VERITAS®

VERITAS WHITE PAPER

VERITAS FlashSnap™

Guidelines for Using VERITAS FlashSnap

VERITAS®

Table of Contents

Summary

Enterprises require continuous availability of their business-critical data. Not only is data required to be online, it is required to be available for business analysis, backups and other administrative operations. This same data is also required by application developers to test new applications.

VERITAS FlashSnap™ resolves data availability issues by making point-in-time copies of business-critical data available off-host for backups, application testing, business analysis and other processing needs for business continuity.

VERITAS FlashSnap uses Dynamic Group Split and Join and FastResync technologies of VERITAS Volume Manager™ to create point-in-time copies, or snapshots, of the diskgroups. These copies can then be made available for performing multiple tasks in parallel while the primary diskgroup remains available for regular business processing on the original host. Since VERITAS technologies are storage-vendor-neutral, VERITAS FlashSnap can work on virtually any disk array, whether it be an expensive high-end disk array or inexpensive JBOD.

Introduction

This paper discusses uses, applications, detailed architectural descriptions, operations and commands for using VERITAS FlashSnap.

VERITAS technologies provide storage-vendor-neutral software solutions supporting disk subsystems ranging from expensive high-end disk arrays to inexpensive JBOD. With VERITAS FlashSnap you can create point-in-time copies at the volume level (referred to as volume snapshots) that can be used to perform business-critical operations. The original diskgroup and the volume snapshot do not have to be on the same type of the disk subsystem. In the storage area network (SAN) environment, the volume snapshots can be accessed from any host machine that has the I/O path to the physical devices. Point-in-time copies at the file-system level, called Storage Checkpoints, can also be made in UNIX environments.

VERITAS Technologies Used in VERITAS FlashSnap™

VERITAS Volume Manager™

VERITAS Volume Manager is a disk management subsystem that supports disk striping, disk mirroring and simplified disk management for improved data availability, manageability and superior performance. VERITAS Volume Manager is required, at a minimum, to enable the use of VERITAS FlashSnap.

VERITAS FlashSnap™

VERITAS FlashSnap is a license-key-enabled option of the VERITAS Foundation Suite™ solutions. The FlashSnap license enables the use of the FastResync and Dynamic Split and Join features of Volume Manager and Storage Checkpoints if used in conjunction with VERITAS File System™ *.

FastResync of Volume Snapshots

VERITAS Volume Manager allows you to take multiple snapshots (up to 32) of your data at the level of a volume. A volume snapshot contains a stable copy of a volume's data at a given moment in time that you can use for operations such as online backup or decision support. If FastResync is enabled on a volume, Volume Manager uses a "FastResync map" to keep track

* The Storage Checkpoint feature available in VERITAS File System is available for UNIX only.

of which blocks are updated (dirty blocks) in the volume and in the snapshot. When the data in the snapshot is not updated, it becomes out-of-date, or stale, with respect to the primary volume. The presence of the FastResync map ensures that only those updates (dirty blocks) that the mirror has missed are reapplied to the stale volume/snapshot once the volume snapshot is resynchronized with the primary volume. A full, and thereby much slower, resynchronization of the mirror with the primary volume is unnecessary. Volume Manager FastResync maps are persistent because the maps are stored in a data change object (DCO) log volume that is associated to the volume. Therefore, the maps survive both system reboots and system crashes. Volume Manager snapshot resynchronization is bidirectional.

Dynamic Split and Join

One or more volume snapshots can be split off into a separate diskgroup and deported using the Dynamic Split and Join feature of FlashSnap. These deported volume snapshots are then ready for importing onto another host that may be dedicated to off-host processing. This host need not be a member of a cluster but it must have access to the disks on which the volumes are configured. At a later stage, when processing is complete, the diskgroup can be deported, re-imported and joined with the original diskgroup, or with a different diskgroup if need be.

VERITAS File System™

VERITAS File System is a high-performance, fast-recovery file system that is optimized for business-critical applications and data-intensive workloads. VERITAS File System offers online administration capabilities, enabling frequently scheduled maintenance tasks (including online backup, resizing, defragmentation and file system changes) to be performed without interrupting data or system availability. The FlashSnap license enables the use of the Storage Checkpoint™ feature of VERITAS File System.

