



# Realize extreme performance and high availability with the HP Universal Database Solution

## HP ProLiant DL580 Universal Database Reference Architecture

### Table of contents

Executive summary .....	3
HP Universal Database Reference Architecture .....	4
Common foundation.....	4
HP ProLiant DL580 Gen8 Server .....	4
HP 3PAR StoreServ 7450 All-flash Storage Array .....	6
Solution architecture overview.....	8
Architectural diagram (non-virtualized and virtualized) .....	8
Capacity and sizing.....	11
HP ProLiant DL580 Gen8 Server configurations .....	11
HP 3PAR StoreServ SSD IOPS.....	11
HP 3PAR StoreServ configuration for the I/O and Oracle database workloads.....	11
Workload characterization.....	12
I/O characterization description for non-virtualized and virtualized.....	12
Oracle database workload for non-virtualized and virtualized.....	12
Oracle OLTP peak transactions for non-virtualized and virtualized.....	13
Workloads tuning considerations .....	14
Large block throughput for business intelligence workloads.....	15
Key features of the HP Universal Database Solution .....	15
High-availability solution deployment scenarios.....	16
Disaster recovery with HP 3PAR Remote Copy, Oracle Data Guard, and VMware vSphere .....	17
HP 3PAR Remote Copy .....	17
Oracle Data Guard .....	18
Oracle Real Application Clusters .....	18
VMware vSphere High Availability.....	18
VMware vSphere vMotion .....	18
VMware vSphere Distributed Resource Scheduler .....	19

Best practices .....	19
Analysis and recommendations.....	19
SAN recommendations.....	19
Server configuration best practices.....	19
Operating system configuration parameters.....	20
Storage configuration best practices .....	20
Database configuration best practices.....	20
ESXi host and virtual machine configuration best practices .....	20
Summary.....	21
Implementing a proof-of-concept.....	21
Reference architecture rack diagram .....	21
Bill of materials.....	22
Reference Architecture Bill of Materials .....	22
Appendices.....	24
Appendix A: Red Hat 6.5 kernel tunables <i>/etc/sysctl.conf</i> .....	24
Appendix B: Grub configuration for disabling C-states (virtual/physical) .....	25
Appendix C: IRQ affinity script for <i>/etc/rc.local</i> (physical only).....	26
Appendix D: HBA NUMA mapping and IRQ map (physical/virtual) .....	27
Appendix E: UDEV configurations .....	28
Appendix F—Storage information.....	32
Appendix G: Check or set operating system tracing parameter (physical/virtual) .....	36
Appendix H: Oracle parameters (physical/virtual) .....	37
Appendix I: Virtual machine configuration file (virtual only).....	38
For more information .....	44

## Executive summary

The demands of database implementations continue to escalate. Faster transaction processing speeds, scalable capacity, and increased flexibility is required to meet the needs of today's business. At the same time, enterprises are looking for cost-effective, open-architecture solutions that don't include vendor lock-in or carry the high price tag attached to single-vendor, proprietary solutions.

Customers today require high performance, highly available flexible database solutions without the high cost and rigidity of an "all-Oracle" solution. The HP Universal Database Solution: HP ProLiant DL580 Gen8 and HP 3PAR StoreServ 7450 All-flash Array combination running Oracle 12c, delivers just that by providing a fully tested flexible, high performance, and high-availability reference architecture. In addition to the performance, availability, and total cost of ownership savings of the non-virtualized HP Universal Database (UDB), solution you can reap additional efficiencies by deploying a large database instance on the virtualization reference architecture (VMware® vSphere 5.5 or higher) described in this paper.

The HP UDB is powered by the new HP ProLiant DL580 Gen8 Server, a highly reliable and scalable four-socket x86 server based on top-of-the-line E7-4800/8800 v2 processors by Intel®, and the HP 3PAR StoreServ 7450 All-flash Storage Array. While the HP UDB can be tailored to run any database software, this white paper focuses on configurations, results and best practices for the Oracle 12c database.

The reference architecture described in the paper delivers:

- Extreme online transaction processing (OLTP) performance (millions of transactions).
- More than one million IOPS with less than one millisecond response time and a throughput of over 10 GB/second.
- High performance, virtualized large database instances.
- Reduced overhead with minimum logical unit number (LUN) paths and inter NUMA-node communication.
- A host of features including high availability (HA) and disaster recovery (DR) for the highest levels of uptime.
- Unmatched configuration flexibility and choice.

Customer performance workload characteristics and requirements vary. HP solutions are tailored to provide maximum performance for various workloads without compromising availability commitments required by the business.

**Target audience:** This HP white paper is designed for IT professionals who use, program, manage, or administer large databases that require high availability and high performance. Specifically, this information is intended for those who evaluate, recommend, or design new IT high performance architectures. It includes details for Oracle 12c deployments requiring:

- Extreme performance and uptime
- Workload mobility and flexibility running as VMware vSphere 5.5 guests.

Oracle 12c database and Red Hat® installations are standard configurations except where explicitly stated in the reference architecture. This white paper describes testing performed in October and December 2014.

## HP Universal Database Reference Architecture

IT departments are under continuous pressure to add value to the business, improve existing infrastructures, enable growth opportunities, and reduce overhead. At the same time, exploding transactional data growth is driving database performance and availability requirements to entirely new levels. The demand for high speed and low latency, along with staggering volumes of transactional data, is prompting the adoption of new storage technologies that range from traditional disk to solid state.

Driven by the creation of new, high-value applications, customers are discovering the Oracle Exadata “one-size-fits-all” approach—one operating system, one database and one vendor—doesn’t do the job. Rather, Exadata requires extensive tuning, leads to high cost, and results in vendor lock-in.

In response, IT departments are looking for an “appliance-like” solution that provides a reliable foundation yet offers high-performance storage flexibility with a choice of OS and database. The HP ProLiant DL580 UDB Solution—optimized for the HP ProLiant DL580 Gen8 Server—delivers increased value with better performance and lower costs.

This white paper outlines the architecture and performance you can expect from the HP Universal Database Solution built on HP ProLiant DL580 Gen8 Servers and the HP 3PAR StoreServ 7450 All-flash Array running Oracle 12c. This solution offers the latest, most robust reference architecture developed under the HP UDB portfolio for mission-critical database solutions.

This reference architecture focuses primarily on the design, configuration, and best practices for deploying a highly available extreme-performance Oracle database solution. The Oracle and Red Hat installations are standard configurations except where explicitly stated in the reference architecture.

## Common foundation

### HP ProLiant DL580 Gen8 Server

**Figure 1.** The HP ProLiant DL580 Gen8 Server



The HP ProLiant DL580 Gen8 Server is a four-socket (4S) scale-up platform designed for performance, efficiency, and reliability. The server architecture and key technologies make the DL580 Gen8 Server an excellent choice for enterprise customers seeking enhanced performance and investment protection.

To meet the requirements for demanding and data-intensive workloads, the architecture includes key advancements in processor performance, memory capacity, I/O bandwidth, and storage performance. The HP ProLiant DL580 Gen8 Server has security and data protection features for system resiliency that your business can depend on. Unique features such as HP Advanced Error Recovery and HP Memory Quarantine increase memory and processor reliability by more than 30 percent. All these features make the DL580 Gen8 Server an ideal choice for handling large OLTP and decision support databases.

The DL580 Gen8 Server accommodates up to four Intel® Xeon® E7 4800/8800 v2 processors, up to 6 TB memory (with 64 GB DIMMs), and nine full length/full height PCI Express (PCIe) Gen3 slots. This expansive ability makes the DL580 Gen8 Server well-suited for the following environments:

- Large enterprise database—Both OLTP and decision support.
- Server consolidation and virtualization—Within a virtual environment, a server’s high-bandwidth and large memory capacity provide increased execution speed of multiple applications.
- High performance numeric computing—The server easily handles the high I/O bandwidth requirements of these applications through its capability of up to five double-wide x16 general-purpose computing on graphics processing units (GPGPUs).

**HP ProLiant DL580 Gen8 NUMA technology**

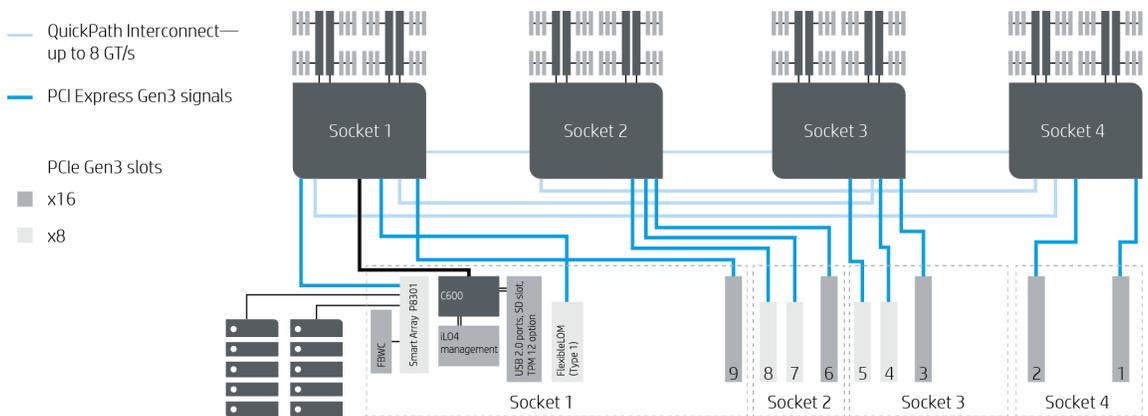
The HP DL580 Gen8 Server is ideal for scale-up database implementations. Its modular design provides the flexibility to adapt the configuration to meet the demands of your dynamic environment. The basis for the HP DL580 Gen8 Server architecture is the Intel Xeon E7-4800/8800 v2 processors that support Non-Uniform Memory Architecture (NUMA). Each processor has its own set of memory that connects directly to a set of peripherals.

Key processes in your system and databases can have individual CPU affinity for most efficient overall processing. Applications, which are NUMA-aware potentially have the capability to optimize system performance through NUMA control. For those applications that are not NUMA-aware, the affinity can be set manually. VMware vSphere is NUMA-aware and makes virtual machine (VM) placement decisions to take the best advantage of NUMA performance. It is also possible to set affinity for vSphere manually.

Figure 2 shows an architectural view of the HP DL580 Gen8 Server. It also shows the four NUMA node boundaries that overlap with each CPU socket. The balanced I/O slots provide direct I/O access from each node. This diagram can be used to determine balanced storage host bus adapters (HBAs) and PCIe flash card placements.

For detailed technical specifications, please refer to the [“HP ProLiant DL580 Gen8 Server technology technical white paper.”](#)

**Figure 2.** The HP ProLiant DL580 Gen8 Server architecture



## HP 3PAR StoreServ 7450 All-flash Storage Array

**Figure 3.** Front view of the HP 3PAR StoreServ 7450



HP 3PAR StoreServ 7450 is an all-flash based storage array that combines accelerated performance with all the enterprise-class, Tier 1 features and functionality expected for mission-critical application environments. This flash-optimized architecture relies on several unique HP 3PAR StoreServ innovations:

- Mesh-active architecture—Fine-grained virtualization and system-wide striping.
- Purpose-built HP 3PAR StoreServ ASIC—Supports mixed workloads with extremely high performance levels.
- HP 3PAR Adaptive Read and Write—Matches host I/O size reads and writes to flash media at a granular level to avoid unnecessary data reads and writes to reduce latency.
- Autonomic cache offload—Reduces cache bottlenecks by automatically changing the frequency at which data is offloaded from cache to flash media based on utilization rate.
- Multi-tenant I/O processing—Enables performance improvement for mixed workloads by breaking large I/O into smaller chunks.

A unique suite of persistent technologies power HP 3PAR StoreServ 7450 Storage in delivering high availability and Tier 1 resiliency to performance-critical applications. These include:

- HP 3PAR Persistent Cache—Preserves service levels so they are not impacted by unplanned component failures.
- HP 3PAR Persistent Ports—Allows non-disruptive upgrades without relying on multipathing software and without initiating failover.
- HP 3PAR Peer Persistence—Ability to federate storage across data centers without being constrained by physical boundaries.
- HP 3PAR StoreServ Data at Rest Encryption—Protects data from both internal and external security breaches.
- Flash-based media failure reconstruction—This enables the system to provide consistent performance levels even under situations of flash media failure.
- HP 3PAR Remote Copy software enables low recovery time objectives (RTOs) and zero data-loss recovery point objectives (RPOs) with complete distance flexibility.

### Solid-state drives

The HP 3PAR StoreServ 7450 All-flash Array offers the following types of solid-state drives (SSDs) in either SFF or LFF profile. The reference architecture uses the SFF enclosures and 100 GB drives.

**Table 1.** Solid-state drives supported in the HP 3PAR StoreServ 7450

Supported Solid-state drives	
100 GB 6G SFF SLC SAS Solid-state drive	480 GB MLC SAS Solid-state drive
200 GB 6G SFF SLC SAS Solid-state drive	920 GB MLC SAS Solid-state drive
100 GB 6G LFF SLC SAS Solid-state drive	400 GB 6G SFF MLC SAS Solid-state drive
200 GB 6G LFF SLC SAS Solid-state drive	1.92 TB 6G SFF cMLC SAS Solid-state drive
400 GB 6G LFF MLC SAS Solid-state drive	

For this reference architecture, any of the above types of drives can be chosen. It is recommended that they be used in groups of four drives per enclosure. This would mean a minimum of 16 drives per drive type.

