



Dell EqualLogic Best Practices Series

# **Sizing and Best Practices for Deploying Oracle 11g Release 2 Based Decision Support Systems with Dell EqualLogic 10GbE iSCSI SAN**

A Dell Technical Whitepaper

THIS WHITE PAPER IS FOR INFORMATIONAL PURPOSES ONLY, AND MAY CONTAIN TYPOGRAPHICAL ERRORS AND TECHNICAL INACCURACIES. THE CONTENT IS PROVIDED AS IS, WITHOUT EXPRESS OR IMPLIED WARRANTIES OF ANY KIND.

© 2012 Dell Inc. All rights reserved. Reproduction of this material in any manner whatsoever without the express written permission of Dell Inc. is strictly forbidden. For more information, contact Dell.

Dell, the DELL logo, and the DELL badge, PowerConnect™, EqualLogic™, PowerEdge™ and PowerVault™ are trademarks of Dell Inc. Broadcom® is a registered trademark of Broadcom Corporation. Intel® is a registered trademark of Intel Corporation in the U.S. and other countries. Microsoft®, Windows®, Windows Server®, and Active Directory® are either trademarks or registered trademarks of Microsoft Corporation in the United States and/or other countries.

# Table of Contents

1	Introduction .....	2
1.1	Audience .....	2
1.2	Terminology .....	2
2	Storage considerations for decision support systems.....	3
2.1	Benefits of EqualLogic 10GbE iSCSI arrays.....	3
2.2	Benefits of EqualLogic Multipath I/O.....	4
3	DSS I/O simulation .....	5
3.1	Test methodology.....	5
3.1.1	Simulation tool and workload definition.....	5
3.1.2	Measurement and monitoring .....	6
3.2	Test configuration .....	8
3.2.1	SAN connectivity .....	8
3.2.2	Storage volume layout .....	9
3.3	Test studies and results .....	10
3.3.1	PS6010XV RAID10 and RAID50 .....	10
3.3.2	PS6010XV and PS6510E .....	10
3.3.3	PS6010XV scaling studies .....	12
4	DSS application simulation .....	13
4.1	Test methodology.....	13
4.1.1	Simulation tool and workload definition.....	13
4.1.2	Measurement and monitoring .....	13
4.2	Test configuration .....	15
4.2.1	SAN connectivity .....	15
4.2.2	Storage volume layout .....	17
4.2.3	Oracle RAC configuration.....	18
4.3	Test studies and results .....	22
4.3.1	PS6010XV scaling studies .....	22
5	Best practices.....	27
5.1	Storage .....	27
5.2	SAN.....	27
5.3	Server.....	28
5.4	Operating system .....	28
5.5	Application – Oracle database.....	29
6	Sizing .....	30

6.1	Capacity .....	30
6.2	Performance.....	30
6.3	Scalable and modular deployment architecture .....	31
7	Conclusion .....	32
Appendix A	System .....	33
A.1	Hardware and software component versions.....	33
Appendix B	Orion.....	34
B.1	OS Disk volumes.....	34
B.2	Matrix test: Determine optimal parameter .....	34
B.3	Test run parameters.....	35
Appendix C	Oracle .....	36
C.1	Oracle ASM udev rules workaround .....	36
C.2	SQL*Plus statements for database tablespaces and files.....	37
C.3	SQL tuning .....	38

# Acknowledgements

This whitepaper was produced by the PG Storage Infrastructure and Solutions of Dell Inc.

The team that created this whitepaper:

**Ananda Sankaran, Chhandomay Mandal, and Camille Daily**

We would like to thank the following Dell team members for providing significant support during development and review:

**Keith Swindell and Chidambara Shashikir**

## Feedback

We encourage readers of this publication to provide feedback on the quality and usefulness of this information by sending an email to [SIFeedback@Dell.com](mailto:SIFeedback@Dell.com).



[SIFeedback@Dell.com](mailto:SIFeedback@Dell.com)

# 1 Introduction

This paper describes sizing and best practices for deploying Decision Support Systems (DSS) based on Oracle® 11g Release 2 database using Dell™ EqualLogic™ storage. DSS solutions require scalable storage platforms that offer high levels of performance and capacity. Deploying the storage platform using validated best practices, recommended settings and scalability guidelines for various system components (storage, server, switch, operating system, and application) helps ensure optimal operation of the entire DSS solution.

The Dell Storage Infrastructure and Solutions engineering lab conducted a series of tests using Dell PowerEdge™ servers, Dell PowerConnect™ network switches, and EqualLogic storage arrays. From those tests, best practices were validated and scalability guidelines have been noted for designing storage architectures that are best suited to support the storage demands of DSS applications running on Oracle 11g Release 2 databases.

The objectives of the tests were:

- Determine the best practices and sizing guidelines for deploying Oracle based DSS applications.
- Determine the I/O throughput of the storage using different RAID configurations and array models via I/O simulation.
- Determine the scalability of the I/O throughput via I/O simulation as storage arrays are added.
- Determine the scalability of the I/O throughput via DSS application simulation as storage arrays are added, while ensuring that the overall configuration is balanced with no resource bottlenecks existing on server.

## 1.1 Audience

This paper is intended for solution architects, storage network engineers, system administrators, and IT managers designing and deploying Oracle 11g Release 2 based DSS solutions on PowerEdge servers and EqualLogic storage.

## 1.2 Terminology

Decision Support System (DSS): a broad class of information processing applications used to support data analysis and decision making activities. Business Intelligence (BI), Online Analytical Processing (OLAP), and Data Warehousing (DW) systems provide decision support functionality and require similar designs. Therefore, the design and configuration best practices presented in this paper could apply to those scenarios as well.

## 2 Storage considerations for decision support systems

DSS applications are designed to support complex analytical query activities using very large data sets. The user queries executed on a DSS database typically take several minutes (or even hours) to complete and require processing large amounts of data. A DSS query is often required to fetch millions of records from the database for processing. To support these queries, the database management system (DBMS) reads table data from the storage devices. The DBMS typically uses a sequential table scan, fetching numerous records from a table with a single request. The resultant read I/O to storage consists of large I/O blocks (approximately 512KB or 1MB in size). The large I/O requests require large I/O throughput rates from storage to the database server to provide optimal performance.

In addition to the significant I/O throughput required, the large data sets that are common to DSS user queries also require substantial processing resources (CPU and RAM), therefore the database server must be provided with sufficient processing and memory resources to accomplish the processing of the raw query results into the desired useable data that the user intends the query to produce.

The large I/O patterns and processing necessary in DSS queries warrant careful system design to ensure that the performance requirements are met by the storage arrays as well as the server and interconnect fabric. The overall performance of a DSS solution is determined by the performance characteristics of each component in the system. These include database and operating system settings, server resources, SAN design and switch settings, storage Multipath I/O (MPIO) software, storage resources, and storage design.

### 2.1 Benefits of EqualLogic 10GbE iSCSI arrays

EqualLogic 10GbE iSCSI storage arrays offer the high I/O throughput that is required for typical DSS applications. To set up the iSCSI SAN in our test configuration, we created a 10 gigabit Ethernet connection fabric using PowerConnect 8024F 10GbE switches as well as EqualLogic PS6010XV and PS6510E storage arrays. The arrays were configured with redundant multi-port 10GbE controllers.

Information about EqualLogic 10GbE PS6010 and PS6510 series storage arrays is available on [equallogic.com](http://www.equallogic.com):

<http://www.equallogic.com/products/ps6010-series.aspx?id=8945&slider6010=1>

Information about PowerConnect 8024F switches is available on [dell.com](http://www.dell.com):

<http://www.dell.com/downloads/global/products/pwcnt/en/switch-powerconnect-8024f-spec.pdf>

Information about EqualLogic PS6010XV storage arrays is available on [dell.com](http://www.dell.com):

<http://www.dell.com/downloads/global/products/pvaul/en/equallogic-ps6010xv-specsheet.pdf>

Information about EqualLogic PS6510E storage arrays is available on [dell.com](http://www.dell.com):

<http://www.dell.com/downloads/global/products/pvaul/en/equallogic-ps6510e-specsheet.pdf>

## 2.2 Benefits of EqualLogic Multipath I/O

Dell provides numerous value added software products as part of the EqualLogic solution. One of these offerings, Host Integration Tool for Linux<sup>®</sup> (HIT/Linux), provides efficient iSCSI session establishment and management to each target volume in Linux environments. In a storage pool consisting of multiple EqualLogic arrays, EqualLogic's group manager software automatically distributes the volume across multiple members. These distributed portions of the volume are referred as slices. HIT/Linux maintains volume layout information, which is used to route I/O requests intelligently to the correct volume slice thereby improving performance. This helps throughput centric workloads like DSS to obtain higher throughput rates by setting up appropriate iSCSI sessions to a volume's member arrays. In addition, network ports of all member arrays are efficiently utilized in parallel for I/O requests. HIT/Linux also provides path failover capabilities during NIC or switch failures.

More information on HIT/Linux can be found in the document titled, "Configuring and Deploying the Dell EqualLogic Host Integration Toolkit for Linux."

<http://www.equallogic.com/WorkArea/DownloadAsset.aspx?id=10311>



## 3 DSS I/O simulation

### 3.1 Test methodology

#### 3.1.1 Simulation tool and workload definition

A series of I/O simulations were conducted using Oracle I/O Numbers (Orion) to understand the performance characteristics of the storage arrays. Orion is a popular I/O simulation tool used for understanding the I/O performance behavior of storage systems for Oracle databases. The tool and documentation can be downloaded from Oracle's website at <http://www.oracle.com/technetwork/topics/index-089595.html>. The tool is installed on the operating system (Red Hat® Enterprise Linux® in this case) and executed to simulate I/O on the external EqualLogic storage arrays. The EqualLogic storage volumes were presented to the OS as block I/O devices.

I/O simulations were conducted with large sized I/O block requests simulating a DSS database I/O to understand the read and write I/O throughput rate offered by the storage arrays. The measured I/O simulation rate provided a throughput baseline for DSS tests conducted later with the actual Oracle database installed. Four types of large block I/O workload were chosen for simulation:

- Sequential Read (SR) I/O in concat mode and 256KB I/O block size

Sequential read I/O streams were simulated on the disk volumes with an I/O block size of 256KB. The disk volumes were simulated in a concat mode treating them together as a single large concatenated volume. This test determines the read I/O throughput of the array(s) with a sequential I/O workload.

- Sequential Read (SR) I/O in RAID 0 mode and 1MB I/O block size

**Note:** Details on Oracle topics can be found at docs.oracle.com. Specific links are listed in the section titled, "Additional resources" at the end of this document.

Sequential read I/O streams were simulated on the disk volumes with an I/O block size of 1MB as DSS workload is typically represented by large block size I/O. 1MB is the default maximum I/O block size used by Oracle for Red Hat Enterprise Linux (RHEL) 5.5, the OS used in this project. This setting is controlled by the database initialization parameter *DB\_FILE\_MULTIBLOCK\_READ\_COUNT*.

The disk volumes were simulated in RAID 0 mode by Orion, treating them as individual volumes and striping the I/O requests across them. This volume configuration simulates I/O behavior when Oracle Automatic Storage Management (ASM) is configured to manage disk volumes. Oracle recommends using ASM for managing OS disk volumes due to the advanced management capabilities. Since EqualLogic arrays provide robust disk redundancy via external RAID, this test configuration was set up to only utilize volume striping (i.e RAID 0 across volumes) with no volume level redundancy or volume mirroring. The Orion stripe depth in the tests run on each volume was set at 1MB which is the default stripe size used by ASM for Oracle data files. It defines how much data is stored on one volume before moving onto the next volume in the RAID 0 volume striping scheme. The eventual I/O block size to storage is determined by the *DB\_FILE\_MULTIBLOCK\_READ\_COUNT* parameter and defaults to 1MB. The

volume stripe will be split into 1MB chunks for storage access. Since the volume stripe size is 1MB in our configuration, it would fit within one I/O block. If volume stripe size is set at 4MB per volume, there will be 4 I/O blocks per volume level stripe.

This test determined the read I/O throughput of the array(s) in an ASM like environment with a sequential I/O workload. This provides a baseline for sequential read workloads like backup and table scans.

- Sequential Write (SW) I/O in RAID 0 mode and 1MB I/O block size

Sequential write I/O streams were simulated on the disk volumes with an I/O block size of 1MB. The disk volumes were utilized in RAID 0 with a stripe depth of 1MB to match ASM behavior. The intent of this test was to determine the write I/O throughput of the arrays in an ASM like environment with a sequential I/O workload. This provides a baseline for sequential write workloads such as backup, bulk, or batch updates to table data.

- Large Random Read (RR) I/O in RAID 0 mode and 1MB I/O block size

Large random read I/O streams were simulated on the disk volumes with an I/O block size of 1MB. The disk volumes were utilized in RAID 0 with stripe depth of 1MB to match ASM behavior. The intent of this test was to identify large random read I/O throughput of the arrays in an ASM like environment. DSS database environments exhibit an I/O pattern of random read I/O with large blocks due to the multiple users submitting large running queries to the system. Even though each user's query generates table scans involving sequential I/O, a collection of queries from multiple users makes the resulting I/O more parallel and random. Hence this test provides a baseline for a random read workload like DSS which was simulated as part of the study and described in later sections of this paper.

