Storage Multi-Pathing choices in HP-UX Serviceguard environments

Version 1.0

Abstract

Serviceguard protects users and applications from a variety of failures (e. g. software, system, and network failure to name a few) by forming a cluster of multiple nodes and providing sufficient redundancy to remove single points of failure from the configuration. In general, Serviceguard directly provides the failure protection while in specific areas it relies on other technology to add redundancy. Highly available storage is usually provided by hardware mirroring with storage arrays or software mirroring functionality of volume managers, while Serviceguard coordinates the access to the highly available storage among the cluster nodes. When physically redundant hardware paths between the cluster nodes and the storage devices are present, multi-pathing software provides protection against path failure. The focus of this whitepaper is to provide the reader with an overview of what multi-pathing solutions are available for various HP-UX Serviceguard environments and decision points for choosing between them if more than one multi-pathing solution is feasible in a specific configuration.

Executive summary

Serviceguard clusters rely on multi-pathing software to protect against storage path failures. Either the Operating System, as in HP-UX 11i v3, or additional software provides this functionality. Customers deploying Serviceguard clusters on HP-UX have several choices to configure systems that are resilient to storage path failures. Depending on the version of HP-UX – either 11i v2 or 11i v3 – those choices are different.

HP-UX 11i v2 itself does not offer a built-in multi-pathing solution, so customers need to utilize additional multi-pathing software that protects against storage path failures and often also provides performance improvements by dynamically balancing the I/O load between the available paths. These solutions are either volume manager or array based:

Table 1.

Volume manager based	**Array based**
LVM Alternate Links **(PVlinks)**	HP StorageWorks Secure Path
VxVM/CVM Dynamic Multipathing **(DMP)**	EMC PowerPath

These multi-pathing solutions can generally coexist on the same server and storage system but multi-pathing functionality to an individual LUN can only be provided by one solution. HP-UX 11i v3 provides built-in native multi-pathing that generally makes add-on multi-pathing solutions redundant.

- While LVM PVlinks are still supported with 11i v3 – they utilize native multi-pathing underneath – it is no longer required nor recommended to configure LVM alternate links.
- DMP [1]
  - with VxVM 4.1 coexists with native multi-pathing; it acts in pass-through mode allowing HP-UX to perform multi-pathing and load-balancing.
  - with VxVM 5.0 will coexist with native multi-pathing;
    - will handle the multi-pathing by default (native multi-pathing is pass-through)
    - can be configured in pass-through mode (native multi-pathing is active)
  - CVM 5.0 volumes require DMP to handle multi-pathing (native multi-pathing is in pass-through mode)
- Secure Path is not supported on HP-UX 11i v3.
- PowerPath is currently not supported on HP-UX 11i v3.

[1] DMP 5.0 on HP-UX 11i v3 information is preliminary since the product was not released at the time of writing this documentation

Definition of terms

Table 2.

Term	Definition
CFS	Serviceguard Cluster File System based on Veritas CFS from Symantec; included in some of the HP Serviceguard Storage Management Suite bundles.
CVM	Veritas Cluster Volume Manager from Symantec; included in some of the HP Serviceguard Storage Management Suite bundles.
DMP	Dynamic Multipathing; Symantec's multi-pathing product for VxVM and CVM volume manager
DSF	Device Special File; a special file used to interface with devices – e. g. storage systems
PP	EMC PowerPath
SP	HP StorageWorks Secure Path
SMS	Serviceguard Storage Management Suite, Serviceguard bundles based on Veritas volume manager and file system products.
VxFS	Veritas File System from Symantec; included in the Serviceguard Storage Management Suite. Also known as Online-JFS individually or part of the HP-UX MCOE
VxVM	Veritas Volume Manger from Symantec; included in the Serviceguard Storage Management Suite.

Multi-pathing product reference documentation

Most of the information in this document is retrieved from other documentation. This section refers to individual manuals and white papers that provide more detailed information on specific multi-pathing solutions and their interoperability with each other. The references point to documents which were current at the time of writing this white paper. The reader should find and read up-to-date versions of these documents which apply to the configuration of interest. The following list will guide through that process:

- **VERITAS Volume Manager 4.1 Administrator's Guide (HP-UX 11i v2)**

  (http://docs.hp.com/en/5991-1838/5991-1838.pdf)

  Chapter 3 describes administering DMP and also provides an overview of its operation. Chapter 2 has a section on **Third Party Driver Coexistence** which overviews interoperability in very general terms, and also has various references to configuring "foreign" devices when use of other multi-pathing software is desired. A section in chapter 10 describes policies which react to path failures from 1 or more nodes in a cluster. Chapter 12 describes a DMP tunable kernel parameter.
- **Dynamic Multipathing,**

  Improving data availability and I/O performance through intelligent management of storage access paths.

