

Veritas Storage Foundation™ and High Availability Solutions 6.0

Solutions Guide

HP-UX

Veritas Storage Foundation and High Availability Solutions Guide

The software described in this book is furnished under a license agreement and may be used only in accordance with the terms of the agreement.

Product version: 6.0

Document version: 6.0.4

Legal Notice

Copyright © 2012 Symantec Corporation. All rights reserved.

Symantec, the Symantec logo, Veritas, Veritas Storage Foundation, CommandCentral, NetBackup, Enterprise Vault, and LiveUpdate are trademarks or registered trademarks of Symantec corporation or its affiliates in the U.S. and other countries. Other names may be trademarks of their respective owners.

The product described in this document is distributed under licenses restricting its use, copying, distribution, and decompilation/reverse engineering. No part of this document may be reproduced in any form by any means without prior written authorization of Symantec Corporation and its licensors, if any.

THE DOCUMENTATION IS PROVIDED "AS IS" AND ALL EXPRESS OR IMPLIED CONDITIONS, REPRESENTATIONS AND WARRANTIES, INCLUDING ANY IMPLIED WARRANTY OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE OR NON-INFRINGEMENT, ARE DISCLAIMED, EXCEPT TO THE EXTENT THAT SUCH DISCLAIMERS ARE HELD TO BE LEGALLY INVALID. SYMANTEC CORPORATION SHALL NOT BE LIABLE FOR INCIDENTAL OR CONSEQUENTIAL DAMAGES IN CONNECTION WITH THE FURNISHING, PERFORMANCE, OR USE OF THIS DOCUMENTATION. THE INFORMATION CONTAINED IN THIS DOCUMENTATION IS SUBJECT TO CHANGE WITHOUT NOTICE.

The Licensed Software and Documentation are deemed to be commercial computer software as defined in FAR 12.212 and subject to restricted rights as defined in FAR Section 52.227-19 "Commercial Computer Software - Restricted Rights" and DFARS 227.7202, "Rights in Commercial Computer Software or Commercial Computer Software Documentation", as applicable, and any successor regulations. Any use, modification, reproduction release, performance, display or disclosure of the Licensed Software and Documentation by the U.S. Government shall be solely in accordance with the terms of this Agreement.

Symantec Corporation
350 Ellis Street
Mountain View, CA 94043
<http://www.symantec.com>

Technical Support

Symantec Technical Support maintains support centers globally. Technical Support's primary role is to respond to specific queries about product features and functionality. The Technical Support group also creates content for our online Knowledge Base. The Technical Support group works collaboratively with the other functional areas within Symantec to answer your questions in a timely fashion. For example, the Technical Support group works with Product Engineering and Symantec Security Response to provide alerting services and virus definition updates.

Symantec's support offerings include the following:

- A range of support options that give you the flexibility to select the right amount of service for any size organization
- Telephone and/or Web-based support that provides rapid response and up-to-the-minute information
- Upgrade assurance that delivers software upgrades
- Global support purchased on a regional business hours or 24 hours a day, 7 days a week basis
- Premium service offerings that include Account Management Services

For information about Symantec's support offerings, you can visit our Web site at the following URL:

www.symantec.com/business/support/index.jsp

All support services will be delivered in accordance with your support agreement and the then-current enterprise technical support policy.

Contacting Technical Support

Customers with a current support agreement may access Technical Support information at the following URL:

www.symantec.com/business/support/contact_techsupp_static.jsp

Before contacting Technical Support, make sure you have satisfied the system requirements that are listed in your product documentation. Also, you should be at the computer on which the problem occurred, in case it is necessary to replicate the problem.

When you contact Technical Support, please have the following information available:

- Product release level

- Hardware information
- Available memory, disk space, and NIC information
- Operating system
- Version and patch level
- Network topology
- Router, gateway, and IP address information
- Problem description:
 - Error messages and log files
 - Troubleshooting that was performed before contacting Symantec
 - Recent software configuration changes and network changes

Licensing and registration

If your Symantec product requires registration or a license key, access our technical support Web page at the following URL:

www.symantec.com/business/support/

Customer service

Customer service information is available at the following URL:

www.symantec.com/business/support/

Customer Service is available to assist with non-technical questions, such as the following types of issues:

- Questions regarding product licensing or serialization
- Product registration updates, such as address or name changes
- General product information (features, language availability, local dealers)
- Latest information about product updates and upgrades
- Information about upgrade assurance and support contracts
- Information about the Symantec Buying Programs
- Advice about Symantec's technical support options
- Nontechnical presales questions
- Issues that are related to CD-ROMs or manuals

Support agreement resources

If you want to contact Symantec regarding an existing support agreement, please contact the support agreement administration team for your region as follows:

Asia-Pacific and Japan customercare_apac@symantec.com

Europe, Middle-East, and Africa semea@symantec.com

North America and Latin America supportsolutions@symantec.com

Documentation

Your feedback on product documentation is important to us. Send suggestions for improvements and reports on errors or omissions. Include the title and document version (located on the second page), and chapter and section titles of the text on which you are reporting. Send feedback to:

doc_feedback@symantec.com

About Symantec Connect

Symantec Connect is the peer-to-peer technical community site for Symantec's enterprise customers. Participants can connect and share information with other product users, including creating forum posts, articles, videos, downloads, blogs and suggesting ideas, as well as interact with Symantec product teams and Technical Support. Content is rated by the community, and members receive reward points for their contributions.

<http://www.symantec.com/connect/storage-management>

Displaying Quick I/O status and file attributes	41
Extending a Quick I/O file in a DB2 or Sybase environment	43
Monitoring tablespace free space with DB2 and extending tablespace containers	45
Disabling Quick I/O	48

Chapter 4

Improving database performance with Veritas Cached Quick I/O	49
About Cached Quick I/O	49
How Cached Quick I/O works in a DB2 environment	49
How Cached Quick I/O works in a Sybase environment	51
How Cached Quick I/O improves database performance	52
Tasks for setting up Cached Quick I/O	53
Enabling Cached Quick I/O on a file system	53
Enabling and disabling the qio_cache_enable flag	53
Making Cached Quick I/O settings persistent across reboots and mounts	54
Using vxtunefs to obtain tuning information	55
Determining candidates for Cached Quick I/O	57
Collecting I/O statistics	57
About I/O statistics for DB2	58
About I/O statistics for Sybase	59
Effects of read-aheads on I/O statistics	60
Other tools for analysis	60
Enabling and disabling Cached Quick I/O for individual files	61
Setting cache advisories for individual files	61
Making individual file settings for Cached Quick I/O persistent	62
Determining individual file settings for Cached Quick I/O using qioadmin	63

Chapter 5

Improving database performance with Veritas Concurrent I/O	65
About Concurrent I/O	65
How Concurrent I/O works	65
Tasks for enabling and disabling Concurrent I/O	66
Enabling Concurrent I/O for DB2	66
Disabling Concurrent I/O for DB2	68
Enabling Concurrent I/O for Sybase	69
Disabling Concurrent I/O for Sybase	70

Section 3	Using point-in-time copies	71
Chapter 6	Understanding point-in-time copy methods	73
	About point-in-time copies	73
	Implementing point-in time copy solutions on a primary host	74
	Implementing off-host point-in-time copy solutions	75
	When to use point-in-time copies	81
	About Storage Foundation point-in-time copy technologies	82
	Volume-level snapshots	83
	Storage Checkpoints	84
	About FileSnaps	85
Chapter 7	Backing up and recovering	87
	About Storage Foundation and High Availability Solutions backup and recovery use cases	87
	Creating and maintaining a full image snapshot and incremental point-in-time copies	88
	Setting up a full image snapshot and incremental point-in-time copies	88
	Refreshing point-in-time copies	91
	Recovering from logical corruption	91
	Off-host processing using refreshed snapshot images	93
	Online database backups	93
	Making a backup of an online database on the same host	94
	Making an off-host backup of an online database	105
	Backing up on an off-host cluster file system	116
	Mounting a file system for shared access	118
	Preparing a snapshot of a mounted file system with shared access	118
	Backing up a snapshot of a mounted file system with shared access	120
	Resynchronizing a volume from its snapshot volume	123
	Reattaching snapshot plexes	124
	Database recovery using Storage Checkpoints	125
	Creating Storage Checkpoints	126
	Rolling back a database	127

Chapter 8	Backing up and recovering in a NetBackup environment	129
	About Veritas NetBackup	129
	About using Veritas NetBackup for backup and restore for DB2	130
	About using NetBackup for backup and restore for Sybase	131
	About using Veritas NetBackup to backup and restore Quick I/O files for DB2	132
	About using Veritas NetBackup to backup and restore Quick I/O files for Sybase	133
	Using NetBackup in an SFHA Solutions product environment	134
	Clustering a NetBackup Master Server	134
	Backing up and recovering a VxVM volume using NetBackup	135
	Recovering a VxVM volume using NetBackup	137
Chapter 9	Off-host processing	139
	Storage Foundation off-host processing methods	139
	Using a replica database for decision support	140
	Creating a replica database on the same host	141
	Creating an off-host replica database	151
	What is off-host processing?	162
	About using VVR for off-host processing	162
Chapter 10	Creating and refreshing test environments	163
	About test environments	163
	Creating a test environment	163
	Refreshing a test environment	164
Chapter 11	Creating point-in-time copies of files	167
	Using FileSnaps to create point-in-time copies of files	167
	Using FileSnaps to provision virtual desktops	167
	Using FileSnaps to optimize write intensive applications for virtual machines	168
	Using FileSnaps to create multiple copies of data instantly	168

Section 4	Maximizing storage utilization	171
Chapter 12	Optimizing storage tiering with SmartTier	173
	About SmartTier	173
	SmartTier building blocks	174
	About VxFS multi-volume file systems	175
	About VxVM volume sets	175
	About volume tags	176
	SmartTier use cases for DB2 or Sybase	176
	Setting up a filesystem for storage tiering with SmartTier	176
	Relocating old archive logs to tier two storage using SmartTier	179
	Relocating inactive tablespaces or segments to tier two storage	181
	Relocating active indexes to premium storage	184
	Relocating all indexes to premium storage	186
Section 5	Migrating data	191
Chapter 13	Understanding data migration	193
	Types of data migration	193
Chapter 14	Offline migration from Logical Volume Manager to Veritas Volume Manager	195
	About migrating Logical Volume Manager to VxVM	195
	About LVM to VxVM conversion	196
	Converting unused LVM physical volumes to VxVM disks	196
	Removing LVM disk information	196
	Initializing disks for VxVM use	197
	Converting LVM volume groups to VxVM disk groups	198
	Volume group conversion limitations	198
	Conversion process summary	200
	Identifying LVM disks and volume groups for conversion	201
	Analyzing an LVM volume group to see if conversion is possible	202
	Taking actions to make conversion possible if analysis fails	202
	Backing up your LVM configuration and user data	203
	Planning for new VxVM logical volume names	205
	Stopping application access to volumes in the volume group to be converted	206
	Converting a volume group	207
	Taking actions if conversion fails	208

Implementing changes for new VxVM logical volume names	208
Restarting applications on the new VxVM volumes	208
Tailoring your VxVM configuration	209
Restoring the LVM volume group configuration	209
Rollback to LVM using vxvmconvert	211
Full LVM restoration	212
Examples	213
Example: displaying the vxvmconvert menu	213
Example: listing disk information	214
Example: listing LVM volume group information	214
Example: analyzing LVM volume groups	216
Example: converting LVM volume groups to VxVM disk groups	219
Example: converting LVM version 2 volume groups to VxVM disk groups	224
Example: list, listvg, and vxprint outputs of an LVM volume group before and after conversion	226
Example: VxVM to LVM rollback	228
General information regarding conversion speed	229
Non-interactive conversion of volume groups	230
Analyzing volume groups for conversion	230
Converting volume groups to disk groups	231
Converting disk groups back to volume groups	232

Chapter 15

Online migration of native LVM volumes to VxVM volumes	233
About online migration from Logical Volume Manager (LVM) volumes to Veritas Volume Manager (VxVM) volumes	233
Online migration from LVM volumes in standalone environment to VxVM volumes	235
Administrative interface for online migration from LVM in standalone environment to VxVM	236
Preparing for online migration from LVM in standalone environment to VxVM	238
Migrating from LVM in standalone environment to VxVM	239
Reconfiguring the application to use VxVM volume device path	242
Backing out online migration of LVM in standalone environment to VxVM	242
Do's and Don'ts for online migration from LVM in standalone environment to VxVM	243

	Scenarios not supported for migration from LVM in standalone environment to VxVM	244
	Online migration from LVM volumes in VCS HA environment to VxVM volumes	245
	About online migration from LVM in VCS HA environment to VxVM	246
	Administrative interface for online migration from LVM in VCS HA environment to VxVM	247
	Preparing for online migration from LVM in VCS HA environment to VxVM	248
	Migrating from LVM in VCS HA environment to VxVM	249
	Migrating configurations with multiple volume groups	251
	Backing out online migration of LVM in VCS HA environment to VxVM	252
	Do's and Don'ts for online migration from LVM in VCS HA environment to VxVM	253
	Scenarios not supported for migration from LVM VCS HA environment to VxVM	254
Chapter 16	Migrating storage arrays	255
	Array migration for storage using Linux	255
	Overview of storage mirroring for migration	256
	Allocating new storage	257
	Initializing the new disk	259
	Checking the current VxVM information	260
	Adding a new disk to the disk group	261
	Mirroring	262
	Monitoring	263
	Mirror completion	264
	Removing old storage	264
	Post-mirroring steps	265
Chapter 17	Migrating data between platforms	267
	Overview of the Cross-Platform Data Sharing (CDS) feature	267
	Shared data across platforms	268
	Disk drive sector size	269
	Block size issues	269
	Operating system data	269
	CDS disk format and disk groups	269
	CDS disk access and format	270
	Non-CDS disk groups	273
	Disk group alignment	273

Setting up your system to use Cross-platform Data Sharing	
(CDS)	275
Creating CDS disks from uninitialized disks	276
Creating CDS disks from initialized VxVM disks	277
Creating CDS disk groups	278
Converting non-CDS disks to CDS disks	279
Converting a non-CDS disk group to a CDS disk group	280
Verifying licensing	282
Defaults files	282
Maintaining your system	284
Disk tasks	285
Disk group tasks	287
Displaying information	293
Default activation mode of shared disk groups	296
Additional considerations when importing CDS disk groups	296
File system considerations	297
Considerations about data in the file system	298
File system migration	298
Specifying the migration target	299
Using the fscdsadm command	300
Migrating a file system one time	303
Migrating a file system on an ongoing basis	303
When to convert a file system	305
Converting the byte order of a file system	305
Alignment value and block size	309
Migrating a snapshot volume	310
Index	313

Solutions for Veritas Storage Foundation and High Availability products

- [Chapter 1. Solutions for Veritas Storage Foundation and High Availability products](#)

Solutions for Veritas Storage Foundation and High Availability products

This chapter includes the following topics:

- [About Storage Foundation and High Availability Solutions](#)
- [Feature support across Veritas Storage Foundation and High Availability Solutions 6.0 products](#)
- [Using SmartMove and Thin Provisioning with DB2 and Sybase](#)
- [Finding Veritas Storage Foundation and High Availability Solutions management feature information](#)

About Storage Foundation and High Availability Solutions

Veritas Storage Foundation and High Availability (SFHA) Solutions product components and features can be used individually and in concert to improve performance, resilience, and ease of management for your storage and applications. This guide documents key use cases for the management features of SFHA Solutions products:

Table 1-1 Key use cases for SFHA Solutions products

Use case	SFHA Solutions feature
<p>Improve DB2 and Sybase database performance using SFHA Solutions database accelerators to enable your database to achieve the speed of raw disk while retaining the management features and convenience of a file system.</p>	<p>Concurrent I/O</p>
<p>Protect your data using SFHA Solutions Flashsnap, Storage Checkpoints, and NetBackup point-in-time copy methods to back up and recover your data.</p>	<p>FlashSnap Storage Checkpoints NetBackup with SFHA Solutions</p>
<p>Process your data off-host to avoid performance loss to your production hosts by using SFHA Solutions volume snapshots.</p>	<p>FlashSnap</p>
<p>Optimize copies of your production database for test, decision modeling, and development purposes by using SFHA Solutions point-in-time copy methods.</p>	<p>FlashSnap</p>
<p>Make file level point-in-time snapshots using SFHA Solutions space-optimized FileSnap when you need finer granularity for your point-in-time copies than file systems or volumes. You can use FileSnap for cloning virtual machines.</p>	<p>FileSnap</p>
<p>Maximize your storage utilization using SFHA Solutions SmartTier to move data to storage tiers based on age, priority, and access rate criteria.</p>	<p>SmartTier</p>
<p>Convert your data from native OS to VxFS using SFHA Solutions Portable Data Containers.</p>	<p>Offline conversion utility Online migration utility</p>
<p>Convert your data from raw disk to VxFS: use SFHA Solutions Portable Data Containers.</p>	<p>Offline conversion utility</p>
<p>Migrate your data from one platform to another (server migration) using SFHA Solutions Portable Data Containers.</p>	<p>Portable Data Containers</p>

Table 1-1 Key use cases for SFHA Solutions products (*continued*)

Use case	SFHA Solutions feature
Migrate your data across arrays using SFHA Solutions Portable Data Containers.	Volume mirroring

Feature support across Veritas Storage Foundation and High Availability Solutions 6.0 products

Storage solutions and use cases are based on the shared management features of Veritas Storage Foundation and High Availability (SFHA) Solutions products. Clustering features are available separately through Veritas Cluster Server (VCS) as well as through the SFHA Solutions products.

[Table 1-2](#) lists the features supported across SFHA Solutions products. [Table 1-3](#) lists the high availability and disaster recovery features available in VCS.

Table 1-2 Storage management features in SFHA Solutions products

Storage management feature	SF Basic	SF Std.	SF Ent.	SF Std. HA	SF Ent. HA	SFCFS HA	SFRAC	SVS	SF Syb CE
Veritas Extension for Oracle Disk Manager	Y	Y	Y	Y	Y	Y	Y	Y	N
Veritas Extension for Cached Oracle Disk Manager	Y	Y	Y	Y	Y	Y	N	Y	Y
Quick I/O	Y	Y	Y	Y	Y	Y	Y	Y	Y
Cached Quick I/O	Y	Y	Y	Y	Y	Y	Y	Y	Y
Concurrent I/O	Y	Y	Y	Y	Y	Y	Y	Y	Y
Compression	N	Y	Y	Y	Y	Y	Y	Y	Y
SmartMove	Y	Y	Y	Y	Y	Y	Y	Y	Y
SmartTier	N	Y	Y	Y	Y	Y	Y	Y	Y
SmartTier for Oracle	N	N	Y	Y	Y	Y	Y	N	N
Thin Reclamation	N	N	Y	Y	Y	Y	Y	Y	Y
Portable Data Containers	N	N	Y	Y	Y	Y	Y	Y	Y
FlashSnap	N	N	Y	Y	Y	Y	Y	Y	Y

Table 1-2 Storage management features in SFHA Solutions products
(continued)

Storage management feature	SF Basic	SF Std.	SF Ent.	SF Std. HA	SF Ent. HA	SFCFS HA	SFRAC	SVS	SF Syb CE
Database FlashSnap	N	N	Y	Y	Y	Y	Y	N	N
Storage Checkpoints	N	N	Y	Y	Y	Y	Y	Y	Y
Database Storage Checkpoints	N	N	Y	Y	Y	Y	Y	N	N
FileSnap	N	N	N	Y	Y	Y	Y	Y	Y
Volume replication	O	O	O	O	O	O	O	O	O
File replication	O	O	O	O	O	O	O	Y	O
Advanced support for virtual storage	Y	Y	Y	Y	Y	Y	N	N	N
Symantec Storage Plug-in for VMWare vCenter	Y	Y	Y	Y	Y	Y	Y	Y	Y
Clustering features for high availability (HA)	N	N	N	Y	Y	Y	Y	Y	Y
Disaster recovery features (HA/DR)	N	N	N	O	O	O	O	N	O

Table 1-3 Availability management features in SFHA Solutions products

Availability management feature	VCS	VCS HA/DR
Clustering for high availability (HA)	Y	Y
Database and application/ISV agents	Y	Y
Advanced failover logic	Y	Y
Data integrity protection with I/O fencing	Y	Y
Advanced virtual machines support	Y	Y
Virtual Business Services	Y	Y
Replication agents	N	Y
Replicated Data Cluster	N	Y
Stretch cluster	N	Y
Campus cluster	N	Y

Table 1-3 Availability management features in SFHA Solutions products
(continued)

Availability management feature	VCS	VCS HA/DR
Global clustering (GCO)	N	Y
Fire Drill	N	Y

Legend:

- Y=Feature is included in the product license.
- O=Feature is not included in the product license but may be licensed separately.
- N=Feature is not supported with the product license.

Notes:

- The Veritas File Replicator license includes file replication.
- The Veritas Replicator license includes both file replication and volume replication (also known as Veritas Volume Replicator, VVR).
- SmartTier is an expanded and renamed version of Dynamic Storage Tiering (DST).
- Symantec VirtualStore (SVS) is available only for the Linux and Solaris operating systems.
- All features listed in [Table 1-2](#) and [Table 1-3](#) are supported on HP-UX. Consult specific product documentation for information on supported operating systems.

Using SmartMove and Thin Provisioning with DB2 and Sybase

You can use SmartMove and Thin Provisioning with Veritas Storage Foundation and High Availability products and your DB2 or Sybase database.

When data files are deleted, you can reclaim the storage space used by these files if the underlying devices are thin reclaimable LUNs.

For information about the Storage Foundation Thin Reclamation feature, see the *Veritas Storage Foundation Administrator's Guide*.

Finding Veritas Storage Foundation and High Availability Solutions management feature information

The following Veritas Storage Foundation and High Availability Solutions management features are illustrated with use case examples in this guide:

- Improving database performance for DB2 and Sybase
- Backing up and recovering your data
- Processing data off-host
- Optimizing test and development environments
- Maximizing storage utilization
- Converting your data from native OS to VxFS
- Converting your data from raw disk to VxFS
- Migrating your data from one platform to another (server migration)
- Migrating your data across arrays

For Veritas Storage Foundations and High Availability Solutions management features concept and administrative information, see the following guides:

- *Veritas Storage Foundation™ Administrator's Guide*
- *Veritas Storage Foundation™ for Cluster File System High Availability Administrator's Guide.*
- *Veritas Storage Foundation™ for Oracle RAC Administrator's Guide.*
- *Veritas Storage Foundation™ for Sybase ASE CE Administrator's Guide.*

For Information on using Veritas Storage Foundations and High Availability Solutions management features with Oracle databases, see: *Veritas Storage Foundation™: Storage and Availability Management for Oracle Databases.*

For Information on using Veritas Storage Foundations and High Availability Solutions replication features, see: *Veritas Storage Foundation and High Availability Solutions Replication Guide*

Improving database performance

- [Chapter 2. Overview of database accelerators](#)
- [Chapter 3. Improving database performance with Veritas Quick I/O](#)
- [Chapter 4. Improving database performance with Veritas Cached Quick I/O](#)
- [Chapter 5. Improving database performance with Veritas Concurrent I/O](#)

Overview of database accelerators

This chapter includes the following topics:

- [About Storage Foundation and High Availability Solutions products database accelerators](#)

About Storage Foundation and High Availability Solutions products database accelerators

The major concern in any environment is maintaining respectable performance or meeting performance service level agreements (SLAs). Veritas Storage Foundation products improve the overall performance of database environments in a variety of ways.

- Quick I/O (QIO) and cached Quick I/O (CQIO) optimized for all database environments
- Concurrent I/O (CIO) optimized for DB2 and Sybase environments
- Oracle Disk Manager (ODM) and Cached Oracle Disk Manager (CODM) optimized specifically for Oracle environments

These database accelerator technologies enable database performance equal to raw disk partitions, but with the manageability benefits of a file system. With the Dynamic Multi-pathing (DMP) feature of Storage Foundation, performance is maximized by load-balancing I/O activity across all available paths from server to array. DMP supports all major hardware RAID vendors, hence there is no need for third-party multi-pathing software, reducing the total cost of ownership.

Storage Foundation database accelerators enable you to manage performance for your database with more precision.

- To achieve raw device performance for databases run on VxFS file systems, use Veritas Quick I/O.
- To further enhance database performance by leveraging large system memory to selectively buffer the frequently accessed data, use Veritas Cached Quick I/O.
- To achieve improved performance for DB2 or Sybase databases run on VxFS file systems, without restrictions on increasing file size, use Veritas Concurrent I/O.
- To improve Oracle performance and manage system bandwidth through an improved Application Programming Interface (API) that contains advanced kernel support for file I/O, use Veritas Oracle Disk Manager (ODM).
- To enable selected I/O to use caching to improve ODM I/O performance, use Veritas Extension for Cached Oracle Disk Manager (Cached ODM).

Improving database performance with Veritas Quick I/O

This chapter includes the following topics:

- [About Quick I/O](#)
- [How to set up Quick I/O](#)
- [Creating DB2 database containers or Sybase files as Quick I/O files using qiomkfile](#)
- [Preallocating space for Quick I/O files using the setext command](#)
- [Accessing regular VxFS files as Quick I/O files](#)
- [About sparse files](#)
- [Displaying Quick I/O status and file attributes](#)
- [Extending a Quick I/O file in a DB2 or Sybase environment](#)
- [Monitoring tablespace free space with DB2 and extending tablespace containers](#)
- [Disabling Quick I/O](#)

About Quick I/O

Veritas Quick I/O is a VxFS feature included in Veritas Storage Foundation Standard and Enterprise products that enables applications access preallocated VxFS files as raw character devices. Quick I/O provides the administrative benefits

of running databases on file systems without the typically associated degradation in performance.

How Quick I/O works

Veritas Quick I/O supports direct I/O and kernel asynchronous I/O and enables dataserverters to access regular files on a VxFS file system as raw character devices.

The benefits of using Quick I/O are:

- Improved performance and processing throughput by having Quick I/O files act as raw devices.
- Ability to manage Quick I/O files as regular files, which simplifies administrative tasks such as allocating, moving, copying, resizing, and backing up DB2 containers.
- Ability to manage Quick I/O files as regular files, which simplifies administrative tasks such as allocating, moving, copying, resizing, and backing up Sybase dataserverters.

How Quick I/O improves database performance

The benefits of using Quick I/O are:

- Improved performance and processing throughput by having Quick I/O files act as raw devices.
- Ability to manage Quick I/O files as regular files, which simplifies administrative tasks such as allocating, moving, copying, resizing, and backing up DB2 containers.
- Ability to manage Quick I/O files as regular files, which simplifies administrative tasks such as allocating, moving, copying, resizing, and backing up Sybase dataserverters.

Quick I/O's ability to access regular files as raw devices improves database performance by:

Table 3-1

Quick I/O feature	Advantage
Supporting direct I/O	I/O on files using <code>read()</code> and <code>write()</code> system calls typically results in data being copied twice: once between user and kernel space, and later between kernel space and disk. In contrast, I/O on raw devices is direct. That is, data is copied directly between user space and disk, saving one level of copying. As with I/O on raw devices, Quick I/O avoids extra copying.
Avoiding kernel write locks on database files	When database I/O is performed using the <code>write()</code> system call, each system call acquires and releases a write lock inside the kernel. This lock prevents multiple simultaneous write operations on the same file. Because database systems usually implement their own locking to manage concurrent access to files, per file writer locks unnecessarily serialize I/O operations. Quick I/O bypasses file system per file locking and lets the database server control data access.
Avoiding double buffering	Most database servers maintain their own buffer cache and do not need the file system buffer cache. Database data cached in the file system buffer is therefore redundant and results in wasted memory and extra system CPU utilization to manage the buffer. By supporting direct I/O, Quick I/O eliminates double buffering. Data is copied directly between the relational database management system (RDBMS) cache and disk, which lowers CPU utilization and frees up memory that can then be used by the database server buffer cache to further improve transaction processing throughput.

Table 3-1 (continued)

Quick I/O feature	Advantage
Supporting AIX Fastpath asynchronous I/O	AIX Fastpath asynchronous I/O is a form of I/O that performs non-blocking system level reads and writes, allowing the system to handle multiple I/O requests simultaneously. Operating systems such as AIX provide support for asynchronous I/O on raw devices, but not on regular files. As a result, even if the database server is capable of using asynchronous I/O, it cannot issue asynchronous I/O requests when the database runs on file systems. Lack of asynchronous I/O significantly degrades performance. Quick I/O lets the database server take advantage of kernel-supported asynchronous I/O on file system files accessed using the Quick I/O interface.
Supporting asynchronous I/O	HP-UX asynchronous I/O is a form of I/O that performs non-blocking system level reads and writes, allowing the system to handle multiple I/O requests simultaneously. Operating systems such as HP-UX provide kernel support for asynchronous I/O on raw devices, but not on regular files. As a result, even if the database server is capable of using asynchronous I/O, it cannot issue asynchronous I/O requests when the database runs on file systems. Lack of asynchronous I/O significantly degrades performance. Quick I/O enables the database server to take advantage of kernel-supported asynchronous I/O on file system files accessed using the Quick I/O interface.

Table 3-1 (continued)

Quick I/O feature	Advantage
Supporting kernel asynchronous I/O	Solaris kernel asynchronous I/O is a form of I/O that performs non-blocking system level reads and writes, allowing the system to handle multiple I/O requests simultaneously. Operating systems such as Solaris provide kernel support for asynchronous I/O on raw devices, but not on regular files. As a result, even if the database server is capable of using asynchronous I/O, it cannot issue asynchronous I/O requests when the database runs on file systems. Lack of asynchronous I/O significantly degrades performance. Quick I/O enables the database server to take advantage of kernel-supported asynchronous I/O on file system files accessed using the Quick I/O interface.

How to set up Quick I/O

Quick I/O is included in the VxFS package shipped with Veritas Storage Foundation Standard and Enterprise products. By default, Quick I/O is enabled when you mount a VxFS file system.

If Quick I/O is not available in the kernel, or a Veritas Storage Foundation Standard or Enterprise product license is not installed, a file system mounts without Quick I/O by default, the Quick I/O file name is treated as a regular file, and no error message is displayed. If, however, you specify the `-o qio` option, the `mount` command prints the following error message and terminates without mounting the file system.

```
VxFDD: You don't have a license to run this program
vxfs mount: Quick I/O not available
```

Depending on whether you are creating a new database or are converting an existing database to use Quick I/O, you have the following options:

If you are creating a new database to use Quick I/O:

- You can use the `qiomkfile` command to preallocate space for database files and make them accessible to the Quick I/O interface.
- You can use the `setext` command to preallocate space for database files and create the Quick I/O files.

If you are converting an existing database:

- You can create symbolic links for existing VxFS files, and use these symbolic links to access the files as Quick I/O files.

Creating DB2 database containers or Sybase files as Quick I/O files using `qiomkfile`

The best way to preallocate space for tablespace containers and to make them accessible using the Quick I/O interface is to use the `qiomkfile`. You can use the `qiomkfile` to create the Quick I/O files for either temporary or permanent tablespaces.

For DB2, you can create Database Managed Space (DMS) containers with the type 'DEVICE' using Quick I/O.

- | | |
|---------------|---|
| Prerequisites | <ul style="list-style-type: none">■ You can create Quick I/O files only on VxFS file systems.■ If you are creating containers on an existing file system, run <code>fsadm</code> (or similar utility) to report and eliminate fragmentation.■ You must have read/write permissions on the directory in which you intend to create database Quick I/O files. |
| Usage notes | <ul style="list-style-type: none">■ The <code>qiomkfile</code> command creates two files: a regular file with preallocated, contiguous space, and a file that is a symbolic link pointing to the Quick I/O name extension.■ See the <code>qiomkfile(1M)</code> manual page for more information. |
| -a | Creates a symbolic link with an absolute path name for a specified file. Use the <code>-a</code> option when absolute path names are required. However, the default is to create a symbolic link with a relative path name. |
| -e | Extends a file by a specified amount to allow tablespace resizing.
See “Extending a Quick I/O file in a DB2 or Sybase environment” on page 43. |
| -r | Increases the file to a specified size to allow DB tablespace resizing.
See “Extending a Quick I/O file in a DB2 or Sybase environment” on page 43. |

-s Specifies the space to preallocate for a file in bytes, kilobytes, megabytes, gigabytes, or sectors (1024 bytes) by adding a k, K, m, M, g, G, s, or S suffix. The default is bytes—you do not need to attach a suffix to specify the value in bytes. The size of the file that is preallocated is the total size of the file (including the header) rounded to the nearest multiple of the file system block size.

Warning: Exercise caution when using absolute path names. Extra steps may be required during database backup and restore procedures to preserve symbolic links. If you restore files to directories different from the original paths, you must change the symbolic links that use absolute path names to point to the new path names before the database is restarted.

To create a DB2 container as a Quick I/O file using `qiomkfile`

- 1 Create a Quick I/O-capable file using the `qiomkfile` command:

```
# /opt/VRTS/bin/qiomkfile -s file_size /mount_point/filename
```

For example, to show how to create a 100MB Quick I/O-capable file named `dbfile` on the VxFS file system `/db01` using a relative path name:

```
# /opt/VRTS/bin/qiomkfile -s 100m /db01/dbfile
# ls -al
-rw-r--r--  1 db2inst1  db2iadml 104857600  Oct 2 13:42  .dbfile
lrwxrwxrwx  1 db2inst1  db2iadml          19  Oct 2 13:42  dbfile -> \
.dbfile::cdev:vxfs:
```

- 2 Create tablespace containers using this file with the following SQL statements:

```
$ db2 connect to database
$ db2 create tablespace tbsname managed by database using \
( DEVICE /mount_point/filename size )
$ db2 terminate
```

In the example from 1, `qiomkfile` creates a regular file named `/db01/dbfile`, which has the real space allocated. Then, `qiomkfile` creates a symbolic link named `/db01/dbfile`. This symbolic link is a relative link to the Quick I/O interface for `/db01/.dbfile`, that is, to the `.dbfile::cdev:vxfs:` file. The symbolic link allows `.dbfile` to be accessed by any database or application using its Quick I/O interface.

We can then add the file to the DB2 database `PROD`:

```
$ db2 connect to PROD
$ db2 create tablespace NEWTBS managed by database using \
( DEVICE '/db01/dbfile' 100m )
$ db2 terminate
```

To create a Sybase database file as a Quick I/O file using qiomkfile

- 1 Create a database file using the `qiomkfile` command:

```
# /opt/VRTS/bin/qiomkfile -s file_size /mount_point/filename
```

For example, to show how to create a 100MB database file named `dbfile` on the VxFS file system `/db01` using a relative path name:

```
# /opt/VRTS/bin/qiomkfile -s 100m /db01/dbfile
$ ls -al
-rw-r--r--  1 sybase  sybase  104857600  Oct 2 13:42  .dbfile
lrwxrwxrwx  1 sybase  sybase           19  Oct 2 13:42  dbfile -> \
.dbfile::cdev:vxfs:
```

- 2 Add a device to the Sybase dataserer device pool for the Quick I/O file using the `disk init` command:

```
$ isql -Usa -Psa_password -Sdataserver_name
> disk init
> name="device_name",
> physname="/mount_point/filename",
> vdevno="device_number",
> size=51200
> go
> alter database production on new_device=file_size
> go
```

The size is in 2K units. The Enterprise Reference manual contains more information on the `disk init` command.

In the example from 1, `qiomkfile` creates a regular file named `/db01/.dbfile`, which has the real space allocated. Then, `qiomkfile` creates a symbolic link named `/db01/dbfile`. The symbolic link is a relative link to the Quick I/O interface for `/db01/.dbfile`, that is, to the `.dbfile::cdev:vxfs: file`. The symbolic link allows `.dbfile` to be accessed by any database or application using its Quick I/O interface.

The device size is a multiple of 2K pages. In the example, 51200 times 2K pages is 104857600 bytes. The `qiomkfile` command must use this size.

An example to show how to add a 100MB Quick I/O file named `dbfile` to the list of devices used by database `production`, using the `disk init` command:

```
$ isql -Usa -Psa_password -Sdataserver_name
> disk init
> name="new_device",
> physname="/db01/dbfile",
> vdevno="device_number",
> size=51200
> go
> alter database production on new_device=100
> go
```

See the *Sybase Adaptive Server Enterprise Reference Manual*.

3 Use the file to create a new segment or add to an existing segment.

To add a new segment:

```
$ isql -Usa -Psa_password -Sdataserver_name
> sp_addsegment new_segment, db_name, device_name
> go
```

To extend a segment:

```
$ isql -Usa -Psa_password -Sdataserver_name
> sp_extendsegment segment_name, db_name, device_name
> go
```

An example to show how to create a new segment, named `segment2`, for device `dbfile` on database `production`:

```
$ isql -Usa_password -Sdataserver_name
> sp_addsegment segment2, production, dbfile
> go
```

An example to show how to extend a segment, named `segment1`, for device `dbfile` on database `production`:

```
$ isql -Usa_password -Sdataserver_name
> sp_extendsegment segment1, production, dbfile
> go
```

See the *Sybase Adaptive Server Enterprise Reference Manual*.

Preallocating space for Quick I/O files using the `setext` command

As an alternative to using the `qiomkfile` command, you can also use the VxFS `setext` command to preallocate space for database files.

Before preallocating space with `setext`, make sure the following conditions have been met:

- | | |
|---------------|---|
| Prerequisites | ■ The <code>setext</code> command requires superuser (<code>root</code>) privileges. |
| Usage notes | ■ You can use the <code>chown</code> command to change the owner and group permissions on the file after you create it.
See the <code>setext</code> (1M) manual page for more information. |

To create a Quick I/O database file using `setext`

- 1 Access the VxFS mount point and create a file:

```
# cd /mount_point  
  
# touch .filename
```

- 2 Use the `setext` command to preallocate space for the file:

```
# /opt/VRTS/bin/setext -r size -f noreserve -f chgsize \  
.filename
```

- 3 Create a symbolic link to allow databases or applications access to the file using its Quick I/O interface:

```
# ln -s .filename::cdev:vxfs: filename
```

4 Change the owner and group permissions on the file.

For example, for DB2:

```
# chmod 660 .dbfile
```

For example, for Sybase:

```
# chown sybase:sybase .filename
```

```
# chmod 660 .filename
```

5 To access the mountpoint for the database:

For example, for */db01*, create a container, preallocate the space, and change the permissions:

```
# cd /db01
# touch .dbfile
# /opt/VRTS/bin/setext -r 100M -f noreserve -f chgsize .dbfile
# ln -s .dbfile::
```

For DB2:

```
# chown db2inst1:db2iadml .dbfile
```

```
# chmod 660 .dbfile
```

For Sybase:

```
# chown sybase:sybase .dbfile
```

```
# chmod 660 .dbfile
```

Accessing regular VxFS files as Quick I/O files

You can access regular VxFS files as Quick I/O files using the `::cdev:vxfs:` name extension.

While symbolic links are recommended because they provide easy file system management and location transparency of database files, the drawback of using symbolic links is that you must manage two sets of files (for instance, during database backup and restore).

- Usage notes
- If possible, use relative path names instead of absolute path names when creating symbolic links to access regular files as Quick I/O files. Using relative path names prevents copies of the symbolic link from referring to the original file when the directory is copied. This is important if you are backing up or moving database files with a command that preserves the symbolic link. However, some applications require absolute path names. If a file is then relocated to another directory, you must change the symbolic link to use the new absolute path. Alternatively, you can put all the symbolic links in a directory separate from the data directories. For example, you can create a directory named `/database` and put all the symbolic links there, with the symbolic links pointing to absolute path names.

To access an existing regular file as a Quick I/O file on a VxFS file system

- 1 Access the VxFS file system mount point containing the regular files:

```
$ cd /mount_point
```

- 2 Create the symbolic link:

```
$ mv filename .filename  
$ ln -s .filename::cdev:vxfs: filename
```

This example shows how to access the VxFS file `dbfile` as a Quick I/O file:

```
$ cd /db01  
$ mv dbfile .dbfile  
$ ln -s .dbfile::cdev:vxfs: dbfile
```

This example shows how to confirm the symbolic link was created:

```
$ ls -lo .dbfile dbfile
```

For DB2:

```
-rw-r--r-- 1 db2inst1 104890368 Oct 2 13:42 .dbfile
```

For Sybase:

```
$ ls -lo .dbfile dbfile  
  
-rw-r--r-- 1 sybase 104890368 Oct 2 13:42 .dbfile  
lrwxrwxrwx 1 sybase 19 Oct 2 13:42 dbfile ->  
.dbfile::cdev:vxfs:
```


About sparse files

Support for sparse files lets applications store information (in inodes) to identify data blocks that have only zeroes, so that only blocks containing non-zero data have to be allocated on disk.

For example, if a file is 10KB, it typically means that there are blocks on disk covering the whole 10KB. Assume that you always want the first 9K to be zeroes. The application can go to an offset of 9KB and write 1KB worth of data. Only a block for the 1KB that was written is allocated, but the size of the file is still 10KB.

The file is now sparse. It has a hole from offset 0 to 9KB. If the application reads any part of the file within this range, it will see a string of zeroes.

If the application subsequently writes a 1KB block to the file from an offset of 4KB, for example, the file system will allocate another block.

The file then looks like:

- 0-4KB - hole
- 4-5KB - data block
- 5-9KB - hole
- 9-10KB - data block

So a 1TB file system can potentially store up to 2TB worth of files if there are sufficient blocks containing zeroes. Quick I/O files cannot be sparse and will always have all blocks specified allocated to them.

Displaying Quick I/O status and file attributes

You can obtain and display information about Quick I/O status and file attributes using various options of the `ls` command:

- | | |
|-------------|--|
| -al | Lists all files on a file system, including Quick I/O files and their links. |
| -lL | Shows if Quick I/O was successfully installed and enabled. |
| -a1L | Shows how a Quick I/O file name is resolved to that of a raw device. |

To list all files on the current file system, including Quick I/O files and their links

- ◆ Use the `ls -al` command with the file names:

```
$ ls -al filename .filename
```

The following example shows how to use the `-a` option to display the absolute path name created using `qiomkfile`:

```
$ ls -al d* .d*
```

For DB2:

```
-rw-r--r--  1 db2inst1 db2iadm1 104890368  Oct 2 13:42  .dbfile  
lrwxrwxrwx  1 db2inst1 db2iadm1   19      Oct 2 13:42  dbfile ->  
  .dbfile::cdev:vxfs:
```

For Sybase:

```
-rw-r--r--  1 sybase  sybase  104890368  Oct 2 13:42  .dbfile  
lrwxrwxrwx  1 sybase  sybase    19      Oct 2 13:42  dbfile ->  
  .dbfile::cdev:vxfs:
```

To determine if a Sybase segment has been converted to Quick I/O

- ◆ Use the `ls` command as follows:

```
$ ls -lL filename
```

The following example shows how to determine if Quick I/O is installed and enabled:

```
$ ls -lL dbfile
```

```
crw-r--r--  1 sybase  dba  45, 1  Oct 2 13:42  dbfile
```

To show a Quick I/O file resolved to a raw device

- ◆ Use the `ls` command with the file names as follows:

```
$ ls -all filename .filename
```

The following example shows how the Quick I/O file name `dbfile` is resolved to that of a raw device:

```
$ ls -all d* .d*
```

For DB2:

```
crw-r--r--  1 db2inst1 db2iadml 45,  1          Oct 2 13:42  dbfile
-rw-r--r--  1 db2inst1 db2iadml 104890368    Oct 2 13:42  .dbfile
```

For Sybase:

```
crw-r--r--  1 sybase  sybase  45,  1          Oct 2 13:42  dbfile
-rw-r--r--  1 sybase  sybase  104890368    Oct 2 13:42  .dbfile
```

Extending a Quick I/O file in a DB2 or Sybase environment

Although Quick I/O files must be preallocated, they are not limited to the preallocated sizes. You can grow or “extend” a Quick I/O file by a specific amount or to a specific size, using options to the `qiomkfile` command. Extending Quick I/O files is a fast, online operation and offers a significant advantage over using raw devices.

Before extending a Quick I/O file, make sure the following conditions have been met:

- Prerequisites
- You must have sufficient space on the file system to extend the Quick I/O file.

- Usage notes
- You can also grow VxFS file systems online (provided the underlying disk or volume can be extended) using the `fsadm` command. You can expand the underlying volume and the filesystem with the `vxresize` command.
 - You must have superuser (`root`) privileges to resize VxFS file systems using the `fsadm` command.
 - For Sybase: although you have the ability to extend a Quick I/O file, you cannot resize a database device in Sybase once it is initialized. However, with the ability to grow the volumes and file systems online, you can easily allocate new database devices to be used for new segments and to extend existing segments.
 - See the `fsadm_vxfs(1M)` and `qiomkfile(1M)` manual pages for more information.

The following options are available with the `qiomkfile` command:

- e Extends the file by a specified amount to allow resizing.
- r Increases the file to a specified size to allow resizing.

To extend a Quick I/O file

- 1 If required, ensure the underlying storage device is large enough to contain a larger VxFS file system (see the `vxassist(1M)` manual page for more information), and resize the VxFS file system using `fsadm` command:

For Sybase, for example:

```
# /opt/VRTS/bin/fsadm -b newsize /mount_point
```

where:

- `-b` is the option for changing size
- `newsize` is the new size of the file system in bytes, kilobytes, megabytes, blocks, or sectors

- *mount_point* is the file system's mount point

2 Extend the Quick I/O file using the `qiomkfile` command:

```
# /opt/VRTS/bin/qiomkfile -e extend_amount /mount_point/filename
```

or

```
# /opt/VRTS/bin/qiomkfile -r newsize /mount_point/filename
```

An example to show how to grow VxFS file system:

/db01 to 500MB and extend the `tbs1_cont001` Quick I/O file by 20MB:

```
# /opt/VRTS/bin/qiomkfile -e 20M /db01/tbs1_cont001
```

```
# /opt/VRTS/bin/fsadm -b 500M /db01
```

An example to show how to grow VxFS file system for DB2:

/db01 to 500MB and resize the `tbs1_cont001` Quick I/O file to 300MB:

```
# /opt/VRTS/bin/fsadm -b 500M /db01
```

```
# /opt/VRTS/bin/qiomkfile -r 300M /db01/tbs1_cont001
```

An example to show how to grow VxFS file system for Sybase:

/db01 to 500MB and resize the `dbfile` Quick I/O file to 300MB:

```
# /opt/VRTS/bin/fsadm -b 500M /db01
```

```
# /opt/VRTS/bin/qiomkfile -r 300M /db01/dbfile
```

Monitoring tablespace free space with DB2 and extending tablespace containers

DB2 does not automatically make use of extended DMS files. When tablespace space needs to be extended, a number of DB2 commands must be run. Unlike raw devices, a Database Administrator can easily extend Quick I/O files online. Using this method, a Database Administrator can monitor the free space available in the DB2 tablespaces and use the `qiomkfile` command to grow the Quick I/O files online as needed (typically when the file is about 80 to 90% full). This method does not require you to lock out unused disk space for Quick I/O files. The free space on the file system is available for use by other applications.

Before extending tablespaces, make sure the following conditions have been met:

- Prerequisites
- Log on as the DB2 instance owner.
- Usage notes
- Monitor the free space available in the Quick I/O file, and grow the file as necessary with the `qiomkfile` command.
 - A Database Administrator can grow underlying VxFS file systems online (provided the underlying disk or volume can be extended) using the `fsadm` command. See the `fsadm (1M)` manual page for more information.
 - It is strongly recommended that if a DB2 tablespace is running out of space, that all containers are grown by the same amount. It is possible to add extra containers to the tablespace, but this results in DB2 performing a relayout of the data to balance usage across all available containers in the tablespace.
Also refer to the DB2 Administration Guide section on *Managing Storage and the DB2 SQL Reference Guide* for information on the `ALTER TABLESPACE` command.

To monitor the free space available in a DB2 tablespace

- ◆ Use the following DB2 commands:

```
$ db2 connect to database
$ db2 list tablespaces show detail
$ db2 terminate
```

To extend a Quick I/O file using `qiomkfile`

- ◆ Use the `qiomkfile` command to extend the Quick I/O file (if the container is running low on free blocks):

```
# /opt/VRTS/bin/qiomkfile -e extend_amount filename
```

To extend a DB2 tablespace by a fixed amount

- ◆ Use the following DB2 commands:

```
$ db2 connect to database
$ db2 alter tablespace tablespace-name extend (ALL amount)
$ db2 terminate
```

This example shows how to monitor the free space on the tablespaces in database `PROD`:

```
$ db2 connect to PROD
$ db2 list tablespaces show detail
$ db2 terminate
```

This example shows how to extend the three DB2 containers owned by tablespace `EMP` by 500MB using the `qiomkfile` command:

```
# /opt/VRTS/bin/qiomkfile -e 500M tbsEMP_cont001
# /opt/VRTS/bin/qiomkfile -e 500M tbsEMP_cont002
# /opt/VRTS/bin/qiomkfile -e 500M tbsEMP_cont003
```

This example shows how to notify DB2 that all containers in tablespace `EMP` have grown by 500MB:

```
$ db2 connect to PROD
$ db2 alter tablespace EMP extend (ALL 500M)
$ db2 terminate
```

This example shows how to verify the newly allocated space on the tablespace `EMP` in database `PROD`:

```
$ db2 connect to PROD
$ db2 list tablespaces show detail
$ db2 terminate
```

Disabling Quick I/O

Before disabling Quick I/O, make sure the following condition has been met:

Prerequisite The file system you are planning to remount must be located in the `/etc/vfstab` file.

If you need to disable the Quick I/O feature, you first need to convert any Quick I/O files back to regular VxFS files. Then, remount the VxFS file system using a special mount option.

To disable Quick I/O for DB2

- 1 If the database is active, make it inactive by either shutting down the instance or disabling user connections.
- 2 To remount the file system with Quick I/O disabled, use the command as follows:

```
# /opt/VRTS/bin/mount -F vxfs -o remount,noqio /mount_point
```

To disable Quick I/O for Sybase

- 1 If the database is running, shut it down.
- 2 To remount the file system with Quick I/O disabled, use the `mount -o noqio` command as follows:

```
# /opt/VRTS/bin/mount -F vxfs -o remount,noqio /mount_point
```


Improving database performance with Veritas Cached Quick I/O

This chapter includes the following topics:

- [About Cached Quick I/O](#)
- [Tasks for setting up Cached Quick I/O](#)
- [Enabling Cached Quick I/O on a file system](#)
- [Determining candidates for Cached Quick I/O](#)
- [Enabling and disabling Cached Quick I/O for individual files](#)

About Cached Quick I/O

Veritas Cached Quick I/O maintains and extends the database performance benefits of Veritas Quick I/O by making more efficient use of large, unused system memory through a selective buffering mechanism. Cached Quick I/O also supports features that support buffering behavior, such as file system read-ahead.

How Cached Quick I/O works in a DB2 environment

Cached Quick I/O is a specialized external caching mechanism suitable for dynamically changing memory allocation available for caching DB2 tablespace data. Cached Quick I/O can be selectively applied to containers that are suffering an undesirable amount of physical disk I/O due to insufficient DB2 BufferPool reservation. Cached Quick I/O works by taking advantage of the available physical memory that is left over after the operating system reserves the amount it needs

and the DB2 BufferPools have been sized to their normal operating capacity. This extra memory serves as a cache to store file data, effectively serving as a second-level cache backing the BufferPool(s).

For example, consider a system configured with 12GB of physical memory, an operating system using 1GB, and a total DB2 BufferPool size of 3.5GB. Unless you have other applications running on your system, the remaining 7.5GB of memory is unused. If you enable Cached Quick I/O, these remaining 7.5GB become available for caching database files.

Note: You cannot allocate specific amounts of the available memory to Cached Quick I/O. When enabled, Cached Quick I/O takes advantage of available memory.

Cached Quick I/O is not beneficial for all containers in a database. Turning on caching for all database containers can degrade performance due to extra memory management overhead (double buffer copying). You must use file I/O statistics to determine which individual database containers benefit from caching, and then enable or disable Cached Quick I/O for individual containers.

If you understand the applications that generate load on your database and how this load changes at different times during the day, you can use Cached Quick I/O to maximize performance. By enabling or disabling Cached Quick I/O on a per-container basis at different times during the day, you are using Cached Quick I/O to dynamically tune the performance of a database.

For example, files that store historical data are not generally used during normal business hours in a transaction processing environment. Reports that make use of this historical data are generally run during off-peak hours when interactive database use is at a minimum. During normal business hours, you can disable Cached Quick I/O for database files that store historical data in order to maximize memory available to other user applications. Then, during off-peak hours, you can enable Cached Quick I/O on the same files when they are used for report generation. This will provide extra memory resources to the database server without changing any database configuration parameters. Enabling file system read-ahead in this manner and buffering read data can provide great performance benefits, especially in large sequential scans.