Initially an Orion matrix type test was conducted for read I/O workloads to determine the optimal outstanding I/O parameter (*num\_large*) that maximizes I/O throughput. After determining the optimal outstanding I/O parameter, the same was used to run Orion point tests. The Orion tool parameters used for simulating each of the above workloads is given in Appendix B . Visit the Oracle website at [http://docs.oracle.com/cd/E11882\\_01/server.112/e16638/iodesign.htm#BABEIFFG](http://docs.oracle.com/cd/E11882_01/server.112/e16638/iodesign.htm#BABEIFFG) for details on Orion parameters.

### 3.1.2 Measurement and monitoring

The test duration was set at 20 minutes per I/O simulation run using Orion. During each I/O simulation test, different metrics were monitored and captured by the following tools:

- Orion

I/O throughput rate in Megabytes per second (MBps) as part of Orion standard test result report was captured.

- EqualLogic SAN HeadQuarters (SAN HQ)

EqualLogic storage array statistics for various objects including network, disk, I/O performance, and I/O capacity were captured. A dedicated server was used to run this tool.

- Custom Network Monitoring Script

EqualLogic storage array network statistics including TCP retransmissions and network throughput in MBps were captured via a custom script using host scripting toolkit. This monitoring script polled the storage arrays every 60 seconds for network statistics. A dedicated server was used to run this tool.

The criteria established for TCP retransmissions by the arrays was under 0.5%. TCP retransmissions during all tests remained under this established limit.

- Server OS Tools

The Red Hat Linux *ifconfig* utility was used for capturing iSCSI NIC statistics (RX and TX bytes). Red Hat Linux *IOstat* and *vmstat* utilities were used for capturing disk, CPU, and memory statistics respectively.

The logical test bed setup (including the monitoring stations) are shown in Figure 1.

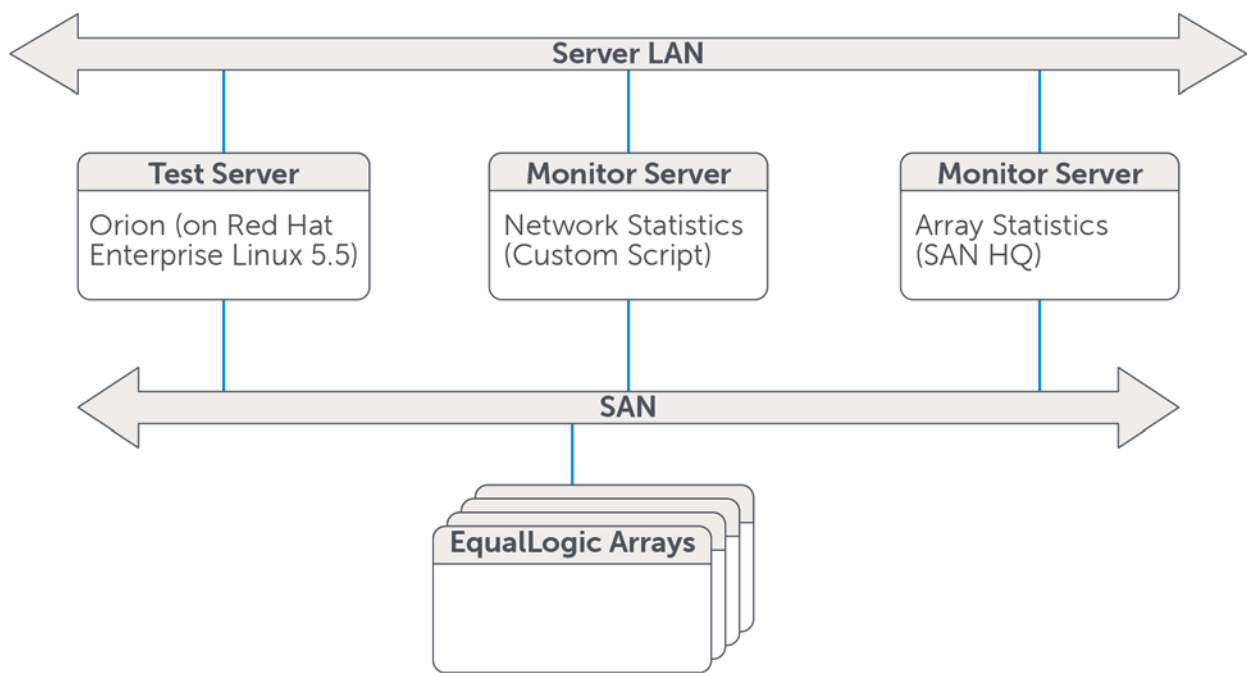
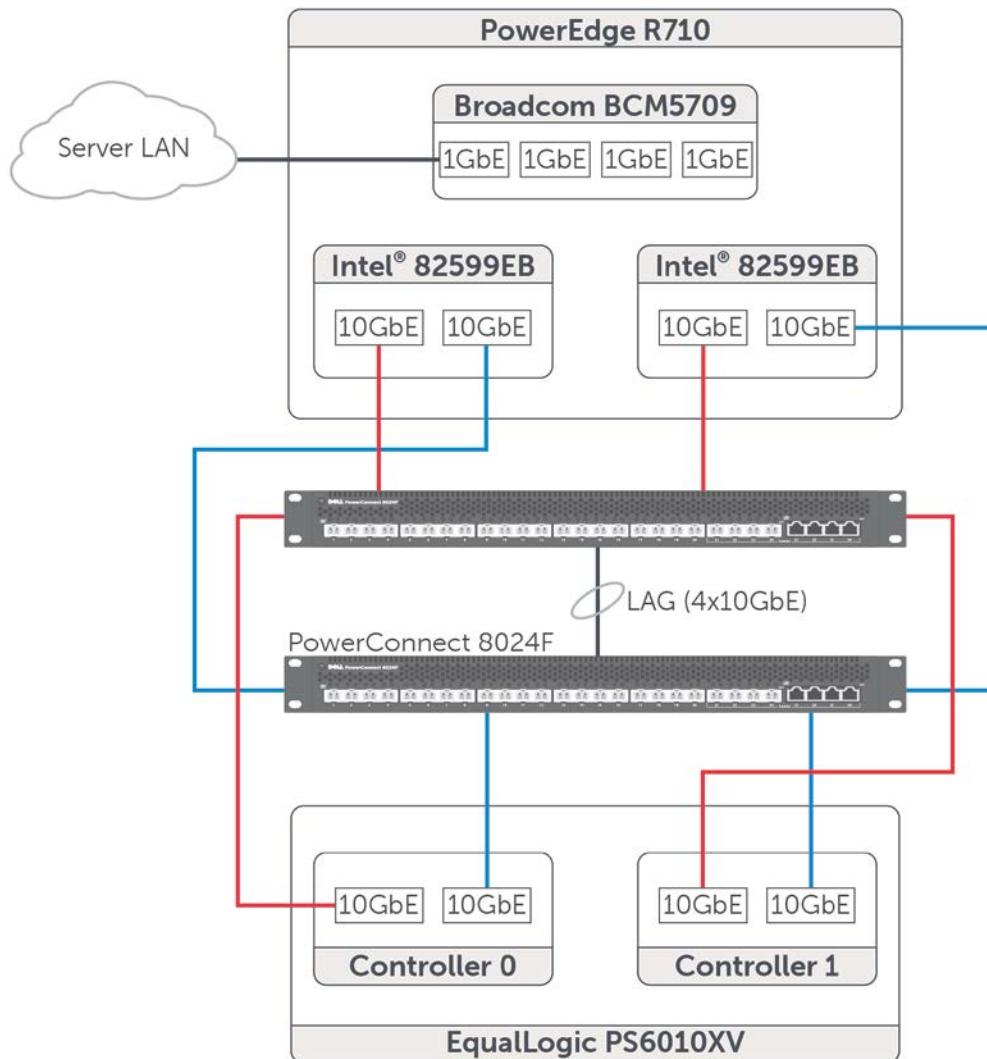


Figure 1 Orion logical test bed setup

## 3.2 Test configuration

### 3.2.1 SAN connectivity

Server connectivity to EqualLogic storage arrays and the details of the iSCSI network are illustrated in Figure 2.



**Figure 2** Test server and storage connectivity

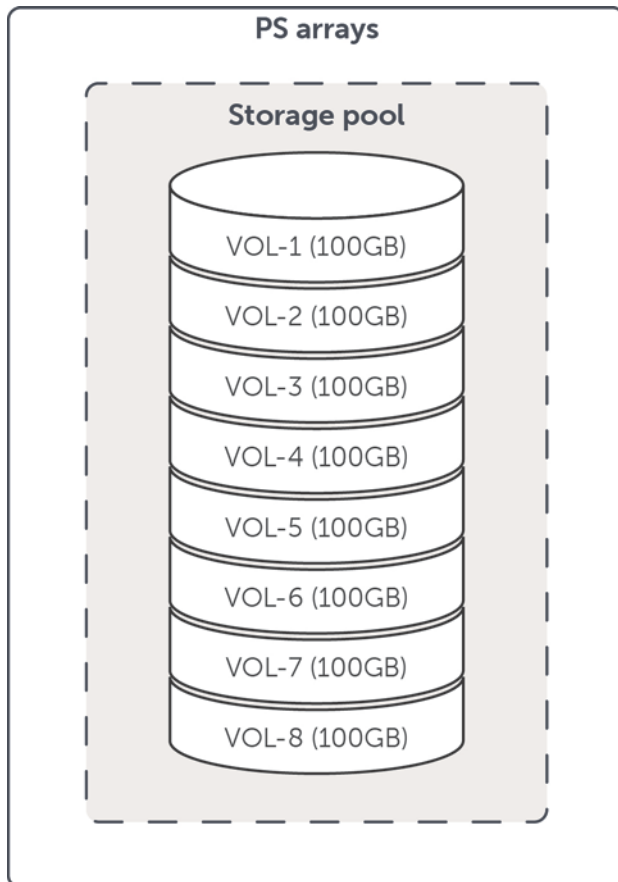
A PowerEdge R710 server was installed with two Intel<sup>®</sup> 82599EB 10 Gigabit Ethernet controllers with 2 x 10GbE SFP+ ports each. These 10GbE ports were used for iSCSI connectivity to the SAN. One port from each controller was used in configurations with up to two storage arrays. All four ports were used in configuration with more than two arrays to ensure that there is sufficient bandwidth available on the server NICs as arrays are scaled. One of the on-board Broadcom NetXtreme II BCM5709 Gigabit Ethernet Quad NIC ports was used for server LAN connectivity. The hardware and driver/firmware versions are listed in Appendix A of this document.

Two PowerConnect 8024F switches were interconnected using a Link Aggregation Group (LAG) via Link Aggregation Control Protocol (LACP). The LAG comprised of four 10GbE ports from each switch. These 2 switches comprised the SAN fabric. The iSCSI network ports from each server NIC and

storage controller were connected to these two switches as shown in Figure 2. One port from each controller was connected to these switches to avoid the switch as the single point of failure (SPOF).

### 3.2.2 Storage volume layout

The configuration of volumes in EqualLogic storage is shown in Figure 3.



**Figure 3** EqualLogic storage volume configuration

Eight EqualLogic storage volumes were created at 100GB size each in a single storage pool. This was kept constant across all Orion test configurations. The number and/or type of EqualLogic PS arrays within the storage pool were changed across configurations. The volumes were recreated to keep the test configuration consistent as the storage arrays were changed.

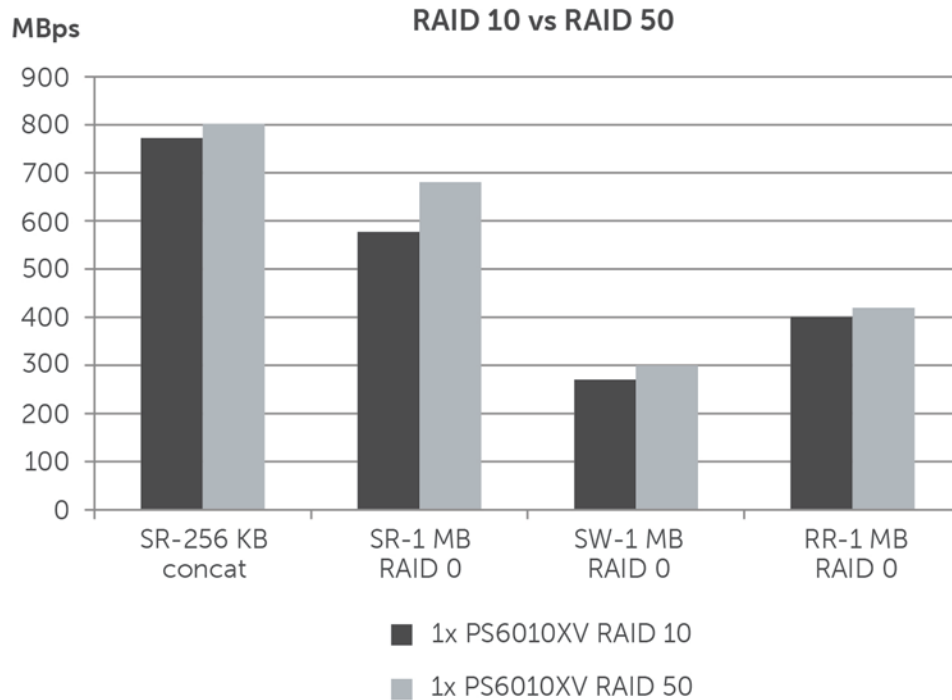
To enable advanced MPIO functionality, the EqualLogic HIT/Linux was installed on the RHEL server. The default connection parameters of two iSCSI sessions per volume slice on an array and a maximum of six iSCSI sessions for the entire volume were used. Also the default load balancing policy and I/Os per path were used for the MPIO. The utility *eqtune* installed as part of the HIT/Linux package was executed and the OS parameter recommendations from the utility were adhered to.