  (http://eval.veritas.com/mktginfo/products/White_Papers/Storage_Server_Management/sf_dmp_wp.pdf) An alternate link: Veritas Storage Foundation White Papers contains a good overview of how DMP works as well as some in-depth discussion of interoperability with other multi-pathing solutions, especially PowerPath. The subsection *PowerPath pseudo mode* does not apply to HP-UX. An updated version of the white paper with DMP 5.0 information can be found at:

http://eval.symantec.com/mktginfo/enterprise/white_papers/ent-whitepaper_vsf_5.0_dynamic_multi-pathing_05-2007.en-us.pdf

- **VERITAS Volume Manager 4.1 Hardware Notes (HP-UX 11i v2)**

  (http://docs.hp.com/en/5991-1839/5991-1839.pdf)

  Identifies supported HP-UX server models and storage devices. Describes conditions and limitations for various devices. Describes coexistence with PowerPath. Describes support procedures for new disk arrays. Provides an overview of the `vxdmpadm` command.

- **Device Support for Veritas Products on HP-UX**

  (http://docs.hp.com/en/5992-0563/5992-0563.pdf)

  This document describes the various devices that support the Veritas products from Hewlett-Packard and Symantec on HP-UX systems. Support for storage devices varies according to the product.

- **Veritas 4.1 Installation Guide (HP-UX 11i v3)**

  Chapter 8 "VxVM 4.1, New Style DSFs and HP-UX Native Multipathing" explains how VxVM and DMP 4.1 interoperate with the new mass storage subsystem.

- **SAN Design Guide**

  (http://spock.corp.hp.com/pages/san_design.aspx)

  Configuration reference for HP storage. The **HP-UX SAN configuration rules** table identifies which HP devices support multiple controllers and which support SP, DMP, and PVlinks. The **HP-UX multipathing coexistence support** table shows which devices support which multi-pathing solutions, and which device/multi-pathing combinations can coexist on the same SAN.

- **HP StorageWorks Secure Path 3.0F Service Pack 2 for HP-UX 11i v1.0 and 11i v2.0, installation and reference guide**

  (http://h20000.www2.hp.com/bc/docs/support/SupportManual/c00248499/c00248499.pdf)

  Other links to this document: Storage Software and Multi-path Software.

  A comprehensive user guide for both Secure Path for Active-Passive disk arrays and Secure Path for Active-Active disk arrays.

- **PowerPath for HP-UX Version 4.5 Installation and Administration Guide (P/N 300-002-838, Rev A03)**

  Describes how to install/uninstall PowerPath with VxVM with specific array types. Describes when to use PVlinks with PowerPath, and migrating from Secure Path to PowerPath. To obtain a copy of this document, registration is required at: http://powerlink.emc.com

- **PowerPath Version 4.5 Product Guide (P/N 300-002-834 REV A03)**

  Describes the concepts and features of PowerPath. To obtain a copy of this document, registration is required at: http://powerlink.emc.com

- **Third party mass storage devices**

  (http://www.hp.com/products1/serverconnectivity/mass_storage_devices.html)

  Identifies the storage devices validated for use with HP-UX. Specifies which of these is not supported with LVM PVlinks, and which require third party multi-pathing software, such as EMC PowerPath for Serviceguard support.

- **New HP-UX 11i v3 Mass Storage Stack**

  To learn more about the new mass storage on HP-UX 11i v3 see the following documents on the HP documentation websites (http://docs.hp.com and http://docs.hp.com/en/oshpux11iv3.html#Storage%20Area%20Management):
  - HP-UX 11i v2 to 11i v3 Mass Storage Stack Update Guide
  - HP-UX 11i v3 Mass Storage Device Naming
  - HP-UX 11i v3 Mass Storage I/O Performance Improvements

- HP-UX 11i v3 Native Multi-Pathing for Mass Storage
- HP-UX 11i v3 Persistent DSF Migration Guide
- LVM Migration from Legacy to Agile Naming Model
- Overview: Next Generation Mass Storage Stack
- Product Brief: The Next Generation Mass Storage Stack
- scsimgr SCSI Management and Diagnostics Utility on HP-UX 11i v3
- Migrating from HP StorageWorks Secure Path for Active-Active Disk Arrays to Native Multipathing in HP-UX 11i v3

  (http://docs.hp.com/en/AA-RWF2A-TE/AA-RWF2A-TE.pdf)

