VERITAS® Volume Manager
Hardware Application Note
Release 3.0.4

READ THIS BEFORE INSTALLING THE SOFTWARE!

Introduction

This document provides hardware support information for the VERITAS Volume Manager™ Release 3.0.4.

This release of the Volume Manager has been tested on:

- Solaris 2.5.1, Solaris 2.6, Solaris 7, and Solaris 8 (32- and 64-bit).

Contents

The following topics are covered in this document:

- Getting Help
- Using Prestoserve With the Volume Manager
- VERITAS Volume Manager Disk Array Configuration Updates
  - Disk Arrays Supported for DMP
  - Using the ATF Driver With the VERITAS Volume Manager
  - Moving From Alternate Pathing to DMP
  - Dynamic Reconfiguration
  - Using the Sun StorEdge A3000 Array With the VERITAS Volume Manager
  - Installation of the Volume Manager With the Sun StorEdge A3000 Array
  - VERITAS Volume Manager and Sun StorEdge A3000 Array Device Nodes
  - Benefits in Using VERITAS Volume Manager With the Sun StorEdge A3000 Array

- Sun StorEdge A3000 Array Supported Configurations
- Sun StorEdge A3000 Array Unsupported Configurations
- Volume Manager Hot-Relocation and Sun StorEdge A3000 Array Hot Sparing

Getting Help

For license information or information about VERITAS® service packages, contact VERITAS Customer Support.

US Customers: 1-800-342-0652
International Customers: +1 (650) 335-8555
Fax: (650) 335-8428
Electronic mail: support@veritas.com

Using Prestoserve With the Volume Manager

Prestoserve is designed to accelerate performance of file systems, particularly when used on a server for NFS advertised file systems. This is accomplished via the use of NVRAM hardware and the Prestoserve drivers. The hardware provides a fast, nonvolatile solid-state writeback cache that can cause writes to a disk device to be returned to the user as completed before the data reaches the disk.

This mechanism can be configured to work below VxVM as direct replacement for the disk device that VxVM uses. This approach presents no particular problems for VxVM, which remains unaware of the underlying cache device. In the event of a failure of the NVRAM devices, it is possible to lose data since the disks backing the NVRAM may not be up-to-date.

Prestoserve can be configured to run above VxVM in such a way that VxVM replaces the disks that Prestoserve controls. In this situation, VxVM has a number of problems to address.

The most serious problem is with the startup of Prestoserve. Following a system failure, the Prestoserve drivers cause a flush of all outstanding dirty buffers to be flushed to disk. If this flush request occurs before VxVM drivers have been loaded into the kernel and before the volume devices can be started and made available for use, then Prestoserve's attempts at flushing to the volumes fails.

> **CAUTION!** This problem could lead to file system corruption and data loss.

To prevent this situation, it is recommended that the order of the starting of Prestoserve with respect to the volumes be altered to occur after the volumes have been started. To achieve this, perform the following steps:

1. Edit the `/etc/system` file to add the line

   ```
   exclude: drv/pr
   ```

   and remove any line similar to this:

   ```
   forceload: drv/pr
   ```

   This causes the Prestoserve driver to be loaded and the flush operation to be started *after* the volume devices have been started.

2. Edit the file `/etc/init.d/vxvm-startup2` and add the following lines to the end of the file:

   ```
   modload /kernel/drv/pr
   presto -p > /dev/null
   ```

   This causes initialization of the Prestoserve driver (and flushing of saved blocks) following the start of all volumes.

   Edit the `/etc/init.d/prestoserve` file and replace the following line:

   ```
   presto -u
   ```

   with the following:

   ```
   presto -u /mnt1 /mnt2 ...
   ```

   Where */mnt1* and */mnt2* should list every file system to be accelerated with Prestoserve. Note that this list must *not* include `/`, `/usr`, `/usr/kvm`, `/var`, or `/var/adm`.

> **CAUTION!** Without the explicit starting of `presto` with the desired file systems, the default action accelerates all file systems, which (on the next reboot) could lead to data corruption since the file system checks for `/`, `/usr`, `/usr/kvm`, `/var`, and `/var/adm` precede the flushing of the Prestoserve buffers.

As is stated in the *Prestoserve User's Guide*, it is important not to reboot the system using the `-r` flag unless the system was cleanly shut down. VxVM can be used to add and remove drives without the need for a `boot -r`.

A second, and more subtle problem, is in Prestoserve's use of disk devices. Some applications (including Prestoserve) maintain device numbers between reboots. VxVM also attempts to maintain device numbers between reboots. However, if a disk group is moved from one machine to another (or reattaching of disks causes some previously non-working disk group to reappear), it is possible for a conflict of minor numbers to be detected. VxVM tries to prevent this by using randomly-assigned device number ranges (which are less likely to result in conflicts), but this policy can fail if two disk groups from different machines end up with the same randomly-assigned range.

The danger of VxVM changing its device numbers on a reboot following a system failure is that Prestoserve may flush its dirty buffers to the wrong volume devices. This may have destructive results.

If two disk groups do end up with volumes that use conflicting device numbers, the device number for one of the volumes is temporarily remapped on reboot (or when the second disk group is imported). Each reboot may choose a different remapping than was chosen on the previous reboot, leading to problems with Prestoserve. Anytime a volume device number is temporarily remapped, the following console message is displayed:

```
vxvm:vxconfigd: WARNING: Group …: Duplicate virtual device number:
        Volume … remapped from old-number to new-number
```

If you see such a message, you should change the numbers for all volumes in the named disk group, using the `vxdg reminor` command.

One method for ensuring that Prestoserve never encounters two volumes with conflicting device numbers is to use Prestoserve only on volumes in the `rootdg` disk group. The `rootdg` disk group is always imported first, so it cannot conflict with a previously-imported disk group.

VERITAS Volume Manager Disk Array Configuration Updates

This section lists some of the disk arrays supported for Dynamic Multipathing (DMP) and describes how to use Volume Manager with the Sun StorEdge A3000 Array. These topics are described in this section.

- The DMP Command Utility (`vxdmpadm`)

- Disk arrays supported for DMP

- Using the Sun StorEdge A3000 Array with the VERITAS Volume Manager

- Installation of the Volume Manager with the Sun StorEdge A3000 Array

- VERITAS Volume Manager and Sun StorEdge A3000 Array device nodes

- Benefits in using VERITAS Volume Manager with the Sun StorEdge A3000 Array

- VERITAS Volume Manager and Sun StorEdge A3000 Array supported and unsupported configurations

- Volume Manager hot relocation and Sun StorEdge A3000 Array hot sparing

The DMP Command Utility (`vxdmpadm`)

The `vxdmpadm` utility administers the multipathing functionality of Volume Manager (DMP). Refer to the manual page of the `vxdmpadm`(1M) utility for a detailed description of the options and attributes of this command.

