Sizing and Best Practices for Online Transaction Processing Applications with Oracle 11g R2 using Dell PS Series

Dell EMC Engineering
March 2017

A Dell EMC Technical White Paper

Revisions

Date	Description
December 2013	Initial release
March 2017	Updated for Dell EMC branding; consolidated two best practices documents (BP1003 and BP1069)

Acknowledgements

This best practice white paper was produced by the following members of the Dell Storage team:

Engineering: Chidambara Shashikiran

Technical Marketing: Chhandomay Mandal

Additional contributors: Darren Miller, Henry Wong, Rob Young, Chuck Farah, Rick Sussman

Table of contents

Executive summary

Online transaction processing (OLTP) applications — ranging from web-based e-commerce sites, to accounting systems, to customer support programs — are at the heart of today's business functions. They are the foundation for a wide range of mission-critical applications, particularly those utilizing Oracle® real application cluster (RAC) database platforms, and depend on exceptional storage performance and reliability. Moreover, data generated by OLTP applications can grow exponentially over time. Thus, designing and managing storage that can effectively accommodate performance, capacity, and future growth requirements for Oracle OLTP databases is a key challenge today.

This paper presents sizing guidelines and best practices for deploying Oracle 11g R2 OLTP databases with Dell™ PS Series arrays based on a series of storage I/O performance test results. A variety of OLTP benchmarking tools including Oracle I/O Numbers (ORION), Vdbench, TPC-C, and TPC-E were run to demonstrate the efficacy of the PS Series arrays for the OLTP workloads. Detailed tests results for Dell PS6110XS hybrid arrays are included in this paper along with some key performance results for PS6010XV and PS6010S arrays.

Key findings include:

- PS6110XS hybrid arrays can deliver high IOPS with low latency for OLTP workloads. For example: 25,745 IOPS with less than 1 ms latency for a TPC-E like workload; and 17,000 IOPS with 5 ms latency for a TPC-C like workload.
- Highly effective, dynamic, automated data tiering by PS Series firmware for OLTP environments. For example: 51% IOPS increase over time due to tiering between SSDs and HDDs in Vdbench test.
- Oracle ASM and PS Series load balancing features complement each other to provide optimal performance.
- A single PS6110XS array produced more than 8,700 IOPS even with very heavy write-intensive operations simulated as part of TPC-E workload on a 2-node Oracle RAC database.
- The PS6010XV array in RAID 10 configuration was able to sustain approximately 5,000 IOPS and PS6010S array produces almost 35,000 IOPS for 100% small random read I/O.

Additionally, this paper presents guidelines for designing Oracle OLTP environments including database volume layout configurations, Oracle Automatic Memory Management (AMM), and Oracle Automatic Storage Management (ASM) considerations for the PS Series platform are also discussed in this paper.

1 Introduction

Different types of database applications have varying performance and capacity needs. Understanding the models for common database application workloads can be useful in predicting the possible application behavior in a given infrastructure environment. The most common database application workload models are OLTP and data warehousing (DW). This paper focuses on OLTP database workloads.

OLTP applications often need to deliver extremely fast response times and rapid access to data that is in high demand. Moreover, significant data growth over time in a typical OLTP system leads to high capacity requirements as well. While the OLTP data set size can grow exponentially over time, the active, frequently-accessed data set size remains relatively small in comparison with the overall OLTP system size.

Storage infrastructure supporting OLTP platforms like Oracle RAC needs to non-disruptively scale both in performance and capacity, and also deliver the required high performance in an optimal, cost-effective way. PS Series hybrid arrays – combining solid-state drives (SSDs) and hard disk drives (HDDs) within a single chassis with automated data tiering – are well-suited for these OLTP environments.

1.1 Objective

The key objectives of testing were:

- Establish baseline I/O performance characteristics of the PS Series platform when running OLTP-like I/O patterns using the ORION tool without deploying Oracle databases.
- Performance characterization using Benchmark Factory® for Databases, simulating Oracle OLTP transactions on the 2-node Oracle RAC database with a TPC-E type workload.
- I/O profiling with Vdbench to demonstrate performance improvements with the automated data tiering of PS6110XS hybrid arrays due to locality of data access in Oracle OLTP environments.
- Determine guidelines and best practices for deploying and achieving optimal performance of Oracle 11g R2 OLTP databases on PS Series arrays.

1.2 Audience

This paper is intended for system administrators, database administrators, and storage architects interested in deploying Oracle OLTP database solutions on Red Hat Enterprise Linux (RHEL) using PS Series SANs. It is assumed that the readers of this document have familiarity with RHEL system administration, PS Series storage array configuration, and Oracle 11g R2 database installation and administration tasks.

1.3 Terminology

The following terms are used throughout this document:

AMM: The Oracle Automatic Memory Management (AMM) feature automates and simplifies the memory-tuning configuration tasks by allowing you to manage the different memory components of Oracle using a single database initialization parameter called MEMORY_TARGET.

ASM: Oracle Automatic Storage Management (ASM) is a volume manager and a file system that supports Oracle database.

Group: One or more PS Series arrays connected to an IP network that work together to provide SAN resources to host servers.

Member: A single physical PS Series array.

OLTP I/O pattern: OLTP workloads tend to select a small number of rows at a time. These transfers happen anywhere in the data, and are each fairly small in size, typically between 8K and 64K. This causes the I/O pattern to be random in nature.[1] The key metric in measuring performance of OLTP workloads is the number of I/Os per second (IOPS) that can be achieved while maintaining a healthy response time.

ORION: Oracle I/O Numbers (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 are included as part of the Oracle 11g installation software.

Quest Benchmark Factory: This is a database performance tool which can simulate actual database transactions such as TPC-C and TPC-E per the TPC specifications. More details on this tool and the download location can be found at http://www.quest.com/benchmark-factory/.

SAN HQ: SAN Headquarters (SAN HQ) monitors one or more PS Series Groups. The tool is a client/server application that runs on a Microsoft® Windows® system and uses the Simple Network Management Protocol (SNMP) to query the groups. Much like a flight data recorder on an aircraft, SAN HQ collects data over time and stores it on the server for later retrieval and analysis. Client systems connect to the server to format and display the data in the SAN HQ GUI.

TPC-E: A benchmark developed by TPC to simulate OLTP database workloads. The TPC-E benchmark simulates the OLTP workload of a brokerage firm. Although the underlying business model of TPC-E is a brokerage firm, the database scheme, data population, transaction mix, and implementation rules broadly represent a modern OLTP database system. More information on TPC-E can be found at http://www.tpc.org/tpce/default.asp.

Vdbench: Vdbench is a disk and tape I/O workload generator specifically created to help engineers and customers generate disk I/O workloads to be used for validating storage performance and storage data integrity. The tool can be downloaded at the following link: http://www.oracle.com/technetwork/server-storage/vdbench-downloads-1901681.html

[1] Data warehouse (DW) workloads, primarily composed of running database queries and generating reports, are sequential in nature. If the same database is used for both transaction processing and report generation, a mix of OLTP and DW workloads will be observed.

2 OLTP applications with PS Series arrays

OLTP database applications are optimal for managing rapidly changing data. These applications typically have many users who are performing transactions simultaneously. Although individual data requests by users usually reference few records, many of these requests are being made at the same time.

To develop appropriate storage performance and capacity sizing guidelines for OLTP workloads, it is essential to understand the following factors:

- Read I/O and write I/O percentage
- I/O block sizes
- Data access rate
- Data locality
- Memory configuration on the database server (this plays a significant role in Oracle database I/O performance)

2.1 PS Series hybrid arrays

The important characteristics of OLTP workload are a higher percentage of read I/O than write I/O, and a high percentage of total I/O being concentrated on a relatively small percentage of the total data set. PS Series hybrid storage arrays that combine SSDs and HDDs within the same chassis with automated, dynamic data tiering are well suited for these OLTP applications. PS Series firmware automatically moves more frequently accessed data to high-performing SSDs and less-frequently accessed data to HDDs, thereby helping optimize the balance between performance and capacity needs of a demanding Oracle OLTP environment.

3 Solution infrastructure

The test system used to conduct ORION, Vdbench, and TPC-C/TPC-E testing for this paper is shown in this section.

3.1 Physical system test configuration

The logical test configuration diagram for simulating Oracle OLTP database I/O using the ORION tool is shown in Figure 1.

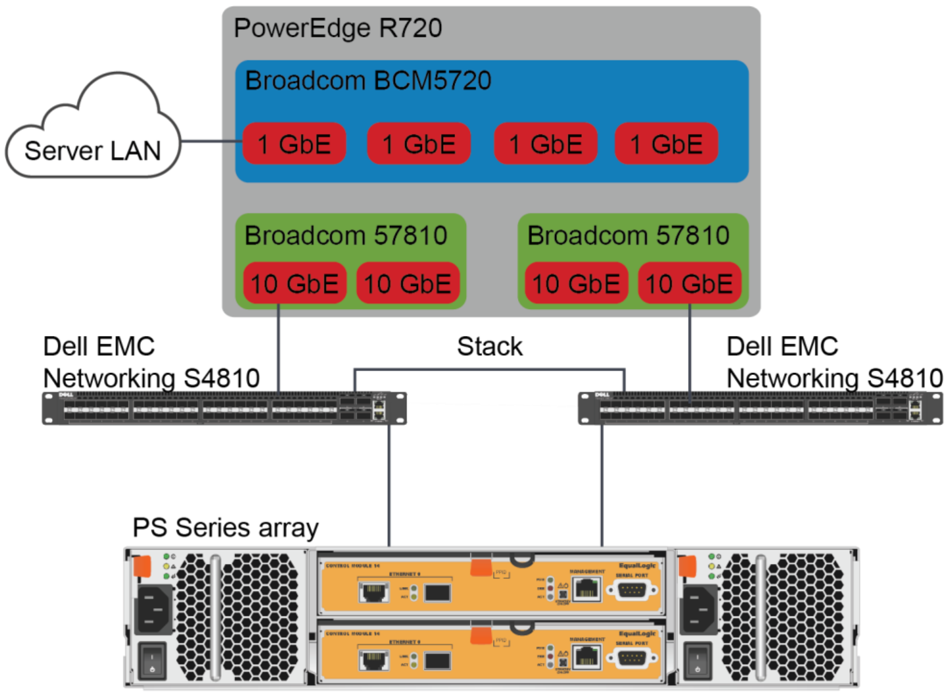


Figure 1 ORION test configuration — LAN and iSCSI SAN connectivity

The logical test configuration diagram for OLTP database application tests (TPC-E) using Quest Benchmark Factory is shown in Figure 2.

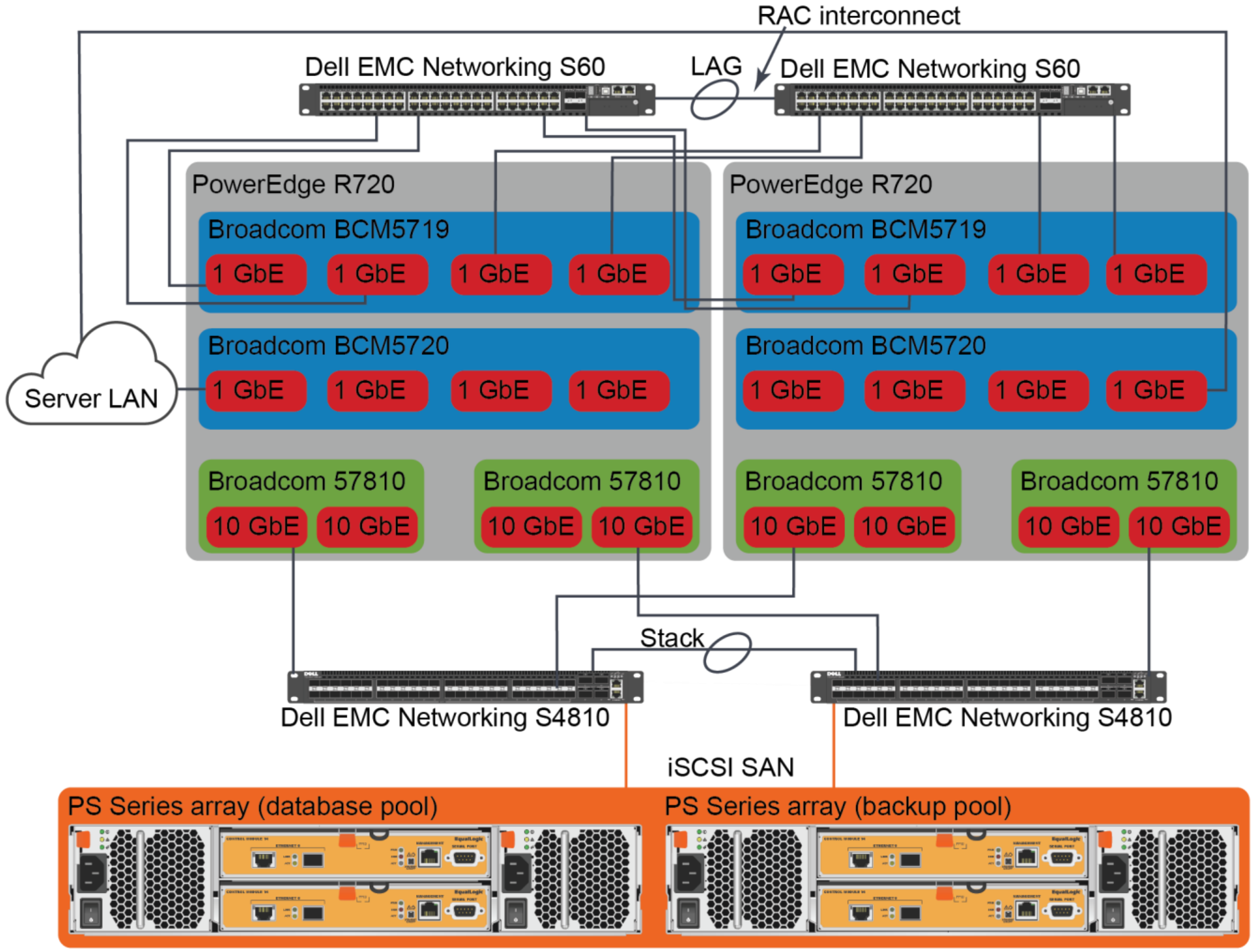


Figure 2 TPC-E test configuration — LAN and iSCSI SAN connectivity

3.2 Storage configuration

The database files, REDO logs, and system-related table spaces such as SYSTEM and UNDO were stored on PS6110XS and PS6010XV arrays in their own pool called **Database pool**. The flashback recovery and archive data was stored on a PS6XXXE array consisting of high-capacity SATA drives which was in a separate pool named **Backup pool**. To enable advanced MPIO functionality, the PS Series Host Integration Tools for Linux (HIT/Linux) were 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 **eqltune**, installed as part of the HIT/Linux package, was run and its OS parameter recommendations were implemented.

