

Veritas Storage Foundation™ for Sybase Administrator's Guide

Linux on IBM Power

5.0 Release Update 3



Veritas Storage Foundation™ for Sybase Administrator's 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.

Documentation version 5.0RU3

Legal Notice

Copyright © 2009 Symantec Corporation.

All rights reserved.

Federal acquisitions: Commercial Software - Government Users Subject to Standard License Terms and Conditions.

Symantec, the Symantec Logo, Veritas, and Veritas Storage Foundation 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.

Third-party software may be recommended, distributed, embedded, or bundled with this Symantec product. Such third-party software is licensed separately by its copyright holder. All third-party copyrights associated with this product are listed in the accompanying release notes.

Linux is a registered trademark of Linus Torvalds.

Sybase is a registered trademark of Sybase, Inc.

Windows is a registered trademark of Microsoft Corporation.

Veritas Storage Foundation™ is a licensed product. See the Veritas Storage Foundation™ Installation Guide for license installation instructions.

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" and "commercial computer software documentation" as defined in FAR Sections 12.212 and DFARS Section 227.7202.

Symantec Corporation 20330 Stevens Creek Blvd. Cupertino, CA 95014 USA

<http://www.symantec.com>

Technical Support

For technical assistance, visit <http://support.veritas.com> and select phone or email support. Use the Knowledge Base search feature to access resources such as TechNotes, product alerts, software downloads, hardware compatibility lists, and our customer email notification service.

Contents

Chapter 1	Introducing Veritas Storage Foundation for Sybase	
	About Veritas Storage Foundation for Sybase	9
	Components of Veritas Storage Foundation for Sybase	10
	How Veritas Volume Manager works	10
	About volumes	12
	About disk groups	13
	About volume layouts	13
	About online relayout	16
	About volume resynchronization	16
	About dirty region logging	17
	About volume sets	17
	About volume snapshots	17
	About Veritas FastResync	17
	About disk group split and join	18
	About hot-relocation	19
	About DMP-supported disk arrays	19
	About dynamic LUN expansion	20
	About Storage Expert	20
	About cluster functionality (optional)	20
	About Veritas Volume Replicator (optional)	20
	How Veritas File System works	21
	About Veritas Concurrent I/O	21
	About extent-based allocation	21
	About fast file system and database recovery	22
	About online system administration	22
	About cross-platform data sharing	23
	Support for multi-volume file systems	23
	About Quality of Storage Service (optional)	23
	Support for large file systems and large files (optional)	23
	About restoring file systems using Storage Checkpoints	24
	About quotas	24
	About cluster functionality (optional)	24

Chapter 2	Setting up dataservers	
	Tasks for setting up new databases	25
	About setting up a disk group	26
	Disk group configuration guidelines	27
	Creating a disk group	27
	Adding disks to a disk group	28
	About selecting a volume layout	29
	How to choose appropriate stripe unit sizes	30
	How to choose between mirroring and RAID-5	30
	Volume configuration guidelines	30
	Creating a volume	31
	Creating a volume set	32
	Adding a volume to a volume set	33
	File system creation guidelines	33
	Creating a VxFS file system	34
	Multi-volume support	35
	Mounting a file system	36
	Unmounting a file system	37
	About fragmentation	38
	How to control fragmentation	38
	Types of fragmentation	38
	How to monitor fragmentation	39
	Defragmenting a file system	39
	Resizing a file system	41
	Resizing a file system and the underlying volume	42
Chapter 3	Using Veritas Concurrent I/O	
	About Concurrent I/O	45
	How Concurrent I/O works	45
	Enabling and disabling Concurrent I/O	46
	Enabling Concurrent I/O	46
	Disabling Concurrent I/O	46
Chapter 4	Converting existing database configurations to VxFS	
	Upgrading from earlier VxFS version layouts	49
	About snapshot volumes	54
	About snapshot volumes	54

Chapter 5	Using volume snapshots for dataserer backup and off-host processing	
	About snapshot volumes	54
	Backup and off-host processing applications	55
	FastResync of snapshot volumes	55
	Disk group split and join	56
	Preparing hosts for database backup or off-host processing	59
	Single-host configuration	59
	Two-host configuration	60
	Upgrading existing volumes to use VxVM 4.0 features	61
	Sybase Adaptive Server Enterprise 12.5 quiesce feature	63
	How to set up volume snapshots with Sybase ASE 12.5 server	64
	Implementing online backup or off-host processing	65
	Creating a warm standby server	70
	Resynchronizing the snapshot to your ASE dataserer	82
	Recovering the database from a backup image	87
	Refreshing a snapshot database image	91
	Dissociating a snapshot volume	93
	Removing a snapshot volume	93
Chapter 6	Tuning for performance	
	Additional documentation	95
	About tuning VxFS	95
	How monitoring free space works	96
	How tuning VxFS I/O parameters works	97
	About tunable VxFS I/O parameters	97
	About tuning VxVM	101
	About obtaining volume I/O statistics	101
	About tuning Sybase dataservers	103
	Sybase tempdb database	103
	Sybase sybsecurity database	104
	Placement of the transaction logs	104
	Database device layout	104
	Nonclustered indexes placement	104
Glossary		
Index		

Introducing Veritas Storage Foundation for Sybase

This chapter includes the following topics:

- [About Veritas Storage Foundation for Sybase](#)
- [How Veritas Volume Manager works](#)
- [How Veritas File System works](#)

About Veritas Storage Foundation for Sybase

There are two versions of this product:

- Veritas Storage Foundation for Sybase Standard Edition
- Veritas Storage Foundation for Sybase Enterprise Edition
The Enterprise Edition contains everything in the Standard Edition plus FastResync, disk group split and join, Quality of Storage Service, and support for large file systems (up to 8 exabytes).

Note: Veritas Storage Foundation/High Availability (HA) for Sybase is available only with the Enterprise Edition.

Unless otherwise noted, features pertain to both the Standard and Enterprise Edition products.

Components of Veritas Storage Foundation for Sybase

Veritas Storage Foundation for Sybase combines the strengths of the core technology products with database-specific enhancements to offer performance, availability, and manageability for Sybase database servers.

Veritas Storage Foundation for Sybase includes the following products:

- **Veritas Volume Manager (VxVM)**

A disk management subsystem that supports disk striping, disk mirroring, and simplified disk management for improved data availability and performance.

- **Veritas File System (VxFS)**

A high-performance, fast-recovery file system that is optimized for business-critical database applications and data-intensive workloads. VxFS offers online administration, letting you perform most frequently scheduled maintenance tasks (including online backup, resizing, and file system changes) without interrupting data or system availability. VxFS also provides support for large file systems (of more than 8 exabytes in a 64-bit environment) and large files (in the exabyte range in a 64-bit environment).

Veritas File System offers the following performance-enhancing features that are of particular interest in a database environment:

- **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 relational databases (such as Sybase), which do not have a sequential read/write requirement, to access files concurrently.

An optional High Availability (HA) version of Veritas Storage Foundation for Sybase Enterprise Edition, which includes Veritas Cluster Server, is available for customers who have high system-availability requirements.

How Veritas Volume Manager works

Databases require their storage media to be robust and resilient to failure. It is vital to protect against hardware and disk failures and to maximize performance using all the available hardware resources. Using a volume manager provides this necessary resilience and eases the task of management. A volume manager can help you manage hundreds of disk devices and makes spanning, striping, and mirroring easy.

Veritas Volume Manager (VxVM) builds virtual devices called volumes on top of physical disks. Volumes are accessed by a file system, a database, or other

applications in the same way physical disk partitions would be accessed. Using volumes, VxVM provides the following administrative benefits for databases:

Table 1-1 Veritas Volume Manager features

Feature	Benefit
Spanning of multiple disks	Eliminates media size limitations.
Striping	Increases throughput and bandwidth.
Mirroring or RAID-5	Increases data availability.
Online relayout	Allows volume layout changes without application or database downtime. Online relayout can be used to change performance or reliability characteristics of unerlying storage.
Volume resynchronization	Ensures that all mirrors contain exactly the same data and that the data and parity in RAID-5 volumes agree.
Dirty Region Logging (DRL)	Speeds the recovery of mirrored volumes after a system crash.
Volume snapshots	Allows backup of volumes based on disk mirroring. VxVM provides full-sized and space-optimized instant snapshots, which are online and off-host point-in-time copy solutions.
FastResync	Separately licensed, optional feature that performs quick and efficient resynchronization of stale mirrors. FastResync is included with the Enterprise Edition and is also included as part of the Veritas FlashSnap option with the Standard Edition.
Disk group split and join	Separately licensed, optional feature that supports general disk group reorganization and allows you to move volume snapshots to another host for off-host backup. Disk group split and join is included with the Enterprise Edition and is also included as part of the Veritas FlashSnap option with the Standard Edition.
Hot-relocation	Automatically restores data redundancy in mirrored and RAID-5 volumes when a disk fails.
Dynamic multipathing (DMP)	Allows for transparent failover, load sharing, and hot plugging of physical disks.

Table 1-1 Veritas Volume Manager features (*continued*)

Feature	Benefit
Volume sets	Allows several volumes to be represented by a single logical mount device.
Dynamic LUN Expansion	Allows you to resize a disk after it has been initialized while preserving the existing data on the disk.
Storage Expert	Helps diagnose configuration problems with VxVM.
Cluster Volume Manager (CVM)	Separately licensed, optional feature that allows you to use VxVM in a cluster environment.
Veritas Volume Replicator (VVR)	Separately licensed, optional feature that provides data replication for disaster recovery solutions.
Free space pool management	Simplifies administration and provides flexible use of available hardware.
Online administration	Allows configuration changes without system or database down time.

For a more detailed description of VxVM and its features, refer to the *Veritas Volume Manager Administrator's Guide*.

About volumes

A volume is a virtual disk device that appears to applications, databases, and file systems like a physical disk partition without the physical limitations of a disk partition. A volume consists of one or more plexes, each holding a copy of the selected data in the volume. Due to its virtual nature, a volume is not restricted to a particular disk or a specific area of a disk. For example, a volume can span multiple disks and can be used to create a large file system.

Volumes consist of other virtual objects that can be manipulated to change the volume's configuration. Volumes and their virtual components are referred to as Volume Manager objects. You can manipulate Veritas Volume Manager objects in a variety of ways to optimize performance, provide redundancy of data, and perform backups or other administrative tasks on one or more physical disks without interrupting applications. As a result, data availability and disk subsystem throughput are improved.

You can change the configuration of a volume without causing disruption to databases or file systems that are using the volume. For example, you can mirror a volume on separate disks or move the volume to use different disk storage.

About disk groups

A disk group is a collection of disks that share a common configuration (for example, configuration objects that belong to a single database). We recommend creating one disk group for each database.

You can move a disk group and its components as a unit from one host to another host. For example, you can move volumes and file systems that belong to the same database and are created within one disk group as a unit. You must configure a given volume from disks belonging to one disk group.

About volume layouts

A Redundant Array of Independent Disks (RAID) is a disk array in which a group of disks appears to the system as a single virtual disk or a single volume. VxVM supports several RAID implementations, as well as spanning.

The following volume layouts are available to satisfy different database configuration requirements:

- Spanning and concatenation
- Striping (RAID-0)
- Mirroring (RAID-1)
- Mirrored-Stripe Volumes (RAID-0+1)
- Striped-Mirror Volumes (RAID-1+0)
- RAID-5

Caution: Spanning or striping a volume across multiple disks increases the chance that a disk failure will result in failure of that volume. Use mirroring or RAID-5 to substantially reduce the chance of a single volume failure caused by a single disk failure.

How spanning and concatenation work

Concatenation maps data in a linear manner onto one or more subdisks in a plex. To access all of the data in a concatenated plex sequentially, data is first accessed in the first subdisk from beginning to end. Data is then accessed in the remaining subdisks sequentially from beginning to end, until the end of the last subdisk.

You can use concatenation with multiple subdisks when there is insufficient contiguous space for the plex on any one disk. This form of concatenation can be

used for load balancing between disks, and for head movement optimization on a particular disk.

Concatenation using subdisks that reside on more than one VxVM disk is called spanning.

Warning: Spanning a plex across multiple disks increases the chance that a disk failure results in failure of the assigned volume. Use mirroring (RAID-1) or striping with parity (RAID-5) to reduce the risk that a single disk failure results in a volume failure.

Spanning is useful when you need to read or write data sequentially (for example, reading from or writing to database redo logs) and there is not sufficient contiguous space.

How striping (RAID-0) works

Striping is a technique of mapping data so that the data is interleaved among multiple physical disks. Data is allocated in equal-sized units (called stripe units) that are interleaved between the disks. Each stripe unit is a set of contiguous blocks on a disk. A stripe consists of the set of stripe units at the same position across all columns. A column is a set of one or more subdisks within a striped plex.

Striping is useful if you need large amounts of data written to or read from physical disks, and performance is important. Striping is also helpful in balancing the I/O load from multi-user applications across multiple disks. By using parallel data transfer to and from multiple disks, striping significantly improves data-access performance.

When striping across multiple disks, failure of any one disk will make the entire volume unusable.

How mirroring (RAID-1) works

Mirroring is a technique of using multiple copies of the data, or mirrors, to duplicate the information contained in a volume. In the event of a physical disk failure, the mirror on the failed disk becomes unavailable, but the system continues to operate using the unaffected mirrors. For this reason, mirroring increases system reliability and availability. A volume requires at least two mirrors to provide redundancy of data. A volume can consist of up to 32 mirrors. Each of these mirrors must contain disk space from different disks for the redundancy to be effective.

How striping plus mirroring (mirrored-stripe or RAID-0+1) works

VxVM supports the combination of mirroring above striping. The combined layout is called a mirrored-stripe layout. A mirrored-stripe layout offers the dual benefits of striping to spread data across multiple disks, while mirroring provides redundancy of data. For mirroring above striping to be effective, the striped plex and its mirrors must be allocated from separate disks.

The layout type of the data plexes in a mirror can be concatenated or striped. Even if only one is striped, the volume is still termed a mirrored-stripe volume. If they are all concatenated, the volume is termed a mirrored-concatenated volume.

How mirroring plus striping (striped-mirror, RAID-1+0 or RAID-10) works

VxVM supports the combination of striping above mirroring. This combined layout is called a striped-mirror layout and mirrors each column of the stripe. If there are multiple subdisks per column, each subdisk can be mirrored individually instead of each column. A striped-mirror volume is an example of a layered volume.

Compared to a mirrored-stripe volume, a striped-mirror volume offers the dual benefits of striping to spread data across multiple disks, while mirroring provides redundancy of data. A striped-mirror volume enhances redundancy, which makes it more tolerant of disk failure, and reduces recovery time after disk failure.

For databases that support online transaction processing (OLTP) workloads, we recommend either mirrored-stripe or striped-mirror volumes to improve database performance and reliability. For highest availability, we recommend striped-mirror volumes (RAID 1+0).

How striping with parity (RAID-5) works

RAID-5 provides data redundancy through the use of parity (a calculated value that the system uses to reconstruct data after a failure). While data is written to a RAID-5 volume, parity is also calculated by performing an exclusive OR (XOR) procedure on data. The resulting parity is then written to another part of the volume. If a portion of a RAID-5 volume fails, the data that was on that portion of the failed volume can be recreated from the remaining data and the parity.

RAID-5 offers data redundancy similar to mirroring, while requiring less disk space. RAID-5 read performance is similar to that of striping but with relatively slow write performance. RAID-5 is useful if the database workload is read-intensive (as in many data warehousing applications). You can snapshot a RAID-5 volume and move a RAID-5 subdisk without losing redundancy.

About online relayout

As databases grow and usage patterns change, online relayout lets you change volumes to a different layout, with uninterrupted data access. Relayout is accomplished online and in place. Use online relayout to change the redundancy or performance characteristics of the storage, such as data organization (RAID levels), the number of columns for RAID-5 and striped volumes, and stripe unit size.

About volume resynchronization

When storing data redundantly, using mirrored or RAID-5 volumes, Veritas Volume Manager ensures that all copies of the data match exactly. However, if the system crashes, small amounts of the redundant data on a volume can become inconsistent or unsynchronized. For mirrored volumes, unsynchronized data can cause two reads from the same region of the volume to return different results if different mirrors are used to satisfy the read request. In the case of RAID-5 volumes, unsynchronized data can lead to parity corruption and incorrect data reconstruction.

In the event of a system crash, Veritas Volume Manager ensures that all mirrors contain exactly the same data and that the data and parity in RAID-5 volumes agree. This process is called volume resynchronization. Not all volumes require resynchronization after a system failure. VxVM notices when a volume is first written and marks it as dirty. Only volumes that are marked dirty when the system reboots require resynchronization.

The process of resynchronization can impact system and database performance. However, it does not affect the availability of the database after system reboot. You can immediately access the database after database recovery although the performance may suffer due to resynchronization. For very large volumes or for a very large number of volumes, the resynchronization process can take a long time. You can significantly reduce resynchronization time by using Dirty Region Logging (DRL) for mirrored volumes or by making sure that RAID-5 volumes have valid RAID-5 logs. However, using logs can slightly reduce the database write performance.

For most database configurations, we recommend using dirty region logs or the RAID-5 logs when mirrored or RAID-5 volumes are used. It is also advisable to evaluate the database performance requirements to determine the optimal volume configurations for the databases.

About dirty region logging

Dirty region logging (DRL), if enabled, speeds the recovery of mirrored volumes after a system crash. DRL keeps track of the regions that have changed due to I/O writes to a mirrored volume. DRL uses this information to recover only those portions of the volume that need to be recovered.

Note: If a version 20 data change object (DCO) volume is associated with a volume, a portion of the DCO volume can be used to store the DRL log. There is no need to create a separate DRL log for a volume that has a version 20 DCO volume.

For more information on DCOs and DCO volumes, see the *Veritas Volume Manager Administrator's Guide*.

About volume sets

Volume sets are an enhancement to VxVM that allow several volumes to be represented by a single logical mount device. 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-device 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 snapshots

A volume snapshot is a point-in-time image of a volume. Veritas Volume Manager provides three volume snapshot features based on disk mirroring:

- Full-sized instant snapshots
- Space-optimized instant snapshots
- Emulation of third-mirror snapshots

About Veritas FastResync

Veritas FastResync (previously called Fast Mirror Resynchronization or FMR) is included with the Enterprise Edition. It is also included as part of the Veritas FlashSnap option with the Standard Edition.

Veritas FastResync performs quick and efficient resynchronization of stale mirrors (mirrors that are not synchronized). This increases the efficiency of the VxVM snapshot mechanism, and improves the performance of operations such as backup

and decision support. Typically, these operations require that the volume is quiescent, and that they are not impeded by updates to the volume by other activities on the system. To achieve these goals, the snapshot mechanism in VxVM creates an exact copy of a primary volume at an instant in time. After a snapshot is taken, it can be accessed independently of the volume from which it was taken.

How non-persistent FastResync works

Non-persistent FastResync allocates its change maps in memory. If non-persistent FastResync is enabled, a separate FastResync map is kept for the original volume and for each snapshot volume. Unlike a dirty region log (DRL), these maps do not reside on disk nor in persistent store. The advantage is that updates to the FastResync map have little impact on I/O performance, as no disk updates need to be performed. However, if a system is rebooted, the information in the map is lost, so a full resynchronization is required when performing a snapback operation. This limitation can be overcome for volumes in cluster-shareable disk groups, provided that at least one of the nodes in the cluster remains running to preserve the FastResync map in its memory.

