



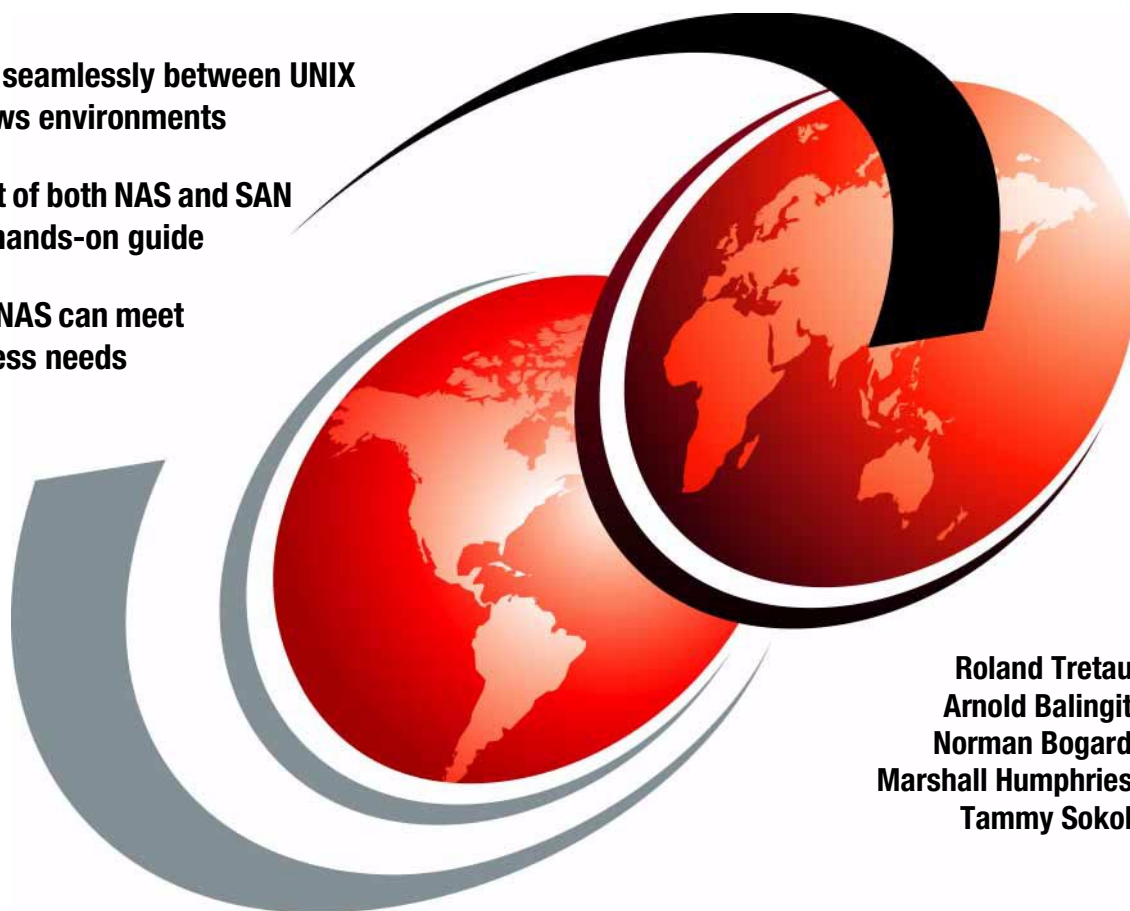
Implementing the IBM TotalStorage NAS 300G

High Speed Cross Platform Storage and Tivoli SANergy!

Share data seamlessly between UNIX
and Windows environments

Get the best of both NAS and SAN
using this hands-on guide

Learn how NAS can meet
your business needs



Roland Tretau
Arnold Balingit
Norman Bogard
Marshall Humphries
Tammy Sokol

ibm.com/redbooks

Redbooks



International Technical Support Organization

**Implementing the IBM TotalStorage NAS 300G:
High Speed Cross Platform Storage and
Tivoli SANergy!**

September 2001

Take Note! Before using this information and the product it supports, be sure to read the general information in “Special notices” on page 323.

First Edition (September 2001)

This edition applies to the IBM TotalStorage Network Attached Storage 300G with Tivoli SANergy 2.2 and running the Windows Powered OS.

Comments may be addressed to:
IBM Corporation, International Technical Support Organization
Dept. QXXE Building 80-E2
650 Harry Road
San Jose, California 95120-6099

When you send information to IBM, you grant IBM a non-exclusive right to use or distribute the information in any way it believes appropriate without incurring any obligation to you.

© Copyright International Business Machines Corporation 2001. All rights reserved.

Note to U.S Government Users – Documentation related to restricted rights – Use, duplication or disclosure is subject to restrictions set forth in GSA ADP Schedule Contract with IBM Corp.

Contents

Figures	ix
Tables	xvii
Preface	xix
The team that wrote this IBM Redbook	xx
Special notice	xxii
IBM trademarks	xxii
Comments welcome	xxiii
Part 1. Storage networking concepts and hardware	1
Chapter 1. NAS or SAN? What's all the fuss about storage networking? .	3
1.1 Local Area Networks	5
1.2 Open Systems Interconnection (OSI) model	7
1.2.1 Device driver and hardware layer	8
1.2.2 Internet Protocol layer	8
1.2.3 TCP layer	10
1.2.4 Application layer	11
1.2.5 Protocol suites	11
1.3 File systems and I/O	12
1.3.1 Network file system protocols	12
1.3.2 Understanding I/O	14
1.4 Network Attached Storage (NAS)	15
1.4.1 File servers	15
1.4.2 Network appliances	16
1.4.3 NAS uses File I/O	17
1.4.4 NAS benefits	18
1.4.5 Other NAS considerations	20
1.4.6 Total cost of ownership	22
1.5 Storage Area Networks	22
1.5.1 Overview of Fibre Channel storage networks	23
1.5.2 Fibre Channel SANs use Block I/O	26
1.5.3 SAN benefits	26
1.5.4 Other SAN considerations	29
1.5.5 Data and SAN management	30
1.6 Getting the best of both worlds: SAN with NAS	31
1.6.1 Tivoli SANergy	31
1.6.2 SANergy uses a mix of File I/O and Block I/O	34

1.6.3 SANergy benefits	35
1.6.4 SANergy considerations	36
1.6.5 Tivoli Storage Manager	37
1.6.6 SANergy with Tivoli Storage Manager	37
1.6.7 Application server-free backup and restore	38
1.6.8 SAN exploitation: LAN-free client data transfer.	38
1.7 Industry standards	38
1.7.1 Storage Networking Industry Association	39
1.7.2 Internet Engineering Task Force.	39
Chapter 2. IBM hardware overview	41
2.1 IBM TotalStorage NAS 300G	42
2.1.1 Hardware configuration for the 300G	43
2.1.2 IBM TotalStorage NAS 300G features and benefits	44
2.1.3 300G optional features	46
2.1.4 300G included software	47
2.1.5 300G preloaded and optional software	48
2.1.6 IBM NAS 300G sample connectivity	51
2.2 IBM Enterprise Storage Server.	52
2.2.1 IBM Enterprise Storage Server Overview	52
2.2.2 ESS models and expansion enclosure	53
2.2.3 ESS benefits	54
2.3 IBM Modular Storage Server	57
2.3.1 IBM Modular Storage Server overview	57
2.3.2 MSS models and expansion enclosures	58
2.3.3 MSS benefits.	59
2.4 IBM Fibre Array Storage Technology (FAStT200)	60
2.4.1 IBM FAStT200 overview	60
2.4.2 FAStT models and expansion enclosure	61
2.4.3 FAStT200 benefits	62
2.5 IBM SAN Fibre Channel Switch	63
2.5.1 Product description	65
Part 2. Implementing the NAS 300G in your storage network	69
Chapter 3. Implementing the IBM TotalStorage NAS 300G	71
3.1 Sharing SAN-based storage through the 300G	72
3.1.1 Getting started.	72
3.1.2 Re-initializing the 300G	72
3.2 To SAN or not to SAN	76
3.2.1 Finding the World Wide Name	77
3.2.2 Zoning the IBM 2109.	78
3.3 Setting up the FAStT200	87
3.4 Setting up the MSS	98

3.4.1	Failover modes and SAN zoning	99
3.4.2	Transparent failover mode and zoning	100
3.4.3	Multiple-bus failover mode and zoning	102
3.4.4	Preferred controller in multiple-bus failover mode	104
3.4.5	Setting up failover modes	106
3.4.6	Create a Logical Unit Number (LUN)	107
3.4.7	Create a RAIDset	107
3.4.8	Initialize a RAIDset	109
3.4.9	Create a partition	109
3.4.10	Assign a unit number	110
3.4.11	Define host(s) and assign logical drive(s)	112
3.4.12	Assigning a connection name	115
3.4.13	Assigning a LUN to a host	115
3.5	Setting up the ESS	116
3.5.1	Regarding SAN zoning	116
3.5.2	Setting up the ESS	118
3.6	Claiming ownership of pooled storage with the 300G	132
3.7	Sharing the SAN-based storage to LAN/WAN clients	148
3.7.1	File sharing for Windows clients	149
3.7.2	File sharing for UNIX clients	154
3.7.3	Accessing the shares from our Windows clients	157
3.7.4	Accessing the shares from our Linux/Solaris/HP-UX clients	159
3.7.5	Accessing the shares from our AIX clients	161
3.7.6	Setting up FTP access permissions on the 300G	162
3.8	User and security management on the 300G	163
3.8.1	Active Directory, NT 4 Domains, and Workgroups	163
3.8.2	UNIX NIS integration	165
3.8.3	Password synchronization	165
	Chapter 4. Clustering for high availability	167
4.1	Our environment	168
4.2	Second node first-time setup	169
4.2.1	Configure the private network adapter	169
4.2.2	Joining the domain	177
4.2.3	Update drive letters	181
4.2.4	Shut down the second node	184
4.3	First node first-time setup	184
4.3.1	Configure private network adapter	184
4.3.2	Joining the domain	186
4.3.3	Update drive letters	186
4.3.4	Restart first node	187
4.3.5	Cluster setup	187
4.4	Second node second-time setup	192

4.4.1	Add the second node to the cluster	192
4.5	Administering the cluster	194
4.5.1	Configure cluster properties	194
4.5.2	Disk group administration	197
4.5.3	Cluster resource balancing	200
4.5.4	Configure file shares	207
4.6	Client connectivity	227
4.6.1	Windows clients	228
4.6.2	UNIX clients	231
4.6.3	AIX clients	232
Chapter 5. Using SANergy to secure high-speed data sharing		235
5.1	A brief overview of Tivoli SANergy	236
5.2	Configuring the 300G as a SANergy MDC	238
5.3	Configuring your other machines to use SANergy	248
5.3.1	Installing and configuring SANergy Windows NT/2000 hosts	249
5.3.2	Installing and configuring SANergy UNIX hosts	257
5.4	Using SANergy on the 300G Model G25	260
5.4.1	Base configuration	261
5.4.2	SANergy and MSCS: mixing two domineering personalities	262
5.4.3	Installing and configuring SANergy with MSCS	262
Chapter 6. Backing up the IBM TotalStorage NAS 300G		277
6.1	The 300G and its native backup solution	278
6.1.1	300G cache and backup	278
6.1.2	Persistent Storage Manager (PSM)	279
6.1.3	Backing up PSM using the W2K Terminal Service	280
6.1.4	Archive, backup, and restoration of the 300G	286
6.1.5	NTBackup	287
6.2	Integrating the 300G with TSM	290
6.2.1	The 300G and LAN-based backup	290
6.3	TSM with SANergy	293
6.3.1	TSM backup using SANergy	295
6.4	Getting backups off the LAN: TSM with SANergy	296
6.4.1	SAN zoning	296
6.4.2	Configuring SANergy	297
6.4.3	Installing the TSM Server version 4.2	297
6.4.4	Configuring the TSM server	300
6.4.5	Installing and configuring a TSM Agent on the 300G	304
6.4.6	Configuring a TSM Agent	305
6.4.7	Installing a TSM client	307
6.4.8	Backup/Restore for the 300G with TSM and SANergy	308
6.4.9	Backup results	313

6.5 Recovering the 300G	314
6.5.1 Recovering the 300G operating system using the recovery CD . . .	314
6.5.2 300G recovery method with TSM	314
6.6 NAS and the Network Data Management Protocol	316
6.6.1 NDMP overview	316
6.6.2 Tape library setup	316
6.6.3 How TDP for NDMP backs up the NAS	318
Related publications	319
IBM Redbooks	319
Other resources	320
Referenced Web sites	321
How to get IBM Redbooks	322
IBM Redbooks collections.	322
Special notices	323
Glossary	325
Abbreviations and acronyms	331
Index	339

Figures

0-1	From left to right: Roland, Tammy, Norman, Marshall, and Arnold	xx
1-1	Bus topology	5
1-2	Ring topology	6
1-3	Star topology	6
1-4	Comparing the Internet protocol suite with the OSI reference model . . .	7
1-5	Layering and encapsulation	12
1-6	The role of the NAS 300G in your storage network	16
1-7	NAS devices use File I/O	18
1-8	SAN — the network behind the servers	23
1-9	SAN uses Block I/O	26
1-10	SANergy configuration	32
1-11	SANergy data flow	34
2-1	The 300G models G00 and G25	42
2-2	Visualization of the 300G's interoperability features	45
2-3	300G connectivity with ESS, FastT200/500, MSS, and 7133	51
2-4	ESS overview	52
2-5	ESS models	53
2-6	Disaster Recovery and Availability	55
2-7	IBM Modular Storage Server overview	58
2-8	MSS models	59
2-9	FastT200 overview	61
2-10	FastT200 models	62
2-11	IBM SAN Fibre Channel Switch 2109-S08 (top), 2109-S16 (bottom) . .	64
2-12	IBM SAN Fibre Channel Switch 2109-S16	65
2-13	Front panel of the IBM SAN Fibre Channel Switch 2109-S16	66
2-14	IBM SAN Fibre Channel Switch 2109-S08	66
2-15	Front panel of the IBM SAN Fibre Channel Switch 2109-S08	66
3-1	The IBM Advanced Appliance Configuration Utility	73
3-2	Initial setup of the remote Web Management application	74
3-3	The remote Web Management application	75
3-4	The FAStT Check application running on the NAS 300G	77
3-5	Error message seen occasionally in the FAStT Check application	78
3-6	The 300G's WWN displayed in the FAStT Check application	78
3-7	Fabric View of the 2109	79
3-8	Zone login	79
3-9	Rename Alias	80
3-10	Locate WWN	80
3-11	Add members	81

3-12	Zone creation	82
3-13	Create Zone	82
3-14	Select alias	83
3-15	Add alias to zone	84
3-16	Select zone to add	85
3-17	Add zone to configuration	86
3-18	Storage Manager 7 main screen	87
3-19	Subsystem management	88
3-20	Create Array	88
3-21	RAID level selection	89
3-22	Array Creation Successful	90
3-23	Specify Logical Drive Parameters	91
3-24	Logical Drive Creation Successful dialog	92
3-25	Storage Partitioning	93
3-26	Configure Storage Partitions	93
3-27	Define host group	94
3-28	Name host group	94
3-29	Define new host	95
3-30	Name new host	95
3-31	Define host port	96
3-32	Enter WWN	96
3-33	Define New Host Group Mapping	97
3-34	Set LUN	97
3-35	Confirm LUN Mapping	98
3-36	Zoning in transparent failover mode	101
3-37	Typical SAN Configuration in multiple-bus failover mode	103
3-38	Zones in multiple-bus failover mode	104
3-39	Zones and preferred pathing in multiple-bus failover mode	105
3-40	The add raidset command	108
3-41	The initialize raidset command	109
3-42	The create partition command	110
3-43	The add unit command	111
3-44	The rename connection command	115
3-45	The set operating system command	115
3-46	The enable access command	116
3-47	ESS internals	117
3-48	ESS preparation for the 300G host	118
3-49	StorWatch ESS Specialist home page	119
3-50	Storage Allocation panel	120
3-51	Open Systems Storage panel	121
3-52	Modify Host Systems panel	122
3-53	Added host systems	123
3-54	Host modification in progress	123

3-55	Configure host adapter ports	124
3-56	Fixed block storage groups	125
3-57	Define fixed block storage (RAID array)	126
3-58	Time intensive action warning	126
3-59	Add volumes panel	127
3-60	Add volumes to selected host	128
3-61	Select number and size of LUNs	129
3-62	New volumes created	130
3-63	Modify volume assignments.	131
3-64	Validate volume assignment modification	132
3-65	Windows Powered Server Appliance Tasks	133
3-66	Disks and Volumes page	134
3-67	Terminal Services Client login	135
3-68	Terminal Services Client	136
3-69	Disk management	137
3-70	Select disks	138
3-71	Do not upgrade disk.	139
3-72	Complete Write Signature Wizard	140
3-73	Create partition	141
3-74	Create Partition Wizard	142
3-75	Select Partition Type	143
3-76	Specify Partition Size	144
3-77	Assign Drive Letter.	145
3-78	Format Partition	146
3-79	Complete the Create Partition Wizard	147
3-80	Healthy disks	148
3-81	Setting up a share	149
3-82	Administrative Share	150
3-83	Getting rid of the administrative share	151
3-84	Establishing a share.	152
3-85	Permissions for the Windows share.	153
3-86	Permissions for the NFS share	154
3-87	Setting share permissions for UNIX clients on the 300G.	155
3-88	Shared directory.	156
3-89	Map Network Drive	157
3-90	Windows mapping information.	158
3-91	One of our shared drives	158
3-92	Adding the 300G's shared disk to the Linux fstab file	159
3-93	Mounting the 300G's shared directory from a Linux client.	160
3-94	User mapping administration	162
3-95	Setting up network identification for Active Directory	164
4-1	Our clustered environment.	168
4-2	LAN connection properties selection	169

4-3	LAN connection properties page	170
4-4	Verify External PHY property	171
4-5	Verify Full Duplex property.	172
4-6	Verify IP Mode property	173
4-7	Configure private network IP settings	174
4-8	Disable NetBIOS over TCP/IP	175
4-9	Empty primary WINS address confirmation	175
4-10	Network connections advanced settings selection	176
4-11	Connections ordering.	177
4-12	Just say No to a reboot	177
4-13	My Computer properties selection	178
4-14	System properties network identification	178
4-15	Change the computer name and update domain.	179
4-16	Enter name and password panel	180
4-17	Node joined the domain successfully.	180
4-18	Disk storage view.	181
4-19	Change Drive Letter selection	182
4-20	Drive letter display	183
4-21	Assign new drive letter.	183
4-22	Change drive letter confirmation	183
4-23	Shut down	184
4-24	Node 1 private network IP properties.	185
4-25	Disk storage view.	186
4-26	Restart Node 1.	187
4-27	Cluster setup selection.	188
4-28	Cluster Configuration Wizard	189
4-29	Select node	189
4-30	Cluster information.	190
4-31	Cluster information confirmation	191
4-32	Cluster administrator window	191
4-33	Cluster setup selection.	192
4-34	Joining node selection	193
4-35	First node information	193
4-36	Cluster administrator	194
4-37	Cluster properties.	195
4-38	Cluster properties.	196
4-39	Update network priority	197
4-40	Change disk groups.	198
4-41	Moved disk.	199
4-42	Delete disk group.	199
4-43	Delete group confirmation	200
4-44	Move group	201
4-45	Moved group	202

4-46	Disk group 1 preferred owners	203
4-47	Disk group 2 preferred owners	204
4-48	Failover options	205
4-49	Failback options	206
4-50	File share resource dependencies	207
4-51	Create new resource selection	208
4-52	Create a new IP resource	209
4-53	Possible owners	210
4-54	Dependencies	211
4-55	Resource parameters	212
4-56	IP resource created successfully	212
4-57	Bring IP resource online	213
4-58	New network name resource	214
4-59	Possible owners	214
4-60	Network name dependencies	215
4-61	Network name parameters	216
4-62	Create network name successful	216
4-63	Bring network name resource online	217
4-64	New file share resource	218
4-65	Possible owners	219
4-66	File share dependencies	220
4-67	File share parameters	221
4-68	File share permissions	222
4-69	File share advanced settings	222
4-70	Create file share successful	223
4-71	Bring file share online	223
4-72	Create new NFS resource	224
4-73	New NFS resource dependencies	225
4-74	New NFS resource parameters	226
4-75	New NFS share permissions	226
4-76	Bring new NFS share online	227
4-77	Map network drive	228
4-78	Select network drive letter and location	229
4-79	Drive connected successfully	230
4-80	Directory listing of files on T:	231
5-1	SANergy configuration	237
5-2	Verifying volume access on the 300G	239
5-3	SANergy bus management check application	240
5-4	Select managed buses for Windows MDC	241
5-5	Checking for the current device owners on the managed bus	242
5-6	Assigning the 300G as the device owner	243
5-7	Displaying the available volumes	244
5-8	Selecting the 300G to be the MDC for the selected volume	245

5-9	Making the 300G the MDC for the shared volumes it owns.	246
5-10	Performance Tester tab in the SANergy Setup Tool on the 300G. . . .	247
5-11	The 300G configured to share SAN-based storage as an MDC	249
5-12	Disk Administrator view before connecting the host to the SAN	250
5-13	Disk Administrator view showing the new device (attached to SAN) . .	251
5-14	Unassigning drive letters	252
5-15	Disk Administrator view after unassigning drive letter	252
5-16	Initial SANergy installation screen	253
5-17	Select components to install	254
5-18	SANergy configuration on a Windows host	255
5-19	The 300G is already identified as the MDC for the shared volume . . .	256
5-20	Verifying the host installation with the performance tester.	257
5-21	Tivoli SANergy main screen on Linux	258
5-22	Mount point before fusing	259
5-23	Fused mount point	259
5-24	The SANergy Performance Tester running on a Linux host	260
5-25	Testing cluster configuration	261
5-26	Deleting Physical Disk resources from MSCS	263
5-27	Special names on volume assignment	265
5-28	Installing SANergy MSCS on the first Node.	266
5-29	Installing SANergy MSCS on the last node in the cluster	267
5-30	Last window of SANergy MSCS install on final cluster node	267
5-31	Validate that the SANergy Volume resource type is available.	268
5-32	Adding a new cluster resource.	269
5-33	Defining a new SANergy Volume resource	270
5-34	Setting SANergy Volume resource parameters	271
5-35	Bringing SANergy Volume resources online	272
5-36	SANergy volume file share dependencies.	273
5-37	Validating the installation using SANergy Setup	274
5-38	Move MSCS group to test failover	275
5-39	MSCS group going offline on current node	275
5-40	MSCS group online on the other node.	276
6-1	The Windows 2000 Terminal Service	281

6-12	The TSM wizard window	301
6-13	TSM server initialization wizard	302
6-14	The Device Configuration Wizard	303
6-15	Tivoli Storage Agent Install Window	304
6-16	Storage Agent initialization wizard	306
6-17	TSM server as the MDC	309
6-18	TSM backup test data	310
6-19	Result with SANergy running	311
6-20	300G as the MDC	312
6-21	The backup results of setup 2	313
6-22	NAS and TSM interaction with NDMP	317

Tables

2-1	Hardware configurations for the G00 and G25	44
2-2	300G features and benefits	45
2-3	300G software	47
3-1	Maximum supported hosts for the MSS.....	114
4-1	Adapter properties	185

Preface

This IBM Redbook describes how to install and configure the very latest IBM storage solution and concept, the IBM TotalStorage Network Attached Storage 300G, in heterogeneous environments.

The 300G series is an innovative Network Attached Storage (NAS) appliance that connects clients and servers on an IP network to Fibre Channel storage, efficiently bridging the gap between LAN storage needs and SAN storage capacities. The 300G is a storage solution for Linux/UNIX and Windows NT/2000 environments. In this book, we show you how to integrate the 300G and explain how it will benefit your company's business needs.

This book is an easy-to-follow guide which describes the market segment that the 300G is aimed at, and explains NAS installation, ease-of-use, remote management, expansion capabilities, high availability (clustering), and backup and recovery techniques. It also explains cross platform storage concepts and methodologies for common data sharing for Linux/UNIX and Windows NT/2000 environments.

This book makes use of the IBM NAS initiative in the marketplace and defines its position and value-add. We show how the reliability, availability, scalability, and security of the 300G have the potential to be at the heart of an enterprise's data storage system. We also provide concrete scenarios for extending the reach of Storage Area Networks (SANs) to IP Networks.

The team that wrote this IBM Redbook

This IBM Redbook was produced by an international team of specialists from around the world working at the International Technical Support Organization, San Jose Center.



Figure 0-1 From left to right: Roland, Tammy, Norman, Marshall, and Arnold

Roland Tretau is a Project Leader at the International Technical Support Organization, San Jose Center. Before joining the ITSO in April 2001, Roland worked in Germany as an IT Architect for Cross Platform Solutions and Microsoft Technologies. He holds a Masters degree in Electrical Engineering with a focus in Telecommunications.

Norman Bogard is a member of the IBM Storage Systems Group's Advanced Technical Support team where he specializes in SAN and NAS solutions. Norman has been working with SANs since 1996 in the NUMA-Q division (formerly Sequent Computer Systems), and prior to that, he spent 15 years as a Customer Engineer supporting various networks and systems.

Tammy Sokol is an Advisory Software Engineer who has been with IBM since 1995. She holds a degree in Computer Science from the University of Texas at Austin. Currently, she is transitioning to become a Project Manager of Midrange Storage Products. Prior to this assignment, she spent 4 years in enterprise subsystems development and test.

Marshall Humphries joined IBM in 1999 and is a Software Engineer for Tivoli Storage. Currently, he is an Information Developer working with the Tivoli Storage Network Manager and Tivoli Storage Manager development teams. Prior to his tenure at IBM, he worked as a Network Administrator in Hawaii. He is interested in usability and interface design and has a Masters degree in English as a Second Language, with a focus on second language reading and testing. His background includes teaching and Japanese studies.

Arnold Balingit works for IBM Global Services in the Philippines. He is currently the Project Manager for the Business Continuity and Recovery Services (BCRS) and concurrently the Level 2 Country Support for Tivoli Storage Manager. His areas of expertise include the implementation of Tivoli Storage Manager and doing consulting engagements for BCRS.

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

Emma Jacobs, Yvonne Lyon, Deanna Polm, Jon Tate, Barry Kadleck
International Technical Support Organization, San Jose Center

Megan E Kirkpatrick, Paula Koziol, Rebecca Witherspoon, Margaret Lukowski,
Eric Dunlap, Bob Moon, Vickie Perris, Rainer Wolafka, Jay Knott, Rich McCartt,
Ken Quarles, Kirk Kashgagian, Gene Graham
IBM US

Arwed Tschoeke, Craig McKenna, Ka Yee Chan, Wayne Gorton, Corine Ross
Fellow ITSO Redbook Residents


Armando Lemos
Maxblue by Deutsche Bank


Special notice

This publication is intended to help network or storage administrators to install and configure the IBM NAS 300G with Tivoli SANergy and Tivoli Storage Manager software. The information in this publication is not intended as the specification of any programming interfaces that are provided by Tivoli SANergy, Tivoli Storage Manager, or the IBM TotalStorage NAS 300G. See the PUBLICATIONS section of the IBM Programming Announcement for Tivoli SANergy, Tivoli Storage Manager, and the IBM TotalStorage NAS 300G for more information about what publications are considered to be product documentation.

IBM trademarks

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

e (logo)® 
IBM ®

Redbooks
Redbooks Logo 

AFS
AIX
AIX 5L
Alert on LAN
APPN
AT
Cross-Site
CT
Current
DATABASE 2
DB2
DFS
DFSMSdss
DYNIX
Early
Enterprise Storage Server
ESCON
FICON
FlashCopy
GDPS
IBM TotalStorage
Magstar
Manage. Anything. Anywhere.
Micro Channel
Netfinity
NetView
NUMA-Q

PAL
Parallel Sysplex
Planet Tivoli
PowerPC
RACF
RAMAC
RMF
RS/6000
S/390
SANergy
Seascape
Sequent
Shark
SP
StorWatch
TCS
Tivoli
Tivoli Certified
Tivoli Enterprise
Tivoli Ready
TME
xSeries
z/OS
z/VM
zSeries
400

Comments welcome

Your comments are important to us!

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

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

ibm.com/redbooks

- ▶ Send your comments in an Internet note to:

redbook@us.ibm.com

- ▶ Mail your comments to the address on page ii.



Part 1

Storage networking concepts and hardware

Part 1 of this book introduces the basic concepts of storage networking. It covers NAS and SAN concepts, provides a high-level overview of the Tivoli SANergy software, and introduces the hardware we used in this book, including the IBM TotalStorage NAS 300G. If you are already familiar with these concepts, or if you are anxious to get to the meat of this book, please jump ahead to Part 2, “Implementing the NAS 300G in your storage network” on page 69.



NAS or SAN? What's all the fuss about storage networking?

Given the expansive growth in both storage and network technology, it is only natural that the two would form a union powerful enough to change the way we think about data storage. However, the advent of networked storage solutions created a tendency to view Networked Attached Storage (NAS) and Storage Area Networks (SAN) as competing technologies within the market.

This is partly due to confusion regarding how to apply each technology. After all, both terms include the words *storage* and *network*. The problem to be solved is how to connect lots of *storage* to lots of servers and have that storage be accessible to everyone.

The natural solution to use to resolve this problem is a network. However, the implementations of NAS and SAN are very different. NAS exploits the existing intermediate speed messaging network, whereas the SAN solution uses high speed network channel technology.

In this book, we focus on NAS as a storage networking solution. Reading this book should adequately equip you to implement a NAS solution using one or more of the products we describe to meet your networked storage requirements.

First we provide the concepts and technical knowledge needed (Chapter 1, “NAS or SAN? What’s all the fuss about storage networking?” on page 3). Next we offer a brief overview of the IBM products we used (Chapter 2, “IBM hardware overview” on page 41). Then we describe how to integrate the 300G into your storage network to help you bridge the gap between your LAN-based and SAN-based storage (Chapter 3, “Implementing the IBM TotalStorage NAS 300G” on page 71). Next we explain how to set up the NAS 300G Model G25 with clustering (Chapter 4, “Clustering for high availability” on page 167). Then we show how you can use the Tivoli SANergy software that is pre-installed on the 300G to securely manage high speed data sharing in your network (Chapter 5, “Using SANergy to secure high-speed data sharing” on page 235). Finally, we integrate Tivoli Storage Manager into this picture to show how it can be used in a LAN and server-free backup solution for this network structure (Chapter 6, “Backing up the IBM TotalStorage NAS 300G” on page 277).

Most of this book is a hands-on guide to implementing the 300G as part of a storage networking solution, but before we can leap into the how-to section, it is important that you understand a few of the basic concepts about networks and storage.

Note: If you are a seasoned storage networking professional and are already very familiar with this subject, feel free to skip ahead to Chapter 3, “Implementing the IBM TotalStorage NAS 300G” on page 71. However, if you would like a quick primer, please read these first two chapters. They provide the background information you need to understand, not only how to proceed with the integration, but also what you stand to gain from doing so.

1.1 Local Area Networks

A Local Area Network (LAN) is simply the connection of two or more computers (nodes) to facilitate data and resource sharing. They proliferated from the mid-1980s to address the problem of “islands of information” which occurred with standalone computers within departments and enterprises. LANs typically reside in a single or multiple buildings confined to a limited geographic area which is spanned by connecting two or more LANs together to form a Wide Area Network (WAN).

The design of LANs are based typically on open systems networking concepts. These concepts are described in the network model of the Open Systems Interconnection (OSI) standards of the International Standards Organization (ISO). The OSI model is described in detail in 1.2, “Open Systems Interconnection (OSI) model” on page 7.

LAN types are defined by their topology, which is simply how nodes on the network are physically connected together. A LAN may rely on a single topology throughout the entire network but typically has a combination of topologies connected using additional hardware. The primary topologies defined for Local Area Networks are:

Bus topology

In a bus topology, all nodes are connected to a central cable, called the bus or backbone. Bus networks are relatively inexpensive and easy to install. Ethernet systems use a bus topology (Figure 1-1).

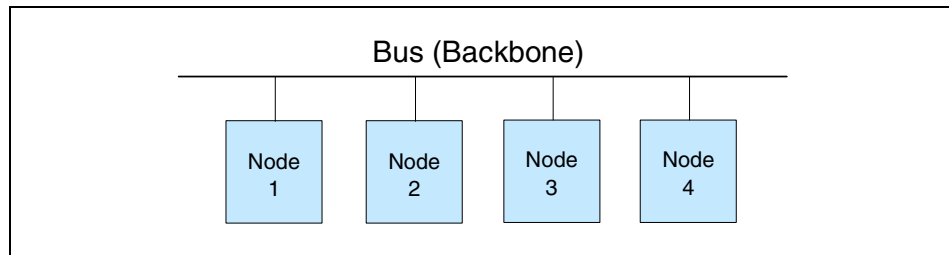


Figure 1-1 Bus topology

Ring topology

Nodes in a ring topology are connected via a closed loop such that each node has two other nodes connected directly to either side of it. Ring topologies are more costly and can be difficult to install. IBM's Token Ring uses a ring topology (Figure 1-2).

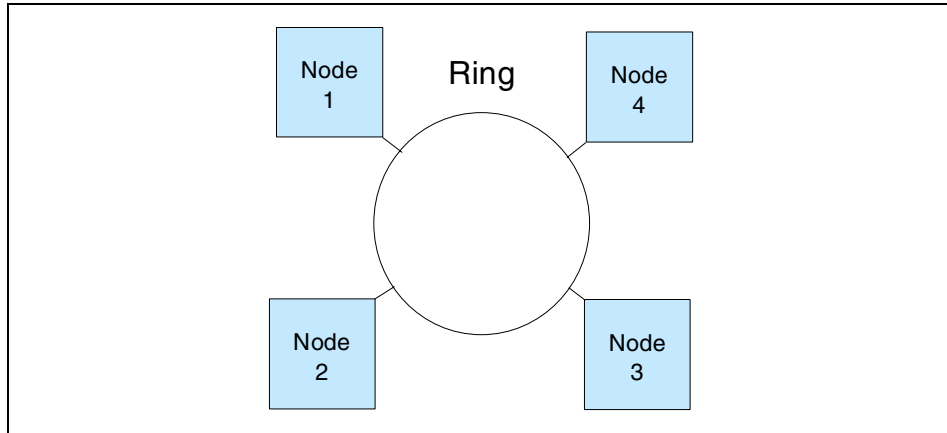


Figure 1-2 Ring topology

Star topology

A star topology uses a centralized hub to connect the nodes in the network together. Star networks are easy to install and manage however, bottlenecks occur since all of the network traffic travels through the hub. Ethernet systems also use a star topology (Figure 1-3).

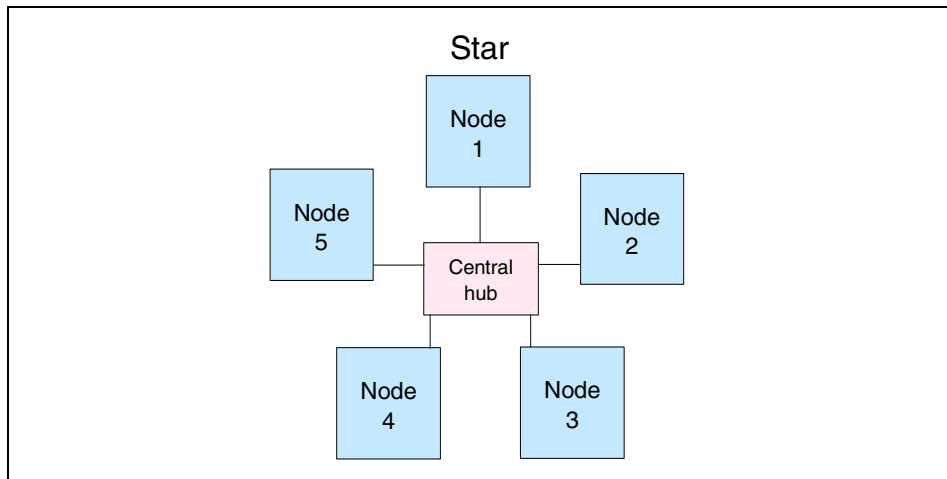


Figure 1-3 Star topology

Today, Ethernet topologies are predominant. International Data Corporation (IDC) estimates more than 85% of all installed network connections worldwide are Ethernet. It is popular due to its simplicity, affordability, scalability, and manageability. Ethernet includes definitions of protocols for addressing, formatting and sequencing of data transmissions across the network and also describes the physical media (cables) used for the network.

1.2 Open Systems Interconnection (OSI) model

The Open Systems Interconnection (OSI) model describes the layers in the network required for communication between computers. OSI is a seven layered model illustrated with the Internet protocol suite (or stack) in Figure 1-4. Each layer is responsible for a certain set of tasks associated with moving data across the network. Most Ethernet networks (including ours) communicate using the TCP/IP protocol. In this section, we discuss TCP/IP and how it relates to the OSI model since it is the default communication protocol for the 300G.

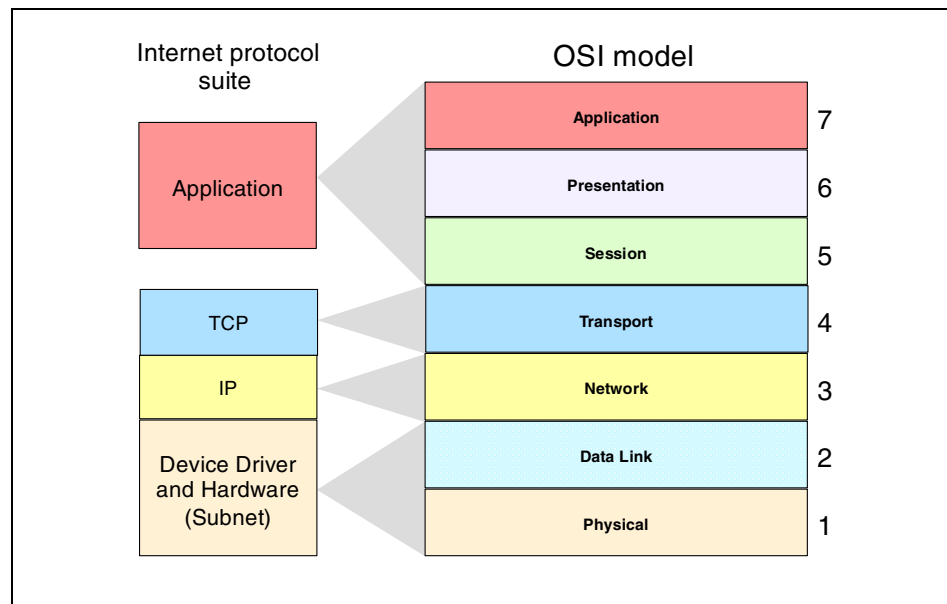


Figure 1-4 Comparing the Internet protocol suite with the OSI reference model

1.2.1 Device driver and hardware layer

Also called the Subnet layer, the device driver and hardware layer comprises both the physical and data link layers of the OSI model. It is considered the hardware that is part of each node on the network. The hardware handles the electrical and mechanical aspects of data transfers, moving the bits across a physical link. The data link layer packages packets of data into frames, ensures that they arrive safely to the target destination, and encompasses error detection and correction.

1.2.2 Internet Protocol layer

In the OSI model, the Network layer finds the best route through the network to the target destination. It has little to do in a single discrete LAN; but in a larger network with subnets, or access to WAN's, the Network layer works with the various routers, bridges, switches, gateways, and software, to find the best route for data packets.

The Internet Protocol (IP) layer in the Internet protocol suite performs the functions of the network layer. It is the common thread running through the Internet and most LAN technologies, including Ethernet. It is responsible for moving data from one host to another, using various "routing" algorithms. Layers above the network layer break a data stream into chunks of a predetermined size, known as packets or datagrams. The datagrams are then sequentially passed to the IP layer.

The job of the IP layer is to route these packets to the target destination. IP packets consist of an IP header, together with the higher level TCP protocol and the application datagram. IP knows nothing about the TCP and datagram contents. Prior to transmitting data, the network layer might further subdivide it into smaller packets for ease of transmission. When all the pieces finally reach the destination, they are reassembled by the network layer into the original datagram.

IP connectionless service

The IP is the standard that defines the manner in which the network layers of two hosts interact. These hosts may be on the same network, or reside on physically remote heterogeneous networks. IP was designed with inter-networking in mind. It provides a connectionless, best-effort packet delivery service. Its service is called connectionless because it is like the postal service rather than the telephone system. IP packets, like telegrams or mail, are treated independently. Each packet is stamped with the addresses of the receiver and the sender.

Routing decisions are made on a packet-by-packet basis. On the other hand, connection-oriented, circuit switched telephone systems explicitly establish a connection between two users before any conversation takes place. They also maintain the connection for the entire duration of conversation.

A best-effort delivery service means that packets might be discarded during transmission, but not without a good reason. Erratic packet delivery is normally caused by the exhaustion of resources, or a failure at the data link or physical layer. In a highly reliable physical system such as an Ethernet LAN, the best-effort approach of IP is sufficient for transmission of large volumes of information. However, in geographically distributed networks, especially the Internet, IP delivery is insufficient. It needs to be augmented by the higher-level TCP protocol to provide satisfactory service.

The IP packet

All IP packets or datagrams consist of a header section and a data section (payload). The payload may be traditional computer data, or it may, commonly today, be digitized voice or video traffic. Using the postal service analogy again, the “header” of the IP packet can be compared with the envelope and the “payload” with the letter inside it. Just as the envelope holds the address and information necessary to direct the letter to the desired destination, the header helps in the routing of IP packets.

The payload has a maximum size limit of 65,536 bytes per packet. It contains error and/or control protocols, like the Internet Control Message Protocol (ICMP). To illustrate control protocols, suppose that the postal service fails to find the destination on your letter. It would be necessary to send you a message indicating that the recipient's address was incorrect. This message would reach you through the same postal system that tried to deliver your letter. ICMP works the same way: It packs control and error messages inside IP packets.

IP addressing

An IP packet contains a source and a destination address. The source address designates the originating node's interface to the network, and the destination address specifies the interface for an intended recipient or multiple recipients (for broadcasting).

Every host and router on the wider network has an address that uniquely identifies it. It also denotes the sub-network on which it resides. No two machines can have the same IP address. To avoid addressing conflicts, the network numbers are assigned by an independent body.

The network part of the address is common for all machines on a local network. It is similar to a postal code, or zip code, that is used by a post office to route letters to a general area. The rest of the address on the letter (i.e., the street and house number) are relevant only within that area. It is only used by the local post office to deliver the letter to its final destination.

The host part of the IP address performs a similar function. The host part of an IP address can further be split into a sub-network address and a host address.

Time to Live (TTL)

The IP packet header also includes Time to Live (TTL) information that is used to limit the life of the packet on the network. It includes a counter that is decremented each time the packet arrives at a routing step. If the counter reaches zero, the packet is discarded.

1.2.3 TCP layer

The transport layer is responsible for ensuring delivery of the data to the target destination, in the correct format in which it was sent. In the event of problems on the network, the Transport layer finds alternative routes. It is also responsible for delivering the sequence of packets in the correct order. In the Internet protocol suite, the protocol operating in the transport layer is the Transmission Control Program (TCP).

The application data has no meaning to the Transport layer. On the source node, the transport layer receives data from the application layer and splits it into data packets or chunks. The chunks are then passed to the network layer. At the destination node, the transport layer receives these data packets and reassembles them before passing them to the appropriate process or application.

The Transport layer is the first end-to-end layer of the TCP/IP stack. This characteristic means that the transport layer of the source host can communicate directly with its peer on the destination host, without concern about 'how' data is moved between them. These matters are handled by the network layer. The layers below the transport layer understand and carry information required for moving data across links and subnetworks. In contrast, at the transport layer or above, one node can specify details that are only relevant to its peer layer on another node. For example, it is the job of the transport layer to identify the exact application to which data is to be handed over at the remote end. This detail is irrelevant for any intermediate router. But it is essential information for the transport layers at both the ends.

1.2.4 Application layer

The functions of the Session, Presentation, and Application layers of the OSI model are all combined in the Application layer of the Internet protocol suite. It encompasses initial logon, security, final termination of the session, interpretation services (compression, encryption, or formatting), and delivery of the network messages to the end user program.

The Application layer is the layer with which end users normally interact. It is responsible for formatting the data so that its peers can understand it. Whereas the lower three layers are usually implemented as a part of the OS, the application layer is a user process. Some application-level protocols that are included in most TCP/IP implementations, include:

- ▶ Telnet for remote login
- ▶ File Transfer Protocol (FTP) for file transfer
- ▶ Simple Mail Transfer Protocol (SMTP) for mail transfer

1.2.5 Protocol suites

A protocol suite (or protocol stack), as we saw in the Internet protocol suite, is organized so that the highest level of abstraction resides at the top layer. For example, the highest layer may deal with streaming audio or video frames, whereas the lowest layer deals with raw voltages or radio signals. Every layer in a suite builds upon the services provided by the layer immediately below it.

Note: You may see the different terms Internet protocol suite, *TCP/IP suite*, or *TCP/IP stack*. These are simply names for the same thing, the group of network layers to describe how two nodes on the Internet communicate.

The terms protocol and service are often confused. A *protocol* defines the exchange that takes place between identical layers of two hosts. For example, in the IP suite, the transport layer of one host talks to the transport layer of another host using the TCP protocol. A *service*, on the other hand, is the set of functions that a layer delivers to the layer above it. For example, the TCP layer provides a reliable byte-stream service to the application layer above it.

Each layer adds a header containing layer-specific information to the data packet. A header for the network layer might include information such as source and destination addresses. The process of appending headers to the data is called encapsulation. Figure 1-5 shows how data is encapsulated by various headers. During de-encapsulation the reverse occurs; the layers of the receiving stack extract layer-specific information and process the encapsulated data accordingly. The process of encapsulation and de-encapsulation increases the overhead involved in transmitting data.

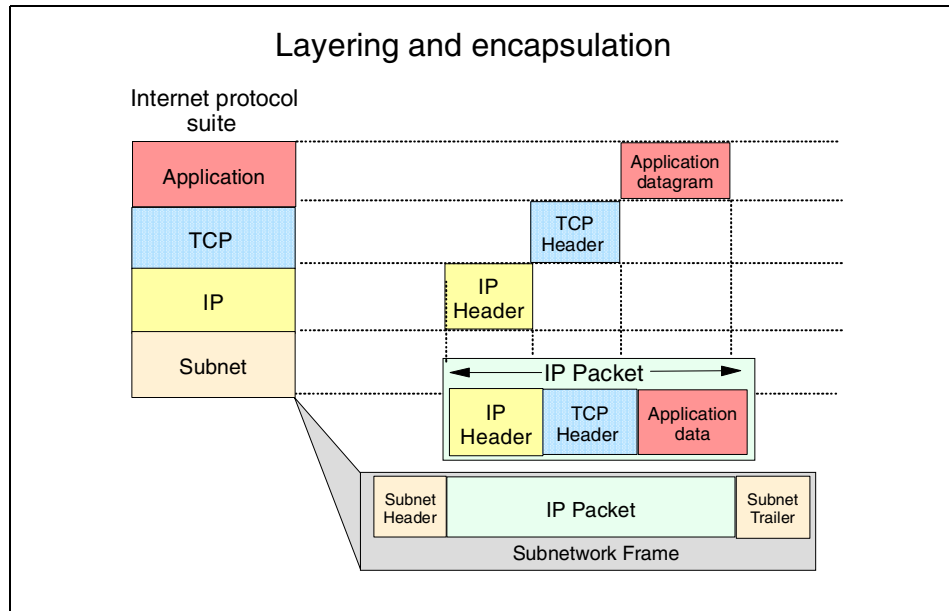


Figure 1-5 Layering and encapsulation

1.3 File systems and I/O

In this section we describe the most common file level protocols and attempt to untangle the confusion surrounding the various I/O concepts.

1.3.1 Network file system protocols

The two most common file level protocols used to share files across networks are Network File System (NFS) for UNIX and Common Internet File System (CIFS) for Windows. Both are network based client/server protocols which enable hosts to share resources across a network using TCP/IP. Users manipulate shared files, directories and devices such as printers, as if they were locally on or attached to the user's own computer. The 300G is preconfigured to support both NFS and CIFS.

Network File System (NFS)

NFS servers make their file systems available to other systems in the network by *exporting* directories and files over the network. Once exported, an NFS client can then “mount” a remote file system from the exported directory location. NFS controls access by giving client-system level user authorization based on the assumption that a user who is authorized to the system must be trustworthy. Although this type of security is adequate for some environments, it is open to abuse by anyone who can access a UNIX system via the network.

For directory and file level security, NFS uses the UNIX concept of file permissions with *User* (the owner’s ID), *Group* (a set of users sharing a common ID), and *Other* (meaning all other user IDs). For every NFS request, the IDs are verified against the UNIX file permissions.

NFS is a *stateless* service. Therefore, any failure in the link will be transparent to both client and server. When the session is re-established the two can immediately continue to work together again.

NFS handles file locking by providing an *advisory lock* to subsequent applications to inform them that the file is in use by another application. The ensuing applications can decide if they want to abide by the lock request or not. This has the advantage of allowing any UNIX application to access any file at any time, even if it is in use. The system relies on “good neighbor” responsibility which, though often convenient, clearly is not foolproof. This is avoided by using the optional Network Lock Manager (NLM). It provides file locking support to prevent multiple instances of open files.

Common Internet File System (CIFS)

Another method used to share resources across a network uses CIFS, which is a protocol based on Microsoft’s Server Message Block (SMB) protocol. Using CIFS, servers create *file shares* which are accessible by authorized clients. Clients subsequently connect to the server’s shares to gain access to the resource.

Security is controlled at both the user and share level. Client authentication information is sent to the server before the server will grant access. CIFS uses access control lists that are associated with the shares, directories, and files, and authentication is required for access.

A *session* in CIFS is oriented and *stateful*. This means that both client and server share a history of what is happening during a session, and they are aware of the activities occurring. If there is a problem, and the session has to be re-initiated, a new authentication process must be completed.

CIFS employs opportunistic locks (*oplocks*) to control file access. Depending on the type of locking mechanism required by the client, CIFS offers nodes the ability to cache read or write data from the file being accessed to improve network performance. Exclusive rights to the file prevents other nodes on the network from gaining access to that file until it is closed. During a CIFS session the lock manager has historical information concerning which client has opened the file, for what purpose, and in which sequence.

1.3.2 Understanding I/O

A major source of confusion regarding NAS is the concept of *File I/O* versus *Block I/O*. We try to shed a little light on this subject here. Understanding the difference between these two forms of data access is crucial to realizing the potential benefits of any SAN-based or NAS-based solution.

When a partition on a hard drive is under the control of an operating system (OS), the OS will format it. Formatting of the partition occurs when the OS lays a file system structure on the partition. This file system is what enables the OS to keep track of where it stores data. The file system is an addressing scheme the OS uses to map data on the partition. Now, when you want to get to a piece of data on that partition, you must request the data from the OS that controls it. For example, suppose that Windows 2000 formats a partition (or drive) and maps that partition to your system. Every time you request to open data on that partition, your request is processed by Windows 2000. Since there is a file system on the partition, it is accessed via File I/O. Additionally, you cannot request access to just the last 10 KB of a file. You must open the entire file, which is another reason that this method is referred to as File I/O.

Block I/O is handled differently: There is no OS format done to lay out a file system on the partition. The addressing scheme that keeps up with where data is stored is provided by the application using the partition. An example of this would be DB2 using its tables to keep track of where data is located rather than letting the OS do that job. That is not to say that DB2 cannot use the OS to keep track of where files are stored. It is just more efficient, for the database to bypass the cost of requesting the OS to do that work.

Using File I/O is like using an accountant. Accountants are good at keeping up with your money for you, but they charge you for that service. For your personal checkbook, you probably want to avoid that cost. On the other hand, for a corporation where many different kinds of requests are made, an accountant is a good idea. That way, checks are not written when they should not be. When sharing files across a network, something needs to control when writes can be done. The operating system fills this role. It does not allow multiple writes at the same time, even though many write requests are made. Databases are able to control this writing function on their own, so they run faster by skipping the OS.

For a more in-depth study of these topics, refer to the redbook, *IP Storage Networking: IBM NAS & iSCSI Solutions*, SG24-6240.

1.4 Network Attached Storage (NAS)

Storage devices which optimize the concept of file sharing across the network have come to be known as Network Attached Storage (NAS). NAS solutions utilize the mature Ethernet IP network technology of the LAN. Data is sent to and from NAS devices over the LAN using TCP/IP.

By making storage devices LAN addressable, the storage is freed from its direct attachment to a specific server and any-to-any connectivity is facilitated using the LAN fabric. In principle, any user running any operating system can access files on the remote storage device. This is done by means of a common network access protocol, for example, NFS for UNIX servers, and CIFS for Windows servers.

A storage device cannot just attach to a LAN. It needs intelligence to manage the transfer and the organization of data on the device. The intelligence is provided by a dedicated server to which the common storage is attached. It is important to understand this concept. NAS comprises a server, an operating system, plus storage which is shared across the network by many other servers and clients. So NAS is a *device*, rather than a *network infrastructure*, and shared storage is either internal to the NAS device or attached to it.

1.4.1 File servers

Early NAS implementations in the late 1980s used a standard UNIX or NT server with NFS or CIFS software to operate as a remote file server. In such implementations, clients and other application servers access the files stored on the remote file server, as though the files are located on their local disks. The location of the file is transparent to the user.

Several hundred users could work on information stored on the file server, each one unaware that the data is located on another system. The file server has to manage I/O requests accurately, queuing as necessary, fulfilling the request and returning the information to the correct initiator. The NAS server handles all aspects of security and lock management. If one user has the file open for updating, no-one else can update the file until it is released. The file server keeps track of connected clients by means of their network IDs, addresses, and so on.

1.4.2 Network appliances

More recent developments use application specific, specialized, “thin server” configurations with customized operating systems, usually comprising a stripped down UNIX kernel, reduced Linux OS, or a specialized Windows 2000 kernel, as with the IBM xSeries 150 integrated NAS appliances. In these reduced operating systems, many of the server operating system functions are not supported. The objective is to improve performance and reduce costs by eliminating unnecessary functions normally found in the standard hardware and software. Some NAS implementations also employ specialized data mover engines and separate interface processors in efforts to further boost performance.

These specialized file servers with a reduced OS are typically known as appliances, describing the concept of an application specific system. The term “appliance” borrows from household electrical devices the idea of a specialized “plug-and-play” application specific tool, such as a coffee maker or a toaster. NAS appliances, like the IBM TotalStorage NAS 300G, typically 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 appliance and manages the disk resources from a remote console.

One of the typical characteristics of a NAS appliance is its ability to be installed rapidly using minimal time and effort to configure the system. It is integrated seamlessly into the network as shown in Figure 1-6. This “plug-and-play” approach makes NAS appliances especially attractive when lack of time and skills are elements in the decision process.

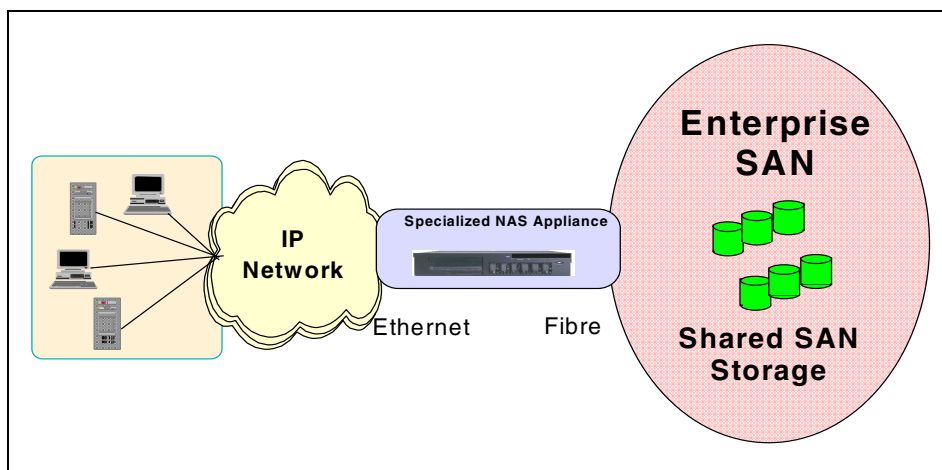


Figure 1-6 The role of the NAS 300G in your storage network

So, a NAS appliance is an easy to use device, which is designed for a specific function, such as serving files to be shared among multiple clients. It performs this task very well. It is important to recognize this when selecting a NAS solution. It is not a general purpose server, and should not be used (indeed, due to its reduced OS, probably cannot be used) for general purpose server tasks. However, it does provide a good solution for appropriately selected shared storage applications.

1.4.3 NAS uses File I/O

One of the key differences of a NAS disk device, compared to direct access storage (DAS) is that all I/O operations use file level I/O protocols. File I/O is a high level type of request that, in essence, specifies only the file to be accessed, but does not directly address the storage device. This is done later by other operating system functions in the remote NAS appliance.

A File I/O request specifies the file and the offset into the file. For instance, the I/O may specify “Go to byte ‘1000’ in the file (as if the file was a set of contiguous bytes), and read the next 256 bytes beginning at that position”. Unlike Block I/O, there is no awareness of a disk volume or disk sectors in a File I/O request. Inside the NAS appliance, the operating system keeps track of where files are located on disk. The OS issues a Block I/O request to the disks to fulfill the File I/O read and write requests it receives.

Network access methods, NFS and CIFS, can only handle File I/O requests to the remote file system. I/O requests are packaged by the node initiating the I/O request into packets to move across the network. The remote NAS file system converts the request to Block I/O and reads or writes the data to the NAS disk storage. To return data to the requesting client application, the NAS appliance software re-packages the data in TCP/IP protocols to move it back across the network. This is illustrated in Figure 1-7.

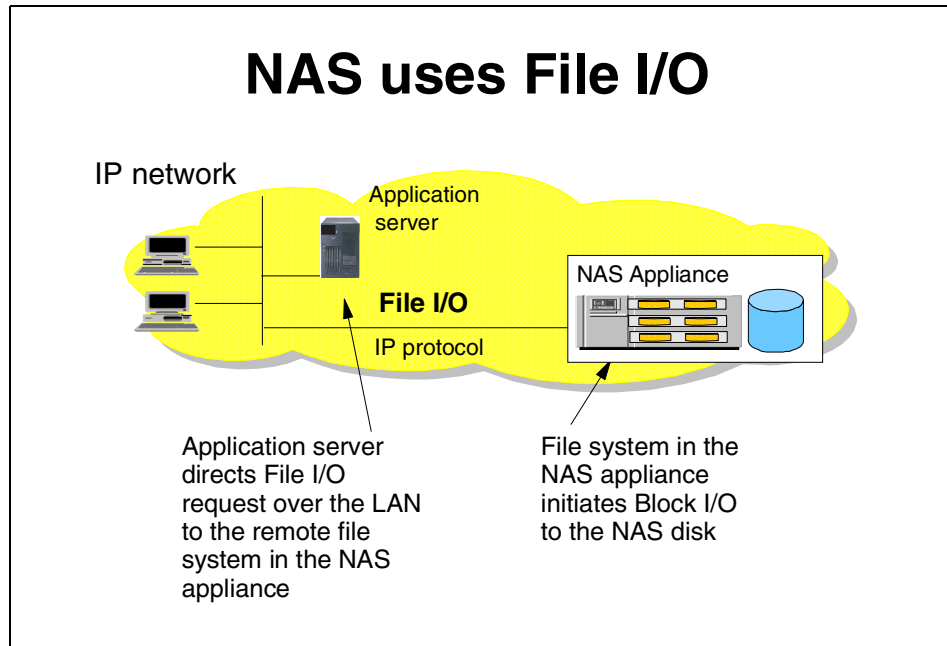


Figure 1-7 NAS devices use File I/O

1.4.4 NAS benefits

NAS offers a number of benefits that address some of the limitations of directly attached storage devices, and that overcome some of the complexities associated with SANs.

Resource pooling

A NAS appliance enables disk storage capacity to be consolidated and pooled on a shared network resource, at great distances from the clients and servers which will share it. Thus a NAS device can be configured as one or more file systems, each residing on specified disk volumes. All users accessing the same file system are assigned space within it on demand. This contrasts with individual DAS storage, when some users may have too little storage, and others may have too much.

Consolidation of files onto a centralized NAS device can minimize the need to have multiple copies of files spread on distributed clients. Thus overall hardware costs can be reduced.

NAS pooling can reduce the need to physically reassign capacity among users. The results can be lower overall costs through better utilization of the storage, lower management costs, increased flexibility, and increased control.

Exploits existing infrastructure

Because NAS utilizes the existing LAN infrastructure, there are minimal costs of implementation. Introducing a new network infrastructure, such as a Fibre Channel SAN, can incur significant hardware costs. In addition, new skills must be acquired, and a project of any size will need careful planning and monitoring to bring it to completion.

Simple to implement

Because NAS devices attach to mature, standard LAN implementations, and have standard LAN addresses, they are typically extremely easy to install, operate, and administer. This plug-and-play operation results in low risk, ease of use, and fewer operator errors, all of which contributes to lower costs of ownership.

Enhanced choice

The storage decision is separated from the server decision, thus enabling the buyer to exercise more choice in selecting equipment to meet the business needs.

Connectivity

LAN implementation allows any-to-any connectivity across the network. NAS appliances may allow for concurrent attachment to multiple networks, thus supporting many users.

Scalability

NAS appliances can scale in capacity and performance within the allowed configuration limits of the individual appliance. However, this may be restricted by considerations such as LAN bandwidth constraints, and the need to avoid restricting other LAN traffic.

Heterogeneous file sharing

Remote file sharing is one of the basic functions of any NAS appliance. Multiple client systems can have access to the same file. Access control is serialized by NFS or CIFS. Heterogeneous file sharing may be enabled by the provision of translation facilities between NFS and CIFS, as with the 300G.

Improved manageability

By providing consolidated storage, which supports multiple application systems, storage management is centralized. This enables a storage administrator to manage more capacity on a NAS appliance than typically would be possible for distributed, directly attached storage.

Enhanced backup

NAS appliance backup is a common feature of most popular backup software packages. For instance, the IBM xSeries 150 and 300G appliances all provide TSM client software support. Some NAS appliances have some integrated, automated backup facility to tape, enhanced by the availability of advanced functions such as the IBM NAS appliance facility called Persistent Storage Manager (PSM). This enables multiple point-in-time copies of files to be created on disk, which can be used to make backup copies to tape in the background. This is similar in concept to features such as IBM's Snapshot function on the IBM RAMAC Virtual Array (RVA).

1.4.5 Other NAS considerations

On the converse side of the storage network decision, you need to take into consideration the following factors regarding NAS solutions.

Proliferation of NAS devices

Pooling of NAS resources can only occur within the capacity of the individual NAS appliance. As a result, in order to scale for capacity and performance, there is a tendency to grow the number of individual NAS appliances over time, which can increase hardware and management costs.

Software overhead impacts performance

As we explained earlier, TCP/IP is designed to bring data integrity to Ethernet-based networks by guaranteeing data movement from one place to another. The trade-off for reliability is a software intensive network design which requires significant processing overheads, which can consume more than 50% of available processor cycles when handling Ethernet connections. This is not normally an issue for applications such as Web-browsing, but it is a drawback for performance intensive storage applications.

Consumption of LAN bandwidth

Ethernet LANs are tuned to favor short burst transmissions for rapid response to messaging requests, rather than large continuous data transmissions. Significant overhead can be imposed to move large blocks of data over the LAN. The maximum packet size for Ethernet is 1518 bytes. A 10 MB file has to be segmented into more than 7000 individual packets. Each packet is sent separately to the NAS device by the Ethernet collision detect access method. As a result, network congestion may lead to reduced or variable performance.

Data integrity

The Ethernet protocols are designed for messaging applications, so data integrity is not of the highest priority. Data packets may be dropped without warning in a busy network, and have to be resent. Since it is up to the receiver to detect that a data packet has not arrived, and to request that it be resent, this can cause additional network traffic.

With NFS file sharing there are some potential risks. Security controls can fairly easily be by-passed. This may be a concern for certain applications. Also the NFS file locking mechanism is not foolproof, so that multiple concurrent updates could occur in some situations.

Impact of backup/restore applications

One of the potential downsides of NAS is the consumption of substantial amounts of LAN bandwidth during backup and restore operations, which may impact other user applications. NAS devices may not suit applications which require very high bandwidth. To overcome this limitation, some users implement a dedicated IP network for high data volume applications, in addition to the messaging IP network. This can add significantly to the cost of the NAS solution.

Suitability for database

Given that their design is for File I/O transactions, NAS appliances are not optimized for the I/O demands of some database applications. They do not allow the database programmer to exploit raw Block I/O for high performance. As a result, typical databases, such as Oracle or UDB, do not perform as well on NAS devices as they would on DAS, or SAN. However, some customers may choose to use NAS for database applications with File I/O, because of their other advantages, including lower cost.

1.4.6 Total cost of ownership

Because it makes use of both existing LAN network infrastructures and network administration skills already employed in many organizations, NAS costs may be substantially lower than for directly attached or SAN-attached storage. Specifically, NAS-based solutions offer the following cost-reducing benefits:

- ▶ They reduce administrative staff requirements.
- ▶ They improve reliability and availability.
- ▶ They bridge the gap between UNIX and Windows environments.

Reduced administrative staff requirements

Implementing single or clustered NAS appliances to manage your networked storage concentrates the administrative tasks and thereby reduces the number of people required to maintain the network. Since the NAS appliance is a headless system, administration is performed via a Web-based GUI interface accessible from anywhere on the network. In addition, more capacity can be managed per administrator, thus resulting in a lower cost of ownership.

Improved reliability and availability

In today's business world, it has become the de facto standard to provide customers with access to information 24 hours per day, 7 days per week, allowing very little time available for unplanned outages. Some IBM NAS appliances offer the ability to provide great availability with options for clustered models.

Bridges the gap between UNIX and Windows environments

Most companies today contain heterogeneous operating environments. A NAS solution offers customers the ability for true cross-platform file sharing between Windows and UNIX clients by offering support for CIFS and NFS. This becomes increasingly important when application data becomes more common across platforms.

1.5 Storage Area Networks

A Storage Area Network (SAN) is a specialized, dedicated high speed network. Servers and storage devices may attach to the SAN. It is sometimes called "the network behind the servers". Like a LAN, a SAN allows any-to-any connection across the network, using interconnect elements such as routers, gateways, hubs, and switches.

Fibre Channel is the de facto SAN networking architecture, although other network standards could be used. Throughout this book, when we refer to SANs, we mean a Fibre Channel SAN. Thus, a decision to implement a SAN is a decision to develop a new storage network infrastructure (see Figure 1-8). Large numbers of customers worldwide are implementing Fibre Channel SANs right now. Industry analysts view this as the storage network infrastructure with the most momentum during the next two or three years.

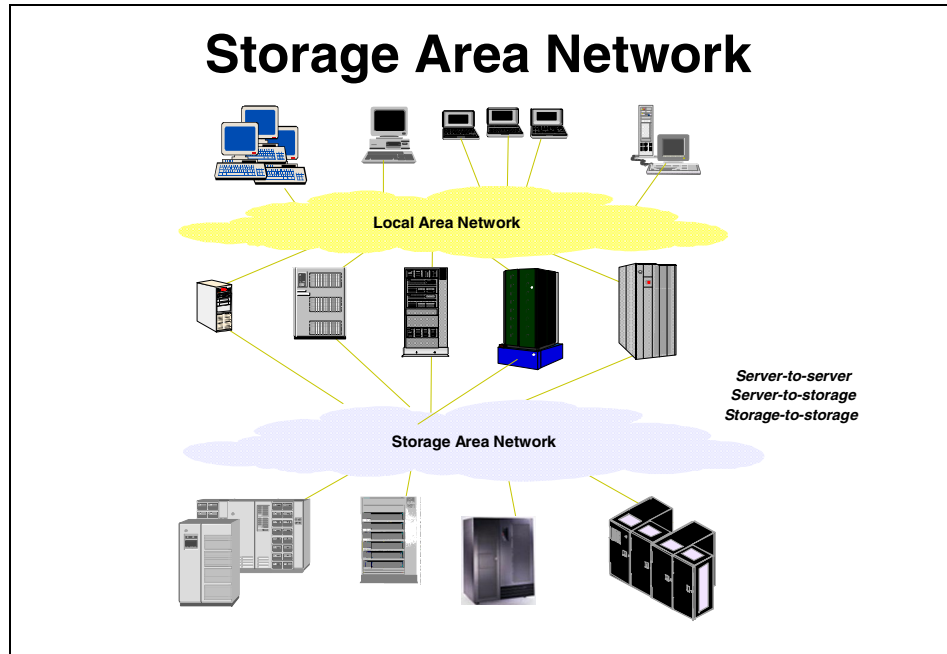


Figure 1-8 SAN — the network behind the servers

1.5.1 Overview of Fibre Channel storage networks

Fibre Channel is an open, technical standard for networking. It incorporates the data delivery (OSI Transport layer) characteristics of an I/O bus, with the flexible connectivity and distance characteristics of a network. One of the fundamental differences of SAN attached storage, compared to NAS, is that SAN storage systems typically attach directly to the network by means of hardware called host bus adapters (HBA). NAS, on the other hand, requires a “front-end” server as part of the appliance, and attaches to the LAN by means of a Network Interface Card (NIC).

A SAN eliminates the traditional dedicated connection between a server and DAS. Individual servers no longer own and manage the storage devices. Restrictions to the amount of data that a server can access is also minimized. Instead, a SAN enables many heterogeneous servers to share a common storage “pool”. This pool may comprise many storage devices, including disk, tape, and optical storage, and may be located many kilometers from the servers which use it. Thus, SAN attached storage has the potential to be highly scalable relative to a typical NAS device.

Because of its channel, or bus-like, qualities, hosts and applications see storage devices attached to the SAN as if they are locally attached storage. Because of its network characteristics, it can support multiple protocols and a broad range of devices, and it can be managed as a network.

Fibre Channel is a multi-layered network based on a series of American National Standards Institute (ANSI) standards. These define characteristics and functions for moving data across the network. Like other networks, information is sent in structured packets or frames, and data is serialized before transmission. But, unlike other networks, the Fibre Channel architecture includes a significant amount of hardware processing. This is oriented to storage Block I/O protocols, such as serial SCSI, also known as Fibre Channel Protocol (FCP). It is capable of delivering very high performance, relative to a NAS device, which is optimized for network File I/O. The speed currently achieved is 100 MBps full duplex, with 200 MBps soon to be delivered.

Measured effective data rates of Fibre Channel have been demonstrated in the range of 60 to 80 MBps over the 1 Gbps implementation. This compares to less than 30 MBps measured over Gigabit Ethernet. The packet size of Fibre Channel is 2,112 bytes. In comparison, an IP packet is 1,518 bytes, although typical IP transfers are much smaller. But for Fibre Channel, a maximum transfer unit sequence of up to 64 frames can be defined, allowing transfers of up to 128 MB without incurring additional overhead due to processor interrupts. As a result, Fiber Channel is unsurpassed for efficiency and high performance in moving large amounts of data at this moment in time.

Transmission is defined in the Fibre Channel standards across three transport topologies:

- ▶ **Point to point:** This is a bi-directional, dedicated interconnection between two nodes. This delivers a topology similar to DAS, but with the added benefits of longer distance.
- ▶ **Arbitrated loop:** This is a uni-directional ring topology, similar to a token ring, supporting up to 126 interconnected nodes. Each node passes data to the next node in the loop, until the data reaches the target node. All nodes share the 100 MBps bandwidth. Devices must arbitrate for access to the loop. FC-AL is suitable for small SAN configurations, or SANlets.

- ▶ **Switched fabric:** This describes an intelligent switching infrastructure which delivers data from any source to any destination. Each node is able to utilize the full 100 MBps bandwidth. Each logical connection receives dedicated bandwidth, so the overall bandwidth is multiplied by the number of connections. Complex fabrics must be managed by software which can exploit SAN management functions which are built into the fabric.

A mix of these three topologies can be implemented to meet specific needs.

SANs support the following direct, high speed transfers:

- ▶ **Server to storage:** This is similar to a DAS connection to a server. The SAN advantage, as with a NAS appliance, is that the same storage device may be accessed serially or concurrently by multiple servers.
- ▶ **Server to server:** High speed communications between servers.
- ▶ **Storage to storage:** Outboard data movement means data can be moved without server intervention. Examples include a disk device moving data directly to a tape device, or remote device mirroring across the SAN.

Fibre Channel combines the characteristic strengths of traditional I/O channels with those of computer networks, including these:

- ▶ A transport mechanism immune to electrical interference
- ▶ High performance for large data transfers by using simple transport protocols and extensive hardware assistance
- ▶ Serial data transmission
- ▶ A physical interface with a low error rate definition
- ▶ Reliable transmission of data with the ability to guarantee or confirm error free delivery of the data
- ▶ Packaging data in packets (frames in Fibre Channel terminology)
- ▶ Flexibility in terms of the types of information which can be transported in frames (such as data, video and audio)
- ▶ Use of existing device-oriented command sets, such as SCSI.
- ▶ A vast expansion in the number of devices which can be addressed when compared to traditional I/O interfaces.

It is this high degree of flexibility, availability, and scalability over long distances, and the broad acceptance of the Fibre Channel standards by vendors throughout the IT industry, which makes the Fibre Channel architecture attractive as the basis for new enterprise storage infrastructures.

1.5.2 Fibre Channel SANs use Block I/O

A SAN is similar to direct access storage to the extent that it is constructed from hardware and software storage interfaces. Fibre Channel uses serial SCSI-3 lower level protocols which use Block I/O access, just like a SCSI bus. Host based file systems and/or database I/O management are used, as with direct attached storage (see 1.1, “Local Area Networks” on page 5). All I/Os across the SAN are Block I/Os. The conversion to blocks takes place in the client or server platform, before transmission of the I/O request over the network to the target storage device (Figure 1-9). For more details of Block I/O, refer to Section 1.3.2, “Understanding I/O” on page 14.

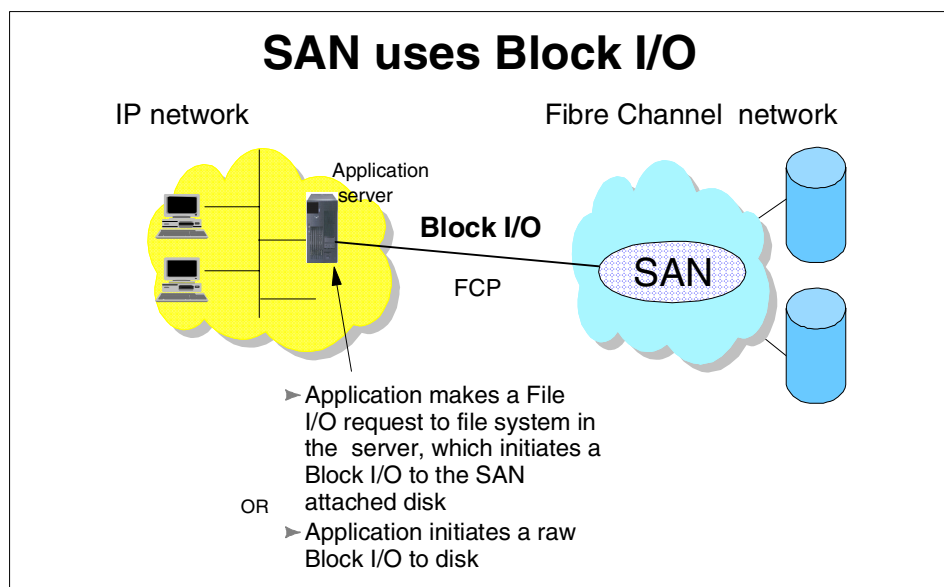


Figure 1-9 SAN uses Block I/O

1.5.3 SAN benefits

Today’s business environment creates many challenges for the enterprise IT planner. SANs can provide solutions to many of their operational problems:

Storage consolidation

By enabling storage capacity to be connected to servers at a greater distance, and by disconnecting storage resource management from individual hosts, a SAN enables disk storage capacity to be consolidated. The results can be lower overall costs through better utilization of the storage, lower management costs, increased flexibility, and increased control.

Data sharing

The term “data sharing” is used somewhat loosely by users and some vendors. It is sometimes interpreted to mean the replication of files (FTP-like). This enables two or more users or applications, possibly running on different host platforms, concurrently to use separate copies of the data. A SAN can ease the creation of such duplicate copies of data by enabling storage consolidation. This is also eased by using techniques found on enterprise class storage subsystems, such as remote mirroring and Flash Copy on the IBM Enterprise Storage Server.

Data sharing may also be used to describe multiple users accessing a single copy of a file. This is the role for which a NAS appliance is optimized. IBM provides a NAS-like file sharing capability across the SAN, for selected heterogeneous server environments, using the Tivoli SANergy File Sharing solution (see 1.6.1, “Tivoli SANergy” on page 31 for more details).

By enabling high speed (100 MBps) data sharing, the SAN solution may reduce traffic on the LAN, and eliminate the need for and cost of extra hardware that would otherwise be required to store duplicate copies of data. A SAN can also minimize the need for duplication of files by enabling storage consolidation. It also enhances the ability to implement cross enterprise applications, such as e-business, which may be inhibited when multiple data copies are stored.

Non-disruptive scalability for growth

A finite amount of disk storage can be connected physically to an individual server. With a SAN, new capacity can be added as required, without disrupting ongoing operations. SANs enable disk storage to be scaled independently of servers.

Improved backup and recovery

With data doubling every year, what effect does this have on the backup window? Backup to tape and recovery operations can increase LAN overheads:

- ▶ **Tape pooling:** SANs allow for greater connectivity of tape drives and tape libraries, especially at greater distances. Tape pooling is the ability for more than one server logically to share tape drives within an automated library.
- ▶ **LAN-free and server-free data movement:** Backup using the LAN may cause very high traffic volume which may be disruptive to normal application access to the network. SANs can minimize the movement of backup and recovery data across the LAN. IBM’s Tivoli Storage Manager (TSM) for LAN-free backup offers the capability for clients to move data directly to tape using the SAN. A server free data movement facility is also provided by Tivoli, allowing data to be read directly from disk to tape (and tape to disk), saving server cycles used for housekeeping. Further discussion of TSM is covered in 1.6.5, “Tivoli Storage Manager” on page 37.

- ▶ **High performance:** Many applications benefit from the more efficient transport mechanism of Fibre Channel. Most of the elements of FCP are implemented in hardware to increase performance and efficiency. Currently, Fibre Channel transfers data at up to 100 MBps full duplex (in practice measured with effective data rates in the range of 60 MBps to 80 MBps), several times faster than typical SCSI capabilities, and many times faster than standard LAN data transfers which operate at 10 Mbps or 100 Mbps. It is also faster than Gigabit Ethernet, which nominally operates at 100 Mbps, but which in practice typically delivers around 30 MBps to 40 MBps when moving storage related data. This is because of the latter's software overhead for large data transfers. Moving storage data transfers from the LAN to the SAN may improve application performance on servers.

High availability server clustering

Reliable and continuous access to information is an essential prerequisite in any business. In response, server and software vendors have developed high availability solutions based on clusters of servers. SCSI cabling tends to limit clusters to no more than two servers. A Fibre Channel SAN allows clusters to scale to 4, 8, 16, and even to 100 or more servers, as required, to provide very large shared data configurations making this more appropriate to database applications.

Data integrity

In Fibre Channel SANs, the class of service setting, such as Class 2, guarantees delivery of frames. Sequence checking and acknowledgement is handled in the hardware, so incurring no additional overhead. This compares to IP networks, where frames may be dropped in the event of network congestion, causing problems for data intensive applications.

Disaster tolerance

Sophisticated functions, like Peer-to-Peer Remote Copy (PPRC) services, address the need for secure and rapid recovery of data in the event of a disaster. A SAN implementation allows multiple open servers to benefit from this type of disaster protection. Additionally, the servers may be located at campus and metropolitan distances of up to 10-20 kilometers (km) from the disk array which holds the primary copy of the data. The secondary site, holding the mirror image of the data, may be located up to an additional 100 km from the primary site.

Allow selection of “best of breed” storage

A SAN enables storage purchasing decisions to be made independently of the server. Buyers are free to choose the best of breed solution to meet their performance, function, and cost needs. Large capacity external disk arrays may provide an extensive selection of advanced functions:

- ▶ Client/server backup solutions often include attachment of low capacity tape drives to individual servers. This introduces a significant administrative overhead as users often have to control the backup and recovery processes manually.
- ▶ A SAN allows the alternative strategy of sharing fewer, highly reliable, centralized tape solutions (such as IBM's Magstar family) between multiple users and departments.

Ease of data migration

Using a SAN, data can be moved non-disruptively from one storage subsystem to another, bypassing the server. The elimination of the use of server cycles may greatly ease the migration of data from old devices when introducing new technology.

Reduced total costs of ownership

Consolidation of storage in a SAN can reduce wasteful fragmentation of storage attached to multiple servers. A single, consistent data and storage resource management solution can be implemented. This can reduce costs of software and human resources for storage management compared to distributed DAS systems.

Storage resources match e-business enterprise needs

By eliminating islands of information, and introducing an integrated storage infrastructure, SAN solutions can be designed to match the strategic needs of today's e-business.

1.5.4 Other SAN considerations

There are pros and cons to most decisions. For example, consider these issues when making a SAN investment:

Costs

SAN entails installation of a new, dedicated Fibre Channel network infrastructure. The cost of the fabric components, such as Fibre Channel HBAs, hubs, and switches, is therefore an important consideration. Today these costs are significantly more expensive than the equivalent Ethernet connections and fabric components.

Inter-operability

Unlike Ethernet LANs, which have been implemented for more than fifteen years, Fibre Channel is still relatively early in its development cycle. A number of important industry standards are in place, but others have yet to be agreed upon. This has implications for ease of inter operability between different vendor's hardware and software, which may cause added complexity to the implementation of multi-vendor, heterogeneous SANs. However, this issue is gradually going away over time owing to industry-wide efforts in interoperability testing, and cooperation on development of standards (see 1.7, "Industry standards" on page 38).

Storage Wide Area Networks (SWANs)

Today, Fibre Channel Protocol SANs are mostly restricted in scope to the size of a LAN, due to the limited distances (10+ kilometers) supported by the Fibre Channel architecture. This has implications when considering the interconnection of multiple SANs into a SWAN. Such interconnections require protocol conversions to other transport technologies such as Asynchronous Transfer Mode (ATM) or TCP/IP, and the costs are high. Future implementations of FCP are expected to enable SANs to network across wider domains than a LAN, but this will likely take a few years.

Skills

Due to its recent introduction and explosive growth (only really beginning to take off in 1998), people with Fibre Channel skills are still relatively scarce. Employment of new staff with appropriate experience may be difficult or costly. It is often necessary, therefore, to invest in extensive education of your own staff, or use external services, such as IBM's Global Services organization, which have developed the necessary skills, and have wide experience with SAN implementations.

1.5.5 Data and SAN management

It is evident that the emergence of open, heterogeneous SAN architectures brings added complexity to storage administrators. Comprehensive management tools are required to enable them to effectively control and coordinate all aspects of data and storage resource utilization. These tools must enable appropriate data backup and recovery routines, as well as control data access, security, and disaster protection. They should also enable exploitation of the new capabilities of the SAN for consolidation, centralized management, LAN-free and server-less data movement, and so on.

IBM has introduced a family of data and SAN resource management tools, namely the IBM StorWatch family of tools, Tivoli Storage Manager, and Tivoli Network Storage Manager. In addition IBM has indicated its strategic direction to develop storage network virtualization solutions, which will allow enterprise-wide, policy driven, open systems management of storage.

1.6 Getting the best of both worlds: SAN with NAS

Most organizations have applications which require SAN performance, and others which will benefit from the lower cost and file sharing of a NAS solution. Recent IBM developments allow you to mix and match storage network solutions to deliver the most cost effective answer to meet your business needs. IBM's Tivoli SANergy software and the IBM NAS 300G appliance, either alone or combined, deliver NAS file sharing functions while exploiting Fibre Channel SAN scalability, high performance, and availability.

1.6.1 Tivoli SANergy

Tivoli SANergy introduces LAN file sharing technologies to SANs. In this section we describe the SANergy architecture and its cooperation with the Tivoli Storage Manager.

Tivoli SANergy is unique SAN software that allows sharing of access to application files and data between a variety of heterogeneous servers and workstations connected to a SAN. In addition, SANergy uses only industry-standard file systems like NFS and CIFS, enabling multiple computers simultaneous access to shared files through the SAN (see Figure 1-10). This allows users to leverage existing technical resources instead of learning new tools or migrating data to a new file system infrastructure. This software allows SAN-connected computers to have the high-bandwidth disk connection of a SAN while keeping the security, maturity, and inherent file sharing abilities of a LAN.

SANergy employs technology to combine the simplicity of LAN-based file sharing with the very high data transfer speeds afforded by today's Fibre Channel, SCSI, and SSA storage networks. This enables the use of high-speed, heterogeneous data sharing without the performance limiting bottlenecks of file servers and traditional networking protocols.

SANergy is unique in that it extends standard file systems and network services provided by the operating systems that it supports (Windows NT, MacOS, AIX, plus various UNIX and Linux platforms). As an OS extension built on standard systems interfaces, SANergy fully supports the user interface, management, access control, and security features native to the host platforms, providing all the file system management, access control, and security required in a network. With SANergy, virtually any network-aware application can access any file at any time, and multiple systems can transparently share common data.

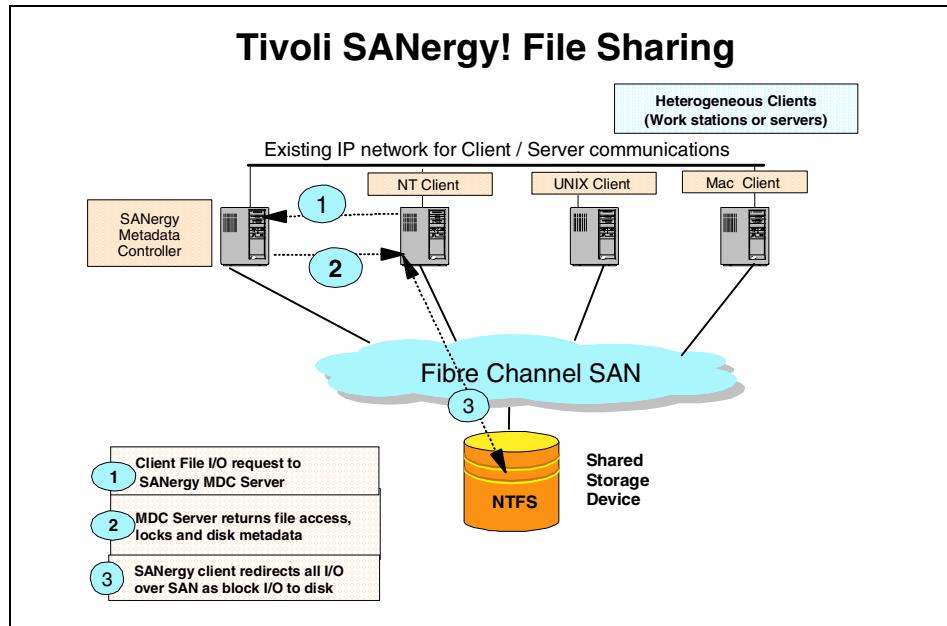


Figure 1-10 SANergy configuration

In addition to the SAN, SANergy also uses a standard LAN for all the metadata associated with file transfers. Because SANergy uses standard file systems, even if the SAN should fail, access to data via the LAN is still possible. Since each system has direct access to the SAN-based storage, SANergy can eliminate the file server as a single point of failure for mission-critical enterprise applications. It can also easily manage all data backup traffic over the storage network, while the users enjoy unimpeded LAN access to the existing file servers.

Tivoli SANergy architecture and data flow

The basic problem in storage area networking at the file level is keeping the separate operating systems up to date with each other's independent and asynchronous use of the storage. Tivoli SANergy is a hybrid of conventional networking and direct attached storage.

Conventional networking is rich with abilities for keeping many computers coherent. That is, if one computer has an open view of a directory, and another changes that directory (adds/deletes a file), the view on all computers will change. Conventional networking allows administrators to establish centralized access control lists and other data management facilities.

Data “about” data is referred to as metadata. Examples include file names, file sizes, and access control lists. The Tivoli SANergyFS architecture lets metadata transactions take place over conventional LAN networking. The actual content of files moves on the high-speed direct SAN connection, as illustrated in Figure 1-11.

SANergy works with Ethernet, ATM, or anything else that carries networking protocols. The network operating system can also be CIFS protocol (Windows NT), Appletalk, NFS (UNIX), or a combination. Similarly, SANergy supports any available disk-attached storage fabric. This includes Fibre Channel, SSA, SCSI, and any other disk-level connection. It is also possible for installations to use one set of physical wiring to carry both the LAN and storage traffic. When you use SANergy, one computer in the workgroup is tagged as the MetaData Controller (MDC) for a particular volume. You can have a single computer as the MDC for all volumes, or it can be spread around. The other computers are SANergy clients. They use conventional networking to “mount” that volume, and SANergy on those clients separates the metadata from the raw data automatically.

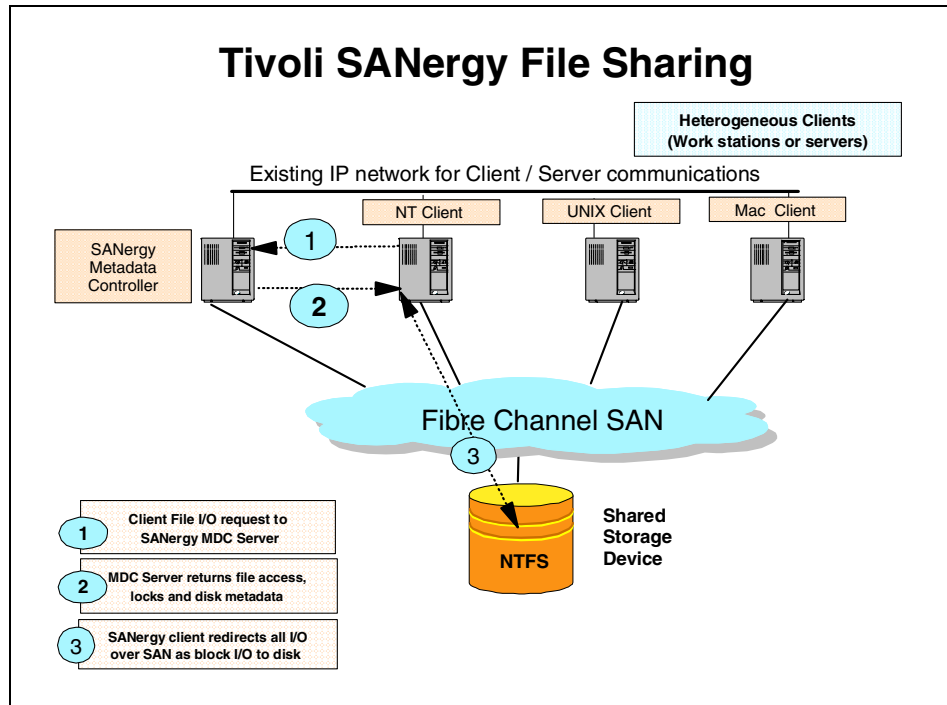


Figure 1-11 SANergy data flow

1.6.2 SANergy uses a mix of File I/O and Block I/O

SANergy is an intelligent, hybrid solution. It combines aspects of the LAN client, requesting file access to information stored on a remote server, with those of a SAN attached client, accessing data directly on the device.

When the initial request to open the file is made, the SANergy client does not own the device, and has no knowledge of the structure of the data on the disk. It therefore follows the standard approach of making an NFS, or CIFS, file call via TCP/IP to the remote server. In this case the server is the SANergy MetaData Controller (MDC). Recognizing that the I/O request is from a SANergy client, the SANergy MDC returns a number of important pieces of information to the client.

