sgi

SGI InfiniteStorage 4000 Series and 5000 Series
Failover Drivers Guide for SANtricity ES

(ISSM 10.86)

007-5886-002   April 2013

The information in this document supports the SGI InfiniteStorage 4000 series and 5000 series storage systems (ISSM 10.86). Refer to the table below to match your specific SGI InfiniteStorage product with the model numbers used in this document.

SGI Model #	NetApp Model
TP9600H	6091
TP9700F	6091
IS4500F	6091
TP9600F	3994 and 3992
IS4000H	3994
IS350	3992
IS220	1932 1333 DE1300
IS4100	4900
IS-DMODULE16-Z	FC4600
IS-DMODULE60	DE6900
IS4600	7091
IS-DMODULE12 & IS2212 (JBOD)	DE1600
IS-DMODULE24 & IS2224 (JBOD)	DE5600
IS-DMODULE60-SAS	DE6600
IS5012	E2600
IS5024	E2600
IS5060	E2600
IS5512	E5400
IS5524	E5400
IS5560	E5400
IS5600	E5500

Copyright information

Trademark information

NetApp, the NetApp logo, Network Appliance, the Network Appliance logo, Akorri, ApplianceWatch, ASUP, AutoSupport, BalancePoint, BalancePoint Predictor, Bycast, Campaign Express, ComplianceClock, Cryptainer, CryptoShred, Data ONTAP, DataFabric, DataFort, Decru, Decru DataFort, DenseStak, Engenio, Engenio logo, E-Stack, FAServer, FastStak, FilerView, FlexCache, FlexClone, FlexPod, FlexScale, FlexShare, FlexSuite, FlexVol, FPolicy, GetSuccessful, gFiler, Go further, faster, Imagine Virtually Anything, Lifetime Key Management, LockVault, Manage ONTAP, MetroCluster, MultiStore, NearStore, NetCache, NOW (NetApp on the Web), Onaro, OnCommand, ONTAPI, OpenKey, PerformanceStak, RAID-DP, ReplicatorX, SANscreen, SANshare, SANtricity, SecureAdmin, SecureShare, Select, Service Builder, Shadow Tape, Simplicity, Simulate ONTAP, SnapCopy, SnapDirector, SnapDrive, SnapFilter, SnapLock, SnapManager, SnapMigrator, SnapMirror, SnapMover, SnapProtect, SnapRestore, Snapshot, SnapSuite, SnapValidator, SnapVault, StorageGRID, StoreVault, the StoreVault logo, SyncMirror, Tech OnTap, The evolution of storage, Topio, vFiler, VFM, Virtual File Manager, VPolicy, WAFL, Web Filer, and XBB are trademarks or registered trademarks of NetApp, Inc. in the United States, other countries, or both.

IBM, the IBM logo, and ibm.com are trademarks or registered trademarks of International Business Machines Corporation in the United States, other countries, or both. A complete and current list of other IBM trademarks is available on the Web at *www.ibm.com/legal/copytrade.shtml.*

Apple is a registered trademark and QuickTime is a trademark of Apple, Inc. in the U.S.A. and/or other countries. Microsoft is a registered trademark and Windows Media is a trademark of Microsoft Corporation in the U.S.A. and/or other countries. RealAudio, RealNetworks, RealPlayer, RealSystem, RealText, and RealVideo are registered trademarks and RealMedia, RealProxy, and SureStream are trademarks of RealNetworks, Inc. in the U.S.A. and/or other countries.

All other brands or products are trademarks or registered trademarks of their respective holders and should be treated as such.

NetApp, Inc. is a licensee of the CompactFlash and CF Logo trademarks.

NetApp, Inc. NetCache is certified RealSystem compatible.

Table of Contents

Overview of Failover Drivers *1*

Failover drivers provide redundant path management for storage devices and cables in the data path from the host bus adapter to the controller. For example, you can connect two host bus adapters in the system to the redundant controller pair in a storage array, with different buses for each controller. If one host bus adapter, one bus cable, or one controller fails, the failover driver automatically reroutes input/output (I/O) to the good path, which permits the storage array to continue operating without interruption.

Failover drivers provide these functions:

- They automatically identify redundant I/O paths.
- They automatically reroute I/O to an alternate controller when a controller fails or all of the data paths to a controller fail.
- They check the state of known paths to the storage array.
- They provide status information on the controller and the bus.
- They check to see if the Service mode is enabled and if the modes have switched between Redundant Dual Active Controller (RDAC) and Auto-Volume Transfer (AVT).
- They check to see if Service mode is enabled on a controller and if the AVT or asymmetric logical unit access (ALUA) mode of operation has changed.

Failover Driver Setup Considerations

Most storage arrays contain two controllers that are set up as redundant controllers. If one controller fails, the other controller in the pair takes over the functions of the failed controller, and the storage array continues to process data. You can then replace the failed controller and resume normal operation. You do not need to shut down the storage array to perform this task.

The redundant controller feature is managed by the failover driver software, which controls data flow to the controller pairs independent of the operating system (OS). This software tracks the current status of the connections and can perform the switch-over without any changes in the OS.

Whether your storage arrays have the redundant controller feature depends on a number of items:

- Whether the hardware supports it. Refer to the hardware documentation for your storage arrays to determine whether the hardware supports redundant controllers.
- Whether your OS supports certain failover drivers. Refer to the installation and support guide for your OS to determine if your OS supports redundant controllers.
- How the storage arrays are connected. The storage array must have two controllers installed in a redundant configuration. Drive trays must be cabled correctly to support redundant paths to all drives.

The I/O Shipping feature implements support for ALUA. With the I/O Shipping feature, a storage array can service I/O requests through either controller in a duplex configuration. However, I/O shipping alone does not guarantee that I/O is routed to the optimized path. With Windows, Linux and VMWare, your storage array supports an extension to ALUA to address this problem so that volumes are accessed through the optimized path unless that path fails. With SANtricity ES 10.86 and subsequent releases, Windows and Linux DM-MP have I/O shipping enabled by default.

Supported Failover Drivers Matrix

Table 1 Matrix of Supported Failover Drivers by Operating System (OS)

	Windows Server 2012, Windows 2008, Vista and Hyper-V	**Red Hat Enterprise Linux (RHEL) 5.8 and 5.9 OSs and SUSE Linux Enterprise (SLES) 10.4 OS**	**RHEL 6.3 OS and SUSE Linux Enterprise 11.2**	**Solaris 10 U10, Solaris 10 u11**
Failover driver type	MPIO with NetApp DSM-ALUA	RDAC	DMMP (with E5500) or RDAC	MPxIO (non-TPGS)
Storage array mode	Mode Select, or ALUA	Either Mode Select or AVT	Mode Select.	Mode Select or Asymmetric Logical Unit Access (ALUA)
Number of volumes supported per partition	255	256	256	255
Cluster support?	Yes	Yes	Yes	Yes
Controller-drive trays supported	E2600 and E5400 E5500 (Windows 2012 only)	E2600 and E5400	E2600, E5400, and E5500	E2600 and E5400

Table 2 Matrix of Supported Failover Drivers by Operating System (OS)

	Solaris 11, Solaris 11.1	HPUX 11.31	VMWare 4.1 u3, 5.1, ad 5.0 u2	Mac OS 10.6 and 10.7
Failover driver type	MPxIO (TPGS/ALUA)	TPGS/ALUA	VMWare native - SATP/ALUA	ATTO driver with TPGS/ALUA
Storage array mode	Mode Select or ALUA			
Number of volumes per partition	255			
Cluster support?	No	No	Yes	NA
Controller-drive trays supported	E2600, E5400, and E5500	E2600 and E5400	E2600 and E5400	E2600 and E5400

I/O Coexistence

A host can have connections to more than one storage array. The connected storage arrays can have different controller firmware levels, provided that they share common components (HBAs, switches, and operating system versions) and common failover drivers.

Failover Configuration Diagrams

You can configure failover in several ways. Each configuration has its own advantages and disadvantages. This section describes these configurations:

- Single-host configuration
- Multi-host configuration

This section also describes how the storage management software supports redundant controllers.

NOTE For best results, use the multi-host configuration. It provides the fullest failover protection and functionality in the event that a problem exists with the connection.

Single-Host Configuration

In a single-host configuration, the host system contains two host bus adapters (HBAs), with each HBA connected to one of the controllers in the storage array. The storage management software is installed on the host. The two connections are required for maximum failover support for redundant controllers.

Although you can have a single controller in a storage array or a host that has only one HBA port, you do not have complete failover data path protection with either of those configurations. The cable and the HBA become a single point of failure, and

any data path failure could result in unpredictable effects on the host system. For the greatest level of I/O protection, provide each controller in a storage array with its own connection to a separate HBA in the host system.

Figure 1 Single-Host-to-Storage Array Configuration
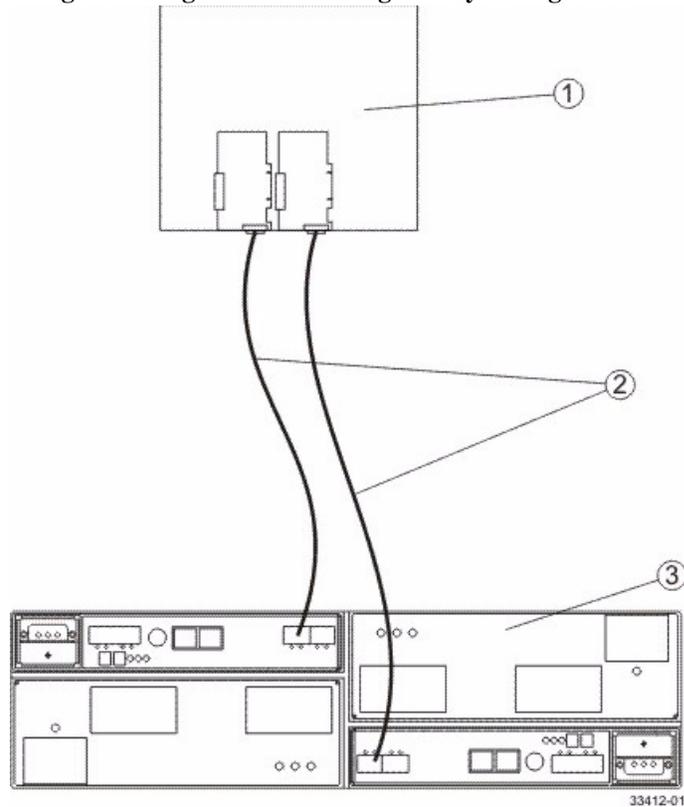
1. Host System with Two Fibre Channel Host Bus Adapters
2. Fibre Channel Connection – Fibre Channel Connection Might Contain One or More Switches
3. Storage Array with Two Fibre Channel Controllers

Host Clustering Configurations

In a clustering configuration, two host systems are each connected by two connections to both of the controllers in a storage array. SANtricity ES Storage Manager, including failover driver support, is installed on each host.

Not every operating system supports this configuration. Consult the restrictions in the installation and support guide specific to your operating system for more information. Also, the host systems must be able to handle the multi-host configuration. Refer to the applicable hardware documentation.

Both hosts have complete visibility of both controllers, all data connections, and all configured volumes in a storage array, plus failover support for the redundant controllers. However, in this configuration, you must use caution when you perform storage management tasks (especially deleting and creating volumes) to make sure that the two hosts do not send conflicting commands to the controllers in the storage arrays.

The following items apply to these clustering configurations:

- Both hosts must have the same operating system version and SANtricity ES Storage Manager version installed.

- The failover driver configuration might require tuning.

- Both host systems must have the same volumes-per-host bus adapter capacity. This capacity is important for failover situations so that each controller can take over for the other and show all of the configured volume groups and volumes.

- If the operating system on the host system can create reservations, the storage management software honors them. Each host might have reservations to specified volume groups and volumes, and only the software on that host can perform operations on the reserved volume group and volume. Without reservations, the software on either host system is able to start any operation. Therefore, you must use caution when you perform certain tasks that need exclusive access. Especially when you create and delete volumes, make sure that you have only one configuration session open at a time (from only one host), or the operations that you perform could fail.

Figure 2 Multi-Host-to-Storage Array Configuration

33415-01
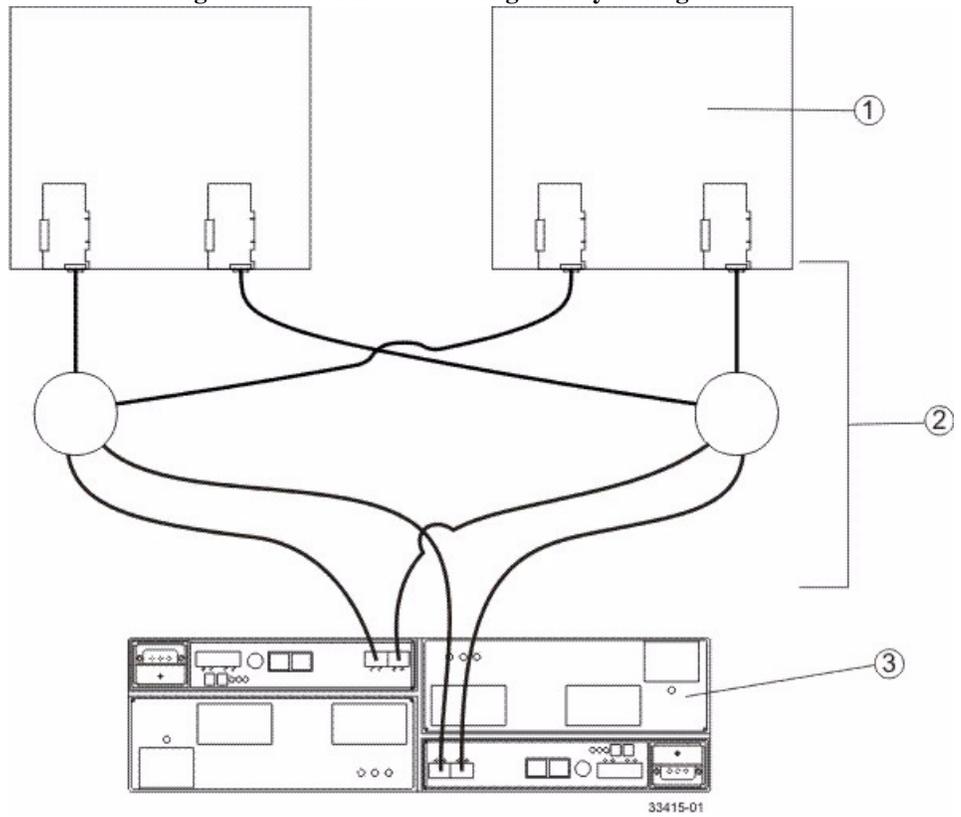
1. Two Host Systems, Each with Two Fibre Channel Host Bus Adapters

2. Fibre Channel Connections with Two Switches (Might Contain Different Switch Configurations)

3. Storage Array with Two Fibre Channel Controllers

Supporting Redundant Controllers

The following figure shows how failover drivers provide redundancy when the host application generates a request for I/O to controller A, but controller A fails. Use the numbered information to trace the I/O data path.

