761 BTU/hr 1,602 BTU/hr	2,594 BTU/hr 1,502 BTU/hr	2,026 BTU/hr 1,272 BTU/hr

Platform Specifications (HA Pair/Single Controller)

	N7900	N7700	N6070	N6060	N6040
Processor (HA/Single)	8 / 4 64-bit dual-core	4 / 2 64-bit	4 / 2 64-bit dual-core	4 / 2 64-bit dual-core	2 / 1 64-bit dual-core
Memory (HA/Single)	64GB / 32GB	32GB / 16GB	32GB / 16GB	16GB / 8GB	8GB / 4GB
NVRAM (HA/Single)	4GB / 2GB	1GB / 512MB	4GB / 2GB onboard	4GB / 2GB onboard	1GB / 512MB onboard
PCI Slots (HA/Single)	10 / 5 (PCIe) 6 / 3 (PCI-X)	10 / 5 (PCIe) 6 / 3 (PCI-X)	8 / 4 (PCIe)	8 / 4 (PCIe)	8 / 4 (PCIe)
Ethernet (HA/Single)	12 / 6 GbE RJ45	12 / 6 GbE RJ45	4 / 2 GbE RJ45	4 / 2 GbE RJ45	4 / 2 GbE RJ45
FC Ports (HA/Single)	16 / 8 4Gb SFP[3]	16 / 8 4Gb SFP[3]	8 / 4 4Gb SFP[3]	8 / 4 4Gb SFP[3]	8 / 4 4Gb SFP[3]
Data ONTAP® (Min Release)	7.2.4 / 8.0	7.2.4 / 8.0	7.2.5 / 8.0	7.2.6 / 7.3.1 / 8.0	7.2.5 / 8.0

Notes

1 System capacity is calculated using base 10 arithmetic (i.e. 1TB = 1012 bytes) and is derived based on the type, size, and number of drives.
2 The thermal dissipation values shown are based on typical system values at 100-120V input voltage.
3 Autosensing ports: 1, 2, 4Gb.
4 Max performance and capacity are calculated using base 10 arithmetic
5 Assumes deduplication at 20:1.
6 Assumes that deduplication is disabled and that data is compressible at 2:1.

Terms and Abbreviations

BTU – British Thermal Unit
Cu – Copper Connector
FC – Fibre Channel
GbE – Gibabit Ethernet
HA – High Availability
NVRAM – Non-Volatile RAM
LUN – Logical Unit Number
Op – Optical Connector
SFP – Small Form-Factor Pluggable
VAC – Volts Alternating Current

Figure C-3 Gateway specs

IBM

IBM Hardware Universe – 4 – 2010/03/10

Adapter Cards A

Type	Bus	Feature Code	Media	Data ONTAP (min release)	N series Appliance 7900	N series Appliance 7700	N series Appliance 6070	N series Appliance 6060	N series Appliance 6040	N series Appliance 3600	N series Gateway 7900	N series Gateway 7700	N series Gateway 6070	N series Gateway 6060	N series Gateway 6040
FIBRE CHANNEL	PCIe	1017 FCP TARGET Dual Port 4Gb	Op - LC	7.2	4	4	2	2	2		4	4	2	2	2
FIBRE CHANNEL	PCIe	1030 FCP TARGET Quad Port 4Gb	Op - LC	7.3	4	4	4	4	4	1	4	4	4	4	4
FIBRE CHANNEL	PCIe	1036 FCP TARGET Dual Port 8Gb	Op - LC	7.3	4	4	4	4	4	1	4	4	4	4	4
FIBRE CHANNEL	PCIe	10156 DISK/TAPE Dual Port 4Gb	Op - LC	7.2 10.0.1	3	3	4	4	4	1	3	3	4	4	4
FIBRE CHANNEL	PCIe	1029 DISK Quad Port 4Gb	Op - LC	7.2.2	5	5	4	4	4	1	5	5	4	4	4
FIBRE CHANNEL	PCIe	1035 DISK/TAPE Quad Port 4Gb	Op - LC	7.2.3	5	5	4	4	4						
FIBRE CHANNEL	PCIe	1014 DISK Dual Port 4Gb	Op - LC	7.2 10.0.1	5	5	4	4	4	1	5	5	4	4	4
ETHERNET	PCIe	1022 TOE Quad Port GbE	Cu – RJ45	7.2.1	5	5	4	4	4		5	5	4	4	4
ETHERNET	PCIe	1031 TOE Dual Port 10GbE	Op - LC	7.2.3 10.0.3	5	5	4	4	4		5	5	4	4	4
ETHERNET	PCIe	1012 NIC Dual Port GbE	Op - LC	7.2	5	5	4	4	4	1	5	5	4	4	4
ETHERNET	PCIe	1013 NIC Dual Port GbE	Cu – RJ45	7.2 10.0.1	5	5	4	4	4	1	5	5	4	4	4
ETHERNET	PCIe	1023 NIC Quad Port GbE	Cu – RJ45	7.2.1 10.0.1	5	5	4	4	4	1	5	5	4	4	4
ETHERNET	PCIe	1062 NIC Single Port 10GbE	Op - LC	7.3.2						1					
ETHERNET	PCIe	1065 NIC Dual Port 10GbE	Op - LC	7.3.2	5	5	4	4	4		5	5	4	4	4
ETHERNET	PCIe	1026 iSCSI Dual Port GbE	Cu – RJ45	7.2.1	5	5	4	4	4	1	5	5	4	4	4
ETHERNET	PCI-X	1021 iSCSI Dual Port GbE target	Op - LC	7.2.1	5	5	4	4	4	1	5	5	4	4	4
ETHERNET	PCI-X	1008 TOE Single Port 10GbE	Op - LC	7.2 10.0.1	2	2					2	2			
ETHERNET	PCI-X	1007 TOE Quad Port GbE	Cu – RJ45	7.2	2	2					2	2			
ETHERNET	PCI-X	1010 iSCSI Dual Port	Cu – RJ45	7.2	3	3					3	3			
ETHERNET	PCI-X	1011 iSCSI Dual Port	Op - LC	7.1.1	3	3					3	3			