You can automate the enabling and disabling of Cached Quick I/O on a per-container basis using scripts, allowing the same job that produces reports to tune the file system behavior and make the best use of system resources. You can specify different sets of containers for different jobs to maximize file system and database performance.

Note: Modifying Cached Quick I/O settings does not require any database level changes. So, unlike modifying existing database global memory or bufferpool settings, Cached Quick I/O does not require the database to be restarted for changes to take effect.

How Cached Quick I/O works in a Sybase environment

Cached Quick I/O is a specialized external caching mechanism specifically suitable to 32-bit ports of the Sybase server. Cached Quick I/O can be used on 64-bit ports of the Sybase server, but the benefits are not as great. Cached Quick I/O can be selectively applied to datafiles that are suffering an undesirable amount of physical disk I/O due to insufficient dataserer buffer caches. Cached Quick I/O works by taking advantage of the available physical memory that is left over after the operating system reserves the amount it needs and the Sybase dataserer buffer cache has been sized to the maximum capacity allowed within a 32-bit virtual address space. This extra memory serves as a cache to store file data, effectively serving as a second-level cache backing the dataserer buffer caches.

For example, consider a system configured with 12GB of physical memory, an operating system using 1GB, and a total Sybase size of 3.5GB. Unless you have other applications running on your system, the remaining 7.5GB of memory is unused. If you enable Cached Quick I/O, these remaining 7.5GB become available for caching database files.

Note: You cannot allocate specific amounts of the available memory to Cached Quick I/O. When enabled, Cached Quick I/O takes advantage of available memory.

Cached Quick I/O is not beneficial for all device files in a database. Turning on caching for all database device files can degrade performance due to extra memory management overhead (double buffer copying). You must use file I/O statistics to determine which individual database device files benefit from caching, and then enable or disable Cached Quick I/O for individual device files.

If you understand the applications that generate load on your database and how this load changes at different times during the day, you can use Cached Quick I/O to maximize performance. By enabling or disabling Cached Quick I/O on a per-file basis at different times during the day, you are using Cached Quick I/O to dynamically tune the performance of a database.

For example, files that store historical data are not generally used during normal business hours in a transaction processing environment. Reports that make use of this historical data are generally run during off-peak hours when interactive database use is at a minimum. During normal business hours, you can disable

Cached Quick I/O for database files that store historical data in order to maximize memory available to other user applications. Then, during off-peak hours, you can enable Cached Quick I/O on the same files when they are used for report generation. This will provide extra memory resources to the database server without changing any database configuration parameters. Enabling file system read-ahead in this manner and buffering read data can provide great performance benefits, especially in large sequential scans.

You can automate the enabling and disabling of Cached Quick I/O on a per-file basis using scripts, allowing the same job that produces reports to tune the file system behavior and make the best use of system resources. You can specify different sets of files for different jobs to maximize file system and database performance.

How Cached Quick I/O improves database performance

Enabling Cached Quick I/O on suitable Quick I/O files improves database performance by using the file system buffer cache to store data. This data storage speeds up system reads by accessing the system buffer cache and avoiding disk I/O when searching for information.

Having data at the cache level improves database performance in the following ways:

- For read operations, Cached Quick I/O caches database blocks in the system buffer cache, which can reduce the number of physical I/O operations and therefore improve read performance.
- For write operations, Cached Quick I/O uses a direct-write, copy-behind technique to preserve its buffer copy of the data. After the direct I/O is scheduled and while it is waiting for the completion of the I/O, the file system updates its buffer to reflect the changed data being written out. For online transaction processing, Cached Quick I/O achieves better than raw device performance in database throughput on large platforms with very large physical memories.
- For sequential table scans, Cached Quick I/O can significantly reduce the query response time because of the read-ahead algorithm used by Veritas File System. If a user needs to read the same range in the file while the data is still in cache, the system is likely to return an immediate cache hit rather than scan for data on the disk.

Tasks for setting up Cached Quick I/O

To set up and use Cached Quick I/O, you should do the following in the order in which they are listed:

- Enable Cached Quick I/O on the underlying file systems used for your database.
- Exercise the system in your production environment to generate file I/O statistics.
- Collect the file I/O statistics while the files are in use.
- Analyze the file I/O statistics to determine which files benefit from Cached Quick I/O.
- Disable Cached Quick I/O on files that do not benefit from caching.

Enabling Cached Quick I/O on a file system

Cached Quick I/O depends on Veritas Quick I/O running as an underlying system enhancement in order to function correctly. Follow the procedures listed here to ensure that you have the correct setup to use Cached Quick I/O successfully.

Prerequisites

- You must have permission to change file system behavior using the `vxtunefs` command to enable or disable Cached Quick I/O. By default, you need superuser (`root`) permissions to run the `vxtunefs` command, but other system users do not.
- You must enable Quick I/O on the file system. Quick I/O is enabled automatically at file system mount time.

If you have correctly enabled Quick I/O on your system, you can proceed to enable Cached Quick I/O as follows:

- Set the file system Cached Quick I/O flag, which enables Cached Quick I/O for all files in the file system.
- Setting the file system Cached Quick I/O flag enables caching for all files in the file system. You must disable Cached Quick I/O on individual Quick I/O files that do not benefit from caching to avoid consuming memory unnecessarily. This final task occurs at the end of the enabling process.

Enabling and disabling the `qio_cache_enable` flag

As superuser (`root`), set the `qio_cache_enable` flag using the `vxtunefs` command after you mount the file system.

To enable the `qio_cache_enable` flag for a file system

- ◆ Use the `vxtunefs` command as follows:

```
# /sbin/fs/vxfs5.0/vxtunefs -s -o qio_cache_enable=1 /  
mount_point
```

For example:

```
# /sbin/fs/vxfs5.0/vxtunefs -s -o qio_cache_enable=1 /db02
```

where `/db02` is a VxFS file system containing the Quick I/O files and setting the `qio_cache_enable` flag to “1” enables Cached Quick I/O. This command enables caching for all the Quick I/O files on this file system.

To disable the flag on the same file system

- ◆ Use the `vxtunefs` command as follows:

```
# /sbin/fs/vxfs5.0/vxtunefs -s -o qio_cache_enable=0 /mount_point
```

For example:

```
# /sbin/fs/vxfs5.0/vxtunefs -s -o qio_cache_enable=0 /db02
```

where `/db02` is a VxFS file system containing the Quick I/O files and setting the `qio_cache_enable` flag to “0” disables Cached Quick I/O. This command disables caching for all the Quick I/O files on this file system.

Making Cached Quick I/O settings persistent across reboots and mounts

You can make the Cached Quick I/O system setting persistent across reboots and mounts by adding a file system entry in the `/etc/vx/tunefstab` file.

Note: The `tunefstab` file is a user-created file. For information on how to create the file and add tuning parameters, see the `tunefstab(4)` manual page.

To enable a file system after rebooting

- ◆ Put the file system in the `/etc/vx/tunefstab` file and set the flag entry:

```
/dev/vx/dsk/dgname/volname qio_cache_enable=1
```

where:

- `/dev/vx/dsk/dgname/volname` is the name of a block device
- `dgname` is the name of the disk group

- *volname* is the name of the volume

For example:

```
/dev/vx/dsk/PRODDg/db01 qio_cache_enable=1  
/dev/vx/dsk/PRODDg/db02 qio_cache_enable=1
```

where `/dev/vx/dsk/PRODDg/db01` is the block device on which the file system resides.

The `tunefstab` (4) manual pages contain information on how to add tuning parameters.

See the `tunefstab` (4) manual page.

Note: `vxtunefs` can specify a mount point or a block device; `tunefstab` must always specify a block device only.

Using `vxtunefs` to obtain tuning information

Check the setting of the `qio_cache_enable` flag for each file system using the `vxtunefs` command.

To obtain information on only the `qio_cache_enable` flag setting

- ◆ Use the `grep` command with `vxtunefs`:

```
# /sbin/fs/vxfs6.0/vxtunefs /mount_point | grep qio_cache_enable
```

For example:

```
# /sbin/fs/vxfs6.0/vxtunefs /db01 | grep qio_cache_enable
```

where `/db01` is the name of the file system. This command displays only the `qio_cache_enable` setting as follows:

```
qio_cache_enable = 0
```

You can also use the `vxtunefs` command to obtain a more complete list of I/O characteristics and tuning statistics.

See the `vxtunefs` (1) manual page.

To obtain information on all `vxtunefs` system parameters

- ◆ Use the `vxtunefs` command without `grep`:

```
# /opt/VRTS/bin/vxtunefs /mount_point
```

For example:

```
# /opt/VRTS/bin/vxtunefs /db01
```

The `vxtunefs` command displays output similar to the following:

```
Filesystem i/o parameters for /db01
read_pref_io = 65536
read_nstream = 1
read_unit_io = 65536
write_pref_io = 65536
write_nstream = 1
write_unit_io = 65536
pref_strength = 10
buf_breakup_size = 131072
discovered_direct_iosz = 262144
max_direct_iosz = 1048576
default_indir_size = 8192
qio_cache_enable = 0
odm_cache_enable = 0
write_throttle = 0
max_diskq = 1048576
initial_extent_size = 1
max_seqio_extent_size = 2048
max_buf_data_size = 8192
hsm_write_prealloc = 0
read_ahead = 1
inode_aging_size = 0
inode_aging_count = 0
fcl_maxalloc = 887660544
fcl_keeptime = 0
fcl_winterval = 3600
fcl_ointerval = 600
oltp_load = 0
delicache_enable = 1
thin_friendly_alloc = 0
dalloc_enable = 1
dalloc_limit = 90
```

The `vxtunefs(1)` manual pages contain a complete description of `vxtunefs` parameters and the tuning instructions.

See the `vxtunefs(1)` manual page.

Determining candidates for Cached Quick I/O

Determining which files can benefit from Cached Quick I/O is an iterative process that varies with each application. For this reason, you may need to complete the following steps more than once to determine the best possible candidates for Cached Quick I/O.

Before determining candidate files for Quick I/O, make sure the following conditions have been met:

- Prerequisites ■ You must enable Cached Quick I/O for the file systems.
- Usage notes ■ See the `qiostat` (1M) manual page for more information.

Collecting I/O statistics

Once you have enabled Cached Quick I/O on a file system, you need to collect statistics to determine and designate the files that can best take advantage of its benefits.

To collect statistics needed to determine files that benefit from Cached Quick I/O

- 1 Reset the `qiostat` counters by entering:

```
$ /opt/VRTS/bin/qiostat -r /mount_point/filenames
```

- 2 Run the database under full normal load and through a complete cycle (24 to 48 hours in most cases) to determine your system I/O patterns and database traffic in different usage categories (for example, OLTP, reports, and backups) at different times of the day.
- 3 While the database is running, run `qiostat -l` to report the caching statistics as follows:

```
$ /opt/VRTS/bin/qiostat -l /mount_point/filenames
```

or, use the `-i` option to see statistic reports at specified intervals:

```
$ /opt/VRTS/bin/qiostat -i n /mount_point/filenames
```

where `n` is time in seconds

For example:

To collect I/O statistics from all database files on file system :

```
$ /opt/VRTS/bin/qiostat -l /db01/*.dbf
```

About I/O statistics for DB2

The output of the `qiostat` command is the primary source of information to use in deciding whether to enable or disable Cached Quick I/O on specific files. Statistics are printed in two lines per object.

The second line of information is defined as follows:

- `CREADS` is the number of reads from the VxFS cache (or total number of reads to Quick I/O files with cache advisory on)
- `PREADS` is the number of reads going to the disk for Quick I/O files with the cache advisory on
- `HIT_RATIO` is displayed as a percentage and is the number of `CREADS` minus the number of `PREADS` times 100 divided by the total number of `CREADS`. The formula looks like this:

$$(\text{CREADS} - \text{PREADS}) * 100 / \text{CREADS}$$

The `qiostat -l` command output looks similar to the following:

```
/db01/tbs1_cont001 6          1    21    4    10.0  0.0
                   6          6    0.0

/db01/tbs2_cont001 62552 38498 250213 153992 21.9  0.4
62567 49060    21.6

/db01/tbs2_cont002 62552 38498 250213 153992 21.9  0.4
62567 49060    21.6
```

Analyze the output to find out where the cache-hit ratio is above a given threshold. A cache-hit ratio above 20 percent on a file for a given application may be sufficient to justify caching on that file. For systems with larger loads, the acceptable ratio may be 30 percent or above. Cache-hit-ratio thresholds vary according to the database type and load.

Using the sample output above as an example, the file `/db01/tbs1_cont001` does not benefit from the caching because the cache-hit ratio is zero. In addition, the file receives very little I/O during the sampling duration.

However, the files `/db01/tbs2_*` have a cache-hit ratio of 21.6 percent. If you have determined that, for your system and load, this figure is above the acceptable threshold, it means the database can benefit from caching. Also, study the numbers reported for the read and write operations. When you compare the number of reads and writes for the `/db01/tbs2_*` files, you see that the number of reads is roughly twice the number of writes. You can achieve the greatest performance

gains with Cached Quick I/O when using it for files that have higher read than write activity.

Based on these two factors, `/db01/tbs2_cont001` and `/db01/tbs2_cont002` are prime candidates for Cached Quick I/O.

About I/O statistics for Sybase

The output of the `qiostat` command is the primary source of information to use in deciding whether to enable or disable Cached Quick I/O on specific files. Statistics are printed in two lines per object.

The second line of information is defined as follows:

- `CREAD` is the number of reads from the VxFS cache (or total number of reads to Quick I/O files with cache advisory on)
- `PREAD` is the number of reads going to the disk for Quick I/O files with the cache advisory on
- `HIT RATIO` is displayed as a percentage and is the number of `CREADS` minus the number of `PREADS` times 100 divided by the total number of `CREADS`. The formula looks like this:

$$(\text{CREADS} - \text{PREADS}) * 100 / \text{CREADS}$$

The `qiostat -l` command output looks similar to the following:

```
/db01/sysprocs.dbf 17128 9634 68509 38536 24.8 0.4
17124 15728 8.2

/db1/master.dbf 6 1 21 4 10.0 0.0
6 6 0.0

/db01/user.dbf 62552 38498 250213 153992 21.9 0.4
62567 49060 21.6
```

Analyze the output to find out where the cache-hit ratio is above a given threshold. A cache-hit ratio above 20 percent on a file for a given application may be sufficient to justify caching on that file. For systems with larger loads, the acceptable ratio may be 30 percent or above. Cache-hit-ratio thresholds vary according to the database type and load.

Using the sample output above as an example, the file `/db01/master.dbf` does not benefit from the caching because the cache-hit ratio is zero. In addition, the file receives very little I/O during the sampling duration.

However, the file `/db01/user.dbf` has a cache-hit ratio of 21.6 percent. If you have determined that, for your system and load, this figure is above the acceptable

threshold, it means the database can benefit from caching. Also, study the numbers reported for the read and write operations. When you compare the number of reads and writes for the `/db01/user.dbf` file, you see that the number of reads is roughly twice the number of writes. You can achieve the greatest performance gains with Cached Quick I/O when using it for files that have higher read than write activity.

Based on these two factors, `/db01/user.dbf` is a prime candidate for Cached Quick I/O.

Effects of read-aheads on I/O statistics

The number of `CREADS` in the `qiostat` output is the total number of reads performed, including Cached Quick I/O, and the number of `PREADS` is the number of physical reads. The difference between `CREADS` and `PREADS` (`CREADS - PREADS`) is the number of reads satisfied from the data in the file system cache. Thus, you expect that the number of `PREADS` would always be equal to or lower than the number of `CREADS`.

However, the `PREADS` counter also increases when the file system performs read-aheads. These read-aheads occur when the file system detects sequential reads. In isolated cases where cache hits are extremely low, the output from `qiostat` could show that the number of `CREADS` is lower than the number of `PREADS`. The cache-hit ratio calculated against these `CREAD/PREAD` values is misleading when used to determine whether Cached Quick I/O should be enabled or disabled.

For DB2, you can make a more accurate decision based on a collective set of statistics by gathering multiple sets of data points. Consequently, you might want to enable Cached Quick I/O for all the container files that contain a particular table, across multiple tablespaces used by a given database, even if the containers in just one of the tablespaces exhibited a high cache hit ratio. In general, we expect all containers in a tablespace to have approximately the same cache hit ratio.

For Sybase, you can make a more accurate decision based on a collective set of statistics by gathering multiple sets of data points. Consequently, you might want to enable Cached Quick I/O for all the data files in a given tablespace, even if just one of the files exhibited a high cache-hit ratio.

Other tools for analysis

While the output of the `qiostat` command is the primary source of information to use in deciding whether to enable Cached Quick I/O on specific files, we also recommend using other tools in conjunction with `qiostat`. For example, benchmarking software that measures database throughput is also helpful. If a benchmark test in which Cached Quick I/O was enabled for a certain set of data

files resulted in improved performance, you can also use those results as the basis for enabling Cached Quick I/O.

Enabling and disabling Cached Quick I/O for individual files

After using `qiostat` or other analysis tools to determine the appropriate files for Cached Quick I/O, you need to disable Cached Quick I/O for those individual files that do not benefit from caching using the `qioadmin` command.

- | | |
|---------------|---|
| Prerequisites | ■ Enable Cached Quick I/O for the file system before enabling or disabling Cached Quick I/O at the individual file level. |
| Usage notes | ■ You can enable or disable Cached Quick I/O for individual files while the database is online.
■ You should monitor files regularly using <code>qiostat</code> to ensure that a file's cache-hit ratio has not changed enough to reconsider enabling or disabling Cached Quick I/O for the file.
■ Enabling or disabling Cached Quick I/O for an individual file is also referred to as setting the cache advisory on or off.
■ See the <code>qioadmin(1)</code> manual page. |

Setting cache advisories for individual files

You can enable and disable Cached Quick I/O for individual files by changing the cache advisory settings for those files.

To disable Cached Quick I/O for an individual file

- ◆ Use the `qioadmin` command to set the cache advisory to `OFF` as follows:

```
$ /opt/VRTS/bin/qioadmin -S filename=OFF /mount_point
```

For example for DB2, to disable Cached Quick I/O for the file `/db01/dbfile`, set the cache advisory to `OFF`:

```
$ /opt/VRTS/bin/qioadmin -S dbfile=OFF /db01
```

For example for Sybase, to disable Cached Quick I/O for the file `/db01/master.dbf`, set the cache advisory to `OFF`:

```
$ /opt/VRTS/bin/qioadmin -S master.dbf=OFF /db01
```

To enable Cached Quick I/O for an individual file

- ◆ Use the `qioadmin` command to set the cache advisory to `ON` as follows:

```
$ /opt/VRTS/bin/qioadmin -S filename=ON /mount_point
```

For example for DB2, running `qiostat` shows the cache hit ratio for the file `/db01/dbfile` reaches a level that would benefit from caching. To enable Cached Quick I/O for the file `/db01/dbfile`, set the cache advisory to `ON`:

```
$ /opt/VRTS/bin/qioadmin -S dbfile=ON /db01
```

For example for Sybase, running `qiostat` shows the cache hit ratio for the file `/db01/master.dbf` reaches a level that would benefit from caching. To enable Cached Quick I/O for the file `/db01/master.dbf`, set the cache advisory to `ON`:

```
$ /opt/VRTS/bin/qioadmin -S master/dbf=ON /db01
```

Making individual file settings for Cached Quick I/O persistent

You can make the enable or disable individual file settings for Cached Quick I/O persistent across reboots and mounts by adding cache advisory entries in the `/etc/vx/qioadmin` file.

Cache advisories set using the `qioadmin` command are stored as extended attributes of the file in the inode. These settings persist across file system remounts and system reboots, but these attributes are not backed up by the usual backup methods, so they cannot be restored. Therefore, always be sure to reset cache advisories after each file restore. This is not necessary if you maintain the cache advisories for Quick I/O files in the `/etc/vx/qioadmin` file.

To enable or disable individual file settings for Cached Quick I/O automatically after a reboot or mount

- ◆ Add cache advisory entries in the `/etc/vx/qioadmin` file as follows:

```
device=/dev/vx/dsk/<diskgroup>/<volume>

filename,OFF

filename,OFF

filename,OFF

filename,ON
```

For example, to make the Cached Quick I/O settings for individual files in the `/db01` file system persistent, edit the `/etc/vx/qioadmin` file similar to the following:

```
#
# List of files to cache in /db01 file system
#
device=/dev/vx/dsk/PRODDg/db01
```

For DB2

```
dbfile01,OFF
dbfile02,OFF
dbfile03,OFF
```

For Sybase

```
user.dbf,ON
sysprocs.dbf,OFF
master.dbf,OFF
```

Determining individual file settings for Cached Quick I/O using `qioadmin`

You can determine whether Cached Quick I/O is enabled or disabled for individual files by displaying the file's cache advisory setting using the `qioadmin` command.

Note: To verify caching, always check the setting of the flag `qio_cache_enable` using `vxtunefs`, along with the individual cache advisories for each file.

To display the current cache advisory settings for a file

- ◆ Use the `qioadmin` command with the `-P` option as follows:

```
$ /opt/VRTS/bin/qioadmin -P filename /mount_point
```

For example for DB2, to display the current cache advisory setting for the file `dbfile` in the `/db01` file system:

```
$ /opt/VRTS/bin/qioadmin -P dbfile /db01
```

```
dbfile,OFF
```

For example for Sybase, to display the current cache advisory setting for the file `sysprocs.dbf` in the `/db01` file system:

```
$ /opt/VRTS/bin/qioadmin -P sysprocs.dbf /db01
```

```
sysprocs.dbf,OFF
```


Improving database performance with Veritas Concurrent I/O

This chapter includes the following topics:

- [About Concurrent I/O](#)
- [Tasks for enabling and disabling Concurrent I/O](#)

About Concurrent I/O

Veritas Concurrent I/O improves the performance of regular files on a VxFS file system without the need for extending namespaces and presenting the files as devices. This simplifies administrative tasks and allows databases, which do not have a sequential read/write requirement, to access files concurrently. This chapter describes how to use the Concurrent I/O feature.

Quick I/O is an alternative solution for DMS tablespaces.

In some cases (for example, if the system has extra memory), Cached Quick I/O may further enhance performance.

See [“About Cached Quick I/O”](#) on page 49.

How Concurrent I/O works

Traditionally, UNIX semantics require that read and write operations on a file occur in a serialized order. Because of this, a file system must enforce strict ordering of overlapping read and write operations. However, databases do not

usually require this level of control and implement concurrency control internally, without using a file system for order enforcement.

The Veritas Concurrent I/O feature removes these semantics from the read and write operations for databases and other applications that do not require serialization.

The benefits of using Concurrent I/O are:

- Concurrency between a single writer and multiple readers
- Concurrency among multiple writers
- Minimalization of serialization for extending writes
- All I/Os are direct and do not use file system caching
- I/O requests are sent directly to file systems
- Inode locking is avoided

Tasks for enabling and disabling Concurrent I/O

Concurrent I/O is not turned on by default and must be enabled manually. You will also have to manually disable Concurrent I/O if you choose not to use it in the future.

You can perform the following tasks:

- Enable Concurrent I/O
- Disable Concurrent I/O

Enabling Concurrent I/O for DB2

Because you do not need to extend name spaces and present the files as devices, you can enable Concurrent I/O on regular files.

For DB2, you can enable an entire file system to use Concurrent I/O or you can enable specific SMS containers to use Concurrent I/O. If you enable a specific SMS container, the rest of the file system will use the regular buffer I/O.

Before enabling Concurrent I/O, review the following:

- Prerequisites
- To use the Concurrent I/O feature, the file system must be a VxFS file system.
 - Make sure the mount point on which you plan to mount the file system exists.
 - Make sure the DBA can access the mount point.

For DB2, */mount_point* is the directory in which you can put data containers of the SMS tablespaces using the Concurrent I/O feature.

Note: This applies to both creating a new tablespace to use Concurrent I/O or enabling an existing tablespace to use Concurrent I/O.

For example for DB2 to mount a file system named */datavol* on a mount point named */db2data*:

To enable Concurrent I/O on a new SMS container using the `namefs -o cio` option

- ◆ Using the `mount` command, mount the directory in which you want to put data containers of the SMS tablespaces using the Concurrent I/O feature.

```
# /usr/sbin/mount -vt namefs -o cio /path_name /new_mount_point
```

where:

- */path_name* is the directory in which the files that will be using Concurrent I/O reside
- */new_mount_point* is the new target directory that will use the Concurrent I/O feature

The following is an example of mounting a directory (where the new SMS containers are located) to use Concurrent I/O.

To mount an SMS container named */container1* on a mount point named */mysms*:

```
# /usr/sbin/mount -vt namefs -o cio /datavol/mysms/container1 /mysms
```

To enable Concurrent I/O on an existing SMS container using the `namefs -o cio` option

- 1 Stop the DB2 instance using the `db2stop` command.
- 2 Make the directory that will have Concurrent I/O turned on available using the `mv` command.

```
# mv /mydb/mysmsdir /mydb/mysmsdir2
```

- 3 Remount `/mydb/mysmsdir2` on `/mydb/mysmsdir` using the `mount` command with the `-o cio` option.

```
# mount -Vt namefs -o cio /mydb/mysmsdir2 /mydb/mysmsdir
```

- 4 Start the DB2 instance using the `db2start` command.

```
# db2stop
# mv /mydb/mysmsdir /mydb/mysmsdir2
# mount -Vt namefs -o cio /mydb/mysmsdir2 /mydb/mysmsdir
# db2start
```

This example shows how to mount a directory for an existing SMS container to use Concurrent I/O.

To enable Concurrent I/O on a DB2 tablespace when creating the tablespace

- 1 Use the `db2 -v "create regular tablespace..."` command with the `no file system caching` option.
- 2 Set all other parameters according to your system requirements.

To enable Concurrent I/O on an existing DB2 tablespace

- ◆ Use the DB2 `no file system caching` option as follows:

```
# db2 -v "alter tablespace tablespace_name no file system caching"
```

where *tablespace_name* is the name of the tablespace for which you are enabling Concurrent I/O.

To verify that Concurrent I/O has been set for a particular DB2 tablespace

- 1 Use the DB2 `get snapshot` option to check for Concurrent I/O.

```
# db2 -v "get snapshot for tablespaces on dbname"
```

where *dbname* is the database name.

- 2 Find the tablespace you want to check and look for the `File system caching` attribute. If you see `File system caching = No`, then Concurrent I/O is enabled.

Disabling Concurrent I/O for DB2

If you need to disable Concurrent I/O, use the DB2 `file system caching` option.

Note: If you used the `namefs -o cio` option with the `mount` command to mount a directory to enable Concurrent I/O, make sure you remount without that option as well. Also, if you follow the directions for enabling Concurrent I/O on an existing SMS container, rename the directory back to the original name.

To disable Concurrent I/O on a DB2 tablespace

- ◆ Use the DB2 `file system caching` option as follows:

```
# db2 -v "alter tablespace tablespace_name file system caching"
```

where *tablespace_name* is the name of the tablespace for which you are disabling Concurrent I/O.

Enabling Concurrent I/O for Sybase

Because you do not need to extend name spaces and present the files as devices, you can enable Concurrent I/O on regular files.

Before enabling Concurrent I/O, review the following:

- Prerequisites
- To use the Concurrent I/O feature, the file system must be a VxFS file system.
 - Make sure the mount point on which you plan to mount the file system exists.
 - Make sure the DBA can access the mount point.

To enable Concurrent I/O on a file system using `mount` with the `-o cio` option

- ◆ Mount the file system using the `mount` command as follows:

```
# /usr/sbin/mount -F vxfs -o cio special /mount_point
```

where:

- *special* is a block special device.
- */mount_point* is the directory where the file system will be mounted.

For example for Sybase, to mount a file system named `/datavol` on a mount point named `/sybasedata`:

```
# /usr/sbin/mount -F vxfs -o cio /dev/vx/dsk/sybasedg/datavol \  
/sybasedata
```

The following is an example of mounting a directory (where the new SMS containers are located) to use Concurrent I/O.

To mount an SMS container named `/container1` on a mount point named `/mysms`:

```
# /usr/sbin/mount -vt namefs -o cio /datavol/mysms/container1 /mysms
```

Disabling Concurrent I/O for Sybase

If you need to disable Concurrent I/O, unmount the VxFS file system and mount it again without the mount option.

To disable Concurrent I/O on a file system using the mount command

- 1 Shutdown the DB instance.
- 2 Unmount the file sytem using the `umount` command.
- 3 Mount the file system again using the `mount` command without using the `-o cio` option.

Using point-in-time copies

- [Chapter 6. Understanding point-in-time copy methods](#)
- [Chapter 7. Backing up and recovering](#)
- [Chapter 8. Backing up and recovering in a NetBackup environment](#)
- [Chapter 9. Off-host processing](#)
- [Chapter 10. Creating and refreshing test environments](#)
- [Chapter 11. Creating point-in-time copies of files](#)

Understanding point-in-time copy methods

This chapter includes the following topics:

- [About point-in-time copies](#)
- [When to use point-in-time copies](#)
- [About Storage Foundation point-in-time copy technologies](#)

About point-in-time copies

Veritas Storage Foundation offers a flexible and efficient means of managing business-critical data. Storage Foundation lets you capture an online image of an actively changing database at a given instant, called a point-in-time copy.

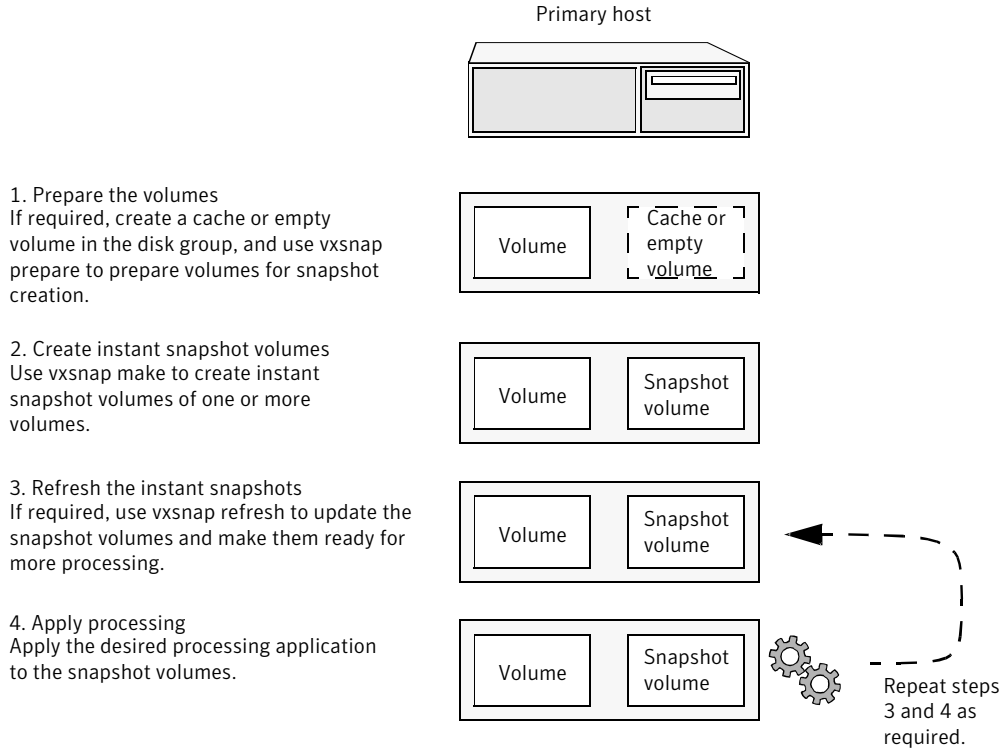
More and more, the expectation is that the data must be continuously available (24x7) for transaction processing, decision making, intellectual property creation, and so forth. Protecting the data from loss or destruction is also increasingly important. Formerly, data was taken out of service so that the data did not change while data backups occurred; however, this option does not meet the need for minimal down time.

A point-in-time copy enables you to maximize the online availability of the data. You can perform system backup, upgrade, or perform other maintenance tasks on the point-in-time copies. The point-in-time copies can be processed on the same host as the active data, or a different host. If required, you can offload processing of the point-in-time copies onto another host to avoid contention for system resources on your production server. This method is called off-host processing. If implemented correctly, off-host processing solutions have almost no impact on the performance of the primary production system.

Implementing point-in time copy solutions on a primary host

Figure 6-1 illustrates the steps that are needed to set up the processing solution on the primary host.

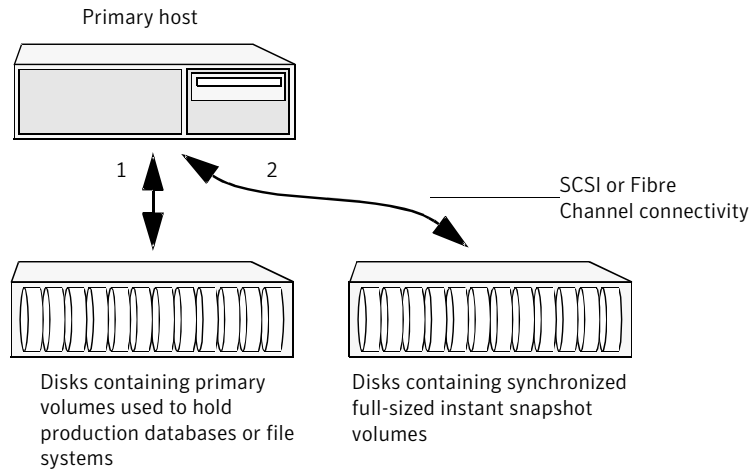
Figure 6-1 Using snapshots and FastResync to implement point-in-time copy solutions on a primary host



Note: The Disk Group Split/Join functionality is not used. As all processing takes place in the same disk group, synchronization of the contents of the snapshots from the original volumes is not usually required unless you want to prevent disk contention. Snapshot creation and updating are practically instantaneous.

Figure 6-2 shows the suggested arrangement for implementing solutions where the primary host is used and disk contention is to be avoided.

Figure 6-2 Example point-in-time copy solution on a primary host



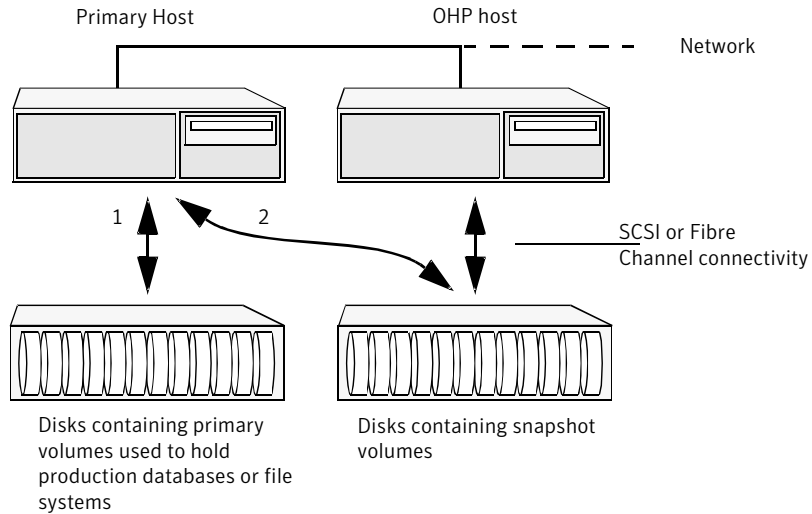
In this setup, it is recommended that separate paths (shown as 1 and 2) from separate controllers be configured to the disks containing the primary volumes and the snapshot volumes. This avoids contention for disk access, but the primary host's CPU, memory and I/O resources are more heavily utilized when the processing application is run.

Note: For space-optimized or unsynchronized full-sized instant snapshots, it is not possible to isolate the I/O pathways in this way. This is because such snapshots only contain the contents of changed regions from the original volume. If applications access data that remains in unchanged regions, this is read from the original volume.

Implementing off-host point-in-time copy solutions

Figure 6-3 illustrates that, by accessing snapshot volumes from a lightly loaded host (shown here as the OHP host), CPU- and I/O-intensive operations for online backup and decision support are prevented from degrading the performance of the primary host that is performing the main production activity (such as running a database).

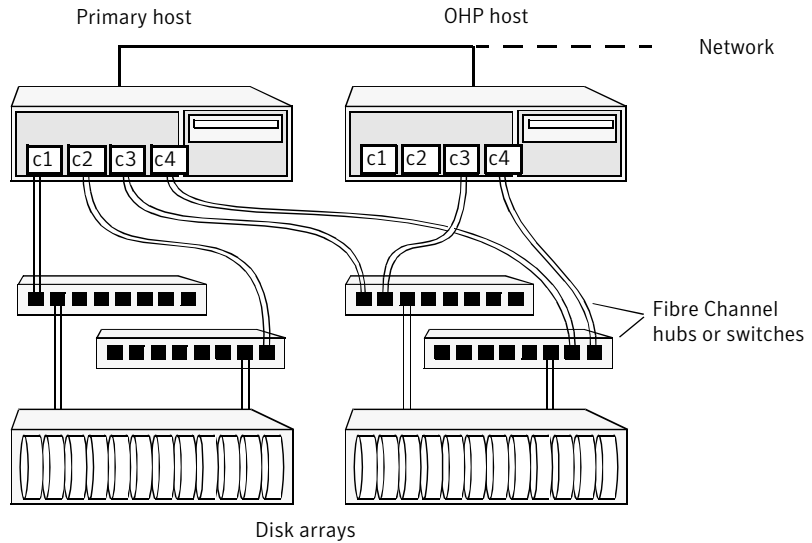
Figure 6-3 Example implementation of an off-host point-in-time copy solution



Also, if you place the snapshot volumes on disks that are attached to host controllers other than those for the disks in the primary volumes, it is possible to avoid contending with the primary host for I/O resources. To implement this, paths 1 and 2 shown in the [Figure 6-3](#) should be connected to different controllers.

[Figure 6-4](#) shows an example of how you might achieve such connectivity using Fibre Channel technology with 4 Fibre Channel controllers in the primary host.

Figure 6-4 Example connectivity for off-host solution using redundant-loop access



This layout uses redundant-loop access to deal with the potential failure of any single component in the path between a system and a disk array.

Note: On some operating systems, controller names may differ from what is shown here.

[Figure 6-5](#) shows how off-host processing might be implemented in a cluster by configuring one of the cluster nodes as the OHP node.

Figure 6-5 Example implementation of an off-host point-in-time copy solution using a cluster node

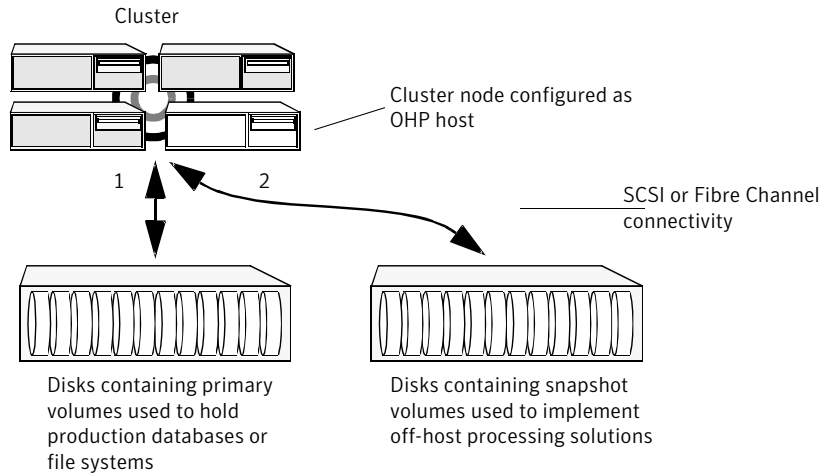
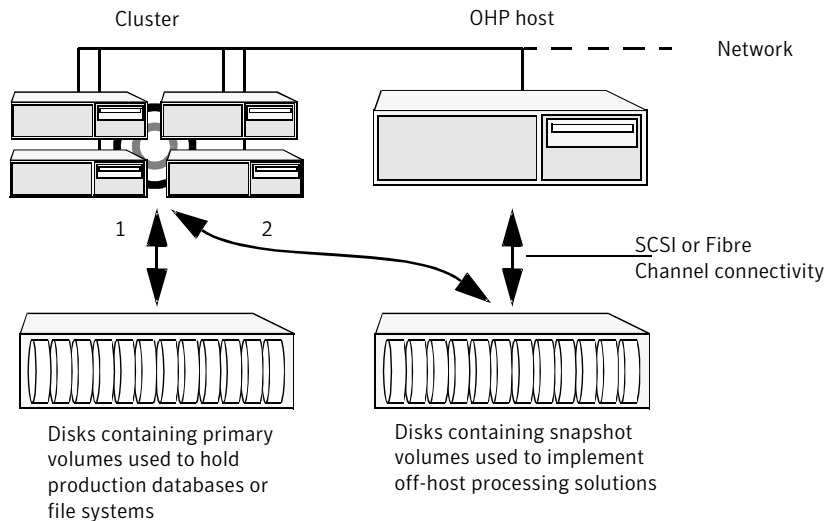


Figure 6-6 shows an alternative arrangement, where the OHP node could be a separate system that has a network connection to the cluster, but which is not a cluster node and is not connected to the cluster's private network.

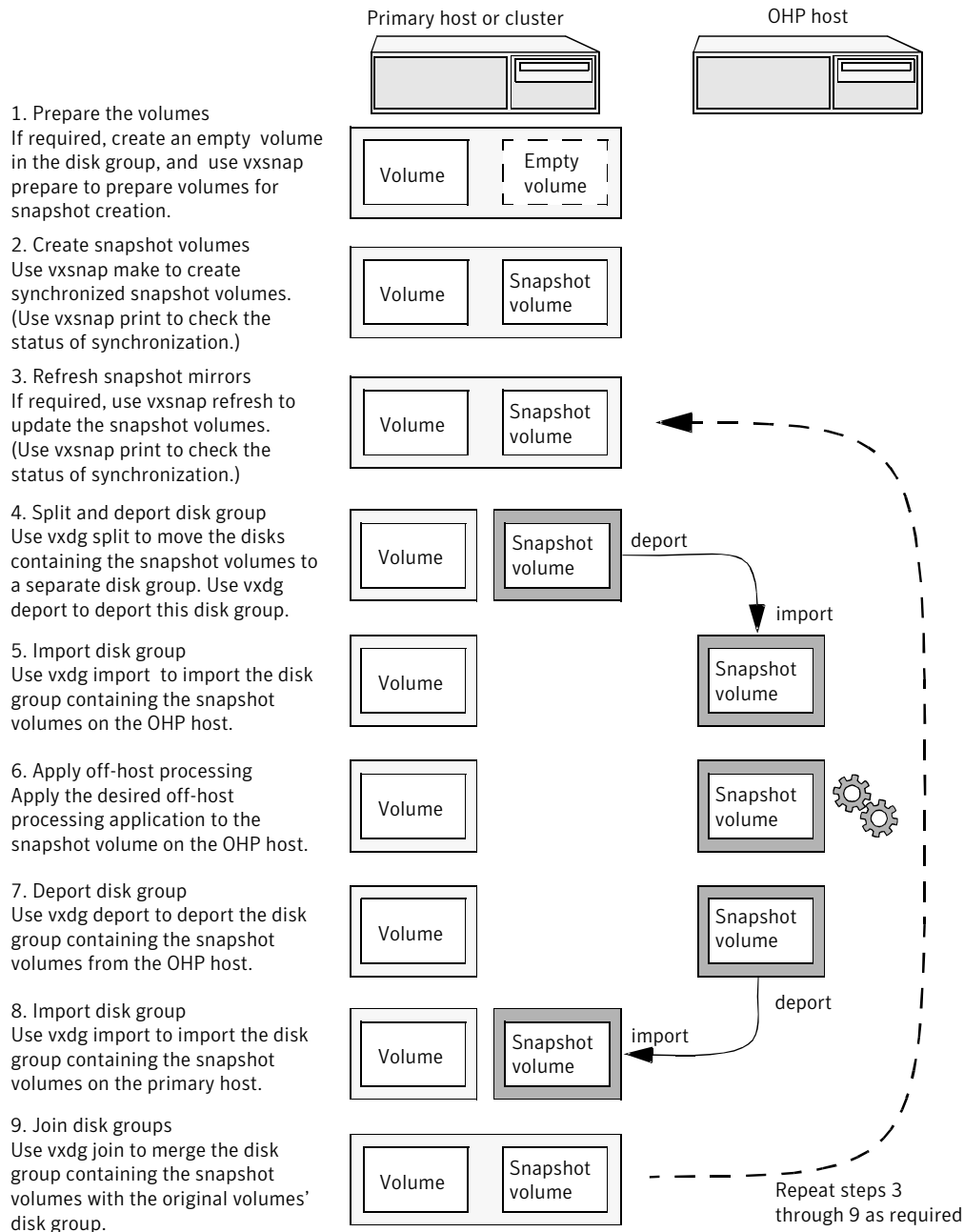
Figure 6-6 Example implementation of an off-host point-in-time copy solution using a separate OHP host



Note: For off-host processing, the example scenarios in this document assume that a separate OHP host is dedicated to the backup or decision support role. For clusters, it may be simpler, and more efficient, to configure an OHP host that is not a member of the cluster.

[Figure 6-7](#) illustrates the steps that are needed to set up the processing solution on the primary host.

Figure 6-7 Implementing off-host processing solutions



Disk Group Split/Join is used to split off snapshot volumes into a separate disk group that is imported on the OHP host.

Note: As the snapshot volumes are to be moved into another disk group and then imported on another host, their contents must first be synchronized with the parent volumes. On reimporting the snapshot volumes, refreshing their contents from the original volume is speeded by using FastResync.

When to use point-in-time copies

The following typical activities are suitable for point-in-time copy solutions implemented using Veritas FlashSnap:

- Data backup –Many enterprises require 24 x 7 data availability. They cannot afford the downtime involved in backing up critical data offline. By taking snapshots of your data, and backing up from these snapshots, your business-critical applications can continue to run without extended downtime or impacted performance.
- Providing data continuity –To provide continuity of service in the event of primary storage failure, you can use point-in-time copy solutions to recover application data. In the event of server failure, you can use point-in-time copy solutions in conjunction with the high availability cluster functionality of Veritas Storage Foundation™ for Cluster File System HA or Veritas Storage Foundation HA.
- Decision support analysis and reporting—Operations such as decision support analysis and business reporting may not require access to real-time information. You can direct such operations to use a replica database that you have created from snapshots, rather than allow them to compete for access to the primary database. When required, you can quickly resynchronize the database copy with the data in the primary database.
- Testing and training—Development or service groups can use snapshots as test data for new applications. Snapshot data provides developers, system testers and QA groups with a realistic basis for testing the robustness, integrity and performance of new applications.
- Database error recovery—Logic errors caused by an administrator or an application program can compromise the integrity of a database. You can recover a database more quickly by restoring the database files by using Storage Checkpoints or a snapshot copy than by full restoration from tape or other backup media.

Use Storage Checkpoints to quickly roll back a database instance to an earlier point in time.

- Cloning data—You can clone your file system or application data. This functionality enable you to quickly and efficiently provision virtual desktops.

All of the snapshot solutions mentioned above are also available on the disaster recovery site, in conjunction with Veritas Volume Replicator.

For more information about snapshots with replication, see the *Veritas Storage Foundation and High Availability Solutions Replication Administrator's Guide*.

Veritas Storage Foundation provides several point-in-time copy solutions that support your needs, including the following use cases:

- Creating a replica database for decision support.
See [“Using a replica database for decision support”](#) on page 140.
- Backing up and recovering a database with snapshots.
See [“Online database backups”](#) on page 93.
- Backing up and recovering an off-host cluster file system
See [“Backing up on an off-host cluster file system”](#) on page 116.
- Backing up and recovering an online database.
See [“Database recovery using Storage Checkpoints”](#) on page 125.

About Storage Foundation point-in-time copy technologies

This topic introduces the point-in-time copy solutions that you can implement using the Veritas FlashSnap™ technology. Veritas FlashSnap technology requires an enterprise product license.

Veritas FlashSnap offers a flexible and efficient means of managing business critical data. It allows you to capture an online image of actively changing data at a given instant: a point-in-time copy. You can perform system backup, upgrade and other maintenance tasks on point-in-time copies while providing continuous availability of your critical data. If required, you can offload processing of the point-in-time copies onto another host to avoid contention for system resources on your production server.

The following kinds of point-in-time copy solution are supported with an enterprise license:

- Volume-level solutions. There are several types of volume-level snapshots. These features are suitable for solutions where separate storage is desirable

to create the snapshot. For example, lower-tier storage. Some of these techniques provided exceptional offhost processing capabilities.

See “[Volume-level snapshots](#)” on page 83.

- File system-level solutions use the Storage Checkpoint feature of Veritas File System. Storage Checkpoints are suitable for implementing solutions where storage space is critical for:
 - File systems that contain a small number of mostly large files.
 - Application workloads that change a relatively small proportion of file system data blocks (for example, web server content and some databases).
 - Applications where multiple writable copies of a file system are required for testing or versioning.See “[Storage Checkpoints](#)” on page 84.
- File level snapshots.
The FileSnap feature provides snapshots at the level of individual files.
See “[About FileSnaps](#)” on page 85.

Volume-level snapshots

A volume snapshot is an image of a Veritas Volume Manager (VxVM) volume at a given point in time. You can also take a snapshot of a volume set.

Volume snapshots allow you to make backup copies of your volumes online with minimal interruption to users. You can then use the backup copies to restore data that has been lost due to disk failure, software errors or human mistakes, or to create replica volumes for the purposes of report generation, application development, or testing.

Volume snapshots can also be used to implement off-host online backup.

Physically, a snapshot may be a full (complete bit-for-bit) copy of the data set, or it may contain only those elements of the data set that have been updated since snapshot creation. The latter are sometimes referred to as allocate-on-first-write snapshots, because space for data elements is added to the snapshot image only when the elements are updated (overwritten) for the first time in the original data set. Storage Foundation allocate-on-first-write snapshots are called space-optimized snapshots.

Persistent FastResync of volume snapshots

If persistent FastResync is enabled on a volume, VxVM uses a FastResync map to keep track of which blocks are updated in the volume and in the snapshot.

When snapshot volumes are reattached to their original volumes, persistent FastResync allows the snapshot data to be quickly refreshed and re-used. Persistent FastResync uses disk storage to ensure that FastResync maps survive both system and cluster crashes. If persistent FastResync is enabled on a volume in a private disk group, incremental resynchronization can take place even if the host is rebooted.

Persistent FastResync can track the association between volumes and their snapshot volumes after they are moved into different disk groups. After the disk groups are rejoined, persistent FastResync allows the snapshot plexes to be quickly resynchronized.

Data integrity in volume snapshots

A volume snapshot captures the data that exists in a volume at a given point in time. As such, VxVM does not have any knowledge of data that is cached in memory by the overlying file system, or by applications such as databases that have files open in the file system. Snapshots are always crash consistent, that is, the snapshot can be put to use by letting the application perform its recovery. This is similar to how the application recovery occurs after a server crash. If the `fsgen` volume usage type is set on a volume that contains a mounted Veritas File System (VxFS), VxVM coordinates with VxFS to flush data that is in the cache to the volume. Therefore, these snapshots are always VxFS consistent and require no VxFS recovery while mounting.

For databases, a suitable mechanism must additionally be used to ensure the integrity of tablespace data when the volume snapshot is taken. The facility to temporarily suspend file system I/O is provided by most modern database software. The examples provided in this document illustrate how to perform this operation. For ordinary files in a file system, which may be open to a wide variety of different applications, there may be no way to ensure the complete integrity of the file data other than by shutting down the applications and temporarily unmounting the file system. In many cases, it may only be important to ensure the integrity of file data that is not in active use at the time that you take the snapshot. However, in all scenarios where application coordinate, snapshots are crash-recoverable.

Storage Checkpoints

A Storage Checkpoint is a persistent image of a file system at a given instance in time. Storage Checkpoints use a copy-on-write technique to reduce I/O overhead by identifying and maintaining only those file system blocks that have changed since a previous Storage Checkpoint was taken. Storage Checkpoints have the following important features:

- Storage Checkpoints persist across system reboots and crashes.

- A Storage Checkpoint can preserve not only file system metadata and the directory hierarchy of the file system, but also user data as it existed when the Storage Checkpoint was taken.
- After creating a Storage Checkpoint of a mounted file system, you can continue to create, remove, and update files on the file system without affecting the image of the Storage Checkpoint.
- Unlike file system snapshots, Storage Checkpoints are writable.
- To minimize disk space usage, Storage Checkpoints use free space in the file system.

Storage Checkpoints and the Storage Rollback feature of Veritas Storage Foundation for Databases enable rapid recovery of databases from logical errors such as database corruption, missing files and dropped table spaces. You can mount successive Storage Checkpoints of a database to locate the error, and then roll back the database to a Storage Checkpoint before the problem occurred.

See “[Database recovery using Storage Checkpoints](#)” on page 125.

