

Linux® Host Utilities 5.2

Release Notes

NetApp, Inc.
495 East Java Drive
Sunnyvale, CA 94089 USA
Telephone: +1 (408) 822-6000
Fax: +1 (408) 822-4501
Support telephone: +1 (888) 4-NETAPP
Documentation comments: doccomments@netapp.com
Information Web: <http://www.netapp.com>

Part number: 215-05187_A0
March 2010

Copyright and trademark information

Copyright information

Copyright © 1994-2010 NetApp, Inc. All rights reserved. Printed in the U.S.A.

Software derived from copyrighted NetApp material is subject to the following license and disclaimer:

THIS SOFTWARE IS PROVIDED BY NETAPP “AS IS” AND WITHOUT ANY EXPRESS OR IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE, WHICH ARE HEREBY DISCLAIMED. IN NO EVENT SHALL NETAPP BE LIABLE FOR ANY DIRECT, INDIRECT, INCIDENTAL, SPECIAL, EXEMPLARY, OR CONSEQUENTIAL DAMAGES (INCLUDING, BUT NOT LIMITED TO, PROCUREMENT OF SUBSTITUTE GOODS OR SERVICES; LOSS OF USE, DATA, OR PROFITS; OR BUSINESS INTERRUPTION) HOWEVER CAUSED AND ON ANY THEORY OF LIABILITY, WHETHER IN CONTRACT, STRICT LIABILITY, OR TORT (INCLUDING NEGLIGENCE OR OTHERWISE) ARISING IN ANY WAY OUT OF THE USE OF THIS SOFTWARE, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGE.

NetApp reserves the right to change any products described herein at any time, and without notice. NetApp assumes no responsibility or liability arising from the use of products described herein, except as expressly agreed to in writing by NetApp. The use or purchase of this product does not convey a license under any patent rights, trademark rights, or any other intellectual property rights of NetApp.

The product described in this manual may be protected by one or more U.S.A. patents, foreign patents, or pending applications.

RESTRICTED RIGHTS LEGEND: Use, duplication, or disclosure by the government is subject to restrictions as set forth in subparagraph (c)(1)(ii) of the Rights in Technical Data and Computer Software clause at DFARS 252.277-7103 (October 1988) and FAR 52-227-19 (June 1987).

Trademark information

NetApp, the Network Appliance logo, the bolt design, NetApp—the Network Appliance Company, Cryptainer, Cryptoshred, DataFabric, DataFort, Data ONTAP, Decru, FAServer, FilerView, FlexClone, FlexVol, Manage ONTAP, MultiStore, NearStore, NetCache, NOW NetApp on the Web, SANscreen, SecureShare, SnapDrive, SnapLock, SnapManager, SnapMirror, SnapMover, SnapRestore, SnapValidator, SnapVault, Spinnaker Networks, SpinCluster, SpinFS, SpinHA, SpinMove, SpinServer, StoreVault, SyncMirror, Topio, VFM, VFM (Virtual File Manager), and WAFL are registered trademarks of NetApp, Inc. in the U.S.A. and/or other countries. gFiler, Network Appliance, SnapCopy, Snapshot, and The evolution of storage are trademarks of NetApp, Inc. in the U.S.A. and/or other countries and registered trademarks in some other countries. The NetApp arch logo; the StoreVault logo; ApplianceWatch; BareMetal; Camera-to-Viewer; ComplianceClock; ComplianceJournal; ContentDirector; ContentFabric; Data Motion; EdgeFiler; FlexShare; FPolicy; Go Further, Faster; HyperSAN; InfoFabric; Lifetime Key Management; LockVault; NOW; ONTAPI; OpenKey, RAID-DP; ReplicatorX; RoboCache; RoboFiler; SecureAdmin; SecureView; Serving Data by Design; Shadow Tape; SharedStorage; Simplicore; Simulate ONTAP; Smart SAN; SnapCache; SnapDirector; SnapFiler; SnapMigrator; SnapSuite; SohoFiler; SpinMirror; SpinRestore; SpinShot; SpinStor; vFiler; VPolicy; and Web Filer are trademarks of NetApp, Inc. in the U.S.A. and other countries. NetApp Availability Assurance and NetApp ProTech Expert are service marks of NetApp, Inc. in the U.S.A.

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 <http://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

- Contents of the Host Utilities. 1
- Recent updates to the Release Notes 4
- Using FCoE. 6
- Known problems and limitations with SUSE Linux Enterprise Server
 hosts on FC 8
- Known problems and limitations with Red Hat Enterprise Linux
 hosts on FC 16
- Known problems and limitations with SUSE Linux Enterprise Server
 hosts on iSCSI 24
- Known problems and limitations with Red Hat Enterprise Linux
 hosts on iSCSI 30
- Known problems and limitations with Veritas Storage Foundation 38
- Information about additional problems. 39
- Where to find additional information. 40

Contents of the Host Utilities

These Release Notes contain the latest information for the Linux® Host Utilities 5.2. The Release Notes include updates about known problems and limitations, information on tunable parameters, and any important cautions. The Release Notes also include any last-minute information about using the Host Utilities as well as any issues that might have been discovered since the Host Utilities were released.

Software

The Host Utilities include a set of software programs to help you use your Linux host with storage systems. You can download this software from the NOW (NetApp on the Web)® site after accepting the NetApp Customer Software License.

The software is distributed as a single file (**netapp_linux_host_utilities_5_2.tar.gz**), which you can install on your Linux host. The installation instructions are provided in the Host Utilities *Installation and Setup Guide* and on the NOW download page.

The same software distribution is shared by a number of Host Utilities. Therefore, it contains some software that is not used for Linux iSCSI. The following table lists each file in the distribution:

File	Purpose
install	Installs the Host Utilities software on a Linux host.
uninstall	Removes the Host Utilities software from a Linux host.
sanlun_bin_netapp.tar	Contains the following files: ◆ sanlun: Displays information about LUNs on connected storage systems. ◆ san_version: Displays the version of the Host Utilities ◆ man page for sanlun

Documentation

The Host Utilities include the following documentation.

Document	Purpose
Release Notes	Lists changes since the last release, the contents of the Host Utilities, and any known problems or last minute updates. Download from the NOW site.
Installation and Setup Guide	Explains how to install and configure the Host Utilities software to enable the Linux host to use LUNs on the storage system. Download from the NOW site.
Quick Command Reference	Lists the most common storage system and Linux commands used to manage both FC and iSCSI. Download from the NOW site.
Manual (man) pages	Displays information about the Host Utilities software. The sanlun.1 man page for the sanlun command is copied to the /opt/netapp/man/man1 directory when you install the Host Utilities software.
NOW pages	The Host Utilities installation instructions, information updates, and links to additional information. Browse on the NOW site.

Recent updates to the Release Notes

March 2010 updates

The Release Notes were updated in March 2010 with the release of the Linux Host Utilities 5.2. The Linux Host Utilities 5.2 release includes the following changes:

- ◆ SUSE Linux Enterprise Server 11 support for software iSCSI.
- ◆ Veritas Storage Foundation 5.1 support for the software iSCSI initiator.
- ◆ Starting from this release the following diagnostic scripts are not bundled along with the toolkit
 - ❖ `linux_info`: Collects configuration information about the Linux host, including Veritas stack information
 - ❖ `filer_info`: Collects information about the storage system (for Data ONTAP versions 7.1 and earlier)
 - ❖ `controller_info`: Collects information about the storage system
 - ❖ `qlogic_info`: Collects information about QLogic switches installed in the network
 - ❖ `brocade_info`: Collects information about Brocade switches installed in the network
 - ❖ `cisco_info`: Collects information about Cisco switches installed in the network
 - ❖ `mcdata_info`: Collects information about McData switches installed in the network
 - ❖ `SHsupport.pm`: Support library for all the above info scripts
 - ❖ `Telnet.pm`: Support library for all the above info scripts
 - ❖ `FTP.pm`: Support library for all the above info scripts
 - ❖ `NaElement.pm`: Support library for all the above info scripts
 - ❖ `NaErrno.pm`: Support library for all the above info scripts
 - ❖ `NaServer.pm`: Support library for all the above info scripts

December 2009 updates

The Release Notes were updated in December 2009 with the release of the Linux Host Utilities 5.1. The Linux Host Utilities 5.1 release includes the following changes:

- ◆ `controller_info` has been added as a diagnostic utility to collect information about the storage system. The Host Utilities continue to support the current storage system diagnostic utility `filer_info`.