Please refer to the IBM configuration guidelines for proper platform and expansion slot assignment information.

Figure C-4 Adapter cards

Figure C-4 shows the currently available expansion/connection cards for both FC (internal and external) and IP connectivity. It is extremely important to be aware of the potential system expandibility when you have the need to add more disk shelves or when connecting the systems to new servers (via LAN or SAN).

IBM

IBM Hardware Universe – 5 – 2010/03/10

Adapter Cards B

		Feature Code	Media	Data ONTAP (min release)	N series appliance 7900	N series appliance 7700	N series appliance 6070	N series appliance 6060	N series appliance 6040	N series appliance 3600	N series Gateway 7900	N series Gateway 7700	N series Gateway 6070	N series Gateway 6060	N series Gateway 6040
SAS / SCSI	PCIe	**1024** TAPE Dual SCSI-LVD/SE	Cu - 68p VHDCI	7.2	3	3	4	4	4	1	3	3	4	4	4
SAS / SCSI	PCIe	**1060** SAS, 2-Port, 3Gb	Cu - miniSAS	7.3.2						1					
SAS / SCSI	PCIe	**1061** SAS, 4-Port, 3Gb	Cu - QSFP	7.3.2	5	5	4	4	4		5	5	4	4	4
OTHER	PCI-X	**1016** TAPE Dual SCSI-LVD/SE	Cu - 68p VHDCI	7.0.5	3	3					3	3			
OTHER	PCIe	**1033** SnapMirror®/Fibre Channel 4Gb	Op - LC	7.2.2	2	2	1	1	1	1	2	2	1	1	1
OTHER	PCIe	**1063** CNA Dual-Port 10Gb FCoE	Op - LC	7.3.2	5	5	4	4	4	1	5	5	4	4	4
OTHER	PCIe	**10646** CNA Dual-Port 10Gb FCoE	Cu - SFP+	7.3.2	5	5	4	4	4	1	5	5	4	4	4
OTHER	PCIe	**1032** MetroCluster FC-VI 4Gb Dual	Op - LC	7.2.3	1	1	1	1	1		1	1	1	1	1
OTHER	PCIe	**1056** PAM Cache Card - 16GB	-	7.3	5	5	4	2	2		5	4	4	2	2
OTHER	PCIe	**1058** PAM II Cache Card - 256GB	-	7.3.2				2	1					2	1
OTHER	PCIe	**1057** PAM II Cache Card - 512GB	-	7.3.2	4	2	2				4	2	2		
OTHER	PCI-X	**1034** SnapMirror®/Fibre Channel 2Gb	Op - LC	6.5	2	2									

Please refer to the IBM configuration guidelines for proper platform and expansion slot assignment information.

Not Supported (empty box)   **Supported (# = max per controller)**

Terms and Abbreviations

CNA – Converged Network Adapter
Cu – Copper Connector
FC – Fibre Channel
FCP – Fibre Channel Protocol
GbE – Gigabit Ethernet
LC – Lucent Connector
NIC – Network Interface Card
NVRAM – Non-Volatile RAM
Op – Optical (connector)
PAM – Performance Acceleration Module
TOE – TCP Offload Engine