How persistent FastResync works

Non-persistent FastResync has been augmented by the introduction of persistent FastResync. Unlike non-persistent FastResync, Persistent FastResync keeps the FastResync maps on disk so that they can survive system reboots and system crashes. When the disk groups are rejoined, this allows the snapshot plexes to be quickly resynchronized. This ability is not supported by non-persistent FastResync.

If persistent FastResync is enabled on a volume or on a snapshot volume, a DCO and a DCO log volume are associated with the volume.

The DCO object is used not only to manage FastResync maps, but also to manage DRL recovery maps and special maps called copy maps that allow instant snapshot operations to be resume following a system crash.

Persistent FastResync can also track the association between volumes and their snapshot volumes after they are moved into different disk groups. When the disk groups are rejoined, this allows the snapshot plexes to be quickly resynchronized. This ability is not supported by non-persistent FastResync.

About disk group split and join

Disk group split and join is included with the Enterprise Edition. It is also included as part of the Veritas FlashSnap option with the Standard Edition.

VxVM provides a disk group content reorganization feature that supports general disk group reorganization and allows you to move volume snapshots to another

host for off-host backup. Additional options to the `vxchg` command enable you to take advantage of the ability to remove all VxVM objects from an imported disk group and move them to a newly created target disk group (`split`), and to remove all VxVM objects from an imported disk group and move them to an imported target disk group (`join`). The move operation enables you to move a self-contained set of VxVM objects between the imported disk groups.

About hot-relocation

In addition to providing volume layouts that help improve database performance and availability, VxVM offers features that you can use to further improve system availability in the event of a disk failure. Hot-relocation is a feature that allows a system to react automatically to I/O failures on mirrored or RAID-5 volumes and restore redundancy and access to those volumes.

VxVM detects I/O failures on volumes and relocates the affected portions to disks designated as spare disks or free space within the disk group. VxVM then reconstructs the volumes that existed before the failure and makes them redundant and accessible again.

The hot-relocation feature is enabled by default and is recommended for most database configurations. After hot-relocation occurs, we recommend verifying the volume configuration for any possible performance impact. It is also a good idea to designate additional disks as spares to augment the spare pool.

While a disk is designated as a spare, you cannot use the space on that disk for the creation of VxVM objects within its disk group. VxVM also lets you free a spare disk for general use by removing it from the pool of hot-relocation disks.

About DMP-supported disk arrays

VxVM provides administrative utilities and driver support for disk arrays that can take advantage of its Dynamic Multipathing (DMP) feature. Some disk arrays provide multiple ports to access their disk devices. These ports, coupled with the host bus adaptor (HBA) controller and any data bus or I/O processor local to the array, make up multiple hardware paths to access the disk devices. Such disk arrays are called multipathed disk arrays. This type of disk array can be connected to host systems in many different configurations, (such as multiple ports connected to different controllers on a single host, chaining of the ports through a single controller on a host, or ports connected to different hosts simultaneously). DMP is available for multiported disk arrays from various vendors and provides improved reliability and performance by using path failover and load balancing.

See the *Veritas Volume Manager Administrator's Guide*.

See the *Veritas Volume Manager Hardware Notes*.

About dynamic LUN expansion

Dynamic LUN expansion allows you to resize a disk after it has been initialized while preserving the existing data on the disk.

See the *Veritas Volume Manager Administrator's Guide*.

About Storage Expert

Storage Expert consists of a set of simple commands that collect VxVM configuration data and compare it with “best practice.” Storage Expert then produces a summary report that shows which objects do not meet these criteria and makes recommendations for VxVM configuration improvements.

These user-configurable tools help you as an administrator to verify and validate systems and non-optimal configurations in both small and large VxVM installations.

Storage Expert components include a set of rule scripts and a rules engine. The rules engine runs the scripts and produces ASCII output, which is organized and archived by Storage Expert's report generator. This output contains information about areas of VxVM configuration that do not meet the set criteria. By default, output is sent to the screen, but you can redirect it to a file using standard UNIX redirection.

See the *Veritas Volume Manager Administrator's Guide*.

About cluster functionality (optional)

VxVM includes an optional, separately licensable clustering feature, known as Cluster Volume Manager, that enables VxVM to be used in a cluster environment. With the clustering option, VxVM supports up to 16 nodes per cluster.

See the *Veritas Volume Manager Administrator's Guide*.

About Veritas Volume Replicator (optional)

Veritas Volume Replicator (VVR) is an optional, separately licensable feature of VxVM. VVR is a data replication tool designed to maintain a consistent copy of application data at a remote site. It is built to contribute to an effective disaster recovery plan. If the data center is destroyed, the application data is immediately available at the remote site, and the application can be restarted at the remote site.

VVR works as a fully integrated component of VxVM. VVR benefits from the robustness, ease of use, and high performance of VxVM and, at the same time, adds replication capability to VxVM. VVR can use existing VxVM configurations

with some restrictions. Any application, even with existing data, can be configured to use VVR transparently.

See the Veritas Volume Replicator documentation.

How Veritas File System works

Veritas File System (referred to as VxFS) is an extent-based, intent logging file system intended for use in UNIX environments that deal with large volumes of data and that require high file system performance, availability, and manageability. VxFS also provides enhancements that make file systems more viable in database environments.

The following sections provide a brief overview of VxFS concepts and features that are relevant to database administration.

See the *Veritas File System Administrator's Guide*.

About Veritas 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.

Veritas Concurrent I/O allows for concurrency between a single writer and multiple readers or between multiple writers. It minimizes serialization for extending writes and sends I/O requests directly to file systems.

See [“About Concurrent I/O”](#) on page 45.

About extent-based allocation

By allocating disk space to files in extents, disk I/O to and from a file can be done in units of multiple blocks. This type of I/O can occur if storage is allocated in units of consecutive blocks. For sequential I/O, multiple block operations are considerably faster than block-at-a-time operations. Almost all disk drives accept I/O operations of multiple blocks.

The VxFS file system allocates disk space to files in groups of one or more extents. VxFS also allows applications to control some aspects of the extent allocation for a given file. Extent attributes are the extent allocation policies associated with a file.

About fast file system and database recovery

Veritas File System begins recovery procedures within seconds after a system failure by using a tracking feature called intent logging. This feature records pending changes to the file system structure in a circular intent log. The intent log recovery feature is not readily apparent to users or a system administrator except during a system failure. During system failure recovery, the VxFS `fsck` utility performs an intent log replay, which scans the intent log and nullifies or completes file system operations that were active when the system failed. The file system can then be mounted without completing a full structural check of the entire file system. Replaying the intent log may not completely recover the damaged file system structure if there was a disk hardware failure; hardware problems may require a complete system check using the `fsck` utility provided with Veritas File System.

About online system administration

The VxFS file system provides online system administration utilities to help resolve certain problems that impact database performance. You can defragment and resize a VxFS file system while it remains online and accessible to users.

How the defragmentation utility works

Free resources are originally aligned in the most efficient order possible and are allocated to files in a way that is considered by the system to provide optimal performance. When a file system is active for extended periods of time, new files are created, old files are removed, and existing files grow and shrink. Over time, the original ordering of free resources is lost and the file system tends to spread across the disk, leaving unused gaps or fragments between areas that are in use. This process, known as fragmentation, leads to degraded performance because the file system has fewer choices when selecting an extent (a group of contiguous data blocks) to assign to a file. You should analyze the degree of fragmentation before creating new database files.

VxFS provides the online administration utility `fsadm` to resolve fragmentation problems. The utility can be run on demand and should be scheduled regularly as a `cron` job.

How the resizing utility works

Changes in database size can result in file systems that are too large or too small for the current database. Without special utilities, expanding or shrinking a file system becomes a matter of stopping applications, offloading the contents of

the file system, rebuilding the file system to a new size, and then restoring the data. Data is unavailable to users while these administrative tasks are performed.

The VxFS file system utility `fsadm` provides a mechanism to resize file systems without unmounting them or interrupting users' productivity. Because the VxFS file system can only be mounted on one device, expanding a file system means that the underlying device must also be expandable while the file system is mounted. Working with VxVM, VxFS provides online expansion capability.

About cross-platform data sharing

Veritas Cross-Platform Data Sharing allows data to be serially shared among heterogeneous systems where each system has direct access to the physical devices that hold the data. This feature can be used only in conjunction with Veritas Volume Manager. Shared or parallel access is possible for read-only data.

See the *Veritas Storage Foundation Cross-Platform Data Sharing Administrator's Guide*.

Support for multi-volume file systems

The multi-volume file system (MVS) feature allows several volumes to be represented by a single logical object. All I/O to and from an underlying logical volume is directed by way of volume sets. A volume set is a container for multiple different volumes. This feature can be used only in conjunction with Veritas Volume Manager.

About Quality of Storage Service (optional)

The Quality of Storage Service (QoS) feature is included with the Enterprise Edition.

The QoS option is built on the multi-volume support technology introduced in this release. Using QoS, you can map more than one device to a single file system. You can then configure policies that automatically relocate files from one device to another, or relocate files by running file relocation commands. Having multiple devices lets you determine where files are located, which can improve performance for applications that access specific types of files and reduce storage-related costs.

Database Dynamic Storage Tiering is built on top of QoS and automates the migration process for Sybase database objects.

Support for large file systems and large files (optional)

Support for large file systems is included with the Enterprise Edition.

In conjunction with VxVM, VxFS can support file systems up to 8 exabytes in size.

For large database configurations, this eliminates the need to use multiple file systems because of the size limitations of the underlying physical devices.

You can create or mount file systems with or without large files by specifying either the `largefiles` or `nolargefiles` option in `mkfs` or `mount` commands.

See “[Creating a VxFS file system](#)” on page 34.

About restoring file systems using Storage Checkpoints

Storage Checkpoints can be used by backup and restore applications to restore either individual files or an entire file system. Restoring from Storage Checkpoints can recover data from incorrectly modified files, but typically cannot be used to recover from hardware damage or other file system integrity problems. File restoration can be done using the `fsckpt_restore(1M)` command.

See the *Veritas File System Administrator's Guide*.

About quotas

VxFS supports quotas, which allocate per-user and per-group quotas and limit the use of two principal resources: files and data blocks. You can assign quotas for each of these resources.

Each quota consists of two limits for each resource:

- The hard limit represents an absolute limit on data blocks or files. A user can never exceed the hard limit under any circumstances.
- The soft limit is lower than the hard limit and can be exceeded for a limited amount of time. This allows users to temporarily exceed limits as long as they fall under those limits before the allotted time expires.

You can use quotas to limit the amount of file system space used by Storage Checkpoints.

See the *Veritas File System Administrator's Guide*.

About cluster functionality (optional)

File system clustering is an optional, separately licensed feature of VxFS, where one system is configured as a primary server for the file system, and the other members of a cluster are configured as secondaries. All servers access shared disks for file data operations. If the primary server fails, one of the secondary servers takes over the file system operations.

See the *Veritas File System Administrator's Guide*.

Setting up dataservers

This chapter includes the following topics:

- [Tasks for setting up new databases](#)
- [About setting up a disk group](#)
- [Creating a disk group](#)
- [Adding disks to a disk group](#)
- [About selecting a volume layout](#)
- [Creating a volume](#)
- [Creating a volume set](#)
- [Adding a volume to a volume set](#)
- [File system creation guidelines](#)
- [Creating a VxFS file system](#)
- [Mounting a file system](#)
- [Unmounting a file system](#)
- [About fragmentation](#)
- [Resizing a file system](#)

Tasks for setting up new databases

If you are using Veritas Storage Foundation for Sybase to set up a new database, complete these tasks in the order listed below:

Determine the number and sizes of file systems you need for the database you want to create. See the *Veritas File System Administrator's Guide*.

Create volumes to meet your file system needs. You can use disk mirroring as a safeguard against disk failures and striping for better performance.

- See [“About selecting a volume layout”](#) on page 29.
- See [“Creating a volume ”](#) on page 31.
- See [“About tuning VxVM ”](#) on page 101.

If you plan to create volume snapshots for the database and use them on either the same host or a secondary one, ensure that your volume layout is consistent with Database FlashSnap requirements.

Create the VxFS file systems you need on the volumes. See [“File system creation guidelines”](#) on page 33.
See [“Creating a VxFS file system ”](#) on page 34.

Install and configure your database.

If you are not currently running on VxVM and VxFS, make sure Veritas Storage Foundation for Sybase is installed and convert your existing database configuration. See the *Veritas Storage Foundation for Sybase Installation Guide*.

For backup and recovery on the same host, you can use the Storage Checkpoint facility to create file system snapshots of the database. A Storage Checkpoint creates an exact image of a database instantly and provides a consistent image of the database from the point in time the Storage Checkpoint was created.

About setting up a disk group

Before creating file systems for a database, set up a disk group for each Sybase dataserver.

A disk group lets you group disks, volumes, file systems, and files that are relevant to a single database into a logical collection for easy administration. Because you can move a disk group and its components as a unit from one machine to another, you can move an entire database when all the configuration objects of the database are in one disk group. This capability is useful in a failover situation.

Disk group configuration guidelines

Follow these guidelines when setting up disk groups:

- Only disks that are online and do not already belong to a disk group can be used to create a new disk group.
- Create one disk group for each Sybase ASE server.
- The disk group name must be unique. Name each disk group using the Sybase dataserver name this disk group belongs to and a `dg` suffix. The dataserver name is the name of the Sybase server as defined in the Sybase interface file. The `dg` suffix helps identify the object as a disk group. Also, each disk name must be unique within the disk group.
- Never create database devices for a dataserver using file systems or volumes that are not in the same disk group.

In earlier releases of Veritas Volume Manager, a system installed with VxVM was configured with a default disk group, `rootdg`, that had to contain at least one disk. VxVM can now function without any disk group having been configured. Only when the first disk is placed under VxVM control must a disk group be configured.

Note: Most VxVM commands require superuser or equivalent privileges.

See [“About tuning VxVM ”](#) on page 101.

Creating a disk group

You can use the `vxvg` command to create a new disk group. A disk group must contain at least one disk at the time it is created. You also have the option to create a shared disk group for use in a cluster environment.

Disks must be placed in disk groups before they can be used by VxVM. You can create disk groups to organize your disks into logical sets of disks.

Before creating a disk group, review the following:

- | | |
|---------------|---|
| Prerequisites | <ul style="list-style-type: none">■ Only disks that are online and do not belong to a disk group can be used to create a disk group.■ The disk group name must be unique in the host or cluster.■ Creating a disk group requires at least one disk. |
|---------------|---|

Usage notes

- Veritas Storage Foundation for Sybase only supports single disk groups.
- New disks must be placed under VxVM control and then added to a dynamic disk group before they can be used for volumes.
- When you place a disk under VxVM control, the disk is either encapsulated or initialized. Encapsulation preserves any existing data on the disk in volumes. Initialization destroys any existing data on the disk.
- If you place the root disk under VxVM control, you must encapsulate the disk. If you want to create an alternate boot disk, you can mirror the encapsulated root disk.
- For information on the `vxldg` command, see the `vxldg(1M)` manual page.

To create a new disk group

- ◆ Use the `vxldg` command as follows:

```
# /opt/VRTS/bin/vxldg init disk_group [disk_name=disk_device]
```

The following is an example of creating a disk group using the `vxldg` command:

To create a disk group named `PRODDg` on a raw disk partition `sda`, where the disk name `PRODDg01` references the disk within the disk group:

```
# /opt/VRTS/bin/vxldg init PRODDg PRODDg01=sda
```

Adding disks to a disk group

When a disk group is first created, it can contain only a single disk. You may need to add more disks to the disk group. Before adding disks, review the following:

Usage notes

- When you place a disk under VxVM control, the disk is initialized. Initialization destroys any existing data on the disk.
- Disks must be under VxVM control and in a disk group before they can be used to create volumes.
- Disks must be online before they can be added to a disk group.
- Disks that already belong to a disk group cannot be added to another disk group.

To add disks to a disk group

- ◆ Use the `vxdg` command as follows:

```
# /opt/VRTS/bin/vxdg -g disk_group adddisk \  
[disk_name=disk_device]
```

The following is an example of adding disks to a disk group using the `vxdg` command:

To add disks named `PRODDg02`, `PRODDg03`, and `PRODDg04` to the disk group `PRODDg`:

```
# /opt/VRTS/bin/vxdg -g PRODDg adddisk PRODDg02=sda  
  
# /opt/VRTS/bin/vxdg -g PRODDg adddisk PRODDg03=sdb  
  
# /opt/VRTS/bin/vxdg -g PRODDg adddisk PRODDg04=sdc
```

About selecting a volume layout

Veritas Volume Manager offers a variety of layouts that allow you to configure your database to meet performance and availability requirements. The proper selection of volume layouts provides optimal performance for the database workload.

An important factor in database performance is the segment placement on the disks.

Disk I/O is one of the most important determining factors of your database's performance. Having a balanced I/O load usually means optimal performance. Designing a disk layout for the database objects to achieve balanced I/O is a crucial step in configuring a database.

Sybase maps each physical file or raw device to its database devices. Devices are grouped into segments; tables are created on segments. When deciding which devices to put in a segment, it is often difficult to anticipate future usage patterns. VxVM provides flexibility in configuring storage for the initial database set up and for continual database performance improvement as needs change. VxVM can split volumes across multiple drives to provide a finer level of granularity in data placement. By using striped volumes, I/O can be balanced across multiple disk drives. For most databases, ensuring that different database devices are distributed across the available disks may be sufficient.

Striping also helps sequential table scan performance. When a table is created on a segment containing database devices striped across multiple disks, a high transfer bandwidth can be achieved.

See [“About tuning VxVM”](#) on page 101.

How to choose appropriate stripe unit sizes

When creating a striped volume, you need to decide the number of columns to form a striped volume and the stripe unit size. You also need to decide how to stripe the volume. You may stripe a volume across multiple disk drives on the same controller or across multiple disks on multiple controllers. By striping across multiple controllers, disk I/O can be balanced across multiple I/O channels. The decision is based on the disk and controller bandwidth and the database workload. In general, for most OLTP databases, use the default stripe unit size of 64 K or smaller for striped volumes and 16 K for RAID-5 volumes.

How to choose between mirroring and RAID-5

VxVM provides two volume configuration strategies for data redundancy: mirroring and RAID-5. Both strategies allow continuous access to data in the event of disk failure. For most database configurations, we recommend using mirrored, striped volumes. If hardware cost is a significant concern, but having higher data availability is still important, use RAID-5 volumes.

RAID-5 configurations have certain performance implications you must consider. Writes to RAID-5 volumes require parity-bit recalculation, which adds significant I/O and CPU overhead. This overhead can cause considerable performance penalties in online transaction processing (OLTP) workloads. If the database has a high read ratio, however, RAID-5 performance is similar to that of a striped volume.

Volume configuration guidelines

Follow these guidelines when selecting volume layouts:

- Put the database log files on a file system created on a striped and mirrored (RAID-0+1) volume separate from the index or data segments. Stripe multiple devices to create larger volumes if needed. Use mirroring to improve reliability. Do not use VxVM RAID-5 for transaction logs.
- When normal system availability is acceptable, put the segments on file systems created on striped volumes for most OLTP workloads.
- It is generally good practice to place frequently used databases such as `tempdb` and `sybsystemprocs` on striped devices.
- Put log segment and default segment for each database on different volumes.
- Create striped volumes across at least four disks. Try to stripe across disk controllers. For sequential scans, do not stripe across too many disks or controllers. The single thread that processes sequential scans may not be able to keep up with the disk speed.

