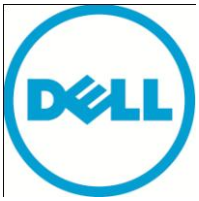
TECHNICAL REPORT



Deploying Red Hat Enterprise Linux, CentOS and Oracle Linux Server with Dell EqualLogic PS Series Arrays

Abstract

This technical report documents the procedure for configuring the Red Hat® Enterprise Linux® (RHEL), CentOS and Oracle® Linux Server® operating systems to operate with Dell™ EqualLogic™ PS Series storage arrays.

This document has been archived and will no longer be maintained or updated. For more information go to the Storage Solutions Technical Documents page on Dell TechCenter or contact support.

TR1062

V1.2

Copyright © 2011 Dell Inc. All
Rights Reserved.

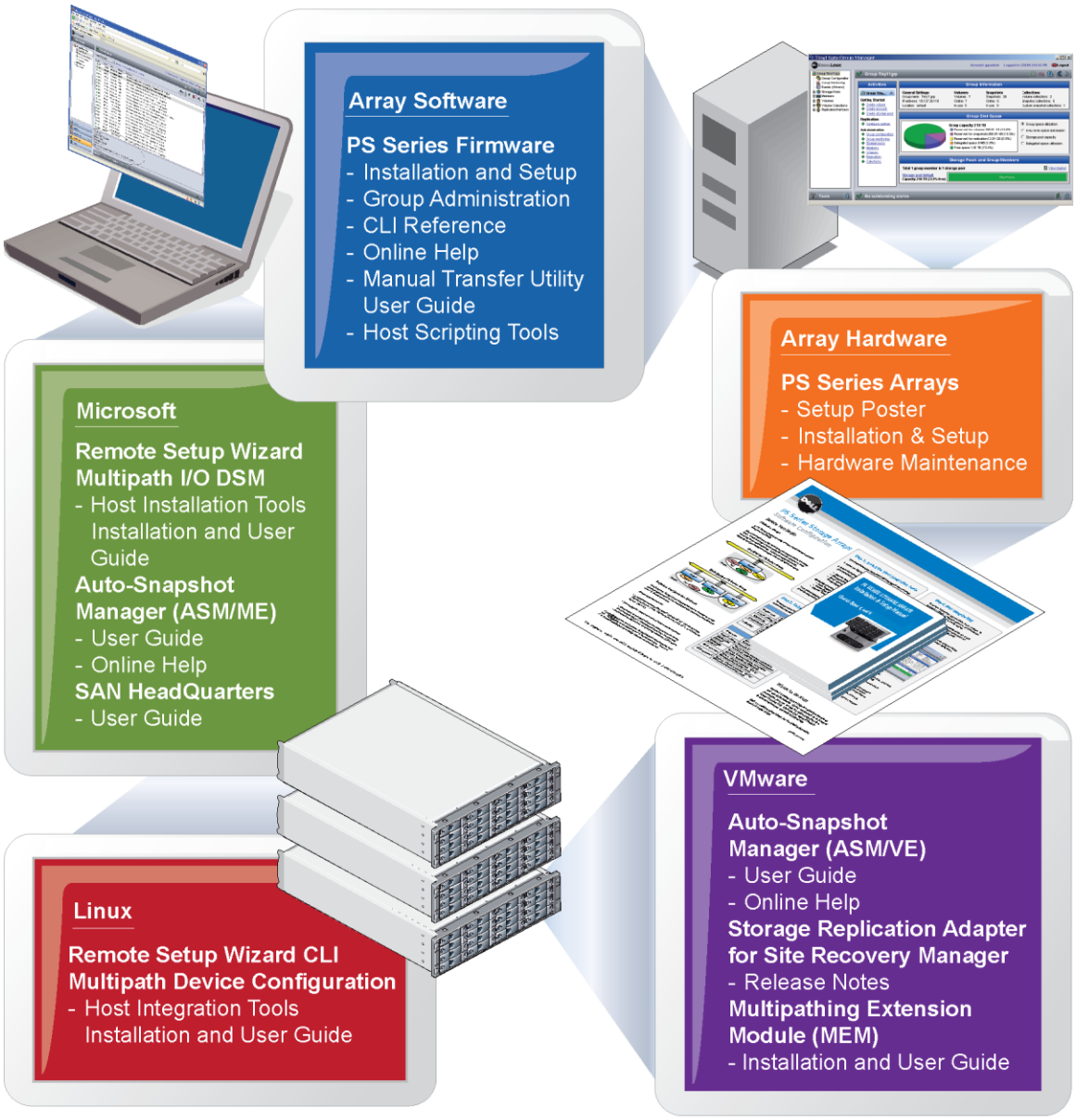
Dell EqualLogic is a trademark of Dell Inc.

All trademarks and registered trademarks mentioned herein are the property of their respective owners.

Possession, use, or copying of the documentation or the software described in this publication is authorized only under the license agreement specified herein.

Dell, Inc. will not be held liable for technical or editorial errors or omissions contained herein. The information in this document is subject to change.

[September 2011]



Current Customers Please Note: You may not be running the latest versions of the tools and software listed below. If you are under valid warranty or support agreements for your PS Series array, you are entitled to obtain the latest updates and new releases as they become available.

To learn more about any of these products, contact your local sales representative or visit the Dell EqualLogic site at http://www.equallogic.com. To set up a Dell EqualLogic support account to download the latest available PS Series firmware and software kits visit: https://www.equallogic.com/secure/login.aspx

Preface

Thank you for your interest in Dell™ EqualLogic™ PS Series storage products. We hope you will find them intuitive and simple to configure and manage.

PS Series arrays optimize resources by automating volume and network load balancing. Additionally, PS Series arrays offer all-inclusive array management software, host software, and firmware updates. The features and products described next are available as part of the all-inclusive software package and are at no additional cost.

PS Series Software

- Firmware: Installed on each array, PS Series firmware software allows you to manage your storage environment and provides capabilities such as volume snapshots, cloning, and replication to ensure data hosted on the arrays is protected in the event of an error or disaster.
  - Group Manager GUI: Provides a graphical user interface for managing a group.
  - Group Manager CLI: Provides a command line interface for managing a group.
- Manual Transfer Utility (MTU): Runs on Windows and Linux systems and enables you to use physical media to securely transfer large amounts of data to a replication partner, facilitating replication and preventing network congestion.
- SAN Headquarters (SANHQ): Provides centralized monitoring, historical performance trending, and event reporting for multiple PS Series groups.

Host Software for Linux

- Remote Setup Wizard Command Line Interface (RSWCLI): Discovers and initializes PS Series arrays, configures and manages access to PS Series arrays, and configures and manages multipathing.
- Multipath device configuration components: Provides the EqualLogic Connection Manager (ehcmd) daemon to manage multipath devices, a Device Mapper kernel module (dm-switch) to optimize routing of multipathing devices, and a command line interface (ehcmcli) that allows users to review the diagnostic state of EqualLogic multipathing.
- EqualLogic Host Performance and Tuning Tool Suite (eqltune), a utility used to validate configurable parameters against practices recommended by Dell.