Figure C-5 Adapter B cards

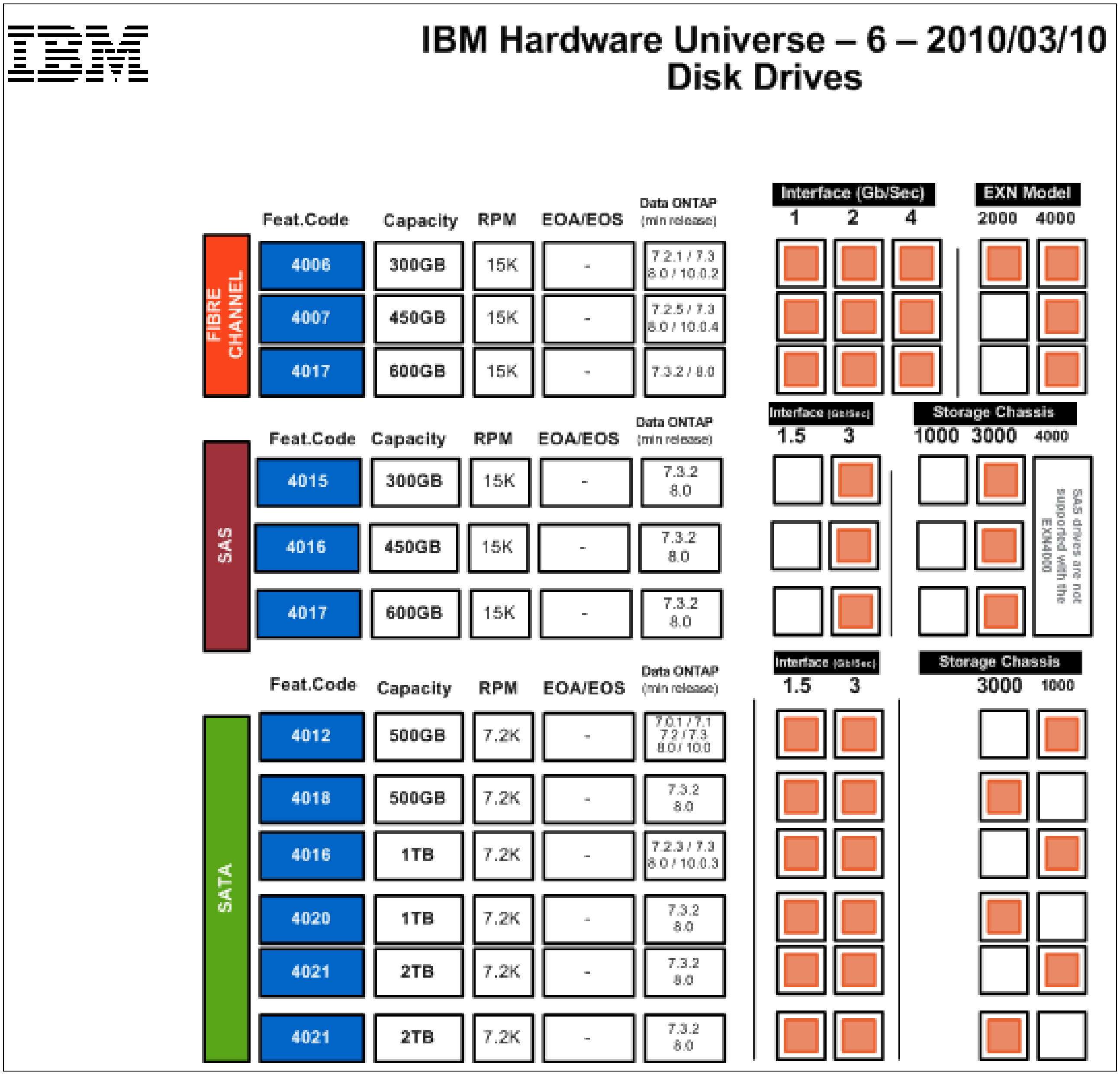
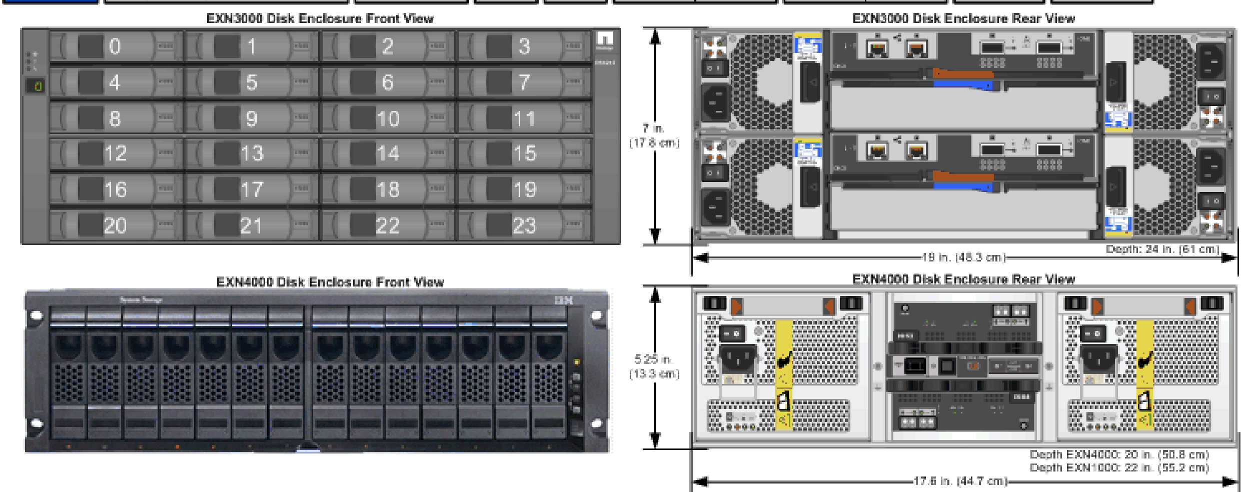