### Unique features of the HP 3PAR 7450 StoreServ Array

*HP 3PAR Thin Suite (thin provisioning, thin persistence, and thin conversion)*

The HP suite of technologies includes.

- Thin provisioning—Thin Provisioning allows users to allocate virtual volumes to servers and provision only a fractional or part of the physical storage in the volume. This allows for maximum efficiency in capacity utilization, saving a considerable amount of investment in storage resources being stranded and as data provisioned but not used.
- Thin conversion—This feature allows users to convert a fully-provisioned set of volumes to thinly-provisioned volumes. For instance, if a volume was created with the intent of using most of the space, but circumstances resulted in most of the space not being used, the volume can be converted to a thin-provisioned volume. This results in tangible space and cost savings.
- Thin persistence—Thin persistence is a technology within the HP 3PAR StoreServ Arrays that detects zero valued data during data transfers. When data not being used in the volume is identified, it can be reallocated to free-to-use status. If data is removed from an application volume and those addresses are set to zero, thin Persistence can free them. Oracle developed an Automatic Storage Management (ASM) Storage Reclamation Utility (ASRU) for zeroing out data in an Oracle ASM disk group. This tool can be run and then thin persistence will detect the zeros and free up the data. For more information about HP 3PAR Thin Provisioning for Oracle and the ASRU utility, see [“Best Practices for Oracle and HP 3PAR StoreServ Storage.”](#)

#### *HP 3PAR Remote Copy*

HP 3PAR Remote Copy software brings a rich set of features and benefits that can be used to design disaster tolerant solutions that cost-effectively address availability challenges of enterprise environments. HP 3PAR Remote Copy is a uniquely easy, efficient, and flexible replication technology that allows you to protect and share data from any application.

Implemented over native IP—through Gigabyte Ethernet (GbE) or Fibre Channel (FC)—users may choose either the asynchronous periodic or synchronous mode of operation to design a solution that meets their requirements for RPO and RTO. With these modes, HP 3PAR Remote Copy allows you to mirror data between any two HP 3PAR StoreServ Storage systems, eliminating the incompatibilities and complexities associated with trying to mirror between the midrange and enterprise array technologies from traditional vendors. Source and target volumes may also be flexibly and uniquely configured to meet your needs, using, for example, different RAID levels, thick or thin volumes or drive types.

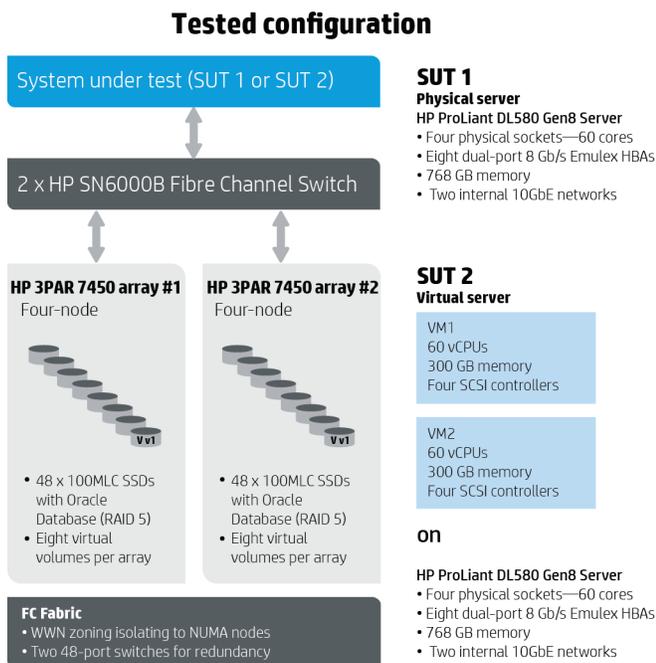
For more information, refer to [“Replication solutions for demanding disaster tolerant environments.”](#)

## Solution architecture overview

### Architectural diagram (non-virtualized and virtualized)

Figure 4 shows an architectural diagram of the tested HP UDB solution using the HP DL580 Gen8 Server and two HP 3PAR StoreServ 7450 All-flash Arrays. The configuration is a good example setup for most scale-up database customers. This configuration is the basis for several other variant configurations which provide the flexibility to meet the need at hand.

**Figure 4.** HP UDB solution architecture



#### HP ProLiant DL580 Gen8 Server

Our testing used two HP ProLiant DL580 Gen8 (one for physical and one for virtual) with 4 x Intel Xeon E7-4890 v2 processors each. Each processor has 15 cores totaling 60 cores per server. Turning on hyper-threading in the HP DL580 Gen8 BIOS will enable two threads per core making 120 logical cores on the physical configuration.

On the virtual configuration we created two virtual machines on the HP DL580 Gen8 dedicated to the virtual environment. We assigned 60 vCPU to be able to use the hyper-threading of the 30 physical cores per each VM 30 vCPU. The physical system was equipped with 768 GB of memory. Out of total system memory, 512 GB of it was allocated to the operating system shared memory (SGA/PGA) to obtain OLTP results in physical environment. In the virtual solution, out of 768 GB of memory, each VM was equipped with 375 GB memory and remaining 18 GB memory has been left for ESXi. Out of the total system memory of the each VM, 512 GB of it was allocated to the operating system shared memory (SGA/PGA) to obtain the required OLTP Results.

Also installed in the HP DL580 Gen8 Server were eight dual-port Emulex 8 Gb/s Fibre Channel cards. The cards are placed within different NUMA nodes for best performance and scalability.

The environment is fairly simple from the standpoint of number of servers and storage units. The entire solution delivers one million IOPS on both physical and virtual and fits into a single rack with room for storage growth. The optional DL380 Gen8 Server for HP 3PAR StoreServ management is not included. See the bill of materials section for a rack view and details.

### HP SN6000B Fibre Channel Switch

The storage connection to the server is accomplished using two HP SN6000B Fibre Channel Switches in a completely redundant setup. For each HBA card, a SAN connection goes from the first port to the first switch and a redundant connection goes from the second HBA port to the second switch.

The switches were tightly zoned using single-initiator to single target World Wide Name (WWN) zoning. Each HBA port is connected to a single port on a single HP 3PAR StoreServ 7450 All-flash Storage node. This was done also considering the NUMA location of the HBA card. The goal is to create multiple paths for HA while also minimizing cross communication between NUMA nodes, the storage nodes and the volumes themselves. Too many paths can create unwanted latencies in the I/O subsystem of the operating system. Tight volume allocation and zoning to nodes improved I/O performance by 20 percent.

### HP 3PAR StoreServ 7450 All-flash—four-node storage arrays

The HP 3PAR StoreServ 7450 All-flash units used for this testing were four-node units. Each node pair has two additional disk enclosures. The SSDs were installed equally across the node pairs and expansion units. There were 12 x 100 GB SLC drives installed in each unit totaling 48 SSDs per HP 3PAR StoreServ 7450 Array. With two arrays, the total maximum number of SSDs tested on a single database was 96 drives in two HP 3PAR StoreServ 7450 All-flash Arrays.

### Server connection layout

The HP DL580 Gen8 Server has both I/O expansion modules installed to accommodate the FC HBA cards needed. For maximum performance, the dual port 8 Gb FC HBA cards are spread across three separate NUMA nodes (1, 2, and 3). The cards are connected only to x8/x16 slots in the server. This provides I/O throughput bandwidth for the tested solution as well as for additional storage array future expansion. Table 2 shows the HBA card NUMA node assignments and local CPUs belonging to each NUMA node.

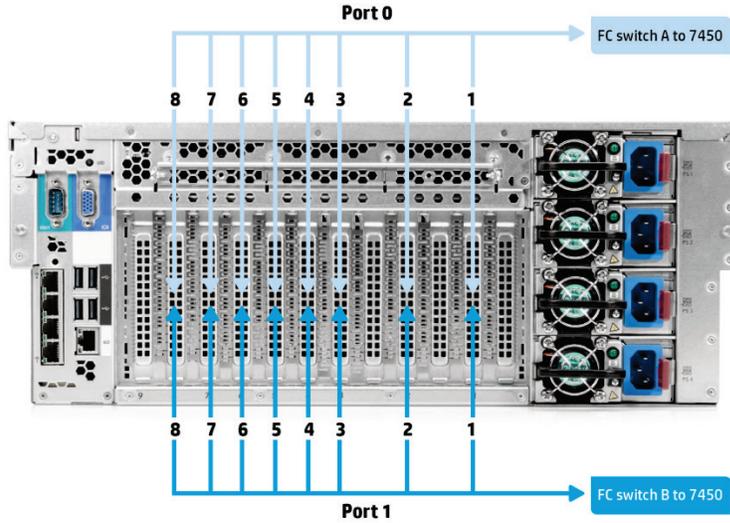
**Table 2.** HBA card placement in the ProLiant DL580 Server

NUMA node #	Card slots	Local CPU list
NUMA node 1	6, 7, 8	15–29
NUMA node 2	3, 4, 5	30–44
NUMA node 3	1, 2	45–59

Figure 5 shows the connection locations and FC connection mapping to the HP 3PAR StoreServ All-flash Array. All port 0 HBA connections go to switch A, and all port 1 HBA connections go to switch B. To achieve the maximum IOPS during I/O characterization testing, the connections to the virtual volumes needed to be isolated to specific NUMA nodes to minimize latencies in the operating system. Each connection has a specific single initiator to single target zone defined. See Appendix F for a zoning example.

The integrated 10GbE is used for connections back to the switches for client access. The user can choose whatever 10GbE infrastructure connection is required for their environment. The iLO connection is available on the HP DL580 Gen8 Server for remote management as needed.

**Figure 5.** The HP ProLiant DL580 Gen8 Server rear view connection diagram

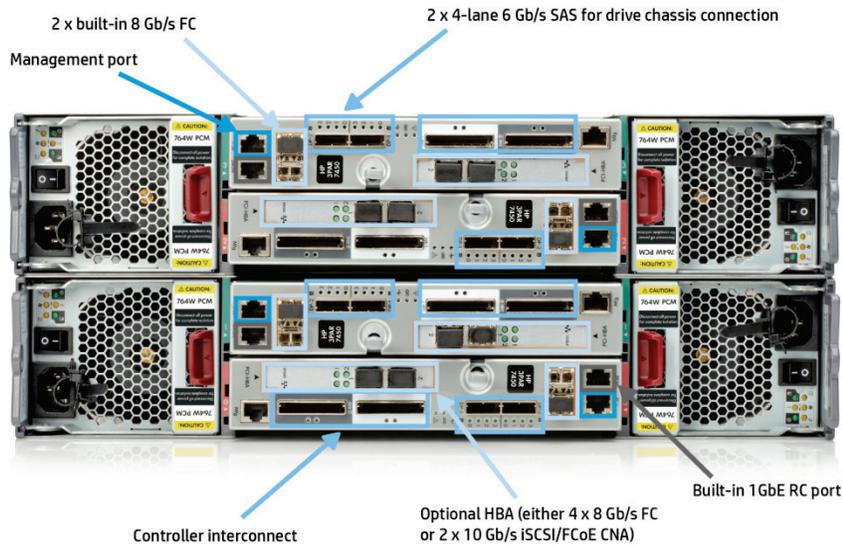


**HP 3PAR StoreServ 7450 All-flash—two-node pairs**

Figure 6 shows the rear view of two HP 3PAR StoreServ 7450 Node pairs. A four-node array has two node pairs. This tested solution used two four-node arrays with each array having two additional disk shelves. The SSDs in the array are evenly distributed across the disk shelves. So for 48 drives in the array, each enclosure holds 12 drives. With two arrays, this would be a total of 96 drives and 12 drives per enclosure. This leaves 12 more slots per enclosure open for capacity expansion.

The HP UDB solution uses the four-port 8 Gb/s FC option for additional FC ports to achieve the one million IOPS.

**Figure 6.** Two node pairs for the HP 3PAR StoreServ 7450 All-flash Array



## Capacity and sizing

### HP ProLiant DL580 Gen8 Server configurations

Depending on your application performance requirements, you have several options of processor, CPU and I/O card configurations. Please refer to [“HP ProLiant DL580 Gen8 Server technology white paper”](#) for available processors, CPU and I/O card configurations.

Additionally, the [“Best Practices When Deploying Linux® on the HP ProLiant DL580 Gen8 Server”](#) white paper outlines recommended HP software components for the DL580 Gen8 Server, and best practices you can use when deploying Linux.

### HP 3PAR StoreServ SSD IOPS

There is flexibility in the size of the SSDs used in the HP UDB. There is very minimal performance impact between the drives. Maximum throughput and IOPS are more dependent on the number of HP 3PAR StoreServ 7450 Arrays used. At least two arrays are recommended and required for one million IOPS and 10 GB/sec throughput. Using one array cuts the maximum IOPS and throughput in half. This solution was tested with two arrays; but additional arrays are supported.