Symantec NetBackup for Oracle Advanced BLI Agent uses Storage Checkpoints to enhance the speed of backing up Oracle databases.

See the *Symantec NetBackup for Oracle Advanced BLI Agent System Administrator's Guide*.

About FileSnaps

A FileSnap is an atomic space-optimized copy of a file in the same name space, stored in the same file system. Veritas File System (VxFS) supports snapshots on file system disk layout Version 8 and later.

FileSnaps provide an ability to snapshot objects that are smaller in granularity than a file system or a volume. The ability to snapshot parts of a file system name space is required for application-based or user-based management of data stored in a file system. This is useful when a file system is shared by a set of users or applications or the data is classified into different levels of importance in the same file system.

All regular file operations are supported on the FileSnap, and VxFS does not distinguish the FileSnap in any way.

Backing up and recovering

This chapter includes the following topics:

- [About Storage Foundation and High Availability Solutions backup and recovery use cases](#)
- [Creating and maintaining a full image snapshot and incremental point-in-time copies](#)
- [Online database backups](#)
- [Backing up on an off-host cluster file system](#)
- [Database recovery using Storage Checkpoints](#)

About Storage Foundation and High Availability Solutions backup and recovery use cases

Storage Foundation and High Availability Solutions (SFHA Solutions) provide point-in-time copy methods which can be applied to multiple database backup use cases.

Example procedures are provided for the following use cases:

- [Creating and maintaining a full image snapshot and incremental point-in-time copies](#)
See [“Creating and maintaining a full image snapshot and incremental point-in-time copies”](#) on page 88.
- [Backing up and recovering a database with snapshots.](#)
See [“Online database backups”](#) on page 93.
- [Backing up and recovering an off-host cluster file system](#)
See [“Backing up on an off-host cluster file system”](#) on page 116.

- Backing up and recovering an online database.
See “[Database recovery using Storage Checkpoints](#)” on page 125.
- Recovering a database with Storage Checkpoints
See “[Database recovery using Storage Checkpoints](#)” on page 125.
- Backing up and recovering a database in a NetBackup environment.

For basic backup and recovery configuration information, see the *Veritas Storage Foundation Administrator's Guide*.

Creating and maintaining a full image snapshot and incremental point-in-time copies

On-disk snapshots are efficient when it comes to recovering a logically corrupted data. Storage Foundation and High Availability Solutions (SFHA Solutions) provide a cost effective and very efficient mechanism to manage multiple copies of production data at different points in time. With Flashsnap, you can create a solution to preserve and manage the whole lifecycle of snapshots for recovery from logical data corruption. You can also use the preserved snapshot image itself for business continuity in case of primary storage failure or for off-host processing.

The following example procedures illustrate how to create a full image snapshot and periodic point-in-time copies for recovery. With multiple point-in-time copies to choose from, you can select the point-in-time to which you want to recover with relative precision.

Setting up a full image snapshot and incremental point-in-time copies

To set up the initial configuration for a full image snapshot and incremental point-in-time copies, set up storage for the point-in-time copies that will be configured over time.

In the example procedures, *disk1*, *disk2*, ..., *diskN* are the LUNs configured on tier 1 storage for application data. A subset of these LUNs *logdisk1*, *logdisk2*, ..., *logdiskN*, will be used to configure DCO. Disks *sdisk1*, *sdisk2*, ..., *sdiskN* are disks from tier 2 storage.

Note: If you have an enclosure or disk array with storage that is backed by write cache, Symantec recommends that you use the same set of LUNS for DCO and for data volume.

If no logdisks are specified by default, Veritas Volume Manager tries to allocate the DCO from the same LUNs used for the data volumes.

See [Figure 6-4](#) on page 77.

You will need to make sure your cache is big enough for the multiple copies with multiple changes. The following guidelines may be useful for estimating your requirements.

To determine your storage requirements, use the following:

Table 7-1 Storage requirements

S_p	Represents the storage requirement for the primary volume
S_b	Represents the storage requirement for the primary break-off snapshot.
N_c	Represents the number of point-in-time copies to be maintained.
S_c	Represents the average size of the changes that occur in an interval before the snapshot is taken.
S_t	Represents the total storage requirement.

The total storage requirement for management of multiple point-in-time copies can be roughly calculated as:

$$S_b = S_p$$

$$S_t = S_b + N_c * S_c$$

To determine the size of the cache volume, use the following:

Table 7-2 Cache volume requirements

N_c	Represents the number of point-in-time copies to be maintained.
S_c	Represents the average size of the changes that occur in an interval .
R_c	Represents the region-size for cache-object.
S_t	Represents the total storage requirement.

The size of cache-volume to be configured can be calculated as:

$$N_c * S_c * R_c$$

This equation assumes that the application IO size granularity is smaller than cache-object region-size by factor of at most R_c

To configure the initial setup for a full image snapshot and incremental point-in-time copies

- 1 If the primary application storage is already configured for snapshots, that is, the DCO is already attached for the primary volume, go to step 2.

If not, configure the primary volumes and prepare them for snapshots.

For example:

```
# vxassist -g appdg make appvol 10T <disk1 disk2 ... diskN >  
# vxsnap -g appdg prepare appvol
```

- 2 Configure your storage for a full image snapshot and the point-in-time copies to be created periodically, in a similar manner to the way you configured your primary volumes. Veritas Storage Foundation does not need allocation from the same (tier 1) storage to configure point-in-time copies. For instance, the allocation could come from tier 2 storage.
- 3 Configure a snapshot volume from tier 2 storage to use as primary snapshot of the primary volume.

```
# vxassist -g appdg make snap-appvol 10T <sdisk1 sdisk2 ... sdiskN >  
# vxsnap -g appdg prepare snap-appvol \  
<alloc=slogdisk1, slogdisk2, ...slogdiskN>
```

- 4 Establish the relationship between primary and snapshot volumes and wait for synchronization of the snapshot to complete.

```
# vxsnap -g appdg make source=appvol/snapvol=snap-appvol/sync=yes  
# vxsnap -g appdg syncwait snap-appvol
```

- 5 Create a volume to use for the cache volume.

```
# vxassist -g appdg make cachevol 1G layout=mirror \  
init=active disk16 disk17
```

- 6 Configure a shared cache object in the disk group from tier 2 storage to use for point-in-time copies created at regular intervals.

```
# vxmake -g appdg cache snapcache cachevolname=cachevol
```

Start the cache object.

```
# vxcache -g appdg start snapcache
```

You now have an initial setup in place to create regular point-in-time copies.

Refreshing point-in-time copies

After configuring your volumes for snapshots, you can periodically invoke a script with steps similar to following to create point-in-time copies at regular intervals.

To identify snapshot age

- ◆ To find the oldest and the most recent snapshots, use the creation time of the snapshots. You can use either of the following commands:

- Use the following command and find the SNAPDATE of snapshot volume.

```
# vxsnap -g appdg list appvol
```

- Use the following command:

```
# vxprint -g appdg -m snapobject_name | grep creation_time
```

Where the *snapobject-name* is *appvol-snp*, *appvol-snp1* *appvol-snpN*.

To refresh the primary snapshot

- ◆ Refresh the primary snapshot from the primary volume.

```
# vxsnap -g appdg refresh snap-appvol source=appvol
```

To create cascaded snapshot of the refreshed snapshot volume

- ◆ Create a cascaded snapshot of the refreshed snapshot volume.

```
# vxsnap -g appdg make source=snap-appvol/new=sosnap-\
appvol${NEW_SNAP_IDX}/cache=snapcache/infrontof=snap-appvol
```

To remove the oldest point-in-time copy

- ◆ If the limit on number of point-in-time copies is reached, remove the oldest point-in-time copy.

```
# vxedit -g appdg -rf rm sosnap-appvol${ OLDEST_SNAP_IDX }
```

Recovering from logical corruption

You can use the preserved snapshot image in case of primary storage corruption. You must identify the most recent snapshot which is not affected by the logical corruption.

To identify the most recent valid snapshot

- 1 For each snapshot, starting from the most recent to the oldest, verify the snapshot image. Create a space-optimized snapshot of the point-in-time copy to generate a synthetic replica of the point-in-time image.

```
# vxsnap -g appdg make source=sosnapappvol${  
CURIDX}/new=syn-appvol/cache=snapcache/sync=no
```

- 2 Mount the synthetic replica and verify the data.

If a synthetic replica is corrupted, proceed to 3.

When you identify a synthetic replica that is not corrupted, you can proceed to the recovery steps.

See “[To recover from logical corruption](#)” on page 92.

- 3 Unmount the synthetic replica, remove it and go back to verify the next most recent point-in-time copy. Use the following command to dissociate the synthetic replica and remove it:

```
# vxsnap -g appdg dis syn-appvol  
# vxedit -g appdg -rf rm syn-appvol
```

When you find the most recent uncorrupted snapshot, use it to restore the primary volume.

To recover from logical corruption

- 1 If the application is running on the primary volume, stop the application.
- 2 Unmount the application volume.
- 3 Restore the primary volume from the synthetic replica.

```
# vxsnap -g appdg restore appvol source=syn-appvol
```

- 4 Resume the application:
 - Mount the primary volume.
 - Verify the content of the primary volume.
 - Restart the application.

Off-host processing using refreshed snapshot images

Preserved point-in-time images can also be used to perform off-host processing. Using preserved point-in-time images for this purpose requires that the storage used for creating the snapshots must be:

- Accessible from the application host
- Accessible from the off-host processing host
- Split into a separate diskgroup

To split the snapshot storage into a separate disk group

- ◆ Split the snapshot storage into a separate disk group.

```
# vxdg split appdg snapdg snap-appvol
```

The *snapdg* disk group can optionally be deported from the application host using the `vxdg deport` command and imported on another host using the `vxdg import` command to continue to perform off-host processing.

To refresh snapshot images for off-host processing

- 1 Deport the *snapdg* disk group from the off-host processing host.

```
# vxdg deport snapdg
```

- 2 Import the *snapdg* disk group on the application host.

```
# vxdg import snapdg
```

- 3 On the application host, join the *snapdg* disk group to *appdg*.

```
# vxdg join snapdg appdg
```

After this step, you can proceed with the steps for managing point-in-time copies.

See [“Refreshing point-in-time copies”](#) on page 91.

Online database backups

Online backup of a database can be implemented by configuring either the primary host or a dedicated separate host to perform the backup operation on snapshot mirrors of the primary host’s database.

Two backup methods are described in the following sections:

- See [“Making a backup of an online database on the same host”](#) on page 94.

- See “[Making an off-host backup of an online database](#)” on page 105.

Note: All commands require superuser (`root`) or equivalent privileges, except where it is explicitly stated that a command must be run by the database administrator.

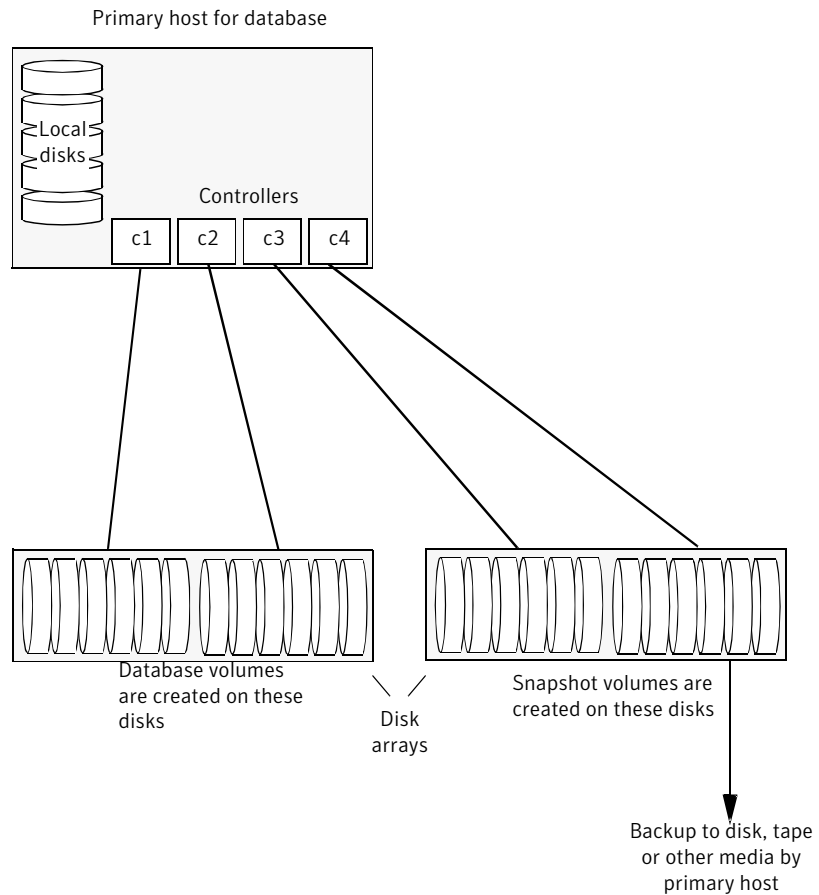
For more information about using snapshots to back up Oracle databases, see the *Veritas Storage Foundation™: Storage and Availability Management for Oracle Databases*.

Note: The sample database scripts in the following procedures are not supported by Symantec, and are provided for informational use only. You can purchase customization of the environment through Veritas Vpro Consulting Services.

Making a backup of an online database on the same host

Figure 7-1 shows an example with two primary database volumes to be backed up, `database_vol` and `dbase_logs`, which are configured on disks attached to controllers `c1` and `c2`, and the snapshots to be created on disks attached to controllers `c3` and `c4`.

Figure 7-1 Example system configuration for database backup on the primary host



Note: It is assumed that you have already prepared the volumes containing the file systems for the datafiles to be backed up as described in the example.

To make an online database backup

- Prepare the snapshot, either full-sized or space-optimized.
 Depending on the application, space-optimized snapshots typically require 10% of the disk space that is required for full-sized instant snapshots.
 See [“Preparing a full-sized instant snapshot for a backup”](#) on page 96.
 See [“Preparing a space-optimized snapshot for a database backup”](#) on page 98.

- Create snapshot mirrors for volumes containing VxFS file systems for database files to be backed up.
- Make the online database backup.
See “[Backing up a DB2 database on the same host](#)” on page 100.
See “[Backing up a Sybase database on the same host](#)” on page 102.
- Resynchronize a volume if required.
See “[Resynchronizing a volume](#)” on page 104.

Some scenarios where full-instant snapshot works better are:

- Off host processing is planned for a databases backup.
- If a space-optimized snapshot is taken for longer duration and modified frequently then it is not much different than the full-snapshot. So, for performance reason full-snapshot will be preferred,

Preparing a full-sized instant snapshot for a backup

You can use a full-sized instant snapshot for your online or off-host database backup.

Warning: To avoid data inconsistencies, do not use the same snapshot with different point-in-time copy applications. If you require snapshot mirrors for more than one application, configure at least one snapshot mirror that is dedicated to each application.

To make a full-sized instant snapshot for a backup of an online database on the same host

- 1 Use the following commands to add one or more snapshot plexes to the volume, and to make a full-sized break-off snapshot, *snapvol*, of the tablespace volume by breaking off these plexes:

```
# vxsnap -g database_dg addmir database_vol [nmirror=N] \  
[alloc=storage_attributes]  
# vxsnap -g database_dg make \  
source=database_vol/newvol=snapvol[/nmirror=N] \  
[alloc=storage_attributes]
```

By default, one snapshot plex is added unless you specify a number using the `nmirror` attribute. For a backup, you should usually only require one plex. You can specify storage attributes (such as a list of disks) to determine where the plexes are created.

You can specify at least N number of disks if the specified number of mirrors is N .

- 2 If the volume layout does not support plex break-off, prepare an empty volume for the snapshot. Create a full-sized instant snapshot for an original volume that does not contain any spare plexes, you can use an empty volume with the required degree of redundancy, and with the same size and same region size as the original volume.

Use the `vxprint` command on the original volume to find the required size for the snapshot volume.

```
# LEN=`vxprint [-g diskgroup] -F%len database_vol`
```

Note: The command shown in this and subsequent steps assumes that you are using a Bourne-type shell such as `sh`, `ksh` or `bash`. You may need to modify the command for other shells such as `csh` or `tcsh`. These steps are valid only for instant snap DCOs.

- 3 Use the `vxprint` command on the original volume to discover the name of its DCO:

```
# DCONAME=`vxprint [-g diskgroup] -F%dco_name database_vol`
```

- 4 Use the `vxprint` command on the DCO to discover its region size (in blocks):

```
# RSZ=`vxprint [-g diskgroup] -F%regionsz $DCONAME`
```

- 5 Use the `vxassist` command to create a volume, `snapvol`, of the required size and redundancy, together with a instant snap DCO volume with the correct region size:

```
# vxassist [-g diskgroup] make snapvol $LEN \
  [layout=mirror nmirror=number] logtype=dco drl=no \
  dconversion=20 [ndcomirror=number] regionsz=$RSZ \
  init=active [storage_attributes]
```

It is recommended that you specify the same number of DCO mirrors (`ndcomirror`) as the number of mirrors in the volume (`nmirror`). The `init=active` attribute is used to make the volume available immediately. You can use storage attributes to specify which disks should be used for the volume.

As an alternative to creating the snapshot volume and its DCO volume in a single step, you can first create the volume, and then prepare it for instant snapshot operations as shown here:

```
# vxassist [-g diskgroup] make snapvol $LEN \
  [layout=mirror nmirror=number] init=active \
  [storage_attributes]
# vxsnap [-g diskgroup] prepare snapvol [ndcomirs=number] \
  regionsz=$RSZ [storage_attributes]
```

- 6 Use the following command to create the snapshot:

```
# vxsnap -g database_dg make source=database_vol/snapvol=snapvol
```

If a database spans more than one volume, specify all the volumes and their snapshot volumes as separate tuples on the same line, for example:

```
# vxsnap -g database_dg make source=database_vol1/snapvol=snapvol1 \
  source=database_vol2/newvol=snapvol2 \
  source=database_vol3/snapvol=snapvol3
```

When you are ready to make a backup, proceed to make a backup of an online database on the same host.

Preparing a space-optimized snapshot for a database backup

If a snapshot volume is to be used on the same host, and will not be moved to another host, you can use space-optimized instant snapshots rather than full-sized instant snapshots. Depending on the application, space-optimized snapshots typically require 10% of the disk space that is required for full-sized instant snapshots.

To prepare a space-optimized snapshot for a backup of an online database

- 1 To save disk space, use the following command to create a space-optimized snapshot:

```
# vxsnap -g database_dg make \
  source=database_vol/newvol=snapvol/cache=snapcache
```

The argument *snapcache* is the name of a pre-existing cache that you have created in the disk group for use with space-optimized snapshots.

- 2 Decide on the following characteristics that you want to allocate to the cache volume that underlies the cache object:

- The size of the cache volume should be sufficient to record changes to the parent volumes during the interval between snapshot refreshes. A suggested value is 10% of the total size of the parent volumes for a refresh interval of 24 hours.
 - If redundancy is a desired characteristic of the cache volume, it should be mirrored. This increases the space that is required for the cache volume in proportion to the number of mirrors that it has.
 - If the cache volume is mirrored, space is required on at least as many disks as it has mirrors. These disks should not be shared with the disks used for the parent volumes. The disks should also be chosen to avoid impacting I/O performance for critical volumes, or hindering disk group split and join operations.
- 3 Having decided on its characteristics, use the `vxassist` command to create the volume that is to be used for the cache volume. The following example creates a mirrored cache volume, `cachevol`, with size 1GB in the disk group, `database_dg`, on the disks `disk16` and `disk17`:

```
# vxassist -g database_dg make cachevol 1g layout=mirror \  
init=active disk16 disk17
```

The attribute `init=active` is specified to make the cache volume immediately available for use.

- 4 Use the `vxmake cache` command to create a cache object on top of the cache volume that you created in the previous step:

```
# vxmake [-g diskgroup] cache cache_object \  
cachevolname=cachevol [regionsize=size] [autogrow=on] \  
[highwatermark=hwmk] [autogrowby=agbvalue] \  
[maxautogrow=maxagbvalue]
```

If you specify the region size, it must be a power of 2, and be greater than or equal to 16KB (16k). If not specified, the region size of the cache is set to 64KB.

Note: All space-optimized snapshots that share the cache must have a region size that is equal to or an integer multiple of the region size set on the cache. Snapshot creation also fails if the original volume's region size is smaller than the cache's region size.

If the cache is not allowed to grow in size as required, specify `autogrow=off`. By default, the ability to automatically grow the cache is turned on.

In the following example, the cache object, *cache_object*, is created over the cache volume, *cachevol*, the region size of the cache is set to 32KB, and the *autogrow* feature is enabled:

```
# vxmake -g database_dg cache cache_object cachevolname=cachevol \  
    regionsize=32k autogrow=on
```

- 5 Having created the cache object, use the following command to enable it:

```
vxcache [-g diskgroup] start cache_object
```

For example, start the cache object *cache_object*:

```
# vxcache -g database_dg start cache_object
```

- 6 Create a space-optimized snapshot with your cache object.

```
# vxsnap -g database_dg make \  
source=database_vol1/newvol=snapvol1/cache=cache_object
```

- 7 If several space-optimized snapshots are to be created at the same time, these can all specify the same cache object as shown in this example:

```
# vxsnap -g database_dg make \  
    source=database_vol1/newvol=snapvol1/cache=cache_object \  
    source=database_vol2/newvol=snapvol2/cache=cache_object \  
    source=database_vol3/newvol=snapvol3/cache=cache_object
```

Note: This step sets up the snapshot volumes, prepares for the backup cycle, and starts tracking changes to the original volumes.

When you are ready to make a backup, proceed to make a backup of an online database on the same host

Backing up a DB2 database on the same host

You can make an online backup of your DB2 database.

To make a backup of an online DB2 database on the same host

- 1 If the volumes to be backed up contain database tables in file systems, suspend updates to the volumes. DB2 provides the `write suspend` command to temporarily suspend I/O activity for a database. As the DB2 database administrator, use a script such as that shown in the example to suspend I/O for a DB2 database.

```
#!/bin/ksh
#
# script: backup_start.sh
#
# Sample script to suspend I/O for a DB2 database.
#
# Note: To recover a database using backups of snapshots, the database
# must be in LOGRETAIN mode.

db2 <<!
connect to database
set write suspend for database
quit
!
```

Note that to allow recovery from any backups taken from snapshots, the database must be in LOGRETAIN RECOVERY mode.

- 2 Refresh the contents of the snapshot volumes from the original volume using the following command:

```
# vxsnap -g database_dg refresh snapvol1 source=database_vol1 \  
[snapvol2 source=database_vol2]...
```

For example, to refresh the snapshots *snapvol1*, *snapvol2* and *snapvol3*:

```
# vxsnap -g database_dg refresh snapvol1 source=database_vol1 \  
snapvol2 source=database_vol2 snapvol3 source=database_vol3
```

- 3 If you have temporarily suspended updates to volumes, release all the tablespaces or databases from suspend mode. As the DB2 database administrator, use a script such as that shown in the example to resume I/O for a DB2 database.

```
#!/bin/ksh
#
# script: backup_end.sh
#
# Sample script to resume I/O for a DB2 database.
#

db2 <<!
connect to database
set write resume for database
```

```
quit  
!
```

- 4 Back up the snapshot volume. If you need to remount the file system in the volume to back it up, first run `fsck` on the volume. The following are sample commands for checking and mounting a file system:

```
# fsck -F vxfs /dev/vx/rdisk/database_dg/snapvol  
# mount -F vxfs /dev/vx/dsk/database_dg/snapvol  
  mount_point
```

Back up the file system at this point using a command such as `bpbackup` in Symantec NetBackup. After the backup is complete, use the following command to unmount the file system.

```
# umount mount_point
```

- 5 Repeat steps in this procedure each time that you need to back up the volume.

Backing up a Sybase database on the same host

You can make an online backup of your Sybase database.

To make a backup of an online Sybase database on the same host

- 1 If the volumes to be backed up contain database tables in file systems, suspend updates to the volumes. Sybase ASE from version 12.0 onward provides the Quiesce feature to allow temporary suspension of writes to a database. As the Sybase database administrator, put the database in quiesce mode by using a script such as that shown in the example.

```
#!/bin/ksh
#
# script: backup_start.sh
#
# Sample script to quiesce example Sybase ASE database.
#
# Note: The "for external dump" clause was introduced in Sybase
# ASE 12.5 to allow a snapshot database to be rolled forward.
# See the Sybase ASE 12.5 documentation for more information.

isql -Usa -Ppassword -SFMR <<!
quiesce database tag hold database1[, database2]... [for external dump]
go
quit
!
```

- 2 Refresh the contents of the snapshot volumes from the original volume using the following command:

```
# vxsnap -g database_dg refresh snapvol source=database_vol \
[snapvol2 source=database_vol2]...
```

For example, to refresh the snapshots *snapvol1*, *snapvol2* and *snapvol3*:

```
# vxsnap -g database_dg refresh snapvol1 source=database_vol1 \
snapvol2 source=database_vol2 snapvol3 source=database_vol3
```

- 3 If you have temporarily suspended updates to volumes, release all the tablespaces or databases from quiesce mode.

As the Sybase database administrator, release the database from quiesce mode using a script such as that shown in the example.

```
#!/bin/ksh
#
# script: backup_end.sh
#
# Sample script to release example Sybase ASE database from
# quiesce mode.

isql -Usa -Ppassword -SFMR <<!
quiesce database tag release
go
quit
!
```

- 4 Back up the snapshot volume. If you need to remount the file system in the volume to back it up, first run `fsck` on the volume. The following are sample commands for checking and mounting a file system:

```
# fsck -F vxfs /dev/vx/rdisk/database_dg/snapvol
# mount -F vxfs /dev/vx/dsk/database_dg/snapvol
  mount_point
```

Back up the file system at this point using a command such as `bpbackup` in Symantec NetBackup. After the backup is complete, use the following command to unmount the file system.

```
# umount mount_point
```

- 5 Repeat steps in this procedure each time that you need to back up the volume.

Resynchronizing a volume

In some instances, such as recovering the contents of a corrupted volume, it may be useful to resynchronize a volume from its snapshot volume (which is used as a hot standby).

To resynchronize a volume from its snapshot volume

◆ Enter:

```
# vxsnap -g diskgroup restore database_vol source=snapvol \  
destroy=yes|no
```

The `destroy` attribute specifies whether the plexes of the snapshot volume are to be reattached to the original volume. For example, to resynchronize the volume `database_vol` from its snapshot volume `snapvol` without removing the snapshot volume:

```
# vxsnap -g database_dg restore database_vol \  
source=snapvol destroy=no
```

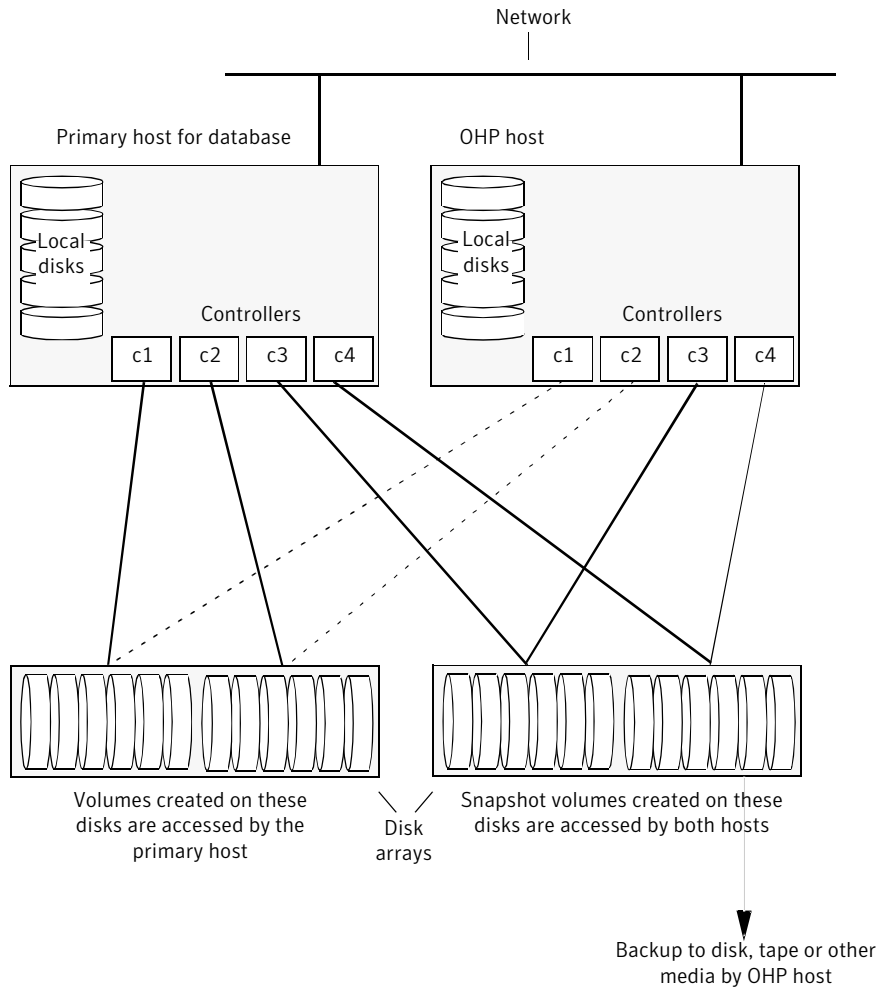
Note: You must shut down the database and unmount the file system that is configured on the original volume before attempting to resynchronize its contents from a snapshot.

Making an off-host backup of an online database

Figure 7-2 shows an example of two primary database volumes to be backed up, `database_vol` and `dbase_logs`, which are configured on disks attached to controllers `c1` and `c2`, and the snapshots to be created on disks attached to controllers `c3` and `c4`.

There is no requirement for the off-host processing host to have access to the disks that contain the primary database volumes.

Figure 7-2 Example system configuration for off-host database backup



If the database is configured on volumes in a cluster-shareable disk group, it is assumed that the primary host for the database is the master node for the cluster. However, if the primary host is not also the master node, most VxVM operations on shared disk groups are best performed on the master node.

The procedure in this section is designed to minimize copy-on-write operations that can impact system performance. You can also implement this procedure on a single host by omitting steps 5 through [To make an off-host backup of an online Sybase database](#) and 10 through 13 that split, deport, reimport and rejoin the snapshot disk group.

To make an off-host database backup of an online database:

- Prepare the full-sized snapshot for backing up.
- Make the off-host database backup using the steps for DB2 or Sybase.
See [“Making an off-host backup of an online DB2 database”](#) on page 108.
See [“Making an off-host backup of an online Sybase database”](#) on page 113.
- Resynchronize a volume if required.
See [“Resynchronizing a volume”](#) on page 111.

Making an off-host backup of an online DB2 database

To start an off-host backup of an online DB2 database

- 1 On the primary host, add one or more snapshot plexes to the volume using this command:

```
# vxsnap -g database_dg addmir database_vol [nmirror=N] \  
[alloc=storage_attributes]
```

By default, one snapshot plex is added unless you specify a number using the `nmirror` attribute. For a backup, you should usually only require one plex. You can specify storage attributes (such as a list of disks) to determine where the plexes are created.

- 2 Suspend updates to the volumes. DB2 provides the `write suspend` command to temporarily suspend I/O activity for a database. As the DB2 database administrator, use a script such as that shown in the example to suspend I/O for a DB2 database.

```
#!/bin/ksh  
#  
# script: backup_start.sh  
#  
# Sample script to suspend I/O for a DB2 database.  
#  
# Note: To recover a database using backups of snapshots,  
# the database must be in LOGRETAIN mode.  
  
db2 <<!  
connect to database  
set write suspend for database  
quit  
!
```

Note that to allow recovery from any backups taken from snapshots, the database must be in LOGRETAIN RECOVERY mode.

- 3 Make a full-sized snapshot, *snapvol*, of the tablespace volume by breaking off the plexes that you added in step 1 from the original volume:

```
# vxsnap -g database_dg make \  
    source=database_vol/newvol=snapvol/nmirror=N \  
    [alloc=storage_attributes]
```

The `nmirror` attribute specifies the number of mirrors, *N*, in the snapshot volume.

If a database spans more than one volume, specify all the volumes and their snapshot volumes as separate tuples on the same line, for example:

```
# vxsnap -g database_dg \  
    make source=database_vol1/snapvol=snapvol1 \  
    source=database_vol2/snapvol=snapvol2 \  
    source=database_vol3/snapvol=snapvol3 alloc=ctlr:c3,ctlr:c4
```

- 4 This step sets up the snapshot volumes ready for the backup cycle, and starts tracking changes to the original volumes.
- 5 Release all the tablespaces or databases from suspend mode. As the DB2 database administrator, use a script such as that shown in the example. to resume I/O for a DB2 database.

```
#!/bin/ksh  
#  
# script: backup_end.sh  
#  
# Sample script to resume I/O for a DB2 database.  
#  
  
db2 <<!  
connect to database  
set write resume for database  
quit  
!
```

- 6 On the primary host, split the disks containing the snapshot volumes into a separate disk group, *snapvoldg*, from the original disk group, *database_dg* using the following command:

```
# vxdg split database_dg snapvoldg snapvol ...
```

- 7 On the primary host, deport the snapshot volume's disk group using the following command:

```
# vxdg deport snapvoldg
```

- 8 On the OHP host where the backup is to be performed, use the following command to import the snapshot volume's disk group:

```
# vxdg import snapvoldg
```

- 9 VxVM recovers the volumes automatically after the disk group import unless it is set to not recover automatically. Check if the snapshot volume is initially disabled and not recovered following the split.

If a volume is in the DISABLED state, use the following command on the off-host processing host to recover and restart the snapshot volume:

```
# vxrecover -g snapvoldg -m snapvol ...
```

- 10 On the off-host processing host, back up the snapshot volumes. If you need to remount the file system in the volume to back it up, first run `fsck` on the volumes. The following are sample commands for checking and mounting a file system:

```
# fsck -F vxfs /dev/vx/rdisk/snapvoldg/snapvol  
# mount -F vxfs /dev/vx/dsk/snapvoldg/snapvol mount_point
```

Back up the file system using a command such as `bpbackup` in Symantec NetBackup. After the backup is complete, use the following command to unmount the file system.

```
# umount mount_point
```

- 11 On the OHP host, use the following command to deport the snapshot volume's disk group:

```
# vxdg deport snapvoldg
```

- 12 On the primary host, re-import the snapshot volume's disk group using the following command:

```
# vxdg [-s] import snapvoldg
```

Note: Specify the `-s` option if you are reimporting the disk group to be rejoined with a shared disk group in a cluster.

- 13 On the primary host, use the following command to rejoin the snapshot volume's disk group with the original volume's disk group:

```
# vxdg join snapvoldg database_dg
```

- 14 VxVM will recover the volumes automatically after the join unless it is set not to recover automatically. Check if the snapshot volumes are initially disabled and not recovered following the join.

If a volume is in the DISABLED state, use the following command on the primary host to recover and restart the snapshot volume:

```
# vxrecover -g database_dg -m snapvol
```

- 15 On the primary host, reattach the snapshot volumes to their original volume. Reattach of multiple volumes in a single command is not supported. Split it into multiple invocations:

```
# vxsnap -g database_dg reattach snapvol1 source=database_vol1
# vxsnap -g database_dg reattach snapvol2 source=database_vol2
# vxsnap -g database_dg reattach snapvol3 source=database_vol3
```

Use the following commands to wait for synchronization to complete:

```
# vxsnap -g database_dg snapwait database_vol1
# vxsnap -g database_dg snapwait database_vol2
# vxsnap -g database_dg snapwait database_vol3
```

For example, to reattach the snapshot volumes `snapvol1`, `snapvol2` and `snapvol3`:

```
# vxsnap -g database_dg reattach snapvol1 source=database_vol1 \
    snapvol2 source=database_vol2 snapvol3 source=database_vol3
```

While the reattached plexes are being resynchronized from the data in the parent volume, they remain in the `SNAPTMP` state. After resynchronization is complete, the plexes are placed in the `SNAPDONE` state. You can use the `vxsnap print` command to check on the progress of synchronization.

Repeat steps 2 through 15 each time that you need to back up the volume.

Resynchronizing a volume

In some instances, such as recovering the contents of a corrupted volume, it may be useful to resynchronize a volume from its snapshot volume (which is used as a hot standby).

To resynchronize a volume

- ◆ Use the following command syntax:

```
vxsnap -g database_dg restore database_vol source=snapvol \  
destroy=yes|no
```

The `destroy` attribute specifies whether the plexes of the snapshot volume are to be reattached to the original volume.

For example, to resynchronize the volume `database_vol` from its snapshot volume `snapvol` without removing the snapshot volume:

```
# vxsnap -g database_dg restore database_vol \  
source=snapvol destroy=no
```

Note: You must shut down the database and unmount the file system that is configured on the original volume before attempting to resynchronize its contents from a snapshot.

Making an off-host backup of an online Sybase database

To make an off-host backup of an online Sybase database

- 1 On the primary host, add one or more snapshot plexes to the volume using this command:

```
# vxsnap -g database_dg addmir database_vol [nmirror=N] \  
[alloc=storage_attributes]
```

By default, one snapshot plex is added unless you specify a number using the `nmirror` attribute. For a backup, you should usually only require one plex. You can specify storage attributes (such as a list of disks) to determine where the plexes are created.

- 2 Suspend updates to the volumes. As the Sybase database administrator, put the database in quiesce mode by using a script such as that shown in the example.

```
#!/bin/ksh  
#  
# script: backup_start.sh  
#  
# Sample script to quiesce example Sybase ASE database.  
#  
# Note: The "for external dump" clause was introduced in Sybase  
# ASE 12.5 to allow a snapshot database to be rolled forward.  
# See the Sybase ASE 12.5 documentation for more information.  
  
isql -Usa -Ppassword -SFMR <<!  
quiesce database tag hold database1[, database2]... [for  
external dump]  
go  
quit  
!
```

- 3 Use the following command to make a full-sized snapshot, *snapvol*, of the tablespace volume by breaking off the plexes that you added in step 1 from the original volume:

```
# vxsnap -g database_dg make \  
    source=database_vol/newvol=snapvol/nmirror=N \  
    [alloc=storage_attributes]
```

The `nmirror` attribute specifies the number of mirrors, *N*, in the snapshot volume.

If a database spans more than one volume, specify all the volumes and their snapshot volumes as separate tuples on the same line, for example:

```
# vxsnap -g database_dg make source=database_vol1/snapvol=snapvol1 \  
    source=database_vol2/snapvol=snapvol2 \  
    source=database_vol3/snapvol=snapvol3 alloc=ctlr:c3,ctlr:c4
```

This step sets up the snapshot volumes ready for the backup cycle, and starts tracking changes to the original volumes.

- 4 Release all the tablespaces or databases from quiesce mode. As the Sybase database administrator, release the database from quiesce mode using a script such as that shown in the example.

```
#!/bin/ksh  
#  
# script: backup_end.sh  
#  
# Sample script to release example Sybase ASE database from quiesce  
# mode.  
  
isql -Uasa -Ppassword -SFMR <<!  
quiesce database tag release  
go  
quit  
!
```

- 5 On the primary host, split the disks containing the snapshot volumes into a separate disk group, *snapvoldg*, from the original disk group, *database_dg* using the following command:

```
# vxdg split database_dg snapvoldg snapvol ...
```

- 6 On the primary host, deport the snapshot volume's disk group using the following command:

```
# vxdg deport snapvoldg
```

- 7 On the OHP host where the backup is to be performed, use the following command to import the snapshot volume's disk group:

```
# vxdg import snapvoldg
```

- 8 VxVM will recover the volumes automatically after the disk group import unless it is set to not recover automatically. Check if the snapshot volume is initially disabled and not recovered following the split.

If a volume is in the DISABLED state, use the following command on the OHP host to recover and restart the snapshot volume:

```
# vxrecover -g snapvoldg -m snapvol ...
```

- 9 On the OHP host, back up the snapshot volumes. If you need to remount the file system in the volume to back it up, first run `fsck` on the volumes. The following are sample commands for checking and mounting a file system:

```
# fsck -F vxfs /dev/vx/rdisk/snapvoldg/snapvol  
# mount -F vxfs /dev/vx/dsk/snapvoldg/snapvol mount_point
```

Back up the file system using a command such as `bbbackup` in Symantec NetBackup. After the backup is complete, use the following command to unmount the file system.

```
# umount mount_point
```

- 10 On the OHP host, use the following command to deport the snapshot volume's disk group:

```
# vxdg deport snapvoldg
```

- 11 On the primary host, re-import the snapshot volume's disk group using the following command:

```
# vxdg [-s] import snapvoldg
```

Note: Specify the `-s` option if you are reimporting the disk group to be rejoined with a shared disk group in a cluster.

- 12 On the primary host, use the following command to rejoin the snapshot volume's disk group with the original volume's disk group:

```
# vxdg join snapvoldg database_dg
```

- 13 VxVM will recover the volumes automatically after the join unless it is set not to recover automatically. Check if the snapshot volumes are initially disabled and not recovered following the join.

If a volume is in the DISABLED state, use the following command on the primary host to recover and restart the snapshot volume:

```
# vxrecover -g database_dg -m snapvol
```

- 14 On the primary host, reattach the snapshot volumes to their original volume using the following command:

```
# vxsnap -g database_dg reattach snapvol source=database_vol \  
[snapvol2 source=database_vol2]...
```

For example, to reattach the snapshot volumes *snapvol1*, *snapvol2* and *snapvol3*:

```
# vxsnap -g database_dg reattach snapvol1 source=database_vol1 \  
snapvol2 source=database_vol2 snapvol3 source=database_vol3
```

While the reattached plexes are being resynchronized from the data in the parent volume, they remain in the `SNAPTMP` state. After resynchronization is complete, the plexes are placed in the `SNAPDONE` state. You can use the `vxsnap print` command to check on the progress of synchronization.

Repeat steps 2 through 14 each time that you need to back up the volume.

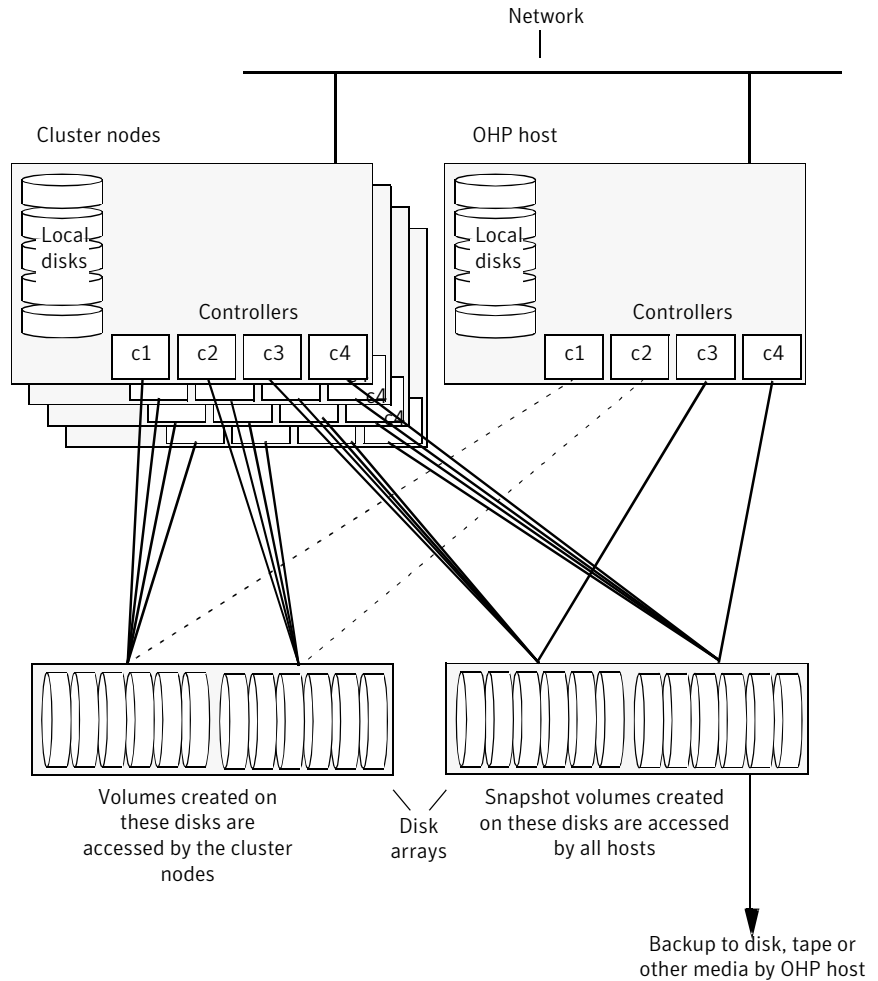
Backing up on an off-host cluster file system

Storage Foundation Cluster File System High Availability (SFCFSHA) allows cluster nodes to share access to the same file system. SFCFSHA is especially useful for sharing read-intensive data between cluster nodes.

Off-host backup of cluster file systems may be implemented by taking a snapshot of the volume containing the file system and performing the backup operation on a separate host.

Figure 7-3 shows an example where the primary volume that contains the file system to be backed up is configured on disks attached to controllers *c1* and *c2*, and the snapshots are to be created on disks attached to controllers *c3* and *c4*.

Figure 7-3 System configuration for off-host file system backup scenarios



To set up an off-host cluster file system backup:

- Mount a VxFS file system for shared access by the nodes of a cluster.
 See “[Mounting a file system for shared access](#)” on page 118.
- Prepare a snapshot of the mounted file system with shared access.
- Back up a snapshot of a mounted file system with shared access
 See “[Backing up a snapshot of a mounted file system with shared access](#)”
 on page 120.
- All commands require superuser (`root`) or equivalent privileges.

Mounting a file system for shared access

To mount a VxFS file system for shared access, use the following command on each cluster node where required:

```
# mount -F vxfs -o cluster /dev/vx/dsk/database_dg/database_vol  
    mount_point
```

For example, to mount the volume *database_vol* in the disk group *database_dg* for shared access on the mount point, */mnt_pnt*:

```
# mount -F vxfs -o cluster /dev/vx/dsk/database_dg/database_vol /mnt_pnt
```

Preparing a snapshot of a mounted file system with shared access

You must use a full-sized snapshot for your off-host backup.

Warning: To avoid data inconsistencies, do not use the same snapshot with different point-in-time copy applications. If you require snapshot mirrors for more than one application, configure at least one snapshot mirror that is dedicated to each application.

To prepare to back up a snapshot of a mounted file system which has shared access

- 1 On the master node, use the following command to make a full-sized snapshot, *snapvol*, of the volume containing the file system by breaking off plexes from the original volume:

```
# vxsnap -g database_dg make \  
    source=database_vol/newvol=snapvol/nmirror=N
```

The `nmirror` attribute specifies the number of mirrors, *N*, in the snapshot volume.

For example, to take a snapshot of the volume *database_vol* in the shared disk group *exempldg*:

```
# vxsnap -g database_dg make source=database_vol/newvol=snapvol\  
/nmirror=1
```

If the volume does not have any available plexes, or its layout does not support plex break-off, prepare an empty volume for the snapshot.

- 2 Use the `vxprint` command on the original volume to find the required size for the snapshot volume.

```
# LEN=`vxprint [-g database_dg] -F%len database_vol`
```

Note: The command shown in this and subsequent steps assumes that you are using a Bourne-type shell such as `sh`, `ksh` or `bash`. You may need to modify the command for other shells such as `csh` or `tcsh`. These steps are valid only for instant snap DCOs.

- 3 Use the `vxprint` command on the original volume to discover the name of its DCO:

```
# DCONAME=`vxprint [-g database_dg] -F%dco_name database_vol`
```

- 4 Use the `vxprint` command on the DCO to discover its region size (in blocks):

```
# RSZ=`vxprint [-g database_dg] -F%regionsz $DCONAME`
```

- 5 Use the `vxassist` command to create a volume, `snapvol`, of the required size and redundancy, together with an instant snap DCO volume with the correct region size:

```
# vxassist [-g database_dg] make snapvol $LEN \  
  [layout=mirror nmirror=number] logtype=dco drl=no \  
  dconversion=20 [ndcomirror=number] regionsz=$RSZ \  
  init=active [storage_attributes]
```

It is recommended that you specify the same number of DCO mirrors (`ndcomirror`) as the number of mirrors in the volume (`nmirror`). The `init=active` attribute is used to make the volume available immediately. You can use storage attributes to specify which disks should be used for the volume.

As an alternative to creating the snapshot volume and its DCO volume in a single step, you can first create the volume, and then prepare it for instant snapshot operations as shown here:

```
# vxassist [-g database_dg] make snapvol $LEN \  
  [layout=mirror nmirror=number] init=active \  
  [storage_attributes]  
# vxsnap [-g database_dg] prepare snapvol [ndcomirs=number] \  
  regionsz=$RSZ [storage_attributes]
```

- 6 Then use the following command to create the snapshot:

```
# vxsnap -g database_dg make source=database_dg/snapvol=snapvol
```

Note: This step actually takes the snapshot and sets up the snapshot volumes ready for the backup cycle, and starts tracking changes to the original volumes.

When you are ready to make a backup, proceed to step 1.

Backing up a snapshot of a mounted file system with shared access

While you can run the commands in the following steps from any node, Symantec recommends running them from the master node.

To back up a snapshot of a mounted file system which has shared access

- 1 On any node, refresh the contents of the snapshot volumes from the original volume using the following command:

```
# vxsnap -g database_dg refresh snapvol source=database_vol \  
[snapvol2 source=database_vol2] ... syncing=yes
```

The `syncing=yes` attribute starts a synchronization of the snapshot in the background.

For example, to refresh the snapshot *snapvol*:

```
# vxsnap -g database_dg refresh snapvol source=database_vol \  
syncing=yes
```

This command can be run every time you want to back up the data. The `vxsnap refresh` command will resync only those regions which have been modified since the last refresh.

- 2 On any node of the cluster, use the following command to wait for the contents of the snapshot to be fully synchronous with the contents of the original volume:

```
# vxsnap -g database_dg syncwait snapvol
```

For example, to wait for synchronization to finish for the snapshots *snapvol*:

```
# vxsnap -g database_dg syncwait snapvol
```

Note: You cannot move a snapshot volume into a different disk group until synchronization of its contents is complete. You can use the `vxsnap print` command to check on the progress of synchronization.

- 3 On the master node, use the following command to split the snapshot volume into a separate disk group, *snapvoldg*, from the original disk group, *database_dg*:

```
# vxdg split volumedg
    snapvoldg
    snapvol
```

For example, to place the snapshot of the volume *database_vol* into the shared disk group *splitdg*:

```
# vxdg split database_dg splitdg snapvol
```

- 4 On the master node, deport the snapshot volume's disk group using the following command:

```
# vxdg deport snapvoldg
```

For example, to deport the disk group *splitdg*:

```
# vxdg deport splitdg
```

- 5 On the OHP host where the backup is to be performed, use the following command to import the snapshot volume's disk group:

```
# vxdg import snapvoldg
```

For example, to import the disk group *splitdg*:

```
# vxdg import splitdg
```

- 6 VxVM will recover the volumes automatically after the disk group import unless it is set not to recover automatically. Check if the snapshot volume is initially disabled and not recovered following the split.

If a volume is in the DISABLED state, use the following command on the OHP host to recover and restart the snapshot volume:

```
# vxrecover -g snapvoldg -m snapvol
```

For example, to start the volume *snapvol*:

```
# vxrecover -g splitdg -m snapvol
```

- 7 On the OHP host, use the following commands to check and locally mount the snapshot volume:

```
# fsck -F vxfs /dev/vx/rdisk/database_dg/database_vol  
  
# mount -F vxfs /dev/vx/dsk/database_dg/database_vol  
  mount_point
```

For example, to check and mount the volume *snapvol* in the disk group *snapvoldg* for shared access on the mount point, */bak/mnt_pnt*:

```
# fsck -F vxfs /dev/vx/rdisk/snapvoldg/snapvol  
# mount -F vxfs /dev/vx/dsk/snapvoldg/snapvol /bak/mnt_pnt
```

- 8 Back up the file system at this point using a command such as `bpbbackup` in Symantec NetBackup. After the backup is complete, use the following command to unmount the file system.

```
# umount mount_point
```

- 9 On the off-host processing host, use the following command to deport the snapshot volume's disk group:

```
# vxdg deport snapvoldg
```

For example, to deport *snapvoldg*:

```
# vxdg deport snapvoldg
```

- 10 On the master node, re-import the snapshot volume's disk group as a shared disk group using the following command:

```
# vxdg -s import snapvoldg
```

For example, to import *snapvoldg*:

```
# vxdg -s import snapvoldg
```

- 11 On the master node, use the following command to rejoin the snapshot volume's disk group with the original volume's disk group:

```
# vxdg join snapvoldg database_dg
```

For example, to join disk group *snapvoldg* with *database_dg*:

```
# vxdg join snapvoldg database_dg
```

- 12 VxVM will recover the volumes automatically after the join unless it is set not to recover automatically. Check if the snapshot volumes are initially disabled and not recovered following the join.