Figure 3 Example of Failover I/O Data Path Redundancy
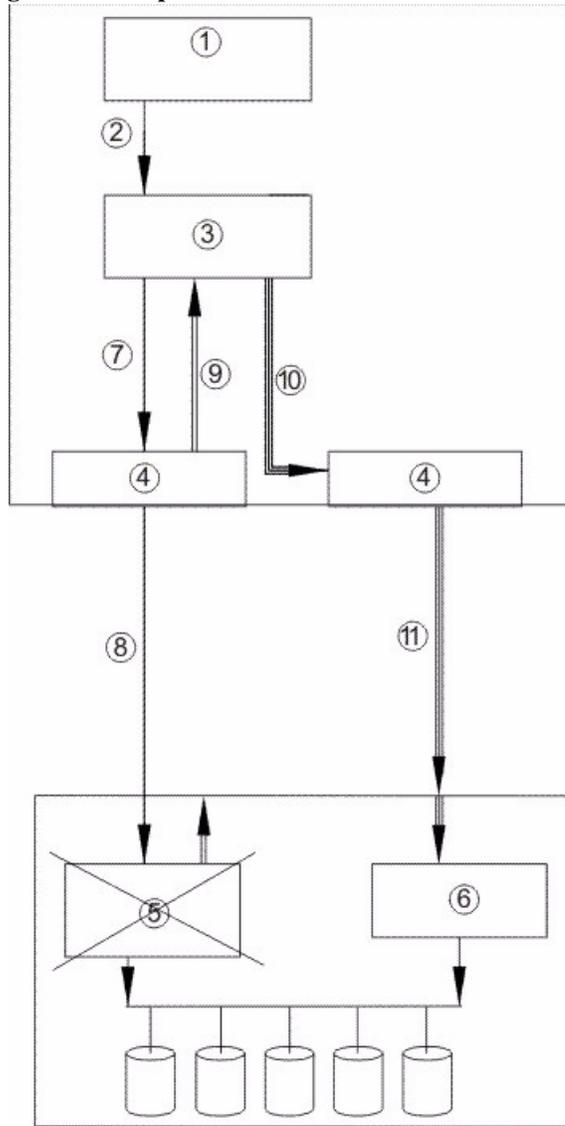
1. Host Application
2. I/O Request
3. Failover Driver
4. Host Bus Adapters
5. Controller A Failure
6. Controller B
7. Initial Request to the HBA

8. Initial Request to the Controller Failed

9. Request Returns to the Failover Driver

10. Failover Occurs and I/O Transfers to Another Controller

11. I/O Request Re-sent to Controller B

How a Failover Driver Responds to a Data Path Failure

One of the primary functions of the failover feature is to provide path management. Failover drivers monitor the data path for devices that are not working correctly or for multiple link errors. If a failover driver detects either of these conditions, the failover driver automatically performs these steps:

- The failover driver checks the pair table for the redundant controller.

- The failover driver forces volumes to the other controller and routes all I/O to the remaining active controller.

- The failover driver performs a path failure if alternate paths to the same controller are available. If all of the paths to a controller fail, RDAC performs a controller failure. The failover driver provides notification of an error through the OS error log facility.

When a drive and a controller fail simultaneously, the event is considered to be a double failure. The storage management software provides data integrity as long as all drive failures and controller failures are detected and fixed before more failures occur.

User Responses to a Data Path Failure

Use the Major Event Log (MEL) to respond to a data path failure. The information in the MEL provides the answers to these questions:

- What is the source of the error?

- What is required to fix the error, such as replacement parts or diagnostics?

Under most circumstances, contact your Technical Support Representative any time a path fails and the storage array notifies you of the failure. Use the Recovery Guru in the storage management software to diagnose and fix the problem, if possible. If your controller has failed and your storage array has customer-replaceable controllers, replace the failed controller. Follow the manufacturer's instructions for how to replace a failed controller.

Load-Balancing Policies

Load balancing is the redistribution of read/write requests to maximize throughput between the server and the storage array. Load balancing is very important in high workload settings or other settings where consistent service levels are critical. The multi-path driver transparently balances I/O workload without administrator intervention. Without multi-path software, a server sending I/O requests down several paths might operate with very heavy workloads on some paths, while other paths are not used efficiently.

The multi-path driver determines which paths to a device are in an active state and can be used for load balancing. The load-balancing policy uses one of three algorithms: round robin, least queue depth, or least path weight. Multiple options for setting the load-balancing policies let you optimize I/O performance when mixed host interfaces are configured. The load-balancing policies that you can choose depend on your operating system. Load balancing is performed on multiple paths to the same controller but not across both controllers.

Table 3 Load-Balancing Policies That Are Supported by the Operating Systems

Operating System	Multi-Path Driver	Load-Balancing Policy
Windows	MPIO DSM	Round robin with subset, least queue depth, weighted paths
Red Hat Enterprise Linux (RHEL)	RDAC	Round robin with subset, least queue depth
	DMMP	Round robin
SUSE Linux Enterprise (SLES)	RDAC	Round robin with subset, least queue depth
	DMMP	Round robin
Solaris	MPxIO	Round robin with subset

Least Queue Depth

The least queue depth policy is also known as the least I/Os policy or the least requests policy. This policy routes the next I/O request to the data path on the controller that owns the volume that has the least outstanding I/O requests queued. For this policy, an I/O request is simply a command in the queue. The type of command or the number of blocks that are associated with the command is not considered. The least queue depth policy treats large block requests and small block requests equally. The data path selected is one of the paths in the path group of the controller that owns the volume.

Round Robin with Subset I/O

The round robin with subset I/O load-balancing policy routes I/O requests, in rotation, to each available data path to the controller that owns the volumes. This policy treats all paths to the controller that owns the volume equally for I/O activity. Paths to the secondary controller are ignored until ownership changes. The basic assumption for the round robin with subset I/O policy is that the data paths are equal. With mixed host support, the data paths might have different bandwidths or different data transfer speeds.

Weighted Paths

The weighted paths policy assigns a weight factor to each data path to a volume. An I/O request is routed to the path with the lowest weight value to the controller that owns the volume. If more than one data path to the volume has the same weight value, the round-robin with subset path selection policy is used to route I/O requests between the paths with the same weight value.

Dividing I/O Activity Between Two RAID Controllers to Obtain the Best Performance

For the best performance of a redundant controller system, use the storage management software to divide I/O activity between the two RAID controllers in the storage array. You can use either the graphical user interface (GUI) or the command line interface (CLI).

To use the GUI to divide I/O activity between two RAID controllers, perform one of these steps:

- **Specify the owner of the preferred controller of an existing volume** – Select **Volume >> Change >> Ownership/Preferred Path** in the Array Management Window.

> **NOTE** You also can use this method to change the preferred path and ownership of all volumes in a volume group at the same time.

- **Specify the owner of the preferred controller of a volume when you are creating the volume** – Select **Volume >> Create** in the Array Management Window.

To use the CLI, go to the "Create RAID Volume (Free Extent Based Select)" online help topic for the command syntax and description.

Changing the Preferred Path Online Without Stopping the Applications

You can change the preferred path setting for a volume or a set of volumes online without stopping the applications. However, the host multipath driver might not immediately recognize that the preferred path has changed.

During the interval when the old path is still used, the driver might trigger a `volume not on preferred path` Needs Attention condition in the storage management software. This condition is removed as soon as the state change monitor is run. A MEL event and an associated alert notification are delivered for the condition. You can configure the AVT alert delay period with the storage management software. The Needs Attention reporting is postponed until the failover driver failback task has had a chance to run.

If the ALUA mode is enabled, the failover driver will continue to use the same data path or paths. The storage array will, however, use its I/O Shipping feature to service I/Os requests even if they are directed to the non-owning controller.

> **NOTE** The newer versions of the RDAC failover driver and the DSM failover driver do not recognize any AVT or ALUA status change (enabled or disabled) until the next cycle of the state change monitor.

Failover Drivers for the Windows Operating System   *2*

The failover driver for hosts with Microsoft Windows operating systems is Microsoft Multipath I/O (MPIO) with a Device Specific Module (DSM) for SANtricity ES Storage Manager.

Microsoft Multipath Input/Output

Microsoft Multipath I/O (MPIO) provides an infrastructure to build highly available solutions for Windows operating systems (OSs). MPIO uses Device Specific Modules (DSMs) to provide I/O routing decisions, error analysis, and failover.

NOTE You can use MPIO for all controllers that run controller firmware version 6.19 or later. MPIO is not supported for any earlier versions of the controller firmware, and MPIO cannot coexist on a server with RDAC. If you have legacy systems that run controller firmware versions earlier than 6.19, you must use RDAC for your failover driver. For SANtricity ES Storage Manager Version 10.10 and later and all versions of SANtricity Storage Manager, the Windows OS supports only MPIO.

Device Specific Module for the Microsoft MPIO Solution

The DSM driver is the hardware-specific part of Microsoft's MPIO solution. This release supports Microsoft's Windows Server 2008 OS, and Windows Server 2008 R2 OS and Windows Server 2012. The Hyper-V role is also supported when running the DSM within the parent partition. The SANtricity ES Storage Manager Version 10.86 works with the DSM to implement supported features.

Windows OS Restrictions

The MPIO DSM failover driver comes in these versions:

- 32-bit (x86)
- 64-bit AMD/EM64T (x64)

These versions are not compatible with each other. Because multiple copies of the driver cannot run on the same system, each subsequent release is backward compatible.

Per-Protocol I/O Timeout Values

The timeout value associated with a non-passthrough I/O requests, such as read/write requests, is based on the Microsoft disk driver's `TimeOutValue` parameter, as defined in the Registry. A feature within the DSM allows a customized timeout value to be applied based on the protocol (such as Fibre Channel, SAS, or iSCSI) that a path uses. Per-protocol timeout values provide these benefits:

- Without per-protocol timeout values, the `TimeOutValue` setting is global and affects all storage.

- The `TimeOutValue` is typically reset when an HBA driver is upgraded.
- For information about the configurable parameters for the customized timeout feature, go to Configuration Settings for the Windows DSM and the Linux RDAC.

The per-protocol timeout values feature slightly modifies the way in which the `SynchTimeout` parameter is evaluated. The `SynchTimeout` parameter determines the I/O timeout for synchronous requests generated by the DSM driver. Examples include the native handling of persistent reservation requests and inquiry commands used during device discovery. It is important that the timeout value for the requests from the DSM driver be at least as large as the per-protocol I/O timeout value. When a host boots, the DSM driver performs these actions:

- If the value of the `SynchTimeout` parameter is defined in the Registry key of the DSM driver, record the current value.
- If the value of the `TimeOutValue` parameter of the Microsoft disk driver is defined in the Registry, record the current value.
- Use the higher of the two values as the initial value of the `SynchTimeout` parameter.
- If neither value is defined, use a default value of 10 seconds.
- For each synchronous I/O request, the higher value of either the per-protocol I/O timeout or the `SynchTimeout` parameter is used. For example:
  — If the value of the `SynchTimeout` parameter is 120 seconds, and the value of the `TimeOutValue` parameter is 60 seconds, 120 seconds is used for the initial value.
  — If the value of the `SynchTimeout` parameter is 120 seconds, and the value of the `TimeOutValue` parameter is 180 seconds, 180 seconds is used for the initial value of the synchronous I/O requests for the DSM driver.
  — If the I/O timeout value for a different protocol (for example, SAS) is 60 seconds and the initial value is 120 seconds, the I/O is sent using a 120-second timeout.

Selective LUN Transfer

This feature limits the conditions under which the DSM driver moves a LUN to the alternative controller to three cases:

1. When a DSM driver with a path to only one controller, the non-preferred path, discovers a path to the alternate controller.
2. When an I/O request is directed to a LUN that is owned by the preferred path, but the DSM driver is attached to only the non-preferred path.
3. When an I/O request is directed to a LUN that is owned by the non-preferred path, but the DSM driver is attached to only the preferred path.

Case 2 and case 3 have these user-configurable parameters that can be set to tune the behavior of this feature.

- The maximum number of times that the LUN transfer is issued. This parameter setting prevents a continual ownership thrashing condition from occurring in cases where the controller tray or the controller-drive tray is attached to another host that requires the LUN be owned by the current controller.

- A time delay before LUN transfers are attempted. This parameter is used to de-bounce intermittent I/O path link errors. During the time delay, I/O requests will be retried on the current controller to take advantage of the possibility that another host might transition the LUN to the current controller.

For further information about these two parameters, go to Configuration Settings for the Windows DSM and the Linux RDAC.

In the case where the host system is connected to both controllers and an I/O is returned with a 94/01 status, in which the LUN is not owned and can be owned, the DSM driver modifies its internal data on which controller to use for that LUN and reissue the command to the other controller. To avoid interfering with other hosts that might be attached to a controller tray or the controller-drive tray, the DSM driver does not issue a LUN transfer command to that controller tray or controller-drive tray.