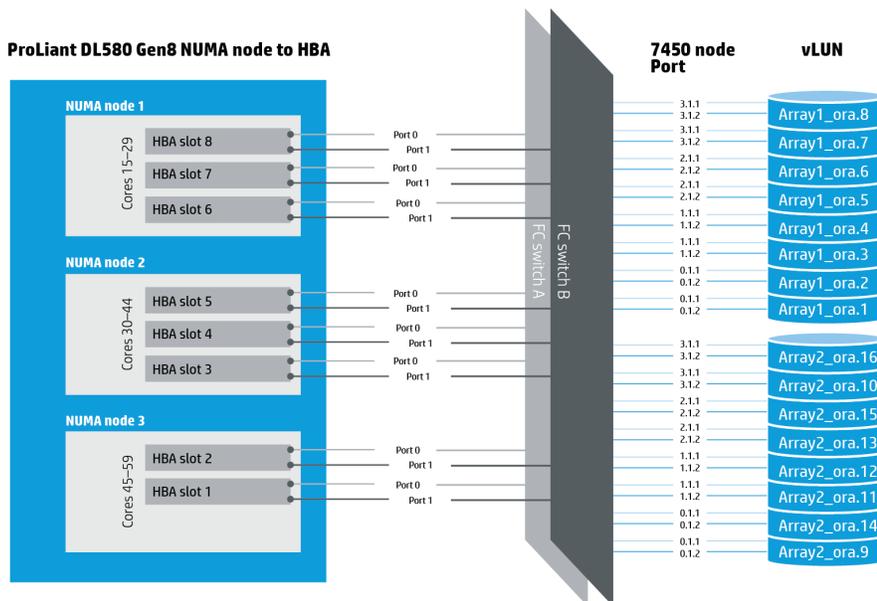
### HP 3PAR StoreServ configuration for the I/O and Oracle database workloads

#### Storage configuration for testing

The storage configuration involves the use of 16 virtual volumes. Each of the two HP 3PAR StoreServ 7450 Arrays had eight virtual volumes. Each virtual volume had two virtual logical unit numbers (vLUNs) (device paths). The device mapper on the server saw two paths for each virtual volume exported to the host. Figure 7 shows how the virtual volumes are mapped to the host for best performance. Extensive tests were run to achieve one million IOPS, as per the storage mapping and zoning shown in figure 7.

- The virtual volumes have two vLUNs paths per volume. Each virtual volume has both of its paths coming from HBAs belonging to the same NUMA node. Any one virtual volume never has two paths to different NUMA nodes, but only to different HBA cards within the node.
- Each port 0 on the HBA goes to switch A and port 1 goes to switch B. The zoning is tightly configured to single initiator to single target for maximum performance.
- Each array is identically configured and connected to the server. Any additional arrays would also be configured and connected the same way.
- For the Oracle configuration, a single ASM disk group was used for both DATA and LOGs.

**Figure 7.** Server to storage volume mapping to NUMA nodes



## Workload characterization

### I/O characterization description for non-virtualized and virtualized

All I/O characterization testing was performed with I/O generator tools capable of producing standard asynchronous I/Os using the Linux libaio libraries which are also used by Oracle and other database solutions. The tools are capable of generating many variations of workloads, thereby allowing flexibility in generating random and sequential access, various block sizes, various queuing and thread counts. In our testing, we demonstrated that the HP 3PAR StoreServ can deliver one million IOPS for small block size (4K) on read.

On each array, 48 drives were configured with RAID 5 capacity. These values are valid in the context of this UDB testing on the HP DL580 Gen8 Server.

The I/O characterization tests were run repeatedly and the storage system, fabric zoning and HP DL580 Gen8 Server were tuned for the purpose of determining maximum I/O performance. Specific information in this paper reflects the best practices for the tested configuration.

### I/O workload results for non-virtualized and virtualized

*I/O characterization results—OLTP random workloads—testing RAID5 48 drives per array*

The results of the characterization tests involving random 4K block reads revealed capabilities of more than one million IOPS using 48 drives per array.

Table 3 shows IOPS for a RAID5 configuration using the 48 drive set. The test used two HP 3PAR StoreServ 7450 All-flash Arrays, 96 drives total for both. RAID5 performs well with a read weighted workload.

**Table 3.** I/O characterization results

Random test	IOPS	Latency
4K Random reads (physical)	1.17 million	< 1 ms
4K Random reads (virtual)	1.12 million	< 1 ms

#### Note

All IOPS results documented in this paper were achieved using the server operating system (Red Hat Enterprise Linux [RHEL] release 6 update 5), NUMA and storage tuning mentioned in the recommendations and best practices. 8K random reads were also tested with physical and achieved more than one million IOPS. For virtual, testing was limited to 4K random reads and writes.

### Oracle database workload for non-virtualized and virtualized

The Oracle workload is tested using HammerDB an open-source tool. The tool implements an OLTP-type workload (60 percent read and 40 percent write) with small I/O sizes. The transaction results have been normalized and are used to compare UDB test configurations. Other metrics measured during the workload come from the operating system and/or standard Oracle Automatic Workload Repository (AWR) stats reports.

Tests performed on a 1 TB database (RAID5), included an I/O intensive OLTP test and a CPU and I/O intensive database test. The environment was tuned for maximum user transactions and maximum percentage database usage efficiency. After the database was tuned, the transactions were recorded at different user count levels. Because many workloads vary so much in characteristics, the measurement was made with maximum transactions.

Oracle Enterprise Database version 12.1.0.1 was used in this test configuration but other databases can be implemented on the HP Universal Database Solution.

## Oracle OLTP peak transactions for non-virtualized and virtualized

The Oracle test on non-virtualized and virtualized and virtual consisted of the creation of an OLTP database on RAID5 storage. The workload used a CPU and I/O intensive OLTP benchmark tool that could stress the server as well as the I/O subsystem. As the series of tests were run, the Oracle database *init* file was adjusted for maximum transactions and minimum physical I/Os.

The Oracle testing for the virtual solution consisted of two databases running on virtual machines, one per virtual machine. These databases were identical to the Oracle database used on the non-virtualized environment. The same workload was run on both configurations. The workload ran in parallel on the two virtual machines to simulate a consolidated solution on a scale-up architecture.

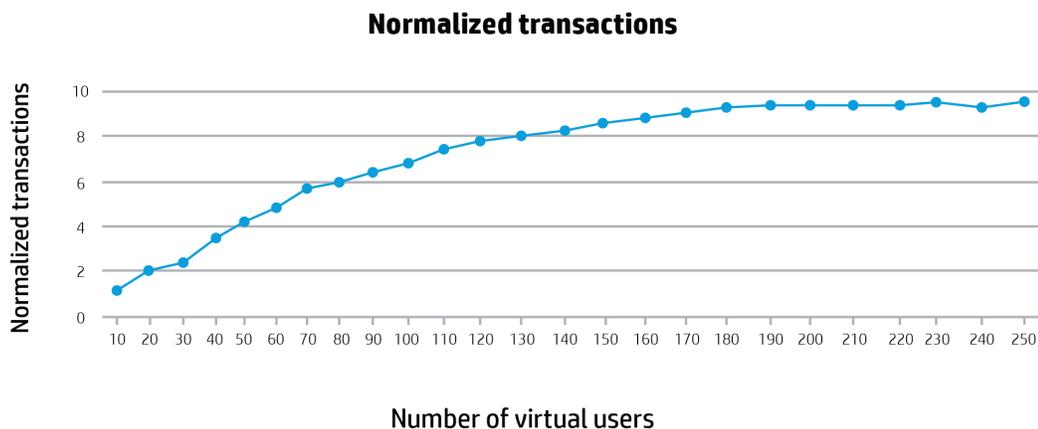
All of the specific Oracle tuning values are documented in Appendix H and best practices section of this paper. The ESXi tuning values are documented in Appendix I.

The databases use an ASM group created with default extent values. The DATA group consists of 16 volumes used for both data and logs on non-virtualized while 16 volumes were created per virtual machine for virtualized configuration. Typical recommendation is to separate DATA and LOG on different volumes, but as storage is all flash, we haven't noticed any significant gain in our tests in the past. This should be tested on individual implementations.

Workload ramped up from 10 users to 250 users on non-virtualized configuration. On the virtual environment we ramped up from 10 users to 125 users per virtual machine. Therefore, both the physical and virtual test platforms support the same number of Oracle users. In real database applications, the HP DL580 Gen8 Server handles tens of thousands of users but with the stress benchmark, each user is doing thousands of transactions per second with no latencies or think times. This is why the user count was not tested with more than 250 users on physical and 125 users per virtual machine. If the benchmark were doing a connection stress test, then the user count would be in the tens of thousands. The benchmark workload generally started ramping at 170 users on physical and 90 users on virtual and peaked at 230 users on physical and 100 users on virtual.

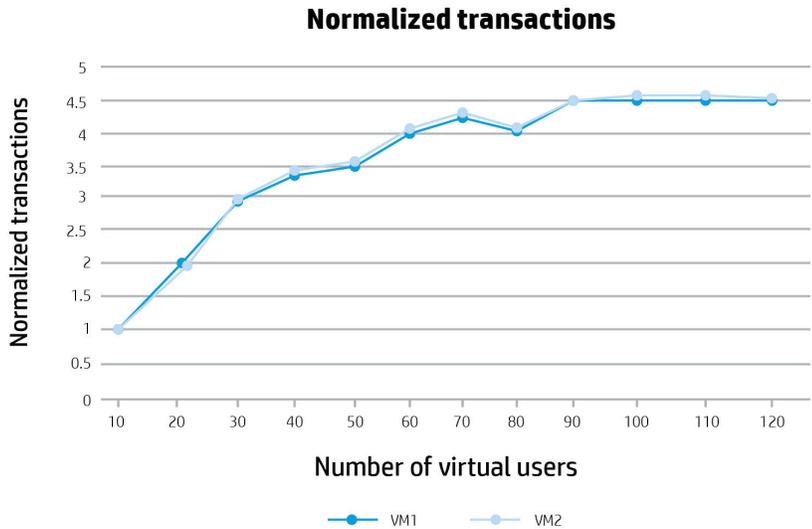
Initially, a single VM with 60 vCPUs was created on top of HP DL580 Gen8 Server with hyper-threading off. Oracle workload showed the transactions peaking at around 150 virtual user and the normalized transaction's counts were 39 percent below the physical results. To utilize the hyper-threading facility and to overcome the vSphere 5.5 limitation of number of CPUs per virtual machine, two VM's were created with 60 vCPU (30 physical core) with hyper-threading on. With this and a couple of ESXi tunings like Set Node affinity, Relocated the vCPU pinning from node 0 to node 1, Set CPU shares to high and Pinned PVSCSI Affinity (see Appendix I) the normalized transaction's counts were 9 percent below the physical results and the virtualized architecture were able to utilize the system resources efficiently to yield the much better results compare to a single VM configuration.

**Figure 8.** Oracle OLTP results for non-virtualized configuration



The Oracle workload shows the transactions peaking at around 230 virtual users on non-virtualized and 100 users on virtualized. On non-virtualized, the operating system utilization was 91 percent for user, four percent for system and less than one percent for I/O wait. The same measurement per virtual machine shows 96 percent for user, three percent for system and less than 1 percent for I/O wait. Figure 8 shows the normalized transactions calculated from transactions per second in Oracle AWR reports for non-virtualized and figure 9 for virtualized.

**Figure 9.** Oracle OLTP results for virtualized configuration



**Note**

All the transactional numbers have been normalized to demonstrate how the transactions scale up as the number of users increase.

**Workloads tuning considerations**

The server, storage, SAN and operating system parameters were adjusted to deliver best I/O and processing performance after several I/O characterization test iterations. The I/O characterization workloads were used to validate the best configuration to deliver the best I/O performance, thus validating the capabilities of the infrastructure. The storage capabilities are validated by HP’s Storage division and are specific to this configuration in these general areas:

- Server NUMA affinity—Minimize communication between NUMA nodes
- BIOS—HP DL580 Gen8 has BIOS optimizations for best performance
- Kernel and operating system—sysfs, sysctl kernel parameters
- Debug/tools—Disable processes or tools that can cause latencies
- I/O tuning—Provisioning, zoning, multipathing, special array settings, etc.

## Large block throughput for business intelligence workloads

Decision Support Systems (DSS) testing was not part of the scope of this paper but I/O throughput tests were run to measure the large block sequential capabilities of the HP 3PAR StoreServ 7450 All-flash Storage Array. We limited these tests to physical only. This was not motivated by any limitation on virtual. We wanted to show the I/O throughput on a four-socket ProLiant DL580 Server.

Table 4 shows the throughput results for sequential reads and writes with a one million block size. The result is certainly useful for considering a DSS implementation on UDB for HP 3PAR StoreServ 7450 All-flash. The DL580 is a proven solution for business intelligence (BI) workloads. The high HP 3PAR StoreServ 7450 All-flash throughput capabilities make it a very good match when doing large block queries with the Oracle database, and other databases as well.

**Table 4.** Sequential and random read access results for a one million block size

1 MB reads	Throughput
Random reads	10.5 GB/s
Sequential reads	10.8 GB/s

## Key features of the HP Universal Database Solution

Table 5 highlights specific data points of interest for the HP ProLiant DL580 Universal Database Solution.

**Table 5.** Specific features of the HP ProLiant DL580 Universal Database Solution

Attribute	Result
IOPS	One million—4K reads
Usable storage capacity data and redo	26.8 TB RAID5—400 GB MLC drive Two HP 3PAR StoreServ 7450 with 48 SSDs—96 drives total
Storage HA	Yes—redundant storage nodes and RAID protection
Server HA without performance impact	Yes—redundancy at server Server protection—HP Serviceguard
Data loss on single failure	No
Oracle Real Application Cluster (RAC) required	No
Duplicate copy of database	Yes (HP 3PAR Remote Copy or Oracle Data Guard)
Disaster recovery	Yes
Query standby database	Yes (Remote Copy)
Data device retention	Yes
Database storage	All flash
Storage thin provisioning	Yes
Storage thin persistence	Yes
Thin conversion	Yes
Volume snapshot	Yes

**Table 5.** Specific features of the HP ProLiant DL580 Universal Database Solution

Integrated Oracle backup solution	Yes
Dynamic optimization	Yes
Ease of management	Open IT tools
Operating system choice	RHEL, SUSE Linux, Oracle Linux, Microsoft® Windows®
Database choice	Flexible for other databases, tested with Oracle 12c using ASM on Grid Infrastructure.