If a volume is in the DISABLED state, use the following command on the primary host to recover and restart the snapshot volume:

```
# vxrecover -g database_dg -m snapvol
```

- 13 When the backup is complete, use the following command to unmount the snapshot volume, and make it ready for its contents to be refreshed from the primary volume:

```
# umount mount_point
```

When synchronization is complete, the snapshot is ready to be re-used for backup.

Warning: Before attempting to unmount the snapshot, shut down all applications that access a file system in the snapshot volume, and also unmount any such file system.

Repeat the entire procedure each time that you need to back up the volume.

Resynchronizing a volume from its snapshot volume

In some instances, such as recovering the contents of a corrupted volume, it may be useful to resynchronize a volume from its snapshot volume (which is used as a hot standby).

To resynchronize a volume from its snapshot volume

- ◆ Enter:

```
vxsnap -g database_dg restore database_vol source=snapvol \  
destroy=yes|no
```

The `destroy` attribute specifies whether the plexes of the snapshot volume are to be reattached to the original volume. For example, to resynchronize the volume `database_vol` from its snapshot volume `snapvol` without removing the snapshot volume:

```
# vxsnap -g database_dg restore database_vol source=snapvol destroy=no
```

Note: You must unmount the file system that is configured on the original volume before attempting to resynchronize its contents from a snapshot.

Reattaching snapshot plexes

Some or all plexes of an instant snapshot may be reattached to the specified original volume, or to a source volume in the snapshot hierarchy above the snapshot volume.

Note: This operation is not supported for space-optimized instant snapshots.

By default, all the plexes are reattached, which results in the removal of the snapshot. If required, the number of plexes to be reattached may be specified as the value assigned to the `nmirror` attribute.

Note: The snapshot being reattached must not be open to any application. For example, any file system configured on the snapshot volume must first be unmounted.

To reattach a snapshot

- ◆ Use the following command to reattach an instant snapshot to the specified original volume, or to a source volume in the snapshot hierarchy above the snapshot volume:

```
vxsnap [-g database_dg] reattach snapvol source=database_vol \  
      [nmirror=number]
```

For example the following command reattaches 1 plex from the snapshot volume, *snapvol*, to the volume, *database_vol*:

```
# vxsnap -g database_dg reattach snapvol source=database_vol nmirror=1
```

While the reattached plexes are being resynchronized from the data in the parent volume, they remain in the `SNAPTMP` state. After resynchronization is complete, the plexes are placed in the `SNAPDONE` state.

The `vxsnap refresh` and `vxsnap reattach` commands have slightly different behaviors.

The `vxsnap reattach` command reattaches a snapshot volume to its source volume and begins copying the volume data to the snapshot volume.

The `vxsnap refresh` command updates the snapshot volumes contents view. The updated snapshot is available immediately with the new contents while synchronization occurs in the background.

Database recovery using Storage Checkpoints

You can use Storage Checkpoints to implement efficient backup and recovery of databases that have been laid out on VxFS file systems. A Storage Checkpoint allows you to roll back an entire database, a tablespace, or a single database file to the time that the Storage Checkpoint was taken. Rolling back to or restoring from any Storage Checkpoint is generally very fast because only the changed data blocks need to be restored.

Storage Checkpoints can also be mounted, allowing regular file system operations to be performed or secondary databases to be started.

For information on how to administer Storage Checkpoints, see *Veritas Storage Foundation™ Administrator's Guide*.

For information on how to administer Database Storage Checkpoints for an Oracle database, see *Veritas Storage Foundation™: Storage and High Availability Management for Oracle Databases*.

Note: Storage Checkpoints can only be used to restore from logical errors such as human mistakes or software faults. You cannot use them to restore files after a disk failure because all the data blocks are on the same physical device. Disk failure requires restoration of a database from a backup copy of the database files kept on a separate medium. Combining data redundancy (for example, disk mirroring) with Storage Checkpoints is recommended for highly critical data to protect against both physical media failure and logical errors.

Storage Checkpoints require space in the file systems where they are created, and the space required grows over time as copies of changed file system blocks are made. If a file system runs out of space, and there is no disk space into which the file system and any underlying volume can expand, VxFS automatically removes the oldest Storage Checkpoints if they were created with the removable attribute.

Creating Storage Checkpoints

To create Storage Checkpoints, select 3 Storage Checkpoint Administration > Create New Storage Checkpoints in the VxDBA utility. This can be done with a database either online or offline.

Note: To create a Storage Checkpoint while the database is online, ARCHIVELOG mode must be enabled in Oracle. During the creation of the Storage Checkpoint, the tablespaces are placed in backup mode. Because it only takes a few seconds to take a Storage Checkpoint, the extra redo logs generated while the tablespaces are in online backup mode are very small. To optimize recovery, it is recommended that you keep ARCHIVELOG mode enabled.

Warning: Changes to the structure of a database, such as the addition or removal of datafiles, make Storage Rollback impossible if they are made after a Storage Checkpoint was taken. A backup copy of the control file for the database is saved under the `/etc/vx/vxdba/ORACLE_SID/checkpoint_dir` directory immediately after a Storage Checkpoint is created. If necessary, you can use this file to assist with database recovery. If possible, both an ASCII and binary copy of the control file are made, with the binary version being compressed to conserve space. Use extreme caution if you attempt to recover your database using these control files. It is recommended that you remove old Storage Checkpoints and create new ones whenever you restructure a database.

Rolling back a database

The procedure in this section describes how to roll back a database using a Storage Checkpoint, for example, after a logical error has occurred.

To roll back a database

- 1 Ensure that the database is offline. You can use the VxDBA utility to display the status of the database and its tablespaces, and to shut down the database:
 - Select 2 `Display Database/VxDBA Information` to access the menus that display status information.
 - Select 1 `Database Administration > Shutdown Database Instance` to shut down a database.
- 2 Select 4 `Storage Rollback Administration > Roll Back the Database to a Storage Checkpoint` in the VxDBA utility, and choose the appropriate Storage Checkpoint. This restores all data files used by the database, except redo logs and control files, to their state at the time that the Storage Checkpoint was made.
- 3 Start up, but do not open, the database instance by selecting 1 `Database Administration > Startup Database Instance` in the VxDBA utility.
- 4 Use one of the following commands to perform an incomplete media recovery of the database:

- Recover the database until you stop the recovery:

```
recover database until cancel;  
...  
alter database [database] recover cancel;
```

- Recover the database to the point just before a specified system change number, scn:

```
recover database until change scn;
```

- Recover the database to the specified time:

```
recover database until time 'yyyy-mm-dd:hh:mm:ss';
```

- Recover the database to the specified time using a backup control file:

```
recover database until time 'yyyy-mm-dd:hh:mm:ss' \  
using backup controlfile;
```

Note: To find out when an error occurred, check the `../bdump/alert*.log` file.

See the Oracle documentation for complete and detailed information on database recovery.

- 5 To open the database after an incomplete media recovery, use the following command:

```
alter database open resetlogs;
```

Note: The `resetlogs` option is required after an incomplete media recovery to reset the log sequence. Remember to perform a full database backup and create another Storage Checkpoint after log reset.

- 6 Perform a full database backup, and use the VxDBA utility to remove any existing Storage Checkpoints that were taken before the one to which you just rolled back the database. These Storage Checkpoints can no longer be used for Storage Rollback. If required, use the VxDBA utility to delete the old Storage Checkpoints and to create new ones.

Backing up and recovering in a NetBackup environment

This chapter includes the following topics:

- [About Veritas NetBackup](#)
- [About using Veritas NetBackup for backup and restore for DB2](#)
- [About using NetBackup for backup and restore for Sybase](#)
- [About using Veritas NetBackup to backup and restore Quick I/O files for DB2](#)
- [About using Veritas NetBackup to backup and restore Quick I/O files for Sybase](#)
- [Using NetBackup in an SFHA Solutions product environment](#)

About Veritas NetBackup

Veritas NetBackup provides backup, archive, and restore capabilities for database files and directories contained on client systems in a client-server network. NetBackup server software resides on platforms that manage physical backup storage devices. The NetBackup server provides robotic control, media management, error handling, scheduling, and a repository of all client backup images.

Administrators can set up schedules for automatic, unattended full and incremental backups. These backups are managed entirely by the NetBackup server. The administrator can also manually back up clients. Client users can

perform backups, archives, and restores from their client system, and once started, these operations also run under the control of the NetBackup server.

Veritas NetBackup can be configured for DB2 in an Extended Edition (EE) or Extended-Enterprise Edition (EEE) environment. For detailed information and instructions on configuring DB2 for EEE, see “Configuring for a DB2 EEE (DPF) Environment” in the *Veritas NetBackup for DB2 System Administrator's Guide for UNIX*.

Veritas NetBackup, while not a shipped component of Veritas Storage Foundation Enterprise products, can be purchased separately.

About using Veritas NetBackup for backup and restore for DB2

With Veritas NetBackup, you can perform high performance, online (hot) backups of databases that must be available on a 24x7 basis. NetBackup supports the Extended Edition (EE) and the Enterprise Extended Edition (EEE) environments. NetBackup also supports Database Partitioning Feature (DPF) for DB2 8.1 and higher.

Veritas NetBackup enables you to back up and restore database files and directories. You can set up schedules for automatic, unattended database backup, as well as full or incremental backup. These backups are managed entirely by the NetBackup server. You can also manually back up database files from any of the NetBackup clients. Client users can perform database backups and restores from their client systems on demand.

Veritas NetBackup can be configured for DB2 in an Extended Edition (EE), Extended-Enterprise Edition (EEE), or Database Partitioning Feature (DPF) environment. Two types of DB2 backup policies are required. One is used to backup the catalog nodes and the other is used to backup all the nodes, including the catalog node. Detailed information and instructions on configuring DB2 for EEE is available in the system administrator's guide.

See the *Veritas NetBackup for DB2 System Administrator's Guide for UNIX*.

Veritas NetBackup for DB2 has the following features:

- Media and device management
- Scheduling facilities
- Multiplexed backups and restores
- Transparent execution of both DB2 and regular file system backup and restore operations

- Shared devices and tapes used during other file backups
- Centralized and networked backup operations
- Parallel backup and restore operations
- Incremental backups of DB2 databases

Table 8-1 Options for backing up DB2 with NetBackup

	Automatically	Manually	DB2 BACKUP DATABASE command
DB2 database log backups	Supported	Supported	Supported
DB2 archive log backups	Supported	Supported	Supported
DB2 policy backups	Supported	Supported	

Setting up schedules for automatic backups is the most convenient way to back up your database.

See 'Performing a Backup' in the *Veritas NetBackup for DB2 System Administrator's Guide for UNIX*.

The procedure for restoring a DB2 database depends on the database involved and the problems that you have on your system. You can browse the backups using the `db2 list history` command or using the NetBackup `bplist` command before restoring.

See the *DB2 UDB Administration Guide Data Recovery and High Availability Guide*.

About using NetBackup for backup and restore for Sybase

Veritas NetBackup for Sybase is not included in the standard Veritas Database Edition. The information included here is for reference only.

Veritas NetBackup for Sybase integrates the database backup and recovery capabilities of Sybase Backup Server with the backup and recovery management capabilities of NetBackup.

Veritas NetBackup works with Sybase APIs to provide high-performance backup and restore for Sybase dataservers. With Veritas NetBackup, you can set up schedules for automatic, unattended backups for Sybase ASE dataservers (NetBackup clients) across the network. These backups can be full database dumps or incremental backups (transaction logs) and are managed by the NetBackup

server. You can also manually backup dataservers. The Sybase `dump` and `load` commands are used to perform backups and restores.

Veritas NetBackup has both graphical and menu driven user interfaces to suit your needs.

For details, refer to *Veritas NetBackup System Administrator's Guide for UNIX*.

About using Veritas NetBackup to backup and restore Quick I/O files for DB2

Veritas NetBackup does not follow symbolic links when backing up files. Typical backup management applications are designed this way to avoid backing up the same data twice. This would happen if both the link and the file it points to were included in the list of files to be backed up.

A Quick I/O file consists of two components: a hidden file with the space allocated for it, and a link that points to the Quick I/O interface of the hidden file. Because NetBackup does not follow symbolic links, you must specify both the Quick I/O link and its hidden file in the list of files to be backed up.

To view all files and their attributes in the `db01` directory:

```
$ ls -la /db01

total 2192

drwxr-xr-x  2 root  root  96 Oct 20 17:39  .
drwxr-xr-x  9 root  root 8192 Oct 20 17:39  ..
-rw-r--r--  1 db2  dba 1048576 Oct 20 17:39  .dbfile
lrwxrwxrwx  1 db2  dba  22 Oct 20 17:39  dbfile ->\
.dbfile::cdev:vxfs:
```

In the example above, you must include both the symbolic link `dbfile` and the hidden file `.dbfile` in the file list of the backup class.

If you want to back up all Quick I/O files in a directory, you can simplify the process by just specifying the directory to be backed up. In this case, both components of each Quick I/O file will be properly backed up. In general, you should specify directories to be backed up unless you only want to back up some, but not all files, in those directories.

Because NetBackup is tightly integrated with the Veritas Database Edition for DB2, NetBackup backs up extent attributes of a Quick I/O file and restores them

accordingly. Quick I/O files can then be backed up and restored as regular files using NetBackup, while preserving the Quick I/O file's extent reservation. Without this feature, restoring the file could cause the loss of contiguous reservation, which can degrade performance.

When restoring a Quick I/O file, if both the symbolic link and the hidden file already exist, NetBackup will restore both components from the backup image. If either one of or both of the two components are missing, NetBackup creates or overwrites as needed.

About using Veritas NetBackup to backup and restore Quick I/O files for Sybase

Veritas NetBackup does not follow symbolic links when backing up files. Typical backup management applications are designed this way to avoid backing up the same data twice. This would happen if both the link and the file it points to were included in the list of files to be backed up.

A Quick I/O file consists of two components: a hidden file with the space allocated for it, and a link that points to the Quick I/O interface of the hidden file. Because NetBackup does not follow symbolic links, you must specify both the Quick I/O link and its hidden file in the list of files to be backed up.

To view all files and their attributes in the db01 directory:

```
$ ls -la /db01

total 2192

drwxr-xr-x    2 root   root   96 Oct 20 17:39  .
drwxr-xr-x    9 root   root  8192 Oct 20 17:39  ..
-rw-r--r--    1 db2    dba   1048576 Oct 20 17:39  .dbfile
lrwxrwxrwx    1 db2    dba    22 Oct 20 17:39  dbfile ->\
.dbfile::cdev:vxfs:
```

In the example above, you must include both the symbolic link `dbfile` and the hidden file `.dbfile` in the file list of the backup class.

If you want to back up all Quick I/O files in a directory, you can simplify the process by just specifying the directory to be backed up. In this case, both components of each Quick I/O file will be properly backed up. In general, you should specify directories to be backed up unless you only want to back up some, but not all files, in those directories.

When restoring a Quick I/O file, if both the symbolic link and the hidden file already exist, NetBackup will restore both components from the backup image. If either one of or both of the two components are missing, NetBackup creates or overwrites as needed.

Using NetBackup in an SFHA Solutions product environment

You can enhance the ease of use and efficiency of your SFHA Solutions product and NetBackup by integrating them as follows:

- Clustering a NetBackup Master Server
- Backing up and recovering a VxVM volume using NetBackup

Clustering a NetBackup Master Server

To enable your NetBackup Master Server to be highly available in a cluster environment, use the following procedure.

To make a NetBackup Master Server, media, and processes highly available

- 1 Verify that your versions of NetBackup and Veritas Cluster Server are compatible. Detailed combination information is included in the NetBackup cluster compatibility list:
 - For NetBackup 7.x cluster compatibility:
See <http://www.symantec.com/business/support/index?page=content&id=TECH126902>
 - For NetBackup 6.x cluster compatibility:
See <http://www.symantec.com/business/support/index?page=content&id=TECH43619>
 - For NetBackup 5.x cluster compatibility:
See <http://www.symantec.com/business/support/index?page=content&id=TECH29272>
 - For more on NetBackup compatibility, see <http://www.symantec.com/business/support/index?page=landing&key=15145>
- 2 The steps to cluster a Master Server are different for different versions of NetBackup. See the applicable NetBackup guide for directions.
 - For NetBackup 7.x:
See *Symantec NetBackup Clustered Master Server Administrator's Guide*
See <http://www.symantec.com/docs/DOC3679>

- For NetBackup 7.0.1:
See *Symantec NetBackup Clustered Master Server Administrator's Guide 7.0.1*
See <http://www.symantec.com/docs/DOC135520>
- For NetBackup 7.0:
See *NetBackup High Availability Guide 7.0*
See <http://www.symantec.com/docs/DOC127045>
- For NetBackup 6.5:
See *NetBackup High Availability Guide 6.5*
See <http://www.symantec.com/docs/DOC52835>
- For more on NetBackup documentation, see
<http://www.symantec.com/business/support/index?page=landing&key=15145>

To verify the robustness of the VCS resources and NetBackup processes

- 1 Verify that you can online the Netbackup master.
- 2 Verify that you can offline the Netbackup master.
- 3 Verify that you can monitor all the NetBackup resources.

Backing up and recovering a VxVM volume using NetBackup

To enable NetBackup to backup objects on a VxVM volume, use the following procedure. This procedure enables an Instant Recovery (IR) using a VxVM volume.

To back up objects in a VxVM volume using NetBackup

- 1 Create a VxVM disk group with six disks. The number of disks may vary depending on the volume size, disk size, volume layout, and snapshot method.

If the system this test is running on is a clustered system, create a shared disk group using the `-s` option.

```
# vxvxdg -s init database_dg disk1 disk2 disk3 \  
disk4 disk5 disk6
```

- 2 Create a "mirror-striped" VxVM volume with a size of 10 Gbytes or the maximum size of the disk, whichever is larger.

```
# vxassist -g database_dg make vol_name 10G \  
layout=mirror-stripe init=active  
# vxvol -g database_dg set fastresync=on vol_name  
# vxassist -g database_dg snapstart nmirror=1 vol_name
```

Note: There are three types of snapshot: mirror, full-size instant, and space-optimized instant snapshots. The example uses an Instant Recovery (IR) snapshot. For snapshot creation details:

See pages 104-107 of the *NetBackup Snapshot Client Administrator's Guide* for 7.1.

See <http://www.symantec.com/docs/DOC3661>

- 3 Make the file system on the volume.
- 4 Mount a VxFS file system on the volume.
If the VxVM volume is a clustered volume, mount the VxFS file system with the `"-o cluster"` option.
- 5 Fill up the VxFS file system up to the desired level. For example, you can fill to 95% full, or to whatever level is appropriate for your file system.
- 6 Store the `cksum(1)` for these files.
- 7 Un-mount the VxFS file system.
- 8 Enable the following Advanced Client option:
 - Perform Snapshot Backup.
 - Set **Advanced Snapshot Options** to `vxvm`.

- Enable **Retain snapshots for instant recovery**.
- 9 Back up the VxVM volume with the NetBackup policy.
See pages 98-101 of the *NetBackup Snapshot Client Administrator's Guide* for 7.1.
See <http://www.symantec.com/docs/DOC3661>

Recovering a VxVM volume using NetBackup

To enable NetBackup to recover objects on a VxVM volume, use the following procedure. This procedure performs an Instant Recovery (IR) using a VxVM volume.

To recover objects in a VxVM volume using NetBackup

- 1 Initialize the VxVM volume to zeros.
- 2 Recover the VxVM volume to the newly initialized VxVM volume.
- 3 Mount the VxFS file system on the empty VxVM volume.
- 4 Verify the cksum(1) values against the files recovered.

Off-host processing

This chapter includes the following topics:

- [Storage Foundation off-host processing methods](#)
- [Using a replica database for decision support](#)
- [What is off-host processing?](#)
- [About using VVR for off-host processing](#)

Storage Foundation off-host processing methods

While backup and recovery is an important use case for Storage Foundation point-in-time copy methods, they can also be used for:

- Periodic analysis (mining) of production data
- Predictive what-if analysis
- Software testing against real data
- Application or database problem diagnosis and resolution

Off-host processing use cases are similar to the backup use case in that they generally require consistent images of production data sets. They differ from backup in the three important respects:

- **Access mode**
Whereas backup is a read-only activity, most off-host processing activities update the data they process. Thus, Snapshot File Systems are of limited utility for off-host processing uses.
- **Multiple uses**
Backup uses each source data image once, after which the snapshot can be discarded. With other use cases, it is often useful to perform several experiments on the same data set. It is possible to take snapshots of both

Storage Checkpoints and Space-Optimized Instant Snapshots of production data. This facility provides multiple identical data images for exploratory applications at almost no incremental overhead. Rather than testing destructively against a snapshot containing the data set state of interest, tests can be run against snapshots of that snapshot. After each test, the snapshot used can be deleted, leaving the original snapshot containing the starting state intact. Any number of tests or analyses can start with the same data, providing comparable alternatives for evaluation. All such tests can be run while production applications simultaneously process live data.

- **Scheduling**

Whereas backup is typically a regularly scheduled activity, allowing storage and I/O capacity needs to be planned, other applications of snapshots must run with little or no notice. Full-sized and space-optimized instant snapshots and Storage Checkpoints provide instantly accessible snapshots, and are therefore more suitable for these applications.

Storage Foundation examples for data analysis and off-host processing use cases:

- Decision support
- Active secondary use-case with VVR

Using a replica database for decision support

You can use snapshots of a primary database to create a replica of the database at a given moment in time. You can then implement decision support analysis and report generation operations that take their data from the database copy rather than from the primary database. The FastResync functionality of Veritas Volume Manager (VxVM) allows you to quickly refresh the database copy with up-to-date information from the primary database. Reducing the time taken to update decision support data also lets you generate analysis reports more frequently.

Two methods are described for setting up a replica database for decision support:

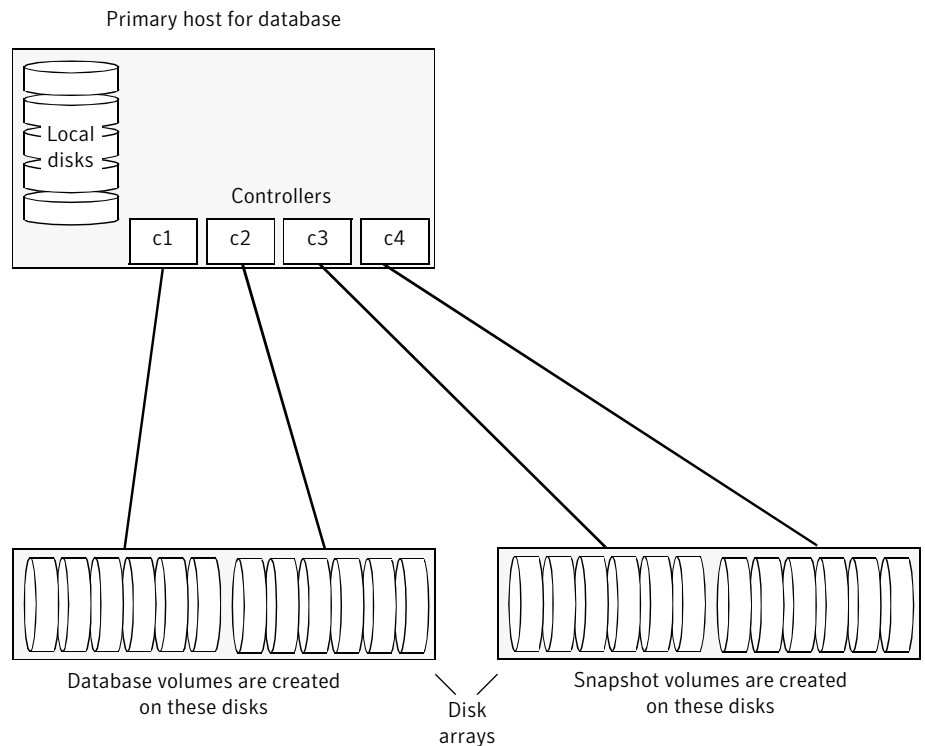
- See [“Creating a replica database on the same host”](#) on page 141.
- See [“Creating an off-host replica database”](#) on page 151.

Note: All commands require superuser (`root`) or equivalent privileges, except where it is explicitly stated that a command must be run by the database administrator.

Creating a replica database on the same host

Figure 9-1 shows an example where the primary database volumes to be backed up, `dbase_vol` and `dbase_logs`, are configured on disks attached to controllers `c1` and `c2`, and the snapshots are to be created on disks attached to controllers `c3` and `c4`.

Figure 9-1 Example system configuration for decision support on the primary host



To set up a replica database to be used for decision support on the primary host

- Prepare the snapshot, either full-sized or space-optimized.
See [“Preparing a full-sized instant snapshot for a backup”](#) on page 96.
- Create snapshot mirrors for volumes containing VxFS file systems for database files to be backed up.
- Make the database replica.
- All commands require superuser (`root`) or equivalent privileges.

Preparing for the replica database

To prepare a snapshot for a replica database on the primary host

- 1 If you have not already done so, prepare the host to use the snapshot volume that contains the copy of the database tables. Set up any new database logs and configuration files that are required to initialize the database.
- 2 Use the following command to make a full-sized snapshot, *snapvol*, of the tablespace volume by breaking off plexes from the original volume:

```
# vxsnap -g database_dg make \  
source=volume/newvol=snapvol/nmirror=N
```

The `nmirror` attribute specifies the number of mirrors, *N*, in the snapshot volume.

If the volume does not have any available plexes, or its layout does not support plex break-off, prepare an empty volume for the snapshot.

- 3 Use the `vxprint` command on the original volume to find the required size for the snapshot volume.

```
# LEN=`vxprint [-g diskgroup] -F%len volume`
```

Note: The command shown in this and subsequent steps assumes that you are using a Bourne-type shell such as `sh`, `ksh` or `bash`. You may need to modify the command for other shells such as `csh` or `tcsh`. These steps are valid only for an instant snap DCO.

- 4 Use the `vxprint` command on the original volume to discover the name of its DCO:

```
# DCONAME=`vxprint [-g diskgroup] -F%dco_name volume`
```

- 5 Use the `vxprint` command on the DCO to discover its region size (in blocks):

```
# RSZ=`vxprint [-g diskgroup] -F%regionsz $DCONAME`
```

- 6 Use the `vxassist` command to create a volume, `snapvol`, of the required size and redundancy, together with an instant snap DCO volume with the correct region size:

```
# vxassist [-g diskgroup] make snapvol $LEN \  
  [layout=mirror nmirror=number] logtype=dco dnl=no \  
  dconversion=20 [ndcomirror=number] regionsz=$RSZ \  
  init=active [storage_attributes]
```

It is recommended that you specify the same number of DCO mirrors (`ndcomirror`) as the number of mirrors in the volume (`nmirror`). The `init=active` attribute is used to make the volume available immediately. You can use storage attributes to specify which disks should be used for the volume.

As an alternative to creating the snapshot volume and its DCO volume in a single step, you can first create the volume, and then prepare it for instant snapshot operations as shown here:

```
# vxassist [-g diskgroup] make snapvol $LEN \  
  [layout=mirror nmirror=number] init=active \  
  [storage_attributes]  
# vxsnap [-g diskgroup] prepare snapvol [ndcomirs=number] \  
  regionsz=$RSZ [storage_attributes]
```

7 To create the snapshot, use the following command:

```
# vxsnap -g database_dg make source=volume/snapvol=snapvol
```

If a database spans more than one volume, specify all the volumes and their snapshot volumes as separate tuples on the same line, for example:

```
# vxsnap -g database_dg make \  
source=vol1/snapvol=svol1/nmirror=2 \  
source=vol2/snapvol=svol2/nmirror=2 \  
source=vol3/snapvol=svol3/nmirror=2
```

If you want to save disk space, you can use the following command to create a space-optimized snapshot instead:

```
# vxsnap -g database_dg make \  
source=volume/newvol=snapvol/cache=cacheobject
```

The argument `cacheobject` is the name of a pre-existing cache that you have created in the disk group for use with space-optimized snapshots.

If several space-optimized snapshots are to be created at the same time, these can all specify the same cache object as shown in this example:

```
# vxsnap -g database_dg make \  
source=vol1/newvol=svol1/cache=dbaseco \  
source=vol2/newvol=svol2/cache=dbaseco \  
source=vol3/newvol=svol3/cache=dbaseco
```

8 Decide on the following characteristics that you want to allocate to the cache volume that underlies the cache object:

- The size of the cache volume should be sufficient to record changes to the parent volumes during the interval between snapshot refreshes. A suggested value is 10% of the total size of the parent volumes for a refresh interval of 24 hours.
- If redundancy is a desired characteristic of the cache volume, it should be mirrored. This increases the space that is required for the cache volume in proportion to the number of mirrors that it has.
- If the cache volume is mirrored, space is required on at least as many disks as it has mirrors. These disks should not be shared with the disks used for the parent volumes. The disks should also be chosen to avoid impacting I/O performance for critical volumes, or hindering disk group split and join operations.

- 9 Having decided on its characteristics, use the `vxassist` command to create the volume that is to be used for the cache volume. The following example creates a mirrored cache volume, `cachevol`, with size 1GB in the disk group, `mydg`, on the disks `disk16` and `disk17`:

```
# vxassist -g mydg make cachevol 1g layout=mirror \  
  init=active disk16 disk17
```

The attribute `init=active` is specified to make the cache volume immediately available for use.

- 10** Use the `vxmake cache` command to create a cache object on top of the cache volume that you created in the previous step:

```
# vxmake [-g diskgroup] cache cache_object \
  cachevolname=volume [regionsize=size] [autogrow=on] \
  [highwatermark=hwmk] [autogrowby=agbvalue] \
  [maxautogrow=maxagbvalue]
```

If you specify the region size, it must be a power of 2, and be greater than or equal to 16KB (16k). If not specified, the region size of the cache is set to 64KB.

Note: All space-optimized snapshots that share the cache must have a region size that is equal to or an integer multiple of the region size set on the cache. Snapshot creation also fails if the original volume's region size is smaller than the cache's region size.

If the cache is not allowed to grow in size as required, specify `autogrow=off`. By default, the ability to automatically grow the cache is turned on.

In the following example, the cache object, `cobjmydg`, is created over the cache volume, `cachevol`, the region size of the cache is set to 32KB, and the `autogrow` feature is enabled:

```
# vxmake -g mydg cache cobjmydg cachevolname=cachevol \
  regionsize=32k autogrow=on
```

- 11** Having created the cache object, use the following command to enable it:

```
# vxcache [-g diskgroup] start cache_object
```

For example to start the cache object, `cobjmydg`:

```
# vxcache -g mydg start cobjmydg
```

Note: This step sets up the snapshot volumes, and starts tracking changes to the original volumes.

Creating a replica database

After you prepare the snapshot, you are ready to create a replica of the database.

To create the replica database

- 1 If the volumes to be backed up contain database tables in file systems, suspend updates to the volumes:

- DB2 provides the `write suspend` command to temporarily suspend I/O activity for a database. As the DB2 database administrator, use a script such as that shown in the example. Note that to allow recovery from any backups taken from snapshots, the database must be in LOGRETAIN RECOVERY mode.

```
#!/bin/ksh
#
# script: backup_start.sh
#
# Sample script to suspend I/O for a DB2 database.
#
# Note: To recover a database using backups of snapshots,
# the database must be in LOGRETAIN mode.

db2 <<!
connect to database
set write suspend for database
quit
!
```

- Sybase ASE from version 12.0 onward provides the Quiesce feature to allow temporary suspension of writes to a database. As the Sybase database administrator, put the database in quiesce mode by using a script such as that shown in the example.

```
#!/bin/ksh
#
# script: backup_start.sh
#
# Sample script to quiesce example Sybase ASE database.
#
# Note: The "for external dump" clause was introduced in Sybase
# ASE 12.5 to allow a snapshot database to be rolled forward.
# See the Sybase ASE 12.5 documentation for more information.

isql -Usa -Ppassword -SFMR <<!
quiesce database tag hold database1[, database2]... [for
external dump]
go
```

```
quit
!
```

If you are using Sybase ASE 12.5, you can specify the `for external dump` clause to the `quiesce` command. This warm standby method allows you to update a replica database using transaction logs dumped from the primary database.

See [“Updating a warm standby Sybase ASE 12.5 database”](#) on page 161.

- 2 Refresh the contents of the snapshot volumes from the original volume using the following command:

```
# vxsnap -g database_dg refresh snapvol source=vol \
  [snapvol2 source=vol2]...
```

For example, to refresh the snapshots `svol1`, `svol2` and `svol3`:

```
# vxsnap -g database_dg refresh svol1 source=vol1 \
  svol2 source=vol2 svol3 source=vol3
```

- 3 If you temporarily suspended updates to volumes in step 1, perform the following steps.

Release all the tablespaces or databases from suspend, hot backup or quiesce mode:

- As the DB2 database administrator, use a script such as that shown in the example.

```
#!/bin/ksh
#
# script: backup_end.sh
#
# Sample script to resume I/O for a DB2 database.
#

db2 <<!
connect to database
set write resume for database
quit
!
```

- As the Sybase database administrator, release the database from quiesce mode using a script such as that shown in the example.

```
#!/bin/ksh
#
# script: backup_end.sh
#
# Sample script to release example Sybase ASE database from
# quiesce mode.

isql -Usa -Ppassword -SFMR <<!
quiesce database tag release
go
quit
!
```

If you are using Sybase ASE 12.5, you can specify the `for external dump` clause to the `quiesce` command. This warm standby method allows you to update a replica database using transaction logs dumped from the primary database.

See [“Updating a warm standby Sybase ASE 12.5 database”](#) on page 161.

- 4 For each snapshot volume containing tablespaces, check the file system that it contains, and mount the volume using the following commands:

```
# fsck -F vxfs /dev/vx/rdisk/diskgroup/snapvol
# mount -F vxfs /dev/vx/dsk/diskgroup/snapvol
  mount_point
```

For example, to check the file system in the snapshot volume `snap1_dbase_vol`, and mount it on `/rep_dbase_vol`:

```
# fsck -F vxfs /dev/vx/rdisk/database_dg/snap1_dbase_vol
# mount -F vxfs /dev/vx/dsk/database_dg/snap1_dbase_vol \
  /rep_dbase_vol
```

- 5 Copy any required log files from the primary database to the replica database.

For a Sybase ASE database, if you specified the `for external dump` clause when you quiesced the database, use the following `isql` command as the database administrator to dump the transaction log for the database:

```
dump transaction to dump_device with standby_access
```

Then copy the dumped transaction log to the appropriate replica database directory.

- 6 As the database administrator, start the new database:
 - For a Sybase ASE database, use a script such as that shown in the example.

```

#!/bin/ksh
#
# script: startdb.sh <list_of_database_volumes>
#
# Sample script to recover and start replica Sybase ASE
# database.

# Import the snapshot volume disk group.

vxdg import $snapvoldg

# Mount the snapshot volumes (the mount points must already
# exist).

for i in $*
do
    fsck -F vxfs /dev/vx/rdisk/$snapvoldg/snap_$i
    mount -F vxfs /dev/vx/dsk/$snapvoldg/snap_$i \
    ${rep_mnt_point}/${i}
done

# Start the replica database.
# Specify the -q option if you specified the "for external
# dump" clause when you quiesced the primary database.
# See the Sybase ASE 12.5 documentation for more information.

/sybase/ASE-12_5/bin/dataserver \
[-q] \
-sdatabase_name \
-d /sybevm/master \
-e /sybase/ASE-12_5/install/dbasename.log \
-M /sybase

# Online the database. Load the transaction log dump and
# specify "for standby_access" if you used the -q option
# with the dataserver command.

isql -Usa -Ppassword -SFMR <<!
[load transaction from dump_device with standby_access
go]
online database database_name [for standby_access]
go

```

```
quit  
!
```

If you are using the warm standby method, specify the `-q` option to the `dataserver` command. Use the following `isql` commands to load the dump of the transaction log and put the database online:

```
load transaction from dump_device with standby_access  
online database database_name for standby_access
```

If you are not using the warm standby method, use the following `isql` command to recover the database, roll back any uncommitted transactions to the time that the quiesce command was issued, and put the database online:

```
online database database_name
```

When you want to resynchronize a snapshot with the primary database, shut down the replica database, unmount the snapshot volume, and go back to step 1 to refresh the contents of the snapshot from the original volume.

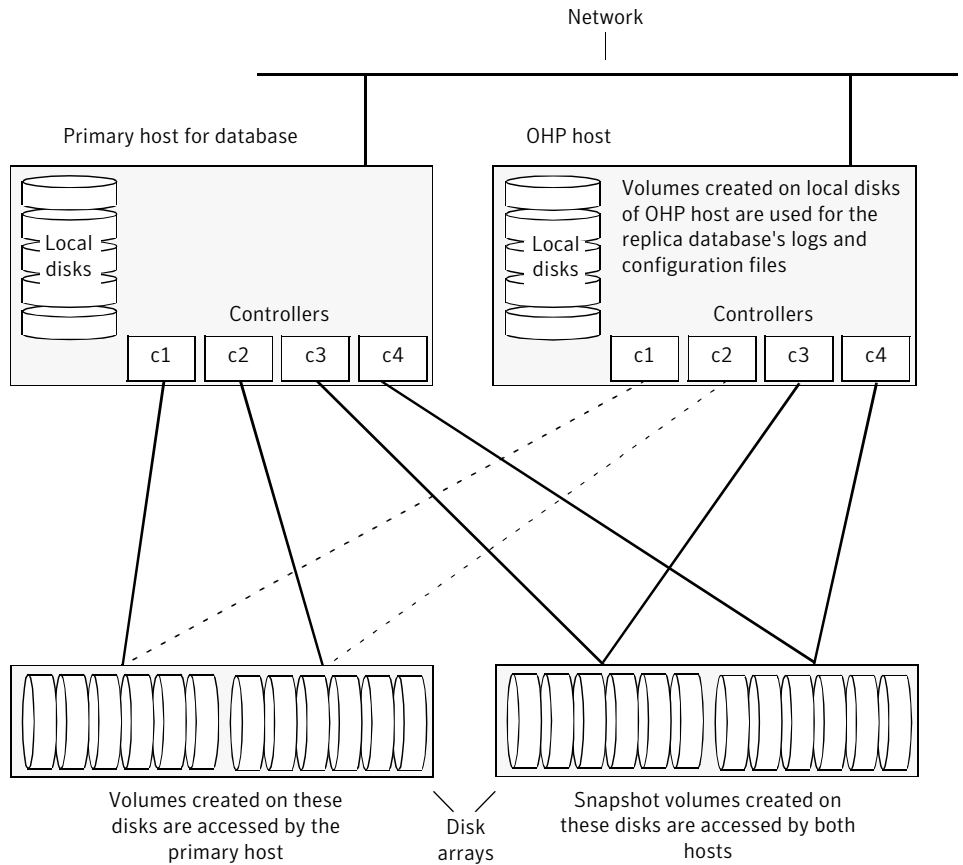
Creating an off-host replica database

Figure 9-2 shows an example where the primary database volumes to be backed up, `dbase_vol` and `dbase_logs`, are configured on disks attached to controllers `c1` and `c2`, and the snapshots are to be created on disks attached to controllers `c3` and `c4`.

There is no requirement for the off-host processing host to have access to the disks that contain the primary database volumes.

Note: If the database is configured on volumes in a cluster-shareable disk group, it is assumed that the primary host for the database is the master node for the cluster. If the primary host is not also the master node, all VxVM operations on shared disk groups must be performed on the master node.

Figure 9-2 Example system configuration for off-host decision support



To set up a replica database to be used for decision support on another host

- Prepare the full-sized snapshot.
See "[Preparing a space-optimized snapshot for a database backup](#)" on page 98.
- Create snapshot mirrors for volumes containing VxFS file systems for database files to be backed up.
- Make the database replica.
- All commands require superuser (`root`) or equivalent privileges.

Setting up a replica database for off-host decision support

To set up a replica database for off-host decision support

- 1 If you have not already done so, prepare the off-host processing host to use the snapshot volume that contains the copy of the database tables. Set up any new database logs and configuration files that are required to initialize the database.
- 2 On the primary host, use the following command to make a full-sized snapshot, `snapvol`, of the tablespace volume by breaking off plexes from the original volume:

```
# vxsnap -g database_dg make \
  source=volume/newvol=snapvol/nmirror=N
```

The `nmirror` attribute specifies the number of mirrors, `N`, in the snapshot volume.

If the volume does not have any available plexes, or its layout does not support plex break-off, prepare an empty volume for the snapshot.

Then use the following command to create the snapshot:

```
# vxsnap -g database_dg make source=volume/snapvol=snapvol
```

If a database spans more than one volume, specify all the volumes and their snapshot volumes as separate tuples on the same line, for example:

```
# vxsnap -g database_dg make source=vol1/snapvol=svol1 \
  source=vol2/snapvol=svol2 source=vol3/snapvol=svol3
```

Note: This step sets up the snapshot volumes, and starts tracking changes to the original volumes.

When you are ready to create the replica database, proceed to step 3.

- 3 If the volumes to be backed up contain database tables in file systems, suspend updates to the volumes:
 - DB2 provides the `write suspend` command to temporarily suspend I/O activity for a database. As the DB2 database administrator, use a script such as that shown in the example. Note that if the replica database must be able to be rolled forward (for example, if it is to be used as a standby database), the primary database must be in LOGRETAIN RECOVERY mode.

```
#!/bin/ksh
#
```

```

# script: backup_start.sh
#
# Sample script to suspend I/O for a DB2 database.
#
# Note: To recover a database using backups of snapshots, the database
# must be in LOGRETAIN mode.

db2 <<!
connect to database
set write suspend for database
quit
!
```

- Sybase ASE from version 12.0 onward provides the Quiesce feature to allow temporary suspension of writes to a database. As the Sybase database administrator, put the database in quiesce mode by using a script such as that shown in the example.

```

#!/bin/ksh
#
# script: backup_start.sh
#
# Sample script to quiesce example Sybase ASE database.
#
# Note: The "for external dump" clause was introduced in Sybase
# ASE 12.5 to allow a snapshot database to be rolled forward.
# See the Sybase ASE 12.5 documentation for more information.

isql -Usa -Ppassword -SFMR <<!
quiesce database tag hold database1[, database2]... [for external dump]
go
quit
!
```

If you are using Sybase ASE 12.5, you can specify the `for external dump` clause to the `quiesce` command. This warm standby method allows you to update a replica database using transaction logs dumped from the primary database.

See ["Updating a warm standby Sybase ASE 12.5 database"](#) on page 161.

- 4 On the primary host, refresh the contents of the snapshot volumes from the original volume using the following command:

```
# vxsnap -g database_dg refresh snapvol source=vol \  
[snapvol2 source=vol2]... syncing=yes
```

The `syncing=yes` attribute starts a synchronization of the snapshot in the background.

For example, to refresh the snapshots `svol1`, `svol2` and `svol3`:

```
# vxsnap -g database_dg refresh svol1 source=vol1 \  
svol2 source=vol2 svol3 source=vol3
```

- 5 If you temporarily suspended updates to volumes in step 3, release all the tablespaces or databases from suspend, hot backup or quiesce mode:
 - As the DB2 database administrator, use a script such as that shown in the example.

```
#!/bin/ksh  
#  
# script: backup_end.sh  
#  
# Sample script to resume I/O for a DB2 database.  
#  
  
db2 <<!  
connect to database  
set write resume for database  
quit  
!
```

- As the Sybase database administrator, release the database from quiesce mode using a script such as that shown in the example.

```
#!/bin/ksh  
#  
# script: backup_end.sh  
#  
# Sample script to release example Sybase ASE database from  
quiesce mode.  
  
isql -Usa -Ppassword -SFMR <<!  
quiesce database tag release
```

```

go
quit
!
```

- 6 Use the following command to wait for the contents of the snapshot to be fully synchronous with the contents of the original volume:

```
# vxsnap -g database_dg syncwait snapvol
```

For example, to wait for synchronization to finish for all the snapshots `svol1`, `svol2` and `svol3`, you would issue three separate commands:

```
# vxsnap -g database_dg syncwait svol1
# vxsnap -g database_dg syncwait svol2
# vxsnap -g database_dg syncwait svol3
```

Note: You cannot move a snapshot volume into a different disk group until synchronization of its contents is complete. You can use the `vxsnap print` command to check on the progress of synchronization.

- 7 On the primary host, use the following command to split the disks containing the snapshot volumes into a separate disk group, `snapvoldg`, from the original disk group, `database_dg`:

```
# vxdg split database_dg snapvoldgsnapvol ...
```

- 8 On the primary host, deport the snapshot volume's disk group using the following command:

```
# vxdg deport snapvoldg
```

- 9 On the off-host processing host where the replica database is to be set up, use the following command to import the snapshot volume's disk group:

```
# vxdg import snapvoldg
```

- 10 VxVM will recover the volumes automatically after the disk group import unless it is set to not recover automatically. Check if the snapshot volume is initially disabled and not recovered following the split.

If a volume is in the DISABLED state, use the following command on the off-host processing host to recover and restart the snapshot volume:

```
# vxrecover -g snapvoldg -m snapvol ...
# vxvol -g snapvoldg start snapvol ...
```

- 11 On the off-host processing host, for each snapshot volume containing tablespaces, check the file system that it contains, and mount the volume using the following commands:

```
# fsck -F vxfs /dev/vx/rdisk/diskgroup/snapvol
# mount -F vxfs /dev/vx/dsk/diskgroup/snapvol
  mount_point
```

For example, to check the file system in the snapshot volume `snap1_dbase_vol`, and mount it on `/rep/dbase_vol`:

```
# fsck -F vxfs /dev/vx/rdisk/snapvoldg/snap1_dbase_vol
# mount -F vxfs /dev/vx/dsk/snapvoldg/snap1_dbase_vol \
  /rep/dbase_vol
```

Note: For a replica DB2 database, the database volume must be mounted in the same location as on the primary host.

- 12 Copy any required log files from the primary host to the off-host processing host.

For a Sybase ASE database on the primary host, if you specified the `for external dump` clause when you quiesced the database, use the following `isql` command as the database administrator to dump the transaction log for the database:

```
dump transaction to dump_device with standby_access
```

Then copy the dumped transaction log to the appropriate database directory on the off-host processing host.

- 13 As the database administrator, start the new database:
- If the replica DB2 database is not to be rolled forward, use the following commands to start and recover it:

```
db2start
db2inidb database as snapshot
```

If the replica DB2 database is to be rolled forward (the primary must have been placed in LOGRETAIN RECOVERY mode before the snapshot was taken), use the following commands to start it, and put it in roll-forward pending state:

```
db2start
db2inidb database as standby
```

Obtain the latest log files from the primary database, and use the following command to roll the replica database forward to the end of the logs:

```
db2 rollforward db database to end of logs
```

- For a Sybase ASE database, use a script such as that shown in the example.

```
#!/bin/ksh
#
# script: startdb.sh <list_of_database_volumes>
#
# Sample script to recover and start replica Sybase ASE
# database.

# Import the snapshot volume disk group.

vxdg import $snapvoldg

# Mount the snapshot volumes (the mount points must already
# exist).

for i in $*
do
    fsck -F vxfs /dev/vx/rdisk/$snapvoldg/snap_$i
    mount -F vxfs /dev/vx/dsk/$snapvoldg/snap_$i \
${rep_mnt_point}/${i}
done

# Start the replica database.
# Specify the -q option if you specified the "for external
# dump" clause when you quiesced the primary database.
# See the Sybase ASE 12.5 documentation for more information.
```

```
/sybase/ASE-12_5/bin/dataserver \  
[-q] \  
-sdatabase_name \  
-d /sybevm/master \  
-e /sybase/ASE-12_5/install/dbasename.log \  
-M /sybase  
  
# Online the database. Load the transaction log dump and  
# specify "for standby_access" if you used the -q option  
# with the dataserver command.  
  
isql -Usa -Ppassword -SFMR <<!  
[load transaction from dump_device with standby_access  
go]  
online database database_name [for standby_access]  
go  
quit  
!
```

If you are using the warm standby method, specify the `-q` option to the `dataserver` command. Use the following `isql` commands to load the dump of the transaction log and put the database online:

```
load transaction from dump_device with standby_access  
online database database_name for standby_access
```

If you are not using the warm standby method, use the following `isql` command to recover the database, roll back any uncommitted transactions to the time that the quiesce command was issued, and put the database online:

```
online database database_name
```

Resynchronizing the data with the primary host

This procedure describes how to resynchronize the data in a snapshot with the primary host.

To resynchronize a snapshot with the primary database

- 1 On the off-host processing host, shut down the replica database, and use the following command to unmount each of the snapshot volumes:

```
# umount mount_point
```

- 2 On the off-host processing host, use the following command to deport the snapshot volume's disk group:

```
# vxdg deport snapvoldg
```

- 3 On the primary host, re-import the snapshot volume's disk group using the following command:

```
# vxdg [-s] import snapvoldg
```

Note: Specify the `-s` option if you are reimporting the disk group to be rejoined with a shared disk group in a cluster.

- 4 On the primary host, use the following command to rejoin the snapshot volume's disk group with the original volume's disk group:

```
# vxdg join snapvoldg  
          database_dg
```

- 5 VxVM will recover the volumes automatically after the join unless it is set to not recover automatically. Check if the snapshot volumes are initially disabled and not recovered following the join.

If a volume is in the DISABLED state, use the following command on the primary host to recover and restart the snapshot volume:

```
# vxrecover -g database_dg -m snapvol
```

- 6 Use the steps in [Creating an off-host replica database](#) to resynchronize the snapshot and make the snapshot available at off-host processing host again.

The snapshots are now ready to be re-used for backup or for other decision support applications.

Updating a warm standby Sybase ASE 12.5 database

If you specified the `for external dump` clause when you quiesced the primary database, and you started the replica database by specifying the `-q` option to the `dataserver` command, you can use transaction logs to update the replica database.

To update the replica database

- 1 On the primary host, use the following `isql` command to dump the transaction log for the database:

```
dump transaction to dump_device with standby_access
```

Copy the transaction log dump to the appropriate database directory on the off-host processing host.

- 2 On the off-host processing host, use the following `isql` command to load the new transaction log:

```
load transaction from dump_device with standby_access
```

- 3 On the off-host processing host, use the following `isql` command to put the database online:

```
online database database_name for standby_access
```

Reattaching snapshot plexes

Some or all plexes of an instant snapshot may be reattached to the specified original volume, or to a source volume in the snapshot hierarchy above the snapshot volume.

Note: This operation is not supported for space-optimized instant snapshots.

By default, all the plexes are reattached, which results in the removal of the snapshot. If required, the number of plexes to be reattached may be specified as the value assigned to the `nmirror` attribute.

Note: The snapshot being reattached must not be open to any application. For example, any file system configured on the snapshot volume must first be unmounted.

To reattach a snapshot

- ◆ Use the following command, to reattach some or all plexes of an instant snapshot to the specified original volume, or to a source volume in the snapshot hierarchy above the snapshot volume:

```
# vxsnap [-g diskgroup] reattach snapvol source=vol \  
[nmirror=number]
```

For example the following command reattaches 1 plex from the snapshot volume, `snapmyvol`, to the volume, `myvol`:

```
# vxsnap -g mydg reattach snapmyvol source=myvol nmirror=1
```

While the reattached plexes are being resynchronized from the data in the parent volume, they remain in the `SNAPTMP` state. After resynchronization is complete, the plexes are placed in the `SNAPDONE` state.

What is off-host processing?

Off-host processing consists of performing operations on application data on a host other than the one where the application is running. Typical operations include Decision Support Systems (DSS) and backup. In a VVR environment, off-host processing operations can be performed on the Secondary of the Replicated Data Set. This reduces the load on the application server, the Primary.

The model for data access on the Secondary is that you break off a mirror from each data volume in the RVG, perform the operation on the mirror, and then reattach the mirror while replication is in progress.

About using VVR for off-host processing

This chapter explains how to use Veritas Volume Replicator (VVR) for off-host processing on the Secondary host. You can use the In-Band Control (IBC) Messaging feature with the FastResync (FMR) feature of Veritas Volume Manager (VxVM) and its integration with VVR to take application-consistent snapshots at the replicated volume group (RVG) level. This lets you perform off-host processing on the Secondary host.

This chapter explains how to perform off-host processing operations using the `vradmin ibc` command. You can also use the `vxibc` commands to perform off-host processing operations.

Creating and refreshing test environments

This chapter includes the following topics:

- [About test environments](#)
- [Creating a test environment](#)
- [Refreshing a test environment](#)

About test environments

Sometimes, there is a need to do some testing or development on a copy of production data. In such scenarios, it is essential to provide isolation of these environments from the production environment. This is required so that there is no interference of testing or development environments with production environment. Veritas Storage Foundation can provide a very efficient and cost effective mechanism to create multiple test setups at the same time from a copy of production data. This is done without affecting performance of the production application, at the same time providing complete isolation.