You can use the vxdmpadm utility command `listctlr` to list all controllers connected to disks that are attached to the host. For example, to list all controllers connected to disks on the host, use the following command:

```
# vxdmpadm listctlr all
```

to display the following example output:

```
CTLR-NAME              DA-TYPE        STATE        DA-SNO

========================================================================

ctlr0                  SEAGATE        ENABLED      SEAGATE_DISKS

        ctlr0=/sbus@1f,0/SUNW,fas@e,8800000

ctlr1                  SEAGATE        ENABLED      SEAGATE_DISKS

        ctlr1=/sbus@1f,0/SUNW,socal@0,0/sf@0,0

ctlr2                  SEAGATE        ENABLED      SEAGATE_DISKS

        ctlr2=/sbus@1f,0/SUNW,socal@0,0/sf@1,0

ctlr3                  SSA            ENABLED      501255200077

        ctlr3=/sbus@1f,0/SUNW,soc@1,0/SUNW,pln@b0005012,55200077

ctlr4                  SSA            ENABLED      501255200077

        ctlr4=/sbus@1f,0/SUNW,soc@1,0/SUNW,pln@b0005012,55200077

ctlr5                  OTHER          ENABLED      OTHER_DISKS

        ctlr5=/sbus@1f,0/espdma@e,8400000/esp@e,8800000
```

The above display output shows that:

- the first controller (ctlr0) is connected to Seagate disk(s)

- the second and third controllers (ctlr1 and ctlr2) are connected to an A5X00 (SEAGATE) disk array

  - the Enclosure Serial Number (Disk Array Serial Number) is 50800200000000c8

- the fourth and fifth controllers (ctlr3 and ctlr4) are connected to an SSA (Sparc Storage Array)

  - the Enclosure Serial Number (Disk Array Serial Number) is 501255200077

- the sixth controller (`ctlr5`) is connected to disk(s) that are not under any recognized DMP category

  - the disk(s) are not multipathed

All the above controllers are in the ENABLED state which indicates that they are available for IO operations.

The state DISABLED is used to indicate that controllers are unavailable for IO operations. The unavailability can be due to a hardware failure or due to IO operations being disabled on that controller by the System Administrator. The controller state can be changed by using the vxdmpadm utility.

To list all the paths that are connected to a particular controller, you can use the getsubpaths option with the ctlr attribute. For example, use the following command:

```
# vxdmpadm getsubpaths ctlr=/sbus@1f,0/SUNW,socal@1,0/sf@1,0:
```

to display the following example output:

```
NAME        STATE      TYPE DMPNODENAME   DA-TYPE DA-SNO
============================================================
c2t48d0s2   ENABLED     - c2t48d0s2      SEAGATE SEAGATE_DISKS
c2t35d0s2   ENABLED     - c2t35d0s2      SEAGATE SEAGATE_DISKS
c2t36d0s2   ENABLED     - c2t36d0s2      SEAGATE SEAGATE_DISKS
c2t49d0s2   ENABLED     - c2t49d0s2      SEAGATE SEAGATE_DISKS
c2t38d0s2   ENABLED     - c2t38d0s2      SEAGATE SEAGATE_DISKS
c2t37d0s2   ENABLED     - c2t37d0s2      SEAGATE SEAGATE_DISKS
c2t53d0s2   ENABLED     - c2t53d0s2      SEAGATE SEAGATE_DISKS
c2t34d0s2   ENABLED     - c2t34d0s2      SEAGATE SEAGATE_DISKS
c2t51d0s2   ENABLED     - c2t51d0s2      SEAGATE SEAGATE_DISKS
c2t54d0s2   ENABLED     - c2t54d0s2      SEAGATE SEAGATE_DISKS
c2t52d0s2   ENABLED     - c2t52d0s2      SEAGATE SEAGATE_DISKS
c2t33d0s2   ENABLED     - c2t33d0s2      SEAGATE SEAGATE_DISKS
c2t50d0s2   ENABLED     - c2t50d0s2      SEAGATE SEAGATE_DISKS
c2t32d0s2   ENABLED     - c2t32d0s2     SEAGATE SEAGATE_DISKS
```

The output display shows the paths that are connected to the controller named:

```
/sbus@1f,0/SUNW,socal@1,0/sf@1,0.
```

Additional information displayed shows that:

- path `c2t48d0s2` (represented by nodes in the `/dev/rdsk` and `/dev/dsk` directories) is in the `ENABLED` state

- path `c2t48d0s2` is represented by the DMP metanode `c2t48d0s2` which is represented by device nodes in the `/dev/vx/dmp` and `/dev/vx/rdmp` directories.

- Each of these paths is in a SEAGATE disk array with Enclosure Serial Number 50800200000000c8

You can use the `getsubpaths` option combined with the `dmpnodename` attribute to list all paths that are connected to a `LUN` (represented by a DMP device). For example, to list information about paths that lead to the LUN named `c2t48d0s2`, use the following command:

```
# vxdmpadm getsubpaths dmpnodename=c2t48d0s2
```

to display the following example output:

```
NAME        STATE      TYPE       CTLR-NAME  DA-TYPE   DA-SNO
===========================================================
c1t48d0s2   DISABLED   -          ctlr0      SEAGATE SEAGATE_DISKS
    ctlr0=/sbus@1f,0/SUNW,socal@1,0/sf@0,0
c2t48d0s2   ENABLED    -          ctlr1      SEAGATE SEAGATE_DISKS
    ctlr1=/sbus@1f,0/SUNW,socal@1,0/sf@1,0
```

The listing above shows that the DMP device `c2t48d0s2` has two paths to it that are named `c1t48d0s2` and `c2t48d0s2`. Additional information indicates that only one of these paths is available for I/O operations. One of these paths is in the ENABLED state and the other is in the DISABLED state. Both paths are in a SEAGATE disk array with Enclosure Serial Number `50800200000000c8`.

To retrieve the name of the DMP device that controls a particular path, the `getdmpnode` option can be combined with the `nodename` attribute by using the following command:

```
# vxdmpadm getdmpnode nodename=c1t48d0s2
```

to display the following example output:

```
NAME        STATE      DA-TYPE    PATHS  ENBL  DSBL  DA-SNO
=========================================================
c2t48d0s2   ENABLED    SEAGATE      2      1     1 SEAGATE_DISKS
```

This example output shows that the path `c1t48d0s2` is owned by the DMP device `c2t48d0s2` which has 2 paths to it.

