

iSCSI SUSE® Linux® Host Utilities 3.1 for Native OS

Release Notes

Network Appliance, 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 215-02956_A0
September 2007

Copyright and trademark information

Copyright information

Copyright © 1994–2007 Network Appliance, Inc. All rights reserved. Printed in the U.S.A.

No part of this file covered by copyright may be reproduced in any form or by any means—graphic, electronic, or mechanical, including photocopying, recording, taping, or storage in an electronic retrieval system—without prior written permission of the copyright owner.

Software derived from copyrighted material of Network Appliance, Inc. is subject to the following license and disclaimer:

THIS SOFTWARE IS PROVIDED BY NETWORK APPLIANCE “AS IS” AND ANY EXPRESS OR IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE ARE DISCLAIMED. IN NO EVENT SHALL NETWORK APPLIANCE 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.

Network Appliance reserves the right to change any products described herein at any time, and without notice. Network Appliance assumes no responsibility or liability arising from the use of products described herein, except as expressly agreed to in writing by Network Appliance. 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 Network Appliance.

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 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, DataFabric, Data ONTAP, FAServer, FilerView, FlexClone, FlexVol, Manage ONTAP, MultiStore, NearStore, NetCache, SecureShare, SnapDrive, SnapLock, SnapManager, SnapMirror, SnapMover, SnapRestore, SnapValidator, SnapVault, Spinnaker Networks, SpinCluster, SpinFS, SpinHA, SpinMove, SpinServer, SyncMirror, Topio, VFM, and WAFL are registered trademarks of Network Appliance, Inc. in the U.S.A. and/or other countries. Cryptainer, Cryptoshred, Datafort, and Decru are registered trademarks, and Lifetime Key Management and OpenKey are trademarks, of Decru, a Network Appliance, Inc. company, in the U.S.A. and/or other countries. gFiler, Network Appliance, SnapCopy, Snapshot, and The evolution of storage are trademarks of Network Appliance, Inc. in the U.S.A. and/or other countries and registered trademarks in some other countries. ApplianceWatch, BareMetal, Camera-to-Viewer, ComplianceClock, ComplianceJournal, ContentDirector, ContentFabric, EdgeFiler, FlexShare, FPolicy, HyperSAN, InfoFabric, LockVault, NOW, NOW NetApp on the Web, ONTAPI, RAID-DP, RoboCache, RoboFiler, SecureAdmin, Serving Data by Design, SharedStorage, Simplicore, Simulate ONTAP, Smart SAN, SnapCache, SnapDirector, SnapFilter, SnapMigrator, SnapSuite, SohoFiler, SpinMirror, SpinRestore, SpinShot, SpinStor,

StoreVault, vFiler, Virtual File Manager, VPolicy, and Web Filer are trademarks of Network Appliance, Inc. in the U.S.A. and other countries. NetApp Availability Assurance and NetApp ProTech Expert are service marks of Network Appliance, Inc. in the U.S.A.

IBM, the IBM logo, AIX, and System Storage are trademarks and/or registered trademarks of International Business Machines Corporation.

Apple is a registered trademark and QuickTime is a trademark of Apple Computer, 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.

Network Appliance is a licensee of the CompactFlash and CF Logo trademarks.

Network Appliance NetCache is certified RealSystem compatible.

Table of Contents

Changes in this release 1

Contents of the Host Utilities. 2

Upgrading the operating system from the previous version 4

Known problems and limitations on SUSE Linux Enterprise Server 10 SP1 . . 5

Known problems and limitations on SUSE Linux Enterprise Server 9 8

For information about additional problems. 14

Where to find additional information. 15

Changes in this release

New features

iSCSI SUSE Linux Host Utilities 3.1 supports Open-iSCSI initiator on SUSE Linux Enterprise Server 10 SP1 hosts. iSCSI SUSE Linux Host Utilities 3.1 also supports root partition on multipath (SAN boot).

Supported packages

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

Linux Clustering Support

iSCSI SUSE Linux Host Utilities 3.1 supports cluster file systems and clustering suite on Linux hosts.

The usage of cluster site as framework for the file system, and the usage of iSCSI devices as shared quorum disk(s) by the cluster suites are supported.

The file system supported is:

- ◆ Oracle® Cluster File System (OCFS2)

The cluster suite supported is:

- ◆ Oracle Clusterware and Oracle Real Application Cluster (RAC) 10g R2

Documentation changes

The documentation changes include:

- ◆ Updates related to SUSE Linux Enterprise Server 10 SP1 iSCSI software initiator support in the *Setup Guide*.
- ◆ Updates related to root partition on mutipath (SAN Boot) and Troubleshooting information in the *Setup Guide*.
- ◆ Updates to Known Problems section of the *Release Notes*.