Creating a test environment

Before you set up a test or development environment, you must have a production application volume already created in the application disk group.

To prepare for a test environment

- ◆ Prepare the application data volume(s) for snapshot operation

```
# vxsnap -g appdg prepare appvol
```

To create a test environment

- 1 Identify disks to create break-off snapshots. These disks need not be from the same array as the application volume. These disks must be visible to the host that will run test/dev environment.
- 2 Use these disks to create a mirror breakoff snapshot:
 - Add the mirror to create a breakoff snapshot. This step copies application volume data into the new mirror added to create the snapshot.

```
# vxsnap -g appdg addmir appvol alloc=<sdisk1,sdisk2,...>
```

- Create a snapshot.

```
# vxsnap -g appdg make src=appvol/nmirror=1/new=snapvol
```

- 3 Split the diskgroup containing the mirror breakoff snapshot.

```
# vxdg split appdg testdevdg snapvol
```

- 4 Deport the diskgroup from the production application host

```
# vxdg deport testdevdg
```

- 5 Import the *testdev* disk group on the host that will run the test environment.

```
# vxdg import testdevdg
```

Once this step is done, the *snapvol* present in the *testdevdg* disk group is ready to be used for testing or development purposes. If required, it is also possible to create multiple copies of *snapvol* using Veritas Storage Foundation's Flashsnap feature by creating a snapshot of *snapvol* using method described above.

Refreshing a test environment

Periodically, it may be required to resynchronize the test or development environment with current production data. This can be efficiently achieved using the Flashsnap feature of Storage Foundation and High Availability Solutions products.

To refresh a test environment

- 1 Deport the *testdevdg* disk group from the test environment. This step requires stopping the usage of *snapvol* in the test environment.

```
# vxdg deport testdevdg
```

- 2 Import *testdevdg* into the production environment.

```
# vxdg import testdevdg
```

- 3 Reattach the *snapvol* to *appvol* in order to synchronize current production data. Note that this synchronization is very efficient since it copies only the changed data.

```
# vxsnap -g appdg snapvol source=appvol
```

- 4 When you need to setup the *testdevdg* environment again, recreate the break-off snapshot.

```
# vxsnap -g appdg make src=appvol/nmirror=1/new=snapvol
```

- 5 Split the diskgroup containing the mirror breakoff snapshot.

```
# vxdg split appdg testdevdg snapvol
```

- 6 Deport the diskgroup from the production application host

```
# vxdg deport testdevdg
```

- 7 Import the *testdev* disk group on the host that will run the test environment.

```
# vxdg import testdevdg
```

Once this step is done, the *snapvol* present in *testdevdg* is ready to be used for testing or development purposes.

You can also create further snapshots of *snapvol* in order to create more test or development environments using the same snapshot. For this purpose, the following mechanisms can be used:

- Mirror breakoff snapshots
See [“Preparing a full-sized instant snapshot for a backup”](#) on page 96.
- Space-optimized snapshot
See [“Preparing a space-optimized snapshot for a database backup”](#) on page 98.
- Veritas File System level checkpoints

See [“Creating Storage Checkpoints”](#) on page 126.

For more detailed information, see the *Veritas Storage Foundation™ Administrator's Guide*

Creating point-in-time copies of files

This chapter includes the following topics:

- [Using FileSnaps to create point-in-time copies of files](#)

Using FileSnaps to create point-in-time copies of files

The key to obtaining maximum performance with FileSnaps is to minimize the copy-on-write overhead. You can achieved this by enabling lazy copy-on-write. Lazy copy-on-write is easy to enable and usually results in significantly better performance. If lazy copy-on-write is not a viable option for the use case under consideration, an efficient allocation of the source file can reduce the need of copy-on-write.

Using FileSnaps to provision virtual desktops

Virtual desktop infrastructure (VDI) operating system boot images are a good use case for FileSnaps. The parts of the boot images that can change are user profile, page files (or swap for UNIX/Linux) and application data. You should separate such data from boot images to minimize unsharing. You should allocate a single extent to the master boot image file.

Example of using FileSnaps to provision a virtual desktop

The following example uses a 4 GB master boot image that has a single extent that will be shared by all snapshots.

```
# touch /vdi_images/master_image
# /opt/VRTS/bin/setext -r 4g -f chgsize /vdi_images/master_image
```

The `master_image` file can be presented as a disk device to the virtual machine for installing the operating system. Once the operating system is installed and configured, the file is ready for snapshots.

Using FileSnaps to optimize write intensive applications for virtual machines

When virtual machines are spawned to perform certain tasks that are write intensive, a significant amount of unsharing can take place. Symantec recommends that you optimize performance by enabling lazy copy-on-write. If the use case does not allow enabling lazy copy-on-write, with careful planning, you can reduce the occurrence of unsharing. The easiest way to reduce unsharing is to separate the application data to a file other than the boot image. If you cannot do this due to the nature of your applications, then you can take actions similar to the following example.

Example of using FileSnaps to optimize write intensive applications

Assume that the disk space required for a boot image and the application data is 20 GB. Out of this, only 4 GB is used by the operating system and the remaining 16 GB is the space for applications to write. Any data or binaries that are required by each instance of the virtual machine can still be part of the first 4 GB of the shared extent. Since most of the writes are expected to take place on the 16 GB portion, you should allocate the master image in such a way that the 16 GB of space is not shared, as shown in the following commands:

```
# touch /vdi_images/master_image
# /opt/VRTS/bin/setext -r 4g -f chgsize /vdi_images/master_image
# dd if=/dev/zero of=/vdi_images/master_image seek=16777215 \
bs=1024 count=1
```

The last command creates a 16 GB hole at the end of the file. Since holes do not have any extents allocated, the writes to hole do not need to be unshared.

Using FileSnaps to create multiple copies of data instantly

It is common to create one or more copies of production data for the purpose of generating reports, mining, and testing. These cases frequently update the copies of the data with the most current data, and one or more copies of the data always exists. FileSnaps can be used to create multiple copies instantly. The application that uses the original data can see a slight performance hit due to the unsharing of data that can take place during updates. This slight impact on performance

can still be present even when all FileSnaps have been deleted. However, you rarely see all FileSnaps being deleted since these use cases usually have one or more copies at any given time.

4

Section

Maximizing storage utilization

- [Chapter 12. Optimizing storage tiering with SmartTier](#)

Optimizing storage tiering with SmartTier

This chapter includes the following topics:

- [About SmartTier](#)
- [SmartTier building blocks](#)
- [SmartTier use cases for DB2 or Sybase](#)
- [Setting up a filesystem for storage tiering with SmartTier](#)
- [Relocating old archive logs to tier two storage using SmartTier](#)
- [Relocating inactive tablespaces or segments to tier two storage](#)
- [Relocating active indexes to premium storage](#)
- [Relocating all indexes to premium storage](#)

About SmartTier

SmartTier matches data storage with data usage requirements. After data matching, the data can then be relocated based upon data usage and other requirements determined by the storage or database administrator (DBA).

As more and more data is retained over a period of time, eventually, some of that data is needed less frequently. The data that is needed less frequently still requires a large amount of disk space. SmartTier enables the database administrator to manage data so that less frequently used data can be moved to slower, less expensive disks. This also permits the frequently accessed data to be stored on faster disks for quicker retrieval.

Tiered storage is the assignment of different types of data to different storage types to improve performance and reduce costs. With SmartTier, storage classes are used to designate which disks make up a particular tier. There are two common ways of defining storage classes:

Note: SmartTier is the expanded and renamed feature previously known as Dynamic Storage Tiering (DST).

- Performance, or storage, cost class: The most-used class consists of fast, expensive disks. When data is no longer needed on a regular basis, the data can be moved to a different class that is made up of slower, less expensive disks.
- Resilience class: Each class consists of non-mirrored volumes, mirrored volumes, and n-way mirrored volumes.
For example, a database is usually made up of data, an index, and logs. The data could be set up with a three-way mirror because data is critical. The index could be set up with a two-way mirror because the index is important, but can be recreated. The redo and archive logs are not required on a daily basis but are vital to database recovery and should also be mirrored.

SmartTier policies control initial file location and the circumstances under which existing files are relocated. These policies cause the files to which they apply to be created and extended on specific subsets of a file systems's volume set, known as placement classes. The files are relocated to volumes in other placement classes when they meet specified naming, timing, access rate, and storage capacity-related conditions.

In addition to preset policies, you can manually move files to faster or slower storage with SmartTier, when necessary. You can also run reports that list active policies, display file activity, display volume usage, or show file statistics.

To use SmartTier, your storage must be managed using the following features:

- VxFS multi-volume file system
- VxVM volume set
- Volume tags
- SmartTier management at the file level
- SmartTier management at the sub-file level

SmartTier building blocks

To use SmartTier, your storage must be managed using the following features:

- VxFS multi-volume file system
- VxVM volume set
- Volume tags
- SmartTier management at the file level
- SmartTier management at the sub-file level

About VxFS multi-volume file systems

Multi-volume file systems are file systems that occupy two or more virtual volumes. The collection of volumes is known as a volume set, and is made up of disks or disk array LUNs belonging to a single Veritas Volume Manager (VxVM) disk group. A multi-volume file system presents a single name space, making the existence of multiple volumes transparent to users and applications. Each volume retains a separate identity for administrative purposes, making it possible to control the locations to which individual files are directed.

This feature is available only on file systems meeting the following requirements:

- The minimum disk group version is 140.
- The minimum file system layout version is 7 for file level SmartTier.
- The minimum file system layout version is 8 for sub-file level SmartTier.

To convert your existing VxFS system to a VxFS multi-volume file system, you must convert a single volume to a volume set.

The VxFS volume administration utility (`fsvoladm` utility) can be used to administer VxFS volumes. The `fsvoladm` utility performs administrative tasks, such as adding, removing, resizing, encapsulating volumes, and setting, clearing, or querying flags on volumes in a specified Veritas File System.

See the `fsvoladm` (1M) manual page for additional information about using this utility.

About VxVM volume sets

Volume sets allow several volumes to be represented by a single logical object. Volume sets cannot be empty. All I/O from and to the underlying volumes is directed via the I/O interfaces of the volume set. The volume set feature supports the multi-volume enhancement to Veritas File System (VxFS). This feature allows file systems to make best use of the different performance and availability characteristics of the underlying volumes. For example, file system metadata could be stored on volumes with higher redundancy, and user data on volumes with better performance.

About volume tags

You make a VxVM volume part of a placement class by associating a volume tag with it. For file placement purposes, VxFS treats all of the volumes in a placement class as equivalent, and balances space allocation across them. A volume may have more than one tag associated with it. If a volume has multiple tags, the volume belongs to multiple placement classes and is subject to allocation and relocation policies that relate to any of the placement classes.

Warning: Multiple tagging should be used carefully.

A placement class is a SmartTier attribute of a given volume in a volume set of a multi-volume file system. This attribute is a character string, and is known as a volume tag.

SmartTier use cases for DB2 or Sybase

Storage Foundation High Availability Solutions include SmartTier, a storage tiering feature which enables you to tier your data to achieve optimal use of your storage.

Example procedures illustrate the following use cases:

- Relocating archive logs older than 2 days to Tier-2 storage
- Relocating inactive tablespaces or segments to Tier-2 storage
- Relocating active indexes to Tier-0 storage
- Relocating all indexes to Tier-0 storage

Setting up a filesystem for storage tiering with SmartTier

In the use case examples, the following circumstances apply:

- The database containers are in the file system */DBdata*
- The database archived logs are in the file system */DBarch*

To create required filesystems for SmartTier

1 List the disks:

```
# vxdisk list
```

DEVICE	TYPE	DISK	GROUP	STATUS
fas30700_0	auto:cdsdisk	fas30700_0	---	online thin
fas30700_1	auto:cdsdisk	fas30700_1	---	online thin
fas30700_2	auto:cdsdisk	fas30700_2	---	online thin
fas30700_3	auto:cdsdisk	fas30700_3	---	online thin
fas30700_4	auto:cdsdisk	fas30700_4	---	online thin
fas30700_5	auto:cdsdisk	fas30700_5	---	online thin
fas30700_6	auto:cdsdisk	fas30700_6	---	online thin
fas30700_7	auto:cdsdisk	fas30700_7	---	online thin
fas30700_8	auto:cdsdisk	fas30700_8	---	online thin

Assume there are 3 LUNs on each tier.

2 Create the disk group.

```
# Vxdg init DBdg fas30700_0 fas30700_1 fas30700_2 \
fas30700_3 fas30700_4 fas30700_5 fas30700_6 fas30700_7 \
fas30700_8
```

3 Create the volumes *datavol* and *archvol*.

```
# vxassist -g DBdg make datavol 200G alloc=fas30700_3,\
fas30700_4,fas30700_5
# vxassist -g DBdg make archvol 50G alloc=fas30700_3,\
fas30700_4,fas30700_5
```

Tag *datavol* and *archvol* as tier-1.

```
# vxassist -g DBdg settag datavol vxfs.placement_class.tier1
# vxassist -g DBdg settag archvol vxfs.placement_class.tier1
```

4 Create the Tier-0 volumes.

```
# vxassist -g DBdg make tier0_vol1 50G alloc=fas30700_0,\
fas30700_1,fas30700_2
# vxassist -g DBdg make tier0_vol2 50G alloc=fas30700_0,\
fas30700_1,fas30700_2
# vxassist -g DBdg settag tier0_vol1 vxfs.placement_class.tier0
# vxassist -g DBdg settag tier0_vol2 vxfs.placement_class.tier0
```

5 Create the Tier-2 volumes.

```
# vxassist -g DBdg make tier2_vol1 50G alloc= fas30700_6,\
fas30700_7,fas30700_8
# vxassist -g DBdg make tier2_vol2 50G alloc= fas30700_6,\
fas30700_7,fas30700_8
# vxassist -g DBdg settag tier2_vol1 vxfs.placement_class.tier2
# vxassist -g DBdg settag tier2_vol2 vxfs.placement_class.tier2
```

6 Convert *datavol* and *archvol* to a volume set.

```
# vxvset -g DBdg make datavol_mvfs datavol
# vxvset -g DBdg make archvol_mvfs archvol
```

7 Add the volumes *Tier-0* and *Tier-2* to *datavol_mvfs*.

```
# vxvset -g DBdg addvol datavol_mvfs tier0_vol1
# vxvset -g DBdg addvol datavol_mvfs tier2_vol1
```

8 Add the volume *Tier-2* to *archvol_mvfs*

```
# vxvset -g DBdg archvol_mvfs tier2_vol2
```

9 Make the file system and mount *datavol_mvfs* and *archvol_mvfs*.

```
# mkfs -F vxfs /dev/vx/rdisk/DBdg/datavol_mvfs
```

10 Mount the *DBdata* file system

```
# mount -F vxfs /dev/vx/rdisk/DBdg/datavol_mvfs /DBdata
```

11 Mount the *DBarch* filesystem

```
# mount -F vxfs /dev/vx/rdisk/DBdg/datavol_mvfs /DBarch
```

12 Migrate the database into the newly created, SmartTier-ready file system. You can migrate the database either by restoring from backup or copying appropriate files into respective filesystems.

See the database documentation for more information.

Relocating old archive logs to tier two storage using SmartTier

A busy database can generate few hundred gigabytes of archive logs per day. Restoring these archive logs from tape backup is not ideal because it increases database recovery time. Regulatory requirements could mandate that these archive logs be preserved for several weeks.

To save storage costs, you can relocate archive logs older than two days (for example) into tier two storage. To achieve this you must create a policy file, for example, `archive_policy.xml`.

Note: The relocating archive logs use case applies for both DB2 and Sybase environments.

To relocate archive logs that are more than two days old to Tier-2

1 Create a policy file. A sample XML policy file is provided below.

```
<?xml version="1.0"?>
<!DOCTYPE PLACEMENT_POLICY SYSTEM "/opt/VRTSvxfs/etc\
  /placement_policy.dtd">
<PLACEMENT_POLICY Version="5.0" Name="access_age_based">
  <RULE Flags="data" Name="Key-Files-Rule">
    <COMMENT>
      This rule deals with key files such as archive logs.
    </COMMENT>

    <SELECT Flags="Data">
      <COMMENT>
        You want all files. So choose pattern as '*'
      </COMMENT>
      <PATTERN> * </PATTERN>
    </SELECT>

    <CREATE>
      <ON>
        <DESTINATION>
          <CLASS> tier1 </CLASS>
        </DESTINATION>
      </ON>
    </CREATE>

    <RELOCATE>
      <TO>
        <DESTINATION>
          <CLASS> tier2 </CLASS>
        </DESTINATION>
      </TO>
      <WHEN>
        <ACCAGE Units="days">
          <MIN Flags="gt">2</MIN>
        </ACCAGE>
      </WHEN>
    </RELOCATE>

  </RULE>
</PLACEMENT_POLICY>
```

Notice the **ACCAGE** units in the **WHEN** clause.

- 2 To locate additional sample policy files, go to `/opt/VRTSvxfs/etc`.

The access age-based policy is appropriate for this use case. Pay attention to the `CREATE ON` and `RELOCATE TO` sections of the XML file.

To apply a policy file

- 1 As root, validate *archive_policy.xml*

```
# fsppadm validate /DBarch archive_policy.xml
```

- 2 If the validation process is not successful, correct the problem. Validate *archive_policy.xml* successfully before proceeding.

- 3 Assign the policy to /DBarch filesystem

```
# fsppadm assign /DBarch archive_policy.xml
```

- 4 Enforce the policy. The relocation of two day old archive logs happens when the enforcement step is performed. The policy enforcements must be done every day to relocate aged archive logs. This enforcement can be performed on demand as needed or by using a cron- like scheduler.

```
# fsppadm enforce /DBarch
```

Relocating inactive tablespaces or segments to tier two storage

It is general practice to use partitions in databases. Each partition maps to a unique tablespace. For example in a shopping goods database, the orders table can be partitioned into orders of each quarter. Q1 orders can be organized into *Q1_order_tbs tablespace* or *Q1_order_tbs segment*, Q2 order can be organized into *Q2_order_tbs*.

As the quarters go by, the activity on older quarter data decreases. By relocating old quarter data into Tier-2, significant storage costs can be saved. The relocation of data can be done when the database is online.

For the following example use case, the steps illustrate how to relocate Q1 order data into Tier-2 in the beginning of Q3. The example steps assume that all the database data is in /DBdata filesystem.

To prepare to relocate Q1 order data into Tier-2 storage for DB2

- 1 Obtain a list of containers belonging to *Q1_order_tbs*.

```
$ db2inst1$ db2 list tablespaces
```

- 2 Find the tablespace-id for the tablespace *Q1_order_tbs*.

```
$ db2inst1$ db2 list tablespace containers for <tablespace-id>
```

- 3 Find the path names for the containers and store them in file *Q1_order_files.txt*.

```
#cat Q1_order_files.txt
      NODE0000/Q1_order_file1.f
      NODE0000/Q1_order_file2.f
      ...
      NODE0000/Q1_order_fileN.f
```

To prepare to relocate Q1 order data into Tier-2 storage for Sybase

- 1 Obtain a list of datafiles belonging to segment *Q1_order_tbs*. System Procedures *sp_helpsegment* and *sp_helpdevice* can be used for this purpose.

```
sybsadmin$ sp_helpsegment Q1_order_tbs
```

Note: In Sybase terminology, a "tablespace" is same as a "segment."

- 2 Note down the device names for the segment *Q1_order_tbs*.
- 3 For each device name use the *sp_helpdevice* system procedure to get the physical path name of the datafile.

```
sybsadmin$ sp_helpdevice <device name>
```

- 4 Save all the datafile path names in *Q1_order_files.txt*

```
# cat Q1_order_files.txt
      NODE0000/Q1_order_file1.f
      NODE0000/Q1_order_file2.f
      ...
      NODE0000/Q1_order_fileN.f
```

To relocate Q1 order data into Tier-2

- 1 Prepare a policy XML file. For the example, the policy file name is *Q1_order_policy.xml*. Below is a sample policy.

This policy is for unconditional relocation and hence there is no `WHEN` clause. There are multiple `PATTERN` statements as part of the `SELECT` clause. Each `PATTERN` selects a different file.

```
<?xml version="1.0"?>
<!DOCTYPE PLACEMENT_POLICY SYSTEM "/opt/VRTSvxfs/etc/\
placement_policy.dtd">
<PLACEMENT_POLICY Version="5.0" Name="selected files">
  <RULE Flags="data" Name="Key-Files-Rule">
    <COMMENT>
      This rule deals with key important files.
    </COMMENT>

    <SELECT Flags="Data">
      <DIRECTORY Flags="nonrecursive" > NODE0000</DIRECTORY>
      <PATTERN> Q1_order_file1.f </PATTERN>
      <PATTERN> Q1_order_file2.f </PATTERN>
      <PATTERN> Q1_order_fileN.f </PATTERN>
    </SELECT>

    <RELOCATE>
      <COMMENT>
        Note that there is no WHEN clause.
      </COMMENT>
      <TO>
        <DESTINATION>
          <CLASS> tier2 </CLASS>
        </DESTINATION>
      </TO>
    </RELOCATE>

  </RULE>
</PLACEMENT_POLICY>
```

- 2 Validate the policy *Q1_order_policy.xml*.

```
# fsppadm validate /DBdata Q1_order_policy.xml
```

- 3 Assign the policy.

```
# fsppadm assign /DBdata Q1_order_policy.xml
```

- 4 Enforce the policy.

```
# fsppadm enforce /DBdata
```

Relocating active indexes to premium storage

Database transaction rate depends upon how fast indexes can be accessed. If indexes reside on slow storage, the database transaction rate suffers. Tier-0 storage is generally too expensive to be practical to relocate the entire table data to Tier-0. Indexes are generally much smaller in size and are created to improve the database transaction rate, hence it is more practical to relocate active indexes to Tier-0 storage. Using SmartTier you can move active indexes to Tier-0 storage.

For the following telephone company database example procedure, assume the *call_details* table has an index *call_idx* on the column *customer_id*.

To prepare to relocate *call_idx* to Tier-0 storage for DB2

- 1 Find the tablespace where *call_idx* resides.

```
$ db2inst1$ db2 connect to PROD
$ db2inst1$ db2 select index_tablespace from syscat.tables \
where tabname='call_details'
```

- 2 In this example, the index is in tablespace *tbs_call_idx*. To get the tablespace id for *tbs_call_idx* and the list of containers:

```
$ db2inst1$ db2 list tablespaces
```

Note the tablespace id for *tbs_call_idx*.

- 3 List the containers and record the filenames in the tablespace *tbs_call_idx*.

```
$ db2inst1$ db2 list tablespace containers for <tablespace-id>
```

- 4 Store the files in *index_files.txt*.

```
# cat index_files.txt
/DB2data/NODE0000/IDX/call11.idx
/DB2data/NODE0000/IDX/call12.idx
/DB2data/NODE0000/IDX/call13.idx
```


To prepare to relocate *call_idx* to premium storage for Sybase

- 1 Obtain a list of datafiles for the *call_idx* segment.

```
$ sybsadmin$ sp_helpsegment call_idx
```

- 2 Note down the device names for the segment *call_idx*.

- 3 For each device name use the `sp_helpdevice` system procedure to get the physical pathname of the datafile.

```
sybsadmin$ sp_helpdevice <device name>
```

- 4 Save all the datafile path names in *index_files.txt*.

```
# cat index_files.txt  
/DB2data/NODE0000/IDX/call11.idx  
/DB2data/NODE0000/IDX/call12.idx  
/DB2data/NODE0000/IDX/call13.idx
```

To relocate *call_idx* to Tier-0 storage

1 Prepare the policy *index_policy.xml*.

Example policy:

```
<?xml version="1.0"?>
<!DOCTYPE PLACEMENT_POLICY SYSTEM "/opt/VRTSvxfs/etc/\
placement_policy.dtd">
<PLACEMENT_POLICY Version="5.0" Name="selected files">
  <RULE Flags="data" Name="Key-Files-Rule">
    <COMMENT>
      This rule deals with key important files.
    </COMMENT>

    <SELECT Flags="Data">
      <DIRECTORY Flags="nonrecursive" > NODE0000</DIRECTORY>
      <PATTERN> call*.idx </PATTERN>
    </SELECT>

    <RELOCATE>
      <COMMENT>
        Note that there is no WHEN clause.
      </COMMENT>
      <TO>
        <DESTINATION>
          <CLASS> tier0 </CLASS>
        </DESTINATION>
      </TO>
    </RELOCATE>

  </RULE>
</PLACEMENT_POLICY>
```

2 Assign and enforce the policy.

```
# fsppadm validate /DBdata index_policy.xml
# fsppadm assign /DBdata index_policy.xml
# fsppadm enforce /DBdata
```

Relocating all indexes to premium storage

It is a common practice for DBAs to name index files with some common extensions. For example, all index files are named with “.inx” extensions. If your

Tier-0 storage has enough capacity, you can relocate all indexes of the database to Tier-0 storage. You can also make sure all index containers created with this special extension are automatically created on Tier-0 storage by using the `CREATE` and `RELOCATE` clause of policy definition.

To relocate all indexes to Tier-0 storage

1 Create a policy such as the following example:

```
# cat index_policy.xml

<?xml version="1.0"?>
<!DOCTYPE PLACEMENT_POLICY SYSTEM "/opt/VRTSvxfs/etc/\
placement_policy.dtd">
<PLACEMENT_POLICY Version="5.0" Name="selected files">
  <RULE Flags="data" Name="Key-Files-Rule">
    <COMMENT>
      This rule deals with key important files.
    </COMMENT>

    <SELECT Flags="Data">
      <PATTERN> *.inx </PATTERN>
    </SELECT>

    <CREATE>
      <COMMENT>
        Note that there are two DESTINATION.
      </COMMENT>
      <ON>
        <DESTINATION>
          <CLASS> tier0 </CLASS>
        </DESTINATION>
        <DESTINATION>
          <CLASS> tier1</CLASS>
        </DESTINATION>
      </ON>
    </CREATE>

    <RELOCATE>
      <COMMENT>
        Note that there is no WHEN clause.
      </COMMENT>
      <TO>
        <DESTINATION>
          <CLASS> tier0 </CLASS>
        </DESTINATION>
      </TO>
    </RELOCATE>

  </RULE>
</PLACEMENT_POLICY>
```

- 2 To make sure file creation succeeds even if Tier-0 runs out of space, add two `ON` clauses as in the example policy in [1](#).
- 3 Assign and enforce the policy.

```
# fsppadm validate /DBdata index_policy.xml  
# fsppadm assign /DBdata index_policy.xml  
# fsppadm enforce /DBdata
```


Migrating data

- [Chapter 13. Understanding data migration](#)
- [Chapter 14. Offline migration from Logical Volume Manager to Veritas Volume Manager](#)
- [Chapter 15. Online migration of native LVM volumes to VxVM volumes](#)
- [Chapter 16. Migrating storage arrays](#)
- [Chapter 17. Migrating data between platforms](#)

Understanding data migration

This chapter includes the following topics:

- [Types of data migration](#)

Types of data migration

This section describes the following types of data migration:

- **Migrating data from LVM to Storage Foundation using offline migration**
When you install Storage Foundation, you may already have some volumes that are controlled by the Logical Volume Manager. You can preserve your data and convert these volumes to Veritas Volume Manager volumes.
See [“About migrating Logical Volume Manager to VxVM”](#) on page 195.
- **Migrating data from native LVM volume to a VxVM volume using online migration**
You may have some volumes that are controlled by the native Logical Volume Manager. You can migrate your data from these volumes to Veritas Volume Manager volumes, with limited application downtime, and an ability to roll back if necessary.
See [“About online migration from Logical Volume Manager \(LVM\) volumes to Veritas Volume Manager \(VxVM\) volumes”](#) on page 233.
You can migrate LVM volumes in a standalone environment to VxVM with online migration.
See [“Online migration from LVM volumes in standalone environment to VxVM volumes”](#) on page 235.
You can also migrate LVM volumes under a VCS HA cluster environment to VxVM with online migration.

See [“Online migration from LVM volumes in VCS HA environment to VxVM volumes”](#) on page 245.

- Migrating data between platforms using Cross-platform Data Sharing (CDS)
Storage Foundation lets you create disks and volumes so that the data can be read by systems running different operating systems. CDS disks and volumes cannot be mounted and accessed from different operating systems at the same time. The CDS functionality provides an easy way to migrated data between one system and another system running a different operating system.
See [“Overview of the Cross-Platform Data Sharing \(CDS\) feature”](#) on page 267.
- Migrating data between arrays
Storage Foundation supports arrays from various vendors. If your storage needs change, you can move your data between arrays.
See [“Array migration for storage using Linux”](#) on page 255.

Note: The procedures are different if you plan to migrate to a thin array from a thick array.

Offline migration from Logical Volume Manager to Veritas Volume Manager

This chapter includes the following topics:

- [About migrating Logical Volume Manager to VxVM](#)
- [About LVM to VxVM conversion](#)
- [Converting unused LVM physical volumes to VxVM disks](#)
- [Converting LVM volume groups to VxVM disk groups](#)
- [Restoring the LVM volume group configuration](#)
- [Examples](#)
- [General information regarding conversion speed](#)
- [Non-interactive conversion of volume groups](#)

About migrating Logical Volume Manager to VxVM

This section provides an overview of migrating Logical Volume Manager to VxVM.

There are benefits of migrating from the HP-UX Logical Volume Manager (LVM) to VxVM.

See [“About LVM to VxVM conversion”](#) on page 196.

About LVM to VxVM conversion

The basic tools for conversion are the VxVM commands, `vxvmconvert` and `vxdiskadm`, and the LVM administrative utilities such as `pvremove` and `vgcfgbackup`. The discussion here details how to use these tools and gives some insights into how these tools work.

The disks on your system that are managed by LVM can be of two types:

- Unused disks, which contain no user data, and are not used by any volume group, but which have LVM disk headers.
For unused LVM disks you can use a combination of `pvremove` and `vxdiskadm`. See the man page `vxdiskadm(1M)`.
- LVM disks in volume groups, and which contain logical volumes and volume groups.
For LVM disks in volume groups, the primary tool for conversion is the `vxvmconvert` command.
The `vxvmconvert` utility is an interactive command. You can also use the `vxautoanalysis` command and the `vxautoconvert` command to perform non-interactive analysis and conversion of LVM volume groups. The `vxautorollback` command also lets you reverse the conversion, and turn a disk group back into a volume group.
See [“Non-interactive conversion of volume groups”](#) on page 230.

Converting unused LVM physical volumes to VxVM disks

LVM disks which are not part of any volume group, and contain no user data, are simply cleaned up, so that there are no LVM disk headers. Then the disks are given over to VxVM through the normal means of initializing disks.

Warning: Exercise caution while using this procedure to give disks over to VxVM. You must be absolutely certain that the disks are not in use in any LVM configuration. If there is any user data on these disks, it will be lost during conversion.

Removing LVM disk information

To remove LVM disk information from the disks use the following command:

```
# pvremove disk_name
```

The `pvremove` command does not allow the removal of disk headers which indicate a Volume Group membership for the disk. If the disk fails `pvremove` for this reason, you should first make certain that the group membership information is stale. Do this by using `pvdisplay`:

```
# pvdisplay disk_name
```

If `pvdisplay` finds no valid group information associated with the disk, you can overwrite the LVM headers using `pvcreate`:

```
# pvcreate disk_name
```

Warning: If `disk_name` is an alternate path to a disk that does not appear in the `lvmtab` file for this system, or is a disk that is in use on another system, but not imported onto this system, then do not `pvcreate -f` the disk. Doing so will destroy the LVM headers. You can use `pvcreate` without the `-f` option safely, as it will only succeed if the disk is not listed in the `/etc/lvmtab` file, and if the LVM headers indicate that it is not a member of any volume group. (That is, `pvcreate` has been run on the disk, but neither `vgcreate` nor `vgextend` have.)

Initializing disks for VxVM use

To initialize the disk for VxVM use, use the `vxdiskadm` command, selecting the option:

```
1) Add or initialize one or more disks
```

Or use the command:

```
# vxdisksetup -i disk_name
```

VxVM utilities will not tamper with any disks that are recognized as owned by LVM (by virtue of the LVM disk headers). If you attempt to use `vxdisksetup` or `vxdiskadm` on an LVM disk without using the `pvremove` command first, the command fails.

Note: The above behavior is displayed on both LVM version 1 and version 2 volume groups.

Converting LVM volume groups to VxVM disk groups

This section outlines the process for converting LVM volume groups to VxVM disk groups.

Note: Symantec recommends that you read through this section carefully before you begin any volume group conversion.

The conversion process involves many steps. Though there are tools to help you with the conversion, some of these steps cannot be automated. You should be sure to understand how the whole conversion process works, and what you will need to do in the process before beginning a volume group conversion.

The tool used for conversion is `vxvmconvert`. This interactive, menu-driven program walks you through many of the steps of the process of converting volume groups for use by VxVM. Using `vxvmconvert` can reduce the downtime associated with converting from LVM to VxVM. Without the `vxvmconvert` tool, the only possible method of conversion would be to take full backups of user data, destroy the existing LVM configuration leaving only raw disks, recreate the configuration in VxVM, and then reload the user data.

The `vxvmconvert` process converts LVM volume groups to VxVM disk groups in place. In reality, the utility changes disks within LVM volume groups to VxVM disks by taking over the areas of the disks used for LVM configuration information, and creating the equivalent VxVM volume configuration information. User data, the portions of the disks used for file systems, databases and so on, are not affected by the conversion.

The act of conversion changes the names by which your system refers to the logical storage. For this reason, the conversion process is necessarily an off-line one. There can be no application access to user data in the volume groups undergoing conversion. Access to the LVM configuration itself (the metadata of LVM) must also be limited to the conversion process.

Volume group conversion limitations

There are certain LVM volume configurations that cannot be converted to VxVM. Some of the reasons a conversion could fail are:

- A volume group with insufficient space for metadata.

In the conversion of LVM to VxVM, the areas of the disks used to store LVM metadata are overwritten with VxVM metadata. If the VxVM metadata that needs to be written will not fit the space occupied by the LVM metadata, the group containing the disk cannot be converted. If you have just enough space

for the conversion, you probably would want to have more space for future configuration changes.

Note: The most likely scenario in which a Volume Group cannot be converted, because of insufficient private space, is when a large HP-UX system using “Extent based Striping” is being used

- A volume group containing the root volume.

The `vxvmconvert` command does not convert any volume group that contains a rootable volume, identified by the presence of the LIF area as created by `mkboot(1M)`. Not only is the current root volume off limits, but any volume that might be used as an alternate root volume is rejected as well.

Note: You can use the `vxcp_lvmroot(1M)` command to create a VxVM root disk on a spare physical disk from the contents of the existing LVM root disk. Similarly, you can use the `vxres_lvmroot(1M)` command to recreate an LVM root disk on a spare disk from the contents of the VxVM root disk.

For more information about the VxVM root disk, see the *Veritas Storage Foundation Administrator's Guide*.

- A volume group containing mirrors using the Mirror Write Cache (MWC) feature for volume consistency recovery.
Users should be aware that when converting mirrored LVM volumes to VxVM, some of these volumes will likely have the Mirror Write Cache consistency recovery method in force on the volume. The `vxvmconvert` utility can convert these volumes, but must use the Dirty Region Logging (DRL) feature to obtain the same level of functionality. However, since Dirty Region Logging requires some user space to be available for the log, a conversion could fail due to an MWC volume being full, leaving no space for the DRL log. However it is very unlikely that this situation would occur. Note that the MWC and DRL are used only when the system crashes or is improperly shut down, to quickly bring all mirrors in the volume back into a consistent state.
- A volume group containing the `/usr` file system.
For this release, a volume group containing the `/usr` file system cannot be converted because `vxvmconvert` needs access to files in `/usr`.
- Volume groups with any `dump` or primary `swap` volumes.
Because this release does not support rootability, `vxvmconvert` will not convert any volume group with `dump` or primary `swap` volumes. These are volumes

known to the boot process. However, `swap` volumes that belong to a different volume group than the root volume can be converted.

- Volume group disks used in MC/ServiceGuard clusters.
The conversion process does not support conversion of any volume group that is marked as a member of a MC/ServiceGuard or OPS Edition high availability cluster. The volume group must be deactivated and removed from membership in the high availability cluster before it can be converted.
- Volume groups and logical volumes with large number of characters in the name.
If you convert an LVM volume group with more than 31 characters in the volume group name, VxVM will retain only the first 31 characters of the LVM volume group name. Conversion of a logical volume containing more than 28 characters in the logical volume name is not supported.
- Volume groups or logical volumes containing special characters
Conversion of volume groups or logical volumes containing special characters like #, \$, *, _ in the name, is not supported.
- Volume groups used for cluster lock disks.
The conversion process does not support conversion of a volume group that contains a disk that is being used for a cluster lock disk for an MC/ServiceGuard cluster.
- Volume groups with any disks that have bad blocks in the bad block directory.
Unlike LVM, VxVM does not support bad block revectoring at the physical volume level. If there appear to be any valid bad blocks in the bad block directory of any disk used in an LVM volume group, the group cannot be converted.
- Volume groups with mirrored volumes.
A conversion fails if the LVM volume group being converted has mirrored volumes, but the system does not have a valid license installed that enables mirroring for VxVM.

The `analyze` option in `vxvmconvert`, which is described in later sections, aids you in identifying which volume groups can be converted.

Conversion process summary

Several steps are used to convert LVM volume groups to VxVM disk groups. Most of these steps can be done with the `vxvmconvert` utility. All the steps are not compulsory, and some may have to be followed only if there are problems during conversion. Some of them (e.g. backing up user data) are left to you to accomplish through your regular administrative processes.

The steps in the conversion process are:

- Identifying LVM volume groups for conversion.
- Analyzing an LVM group to see if conversion is possible.
- Taking actions to make conversion possible if analysis fails.
- Backing up your LVM configuration and user data.
- Planning for new VxVM logical volume names.
- Stopping application access to volumes in the volume group to be converted.
- Converting a volume group.
- Taking actions if conversion fails.
- Implementing changes for new VxVM logical volume names
- Restarting applications on the new VxVM volumes.
- Tailoring your VxVM configuration.

These steps are described in detail in later sections of this chapter. Also available are annotated examples on how to use `vxvmconvert`.

See [“Examples”](#) on page 213.

Additional information is available on restoring back to your original LVM configuration.

See [“Restoring the LVM volume group configuration”](#) on page 209.

Identifying LVM disks and volume groups for conversion

The obvious first step in the conversion process is to identify what you want to convert. The native LVM administrative utilities like `vgdisplay` and `SMH` can help you identify candidate LVM volume groups as well as the disks that comprise them.

You can also use the `vxvmconvert` command and the `vxdisk` command to examine groups and their member disks. The information presented through the `vxvmconvert` command and the `vxdisk` command and their interpretation are available.

See [“Examples”](#) on page 213.

You can also list the LVM disks with the following VxVM command:

```
# vxdisk list
```

Analyzing an LVM volume group to see if conversion is possible

After you have selected a volume group for conversion, you need to analyze it to determine if conversion for VxVM use is possible.

Use the analyze option of `vxvmconvert` to check for problems that would prevent the conversion from completing successfully. Additional information is available on all the conditions that this option checks for.

See [“Volume group conversion limitations”](#) on page 198.

The analysis calculates the space required to add the volume group disks to a VxVM disk group, and to replace any existing disks and volumes with VxVM volumes, plexes, and subdisks. If you don't have the required space to convert the disks, the conversion would fail.

Analysis can be run on a live system while users are accessing their data. To analyze LVM volume groups, choose option 1 of the `vxvmconvert` utility.

Note: The analysis option is presented as a separate menu item in `vxvmconvert`, but there is an implicit analysis with any conversion. If you simply select the “Convert LVM Volume Groups to VxVM” menu option, `vxvmconvert` will go through analysis on any group you specify. When you are using the convert option directly, you are given a chance to abort the conversion after analysis, and before any changes are committed to disk.

The analysis option is useful when you have a large number of groups/disks for conversion and some amount of planning is needed before the actual conversion. Installations with many users or critical applications can use the analyze option on a running system. Then conversion downtime can be better planned and managed. Smaller configurations may be better served by using the convert option directly while in a downtime period.

More information is available on converting disk groups.

See [“Converting LVM volume groups to VxVM disk groups”](#) on page 198.

Sample examples of the analyze option are available.

See [“Examples”](#) on page 213.

Taking actions to make conversion possible if analysis fails

Available is a list of reasons that may cause analysis to fail.

See [“Volume group conversion limitations”](#) on page 198.

Messages from `vxvmconvert` will explain the type of failure and any actions that can be taken before retrying the analysis.

Complete details on specific error messages and actions are available.

Backing up your LVM configuration and user data

After analysis you know which volume group or groups you want to convert to VxVM disk groups. Up to this point, you have not altered your LVM configuration.

By taking the next step (completing the conversion to VxVM), you are significantly changing access to your storage.

Although the conversion process does not move, or in any other way affect user data, you are strongly encouraged to back up all data on the affected disks. Similarly, you should back up the LVM configuration itself.

During a conversion, any spurious reboots, power outages, hardware errors or operating system bugs can have unpredictable and undesirable consequences. You are advised to be on guard against disaster with a set of verified backups.

Backing up an LVM configuration

Use the `vgcfgbackup(1M)` utility before running `vxvmconvert` to save a copy of the LVM configuration.

You can back up the LVM volumes using the following command:

```
# vgcfgbackup -f pathname/filename vol_grp_name
```

Be sure to use the `-f` option to save the data into a file other than the default. `vxvmconvert` uses LVM utilities which themselves save the configuration using `vgcfgbackup`. If you do not use the `-f` option when you attempt to back up the configuration, the conversion process overwrites your attempted backup.

Keep a copy of this LVM configuration offline on tape or some other medium for use in the event of a disaster during conversion.

For example, to put a copy on tape, use the following command:

```
# tar cvf /dev/rmt/c3t0d0BEST /vgbackups/vg08
```

Note: The `vxvmconvert` utility itself also saves a snapshot of the LVM metadata in the process of conversion for each disk. This data is saved in a different format from that of `vgcfgbackup`. It can only be used by the `vxvmconvert` program. With certain limitations, you can reinstate the LVM volumes after they have been converted to VxVM using this data. Even though `vxvmconvert` provides this level of backup of the LVM configuration, you are advised to use `vgcfgbackup` before running `vxvmconvert`.

See [“Example: displaying the vxvmconvert menu”](#) on page 213.

Backing up user data

To back up user data, use your regular backup processes.

Warning: Before you do the backup, you should carefully review how to implement changes for new VxVM logical volume names. Backup processes and systems themselves may have dependencies on the volume names currently in use on your system. The conversion to VxVM changes those names. You are advised to understand the implications name changes have for restoring from the backups you are about to make.

See [“Implementing changes for new VxVM logical volume names”](#) on page 208.

File system back up of user data

You can use the backup utility that you normally use to back up data on your logical volumes. For example, to back up logical volumes that contain file systems, the `fbbackup(1M)` command can be used to back up the data to tape.

For example, to back up the data on `/dev/vg01/lvol3` mounted on `/foodir`, use the following command:

```
# fbbackup -0i /foodir -f /dev/rmt/c0t0d0BEST
```

Non-file system back up

If a logical volume you are converting does not contain a file system, and is being used directly by an application (such as a database application), use the backup facilities provided by the application. If no such facility exists, consider using the `dd` command.

Planning for new VxVM logical volume names

When you change from LVM volumes to VxVM volumes, the device names by which your system accesses data are changed. LVM creates device nodes for its logical volumes in `/dev` under directories named for the volume group. VxVM creates its device nodes in `/dev/vx/dsk` and `/dev/vx/rdsk`. When conversion is complete, the old LVM device nodes are gone from the system, and the system will access data on the device nodes in `/dev/vx`.

This change in names can present problems. Any application that refers to specific device node names will be at risk when these names change. Similarly, any files that record specific device node names for use by applications can be problematic.

The most obvious area where this problem arises is in `/etc/fstab`. To handle this problem, `vxvmconvert` will rewrite the `fstab` with the new VxVM names when conversion is done so that `fsck`, `mount`, and related utilities will behave as they did before the conversion.

There are potentially many other applications, though, that may be put at risk by the name changes in conversion. `vxvmconvert` cannot help with these. The system administrator must examine the mechanisms used in each of the following areas to see if they reference LVM device names:

- Databases run on raw logical devices may record the name of that device node.
- Backup systems may do device level backups based on device node names recorded in private files. Also labeling of the backups may record device names.
- Scripts run by `cron(1M)`.
- Other administrative scripts.

Workaround

`vxvmconvert` records a mapping between the names of the LVM device nodes and VxVM device nodes. This data can be used to create symbolic links from the old LVM volume to the new VxVM device names. The mapping is recorded in the file:

```
/etc/vx/reconfig.d/vgrecords/vol_grp_name/vol_grp_name.trans
```

This file provides information on how to proceed further to link the old LVM volume names to the new VxVM device names.

Warning: This method of resolving the naming problem has risks. The symbolic links can become stale. For example, if a database refers to `/dev/vx/rdisk/vol1` through a symbolic link `/dev/vg00/rvol1` (“the old LVM name”), and if the underlying VxVM volume configuration is changed in any way, the database could refer to a missing or different volume.

Note: You may want to use this symbolic link approach to ease the transition to VxVM. You can set up the symbolic links after the successful conversion to VxVM. Then, you can do the investigation on a case by case basis for each volume. When you are satisfied that there are no problems introduced by the name change, the symbolic link to that volume can be removed. You must be careful to maintain a static VxVM volume configuration during this transition period.

Over time, the ultimate goal should be that the underlying VxVM naming is used by all applications, and that there are no indirect references to those volumes.

Stopping application access to volumes in the volume group to be converted

No applications can be active on the LVM volume group undergoing conversion. Before attempting to convert any volume group, you must ensure that applications using that group are down. This involves stopping databases, unmounting file systems, etc.

Note: If you are converting a volume with swap space on it, the conversion requires a reboot. The swap space cannot be taken out of control of the operating system with a shutdown to single user mode.

`vxvmconvert` tries to unmount mounted file systems during the conversion. Bear in mind though, that `vxvmconvert` makes no attempt to close down running applications on those file systems, nor does it attempt to deal with applications (e.g., databases) running on raw LVM volumes.

See “[Conversion and reboot](#)” on page 207.

Note: It is strongly recommended that you do not rely on `vxvmconvert`'s mechanisms for unmounting file systems. Conversion will be simpler if you close applications, and unmount file systems before running `vxvmconvert`.

To unmount a file system, use the following command:

```
# umount file-system
```

Conversion and reboot

During conversion, after the analysis phase is complete, the disks to be converted are deemed to be conversion ready. The `vxvmconvert` program asks if you are ready to commit to the conversion changes. If you choose to complete the conversion, the system will try to unmount all of the associated mounted file systems, stop and export the volume group, and then install the VxVM configuration.

If `vxvmconvert` is unable to stop and export volume groups or unmount file systems, the conversion cannot be completed without rebooting the system. You will have the option of aborting the conversion or completing the conversion by rebooting the system. If you choose to reboot, `vxvmconvert` will trigger the completion of the conversion automatically, during reboot, when it can be guaranteed that no processes have access to the volumes that are being converted.

If you choose to abort rather than reboot to complete the conversion, `vxvmconvert` will return to the main menu.

Note: The LVM logical volumes to be converted must all be available to the `vxvmconvert` process. You should not deactivate the volume group or any logical volumes before running `vxvmconvert`.

To activate a volume group

If you are not certain if the LVM volumes or the corresponding volume groups are active, you can activate them with the following command:

```
# vgchange -a y vol_grp_name
```

Converting a volume group

To do the actual conversion of LVM volume groups to VxVM disk groups, choose option 2 of the `vxvmconvert` utility.

`vxvmconvert` will prompt for a name for the VxVM disk group that will be created to replace the LVM volume group you are converting. This is the only object naming that is done through `vxvmconvert`. Additional details are available on modifying VxVM volume names.

See [“Tailoring your VxVM configuration”](#) on page 209.

As described earlier, the volume groups selected for conversion are analyzed to ensure that conversion is possible.

See [“Analyzing an LVM volume group to see if conversion is possible”](#) on page 202.

After a successful analysis phase, `vxvmconvert` prompts you to commit to the change or abort the conversion. When you select to commit the conversion, the new VxVM metadata is written. Before the conversion is committed, the `vxvmconvert` operation displays the estimated required time as:

```
VxVM INFO V-5-2-4906
The expected time for convert is: 0 hrs 0 mins 7 secs.
```

Note: The time required for conversion is an estimate and is not a calculated time. The actual conversion time may differ depending on factors like number of CPUs, Memory, I/O throughput, etc.

More information is available on the details of the conversion process.

See [“Examples”](#) on page 213.

Taking actions if conversion fails

There are several reasons why conversion can fail.

See [“Volume group conversion limitations”](#) on page 198.

Messages from `vxvmconvert` explain the type of failure, and any actions you can take before retrying the conversion.

Complete details of specific error messages are available.

Implementing changes for new VxVM logical volume names

You must be sure that all applications and configuration files refer properly to the new VxVM logical volumes.

See [“Planning for new VxVM logical volume names”](#) on page 205.

Restarting applications on the new VxVM volumes

After the conversion to VxVM is complete, file systems can be mounted on the new devices and applications can be restarted.

If you unmounted file systems before you ran `vxvmconvert`, you need to remount them by the new volume names. `vxvmconvert` updated `/etc/fstab` with the new names. When you started `vxvmconvert`, you may have left file systems mounted

that are associated with the volumes you converted. `vxvmconvert` remounts these with the new VxVM volume names.

Tailoring your VxVM configuration

`vxvmconvert` provides a default name for naming the newly formed VxVM disk group during conversion only as an option. However, you will be given the choice of choosing your own VxVM disk group name. By default, `vxvmconvert` renames the LVM volume group by replacing the prefix `vg` in the volume group name with the prefix `dg`. For example, `vg08` would become `dg08`. If there is no `vg` in the LVM volume group name, `vxvmconvert` simply uses the same volume group name for its disk group.

The disks in the new VxVM disk group are given VxVM disk media names () based on this disk group name. Additional information is available on VxVM disk media names.

See `vxintro(1M)`.

If your new VxVM disk group is `dg08`, it will have VxVM disks with names like `dg0801`, `dg0802`, etc. The VxVM plexes within the logical volumes will be `dg0801-01`, `dg0801-02`, etc.

If you do not like the default object names generated by the conversion, use the standard VxVM utilities to rename these objects. See the `rename` option in the `vxedit(1M)` man page for more details on renaming the disk groups.

Note: You must only rename objects in the VxVM configuration after you are fully satisfied with that configuration. In particular, you should never use menu option 3 of `vxvmconvert` (Roll back) after name changes. If you have chosen to set up symbolic links to the VxVM volumes, avoid renaming VxVM objects. Additional information is available on setting up symbolic links. These symbolic links are made invalid if the underlying VxVM device node name changes.

See “[Planning for new VxVM logical volume names](#)” on page 205.

Restoring the LVM volume group configuration

In some circumstances, you may need to restore the LVM configuration that existed before you converted to VxVM with `vxvmconvert`. For example:

- If something went wrong during the conversion, such as a system crash or a disk crash that caused the conversion to be unworkable.

- If during a conversion only some of a set of volume groups converted successfully, then you may want to restore the LVM configuration for the entire set.

It is possible to restore the original LVM configuration in one of two ways, but both have limitations and restrictions. The method you use depends on if any changes have been made to the VxVM configuration since the conversion occurred. Any of the following actions changes the VxVM configuration:

- adding or removing disks
- adding or removing volume groups
- changing the names of VxVM objects

Restoration methods include:

- **rollback using `vxvmconvert`**
Use rollback only if the VxVM configuration has not changed since the conversion. This method restores the LVM configuration without the need for user data restoration. Additional information is available on the rollback method.
See [“Rollback to LVM using `vxvmconvert`”](#) on page 211.
- **restore user data using `vgrestore` and `frecover`**
This method is a full LVM restoration which is used to restore your user data from backup when the VxVM configuration has changed since the conversion was made. First of all, this method restores the original LVM configuration information (`vgrestore`), and then restores the original user data from the backup that was made before the conversion was done (`frecover`). Additional information is available on full LVM restoration.
See [“Full LVM restoration”](#) on page 212.

Note: Restoring user data using the `vgrestore` and `frecover` method will result in the loss of all user data changes made since the conversion, and the loss of all new volumes created since the conversion.

In other words, this method of restoring data will take you back to exactly where you were before the conversion was done.

However, if no new volumes have been created, and if none of the existing volumes have been resized, you can use the `vxvmconvert rollback` option to restore the original LVM configuration. If you use this method, any user data changes made since the conversion will be retained, and you will not need to carry out a user level data restore (`frecover`).

The name changes that `vxvmconvert` makes as part of the conversion are managed by rollback, and do not count as VxVM configuration changes for the purposes of choosing a restoration method.