- For most workloads, use the default 64 K stripe-unit size for striped volumes and 16 K for RAID-5 volumes.
- When system availability is critical, use mirroring for most write-intensive OLTP workloads. Turn on Dirty Region Logging (DRL) to allow fast volume resynchronization in the event of a system crash.
- When system availability is critical, use RAID-5 for read-intensive OLTP workloads to improve database performance and availability. Use RAID-5 logs to allow fast volume resynchronization in the event of a system crash.
- For most decision support system (DSS) workloads, where sequential scans are common, experiment with different striping strategies and stripe-unit sizes. Put the most frequently accessed tables or tables that are accessed together on separate striped volumes to improve the bandwidth of data transfer.

See [“About tuning VxVM ”](#) on page 101.

Creating a volume

Veritas Volume Manager uses logical volumes to organize and manage disk space. A volume is made up of portions of one or more physical disks, so it does not have the limitations of a physical disk.

For databases where the data storage needs to be resilient and the data layout needs to be optimized for maximum performance, we recommend using VxVM. The striping and mirroring capabilities offered by a volume manager will help you achieve your manageability, availability, and performance goals.

After you decide on a volume layout, you can use the `vxassist` command to create the volume.

Before creating a volume, make sure the following conditions are met:

- Usage notes
- Creating a volume requires a disk group name, volume name, volume size, and volume layout. You will also need to know subdisk names if you are creating a striped volume.
 - Striped or mirrored volumes require at least two disks.
 - Striped pro and concatenated pro volumes are mirrored by default, so a striped pro volume requires more disks than an unmirrored striped volume and a concatenated pro volume requires more disks than an unmirrored concatenated volume.
 - You cannot use a striped pro or concatenated pro volume for a `root` or `swap` volume.
 - A RAID-5 volume requires at least three disks. If RAID-5 logging is enabled, a RAID-5 volume requires at least four disks. RAID-5 mirroring is not supported.

To create a volume

- ◆ Use the `vxassist` command as follows:

```
# /opt/VRTS/bin/vxassist -g disk_group make volume_name \  
size disk_name
```

The following is an example of creating a volume using the `vxassist` command:

To create a 1 GB mirrored volume called `db01` on the `PRODdg` disk group:

```
# /opt/VRTS/bin/vxassist -g PRODdg make db01 1g PRODdg01
```

Creating a volume set

Volume Sets enable the use of the Multi-Volume Support feature with Veritas File System (VxFS). It is also possible to use the Veritas Enterprise Administrator (VEA) to create and administer volumes sets. For more information, see the VEA online help.

Before creating a volume set, make sure the following conditions are met:

- | | |
|-------------|--|
| Usage notes | <ul style="list-style-type: none"> ■ Before creating a volume set, you must have at least one volume created.
See “Creating a volume” on page 31. ■ A maximum of 256 volumes may be configured in a volume set. ■ Only Veritas File System is supported on a volume set. ■ The first volume in a volume set must be larger than 20MB. ■ Raw I/O from and to a volume set is not supported. ■ Volume sets can be used instead of volumes with the following <code>vxsnap</code> operations on instant snapshots: <code>addmir</code>, <code>dis</code>, <code>make</code>, <code>prepare</code>, <code>reattach</code>, <code>refresh</code>, <code>restore</code>, <code>rmir</code>, <code>split</code>, <code>syncpause</code>, <code>syncresume</code>, <code>syncstart</code>, <code>syncstop</code>, <code>syncwait</code>, and <code>unprepare</code>. The third-mirror break-off usage model for full-sized instant snapshots is supported for volume sets provided that sufficient plexes exist for each volume in the volume set. See the <i>Veritas Volume Manager Administrator's Guide</i>. ■ Most VxVM commands require superuser or equivalent privileges. ■ For details regarding usage of the <code>vxvset</code> command, see the <code>vxvset (1M)</code> manual page. |
|-------------|--|

To create a volume set for use by Veritas file system (VxFS)

- ◆ Use the following command:

```
# /opt/VRTS/bin/vxvset [-g diskgroup] -t vxfs make volset volume
```

where:

- `volset` is the name of the volume set
- `volume` is the name of the first volume in the volume set
- `-t` defines the content handler subdirectory for the application that is to be used with the volume. This subdirectory contains utilities that an application uses to operate on the volume set. The operation of these utilities is determined by the requirements of the application and not by VxVM.

For example, to create a volume set named `db01vset` that contains the volume `db01`, in the disk group `PRODDg`, you would use the following command:

```
# /opt/VRTS/bin/vxvset -g PRODDg -t vxfs make db01vset db01
```

Adding a volume to a volume set

After creating a volume set, you can add additional volumes.

To add a volume to a volume set

- ◆ Use the `vxvset` command as follows:

```
# /opt/VRTS/bin/vxvset [-g diskgroup] [-f] addvol volset \  
volume
```

Warning: The `-f` (force) option must be specified if the volume being added, or any volume in the volume set, is either a snapshot or the parent of a snapshot. Using this option can potentially cause inconsistencies in a snapshot hierarchy if any of the volumes involved in the operation is already in a snapshot chain.

See the *Veritas Volume Manager Administrator's Guide*.

For example, to add the volume `db02`, to the volume set `db01vset`, use the following command:

```
# /opt/VRTS/bin/vxvset -g PRODDg addvol db01vset db02
```

File system creation guidelines

Follow these guidelines when creating VxFS file systems:

- Specify the maximum log size when creating file systems for databases.
- Except for specifying the maximum log size and support for large files as required, use the VxFS defaults when creating file systems for databases.
- Never disable the intent logging feature of the file system.
- For log segments in user databases, use database devices created on file systems with simple (and mirrored, if necessary) volumes. Put the other database devices on file systems created on striped, striped and mirrored, or mirrored and striped volumes.
- When using the command line, use the mount points to name the underlying volumes. For example, if a file system named `/db01` is to be created on a mirrored volume, name the volume `db01` and the mirrors `db01-01` and `db01-02` to relate to the configuration objects. If you are using the `vxassist` command or the GUI, this is transparent.

Creating a VxFS file system

You can create a file system on a volume, as long as the volume is large enough to accommodate the file system. We recommend creating a VxFS file system.

Before creating a file system, see the following notes:

- Usage notes
- See the `mkfs(1M)` and `mkfs_vxfs(1M)` manual pages for more information about the options and variables available for use with the `mkfs` command.
 - See the `mount(1M)` and `mount_vxfs(1M)` manual pages for more information about mount settings.

To create a VxFS file system on an existing volume

- ◆ Use the `mkfs` command as follows:

```
# /usr/sbin/mkfs -t vxfs [generic_options] \  
[-o specific_options] special [size]
```

where:

- `vxfs` is the file system type
- `generic_options` are the options common to most file systems
- `specific_options` are options specific to the VxFS file system

- `special` is the full path name of the raw character device or VxVM volume on which to create the file system (for example, `/dev/vx/rdisk/PRODDg/db01`)
- `size` is the size of the new file system (optional)

If you do not specify `size`, the file system will be as large as the underlying volume.

For example, to create a VxFS file system that has an 8 KB block size and supports files larger than 2 GB on the newly created `db01` volume:

```
# /usr/sbin/mkfs -t vxfs -o largefiles,bsize=8192,\
logsize=2000 /dev/vx/rdisk/PRODDg/db01
```

The `-o largefiles` specific option allows you to create files larger than 2 GB.

Note: Because `size` is not specified in this example, the size of the file system will be calculated automatically to be the same size as the volume on which the file system is created.

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

```
version 6 layout

20480 sectors, 10240 blocks of size 1024, log size 1024 blocks
```

You can now mount the newly created file system.

See “[Mounting a file system](#)” on page 36.

Multi-volume support

The multi-volume support feature enabled by VxFS Version 6 disk layout allows several volumes to be represented by a single logical object, known as a volume set. The `vxvset` command can be used to create and administer volume sets in Veritas Volume Manager.

VxFS's multi-volume support feature can be used with volume sets. There are two VxFS commands associated with multi-volume support:

- `fsapadm` - VxFS allocation policy administration utility
- `fsvoladm` - VxFS device administration utility

See the *Veritas File System Administrator's Guide*.

Mounting a file system

After creating a VxFS file system, mount the file system using the `mount` command.

Before mounting a file system, review the following:

- | | |
|---------------|---|
| Prerequisites | <ul style="list-style-type: none">■ A file system must exist in order to be mounted.■ DBAs should log in as the Sybase DBA user. |
| Usage notes | <ul style="list-style-type: none">■ The mount point must be an absolute path name (that is, it must begin with <code>/</code>).■ See the <code>mount_vxfs (1M)</code> manual page for more information about mount settings.■ See the <code>mount (1M)</code> manual page for more information about generic mount options. |

To mount a file system

- ◆ Use the `mount` command as follows:

```
# /usr/sbin/mount -t vxfs [generic_options] [-r] \  
[-o specific_options] special /mount_point
```

where:

- `generic_options` are the options common to most file systems
- `-r` mounts the file system as read only
- `specific_options` are options specific to the VxFS file system
- `special` is a block special device
- `/mount_point` is the directory where the file system will be mounted

For example, to mount a file system named `/db01` that supports large files on volume `/dev/vx/dsk/PRODDg/db01`:

```
# mkdir /db01  
  
# /usr/sbin/mount -t vxfs -o largefiles /dev/vx/dsk/PRODDg/db01 \  
/db01
```

If you would like `/db01` to be mounted automatically after rebooting, add an entry for it in `/etc/fstab` as follows:

```
/dev/vx/dsk/PRODDg/db01 /db01 vxfs largefiles,qio 0 2
```

If you would like `/db01` to be mounted automatically after rebooting, add a tab-separated entry for it in `/etc/fstab` as follows:

```
/dev/vx/dsk/PRODDg/db01 /dev/vx/rdisk/PRODDg/do1 /db01 vxfs \  
largefiles 0 2
```

Unmounting a file system

If you no longer need to access the data in a file system, you can unmount the file system using the `umount` command.

Before unmounting a file system, review the following:

- | | |
|---------------|---|
| Prerequisites | ■ A file system must exist and be mounted in order to be unmounted. |
| Usage notes | ■ You cannot unmount a file system that is in use.
See the <code>umount (1M)</code> manual page for more information on mounting file systems. |

To unmount a file system

- 1 Use the `fuser` command to make sure that the file system is not being used:

```
# fuser -c /mount_point
```

where the `-c` option provides information on file system mount points and any files within mounted file systems.

If the file system is being used and you need to unmount it, use the `fuser -ck` command. See the `fuser(1M)` man page for more information.

- 2 Unmount the file system using one of the `umount` command options:

- `umount special`
- `umount /mount_point`
- `umount -f /mount_point`

where:

- `special` is a block special device
- `/mount_point` is the location where the file system is mounted
- `-f` forcibly unmounts the mount point

The following is an example of unmounting a file system:

To verify that the file system `/db01` is not in use and then unmount the file system:

```
# fuser -c /db01
/db01:
# umount /db01
```

About fragmentation

When free resources are initially allocated to files in a Veritas file system, they are aligned in the most efficient order possible to provide optimal performance. On an active file system, the original order is lost over time as files are created, removed, or resized. As space is allocated and deallocated from files, the available free space becomes broken into fragments. This means that space must be assigned to files in smaller and smaller extents. This process is known as fragmentation. Fragmentation leads to degraded performance and availability. The degree of fragmentation depends on file system usage and activity.

How to control fragmentation

It is very rare to have a badly fragmented VxFS file system in an ASE environment. However, fragmentation can occur when many database devices are created and deleted.

VxFS provides online reporting and optimization utilities to enable you to monitor and defragment a mounted file system. These utilities are accessible through the file system administration command, `fsadm`. Using the `fsadm` command, you can track and eliminate fragmentation without interrupting user access to the file system.

Types of fragmentation

VxFS addresses two types of fragmentation:

- **Directory Fragmentation**
As files are created and removed, gaps are left in directory inodes. This is known as directory fragmentation. Directory fragmentation causes directory lookups to become slower.
- **Extent Fragmentation**
As files are created and removed, the free extent map for an allocation unit changes from having one large free area to having many smaller free areas. Extent fragmentation occurs when files cannot be allocated in contiguous chunks and more extents must be referenced to access a file. In a case of extreme fragmentation, a file system may have free space that cannot be allocated.

How to monitor fragmentation

You can monitor fragmentation in VxFS by running reports that describe fragmentation levels. Use the `fsadm` command to run reports on directory fragmentation and extent fragmentation. The `df` command, which reports on file system free space, also provides information useful in monitoring fragmentation.

Use the following commands to report fragmentation information:

- `fsadm -D`, which reports on directory fragmentation.
- `fsadm -E`, which reports on extent fragmentation.
- `/opt/VRTS/bin/fsadm [-t vxfs] -o s`, which prints the number of free extents of each size.

Defragmenting a file system

You can use the online administration utility `fsadm` to defragment or reorganize file system directories and extents.

The `fsadm` utility defragments a file system mounted for read/write access by:

- Removing unused space from directories.
- Making all small files contiguous.
- Consolidating free blocks for file system.

The following options are for use with the `fsadm` utility:

- | | |
|-----------------|---|
| <code>-d</code> | Reorganizes directories. Directory entries are reordered to place subdirectory entries first, then all other entries in decreasing order of time of last access. The directory is also compacted to remove free space.
Note: If you specify <code>-d</code> and <code>-e</code> , directory reorganization is always completed first. |
| <code>-a</code> | Use in conjunction with the <code>-d</code> option to consider files not accessed within the specified number of days as “aged” files. Aged files are moved to the end of the directory. The default is 14 days. |
| <code>-e</code> | Reorganizes extents. Files are reorganized to have the minimum number of extents.
Note: If you specify <code>-d</code> and <code>-e</code> , directory reorganization is always completed first. |

- `-D -E` Produces reports on directory and extent fragmentation, respectively.
Note: If you use both `-D` and `-E` with the `-d` and `-e` options, the fragmentation reports are produced both before and after reorganization.
- `-v` Specifies verbose mode and reports reorganization activity.
- `-l` Specifies the size of a file that is considered large. The default is 64 blocks.
- `-t` Specifies a maximum length of time to run, in seconds.
Note: The `-t` and `-p` options control the amount of work performed by `fsadm`, either in a specified time or by a number of passes. By default, `fsadm` runs five passes. If both `-t` and `-p` are specified, `fsadm` exits if either of the terminating conditions are reached.
- `-p` Specifies a maximum number of passes to run. The default is five.
Note: The `-t` and `-p` options control the amount of work performed by `fsadm`, either in a specified time or by a number of passes. By default, `fsadm` runs five passes. If both `-t` and `-p` are specified, `fsadm` exits if either of the terminating conditions are reached.
- `-s` Prints a summary of activity at the end of each pass.
- `-r` Specifies the pathname of the raw device to read to determine file layout and fragmentation. This option is used when `fsadm` cannot determine the raw device.

Note: You must have superuser (`root`) privileges to reorganize a file system using the `fsadm` command.

To defragment a file system

- ◆ Run the `fsadm` command followed by the options specifying the type and amount of defragmentation. Complete the command by specifying the mount point or raw device to identify the file system.

```
# /opt/VRTS/bin/fsadm [-d] [-D] [-e] [-E] [-s] [-v] \  
[-l largesize] [-a days] [-t time] [-p pass_number] \  
[-r rawdev_path] mount_point
```

Refer to the *Veritas File System Administrator's Guide* for instructions and information on scheduling defragmentation. *Veritas File System Administrator's Guide* for instructions and information on scheduling defragmentation.

For example, to defragment a file system:

```
# /opt/VRTS/bin/fsadm -d -D /sybdata_qiovm
```

```
Directory Fragmentation Report
```

	Dirs	Total	Immed	Immeds	Dirs to	Blocks
	Searched	Blocks	Dirs	to Add	Reduce	Reduce
total	5	1	4	0	0	0

```
Directory Fragmentation Report
```

	Dirs	Total	Immed	Immeds	Dirs to	Blocks
	Searched	Blocks	Dirs	to Add	Reduce	Reduce
total	5	1	4	0	0	0

Resizing a file system

If you need to extend or shrink a VxFS file system, you can use the `fsadm` command.

If a VxFS file system requires more space, you can use this procedure to extend the size of the file system. If a VxFS File System is too large and you need the space elsewhere, you can use this procedure to shrink the file system.

Remember to increase the size of the underlying device or volume before increasing the size of the file system.

See the *Veritas Volume Manager Administrator's Guide*. *Veritas Volume Manager Administrator's Guide*.

Before resizing a file system, review the following:

- Prerequisites
- This task requires a mounted file system. You must know either the desired size or the amount of space to add to or subtract from the file system size.
- Usage notes
- See the (1M) manual page. See the `fsadm_vxfs(1M)` manual page.

To resize a file system

- ◆ Use `fsadm` command as follows:

```
# /opt/VRTS/bin/fsadm -t vxfs [-b newsize] \  
[-r rawdev] /mount_point
```

where:

- `newsiz` is the size (in sectors) to which the file system will increase or shrink
- `/mount_point` is the location where the file system is mounted

For example, to extend the file system `/db01` to 2 GB:

```
# /opt/VRTS/bin/fsadm -t vxfs -b 2g /db01
```

Note: See the *Veritas File System Administrator's Guide* and `fsadm_vxfs(1M)` manual page for information on how to perform common file system tasks using `fsadm`.

Resizing a file system and the underlying volume

The `fsadm` command resizes the file system only. If you attempt to use `fsadm` to make the file system the same size or larger than the underlying volume, the `fsadm` command will fail. To resize the file system and its underlying volume, use the `vxresize` command instead.

Warning: Resizing a volume with a usage type other than FSGEN or RAID5 can result in data loss. If such an operation is required, use the `-f` option to forcibly resize such a volume.

Before resizing a file system and the underlying volume, review the following:

- Prerequisites
- You must know the new desired size of the file system.

- Usage notes
- `vxresize` works with VxFS, ext2, ext3, and reiserfs file systems only.
 - If the file system is mounted or VxFS, you can grow or shrink the size. If the file system is Lite-VxFS, ext2, ext3, or reiserfs, you cannot grow or shrink the size.
 - If the file system is unmounted and VxFS or Lite-VxFS, you cannot grow or shrink the size. If the file system is ext2, ext3, or reiserfs, you can grow or shrink the size.
 - When resizing large volumes, `vxresize` may take a long time to complete.
 - Resizing a volume with a usage type other than FSGEN or RAID5 can result in data loss. If such an operation is required, use the `-f` option to forcibly resize such a volume.
 - You cannot resize a volume that contains plexes with different layout types.
 - See the `vxresize (1M)` manual page for more details.

To resize a file system and the underlying volume

- ◆ Use the `vxresize` command as follows:

```
# /etc/vx/bin/vxresize -g disk_group -b -F vxfs -t \  
homevolresize homevol volume_size disk_name disk_name
```

For example, to extend a 1-gigabyte volume, `homevol`, that contains a VxFS file system, to 10 gigabytes using the spare disks `disk10` and `disk11`, enter:

```
# /etc/vx/bin/vxresize -b -F vxfs -t homevolresize homevol 10g \  
disk10 disk11
```

The `-b` option specifies that this operation runs in the background. Its progress can be monitored by specifying the task tag `homevolresize` to the `vxtask` command.