## High-availability solution deployment scenarios

### High-availability clustering with HP Serviceguard for Linux or Oracle RAC

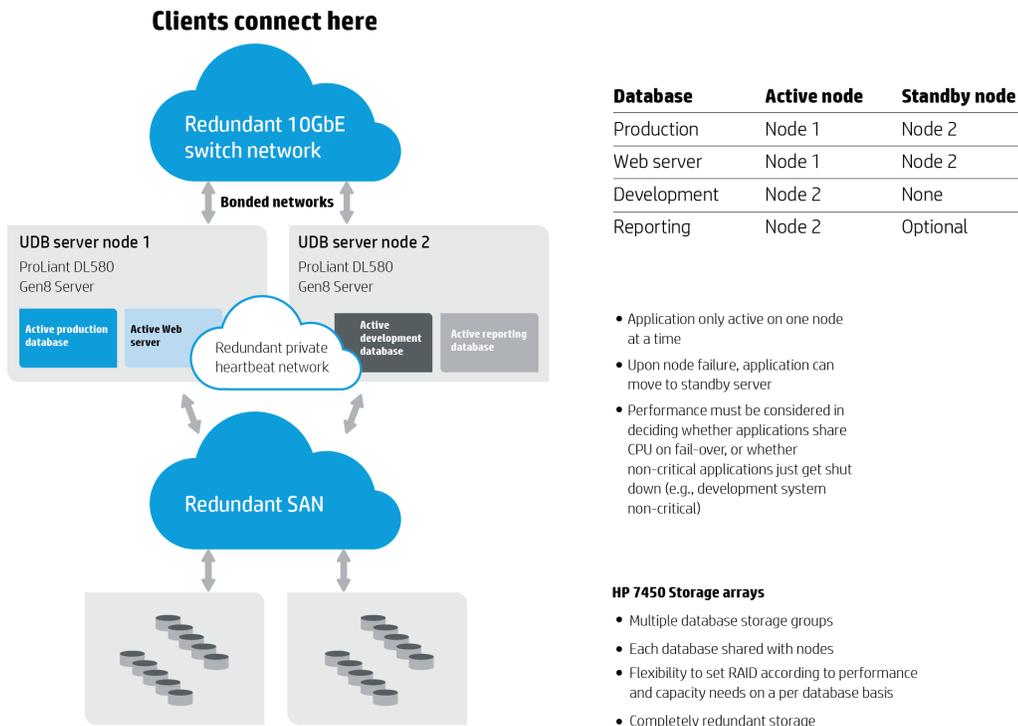
Three topologies for deploying a high availability UDB Oracle database solution are:

1. Using HP Serviceguard for Linux cluster solution
2. Disaster Recovery solution using HP 3PAR Remote Copy or Oracle Data Guard
3. Oracle Real Application Clusters (RAC)

### HP Serviceguard for Linux

HP Serviceguard for Linux, the high availability clustering software used in this solution, is designed to protect applications and services from planned and unplanned downtime. HP Serviceguard employs an active-standby cluster and provides great flexibility. Serviceguard can be configured to run multiple databases and has many features that integrate with not only the database but other components of the environment, such as applications and Web servers.

**Figure 10.** Oracle and HP 3PAR StoreServ 7450 optional integration with Serviceguard for Linux



HP Serviceguard Solutions for Linux portfolio also includes numerous implementation toolkits that allow you to easily integrate various databases and open source applications into a Serviceguard cluster with three distinct disaster recovery options. For additional information, see the [HP Serviceguard for Linux](#) website.

Key features of HP Serviceguard for Linux include:

- Robust monitoring protection against system, software, network, and storage faults.
- Advance cluster arbitration and fencing mechanisms preventing data corruption or loss.
- GUI and CLI management interfaces.
- Fast, accurate cluster package creation.

Figure 10 is an example of how Oracle and HP 3PAR StoreServ 7450 All-flash could be implemented in an HP Serviceguard for Linux cluster. This example is a two-node active-standby setup in which both servers can be used concurrently by multiple database instances, and also be configured to fail-over critical databases in case of failures.

## Disaster recovery with HP 3PAR Remote Copy, Oracle Data Guard, and VMware vSphere

The UDB Reference Architecture also includes disaster recovery options from both HP and Oracle.

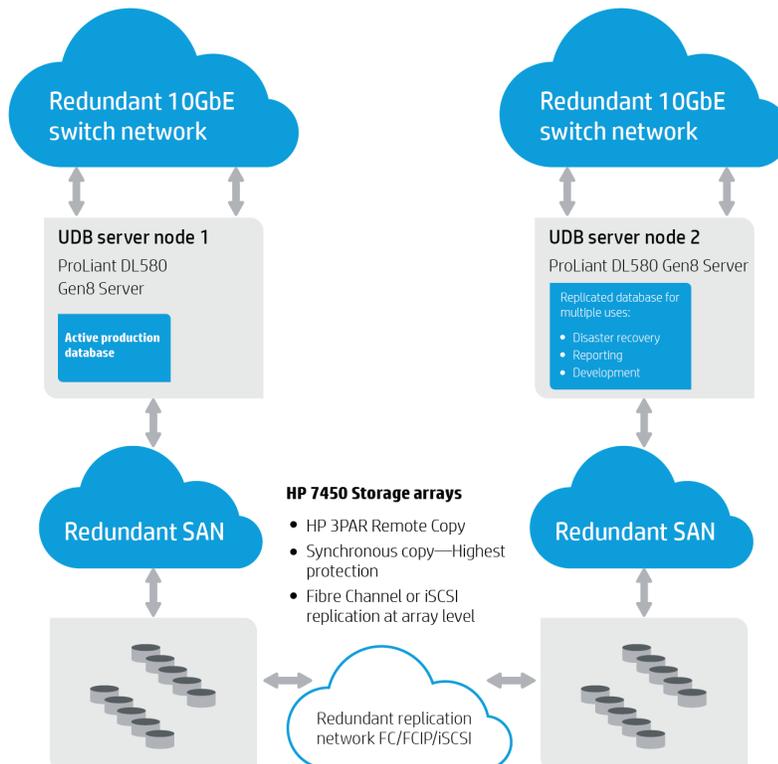
### HP 3PAR Remote Copy

The HP 3PAR Remote Copy software product provides an array based data replication solution on HP 3PAR StoreServ for replicating using HP 3PAR Remote Copy software.

Both synchronous and asynchronous replication options are supported. Figure 11 shows an example scenario for disaster recovery replication between two Oracle and HP 3PAR StoreServ 7450 All-flash environments. Replication to a remote site can be used for more than disaster recovery. The secondary site can be used for remote database reporting or database development. Use with HP 3PAR Snapshot technology allows for making database copies or even volume copies for remote backup. HP 3PAR StoreServ All-flash Array has the unique ability to provide flash-level performance and many of the desirable HP 3PAR management features.

**Figure 11.** An HP 3PAR Remote Copy environment for Oracle and HP 3PAR StoreServ 7450 All-flash Array

### Site A—Clients connect here      Site B—Clients connect here



## Oracle Data Guard

Oracle Data Guard is an Oracle product that provides data protection and disaster recovery for enterprise environments. Data Guard synchronizes a remote standby database, keeping the data consistent on the standby database. If the production database fails or needs to be taken down for service, Data Guard can switch the standby database to the production role. Data Guard can also be used for database backup and recovery.

## Oracle Real Application Clusters

Another supported clustering option is Oracle's RAC with Oracle Enterprise Database and Grid Infrastructure. Oracle RAC clustering technology is a scale out active-active cluster where multiple nodes are running their own instance of the same database allowing multiple server processing on the same database. Scaling out with Oracle RAC is a high availability and performance option.

## VMware vSphere High Availability

VMware HA delivers the high availability that many applications running in virtual machines require, independent of the operating system or underlying hardware configuration. VMware vSphere HA provides failover protection from hardware and operating system failures in the virtualized IT environment by:

- Monitoring virtual machines to detect operating system and hardware failures.
- Restarting virtual machines on other physical servers in the resource pool, without manual intervention when a server failure is detected.
- Protecting applications from operating system failures by automatically restarting virtual machines when an operating system failure is detected.

For more information, see [vmware.com/in/products/vsphere/features/availability](https://vmware.com/in/products/vsphere/features/availability).

## VMware vSphere vMotion

VMware vSphere vMotion enables the live migration of running virtual machines from one physical server to another with zero downtime, continuous service availability, and complete transaction integrity. vMotion is a key enabling technology for creating the dynamic, automated, and self-optimizing data center. This capability makes hardware maintenance possible at any time, and vMotion does not require clustering or redundant servers. vMotion:

- Moves entire running virtual machines instantly. Performs live migrations with zero downtime, undetectable to the user.
- Manages and schedules live migrations with ease at pre-defined times without an administrator's presence, with the reliability and manageability that is derived from a production-proven product.
- Performs multiple concurrent migrations of a virtual machine running any operating system, across any type of hardware and storage that is supported by vSphere, complete with an audit trail.
- Moves online workloads from one ESXi server host machine to another in order to maintain service levels and performance goals.
- Continuously and automatically optimizes virtual machine placement within resource pools. Proactively moves virtual machines away from failing or underperforming servers.
- Performs hardware maintenance without the need to schedule downtime and disrupt business operations.

For more information, see [vmware.com/in/products/vsphere/features/vmotion](https://vmware.com/in/products/vsphere/features/vmotion).

## VMware vSphere Distributed Resource Scheduler

VMware Distributed Resource Scheduler (DRS) is an automated load balancing technology that aligns resource usage with business priority by enabling automated load balancing across hosts, balances computing capacity, and it optimizes power consumption in the data center by turning off hosts during lower load periods. DRS continuously monitors utilization across vSphere servers and intelligently allocates available resources among virtual machines according to business needs. DRS takes advantage of vMotion to migrate virtual machines among a set of ESXi hosts.

DRS enables automatic initial virtual machine placement on any of the hosts in the cluster. It also makes automatic resource relocation and optimization decisions as hosts and virtual machines are added or removed from the cluster, or the load on individual virtual machines goes up or down. VMware DRS makes cluster wide resource pools possible. When DRS is configured for manual control, it makes recommendations for review and later implementation only; there is no automated activity.

DRS also provides the capability to control the placement of virtual machines based on different affinity rules. This enables virtual machines to be placed based on high availability or other business criticality requirement.

For more information, see [vmware.com/products/vsphere/features/drs-dpm](http://vmware.com/products/vsphere/features/drs-dpm).

## Best practices

### Analysis and recommendations

For best I/O performance on the HP ProLiant DL580 Gen8 Server, HP recommends using multiple paths to maintain high availability while also maximizing performance and minimizing latencies. A way to achieve better performance in extreme performance environments is to minimize inter-communication between NUMA nodes. This can be achieved using tightly-zoned hardware configurations and operating system-to-hardware configurations, such as setting CPU affinity to minimize latencies across the NUMA nodes.

The approach taken to achieve maximum I/Os and throughput was to connect and zone the HP DL580 Gen8 Server to storage in such a way that cross node activity is minimized both from the server and storage. By dedicating virtual volumes to specific HBAs and NUMA nodes, all of the I/O for a specific volume stays local to specific storage and server nodes. For applications that do a good job with NUMA awareness, this can deliver extremely good performance. For those applications that are not as good with NUMA awareness, more manual tuning may be required, but the flexibility to tune the environment exists.

### SAN recommendations

Each dedicated port has its own zone. For each virtual volume, two ports are connected to any single virtual volume. Zoning too many paths on a single volume can create latencies across the NUMA nodes. Improvements observed were as high as 25 percent during IOPS testing. At the very least, the paths on a single volume should all come from HBAs within a single NUMA node on the HP DL580 Gen8 Server.

### Server configuration best practices

For the best I/O performance of the HP ProLiant DL580 Gen8 Server, HP recommends using multiple paths to maintain high availability while also maximizing performance and minimizing latencies.

#### HP ProLiant DL580 Gen8 Server BIOS

- Virtual technology—Disabled
- Hyper-threading—Enabled
- Intel Turbo Boost—Enabled
- HP Power Profile—Max Performance
- Minimum processor idle power states—No C-states in the BIOS

## Operating system configuration parameters

- NUMA configuration—Map the location of the HBAs NUMA node owner. Map out the interrupt numbers in the server's Linux directory */proc/irq*, then assign the interrupt affinity to a core owned by that NUMA node (See Appendices C and D for details).
- OS and kernel configuration
  - Set *cstates* to 1 at kernel boot (see details in Appendix B)
  - Set *sysctl.conf* values as indicated in Appendix A
  - Ensure *debug* is not enabled in *sysfs*. Remove any packages in the OS that may be enabling tracing. To check to see if tracing is disabled, see Appendix G.

## Storage configuration best practices

- UDEV settings for performance: Set UDEV parameters per values in Appendix E.
- Set the *sysfs* “rotational” value for disks to 0.
- Set the *sysfs* value *rq\_affinity* to 2 for each device. Request completions all occurring on core 0 cause a bottleneck. Setting *rq\_affinity* to a value of 2 resolves this problem.
- Set scheduler to Deadline (no operation).
- Set permissions and ownership for Oracle volumes.
- SSD loading—Load SSDs in groups of four per enclosure at a minimum.
- Volume size—Virtual volumes should all be the same size and SSD type for each Oracle ASM group.
- vLUNs—HP recommends that any volume path originate from the same NUMA node on the DL580. It's best to keep the number of vLUNs per volume down to two.