Disk Arrays Supported for DMP

> **Note:** Arrays supported by the DMP feature are a subset of the arrays supported by Volume Manager. The following arrays can have multiple paths to the same disk. Other arrays not listed here can only have a single path to the physical disk.

The Dynamic Multipathing feature of the Volume Manager supports the following arrays:

- EMC Symmetrix (described in this section)
- Hitachi Data Systems™ 5700E/ 5800E Disk Array Subsystem™ (described in this section)
- Hitachi Data Systems 7700E Disk Array Subsystem (described in this section)
- Sun StorEdge A5x00 Array (described in this section)
- JBOD (Just a Bunch of Disks) (described in this section)
- SEAGATE disks that return unique serial numbers in standard SCSI inquiry data
- Storage Computer™ OmniRaid™ disk array. To multipath Storage Computer disk arrays connected to the system while using Volume Manager 3.0.4, you must assign a unique *system name* for each disk array connected to a machine. The RAID-7 Users Manual at the ftp site *www.storage.com* describes how to set a system name for Storage Computer disk arrays.
- ECCS™ Synchronix™ Array

DMP Co-existence With Drivers

The Dynamic Multipathing feature of the Volume Manager also supports co-existence with the following arrays:

- DG CLARiiON with the ATF driver installed on the system

- SYMBIOS Sun StorEdge A3000 Array, only when the RDAC driver is installed on the system

DMP and EMC Symmetrix Disk Array Support

The Dynamic Multipathing feature of the Volume Manager works with the EMC Symmetrix disk array only if the disk array is configured in the *Common Serial Number Mode.* This is because only the *Common Serial Number Mode* configuration provides unique device identification on different paths for DMP to configure its database. Without this mode, DMP cannot provide multipathing functionality. DMP cannot access EMC Symmetrix disks as individual disks with a single path it found by scanning the device information tree in the kernel.

For the DMP functionality to work with the EMC Symmetrix disk array, the following conditions must be met:

- It is necessary that the EMC Symmetrix users program their disk arrays in the *Common Serial Number Mode* before installing the Volume Manager package.

- The minimum level of microcode on the EMC Symmetrix should be:

    -5063 for the 31XX, 32XX, 3500, 51XX, 52XX, and 5500 models

    -5263 for the 33XX, 34XX, 3700, 53XX, 54XX, and 5700 models

- The *Common Serial Number Mode* should be enabled for all channels that participate in DMP. This is done at installation time in the Symmetrix configuration by the EMC Customer Engineer.

- In addition, the *disable queue reset on unit attention mode* must be set for all channels that participate in the DMP. This also is done at installation time in the Symmetrix configuration by the EMC Customer Engineer.

Contact your EMC Technical Support Representative for instructions.

DMP supports the Ultra SCSI and the fast-wide differential (FWD) SCSI controllers used with the EMC Symmetrix disk array.

- Ultra SCSI and fast-wide differential (FWD) SCSI support has been tested on the EMC OSD (Open System Director) controller.

- DMP was tested on and supports the Jaycor fiber controllers: firmware release level F, and driver level V2.0.1EMC and V2.0.2EMC. You should contact Jaycor and EMC for the latest supported driver for the specific controller being used on the EMC Symmetrix.

Note: Uncommenting the `failover=` line in `/kernel/drv/fca.conf line` is important for the DMP failover as it reduces the amount of time it takes for a path to be considered failed. The default time is over five minutes and setting this value reduces the time considerably.

See your sales representative for further updates on hardware support.

DMP and Hitachi Data Systems 5700E/5800E Disk Array Subsystem Support

The Dynamic Multipathing feature of the Volume Manager supports the Hitachi Data Systems 5700E/5800E Disk Array Subsystem under the following conditions:

- The optional second controller is installed in the array.

- One controller is configured as *Active* and the other controller is configured as *Passive* (for improved performance).

- The array is connected to two SCSI cards in the host machine.

- The array has a firmware version that supports serial numbers in the SCSI inquiry data and they are configured correctly.

Table 1   Supported Microcode Levels

Unit	Microcode Version	Flash Revision
DF400 5800	x407/M	a03, b13
DF400 5800	x457/H	a03, b13
DF350 5700E	x308/E	pb28
DF350F 5700E	x356/F	pe27
DF300 5700	x109/L	dq3x/4x (x=89)

Contact your Hitachi Data Systems Technical Support Representative for instructions.

DMP and Hitachi Data Systems 7700E Disk Array Subsystem Support

The Dynamic Multipathing feature of the Volume Manager supports the Hitachi Data Systems 7700E Disk Array Subsystem in *Active/Active* mode only. The 7700E is treated as a jbod by DMP when the string "HITACHI" is added to the /kernel/drv/vxdmp.conf file in the dmp_jbods[] array. This entry is present in the /kernel/drv/vxdmp.conf file by default.

If more than 500 volumes are to be created in a single disk group, the private region for each disk must be increased from the default to 725k. This can be accomplished by initializing each disk with the following command:

```
vxdisksetup -i cXtXdX privlen=725k
```

Sun StorEdge A5x00 Array

This section pertains to the A5000, A5100, and A5200 disk arrays. These disk arrays are claimed by VxVM (DMP) under the Seagate category.

When a cable connecting any of these disk arrays to the host is removed and the commands drvconfig, disks, and vxdctl enable are executed on the host, DMP does not display the pathnames in the standard c#t#d#s# format. Instead, the NONAME string is displayed to represent the name of the removed path.

For example, assume that a host is connected to one of the A5x00 disk arrays. That disk array has two fibre cables connecting it to the host. There is one disk on the disk array. The two paths to this disk are called /dev/[r]dsk/c1t33d0s2 and /dev/[r]dsk/c2t33d0s2. When the disk is used by VxVM, the vxdisk list <*daname*> command executed on this disk displays the following:

```
# vxdisk list c1t33d0s2
Device:    c1t33d0s2
devicetag: c1t33d0
type:      sliced
hostid:    gluon.veritas.com
```

```
disk:      name=rootdisk id=933470913.1042.gluon.veritas.com
group:     name=rootdg id=933470911.1025.gluon.veritas.com
flags:     online ready autoconfig autoimport imported
pubpaths:  block=/dev/vx/dmp/c1t33d0s6 char=/dev/vx/rdmp/c1t33d0s6
privpaths: block=/dev/vx/dmp/c1t33d0s7 char=/dev/vx/rdmp/c1t33d0s7
version:   2.1
iosize:    min=512 (bytes) max=2048 (blocks)
public:    slice=6 offset=1 len=35363560
private:   slice=7 offset=1 len=4711
update:    time=933471058 seqno=0.9
headers:   0 248
configs:   count=1 len=3448
logs:      count=1 len=522
Defined regions:
config   priv 000017-000247[000231]: copy=01 offset=000000 enabled
config   priv 000249-003465[003217]: copy=01 offset=000231 enabled
log      priv 003466-003987[000522]: copy=01 offset=000000 enabled
Multipathing information:
numpaths:   2
c1t33d0s2  state=enabled
c2t33d0s2  state=enabled
```