The `vgrestore` command should not be confused with the LVM command, `vgcfgrestore`. `vgcfgrestore` is used to restore the LVM configuration information saved by `vgcfgbackup`, but it will not restore your device files and `/etc/fstab` entries. It also will not import and activate the volume group, nor will it clean up any VxVM information left around. However, `vgrestore` will do all of this for you.

Rollback to LVM using `vxvmconvert`

Rollback replaces the VxVM disk groups with the original LVM volume groups. During conversion, `vxvmconvert` saves a “snapshot” of the original LVM metadata and associated configuration files, such as `/etc/fstab` and LVM device files. It restores only the LVM metadata and configuration files from this snapshot; user data is not changed. This method can only be used if no changes have been made to the configuration since the conversion.

For example, if a disk has been added to the disk group or if the names of any logical volumes have changed, you cannot use the rollback method.

Note: In many cases, if you choose the rollback method and the configuration has changed, you receive an error and must use the full restore method..

If you used the workaround of creating symbolic links from the old LVM names to the new VxVM names, you must remove the symbolic links you created before beginning the rollback. Additional information is available on creating symbolic links.

See [“Planning for new VxVM logical volume names”](#) on page 205.

This “snapshot” is kept on the root file system. The presence of this snapshot should not be taken as assurance that full off-line backups will not be needed. Specific information is available on backups.

See [“Backing up your LVM configuration and user data”](#) on page 203.

To rollback to LVM from the VxVM conversion, run `vxvmconvert` and choose option 3. An illustration of this method is available.

See [“Example: VxVM to LVM rollback”](#) on page 228.

Warning: Do not use this option unless you are certain that you want to restore LVM volume groups. After this is run, the VxVM disks that were created as a result of the original conversion from LVM to VxVM no longer exists. This option is not a full complement to `vxvmconvert`. It simply writes the saved LVM metadata back on top of the disks. Those data can only be considered valid for the period of time when the logical volumes are offline. If the VxVM configuration has been brought online, the metadata in the rollback snapshot should be considered obsolete. Specific information on full LVM restoration is available.

See [“Full LVM restoration”](#) on page 212.

Full LVM restoration

If you need to restore the original LVM configuration, but changes have been made to the VxVM configuration, you cannot use the rollback option of `vxvmconvert`. In this case, you must restore the user data in addition to restoring the old LVM metadata and associated configuration files. You may need to use this method if the disks in use by the LVM/VxVM volumes were corrupted during or after conversion.

Note: The snapshot of LVM internal data is kept on the root file system.

To use this method, you must have backed up data located on all the volume groups' logical volumes before conversion to VxVM.

Restoration of LVM volume groups is a two-step process consisting of a restoration of LVM internal data (metadata and configuration files), and a restoration of user or application data.

The process is limited to restoring the state of the logical volumes as they existed before conversion to VxVM disks. If the data has changed on the volumes during the time they were VxVM volumes, those changes are lost after you restore the LVM configuration and saved user data.

To do a full restoration of the original LVM configuration, do the following:

- 1 Use `vgrestore` to restore LVM internal data.

```
# vgrestore vol_grp_name
```

- 2 Use the recovery method to restore user or application data. In preparation for conversion, the recovery method should have been done with the standard backups you made in preparation for conversion. The following example shows an `frecover` from the `fbackup` example. Additional information on the `fbackup` example is available.

See [“Backing up your LVM configuration and user data”](#) on page 203.

```
# mount -F vxfs /dev/vg01/lvol3 /foodir  
# frecover -r -f /dev/rmt/c0t0d0BEST
```

Examples

The following are annotated examples on how to use `vxvmconvert` command.

Example: displaying the `vxvmconvert` menu

To display the `vxvmconvert` menu, use the following command:

```
# vxvmconvert
```

The following menu is displayed:

```
Volume Manager Support Operations  
Menu: Volume Manager/LVM_Conversion  
  
1      Analyze LVM Volume Groups for Conversion  
2      Convert LVM Volume Groups to VxVM
```

```
3      Roll back from VxVM to LVM
list   List disk information
listvg List LVM Volume Group information
?      Display help about menu
??     Display help about the menuing system
q      Exit from menus
```

Example: listing disk information

The `list` option of `vxvmconvert` displays information about the disks on a system. Select the `list` option from the `vxvmconvert` Main Menu:

```
Menu: Volume Manager/LVM_Conversion/list
```

```
# list
```

Use this menu option to display a list of disks. You can also choose to list detailed information about a disk by entering a specific disk device address.

```
Enter disk device or "all" [<address>,all,q,?](default: all) x
```

DEVICE	DISK	GROUP	STATUS
disk10	-	-	online
disk11	-	-	LVM
disk12	-	-	LVM
disk13	mydg01	mydg	online
disk14	-	-	online

```
Device to list in detail [<address>,none,q,?] (default: none)
```

```
none
```

Example: listing LVM volume group information

To list LVM volume group information, use the `listvg` option of `vxvmconvert`.

Select the `listvg` option from the `vxvmconvert` Main Menu:

```
Menu: Volume Manager/LVM_Conversion/ListLVMVolumeGroups
```

```
# listvg
```

Use this menu option to display a list of LVM volume groups. You can also choose to list detailed information about the LVM volume groups at a specific disk device address.

Select the Volume Group as follows:

```
Enter Volume Group (i.e.- vg08) or "all" [<address>,all,q,?]
```

```
default: all)
```

```
LVM VOLUME GROUP INFORMATION
NAME   VERSION   TYPE      PHYSICAL VOLUME
vg00   1.0       ROOT      disk10
vg09   2.0       Non-Root  disk11
vg08   1.0       Non-Root  disk12
```

Volume Group to list in detail

```
[<address>,none,q,?] (default: none) none
```

To display detailed information about a volume group, select any of the volume groups from the above list.

Volume Group to list in detail \

```
[<address>,none,q,?] (default: none) vg08
```

--- Volume groups ---

```
VG Name                /dev/vg08
VG Write Access        read/write
VG Status               available
Max LV                 255
Cur LV                 2
Open LV                 2
Max PV                 16
Cur PV                 1
Act PV                 1
Max PE per PV          1016
VGDA                   2
PE Size (Mbytes)       4
Total PE               250
Alloc PE                250
Free PE                 0
Total PVG               0
```

--- Logical volumes ---

```
LV Name                /dev/vg08/lvol1
LV Status               available/syncd
LV Size (Mbytes)       500
Allocated PE            125
Used PV                 1
```

--- Physical Extent ---

```
LV Name                /dev/vg08/lvol2
LV Status               available/syncd
LV Size (Mbytes)       500
```

```
Current LE                125
Allocated PE              125
Used PV                   1

--- Physical volumes ---
PV Name                   /dev/disk/disk12
PV Status                  available
Total PE                  250
Free PE                   0
List another LVM Volume Group? [y,n,q,?] (default: n)
Select an operation to perform:
```

Note: The volume groups you want to convert must not be a root volume group or have bootable volumes in the group.

Example: analyzing LVM volume groups

To analyze one or more LVM volume groups:

```
# vxvmconvert
```

```
Volume Manager Support Operations
Menu: VolumeManager/LVM_Conversion
1      Analyze LVM Volume Groups for Conversion
2      Convert LVM Volume Groups to VxVM
3      Roll back from VxVM to LVM
list   List disk information
listvg List LVM Volume Group information
?      Display help about menu
??     Display help about the menuing system
q      Exit from menus

Select an operation to perform: 1
Analyze one or more LVM Volume Groups
Menu: Volume Manager/LVM_Conversion/Analyze_LVM_VGs
Use this operation to analyze one or more LVM volume groups for
possible conversion using the VxVM Volume Manager. This
operation checks for problems that would prevent the conversion
from completing successfully. It calculates the space required
to add the volume groups disks to a Volume Manager disk group,
and to replace any existing partitions and volumes with Volume
Manager volumes, plexes, and sub-disks.
```


More than one volume group or pattern may be entered at the prompt.

Here are some LVM volume group selection examples:

```
all:      analyze all LVM Volume Groups (all except Root VG)
listvg:  list all LVM Volume Groups
list:    list all disk devices
vg_name: a single LVM Volume Group, named vg_name
<pattern>: for example vg08 vg09 vg05
```

Select volume groups to analyze:

```
[<pattern list>,all,list,listvg,q,?] vg08
```

Name a new disk group [<group>,list,q,?] (default: **dg08**)

Each volume group will be analyzed one at a time. If there are any in this list that you do not want to analyze, you can either abort now or wait until a later time when you will be given an opportunity to skip the analysis of any group(s) in this list. The following disk has been found in the vg08 volume group and will be analyzed for VxVM conversion.

disk12

To allow analysis, a new VxVM disk group, dg08, will be fabricated and the disk device disk12 will be added to the disk group with the disk name dg0801.

The disk12 disk has been configured for conversion.

The first stage of the Analysis process has completed successfully.

Second Stage Conversion Analysis of vg08

Analysis of vg08 found sufficient Private Space for conversion
Conversion Analysis of disk12 indicates that the Volume Group is still in use, which may prevent the completion of the conversion without having to reboot the system. You may want to double check that none of the volumes in the volume group are in use before continuing with the conversion.

Volume Group Analysis Completed

Hit RETURN to continue.

Example of a failed analysis:

```
# vxvmconvert
```

Volume Manager Support Operations

Menu: VolumeManager/LVM_Conversion

```
1      Analyze LVM Volume Groups for Conversion
2      Convert LVM Volume Groups to VxVM
3      Roll back from VxVM to LVM
```

```
list      List disk information
listvg   List LVM Volume Group information
?        Display help about menu
??       Display help about the menuing system
q        Exit from menus
```

Select an operation to perform: **1**

Analyze one or more LVM Volume Groups

Menu: Volume Manager/LVM_Conversion/Analyze_LVM_VGs

Use this operation to analyze one or more LVM volume groups for possible conversion using the VxVM Volume Manager. This operation checks for problems that would prevent the conversion from completing successfully. It calculates the space required to add the volume groups disks to a Volume Manager disk group, and to replace any existing partitions and volumes with Volume Manager volumes, plexes, and sub-disks.

More than one volume group or pattern may be entered at the prompt.

Here are some LVM volume group selection examples:

```
all:      analyze all LVM Volume Groups (all except Root VG)
listvg:   list all LVM Volume Groups
list:     list all disk devices
vg_name:a single LVM Volume Group, named vg_name
<pattern>: for example vg08 vg09 vg05
```

Select Volume Groups to analyze :

```
[<pattern-list>,all,list,listvg,q,?] vg08
```

Name a new disk group [<group>,list,q,?] (default: **dg08**)

The following disk has been found in the vg08 volume group and will be analyzed for VxVM conversion.

disk12

To allow analysis, a new VxVM disk group, dg08, will be fabricated and the disk device disk12 will be added to the disk group with the disk name dg0801.

The disk12 disk has been configured for conversion.

The first stage of the Analysis process has completed successfully.

Second Stage Conversion Analysis of vg08

Analysis of vg08 found insufficient Private Space for conversion

```
SMALLEST VGRA space = 176
RESERVED space sectors = 78
PRIVATE SPACE/FREE sectors = 98
AVAILABLE sector space = 49
AVAILABLE sector bytes = 50176
RECORDS neededs to convert = 399
MAXIMUM records allowable = 392
```

The smallest disk in the Volume Group (vg08) does not have sufficient private space for the conversion to succeed. There is only enough private space for 392 VM Database records and the conversion of Volume Group (vg08) would require enough space to allow 399 VxVM Database records. This would roughly translate to needing an additional 896 bytes available in the private space. This can be accomplished by reducing the number of volumes in the (vg08) Volume Group, and allowing that for every volume removed, the number of Database records required would be reduced by three. This is only a rough approximation, however.

Hit RETURN to continue.

Example: converting LVM volume groups to VxVM disk groups

To convert LVM volume groups to VxVM disk groups:

```
# vxvmconvert
```

```
Volume Manager Support Operations
Menu: VolumeManager/LVM_Conversion
1      Analyze LVM Volume Groups for Conversion
2      Convert LVM Volume Groups to VxVM
3      Roll back from VxVM to LVM
list   List disk information
listvg List LVM Volume Group information
?      Display help about menu
??     Display help about the menuing system
q      Exit from menus
```

Select an operation to perform: **2**

Convert one or more LVM Volume Groups

```
Menu: VolumeManager/LVM_Conversion/Convert_LVM_VGs
```

Use this operation to convert one or more LVM Volume Groups to

one or more VxVM disk groups. This adds the disks to a disk group and replaces existing partitions with volumes. LVM-VxVM Volume Group conversion may require a reboot for the changes to take effect. For this release, only Non-root LVM Volume Groups are allowed to be converted.

More than one Volume Group or pattern may be entered at the prompt.

Here are some LVM Volume Group selection examples:

```
all:      analyze all LVM Volume Groups (all except Root VG)
listvg:  list all LVM Volume Groups
list:    list all disk devices
vg_name: a single LVM Volume Group, named vg_name
<pattern>:  for example vg08 vg09 vg05
```

Select Volume Groups to convert :

```
[<pattern-list>,all,list,listvg,q,?] listvg
```

LVM VOLUME GROUP INFORMATION

NAME	VERSION	TYPE	PHYSICAL VOLUME
vg00	1.0	ROOT	disk10
vg05	1.0	Non-Root	disk11
vg03	2.0	Non-Root	disk14 disk15
vg08	1.0	Non-Root	disk12

Select Volume Groups to convert :

```
[<pattern-list>,all,list,listvg,q,?] vg08
```

vg08

Convert this Volume Group? [y,n,q,?] (default: **y**)

Name a new disk group [<group>,list,q,?] (default: **dg08**)

The following disk has been found in the vg08 volume group and will be configured for conversion to a VxVM disk group.

disk12

A new disk group dg08 will be created and the disk device disk12 will be converted and added to the disk group with the disk name dg0801.

The disk12 disk has been configured for conversion.

The first stage of the conversion operation has completed successfully. If you commit to the changes hereafter, the system will attempt to umount all of the associated file systems, stop and export each Volume Group, and then attempt to complete the conversion without having to reboot the system. If we are unable to stop and export any of the Volume Groups, then the conversion

process will not be able to complete without a reboot. You would then be given the choice to either abort the conversion, or finish the conversion by rebooting the system.

The conversion process will update the `/etc/fstab` file so that volume devices are used to mount the file systems on this disk device. You will need to update any other references such as backup scripts, databases, or manually created swap devices. If you do not like the default names chosen for the corresponding logical volumes, you may change these to whatever you like using `vxedit`.

Second Stage Conversion Analysis of `vg08`
Analysis of `vg08` found sufficient Private Space for conversion

Conversion Analysis of `disk12` indicates that the Volume Group is still in use, which may prevent the completion of the conversion without having to reboot the system. You may want to double check that none of the volumes in the volume group are in use before continuing with the conversion.

Hit RETURN to continue.
VxVM INFO V-5-2-4906
The expected time for convert is: 0 hrs 0 mins 7 secs.

Are you ready to commit to these changes? [y,n,q,?] (default: **y**)
Saving LVM configuration records for Volume Group `vg08`

Volume Group configuration for `/dev/vg08` has been saved in
`/etc/vx/reconfig.d/vgrecords/vg08/vg08.backup`
Beginning the export process on Volume Group `"/dev/vg08"`.
Volume group `"/dev/vg08"` is still active.
`/dev/disk/disk12`
`/dev/vg08/rlv1` will convert to `/dev/vx/dsk/dg08/dg08lv1`
`/dev/vg08/rlv1` will convert to `/dev/vx/rdisk/dg08/dg08lv1`
`/dev/vg08/lv2` will convert to `/dev/vx/dsk/dg08/dg08lv2`
`/dev/vg08/rlv2` will convert to `/dev/vx/rdisk/vg08dg/dg08lv2`
LVM Volume Group `vg08` Records Saved
Unmounting `vg08` file systems
Volume group `"/dev/vg08"` has been successfully changed.

The Volume Manager is now reconfiguring (partition phase)...
Volume Manager: Initializing `disk12` as a converted LVM disk.

```
The system reconfiguration will now be done without rebooting.
The Volume Manager is now reconfiguring (initialization
phase)...
Volume Manager: Adding dg0801 (disk12) as a converted LVM disk.
Adding volumes for disk12...
Starting new volumes...
Updating /etc/fstab...
The system will now Convert the LVM Volume Groups over to VxVM
disk groups.
Convert other LVM Volume Groups? [y,n,q,?] (default: n)
```

Example of a failed conversion:

```
# vxvmconvert
```

```
Volume Manager Support Operations
Menu: VolumeManager/LVM_Conversion
```

```
1      Analyze LVM Volume Groups for Conversion
2      Convert LVM Volume Groups to VxVM
3      Roll back from VxVM to LVM
list   List disk information
listvg List LVM Volume Group information
?      Display help about menu
??     Display help about the menuing system
q      Exit from menus
```

```
Select an operation to perform: 2
```

```
Convert one or more LVM Volume Groups
```

```
Menu: VolumeManager/LVM_Conversion/Convert_LVM_VGs
```

```
Use this operation to convert one or more LVM Volume Groups to
one or more VxVM disk groups. This adds the disks to a disk group
and replaces existing partitions with volumes. LVM-VxVM Volume
Group conversion may require a reboot for the changes to take
effect. For this release, only Non-root LVM Volume Groups are
allowed to be converted.
```

```
More than one Volume Group or pattern may be entered at the
prompt.
```

```
Here are some LVM Volume Group selection examples:
```

```
all:    analyze all LVM Volume Groups (all except Root VG)
listvg: list all LVM Volume Groups
list:   list all disk devices
vg_name: a single LVM Volume Group, named vg_name
<pattern>: for example vg08 vg09 vg05
```

```
Select Volume Groups to convert :
[<pattern-list>,all,list,listvg,q,?] listvg
LVM VOLUME GROUP INFORMATION
NAME      VERSION    TYPE        PHYSICAL  VOLUME
vg00      1.0        ROOT        disk10
vg05      1.0        Non-Root    disk11
vg03      2.0        Non-Root    disk14  disk15
vg08      1.0        Non-Root    disk12
Select Volume Groups to convert :
[<pattern-list>,all,list,listvg,q,?] vg08
vg08
Convert this Volume Group? [y,n,q,?] (default: y)
Name a new disk group [<group>,list,q,?] (default: dg08)
```

The following disk has been found in the vg08 volume group and will be configured for conversion to a VxVM disk group.

```
disk12
A new disk group dg08 will be created and the disk device disk12 will be converted and added to the disk group with the disk name dg0801.
```

The disk12 disk has been configured for conversion. The first stage of the conversion operation has completed successfully. If you commit to the changes hereafter, the system will attempt to unmount all of the associated file systems, stop and export each Volume Group, and then attempt to complete the conversion without having to reboot the system. If we are unable to stop and export any of the Volume Groups, then the conversion process will not be able to complete without a reboot. You would then be given the choice to either abort the conversion, or finish the conversion by rebooting the system.

The conversion process will update the /etc/fstab file so that volume devices are used to mount the file systems on this disk device. You will need to update any other references such as backup scripts, databases, or manually created swap devices. If you do not like the default names chosen for the corresponding logical volumes, you may change these to whatever you like using vxedit.

Second Stage Conversion Analysis of vg08

```
Analysis of vg08 found insufficient Private Space for conversion
SMALLEST VGRA space          = 176
```

```
RESERVED space sectors = 78
PRIVATE SPACE/FREE sectors = 98
AVAILABLE sector space = 49
AVAILABLE sector bytes = 50176
RECORDS neededs to convert= 399
MAXIMUM records allowable = 392
```

The smallest disk in the Volume Group (vg08) does not have sufficient private space for the conversion to succeed. There is only enough private space for 392 VM Database records and the conversion of Volume Group (vg08) would require enough space to allow 399 VxVM Database records. This would roughly translate to needing an additional 896 bytes available in the private space. This can be accomplished by reducing the number of volumes in the (vg08) Volume Group, and allowing that for every volume removed, the number of Database records required would be reduced by three. This is only a rough approximation, however. Hit RETURN to continue.

Example: converting LVM version 2 volume groups to VxVM disk groups

To convert LVM volume groups to VxVM disk groups:

```
#vxvmconvert
```

```
Volume Manager Support Operations
Menu: VolumeManager/LVM_Conversion
1      Analyze LVM Volume Groups for Conversion
2      Convert LVM Volume Groups to VxVM
3      Roll back from VxVM to LVM
list   List disk information
listvg List LVM Volume Group information
?      Display help about menu
??     Display help about the menuing system
q      Exit from menus
```

```
Select an operation to perform: listvg
```

```
List LVM Volume Group information
```

```
Menu: VolumeManager/LVM_Conversion/ListLVMVolumeGroups
```

Use this menu operation to display a list of LVM volume Groups. You can also choose to list detailed information about the LVMVG at a specific disk device address.

Enter Volume Group (i.e.- vg04) or "all"
 [<address>,all,q,?] (default:**all**)

LVM VOLUME GROUP INFORMATION

Name	Version	Type	Physical Volumes
vg00	1.0	ROOT	c2t0d0
vg01	2.0	Non-Root	c9t3d4

Volume Group to list in detail
 [<address>,none,q,?] (default:**none**)

Volume Manager Support Operations
 Menu: VolumeManager/LVM_Conversion

- 1 Analyze LVM Volume Groups for Conversion
- 2 Convert LVM Volume Groups to VxVM
- 3 Roll back from VxVM to LVM
- list List disk information
- listvg List LVM Volume Group information

- ? Display help about menu
- ?? Display help about the menuing system
- q Exit from menus

Select an operation to perform:**2**

Convert one or more LVM Volume Groups
 Menu: VolumeManager/LVM_Conversion/Convert_LVM_VGs

Use this operation to convert LVM Volume Groups to VxVM disk groups. This adds the disks to a disk group and replaces existing LVM volumes with VxVM volumes. LVM-VxVM Volume Group conversion may require a reboot for the changes to take effect. For this release, only Non-root LVM Volume Groups can be converted.

More than one Volume Group or pattern may be entered at the prompt. Here are some LVM Volume Group selection examples:

all: convert all LVM Volume Groups (all except Root VG)

```
listvg:      list all LVM Volume Groups
list:       list all disk devices
vg_name:    a single LVM Volume Group, named vg_name
<pattern>:  for example:  vg04 vg08 vg09
```

```
Select Volume Groups to convert :
[<pattern-list>,all,list,listvg,q,?] vg01
Convert LVM Volume Group vg01
```

```
Convert this Volume Group? [y,n,q,?] (default:y)
```

```
VxVM NOTICE V-5-2-0
LVM Volume Group vg01 reports a version 2.0
LVM Volume Group other than version 1.0 is not supported for conversion.

Hit RETURN to continue.
```

Example: list, listvg, and vxprint outputs of an LVM volume group before and after conversion

The examples below show the `vxvmconvert listvg`, `list`, and `vxprint` output for an LVM volume group `vg08` converted to a VxVM disk group `dg08`.

Example of `vxvmconvert listvg` output before conversion of volume group `vg08`

```
LVM VOLUME GROUP INFORMATION
NAME      VERSION   TYPE      PHYSICAL VOLUME
vg00      1.0      ROOT      c0t5d0
vg08      1.0      Non-Root  c0t8d0
vg09      2.0      Non-Root  c0t9d0
```

Example of the `vxvmconvert list` output which shows the disk devices on a system

```
DEVICE    DISK      GROUP    STATUS
c0t5d0    -         -        online
c0t8d0    -         -        LVM
c0t9d0    -         -        LVM
c0t10d0   disk01   rootdg   online
c0t11d0   -         -        online
Disk group: rootdg
```

What does vxvmconvert list display?

The device indicates a physical disk, a disk with a name indicates if the disk is under VxVM control, a group shows the disk group name, and the status indicates if it is an LVM disk. If the status is online, that means VxVM acknowledges the disk and has it under its control.

Example vxprint output before conversion

The `list` and `listvg` output is from within the `vxvmconvert` command. `vxprint` is a command line command.

```
TY NAME          ASSOC   KSTATE  LENGTH  PLOFFS  STATE  TUTILO  PUTILO
dg rootdg       rootdg  -        -        -        -        -        -
dm disk01       c0t10d0 -      2079468 -        -        -        -
```

Example of the listvg output after conversion of vg08

```
LVM VOLUME GROUP INFORMATION
Name      Version  Type      Physical Volumes
vg00      1.0     ROOT      c0t5d0
vg09      2.0     Non-Root  c0t9d0
ourvg     2.0     Non-Root  c5t3d5
Volume Group to list in detail: None
```

Note: Note that `vg08` is no longer listed under LVM information.

Example of the vxvmconvert list output after conversion of volume group vg08 to dg08

```
DEVICE  DISK   GROUP  STATUS
c0t5d0  -     -      online invalid
c0t8d0  dg0801 dg08   online
c0t9d0  -     -      LVM
c0t10d0 disk01 rootdg online
c0t11d0 -     -      online
Disk group: rootdg
```

Example of the vxprint output after conversion

```
TY NAME          ASSOC   KSTATE  LENGTH  PLOFFS  STATE  TUTILO  PUTILO
dg dg08         dg08    -        -        -        -        -        -
dm dg0801       c0t8d0  -      2080768 -        -        -        -
```

```
v dg081v1 fsgen ENABLED 102400 - ACTIVE - -
pl dg081v1-01 dg081v1 ENABLED 102400 - ACTIVE - -
sd dg0801-01 dg081v1-01 ENABLED 102400 - - -
```

The `vxprint` output provides the following information:

- The disk group `dg08` contains the VxVM disk `dg0801` and the volume `dg081v1`. The VxVM disk `dg0801` is associated with disk device `c0t8d0` and is 2080768 blocks in length. The volume `dg081v1` is of type `fsgen`, is enabled in the VxVM kernel driver, is of length 102400, and is in the ACTIVE state. This means that the volume is started, and the plex is enabled. Operations to the volume such as recovery and data access will be governed by the usage type `fsgen`.
- The plex `dg081v1-01` is associated with volume `dg081v1`, and maps the entire address range of the volume. Associated with the plex is one subdisk, `dg0801-01` which maps the plex address range from 0 to the entire length of the plex, i.e. 102400 blocks. As implied by the root of its name, the subdisk `dg0801-01` uses an extent from the VxVM disk `dg0801`.

Example: VxVM to LVM rollback

```
Volume Manager Support Operations
Menu: VolumeManager/LVM_Conversion

1 Analyze LVM Volume Groups for Conversion
2 Convert LVM Volume Groups to VxVM
3 Roll back from VxVM to LVM
list List disk information
listvg List LVM Volume Group information
? Display help about menu
?? Display help about the menuing system
q Exit from menus

Select an operation to perform: 3
Rollback one or more LVM Volume Groups
Menu: VolumeManager/LVM_Conversion/Rollback_LVM_VG
```

Use this operation to rollback from a conversion of an LVM Volume Group. This operation will tear down the VxVM disk group and recreate the LVM volume group in its original form. User data is untouched by rollback.

Warning: If the VxVM configuration has changed since the conversion you should NOT use this operation. Additional information is available on restoring configuration.

See [“Restoring the LVM volume group configuration”](#) on page 209.

More than one Volume Group or pattern may be entered at the prompt. Here are some LVM Volume Group selection examples:

```
all:      Rollback all converted LVM Volume Groups
listvg:  list all LVM Volume Groups eligible for rollback
list:    list all disk devices
vg_name: a single LVM Volume Group, named vg_name
<pattern>: for example vg08 vg09 vg05
```

```
Select Volume Group(s) to rollback :
[<pattern-list>,all,list,listvg,q,?]
vg08
Roll back this Volume Group? [y,n,q,?] (default: y)
Rolling back LVM configuration records for Volume Group vg08
Selected Volume Groups have been restored.
Hit RETURN to continue.
Rollback other LVM Volume Groups? [y,n,q,?] (default: n)
```

General information regarding conversion speed

The speed of the process of converting an existing LVM volume group to a similar VxVM disk group is largely dependent upon the size of the volume group being converted, as well as on the complexity of the volumes within that volume group.

Note: The `vxvmconvert` operation displays an estimated time before the actual conversion is committed.

Factors affecting conversion speed include:

- Size of volume groups. The larger the volume groups, the larger the LVM metadata on each disk. A copy must be made of the LVM metadata for each physical disk. Some areas are greater than 2MB; therefore, a 50-disk volume group requires 50 2MB reads and writes (i.e., 100 large I/Os) to complete.
- Individual size of a logical volume in a volume group, and the complexity of the logical volume layout. For example, for a system with 50 9GB drives, a simple 50GB logical volume of the first 5 1/2 disks can be created. But a 50GB striped logical volume that takes the first 1GB of all 50 disks can also be created.

The first and simple logical volume takes less time to convert than the striped volume. However, for the striped volume, 50 disks need to be checked. Also, the complexity of reproducing the VxVM commands to set up the striped volumes requires more VxVM commands to be generated to represent more smaller sub-disks representing the same amount of space.

Another factor in converting stripes is that stripes create more work for the converter. In some cases, stripes require 1GB volume, although only the metadata is being changed. In other cases, where there are more physical disks in one volume than another, there is more metadata to deal with. The converter has to read every physical extent map to ensure there are no holes in the volume; if holes are found, the converter maps around them.

- Number of volumes. While it takes longer to convert one 64GB volume than one 2GB volume, it also takes longer to convert 64 1GB volumes than one 64GB volume, providing that the volumes are of similar type.
- Mirrored volumes. Mirrored volumes typically do not take more time to convert than simple volumes. Volumes that are mirrored and striped at the same time would take longer, but LVM currently does not allow this.

Currently, after conversion, mirrored volumes are not automatically synchronized because a large mirror could take hours to complete.

For example, in tests, a 150GB volume group consisting of 20 simple logical volumes takes approximately 35-40 minutes to convert. In contrast, the same volume group (150GB) consisting of mirrored volumes that need to be synchronized can take 30-40 hours to convert.

Note: If you convert mirrored volumes, you must synchronize them in a separate step.

Non-interactive conversion of volume groups

The `vxvmconvert` utility is an interactive command. You can also use the `vxautoanalysis` command and the `vxautoconvert` command to perform non-interactive analysis and conversion of LVM volume groups. The `vxautorollback` command also lets you reverse the conversion, and turn a VxVM disk group back into an LVM volume group.

Analyzing volume groups for conversion

The `vxautoanalysis` utility analyses one or more LVM volume groups, and reports on their suitability for conversion to VxVM disk groups. Only analysis of the

suitability of the volume groups for conversion is performed. The actual conversion can be performed by using the `vxautoconvert` utility or the `vxvmconvert` utility.

Note: The VxVM configuration daemon (`vxconfigd`) must be running in order for the analysis to succeed.

To analyze volume groups for conversion

- ◆ Run the `vxautoanalysis` command:

```
# /usr/sbin/vxautoanalysis [-f] [vgname ...]
```

The volume groups may be specified by their names or full pathnames. If no volume groups are specified, analysis of all volume groups on the system is attempted.

If the value of the system tunable, `nproc`, is too low, the analysis will report that the conversion analysis of the volume groups cannot be performed in parallel. In that case, you can use the `-f` option to specify that the volume groups are to be analyzed one at a time.

See the `vxautoanalysis(1M)` manual page.

Converting volume groups to disk groups

The `vxautoconvert` utility converts one or more LVM volume groups to VxVM disk groups. Any LVM extent-based striped volumes are converted to stripe-mirror (non-layered) VxVM volumes.

Note: The VxVM configuration daemon (`vxconfigd`) must be running in order for the conversion to succeed. Conversion of the root volume group is not permitted. An error results if this is attempted.

To convert volume groups to disk groups

- ◆ Run the `vxautoconvert` command:

```
# /usr/sbin/vxautoconvert [-f] [vgname ...]
```

The volume groups may be specified by their names or full pathnames. If no volume groups are specified, conversion of all volume groups on the system is attempted.

If the value of the system tunable, `nproc`, is too low, the analysis will report that the conversion of the volume groups cannot be performed in parallel. In that case, you can use the `-f` option to specify that the volume groups are to be converted one at a time.

See the `vxautoconvert(1M)` manual page.

Converting disk groups back to volume groups

The `vxautorollback` utility converts one or more VxVM disk groups back to the LVM volume groups from which they had previously been converted.

Note: The VxVM configuration daemon (`vxconfigd`) must be running in order for the analysis to succeed. Reverse conversion is performed on each disk group in turn. Parallel conversion is not supported. If the configuration or layout of the volumes in the converted disk group has been changed since conversion from LVM to VxVM, reverse conversion is not possible. The records of the converted volume groups are stored in the directory `/etc/vx/reconfig.d/vgrecords`. Reverse conversion is not possible if these records are removed or modified.

To convert disk groups back to volume groups

- ◆ Run the `vxautorollback` command:

```
# /usr/sbin/vxautorollback [dgname ...]
```

The disk groups may be specified by their names or full pathnames. If no disk groups are specified, conversion is attempted for all disk groups on the system that have volume group records stored in the directory `/etc/vx/reconfig.d/vgrecords`.

See the `vxautorollback(1M)` manual page.

Online migration of native LVM volumes to VxVM volumes

This chapter includes the following topics:

- [About online migration from Logical Volume Manager \(LVM\) volumes to Veritas Volume Manager \(VxVM\) volumes](#)
- [Online migration from LVM volumes in standalone environment to VxVM volumes](#)
- [Online migration from LVM volumes in VCS HA environment to VxVM volumes](#)

About online migration from Logical Volume Manager (LVM) volumes to Veritas Volume Manager (VxVM) volumes

The online migration feature provides a convenient way to migrate existing applications running on native LVM volumes to VxVM volumes, with a limited downtime. The native LVM volume is referred to as the source volume and the VxVM volume as the target volume. This feature migrates the source LVM volume data to the target VxVM volumes on new storage, with the flexibility of different storage configuration and layouts. Once the migration is set up, the application can be resumed, while data synchronization from LVM to the VxVM volumes continues in the background. The migration configuration is set up such that the application does not require immediate reconfiguration to the new VxVM device paths. You can choose the point of committing the migration, when data synchronization is complete for all required volumes. In case of errors, it provides

a way to abort the migration and safely revert to the original LVM configuration. Source volumes are kept up-to-date till you commit or abort the migration.

Online migration can utilize the SmartMove capability for data synchronization, if VxFS file system is present on the source LVM volumes.

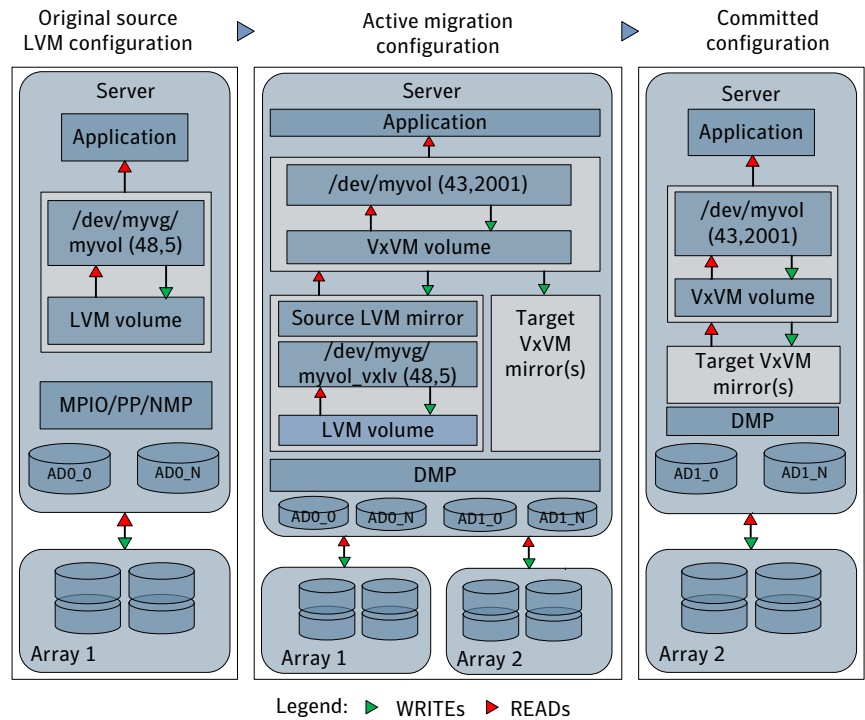
Online migration feature enables you to migrate from LVM to VxVM, in the following environments.

- From LVM volumes in a standalone environment. This supports a single host, in a non high-availability environment.
See [“Online migration from LVM volumes in standalone environment to VxVM volumes”](#) on page 235.
- From LVM volumes in a VCS HA environment. This supports the VCS cluster system with failover type high availability.
See [“Online migration from LVM volumes in VCS HA environment to VxVM volumes”](#) on page 245.

This feature is certified with Oracle as the primary database application, particularly with the versions 10gr2, 11gr1.

[Figure 15-1](#) illustrates the online migration process.

Figure 15-1 Migration process



The first stage shows the initial LVM source configuration.

The second stage shows the migration configuration. The renamed LVM volume is configured as a source mirror under the VxVM volume, from which data is synchronized to the VxVM target mirrors. Application accesses the VxVM volume through the original LVM device path. WRITES go to LVM as well as VxVM mirrors, and READs are served by the LVM mirror till the VxVM mirrors get synchronized.

The third stage shows the committed setup, where the LVM mirror is dissociated from the VxVM configuration. The LVM volume is not updated with the application IO any further.

Online migration from LVM volumes in standalone environment to VxVM volumes

The online migration feature provides a method to migrate native LVM volumes in a standalone environment to VxVM volumes.

See [“Administrative interface for online migration from LVM in standalone environment to VxVM”](#) on page 236.

Review the following prerequisites, do's and don'ts, and the scenarios not supported, before you start online migration in standalone environment.

See [“Preparing for online migration from LVM in standalone environment to VxVM”](#) on page 238.

See [“Do's and Don'ts for online migration from LVM in standalone environment to VxVM ”](#) on page 243.

See [“Scenarios not supported for migration from LVM in standalone environment to VxVM”](#) on page 244.

Refer to the following sections for the procedure to migrate, or to backout from a migration.

See [“Migrating from LVM in standalone environment to VxVM ”](#) on page 239.

See [“Backing out online migration of LVM in standalone environment to VxVM ”](#) on page 242.

Administrative interface for online migration from LVM in standalone environment to VxVM

Online migration of native LVM volumes to VxVM volumes can be managed using the `vxmigadm` administrative command. This command is available in `/etc/vx/bin`, `/opt/VRTS/bin`, and `/usr/lib/vxvm/bin`.

The syntax is as follows:

```
vxmigadm [options] operation [input-set]
```

The input-set can be a volume group, or a subset of volumes in a volume group, specified with the original source or target names. This input-set indicates the volumes or volume group the application is working on, and must be the same for all operations intended on that set, except for the recover operation.

Details of operation execution are logged to the `vxmigadm.log` file in the `/var/adm/vx` directory, which can also be accessed through `/etc/vx/log`.

You can use the verbose option `-v`, to get an indication of the operation progress.

[Table 15-1](#) describes the `vxmigadm` operations.

Note: The same operations are used for online migration in the VCS HA environment. For their additional description in the VCS HA environment:

See [Table 15-2](#) on page 248.

Table 15-1

Operation	Description
analyze	<p>Conducts sanity checks for the LVM and VxVM configuration intended for migration. Displays the detected irregularities, if any, that need to be fixed. Does not make any configuration changes in the setup. Performs checks for the environment, source volume group, source volumes, target diskgroup, and target volumes.</p> <p>You must fix the reported problems, and re-run the analyze operation until no issues are reported.</p> <p>Use this operation for analysis only prior to the start operation.</p>
start	<p>Starts the migration. Shut down the application before executing the start operation.</p> <p>Displays error and exits if anomalies are detected.</p> <p>Sets up the migration configuration, by adding the LVM volume as a source mirror, and the VM volume plex(es) as the target mirror(s).</p> <p>Starts data synchronization from LVM to VxVM volumes, in the background.</p> <p>The application, when resumed, works with the target VxVM volumes through the original LVM volume device paths. Source LVM volumes are renamed to <lvname>_vxl. Source volumes are kept up to date till you commit or abort the migration.</p> <p>Use this operation only after the analyze operation is successful.</p>
status	<p>Shows a detailed status of data synchronization.</p> <p>Lists volumes in different categories like "data synchronization not yet complete", "active data synchronization", and "data synchronization complete". For volumes in active synchronization, it displays details including the volume name, task id, task state, and the percentage of completion of that task. A volume can have multiple tasks, and it can switch between the categories of active data synchronization and data synchronization not yet complete, till all its tasks are done.</p> <p>Displays status only for the live migration setup, before you commit or abort migration. Also shows the current config type as STANDALONE or SFHA.</p>

Table 15-1 (continued)

Operation	Description
commit	<p>Commits the migration. Verifies that the migration is successfully completed for all input volumes. Dissociates the LVM configuration from the target VxVM configuration.</p> <p>The application continues to use the target VxVM volumes through the original LVM volume device paths. Source LVM volumes are retained with renamed names as <lvlname>_vxlv. The LVM volumes have the data up to the commit point, and are not updated any further.</p> <p>Commit the migration only after the status operation indicates that data synchronization is completed for all volumes.</p>
abort	<p>Aborts the migration and restores the LVM and VxVM configurations to the original state. Shut down the application before executing the abort operation.</p> <p>The application, when resumed, works with the source LVM volumes through the original LVM volume device paths. Source LVM volumes are renamed back to original names. The source LVM volumes contain data up to the abort point.</p> <p>Use this operation in case of errors, before you commit the migration.</p>
recover	<p>Recovers all existing migration configurations after a system crash or reboot. For any other disturbances, refer to the usual VxVM recovery procedures.</p> <p>Do this operation for post-reboot recovery, only if the migration is not already committed or aborted.</p>

See the `vxmigadm(1M)` manual page for more details.

Preparing for online migration from LVM in standalone environment to VxVM

For online migration of native LVM volumes to VxVM volumes, make sure the standalone system satisfies the following prerequisites.

- The system has a valid license for Mirroring.
- LVM volume group does not have bootable, paging space, or swap volumes.
- LVM volume group is of version 1.0.
- LVM volumes are in a usable state.
- LVM volume names are within the permissible length of 22 characters.

- Original LVM volume names do not have the `_vxl` string.
- To avoid conflicts for the LVM volume devices to be added in VxVM during migration, make sure that there are no VxVM objects with names of the form `<source lvolname>_vxl`.
- If the VxVM disk group is CDS type, temporarily turn off CDS to allow for the migration configuration to be set up. You can enable CDS again, after the commit or abort operation.
- The VxVM disk group is not a shared disk group.
- To avoid relocating data during migration, turn off hot relocation for the VxVM diskgroup, using the following method.
 In the `/etc/default/vxassist` file, set `spare=only`, and in the VxVM diskgroup, do not mark any disks as spare.
- VxVM volumes have the desired layout, except RAID5. RAID5 layout is not supported.
- Target VxVM disk group and volumes have the same names as the corresponding source LVM volume group and volumes.
- Target VxVM volumes have the same sizes, and credentials such as permissions, owner, and group, as that of the source LVM volumes.
- For optimized synchronizations and recoveries, add DCO, with version 20 or higher, to all the target VxVM volumes.

```
vxsnap -g mydg prepare volume_name drl=yes ndcomirs=num_mirrors
```

Otherwise, reserve sufficient free space for the start operation to create mirrored DCO automatically for all VxVM volumes.

- Comment fields for the VM volume, plexes, and DCO volumes are empty.
- The source and target volumes are not open or in use.

Note: Any source volume-specific features are not preserved on the target volume.

Additional prerequisites apply to the online migration in VCS HA environment.

See [“Preparing for online migration from LVM in VCS HA environment to VxVM”](#) on page 248.

Migrating from LVM in standalone environment to VxVM

Perform the following procedure.

To migrate native LVM volume to VxVM volume

- 1 Install Veritas Storage Foundation on the application host.
See the *Veritas Storage Foundation Installation Guide*.
- 2 Add new storage to the application host and configure it under Veritas Volume Manager (VxVM).
- 3 It is recommended to migrate all storage to DMP, as a common host-wide multi-pathing solution.
See the *Veritas Dynamic Multi-pathing Administrator's Guide*.
- 4 Create VxVM diskgroups and volumes on the newly added storage, according to your desired configuration. Online migration does not automatically create target VxVM configuration.
- 5 Make sure that the prerequisites for online migration are satisfied.
See [“Preparing for online migration from LVM in standalone environment to VxVM”](#) on page 238.

Note: Stop all operations or configuration changes that affect the VxVM or LVM volumes in the migration set, until migration is committed or aborted.

- 6 Analyze the configuration. Fix any discrepancies found, and re-analyze until no errors are reported.

```
# vxmigadm [-v] [-L logfile] [-t type] [-S] analyze -g grpname [-l vol_list]
```

If the application uses multiple volume groups, perform the analyze operation for each volume group.

- 7 If the application is online, then shut down the application. Make sure that the source and target volumes are not in use.
- 8 Start the migration.

```
# vxmigadm [-v] [-L logfile] [-t type] start -g grpname [-l vol_list]
```

If the application uses multiple volume groups, start the migration for each volume group separately, one by one.

Migrate all volumes of a particular application before you proceed to the next step of resuming the application.

9 When alerted, bring the application online.

The application can remain online while data synchronization continues in the background.

10 Use the status operation to track the progress of migration and to check if the data synchronization is complete.

```
# vxmigadm [-v] [-L logfile] status -g grpname [-l vol_list]
```

If the application uses multiple volume groups, track the migration status for each volume group. Ensure that data synchronization is complete for all required volumes, before you proceed to the next step of committing the migration.

You can administer the data synchronization tasks using the `vxtask` command. For example,

```
■ # vxtask pause <task-id>
```

```
■ # vxtask resume <task-id>
```

11 Commit the migration:

```
# vxmigadm [-v] [-L logfile] [-t type] commit -g grpname [-l vol_list]
```

In the presence of the FastResync license, data consistency is maintained by default for the source LVM set in a volume group.

In the absence of FastResync license, you must use forced commit:

If you want to maintain data consistency in this case, shut down the application before the forced commit.

```
# vxmigadm [-v] [-L logfile] [-t type] -f commit -g grpname [-l vol_list]
```

If the application was shutdown, resume it after the forced commit is completed.

If the application uses multiple volume groups, commit the migration for each volume group separately, one by one. If data consistency for the source LVM volumes is to be maintained across multiple volume groups, shut down the application before commit. Once all the volume groups are committed, resume the application.

Note: After the commit operation, reconfigure the application to directly use the target VxVM volume device path, as early as possible.

See [“Reconfiguring the application to use VxVM volume device path”](#) on page 242.

Reconfiguring the application to use VxVM volume device path

The LVM volume device path exists in the LVM volume group directory. Operations, such as volume group export or deletion, would disrupt the LVM device path, affecting the application's access to the VM volume in the migrated or committed configuration.

Also, if the source LVM volumes have VxFS file system present, special features provided by integration between VxVM and VxFS, cannot be used before the VxFS mount path is moved to the VxVM device path. These are operations like `vxresize`, `vxrelayout`, snapshot operations involving file system freeze and thaw, and so on.

After the commit operation, reconfigure the application, as early as possible, to use the standard device path of the target VxVM volume `/dev/vx/(r)disk/<dgname>/<volname>`. This would be as per the recommended reconfiguration procedure for the application. To discontinue the maintenance of the original LVM volume device path mapping to the VM volume, clear the `vxmig_end` marker from the VxVM volumes' comment field. Also remove the block and char device entries of the original LVM device path initially used by application.

Backing out online migration of LVM in standalone environment to VxVM

Perform the following procedure to abort the online migration in case of any errors, before the point of committing the migration.

The abort operation rolls back any LVM and VxVM configuration changes performed for migration setup, and restores the original source LVM configuration.

To back out online migration of LVM in standalone environment to VxVM

- 1 If the application is running, shut it down.
- 2 If the start operation is in progress, exit it.

3 Abort the migration:

```
# vxmigadm [-v] [-L logfile] [-t type] abort -g grpname [-l vol_list]
```

If the application uses multiple volume groups, abort the migration for each volume group separately, one by one.

Abort all volumes of a particular application before you proceed to the next step of resuming the application.

4 When alerted, bring the application online.

Do's and Don'ts for online migration from LVM in standalone environment to VxVM

For online migration of native LVM volumes to the VxVM volumes, follow these do's and don'ts in a standalone system.

- To avoid data corruption, do not use the original LVM device path and the VxVM device path simultaneously, in the migrated or committed config. To maintain application access to VxVM volume through the original LVM volume device path, in the migrated or committed config, do not cause any disturbances to the LVM volume group directory contents and availability, or to the comment fields of VxVM volumes. To avoid such disruptions, reconfigure the application as early as possible to directly use the target VxVM volume device path, after the commit operation. See [“Reconfiguring the application to use VxVM volume device path”](#) on page 242.
- Perform the migration operations in the right supported sequence. Also, do not interrupt the vxmigadm operations as far as possible. Else the configuration might need manual intervention for cleanup, for example in the case of commit.
- In the status operation output, if a volume shows in the category of "Data synchronization not yet complete" for a considerable time, then user intervention might be required. Check for any target VxVM mirrors in erroneous state, and manually rectify the issue for data synchronization to proceed to completion. Do not perform any operations on, or disturb, the source LVM mirror under any circumstances.
- Do not perform configuration changes on, or disturb the availability of, the source LVM and the target VxVM objects in the migration set, till the migration is aborted or committed. The migration setup, including the original LVM volume device path through which the VxVM volume is made available, must not be disrupted. Vxmigadm operations expect the setup must be in appropriate state for their further execution.

- For migrated or committed configuration, maintain the same volume name, permissions, owner, and group for the VxVM volume and the original LVM device path used by application. If these parameters need to be changed, do the changes onto the VxVM target volume. These changes done for the VxVM volume will also get applied to the original LVM device path used by application.
- Do not execute configuration change operations like start, abort, or commit simultaneously on multiple volume groups.
- Do not perform reconfiguration operations like `vxdisk scandisks` or `vold` restarts while configuration change operations like start, abort, or commit are being executed during migration.
- Migration configuration is intended only for the special process of migration. Do not use it for exhaustive or general purpose.
- For SmartMove capable volumes, the `vxmigadm recover` operation used for post-reboot recovery, gives the following message. It prompts to mount the VxFS file system for SmartMove enabled data synchronization for the listed volumes, else it proceeds with full synchronization.

```
VxVM vxmigadm NOTICE Mount the file system to enable synchronization
with SmartMove for these volumes. If file system is not
mounted, then full synchronization will be done : <volume names>
VxVM vxmigadm NOTICE Press y to continue.
```

Note that if the file system is not mounted here, then full synchronization will not be automatically started for volumes on thin LUNs, similar to the usual VxVM recovery for volumes through `vxrecover`. For such volumes, after the `vxmigadm recover` operation completes, follow the method of executing the `vxrecover` command with `-o force` option, if you want to trigger full synchronization.

Additional recommendations apply to the online migration in VCS HA environment.

See [“Do's and Don'ts for online migration from LVM in VCS HA environment to VxVM”](#) on page 253.

Scenarios not supported for migration from LVM in standalone environment to VxVM

Online migration of LVM in standalone environment to VxVM is not supported in the following scenarios.

- Moving or importing the migration configuration in migrated or committed state, onto another host is not supported, until you perform the entire

procedure to reconfigure the application to use VxVM volume device path after commit.

See [“Reconfiguring the application to use VxVM volume device path”](#) on page 242.

- LVM volume group versions other than 1.0 are not supported.
- LVM or VxVM configurations with advanced features like snapshots, replication, and such other features are not supported.

Some additional scenarios are not supported for online migration in VCS HA environment.

See [“Scenarios not supported for migration from LVM VCS HA environment to VxVM”](#) on page 254.

Online migration from LVM volumes in VCS HA environment to VxVM volumes

The online migration feature is also integrated with VCS to provide online migration in High Availability (HA) environment. This enables VCS to monitor and maintain high availability of the updated configuration, during the migration process.

See [“About online migration from LVM in VCS HA environment to VxVM”](#) on page 246.

See [“Administrative interface for online migration from LVM in VCS HA environment to VxVM”](#) on page 247.

Review the following prerequisites, do's and don'ts, and the scenarios not supported, before you start online migration in SFHA environment.

See [“Preparing for online migration from LVM in VCS HA environment to VxVM”](#) on page 248.

See [“Do's and Don'ts for online migration from LVM in VCS HA environment to VxVM”](#) on page 253.

See [“Scenarios not supported for migration from LVM VCS HA environment to VxVM”](#) on page 254.

Refer to the following sections for the procedure to migrate, or to backout from a migration.

See [“Migrating from LVM in VCS HA environment to VxVM ”](#) on page 249.

See [“Backing out online migration of LVM in VCS HA environment to VxVM ”](#) on page 252.