Refer to appendix A for configuration details.

3.3 Database layout

The PS Series arrays were placed in different pools, and Oracle ASM disks were configured using the ASM disk layout described as follows.

DB1DATA: Database files; temporary table space; online REDO logs; and system-related table spaces such as SYSTEM and UNDO

DB1OG1/DB1LOG2: REDO logs

DB1FLASH: Archive logs and flash recovery area

OCRVOTE: ASM Cluster File System (ACFS) for storing clusterware-related information such as the OCR and voting disks

Table 1 shows the ASM disk and disk group configuration.

Table 1 Oracle ASM disk and disk group configuration

DB1DATA	DB1LOG1	DB1LOG2	DB1FLASH	OCRVOTE
ORADB1 (250GB) ORADB2 (250GB) ORADB3 (250GB) ORADB4 (250GB) ORADB5 (250GB) ORADB6 (250GB)	ORALOG1 (100GB)	ORALOG2(100GB)	ORAFLASH1 (1.5TB) ORAFLASH2 (1.5TB)	ORACRS (20GB)

Figure 3 shows the containment model and relationships between the PS Series pool and volumes, and the Oracle ASM disk groups and disks.

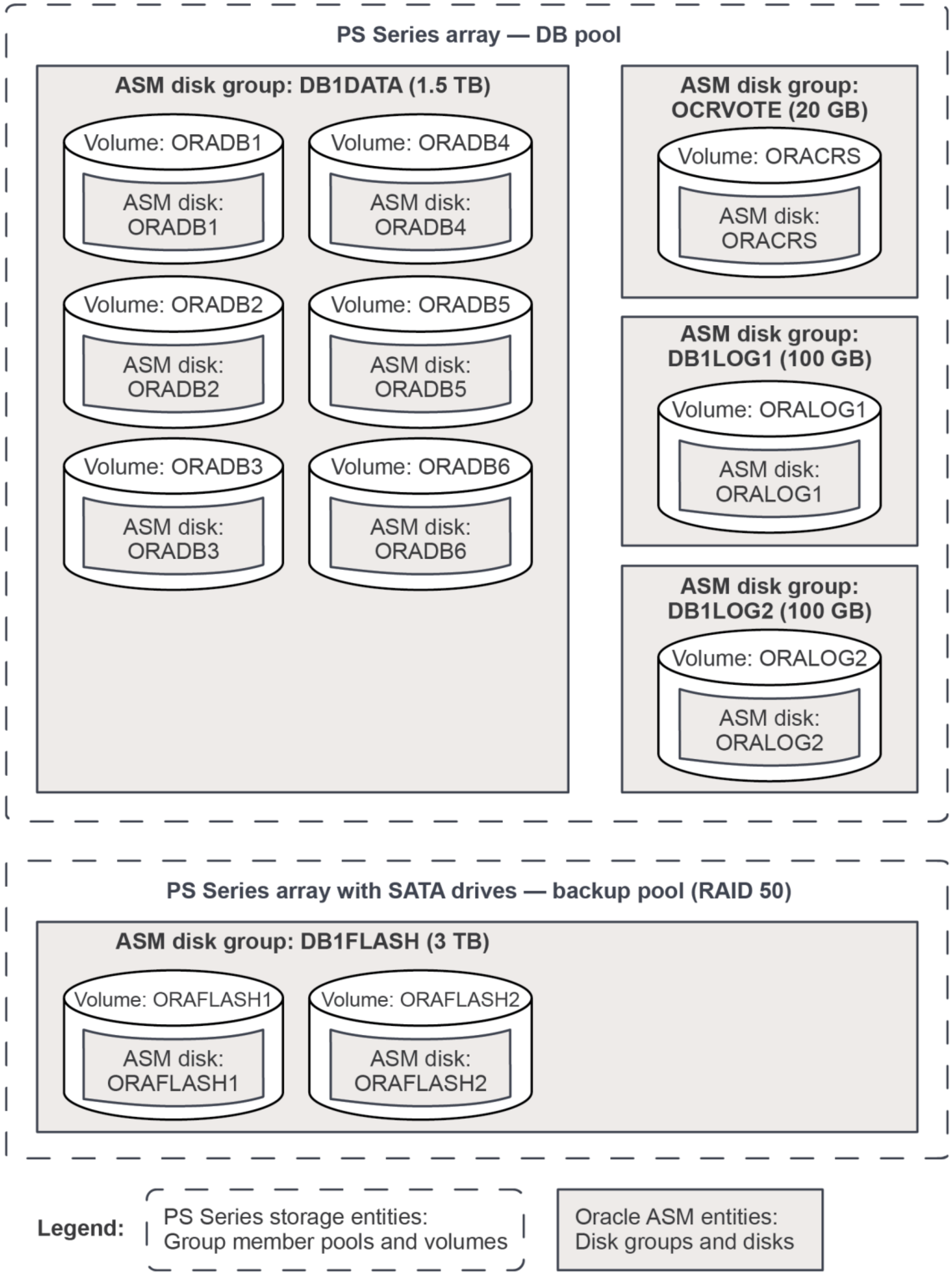


Figure 3 PS Series volume and Oracle ASM disk configuration

4 Test methodology

A series of I/O simulations were conducted using ORION to understand the performance characteristics of the PS Series XV, S, and XS hybrid storage arrays.

The ORION tool was installed on the operating system (Red Hat® Enterprise Linux® in this case) and used to simulate I/O on the external PS Series storage arrays. The PS Series storage volumes were presented to the OS as block I/O devices. I/O simulations were conducted with small-sized 8K I/O block requests simulating Oracle OLTP database I/O to understand the IOPS offered by the storage arrays. The Vdbench I/O tool was used to evaluate the performance characteristics of the array when a specific locality of data on the disks was frequently accessed as observed in a typical database environment.

Typical OLTP I/O is represented by a ratio of 70% read and 30% write I/O transactions. However, modern OLTP I/O is typically represented as 90% read and 10% write, specifically for workloads simulating financial organizations like banks and brokerage firms. Find more detailed information about TPC-E characteristics in the document, *TPC-E vs. TPC-C: Characterizing the New TPC-E Benchmark via an I/O Comparison Study*.

The I/O simulation using the ORION tool was performed to characterize both TPC-C and TPC-E I/O ratios defined previously, but the actual database performance characterization was performed by executing TPC-C-like database transactions on the PS Series XV array and TPC-E style database transactions on PS Series hybrid arrays using Quest Benchmark Factory.

4.1 Measurement and monitoring

Detailed performance metrics were captured from storage, operating system, and database application layers as listed in the following.

- **ORION:** Total IOPS and average latency as part of ORION standard test results were captured.
- **Quest Benchmark Factory:** The following metrics were captured from Quest Benchmark Factory while executing TPC-E database transactions:
  - **TPS**: Transactions per second, based on the simulated load.
  - **Response time**: End-to-end SQL execution time, measured from SQL submission until the end of the fetch of the SQL result set.
  - **Transaction time**: Average time to run a single transaction.
- **SAN HeadQuarters (SAN HQ):** PS Series 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.
- **Oracle AWR reports:** Oracle Automatic Workload Repository (AWR) reports were collected at one-minute intervals while executing TPC-E transactions. The top five events by total wait time were analyzed to determine if there were any bottlenecks related to RAC, network, or other factors.
- **Oracle OSWatcher Black Box (OSWbb):** Oracle OSWatcher Black Box utility captures and archives operating system and network statistics. This utility was used to capture detailed operating system statistics such as vmstat, iostat, netstat, and others at regular intervals while executing the tests.

4.2 Test criteria

The test criteria used for the study includes:

- The storage array disk access latencies (read and write) remain below 20 ms per volume.
- Database server CPU utilization remains below an 80% average.
- TCP retransmits on the storage network remain below 0.5%.
- Application response times remain below two seconds, which is the industry-standard acceptable latency for Oracle OLTP database applications.
- No RAC related bottlenecks (such as database locking, user and system I/O wait, RAC interconnect response, or others) while executing the peak I/O load.

5 I/O profiling study

The ORION tool uses the Oracle I/O software stack to generate simulated I/O workloads without having to create an Oracle database, load the application, and simulate users. ORION is commonly used for benchmarking the I/O performance characteristics of Oracle database storage systems.

A series of ORION tests were executed on PS6010XV, PS6010S, and PS6110XS hybrid arrays for measuring the storage I/O performance and system limits for pure OLTP workloads.

5.1 PS6010XV array test

The key results from ORION tests executed on the PS6010XV array are shown in Figure 4.

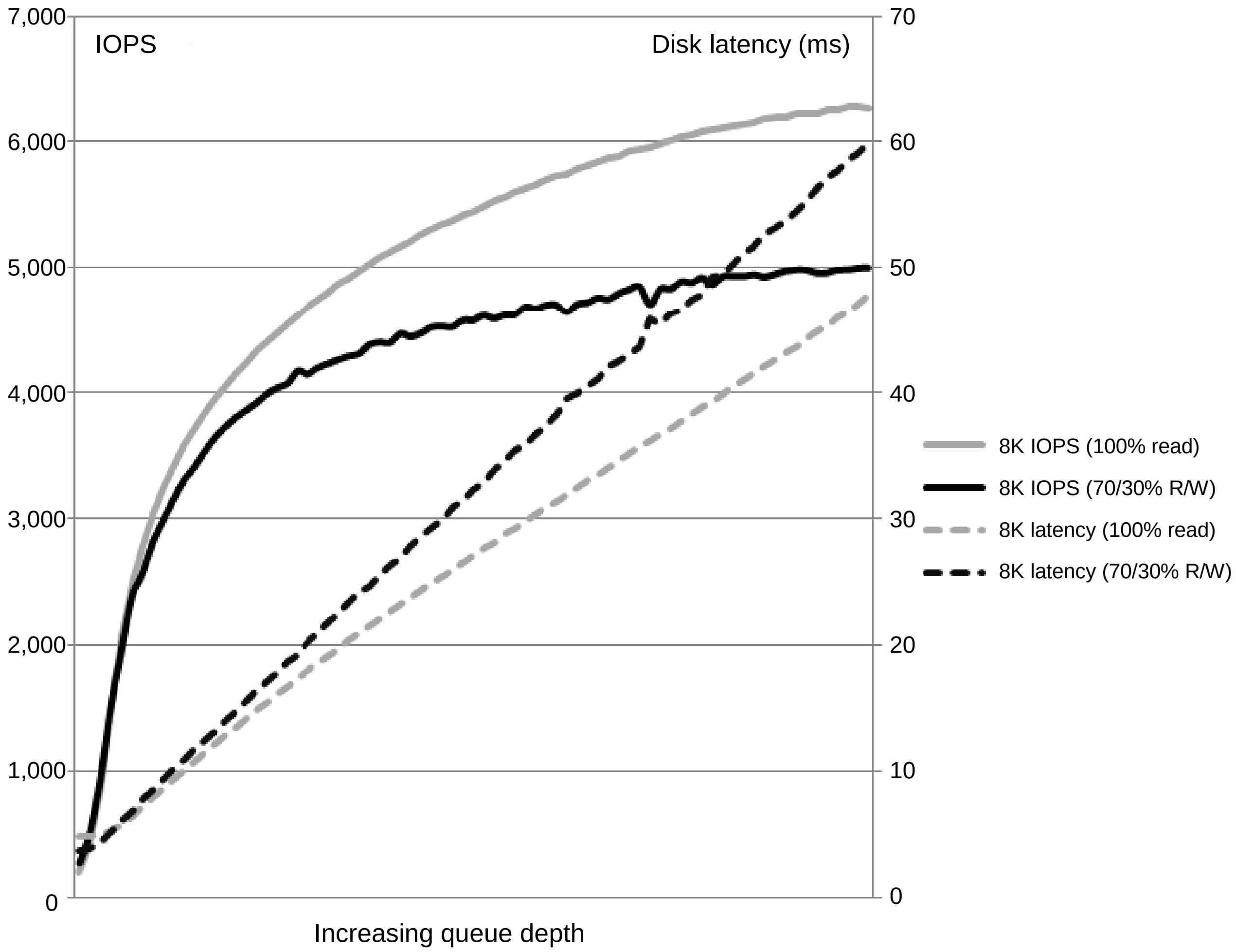


Figure 4 ORION test performance on PS6010XV (15K SAS)

Key findings from the test results are summarized as follows:

- As expected, average IOPS increased with the queue depth.
- The PS6010XV array in a RAID 10 configuration produced approximately 4,000 IOPS for OLTP-type transactions while staying within the generally accepted disk latency limit of 20 ms (for both read and write IOPS).
- For block sizes as large as 64K, a single PS6010XV array is able to sustain approximately 3,000 IOPS for a typical OLTP workload.