Using Veritas Concurrent I/O

This chapter includes the following topics:

- [About Concurrent I/O](#)
- [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.

How Concurrent I/O works

Traditionally, Linux 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

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.

Enabling Concurrent I/O

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
```

```
# /usr/sbin/mount -t 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.

Disabling Concurrent I/O

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 Sybase instance.
- 2** Unmount the file system using the `umount` command.
- 3** Mount the file system again using the `mount` command without using the `-o cio` option.

Converting existing database configurations to VxFS

This chapter includes the following topics:

- [Upgrading from earlier VxFS version layouts](#)
- [About snapshot volumes](#)
- [About snapshot volumes](#)

Upgrading from earlier VxFS version layouts

Before starting the upgrade process, make sure the following conditions have been met:

- | | |
|---------------|--|
| Prerequisites | ■ Perform a full backup of the file system before upgrading to a new disk layout. |
| Usage notes | <ul style="list-style-type: none">■ The <code>vxupgrade</code> command lets you to upgrade the VxFS file system disk layout while the file system is mounted. See the <code>vxupgrade(1M)</code> manual page for more details.■ VxFS supports two file system disk layouts: Versions 4 and 6. New file systems are created with the Version 6 (for large file systems) disk layout by default when the current version of Veritas Storage Foundation for Sybase is installed on a system. You must minimally upgrade to Version 4 disk layout if you want to use the Storage Rollback or Veritas NetBackup BLI Backup features. |

To upgrade an existing VxFS file system to a new file system disk layout version

- ◆ Use the `vxupgrade` command to upgrade to Version 4, 6, or 7 disk layout:

```
# /opt/VRTS/bin/upgrade -n new_version new_version/mount_point
```

where:

- `new_version` is the version of the file system disk layout you want to upgrade to
- `/mount_point` is the location where the file system is mounted

This is an example of upgrading to disk layout Version 7:

```
# /opt/VRTS/bin/vxupgrade -n 7 /db01
```

- 1 Shut down the database.
- 2 Restart the database.

About snapshot volumes

Veritas Volume Manager (VxVM) provides a snapshot facility to create a point-in-time image of a volume to use as a source for taking a backup or for off-host processing. This capability is provided through the `vxsnap` command, and the FastResync and disk group move, split, and join features of VxVM, which are described in this chapter.

A snapshot volume can be used to create a database clone, which can be used on a secondary host for off-host processing, including decision-support analysis and reporting, application development and testing, database backup, and logical error recovery. A snapshot volume can be resynchronized with the primary database volumes. In addition, in the event of a failure, the primary database can be recovered by resynchronizing it with the snapshot volume.

VxVM's snapshot operation creates a new volume that is an exact point-in-time copy of an existing volume. This is done by creating a mirror of the existing volume using disk space from the pool of free disk space. The mirror is brought up-to-date and a separate snapshot volume is then created. You can use the snapshot volume to make a backup of the original volume at a convenient time without stopping the original volume. A volume can be moved from one disk group to another, provided that there is no disk that is shared between the two disk groups. You can split the snapshot volumes to a new disk group, and import it onto another host to perform off-host backup and decision support operations.

To ensure the mirrors are a consistent and recoverable image of the database, the snapshot functionality must be used in conjunction with Sybase ASE 12.5's

database quiesce feature. The Sybase database must be put into the quiescent mode to temporarily suspend I/O to the dataserver before attempting to break off mirrors. Taking a volume snapshot occurs quickly, so the ASE server can be released from the quiescent state immediately after the mirrors are broken off.

Snapshot volumes can also be used in conjunction with the cluster functionality of VxVM.

After the snapshot mirror is synchronized, it continues being updated until it is detached. You can then select a convenient time at which to create snapshot volumes for all the volumes used by the database to represent a valid backup image of the database. You need to either shut down the database for an offline backup or suspend I/O writes to the database for an online backup during the brief time required to detach the snapshot volume (typically less than a minute). In contrast to the brief amount of time that it takes to detach a mirror and create a snapshot volume, the amount of time involved in creating a snapshot mirror is long and directly proportional to the size of the original volume.

About snapshot volumes

Veritas Volume Manager (VxVM) provides a snapshot facility to create a point-in-time image of a volume to use as a source for taking a backup or for off-host processing. This capability is provided through the `vxsnap` command, and the FastResync and disk group move, split, and join features of VxVM, which are described in this chapter.

A snapshot volume can be used to create a database clone, which can be used on a secondary host for off-host processing, including decision-support analysis and reporting, application development and testing, database backup, and logical error recovery. A snapshot volume can be resynchronized with the primary database volumes. In addition, in the event of a failure, the primary database can be recovered by resynchronizing it with the snapshot volume.

VxVM's snapshot operation creates a new volume that is an exact point-in-time copy of an existing volume. This is done by creating a mirror of the existing volume using disk space from the pool of free disk space. The mirror is brought up-to-date and a separate snapshot volume is then created. You can use the snapshot volume to make a backup of the original volume at a convenient time without stopping the original volume. A volume can be moved from one disk group to another, provided that there is no disk that is shared between the two disk groups. You can split the snapshot volumes to a new disk group, and import it onto another host to perform off-host backup and decision support operations.

To ensure the mirrors are a consistent and recoverable image of the database, the snapshot functionality must be used in conjunction with Sybase ASE 12.5's

database quiesce feature. The Sybase database must be put into the quiescent mode to temporarily suspend I/O to the dataserver before attempting to break off mirrors. Taking a volume snapshot occurs quickly, so the ASE server can be released from the quiescent state immediately after the mirrors are broken off.

Snapshot volumes can also be used in conjunction with the cluster functionality of VxVM.

After the snapshot mirror is synchronized, it continues being updated until it is detached. You can then select a convenient time at which to create snapshot volumes for all the volumes used by the database to represent a valid backup image of the database. You need to either shut down the database for an offline backup or suspend I/O writes to the database for an online backup during the brief time required to detach the snapshot volume (typically less than a minute). In contrast to the brief amount of time that it takes to detach a mirror and create a snapshot volume, the amount of time involved in creating a snapshot mirror is long and directly proportional to the size of the original volume.

Using volume snapshots for dataserver backup and off-host processing

This chapter includes the following topics:

- [About snapshot volumes](#)
- [Backup and off-host processing applications](#)
- [FastResync of snapshot volumes](#)
- [Disk group split and join](#)
- [Preparing hosts for database backup or off-host processing](#)
- [Sybase Adaptive Server Enterprise 12.5 quiesce feature](#)
- [How to set up volume snapshots with Sybase ASE 12.5 server](#)
- [Implementing online backup or off-host processing](#)
- [Creating a warm standby server](#)
- [Resynchronizing the snapshot to your ASE dataserver](#)
- [Recovering the database from a backup image](#)
- [Refreshing a snapshot database image](#)
- [Dissociating a snapshot volume](#)
- [Removing a snapshot volume](#)

About snapshot volumes

Veritas Volume Manager (VxVM) provides a snapshot facility to create a point-in-time image of a volume to use as a source for taking a backup or for off-host processing. This capability is provided through the `vxsnap` command, and the FastResync and disk group move, split, and join features of VxVM, which are described in this chapter.

A snapshot volume can be used to create a database clone, which can be used on a secondary host for off-host processing, including decision-support analysis and reporting, application development and testing, database backup, and logical error recovery. A snapshot volume can be resynchronized with the primary database volumes. In addition, in the event of a failure, the primary database can be recovered by resynchronizing it with the snapshot volume.

VxVM's snapshot operation creates a new volume that is an exact point-in-time copy of an existing volume. This is done by creating a mirror of the existing volume using disk space from the pool of free disk space. The mirror is brought up-to-date and a separate snapshot volume is then created. You can use the snapshot volume to make a backup of the original volume at a convenient time without stopping the original volume. A volume can be moved from one disk group to another, provided that there is no disk that is shared between the two disk groups. You can split the snapshot volumes to a new disk group, and import it onto another host to perform off-host backup and decision support operations.

To ensure the mirrors are a consistent and recoverable image of the database, the snapshot functionality must be used in conjunction with Sybase ASE 12.5's database quiesce feature. The Sybase database must be put into the quiescent mode to temporarily suspend I/O to the dataserver before attempting to break off mirrors. Taking a volume snapshot occurs quickly, so the ASE server can be released from the quiescent state immediately after the mirrors are broken off.

Snapshot volumes can also be used in conjunction with the cluster functionality of VxVM.

After the snapshot mirror is synchronized, it continues being updated until it is detached. You can then select a convenient time at which to create snapshot volumes for all the volumes used by the database to represent a valid backup image of the database. You need to either shut down the database for an offline backup or suspend I/O writes to the database for an online backup during the brief time required to detach the snapshot volume (typically less than a minute). In contrast to the brief amount of time that it takes to detach a mirror and create a snapshot volume, the amount of time involved in creating a snapshot mirror is long and directly proportional to the size of the original volume.

Backup and off-host processing applications

The following are typical backup and off-host processing applications made possible using the vxassist command, FastResync, and disk group move, split, and join features of VxVM:

- **Database Backup and Restore:** Many enterprises require 24/7 online data availability. They cannot afford the downtime involved in backing up critical data offline. By taking a snapshot of the data and then using it to back up your data, your business-critical applications can continue to run without extended down time or impacted performance. After a snapshot volume is created, it can be used as a source to backup the volume.
- **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 clone database that you have created from snapshot volumes, rather than allowing them to compete for access to the primary volume or database. When required, you can quickly resynchronize the clone database with the primary database to get up-to-date information.
- **Application Development and Testing:** Development or service groups can use a clone database created with snapshot volumes as a test database for new applications. A clone database provides developers, system testers, and quality assurance groups with a realistic basis for testing the robustness, integrity, and performance of new applications.
- **Logical Error Recovery:** Logical errors caused by an administrator or an application program can compromise the integrity of a database. You can recover a database by restoring the database files from a snapshot volume or by recovering logical objects (such as tables, for example) from a clone database created from snapshot volumes. These solutions are faster than fully restoring database files from tape or other backup media.

FastResync of snapshot volumes

Note: You may need an additional license to use this feature.

FastResync optimizes mirror resynchronization by keeping track of updates to stored data that have been missed by a mirror. If 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. If the data in one mirror is not updated for some reason, it becomes out-of-date, or stale, with respect to the other mirrors

in the volume. The presence of the FastResync map means that only those updates that the mirror has missed need be reapplied to resynchronize it with the volume. A full (and therefore much slower) resynchronization of the mirror from the volume is unnecessary. The FastResync feature increases the efficiency of the VxVM snapshot mechanism to better support operations such as backup and decision support.

The persistent form of FastResync ensures that FastResync maps survive both system crashes and cluster restarts. When snapshot volumes are reattached to their original volumes, FastResync allows the snapshot data to be quickly refreshed and re-used. If Persistent FastResync is enabled on a volume in a private disk group, such incremental resynchronization can happen 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. When the disk groups are rejoined, this allows the snapshot plexes to be quickly resynchronized. Non-Persistent FastResync cannot be used for this purpose.

FastResync allows you to refresh and re-use snapshots rather than discard them. You can quickly resynchronize a snapshot volume with its original volume. This reduces the system overhead required to perform cyclical operations, such as backups, that rely on the snapshot functionality of VxVM. You can also resynchronize the original volume from the snapshot volume. In this case, the database must be shut down and all of the file systems on the original volumes must be unmounted.

Up to 31 snapshot mirrors can be taken and tracked via FastResync.

For more information about FastResync, see the *Veritas Volume Manager Administrator's Guide*.

Disk group split and join

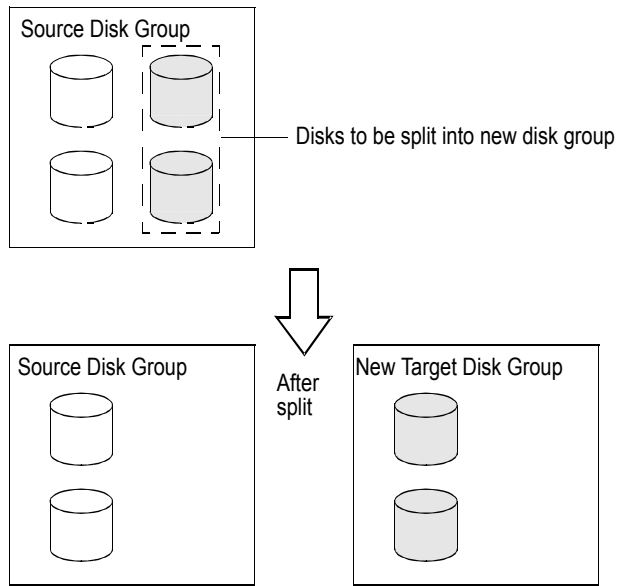
Note: You may need an additional license to use these features.

A snapshot volume can be split off into a separate disk group and deported. It is then ready for importing on another host that is dedicated to off-host processing. At a later stage, the disk group can be deported, re-imported, and joined with the original disk group or with a different disk group.

The `split` and `join` operations allow you to move VxVM objects such as disks or top-level volumes from one disk group to another.

The `split` operation is illustrated in [Figure 5-1](#).

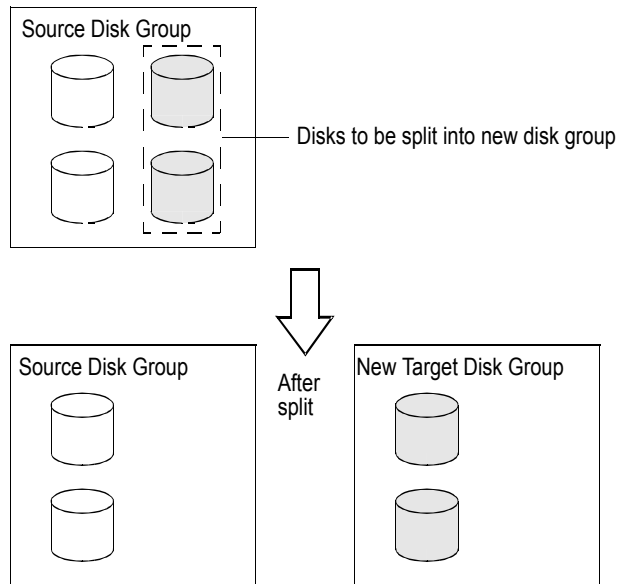
Figure 5-1 Disk group split operation



The `join` operation allows you to remove all VxVM objects from an imported disk group and move them to an imported target disk group. The source disk group is removed when the `join` is complete.

The `join` operation is illustrated in [Figure 5-2](#).

Figure 5-2 Disk group join operation



Caution: Before moving volumes between disk groups, you must stop all applications that are accessing the volumes and unmount all file systems that are configured in the volumes.

If the system crashes or a hardware subsystem fails, VxVM attempts to complete or reverse an incomplete disk group reconfiguration when the system is restarted or the hardware subsystem is repaired, depending on how far the reconfiguration had progressed. If one of the disk groups is no longer available because it has been imported by another host or because it no longer exists, you must recover the disk group manually.

The disk group `move`, `split` and `join` features have the following limitations:

- Disk groups involved in a `move`, `split` or `join` must be version 90 or greater. If needed, you can upgrade your volume. Refer to the *Veritas Volume Manager Administrator's Guide*.
- The reconfiguration must involve an integral number of physical disks.
- Objects to be moved must not contain open volumes.
- Moved volumes are initially disabled following a disk group `move`, `split` or `join`. If required, use either `vxrecover -m` or `vxvol startall` to restart the volumes.

- Data change objects (DCOs) and snap objects that have been dissociated by persistent FastResync cannot be moved between disk groups.
- Veritas Volume Replicator (VVR) objects cannot be moved between disk groups.
- For a disk group `move` to succeed, the source disk group must contain at least one disk that can store copies of the configuration database after the move.
- For a disk group `split` to succeed, both the source and target disk groups must contain at least one disk that can store copies of the configuration database after the split.
- For a disk group `move` or `join` to succeed, the configuration database in the target disk group must be able to accommodate information about all the objects in the enlarged disk group.
- Splitting or moving a volume into a different disk group changes the volume's record ID.

For more information, see the *Veritas Volume Manager Administrator's Guide*.

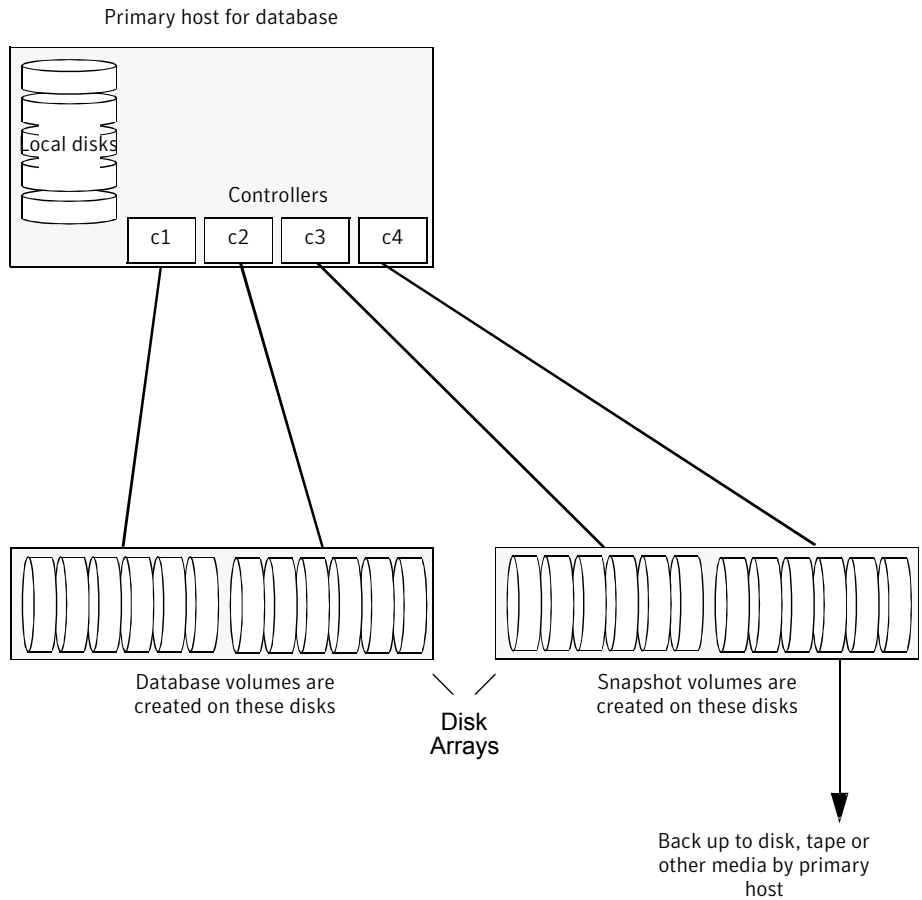
Preparing hosts for database backup or off-host processing

Snapshot volumes can be used on the same host that the database resides on (the primary host) or on a secondary host. On a secondary host, a snapshot volume can be used to implement regular online backup of a volume in a private disk group or set up a clone of the production database for decision support or off-host processing. Snapshot volumes can also be used on the primary host to create a backup image of a database.

Single-host configuration

[Figure 5-3](#) shows the suggested arrangement for using snapshot volumes on the primary host to avoid disk contention.

Figure 5-3 Example of a single-host configuration



Two-host configuration

Figure 5-4 shows the suggested arrangement for using snapshot volumes on a secondary host so that CPU- and I/O-intensive operations can be performed for online backup and decision support without degrading the performance of the primary host running the production database.