Host Software for Windows®

- Remote Setup Wizard (RSW): Initializes new PS Series arrays, configures host connections to PS Series groups, and configures and manages multipathing.
- Remote Setup Wizard Command Line Interface (RSWCLI): Provides a command-line alternative to the Remote Setup Wizard.
- Multipath I/O Device Specific Module (MPIO DSM): Includes a connection awareness module that understands PS Series network load balancing and facilitates host connections to PS Series volumes.
- VSS and VDS Provider Services: Allows 3rd party backup software vendors to perform off-host backups.
- Auto-Snapshot Manager/Microsoft Edition (ASM/ME): VSS Requestor interface that uses PS Series snapshots, cloning, and replication to provide point-in-time application-consistent

protection of critical data for supported applications (including SQL Server, Exchange Server, Hyper-V, and NTFS file shares).

– PowerShell Tools: Provides a comprehensive set of PowerShell cmdlets for managing one or many PS Series groups.

Host Software for VMware

– Host Integration Tools for VMware®

  - Auto-Snapshot Manager/VMware Edition (ASM/VE): Integrates with VMware Virtual Center and PS Series snapshots to allow administrators to enable Smart Copy protection of Virtual Center folders, datastores, and virtual machines.

  - EqualLogic Datastore Manager: Allows administrators to create and manage datastore on EqualLogic PS Series arrays from within vCenter.

  - Virtual Desktop Deployment Utility: Automates the deployment of virtual desktops in a VMware View environment.

– Storage Adapter for Site Recovery Manager (SRM): Allows SRM to understand and recognize PS Series replication for full SRM integration.

– Multipathing Extension Module (MEM): Provides enhancements to existing VMware multipathing functionality. Includes a connection awareness module that understands PS Series network load balancing and facilitates host connections to PS Series volumes.

Table of Contents

Revision information vii
Introduction 1
Terminology 1
Deployment scenarios 3
    Single-path I/O connection environment: 3
    Multi-path I/O connection environment: 3
Server setup – pre-requisite steps 4
    Ethernet interfaces 5
    iSCSI Initiator 5
    Configuring disk timeout for Linux 2.6 Kernel SCSI devices 6
    Web-browser configuration 6
Procedures – how do I? 7
Summary 19
Technical Support and Customer Service 20

Revision information

The following table describes the release history of this Technical Report.

Report	Date	Document Revision
1.0	August 2010	Initial Release
1.1	September 2010	Added 'service multipathd reload' - step #14 in Task 5
1.2	September 2011	Included CentOS and Oracle Linux Server

The following table shows the software and firmware used for the preparation of this Technical Report.

Vendor	Model	Software Revision
Red Hat	Red Hat Enterprise Linux	RHEL 5 Update 4
Oracle	Oracle Linux Server	Version 5 Update 7
Community-supported	Cent0S	Cent0S 5 Update 6
Dell	PS Series Firmware	Version 3.4 and later

The following table lists the documents referred to in this Technical Report. All PS Series Technical Reports are available on the Customer Support site at: *support.dell.com*

Vendor	Document Title
Dell	Red Hat Linux v5.x Software iSCSI Initiator Configuration and Tuning Guide http://www.equallogic.com/resourcecenter/assetview.aspx?id=8727
Red Hat	DM Multipath Configuration and Administration, Edition 3 http://www.redhat.com/docs/en-US/Red_Hat_Enterprise_Linux/5.4/html/DM_Multipath/index.html
Red Hat	Knowledgebase article on device-mapper device naming https://access.redhat.com/kb/docs/DOC-5551 (login required)
Red Hat	MultipathUsageGuide - Device-mapper and LVM2 http://sources.redhat.com/lvm2/wiki/MultipathUsageGuide
Dell	PS Series Storage Arrays – iSCSI Initiator and Operating System Considerations https://www.equallogic.com/support/download_file.aspx?id=928 (login req'd)

Introduction

This technical report documents the configuration procedures for a server running Red Hat® Enterprise Linux® 5 (RHEL 5) with a Dell™ EqualLogic™ PS Series storage array. This report has been updated to include CentOS 5 and Oracle Linux Server 5. The procedures to deploy Cent0S and Oracle Linux Server with a Dell EqualLogic PS Series storage array are mostly identical with RHEL 5. The instances where the procedures are slightly different have been noted. The report begins by introducing relevant terminology and detailing the hardware, software and network configuration used in the development of this report. Then, two deployment scenarios are discussed: single-path I/O and multi-path I/O. The next section, titled 'Server Setup – Pre-requisite Steps', guides the reader through configuration of the RHEL server to work with a PS Series Array. The final section of the report titled 'Procedures – How Do I?" is a cookbook for common storage administrative tasks. This section covers tasks such as: "how do I add a volume?"; "how do I delete a volume?" and others.
This technical report only addresses configuration procedures in Linux native mode. If you prefer using the EqualLogic Host Integration Tools for Linux, download TR1065: Configuring and Deploying EqualLogic Host Integration Toolkit for Linux.

Note that this technical report does not address performance tuning. For that information, the reader is referred to the Dell document "Red Hat Linux v5.x Software iSCSI Initiator Configuration and Tuning Guide". Please refer to the 'Revision Information' section of this document for the URL.

Terminology

Multipath: technique which provides more than one physical path between a server system and storage devices. See http://en.wikipedia.org/wiki/Multipath_I/O

Device Mapper: Linux framework to map one block device onto another. Essential part of Linux Multipath I/O. For additional information, the reader is referred to http://en.wikipedia.org/wiki/Device_mapper

Round-Robin: Multipath policy which routes I/O requests, in rotation, to each available data path.

NIC: Network Interface Controller, a network interface card or Ethernet adapter.

RFC 3270: Internet Standards Track which defines transporting of the SCSI protocol over a TCP/IP network that allows block access to target devices. This standard is implemented within the PS Series Array.

iSCSI: Internet Small Computer System Interface, an IP-based storage networking standard.

iSCSI initiator: The iSCSI client, which runs on the RHEL server.

iSCSI target: The iSCSI storage device, in the test configuration this is the PS Series storage array.

WWID: World Wide Identifier, used to uniquely identify a particular scsi device (e.g. iSCSI target).

iqn: iscsi qualified name, a unique identifier used to identify both iscsi initiators and iscsi target volumes. Here is a sample iscsi qualified name, for an initiator on a RHEL server:

iqn.1994-05.com.redhat:938c6b6da1c2

Here is a sample iscsi qualified name, for a target volume named 'rhvol' on a PS Series storage array:

iqn.2001-05.com.equallogic:0-8a0906-fdce66602-95e792190014c000-rhvol

RHEL 5 (Red Hat Enterprise Linux 5)

RHEL 5: iSCSI Components (also includes Cent0S 5 and Oracle Linux 5)

- iscsiadm: Open-iSCSI administration utility, included in the 'iscsi-initiator-utils' package within the RHEL software distribution. Command line tool which allows: discovery of iscsi targets; login/logout of iscsi targets; and management of the iscsi records database.
- iscsid: Open-iSCSI daemon. Runs on the RHEL server and provides block level access to the PS Series storage array. Managed by the iscsi service.