After the cable connected to controller c2 is removed and a vxdctl enable command is executed on the system, the display changes as follows:

```
# vxdisk list c1t33d0s2
Device:    c1t33d0s2
devicetag: c1t33d0
type:      sliced
hostid:    gluon.veritas.com
disk:      name=rootdisk id=933470913.1042.gluon.veritas.com
group:     name=rootdg id=933470911.1025.gluon.veritas.com
flags:     online ready autoconfig autoimport imported
pubpaths:  block=/dev/vx/dmp/c1t33d0s6 char=/dev/vx/rdmp/c1t33d0s6
privpaths: block=/dev/vx/dmp/c1t33d0s7 char=/dev/vx/rdmp/c1t33d0s7
version:   2.1
iosize:    min=512 (bytes) max=2048 (blocks)
public:    slice=6 offset=1 len=35363560
```

```
private:    slice=7 offset=1 len=4711
update:     time=933471058 seqno=0.9
headers:    0 248
configs:    count=1 len=3448
logs:       count=1 len=522
Defined regions:
 config    priv 000017-000247[000231]: copy=01 offset=000000 enabled
 config    priv 000249-003465[003217]: copy=01 offset=000231 enabled
 log       priv 003466-003987[000522]: copy=01 offset=000000 enabled
Multipathing information:
numpaths:   2
c1t33d0s2  state=enabled
NONAME     state=disabled
```

Similarly, the output of the vxdmpadm command also displays the NONAME string for the names of the paths that have been removed. For example, if the controller /sbus@1f,0/SUNW,socal@1,0/sf@1,0 is removed from the host, the vxdmpadm command displays the following:

```
# vxdmpadm getsubpaths ctlr=/sbus@1f,0/SUNW,socal@1,0/sf@1,0
NAME        STATE     TYPE       DMPNODENAME DA-TYPE  DA-SNO
========================================================
NONAME      DISABLED  -          c1t36d0s2 SEAGATE  SEAGATE_DISKS
NONAME      DISABLED  -          c1t49d0s2 SEAGATE  SEAGATE_DISKS
NONAME      DISABLED  -          c1t38d0s2 SEAGATE  SEAGATE_DISKS
NONAME      DISABLED  -          c1t37d0s2 SEAGATE  SEAGATE_DISKS
NONAME      DISABLED  -          c1t53d0s2 SEAGATE  SEAGATE_DISKS
NONAME      DISABLED  -          c1t34d0s2 SEAGATE  SEAGATE_DISKS
NONAME      DISABLED  -          c1t51d0s2 SEAGATE  SEAGATE_DISKS
NONAME      DISABLED  -          c1t54d0s2 SEAGATE  SEAGATE_DISKS
NONAME      DISABLED  -          c1t52d0s2 SEAGATE  SEAGATE_DISKS
NONAME      DISABLED  -          c1t33d0s2 SEAGATE  SEAGATE_DISKS
NONAME      DISABLED  -          c1t50d0s2 SEAGATE  SEAGATE_DISKS
NONAME      DISABLED  -          c1t32d0s2 SEAGATE  SEAGATE_DISKS
NONAME      DISABLED  -          c1t48d0s2 SEAGATE  SEAGATE_DISKS
NONAME      DISABLED  -          c1t35d0s2 SEAGATE  SEAGATE_DISKS
```

You should also note that when the path used to name the DMP metadevice (that represents a LUN) is removed, DMP uses the name of the alternate path to name the DMP metadevice.

For example, assume that there are two paths to a LUN called `c1t33d0s2` and `c2t33d0s2`. The DMP metadevice that represents the LUN is named `c1t33d0s2` and is in the `online` state. If the cable connected to controller `c1` is removed, followed by the `vxdctl enable` command, the DMP metadevice is renamed to `c2t33d0s2`. As a result, there are two metadevices corresponding to the same LUN visible to the user. Any one of these could be in the `error` state and the other in the `online` state. Note that this is not a problem. However, this results in some error messages when the `vxdisk list <daname>` command is used on the original *daname* that was used to name the DMP device. The `Multipathing Information` is not displayed to the user but the following error message is displayed:

```
vxvm:vxdisk: ERROR: <da_name> not a valid dmpnode disk access name.
```

The disk paths that are no longer accessible can be removed using the `rm` command.

DMP and JBOD Disk Array Support

The dynamic multipathing feature on JBOD drives enables DMP to multipath JBOD disk arrays connected to a system using VxVM. These JBOD disk arrays/disks can contain disks from any vendor. However, DMP can correctly detect multiple paths to these disks/disk arrays *only* if the following conditions are satisfied:

- The serial number field of the Standard SCSI INQUIRY DATA (bytes 36—47) should be "world wide unique" for any disk that is connected to a JBOD. This fact must be known to the user before the user attempts to configure DMP to add these disks to the JBOD category of DMP.

Note: If any disk DOES NOT provide a unique serial number, DMP *cannot* detect multiple paths to the disk correctly and results are unpredictable. Contact the JBOD disk/disk array vendors to get correct information about uniqueness of serial numbers on the disks.

• The JBOD disk arrays that require multipathing functionality from DMP should be of the Active/Active type. This means that all paths to all LUNs on the JBOD disk array should be simultaneously accessible through all paths, without causing any performance degradation.

Once it is known that disks connected to a JBOD disk array satisfy the above conditions, the following four steps must be followed to get these disks under the JBOD category of DMP:

1. Go to the `/kernel/drv` directory

2. Edit the file `vxdmp.conf` in that directory

3. You will see the following default entry in the `vxdmp.conf`:

   ```
   dmp_jbod="HITACHI"
   ```

   For example, assume that there is a JBOD connected to a machine that contains DG Clariion disks. You need to know the *exact* Vendor ID string returned by these disks to configure DMP. This is to ensure that the required disks are claimed under the JBOD category.

   The `format` command could be used to retrieve the Vendor ID string for a disk (for more details about the `format` command, check the `format`(1M) manual page on the system). The Vendor ID string is contained in bytes 8—15 of standard SCSI INQUIRY data returned by the disk. The `inquiry` option of the `format` command (when executed on a disk) returns the Vendor ID associated with a disk under the `Vendor`: field of the output. For example, a DG Clariion disk might return the Vendor ID

   ```
   string "DGC".
   ```

   This string must be included in the `vxdmp.conf` file as follows:

   ```
   dmp_jbod="DGC";
   ```

   If disks from multiple vendors are present in one JBOD disk array, vendor IDs of all vendors should be specified in the `vxdmp.conf` file. If this is not done, disks belonging to vendors that are not mentioned in the `dmp_jbod[]` array, are not multipathed by DMP and come under the category of OTHER DISKS. (Other Disks is a separate category of disks not multipathed by DMP and does not come under any of the supported categories in DMP.)