A two-host configuration also allows the snapshot volumes to avoid contending for I/O resources on the primary host.

For off-host processing applications, both the primary and secondary hosts need to be able to access the disks containing the snapshot volumes.

Figure 5-4 Example of a two-host configuration

Note: A snapshot volume represents 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 by the overlying file system, or by applications such as databases that have files open in the file system. If the `fsген` volume usage type is set on a volume that contains a Veritas File System (VxFS), intent logging of the file system metadata ensures the internal consistency of the file system that is backed up. For other file system types, depending on the intent logging capabilities of the file system, there may potentially be inconsistencies between in-memory data and the data in the snapshot image.

For databases, a suitable mechanism must additionally be used to ensure the integrity of segment data when the snapshot volume is taken. The facility to temporarily suspend file system I/O is provided by most modern database software. 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.

Upgrading existing volumes to use VxVM 4.0 features

This section describes how to upgrade a volume created before VxVM 4.0 so that it can take advantage of new features.

Note: The plexes of the DCO volume require persistent storage space on disk to be available. To make room for the DCO plexes, you may need to add extra disks to the disk group, or reconfigure existing volumes to free up space in the disk group. Another way to add disk space is to use the disk group move feature to bring in spare disks from a different disk group.

To upgrade an existing volume

- 1 Upgrade the disk group that contains the volume to the latest version before performing the remainder of the procedure described in this section. Use the following command to check the version of a disk group:

```
# vxdg list diskgroup
```

To upgrade a disk group to the latest version, use the following command:

```
# vxdg upgrade diskgroup
```

- 2 If the volume to be upgraded has an old-style DRL plex or subdisk, remove it:

```
# vxassist -g diskgroup remove log volume [nlog=n]
```

Use the optional attribute `nlog=` to specify the number, `n`, of logs to be removed. By default, the `vxassist` command removes one log.

- 3 For a volume that has one or more associated snapshot volumes, reattach and resynchronize each snapshot:

```
# vxassist -g diskgroup snapback snapvol
```

If persistent FastResync was enabled on the volume before the snapshot was taken, the data in the snapshot plexes is quickly resynchronized from the original volume. If persistent FastResync was not enabled, a full resynchronization is performed.

- 4 Turn off persistent FastResync for the volume:

```
# vxvol -g diskgroup set fastresync=off volume
```

- 5 Dissociate an old-style DCO object, DCO volume, and snap objects from the volume:

```
# vxassist -g diskgroup remove log volume logtype=dc0
```

- 6 Upgrade the volume:

```
# vxsnap [-g diskgroup] prepare volume [ndcomirs=number] \  
[regionsize=size] [drl=yes|no|sequential] \  
[storage_attribute ...]
```

The `ndcomirs` attribute specifies the number of DCO plexes that are created in the DCO volume. It is recommended that you configure as many DCO plexes as there are data and snapshot plexes in the volume. The DCO plexes are used to set up a DCO volume for any snapshot volume that you subsequently create from the snapshot plexes. For example, specify `ndcomirs=5` for a volume with 3 data plexes and 2 snapshot plexes.

The value of the `regionsize` attribute specifies the size of the tracked regions in the volume. A write to a region is tracked by setting a bit in the change map. The default value is `64k` (64 KB). A smaller value requires more disk space for the change maps, but the finer granularity provides faster resynchronization.

To enable DRL logging on the volume, specify `drl=yes`. If sequential DRL is required, specify `drl=sequential`.

You can also specify `vxassist`-style storage attributes to define the disks that can or cannot be used for the plexes of the DCO volume.

Note: The `vxsnap prepare` command automatically enables persistent FastResync on the volume and on any snapshots that are generated from it. If the volume is a RAID-5 volume, it is converted to a layered volume that can be used with snapshots and FastResync.

Sybase Adaptive Server Enterprise 12.5 quiesce feature

The Sybase Adaptive Server Enterprise (ASE) 12.5 `quiesce database` command allows you to temporarily suspend all writes to one or more databases. Background tasks will also skip any database that is in the suspended state. An ASE database

will stay in the quiesce state until a `quiesce database release` command is issued.

Without shutting down the ASE server, a database administrator (DBA) can take VxVM volume snapshots while the database is in the quiesce state to ensure that break-off mirrors contain a consistent and recoverable database image. Quiescing can be done at the database level and ASE databases that are not in the quiesce state can be updated as usual.

Note: With ASE 12.0, snapshot images taken when the database is in the quiesce state cannot be rolled forward. With ASE 12.5, they can be. Also, the `for external dump` clause was added to the `quiesce database` command in ASE 12.5 to permit rolling forward a snapshot database.

How to set up volume snapshots with Sybase ASE 12.5 server

Follow these guidelines when setting up the Sybase ASE 12.5 server to use volume snapshots:

- Veritas recommends that you use Quick I/O files instead of raw volumes.
- Do not share file systems and volumes between two ASE servers.
- Do not share the same disk group between two ASE servers.
- Avoid sharing file systems and volumes between user databases if the database administrator intends to create a warm standby for a single database only.
- When taking volume snapshots, create them on disks separate from the original volumes and preferably on the disks that belong to a separate array box. The `disk group split` command cannot move volumes if the move results in two disks sharing the same disk. This rule also applies when creating a DCO log volume for storing FastResync maps.
- Do not put database devices (Quick I/O files, for example) in a Sybase file system. Database files should not be in the Sybase software installation area.
- Do not turn on FastResync tracking for `tempdb` devices because they will be cleared every time the ASE server starts. In addition, the default data and log segments for `tempdb` on the master device should be dropped. `tempdb` should have its own file systems and volumes.

[Table 5-1](#) shows the storage layout for an example ASE server with FastResync running on the primary host. In this storage configuration:

- All volumes are mirrored and belong to the disk group `syb`.
- A secondary host is in the network and has I/O paths to disks in `syb`.
- Primary and secondary hosts access the same set of disks via fibre channel switches. They can also be configured to attach to a dual-hosted RAID box or JBOD.
- The production server is performing I/O in the example.

Table 5-1 Example of storage layout for an ASE server with FastResync

Volume Name	Mount Point	Sybase Device Name	Physical File Name
sybvoll	/sybasedata	Master	/sybasedata/snaptest/master
		Sysprocsdev	/sybasedata/snaptest/sysprocs

Implementing online backup or off-host processing

To implement online backup of an ASE database on a secondary host, you back up its database devices. You can set up a regular backup cycle on another host under Sybase ASE 12.0.0.1 ESD 1 or higher by combining the Persistent FastResync and disk group split and join features of VxVM.

This procedure can also be used to create a second Sybase ASE installation on a secondary host for off-host processing purposes, such as decision-support, reporting, and testing.

Before implementing off-host processing activities, make sure the following conditions have been met:

- Prerequisites
- You must be logged in as superuser (root).
 - The disk group must be version 90 or later. For more information on disk group versions, see the `vxdg(1M)` manual page.

Usage notes

When creating snapshot mirrors used by databases:

- Ensure that your Sybase instance owner directory is on a volume included in the snapshot.
- Create a separate disk group for Sybase database-related files.
- Do not share file systems and volumes between two ASE servers.
- Do not share the same disk group between two ASE servers.
- If you intend to move a snapshot volume to another host for off-host processing, do not share any disks between the original mirror plex and the snapshot mirror.
- Create snapshot mirrors on a separate controller and separate disk from the primary volume.
- Create snapshot mirrors for datafiles and archive logs so that they do not share any disks with the data of the original volumes. If they are not created in this way, they cannot be split and moved to a secondary host.
- Allocate separate volumes for logs and do not create snapshots on those volumes.
- Resynchronization speed varies based on the amount of data changed in both the primary and secondary volumes during the break-off time.

To back up database devices

- 1 On the primary host, see if the volume is associated with a version 20 data change object (DCO) and DCO volume that allow instant snapshots and Persistent FastResync to be used with the volume:

```
# vxprint -g diskgroup -F%instant volume_name
```

This command returns `on` if the volume can be used for instant snapshot operations; otherwise, it returns `off`.

If the volume was created under VxVM 4.0, and it is not associated with a new-style DCO object and DCO volume... Prepare the volume for DRL and instant snapshots.
See the *Veritas Volume Manager Administrator's Guide*.

If the volume was created before release 4.0 of VxVM, and it has any attached snapshot plexes or it is associated with any snapshot volumes... See "[Upgrading existing volumes to use VxVM 4.0 features](#)" on page 61.

- 2 Prepare the volume for being snapshot using the `vxsnap prepare` command:

```
# vxsnap -g diskgroup prepare volume_name [alloc=disk_name]
```

The `vxsnap prepare` command creates a DCO and DCO volumes and associates them with the volume. It also enables persistent FastResync on the volume.

- 3 On the primary host, verify that FastResync is enabled on the volume:

```
# vxprint -g diskgroup -F%fastresync volume_name
```

This command returns `on` indicating that FastResync is enabled.

- 4 Create a snapshot mirror of a volume:

```
# vxsnap -g diskgroup addmir volume_name alloc=disk_name \  
[nmirror=N]
```

where *N* specifies the number of mirrors.

Ensure that the mirror is created on a disk with no other volumes.

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.

Alternatively, create a new volume for use as the snapshot volume as described in *Veritas Volume Manager Administrator's Guide*.

- 5 Suspend I/O updates on the primary database by using the `quiesce database` command:

```
$ isql -Usa -Ppassword -Sserver
> quiesce database tag1 hold testdb, sybssystemdb, sybssystemprocs
> go
> quiesce database tag2 hold master
> go
> quit
```

- 6 If you added a snapshot plex to the volume earlier, create a full-sized instant snapshot volume by running the following command on the primary host:

```
# vxsnap -g diskgroup make \
source=volume/newvol=snapshot_volume/plex=plex_name
```

By default, VxVM attempts to avoid placing snapshot mirrors on a disk that already holds any plexes of a data volume. However, this may be impossible if insufficient space is available in the disk group. In this case, VxVM uses any available space on other disks in the disk group. If the snapshot plexes are placed on disks which are used to hold the plexes of other volumes, this may cause problems when you subsequently attempt to move a snapshot volume into another disk group. To override the default storage allocation policy, you can use storage attributes to specify explicitly which disks to use for the snapshot plexes.

See the *Veritas Volume Manager Administrator's Guide*.

If a database spans more than one volume, you can specify all the volumes and their snapshot volumes using one command, as shown here:

```
# vxsnap -g diskgroup make \
source=vol1/newvol=snapvol1/nmirror=1 \
source=vol2/newvol=snapvol2/nmirror=1 \
source=vol3/newvol=snapvol3/nmirror=1
```

Note: This step 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 the next step.

- 7** On the primary host, resume I/O updates on the primary database. To resume the updates, release the databases from quiesce mode:

```
$ isql -Usa -Ppassword -Sserver
> quiesce database tag2 release
> go
> quiesce database tag1 release
> go
> quit
```

- 8** On the primary host, use the following command to move the snapshot volume to another disk group from the original disk group:

```
# vxdg split diskgroup new_diskgroup snapshot_volume
```

The `split` (move) command will fail if the move will cause disks to be shared by two disk groups. The `split` (move) command will fail if the result of the move causes disks to be shared by two disk groups.

- 9** On the primary host, deport the snapshot volume's disk group:

```
# vxdg deport new_diskgroup
```

- 10** After the split, the snapshot volume is initially disabled. Use the following commands on the secondary host to recover and restart the snapshot volume:

```
# vxrecover -g new_diskgroup
# vxvol -g new_diskgroup start snapshot_volume
```

- 11** On the secondary host, 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 the commands for checking and mounting a file system:

```
# fsck -F vxfs /dev/vx/rdisk/new_diskgroup/snapshot_volume
# mount -F vxfs /dev/vx/dsk/new_diskgroup/snapshot_volume \
/mount_point
```

After the file system is mounted, administrators can back up the files using `cp` commands.

- 12** When the backup is finished, unmount the file system:

```
# umount /mount_point
```

- 13** On the secondary host, deport the snapshot volume's disk group:

```
# vxdg deport new_diskgroup
```

- 14** On the primary host, re-import the snapshot volume's disk group:

```
# vxdg import new_diskgroup
```

- 15** On the primary host, join the snapshot volume's disk group with the original volume's disk group:

```
# vxdg join new_diskgroup diskgroup
```

- 16** After the join, the snapshot volume is initially disabled. Use the following commands on the primary host to recover and restart the snapshot volume:

```
# vxrecover -g diskgroup -m snapshot_volume
```

The contents of the snapshot volume are now ready to be re-attached and resynchronized with the original volume. VxVM offers a `snaprint` command for users to check the percentage of volumes that need to be synchronized. In the `vxprint` output, this is referred to as `% DIRTY`.

- 17** To check the percentage of volumes that need to be synchronized:

```
# vxsnap -g diskgroup print snapshot_volume
```

snaprint<snapshot_volume>

- 18** To resynchronize with the original volume:

```
# vxsnap -g diskgroup reattach snapvol source=vol
```

Each time you need to back up the volume, repeat this procedure from the `disk split` command to move the snapshot volume to another disk group from the original disk group.

Creating a warm standby server

With volume snapshots, you can use ASE 12.5 `quiesce database` and `dump/load transaction` commands to create a standby database and roll it forward with a transaction dump from the production database.

In some Sybase environments, two installations are maintained: one for production and one for DSS or failover. The DBA creates the initial standby database server by first quiescing the databases on the primary server. Protected by the database

quiesce framework, the DBA creates snapshots for all the volumes used by the database server and then releases the databases. Next, the DBA can deport the volume snapshots to another host and start the standby database on that host. After the initial setup, the DBA periodically performs `dump transaction` commands on the production database and loads them onto the standby database. Between the load transactions, the standby database is available as read-only and users can run queries against the standby database that do not update it. If the production database fails, a DBA can switch the standby database to read-write mode to take over the production workload.

The following is a summary of the steps to create a warm standby server in the ASE 12.5 environment:

- Use VxVM volume snapshots to move the entire ASE server to the secondary host. After users perform the first load transaction step, the failover database will be online standing by.
- Perform periodic `dump transaction` commands for databases to a dedicated disk dump device. This disk dump device can reside on a separate disk group, for example, `dump_diskgroup`, which will be able to deport/import between the primary host and the secondary host. After the dump is finished, deport the disk group `dump_diskgroup` on the primary host.
- Import the disk group `dump_diskgroup` onto the secondary host and perform `load transaction` commands on the failover database.
- Repeat the `dump/load transaction` commands (steps 2 and 3 above) in fixed intervals. For example, repeat the process every hour.
- Repeat the entire process in fixed intervals. For example, repeat the process every day.

Before creating a warm standby server, make sure the following conditions have been met:

- Prerequisites
- You must be logged in as superuser (root).
 - The disk group must be version 90 or later. For more information on disk group versions, see the `vxldg(1M)` manual page.

Usage notes

- Ensure that your Sybase instance owner directory is on a volume included in the snapshot.
- Create a separate disk group for Sybase database-related files.
- Do not share file systems and volumes between two ASE servers.
- Do not share the same disk group between two ASE servers.
- If you intend to move a snapshot volume to another host for off-host processing, do not share any disks between the original mirror plex and the snapshot mirror.
- Create snapshot mirrors on a separate controller and separate disk from the primary volume.
- Create snapshot mirrors for datafiles and archive logs so that they do not share any disks with the data of the original volumes. If they are not created in this way, they cannot be split and moved to a secondary host.
- Allocate separate volumes for logs and do not create snapshots on those volumes.
- Resynchronization speed varies based on the amount of data changed in both the primary and secondary volumes during the break-off time.

To create a warm standby server

- 1 On the primary host, use the following command to see if the volume is associated with a version 20 data change object (DCO) and DCO volume that allow instant snapshots and Persistent FastResync to be used with the volume:

```
# vxprint -g diskgroup -F%instant volume_name
```

This command returns `on` if the volume can be used for instant snapshot operations; otherwise, it returns `off`.

If the volume was created under VxVM 4.0, and it is not associated with a new-style DCO object and DCO volume...

Prepare the volume for DRL and instant snapshots.

See the *Veritas Volume Manager Administrator's Guide*.

If the volume was created before release 4.0 of VxVM, and it has any attached snapshot plexes or it is associated with any snapshot volumes...

See "[Upgrading existing volumes to use VxVM 4.0 features](#)" on page 61.

Note: If the volume was created under PRODUCTNAME 4.0, and it is not associated with a new-style DCO object and DCO volume, follow the procedure for preparing a volume for DRL and instant snapshots in the *Veritas Volume Manager Administrator's Guide*.

If the volume was created before release 4.0 of PRODUCTNAME, and it has any attached snapshot plexes or it is associated with any snapshot volumes, follow the procedure given in .

- 2 Prepare the volume for taking a snapshot using the `vxsnap prepare` command:

```
# vxsnap -g diskgroup prepare volume_name [alloc=disk_name]
```

The `vxsnap prepare` command creates a DCO and DCO volumes and associates them with the volume. It also enables persistent FastResync on a volume.

- 3 On the primary host, verify that FastResync is enabled on the volume:

```
# vxprint -g diskgroup -F%fastresync volume_name
```

This command returns `on` indicating that FastResync is enabled.

- 4 Create a snapshot mirror of the volume.

Note: Ensure that the mirror is created on a disk with no other volumes.

```
# vxsnap -g diskgroup addmir volume_name alloc=disk_name
```

- 5 Prepare the secondary host to receive the snapshot volume that contains the copy of the database tables. This may involve setting up private volumes to contain any redo logs, and configuring any files that are used to initialize the database.
- 6 List the plexes and determine the name of the snapshot plex. The snapshot plex appears directly after the snapshot volume:

```
# vxprint -g diskgroup
```

- 7 On the primary host, suspend I/O updates by using the `quiesce database` command:

```
$ isql -Usa -Ppassword -Sserver
> quiesce database tag1 hold userdb, sybssystemdb, \
sybssystemprocs for external dump
> go
> quiesce database tag2 hold master for external dump
> go
> quit
```

Note: The `for external dump` clause is new in ASE 12.5 and it tells the ASE that a physical copy of the database device(s) will be made during the quiesce state, and that the copy will serve as the foundation for a new dump sequence. A dump made with this clause can be used for starting a secondary server and the secondary server can be rolled forward with transaction logs.

The `for external dump` clause has no effect on system databases.

- 8 Create a full-sized instant snapshot volume on the primary host by specifying the snapshot plex identified in 6:

```
# vxsnap -g diskgroup make \  
source=volume_name/newvol=snapshot_volume/plex=plex_name
```

By default, PRODUCTNAME attempts to avoid placing snapshot mirrors on a disk that already holds any plexes of a data volume. However, this may be impossible if insufficient space is available in the disk group. In this case, PRODUCTNAME uses any available space on other disks in the disk group. If the snapshot plexes are placed on disks that are used to hold the plexes of other volumes, this may cause problems when you subsequently attempt to move a snapshot volume into another disk group. To override the default storage allocation policy, you can use storage attributes to specify explicitly which disks to use for the snapshot plexes.

See the *Veritas Volume Manager Administrator's Guide*.

If a database spans more than one volume, you can specify all the volumes and their snapshot volumes using one command, as shown here:

```
# vxsnap -g diskgroup make \  
source=volume1/newvol=snapshot_volume1/nmirror=2 \  
source=volume2/newvol=snapshot_volume2/nmirror=2 \  
source=volume3/newvol=snapshot_volume3/nmirror=2
```

Note: This step sets up the snapshot volumes, and starts tracking changes to the original volumes. When you are ready to create a clone database, proceed to the next step.