Storage Checkpoint™

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

- A Storage Checkpoint preserves not only the directory hierarchy of the file system, but also the user data as it existed when the Storage Checkpoint was taken.
- After creating a Storage Checkpoint of a mounted file system, you can continue to create, remove and update files on the file system without affecting the image of the Storage Checkpoint.
- Storage Checkpoints are persistent across system reboots and crashes.
- Unlike file system snapshots, users have the option to make Storage Checkpoints writable.
- To minimize usage of disk space, Storage Checkpoints use free space that is available to the file system.

Applications of FlashSnap™

The following activities are suitable to be performed using point-in-time copy solutions using VERITAS FlashSnap:

Decision Support Analysis and Reporting

Operations such as decision-support analysis and business reporting may not require access to real-time information. Such operations can be directed to use a copy or replica database that can be created from snapshots using VERITAS FlashSnap, rather than competing for access to the primary volume or database. When required, the snapshot can be quickly resynchronized with the data in the primary database using the FastResync technology in FlashSnap.

Application Testing and Training

Development groups and/or service groups can use volume snapshots as test data for new applications. Volume snapshot data provides developers, system testers and QA groups with a realistic basis for testing the robustness, integrity and performance of new applications.

Data Backup

Many enterprises require 24/7 online data availability. They cannot afford the downtime involved in backing up critical data. Creating volume snapshots of data and using them to conduct backups allows business-critical applications to continue to run without impacted performance or extended down time.

Logical Error Recovery

Logical errors caused by an administrator or an application program can compromise the integrity of a database. A database can be quickly recovered by restoring the database files from a volume snapshot or Storage Checkpoint. The alternative would require a full restoration from tape or other backup media.

How Does FlashSnap™ Work?

FlashSnap technology has evolved from a Volume Manager concept called a volume snapshot (mirror). As show in Figure 1, for a mirror to be used as a volume snapshot, it needs to be created by issuing a **"`vxassist snapstart`"** command, which attaches the mirror to the volume as a snapshot.

When the attachment is complete, the `vxassist snapshot` command is used to create a new snapshot volume by taking one or more snapshot mirrors to use as its data plexes (mirrors). The volume snapshot contains a copy of the primary volume's data at the time that you took the snapshot. If more than one volume snapshot is used, the snapshot volume is itself mirrored.

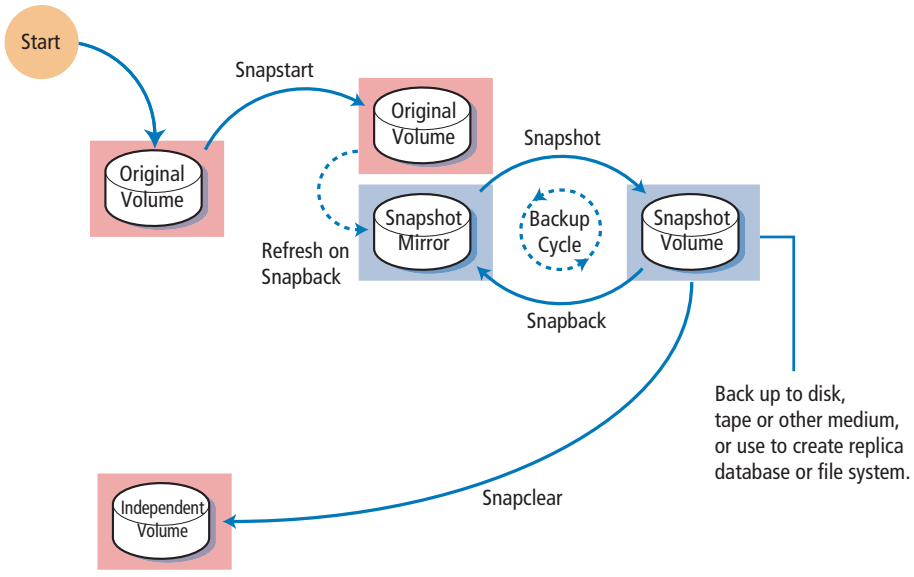

Figure 1: How a snapshot works.