- ◆ SUSE Linux Enterprise Server 11 support for Fibre Channel (FC) and Fibre Channel over Ethernet (FCoE).
- ◆ SUSE Linux Enterprise Server 10 SP3 support for FC, FCoE, and software iSCSI.

July 2009 updates (Host Utilities 5.0)

The Release Notes were updated in July 2009 to add the following information:

- ◆ Support for Red Hat Enterprise Linux 4 Update 8
The specific configurations that support Red Hat Enterprise Linux 4 Update 8 are listed in the NetApp® Interoperability Matrix at <http://now.netapp.com/NOW/products/interoperability>.

The required changes to the `/etc/multipath.conf` file on Red Hat Enterprise Linux 4 Update 8 are:

- ◆ `max_fds` parameter is now set to `max`.
- ◆ `rr_min_io` parameter is now added in device section, previously it was in default section.

April 2009 updates (Host Utilities 5.0)

The Release Notes were updated in April 2009 to add the following information:

- ◆ Support for both iSCSI and FC protocols in a single Host Utilities kit
- ◆ Support for hardware iSCSI host bus adapters (HBAs)
- ◆ Support for Red Hat Enterprise Linux 5 Update 3 for FC. The following parameters are added to the `/etc/multipath.conf` file:
 - ❖ `max_fds`
The `max_fds` parameter is now set to `max`, which sets the maximum file descriptors for the `multipathd` process. The value of this parameter must be set to a higher value than the number of paths to be configured.
 - ❖ `flush_on_last_del`
The `flush_on_last_del` parameter is set to `yes`. The `multipathd` process disables queueing when the last path to a device is deleted.
 - ❖ `queue_without_daemon`
The `queue_without_daemon` parameter is set to `no`. The `multipathd` process disables queueing when the `multipathd` is stopped.
- ◆ Support for Veritas Storage Foundation 5.0 Maintenance Pack 3 Rolling Patch 2 Hot Fix 1

Latest versions supported

New configuration components might be qualified between the Host Utilities releases. For information about the correct version and supported configuration, see the NetApp Interoperability Matrix at <http://now.netapp.com/NOW/products/interoperability>.

Supported packages

Use the version of the packages that are shipped along with the specific operating system release, unless explicitly mentioned.

Using FCoE

FCoE support

FCoE is a new model for connecting hosts to storage systems. FCoE is similar to traditional FC, as it maintains existing FC management and controls, but the hardware transport is a 10-gigabit Ethernet network.

Setting up an FCoE connection requires one or more supported converged network adapters (CNAs) in the host, connected to a supported data center bridging (DCB) Ethernet switch. The CNA is a consolidation point and effectively serves as both an HBA and an Ethernet adapter.

As an HBA, the presentation to the host is FC targets and all FC traffic is sent out as FC frames mapped into Ethernet packets (FC over Ethernet). The 10-gigabit Ethernet adapter is used for IP traffic, such as iSCSI, NFS, and HTTP. Both FCoE and IP communications through the CNA run over the same 10-gigabit Ethernet port, which connects to the DCB switch.

In general, you can configure and use FCoE connections just like traditional FC connections.

Installing and cabling the CNA

Install the CNA following the CNA manufacturer's instructions. The CNA must be directly cabled to a supported DCB switch. No intermediate Ethernet switches should be connected between the CNA end point and the DCB switch.

Obtaining CNA drivers and firmware

Updating the drivers and firmware for a CNA is just like updating them for a traditional FC HBA. For the supported firmware versions, check the NetApp Interoperability Matrix at <http://now.netapp.com/NOW/products/interoperability>.

The supported drivers and firmware are available directly from the CNA manufacturer:

- ◆ Emulex
<http://www.emulex.com/ts/index.html>
- ◆ QLogic
http://support.qlogic.com/support/drivers_software.asp

Updating the FCoE engine

For the supported versions of the FCoE engine or Menlo firmware, see the NetApp Interoperability Matrix at <http://now.netapp.com/NOW/products/interoperability>.

Also, see the CNA documentation on the respective QLogic and Emulex Web sites on how to update the FCoE engine or Menlo firmware.

Setting the CNA parameters

The CNA uses the same timeout parameters as a traditional FC HBA.

Configuring the DCB switch

The DCB switch requires special setup steps for FCoE. See the documentation supplied by the switch manufacturer. For example, the steps for a Cisco 5020 are included in the *Cisco Nexus 5000 Series Switch Fabric Manager Software Configuration Guide* available at <http://www.cisco.com/en/US/products/>.

You zone the DCB switch for FCoE just like you zone a traditional FC switch.

Known problems and limitations with SUSE Linux Enterprise Server hosts on FC

The qlinstall script fails, displaying an error message

Problem: On a Linux FC host that has the bundled QLogic driver loaded and in use, running the `uninstall` option of the SANsurfer scli `qlinstall` script fails, displaying an error message similar to the following:

```
ERROR:qla2xxx does not seem to be installed
```

Workaround: Before you install the recommended QLogic driver on a Linux FC host by using the SANsurfer scli package, complete the following steps.

1. Ensure that the QLogic driver is not in use.
2. Uninstall the QLogic driver by the following method:
 - a. Run the `uninstall` option of `qlinstall` for a SANsurfer package installed on the host.
 - b. Run the following command on the Linux host where the bundled drivers are loaded:

```
modprobe -r {qla2300 | qla2400}
```
3. Ensure that the QLogic driver is successfully unloaded.
4. Run the `qlinstall` script (with no options) of the SANsurfer scli package to install the requisite driver.

Commands hang when interconnect cable is pulled

Problem: When the NetApp controller interconnect cable is pulled, commands such as `multipath -ll` and `sanlun lun show` hang.

Workaround: Ensure that the interconnect cable is properly connected and working.

Problems when creating partitions on Linux DM-Multipath devices

Problem: In some versions of Linux, creating partitions on DM-Multipath devices might cause problems. The partitioning software might report an error after a new partition is created. This error message indicates that the new partition is not recognized. As a result, a new DM-Multipath device is not created automatically for the new partition.

Workaround: Complete the following steps.

Step	Action
1	Run the following command to create the new partition: fdisk
2	Run the following tool manually to read the partition table on the multipath device: kpartx
3	Create new multipath devices for the partitions.

This workaround does not disturb other multipath maps and is required only after creating a new partition.

For more information, see bug number 196448 on Bugs Online at <http://now.netapp.com/NOW/cgi-bin/bol>.

Configuring multipath devices on a Linux host fails

Problem: While configuring multipath devices on a Linux host, the DM-Multipath driver fails to create appropriate device maps if stale sd entries are present in the sysfs file. Configuring multipath devices on such a host fails. As a result, no DM-Multipath entries are displayed in the /dev/mapper/ directory or in the `multipath -l/-ll` output.

Workaround: Identify the stale sd entries and blacklist (exclude) them.

To identify and blacklist (exclude) the stale sd entries, use the following methods.

- ◆ Blacklist the sd entries in the sysfs (or /sys/block/) file that are not displayed by the `sanlun lun show` command.
- ◆ Blacklist the sd entries that display errors when the `scsi_id -gus` command is run.

Using the devnode method causes incorrect blacklisting on reboot

Problem: In the `multipath.conf` file for DM-Multipath, using the devnode method causes incorrect blacklisting. This is because the device that is registered as `sda` during one boot process might not remain as `sda` in the next boot process.

Workaround: Use the WWID method instead of the devnode method for blacklisting devices in the `multipath.conf` file.

LUNs not recognized by DM-Multipath

Problem: LUNs are mapped to the Linux host and are correctly reported by the `sanlun lun show all` command. However, the LUNs are not reported by the `multipath -ll` command. Multipath devices are not created for the LUNs.