   The following example shows how to add entries for multiple vendors in the `dmp_jbod[]` array.

   ```
   dmp_jbod="DGC",
   ```

```
          "IBM",
          "FUJITSU";
```

The specified Vendor ID in the `vxdmp.conf` file cannot be more than eight characters long. If a Vendor ID string more than eight characters long is specified in the `vxdmp.conf` file, the following message is displayed on the console:

```
NOTICE: vxvm:vxdmp: Invalid Vendor ID = <vendor_id> specified
in vxdmp.conf file
Action: Remove or correct this entry in the vxdmp.conf file.
```

4. After the `vxdmp.conf` file has been edited to add the required entries, the system should be rebooted so that DMP can multipath the specified JBODs.

Using the ATF Driver With the VERITAS Volume Manager

In previous releases, the CLARiiON ATF driver and DMP could not co-exist. Co-existence is allowed in VM 3.0 and above.

For DMP to co-exist with the ATF driver, the `CLatf` package (Application Transparent Failover Driver for CLARiiON) must be installed before the `VRTSvxvm/SUNWvxvm` packages are installed on the system.

Note: In VM 3.0.4, ATF/DMP co-existence works *only* with the CLatf 1.3.1 package.

For DMP to co-exist with the ATF driver, the boot disk should not be under control of the ATF driver.

You can get DMP to co-exist with the ATF driver by following the procedure described here. Here is a summary of steps to be taken for three scenarios, followed by a detailed procedure:

1. Volume Manager is not installed.

    • Add ATF

    • Add Volume Manager

2. Volume Manager installed with DMP currently enabled.

    • Disable DMP

- Add the ATF driver

- Add Volume Manager 3.0.4 if appropriate

- Re-enable DMP

3. Volume Manager installed with DMP currently disabled:

- Add the ATF driver

- Add Volume Manager 3.0.4 if appropriate

- Enable DMP

DMP and ATF Driver Co-existence Procedure

The following warning messages may be displayed after the ATF driver is forceloaded by the operating system (this is required for DMP co-existence with the ATF driver). The messages are similar to the following:

```
WARNING: atf[ 0/ 0/ 1/ 0]:No peer found.
WARNING: atf[ 0/ 0/ 1/ 1]:No peer found.
WARNING: atf[ 0/ 0/ 1/ 2]:No peer found.
```

These messages are normal and appear only when a DG CLARiiON connected to the system is single pathed. If it is dual pathed, these messages should not appear. You can safely ignore these warnings. Co-existence of DMP and ATF is in no way affected by these messages.

To Disable DMP

1. `umount` all file systems created on Volume Manager volumes.

2. Stop the Volume Manager (use `vxdctl stop`).

3. Remove the `vxdmp` driver from the `/kernel/drv` directory:

   ```
   rm /kernel/drv/vxdmp
   ```

4. 2. Edit `/etc/system`, and remove the line:

   ```
   forceload: drv/vxdmp
   ```

5. 3. Remove the Volume Manager DMP files:

   ```
   rm -rf /dev/vx/dmp /dev/vx/rdmp
   ```

6. 4. Symbolically link `/dev/vx/dmp` to `/dev/dsk`:

    ```
    ln -s /dev/dsk /dev/vx/dmp
    ```

7. 5. Symbolically link `/dev/vx/rdmp` to `/dev/rdsk`:

    ```
    ln -s /dev/rdsk /dev/vx/rdmp
    ```

8. 6. Shut down the system to disable the DMP functionality:

    ```
    /usr/sbin/shutdown
    ```

9. 7. Reboot the system.

To Re-enable DMP After it has Been Disabled

Follow these instructions to re-enable the DMP driver once it has been removed:

1. Halt the system:

    ```
    shutdown -g0 -y -i0
    ```

2. Boot to single-user mode:

    ```
    boot -s
    ```

3. Copy `vxdmp` driver back:

    ```
    cp /kernel/drv/vxdmp.SunOS_`uname -r` \

    /kernel/drv/vxdmp

    cp /kernel/drv/sparcv9/vxdmp.SunOS_`uname -r` \

    /kernel/drv/sparcv9/vxdmp
    ```

4. Edit `/etc/system` and restore DMP entry:

Put the line `forceload: drv/vxdmp` just above the line
`forceload: drv/vxio`

5. Remove DMP links:

    ```
    rm /dev/vx/rdmp /dev/vx/dmp
    ```

6. Restore DMP directories:

```
mkdir /dev/vx/dmp /dev/vx/rdmp
```

7. Halt the system:

```
shutdown -g0 -y -i0
```

8. Perform a reconfiguration boot: `boot -r`

Moving From Alternate Pathing to DMP

If your system had been using Alternate Pathing (AP) software from Sun, DMP would not have been installed on your system. If you would like to move from using AP to using DMP, follow this procedure:

1. Remove the AP software from your system. (Refer to the documentation from Sun for the proper method to remove the AP software.)

2. Perform a reconfiguration reboot:

   ```
   boot -r
   ```

3. Copy the appropriate DMP driver into place:

   ```
   cp /kernel/drv/vxdmp.SunOS_`uname -r` /kernel/drv/vxdmp
   ```

   ```
   cp /kernel/drv/sparcv9/vxdmp.SunOS_`uname -r` \
   ```

   ```
   /kernel/drv/sparcv9/vxdmp
   ```

4. Add the line:

   ```
   forceload: drv/vxdmp
   ```

   into your `/etc/system` file, just above the line:

   ```
   forceload: drv/vxio
   ```

5. Remove the DMP links:

   ```
   rm /dev/vx/rdmp /dev/vx/dmp
   ```

6. Create the DMP directories:

   ```
   mkdir /dev/vx/rdmp /dev/vx/dmp
   ```

7. Halt the system:

   ```
   shutdown -g0 -y -i0
   ```

8. Perform a reconfiguration reboot:

```
boot -r
```

Dynamic Reconfiguration

Dynamic Reconfiguration (DR) is a feature available on some high end SUN Enterprise systems. The *board* to be reconfigured is a system board that contains disks controlled by Volume Manager (in addition to cpus, memory, and other controllers or I/O boards) that can be taken offline while the system is still running. You can dynamically reconfigure your system using one of the relevant procedures described below.