➢ Required RHEL 5 packages:

  - iscsi-initiator-utils-*: iscsi daemon and utility programs

RHEL 5 Device-mapper, Multipath I/O and Volume Management Components

- LVM2: provides logical volume management facilities on Linux.
- Device-mapper: component of the 2.6 linux kernel. An integral part of Multipath I/O (Device Mapper Multipath).
- dmsetup: user space configuration tool for device mapper.
- multipath: the administration utility, part of the RHEL software distribution. Command line tool which allows configuration, removal and listing of multipathed devices.
- multipathd: the multipath daemon. Constantly monitors the devices under its control, and their paths. Activates or deactivates paths based on path condition and failback policies.

➢ Required RHEL 5 packages:

  - device-mapper-*: device mapper library
  - device-mapper-event-*: the event daemon
  - device-mapper-multipath-*: tools to manage multipath
  - lvm2-2-*: logical volume management tools
  - system-config-lvm-*: GUI for configuring Logical Volumes (optional)

Test configuration

Hardware

- Dell PowerEdge™ 1950 server
- Four 1GbE interfaces: eth0 -> eth3
- Dell EqualLogic PS5000 Storage Array
- GbE Switch and cables

Network/IP Addresses

- Server eth0 = 172.17.5.114 (data network)
- Server eth1 = not connected
- Server eth2 = 10.10.5.16 (storage network – path1)
- Server eth3 = 10.10.5.17 (storage network – path2)
- PS5000 Group IP = 10.10.5.10
- Gateway = 172.17.5.1
- DNS: primary = 172.16.1.27 : secondary = 172.16.1.7
- Netmask = 255.255.255.0 (all networks)

Software

- Red Hat Enterprise Linux Server 5.4, also known as RHEL 5 Update 4
- PS Series firmware version 5.1

Deployment scenarios

PS Series arrays connect to servers through physical Ethernet connections, either 1GbE or 10GbE. If the server has multiple NIC's , then one can be used for data traffic and one (or more) can be used for iSCSI storage traffic.

Single-path I/O connection environment:

In a single-path I/O configuration one physical Ethernet connection exists between the iSCSI initiator and the PS Series array. Refer to Figure 1 for a visual. Since there is only one physical connection, no redundancy exists. If the connection path fails, communication between the server (iSCSI initiator) and the PS Series array (iSCSI target volume) is lost.

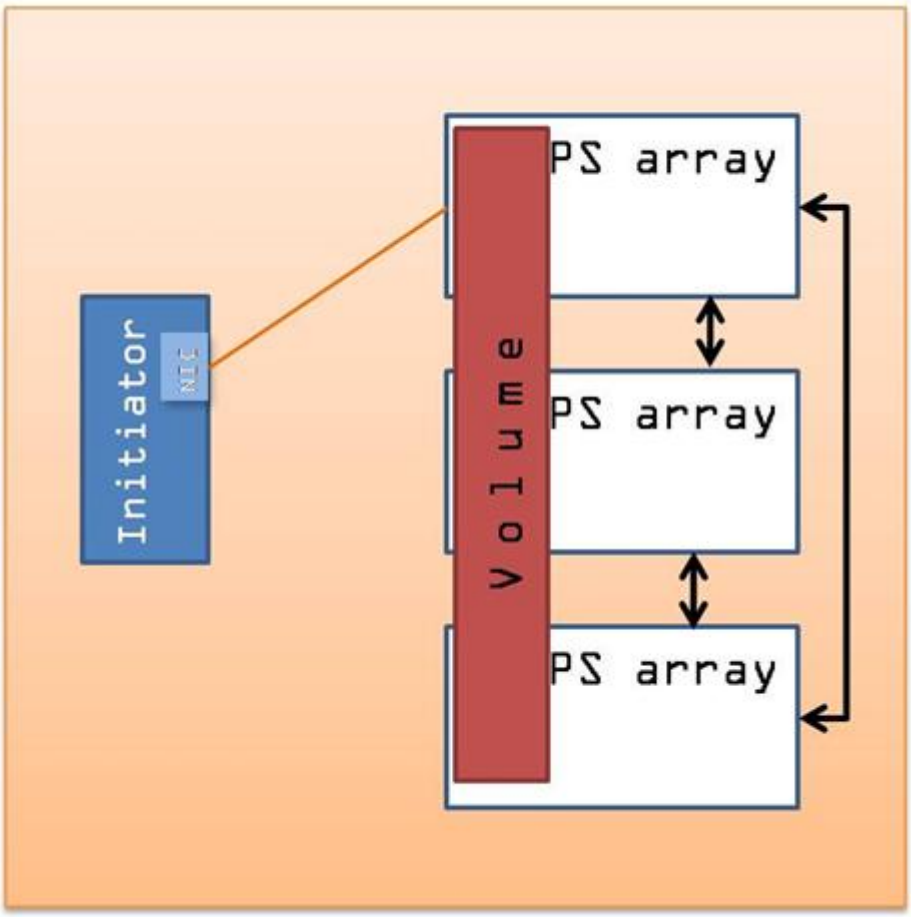


Figure 1: Single-path I/O connection environment

Multi-path I/O connection environment:

In a multipath I/O configuration two (or more) physical Ethernet connections exist between the iSCSI initiator and the PS Series array. EqualLogic multipath I/O provides high availability and performance load balancing across multiple network ports, allowing the iscsi initiator software on

the server to utilize multiple network interfaces with a single iscsi target. Refer to Figure 2 for a visual. The multiple paths can be utilized in one of two ways, either to improve the throughput between the initiator and the target or to implement redundancy for high availability. The multipath policy set on the initiator (RHEL server) determines how the multiple connection paths are utilized.