## Database configuration best practices

Oracle parameters (See Appendix H)

- Set HUGE pages only.
- Disable automatic memory management if applicable.
- Set buffer cache memory size large enough per your implementation to avoid physical reads.
- Enable NUMA support.

## ESXi host and virtual machine configuration best practices

We do not consider the ESXi Host and Virtual machine settings listed to be best practice settings.

We consider most of these settings to be specific to obtaining extreme performance. These settings were used to obtain the 1 million IOPS result, but we would not recommend them to customers in general.

- Set Node affinity
- Relocated the vCPU pinning from node 0 to node 1
- Set CPU shares to high
- Pinned PVSCSI Affinity
- Set *numa.vcpu.preferHT* to true (in the case that the number of active vCPUs is greater than the number of available cores, there is more than one VM)

Please note that these settings are not recommended for general use. For Virtual machine configurations and settings for high performance, please see Appendix I and J.

## Summary

The HP Universal Database Solution is a new significant part of the overall HP performance reference architecture portfolio. It was developed to provide high performance I/O throughput for transactional databases in a package that delivers business continuity, extreme IOPS, faster user response times and increased throughput vs. comparable traditional server/storage configurations. This solution integrates with high availability options and disaster recovery options like HP 3PAR Remote Copy and HP Serviceguard for Linux.

Our extensive testing found the tested configuration to prove the following:

- An Oracle database environment using the HP ProLiant DL580 Gen8 Server and two HP 3PAR StoreServ 7450 All-flash Arrays is capable of delivering over one million IOPS.
- The reference architecture supports stable OLTP CPU or I/O stressed workloads in both non-virtual and virtualized configurations.
- The HP DL580 Gen8 Server is a suitable scale-up platform for consolidating large virtual machines. The aggregated performance throughput of the two virtual machines under a heavy transactional workload was measured nine percent below the throughput on physical.
- The HP ProLiant DL580 UDB is a flexible, mission-critical, extreme database performance solution with deployment flexibility to meet the needs of the customers.

## Implementing a proof-of-concept

As a best practice for all deployments, HP recommends implementing a proof-of-concept using a test environment that closely matches the planned production environment. This allows you to obtain appropriate performance and scalability characterizations.

For help with a proof-of-concept, contact your HP partner or an HP Services representative ([hp.com/large/contact/enterprise/index.html](http://hp.com/large/contact/enterprise/index.html)).

## Reference architecture rack diagram

Diagram of the single rack reference architecture.

**Figure 12.** HP ProLiant DL580 UDB rack diagram

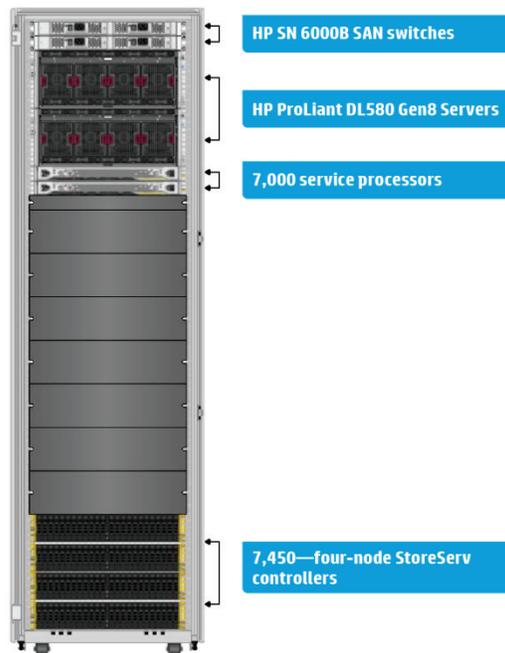
**2 x HP ProLiant DL580 Gen8 Servers each with:**

- 4 x Intel Xeon E7-4890 v2 processors
- 96 x 32 GB DIMMs (1.5 TB memory)
- 8 x 8 Gbps 2P FC HBA per server
- RHEL 6.5
- Oracle 12c Enterprise Edition

For virtualized Oracle: 60 vCPUs and 1 TB memory

**2 x HP 3PAR 7450 four-node storage array**

- 4 x drive shelves (Including the controller node pairs)
- 96 x 100 GB SLC SSD
- 48 x SSD drives per 7450 array
- 9.6 TB as configured



## Bill of materials

Below is the bill of materials (BOM) for the tested configuration. Variations of the configuration based on customer needs are possible, but would require using a separate BOM. Talk to your HP Sales representative for detailed quotes. See figure 12 for the reference architecture diagram of the tested environment.

## Reference architecture bill of materials

### Note

Part numbers are at time of publication and subject to change. The BOM does not include complete support options or other rack and power requirements. If you have questions regarding ordering, please consult with your HP Sales Representative or HP Reseller for more details. See [hp.com/large/contact/enterprise/index.html](http://hp.com/large/contact/enterprise/index.html).

### Rack configuration

**Table 6.** BOM for the HP ProLiant DL580 UDB rack

Quantity	Part number	Description
1	BW910A	HP 642 Intelligent Air Duct Rack
5	AF547A	HP 5xC13 Intlght PDU Ext Bars G2 Kit
1	BW909A	HP 42U 1,200 mm Side Panel Kit
13	C7535A	HP Ethernet 7 ft CAT5e RJ45 M/M Cable
1	BW930A	HP Air Flow Optimization Kit
1	BW937A	HP Air Duct Kit
1	BW891A	HP Rack Grounding Kit
<b>Select appropriate PDU option</b>		
2	H5M58A	HP 4.9 kVA 208 V 20 out NA/JP bPDU

### HP ProLiant DL580 Gen8 Servers

**Table 7.** BOM for the HP ProLiant DL580 UDB servers

Quantity	Part number	Description
2	728544-001	HP DL580 Gen8 E7-4890 v3 UP 128 GB US Sir
96	708641-B2B	HP 16 GB (1 x 16 GB) Dual Rank x4 PC3-14900R (DDR3-1866) Registered CAS-13 Memory Kit
8	652611-B21	HP 300 GB G6 SAS 25K 2.5 in SC END HDD
16	AJ763B	HP 82E 8 Gb Dual-port PCIe FC HBA
4	G5J64AAE	Red Hat Enterprise Linux Server 2 sockets, 4 guests, 3-year subscription 24x7 support E-LTU
8	BD713AAE	VMware vSphere Enterprise 1 processor, 3-year E-LTU *For virtualized configurations only

**Table 7.** BOM for the HP ProLiant DL580 UDB servers (continued)

8	U7D46E	HP 3 year Proactive Care VMware vSphere Ent 1P 3-year Software Service *For virtualized configurations only
---	--------	--

**HP 3PAR StoreServ 7450 Storage arrays**

**Table 8.** BOM for the HP ProLiant DL580 UDB storage arrays

Quantity	Part number	Description
2	C8R37A	HP 3PAR StoreServ 7450 4N Base
4	QR486A	HP 3PAR 7000 4-pt 8 Gb/s FC Adapter
96	QR502A	HP M6710 100 GB 6G SAS 2.5 in SLC SSD
2	QR516B	HP 3PAR 7000 Service Processor
2	BC890BAE	HP 3PAR 7450 OS Suite Base E-LTU
96	BC891AAE	HP 3PAR 7450 OS Suite Drive E-LTU
2	BC914BAE	HP 3PAR 7450 Reporting Suite E-LTU
1	BD362AAE	HP 3PAR StoreServ Mgmt/Core SW E-Media
1	BD365AAE	HP 3PAR 7000 Service Proc SW E-Media
1	BD373AAE	HP 3PAR Reporting Suite E-Media

**HP SN6000B switches**

**Table 9.** BOM for the HP ProLiant DL580 UDB network

Quantity	Part number	Description
2	QR480B	HP SN6000B 16 Gb 48-port/48-port Active Fibre Channel Switch
4	QK724A	HP B-series 16 Gb SFP+SW XCVR
96	QK733A	HP Premier Flex LC/LC OM4 2f 2m Cbl

**Notes**

- Refer to [HP SN6000B Network switch QuickSpecs](#) to determine proper transceivers and accessories for your specific network environment.
- Refer to [HP 3PAR StoreServ 7450 All-flash QuickSpecs](#) for the service processor (SP) and HP 3PAR StoreServ 7450 OS Suite options.
- Refer to [HP ProLiant DL580 Gen8 QuickSpecs](#) to determine the desired options for your environment.
- Refer to [HP 3PAR Software Products QuickSpecs](#) for details on HP 3PAR software options.

## Appendices

### Appendix A: Red Hat 6.5 kernel tunables */etc/sysctl.conf*

```
net.ipv4.ip_forward = 0
net.ipv4.conf.default.rp_filter = 1
net.ipv4.conf.default.accept_source_route = 0
kernel.sysrq = 0
kernel.core_uses_pid = 1
net.ipv4.tcp_syncookies = 1
net.bridge.bridge-nf-call-ip6tables = 0
net.bridge.bridge-nf-call-iptables = 0
net.bridge.bridge-nf-call-arptables = 0
kernel.msgmnb = 65536
kernel.msgmax = 65536
kernel.shmmax = 758959574016
kernel.shmall = 529408185
kernel.shmmni = 4096
kernel.sem = 500 64000 200 256

net.ipv4.ip_local_port_range = 9000 65500

fs.file-max = 8388608
fs.aio-max-nr = 4194304

net.core.rmem_default = 1048576
net.core.wmem_default = 1048576
net.core.rmem_max = 4194304
net.core.wmem_max = 1048576

net.ipv4.tcp_rmem = 1048576 1048576 4194304
net.ipv4.tcp_wmem = 1048576 1048576 1048576
net.ipv4.ip_local_port_range = 9000 65500

vm.swappiness=0

vm.dirty_background_ratio=3
vm.dirty_ratio=15
vm.dirty_expire_centisecs=500
vm.dirty_writeback_centisecs=100

vm.hugetlb_shm_group = 1000
vm.nr_hugepages = 375557
```

### **Red Hat 6.5 kernel tunables /etc/sysctl.conf (virtual)**

```
net.ipv4.ip_forward = 0
net.ipv4.conf.default.rp_filter = 1
net.ipv4.conf.default.accept_source_route = 0
kernel.sysrq = 0
kernel.core_uses_pid = 1
net.ipv4.tcp_syncookies = 1
net.bridge.bridge-nf-call-ip6tables = 0
net.bridge.bridge-nf-call-iptables = 0
net.bridge.bridge-nf-call-arptables = 0
kernel.msgmnb = 65536
kernel.msgmax = 65536
kernel.shmall = 4294967296
kernel.shmmni = 4096
kernel.shmmax = 4398046511104
kernel.sem = 250 32000 100 128
net.ipv4.ip_local_port_range = 9000 65500
vm.nr_hugepages = 125535
fs.file-max = 6815744
fs.aio-max-nr = 1048576
```

### **Appendix B: Grub configuration for disabling C-states (virtual/physical)**

```
kernel /vmlinuz-2.6.32-431.el6.x86_64 ro root=/dev/mapper/vg_aps4673-lv_root
rd_NO_LUKS LANG=en_US.UTF-8 rd_LVM_LV=vg_aps4673/lv_root rd_NO_MD
SYSFONT=latarcyrheb-sun16 crashkernel=auto rd_LVM_LV=vg_aps4673/lv_swap
KEYBOARDTYPE=pc KEYTABLE=us rd_NO_DM rhgb intel_idle.max_cstate=1 quiet
```

## Appendix C: IRQ affinity script for /etc/rc.local (physical only)

---

### Note

HBA card interrupt numbers must be verified with each specific implementation. See file /proc/interrupts in sysfs of the Linux operating system.

---

```
/etc/rc.local
```

```
# This script will be executed *after* all the other init scripts.# You can put your own initialization stuff in here if you don't  
#want to do the full Sys V style init stuff.
```

```
touch /var/lock/subsys/local
```

```
echo "15" > /proc/irq/172/smp_affinity_list  
echo "16" > /proc/irq/173/smp_affinity_list  
echo "17" > /proc/irq/174/smp_affinity_list  
echo "18" > /proc/irq/175/smp_affinity_list  
echo "19" > /proc/irq/176/smp_affinity_list  
echo "20" > /proc/irq/177/smp_affinity_list  
echo "21" > /proc/irq/178/smp_affinity_list  
echo "22" > /proc/irq/179/smp_affinity_list  
echo "23" > /proc/irq/180/smp_affinity_list  
echo "24" > /proc/irq/181/smp_affinity_list  
echo "25" > /proc/irq/182/smp_affinity_list  
echo "26" > /proc/irq/183/smp_affinity_list
```

```
echo "30" > /proc/irq/184/smp_affinity_list  
echo "31" > /proc/irq/185/smp_affinity_list  
echo "32" > /proc/irq/186/smp_affinity_list  
echo "33" > /proc/irq/187/smp_affinity_list  
echo "34" > /proc/irq/188/smp_affinity_list  
echo "35" > /proc/irq/189/smp_affinity_list  
echo "36" > /proc/irq/190/smp_affinity_list  
echo "37" > /proc/irq/191/smp_affinity_list  
echo "38" > /proc/irq/192/smp_affinity_list  
echo "39" > /proc/irq/193/smp_affinity_list  
echo "40" > /proc/irq/194/smp_affinity_list  
echo "41" > /proc/irq/195/smp_affinity_list
```

```
echo "45" > /proc/irq/196/smp_affinity_list  
echo "46" > /proc/irq/197/smp_affinity_list  
echo "47" > /proc/irq/198/smp_affinity_list  
echo "48" > /proc/irq/199/smp_affinity_list  
echo "49" > /proc/irq/200/smp_affinity_list  
echo "50" > /proc/irq/201/smp_affinity_list  
echo "51" > /proc/irq/202/smp_affinity_list  
echo "52" > /proc/irq/203/smp_affinity_list
```

## Appendix D: HBA NUMA mapping and IRQ map (physical/virtual)

### Note

Host values and WWN values are specific to each implementation and must be obtained for each implementation.