Workaround: Ensure that the LUNs are not excluded (blacklisted) from the DM-Multipath configuration. Enter the following command on the Linux host to display the devices that are blacklisted:

```
multipath -v3 -d | grep blacklist
```

If any FC devices are blacklisted, check the `devnode_blacklist` or the `blacklist` section of the `/etc/multipath.conf` file. Ensure that the entries are correctly specified.

If the FC devices are not blacklisted but are still not recognized by the `multipath` command, regenerate the multipath maps by entering the following command:

```
multipath -v3
```

For more information, see bug number 228744 on Bugs Online at <http://now.netapp.com/NOW/cgi-bin/bol>.

File systems on partitions might have performance problems

Problem: The default parameters of the `fdisk` command that are used to create a partition on a LUN might result in poor performance for that LUN.

The default parameters of the `fdisk` command can create a misalignment between the LUN block boundaries and the underlying WAFL® boundaries. This causes the storage system to double its I/O workload for that LUN.

Workaround: When creating a partition, you can explicitly specify the `fdisk` parameters to use values that guarantee the correct block alignment.

For more information, see bug number 156121 on Bugs Online at <http://now.netapp.com/NOW/cgi-bin/bol>.

Unloading of the QLogic module fails

Problem: On the Linux host, unloading the QLogic module sometimes fails.

When Linux FC initiators are included in multiple initiator FC switch zones, the HBA module usage count on the specified initiator host always shows a higher value than expected (more than 1), even when there is no apparent usage of the HBA module.

For example, for a SUSE Linux Enterprise Server 9 SP3 host with its QLogic initiators connected to a multiple initiator zone on the FC switch, the QLogic module usage count is sometimes shown as follows:

```
qla2400 207360 0
qla2xxx 3654906 5 qla2400
```

Here, the qla2xxx usage count is shown as 5 (instead of 1), even though the QLogic module is not being used. As a result, the unloading of the QLogic module can fail.

Workaround: While connecting Linux FC hosts to a switch fabric, always ensure that each FC host initiator is connected to a separate individual initiator zone instead of a multiple initiator zone.

LUN resets observed during interconnect cable pull

Problem: When the interconnect cable that is connected to NetApp controllers running versions prior to Data ONTAP® 7.2.4 is pulled, LUNs resets occur on Linux hosts with Emulex adapters. This also leads to I/O stall on the host and the stalling continues till the interconnect cable is pulled back.

The Emulex initiator on a Linux host sends queries on the proxy paths, which fail during an interconnect cable pull, and this leads to the LUN resets. This is because Data ONTAP 7.2 does not respond to queries on the proxy paths during an interconnect cable pull.

Workaround: Set the following options on both the clustered controllers by running the following commands:

```
options lun.partner_unreachable.linux.behavior error
options lun.partner_unreachable.linux.hold_time 5
options lun.partner_unreachable.linux.skey 0x2
options lun.partner_unreachable.linux.asc 0x4
options lun.partner_unreachable.linux.ascq 0x1
options lun.partner_unreachable.linux.scsi_status 0x2
```

These options are applicable for all Data ONTAP versions from 7.2.0 to 7.2.3.

Linux host freezes after unloading QLogic HBA FC driver

Problem: The Linux host freezes after the QLogic HBA FC driver is unloaded, despite the DM-Multipath devices mounted on it.

Only the common `qla2xxx` module usage count is properly incremented or decremented (and not the `qla2300/qla2400` usage count) when you use the DM-Multipath utility on a Linux host. This can lead to severe repercussions.

When commands such as `umount /mnt/tmp`, `/etc/init.d/boot.multipath start`, and so on, are run on the host, the machine freezes, and this requires a hard boot.

Workaround: Always ensure that the `qla2xxx` module is not in use before you unload or reload the QLogic driver. That is, verify that the usage count of the `qla2xxx` module is 1 before you unload or reload the QLogic driver on the Linux host.

Error message from `sanlun fcp show adapter`

Problem: When you run `sanlun fcp show adapter` command on a Linux host with Emulex adapters, with the 3.2a5 Application Kit or the 3.2a11 Application Kit installed, you get error messages. The error messages that are displayed for the two application kits are different. This is due to a problem in the `hbaapi` libraries provided by both the kits.

In addition, a warning message is displayed when both these kits are installed on an `x86_64` Linux host.

This also causes `sanlun fcp show adapter` to fail.

Workaround: Instead of the 3.2a5 or the 3.2a11 kit, use the 3.2a14 kit on the Linux host. For a `x86_64` host, install the `32bit libnl.so` package first before installing the 3.2a14 kit to avoid the warning message. This also ensures that `sanlun fcp show adapter` functions properly on the `x86_64` host.

Linux FC hosts with QLogic initiators fail to detect LUNs

Problem: Linux FC hosts with QLogic initiators fail to detect LUNs mapped to them when configured to autonegotiate port topology and speed. On hosts that are using QLogic initiators, the host initiator fails to establish a consistent topology and speed, even after setting both to autonegotiate.

This might cause the host to miss a few LUNs that are mapped to it, during the HBA scan. This behavior persists even after rebooting the host.

Workaround: When using QLogic initiators on a Linux FC host, ensure that the following settings are set for port topology and speed by using the `SANsurfer scli` package as follows:

Step	Action	
1	Perform the following step:	
	If you are connecting to a NetApp controller operating in...	Then, set the QLogic HBA port topology to...
	SSI, Standby, or Partner CF modes	Point to Point Only
	Dual Fabric or Mixed CF modes	Loop Only
	a. Run the following command: <code>/opt/QLogic_Corporation/SANsurferCLI/sc1i</code> b. Select Configure HBA Settings from the Main Menu. c. For each of the worldwide port names (WWPNs) listed here, select ConnectionOptions and set it to Point to Point Only or Loop Only, depending on the mode.	
2	Set the QLogic port speed ideally to the highest speed possible, depending on its ceiling and that of the switch/target port it is connected to: a. Run the following command: <code>/opt/QLogic_Corporation/SANsurferCLI/sc1i</code> b. Select Configure HBA Settings from the Main Menu. c. For each of the WWPNs listed here, select the Data Rate option and set it to the specified speed.	

Host creates /dev/sd* if first LUN not mapped as LUN 0

Problem: If the first LUN mapped to a host has an ID other than 0, the host creates a /dev/sd* for the missing LUN 0. Adding LUN 0 later does not solve the problem.

Workaround: Always map the first LUN to a host as LUN 0. This is the default behavior of Data ONTAP when you do not explicitly specify the LUN identifier.

Map LUN 0 to the host and reload the HBA driver.

Using default partitions on Linux with Data ONTAP LUNs might require alignment for best performance

Problem: Using default partition settings on LUNs might result in performance problems. This problem can occur if the `sectors` value for a partition is not a multiple of 8.

Workaround: Manually enter the disk geometry parameters when creating partitions with the `fdisk` command.

For more information, see bug number 156121 on Bugs Online at <http://now.netapp.com/NOW/cgi-bin/bol>.

Zoning with standby or mixed cfmode

Problem: If the storage system is using standby or mixed cfmode, your switches must be zoned by WWPEN.

Workaround: To prevent failover problems, create zones that include the host HBAs and the primary port WWPENs on the storage system. Do not include the standby port WWPENs on the storage system in the zones with the host HBA ports.

The host should never access the standby port WWPENs. During failover, the standby ports of the takeover system assume the WWPENs of the primary ports of the system that is down. The primary ports are already included in existing zones, therefore the host can access them.

Delayed DM-Multipath failback on SUSE Linux Enterprise Server 10 and later hosts

Problem: On a SUSE Linux Enterprise Server 10 or SUSE Linux Enterprise Server 10 SP1 FC host, DM-Multipath failback is sometimes delayed by about 4 to 5 minutes after a controller giveback when using the `readsector0 path_checker` in the `/etc/multipath.conf` file. During this time, there is a complete I/O stall on the host before I/O resumes on the primary paths.