Contents of the Host Utilities

Software

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

The software is distributed as a single file (`netapp_linux_host_utils_3_1.tar.gz`), which you install on your Linux host. The installation instructions are provided in the 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 SUSE Linux Enterprise Server. The following table lists the important files in the distribution:

File	Purpose
<code>install</code>	Installs the Host Utilities software on a Linux host.
<code>uninstall</code>	Removes the Host Utilities software from the Linux host.
<code>sanlun_bin_netapp.tar</code>	Contains the files below.
<code>sanlun</code>	Displays information about LUNs on connected NetApp storage systems.
<code>filer_info</code>	Collects information about the storage system in a file that can be submitted to NetApp for problem analysis.
<code>linux_info</code>	Collects information about the Linux host in a file that can be submitted to NetApp for problem analysis.
<code>mpath_prio_ontap</code>	path priority callout program used in <code>/etc/multipath.conf</code> .

Documentation

The Host Utilities includes the following documentation.

Document	Purpose
Release Notes	Lists the contents of the Host Utilities, and any known problems or last minute updates. Download from the NOW Web site.
Setup Guide	Explains how to install and configure software to enable the Linux host to use LUNs on the storage system. Download from the NOW Web site.
Command Quick Reference	Lists the most common storage system and Linux commands used to manage iSCSI. Download from the NOW Web site.
man pages	Displays information about the Host Utilities software. The man pages for the <code>sanlun</code> , <code>filer_info</code> , and <code>linux_info</code> commands are copied to the <code>/opt/netapp/man/man1</code> directory when you install the host utilities software. Included in Host Utilities software distribution.
NOW pages	Host Utilities installation instructions, information updates, links to additional information. Browse on NOW Web site.

Upgrading the operating system from the previous version

Overview of upgrade procedures

If you are running a version of the operating system, and is upgrading to a later version, unmount iSCSI LUNs and stop the iSCSI service before you upgrade.

For general information on SUSE Linux Enterprise Server 10 SP1, see

<http://www.novell.com/linux/sp1highlights.html>.

For general information on SUSE Linux Enterprise Server 9 SP2, see

<http://www.novell.com/products/server9/sp2.html>.

For general information on SUSE Linux Enterprise Server 9 SP3, see

<http://www.novell.com/products/server9/sp3.html>.

Before upgrading to SUSE Linux Enterprise Server 9 SP2, and SP3

Complete the following steps before upgrading the operating system from SUSE Linux Enterprise Server 9 SP2 and SP3.

Step	Action
1	Unmount all existing iSCSI LUNs by running the following command on the SUSE Linux Enterprise Server host: <code>/sbin/iscsi-umountall</code>
2	Stop the iSCSI service on the SUSE Linux Enterprise Server host: <code>/etc/init.d/iscsi stop</code>
3	Continue with the upgrade.

Known problems and limitations on SUSE Linux Enterprise Server 10 SP1

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 can become unavailable, and none of the nodes in the OCFS2 cluster will be able to write or read the heartbeat values. As a result, the nodes may 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 will not self-fence.

NetApp recommends that, with a dedicated TCP/IP network for the iSCSI link between Linux hosts and the NetApp storage system, a reasonable value in most cases is 181. This corresponds to a total of 360 seconds before a node self-fences.

Depending on the load in the iSCSI network, the value should be raised. NetApp recommends that a value in the range of 61 to 201 would be ideal for a configuration with OCFS2 on mutipath, and iSCSI devices.

For more information, see the following:

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

iSCSI login errors with time out and the host is unable to see LUNs

Problem: When Jumbo frames are enabled on SUSE Linux Enterprise Server10 SP1 hosts, switch and filer logging into discovered targets causes login errors with timeout. Login is successful but hosts are unable to map any of the LUNs.

File system becomes read only after iSCSI stop/start faults.

Problem: File systems becomes read only after iSCSI stop/start faults.