### 3.3 Test studies and results

This section describes the Orion tests conducted and the results observed.

#### 3.3.1 PS6010XV RAID10 and RAID50

The four Orion workloads described earlier were executed on an EqualLogic PS6010XV array in RAID 10 and in RAID 50. The PS6010XV included 16 x 600GB 15K RPM SAS drives, with two drives being configured as hot spares in the RAID layout. The I/O throughput reported by Orion is shown in Figure 4.

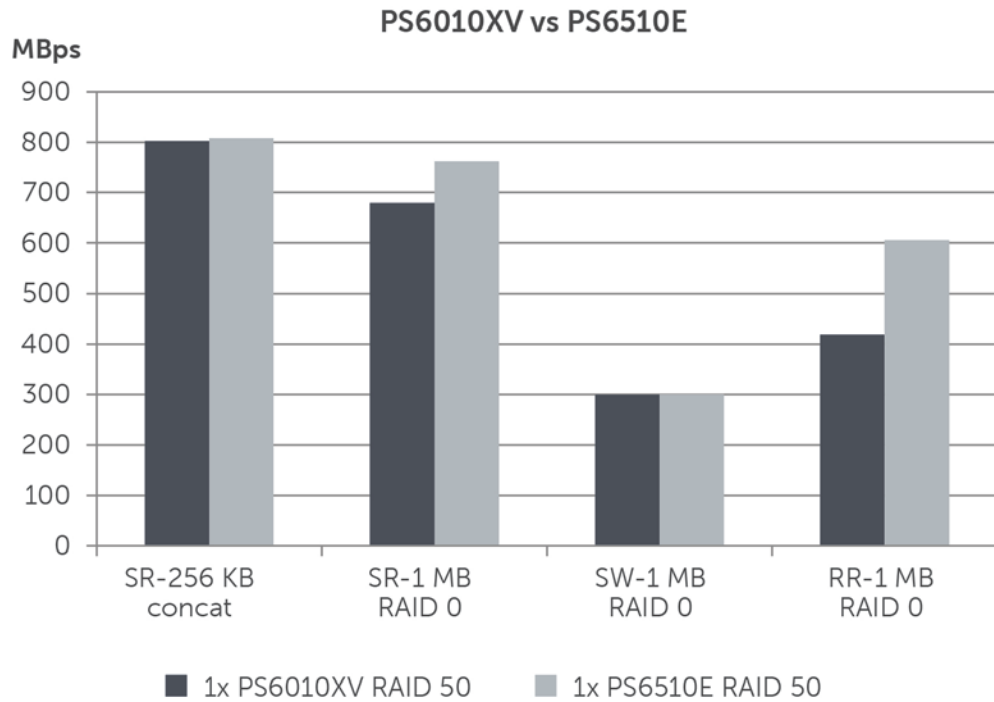


**Figure 4** Orion I/O throughput: RAID 10 versus RAID 50

The RAID 50 provided slightly better throughput than the RAID 10 configuration. This behavior can be primarily attributed to respective RAID implementation differences such as more disks hosting active data than redundant data in RAID 50 which can improve read I/O performance. Throughput increased by 3% with the sequential read test in the concat mode. Using RAID 0 mode, there was a 17% increase with sequential reads, 10% more with sequential writes, and 4% more for large random reads.

#### 3.3.2 PS6010XV and PS6510E

The four Orion workloads were executed on EqualLogic PS6010XV and PS6510E arrays in RAID 50. The PS6010XV included 16 x 600GB 15K RPM SAS drives, with 2 drives being configured as hot spares in the RAID layout. The PS6510E included 48 x 2TB 7.2K RPM SATA drives, with 2 drives being configured as hot spares in the RAID layout. The I/O throughput reported by Orion is shown in Figure 5.



**Figure 5 Orion I/O throughput: PS6010XV versus PS6510E**

PS6510E offered the same levels of throughput as PS6010XV with sequential reads in concat mode and sequential writes using RAID 0 mode. The sequential reads in RAID 0 mode provided 12% more throughput than PS6010XV. With large random reads in RAID 0 mode, the PS6510E provided about 44% more throughput. This increase with random reads provided by the PS6510E array can be attributed to more physical drives servicing the large block I/O requests resulting in significant throughput improvement.

### 3.3.3 PS6010XV scaling studies

The four Orion workloads were executed on EqualLogic PS6010XV arrays using RAID 50. The storage pool was configured with one array initially and then scaled up to four arrays in the same storage pool. Test volumes were recreated and mounted on a server for each scaled configuration. The I/O throughput reported by Orion across the scaled configurations is shown in Figure 6.

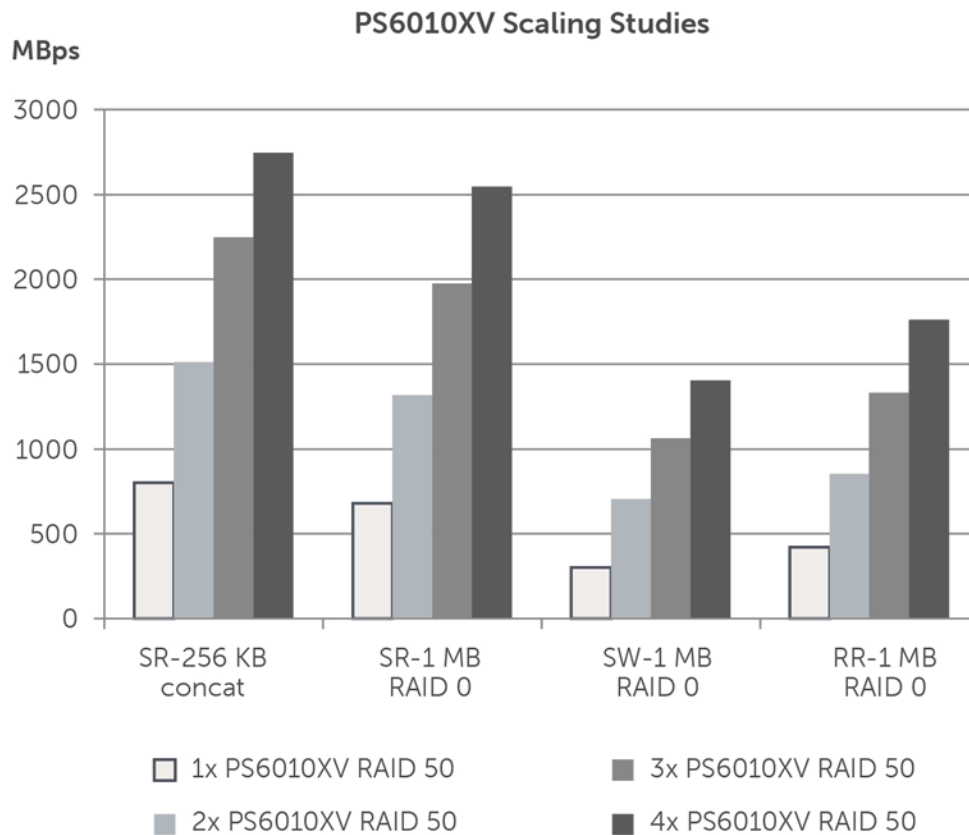


Figure 6 Orion I/O throughput: PS6010XV versus PS6510E

With sequential reads in concat mode, the throughput scaled at 188%, 280% and 342% respectively with 2, 3 and 4 arrays as compared to one. With sequential reads in RAID 0, the throughput scaled at 193%, 290% and 374% respectively. With 2 and 3 arrays sequential reads scaled almost linearly and with four arrays it was slightly less than linear. By default the capacity load balancer distributes a volume across three array members and uses more than three only when there is not sufficient space for the volume within the three arrays. However, all members within the pool are used for volumes with any one volume spanning only three arrays.

With sequential writes in RAID 0 mode, the throughput scaled at 235%, 354% and 468% respectively. With large random reads in RAID 0 mode, the throughput scaled at 203%, 317% and 420% respectively. In both cases, the throughput significantly increased. This increase can be attributed to additional physical drives and array resources servicing the large block I/O requests resulting in throughput improvement. The random read throughput can be considered as the I/O baseline for DSS application tests conducted with Oracle database installed.



## 4 DSS application simulation

### 4.1 Test methodology

#### 4.1.1 Simulation tool and workload definition

A series of DSS application simulations were conducted using Benchmark Factory<sup>®</sup> for Databases by Quest Software<sup>®</sup> as the simulation tool running a TPC-H like workload. TPC-H is an industry standard benchmark for DSS. The DSS tests were conducted to understand the throughput behavior of the storage arrays when the Oracle Real Application Cluster (RAC) database executed DSS queries.

Benchmark Factory was installed on a standalone client system to execute the workload on the database server over the network. During execution, DSS SQL queries via simulated database users were submitted to the Oracle RAC database via the server LAN. The database processes the queries and sends the results back to the simulated client users. For processing these queries, the database creates I/O requests that are predominantly large block sized read requests. The database server was configured with volumes from the external EqualLogic storage arrays. The EqualLogic storage volumes were presented to the server OS (RHEL 5.5) as block I/O devices. Oracle ASM was configured to manage the OS devices as ASM disk groups and the database was configured to use the ASM disk groups for storage of database files.

We conducted DSS simulations generating large I/O block requests to understand the read throughput rate (MBps) offered by the storage arrays under a database workload. The DSS dataset workload was created using:

- A total database size of approximately 130GB,
- A set of five database users running DSS queries simultaneously, and
- Each database user submitting a series of 22 standard DSS queries to the database.

#### 4.1.2 Measurement and monitoring

During DSS database query simulations, different metrics were monitored and captured by the following tools:

- Benchmark Factory

Total and average query completion times for users were captured. This measurement is important since the workload involves long running queries processing large amounts of data.

- EqualLogic SAN HeadQuarters (SAN HQ)

EqualLogic storage array statistics for various objects including network, disk, I/O performance, and I/O capacity were captured. A dedicated server was used to run this tool.

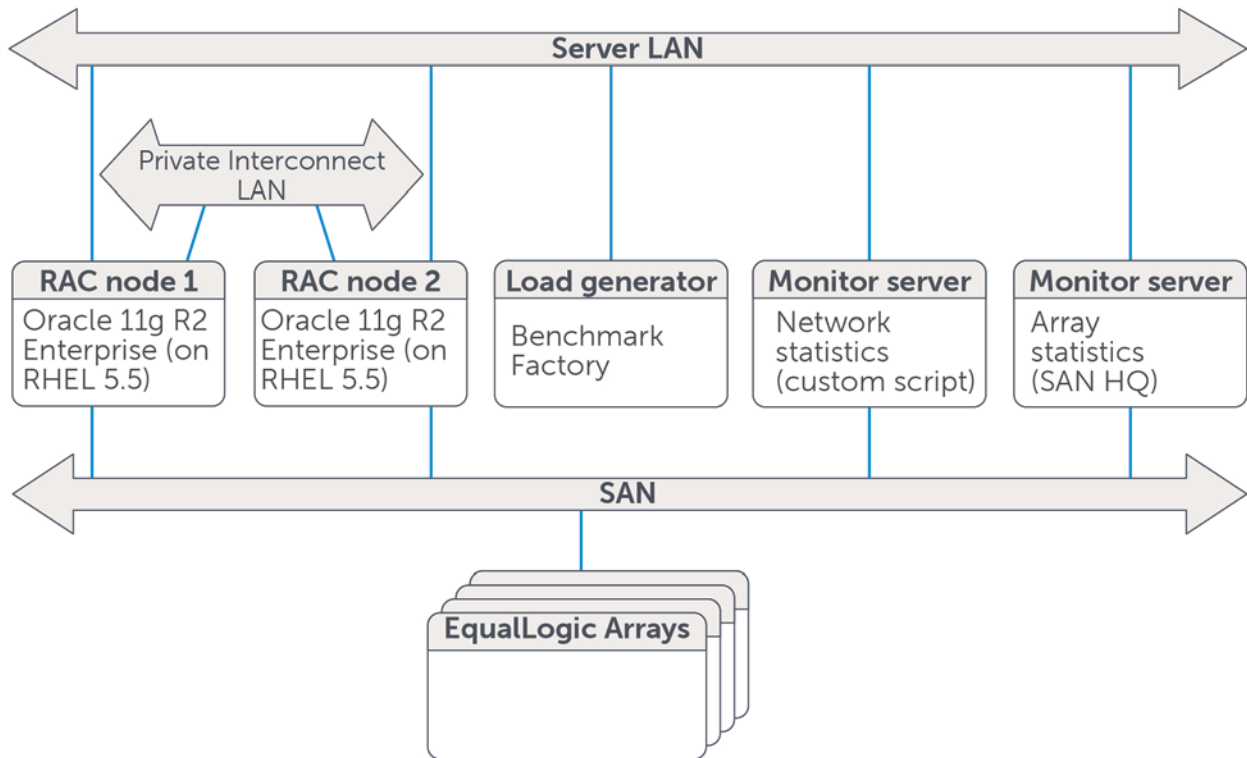
- Custom Network Monitoring Script

EqualLogic storage array network statistics including TCP retransmissions and network throughput in MBps were captured via a custom script using host scripting toolkit. This monitoring script polled the storage arrays every 60 seconds for network statistics. A separate monitoring server was used to run this tool. The criteria established for TCP retransmissions by the arrays was under 0.5%. TCP retransmissions during all tests remained under this established limit.