In certain cases after controller giveback, I/O continues on the secondary paths for a similar delayed period before resuming on the primary paths.

During this I/O stall, the `multipath -ll` command displays the primary paths status as `[failed] [ready]`. Also, there are no `multipathd` messages logged in the `syslog` during this time, despite the default polling interval of 5 seconds by the `multipathd` on the Linux host.

Workaround: Use one of the following two workarounds:

- ◆ Use `directio path_checker` instead of `readsector0` or `tur` in the `/etc/multipath.conf` file as follows:
path_checker directio
- ◆ During the delayed failback, update the primary path status by restarting the multipathd by running the following command:
`/etc/init.d/multipathd restart`

After this is done, I/O immediately resumes on the primary paths.

Known problems and limitations with Red Hat Enterprise Linux hosts on FC

Root device DM-Multipath support fails during operating system installation, if only single path to boot LUN is available

Problem: Root device DM-Multipath support fails during the installation of Red Hat Enterprise Linux 5 Update 1, if there is only a single physical path to the boot LUN. This happens despite specifying *mpath* on the kernel boot line during the operating system installation.

Even after enabling multiple physical paths to the boot LUN and then re-creating the initrd image, Red Hat Enterprise Linux 5 Update 1 boots from the SCSI device instead of the overlying DM-Multipath device.

The boot LUN shows up as a SCSI device instead of a DM-Multipath device.

Workaround: Ensure that multiple physical paths are available to the boot LUN during the installation of Red Hat Enterprise Linux 5 Update 1.

Root device DM-Multipath support fails if *user_friendly_names* is disabled

Problem: Root device DM-Multipath support fails on Red Hat Enterprise Linux 5 Update 1 if *user_friendly_names* is disabled. For root device DM-Multipath support, the *user_friendly_names* is set to *yes* by default during the installation of Red Hat Enterprise Linux 5 Update 1 using Anaconda.

As a result, the operating system fails to boot and displays an error message similar to the following:

```
Checking filesystems fsck.ext3: No such file or directory while
trying to open /dev/mapper/mpath0p1
```

The boot operation fails due to the `/dev/mapper/mpathp1` entry for `/boot` in `fstab`.

Workaround: To disable *user_friendly_names* on the Red Hat Enterprise Linux 5 Update 1 root on a multipath host, complete the following steps:

Step	Action
1	Set <i>user_friendly_names</i> to <i>no</i> or remove it from the <code>/etc/multipath.conf</code> file.
2	Change the <code>/boot</code> entry in <code>/etc/fstab</code> from the <code>/dev/mapper/mpath</code> name to the corresponding <code>/dev/mapper/<DevId></code> name.

Step	Action
3	Re-create the initrd by running the following command: mkinitrd

Problems when creating partitions on Linux DM-Multipath devices

Problem: In some versions of Linux, creating partitions on DM-Multipath devices might cause problems. The partitioning software might report an error after a new partition is created. This error message indicates that the new partition is not recognized. As a result, a new DM-Multipath device is not created automatically for the new partition.

Workaround: To resolve this problem, complete the following steps.

Step	Action
1	Run the following command to create the new partition: fdisk
2	Run the following tool manually to read the partition table on the multipath device: kpartx
3	Create new multipath devices for the partitions.

This workaround does not disturb other multipath maps and is required only after creating a new partition.

For more information, see bug number 196448 on Bugs Online at <http://now.netapp.com/NOW/cgi-bin/bol>.

Linux host freezes after unloading QLogic HBA FC driver

Problem: The Linux host freezes after the QLogic HBA FC driver is unloaded, despite the DM-Multipath devices mounted on it.

Only the common qla2xxx module usage count is properly incremented or decremented (and not the qla2300/qla2400 usage count) when you use the DM-Multipath utility on a Linux host. This can lead to severe repercussions.

When commands such as `umount /mnt/tmp`, `/etc/init.d/boot.multipath start`, and so on, are run on the host, the machine freezes which requires a hard boot.

Workaround: Always ensure that the `qla2xxx` module is not in use before you unload or reload the QLogic driver. That is, verify that the usage count of the `qla2xxx` module is 1 before you unload or reload the QLogic driver on the Linux host.

Delayed DM-Multipath failback on Red Hat Enterprise Linux 5.0 hosts

Problem: On a Red Hat Enterprise Linux 5 FC host, after a storage system giveback, the DM-Multipath failback sometimes gets delayed by about 4 to 5 minutes. During this time, there is a complete I/O stall on the host before I/O resumes on the primary paths.

In certain cases after the storage system giveback, I/O continues on the secondary paths for a similar delayed period before resuming on the primary paths.

During the I/O stall, the `multipath -ll` command displays the primary paths status as `[failed] [ready]`.

Workaround: During the delay, you can update the primary paths status by restarting the `multipathd` by running the following command:

```
/etc/init.d/multipathd restart
```

The I/O immediately resumes on the primary paths.

LUN resets observed on interconnect cable pull

Problem: When the interconnect cable that is connected to NetApp controllers running Data ONTAP 7.2.4 is pulled, LUNs resets occur on Linux hosts with Emulex adapters. This also leads to I/O stall on the host and stalling continues till the interconnect cable is plugged back.

The Emulex initiator on a Linux host sends queries on the proxy paths, which fail during an interconnect cable pull and result in the LUN resets. This is because Data ONTAP 7.2 does not respond to queries on the proxy paths during an interconnect cable pull.

Workaround: Set the following options on both the clustered controllers by running the following commands:

```
options lun.partner_unreachable.linux.behavior error
options lun.partner_unreachable.linux.hold_time 5
options lun.partner_unreachable.linux.skey 0x2
```

```
options lun.partner_unreachable.linux.asc 0x4
options lun.partner_unreachable.linux.ascq 0x1
options lun.partner_unreachable.linux.scsi_status 0x2
```

These options are applicable for all Data ONTAP versions from 7.2.0 to 7.2.3.

Commands hang when interconnect cable is pulled

Problem: When the NetApp controller interconnect cable is pulled, commands such as `multipath -ll` and `sanlun lun show` hang.

Workaround: Ensure that the interconnect cable is properly connected and working.

Error message from `sanlun fcp show adapter`

Problem: When you run `sanlun fcp show adapter` command on a Linux host with Emulex adapters, with the 3.2a5 Application Kit or the 3.2a11 Application Kit installed, you get error messages. The error messages that are displayed on the two application kits are different. This is due to the problem in the hbaapi libraries provided by both the kits.

In addition, a warning message is displayed when both these kits are installed on a x86_64 Linux host.

This also causes `sanlun fcp show adapter` to fail.

Workaround: Instead of the 3.2a5 or the 3.2a11 kit, use the 3.2a14 kit on the Linux host. For a x86_64 host, install the 32bit `libnl.so` package first before installing the 3.2a14 kit to avoid the warning message. This also ensures that `sanlun fcp show adapter` functions properly on the x86_64 host.

File systems on partitions might have performance problems

Problem: The default parameters of the `fdisk` command that are used to create a partition on a LUN might result in poor performance for that LUN.

The default parameters of the `fdisk` command can create a misalignment between the LUN block boundaries and the underlying WAFL boundaries. This causes the storage system to double its I/O workload for that LUN.

Workaround: When creating a partition, you can explicitly specify the `fdisk` parameters to use values that guarantee the correct block alignment.

For more information, see bug number 156121 on Bugs Online at <http://now.netapp.com/NOW/cgi-bin/bol>.

Unloading of the QLogic module fails

Problem: On the Linux host, unloading the QLogic module sometimes fails.

When Linux FC initiators are included in multiple initiator FC switch zones, the HBA module usage count on the specified initiator host always shows a higher value than expected (more than 1), even when there is no apparent usage of the HBA module.

For example, for a Red Hat Enterprise Linux host with its QLogic initiators connected to a multiple initiator zone on the FC switch, the QLogic module usage count is sometimes shown as follows:

```
qla2400 207360 0
qla2xxx 3654906 5 qla2400
```

Here, the `qla2xxx` usage count is shown as 5 (instead of 1), even though the QLogic module is not being used. This can cause unloading of the QLogic module to fail.