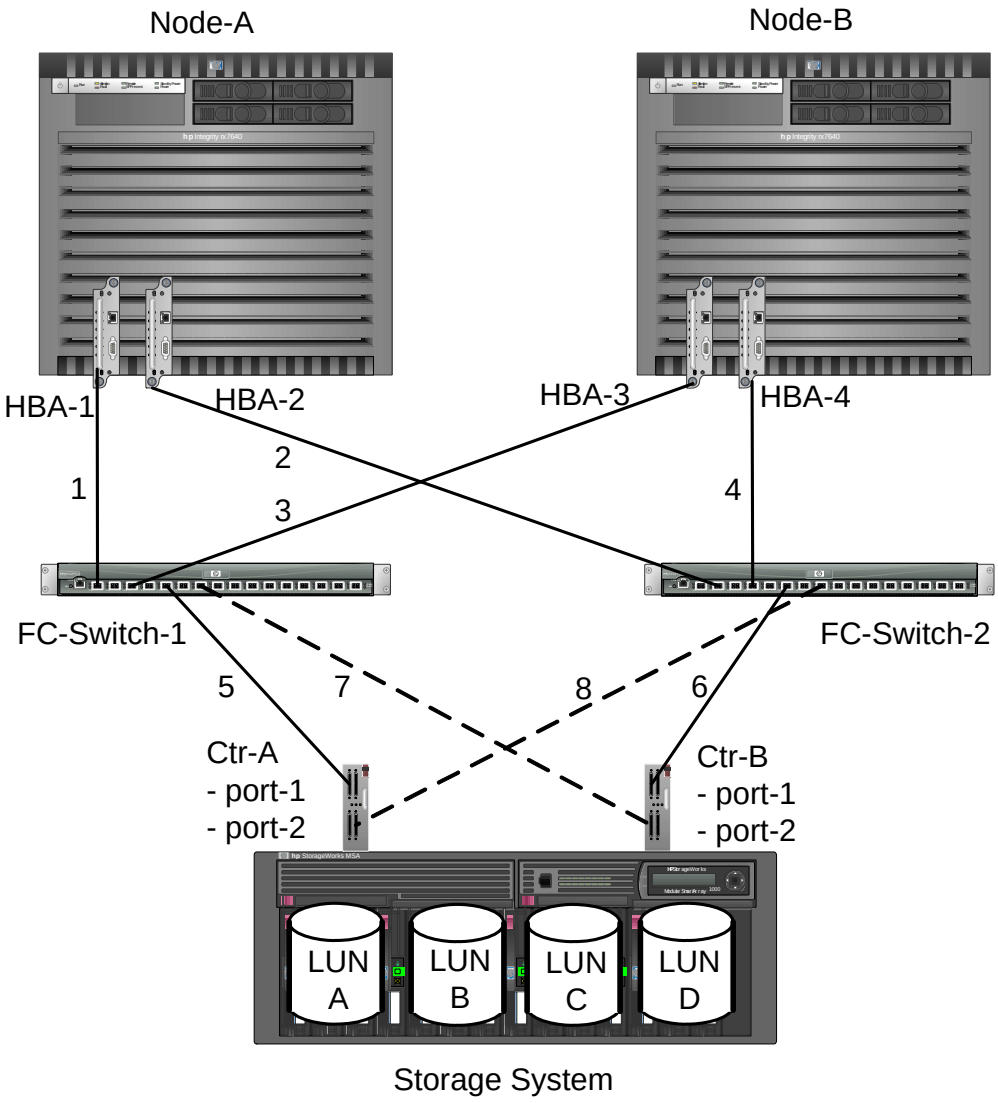
Typical Cluster Storage Configuration

Physically redundant hardware paths between the cluster nodes and the storage devices are required for multi-pathing software to provide path failure protection. Redundancy should be implemented on all 3 layers as shown in the configuration below:

- Each node is equipped with two Fiber Channel Host Bus Adapters (HBA).
  HBA-1, HBA-2, HBA-3, and **HBA-4**
- The Storage Area Network (SAN) is built with 2 Fiber Channel switches.
  FC-Switch-1 and **FC-Switch-2**

  The fiber channel connections are numbered 1 through 8 (connection 7 and 8 shown in dashed lines are only available on arrays with multiple ports per controller)
- The high available storage system is connected to the SAN with 2 controllers
  Ctr-A and **Ctr-B**
  - Each controller might have multiple ports connected to the SAN.**port-1** and **port-2**

Figure 1: Typical Cluster Storage Configuration



While this configuration examples consists of sufficient FC connectivity from a high availability point of view, additional paths might be required if the application has specific storage performance requirements.

Individual storage systems distinguish from each other in several ways. Some have only 2 controllers, others have more. Some have controllers with only one port to the SAN while others have more. Depending on the specific storage system the optimal cabling might look different than in the example shown in figure 1.

An important attribute of a storage system is whether I/O operations to a specific LUN are handled concurrently through multiple controllers or just a single controller, while the other controller(s) act as standby.

Active/Passive Storage Systems

With Active/Passive storage systems, one controller is assigned to a LUN as primary controller (owner of the LUN) and handles all the I/O requests to it. The other controller – or multiple other controllers, if

available – acts as standby controller. The standby controller of a LUN only issues I/O requests to it, if their primary controller failed. To achieve load balancing it makes sense to distribute the LUNs evenly among the controllers. In the configuration example above, each of the two controllers would "own" two LUNs.

Most active/passive storage systems require a controller failover to switch between controllers. A failover can be caused by simply accessing a LUN via the controller that is not currently the owner. This is not a big issue for clusters which use LVM or VxVM with exclusive activation by only one node at a time. However for concurrent access from multiple nodes with SLVM, CVM or CFS, all nodes must coordinate primary LUN ownership so they use the same controller for the same LUNs. HP SecurePath does this coordination for EVA active/passive models (EVA 3000 [HSV100] and EVA 5000 [HSV110]). EMC PowerPath coordinates primary LUN ownership among the cluster nodes for CLARiiON arrays.

Active/Active Storage Systems

With Active/Active storage systems, multiple controllers can issue I/O requests to an individual LUN concurrently. This enables multi-pathing software to provide load balancing between the paths in addition to path failure protection. The multi-pathing solutions offer different load balancing policies. Depending on the type of array and specific I/O pattern of the application, users can choose an appropriate policy.