5.2 PS6010S array test

The key results from ORION tests executed on the PS6010S array are shown in Figure 5.

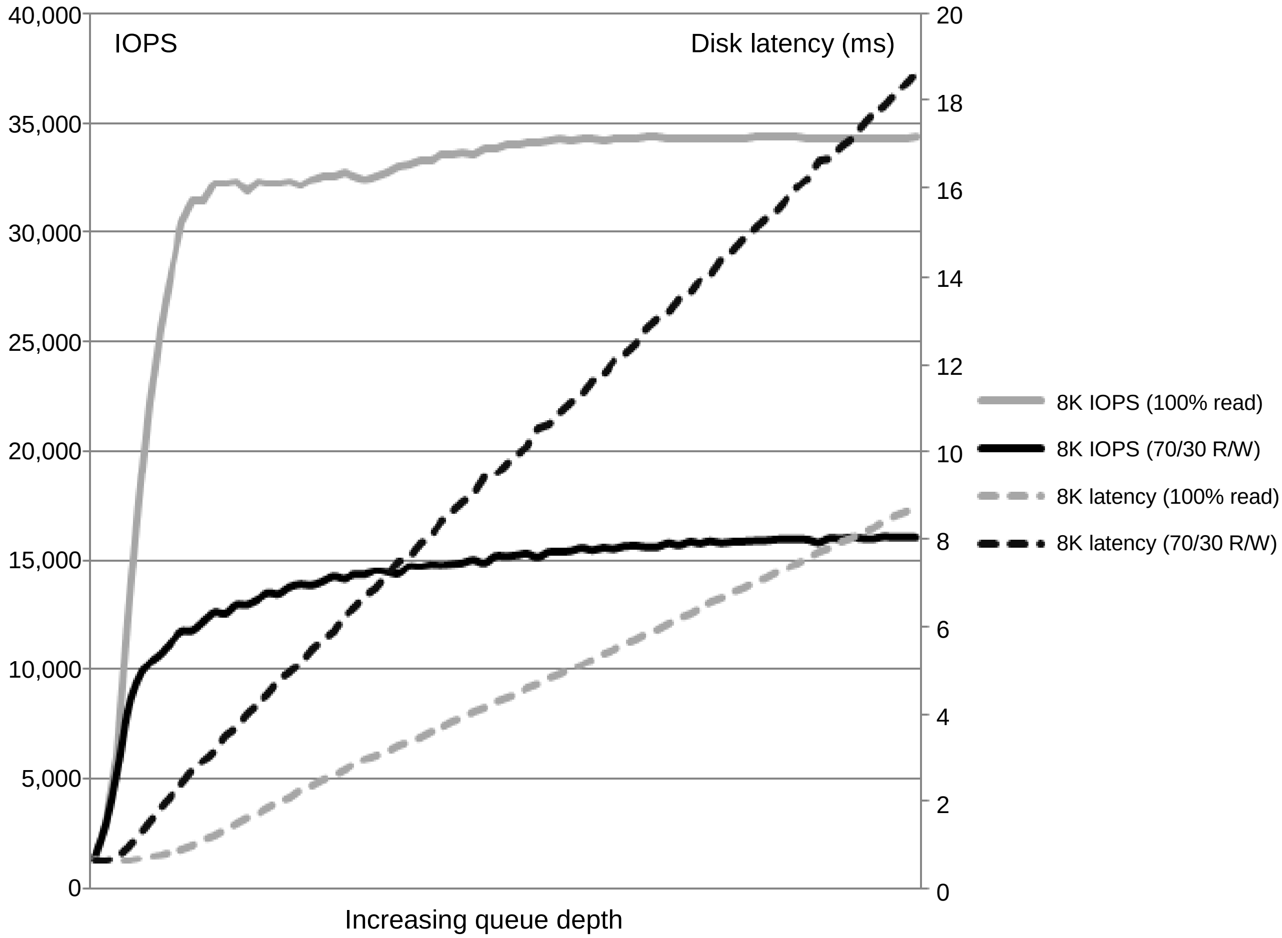


Figure 5 ORION test performance on PS6010S (SSD)

Key findings from the test results are summarized as follows:

- For 100% read I/O with an 8K block size, the PS6010S array sustained a maximum of approximately 34,000 IOPS.
- For 70/30 read/write I/O with an 8K block size, the PS6010S array sustained a maximum of approximately 16,000 IOPS.
- At the maximum IOPS level measured for the PS6010XV array before exceeding the 20 ms disk latency threshold (4,047), the disk latency for the PS6010S array at the same IOPS level was < 2 ms.

5.3 PS6110XS hybrid array test

PS6110XS arrays offer high performance and large usable capacity due to the unique combination of SSD and 10K SAS drives within the same chassis. The array performance in a specific OLTP environment depends on a number of factors including amount of data stored on the array, frequency of access, and most importantly, the locality of data reference.

First, I/O profiling tests were run using ORION to establish baseline I/O performance characteristics of the storage array under test at different capacity configurations.

Next, the Vdbench I/O generation tool was used for simulating frequently accessed data with a specific locality as observed in a typical database environment. The ORION tool was not appropriate for this study because it does not provide an option to specify locality of data reference.

The tests listed in Table 2 were run with an 8K I/O block size and at increasing queue depths (number of outstanding I/Os) to determine the maximum IOPS the storage array would sustain within the 20 ms latency (read and write latencies were measured separately).

The configuration parameters for the tests are shown in Table 2.

Table 2 Test parameters: I/O workload studies

Configuration parameters	
PS Series SAN	PS6110XS (7 x 400GB SSD, 17 x 600GB 10K SAS drives)
Test configuration #1	**ORION** (1TB array capacity utilization, 100% random I/O locality) **I/O mix**: 8K 70/30 read/write (simulate TPC-C)
Test configuration #2	**ORION** (2TB array capacity, 100% random I/O locality) **I/O mix**: 8K 70/30 read/write (simulate TPC-C)
Test configuration #3	**ORION** (4TB array capacity,100% random I/O locality) **I/O mix**: 8K 92/8 read/write (simulate TPC-E)
Test configuration #4	**Vdbench** (2TB array capacity, 20% random I/O locality) **I/O mix**: 70/30 read/write (simulate I/O with 20% I/O locality)
RAID type	RAID 6 Accelerated

The results collected from these tests are illustrated in sections 5.3.1 to 5.3.4. Additional I/O simulation tests were run to evaluate Oracle ASM benefits and also to determine the volume configuration which produces optimal performance. These test results are discussed in sections 5.3.5 and 5.3.6.

5.3.1 ORION (1 TB array capacity utilization, 100% random I/O and no data locality)

ORION tests were run with a range of user loads on 2 x 500 GB volumes. The total capacity utilized on the array was 1 TB and the I/O access was completely random in nature (100% random without any specific locality of I/O access).

This test on the array produced approximately 17,000 IOPS for an 8K block size with a 70/30% read/write mix workload while staying within the generally accepted disk latency limit of 20 ms (for both read and write IOPS measured separately).

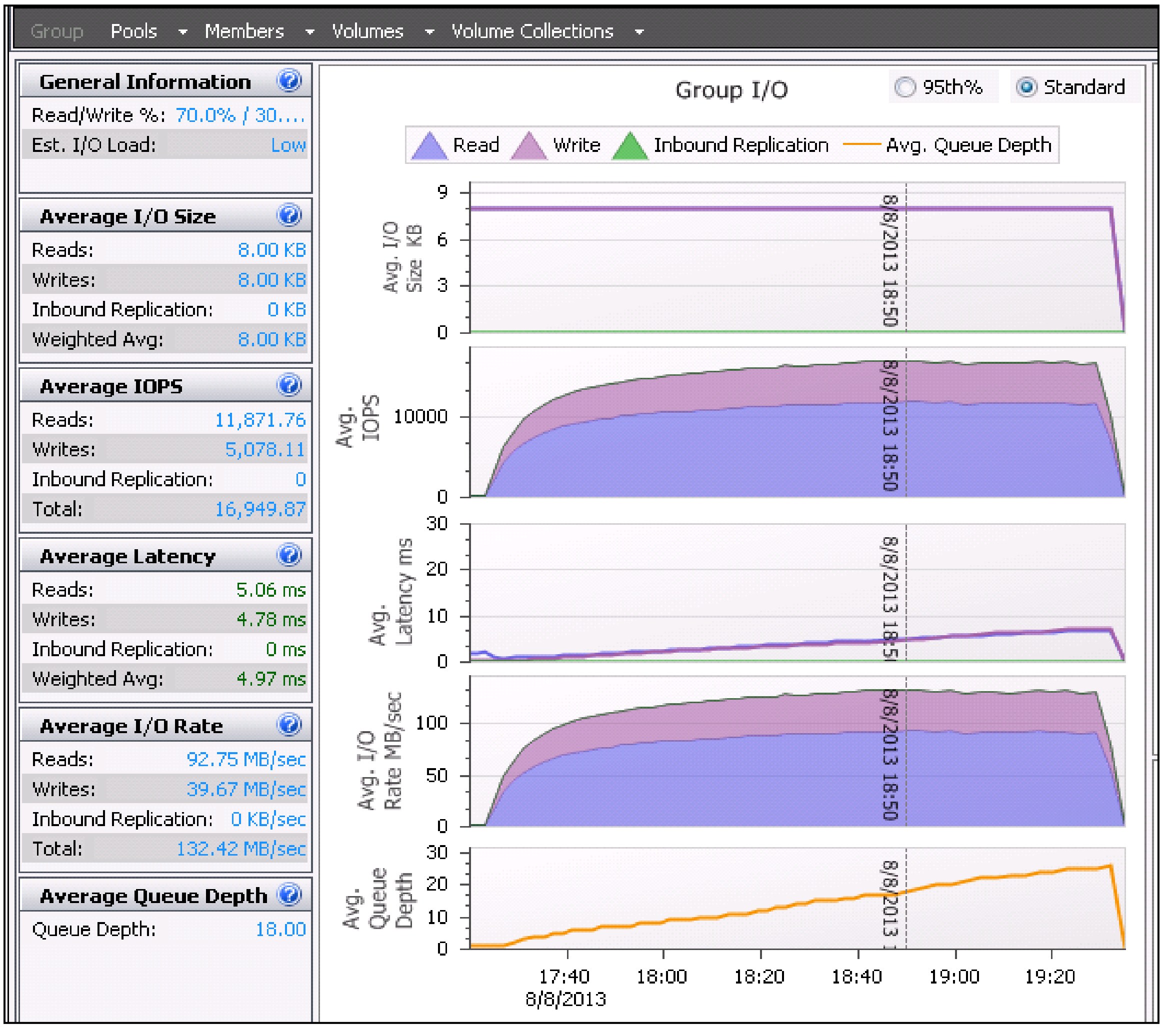


Figure 6 IOPS and latency reported by SAN HQ (1 TB capacity utilization and no data locality)

Because the total capacity of volumes was 1 TB, the seven SSDs were able to handle all the I/O activity and there was not much activity on 10K SAS drives.

The SAN HQ chart in Figure 7 shows that each SSD drive was producing more than 5,300 IOPS during peak I/O activity.