- Server OS Tools

Red Hat Linux *ifconfig* utility was used for capturing iSCSI NIC statistics (RX and TX bytes). Red Hat Linux *iostat* and *vmstat* utilities were used for capturing disk, CPU and memory statistics respectively. A server CPU utilization level of 75% was set as the threshold for DSS test runs.

The logical test bed for DSS simulations including monitoring stations are shown in Figure 7.

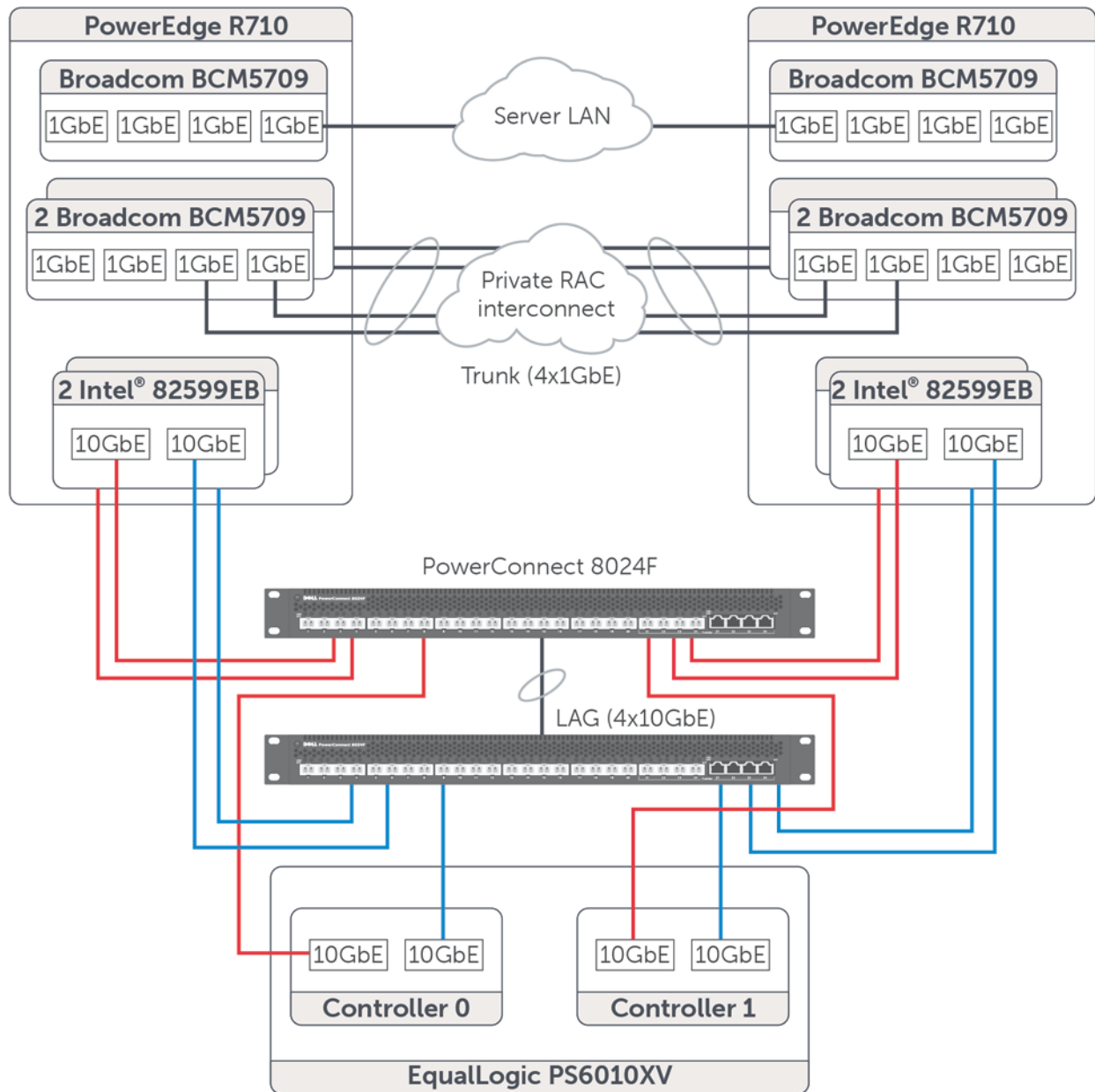


**Figure 7** Benchmark Factory logical test bed setup

## 4.2 Test configuration

### 4.2.1 SAN connectivity

The server connectivity to EqualLogic storage arrays and the details of the iSCSI network are illustrated in Figure 8.



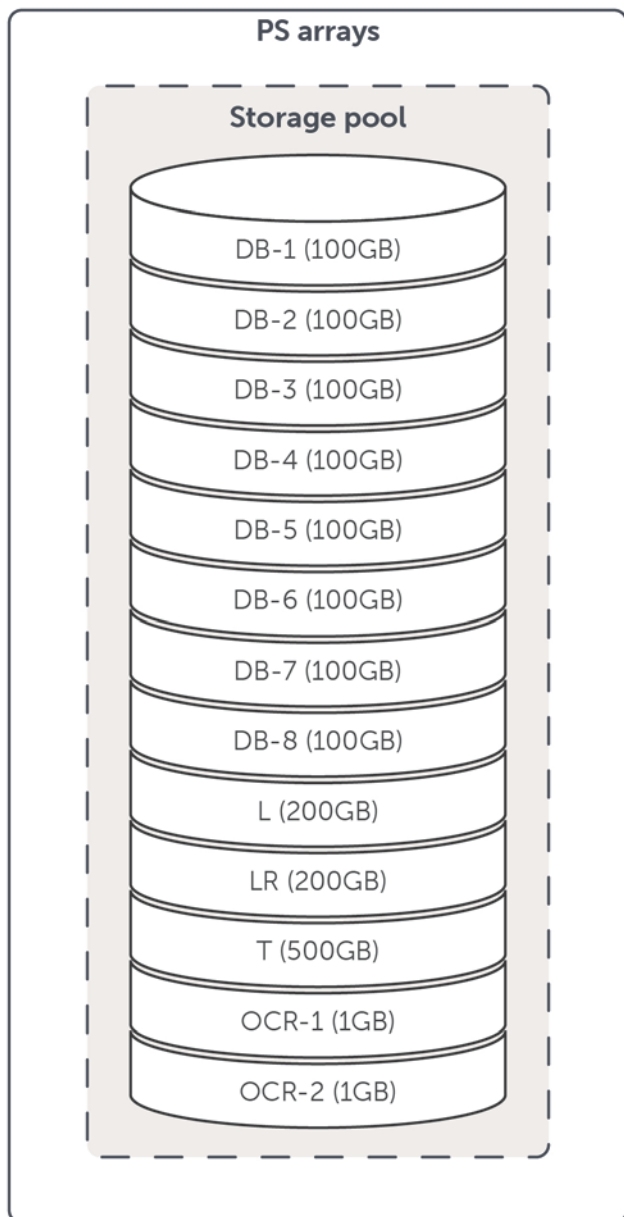
**Figure 8** Test server and storage connectivity

- Two PowerEdge R710 Servers were installed with two Intel® 82599EB 10 Gigabit Ethernet Controllers with 2 x 10GbE SFP+ ports each. These 10GbE ports were used for iSCSI connectivity to the SAN. One port from each NIC controller on a server was used in test configurations. The hardware and driver/firmware versions are listed in Appendix A .
- One of the on-board Broadcom NetXtreme II BCM5709 Gigabit Ethernet Quad Port NICs was used for server LAN connectivity.

- Two NetXtreme II BCM5709 Gigabit Ethernet Quad Port NICs were installed per server for private RAC interconnect. Two ports from each quad port NIC were used to connect to a private interconnect switch, for a total of four ports per server. The four interconnect ports were configured as a bond.
- Two PowerConnect 8024F switches were interconnected using a LAG via LACP. The LAG comprised of four 10GbE ports from each switch. These two switches comprised the SAN. The iSCSI network ports from each server NIC controller and each storage controller from the array were connected to these two switches in a highly available manner as show in Figure 1. One port from each controller was connected to these switches to avoid the switch as the single point of failure (SPOF).

## 4.2.2 Storage volume layout

The configuration of volumes in EqualLogic storage is shown in Figure 9.



**Figure 9** EqualLogic storage volume configuration

- Eight EqualLogic storage volumes were created at 100GB each in a single storage pool for database data. Two volumes were created at 200GB each for database logs and one volume at 500GB was created for temporary database files. Two additional volumes were created at 1GB each for Oracle Clusterware components. This was kept constant across all DSS test configurations. Only the number of EqualLogic PS arrays comprising the storage pool was changed across configurations. The volumes were recreated whenever the storage arrays were changed in the configuration.
- EqualLogic HIT/Linux was installed on the database servers running RHEL 5.5 to provide EqualLogic aware MPIIO to these eight volumes. Default parameters were used including the default connection parameters of two iSCSI sessions per volume slice on an array and a

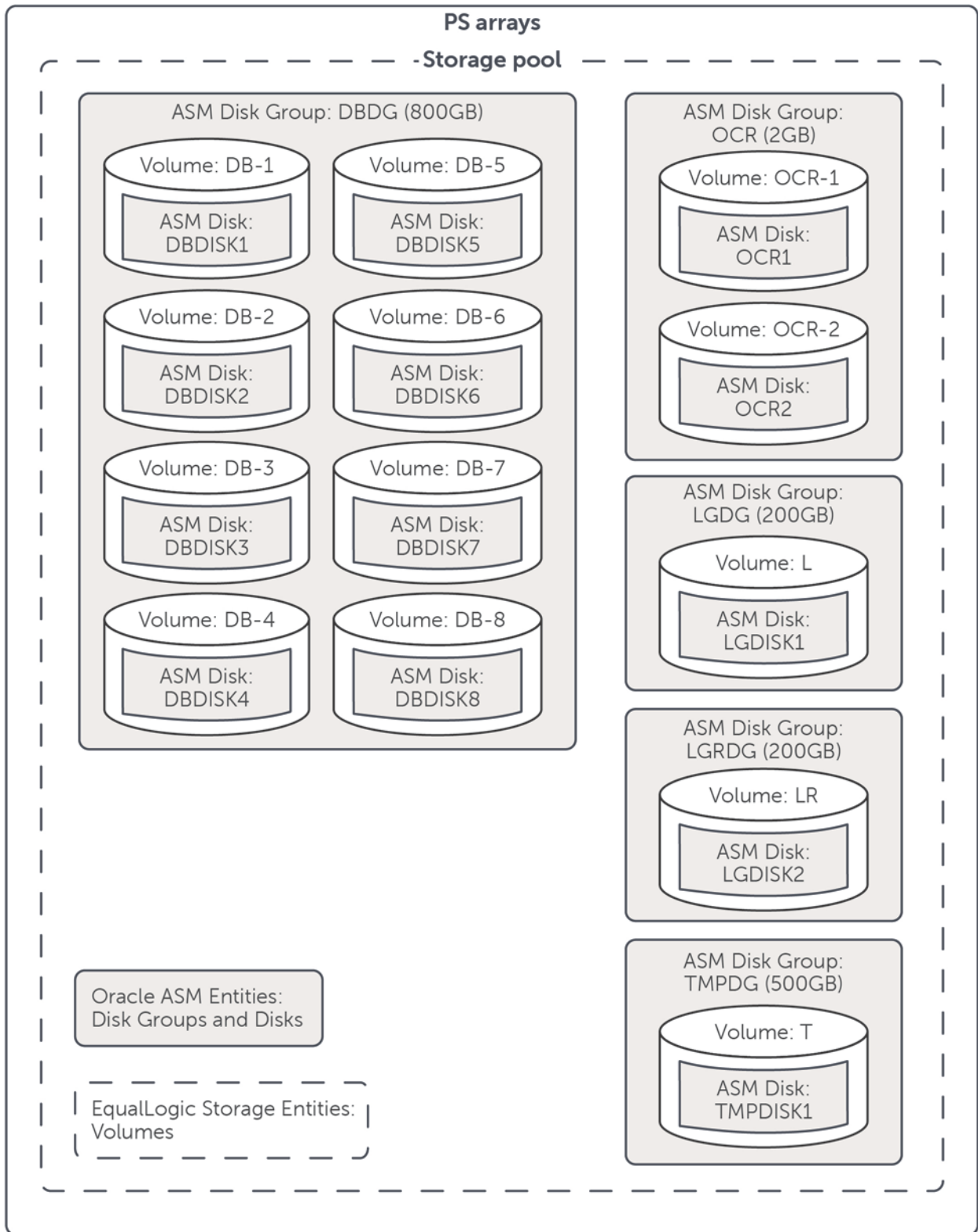
maximum of six iSCSI sessions for the entire volume. The utility *eqltune* installed as part of the HIT/Linux was executed, and the OS parameter recommendations from the utility were adhered to.

### 4.2.3 Oracle RAC configuration

- Oracle Clusterware was installed and configured on the two PowerEdge R710 servers running RHEL. This enabled them to function as two nodes of an Oracle RAC.
- The storage volumes shown in Figure 9 were mounted on the OS as block devices accessed via the EqualLogic HIT/Linux MPIO. Oracle ASM was configured to manage and utilize these OS disk volumes.