First, if the file is available for use, permission is granted to access the file (with either read or read/write capability). Second, the MDC provides file locking procedures, which prevents another client from accessing and updating the file while it is in use by the requestor. Finally, the MDC provides metadata about the file location and format on the disk.

With this information, the client now has the requisite information with which to access the disk device directly over the SAN. All subsequent I/O requests are redirected by the SANergy client as Block I/Os, over the Fibre Channel SAN, directly to the device, as we illustrated in Figure 1-11 on page 34.

1.6.3 SANergy benefits

In summary, Tivoli SANergy provides the following benefits:

File sharing at SAN speeds

SANergy software provides NAS-like file sharing, with data sent over the SAN rather than the LAN for higher performance. Applications which would benefit from remote file sharing, but which might previously have achieved poor performance over a LAN, can now participate in the benefits of pooled SAN storage, while delivering excellent performance.

True heterogeneous file sharing

SANergy file sharing is independent of the network file protocol. Once access to a file has been given to any client, subsequent disk I/O is done in serial SCSI block format. Multiple unlike platforms such as Windows, UNIX, AIX, and Macintosh may therefore concurrently share the file. This greatly increases user flexibility, and allows important information to be made available to user departments which have been equipped with a variety of host platforms.

Storage hardware flexibility

SANergy has the attributes of NAS and SAN with added flexibility. SANergy supports the NFS and CIFS protocols, but allows for the selection of enterprise class scalable disk systems like the IBM ESS, or other SAN attached disk storage required to suit the business need.

LAN-free and server-less data movement

SANergy automates the capability to move large data transfers like back up and recovery across the high speed SAN rather than over the LAN. These applications are among the most seductive for enterprise data managers.

- ▶ Using SANergy together with TSM lets you transfer your data through the SAN. It supports both LAN-free and server-less types of backup/restore. In both cases, the data transfer will be off-loaded to the SAN.
- ▶ These applications provide some of the most attractive benefits of SAN implementation because they eliminate so much traffic which currently moves across the LAN. We describe two possible scenarios for using TSM and SANergy in concert to provide these solutions in Chapter 6, “Backing up the IBM TotalStorage NAS 300G” on page 277.

Reduced hardware costs

SANergy supports the protocols of a conventional NAS appliance, but with significantly higher performance. At the same time, it does not require the dedicated NAS processor front-end to the disk storage. Instead, SANergy software sits in client hosts and in the SANergy MetaData controller. This may be a standard server or the 300G.

1.6.4 SANergy considerations

A number of considerations must be taken into account when implementing SANergy. These include:

File opening overheads

The remote file call across the LAN to the SANergy MDC entails an overhead every time a file is opened. Applications which open and close many small files for short periods of time, and issue a small number of I/O requests while the file is open, will not perform well. SANergy is optimized to give most benefit to applications which utilize relatively large files, keep them open for long periods, and issue large numbers of I/Os while the file is open.

File fragmentation

Metadata regarding the files to be accessed is normally very small. It takes little time to send this from the MDC to the SANergy client. However, if a file is fragmented across many sectors and disk devices, the volume of the metadata, and the time needed to send it to the client, may impact SANergy's performance. Storage administrators should ensure that de-fragmentation is carried out regularly, in order to minimize the file opening and file access overheads.

Database applications

Although SANergy is using Block I/O, it is NOT using the raw partition processing required by some database applications. For this reason, SANergy is not suitable for database applications unless the database I/O is processed via the client's file system. Also, some database vendors do not support access via redirected I/O.

1.6.5 Tivoli Storage Manager

The Tivoli Storage Management product set covers the various aspects of storage management within an enterprise environment. In general, the product set can be structured in three parts: The Tivoli Storage Manager, which provides the storage management backbone; the complementary products, which add special storage management functions (like disaster recovery management or hierarchical storage management) in conjunction with Tivoli Storage Manager; and the Tivoli Data Protection group of products, which integrate application data management into the Tivoli Storage Manager storage management environment.

Since the announcement of Tivoli Storage Manager in September 1999, there have been two major new versions of Tivoli Storage Manager introduced, Version 3.7.3 in April 2000 and Version 4.1 in July 2000.

In the following section we provide an introduction to these major new functions and features of both versions:

- ▶ Windows 2000 exploitation and support of client and server
- ▶ New AIX and Windows 2000 platform support tape library sharing in a SAN environment for SCSI connected libraries
- ▶ Client tape support for backup set restore on the local client system without interaction with the Tivoli Storage manager server
- ▶ New V3.7 backup-archive clients, including the newly supported Linux platform
- ▶ Backup-archive client and server enhancements supporting the special backup requirements of mobile systems
- ▶ LAN-free client data transfer for the API client in a SAN environment
- ▶ Tivoli Storage Manager hardware integration to implement server-free backup solutions for enhanced disk subsystems like IBM ESS and EMC Symmetrix

1.6.6 SANergy with Tivoli Storage Manager

Tivoli Storage Manager operates in a client-server architecture where a server system (with various storage media attached) provides backup services to the clients. The trivial case of using TSM with SANergy involves placing a TSM client on the MDC to perform normal Backup/Archive client operations for the MDC data. Setup and considerations here are the same as for a Tivoli Storage Manager client without SANergy. Beyond this, various scenarios are presented in the following sections, including application server-free backup and restore, implementing Tivoli Storage Manager backup sets on SANergy volumes, backing up a commercial database from a SANergy host, and using Tivoli Storage Manager in environments where there are different operating systems for the MDC and SANergy host.

1.6.7 Application server-free backup and restore

The purpose of this scenario is to show how to successfully perform LAN-free backup and restore operations where the Tivoli Storage Manager client does not run on the system which owns the data, that is, it is not installed on the SANergy MDC, but rather on one or more of the SANergy hosts. This configuration is sometimes referred to as *application server-free*, since there is no Tivoli Storage Manager code installed and no backup processing takes place on the MDC/application server itself. The operating system or platform of the SANergy host is the same as that of the SANergy MDC. The reason we chose to run the client on a SANergy host whose platform is the same as that of the SANergy MDC is explained in Chapter 4, “Clustering for high availability” on page 167.

1.6.8 SAN exploitation: LAN-free client data transfer

The advent of SAN technology provides an alternative path for data movement between the Tivoli Storage Manager client and the server. Shared storage resources are accessible to both the client and the server through the storage area network. It is now possible for the client to write to tape storage managed by the server. Data movement is off-loaded from the LAN and from the server processor for greater scalability.

LAN-free client data transfers are a new feature of Tivoli Storage Manager V4.1 that allows a Tivoli Data Protection for Application Client to transfer data directly to a SAN attached tape device that is known to both the client and server. The TSM API client in conjunction with the enhanced TSM Server and a new TSM Storage Agent have been enhanced with the ability to write directly to server-owned tape storage media in a format that is consistent with that written by the server today. TDP for Exchange and TDP for R/3 are the only supported applications in the initial release.

1.7 Industry standards

There is a clear customer need for standardization within the storage networking industry to allow users to freely select equipment and solutions, knowing that they are not tying themselves to a proprietary or short term investment. To this end, there are extensive efforts among the major vendors in the storage networking industry to cooperate in the early agreement, development, and adoption of standards. A number of industry associations, standards bodies, and company groupings are involved in developing and publishing storage networking standards. The most important of these are the SNIA and the Internet Engineering Task Force (IETF).

In addition, IBM, IBM Business Partners, and other major vendors in the industry, have invested heavily in inter-operability laboratories. The IBM laboratories in Gaithersburg (Maryland, USA), Mainz (Germany), and Tokyo (Japan) are actively testing equipment from IBM and many other vendors, to facilitate the early confirmation of compatibility between multiple vendors servers, storage, and network hardware and software components.

1.7.1 Storage Networking Industry Association

The Storage Networking Industry Association (SNIA) is an international computer industry forum of developers, integrators, and IT professionals who evolve and promote storage networking technology and solutions. SNIA was formed to ensure that storage networks become efficient, complete, and trusted solutions across the IT community.

SNIA is accepted as the primary organization for the development of SAN and NAS standards, with over 150 companies and individuals as its members, including all the major server, storage, and fabric component vendors. SNIA is committed to delivering architectures, education, and services that will propel storage networking solutions into a broader market.

IBM is one of the founding members of SNIA, and has senior representatives participating on the board and in technical groups. For additional information on the various activities of SNIA, see its Web site at:

<http://www.snia.org>

1.7.2 Internet Engineering Task Force

The Internet Engineering Task Force (IETF) is a large open international community of network designers, operators, vendors, and researchers concerned with the evolution of the Internet architecture and the smooth operation of the Internet. It is open to any interested individual. The actual technical work of the IETF is done in its working groups, which are organized by topic into several areas (for example, routing, transport, and security).

For more information on the IETF and its work groups, refer to:

<http://www.ietf.org>



IBM hardware overview

This chapter provides a brief overview of each of the IBM products used during the development of this redbook. Specifically, we describe the key features and benefits of the IBM TotalStorage NAS 300G (300G), the IBM Enterprise Storage Server (ESS), the IBM Modular Storage Server (MSS), the IBM Fibre Array Storage Technology (FAST) 200, and the IBM SAN Fibre Channel Switch. The hardware configurations of each of the products used in our environment are not covered here, but are addressed in the subsequent chapters.

2.1 IBM TotalStorage NAS 300G

Designed to be a plug-in appliance, the 300G is an optimized, high performance server designed to provide shared data to both clients and servers in Windows, UNIX, and mixed environments. Attached directly to both the LAN and the SAN, the 300G off-loads general file sharing from other servers, freeing those servers to handle more resource-intensive application processing. Adding the 300G to the LAN and SAN does not affect any other systems in either of these networks, and upgrading other servers, clients, or applications does not impact the 300G.

The 300G high-speed appliance connects your Ethernet LAN to storage resources on your SAN. These are high-performance models which are designed to link application servers, transaction servers, file servers, and end-user clients to storage resources located on the SAN, 24 hours a day, 7 days a week.

Two different types of configurations are available for this product: the single-node Model G00 and the dual-node Model G25. The dual node Model G25 also provides clustering and failover protection for top performance and availability. In this book, we discuss implementing both configurations. Both the single and dual nodes are depicted in Figure 2-1.

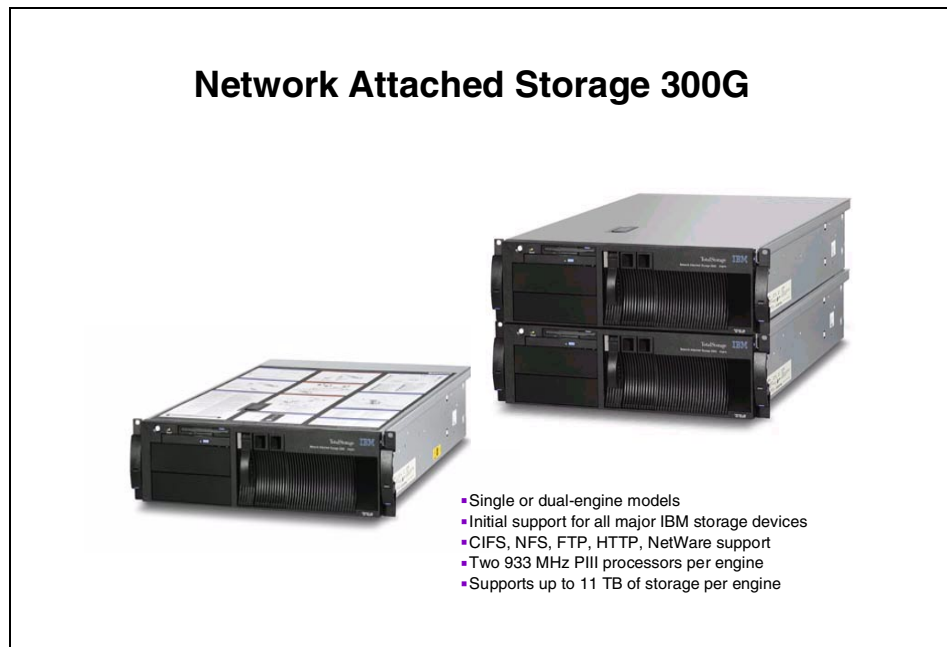


Figure 2-1 The 300G models G00 and G25

To remain competitive, your information systems must be flexible enough to accommodate evolving needs and must be available around the clock. The G00 and G25 have been designed to meet these challenges head on. Not only do they feature a modular design for flexible growth in processing power and connectivity (to provide a lower overall total cost of ownership), they also provide high availability and reliability with hot-swappable and redundant power supplies.

The 300G G00 and G25 models are specialized NAS appliances acting as a high-bandwidth conduit. They connect LAN-attached clients and servers to the SAN through high-speed Fibre Channel paths.

The 300G can be a valuable addition to your storage network strategy because:

- ▶ It is easy to use and install.
- ▶ It is a headless appliance — it requires no keyboard, mouse, or display to configure and maintain.
- ▶ It supports CIFS, NFS, NetWare File Systems, FTP, and HTTP.
- ▶ It provides persistent image file server backup — a point-in-time backup accessible by users without administrator intervention.
- ▶ It includes Web-based GUI administration tools.
- ▶ It is preconfigured to support Windows Terminal Services for remote administration and configuration.
- ▶ It includes the IBM Director agent.
- ▶ It includes a Tivoli Storage Manager client.
- ▶ It comes bundled with Tivoli SANergy.

2.1.1 Hardware configuration for the 300G

The 300G G00 and G25 come shipped with standard hardware. Since the 300G is a NAS appliance designed for a specific purpose, it is equipped with the hardware required to integrate it directly into your network. For each model, there is optional hardware available to improve performance. Table 2-1 lists the hardware components for both models.

Table 2-1 Hardware configurations for the G00 and G25

Hardware (per engine)	Description
Processor	Dual 933 MHz Pentium III
L2 Cache	512 KB
ECC SDRAM memory	1 GB
Available PCI slots	4 (2x32 bit and 2x64 bit)
Ethernet connections	1 1-port 10/100 Mbps Ethernet (single node) 2 1-port 10/100 Mbps Ethernet (dual node)
Fibre channel adapter	QLogic F-port 2200
Optional adapters	1-port 10/100 Mbps Ethernet 1-port Gigabit Ethernet Advanced System Management Adapter 1-port Fibre Channel Adapter

For higher availability and redundancy, you can turn to the dual-node Model G25 which provides a dual redundant path to data access. This model has been configured for clustering takeover should there be a failure on any one of its nodes.

The IBM TotalStorage NAS 300G Model G25 is made up of 2 individual rack mounted single node units. The hardware in each of these units is identical to the single node configuration.

2.1.2 IBM TotalStorage NAS 300G features and benefits

The 300G is an appliance that is designed to work in heterogeneous environments right out of the box. Figure 2-2 visually demonstrates how the 300G's built-in features allow it to plug into almost any environment.

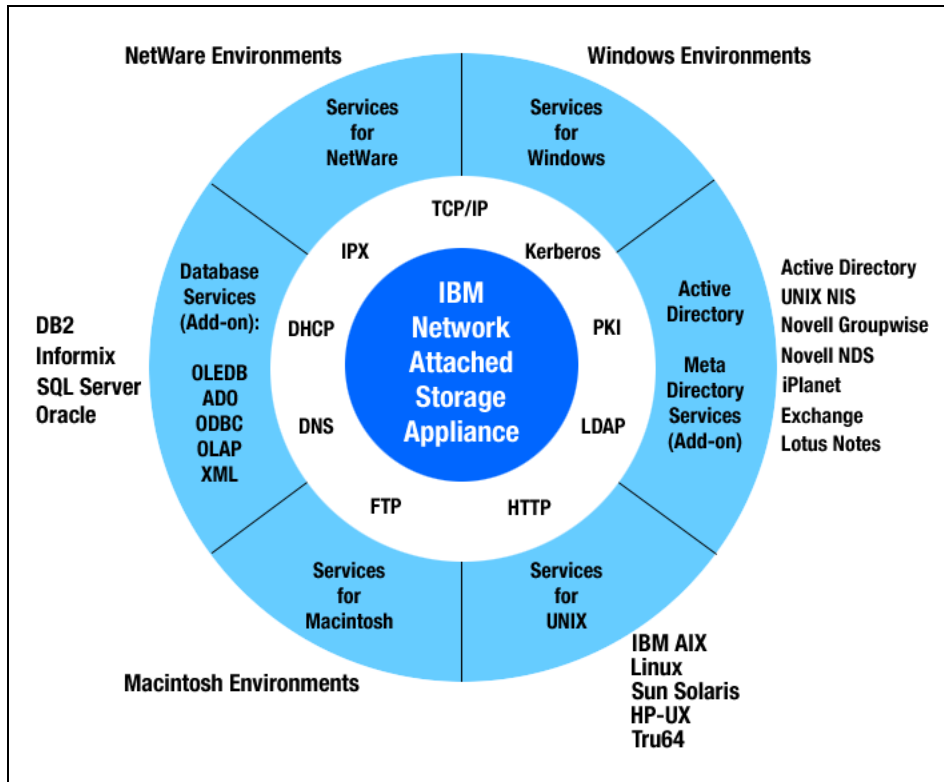


Figure 2-2 Visualization of the 300G's interoperability features

Table 2-2 summarizes the features and benefits of the NAS 300G model.

Table 2-2 300G features and benefits

Features	Benefits
Capacity supported per engine	Up to 11 TB of Fibre-based storage
Dual Node Configuration — G25 only	Clustered Failover support for increased availability and performance
Two 933 Mhz Pentium III Processors per node	Very high performance
Open standards	Easy integration into existing networks Smooth migration paths for business growth

Features	Benefits
Multiple file protocol support	Supports heterogeneous client/server environments Windows (CIFS), UNIX (NFS), Netware, FTP, HTTP
Simplified data management	Heterogeneous and consolidated file serving management
Web browser interface	Simplifies appliance installation
Provides remote LAN users access to SAN storage	Access to pooled storage on SAN without individual Fibre Channel connections
Combines NAS and SAN	Provides SAN scalability and performance on the IP network
Preloaded SANergy software	Enables SANergy clients to simultaneously share storage volumes
Systems management via IBM Director, Tivoli SANergy, Tivoli Storage Manager, Microsoft Systems Management	Comprehensive management facilities are preloaded and easy to use

2.1.3 300G optional features

In addition to the features listed above, the following are optional features for the 300G:

- ▶ Fibre Channel Adapter (#0002) — plant and field installable

The Fibre Channel Adapter is an intelligent, DMA bus mastering adapter that has been optimized for high throughput. It contains a powerful RISC processor, Fibre Channel protocol module with 1-Gb/s transceivers, and a PCI local bus interface. The Fibre Channel Adapter has a duplex type SC connector for attaching multi-mode fiber cable. The adapter supports 50 μ m or 62.5 μ m multi-mode fiber cable lengths up to 500 meters. Key features include:

- 100 MBps data rate over Fibre Channel connections
- PCI bus operation at 33 MHz or 66 MHz
- 64-bit, 32-bit PCI bus interfaces
- Supports:
 - FCP-SCSI protocols
 - FC-AL public loop profile (FL-PORT FABRIC LOGIN)
 - Point-to-point fabric connection (F-PORT FABRIC LOGIN)
 - Fibre Channel service classes 2 and 3

- ▶ Advanced System Management (ASM) Adapter For 5196-G00 (#0003) And Interconnect Cable for 5196-G25 (#0004) — plant and field installable

The Advanced System Management PCI Adapter, in conjunction with the ASM processor that is integrated into the base planar board of the servers, allows you to connect via LAN or modem from virtually anywhere for extensive remote management of the 300G. The ASM adapter enables more flexible management through a Web browser interface, in addition to ANSI terminal, Telnet, and IBM Director.

- ▶ 10/100 Ethernet Adapter (#0005) — plant and field installable

The 10/100 Ethernet Adapter (34L1501) provides IEEE 802.3-compliant 100BASE-TX and 10BASE-T Ethernet connectivity for servers over an unshielded twisted pair link through a single RJ-45 connector. Its 32-bit PCI 2.1 bus mastering architecture offers outstanding performance with low server CPU utilization. It provides half-duplex and full-duplex operation at both 10 Mbps and 100 Mbps. Auto-negotiation of speed and duplex mode facilitates the use of the adapter in shared or switched environments running at either speed.

- ▶ Gigabit Ethernet SX Adapter (#0006) — plant and field installable

The Gigabit Ethernet SX Adapter provides 1000BASE-SX connectivity to a Gigabit Ethernet network for servers over a 50 or 62.5 micron multimode fiber optic link attached to its duplex SC connector. Its 1000 Mbps data rate and 32- or 64-bit PCI bus mastering architecture enable the highest Ethernet bandwidth available in an adapter. It is compliant with IEEE 802.3z Ethernet and PCI 2.1 standards, ensuring compatibility with existing Ethernet installations. It also supports 802.1p packet prioritization and 802.1q VLAN tagging.

- ▶ 250W Hot-Swap Redundant Power Supply (33L3760)

2.1.4 300G included software

The software in Table 2-3 is included in the 300G.

Table 2-3 300G software

Software	5196-G00 and 5196-G25
Operating system	Windows Powered OS
Backup	Columbia Data Products Persistent Storage Manager enables point-in-time backup 250 Persistent Images
Storage management	Tivoli Storage Manager Client(V3.7201)
Systems management	IBM Director 2.12 agent

Software	5196-G00 and 5196-G25
Performance Management	Tivoli SANergy Exec Agent (V2.2)
Remote Administration	Web-based GUI Microsoft Terminal Services
Configuration tools	IBM Advanced Appliance Configuration Utility

2.1.5 300G preloaded and optional software

Each 300G Model G00 and G25 is preloaded at the factory with its base operating system and applications. The code is loaded to the system's hard disk with a backup copy provided on an emergency recovery CD-ROM. The operating system and NAS application code have been specifically tuned to enable the Model G00 and G25 as high performance NAS server appliances.

In addition to the operating system and application software, each unit contains tools which simplify remote configuration and administration tasks. Additionally, included network management agents provide options for managing the units.

Specifically, the units come preconfigured with the following functions:

- ▶ Windows Powered OS
 - Windows 2000 Advanced Server code optimized for the IBM TotalStorage NAS 300G Models G00 and G25
- ▶ Multiple file systems support
 - CIFS
 - NFS
 - Netware
 - Macintosh
- ▶ Multiple file transfer services
 - FTP
 - HTTP
- ▶ Remote NAS system administration
 - Administrative tasks can be performed in the Web-based GUI
 - IBM Advanced Appliance Configuration Utility
 - Alternate administrative task performed using Windows Terminal Service
 - Advanced management functions available via Windows Terminal Service
 - Simple point-and-click for restores using NT Backup
 - NAS Backup Assistant MMC Snap-in Web page

- ▶ UNIX services
 - Pre-configured NFS support
 - Web-based GUI for performing administrative tasks
 - Microsoft Services for UNIX V2.2
 - NFS V3.0 (IETF RFC 1830)
- ▶ Automatic disaster recovery of operating system
 - The IBM Snap command-line utility creates a point-in-time persistent image of the specified drive letter, and then backs up that persistent image (by calling a custom-written batch file that invokes the backup software via command line — batch file is passed to IBM Snap as command-line parameter); supports full and incremental backups
 - The NAS Backup Assistant is a GUI front end to IBM Snap that generates a batch file that invokes NT Backup from settings configured by the user in the GUI
 - Fifteen minute original factory CD-ROM reload of operating system
 - Prevention of accidental reloads via reload enablement diskette
- ▶ IBM Fibre Management utility
 - IBM Fibre Stand-Alone Management utility
 - MMC snap-in that launches the utility
 - Users can use terminal services to remotely monitor the fibre adapter configuration
- ▶ Advanced Aggregate Management
 - IBM Director Agent
- ▶ Columbia Data Products Persistent Storage Manager for IBM NAS

Persistent Storage Manager (PSM) creates True Images (tm) (multiple point-in-time persistent images of any or all system and data volumes). All persistent images survive system power loss or a planned or unplanned reboot. Each instance of PSM seamlessly handles 250 concurrent images of up to 255 independent volumes for a total of 63,750 independent data images. Any image can be easily managed through the Microsoft Web user interface, and accessed the same as any other active volume. In case of data corruption or loss, any persistent image can be used for manual retrieval of individual files (by the administrator or end users), or more importantly, for instant restoration (by a PSM function initiated by the administrator, in the Web user interface) of the entire volume from image, which can substantially reduce the amount of system down time.

- Persistent Storage Manager creates and keeps multiple point-in-time persistent images (maximum of 250 concurrent images of up to 255 independent volumes)
 - All images for a volume are mounted under a single directory (in the root directory of the volume), with each image under its own mount point
 - User-level access can be granted to one or more of the images, to allow users to restore their own files from the images (users automatically have the same access privileges to individual files and directories in the images that they would have on the actual volume)
 - Images can be read-only, or read-write (with ability to reset (undo changes to) read-write images)
 - Images can be assigned retention levels (if an image needs to be automatically deleted by PSM, the highest priority images can be kept)
 - Any image can be used to restore an entire volume instantly (for data volumes, typically within seconds; for system volume, system reboot is required)
 - Flexible, configurable image access and administration via Web-based user interface
 - Schedule images (for each schedule entry (image group), specify interval, number to keep, image name, properties (read-only or read-write, retention level))
 - Create a new image immediately (specify name and properties)
 - Delete images
 - View and change properties of images (also reset read-write images)
 - Restore a volume (from any image of that volume)
 - Configure advanced parameters:
 - Maximum number of images to keep concurrently
 - Name of image root directory
 - Quiescent period and quiescent period wait time-out
 - Size of the image cache file (per volume)
 - Image cache file usage warning and automatic image deletion thresholds (per volume)
- ▶ IBM Director with Universal Manageability (UM) Services V2.12
- The IBM TotalStorage NAS 300G contains a IBM Director agent and can be managed by this powerful, highly-integrated, systems management software solution that is built upon industry standards and designed for ease-of-use. Using its intuitive Java-based GUI, an administrator can centrally manage

individual or large groups of IBM and non-IBM PC-based servers. IT administrators can view the hardware configuration of remote systems in detail and monitor the usage and performance of crucial components, such as processors, disks, and memory.

The following functions have been added in V2.12:

- Windows 2000 server, console, and agent
- SCO UnixWare agent
- Alert on LAN — (AOL) configuration enhancements
- Wired for Management — (WfM) — compliant CIM to DMI Mapper
- SNMP device listener for Netfinity hardware

IBM Director with UM Services V2.12 is the latest update to IBM world-class systems manageability solutions. V2.12 replaces all earlier versions of NF Director and UM Services and adds support for Windows 2000, SCO UnixWare, and new IBM hardware systems.

2.1.6 IBM NAS 300G sample connectivity

The 300G supports connectivity to a variety of IBM storage products. Figure 2-3 shows sample connectivity between the 300G and the ESS, MSS, FASTt200, FASTt500, or the 7133 Serial Disk System.

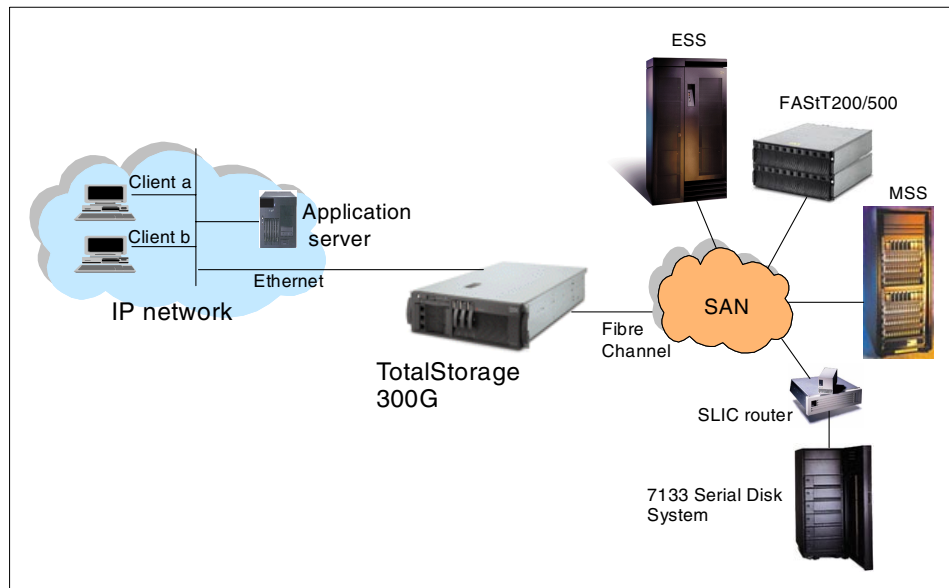


Figure 2-3 300G connectivity with ESS, FastT200/500, MSS, and 7133

2.2 IBM Enterprise Storage Server

This topic provides a brief overview of the major IBM Enterprise Storage Server (ESS) components, features and benefits. For more detailed information, please refer to the redbook, *The IBM Enterprise Storage Server*, SG24-5645.

2.2.1 IBM Enterprise Storage Server Overview

The heterogeneous servers on which the e-business applications are deployed require high-function, high-performance, scalable storage servers to meet the demanding requirements of Enterprise Resource Planning, Data Warehousing, and Business Intelligence applications. The IBM Enterprise Storage Server has set new standards in function, performance, and scalability in these most challenging environments.

The IBM Enterprise Storage Server (ESS) is a high performance, high availability, high capacity disk storage subsystem. It is a member of the Seascope family of storage servers. It is a solution that provides the outboard intelligence required by Storage Area Network (SAN) solutions, off-loading key functions from host servers and freeing up valuable processing power for applications.

The key features on the IBM ESS can be seen in Figure 2-4.


<ul style="list-style-type: none">Two 4-way RISC processors64-bit 255MHzUp to 11 TB capacity8 x 160 MB/sec SSA loops8GB, 16GB, 24GB or 32GB cache384 MB NVSUp to 32 ESCON channel portsUp to 16 FICON (lw / sw) channel portsUp to 32 SCSI portsUp to 16 FC (lw / sw) portsConnects to Storage Area Networks	<p><i>Enterprise Storage Server</i></p> 
---	--

Figure 2-4 ESS overview

The ESS contains two clusters each having a 4-way 64-bit RISC processor. It also has 384 MB of non-volatile storage (NVS), and options of either 8 GB, 16 GB, 24 GB, or 32 GB of cache. The ESS scales up to 11.2 TB of usable capacity with 36 GB disks. For servers with ESCON or FICON channels like the IBM zSeries 900 family from the IBM @server brand, the ESS can connect to them via 32 ESCON links or 16 FICON links. For servers with SCSI bus or Fibre Channel link adapters, the ESS can connect to them either via 32 SCSI ports or 16 Fibre Channel ports. Many more hosts can be accommodated if a SAN fabric is implemented to provide connectivity between the hosts and the ESS.

The ESS also provides Advanced Copy Services for backup-recovery and disaster-recovery situations. As a comprehensive SAN-based storage solution, the ESS provides the management flexibility to meet the fast-paced changing requirement of today's business.

2.2.2 ESS models and expansion enclosure

The ESS has two models available, the model F10 and the model F20. Both models can be populated with the same internal disks. Both models also have the same clusters, RISC processors, host attachments and SSA Adapters. The difference between the two models is in the power supply, the resulting ability to attach an expansion rack, and the final resulting scalability. Figure 2-5 shows the differences between the models.

The expansion rack attaches only to the model F20. It provides for an additional 256 disks in four cages. This brings the total disk capacity of the model F20 to 384 disks.

<p style="text-align: center;">Enterprise Storage Server models</p> <ul style="list-style-type: none">• 2105-F20 Enterprise Storage Server<ul style="list-style-type: none">– Three phase power supply– Supports maximum capacity of 128 disks in base rack– Feature for expansion rack for additional capacity• 2105-F10 Enterprise Storage Server<ul style="list-style-type: none">– Single-phase power supply– Restricted to a maximum of 64 disks in base rack– No feature for expansion rack <p style="text-align: center;">ESS expansion enclosure</p> <ul style="list-style-type: none">• 2105-F20 ESS expansion rack feature<ul style="list-style-type: none">– Three-phase power supply– Supports up to 256 disks in four ESS cages

Figure 2-5 ESS models

The Enterprise Storage Server initially was introduced with two earlier models, the E10 and the E20. The old E models and the new F models do not differ in external appearance, but they do differ in throughput and performance due to some internal changes such as the introduction of 64-bit RISC processors, the increase in the number of PCI buses, and the increased cache sizes available in the F models. The increased bandwidth allows for up to 100% greater throughput for sequential workloads or 25% greater throughput for database workloads. Thanks to its Seascapes Architecture design, we are able to upgrade currently installed E models to become F models.

2.2.3 ESS benefits

The ESS can help achieve business objectives in many areas; it provides a high-performance, high-availability subsystem with flexible characteristics that can be configured according to any variety of requirements.

Storage consolidation

The ESS attachment flexibility, and large capacity, enable consolidation of data from different platforms onto a single high performance, high availability box. Storage consolidation can be the first step towards server consolidation, reducing the number of boxes to manage and allowing the flexibility to add or assign capacity when and where it is needed. The ESS supports all the major server platforms, from all the IBM @server series of servers, to the non-IBM Windows NT, Netware, Linux, and different flavors of UNIX servers. With a capacity of up to 11.2 TB, and up to 32 host connections, an ESS can meet high capacity requirements and performance expectations.

Performance

The ESS is a high performance design which takes advantage of IBM's leading technologies. In today's world, business solutions need to deliver high levels of performance continuously every day, day after day. They also need to handle different workloads simultaneously, so as to enable the running of Business Intelligence models, large databases for Enterprise Resource Planning (ERP), and online and Internet transactions alongside each other.

The ESS is designed to provide the highest performance, for the different type of workloads that may be found, even when mixing those dissimilar workload demands. For example, the zSeries 900 servers and open systems servers put very different workload demands on the storage subsystem. A server like the zSeries 900 typically has an I/O profile that is considered very *cache-friendly*, taking full advantage of the large cache sizes available in the ESS. On the other hand, an open systems server has an I/O profile that is often very *cache-unfriendly*, because most of the hits are made to the buffers defined in the processing server itself.

So, when an open systems server does I/O to the ESS (that is, a cache miss in its own buffers), the chances of the required data being in the ESS cache are somewhat reduced. In this case, as the data is more likely to be on the physical disk, actual disk speed is much more important. For the zSeries 900 type of workload, the ESS has a large cache and, most important, it has very sophisticated cache algorithms. The ESS has SSA high performance disk adapters that address the needs of workloads that do not benefit so much from cache size.

Parallel Sysplex I/O management

In the zSeries 900 Parallel Sysplex environments, the Workload Manager (WLM) controls where work is run and optimizes the throughput and performance of the total system. The ESS provides the WLM with more sophisticated ways to control the I/O across the Sysplex. These functions include parallel access to both single system and shared volumes and the ability to prioritize the I/O based upon WLM goals. The combination of these features significantly improves performance in a wide variety of workload environments.

Disaster Recovery and Availability (PPRC)

The ESS is designed with no single points of failure. It is a fault tolerant storage subsystem which can be maintained and upgraded in a live production environment without interrupting operation. Some of the functions that contribute to this capability in the ESS are shown in Figure 2-6.

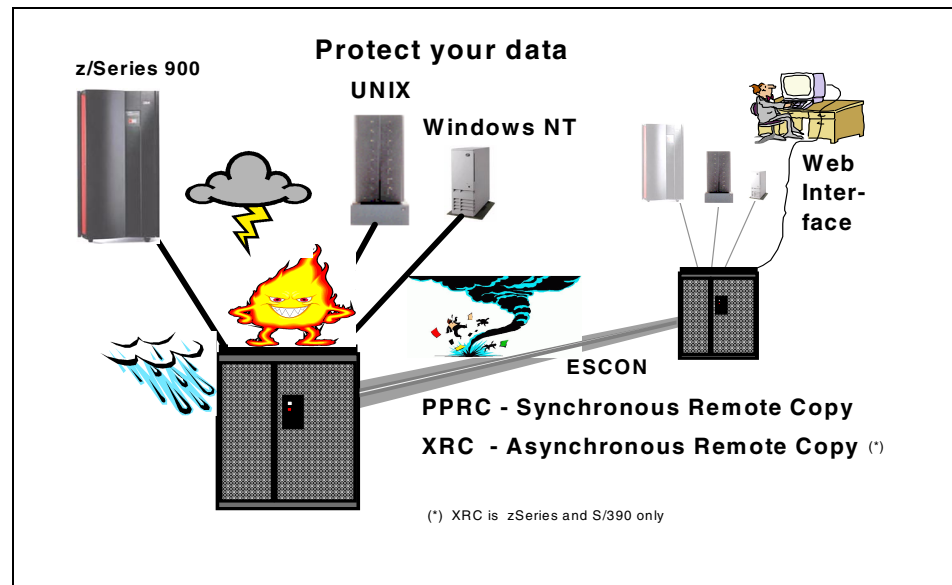


Figure 2-6 Disaster Recovery and Availability

The Peer-to-Peer Remote Copy (PPRC) function is a hardware solution that enables the shadow-mirroring of application system data from one site and its associated logical volumes, to a secondary site subsystem. Updates made on the primary logical volumes are synchronously shadowed on the secondary site logical volumes.

For UNIX and Windows environments, management of the PPRC setup is done through the StorWatch ESS Specialist Web interface. The ESS also provides a command line interface (CLI) for invocation and management of the PPRC functions through batch processes and scripts. The CLI can be run from AIX, HP-UX, Solaris, and Windows servers. This way, solutions for disaster are available for open systems platforms using a simple and easy-to-use interface.

For the zSeries 900 servers, the PPRC setup can also be managed using TSO commands or the ICKDSF utility. PPRC, together with Geographically Dispersed Parallel Sysplex (GDPS) for zSeries 900 servers, lead the industry in high availability solutions.

Note: PPRC is not currently supported for native Linux environments. Linux running as a guest operating system(s) within VM or z/VM has support as per VM or z/VM.

Extended Remote Copy (XRC), the z/OS disaster recovery solution, can be used over very long distances. XRC is an asynchronous remote copy solution offered on the ESS. It is a combined hardware and software solution that offers--over the longest distances--the highest levels of data integrity and availability for disaster recovery, workload movement, and disk migration.

FlashCopy

Customers still need to take backups to protect data from logical errors and disasters. For most environments, taking backups of user data takes a long time. Backups are often taken outside prime shift because of the impact to normal operations. Databases must be closed to create consistency and data integrity, and the online systems are normally shut down. To minimize the impact of these type of activities, the ESS has FlashCopy.

FlashCopy creates a physical point-in-time copy of data and makes it possible to access both the source and target copies immediately. By creating an “instant” copy, it enables applications using either the source or the target to operate with only a minimal interruption to I/O. On all server platforms, FlashCopy may be controlled using the StorWatch ESS Specialist. Under z/OS, FlashCopy can also be invoked using DFSMSdss. In addition to the ESS Specialist and z/OS interfaces, the ESS also provides a command line interface (CLI) for invocation and management of FlashCopy functions through batch processes and scripts.

Note: FlashCopy is not currently supported for native Linux environments. Linux running as a guest operating system within VM or z/VM has support as per VM or z/VM.

Storage Area Networks

The IBM ESS is a comprehensive Storage Area Network (SAN) disk storage subsystem providing solutions to today's most demanding Business Intelligence, e-business, server consolidation and ERP requirements.

The ESS continues to deliver on its SAN strategy, as was previewed in the July 27, 1999, announcement of the ESS Models E10 and E20. The ESS Models F10 and F20 now provide up to sixteen 100 MB/s native Fibre Channel short-wave or long-wave adapters. Each single port adapter supports Fibre Channel Protocol (FCP) in a direct point-to-point configuration, point-to-point to a switch (fabric) configuration, or Fibre Channel-Arbitrated Loop (FC-AL) in a private loop configuration.

Fabric support includes the IBM SAN Fibre Channel switch (2109 Model S08 and S16), McDATA Loop Switch (ES-1000), McDATA Switches (ES-3016, ES-3032), McDATA Enterprise Fibre Channel Directors (ED-5000, ED-6064), INRANGE Fibre Channel Director (FC/9000) and IBM Fibre Channel Storage Hub (2103-H07). All these SAN connectivity options make the ESS the unquestionable choice when customers plan for their SANs.

2.3 IBM Modular Storage Server

This topic provides a brief overview of the major IBM Modular Storage Server (MSS) components, features and benefits. For more detailed information, please refer to the redbook, *IBM Modular Storage Server - An Introduction Guide*, SG24-6103.

Note: Linux is not currently supported on the MSS but is expected by year end 2001. The information contained in this redbook is intended for people who wish to experiment with Linux on the MSS without IBM support.

2.3.1 IBM Modular Storage Server overview

The IBM 2106 Modular Storage Server (MSS) combines the high-performance of Fibre Channel technology with the flexibility and scalability of modular components.

The key features on the IBM MSS can be seen in Figure 2-7.

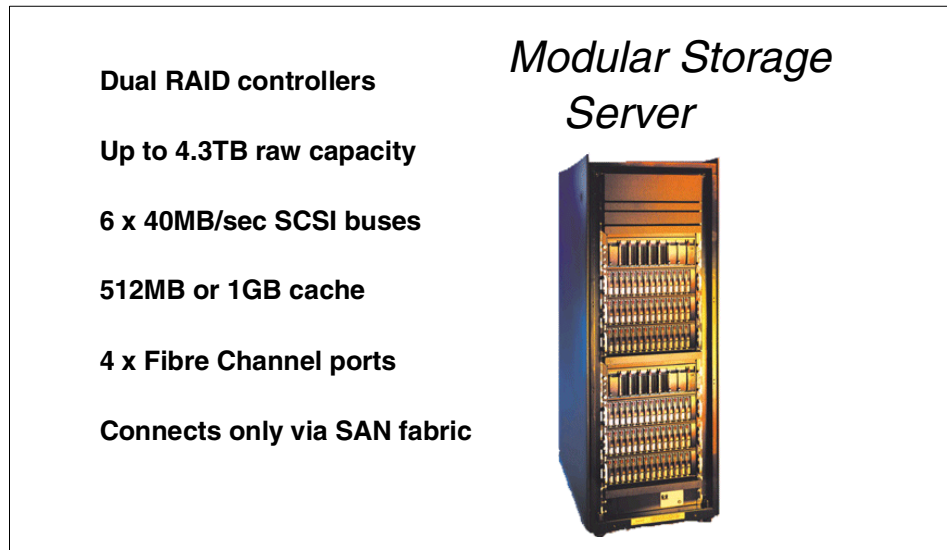


Figure 2-7 IBM Modular Storage Server overview

The Modular Storage Server is a fibre-attached RAID controller. It uses six SCSI buses on the backend to attach hard drive storage into a Fibre Channel (FC) network. Supported RAID levels include RAID 0, RAID 1, RAID 3/5 (automatically uses synchronous data transfer from the RAID 3 specification when applicable), and RAID 1+0.

Advanced features include FlashCopy, Peer-to-Peer Remote Copy (PPRC), and device cloning (copying a given Logical Unit Number (LUN) and breaking it as an independent device).

The MSS is similar to the Compaq MA8000. While many of the components are the same, they are not the same offering. IBM disk drives are used extensively in the MSS. Keep this in mind if referencing documentation for the MA8000 from Compaq.

2.3.2 MSS models and expansion enclosures

The IBM Modular Storage Server (MSS) is a high-availability, high-performance Fibre Channel storage subsystem providing scalable capacity, RAID protection, and ease of management.

Modular Storage Server models

- 2106-200 Modular Storage Server
 - supports up to 4.3TB raw capacity with 6 expansion enclosures

MSS expansion enclosures

- 2106-D10
 - single SCSI bus configuration
 - up to 6 per controller
 - up to 10 x 72GB drives
- 2106-D20
 - dual SCSI bus configuration
 - up to 3 per controller
 - up to 10 x 72GB drives
- 2106-D14
 - single SCSI bus configuration
 - up to 6 per controller
 - up to 12 x 18GB or 36GB drives
- 2106-D24
 - dual SCSI bus configuration
 - up to 3 per controller
 - up to 12 x 18GB or 36GB drives

Figure 2-8 MSS models

As shown in Figure 2-8, there are two major components to the modularity of the MSS — the controller enclosure and the disk drive enclosures. The controller enclosure (Model 200) contains two RAID controllers each with 256 MB or 512 MB of read/write cache, a total maximum cache size of 1 GB. There are four models of disk drive enclosures (Models D10, D14, D20, and D24), and up to six of these disk drive enclosures can be attached to the controller enclosure providing a maximum storage capacity of 4.32 TB.

2.3.3 MSS benefits

The MSS is one of several IBM disk subsystems available to solve storage needs. Choosing the right disk subsystem requires some knowledge of the customer storage environment, preferences, performance, data protection, host attachment, availability requirements, and affordable price range. The MSS is designed to offer flexible, highly-available, and fully-protected disk storage for a variety of UNIX, Intel, or Alpha-based servers.

Flexibility

Combined with IBM's Fibre Channel switches and hubs, flexible Storage Area Networks (SANs) can be created with a variety of servers attached, including those with UNIX, Windows NT, Windows 2000, Novell NetWare, or OpenVMS operating systems (Linux soon). Depending on specific requirements, we can start with just a few disk modules and grow the system to over 4 TB, using 72-GB disk modules.

Disaster Recovery and Availability

To reduce downtime, the MSS includes hot-swappable redundant components such as controllers, power supplies and fans. Support for clustered servers running Windows NT, Novell NetWare, and UNIX provide host-level redundancy.

In case of a disaster, Peer-to-Peer Remote Copy (PPRC) is designed to provide rapid recovery. PPRC retains data on-line, in real time, on a remote MSS system that can be miles away.

FlashCopy

FlashCopy provides point-in-time copy for backups. This routine operation can be conducted without compromising production systems. StorWatch MSS Specialist provides a single management console so as to easily configure, monitor and manage all MSSs in the SAN.

2.4 IBM Fibre Array Storage Technology (FAStT200)

This topic provides a brief overview of the major IBM FAStT200 components, features and benefits. For more detailed information, please refer to the redbook, *Fibre Array Storage Technology - A FAStT Introduction*, SG24-6246.

2.4.1 IBM FAStT200 overview

The IBM FAStT200 Storage Server provides flexible, affordable storage for Intel-based servers. The FAStT200 is designed for workgroup and departmental servers that require an external storage solution. The single controller model provides a cost-effective solution, while the FAStT200 High Availability (HA) model features a fully redundant configuration with dual redundant controllers. As storage requirements grow, additional storage capacity is easily added with the IBM FAStT EXP500 Expansion Units.

The key features of the IBM FAStT200 can be seen in Figure 2-9.


<p>Up to 2 x RAID controllers</p> <p>Up to 4.3TB raw capacity</p> <p>Up to 2 x 100MB/sec FC loops</p> <p>128MB or 256MB cache</p> <p>Up to 2 x Fibre Channel host connections</p> <p>Direct and SAN fabric connect</p>	<p><i>FAST200</i></p> 
--	--

Figure 2-9 FastT200 overview

The FAST200 features industry-leading Fibre Channel host attachment with Fibre Channel drives. The FAST200 HA model provides dual Fibre Channel host attachment with an aggregate bandwidth of 200 MB/s. By connecting the FAST200 to an IBM SAN Fibre Channel switch or hub, data transfer at distances over several kilometers is possible.

As a rack-mount solution, the FAST200 can support up to ten Fibre Channel hard disk drives internally in a 3U (EIA units) space. It supports 18 GB, 36 GB, and 73 GB 10,000 RPM drives, and 18 GB 15,000 RPM drives. The attachment of five FAST EXP500 expansion units (containing up to 50 drives) increases storage capacity to more than 4 TB.

At the time of producing this redbook, the FAST200 supported Windows NT and Windows 2000. By the time this book is published, Netware, Linux, HP/UX and Solaris support are expected to be announced. AIX support is expected towards the end of 2001 or early in 2002.

2.4.2 FAST models and expansion enclosure

The FAST200 units are ideal for creating a cost-effective storage area network (SAN) using fibre-to-fibre technology. The storage servers are designed for easy service. Hot-plug components, Light-Path Diagnostics, and customer replaceable units (CRUs) are used throughout to minimize downtime and service costs.

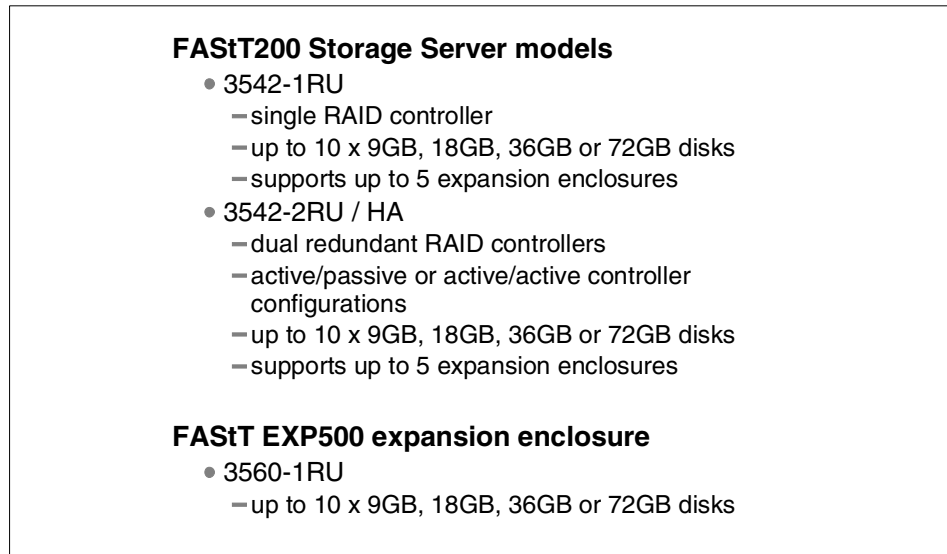


Figure 2-10 *FastT200 models*

As seen in Figure 2-10, the FAST200 Storage Server is an entry-level machine containing a single RAID controller and dual redundant hot-swap power supplies and fans. This unit can be upgraded for high-availability applications by installing an optional FAST200 redundant RAID controller which effectively makes it a FAST200 HA Storage Server.

The FAST200 HA Storage Server is designed for high-availability applications requiring a high degree of component redundancy. It is identical to the FAST200 Storage Server except that it contains dual redundant hot-plug RAID controllers.

2.4.3 FAST200 benefits

The FAST200 is very flexible unit. The main features and benefits are described in the following section.

Flexibility

The IBM FAST Storage Manager is a network-based, integrated storage management tool that enables administrators to configure, monitor, dynamically change, and manage multiple FAST200 Storage Servers from a single Windows 95 or Windows NT workstation. The software provides remote management capabilities and the ability to share storage resources among servers through storage partitioning. The FAST Storage Manager supports up to 128 logical units to increase configuration flexibility for large storage arrays. The FAST200 supports up to sixteen host systems or eight two-node clusters.

The Netfinity FAStT Storage Manager V7.10 is standard with both the FAStT200 HA Storage Server and FAStT200 Storage Server. This Java-based tool for managing Fibre Channel storage products provides support for up to sixteen storage partitions. Benefits include:

- ▶ Through storage partitioning, we can consolidate storage management functions by defining storage access for up to sixteen hosts to create a single storage subsystem. This provides greater flexibility to manage the storage array within the boundaries of:
 - Enterprise security
 - Data access models
 - Operating system driver limitations
- ▶ Amortization of costs can be realized through storage partitioning as the costs of sharing highly available data storage can be spread across multiple server environments.
- ▶ With the dynamic storage allocation, greater storage capacity requirements can be easily handled by adding more hard drives to the array without taking the storage subsystem down.
- ▶ Improved performance is achieved by using greater numbers of large capacity drives (with lower cost per megabyte) to support multiple servers. This increases the number of spindles to improve performance over installations where each server has its own storage subsystem.

Disaster recovery and availability

The FAStT200 helps ensure high availability by utilizing redundant, hot-swappable components such as power supplies and fans. The FAStT200 HA model features dual redundant RAID controllers with transparent failover support to further increase availability. With multiple RAID levels (0, 1, 3, 5, and 10), the FAStT200 can help protect valuable data and keep business-critical applications up and running.

2.5 IBM SAN Fibre Channel Switch

The IBM Enterprise Storage Area Network (SAN) is a high-speed, interconnected switched fabric of centrally managed, multi-vendor heterogeneous servers and storage systems. The IBM Enterprise SAN can help companies derive greater value from their business information by enabling IT resource management and information sharing anytime, anywhere across the enterprise.

The IBM SAN Fibre Channel Switch 2109-S08 and 2109-S16 provides Fibre Channel connectivity to Intel-based servers running Windows NT/2000 or DYNIX-based and UNIX-based servers, Fibre Channel-attached disk storage including IBM Enterprise Storage Server, IBM Modular Storage Server, IBM FAStT200 and FAStT500, and tape subsystems supporting IBM Magstar 3590 Fibre Channel drives, IBM SAN Data Gateways for attachment of IBM Enterprise Storage Server disk systems, Magstar 3590, Ultrium 358X, Magstar MP 3570, or 7337 and 3502 DLT tape systems and libraries, and IBM Fibre Channel Hubs and Netfinity Fibre Channel Hubs.

The IBM SAN Fibre Channel Switch interconnects multiple host servers with storage servers and devices, creating a storage area network or SAN. An IBM SAN Fibre Channel Switch can be used either as a standalone device to build a simple SAN fabric, or interconnected with other switches to build a larger SAN fabric. The interconnection of IBM and IBM-compatible switches and hubs creates a switched fabric containing hundreds of Fibre Channel ports. The SAN fabric provides the high performance, scalability, and fault tolerance required by the most demanding e-business applications and enterprise storage management applications such as LAN-free backup, server-less backup, disk and tape pooling, and data sharing.

The IBM SAN Fibre Channel Switch operates up to 100 MB/s per port with full-duplex data transfer. Unlike hub-based Fibre Channel Arbitrated Loop (FC-AL) solutions which reduce performance as devices are added, the SAN fabric performance increases as additional switches are interconnected.

IBM offers two different types of IBM SAN Fibre Channel Switches which are OEM products from the Brocade SilkWorm family:

- ▶ IBM SAN Fibre Channel Switch Model 2109-S08 is an 8-port model.
- ▶ IBM SAN Fibre Channel Switch Model 2109-S16 is a 16-port model.

Figure 2-11 shows both the 8-port and the 16-port models.



Figure 2-11 IBM SAN Fibre Channel Switch 2109-S08 (top), 2109-S16 (bottom)

In the following sections, we describe the IBM SAN Fibre Channel Switches in greater detail in terms of features including high availability, system components, zoning, inter-switch links (ISLs) and performance.

2.5.1 Product description

The IBM SAN Fibre Channel Switch Model 2109-S16, as shown in Figure 2-12, is a 16-port, Fibre Channel gigabit switch.



Figure 2-12 IBM SAN Fibre Channel Switch 2109-S16

The IBM SAN Fibre Channel Switch Model 2109-S16 Switch consists of a system board with connectors for supporting up to 16 ports, and a SAN fabric operating system for creating and managing a SAN fabric.

SAN fabrics

A SAN fabric is an active, intelligent, and non-shared interconnection of multiple SAN Fibre Channel switches, and is also known as Cascaded Fabric. It is used to increase the amount of connectivity in a SAN, due to the limitation of number of ports per IBM SAN Fibre Channel Switch that supports up to 16 ports.

The Cascaded Fabric is also used to support fault tolerant fabric topologies, which eliminates single points of failure, and increases the maximum possible distance between interconnected devices. The high-end industry standard supports up to 7 switches in a cascade due to SAN operation latency. In a SAN fabric environment, multistage or mesh topologies should be considered.

Figure 2-13 shows the front view of the IBM SAN Fibre Channel Switch Model 2109-S16.

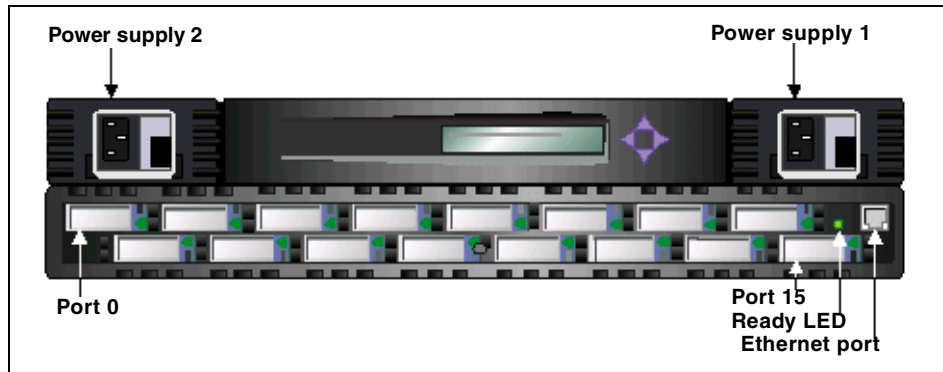


Figure 2-13 Front panel of the IBM SAN Fibre Channel Switch 2109-S16

The ports are numbered sequentially starting with zero for the left-most port. The switch faceplate includes a silk screen imprint of the port numbers. Up to two power supplies are supported; these are shown to the above left and right of the switch ports.

The IBM SAN Fibre Channel Switch Model 2109-S08, as shown in Figure 2-14, is an 8-port, Fibre Channel, gigabit switch.



Figure 2-14 IBM SAN Fibre Channel Switch 2109-S08

As in the IBM SAN Fibre Channel Switch 2109-S16, the Model 2109-S08 also consists of a system board with connectors for supporting up to 8 ports and a fabric operating system for building and managing a SAN fabric.

Figure 2-15 shows the front view of the IBM SAN Fibre Channel Switch 2109-S08.

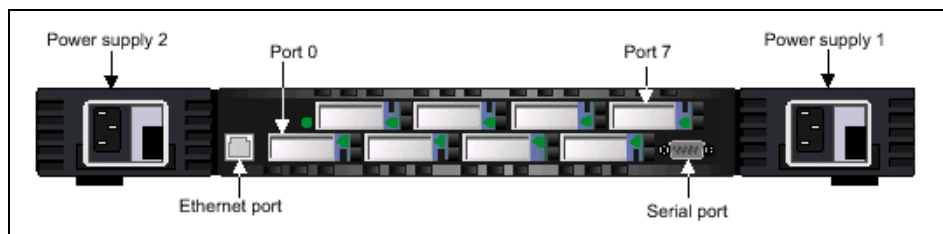


Figure 2-15 Front panel of the IBM SAN Fibre Channel Switch 2109-S08

The ports are numbered sequentially, starting with zero for the left port. The switch faceplate includes a silk screen imprint of the port number. Up to two power supplies are supported; these are shown to the left and right of the switch ports.

SAN switch selection

The IBM SAN Fibre Channel Model 2109-S08 switch is targeted at applications that include small NT/2000 or UNIX systems connecting three to five servers, storage, and tape together.

The IBM SAN Fibre Channel Model 2109-S16 switch is targeted at high-throughput and demanding applications.

The IBM SAN Fibre Channel Switch Model 2109-S08 comes with two ASICs per system board, whereas model S16 has four ASICs per system board with continuous and sustained performance, any port to any other port. Each ASIC provides four Fibre Channel ports that may be used to connect to external N_Ports (as an F_Port), external loop devices (as an FL_Port) or to other IBM SAN Fibre Channel switches (as an E_Port). Each port operates at 1.0625 Gb/s full-duplex, and its system architecture supports non-blocking shared memory with throughput of up to 100 MB/s with a total bandwidth of 800 MB/s (0.8 GB/s) for Model 2109-S08 and 1.6 GB/s for Model 2109-S16.

Each port comes with 16 buffers at 2112 bytes per frame. The switches support Fibre Channel class 2, 3 and F connectionless service operation and have a minimum latency of 1 μ s and a maximum of 2 μ s.

Classes of service

Class F is a connectionless service for interswitch control traffic. It provides notification of delivery or nondelivery between two E_Ports.

Class 2 is a connectionless service between ports with notification of delivery or non-delivery.

Class 3 is a connectionless service between ports without notification of delivery. Other than notification, the transmission and routing of Class-3 frames is the same as Class 2 frames.



Part 2

Implementing the NAS 300G in your storage network

Part 2 of this book starts the step-by-step walkthrough we promised. Chapter 3, “Implementing the IBM TotalStorage NAS 300G” on page 71 describes how to integrate your NAS 300G into a storage network for heterogeneous clients. Chapter 4, “Clustering for high availability” on page 167 covers implementing Tivoli SANergy to maximize the benefits of a SAN using your NAS 300G. Chapter 5, “Using SANergy to secure high-speed data sharing” on page 235 explains how to use Tivoli Storage Manager to back it all up.



Implementing the IBM TotalStorage NAS 300G

This chapter provides a step-by-step walkthrough, explaining how to integrate the IBM TotalStorage NAS 300G into a storage network for heterogeneous clients. We show how to connect the NAS 300G to various storage devices, how to access the storage from Windows and UNIX clients, and how to integrate the NAS 300G in your user and security management.

3.1 Sharing SAN-based storage through the 300G

We start with a step-by-step walkthrough which shows how to connect to NAS 300G to the ESS, MSS, and FAStT after re-initializing the system.

3.1.1 Getting started

The IBM TotalStorage Network Attached Storage 300G (300G) is designed and pre-configured to function as a headless appliance; however, you can also connect a keyboard, a monitor, and a mouse to it if you so desire. In fact, if you are installing only one or two units, you may prefer to work directly with the unit rather than going through the remote interface. Should you choose to configure it in this manner, you may be pleasantly surprised to note that, unlike some servers or headless appliances which do not bother to include much video support, the 300G has a fine video card that easily supports resolutions of 1024x768x32. This makes working with the system from a directly attached monitor a pleasure. In this section, we will describe both methods for initially configuring the 300G.

3.1.2 Re-initializing the 300G

To demonstrate how easy this process is, and to ensure that our system is in its pristine factory-shipped state, we will begin our work by re-initializing our system.

Re-initializing the unit is very simple, as it comes with a CD-ROM for this purpose. However, the unit also has a protection system to prevent it from being accidentally re-initialized if the CD-ROM is left in the drive during a reboot. To circumvent this protection, you must use the Recovery Enablement Diskette and follow this procedure:

Place the 300G System Recovery CD-ROM in the CD-ROM drive, insert the Recovery Enablement Diskette in the floppy drive, reset the system, and wait for the system to start beeping. This signal lets you know that the initialization protection has been overridden. Next, remove the Recovery Enablement Diskette from the floppy drive and press the reset button (*or **CTRL+ALT+ DEL** if you have attached a keyboard*). It will reboot itself a couple of times during this procedure, but approximately 15 minutes later, the unit will be restored to a pristine state.

Note: A Recovery Enablement Diskette can be made from the supplementary CD-ROM if one cannot be located. Simply follow the instructions in the readme file located in the DiskImages folder.

Remotely managing the 300G

Once the 300G is up and running, you need to install the IBM Advanced Appliance Configuration Utility (IAACU) software on the workstation that you will be using as your administrative console. This software uses Internet Explorer to remotely manage your 300G using a Web interface. It can be found on the 300G's Supplementary CD. Follow these steps to install the IAACU software:

1. Go to the IBM Advanced Appliance Configuration folder and run the setup.BAT file. This will launch a command window and prompt you to press any key to continue.
2. In the installation GUI, accept the default installation directory.
3. Choose **No, I will reboot my computer later** at the end of the GUI-based install.
4. Press any key to complete the batch installation from the command prompt window.
5. Reboot your workstation.

To begin remotely managing your 300G, start the IAACU by clicking **Start -> IBM Advanced Appliance Configuration Utility -> IBM Advanced Appliance Configuration Utility**. This will launch a window, as shown in Figure 3-1.

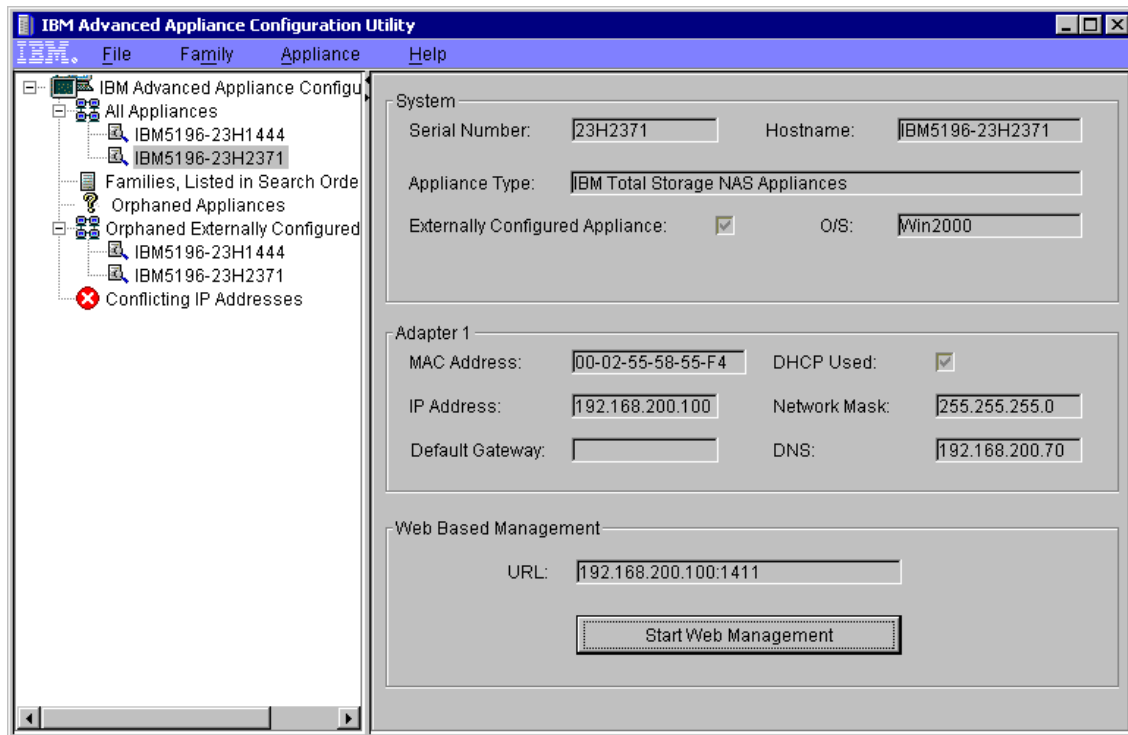


Figure 3-1 The IBM Advanced Appliance Configuration Utility

Select the 300G you want to manage from the list of **All Appliances** in the left pane and click the **Start Web Management** button this will launch Internet Explorer. If this is the first time you have run the Web management application, you will see the screen shown in Figure 3-2.

Note: The 300G will initially use its serial number as its name. If you do not know the serial number, it can be found on the front of the box in the lower right corner.

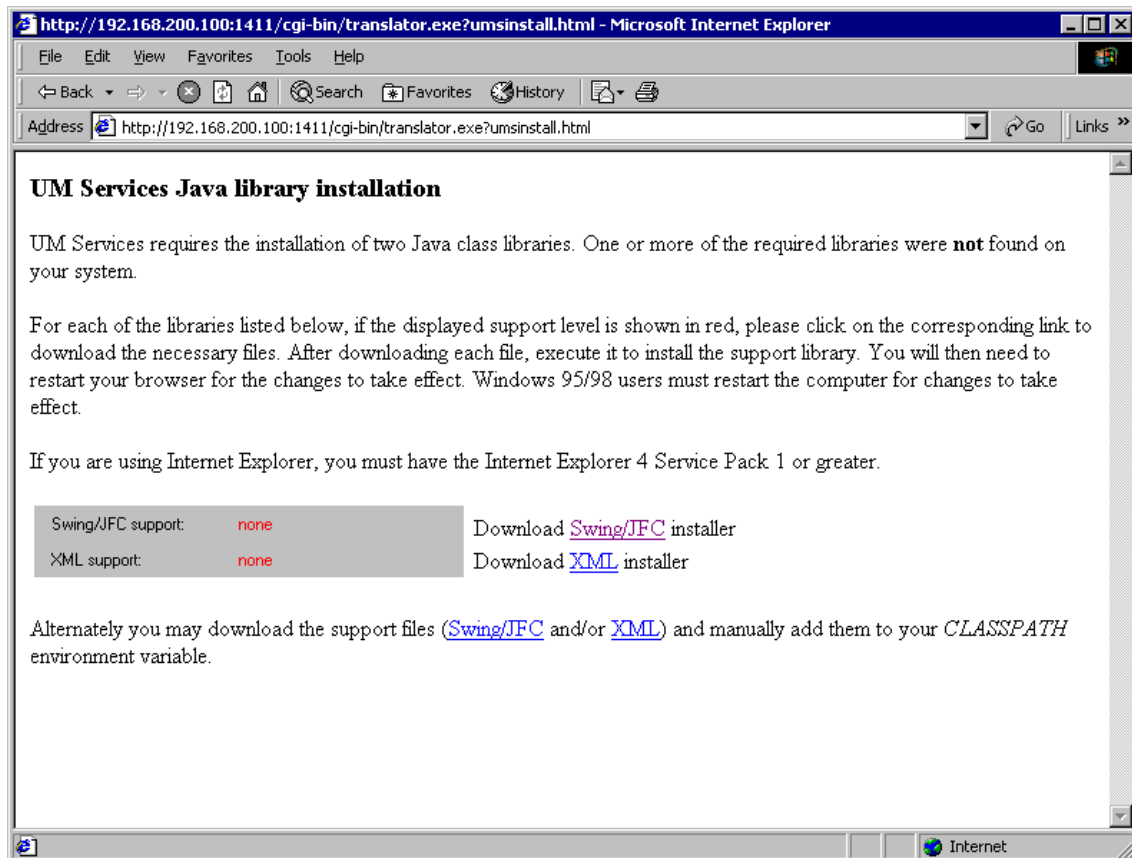


Figure 3-2 Initial setup of the remote Web Management application

Follow the on-screen directions to install the Swing and XML components this application requires. Then shut your browser down and launch it again (or reboot your computer if you are using a Windows 95 or 98 machine). Now you should finally see the remote management interface, as shown in Figure 3-3.

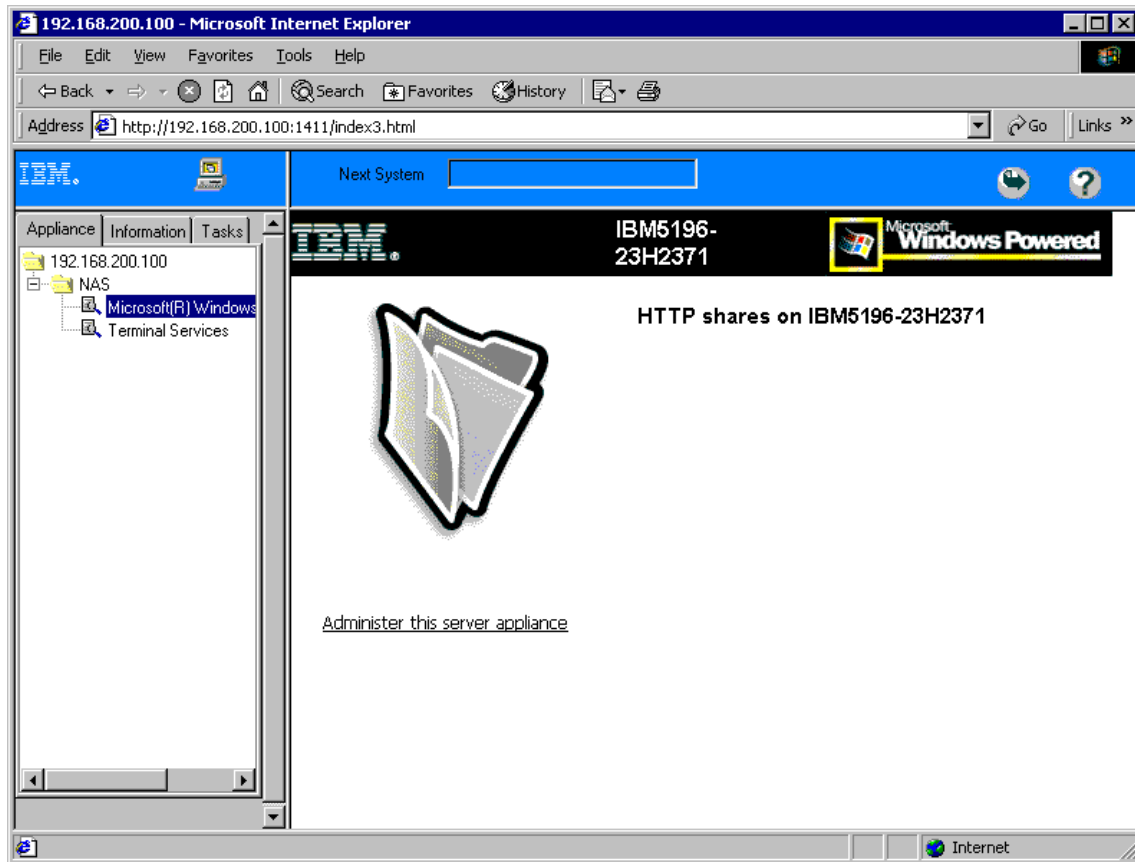


Figure 3-3 The remote Web Management application

Clicking the **Administer this server appliance** link will take you to the main menu screen. From here, you can set the date and time on your 300G by clicking **Maintenance -> Date and Time**, or change its name to something you think is more meaningful, change the administrative password, and change your network settings by clicking **Network Setup**.

If you prefer, you can also click **Maintenance -> Terminal Services Advanced Client** to remotely administer the 300G from its Windows desktop. The Windows Terminal Services client will let you interact with the 300G as if you had connected a keyboard, mouse, and monitor up to it and were sitting directly in front of it, as we describe in the following section.

Locally configuring the 300G

If you have directly attached a mouse, monitor, and keyboard to your 300G, you can configure it in exactly the same way you would configure a Windows 2000 workstation. Just login using the administrator account, right-click **My Network Places**, and choose **Properties**. Now highlight the adapter that will be used for remote management.

Note: The management adapter is the Local Area Connection associated to the IBM Netfinity Fault Tolerance PCI Adapter. This is the on-board port built into the planar board. However, if the system is clustered to another 300G, the on-board port is used for clustering only and the Local Area Connection 2 associated to the IBM 10/100 Ethernet Server Adapter is to be used for remote administration.

Right-click this adapter and chose **Properties**. Choose **TCP/IP** from the list of components and then click **Properties**. Then assign an IP address or the IP address of the DHCP server you want to use.

Note: The 300G will automatically try to find a DHCP server if no IP address is assigned. Assigning an IP address for the DHCP server is only to ensure the server assigning the address is the desired server. Additionally, the 300G will automatically default to the 169.254.X.X network with a subnet mask of 255.255.0.0 if no IP address is assigned and no DHCP server is located.

We are now ready to begin configuring our storage devices so the 300G can take ownership of them and subsequently share them with the LAN clients.

3.2 To SAN or not to SAN

The MSS cannot be connected directly to the 300G, so we only discuss connecting to it via our fabric. However, for both the FAStT200 and the ESS, we describe both connecting via the SAN and a point-to-point connection (although we did not do both methods simultaneously). The configuration of the FAStT200, the ESS, and the 300G are all identical regardless of how they are attached. We will configure the switch early on, to show the steps specific to setting up our SAN environment. We will be using the IBM 2109 as our fabric device. Additionally, we will be setting up a zone to isolate our devices from devices being used for other purposes.

3.2.1 Finding the World Wide Name

Getting the World Wide Name (WWN) from some units can be somewhat difficult, but the 300G conveniently provides a Microsoft Management Console (MMC) snap-in that makes this operation exceedingly simple. All you have to do is double-click the **IBM NAS Admin.msc** shortcut on the desktop of the 300G. This launches the MMC snap-in. Then just click **Storage -> NAS Utilities -> FAST Check**. This will launch the FAST Check application, as shown in Figure 3-4.

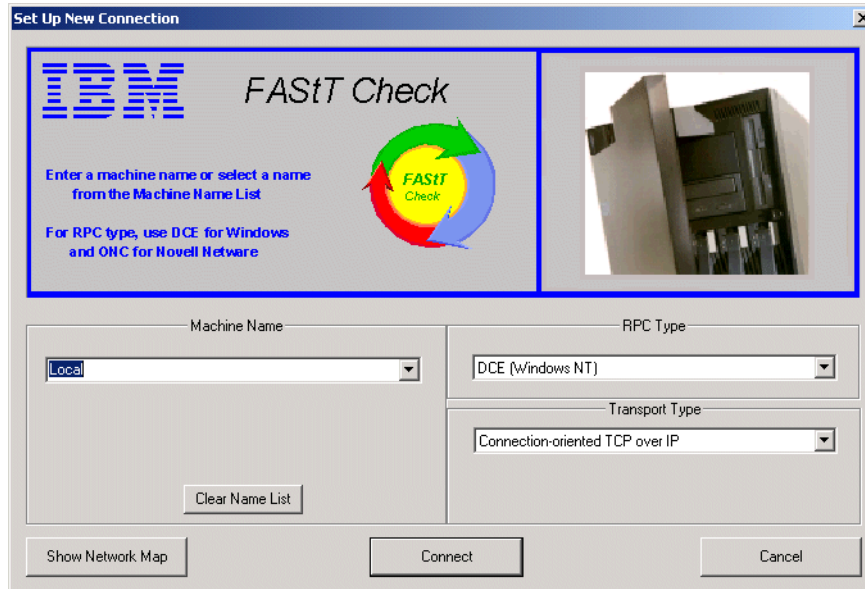


Figure 3-4 The FAST Check application running on the NAS 300G

Once the window launches, just click the **Connect** button.

If you are running the 2.0 version of the software, you may receive an error message like the one in Figure 3-5. You can safely ignore this error. Just click **OK** to continue past it.



Figure 3-5 Error message seen occasionally in the FAST Check application

The next window displays the WWN next to the QLA2200, as shown in Figure 3-6.

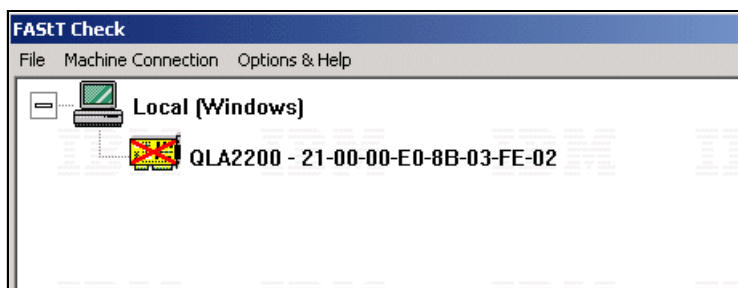


Figure 3-6 The 300G's WWN displayed in the FAST Check application

Tip: If your system has two adapters, you can make it simple to determine which adapter has which WWN by disconnecting one of the devices before retrieving the WWNs.

3.2.2 Zoning the IBM 2109

It is not our intention to point out *every* detail required to set up and/or configure the IBM 2109, but we will hit the highlights. To learn more about this subject, please refer to the excellent description in the *IBM SAN Survival Guide*, SG24-6143.

Note: Many other IBM redbooks which provide details on the equipment we used during this work are listed under “Related publications” on page 319.

Using a browser, go to the **Fabric View** of your 2109 by typing its IP address into the address bar of your browser. This should bring up a screen similar to the one in Figure 3-7.

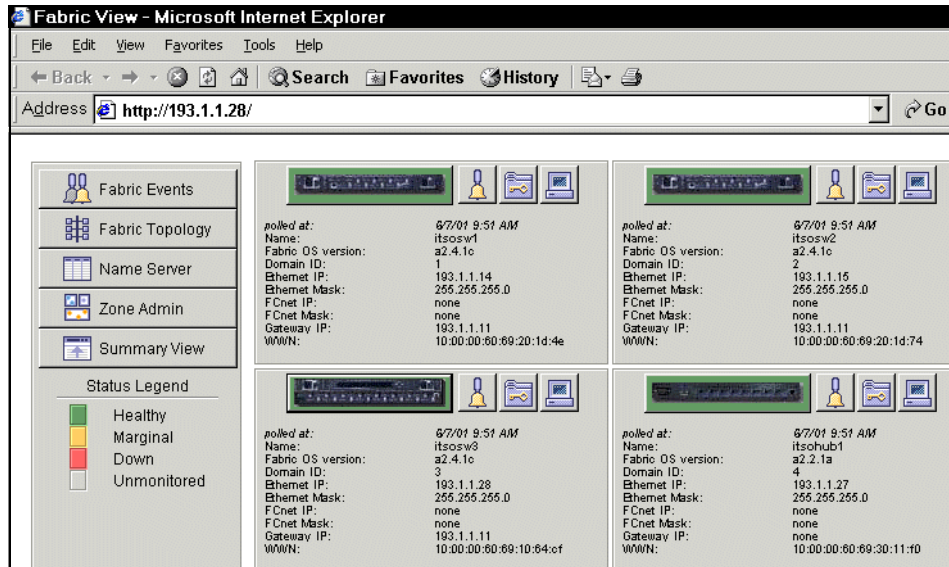


Figure 3-7 Fabric View of the 2109

Now select the switch to configure and press the **Zone Admin** button in the left pane. You will need to supply the user name and password in the dialog box shown in Example 3-8, and then click **OK**.



Figure 3-8 Zone login

Once you've logged in, create an alias. We are simply re-naming an existing alias, so our screen (Figure 3-9) may be different from yours.

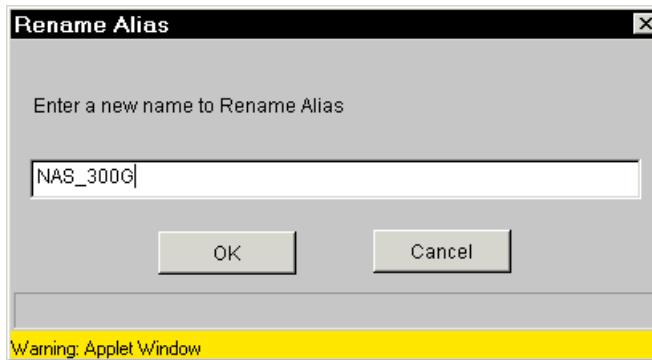


Figure 3-9 Rename Alias

Next, we will add the World Wide Name (WWN) of the devices to be associated to this alias. **Highlight** the WWN of the devices in the list on the left (Figure 3-10).

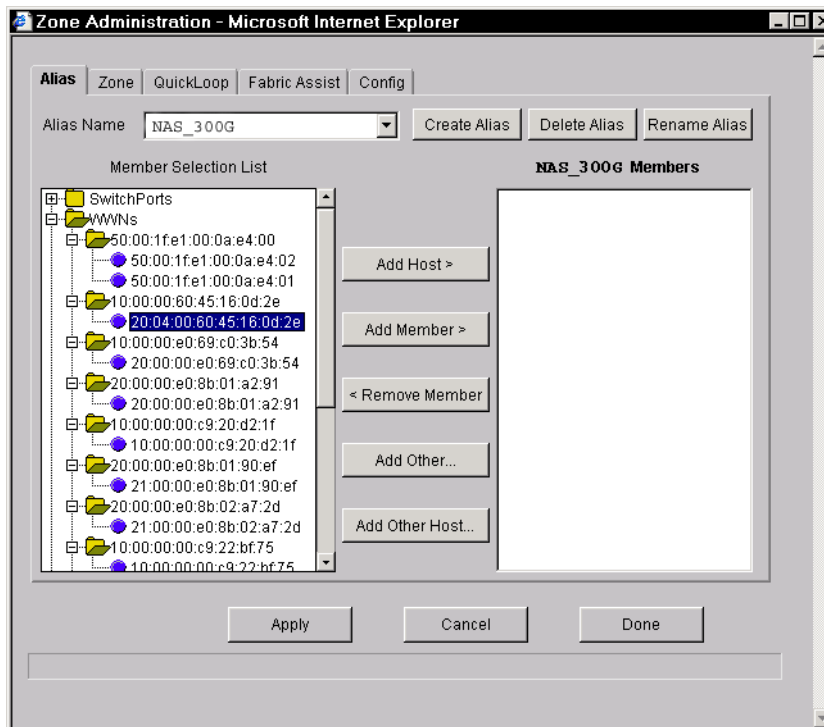


Figure 3-10 Locate WWN

Now associate the WWN to the alias you have created by pressing the **Add Member** button (Figure 3-11).

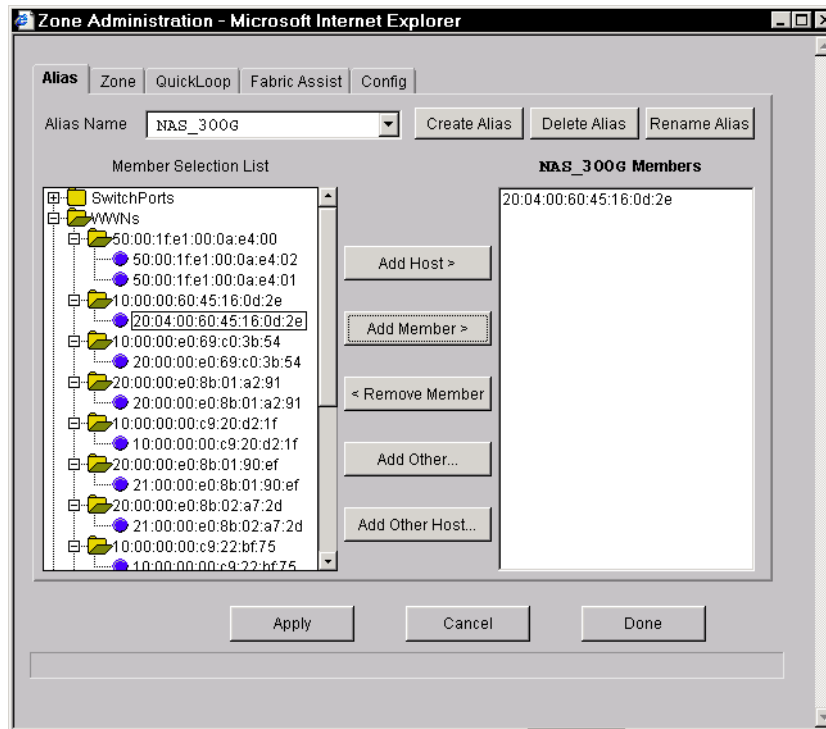


Figure 3-11 Add members

Next, we will make a zone to put the alias in. Select the **Zone** tab and click **Create Zone** (Figure 3-12).

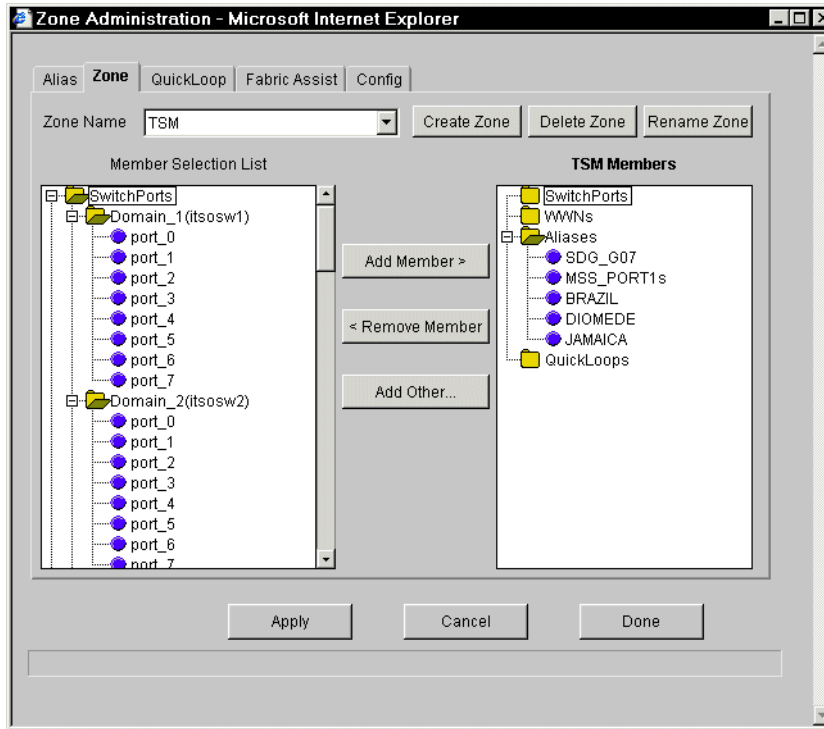


Figure 3-12 Zone creation

Enter the name of the zone you are going to add and click **OK** (Figure 3-13).

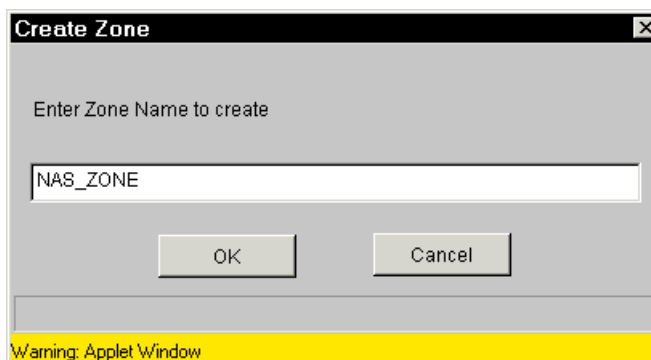


Figure 3-13 Create Zone

Now we can add the alias we created earlier to this zone. Open the **Aliases** folder and **highlight** the alias to add (Figure 3-14).

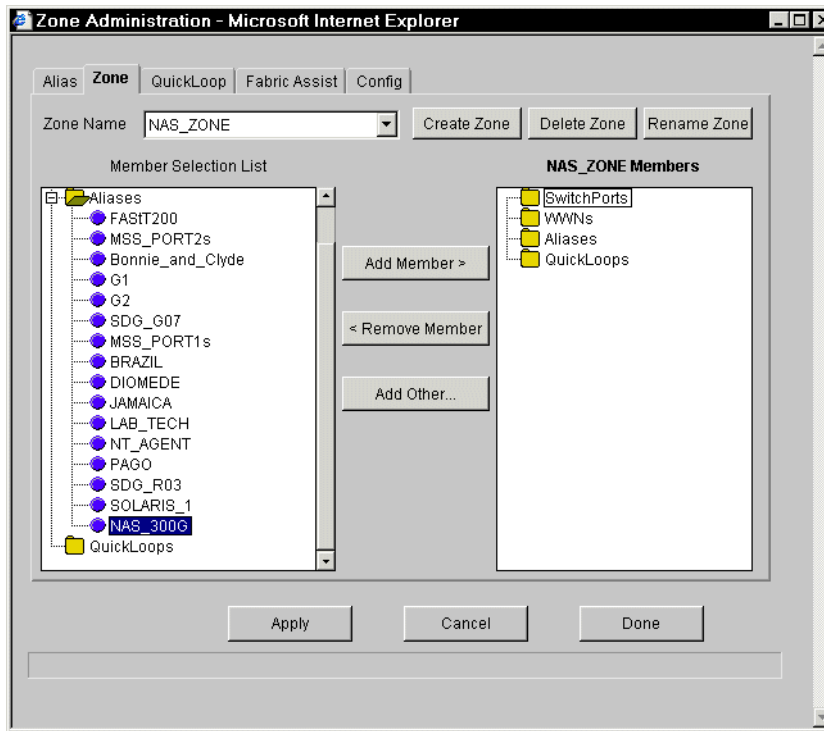


Figure 3-14 Select alias

Click the **Add Member** button. Your screen will look similar to the one shown in Figure 3-15.