The command **"vxassist snapback"** can be used to return the volume snapshot to the volume from which it was snapped. When the volume snapshot is attached back to the primary volume, resynchronization of the data takes place. Because of the FastResync technology in FlashSnap, the resynchronization becomes faster because only the changes that took place after the volume snapshot was created are resynchronized.

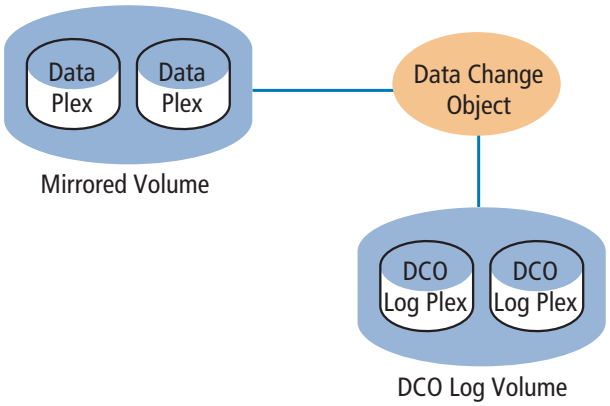

Figure 2: Mirrored volume with persistent FastResync enabled.

Persistent FastResync stores the FastResync maps on disk so that they can survive both system reboots, system crashes and cluster crashes. If Persistent FastResync is enabled on a volume or on a snapshot volume, a data change object (DCO) and a DCO log volume are associated with the volume.

DCO Log

Persistent FastResync uses a map in a DCO log volume on disk to implement change tracking. These maps track writes to the original volume and to each of the volume snapshots (up to 32 snapshots supported) since the last snapshot operation. The DCO log volume on disk holds the 33 maps, each of which is four blocks in size by default (i.e., by default dcologlen is 33x4=132 blocks). To use a larger map size, multiply the desired map size by 33 to calculate dcologlen that needs to be specified. The maximum DCO log length is 2,112 blocks. Persistent FastResync can also track the association between volumes and their volume snapshots after they are moved into different diskgroups. When the diskgroups are rejoined, this allows the volume snapshots to be quickly resynchronized.

Figure 3 shows an example of a mirrored volume with two plexes on which Persistent FastResync is enabled. Associated with the volume are a DCO object and a DCO log volume with two plexes. When the **"snapstart"** operation is performed on the volume, it sets up a snapshot plex in the volume and associates a disabled DCO log plex with it.

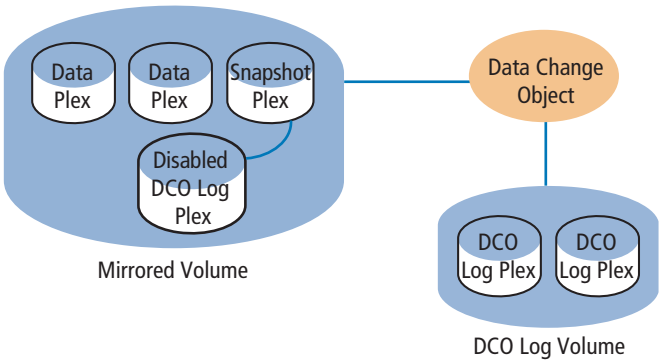

Figure 3: Mirrored volume after completion of a snapstart operation.

A volume snapshot can now be created from a primary volume by running the `snapshot` operation on the volume. As illustrated in Figure 4, this also sets up a DCO object and a DCO log volume for the volume snapshot. The new DCO log volume contains the single DCO log plex that was associated with the primary volume.