**Note:** A custom Linux *udev* rules file was needed to create Linux DM device handles for EqualLogic multipath devices in a path where the Oracle ASM utility will locate them. More information on this is listed in Appendix C .

The ASM Disk and Disk Groups were created following appropriate best practices applicable to Oracle RAC deployments. Separate ASM Disk Groups were created for database data, logs, temporary files and RAC Clusterware components. These Disk Groups were backed by dedicated disk volumes. Configuration details are shown in Figure 10.



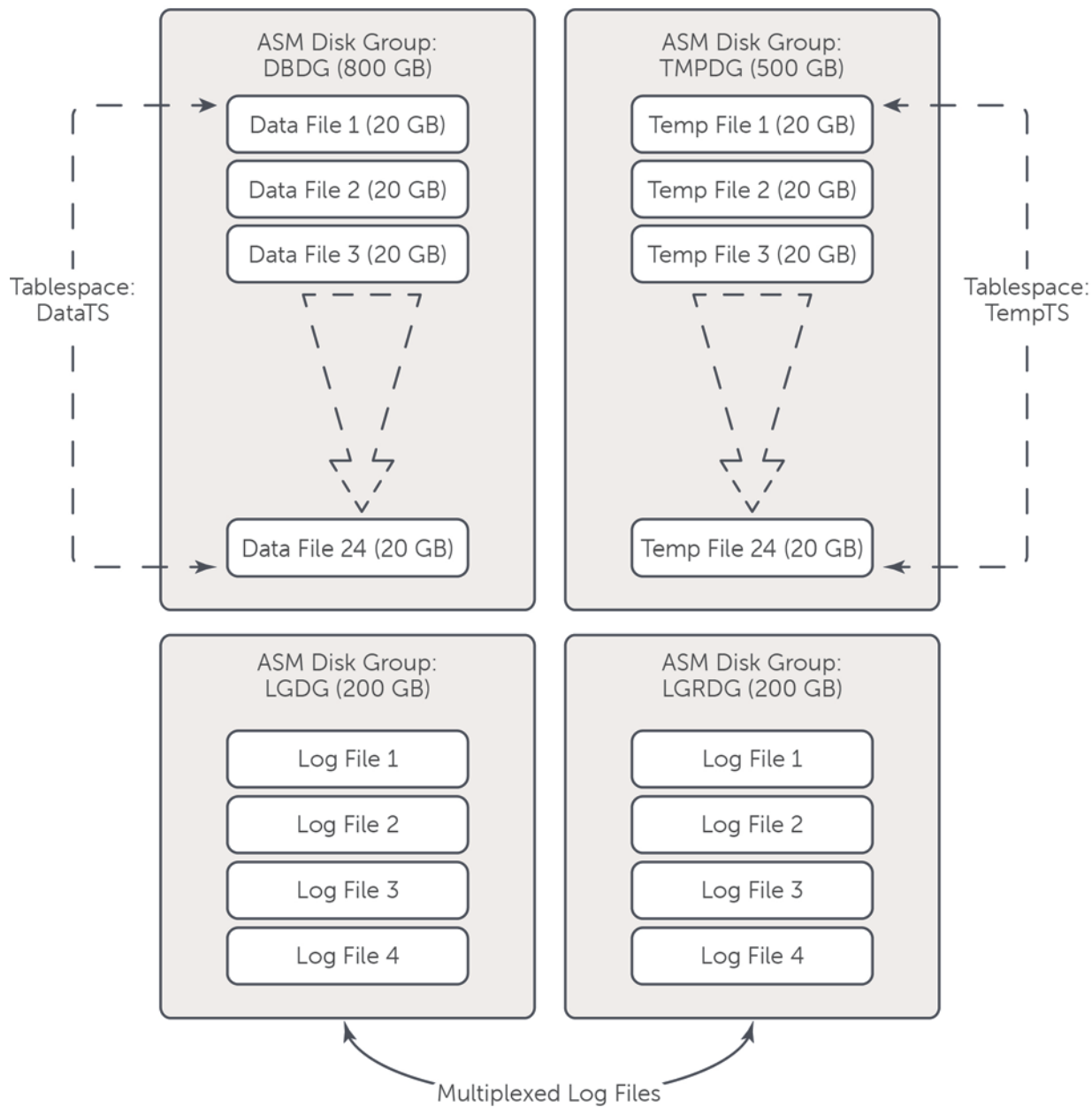
**Figure 10** ASM disk and disk group configuration

- Each EqualLogic storage volume was logically mounted and mapped as an ASM Disk as shown in Figure 10. Five ASM Disk Groups were created:
  - Disk Group with eight 100GB volumes for hosting database data files and total size of 800GB
  - Disk Group with one 200GB volume for hosting log files and total size of 200GB
  - Disk Group with one 200GB volume for hosting multiplexed log files and total size of 200GB
  - Disk Group with one 500GB volume for temporary database files and total size of 500GB
  - Disk Group with two 1GB volumes for hosting Clusterware components – voting disk and Oracle Cluster Registry (OCR). Total size of 2GB.
- Once the ASM Disk Groups were created, the cluster database with two nodes as two database instances was installed with the following parameters:
  - Total memory size for database instance was set as 48,284MB and automatic memory tuning was enabled to maximize and efficiently utilize available server memory.
  - The number of processes per database instance was increased to 1,100 to support larger parallel processing needs.
  - Client connection mode was specified as *dedicated server mode* to support few users running long running queries.
  - The log files were configured to be multiplexed across the two ASM disk groups created for database logs.
  - Other database initialization parameters were left with default values.

After cluster database installation, tablespaces for hosting the database data files and temporary database files were created. 24 data files of size 20GB each were created on the data tablespace and 24 data files of size 20GB each were created on the temporary tablespace. 24 data files were created for each tablespace to match the number of logical processors available on each RAC node. The system temporary tablespace was also moved to the new temporary tablespace created. The tablespace and file layout of the database is shown in Figure 11. The SQL statements used for creating the tablespaces and files via SQL\*Plus utility is given in Appendix C .



### Tablespace and File Layout



**Figure 11 Database tablespace and file layout**

Once the user schema, tablespaces and data files were created, the schema objects including tables and indices were created using Benchmark Factory by submitting SQL queries via the network to the database server. The *create database* feature of the tool enabled this object creation and also the population of the tables/indices with test data.

## 4.3 Test studies and results

### 4.3.1 PS6010XV scaling studies

DSS query simulations, with the workload described in “Test methodology” on page 13, were executed on the EqualLogic PS6010XV array. As demonstrated by the Orion test studies in the earlier section, RAID 50 offered slightly higher throughput than RAID 10 and also provides higher useable capacity than RAID 10. Therefore, RAID 50 was chosen due to the balanced capacity and performance levels offered by this RAID type. The PS6010XV was chosen over PS6510E for the scalability studies, even though PS6510E offered higher throughput in the Orion tests. The primary reason for this choice is that the PS6010XV offers a good balance in cost, capacity, and performance levels needed for DSS applications with finer granularity to achieve the required capacity needs of the application. Larger capacity requirements would drive choosing a PS6510E instead.

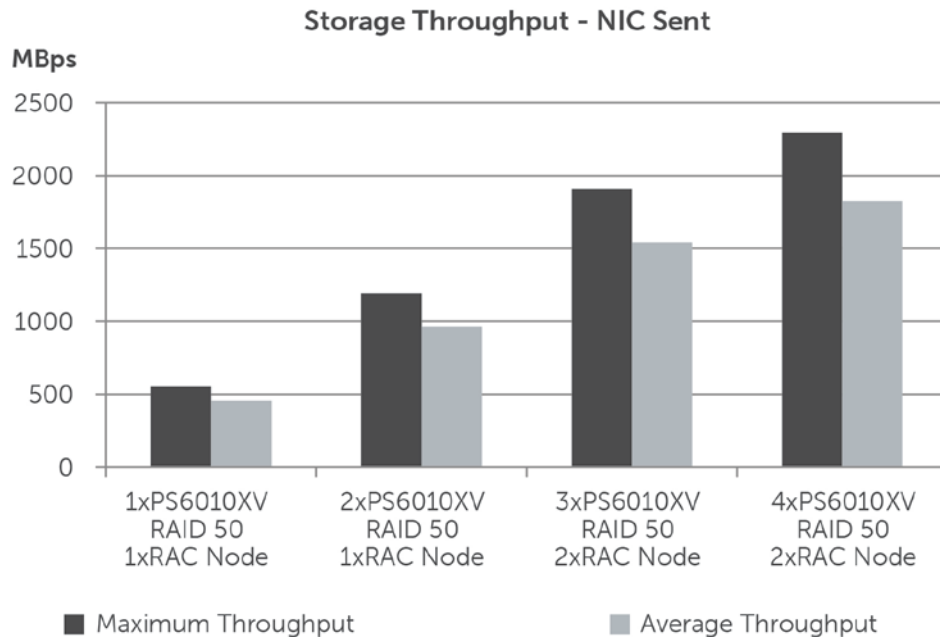
First the simulation was executed on one array in a single storage pool with the volume layout, Oracle ASM layout, and database layout described in “Test configuration” earlier. The number of arrays was then increased to two, three and four within the same storage pool and the same DSS test rerun respectively on each configuration. Only one of the Oracle RAC nodes was used in configurations with 1 and 2 arrays. The database instance on the second node was shutdown. With three and four arrays, both the RAC nodes along with their database instances were used for the test run. This is because during the initial test runs, one server instance was sufficient to drive I/O throughput for up to two arrays within established criteria for CPU utilization (under 75%).

For each storage configuration, the storage volumes, ASM Disks/Disk Groups were re-created. The cluster database was reinstalled; the tablespaces, files, and user schema were recreated. Benchmark Factory was rerun to recreate schema objects and fill tables/indices with test data. It must be noted that one of the 22 user queries in the DSS test required custom SQL tuning for improved execution. The remaining 21 queries ran with the default optimization provided by Oracle 11g Release 2 optimizer. Custom tuning details are described in Appendix C .

The various metrics monitored during the DSS tests when arrays were scaled from one to four arrays are described in the following sections.

#### Storage I/O throughput

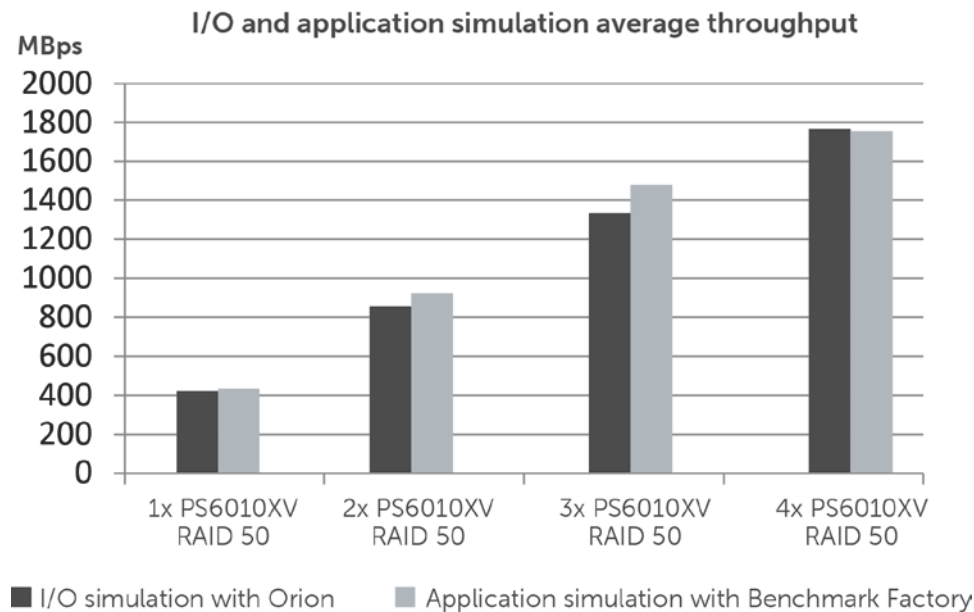
The storage I/O throughput measured during the DSS tests with one to four arrays is shown in Figure 12 below. The I/O profile of the database during DSS tests was large block sized I/O read requests which were random in nature. Both the maximum and average throughput measured is shown. The throughput was measured as the rate of data sent by the array NICs in MBps.



**Figure 12 Storage I/O throughput – PS6010XV scalability**

As seen in the bar graph, both the average and maximum throughput measured scaled linearly with the number of arrays. In the two and three array cases, it was better than linear scalability. The maximum throughput scaled at 216%, 346% and 415% with two, three, and four arrays as compared to a single baseline array. The average throughput scaled at 212%, 338% and 401% with two, three, and four arrays as compared to a single baseline array. This I/O throughput scaling is possible because of EqualLogic’s scale-out architecture – an array adds not only disks but also controllers and NIC ports.

The I/O throughput measured during DSS test runs using the actual Oracle database matched well with the random read I/O test conducted with Orion as shown in Figure 13.



**Figure 13 Storage I/O throughput – I/O (Orion) and application simulation**

The I/O throughput across storage array members as reported by EqualLogic SAN HQ in the four array configuration is shown in Figure 14.