Most storage systems implement a read-ahead mechanism to improve sequential read performance by efficiently using the array's cache and internal components. Key for those read-ahead mechanisms is to successfully identify sequential read operations. If the disk array recognizes sequential access patterns and reads ahead, subsequent requests may be satisfied from data pre-read into cache.

I/O requests travel through multiple layers of software and hardware from the application which issued the initial request to the storage system and back. Load balancing multi-pathing software and volume managers are two of the layers which need to be configured carefully to allow those read-ahead mechanisms to work successfully with sequential read operations. The key during the configuration is not to divide the sequential read requests coming from a layer above into multiple smaller requests – which might then be too small for the array to be recognized or handled as sequential I/O – before sending it down to the next layer. The following two paragraphs illustrate this matter based on the example configuration from figure 1:

- If you have two active paths to a LUN (e. g. from Node-A to LUN-A)

  Node-A -> HBA-1 -> FC-1 -> Ctr-A; port-1 -> LUN-A (FC cable 1 and 5)

  Node-A -> HBA-2 -> FC-2 -> Ctr-B ; port-1 -> LUN-A (FC cable 2 and 6)

  you want to choose a multi-pathing load balancing policy which sends sufficient I/Os down one path before it switches to the other path. Each storage system might have a different threshold in regards of how many sequential read requests (I/O blocks) it needs to receive in order to trigger a read-ahead. Symantec DMP for instances addresses this matter by allowing one to configure the number of blocks (DMP_PATHSWITCH_BLKS_SHIFT) being send over one path before switching to the next path with the "balanced path routing" policy. For HP StorageWorks XP 10000 / 12000 arrays, tests have shown that increasing this parameter from the default of 2 MB to a value of 32 MB increased the sequential read performance.
- If you have 4 LUNs on different physical media (array groups on HP StorageWorks XP Disk Arrays) you would want to stripe a logical volume over all 4 LUNs with a stripe size that is large enough to allow the array to recognize sequential reads and efficiently use the read-ahead. Setting the stripe size too small would actually not just prevent the read-ahead, it would actually convert the sequential I/O into multiple random I/Os. A stripe size of 4 MB proved to be nondestructive for the read-ahead algorithm of the HP StorageWorks XP 10000 / 12000 arrays.

The implementation of read-ahead mechanism varies between different types of arrays. It is advisable to research this topic and consider the array specific characteristics only for environments with a predominant sequential read I/O pattern.

Active/Active Storage Systems further divide into two categories:

- **Asymmetric Active/Active (AAA)** storage systems;

  With theses type of arrays, one controller is assigned to each LUN as a **preferred** controller. Each controller can have multiple ports – our example above showed 2 ports per controller. I/O requests reaching the storage system through ports of the preferred controller of a LUN will be sent directly to the LUN. I/O requests arriving at the non-preferred controller of a LUN will be first forwarded to the preferred controller of the LUN. These arrays are also called Asymmetric Logical Unit Access (ALUA) compliant devices.

  The EVA 4x00/6x00/8x00 and the EVA3000/5000 (HSV101/111) with active/active firmware are ALUA compliant arrays. Multi-pathing software can query ALUA compliant arrays to load balance only between paths connected to the preferred controller and use the paths to the non-preferred controller for automatic path failover if all of the paths to the primary controller fail.

- **Symmetrical Active/Active (SAA)** storage systems;

  These type of arrays do not have a primary or preferred controller per LUN. I/O requests can be issued over all paths mapped to a LUN. The current models of the HP StorageWorks XP Disk Array family are symmetrical active/active arrays.

Multi-Pathing solutions on HP-UX 11i v2

HP-UX 11i v2 does not include native (as part of the OS) multi-pathing software. However, there are a number of choices available in a Serviceguard cluster to protect against storage path failure. These multi-pathing solutions are either tightly integrated with a volume manager or they are storage system specific. Both types have their pros and cons.

On HP-UX 11i v2 a device special file (DSF) represents a path to a device. Devices that are connected to the server via multiple paths have multiple device files – one per path.

The following table provides an overview of the multi-pathing solutions available in Serviceguard clusters on HP-UX 11i v2.

Table 3: Multi-pathing solutions available on HP-UX 11i v2 Serviceguard clusters

	automatic path failover	load balancing across multiple active paths	Volume Manager independent	Storage System independent	Included in the software stack[2]
LVM PVlinks	yes	no[3]	no	yes[4]	yes
Symantec DMP	yes	yes	no	yes[5]	yes
HP Storage Works Secure Path	yes	yes	yes	no	no
EMC PowerPath	yes	yes	yes	no	no

[2] Software stack in this regard is either the Serviceguard Storage Management Suite or the volume manager included in the OS.
[3] PVLinks don't provide automatic load balancing, but static load balancing can be implemented as shown in the example in Figure 2.
[4] PVLinks are not supported in Serviceguard clusters with active/passive devices.
[5] DMP is only available with arrays supported by VxVM/CVM; some arrays might require an ASL for support

Volume Manager based

Serviceguard supports two different volume managers which also provide a multi-pathing solution. If customers choose LVM or SLVM as their volume manager, Alternate Links (PVlinks) are available. If they choose VxVM or CVM, Symantec Veritas Dynamic Multipathing (DMP) protects against path failures.