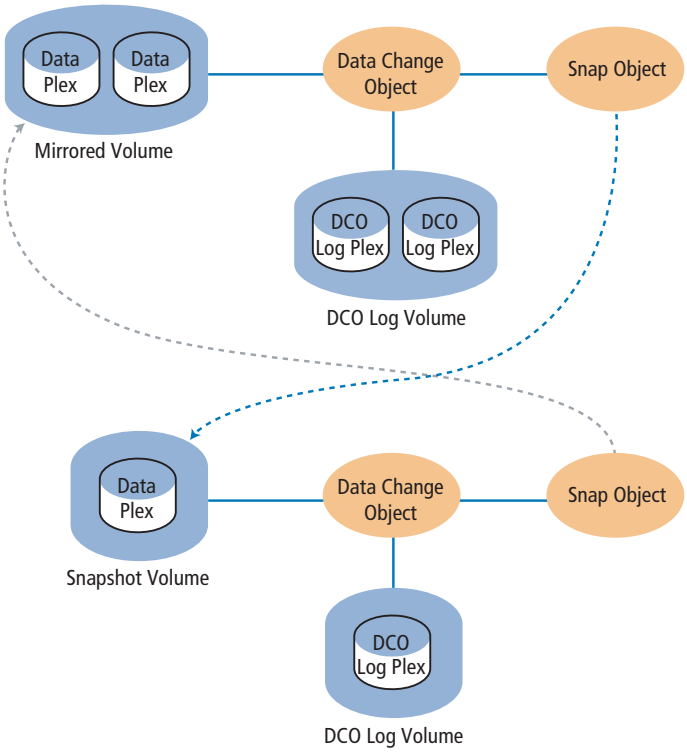

Figure 4: Mirrored volume and snapshot volume after completion of a snapshot operation.

Associated with both the primary volume and the volume snapshot are **"snap objects."** The snap object for the primary volume points to the volume snapshot, and the snap object for the volume snapshot points to the primary volume. This allows Volume Manager to track the relationship between volumes and their snapshots even if they are moved into different diskgroups.

The snap objects in the primary volume and volume snapshot are automatically deleted in either of the following circumstances:

- The **"snapback"** operation is run to return all of the volume snapshots to the primary volume.
- The **"snapclear"** operation is run on a volume to break the association between the primary volume and the volume snapshot. If the volumes are in different diskgroups, snapclear must be run separately on each volume.

By default, the volume snapshot is resynchronized to the data in the primary volume during a snapback operation. Alternatively, you can choose the volume snapshot as the preferred copy of the data when performing a snapback as illustrated in Figure 5. Specifying the option **"-o resyncfromreplica to vxassist"** resynchronizes the primary volume from the data in the snapshot. The primary volume must not be in use during a snapback operation in which it is resynchronized from a snapshot. Stop any application, such as a database, and unmount any file systems that are configured to use the volume.

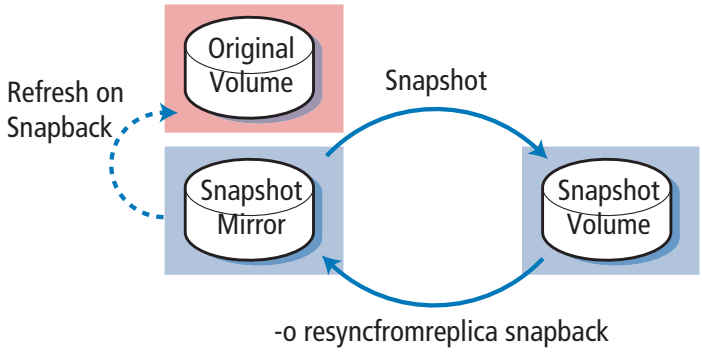

Figure 5: Resynchronizing the original volume from the snapshot.

Implementing FlashSnap™

As shown in Figure 6, by accessing volume snapshots from a lightly loaded host (shown here as the "OHP host"), CPU- and I/O-intensive operations for online backup and decision support do not degrade the performance of the primary host that is performing the main production activity (such as running a database).