Member	Pool	Disk	Description	Average IOPS	Read I/O Rate	Write I/O Rate	Status	Average Disk Queue Depth
6110XS-2	default	0	SSD 400GB SAS	5,344.97	31.78 MB/sec	20.73 MB/sec	online	4.90
6110XS-2	default	1	SSD 400GB SAS	5,345.03	31.78 MB/sec	20.74 MB/sec	online	4.89
6110XS-2	default	2	SSD 400GB SAS	5,343.17	31.77 MB/sec	20.73 MB/sec	online	4.89
6110XS-2	default	3	SSD 400GB SAS	5,342.65	31.76 MB/sec	20.72 MB/sec	online	4.89
6110XS-2	default	4	SSD 400GB SAS	5,343.43	31.77 MB/sec	20.73 MB/sec	online	4.89
6110XS-2	default	5	SSD 400GB SAS	5,342.51	31.77 MB/sec	20.73 MB/sec	online	4.89
6110XS-2	default	6	SSD 400GB SAS	5,342.48	31.77 MB/sec	20.73 MB/sec	online	4.89
6110XS-2	default	7	10K 600GB SAS	370.04	2.89 MB/sec	<1.0KB/sec	online	1.22
6110XS-2	default	8	10K 600GB SAS	<1.0	<1.0KB/sec	1.18 KB/sec	online	0.13
6110XS-2	default	9	10K 600GB SAS	<1.0	<1.0KB/sec	1.18 KB/sec	online	0.13
6110XS-2	default	10	10K 600GB SAS	<1.0	<1.0KB/sec	1.18 KB/sec	online	0.12
6110XS-2	default	11	10K 600GB SAS	1.73	127.56 KB/...	51.98 KB/sec	online	0.94
6110XS-2	default	12	10K 600GB SAS	<1.0	<1.0KB/sec	<1.0KB/sec	online	0.14
6110XS-2	default	13	10K 600GB SAS	0	0 KB/sec	0 KB/sec	spare	0
6110XS-2	default	14	10K 600GB SAS	<1.0	<1.0KB/sec	1.31 KB/sec	online	0.20
6110XS-2	default	15	10K 600GB SAS	<1.0	<1.0KB/sec	1.31 KB/sec	online	0.21
6110XS-2	default	16	10K 600GB SAS	<1.0	<1.0KB/sec	1.18 KB/sec	online	0.15
6110XS-2	default	17	10K 600GB SAS	<1.0	<1.0KB/sec	1.18 KB/sec	online	0.13
6110XS-2	default	18	10K 600GB SAS	<1.0	<1.0KB/sec	<1.0KB/sec	online	0.18

Figure 7 Disk IOPS activity reported by SAN HQ (1 TB capacity utilization and no data locality)

5.3.2 ORION (2 TB array capacity utilization, 100% random I/O and no data locality)

In this test, 4 x 500 GB volumes were used to run the ORION test. The total capacity utilization was 2 TB. Because there was no locality of data reference and the utilized capacity exceeded total SSD capacity, I/O activity was observed on both the SSD and 10K SAS drives for this completely random workload as shown in Figure 8.

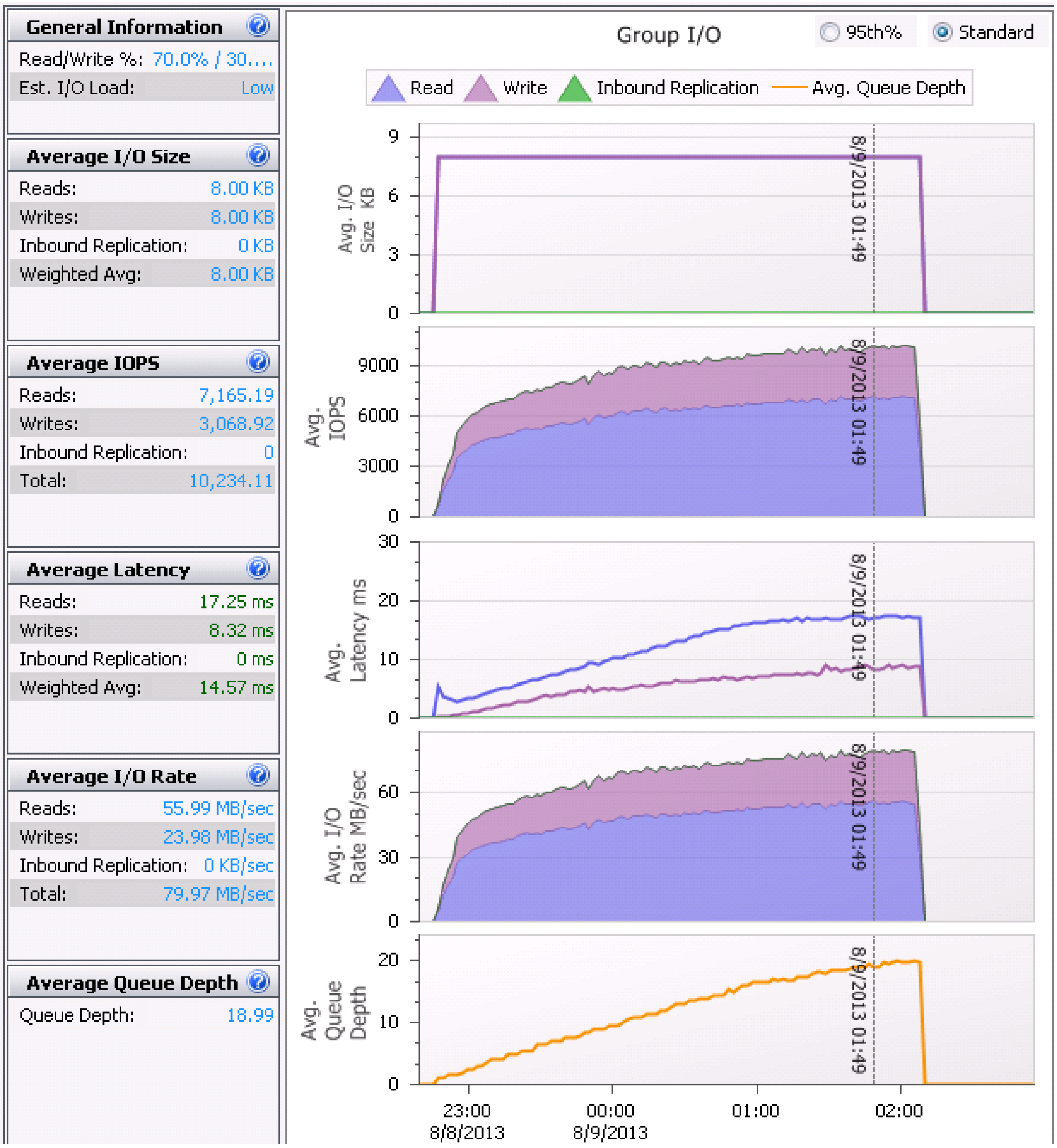


Figure 8 IOPS and latency reported by SAN HQ (2 TB capacity utilization and no data locality)

The peak load was around 10,234 IOPS compared to 17,000 IOPS in the previous test configuration.

The entire capacity of the SSDs was saturated when the space utilization was increased to 2 TB. The simulated I/O was completely random with no locality of data access. Consequently, the SSDs had to wait for I/O activity to complete on the 10K SAS drives and the maximum IOPS delivered was decreased to around 10,000 IOPS.

5.3.3 ORION (4 TB array capacity utilization, 100% random I/O and no data locality)

The ORION and Vdbench tests described in sections 5.3.1 and 5.3.2 were run using 8K I/O block size and a read-write ratio of 70/30. This represented a typical TPC-C like workload.

TPC-E is a highly read-intensive workload typically represented by 90% read and 10% write I/O. To determine the baseline performance, the I/O ratio was fixed to 92/8 read/write with 8K I/O block size in this ORION test. Also, 4 TB of the total array capacity was used while running this test.

This I/O ratio was fixed based on the I/O ratio observed when the actual TPC-E transactions were run on the Oracle RAC database. Refer to section 7.3 for more details.

This test on the array produced approximately 7,167 IOPS with less than 12ms latency on the storage array as shown in the SAN HQ chart in Figure 9.

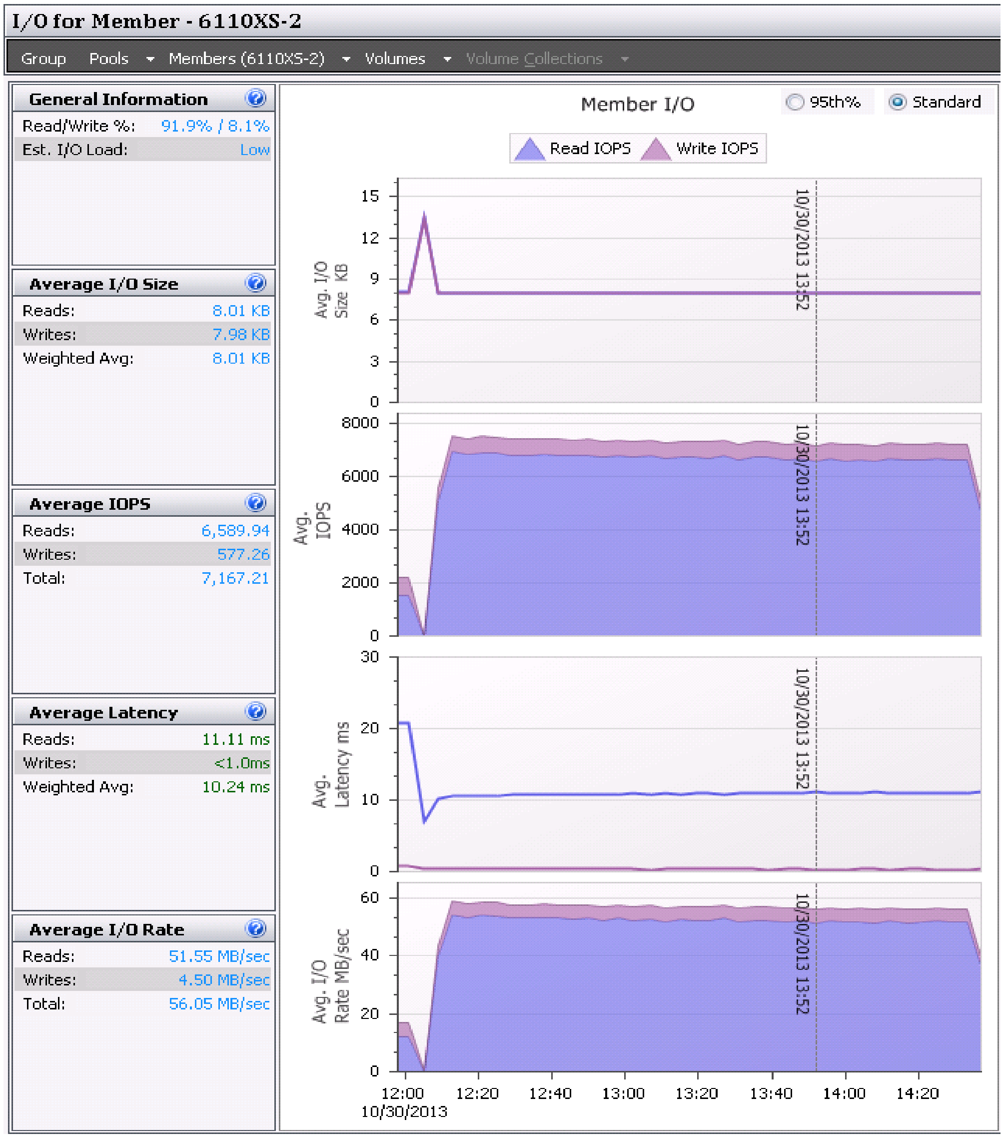


Figure 9 IOPS and latency reported by SAN HQ (4 TB capacity utilization and no data locality)

The raw I/O performance from this simulation test is compared with the I/O performance when the TPC-E-like workload was run on the actual database in section 6.3.

5.3.4 Vdbench (2 TB array capacity utilization, 100% random I/O and 20% data locality)

Applications such as Oracle OLTP databases typically contain both infrequently accessed and highly dynamic data within a single volume. The advanced tiering feature on PS6110XS arrays categorizes and operates on data workloads within each volume.

These workloads are automatically categorized and placed as appropriate onto either SSD or conventional hard-drive tiers based on the performance requirements on a PS6110XS hybrid array. Typically, in database applications, there are always specific areas on the disks such as REDO logs which are accessed more frequently than the other database objects. These frequently accessed objects usually need higher performance than other database objects.

In order to test the I/O scenario with a specific locality and frequency of I/O access, the Vdbench I/O generation tool was used to simulate I/O. The ORION tool was not appropriate for this study because there is no option to specify that the I/O workload accesses a specific locality of data on the disk. A Vdbench test was run with 20% locality of reference to evaluate the tiering benefits.

The test result showed 7,000 IOPS at the beginning and then gradually increasing to 12,000 over the test time span. As the test was run, specific portions of the disks were accessed more frequently, and as a result, more hot pages got created and automatically moved to the SSD tier. This resulted in significant performance improvement.

The SAN HQ chart in Figure 10 shows the gradual increase in IOPS from 7,000 to 12,000 over the test time span due to the automatic tiering feature of PS6110XS arrays.