When the DSM detects that a volume transfer operation is required, the DSM does not immediately issue the failover/failback command. The DSM will, with the default settings, delay for three seconds before sending the command to the storage array. This delay provides time to batch together as many volume transfer operations for other LUNs as possible. The controller single-threads volume transfer operations and therefore rejects additional transfer commands until the controller has completed the operation that it is currently working on. This action results in a period during which I/Os are not successfully serviced by the storage array. By introducing a delay during which volume transfer operations can be aggregated into a batch operation, the DSM reduces the likelihood that a volume transfer operation will exceed the retry limit. In large system configurations, you might need to increase the default three-second delay value because, with more hosts in the configuration, more volume transfer commands might be sent.

The delay value is located at:
`HKEY_LOCAL_MACHINE\SYSTEM\CurrentControlSet\Services\<DSM _Driver>\Parameters\ LunFailoverDelay`.

This feature is enabled when these conditions exist:

- The controller tray or the controller-drive tray does not have AVT enabled.
- The DSM driver configurable parameter `ClassicModeFailover` is set to `1`.
- The DSM driver configurable parameter `DisableLunRebalance` is set to `4`.

Windows Failover Cluster

Clustering for Windows Server 2008 Operating Systems and subsequent releases of Windows uses SCSI-3 persistent reservations natively. As a result, you can use one of the previously mentioned load-balancing policies across all controller paths. If you are operating in a clustered or shared LUN environment and are not using the I/O

Shipping feature of the CFW or Selective LUN Transfer feature of the DSM, set the `DisableLunRebalance` parameter to 3. For information about this parameter, go to Configuration Settings for the Windows DSM and the Linux RDAC.

I/O Shipping Feature for Asymmetric Logical Unit Access (ALUA)

The I/O Shipping feature implements support for ALUA. With earlier releases of the controller firmware (CFW), the device specific module (DSM) had to send input/output (I/O) requests for a particular volume to the controller that owned that volume. A controller would reject requests it received for a volume that it did not own. This behavior was necessary in order for the storage array to maintain data consistency within the volume. This same behavior, however, was responsible for several areas of contention during system boot and during multi-host\path failure conditions.

With the I/O Shipping feature, a storage array can service I/O requests through either controller in a duplex configuration. There is a performance penalty when the non-owning controller accesses a volume. To maintain the best I/O subsystem performance, the DSM interacts with the CFW to insure that I/O requests are sent to the owning controller if that controller is available.

Verifying that the I/O Shipping Feature for ALUA Support is Installed on the Windows OS

When you install or update host software to SANtricity ES version 10.83 or more recent and install or update controller firmware to CFW version 7.83 or more recent on the Windows OS, support for ALUA is enabled by default. Perform the following steps to verify that ALUA support is installed.

1. From the Start menu, select Command Prompt.
2. At the command prompt, type `dsmUtil -g <target_id>`.

   The appearance of `ALUAEnabled` properties in the output indicates that the failover driver is upgraded.
3. At the command prompt, type `SMdevices`.

   The appearance of `active optimized` or `active non-optimized` (rather than `passive` or `unowned`) in the output indicates that the host can see the LUNs that are mapped to it.

Disabling the I/O Shipping Feature for ALUA Support on the Windows OS

When you install or update the DSM, by default, the Selective LUN Transfer (SLT) feature is enabled to support I/O Shipping. Some registry values are modified during the DSM update if the previous version did not enable SLT. To prevent enabling SLT so that your storage array operates without the I/O Shipping feature, edit the registry to have the following settings.

1. From the Start menu, select Command Prompt.
2. At the command prompt, type the following command: `Regedit.exe`

3. In the Registry Editor, expand the path to `HKEY_LOCAL_MACHINE\SYSTEM\CurrentControlSet\Services\<DSM_Driver>\Parameters`. The expression <DSM_Driver> is the name of the DSM driver used in your storage array.

4. Set the parameter `DisableLunRebalance`.

5. Set the parameter `ClassicModeFailover`.

6. Close the Registry Editor.

Reduced Failover Timing

Settings related to drive I/O timeout and HBA connection loss timeout are adjusted in the host operating system so that failover does not occur when a controller is restarted. These settings provide protection from exception conditions that might occur when both controllers in a controller tray or a controller-drive tray are restarted at the same time, but they have the unfortunate side-effect of causing longer failover times than might be tolerated by some application or clustered environments. Support for the reduced failover timing feature includes support for reduced timeout settings, which result in faster failover response times.

The following restrictions apply to this feature:

- Only the Windows Server 2008 OS and subsequent releases of the Windows OS support this feature.

- Non-enterprise products attached to a host must use controller firmware release 7.35 or higher. Enterprise products attached to a host must use controller firmware release 7.6 or higher. For configurations where a mix of earlier releases is installed, older versions are not supported.

- When this feature is used with Windows Server Failover Cluster (WSFC) on the Windows Server 2008 OS, MPIO HotFix 970525 is required. The required HotFix is a standard feature for the Windows Server 2008 R2 OS.

Additional restrictions apply to storage array brownout conditions. Depending on how long the brownout condition lasts, PR registration information for volumes might be lost. By design, WSFC periodically polls the cluster storage to determine the overall health and availability of the resources. One action performed during this polling is a PRIN READ_KEYS request, which returns registration information. Because a brownout condition can cause blank information to be returned, WSFC interprets this as a loss of access to the drive resource and attempts recovery by first failing that drive resource, and then performing a new arbitration.

The arbitration recovery process happens almost immediately after the resource is failed. This situation, along with the PnP timing issue, might result in a failed recovery attempt. Fortunately, you can modify the timing of the recovery process by using the cluster.exe command-line tool. Microsoft recommends changing the following, where *resource_name* is a cluster disk resource, such as Cluster Disk 1:

```
cluster.exe resource "resource_name" /prop
RestartDelay=4000

cluster.exe resource "resource_name" /prop
RestartThreshold=5
```

The previous example changes the disk-online delay to four seconds and the number of online restarts to five. The changes must be repeated for each drive resource. The changes persist across reboots. To display either the current settings or the changed settings, run the following command:

```
cluster.exe resource "resource_name" /prop
```

Another option exists to prevent the storage array from returning blank registration information. This option takes advantage of the Active Persist Through Power Loss (APTPL) feature found in Persistent Reservations, which ensures that the registration information persists through brownout or other conditions related to a power failure. APTPL is enabled when a registration is initially made to the drive resource. WSFC does not use the APTPL feature, but an option is provided in the DSM to set this feature when a registration request is made.

NOTE The APTPL feature within the DSM driver is enabled using the DSM utility with the `- o` (feature) option by setting the `SetAPTPLForPR` option to `1`. According to the SCSI specification, you must set this option before PR registration occurs. If you set this option after a PR registration occurs, take the disk resource offline, and then bring the disk resource back online. If the DSM driver has set the APTPL option during registration, an internal flag is set, and the DSM utility output from the `-g` option indicates this condition. The SCSI specification does not provide a means for the initiator to query the storage array to determine the current APTPL setting. As a result, the `-g` output from one node might show the option set, but another node might not. Interpret output from the `-g` option with caution. By default, the DSM driver is released without this option enabled.

Path Congestion Detection and Online/Offline Path States

The path congestion detection feature allows the DSM driver to place a path `offline` based on the path I/O latency. The DSM automatically sets a path `offline` when I/O response times exceed user-definable congestion criteria. An administrator can manually place a path into the `Admin Offline` state. When a path is either set `offline` by the DSM driver or by an administrator, I/O is routed to a different path. The `offline` or `admin offline` path is not used for I/O until the system administrator sets the path `online`.

For more information on path congestion configurable parameters, go to Configuration Settings for the Windows DSM and the Linux RDAC.

Configuration Settings for the Windows DSM and the Linux RDAC

This topic applies to both the Windows OS and the Linux OS. The failover driver that is provided with the storage management software contains configuration settings that can modify the behavior of the driver.

- For the Linux OS, the configuration settings are in the `/etc/mpp.conf` file.
- For the Windows OS, the configuration settings are in the
  `HKEY_LOCAL_MACHINE\System\CurrentControlSet\`
  `Services\<DSM_Driver>\Parameters` registry key, where
  `<DSM_Driver>` is the name of the OEM-specific driver.

The default Windows failover driver is `mppdsm.sys`. Any changes to the settings take effect on the next reboot of the host.

The default values listed in the following table apply to both the Windows OS and the Linux OS unless the OS is specified in parentheses. Many of these values are overridden by the failover installer for the Linux OS or the Windows OS.

ATTENTION **Possible loss of data access** – If you change these settings from their configured values, you might lose access to the storage array.

Table 1 Configuration Settings for the Windows DSM and the Linux RDAC

Parameter Name	Default Value (Operating System)	Description
MaxPathsPerController	4	The maximum number of paths (logical endpoints) that are supported per controller. The total number of paths to the storage array is the MaxPathsPerController value multiplied by the number of controllers. The allowed values range from 0x1 (1) to 0x20 (32) for the Windows OS, and from 0x1 (1) to 0xFF (255) for Linux RDAC. For use by Technical Support Representatives only.
ScanInterval	1 (Windows) 60 (Linux)	The interval time, in seconds, that the failover driver checks for these conditions: ■ A change in preferred ownership for a LUN ■ An attempt to rebalance LUNs to their preferred paths ■ A change in AVT enabled status or AVT disabled status ■ A change in ALUA enabled status or ALUA disabled status For the Windows OS, the allowed values range from 0x1 to 0xFFFFFFFF and must be specified in minutes. For the Linux OSs, the allowed values range from 0x1 to 0xFFFFFFFF and must be specified in seconds. For use by Technical Support Representatives only.

Parameter Name	Default Value (Operating System)	Description
ErrorLevel	3	This setting determines which errors to log. These values are valid: ■ 0 – Display all errors ■ 1 – Display path failover errors, controller failover errors, re-triable errors, fatal errors, and recovered errors ■ 2 – Display path failover errors, controller failover errors, re-triable errors, and fatal errors ■ 3 – Display path failover errors, controller failover errors, and fatal errors ■ 4 – Display controller failover errors, and fatal errors For use by Technical Support Representatives only.
SelectionTimeoutRetryCount	0	The number of times a selection timeout is retried for an I/O request before the path fails. If another path to the same controller exists, the I/O is retried. If no other path to the same controller exists, a failover takes place. If no valid paths exist to the alternate controller, the I/O is failed. The allowed values range from 0x0 to 0xFFFFFFFF. For use by Technical Support Representatives only.
CommandTimeoutRetryCount	1	The number of times a command timeout is retried for an I/O request before the path fails. If another path to the same controller exists, the I/O is retried. If another path to the same controller does not exist, a failover takes place. If no valid paths exist to the alternate controller, the I/O is failed. The allowed values range from 0x0 to 0xa (10) for the Windows OS, and from 0x0 to 0xFFFFFFFF For Linux RDAC. For use by Technical Support Representatives only.

Parameter Name	Default Value (Operating System)	Description
UaRetryCount	10	The number of times a Unit Attention (UA) status from a LUN is retried. This parameter does not apply to UA conditions due to Quiescence In Progress. The allowed values range from `0x0` to `0x64` (100) for the Windows OS, and from `0x0` to `0xFFFFFFFF` For Linux RDAC. For use by Technical Support Representatives only.
SynchTimeout	120	The timeout, in seconds, for synchronous I/O requests that are generated internally by the failover driver. Examples of internal requests include those related to rebalancing, path validation, and issuing of failover commands. The allowed values range from `0x1` to `0xFFFFFFFF`. For use by Technical Support Representatives only.
DisableLunRebalance	0	This parameter provides control over the LUN failback behavior of rebalancing LUNs to their preferred paths. These values are possible: ■ `0` – LUN rebalance is enabled for both AVT and non-AVT modes. ■ `1` – LUN rebalance is disabled for AVT mode and enabled for non-AVT mode. ■ `2` – LUN rebalance is enabled for AVT mode and disabled for non-AVT mode. ■ `3` – LUN rebalance is disabled for both AVT and non-AVT modes. ■ `4` – The selective LUN Transfer feature is enabled if AVT mode is off and `ClassicModeFailover` is set to LUN level 1.
S2ToS3Key	Unique key	This value is the SCSI-3 reservation key generated during failover driver installation. For use by Technical Support Representatives only.

Parameter Name	Default Value (Operating System)	Description
LoadBalancePolicy	1	This parameter determines the load-balancing policy used by all volumes managed by the Windows DSM and Linux RDAC failover drivers. These values are valid: ■ 0 – Round robin with subset. ■ 1 – Least queue depth with subset. ■ 2 – Least path weight with subset (Windows OS only)
ClassicModeFailover	0	This parameter provides control over how the DSM handles failover situations. These values are valid: ■ 0 – Perform controller-level failover (all LUNs are moved to the alternate controller). ■ 1 – Perform LUN-level failover (only the LUNs indicating errors are transferred to the alternate controller).
SelectiveTransfer MaxTransferAttempts	3	This parameter sets the maximum number of times that a host transfers the ownership of a LUN to the alternate controller when the Selective LUN Transfer mode is enabled. This setting prevents multiple hosts from continually transferring LUNs between controllers.
SelectiveTransferMinIO WaitTime	3	This parameter sets the minimum wait time (in seconds) that the DSM waits before it transfers a LUN to the alternate controller when the Selective LUN Transfer mode is enabled. This parameter is used to limit excessive LUN transfers due to intermittent link errors.

Wait Time Settings

When the failover driver receives an I/O request for the first time, the failover driver logs timestamp information for the request. If a request returns an error and the failover driver decides to retry the request, the current time is compared with the original timestamp information. Depending on the error and the amount of time that has elapsed, the request is retried to the current owning controller for the LUN, or a failover is performed and the request sent to the alternate controller. This process is

known as a wait time. If the `NotReadyWaitTime` value, the `BusyWaitTime` value, and the `QuiescenceWaitTime` value are greater than the `ControllerIoWaitTime` value, they will have no effect.