Workaround: While connecting Linux FC hosts to a switch fabric, always ensure that each FC host initiator is connected to a separate individual initiator zone instead of a multiple initiator zone.

Automatic mounting of DM-Multipath devices fails

Problem: Automatic mounting of DM-Multipath devices by using `/dev/mpath/` entries in `fstab` fails on Red Hat Enterprise Linux 5 hosts.

If `user_friendly_names` is set to `yes` in the `multipath.conf` file on a Red Hat Enterprise Linux 5 host, it causes problems during the boot process. No `/dev/mpath` entries are available during the boot process, which causes the automatic mounting to fail.

Workaround: Use the `/dev/mapper/` entries for automatically mounting DM-Multipath devices on a Red Hat Enterprise Linux host instead of the `/dev/mpath/` entries.

The qlinstall script fails, displaying an error message

Problem: On a Linux FC host that has the bundled QLogic driver loaded and in use, running the `uninstall` option of the SANsurfer scli `qlinstall` script fails, displaying an error message similar to the following:

```
ERROR:qla2xxx does not seem to be installed
```

Workaround: Before you install the recommended QLogic driver on a Linux FC host using the SANsurfer scli package, complete the following steps.

1. Ensure that the QLogic driver is not in use.
2. Uninstall the QLogic driver by the following method:
 - a. Run the `uninstall` option of `qlinstall` for a SANsurfer package installed on the host.
 - b. Run the following command on the Linux host where the bundled drivers are loaded:


```
modprobe -r {qla2300|qla2400}
```
3. Ensure that the QLogic driver is successfully unloaded. Run the `qlinstall` script (with no options) of the SANsurfer `scli` package to install the requisite driver.

Host creates /dev/sd* if first LUN not mapped as LUN 0

Problem: If the first LUN mapped to a host has an ID other than 0, the host creates a `/dev/sd*` for the missing LUN 0. Adding LUN 0 later does not solve the problem.

Workaround: Always map the first LUN to a host as LUN 0. This is the default behavior of Data ONTAP when you do not explicitly specify the LUN identifier.

Map LUN 0 to the host and reload the HBA driver.

Configuring multipath devices on a Linux host fails

Problem: While configuring multipath devices on a Linux host, the DM-Multipath driver fails to create appropriate device maps if stale `sd` entries are present in the `sysfs` file. Configuring multipath devices on such a host fails. As a result, no DM-Multipath entries are displayed in the `/dev/mapper/` directory or in the `multipath -l/-ll` output.

Workaround: Identify the stale `sd` entries and blacklist (exclude) them.

To identify the stale `sd` entries for blacklisting, use the following methods.

- ◆ Blacklist those `sd` entries in the `sysfs` (or `/sys/block/`) file that are not displayed by the `sanlun lun show` command.
- ◆ Blacklist the `sd` entries that display errors when the `scsi_id -gus` command is run.

Using the devnode method causes incorrect blacklisting on reboot

Problem: In the `multipath.conf` file for DM-Multipath, using the devnode method causes incorrect blacklisting. This is because the device that is registered as `sda` during one boot process might not remain as `sda` in the next boot process.

Workaround: Use the WWID method rather than the devnode method for blacklisting devices in the `multipath.conf` file.

Using default partitions on Linux with Data ONTAP LUNs might require alignment for best performance

Problem: Using default partition settings on LUNs might result in performance problems. This problem can occur if the `sectors` value for a partition is not a multiple of 8.

Workaround: Manually enter the disk geometry parameters when creating partitions with the `fdisk` command.

For more information, see bug number 156121 on Bugs Online at <http://now.netapp.com/NOW/cgi-bin/bol>.

Zoning with standby or mixed cfmodes

Problem: If the storage system is using standby or mixed cfmodes, your switches must be zoned by WWPN.

Workaround: To prevent failover problems, create zones that include the host HBAs and the primary port WWPNs on the storage system. Do not include the standby port WWPNs on the storage system in the zones with the host HBA ports.

The host should never access the standby port WWPNs. During failover, the standby ports of the takeover system assume that the WWPNs of the primary ports of the system that is down. The primary ports are already included in existing zones, therefore the host can access them.

LUNs not recognized by DM-Multipath

Problem: LUNs are mapped to the Linux host and are correctly reported by the `sanlun lun show all` command. However, the LUNs are not reported by the `multipath -ll` command. Multipath devices are not created for the LUNs.

Workaround: Ensure that the LUNs are not excluded (blacklisted) from the DM-Multipath configuration. Enter the following command on the Linux host to display the devices that are blacklisted:

```
multipath -v -d | grep blacklist
```

If any FC devices are blacklisted, check the `devnode_blacklist` or `blacklist` section of the `/etc/multipath.conf` file. Ensure that the entries are correctly specified.

If the FC devices are not blacklisted, but are still not recognized by the multipath command, regenerate the multipath maps by entering the following command:

```
multipath -v3
```

For more information, see bug number 228744 on Bugs Online at <http://now.netapp.com/NOW/cgi-bin/bol>.

Delayed DM-Multipath failback on Red Hat Enterprise Linux 5 Update 1 and later

Problem: On Red Hat Enterprise Linux 5 FC hosts, DM-Multipath failback is sometimes delayed by about 4 to 5 minutes after a controller giveback, when *readsector0* or *tur* is used as the path checker in the `multipath.conf` file. During this delay, there is a complete I/O stall on the host before I/O resumes on the primary paths.

In certain cases after controller giveback, I/O continues on the secondary paths for a similar delayed period before resuming on the primary paths.

During this I/O stall, the `multipath -ll` command displays the primary paths status as `[failed] [ready]`. Also, there are no `multipathd` messages logged in the `syslog` during this time, despite the default polling interval of 5 seconds by the `multipathd` on the Linux host.

Workaround:

- ◆ On Red Hat Enterprise Linux 5.0: During this delay, you can update the primary path's status can be updated by restarting the `multipathd` by running the following command:

```
/etc/init.d/multipathd restart
```

- ◆ On Red Hat Enterprise Linux 5 Update 1 and later: Use `directio` as the path checker in the `multipath.conf` file instead of *readsector0* or *tur* as shown below:

```
path_checker directio
```

This ensures that there is no delay during failback.

Known problems and limitations with SUSE Linux Enterprise Server hosts on iSCSI

OCFS2 self-fencing during cluster failover

Problem: During a clustered failover or network disruption, nodes fence themselves by panicking.

Oracle Cluster File System (OCFS2) has a fencing mechanism that relies on each node being able to write a heartbeat value to a specific area on each OCFS2 volume. Every two seconds, each node writes a new heartbeat value and reads the value of the other node. If a node is unable to write or read the heartbeat value, it retries the operation a specific number of times, and then self-fences by panicking the Linux kernel as it fails to obtain quorum.

During a clustered failover or network disruption, all paths to the storage system can become unavailable, and none of the nodes in the OCFS2 cluster is able to write or read the heartbeat values. As a result, the nodes might fence by panicking.

Workaround: The workaround for this problem is to set the `O2CB_HEARTBEAT_THRESHOLD` value in `/etc/sysconfig/o2cb` to a value large enough to ensure that the nodes do not self-fence.

With a dedicated TCP/IP network for the iSCSI link between Linux hosts and the storage system, a reasonable value usually is 181. This value corresponds to a total of 360 seconds before a node self-fences.

Depending on the load in the iSCSI network, you should raise the value. It is best to use a value in the range of 61 to 201, which is ideal for a configuration with OCFS2 on multipath and iSCSI devices.

For more information, see the OCFS FAQs at http://oss.oracle.com/projects/ocfs2/dist/documentation/ocfs2_faq.html#HEARTBEAT.

Note

OCFS2 is not supported in SUSE Linux Enterprise Server 10 SP2.

LUN status undefined

Problem: `multipath -l` displays the LUN status as `undef`.

Workaround: Use `multipath -ll` to get the correct information.