Note: The system board, disks and controllers could be configured in three different ways. Each of these configurations has a different procedure and may or may not require data to be offlined.

Domain Configuration Type 1

In this configuration, each of the disk arrays connected to the system board to be detached satisfy the following criteria:

- The disk array is multipathed.

- Paths to the disk array are connected to more than one system board that constitute the *domain*. In other words, even if one board is removed, there will still be one or more access paths to the disk array, from other system boards.

An example of such a configuration is shown in the illustration, below.

```
┌─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ┐
  UE10K Domain
│   ┌────────────────────┐          ┌────────────────────┐     │
    │ System Board 8     │          │ System Board 15    │
│   │                    │          │                    │     │
    │  ┌─────┐ ┌─────┐   │          │   ┌─────┐ ┌─────┐  │
│   │  │HBA1 │ │HBA2 │   │          │   │HBA3 │ │HBA4 │  │     │
    │  └─────┘ └─────┘   │          │   └─────┘ └─────┘  │
│   └────────────────────┘          └────────────────────┘     │

│                  ┌──────────┐  ┌──────────┐                  │
                   │  T300    │  │  A5000   │
│                  │Disk Array│  │Disk Array│                  │
                   └──────────┘  └──────────┘
├ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ┤
  Full controller pathnames:
│ HBA1=/sbus@61,0/SUNW,socal@1,0/sf@1,0  HBA3=/sbus@7d,0/SUNW,socal@1,0/sf@1,0
  HBA2=/sbus@61,0/SUNW,socal@2,0/sf@2,0  HBA4=/sbus@7d,0/SUNW,socal@2,0/sf@2,0
└ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ┘
```
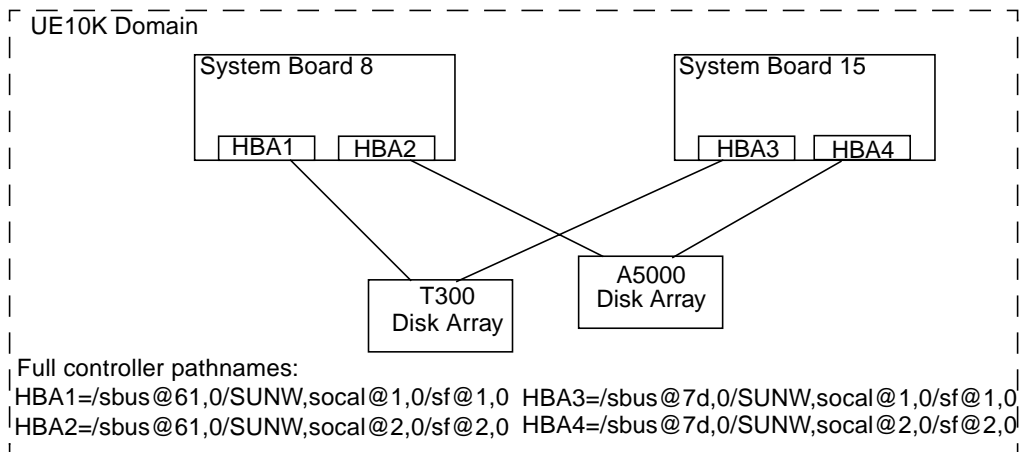
Follow these steps while reconfiguring a system board in the domain, for example the System Board 8 in UE10K domain in the figure, below.

1. Identify all the disk controllers that are on the system board

   To identify the full controller names of all disk controllers connected to the board being detached, use the commands:

   ```
   vxdmpadm listctlr all
   ```

   ```
   drshow board_no IO
   ```

   where *board_no* is the system board to be detached.

   For example, in Configuration 1, on the System Board 8, the disk controllers are: `/sbus@61,0/SUNW,socal@1,0/sf@1,0` and `/sbus@61,0/SUNW,socal@2,0/sf@2,0`.

2. For each disk controller that is on the System Board, run the following command:

   ```
   vxdmpadm disable ctlr=ctlr_name
   ```

   For example, in case of system board 8 in Figure 12, run the commands:

   ```
   vxdmpadm disable ctlr=/sbus@61,0/SUNW,socal@1,0/sf@1,0
   ```

   ```
   vxdmpadm disable ctlr=/sbus@61,0/SUNW,socal@2,0/sf@2,0
   ```

   This command insures that I/Os are shifted over from the specified controller to an alternate controller connected to one of the other system boards (System Board 15 in the figure, above). This is a blocking command and may take some time to return to the caller depending on the number of I/Os currently active through this controller.

   Once a controller has been disabled using the `vxdmpadm disable` command, the controller is no longer monitored. As a result, if the controller is subsequently removed as part of a Dynamic Reconfiguration operation, DMP will not be aware of it and subsequent `vxdmpadm` commands will continue to show the controller as DISABLED (even though it no longer exists.

3. Follow the steps in the Sun documentation to dynamically reconfigure the system board.

When the board is replaced (or attached):

4. Re-enable the Volume Manager configuration daemon by running the command:

   `vxdctl enable`

   At this point, `vxdmpadm` will still continue to show it as DISABLED.

5. Enable all the controllers on the board being attached using:

   `vxdmpadm enable ctlr=`*ctlr_name*

   for each controller on the board being replaced.

   For example, in the figure, above, for attaching System Board 8, run the following commands:

   `vxdmpadm enable ctlr=/sbus@61,0/SUNW,socal@1,0/sf@1,0`
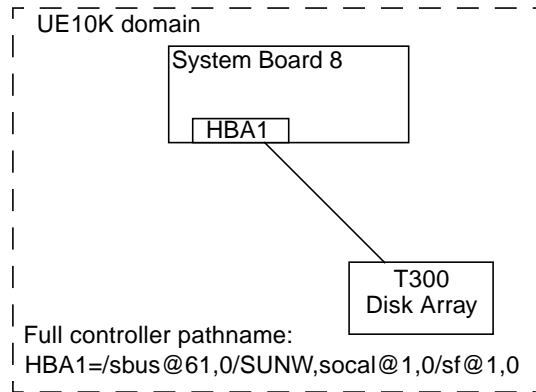
   `vxdmpadm enable ctlr=/sbus@61,0/SUNW,socal@2,0/sf@2,0`

This command insures that the controllers are now available for I/O to the Volume Manager.

Domain Configuration Type 2

In this case, the disk array(s) connected to the system board to be detached are not accessible via other system boards in the same domain. For example, see the illustration, below.