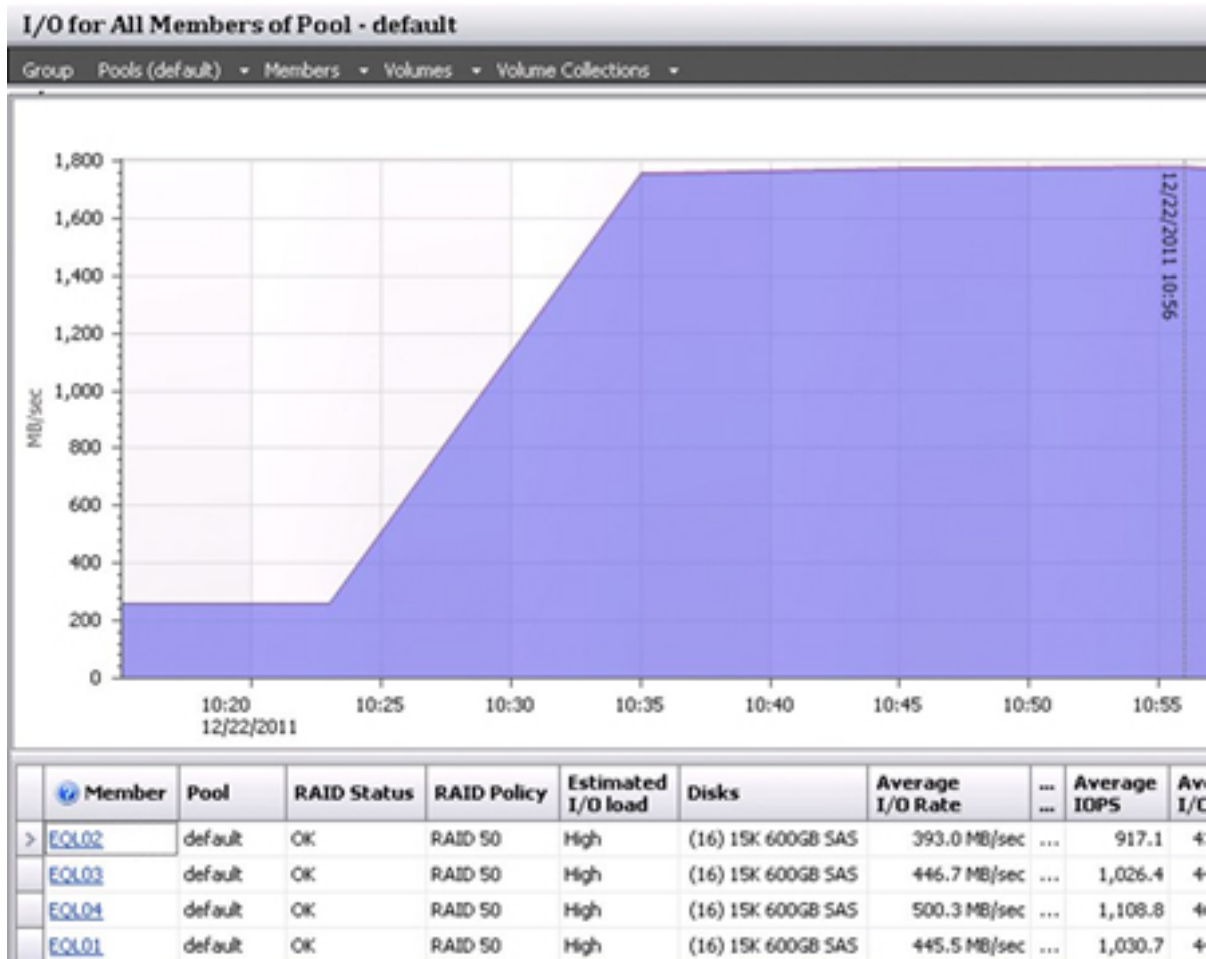
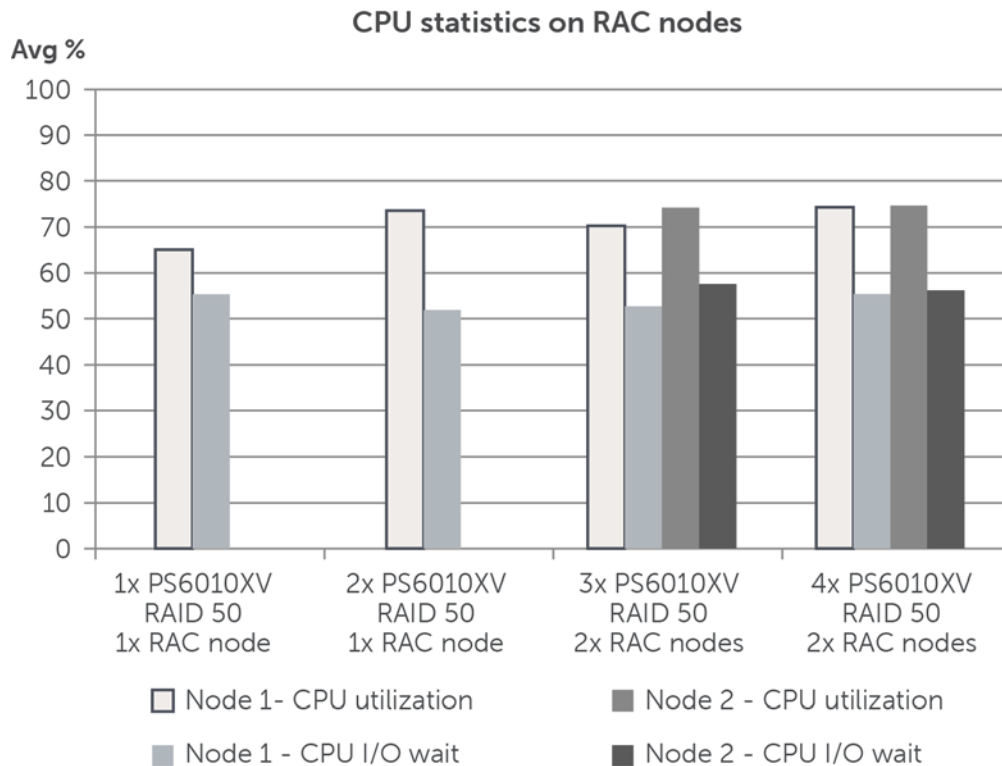


Figure 14 Storage I/O throughput – SAN Headquarters

The initial ramp up in Figure 14 is during the start of the test when all 5 users login and start submitting their queries. As the users login, the database server commences query processing and generates increasing I/O requests resulting in a ramp up of I/O throughput. After this initial ramp up period, the I/O throughput achieved stabilizes at a constant level based on I/O throughput available from storage.

## CPU Statistics on RAC nodes

The average CPU utilization percent and percent I/O wait time were measured during the DSS tests and are shown in Figure 15. Only one RAC node database instance was used with one and two arrays. With three and four arrays both RAC node instances were used. This is because with one and two arrays, the average CPU utilization on a single instance was below the threshold of 75%. With three and four arrays, two RAC nodes were needed due to the CPU utilization constantly above 90% on a single node.

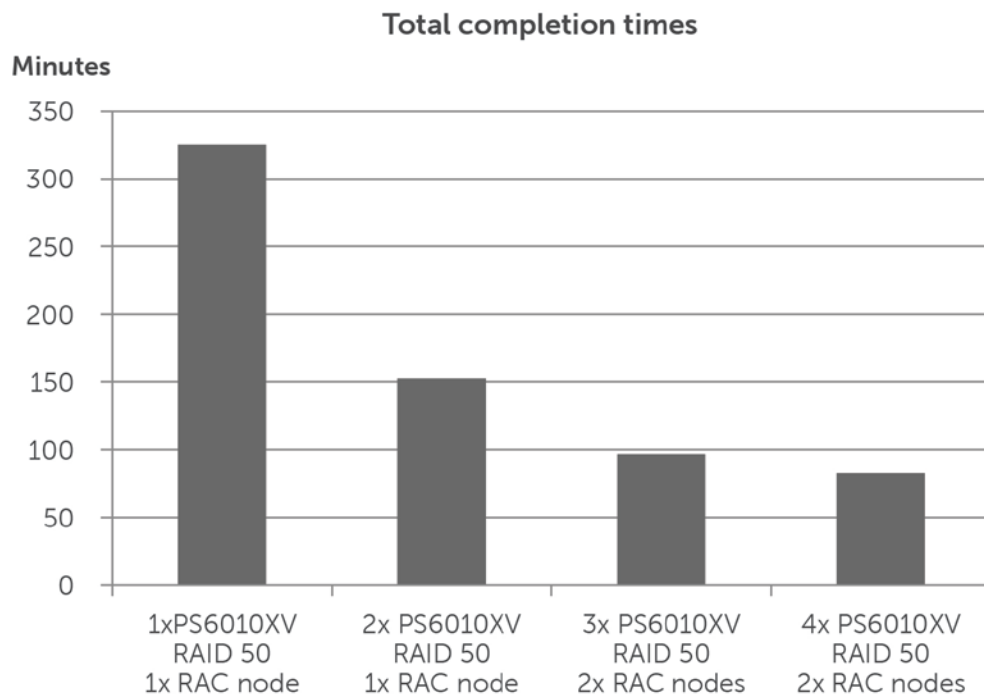


**Figure 15** CPU statistics on RAC nodes

The average CPU utilization was under 75% across all tests. % I/O wait time occupied a significant portion of the CPU utilization. I/O wait time was at 70% to 85% of the CPU utilization. This indicated that the database server reached I/O limits with sufficient CPU resources available. This supported the objective of these tests to characterize the storage I/O throughput impacts to the database workload without any resource bottlenecks on the server.

## Query Completion Time

The end-to-end completion time for all 22 queries submitted by the 5 simulated users is shown in Figure 16 below. This is the duration between the time when the first query was submitted to the database and when the last query was completed.



**Figure 16** Query streams completion time

Due to the increased storage throughput available as the number of arrays increased, data was fetched at a much faster rate by the database. The queries were completed faster resulting in decreased completion times with increasing arrays. With two arrays, the completion time was 46% less than with a single array. With three arrays, it was 29% and with four it was 25%. The nonlinear decrease in completion time can be attributed to other database processing factors including workload characteristics and database size.

# 5 Best practices

Best practices apply to each of the component layers in the solution stack. The considerations at each layer of the solution stack are described in the following sections.

## 5.1 Storage

- It is an important best practice to physically run DSS workloads on a separate storage system that is optimized for DSS query I/O activity. This will prevent DSS workloads from interfering with the operation of other workloads.
- The storage arrays must be configured to support capacity and performance requirements. Considerations include array controller capabilities, I/O port specifications, RAID policies, hard drive types and spindle speeds.

**Drives:** For maximum throughput with the best capacity granularity, SAS 15K drives are recommended because they offer the highest level of disk I/O throughput. SAS 10K drives (EqualLogic PS6010X) or SATA drives (EqualLogic PS6010E) could also be considered for environments requiring lower levels of workload storage I/O throughput. Alternately, the 48 drive array models PS6510E and PS6510X offer higher throughput levels due to larger drive count. Choice of drive size depends on the capacity requirements of the application workload.

**RAID:** RAID level 50 offers a good balance between capacity utilization and performance levels and is recommended for DSS workloads. Since DSS environments are highly read intensive, RAID 5 can also be considered if you want to optimize for capacity utilization. However, the larger RAID 5 layout compared to the RAID 50 layout can lead to significantly longer RAID rebuild times when drives fail. RAID 10 is recommended for environments that include a mix of both read and write operations such as batch database updates.

**Storage Pool:** DSS workloads are primarily read intensive and typically generate minimal transaction log activity. The transaction log volumes can be hosted on the same pool and disk drives as the data volumes. EqualLogic storage arrays provide redundant hardware features and RAID levels to ensure data protection. Typical database environments also include data backup implementations to periodically backup data to secondary storage. Environments requiring higher levels of data protection and recoverability, as determined by individual IT department best practices, can consider separating data and log volumes to separate disk drives and pools. If the temporary database incurs heavy I/O based on workload needs, then it can be moved to a separate pool backed by dedicated drives.

## 5.2 SAN

- Use redundant switch paths between servers and storage arrays. Redundant switch paths in the SAN will ensure uninterrupted access to storage controllers in the event of a switch failure or scheduled switch downtime for maintenance.
  - NIC ports installed on the server for iSCSI SAN connectivity should be equally distributed across the redundant switches.
  - Ethernet ports from each controller of the storage array should be equally distributed across the redundant switches.
  - The two switches must be interconnected with appropriate interswitch bandwidth based on the total I/O bandwidth requirement between the servers and storage arrays plus I/O workload characteristics.

- A separate switching infrastructure for other non-SAN connections should be in place to isolate the iSCSI SAN switching infrastructure exclusively for storage traffic
- Enable flow control on all iSCSI SAN switch ports (including LAG ports) for optimal throughput of iSCSI traffic. Flow control will help regulate packet flow across SAN fabric between nodes (server, switches, and storage) during high I/O traffic conditions.
- Enable jumbo frames on all ports (including LAG) to support larger MTU frame size (9000 bytes). Jumbo frames enable DSS workloads to perform optimally by enabling large data transfers.
- Disable Spanning Tree Protocol (STP) on all switch ports connecting to end devices like server NICs and storage network ports. Instead, PortFast setting (for PowerConnect /Cisco switches, and similar functionality in other switches) is recommended on those ports to enable quick initialization of iSCSI sessions.
- Select appropriate physical interconnects for Ethernet connectivity. Use either transceivers along with optical cables or copper cables for interconnecting switches and end devices to the switches. Optical connectivity will support longer distances.