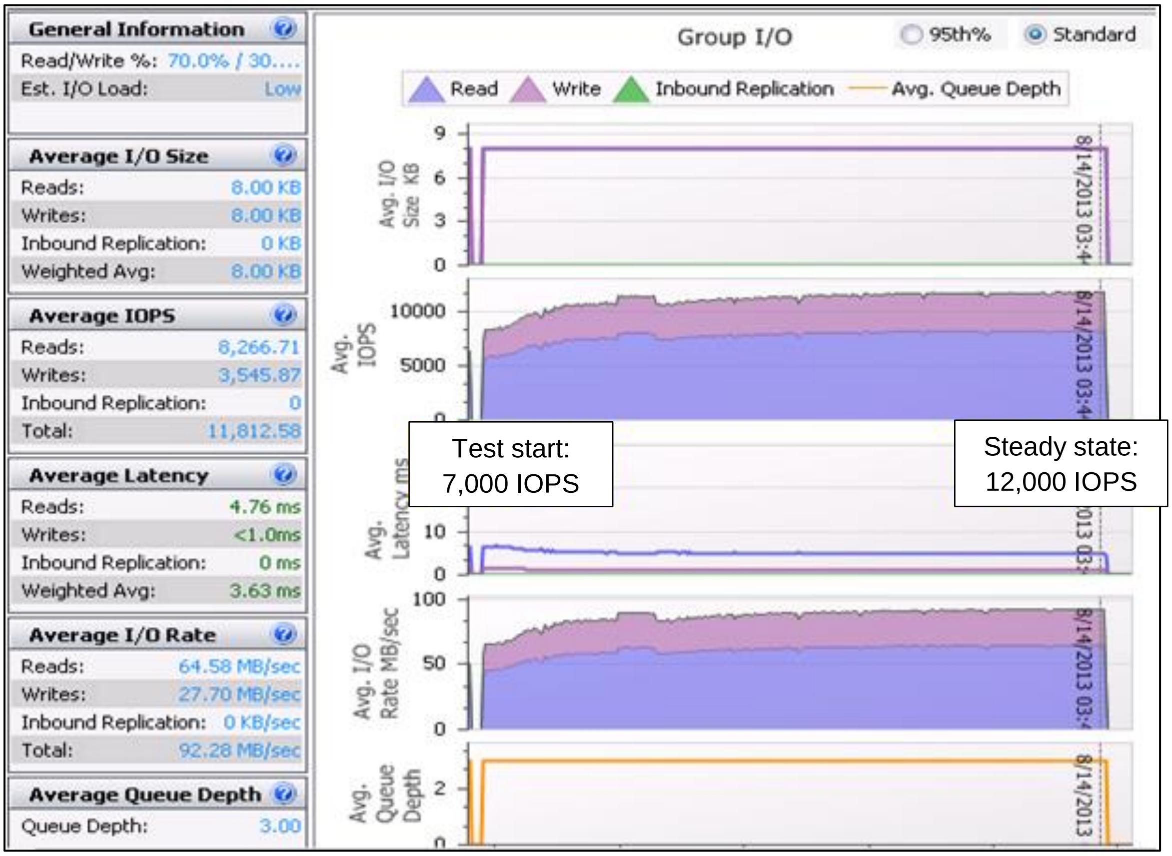


Figure 10 Increase in IOPS due to PS6110XS internal tiering (as reported by SAN HQ)

5.3.5 Number of volumes studies

The primary objective of these tests was to determine the PS Series volume configuration which produces optimal performance for the Oracle OLTP-type of workload.

These tests were run with 8K I/O block size to determine the maximum IOPS the storage array would sustain within the 20 ms latency (read and write latencies measured separately).

The total capacity utilization on the array was kept constant at 2 TB and only the volume configuration was modified as shown in Table 3.

Table 3 Test parameters: Volume configuration studies

Configuration parameters	
PS Series SAN	One PS6110XS (7 x 400GB SSDs and 17 x 600GB 10K SAS drives)
Total capacity utilization	2 TB
Volume configuration #1	Two volumes, 1 TB each
Volume configuration #2	Four volumes, 500 GB each
Volume configuration #3	Eight volumes, 250 GB each
RAID type	RAID 6 (Accelerated)
OLTP workload parameters	
I/O mix	I/O block size (KB)
70/30 read/write	8K (to simulate TPC-C database transactions)

Performance increased by more than 6% when the number of volumes was increased from two to four. Similarly, almost a 23% increase in performance was observed when eight volumes were configured on the array. The total capacity on the array was kept constant at 2 TB for all the volume configurations.

The results displayed in Figure 11 confirm higher performance when there are eight or more volumes. Increasing the number of volumes beyond eight did not result in any significant increase in IOPS.

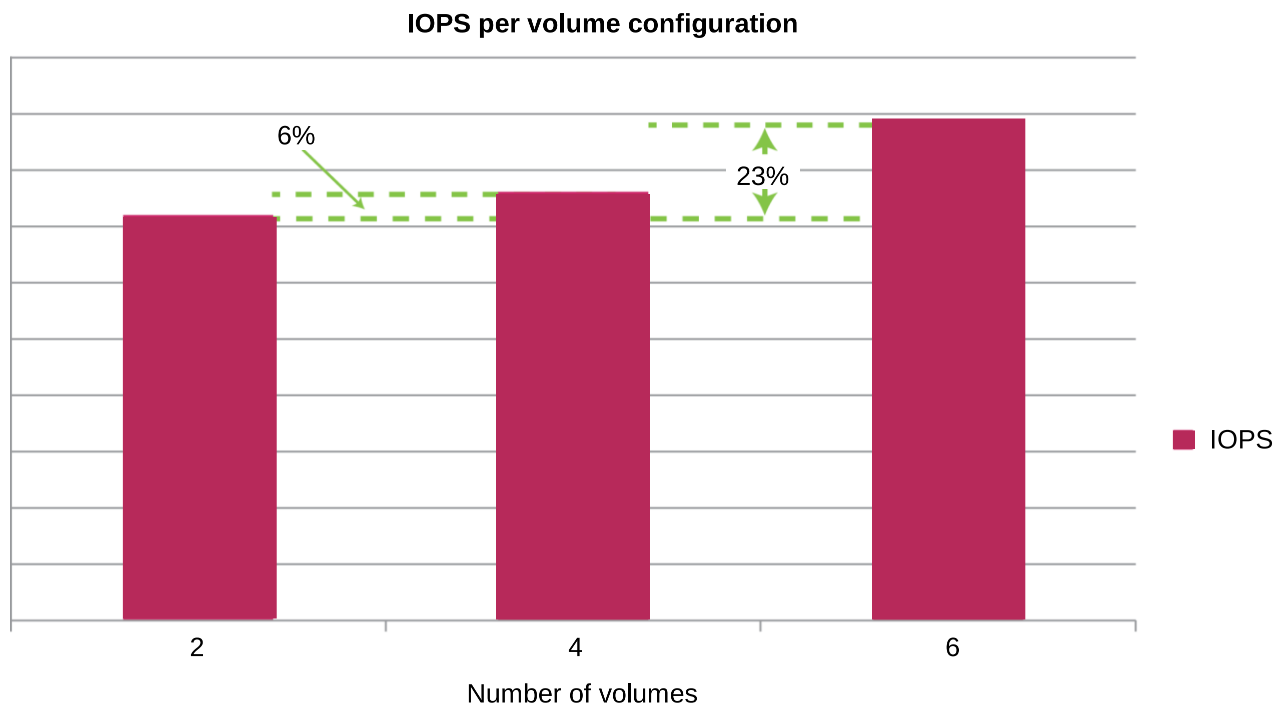


Figure 11 IOPS per volume configuration

> **Note:** A configuration with more small volumes is preferred over a low number of large volumes for better performance and simpler management of OLTP databases. As illustrated in Figure 11, there was 23% improvement in performance when the number of volumes were increased from 2 to 8.

5.3.6 Oracle ASM and non-ASM environments

Oracle recommends using ASM for managing OS disk volumes due to the advanced management capabilities. PS Series arrays provide load balancing among all available disk drives across the volumes within a storage pool. These two features complement each other to provide the most flexible and optimal performance configuration.

The primary objective of these tests was to determine the baseline I/O performance of the PS6110XS array in an ASM and non-ASM environment.

The ORION default option, CONCAT mode, was used to simulate non-ASM test scenarios. To evaluate the ASM configuration, disk volumes were simulated in RAID 0 mode by ORION, treating them as individual volumes and striping the I/O requests across them. Test results from both non-ASM and ASM configurations were compared.

Slightly higher IOPS (4,300 compared to 4,200) were observed with the ASM configuration. The IOPS numbers maintained the generally accepted disk latency limit of 20 ms (for both read and write latencies measured separately) for a random workload.

5.4 ORION/Vdbench on PS6110XS

The tests described in sections 5.3.1 to 5.3.3 were run with different capacity utilizations on the PS6110XS array. Also, the ORION tool was used to simulate 8K-block-sized I/Os with 70/30 read/write mix which are completely random in nature with no locality of data access.

The test results show that the array tries to use the SSDs first to provide the best performance. If the entire database does not fit into SSDs, then the data is spread across both SSDs and 10K SAS drives. The maximum IOPS on the storage array with 1 TB capacity utilization was 17, 000 IOPS. When the capacity utilization on the array was increased to 2 TB, the maximum IOPS on the array was 10,230. One more ORION test was run to simulate TPC-E like I/O transactions with heavy write intensive transactions and more than 7,000 IOPS were observed as illustrated in section 5.3.3.

The decrease in IOPS as the capacity utilization increased is expected due to the large dataset that is randomized and spread across more disk space. I/O generated by the ORION tool is completely randomized and there is no mechanism to specify locality of data access, so there was not much performance optimization to offer via the internal data tiering mechanism of the PS6110XS array.

However in typical OLTP database applications, most of the active data set is usually accessed frequently and contained within a relatively small locality on disks. In these scenarios, the internal tiering feature of PS6110XS array will identify the frequently accessed data and move the hot pages automatically to SSDs. This can create significant performance improvements. The Vdbench I/O generation tool was used to evaluate this scenario.

The Vdbench test was run to simulate 8K block sized I/Os with 70/30 read/write mix and 20% locality of I/O access. When a specific locality of I/O got accessed frequently, it was clear that the PS6110XS array was providing tiering benefits and the IOPS increased from 7,000 to 12,000 (refer to section 5.3.4 for more details).

Additional tests were run as described in sections 5.3.5 and 5.3.6 to evaluate the benefits of Oracle ASM and also to determine the volume configuration which produces optimal performance. Test results from sections 5.3.5 and 5.3.6 confirmed that using at least eight or more volumes per PS6110XS array member and ASM is recommended for achieving optimal performance for Oracle OLTP database workloads.

6 OLTP performance studies using TPC-E like workload

The ORION and Vdbench test results described in section 5 helped in determining the baseline performance of a typical Oracle OLTP database on PS Series arrays. Also, ORION and Vdbench are both synthetic I/O workload generation tools and they do not use the actual Oracle stack.

In this phase of testing, Quest Benchmark Factory for Databases simulated a real TPC-E-style workload on the test systems using an actual 2-node Oracle 11g R2 RAC database. The TPC-E-style workload was executed using PS6110XS hybrid arrays.

Some of the key configuration details for the database are described below:

- Used a single PS6110XS array for hosting Oracle database and REDO log files. One more PS6110E array was used to configure the archive and flashback recovery area.
- Installed Oracle 11g R2 clusterware on the two Dell EMC PowerEdge™ R720 database servers. The logical test configuration is described section 3.1.
- The database elements, REDO logs, archive, and flash data were managed using Oracle ASM. More details on disk layout are available in section 3.3.
- A TPC-E-like database was populated with a scaling factor of 50 using Quest Benchmark Factory.

6.1 Oracle memory management study

The most important and complex aspect of system tuning for maximizing Oracle database performance is memory management. With the Oracle 11g release, the Automatic Memory Management (AMM) feature was introduced to help automate and simplify this configuration tuning task by allowing you to manage the System Global Area (SGA) and Program Global Area (PGA) components using a single database initialization parameter called **MEMORY_TARGET**.

However, it remains a challenge for database administrators to determine the recommended value to be set for MEMORY_TARGET to achieve optimal database performance. Also, the memory requirements of a database server vary depending on the type of I/O workload. These tests were run using different AMM memory settings to determine the configuration which produces optimal performance for TPC-E-like I/O.

Table 4 Test parameters: Oracle memory management study

Configuration parameters	
PS Series SAN	1 x PS6010XV: • 16 x 300GB 15K SAS disks with dual 2-port 10GbE controllers • 14 SAS disks configured as RAID 10, two hot spares • Controller firmware version: 5.0.0.0 (R122845) 1 x PS6010S: • 16 x 100GB SSD disks with dual 2-port 10GbE controllers • 14 SSDs configured as RAID 10, with two hot spares • Controller firmware version: 5.0.0.0 (R122845) 1 x PS6110XS: • 7 x 400GB SSDs and 17 x 600GB 10K SAS drives • RAID 6 Accelerated • Hosting the database files and REDO logs 1 x PS6110E: • 24 x 2 TB SATA drives as RAID 50 • Hosting the ARCH logs and flash recovery area
Test configuration details	
Test configuration #1	AMM (MEMORY_TARGET = 38GB)
Test configuration #2	AMM (MEMORY_TARGET = 48GB)
Test configuration #3	AMM (MEMORY_TARGET = 64GB)
Database workload parameters	
Database	2-node RAC running on two R720 servers
User transactions	TPC-E from Benchmark Factory for Databases
Test duration	5 hours
Database size	~900GB (including tables, indexes)

By default, Oracle will set the MEMORY_TARGET value to 40% of the total memory available on the server. In the test configuration, each node (R720 server) had 96GB of memory. The first test assigned 38 GB of memory to Oracle. Later, more tests were run with different memory settings (38 GB, 48 GB, and 64 GB) as shown in Table 5 to determine the configuration which produces optimal performance.

A 9% increase in IOPS was observed when the memory was increased to 48 GB from the default setting of 38 GB. At the same time, the performance dropped when the memory was increased to a larger value such as 64 GB. A slight increase in IOPS was observed until the memory setting of 54 GB. After that, no additional improvements in performance were seen.

As illustrated in Figure 12, setting MEMORY_TARGET to a value between 50% and 60% of the total physical memory on the database server resulted in optimal database performance for the simulated TPC-E-like workload.

Note: Setting the MEMORY_TARGET value to 50–60% of the total memory on the server produced the optimal performance for TPC-E-style workload on the 2-node Oracle RAC database. This value can be treated as a guideline for tuning database memory but memory setting may vary depending on various factors such as type of I/O workload, number of databases, and database characteristics. It is recommended to monitor and tune the value depending on the type of I/O workload.