LVM Alternate Links (PVlinks)

LVM supports PVlinks to a device to allow continued access to the device if the primary link fails. Such a device must have multiple active paths from the host. In HP-UX 11i v2, LVM's multi-pathing solution increases the data availability by providing automatic path failover but does not allow multiple paths to be used simultaneously to increase I/O performance.

LVM uses the device files created during HP-UX I/O boot phase in `/dev/rdsk` and `/dev/dsk` directories (e. g.: `/dev/dsk/c6t0d2` and `/dev/dsk/c8t0d2`). The path named first at volume group creation time becomes the primary path. All paths added later become alternate paths. If the primary path fails, LVM automatically re-routes I/Os to an alternate path. After a failed primary path becomes available again, LVM automatically switches back to it.

A good practice is to manually distribute the load statically between multiple links at volume group creation time. This is achieved by alternating the paths when naming the physical volumes. Figure 2 shows an example in which 4 devices are accessible through controller c6 and c8.

Figure 2: Static load balancing with PVlinks

```
vgcreate vgdb  /dev/dsk/c6t0d0 /dev/dsk/c8t0d1 /dev/dsk/c6t0d2 /dev/dsk/c8t0d3
vgextend vgdb  /dev/dsk/c8t0d0 /dev/dsk/c6t0d1 /dev/dsk/c8t0d2 /dev/dsk/c6t0d3

vgdisplay -v vgdb | grep "PV Name"
   PV Name                     /dev/dsk/c6t0d0
   PV Name                     /dev/dsk/c8t0d0       Alternate Link
   PV Name                     /dev/dsk/c8t0d1
   PV Name                     /dev/dsk/c6t0d1       Alternate Link
   PV Name                     /dev/dsk/c6t0d2
   PV Name                     /dev/dsk/c8t0d2       Alternate Link
   PV Name                     /dev/dsk/c8t0d3
   PV Name                     /dev/dsk/c6t0d3       Alternate Link
```

The PVlinks feature comes with LVM at no additional charge and works with all devices supported by Serviceguard, but only for those that are configured to be under the control of LVM.

VxVM/CVM Dynamic Multi-pathing (DMP)

Symantec Veritas DMP increases the data availability and I/O performance by dynamically load balancing I/O requests across multiple storage paths for faster throughput. There are several different load balancing policies available. Chapter 2 and 3 of VERITAS Volume Manager 4.1 Administrator's Guide (HP-UX 11i v2) give an overview of DMP and describe the administration of it. The Dynamic Multipathing white paper goes into further details about the different load balancing policies and interoperability with other multi-pathing solutions.

VxVM's configuration daemon creates its own path-independent device files in `/dev/vx/rdmp` and `/dev/vx/dmp` at boot up time. It uses one device file per LUN or disk – not per path. Multiple paths (SUBPATH) are associated with VxVM/DMP devices files. Figure 3 shows how to identify which paths are associated with a specific device. In the example below, the disk group "`dbDG`" consists of two subdisks "`dbDG01`" and "`dbDG02`" which reside on disk "`c6t0d0`" and "`c6t0d1`" (DANAME column).

Each of these disks has two active paths – one through c6 and the other through c8 (SUBPATH column).

Figure 3: VxVM device file to subpath mapping

```
# vxdisk -g dbDG path
SUBPATH         DANAME          DMNAME          GROUP           STATE
c6t0d0          c6t0d0          dbDG01          dbDG            ENABLED
c8t0d0          c6t0d0          dbDG01          dbDG            ENABLED
c6t0d1          c6t0d1          dbDG02          dbDG            ENABLED
c8t0d1          c6t0d1          dbDG02          dbDG            ENABLED
```

The column "SUBPATH" lists HP-UX device files (paths) in directory "`/dev/(r)dsk`" and the column "DANAME" lists VxVM/DMP device files from "`/dev/vx/(r)dmp`" directory. VxVM 4.1 also offers the option to use "enclosure-based" naming schema for VxVM device files. With enclosure-based naming the "DANAME" column could have entries like `XP12K0_1` and `XP12K0_2`, or `EVA80000_1` and `EVA80000_2` instead of `c6t0d0` and `c6t0d1`.

DMP is tightly integrated with VxVM and CVM. As part of HP's Serviceguard Storage Management Suite, DMP comes at no additional charge and works with most storage devices supported with Serviceguard and VxVM/CVM. The VERITAS Volume Manager 4.1 Hardware Notes (HP-UX 11i v2) and Device Support for Veritas Products on HP-UX provide details about the disk array configurations supported with DMP.

Active/active arrays that fulfill certain requirements are supported without any additional software. However, for those arrays that require more specialized handling, DMP's architecture provides for array-specific array support libraries (ASLs) for discovery and configuration, and kernel mode array policy modules (APMs) to perform array-specific functions in the I/O path. These libraries/modules allow DMP to also support some active/passive storage systems in a cluster. ASLs and APMs for arrays that are already on the market are usually included in and ship with DMP. To provide the ability to support new arrays as they become available, ASLs and APMs can be added later. New ASLs can be downloaded from Symantec's support website.