### 5.3 Server

- Hardware configuration to meet performance and capacity requirements is critical. In general, the Dell provided tools for sizing server resources for Oracle DSS deployments are recommended. The Dell Oracle Advisor tool is available at [www.dell.com/oracle](http://www.dell.com/oracle).
- Besides CPU and Memory, it is recommended that a server system with appropriate PCIe technology and required number of PCIe lanes is selected to provide for the large bandwidth required for data transfers with DSS workloads.
- Dedicate specific network interface cards on the server for iSCSI storage connectivity. Appropriately size the number of NIC ports used for iSCSI connectivity to storage, based on the throughput requirements of the workload. Also, to provide redundant paths in a high availability configuration, at least two physical network cards are recommended for iSCSI connectivity. Install and provision separate network cards for server side LAN traffic.

### 5.4 Operating system

- Enable flow control and jumbo frames on the Server NIC ports used for connecting to storage and match those settings with your switch and array controller port settings. Jumbo frames provide a more efficient transport for large data transfers typically generated DSS workloads.
- Since MPIO will be in use, disable NIC teaming or bonding on all ports used for iSCSI connections. Also, to simplify system management and reduce the chances of making errors during system configuration, disable any unused NIC ports on the system.
- It is recommended to utilize EqualLogic aware MPIO functionality provided by the HIT kit for the server operating system. HIT/Linux can be obtained from EqualLogic support at [www.equallogic.com/support](http://www.equallogic.com/support). Configuring EqualLogic aware MPIO provides enhanced performance benefits for bandwidth intensive applications like DSS. It also enables easier configuration by automating setup of multiple required iSCSI sessions (connections) to the target volumes.
- It is recommended to run the utility *eqltune* included with the HIT/Linux and follow the OS specific parameter recommendations listed by the utility.



## 5.5 Application – Oracle database

- Storage volumes need to be appropriately sized for capacity and performance to host the data, transaction logs, and temporary database files. Split the database into at least five volumes. The temporary database will need to be appropriately provisioned with storage space, since DSS workloads often create large temporary data objects during query processing. Consider separating log volumes into their own EqualLogic storage pool backed by dedicated arrays/disk drives if significant log activity is incurred (i.e. batch updates) or if log recoverability is critical. Similarly, consider separating temporary database volumes into a separate storage pool based on capacity and performance needs. I/O performance for different database objects can be optimized based on which RAID level is used:

**Database Volumes:** DSS workloads are read intensive, therefore RAID 50 is recommended to optimize for both capacity and performance. RAID 5 can also be used for capacity optimization since it offers similar performance compared to RAID 50, but with lower data protection.

**Temporary database file volumes:** Use RAID 50 to optimize both capacity and performance if write activity is low. If write activity is high, RAID 10 is recommended instead.

**Log Volumes:** DSS workloads have relatively low log activity except during database refreshes. Therefore, RAID 50 is recommended to optimize for both capacity and performance. If DSS DB refreshes include very large bulk updates, then RAID 10 is recommended.

- Database layout on storage volumes is an important factor for optimal operation of DSS workloads.
  - Map storage volumes to individual ASM disks and create separate ASM Disk Groups for database data files, temporary data files and transactions logs.
  - Create a separate tablespace for database data files and temporary database files, ensuring hosting on respective ASM disk groups.
  - Ensure the number of files created both in the data and temporary tablespaces is at least equal to the number of logical CPUs on the server node.
  - Isolate the DSS database under consideration to its own tablespace and ASM disk group backed by physically separate storage arrays; not shared with other databases.
- Use Oracle provided features and utilities to improve query processing performance.
  - Use query tuning utilities such as SQL Tuning Advisor provided with Oracle 11g Release 2 Enterprise Manager to tune query execution plans and improve execution efficiency.
  - Consider table partitioning schemes where it can improve throughput via parallel processing.

## 6 Sizing

DSS application databases need to be sized for both capacity and performance.

### 6.1 Capacity

The useable capacity of an array depends on the drive size, drive count, and RAID scheme chosen. The PS6010XV array tested has a useable capacity of 6.28 TB with RAID 50. Refer to respective array specification documents at [equallogic.com](http://equallogic.com) for useable capacity of other array models.

To determine the array count based on capacity needs:

1. Determine the capacity needed for database, logs and temporary data files.
2. Determine the storage pool layout.
  - a. All three above database entities hosted on same storage pool,
  - b. Each hosted on separate storage pool, or
  - c. A combination of two hosted on one pool and remaining one on another pool.
3. For each storage pool:
  - a. Determine the total pool capacity needed by adding the individual capacity of database entities in the pool.
  - b. Add a 20% overhead capacity factor per pool.
  - c. Determine the number of arrays based on capacity needed and RAID chosen.
  - d. Ensure the arrays chosen offer required performance per pool as described in "Performance."

### 6.2 Performance

The I/O throughput in MBps of the arrays depend on the drive speed (RPM), drive count, and RAID chosen. The average read I/O throughput measured on the PS6010XV array was 450 MBps when tested with RAID 50.

To determine the array based on throughput needs:

1. Determine the I/O activity and read throughput needed (MBps) for database, logs, and temporary data files.

Logs incur minimal activity during DSS query access, but may incur load for other operations such as bulk batch updates or backup.
2. For each storage pool (as determined in capacity sizing step above) with multiple member arrays:
  - a. Determine the total pool read throughput needed by adding the individual read throughput needs of database entities in the pool.
  - b. Add a 20% overhead throughput factor per pool.
  - c. Determine the number of arrays based on throughput needed and RAID chosen.
  - d. Ensure the arrays chosen have required capacity per pool as described in section 6.1.

### 6.3 Scalable and modular deployment architecture

The Oracle RAC server nodes under test were configured with two Intel Xeon CPUs (X5690) at 3.47GHz, six cores, and 48 GB RAM. One node was able to drive two PS6010XV arrays in RAID 50 within a reasonable average CPU utilization limit of 75%. Two nodes were able to drive four PS6010XV arrays in RAID 50. A single server and two storage arrays with similar specifications can be viewed as a single server/storage block driving a certain amount of I/O throughput. This modular block concept is illustrated in Figure 17.

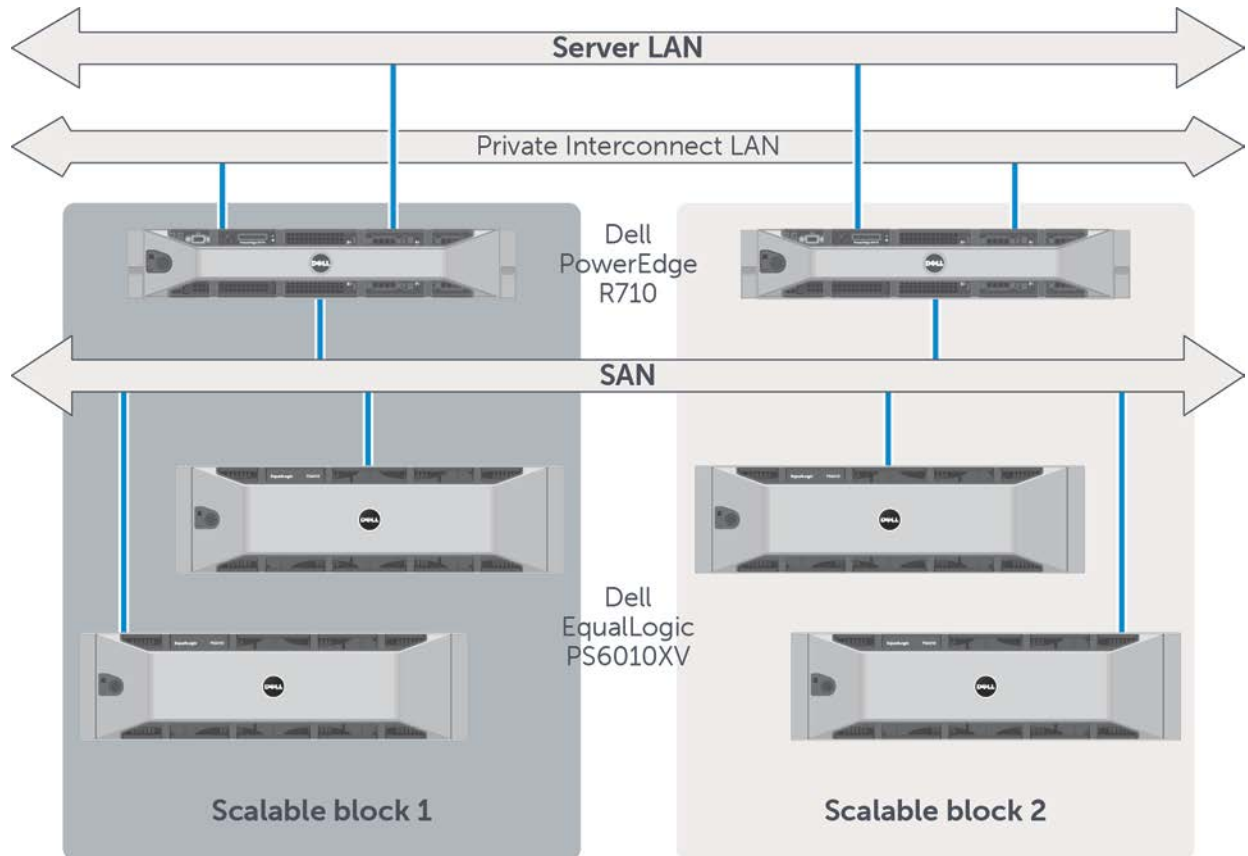
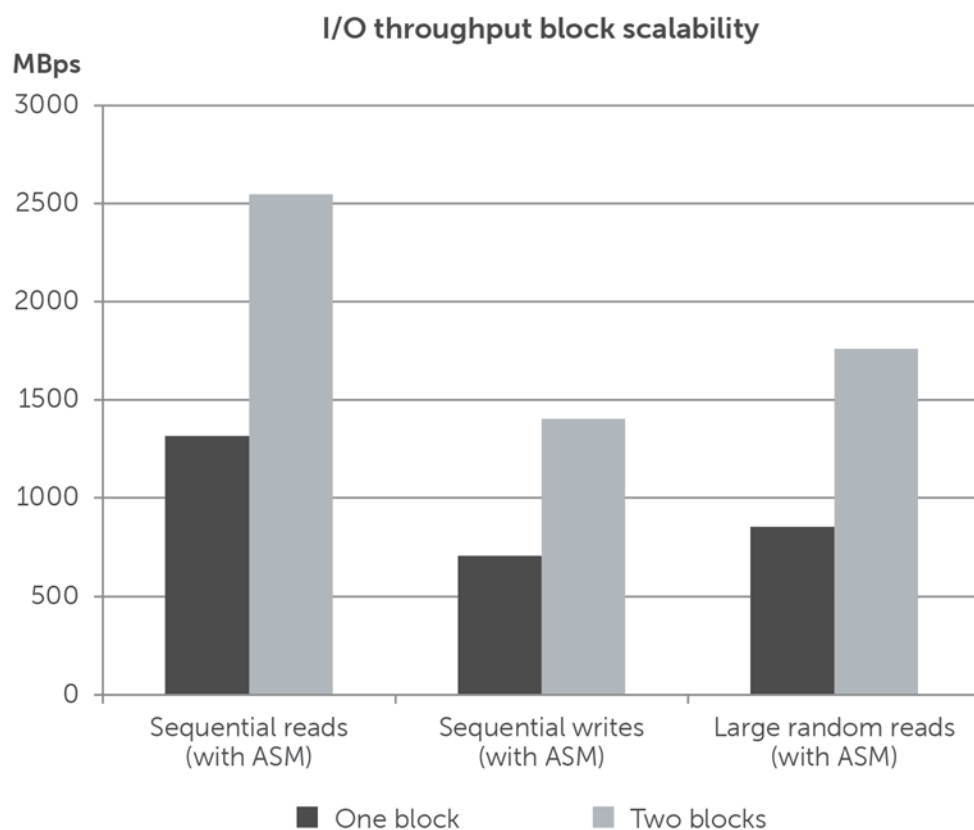


Figure 17 Scalable modular block

The I/O throughput achieved scales with the server/storage blocks. Based on total I/O throughput needed, multiple such blocks can be deployed rendering a scalable and modular architecture for deploying Oracle RAC servers with EqualLogic storage. The I/O throughput per deployment block would map to a certain DSS query rate per hour. The I/O throughput scalability for up to two blocks is shown in Figure 18.



**Figure 18** I/O throughput block scalability

As seen in the figure above, with sequential reads one block provides approximately 1300 MBps of I/O throughput and two blocks provide 2600 MBps. With sequential writes, one block provides 700 MBps and two blocks provide 1400 MBps. Large random reads map to I/O needs of DSS applications and one block provides about 900 MBps of read I/O throughput. Two blocks provide 1800 MBps of throughput.

## 7 Conclusion

EqualLogic 10 GbE iSCSI PS series arrays provide high levels of I/O throughput warranted by DSS. By adding arrays, PS series scale not only capacity but also I/O throughput as well. This is due to the scale-out architecture where all array resources, including controllers and NICs, being scaled proportionately. Optimal operation of DSS can be achieved when the applicable best practices laid out in this paper are adhered to. It must be ensured that the entire ecosystem including server resources and network switches are sized and configured appropriately to meet the workload performance requirements. It must also be ensured that the operating system and database are configured with the optimal settings and design recommendations mentioned in this paper.