LUNs not recognized by DM-Multipath

Problem: LUNs are mapped to the Linux host and are correctly reported by the `iscsi-ls -l` and `sanlun lun show all` commands. However, the LUNs are not reported by the `multipath -ll` command. Multipath devices are not created for the LUNs.

Workaround: Ensure that the LUNs are not excluded (blacklisted) from the DM-Multipath configuration. Enter the following command on the Linux host to display the devices that are blacklisted:

```
multipath -v3 -d | grep blacklist
```

If any of the iSCSI devices are blacklisted check the `blacklist` section of the `/etc/multipath.conf` file. Ensure that the entries are correctly specified.

If the iSCSI devices are not blacklisted but are still not recognized by the `multipath` command, regenerate the multipath maps by entering the following command:

```
multipath -v3
```

File systems on partitions might have performance problems

Problem: The default `fdisk` command parameters used to create a partition on a LUN might result in poor performance for that LUN.

The default `fdisk` parameters can create a misalignment between the LUN block boundaries and the underlying WAFL boundaries. This misalignment causes the storage system to double its I/O workload for that LUN.

Workaround: When creating partitions, you can explicitly specify the `fdisk` parameters to use values that guarantee the correct block alignment.

For specific steps, see bug number 156121 on Bugs Online at <http://now.netapp.com/NOW/cgi-bin/bol>.

Online resize of file system is not supported

Problem: Currently, a resize of the LUN on the NetApp storage controller is reflected on the host when you rescan the iSCSI sessions. The rescan reflects the new size of the LUN. However, the resize of the file system is not reflected while it is online.

See the Novell bugzilla number 386325 for further details.

Workaround: To reflect the new size of the LUN on the file system, unmount the file system and use the `resize` utility to extend or shrink the file system.

Automatically mounting and automatically unmounting the file system on multipath devices do not work on SUSE Linux Enterprise Server 9 SP3

Problem: Automatic mounting of multipath-based file systems at boot time using entries in `/etc/fstab.conf` or `/etc/fstab.iscsi` does not work. This is because the placement of the multipath boot and shutdown scripts in the rc script list in SUSE Linux Enterprise Server 9 SP3 is out of order with respect to the mount and unmount scripts for iSCSI and non-iSCSI devices.

The `multipathd` and `iscsid` start-up scripts are at the same serial number for a run level in `/etc/init.d/rc3.d`. Because alphabetically *i* comes before *m*, `iscsid` starts before `multipathd`.

As a result, the multipath boot and shutdown sequence is not synchronized with the iSCSI boot and shutdown sequence. Multipath devices are not available when the mount command is run, and the mount command fails to mount these devices.

Similarly in the shutdown path, by the time the unmount script runs, the multipath utility would have already been shut down.

Workaround: Manual mounting and unmounting is required for multipath-based file systems on SUSE Linux Enterprise Server 9 SP3. After the multipath devices are available on booting, the file systems can be mounted.

For unmounting, ensure that any multipath-based file systems are unmounted before the multipath utility is shut down.

To ensure the proper shutdown of multipath-based iSCSI file systems, complete the following steps.

Step	Action
1	Unmount all file systems on iSCSI LUNs.
2	Flush the multipath maps by entering the following command: <code>multipath -F</code>
3	Stop <code>multipathd</code> by entering the following command: <code>/etc/init.d/multipathd stop</code>
4	Stop the iSCSI service by entering the following command: <code>/etc/init.d/iscsi stop</code>

For more information, see bug number 196997 on Bugs Online at <http://now.netapp.com/NOW/cgi-bin/bol>.

iscsi-ls command displays error message on SUSE Linux Enterprise Server 9 SP3

Problem: For some SUSE Linux versions, the `iscsi-ls -l` command displays an error message similar to the following for each LUN:

```
Cannot open /etc/sysconfig/devlabel: No such file or directory
```

This message is displayed because of a bug in the command. The message does not indicate an error and can be safely ignored.

One or more ext3 file systems not mounted after reboot

Problem: After a reboot of the host, one or more ext3 file systems are not mounted, and an error message similar to the following is displayed:

```
fs type ext3 not supported by kernel
```

The module that supports ext3 file systems was not loaded before the system tried to mount the file systems.

Workaround: Run the `iscsi-mountall` command to mount the file systems.

sanlun does not display all iSCSI LUNs

Problem: The `sanlun` command does not display all iSCSI LUNs mapped to the initiator.

There are two possible explanations for `sanlun` not displaying all iSCSI LUNs:

◆ Number of SCSI devices

The `sanlun` command checks the number of `/dev/sg*` entries on the host. The default number of `/dev/sg*` entries is 32. If you have more than 32 SCSI devices (for example, two physical SCSI hard drives, one SCSI CD-ROM, and 30 LUNs), `sanlun` does not list all the LUNs.

◆ Storage system is unreachable for a long time

If the time period during which the storage system is unreachable exceeds the timeouts in the `/etc/iscsi.conf` file, the `sanlun lun show all` command might not provide a complete list of mapped LUNs the first time you run the command.

Workaround: The following table lists the workaround for this problem.

Step	Action	
1	Depending on why the iSCSI LUNs are not displayed, complete the following step:	
	If...	Then...
	You have more than 32 SCSI devices	Create more <code>/dev/sg*</code> entries by using the Linux <code>mknod</code> command. For details about the <code>mknod</code> command, see the Linux man page. Example: <pre>mknod --mode=660 /dev/sg32 c 21 32</pre>
	The storage system was unreachable	1. Run the following command for the second time: sanlun lun show all 2. Ensure that the LUNs listed match the devices listed in the <code>/dev/iscsi</code> directory or in the output of the <code>iscsi-ls</code> command.

File system errors and system shutdown interruption on SUSE Linux Enterprise Server 9 SP3

Problem: If the initiator is accessing LUNs on a storage system and I/O takes a long time (for example, the target is unreachable, a CFO is in progress, or the target is heavily loaded), file system errors or system shutdown interruption might occur.

File system errors and system shutdown interruption depend on the disk command timeout and connection failure timeout settings in the `/etc/iscsi.conf` file.

By default, the disk command and connection failure timeouts are not listed in the `/etc/iscsi.conf` file. As a result, these timeout values are infinite and the host infinitely retries these I/O commands. The infinite timeouts prevent file system errors, but stopping the iSCSI service hangs or a shutdown is interrupted if the storage system I/O takes a long time.

Workaround: If you did not create Linux file systems on your LUNs, and you are using raw devices instead, you can set the timeouts to a value of 180 seconds to prevent system shutdown interruption. However, any application using these LUNs might receive SCSI I/O errors and file system errors if a storage system is unreachable for a long time.

To set the I/O command timeout values, complete the following steps:

Note

Do not change the timeout values, if you are using any Linux file systems on your storage system LUNs. This includes Reiser, ext2, and ext3 file systems.

Step	Action
1	Ensure that you are logged in to the host with root privileges.
2	With a text editor, open the host's <code>/etc/iscsi.conf</code> file.
3	At the beginning the file, add the following lines. Ensure that you add these lines before the storage system's <code>DiscoveryAddress</code> settings: <code>MaxDiskCommandTimeout=180</code> <code>ConnFailTimeout=180</code>

iSCSI node names

The iSCSI Linux initiator does not support node names with more than 255 characters.

If you use `iscsi nodename` to change the default node name of the storage system, ensure that the new node name does not exceed 255 characters.

Known problems and limitations with Red Hat Enterprise Linux hosts on iSCSI

GFS growth is not possible on a mounted file system on an iSCSI LUN

Problem: The online grow feature in the GFS enables you to extend the GFS without unmounting the file system. However, you cannot use this feature on a file system that is created over an iSCSI device or a multipath device on a NetApp LUN. This feature cannot be used on a file system created over an iSCSI and multipath device on a NetApp LUN.

When GFS is used on the LVM volume, the volume can be extended without unmounting the file system by adding new LUNs to the volume group and by expanding the logical volume on which the GFS is created.

Though storage systems provide dynamic resizing of the LUN, the iSCSI layer in Linux is not capable of detecting the LUN resize. The new size of the LUN is reflected on Linux only after the iSCSI service is restarted on Linux.