- 9 On the primary host, resume I/O updates on the primary database. To resume the updates, release the databases from quiesce mode:

```
# vxdg split diskgroup new_diskgroup snapshot_volume
```

- 10 On the primary host, move the snapshot volume to another disk group from the original disk group:

```
# vxdg split diskgroup new_diskgroup snapshot_volume
```

The `split` (move) command will fail if the result of the move causes disks to be shared by two disk groups.

- 11** On the primary host, deport the snapshot volume's disk group:

```
# vxdg deport new_diskgroup
```

- 12** On the secondary host, import the snapshot volume's disk group:

```
# vxdg import new_diskgroup
```

- 13** After the split, the snapshot volume is initially disabled. Use the following commands on the secondary host to recover and restart the snapshot volume:

```
# vxrecover -g new_diskgroup -m snapshot_volume  
# vxvol -g new_diskgroup start snapshot_volume
```

- 14** On the secondary host, check and mount all of the file systems used by the ASE servers. The following are the commands for checking and mounting a file system:

```
# fsck -F vxfs /dev/vx/dsk/new_diskgroup/snapshot_volume  
# mount -F vxfs /dev/vx/dsk/new_diskgroup/snapshot_volume \  
/mount_point1
```

- 15** Change the ownership of the mount point specified in step 14 to this instance:

```
# chown -R sybase:sybase mount_point
```

Ensure that redo logs are part of the snapshot volume so that Sybase can perform a crash recovery if needed.

- 16** Create a server on the secondary host:

```
$ server_name
```

- 17** Modify the `interfaces` file on the secondary host to change the host name from the primary to the secondary host. The `interfaces` file is located under the `/$SYBASE` directory.

- 18** Start the ASE server on the secondary host with the ASE 12.5 `-q` flag:

```
$ /sybase/ASE-12_5/bin/dataserver -Sserver_name \  
-d/master_device_path -e/error_log_file_path \  
-M/sybase_software_path -q
```

- 19** After transactions have occurred on the primary host, deport the snapshot volume's disk group on the secondary host:

```
$ isql -Usa -Ppassword -Sserver
> dump tran userdb to dump_device with standby_access
> go
> quit
```

- 20** On the primary host, import the snapshot volume's disk group:

```
# vxdg split diskgroup dump_diskgroup snapshot_dump_volume
```

- 21** Join new_diskgroup back to the original disk group:

```
# vxdg join new_diskgroup diskgroup
```

- 22** Recover the snapshot volume:

```
# vxrecover -g new_diskgroup
```

- 23** Dump the transactions to the dump device (dump_diskgroup):

```
$ isql -Usa -Ppassword -Sserver
> dump tran userdb to dump_diskgroup with standby_access
> go
> quit
```

- 24** Deport the disk group dump_diskgroup on the primary host:

```
# vxdg deport dump_diskgroup
```

- 25** Import the disk group dump_diskgroup on the secondary host:

```
# vxdg import dump_diskgroup
```

- 26** On the secondary host, recover and restart the snapshot volume:

```
# vxrecover -g dump_diskgroup -m snapshot_volume
# vxvol -g dump_diskgroup start snapshot_volume
```

- 27** Load the transaction dumps to the failover database and put the database on the secondary host online:

```
$ isql -Usa -Ppassword -Sserver
> load tran userdb from dump_device
> go
> online database userdb for standby_access
> go
```

- 28** After the load is finished, `deport dump_diskgroup` and import it back to the primary host disk group. It is recommended that [19](#) to [27](#) be repeated regularly (for example, every hour).

- 29** To refresh all ASE servers, join new_diskgroup back to the original disk group and restart all volume snapshots:

```
# vxdg join new_diskgroup diskgroup  
# vxrecover -g diskgroup
```

30 Resynchronize the snapshot volume with the original volume:

```
# vxsnap -g diskgroup reattach snapshot_volume source=volume
```

The entire cycle can be repeated every day, if desired.

In this example, a warm standby dataserer is created in an ASE 12.5 environment on a secondary host.

Log in as root and create the disk group, `syb`, for database use.

```
# vxdg init syb disk1=c2t130d0s2
# vxdg -g syb adddisk disk2=c2t131d0s2
```

Log in as the instance owner and create a volume and VxFS file system.

```
# vxassist -g syb make sybvoll 5g disk1
# mkfs -F vxfs -o largefiles /dev/vx/rdsk/syb/sybvoll
# mkdir /sybasedata
# mount -F vxfs -o largefiles /dev/vx/dsk/syb/sybvoll /sybasedata
# chown -R sybase:sybase /sybasedata
```

Create an instance, `inst1`.

```
# /sybase/ASE-12.5/bin/dataserver -d/sybasedata/snaptest/master \
-e/sybasedata/snaptest/errorlog -M/sybase/ASE-12.5/ -s/mysrv -q
```

Create an ASE dataserer on the VxFS file system.

```
# srvbuild -r srvbuild.mysrv.rs
```

Modify the `interfaces` file on the secondary host to change the host name from the primary to the secondary host. The `interfaces` file is located under the `/$SYBASE` directory.

Log in to the dataserer and create a database.

```
$ isql -Usa -P -Smysrv
> disk init
> name=db1
> physname=/sybasedata/testdb/testfile
> vdevno=3
> size=5m
> dsync=true
> go
> disk init
> name=db1log
> physname=/sybasedata/testdb/testlog
```

```
> vdemo=4
> size=3m
> dsync=true
> go
> create database dbtest1 on db1=3m log db1log=1m
> go
> create table dept ( chat# char(2), name char(2), talk decimal(5),
> go
> insert into dept values (ab, cd, 12345);
> go
```

Log in as root and prepare the volume, `sybvoll`, for taking a snapshot.

```
# vxsnap -g syb prepare sybvoll alloc=syb01
```

Verify that DCO and FastResync are enabled on the volume.

```
# vxprint -g syb -F%instant sybvoll
on
on
```

Create a snapshot mirror.

```
# vxsnap -g syb addmir sybvoll alloc=syb02
```

Create a full-sized instant snapshot, `snap_voll`, on the primary host.

```
# vxsnap -g syb make \
source=sybvoll/newvol=snap_voll/plex=sybvoll-02
```

Move the snapshot volume, `snap_voll`, into a separate disk group, `new_syb`, from the original disk group, `syb`.

```
# vxdg split syb new_syb snap_voll
```

Deport the snapshot volume's disk group on the primary host.

```
# vxdg deport new_syb
```

Import the snapshot volume's disk group on the secondary host.

```
# vxdg import new_syb
```

After the split, the snapshot volume is initially disabled. Recover and restart the snapshot volume.

```
# vxrecover -g new_syb -m snap_voll  
# vxvol -g new_syb start snap_voll
```

On the secondary host, check and mount the snapshot volume.

```
# mkdir /sybasedata  
# fsck -F vxfs /dev/vx/rdisk/new_syb/snap_voll  
# mount -F vxfs /dev/vx/dsk/new_syb/snap_voll /sybasedata  
# chown -R sybase:sybase /sybasedata
```

Log in as the instance owner.

```
# su - sybase
```

Start the ASE server on the secondary host with the ASE 12.5 -q flag:

```
$ /sybase/ASE-12.5/bin/dataserver -d/sybasedata/snaptest/maste \  
-e/sybasedata/snapshot/errorlog -M/sybasedata/ASE-12.5/ -s mysrv \  
-q  
> isql -Usa -P -Smysrv  
> use dbtest1  
> go
```

The warm standby dataserver is now ready to use.

Resynchronizing the snapshot to your ASE dataserver

When you want to refresh a clone database, you can resynchronize it with the original database. This is also known as refreshing the snapshot volume or merging the split snapshot image back to the current database image. After resynchronizing, the snapshot is ready to be used for backup or decision-support purposes.

There are two choices when resynchronizing the data in a volume:

- Resynchronizing the snapshot from the original volume. This option is explained in this section.
- Resynchronizing the original volume from the snapshot. This choice is known as reverse resynchronization. Reverse resynchronization may be necessary to restore a corrupted database or file system, or to implement upgrades to production software, and is usually much quicker than using alternative approaches such as full restoration from backup media.

See [“Recovering the database from a backup image”](#) on page 87.

When you want to resynchronize the snapshot volume's data with the primary database, you can refresh the snapshot plexes from the original volume as described below.

To resynchronize the snapshot image

- 1 On the secondary host, shut down the clone database.

```
> isql -Usa -P -Sserver  
> shutdown  
> go
```

- 2 Log in as root and unmount the file system.

```
# umount mount_point
```

- 3 On the secondary host, deport the snapshot volume's disk group:

```
# vxdg deport new_diskgroup
```

- 4 On the primary host, re-import the snapshot volume's disk group:

```
# vxdg import new_diskgroup
```

- 5 On the primary host, rejoin the snapshot volume's disk group with the original volume's disk group:

```
# vxdg join new_diskgroup diskgroup
```

- 6 After the join, the snapshot volume is initially disabled. Use the following commands on the primary host to recover and restart the snapshot volume:

```
# vxvol -g diskgroup startall
```

The contents of the snapshot volume are now ready to be refreshed.

7 Reattach the snapshot volume's disk group to the original volume's disk group:

```
# vxsnap -g diskgroup reattach <snapshot_volume> \  
source=volume
```

- 8 Follow step 10 through step 18 to split the snapshot volume from the disk group when synchronization is again complete.

The snapshot image is now ready to be re-used for backup or decision-support applications.

Example

On the secondary host:

Log in as the instance owner and shut down the clone database:

```
> isql -Usa -P -Smysrv
> shutdown
> go
```

Log in as root and unmount the file system:

```
# umount /sybasedata
```

As root, deport the snapshot volume's disk group:

```
# vxdg deport new_syb
```

On the primary host:

Log in as root and import the snapshot volume's disk group:

```
# vxdg import new_syb
```

As root, rejoin the snapshot volume's disk group to the original volume's disk group:

```
# vxdg join new_syb syb
```

As root, recover and restart the snapshot volume:

```
# vxrecover -g syb
# vxvol -g syb start snap_voll
```

As root, reattach the snapshot:

```
# vxsnap -g syb reattach snap_voll source=sybvoll
```

Split the snapshot volume into a separate disk group from the original disk group:

```
# vxdg split syb new_syb snap_voll
```

Note: The `split` (move) command will fail if the result of the move causes disks to be shared by two disk groups.

Deport the snapshot volume's disk group:

```
# vxdg deport new_syb
```

On the secondary host:

Log in as root and import the snapshot volume's disk group:

```
# vxdg import new_syb
```

After the split, the snapshot volume is initially disabled. As root, use the following commands on to recover and restart the snapshot volume:

```
# vxrecover -g syb -m snap_voll
```

As root, check and mount all of the file systems used by the ASE servers. The following are the commands for checking and mounting a file system:

```
# fsck -F vxfs /dev/vx/dsk/syb/snap_voll
# mount -F vxfs /dev/vx/dsk/syb/snap_voll /sybasedata
```

Ensure that the dataserver is up and running on the secondary host.

As root, change the ownership of the mount point `sybasedata`, if needed:

```
# chown -R sybase:sybase sybasedata
```

Note: Ensure that redo logs are part of the snapshot volume so that Sybase can perform a crash recovery if needed.

Modify the `interfaces` file on the secondary host to change the host name from the primary to the secondary host. The `interfaces` file is located under the `/$SYBASE` directory.

Start the ASE server on the secondary host with the ASE 12.5 `-q` flag:

```
$ /sybase/ASE-12.5/bin/dataserver -d/sybasedata/snaptest/master \
-e/sybasedata/snapshot/errorlog -M/sybasedata/ASE-12.5/ -s mysrv \
-q
> isql -Usa -P -Smysrv
> use dbtest1
> go
```

Recovering the database from a backup image

A backup image on the primary host can be used to restore the primary database if it becomes corrupted. This section describes the procedure for using snapshot volumes to create a backup image of a database on the primary host. It also explains how to recover the primary database from the backup image.

- | | |
|---------------|--|
| Prerequisites | <ul style="list-style-type: none">■ You must be logged in as superuser (root).■ The disk group must be version 90 or later. For more information on disk group versions, see the <code>vxddg(1M)</code> manual page. |
| Usage notes | <ul style="list-style-type: none">■ Ensure that your Sybase instance owner directory is on a volume included in the snapshot.■ Create a separate disk group for Sybase database-related files.■ Do not share file systems and volumes between two ASE servers.■ Do not share the same disk group between two ASE servers.■ If you intend to move a snapshot volume to another host for off-host processing, do not share any disks between the original mirror plex and the snapshot mirror.■ Create snapshot mirrors on a separate controller and separate disk from the primary volume.■ Create snapshot mirrors for datafiles and archive logs so that they do not share any disks with the data of the original volumes. If they are not created in this way, they cannot be split and moved to a secondary host.■ Allocate separate volumes for logs and do not create snapshots on those volumes.■ Resynchronization speed varies based on the amount of data changed in both the primary and secondary volumes during the break-off time. |

To create a backup image on the primary host

- 1 Follow steps 1 to 4 and then steps 6 to 8 in:

See [“Resynchronizing the snapshot to your ASE dataserver”](#) on page 82.

to create a snapshot volume on the primary host.

A failure occurs, requiring you to restore the primary database from the backup image.

- 2 Shut down the primary database if it is still active.

```
> isql -Usa -P -Sserver
> shutdown
> go
```

3 Unmount the file system for the primary database

```
# umount -f /mount_point
```

4 Restore the volume from the snapshot volume. You must be logged in as root to execute this command.

```
# vxsnap -g diskgroup restore volume_name \  
source=snapshot_volume destroy=yes
```

5 Check and mount the file system for the primary database.

```
# fsck -F vxfs /dev/vx/dsk/diskgroup/volume  
# mount -F vxfs /dev/vx/dsk/diskgroup/volume /mount_point
```

6 Ensure that the ASE dataserer has the correct permissions:

```
# chown -R sybase:sybase /mount_point
```

7 As the instance owner, start the primary database.

```
> isql -Usa -P -Sserver
```

- 8 Bring the backup image of the database online by applying the transaction logs and roll it forward.

The Sybase clients are now ready to reconnect.

Note: The archive logs are needed to roll the database forward. In case the archive logs are corrupted on the primary database, they need to be restored from backup.

In this example the primary database is restored from a backup image on the primary host.

After the corruption occurs, log in as the instance owner and terminate all active connections to the primary database:

```
> isql -Usa -P -Smysrv
> shutdown
> go
```

Log in as root and unmount the primary database's file systems:

```
# umount -f /sybasedata
```

As root, join the snapshot volume's disk group with the primary disk group:

```
# vxdg join new_syb syb
```

After the join, the snapshot volume is initially disabled. Recover and restart the snapshot volume:

```
# vxrecover -g syb -m snap_voll
```

As root, restore the volume from the snapshot volume:

```
# vxsnap -g syb restore sybvoll source=snap_voll destroy=yes
```

As root, check and mount the file system for the primary database:

```
# fsck -F vxfs /dev/vx/rdisk/syb/sybvoll
# mount -F vxfs /dev/vx/dsk/syb/sybvoll /sybasedata
```

Ensure that the ASE dataserer has the correct permissions:

```
# chown -R sybase:sybase /sybasedata
```

Log in as the instance owner and start the primary database:

```
# su - sybase
$ /sybase/ASE-12.5/install/startserver -f RUN_mysrv
```

The database has now been restored to where it was when the snapshot was taken or refreshed last.

Refreshing a snapshot database image

Refresh a snapshot database image so that it can later be used to restore the primary database.

Before refreshing a snapshot database image, make sure the following conditions have been met:

- | | |
|---------------|--|
| Prerequisites | <ul style="list-style-type: none">■ You must be logged in as superuser (root).■ The disk group must be version 90 or later. For more information on disk group versions, see the <code>vxdg(1M)</code> manual page. |
| Usage notes | <p>When creating snapshot mirrors used by databases refer to the following:</p> <ul style="list-style-type: none">■ Ensure that your Sybase instance owner directory is on a volume included in the snapshot.■ Create a separate disk group for Sybase database-related files.■ Do not share file systems and volumes between two ASE servers.■ Do not share the same disk group between two ASE servers.■ If you intend to move a snapshot volume to another host for off-host processing, do not share any disks between the original mirror plex and the snapshot mirror.■ Create snapshot mirrors on a separate controller and separate disk from the primary volume.■ Create snapshot mirrors for datafiles and archive logs so that they do not share any disks with the data of the original volumes. If they are not created in this way, they cannot be split and moved to a secondary host.■ Allocate separate volumes for logs and do not create snapshots on those volumes.■ Resynchronization speed varies based on the amount of data changed in both the primary and secondary volumes during the break-off time. |

To refresh a snapshot image of the database on the primary host

- 1** As root, join the snapshot disk group with the primary disk group:

```
# vxdg join new_diskgroup diskgroup
```

- 2** After the join, the snapshot volume is initially disabled. Recover and restart the snapshot volume:

```
# vxrecover -g diskgroup -m snapshot_volume
```

- 3** Suspend I/O updates on the primary database by using the `quiesce database` command:

```
$ isql -Usa -Ppassword -Sserver  
> quiesce database tag1 hold testdb, sybssystemdb, \  
sybssystemprocs  
> go  
> quiesce database tag2 hold master  
> go  
> quit
```

- 4** Refresh the snapshot image:

```
$ vxsnap -g diskgroup refresh snapshot_volume \  
source=volume_name syncing=yes
```

- 5** Resume I/O updates on the primary database. To resume the updates, release the databases from quiesce mode:

```
$ isql -Usa -Ppassword -Sserver  
> quiesce database tag2 release  
> go  
> quiesce database tag1 release  
> go  
> quit
```

- 6** On the primary host, split the snapshot volume into a separate disk group from the original disk group:

```
# vxdg split diskgroup new_diskgroup snapshot_volume
```

Dissociating a snapshot volume

You can permanently break the link between a snapshot and its original volume so that the snapshot volume becomes an independent volume.

- You must be logged in as superuser (root).
- The snapshot volume must be associated with its original volume.

You can permanently break the link between a snapshot and its original volume so that the snapshot volume becomes an independent volume.

To dissociate a snapshot from its original volume

- 1 Use the `vxsnap dis` command as follows:

```
# vxsnap -g <diskgroup> [-f] dis <snapshot_volume>
```

- 2 Remove the DCO logs from the snapshot volume:

```
# vxsnap -g <diskgroup> [-f] unprepare <snapshot_volume>
```

- 3 Remove the DCO logs from the primary volume, if necessary:

```
# vxsnap -g <diskgroup> [-f] unprepare <volume_name>
```

This example shows how to dissociate a snapshot volume from its original volume:

```
# vxsnap -g syb dis snap_vol
```

This example shows how to remove the DCO logs from the snapshot volume:

```
# vxsnap -g syb -f unprepare snap_vol
```

This example shows how to remove the DCO logs from the primary volume:

```
# vxsnap -g syb -f unprepare sybvol
```

Removing a snapshot volume

If a volume is no longer necessary, you can remove the volume and free up the disk space for other uses by using the `vxedit rm` command.

Note: Removing a volume destroys all of the data in that volume. After a volume is removed, the space it occupied is returned to the free space pool.