Bus Address	Slot	Local CPU List	NUMA Node	Host	Port WWN
41:00.0	6	15-29,75-89	1	host5	0x10000000c9b74b20
41:00.1	6	15-29,75-89	1	host6	0x10000000c9b74b21
44:00.0	7	15-29,75-89	1	host1	0x10000000c9e89b84
44:00.1	7	15-29,75-89	1	host2	0x10000000c9e89b85
47:00.0	8	15-29,75-89	1	host3	0x10000000c9e89cbe
47:00.1	8	15-29,75-89	1	host4	0x10000000c9e89cbf
81:00.0	3	30-44,90-104	2	host11	0x10000000c9e8ea62
81:00.1	3	30-44,90-104	2	host12	0x10000000c9e8ea63
84:00.0	4	30-44,90-104	2	host7	0x10000000c9a1a156
84:00.1	4	30-44,90-104	2	host8	0x10000000c9a1a157
87:00.0	5	30-44,90-104	2	host9	0x10000000c9e8eb46
87:00.1	5	30-44,90-104	2	host10	0x10000000c9e8eb47
c1:00.0	1	45-59,105-119	3	host15	0x10000000c9b752c2
c1:00.1	1	45-59,105-119	3	host16	0x10000000c9b752c3
c4:00.0	2	45-59,105-119	3	host13	0x10000000c9e8ea68
c4:00.1	2	45-59,105-119	3	host14	0x10000000c9e8ea69

### Emulex Interrupt Affinity

Interrupt Number	NUMA Node	Affinity
172	1	0000,00000000,00008000,00000000,00080000 lpfc:sp
173	1	0000,00000000,00010000,00000000,00100000 lpfc:fp
174	1	0000,00000000,00020000,00000000,00200000 lpfc:sp
175	1	0000,00000000,00040000,00000000,00400000 lpfc:fp
176	1	0000,00000000,00080000,00000000,00800000 lpfc:sp
177	1	0000,00000000,00100000,00000000,01000000 lpfc:fp
178	1	0000,00000000,00200000,00000000,02000000 lpfc:sp
179	1	0000,00000000,00020000,00000000,00200000 lpfc:fp
180	1	0000,00000000,01000000,00000000,10000000 lpfc:sp
181	1	0000,00000000,00008000,00000000,00080000 lpfc:fp
182	1	0000,00000000,02000000,00000000,20000000 lpfc:sp
183	1	0000,00000000,00001000,00000000,00010000 lpfc:fp
184	2	0000,00000000,40000000,00000004,00000000 lpfc:sp
185	2	0000,00000000,80000000,00000008,00000000 lpfc:fp
186	2	0000,00000001,00000000,00000010,00000000 lpfc:sp
187	2	0000,00000002,00000000,00000020,00000000 lpfc:fp
188	2	0000,00000004,00000000,00000040,00000000 lpfc:sp
189	2	0000,00000008,00000000,00000080,00000000 lpfc:fp
190	2	0000,00000010,00000000,00000100,00000000 lpfc:sp
191	2	0000,00000010,00000000,00000100,00000000 lpfc:fp
192	2	0000,00000100,00000000,00001000,00000000 lpfc:sp

193	2	0000,00000080,00000000,00000800,00000000	lpfc:fp
194	2	0000,00000000,10000000,00000001,00000000	lpfc:sp
195	2	0000,00000000,08000000,00000000,80000000	lpfc:fp
196	3	0000,00020000,00000000,00200000,00000000	lpfc:sp
197	3	0000,00040000,00000000,00400000,00000000	lpfc:fp
198	3	0000,00080000,00000000,00800000,00000000	lpfc:sp
199	3	0000,00100000,00000000,01000000,00000000	lpfc:fp
200	3	0000,00200000,00000000,02000000,00000000	lpfc:sp
201	3	0000,00400000,00000000,04000000,00000000	lpfc:fp
202	3	0000,00800000,00000000,08000000,00000000	lpfc:sp
203	3	0000,00000400,00000000,00004000,00000000	lpfc:fp

## Appendix E: UDEV configurations

```
/etc/udev/rules.d/99-3par.rules
ACTION=="add|change", SUBSYSTEM=="block", ATTR{device/vendor}=="3PARdata", ATTR{queue/rotational}="0",
ATTR{queue/scheduler}="noop", ATTR{queue/rq_affinity}="2", ATTR{queue/nr_requests}="2048"
ACTION=="add|change", KERNEL=="dm-*", PROGRAM="/bin/bash -c 'cat /sys/block/$name/slaves/*/device/vendor | grep
3PARdata'", ATTR{queue/rotational}="0", ATTR{queue/scheduler}="noop", ATTR{queue/rq_affinity}="2",
ATTR{queue/nr_requests}="2048"
```

```
/etc/udev/rules.d/99-asm.rules
```

```
KERNEL=="dm*", \
PROGRAM=="scsi_id --page=0x83 --whitelisted --device=/dev/percentk", \
RESULT=="360002ac000000000000002000005f98", \
NAME="oracleasm/asm3par01", OWNER="oracle", GROUP="oinstall", MODE="0660"
```

```
KERNEL=="dm*", \
PROGRAM=="scsi_id --page=0x83 --whitelisted --device=/dev/percentk", \
RESULT=="360002ac0000000000000001c00005f98", \
NAME="oracleasm/asm3par02", OWNER="oracle", GROUP="oinstall", MODE="0660"
```

```
KERNEL=="dm*", \
PROGRAM=="scsi_id --page=0x83 --whitelisted --device=/dev/percentk", \
RESULT=="360002ac0000000000000001f00005f98", \
NAME="oracleasm/asm3par03", OWNER="oracle", GROUP="oinstall", MODE="0660"
```

```
KERNEL=="dm*", \
PROGRAM=="scsi_id --page=0x83 --whitelisted --device=/dev/percentk", \
RESULT=="360002ac0000000000000001e00005f98", \
NAME="oracleasm/asm3par04", OWNER="oracle", GROUP="oinstall", MODE="0660"
```

```
KERNEL=="dm*", \
PROGRAM=="scsi_id --page=0x83 --whitelisted --device=/dev/percentk", \
RESULT=="360002ac0000000000000001700005f9a", \
NAME="oracleasm/asm3par05", OWNER="oracle", GROUP="oinstall", MODE="0660"
```

```
KERNEL=="dm*", \
PROGRAM=="scsi_id --page=0x83 --whitelisted --device=/dev/percentk", \
RESULT=="360002ac0000000000000001800005f9a", \
NAME="oracleasm/asm3par06", OWNER="oracle", GROUP="oinstall", MODE="0660"
```

```
KERNEL=="dm*", \
PROGRAM=="scsi_id --page=0x83 --whitelisted --device=/dev/ percentk", \
RESULT=="360002ac000000000000001600005f9a", \
NAME="oracleasm/asm3par07", OWNER="oracle", GROUP="oinstall", MODE="0660"
```

```
KERNEL=="dm*", \
PROGRAM=="scsi_id --page=0x83 --whitelisted --device=/dev/ percentk", \
RESULT=="360002ac000000000000001500005f9a", \
NAME="oracleasm/asm3par08", OWNER="oracle", GROUP="oinstall", MODE="0660"
```

```
KERNEL=="dm*", \
PROGRAM=="scsi_id --page=0x83 --whitelisted --device=/dev/ percentk", \
RESULT=="360002ac000000000000001300005f9a", \
NAME="oracleasm/asm3par09", OWNER="oracle", GROUP="oinstall", MODE="0660"
```

```
KERNEL=="dm*", \
PROGRAM=="scsi_id --page=0x83 --whitelisted --device=/dev/ percentk", \
RESULT=="360002ac000000000000001400005f9a", \
NAME="oracleasm/asm3par10", OWNER="oracle", GROUP="oinstall", MODE="0660"
```

```
KERNEL=="dm*", \
PROGRAM=="scsi_id --page=0x83 --whitelisted --device=/dev/ percentk", \
RESULT=="360002ac000000000000001900005f98", \
NAME="oracleasm/asm3par11", OWNER="oracle", GROUP="oinstall", MODE="0660"
```

```
KERNEL=="dm*", \
PROGRAM=="scsi_id --page=0x83 --whitelisted --device=/dev/ percentk", \
RESULT=="360002ac000000000000001d00005f98", \
NAME="oracleasm/asm3par12", OWNER="oracle", GROUP="oinstall", MODE="0660"
```

```
KERNEL=="dm*", \
PROGRAM=="scsi_id --page=0x83 --whitelisted --device=/dev/ percentk", \
RESULT=="360002ac000000000000001a00005f98", \
NAME="oracleasm/asm3par13", OWNER="oracle", GROUP="oinstall", MODE="0660"
```

```
KERNEL=="dm*", \
PROGRAM=="scsi_id --page=0x83 --whitelisted --device=/dev/ percentk", \
RESULT=="360002ac000000000000001b00005f98", \
NAME="oracleasm/asm3par14", OWNER="oracle", GROUP="oinstall", MODE="0660"
```

```
KERNEL=="dm*", \
PROGRAM=="scsi_id --page=0x83 --whitelisted --device=/dev/ percentk", \
RESULT=="360002ac000000000000001900005f9a", \
NAME="oracleasm/asm3par15", OWNER="oracle", GROUP="oinstall", MODE="0660"
```

```
KERNEL=="dm*", \
PROGRAM=="scsi_id --page=0x83 --whitelisted --device=/dev/ percentk", \
RESULT=="360002ac000000000000001a00005f9a", \
NAME="oracleasm/asm3par16", OWNER="oracle", GROUP="oinstall", MODE="0660"
```

### UDEV configurations (virtual)

/etc/udev/rules.d/99-oracleasm.rules

#3Par Luns

```
KERNEL=="sd*", \
PROGRAM=="scsi_id --page=0x83 --whitelisted --device=/dev/ percentk", \
RESULT=="360002ac000000000000006200005f98", \
NAME="oracleasm/asm3parocr80", OWNER="oracle", GROUP="oinstall", MODE="0660"
```

#3Par Luns

```
KERNEL=="sd*", \
PROGRAM=="scsi_id --page=0x83 --whitelisted --device=/dev/ percentk", \
RESULT=="360002ac000000000000005700005f9a", \
NAME="oracleasm/asm3par01", OWNER="oracle", GROUP="oinstall", MODE="0660"
```

```
KERNEL=="sd*", \
PROGRAM=="scsi_id --page=0x83 --whitelisted --device=/dev/ percentk", \
RESULT=="360002ac000000000000005600005f9a", \
NAME="oracleasm/asm3par02", OWNER="oracle", GROUP="oinstall", MODE="0660"
```

```
KERNEL=="sd*", \
PROGRAM=="scsi_id --page=0x83 --whitelisted --device=/dev/ percentk", \
RESULT=="360002ac000000000000006a00005f98", \
NAME="oracleasm/asm3par03", OWNER="oracle", GROUP="oinstall", MODE="0660"
```

```
KERNEL=="sd*", \
PROGRAM=="scsi_id --page=0x83 --whitelisted --device=/dev/ percentk", \
RESULT=="360002ac000000000000006900005f98", \
NAME="oracleasm/asm3par04", OWNER="oracle", GROUP="oinstall", MODE="0660"
```

```
KERNEL=="sd*", \
PROGRAM=="scsi_id --page=0x83 --whitelisted --device=/dev/ percentk", \
RESULT=="360002ac000000000000005b00005f9a", \
NAME="oracleasm/asm3par05", OWNER="oracle", GROUP="oinstall", MODE="0660"
```

```
KERNEL=="sd*", \
PROGRAM=="scsi_id --page=0x83 --whitelisted --device=/dev/ percentk", \
RESULT=="360002ac000000000000005800005f9a", \
NAME="oracleasm/asm3par06", OWNER="oracle", GROUP="oinstall", MODE="0660"
```

```
KERNEL=="sd*", \
PROGRAM=="scsi_id --page=0x83 --whitelisted --device=/dev/ percentk", \
RESULT=="360002ac000000000000005900005f9a", \
NAME="oracleasm/asm3par07", OWNER="oracle", GROUP="oinstall", MODE="0660"
```

```
KERNEL=="sd*", \
PROGRAM=="scsi_id --page=0x83 --whitelisted --device=/dev/ percentk", \
RESULT=="360002ac000000000000005c00005f9a", \
NAME="oracleasm/asm3par08", OWNER="oracle", GROUP="oinstall", MODE="0660"
```

```
KERNEL=="sd*", \
PROGRAM=="scsi_id --page=0x83 --whitelisted --device=/dev/ percentk", \
RESULT=="360002ac0000000000000005d00005f9a", \
NAME="oracleasm/asm3par09", OWNER="oracle", GROUP="oinstall", MODE="0660"
```

```
KERNEL=="sd*", \
PROGRAM=="scsi_id --page=0x83 --whitelisted --device=/dev/ percentk", \
RESULT=="360002ac0000000000000005a00005f9a", \
NAME="oracleasm/asm3par10", OWNER="oracle", GROUP="oinstall", MODE="0660"
```

```
KERNEL=="sd*", \
PROGRAM=="scsi_id --page=0x83 --whitelisted --device=/dev/ percentk", \
RESULT=="360002ac0000000000000006800005f98", \
NAME="oracleasm/asm3par11", OWNER="oracle", GROUP="oinstall", MODE="0660"
```

```
KERNEL=="sd*", \
PROGRAM=="scsi_id --page=0x83 --whitelisted --device=/dev/ percentk", \
RESULT=="360002ac0000000000000006700005f98", \
NAME="oracleasm/asm3par12", OWNER="oracle", GROUP="oinstall", MODE="0660"
```

```
KERNEL=="sd*", \
PROGRAM=="scsi_id --page=0x83 --whitelisted --device=/dev/ percentk", \
RESULT=="360002ac0000000000000006600005f98", \
NAME="oracleasm/asm3par13", OWNER="oracle", GROUP="oinstall", MODE="0660"
```

```
KERNEL=="sd*", \
PROGRAM=="scsi_id --page=0x83 --whitelisted --device=/dev/ percentk", \
RESULT=="360002ac0000000000000006500005f98", \
NAME="oracleasm/asm3par14", OWNER="oracle", GROUP="oinstall", MODE="0660"
```

```
KERNEL=="sd*", \
PROGRAM=="scsi_id --page=0x83 --whitelisted --device=/dev/ percentk", \
RESULT=="360002ac0000000000000006400005f98", \
NAME="oracleasm/asm3par15", OWNER="oracle", GROUP="oinstall", MODE="0660"
```

```
KERNEL=="sd*", \
PROGRAM=="scsi_id --page=0x83 --whitelisted --device=/dev/ percentk", \
RESULT=="360002ac0000000000000006300005f98", \
NAME="oracleasm/asm3par16", OWNER="oracle", GROUP="oinstall", MODE="0660"
```