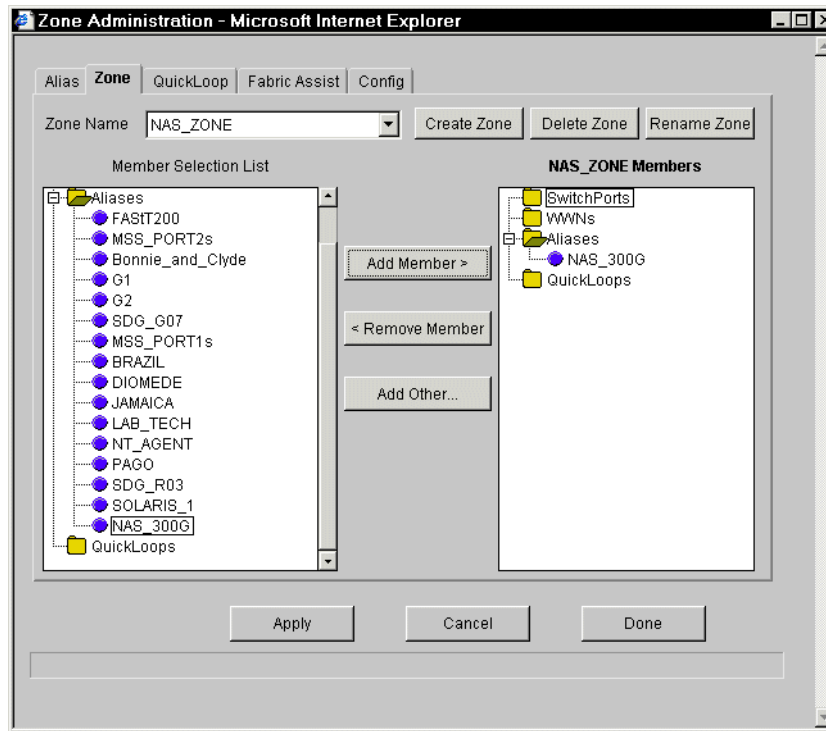


Figure 3-15 Add alias to zone

To finish this zoning process, choose the **Config** tab, **highlight** the configuration to add to the new zone to by using the **Config Name** pull-down menu, open the **Zones** folder, and **highlight** the newly created zone (Figure 3-16).

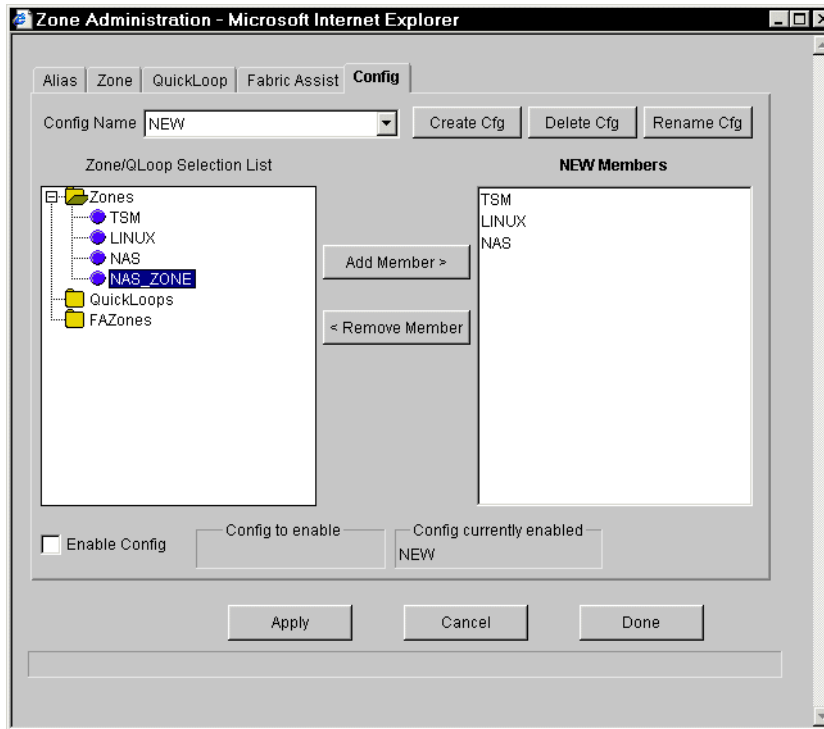


Figure 3-16 Select zone to add

Click the **Add Member** button, and then click the **Apply** button (Figure 3-17).

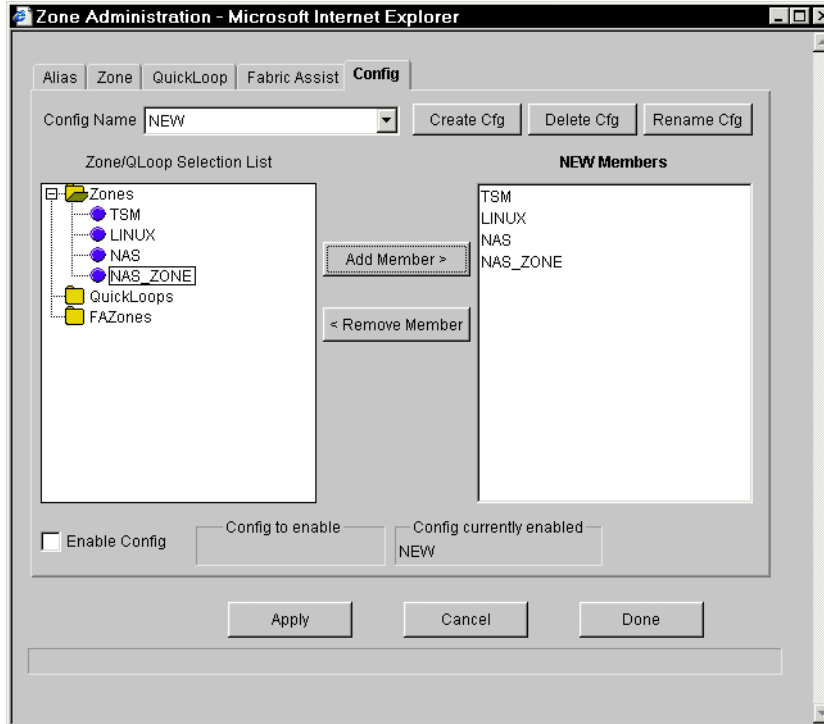


Figure 3-17 Add zone to configuration

Note: Only the zones listed in the NEW Members panel on the right will be active once the Apply button is pressed.

3.3 Setting up the FAStT200

If you are performing this setup in a production environment, there are many planning steps you would need to go through before proceeding. We do not cover all of those steps here; we just focus on the setup steps which occur after planning. If you need more comprehensive background information, we recommend that before proceeding, you take a look at the redbook, *Fibre Array Storage Technology - A FAStT Introduction*, SG24-6246.

Getting storage space in the FAStT200 to appear as drives to the 300G is a detailed process. Again, you will have to do some planning to make the FAStT200's storage space meet your requirements. Just as an example, we will create RAID5 drives from the free space in our FAStT200.

We are launching the **IBM FAStT Storage Manager 7** client from the system that is used to perform work on the FAStT200. Our opening screen appears in Figure 3-18.

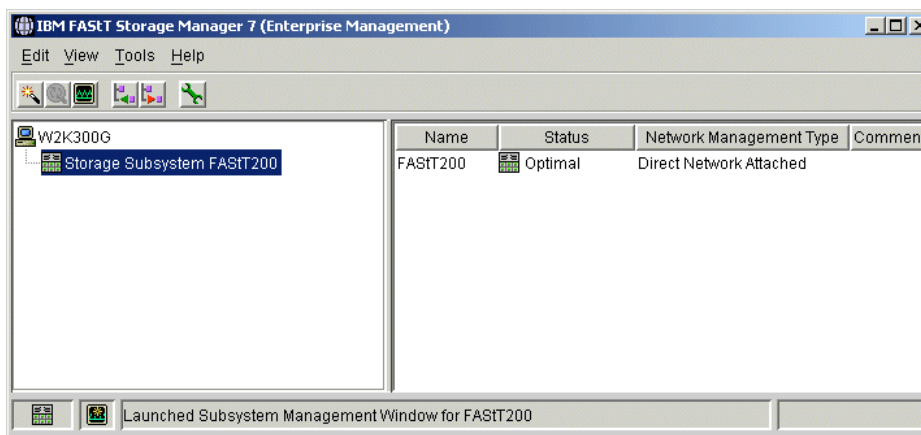


Figure 3-18 Storage Manager 7 main screen

Now we **highlight** the device we are going to create drives on and, using the **Tools** pull-down menu, select **Manage Device**. A new window will open. This window is shown in Figure 3-19.

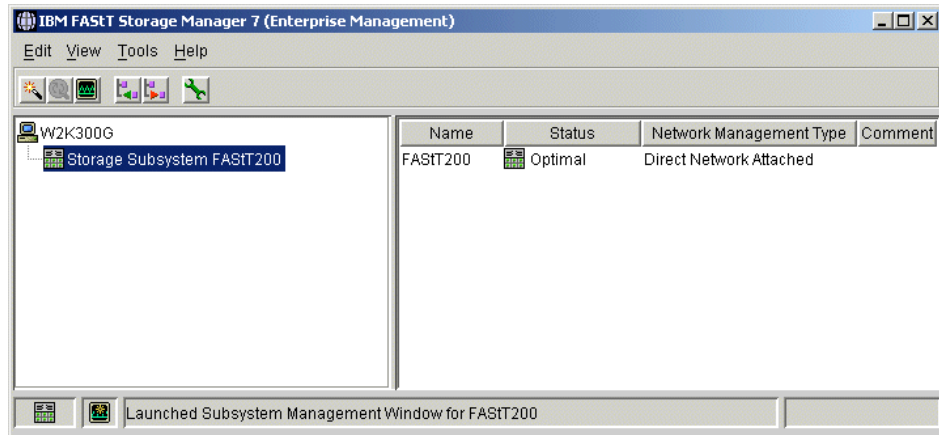


Figure 3-19 Subsystem management

Now we are going to create a new array. **Highlight** the unconfigured capacity available, then use the **Configure** pull-down menu and select **Create Array:Logical Drive** (Figure 3-20).

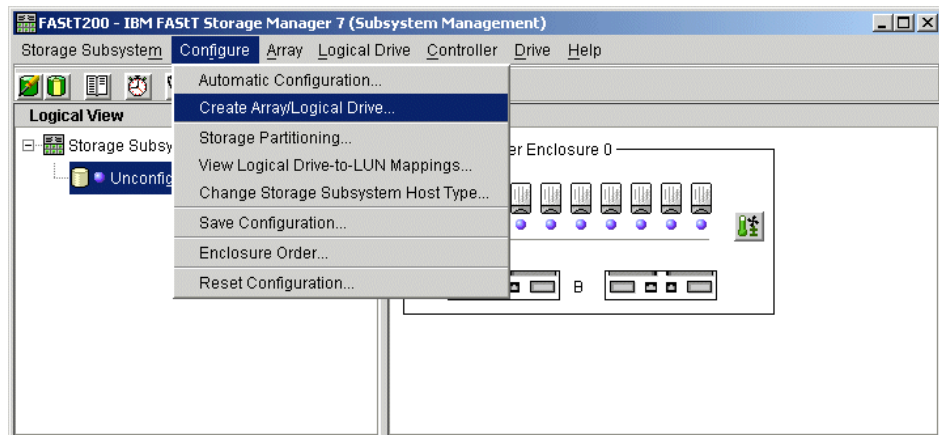


Figure 3-20 Create Array

We are now in a configurator that will walk us through the steps required to create the Array and Logical Drive, select the RAID level, and choose the capacity allocation method. We have selected RAID5 and let the configurator use its default allocation method of automatic. See Figure 3-21.

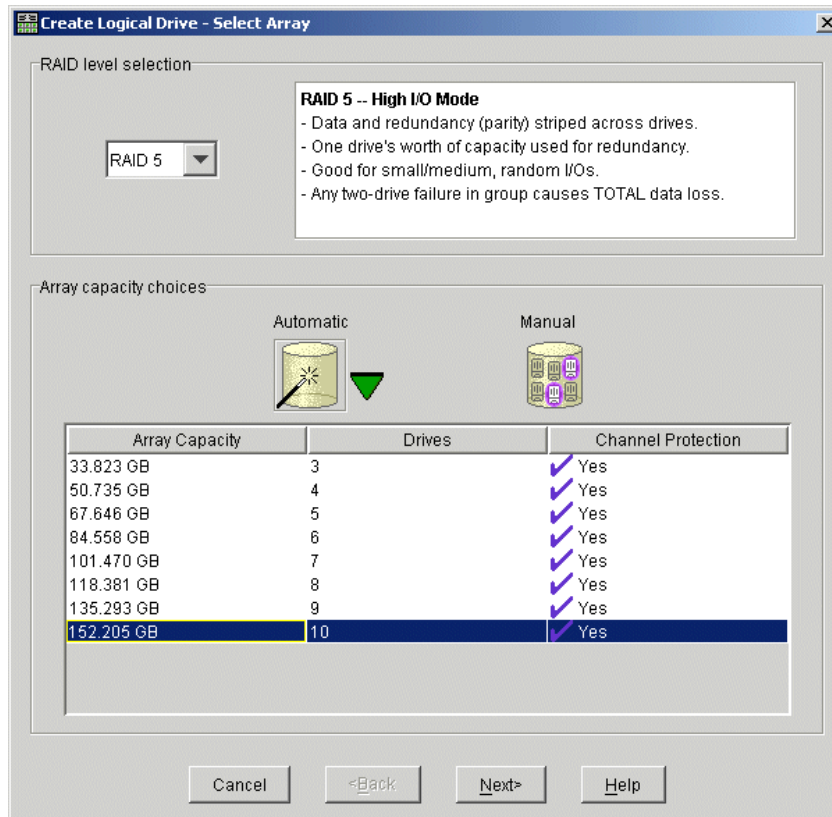


Figure 3-21 RAID level selection

If we did not have the necessary number of drives to support the RAID5 we chose, the configurator would not have shown any available space in the lower pane of the window. If we had selected the manual allocation method and tried to do a RAID5 array anyway, we would not have seen the array creation successful window, as shown in Figure 3-22.

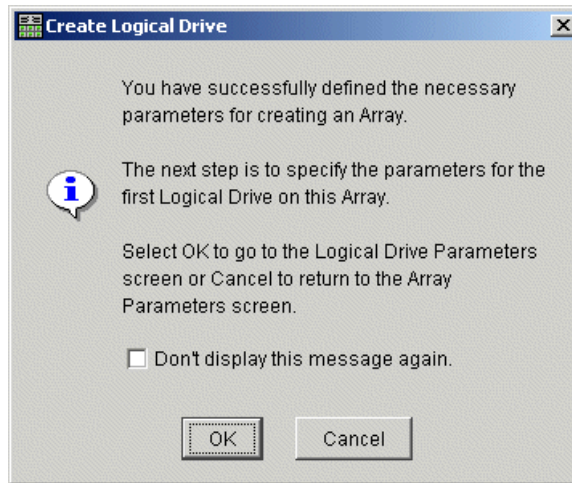


Figure 3-22 Array Creation Successful

Now that the array has been established, we need to continue on and this is done by selecting the **Logical Drive Parameters**. We have opted to see the capacity in megabytes and set the size of the drive. This screen also allows us to name our drive. Since we will be doing file sharing through the 300G, we set the type of usage we expect for this drive to “file system”. We pointed our segment size back to our expected usage setting and let the FAStT200 determine that automatically. We did not have a choice about the preferred ownership of our drive, due to the configuration of our FAStT200. Finally, we decided to do our Logical Drive to Logical Unit Number (LUN) mapping at the time of partitioning. This can all be seen in Figure 3-23. To complete this configurator, click **Finish**.

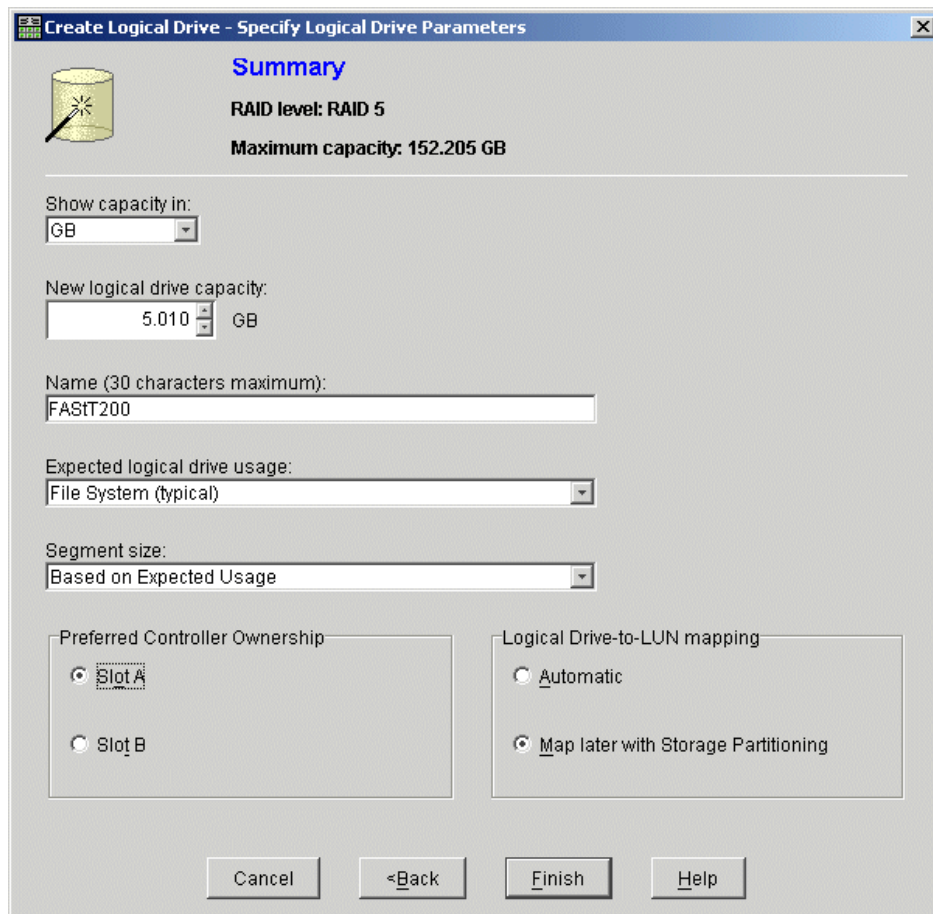


Figure 3-23 Specify Logical Drive Parameters

All of our parameters were acceptable and we successfully completed the Logical Drive Creation, as shown in Figure 3-24.

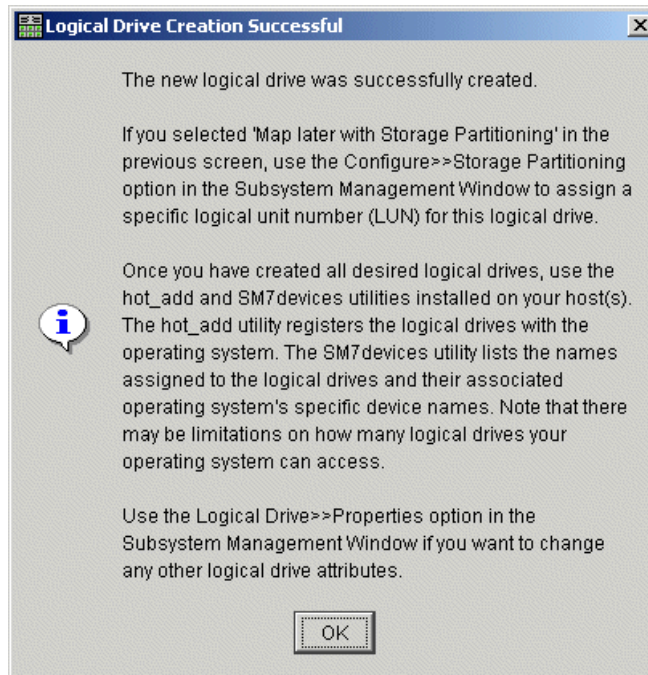


Figure 3-24 Logical Drive Creation Successful dialog

In order to set up our storage partitioning, we used the Subsystem Management window's **Configure** pull-down menu and selected **Storage Partitioning**, as depicted in Figure 3-25.

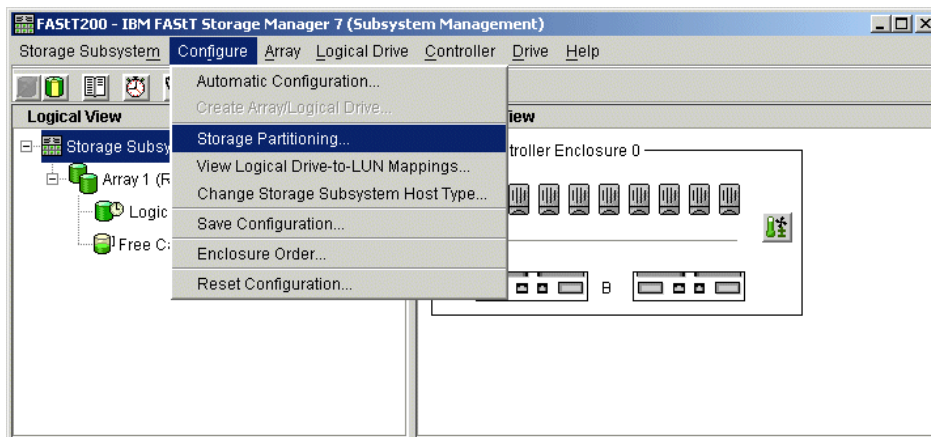


Figure 3-25 Storage Partitioning

Once we enter storage partitioning, we get a box that gives some hints as to what to do next, based on what has been done before, if anything (Figure 3-26). We want to create a new host group.

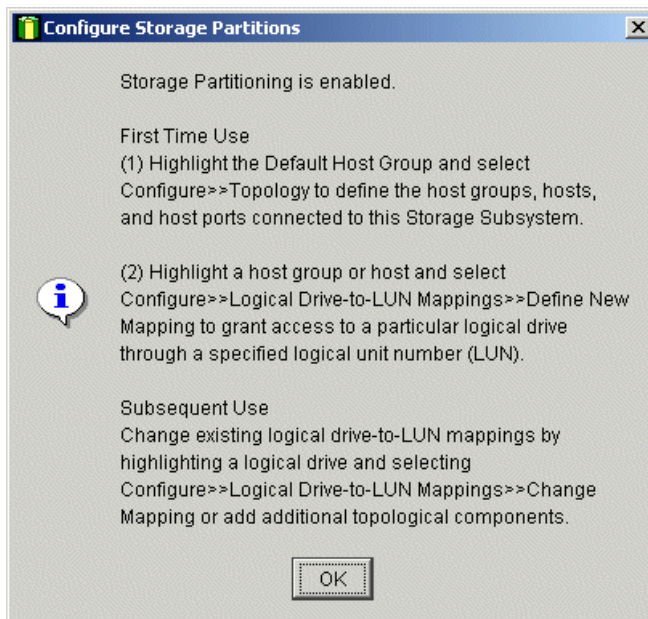


Figure 3-26 Configure Storage Partitions

In the FAST200's Mappings window, we used **Configure ->Topology -> Define New Host Group** (Figure 3-27).

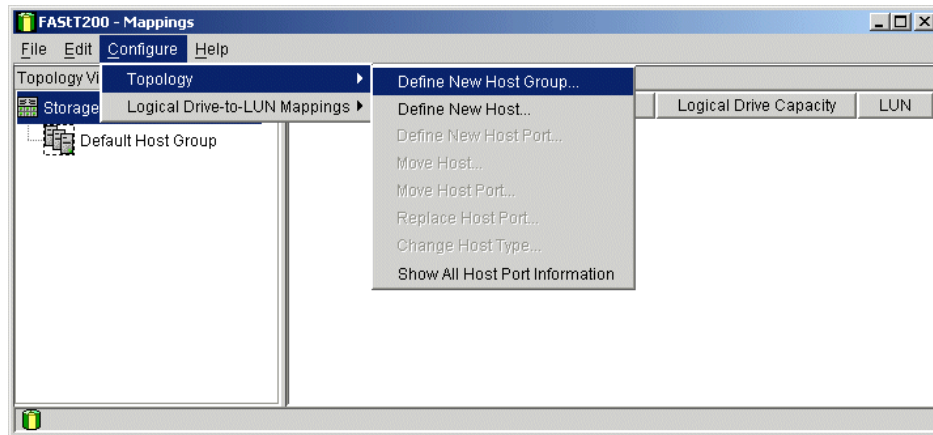


Figure 3-27 Define host group

Enter the name for the host group and click **Add** (Figure 3-28).

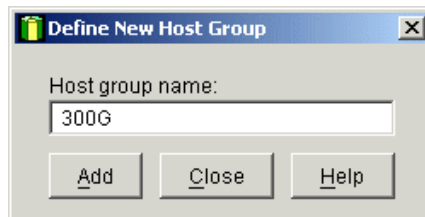


Figure 3-28 Name host group

Highlight the host group created. Now we will define a host for that group by navigating to **Configure -> Topology->Define New Host** (Figure 3-29).

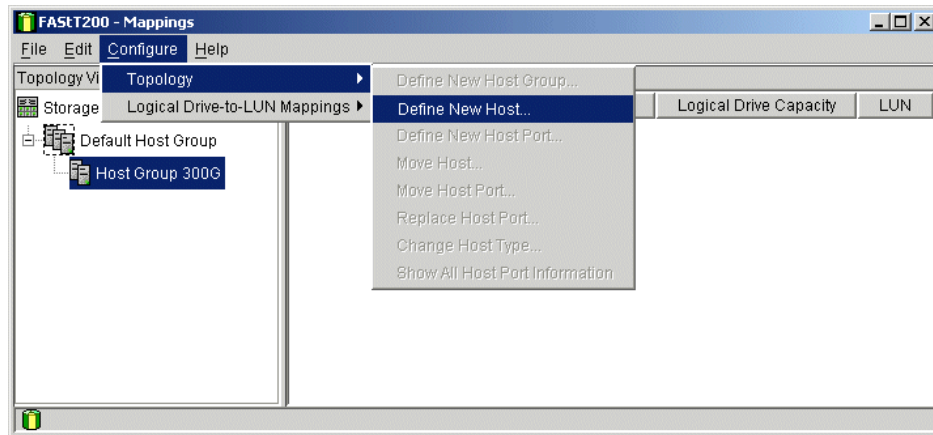


Figure 3-29 Define new host

Enter the name of the host and press **Add** (Figure 3-30).

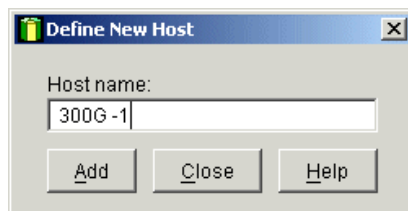


Figure 3-30 Name new host

Now we can create a port for this host. This is accomplished by choosing **Configure -> Topology -> Define New Host Port** (Figure 3-31).

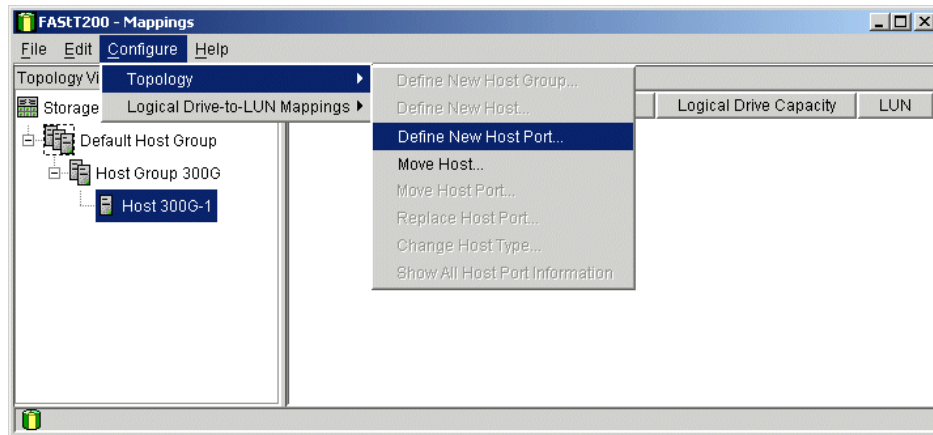


Figure 3-31 Define host port

Select the WWN of the 300G in the Host port identifier box. Select Fibre_Adapter in the Host port name box. The Host type box gets set to the appropriate setting for your 300G. Click **Add** (Figure 3-32).

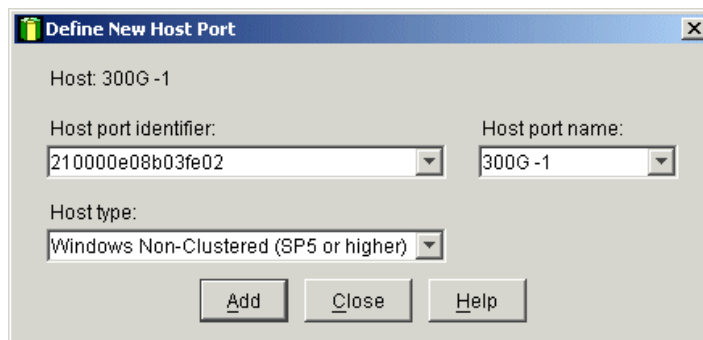


Figure 3-32 Enter WWN

Note: If your FAST200 has not actually “seen” the 300G on its Fibre port, the 300G’s WWN will not be available in the Host port identifier box.

Next, in order to define port mapping, **highlight** the host in the left pane and select **Configure -> Logical Drive to LUN Mappings -> Define New Mapping** (Figure 3-33).

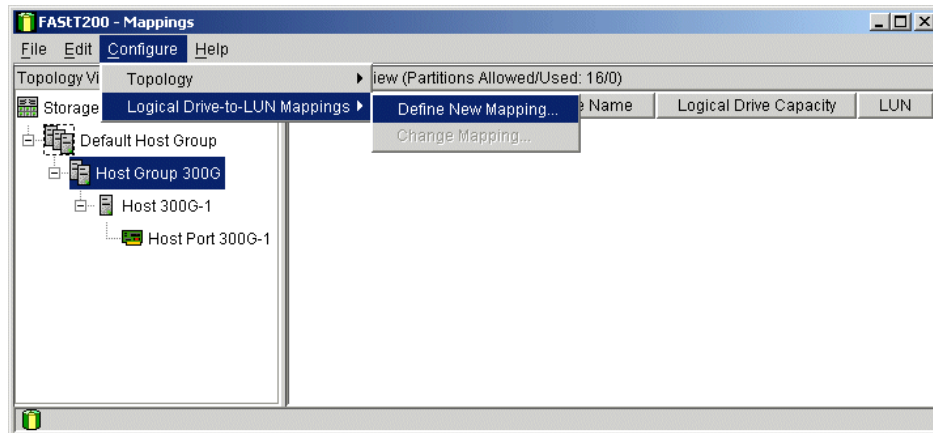


Figure 3-33 Define New Host Group Mapping

In the Define New Mapping screen, **highlight** the drive and select a LUN, then click **Add** (Figure 3-34).

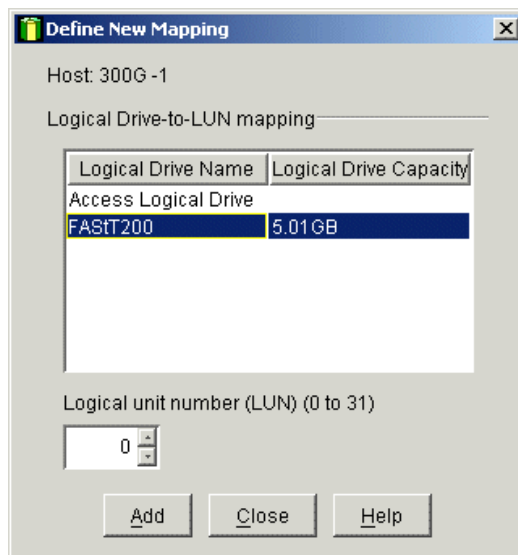


Figure 3-34 Set LUN

Finally, to confirm that the assignment is correct, highlight the Host in the left pane and check the right pane to see the assignment (Figure 3-35).

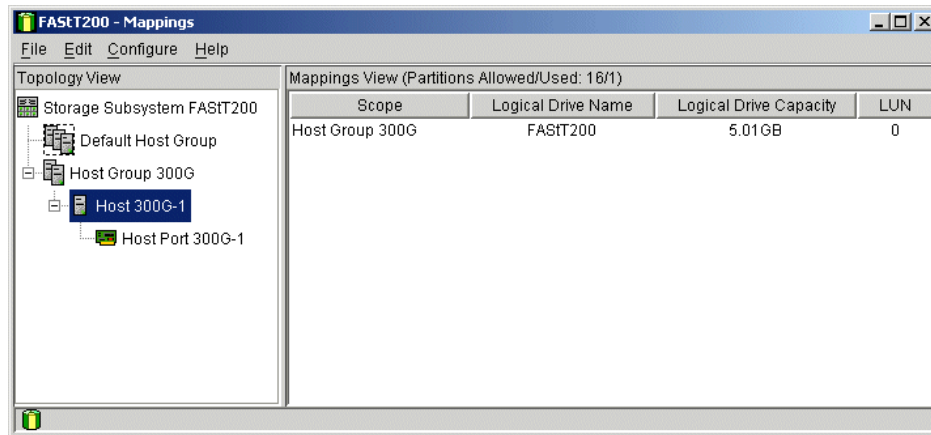


Figure 3-35 Confirm LUN Mapping

Now we have completed all the tasks required on the FAST200 side of things. We have allocated storage for use by the 300G, so all we have to do is claim it, as shown in Section 3.6, “Claiming ownership of pooled storage with the 300G” on page 132. Before we get into that, though, we will cover the procedures we used to allocate storage from the IBM Modular Storage Server (MSS) and the IBM Enterprise Storage Server.

3.4 Setting up the MSS

If you are working in a production environment, there are many planning steps you should go through before proceeding with setup. This section does not cover all of them. If you need comprehensive background information on this subject, before proceeding, we recommend reviewing the redbook, *IBM Modular Storage Server - An Introduction Guide*, SG24-6103.

For the purpose of this redbook, we will be using the Command Line Interface (CLI) by serial connection to set up the MSS for access by the 300G. The CLI is actually quite a simple tool to use, but perhaps not as pretty as some of the other tools available (such as IBM StorWatch MSS Specialist). For information on other management methods and tools, please refer to the redbook, *IBM Modular Storage Server - An Introduction Guide*, SG24-6103.

The CLI engine, which is built into the MSS, is a rich text-based interface that can be accessed by a VT terminal emulator using a serial connection. It can also be accessed by the StorWatch MSS Specialist Graphical User Interface (GUI), but that would not be half the adventure!

Note: The Command Line Interface (CLI) is *not* case sensitive.

More technical information on the CLI can be found in the Storage Works CLI Reference Guide on the Compaq Web site at:

<http://www.compaq.com>

This is a very feature-rich product. Having the CLI Reference Guide available as a reference is critical to conducting a successful installation.

3.4.1 Failover modes and SAN zoning

There are two failover modes available with the MSS: transparent and multiple-bus. This section will provide an overview of these failover modes, explain when each should be used, and describe some of the methods for zoning the SAN fabric to best take advantage of these modes while minimizing impact on other hosts utilizing the disk subsystem.

Failover is a way to keep a storage array available to a host in the event of a controller becoming unresponsive. A controller can become unresponsive due to a hardware failure (such as a failure of a host bus adapter or of the controller) or due to failure of a link between a host and a controller. Failover keeps the storage array available to the host or hosts by allowing the surviving controller to take over total control of the subsystem. There are two failover modes:

- ▶ *Transparent mode* is handled by the surviving controller and is transparent (invisible) to the host.
- ▶ *Multiple-bus mode* is handled by the host or hosts.

Either failover mode can work with either Fibre Channel topology (FC-AL loop or switched fabric). Note that the controllers must be configured for either transparent failover or multiple-bus failover, but not both. Also, we cannot configure the system without selecting one failover mode or the other. Both modes differ in the failover behavior and in the way LUNs are presented to the hosts. Both configurations also strongly differ in external topology, particularly in the way the MSS, hosts, and switches or hubs are interconnected.

3.4.2 Transparent failover mode and zoning

Transparent failover mode has the following characteristics:

- ▶ Hosts do not know that failover has taken place.
- ▶ Units are divided between host Fibre Channel ports.
- ▶ It only compensates for controller failure.

In transparent failover mode, host port 1 of controller A and host port 1 of controller B must be on the same Fibre Channel link to a host. Likewise, host port 2 of controller A and host port 2 of controller B must be on the same Fibre Channel link to a host. This is so because when a failure occurs, the LUNs will be transferred to the standby port of the same number on the alternate controller.

From a SAN fabric point of view, we can have all ports into one switch (or fabric) with no zoning and still maintain a single path from a host to a LUN, as a LUN is only visible over one active port at a time. However, from the MSS's point of view, there will be two active physical connections to the host adapter card (port 1 controller A, port 2 controller B). This is not necessarily disruptive but, as the MSS has a maximum of 64 connections, we immediately limit ourselves to fewer if host are connected in this manner (actual support differs, as listed in Table 3-1 on page 114).

Should we eliminate the second path? As is always with IT, the answer is, it depends! To help us decide, we need a little more background on how unit numbers are allocated within the MSS in transparent failover mode.

In transparent failover mode, a unit number is assigned to either port 1 of both controllers or to port 2 of both controllers. Unit numbers are assigned to ports as follows:

- ▶ Numbers 0 through 99 are assigned to host port 1 of both controllers.
- ▶ Numbers 100 through 199 are assigned to host port 2 of both controllers.

Given these conditions, there are two ways we can set up our SAN fabric depending on how we allocate unit numbers (we will discuss how to allocate unit numbers in the next section) and both have administrative implications.

For example, if we allocate unit numbers 1-10 to a given host, then all of the associated LUNS will be accessed over the active port 1 (controller A). However, if we allocate 100-110, then all associated LUNs will be accessed over the active port 2 (controller B). If unit numbers are assigned in this way, then it is recommended that we use zoning, as shown in Figure 3-36.

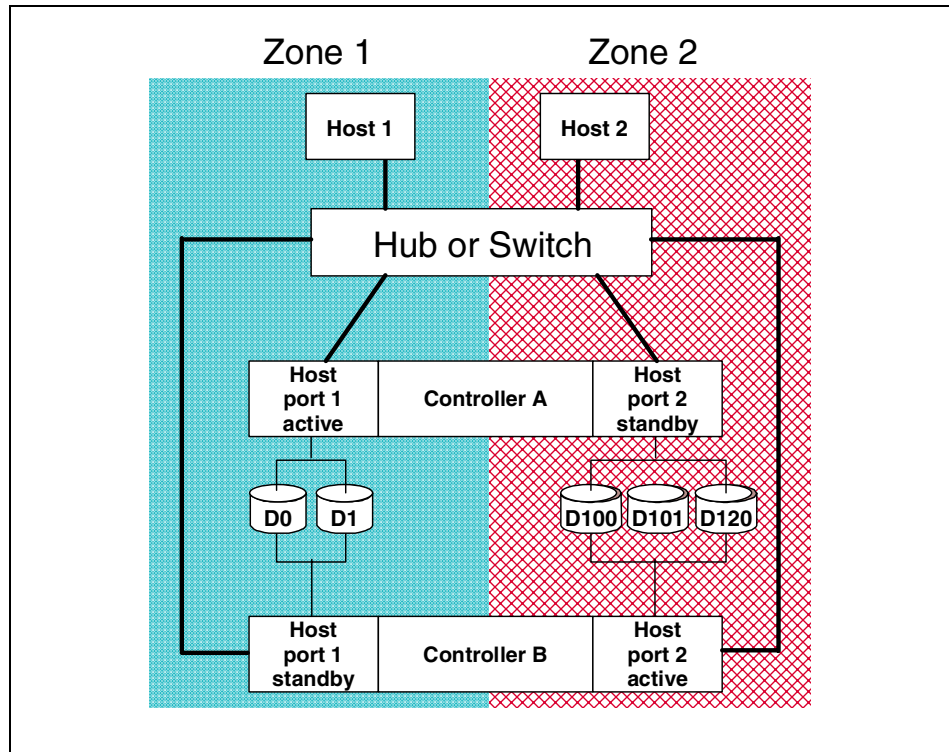


Figure 3-36 Zoning in transparent failover mode

Note: the above diagram shows all port connections into a single fabric device, divided into two distinct zones. This could just as easily be done with two fabric devices with no zoning (provided they are not interlinked).

The advantage here is that we have only used up one connection to the MSS for each host adapter. For example, in the above diagram, host 1 would only be visible to the MSS over controller A, port 1. Host 2 would only be visible to the MSS over controller B, port 2. Two servers, two connections! The disadvantage is that all I/O traffic from each host will be driven through a single controller on the MSS.

An alternative is to allocate unit numbers across controllers/ports. For example, if we allocate unit numbers 1-5 and 100-105 to a given host, then LUNs associated with unit numbers 1-5 will be accessed over the active port 1 while LUNs associated with unit numbers 100-105 will be accessed over the active port 2. If unit numbers are assigned in this way, then we *cannot* zone our fabric as shown in Figure 3-36, because we would lose connections to half of our LUNs (the ones in the other zone).

The advantage of this configuration is that we have, in effect, done some manual load balancing across MSS controllers from our host to its LUNs. The disadvantage is that we have used up two connections for every host we connect in this way (remember, we only have 64).

Note: You need not be concerned about the large unit numbers shown above. With connection offsets (which we discuss in the next section), these large unit numbers can still be made visible within an operating system's limitations.

3.4.3 Multiple-bus failover mode and zoning

Multiple-bus failover mode has the following characteristics:

- ▶ The host controls the failover process by moving LUNs from one controller to the other.
- ▶ All units (0 through 199) are visible on all host ports, but accessible only through one controller at any specific time.
- ▶ Each host has two or more paths to the logical units.
- ▶ All host ports are active and have their own port identity (port number).
- ▶ It compensates for controller, fabric, cable and adapters failures.

Multiple-bus failover must be used only in conjunction with Subsystem Device Driver (SDD), *formerly known as Data Path Optimizer (DPO)* or operating systems, such as Tru64 V5.x, OpenVMS V7.2-1 and higher, which provide their own dual pathing functionality.

When one path fails, a host can issue commands to move the units from one path to another.

A typical multiple-bus failover configuration is shown in Figure 3-37. The diagram shows a typical configuration when we have hosts capable of multi-pathing. As we can see, there is no zoning required, as the hosts are handling the failover.

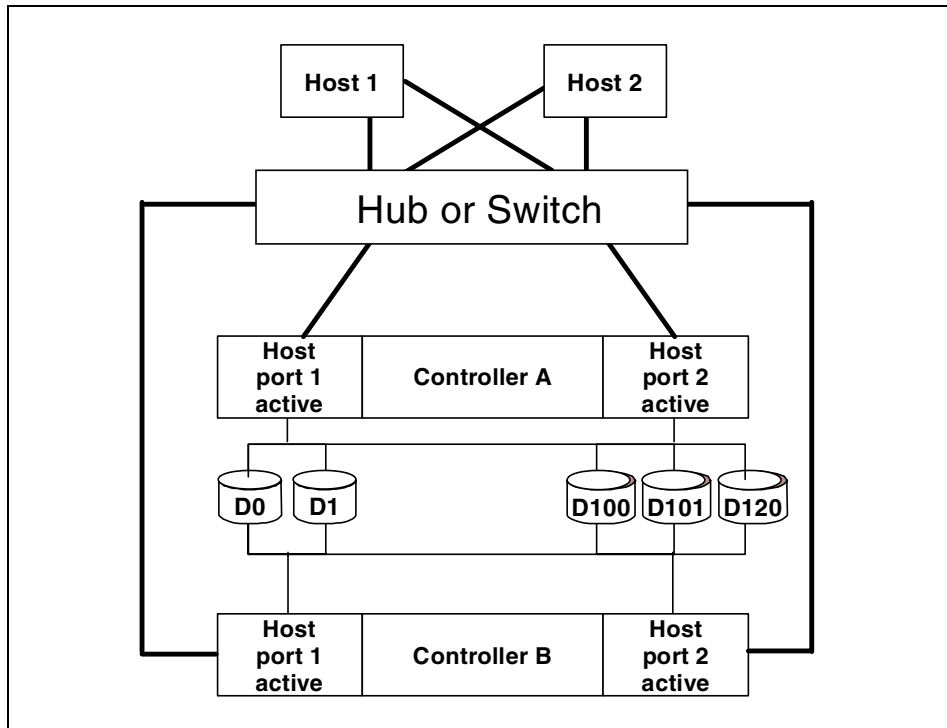


Figure 3-37 Typical SAN Configuration in multiple-bus failover mode

Not all host operating systems are capable of multi-pathing, so what we generally see is a mix of hosts — some capable of multi-pathing and some not. As we know, we can only set one failover mode for the MSS, so how do we decide which one to use? Must we use the lowest common denominator (that is, transparent failover) or can we allow our multi-path capable hosts to take advantage of the multiple-bus failover mode without eliminating access for those hosts not capable of multi-pathing?

The answer lies with zoning and the preferred path setup, as can be seen in Figure 3-38. Here we see that although we are in multiple-bus failover mode, our single path capable hosts still only have one path to a LUN because a LUN is only active on one controller at a time. As we have zoned one port from each controller, our single path hosts will be able to see the LUN no matter which controller is the preferred one. The only requirement is that the preferred path does not change, as this would result in the LUN identifier changing and our host losing connection.

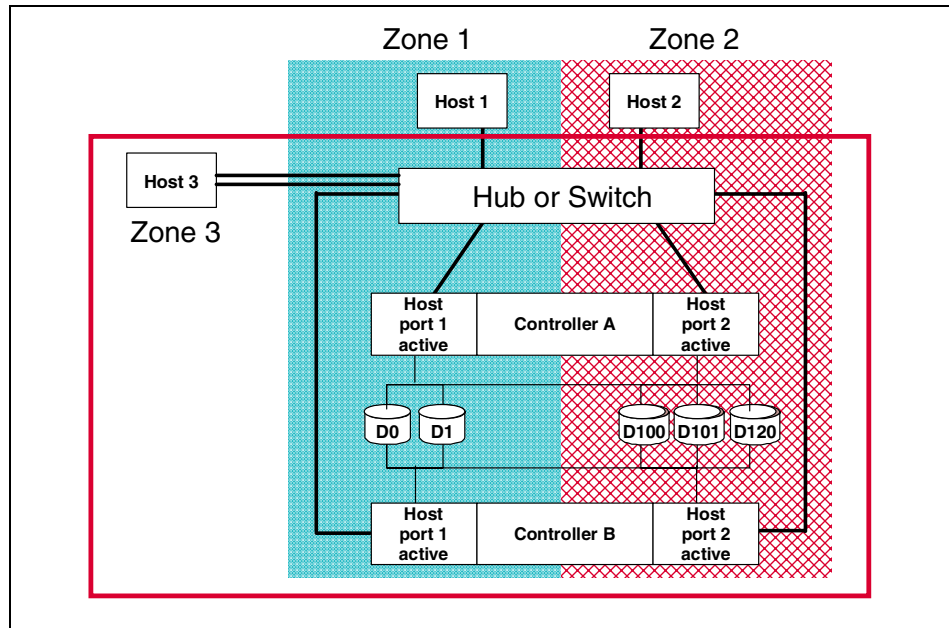


Figure 3-38 Zones in multiple-bus failover mode

3.4.4 Preferred controller in multiple-bus failover mode

When an MSS is initialized in multiple-bus failover mode, all defined LUNs are allocated to a particular controller based on which controller sees them first during the boot up process. If all hosts support multi-pathing, then there is no need to make any changes to this, as the hosts will find the LUN no matter which controller owns it. If we wish to have a little more control over which controller owns a particular LUN, then there is a parameter we can enter, as follows:

```
set unit_number preferred_path=controller (this or other)
```

Note: We will explain unit numbers in the next section.

In addition, if we have hosts which do not support multi-pathing, we *must* set a preferred path for each LUN. For example, in the diagram just shown in Figure 3-38, hosts 1 and 2 do not support multi-pathing. If we do not set a preferred path, a particular LUN could be owned by a different controller each time the MSS is re-initialized (granted we should not be doing this very often).

As the ports in the MSS do not share identities in multiple-bus failover mode (they do in transparent failover mode), the LUN identifiers would change if preferred ownership shifted from one controller to the other. If this were to occur, hosts 1 and 2 will lose connectivity, as they have no method of resolving this change of identifiers.

If we assign a preferred path, our LUN will not move after each initialization so our single path hosts are not exposed by this. In the event of a controller failure our LUNs will indeed move to the alternate controller, which will cause our single path hosts to lose connectivity to their LUNs. Since this is the case, we might as well take our zoning one step further, as shown in Figure 3-39.

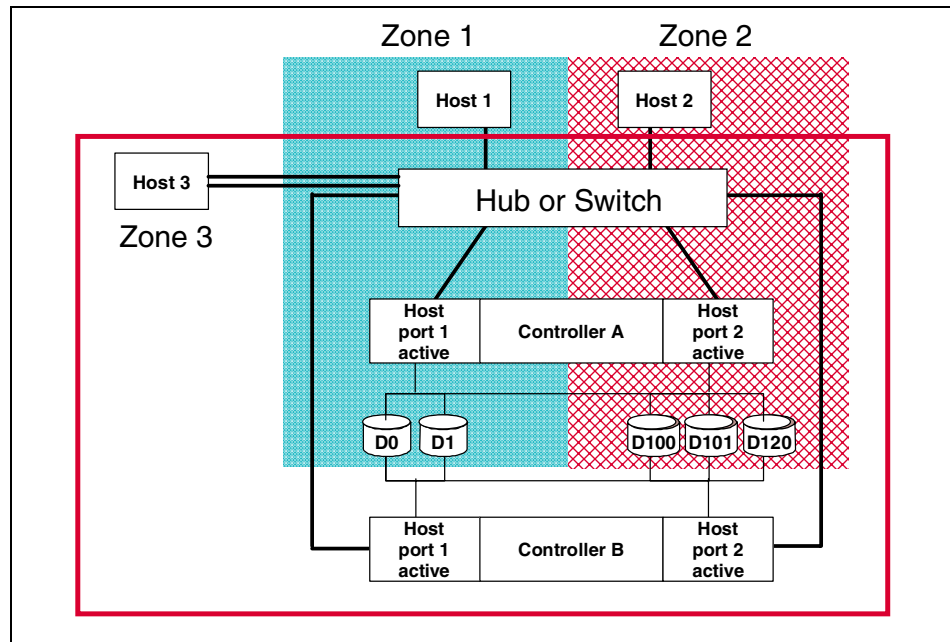


Figure 3-39 Zones and preferred pathing in multiple-bus failover mode

In Figure 3-39, the LUNs associated with hosts 1 and 2 should be preferred by controller A, while LUNs associated with hosts 3 would have no preferred path set (it is multi-pathing capable). If controller A fails, hosts 1 and 2 will still lose connectivity to their LUNs, but host 3 will continue to function. So what was the benefit of changing the zones? Well, in the zone setup seen in Figure 3-38 on page 104, the MSS would see 8 connections, 4 from the multi-path host and 2 each from our single path hosts (even though one of the connection from each of our single path hosts could never be used). In Figure 3-39, the MSS would see only 6 connections, 4 from our multi-path host and 1 each from our single path hosts. With a total of 64 available in the MSS, this may be a consideration.

In the final analysis, when using multiple-bus failover mode combined with the zoning shown, our single-path hosts have less failover protection than they would get in transparent failover but our multi-path capable hosts have all the benefits of multiple-bus failover.

The question is, which situation do we prefer? In transparent mode, all hosts have some sort of failover protection (against controller failure at least). In multiple-bus mode with zoning, our single-path hosts have no failover protection but our multi-path hosts have extra protection (against controller, fabric, cable and adapters). The answer is, once again — it depends. Which hosts are critical? If they are equally important, go for transparent failover. If only our multi-path capable hosts are critical, then go for multiple-bus failover. Of course, the ideal solution would be to have two MSS boxes, one for each failover mode and its respective hosts. The alternative is to wait patiently for multi-path support for the operating system of choice.

3.4.5 Setting up failover modes

If working with anything but factory-new controllers, enter the following command to remove any failover mode that may have been previously configured:

```
set nofailover
```

Enter the following command to stop the CLI from reporting a misconfiguration error that results from having no failover mode specified:

```
clear cli
```

Put the controller pair into transparent failover mode or multiple-bus failover mode, using one of the following commands:

```
set failover copy=this (transparent failover mode)
```

```
set multibus copy=this (multiple bus failover mode)
```

The switch, **this**, simply tells the controller which controller has the valid failover configuration and restarts the **other** controller to ensure they are both synchronized.

For more information on failover modes and their use, please see the redbook, *IBM Modular Storage Server - An Introduction Guide*, SG24-6103.

3.4.6 Create a Logical Unit Number (LUN)

The first step in creating a LUN is to create an array of disks (called a storageset in MSS terminology). There are a number of different types of storagesets available to us with the MSS. They are: disk drive (JBOD); stripesets (RAID-0); mirrorsets (RAID-1); RAIDsets (RAID-5); striped mirrorsets (RAID-0+1); and sparesets (hotspare disks). We will configure a RAIDset. For information on the commands used to create other storageset types, please see the *IBM Modular Storage Server - An Introduction Guide*, SG24-6103 or the CLI Reference Guide from the Compaq Web site at:

<http://www.compaq.com>

If you already have a defined storageset we wish to use, then skip this step.

3.4.7 Create a RAIDset

To create a RAIDset, we must first know what disk we have available to put in that RAIDset. We do this by typing the following command:

```
show disks
```

The output should resemble that shown in Example 3-1.

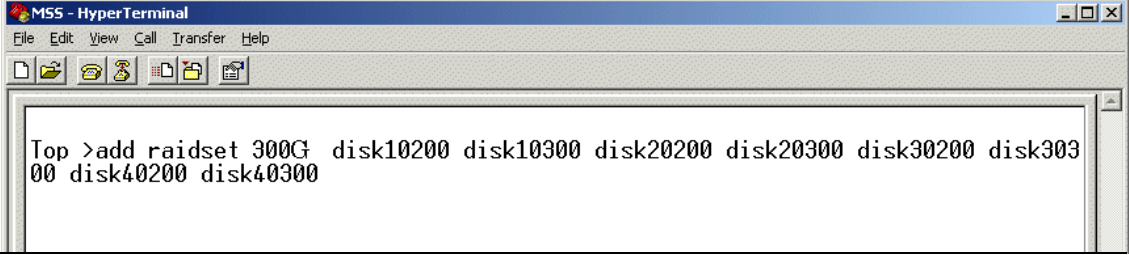
Example 3-1 show disks command output

Name	Type	Port	Targ	Lun	Used by
DISK10000	disk	1	0	0	
DISK10100	disk	1	1	0	
DISK10200	disk	1	2	0	
DISK10300	disk	1	3	0	
DISK20000	disk	2	0	0	
DISK20100	disk	2	1	0	
DISK20200	disk	2	2	0	
DISK20300	disk	2	3	0	
DISK30000	disk	3	0	0	
DISK30100	disk	3	1	0	
DISK30200	disk	3	2	0	
DISK30300	disk	3	3	0	
DISK40000	disk	4	0	0	
DISK40100	disk	4	1	0	
DISK40200	disk	4	2	0	
DISK40300	disk	4	3	0	

Once we know what disks are available, we are ready to create a RAIDset by typing in the following command:

```
add raidset raidset_name diskn
```

In our example we created a RAIDset called 300G consisting of two disks from each SCSI channel in use (total 8 disks), as shown in Figure 3-40.

A screenshot of an MSS HyperTerminal window. The window title is "MSS - HyperTerminal". The menu bar includes "File", "Edit", "View", "Call", "Transfer", and "Help". Below the menu bar is a toolbar with icons for file operations. The main text area shows the command: "Top >add raidset 300G disk10200 disk10300 disk20200 disk20300 disk30200 disk30300 disk40200 disk40300".

```
Top >add raidset 300G disk10200 disk10300 disk20200 disk20300 disk30200 disk30300
disk40200 disk40300
```

Figure 3-40 The add raidset command

We get a warning from this configuration, as we have chosen more than one disk from each SCSI channel. If we were to lose a disk enclosure in this configuration, we would lose access to all LUNs defined on the RAIDset, because we would lose more than one disk from the array. This obviously something we would need to consider for a production environment but for now, we will just keep going.

Note: Leave at least one disk for a spareset.

Now we need to initialize the RAIDset. At this point we could specify a block size for the RAIDset (chunksize in MSS terminology). The defaults (256 blocks for RAIDsets of less than or equal to 9 physical disks and 128 blocks for RAIDsets of greater than 9 physical disks) are normally adequate, but if we know the workload attributes of the particular application which will use this logical drives within the RAIDset, we can set the chunksize here.

Note: The chunksize must be specified in blocks, not Kb.

The bad news is that we cannot change the chunksize without doing a backup, destroying and recreating our RAIDset with a new chunksize, and restoring our data.

3.4.8 Initialize a RAIDset

To initialize a RAIDset, type in the following command and wait for it to complete (it may take a few minutes).

```
initialize raidset_name chunksize=blocks (optional)
```

In our example, we did not bother setting a chunksize (it will default to 256 blocks) so we simply substituted **300G** for the `raidset_name`, as shown in Figure 3-41.

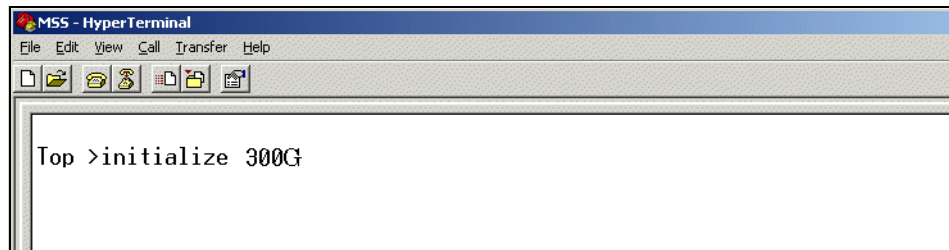


Figure 3-41 The initialize raidset command

3.4.9 Create a partition

The next step is to create a partition within our RAIDset, as follows:

```
create_partition raidset_name size=%
```

That is right, we need to enter a percentage of total capacity not MB or GB. This presents a bit of a mathematical challenge and it is important to know that it works on usable capacities, not raw capacity. As we used eight 18.2 GB disks in our RAIDset (RAID-5), we have seven disks (127 GB) of usable capacity available (losing one disk to parity). Now to work out the percentage by dividing the size of the partition we want by the total usable capacity. For example, if we wanted a 50 GB partition, we'd need to divide 50 GB by 127 GB to get a percentage equivalent of 39%.

We found this too difficult, so we simply chose a partition size of 50% within the 300G RAIDset. This gives us approximately 63 GB in our partition (some capacity is always lost to metadata, so this is not an exact science). See Figure 3-42.

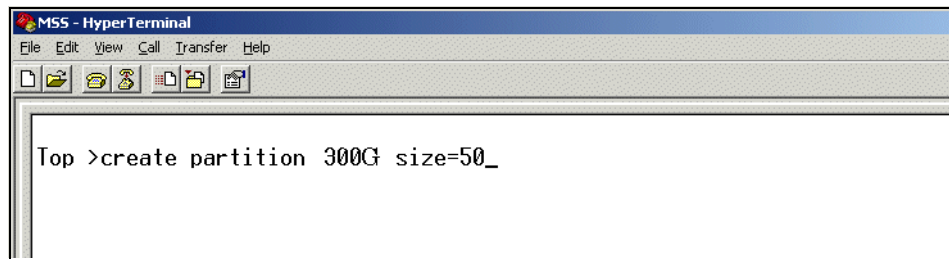


Figure 3-42 The create partition command

Subsequent partitions use a percentage of the total usable RAIDset size (not percentage of remaining) so if we create another 50% partition, we would use all the remaining space in our RAIDset (plus or minus a few bytes).

Note: If we wish to use the entire RAIDset (or remaining space within the RAIDset) for our partition, we could enter **size=largest** instead of a percentage.

3.4.10 Assign a unit number

Finally, we need to assign a unique unit number (00-199) to each partition we created in order for it to be allocated to a host and subsequently made available for connection by that host.

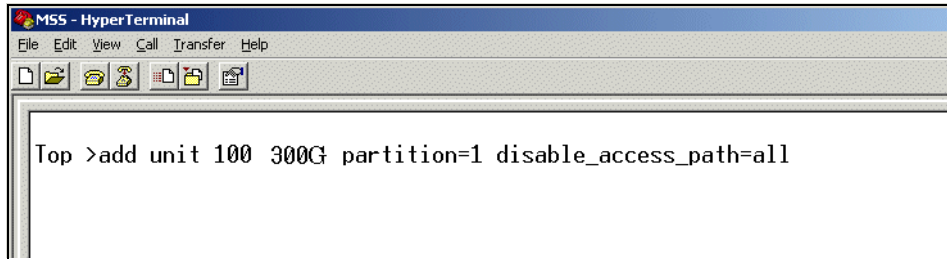
To assign a unit number to our partition, enter the following command:

```
add unit unit_number raidset_name partition=partition_number  
disable_access_path=all (optional)
```

The `disable_access_path` switch is optional, but if we do not specify it, the new unit number will be visible to all hosts, which *could* cause problems. Specifying that we wish to disable access means we can assign to a specific host at a later date. Alternatively, we could enable access to a specific host by typing:

```
add unit unit_number raidset_name partition=partition_number  
enable_access_path=host_name
```


In our example, we assigned unit number 100 to partition 1 (the only one we created) of the **300G** RAIDset and disabled access to the unit number (we will assign later), as shown in Figure 3-43. The unit number will automatically be prefixed with a D (which stands for disk drive).



```
Top >add unit 100 300G partition=1 disable_access_path=all
```

Figure 3-43 The add unit command

Note: The unit number is not the same as a LUN.

A LUN is determined by the unit number minus any offset specified for the host connection. If no offset is specified, then the default offset for each MSS port will apply. The default offset depends on the failover mode being used.

As we learned earlier, in transparent failover mode, a unit number is assigned to either port 1 of both controllers or to port 2 of both controllers (port 1 on controller A and port 2 of controller B being active, the other ports are in standby mode). Unit numbers are assigned to ports, as follows:

- ▶ Numbers 0 through 99 are assigned to host port 1 of both controllers
- ▶ Numbers 100 through 199 are assigned to host port 2 of both controllers

The default offset for port 1 is 0 and the default for port 2 is 100. So, for example, a unit number of 10 would only be visible over port 1 of each controller. With an offset of 0, the LUN would be presented as 10. A unit number of 100 (as in our example) is only visible over port 2 on each controller. With an offset of 100, the LUN would be 0.

In multiple-bus failover mode, a unit number is not assigned to a specific pair of ports, rather it is assigned to a preferred controller (all ports on both controllers being active). The default offset for all four ports is 0. So, for example, a unit number of 10 is visible over all ports. With an offset of 0, the LUN would also be 10. A unit number of 100 (as in our example) is visible over all ports. With an offset of 0, the LUN would also be 100.

So, offsets can be used as a method of LUN masking. By specifying a unit number which, when coupled with a particular host's offset (default or otherwise), puts a LUN outside the address range of that host, we can effectively mask that LUN from that host. For example, a unit number 100 (as in our example) would not be visible to a host with an offset of 110, as this would equate to a LUN of -10 to that host (negative LUNs are typically not recognized by hosts).

Offsets can also be used as a method of making sure unit numbers are accessible by hosts which have LUN addressing limitations. For example, an MSS has three host connections, each capable of seeing 8 LUNs per target. The operating system on all three hosts designates that it can only see LUNs as 0 through 7. With the default offsets in transparent failover mode, only two of these hosts would have access to LUNs on the MSS (unit 0-7 with offset of 0 and unit 100-107 with offset 100). So, in order to allocate LUNs to the second host, we define offsets which make the larger unit numbers visible as LUNs 0-7. For example, we assign unit numbers 20-27 to eight partitions within a RAIDset. Then we change our third host connection offset to 20, as follows:

```
set host_name offset=20
```

This means the third host will see LUNs 0-7, as it expects, despite the higher unit numbers assigned in the MSS.

For more details on offsets or failover modes, please refer to the redbook, *IBM Modular Storage Server - An Introduction Guide*, SG24-6103.

3.4.11 Define host(s) and assign logical drive(s)

Simply plugging our host into the same SAN fabric zone, as the MSS should be enough to establish a connection. When a new host is added to a SAN fabric, a Loop Initialization or Fabric Login takes place, this process causes the fabric device to broadcast an update to its configuration, which is how the MSS detects a new host connection.

Note: If by chance the MSS did not auto-discover the host, it is likely that there is a problem with the connectivity. Check all physical connections, media, and SAN zones to make sure that a physical connection exists and the host is in the same zone as the MSS. Once the problem is fixed, the connection should become visible to the MSS.

To see if the MSS has found our host, enter the following command:

show connections

The output should be all existing defined host connections and any new hosts which have been found, as shown in Example 3-2.

Example 3-2 show connections command output

Connection Name	Operating system	Controller	Port	Address	Status	Unit Offset
!NEWCON48	WINNT	OTHER	2	031C00	0L	other 100
		HOST_ID=2000-00E0-8B03-FE02		ADAPTER_ID=2100-00E0-8B03-FE02		
BRAZIL	IBM	OTHER	2		0L	other 100
		HOST_ID=1000-0000-C922-BF75		ADAPTER_ID=1000-0000-C922-BF75		

Note: The controller tags, THIS and OTHER, will remain relative to the controller we are serially connected to for management. In our case, we connected to controller A, so the controller tag OTHER applies to controller B.

If we did not zone our SAN fabric, as recommended in the section “Failover modes and SAN zoning” on page 99, then we will see multiple connections for each host adapter, as follows:

- ▶ In transparent failover mode, each host adapter port will be seen twice (once over port 1 controller A and once over port 2 of controller B).
- ▶ In multiple-bus failover mode, each host adapter port will be seen four times (once over each port in both controllers).

In either case, do not panic, this is simply a table of the physical connections to the MSS, not a table of connection to the partitions (LUNs) within it. Connections to the actual partitions is dependent on both SAN fabric zoning *and* LUN masking, as detailed in the section entitled, “Failover modes and SAN zoning” on page 99.

It should be noted however, that this table of connections is maintained in the memory of the MSS, and it has a maximum of 64 entries. Once reached, the only way to add more connections is to delete others. What does this mean? Well, if we have not done any SAN zoning, in transparent failover mode there are 2 physical connections for every host connected, which means we limit ourselves to 32 hosts. In multiple-bus failover, there are 4 physical connections for every host adapter, which means a maximum of 16 hosts.

It is therefore highly recommended that we zone our SAN fabric, as detailed in the section entitled, "Failover modes and SAN zoning" on page 99 to avoid this issue.

Important: Actual supported host numbers are dependent on operating system type, whether we are operating the MSS in a homogenous or heterogeneous host environment and Compaq/IBM testing. Please see Table 3-1 for information pertaining to host numbers supported at the time of producing this redbook.

Zoning is also recommended because, without it, the possibility of errors when assigning LUNs to a connection increases. With so many connections to choose from, it is very easy to assign a LUN to a connection which is either not on the preferred controller and/or port. Doing so would most likely result in the LUN being invisible to the desired host (this is not destructive, just confusing for an administrator).

Table 3-1 Maximum supported hosts for the MSS

Operating system	Maximum supported hosts
Tru64 UNIX	8
OpenVMS	32
Windows NT	16
Windows 2000	16
SUN Solaris	8
HP-UX	8
Novell Netware	16
Heterogeneous Access	8

3.4.12 Assigning a connection name

As we saw in Example 3-2 on page 113, the default names assigned to new connections are somewhat meaningless. To change the label from !newconxx to something more useful, enter the following command:

```
rename !newconxx connection_name
```

The command we entered is shown in Figure 3-44.

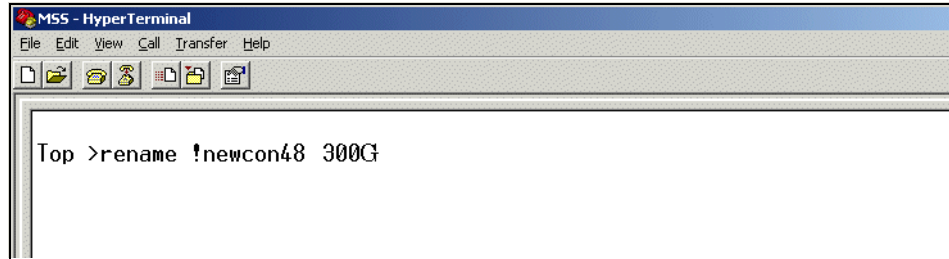


Figure 3-44 The rename connection command

Now, to assign an operating system type to the connection, enter the following:

```
set connection_name operating_system=os_name
```

The command we entered is shown in Figure 3-45.

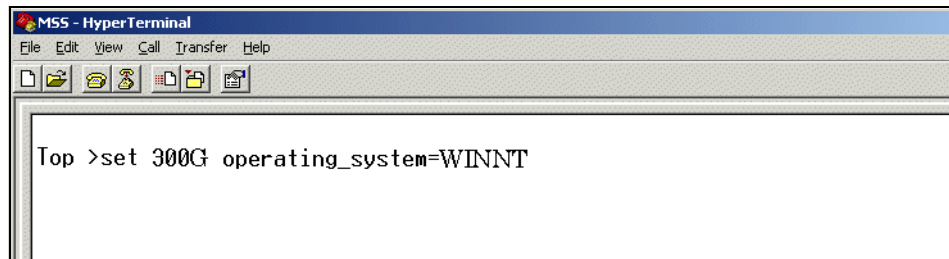


Figure 3-45 The set operating system command

3.4.13 Assigning a LUN to a host

Now that we have a defined host connection, we are able to assign a LUN to one or more of those host connections by using the command:

```
set unit_number enable_access_path=host_name
```

The command we entered is shown in Figure 3-46.

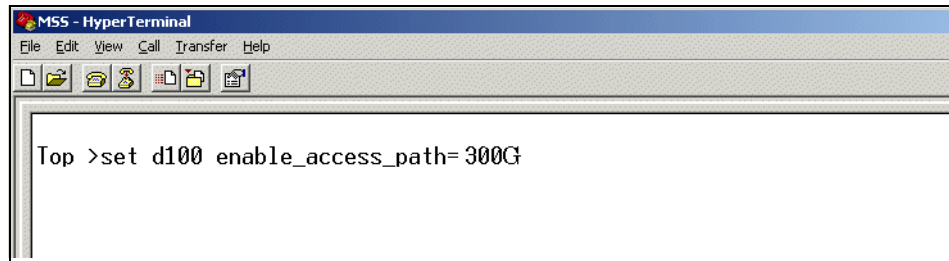


Figure 3-46 The enable access command

We have now completed all the tasks required on the MSS side of things. We now have storage inside the MSS allocated for use by the 300G. All we have to do is claim it. We show how to do this in Section 3.6, “Claiming ownership of pooled storage with the 300G” on page 132. First, however, we will round out our storage environment by detailing the steps we took to allocate even more storage for our use from an IBM Enterprise Storage Server.

3.5 Setting up the ESS

If you are working in a production environment, there are many planning steps you should go through before proceeding with setup. This section does not cover all of them. If you need comprehensive background information on this subject, before proceeding, we recommend that you review either of the redbooks, *The IBM Enterprise Storage Server, SG24-5645*, or *ESS Solutions for Open Systems Storage: Compaq AlphaServer, HP, and SUN, SG24-6119*.

3.5.1 Regarding SAN zoning

With the ESS, we do not have to worry about setting failover modes like we do with the FAStT and MSS. Multiple failover modes are not required in the ESS, as the RAID controllers are separated from the connected hosts by a layer of management. The RAID controllers in the ESS are within the RS/6000 nodes inside the enclosure and, as these two nodes are in a clustered configuration (HACMP), failure of a controller or entire node is handled internally.

Because failover is handled internally to the cluster, hosts connected to the ESS do not need to be aware of the internal workings of the disk subsystem — this is all provided by the cluster. If a failure occurs, although many changes take place within the cluster, the connected hosts see no differences, as their LUNs maintain their device and path identifiers on the alternate node — this is one of the features of HACMP.

Note: This description does not account for adapter or link failure in the connected hosts.

Figure 3-47 shows a logical view of the internals of the ESS. Connectivity to the disk is through the “intelligent” RS/6000 controllers (Nodes A and B).

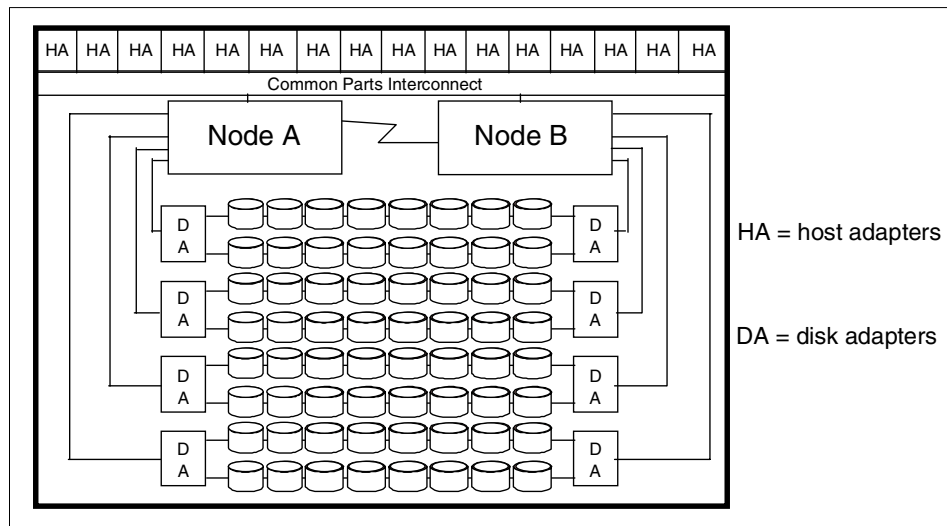


Figure 3-47 ESS internals

For more information on zoning and the switch products available from IBM, please see the redbook, the *IBM SAN Survival Guide*, SG24-6143.

3.5.2 Setting up the ESS

In this section, we document the steps involved in setting up the ESS for access by the 300G. A high-level view of the process is shown in Figure 3-48.

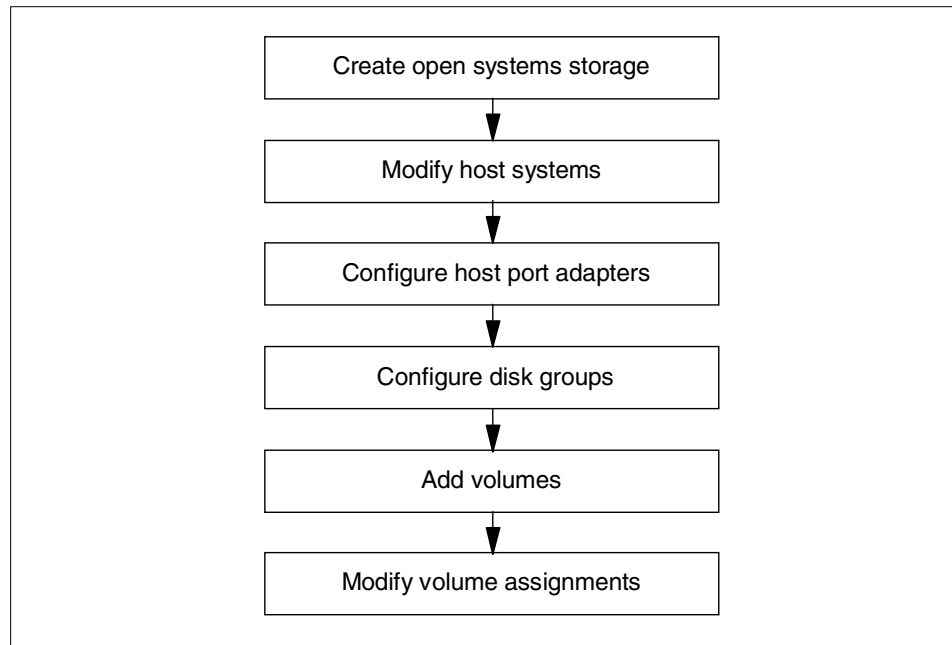


Figure 3-48 ESS preparation for the 300G host

For the purpose of configuring the ESS, we will be using the IBM StorWatch ESS Specialist. It is a simple yet powerful administration tool that comes with the ESS at no additional cost, and it is the only supported way in which to manage the allocation of storage within the ESS storage subsystem.

More information on the IBM StorWatch ESS Specialist can be found in the redbook, *Implementing the Enterprise Storage Server in Your Environment*, SG24-5420.

Create open systems storage

To configure our ESS, we must first launch the IBM StorWatch ESS Specialist by entering the hostname or IP address for either of the ESS clusters in the Location or URL window of the browser. Doing so will bring us to the Specialist home page, as shown in Figure 3-49.

StorWatch Solutions

Enterprise Storage Server Specialist

Introduction

Status

Problem Notification

Communications

Storage Allocation

Users

Licensed Internal Code

Welcome to
IBM StorWatch
Enterprise Storage
Server Specialist

Machine Type: 2105
Machine Model: F20
Serial Number: 075-18540
WWNN: 5005076300C08604

View README

© Copyright IBM Corp. 1998, 2000. Licensed Materials. Property of IBM. All Rights Reserved. IBM is a registered trademark of the IBM Corp. U.S. Government Users Restricted Rights - Use, duplication, or disclosure restricted by GSA ADP Schedule Contract with IBM Corp.

Figure 3-49 StorWatch ESS Specialist home page

Selecting the **Storage Allocation** button on the left of the home page will present the administrator with a number of security certificates and a login pop-up window. After successfully entering a valid login and password, the administrator will see the **Storage Allocation - Graphical View** window (as shown in Figure 3-50).

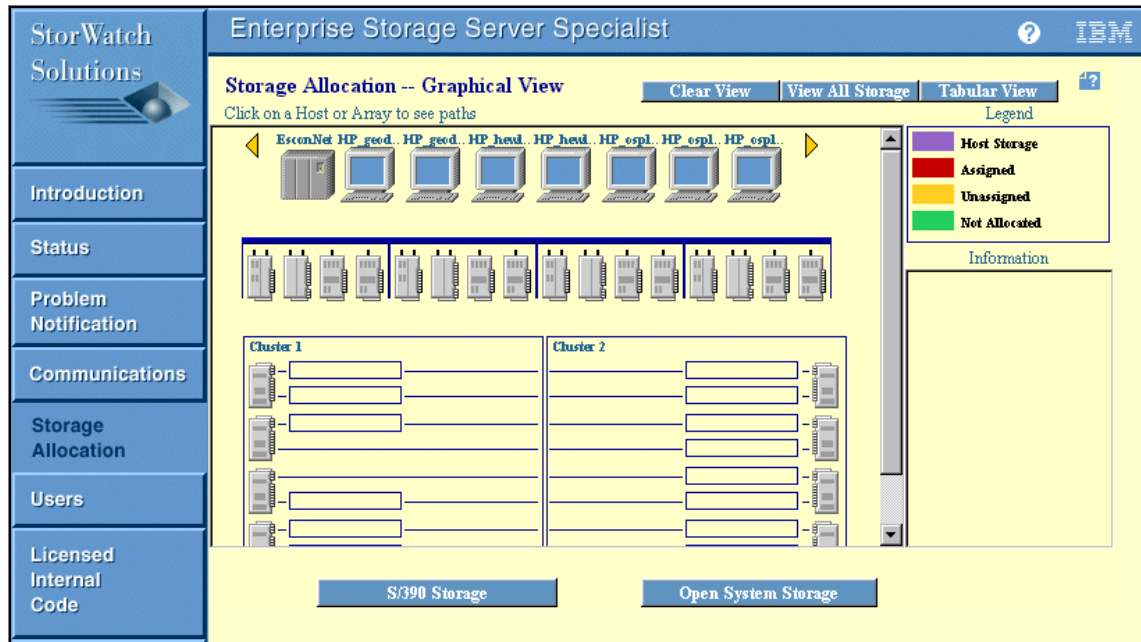


Figure 3-50 Storage Allocation panel

The Storage Allocation panel is the most important and detailed of the ESS Specialist panels. It provides an authorized user with a graphical representation of the hosts, host adapters (ESCON, SCSI, Fibre Channel and — soon — FICON), Device adapters (SSA Adapters), Clusters, and Arrays on the ESS machine.

From the Storage Allocation panel (seen in Figure 3-50), we can choose to create either **S/390 Storage** or **Open System Storage**. As we are creating a storage partition for the 300G, we should click the **Open System Storage** button in the bottom right of the panel.

Clicking the **Open System Storage** button will produce a screen similar to the one depicted in Figure 3-51. From this view, it is possible to add and remove hosts, configure HAs, setup disk groups, add and remove volumes, and modify how existing volumes are assigned.

The basic setup procedure is to click the buttons on the bottom of the screen in the order shown in the following pages.

Enterprise Storage Server Specialist

Open System Storage

Host Systems

Nidname	Host Type	Attachment	WWPN	Hostname/IP Address
HP_geode_5158_0_2_0_0	HP 9000 Series 800	FC	50060B0000068B84	
HP_geode_5158_0_7_0_0	HP 9000 Series 800	FC	50060B0000068172	
HP_hewlett_6685_8_12_1_0	HP 9000 Series 800	FC	50060B00000902AE	
HP_hewlett_6685_8_8_1_0	HP 9000 Series 800	FC	50060B00000902A8	

Assigned Volumes (Total: 0 volumes)

Volume	Vol Type	Size	Storage Type	Location	LSS	Shared
Select one host in the Host Systems table, to view its currently assigned volumes						

Modify Host Systems Configure Host Adapter Ports Configure Disk Groups
 Add Volumes Modify Volume Assignments

Figure 3-51 Open Systems Storage panel

Modify Host Systems

To add, remove, or modify an existing host, select the **Modify Host Systems** button from the bottom left of the Open Systems Storage panel. The display in Figure 3-52 will appear.

The screenshot shows the 'Modify Host Systems' panel in the Enterprise Storage Server Specialist interface. The panel is divided into two main sections: 'Host Attributes' on the left and 'Host Systems List' on the right. The 'Host Attributes' section includes the following fields:

- Nickname:** ITSO_300G
- Host Type:** PC Server (Win NT 4.0 or higher)
- Host Attachment:** Fibre Channel attached
- Hostname/IP Address:** (empty field)
- World-Wide Port-Name:** 210000E08E03FED2
- Fibre-Channel Ports:** All installed ports

The 'Host Systems List' section shows a table of existing hosts:

Nickname	Host Type
HP_geode_5158_0_2_0_0	HP 9000 Series 800
HP_geode_5158_0_7_0_0	HP 9000 Series 800
HP_hewlett_6685_8_12_1_0	HP 9000 Series 800
HP_hewlett_6685_8_8_1_0	HP 9000 Series 800
HP_osp1hp1_5158_0_3_0_0	HP 9000 Series 800
HP_osp1hp1_5158_0_4_0_0	HP 9000 Series 800
HP_osp1sm2_6684_8_4_1_0	HP 9000 Series 800
HP_osp1sm2_6684_8_8_1_0	HP 9000 Series 800
ITSO_linux_f	PC Server (Win NT)
ITSO_linux_scsi	PC Server (Win NT)
Linux_SCSI_x230	PC Server (Win NT)
MOUNTAINDEW	PC Server (Win NT)
Netfy-hba1	PC Server (Win NT)
Netfy-hba2	PC Server (Win NT)
Netfy2-hba1	PC Server (Win NT)
Netfy2-hba2	PC Server (Win NT)
pc17	PC Server (Win NT)

At the bottom of the panel, there are two buttons: 'Perform Configuration Update' and 'Cancel Configuration Update'.

Figure 3-52 Modify Host Systems panel

To add a new host, fill in the fields within the Host Attributes table on the left. The Host Type configured is extremely important, as the ESS determines drive geometry, labels, targets, and LUNs available, and so on, based on the Host Type field. We are using the **PC Server (Win NT 4.0 or higher)** host type.

The Hostname/IP Address field is optional in all cases. However, when configuring a Fibre Channel host, we must enter the WWN of the host adapter. Be careful, as mistakes in typing this number will lead to hard-to-trace connectivity failure.

Once you are satisfied with the settings, click the **Add** button and the newly created host will be displayed within the Host Systems List table on the right, as shown in Figure 3-53.

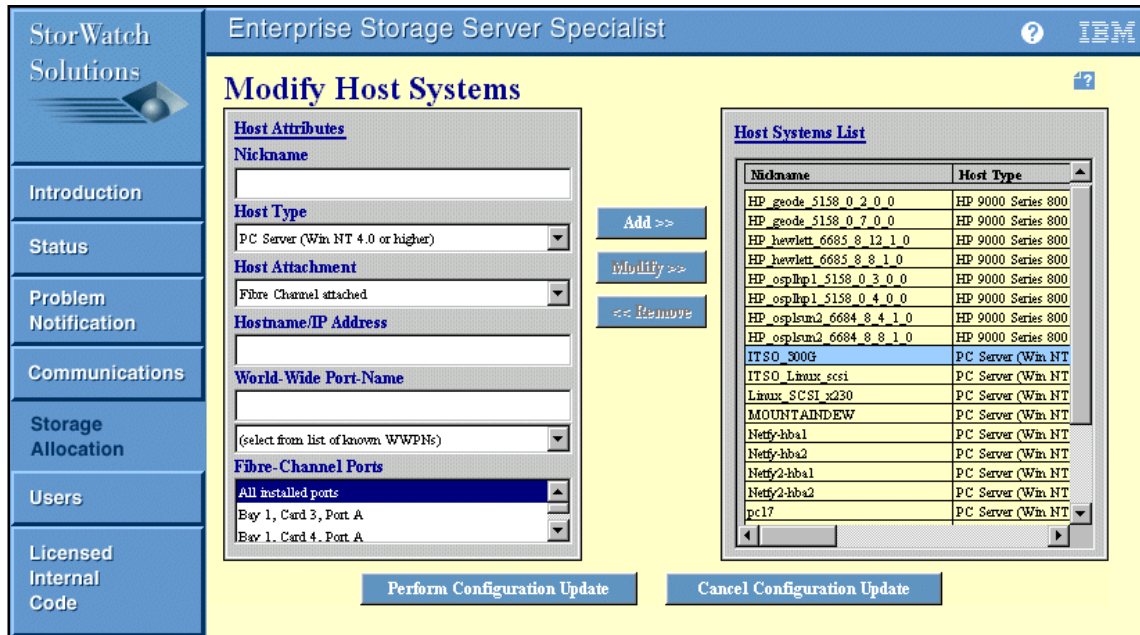


Figure 3-53 Added host systems

None of the changes we have made to this point will take effect until the **Perform Configuration Update** button at the bottom of the screen is selected. Once we select this button, we see a progress window, as shown in Figure 3-54.

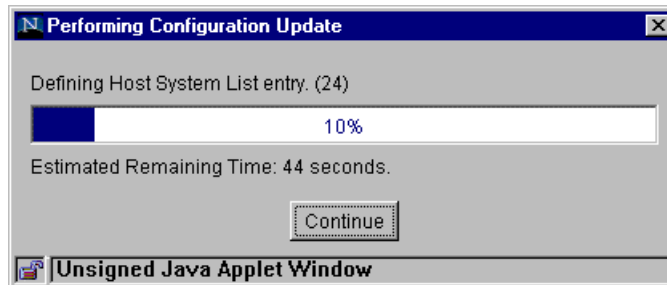


Figure 3-54 Host modification in progress

Configure host adapter ports

Now we have configured our hosts, we need to configure the host adapter ports in the ESS, as follows.

Select the **Configure Host Adapter Ports** button. This will bring up a display similar to the one in Figure 3-55. The HAs are graphically represented beneath the Configure Host Adapter Ports title. ESCON and SCSI HAs are identified by two ports located on the top of each HA in the view while Fibre Channel HAs have only a single port. ESCON and SCSI can be differentiated by the detail on the HA representation. Also, by clicking the HAs icon or by selecting the bay-adapter-port in the Host Adapter Port pull-down, different attributes will be displayed below the row of HAs.

Finally, only those adapter slots that are occupied will be visible. Empty adapter slots will not be visible.

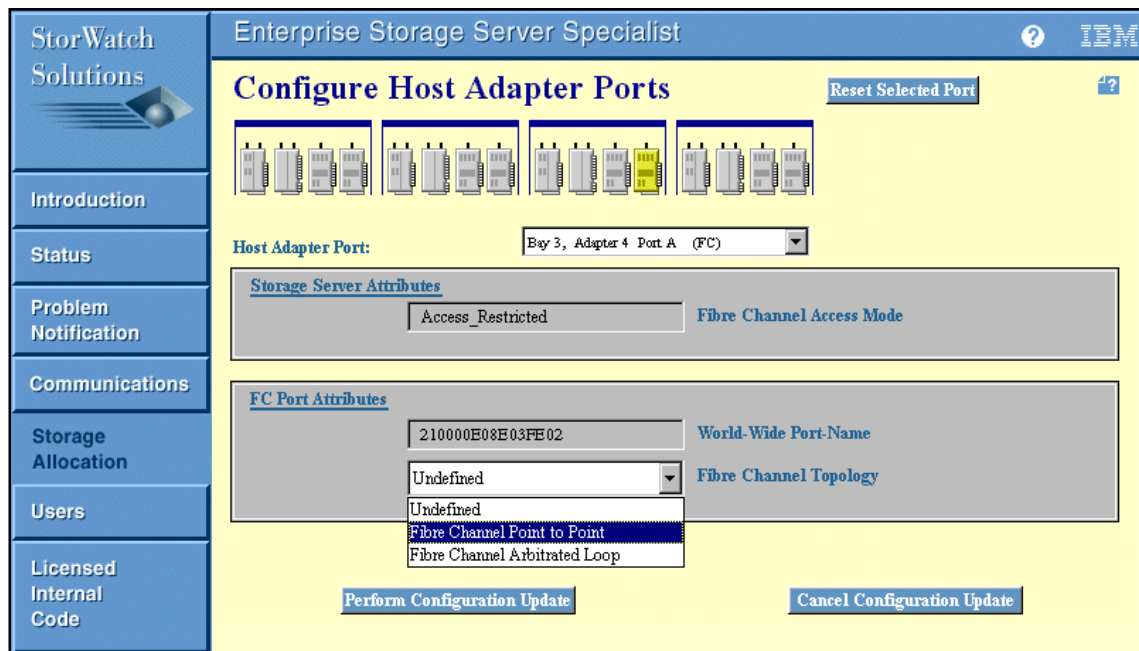


Figure 3-55 Configure host adapter ports

Unlike SCSI, Fibre Channel allows all hosts in a FC-SW or FC-AL environment to view all storage available within the environment (assuming no zoning has been established within a switch or restrictions put in place elsewhere). In order to get around the “see everything” issue, the Fibre Channel HAs within the ESS can be configured with **Access Restricted** attributes, as shown in Figure 3-55.

Using **Access Restricted** mode, the ESS limits the visibility of the LUNs to only those WWNs associated with each LUN. In effect, the ESS performs LUN masking to prevent other hosts from gaining access to LUNs that are not defined as available to those hosts.

Select a Fibre Channel host adapter card at the top of screen and we see whether it is configured as **Undefined**, **Fibre Channel Point to Point**, or **Fibre Channel Arbitrated Loop**. We can either use a predefined adapter or define our own by finding an adapter that is **Undefined** and selecting the appropriate Fibre Channel topology. When connecting direct from the host to the ESS or via a hub, we must select **Fibre Channel Arbitrated Loop**. If connecting via a switch, we must select **Fibre Channel Point to Point**.

At this point we can click **Perform Configuration Update**.

Configure Disk Groups

After performing the above update, we will automatically be returned to the Open System Storage panel. From here, click the **Configure Disk Groups** button. This brings up the window shown in Figure 3-56.

Note: This step is only necessary if no disk groups exist, or if there is insufficient space on the currently available disk groups.

The screenshot shows the 'Fixed Block Storage' configuration window in the StorWatch Solutions Enterprise Storage Server Specialist. The window has a sidebar on the left with navigation options: Introduction, Status, Problem Notification, Communications, Storage Allocation, Users, and Licensed Internal Code. The main area is titled 'Fixed Block Storage' and contains a table of 'Available Storage' groups. Below the table are 'Disk Group Attributes' and two buttons: 'Perform Configuration Update' and 'Cancel Configuration Update'.