About online migration from LVM in VCS HA environment to VxVM

The online migration feature provides a method to migrate native LVM volumes in VCS HA environment to VxVM volumes. The `vxmigadm` command can be used for migration in both standalone or VCS HA environment. For VCS HA environment, execute the `vxmigadm` commands on the node having the LVM volume group's service group online.

All the `vxmigadm` operations detect the current configuration type as VCS HA or standalone. For VCS HA environment, `vxmigadm` operations also do the required VCS configuration changes, along with the LVM and VxVM changes.

For a migrated, committed, or aborted setup, all the standard VCS operations like service group offline, online, switch, and so on, can be performed as usual. In case of service group failover or switch, all the requisite configuration is established or recovered on the target node, for the application to work.

In VCS HA environment, the `vxmigadm` command operations automatically perform the required VCS configuration changes like the following.

- Freezes the Service Group temporarily during the operation duration, and then unfreezes it.
- Takes the relevant resources offline before performing the operation, and brings the resources online towards the end.
- Adds or removes required resources like VxVM DiskGroup, VxVM Volume, and other temporary resources, under VCS control.
- Changes the VCS resource dependencies, so that the application depends on LVM and/or VxVM configuration as appropriate.
- Modifies the required resource attributes.

VCS resource hierarchy changes:

During migration, the start operation adds VxVM resources for the respective VxVM objects, corresponding to the LVM resources.

VxVM DiskGroup resource is added for the LVMVolumeGroup or LVMCombo resources.

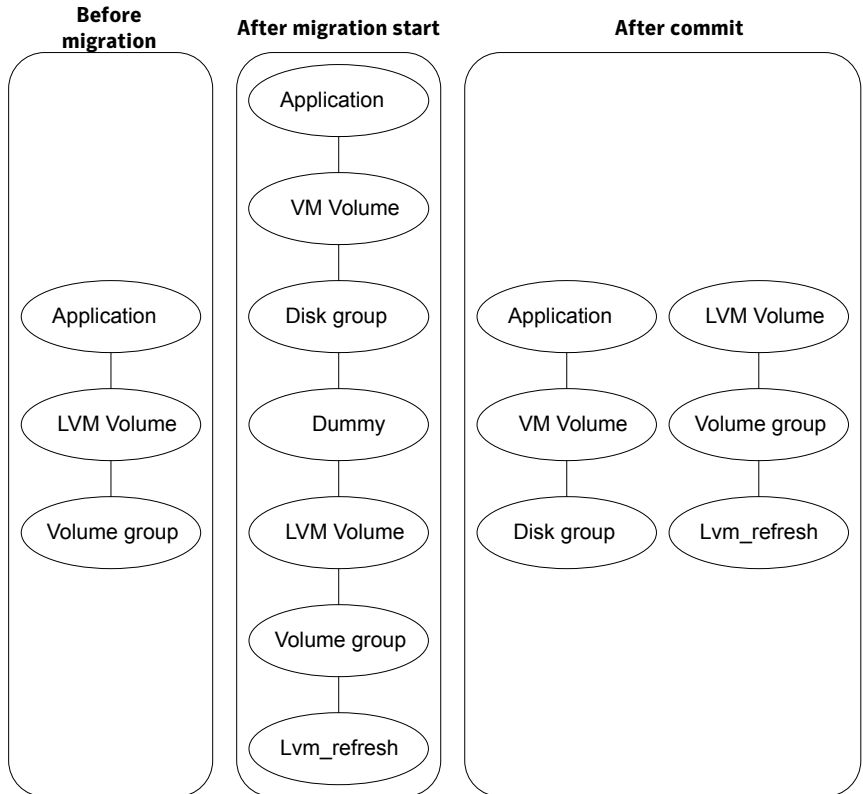
VxVM Volume resources are added for their corresponding LVMLogicalVolume or LVMCombo resources.

During the start operation, a temporary resource is also added to the VCS hierarchy, below the VxVM DiskGroup resource. This resource is specific to the live migration setup, for managing the LVM volume device configuration in VxVM, during service group online and offline operations. It is removed once the setup is aborted or committed.

Another resource is added below the LVMVolumeGroup resource. It reflects the LVM volume naming changes, during failover.

Figure 15-2 illustrates the changes to VCS hierarchy, during and after online migration.

Figure 15-2 VCS resource hierarchy for online migration in VCS HA environment



Administrative interface for online migration from LVM in VCS HA environment to VxVM

Online migration from LVM in VCS HA environment utilizes the `vxmigadm` administrative command. Details of VCS operations are logged to the `vc_s_vxvm_aom.log` file in the `/var/VRTSvcs/log` directory.

See “[Administrative interface for online migration from LVM in standalone environment to VxVM](#)” on page 236.

The same operations, as used in standalone environment, are used during online migration from LVM in VCS HA environment.

See [Table 15-1](#) on page 237.

[Table 15-2](#) gives additional description of the `vxmigadm` operations during online migration in VCS HA environment.

Table 15-2

Operation	Description
<code>analyze</code>	Performs VCS sanity checks. The <code>-s</code> option is used for analysis on all standby nodes which do not have the service group online.
<code>start</code>	New VxVM resources are added in the volume group's service group under VCS control. The resource dependencies are updated so that application will depend on both LVM and VxVM resources. Also some temporary resources are added to manage the LVM and VxVM migration configuration during service group failover.
<code>commit</code>	Resource dependencies are updated so that the application depends only on VxVM resources. The LVM resources are still present in the service group, but are disconnected from the application resource tree, and marked non-critical. Also removes the unrequired temporary resources added during the start operation.
<code>abort</code>	All the VxVM resources that were added in the start operation are deleted. The original VCS resource dependencies are restored back so that application depends only on LVM resources. Also removes the unrequired temporary resources added during the start operation.

Note: The `recover` operation is required only in the standalone environment. In VCS HA environment, the required recovery is internally done by VCS.

See the `vxmigadm(1M)` manual page for more details.

Preparing for online migration from LVM in VCS HA environment to VxVM

The prerequisites for online migration in a standalone system apply to online migration in VCS HA environment.

See [“Preparing for online migration from LVM in standalone environment to VxVM”](#) on page 238.

In addition, the following prerequisites must be satisfied for migration from LVM volumes in VCS HA environment.

- VCS is installed and running on all the cluster nodes.
- LVM volume group, logical volumes, and all the related resources to be migrated, are configured under VCS control in a single service group of failover type.
- Target VxVM objects, such as diskgroup and volumes, are not already configured under VCS control.
- The service group having volume group to be migrated, is not in frozen state.
- VCS configuration is not in read-write mode.

Migrating from LVM in VCS HA environment to VxVM

Perform the following procedure.

To migrate native LVM volume to VxVM volume

- 1 Install Veritas Storage Foundation High Availability on the nodes of the cluster.
See the [Veritas Storage Foundation and High Availability Installation and Configuration Guide](#).
- 2 Add new storage to the application host and configure it under Veritas Volume Manager (VxVM).
- 3 It is recommended to migrate all storage to DMP, as a common host-wide multi-pathing solution.
See the [Veritas Dynamic Multi-pathing Administrator's Guide](#).
- 4 Create VxVM diskgroups and volumes on the newly added storage, according to your desired configuration. Online migration does not automatically create target VxVM configuration.
- 5 Make sure that the prerequisites for online migration are satisfied.
[See “Preparing for online migration from LVM in VCS HA environment to VxVM” on page 248.](#)

Note: Stop all operations or configuration changes that affect the VxVM or LVM volumes in the migration set, and the VCS configuration, until migration is committed or aborted.

- 6 Analyze the configuration. Fix any discrepancies found, and re-analyze until no errors are reported. Execute analyze with `-s` option on all standby nodes which do not have the service group online.

```
# vxmigadm [-v] [-L logfile] [-t type] [-S] analyze -g grpname [-l vol_list]
```

- 7 If the application is online, then shut down the application resource. Make sure that the source and target volumes are not in use.

- 8 Start the migration.

```
# vxmigadm [-v] [-L logfile] [-t type] start -g grpname [-l vol_list]
```

- 9 When alerted, bring the application resources online.

The application can remain online while data synchronization continues in the background.

- 10 Use the status operation to track the progress of migration and to check if the data synchronization is complete.

```
# vxmigadm [-v] [-L logfile] status -g grpname [-l vol_list]
```

You can administer the data synchronization tasks using the `vxtask` command. For example,

```
■ # vxtask pause <task-id>
```

- # `vxtask resume <task-id>`

11 Commit the migration:

```
# vxmigadm [-v] [-L logfile] [-t type] commit -g grpname [-l vol_list]
```

In the presence of the FastResync license, data consistency is maintained by default for the source LVM set in a volume group.

In the absence of FastResync license, you must use forced commit:

If you want to maintain data consistency in this case, shut down the application before the forced commit.

```
# vxmigadm [-v] [-L logfile] [-t type] -f commit -g grpname [-l vol_list]
```

If the application was shutdown, resume it after the forced commit is completed.

For migrating configurations with multiple volume groups:

See [“Migrating configurations with multiple volume groups”](#) on page 251.

Note: After the commit operation, reconfigure the application to directly use the target VxVM volume device path, as early as possible.

See [“Reconfiguring the application to use VxVM volume device path”](#) on page 242.

Additionally, also modify the BlockDevice attributes of Mount resource to reflect the VxVM volume device path.

Migrating configurations with multiple volume groups

In a VCS HA system, only one volume group can be in a migration state at a time. Hence for migrating multiple volume groups present on the system, a volume group must be migrated and committed, before proceeding with the migration of the next volume group.

Perform the following procedure, to migrate configurations having multiple volume groups:

- One or more applications using multiple volume groups configured in a single service group.

- One or more applications using one or more volume groups in service group 1, another set of one or more applications using one or more volume groups in service group 2, and so on.

To migrate a configuration with multiple volume groups

- 1 Select a volume group to be migrated.
- 2 If its application is online, then make offline the application resources dependent on the LVM resources.
- 3 Start migration of the volume group using the `vxmigadm start` operation.
- 4 When alerted, bring the application resources online, that were made offline (in step 2).
- 5 Use the `vxmigadm status` operation to track progress till it indicates that the data synchronization is completed for the volume group.
- 6 If application data consistency is required for the source LVM volumes after commit, and the application is online, then make offline the application resources dependent on that volume group.
- 7 Commit migration of the volume group using the `vxmigadm commit` operation.
- 8 If you had made the application resources offline (in step 6), then bring the application resources online.
- 9 Repeat the above steps for the next volume group. Continue in this way for all other volume groups one by one.

Backing out online migration of LVM in VCS HA environment to VxVM

Perform the following procedure to abort the online migration in case of any errors, before the point of committing the migration.

The abort operation cancels any LVM, VxVM, and VCS configuration changes performed for migration setup, and restores the original source LVM configuration.

To back out online migration of LVM in VCS HA environment to VxVM

- 1 If the application is running, offline the application resources.
- 2 If the start operation is in progress, exit it.
- 3 Abort the migration:

```
# vxmigadm [-v] [-L logfile] [-t type] abort -g grpname [-l vol_list]
```

- 4 When alerted, bring the application resource online.

Do's and Don'ts for online migration from LVM in VCS HA environment to VxVM

The do's and don'ts for online migration in a standalone system apply to online migration in VCS HA environment.

See [“Do's and Don'ts for online migration from LVM in standalone environment to VxVM”](#) on page 243.

In addition, follow these do's and don'ts for online migration from LVM volumes in VCS HA environment.

- All volumes of a volume group that have LVM volume resources configured in VCS should be migrated together. Remaining volumes of that volume group cannot be migrated later separately.
- The volumes that are to be migrated but are not configured as LVM volume resources in VCS should be migrated along with all those having volume resources.
- If the start operation fails, the service group is left in a frozen state. Unfreeze the service group manually and run the abort operation on the same node.
- Do not perform any VCS configuration change while migration operations are in progress.
- Do not perform any changes in the VCS service group involved in migration till the migration is either committed or aborted.
- For an active migration setup, the service group failover may require disk re-scanning on the failover node. This may require tuning the value of OnlineTimeout attribute of temporary resource added below DiskGroup resource, as per time required for the `vxdisk scandisks` operation. The default value of OnlineTimeout is 1hr.
- On committing the migration, the LVM resources like LVMVolumeGroup, LVMLogicalVolume, and LVMCombo are present in the service group as non-critical resources. If these resources are not required, remove them from the service group.
- During abort or commit operation on a cluster node, the LVM volume devices are cleaned up from VxVM configuration on this node. For all other cluster nodes previously traversed by migration setup, stray entries of LVM volume devices in VxVM configuration remain. You should clean up these stray entries, to prevent interference in any further online migration functioning on that node. Check and clean up the stray entries for LVM volume <lvol>, as follows:

```
# vxdisk list <lvol>_vxlv
```

If the device exists, remove it as:

```
# vxdisk rm <lvol>_vxlv  
# vxddladm listforeign | grep <lvol>_vxlv
```

If the device entry exists, remove it as:

```
# vxddladm rmforeign charpath= \  
/dev/<vg>/r<lvol>_vxlv blockpath=/dev/<vg>/<lvol>_vxlv
```

Scenarios not supported for migration from LVM VCS HA environment to VxVM

Scenarios not supported for online migration in standalone environment apply to online migration in VCS HA environment.

See [“Scenarios not supported for migration from LVM in standalone environment to VxVM”](#) on page 244.

In addition, the following scenarios are not supported.

- Moving or importing the configuration in migrated or committed state, onto another independent system is not supported, until you perform the entire procedure to reconfigure the application to use VxVM volume device path after commit.

See [“Reconfiguring the application to use VxVM volume device path”](#) on page 242.

Additionally, also modify the BlockDevice attributes of Mount resource to reflect the VxVM volume device path.

- Configurations with advanced VCS features like GCO, RDC, and such other features are not supported.
- Clustering software other than Veritas Cluster Server is not supported.
- Configurations with parallel type service group are not supported.

Migrating storage arrays

This chapter includes the following topics:

- [Array migration for storage using Linux](#)
- [Overview of storage mirroring for migration](#)
- [Allocating new storage](#)
- [Initializing the new disk](#)
- [Checking the current VxVM information](#)
- [Adding a new disk to the disk group](#)
- [Mirroring](#)
- [Monitoring](#)
- [Mirror completion](#)
- [Removing old storage](#)
- [Post-mirroring steps](#)

Array migration for storage using Linux

The array migration example documented for this use case uses a Linux system. The example details would be different for AIX, Solaris, or Windows systems.

Veritas Storage Foundation and High Availability Solutions (SFHA Solutions) products provide enterprise-class software tools which enable companies to achieve data management goals which would otherwise require more expensive hardware or time-consuming consultative solutions.

For many organizations, both large and small, storage arrays tend to serve as useful storage repositories for periods of 3-6 years. Companies are constantly

evaluating new storage solutions in their efforts to drive down costs, improve performance and increase capacity. The flexibility of Storage Foundation and High Availability Solutions enable efficient migration to new storage and improve the overall availability of data.

While there are several methods for accomplishing the migration, the most basic and traditional method is using a volume level mirror. The example procedures:

- Provide system administrators responsible for SFHA Solutions systems within their organization a demonstration of the steps required for performing an online storage array migration from one array to another.
- Illustrate the migration process using a Linux system which is connected to two different storage arrays through a SAN.
- Provide steps for starting with a file system with a single volume, mirroring the volume to a volume to another array, and then detaching the original storage.
- Are performed from the command prompt.
- Use Operating System Based Naming (OSN) for disk devices (sdb, sdc, etc).

There are two user interface options:

- The SFHA Solutions command line interface (CLI).
- The Veritas Operations Manager (VOM) graphical user interface (GUI) has a storage migration wizard.

See the Veritas Operations Manager documentation for details:

https://sots.symantec.com/public/documents/vom/40ru1/windowsandunix/productguides/pdf/vom_sp_addon_users.pdf

Note: Symantec's NetBackup PureDisk comes bundled with the Veritas Storage Foundation and High Availability Solutions software for the purpose of enhanced storage management and high availability. Storage arrays used with PureDisk can also be migrated using the SFHA Solutions methodologies.

Overview of storage mirroring for migration

The migration example occurs between a Hitachi 9900 array, 350 GB disk/LUN and a NetApps 3050 Fibre-Attached 199GB disk/LUN.

To migrate storage using storage mirroring

- 1 Connect the new array to the SAN.
- 2 Zone the array controller ports(s) to the server HBA port(s).
- 3 Create or present the new LUN(s) on the array.

- 4 Map the new LUN(s) to the server HBA port(s).
- 5 Stop any processes that are running on this volume or file system.
- 6 Rescan hardware using `rescan-scsi-bus.sh` and `scsidev` commands, or reboot (optional).
- 7 Confirm that the operating system has access to the new target storage (*Array-B*).
- 8 Bring new disks under Veritas Volume Manager (VxVM) control.
- 9 Start the VxVM mirroring process to synchronize the plexes between the source and target array enclosures.
- 10 Monitor the mirroring process.
- 11 After mirroring is complete, logically remove disks from VxVM control.

Note: The example Linux system happens to be running as a Symantec NetBackup Puredisk server which includes the Storage Foundation software. Puredisk also supports this mode of storage array migration.

- 12 Disconnect the old storage array (enclosure).

Allocating new storage

The first step to migrating array storage is to allocate new storage to the server.

To allocate new storage as in the example

- 1 Create the LUN(s) on the new array.
- 2 Map the new LUN(s) to the server.
- 3 Zone the new LUN(s) to the server.
- 4 Reboot or rescan using a native OS tool such as “fdisk” and the new external disk is now visible.

In the example, the original disk (`/dev/sdb`) has already been initialized by Veritas Volume Manager (VxVM). Note that it has a partition layout already established. Note also the different disk sizes. It may turn out that you want to use smaller or larger LUNs. This is fine, but if you are going to mirror to a smaller LUN you will need to shrink the original volume so that it can fit onto the physical disk device or LUNs.

```

root@ny-puredisk: fdisk -l

Disk /dev/sda: 80.0 GB, 80026361856 bytes
255 heads, 63 sectors/track, 9729 cylinders
Units = cylinders of 16065 * 512 = 8225280 bytes

   Device Boot      Start         End      Blocks   Id  System
/dev/sda1  *           1           13        104391   83  Linux
/dev/sda2                14          1580       12586927+  82  Linux swap / Solaris
/dev/sda3           1581          9729       65456842+  83  Linux

Disk /dev/sdb: 375.0 GB, 375813308416 bytes
255 heads, 63 sectors/track, 45690 cylinders
Units = cylinders of 16065 * 512 = 8225280 bytes

   Device Boot      Start         End      Blocks   Id  System
/dev/sdb4                1          45683       366948666    5  Extended
/dev/sdb5                1           1           1165+    7f  Unknown
/dev/sdb6                1          45683       366947343    7e  Unknown

Disk /dev/sdc: 213.6 GB, 213676323840 bytes
255 heads, 63 sectors/track, 25978 cylinders
Units = cylinders of 16065 * 512 = 8225280 bytes

   Device Boot      Start         End      Blocks   Id  System
/dev/sdc1                1          25976       208652188+    5  Extended
root@ny-puredisk:

```

To shrink the original volume

- ◆ You can shrink the new volume size to *n* gigabytes:

```
# vxassist -g diskgroup_name shrinkto volume_name ng
```

Then resize the file system:

```
# /opt/VRTS/bin/fsadm -F vxfs -b new_size_in_sectors /Storage
```

Alternately, use the `vxresize` command to resize both the volume and the file system.

To grow the original volume

- ◆ You can increase the new volume size to *n* gigabytes:

```
# vxassist -g diskgroup_name growto volume_name ng
```

Then resize the file system:

```
# /opt/VRTS/bin/fsadm -F vxfs -b new_size_in_sectors /Storage
```

Alternately, use the `vxresize` command to resize both the volume and the file system.

Note: SmartMove enables you to migrate from thick array LUNs to thin array LUNs on those enclosures that support Thin Provisioning.

Initializing the new disk

Now that the operating system recognizes the new disk, the next step is to initialize it.

To initialize the disk as in the example

- ◆ Use the `vxdisksetup` command to establish the appropriate VxVM-friendly partition layout for Veritas Volume Manager.

Note below that the internal name `OLDDISK` is assigned to the old disk. The new disk is assigned a unique name later for the sake of clarity.

```
root@ny-puredisk:~# vxdisk list
DEVICE      TYPE      DISK      GROUP      STATUS
sda         auto:none -          -          online invalid
sdb         auto:sliced OLDDISK   PDDG       online
sdc         auto:none -          -          online invalid
root@ny-puredisk:~#
root@ny-puredisk:~#
root@ny-puredisk:~# /usr/lib/vxvm/bin/vxdisksetup sdc
root@ny-puredisk:~# vxdisk list
DEVICE      TYPE      DISK      GROUP      STATUS
sda         auto:none -          -          online invalid
sdb         auto:sliced OLDDISK   PDDG       online
sdc         auto:none -          -          online invalid
root@ny-puredisk:~#
```

The disk is now initialized under VxVM control. Note below, that the disk has a new partition table similar to the existing disk (`sdb`) and is ready to be joined to the Veritas disk group `PDDG` (name of the example disk group).

```
root@ny-puredisk:~# fdisk -l /dev/sdc
Disk /dev/sdc: 213.6 GB, 213676323840 bytes
255 heads, 63 sectors/track, 25978 cylinders
Units = cylinders of 16065 * 512 = 8225280 bytes

   Device Boot      Start         End      Blocks   Id  System
/dev/sdc4          1         25976   208652180+    5  Extended
/dev/sdc5          1           1        1165+    7f  Unknown
/dev/sdc6          1         25976   208650865+    7e  Unknown
root@ny-puredisk:~#
```

Checking the current VxVM information

Check the VxVM information after initializing the new disk. The screen shot below illustrates all the disks on the server along with their corresponding partition tables. Note that disks *sdb* and *sdc* are partitioned in the same manner since they were both set up with the `vxdisksetup` command.

```

root@ny-puredisk: fdisk -l

Disk /dev/sda: 80.0 GB, 80026361856 bytes
255 heads, 63 sectors/track, 9729 cylinders
Units = cylinders of 16065 * 512 = 8225280 bytes

   Device Boot      Start         End      Blocks   Id  System
/dev/sda1 *          1           13        104391   83  Linux
/dev/sda2            14          1580     12586927+  82  Linux swap / Solaris
/dev/sda3           1581          9729     65456842+  83  Linux

Disk /dev/sdb: 375.8 GB, 375813308416 bytes
255 heads, 63 sectors/track, 45690 cylinders
Units = cylinders of 16065 * 512 = 8225280 bytes

   Device Boot      Start         End      Blocks   Id  System
/dev/sdb4            1          45683     366948666    5  Extended
/dev/sdb5            1           1         1165+     7f  Unknown
/dev/sdb6            1          45683     366947343    7e  Unknown

Disk /dev/sdc: 213.6 GB, 213676323840 bytes
255 heads, 63 sectors/track, 25978 cylinders
Units = cylinders of 16065 * 512 = 8225280 bytes

   Device Boot      Start         End      Blocks   Id  System
/dev/sdc4            1          25976     208652188+    5  Extended
/dev/sdc5            1           1         1165+     0  Empty
/dev/sdc6            1          25976     208650865+    0  Empty
root@ny-puredisk: █
    
```

The screen shot below shows the VxVM hierarchy for existing storage objects. Remember that we are working with a live and running server. We are using a logical disk group called *PDDG* which has other storage objects subordinate to it. The most important storage object here is the volume which is called *Storage*. The volume name can be any arbitrary name that you want, but for this example, the volume name is “*Storage*”. The volume object is denoted by “v” in the output of the `vxprint` command. Other objects are subdisks (*sd*) which represents a single contiguous range of blocks on a single LUN. The other object here is a plex (“pl”) which represents the virtual object or container to which the OS reads and writes. In `vxprint`, the length values are expressed in sectors, which in Linux are 512 bytes each. The raw volume size is 377487360 sectors in length, or when multiplied by 512 bytes (512*377487360) is 193273528320 bytes, or about 193 GB(2).

Notice that when the new disk was added it was 213GB yet the original existing *Storage* volume was 250GB. The *Storage* volume had to first be shrunk to a size equal the same (or smaller) number of sectors as the disk to which it would be mirrored.

```

root@ny-puredisk:vxprint
Disk group: PDDG

TY NAME          ASSOC      KSTATE  LENGTH  PLOFFS  STATE  TUTILO
PUTIL0
dg PDDG          PDDG      -        -        -        -        -
dm OLDDISK       sdb       -        733894672 -        -        -
dm sdc           sdc       -        417301712 -        -        -

v Storage       fsngen    ENABLED  377487360 -        ACTIVE  -
pl Storage-02   Storage  ENABLED  377487360 -        ACTIVE  -
sd OLDDISK-01   Storage-02 ENABLED  377487360 0        -        -
root@ny-puredisk:vxedit -g PDDG rename sdc NEWDISK

```

To shrink a volume as in the example *Storage* volume

◆ Use the `vxresize` command:

```
# vxresize -f -t my-shrinktask -g PDDG Storage 193g
```

The original physical disk (“dm”) that has been grouped into the *PDDG* diskgroup is called *sdb* but we have assigned the internal name *OLDDISK* for the purpose of this example. This can be done with the `vxedit` command using the `rename` operand. We also see the new disk (*sdc*) under VxVM control. It has been initialized but not yet assigned to any disk group.

```

root@ny-puredisk:vxdisk list
DEVICE  TYPE          DISK      GROUP    STATUS
sda     auto:none    -         -        online invalid
sdb     auto:sliced  OLDDISK  PDDG    online
sdc     auto:none    -         -        online invalid
root@ny-puredisk:

```

Adding a new disk to the disk group

The next step is adding the new disk into the *PDDG* disk group and assigning the name of *NEWDISK* to the disk.

To add a new disk to the example disk group

- 1 Initialize the disk.
- 2 Add the disk into the *PDDG* disk group

```

root@ny-puredisk:~#vxedit init sdc
root@ny-puredisk:~#vxchg -g PDDG adddisk sdc

root@ny-puredisk:~#vxprint
Disk group: PDDG

```

TY NAME	ASSOC	KSTATE	LENGTH	PLOFFS	STATE	TUTILO	POTILO
dg PDDG	PDDG	-	-	-	-	-	-
dm OLDDISK	sdb	-	733894672	-	-	-	-
dm sdc	sdc	-	417301712	-	-	-	-
v Storage	fsgen	ENABLED	377487360	-	ACTIVE	-	-
pl Storage-02	Storage	ENABLED	377487360	-	ACTIVE	-	-
sd OLDDISK-01	Storage-02	ENABLED	377487360	0	-	-	-

The internal VxVM name of the new disk is changed from the default *disk* to *NEWDISK*.

```

root@ny-puredisk:~#vxedit -g PDDG rename sdc NEWDISK
root@ny-puredisk:~#vxprint
Disk group: PDDG

```

TY NAME	ASSOC	KSTATE	LENGTH	PLOFFS	STATE	TUTILO	POTILO
dg PDDG	PDDG	-	-	-	-	-	-
dm OLDDISK	sdb	-	733894672	-	-	-	-
dm NEWDISK	sdc	-	417301712	-	-	-	-
v Storage	fsgen	ENABLED	377487360	-	ACTIVE	-	-
pl Storage-02	Storage	ENABLED	377487360	-	ACTIVE	-	-
sd OLDDISK-01	Storage-02	ENABLED	377487360	0	-	-	-

Mirroring

The next step is to start the mirroring process. We used the `vxassist` command to transform the *Storage* volume from a simple, concatenated volume into a mirrored volume. Optionally, a DRL (Dirty Region Log) can be added to the volume. If enabled, the DRL speeds recovery of mirrored volumes after a system crash. It requires an additional 1 Megabyte of extra disk space.

```

root@ny-puredisk:~# /usr/sbin/vxassist -b -g PDDG mirror Storage layout=nostripe alloc=NEWDISK
root@ny-puredisk:~# vxprint
Disk group: PDDG

```

TY NAME	ASSOC	KSTATE	LENGTH	PLOFFS	STATE	TUTILO	POTILO
dg PDDG	PDDG	-	-	-	-	-	-
dm OLDDISK	sdb	-	733894672	-	-	-	-
dm NEWDISK	sdc	-	417301712	-	-	-	-
v Storage	fsgen	ENABLED	377487360	-	ACTIVE	ATT1	-
pl Storage-01	Storage	ENABLED	377487360	-	TEMPMED	ATT	-
sd NEWDISK-01	Storage-01	ENABLED	377487360	0	-	-	-
pl Storage-02	Storage	ENABLED	377487360	-	ACTIVE	-	-
sd NEWDISK-01	Storage-02	ENABLED	377487360	0	-	-	-

To add a DRL as in the example *Storage* volume

- ◆ Use:

```
# vxassist -g PDDG addlog Storage logtype=drl
```

For more information about DRL logging,:

See the *Veritas Storage Foundation Administrator's Guide*

Monitoring

The mirroring process must complete before you can proceed. During this time there may be a heavy I/O load on the system as the mirroring process reads from one disk and writes to another.

To monitor the mirroring progress

- ◆ Use the `vxtask list` command.

Raw I/O statistics can also be monitored with the `vxstat` command. Mirroring should be done either during times of low demand on the server, or, optionally, to have the services stopped completely. While the initial synchronization is underway, the STATE of the new plex is still TEMPRMSD.

```
root@ny-puredisk:~# vxprint
Disk group: PDDG
```

TY	NAME	ASSOC	RSTATE	LENGTH	PLOFFS	STATE	TUTILO	PUTILO
dg	PDDG	PDDG	-	-	-	-	-	-
dm	OLDDISK	sdb	-	733894672	-	-	-	-
dm	NEWDISK	sdc	-	417301712	-	-	-	-
v	Storage	fsagen	ENABLED	377487360	-	ACTIVE	ATT1	-
pl	Storage-01	Storage	ENABLED	377487360	-	TEMPRMSD	ATT	-
sd	NEWDISK-01	Storage-01	ENABLED	377487360	0	-	-	-
pl	Storage-02	Storage	ENABLED	377487360	-	ACTIVE	-	-
sd	OLDDISK-01	Storage-02	ENABLED	377487360	0	-	-	-

To pause and resume mirroring

- ◆ Use the `vxtask` command.

To throttle the mirroring process and free up I/O if needed

- ◆ Use the `vxtask` command.

```
root@ny-puredisk:~# vxtask list
TASKID  PTID  TYPE/STATE  PCT  PROGRESS
228          ATCOPY/R  00.12%  0/377487360/449792  PLXATT Storage Storage-01 PDDG
```

The TEMPRMSD plex state is used by `vxassist` when attaching new data plexes to a volume. If the synchronization operation does not complete, the plex and its subdisks are removed.

See <http://www.symantec.com/docs/TECH19044>

Mirror completion

When the mirroring completes, you can see that the output from `vxprint` shows the volume now has two active plexes associated with it. This is the mirrored volume comprised of two plexes, each plex residing on separate physical storage arrays.

```

root@ny-puredisk:~# vxtask list
root@ny-puredisk:~# vxprint
Disk group: PDDG

```

TY	NAME	ASSOC	HSTATE	LENGTH	PLOFFS	STATE	TUTIL0	PUTIL0
dq	PDDG	PDDG	-	-	-	-	-	-
dm	OLDDISK	sdb	-	733894672	-	-	-	-
dm	NEWDISK	sdc	-	417301712	-	-	-	-
v	Storage	fsgen	ENABLED	377487360	-	ACTIVE	-	-
pl	Storage-01	Storage	ENABLED	377487360	-	ACTIVE	-	-
sd	NEWDISK-01	Storage-01	ENABLED	377487360	0	-	-	-
pl	Storage-02	Storage	ENABLED	377487360	-	ACTIVE	-	-
sd	OLDDISK-01	Storage-02	ENABLED	377487360	0	-	-	-

To confirm completion of the mirror task

- ◆ Use the `vxtask` command.

Removing old storage

After the mirroring process completes, you can remove the old storage.

To remove the old storage

- 1 Break the mirror.
- 2 Logically detach the old disk (*OLDDISK*, *sdb*).
- 3 Check the viability of the volume. Services do not need to be stopped during this phase.
- 4 Remove the old storage from the diskgroup.


```

root@ny-puredisk:~# vxpdg -g PDDG rmdisk OLDDISK
root@ny-puredisk:~# vxdisk list
DEVICE      TYPE      DISK      GROUP      STATUS
sda         auto:none -          -          online invalid
sdb         auto:sliced NEWDISK    PDDG       online
sdc         auto:sliced NEWDISK    PDDG       online

root@ny-puredisk:~# vxassist -g PDDG remove mirror Storage alloc=\!OLDDISK ←(note)
root@ny-puredisk:~# vxprint
Disk group: PDDG
TY NAME      ASSOC      KSTATE  LENGTH  PLOFFS  STATE  TUTILO  FUTILO
dg PDDG      PDDG      -        -        -        -        -        -
dm NEWDISK    sdc       -        417301712 -        -        -        -
dm OLDDISK    sdb       -        733894672 -        -        -        -
v Storage    fsgen     ENABLED  377487360 -        ACTIVE  -        -
pl Storage-01 Storage    ENABLED  377487360 -        ACTIVE  -        -
sd NEWDISK-01 Storage-01 ENABLED  377487360 0        -        -        -

```

The use of the backslash is necessary to override the significance of "!" to the bash Shell which is the default shell for root user. Without the "\", the bash (or C Shell) command line interpreter would look for some history of command event.

Post-mirroring steps

The last step is to check on application services by running whatever utilities you have to ensure the applicatio is up. At some point, a reboot should be done at this point to ensure that the system properly starts and can access the disks during a reboot. No additional modifications need to be made to the file system mount table (etc/fstab, for example) since all storage, disk group, and volume object names remain unchanged.

Migrating data between platforms

This chapter includes the following topics:

- [Overview of the Cross-Platform Data Sharing \(CDS\) feature](#)
- [CDS disk format and disk groups](#)
- [Setting up your system to use Cross-platform Data Sharing \(CDS\)](#)
- [Maintaining your system](#)
- [File system considerations](#)
- [Alignment value and block size](#)
- [Migrating a snapshot volume](#)

Overview of the Cross-Platform Data Sharing (CDS) feature

This section presents an overview of the Cross-Platform Data Sharing (CDS) feature of Symantec's Veritas Storage Foundation™ software. CDS provides you with a foundation for moving data between different systems within a heterogeneous environment. The machines may be running HP-UX, AIX, Linux or the Solaris™ operating system (OS), and they may all have direct access to physical devices holding data. CDS allows Symantec's Veritas products and applications to access data storage independently of the operating system platform, enabling them to work transparently in heterogeneous environments.

The Cross-Platform Data Sharing feature is also known as Portable Data Containers (PDC). For consistency, this document uses the name Cross-Platform Data Sharing throughout.

The following levels in the device hierarchy, from disk through file system, must provide support for CDS to be used:

End-user applications	Application level.
Veritas™ File System (VxFS)	File system level.
Veritas™ Volume Manager (VxVM)	Volume level.
Operating system	Device level.

CDS is a license-enabled feature that is supported at the disk group level by VxVM and at the file system level by VxFS.

CDS utilizes a new disk type (`auto:cdsdisk`). To effect data sharing, VxVM supports a new disk group attribute (`cds`) and also supports different OS block sizes.

Note: CDS allows data volumes and their contents to be easily migrated between heterogeneous systems. It does not enable concurrent access from different types of platform unless such access is supported at all levels that are required.

Shared data across platforms

While volumes can be exported across platforms, the data on the volumes can be shared only if data sharing is supported at the application level. That is, to make data sharing across platforms possible, it must be supported throughout the entire software stack.

For example, if a VxFS file system on a VxVM volume contains files comprising a database, then the following functionality applies:

- Disks can be recognized (as `cds` disks) across platforms.
- Disk groups can be imported across platforms.
- The file system can be mounted on different platforms.

However, it is very likely that, because of the inherent characteristics of databases, you may not be able to start up and use the database on a platform different from the one on which it was created.

An example is where an executable file, compiled on one platform, can be accessed across platforms (using CDS), but may not be executable on a different platform.

Note: You do not need a file system in the stack if the operating system provides access to raw disks and volumes, and the application can utilize them. Databases and other applications can have their data components built on top of raw volumes without having a file system to store their data files.

Disk drive sector size

Sector size is an attribute of a disk drive (or SCSI LUN for an array-type device), which is set when the drive is formatted. Sectors are the smallest addressable unit of storage on the drive, and are the units in which the device performs I/O. The sector size is significant because it defines the atomic I/O size at the device level. Any multi-sector writes which VxVM submits to the device driver are not guaranteed to be atomic (by the SCSI subsystem) in the case of system failure.

Block size issues

The block size is a platform-dependent value that is greater than or equal to the sector size. Each platform accesses the disk on block boundaries and in quantities that are multiples of the block size.

Data that is created on one platform, and then accessed by a platform of a different block size, can suffer from the following problems:

- Addressing issues
- The data may not have been created on a block boundary compatible with that used by the accessing platform.
 - The accessing platform cannot address the start of the data.

Bleed-over issues The size of the data written may not be an exact multiple of the block size used by the accessing platform. Therefore the accessing platform cannot constrain its I/O within the boundaries of the data on disk.

Operating system data

Some operating systems (OS) require OS-specific data on disks in order to recognize and control access to the disk.

CDS disk format and disk groups

This section provides additional information about CDS disk format and CDS disk groups.

CDS disk access and format

For a disk to be accessible by multiple platforms, the disk must be consistently recognized by the platforms, and all platforms must be capable of performing I/O on the disk. CDS disks contain specific content at specific locations to identify or control access to the disk on different platforms. The same content and location are used on all CDS disks, independent of the platform on which the disks are initialized.

In order for a disk to be initialized as, or converted to a CDS disk, it must satisfy the following requirements:

- Must be a SCSI disk
- Must be the entire physical disk (LUN)
- Only one volume manager (such as VxVM) can manage a physical disk (LUN)
- There can be no disk partition (slice) which is defined, but which is not configured on the disk
- Cannot contain a volume whose use-type is either `root` or `swap` (for example, it cannot be a boot disk)

The CDS conversion utility, `vxcdsconvert`, is provided to convert non-CDS VM disk formats to CDS disks, and disk groups with a version number less than 110 to disk groups that support CDS disks.

See [“Converting non-CDS disks to CDS disks”](#) on page 279.

CDS disk types

The CDS disk format, `cdsdisk`, is recognized by all VxVM platforms. The `cdsdisk` disk format is the default for all newly-created VM disks unless overridden in a defaults file. The `vxcdsconvert` utility is provided to convert other disk formats and types to CDS.

See [“Defaults files”](#) on page 282.

Note: Disks with format `cdsdisk` can only be added to disk groups with version 110 or later.

Private and public regions

A VxVM disk usually has a private and a public region.

The private region is a small area on the disk where VxVM configuration information is stored, such as a disk header label, configuration records for VxVM objects (such as volumes, plexes and subdisks), and an intent log for the

configuration database. The default private region size is 32MB, which is large enough to record the details of several thousand VxVM objects in a disk group.

The public region covers the remainder of the disk, and is used for the allocation of storage space to subdisks.

The private and public regions are aligned and sized in multiples of 8K to permit the operation of CDS. The alignment of VxVM objects within the public region is controlled by the disk group alignment attribute. The value of the disk group alignment attribute must also be 8K to permit the operation of CDS.

Note: With other (non-CDS) VxVM disk formats, the private and public regions are aligned to the platform-specific OS block size.

Disk access type auto

The disk access (DA) type `auto` supports multiple disk formats, including `cdsdisk`, which is supported across all platforms. It is associated with the DA records created by the VxVM auto-configuration mode. Disk type `auto` automatically determines which format is on the disk.

Platform block

The platform block resides on disk sector 0, and contains data specific to the operating system for the platforms. It is necessary for proper interaction with each of those platforms. The platform block allows a disk to perform as if it was initialized by each of the specific platforms.

AIX coexistence label

The AIX coexistence label resides on the disk, and identifies the disk to the AIX logical volume manager (LVM) as being controlled by VxVM.

HP-UX coexistence label

The HP-UX coexistence label resides on the disk, and identifies the disk to the HP logical volume manager (LVM) as being controlled by VxVM.

VxVM ID block

The VxVM ID block resides on the disk, and indicates the disk is under VxVM control. It provides dynamic VxVM private region location and other information.

About Cross-platform Data Sharing (CDS) disk groups

A Cross-platform Data Sharing (CDS) disk group allows cross-platform data sharing of Veritas Volume Manager (VxVM) objects, so that data written on one of the

supported platforms may be accessed on any other supported platform. A CDS disk group is composed only of CDS disks (VxVM disks with the disk format `cdsdisk`), and is only available for disk group version 110 and greater.

Starting with disk group version 160, CDS disk groups can support disks of greater than 1 TB.

Note: The CDS conversion utility, `vxcdsconvert`, is provided to convert non-CDS VxVM disk formats to CDS disks, and disk groups with a version number less than 110 to disk groups that support CDS disks.

See [“Converting non-CDS disks to CDS disks”](#) on page 279.

All VxVM objects in a CDS disk group are aligned and sized so that any system can access the object using its own representation of an I/O block. The CDS disk group uses a platform-independent alignment value to support system block sizes of up to 8K.

See [“Disk group alignment”](#) on page 273.

CDS disk groups can be used in the following ways:

- Initialized on one system and then used “as-is” by VxVM on a system employing a different type of platform.
- Imported (in a serial fashion) by Linux, Solaris, AIX, and HP-UX systems.
- Imported as private disk groups, or shared disk groups (by CVM).

You cannot include the following disks or volumes in a CDS disk group:

- Volumes of usage type `root` and `swap`. You cannot use CDS to share boot devices.
- Encapsulated disks.

Note: On Solaris and Linux systems, the process of disk encapsulation places the slices or partitions on a disk (which may contain data or file systems) under VxVM control. On AIX and HP-UX systems, LVM volumes may similarly be converted to VxVM volumes.

Device quotas

Device quotas limit the number of objects in the disk group which create associated device nodes in the file system. Device quotas are useful for disk groups which to be transferred between Linux with a pre-2.6 kernel and other supported platforms. Prior to the 2.6 kernel, Linux supported only 256 minor devices per major device.

You can limit the number of devices that can be created in a given CDS disk group by setting the device quota.

See [“Setting the maximum number of devices for CDS disk groups”](#) on page 290.

When you create a device, an error is returned if the number of devices would exceed the device quota. You then either need to increase the quota, or remove some objects using device numbers, before the device can be created.

See [“Displaying the maximum number of devices in a CDS disk group”](#) on page 294.

Minor device numbers

Importing a disk group will fail if it will exceed the maximum devices for that platform.

Note: There is a large disparity between the maximum number of devices allowed for devices on the Linux platform with a pre-2.6 kernel, and that for other supported platforms.

Non-CDS disk groups

Any version 110 (or greater) disk group (DG) can contain both CDS and non-CDS disks. However, only version 110 (or greater) disk groups composed entirely of CDS disks have the ability to be shared across platforms. Whether or not that ability has been enabled is controlled by the `cds` attribute of the disk group. Enabling this attribute causes a non-CDS disk group to become a CDS disk group.

Although a non-CDS disk group can contain a mixture of CDS and non-CDS disks having dissimilar private region alignment characteristics, its disk group alignment will still direct how all subdisks are created.

Disk group alignment

One of the attributes of the disk group is the block alignment, which represents the largest block size supported by the disk group.

The alignment constrains the following attributes of the objects within a disk group:

- Subdisk offset
- Subdisk length
- Plex offset
- Volume length
- Log length
- Stripe width

The offset value specifies how an object is positioned on a drive.
The disk group alignment is assigned at disk group creation time.
See [“Disk group tasks”](#) on page 287.

Alignment values

The disk group block alignment has two values: 1 block or 8k (8 kilobytes).

All CDS disk groups must have an alignment value of 8k.

All disk group versions before version 110 have an alignment value of 1 block, and they retain this value if they are upgraded to version 110 or later.

A disk group that is not a CDS disk group, and which has a version of 110 and later, can have an alignment value of either 1 block or 8k.

The alignment for all newly initialized disk groups in VxVM 4.0 and later releases is 8k. This value, which is used when creating the disk group, cannot be changed. However, the disk group alignment can be subsequently changed.

See [“Changing the alignment of a non-CDS disk group”](#) on page 288.

Note: The default usage of `vxassist` is to set the `layout=diskalign` attribute on all platforms. The `layout` attribute is ignored on 8K-aligned disk groups, which means that scripts relying on the default may fail.

Dirty region log alignment

The location and size of each map within a dirty region log (DRL) must not violate the disk group alignment for the disk group (containing the volume to which the DRL is associated). This means that the region size and alignment of each DRL map must be a multiple of the disk group alignment, which for CDS disk groups is 8K. (Features utilizing the region size can impose additional minimums and size increments over and above this restriction, but cannot violate it.)

In a version 110 disk group, a traditional DRL volume has the following region requirements:

- Minimum region size of 512K
- Incremental region size of 64K

In a version 110 disk group, an instant snap DCO volume has the following region requirements:

- Minimum region size of 16K
- Incremental region size of 8K

Object alignment during volume creation

For CDS disk groups, VxVM objects that are used in volume creation are automatically aligned to 8K. For non-CDS disk groups, the `vxassist` attribute, `dgaligned_checking`, controls how the command handles attributes that are subject to disk group alignment restrictions. If set to `strict`, the volume length and values of attributes must be integer multiples of the disk group alignment value, or the command fails and an error message is displayed. If set to `round` (default), attribute values are rounded up as required. If this attribute is not specified on the command-line or in a defaults file, the default value of `round` is used.

The `diskalign` and `nodiskalign` attributes of `vxassist`, which control whether subdisks are aligned on cylinder boundaries, is honored only for non-CDS disk groups whose alignment value is set to 1.

Setting up your system to use Cross-platform Data Sharing (CDS)

In order to migrate data between platforms using Cross-platform Data Sharing (CDS), set up your system to use CDS disks and CDS disk groups. The CDS license must be enabled. You can use the default files to configure the appropriate settings for CDS disks and disk groups.

[Table 17-1](#) describes the tasks for setting up your system to use CDS.

Table 17-1 Setting up CDS disks and CDS disk groups

Task	Procedures
Create the CDS disks.	You can create a CDS disk in one of the following ways: <ul style="list-style-type: none"> ■ Creating CDS disks from uninitialized disks See “Creating CDS disks from uninitialized disks” on page 276. ■ Creating CDS disks from initialized VxVM disks See “Creating CDS disks from initialized VxVM disks” on page 277. ■ Converting non-CDS disks to CDS disks See “Converting non-CDS disks to CDS disks” on page 279.

Table 17-1 Setting up CDS disks and CDS disk groups (*continued*)

Task	Procedures
Create the CDS disk groups.	<p>You can create a CDS disk group in one of the following ways:</p> <ul style="list-style-type: none"> ■ Creating CDS disk groups See “Creating CDS disk groups” on page 278. ■ Converting a non-CDS disk group to a CDS disk group See “Converting a non-CDS disk group to a CDS disk group” on page 280.
Verify the CDS license.	Verifying licensing See “Verifying licensing” on page 282.
Verify the system defaults related to CDS.	Defaults files See “Defaults files” on page 282.

Creating CDS disks from uninitialized disks

You can create a CDS disk from an uninitialized disk by using one of the following methods:

- [Creating CDS disks by using vxdisksetup](#)
- [Creating CDS disks by using vxdiskadm](#)

Creating CDS disks by using vxdisksetup

To create a CDS disk by using the `vxdisksetup` command

- Type the following command:

```
# vxdisksetup -i disk [format=disk_format]
```

The format defaults to `cdsdisk` unless this is overridden by the `/etc/default/vxdisk` file, or by specifying the disk format as an argument to the `format` attribute.

See [“Defaults files”](#) on page 282.

See the `vxdisksetup(1M)` manual page.

Creating CDS disks by using vxdiskadm

To create a CDS disk by using the `vxdiskadm` command

- Run the `vxdiskadm` command, and select the “Add or initialize one or more disks” item from the main menu. You are prompted to specify the format.

Warning: On CDS disks, the CDS information occupies the first sector of that disk, and there is no `fdisk` partition information. Attempting to create an `fdisk` partition (for example, by using the `fdisk` or `format` commands) erases the CDS information, and can cause data corruption.

Creating CDS disks from initialized VxVM disks

How you create a CDS disk depends on the current state of the disk, as follows:

- [Creating a CDS disk from a disk that is not in a disk group](#)
- [Creating a CDS disk from a disk that is already in a disk group](#)

Creating a CDS disk from a disk that is not in a disk group

To create a CDS disk from a disk that is not in a disk group

- 1 Run the following command to remove the VM disk format for the disk:

```
# vxdiskunsetup disk
```

This is necessary as non-auto types cannot be reinitialized by `vxdisksetup`.

- 2 If the disk is listed in the `/etc/vx/darecs` file, remove its disk access (DA) record using the command:

```
# vxdisk rm disk
```

(Disk access records that cannot be configured by scanning the disks are stored in an ordinary file, `/etc/vx/darecs`, in the root file system. Refer to the `vxintro(1M)` manual page for more information.)

- 3 Rescan for the disk using this command:

```
# vxdisk scandisks
```

- 4 Type this command to set up the disk:

```
# vxdisksetup -i disk
```

Creating a CDS disk from a disk that is already in a disk group

To create a CDS disk from a disk that is already in a disk group

- Run the `vxcdsconvert` command.
See “[Converting non-CDS disks to CDS disks](#)” on page 279.

Creating CDS disk groups

You can create a CDS disk group in the following ways:

- [Creating a CDS disk group by using `vx dg init`](#)
- [Creating a CDS disk group by using `vx diskadm`](#)

Creating a CDS disk group by using `vx dg init`

Note: The disk group version must be 110 or greater.

To create a CDS disk group by using the `vx dg init` command

- Type the following command:

```
# vx dg init diskgroup disklist [cds={on|off}]
```

The format defaults to a CDS disk group, unless this is overridden by the `/etc/default/vx dg` file, or by specifying the `cds` argument.

See the `vx dg(1M)` manual page for more information.

Creating a CDS disk group by using `vx diskadm`

You cannot create a CDS disk group when encapsulating an existing disk, or when converting an LVM volume.

When initializing a disk, if the target disk group is an existing CDS disk group, `vx diskadm` will only allow the disk to be initialized as a CDS disk. If the target disk group is a non-CDS disk group, the disk can be initialized as either a CDS disk or a non-CDS disk.

If you use the `vx diskadm` command to initialize a disk into an existing CDS disk group, the disk must have been added with the `cdsdisk` format.

The CDS attribute for the disk group remains unchanged by this procedure.

To create a CDS disk group by using the `vx diskadm` command

- Run the `vxdiskadm` command, and select the “Add or initialize one or more disks” item from the main menu. Specify that the disk group should be a CDS disk group when prompted.

Converting non-CDS disks to CDS disks

Note: The disks must be of type of `auto` in order to be re-initialized as CDS disks.

To convert non-CDS disks to CDS disks

- 1 If the conversion is not going to be performed on-line (that is, while access to the disk group continues), stop any applications that are accessing the disks.
- 2 Make sure that the disks have free space of at least 256 sectors before doing the conversion.
- 3 Add a disk to the disk group for use by the conversion process. The conversion process evacuates objects from the disks, reinitializes the disks, and relocates objects back to the disks.

Note: If the disk does not have sufficient free space, the conversion process will not be able to relocate objects back to the disk. In this case, you may need to add additional disks to the disk group.

- 4 Type one of the following forms of the CDS conversion utility (`vxcdsconvert`) to convert non-CDS disks to CDS disks.

```
# vxcdsconvert -g diskgroup [-A] [-d defaults_file] \  
  [-o novolstop] disk name [attribute=value] ...  
# vxcdsconvert -g diskgroup [-A] [-d defaults_file] \  
  [-o novolstop] alldisks [attribute=value] ...
```

The `alldisks` and `disk` keywords have the following effect

<code>alldisks</code>	Converts all non-CDS disks in the disk group into CDS disks.
-----------------------	--

`disk`

Specifies a single disk for conversion. You would use this option under the following circumstances:

- If a disk in the non-CDS disk group has cross-platform exposure, you may want other VxVM nodes to recognize the disk, but not to assume that it is available for initialization.
- If the native Logical Volume Manager (LVM) that is provided by the operating system needs to recognize CDS disks, but it is not required to initialize or manage these disks.
- Your intention is to move the disk into an existing CDS disk group.

Specify the `-o novolstop` option to perform the conversion on-line (that is, while access to the disk continues). If the `-o novolstop` option is not specified, stop any applications that are accessing the disks, and perform the conversion off-line.

Warning: Specifying the `-o novolstop` option can greatly increase the amount of time that is required to perform conversion.

Before you use the `vxcdsconvert` command, make sure you understand its options, attributes, and keywords.

See the `vxcdsconvert(1M)` manual page.