# Appendix A System

## A.1 Hardware and software component versions

Component	Details / Version
<b>Server</b>	
Model	PowerEdge R710
CPU	2 x Intel(R) Xeon(R) CPU X5690 @ 3.47GHz with 6 cores
memory	48GB RAM
Server BIOS	6.0.7
iSCSI NIC	Intel® 82599EB 10 Gigabit Ethernet Controller with 2 x 10GbE SFP+ ports
iSCSI NIC Driver / Firmware	3.2.9-NAPI / 0.9-3
Operating System	Red Hat Enterprise Linux 5.5
<b>iSCSI Switch</b>	
Model	PowerConnect 8024F
Switch Firmware	4.1.0.19
<b>Storage</b>	
Model(s)	EqualLogic PS6010XV - 16 x 600GB 15K SAS EqualLogic PS6510E - 48 x 2TB 7.2K SATA
Storage Firmware	V5.1.2 (R197668)
Server MPIO	Host Integration Toolkit for Linux Version 1.0.0

# Appendix B Orion

Detailed information on Orion tools parameters is listed here -

[http://docs.oracle.com/cd/E11882\\_01/server.112/e16638/iodesign.htm#BABEIFFG](http://docs.oracle.com/cd/E11882_01/server.112/e16638/iodesign.htm#BABEIFFG)

## B.1 OS Disk volumes

Target OS disk device volumes under test were listed in a test file and provided as a parameter input for the Orion test runs. Example parameter file *test.lun* is listed below:

```
/dev/mapper/eql-0-8a0906-d7c95290b-521000000224ec67-vol-1  
/dev/mapper/eql-0-8a0906-d7c95290b-5cf0000002a4ec67-vol-5  
/dev/mapper/eql-0-8a0906-d7c95290b-63b000000284ec67-vol-4  
/dev/mapper/eql-0-8a0906-d7c95290b-ba2000000264ec67-vol-3  
/dev/mapper/eql-0-8a0906-d7c95290b-e8c000000244ec67-vol-2  
/dev/mapper/eql-0-8a0906-d7d95290b-1fe0000002c4ec67-vol-6  
/dev/mapper/eql-0-8a0906-d7d95290b-4780000002e4ec67-vol-7  
/dev/mapper/eql-0-8a0906-d7d95290b-88a000000304ec67-vol-8
```

## B.2 Matrix test: Determine optimal parameter

**Note:** The matrix test was required to find optimal parameter *num\_large* for sequential read tests only. All other orion parameters were kept constant. This parameter affects number of outstanding parallel I/O and was critical in maximizing sequential throughput. Too little value for a configuration does not generate sufficient load and too high of a value generates more than required I/O resulting in parallelization of the sequential load.

### Sequential Read in concat mode and 256KB block size

```
./orion_linux_x86-64 -run advanced -testname test -matrix col -num_small 0 -size_large 256 -  
num_disks 8 -type seq -num_streamIO 48 -simulate concat -verbose
```

### Sequential Read in raid0 mode in raid0 mode and 1MB block size

```
./orion_linux_x86-64 -run advanced -testname test -matrix col -num_small 0 -size_large 1024 -  
num_disks 8 -type seq -num_streamIO 48 -simulate raid0 -verbose
```

## B.3 Test run parameters

**Note:** The below parameters apply to tests with 1 array as target storage. Configurations with more than one array had the *num\_large* multiplied by the number of arrays under test. This change only applies to read tests and not write tests. For sequential write tests this parameter was kept constant at 1 across all test configurations as this offered the best throughput.

The parameter *num\_disks* was kept constant at 8 since the number of logical disk device volumes (EqualLogic storage volumes) across all configurations was kept constant at 8. The parameter *num\_streamIO* was kept constant at 48 across all sequential read/write tests. The value is equal to 2 times the number of logical CPUs on the server under test. The server under test had 2 x 6 Core Intel CPUs with hyper-threading enabled. This equals to 24 logical CPUs (2 CPUs x 6 cores x 2 for hyper-threading) on the server. The factor of 2 times logical CPUs is derived from the default value used by Oracle for parallelizing thread per CPU which is 2 (Oracle parameter *PARALLEL\_THREADS\_PER\_CPU*). More information can be found here: [http://docs.oracle.com/cd/E14072\\_01/server.112/e10820/initparams183.htm](http://docs.oracle.com/cd/E14072_01/server.112/e10820/initparams183.htm).

### Sequential Read in concat mode and 256KB block size

```
./orion_linux_x86-64 -run advanced -testname test -matrix point -num_small 0 -size_large 256 -num_large 7 -num_disks 8 -type seq -num_streamIO 48 -simulate concat -verbose -duration 1200
```

### Sequential Read in raid0 mode and 1MB block size

```
./orion_linux_x86-64 -run advanced -testname test -matrix point -num_small 0 -size_large 1024 -num_large 2 -num_disks 8 -type seq -num_streamIO 48 -simulate raid0 -verbose -duration 1200
```

### Sequential Write in raid0 mode and 1MB block size

```
./orion_linux_x86-64 -run advanced -testname test -matrix point -num_small 0 -size_large 1024 -num_large 1 -num_disks 8 -type seq -num_streamIO 48 -simulate raid0 -verbose -write 100 -duration 1200
```

### Random Read in raid0 mode and 1MB block size

```
./orion_linux_x86-64 -run advanced -testname test -matrix point -num_small 0 -size_large 1024 -num_large 128 -num_disks 8 -type rand -simulate raid0 -verbose -duration 1200
```

# Appendix C Oracle

## C.1 Oracle ASM udev rules workaround

The udev rules file content laid out below creates DM device handles for EqualLogic multipath devices in a location where the Oracle ASM utility will locate them. This is a temporary workaround for Oracle ASM with HIT/Linux 1.0.0. Future version of the HIT/Linux will address this as part of the toolkit.

1. Install this file to location: `/etc/udev/rules.d/39-eqloracle.rules`
2. Ensure the following is set on `/etc/sysconfig/oracleasm`: `ORACLEASM_SCANORDER=dm` and `ORACLEASM_SCANEXCLUDE=sd`
3. Reboot server

Contents of `39-eqloracle.rules`:

```
#-----  
-----  
  
#Copyright (c) 2011 by Dell, Inc.  
  
#  
  
# All rights reserved. This software may not be copied, disclosed,  
# transferred, or used except in accordance with a license granted  
# by Dell, Inc. This software embodies proprietary information  
# and trade secrets of Dell, Inc.  
  
#  
  
#-----  
-----  
  
#  
  
# This udev rules file creates DM device handles for Equallogic  
multipath  
  
# devices in a location where the Oracle ASM utility will locate them.  
  
#  
  
# You should also set the following values in /etc/sysconfig/oracleasm:  
  
# ORACLEASM_SCANORDER="dm"  
  
# ORACLEASM_SCANEXCLUDE="sd"  
  
#  
  
# Note: It is important that you have blacklisted all Equallogic devices  
in
```



```

# /etc/multipath.conf so duplicate devices are not accidentally created.
See

# the Equallogic HIT Kit user guide for details.

#

# Skip non-DM devices:
KERNEL!="dm-[0-9]*", GOTO="end_eqloracle"

# For DM devices, query the table name:
PROGRAM="/sbin/dmsetup info -c --noheadings -o name -j %M -m %m"

# Based on the result of DM name query above, creates a /dev/dm-* device
for

# each user-facing (ie, mountable) equallogic device.
RESULT=="eql-?-??????-????????-????????????????????-?*",
OPTIONS="last_rule"

LABEL="end_eqloracle"

```

## C.2 SQL\*Plus statements for database tablespaces and files

### Database Entity names:

Database name = *TPCH1*

User Schema name = *DB1*

Data Disk Group = *DB1AUDG*

Temporary data Disk Group = *TMP1AUDG*

### Statements:

```
CREATE TABLESPACE TS1 DATAFILE +DB1AUDG SIZE 20480M AUTOEXTEND ON NEXT 100M
MAXSIZE UNLIMITED;
```

```
ALTER TABLESPACE TS1 ADD DATAFILE +DB1AUDG SIZE 20480M AUTOEXTEND ON NEXT 100M
MAXSIZE UNLIMITED;
```

*(REPEAT ABOVE ALTER STATEMENT 22 TIMES)*

```
CREATE TEMPORARY TABLESPACE TEMPS TEMPFIL +TMP1AUDG SIZE 20480M AUTOEXTEND ON
NEXT 100M MAXSIZE UNLIMITED;
```

```
ALTER TABLESPACE TEMPS ADD TEMPFILE +TMP1AUDG SIZE 20480M AUTOEXTEND ON NEXT
100M MAXSIZE UNLIMITED;
```

*(REPEAT ABOVE ALTER STATEMENT 22 TIMES)*

```
CREATE USER DB1 IDENTIFIED BY DB1 DEFAULT TABLESPACE TS1 TEMPORARY TABLESPACE TEMPS;
```

```
GRANT DBA TO DB1 IDENTIFIED BY DB1;
```

```
ALTER DATABASE DEFAULT TEMPORARY TABLESPACE TEMPS;
```

### **C.3 SQL tuning**

One of the 22 user queries required custom SQL Tuning. The SQL Tuning Advisor utility included within Oracle 11g R2 Enterprise Manager was used to tune this query. A SQL Profile was created to capture this tuning output and execution plan. The following procedure was used to export this custom SQL Profile and later imported across configurations to consistently apply this tuning.

1. SQL\*PLUS statement to pack SQL Profile in a table:  
EXEC DBMS\_SQLTUNE.PACK\_STGTAB\_SQLPROF (staging\_table\_name => *PROFTAB*);
2. Export table using *EXP* utility to a file
3. Import table from file to new database using *IMP* utility
4. SQL\*PLUS statement to unpack SQL Profile from table:  
EXEC DBMS\_SQLTUNE.UNPACK\_STGTAB\_SQLPROF(replace => FALSE, staging\_table\_name  
=> *PROFTAB*);

Verify imported SQL Profile in system view *dba\_sql\_profiles*

## Additional resources

Support.dell.com is focused on meeting your needs with proven services and support.

DellTechCenter.com is an IT Community where you can connect with Dell Customers and Dell employees for the purpose of sharing knowledge, best practices, and information about Dell products and your installations.

Referenced or recommended Dell publications:

- Dell EqualLogic Configuration Guide:  
<http://en.community.dell.com/techcenter/storage/w/wiki/equallogic-configuration-guide.aspx>
- Dell EqualLogic PS Series Network Performance Guidelines:  
<http://www.equallogic.com/resourcecenter/assetview.aspx?id=5229>

Referenced or recommended Oracle publications:

- I/O configuration and Design:  
[http://docs.oracle.com/cd/E11882\\_01/server.112/e16638/iodesign.htm](http://docs.oracle.com/cd/E11882_01/server.112/e16638/iodesign.htm)
- Administering Oracle ASM Disk Groups:  
[http://docs.oracle.com/cd/E16338\\_01/server.112/e10500/asmdiskgrps.htm](http://docs.oracle.com/cd/E16338_01/server.112/e10500/asmdiskgrps.htm)
- Administering Oracle ASM Files, Directories, and Templates:  
[http://docs.oracle.com/cd/E16338\\_01/server.112/e10500/asmfiles.htm](http://docs.oracle.com/cd/E16338_01/server.112/e10500/asmfiles.htm)
- Setting up a Data Warehouse system:  
[http://docs.oracle.com/cd/E14072\\_01/server.112/e10578/tdpdw\\_system.htm](http://docs.oracle.com/cd/E14072_01/server.112/e10578/tdpdw_system.htm)
- Using ASM for managing OS disk volumes due to the advanced management capabilities:  
[http://docs.oracle.com/cd/E11882\\_01/server.112/e16638/iodesign.htm#i12331](http://docs.oracle.com/cd/E11882_01/server.112/e16638/iodesign.htm#i12331)
- Utilizing ASM external redundancy or striping  
[http://docs.oracle.com/cd/E11882\\_01/server.112/e18951/asmdiskgrps.htm#BABBEJAF](http://docs.oracle.com/cd/E11882_01/server.112/e18951/asmdiskgrps.htm#BABBEJAF)
- ASM for Oracle data files:  
[http://docs.oracle.com/cd/E16338\\_01/server.112/e10500/asmfiles.htm#CIHIEGBI](http://docs.oracle.com/cd/E16338_01/server.112/e10500/asmfiles.htm#CIHIEGBI)
- Database initialization parameter *DB\_FILE\_MULTIBLOCK\_READ\_COUNT*:  
[http://docs.oracle.com/cd/B19306\\_01/server.102/b14237/initparams047.htm](http://docs.oracle.com/cd/B19306_01/server.102/b14237/initparams047.htm)

For EqualLogic best practices white papers, reference architectures, and sizing guidelines for enterprise applications and SANs, refer to Storage Infrastructure and Solutions Team Publications at:

- <http://dell.to/sM4hJT>



THIS WHITE PAPER IS FOR INFORMATIONAL PURPOSES ONLY, AND MAY CONTAIN TYPOGRAPHICAL ERRORS AND TECHNICAL INACCURACIES. THE CONTENT IS PROVIDED AS IS, WITHOUT EXPRESS OR IMPLIED WARRANTIES OF ANY KIND.