Modification	Disk Group	Storage Type	Track Format	Capacity
	Device Adapter Pair: 2, Cluster: 2, Loop: A, Array: 1	RAID Array	Fixed Block (FB)	Formatted: 210.48 GB
	Device Adapter Pair: 2, Cluster: 1, Loop: A, Array: 2	RAID Array	Fixed Block (FB)	Formatted: 210.44 GB
	Device Adapter Pair: 2, Cluster: 2, Loop: B, Array: 1	RAID Array	Fixed Block (FB)	Formatted: 210.48 GB
	Device Adapter Pair: 2, Cluster: 1, Loop: B, Group: 2	Undefined		Unformatted: 254.80 GB
	Device Adapter Pair: 3, Cluster: 2, Loop: A, Array: 1	RAID Array	Fixed Block (FB)	Formatted: 210.48 GB
	Device Adapter Pair: 3,	Undefined		Unformatted: 254.80 GB

Disk Group Attributes

Storage Type: Undefined
Track Format: None (unused disk)

Buttons: Perform Configuration Update, Cancel Configuration Update

Figure 3-56 Fixed block storage groups

Scrolling through the table, we select an undefined disk group, select a Storage Type from the drop-down list (we chose **RAID ARRAY**) and **Fixed Block** as our Track Format, as shown in Figure 3-57.

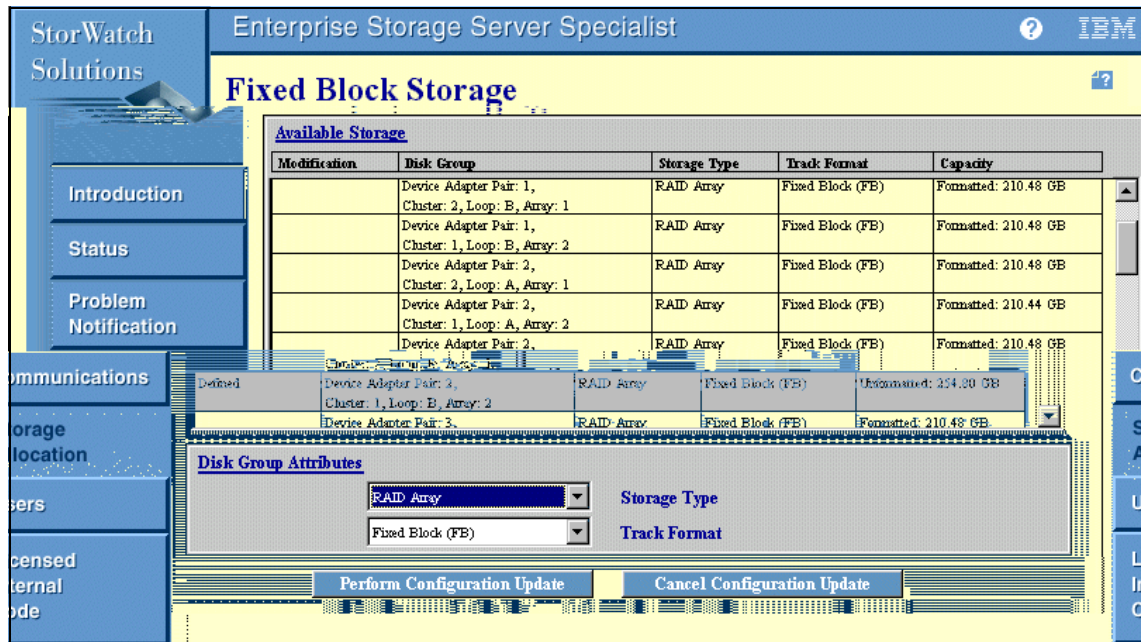


Figure 3-57 Define fixed block storage (RAID array)

Clicking the **Perform Configuration Update** button will reveal a warning message, as shown in Figure 3-58. This configuration update initializes the disk group ready for volumes to be created on it.



Figure 3-58 Time intensive action warning

Clicking **Yes** will begin the process. (We recommend either taking a lunch break now or saving this action until just before going home and then picking up again the following day, because, depending on how busy the subsystem is, this process can take a couple of hours.)

Add volumes

Now that we have a defined Open System disk group, we are ready to add volumes to it and assign them to our hosts. From the Open Systems Storage panel, click the **Add Volumes** button on the bottom left of the screen. Doing so will reveal the Add Volumes (1 of 2) panel, as shown in Figure 3-59.

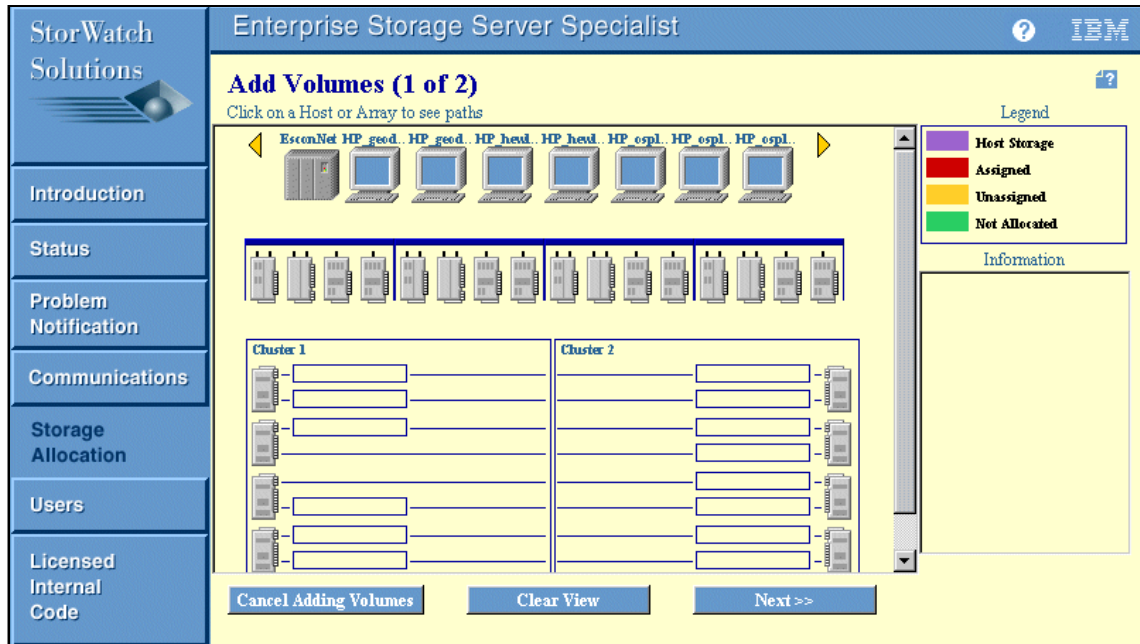


Figure 3-59 Add volumes panel

Scroll through the list of hosts at the top of the screen and select the Fibre Channel host we defined earlier. Once selected, lines will be drawn to all host adapters through which that host can access LUNs, as shown in Figure 3-60. Select the appropriate host adapter.

Note: After selecting a fibre-attached host, all Fibre Channel host adapters in the ESS will be highlighted as being valid access paths to LUNs. This indicates *possible* connections rather than actual connections as, in the case when we are direct connected, we can actually only see LUNs through the host adapter we are physically connected to. Be sure to select a Fibre Channel adapter we are *actually* able to connect through, as dictated by direct connection or SAN zoning.

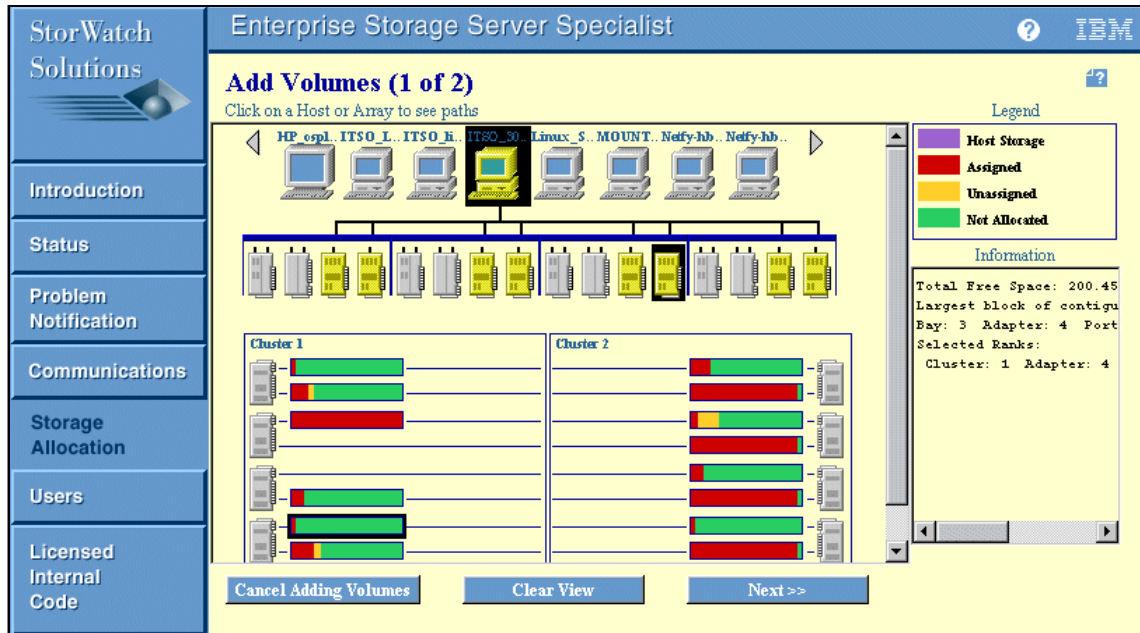


Figure 3-60 Add volumes to selected host

After selecting a valid Fibre Channel host adapter, we highlight the disk group (or groups) on which we wish to create LUNs. Then, click the **Next** button to reveal the second configuration page, as shown in Figure 3-61.

Note: We are showing the process of setting up LUNs on disk groups, but this should not be done in isolation from, or prior to, an end-to-end storage plan for the subsystem.

StorWatch Solutions Enterprise Storage Server Specialist ? IBM

Add Volumes (2 of 2) ?

Available Free Space

Storage Type	Available Capacity	Maximum Volume Size
RAID-5 Array	200.45 GB	200.45 GB
Non-RAID	0.00 GB	0.00 GB

Volume Attributes

Select a Volume Size

10.0 GB

10.1 GB

10.2 GB

Number of Volumes (Enter 1 to 20)

2

Storage Type

RAID-5 Array

New Volumes

Number	Volume Size	Storage Type
Total: 0 GB		

Volume Placement

Place volumes sequentially, starting in first selected storage area

Spread volumes across all selected storage areas

Figure 3-61 Select number and size of LUNs

On this panel we select an item from the list of available free space in the top window (typically RAID-5 Array), then select the size of the volume(s) we wish to create in the Volume Attributes window and the number of volumes of that size we wish to create. We have selected to create two 10 GB LUNs. If we had selected more than one disk group in the step before this, we could also select to have our LUNs distributed evenly across those groups. As we did not do so, we will leave the default Volume Placement selection as is.

We click the **Add** button, and the new LUNs appear in the New Volumes window on the right of the screen, as shown in Figure 3-62.

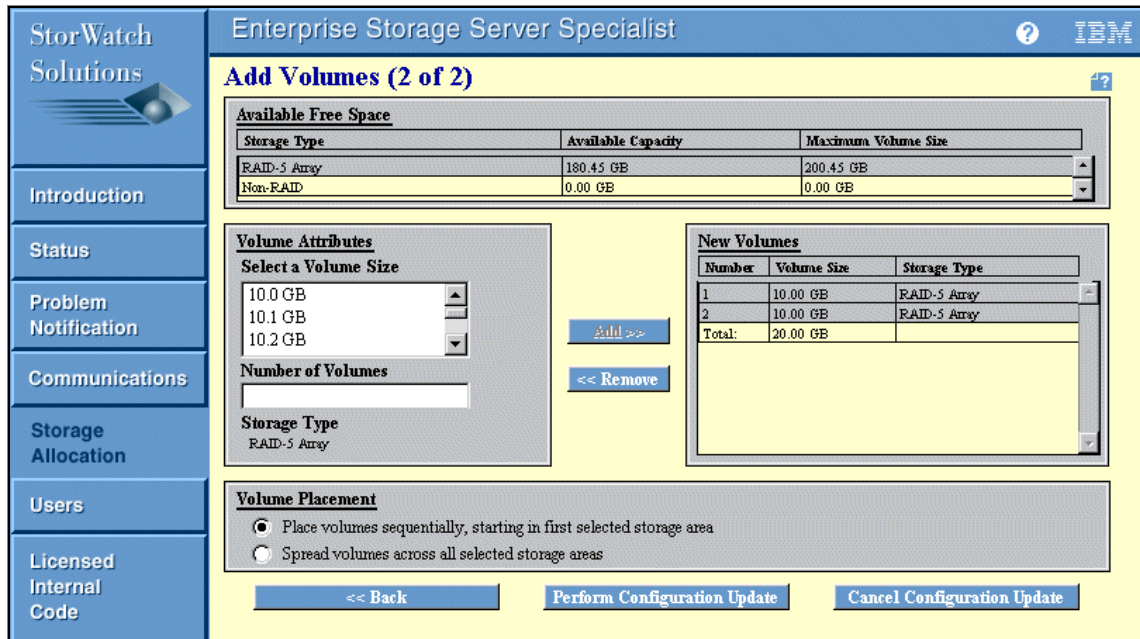


Figure 3-62 New volumes created

Create additional LUNs as desired. When finished, click **Perform Configuration Update**. Once again, we see a “time intensive activity” warning window followed by the progress indicator and, finally, the “Volumes Successfully Added” message box.

When this is done, we are returned to the Add Volumes panel.

Modify volume assignments

From time-to-time, it may become necessary to modify the assignment of volumes. For example, if we wish to assign a LUN to the second host in a cluster or assign a LUN to a second adapter in an existing host (which must have multi-path support in host operating system). We are going to assign LUNs to a second adapter in our host.

First we need to define our new host as in “Modify Host Systems” on page 122 and the host adapter port in the ESS, as described in “Configure host adapter ports” on page 123. Then we are ready to assign our LUN using the Modify Volume Assignments display.

From the Open System Storage window (see Figure 3-51 on page 121), select the **Modify Volume Assignments** button. A display similar to the one in Figure 3-63 should appear.

StorWatch Solutions Enterprise Storage Server Specialist

Modify Volume Assignments

Volume Assignments Refresh Status Print Table Perform Sort

no sort no sort no sort no sort no sort no sort no sort no sort

Volume	Location	LSS	Volume Type	Size	Storage Type	Host Port	Host Nicknames
	Cluster 1, Loop B Array 2, Vol 025					ID 00, LUN 0005	
61A-18540	Device Adapter Pair 4 Cluster 1, Loop B Array 2, Vol 026	16	Open System	004.0 GB	RAID Array	Fibre Channel ID 00, LUN 0006	ITSO_300G
61B-18540	Device Adapter Pair 4 Cluster 1, Loop B Array 2, Vol 027	16	Open System	004.0 GB	RAID Array	Fibre Channel ID 00, LUN 0007	ITSO_linux_fc
61C-18540	Device Adapter Pair 4 Cluster 1, Loop A Array 2, Vol 028	16	Open System	010.0 GB	RAID Array	Unassigned	

Action

Assign selected volume(s) to target hosts
 Use same ID/Lun in source and target
 Unassign selected volume(s) from target hosts

Target Hosts

HP_osplsun2_6684_8_8_1_0
ITSO_300G
MOUNTAINDEW

Perform Configuration Update Cancel Configuration Update

Figure 3-63 Modify volume assignments

Select the volume(s) that we wish to modify from the table. When we select the Action radio button (either assign or unassign) notice the list of hosts in the Target Host window on the bottom right of screen changes. Select the host or hosts to perform the action on, and select **Perform Configuration Update**.

Note: The check box, Use same ID/Lun in source and target, is optionally selected to allow some control over the ID and LUN used for the new assignment.

We should now see the familiar progress indicator followed by the “volume assignment successful” message box. If we now scroll back through the Modify Volume Assignments table (easier if we sort first), we see that our LUNs are indeed now assigned to our host (Figure 3-64).

The screenshot shows the 'Modify Volume Assignments' window in the StorWatch Enterprise Storage Server Specialist interface. The window has a sidebar on the left with navigation options: Introduction, Status, Problem Notification, Communications, Storage Allocation, Users, and Licensed Internal Code. The main area is titled 'Modify Volume Assignments' and contains a 'Volume Assignments' table. Above the table are buttons for 'Refresh Status', 'Print Table', and 'Perform Sort', along with several 'no sort' dropdown menus. The table has columns for Volume, Location, LSS, Volume Type, Size, Storage Type, Host Port, and Host Nicknames. Below the table is an 'Action' section with radio buttons for 'Assign selected volume(s) to target hosts' (selected), 'Unassign selected volume(s) from target hosts', and a checkbox for 'Use same ID/Lun in source and target'. To the right is a 'Target Hosts' text area. At the bottom are buttons for 'Perform Configuration Update' and 'Cancel Configuration Update'.

Volume	Location	LSS	Volume Type	Size	Storage Type	Host Port	Host Nicknames
61A-18540	Device Adapter Pair 4 Cluster 1, Loop B Array 2, Vol 026	16	Open System	004.0 GB	RAID Array	Fibre Channel ID 00, LUN 0006	ITSO_300G
61A-18540	Device Adapter Pair 4 Cluster 1, Loop B Array 2, Vol 026	16	Open System	004.0 GB	RAID Array	Fibre Channel ID 00, LUN 0006	LINUX_ARWED
61B-18540	Device Adapter Pair 4 Cluster 1, Loop B Array 2, Vol 027	16	Open System	004.0 GB	RAID Array	Fibre Channel ID 00, LUN 0007	ITSO_linux_fc
61B-18540	Device Adapter Pair 4 Cluster 1, Loop B	16	Open System	004.0 GB	RAID Array	Fibre Channel ID 00, LUN 0007	LINUX_ARWED

Figure 3-64 Validate volume assignment modification

We have now completed all the tasks required on the ESS side of things. We have allocated storage to our 300G, and all we have to do now is claim it. In the next section, we return to the 300G and explain how to start using all this storage we have set aside for it.

3.6 Claiming ownership of pooled storage with the 300G

Since (at the end of the section “Re-initializing the 300G” on page 72) we brought the 300G back up with the keyboard, monitor, and mouse removed, we need to access it remotely in order to configure our ready-and-waiting storage. To do this, we go to our administration station and launch Internet Explorer. In the address bar of Internet Explorer, we enter the IP address of the 300G followed by a colon (:), and 8099. Next, we login to the 300G using the administrator account and password. The default password for the administrator account is *password*.

Now we are in the home page of the **Windows Powered Server Appliance Tasks** screen (Figure 3-65).

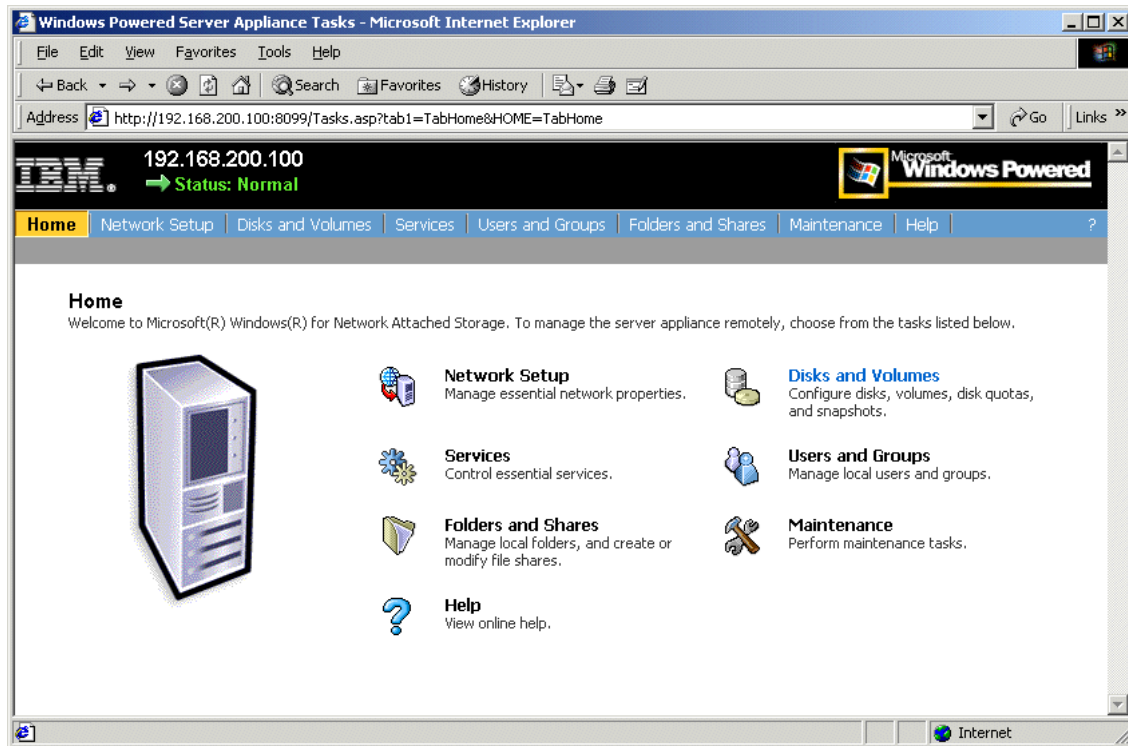


Figure 3-65 Windows Powered Server Appliance Tasks

Select to enter **Disks and Volumes** from either the tab on the top of the window or the link in the body of the window.

Now we need to select **Disks and Volumes** again from the link in the body of the Disks and Volumes page, as shown in Figure 3-66.

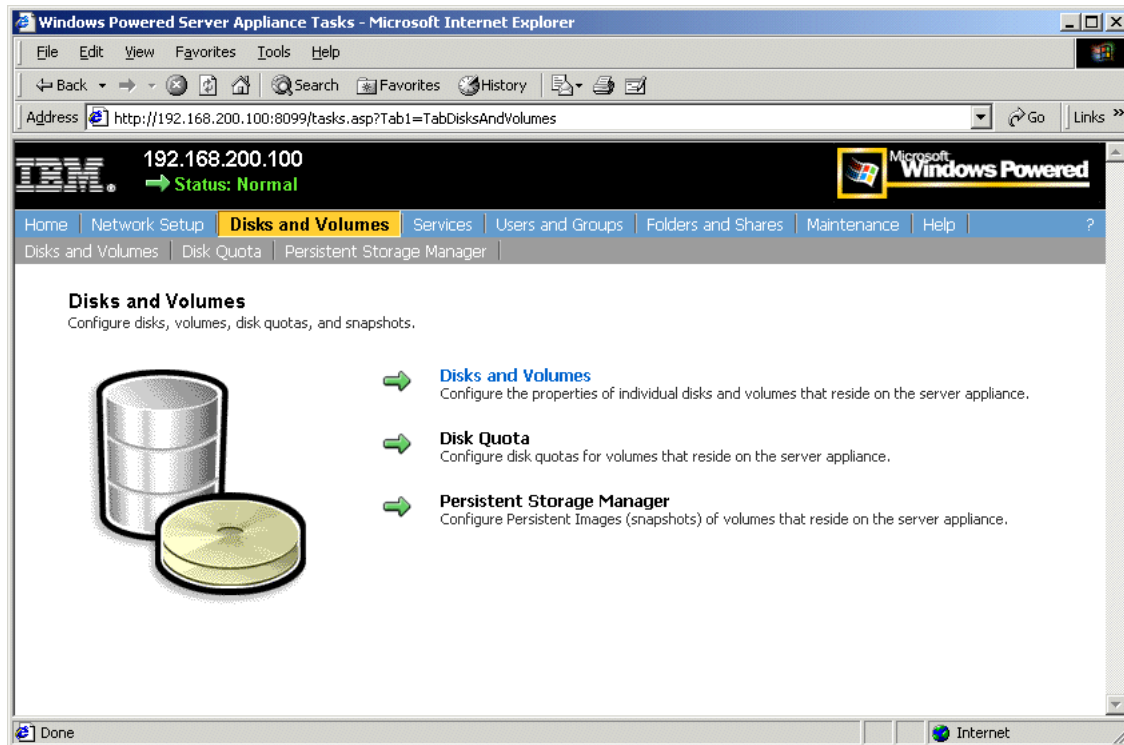


Figure 3-66 Disks and Volumes page

Now we see the login for the Terminal Services Client. We login using the administrator account again (Figure 3-67).

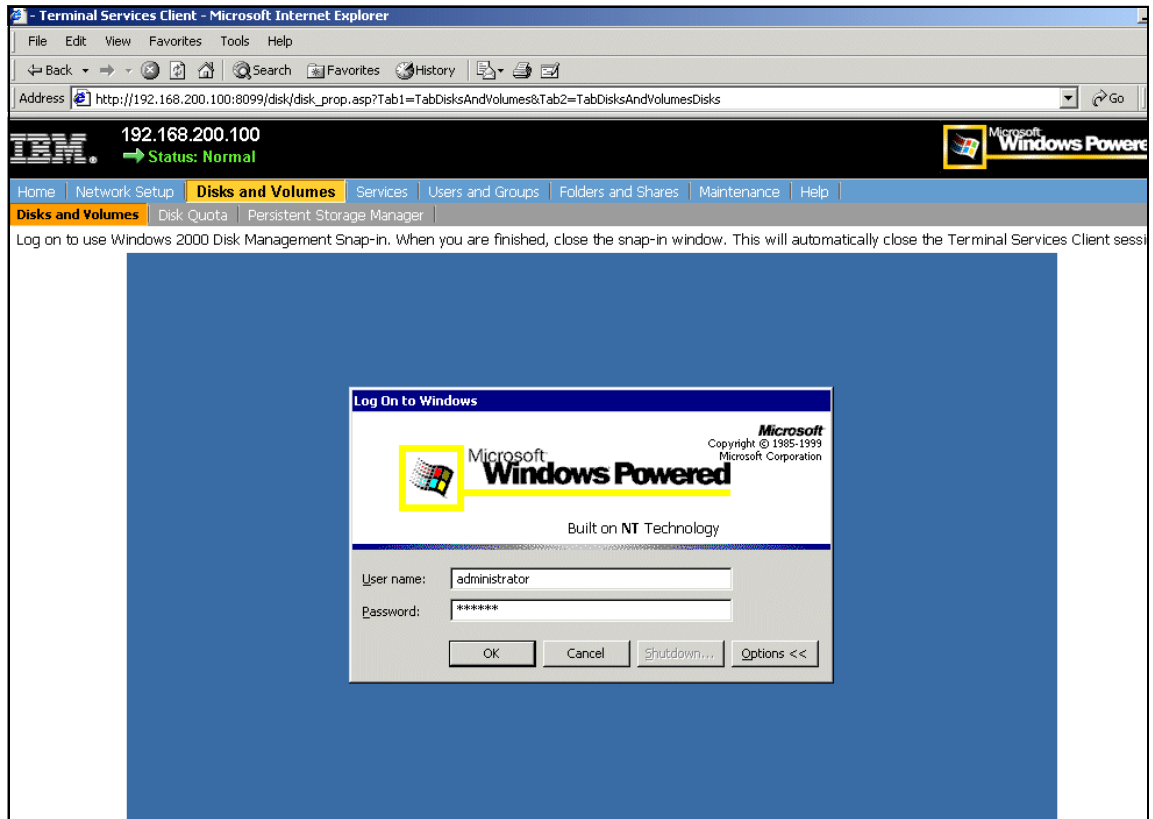


Figure 3-67 Terminal Services Client login

We are looking at the desktop of the 300G via a Terminal Services client. Now we are going to launch the **IBM NAS Admin.msc** from the shortcut provided (Figure 3-68).

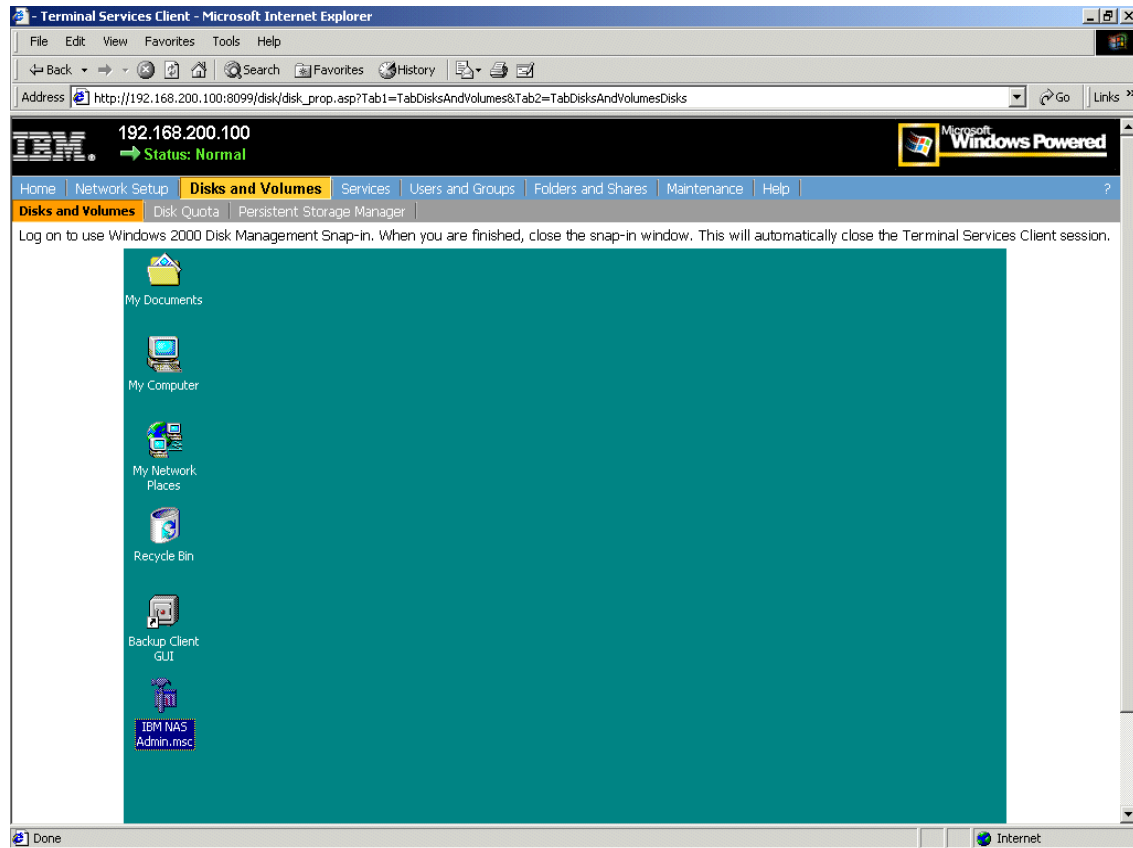


Figure 3-68 Terminal Services Client

We now open the **Storage** folder and click the **Disk Management** folder. After a few moments, the system will bring up the **Write Signature and Upgrade Disk Wizard**, as shown in Figure 3-69. Press **Next**.

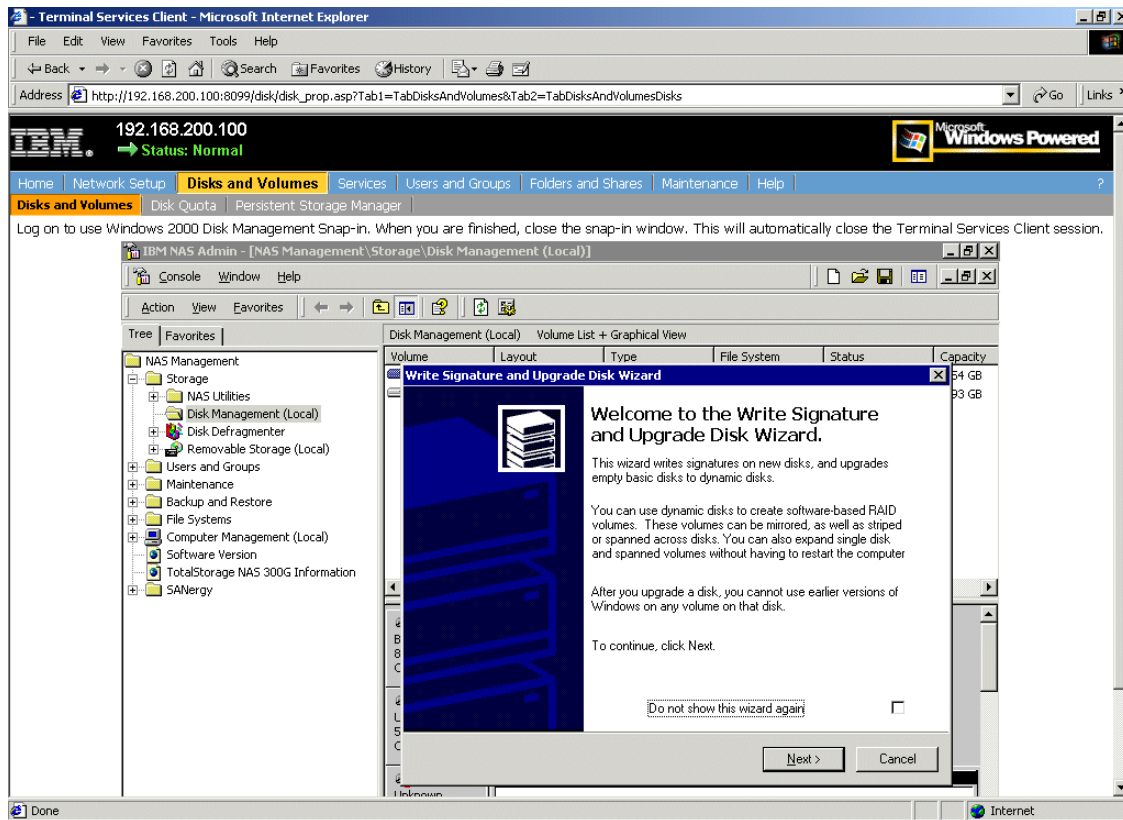


Figure 3-69 Disk management

For our purposes, we want the 300G to own the disks, but they should not be dynamic disks (since dynamic disks are not supported on the 300G). To accomplish this, select the disk to write a signature on and press **Next** (Figure 3-70).

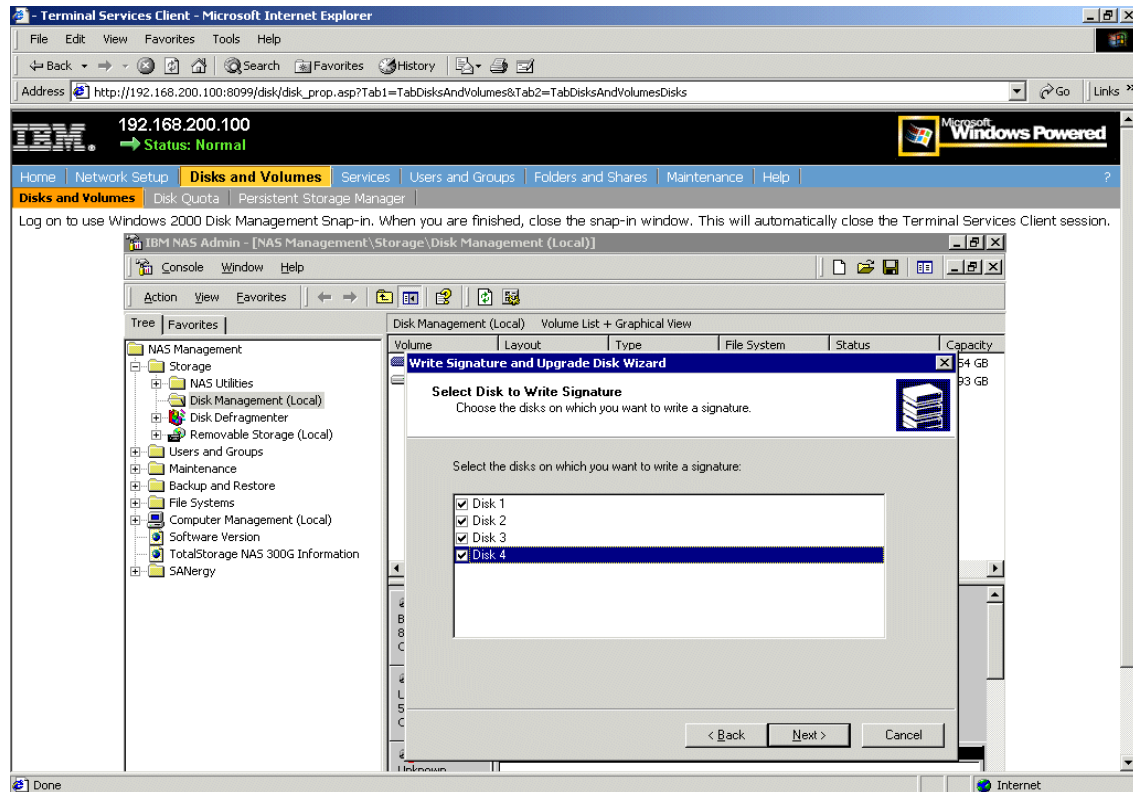


Figure 3-70 Select disks

Next, we are asked which disks to upgrade. Figure 3-71 shows that we deselected all of the disks before pressing **Next**.

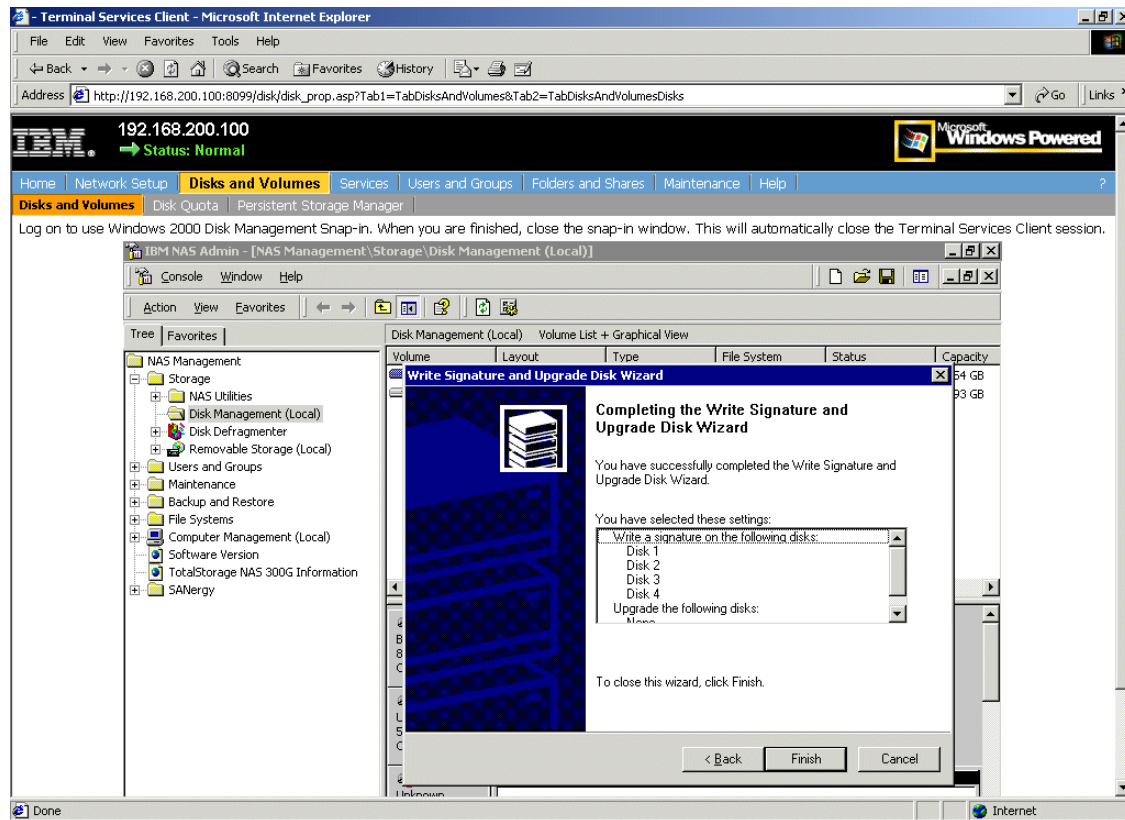


Figure 3-71 Do not upgrade disk

Now, to complete the Write Signature Wizard, we press **Next** (Figure 3-72).

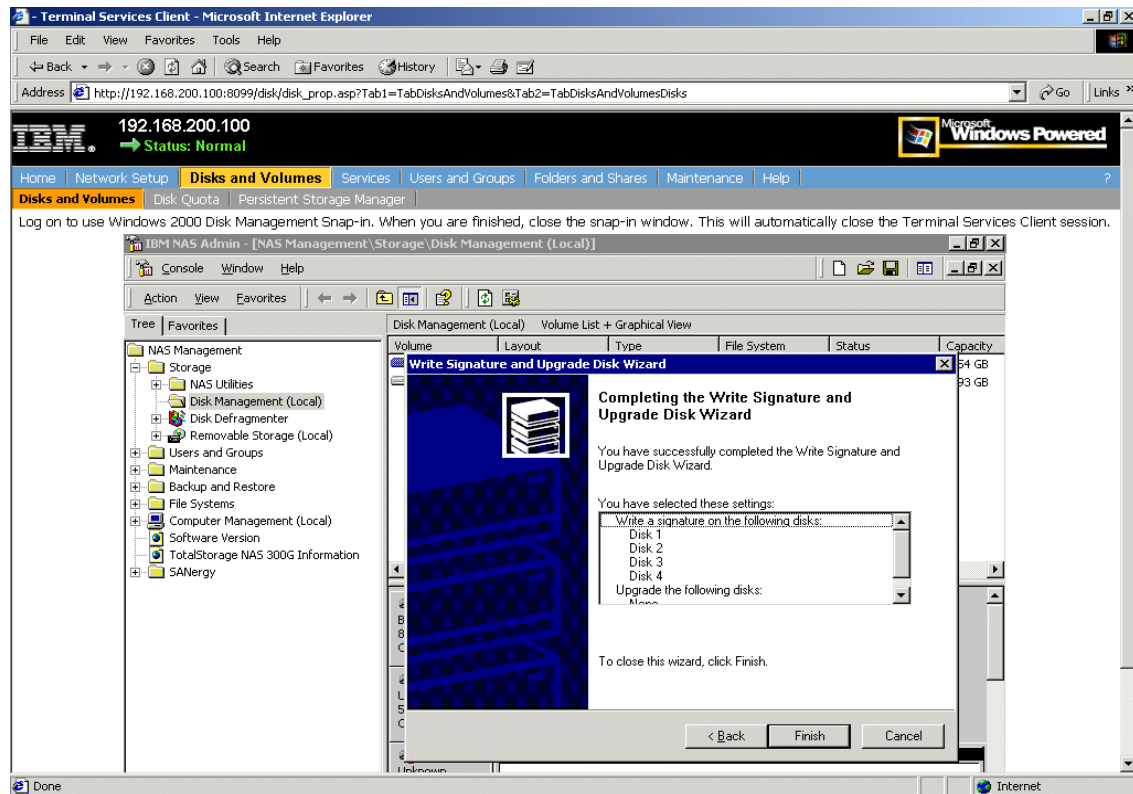


Figure 3-72 Complete Write Signature Wizard

Once the wizard completes, we begin creating partitions by right-clicking on a new disk and selecting **Create Partition** (Figure 3-73).

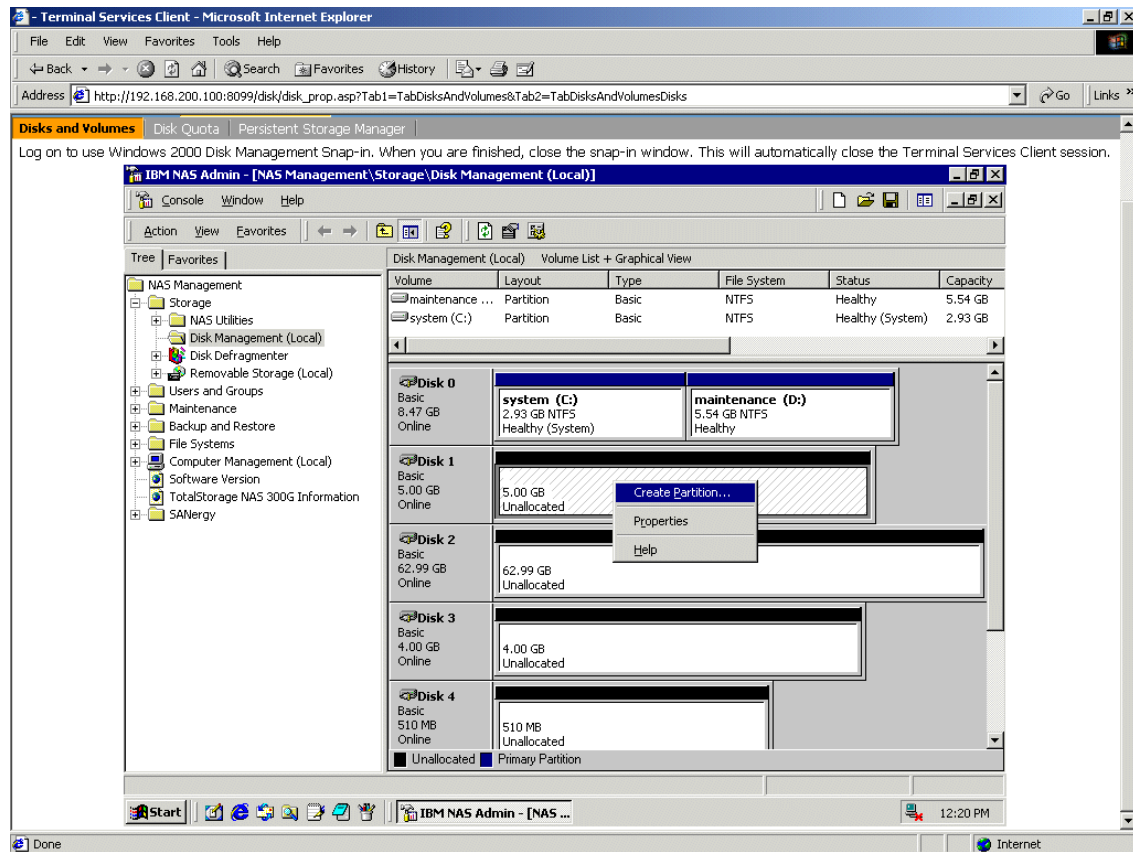


Figure 3-73 Create partition

Click **Next** on the Create Partition Wizard (Figure 3-74).

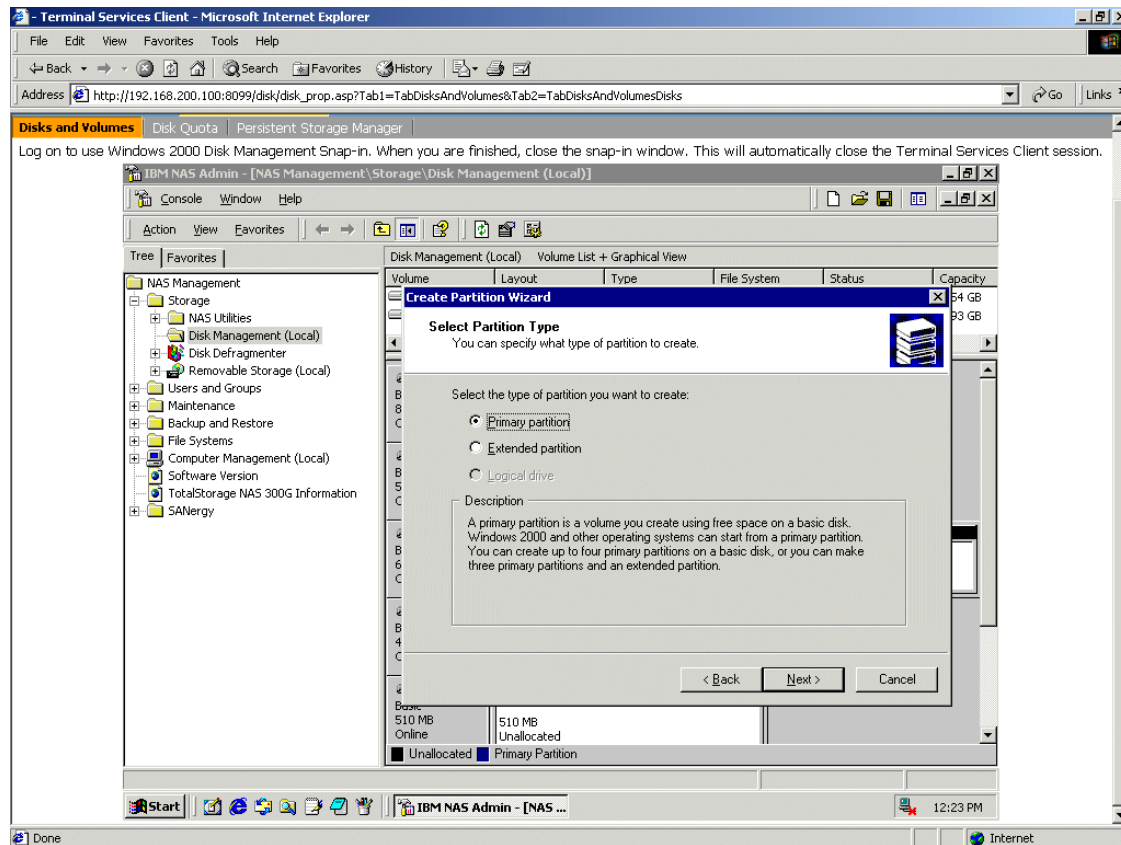


Figure 3-74 Create Partition Wizard

We chose to make a primary partition and clicked **Next** (Figure 3-75).

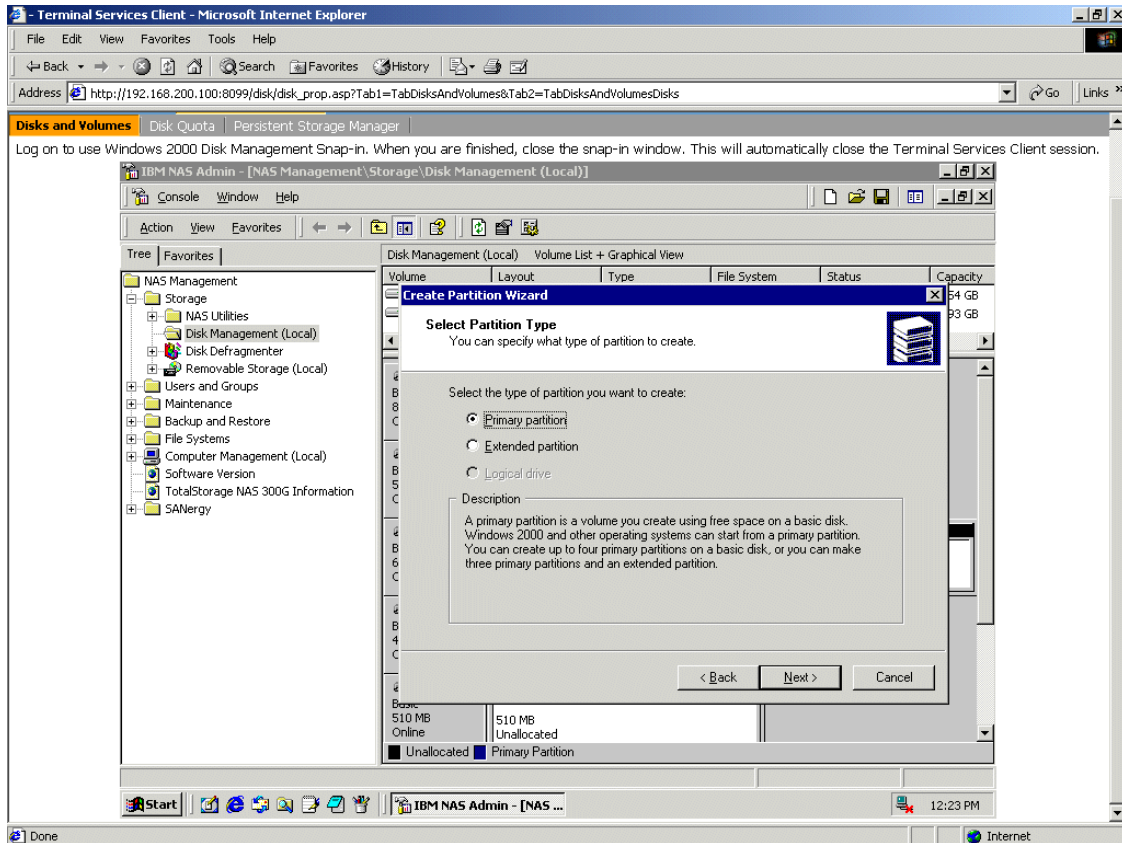


Figure 3-75 Select Partition Type

We used all available space and clicked **Next** (Figure 3-76).

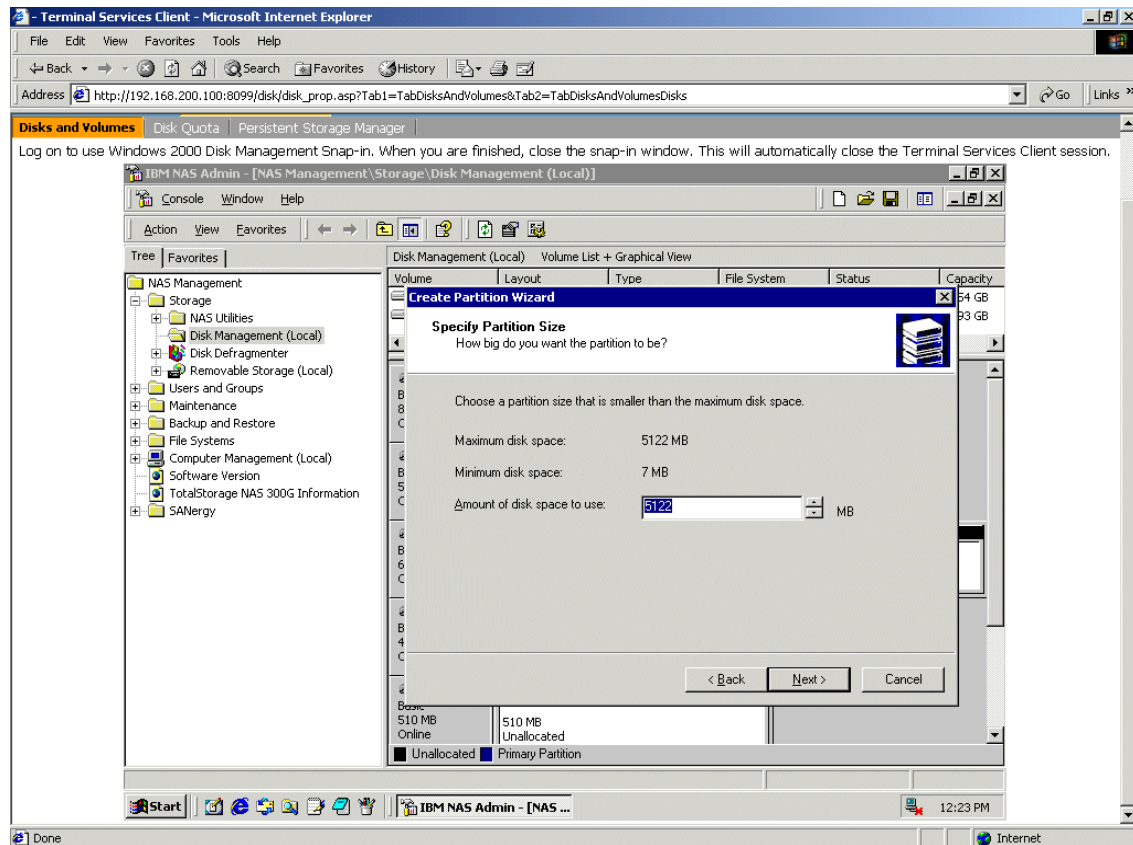


Figure 3-76 Specify Partition Size

Now we select drive letter **H:** and click **Next**, as shown in Figure 3-77.

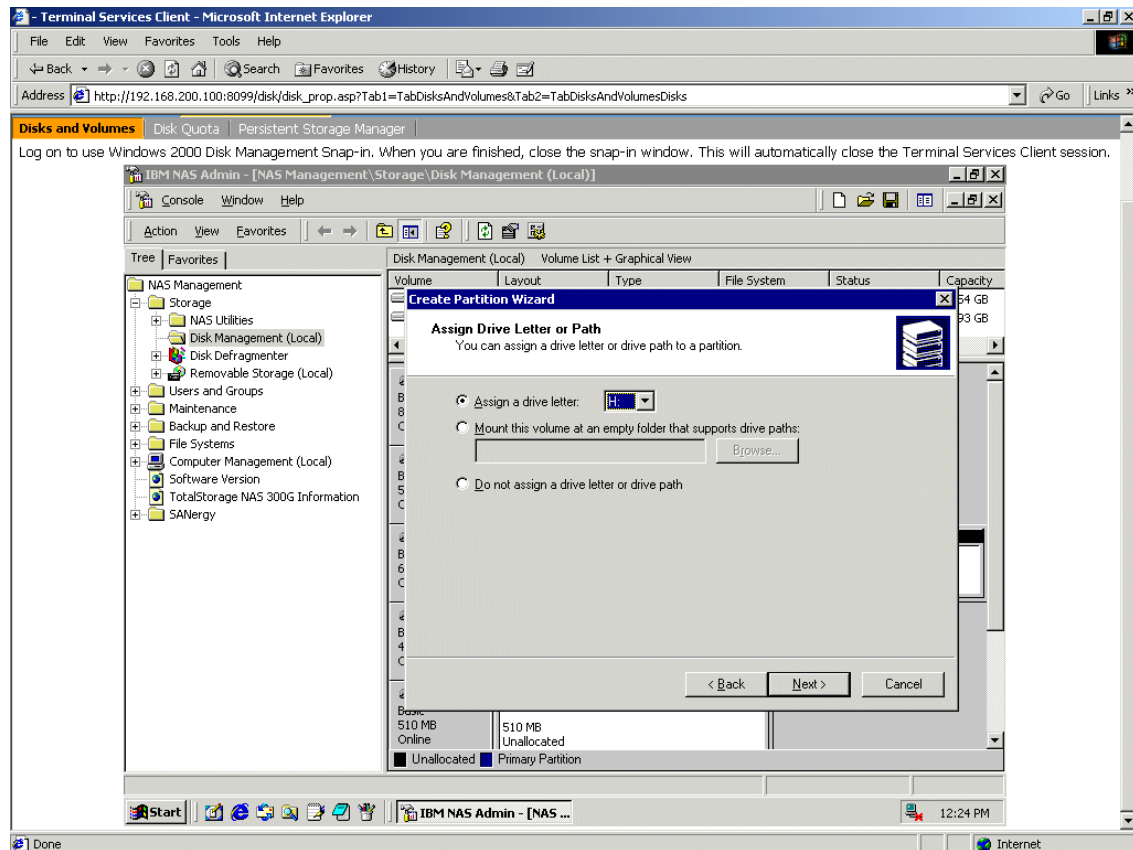


Figure 3-77 Assign Drive Letter

The Create Partition Wizard is almost complete. We left the defaults of NTFS for the file system and default for allocation unit size. We entered our volume name. We did not do a quick format, and did not enable compression.

With these choices made, we clicked **Next** (Figure 3-78).

Note: File and folder compression is not supported on the 300G.

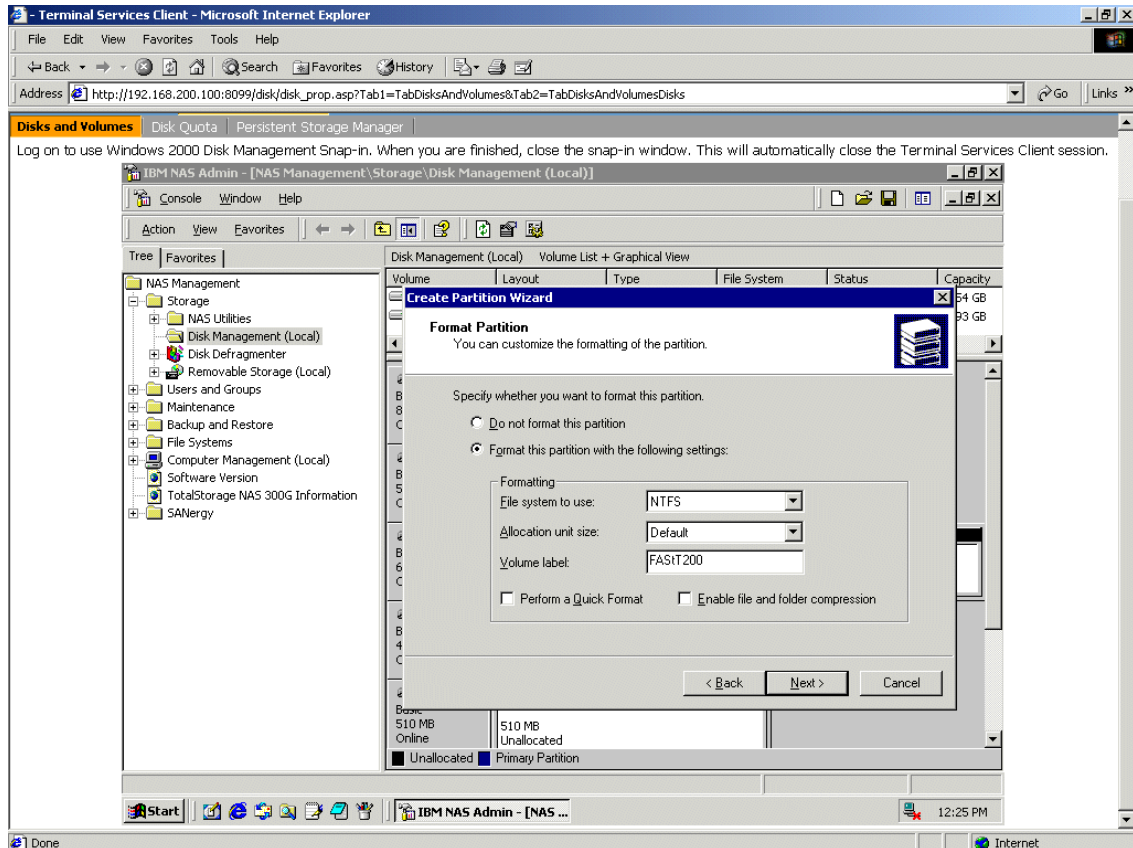


Figure 3-78 Format Partition

At this point, a final window confirming partition creation appeared, and we clicked **Finish** (Figure 3-79).

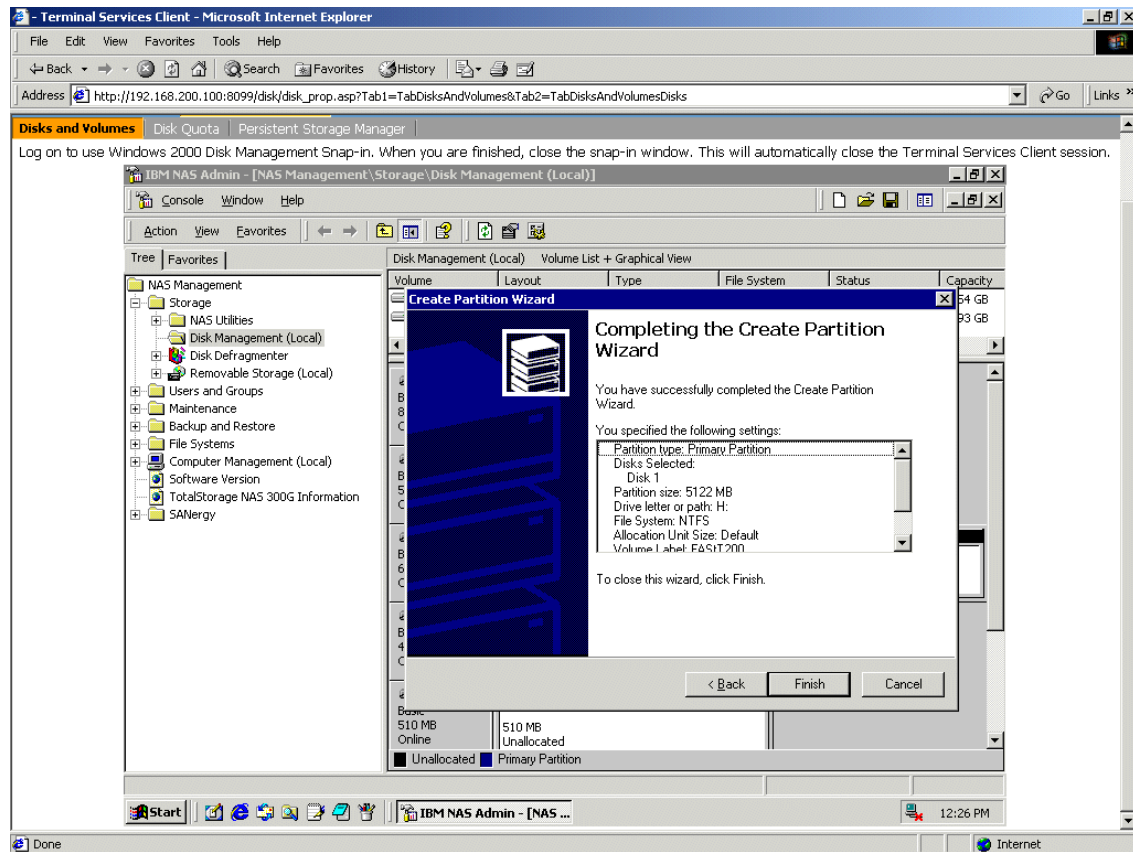


Figure 3-79 Complete the Create Partition Wizard

Figure 3-80 shows that once we completed the partitioning of all of our SAN disks, and the formatting was completed, we saw all of the disks as healthy.

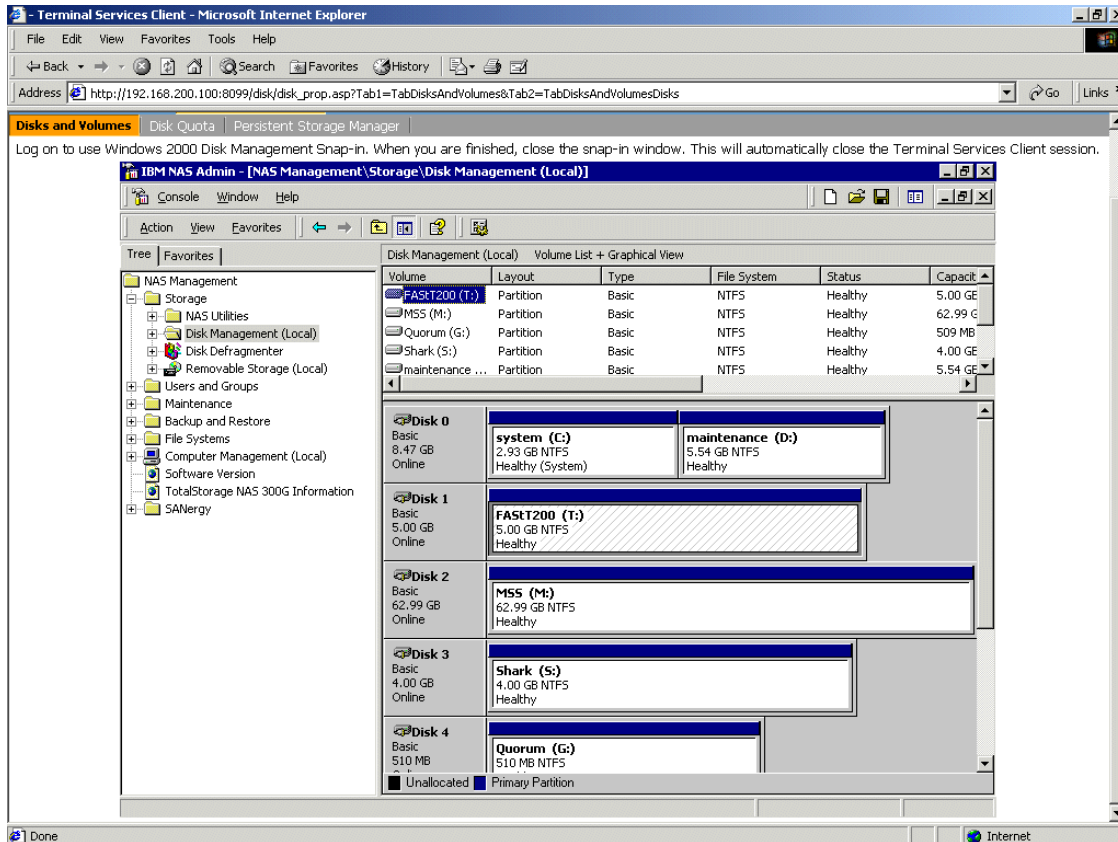


Figure 3-80 Healthy disks

All that remains is to make the storage available to other client machines. The next section explains how to share this storage over the LAN.

3.7 Sharing the SAN-based storage to LAN/WAN clients

We now have lots of storage space that is owned by our 300G, and we want to give the other machines in our network access to it. Normally, this would require a great deal of effort because, while the storage devices are connected to our SAN, none of our client machines are. Additionally, all of our client machines are connected to our LAN, but our storage devices are not. Fortunately, the 300G is connected to both networks, so it can serve as the link between the storage devices and our client machines.

Since Windows Services for UNIX comes pre-installed on the 300G, configuring the 300G to share its available storage across both Windows and UNIX platforms was very simple.

3.7.1 File sharing for Windows clients

All we had to do to set up sharing was launch the Windows Explorer, select a disk, right-click it, and chose **Sharing**, as shown in Figure 3-81.

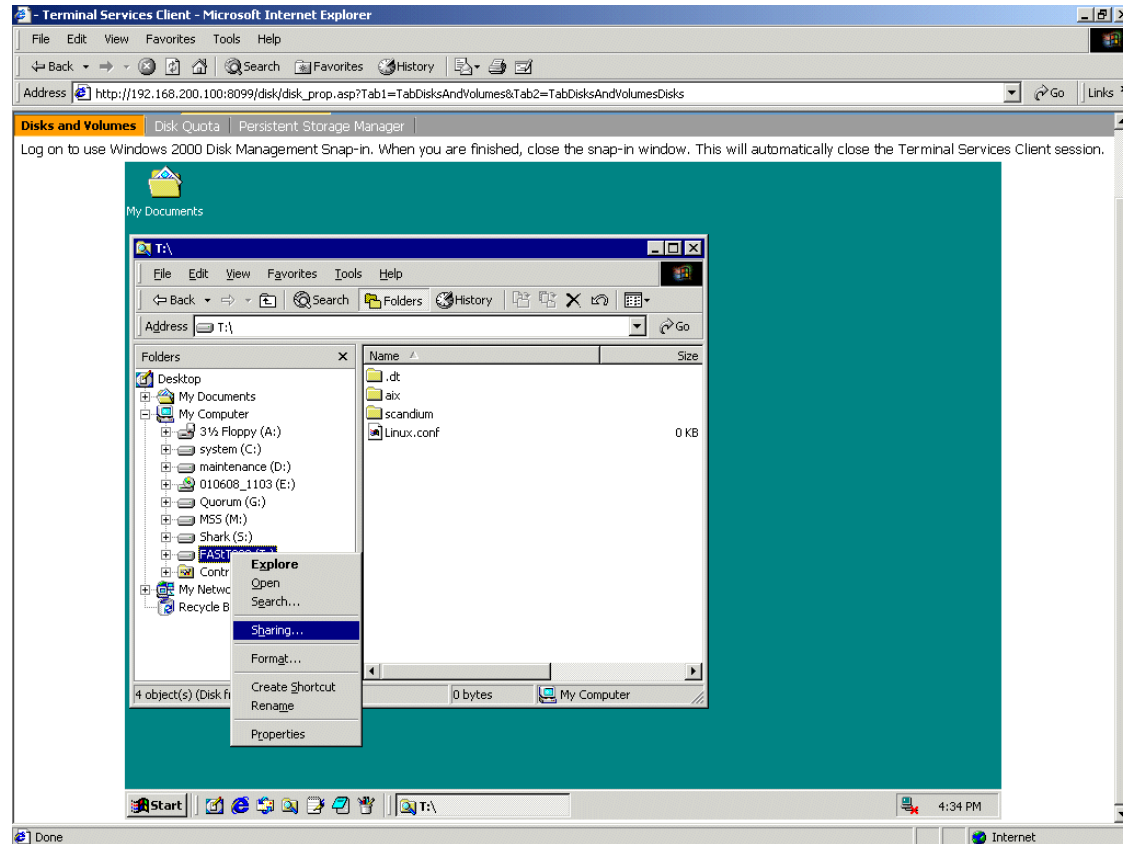


Figure 3-81 Setting up a share

When the Properties window for the selected drive opens, there is an administrative share name in place. The administrative share name ends with a dollar sign (\$), as shown in Figure 3-82.

Important: To ensure access for Windows users other than administrators, you should always define an explicit share. This is especially important if you plan on using Tivoli SANergy. Although on Windows NT and 2000 there is already a hidden administrative share present by default, you should create a separate one, because using the standard administrative shares can cause problems with shared SANergy volumes. See Chapter 4, “Clustering for high availability” on page 167 for more information.

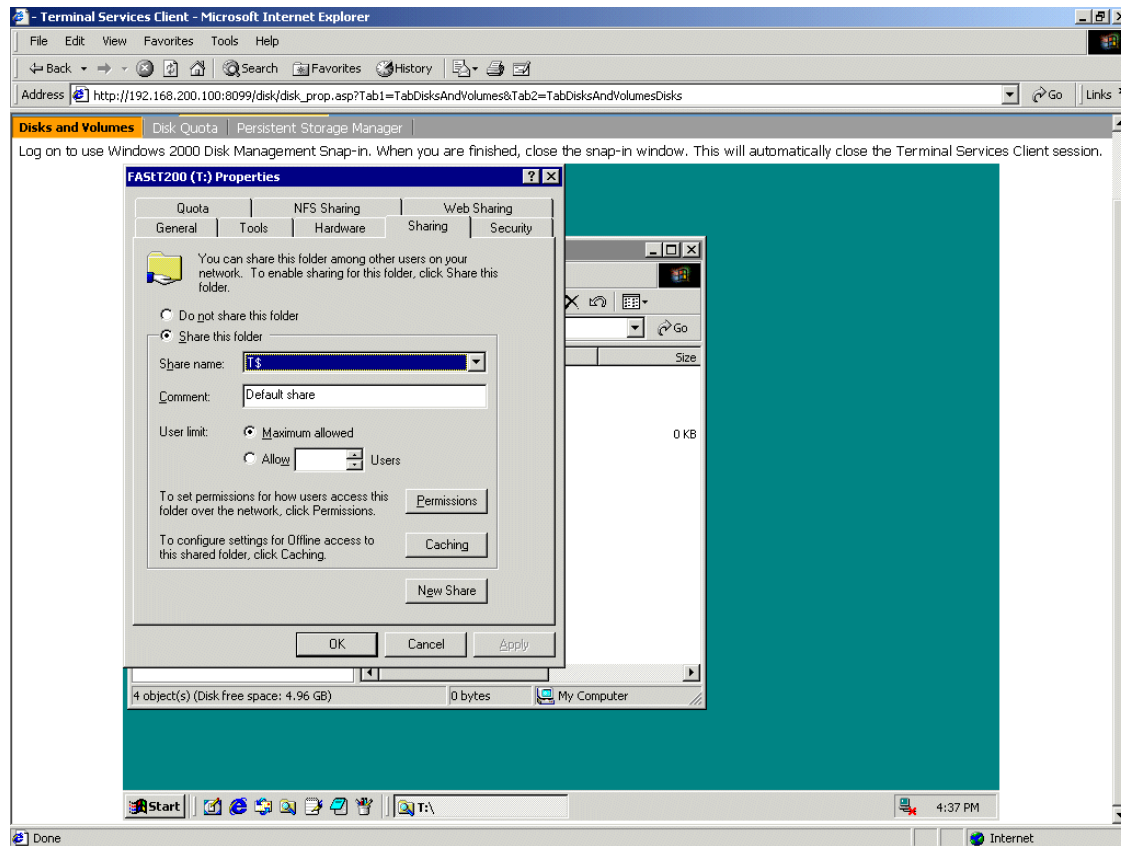


Figure 3-82 Administrative Share

In order to set up a share that would work well for us later on, we selected the “Do not share this folder” option and pressed **Apply** (Figure 3-83).

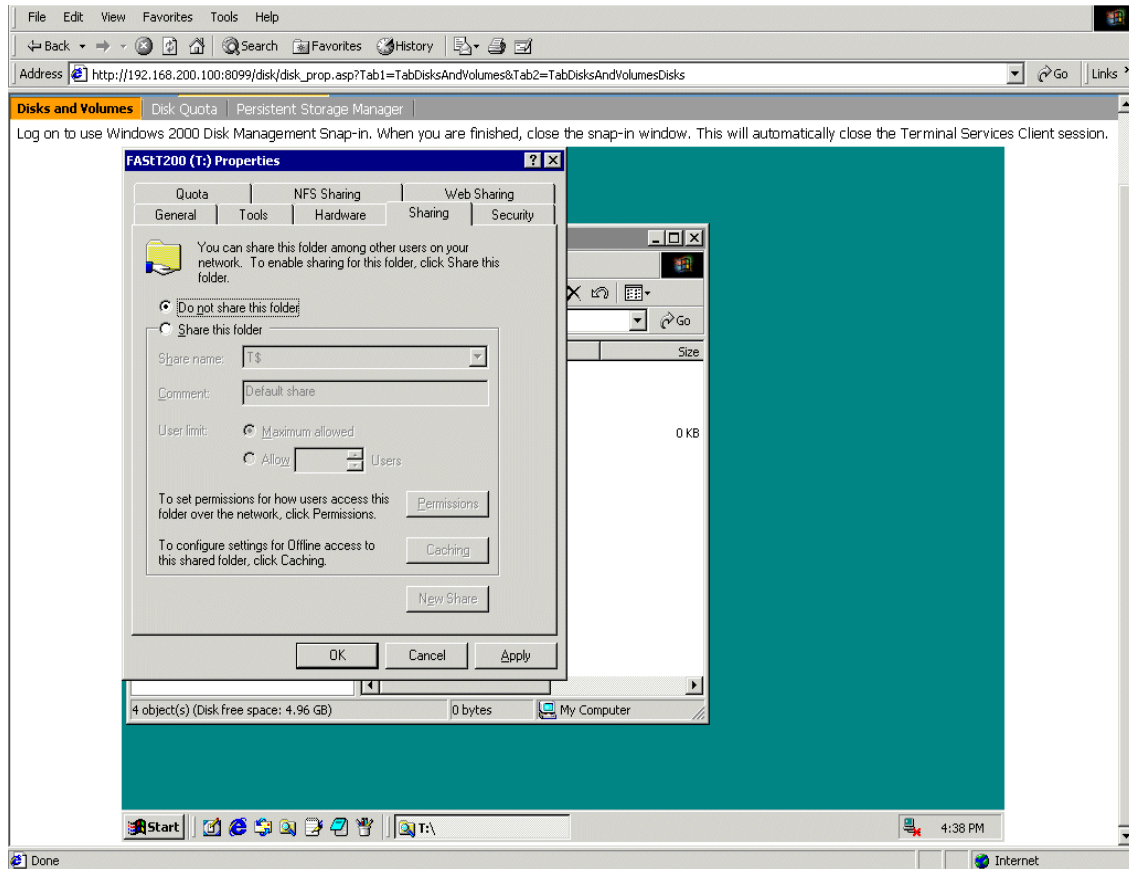


Figure 3-83 Getting rid of the administrative share

Next we selected the “Share this folder” option and supplied a share name (Figure 3-84).

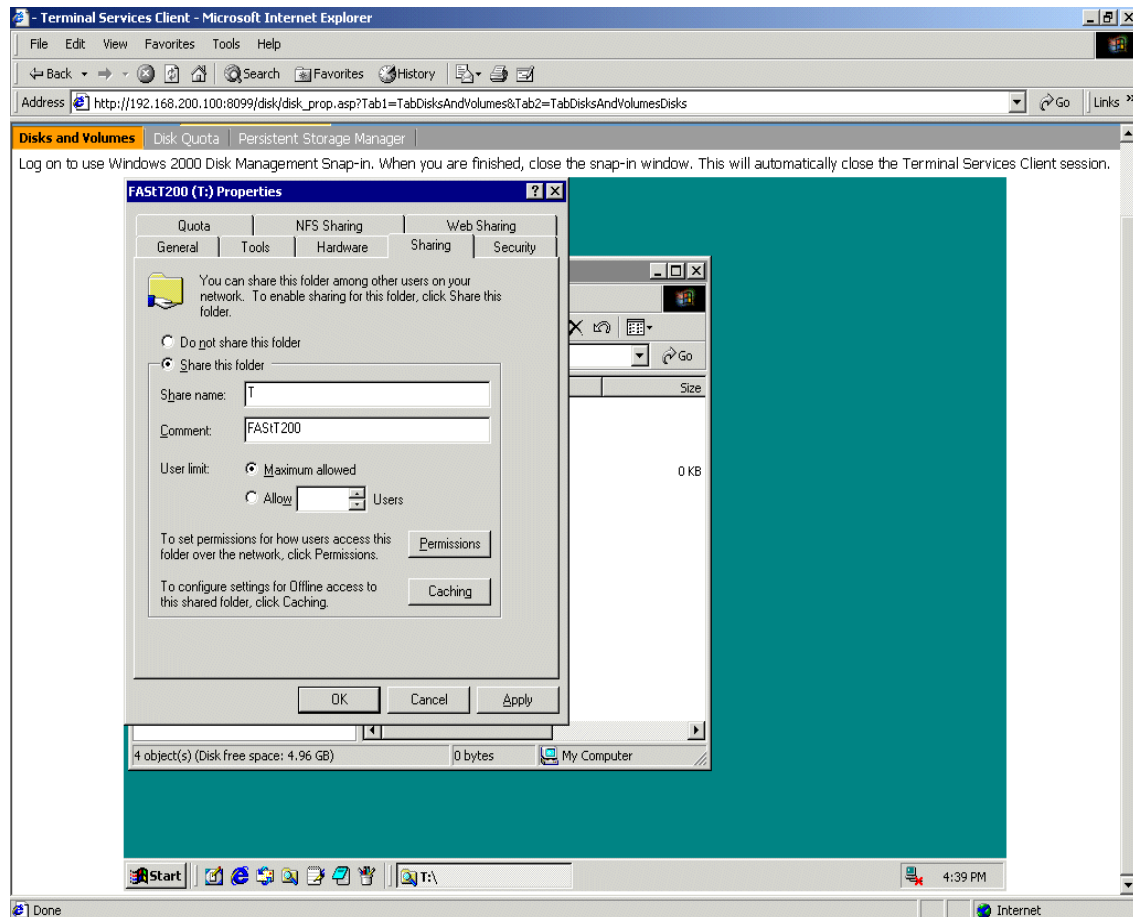


Figure 3-84 Establishing a share

Before accepting this share, we pressed the **Permissions** button so security for the share could be adjusted to meet our needs (Figure 3-85).

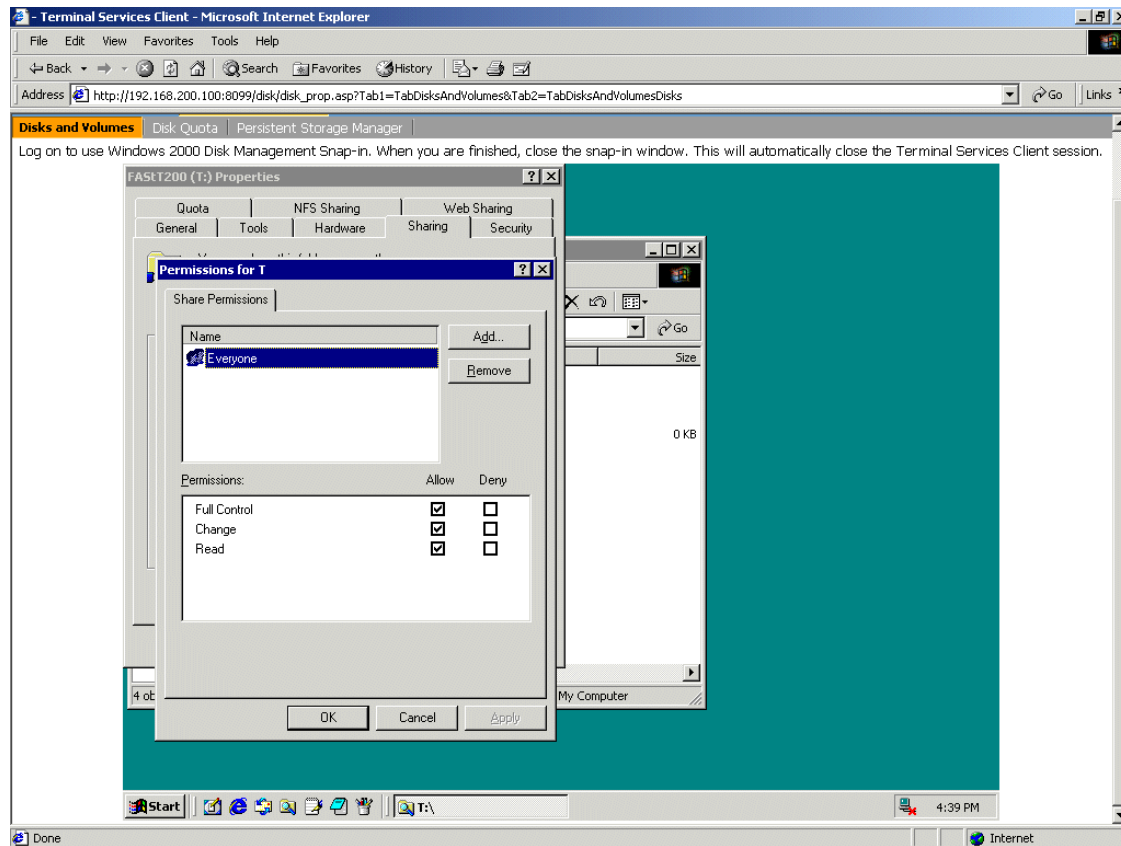


Figure 3-85 Permissions for the Windows share

3.7.2 File sharing for UNIX clients

Enabling access for UNIX systems required just one more step. From the same dialog, we clicked the **NFS Sharing** tab and set it up as well. This is shown in Example 3-86.

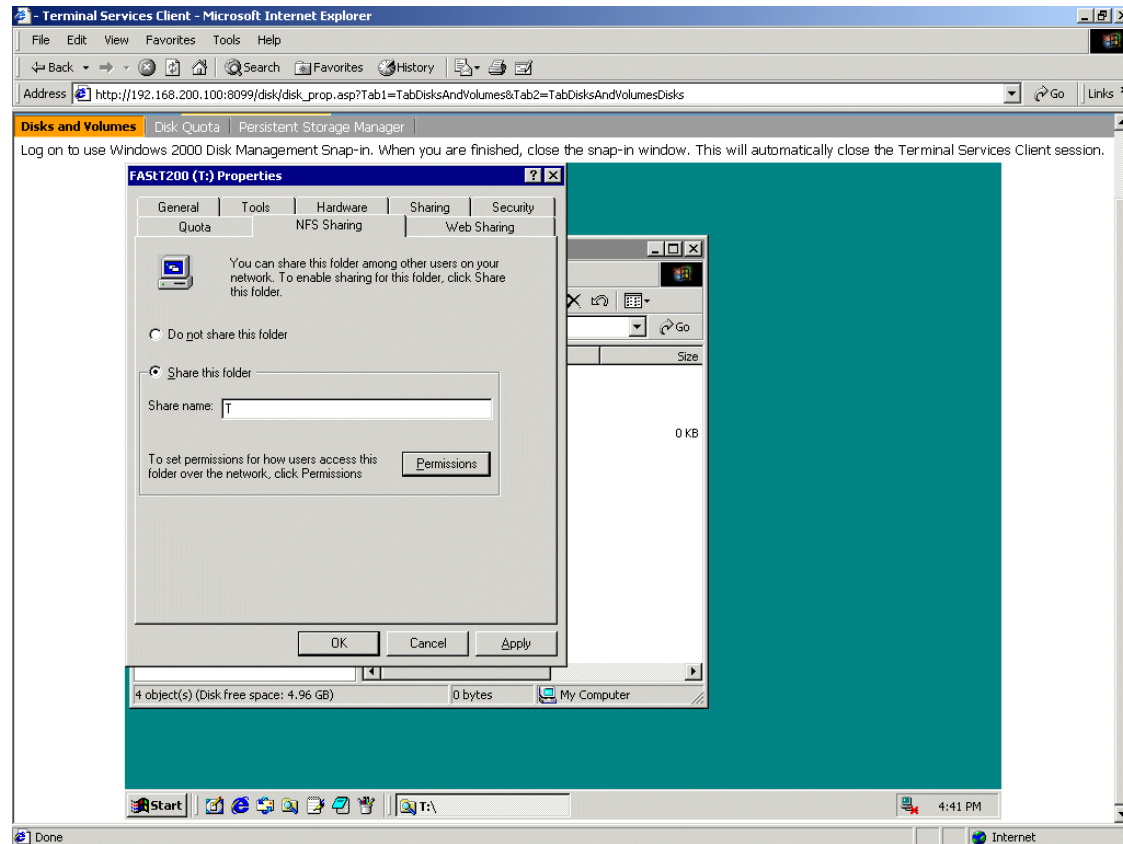


Figure 3-86 Permissions for the NFS share

Once again, we assigned a name to the share. We chose to use the same name as we used for the Windows clients. This conveniently allows the shared directory to be mapped/mounted in the same way from both UNIX and Windows clients. Since access permissions in Windows and UNIX are significantly different, however, we checked the **Permissions** dialog for NFS Sharing.

The default configuration is shown in Figure 3-87.

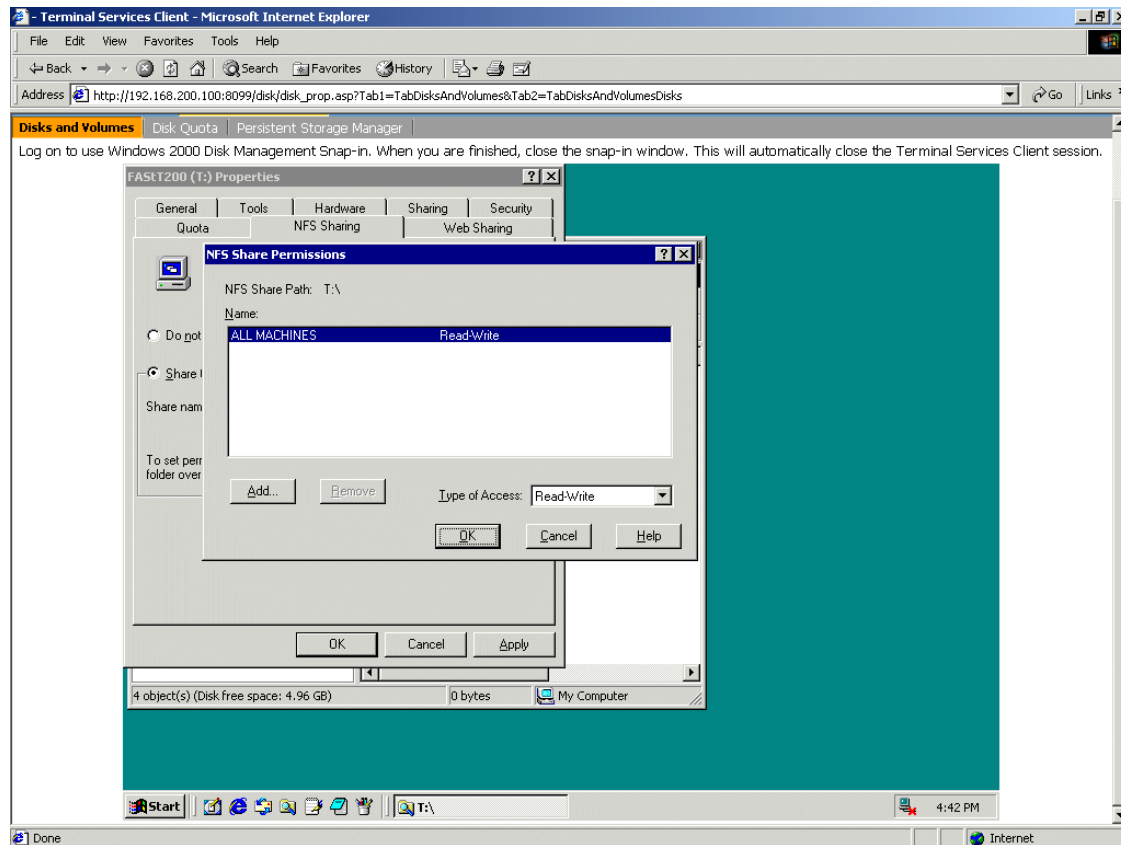


Figure 3-87 Setting share permissions for UNIX clients on the 300G

We chose to keep these settings because they allow all of our UNIX clients full access to the shared directory and any files it may contain. To be a little more picky about security on your systems, see “Setting up FTP access permissions on the 300G” on page 162 for more on this subject.