For the Linux OS, the configuration settings can be found in the `/etc/mpp.conf` file. For the Windows OS, the configuration settings can be found in the Registry under:

```
HKEY_LOCAL_MACHINE\System\CurrentControlSet\
Services\<DSM_Driver>
```

In the preceding setting, `<DSM_Driver>` is the name of the OEM-specific driver. The default Windows driver is named `mppdsm.sys`. Any changes to the settings take effect the next time the host is restarted.

ATTENTION Possible loss of data access – If you change these settings from their configured values, you might lose access to the storage array.

Table 2 Configuration Settings for the Path Congestion Detection Feature

Parameter Name	Default Values	Description
NotReadyWaitTime	Linux: 270 Windows: NA	The time, in seconds, a Not Ready condition (SK 0x06, ASC/ASCQ 0x04/0x01) is allowed before failover is performed. Valid values range from `0x1` to `0xFFFFFFFF`.
BusyWaitTime	Linux: 270 Windows: NA	The time, in seconds, a Busy condition is allowed before a failover is performed. Valid values range from `0x1` to `0xFFFFFFFF`.
QuiescenceWaitTime	Linux: 270 Windows: NA	The time, in seconds, a Busy condition is allowed before a failover is performed. Valid values range from `0x1` to `0xFFFFFFFF`.
ControllerIoWait Time	Linux: 120 Windows: NA	Provides an upper-bound limit, in seconds, that an I/O is retried on a controller regardless of the retry status before a failover is performed. If the limit is exceeded on the alternate controller, the I/O is again attempted on the original controller. This process continues until the value of the `ArrayIoWaitTime` limit is reached. Valid values range from `0x1` to `0xFFFFFFFF`.
ArrayIoWaitTime	Linux RDAC: 600 Windows DSM: 600	Provides an upper-bound limit, in seconds, that an I/O is retried to the storage array regardless of to which controller the request is attempted. After this limit is exceeded, the I/O is returned with a failure status. Valid values range from `0x1` to `0xFFFFFFFF`.

Configuration Settings for Path Congestion Detection and Online/Offline Path States

The following configuration settings are applied using the utility mppUtil -o option parameter.

Table 3 Configuration Settings for the Path Congestion Detection Feature

Parameter Name	Default Value	Description
CongestionDetectionEnabled	0x0	A Boolean value that indicates whether the path congestion detection is enabled. If this parameter is not defined or is set to 0x0, the value is false, the path congestion feature is disabled, and all of the other parameters are ignored. If set to 0x1, the path congestion feature is enabled. Valid values are 0x0 or 0x1.
CongestionResponseTime	0x0	If CongestionIoCount is 0x0 or not defined, this parameter represents an average response time in seconds allowed for an I/O request. If the value of the CongestionIoCount parameter is nonzero, this parameter is the absolute time allowed for an I/O request. Valid values range from 0x1 to 0x10000 (approximately 18 hours).
CongestionIoCount	0x0	The number of I/O requests that have exceeded the value of the CongestionResponseTime parameter within the value of the CongestionTimeFrame parameter. Valid values range from 0x0 to 0x10000 (approximately 4000 requests).
CongestionTimeFrame	0x0	A sliding window that defines the time period that is evaluated in seconds. If this parameter is not defined or is set to 0x0, the path congestion feature is disabled because no time frame has been defined. Valid values range from 0x1 to 0x1C20 (approximately two hours).

Parameter Name	Default Value	Description
CongestionSamplingInterval	0x0	The number of I/O requests that must be sent to a path before the nth request is used in the average response time calculation. For example, if this parameter is set to 100, every 100th request sent to a path is used in the average response time calculation. If this parameter is set to 0x0 or not defined, the path congestion feature is disabled for performance reasons—every I/O request would incur a calculation. Valid values range from 0x1 to 0xFFFFFFFF (approximately 4 billion requests).
CongestionMinPopulationSize	0x0	The number of sampled I/O requests that must be collected before the average response time is calculated. Valid values range from 0x1 to 0xFFFFFFFF (approximately 4 billion requests).
CongestionTakeLastPathOffline	0x0	A Boolean value that indicates whether the DSM driver takes the last path available to the storage array offline if the congestion thresholds have been exceeded. If this parameter is not defined or is set to 0x0, the value is false. Valid values are 0x0 or 0x1.
		NOTE Setting a path offline with the dsmUtil utility succeeds regardless of the setting of this value.

Example Configuration Settings for the Path Congestion Detection Feature

NOTE Before path congestion detection can be enabled, you must set the CongestionResponseTime, CongestionTimeFrame, and CongestionSamplingInterval parameters to valid values.

To set the path congestion IO response time to 10 seconds:

```
dsmUtil -o CongestionResponseTime=10,SaveSettings
```

To set the path congestion sampling interval to one minute:

```
dsmUtil -o CongestionSamplingInterval=60,SaveSettings
```

To enable path congestion detection:

```
dsmUtil -o CongestionDetectionEnabled=0x1,SaveSettings
```

To use the dsmUtil command to set a path to Admin Offline:

```
dsmUtil -o SetPathOffline=0x77070001
```

To use the dsmUtil command to set a path to online:

```
dsmUtil -o SetPathOnline=0x77070001
```

Running the DSM Failover Driver in a Hyper-V Guest with Pass-Through Disks

Consider a scenario where you map storage to a Windows Server 2008 'R2' (or above) parent partition. You then use the Settings -> SCSI Controller -> Add Hard Drive command to attach that storage as a pass-through disk to the SCSI controller of a Hyper-V guest running Windows Server 2008. By default, some SCSI commands are filtered by Hyper-V, so the DSM Failover driver will fail to run properly.

To work around this issue, SCSI command filtering must be disabled. Run the following PowerShell script in the Parent partition to determine if SCSI pass-through filtering is enabled or disabled:

```
# Powershell Script:  Get_SCSI_Passthrough.ps1
$TargetHost=$args[0]
foreach ($Child in Get-WmiObject -Namespace
root\virtualization Msvm_ComputerSystem -Filter
"ElementName='$TargetHost'")
{
$vmData=Get-WmiObject -Namespace root\virtualization
-Query "Associators of {$Child} Where
ResultClass=Msvm_VirtualSystemGlobalSettingData
AssocClass=Msvm_ElementSettingData"
Write-Host "Virtual Machine:" $vmData.ElementName
Write-Host "Currently Bypassing SCSI Filtering:"
$vmData.AllowFullSCSICommandSet
}?
```

If necessary, run the following PowerShell script in the Parent partition to disabled the SCSI filtering:

```
# Powershell Script:  Set_SCSI_Passthrough.ps1
$TargetHost=$args[0]
$vsManagementService=gwmi
MSVM_VirtualSystemManagementService -namespace
"root\virtualization"
foreach ($Child in Get-WmiObject -Namespace
root\virtualization Msvm_ComputerSystem -Filter
"ElementName='$TargetHost'")
{
$vmData=Get-WmiObject -Namespace root\virtualization
-Query "Associators of {$Child} Where
ResultClass=Msvm_VirtualSystemGlobalSettingData
AssocClass=Msvm_ElementSettingData"
$vmData.AllowFullSCSICommandSet=$true
$vsManagementService.ModifyVirtualSystem($Child,$vmDat
a.PSBase.GetText(1))|out-null
}?
```

dsmUtil Utility

The dsmUtil utility is a command-line driven utility that works only with the Multipath I/O (MPIO) Device Specific Module (DSM) solution. The utility is used primarily as a way to instruct the DSM driver to perform various maintenance tasks, but the utility can also serve as a troubleshooting tool when necessary.

To use the dsmUtil utility, type this command, and press Enter:

```
dsmUtil [[-a [target_id]]
[-c array_name | missing]
[-d debug_level] [-e error_level] [-g
virtual_target_id]
[-o [[feature_action_name[=value]] |
[feature_variable_name=value]][, SaveSettings]] [-M]
[-P [GetMpioParameters | MpioParameter=value | ...]]
[-R]
[-S]
[-s "failback" | "avt" | "busscan" | "forcerebalance"]
[-w target_wwn, controller_index]
```

NOTE You must enter the quotation marks as shown around parameter names.

Typing dsmUtil without any parameters shows the usage information.

The following table shows the dsmUtil parameters.

Table 4 dsmUtil Parameters

Parameter	Description
`-a [target_id]`	Shows a summary of all storage arrays seen by the DSM. The summary shows the `target_id`, the storage array WWID, and the storage array name. If `target_id` is specified, DSM point-in-time state information appears for the storage array. On UNIX operating systems, the virtual HBA specifies unique target IDs for each storage array. The Windows MPIO virtual HBA driver does not use target IDs. The parameter for this option can be viewed as an offset into the DSM information structures, with each offset representing a different storage array. For use by Technical Support Representatives only.
`-c array_name | missing`	Clears the WWN file entries. This file is located in the `Program Files\DSMDrivers\mppdsm\WWN_FILES` directory with the extension `.wwn`. If the `array_name` keyword is specified, the WWN file for the specific storage array is deleted. If the `missing` keyword is used, all WWN files for previously attached storage arrays are deleted. If neither keyword is used, all of the WWN files, for both currently attached and previously attached storage arrays, are deleted.
`-d debug_level`	Sets the current debug reporting level. This option only works if the RDAC driver has been compiled with debugging enabled. For use by Technical Support Representatives only.
`-e error_level`	Sets the current error reporting level to `error_level`, which can have one of these values: ■ `0` – Show all errors. ■ `1` – Show path failover, controller failover, retryable, fatal, and recovered errors. ■ `2` – Show path failover, controller failover, retryable, and fatal errors. ■ `3` – Show path failover, controller failover, and fatal errors. This is the default setting. ■ `4` – Show controller failover and fatal errors. ■ `5` – Show fatal errors. For use by Technical Support Representatives only.
`-g target_id`	Displays detailed information about the state of each controller, path, and LUNs for the specified storage array. You can find the `target_id` by running the `dsmUtil -a` command.
`-M`	Shows the MPIO disk-to-drive mappings for the DSM. The output is similar to that found with the SMdevices utility. For use by Technical Support Representatives only.

Parameter	Description

`-o`
`[[`*`feature_action_nam`*
`e``[=`*`value`*`]] |`
`[`*`feature_variable_na`*
`me``=value]][,`
`SaveSettings]`
	Troubleshoots a feature or changes a configuration setting. Without the `SaveSettings` keyword, the changes only affect the in-memory state of the variable. The `SaveSettings` keyword changes both the in-memory state and the persistent state. Some example commands are:
	
	■ `dsmUtil -o` – Displays all the available feature action names.
	
	■ `dsmUtil -o DisableLunRebalance=0x3` – Turns off the DSM-initiated storage array LUN rebalance (affects only the in-memory state).

`-P`
`[GetMpioParameters |`
`MpioParameter=` *`value`*
`| ...]`
	Displays and sets MPIO parameters.
	
	For use by Technical Support Representatives only.

`-R`	Remove the load-balancing policy settings for inactive devices.
`-S`	Reports the Up state or the Down state of the controllers and paths for each LUN in real time.
`-s ["failback" |` `"avt"	"busscan" |` `"forcerebalance"]`	Manually initiates one of the DSM driver's scan tasks. A "failback" scan causes the DSM driver to reattempt communications with any failed controllers. An "avt" scan causes the DSM driver to check whether AVT has been enabled or disabled for an entire storage array. A "busscan" scan causes the DSM driver to go through its unconfigured devices list to see if any of them have become configured. A "forcerebalance" scan causes the DSM driver to move storage array volumes to their preferred controller and ignores the value of the `DisableLunRebalance` configuration parameter of the DSM driver.
`-w` *`target_wwn`*`,` `controller_index`	For use by Technical Support Representatives only.

Device Manager

Device Manager is part of the Windows operating system. Select **Control Panel** from the **Start** menu. Then select **Administrative Tools >> Computer Management >> Device Manager**.

Scroll down to System Devices to view information about the DSM driver itself. This name might be different based on your network configuration.

The Drives section shows both the drives identified with the HBA drivers and the volumes created by MPIO. Select one of the MPIO volumes, and right-click it. Select **Properties** to open the **Multi-Path Disk Device Properties** window.

This properties window shows if the device is working correctly. Select the **Driver** tab to view the driver information.

Determining if a Path Has Failed

If a controller supports multiple data paths and one of those paths fails, the failover driver logs the path failure in the OS system log file. In the storage management software, the storage array shows a Degraded status.

If all of the paths to a controller fail, the failover driver makes entries in the OS system log that indicate a path failure and failover. In the storage management software, the storage array shows a Needs Attention condition of Volume not on preferred path. The failover event is also written to the Major Event Log (MEL). In addition, if the administrator has configured alert notifications, email messages, or SNMP traps, messages are posted for this condition. The Recovery Guru in the storage management software provides more information about the path failure, along with instructions about how to correct the problem.

NOTE Alert reporting for the Volume not on preferred path condition can be delayed if you set the failover alert delay parameter through the storage management software. When you set this parameter, it imposes a delay on the setting of the Needs Attention condition in the storage management software.

Installing or Upgrading SANtricity ES and DSM on the Windows OS

Perform the steps in this task to install SANtricity ES Storage Manager and the DSM or to upgrade from an earlier release of SANtricity ES Storage Manager and the DSM on a system with a Windows OS. For a clustered system, perform these steps on each node of the system, one node at a time. See I/O Shipping Feature for Asymmetric Logical Unit Access (ALUA) for the procedure to verify or disable ALUA with the Windows OS.

1. Open the SANtricity ES Storage Manager SMIA installation program, which is available from either the SANtricity ES DVD or from your storage vendor's website.

2. Click **Next**.

3. Accept the terms of the license agreement, and click **Next**.

4. Select **Custom**, and click **Next**.

5. Select the applications that you want to install.

   a. Click the name of an application to see its description.

   b. Select the check box next to an application to install it.