LUN status undefined

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

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: Make sure 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 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
```

iSCSI node names

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

File system errors and system shutdown hangs

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 hangs might occur.

File systems on partitions may have performance problems

Problem: The default `fdisk` command parameters used to create a partition on a LUN may 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 causes the storage system to double its I/O workload for that LUN.

Workaround: You can explicitly specify the `fdisk` parameters when creating a partition to use values that guarantee the correct block alignment. Refer to bug number 156121 on Bugs Online at <http://now.netapp.com/NOW/cgi-bin/bol> for specific steps.

Known problems and limitations on SUSE Linux Enterprise Server 9

OCFS2 self fencing during cluster failover on SUSE Linux Enterprise Server 9 SP3

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 can become unavailable, and none of the nodes in the OCFS2 cluster will be able to write or read the heartbeat values. As a result, the nodes may 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 will not self-fence.

NetApp recommends that, with a dedicated TCP/IP network for iSCSI link between Linux hosts and NetApp filer, a reasonable value in most cases is 61. This corresponds to a total of 120 seconds before a node gives up, and self-fences.

Depending on the load in the iSCSI network used between the Linux host and NetApp filer, the value should be raised. Netapp recommends that a value in the range of 61 to 201 would be ideal for a configuration with OCFS2 on mutipath, and iSCSI devices.

For more information, see:

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

See also, bug 210340 on Bugs Online at:

<http://now.netapp.com/NOW/cgi-bin/bol>

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

Problem: Automated 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 multipath bootup 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.

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

As a result, the multipath bootup and shutdown sequence is not synchronized with iSCSI bootup and shutdown sequence. Multipath devices will not be 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 shutdown.

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

For unmounting, make sure that any multipath-based file systems are unmounted before the multipath utility is shutdown.

The following sequence of actions should be performed to ensure proper shutdown of multipath-based iSCSI file systems:

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 the multipath daemon 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 196997 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: Make sure 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 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 may have performance problems

Problem: The default `fdisk` command parameters used to create a partition on a LUN may 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 causes the storage system to double its I/O workload for that LUN.

Workaround: You can explicitly specify the `fdisk` parameters when creating a partition to use values that guarantee the correct block alignment. Refer to bug number 156121 on Bugs Online at <http://now.netapp.com/NOW/cgi-bin/bol> for specific steps.

iscsi-ls command displays error message

Problem: For some SUSE Linux versions, the `iscsi-ls -l` command displays the following error 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 actual 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 the following error message is issued:
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: Issue 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 this problem:

- ◆ 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 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 workarounds for this problem.

If...	Then...
You have more than 32 SCSI devices	Create more <code>/dev/sg*</code> entries using the Linux <code>mknod</code> command. For details about the <code>mknod</code> command, see the Linux man page. Example: <code>mknod --mode=660 /dev/sg32 c 21 32</code>
The storage system was unreachable	1. Run the <code>sanlun lun show all</code> command a second time. 2. Make sure 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 hangs**

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 hangs might occur.

File system errors and system shutdown hangs depend on 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 or a shutdown hangs if storage system I/O takes a long time.

If you did not create Linux file systems on your LUNs, and are using raw devices instead, you can set the timeouts to a value of 180 seconds to prevent system shutdown hangs. 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:

Caution _____
Do not change the timeout values in the following section 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. Make sure 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 the `iscsi nodename` to change the default node name of the storage system, ensure the new node name does not exceed 255 characters.

For information about additional problems

Occasionally, additional problems are reported with the Host Utilities 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 iSCSI-Linux from the Bug Type(s) menu.
5	Click Go. Result: The Bugs Online - Advanced Search Results lists all bugs for Linux iSCSI initiators.

Where to find additional information

Where to go for more information

The following table describes where to find additional information about using iSCSI on your storage system.

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 , or the NetApp SAN Support Matrix at http://now.netapp.com/NOW/knowledge/docs/san/fcp_iscsi_config/index.shtml .
Best practices or configuration issues	The NetApp Knowledgebase at ◆ https://now.netapp.com/eservice/kbbrowse ◆ http://now.netapp.com/Knowledgebase/solutionarea.asp?id=kb8190
The supported storage system models for Data ONTAP licensed with iSCSI	◆ The NetApp Interoperability Matrix at http://now.netapp.com/NOW/products/interoperability, or the NetApp SAN Support Matrix at http://now.netapp.com/NOW/knowledge/docs/san/fcp_iscsi_config/index.shtml. ◆ <i>System Configuration Guide</i> for Data ONTAP at http://now.netapp.com.
Information about how to configure and manage the iSCSI service on the storage system	The following documents for your Data ONTAP release: ◆ <i>Data ONTAP Block Access Management Guide for iSCSI</i> ◆ <i>Data ONTAP Release Notes</i>
Information about configuring multipathing on SUSE Linux Enterprise Server	The Novell knowledgebase article <i>How to setup / use multipathing on SLES</i> at: http://support.novell.com/techcenter/sdb/en/2005/04/sles_multipathing.html .

If you want...	Go to...
Information on OCFS2	◆ Oracle Cluster File System (OCFS2) User's Guide http://oss.oracle.com/projects/ocfs2/dist/documentation/ocfs2_users_guide.pdf. ◆ OCFS2 - Frequently Asked Questions http://oss.oracle.com/projects/ocfs2/dist/documentation/ocfs2_faq.html.
Information on Oracle RAC	http://www.oracle.com/technology/pub/articles/smiley_rac10g_install.html .
NetApp Knowledge Base for Oracle RAC	http://www.netapp.com/library/tr/3369.pdf http://www.netapp.com/library/tr/3423.pdf
Information about configuring iSCSI initiator on SUSE Linux Enterprise Server 10 SP1	SUSE Linux Enterprise Server 10 SP1 Installation and Administration Guide at: http://www.novell.com/documentation/sles10/sles_admin/data/book_in_fo_book_sles_admin.html .