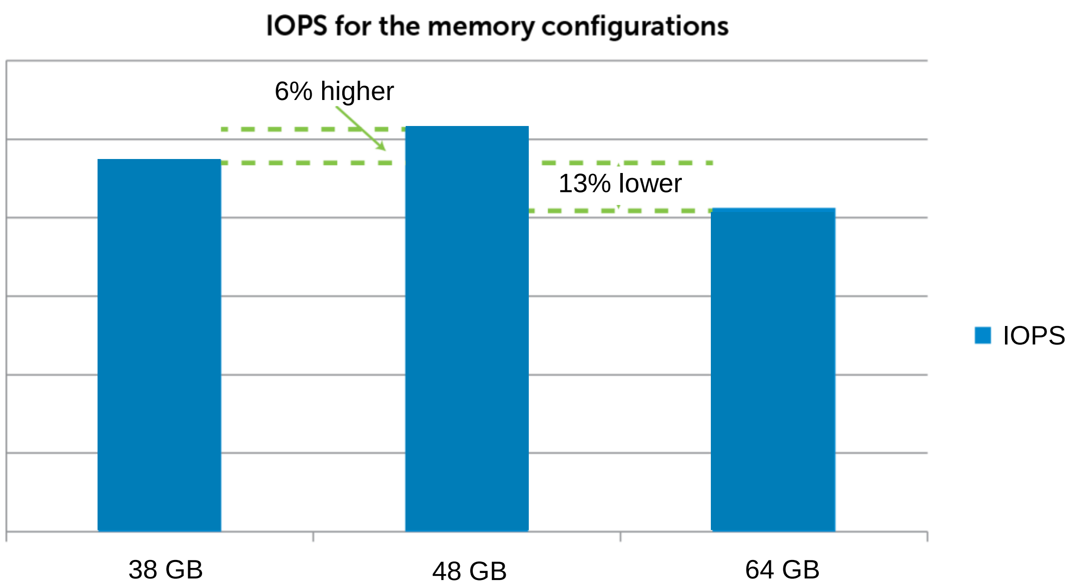


Figure 12 IOPS versus memory configuration

6.2 TPC-E (array capacity utilization study)

The performance from the storage array differed based on the capacity utilization as illustrated by the ORION test results in sections 5.3.1 to 5.3.3. The TPC-E tests were also run on two capacity configurations (array capacity utilization: 1 TB and 4 TB) to evaluate the performance implications. The two test configuration details are listed in Table 5.

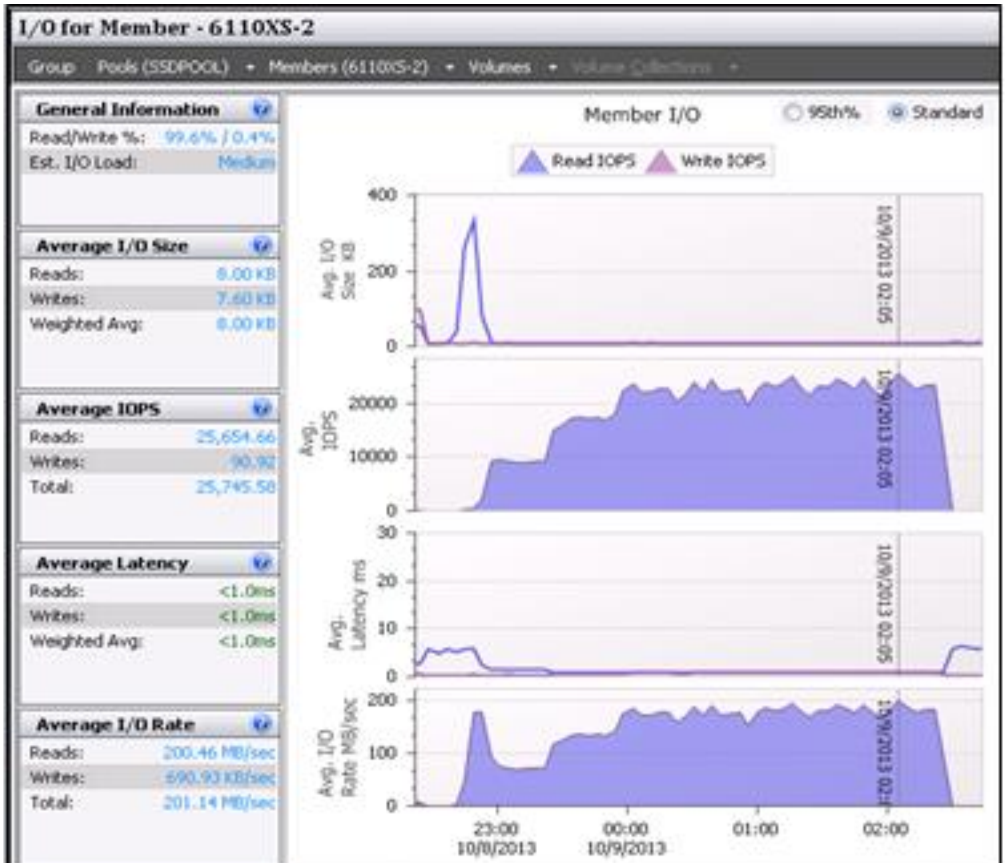
Table 5 Test parameters: Database file and volume layout studies

Configuration parameters	
PS Series SAN	One PS6110XS • 7 x 400GB SSDs, 17 x 600GB 10K SAS drives • RAID 6 Accelerated • Hosting the database files and REDO logs One PS6110E • 24 x 2 TB SATA drives as RAID 50 • Hosting the ARCH logs and flash recovery area
Test configuration details	
Test configuration #1	1TB array capacity utilization (4 x 250GB ASM disks)
Test configuration #2	4TB array capacity utilization (6 x 250GB ASM disks for database, 5 x 500GB with other data)
Database Workload Parameters	
Database	2-node RAC running on two R720 servers
User transactions	TPC-E from Benchmark Factory for Databases
Test duration	5 hours

The results from these two tests are analyzed in the sections 6.2.1 and 6.2.2.

6.2.1 1 TB array capacity utilization

Quest Benchmark Factory software simulated TPC-E transactions on the 2-node RAC database at varying user loads. The SAN HQ chart in Figure 13 shows the peak IOPS of 25,745 with less than 1.0 ms latency.



Member	Pool	Disk	Description	Average IOPS	Read I/O Rate	Write I/O Rate	Status	Average Disk Queue Depth	Model	Firmware Revision	Serial Number
6110XS-2	default	0	SSD 400GB SAS	3,623.13	28.59 MB/sec	300.44 KB/sec	online	2.27	LB406M	E312	40311296
6110XS-2	default	1	SSD 400GB SAS	3,634.06	28.68 MB/sec	304.61 KB/sec	online	2.28	LB406M	E312	40261744
6110XS-2	default	2	SSD 400GB SAS	3,630.09	28.65 MB/sec	300.44 KB/sec	online	2.28	LB406M	E312	40311300
6110XS-2	default	3	SSD 400GB SAS	3,630.05	28.64 MB/sec	287.92 KB/sec	online	2.28	LB406M	E312	40311928
6110XS-2	default	4	SSD 400GB SAS	3,634.83	28.68 MB/sec	304.61 KB/sec	online	2.28	LB406M	E312	40310120
6110XS-2	default	5	SSD 400GB SAS	3,633.18	28.68 MB/sec	300.44 KB/sec	online	2.28	LB406M	E312	40261612
6110XS-2	default	6	SSD 400GB SAS	3,625.80	28.61 MB/sec	300.44 KB/sec	online	2.27	LB406M	E312	40312252
6110XS-2	default	7	10K 600GB SAS	0	0 KB/sec	0 KB/sec	online	0	ST9600205SS	CN04	6XR3C2SS
6110XS-2	default	8	10K 600GB SAS	0	0 KB/sec	0 KB/sec	online	0	ST9600205SS	CN04	6XR3C5LS
6110XS-2	default	9	10K 600GB SAS	0	0 KB/sec	0 KB/sec	online	0	ST9600205SS	CN04	6XR3C6C7
6110XS-2	default	10	10K 600GB SAS	0	0 KB/sec	0 KB/sec	online	0	ST9600205SS	CN04	6XR3C60S
6110XS-2	default	11	10K 600GB SAS	0	0 KB/sec	0 KB/sec	online	0	ST9600205SS	CN04	6XR3D29X
6110XS-2	default	12	10K 600GB SAS	0	0 KB/sec	0 KB/sec	online	0	ST9600205SS	CN04	6XR3C6CZ
6110XS-2	default	13	10K 600GB SAS	0	0 KB/sec	0 KB/sec	spare	0	ST9600205SS	CN04	6XR3D20S
6110XS-2	default	14	10K 600GB SAS	0	0 KB/sec	0 KB/sec	online	0	ST9600205SS	CN04	6XR3C205
6110XS-2	default	15	10K 600GB SAS	0	0 KB/sec	0 KB/sec	online	0	ST9600205SS	CN04	6XR3BX4P
6110XS-2	default	16	10K 600GB SAS	0	0 KB/sec	0 KB/sec	online	0	ST9600205SS	CN04	6XR3BXA4
6110XS-2	default	17	10K 600GB SAS	0	0 KB/sec	0 KB/sec	online	0	ST9600205SS	CN04	6XR3C776
6110XS-2	default	18	10K 600GB SAS	0	0 KB/sec	0 KB/sec	online	0	ST9600205SS	CN04	6XR3BWQ2
6110XS-2	default	19	10K 600GB SAS	0	0 KB/sec	0 KB/sec	online	0	ST9600205SS	CN04	6XR3CYCJ
6110XS-2	default	20	10K 600GB SAS	0	0 KB/sec	0 KB/sec	online	0	ST9600205SS	CN04	6XR3BWXH
6110XS-2	default	21	10K 600GB SAS	0	0 KB/sec	0 KB/sec	online	0	ST9600205SS	CN04	6XR3C2F2
[illegible]XS-2	default	22	10K 600GB SAS	0	0 KB/sec	0 KB/sec	online	0	ST9600205SS	CN04	6XR3D08Z
6110XS-2	default	23	10K 600GB SAS	0	0 KB/sec	0 KB/sec	online	0	ST9600205SS	CN04	6XR3D289

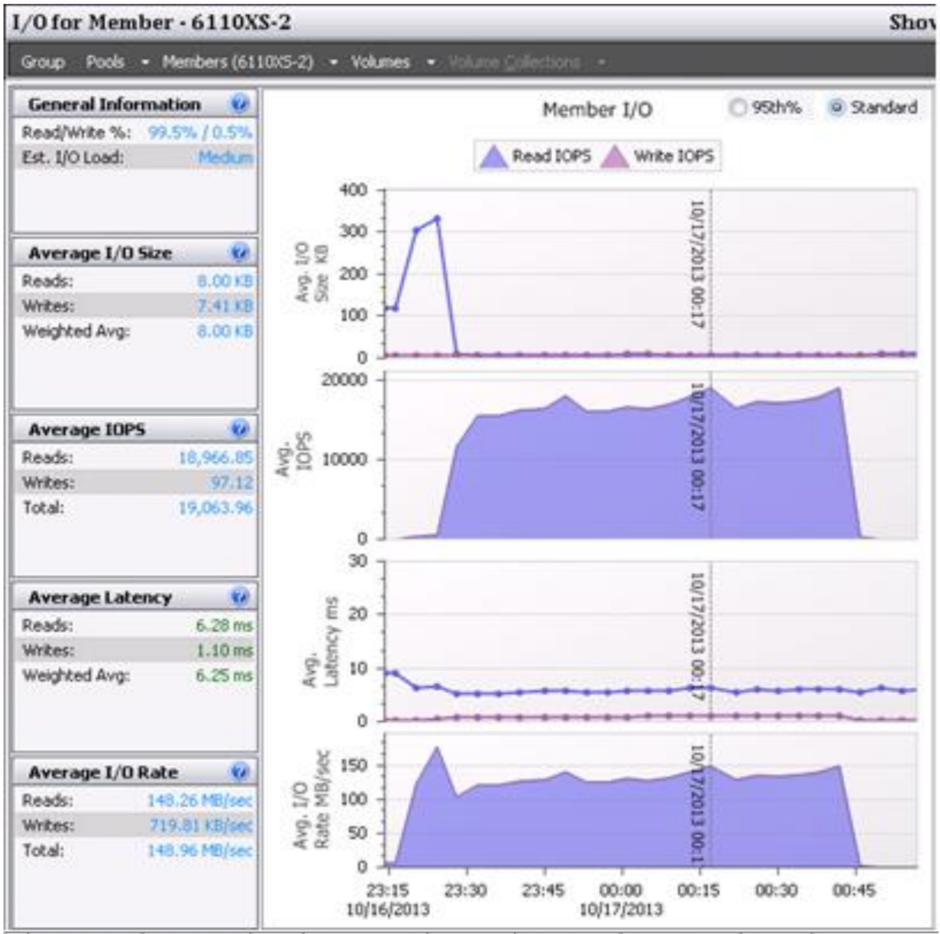
Figure 13 IOPS, latency, and disk IOPS reported by SAN HQ (TPC-E I/O with 1 TB array capacity utilization)

Also, as you can see from the IOPS activity on the disks, complete I/O activity was handled by the SSDs and there was no I/O activity on the 10K SAS drives. This is because the array capacity utilization was 1 TB and all the data was able to fit within the SSDs.