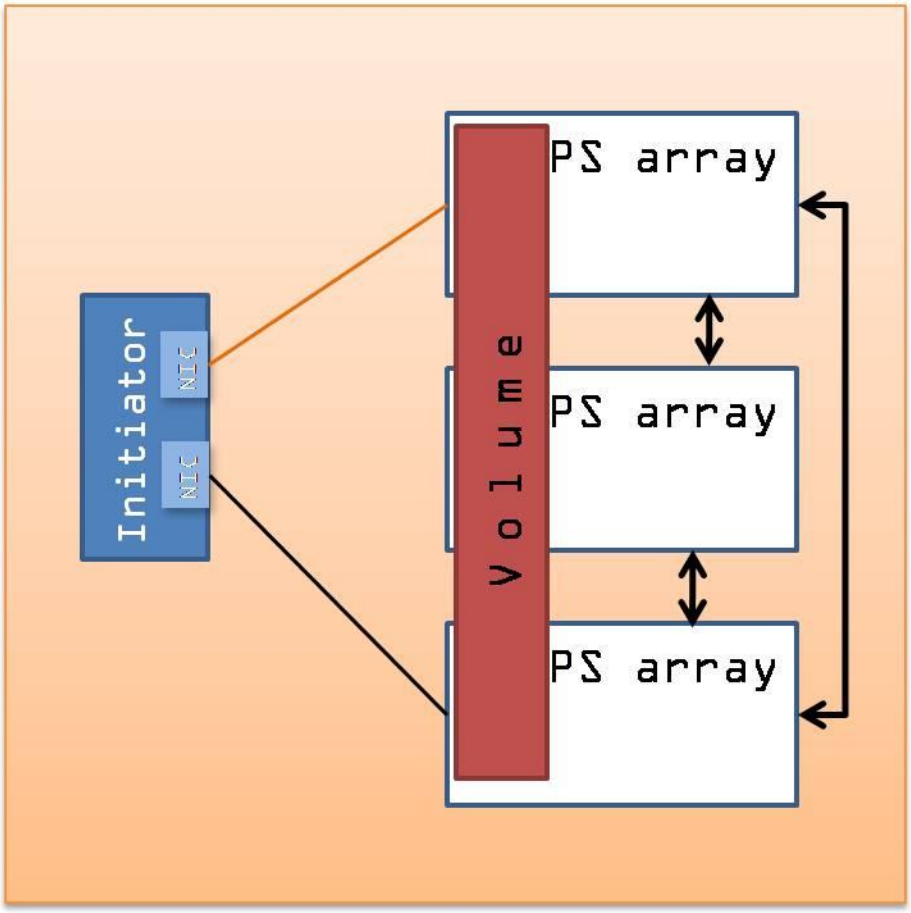


Figure 2: Multi-path I/O connection environment

Red Hat Enterprise Linux 5.2 and later supports multipath policies of 'round-robin' and 'failover'. A multipath policy of 'round-robin' will utilize multiple active connections between the iSCSI initiator and a target. A policy of 'round-robin' in the configuration shown in Figure 2 would double the potential network throughput between the iSCSI initiator and target. When a multipath policy of 'fail-over' is chosen, only a single connection will be active between the iSCSI initiator and target. Therefore the network throughput between the initiator and target is not increased. Instead, a redundant connection path is available to be activated in the event of a connection failure. If the multipath daemon on the RHEL server detects a failure in the primary connection path, it will redirect traffic to the alternate connection path.

Server setup – pre-requisite steps

This section of the document discusses the basic server setup requirements and describes how to configure the server to meet them. The areas of consideration are: Ethernet interfaces; iSCSI Initiator and web browser configuration.

Ethernet interfaces

The server and the storage array communicate over Ethernet. While the basic functionality of the storage array can be demonstrated over a single Ethernet connection, the advanced features (e.g. Multipath I/O) require that the server be configured with multiple Ethernet interfaces. For this test configuration, the server has four 1 gigabit Ethernet interfaces, three of which are utilized. It is recommended to verify the proper configuration of the Ethernet interfaces in the server.

Verify the connection paths between the server and the PS Series storage array. Note that in this test configuration, the eth2 and eth3 interfaces are connected to the storage network and the Group IP Address of the PS Series storage array is 10.10.5.10

```
# ping –I eth2 10.10.5.10
PING 10.10.5.10 (10.10.5.10) from 10.10.5.16 eth2: 56(84) bytes
of data.
64 bytes from 10.10.5.10: icmp_seq=1 ttl=255 time=0.106 ms
64 bytes from 10.10.5.10: icmp_seq=2 ttl=255 time=0.083 ms

# ping –I eth3 10.10.5.10
PING 10.10.5.10 (10.10.5.10) from 10.10.5.17 eth3: 56(84) bytes
of data.
64 bytes from 10.10.5.10: icmp_seq=1 ttl=255 time=0.614 ms
64 bytes from 10.10.5.10: icmp_seq=2 ttl=255 time=0.083 ms
```

iSCSI Initiator

An iSCSI initiator is included in RHEL 5. The iSCSI initiator kernel modules and utilities have been improved throughout the RHEL 5 releases. PS Series Storage Arrays require servers running RHEL 5.2 or later.

Note: For greatest functionality and stability, it is strongly recommend that servers connected to PS Series storage arrays be running RHEL 5 Update 4 or later. There are known issues with earlier versions. For example, the SE Linux subsystem in RHEL 5.2, and earlier versions, can interfere with iSCSI operations.

The iSCSI initiator is included as a kernel module in RHEL 5. However, the utilities to manage the initiator are in the optional 'iscsi-initiator-utils' package. Follow these steps to check installation status of the iscsi-initiator-utils package and configure the iSCSI Initiator daemon on the server:

1) Check the installation status of the iscsi-initiator-utils package. Install the package if it is not already installed:

```
# yum list | grep iscsi
iscsi-initiator-utils.x86_64   6.2.0.871-0.10.el5   installed
```

2) Start the service:

```
# service iscsi start
iscsid is stopped
Turning off network shutdown. Starting iSCSI daemon:   [OK]
                                                       [OK]
Setting up iSCSI targets: iscsiadm: No records found!  [OK]
```

3) Verify the service is enabled for boot:

```
# chkconfig --list iscsi
```

```
iscsi    0:off    1:off    2:off    3:on     4:on     5:on     6:off
```

4) In this test configuration the iscsi service was already enabled to execute on boot. If the system displays "off" for runlevels 3, 4, 5 - then use the following command to enable the service for boot:

```
# chkconfig iscsi on
```

Configuring disk timeout for Linux 2.6 Kernel SCSI devices

The following text was extracted from the Dell document:
"PS Series Storage Arrays – iSCSI Initiator and Operating System Considerations"
(Please refer to the 'Revision Information' section of this document for the URL).

The default disk timeout setting for Linux 2.6 kernels is 30 seconds. The desired value is 60 seconds.

Make the appropriate change for your operating system:

•For most Linux operating systems, use the following script to update the timeout value:

```
for i in 'ls /sys/class/scsi_disk';do echo "60" >
/sys/class/scsi_disk/$i/device/timeout;done
```

•For Red Hat operating systems, use the following script to update the timeout value:

```
for i in 'ls /sys/class/scsi_device';do echo "60" >
/sys/class/scsi_device/$i/device/timeout;done
```

You can add the appropriate script to your system start up script. The location of the startup script varies with different Linux distributions; for example, /etc/rc.local, /etc/init.d/rc.local.

If devices are connected after startup, run this script again to update the new devices' timeout value.

Web-browser configuration

The group manager GUI application requires a java-enabled web browser. If java is not installed, the group manager GUI will detect that condition and instruct the user to download java.

Depending on the installed browser, the java plug-in may have to be added manually. In this test configuration, that was the case. This setup used firefox version 3.0.12 and jre v1.6.0_20. Follow these steps to manually add the java plug-in to the browser:

1) Download the latest java plug-in and install in /usr/java.

   http://www.java.com/en/download/manual.jsp

   In this configuration, the file downloaded was 'jre-6u20-linux-x64.bin'

2) Shutdown all instances of the web browser

3) Make the java installer file executable:

```
# chmod 755 jre-6u20-linux-x64.bin
```

4) Execute the installer. This creates the java release directory (e.g. jre1.6.0_20):

```
# ./jre-6u20-linux-x64.bin
```

5) Create the /usr/java directory, if it doesn't already exist:

```
# mkdir /usr/java
```

6) Move the java release directory to the /usr/java directory:

```
# mv jre1.6.0_20/ /usr/java/.
```

7) Create a symbolic link from the firefox plugins directory to the java plug-in shared library:

```
# cd /usr/lib64/firefox-3.0.12
# mkdir plugin
# cd plugin
# ln –s /usr/java/jre1.6.0_20/lib/amd64/libnpjp2.so libnpjp2.so
```

8) Start a web browser and open the PS Group Manager GUI:

```
http://10.10.5.10   (the IP Address of the PS Group Mgr)
```

Procedures – how do I?