## Appendix F: Storage information

### Zoning example—WWNs specific to implementation (example only)

```

Switch Top
Effective configuration:
cfg:  ORACLE
zone:  H1_I1_A1_T2
      10:00:00:00:c9:b7:52:c3
      20:12:00:02:ac:00:5f:98
zone:  H1_I2_A1_T2
      10:00:00:00:c9:e8:ea:69
      21:12:00:02:ac:00:5f:98
zone:  H1_I3_A1_T2
      10:00:00:00:c9:e8:ea:63
      22:12:00:02:ac:00:5f:98
zone:  H1_I4_A1_T2
      10:00:00:00:c9:a1:a1:57
      23:12:00:02:ac:00:5f:98
zone:  H1_I5_A2_T2
      10:00:00:00:c9:e8:eb:47
      20:12:00:02:ac:00:5f:9a
zone:  H1_I6_A2_T2
      10:00:00:00:c9:b7:4b:21
      21:12:00:02:ac:00:5f:9a
zone:  H1_I7_A2_T2
      10:00:00:00:c9:e8:9b:85
      22:12:00:02:ac:00:5f:9a
zone:  H1_I8_A2_T2
      10:00:00:00:c9:e8:9c:bf
      23:12:00:02:ac:00:5f:9a
zone:  H2_I1_A1_T4
      10:00:00:00:c9:b7:28:f9
      20:22:00:02:ac:00:5f:98
zone:  H2_I2_A1_T4
      10:00:00:00:c9:b7:2a:87
      21:22:00:02:ac:00:5f:98
zone:  H2_I3_A1_T4
      10:00:00:00:c9:b7:51:6b
      22:22:00:02:ac:00:5f:98
zone:  H2_I4_A1_T4
      10:00:00:00:c9:c0:8f:cd
      23:22:00:02:ac:00:5f:98
zone:  H2_I5_A2_T4
      10:00:00:00:c9:a1:a2:9d
      20:22:00:02:ac:00:5f:9a
zone:  H2_I6_A2_T4
      10:00:00:00:c9:e8:e8:ad
      21:22:00:02:ac:00:5f:9a
zone:  H2_I7_A2_T4
      10:00:00:00:c9:b7:28:ef
      22:22:00:02:ac:00:5f:9a
zone:  H2_I8_A2_T4
      10:00:00:00:c9:a1:ae:b7
      23:22:00:02:ac:00:5f:9a
    
```

```

Effective configuration:
cfg:  ORACLE
zone:  H1_I1_A1_T1
      10:00:00:00:c9:b7:52:c2
      20:11:00:02:ac:00:5f:98
zone:  H1_I2_A1_T1
      10:00:00:00:c9:e8:ea:68
      21:11:00:02:ac:00:5f:98
zone:  H1_I3_A1_T1
      10:00:00:00:c9:e8:ea:62
      22:11:00:02:ac:00:5f:98
zone:  H1_I4_A1_T1
      10:00:00:00:c9:a1:a1:56
      23:11:00:02:ac:00:5f:98
zone:  H1_I5_A2_T1
      10:00:00:00:c9:e8:eb:46
      20:11:00:02:ac:00:5f:9a
zone:  H1_I6_A2_T1
      10:00:00:00:c9:b7:4b:20
      21:11:00:02:ac:00:5f:9a
zone:  H1_I7_A2_T1
      10:00:00:00:c9:e8:9b:84
      22:11:00:02:ac:00:5f:9a
zone:  H1_I8_A2_T1
      10:00:00:00:c9:e8:9c:be
      23:11:00:02:ac:00:5f:9a
zone:  H2_I1_A1_T3
      10:00:00:00:c9:b7:28:f8
      20:21:00:02:ac:00:5f:98
zone:  H2_I2_A1_T3
      10:00:00:00:c9:b7:2a:86
      21:21:00:02:ac:00:5f:98
zone:  H2_I3_A1_T3
      10:00:00:00:c9:b7:51:6a
      22:21:00:02:ac:00:5f:98
zone:  H2_I4_A1_T3
      10:00:00:00:c9:c0:8f:cc
      23:21:00:02:ac:00:5f:98
zone:  H2_I5_A2_T3
      10:00:00:00:c9:a1:a2:9c
      20:21:00:02:ac:00:5f:9a
zone:  H2_I6_A2_T3
      10:00:00:00:c9:e8:e8:ac
      21:21:00:02:ac:00:5f:9a
zone:  H2_I7_A2_T3
      10:00:00:00:c9:b7:28:ee
      22:21:00:02:ac:00:5f:9a
zone:  H2_I8_A2_T3
      10:00:00:00:c9:a1:ae:b6
      23:21:00:02:ac:00:5f:9a

```

### HP 3PAR StoreServ 7450 All-flash CLI examples

3PAR #1

Osorno cli percent showcpg

```

----- (MB) -----
      -Volumes- -Usage- --- Usr --- -- Snp --- -- Adm ---
Id Name      Warn  percent  VVs  TPVVs  Usr  Snp    Total  Used  Total  Used  Total  Used
0 SSD_r1     -    0         0    0       0    0      0      0    0     0    0     0
1 SSD_r5     -    1         1    1       0   20480  20480  0     0    16384 384
2 SSD_r6     -    0         0    0       0    0      0      0    0     0    0     0
    
```

Osorno cli percent showvlun

Active VLUNs

Lun	VVName	HostName	-Host_WWN/iSCSI_Name-	Port	Type	Status	ID
0	Array1_ora.2_nd0	EB03_APS46_073_n0	10000000C9E8EB46	0:1:1	host	active	1
1	Array1_ora.1_nd0	EB03_APS46_073_n0	10000000C9E8EB46	0:1:1	host	active	1
0	Array1_ora.2_nd0	EB03_APS46_073_n0	10000000C9E8EB47	0:1:2	host	active	1
1	Array1_ora.1_nd0	EB03_APS46_073_n0	10000000C9E8EB47	0:1:2	host	active	1
0	Array1_ora.4_nd1	EB03_APS46_073_n1	10000000C9B74B20	1:1:1	host	active	1
1	Array1_ora.3_nd1	EB03_APS46_073_n1	10000000C9B74B20	1:1:1	host	active	1
0	Array1_ora.4_nd1	EB03_APS46_073_n1	10000000C9B74B21	1:1:2	host	active	1
1	Array1_ora.3_nd1	EB03_APS46_073_n1	10000000C9B74B21	1:1:2	host	active	1
0	Array1_ora.6_nd2	EB03_APS46_073_n2	10000000C9E89B84	2:1:1	host	active	1
1	Array1_ora.5_nd2	EB03_APS46_073_n2	10000000C9E89B84	2:1:1	host	active	1
0	Array1_ora.6_nd2	EB03_APS46_073_n2	10000000C9E89B85	2:1:2	host	active	1
1	Array1_ora.5_nd2	EB03_APS46_073_n2	10000000C9E89B85	2:1:2	host	active	1
0	Array1_ora.8_nd3	EB03_APS46_073_n3	10000000C9E89CBE	3:1:1	host	active	1
1	Array1_ora.7_nd3	EB03_APS46_073_n3	10000000C9E89CBE	3:1:1	host	active	1
0	Array1_ora.8_nd3	EB03_APS46_073_n3	10000000C9E89CBF	3:1:2	host	active	1
1	Array1_ora.7_nd3	EB03_APS46_073_n3	10000000C9E89CBF	3:1:2	host	active	1

VLUN Templates

Lun	VVName	HostName	-Host_WWN/iSCSI_Name-	Port	Type
0	Array1_ora.2_nd0	EB03_APS46_073_n0	-----	---	host
1	Array1_ora.1_nd0	EB03_APS46_073_n0	-----	---	host
0	Array1_ora.4_nd1	EB03_APS46_073_n1	-----	---	host
1	Array1_ora.3_nd1	EB03_APS46_073_n1	-----	---	host
0	Array1_ora.6_nd2	EB03_APS46_073_n2	-----	---	host
1	Array1_ora.5_nd2	EB03_APS46_073_n2	-----	---	host
0	Array1_ora.8_nd3	EB03_APS46_073_n3	-----	---	host
1	Array1_ora.7_nd3	EB03_APS46_073_n3	-----	---	host
0	Datastore_74	EB03_APS46_074_n0	-----	---	host
1	Array1_ora.2_nd0	EB03_APS46_074_n0	-----	---	host
2	Array1_ora.1_nd0	EB03_APS46_074_n0	-----	---	host
0	Array1_ora.4_nd1	EB03_APS46_074_n1	-----	---	host
1	Array1_ora.3_nd1	EB03_APS46_074_n1	-----	---	host
0	Array1_ora.6_nd2	EB03_APS46_074_n2	-----	---	host
1	Array1_ora.5_nd2	EB03_APS46_074_n2	-----	---	host
0	Array1_ora.8_nd3	EB03_APS46_074_n3	-----	---	host
1	Array1_ora.7_nd3	EB03_APS46_074_n3	-----	---	host

Osorno cli percent showport

N:S:P	Mode	State	----Node_WWN----	Port_WWN/HW_Addr-	Type	Protocol	Label	Partner	FailoverState
0:0:1	initiator	ready	50002ACFF7005F9A	50002AC001005F9A	disk	SAS	DP-1	-	-
0:0:2	initiator	ready	50002ACFF7005F9A	50002AC002005F9A	disk	SAS	DP-2	-	-
0:1:1	target	ready	2FF70002AC005F9A	20110002AC005F9A	host	FC	-	1:1:1	none
0:1:2	target	ready	2FF70002AC005F9A	20120002AC005F9A	host	FC	-	1:1:2	none
0:2:1	target	ready	2FF70002AC005F9A	20210002AC005F9A	host	FC	-	1:2:1	none
0:2:2	target	ready	2FF70002AC005F9A	20220002AC005F9A	host	FC	-	1:2:2	none
0:2:3	target	ready	2FF70002AC005F9A	20230002AC005F9A	free	FC	-	1:2:3	none
0:2:4	target	ready	2FF70002AC005F9A	20240002AC005F9A	free	FC	-	1:2:4	none
0:3:1	peer	offline-		AC162D360385	rcip	IP	RCIP0	-	-
1:0:1	initiator	ready	50002ACFF7005F9A	50002AC101005F9A	disk	SAS	DP-1	-	-
1:0:2	initiator	ready	50002ACFF7005F9A	50002AC102005F9A	disk	SAS	DP-2	-	-
1:1:1	target	ready	2FF70002AC005F9A	21110002AC005F9A	host	FC	-	0:1:1	none
1:1:2	target	ready	2FF70002AC005F9A	21120002AC005F9A	host	FC	-	0:1:2	none
1:2:1	target	ready	2FF70002AC005F9A	21210002AC005F9A	host	FC	-	0:2:1	none
1:2:2	target	ready	2FF70002AC005F9A	21220002AC005F9A	host	FC	-	0:2:2	none
1:2:3	target	ready	2FF70002AC005F9A	21230002AC005F9A	free	FC	-	0:2:3	none
1:2:4	target	ready	2FF70002AC005F9A	21240002AC005F9A	free	FC	-	0:2:4	none
1:3:1	peer	offline-		AC162D360845	rcip	IP	RCIP1	-	-
2:0:1	initiator	ready	50002ACFF7005F9A	50002AC201005F9A	disk	SAS	DP-1	-	-
2:0:2	initiator	ready	50002ACFF7005F9A	50002AC202005F9A	disk	SAS	DP-2	-	-
2:1:1	target	ready	2FF70002AC005F9A	22110002AC005F9A	host	FC	-	3:1:1	none
2:1:2	target	ready	2FF70002AC005F9A	22120002AC005F9A	host	FC	-	3:1:2	none
2:2:1	target	ready	2FF70002AC005F9A	22210002AC005F9A	host	FC	-	3:2:1	none
2:2:2	target	ready	2FF70002AC005F9A	22220002AC005F9A	host	FC	-	3:2:2	none
2:2:3	target	ready	2FF70002AC005F9A	22230002AC005F9A	free	FC	-	3:2:3	none
2:2:4	target	ready	2FF70002AC005F9A	22240002AC005F9A	free	FC	-	3:2:4	none
2:3:1	peer	offline-		0002AC800577	rcip	IP	RCIP2	-	-
3:0:1	initiator	ready	50002ACFF7005F9A	50002AC301005F9A	disk	SAS	DP-1	-	-
3:0:2	initiator	ready	50002ACFF7005F9A	50002AC302005F9A	disk	SAS	DP-2	-	-
3:1:1	target	ready	2FF70002AC005F9A	23110002AC005F9A	host	FC	-	2:1:1	none
3:1:2	target	ready	2FF70002AC005F9A	23120002AC005F9A	host	FC	-	2:1:2	none
3:2:1	target	ready	2FF70002AC005F9A	23210002AC005F9A	host	FC	-	2:2:1	none
3:2:2	target	ready	2FF70002AC005F9A	23220002AC005F9A	host	FC	-	2:2:2	none
3:2:3	target	ready	2FF70002AC005F9A	23230002AC005F9A	free	FC	-	2:2:3	none
3:2:4	target	ready	2FF70002AC005F9A	23240002AC005F9A	free	FC	-	2:2:4	none
3:3:1	peer	offline-		0002AC8005DF	rcip	IP	RCIP3	-	-

## Appendix G: Check or set operating system tracing parameter (physical/virtual)

If tracing is enabled on the operating system, latencies from events can be introduced into the kernel causing delays in I/O operations. During I/O characterization testing as much as 10 percent I/O performance degradation was observed. Ensure any tools that enable tracing have been disabled or removed unless they are needed for specific support purposes.