Before removing a snapshot volume, make sure the following requirements have been met:

- Prerequisites
- You must be logged in as superuser (root).
 - You must have an existing snapshot volume.

To remove a snapshot volume using the command line

- 1 Dissociate the snapshot from its original volume:

```
# vxsnap -g diskgroup [-f] dis snapshot_volume
```

- 2 Remove the snapshot volume:

```
# vxedit -g diskgroup -rf rm snapshot_volume
```

where `-r` recursively removes all plexes and subdisks.

Warning: If the volume is on a mounted file system, you must unmount it before removing the volume.

This example shows how to remove a snapshot volume from a disk group:

```
# vxsnap -g syb dis snap_voll  
# vxedit -g syb -rf rm snap_voll
```

Tuning for performance

This chapter includes the following topics:

- [Additional documentation](#)
- [About tuning VxFS](#)
- [About tuning VxVM](#)
- [About tuning Sybase dataservers](#)

Additional documentation

Use the tuning tips and information provided in this chapter in conjunction with other more in-depth publications, such as:

- *Sybase Adaptive Server Enterprise Performance and Tuning Guide*— covers general tuning tips
- Other generic Sybase documentation that deals with Sybase tuning issues
- *Veritas Volume Manager Administrator's Guide*, chapter on “VxVM Performance Monitoring”

About tuning VxFS

Veritas File System provides a set of tuning options to optimize file system performance for different application workloads. VxFS provides a set of tunable I/O parameters that control some of its behavior. These I/O parameters help the file system adjust to striped or RAID-5 volumes that could yield performance far superior to a single disk. Typically, data streaming applications that access large files see the largest benefit from tuning the file system.

How monitoring free space works

In general, VxFS works best if the percentage of free space in the file system is greater than 10 percent. This is because file systems with 10 percent or more of free space have less fragmentation and better extent allocation. Regular use of the `df` command to monitor free space is desirable. Full file systems may have an adverse effect on file system performance. Full file systems should therefore have some files removed or should be expanded.

See the `fsadm_vxfs(1M)` manual page.

About monitoring fragmentation

Fragmentation reduces performance and availability. Regular use of `fsadm`'s fragmentation reporting and reorganization facilities is therefore advisable.

The easiest way to ensure that fragmentation does not become a problem is to schedule regular defragmentation runs using the `cron` command.

Defragmentation scheduling should range from weekly (for frequently used file systems) to monthly (for infrequently used file systems). Extent fragmentation should be monitored with `fsadm` command.

There are three factors that can be used to determine the degree of fragmentation:

- Percentage of free space in extents that are less than eight blocks in length
- Percentage of free space in extents that are less than 64 blocks in length
- Percentage of free space in extents that are 64 or more blocks in length

An unfragmented file system will have the following characteristics:

- Less than 1 percent of free space in extents that are less than eight blocks in length
- Less than 5 percent of free space in extents that are less than 64 blocks in length
- More than 5 percent of the total file system size available as free extents that are 64 or more blocks in length

A badly fragmented file system will have one or more of the following characteristics:

- More than 5 percent of free space in extents that are less than 8 blocks in length
- More than 50 percent of free space in extents that are less than 64 blocks in length

- Less than 5 percent of the total file system size available as free extents that are 64 or more blocks in length

The optimal period for scheduling extent reorganization runs can be determined by choosing a reasonable interval, scheduling `fsadm` runs at the initial interval, and running the extent fragmentation report feature of `fsadm` before and after the reorganization.

The “before” result is the degree of fragmentation prior to the reorganization. If the degree of fragmentation approaches the percentages for bad fragmentation, reduce the interval between `fsadm`. If the degree of fragmentation is low, increase the interval between `fsadm` runs.

How tuning VxFS I/O parameters works

VxFS provides a set of tunable I/O parameters that control some of its behavior. These I/O parameters are useful to help the file system adjust to striped or RAID-5 volumes that could yield performance far superior to a single disk. Typically, data streaming applications that access large files see the biggest benefit from tuning the file system.

If VxFS is being used with Veritas Volume Manager, the file system queries VxVM to determine the geometry of the underlying volume and automatically sets the I/O parameters. VxVM is queried by `mkfs` when the file system is created to automatically align the file system to the volume geometry. If the default alignment from `mkfs` is not acceptable, the `-o align=n` option can be used to override alignment information obtained from VxVM. The `mount` command also queries VxVM when the file system is mounted and downloads the I/O parameters.

If the default parameters are not acceptable or the file system is being used without VxVM, then the `/etc/vx/tunefstab` file can be used to set values for I/O parameters. The `mount` command reads the `/etc/vx/tunefstab` file and downloads any parameters specified for a file system. The `tunefstab` file overrides any values obtained from VxVM. While the file system is mounted, any I/O parameters can be changed using the `vxtunefs` command, which can have tunables specified on the command line or can read them from the `/etc/vx/tunefstab` file.

The `vxtunefs` command can be used to print the current values of the I/O parameters.

See the `vxtunefs(1M)` and `tunefstab(4)` manual pages.

About tunable VxFS I/O parameters

The following are tunable VxFS I/O parameters:

<code>read_pref_io</code>	The preferred read request size. The file system uses this parameter in conjunction with the <code>read_nstream</code> value to determine how much data to read ahead. The default value is 64K.
<code>write_pref_io</code>	The preferred write request size. The file system uses this parameter in conjunction with the <code>write_nstream</code> value to determine how to do flush behind on writes. The default value is 64K.
<code>read_nstream</code>	The number of parallel read requests of size <code>read_pref_io</code> that you can have outstanding at one time. The file system uses the product of <code>read_nstream</code> multiplied by <code>read_pref_io</code> to determine its read ahead size. The default value for <code>read_nstream</code> is 1.
<code>write_nstream</code>	The number of parallel write requests of size <code>write_pref_io</code> that you can have outstanding at one time. The file system uses the product of <code>write_nstream</code> multiplied by <code>write_pref_io</code> to determine when to do flush behind on writes. The default value for <code>write_nstream</code> is 1.
<code>discovered_direct_iosz</code>	Any file I/O requests larger than the <code>discovered_direct_iosz</code> are handled as discovered direct I/O. A discovered direct I/O is unbuffered similar to direct I/O, but does not require a synchronous commit of the inode when the file is extended or blocks are allocated. For larger I/O requests, the CPU time for copying the data into the page cache and the cost of using memory to buffer the I/O data becomes more expensive than the cost of doing the disk I/O. For these I/O requests, using discovered direct I/O is more efficient than regular I/O. The default value of this parameter is 256K.

<code>initial_extent_size</code>	Changes the default initial extent size. VxFS determines the size of the first extent to be allocated to the file based on the first write to a new file. Normally, the first extent is the smallest power of 2 that is larger than the size of the first write. If that power of 2 is less than 8K, the first extent allocated is 8K. After the initial extent, the file system increases the size of subsequent extents (see <code>max_seqio_extent_size</code>) with each allocation. Since most applications write to files using a buffer size of 8K or less, the increasing extents start doubling from a small initial extent. <code>initial_extent_size</code> can change the default initial extent size to be larger, so the doubling policy will start from a much larger initial size and the file system will not allocate a set of small extents at the start of file. Use this parameter only on file systems that will have a very large average file size. On these file systems, it will result in fewer extents per file and less fragmentation. <code>initial_extent_size</code> is measured in file system blocks.
<code>max_direct_iosz</code>	The maximum size of a direct I/O request that will be issued by the file system. If a larger I/O request comes in, then it is broken up into <code>max_direct_iosz</code> chunks. This parameter defines how much memory an I/O request can lock at once, so it should not be set to more than 20 percent of memory.
<code>max_diskq</code>	Limits the maximum disk queue generated by a single file. When the file system is flushing data for a file and the number of pages being flushed exceeds <code>max_diskq</code> , processes will block until the amount of data being flushed decreases. Although this doesn't limit the actual disk queue, it prevents flushing processes from making the system unresponsive. The default value is 1MB.

`max_seqio_extent_size` Increases or decreases the maximum size of an extent. When the file system is following its default allocation policy for sequential writes to a file, it allocates an initial extent that is large enough for the first write to the file. When additional extents are allocated, they are progressively larger (the algorithm tries to double the size of the file with each new extent) so each extent can hold several writes' worth of data. This is done to reduce the total number of extents in anticipation of continued sequential writes. When the file stops being written, any unused space is freed for other files to use. Normally, this allocation stops increasing the size of extents at 2048 blocks, which prevents one file from holding too much unused space. `max_seqio_extent_size` is measured in file system blocks.

`write_throttle`

Warning: The `write_throttle` parameter is useful in special situations where a computer system has a combination of a lot of memory and slow storage devices. In this configuration, sync operations (such as `fsync()`) may take so long to complete that the system appears to hang. This behavior occurs because the file system is creating dirty pages (in-memory updates) faster than they can be asynchronously flushed to disk without slowing system performance.

Lowering the value of `write_throttle` limits the number of dirty pages per file that a file system will generate before flushing the pages to disk. After the number of dirty pages for a file reaches the `write_throttle` threshold, the file system starts flushing pages to disk even if free memory is still available. The default value of `write_throttle` typically generates a lot of dirty pages, but maintains fast user writes. Depending on the speed of the storage device, if you lower `write_throttle`, user write performance may suffer, but the number of dirty pages is limited, so sync operations will complete much faster.

Because lowering `write_throttle` can delay write requests (for example, lowering `write_throttle` may increase the file disk queue to the `max_diskq` value, delaying user writes until the disk queue decreases), it is recommended that you avoid changing the value of `write_throttle` unless your system has a large amount of physical memory and slow storage devices.

If the file system is being used with VxVM, it is recommended that you set the VxFS I/O parameters to default values based on the volume geometry.

If the file system is being used with a hardware disk array or volume manager other than VxVM, align the parameters to match the geometry of the logical disk. With striping or RAID-5, it is common to set `read_pref_io` to the stripe unit size and `read_nstream` to the number of columns in the stripe. For striping arrays, use the same values for `write_pref_io` and `write_nstream`, but for RAID-5 arrays, set `write_pref_io` to the full stripe size and `write_nstream` to 1.

For an application to do efficient disk I/O, it should issue read requests that are equal to the product of `read_nstream` multiplied by `read_pref_io`. Generally, any multiple or factor of `read_nstream` multiplied by `read_pref_io` should be a good size for performance. For writing, the same rule of thumb applies to the `write_pref_io` and `write_nstream` parameters. When tuning a file system, the best thing to do is try out the tuning parameters under a real-life workload.

If an application is doing sequential I/O to large files, it should issue requests larger than the `discovered_direct_iosz`. This causes the I/O requests to be performed as discovered direct I/O requests, which are unbuffered like direct I/O but do not require synchronous inode updates when extending the file. If the file is too large to fit in the cache, then using unbuffered I/O avoids throwing useful data out of the cache and lessons CPU overhead.

About tuning VxVM

Veritas Volume Manager (VxVM) is tuned for most configurations ranging from small systems to larger servers. On smaller systems with less than a hundred drives, tuning should not be necessary and Veritas Volume Manager should be capable of adopting reasonable defaults for all configuration parameters. On very large systems, however, there may be configurations that require additional tuning of these parameters, both for capacity and performance reasons.

For more information on tuning VxVM, see the *Veritas Volume Manager Administrator's Guide*.

About obtaining volume I/O statistics

If your database is created on a single file system that is on a single volume, there is typically no need to monitor the volume I/O statistics. If your database is created on multiple file systems on multiple volumes, or the volume configurations have changed over time, it may be necessary to monitor the volume I/O statistics for the databases.

Use the `vxstat` command to access information about activity on volumes, plexes, subdisks, and disks under VxVM control, and to print summary statistics to the standard output. These statistics represent VxVM activity from the time the

system initially booted or from the last time the counters were reset to zero. If no VxVM object name is specified, statistics from all volumes in the configuration database are reported. Use the `-g` option to specify the database disk group to report statistics for objects in that database disk group.

VxVM records the following I/O statistics:

- count of operations
- number of blocks transferred (one operation can involve more than one block)
- average operation time (which reflects the total time through the VxVM interface and is not suitable for comparison against other statistics programs)

VxVM records the preceding three pieces of information for logical I/Os, including reads, writes, atomic copies, verified reads, verified writes, plex reads, and plex writes for each volume. VxVM also maintains other statistical data such as read failures, write failures, corrected read failures, corrected write failures, and so on. In addition to displaying volume statistics, the `vxstat` command is capable of displaying more detailed statistics on the components that form the volume. For detailed information on available options, refer to the `vxstat(1M)` manual page.

To reset the statistics information to zero, use the `-r` option. You can reset the statistics information for all objects or for only those objects that are specified. Resetting just prior to an operation makes it possible to measure the impact of that particular operation.

The following is an example of output produced using the `vxstat` command:

OPERATIONS		BLOCKS		AVG TIME (ms)			
TYP	NAME	READ	WRITE	READ	WRITE	READ	WRITE
vol	log2	0	6312	0	79836	.0	0.2
vol	db02	2892318	3399730	0283759	7852514	20.6	25.5

Additional information is available on how to use the `vxstat` output to identify volumes that have excessive activity and how to reorganize, change to a different layout, or move these volumes.

Additional volume statistics are available for RAID-5 configurations.

See the `vxstat(1M)` manual page.

See the “Performance Monitoring” section of the “Performance Monitoring and Tuning” chapter in the *Veritas Volume Manager Administrator's Guide*.

About tuning Sybase dataservers

To achieve optimal performance on your Sybase dataserver, the server may need to be tuned to work together with VxFS. This section lists some general suggestions.

Sybase tempdb database

Sybase `tempdb` is used quite frequently so it should be placed on a separate file system mounting on a dedicated volume. The volume should be striped and its disks should not be shared with other high activity volumes. This database should also bind to its own cache space with the Sybase ASE-named cache feature to reduce paging.

The `tempdb` database needs to be large enough to contain all the work tables and temporary tables created by the dataserver. When the Adaptive Server is installed, `tempdb` is created entirely on the master device. The database administrator needs to move `tempdb` on to larger, dedicated devices. (The default size is 2 MB only.)

To do so, first alter `tempdb` onto the new device. By default, the master device is included in `tempdb`'s `logsegment` and `defaultsegment`. To have control on the placement of the log segment and default segment, you need to drop those segments from the master device as shown in the example below.

See the `tempdb` performance chapter in the *Sybase ASE Performance and Tuning Guide*.

To change tempdb to a dedicated 200MB device

- ◆ Execute these commands on the Sybase Adaptive Server:

```
$ isql -Usa -P<sa_password> -S<dataserver_name>
> disk init
> name="newtempdb",
> physname="/newtempdb_dev",
> vnevno=<next_available_number>,
> size=102400
> go
> alter database tempdb on newtempdb=200
> go
> sp_dropsegment "default", tempdb, master
> go
> sp_dropsegment logsegment, tempdb, master
> go
```

Work tables and other temporary tables in `tempdb` will now be created on the device `newtempdb` instead of on the `tempdb` master device.

Sybase sybsecurity database

If you use auditing on your dataserver, the auditing system performs frequent input and output to the `sysaudits` table in the `sybsecurity` database. Follow the same recommendation on the placement of this database as that for the `tempdb`.

Placement of the transaction logs

You should place the transaction log on a separate volume to reduce contentions. Do not put log devices and others database devices on the same file system. For log devices, you should use mirroring instead of RAID-5 for high availability.

Database device layout

Stripe across as many disk drives as possible. For heavily updated tables, use mirroring for high availability instead of RAID-5. Use user-defined segments to achieve the exact placements for your database objects.

Nonclustered indexes placement

Data are usually being accessed at the same time the nonclustered indexes are accessed. To reduce contention, you should separate the data and their nonclustered indexes.

Glossary

address-length pair	Identifies the starting block address and the length of an extent (in file system or logical blocks).
asynchronous I/O	A format of I/O that performs non-blocking reads and writes. This enables the system to handle multiple I/O requests simultaneously.
atomic operation	An operation that either succeeds completely or fails and leaves everything as it was before the operation was started. If the operation succeeds, all aspects of the operation take effect at once and the intermediate states of change are invisible. If any aspect of the operation fails, then the operation aborts without leaving partial changes.
block map	A file system is divided into fixed-size blocks when it is created. As data is written to a file, unused blocks are allocated in ranges of blocks, called extents. The extents are listed or pointed to from the inode. The term used for the data that represents how to translate an offset in a file to a file system block is the “block map” for the file.
boot disk	A disk used for booting an operating system.
buffered I/O	A mode of I/O operation (where I/O is any operation, program, or device that transfers data to or from a computer) that first transfers data into the Operating System buffer cache.
cache	Any memory used to reduce the time required to respond to an I/O request. The read cache holds data in anticipation that it will be requested by a client. The write cache holds data written until it can be safely stored on non-volatile storage media.
cluster	A set of hosts that share a set of disks.
cluster-shareable disk group	A disk group in which the disks are shared between more than one host.
cold backup	The process of backing up a database that is not in active use.
command launcher	A graphical user interface (GUI) window that displays a list of tasks that can be performed by Veritas Volume Manager or other objects. Each task is listed with the object type, task (action), and a description of the task. A task is launched by clicking on the task in the Command Launcher. concatenation A Veritas Volume Manager layout style characterized by subdisks that are arranged sequentially and contiguously.

concurrent I/O	A form of Direct I/O that does not require file-level write locks when writing to a file. Concurrent I/O allows the relational database management system (RDBMS) to write to a given file concurrently.
configuration database	A set of records containing detailed information on existing Veritas Volume Manager objects (such as disk and volume attributes). A single copy of a configuration database is called a configuration copy.
copy-on-write	A technique for preserving the original of some data. As data is modified by a write operation, the original copy of data is copied.
database	A database is a collection of information that is organized in a structured fashion. Two examples of databases are Relational Databases (such as Oracle, Sybase, or DB2), where data is stored in tables and generally accessed by one or more keys and Flat File Databases, where data is not generally broken up into tables and relationships. Databases generally provide tools and/or interfaces to retrieve data.
dataserver	A logical concept of a Sybase instance. A Sybase instance contains databases and daemon processes that manage the data. A Sybase dataserver manages Sybase system databases and user created databases. Each Sybase dataserver is uniquely named when it is created.
Decision Support Systems	Decision Support Systems (DSS) are computer-based systems used to model, identify, and solve problems, and make decisions.
defragmentation	The act of reorganizing data to reduce fragmentation. Data in file systems become fragmented over time.
device file	A block- or character-special file located in the /dev directory representing a device.
device name	The name of a device file, which represents a device. AIX syntax is Disk_#; HP-UX syntax is c#t#d#; Linux syntax is sda, where "a" could be any alphabetical letter; Solaris syntax is c#t#d#s#.
direct I/O	An unbuffered form of I/O that bypasses the kernel's buffering of data. With direct I/O, data is transferred directly between the disk and the user application.
Dirty Region Logging	The procedure by which the Veritas Volume Manager monitors and logs modifications to a plex. A bitmap of changed regions is kept in an associated subdisk called a log subdisk.
disk access name	The name used to access a physical disk, such as Disk_1 on an AIX system, c1t1d1 on an HP-UX system, sda on a Linux system, or c0t0d0s0 on a Solaris system. The term device name can also be used to refer to the disk access name.
disk array	A collection of disks logically and physically arranged into an object. Arrays provide benefits including data redundancy and improved performance.
disk cache	A section of RAM that provides a cache between the disk and the application. Disk cache enables the computer to operate faster. Because retrieving data from hard

disk can be slow, a disk caching program helps solve this problem by placing recently accessed data in the disk cache. Next time that data is needed, it may already be available in the disk cache; otherwise a time-consuming operation to the hard disk is necessary.