The `vxddladm listsupport` command lists all active ASLs installed on a system. It is a best practice for DMP 4.1 and earlier versions to deactivate unused ASLs to improve the speed of the device discovery process. A faster device discovery process also leads to a faster system boot and cluster join time. An unused ASL is an ASL for an array which is not connected to the system. Symantec refers to this process as *ASL Tuning* and strongly recommends it on all pre 5.0 releases of DMP. The DMP 5.0 device discovery algorithms have been enhanced so that optimum boot time performance is achieved out of the box, without any *ASL Tuning*.

Array based multi-pathing solutions

Multi-pathing software from storage array vendors differ from volume manager based solutions mainly in two areas. First, they are independent from the volume managers which are providing a logical view of the storage devices to users and applications. This can be seen as an advantage. Second, since they are tightly integrated with the array vendor's storage systems, their support is often limited to exactly those arrays.

HP StorageWorks Secure Path

HP StorageWorks Secure Path for HP-UX is a purchasable software product and the current version (3.0FSP2 Patch2) provides highly available access to active/active and active/passive HP disk arrays

in HP-UX 11i v2 or earlier. Secure Path provides automatic failover across the paths of the individual storage device (LUN).

Currently supported active/passive HP arrays:

- EVA 3000 (HSV100)
- EVA 5000 (HSV110)

These arrays are only supported in Serviceguard clusters, if Secure Path is being used. For active/passive arrays, Secure Path facilitates static load balancing on active/passive arrays.

For active/active HP arrays Secure Path provides automatic load balancing across the paths of the storage device and thereby improves the I/O throughput.

Currently supported active/active HP arrays:

- EVA3000 (HSV101); EVA5000 (HSV111)
- EVA4x00; EVA6x00; EVA8x00
- XP256
- XP48; XP512
- XP128; XP1024
- XP10000; XP12000
- XP24000

After installing Secure Path, the normal device files in `/dev/dsk/` and `/dev/rdsk` become path independent. For example if a LUN is accessible through two paths (`/dev/dsk/6t0d0` and `/dev/dsk/8t0d0`) it does not matter which path is specified by the application for the I/O request, Secure Path will make sure that the I/O reaches the device, even if the request was issued to a failed path.

HP Storage Works Secure Path for HP-UX v 3.0F SP2 introduced virtual device special files (Vdsf) for active/active arrays which can be created for each LUN under its control. Those Vdfs's are created in the directory `/hpap/dsk/` for block access and in the directory `/hpap/rdsk/` for raw access. Figure 4 shows how to map Secure Path Vdsf to path representing OS device files. In this example, the Vdsf is: `/hpap/dsk/hpap1`.

Figure 4: Secure Path Vdsf to path mapping for a specific LUN

```
# autopath display -l 50_0-2747-0301
=================================================================
 HPswsp Version        : A.3.0F.02F.00F
=================================================================
 Array WWN             : 1005-5
=================================================================
 Lun WWN               : 50_0-2747-0301
 Virtual Device File   : /hpap/dsk/hpap1
 Load Balancing Policy : No Load Balancing
 Lun Timeout           : Infinite Retry (-1)
=================================================================
 Device Path                   Status
=================================================================
 /dev/dsk/c6t0d1              Active
 /dev/dsk/c8t0d1              Active
=================================================================
```

In addition to path failover and load balancing, Secure Path also monitors the path's health to allow automatic fail-back after a path is repaired and to proactively identify issues with idle paths. For more

information about HP StorageWorks Secure Path please see the "HP StorageWorks Secure Path 3.0F Service Pack 2 for HP-UX 11i v1.0 and 11i v2.0, installation and reference guide".

EMC PowerPath

EMC PowerPath is a purchasable software product that provides path management. Individual features are automatic path failover in the event of a hardware failure, dynamic load balancing between the paths, proactive I/O path testing and automatic path recovery upon restoration of a failed path. PowerPath offers various different failover and load balancing policies.

Besides Symmetrix and CLARiiON storage systems PowerPath also extended support to a range of HP StorageWorks arrays including the XP and EVA systems. While PowerPath is optional for most storage systems in a Serviceguard cluster, for CLARiiON storage systems the use of PowerPath is mandatory in a Serviceguard cluster if the nodes are on HP-UX 11i v2.

Further information about PowerPath can be found in the product specific documentation like "PowerPath for HP-UX Version 4.5 Installation and Administration Guide (P/N 300-002-383, Rev A03) " and "PowerPath Version 4.5 Product Guide (P/N 300-002-834 REV A03) "

Multi-pathing solutions for HP Virtual Machine

All multi-pathing solutions discussed above are supported with the current version – A.03.00 – of HP Virtual Machines (HPVM) with the multi-pathing software running on the HPVM host. Multi-pathing solutions must be implemented on the HPVM host and not the guest.