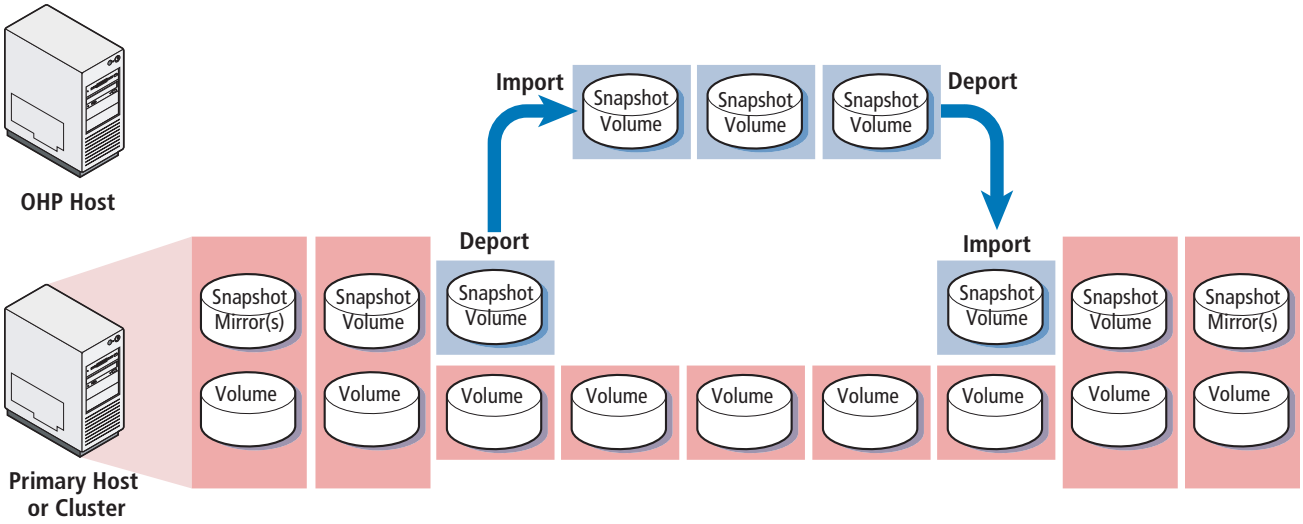

Figure 6: Using VERITAS FlashSnap™ to implement off-host processing solutions.

The steps involved along with the commands are:

1. Create diskgroup: Use vxdg init to create diskgroup with at least two disks.

```
vxdg init <diskgroup name>  <disk name>=<device name>
Example: vxdg init flash disk01=c2t8d0 disk02=c2t9d0
```

2. Create simple volume in the diskgroup just created.

```
vxassist –g <diskgroup name> make <volume name> size  layout=<layout type>
Example: vxdg init flash disk01=c2t8d0
```

3. Create DCO log and DCO log volume to store the FastResync map for each of the primary volumes used.

```
vxassist –g <diskgroup name> addlog <volume name> logtype=dco \
[ dcologlen=<size of the map in blocks>]  [ storage specification]
Example: vxassist -g flash addlog f_vol logtype=dco loglen=264 disk01 disk02
```

The default number of plexes in the mirrored DCO log volume is 2. It is recommended to configure as many DCO plexes as there are existing data and snapshot plexes in the volume.

The default size of each plex is 132 blocks. You can use the dcologlen attribute to specify a different size. If specified, the size of the plex must be an integer multiple of 33 blocks from 33 up to a maximum of 2,112 blocks.

4. Turn on FastResync flag on the volume.

```
vxvol –g <diskgroup name> set fastresync=on <volume name>
Example: vxvol -g flash set fastresync=on f_vol
```

5. Create snapshot mirrors on one or more volumes.

```
vxassist –g <diskgroup name> –b snapstart  <volume name> \
alloc=[ storage specification]
Example: vxassist -g flash -b snapstart f_vol alloc=disk03
```

Starts vxassist snapstart command in the background with the –b option; otherwise the command does not return until the snapshot mirror has been synchronized with the volume and set to SNAPDONE state. The snapstart command also adds another DCO log plex that is DISABLED with a DCOSNP state. The snapshot volume will use this DCO log plex to store the FastResync map with respect to the original volume. The vxassist snapwait command will wait until the snapstart command is done. This command is useful in the following step.

6. Create snapshot volumes from snapshot mirror.

```
vxassist –g <diskgroup name> snapshot  <volume1 name>
Example: vxassist -g flash snapshot f_vol f_vol_snap
```

If this is done before the snapstart completes, then the snapshot command will fail. Using the snapwait command described in Step 5 will allow a script that did a snapstart with -b to wait until it is done before proceeding with the snapshot.