Now that we had named our shares and set up access permissions, we clicked **OK** to save our changes. Once the shares were ready, they were seen, as “handed out” in the Windows Explorer (Figure 3-88).

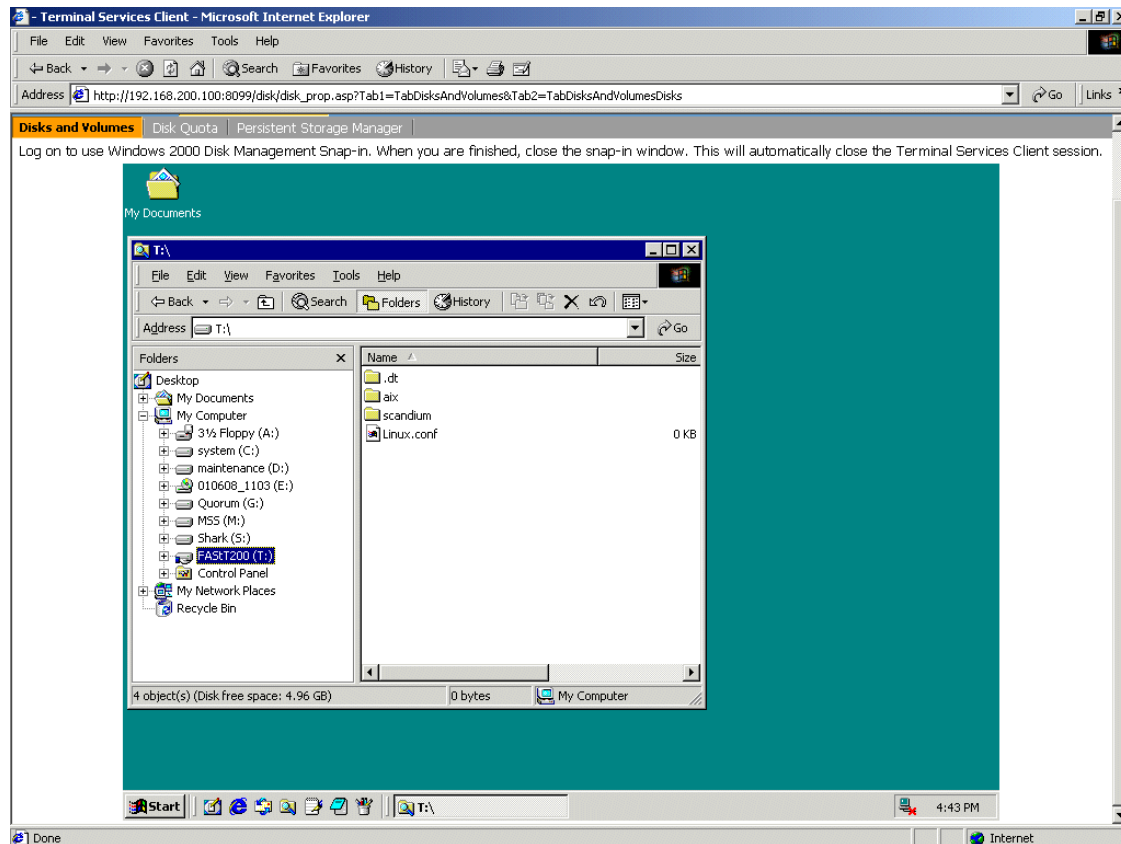


Figure 3-88 Shared directory

The 300G had access to the SAN-based storage, and we easily reached that storage from our LAN/WAN clients. All that remained was to map or mount the shared directory on our client machines.

3.7.3 Accessing the shares from our Windows clients

From Windows, accessing the share was extremely straightforward. We just went into the Network Neighborhood (or My Network Places, as Windows 2000 prefers to call it), drilled down to the 300G, supplied a user name and password, right-clicked the shared directory, and chose Map Network Drive (Figure 3-89).

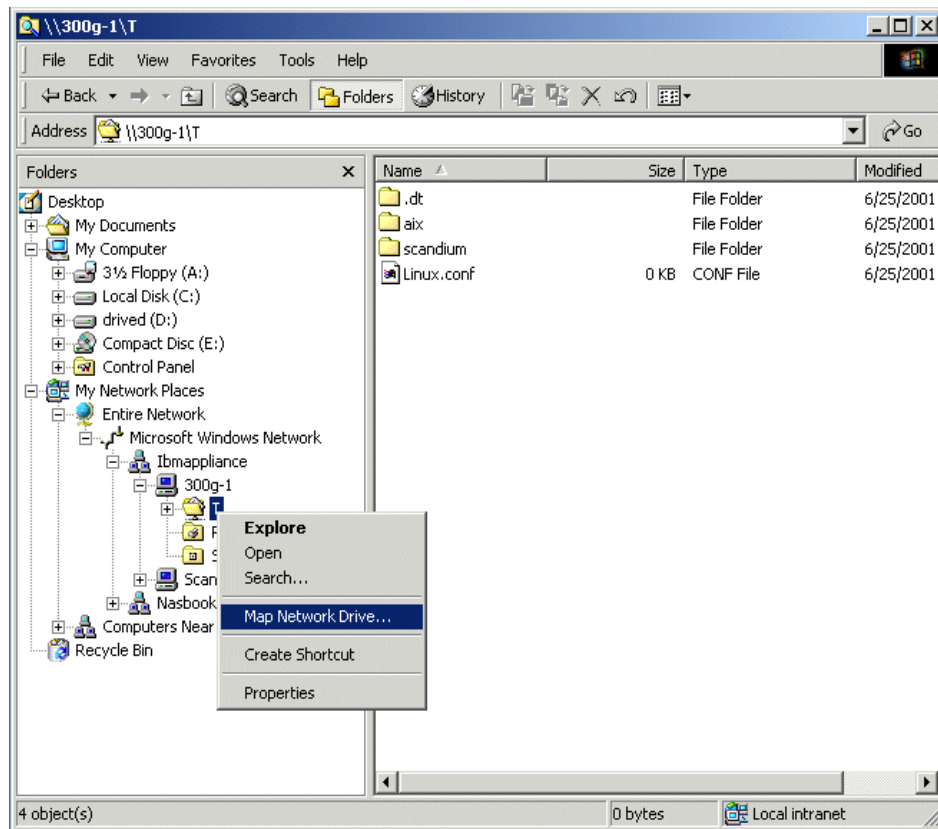


Figure 3-89 Map Network Drive

We were presented with a window requesting a drive, folder, and whether or not we wanted to reconnect at login (Figure 3-90).

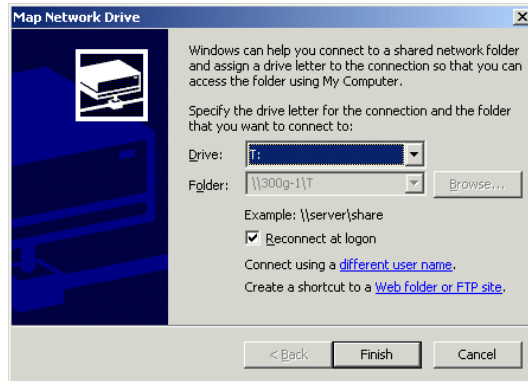


Figure 3-90 Windows mapping information

Once we supplied that information, we were able to see the shared disk, read from it, and write to it (Figure 3-91).

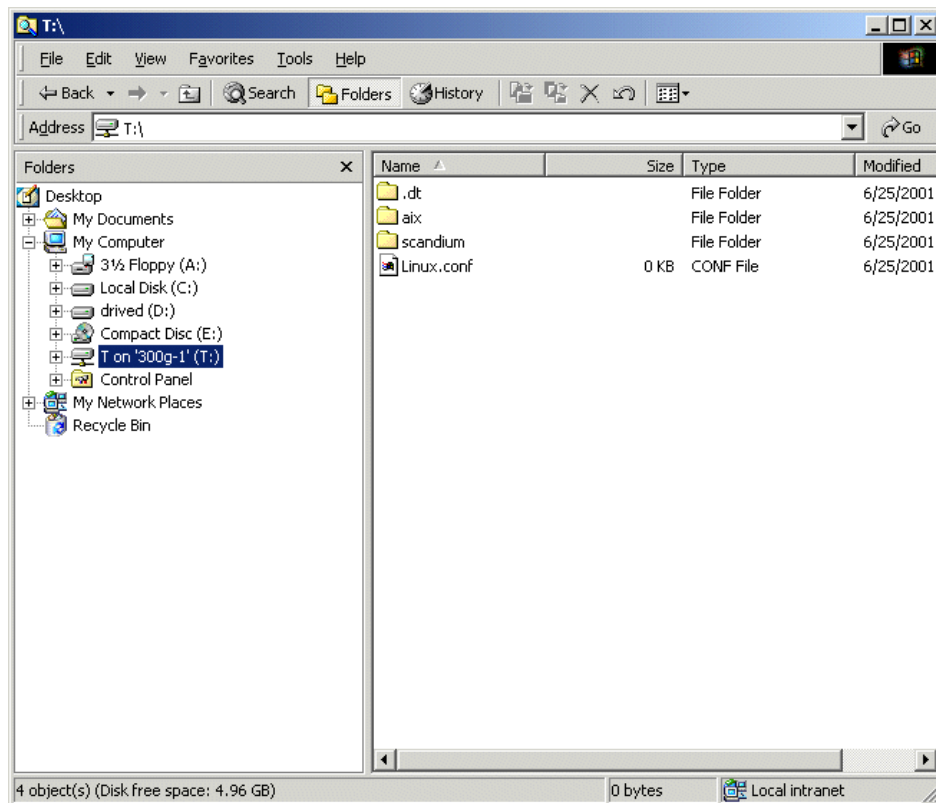


Figure 3-91 One of our shared drives

From the Windows 2000 client, Scandium, we were able to copy an 830 MB zip file onto the FASTT's partition in approximately 7 minutes even though Scandium was not connected to the SAN.

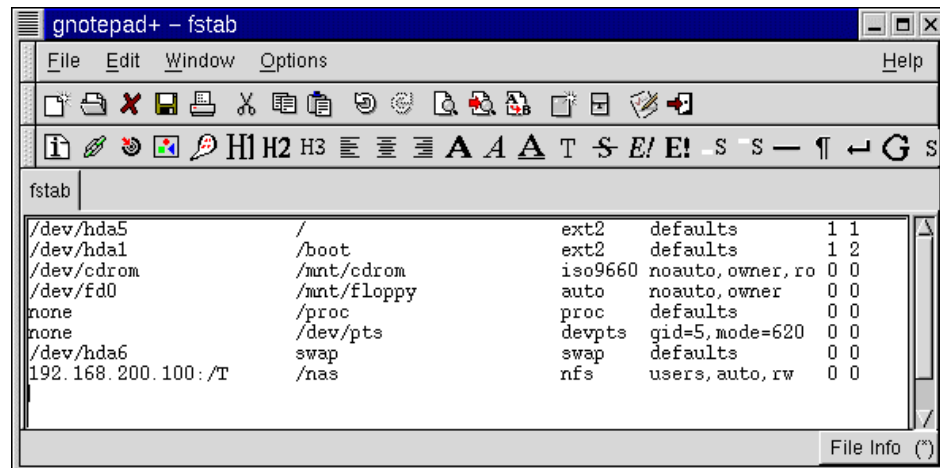
Remember: In Chapter 4, “Clustering for high availability” on page 167, we explain how to modify this file sharing environment to use Tivoli SANergy. To be compatible with SANergy, it is important that you not map your drives to the default administrative shares such as C\$. Define and use your own share name.

3.7.4 Accessing the shares from our Linux/Solaris/HP-UX clients

Connecting to the shared disks from our Linux client, Pagopago, was just as easy. First, we modified the `/etc/fstab` file to include a listing for the shared disk.

Note: Under Solaris, the `/etc/vfstab` file is updated rather than the `/etc/fstab`.

This is shown on the last line in Figure 3-92 below.

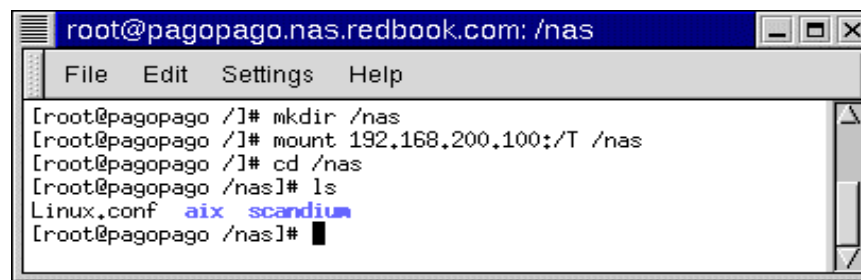


```
gnotepad+ - fstab
File Edit Window Options Help
[Icons]
[Icons]
fstab
/dev/hda5 / ext2 defaults 1 1
/dev/hda1 /boot ext2 defaults 1 2
/dev/cdrom /mnt/cdrom iso9660 noauto, owner, ro 0 0
/dev/fd0 /mnt/floppy auto noauto, owner 0 0
none /proc proc defaults 0 0
none /dev/pts devpts gid=5, mode=620 0 0
/dev/hda6 swap swap defaults 0 0
192.168.200.100: /T /nas nfs users, auto, rw 0 0
File Info (*)
```

Figure 3-92 Adding the 300G's shared disk to the Linux `fstab` file

Once that was done, we created a directory named `/nas` and mounted the shared directory to it normally using the `mount` command, as shown in Figure 3-93. We then changed directory to `/nas` and were immediately able to see all of the data on the shared disk.

Just for fun, we tried opening some of the files and creating new ones. This worked beautifully and we were able to see the changes from the Windows clients. As a further test, we created a text file from the Linux client, saved it, and left it open. We then tried accessing the file from one of the Windows clients. While we were able to open the file normally, we were pleased to note that Windows recognized that the file was still in use on the Linux system and did not let us overwrite it.

A terminal window titled "root@pagopago.nas.redbook.com: /nas" with a menu bar containing "File", "Edit", "Settings", and "Help". The terminal output shows the following commands and their results:

```
[root@pagopago /]# mkdir /nas
[root@pagopago /]# mount 192.168.200.100:/T /nas
[root@pagopago /]# cd /nas
[root@pagopago /nas]# ls
Linux.conf  aix  scandium
[root@pagopago /nas]#
```

Figure 3-93 Mounting the 300G's shared directory from a Linux client

With a minimum of effort, we were able to use the 300G to safely share SAN-based storage among heterogeneous LAN/WAN clients.

3.7.5 Accessing the shares from our AIX clients

Now we will show you the few commands we used to get AIX ready to use the 300G's shared disks.

First we needed to update /etc/filesystems. This was accomplished by using the `crfs` command, as shown in Example 3-3.

Example 3-3 Using the crfs command

```
# crfs -v nfs -m /FastT200 -n 300g-1 -d T -A yes
# cat /etc/filesystems
/FastT200:
    dev           = T
    vfs           = nfs
    nodename      = 300g-1
    mount         = true
    account       = false
```

Finally, we mounted the share, as shown in Example 3-4.

Example 3-4 Mounting the share

```
# mount -v nfs 300g-1:T /FastT200
# mount
```

node	mounted	mounted over	vfs	date	options
/dev/hd4	/		jfs	Jun 25 15:50	rw,log=/dev/hd8
/dev/hd2	/usr		jfs	Jun 25 15:50	rw,log=/dev/hd8
/dev/hd9var	/var		jfs	Jun 25 15:50	rw,log=/dev/hd8
/dev/hd3	/tmp		jfs	Jun 25 15:50	rw,log=/dev/hd8
/dev/hd1	/home		jfs	Jun 25 15:51	rw,log=/dev/hd8
/dev/lv00	/usr/welcome_arcade		jfs	Jun 25 15:51	
rw,log=/dev/hd8	/dev/lv01	/usr/welcome	jfs	Jun 25 15:51	rw,log=/dev/hd8
300g-1	T	/FastT200	nfs3	Jun 25 17:17	

3.7.6 Setting up FTP access permissions on the 300G

We breezed through setting up access permissions for UNIX clients in “File sharing for UNIX clients” on page 154. This section covers the concept in a little more detail.

By default, all machines have read/write permissions on the share, however it is possible to grant authority for individual machines. Use the **Add** and **Remove** buttons and select the appropriate **Type of Access**. Setting permissions on the user level requires more administrative overhead. To give UNIX users the appropriate permissions on the imported NTFS partition, the Windows system must identify the user and group names from the UNIX machines. Users in a Windows domain should be mapped with users on UNIX machines. To set this up, double-click the **IBM NAS Admin.msc** shortcut on the desktop of the 300G and click **File Systems -> UNIX Services**.

If your UNIX users are the same as your Windows users, then you can use the simple mapping option (Figure 3-94) to easily set up shared permissions.

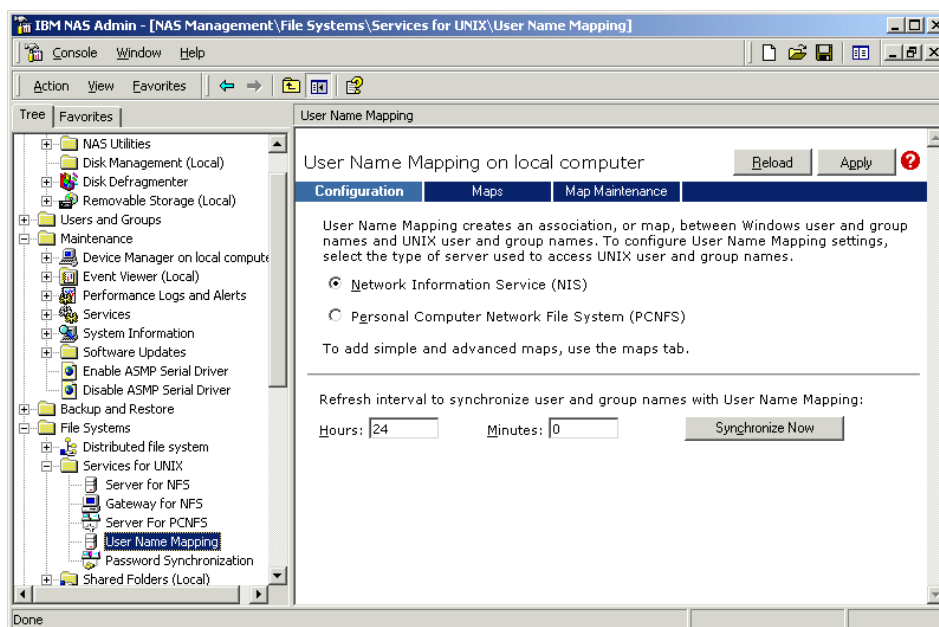


Figure 3-94 User mapping administration

Select the **User Name Mapping** option in the left hand panel, check the **Simple maps** box, and specify the name of your Windows domain (or the name of your MDC server if you are not in a domain) in the **Windows domain name** section.

If the users on both systems are not identical, you will need to use **Advanced maps** to map each user on the Windows side to the corresponding user on the UNIX side. In this case, it is necessary to have a NIS server (Network Information Service) running in your network. For further information on setting up user name mapping with NIS, see:

<http://www.microsoft.com/windows2000/sfu/>

3.8 User and security management on the 300G

This section describes integrating the 300G into a secure environment. We have skimmed over this subject pretty lightly so far, but in keeping with its role as an appliance, the 300G is designed to plug right into your existing user and security management system.

3.8.1 Active Directory, NT 4 Domains, and Workgroups

The 300G will integrate with all of the Microsoft Operating System versions that you have in your current network environment. It will work with any existing user and security management for those systems, including:

- ▶ Windows Workgroup Computing
- ▶ Windows NT 4 Domains
- ▶ Windows 2000 Active Directory (mixed and native mode)

Fully describing user and security management for Windows is beyond the scope of this book, so we will just provide you with a quick overview. For detailed information, please refer to the plethora of literature regarding Microsoft Operating Systems. Some examples are listed in “Related publications” on page 319.

To change the security environment for your 300G, from the desktop (or from a Windows Terminal Service session of your remote administrative console), right-click **My Network Places** and choose **Properties**. In the resulting window, select the **Advanced** menu and choose **Network Identification**.

Now you have the option to join the 300G to your environment, as shown in Figure 3-95.

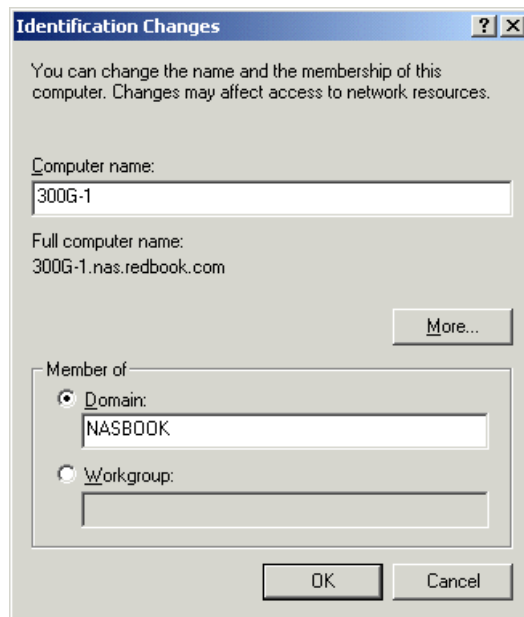


Figure 3-95 Setting up network identification for Active Directory

If you want the 300G to be a storage device for your workgroup, check **Workgroup** and type in the name of your workgroup. Please be aware that all security management within such a workgroup is local. This means you have to administer all user accounts on the 300G. Please also be aware that the default behavior when sharing a network drive with Windows is to grant all users full access to the data.

When you want to join an existing Windows NT 4.0 Domain check the **Domain** box and type in the name of your domain. Press **OK** to confirm your choice and you will soon be a happy member of an existing Windows NT 4.0 Domain,

Tip: When joining an existing NT4 Domain across subnets or via routed paths, define the PDC as the primary WINS server, even if the WINS service is not running on the PDC. This way, the joining client will find the PDC easily.

Joining an Active Directory tree is almost identical to joining an NT 4 Domain. Check **Domain** and type in the name of your Active Directory tree.

Important: When joining an Active Directory, it is essential that your TCP/IP configuration and DNS name resolution be working properly. Make sure both machines can ping each other using the IP address and the fully qualified domain name before joining the domain. For example, type:

```
ping 300g-1.nas.redbook.com
```

3.8.2 UNIX NIS integration

The UNIX Network Information System (NIS) services work like using the yellow pages. While the 300G's feature set includes support for NIS, the security standard of NIS is not very high. Therefore we do not recommend the use of NIS with this product.

The 300G comes preconfigured with Microsoft Services for UNIX 2.0. Within these services you have an NIS migration wizard. This tool allows you to migrate an NIS. The tool takes your NIS source files and migrates them into Active Directory.

The Server for NIS feature allows a Windows Domain controller to be an NIS master server or an NIS subordinate (slave) by integrating NIS into Active Directory. When using the NIS server as a slave, the NIS master server must be a Windows 2000 Server.

For detailed information, please check the following Web site:

<http://www.microsoft.com/WINDOWS2000/sfu/default.asp>

3.8.3 Password synchronization

Another tool that is included within the Microsoft Services for UNIX 2.0 is a password synchronization tool (2-way). It allows you to synchronize password changes between Windows NT or Windows 2000 and UNIX. Pre-compiled single sign-on demons are available for:

- ▶ IBM AIX 4.3+
- ▶ Linux (Red Hat 5.2 and 6.0)
- ▶ Sun Solaris 2.6+
- ▶ HP-UX 10.3+
- ▶ Compaq Tru64 UNIX

Note: Even if your UNIX version is not on the list — it may still work. Microsoft provides the source code for the password synchronization tool.



Clustering for high availability

The IBM TotalStorage NAS 300G model G25 enables you to cluster to maintain high availability. Clustering the 300G is similar to clustering in a client/server environment. Two or more 300G model G25 nodes are connected together via an Ethernet cross-over cable. By connecting each of the nodes, this cable acts as a lifeline of the cluster and carries the heartbeat, continually checking to see if all nodes in the cluster are functional.

This chapter guides you through the configuration and setup of a dual-node 300G cluster using Microsoft's Cluster Server (MSCS) software, which is included with the 300G model G25. Our environment is depicted in Figure 4-1. In a slight departure from the release notes, we chose to start by setting up the second node. While logically backwards, starting with the second node eliminates one reboot and makes the whole process easier.

Our goal in this chapter is to flesh out the release notes and clarify a few steps. By following our procedures carefully, you should be able to implement the sometimes dreadfully complicated cluster service, relatively easily and painlessly. Pleasant clustering!

4.1 Our environment

We implemented a dual node cluster of a NAS 300G model G25. Figure 4-1 is a graphical representation of our setup.

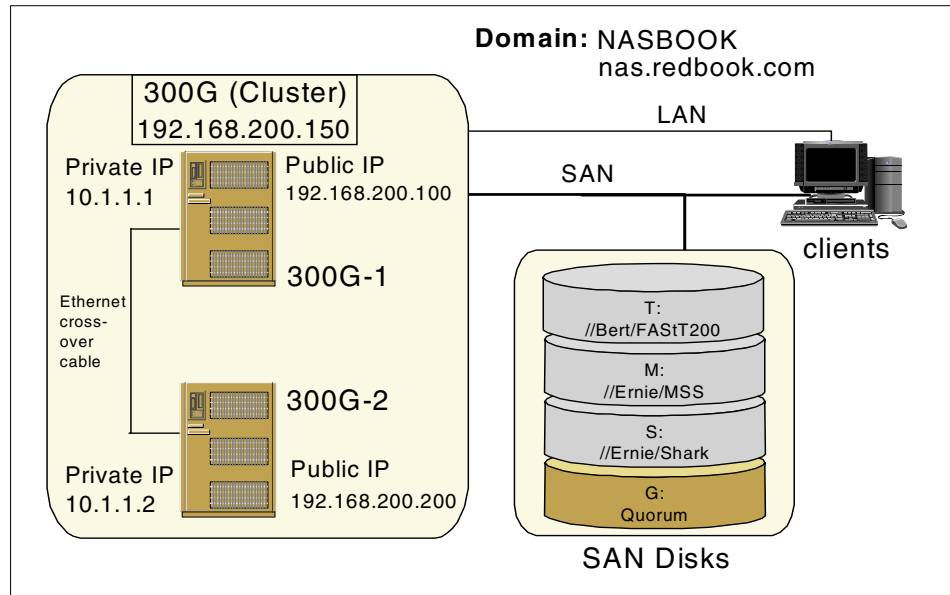


Figure 4-1 Our clustered environment

Note: We explain these procedures from the point of view of having a monitor and keyboard hooked up to both nodes (just because it is simpler and makes screen grabs a lot more readable). If, in reality, you are doing this setup using the remote management interface, we recommend using the Advanced Windows Terminal Service Client so that your screens and procedures will exactly match ours. Using the Windows Terminal Service client is explained in “Remotely managing the 300G” on page 73.

4.2 Second node first-time setup

This section describes the initial setup of the second node in our cluster.

Important: Ensure that the other node is powered off. Otherwise, you risk corrupting data on shared volumes if both nodes should try to simultaneously write to it.

4.2.1 Configure the private network adapter

From the Network and Dial-up Connections window, right-click the Local Area Connection icon and select **Properties**, as shown in Figure 4-2.

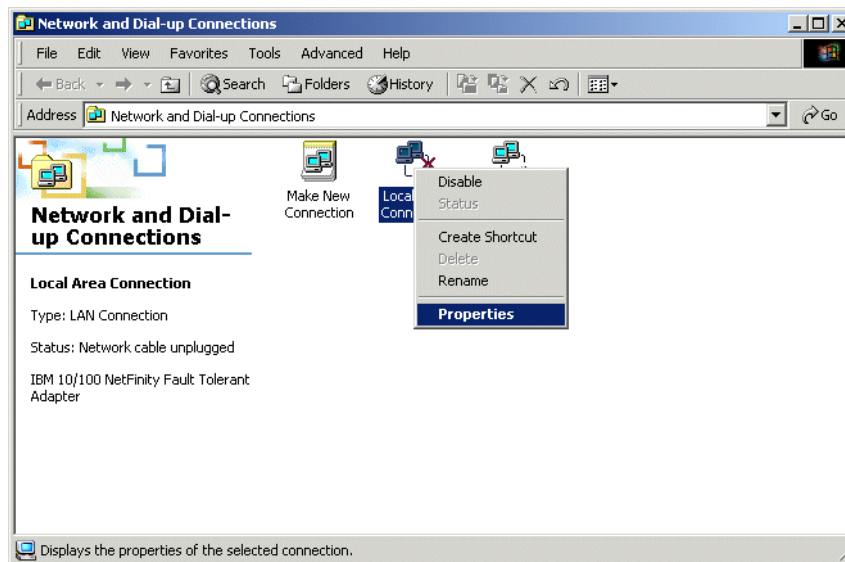


Figure 4-2 LAN connection properties selection

This brings up the Local Area Connection Properties page depicted in Figure 4-3. The private network uses the IBM 10/100 NetFinity Fault Tolerant Adapter, so if that is not the name of the adapter you have open, select **Cancel** and open the properties page on the next Local Area Connection icon (and repeat until you find the right one).

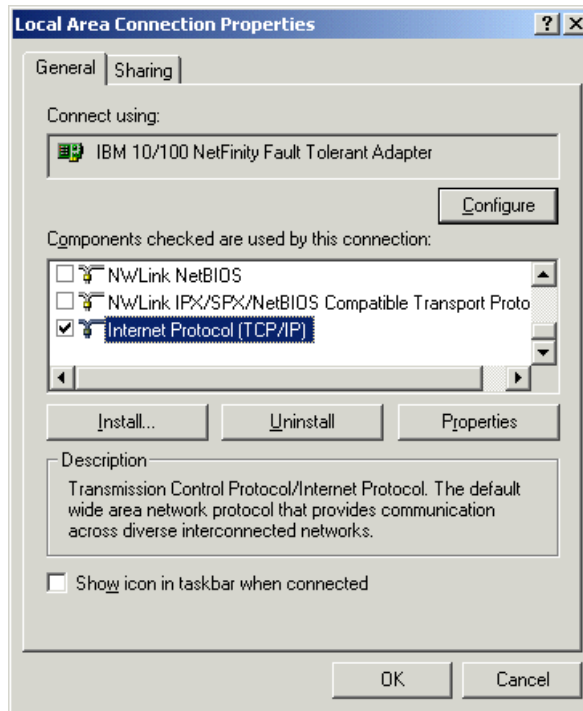


Figure 4-3 LAN connection properties page

From the General page, press the **Configure** button just below the adapter name. This displays the adapter properties page. From here, we click on the **Advanced** tab. We first verify that the External PHY property is set to 100Mbps Full Duplex, as shown in Figure 4-4.

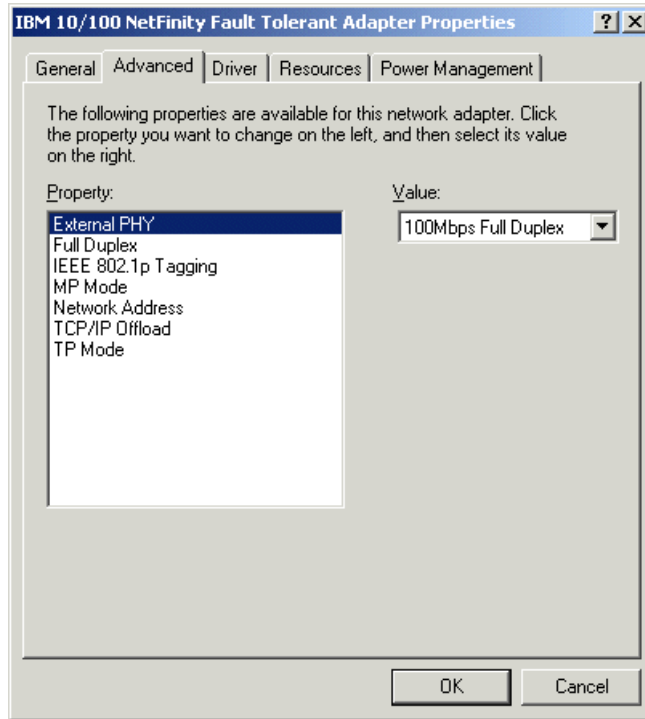


Figure 4-4 Verify External PHY property

If the value is not 100Mbps Full Duplex, then select the value from the drop-down list. We next verify the Full Duplex property is UTP - Full Duplex (Figure 4-5).

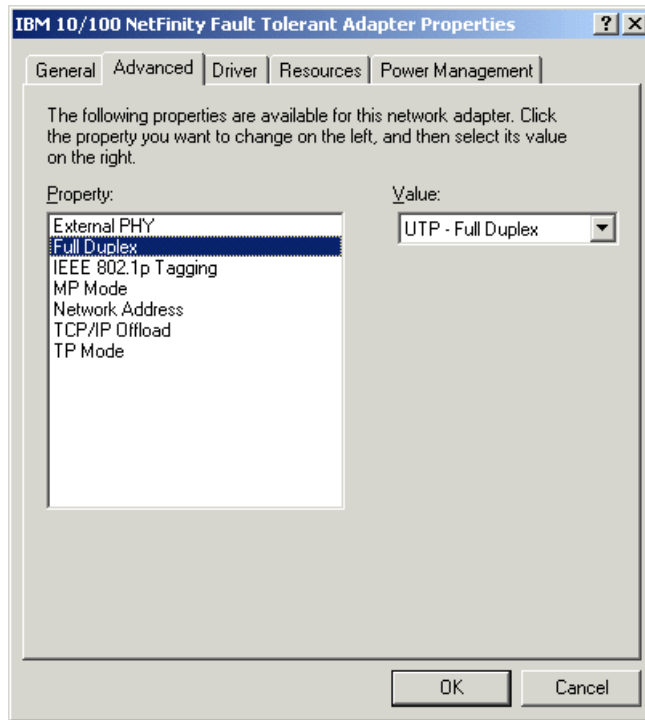


Figure 4-5 Verify Full Duplex property

Again, if the value is not UTP - Full Duplex, select it from the drop-down list. Finally, we verify the IP mode is On (Default), as shown in Figure 4-6.

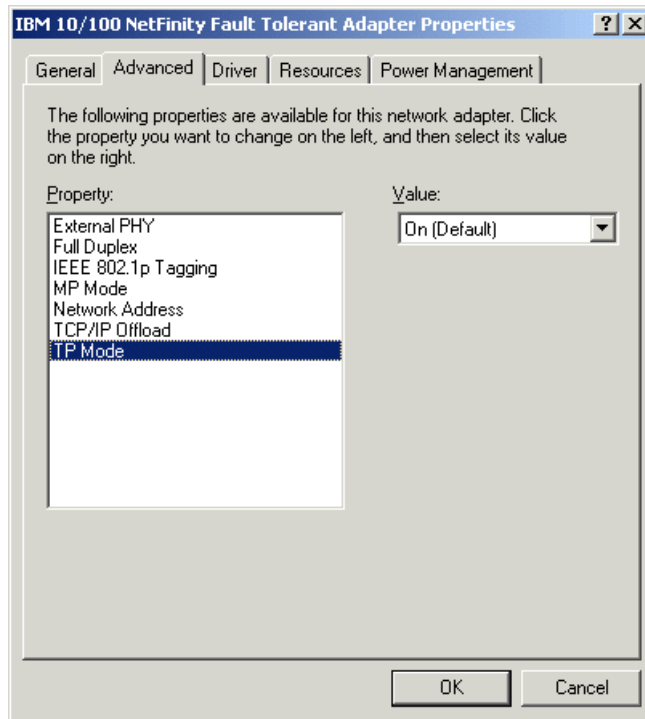


Figure 4-6 Verify IP Mode property

If the value is incorrect, again use the drop-down list to select the correct value. Click **OK**. This brings us back to the Local Area Connection Properties page (refer to Figure 4-3 on page 170).

Now it is time to configure the TCP/IP settings. Select **Internet Protocol (TCP/IP)** in the component list and click the **Properties** button (or just double-click the **Internet Protocol (TCP/IP)** selection in the component list). This displays the properties page shown in Figure 4-7. Since this is actually the Node 2 in the cluster, we recommend that you set the joining node's IP address to 10.1.1.2 and use the subnet mask of 255.0.0.0.

Note: The IP address 10.1.1.2 and the subnet 255.0.0.0 are the default settings for the 300G, so you should not have to modify these on Node 2.

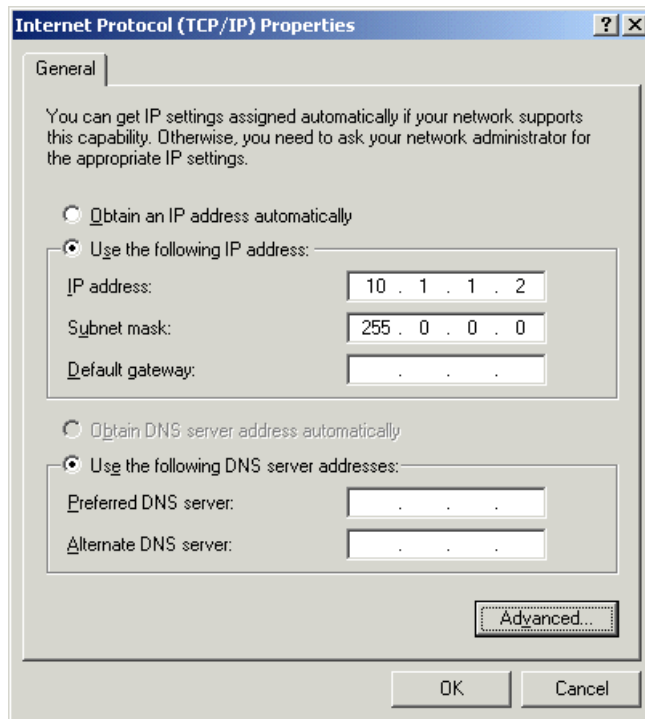


Figure 4-7 Configure private network IP settings

Once the IP address and subnet mask are entered, we click the **Advanced** button and then select **Disable NetBIOS over TCP/IP** on our private network (Example 4-8).

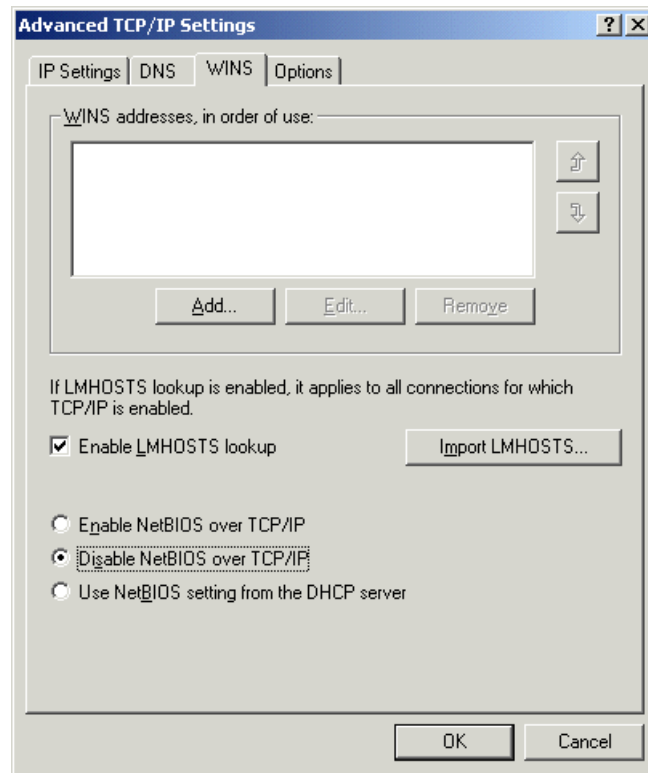


Figure 4-8 Disable NetBIOS over TCP/IP

Once you have done this, click **OK**. A message window will pop up and warn you that this connection has an empty primary WINS address (see Figure 4-9). This is not a problem, because a WINS server is not used for name resolution in our test environment. So just click **Yes** to close this dialog box.

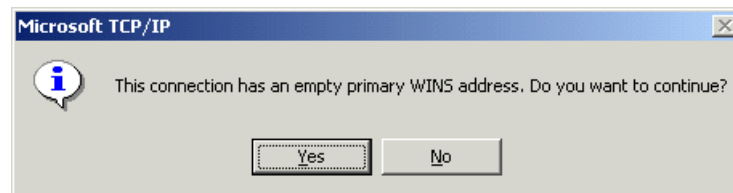


Figure 4-9 Empty primary WINS address confirmation

Click **OK** on the Internet Protocol (TCP/IP) Properties panel and then again on the Local Area Connection panel. Now rename this network connection *Private* and the other connection *Public* (Figure 4-10).

Note: *Private* and *Public* are not required names. You can use anything you like (perhaps *Heartbeat* and *LAN* make more sense to you). It is important that you rename them to something more meaningful than *Local Area Connection 1*, however, because later on, the Cluster Configuration wizard will ask you to tell it which connection to use for what. A meaningful name will help to prevent your choosing the wrong one.

Now we must ensure that our Private network is the first in line to be accessed by all network services. From the Network and Dial-up Connections window, shown in Figure 4-10, select **Advanced -> Advanced Settings...**

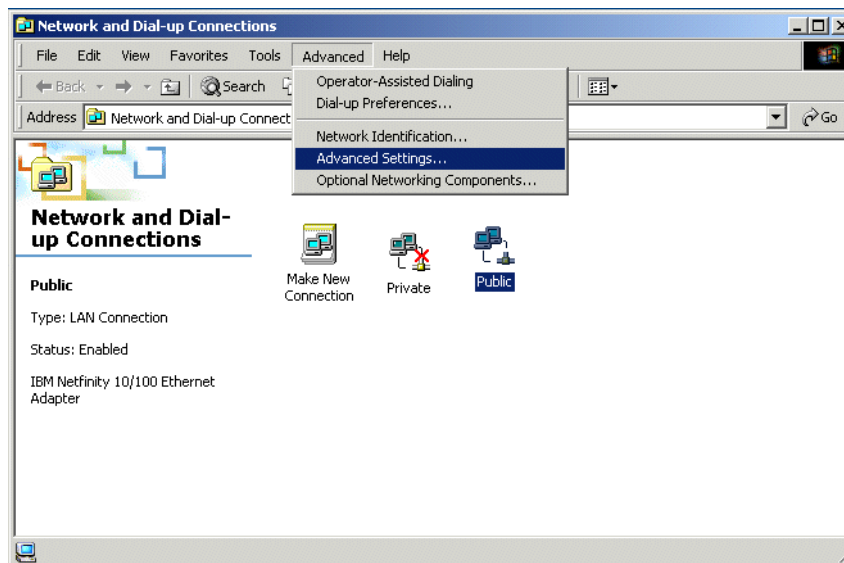


Figure 4-10 Network connections advanced settings selection

Ensure that the Private network connection is the first displayed in the Connections window list, as shown in Figure 4-11. If it is not, use the up arrow on the right side of the Connections list window to move it to the top.

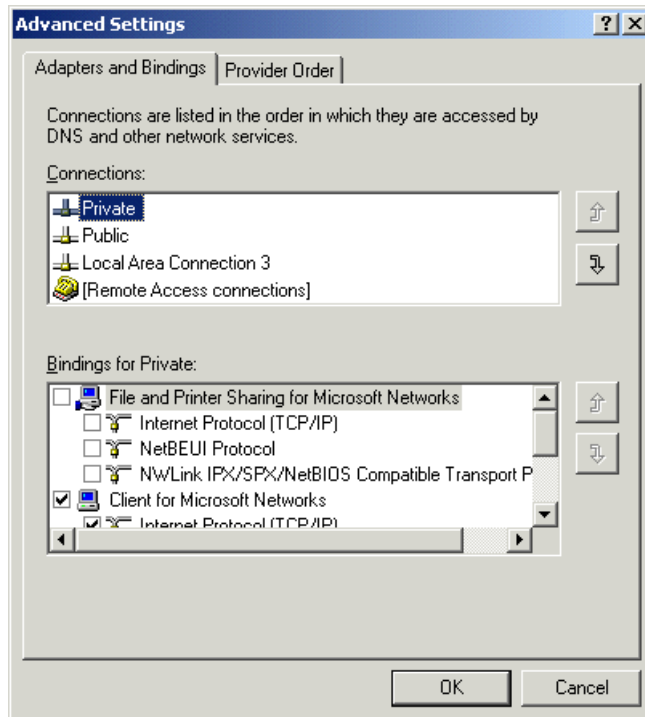


Figure 4-11 Connections ordering

If we were to select **OK**, we would be prompted to reboot now (Figure 4-12). However, we have a few more changes to make before we power off the second node, so we recommend that you answer **No**.

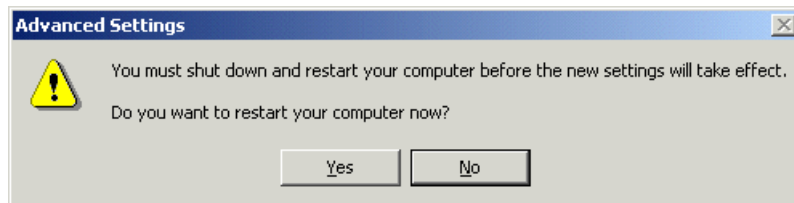


Figure 4-12 Just say **No** to a reboot

4.2.2 Joining the domain

Now we need to change our node name. From the desktop, right-click **My Computer** and select **Properties**, as shown in Figure 4-13.

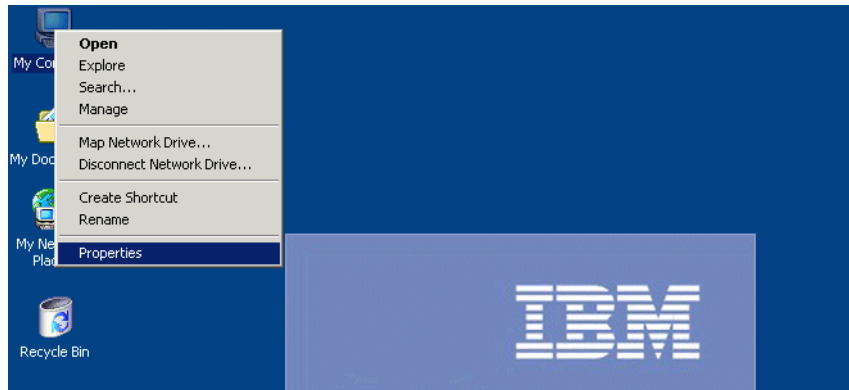


Figure 4-13 My Computer properties selection

After the System Properties panel opens, select the **Network Identification** tab, as shown in Figure 4-14.

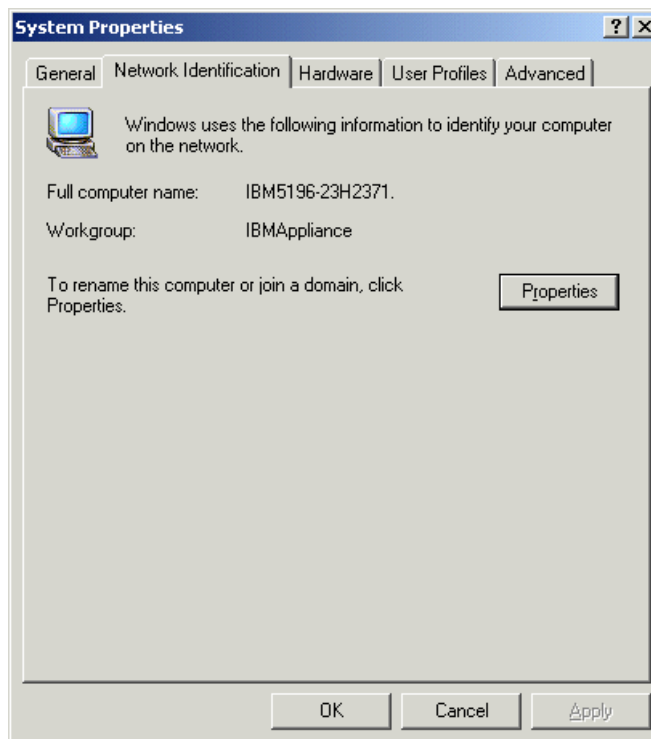


Figure 4-14 System properties network identification

Now click the **Properties** button to change the computer name to something more meaningful. While you are here, go ahead and click the **Domain** radio button and type in the name of the domain your cluster will be a member of (Figure 4-15). This is an important step, since the cluster will use the domain security policies for administering the file shares and other resources it controls.

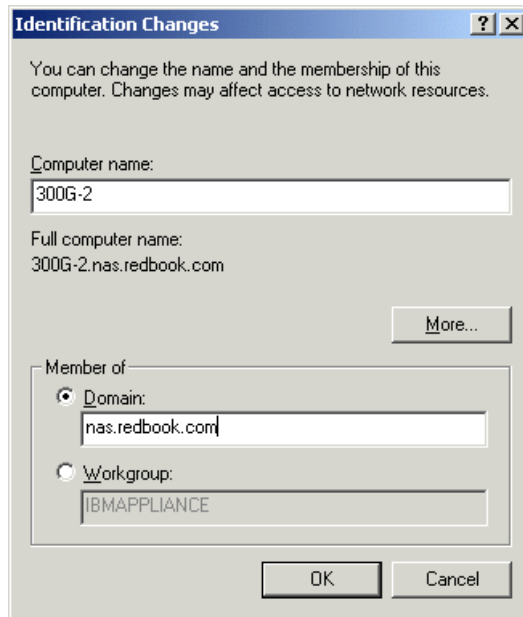


Figure 4-15 Change the computer name and update domain

You will be prompted, as we were, to enter in the user name and password of a domain administrator to ensure that not just anyone is joining your domain (Figure 4-16).



Figure 4-16 Enter name and password panel

You should now see a welcome screen similar to Figure 4-17. (If you get an error message back, you probably just mistyped either the user name or password, so try it again, more carefully!) Now keep going; we are almost done with this part.

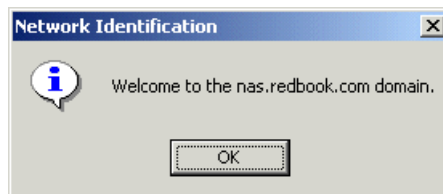


Figure 4-17 Node joined the domain successfully

4.2.3 Update drive letters

Take a quick look at the storage view (Figure 4-18) by double-clicking the **IBM NAS Admin.msc** icon on the desktop and then clicking **Storage -> Disk Management (Local)**. If you do not like the way the drive letters were assigned automatically, just change them to suit your needs.

Important: Whether or not you change the drive letters for your storage, make absolutely sure that the other node uses the *same drive letters* as this one. If you just leave it up to Windows, the letters probably will not match, and that will create access problems for your client machines if the primary node ever becomes unavailable.

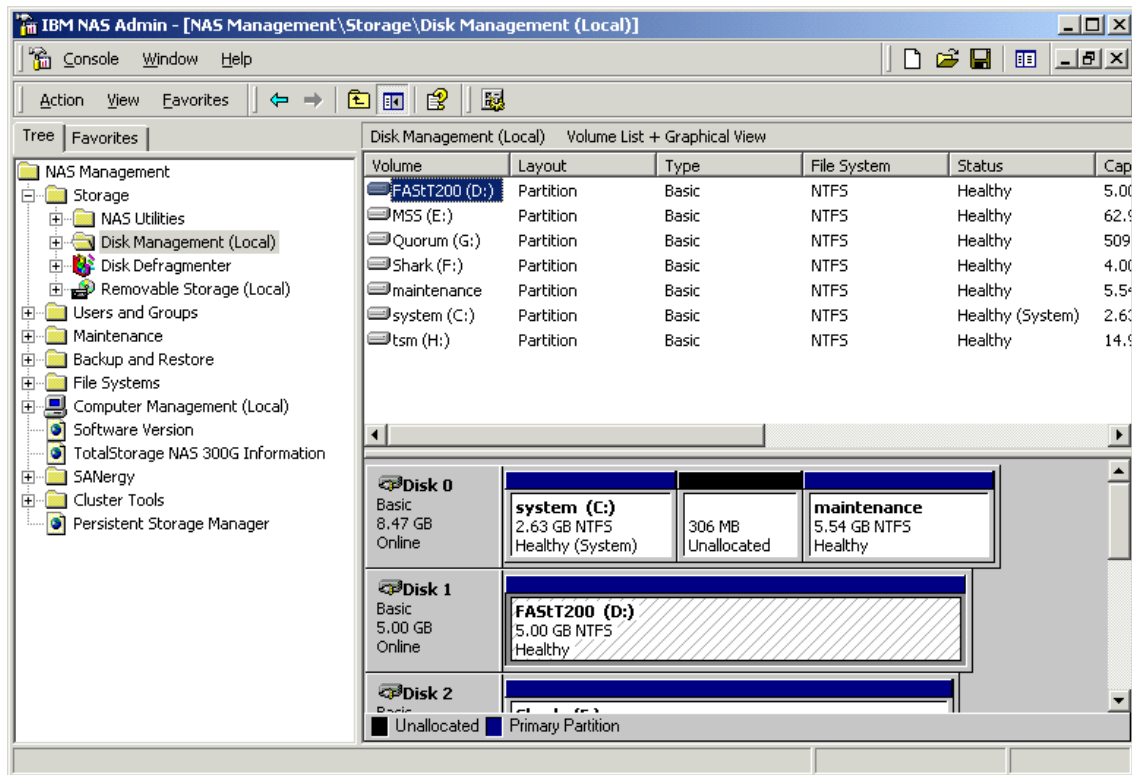


Figure 4-18 Disk storage view

Right-click on a disk you want to change (in either the top or bottom pane) and select the **Change Drive Letter and Path...** option, as shown in Figure 4-20.

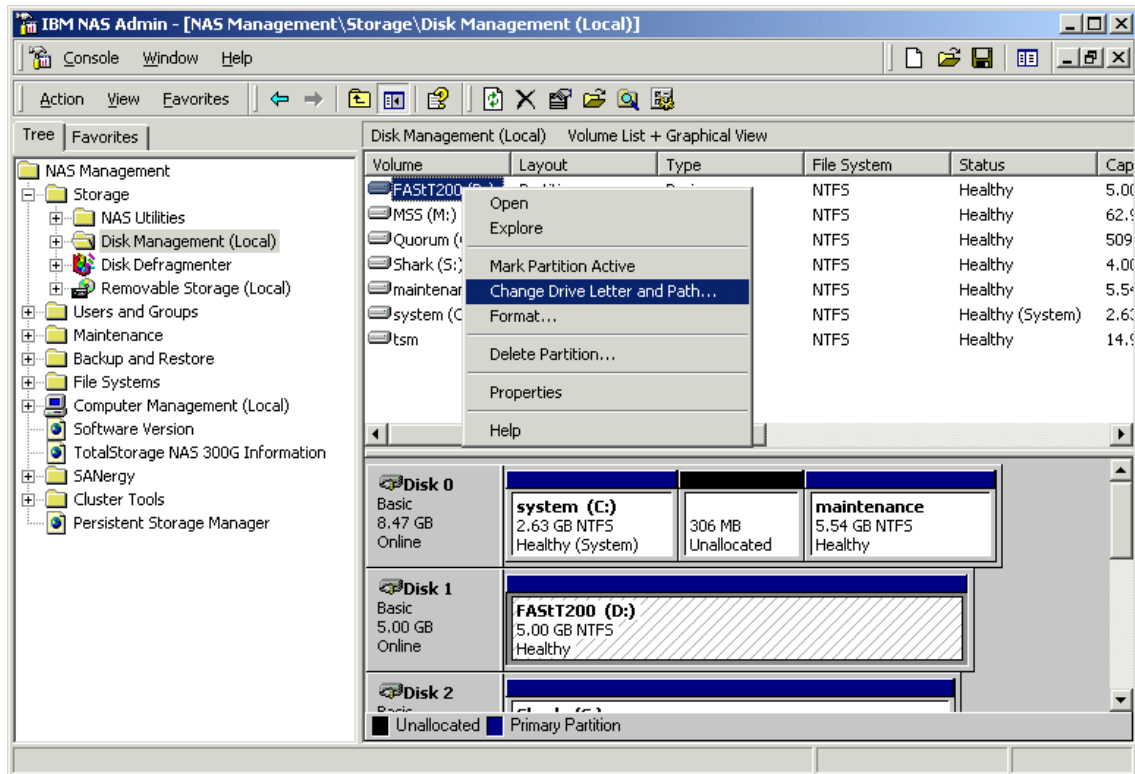


Figure 4-19 Change Drive Letter selection

The Change Drive Letter and Paths dialog (Figure 4-20) shows the current drive letter assigned.

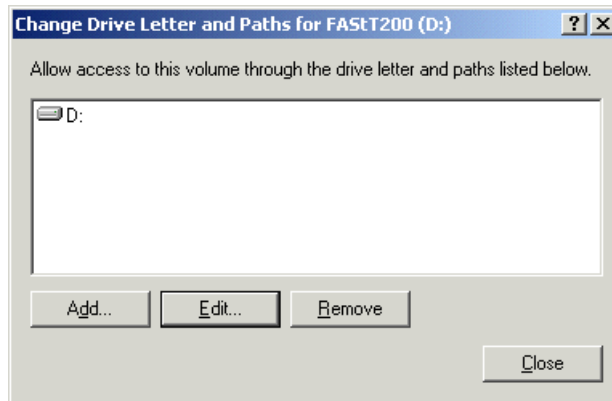


Figure 4-20 Drive letter display

Click the **Edit...** button to display Figure 4-21.

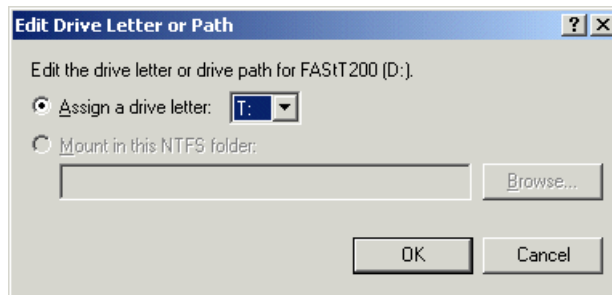


Figure 4-21 Assign new drive letter

Use the pull-down menu to select a new drive letter. We chose **T:** as our drive letter of choice for our FAST200 disk. Once you have settled on a drive letter, just select **OK**. A confirmation window will then appear (Figure 4-22).

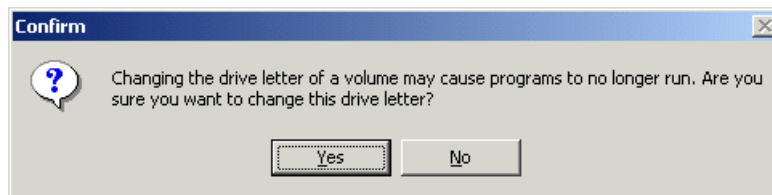


Figure 4-22 Change drive letter confirmation

Your drive letter is now changed. It is that simple! Now repeat this process for every disk you want to change and make a note of what volumes are assigned to what letter. You will need this information in a minute when you start configuring Node 1. Once you have updated the drive letters for each of the disks, you are ready to continue.

4.2.4 Shut down the second node

Now that the initial setup of our second node is complete, we can shut it down (Figure 4-23) and leave it powered off until we are ready to bring it into the cluster.



Figure 4-23 Shut down

This completes the initial setup of Node 2 (which we creatively named 300G-2) for clustering.

4.3 First node first-time setup

We now begin the initial setup for Node 1. The first few steps are, with minor exceptions, identical to the ones we performed on Node 2, so please bear with the repetition here!

4.3.1 Configure private network adapter

Configure the private network adapter just like we did for the second node in Section 4.2.1, “Configure the private network adapter” on page 169.

Verify that the settings shown in Table 4-1 are set to the correct values.

Table 4-1 Adapter properties

Property	Value
External PHY	100Mbps Full Duplex
Full Duplex	UPT - Full Duplex
IP mode	On (Default)

If any of the values are incorrect, use the drop-down list to select the correct value. Once done, press **OK**. This brings us back to the Local Area Connection Properties page (refer to Figure 4-3 on page 170).

We recommend that you set the IP address of Node 1 (the primary node) to 10.1.1.1 and use the subnet mask of 255.0.0.0, as shown in Figure 4-24.

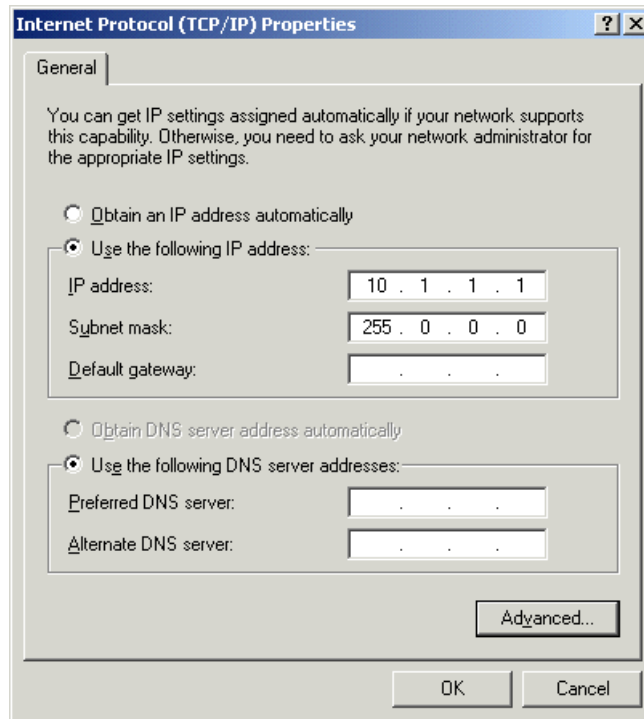


Figure 4-24 Node 1 private network IP properties

Other than the IP address, everything should be the same here as it was on Node 2. When you have finished making these changes, you will be prompted to reboot. Once again, we recommend you postpone this for a little while and continue configuring Node 1.

4.3.2 Joining the domain

This procedure is exactly the same as the one given in “Joining the domain” on page 177, except that you should give Node 1 a different name than you gave Node 2. (We used 300G-2 and 300G-1.) When you have finished, the reboot information screen will pop up. Click **OK**, but do not reboot yet.

4.3.3 Update drive letters

This procedure is identical to Section 4.2.3, “Update drive letters” on page 181. Remember to assign exactly the same drive letters on Node 1 as you assigned on Node 2. Once we updated all of our drive letters, our storage view looks like Figure 4-25.

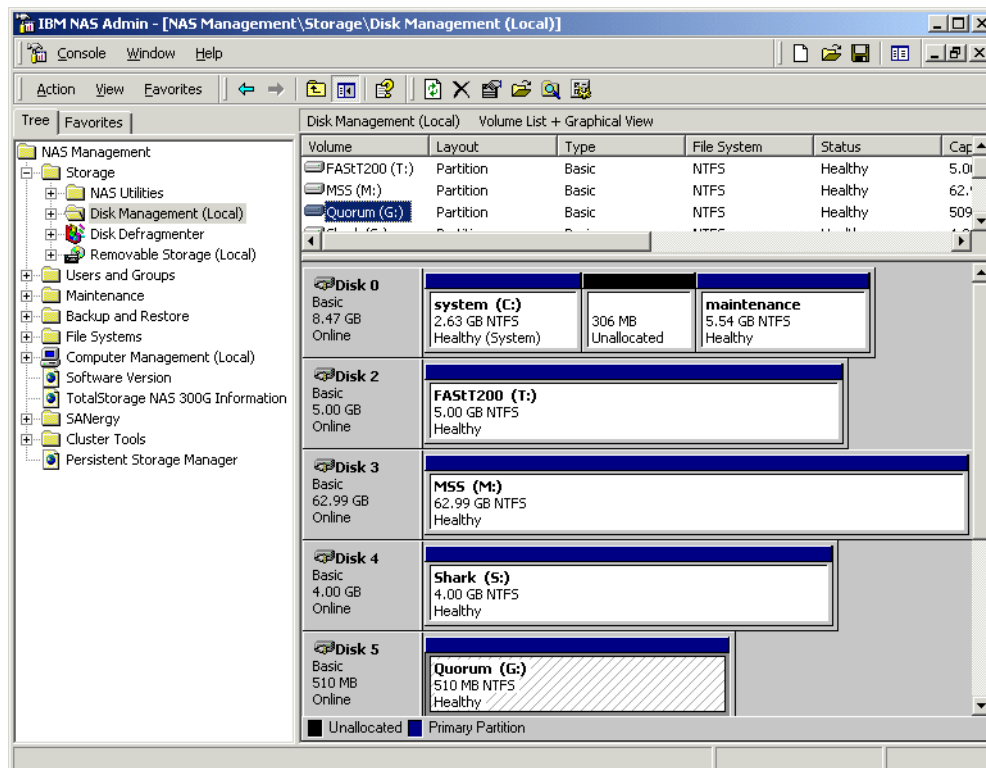


Figure 4-25 Disk storage view

4.3.4 Restart first node

So far, we have been postponing the reboot, but we must now restart Node 1 before proceeding.



Figure 4-26 Restart Node 1

4.3.5 Cluster setup

Open the IBM NAS Admin snap-in using the icon on the desktop. Click **Cluster Tools** -> **Cluster Setup** in the left pane (Figure 4-27).

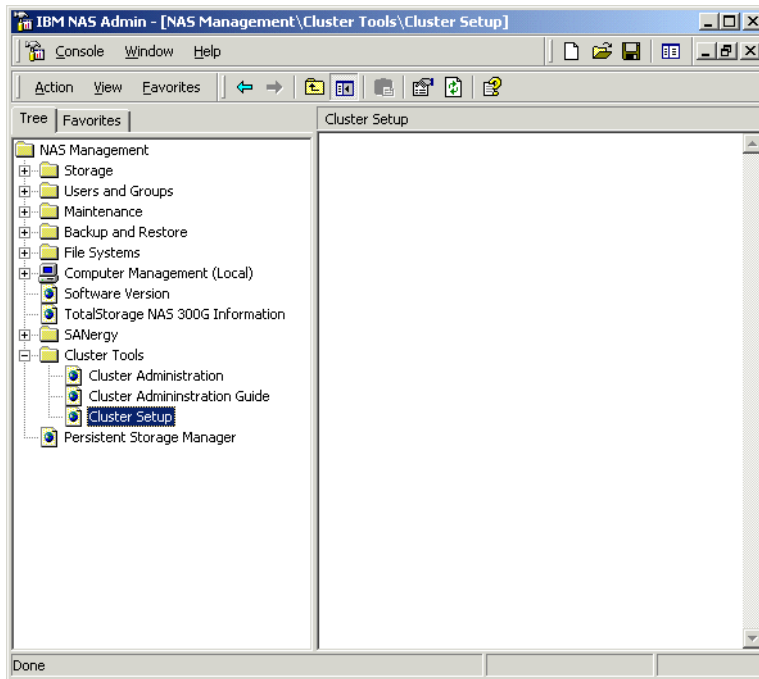


Figure 4-27 Cluster setup selection

This will launch the **TotalStorage Cluster Configuration Wizard** shown in Figure 4-28.



Figure 4-28 Cluster Configuration Wizard

Click **Continue**. We are now asked if this is the first node in the cluster or a joining node. Select **First Node** (Figure 4-29).

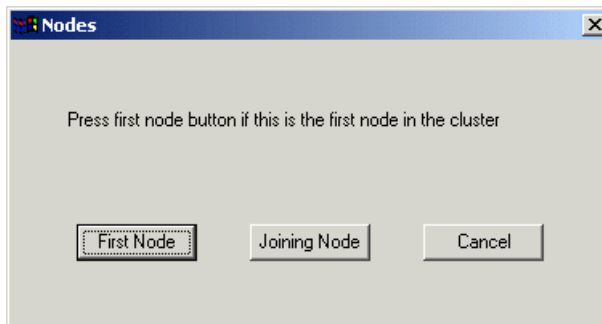
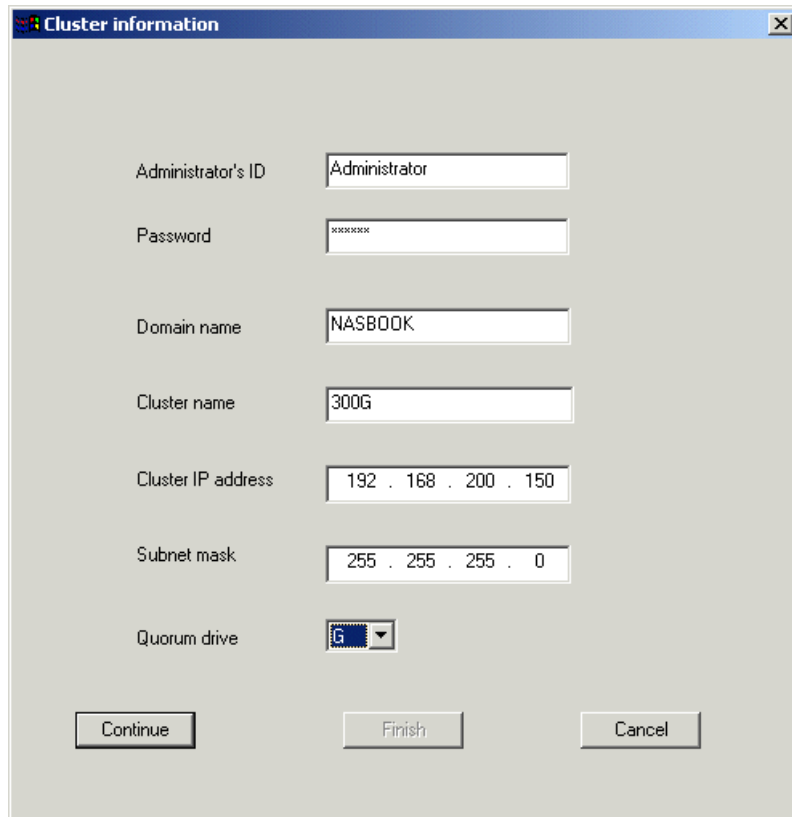


Figure 4-29 Select node

The next window (Figure 4-30) gathers the cluster information.



Cluster information

Administrator's ID Administrator

Password xxxxxxx

Domain name NASBOOK

Cluster name 300G

Cluster IP address 192 . 168 . 200 . 150

Subnet mask 255 . 255 . 255 . 0

Quorum drive G

Continue Finish Cancel

Figure 4-30 Cluster information

Fill in the fields with the appropriate information and click **Continue**.

Important: The current version of the release notes recommends the Quorum drive be Q: We recommend you use G: instead. However, the 300G does not really care.

A funny thing happens here. The next window displayed asks if the information you entered is correct (Figure 4-31). Unfortunately, it also hides the information from view, but you can drag this little pop-up window out of the way. Go ahead and double-check your entries just to be sure.

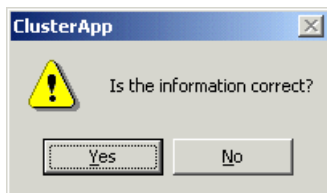


Figure 4-31 Cluster information confirmation

Selecting **Yes** brings the cluster information panel back to the foreground, but now we have to click the **Finish** button in order to save the cluster information settings.

The initial setup of Node 1 (300G-1) is now complete.

We have now successfully created our cluster, even though so far it only has one node in it, as shown in Figure 4-32. It is now time to assimilate the second node into the cluster.

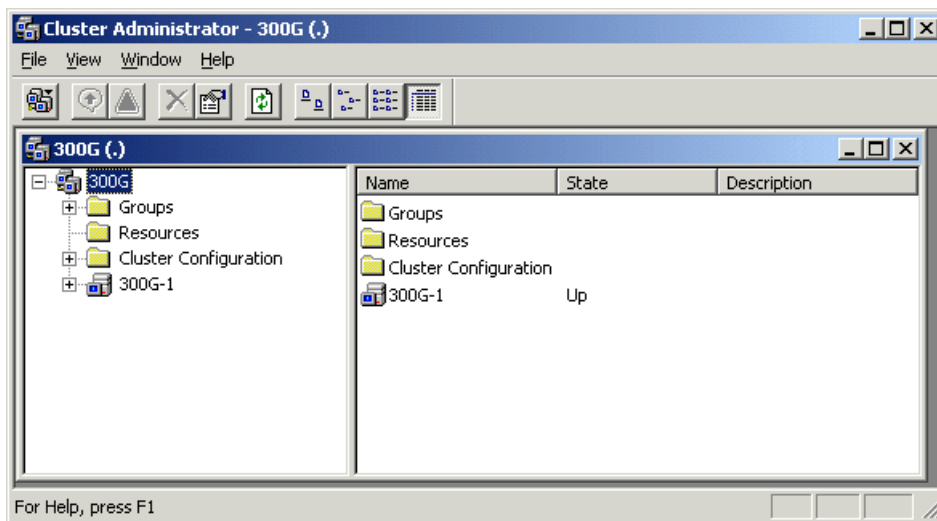


Figure 4-32 Cluster administrator window

Note: Do not power Node 1 off. You can leave it up, because it needs to be running in order for Node 2 to join the cluster.

4.4 Second node second-time setup

It is now time to power-on the second node and add it to the cluster.

4.4.1 Add the second node to the cluster

Once the system is powered on, open the **IBM NAS Admin** application and click **Cluster Tools -> Cluster Setup** in the left pane to begin configuring this node to join our cluster (Figure 4-33).

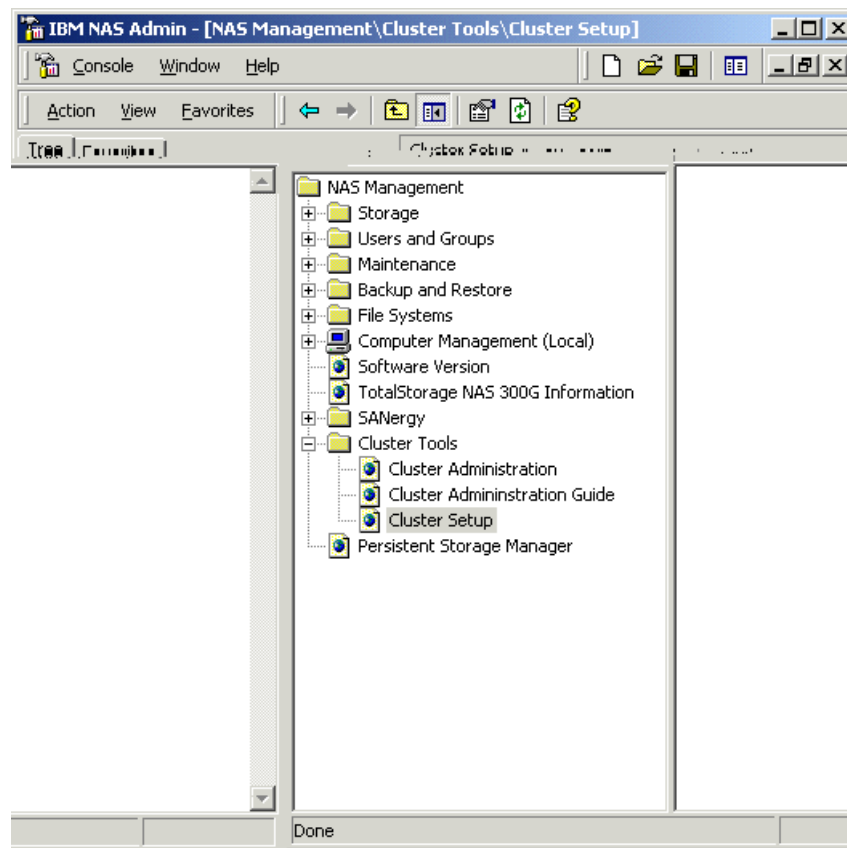


Figure 4-33 Cluster setup selection

This starts the **TotalStorage Cluster Configuration Wizard**. Click **Continue**.

This time, when we are asked which node we are adding, we click the **Joining node** button, as shown in Figure 4-34.

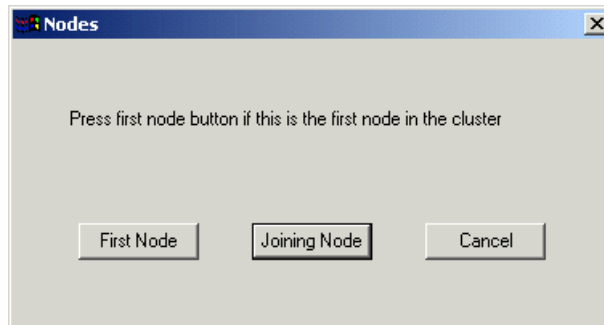


Figure 4-34 Joining node selection

We are then prompted to enter the name of the first node in the cluster. We named Node 1 **300G-1** so that is what we typed in (Figure 4-35).

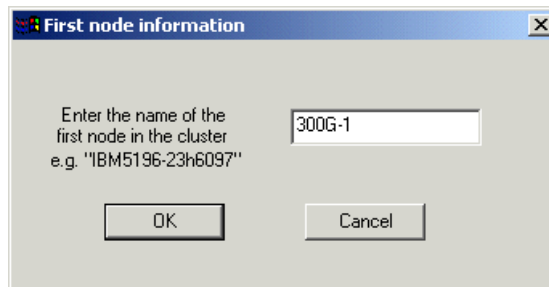


Figure 4-35 First node information

Then select **OK**. This is all that is required to add the second node! You can see in Figure 4-36 that both nodes report a status of 'Up'.

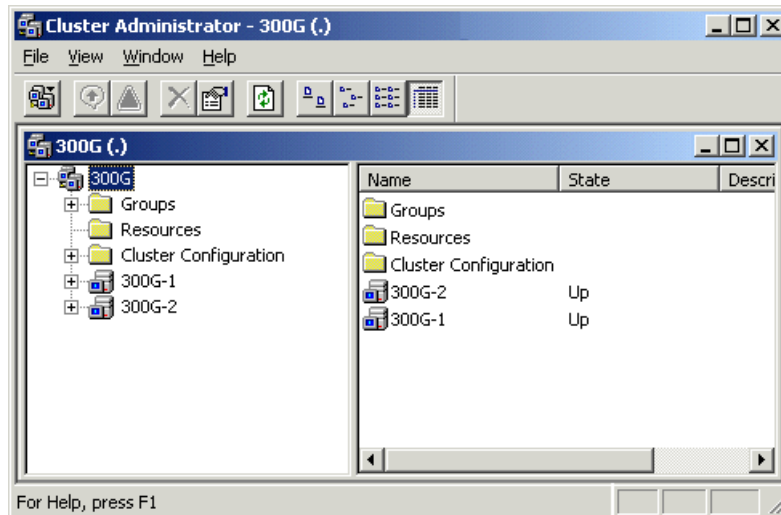


Figure 4-36 Cluster administrator

We have successfully brought the second node into our cluster. Now the real fun begins!

4.5 Administering the cluster

Now it is time to configure the storage for our cluster. All of the steps we performed were completed from the primary node (300G-1).

4.5.1 Configure cluster properties

From the Cluster Administrator window, we right-click on the cluster named 300G and select **Properties**, as shown in Figure 4-37.

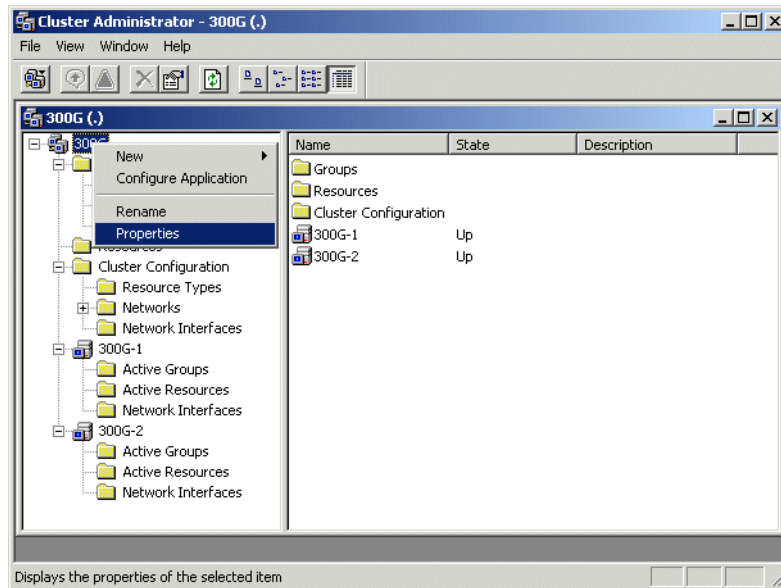


Figure 4-37 Cluster properties

The cluster properties panel is displayed. Select the Quorum tab (Figure 4-38).

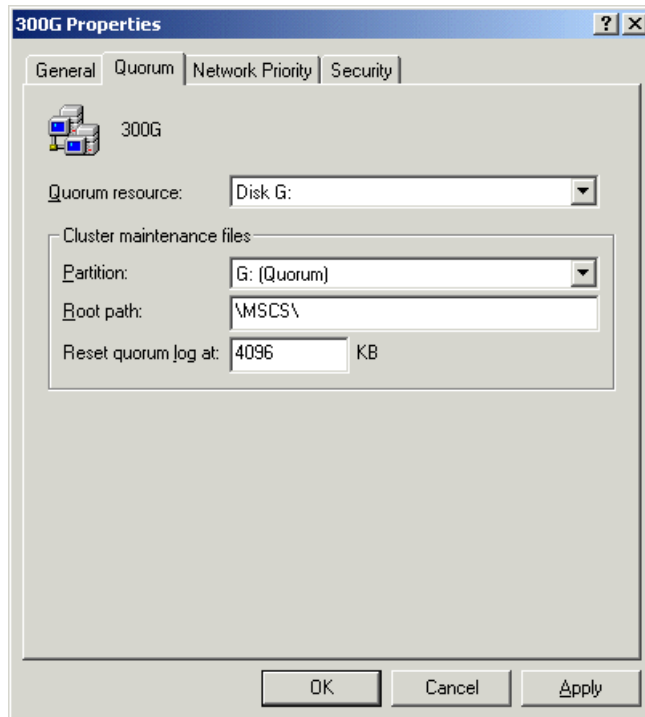


Figure 4-38 Cluster properties

Increase the size of the quorum log and press **Apply**. Now select the Network Priority tab and ensure that the Private network is located at the top of the window list as shown in Figure 4-39 using the **Move up** and **Move down** buttons.

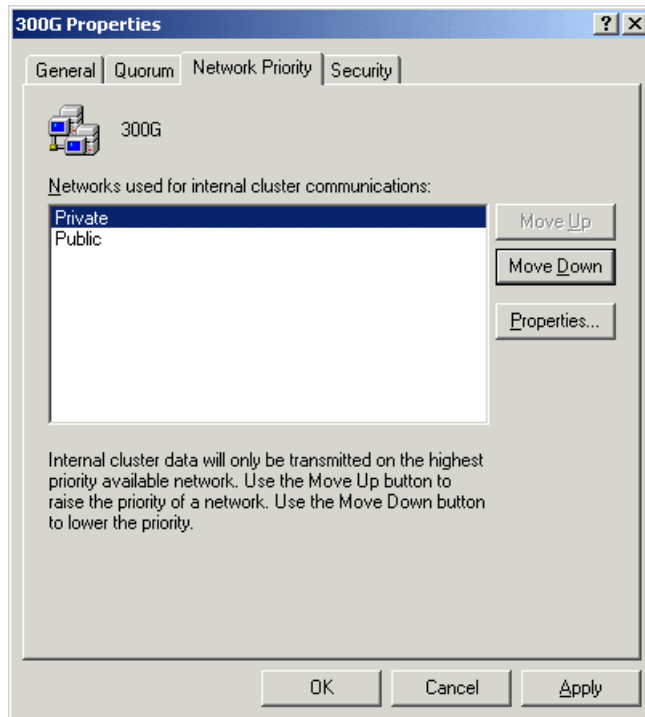


Figure 4-39 Update network priority

Press **OK** to continue.

4.5.2 Disk group administration

When you configure cluster resources, you should manually balance them between the disk groups to distribute the cluster resource functions among the two nodes. This allows for a more efficient response time for the clients and users accessing these resources.

As part of the setup we just went through, each of our physical disks were put in an individual disk group folder, as shown in Figure 4-40. We chose to consolidate all of our physical disks into two disk groups, Disk Group 1 and Disk Group 2.

Delete disk group

To delete a disk group, we first move the physical disk to Disk Group 2, and then delete the Disk Group resource.

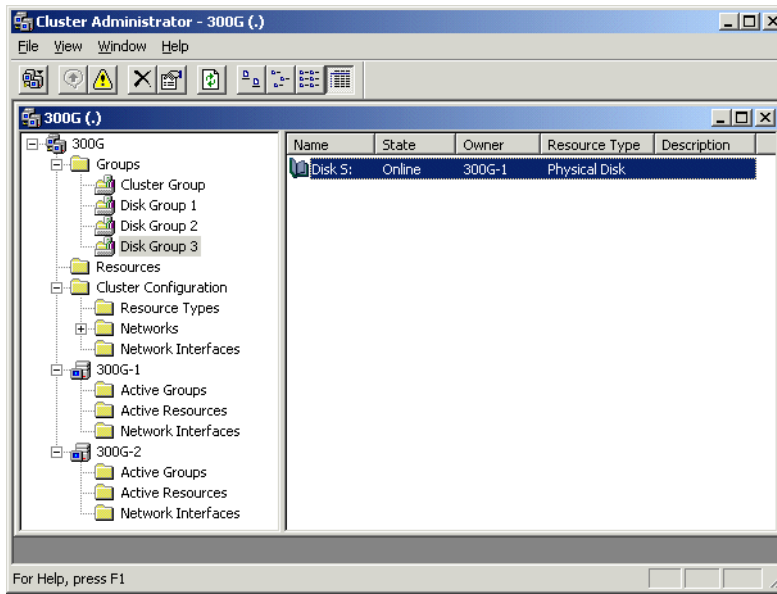


Figure 4-40 Change disk groups

Disk **S:** in Disk Group 3 is highlighted and then moved (via drag-and-drop) to Disk Group 2. Figure 4-41 shows that it has moved.

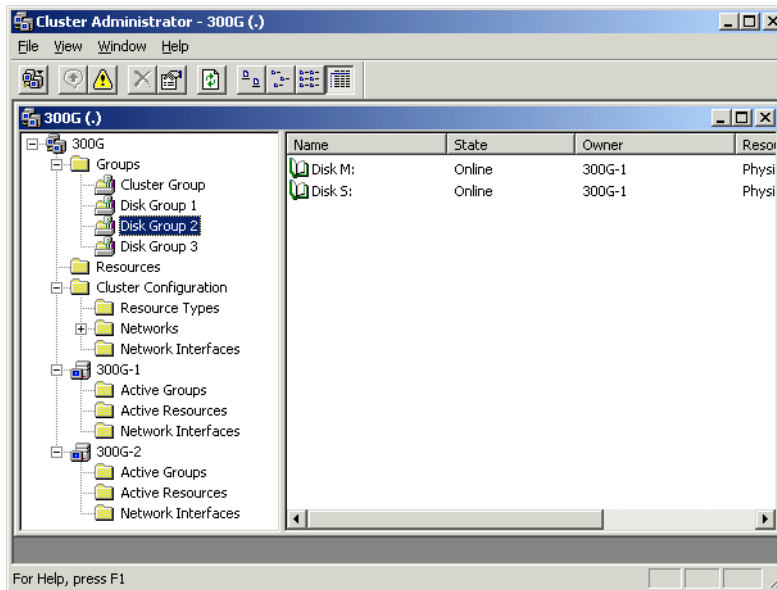


Figure 4-41 Moved disk

We now delete Disk Group 3 by right-clicking Disk Group 3 and selecting **Delete** (Figure 4-42).

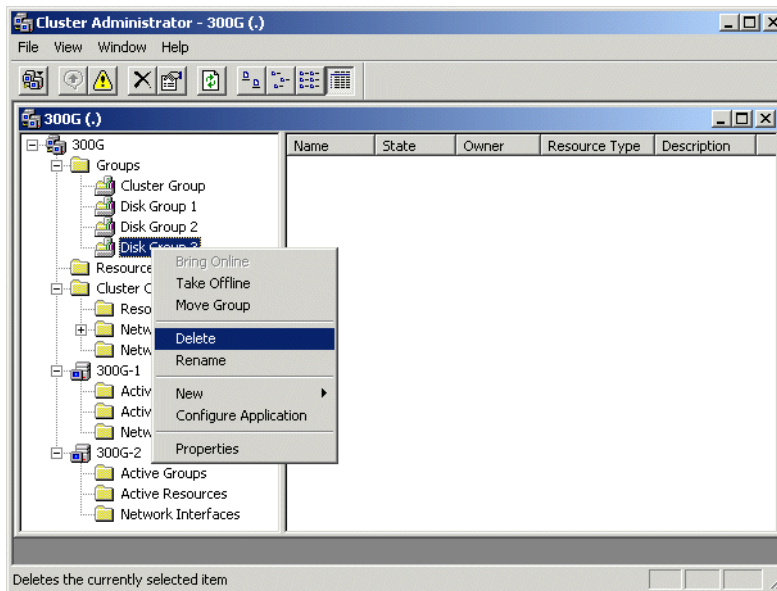


Figure 4-42 Delete disk group

A confirmation window like the one shown in Figure 4-43 is displayed.

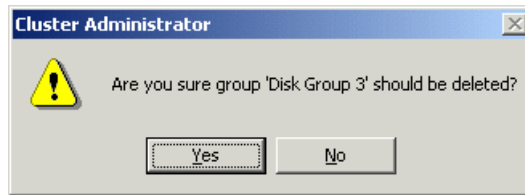


Figure 4-43 Delete group confirmation

We click **Yes** to confirm the deletion of Disk Group 3.

4.5.3 Cluster resource balancing

Each disk group has a preferred owner so that, when both nodes are running, all resources contained within each disk group will have a node defined as the owner of those resources. Even though a disk group has a preferred owner, its resources can run on the other node in the cluster following a failover. If you restart a cluster node, ownership of those resources is transferred to the preferred node once failback is initiated and completes successfully.

Consolidate disk groups

In our configuration, the first node in the cluster (300G-1), owns all of the disk resources. Disk Group 2 should be owned by 300G-2 in order to provide some balance, so we right-click on Disk Group 2 and select **Move Group** (Figure 4-44).