The remainder of the document follows a procedural format, guiding the reader through common storage administration tasks. These procedures assume that the reader has already completed the 'Server Setup – Pre-requisite Steps' section, earlier in this document. Please verify that the pre-requisites have been met prior to attempting the next steps.

The tasks begin with operations in a single-path I/O connection environment (Figure 1) and then document configuration of a multi-path I/O (MPIO) connection environment (Figure 2).

Here is a list of the common storage administration tasks covered in this document:

- Add a Volume
- Grow an Existing Volume
- Access a Snapshot
- Delete a Volume
- Configure Multi-path I/O (MPIO) and Verify
- Mount an MPIO Volume using a Persistent Name

Task #1 - Add a volume

1) Login to the PS Series storage array using the PS Group Manager GUI and create a new volume:

Example settings:

- `Volume name = rhel1`
- `Volume size = 40G`
- `iSCSI Access = 10.10.5.16 (IP Address of the eth2 interface)`
- `Access Type = set read/write (default)`

2) Discover the newly added volume from the server:

```
# iscsiadm -m discovery -t sendtargets -p 10.10.5.10
10.10.5.10:3260,1 iqn.2001-05.com.equallogic:0-8a0906-
f57496702-6a5ebc924194b842-vss-control
10.10.5.10:3260,1 iqn.2001-05.com.equallogic:0-8a0906-
beee66602-27e79218c634bfec-rhel1
```

3) Login to the discovered target, use the iscsi qualified name <iqn> returned by the discovery command (above):

```
# iscsiadm -m node -T <iqn> -p 10.10.5.10 --login
Logging in to [iface: default, target: iqn.2001-
05.com.equallogic:0-8a0906-beee66602-27e79218c634bfec-rhel1,
portal: 10.10.5.10,3260]
Login to [iface: default, target: iqn.2001-05.com.equallogic:0-
8a0906-beee66602-27e79218c634bfec-rhel1, portal:
10.10.5.10,3260]: successful
```

4) Find the newly created device name, using the iscsiadm command. In this test configuration, the new volume is /dev/sdb:

```
# iscsiadm -m session -P3
iSCSI Transport Class version 2.0-871
version 2.0-871
Target: iqn.2001-05.com.equallogic:0-8a0906-beee66602-
27e79218c634bfec-rhel1
   Current Portal: 10.10.5.13:3260,1
    Persistent Portal: 10.10.5.10:3260,1
    **********
    Interface:
    **********
     Iface Name: default
     Iface Transport: tcp
     Iface Initiatorname: iqn.1994-05.com.redhat:6891f663d4ce
     Iface IPaddress: 10.10.5.16
     Iface HWaddress: <empty>
     Iface Netdev: <empty>
     SID: 4
     iSCSI Connection State: LOGGED IN
     iSCSI Session State: LOGGED_IN
     Internal iscsid Session State: NO CHANGE
     ************************
     Negotiated iSCSI params:
     ************************
     HeaderDigest: None
     DataDigest: None
     MaxRecvDataSegmentLength: 262144
     MaxXmitDataSegmentLength: 65536
     FirstBurstLength: 65536
     MaxBurstLength: 262144
     ImmediateData: Yes
     InitialR2T: No
     MaxOutstandingR2T: 1
```

```
     ************************
     Attached SCSI devices:
     ************************
     Host Number: 8  State: running
     scsi8 Channel 00 Id 0 Lun: 0
        Attached scsi disk sdb          State: running
```

The command output indicates that for this test configuration /dev/sdb is the device name of the volume, as highlighted in RED above.

5) Make the filesystem using the 'mkfs' command. This example uses the '-j' option to create an ext3 filesystem with journaling and a volume label name of 'rhel1'. Of course you can partition the disk first if you'd like (refer to 'parted(8)'). Alternatively, the LVM tool could be used to initialize the volume and assign the space to a logical volume:

```
# mke2fs -L rhel1 -j /dev/sdb
   mke2fs 1.39 (29-May-2006)
   /dev/sdb is entire device, not just one partition!
   Proceed anyway? (y,n) y
   Filesystem label=rhel1
   OS type: Linux
   Block size=4096 (log=2)
   Fragment size=4096 (log=2)
   5248992 inodes, 10487040 blocks
   524352 blocks (5.00%) reserved for the super user
   First data block=0
   Maximum filesystem blocks=4294967296
   321 block groups
   32768 blocks per group, 32768 fragments per group
   16384 inodes per group
   Superblock backups stored on blocks:
32768, 98304, 163840, 229376, 294912, 819200, 884736, 1605632,
2654208, 4096000, 7962624

   Writing inode tables: done
   Creating journal (32768 blocks): done
   Writing superblocks and filesystem accounting information:
done

This filesystem will be automatically checked every 28 mounts
or 180 days, whichever comes first.  Use tune2fs -c or -i to
override.
```

6) Mount the filesystem using the 'mount' command:

```
# mkdir /mnt/rhel1

# mount -t ext3 /dev/sdb /mnt/rhel1
```

7) Add the filesystem to the /etc/fstab file to mount it at boot. Here is the edited fstab file, with the new entry displayed in RED text:

```
# cat /etc/fstab
```

```
/dev/VolGroup00/LogVol00 /          ext3    defaults          1 1
LABEL=/boot              /boot      ext3    defaults          1 2
tmpfs                    /dev/shm   tmpfs   defaults          0 0
devpts                   /dev/pts   devpts  gid=5,mode=620    0 0
sysfs                    /sys       sysfs   defaults          0 0
proc                     /proc      proc    defaults          0 0
/dev/VolGroup00/LogVol01 swap       swap    defaults          0 0
/dev/sdb                 /mnt/rhel1 ext3    _netdev,defaults  0 0
```

Note: the '_netdev" mount option ensures that the PS Series volume is mounted at the correct time in the boot process, after the iscsi and iscsid daemons have initialized. The '0 0' settings at the end of the line direct the dump and fsck programs to skip this volume. These settings are recommended for PS Series mounted volumes.

Task #2 - Grow an existing volume

1) Observe existing mounted volume size(s) from the server:

```
# df
Filesystem       1K-blocks      Used Available Use% Mounted on
/dev/mapper/VolGroup00-LogVol00
                  50786940   3710672  44454828   8% /
/dev/sda1           101086     12538     83329  14% /boot
tmpfs              8216020         0   8216020   0% /dev/shm
/dev/sdb          41289276    180240  39011628   1% /mnt/rhel1
```

2) Use the PS Group Manager's 'Modify Settings' feature to change the volume size. In this test configuration, the volume size was increased from 40G to 50G:

```
Volume -> Modify settings -> Select the Space tab -> Volume size
```

3) Rescan the device:

```
# iscsiadm -m node -p 10.10.5.10 --rescan
Rescanning session [sid: 1, target: iqn.2001-05.com.equallogic:
0-8a0906-beee66602-27e79218c634bfec-rhel1, portal:
10.10.5.10,3260]
```

4) Confirm the modified volume size was recognized:

```
# dmesg | tail
sdb: detected capacity change from 42954915840 to 53697576960
SCSI device sdb: 104878080 512-byte hdwr sectors (53698 MB)
sdb: Write Protect is off
sdb: Mode Sense: 91 00 00 00
SCSI device sdb: drive cache: write through
```

5) Resize the filesystem:

```
# resize2fs /dev/sdb
resize2fs 1.39 (29-May-2006)
Filesystem at /dev/sdb is mounted on /mnt/rhel1; on-line
resizing required
```

```
Performing an on-line resize of /dev/sdb to 13109760 (4k)
blocks.
The filesystem on /dev/sdb is now 13109760 blocks long.
```

Note: In order to follow this procedure, the filesystem and kernel must support online resizing. Otherwise the filesystem must be unmounted before the resize operation can be performed.