Workaround: Complete the following steps on all the nodes.

Step	Action
1	Unmount the file system.
2	Resize the LUN on the storage system.
3	Run the following command to flush the multipath: multipath -F
4	Run the following command to stop the multipath service: service multipath stop
5	Run the following command to start the iSCSI session: service iscsi restart
6	Run the following command to start the multipath service: service multipath start
7	Mount the file system.

Step	Action
8	Run the following command on one of the nodes: gfs_grow
9	Run the following command on all nodes to verify the new size: running gfs_grow -Tv mountpoint

For more information, see bug number 215923 on Bugs Online at <http://now.netapp.com/NOW/cgi-bin/bol>.

Using default partitions on Linux with Data ONTAP LUNs might require alignment for best performance

Problem: Using default partition settings on LUNs might result in performance problems. This problem can occur if the `sectors` value for a partition is not a multiple of 8.

Workaround: Manually enter the disk geometry parameters when creating partitions with the `fdisk` command.

For specific steps, see bug number 156121 on Bugs Online at <http://now.netapp.com/NOW/cgi-bin/bol>.

Host freezes intermittently if multipathd is not running in SAN boot scenario on Red Hat Enterprise Linux 5 Update 2

Problem: During controller faults, all paths stop operating for some time. Also, if the multipathd is not running, the root volume is unavailable, which leads to a system freeze. Another scenario can be a host shutdown. If paths stop operating after the multipath service is stopped, the host freezes, and you should perform a hard reboot.

Workaround: Ensure that the multipathd does not stop when the root volume is on the iSCSI multipathed device. If the multipathd is up and running, the `queue_if_no_path` feature stalls I/O on the root volume till some paths get reinstated. After one path gets reinstated, the root volume is accessible.

To ensure that the daemon does not stop, you have to modify the multipathd init script.

You can modify the script by applying the patch, which is available at the following Web site:

<https://bugzilla.redhat.com/attachment.cgi?id=314924>

For more information, see bug 295559 on Bugs Online at <http://now.netapp.com/NOW/cgi-bin/bol>.

Red Hat Enterprise Linux 5 Update 2 host intermittently stops while booting in SAN boot scenario

Problem: There is a problem with the iscsid and multipathd sequence during the boot process. Because the iscsid service is started before the multipathd service, there is a small window when the host sometimes loses connection with the iSCSI root volume, which leads to a system freeze.

Workaround: Configure the system after postinstallation by completing the following steps.

Step	Action
1	During the first boot, switch to single user mode by modifying the grub parameter.
2	Add single to the kernel parameters. kernel /vmlinuz-2.6.18-92.el5 ro root=/dev/VolGroup00/LogVol100 rhgb quiet single:
3	Create the /etc/multipath.conf file as described in “Configuring the /etc/multipath.conf file” in the <i>Installation and Setup Guide</i> .
4	Modify the /etc/iscsi/iscsid.conf file as described in “Editing the host’s configuration file” in the <i>Installation and Setup Guide</i> .
5	Ensure that the multipathd service is started before the iscsid service. This requires you to modify the /etc/init.d/multipathd script and change the start priority from SXX to S06. This assumes that the default iSCSI installation sets iscsid start priority at S07. If not, you need to ensure that the start priority of multipathd is less than the start priority of iscsid. Specify the line as shown in the following example: # chkconfig: - 06 87

Step	Action
6	Run the following commands in the order given: a. chkconfig multipathd off b. chkconfig multipathd on
7	Edit the /etc/fstab file and add the _netdev mount option for the root device. Specify the line as shown in the following example: <pre>/dev/mapper/mpath0p1 / ext3 _netdev,defaults 1 1</pre>
8	Reboot the system.

For more information, see bug number 215923 on Bugs Online at <http://now.netapp.com/NOW/cgi-bin/bol>.

Service clvmd restart fails with connect fail on a local socket

Problem: GFS and the underlying cluster LVM layer are used for providing cluster volume manager. In rare cases, the volume group does not come up after the boot process. This might happen if iSCSI LUNs are not available before the LVM start-up script is run.

In such cases, you might have to run `vgscan` after the boot process to read the volume groups from the iSCSI LUNs. This might also need a `clvmd` restart. If you run `service clvmd restart` to restart `clvmd`, in GFS 4.0, it fails. This failure is due to a timing issue in the `clvmd` service start script.

Workaround: To avoid the `clvmd` restart failure, you must stop and start the `clvmd`. To do this, completing the following steps.

Step	Action
1	Run the following command to stop clvmd: <pre>service clvmd stop</pre>
2	Run the following command to start clvmd: <pre>service clvmd start</pre>

OCFS2 self-fencing during cluster failover on Red Hat Enterprise Linux 4 Update 3

Problem: During a clustered failover or network disruption, nodes fence themselves by panicking.

Oracle Cluster File System (OCFS2) has a fencing mechanism that relies on each node being able to write a heartbeat value to a specific area on each OCFS2 volume. Every two seconds, each node writes a new heartbeat value, and reads the value of the other node. If a node is unable to write or read the heartbeat value, it retries the operation a specific number of times, and then self-fences by panicking the Linux kernel as it fails to obtain quorum.

During a clustered failover or network disruption, all paths to the storage system can become unavailable, and none of the nodes in the OCFS2 cluster is able to write or read the heartbeat values. As a result, the nodes might fence themselves by panicking.

Workaround: The workaround for this problem is to set the `O2CB_HEARTBEAT_THRESHOLD` value in `/etc/sysconfig/o2cb` to a value large enough to ensure that the nodes do not self-fence.

With a dedicated TCP/IP network for the iSCSI link between Linux hosts and storage systems, a reasonable value usually is 61. This corresponds to a total of 120 seconds before a node self-fences.

Depending on the load in the iSCSI network used between the Linux host and the storage system, you should raise the value. A value in the range of 61 to 201 is ideal for a configuration with OCFS2 on multipath and iSCSI devices.

For more information, see also:

http://oss.oracle.com/projects/ocfs2/dist/documentation/ocfs2_faq.html#HEARTBEAT

System stops responding when shutting down or rebooting on Red Hat Enterprise Linux 4 Update 3

Problem: The Linux system stops responding when shutting down or rebooting. The `/etc/init.d/iscsi stop` command might also report errors. This error occurs only in systems running DM-Multipath with iSCSI.

Workaround: This error is a known problem with the shutdown sequence in Red Hat Enterprise Linux 4 Update 3. If the system is currently not responding, disconnect power to force the shutdown to complete.

To prevent shutdown problems, complete the following steps before shutting down or rebooting.

Step	Action
1.	Unmount all file systems on iSCSI LUNs.
2.	Flush the multipath maps by using the following command: <code>multipath -F</code>
3.	Stop the iSCSI service by using the following command: <code>/etc/init.d/iscsi stop</code>
4.	Enter the <code>shutdown</code> or <code>reboot</code> command.

For more information, see Red Hat Bugzilla 177355 at <https://bugzilla.redhat.com/bugzilla/>.

See also bug number 183045 on Bugs Online at <http://now.netapp.com/NOW/cgi-bin/bol>.

Multiple devices shown for each LUN on Red Hat Enterprise Linux 4 Update 3

Problem: The `iscsi-ls -l` and `sanlun lun show` commands display multiple devices for each LUN that is mapped to the Linux host. This problem occurs only with Red Hat Enterprise Linux 4 Update 2 and Red Hat Enterprise Linux 4 Update 3 without DM-Multipath.

Workaround: The iSCSI initiator creates a separate iSCSI session to each iSCSI portal discovered on the storage system. If you are running Red Hat Enterprise Linux 4 Update 3, you can use DM-Multipath to manage multiple paths.

Otherwise, configure the storage system or network routing so that only one portal (interface) is visible to the Linux host. See the Host Utilities *Installation and Setup Guide* for more information.

Always map first LUN as LUN 0 on Red Hat Enterprise Linux 4 Update 4

Problem: If the first LUN mapped to a host has an ID other than 0, the `iscsi-ls -l` command shows incorrect information. Adding LUN 0 later does not solve the problem.