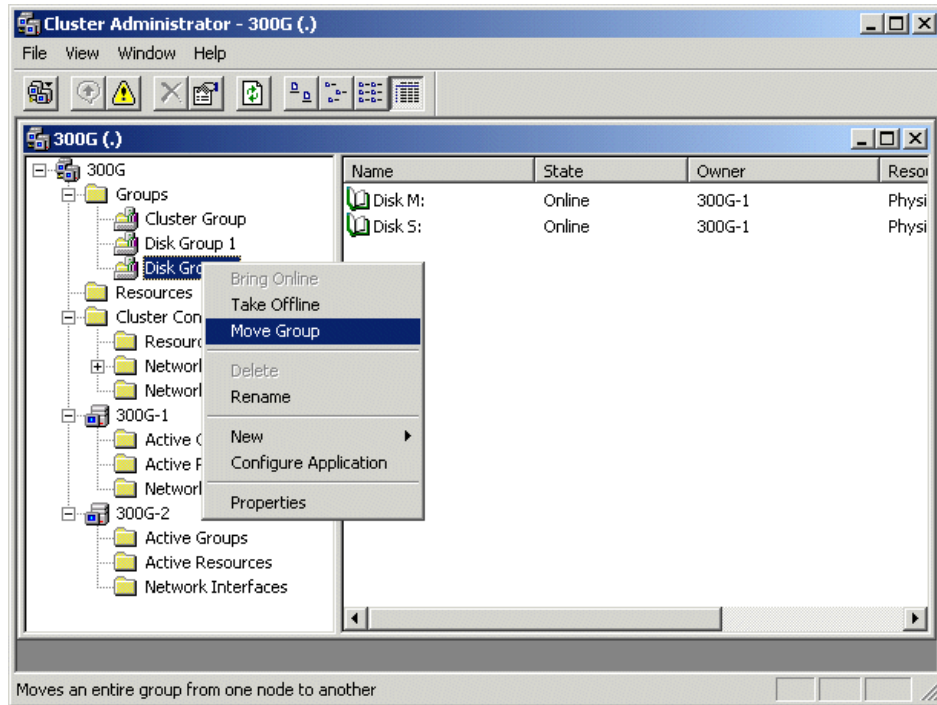


Figure 4-44 Move group

You will notice that moving disk group ownership will automatically take the physical disks offline, update the ownership of each, and finally bring all of them online. Figure 4-45 shows the physical disks with 300G-2 as the owner of each.

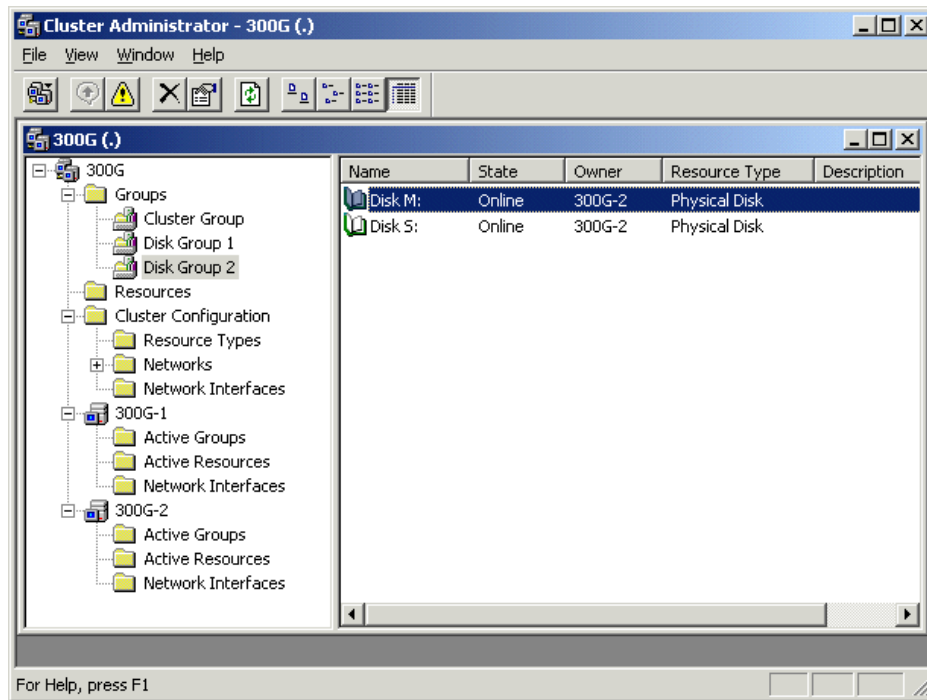


Figure 4-45 Moved group

Preferred ownership

We now must set the preferred ownership for each of the disk groups. We started with Disk Group 1. Right-click on **Disk Group 1** in the Groups folder and highlight **Properties** to display the properties page (Figure 4-46). Preferred owners are displayed in the window and can be modified by clicking the **Modify** button.

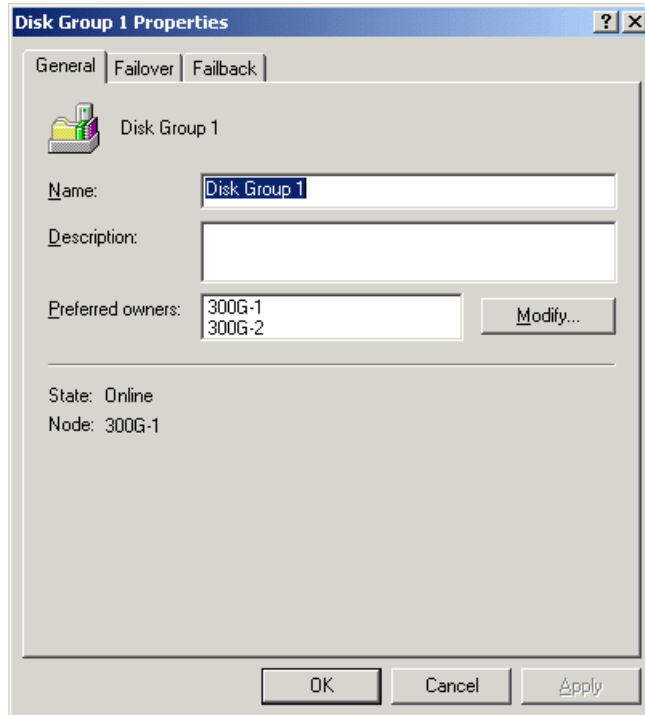


Figure 4-46 Disk group 1 preferred owners

We did the same for Disk Group 2, but now, in Figure 4-47, notice that 300G-2 is first in the list.

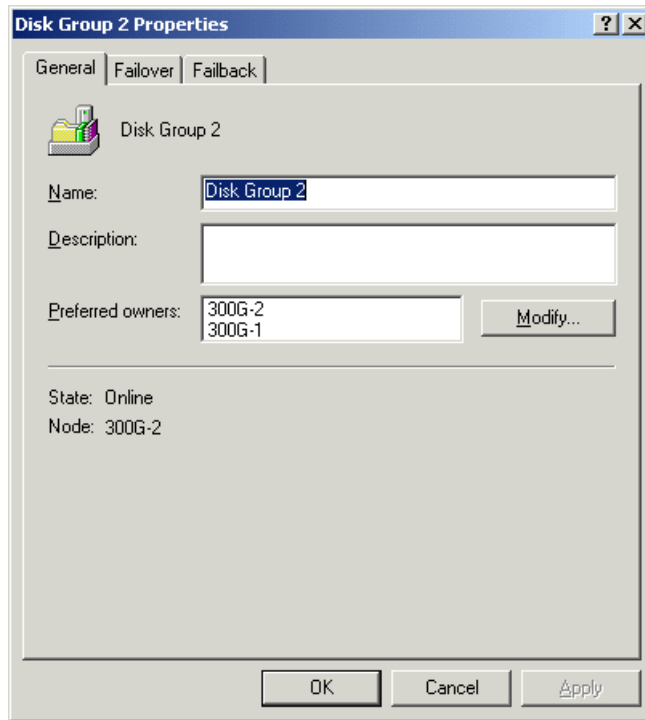


Figure 4-47 Disk group 2 preferred owners

Failover setup

The failover of resources under a disk group on a node allows users to continue accessing the resources if the node goes down. If a disk group contains a large number of resources and any one of those resources fail, then the whole group will failover according to the group's failover policy.

Note: Individual resources contained in a group cannot be moved to the other node; rather the group it is contained in is moved.

From the disk group's property page, select the **Failover** tab to view the failover options, shown in Figure 4-48.

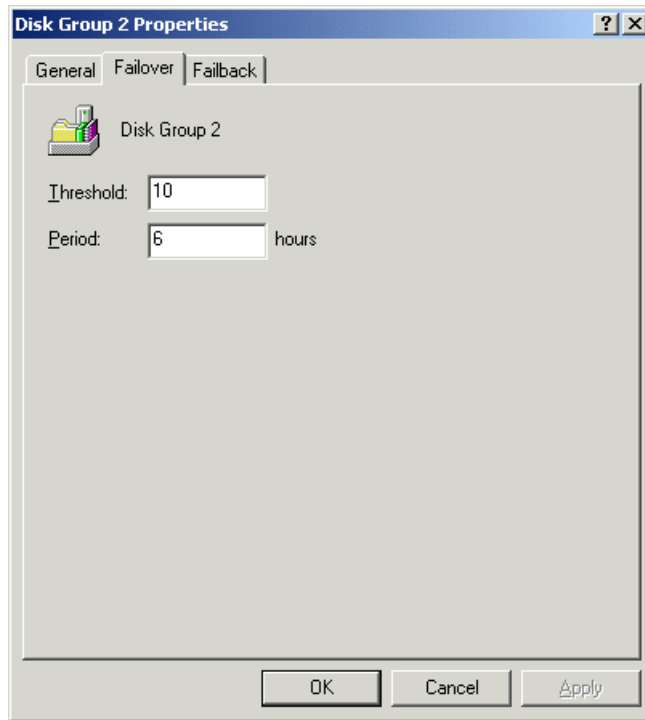


Figure 4-48 Failover options

The threshold and period determine how many times and for how long clustering services will attempt to failover a group. Using the values from Figure 4-48, if a network name fails, clustering services attempts to failover the group 10 times within 6 hours, but if the resource fails an eleventh time, the resource will remain in a failed state, and Administrator action is required to correct the failure.

Failback setup

In allowing failback of groups, there is a slight delay in the resources moving from one node to the other. The group can also be instructed to allow failback when the preferred node becomes available or to failback during specific, off-peak usage hours.

Choosing the right failover policy is not an easy task. We recommend that you either prevent automatic failback or only allow failback during off-peak hours. Figure 4-49 shows the failback options selection.

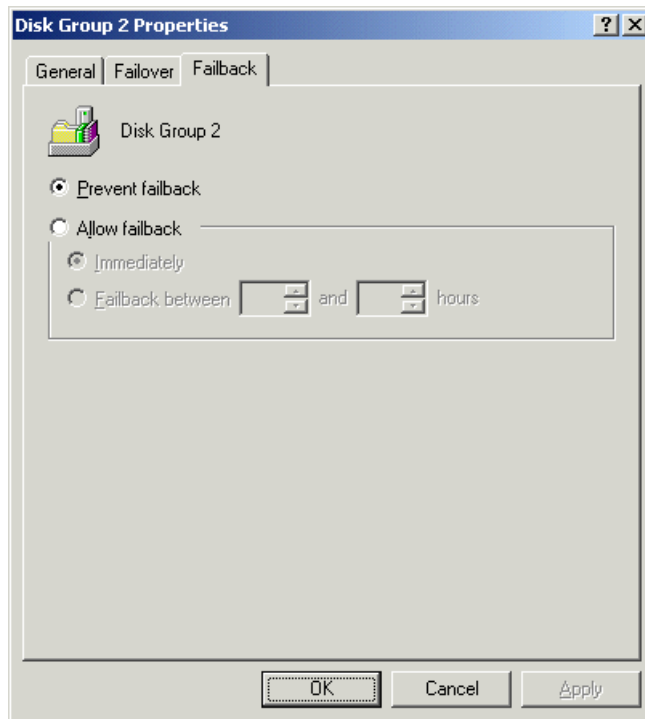


Figure 4-49 Failback options

Select **Prevent failback** and click **OK** to continue.

4.5.4 Configure file shares

File shares are resources which allow you to connect to a physical disk. The creation of file shares involves dependencies on a physical disk, a static IP address and a network name. The various dependencies allow resources that are defined to the same disk group to move as a group. The dependencies also assure necessary access for the given resource. Figure 4-50 provides a pictorial view of how file share dependencies should be configured.

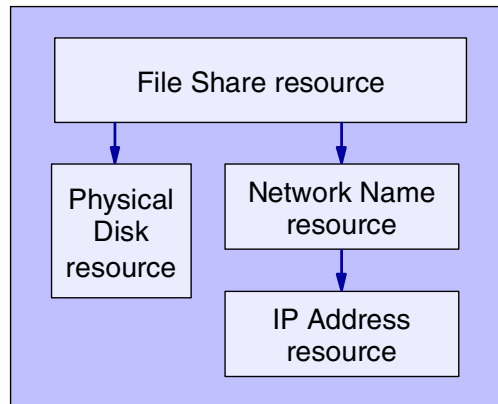


Figure 4-50 File share resource dependencies

As Figure 4-50 suggests, here is the order for creating File Share resources:

1. Create an IP Address resource.
2. Create a Network Name resource and make it dependent on the IP Address.
3. Create a File Share resource and make it dependent on both the Physical Disk and the Network Name.

Now that you know what you have to do, we will show you how to do it.

To create a new resource, from the Cluster Administrator window, select **File -> New -> Resource** (Figure 4-51).

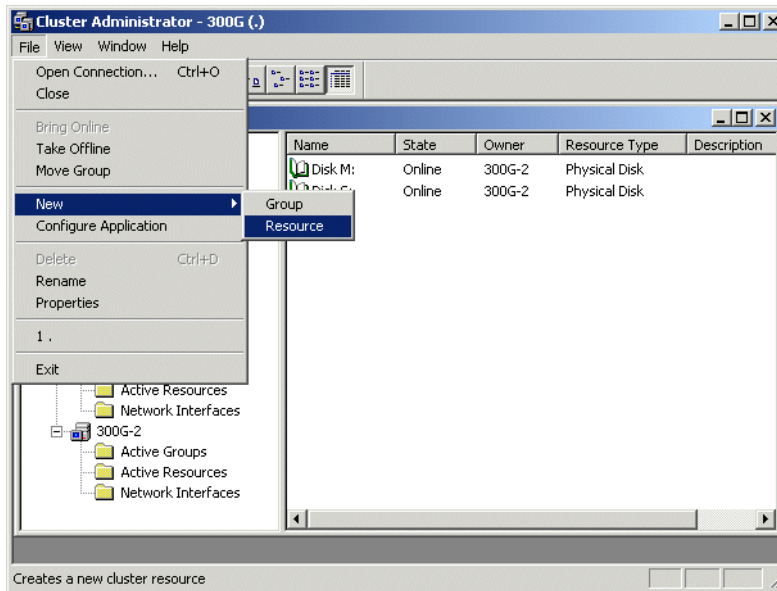


Figure 4-51 Create new resource selection

Create the IP resource

The new resource configuration menu is displayed as shown in Figure 4-52. Enter the name of the IP resource (we chose FastIP). Select **IP Address** as the Resource type and select the Disk Group the IP resource should reside in. Then click **Next**.

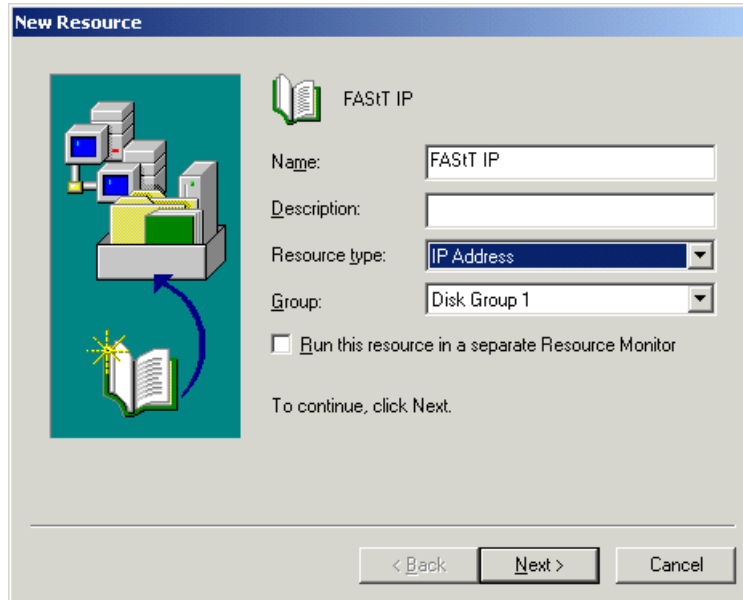


Figure 4-52 Create a new IP resource

Select the possible owners from the available nodes (Figure 4-53) and click **Next**. Remember, these are just *possible* owners. Ownership of resources is based on the preferred ownership set for each disk group.

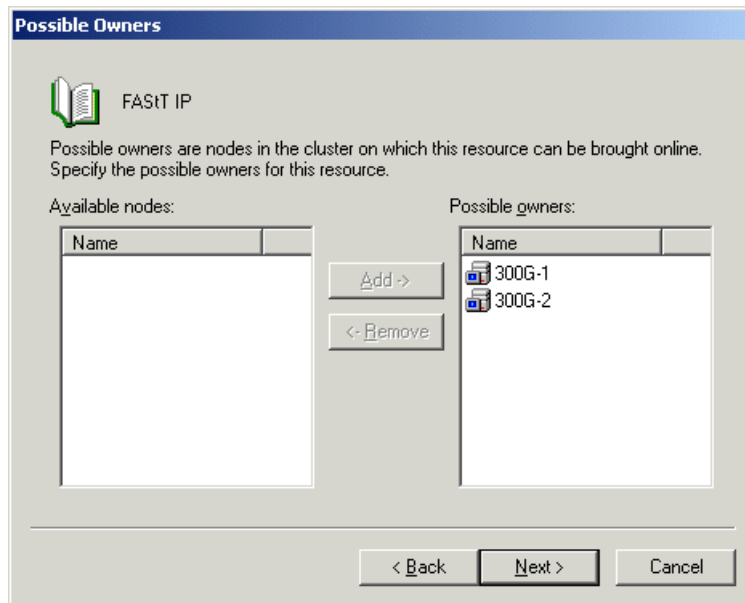


Figure 4-53 Possible owners

The IP resource does not need to have any dependencies, so leave the Resource Dependencies window empty (see Figure 4-54) and just click **Next**.

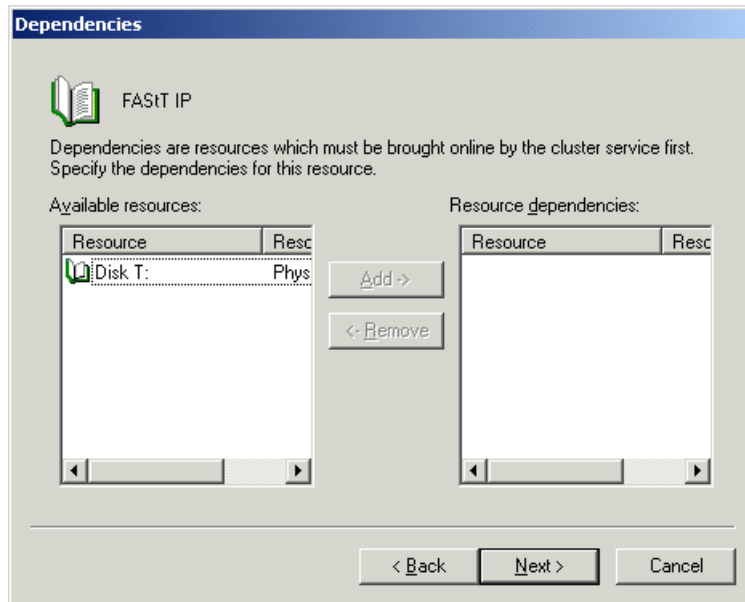


Figure 4-54 Dependencies

The IP resource requires an IP address and subnet mask. Our settings are displayed in Figure 4-55. Enter yours and click **Finish**.

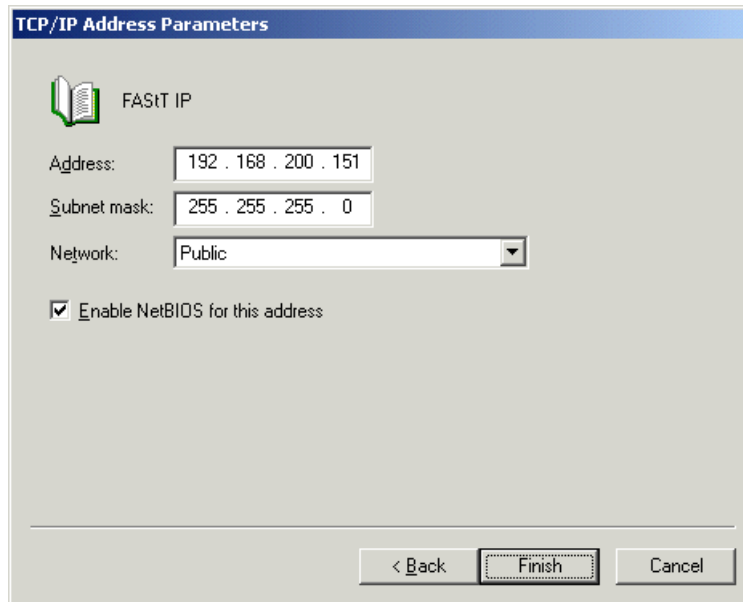


Figure 4-55 Resource parameters

That is all there is to creating an IP resource. You will see a Resource Created Successfully window similar to the one shown in Figure 4-56.



Figure 4-56 IP resource created successfully

The resource needs to be brought online manually before use, so right-click the resource and select **Bring Online**, as shown in Figure 4-57.

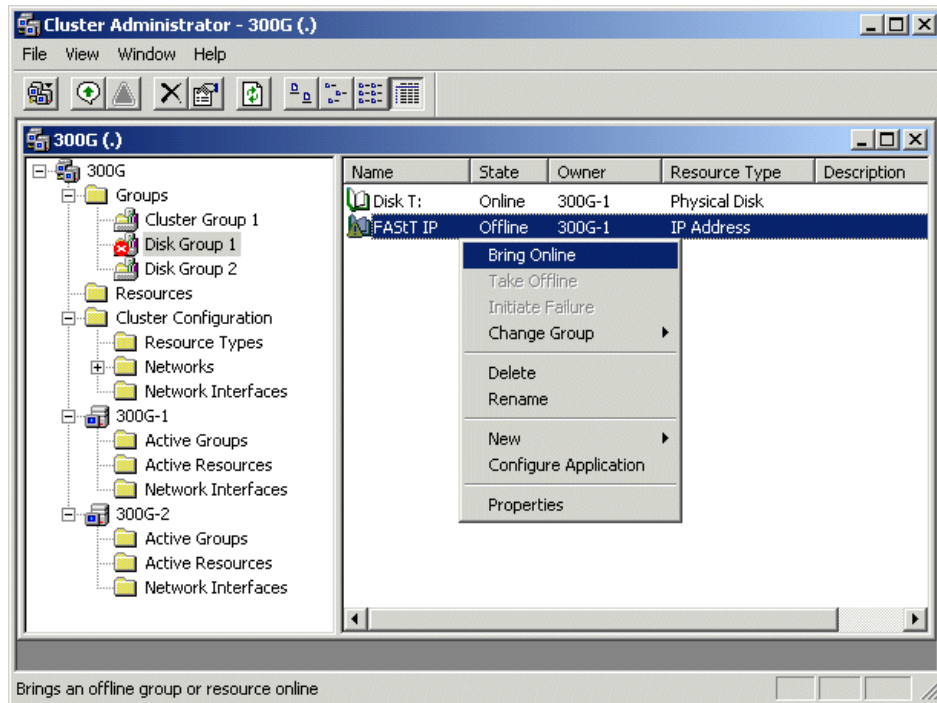


Figure 4-57 Bring IP resource online

The IP resource is now online.

Create network name resource

Select **File -> New -> Resource** (Figure 4-51) to bring up the new resource configuration tool. Enter the network resource name, select the resource type as **Network Name** and select the group from the drop-down list. Our network name is Bert, the resource type is Network Name, and the group is Disk Group 1, as shown in Figure 4-58. Then click **Next**.

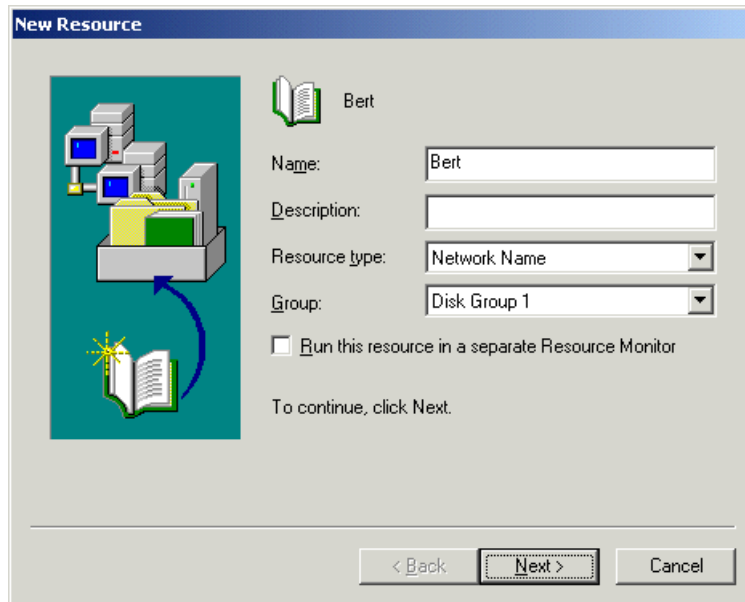


Figure 4-58 New network name resource

Add both nodes as possible owners using the **Add** button (Figure 4-59).

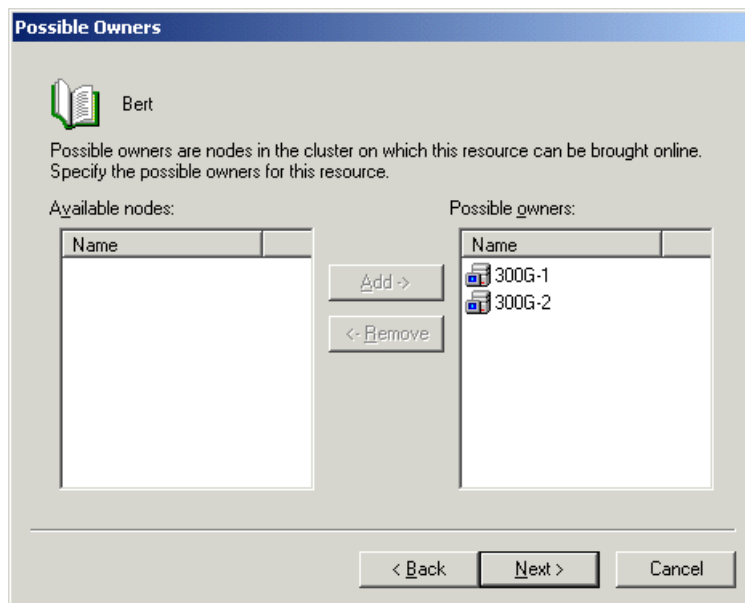


Figure 4-59 Possible owners

Now it is time to configure the dependencies. As described in Figure 4-54, the only dependency for the network name is the IP address. Select the IP resource name in the **Available resources** window and click the **Add ->** button as shown in Figure 4-60.

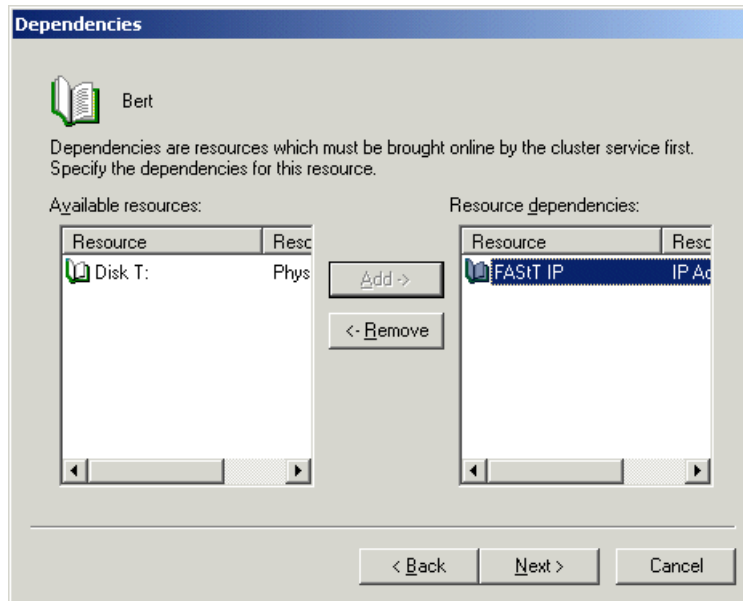


Figure 4-60 Network name dependencies

The last step in adding the network name resource is to enter the actual network name (Figure 4-61). This name is what your clients will use when connecting to your disk resources. Click **Finish** to create the network name resource.

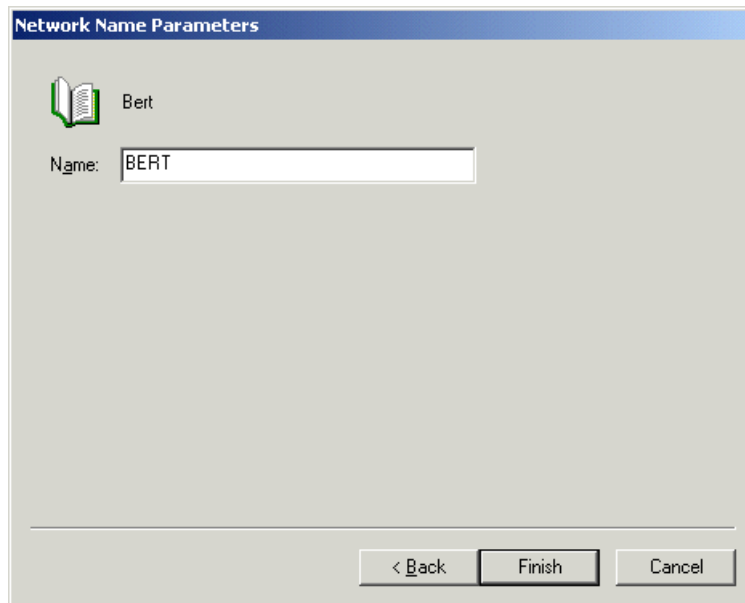


Figure 4-61 Network name parameters

You should now see a Resource Created Successfully window, as shown in Figure 4-62.



Figure 4-62 Create network name successful

The final step is to bring your network name resource online, shown in Figure 4-63.

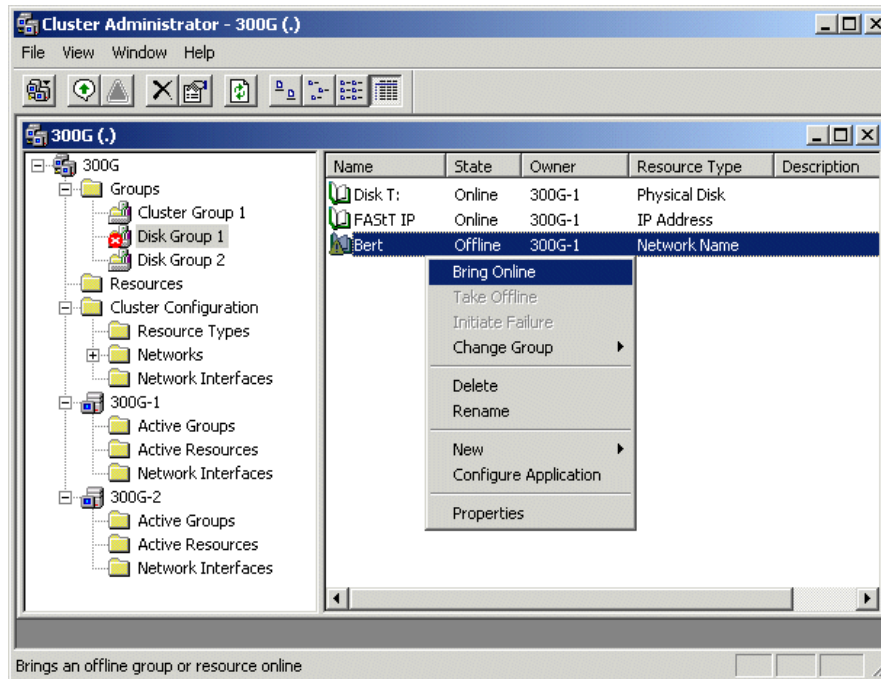


Figure 4-63 Bring network name resource online

Network name creation is now complete.

Create a new file share resource

Creating the file share resource is the final step. Select **File -> New Resource**. The new resource configuration tool opens (see Figure 4-64). Enter the resource name, select **File Share** from the drop-down menu as the Resource Type, and also decide which Disk Group this resource should belong to. (This should be the same Disk Group that you assigned the IP address and network name to.) Then click **Next**.

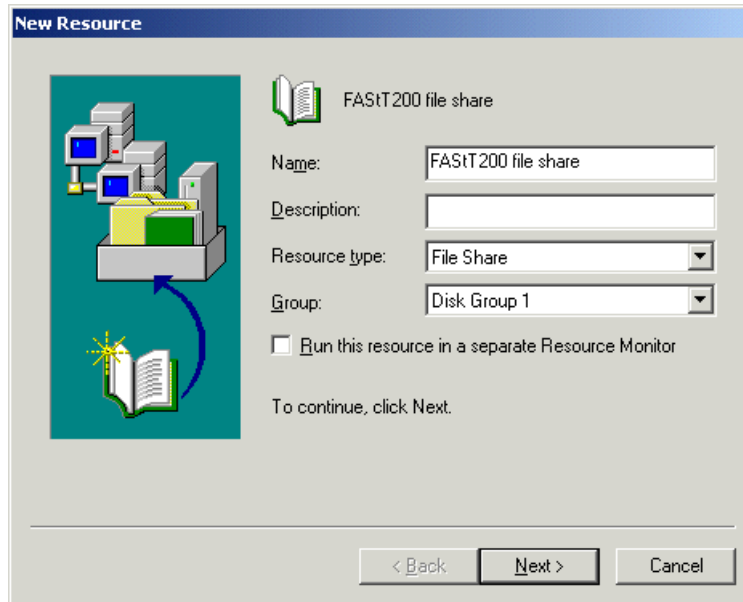


Figure 4-64 New file share resource

Select the possible owners by using the **Add ->** button and then click **Next**, as shown in Figure 4-65.

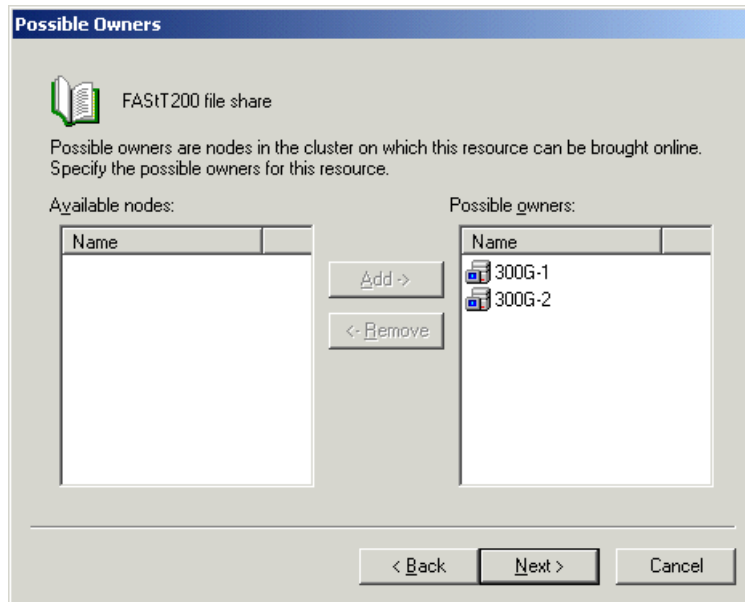


Figure 4-65 Possible owners

The dependencies, as illustrated in Figure 4-54 on page 211, are the network name and the physical disk. Highlight the network name and physical disk and click the **Add ->** button (Figure 4-66) to add these resources as dependencies and then click **Next**.

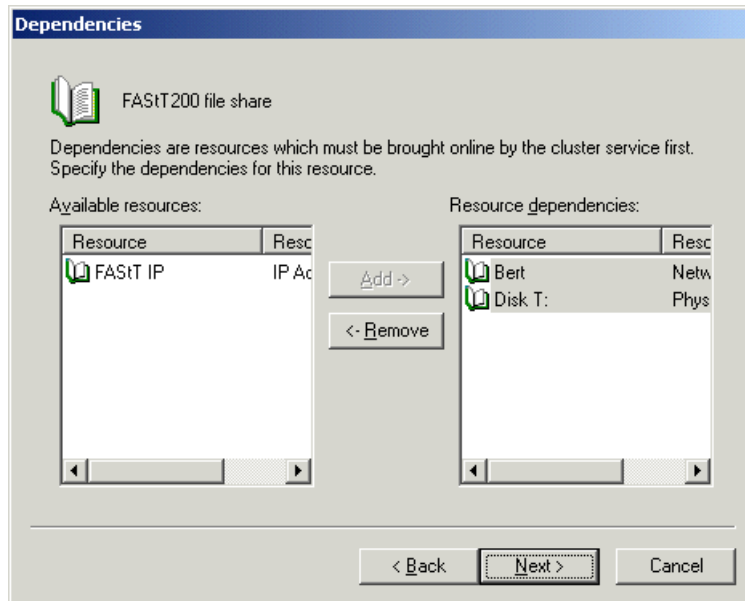


Figure 4-66 File share dependencies

Now enter the file share parameters as shown in Figure 4-67. These include the share name, the path, and an optional comment. Click the **Permissions** button to set up access permissions.

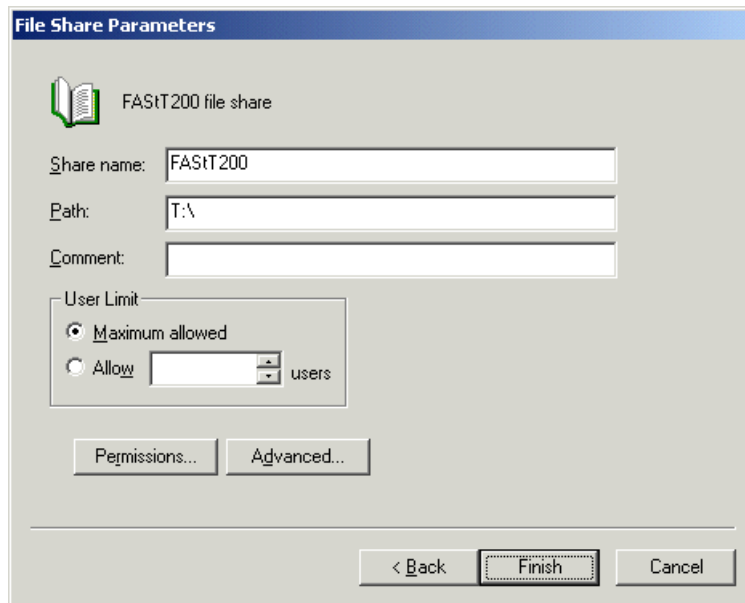


Figure 4-67 File share parameters

Change the permissions of the File Share and then click **OK** (Figure 4-68).

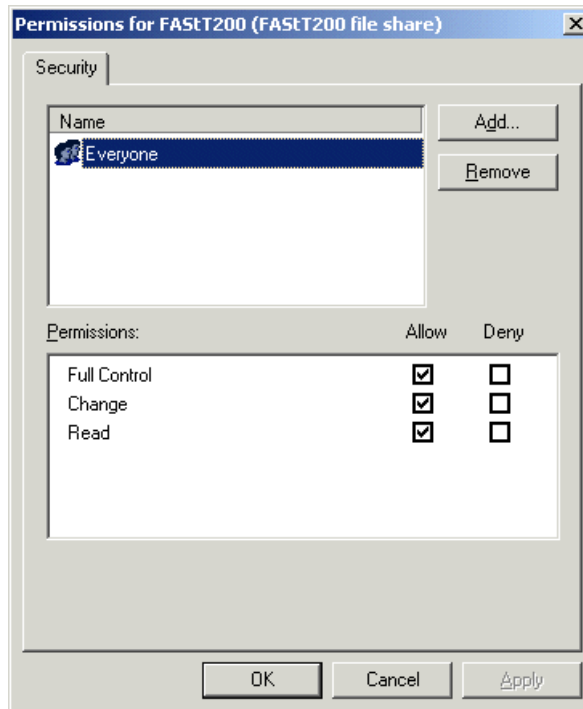


Figure 4-68 File share permissions

From the File Share Parameters window, select **Advanced**. Since this is a windows share, select the **Normal share** radio button, as shown in Figure 4-69, and click **OK**.

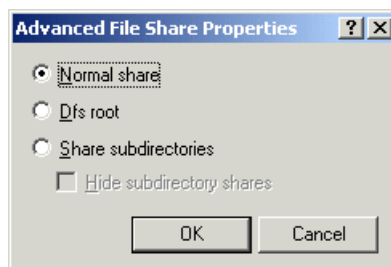


Figure 4-69 File share advanced settings

Finally, click **Finish** from the File Share Parameters window and you should see a Resource Created Successfully window, as shown in Figure 4-70.

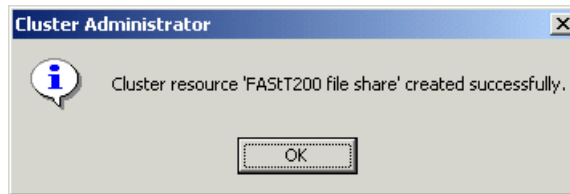


Figure 4-70 Create file share successful

The final step to use the file share is to bring it online. Right-click the new file share resource (Figure 4-71) and select **Bring Online**.

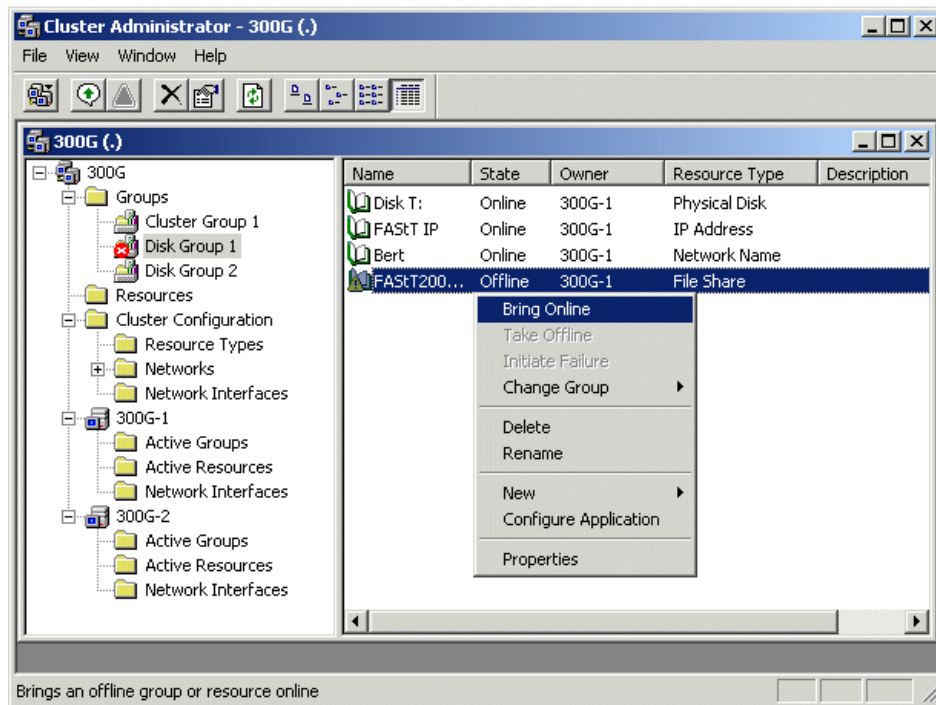


Figure 4-71 Bring file share online

Create new NFS resource

Creating an NFS file share resource is similar to creating a normal file share as described in “Create a new file share resource” on page 218. Using **File -> New Resource**, enter the name of the NFS share, select **NFS Share** from the pull-down menu as the Resource Type, Assign a Disk Group, and finally click the **Next** button, as shown in Figure 4-72.

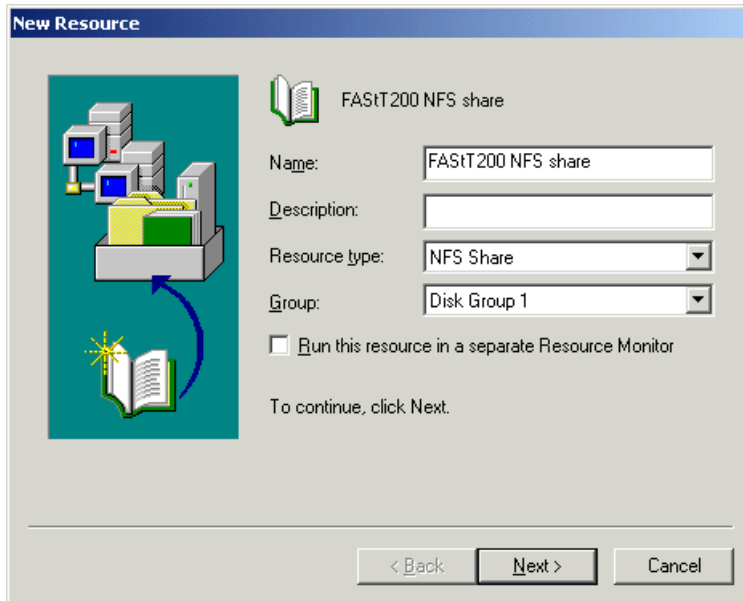


Figure 4-72 Create new NFS resource

Select both nodes as the possible owners and click **Next**. Add the physical disk and the Network Name as dependencies using the **Add ->** button (Figure 4-73).

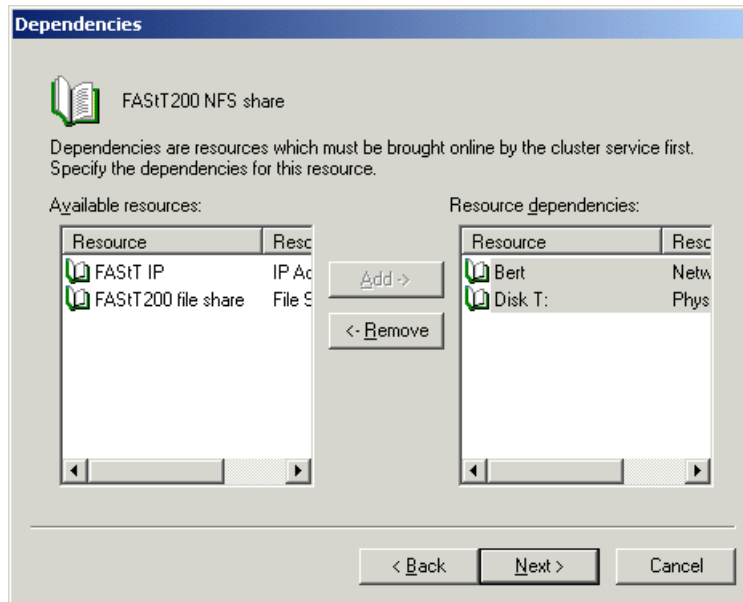


Figure 4-73 New NFS resource dependencies

As the NFS share parameters, enter the share name and path, and then choose your share directory preferences (Figure 4-74). If you do not have a NIFS structure in place you will probably need to click **allow anonymous access** since you will not have any way to validate users from UNIX clients.

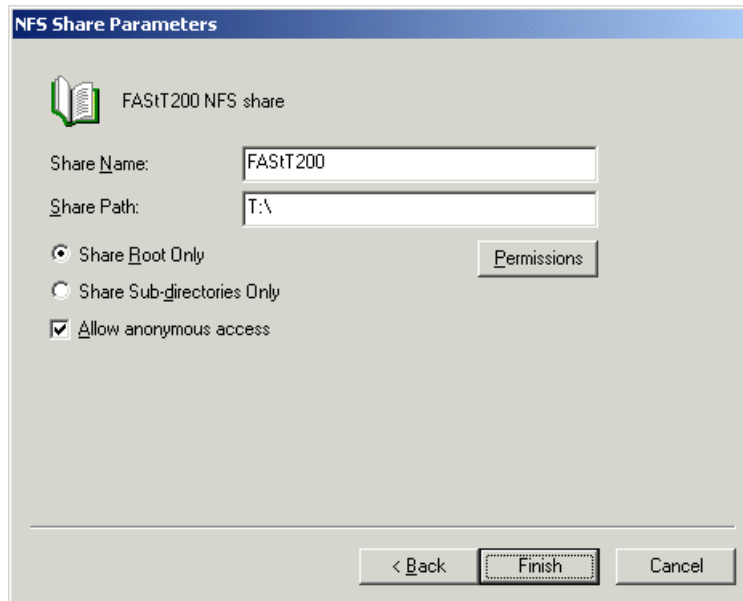


Figure 4-74 New NFS resource parameters

Select **Permissions** to specify specific access as in Figure 4-75, then select **OK**.

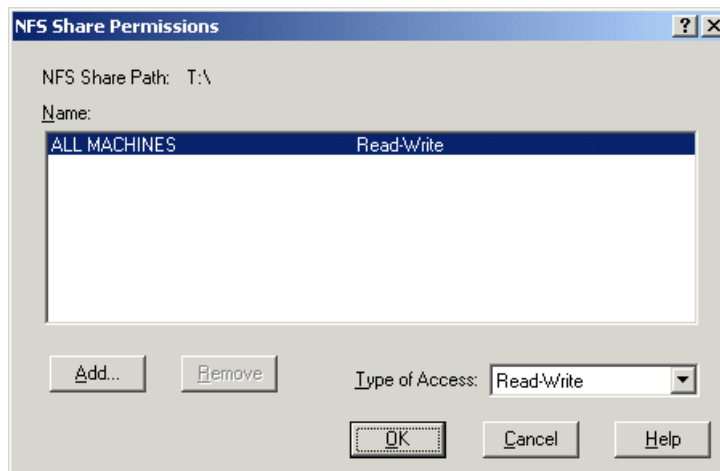


Figure 4-75 New NFS share permissions

Finally, click **Finish**, and the Resource Created Successfully window is displayed. Click **OK** and right-click the newly created NFS resource and select **Bring Online**, as shown in Figure 4-76.

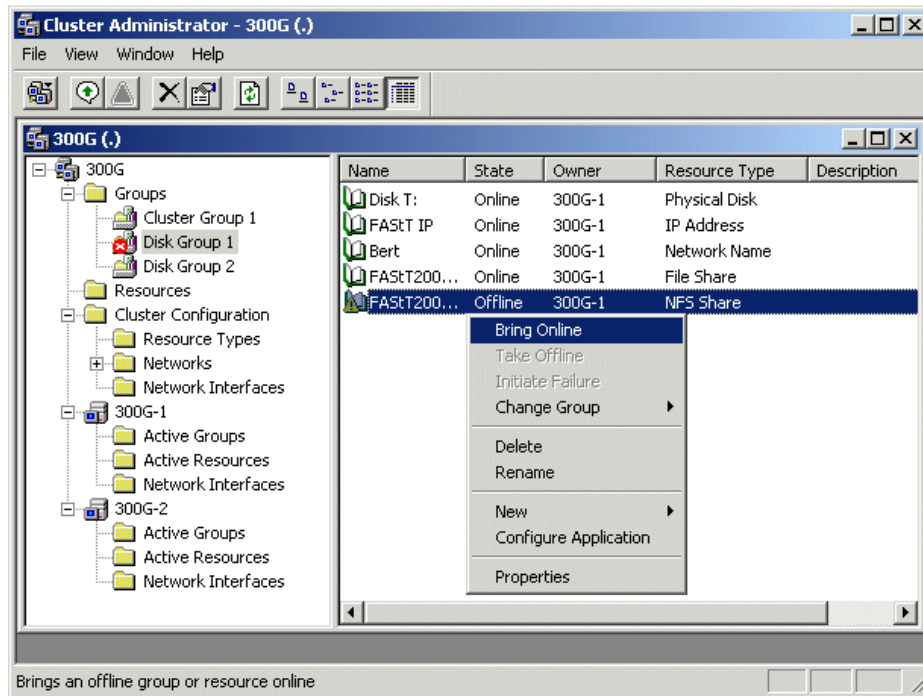


Figure 4-76 Bring new NFS share online

4.6 Client connectivity

In this section we explain how to set up client connectivity.

4.6.1 Windows clients

Figure 4-78 illustrates initiating connectivity to our file resource. From the My Computer window, select **Tools -> Map Network Drive**, as shown in Figure 4-80.

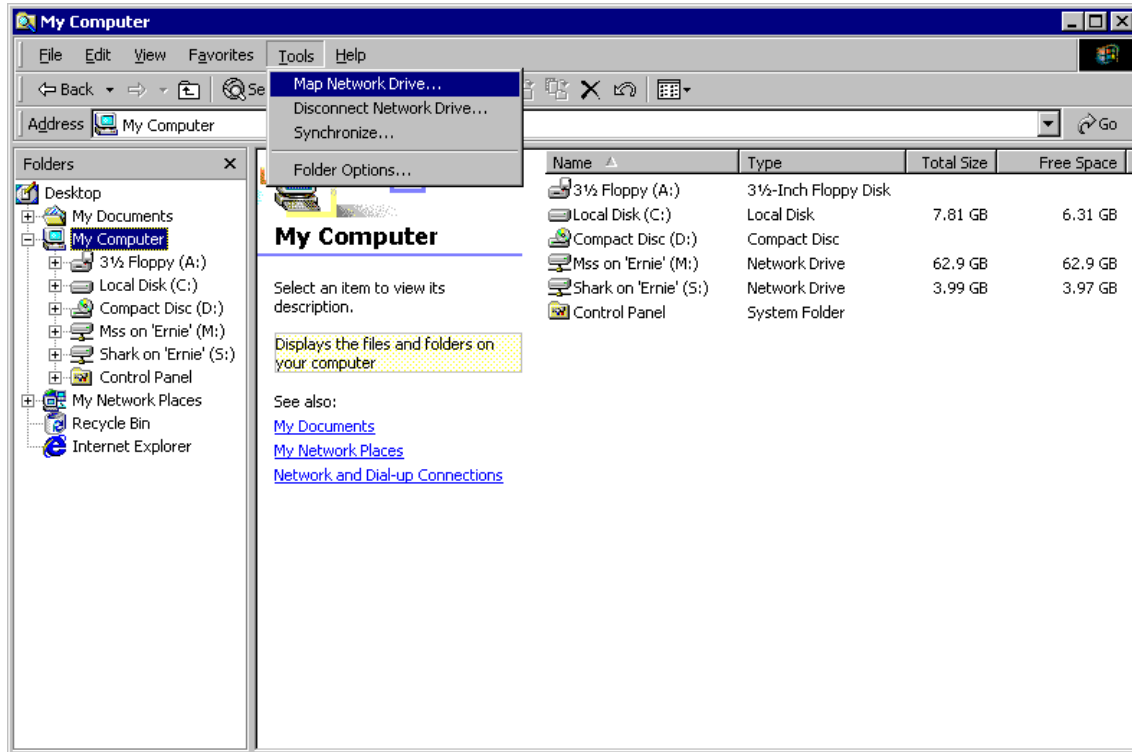


Figure 4-77 Map network drive

Remember, from Figure 4-61 on page 216 and Figure 4-67 on page 221, that the network name is 'Bert' and the file resource is 'FAST200'. After selecting the appropriate drive letter, enter the information as shown in Figure 4-78 and click **Finish**.

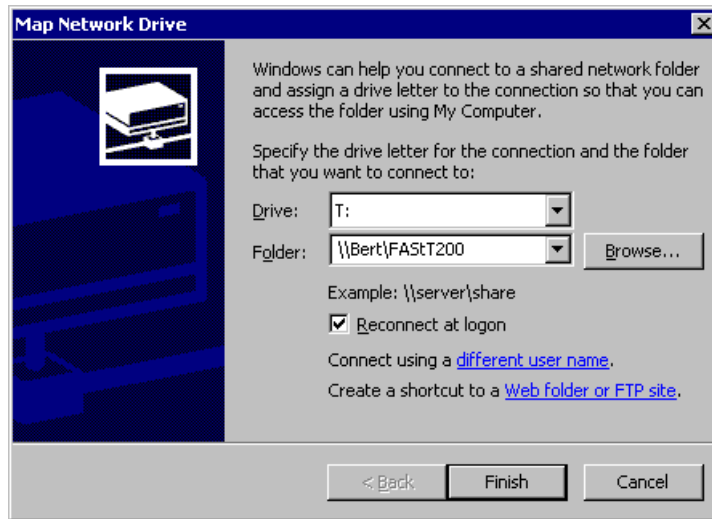


Figure 4-78 Select network drive letter and location

You should now see the newly attached volume (Figure 4-80) and be able to utilize it immediately.

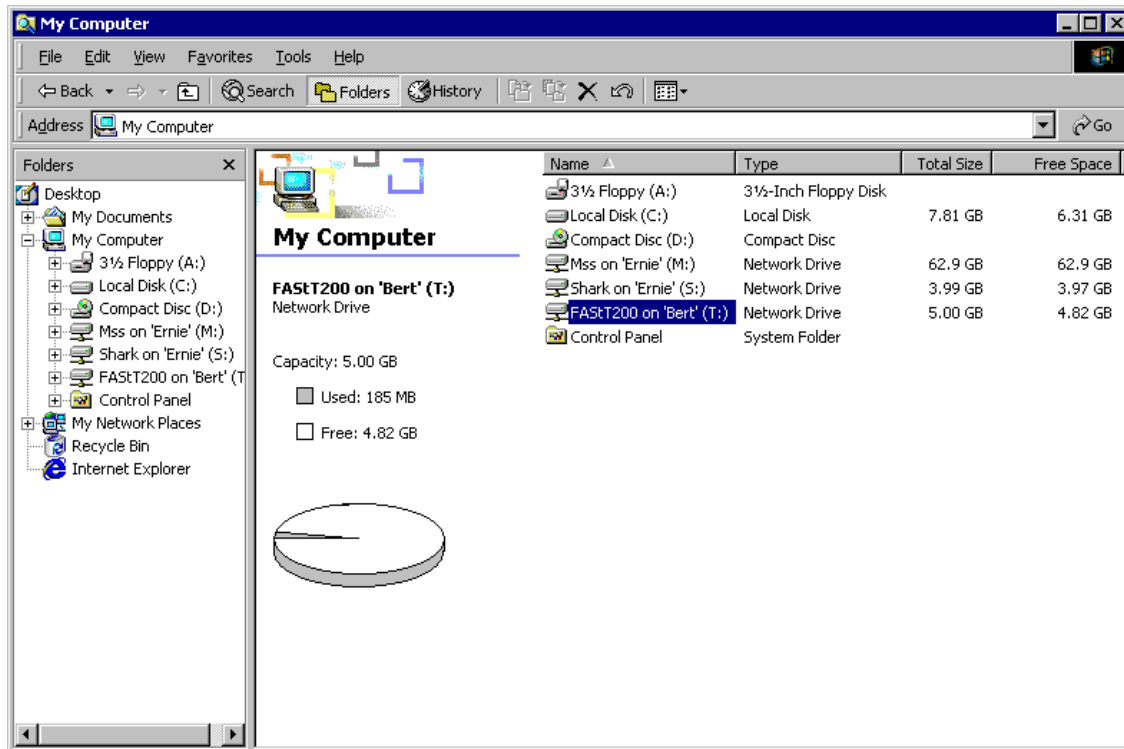


Figure 4-79 Drive connected successfully

Figure 4-80 shows the directory listing of the files located in \\Bert\FAStT200.

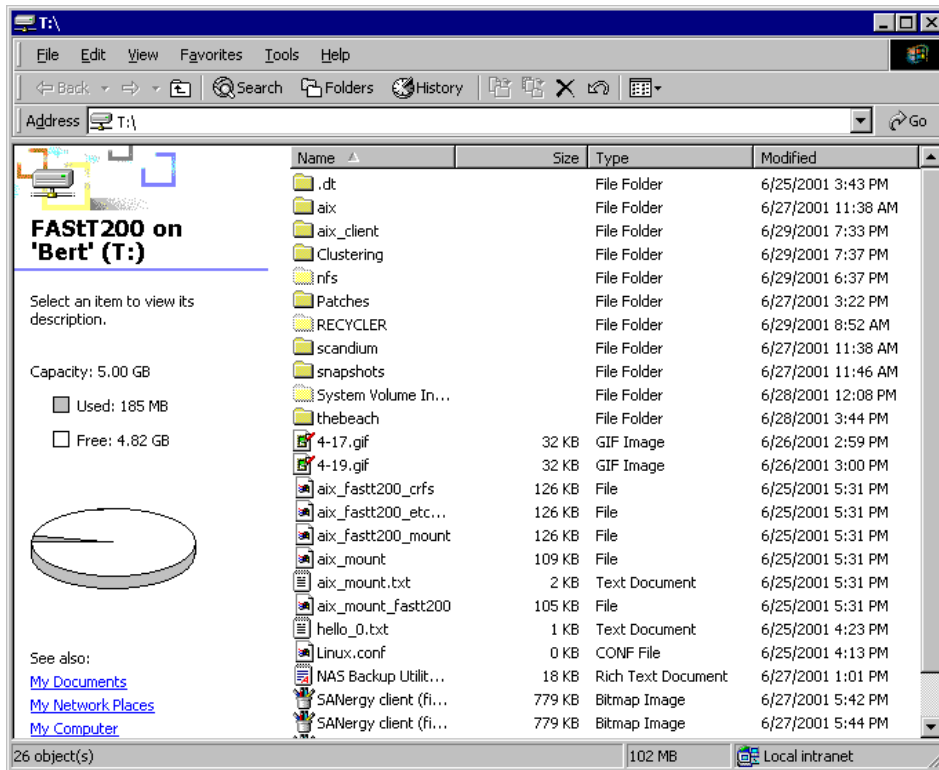


Figure 4-80 Directory listing of files on T:

Congratulations, you have just successfully configured clustering!

4.6.2 UNIX clients

We will use our Red Hat Linux client to show how to connect to the NFS share from a generic UNIX client. This procedure is exactly the same as we described in Section 3.7.4, “Accessing the shares from our Linux/Solaris/HP-UX clients” on page 159 for a single-node installation. To review, all you need to do is add the connection information to the `/etc/fstab` file and then mount the volume. For example, you could add the following line to `/etc/fstab`:

```
bert:/T /FASTt200 nfs user,auto,rw 0 0
```

Then issue the following mount command:

```
mount bert:/T /FASTt200
```

4.6.3 AIX clients

AIX connectivity is as straightforward as the non-clustered version. Example 4-1 lists the commands required to create the entry in the `/etc/filesystems` file, mount the NFS share, and display directory contents.

Example 4-1 AIX connectivity

```
# crfs -v nfs -m /FAStT200 -n Bert -d FAStT200 -A yes
# tail -n 7 /etc/filesystems

/FAStT200:
    dev           = FAStT200
    vfs           = nfs
    nodename      = Bert
    mount         = true
    account       = false

# mount -v nfs Bert:FAStT200 /FAStT200
# mount
  node          mounted          mounted over    vfs      date      options
-----
                /dev/hd4          /               jfs      Jun 25 15:50 rw,log=/dev/hd8
                /dev/hd2          /usr            jfs      Jun 25 15:50 rw,log=/dev/hd8
                /dev/hd9var       /var            jfs      Jun 25 15:50 rw,log=/dev/hd8
                /dev/hd3          /tmp            jfs      Jun 25 15:50 rw,log=/dev/hd8
                /dev/hd1          /home           jfs      Jun 25 15:51 rw,log=/dev/hd8
                /dev/lv00       /usr/welcome_arcade jfs      Jun 25 15:51
rw,log=/dev/hd8
                /dev/lv01          /usr/welcome    jfs      Jun 25 15:51 rw,log=/dev/hd8
Bert  FAStT200      /FAStT200      nfs3     Jun 29 19:34
# cd /FAStT200
# ls
.dt
4-17.gif
4-19.gif
Clustering
Linux.conf
NAS Backup Utility.rtf
Patches
RECYCLER
SANergy client (fig 4-18).bmp
SANergy client (fig 4-19).bmp
SFU-User Name Mapping (end of chapter 3).bmp
System Volume Information
aix
aix_client
aix_fastt200_crfs
aix_fastt200_etc_filesystems
aix_fastt200_mount
aix_mount
```

```
aix_mount.txt  
aix_mount_fastt200  
hello_0.txt  
nfs  
scandium  
snapshots  
thebeach  
x.dat  
#
```



Using SANergy to secure high-speed data sharing

In Chapter 3, “Implementing the IBM TotalStorage NAS 300G” on page 71 we described how to use the IBM TotalStorage NAS 300G (300G) to share SAN-based storage.

However, if you need to have high-speed access to shared data from several client machines, then you may need a more complex network structure that connects those client machines directly into the SAN.

In this chapter, we explain how the 300G, with Tivoli SANergy, can help you get the most out of such an environment.

5.1 A brief overview of Tivoli SANergy

Using a SAN promises high-speed transfers of shared data that do not bog down the LAN. While this sounds wonderful, in practice, it is hard to achieve. The technology that enables a SAN is available, and it is relatively easy to configure multiple host systems to access data from a shared volume in the SAN. Unfortunately, today's operating systems are not SAN-aware.

Giving multiple host systems access to the same data on a shared volume can cause plenty of headaches. Because of the "it's mine, all mine" nature of Windows, UNIX clients may suddenly lose access to a volume after a Windows client happily grabs the whole volume as its own. Worse, systems may overwrite or corrupt data in use by another system because no one host system is controlling the locks on the files in the shared volume.

However, do not despair yet! There is a solution to this problem, which comes bundled with the 300G: Tivoli SANergy. Using this software, you can safely share volumes across multiple host systems at high speeds without risking data corruption. This chapter explains how to improve on the data sharing model we presented in Chapter 3, "Implementing the IBM TotalStorage NAS 300G" on page 71 by taking advantage of the SANergy software that comes pre-installed on the 300G.

Using the 300G as a SANergy MetaData Controller (MDC) makes it possible for you to securely share an NTFS partition on a disk system attached to the SAN out to both UNIX and Windows hosts at the high speeds of Fibre Channel. For SANergy to be useful to you, however, you must meet several prerequisites in addition to having your systems set up for basic file sharing as described in Section 3.1, "Sharing SAN-based storage through the 300G" on page 72. These prerequisites include having:

- ▶ A license for each machine that will be running the Tivoli SANergy software
- ▶ IP connectivity between all of the machines that will participate in the SANergy file sharing and the 300G
- ▶ An HBA card in each machine that gives it access to the shared storage volume through the fibre network

Tivoli SANergy is unique SAN software that allows sharing of access to application files and data between a variety of heterogeneous servers and workstations connected to a SAN. In addition, Tivoli SANergy uses only industry-standard file systems like NFS and CIFS, enabling multiple computers simultaneous access to shared files through the SAN (shown in Figure 5-1).

This software allows users to leverage existing technical resources instead of learning new tools or migrating data to a new file system infrastructure. It also allows SAN-connected computers to have the high-bandwidth disk connection of a SAN while keeping the security, maturity, and inherent file sharing abilities of a LAN.

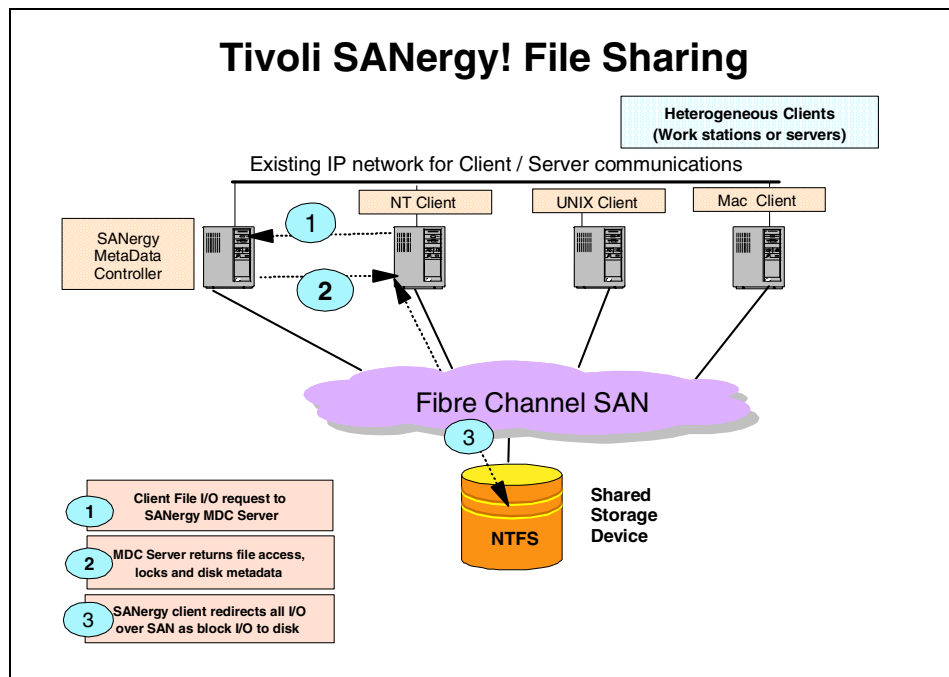


Figure 5-1 SANergy configuration

SANergy employs technology to combine the simplicity of LAN-based file sharing with the very high data transfer speeds afforded by today's Fibre Channel, SCSI, and SSA storage networks. This enables the use of high-speed, heterogeneous data sharing without the performance limiting bottlenecks of file servers and traditional networking protocols. Tivoli SANergy is unique in that it extends standard file systems and network services provided by the operating systems that it supports (Windows NT, MacOS, AIX, as well as various versions of UNIX and Linux).

As an OS extension built on standard systems interfaces, SANergy fully supports the user interface, management, access control, and security features native to its host platforms, providing all the file system management, access control and security required in a network. With SANergy, virtually any network-aware application can access any file at any time, and multiple systems can transparently share common data.

In addition to the SAN, Tivoli SANergy also uses a standard LAN for all the metadata associated with file transfers. Because Tivoli SANergy uses standard file systems, even if the SAN should fail, access to data via the LAN is still possible. Also, because each system has direct access to the SAN-based storage, Tivoli SANergy can eliminate the file server as a single point of failure for mission-critical enterprise applications. Tivoli SANergy can also easily manage all data backup traffic over the storage network, while the users enjoy unimpeded LAN access to the existing file servers.

For more information on Tivoli SANergy, please see Section 1.6.1, “Tivoli SANergy” on page 31, or refer to the SANergy redbook: *A Practical Guide to Tivoli SANergy*, SG24-6146.

5.2 Configuring the 300G as a SANergy MDC

In order for SANergy to work, all of the computers that will be using it will eventually need to be connected to the SAN. When setting up the 300G to be the MDC, however, it is best to leave the other machines disconnected and have only the 300G connected to the SAN. In our case, we started off with the same environment that we described in Section 3.7, “Sharing the SAN-based storage to LAN/WAN clients” on page 148.

Once you have verified that the other machines are not attached to the SAN, check that the LAN connection between the 300G and your other machines is working properly. Next, you should make sure the 300G is connected to the SAN and has access to the device you want to share because, in order for the 300G to be successfully configured as a SANergy MetaData Controller, it has to have access to and be the owner of the partition you want it to manage. That means there must be connectivity over the SAN fabric to the storage system containing the partition and a drive letter must be assigned to that partition.

You can verify the status of your connection with the **IBM NAS admin** snap-in tool. Just double-click the snap-in shortcut on the desktop to launch it, then select **Storage-> Disk Management** to review the settings (Figure 5-2). Note the label of the drive you want to share with this MDC. In our installation, the shared SAN disk will be **Disk1** with the label **FAStT200** and the drive letter **T**.

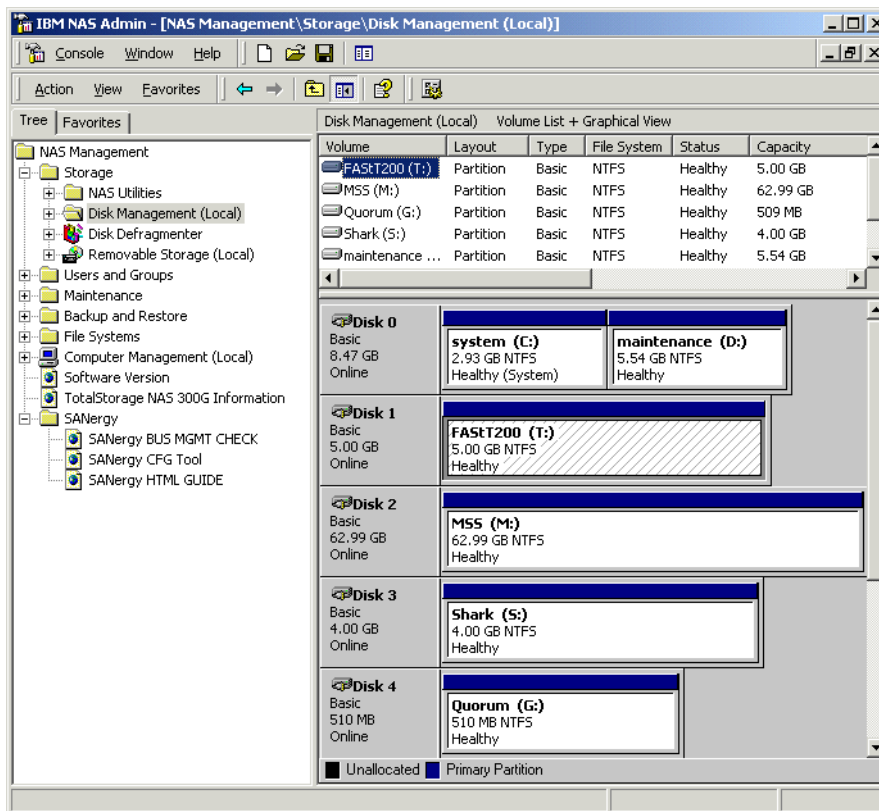


Figure 5-2 Verifying volume access on the 300G

Tivoli SANergy comes pre-installed on the 300G, so you only need to configure it. You can do this by launching the SANergy setup tool from the Windows start menu, but you can also handle this configuration from the same administrative snap-in we used in Chapter 3, “Implementing the IBM TotalStorage NAS 300G” on page 71 to manage the 300G’s connectivity to the FASi200.

Just double-click the **IBM NAS admin** shortcut on the desktop to launch the snap-in. Then open the **SANergy folder** and click on **SANergy Bus Mgmt Check**, as shown in Figure 5-3.

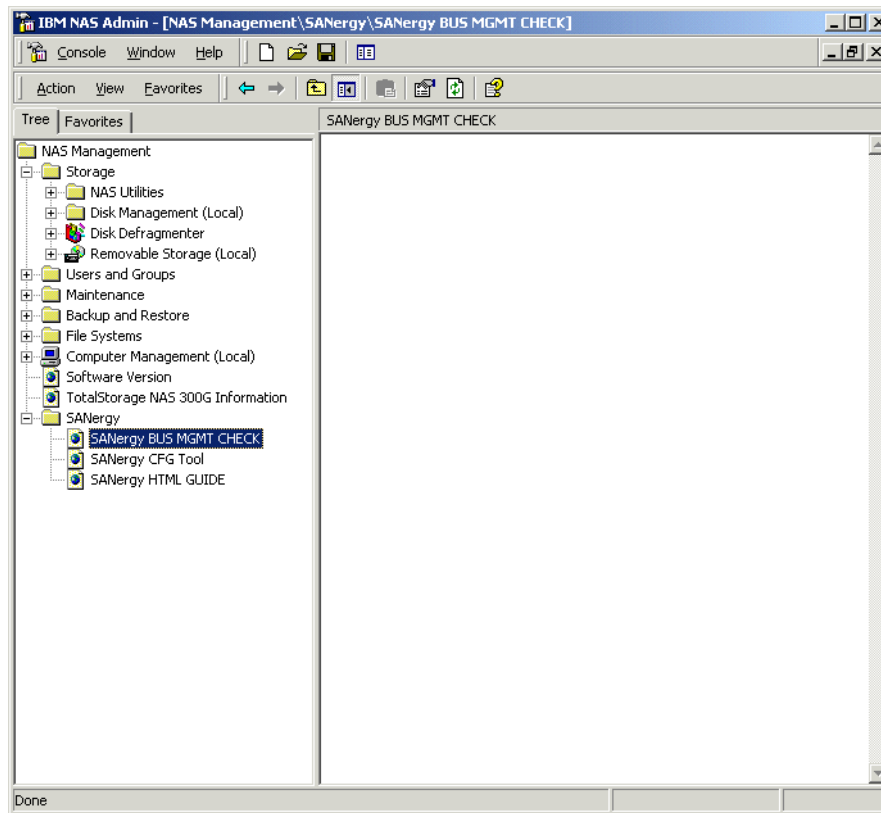


Figure 5-3 SANergy bus management check application

This will launch a command window that prompts you to hit any key. Once you hit a key, a silent install of SANergy will run in the background. When it is done, the SANergy setup tool shown in Figure 5-4 will appear.

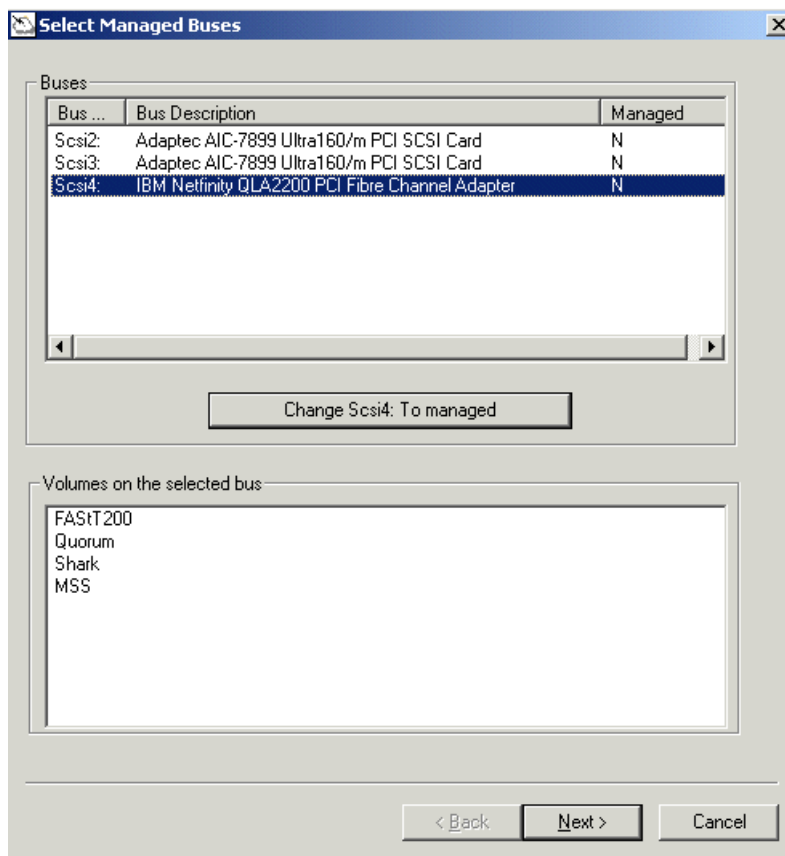


Figure 5-4 Select managed buses for Windows MDC

First select the buses you want to be managed by SANergy. In our case, we have only one Fibre Channel bus, but there might be multiple buses in your environment if you have a custom-configured 300G. Select only those buses which have access to the disk devices containing volumes you want to share.

When you click on the bus in the upper window, you see in the lower screen the volume labels assigned to that specific bus. Only volumes with a drive letter assigned by the operating system are listed. Select the bus that is connected to the storage which is to be shared and click on the **Change** button. After changing the bus from unmanaged to managed, a window will pop up and tell you the 300G needs to reboot. Click **OK** to proceed with the configuration. The restart will occur at the end of the configuration process.

In the next window (Figure 5-5), you can assign the 300G to be the MDC for all of the volumes it owns.

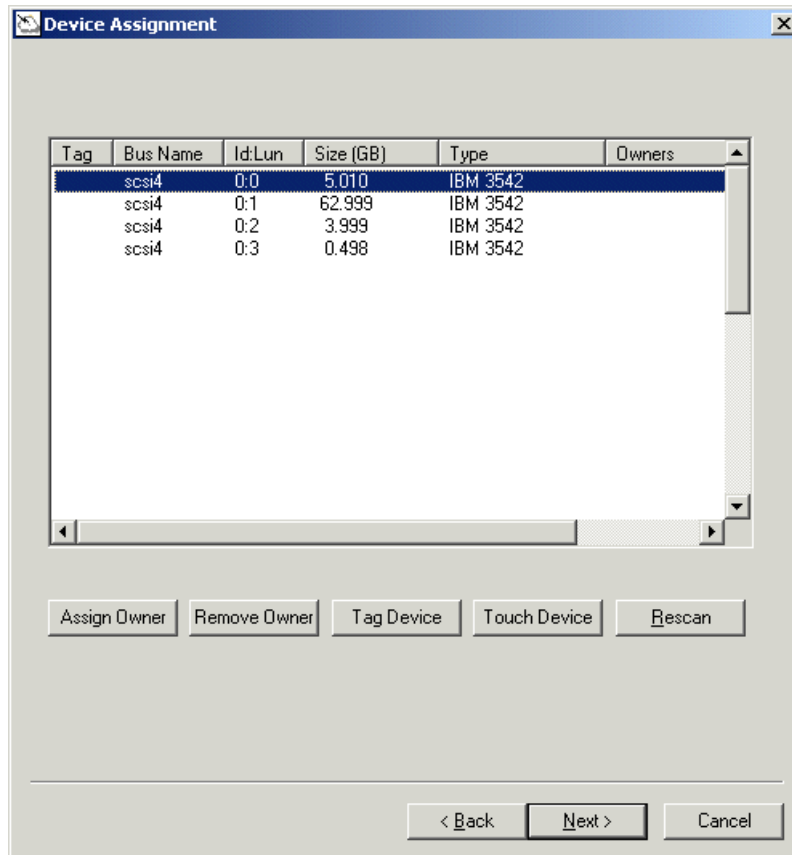


Figure 5-5 Checking for the current device owners on the managed bus

In this list, identify the shared storage devices that are owned by the 300G. If some of the devices have already been assigned to other machines, you can change the owner of a device by first removing the current owner and then assigning the 300G as the new owner. You can do this by selecting the device, clicking the **Remove Owner** button, and then clicking the **Assign Owner** button.

In this example, there are five disks available on our managed bus. We set the 300G as the owner of all of them, as shown in Figure 5-6.

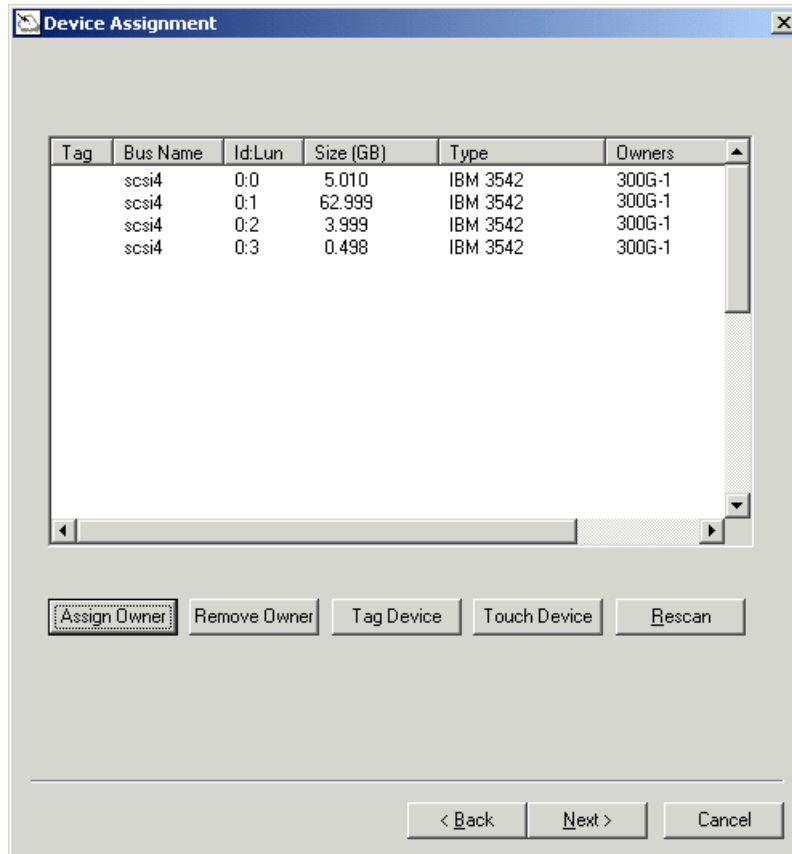


Figure 5-6 Assigning the 300G as the device owner

At this point you can label the devices by clicking on **Tag Device**. The tag has no effect on your device. It is simply a label that SANergy will associate with your device to help you identify it. Another feature of SANergy, the **Touch Device** button, is intended to help you make the mental connection between the physical devices in your storage network and the LUNs listed in this panel. Select a device and click on this button to initiate I/O on the specified disk. This causes the selected disk's LED to flash.

The final configuration step is to make the 300G the MDC for the volumes it will share. Clicking **Next** brings us to the window shown in Figure 5-7.

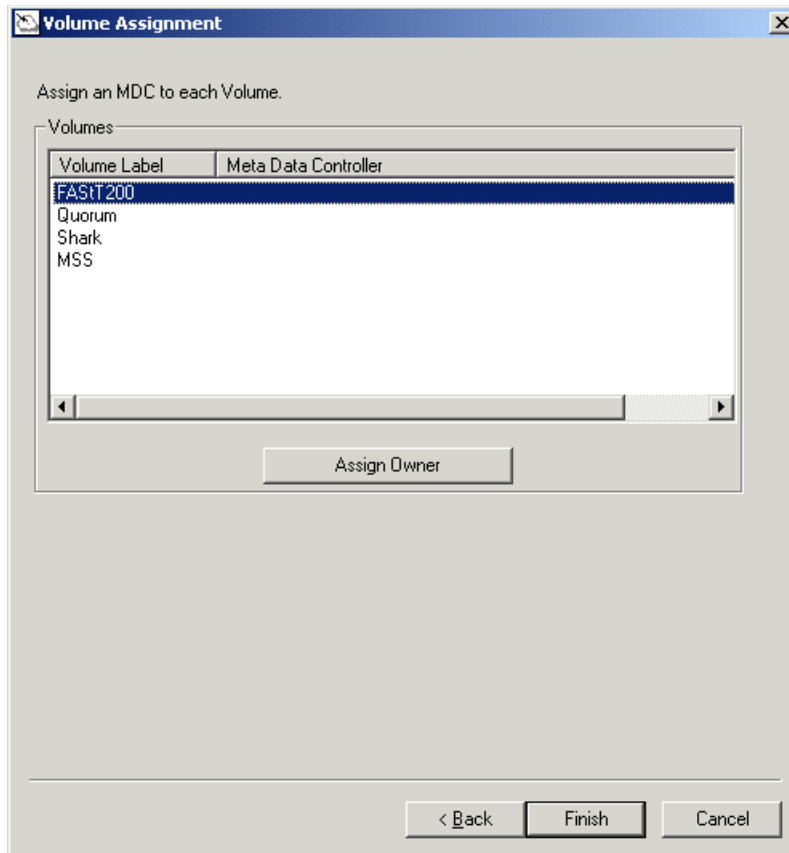


Figure 5-7 Displaying the available volumes

To make the 300G the MDC for each of these volumes, we simply selected them one at a time and clicked on **Assign Owner**. By default, the owner of a storage device will become the MDC for its volume, so the dialog that pops up here will have the name of the device owner in it. In our case, this is the 300G, as shown in Figure 5-8.

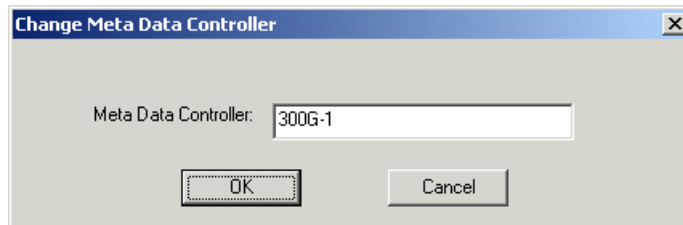


Figure 5-8 Selecting the 300G to be the MDC for the selected volume

Important: You may have SAN storage devices attached to the 300G (the MDC) that you do not want to share for whatever reason. Even so, you still need to assign ownership of those volumes to the 300G. Otherwise, it will lose access to those volumes.

We assigned the 300G to be the MDC for all of the volumes it was sharing, as shown in Figure 5-9.

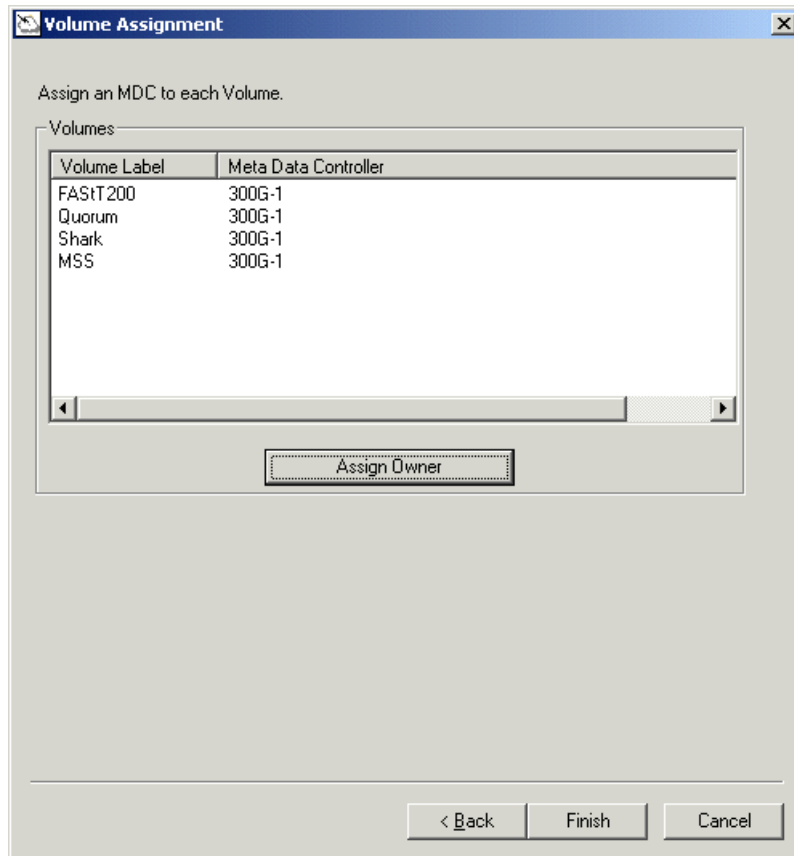


Figure 5-9 Making the 300G the MDC for the shared volumes it owns

After you have completed the volume assignment dialog box by clicking on **Finish**, the system will ask you to reboot. After the reboot, the 300G will be the SANergy MDC for the volumes you assigned to it. You can modify the configuration settings at any later date by re-running the SANergy Setup Tool.

It is a good idea to go ahead and run the tool right now and verify the installation. The main window of the Setup Tool (Figure 5-10) contains tabs with the menu items you have configured during installation (**Select Managed Buses**, **Volume Assignment**, and **Device Assignment**). Furthermore there are two additional tabs, called **Options** and **Performance Tester**, that allow you to tune, customize, and test the performance of the SANergy implementation. An easy way to verify the MDC installation at this stage is to test the access to the SAN storage device by selecting **Performance Tester**.