6. Click **Next**.

   If you have a previous version of the software installed, you receive a warning message:

   ```
   Existing versions of the following software
   already reside on
   this computer ... If you choose to continue,
   the existing
   versions will be overwritten with new
   versions ....
   ```

7. If you receive this warning and want to update SANtricity ES Storage Manager, click **OK**.

8. Select whether to automatically start the Event Monitor. Click **Next**.

    — Start the Event Monitor for the one I/O host on which you want to receive alert notifications.

    — Do not start the Event Monitor for all other I/O hosts attached to the storage array or for computers that you use to manage the storage array.

9. Click **Next**.

10. If you receive a warning about antivirus or backup software that is installed, click **Continue**.

11. Read the pre-installation summary, and click **Install**.

12. Wait for installation to complete, and click **Done**.

Removing SANtricity ES Storage Manager and the DSM Failover Driver from the Windows OS

NOTE To prevent loss of data, the host from which you are removing SANtricity ES Storage Manager and the DSM must have only one path to the storage array. Reconfigure the connections between the host and the storage array to remove any redundant connections before you uninstall SANtricity ES Storage Manager and the DSM failover driver.

1. From the Windows **Start** menu, select **Control Panel**.

    The **Control Panel** window appears.

2. In the **Control Panel** window, double-click **Add or Remove Programs**.

    The **Add or Remove Programs** window appears.

3. Select **SANtricity ES Storage Manager**.

4. Click the **Remove** button to the right of the **SANtricity ES Storage Manager** entry.

WinObj

You can use WinObj to view the Object Manager namespace that is maintained by the OS. Every Windows OS driver that creates objects in the system can associate a name with the object that can be viewed from WinObj. With WinObj, you can view the volumes and paths that the host bus adapters (HBAs) have identified. You can also view what the failover driver identifies from a storage array.

Device Specific Module Driver Directory Structures

NOTE The name MPPDSM in the directory structures might be different based on your network configuration.

The directory structures for the DSM driver include these paths:

■ \Device\MPPDSM – This structure contains information that is maintained by the DSM driver. The objects shown in the \Device\MPPDSM directory in the following table show the items that are reported by MPIO to the DSM driver. If a device is not in this list, MPIO has not notified the DSM driver.

- \Device\Scsi – This structure contains information that is maintained by the ScsiPort driver. The objects shown in the \Device\Scsi directory in the following table show the physical volumes that are identified by the HBAs. If a specific volume is not in this list, the DSM driver cannot detect the volumes.

Table 5 Object Path and Descriptions of the WinObj DSM

Object Path	Description
Device\MPPDSM	The root directory for all named objects that are created by the DSM driver.
\Device\MPPDSM\<*storage array*>	The root directory for all named objects that are created by the storage array named <*storage array*>.
\Device\MPPDSM\<*storage array*>\<*ctlr*>	The root directory for all named objects that are created for a given controller. The <*ctlr*> value can either be A or B.
\Device\MPPDSM\<*storage array*>\<*ctlr*>\ P<*port*>P<*path*>I<*id*>	The root directory for all named objects that are created for a given path to a controller. The <*port*> value, the <*path*> value, and the <*id*> value represent a SCSI address from a given HBA port.
\Device\MPPDSM\<storage array>\<*ctlr*>\ P<*port*>P<*path*>I<*id*>\<*lun*>	The <*lun*> value represents the volume number assigned to the device for a given controller/path combination.
\Device\Scsi	The root directory for all of the named objects created by the ScsiPort driver. Each object represents a physical path found by a given HBA.
\Device\Scsi\<*adapter*>Port<*port*>\ Path<*path*>Target<*target*>Lun<*lun*>	ScsiPort-based HBA drivers. A named device object that represents a drive. The <*adapter*> value represents the HBA vendor. For QLogic, this value is based on the HBA model number (for example, ql2300). The <*port*> value, the <*path*> value, and the <*target*> value represent the location of the volume on the HBA.
\Device\<*auto-generated id*>	StorPort-based HBA drivers. An auto-generated named device object representing a drive.

The following general statements apply to the object paths:

- The objects shown in the \Device\Scsi directory show the physical volumes that are identified by the HBAs. If a specific volume is not in this list, the DSM driver cannot detect the volumes.

- The objects shown in the `\Device\MPPDSM` directory show the items that are reported by MPIO to the DSM driver. If a device is not in this list, MPIO has not notified the DSM driver.

Frequently Asked Questions about Windows Failover Drivers

The following table lists answers to questions about Windows failover drivers.

Table 6 Frequently Asked Questions about Windows Failover Drivers

Question	Answer
My disk devices or host bus adapters (HBAs) show a yellow exclamation point. What does this mean?	When you use Device Manager, you might observe that a disk device icon or an HBA icon has a yellow exclamation point on it. If new volumes have been mapped to the host, the exclamation point might appear on the icon for a few seconds. This action occurs because the PnP Manager is reconfiguring the device, and, during this time, the device or the HBA might not be used. If the exclamation point stays for more than one minute, a configuration error has occurred.
My disk devices or HBAs show a red X. What does this mean?	When you use Device Manager, you might notice that a disk device icon or an HBA icon has a red X on it. This X indicates that the device has been disabled. A disabled device cannot be used or communicated with until it is re-enabled. If the disabled device is an adapter, any disk devices that were connected to that adapter are removed from Device Manager.

Question	Answer
Why does the SMdevices utility not show any volumes?	If the SMdevices utility does not show any volumes, perform these steps: 1. Make sure that all cables are seated correctly. Make sure that all gigabit interface converters (GBICs) are seated correctly. 2. Determine the HBA BIOS and driver versions that the system uses, and make sure that the HBA BIOS and driver versions are correct. 3. Make sure that your mappings are correct. Do not use any HBA mapping tools. 4. Use WinObj to determine if the host has detected the volumes. If the host has not detected the volumes, an HBA problem or a controller problem has occurred. Make sure that the HBAs are logging into the switch or the controller. If they are not logging in, the problem is probably HBA related. If the HBAs have logged into the controller, a controller issue might be the problem.
The SMdevices utility shows duplicate entries for some or all of my disks.	You see that some of your disks show up twice when you run the SMdevices utility. For the Windows Server OS, something went wrong with the device-claiming process.
I run the hot_add utility, but my disks do not appear.	See "Why does the SMdevices utility not show any volumes?
I have mapped new volumes to my host, but I cannot see them.	Run the hot_add utility. See "Why does the SMdevices utility not show any volumes?"
How do I know if a host has detected my volumes?	Use WinObj to determine if the host can see the volumes. ■ If the host cannot see the volumes, an HBA problem or a controller problem has occurred. ■ Make sure that the HBAs log into the switch or the controller. If they are not logging in correctly, the problem is probably HBA related. ■ If the HBAs have logged into the controller, the problem might be a controller issue.

Question	Answer
After I install the DSM driver, my system takes a long time to start. Why?	You might still experience long start times after you install the DSM driver because the Windows OS is completing its configuration for each device. For example, you install the DSM driver on a host with no storage array attached, and you restart the host. Before the Windows OS actually starts, you plug in a cable to a storage array with 32 volumes. In the start-up process, PnP detects the configuration change and starts to process it. After the configuration change has completed, subsequent restarts do not experience any delays unless additional configuration changes are detected. The same process can occur even if the host has already started.
What host type must I use for the MPIO solution?	If you use Windows Server Failover Cluster, select a host type of **Windows Clustered** otherwise, select a host type of **Windows**.
How can I tell if MPIO is installed?	Perform these steps: 1. Go to the **Control Panel** on the **Start** menu, and double-click **Administrative Tools**. 2. Select **Computer Management >> Device Manager >> SCSI and RAID controllers**. 3. On the Windows Server 2008 OS, look for Microsoft Multi-Path Bus Driver. If this item is present, MPIO is installed.
How can I tell if the DSM driver is installed?	Perform these steps: 1. Go to the **Control Panel** on the **Start** menu, and double-click **Administrative Tools**. 2. Select **Computer Management >> Device Manager >> System Devices**. 3. Look for the NetApp-supported DSM. The name ends with the text Device-Specific Module for Multi-Path. If it is present, the DSM is installed.
What is the default vendor ID string and the product ID string?	By default, the vendor ID string and the product ID string configured for NetApp storage arrays are named NETAPP/INF-01-00. If not, the PnP manager cannot choose the failover driver to manage the volume. The failover driver takes over, which causes delays. If you suspect that this event has occurred, check the non-user configuration region of the controller firmware.

Question	Answer

What should I do if I receive this message?
 	
`Warning: Changing the storage array name can cause host applications to lose access to the storage array if the host is running certain path failover drivers.`
 	
`If any of your hosts are running path failover drivers, please update the storage array name in your path failover driver's configuration file before rebooting the host machine to insure uninterrupted access to the storage array. Refer to your path failover driver documentation for more details.`
	You do not need to update files. The information is dynamically created only when the storage array is found initially. Use one of these two options to correct this behavior:
	
	■ Restart the host server.
	
	■ Unplug the storage array from the host server, and perform a rescan of all of the devices. After the devices have been removed from the storage array, you can re-attach them. Another rescan takes place, which rebuilds the information with the updated names.

Failover Drivers for the Linux Operating System 3

Redundant Dual Active Controller (RDAC) is the supported failover driver for SANtricity ES Storage Manager with Linux operating systems.

Linux OS Restrictions

This version of the Linux OS RDAC failover driver does not support any Linux OS 2.4 kernels, such as the following:

- SUSE SLES 8 OS
- Red Hat 3 Linux OS
- SLES 8 OS and Red Hat 3 Linux OS on POWER (LoP) servers

The Linux OS RDAC failover driver cannot coexist with failover or failback drivers that run at the level or a Fibre Channel host bus adapter (HBA). Examples of HBA-level failover or failback drivers include:

- The 8.00.00-fo or 8.00.02-fo Linux OS device driver for the IBM DS4000 fc2-133 HBAs driver on servers with Intel architecture processors or AMD architecture processors
- The QLA driver for the Fibre Channel expansion adapters on the IBM LoP blade servers

You might have to modify the makefile of the HBA driver for it to be compiled in the non-failover mode.

Auto-Volume Transfer (AVT) mode is automatically enabled as the SANshare Storage Partitioning host type for the Linux OS.

Features of the RDAC Failover Driver from NetApp

Redundant Dual Active Controller (RDAC) is the failover driver for the Linux OS that is included in SANtricity ES Storage Manager. The RDAC failover driver includes these features:

- On-the-fly path validation.
- Cluster support.
- Automatic detection of path failure. The RDAC failover driver automatically routes I/O to another path in the same controller or to an alternate controller, in case all paths to a particular controller fail.
- Retry handling is improved, because the RDAC failover driver can better understand vendor-specific statuses returned from the controller through sense key/ASC/ASCQ.
- Automatic rebalance is handled. When the failed controller obtains Optimal status, storage array rebalance is performed automatically without user intervention.

- Three load-balancing policies are supported: round robin subset, least queue depth, and path weight.

Prerequisites for Installing RDAC on the Linux OS

Before installing RDAC on the Linux OS, make sure that your storage array meets these conditions:

- Make sure that the host system on which you want to install the RDAC driver has supported HBAs.
- Refer to the installation electronic document topics for your controller tray or controller-drive tray for any configuration settings that you need to make.
- Although the system can have Fibre Channel HBAs from multiple vendors or multiple models of Fibre Channel HBAs from the same vendor, you can connect only the same model of Fibre Channel HBAs to each storage array.
- Make sure that the low-level HBA driver has been correctly built and installed before RDAC driver installation.
- The standard HBA driver must be loaded before you install the RDAC driver. The HBA driver has to be a non-failover driver.
- For LSI HBAs, the port driver is named `mptbase`, and the host driver is named `mptscsi` or `mptscsih`, although the name depends on the driver version. The Fibre Channel driver is named `mptfc`, the SAS driver is named `mptsas`, and the SAS2 driver is named `mpt2sas`.
- For QLogic HBAs, the base driver is named `qla2xxx`, and host driver is named `qla2300`. The 4-GB HBA driver is named `qla2400`.
- For IBM Emulex HBAs, the base driver is named `lpfcdd` or `lpfc`, although the name depends on the driver version.
- For Emulex HBAs, the base driver is named `lpfcdd` or `lpfc`, although the name depends on the driver version.
- Make sure that the kernel source tree for the kernel version to be built against is already installed. You must install the kernel source rpm on the target system for the SUSE SLES OS. You are not required to install the kernel source for the Red Hat OS.
- Make sure that the necessary kernel packages are installed: `source rpm` for the SUSE OS and `kernel headers/kernel devel` for the Red Hat Enterprise Linux OS.

In SUSE OSs, you must include these items for the HBAs mentioned as follows:

- For LSI HBAs, INITRD_MODULES includes `mptbase` and `mptscsi` (or `mptscsih`) in the `/etc/sysconfig/kernel` file. The Fibre Channel driver is named `mptfc`, the SAS driver is named `mptsas`, and the SAS2 driver is named `mpt2sas`.
- For QLogic HBAs, INITRD_MODULES includes a `qla2xxx` driver and a `qla2300` driver in the `/etc/sysconfig/kernel` file.

- For IBM Emulex HBAs, INITRD_MODULES includes an `lpfcdd` driver or an `lpfc` driver in the `/etc/sysconfig/kernel` file.
- For Emulex HBAs, INITRD_MODULES includes an `lpfcdd` driver or an `lpfc` driver in the `/etc/sysconfig/kernel` file.

Installing SANtricity ES Storage Manager and RDAC on the Linux OS

NOTE SANtricity ES Storage Manager requires that the different Linux OS kernels have separate installation packages. Make sure that you are using the correct installation package for your particular Linux OS kernel.

1. Open the SANtricity ES Storage Manager SMIA installation program, which is available from either the SANtricity ES DVD or from your storage vendor's website.