disk group	<p>A collection of disks that share a common configuration.</p> <p>A disk group configuration is a set of records containing detailed information on existing Veritas Volume Manager objects (such as disk and volume attributes) and their relationships. Each disk group has an administrator-assigned name and an internally defined unique ID. The root disk group (rootdg) is a special private disk group</p>
disk name	<p>A Veritas Volume Manager logical or administrative name chosen for the disk, such as disk03. The term disk media name is also used to refer to the disk name.</p>
DMP	<p>See “Dynamic Multipathing.”</p>
DSS	<p>See “Decision Support Systems.”</p>
Dynamic Multipathing	<p>Dynamic Multipathing (DMP) is a Veritas Volume Manager feature that allows the use of multiple paths to the same storage device for load balancing and redundancy.</p>
error handling	<p>Routines in a program that respond to errors. The measurement of quality in error handling is based on how the system informs the user of such conditions and what alternatives it provides for dealing with them.</p>
evacuate	<p>Moving subdisks from the source disks to target disks.</p>
exabyte	<p>A measure of memory or storage. An exabyte is approximately 1,000,000,000,000,000 bytes (technically 2 to the 60th power, or 1,152,921,504,606,846,976 bytes). Also EB.</p>
extent	<p>A logical database attribute that defines a group of contiguous file system data blocks that are treated as a unit. An extent is defined by a starting block and a length.</p>
extent attributes	<p>The extent allocation policies associated with a file and/or file system. For example, see “address-length pair.”</p>
failover	<p>The act of moving a service from a failure state back to a running/available state. Services are generally applications running on machines and failover is the process of restarting these applications on a second system when the first has suffered a failure.</p>
file system	<p>A collection of files organized together into a structure. File systems are based on a hierarchical structure consisting of directories and files.</p>
file system block	<p>The fundamental minimum size of allocation in a file system.</p>
fileset	<p>A collection of files within a file system.</p>

fixed extent size	An extent attribute associated with overriding the default allocation policy of the file system.
fragmentation	Storage of data in non-contiguous areas on disk. As files are updated, new data is stored in available free space, which may not be contiguous. Fragmented files cause extra read/write head movement, slowing disk accesses.
gigabyte	A measure of memory or storage. A gigabyte is approximately 1,000,000,000 bytes (technically, 2 to the 30th power, or 1,073,741,824 bytes). Also GB, Gbyte, and G-byte.
high availability (HA)	The ability of a system to perform its function continuously (without significant interruption) for a significantly longer period of time than the combined reliabilities of its individual components. High availability is most often achieved through failure tolerance and inclusion of redundancy; from redundant disk to systems, networks, and entire sites.
hot backup	The process of backing up a database that is online and in active use.
hot pluggable	To pull a component out of a system and plug in a new one while the power is still on and the unit is still operating. Redundant systems can be designed to swap disk drives, circuit boards, power supplies, CPUs, or virtually anything else that is duplexed within the computer. Also known as hot swappable.
inode list	An inode is an on-disk data structure in the file system that defines everything about the file, except its name. Inodes contain information such as user and group ownership, access mode (permissions), access time, file size, file type, and the block map for the data contents of the file. Each inode is identified by a unique inode number in the file system where it resides. The inode number is used to find the inode in the inode list for the file system. The inode list is a series of inodes. There is one inode in the list for every file in the file system.
intent logging	A logging scheme that records pending changes to a file system structure. These changes are recorded in an intent log.
interrupt key	A way to end or break out of any operation and return to the system prompt by pressing Ctrl-C.
kilobyte	A measure of memory or storage. A kilobyte is approximately a thousand bytes (technically, 2 to the 10th power, or 1,024 bytes). Also KB, Kbyte, kbyte, and K-byte.
large file	A file more than two gigabytes in size. An operating system that uses a 32-bit signed integer to address file contents will not support large files; however, the Version 4 disk layout feature of VxFS supports file sizes of up to two terabytes.
large file system	A file system more than two gigabytes in size. VxFS, in conjunction with VxVM, supports large file systems.

latency	The amount of time it takes for a given piece of work to be completed. For file systems, this typically refers to the amount of time it takes a given file system operation to return to the user. Also commonly used to describe disk seek times.
load balancing	The tuning of a computer system, network tuning, or disk subsystem in order to more evenly distribute the data and/or processing across available resources. For example, in clustering, load balancing might distribute the incoming transactions evenly to all servers, or it might redirect them to the next available server.
load sharing	The division of a task among several components without any attempt to equalize each component's share of the load. When several components are load sharing, it is possible for some of the shared components to be operating at full capacity and limiting performance, while others components are under utilized.
Logical Unit Number	A method of expanding the number of SCSI devices that can be placed on one SCSI bus. Logical Unit Numbers address up to seven devices at each SCSI ID on an 8-bit bus or up to 15 devices at each ID on a 16-bit bus.
logical volume	See "volume."
LUN	See "Logical Unit Number."
master node	A computer which controls another computer or a peripheral.
megabyte	A measure of memory or storage. A megabyte is approximately 1,000,000 bytes (technically, 2 to the 20th power, or 1,048,576 bytes). Also MB, Mbyte, mbyte, and K-byte.
metadata	Data that describes other data. Data dictionaries and repositories are examples of metadata. The term may also refer to any file or database that holds information about another database's structure, attributes, processing, or changes.
mirror	A duplicate copy of a volume and the data therein (in the form of an ordered collection of subdisks). Each mirror is one copy of the volume with which the mirror is associated. The terms mirror and plex can be used synonymously.
mirroring	A layout technique that mirrors the contents of a volume onto multiple plexes. Each plex duplicates the data stored on the volume, but the plexes themselves may have different layouts.
mount point	The directory path name at which a file system attaches to the file system hierarchy.
multithreaded	Having multiple concurrent or pseudo-concurrent execution sequences. Used to describe processes in computer systems. Multithreaded processes are one means by which I/O request-intensive applications can use independent access to volumes and disk arrays to increase I/O performance.
NBU	See "Veritas NetBackup (NBU)."
node	One of the hosts in a cluster.

object (VxVM)	An entity that is defined to and recognized internally by the Veritas Volume Manager. The VxVM objects include volumes, plexes, subdisks, disks, and disk groups. There are two types of VxVM disk objects—one for the physical aspect of the disk and the other for the logical aspect of the disk.
OLTP	See “Online Transaction Processing.”
online administration	An administrative feature that allows configuration changes without system or database down time.
Online Transaction Processing	A type of system designed to support transaction-oriented applications. OLTP systems are designed to respond immediately to user requests and each request is considered to be a single transaction. Requests can involve adding, retrieving, updating or removing data.
paging	The transfer of program segments (pages) into and out of memory. Although paging is the primary mechanism for virtual memory, excessive paging is not desirable.
parity	A calculated value that can be used to reconstruct data after a failure. While data is being written to a RAID-5 volume, parity is also calculated by performing an exclusive OR (XOR) procedure on data. The resulting parity is then written to the volume. If a portion of a RAID-5 volume fails, the data that was on that portion of the failed volume can be recreated from the remaining data and the parity.
partition	The logical areas into which a disk is divided.
persistence	Information or state that will survive a system reboot or crash.
petabyte	A measure of memory or storage. A petabyte is approximately 1,000 terabytes (technically, 2 to the 50th power).
plex	A duplicate copy of a volume and its data (in the form of an ordered collection of subdisks). Each plex is one copy of a volume with which the plex is associated. The terms mirror and plex can be used synonymously.
preallocation	Prespecifying space for a file so that disk blocks will physically be part of a file before they are needed. Enabling an application to preallocate space for a file guarantees that a specified amount of space will be available for that file, even if the file system is otherwise out of space.
RAID	A Redundant Array of Independent Disks (RAID) is a disk array set up with part of the combined storage capacity used for storing duplicate information about the data stored in that array. This makes it possible to regenerate the data if a disk failure occurs.
repository	A repository holds the name, type, range of values, source, and authorization for access for each data element in a database. The database maintains a repository for administrative and reporting use.
root disk	The disk containing the root file system.

root disk group	A special private disk group on the system. The root disk group is named rootdg. However, starting with the 4.1 release of Veritas Volume Manager, the root disk group is no longer needed.
root file system	The initial file system mounted as part of the UNIX kernel startup sequence.
script	A file, containing one or more commands that can be run to perform processing.
shared disk group	A disk group in which the disks are shared by multiple hosts (also referred to as a cluster-shareable disk group).
sector	A minimal unit of the disk partitioning. The size of a sector can vary between systems. A sector is commonly 512 bytes.
segment	Any partition, reserved area, partial component, or piece of a larger structure.
SGA	See "System Global Area."
single threading	The processing of one transaction to completion before starting the next.
slave node	A node that is not designated as a master node.
snapped file system	A file system whose exact image has been used to create a snapshot file system.
snapped volume	A volume whose exact image has been used to create a snapshot volume.
snapshot	A point-in-time image of a volume or file system that can be used as a backup.
snapshot file system	An exact copy of a mounted file system, at a specific point in time, that is used for online backup. A snapshot file system is not persistent and it will not survive a crash or reboot of the system.
snapshot volume	An exact copy of a volume, at a specific point in time. The snapshot is created based on disk mirroring and is used for online backup purposes.
spanning	A layout technique that permits a volume (and its file system or database) too large to fit on a single disk to distribute its data across multiple disks or volumes.
storage class	Set of volumes with the same volume tag.
stripe	A set of stripe units that occupy the same positions across a series of columns in a multi-disk layout.
stripe unit	Equally sized areas that are allocated alternately on the subdisks (within columns) of each striped plex. In an array, this is a set of logically contiguous blocks that exist on each disk before allocations are made from the next disk in the array.
stripe unit size	The size of each stripe unit. The default stripe unit size for VxVM is 32 sectors (16K). For RAID 0 striping, the stripe unit size is 128 sectors (64K). For VxVM RAID 5, the stripe unit size is 32 sectors (16K). A stripe unit size has also historically been referred to as a stripe width.

striping	A layout technique that spreads data across several physical disks using stripes. The data is allocated alternately to the stripes within the subdisks of each plex.
subdisk	A consecutive set of contiguous disk blocks that form a logical disk segment. Subdisks can be associated with plexes to form volumes.
superuser	A user with unlimited access privileges who can perform any and all operations on a computer. In UNIX, this user may also be referred to as the “root” user. On Windows/NT, it is the “Administrator.”
tablespace	A tablespace is a storage structure (containing tables, indexes, large objects, and long data) that allows you to assign the location of database and table data directly onto containers. Tablespaces reside in database partition groups.
terabyte	A measure of memory or storage. A terabyte is approximately 1,000,000,000,000 bytes (technically, 2 to the 40th power, or 1,000 GB). Also TB.
throughput	A measure of work accomplished in a given amount of time. For file systems, this typically refers to the number of I/O operations in a given period of time.
unbuffered I/O	I/O that bypasses the file system cache for the purpose of increasing I/O performance (also known as direct I/O).
Veritas Enterprise Administrator	Application that is required to access graphical user interface (GUI) functionality.
Veritas NetBackup (NBU)	A product that lets you back up, archive, and restore files, directories, or raw partitions that reside on your client system.
Veritas Volume Replicator (VVR)	A feature of Veritas Volume Manager, VVR is a data replication tool designed to contribute to an effective disaster recovery plan.
volume	A logical disk device that appears to applications, databases, and file systems as a physical disk partition. A logical disk can encompass multiple or one to many physical volumes.
volume layout	A variety of layouts that allows you to configure your database to meet performance and availability requirements. This includes spanning, striping (RAID-0), mirroring (RAID-1), mirrored stripe volumes (RAID-0+1), striped mirror volumes (RAID-1+0), and RAID 5.
volume manager objects	Volumes and their virtual components. See “object (VxVM).”
VVR	See “Veritas Volume Replicator (VVR).”
vxfs or VxFS	The acronym for Veritas File System.
vxvm or VxVM	The acronym for Veritas Volume Manager.

Index

A

availability
 using mirroring for 14

B

backups
 implementing online 65
benefits of Concurrent I/O 45

C

Cluster Volume Manager 20
commands
 fsadm 23, 41
 mkfs 24, 34
 mount 24
 umount 37
 vxupgrade 50
concatenation 13
Concurrent I/O
 benefits 45
 disabling 46
 enabling 46
creating
 a volume 31
cron 96
cross-platform data sharing 23

D

data change object 18
data redundancy 14
data warehousing 15
DCO 18
 log volume 18
decision support
 implementing 87
defragmentation 22
 extent 96
 scheduling 96
 utility 22
device interface 19

Dirty Region Logging 16
dirty region logging 17, 31
dirty volumes 16
disabling Concurrent I/O 46
discovered_direct_iosize tunable parameter 98
disk arrays 13
 DMP-supported 19
disk group
 naming a disk group 27
disk groups
 about 13
 adding disks 29
 configuration guidelines 27
 creating
 using the command line 28
 defined 13
 join 58
 joining 57
 limitations of move 58
 recovery from failed reconfiguration 58
 split 58
 split and join 18
disk space allocation 21
disks
 adding to a disk group 29
 failure and hot-relocation 19
dissociating snapshot volumes 93
DMP 19
DMP-supported disk arrays 19
DRL 16, 31. *See* Dirty Region Logging
DSS workloads
 guidelines 31
dynamic LUN expansion 20
Dynamic Multipathing 19

E

enabling Concurrent I/O 46
exclusive OR 15
expansion
 file system 96
extent-based allocation 21

- F**
- fast file system 22
 - fast recovery 31
 - FastResync
 - non-persistent 18
 - operation with off-host processing 56
 - persistent 18
 - use with snapshots 18
 - FastReysnc 17
 - file systems
 - configuration guidelines 33
 - increasing the size 41
 - mounting 36
 - overview 21
 - resizing 23, 41
 - unmounting 37
 - fragmentation 22, 38
 - controlling 38
 - monitoring 39, 96
 - reorganization facilities 96
 - reporting 96
 - types 38
 - fragmented file system
 - characteristics 97
 - free space 19, 96
 - monitoring 96
 - fsadm
 - reporting extent fragmentation 97
 - scheduling 97
 - fsadm command 23, 41
 - fsadm utility 23, 42
 - fsapadm command 35
 - fsvoladm command 35
- G**
- guidelines
 - creating file systems 33
 - disk groups 27
 - for DSS workloads 31
 - for OLTP workloads 30
 - striped volumes 30
 - volumes 30
- H**
- High Availability (HA)
 - overview 10
 - hot-relocation 19
- I**
- I/O
 - sequential 21
 - statistics
 - obtaining 101
 - initial_extent_size tunable parameter 99
- L**
- large file systems 23
 - large files
 - support for 24
- M**
- max_direct_iosize tunable parameter 99
 - max_diskq tunable parameter 99
 - max_seqio_extent_size tunable parameter 100
 - memory
 - persistence of FastResync in 18
 - mirrored volume snapshots 18
 - mirrored-stripe volumes 15
 - mirroring 10, 14
 - choosing 30
 - defined 14
 - mirroring and striping data 15
 - mkfs command 24, 34
 - monitoring fragmentation 96
 - mount command 24
 - mounting
 - file systems 24
 - mounting file systems 36
 - multi-volume support 23, 35
 - fsvoladm command 35
 - mutli-volume support
 - fsapadm command 35
- N**
- nolargefiles option 24
 - non-persistent FastResync 18
- O**
- OLTP. *See* online transaction processing
 - OLTP workloads
 - guidelines 30
 - online backups
 - implementing 65
 - online relayout 16
 - online transaction processing 15

options
 largefiles and nolargefiles 24

P

parallel data transfer
 using striping for 14
 parameters
 default 97
 tunable 97
 tuning 97
 parity 15
 performance
 obtaining statistics for volumes 102
 RAID-5 15
 performance tuning
 list of guides 95
 persistent FastResync 18

Q

QoS. *See* quality of storage service
 quality of storage service 23
 quotas 24

R

RAID 13
 RAID-0 14
 RAID-0+1 15
 RAID-1 14
 RAID-1+0 15
 RAID-5 15, 30
 choosing 30
 performance 30
 RAID-5 log 31
 read_nstream tunable parameter 98
 read_pref_io tunable parameter 98
 recreating
 data using RAID-5 15
 redo logs
 configuration guidelines 34
 creating a file system 34
 relay layout 16
 reliability
 using mirroring for 14
 removing snapshot volumes 93
 report
 extent fragmentation 96
 resizing file systems 41
 resizing utility 22-23

resynchronization
 using DRL logs 31
 using RAID-5 logs 31
 resynchronizing
 volumes 16

S

SCSI devices 19
 selecting volume layouts 29
 sequential I/O
 using extent-based allocation 21
 sequential read/writes
 using spanning for 14
 snapshot volumes
 dissociating 93
 using the command line 93
 removing 93
 snapshots
 and FastResync 18
 spanning 10, 13
 defined 14
 spare disks 19
 statistics
 volume I/O 101
 Storage Checkpoints
 file system restores 24
 Storage Expert 20
 stripe unit sizes
 choosing 30
 stripe units 14
 striped volumes 30
 configuration guidelines 30
 striping 10, 14
 defined 14
 striping and mirroring data 15
 support for large files 24

T

tunable I/O parameters 97
 discovered_direct_iosize 98
 initial_extent_size 99
 max_direct_iosize 99
 max_diskq 99
 max_seqio_extent_size 100
 read_nstream 98
 read_pref_io 98
 write_nstream 98
 write_pref_io 98

tunable I/O parameters (*continued*)
 write_throttle 100

Tuning
 VxFS 95
 VxFS I/O parameters 97

tuning
 vxfs 95
 VxVM 101

tuning I/O parameters 97

U

umount command 37
 unmounting
 a file system 37
 unmounting file systems 37
 upgrading
 from earlier VxFS layout versions 49
 utilities
 fsadm 23, 42
See also commands
 online administration 22

V

Veritas FastResync. *See* FastResync
 Veritas File System
 cluster functionality 24
 cross-platform data sharing 23
 defragmentation utility 22
 fast file system 22
 multi-volume support 23
 online administration utilities 22
 overview 21
 quality of storage service 23
 quotas 24
 resizing utility 22–23
 support for large file systems 23
 Veritas Volume Manager 11
 and RAID 13
 objects 12
 overview 10–11
 Veritas Volume Replicator 20
 volume layouts 13
 changing 16
 concatenation 13
 mirrored-stripe 15
 mirroring 14
 RAID-5 15
 selecting 29

volume layouts (*continued*)
 spanning 13
 striping 14
 volume resynchronization 16
 volume snapshots. *See* snapshots
 volumes
 about 12
 configuration guidelines 30
 creating 31
 using the command line 32
 definition 12
 layouts 13
 marked as dirty 16
 obtaining performance statistics 102
 resynchronizing 16
 vxassist
 used to remove DCOs from volumes 63
 VxFS
 resizing utility 22
 tuning 95
 VxFS.. *See* Veritas File System
 vxstat
 used to obtain volume performance
 statistics 102
 vxupgrade command 50
 VxVM. *See* Veritas Volume Manager
 overview 11
 tuning 101

W

workloads
 write-intensive 31
 write_nstream tunable parameter 98
 write_pref_io tunable parameter 98
 write_throttle tunable parameter 100

X

XOR. *See* exclusive OR