6) Verify the volume's new size:

```
# df
Filesystem    1K-blocks      Used Available Use% Mounted on
/dev/mapper/VolGroup00-LogVol00
              50786940   3710700    44454800   8%   /
/dev/sda1     101086     12538      83329      14%  /boot
tmpfs         8216020    0          8216020    0%   /dev/shm
/dev/sdb      51615928   184272     48816128   1%   /mnt/rhel1
```

Task #3 - Access a Snapshot

1) Use the Group Manager GUI to take a snapshot and set the snapshot online:

```
Select the volume (rhel1) -> Snapshots -> Create snapshot now
```

2) Discover the new volume (rhel-date-timestamp):

```
# iscsiadm -m discovery -t sendtargets -p 10.10.5.10
10.10.5.10:3260,1 iqn.2001-05.com.equallogic:0-8a0906-
f57496702-6a5ebc924194b842-vss-control
10.10.5.10:3260,1 iqn.2001-05.com.equallogic:0-8a0906-
beee66602-27e79218c634bfec-rhel1
10.10.5.10:3260,1 iqn.2001-05.com.equallogic:0-8a0906-
ae2e66602-b156e11d6b04c0e7-rhel1-2010-06-08-13:16:18.6512
```

3) Login to the discovered target, use the iscsi qualified name <iqn> returned by the discovery command (above):

```
# iscsiadm -m node -T <iqn> -p 10.10.5.10 --login
Logging in to [iface: default, target: iqn.2001-
05.com.equallogic:0-8a0906-ae2e66602-b156e11d6b04c0e7-rhel1-
2010-06-08-13:16:18.6512, portal: 10.10.5.10,3260]
Login to [iface: default, target: iqn.2001-05.com.equallogic:0-
8a0906-ae2e66602-b156e11d6b04c0e7-rhel1-2010-06-08-
13:16:18.6512, portal: 10.10.5.10,3260]: successful
```

4) Find the iSCSI device name of the snapshot:

```
# iscsiadm -m session -P3
iSCSI Transport Class version 2.0-871
version 2.0-871
Target: iqn.2001-05.com.equallogic:0-8a0906-ae2e66602-
b156e11d6b04c0e7-rhel1-2010-06-08-13:16:18.6512
   Current Portal: 10.10.5.13:3260,1
   Persistent Portal: 10.10.5.10:3260,1
   < SNIP – Lines omitted >
```

```
        ************************
        Attached SCSI devices:
        ************************
        Host Number: 15 State: running
        scsi15 Channel 00 Id 0 Lun: 0
          Attached scsi disk sdc          State: running
Target: iqn.2001-05.com.equallogic:0-8a0906-1b9e66602-
cbb6e11d2a34c0d1-jthrh2
     Current Portal: 10.10.5.13:3260,1
     Persistent Portal: 10.10.5.10:3260,1
        < SNIP - Lines omitted >
        ************************
        Attached SCSI devices:
        ************************
        Host Number: 8  State: running
        scsi8 Channel 00 Id 0 Lun: 0
          Attached scsi disk sdb          State: running
```

The command output indicates that for this test configuration /dev/sdb is the device name of the original volume and /dev/sdc is the device name of the snapshot volume, both highlighted in RED above.

5) Mount the volume:

```
# mkdir /mnt/rhelsnap
# mount –t ext3 /dev/sdc /mnt/rhelsnap
```

Task #4 - Delete the volumes

1) Unmount the previously created filesystems:

```
# umount /mnt/rhelsnap
# umount /mnt/rhel1
```

2) Logoff the iscsi target volumes, using the complete ‘iqn’ strings for your targets:

```
# iscsiadm -m node -T <iqn-rhel1> -p 10.10.5.10 --logout
Logging out of session [sid: 1, target: iqn.2001-
05.com.equallogic:0-8a0906-beee66602-27e79218c634bfec-rhel1,
portal: 10.10.5.10,3260]
Logout of [sid: 1, target: iqn.2001-05.com.equallogic:0-8a0906-
beee66602-27e79218c634bfec-rhel1, portal: 10.10.5.10,3260]:
successful
# iscsiadm -m node -T <iqn-snap> -p 10.10.5.10 --logout
Logging out of session [sid: 11, target: iqn.2001-
05.com.equallogic:0-8a0906-ae2e66602-b156e11d6b04c0e7-rhel1-
2010-06-08-13:16:18.6512, portal: 10.10.5.10,3260]
Logout of [sid: 11, target: iqn.2001-05.com.equallogic:0-
8a0906-ae2e66602-b156e11d6b04c0e7-rhel1-2010-06-08-
13:16:18.6512, portal: 10.10.5.10,3260]: successful
```

Note: The 'logout all' method could also be used. The following command will logout the server from all target volumes on the PS Series array with the group IP of 10.10.5.10

```
# iscsiadm -m node -p 10.10.5.10 --logoutall=all
```

3) Use the Group Manager GUI to take the volume 'offline' and delete both the volume (rhel1) and it's snapshot (rhel1-date-timestamp):
   - Select the volume (rhel1)
   - Volume -> Set offline
   - Volume -> Delete volume (this will also delete the snapshot)

4) Be sure to remove the appropriate mount entry line from /etc/fstab or the server will attempt to mount the missing volume's filesystem at reboot:

```
/dev/sdc      /mnt/rhel  ext3  _netdev,defaults      0 0
```

Task #5 - Configure multi-path I/O and verify

Enabling Multi-path I/O functionality on RHEL 5 involves the iSCSI initiator, the multipath subsystem and device mapper. This configuration procedure details the steps required to utilize multiple Ethernet interface connections between a RHEL 5.4 server and a PS Series storage array.

Before beginning the multi-path I/O configuration, it is recommended that existing iSCSI volumes be unmounted and the iscsiadm record database be empty. Be sure to complete the steps listed in "Task #4 – Delete the Volumes", above.

1) Verify there are no active sessions in the iscsi record database:

```
# iscsiadm --mode session
iscsiadm: No active sessions.
```

Note: If there are volumes listed, use the 'iscsiadm –o delete' command to manually delete the records. Use 'iscsiadm –-logoutall=all' to logout of all targets before manually deleting the records. Be sure to enter the appropriate iscsi targetname (-T <iqn>).