   The SANtricity ES Storage Manager installation window appears.
2. Click **Next**.
3. Accept the terms of the license agreement, and click **Next**.
4. Select one of the installation packages:
   — **Typical** – Select this option to install all of the available host software.
   — **Management Station** – Select this option to install software to configure, manage, and monitor a storage array. This option does not include RDAC. This option only installs the client software.
   — **Host** – Select this option to install the storage array server software.
   — **Custom** – Select this option to customize the features to be installed.
5. Click **Next**.

NOTE For this procedure, **Typical** is selected. If the **Host** installation option is selected, the Agent, the Utilities, and the RDAC driver will be installed.

You might receive a warning after you click **Next**. The warning states:

```
Existing versions of the following software
already reside on
this computer ... If you choose to continue,
the existing
versions will be overwritten with new
versions ....
```

If you receive this warning and want to update to SANtricity ES Storage Manager Version 10.80, click **OK**.

6. Click **Install**.

You will receive a warning after you click **Install**. The warning tells you that the RDAC driver is not automatically installed. You must manually install the RDAC driver.

The RDAC source code is copied to the specified directory in the warning message. Go to that directory, and perform the steps in Installing RDAC Manually on the Linux OS.

7. Click **Done**.

Installing RDAC Manually on the Linux OS

1. To unzip the RDAC `tar.gz` file and enter the RDAC tar file, type this command, and press **Enter**:

       tar -zxvf <*filename*>

2. Go to the Linux RDAC directory.

3. Type this command, and press **Enter**.

       make uninstall

4. To remove the old driver modules in that directory, type this command, and press **Enter**:

       make clean

5. To compile all driver modules and utilities in a multiple CPU server (SMP kernel), type this command, and press **Enter**:

       make

6. Type this command, and press **Enter**:

       make install

These actions result from running this command:

— The driver modules are copied to the kernel module tree.
— The new RAMdisk image (mpp-`uname -r`.img) is built, which includes the RDAC driver modules and all driver modules that are needed at boot.

7. Follow the instructions shown at the end of the build process to add a new boot menu option that uses /boot/mpp-`uname -r`.img as the initial RAMdisk image.

Making Sure that RDAC Is Installed Correctly on the Linux OS

1. Restart the system by using the new boot menu option.
2. Make sure that these driver stacks were loaded after restart:
   — `scsi_mod`
   — `sd_mod`
   — `sg`
   — `mppUpper`
   — The physical HBA driver module
   — `mppVhba`
3. Type this command, and press **Enter**:

   ```
   /sbin/lsmod
   ```
4. To make sure that the RDAC driver discovered the available physical volumes and created virtual volumes for them, type this command, and press **Enter**:

   ```
   /opt/mpp/lsvdev
   ```

   You can now send I/O to the volumes.
5. If you make any changes to the RDAC configuration file (`/etc/mpp.conf`) or the persistent binding file (`/var/mpp/devicemapping`), run the `mppUpdate` command to rebuild the RAMdisk image to include the new file. In this way, the new configuration file (or persistent binding file) can be used on the next system restart.
6. To dynamically reload the driver stack (`mppUpper`, physical HBA driver modules, `mppVhba`) without restarting the system, perform these steps:
   a. Remove all of the configured scsi devices from the system.
   b. To unload the `mppVhba` driver, type this command, and press **Enter**:

      ```
      rmmod mppVhba
      ```
   c. To unload the physical HBA driver, type this command, and press **Enter**:

      ```
      modprobe -r "physical hba driver modules"
      ```
   d. To unload the `mppUpper` driver, type this command, and press **Enter**:

      ```
      rmmod mppUpper
      ```
   e. To reload the `mppUpper` driver, type this command, and press **Enter**:

      ```
      modprobe mppUpper
      ```
   f. To reload the physical HBA driver, type this command, and press **Enter**:

      ```
      modprobe "physical hba driver modules"
      ```
   g. To reload the `mppVhba` driver, type this command, and press **Enter**:

      ```
      modprobe mppVhba
      ```
7. Restart the system whenever there is an occasion to unload the driver stack.

8. Use a utility, such as `devlabel`, to create user-defined device names that can map devices based on a unique identifier, called a UUID.

9. Use the `udev` command for persistent device names. The `udev` command dynamically generates device name links in the `/dev/disk` directory based on path, ID or UUID.

```
linux-kbx5:/dev/disk # ls /dev/disk by-id
by-path  by-uuid
```

For example, the `/dev/disk/by-id` directory links volumes that are identified by WWIDs of the volumes to actual disk device nodes.

```
lrwxrwxrwx 1 root root 10 Feb 23 12:15
scsi-3600a0b80000c2df9000003b141417799 ->
../../sdda

lrwxrwxrwx 1 root root  9 Feb 23 12:15
scsi-3600a0b80000f27030000000d416b94fd ->
../../sdc

lrwxrwxrwx 1 root root  9 Feb 23 12:15
scsi-3600a0b80000f270300000015416b958f ->
../../sdg
```

Configuring Failover Drivers for the Linux OS

The Windows OS and the Linux OS share the same set of tunable parameters to enforce the same I/O behaviors. For a description of these parameters, go to Configuration Settings for the Windows DSM and the Linux RDAC.

Table 1 Configuration Settings for Failover Drivers for the Linux OS

Parameter Name	Default Value	Description
ImmediateVirtLunCreate	0	This parameter determines whether to create the virtual LUN immediately if the owning physical path is not yet discovered. This parameter can take the following values: ■ `0` – Do not create the virtual LUN immediately if the owning physical path is not yet discovered. ■ `1` – Create the virtual LUN immediately if the owning physical path is not yet discovered.
BusResetTimeout		The time, in seconds, for the RDAC driver to delay before retrying an I/O operation if the DID_RESET status is received from the physical HBA. A typical setting is `150`.

Parameter Name	Default Value	Description
AllowHBAsgDevs	0	This parameter determines whether to create individual SCSI generic (SG) devices for each I:T:L for the end LUN through the physical HBA. This parameter can take the following values: ■ `0` – Do not allow creation of SG devices for each I:T:L through the physical HBA. ■ `1` – Allow creation of SG devices for each I:T:L through the physical HBA.

Compatibility and Migration

Controller firmware – The Linux OS RDAC driver is compatible with the controller firmware. However, the Linux OS RDAC driver does not support SCSI-2 to SCSI-3 reservation translation unless the release is version 8.40.*xx* or later.

Linux OS distributions – The Linux OS RDAC driver is intended to work on any Linux OS distribution that has the standard SCSI I/O storage array (SCSI middle-level and low-level interfaces). This release is targeted specifically at SUSE Linux OS Enterprise Server and Red Hat Advanced Server.

mppUtil Utility

The mppUtil utility is a general-purpose command-line driven utility that works only with MPP-based RDAC solutions. The utility instructs RDAC to perform various maintenance tasks but also serves as a troubleshooting tool when necessary.

To use the mppUtil utility, type this command, and press **Enter**:

```
mppUtil [-a target_name] [-c wwn_file_name] [-d
debug_level]
[-e error_level] [-g virtual_target_id] [-I host_num]
[-o feature_action_name[=value][, SaveSettings]]
[-s "failback" | "avt" | "busscan" | "forcerebalance"]
[-S] [-U]
[-V] [-w target_wwn,controller_index]
```

NOTE The quotation marks must surround the parameters.

The mppUtil utility is a cross-platform tool. Some parameters might not have a meaning in a particular OS environment. A description of each parameter follows.

Table 2 mppUtil Parameters

Parameter	Description
-a *target_name*	Shows the RDAC driver's internal information for the specified virtual *target_name* (storage array name). If a *target_name* value is not included, the -a parameter shows information about all of the storage arrays that are currently detected by this host.
-c *wwn_file_name*	Clears the WWN file entries. This file is located at /var/mpp with the extension .wwn.
-d *debug_level*	Sets the current debug reporting level. This option works only if the RDAC driver has been compiled with debugging enabled. Debug reporting is comprised of two segments. The first segment refers to a specific area of functionality, and the second segment refers to the level of reporting within that area. The *debug_level* is one of these hexadecimal numbers: ■ 0x20000000 – Shows messages from the RDAC driver's init() routine. ■ 0x10000000 – Shows messages from the RDAC driver's attach() routine. ■ 0x08000000 – Shows messages from the RDAC driver's ioctl() routine. ■ 0x04000000 – Shows messages from the RDAC driver's open() routine. ■ 0x02000000 – Shows messages from the RDAC driver's read() routine. ■ 0x01000000 – Shows messages related to HBA commands. ■ 0x00800000 – Shows messages related to aborted commands. ■ 0x00400000 – Shows messages related to panic dumps. ■ 0x00200000 – Shows messages related to synchronous I/O activity. ■ 0x00000001 – Debug level 1. ■ 0x00000002 – Debug level 2. ■ 0x00000004 – Debug level 3. ■ 0x00000008 – Debug level 4. These options can be combined with the logical AND operator to provide multiple areas and levels of reporting as needed. For use by Technical Support Representatives only.

Parameter	Description
`-e error_level`	Sets the current error reporting level to *error_level*, which can have one of these values: ■ `0` – Show all errors. ■ `1` – Show path failover, controller failover, retryable, fatal, and recovered errors. ■ `2` – Show path failover, controller failover, retryable, and fatal errors. ■ `3` – Show path failover, controller failover, and fatal errors. This is the default setting. ■ `4` – Show controller failover and fatal errors. ■ `5` – Show fatal errors. For use by Technical Support Representatives only.
`-g virtual_target_id`	Display the RDAC driver's internal information for the specified *virtual_target_id*.
`-I host_num`	Prints the maximum number of targets that can be handled by that host. Here, host refers to the HBA drivers on the system and includes the RDAC driver. The host number of the HBA driver is given as an argument. The host numbers assigned by the Linux middle layer start from 0. If two ports are on the HBA card, host numbers 0 and 1 would be taken up by the low-level HBA driver, and the RDAC driver would be at host number 2. Use `/proc/scsi` to determine the host number.
`-o feature_action_name[=value][, SaveSettings]`	Troubleshoots a feature or changes a configuration setting. Without the `SaveSettings` keyword, the changes affect only the in-memory state of the variable. The `SaveSettings` keyword changes both the in-memory state and the persistent state. You must run `mppUpdate` to reflect these changes in the inird image before rebooting the server. Some example commands are: ■ `mppUtil -o` – Displays all the available feature action names. ■ `mppUtil -o ErrorLevel=0x2` – Sets the `ErrorLevel` parameter to `0x2` (affects only the in-memory state).

Parameter	Description
`-s ["failback" | "avt" |` `"busscan" | "forcerebalance"]`	Manually initiates one of the RDAC driver's scan tasks. ■ A "failback" scan causes the RDAC driver to reattempt communications with any failed controllers. ■ An "avt" scan causes the RDAC driver to check whether AVT has been enabled or disabled for an entire storage array. ■ A "busscan" scan causes the RDAC driver to go through its unconfigured devices list to see if any of them have become configured. ■ A "forcerebalance" scan causes the RDAC driver to move storage array volumes to their preferred controller and ignore the value of the `DisableLunRebalance` configuration parameter of the RDAC driver.
`-S`	Reports the Up state or the Down state of the controllers and paths for each LUN in real time.
`-U`	Refreshes the Universal Transport Mechanism (UTM) LUN information in MPP driver internal data structure for all the storage arrays that have already been discovered.
`-V`	Prints the version of the RDAC driver currently running on the system.
`-w target_wwn,controller_index`	For use by Technical Support Representatives only.

Frequently Asked Questions about Linux Failover Drivers

Table 3 Frequently Asked Questions about Linux Failover Drivers

Question	Answer
How do I get logs from RDAC in the Linux OS?	Use the `mppSupport` command to obtain several logs related to RDAC. The `mppSupport` command is found in the `/opt/mpp/mppSupport` directory. The command creates a file named `mppSupportdata_hostname_RDAC version_datetime`.`tar.gz` in the `/tmp` directory.
How does persistent naming work?	The Linux OS SCSI device names can change when the host system restarts. Use a utility, such as devlabel, to create user-defined device names that will map devices based on a unique identifier. The udev method is the preferred method.

Question	Answer
What must I do after applying a kernel update?	After you apply the kernel update and start the new kernel, perform these steps to build the RDAC Initial Ram Disk image (initrd image) for the new kernel: 1. Change the directory to the Linux RDAC source code directory. 2. Type `make uninstall`, and press **Enter**. 3. Reinstall RDAC. Go to Installing RDAC Manually on the Linux OS on page Installing RDAC Manually on the Linux OS.
What is the Initial Ram Disk Image (initrd image), and how do I create a new initrd image?	The initrd image is automatically created when the driver is installed by using the `make install` command. The boot loader configuration file must have an entry for this newly created image. The initrd image is located in the boot partition. The file is named `mpp'-uname -r'.img`. For a driver update, if the system already has a previous entry for RDAC, the system administrator must modify the existing RDAC entry in the boot loader configuration file. In most of the cases, no change is required if the kernel version is the same. To create a new initrd image, type `mppUpdate`, and press **Enter**. The old image file is overwritten with the new image file. For the SUSE OS, if third-party drivers need to be added to the initrd image, change the `/etc/sysconfig/ kernel` file with the third-party driver entries. Run the `mppUpdate` command again to create a new initrd image.
How do I remove unmapped or disconnected devices from the existing host?	Run `hot_add -d` to remove all unmapped or disconnected devices.
What if I remap a LUN from the storage array?	Run `hot_add -u` to update the host with the changed LUN mapping.
What if I change the size of the LUN on the storage array?	Run `hot_add -c` to change the size of the LUN on the host.

How do I make sure that RDAC finds the available storage arrays?	To make sure that the RDAC driver has found the available storage arrays and created virtual storage arrays for them, type these commands, and press **Enter** after each command.
	
	`ls -lR /proc/mpp`
	
	`mppUtil -a`
	
	`/opt/mpp/lsvdev`
	
	To show all attached and discovered volumes, type `cat /proc/scsi/scsi,` and press **Enter**.

Question	Answer

What should I do if I receive this message?
 	