Figure C-6 Disk drive specs

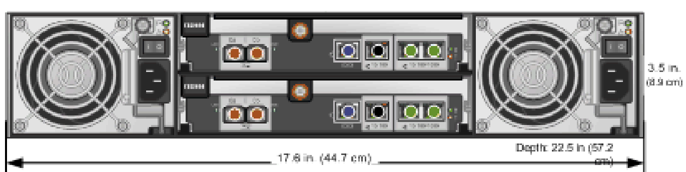
IBM Hardware Universe – 7 – 2010/03/10

Shelf Enclosures

Model	Feature Code	Weight	Disk	Rack	Power (Amps @100-120V)	Thermal (BTU/hr)	EOA/EOS	Data ONTAP (min release)
EXN2000		**Empty** 50.06 lb. (23 kg) **With Drives** 77 lb. (35 kg)	14	3U	**10K RPM** 72GB: 3.43A 144GB: 3.43A 300GB: 3.89A **15K RPM** 72GB: 3.63A 144GB: 3.75A 300GB: 4.32A	**10K RPM** 72GB: 1,167 144GB: 1,167 300GB: 1,320 **15K RPM** 72GB: 1,234 144GB: 1,272 300GB: 1,470 450GB: 1,470	05.2008 06.2013	6.4.1 / 7.X 8.0 / 10.0
EXN4000	**4006** (300GB drives) **4007** (450GB drives) **4008** (600GB drives)	**Empty** 50.06 lb. (23 kg) **W/O Drives:** 68 lb. (30.8 kg) **With Drives** 77 lb. (35 kg)	14	3U	**10K RPM** 72GB: 3.01A 144GB: 3.36A 300GB: 3.78A **15K RPM** 72GB: 3.66A 144GB: 3.45A 300GB: 4.27A 450GB: 4.27A 600GB: 4.27A	**10K RPM** 72GB: 1,020 144GB: 1,140 300GB: 1,287 **15K RPM** 72GB: 1,238 144GB: 1,174 300GB: 1,452 450GB: 1,452 600GB: 1,452	-	7.2.1 / 7.3 8.0 / 10.0.1
EXN1000	**4012** (500GB drives) **4019** (750GB drives) **4016** (1TB drives) **4021** (2TB drives)	**Empty** 50.06 lb. (23 kg) **W/O Drives:** 68 lb. (30.8 kg) **With Drives** 77 lb. (35 kg)	14	3U	290GB: 2.72A 500GB: 2.90A 750GB: 3.22A 1000GB: 3.10A 2000GB: 3.10A	290GB: 923 500GB: 1,095 750GB: 1,050 1000GB: 1,050 2000GB: 1,050	-	6.5 / 7.X 8.0 / 10.0
EXN3000	**4012** (300GB SAS) **4016** (450GB SAS) **4017** (600GB SAS) **4018** (500GB SATA) **4020** (1TB SATA) **4021** (2TB SATA)	**Empty:** 21.1 lb. (9.6 kg) **W/O Drives:** 53.7 lb. (24.4 kg) **With Drives:** 110 lb. (49.9 kg)	24	4U	**SAS** 300GB: 6.0A 450GB: 6.3A 600GB: 6.3A **SATA** 500GB: 4.4A 1000GB: 4.4A 2000GB: 4.4A	**SAS** 300GB: 2,048 450GB: 2,150 600GB: 2,150 **SATA** 500GB: 1,501 1000GB: 1,501 2000GB: 1,501	-	7.3.2 8.0



N3300 (HA configuration shown)



- Fibre Channel SFP (4)
- Gigabit Ethernet RJ45 (4)
- Remote LAN Management RJ45 (2)
- Serial Console RJ45 (2)

Figure C-7 Shelf enclosures

Related publications

The publications listed in this section are considered particularly suitable for a more detailed discussion of the topics covered in this book.

IBM Redbooks

For information about ordering these publications, see “How to get Redbooks” on page 343. Note that some of the documents referenced here may be available in softcopy only.