Converting a non-CDS disk group to a CDS disk group

To convert a non-CDS disk group to a CDS disk group

- 1 If the disk group contains one or more disks that you do not want to convert to CDS disks, use the `vxchg move` or `vxchg split` command to move the disks out of the disk group.
- 2 The disk group to be converted must have the following characteristics:
 - No dissociated or disabled objects.
 - No sparse plexes.
 - No volumes requiring recovery.
 - No volumes with pending snapshot operations.
 - No objects in an error state.

To verify whether a non-CDS disk group can be converted to a CDS disk group, type the following command:


```
# vxcdsconvert -g diskgroup -A group
```

- 3 If the disk group does not have a CDS-compatible disk group alignment, the objects in the disk group must be relayed out with a CDS-compatible alignment.
- 4 If the conversion is not going to be performed on-line (that is, while access to the disk group continues), stop any applications that are accessing the disks.
- 5 Type one of the following forms of the CDS conversion utility (`vxcdsconvert`) to convert a non-CDS disk group to a CDS disk group.

```
# vxcdsconvert -g diskgroup [-A] [-d defaults_file] \
  [-o novolstop] alignment [attribute=value] ...
# vxcdsconvert -g diskgroup [-A] [-d defaults_file] \
  [-o novolstop] group [attribute=value] ...
```

The `alignment` and `group` keywords have the following effect:

<code>alignment</code>	Specifies alignment conversion where disks are not converted, and an object relayout is performed on the disk group. A successful completion results in an 8K-aligned disk group. You might consider this option, rather than converting the entire disk group, if you want to reduce the amount of work to be done for a later full conversion to CDS disk group.
<code>group</code>	Specifies group conversion of all non-CDS disks in the disk group before relaying out objects in the disk group.

The conversion involves evacuating objects from the disk, reinitializing the disk, and relocating objects back to disk. You can specify the `-o novolstop` option to perform the conversion on-line (that is, while access to the disk group continues). If the `-o novolstop` option is not specified, stop any applications that are accessing the disks, and perform the conversion off-line.

Warning: Specifying the `-o novolstop` option can greatly increase the amount of time that is required to perform conversion.

Conversion has the following side effects:

- Non-CDS disk groups are upgraded by using the `vxvg upgrade` command. If the disk group was originally created by the conversion of an LVM volume group (VG), rolling back to the original LVM VG is not possible. If you decide to go through with the conversion, the rollback records for the

disk group will be removed, so that an accidental rollback to an LVM VG cannot be done.

- Stopped, but startable volumes, are started for the duration of the conversion .
- Any volumes or other objects in the disk group that were created with the `layout=diskalign` attribute specified can no longer be disk aligned.
- Encapsulated disks may lose the ability to be unencapsulated.
- Performance may be degraded because data may have migrated to different regions of a disk, or to different disks.

In the following example, the disk group, `mydg`, and all its disks are converted to CDS while keeping its volumes are still online:

```
# vxcdsconvert -g mydg -o novolstop group \  
  move_subdisks_ok=yes evac_subdisks_ok=yes \  
  evac_disk_list=disk11,disk12,disk13,disk14
```

The `evac_disk_list` attribute specifies a list of disks (`disk11` through `disk14`) to which subdisks can be evacuated to disks if required.

Before you use the `vxcdsconvert` command, make sure you understand its options, attributes, and keywords.

See the `vxcdsconvert(1M)` manual page.

Verifying licensing

The ability to create or import a CDS disk group is controlled by a CDS license. CDS licenses are included as part of the Veritas Storage Foundation license.

To verify the CDS enabling license

- Type the following command:

```
# vxlicrep
```

Verify the following line in the output:

```
Cross-platform Data Sharing = Enabled
```

Defaults files

The following system defaults files in the `/etc/default` directory are used to specify the alignment of VxVM objects, the initialization or encapsulation of VM

disks, the conversion of LVM disks, and the conversion of disk groups and their disks to the CDS-compatible format

`vxassist` Specifies default values for the following parameters to the `vxcdsconvert` command that have an effect on the alignment of VxVM objects: `dgalign_checking`, `diskalign`, and `nodiskalign`.
 See [“Object alignment during volume creation”](#) on page 275.
 See the `vxassist(1M)` manual page.

`vxcdsconvert` Specifies default values for the following parameters to the `vxcdsconvert` command: `evac_disk_list`, `evac_subdisks_ok`, `min_split_size`, `move_subdisks_ok`, `privlen`, and `split_subdisks_ok`.

The following is a sample `vxcdsconvert` defaults file:

```
evac_subdisks_ok=no
min_split_size=64k
move_subdisks_ok=yes
privlen=2048
split_subdisks_ok=move
```

An alternate defaults file can be specified by using the `-d` option with the `vxcdsconvert` command.

See the `vxcdsconvert(1M)` manual page.

`vx dg` Specifies default values for the `cds`, `default_activation_mode` and `enable_activation` parameters to the `vx dg` command. The `default_activation_mode` and `enable_activation` parameters are only used with shared disk groups in a cluster.

The following is a sample `vx dg` defaults file:

```
cds=on
```

See the `vx dg(1M)` manual page.

`vxdisk` Specifies default values for the `format` and `privlen` parameters to the `vxdisk` and `vxdisksetup` commands. These commands are used when disks are initialized by VxVM for the first time. They are also called implicitly by the `vxdiskadm` command and the Veritas Operations Manager (VOM) GUI.

The following is a sample `vxdisk` defaults file:

```
format=cdsdisk
privlen=2048
```

See the `vxdisk(1M)` manual page.

See the `vxdisksetup(1M)` manual page.

`vxencap` Specifies default values for the `format`, `privlen`, `privoffset` and `puboffset` parameters to the `vxencap` and `vxlvencap` commands. These commands are used when disks with existing partitions or slices are encapsulated, or when LVM disks are converted to VM disks. It is also called implicitly by the `vxdiskadm`, `vxconvert` (on AIX) and `vxvmconvert` (on HP-UX) commands, and by the VOM.

The following is a sample `vxencap` defaults file:

```
format=sliced
privlen=4096
privoffset=0
puboffset=1
```

See the `vxencap(1M)` manual page.

See the `vxconvert(1M)` manual page.

See the `vxvmconvert(1M)` manual page.

In the defaults files, a line that is empty, or that begins with a “#” character in the first column, is treated as a comment, and is ignored.

Apart from comment lines, all other lines must define attributes and their values using the format `attribute=value`. Each line starts in the first column, and is terminated by the value. No white space is allowed around the = sign.

Maintaining your system

You may need to perform maintenance tasks on the CDS disks and CDS disk groups. Refer to the respective section for each type of task.

- Disk tasks

See “[Disk tasks](#)” on page 285.

- Disk group tasks
See “[Disk group tasks](#)” on page 287.
- Displaying information
See “[Displaying information](#)” on page 293.
- Default activation mode of shared disk groups
See “[Default activation mode of shared disk groups](#)” on page 296.
- Additional considerations when importing CDS disk groups
See “[Defaults files](#)” on page 282.

Disk tasks

The following disk tasks are supported:

- [Changing the default disk format](#)
- [Restoring CDS disk labels](#)

Changing the default disk format

When disks are put under VxVM control, they are formatted with the default `cdsdisk` layout. This happens during the following operations:

- Initialization of disks
- Encapsulation of disks with existing partitions or slices (Linux and Solaris systems)
- Conversion of LVM disks (AIX, HP-UX and Linux systems)

You can override this behavior by changing the settings in the system defaults files. For example, you can change the default format to `sliced` for disk initialization by modifying the definition of the `format` attribute in the `/etc/default/vxdisk` defaults file.

To change the default format for disk encapsulation or LVM disk conversion

- Edit the `/etc/default/vxencap` defaults file, and change the definition of the `format` attribute.
See “[Defaults files](#)” on page 282.

Restoring CDS disk labels

CDS disks have the following labels:

- Platform block

- AIX coexistence label
- HP-UX coexistence or VxVM ID block

There are also backup copies of each. If any of the primary labels become corrupted, VxVM will not bring the disk online and user intervention is required.

If two labels are intact, the disk is still recognized as a `cdsdisk` (though in the error state) and `vxdisk flush` can be used to restore the CDS disk labels from their backup copies.

Note: For disks larger than 1 TB, `cdsdisks` use the EFI layout. The procedure to restore disk labels does not apply to `cdsdisks` with EFI layout.

Primary labels are at sectors 0, 7, and 16; and a normal flush will not flush sectors 7 and 16. Also, the private area is not updated as the disk is not in a disk group. There is no means of finding a “good” private region to flush from. In this case, it is possible to restore the CDS disk labels from the existing backups on disk using the flush operation.

If a corruption happened after the labels were read and the disk is still online and part of a disk group, then a flush operation will also flush the private region.

Warning: Caution and knowledge must be employed because the damage could involve more than the CDS disk labels. If the damage is constrained to the first 128K, the disk flush would fix it. This could happen if another system on the fabric wrote a disk label to a disk that was actually a CDS disk in some disk group.

To rewrite the CDS ID information on a specific disk

- Type the following command:

```
# vxdisk flush disk_access_name
```

This rewrites all labels except sectors 7 and 16.

To rewrite all the disks in a CDS disk group

- Type the following command:

```
# vxdg flush diskgroup
```

This rewrites all labels except sectors 7 and 16.

To forcibly rewrite the AIX coexistence label in sector 7 and the HP-UX coexistence label or VxVM ID block in sector 16

- Type the following command:

```
# vxdisk -f flush disk_access_name
```

This command rewrites all labels if there exists a valid VxVM ID block that points to a valid private region. The `-f` option is required to rewrite sectors 7 and 16 when a disk is taken offline due to label corruption (possibly by a Windows system on the same fabric).

Disk group tasks

The following disk group tasks are supported:

- [Changing the alignment of a disk group during disk encapsulation](#)
- [Changing the alignment of a non-CDS disk group](#)
- [Determining the setting of the CDS attribute on a disk group](#)
- [Splitting a CDS disk group](#)
- [Moving objects between CDS disk groups and non-CDS disk groups](#)
- [Moving objects between CDS disk groups](#)
- [Joining disk groups](#)
- [Changing the default CDS setting for disk group creation](#)
- [Creating non-CDS disk groups](#)
- [Upgrading an older version non-CDS disk group](#)
- [Replacing a disk in a CDS disk group](#)
- [Setting the maximum number of devices for CDS disk groups](#)

Changing the alignment of a disk group during disk encapsulation

If you use the `vxdiskadm` command to encapsulate a disk into a disk group with an alignment of 8K, the disk group alignment must be reduced to 1.

If you use the `vxencap` command to perform the encapsulation, the alignment is carried out automatically without a confirmation prompt.

To change the alignment of a disk group during disk encapsulation

- Run the `vxdiskadm` command, and select the “Add or initialize one or more disks” item from the main menu. As part of the encapsulation process, you are asked to confirm that a reduction of the disk group alignment from 8K to 1 is acceptable.

Changing the alignment of a non-CDS disk group

The alignment value can only be changed for disk groups with version 110 or greater.

For a CDS disk group, `alignment` can only take a value of `8k`. Attempts to set the alignment of a CDS disk group to `1` fail unless you first change it to a non-CDS disk group.

Increasing the alignment may require `vxcdsconvert` to be run to change the layout of the objects in the disk group.

To display the current alignment value of a disk group, use the `vxprint` command.

See [“Displaying the disk group alignment”](#) on page 295.

To change the alignment value of a disk group

- Type the `vx dg set` command:

```
# vx dg -g diskgroup set align={1|8k}
```

The operation to increase the alignment to 8K fails if objects exist in the disk group that do not conform to the new alignment restrictions. In that case, use the `vxcdsconvert alignment` command to change the layout of the objects:

```
# vxcdsconvert -g diskgroup [-A] [-d defaults_file] \  
  [-o novolstop] alignment [attribute=value] ...
```

This command increases the alignment value of a disk group and its objects to 8K, without converting the disks.

The sequence 8K to 1 to 8K is possible only using `vx dg set` as long as the configuration does not change after the 8K to 1 transition.

See [“Converting a non-CDS disk group to a CDS disk group”](#) on page 280.

Splitting a CDS disk group

You can use the `vx dg split` command to create a CDS disk group from an existing CDS disk group. The new (target) disk group preserves the setting of the CDS attribute and alignment in the original (source) disk group.

To split a CDS disk group

- Use the `vx dg split` command to split CDS disk groups.
See the *Veritas Volume Manager Administrator’s Guide*.

Moving objects between CDS disk groups and non-CDS disk groups

The alignment of a source non-CDS disk group must be 8K to allow objects to be moved to a target CDS disk group. If objects are moved from a CDS disk group to a target non-CDS disk group with an alignment of 1, the alignment of the target disk group remains unchanged.

To move objects between a CDS disk group and a non-CDS disk group

- Use the `vxdg move` command to move objects between a CDS disk group and a non-CDS disk groups.

See the *Veritas Volume Manager Administrator's Guide*.

Moving objects between CDS disk groups

The disk group alignment does not change as a result of moving objects between CDS disk groups.

To move objects between CDS disk groups

- Use the `vxdg move` command to move objects between CDS disk groups.

See the *Veritas Volume Manager Administrator's Guide*.

Joining disk groups

Joining two CDS disk groups or joining two non-CDS disk groups is permitted, but you cannot join a CDS disk group to a non-CDS disk group. If two non-CDS disk groups have different alignment values, the alignment of the resulting joined disk group is set to 1, and an informational message is displayed.

To join two disk groups

- Use the `vxdg join` command to join two disk groups.

See the *Veritas Volume Manager Administrator's Guide*.

Changing the default CDS setting for disk group creation

To change the default CDS setting for disk group creation

- Edit the `/etc/default/vxdg` file, and change the setting for the `cds` attribute.

Creating non-CDS disk groups

A disk group with a version lower than 110 is given an alignment value equal to 1 when it is imported. This is because the `dg_align` value is not stored in the configuration database for such disk groups.

To create a non-CDS disk group with a version lower than 110

- Type the following `vxdg` command:

```
# vxdg -T version init diskgroup disk_name=disk_access_name
```

Upgrading an older version non-CDS disk group

You may want to upgrade a non-CDS disk group with a version lower than 110 in order to use new features other than CDS. After upgrading the disk group, the `cds` attribute is set to `off`, and the disk group has an alignment of 1.

Note: You must also perform a disk group conversion (using the `vxcdsconvert` utility) to use the CDS feature.

To upgrade the non-CDS pre-version 110 disk group

- Type the following `vxdg` command:

```
# vxdg upgrade diskgroup
```

Replacing a disk in a CDS disk group

Note: When replacing a disk in a CDS disk group, you cannot use a non-CDS disk as the replacement.

To replace a disk in a CDS disk group

- Type the following commands:

```
# vxdg -g diskgroup -k rmdisk disk_name  
# vxdg -g diskgroup -k adddisk disk_name=disk_access_name
```

The `-k` option retains and reuses the disk media record for the disk that is being replaced. The following example shows a disk device `disk21` being reassigned to disk `mydg01`.

```
# vxdg -g diskgroup -k rmdisk mydg01  
# vxdg -g diskgroup -k adddisk mydg01=disk21
```

For other operating systems, use the appropriate device name format.

Setting the maximum number of devices for CDS disk groups

To set the maximum number of devices that can be created in a CDS disk group

- Type the following `vxchg set` command:

```
# vxchg -g diskgroup set maxdev=max-devices
```

The `maxdev` attribute can take any positive integer value that is greater than the number of devices that are currently in the disk group.

Changing the DRL map and log size

If DRL is enabled on a newly-created volume without specifying a log or map size, default values are used. You can use the command line attributes `logmap_len` and `loglen` in conjunction with the `vxassist`, `vxvol`, and `vxmake` commands to set the DRL map and DRL log sizes. The attributes can be used independently, or they can be combined.

You can change the DRL map size and DRL log size only when the volume is disabled and the DRL maps are not in use. Changes can be made to the DRL map size only for volumes in a CDS disk group.

The `logmap_len` attribute specifies the required size, in bytes, for the DRL log. It cannot be greater than the number of bytes available in the map on the disk.

To change the DRL map and log size

- Use the following commands to remove and rebuild the logs:

```
# vxassist -g diskgroup remove log volume nlog=0
# vxassist -g diskgroup addlog volume nlog=nlogs \
  logtype=drl logmap_len=len-bytes [loglen=len-blocks]
```

Note the following restrictions

If only `logmap_len` is specified

The DRL log size is set to the default value (33 * disk group alignment).

If `logmap_len` is greater than (DRL log size) / 2

The command fails, and you need to either provide a sufficiently large `loglen` value or reduce `logmap_len`.

For CDS disk groups

The DRL map and log sizes are set to a minimum of 2 * (disk group alignment).

Creating a volume with a DRL log

To create a volume with a traditional DRL log by using the `vxassist` command

- Type the following command:

```
# vxassist -g diskgroup make volume length mirror=2 \  
  logtype=drl [loglen=len-blocks] [logmap_len=len-bytes]
```

This command creates log subdisks that are each equal to the size of the DRL log.

Note the following restrictions

If neither `logmap_len` nor `loglen` is specified

- `loglen` is set to a default value that is based on disk group alignment.
- `maplen` is set to a reasonable value.

If only `loglen` is specified

- For pre-version 110 disk groups, `maplen` is set to zero.
- For version 110 and greater disk groups, `maplen` is set to use all the bytes available in the on-disk map.

If only `logmap_len` is specified

- For pre-version 110 disk groups, `logmap_len` is not applicable.
- For version 110 and greater disk groups, `maplen` must be less than the number of available bytes in the on-disk map for the default log length.

Setting the DRL map length

To set a DRL map length

- 1 Stop the volume to make the DRL inactive.
- 2 Type the following command:

```
# vxvol -g diskgroup set [loglen=len-blocks] \  
  [logmap_len=len-bytes] volume
```

This command does not change the existing DRL map size.

Note the following restrictions

If both `logmap_len` and `loglen` are specified

- if `logmap_len` is greater than `loglen/2`, `vxvol` fails with an error message. Either increase `loglen` to a sufficiently large value, or decrease `logmap_len` to a sufficiently small value.
- The value of `logmap_len` cannot exceed the number of bytes in the on-disk map.

If `logmap_len` is specified

- The value is constrained by size of the log, and cannot exceed the size of the on-disk map. The size of the on-disk map in blocks can be calculated from the following formula:

$$\text{round}(\text{loglen}/\text{nmaps}) - 24$$

where `nmaps` is 2 for a private disk group, or 33 for a shared disk group.

- The value of `logmap_len` cannot exceed the number of bytes in the on-disk map.

If `loglen` is specified

- Specifying a value that is less than twice the disk group alignment value results in an error message.
- The value is constrained by size of the logging subdisk.

Displaying information

This section describes the following tasks:

- [Determining the setting of the CDS attribute on a disk group](#)
- [Displaying the maximum number of devices in a CDS disk group](#)
- [Displaying map length and map alignment of traditional DRL logs](#)
- [Displaying the disk group alignment](#)
- [Displaying the log map length and alignment](#)
- [Displaying offset and length information in units of 512 bytes](#)

Determining the setting of the CDS attribute on a disk group

To determine the setting of the CDS attribute on a disk group

- Use the `vxdg list` command or the `vxprint` command to determine the setting of the CDS attribute, as shown in the following examples:

```
# vxdg list

NAME                STATE                ID
dgTestSol2         enabled,cds         1063238039.206.vmescl

# vxdg list dgTestSol2

Group:      dgTestSol2
dgid:      1063238039.206.vmescl
import-id: 1024.205
flags:     cds
version:   110
alignment: 8192 (bytes)
.
.
.

# vxprint -F %cds -G -g dgTestSol2

on
```

The disk group, `dgTestSol2`, is shown as having the CDS flag set.

Displaying the maximum number of devices in a CDS disk group

To display the maximum number of devices in a CDS disk group

- Type the following command:

```
# vxprint -g diskgroup -G -F %maxdev
```

Displaying map length and map alignment of traditional DRL logs

To display the map length and map alignment of traditional DRL logs

- Type the following commands

```
# vxprint -g diskgroup -vl volume
# vxprint -g diskgroup -vF '%name %logmap_len %logmap_align' \
volume
```

Displaying the disk group alignment

To display the disk group alignment

- Type the following command:

```
# vxprint -g diskgroup -G -F %align
```

Utilities such as `vxprint` and `vxdg list` that print information about disk group records also output the disk group alignment.

Displaying the log map length and alignment

To display the log map length and alignment

- Type the following command:

```
# vxprint -g diskgroup -lv volume
```

For example, to print information for the volume `vol1` in disk group `dg1`:

```
# vxprint -g dg1 -lv vol1
```

The output is of the form:

```
logging: type=REGION loglen=0 serial=0/0 mapalign=0
maplen=0 (disabled)
```

This indicates a log map alignment (`logmap_align`) value of 0, and a log map length (`logmap_len`) value of 0.

If the log map is set and enabled, the command and results may be in the following form:

```
# vxprint -lv drlvol
```

```
Disk group: dgTestSol
Volume:    drlvol
info:      len=20480
type:      usetype=fsgen
state:     state=ACTIVE kernel=ENABLED cdsrecovery=0/0 (clean)
assoc:     plexes=drlvol-01,drlvol-02,drlvol-03
policies:  read=SELECT (round-robin) exceptions=GEN_DET_SPARSE
flags:     closed writecopy writeback
```

```

logging:  type=REGION loglen=528 serial=0/0 mapalign=16
maplen=512 (enabled)
apprecov: seqno=0/0
recovery: mode=default
recov_id=0
device:  minor=46000 bdev=212/46000 cdev=212/46000
path=/dev/vx/dsk/dgTestSol/drlvol
perms:   user=root group=root mode=0600
guid:    {d968de3e-1dd1-11b2-8fc1-080020d223e5}

```

Displaying offset and length information in units of 512 bytes

To display offset and length information in units of 512 bytes

- Specify the `-b` option to the `vxprint` and `vxdisk` commands, as shown in these examples:

```

# vxprint -bm
# vxdisk -b list

```

Specifying the `-b` option enables consistent output to be obtained on different platforms. Without the `-b` option, the information is output in units of sectors. The number of bytes per sector differs between platforms.

When the `vxprint -bm` or `vxdisk -b list` command is used, the output also contains the `b` suffix, so that the output can be fed back to `vxmake`.

Default activation mode of shared disk groups

The default activation mode of shared disk groups involves a local in-kernel policy that differs between platforms. This means that, regardless of the platform on which the disk group was created, the importing platform will have platform-specific behavior with respect to activation of shared disk groups. Specifically, with the exception of HP-UX, importing a shared disk group results in the volumes being active and enabled for shared-write. In the case of HP-UX, the shared volumes will be inactive and require other actions to activate them for shared-write operations.

Additional considerations when importing CDS disk groups

Before you attempt to use CDS to move disk groups between different operating systems, and if the configuration of the disks has changed since the target system was last rebooted, you should consider the following points

Does the target system know about the disks?

For example, the disks may not have been connected to the system either physically (not cabled) or logically (using FC zoning or LUN masking) when the system was booted up, but they have subsequently been connected without rebooting the system. This can happen when bringing new storage on-line, or when adding an additional DMP path to existing storage. On the target system, both the operating system and VxVM must be informed of the existence of the new storage. Issue the appropriate command to tell the operating system to look for the storage. (On Linux, depending on the supported capabilities of the host adapter, you may need to reboot the target system to achieve this.) Having done this, run either of the following commands on the target system to have VxVM recognize the storage:

```
# vxdctl enable
# vxdisk scandisks
```

Do the disks contain partitions or slices?

Both the Solaris and Linux operating systems maintain information about partitions or slices on disks. If you repartition a disk after the target system was booted, use the appropriate command to instruct the operating system to rescan the disk's TOC or partition table. For example, on a target Linux system, use the following command:

```
# blockdev --rereadpt
```

Having done this, run either of the following commands on the target system to have VxVM recognize the storage:

```
# vxdctl enable
# vxdisk scandisks
```

Has the format of any of the disks changed since the target system was last booted?

For example, if you use the `vxdisksetup -i` command to format a disk for VxVM on one system, the `vxdisk list` command on the target system may still show the format as being `auto:none`. If so, use either of the following commands on the target system to instruct VxVM to rescan the format of the disks:

```
# vxdctl enable
# vxdisk scandisks
```

File system considerations

To set up or migrate volumes with VxFS file systems with CDS, you must consider the file system requirements. This section describes these requirements. It also describes additional tasks required for migrating or setting up in CDS.

Considerations about data in the file system

Data within a file system might not be in the appropriate format to be accessed if moved between different types of systems. For example, files stored in proprietary binary formats often require conversion for use on the target platform. Files containing databases might not be in a standard format that allows their access when moving a file system between various systems, even if those systems use the same byte order. Oracle 10g's Cross-Platform Transportable Tablespace is a notable exception; if used, this feature provides a consistent format across many platforms.

Some data is inherently portable, such as plain ASCII files. Other data is designed to be portable and the applications that access such data are able to access it irrespective of the system on which it was created, such as Adobe PDF files.

Note that the CDS facilities do not convert the end user data. The data is uninterpreted by the file system. Only individual applications have knowledge of the data formats, and thus those applications and end users must deal with this issue. This issue is not CDS-specific, but is true whenever data is moved between different types of systems.

Even though a user might have a file system with data that cannot be readily interpreted or manipulated on a different type of system, there still are reasons for moving such data by using CDS mechanisms. For example, if the desire is to bring a file system off line from its primary use location for purposes of backing it up without placing that load on the server or because the system on which it will be backed up is the one that has the tape devices directly attached to it, then using CDS to move the file system is appropriate.

An example is a principal file server that has various file systems being served by it over the network. If a second file server system with a different operating system was purchased to reduce the load on the original server, CDS can migrate the file system instead of having to move the data to different physical storage over the network, even if the data could not be interpreted or used by either the original or new file server. This is a scenario that often occurs when the data is only accessible or understood by software running on PCs and the file server is UNIX or Linux-based.

File system migration

File system migration refers to the system management operations related to stopping access to a file system, and then restarting these operations to access the file system from a different computer system. File system migration might be required to be done once, such as when permanently migrating a file system to another system without any future desire to move the file system back to its original system or to other systems. This type of file system migration is referred

to as one-time file system migration. When ongoing file system migration between multiple systems is desired, this is known as ongoing file system migration. Different actions are required depending on the kind of migration, as described in the following sections.

Specifying the migration target

Most of the operations performed by the CDS commands require the target to which the file system is to be migrated to be specified by target specifiers in the following format:

```
os_name=name[,os_rel=release][,arch=arch_name]
[,vxfs_version=version][,bits=nbits]
```

The CDS commands require the following target specifiers:

<code>os_name=name</code>	Specifies the name of the target operating system to which the file system is planned to be migrated. Possible values are HP-UX, AIX, SunOS, or Linux. The <code>os_name</code> field must be specified if the target is specified.
<code>os_rel=release</code>	Specifies the operating system release version of the target. For example, 11.31.
<code>arch=arch_name</code>	Specifies the architecture of the target. For example, specify <code>ia</code> or <code>pa</code> for HP-UX.
<code>vxfs_version=version</code>	Specifies the VxFS release version that is in use at the target. For example, 4.1 or 5.0.
<code>bits=nbits</code>	Specifies the kernel bits of the target. <code>nbits</code> can have a value of 32 or 64 to indicate whether the target is running a 32-bit kernel or 64-bit kernel.

While `os_name` must be specified for all `fscdsadm` invocations that permit the target to be specified, all other target specifiers are optional and are available for the user to fine tune the migration target specification.

The CDS commands use the limits information available in the default CDS limits file, `/etc/vx/cdslimitstab`. If the values for the optional target specifiers are not specified, `fscdsadm` will choose the defaults for the specified target based on the information available in the limits file that best fits the specified target, and proceed with the CDS operation. The chosen defaults are displayed to the user before proceeding with the migration.

Note: The default CDS limits information file, `/etc/vx/cdslimitstab`, is installed as part of the VxFS package. The contents of this file are used by the VxFS CDS commands and should not be altered.

Examples of target specifications

The following are examples of target specifications:

<code>os_name=AIX</code>	Specifies the target operating system and use defaults for the remainder.
<code>os_name=HP-UX, os_rel=11.23, arch=ia, vxfs_version=5.0, bits=64</code>	Specifies the operating system, operating system release version, architecture, VxFS version, and kernel bits of the target.
<code>os_name=SunOS, arch=sparc</code>	Specifies the operating system and architecture of the target.
<code>os_name=Linux, bits=32</code>	Specifies the operating system and kernel bits of the target.

Using the `fscdsadm` command

The `fscdsadm` command can be used to perform the following CDS tasks:

- [Checking that the metadata limits are not exceeded](#)
- [Maintaining the list of target operating systems](#)
- [Enforcing the established CDS limits on a file system](#)
- [Ignoring the established CDS limits on a file system](#)
- [Validating the operating system targets for a file system](#)
- [Displaying the CDS status of a file system](#)

Checking that the metadata limits are not exceeded

To check that the metadata limits are not exceeded

- Type the following command to check whether there are any file system entities with metadata that exceed the limits for the specified target operating system:

```
# fscdsadm -v -t target mount_point
```

Maintaining the list of target operating systems

When a file system will be migrated on an ongoing basis between multiple systems, the types of operating systems that are involved in these migrations are maintained in a `target_list` file. Knowing what these targets are allows VxFS to determine file system limits that are appropriate to all of these targets. The file system limits that are enforced are file size, user ID, and group ID. The contents of the `target_list` file are manipulated by using the `fscdsadm` command.

Adding an entry to the list of target operating systems

To add an entry to the list of target operating systems

- Type the following command:

```
# fscdsadm -o add -t target mount_point
```

See [“Specifying the migration target”](#) on page 299.

Removing an entry from the list of target operating systems

To remove an entry from the list of target operating systems

- Type the following command:

```
# fscdsadm -o remove -t target mount_point
```

See [“Specifying the migration target”](#) on page 299.

Removing all entries from the list of target operating systems

To remove all entries from the list of target operating systems

- Type the following command:

```
# fscdsadm -o none mount_point
```

Displaying the list of target operating systems

To display a list of all target operating systems

- Type the following command:

```
# fscdsadm -o list mount_point
```

Enforcing the established CDS limits on a file system

By default, CDS ignores the limits that are implied by the operating system targets that are listed in the `target_list` file.

To enforce the established CDS limits on a file system

- Type the following command:

```
# fscdsadm -l enforce mount_point
```

Ignoring the established CDS limits on a file system

By default, CDS ignores the limits that are implied by the operating system targets that are listed in the `target_list` file.

To ignore the established CDS limits on a file system

- Type the following command:

```
# fscdsadm -l ignore mount_point
```

Validating the operating system targets for a file system

To validate the operating system targets for a file system

- Type the following command:

```
# fscdsadm -v mount_point
```

Displaying the CDS status of a file system

The CDS status that is maintained for a file system includes the following information:

- the `target_list` file
- the limits implied by the `target_list` file
- whether the limits are being enforced or ignored
- whether all files are within the CDS limits for all operating system targets that are listed in the `target_list` file

To display the CDS status of a file system

- Type the following command:

```
# fscdsadm -s mount_point
```

Migrating a file system one time

This example describes a one-time migration of data between two operating systems. Some of the following steps require a backup of the file system to be created. To simplify the process, you can create one backup before performing any of the steps instead of creating multiple backups as you go.

To perform a one-time migration

- 1 If the underlying Volume Manager storage is not contained in a CDS disk group, it must first be upgraded to be a CDS disk group, and all other physical considerations related to migrating the storage physically between systems must first be addressed.

See [“Converting a non-CDS disk group to a CDS disk group”](#) on page 280.

- 2 If the file system is using a disk layout version prior to 7, upgrade the file system to Version 7.

See the *Veritas Storage Foundation Installation Guide*.

- 3 Use the following command to ensure that there are no files in the file system that will be inaccessible after migrating the data due to large file size or to differences in user or group ID between platforms:

```
# fscdsadm -v -t target mount_point
```

If such files exist, move the files to another file system or reduce the size of the files.

- 4 Unmount the file system:

```
# umount mount_point
```

- 5 Use the `fscdsconv` command to convert the file system to the opposite endian.

See [“Converting the byte order of a file system”](#) on page 305.

- 6 Make the physical storage and Volume Manager logical storage accessible on the Linux system by exporting the disk group from the source system and importing the disk group on the target system after resolving any other physical storage attachment issues.

See [“Disk tasks”](#) on page 285.

- 7 Mount the file system on the target system.

Migrating a file system on an ongoing basis

This example describes how to migrate a file system between platforms on an ongoing basis. Some of the following steps require a backup of the file system to

be created. To simplify the process, you can create one backup before performing any of the steps instead of creating multiple backups as you go.

To perform an ongoing migration

- 1 Use the following command to ensure that there are no files in the file system that will be inaccessible after migrating the data due to large file size or to differences in user or group ID between platforms:

```
# fscdsadm -v -t target mount_point
```

If such files exist, move the files to another file system or reduce the size of the files.

- 2 Add the platform on the `target_list` file:

- If migrating a file system between the Solaris and Linux, add `SunOS` and `Linux` to the `target_list` file:

```
# fscdsadm -o add -t os_name=SunOS /mnt1
# fscdsadm -o add -t os_name=Linux /mnt1
```

- If migrating a file system between the HP-UX and Linux, add `HP-UX` and `Linux` to the `target_list` file:

```
# fscdsadm -o add -t os_name=HP-UX /mnt1
# fscdsadm -o add -t os_name=Linux /mnt1
```

- 3 Enforce the limits:

```
# fscdsadm -l enforce mount_point
```

This is the last of the preparation steps. When the file system is to be migrated, it must be unmounted, and then the storage moved and mounted on the target system.

- 4 Unmount the file system:

```
# umount mount_point
```

- 5 Make the file system suitable for use on the specified target.

See [“Converting the byte order of a file system”](#) on page 305.

- 6 Make the physical storage and Volume Manager logical storage accessible on the target system by exporting the disk group from the source system and importing the disk group on the target system after resolving any other physical storage attachment issues.
See “Disk tasks” on page 285.
- 7 Mount the file system on the target system.

Stopping ongoing migration

To stop performing ongoing migration

- ◆ Type the following commands:

```
# fscdsadm -l ignore mount_point  
# fscdsadm -o none mount_point
```

The file system is left on the current system.

When to convert a file system

When moving a file system between two systems, it is essential to run the `fscdsconv` command to perform all of the file system migration tasks. The `fscdsconv` command validates the file system to ensure that it does not exceed any of the established CDS limits on the target, and converts the byte order of the file system if the byte order of the target is opposite to that of the current system.

Warning: Prior to VxFS 4.0 and disk layout Version 6, VxFS did not officially support moving file systems between different platforms, although in many cases a user may have successfully done so. Do not move file systems between platforms when using versions of VxFS prior to Version 4, or when using disk layouts earlier than Version 6. Instead, upgrade to VxFS 4.0 or higher, and disk layout Version 6 or later. Failure to upgrade before performing cross-platform movement can result in data loss or data corruption.

Note: If you replicate data from a little-endian to a big-endian system (or vice versa), you must convert the application after the replication completes.

Converting the byte order of a file system

Use the `fscdsconv` command to migrate a file system from one system to another.

To convert the byte order of a file system

- 1 Determine the disk layout version of the file system that you will migrate:

```
# fstyp -v /dev/vx/rdisk/diskgroup/volume | grep version  
  
magic a501fcf5 version 9 ctime Thu Jun 1 16:16:53 2006
```

Only file systems with disk layout Version 7 or later can be converted. If the file system has an earlier disk layout version, convert the file system to disk layout Version 7 or later before proceeding.

See the `vxfsconvert(1M)` manual page.

See the `vxupgrade(1M)` manual page.

- 2 Perform a full file system back up. Failure to do so could result in data loss or data corruption under some failure scenarios in which restoring from the backup is required.
- 3 Designate a file system with free space where `fscdsconv` may create a file that will contain recovery information for usage in the event of a failed conversion.

Depending on the nature of the file system to be converted, for example if it is mirrored, you may wish to designate the recovery file to reside in a file system with the same level of failure tolerance. Having the same level of failure tolerance reduces the number of failure scenarios that would require restoration from the backup.

- 4 Unmount the file system to be converted:

```
# umount mount_point
```

- 5 Use the `fscdsconv` command to export the file system to the required target:

```
# fscdsconv -f recovery_file -t target_OS -e special_device
```

target_OS specifies the operating system to which you are migrating the file system.

See [“Specifying the migration target”](#) on page 299.

recovery_file is the name of the recovery file to be created by the `fscdsconv` command.

special_device is the raw device or volume that contains the file system to be converted.

Include the file system that you chose in 3 when designating the recovery file.

For example, if the file system chosen to contain the recovery file is mounted on `/data/fs3`, the recovery file could be specified as `/data/fs3/jan04recovery`. If there is not enough disk space on the chosen file system for the recovery file to be created, the conversion aborts and the file system to be converted is left intact.

The recovery file is not only used for recovery purposes after a failure, but is also used to perform the conversion. The directory that will contain the recovery file should not allow non-system administrator users to remove or replace the file, as this could lead to data loss or security breaches. The file should be located in a directory that is not subject to system or local scripts will remove the file after a system reboot, such as that which occurs with the `/tmp` and `/var/tmp` directories on the Solaris operating system.

The recovery file is almost always a sparse file. The disk utilization of this file can best be determined by using the following command:

```
# du -sk filename
```

The recovery file is used only when the byte order of the file system must be converted to suit the specified migration target.

- 6 If you are converting multiple file systems at the same time, which requires the use of one recovery file per file system, record the names of the recovery files and their corresponding file systems being converted in the event that recovery from failures is required at a later time.

- 7 Based on the information provided regarding the migration target, `fscdsconv` constructs and displays the complete migration target and prompts the use to verify all details of the target. If the migration target must be changed, enter `n` to exit `fscdsconv` without modifying the file system. At this point in the process, `fscdsconv` has not used the specified recovery file.
- 8 If the byte order of the file system must be converted to migrate the file system to the specified target, `fscdsconv` prompts you to confirm the migration. Enter `y` to convert the byte order of the file system. If the byte order does not need to be converted, a message displays indicating this fact.
- 9 The `fscdsconv` command indicates if any files are violating the maximum file size, maximum UID, or maximum GID limits on the specified target and prompts you if it should continue. If you must take corrective action to ensure that no files violate the limits on the migration target, enter `n` to exit `fscdsconv`. At this point in the process, `fscdsconv` has not used the specified recovery file.

If the migration converted the byte order of the file system, `fscdsconv` created a recovery file. The recovery file is not removed after the migration completes, and can be used to restore the file system to its original state if required at a later time.

- 10 If a failure occurs during the conversion, the failure could be one of the following cases:

- System failure.
- `fscdsconv` failure due to program defect or abnormal termination resulting from user actions.

In such cases, the file system being converted is no longer in a state in which it can be mounted or accessed by normal means through other VxFS utilities. To recover the file system, invoke the `fscdsconv` command with the recovery flag, `-r`:

```
# fscdsconv -r -f recovery_file special_device
```

When the `-r` flag is specified, `fscdsconv` expects the recovery file to exist and that the file system being converted is the same file system specified in this second invocation of `fscdsconv`.

- 11 After invoking `fscdsconv` with the `-r` flag, the conversion process will restart and complete, given no subsequent failures.

In the event of another failure, repeat 10.

Under some circumstances, you will be required to restore the file system from the backup, such as if the disk fails that contains the recovery file. Failure to have created a backup would then result in total data loss in the file system. I/O errors on the device that holds the file system would also require a backup to be restored after the physical device problems are addressed. There may be other causes of failure that would require the use of the backup.

Importing and mounting a file system from another system

The `fscdsconv` command can be used to import and mount a file system that was previously used on another system.

To import and mount a file system from another system

- ◆ Convert the file system:

```
# fscdsconv -f recovery_file -i special_device
```

If the byte order of the file system needs to be converted

Enter `y` to convert the byte order of the file system when prompted by `fscdsconv`. If the migration converted the byte order of the file system, `fscdsconv` creates a recovery file that persists after the migration completes. If required, you can use this file to restore the file system to its original state at a later time.

If the byte order of the file system does not need to be converted

A message displays that the byte order of the file system does not need to be converted.

Alignment value and block size

On the AIX, Linux and Solaris operating systems, an alignment value of 1 is equivalent to a block size of 512 bytes. On the HP-UX operating system, it is equivalent to a block size of 1024 bytes.

The block size on HP-UX is different from that on other supported platforms. Output from commands such as `vxdisk` and `vxprint` looks different on HP-UX for the same disk group if the `-b` option is not specified.

Migrating a snapshot volume

This example demonstrates how to migrate a snapshot volume containing a VxFS file system from a Solaris SPARC system (big endian) to a Linux system (little endian) or HP-UX system (big endian) to a Linux system (little endian).

To migrate a snapshot volume

- 1 Create the instant snapshot volume, `snapvol`, from an existing plex in the volume, `vol`, in the CDS disk group, `datadg`:

```
# vxsnap -g datadg make source=vol/newvol=snapvol/nmirror=1
```

- 2 Quiesce any applications that are accessing the volume. For example, suspend updates to the volume that contains the database tables. The database may have a hot backup mode that allows you to do this by temporarily suspending writes to its tables.
- 3 Refresh the plexes of the snapshot volume using the following command:

```
# vxsnap -g datadg refresh snapvol source=yes syncing=yes
```

- 4 The applications can now be unquiesced.

If you temporarily suspended updates to the volume by a database in 2, release all the tables from hot backup mode.

- 5 Use the `vxsnap syncwait` command to wait for the synchronization to complete:

```
# vxsnap -g datadg syncwait snapvol
```

- 6 Check the integrity of the file system, and then mount it on a suitable mount point:

```
# fsck -F vxfs /dev/vx/rdisk/datadg/snapvol  
# mount -F vxfs /dev/vx/dsk/datadg/snapvol /mnt
```

- 7 Confirm whether the file system can be converted to the target operating system:

```
# fscdstask validate Linux /mnt
```

8 Unmount the snapshot:

```
# umount /mnt
```

9 Convert the file system to the opposite endian:

```
# fscdsconv -e -f recoveryfile -t target_specifiers special
```

For example:

```
# fscdsconv -e -f /tmp/fs_recov/recov.file -t Linux \  
/dev/vx/dsk/datadg/snapvol
```

This step is only required if the source and target systems have the opposite endian configuration.

10 Split the snapshot volume into a new disk group, `migdg`, and deport that disk group:

```
# vxdg split datadg migdg snapvol  
# vxdg deport migdg
```

11 Import the disk group, `migdg`, on the Linux system:

```
# vxdg import migdg
```

It may be necessary to reboot the Linux system so that it can detect the disks.

12 Use the following commands to recover and restart the snapshot volume:

```
# vxrecover -g migdg -m snapvol  
# vxvol -g migdg start snapvol
```

13 Check the integrity of the file system, and then mount it on a suitable mount point:

```
# fsck -t vxfs /dev/vx/dsk/migdg/snapvol  
# mount -t vxfs /dev/vx/dsk/migdg/snapvol /mnt
```


Index

Symbols

/etc/default/vxassist defaults file 283
/etc/default/vxcdsconvert defaults file 283
/etc/default/vxdg defaults file 283
/etc/default/vxdisk defaults file 284
/etc/default/vxencap defaults file 284
/etc/vx/darecs file 277

A

absolute pathnames
 use with symbolic links 39
access type 271
accessing
 Quick I/O files with symbolic links 39
activation
 default 294
AIX coexistence label 271
alignment 273
 changing 288
allocating file space 33
analyzing I/O statistics 58–59
ARCHIVELOG mode 126
archiving
 using NetBackup 129
asynchronous I/O 28
attribute
 CDS 294
attributes
 init 96, 118
 ndcomirror 96, 118
 nmirror 96, 118
 regionsize 99, 146
auto disk type 271

B

backing up
 using NetBackup 129
Backup
 vgcfgbackup 203

backup
 of cluster file systems 116
 of online databases 93
backups
 creating for volumes 83
benefits of Concurrent I/O 66
benefits of Quick I/O 28
block size 269
blockdev --rereadpt 297

C

cache advisory
 checking setting for 64
cache hit ratio
 calculating 58–59
Cached Quick I/O
 customizing 61
 determining files to use 57
 disabling individual files 61
 enabling individual files 62
 making settings persistent 62
 prerequisite for enabling 53
calculating cache hit ratio 58–59
CDS
 attribute 294
 changing setting 289
 creating DGs 278
 creating disks 277
 disk group alignment 270
 disk group device quotas 272
 disks 270
CDS disk groups
 alignment 295
 joining 289
 moving 289
 setting alignment 288
CDS disks
 creating 276
changing CDS setting 289
changing default CDS setting 289
changing disk format 285

- changing file sizes 32
- chgrp command 37
- chmod command
 - commands
 - chmod 53
- chown command 37
 - commands
 - chown 53
- cluster file systems
 - off-host backup of 116
- co-existence label 271
- collecting I/O statistics 57
- commands
 - chgrp 37
 - chown 37
 - fsadm command 43
 - grep 55
 - ls 41
 - mount 31
 - qioadmin 61
 - qiomkfile 43, 45
 - qiostat 57
 - setext 37
 - vxtunefs 63
- concepts 267
- Concurrent I/O
 - benefits 66
 - disabling 68, 70
 - enabling 66, 69
- configuration
 - LVM 196
- configuration VxVM 196
- conversion
 - non-interactive 230
 - speed 229
 - vxxmconvert 198
- converting non-CDS disks to CDS 278
- converting non-CDS disks to CDS disks 279
- CREADs 60
- creating
 - Quick I/O files 34–35
 - symbolic links to access Quick I/O files 32
- creating a DRL log 291
- creating CDS disk groups 278
- creating CDS disks 276–277
- creating DRL logs 291
- creating non-CDS disk groups 289
- creating pre-version 110 disk groups 289

- cross-platform data sharing
 - recovery file 306
- current-rev disk groups 273
- customizing Cached Quick I/O 61

D

- data on Secondary
 - using for off-host processing 162
- database performance
 - using Quick I/O 28
- databases
 - incomplete media recovery 127
 - integrity of data in 84
 - online backup of 93
 - rolling back 127
 - using Storage Checkpoints 125
- dataserver buffer cache 51
- DB2 BufferPool 50
- decision support
 - using point-in-time copy solutions 140
- default activation 294
- default CDS setting
 - changing 289
- defaults files 279, 282
- determining
 - if Quick I/O installed and enabled 42
- device quotas 272, 294
 - displaying 294
 - setting 290
- direct I/O 28
- direct-write
 - copy-behind 52
- disabling Cached Quick I/O for a file 61
- disabling Concurrent I/O 68, 70
- disabling qio_cache_enable flag 54
- disabling Quick I/O 48
- disk
 - access type 271
 - change format 285
 - labels 285
 - LVM 285
 - replacing 290
- disk access 269
- disk format 270
- disk group alignment 288
 - displaying 295
- disk groups 271
 - alignment 273
 - creating 289

- disk groups (*continued*)
 - joining 289
 - non-CDS 273
 - upgrading 290
- disk headers 197
- disk quotas
 - setting 290
- disk types 270
- disks
 - effects of formatting or partitioning 296
- displaying device quotas 294
- displaying disk group alignment 295
- displaying DRL log size 294
- displaying DRL map size 294
- displaying log map values 295
- displaying log size 294
- displaying v_logmap values 295
- displaying volume log map values 295
- double buffering 28, 50–51
- DRL log size
 - displaying 294
 - setting 291
- DRL logs
 - creating 291
- DRL map length 292
- DRL map size
 - displaying 294
 - setting 291
- DSS. *See* Decision Support

E

- enabling
 - Quick I/O 31
- enabling Cached Quick I/O for a file 62
- enabling Concurrent I/O 66, 69
- enabling qio_cache_enable flag 54
- encapsulation 285
- Example
 - analyze LVM groups 214
 - conversion 214
 - failed conversion 214
 - list 214
 - list disk information 214
 - list LVM volume group information 214
 - listvg 214
 - LVM to VxVM 214
 - vxprint output 214
- example
 - Failed Analysis 214

- extending a file 32
- extending Quick I/O files 43

F

- FastResync
 - Persistent 83
- file
 - space allocation 33
- File System 204
- file system locking 28
- file systems
 - growing to accommodate Quick I/O files 43
 - mounting for shared access 118
- FileSnaps
 - about 85
 - data mining, reporting, and testing 168
 - virtual desktops 167
 - write intensive applications 168
 - best practices 167
- FlashSnap 82
- fsadm command 43
- fscdsadm 300
- fscdsconv 305

G

- grep command 55
- growing
 - file systems 43
 - Quick I/O files 43

I

- I/O
 - asynchronous 28
 - direct 28
 - kernel asynchronous 28
- I/O block size 269
- ID block 271
- improving
 - database performance 28
- init attribute 96, 118
- instant snapshots
 - reattaching 125, 162
- intent logging 84

J

- joining CDS disk groups 289
- joining disk groups 289

K

kernel asynchronous I/O 28
kernel write locks 28

L

length listing 296
licensing 282
list LVM 214
listing disk groups 296
listing disks 296
listing offset and length information 289
log size
 displaying 294
 setting 291
ls command 41
LVM
 metadata 199
LVM disks 285
LVM names
 symbolic names 206
LVM VGRA 197

M

mapping
 LVM device nodes 205
 VxVM device nodes 205
minor device numbers 273
mount command 31
mounting
 shared-access file systems 118
moving CDS disk groups 289
moving disk group objects 289

N

ndcomirror attribute 96, 118
NetBackup
 overview 129
nmirror attribute 96, 118

O

objects
 moving 289
off-host backup of cluster file systems
 using point-in-time copy solutions 116
offset
 listing 296
offset information 296

online database backup
 using point-in-time copy solutions 93
online migration
 About LVM to VxVM
 VCS HA environment 246
operating system data 269

P

persistence
 for Cached Quick I/O settings 62
Persistent FastResync 83
platform block 271
point-in-time copy solutions
 applications 81
 for decision support 140
 for off-host cluster file system backup 116
 for online database backup 93
PREADs 60
preallocating space for Quick I/O files 37
private region 270
public region 270

Q

qio_cache_enable flag
 disabling 54
 enabling 54
qioadmin command 61
qiomkfile command 43, 45
 options for creating files
 symbolic links 32
qiostat
 output of 58–59
qiostat command 57
Quick I/O
 accessing regular VxFS files as 39
 benefits 28
 determining status 42
 disabling 48
 enabling 31
 extending files 43
 improving database performance with 28
 performance improvements 52
 preallocating space for files 37
 showing resolution to a raw device 43
 using relative and absolute pathnames 39

R

- read-ahead algorithm
 - for Cached Quick I/O 52
- recovery
 - using Storage Checkpoints 125
- recovery file, cross-platform data sharing 306
- regionsize attribute 99, 146
- relative pathnames
 - use with symbolic links 39
- replacing disks 290
- resetlogs option 128
- resizing a file 32
- restoring
 - using NetBackup 129
- restoring CDS disk labels 285
- restoring disk labels 285
- root disk
 - configuring for VxVM 199
 - reconfiguring for LVM 199
- rootability
 - configuring a VxVM root disk 199
 - restoring a LVM root disk 199
 - root disk 199
 - root volume 199

S

- SAM
 - vgdisplay 201
- Secondary
 - using data 162
- setext command 37
- setting CDS disk group alignment 288
- setting device quotas 290
- setting disk quotas 290
- setting DRL log size 291
- setting DRL map length 292
- setting DRL map size 291
- setting log size 291
- settings
 - making Cached Quick I/O persistent 54
- shared access
 - mounting file systems for 118
- showing
 - Quick I/O file resolved to raw device 43
- snapshots
 - reattaching instant 125, 162
- sparse files 41
- Storage Checkpoints 84
 - creating 126

Storage Checkpoints (*continued*)

- database recovery 125
- Storage Rollback
 - implementing using Storage Checkpoints 125
 - using VxDBA 127
- symbolic links
 - advantages and disadvantages 39
 - to access Quick I/O files 39
- system buffer cache 52

T

- tool
 - vxvmconvert 198
- tools
 - vxdiskadm 196
 - vxvmconvert 196
- tunefstab file
 - adding tuning parameters to 54
- tuning parameters
 - adding to tunefstab file 54

U

- upgrading disk groups 290
- upgrading pre-version 110 disk groups 290
- utilities. *See* commands

V

- v_logmap
 - displaying 295
- verifying caching using vxfstune parameters 55
- verifying vxtunefs system parameters 55
- volumes
 - backing up 83
- vradmin utility
 - ibc
 - using off-host processing 162
- vxcdsconvert 279
- vxctl enable 297
- vx dg init 278
- vx dg split 288
- vx disk scandisks 297
- vx disk adm 276, 278
- vx disk setup 276
- vx snap
 - reattaching instant snapshots 125, 162
- vxtunefs command 63
 - commands
 - vxtunefs 55

VxVM

 devices 269

 metadata 199

VxVM names

 symbolic link 206

vxvol 292