`Warning: Changing the storage array name can cause host applications to lose access to the storage array if the host is running certain path failover drivers.`
 	
`If any of your hosts are running path failover drivers, please update the storage array name in your path failover driver's configuration file before rebooting the host machine to insure uninterrupted access to the storage array. Refer to your path failover driver documentation for more details.`
	The path failover drivers that cause this warning are the RDAC drivers on both the Linux OS and the Windows OS.
	
	The storage array user label is used for storage array-to-virtual target ID binding in the RDAC driver. For the Linux OS, change this file to add the storage array user label and its virtual target ID.
	`.~ # more /var/mpp/devicemapping`

Device Mapper Multipath for the Linux Operating System   *4*

Device Mapper (DM) is a generic framework for block devices provided by the Linux operating system. It supports concatenation, striping, snapshots (legacy), mirroring, and multipathing. The multipath function is provided by the combination of the kernel modules and user space tools.

Device Mapper Features

- Provides a single block device node for a multipathed logical unit
- Ensures that I/O is re-routed to available paths during a path failure
- Ensures that the failed paths are revalidated as soon as possible
- Configures the multipaths to maximize performance
- Reconfigures the multipaths automatically when events occur
- Provides DMMP features support to newly added logical unit
- Provides device name persistency for DMMP devices under `/dev/mapper/`
- Configures multipaths automatically at an early stage of rebooting to permit the OS to install and reboot on a multipathed logical unit

Known Limitations and Issues of the Device Mapper

- In certain error conditions, with `no_path_retry` or `queue_if_no_path` feature set, applications might hang forever. To overcome these conditions, you must enter the following command to all the affected multipath devices: `dmsetup message device 0 "fail_if_no_path"`, where `device` is the multipath device name (for example, mpath2; do not specify the path).
- Normally, the scsi_dh_rdac module is not included in initrd image. When this module is not included, boot-up might be very slow with large configurations and the `syslog` file might contain a large number of buffer I/O errors. You should include scsi_dh_rdac in the initrd to avoid these problems. The installation procedures for each operating system describe how to include the driver in initrd image. Refer to the appropriate procedure in Device Mapper Operating Systems Support.
- If the storage vendor and model are not included in scsi_dh_rdac device handler, device discovery might be slower, and the `syslog` file might get populated with buffer I/O error messages.
- Use of the DMMP and RDAC failover solutions together on the same host is not supported. Use only one solution at a time.

Device Mapper Operating Systems Support

NetApp supports device mapper for SLES 11 and RHEL 6.0 onwards. All future updates of these OS versions are also supported. The following sections provide specific information on each of these operating systems.

I/O Shipping Feature for Asymmetric Logical Unit Access (ALUA) with Linux Operating Systems

The I/O Shipping feature implements support for ALUA. With ALUA, a duplex storage array can service I/O requests to any LUN through either controller. Each LUN is owned by only one of the controllers, and the optimized path to the LUN is through the owning controller. There is, however, an unoptimized path to each LUN through the other (non-owning) controller. There is a performance penalty when the non-owning controller accesses a volume. This behavior is optimized with combination of DMMP and storage firmware.

The ALUA feature is supported from Linux versions SLES11.1 and RHEL 6.1 onwards. You must download and install the Linux RPM packages provided by NetApp to make use of this feature on SLES 11.1 and RHEL6.1. If your system shipped with DVDs, the rpm packages may be found on the DVD. If your system did not ship with DVDs, the rpm packages may be found on your storage vendor's website. Note that these packages are applicable only to SLES11.1 and RHEL 6.1. These packages will not be required in SLES11.2, RHEL6.2 and subsequent releases for SLES and RHEL.

Installing the Device Mapper Multi-Path for the SUSE Linux Enterprise Server Version 11 Service Pack 2

All of the components required for DMMP are included on the SUSE Linux Enterprise Server (SLES) version 11.2 installation media. By default, DMMP is disabled in SLES. Complete the following steps to enable DMMP components on the host.

1. If you have not already installed the operating system, use the media supplied by your operating system vendor to install SLES 11.2.

2. If your storage array includes an E5512 controller-drive tray, an E5524 controller-drive tray, or an E5560 controller-drive tray, perform substeps a through e; otherwise, continue with step 3.

   With the E5500 controller-drive trays, the following substeps are required to install maintenance kernel 3.0.34-0.7.9 for the operating system.

   a. Download the kernel-default-base package and the kernel-default package for the maintenance kernel from the Novell website.

   b. On the command line, type `rpm -ivh` *kernel-default-base*, where *kernel-default-base* is the file path for the kernel-default-base package.

   c. On the command line, type `rpm -ivh` *kernel-default*, where *kernel-default* is the file path for the kernel-default package.

   d. Make sure that the boot loader configuration file, (`grub.conf` or `lilo.conf` or `yaboot.conf`) is updated with the newly built kernel and initrd.

      e.   Reboot the host.

3. Use the procedures under Setting Up the multipath.conf File on page 53 to update and configure the `/etc/multipath.conf` file.

4. On the command line, type `chkconfig multipath on`.

   The multipathd daemon is enabled when the system starts again.

5. To add the scsi_dh_rdac to initrd image, edit the `/etc/sysconfig/kernel` file to add the directive scsi_dh_rdac to the INITRD_MODULES section of the file.

6. On the command line, type `mkinitrd -k /boot/vmlinux-$(uname -r) -i /boot/initrd-$(uname -r)-scsi_dh -M /boot/System.map-$(uname -r)` to create the new initrd image.

7. Update the boot loader configuration file, (`grub.conf` or `lilo.conf` or `yaboot.conf`) with the newly built initrd.

ATTENTION You must check the value for host type and, if necessary, change the setting of that value to ensure that ALUA can be enabled.

---

8. Do one of the following to verify and, if necessary, change the host type.
   — If you have hosts defined, go to step 9.
   — If you do not have hosts defined, right-click the default host group and then set the default host type to **Linux (DM-MP)**. Go to step 11.

9. In the SANtricity ES Storage Manager mappings view, right-click on the host and select **Change Host Operating System.**

10. Verify that the selected host type is **Linux (DM-MP)**. If necessary, change the selected host type to **Linux (DM-MP)**.

11. Reboot the host.

Installing the Device Mapper Multi-Path for the Red Hat Enterprise Linux Operating System Version 6.3

All of the components required for DMMP are included on the Red Hat Enterprise Linux (RHEL) version 6.3 installation media. By default, DMMP is disabled in RHEL. Complete the following steps to enable DMMP components on the host.

1. If you have not already installed the operating system, use the media supplied by your operating system vendor to install RHEL.

2. Use the procedures under Setting Up the multipath.conf File on page 53 to update and configure the `/etc/multipath.conf` file.

3. On the command line, type `chkconfig multipath on`.

   The multipathd daemon is enabled when the system starts again.

4. Perform one of the following options to preload `scsi_dh_rdac` during boot up.

   — Use a text editor to add the command `rdloaddriver=scsi_dh_rdac` at the end of the kernel options in your boot loader configuration file.

   — Enter the following commands to create the file `initramfs` and insert the command `scsi_dh_rdac` into the file `/etc/cmdline`.

   ```
   #echo "rdloaddriver=scsi_dh_rdac" >
   /etc/cmdline

   #dracut --install  '/etc/cmdline' -f
   /boot/initramfs-'uname -r' -rdac.img
   ```

5. Update the boot loader configuration file according to the option selected in the preceding step.

ATTENTION You must check the value for host type and, if necessary, change the setting of that value to ensure that ALUA can be enabled.

---

6. Do one of the following to verify and, if necessary, change the host type.

   — If you have hosts defined, go to step 7.

   — If you do not have hosts defined, right-click the default host group and then set the default host type to **Linux (DM-MP)**. Go to step 11.

7. In the SANtricity ES Storage Manager mappings view, right-click the host and select **Change Host Operating System.**

8. Verify that the selected host type is **Linux (DM-MP)**. If necessary, change the selected host type to **Linux (DM-MP)**.

9. Reboot the host.

Verifying that the I/O Shipping Feature for ALUA Support is Installed on the Linux OS

When you install or update host software and controller firmware to SANtricity ES version 10.83 and above and CFW version 7.83 and above and install DMMP on the Linux OS, support for ALUA is enabled. Installation must include the host software on the management station. Perform the following steps to verify that ALUA support is installed.

1. Perform one of the following actions to confirm that the host can see the LUNs that are mapped to it.

   — At the command prompt, type `SMdevices`. The appearance of `active optimized` or `active non-optimized` (rather than `passive` or `unowned`) in the output indicates that the host can see the LUNs that are mapped to it.

   — At the command prompt, type `multipath -ll`. If the host can see the LUNs that are mapped to it, both the path groups are displayed as `active ready` instead of `active ghost`.

2. Check the log file at /var/log/messages for entries similar to scsi 3:0:2:0: rdac: LUN 0 (IOSHIP).

   These entries indicate that the scsi_dh_rdac driver correctly recognizes ALUA mode. The keyword IOSHIP refers to ALUA mode. These messages are displayed when the devices are discovered in the system. These messages might also be displayed in dmesg logs or boot logs.

Setting Up the multipath.conf File

The multipath.conf file is the configuration file for the multipath daemon, multipathd. The multipath.conf file overrides the built-in configuration table for multipathd. Any line in the file whose first non-white-space character is # is considered a comment line. Empty lines are ignored.

Example multipath.conf are available in the following locations:

- For SLES, /usr/share/doc/packages/multipath-tools/multipath.conf.synthetic
- For RHEL, /usr/share/doc/device-mapper-multipath-0.4.9/multipath.conf

All the lines in the sample multipath.conf file are commented out. The file is divided into five sections:

- **defaults** – Specifies all default values.
- **blacklist** – All devices are blacklisted for new installations. The default blacklist is listed in the commented-out section of the /etc/multipath.conf file. Blacklist the device mapper multipath by WWID if you do not want to use this functionality.
- **blacklist_exceptions** – Specifies any exceptions to the items specified in the section blacklist.
- **devices** – Lists all multipath devices with their matching vendor and product values.
- **multipaths** – Lists the multipath device with their matching WWID values.

In the following tasks, you modify the default, blacklist and devices sections of the multipath.conf file. Remove the initial # character from the start of each line you modify.

**Updating the
Blacklist Section**

With the default settings, UTM LUNs might be presented to the host. I/Os operations,
however, are not supported on UTM LUNs. To prevent I/O operations on the UTM
LUNs, add the vendor and product information for each UTM LUN to the blacklist
section of the `/etc/multipath.conf` file. The entries should follow the pattern
of the following example.

```
blacklist {
device {
        vendor "*"
        product "Universal Xport"
    }
}
```

**Updating the
Devices Section of
the multipath.conf
File**

The following example shows part of the `devices` section in the
`/etc/multipath.conf` file. The the example shows the vendor ID as `NETAPP`
or `LSI` and the product ID as `INF-01-00`. Modify the `devices` section with
product and vendor information to match the configuration of your storage array. If
your storage array contains devices from more than one vendor, add additional
`device` blocks with the appropriate attributes and values under the `devices`
section.

```
devices {
  device {
    vendor              "(LSI|NETAPP)"
    product             "INF-01-00"
    path_grouping_policy  group_by_prio
    prio               rdac
    getuid_callout    "/lib/udev/scsi_id -g -u -d
/dev/%n"
    polling_interval 5
    path_checker      rdac
    path_selector     "round-robin 0"
    hardware_handler "1 rdac"
    failback          immediate
    features          "2 pg_init_retries 50"
    no_path_retry     30
    }
}
```

The following table explains the attributes and values in the devices section of the
`/etc/multipath.conf` file.

Table 1 Attributes and Values in the multipath.conf File

Attribute	Parameter Value	Description
path_grouping_poli cy	group_by_prio	The path grouping policy to be applied to this specific vendor and product storage.
prio	rdac	The program and arguments to determine the path priority routine. The specified routine should return a numeric value specifying the relative priority of this path. Higher numbers have a higher priority.
getuid_callout	"/lib/udev/scsi_i d -g -u -d /dev/%n"	The program and arguments to call out to obtain a unique path identifier.
polling_interval	5	The interval between two path checks, in seconds.
path_checker	rdac	The method used to determine the state of the path.
path_selector	"round-robin 0"	The path selector algorithm to use when there is more than one path in a path group.
hardware_handler	"1 rdac"	The hardware handler to use for handling device-specific knowledge.
failback	immediate	A parameter to tell the daemon how to manage path group failback. In this example, the parameter is set to 10 seconds, so failback occurs 10 seconds after a device comes online. To disable the failback, set this parameter to manual. Set it to immediate to force failback to occur immediately. When clustering or shared LUN environments are used, set this parameter to manual.
features	"2 pg_init_retries 50"	Features to be enabled. This parameter sets the kernel parameter pg_init_retries to 50. The pg_init_retries parameter is used to retry the mode select commands.
no_path_retry	30	Specify the number of retries before queuing is disabled. Set this parameter to fail for immediate failure (no queuing). When this parameter is set to queue, queuing continues indefinitely.
retain_attached_hw _handler	Yes	DM-MP will use the currently attached hardware handler instead of trying to attach the one specified during table load. If no hardware handler is attached the specified hardware handler will be used.

Attribute	Parameter Value	Description
rr_min_io	100	The number of I/Os to route to a path before switching to the next path in the same path group, using bio-based device-mapper-multipath. This is only for systems running kernels older that 2.6.31. Smaller value for rr_min_io would give better performance for large I/O block size.
rr_min_io_rq	1	The number of I/O to route to a path before switching to the next in the same path group, using request-based device-mapper-multipath. This setting should be used on systems running kernels 2.6.31 or older, use rr_min_io. The default value is 1.
rr_weight	uniform	If set to priorities the multipath configurator will assign path weights as "path prio * rr_min_io". Possible values are priorities or uniform.