- *Managing Unified Storage with IBM System Storage N series Operation Manager,* SG24-7734
- *IBM System Storage N series Software Guide*, SG24-7129
- *Data Protection Strategies in IBM System Storage N Series*, SG24-7591
- *Multiprotocol Data Access with IBM System Storage N series*, REDP-4176

Other publications

These publications are also relevant as further information sources:

- *IBM System Storage N series Data ONTAP 7.3.3 Release Notes,* GC53-1175
- *IBM System Storage N series Data ONTAP 7.3 Upgrade Guide, GC27-2200*
- *IBM System Storage N series Data ONTAP 7.3 System Administration Guide,* GC52-1279

Online resources

These websites are also relevant as further information sources:

- IBM NAS support website

  `http://www.ibm.com/storage/support/nas/`
- NAS product information

  `http://www.ibm.com/storage/nas/`
- IBM Integrated Technology Services

  `http://www.ibm.com/planetwide/`

How to get Redbooks

You can search for, view, or download Redbooks, Redpapers, Technotes, draft publications and Additional materials, as well as order hardcopy Redbooks publications, at this website:

ibm.com/redbooks

Help from IBM

IBM Support and downloads

ibm.com/support

IBM Global Services

ibm.com/services

Index

A

access after failover 97
ACM 47
ACP 16
active/active configuration 51, 58, 141, 146
active/active configuration interconnect 101
active/active controller 52–53
active/active failover 106
active/active nodes 92
active/active pairs 51
active/passive configuration 56, 146
adapter cards 339
adapter settings 107
adapters 12, 98
Advanced Single Instance Storage (ASIS) 7
AIX 182, 186
AIX Host Utilities 5.0 179
AIX MPIO 179, 184
AIX System Management Interface Tool (SMIT) 182, 186
Alternate Control Path 16
alternative load-balance policy 148
architecture compatibility 54
ASCII terminal console 304
ASUP 29
asymmetrical configurations 56
asymmetrical standard active/active configuration 56
AT-FCX 45
AutoSupport 6

B

backup application 11
backup/recovery 1, 12
block and file services 19
bundled HBA driver 174

C

cache memory 9
cf status 92
CFE 19
cfmode 80, 94, 96, 98–99, 101, 103, 111, 116
  setting 94–95, 99, 109
CFO 111
CFO + FCP 56
CFO event 115
changing IP address 320
changing network configuration 319
Check 322
checking cluster status 92
CLI 157, 160–161, 173, 191
cluster 31, 50
  management 98
  node interaction 55
  status and management 49
cluster configurations
  best practices 56
cluster failover 8, 31, 50, 56
  basic operations 91
  event 106, 114
cluster interconnect card 37
cluster status 88
clustering 31
  benefits 53
  eliminating single points of failure 63
  local 249
  reasons for failover 91
Common Internet File System 7
  services 222
console port 36
controller 22, 30
controller board components 27
controller chassis 14
controller failover 51
controller LEDs 24
cooling configuration 28
cost advantages 12
CPU
  utilization 244

D

data management capabilities 39
Data ONTAP 40, 117, 146
  checking software version 228
  clustering 55
  install from Windows client 313
  updates 309
Data ONTAP 7.2 96
Data ONTAP 7.2.5 32
Data ONTAP 7.3 3
Data ONTAP 7.3 Block Access Management Guide 173
Data ONTAP Block Access Management Guide for FCP 185
Data ONTAP DSM 148, 152, 154, 156–158, 161–162, 164
Data ONTAP DSM for Windows MPIO 157
Data ONTAP DSM Management 160
Data ONTAP DSM service 156
Data ONTAP FCP 73, 84
Data ONTAP LUN 180
Data ONTAP Release Notes 54
Data ONTAP SAN Toolkit 182
data ONTAP update procedure 318–319
Data ONTAP version 167
data port 28
data protection 40
Data Rate option 170
dedicated interface 60
dedicated local interface 60

default cfmode 98
default load-balance policy 161–162
device queue 178
dialog 316
DIMMs 38
direct-attached systems 51
disk resources 10
disk sanitization 6
disk shelves 106
dm-multipath 175
dm-multipath code 177
dm-multipath devices 175
dm-multpath 178
DNS setup 321–322
Domain Name System (DNS) 6
DRAM 142
drive flexibility 10
drive types supported 10–11
DSM 146–149, 153, 155, 157, 161, 167
DSM 3.2 147
DSM ID 147, 166
dual fabric cfmode 109
dual fabric configurations 65
dual fabric mode 109
Dual Fibre Channel target cards 104
Dual gigabit Ethernet ports 15, 17
dual-controller configuration 22
dual-node clustered models 13
dynamic virtualization 38

E
e0M 28–29
Emulex 170
Emulex applications package 170
Emulex driver 169
Emulex HBA 169, 173
Emulex HBA driver 174
Emulex HBAnywhere 173
enterprise data management 32, 42
enterprise serviceability 18, 32
Ethernet internals 30
Ethernet Media Access Control 28
Ethernet ports 15
Ethernet switch 28, 31
EXN expansion units 44–47
expansion HBAs 80
expansion units 14–15, 17, 34, 40
external port 28
external QLogic HBA driver 174