Workaround: Always map the first LUN to a host as LUN 0. This is the default behavior of Data ONTAP when you do not explicitly specify the LUN identifier.

For more information, see bug number 164689 on Bugs Online at <http://now.netapp.com/NOW/cgi-bin/bol>.

LUNs not recognized by DM-Multipath

Problem: LUNs are mapped to the Linux host and are correctly reported by the `iscsi-ls -l` and `sanlun lun show all` commands. However, the LUNs are not reported by the `multipath -ll` command. Multipath devices are not created for the LUNs.

Workaround: Ensure that the LUNs are not excluded (blacklisted) from the DM-Multipath configuration. Enter the following command on the Linux host to display the devices that are blacklisted:

```
multipath -v -d | grep blacklist
```

If any of the iSCSI devices are blacklisted check the `devnode_blacklist` section of the `/etc/multipath.conf` file. Pay special attention to any statements that are added to blacklist local SCSI disks.

Note

`devnode "sd[a-b]"` excludes not only `/dev/sda` and `/dev/sdb`, but also `/dev/sdaa`, `/dev/sdba`, and so on. Ensure that you use `devnode "sd[a-b]"` instead.

If the iSCSI devices are not blacklisted but are still not recognized by the `multipath` command, regenerate the multipath maps by entering the following command:

```
multipath -v3
```

For more information, see bug number 201302 on Bugs Online at <http://now.netapp.com/NOW/cgi-bin/bol>.

Storage system console displays continuous ISID Rule messages on Red Hat Enterprise Linux 4 host

Problem: The storage system console displays continuous ISID Rule messages, similar to the following:

```
Wed Jun 11 13:42:47 GMT [ss1: iscsi.notice:notice]: ISCSI: iswta,
ISID Rule: new connection from same initiator, shutting down old
session 0

Wed Jun 11 13:42:48 GMT [ss1: iscsi.notice:notice]: ISCSI: iswta,
ISID Rule: new connection from same initiator, shutting down old
session 1

Wed Jun 11 13:42:48 GMT [ss1: iscsi.notice:notice]: ISCSI: iswta,
New session from initiator ign.1987-05.com.cisco:rachman at IP addr
10.10.10.10

Wed Jun 11 13:42:50 GMT [ss1: iscsi.notice:notice]: ISCSI: iswta,
ISID Rule: new connection from same initiator, shutting down old
session 0

Wed Jun 11 13:42:50 GMT [ss1: iscsi.notice:notice]: ISCSI: iswta,
New session from initiator ign.1987-05.com.cisco:rachman at IP addr
10.10.10.10

Wed Jun 11 13:42:51 GMT [ss1: iscsi.notice:notice]: ISCSI: iswta,
ISID Rule: new connection from same initiator, shutting down old
```

Workaround: Add the following line to the `/etc/iscsi.conf` file. Ensure that this line is at the beginning of the file, before any “DiscoveryAddress” entries.

```
Continuous=no
```

Then restart the iSCSI service by entering the following command:

```
/etc/init.d/iscsi restart
```

Known problems and limitations with Veritas Storage Foundation

Commands hang when interconnect cable is pulled

Problem: When the NetApp controller interconnect cable is pulled, commands such as `multipath -ll`, `sanlun lun show`, and `vxdctl enable` hang.

Workaround: Ensure that the interconnect cable is properly connected and working.

Some paths are not reenabled after giveback when there are 1024 or more paths under Veritas DMP control

Problem: With Red Hat Enterprise Linux 4 Update 5 and Veritas Storage Foundation Cluster File System for Oracle RAC (SFCFSRAC) 5.0 MP3, some paths are not reenabled after the giveback process is complete. This is seen intermittently when there are 1024 or more paths under Veritas Dynamic Multipathing control. This issue has not been encountered when there are less than 1024 paths under Veritas DMP control.

For more information, see bug number 322127 on Bugs Online at <http://now.netapp.com/NOW/cgi-bin/bol>.

Workaround: Run the following command to reenable the previously disabled paths:

`vxdctl enable`

Delayed I/O seen on hosts during storage controller faults with iSCSI protocol

Problem: With Veritas Storage Foundation 5.1 using iSCSI protocol, it is observed that there is delayed IO on the host.

Since the iSCSI parameter `node.session.timeo.replacement_timeout` is set to 120 seconds (please refer BURT # 321376 for more details), whenever there are iSCSI session failures for more than 120 seconds, there will be a delay of more than 120 seconds in resuming IO on other active DMP paths.

This issue is seen when there will be an Ethernet cable failure or Ethernet port failure.

NetApp is working with Symantec to resolve this issue.

Sometimes Veritas volumes go off-line on Oracle Enterprise Linux 5.4 with Emulex 4Gb FC HBA

Problem: It is observed that on Oracle Enterprise Linux 5.4 with Emulex 4Gb FC HBA, FC ports go off-line intermittently due to bug in driver version 8.2.0.48.2p with msi interrupt handler. As a result of this Veritas volumes go off-line if there is a coincident storage controller fault.

Workaround: To overcome this issue update the Oracle Enterprise Linux 5.4 with latest kernel or with Emulex driver version 8.2.0.48.2p change the Emulex driver parameter "lpfc_use_msi" value to 0, the default value of this parameter is 2.

Perform the following steps to set Emulex driver parameter "lpfc_use_msi" to 0.

Step	Action
1.	In the /etc/modprobe.conf file, append the following line: options lpfc lpfc_use_msi=0
2.	Re-create the initrd by running the following command: mkinitrd -f /boot/initrd-"`uname -r`".img `uname -r` Note _____ Please take backup of the current initrd before running the mkinitrd command.
3.	Re-boot the host.
4.	Verify the value of lpfc_use_msi by using the following command: cat /sys/class/scsi_host/host<adapter_number>/lpfc_use_msi

Information about additional problems

Occasionally, additional problems are reported with the Support Kit or known bugs are fixed. The problems are listed in an online system called Bugs Online.

For a list of known and fixed bugs, access Bugs Online by completing the following steps.

Step	Action
1	Use a Web browser to go to: http://now.netapp.com
2	Click Service & Support.
3	Click Bugs Online & Release Tools.
4	In the Bug Search Tools section, choose Linux from the Bug Type(s) menu.
5	Click Go. Result: Bugs Online - Advanced Search Results lists all the bugs in Linux hosts.

Where to find additional information

Where to go for more information

The following table describes documents with the most current information about host and storage system requirements and additional documentation.

If you want...	Go to...
The most current system requirements for the Linux host	The NetApp Interoperability Matrix at http://now.netapp.com/NOW/products/interoperability
The supported storage system models for Data ONTAP licensed with FC	The following documents for your Data ONTAP release at the NOW site: ◆ <i>FC and iSCSI Configuration Guide</i> ◆ <i>System Configuration Guide</i>
Information about how to configure and manage the FC and iSCSI service on the storage system	The following documents for your Data ONTAP release: ◆ <i>Data ONTAP Block Access Management Guide for iSCSI and FC</i> ◆ <i>Data ONTAP Release Notes</i>
Best practices or configuration issues	The NetApp Knowledgebase
Information about DM-Multipath	The Multipath example files included with the device-mapper multipath RPM are in /usr/share/doc/device-mapper-multipath-<version>. <version> is the latest version number shipped with the Red Hat Enterprise Linux release.
Information about Red Hat Cluster and GFS	The Red Hat Web site
Information about Oracle Cluster File System	Oracle Cluster File System (OCFS2) User's Guide and OCFS2 - Frequently Asked Questions
Information about Oracle Real Application Cluster (RAC)	Oracle RAC installation information from the Oracle Web site.

If you want...	Go to...
NetApp Knowledge Base for Oracle RAC	Knowledge base for Oracle RAC at the following sites: ♦ http://www.netapp.com/library/tr/3339.pdf ♦ http://www.netapp.com/library/tr/3423.pdf
Information about configuring multipathing and the iSCSI initiator on SUSE Linux Enterprise Server	The Novell knowledgebase at the Novell Web site