7. Split the diskgroup to move the disk containing the snapshot volumes to a separate diskgroup (using the Dynamic Split and Join feature).

```
vxdg split <original diskgroup> <target diskgroup> \
<snapshot volume name>
Example: vxdg split flash flash_snapdg f_vol_snap
```

8. Deport the diskgroup containing the volume snapshot.

```
vxdg deport <target diskgroup>
Example: vxdg deport flash_snapdg
```

9. Import the diskgroup containing the volume snapshot on the off-host processing host.

```
vxdg import <target diskgroup>
Example: vxdg import flash_snapdg
```

10. Start all volumes in the new diskgroup.

```
vxvol –g <target diskgroup> startall
Example: vxvol -g flash_snapdg startall
```

11. Apply off-host processing application such as reporting, DSS, backups, etc., to the volume snapshot on the off-host processing host.

12. Deport the diskgroup containing the snapshot from the off-host processing host.

```
vxdg deport <target diskgroup>
Example: vxdg deport flash_snapdg
```

13. Import the diskgroup containing the volume snapshots on the primary host.

```
vxdg import <target diskgroup>
Example: vxdg import flash_snapdg
```

14. Join (or merge) the diskgroup containing the volume snapshots with the primary volume's diskgroup.

```
vxdg join <target diskgroup> <original diskgroup>
Example: vxdg join flash_snapdg flash
```

15. Start the volume snapshots.

```
vxvol –g <original diskgroup> start <snapshot volume name>
Example: vxvol -g flash start f_vol_snap
```

16. Reattach and resynchronize the volume snapshot with the primary volume.

```
vxassist –g <original diskgroup> snapback <snapshot volume1 name>
Example: vxassist -g flash snapback f_vol_snap
```

As discussed earlier, when a volume snapshot is reattached to its primary volume within a shared diskgroup, there are two choices for resynchronizing the data in the volume:

- Resynchronize the snapshot from the primary volume
- Resynchronize the original volume from the snapshot

Considerations for Placing DCO Log Plexes

In order to access a split-off volume snapshot from another host, the snapshot and its corresponding DCO log plex cannot share physical disks with the primary volume and its DCO log plex. If you use the vxassist command or the GUI to create a volume, or to enable Persistent FastResync (DCO logging) on a volume, the DCO log plexes are automatically placed on the same disks as the data plexes of the parent volume. When you move the parent volume (such as a snapshot volume) to a different diskgroup, this ensures that the DCO log volume automatically accompanies it. If you use the `vxmake` and `vxdco` commands to set up DCO logging, you must ensure that the disks that contain the plexes of the DCO log volume can accompany their parent volume during the move. Use the `vxprint` command on a volume to examine the configuration of its associated DCO log volume.

Example 1: When volume snapshots are split from their associated diskgroup, the DCO log plexes are on the same disk as the volumes and will therefore ensure that they will survive a system failure.

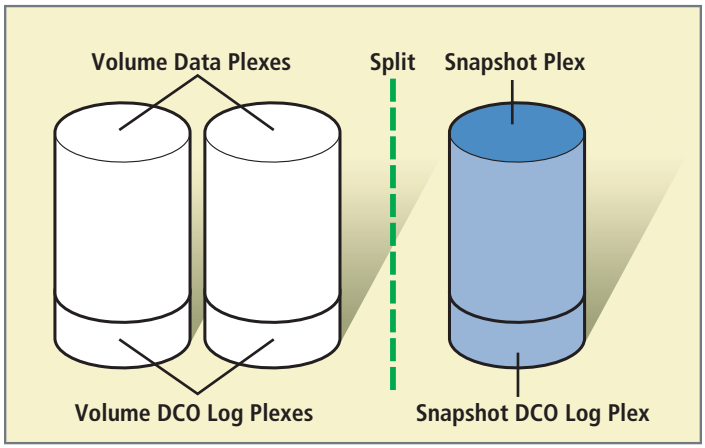

Figure 7: Example of a disk group that can be split.

As illustrated in Figure 8, the diskgroup cannot be split because the DCO log plexes have been separated from their data plexes and cannot accompany their volumes.