F
fabric 102
fabric build sequence 65
fabric MetroCluster 124
fabric rebuild times 65
failed-over node 58
failover 52, 61
    connectivity 91
    disk mismatch 92
    performance 249
failover events 61
fault LED 22
FC
    configurations 65
    HBA for disk dual-port 126
    hop count 64
    Linux Host 167
    Linux Host Utilities kit 167
    port 106
    port LUNs 106
    switch 167
    target port 69, 80
FCoE 20
FCP 148–149, 153, 162, 182
FCP adapters 191
FCP AIX Host Utilities kit 179, 182
FCP cfmode 183
FCP cfmode single_image 183
fcp config mediatype 100
FCP SAN 109
fcstat device_map 63
feature code 9551 43
Fibre Channel 1, 23, 109, 146
    disk drives 10
    host bus adapters 151
    loop 44
    node 96
    port 14–15, 36, 102
    target 109
File Storage Resource Manager (FSRM) 8
File Transfer Protocol (FTP) 6
Filer
    booting 225
    halting 224
FilerView 6, 88, 90, 92
firmware upgrades 13
FlexCache for NFS 7
FlexClone 7
FlexScale 7, 9, 137–139, 143
FlexScale usage 143
FlexShare 6
FlexVol 6
footprint 12
FRU 22, 31–32
FRU LEDs 25

G
Gateway 42–43
    back-end implementation 40
Gateway models 42
gateway option 20
Gateway RAID 41
Gateway specs 338
GC26-7973 173
gigabit Ethernet 33
giveback operation 52

H

hard (port) zoning 99
hardware
  I/O ports 23
  LEDs 22
  N3000 13–17
  N6000 21
  system port 80 LEDs 25
hardware reference 4–5
hardware specifications 336–339
HBA 46, 73, 84, 94, 97, 102, 106, 152, 167, 170–172, 181
  brands 179
  driver 167, 169, 173
  initiator queue depth 183
  interoperability 170
  model 167
  parameter 170–171
  port 170
  settings 175
  settings option 170
hdisk 184
headless system 10
heterogeneous server 19
heterogeneous storage 38
heterogeneous storage management 39
high-performance SAS 15, 17
home directories 245
hop count 64–65
host device filename 189
Host Utilities 182
Host Utilities kit 167–168, 179–180, 182
Host Utilities kit for Native OS 181
Host Utilities kit software package 183
host-side application data 29

I

I/O ports 23
I/O throughput 142
IBM 49
IBM Path Control Module 179
ifconfig command 59
igroup add command 101
igroup set command 101
igroups 191
initiators 80
install Data ONTAP 312
install software option 182
install software screen 182
installation planning 288
interconnect 96
interconnect failure 53
internal LEDs 25
Introduction 1
IP addresses 58
IP failover 106
iSCSI 9, 146–147
iSCSI Host Attach Kit for AIX 6
iSCSI Host Attach Kit for Linux 6
iSCSI Host Attach Kit for Windows 6
iSCSI LUNS 146
iSCSI LUNs 161
iSCSI path 149
iSCSI protocol 155

L

LAN Module 25, 35
least queue depth 146–147
  I/O 148
least weighted paths 147–148
legacy Host Utilities kit 180
Link Down Timeout option 170
Linux 167
Linux host 169
Linux OS version 167
Linux parameters 178
load-balance policy 161
loader prompt 25
local access 116
local host 107
local LUNs 102
local node 51–52, 59, 63
local node console 57
local_address 59
locating a failed disk 229
logical volumes 186
Loop 170
Loop Only 171
loop storm, preventing 31
loopback interfaces 3
LUN 40, 43, 96, 101–102, 146, 148–149, 157, 161, 163–164, 167, 172, 175, 177–179, 183–184, 189–191
LUN configuration 183
LUN ID of 5 101
LUN map information 96, 101
LUN mapping 43, 101
LUN online 101
LUN size, maximum 43
LUN sizing 43

M

mailbox 53
main memory 9
manageability features 42
managed systems 12
management capabilities 13
management complexity 39
mapping conflicts 101
MaxRetriesPerIo 175
mediatype 100
metadata 9, 142
metadata blocks 142
MetroCluster 8, 117
  and N series failure 131
  benefits 119
  host failure 131
  interconnect failure 133
  site failure 134

site recovery 135
stretch 123
synchronous mirroring with SyncMirror 120
Microsoft Exchange 246
Microsoft Management Console 148
Microsoft Windows 2003 cluster 164
Microsoft Windows domain 55
midrange storage solution 146
mirrored active/active configuration 52
mirrors 24
mixed mode path access 113
MMC 158–159
Device Manager 154
services 156
model numbers 20
MPIO 147, 155, 167, 182, 184
MPIO parameter value 147
MPIO setup 185
MPIO support 167
multi-attached host 96
Multi-ID 116
Multi-ID support 116
multipath 176
multipath daemon 175, 177
multipath devices 173–174, 177
multipath information 177
multipath LUNs 177
multipath maps 175, 177
multipath service 175
multipath status 177
multipath support 178
multipath tools 168
multipathing 145
multipathing I/O 146
multipathing packages 169
multipathing software 97–98, 109, 116, 180
multipathing solution 146
multiple HBAs 179
multiprotocol connectivity 9
multiprotocol LUN 147
Multistore 7