2) Create the interface files for the two Ethernet paths:

```
# iscsiadm --mode iface --interface eth2 -o new
New interface eth2 added
# iscsiadm --mode iface --interface eth3 -o new
New interface eth3 added
```

3) Add an interface name for the two Ethernet ports. In the test configuration, the interface names are 'eth2 and 'eth3':

```
# iscsiadm --mode iface --interface eth2 -o update --name
iface.net_ifacename --value=eth2
eth2 updated.
# iscsiadm --mode iface --interface eth3 -o update --name
iface.net_ifacename --value=eth3
eth3 updated.
```

These commands have created two iscsi interface files, the contents of which should look like these (eth2 and eth3):

```
# iscsiadm -m iface -I eth2
# BEGIN RECORD 2.0-871
iface.iscsi_ifacename = eth2
iface.net_ifacename = eth2
iface.ipaddress = <empty>
iface.hwaddress = <empty>
iface.transport_name = tcp
iface.initiatorname = <empty>
# END RECORD
# iscsiadm -m iface -I eth3
# BEGIN RECORD 2.0-871
iface.iscsi_ifacename = eth3
iface.net_ifacename = eth3
iface.ipaddress = <empty>
iface.hwaddress = <empty>
iface.transport_name = tcp
iface.initiatorname = <empty>
# END RECORD
```

4) Acquire the iscsi initiator name for your RHEL server:

```
# cat /etc/iscsi/initiatorname.iscsi
InitiatorName=iqn.1994-05.com.redhat:6891f663d4ce
```

5) At this point, the two Ethernet interfaces should be accessible to the iSCSI initiator software running on the RHEL server. In order to utilize both connection paths, the iSCSI volume must have appropriate iSCSI Access settings. Use the PS Group Manager to create a new volume with these settings:

- Volume name = mpio
- Size of 50G
- Access set to 'iscsi initiator' (use name from above 'cat' command)

6) Discover the new volume from the RHEL server:

```
# iscsiadm –m discovery –t sendtargets –p 10.10.5.10
10.10.5.10:3260,1 iqn.2001-05.com.equallogic:0-8a0906-
faae66602-9316e11da434c126-mpio
10.10.5.10:3260,1 iqn.2001-05.com.equallogic:0-8a0906-
faae66602-9316e11da434c126-mpio
```

Note: two iscsiadm records are listed for the 'mpio' volume. This is because there are two Ethernet interface records in the iscsiadm DB.

7) Login to the newly discovered volume:

```
# iscsiadm -m node -T <iqn> -p 10.10.5.10 --login
Logging in to [iface: eth2, target: iqn.2001-
05.com.equallogic:0-8a0906-faae66602-9316e11da434c126-mpio,
portal: 10.10.5.10,3260]
Logging in to [iface: eth3, target: iqn.2001-
05.com.equallogic:0-8a0906-faae66602-9316e11da434c126-mpio,
portal: 10.10.5.10,3260]
```

```
Login to [iface: eth2, target: iqn.2001-05.com.equallogic:0-
8a0906-faae66602-9316e11da434c126-mpio, portal:
10.10.5.10,3260]: successful

Login to [iface: eth3, target: iqn.2001-05.com.equallogic:0-
8a0906-faae66602-9316e11da434c126-mpio, portal:
10.10.5.10,3260]: successful
```

Note: that both logins succeeded, one on 'eth2' and the other on 'eth3'. Completion of this step verifies that the multiple connection paths between the RHEL server and the PS Storage array are active.

8) In order to allow multi-path I/O access from the RHEL server, the multipath software on the server needs to be configured. Check the installation status of the multipath package and install (if needed). Then start the multipathd process and startup at boot.

```
# yum list | grep multi
device-mapper-multipath.x86_64    0.4.7-34.el5_5.4   installed

# service multipathd start
# chkconfig multipathd on
```

9) Setup the server's multipath software for the PS Series array by editing the multipath configuration file and starting the multipath daemon. The multipath configuration file (/etc/multipath.conf) contains these sections: blacklist; defaults; multipaths and devices. Devices included in the blacklist section are excluded as multipath candidates. The default multipath configuration file blacklists all devices with a blacklist section that looks like this:

Note: with Oracle Linux Server only, you must: **cp /usr/share/doc/device-mapper-multipath-0.4.9/multipath.conf /etc** to get the default multipath.conf file.

```
# Blacklist all devices by default. Remove this to enable
# multipathing on the default devices.
#
blacklist {
        devnode "*"
}
```

The servers' multipath daemon should be enabled on the PS Series array, but not on the local scsi disk. In order to accomplish this, the blacklist section of the multipath configuration file needs to be modified. The test configuration had one direct attach scsi drive, /dev/sda. The World Wide Identifier (WWID) is a unique and persistent name for a scsi device. The WWID is used to specify this particular device in the multipath configuration file. The WWID for /dev/sda on the test configuration was acquired using this command:

```
# scsi_id -g -u -s /block/sda
3600508e0000000009d26af2c9525700d
```

Note: an alternative method to obtain the WWID is: "# ls –l /dev/disk/by-id". The output of which is shown here, with the relevant WWID in RED text:

```
# ls -l /dev/disk/by-id
total 0
```

```
lrwxrwxrwx 1 root root  9 Jul 13 12:43 scsi-
3600508e0000000009d26af2c9525700d -> ../../sda
lrwxrwxrwx 1 root root 10 Jul 13 12:43 scsi-
3600508e0000000009d26af2c9525700d-part1 -> ../../sda1
lrwxrwxrwx 1 root root 10 Jul 13 12:43 scsi-
3600508e0000000009d26af2c9525700d-part2 -> ../../sda2
lrwxrwxrwx 1 root root  9 Jul 13 12:44 scsi-
36090a0286066ee8aa9c124021ee11685 -> ../../sdb
```

Using this information, the blacklist section of the configuration file was changed to this, excluding the local scsi disk from multipathing:

```
blacklist {
   wwid 3600508e0000000009d26af2c9525700d
}
```

Note: Edit the blacklist section for your server, using either a regular expression or WWID to exclude the appropriate devices. Refer to the "Red Hat - DM Multipath Configuration and Administration" document listed in the Revision section of this document for more information on blacklisting devices.

10) Continue the multipath software configuration on the server. Edit the 'devices' section of the multipath configuration file. In the default multipath configuration file, the entire devices section is commented out.

Add the following lines for the PS Series array:

```
# BEGIN - PS Series device edits
devices {
        device {
              vendor "EQLOGIC"
              product "100E-00"
              path_grouping_policy    multibus
              getuid_callout "/sbin/scsi_id -g -u -s /block/%n"
              features "1 queue_if_no_path"
              path_checker readsector0
              failback immediate
              path_selector "round-robin 0"
              rr_min_io 10
              rr_weight priorities
        }
}
# END - PS Series device edits
```

11) Now that the initial edits to the multipath.conf file are complete, restart the multipath daemon:

```
# service multipathd restart
Stopping multipathd daemon:                                [  OK  ]
Starting multipathd daemon:                                [  OK  ]
```

12) The Linux device mapper dynamically names it's entries (e.g. dm-3, dm-5, …). This means that the device name for an iSCSI volume could vary across reboots, which presents an issue for consistently mounting the filesystem. Fortunately, the multipath subsystem provides a persistent naming scheme in which device mapper names can be consistently associated with an iSCSI volume.
Use the multipath command to identify the persistent name for the iSCSI volume:

```
# multipath -ll
mpath0 (36090a0286066eebeecbf34c61892e727) dm-3 EQLOGIC,100E-00
[size=50G][features=1 queue_if_no_path][hwhandler=0][rw]
\_ round-robin 0 [prio=2][active]
 \_ 19:0:0:0 sdc 8:32  [active][ready]
 \_ 18:0:0:0 sdd 8:48  [active][ready]
```