Selection and coexistence of different multi-pathing solutions

In general, customers can choose among the multi-pathing solution available on HP-UX 11i v2. The selection of a storage system and a volume manager used in a Serviceguard cluster reduces the number of available multi-pathing solution in a specific configuration. For example a cluster which uses LVM as a volume manager and an EMC Symmetrix as shared storage, can't deploy DMP or Secure Path as multi-pathing solution. Customers which implement one of the Serviceguard Storage Management Suite bundles most likely will choose DMP since it is included in all of those bundles. If multiple solutions are feasible, the following criteria are often used to select a multi-pathing solution:

- Cost: Does the multi-pathing software incur additional cost? PVlinks and DMP either come with the HP-UX or Serviceguard SMS. Secure Path and EMC PowerPath need to be purchased additionally.
- Multi-platform support: Often customers standardize on certain technologies. This might be the host operating system, the volume manager, or the storage array and they prefer to use the same multi-pathing solution throughout their computing environment. If this criteria has the highest priority for a customer, their standardization strategy predetermines the multi-pathing solution. Customers standardizing on VxVM/CVM volume manager might prefer DMP; others that standardize on EMC storage systems might prefer PowerPath.
- Dynamic load balancing: If dynamic load balancing is of great importance, a customer might rather invest in an array based multi-pathing solution or DMP than implementing LVM PVlinks.

Typically only one multi-pathing solution is installed on a system. In rare situations several multi-pathing solutions could be present on a system. However it does not add an extra level of high availability to have more than one multi-pathing solution providing services for the same device or LUN. It might even be counterproductive having multiple multi-pathing solutions installed. DMP, Secure Path, and PowerPath all try to provide multi-pathing to all devices they find during a discovery process, therefore the administrator has to ensure that only one product is active for specific LUN or sometimes a specific pair of controllers. Each of these solutions provides commands to reduce its scope on the different levels of granularity (LUN, controller, or array). Please see the product specific documentation referenced to in the "Multi-pathing product reference documentation" section above.

Multi-Pathing solutions on HP-UX 11i v3

The mass storage subsystem in HP-UX 11i v3 has been re-architected to increase configuration limits and provide an enriched feature set. It is far beyond the scope of this white paper to discuss all the features and benefits of it. Please see "Overview: Next Generation Mass Storage Stack" for an introduction of the HP-UX 11i v3 mass storage subsystem and see the list at: http://docs.hp.com/en/oshpux11iv3.html#Storage%20Area%20Management for additional information on various areas of the HP-UX 11i v3 mass storage subsystem.

A new representation of mass storage devices called the "agile view" has been introduced along with the new storage subsystem. The central idea of the agile view is that a DSF identifies a disk device or LUN and not a specific path to a device as in HP-UX 11i v2. Paths to a device can change dynamically, and multiple paths to a single device can be transparently treated as a single virtualized path, with I/O distributed across those multiple paths.

These new DSFs which represent LUNs instead of paths are also called "persistent" DSFs in contrast to the "legacy" DSFs since they do not change even if the path to a LUN changes. The legacy DSFs – as known from HP-UX 11i v2 – are still available on HP-UX 11i v3 for backward compatibility. Users are encouraged to move to the persistent DSFs wherever possible, because legacy DSFs are subject to HP-UX 11i v2 limits. The persistent DSFs are stored in a different directory structure:

- `/dev/disk` for block device files and
- `/dev/rdisk` for character device files.

Figure 5 shows how to map between legacy and persistent DSFs.

Figure 5: Map between persistent and legacy DSFs

```
# ioscan -m dsf /dev/dsk/c6t0d1
Persistent DSF          Legacy DSF(s)
========================================
/dev/disk/disk87        /dev/disk/c6t0d1

# ioscan -m dsf /dev/disk/disk87
Persistent DSF          Legacy DSF(s)
========================================
/dev/disk/disk87        /dev/dsk/c6t0d1
                        /dev/dsk/c8t0d1
```

Even though legacy DSFs are subject to HP-UX 11i v2 limits (e. g. number of LUNs and LUN paths), native multi-pathing is available using legacy DSFs on HP-UX 11i v3 and is enabled per default. The white paper: "HP-UX 11i v3 Native Multi-Pathing for Mass Storage" drills down on the multi-pathing related features in the new mass storage subsystem.

Plug-ins are needed to support active/passive arrays with native multi-pathing on HP-UX 11i v3. Asymmetrical active/active arrays (active/active arrays with a preferred controller) that are capable of Asymmetric Logical Unit Access (ALUA) don't need a plug-in to work with native multipathing. However prior to HP-UX 11i v3 0709, the "preferred path" policy might work best for asymmetrical active/active arrays like the HP StorageWorks EVA series. In a cluster, the administrator is responsible for consistent configuration of the preferred path on all nodes.

With the introduction of native multi-pathing in HP-UX 11i v3 which is available at no extra cost, volume manager and arrays based solutions become less dominant or disappear.