N

N 5000 models 18
N 7000 series models 32
N series
advanced data management capabilities 38
business continuity 123
expansion unit failure 132
hardware 3
hardware quick reference
A and G models 5
A models 4
standard software features 6
starting the system 221
stopping the system 222
storage system administration 219
storage systems A models 9
N series appliance models 336
N series Gateway 38–40
N series Gateway features 40
N series hardware 3
N series LUNs 180
N series model numbers 20
N series software features 6
N_Port ID 105
N3000 11–12, 15, 17, 106
N3000 family hardware specification 14–18
N3000 system advantages 12
N3000 system features 12–13
N3000 system models 12
N3300 12–15, 17
N3300/3600 12
N3400 12, 15–16, 47
N3600 12, 14–15
N3700 3, 102, 109, 111, 116
N5000 3, 18–19, 32, 80, 116
N5000 series hardware 18–19
N5200 80
N5300 18
N5500 42
N5600 18–19, 141
N6000 9, 19–21, 27–29, 31–32, 42, 47, 69
N6000 availability 31
N6000 Gateway models 42
N6000 hardware 19, 21
N6000 models 19
N6000 reliability 31
N6000 serviceability 32
N6040 19, 69, 141
N6060 19, 42, 141
N6070 19, 42, 141
N7000 3, 9, 32–34, 42, 65, 94
N7000 Gateway models 42
N7000 hardware 33, 35, 37
N7000 models 32
N7600 32, 37, 141
N7700 32, 42–43, 65, 141
N7800 32, 37, 141
N7900 34, 37, 43, 141
native OS 182
near-line storage 11
NearStore 7
network 10, 12
network environment 60
Network Information Service (NIS) 6
network interface 3
network interface capacity 3–4
network interface limits 3
network interfaces 3–4
network-attached storage 142
network-based storage controller administration 29
NFS 1, 142
NIC 60
node status 53
Non Disruptive Update (NDU) 17
non-proxy FCP 147
NVRAM 25, 27, 31, 33, 43, 52
NVRAM LED 25
NVRAM6 37

NVRAM7 32
NVRAM7 LED 25

O

obtain Data ONTAP 310
OEM Models 170
online storage resource 2
Open System SnapVault (OSSV) 7
Operations Manager Core 8
optical jukebox 11
optimum path selection 146
optional software 7
other_options 59–60
our 318
out of sync 53
ow 315

P

PAM 9, 138, 140, 142
 configuration modes 9
 family 142
 implementation 142
 module 140
PAM II 9, 141
PAM II cards 9
partner access 116
partner cfmode 102, 109
partner command 98
partner head 106
partner LUN 96
partner LUNs 106
partner node 52–53, 58, 63
partner port 116
partner storage system 98, 111
partner_address 59
passive node 56
path failure 109
PCI 12, 37
PCI Express 37
PCI Express slots 142
PCI slots 23, 36
PCIe 12, 15, 69, 141
PCIe card 142
PCI-X 37
Performance Accelerator Module 9
performance tests 178
physical disk operations 41
physical interfaces 4
planning
 pre-installation 239
 primary issues 240
planning worksheets 289–292, 303
point-to-point mode 100
port MetroCluster adapter 126
port pairs 65, 69
PortDownRetryCount 175
PortDownRetryCount option 170
portnames 111
power cable specification 23
power cord specification 332
power on procedure 305–309
Predictive Cache Statistics 9
preferred path 147
Provisioning Manager 8
PSU 28
public loop support 116

Q

qla2xxx 171
qla2xxx directory 175
QLogic 170
QLogic bundled drivers 172
Qlogic card 107
QLogic driver 169, 171
QLogic HBA 169–171, 173
QLogic port 171
queue_depth 178

R

RAID 1, 46
RAID 4 42
RAID group 46
RAID group size 38
RAID group size, N5000 19
RAID group size, N7000 38
ramdisk 171
raw capacity 240
Redbooks Web site 343
 Contact us xii
Redundant Array of Independent Disks (see RAID) 6
reliability 18, 32
remote scripts 98
removal and insertion of the controller 28
removing hot spare disks 233
resiliency to failure 249
riser board 26–27
RLM 28–29, 31–32, 50
RLM port 33
RLM-assisted cluster failover 31
round robin 146–148
round robin with subset 148
RPM 9
rsh command 98
RTC 25