Setting Up DM-MP for Large I/O Blocks

When a single I/O operation request a block larger than 512 KB, this is considered to be a large block. You must tune certain parameters for a device that uses Device Mapper Multipath (DMMP) in order for the device to perform correctly with large I/O blocks. The following parameters affect performance with large I/O blocks:

- max_sectors_kb (RW) - This parameter sets the maximum number of kilobytes that the block layer will allow for a file system request. The value of this parameter must be less than or equal to the maximum size allowed by the hardware.

- max_segments (RO) - This parameter enables low level driver to set an upper limit on the number of hardware data segments in a request.

1. Set the value of the max_segments parameter for the respective HBA driver as load a time module parameter.

   The following table lists HBA drivers which provide module parameters to set the value for max_segments.

HBA	Module Parameter
LSI SAS mpt2sas	max_sgl_entries
Emulex	lpfc_sg_seg_cnt
Infiniband	srp_sg_tablesize

2. On the command line, enter the command echo *N* >/sys/block/*sd device name*/queue/max_sectors_kb to set the value for the max_sectors_kb parameter for all physical paths for dm device in sysfs. In the command, *N* is an unsigned number less than the max_hw_sectors_kb value for the device; *sd device name* is the name of the sd device.

3. On the command line, enter the command echo *N* >/sys/block/*dm device name*/queue/max_sectors_kb to set the value for the max_sectors_kb parameter for all dm device in sysfs. In the command, *N* is an unsigned number less than the max_hw_sectors_kb value for the device; *dm device name* is the name of the dm device represented by dm-X.

Using the Device Mapper Devices

Multipath devices are created under /dev/ directory with the prefix dm-. These devices are the same as any other block devices on the host. To list all of the multipath devices, run the multipath -ll command. The following example shows system output from the multipath -ll command for one of the multipath devices.

```
mpathp (3600a0b80005ab177000017544a8d6b92) dm-0
LSI,INF-01-00
[size=5.0G][features=3 queue_if_no_path p
g_init_retries 50][hwhandler=1 rdac][rw]
\_ round-robin 0 [prio=6][active]
\_ 5:0:0:0 sdc 8:32 [active][ready]
\_round-robin 0 [prio=1][enabled]
\_ 4:0:0:0 sdb 8:16 [active][ghost]
```

In this example, the multipath device nodes for this device are /dev/mapper/mpathp and /dev/dm-0. The following table lists some basic options and parameters for the multipath command.

Table 2 Options and Parameters for the multipath Command

Command	Description
multipath -h	Prints usage information
multipath -ll	Shows the current multipath topology from all available information, such as the sysfs, the device mapper, and path checkers
multipath -f map	Flushes the multipath device map specified by the map option, if the map is unused
multipath -F	Flushes all unused multipath device maps

Trouble-shooting the Device Mapper

Table 3 Troubleshooting the Device Mapper

Situation	Resolution
Is the multipath daemon, multipathd, running?	At the command prompt, enter the command: `/etc/init.d/multipathd status.`
Why are no devices listed when you run the `multipath -ll` command?	At the command prompt, enter the command: `#cat /proc/scsi/scsi.` The system output displays all of the devices that are already discovered. Verify that the `multipath.conf` file has been updated with proper settings.

Failover Drivers for the Solaris Operating System

5

MPxIO is the supported failover driver for the Solaris operating system.

Solaris OS Restrictions

SANtricity ES Storage Manager no longer supports or includes RDAC for these Solaris OS:

- Solaris 11
- Solaris 10 OS
- Solaris 9 OS
- Solaris 8 OS

> **NOTE** MPxIO is not included on the SANtricity ES Storage Manager Installation DVD.

Prerequisites for Installing MPxIO on the Solaris OS for the First Time

Perform these prerequisite tasks.

1. Install the hardware.
2. Map the volumes.
3. Make sure that RDAC is not on the system, because you cannot run both RDAC and MPxIO.

NOTE RDAC and MPxIO cannot run on the same system.

Prerequisites for Installing MPxIO on a Solaris OS that Previously Ran RDAC

Perform these prerequisite tasks.

1. Make sure that there are no major problems in the current RDAC.

ATTENTION Potential loss of data access – Some activities, such as adding and removing storage arrays, can lead to stale information in the RDAC module name file. These activities might also render data temporarily inaccessible.

2. To make sure that no leftover RDAC files exist, type this command, and press **Enter**:

       ls -l /var/symsm/directory

3. To make sure that the RDAC directory does not exist, type this command, and press **Enter**:

       ls -l /dev/rdsk/*s2 >>filename

4. Examine the `/etc/symsm/mnf` file.

    Each currently connected storage array should be on one line. An example line is:

    ```
    infiniti23/24~1T01610104~ 0 1 7~1T04110240~
    7~0~3~~c6t3d0~c4t2d7~
    ```

5. Make sure that there are no extra lines for disconnected storage arrays.
6. Make sure that two controllers are listed on each line. (The example shows c6t3 and c4t2.)
7. Make sure that these controllers are the correct controllers.
8. Make sure there are no identical cXtX combinations. For example, if you see c6t3d0 and c6t3d4, a problem exists.
9. Make sure that no major problems exist in the Solaris OS.

    When you reset or power-cycle a controller, it might take up to three minutes to become fully ready. Older storage arrays might take longer. The default Solaris Not Ready retry timer is only 30 seconds long, so spurious controller failovers might result. Sun Microsystems has already increased the timer for NetApp-branded storage arrays to two minutes.

10. Make sure that the VERITAS Volume Manager can handle the MPxIO upgrade.
11. Capture the current configuration by performing these steps:
    a. Save this file.

        `/etc/symsm/mnf`
    b. Save this file.

        `/etc/raid/rdac_address`
    c. Type this command, and press **Enter**:

        ```
        ls -l /dev/rdsk/*s2 >>rdsk.save
        ```
    d. Type this command, and press **Enter**:

        ```
        ls -l /dev/symsm/dev/rdsk/*s2 >>symsm.save
        ```
    e. Type this command, and press **Enter**:

        ```
        lad -y >>lad.save
        ```
12. To remove RDAC, type this command, and press **Enter**:

    ```
    pkgrm RDAC
    ```

NOTE Do not restart the system at this time.

Installing MPxIO on the Solaris 9 OS

MPxIO is not included in the Solaris 9 OS. To install MPxIO on the Solaris 9 OS, perform these steps.

1. Download and install the *SAN 4.4x release Software/Firmware Upgrades and Documentation* from this website:

   ```
   http://www.sun.com/storage/san/
   ```

2. Install recommended patches.

3. To enable MPxIO on the Solaris 9 OS, perform these steps:

   a. Open this file:

   ```
   /kernel/drv/scsi_vhci.conf
   ```

   b. Change the last line in the script to this command:

   ```
   mpxio-disable="no";
   ```

NOTE Make sure that `"no"` is enclosed in double quotation marks.

---

4. Disable MPxIO on any Fibre Channel drives, such as internal drives, that should not be MPxIO enabled. To disable MPxIO on specific drives, perform these steps:

   a. Open the Fibre Channel port driver configuration file:

   ```
   /kernel/drv/qlc.conf
   ```

   b. Add a line similar to the following:

   ```
   name="qlc" parent="/pci@8,600000" port=0
   unit-address="2"
   mpxio-disable="yes";
   ```

NOTE To find the correct parent and port numbers, look at the device entry for the internal drives, found in `/dev/dsk`.

---

5. To update vfstab and the dump configuration, type this command:

   ```
   stmsboot -u
   ```

6. Reboot the system.

Enabling MPxIO on the Solaris 10 OS

MPxIO is included in the Solaris 10 OS. Therefore, MPxIO does not need to be installed; It only needs to be enabled.

NOTE MPxIO for iSCSI is enabled by default.

To disable MPxIO on specific drives, add a line similar to the following line to the `/kernel/drv/fp.conf` Fibre Channel port driver configuration file:

```
name="fp" parent="/pci@8,600000/SUNW,qlc@2" port=0
mpxio-disable="yes";
```

1. To enable MPxIO, run one of these commands:
   — To enable FC drives, run the `stmsboot -D fp -e`command.
   — To enable 3-GB SAS drives, run the `stmsboot -D mpt -e` command.
   — To enable 6-GB SAS drives, run the `stmsboot -D mpt sas -e` command.

   To find the correct parent and port numbers, look at the device entry for the internal drives, found in the `/dev/dsk` path.

2. Run the `stmsboot -u` command to update vfstab and the dump configuration.

3. Reboot the system.

Enabling MPxIO on the Solaris 11 OS

MPxIO is included in the Solaris 11 OS. Therefore, MPxIO does not need to be installed. It only needs to be enabled.

NOTE MPxIO for the x86 architecture is, by default, enabled for the Fibre Channel (FC) protocol.

1. To enable MPxIO for FC drives, run the following command: `stmsboot -D fp -e`

   To disable MPxIO on specific drives, add a line similar to the following line to the `/kernel/drv/fp.conf` Fibre Channel port driver configuration file:

   ```
   name="fp" parent="/pci@8,600000/SUNW,qlc@2"
   port=0
   mpxio-disable="yes";
   ```

   To find the correct parent and port numbers, look at the device entry for the internal drives, found in the `/dev/dsk` path.

2. Run the `stmsboot -u` command to update vfstab and the dump configuration.

3. Reboot the system.

Configuring Failover Drivers for the Solaris OS

Use the default settings for all Solaris OS configurations.

Frequently Asked Questions about Solaris Failover Drivers

Table 1 Frequently Asked Questions about Solaris Failover Drivers

Question	Answer

Where can I find MPxIO-related files?	You can find MPxIO-related files in these directories:
	`/etc/`
	`/kernel/drv`

Where can I find data files?	You can find data files in these directories:
	`/etc/raid ==> /usr/lib/symsm`
	`/var/symsm`
	`/var/opt/SM`

Where can I find the command line interface (CLI) files?	You can file CLI files in this directory: `/usr/sbin`

Where can I find the bin files?	You can find the bin files in these directories:
	`etc/raid/bin ==> /usr/lib/symsm/bin`
	`/usr/sbin/symsm ==> /usr/lib/symsm/bin`

Where can I find device files?	You can find device files in these directories:
	`/dev/rdsk`
	`/dev/dsk`

Where can I find the SANtricity ES Storage Manager files?	You can find the SANtricity ES Storage Manager files in these directories:
	`/opt/SMgr`
	`/opt/StorageManager`

Question	Answer
How can I get a list of controllers and their volumes?	Use the `lad -y` command. The command uses LUN 0 to get the information and is located in the `/usr/lib/symsm/bin` directory. It can be reached through `/etc/raid/bin`. This command updates the `mnf` file.
Where can I get a list of storage arrays, their volumes, LUNs, WWPNs, preferred paths, and owning controller?	Use the SMdevices utility, which is located in the `/usr/bin` directory. You can run the `SMdevices` utility from any command prompt.
How can I see whether volumes have been added?	Use the devfsadm utility to scan the system. Then run the `mpathad list lu` command to list all volumes and their paths. If you still cannot see any new volumes, either reboot the host and run the `mpathad list lu` command again, or use the SMdevices utility.
What file holds the storage array identification information?	Go to the `/etc/[osa|symsm]/mnf` directory. The `mnf file` identifies storage arrays in these ways: ■ Lists their ASCII names ■ Shows their controller serial numbers ■ Indicates the current LUN distribution ■ Lists controller system names ■ Lists the storage array numbers

Question	Answer
Why might the rdriver fail to attach and what can I do about it?	The rdriver might not attach when no entry exists in the `rdriver.conf` file to match the device, or whether rdnexus runs out of buses.
	If no physical devnode exists, then the following actions must occur:
	■ The `sd.conf` file must specify LUNs explicitly.
	■ The `ssd.conf` file with the itmpt HBA must specify LUNs explicitly.
	■ With Sun driver stacks, underlying HBA drivers dynamically create devnodes for ssd.
	You must restart the system with the reconfigure option after you update the `rdriver.conf` file.
	In the Solaris 9 OS, you can use the `update_drv` command whether the HBA driver supports it.
How can I determine whether the resolution daemon is working?	Type this command, and press **Enter**:
	`ps -ef	grep rd`
How do I find which failover module manages a volume in Solaris 11?	Check the host log messages for the volume. Storage arrays with Asymmetric Logical Unit Access (ALUA) are managed by the `lsi_tpgs` module. Storage arrays with earlier version of firmware are managed by the `lsi_sysm` module.
	As an alternative, list this information by entering the command `mpathadm show lu "/dev/rdsk/c0t0d0s2`.

Installing ALUA Support for VMware Versions ESX4.1U3, ESXi5.0U1, and Subsequent Versions

6

Starting with ESXi5.0 U1 and ESX4.1U3, VMware will automatically have the claim rules to select the VMW_SATP_ALUA plug-in to manage storage arrays that have the target port group support (TPGS) bit enabled. All arrays with TPGS bit disabled are still managed by the VMW_SATP_LSI plug-in.

1. Make sure that the host software on the management station is upgraded to version 10.86.

2. Upgrade the controllers in the storage array to controller firmware version 7.86 and the corresponding NVSRAM version.

3. From host management client, verify that the host OS type is set to *VMWARE*. Starting with storage management software version 10.84, the *VMWARE* host type will have the ALUA and TPGS bits enabled by default.

4. Use one of the following command sequences to verify that the TPGS/ALUA enabled devices are claimed by the VMW_SATP_ALUA plug-in.

   — For ESX4.1, enter the command `#esxcli nmp device list` on the command line of the host. Check that the output shows `VMW_SATP_ALUA` as the value of `Storage Array Type` for every storage array whose host software level is 10.83 or higher. Storage arrays with lower level host software show `VMW_SATP_LSI` as the value of `Storage Array Type`.

   — For ESXi5.0, enter the command `#esxcli storage nmp device list` on the command line of the host. Check that the output shows `VMW_SATP_ALUA` as the value of `Storage Array Type` for every storage array whose host software level is 10.83 or higher. Storage arrays with lower level host software show `VMW_SATP_LSI` as the value of `Storage Array Type`.

210- 05840_A0