6.2.2 4 TB array capacity utilization

The entire database was able to fit within the SSDs in the scenario described in section 6.2.1. This is not always the case, and the data will be spread across both SSDs and 10K SAS drives when more capacity is used on the array. Obviously, the performance numbers will be different in this case because the I/O activity is randomly spread across both SSDs and 10K SAS drives.

This test evaluated the scenario in which the database size was increased from 900 GB to 1.5 TB, and it included five additional 500 GB volumes hosting other data resulting in a capacity utilization of almost 4 TB. As a result, the data was spread across both the SSD and 10K SAS drives. The SAN HQ chart in Figure 14 shows the peak IOPS along with read and write latencies.



Member	Pool	Disk	Description	Average IOPS	Read I/O Rate	Write I/O Rate	Status	Average Disk Queue Depth
6110XS-2	default	0	SSD 400GB SAS	1,836.53	14.43 MB/sec	232.68 KB/sec	online	1.68
6110XS-2	default	1	SSD 400GB SAS	1,837.67	14.43 MB/sec	232.68 KB/sec	online	1.68
6110XS-2	default	2	SSD 400GB SAS	1,840.72	14.46 MB/sec	236.84 KB/sec	online	1.68
6110XS-2	default	3	SSD 400GB SAS	1,841.36	14.46 MB/sec	232.68 KB/sec	online	1.68
6110XS-2	default	4	SSD 400GB SAS	1,836.48	14.42 MB/sec	236.84 KB/sec	online	1.69
6110XS-2	default	5	SSD 400GB SAS	1,833.01	14.40 MB/sec	236.84 KB/sec	online	1.68
6110XS-2	default	6	SSD 400GB SAS	1,841.03	14.46 MB/sec	236.84 KB/sec	online	1.68
6110XS-2	default	7	10K 600GB SAS	369.14	2.88 MB/sec	24.93 KB/sec	online	7.21
6110XS-2	default	8	10K 600GB SAS	374.86	2.94 MB/sec	33.24 KB/sec	online	7.85
6110XS-2	default	9	10K 600GB SAS	376.08	2.95 MB/sec	29.09 KB/sec	online	7.64
6110XS-2	default	10	10K 600GB SAS	375.79	2.94 MB/sec	29.09 KB/sec	online	7.54
6110XS-2	default	11	10K 600GB SAS	376.34	2.94 MB/sec	29.09 KB/sec	online	7.84
6110XS-2	default	12	10K 600GB SAS	378.99	2.97 MB/sec	37.40 KB/sec	online	7.70
6110XS-2	default	13	10K 600GB SAS	0	0 KB/sec	0 KB/sec	spare	0
6110XS-2	default	14	10K 600GB SAS	383.17	3.00 MB/sec	33.24 KB/sec	online	7.95
6110XS-2	default	15	10K 600GB SAS	380.84	3.02 MB/sec	41.55 KB/sec	online	7.84
6110XS-2	default	16	10K 600GB SAS	379.81	2.97 MB/sec	29.09 KB/sec	online	8.20
6110XS-2	default	17	10K 600GB SAS	377.83	2.96 MB/sec	37.40 KB/sec	online	8.04

Figure 14 IOPS, latency, and disk IOPS reported by SAN HQ (TPC-E I/O with 4 TB array capacity utilization)

As seen in the SAN HQ chart in Figure 14, I/O activity was observed on both the SSD and 10K SAS drives. A peak IOPS of 19,000 was observed compared to almost 27,000 as described in section 6.2.1 (1 TB array capacity utilization).

The total capacity utilization on the array was about 4 TB, so the data was spread across both the SSD and 10K SAS drives. The SSDs had to wait for the I/O to complete on 10K SAS drives, and as a result the average IOPS produced from individual SSDs dropped from 3,600 to 1,800 in this scenario.

The 19,000 IOPS represent the more realistic performance in customer environments, because most of the time at least 50% to 60% capacity of the array is utilized, which will force the data to be spread across both the SSD and 10K SAS drives on a PS6110XS array.

6.3 TPC-E I/O with increased write I/O

The I/O characteristics were studied when running the TPC-E workload with two database configurations in sections 6.2.1 and 6.2.2. Almost 100% read I/O was observed in both cases as shown in the SAN HQ charts (Figure 13 and Figure 14). The reasons for this behavior are:

- TPC-E is more read intensive with a 9.7:1 I/O read-to-write ratio, while TPC-C typically represents 1.9:1 read-to-write ratio. More detailed information about TPC-E characteristics is available in the document, *TPC-E vs. TPC-C: Characterizing the New TPC-E Benchmark via an I/O Comparison Study*.
- 48 GB of memory was allocated to the Oracle process on each node. This also helped in caching much of the I/O at the server memory.

Typically, storage arrays perform well with read I/O. Also, the memory on the server will help by caching the read I/O and helps to improve overall read performance. A more realistic test for any storage system is when we introduce write I/O transactions.

In this test, the write transactions were significantly increased using Quest Benchmark Factory software. Figure 15 shows the write-intensive transaction mix used to run this test compared to the default transactions.

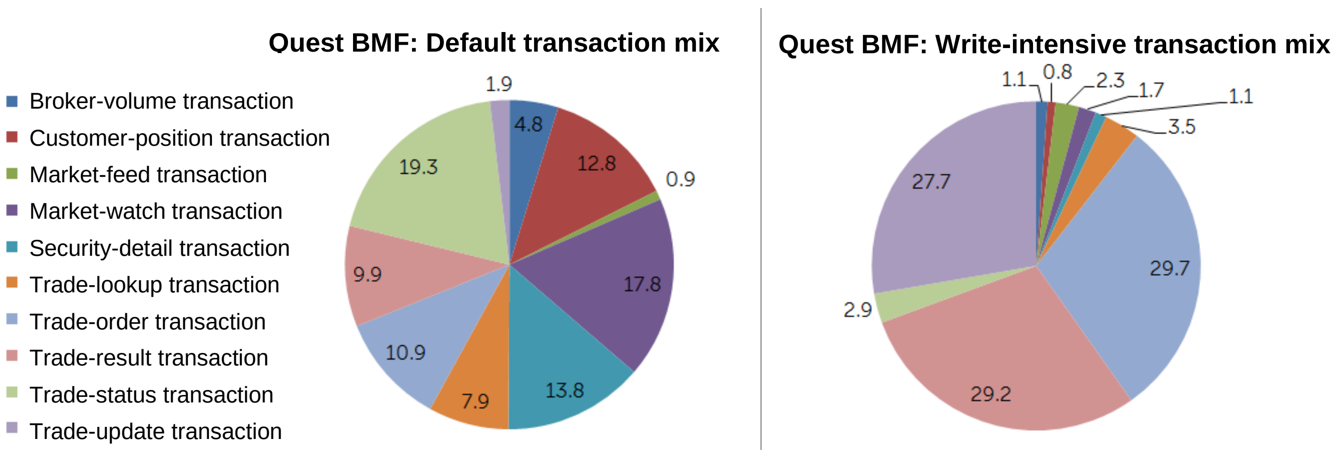


Figure 15 Quest BMF transaction mixes

TPC-E transactions were simulated using this modified transaction mix to determine the performance from a PS6110XS array when exposed to heavy write-intensive I/O workload, which is a worst-case scenario. No SQL queries were modified in this process.

Around 8,704 IOPS were observed with the read/write ratio of almost 90/10. The SAN HQ chart for this test is shown in Figure 16.

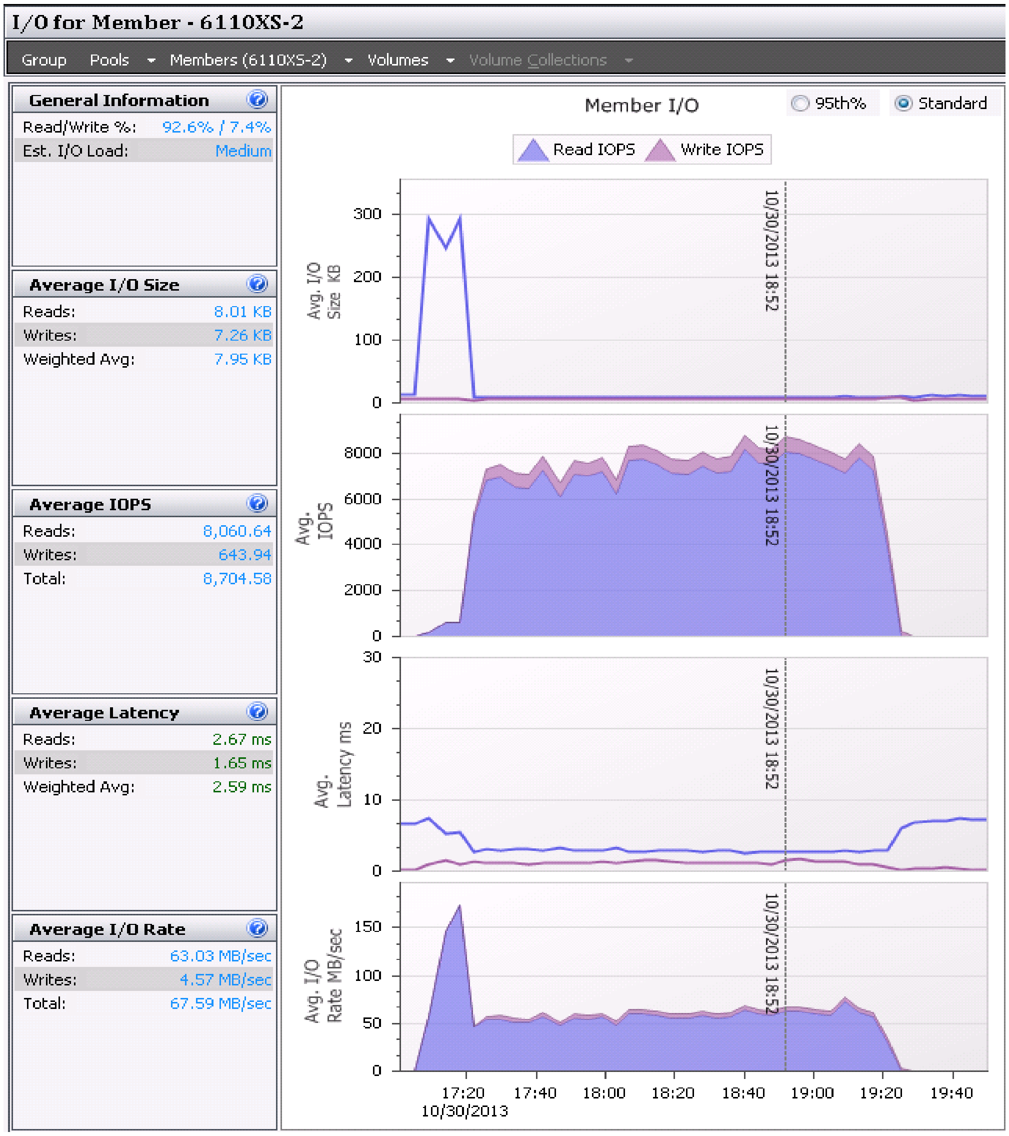


Figure 16 IOPS and latency reported by SAN HQ (TPC-E I/O with heavy write-intensive I/O transactions)

The average latency on the array was less than 3 ms at the peak I/O. These results show the PS6110XS arrays perform extremely well even for a modified heavy write-intensive TPC-E-like workload.

Oracle AWR reports were captured while running these tests and constantly monitored for any RAC or database related bottlenecks. The top 10 events of the AWR report are shown in Figure 17.

Top 10 Foreground Events by Total Wait Time

Event	Waits	Total Wait Time (sec)	Wait Avg(ms)	% DB time	Wait Class
db file sequential read	12,071,681	30.1K	2	75.3	User I/O
DB CPU		5246.1		13.1	
db file parallel read	137,735	2272.2	16	5.7	User I/O
enq: TX - row lock contention	23,043	1480.3	64	3.7	Application
SQL*Net break/reset to client	29,580	1216.5	41	3.0	Application
gc current block 2-way	510,664	183.5	0	.5	Cluster
gc current grant busy	438,666	152	0	.4	Cluster
gc cr block 2-way	337,532	113.6	0	.3	Cluster
log file sync	150,092	93.8	1	.2	Commit
gc cr block busy	26,505	33.8	1	.1	Cluster

Wait Classes by Total Wait Time

Wait Class	Waits	Total Wait Time (sec)	Avg Wait (ms)	% DB time	Avg Active Sessions
User I/O	12,211,912	32,401	3	81.1	9.0
DB CPU		5,246		13.1	1.5
Application	52,647	2,697	51	6.8	0.7
System I/O	1,015,736	2,155	2	5.4	0.6
Cluster	1,490,415	568	0	1.4	0.2
Commit	150,101	94	1	.2	0.0
Other	196,082	46	0	.1	0.0
Network	7,272,105	31	0	.1	0.0
Concurrency	45,408	10	0	.0	0.0
Configuration	35	0	1	.0	0.0
Scheduler	81	0	0	.0	0.0

Figure 17 Top 10 events in AWR report

As you can see from the AWR report, the top event is **db file sequential read,** which means that the disks were saturated and there were no other bottlenecks related to RAC or database configuration parameters.

6.4 TPC-E on PS6110XS

The PS6110XS arrays consists of 7 x SSDs and 17 x 10K SAS drives. The array has built-in intelligence to automatically move the data between SSD and SAS tiers within a volume based on the performance requirements.