Note: This procedure may require data to be taken offline during the operation if it is not on redundant VxVM volumes (example: RAID5, Mirror etc.).

```
┌ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ┐
│  UE10K domain                        │
│         ┌──────────────────┐         │
│         │ System Board 8   │         │
│         │                  │         │
│         │ ┌──────┐         │         │
│         │ │ HBA1 │         │         │
│         └─┴──────┴───\──────┘         │
│                      \               │
│                       \              │
│                        \ ┌─────────┐ │
│                          │  T300   │ │
│                          │Disk Array││
│                          └─────────┘ │
│  Full controller pathname:            │
│  HBA1=/sbus@61,0/SUNW,socal@1,0/sf@1,0│
└ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ┘
```

The following procedure is used for this configuration:

1. All disks connected to the board being detached must be removed entirely from the VxVM control. Run the command `vxdiskadm` and choose the option:

   ```
   Remove a disk for replacement
   ```

   Run the following command on the disk:

   ```
   vxdisk rm daname
   ```

2. Follow the steps in the Sun documentation in order to detach the system board.

When the board is replaced (or attached):

3. Follow the steps in the SUN documentation to see the new drives/paths on this domain (for example, this can include the need to run utilities like `drvconfig` followed by `disks`, etc).

   Run the Volume Manager configuration daemon using the command

   ```
   vxdctl enable
   ```

4. Re-add the disks to Volume Manager by entering the command `vxdiskadm` and choose

   ```
   Replace a failed disk or removed disk
   ```

5. Restart all the volumes that were stopped as a result of the system board reconfiguration.

Domain Configuration Type 3

In this case, the system board being detached has disk arrays connected in the following fashion:

- Some disk arrays are accessible only through the system board being reconfigured.
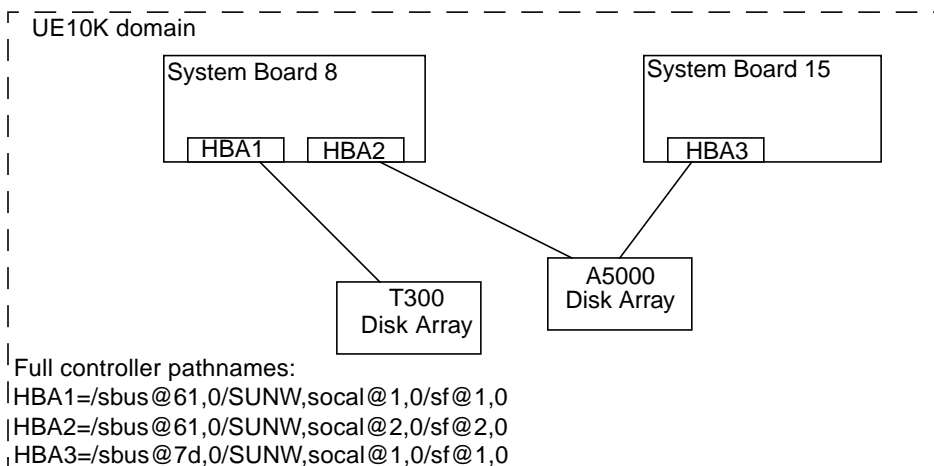
  For disk arrays that are accessible through one system board, use the procedure for Domain Configuration Type 2.

- The remaining disk arrays are accessible through the other system board(s) as well.

For disk arrays that are accessible through multiple system boards, use the procedure for Domain Configuration Type 1.

An example configuration is shown in the illustration, below.


```
UE10K domain
   System Board 8                          System Board 15


     HBA1    HBA2                            HBA3


                                      A5000
                                      Disk Array
                    T300
                    Disk Array
Full controller pathnames:
HBA1=/sbus@61,0/SUNW,socal@1,0/sf@1,0
HBA2=/sbus@61,0/SUNW,socal@2,0/sf@2,0
HBA3=/sbus@7d,0/SUNW,socal@1,0/sf@1,0
```

In the illustration, above, to detach or attach System Board 8, the procedure is:

1. For the T300 disk array which is accessible only through System Board 8, follow the procedure for Domain Configuration Type 2.

2. For the A5000 disk array which is accessible through multiple system boards, that is System Board 8 and System Board 15, follow the procedure for Domain Configuration Type 1.

Using the Sun StorEdge A3000 Array With the VERITAS Volume Manager

Note: The information in this section is current at the time of this printing, but is subject to change with new versions of the Sun StorEdge A3000 Array software.

The Sun StorEdge A3000 Array subsystem with the Sun StorEdge A3000 Array RAID Manager software is compatible with Volume Manager Release 3.0.4. To ensure compatibility between the Sun StorEdge A3000 Array and Volume Manager, you must do the following:

- The Sun StorEdge A3000 Array and Volume Manager installation sequence must be followed exactly as documented in the section "Installation of the Volume Manager With the Sun StorEdge A3000 Array".

- Volume Manager volumes configured using devices from the Sun StorEdge A3000 Array subsystem cannot be part of the root disk group, `rootdg`. Configure StorEdge A3000 Array devices only to non-`rootdg` disk groups.

- When Volume Manager encapsulates a device that has a mounted file system and entry in the `/etc/vfstab` file, VxVM updates the `/etc/vfstab` entry for that device to contain the Volume Manager device node name. Volume Manager is not compatible with the Sun StorEdge A3000 Array RAID Manager device node naming convention, but is compatible with the Solaris operating system device node naming convention.

  To encapsulate the Sun StorEdge A3000 Array devices that have been configured with file systems and have entries in the `/etc/vfstab` file, you should not use the Sun StorEdge A3000 Array RAID Manager generated disk device nodes. For example (see the StorEdge A3000 Array documentation for the device nodes):

      /dev/rRAID_module01/0s[0-7]

  and

      /dev/RAID_module01/0s[0-7].

  Instead, use the standard Solaris operating system device nodes. For example (see the Solaris operating system documentation for the standard device nodes):

      /dev/rdsk/c3t4d0s0 and /dev/dsk/c3t4d0s0.

  For example, if you have an `/etc/vfstab` entry similar to this:

  /dev/RAID_module01/0s0 /dev/rRAID_module01/0s0 /mountpt ufs 2 yes -

  use the Solaris operating system device node that represents this device:

      /dev/dsk/c3t4d0s0 /dev/rdsk/c3t4d0s0 /mountpt ufs 2 yes -

- Follow the Sun StorEdge A3000 Array controller error recovery procedures as documented by the Sun StorEdge A3000 Array subsystem documentation.

> **CAUTION!** Sun StorEdge A3000 Array error recovery procedures must be followed according to the Sun StorEdge A3000 Array subsystem documentation. Improper error recovery procedures can cause incompatibility with the VERITAS Volume Manager.

Installation of the Volume Manager With the Sun StorEdge A3000 Array

The Sun StorEdge A3000 Array installation sequence must be followed as documented in the Sun StorEdge A3000 Array Subsystem Release Notes, the Sun StorEdge A3000 Array System Manual, and the Sun StorEdge A3000 Array RAID Manager Installation Guide.