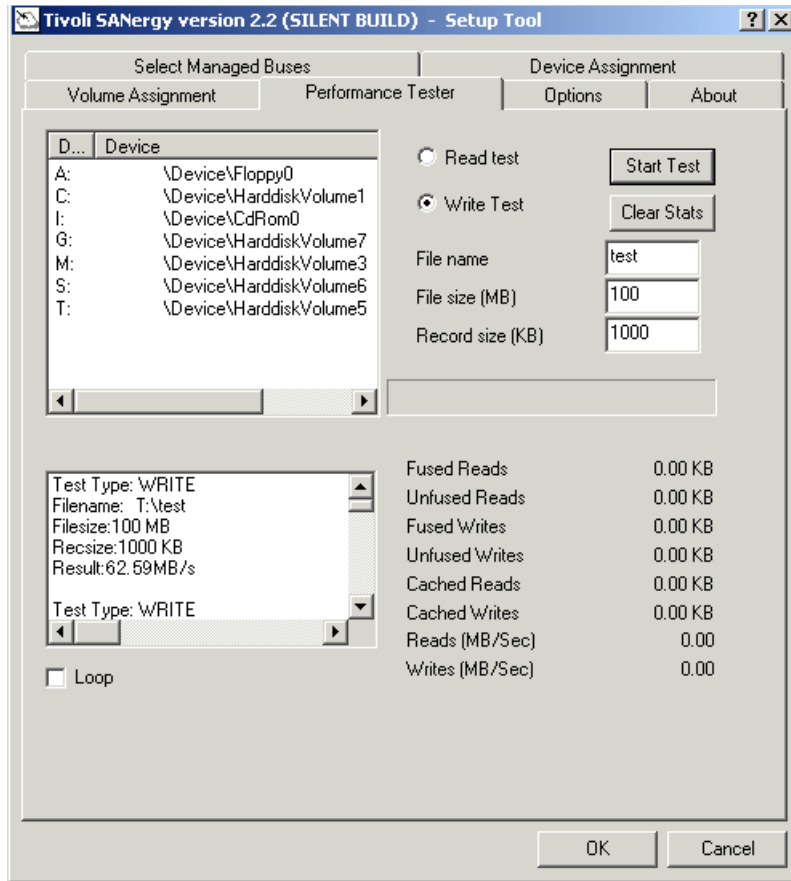


Figure 5-10 Performance Tester tab in the SANergy Setup Tool on the 300G

In the upper left window, select the device owned by the 300G (the SANergy MDC) which will be shared to your other SANergy hosts. To make sure that SANergy can write to this device, select the **Write Test** option, specify a file size and file name, and start the test with the **Start Test** button. This will run a single write test, or you can run a test in a continuous loop by checking the **Loop** box.

During the test, SANergy will try to create and write the file you defined to the device. The bottom left hand window shows the result of the test. You will not see any statistics on the right hand side (for example **Fused Writes**), as these are only collected for SANergy host access (see Figure 5-20 on page 257 for an example). If this test is successful, the 300G has been successfully configured as the MDC for the shared volumes and you are ready to being installing SANergy on your other machines.

While this should be a straightforward procedure, if you do encounter any problems, please consult the *Tivoli SANergy Administrator's Guide*, GC26-7389 for assistance.

Note: The performance shown in the Performance Tester seems to be mainly dependent on the I/O throughput of the storage device. In addition, cache settings on the storage device can have a significant impact on data flow, because the first portions of (cached) data have high performance, but later throughput will fall back to the speed of physical disk performance after the cache is filled. To determine what kind of performance you can expect from hosts accessing the shared storage, run this test with a file size bigger than the cache settings on the storage device.

5.3 Configuring your other machines to use SANergy

Now that we have set up the 300G as our SANergy MDC, both Windows and UNIX hosts can be easily configured to share data through it (Figure 5-11). In order for the client machines in your network to become hosts to your 300G MDC, those machines must meet the following prerequisites:

1. They must have an IP connection to the MDC (in our case, the 300G).
2. They must be able to map/mount the storage that is being shared from the MDC.
3. They must have their own HBA card and have access through the fibre network to the storage that is being shared through the MDC.

We already set up IP file sharing to those hosts in Section 3.7, “Sharing the SAN-based storage to LAN/WAN clients” on page 148, so that takes care of steps 1 and 2. For step 3, you must also install and configure HBA cards in the other machines that you wish to give high speed access to the SAN-based shared storage. This is a simple matter of installing the card and the drivers for it. Once this is done, you also have to make sure that your machines have access through the fibre network to the storage that is being shared.

This process is a little different between UNIX and Windows clients, so we describe them separately. The method we used for setting up the Windows clients is explained in Section 5.3.1, “Installing and configuring SANergy Windows NT/2000 hosts” on page 249, while the method for UNIX clients is explained in Section 5.3.2, “Installing and configuring SANergy UNIX hosts” on page 257.

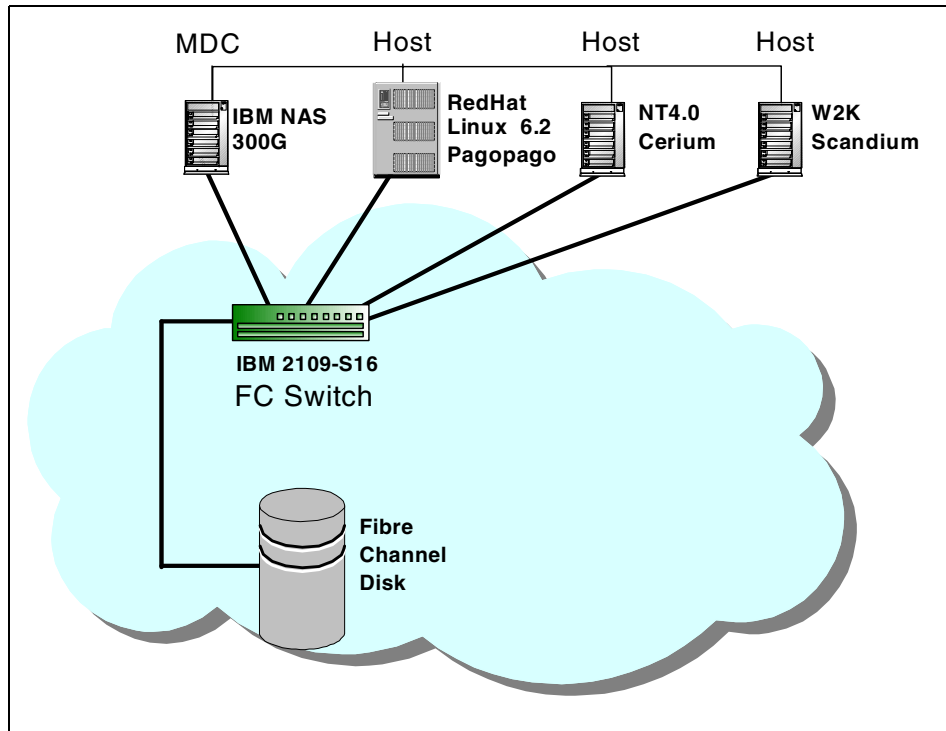


Figure 5-11 The 300G configured to share SAN-based storage as an MDC

5.3.1 Installing and configuring SANergy Windows NT/2000 hosts

This section documents the installation of SANergy on both a Windows NT (machine name: Cerium) and a Windows 2000 (Scandium) machine, as the procedures are essentially identical. We configured our machines to be SANergy hosts for a partition owned by the MDC (the 300G) which we configured in Section 5.2, “Configuring the 300G as a SANergy MDC” on page 238.

As we have mentioned several times now, before you install SANergy, it is important to make sure that the machine has connectivity to the shared SAN device. You can do this easily enough by checking the Windows **Disk Administrator** tool to see if the specific disk is listed. On Window NT, go to **Start-> Programs-> Administrative Tools-> Disk Administrator**. On Windows 2000, choose **Start-> Programs-> Administrative Tools-> Computer Management** and in the tree view select **Storage-> Disk Management** to see your currently accessible drives. You should have access to the drive, but you should not assign a drive letter to it, because the partition is already owned by the 300G.

Unfortunately, as soon as you establish physical connectivity from the host to the shared storage device and reboot, both Windows NT and 2000 will immediately assign a drive letter to that new device automatically without even bothering to ask if you wanted a new drive or not. In Figure 5-12, you see the **Disk Administrator** tool running on Windows 2000 before we connected the machine named **Scandium** to the SAN. As you can see, only the local hard drive is visible.

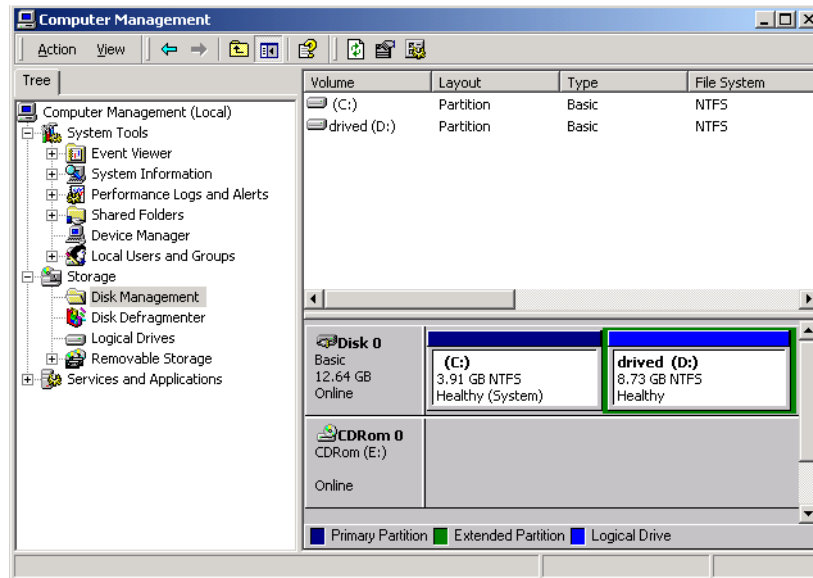


Figure 5-12 Disk Administrator view before connecting the host to the SAN

After we connected the Fibre Channel cable from the HBA to the switch and rebooted, the system automatically added another device and provided it with a drive letter as shown in Figure 5-13. This is because the MDC (the 300G) has already formatted that partition, so the volume is instantly recognized as a usable file system by this host.

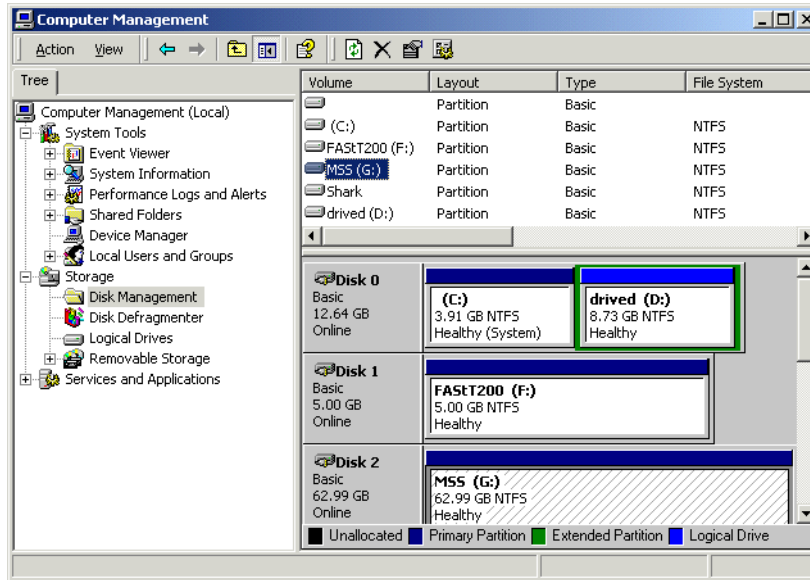


Figure 5-13 Disk Administrator view showing the new device (attached to SAN)

We do not want this disk to be assigned a local drive letter because that would undermine the security offered by SANergy. To prevent Windows from accessing and destroying data on this disk without the MDC's permission, we will unassign the drive letter. On our Windows 2000 machine, we right-clicked each drive and chose **Change Drive Letter and Path...** as shown in Figure 5-14. (On Windows NT, you would right click on the specified disk and choose **Assign Drive Letter** and then select the option **Do Not Assign Drive Letter**.)

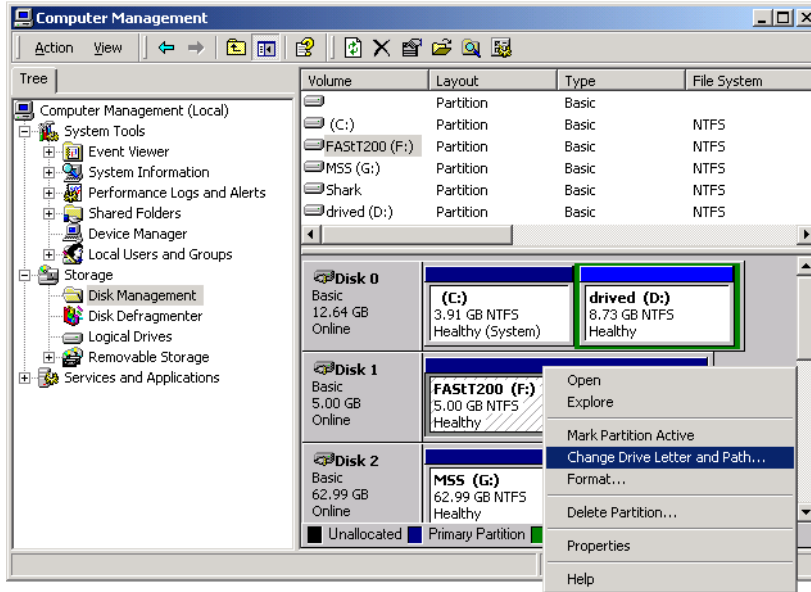


Figure 5-14 Unassigning drive letters

The action will take effect immediately, as shown in Figure 5-15, and should not require a reboot.

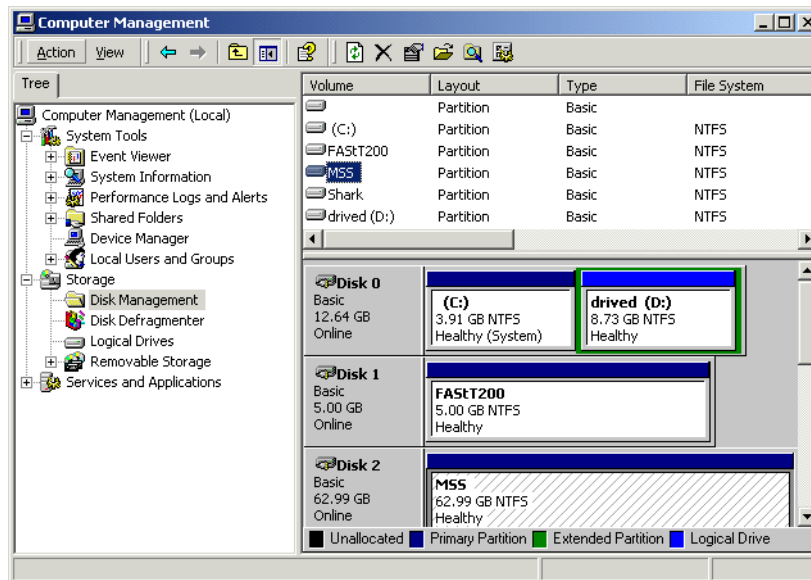


Figure 5-15 Disk Administrator view after unassigning drive letter

In the Explorer view, you can now verify that you no longer have access to the disk as the disk icon for this device will have disappeared. Now you are ready to begin installing SANergy.

You will first install the code from the original product CD which will include the required product key information for the license. Depending on release levels, you may then have to install a patch. You can find information on available patches and download them from the Web site:

<http://www.tivoli.com/support/sanergy/maintenance.html>

Before installing any patch, you should check the README file provided to see if there are any special instructions associated with it.

Once you put the SANergy product CD into the CD-ROM drive, the installation setup should start automatically. Select the **Tivoli SANergy** option.



Figure 5-16 Initial SANergy installation screen

Follow the setup program and perform these steps:

1. Acknowledge the license agreement.
2. Choose your installation directory.
3. Choose Start program folder name.
4. Type in user information.

A more detailed description of this procedure can be found in the *Tivoli SANergy Administrators's Guide*, GC26-7389. After these preliminary steps, you should see the dialog shown in Figure 5-17.

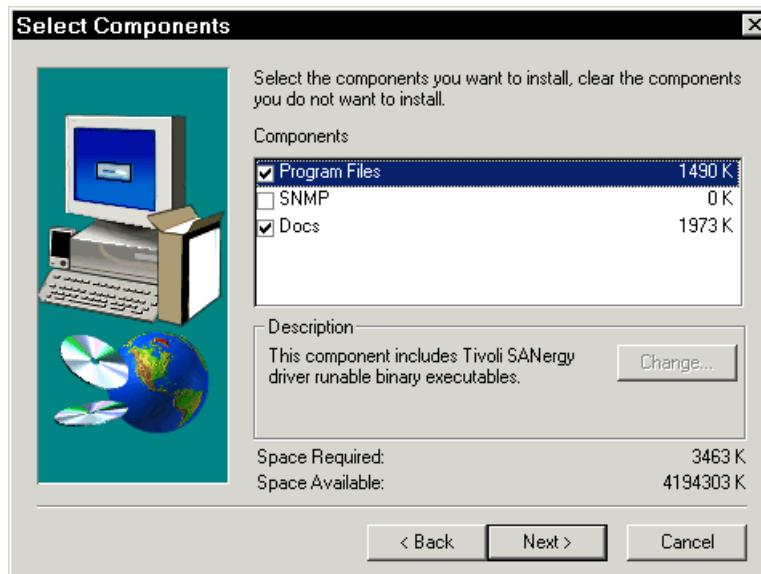


Figure 5-17 Select components to install

Here you can choose the specific program components to be installed. The SANergy software and documentation files are preselected by default. To enable SANergy for an environment monitored by SNMP, you need to explicitly install the SNMP package by checking the box next to it. The default setting is to *not* install it.

Once the install completes, the configuration process begins. The only difference between the configuration for this SANergy host and the configuration of the MDC which we performed on the 300G is that we will not give this machine ownership of any volumes and will not assign to be the MDC for them either. An example of this is shown in Figure 5-18.

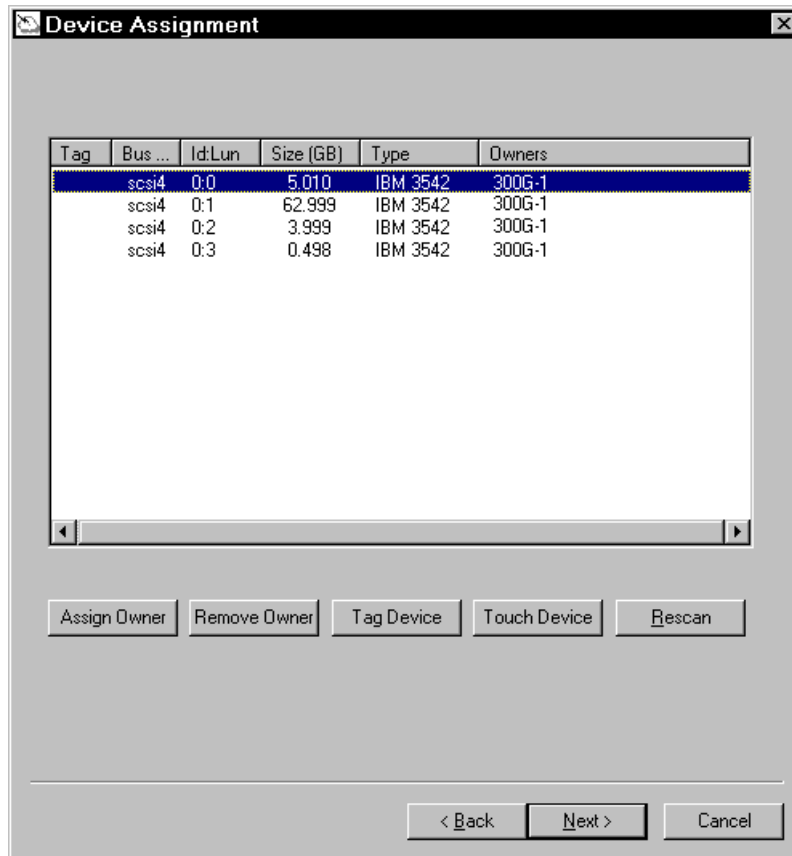


Figure 5-18 SANergy configuration on a Windows host

If everything is going well, the configuration should show that the 300G already owns the shared device. Once you click **Next**, it should also show the 300G as the MDC for the shared volume (as shown in Figure 5-19). Do not change this. If for some reason the 300G is not showing up as the device owner, or if the device is not showing up, then you probably have SAN connectivity problems which you will need to resolve before you continue.

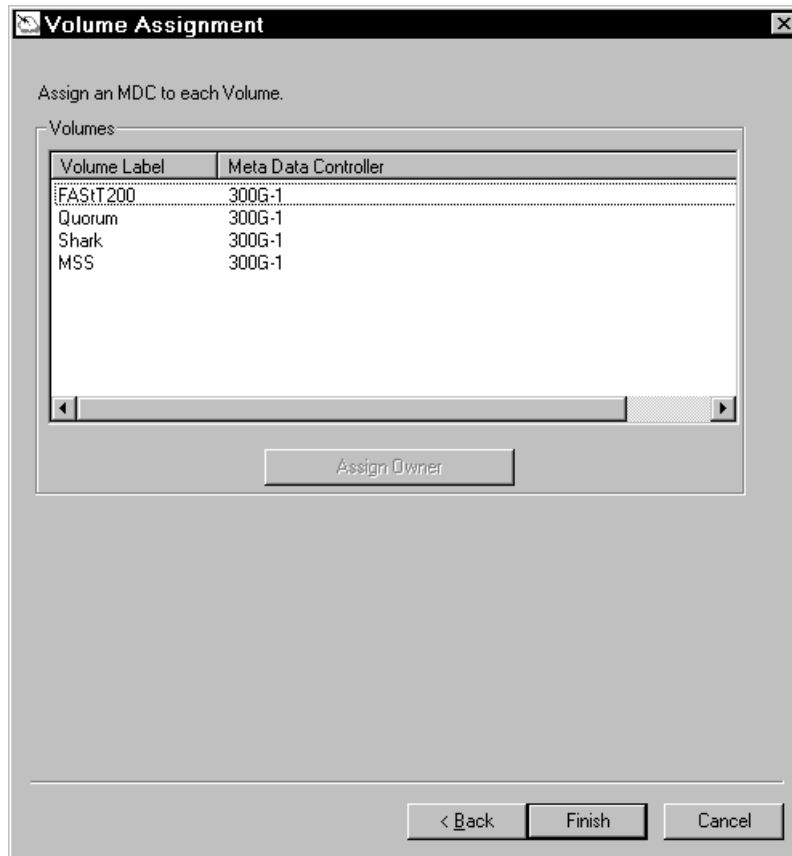


Figure 5-19 The 300G is already identified as the MDC for the shared volume

Once you click **Finish**, the installation will complete and the host **Cerium** will be able to fuse the volume being shared by the 300G. From now on, when Cerium tries to read or write data to the shared volume, the 300G will redirect its I/O so that it communicates with the storage device over the SAN. Just as with the MDC configuration, we now need to test and make sure that the host can directly read and write data to the shared volume over the SAN. Once again, open **Start-> Programs-> SANergy-> SANergy Setup Tool** and select the **Performance Tester** tab as shown in Figure 5-20.

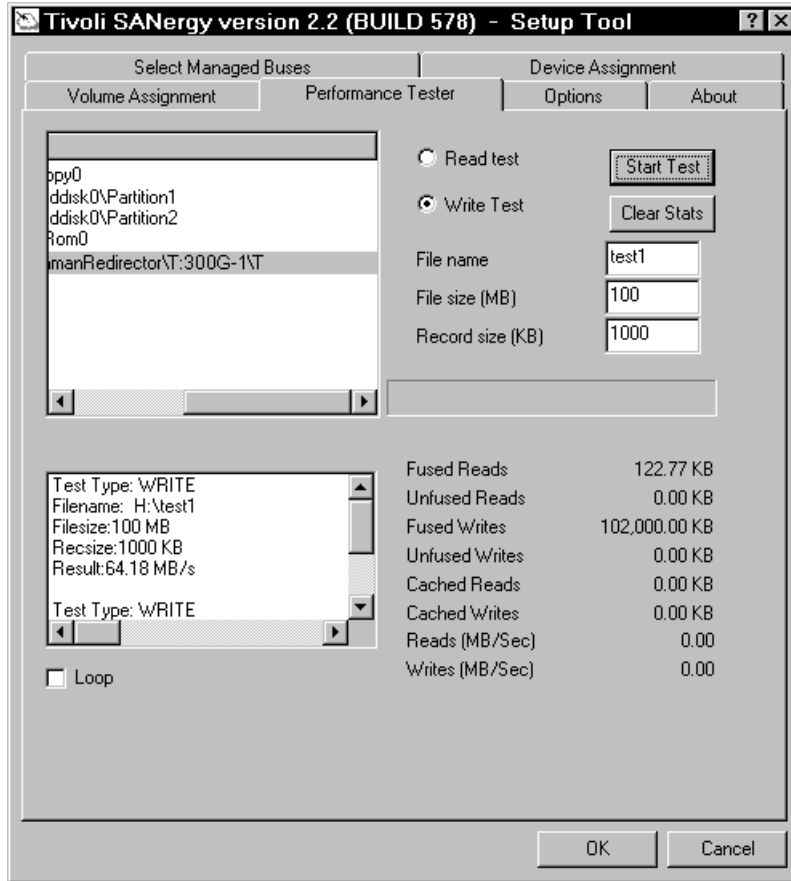


Figure 5-20 Verifying the host installation with the performance tester

If the shared drive on the MDC is properly mounted on the host, you will see the mapped partition in the upper left window. Highlight this partition to perform a write test on the device. Select a dummy file name and a representative file size and press the **Start Test** button. On the bottom right side you can see the measured values for fused (SAN-redirected I/O) data throughput for this attempt. If you do not see any updated statistics for reads and writes, this means that the SAN is not being used for access.

5.3.2 Installing and configuring SANergy UNIX hosts

This section will cover the installation and configuration of SANergy on a UNIX host. This example will be given using the Linux client and the GUI interface, where appropriate. As a prerequisite for installing SANergy on UNIX, you should have a Netscape browser with JavaScript support installed.

Install the SANergy base code

Insert and mount the SANergy CD-ROM (see your operating system documentation or `mount` man page for details on how to mount a CD on your version of UNIX). Change to the directory where you mounted the CD-ROM, and from there change to the directory containing the Unix install script:

```
cd file_sharing/unix
```

Run the install script located in the `./install` directory.

You will be prompted to validate the operating system detected. The installation runs automatically and starts the configuration tool in your Netscape Web browser. The initial screen will ask you to enter or validate the detected product key. Choose **Accept** to proceed to the main screen (Figure 5-21).

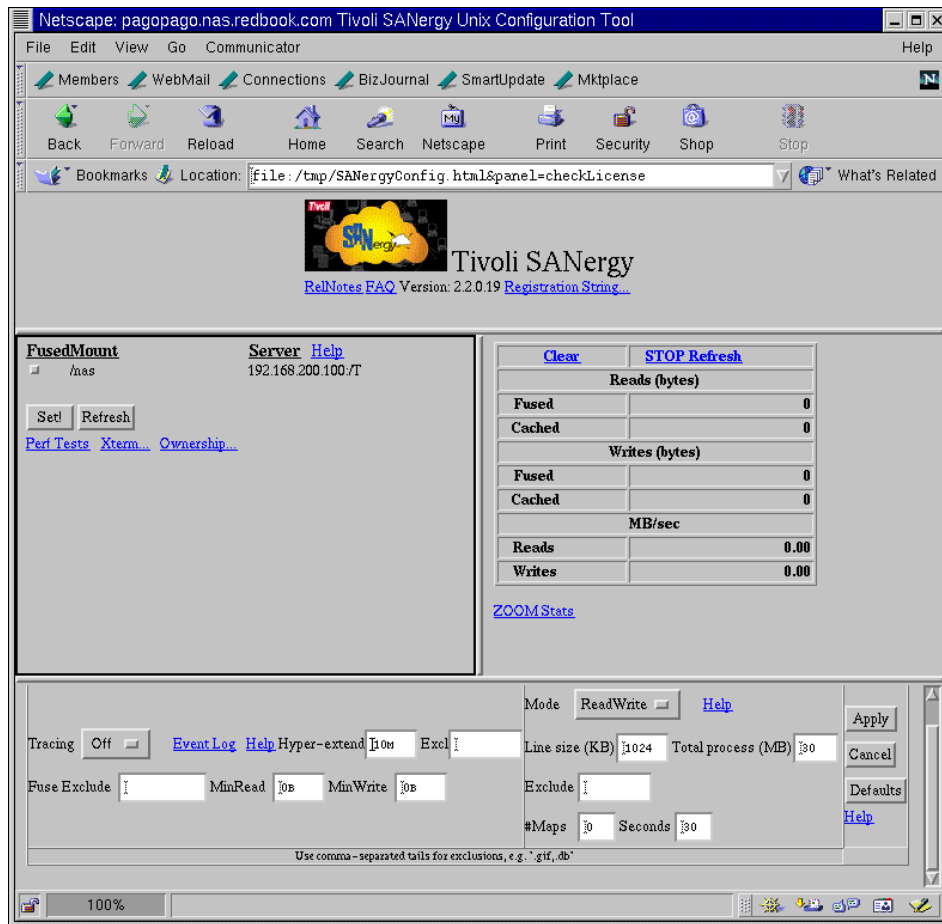


Figure 5-21 Tivoli SANergy main screen on Linux

For more detailed examples of this basic installation, please refer to the *Tivoli SANergy Administrator's Guide*, GC26-7389.

Select the mount points to fuse

The upper left quadrant of the SANergy main window contains a list of all NFS shares currently mounted on the system (see Figure 5-22). Select the volume that is being shared from the 300G and click **Set!**.

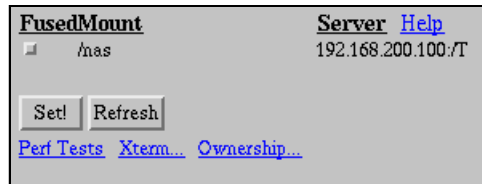


Figure 5-22 Mount point before fusing

If you are not used to Linux GUIs, it is often hard to tell if something is checked or not, so be careful here. Once you have set a mount point, the button beside it will be recessed, but there is no other indication that the file system has been fused for use by SANergy. To make this distinction clearer, you need to compare the difference in the unset mount point shown in Figure 5-22 and the set mount point, as shown in Figure 5-23.

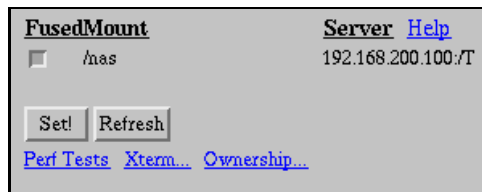


Figure 5-23 Fused mount point

Just as with the MDC and Windows host configurations, you now need to verify that SANergy is working correctly. Again, you can use the performance tester to do this, but this version of the configuration GUI does not have a separate tab. Instead, you just need to click the **Perf Tests** hyperlink (right below the **Set!** button). Next, highlight the mount point to be tested from the **Select Volume** list box (see Figure 5-24), and then click the **Write** button to initiate a write performance test. When the test is complete, the panel will show the results of the test. Also note that the statistics on the upper right quadrant will refresh to indicate the new values.

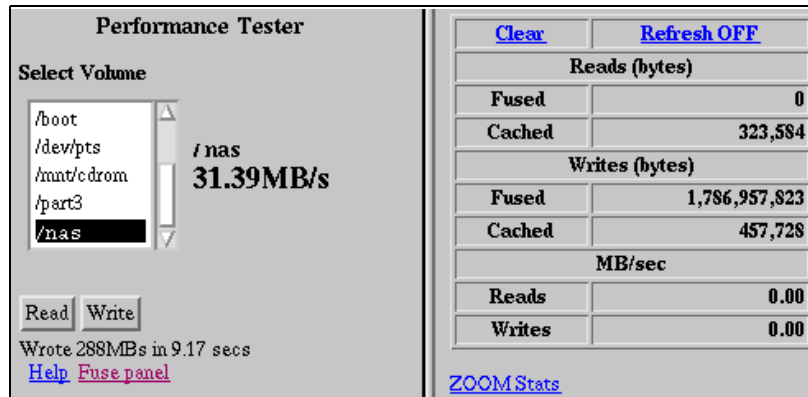


Figure 5-24 The SANergy Performance Tester running on a Linux host

If these tests are successful for all of the mount points you set, then you have successfully installed and configured SANergy on your Linux host.

Using SANergy with a process

In order to run a process or application which will access SANergy managed file systems, it is necessary to first set some environment variables so that the SANergy libraries are used. In the SANergy installation directory (which defaults to /usr/SANergy) there are two scripts to help you set these environment variables. If you are running a C-shell, use SANergycshsetup. If you are using the Korn shell, use SANergyshsetup. We recommend including the appropriate invocation scripts in the startup scripts for all applications requiring access to SANergy-managed file systems.

5.4 Using SANergy on the 300G Model G25

While the previous sections discussed running SANergy on a single node 300G, as we discussed in Chapter 4, “Clustering for high availability” on page 167, the 300G also comes in a dual node configuration (the Model G25).

It is often desirable to configure a SANergy MDC for high availability. Where the data being shared or applications being served are highly critical to the enterprise, it is vital that access to them over the SAN is always available by eliminating single points of failure. Making an application or service highly available typically means duplicating the important resources so that the impact to the end users of any planned or unplanned failure in an individual system or subsystem failure is minimized. Workload on a failed system can be transferred or passed over (transparently or nearly transparently) to another system which is still available.

While SANergy actually has its own special high availability component, if you are working with the clustered 300G Model G25, you will be using the native high availability product of the Windows Powered OS, the Microsoft Cluster Server, to provide high availability to your MetaData Controller (MDC).

This section documents the steps necessary to install SANergy on the 300G Model G25. We assume that Microsoft Cluster Server (MSCS) is already up and running smoothly on your dual-node 300G. If this is not the case and you need help with setting up MSCS, please see Chapter 4, “Clustering for high availability” on page 167 for step-by-step instructions.

5.4.1 Base configuration

Figure 5-25 shows our basic testing and validation configuration. Our cluster was named **300G** (demonstrating our creativity) and it consisted of the two individual nodes **300G-1** and **300G-2**. Our SANergy host in these tests was a Windows 2000 Advanced Server machine named **Scandium**. Each of these machines could see four drives in our SAN. One of these drives was used as the quorum resource for the MSCS cluster (Quorum). The other three were the same shared volumes we set up in Chapter 3, “Implementing the IBM TotalStorage NAS 300G” on page 71. These volumes, creatively named to match the device they were shared from (FAStT200, MSS, and Shark), were used as SANergy shared volumes.

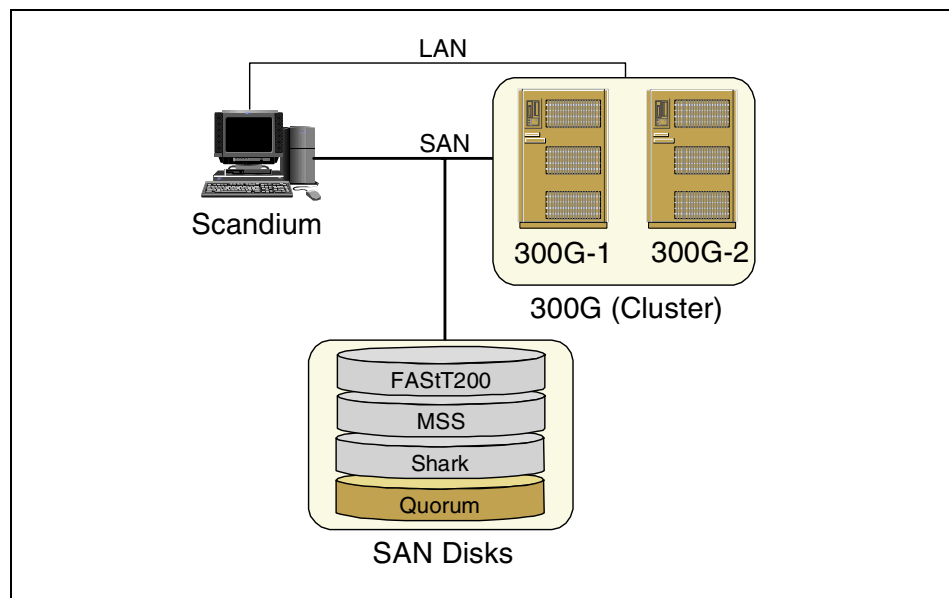


Figure 5-25 Testing cluster configuration

MSCS defines one of its shared drives as a *quorum resource*. This resource contains data vital for the operation of the cluster and is always reserved by the node hosting the cluster group. It is generally recommended that the quorum resource be an entire volume, instead of just one partition on a volume. This prevents potential problems that could occur if the quorum resource had to be moved between nodes - requiring the entire volume to be relocated.

The other shared disks are used by different applications hosted on the cluster. Normally, these are defined as physical disk resources and are controlled by MSCS. However, the SANergy MDC normally controls the disks it shares. This means that to integrate MSCS and SANergy, we have to take special steps to convince these two pieces of software to work together.

5.4.2 SANergy and MSCS: mixing two domineering personalities

MSCS works by allowing you to define resources of various types that require fail-over protection. These resources can be grouped together so that they operate as a single unit. An example of resource types are virtual server names (which can be accessed from the network just like physical machines), disks and applications. A group could include a server name, an application and the disks that support the application. MSCS will guarantee the availability of this group of resources by monitoring them and making certain they are operating together on one of the physical machines in the cluster (a node).

In addition to making certain that the group of resources is always available, MSCS ensures that they are not running on more than one machine at any given time. This is necessary because if a group were active on more than one machine, many sorts of problems and failures would occur. It accomplishes this by preventing the mounting of a physical disk resource on more than a single machine at a time by using the SCSI reserve/release commands.

Due to the fact that MSCS attempts to limit access to physical disk resources to a single host, a special type of resource must be used for those disks which will be shared through SANergy. This allows SANergy to fill its normal role of allowing sharing of its managed resources while preventing data corruption.

The software component that makes SANergy cluster-aware is the SANergy MSCS module. This module is included with the base SANergy package and enables the definition of SANergy volumes within MSCS.

5.4.3 Installing and configuring SANergy with MSCS

The following steps describe how to install and configure MSCS and SANergy.

Prevent MSCS from managing physical disk resources

It is critical that SANergy and MSCS do not both attempt to manage the same disks. To accomplish this, just delete the physical disk resources from MSCS that you wish to be shared by SANergy. Before deleting these resources, be certain to note which groups the physical disk resources were members of as well as any dependencies related to those resources. You will need this information later when you re-define these resources as SANergy Volumes. You can delete the physical disk resources by right-clicking them in **Cluster Administrator** and selecting **Delete** as shown in Figure 5-26.

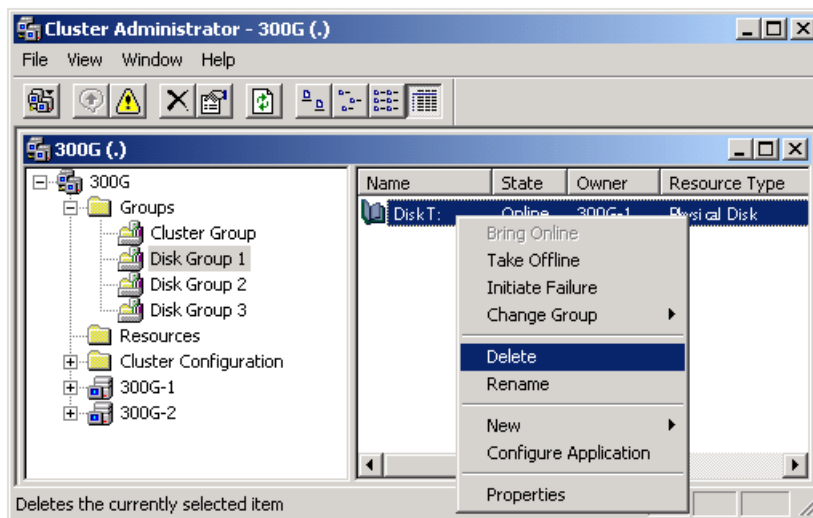


Figure 5-26 Deleting Physical Disk resources from MSCS

Important: Do **NOT** delete the physical disk resource for the quorum disk. This resource should always be managed by MSCS. For SANergy purposes, it is not necessary to touch any of the resources inside the Cluster Group.

Make certain all nodes have access to SAN disks

Both of the nodes in the 300G cluster must have access to the volumes to be shared. The other machines in your network that will be using SANergy should also be able to see those volumes. Configure your SAN components so that the machines have access and can mount the volumes. Once this is done, validate that the volumes are visible in the *Disk Management* applet. This process is the same as the one we described for SANergy running on a single-node 300G. Please see Section 5.2, “Configuring the 300G as a SANergy MDC” on page 238 for the complete details.

All nodes in the cluster should use the same drive letters to represent each disk. This is necessary to later build other cluster resources that access the disk, such as file share resources. We describe this set up step in Section 4.2.3, “Update drive letters” on page 181.

You will also need to note the label and volume serial number assigned to the volumes to be shared later. Open a command prompt (select **Start-> Programs-> Accessories-> Command Prompt**). Issue the `vol` command for each drive. Note both the label and volume serial number of the disk. If the `vol` command gives an error, you may need to run this command from whatever machine currently owns the volume. See Example 5-1.

Example 5-1 Issuing the vol command

```
C:\>vol t:
Volume in drive T is FASTT200
Volume Serial Number is B4C5-083B
```

Important: The `vol` command returns the volume serial number in a four-character—dash—four-character pattern (####-####). When MSCS asks for the volume serial number later, do **NOT** supply the dash in the middle. Just give it the letters and numbers in one long string (#####). If you supply the dash, MSCS will not be able to bring the volume online.

Install SANergy on both cluster nodes

Now that we have the cluster all set up, our next move is to install the SANergy base-code on both cluster nodes. With one important exception, this is exactly like installing SANergy on a single-node machine, so you can just follow the procedures we used in Section 5.2, “Configuring the 300G as a SANergy MDC” on page 238. The difference between installing on a dual-node and single-node machine is the name used to identify which MDC manages a volume.

The name assigned to the MDC for volumes to be shared is critical. By default, SANergy will assign the current owner of the device to be the MDC, but you have to change this in order to use SANergy in a cluster. To change the volume assignments, follow these steps:

1. Select the volume you want to change and click **Assign Owner**.
2. In the pop-up dialog, highlight all the text, hit the **Delete** key, and click **OK**.
You have to delete the current owner first, then assign a new owner because SANergy does not support reassigning an owner in one step.

Note: You have to delete the owner from the machine that currently owns the volume.

3. Select the volume you want to change again and click **Assign Owner**.
4. In the pop-up dialog, type in the appropriate special name for this volume (see Figure 5-27):
 - For the quorum drive, use the special name **?FREE** to prevent SANergy from attempting to manage this disk. The MSCS cluster must manage the quorum resource.
 - For the remaining disks (at least for all of those that are to be shared by SANergy), use the special name **?CLUS** to inform SANergy that these are cluster resources and are not owned by any specific machine. SANergy will then automatically use whichever node currently owns the SANergy Volume resource as the MDC for that volume.

Important: The special names **?FREE** and **?CLUS** must be entered in all caps as shown. Even though this is Windows, these names are case-sensitive.

We recommend that you first install SANergy on Node 1 and make the volume assignments there. When you are done, SANergy will ask you to reboot the machine, but we recommend that you not reboot just yet (there is one more step you will need to perform first). When you have finished installing SANergy on Node 1, go ahead and install it on Node 2, but again, do not reboot yet.

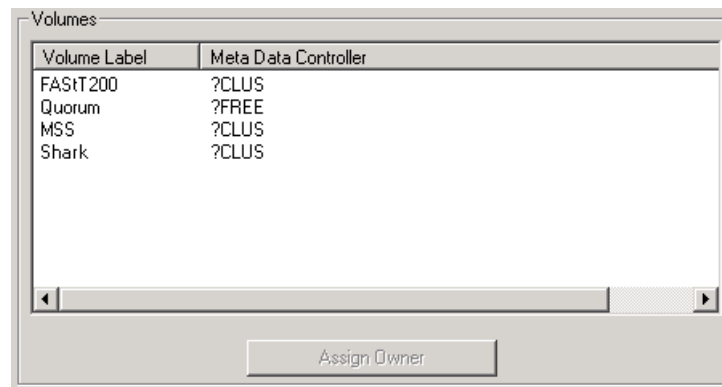


Figure 5-27 Special names on volume assignment

Install SANergy MSCS

Now you need to install the SANergy MSCS component on both nodes. To do this, navigate to C:\ibm\sanergy and double-click the file named **SANergyMSCS 1.0 Build 2.exe**. We recommend that you install on Node 2 first, then install on Node 1. Wait until you have installed this component on both nodes before rebooting. The procedure for the two installs is slightly different as we explain hereafter.

Important: It is very important that you install the component in the same location (drive and folder) on both nodes. We recommend that you accept the defaults on both nodes.

On the first node you are installing this component on (Node 2), select **Cluster Node Setup** as the installation type (see Figure 5-28). Once you complete this install wizard, you will be finished with Node 2. Now go ahead and start the install on Node 1.

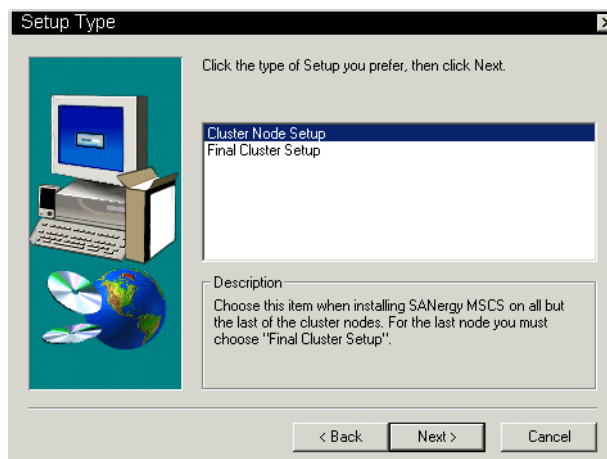


Figure 5-28 Installing SANergy MSCS on the first Node

When installing on the last node (Node 1), select **Final Cluster Setup** as the install type (see Figure 5-29).

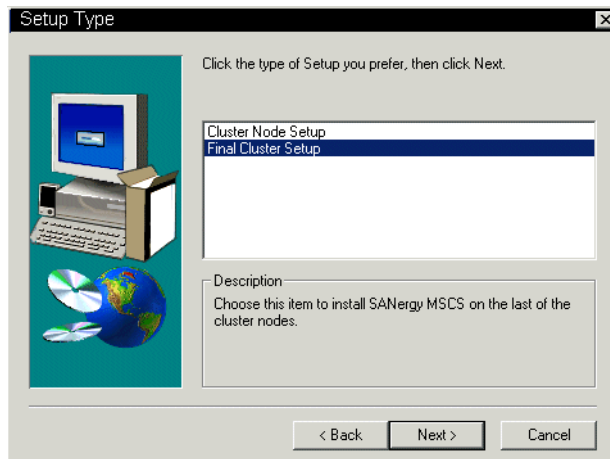


Figure 5-29 Installing SANergy MSCS on the last node in the cluster

There is a bit more to the installation on the last node. After you have completed the initial install wizard, a window will be displayed to register the SANergy Volume resource type to MSCS. Select **Install** from this window to proceed (see Figure 5-30).



Figure 5-30 Last window of SANergy MSCS install on final cluster node

Once the installation on Node 1 is complete, you will need to reboot both nodes. We recommend shutting Node 2 down, rebooting Node 1, and not bringing Node 2 back up until you have tested the installation.

Validate that the SANergy Volume resource is available in your cluster by listing the **Resource Types** from the Cluster Administrator (see Figure 5-31). Start Cluster Administrator by double-clicking the **IBM NAS Admin.msc** shortcut on the desktop and then selecting **Cluster Tools -> Cluster Administration**.

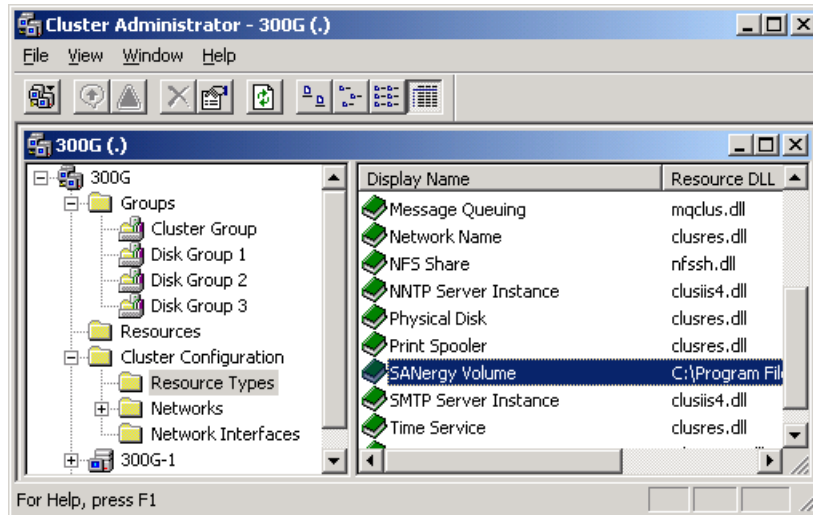


Figure 5-31 Validate that the SANergy Volume resource type is available

If everything looks okay so far, go ahead and bring Node 2 online. Once it is up and has joined the cluster again, you are ready to define your SANergy volumes for sharing.

Define SANergy Volume cluster resources

To allow shared SANergy volumes to be shared from an MSCS cluster, you must define them as a resource to a cluster group. Start Cluster Administrator and right-click the disk group that you want to add the volume into. Select **New -> Resource** from the pop-up menu (see Figure 5-32).

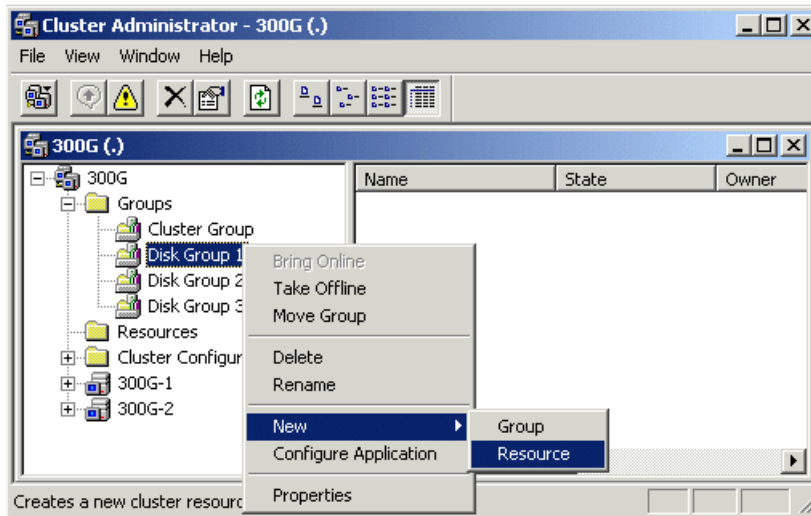


Figure 5-32 Adding a new cluster resource

The **New Resource wizard** will then start. The first dialog will ask you to name the new resource (see Figure 5-33). This can be any valid resource name, since SANergy does not require a special naming convention. In our examples, for clarity, we gave the resource the same name as the volume's label. Make sure you set the resource type to **SANergy Volume**.

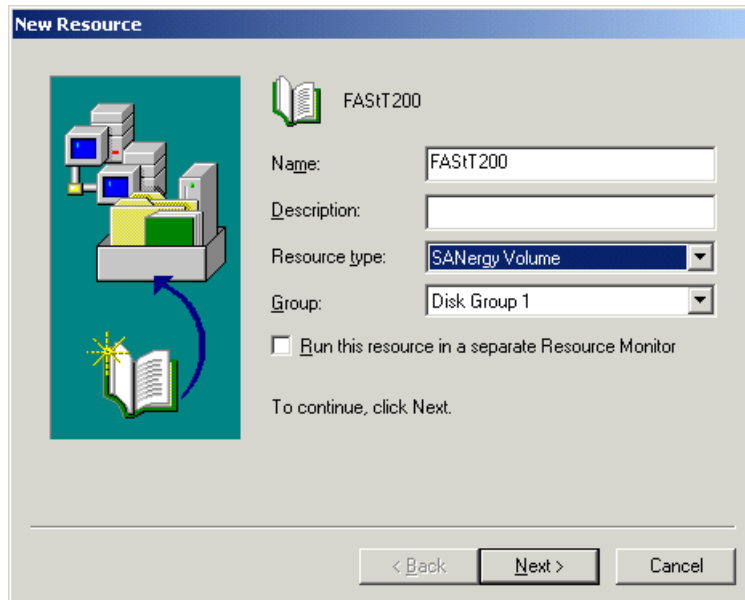


Figure 5-33 Defining a new SANergy Volume resource

You will now see a window which allows you to specify which nodes can host the resource. By default, all nodes are selected (so you should see both Node 1 and Node 2). Leave both nodes in this list and then click **Next**.

The next window asks you to specify any dependencies that this resource may have on other resources. We will handle this later, so do not assign any dependencies now. Just click **Next** and move on.

Finally, the New Resource wizard will prompt you for information about the volume being shared (see Figure 5-34). You need to enter the volume label and serial number which you gathered before installing SANergy (see “Make certain all nodes have access to SAN disks” on page 263).

Important: Just in case you missed this warning before, do **NOT** include the dash in the volume serial number. Just type in the 8 letters or numbers without the dash. If you include the dash in the serial number, MSCS will not be able to bring the volume online.

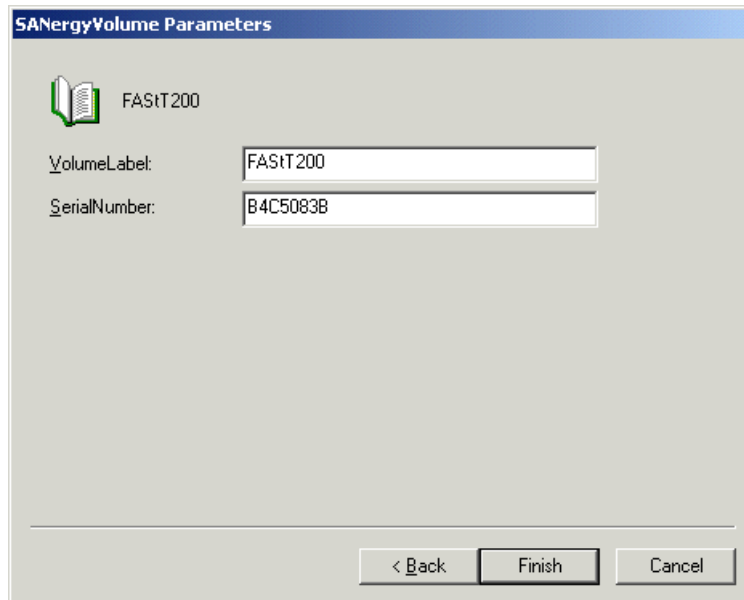


Figure 5-34 Setting SANergy Volume resource parameters

Now click **Finish** to complete the task of creating the new SANergy Volume resource. Repeat this process to create a resource for each volume you wish to share.

Remember: You should not create a SANergy Volume resource for the quorum disk, as that disk is exclusively managed by MSCS.

After you have defined the resources, you can test them by bringing them online. In Cluster Administrator, right-click the resource and select **Bring Online** from the menu (see Figure 5-35).

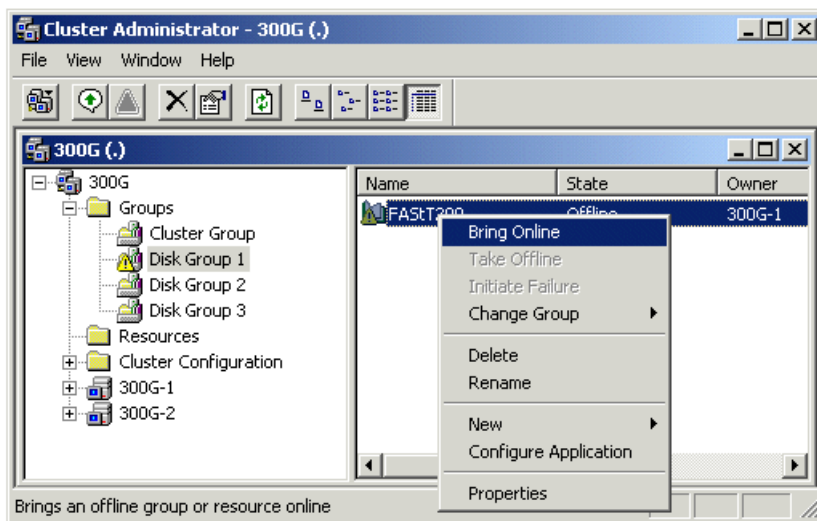


Figure 5-35 Bringing SANergy Volume resources online

If you have any problems, Cluster Administrator should identify the cause. The computer's event logs may also provide diagnostic information. The most common cause of problems are incorrect values in the **Volume Label** and **Serial Number** parameters for the resource. You can edit these by right-clicking on the resource and selecting **Properties -> Parameters**. As we have warned before, including the dash in the serial number will cause this step to fail.

Define File Share cluster resources

Now that you have SANergy Volume resources available, you will need to define file shares to allow SANergy hosts to access the volumes. When defining a share for a cluster, you will need to create File Share and NFS Share resources. By defining the share as a cluster resource, ownership of the share can be relocated by MSCS as needed. We explain the procedures for creating both CIFS and NFS File Share resources in Section 4.5.4, "Configure file shares" on page 207.

The procedure is exactly the same here with one important exception. You will be substituting SANergy Volume resources for Physical Disk resources. In other words, you will still create an IP Address resource and use it as a dependency for a Network Name resource, but when you create the File Share resource, it will be dependent on a SANergy Volume resource and the Network Name. This relationship is shown in Figure 5-36.

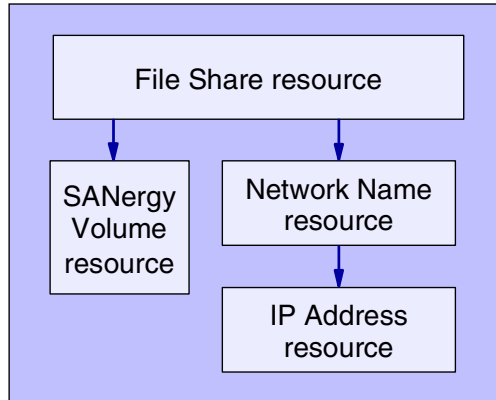


Figure 5-36 SANergy volume file share dependencies

We will not run through the entire process of creating shares again here, so please see Section 4.5.4, “Configure file shares” on page 207 for a full walkthrough of these procedures.

Map the file shares from a SANergy host

Once you have created the file shares on the cluster, you must mount the shared volumes being served from the SANergy MDC (the cluster). To do this, just mount the volume using the UNC name of the Network Name resource and File Share resource you defined. For example, we defined a network name of **FASTT** and a File Share name **FASTT200**, so from our Windows host (Scandium), we typed `net use t: \\FASTT\FASTT200` to mount the volume.

Tip: While it is also possible to use the name of the cluster (300G in our case) rather than the Network Name resource name to access the file share, this is not best practice. Due to an inconsistency in MSCS, you may lose access to the file share if you have mounted a volume through the cluster name but for some reason, the node that controls the file share loses control of the cluster. This will probably only be a concern if you are using the dual-node 300G for load balancing. However, you created the Network Name resource, so you might as well use it. The only disadvantage to this is that, while a search of the network will return the Network Name, the resource is not always visible from the GUI.

After you have mapped the file share, validate both functionality and security settings by creating, modifying, and deleting a file on that share.

Validate configuration

At this point, I/O to the mounted file share should now be fused and operating at SAN-level performance instead of LAN speeds. Double-check this by starting the SANergy Setup application and performing write and read tests to the mapped drive (see Figure 5-37). Under ideal circumstances you should see the same performance to the mapped drive that you find from the MDC node that owns the disk and does direct I/O to the disk.

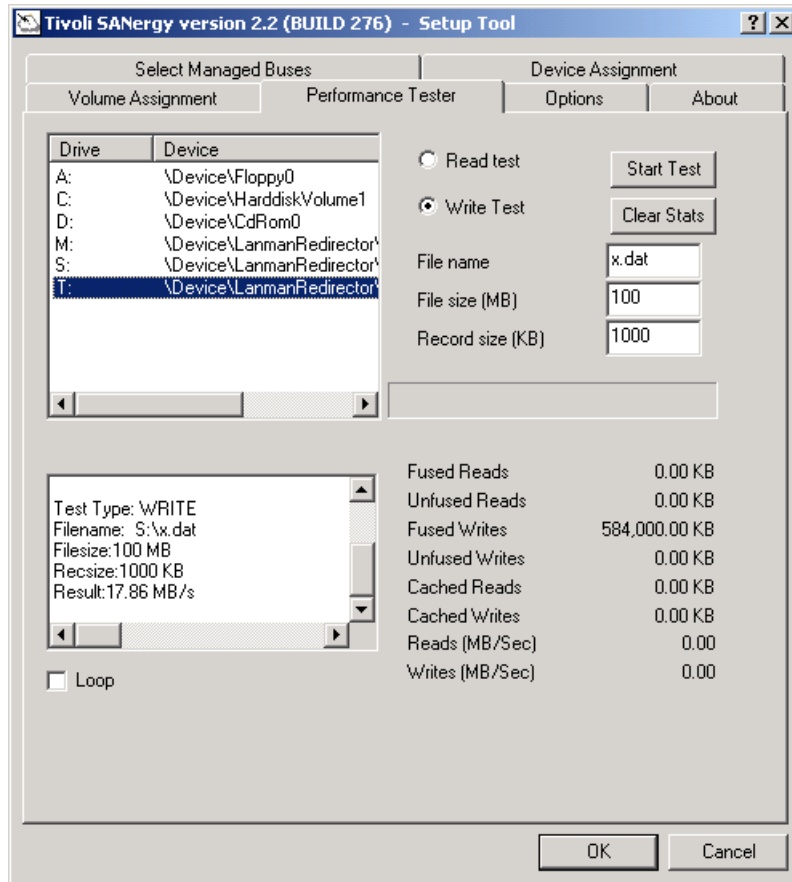


Figure 5-37 Validating the installation using SANergy Setup

You should also validate that the resources fail over successfully. To test this, from Cluster Administrator, right-click the group that contains your file share and select **Move Group** (see Figure 5-38). This will force a failover and cause the other node to take control of the SANergy volume and the file share.

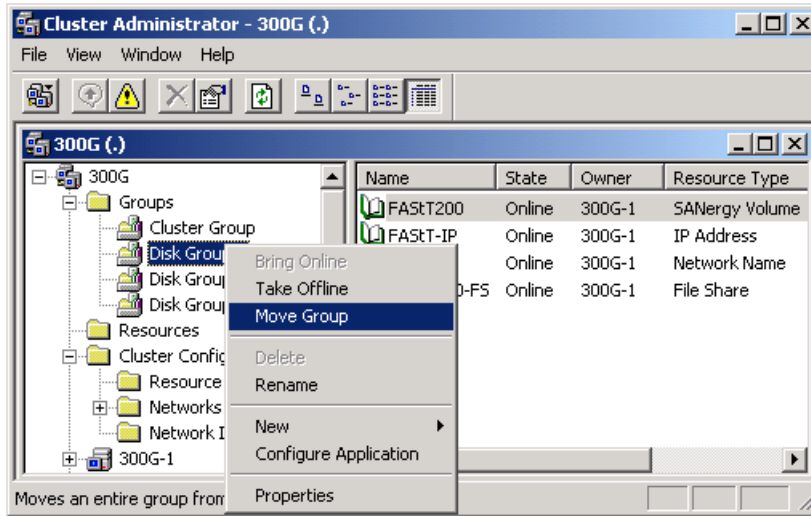


Figure 5-38 Move MSCS group to test failover

You will see the individual resources go offline on the current node (300G-1) and then come online on the new node (300G-2). Figure 5-39 and Figure 5-40 illustrate this.

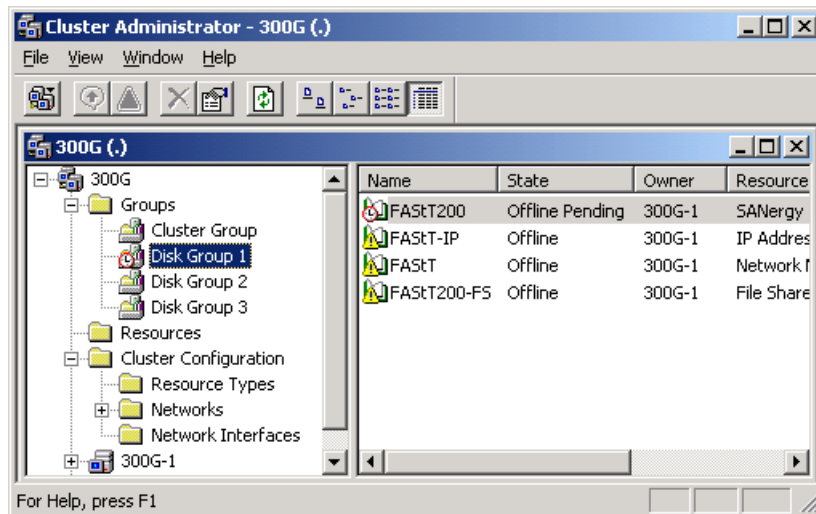


Figure 5-39 MSCS group going offline on current node

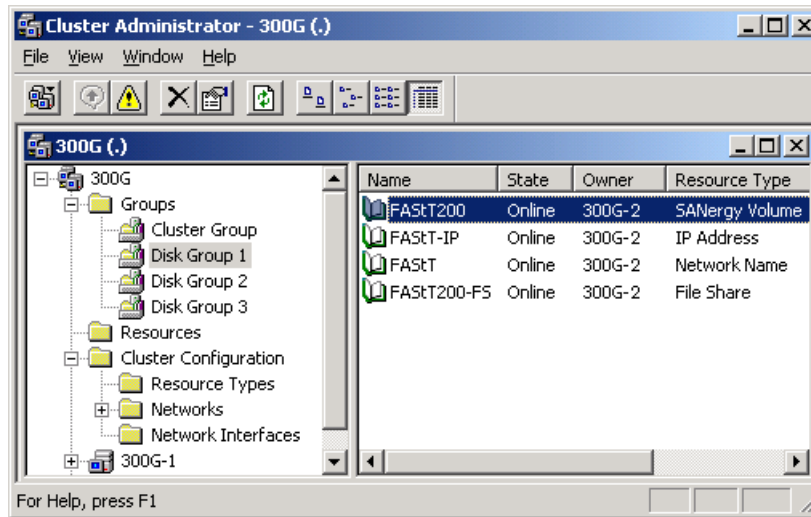


Figure 5-40 MSCS group online on the other node

We tested our configuration by running a continuous performance test on the SANergy host (Scandium) and then initiating a failover. The resources moved successfully, and we could continue the performance test immediately.



Backing up the IBM TotalStorage NAS 300G

In our computing world today, data is considered the most important competitive differentiating factor. Temporary inaccessibility or the complete loss of data has a huge financial impact, and can drive companies out of business. The inability to manage data can have a negative impact on a company's profitability and limit its ability to grow. Storing, protecting, and managing data growth has become one of the major challenges of today's businesses. For these reasons, it is essential that disaster recovery and backup/restore procedures be properly planned and implemented to meet the high demand for security and safety of data.

The 300G is designed to plug into your current data protection scheme. Unlike many NAS appliances, the 300G does not depend on special vendor-provided versions of software but works with out-of-the-box backup software. If you use Tivoli Storage Manager (TSM) as your enterprise backup solution, you will be pleased to know that the 300G is shipped with a TSM client already installed on it. Also, if you do not already have a data protection scheme in place, the 300G comes bundled with its own complete backup solution.

This chapter describes various configurations for backing up and restoring the 300G:

- ▶ Using the native backup solution
- ▶ Integrating the 300G into an existing TSM environment
- ▶ Combining TSM with SANergy to move backups off of the LAN
- ▶ Restoring the 300G from scratch

6.1 The 300G and its native backup solution

The 300G comes with a rich set of utilities for data management. One of the key advantages of using the 300G is the ability to capture point-in-time image copies without the need for a long downtime window by means of the Persistent Storage Manager (PSM) software. The following sections will describe the use of the PSM, and how it can be used in conjunction with NTBackup to help increase productivity in backup and recovery of your mission critical data.

6.1.1 300G cache and backup

Cache is often implemented in computer systems to improve performance. The 300G uses large cache to optimize performance. When enhanced RAID systems with battery backup are used, the RAID caches can be run in write-back mode to dramatically improve file write operations. The 300G can take advantage of these features to improve backup performance.

The 300G uses two types of backup: “point-in-time” image copies and “archival backup”.

Point in time backup

“Point-in-time” images provide a near-instant virtual copy of an entire storage volume. These point-in-time copies are referred to as “persistent images” and are managed by the Persistent Storage Manager (PSM) software.

These instant virtual copies have the following characteristics:

- ▶ Normal reads and writes to the disk continue as usual, as if the copy had not been made.
- ▶ Virtual copies are created very quickly and with little performance impact, as the entire volume is not truly copied at that time.
- ▶ Virtual copies appear exactly as the original volume when the virtual copy was made.
- ▶ Virtual copies typically take up only a fraction of the space of the original volume.

These virtual copies are created very quickly and are relatively small in size. As a result, functions that would otherwise have been too slow, or too costly, are now made possible. Use of these persistent images may allow individual users to restore their own files without any system administrator’s intervention. With the pre-loaded code, the NAS administrator can schedule the PSM to automatically perform an instant virtual copy at regular intervals.

The administrator can also grant end users access to their specific virtual copies. If a particular user accidentally deletes or corrupts a file, he or she can just drag-and-drop the virtual copy of that file to their storage without any administrator involvement. If you would like to know more about this topic, please refer to *IP Storage Networking: IBM NAS and iSCSI Solutions*, SG24-6240.

Archival backup

“Archival backup” is used to make full, incremental, or differential backup copies, which are typically stored to tape. A common problem with these backups is that files that were open at the time the backup ran often fail get backed up. The 300G’s PSM is not hindered by open files, so it can successfully make backup copies in a 24x7 operation.

6.1.2 Persistent Storage Manager (PSM)

PSM provides a “point-in-time” image of the file system. The PSM function in the 300G is similar to the following functions in other products:

- ▶ FlashCopy on the IBM ESS

On the IBM NAS products, all of the following terms refer to the same functionality:

- ▶ persistent image,
- ▶ True Image (Columbia Data Products)
- ▶ Point-in-time image
- ▶ Instant virtual copy

PSM not only solves the “open file” backup problem, it also provides very quick volume copying, eliminates long backup windows allowing continued system access during the backup, and provides easy end-user restorations of individual files.

Usually, after a backup is made, users will continue to update the files on the disk. In time, these backups will become outdated. However, it is very important that the backup data remain exactly as it was when the backup was made.

Unfortunately, making a backup copy while the data is still changing is rather difficult. While data is changing at any given point-in-time, multiple sectors are being written to disk. Write-back caches may not have completed writing to disk. And an application that is changing two or more files “at the same time” will not truly update both at the exact same instant. Therefore, for a good backup in which all data written is consistent in all changed files, these kinds of file writes must not occur while the backup is being made.

Historically, this problem has been solved by disabling all users while the backup occurs, however, this may take several hours. In today’s 24x7 environment, having such a large backup window is simply not acceptable. In the 300G, this problem is solved by making a very quick “instant virtual copy” (also known as a “persistent image”) of a volume.

Tip: After PSM images are created, you might have to wait for a few seconds or minutes in order for PSM to update its write-back queues and caches. In particular, the very first image will generally take much longer than subsequent images. Hence, if the system is heavily utilized, this update may take a while. After this, you should be able to access the images on a read basis only. One other thing to keep in mind is that by design, PSM will run at a lower priority than regular traffic.

6.1.3 Backing up PSM using the W2K Terminal Service

In keeping with its role as a plug-in appliance, the 300G is shipped without any I/O peripheral. However, Windows 2000 Terminal Services is configured and running on it so that you can remotely manage the machine just as if it were your own workstation. This web-based GUI (accessed via port 8099) is called the Windows 2000 for Network attached Storage (NAS) user interface. In this section, we will discuss using the interface to manage PSM backup and restore operations.

PSM is a great tool for getting persistent images of volumes. It is very easy and convenient to recover data using this tool. In this section we will be using Windows 2000 Terminal Service to access NTBackup and Restore.

Backup

1. From your remote management workstation, launch Internet Explorer and point it to the IP address of your 300G with the port number 8099. In our case, this was:

`http:\\192.168.200.100:8099`

2. Click **Disk and Volumes->Disk and Volumes** (Figure 6-1).

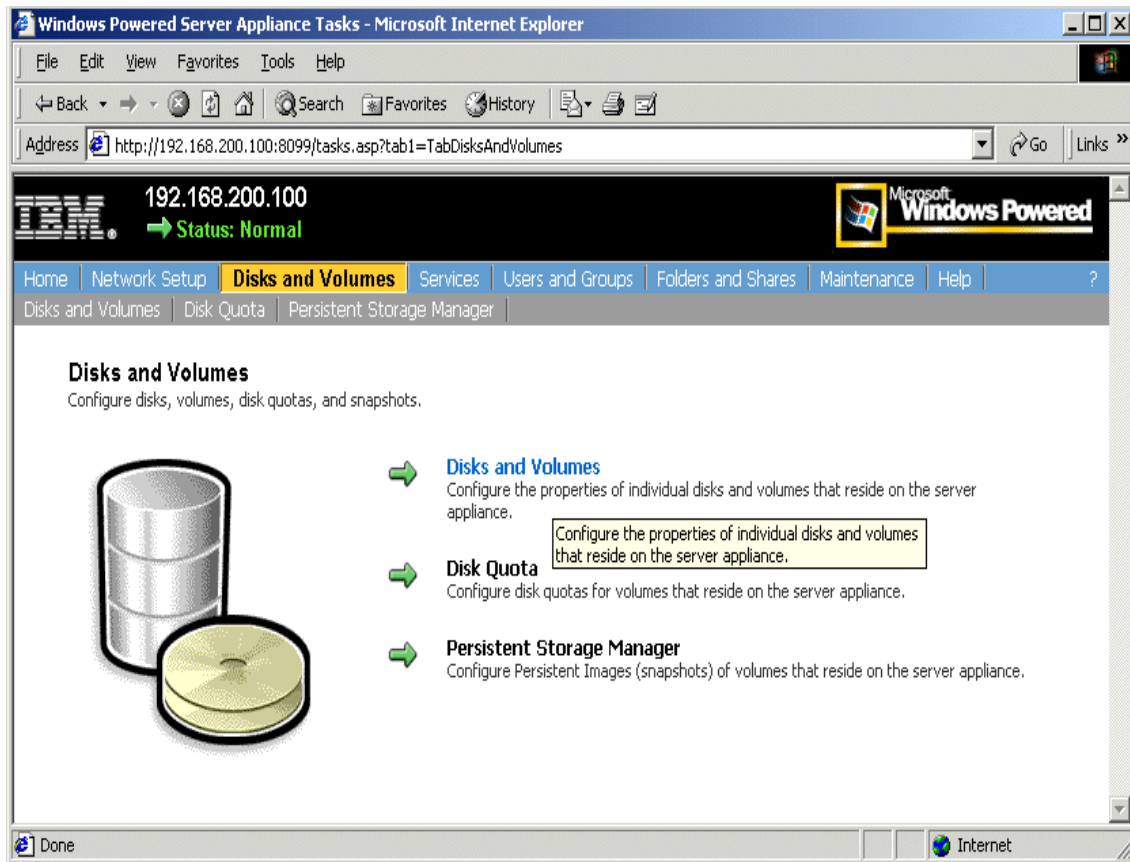


Figure 6-1 The Windows 2000 Terminal Service

3. Click **OK**, then provide the administrator password and click **OK** again.
4. From the Terminal Services client desktop, double-click the **IBM NAS Admin.msc** shortcut.

5. Click **Backup and Restore** as shown in Figure 6-2.
6. Click **Restore Using NTBackup**.

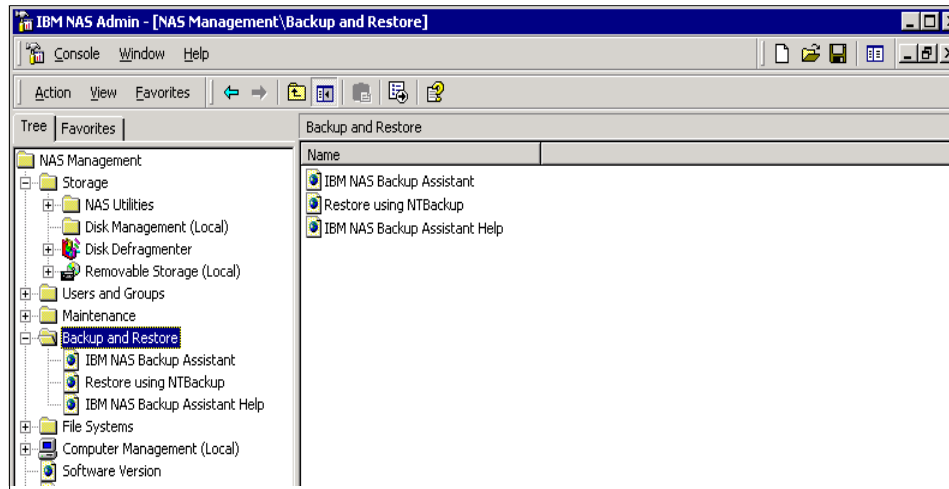


Figure 6-2 NTBackup window

7. Enter the password for the administrator.
8. Click the **Backup Wizard** button then **Next** to proceed to the next window.
9. Since we will be backing up files we have to choose the **Backup selected files, drives or network data** option and click **Next** to proceed.

10. Select the directory you want to back up (Figure 6-3).

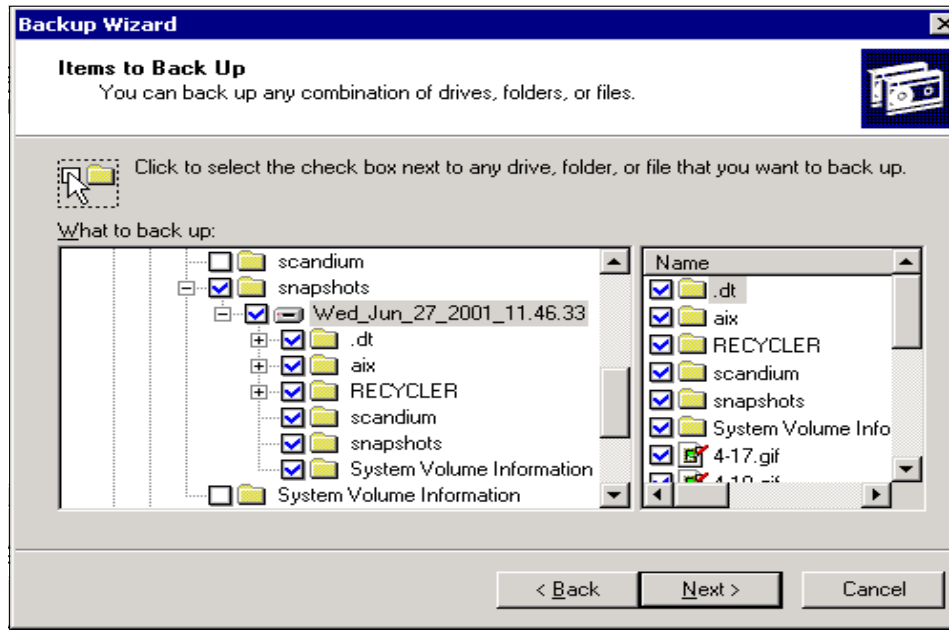


Figure 6-3 The Backup Wizard

Important: Click on your snapshot directory and look in the right pane. Make sure the boxes preceding the files and subdirectories you want to back up are checked. Otherwise, even though the top-level folder is checked in the left pane, none of the files or subdirectories under it will be backed up.

11. After choosing the files you want to back up, click **Next** to proceed to the next window.

12. Provide a file name for the backup file.

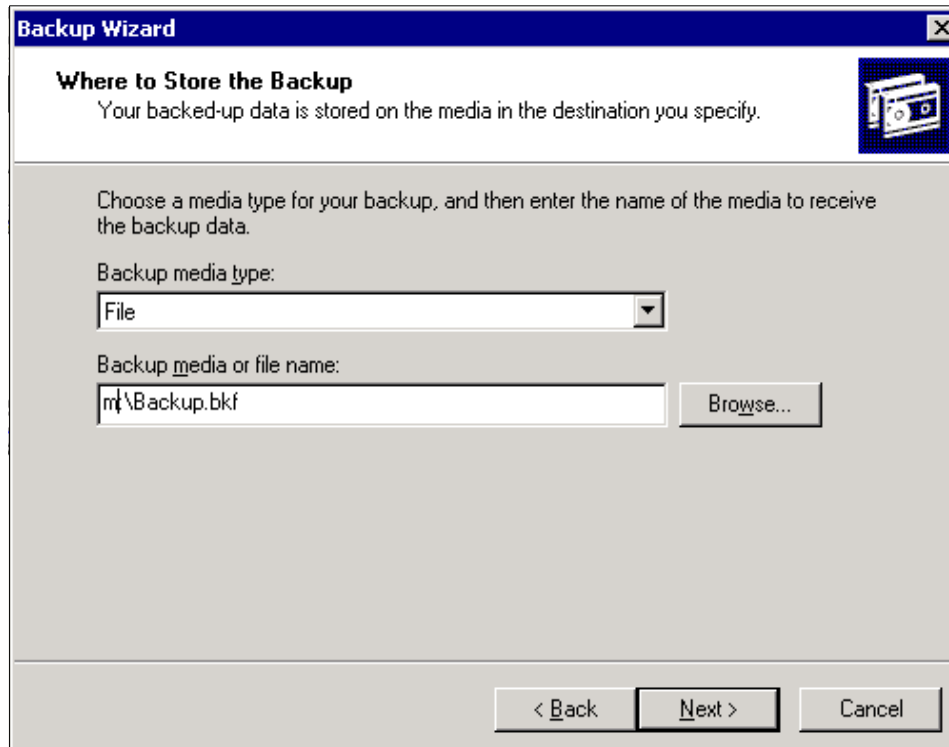


Figure 6-4 Window for the backup file

13. Click **Next** to proceed with the backup.
14. After the backup is done, click **Finish** to proceed, and then click **Close** to close the backup window.

Restore

1. Instead of the Backup Wizard, click the **Restore Wizard**.
2. Click **Next** to proceed to the next window.
3. Choose the files and directories you want to restore, as shown in Figure 6-5.

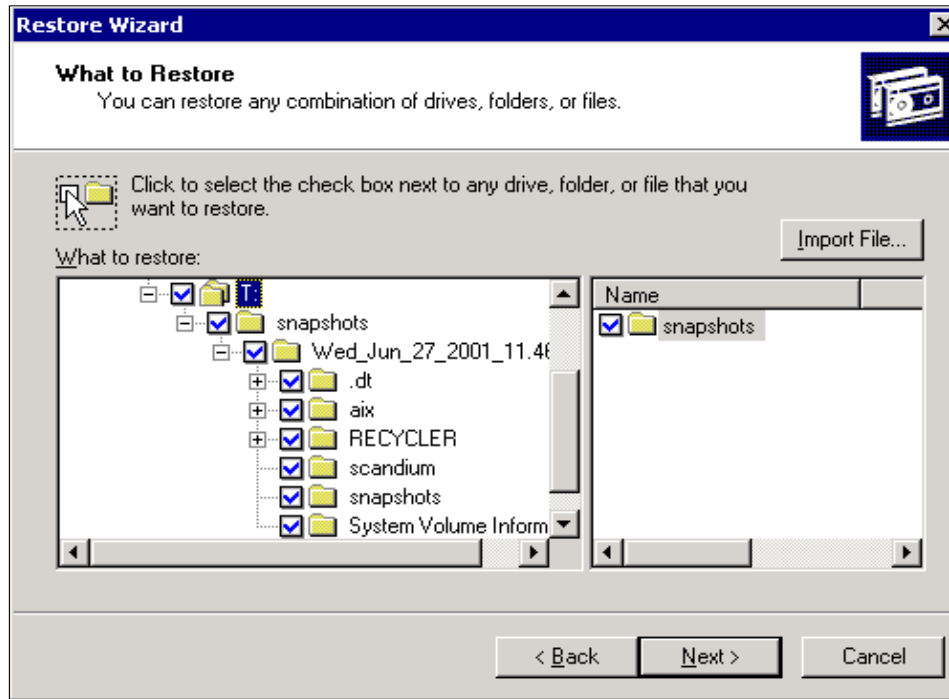


Figure 6-5 The Restore Wizard

4. After carefully choosing the files we need to restore, we can now click the **Next** button to proceed,
5. Click on the **Advanced** button if you want to restore the files to a location other than the original source or click **Finish** to start restoring data,
6. Provide the name of the backup file you want to restore from (Figure 6-6),

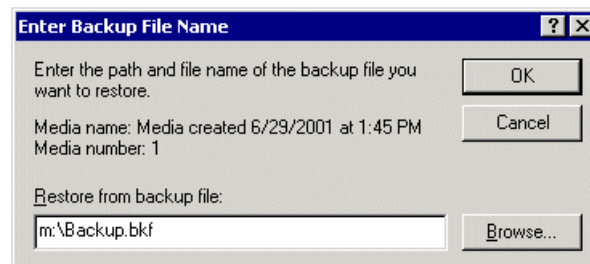


Figure 6-6 Backup file name dialog

7. After the restore, click **OK** to continue and then choose **Close** to close the restore window.

Important: Please be aware if you back up files directly from a PSM persistent image: the entire path name of each backed up file is preserved, with the result being that when you restore such a file, it will attempt to restore to the persistent image and not to the original volume. "Restore Using NTBackup" should only be used in situations where standard backup (that is, not open file) is deemed sufficient, and you only want to back up a few selected files (as opposed to an entire volume). For all other backups, using NT Backup, the NAS Backup Assistant should be used.

6.1.4 Archive, backup, and restoration of the 300G

System administrators will want to ensure that data stored in the 300G has adequate protection against data loss from accidental erasure or replacement, disk crashes, and even disaster scenarios. This section discusses the included options for doing so and the purpose of each option.

Note: The 300G does not support making an archival copy of the PSM cache itself. Therefore, when using these recovery approaches, all PSM persistent images and PSM caches should be deleted.

Archiving the 300G's operating system

The IBM xSeries 150 and Network Attached Storage 300G products are shipped with a Recovery CDROM that allows the NAS administrator to restore the system to its factory default configuration. Therefore, no matter what happens to the operating system or maintenance partition, the NAS administrator can restore the operating system software from this Recovery CD-ROM. However, if the Administrator has applied any fixes to the NAS product, these must be reapplied after the Recovery CD-ROM is used.

Archival backup of the 300G's maintenance partition

The 300G is pre-configured with a 3GB operating system partition and a 6GB maintenance partition. Using the pre-loaded NTBackup software, the administrator can make a backup of the operating system to the maintenance partition. The 300G has a wizard assistant to make this simple. Since the 300G's operating system may be in use when performing the backup, it is important to use a persistent image function to resolve the "open file" problem when making a backup. The NAS backup assistant will invoke the persistent image function before the NTBackup is started to ensure that the backup is complete and valid.

The 300G requires little configuration. Furthermore, the permissions and access control lists are stored with the file systems in the user partition, not with the operating system partition. Therefore, if the operating system must be restored from a back-level copy, there are only a few parameters which need to be updated.

Archival backup of the 300G's OS to tape using NTBackup

The operating system can be backed-up to tape, using the included NTBackup program and the Backup Assistant. Again, to resolve the "open file" problem, the backup should be of a persistent image which can easily be accomplished by using the included NAS Backup Assistant wizard.

Archival backup of the 300G's user (client) data

Systems administrators need to make archival copies of their critical data. Typically, these copies are made to tape, and then these tape cartridges may be taken off-site to protect against site disaster incidents.

In virtually all cases, PSM will be used immediately before the backup is started. While persistent images are retained across reboots, these persistent images are *not* a replacement for tape backup, as they do not provide the ability to have off-site copies. Therefore, NAS administrators should not use persistent images as their disaster recovery approach.

How archival backup accesses a PSM persistent image

The PSM user interface is accessed via the Windows 2000 for NAS user interface, in addition to Windows 2000 Terminal Services. This is where the PSM images are created. They are either executed immediately or scheduled for single or periodic execution.

6.1.5 NTBackup

The 300G is pre-loaded with Windows NTBackup and the NAS Backup Assistant. This approach can be used to back up operating system data or user data. Backups can be made to disk or tape. The pre-loaded PSM function is the recommended method of resolving the "open file" problem.

There are two ways to back up the files in the 300G when you use the NTBackup method. One option is to drive the process manually using either the Windows 2000 for NAS user interface or Windows Terminal Services (by clicking **Maintenance-> System Backup and Restore->Backup** from your remote management machine). For this approach, you should first create a Persistent Image before the NTBackup is started.

This is the best method to use if you only want to back up a selected folder or set of files from one of the persistent images or the system partition. Please be aware that backing up from a persistent image preserves the path. (See the important notice at the end of “Backing up PSM using the W2K Terminal Service” on page 280)

The other option is to use the NAS Backup Assistant tool. The NAS Backup Assistant will automatically create a Persistent Image and start the NTBackup program. This is the best method for backing up data at a volume or file system basis. To use the NAS Backup Assistant, follow these steps:

1. Use **Windows Terminal Services** from any NAS client to access the 300G
2. Double-click the **IBM NAS Admin.msc** shortcut on the desktop
3. This launches the **IBM NAS Admin** console
4. Select **Backup and Restore->IBM NAS Backup Assistant** from the left pane
5. Select one of the following options from the right pane:
 - Backup Operations: Select drive, schedules, backup types, backup methods, destination type, file path or tape name
 - Schedule Jobs: List jobs scheduled for backups. You can also delete jobs that have been scheduled but not yet executed.
 - Backup Logs: Shows logs of all backups. Can view or delete logs here.
 - Display Logs: Displays the logs.

See Figure 6-7 for a detailed illustration of the NAS backup assistant.

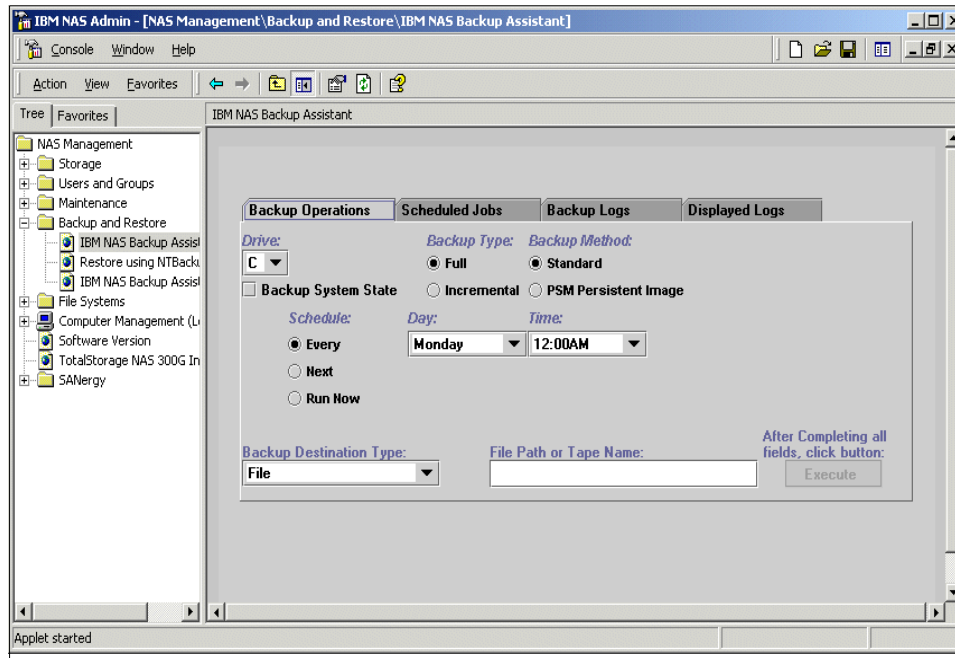


Figure 6-7 The IBM NAS backup assistant

To restore, just follow the preceding steps, but select **Restore Using NTBackup** in step 4 instead of **IBM NAS Backup Assistant**. You can see the details of this process described in the part “Backup” on page 280.