The test configuration described in section 6.2.1 used a database of size 1 TB in which the entire database could fit in the SSDs and consequently, all the I/O activity was handled entirely by the SSDs. The PS6110XS array produced more than 25,500 IOPS in this scenario.

When the capacity utilization was increased to 4 TB, the IOPS changed from 25,675 to 19,000 (see section 6.2.2). This is because the entire data set could not be stored within the SSDs and had to be spread across all of the SSDs and 10K SAS drives. The SSDs also had to wait for the I/O activity to complete on the 10K SAS drives.

The PS6110XS arrays performed well, even with the modified heavy write-intensive TPC-E-style I/O operations as described in section 6.3. The PS6110XS arrays produced more than 8,700 IOPS with very heavy write operations introduced as part of TPC-E-modified workload.

Figure 18 compares the performance numbers produced by the ORION I/O performance numbers (see section 5.3.3) and the TPC-E-style workload with write-intensive I/O transactions simulated using Quest Benchmark Factory (see section 6.3).

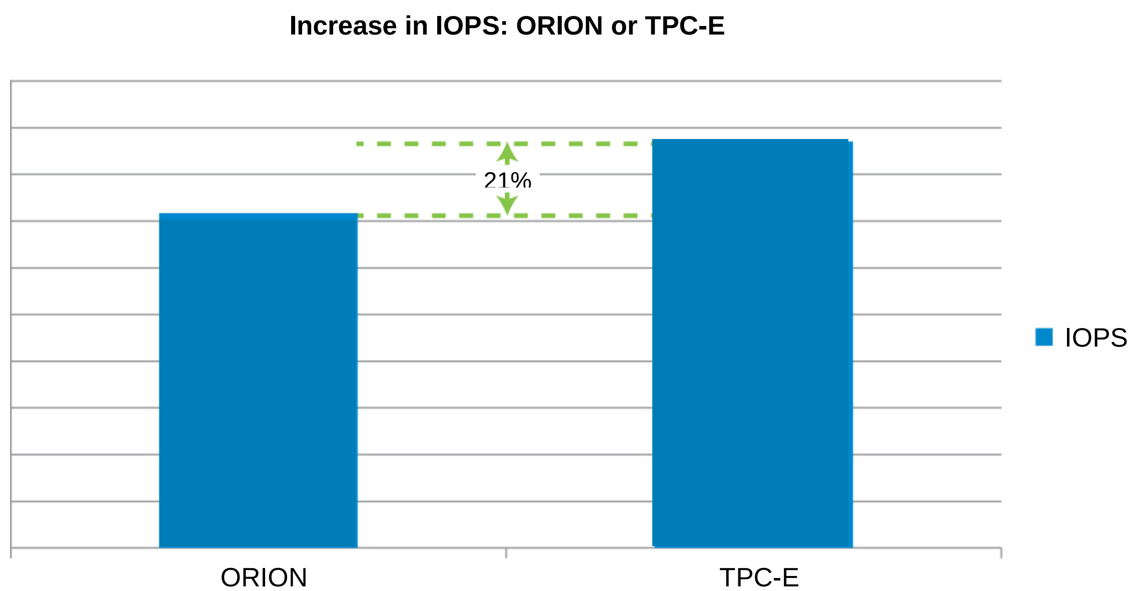


Figure 18 Increase in IOPS: ORION or TPC-E

ORION is a synthetic tool that provides an estimate of Oracle OLTP performance without actually installing an Oracle database, whereas TPC-E provides estimates of Oracle OLTP performance by executing the workload on actual Oracle databases.

Typically in database applications, there are specific areas on the disks such as REDO logs which are accessed more frequently than the other database objects. These frequently accessed objects usually need higher performance than other database objects.

Figure 18 shows more than a 21% increase in IOPS with the TPC-E workload. This is due to TPC-E reflecting the typical database I/O characteristics where a specific locality is being accessed frequently. On the other hand, the ORION tool could not simulate this behavior.

When a specific locality is accessed frequently, the PS6110XS internal tiering mechanism identifies the hot pages and then moves them to the SSD tier as demonstrated in section 5.3.4. These results clearly show how the internal data tiering mechanism of the PS6110XS hybrid array helps in Oracle OLTP database performance improvements while cost-effectively satisfying the capacity requirements simultaneously.

7 Best practice recommendations

7.1 Storage

A configuration with more small volumes is preferred over a low number of large volumes for better performance and simpler management of OLTP databases. It is recommended to use at least eight or more volumes per PS6110XS array member for achieving optimal performance. Increasing the number of volumes beyond eight may not result in significant performance improvements.

7.2 Network infrastructure

The following are network infrastructure design best practices:

- Design separate network infrastructures to isolate the LAN traffic from the SAN traffic (iSCSI).
- Design redundant SAN component architectures. This includes the NICs on the servers and switches for the storage network (including server blade chassis switches and external switches).
- Make sure that the server NIC ports and storage NIC ports are connected so that any single component failure in the SAN will not disable access to any storage array volumes.
- Enable flow control on both the server NICs and switch ports connecting to the server and storage ports.
- Enable Jumbo frames on the server ports and switch ports.
- On iSCSI SAN switches, spanning tree should be disabled on switch ports connecting to end devices like server and storage ports. The **Portfast** setting should be enabled in the switch configuration for these ports.

Note: General recommendations for PS Series array network configuration and performance are provided in the *Dell PS Series Configuration Guide*.

7.3 Operating system

The following are operating system best practices:

- 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 data transfers.
- It is recommended to use PS Series-aware MPIO functionality provided by the Host Integration Tools for the server operating system. HIT/Linux can be obtained from PS Series support. Configuring PS Series-aware MPIO provides enhanced performance benefits and also enables easier configuration by automating configuration of multiple required iSCSI sessions (connections) to the target volumes. 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 run the **eqltune** utility included with HIT/Linux and follow the OS-specific parameter recommendations listed by the utility.

7.4 Oracle database application

7.4.1 Database volume layout

The following are database volume layout best practices:

- Oracle recommends using ASM for simplified administration and optimal I/O balancing across disk groups.
- ASM external redundancy is preferred while configuring ASM disk groups. There is no need to use other types of redundancy because the PS Series array will provide the required redundancy.
- Use same-sized ASM disks within ASM disk groups. This helps Oracle ASM to optimally perform load balancing.
- 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.

7.4.2 Oracle AMM

The following are Oracle AMM best practices:

- The Oracle AMM feature assigns MEMORY_TARGET to be 40% of the total physical memory on the server. Based on the test results described in section 6.1, setting MEMORY_TARGET to a value 50% to 60% of the total server memory can be used as a guideline for TPC-E like workloads. The memory setting can vary based on various database factors, so careful monitoring and analysis should be done before tuning this parameter.
- The Oracle AMM feature requires the /dev/shm file system size to be at least equal to the value of the MEMORY_TARGET parameter. By default, the size of this file system will be set to 50% of the total physical memory.
- You need to change the size of the /dev/shm file system to support custom MEMORY_TARGET settings. Make sure that the database instances are shut down before making any changes to the size of /dev/shm. You can use the **mount** command as shown in the following command, which allows the MEMORY_TARGET value to be set up to 64 GB.

```
[root@oranode1 oswbb]# mount -t tmpfs shmfs -o size=65g /dev/shm
```

- Verify that the size of the /dev/shm file system is modified using the following command.

```
[root@oranode1 oswbb]# df -k
Filesystem           1K-blocks      Used Available Use% Mounted on
/dev/mapper/vg_oranode1-lv_root
                      51606140  26696252  22288448  55% /
tmpfs                 68157440         0  68157440   0% /dev/shm
/dev/sda1               495844     68132    402112  15% /boot
/dev/mapper/vg_oranode1-lv_home
                     201585632    195108 191150524   1% /home
shmfs                 68157440         0  68157440   0% /dev/shm
```

- The ASM instance on nodes will terminate as soon as the /dev/shm size is modified. Complete the following to bring up ASM instances on database nodes.

  Log in to each node as **grid** user and export ORACLE_SID with ASM instance. Enter the following commands to bring up the ASM instance. Repeat the same steps on all database nodes.

```
[grid@oranode1 bin]$ export ORACLE_SID=+ASM1

[grid@oranode1 bin]$ pwd
/u01/app/11.2.0/grid/bin

[grid@oranode1 bin]$ ./ASMCMD
-bash: ./ASMCMD: No such file or directory

[grid@oranode1 bin]$ ./asmcmd
Connected to an idle instance.

ASMCMD> startup
ASM instance started

Total System Global Area 1135747072 bytes
Fixed Size                  2260728 bytes
Variable Size            1108320520 bytes
ASM Cache                  25165824 bytes
ASM diskgroups mounted
ASM diskgroups volume enabled
ASMCMD> ls
DB1DATA/
DB1FLASH/
DB1LOG1/
DB1LOG2/
OCRVOTE/
ASMCMD> exit
```

7.4.3 Oracle ASM configuration on RHEL 6.3

ASMLib is a free, optional software feature from Oracle that simplifies the management and discovery of ASM disks, but Oracle changed the policy for RHEL 6, providing ASMLib software and support only for Unbreakable Enterprise Kernel and the Red Hat compatible kernel for Oracle Linux.

Red Hat released an open-source package, **kmod-oracleasm**, beginning with RHEL 6.4 that can be installed from the RHEL server supplementary channel. However, updates and support for this module must be received from Red Hat and this is not supported by Oracle.

Note: To deploy Oracle ASM on RHEL 6.x kernels without ASMLib support from Oracle, use the procedure described in the article, Configure Oracle ASM on EqualLogic without ASMLib on RHEL 6.x. The described procedure uses *udev* (device manager for Linux which manages all device nodes in /dev) rules on a RHEL 6 server to configure Oracle ASM.

8 Conclusion

Based on the tests explained in sections 5.1 through 5.4, the key findings while comparing performance of PS6010S and 6010XV arrays are summarized in the following.

- A 7x IOPS performance increase was measured for the PS6010S over the PS6010XV for 100% read 8K block I/O.
- The PS6010S can sustain a 4x IOPS performance increase over the PS6010XV at the 20ms disk latency threshold (16,000 compared to 4,000) for 70/30 read/write workloads.
- For block sizes as large as 64K, the single PS6010S array still provided a 3.3x performance increase over the PS6010XV.

The ORION I/O simulation tests and the TPC-E tests using Quest Benchmark Factory (executed with different array capacity utilization) illustrated the benefits of having both SSDs and 10K SAS HDDs. providing performance and as well as sufficient capacity needed by typical OLTP database applications. The Vdbench test results clearly demonstrated how the built-in internal tiering intelligence can help improve the database performance over time.

Optimal database performance 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 also must be ensured that the operating system and databases are configured with the optimal settings and design recommendations mentioned in this paper.

A Configuration details

This section contains an overview of the configurations used for the testing described in this document.

Table 6 Test configuration: hardware components

Solution configuration: hardware components		**Description**
Servers	2 x PowerEdge R720 servers: • Broadcom 5719 quad-port 1Gb NIC • Broadcom 57810 dual-port 10Gb NIC	• Two R720 servers as part of the Oracle RAC configuration • 4 x 1Gb NICs (B5719) trunked using Linux kernel bonding used as RAC interconnect • 2 x 10Gb NICs used for iSCSI SAN connectivity
Management server	1 PowerEdge R710 server: • Windows Server 2008	• Management server for the entire infrastructure • PS Series manager and SAN HQ
Network	2 x Dell EMC Networking S4810 switches: • Firmware: 9.1.0.0 2 x Dell EMC Networking S60 switches: • Firmware: 8.3.3.9	• S4810: Two switches stacked together for iSCSI SAN connectivity • S60: Two switches lagged together and used for RAC interconnect
Storage	1 x PS6110XS storage array: • 7 x 400GB SSDs, 17 x 600GB 10K SAS drives • Firmware: 6.2.1 • RAID 10 1 x PS6110E storage array: • 24 x 2 TB SATA drives • Firmware: 6.2.1 • RAID 50	• PS6110XS: Backend storage for Oracle database • PS6110E: Backend storage for Oracle flashback recovery area and VM repository

Table 7 Test configuration: software components

Test configuration: software components	
Database servers	• Red Hat Enterprise Linux 6.3 • Host Integration Tools for Linux 1.2.0 • Oracle Database 11g R2 Enterprise Edition (version 11.2.0.4) • Two-node Oracle RAC database • ASM for Clusterware and Database
Quest Benchmark Factory	• 2 x virtual machines used (one as the Benchmark Factory console and one for launching agents) • 2 x Windows Server 2008 R2 Enterprise Edition • Quest Benchmark Factory 6.9.2 • 2 agents from each VM • Oracle 11.2.0.3 client installed • Oracle 11.2.0.3 client installed • VMware vSphere ESX version 5.1 • Bare metal hypervisor directly installed on PowerEdge R810 servers • Managed by vCenter.
Monitoring tools	SAN Headquarters version 2.6.0 Oracle OS Watcher utility (installed on database servers) Oracle 11g R2 Automatic Workload Repository (AWR)

Additional resources

Dell.com/Support is focused on meeting customer needs with proven services and support.

Dell TechCenter is an online technical community where IT professionals have access to numerous resources for Dell EMC software, hardware, and services.

Storage Solutions Technical Documents on Dell TechCenter provide expertise that helps to ensure customer success on Dell EMC storage platforms.