Note: the string highlighted in RED text is the WWID, a unique identifier and persistent name for this volume,

13) Edit the 'multipaths' section of the /etc/multipath.conf file to associate this WWID with a more user friendly name for this volume. Add these lines to your multipath configuration file, immediately after the 'devices' section that was added earlier. Of course, use the WWID for your iSCSI volume:

```
multipaths {
        multipath {
                wwid    36090a0286066eebeecbf34c61892e727
                alias   mpio
        }
}
```

Note: These lines will cause the device-mapper to consistently name the volume 'mpio, or more specifically '/dev/mapper/mpio'. This functionality requires that device mapper 'user friendly names' are enabled in the defaults section of the multipath.conf file. This is the default setting but you should confirm it by looking for the following setting in the multipath.conf file:

```
## Use user friendly names, instead of using WWIDs as names.

defaults {

        user_friendly_names yes

}
```

14) In order to make the multipath daemon recognize these configuration file changes, use the 'service reload' command:

```
# service multipathd reload
Reloading multipathd:                                  [  OK  ]
```

15) Run the 'multipath –v2' command and verify the persistent name entry in the /dev/mapper directory:

```
# multipath -v2
36090a0286066eebeecbf34c61892e727: rename mpath0 to mpio
```

```
: mpio (36090a0286066eebeecbf34c61892e727)  EQLOGIC,100E-00
[size=50G][features=1 queue_if_no_path][hwhandler=0][n/a]
\_ round-robin 0 [prio=2][undef]
 \_ 19:0:0:0 sdc 8:32  [active][ready]
 \_ 18:0:0:0 sdd 8:48  [active][ready]
```

```
# ls -l /dev/mapper
total 0
crw------- 1 root root  10, 63 Jun  4 14:21 control
brw-rw---- 1 root disk 253,  3 Jun 10 15:31 mpio
brw------- 1 root root 253,  0 Jun  4 14:22 VolGroup00-LogVol00
brw------- 1 root root 253,  1 Jun  4 14:21 VolGroup00-LogVol01
```

16) Verify the configuration using the 'dd' command. Test that an I/O load is distributed across the two Ethernet paths. Open the PS Group Manager and view the volume named mpio. Open the 'Connections' tab and verify that both Ethernet IP addresses are listed. Take note of the 'MB Read' column, which should be '0 MB' for each interface.
Execute this dd command:

```
# dd if=/dev/mapper/mpio of=/dev/null bs=4k count=1k
iflag=direct
1024+0 records in
1024+0 records out
4194304 bytes (4.2 MB) copied, 0.231391 seconds, 18.1 MB/s
```

This command moves 4MB through the iSCSI connection paths. Check the 'MB Read' column again in the PS Group Manager. Each IP address should read '2 MB', indicating that the traffic was split across the two connection paths.

Debug tips:

- Start with an empty iscsiadm database = no records
- Logout of targets and delete records
- Start with an empty device-mapper database
- multipath –F (if it fails, use 'dmsetup suspend', 'dmsetup remove')

Task #6 - Mount an MPIO volume using the persistent name

In the previous section, multi-path I/O was configured on the RHEL server. This involved changes to the iscsid and multipathd configurations. A persistent name was defined for the volume and the multipath connection was verified using the 'dd' command, which performs I/O operations on block devices. Now it's time to mount the 'mpio' volume as a filesystem.

Note: To ensure that iSCSI volumes being managed by device mapper are correctly and consistently associated with a mount point at boot time, a persistent name should be used. The procedure for defining a persistent name was covered in the previous task. In addition, the reader is referred to the Red Hat document "DM Multipath Configuration and Administration" for more complete coverage of the topic.

1) Configure the volume to be mounted at boot time by adding the following line to /etc/fstab:

```
/dev/mapper/mpio    /mnt/mpio    ext3    _netdev,defaults 0 0
```

2) Make a filesystem on the new volume. As stated before, the LVM tools could also be used to initialize the volume and assign the space to a logical volume:

```
# mke2fs -j -L mpio -v /dev/mapper/mpio
mke2fs 1.39 (29-May-2006)
Filesystem label=mpio
OS type: Linux
Block size=4096 (log=2)
Fragment size=4096 (log=2)
6557152 inodes, 13109760 blocks
655488 blocks (5.00%) reserved for the super user
First data block=0
Maximum filesystem blocks=4294967296
401 block groups
32768 blocks per group, 32768 fragments per group
16352 inodes per group
Superblock backups stored on blocks:
        32768, 98304, 163840, 229376, 294912, 819200, 884736,
1605632, 2654208,
        4096000, 7962624, 11239424
Writing inode tables: done
Creating journal (32768 blocks): done
Writing superblocks and filesystem accounting information: done

This filesystem will be automatically checked every 37 mounts
or 180 days, whichever comes first.  Use tune2fs -c or -i to
override.
```

3) Mount the volume and verify the volume size with the 'df' command. You could also reboot the server and verify that the volume is mounted correctly at boot time:

```
# mkdir /mnt/mpio

# mount -t ext3 /dev/mapper/mpio /mnt/mpio

# df
Filesystem       1K-blocks      Used Available Use% Mounted on
/dev/mapper/VolGroup00-LogVol00
                  50786940   3615548  44549952   8% /
/dev/sda1           101086     12538     83329  14% /boot
tmpfs              8216020         0   8216020   0% /dev/shm
/dev/mapper/mpio  51615928    184272  48809704   1% /mnt/mpio
```

Summary

This technical report described the configuration of a Red Hat Enterprise Linux server, Cent0S server, Oracle Linux server and a Dell EqualLogic PS Series storage array in both single-path I/O and multi-path I/O deployment environments. The test configuration was described and procedures for common storage administration tasks were provided.

TECHNICAL SUPPORT AND CUSTOMER SERVICE

Dell's support service is available to answer your questions about PS Series SAN arrays. If you have an Express Service Code, have it ready when you call. The code helps Dell's automated-support telephone system direct your call more efficiently.

Contacting Dell

Dell provides several online and telephone-based support and service options. Availability varies by country and product, and some services might not be available in your area.

For customers in the United States, call 800-945-3355.

Note: If you do not have access to an Internet connection, contact information is printed on your invoice, packing slip, bill, or Dell product catalog.

Use the following procedure to contact Dell for sales, technical support, or customer service issues:

1. Visit support.dell.com or the Dell support URL specified in information provided with the Dell product.

2. Select your locale. Use the locale menu or click on the link that specifies your country or region.

3. Select the required service. Click the "Contact Us" link, or select the Dell support service from the list of services provided.

4. Choose your preferred method of contacting Dell support, such as e-mail or telephone.

Online Services

You can learn about Dell products and services using the following procedure:

1. Visit www.dell.com (or the URL specified in any Dell product information).

2. Use the locale menu or click on the link that specifies your country or region