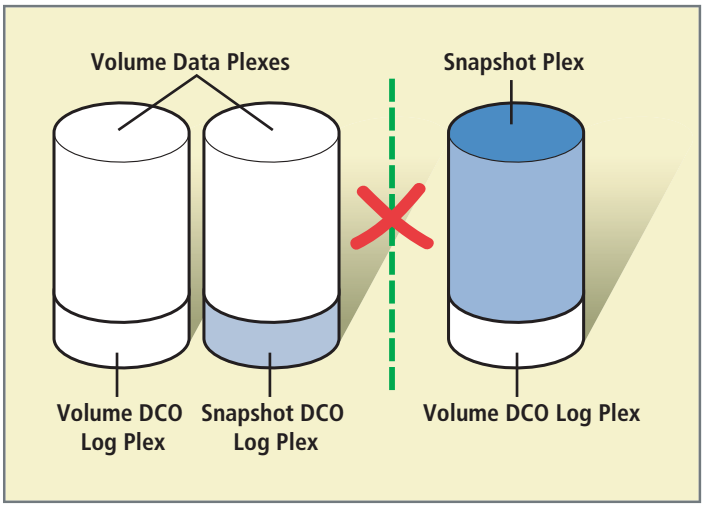

Figure 8: Example of a disk group that cannot be split.

Example 2: The diskgroup illustrated in Figure 9 can be split and DCO logs can accompany their volumes even though they are on separated disk. However it is not a good practice since some portions of the disk that are not required will also be moved.

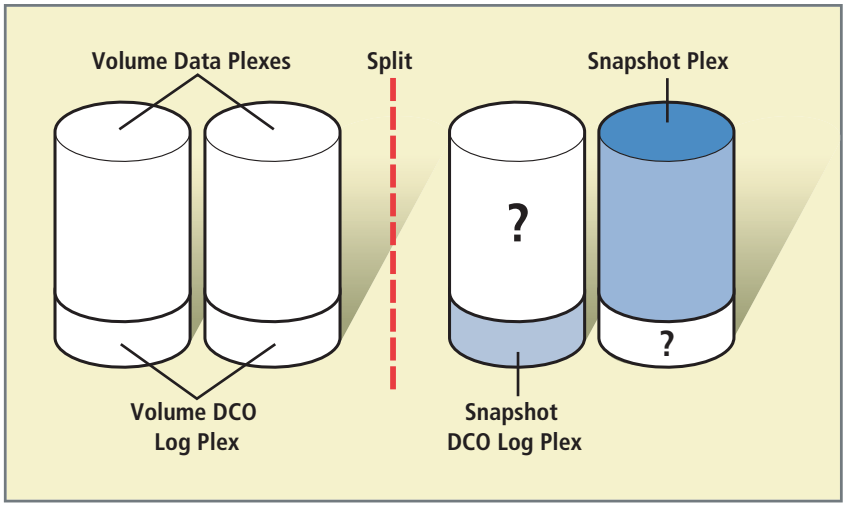

Figure 9: Example of a disk group that cannot be split.

Summary

VERITAS FlashSnap, an option of VERITAS Foundation Suite, is a flexible storage management solution that enables administrators to create point-in-time copies with minimal impact to applications and users. It also addresses issues such as shrinking maintenance and backup windows. These snapshots can be accessed from the same server or easily imported to another host, which allows users to perform resource-intensive processes such as backups, decision support and reporting without disrupting the performance of production systems. To greatly reduce the resynchronization time and performance impact on the server when the volume snapshot is reattached, FastResync technology synchronizes only the changes that occurred while the volume snapshots were split. These point-in-time copies can also serve as on-disk backup images for protecting mission-critical data. FlashSnap works with a wide range of disk subsystems, delivering a powerful and flexible solution that frees you from hardware and application restrictions.

Reference

VERITAS Volume Manager 3.2 Administrator's Guide
VERITAS FlashSnap Point-In-Time Copy Solutions 1.0
VERITAS File System 3.4 Administrator's Guide

VERITAS®

V E R I T A S W H I T E P A P E R

VERITAS Software Corporation
Corporate Headquarters
350 Ellis Street
Mountain View, CA 94043
650-527-8000 or 866-837-4827

For additional information about VERITAS Software, its products, or the location of an office near you, please call our corporate headquarters or visit our Web site at www.veritas.com

02-20483