> **CAUTION!** Any deviation of the installation sequence causes the StorEdge A3000 Array to be incompatible with the VERITAS Volume Manager.

The Volume Manager should be installed only after the following steps have been completed and validated:

1. The Sun StorEdge A3000 Array subsystem is properly attached to the host computer.

2. The Sun StorEdge A3000 Array RAID Manager software is properly installed.

3. The Sun StorEdge A3000 Array Logical Unit Numbers (LUNs) are properly configured using the Sun StorEdge A3000 Array RAID Manager software.

4. The host system is rebooted using the `-r` flag (`reboot`(1M)) to rescan for attached devices upon boot.

5. Upon reboot, the Sun StorEdge A3000 Array RAID Manager has recognized the Sun StorEdge A3000 Array configured devices and has created the appropriate Solaris operating system device nodes (see "VERITAS Volume Manager and Sun StorEdge A3000 Array Device Nodes").

6. Install the Volume Manager software package.

VERITAS Volume Manager and Sun StorEdge A3000 Array Device Nodes

After installing and configuring the Sun StorEdge A3000 Array subsystem, the host needs to be rebooted using the `-r` flag (`reboot`(1M)) to rescan for attached devices. Upon reboot, the Sun StorEdge A3000 Array scripts and Solaris operating system startup scripts create device nodes.

After the host system reboot, the original set of device nodes should be seen (the device nodes that existed before the Sun StorEdge A3000 Array subsystem was installed), plus the new device nodes for each StorEdge A3000 Array LUN. If there are more device nodes created than attached devices (including StorEdge A3000 Array LUNs), then do not install the VERITAS Volume Manager until the proper device nodes exist on the host system.

CAUTION! The Sun StorEdge A3000 Array startup scripts must properly configure and create device nodes. Improper host system device nodes cause incompatibility with the VERITAS Volume Manager.

Benefits in Using VERITAS Volume Manager With the Sun StorEdge A3000 Array

This section describes some of the benefits in using the host-based Volume Manager software and controller-based StorEdge A3000 Array hardware. When the VERITAS Volume Manager is used in conjunction with the Sun StorEdge A3000 Array, availability, performance and manageability of the combined configuration can be improved.

Improved Manageability

The online reconfiguration capabilities of the VERITAS Volume Manager can be used to enhance the performance monitoring, tuning, and reconfiguration of the Sun StorEdge A3000 Array. These capabilities can simplify administration while increasing uptime and performance.

Performance Analysis and Online Load Balancing Between LUNs

Using the Volume Manager statistics-gathering capability allows administrators to analyze the I/O loads and responsiveness of volumes in the configuration and to move storage between disks. Since each LUN within the Sun StorEdge A3000 Array looks like a single disk to the host, and therefore to Volume Manager, it is not possible to identify the load on each spindle, or to tune the layout of the LUN itself. However, performance imbalances between arrays, or between LUNs in the same array, may be detected and alleviated. This can be done by using the Volume Manager performance analysis and online reconfiguration capability.

Increasing Capacity Online by LUN Concatenation

The size of a single file system or database tablespace is limited by the maximum size of a single LUN in a controller-based RAID subsystem. To create very large file systems or tablespaces, administrators can use Volume Manager to combine LUNs from multiple RAID controllers into a larger "super-LUN" volume.

Online Data Movement Between LUNs

The backup/reconfigure/reload cycle required to change the layout of hardware RAID LUNs causes data access interruption and possible data loss. To reduce the likelihood of data loss, if extra disks are available, administrators can choose to construct the destination LUN according to the desired parameters, then copy data directly from the old location to the new one. However, even this requires the interruption of data access, since it is necessary to prevent changes to the old copy of the data after it has been copied to the new location.

If the data copying is performed by adding the new Sun StorEdge A3000 Array LUN as a Volume Manager mirror of the data, however, all writes are delivered to all mirrors, keeping all copies up-to-date at all times. Once the mirror synchronization is complete, the mirror set can be separated, and the original LUN can be removed or used for other data. The data is kept online without interrupting access throughout the operation.

Stable Backup

The method described above for data movement can also be used to provide consistent stable backup without interrupting user access. An additional mirror of data can be created, or an existing mirror can be detached from the updating applications, and that data instance can be backed-up to tape or other offline storage pool (see "Preparing a Volume to Restore From Backup" in the *VERITAS Volume Manager Command Line Interface Administrator's Guide.* This guarantees a backup that is internally consistent at a single point in time. Upon completion of backups, the additional mirror space can be freed, or the mirror reattached to the live data volume and resynchronized.

Dynamic Partitioning

The dynamic partitioning capability of VERITAS Volume Manager is useful when used with the large disks presented by the system for each LUN. The Solaris operating systems have hard limits on the number of slices into which a disk can be partitioned (7 partitions plus the full-disk partition). These limits can make it impossible to divide a large LUN into the number of separate slices needed for file systems or databases.

Sun StorEdge A3000 Array Supported Configurations

VERITAS Volume Manager supports the Sun StorEdge A3000 Array subsystem in the following configurations:

- VxVM RAID-1 (Mirroring) and Sun StorEdge A3000 Array
- VxVM RAID-0 (Striping) and Sun StorEdge A3000 Array

For more information about mirroring and striping, see the *VERITAS Volume Manager Administrator's Reference Guide* and the *VERITAS Volume Manager Getting Started Guide.*

Sun StorEdge A3000 Array Unsupported Configurations

The following are Volume Manager and Sun StorEdge A3000 Array configurations that are not supported:

Volume Manager RAID-5 and Sun StorEdge A3000 Array RAID-5 Configuration

Using RAID-5 at both levels in the I/O subsystem can result in poor performance for no significant gain in reliability or availability. Use of this configuration is not supported.

Volume Manager Hot-Relocation and Sun StorEdge A3000 Array Hot Sparing

Volume Manager hot-relocation allows the host system to automatically react to I/O failures on redundant (mirrored or RAID-5) VxVM objects and restore redundancy and access to these objects. Sun StorEdge A3000 Array hot sparing allows the StorEdge A3000 Array to automatically react to I/O failures internal to the StorEdge A3000 Array to restore access to StorEdge A3000 Array LUNs.

If a disk failure occurs within the Sun StorEdge A3000 Array, and the array hot sparing is configured, then the StorEdge A3000 Array provides the disk failure redundancy. If data redundancy is provided by a Volume Manager RAID-5 or mirror configuration, then Volume Manager hot-relocation can also provide disk or partial disk failure redundancy protection. The most complete level of disk redundancy is achieved with both Volume Manager hot-relocation and Sun StorEdge A3000 Array hot sparing enabled.