Native Multi-Pathing and LVM Alternate Links (PVlinks)

On HP-UX 11i v3, LVM relies on the native multi-pathing functionality of the OS, which is automatically enabled after an upgrade to HP-UX 11i v3 on existing volume groups with legacy DSFs. LVM supports both legacy and persistent DSFs on a system and even within a volume group. This allows the transition to persistent DSFs at the granularity of a pvol. Regardless of which type of device file is being used, native multi-pathing will be active by default. It is recommended to convert existing legacy DSFs to persistent DSFs in existing environments, to get access to all the new features of the HP-UX 11i v3 mass storage subsystem. Further details about native multi-pathing with LVM and LVM migration from legacy to persistent DSFs can be found in: LVM Migration from Legacy to Agile Naming Model.

Native Multi-Pathing and VxVM/CVM Dynamic Multi-pathing (DMP)

The details about DMP on HP-UX 11i v3 depend on the version of the volume manager and whether it is VxVM or CVM.

At the time HP-UX 11i v3 released, Veritas Volume Manager (VxVM) 4.1 was the only version of VxVM available. With this version, VxVM and Dynamic Multipathing (DMP) from Symantec are supported on HP-UX 11i v3, but do not provide multi-pathing and load balancing. Veritas DMP acts as a pass-through driver, allowing multi-pathing and load balancing to be controlled by the HP-UX mass storage subsystem instead.

VxVM 4.1 does not support the agile view. Only legacy DSFs are supported and subsequently, VxVM configurations are subject to HP-UX 11i v2 limits. Further details are described in: Veritas 4.1 Installation Guide (HP-UX 11i v3).

At a later time, VxVM and CVM 5.0 with DMP will become available on HP-UX 11i v3. At this point, the following is expected:

- VxVM/CVM 5.0 will be able to use the persistent DSFs.
- DMP on VxVM 5.0 volumes
  - will handle the multi-pathing by default (native multi-pathing is pass-through)
  - can be configured as pass-through (native multi-pathing is active)
- CVM 5.0 volumes require DMP to handle multi-pathing (native multi-pathing is pass-through)

HP StorageWorks Secure Path

HP StorageWorks Secure Path is not supported on HP-UX 11i v3 since the operating system provides native multi-pathing.

HP-UX 11i v3 supports the HP StorageWorks active/active disk arrays. The Secure Path virtual aliases (`/hpap/dsk/hpapX`) for active/active disk arrays provided for the Secure Path managed LUNs are not available after the OS upgrades to HP-UX 11i v3. Instead, the active/active disk arrays use the new persistent DSFs that are introduced with HP-UX 11i v3.

HP-UX 11i v3 does not support HP StorageWorks active/passive disk arrays. If you are using an active/passive EVA 3000 or 5000 disk arrays, you must upgrade the firmware to an active/active version.

Additional information on the migration from HP StorageWorks Secure Path to native multi-pathing in HP-UX 11i v3 can be found in the following white paper: Migrating from HP StorageWorks Secure Path for Active-Active Disk Arrays to Native Multipathing in HP-UX 11i v3.

EMC PowerPath

EMC PowerPath is not available on HP-UX 11i v3. Native multi-pathing on HP-UX 11i v3 supports EMC active/active disk arrays (Symmetrix and DMX).

Metrocluster SRDF was dependent on EMC PowerPath for configuring consistency groups. On HP-UX 11i v3, it will be possible to configure consistency groups through the EMC Enginuity Consistency Assist (ECA) product and thus removing the dependency of Metrocluster SRDF on EMC PowerPath.

Summary and Recommendation

On HP-UX 11i v2 several multi-pathing solutions exist and the section Selection and coexistence of different multi-pathing solutions provides some information on how to select between them.

For HP-UX 11i v3 the selection is straight forward. Since HP StorageWorks Secure Path and EMC PowerPath are not available on HP-UX 11i v3, customers that used those solutions on HP-UX 11i v2 will most probably move to native multi-pathing. The same holds true for customers that implemented PVlinks prior to HP-UX 11i v3 or those that want to deploy VxVM 4.1 on HP-UX 11i v3. For VxVM 5.0 on HP-UX 11i v3 the customers will have the choice between DMP and native multi-pathing while CVM 5.0 will require DMP according to the information that was available at the time this white paper was written.

For more information

To learn more about the new mass storage on HP-UX 11i v3 see the following documents on the HP documentation websites (http://docs.hp.com and http://docs.hp.com/en/oshpux11iv3.html#Storage%20Area%20Management):

- HP-UX 11i v2 to 11i v3 Mass Storage Stack Update Guide
- HP-UX 11i v3 Mass Storage Device Naming
- HP-UX 11i v3 Mass Storage I/O Performance Improvements
- HP-UX 11i v3 Native Multi-Pathing for Mass Storage
- HP-UX 11i v3 Persistent DSF Migration Guide
- LVM Migration from Legacy to Agile Naming Model
- Overview: Next Generation Mass Storage Stack
- Product Brief: The Next Generation Mass Storage Stack
- scsimgr SCSI Management and Diagnostics Utility on HP-UX 11i v3
- Migrating from HP StorageWorks Secure Path for Active-Active Disk Arrays to Native Multipathing in HP-UX 11i v3 (http://docs.hp.com/en/AA-RWF2A-TE/AA-RWF2A-TE.pdf)

4AA1-4403ENW, August 2007