S

SAN 39, 65, 102
SAN capacity 39
SAN targets 69
SAN Toolkit files 182
sanlun command 189
SANsurfer FC 170
SANsurfer FC HBA CLI 172
SANsurfer package 169
SANsurfer Version 1.7.1 170
SAS 12, 15–17, 47
SAS disk expansion units 45

SAS Drive Bays 17
SAS firmware 17
SAS loop 44
SAS technology 47
SATA drives 11–12, 16, 18, 32, 44–45
schematic view 28
SCSI 12
secondary (proxy path) 106
SecureAdmin 6
see FC 109
see PAM 9
segregate management subnets 28
sequential scan 142
sequential workload 46
Serial-Attached SCSI (see SAS) 16
server-based flash cards 142
shared loops 56
shell connection 98
Shut 319
simple replication and disaster recovery 12
single controller 15, 17
single fabric 65
single I/O 178
Single Mailbox Recovery for Exchange (SMBR) 8
single system image 96
single_image 94, 99–100
single_image cfmode setting 96
SLC 142
SMIT 180, 182, 186–187
SMIT tool 187
smit.script 183
snap features 40
SnapDrive 7, 154–155
SnapLock 7
SnapManager 8
SnapMirror 7, 12
SnapMover 6
SnapRestore 7
  recovery 199
Snapshot 6, 12, 143
  protection consideration 242
SnapValidator 8
SnapVault 7, 40
soft zoning 99
software 7
Software Installation and Maintenance menu 180
Software Maintenance and Utilities 180
spare servers 249
split NVRAM 52
split-brain 53
SPOFs 63
staging 11
standard software, N series 6
standby access 116
standby cfmode 106
standby mode 106, 108
storage environment 1–2
storage expansion units 44
storage infrastructure 42
storage investment 10
storage management 40, 229
storage provisioning 40
storage resources 2
storage utilization 40
strong data protection 2
support matrix 170
switch 29
switch failure 98
switch failure partner mode 104
switch zoning 99
SyncMirror 7
  setup 130
sysconfig 3, 20
system activity 32
system board 32
System port 80 LEDs 25
Systems Manager 7

T

t 317
target port 65, 71, 79, 96, 170
TCP/IP 28
Telnet 6
telnet 98
thermal margin 31
thin server 10
topology 67, 71, 79, 81–82

U

unified model numbers 20
unified storage 15, 17, 19
UNIX 19
unsupported cfmode 94
update 317
upgrade process 2
UPS link 170
usable capacity 240

V

VIF 3–4, 28
VIO 181
VIO environment 182
virtual disk 162, 164
Virtual File Manager (VFM) 8
virtual local port 109
virtual partner port 109
virtual port 109, 116
virtual standby port 109
virtual storage system 53
virtualization 39
VLAN 28

W

WAFL 52, 142
WAFL cache 138–139
WAFL extended cache 142
WAFL file system
  impact of 242

WAN 12
Weighted path 146
Windows 145
windows 307
Windows 2003 148
Windows 2008 146, 148
Windows Disk Manager 163
Windows Host Utilities 152
Windows MPIO layer 148
Windows Server 2003 154
Write Anywhere File Layout (WAFL) 137
WWNN 96, 99–100, 103
WWPN 103

IBM System Storage N Series Hardware Guide

(0.5” spine)
0.475”<->0.873”
250 <-> 459 pages

IBM System Storage N Series Hardware Guide

Selecting the right N series hardware for your environment

Installing and configuring your storage systems

Managing the system and protecting your data

This IBM Redbooks publication presents a detailed look at the features, benefits, and capabilities of the IBM System Storage N series hardware offerings.

The IBM System Storage N series systems can help you tackle the challenge of effective data management using virtualization technology and a unified storage architecture. The N series delivers low- to high-end enterprise storage and data management capabilities with midrange affordability. Built-in serviceability and manageability features help support your efforts to increase reliability; simplify and unify storage infrastructure and maintenance; and deliver exceptional economy.

The IBM System Storage N series systems provide a range of reliable, scalable storage solutions to meet a variety of storage requirements. These capabilities are achieved by using network access protocols such as Network File System (NFS), Common Internet File System (CIFS), HTTP, and iSCSI, as well as storage area network technologies such as Fibre Channel (FC). Utilizing built-in Redundant Array of Independent Disks (RAID) technologies, all data is well protected with options to enhance protection through mirroring, replication, Snapshots, and backup. These storage systems are also characterized by simple management interfaces that make installation, administration, and troubleshooting straightforward.

INTERNATIONAL TECHNICAL SUPPORT ORGANIZATION

BUILDING TECHNICAL INFORMATION BASED ON PRACTICAL EXPERIENCE

IBM Redbooks are developed by the IBM International Technical Support Organization. Experts from IBM, Customers and Partners from around the world create timely technical information based on realistic scenarios. Specific recommendations are provided to help you implement IT solutions more effectively in your environment.

For more information:
ibm.com/redbooks

SG24-7840-00 ISBN 0738434744