While the NTBackup program is a powerful and useful tool, it does have some limitations, specifically:

- ▶ NTBackup is limited to writing to locally attached devices. This reduces backup and restore flexibility by requiring hardware to be directly attached to the machine being backed up.
- ▶ NTBackup has no policy management for maintaining and expiring data. These functions have to be managed manually outside of NTBackup by the administrator.

6.2 Integrating the 300G with TSM

While the 300G offers its own, independent backup solution, whenever possible, best practice is to have a consistent data protection scheme applied across your entire enterprise. Tivoli Storage Manager (TSM) provides just such a solution.

TSM, together with its complementary products, is designed to provide a comprehensive data protection system including:

- ▶ **Operational Backup and Restore of Data:** The backup process creates a copy of the data to protect against the operational loss or destruction of file or application data. You can define how often to back up (frequency) and how many copies (versions) to maintain. The restore process places the backup copy of the data onto any system or workstation you designate.
- ▶ **Disaster Recovery:** This refers to all activities having to do with organizing, managing, and automating the recovery process from a major loss of IT infrastructure and data across the enterprise. This includes processes to move data off site into a secure vault location, to rebuild the IT infrastructure, and to reload data successfully within an acceptable time frame.

The 300G has been pre-installed with TSM client code v3.7. To begin using it, all you have to do is configure it to work with a TSM server currently running or implement a TSM environment to work with the 300G. The TSM server can be any other server in the network, and based on the TSM server's configuration, the final destination of the 300G's backup may either be located in the TSM server's disk storage or an attached tape subsystem. The latter is the preferred target location.

6.2.1 The 300G and LAN-based backup

The 300G is shipped with TSM client code installed. This client code when activated should have a TSM server to connect to, so it is able to do a backup whether it be via the LAN or LAN-free. In our case the NAS box is pre-installed only with the TSM client v3.7, and as such we still need an additional server to back it up LAN-free. For additional details regarding this, please refer to *Using Tivoli Storage Manager in a SAN Environment*, SG24-6132.

Here are some of the steps that will help us in configuring the 300G to work with an existing ADSM/TSM server.

Setting it up

An ADSM or TSM server is currently configured to accept data from a number of clients via LAN. The communication method *COMMMethod* should be set to *tcpip*. For more detailed discussion on the proper configuration of the TSM server, please refer to *Getting Started with Tivoli Storage Manager: Implementation Guide*, SG24-5416-01.

Client configuration

The TSM client uses an option file to store its configuration. Once the setup is completed, it will create an option file on the 300G in the following directory and file name:

```
C:\Program Files\Tivoli\TSM\baclient\dsm.opt
```

Below are the steps to configure the TSM Client.

1. You will need to access the 300G from the IAACU.
2. Open the **Windows Terminal Services**.
3. Select **Start -> Programs -> Tivoli Storage Manager -> Backup Client GUI**.
4. This will produce the Tivoli Storage Manager window.
5. Select **Utilities -> Setup Wizard**.
6. You will get the TSM Client Configuration Wizard, and check the following:
 - Help me configure the TSM Backup Archive Client
7. Select **Next**, and check the following:
 - **Create a new option file:** Select this option for a new setup. Take this option if you are setting up the first time.
 - **Import an existing option file for use:** Select this option only if the dsm.opt file was previously created by the system administrator on some other machine.
 - **Update my options file:** Select this option if you want to update a previously configured dsm.opt on the same machine.
8. Select **Next**, and you will be asked to enter the TSM Node Name to use. This should be the name of the TSM Client, that is., 300G.
9. Select **Next**, and it will display the TSM Client/Server Communications screen. Select **TCP/IP**.
10. Select **Next**, and it will ask for the TCP/IP Parameters.
11. Enter the **Server Address**. This is the TSM Server's IP address, for example: 192.168.200.101
12. Enter the **Port Address** as: **1500**. This is the default value for TSM.

13. Select **Next**, and check the following:

- **Domain List:** Click **Edit** to select the directory to be backed up.
- **Include/exclude List:** Click **Edit** to either include or exclude some files from the list.

14. Select **Next** and then **Finish** to complete the TSM client configuration.

Example 6-1 shows the sample dsm.opt file used to configure the 300G to work with an existing TSM server.

Example 6-1 Sample dsm.opt file on the 300G

```
NODENAME ibm5196
PASSWORDACCESS GENERATE
DOMAIN.IMAGE T:
Exclude "T:\tsmdata\data2.dsm"
Exclude "T:\tsmdata\data4.dsm"
Include "T:\tsmdata\test1" STANDARD
Include "T:\tsmdata\test10" STANDARD
Include "T:\tsmdata\test11" STANDARD
Include "T:\tsmdata\test2" STANDARD
Include "T:\tsmdata\test3" STANDARD
Include "T:\tsmdata\test4" STANDARD
Include "T:\tsmdata\test6" STANDARD
Include "T:\tsmdata\test5" STANDARD
Include "T:\tsmdata\test7" STANDARD
Include "T:\tsmdata\test8" STANDARD
Include "T:\tsmdata\test9" STANDARD
DOMAIN ALL-LOCAL
TCPSERVERADDRESS 192.168.200.101
```

Note: This setup is intended to run in a Local Area Network (LAN) backup/archive environment.

For the backup to work, the TSM Server must have its client's nodename registered in its configuration files. In this case, it will be the 300G's nodename.

To back up the files from the TSM Client, follow these steps:

1. Use **Windows Terminal Services** from any NAS client to access the 300G.
2. Select **Start -> Programs -> Tivoli Storage Manager -> Backup Client GUI**.
3. This will lead you to the **Tivoli Storage Manager GUI**.
4. Select **Backup**.

5. On the left pane, you select the directory to back up or use the right pane to select individual files for the backups.
6. To restore, just follow the above steps, but select Restore in step 4 instead of Backup.

6.3 TSM with SANergy

Tivoli SANergy introduced LAN File sharing technology to Storage Area Networks (SANs).

Tivoli SANergy file sharing

SANs are fast becoming the preferred method for companies to manage the vast amounts of data created by their e-business systems — by creating networks to connect "islands of information" that can be shared quickly across the enterprise. SANs can link data centers and manage data across heterogeneous computing platforms for real-time backup and recovery, data migration, and resource sharing (Figure 6-8).

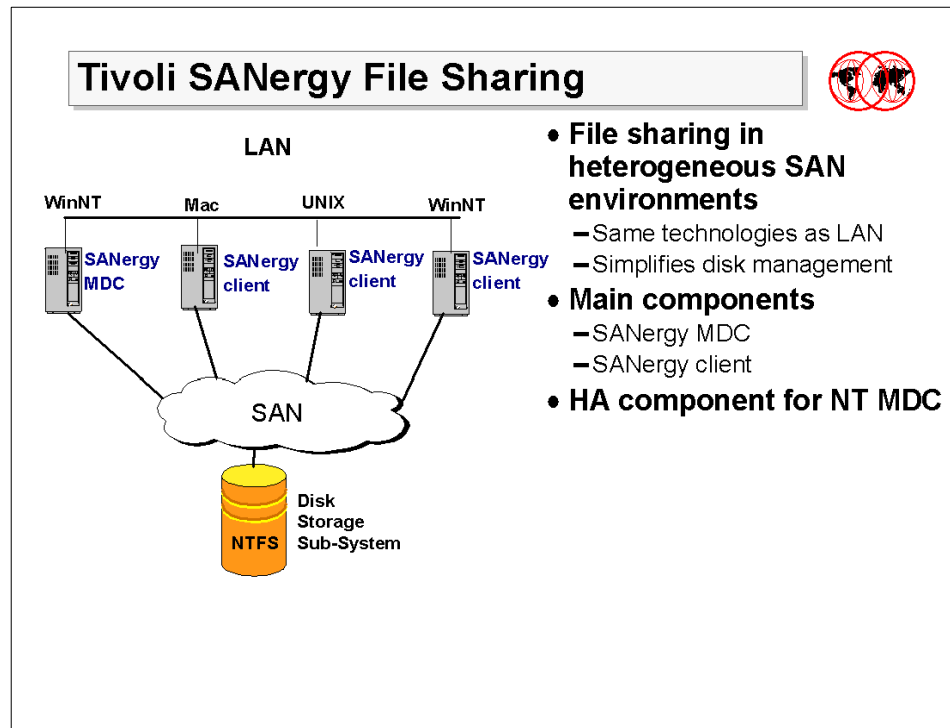


Figure 6-8 File sharing with Tivoli SANergy

SANergy is the only SAN software that allows for the sharing of application files and data between a variety of heterogeneous servers and workstations connected to a SAN. In addition, SANergy software uses only industry-standard file systems, enabling multiple computers simultaneous access to shared files through the SAN. This allows users to leverage existing technical resources instead of learning new tools or migrating data to a new file system infrastructure. This software allows SAN-connected computers to have the high-bandwidth disk connection of a SAN while keeping the security, maturity and inherent file sharing abilities of a LAN.

SANergy employs technology to combine the simplicity of LAN-based file sharing with the very high data transfer speeds afforded by modern Fibre Channel, SCSI, and SSA storage networks. This enables the use of high-speed, heterogeneous data sharing without the performance-limiting bottlenecks of file servers and traditional networking protocols.

SANergy is unique in that it extends standard file systems and network services provided by the operating systems that it supports (Windows NT, MacOS, and UNIX). As an O/S extension built on standard systems interfaces, SANergy fully supports the user interface, management, access control, and security features native to the host platforms, providing all the file system management, access control, and security expected in a network. With SANergy, virtually any network-aware application can access any file at any time, and multiple systems can transparently share common data.

In addition to the SAN, SANergy also uses a standard LAN for all the meta data associated with file transfers. Because SANergy is NTFS-based, should the SAN fail, access to data via the LAN is still possible. Since each system has direct access to the Tivoli SAN-based storage, SANergy can eliminate the file server as a single point of failure for mission-critical enterprise applications. SANergy can also easily manage all data backup traffic over the storage network, while the users enjoy unimpeded LAN access to the existing file servers.

SANergy allows file sharing between different heterogeneous platforms, including Windows NT, SUN Solaris, IBM AIX, and Apple Macintosh. Without SANergy, a piece of data held on disk to be used for Windows NT is often copied to another platform to perform some kind analysis, such that systems could not take full advantage of the capabilities of the SAN.

SANergy helps computer users in networked environments speed transaction time by taking full advantage of the capabilities of the SAN.

Now, why is SANergy important in this chapter? TSM without SANergy will let us use the SAN for backup but the data will go straight to the tape library. With SANergy configured, we can now have a transfer of data from the source volume to a disk storage pool and finally migrate the data to a tape library. This will be a much better solution if and when a backup window is critical.

6.3.1 TSM backup using SANergy

SANergy, when used in conjunction with Tivoli Storage Manager, supports both LAN-free and server-less backup across a SAN. LAN-free is the movement of data from the source to the backup media without using the LAN. This transfer of data is much faster since the path it takes is via the Fibre Channel connection and goes straight to the storage media. Server-less backup is the transfer of data from the source to the storage media via the same route as the LAN-free backup but without any CPU usage from the owner of the data.

One example might be managing CPU utilization so as not to hamper end-user connections. LAN-free and server-less backups allow the data traffic to be off-loaded to the SAN, instead of moving the data through file servers over the LAN. This reduces LAN and server overhead, minimizing the impact of backups on the efficiency of the network. Similarly, when using SANergy, data migration from legacy storage on servers to new storage on the SAN has no impact on the LAN performance.

In the following sections, we discuss two possible configuration scenarios.

TSM with SANergy

TSM and SANergy can be used to achieve a backup where the I/O is LAN-free, and also, no application server resources are used. For applications with NFS or CIFS mounted volumes, SANergy can be used to remove the backup traffic from the LAN.

Application servers run SANergy, and each server is configured as an MDC. These systems are sharing data between them, so we want the SANergy configuration that has the least overhead for them to access their data. These servers read their data as locally attached drives (no-NFS mounts are required).

The system running the TSM server and client runs SANergy code in client mode in order to have shared access to the other system's files. This system NFS mounts the volumes and then uses the Backup-Archive client to back up the data over the SAN using the SANergy technology. During the backup, the TSM client and SANergy client do LAN-free reads of application files from the file server with only the meta data flowing over the LAN.

The main issues revolve around database backups. A customer will NOT be able to back up a database dynamically. For example, Oracle will not allow this backup to take place. The only way this scenario would work with Oracle is for the TSM connect agent to reside on the MDC that owns the Oracle database. Oracle Parallel Server and Oracle Fail Safe (Windows NT only) will allow the TSM Backup-Archive client to back up an off-line copy of the database while the on-line copy can continue to be updated. This information is true for any database package available.

Another potential issue involves multiple backups of open files. If a client currently has a file open and it is backed up, you could potentially get three copies of the same file backed up from each client accessing the file. This occurs automatically with Windows, but the issue is magnified with SANergy.

There is no TSM client on the application server. The SANergy client maps the volumes from the MDC, and TSM sees them as locally attached drives/file systems.

Another point to note is that the SANergy MDC does utilize some CPU resources, and therefore will have some impact on the application server.

6.4 Getting backups off the LAN: TSM with SANergy

In this section, we configure TSM to work with SANergy and do a (mostly) LAN-free backup. We will do this in two types of setups:

1. A setup where the TSM server is residing on the MetaData Controller (MDC), running a TSM client, and the 300G (the client node) is running the SANergy agent.
2. Another setup where the TSM server acting as a SANergy client and the 300G being the MDC and running a TSM agent and client.

6.4.1 SAN zoning

Since we are in a SAN environment, connected via Fibre Channel, we will have to configure the proper zoning before we can go forward in installing everything. Zoning is done to introduce the different components of the SAN environment to each other. For a detailed zoning configuration, please see 3.2.2, "Zoning the IBM 2109" on page 78.

6.4.2 Configuring SANergy

SANergy should be configured in both boxes that will play major roles in the backup and restore of data in a SAN environment.

The configuration of the pre-installed SANergy agent on the 300G is covered in detail in “A brief overview of Tivoli SANergy” on page 236. Since the 300G is pre-installed with the SANergy agent, we just have to customize it for our setup.

The installation and configuration of SANergy on a new box is covered in detail in “Configuring your other machines to use SANergy” on page 248. This should be done on the TSM server if you decide to have a different TSM server other than the 300G.

We highly recommend that a separate TSM server be used for this purpose.

6.4.3 Installing the TSM Server version 4.2

We recommend that the TSM server should reside in another box, whether it be NT, Windows 2000 Advanced Server or UNIX. In our laboratory, we installed the latest version of TSM which is version 4.2. This version, with all the added enhancements from the previous version works well with SANergy version 2.2. For more details on the enhancements, please refer to *Tivoli Storage Manager Version 4.2: Technical Guide*, SG24-6277. This is also a good book to refer to for the installation and configuration of TSM version 4.2.

1. Insert the TSM server CD.
2. Choose the appropriate language.
3. Click on **Install Products**. The TSM server wizard is shown in Figure 6-9.



Figure 6-9 Tivoli Storage Manager server welcome menu

4. Click **TSM Server** to install the server code.
5. Choose the desired language when the language box re-appears.
6. For the succeeding dialog boxes we simply chose the default values and clicked **Next**. For more detailed instructions in installing the server code please refer to *Tivoli Storage Manager Version 4.2: Technical Guide*, SG24-6277.
7. Click **Install** to continue the installation.
8. Click **Finish** to end the installation of the server code.

This will complete the copying and installation process.

Installing the TSM Server Licenses

1. Click the **TSM Server Licenses** in the **Install Products** wizard as shown in Figure 6-10.



Figure 6-10 The install products wizard window

2. Choose the language you want.
3. We just accepted the default values in the next three dialog boxes and clicked on **Next** to continue.
4. Click **Install** to proceed with the installation.
5. Click **Finish** to end the process.

Installing the TSM Device Drivers

1. Click the **TSM Device Driver** option in the **Install Products** wizard window shown in Figure 6-10 above.
2. Choose the language you want.
3. We accepted all the defaults in the succeeding three dialog boxes and clicked **Next**.
4. Click **Install** to continue installing the device drivers.
5. Click **Finish** to end the installation.

6.4.4 Configuring the TSM server

1. After the startup, you will have to double click on the **TSM Management Console** on your desktop (see Figure 6-11).

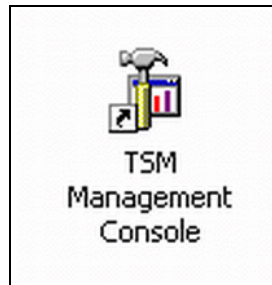


Figure 6-11 The Tivoli Storage Manager utility as shown on the desktop

2. Click **Initial Configuration**.
3. Click **Wizards**. The TSM wizards window is shown in Figure 6-12.
4. Click **Start**.
5. On the **Initial Configuration Environment Wizard** click **Next** to accept the default value then click **Next** to continue.
6. Click **Network**, accept the default value, and click **Next**.
7. Click **Finish** to complete the process.
8. At the **Performance Configuration Wizard**, click **Next** to accept the default.
9. Provide the number of clients and the size of files, then click **Next**.
10. Accept the default value on the next dialog box and click **Next**.
11. Click **Finish** to complete the installation.
12. At the **Server Initialization Wizard**, click **Next** to accept the default, then click **Next** to continue.
13. Accept the default in the next two dialog boxes and click **Next**.
14. Supply the System Administrator password as shown in Figure 6-13, then click **Next** to continue.
15. Choose **Finish**, then choose **OK** to complete the installation.
16. At the **License Wizard**, click **Next**.

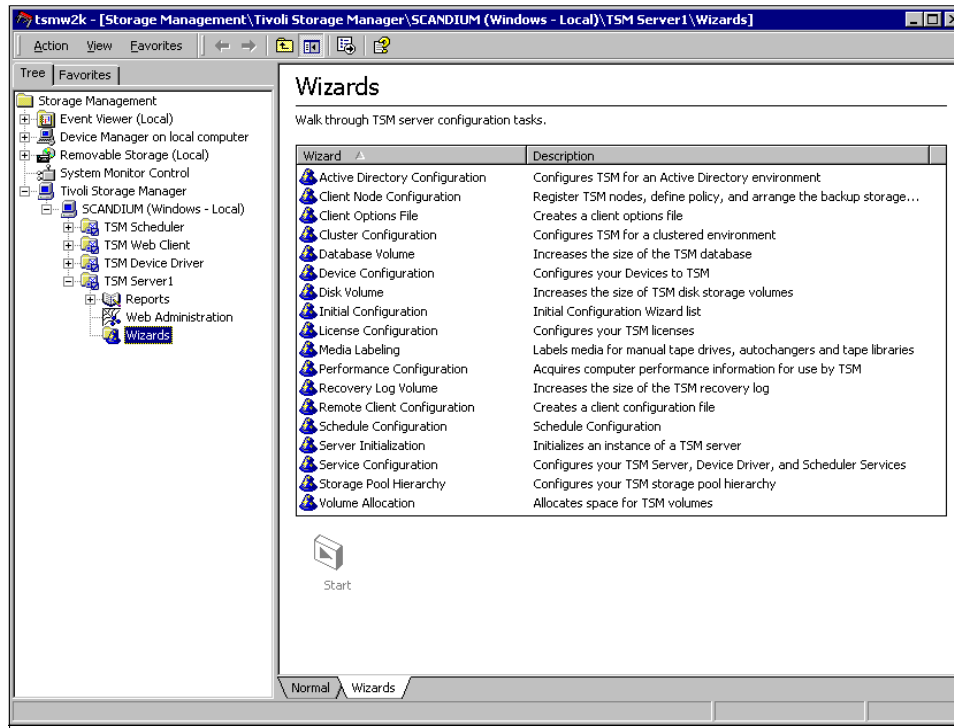


Figure 6-12 The TSM wizard window

17. Select the number of licenses you want to register for each module and click **Apply** before choosing another item.
18. Click **Next** to continue.
19. Click **Finish** to complete the installation and click **OK** to end the process.

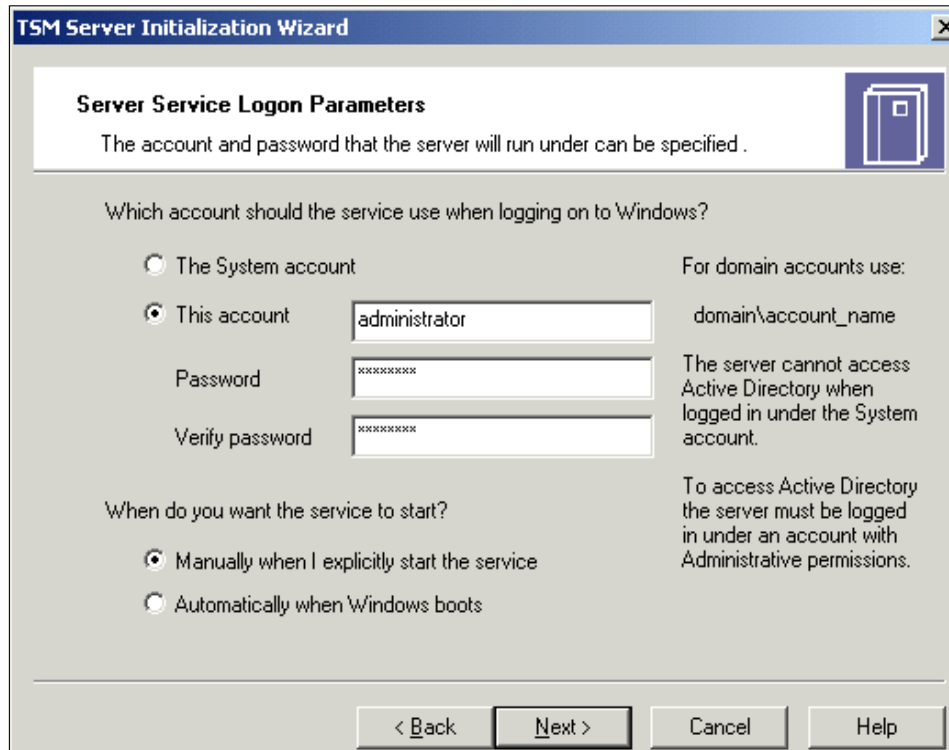


Figure 6-13 TSM server initialization wizard

20. At the **Device Configuration Wizard**, click **Next**.

21. Design the hierarchy of the devices you want to configure.

See our laboratory configuration in Figure 6-14. Properly configuring the devices will make the data flow more smoothly.



Figure 6-14 The Device Configuration Wizard

22. Click **Finish** to complete installation, then **OK** to end.
23. At the **Client Node Configuration Wizard**, click **Next**.
24. Click **Add Node** if you want to add a node, or click **Next** to proceed.
25. Click **Finish** to complete the process and **OK** to proceed to the next item.
26. At the **Client Scheduling Wizard**, click **Next** to accept the defaults for the next two dialog boxes.
27. Click on **Finish**, then **OK** to complete the process.

We accepted the default value, since we recommend that the schedules be set up after configuring everything else.

28. Define server definitions for the storage agent (Example 6-2).

Example 6-2 Defining server definition for the storage agent

```

tsm: SCANDIUM>define server ibm5196 serverpassword=ibm5196
hladdress=192.168.200.101 lladdress=1500 comm=tcPIP
ANR1660I Server ibm5196 defined successfully

```

29. Check server definition by typing **query server** on the TSM prompt.

30. Define drive mapping (Example 6-3 on page 304).

Example 6-3 Defining drive mapping on the TSM server

```
tsm: SCANDIUM>def drive mapping ibm5196 drive0 devi=mt0.2.0.3
```

31. Check drive mapping by issuing **query drivemapping** at the TSM prompt.

Note: We advise you to configure scheduling only after you have configured everything else.

6.4.5 Installing and configuring a TSM Agent on the 300G

1. At the main installation screen as shown Figure 6-9, click on **Install Products->TSM Storage Agent** (Figure 6-15).

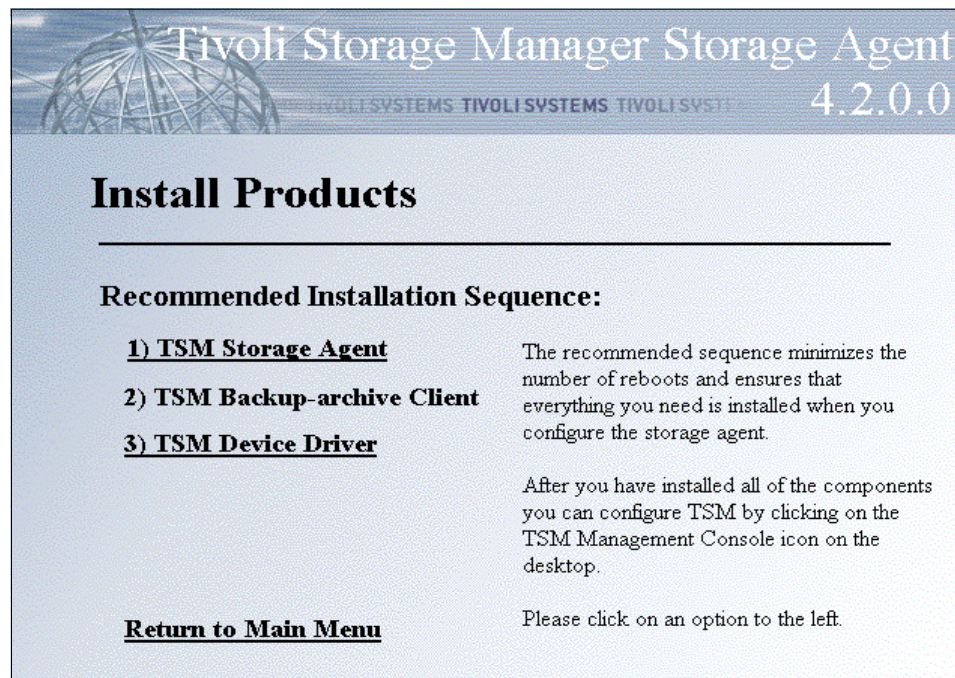


Figure 6-15 Tivoli Storage Agent Install Window

2. Choose the language for the install and click **Next**.
3. Accept the parameters for the default directory and click **Next**.
4. Choose to install the complete software, then click **Next**.
5. Choose **Install** to complete the installation.
6. Click **Finish** and then proceed to installing the device drivers.
7. At the installation screen, select **Install Products** and choose the **TSM Device Driver** (Figure 6-15).
8. Choose the language for the install.
9. Accept the default parameters for the next two dialog boxes and click **Next**.
10. Click **Finish** to complete the installation.

When the install completes, you will need to reboot your system.

6.4.6 Configuring a TSM Agent

1. Once the machine reboots, use the `tmsmcsi` command to see if the device driver is running. If it is not running, type `tmsmcsi /enable` at the “c:\Program Files\tivoli\tsm\storageagent>” path. This will enable Optical support for Windows 2000.
2. To display the device information, enter the `tsmd1st` command at the command prompt. This command will give us the information we need to complete the drive mapping.
3. Run the `dsmsta` command to put the necessary settings in the **devconfig.txt** and **dsmsta.opt** files.
4. Initialize the Storage Agent as a Windows 2000 service. You can launch the Storage Agent wizard from the TSM Console.

Click **Wizards-> Storage Agent Initialization->start** (at the bottom of the window). The **TSM Storage Agent Initialization** wizard is shown in Figure 6-16.



Figure 6-16 Storage Agent initialization wizard

5. Type the Storage Agent information in the next window and click **Next**.
6. Input the TSM server information when the next window appear and click **Next**.
7. Provide the system administrator password in the following window.
8. Click **Next** to proceed, then choose **Finish** to complete the installation.

For more a detailed description, please refer to Chapter 4 of *Tivoli Storage Manager 4.2: Technical Guide*, SG24-6277.

6.4.7 Installing a TSM client

For the installation of the client, please refer to “The 300G and LAN-based backup” on page 290. The only change you need to make from the description there is to the client options file. For this setup, you should use one like that shown in Example 6-4.

Example 6-4 TSM client options file

```
commethod      TCPIP
tcpport        1500
tcpserveraddress 192.168.200.101
ipxsocket       0005
ipxserveraddress 0000000000409512588A
netbiosname     client1
netbiosservername ntserver1
namedpipename  \\.\pipe\adsmpipe
DOMAIN.IMAGE T:
Exclude "T:\tsmdata\data2.dsm"
Exclude "T:\tsmdata\data4.dsm"
Include "T:\tsmdata\test1" STANDARD
Include "T:\tsmdata\test10" STANDARD
Include "T:\tsmdata\test11" STANDARD
Include "T:\tsmdata\test2" STANDARD
Include "T:\tsmdata\test3" STANDARD
Include "T:\tsmdata\test4" STANDARD
Include "T:\tsmdata\test6" STANDARD
Include "T:\tsmdata\test5" STANDARD
Include "T:\tsmdata\test7" STANDARD
Include "T:\tsmdata\test8" STANDARD
Include "T:\tsmdata\test9" STANDARD
DOMAIN "G:"
DOMAIN "T:"
ENABLELANFREE YES
NODENAME ibm5196
PASSWORDACCESS GENERATE
```

Note: You must check the boxes for the directory or individual files you want to back up during the selection process. Otherwise, nothing will be backed up or restored.

6.4.8 Backup/Restore for the 300G with TSM and SANergy

Now that the hardware and software are all in place, we can now proceed with the backup scenarios we want to test.

To back up the files from the TSM Client, follow these steps:

1. Use **Windows Terminal Services** from any NAS client or from any Windows 2000 desktop to access the 300G.
2. Select **Start -> Programs -> Tivoli Storage Manager -> Backup Client GUI**.
3. This will lead you to the **Tivoli Storage Manager GUI**.
4. Select **Backup**.
5. On the left pane, you select the directory to back up, or use the right pane to select individual files for the backups.

To restore, just follow the above steps, but select **Restore** in step 4 instead of **Backup**.

LAN-free backup setup 1

In this environment, we set up a TSM server on the MDC, and then ran our backup using the TSM client installed on the machine. See Figure 6-17 for the configuration of TSM LAN-free backup/archive setup. The FAST is the volume shared by the 300G to its clients, the MSS acts as the TSM disk storage pool. The 3570 library is the final destination of the data being backed up. The flow of the data based on this setup is from the FAST to the MSS to the 3570 library.

Data can also be backed up from the FAST direct to the 3570, but this is the usual backup being done LAN-free and without the benefit of SANergy. For this kind of scenario, please refer to *Using Tivoli Storage Manager in A SAN Environment*, SG24-6132.

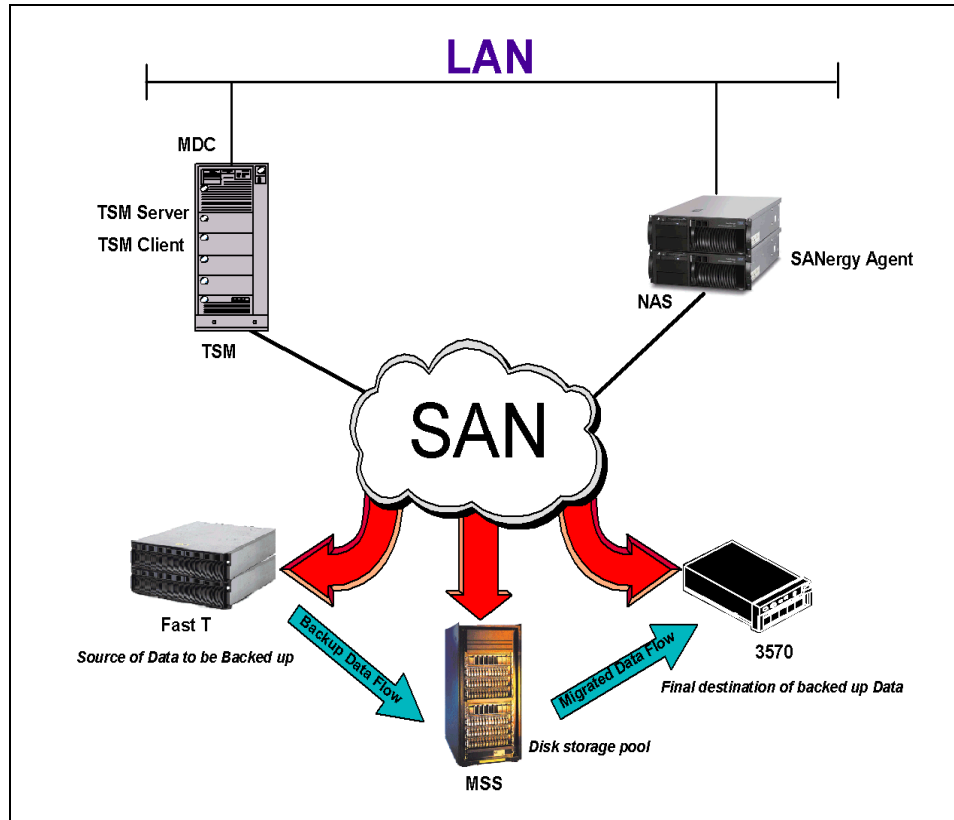


Figure 6-17 TSM server as the MDC

1. The 300G box is the owner of the FASTT volumes. These same volumes were shared to the TSM server running Windows 2000. Though shared, primary ownership still resides in the 300G.
2. The TSM server owns the volumes shared by the MSS. These same volumes are shared to the 300G, but the primary owner is still the TSM server.

Test scenario — without SANergy

1. We created 11 files as our test data. The files are all located in Drive T: the drive owned by the 300G. The volumes of this drive are from the FASTt. They are also shared to the TSM server but are owned by the 300G.
2. After the installation and configuration of the TSM server, SANergy, TSM Agent and the TSM client, we formatted additional disk storage pools using the disk owned by the TSM server which are located on the MSS subsystem.
3. We backed up the files using the Backup/Archive GUI client as shown in Figure 6-18.

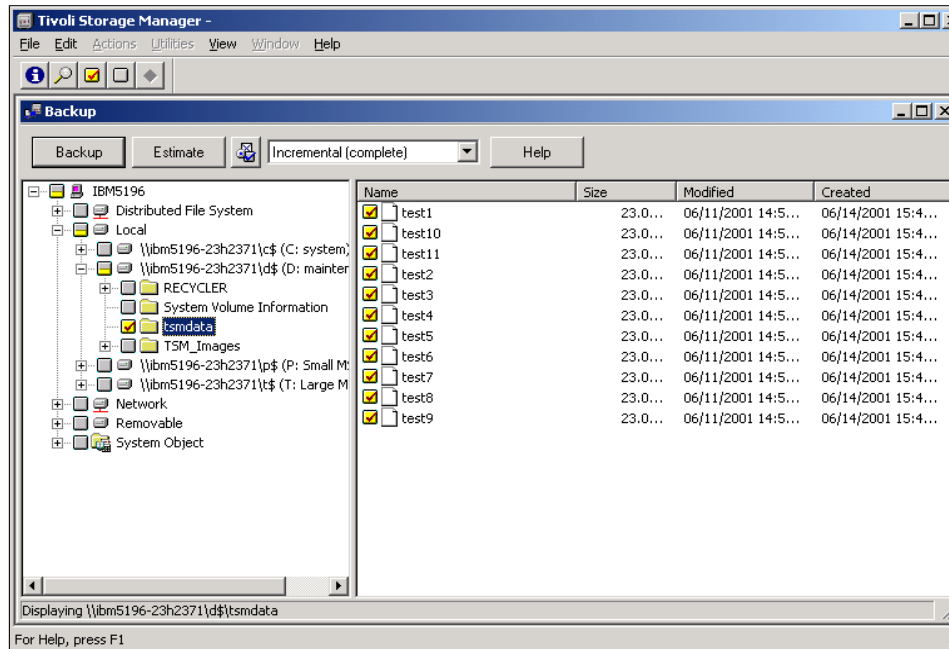


Figure 6-18 TSM backup test data

4. The first back up we tried, the SANergy agent was down. The data thus went through the LAN. After checking on the network status of the LAN card, prior and during the backup, we found a big discrepancy in the number of sent and received packets.

Test scenario — with SANergy

1. We selected the same set of files and tried to back them up with the Backup Always option active. This time we started SANergy.
2. The backup did not go through the LAN, as expected, but instead passed through the SAN, thus making the data flow from the FASTT to the MSS (with the FASTT being the home of the data, and the MSS being the home of the disk storage pools of the TSM server). The data transfer rate for the backup is shown in Figure 6-19.

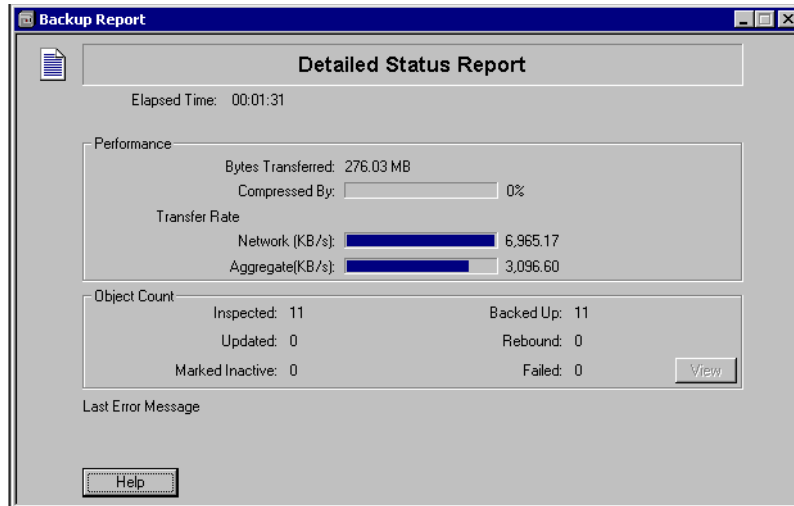


Figure 6-19 Result with SANergy running

3. With this type of configuration, not only did we do a LAN-less backup, we also did not utilize any CPU and memory on the owner of the data (the 300G).
4. We concluded by successfully testing migration of the data from the disk storage pools to the 3570 library.

LAN-free backup setup 2

On this second setup, we installed the TSM Agent and the TSM client on the 300G. We also configured the 300G to be our MDC. See Figure 6-20.

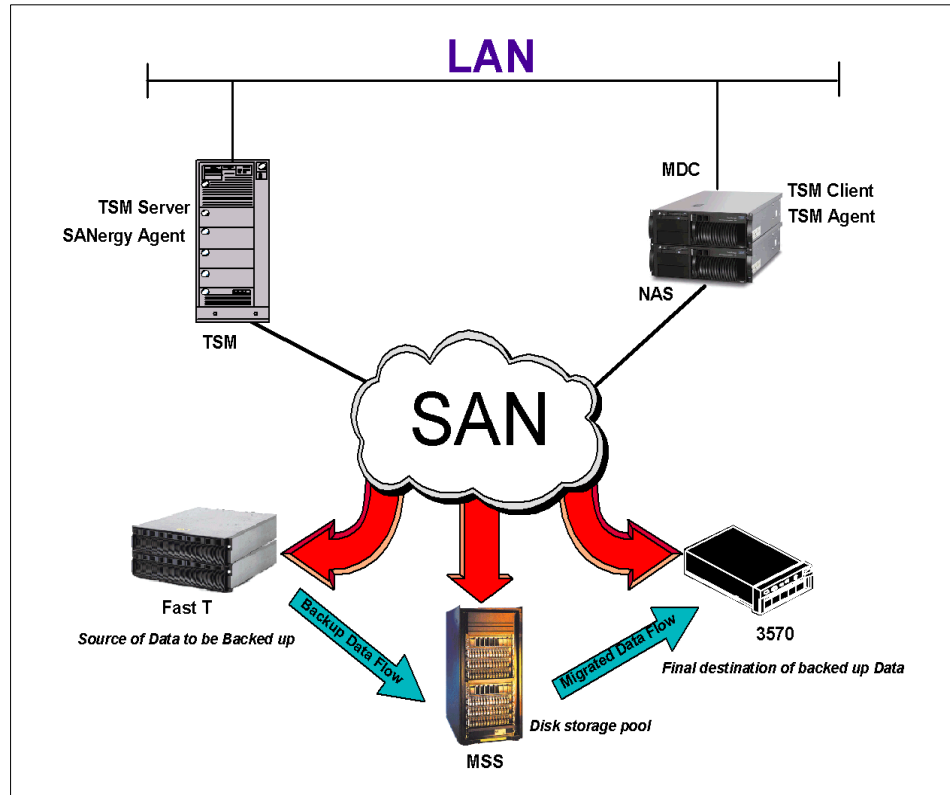


Figure 6-20 300G as the MDC

Test scenario

1. We selected the same files shared by the 300G.
2. We used the backup client GUI to back up the test data with the **backup always** option active.
3. The resulting transfer rates are shown in Figure 6-21. It is evident that the transfer rate is faster than what we saw in setup 1, but we noticed that in setup 2 the CPU utilization average jumped from 1% to 45%.

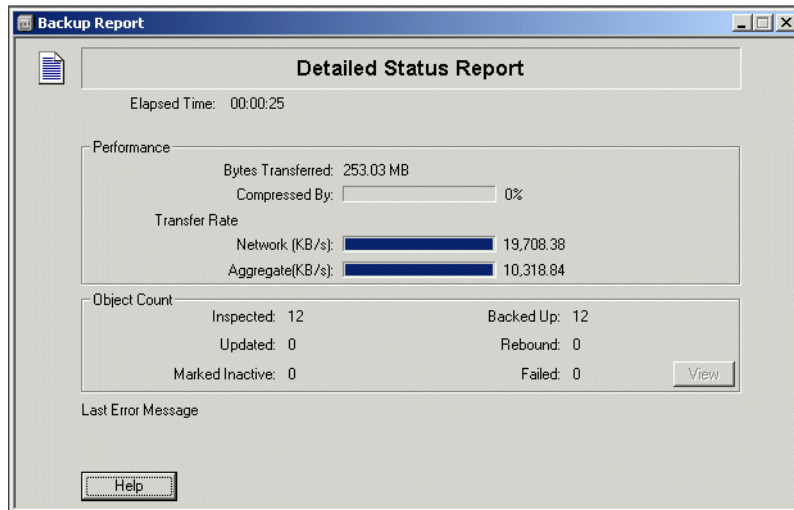


Figure 6-21 The backup results of setup 2

Although we have 11 files, TSM reports 12 files backed up, since it has also considered the directory entry as one additional backup object.

6.4.9 Backup results

It was evident that backup scenario 2 did an outstanding job transferring data. The transfer rate was more than 10 MBps. The only problem with this kind of setup is that, the CPU utilization increases by more than 40% when the backup is being run. The backup achieved these numbers because it went LAN-free.

For setup 1, the average data transfer for the same files used for setup 2 was almost 4 MBps. The transfer rate is 60% better in setup 1, but with setup 1, there was no CPU utilization increase in the 300G. While users who were accessing the 300G in setup 2 may have noticed that it was responding a little more slowly than usual, users in setup 1 noticed that the entire LAN was bogged down. While it is true that this is a resource trade-off, getting backups off the LAN is probably worth the CPU hit.

6.5 Recovering the 300G

As a critical gateway in your whole IT infrastructure, being able to quickly recover this system is vital, so in this section we will discuss how to go about it.

6.5.1 Recovering the 300G operating system using the recovery CD

The 300G box comes with a recovery CD that is used to recover the Operating System in the event of a total breakdown. This is discussed in detail in “Re-initializing the 300G” on page 72.

6.5.2 300G recovery method with TSM

With TSM in the picture to produce a LAN-free backup, we have to discuss the recovery of the 300G using the methods developed in recovering a Windows 2000 TSM client. We will be focusing on the recovery of Windows 2000 simply because the 300G’s operating system is Windows 2000. A prerequisite of this kind of recovery is the full system backup of the 300G using TSM. The steps that we will follow in recovering Windows 2000 are:

1. Perform a minimal installation of Windows 2000 server in a workgroup with network connectivity to the TSM server.
2. Install the Windows 2000 service pack that was running on the original system.
3. Create any additional disk partitions that were on the original system.
4. Install Tivoli Storage Manager client.
5. Restore file level data (files that reside in the Winnt directory) to the Windows 2000 boot / system partition using the TSM backup/restore client.
6. Restore the *entire* System Object as a single entity.
7. Reboot the system.
8. Restore data back onto other drives.
9. Reboot the system.
10. Check the system.

Recovery procedure

1. Perform any vendor specific configuration on the system. For example, define disk arrays, RAID partitions, and so on. If the 300G has a direct attached storage subsystem, it is better that the vendor specific configurations be done first.

2. Install Windows 2000 Advanced Server (the product must match the system being recovered).
3. Install any service packs, patches or drivers that were running on the original system that directly interact with components used by the restore process. For example, network card drivers, disk controller drivers, operating system patches.
4. Configure the Windows 2000 system to contact the Tivoli Storage Manager server. There may be a requirement to place entries in the TCP/IP hosts file or point the Windows 2000 system to a DNS server to achieve this.

Note: Before restoring using TSM backup/archive client, it is important that the original service packs be installed first.

5. Recreate the same number of disk partitions that were on the original system. Ensure the following partition properties match the original system:
6. Install the Tivoli Storage Manager client software.
7. Run the Tivoli Storage Manager client configuration wizard.
8. Before starting the restore, confirm the consistency of the System Object backup by running the command **query systemobject** from the Tivoli Storage Manager client command line.
9. Start the Tivoli Storage Manager Back Archive client and select the **Restore** tab.
10. Restore the boot / system partition.
11. Restore the entire System Object.
12. Restart the system.
13. Restore any other data onto other drives on the system.
14. Restart the system.
15. Confirm that the system restore has been successful.
16. If the system checks do not highlight any problems, the restore can now be considered complete.

For complete and more detailed discussion of these procedures, please see *Chapter 6* of the redbook *Deploying Tivoli Storage Manager for Windows 2000*, SG24-6141.

To do a bare metal restore of your 300G, see Section 3.1.2, “Re-initializing the 300G” on page 72.

6.6 NAS and the Network Data Management Protocol

In this section we provide a short overview how NAS and the Network Data Management Protocol (NDMP) work together. This protocol makes it possible for the NAS device to perform backup and restore operations to a direct attached tape device. So far, there is nothing new here — but now the whole procedure can be integrated and controlled by a TSM server.

6.6.1 NDMP overview

The NDMP function is included in the TSM 4.2.1 client and server. The exact product name is *TDP for NDMP*.

The NDMP controlled backups/restores can be performed in either two modes: Full file system image and differential file system image.

These are the benefits of NDMP:

- ▶ High performance, scalable backups and restore
- ▶ Backup to local tape devices without LAN network data movement
- ▶ NDMP isolates backup software from hardware/software changes on NAS appliances
- ▶ No special backup/restore code on NAS appliances

These are the NDMP limitations:

- ▶ No file level restore

6.6.2 Tape library setup

There are two options to set up the tape library:

- ▶ **Option 1:** The tape library is controlled directly by the TSM server, and the library robotic mechanism is controlled by direct attachment to the TSM server. Due to this, the NAS device, tape library, and TSM server must be in close proximity.
- ▶ **Option 2:** The tape library is controlled via the NAS file server. The robotic mechanism is controlled by passing SCSI commands through the NAS device. The control commands are initiated by the TSM server via TCP/IP.

The setup and the interaction with TSM and NAS using the new NDMP feature is shown in Figure 6-22.

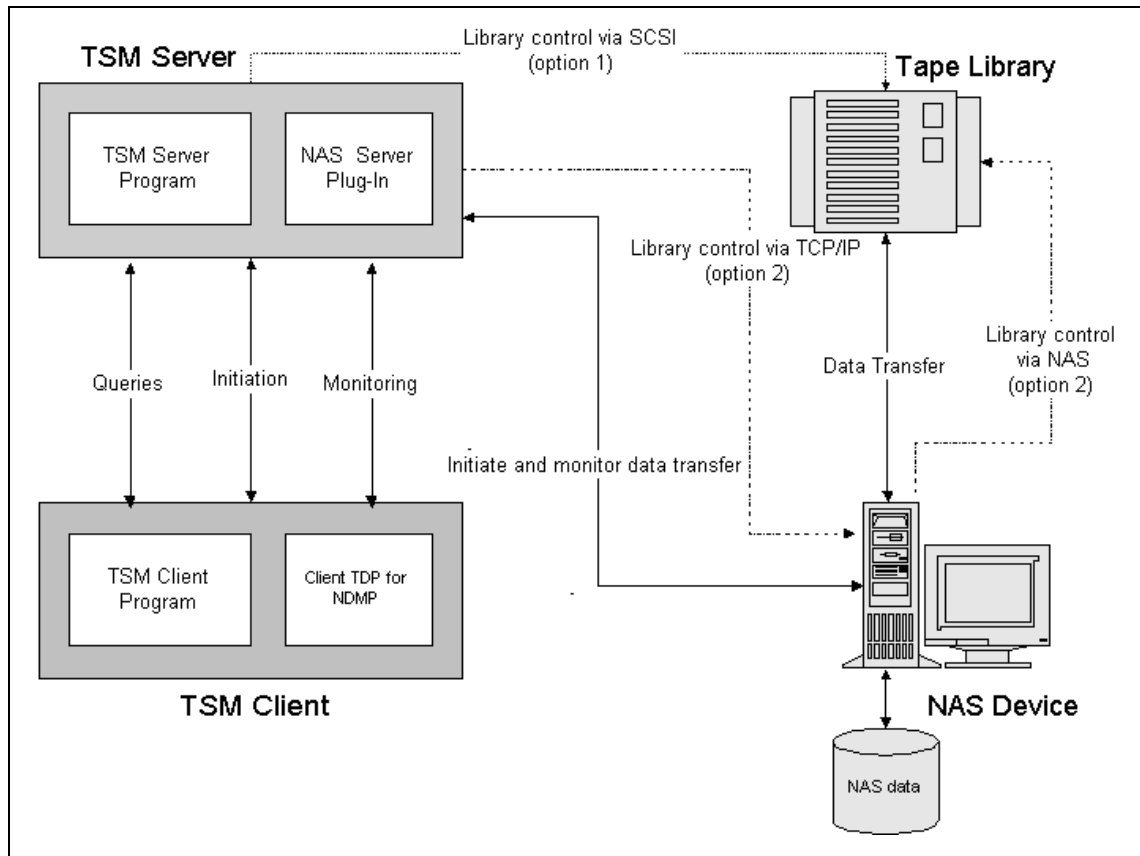


Figure 6-22 NAS and TSM interaction with NDMP

6.6.3 How TDP for NDMP backs up the NAS

The NAS device performs the following tasks:

- ▶ Accepts NDMP requests from TSM server
- ▶ Passes SCSI Cutups to library robotics
- ▶ Transfers data from file system to output tape
- ▶ Reports info during backup

The TSM Server performs the following tasks:

- ▶ Provides server commands for backup/restore operations
- ▶ Initiates and controls NDMP sessions with the NAS device
- ▶ Performs library options directly or via the NAS device
- ▶ Maintains tape library inventory
- ▶ Stores meta-data regarding stored images

The TSM client performs the following tasks:

- ▶ Provides user interface for initiating backup/restore operations
- ▶ Displays information regarding:
 - NAS devices
 - NAS filesystems
 - Existing file systems images stored by the TSM server
 - Progress/outcome of NAS backup/restore

Related publications

The publications listed in this section are considered particularly suitable for a more detailed discussion of the topics covered in this redbook.

IBM Redbooks

For information on ordering these publications, see “How to get IBM Redbooks” on page 322.

- ▶ *IP Storage Networking: NAS and iSCSI Solutions*, SG24-6240
- ▶ *A Practical Guide to Tivoli SANergy*, SG24-6146
- ▶ *Tivoli SANergy Administrator's Guide*, GC26-7389
- ▶ *IBM SAN Survival Guide*, SG24-6143
- ▶ *IBM Storage Solutions for Server Consolidation*, SG24-5355
- ▶ *Tivoli Storage Management Concepts*, SG24-4877
- ▶ *Getting Started with Tivoli Storage Manager: Implementation Guide*, SG24-5416
- ▶ *Using Tivoli Storage Manager in a SAN Environment*, SG24-6132
- ▶ *Tivoli Storage Manager Version 4.2: Technical Guide*, SG24-6277
- ▶ *Red Hat Linux Integration Guide for IBM eServer xSeries and Netfinity*, SG24-5853
- ▶ *AIX 5L and Windows 2000: Side by Side*, SG24-4784
- ▶ *Migrating IBM Netfinity Servers to Microsoft Windows 2000*, SG24-5854
- ▶ *Using TSM in a Clustered NT Environment*, SG24-5742
- ▶ *Planning and Implementing an IBM SAN*, SG24-6116
- ▶ *The IBM Enterprise Storage Server*, SG24-5645
- ▶ *ESS Solutions for Open Systems Storage: Compaq Alpha Server, HP and SUN*, SG24-6119
- ▶ *IBM Modular Storage Server - An Introduction Guide*, SG24-6103

Other resources

These publications are also relevant as further information sources:

- ▶ Peter Gerdson and Peter Kroeger *Kommunikationssysteme*, Band 1&2, Springer Verlag, 1994, ISBN 3540570047
- ▶ A. S. Tanenbaum, *Computer Networks*, Prentice Hall, 1996, ISBN 0133499456
- ▶ M. Schwartz, *Telecommunication Networks: Protocols, Modeling and Analysis*, Addison-Wesley, 1986, ISBN 020116423X
- ▶ Matt Welsh, Mathias Kalle Dalheimer, and Lar Kaufman, *Running Linux (3rd Edition)*, O'Reilly, 1999, ISBN 156592469X
- ▶ Scott M. Ballew, *Managing IP Networks with CISCO Routers*, O'Reilly, 1997, ISBN 1565923200
- ▶ Ellen Siever, et al., *Linux in a Nutshell (3rd Edition)*, O'Reilly, 2000, ISBN 0596000251
- ▶ Andreas Siegert, *The AIX Survival Guide*, Addison-Wesley, 1996, ISBN 0201593882
- ▶ William Boswell, *Inside Windows 2000 Server*, New Riders, 1999, ISBN 1562059297
- ▶ Paul Albitz and Cricket Liu, *DNS and BIND (4th Edition)*, O'Reilly, 2001, ISBN 0596001584
- ▶ Gary L. Olsen and Ty Loren Carlson, *Windows 2000 Active Directory Design and Deployment*, New Riders, 2000, ISBN 1578702429
- ▶ *Microsoft Windows 2000 Professional Resource Kit*, Microsoft Press, 2000, ISBN 1572318082
- ▶ D. Libertone, *Windows 2000 Cluster Server Guidebook*, Prentice Hall, 2000, ISBN 0130284696
- ▶ *Microsoft Services for UNIX version 2 white paper*, found at: <http://www.microsoft.com/WINDOWS2000/sfu/sfu2wp.asp>
- ▶ W. Richard Stevens, *UNIX Network Programming*, Prentice Hall, 1998, ISBN 013490012X

Referenced Web sites

These Web sites are also relevant as further information sources:

- ▶ IBM Storage
<http://www.storage.ibm.com/>
- ▶ IBM TotalStorage
<http://www.storage.ibm.com/ssg>
- ▶ IBM NAS
<http://www.storage.ibm.com/snetwork/nas/index.html>
- ▶ IBM TotalStorage 300G
http://www.storage.ibm.com/snetwork/nas/300g_product_page.htm
- ▶ IBM FastT200
<http://www.storage.ibm.com/hardsoft/products/fast200/fast200.htm>
- ▶ IBM Enterprise Storage Server (Shark)
<http://www.storage.ibm.com/hardsoft/products/ess/ess.htm>
- ▶ IBM Modular Storage Server
<http://www.storage.ibm.com/hardsoft/products/mss/mss.htm>
- ▶ Microsoft Technical Library
<http://www.microsoft.com/windows2000/techinfo/default.asp>
- ▶ Microsoft Services for UNIX
<http://www.microsoft.com/WINDOWS2000/sfu/default.asp>
- ▶ Tivoli
<http://www.tivoli.com/>
- ▶ Tivoli Sanergy Support
<http://www.tivoli.com/support/sanergy>
- ▶ Brocade
<http://www.brocade.com/>
- ▶ Storage Networking Industry Association
<http://www.snia.org/>
- ▶ Fibre Channel Industry Association
<http://www.fibrechannel.com/>
- ▶ Sysinternals Microsoft Tools
<http://www.sysinternals.com/>
- ▶ Linux Documentation
<http://www.linuxdoc.org/>
- ▶ Linux Kernel Resource
<http://www.kernel.org/>

- ▶ Red Hat Linux
<http://www.redhat.com/>
- ▶ SUSE Linux
http://www.suse.com/index_us.html

How to get IBM Redbooks

Search for additional Redbooks or redpieces, view, download, or order hardcopy from the Redbooks Web site:

ibm.com/redbooks

Also download additional materials (code samples or diskette/CD-ROM images) from this Redbooks site.

Redpieces are Redbooks in progress; not all Redbooks become redpieces and sometimes just a few chapters will be published this way. The intent is to get the information out more quickly than the formal publishing process allows.

IBM Redbooks collections

Redbooks are also available on CD-ROMs. Click the CD-ROMs button on the Redbooks Web site for information about all the CD-ROMs offered, as well as updates and formats.

Special notices

References in this publication to IBM products, programs or services do not imply that IBM intends to make these available in all countries in which IBM operates. Any reference to an IBM product, program, or service is not intended to state or imply that only IBM's product, program, or service may be used. Any functionally equivalent program that does not infringe any of IBM's intellectual property rights may be used instead of the IBM product, program or service.

Information in this book was developed in conjunction with use of the equipment specified, and is limited in application to those specific hardware and software products and levels.

IBM may have patents or pending patent applications covering subject matter 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 the IBM Director of Licensing, IBM Corporation, North Castle Drive, Armonk, NY 10504-1785.

Licensees of this program who wish to have information about it for the purpose of enabling: (i) the exchange of information between independently created programs and other programs (including this one) and (ii) the mutual use of the information which has been exchanged, should contact IBM Corporation, Dept. 600A, Mail Drop 1329, Somers, NY 10589 USA.

Such information may be available, subject to appropriate terms and conditions, including in some cases, payment of a fee.

The information contained in this document has not been submitted to any formal IBM test and is distributed AS IS. The use of this information or the implementation of any of these techniques is a customer responsibility and depends on the customer's ability to evaluate and integrate them into the customer's operational environment. While each item may have been reviewed by IBM for accuracy in a specific situation, there is no guarantee that the same or similar results will be obtained elsewhere. Customers attempting to adapt these techniques to their own environments do so at their own risk.

Any pointers in this publication to external Web sites are provided for convenience only and do not in any manner serve as an endorsement of these Web sites.

The following terms are trademarks of other companies:

Tivoli, Manage. Anything. Anywhere., The Power To Manage., Anything. Anywhere., TME, NetView, Cross-Site, Tivoli Ready, Tivoli Certified, Planet Tivoli, and Tivoli Enterprise are trademarks or registered trademarks of Tivoli Systems Inc., an IBM company, in the United States, other countries, or both. In Denmark, Tivoli is a trademark licensed from Kjøbenhavns Sommer - Tivoli A/S.

C-bus is a trademark of Corollary, Inc. in the United States and/or other countries.

Java and all Java-based trademarks and logos are trademarks or registered trademarks of Sun Microsystems, Inc. in the United States and/or other countries.

Microsoft, Windows, Windows NT, Windows 2000 and the Windows logo are trademarks of Microsoft Corporation in the United States and/or other countries.

PC Direct is a trademark of Ziff Communications Company in the United States and/or other countries and is used by IBM Corporation under license.

ActionMedia, LANDesk, MMX, Pentium and ProShare are trademarks of Intel Corporation in the United States and/or other countries.

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

SET, SET Secure Electronic Transaction, and the SET Logo are trademarks owned by SET Secure Electronic Transaction LLC.

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

Glossary

A

Agent A software entity that runs on endpoints and provides management capability for other hardware or software. An example is an SNMP agent. An agent has the ability to spawn other processes.

AL See arbitrated loop.

Allocated storage The space that is allocated to volumes, but not assigned.

Allocation The entire process of obtaining a volume and unit of external storage, and setting aside space on that storage for a data set.

Arbitrated loop A Fibre Channel interconnection technology that allows up to 126 participating node ports and one participating fabric port to communicate. See also Fibre Channel Arbitrated Loop and loop topology.

Array An arrangement of related disk drive modules that have been assigned to a group.

B

Bandwidth A measure of the data transfer rate of a transmission channel.

Bridge Facilitates communication with LANs, SANs, and networks with dissimilar protocols.

C

Client A function that requests services from a server, and makes them available to the user. A term used in an environment to identify a machine that uses the resources of the network.

Client authentication The verification of a client in secure communications where the identity of a server or browser (client) with whom you wish to communicate is discovered. A sender's authenticity is demonstrated by the digital certificate issued to the sender.

Client-server relationship Any process that provides resources to other processes on a network is a server. Any process that employs these resources is a client. A machine can run client and server processes at the same time.

Console A user interface to a server.

D

DATABASE 2 (DB2) A relational database management system. DB2 Universal Database is the relational database management system that is Web-enabled with Java support.

Device driver A program that enables a computer to communicate with a specific device, for example, a disk drive.

Disk group A set of disk drives that have been configured into one or more logical unit numbers. This term is used with RAID devices.

E

Enterprise network A geographically dispersed network under the backing of one organization.

Enterprise Storage Server Provides an intelligent disk storage subsystem for systems across the enterprise.

Event In the Tivoli environment, any significant change in the state of a system resource, network resource, or network application. An event can be generated for a problem, for the resolution of a problem, or for the successful completion of a task. Examples of events are: the normal starting and stopping of a process, the abnormal termination of a process, and the malfunctioning of a server.

F

Fabric The Fibre Channel employs a fabric to connect devices. A fabric can be as simple as a single cable connecting two devices. The term is often used to describe a more complex network utilizing hubs, switches, and gateways.

FC See Fibre Channel.

FCS See Fibre Channel standard.

Fiber optic The medium and the technology associated with the transmission of information along a glass or plastic wire or fiber.

Fibre Channel A technology for transmitting data between computer devices at a data rate of up to 1 Gb. It is especially suited for connecting computer servers to shared storage devices and for interconnecting storage controllers and drives.

Fibre Channel Arbitrated Loop A reference to the FC-AL standard, a shared gigabit media for up to 127 nodes, one of which can be attached to a switch fabric. See also arbitrated loop and loop topology. Refer to American National Standards Institute (ANSI) X3T11/93-275.

Fibre Channel standard An ANSI standard for a computer peripheral interface. The I/O interface defines a protocol for communication over a serial interface that configures attached units to a communication fabric. Refer to ANSI X3.230-199x.

File system An individual file system on a host. This is the smallest unit that can monitor and extend. Policy values defined at this level override those that might be defined at higher levels.

G

Gateway In the SAN environment, a gateway connects two or more different remote SANs with each other. A gateway can also be a server on which a gateway component runs.

H

Hardware zoning Hardware zoning is based on physical ports. The members of a zone are physical ports on the fabric switch. It can be implemented in the following configurations: one to one, one to many, and many to many.

HBA See host bus adapter.

Host Any system that has at least one internet address associated with it. A host with multiple network interfaces can have multiple internet addresses associated with it. This is also referred to as a server.

Host bus adapter (HBA) A Fibre Channel HBA connection that allows a workstation to attach to the SAN network.

Hub A Fibre Channel device that connects up to 126 nodes into a logical loop. All connected nodes share the bandwidth of this one logical loop. Hubs automatically recognize an active node and insert the node into the loop. A node that fails or is powered off is automatically removed from the loop.

IP Internet protocol.

J

Java A programming language that enables application developers to create object-oriented programs that are very secure, portable across different machine and operating system platforms, and dynamic enough to allow expandability.

Java runtime environment (JRE) The underlying, invisible system on your computer that runs applets the browser passes to it.

Java Virtual Machine (JVM) The execution environment within which Java programs run. The Java virtual machine is described by the Java Machine Specification which is published by Sun Microsystems. Because the Tivoli Kernel Services is based on Java, nearly all ORB and component functions execute in a Java virtual machine.

JBOD Just a Bunch Of Disks.

JRE See Java runtime environment.

JVM See Java Virtual Machine.

L

Logical unit number (LUN) The LUNs are provided by the storage devices attached to the SAN. This number provides you with a volume identifier that is unique among all storage servers. The LUN is synonymous with a physical disk drive or a SCSI device. For disk subsystems such as the IBM Enterprise Storage Server, a LUN is a logical disk drive. This is a unit of storage on the SAN which is available for assignment or unassignment to a host server.

Loop topology In a loop topology, the available bandwidth is shared with all the nodes connected to the loop. If a node fails or is not powered on, the loop is out of operation. This can be corrected using a hub. A hub opens the loop when a new node is connected and closes it when a node disconnects. See also Fibre Channel Arbitrated Loop and arbitrated loop.

LUN See logical unit number.

LUN assignment criteria The combination of a set of LUN types, a minimum size, and a maximum size used for selecting a LUN for automatic assignment.

LUN masking This allows or blocks access to the storage devices on the SAN. Intelligent disk subsystems like the IBM Enterprise Storage Server provide this kind of masking.

M

Managed object A managed resource.

Managed resource A physical element to be managed.

Management Information Base (MIB) A logical database residing in the managed system which defines a set of MIB objects. A MIB is considered a logical database because actual data is not stored in it, but rather provides a view of the data that can be accessed on a managed system.

MIB See Management Information Base.

MIB object A MIB object is a unit of managed information that specifically describes an aspect of a system. Examples are CPU utilization, software name, hardware type, and so on. A collection of related MIB objects is defined as a MIB.

N

Network topology A physical arrangement of nodes and interconnecting communications links in networks based on application requirements and geographical distribution of users.

N_Port node port A Fibre Channel-defined hardware entity at the end of a link which provides the mechanisms necessary to transport information units to or from another node.

NL_Port node loop port A node port that supports arbitrated loop devices.

O

Open system A system whose characteristics comply with standards made available throughout the industry, and therefore can be connected to other systems that comply with the same standards.

P

Point-to-point topology It consists of a single connection between two nodes. All the bandwidth is dedicated for these two nodes.

Port An end point for communication between applications, generally referring to a logical connection. A port provides queues for sending and receiving data. Each port has a port number for identification. When the port number is combined with an Internet address, it is called a socket address.

Port zoning In Fibre Channel environments, port zoning is the grouping together of multiple ports to form a virtual private storage network. Ports that are members of a group or zone can communicate with each other but are isolated from ports in other zones. See also LUN masking and subsystem masking.

Protocol The set of rules governing the operation of functional units of a communication system if communication is to take place. Protocols can determine low-level details of machine-to-machine interfaces, such as the order in which bits from a byte are sent. They can also determine high-level exchanges between application programs, such as file transfer.

R

RAID Redundant array of inexpensive or independent disks. A method of configuring multiple disk drives in a storage subsystem for high availability and high performance.

S

SAN See storage area network.

SAN agent A software program that communicates with the manager and controls the subagents. This component is largely platform independent. See also subagent.

SCSI Small Computer System Interface. An ANSI standard for a logical interface to computer peripherals and for a computer peripheral interface. The interface utilizes a SCSI logical protocol over an I/O interface that configures attached targets and initiators in a multi-drop bus topology.

Server A program running on a mainframe, workstation, or file server that provides shared services. This is also referred to as a host.

Shared storage Storage within a storage facility that is configured such that multiple homogeneous or divergent hosts can concurrently access the storage. The storage has a uniform appearance to all hosts. The host programs that access the storage must have a common model for the information on a storage device. You need to design the programs to handle the effects of concurrent access.

Simple Network Management Protocol (SNMP) A protocol designed to give a user the capability to remotely manage a computer network by polling and setting terminal values and monitoring network events.

SNMP See Simple Network Management Protocol.

SNMP agent An implementation of a network management application which is resident on a managed system. Each node that is to be monitored or managed by an SNMP manager in a TCP/IP network, must have an SNMP agent resident. The agent receives requests to either retrieve or modify management information by referencing MIB objects. MIB objects are referenced by the agent whenever a valid request from an SNMP manager is received.

SNMP manager A managing system that executes a managing application or suite of applications. These applications depend on MIB objects for information that resides on the managed system.

SNMP trap A message that is originated by an agent application to alert a managing application of the occurrence of an event.

Software zoning Is implemented within the Simple Name Server (SNS) running inside the fabric switch. When using software zoning, the members of the zone can be defined with: node WWN, port WWN, or physical port number. Usually the zoning software also allows you to create symbolic names for the zone members and for the zones themselves.

SQL Structured Query Language.

Storage administrator A person in the data processing center who is responsible for defining, implementing, and maintaining storage management policies.

Storage area network (SAN) A managed, high-speed network that enables any-to-any interconnection of heterogeneous servers and storage systems.

Subagent A software component of SAN products which provides the actual remote query and control function, such as gathering host information and communicating with other components. This component is platform dependent. See also SAN agent.

Subsystem masking The support provided by intelligent disk storage subsystems like the Enterprise Storage Server. See also LUN masking and port zoning.

Switch A component with multiple entry and exit points or ports that provide dynamic connection between any two of these points.

Switch topology A switch allows multiple concurrent connections between nodes. There can be two types of switches, circuit switches and frame switches. Circuit switches establish a dedicated connection between two nodes. Frame switches route frames between nodes and establish the connection only when needed. A switch can handle all protocols.

T

TCP See Transmission Control Protocol.

TCP/IP Transmission Control Protocol/Internet Protocol.

Topology An interconnection scheme that allows multiple Fibre Channel ports to communicate. For example, point-to-point, arbitrated loop, and switched fabric are all Fibre Channel topologies.

Transmission Control Protocol (TCP) A reliable, full duplex, connection-oriented, end-to-end transport protocol running on top of IP.

W

WAN Wide Area Network.

Z

Zoning In Fibre Channel environments, zoning allows for finer segmentation of the switched fabric. Zoning can be used to instigate a barrier between different environments. Ports that are members of a zone can communicate with each other but are isolated from ports in other zones. Zoning can be implemented in two ways: hardware zoning and software zoning.

For more information on IBM terminology, see the IBM Storage Glossary of Terms at:

<http://www.storage.ibm.com/glossary.htm>

For more information on Tivoli terminology, see the Tivoli Glossary at:

<http://www.tivoli.com/support/documents/glossary/termsm03.htm>

Abbreviations and acronyms

ABI	Application Binary Interface	BDC	Backup Domain Controller
ACE	Access Control Entries	BIND	Berkeley Internet Name Domain
ACL	Access Control List	BNU	Basic Network Utilities
AD	Microsoft Active Directory	BOS	Base Operating System
ADSM	ADSTAR Distributed Storage Manager	BRI	Basic Rate Interface
AFS	Andrew File System	BSD	Berkeley Software Distribution
AIX	Advanced Interactive eXecutive	BSOD	Blue Screen of Death
ANSI	American National Standards Institute	BUMP	Bring-Up Microprocessor
APA	All Points Addressable	CA	Certification Authorities
API	Application Programming Interface	CAL	Client Access License
APPC	Advanced Program-to-Program	C-SPOC	Cluster single point of control
APPN	Advanced Peer-to-Peer Networking	CDE	Common Desktop Environment
ARC	Advanced RISC Computer	CDMF	Commercial Data Masking Facility
ARPA	Advanced Research Projects Agency	CDS	Cell Directory Service
ASCII	American National Standard Code for Information Interchange	CERT	Computer Emergency Response Team
ATE	Asynchronous Terminal Emulation	CGI	Common Gateway Interface
ATM	Asynchronous Transfer Mode	CHAP	Challenge Handshake Authentication
AVI	Audio Video Interleaved	CIDR	Classless InterDomain Routing
		CIFS	Common Internet File System
		CMA	Concert Multi-threaded Architecture
		CO	Central Office

COPS	Computer Oracle and Password System	EFS	Encrypting File Systems
CPI-C	Common Programming Interface for Communications	EGID	Effective Group Identifier
CPU	Central Processing Unit	EISA	Extended Industry Standard Architecture
CSNW	Client Service for NetWare	EMS	Event Management Services
CSR	Client/server Runtime	EPROM	Erasable Programmable Read-Only
DAC	Discretionary Access Controls	ERD	Emergency Repair Disk
DARPA	Defense Advanced Research Projects Agency	ERP	Enterprise Resources Planning
DASD	Direct Access Storage Device	ERRM	Event Response Resource Manager
DBM	Database Management	ESCON	Enterprise System Connection
DCE	Distributed Computing Environment	ESP	Encapsulating Security Payload
DCOM	Distributed Component Object Model	ESS	Enterprise Storage Server
DDE	Dynamic Data Exchange	EUID	Effective User Identifier
DDNS	Dynamic Domain Name System	FAT	File Allocation Table
DEN	Directory Enabled Network	FC	Fibre Channel
DES	Data Encryption Standard	FDDI	Fiber Distributed Data Interface
DFS	Distributed File System	FDPR	Feedback Directed Program Restructure
DHCP	Dynamic Host Configuration Protocol	FIFO	First In/First Out
DLC	Data Link Control	FIRST	Forum of Incident Response and Security
DLL	Dynamic Load Library	FQDN	Fully Qualified Domain Name
DS	Differentiated Service	FSF	File Storage Facility
DSA	Directory Service Agent	FTP	File Transfer Protocol
DSE	Directory Specific Entry	FtDisk	Fault-Tolerant Disk
DNS	Domain Name System	GC	Global Catalog
DTS	Distributed Time Service	GDA	Global Directory Agent
		GDI	Graphical Device Interface

GDS	Global Directory Service	I/O	Input/Output
GID	Group Identifier	IP	Internet Protocol
GL	Graphics Library	IPC	Interprocess Communication
GSNW	Gateway Service for NetWare	IPL	Initial Program Load
GUI	Graphical User Interface	IPsec	Internet Protocol Security
HA	High Availability	IPX	Internetwork Packet eXchange
HACMP	High Availability Cluster Multiprocessing	ISA	Industry Standard Architecture
HAL	Hardware Abstraction Layer	iSCSI	SCSI over IP
HBA	Host Bus Adapter	ISDN	Integrated Services Digital Network
HCL	Hardware Compatibility List	ISNO	Interface-specific Network Options
HSM	Hierarchical Storage Management	ISO	International Organization for Standardization
HTTP	Hypertext Transfer Protocol	ISS	Interactive Session Support
IBM	International Business Machines Corporation	ISV	Independent Software Vendor
ICCM	Inter-Client Conventions Manual	ITSEC	Initial Technology Security Evaluation
IDE	Integrated Drive Electronics	ITSO	International Technical Support Organization
IDL	Interface Definition Language	ITU	International Telecommunications Union
IDS	Intelligent Disk Subsystem	IXC	Inter Exchange Carrier
IEEE	Institute of Electrical and Electronic Engineers	JBOD	Just a Bunch of Disks
IETF	Internet Engineering Task Force	JFS	Journaled File System
IGMP	Internet Group Management Protocol	JIT	Just-In-Time
IIS	Internet Information Server	L2F	Layer 2 Forwarding
IKE	Internet Key Exchange	L2TP	Layer 2 Tunneling Protocol
IMAP	Internet Message Access Protocol	LAN	Local Area Network

LCN	Logical Cluster Number	MOCL	Managed Object Class Library
LDAP	Lightweight Directory Access Protocol	MPTN	Multi-protocol Transport Network
LFS	Log File Service (Windows NT)	MS-DOS	Microsoft Disk Operating System
LFS	Logical File System (AIX)	MSCS	Microsoft Cluster Server
LFT	Low Function Terminal	MSS	Maximum Segment Size
JNDI	Java Naming and Directory Interface	MSS	Modular Storage Server
LOS	Layered Operating System	MWC	Mirror Write Consistency
LP	Logical Partition	NAS	Network Attached Storage
LPC	Local Procedure Call	NBC	Network Buffer Cache
LPD	Line Printer Daemon	NBF	NetBEUI Frame
LPP	Licensed Program Product	NBPI	Number of Bytes per I-node
LRU	Least Recently Used	NCP	NetWare Core Protocol
LSA	Local Security Authority	NCS	Network Computing System
LTG	Local Transfer Group	NCSC	National Computer Security Center
LUID	Login User Identifier	NDIS	Network Device Interface Specification
LUN	Logical Unit Number	NDMP	Network Data Management Protocol
LVCB	Logical Volume Control Block	NDS	NetWare Directory Service
LVDD	Logical Volume Device Driver	NETID	Network Identifier
LVM	Logical Volume Manager	NFS	Network File System
MBR	Master Boot Record	NIM	Network Installation Management
MCA	Micro Channel Architecture	NIS	Network Information System
MDC	Meta Data Controller	NIST	National Institute of Standards and Technology
MFT	Master File Table	NLS	National Language Support
MIPS	Million Instructions Per Second		
MMC	Microsoft Management Console		

NNS	Novell Network Services	PDC	Primary Domain Controller
NSAPI	Netscape Commerce Server's Application	PDF	Portable Document Format
NTFS	NT File System	PDT	Performance Diagnostic Tool
NTLDR	NT Loader	PEX	PHIGS Extension to X
NTLM	NT LAN Manager	PFS	Physical File System
NTP	Network Time Protocol	PHB	Per Hop Behavior
NTVDM	NT Virtual DOS Machine	PHIGS	Programmer's Hierarchical Interactive Graphics System
NVRAM	Non-Volatile Random Access Memory	PID	Process Identification Number
NetBEUI	NetBIOS Extended User Interface	PIN	Personal Identification Number
NetDDE	Network Dynamic Data Exchange	PMTU	Path Maximum Transfer Unit
OCS	On-Chip Sequencer	POP	Post Office Protocol
ODBC	Open Database Connectivity	POSIX	Portable Operating System Interface for Computer Environment
ODM	Object Data Manager	POST	Power-On Self Test
OLTP	OnLine Transaction Processing	PP	Physical Partition
OMG	Object Management Group	PPP	Point-to-Point Protocol
ONC	Open Network Computing	PPTP	Point-to-Point Tunneling Protocol
OS	Operating System	PreP	PowerPC Reference Platform
OSF	Open Software Foundation	PSM	Persistent Storage Manager
PAL	Platform Abstract Layer	PSN	Program Sector Number
PAM	Pluggable Authentication Module	PSSP	Parallel System Support Program
PAP	Password Authentication Protocol	PV	Physical Volume
PBX	Private Branch Exchange	PVID	Physical Volume Identifier
PCI	Peripheral Component Interconnect	QoS	Quality of Service
PCMCIA	Personal Computer Memory Card		

RACF	Resource Access Control Facility	SDK	Software Developer's Kit
RAID	Redundant Array of Independent Disks	SFG	Shared Folders Gateway
RAS	Remote Access Service	SFU	Services for UNIX
RDBMS	Relational Database Management System	SID	Security Identifier
RFC	Request for Comments	SLIP	Serial Line Internet Protocol
RGID	Real Group Identifier	SMB	Server Message Block
RISC	Reduced Instruction Set Computer	SMIT	System Management Interface Tool
RMC	Resource Monitoring and Control	SMP	Symmetric Multiprocessor
RMSS	Reduced-Memory System Simulator	SMS	Systems Management Server
ROLTP	Relative OnLine Transaction Processing	SNA	Systems Network Architecture
ROS	Read-Only Storage	SNAPI	SNA Interactive Transaction Program
RPC	Remote Procedure Call	SNMP	Simple Network Management Protocol
RRIP	Rock Ridge Internet Protocol	SP	System Parallel
RSCT	Reliable Scalable Cluster Technology	SPX	Sequenced Packet eXchange
RSM	Removable Storage Management	SQL	Structured Query Language
RSVP	Resource Reservation Protocol	SRM	Security Reference Monitor
SACK	Selective Acknowledgments	SSA	Serial Storage Architecture
SAK	Secure Attention Key	SSL	Secure Sockets Layer
SAM	Security Account Manager	SUSP	System Use Sharing Protocol
SAN	Storage Area Network	SVC	Serviceability
SASL	Simple Authentication and Security Layer	SWS	Silly Window Syndrome
SATAN	Security Analysis Tool for Auditing	TAPI	Telephone Application Program Interface
SCSI	Small Computer System Interface	TCB	Trusted Computing Base

TCP/IP	Transmission Control Protocol/Internet Protocol	VGDA	Volume Group Descriptor Area
TCSEC	Trusted Computer System Evaluation	VGSA	Volume Group Status Area
TDI	Transport Data Interface	VGID	Volume Group Identifier
TDP	Tivoli Data Protection	VIPA	Virtual IP Address
TLS	Transport Layer Security	VMM	Virtual Memory Manager
TOS	Type of Service	VP	Virtual Processor
TSM	Tivoli Storage Manager	VPD	Vital Product Data
TTL	Time to Live	VPN	Virtual Private Network
UCS	Universal Code Set	VRMF	Version, Release, Modification, Fix
UDB	Universal Database	VSM	Virtual System Management
UDF	Universal Disk Format	W3C	World Wide Web Consortium
UDP	User Datagram Protocol	WAN	Wide Area Network
UFS	UNIX File System	WFW	Windows for Workgroups
UID	User Identifier	WINS	Windows Internet Name Service
UMS	Ultimedia Services	WLM	Workload Manager
UNC	Universal Naming Convention	WOW	Windows-16 on Win32
UPS	Uninterruptable Power Supply	WWW	World Wide Web
URL	Universal Resource Locator	WYSIWYG	What You See Is What You Get
USB	Universal Serial Bus	WinMSD	Windows Microsoft Diagnostics
UTC	Universal Time Coordinated	XCMF	X/Open Common Management Framework
UUCP	UNIX to UNIX Communication Protocol	XDM	X Display Manager
UUID	Universally Unique Identifier	XDMCP	X Display Manager Control Protocol
VAX	Virtual Address eXtension	XDR	eXternal Data Representation
VCN	Virtual Cluster Name	XNS	XEROX Network Systems
VFS	Virtual File System	XPG4	X/Open Portability Guide
VG	Volume Group		

Index

Symbols

?CLUS 265
?FREE 265

Numerics

2109-S08 64
2109-S16 64

A

Active Directory 163
Admin.msc 162, 288
AIX clients 161, 232
arbitrated loop 24
ASIC 67

B

backup 277
 archival 279
 data 287
 database 296
 LAN-based 290
 LAN-free 64, 296
 maintenance partition 286
 NAS 300G 286
 native 278
 operating system 286
 point in time 278
 port address 291
 result 313
 setup LAN-free 308
 using terminal services 280
 with TSM and SANergy 308
 wizard 282
best of breed storage 29
Block I/O 14, 17, 21, 24, 26, 34, 36

C

cache 278
CIFS 13, 19, 237
CIFS share
 with MSCS 272

Class 2 67
Class 3 67
Class F 67
Classes of Service 67
Common Internet File System 13
connectivity 19
crfs command 161

D

data integrity 21, 28
data migration 29
database 21, 36, 54
devconfig.txt 305
disaster tolerance 28
disk drive
 identifying 243
disk group 197
disk management 137
disk signature 139
disk volumes
 fusing 256, 259
 setting MDC 245
 setting SANergy ownership 242
dsmsta 305
dsmsta.opt 305
DYNIX 64

E

e-business 29
Enterprise Resource Planning 54
Enterprise Storage Server. see ESS
ERP 54
ESCON 53
ESS 52, 261
 access 124
 add volume 127
 benefits 54
 databases 54
 disaster recovery 55
 disk groups 125
 fabric support 57
 host adapter ports 123

- host type 122
- models 53
- modify volume assignments 130
- overview 52
- performance 54
- PPRC 56
- RAID array 126
- SAN 57, 116
- setup 116
- storage allocation 120
- volume attributes 129
- zoning 116

F

- F_Port 67
- failover 260
- FAStT 60
 - benefits 62
 - disaster recovery 63
 - expansion enclosure 61
 - models 61
 - overview 60
- FAStT Check 77
- FAStT200 64, 261
 - create array 88
 - host group 94
 - logical drive parameters 91
 - LUN mapping 97
 - new host 95
 - new host port 96
 - setup 87
 - storage partitioning 93
 - WWN 96
- FAStT500 64
- Fibre Array Storage Technology 60
- Fibre Channel 23
- fibre channel switch 63
- FICON 53
- File I/O 14, 17, 21, 24, 34
- file servers 15
- file sharing 19
- filesystems 12
- FL_Port 67
- FlashCopy 56, 279
- fstab 159
- FTP
 - access 162

H

- HACMP 116
- high availability 260
- HP-UX clients 159

I

- I/O 12
- IBM Director 50
- IETF 38, 39
- IP mode 173

L

- LAN bandwidth 21
- LAN free 38
- Linux clients 159
- Local Area Networks 5
- LUN identifiers 105

M

- managed buses 241
- management 30
 - data 30
 - SAN 30
- MDC 33, 238, 308
 - SANergy installation 253
- Meta Data Controller. *see* MDC
- Microsoft Cluster Server. *see* MSCS
- Modular Storage Server. *see* MSS
- MSCS 167
 - add second node 192
 - administration 194
 - client connectivity 227
 - cluster administrator 269
 - defining CIFS share 272
 - disk group 197
 - drive letter 181
 - failback 205
 - file share 218, 272
 - file shares 207
 - first node setup 184
 - heartbeat 176
 - IP resource 209
 - join Domain 178
 - module for SANergy 262
 - move group 201
 - network name resource 213
 - NFS resource 224

- preferred ownership 203
- private network 176, 184
- public network 176
- quorum 190
- quorum resource 262
- resource balancing 200
- resource groups 262
- resource types 262
- SANergy 260
- SANergy volume 269
- second node setup 169
- setup 187
- share permissions 222
- with SANergy 262
- wizard 188
- MSS 261
 - assign logical drive 112
 - assign LUN to host 115
 - assign unit number 110
 - benefits 59
 - command line interface 98
 - create LUN 107
 - create partition 109
 - create RAIDset 107
 - define host 112
 - disaster recovery 60
 - expansion enclosures 58
 - failover 99
 - FlashCopy 60
 - initialize RAIDset 109
 - models 58
 - multiple-bus failover mode 102
 - offsets 112
 - overview 57
 - SAN 99
 - setup 98
 - setup failover modes 106
 - transparent failover mode 100
 - zoning 99
- multiple-bus failover mode 111

N

N_Ports 67

NAS

- appliances 16
- benefits 18
- enhanced backup 20
- File I/O 17

- manageability 20
- Network Attached Storage 15
 - SAN 31
 - user interface 287
- NAS 300G 42
 - configure locally 76
 - disks and volumes 133
 - dual-node 42
 - G00 42
 - G25 42, 167
 - ownership of pooled storage 132
 - remote management 73
 - sample connectivity 51
 - single-node 42
- NAS backup assistant 287
- NDMP 316
 - overview 316
- network appliances 16
- Network Data Management Protocol 316
- network file system protocols 12
- NFS 12, 13, 19, 237
 - client 154
 - user name mapping 162
- NFS shares 259
- NIS
 - Integration 165
 - master 165
 - slave 165
- non-blocking shared memory 67
- NTBackup 287
- NTFS 236

O

- Open System Storage 121
- open system storage
 - ESS 119
- open systems management 31
- oplocks 14
- OSI
 - compared to TCP/IP 7
 - model 7

P

- Parallel Sysplex 55
- partition 141
- performance 20
- persistent image 279
- point to point 24

Protocols 12
protocols 12
PSM 278, 279
 access 287

Q

quorum 261
quorum disk 262

R

Recovery CD 314
Redbooks Web site 322
 Contact us xxiii
resource pooling 18
restore 277
 NAS 300G 314
 operating system 314
 procedure 314
 using NT Backup 282
 with TSM 314
 with TSM and SANergy 308
 wizard 284
RS/6000 116

S

SAN 22, 76
 backup and recovery 27
 benefits 26
 Block I/O 26
 costs 29
 data gateways 64
 data movement 27
 data sharing 27
 fabrics 65
 high availability 28
 inter-operability 30
 LAN-free 27
 non-disruptive scalability for growth 27
 performance 28
 Server to server 25
 Server to storage 25
 server-free 27
 skills required 30
 storage consolidation 26
 storage to storage 25
 SWANs 30
 tape pooling 27

SAN with NAS 31
SANergy 235
 architecture 33
 clients 248
 cluster configuration 274
 cluster nodes 264
 cluster volume 269
 data flow 33
 data movement 35
 database applications 36
 file fragmentation 36
 file opening overheads 36
 file sharing 35, 293, 294
 hardware costs 36
 hardware flexibility 35
 heterogeneous environments 35
 high availability 260
 I/O 34
 installing on UNIX 257
 installing on Windows NT/2000 249
 managed buses 241
 MDC installation 253
 mount points 259
 MSCS 260
 MSCS component 266
 MSCS module 262
 overview 236
 patches 253
 performance tester 246, 256
 SANergy host installation 257
 SANergy volume 269
 setting MDC for disks 245
 setting volume ownership 242
 supported operating system 237
 UNIX startup scripts 260
 volume serial number 264
 with a process 260
 with MSCS 262
 with TSM 37
SANergycshsetup 260
SANergyshsetup 260
scalability 19
SCSI reserve/release 262
security management 163
sharing SAN storage 148
Shark. see ESS
SNIA 39
SNMP 254
Solaris clients 159

- Storage Area Networks 22
- Storage Wide Area Networks 30
- StorWatch 31, 119
- Swing 74
- switched fabric 25
- system board 67

T

- tape library 316
- TCP/IP
 - addressing 9
 - application layer 11
 - device driver and hardware layer 8
 - Internet Protocol layer 8
 - IP addressing 9
 - IP connectionless service 8
 - packet 9
 - protocol suites 11
 - TCP layer 10
 - time to live 10
- thin server 16
- Tivoli Storage Manager. *see* TSM
- topology
 - bus 5
 - ring 5
 - star 6
- total cost of ownership 22, 29
- TSM 277
 - agent configuration 305
 - agent installation 304
 - client node configuration 303
 - client scheduling 303
 - device configuration 302
 - device drivers 299
 - integration NAS 300G 290
 - licenses 298
 - server configuration 300
 - server install 297
 - using SANergy 295
 - with SANergy 293
 - wizard 300

U

- UNIX
 - fusing disk volumes 259
 - installing SANergy host 257
 - password synchronization 165
 - processes with SANergy 260

- SANergy startup scripts 260
- UNIX clients 154, 231
- user management 163

V

- vfstab 159
- virtual copy 278
- virtualization solutions 31

W

- Windows
 - fusing disk volumes 256
 - NT 4 Domain 163
 - preventing drive letter assignment 251
 - terminal service 280
 - Workgroups 163
- Windows 2000 for NAS user interface 287
- Windows clients 149, 157
- Windows Powered Server Appliance 133
- World Wide Name 77
- WWN 77
 - finding 77

X

- XML 74
- XRC 56

Z

- zoning 65, 78



Redbooks

Implementing the IBM TotalStorage NAS 300G

(0.5" spine)

0.475" <-> 0.875"

250 <-> 459 pages



Implementing the IBM TotalStorage NAS 300G

High Speed Cross Platform Storage and Tivoli SANergy!



Redbooks

Share data seamlessly between UNIX and Windows environments

Get the best of both NAS and SAN using this hands-on guide

Learn how NAS can meet your business needs

This IBM Redbook describes how to install and configure the very latest IBM storage solution and concept, the IBM TotalStorage Network Attached Storage 300G, in heterogeneous environments.

The 300G series is an innovative Network Attached Storage (NAS) appliance that connects clients and servers on an IP network to Fibre Channel storage, efficiently bridging the gap between LAN storage needs and SAN storage capacities. The NAS 300G is a storage solution for Linux/UNIX and Windows NT/2000 environments. In this book, we show you how to integrate the NAS 300G and explain how it can benefit your company's business needs.

This book is an easy-to-follow guide which describes the market segment that the 300G is aimed at, and explains NAS installation, ease-of-use, remote management, expansion capabilities, high availability (clustering), and backup and recovery techniques. It also explains cross platform storage concepts and methodologies for common data sharing for Linux/UNIX and Windows NT/2000 environments.

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

SG24-6278-00

ISBN 0738423084