To check the state of tracing on the system, run the following commands:

```
cat
/sys/kernel/debug/tracing/tracing_enable
d cat
/sys/kernel/debug/tracing/tracing_on
```

The result of both these commands should be 0. To disable tracing temporarily run the following commands:

```
echo "0" >
/sys/kernel/debug/tracing/tracing_enabled
echo "0" >
/sys/kernel/debug/tracing/tracing_on
```

To permanently disable tracing, remove the application on the system that is enabling debug or add the above commands to the `/etc/rc.local` file.

A debug tool called JRockit Flight Recorder can cause debug to be enabled. To determine if Flight Recorder is installed check for it on your Linux server using this command:

```
rpm -qa | grep flightrecorder
```

If the package exists, delete it using `rpm -e` or run the following command:

```
service trace-cmd stop
```

## Appendix H: Oracle parameters (physical/virtual)

Some parameters will be different between physical and virtual due to the memory size differences.

_disable_logging	FALSE
_disable_selftune_checkpointing	TRUE
_enable_NUMA_support	TRUE
_in_memory_undo	TRUE
_kgl_hot_object_copies	4
aq_tm_processes	0
audit_file_dest	/home/u01/app/oracle/admin/OLTP1/adump
audit_trail	NONE
commit_logging	BATCH
commit_wait	NOWAIT
compatible	12.1.0.0.0
control_files	+DATA/OLTP1/CONTROLFILE/current.267.861893211
db_16k_cache_size	53687091200
db_block_checking	FALSE
db_block_checksum	FALSE
db_block_size	8192
db_cache_size	322122547200
db_create_file_dest	+DATA
db_domain	
db_file_multiblock_read_count	32
db_name	OLTP1
db_writer_processes	4
diagnostic_dest	/home/u01/app/oracle
disk_asynch_io	TRUE
dispatchers	(PROTOCOL=TCP) (SERVICE=oltp1XDB)
dml_locks	500
fast_start_mttr_target	0
filesystemio_options	setall
large_pool_size	3758096384
local_listener	LISTENER_OLTP1
lock_sga	TRUE
log_checkpoint_interval	0
log_checkpoint_timeout	0
log_checkpoints_to_alert	TRUE
open_cursors	2000
parallel_max_servers	0
parallel_min_servers	0
pga_aggregate_target	134840582144
plsql_code_type	NATIVE
plsql_optimize_level	3
processes	1000
query_rewrite_enabled	TRUE
remote_login_passwordfile	EXCLUSIVE
replication_dependency_tracking	FALSE
result_cache_max_size	0
sessions	1560

```

shared_pool_size          17716740096
statistics_level          BASIC
timed_statistics          FALSE
trace_enabled             FALSE
transactions_per_rollback_segment  1
undo_management           AUTO
undo_retention            0
undo_tablespace           UNDOTBS1
use_large_pages           ONLY
    
```

## Appendix I: Virtual machine configuration file (virtual only)

Some parameters will vary slightly between the two virtual machines.

```

.encoding = "UTF-8"
config.version = "8"
virtualHW.version = "10"
displayName = "OLTP2"
extendedConfigFile = "OLTP2.vmx"
svg.vramSize = "8388608"
numvcpus = "60"
memSize = "307200"
sched.cpu.units = "mhz"
tools.upgrade.policy = "manual"
scsi0.virtualDev = "pvscsi"
scsi0.present = "TRUE"
scsi1.virtualDev = "pvscsi"
scsi1.present = "TRUE"
scsi2.virtualDev = "pvscsi"
scsi2.present = "TRUE"
scsi3.virtualDev = "pvscsi"
scsi3.present = "TRUE"
sata0.present = "TRUE"
floppy0.startConnected = "FALSE"
floppy0.clientDevice = "TRUE"
floppy0.fileName = "vmware-null-remote-floppy"
sata0:0.startConnected = "FALSE"
sata0:0.deviceType = "cdrom-image"
sata0:0.fileName = "/usr/lib/vmware/isoimages/linux.iso"
sata0:0.present = "TRUE"
scsi0:0.deviceType = "scsi-hardDisk"
scsi0:0.fileName = "OLTP2.vmdk"
sched.scsi0:0.shares = "normal"
scsi0:0.present = "TRUE"
ethernet0.virtualDev = "vmxnet3"
ethernet0.networkName = "VM Network"
ethernet0.addressType = "vpx"
ethernet0.generatedAddress = "00:50:56:9b:47:07"
ethernet0.present = "TRUE"
guestOS = "rhel6-64"
toolScripts.afterPowerOn = "TRUE"
    
```

```
toolScripts.afterResume = "TRUE"
toolScripts.beforeSuspend = "TRUE"
toolScripts.beforePowerOff = "TRUE"
tools.syncTime = "FALSE"
uuid.bios = "42 1b f7 b5 0b d8 a2 35-4f e4 82 32 1d fd 85 47"
vc.uuid = "50 1b a0 6a 6a e3 19 82-0c 0f 30 22 d8 cb 58 77"
sched.cpu.min = "0"
sched.cpu.shares = "normal"
sched.mem.min = "0"
sched.mem.minSize = "0"
sched.mem.shares = "normal"
uuid.location = "56 4d 89 58 2c fc 49 b4-82 93 a1 8b 1a dc 6b c0"
cleanShutdown = "FALSE"
scsil:1.deviceType = "scsi-hardDisk"
scsil:1.fileName = "OLTP2_2.vmdk"
scsil:1.mode = "independent-persistent"
sched.scsil:1.shares = "normal"
scsil:1.present = "TRUE"
scsil:2.deviceType = "scsi-hardDisk"
scsil:2.fileName = "OLTP2_3.vmdk"
scsil:2.mode = "independent-persistent"
sched.scsil:2.shares = "normal"
scsil:2.present = "TRUE"
scsil:3.deviceType = "scsi-hardDisk"
scsil:3.fileName = "OLTP2_4.vmdk"
scsil:3.mode = "independent-persistent"
sched.scsil:3.shares = "normal"
scsil:3.present = "TRUE"
scsil:4.deviceType = "scsi-hardDisk"
scsil:4.fileName = "OLTP2_5.vmdk"
scsil:4.mode = "independent-persistent"
sched.scsil:4.shares = "normal"
scsil:4.present = "TRUE"
scsil:5.deviceType = "scsi-hardDisk"
scsil:5.fileName = "OLTP2_6.vmdk"
scsil:5.mode = "independent-persistent"
sched.scsil:5.shares = "normal"
scsil:5.present = "TRUE"
scsi2:0.deviceType = "scsi-hardDisk"
scsi2:0.fileName = "OLTP2_7.vmdk"
scsi2:0.mode = "independent-persistent"
sched.scsil:0.shares = "normal"
scsi2:0.present = "TRUE"
scsi2:1.deviceType = "scsi-hardDisk"
scsi2:1.fileName = "OLTP2_8.vmdk"
scsi2:1.mode = "independent-persistent"
sched.scsil:1.shares = "normal"
scsi2:1.present = "TRUE"
scsi2:2.deviceType = "scsi-hardDisk"
```

```
scsi2:2.fileName = "OLTP2_9.vmdk"
scsi2:2.mode = "independent-persistent"
sched.scsi2:2.shares = "normal"
scsi2:2.present = "TRUE"
scsi2:3.deviceType = "scsi-hardDisk"
scsi2:3.fileName = "OLTP2_10.vmdk"
scsi2:3.mode = "independent-persistent"
sched.scsi2:3.shares = "normal"
scsi2:3.present = "TRUE"
scsi2:4.deviceType = "scsi-hardDisk"
scsi2:4.fileName = "OLTP2_11.vmdk"
scsi2:4.mode = "independent-persistent"
sched.scsi2:4.shares = "normal"
scsi2:4.present = "TRUE"
scsi2:5.deviceType = "scsi-hardDisk"
scsi2:5.fileName = "OLTP2_12.vmdk"
scsi2:5.mode = "independent-persistent"
sched.scsi2:5.shares = "normal"
scsi2:5.present = "TRUE"
scsi3:0.deviceType = "scsi-hardDisk"
scsi3:0.fileName = "OLTP2_13.vmdk"
scsi3:0.mode = "independent-persistent"
sched.scsi3:0.shares = "normal"
scsi3:0.present = "TRUE"
scsi3:1.deviceType = "scsi-hardDisk"
scsi3:1.fileName = "OLTP2_14.vmdk"
scsi3:1.mode = "independent-persistent"
sched.scsi3:1.shares = "normal"
scsi3:1.present = "TRUE"
scsi3:2.deviceType = "scsi-hardDisk"
scsi3:2.fileName = "OLTP2_15.vmdk"
scsi3:2.mode = "independent-persistent"
sched.scsi3:2.shares = "normal"
scsi3:2.present = "TRUE"
scsi3:3.deviceType = "scsi-hardDisk"
scsi3:3.fileName = "OLTP2_16.vmdk"
scsi3:3.mode = "independent-persistent"
sched.scsi3:3.shares = "normal"
scsi3:3.present = "TRUE"
scsi3:4.deviceType = "scsi-hardDisk"
scsi3:4.fileName = "OLTP2_17.vmdk"
scsi3:4.mode = "independent-persistent"
sched.scsi3:4.shares = "normal"
scsi3:4.present = "TRUE"
scsi1:0.deviceType = "scsi-hardDisk"
scsi1:0.fileName = "OLTP2_1.vmdk"
scsi1:0.mode = "independent-persistent"
sched.scsi1:0.shares = "normal"
sched.scsi1:0.vFlash.enabled = "FALSE"
scsi1:0.present = "TRUE"
```

```
scsi0:1.deviceType = "scsi-hardDisk"
scsi0:1.fileName = "OLTP2_18.vmdk"
sched.scsi0:1.shares = "normal"
sched.scsi0:1.vFlash.enabled = "FALSE"
scsi0:1.present = "TRUE"
sched.cpu.latencySensitivity = "normal"
tools.guest.desktop.autolock = "FALSE"
pciBridge7.functions = "8"
hpet0.present = "TRUE"
virtualHW.productCompatibility = "hosted"
cpuid.coresPerSocket = "30"
nvram = "OLTP2.nvram"
sched.scsi0:0.throughputCap = "off"
pciBridge0.present = "true"
vmci.filter.enable = "true"
svga.present = "TRUE"
replay.supported = "FALSE"
pciBridge0.pciSlotNumber = "17"
pciBridge4.pciSlotNumber = "21"
pciBridge5.pciSlotNumber = "22"
pciBridge6.pciSlotNumber = "23"
pciBridge4.present = "true"
pciBridge7.pciSlotNumber = "24"
scsi0.pciSlotNumber = "160"
scsi1.pciSlotNumber = "192"
scsi2.pciSlotNumber = "224"
scsi3.pciSlotNumber = "256"
ethernet0.pciSlotNumber = "1184"
sata0.pciSlotNumber = "33"
scsi0.sasWWID = "50 05 05 65 0b d8 a2 30"
scsi1.sasWWID = "50 05 05 65 0b d8 a3 30"
scsi2.sasWWID = "50 05 05 65 0b d8 a0 30"
scsi3.sasWWID = "50 05 05 65 0b d8 a1 30"
vmotion.checkpointFBSize = "8388608"
softPowerOff = "FALSE"
pciBridge4.virtualDev = "pcieRootPort"
toolsInstallManager.lastInstallError = "0"
tools.remindInstall = "FALSE"
toolsInstallManager.updateCounter = "1"
pciBridge4.functions = "8"
pciBridge5.present = "true"
pciBridge5.virtualDev = "pcieRootPort"
pciBridge5.functions = "8"
pciBridge6.present = "true"
pciBridge6.virtualDev = "pcieRootPort"
pciBridge6.functions = "8"
pciBridge7.present = "true"
pciBridge7.virtualDev = "pcieRootPort"
sched.swap.derivedName = "/vmfs/volumes/5441d179-885c1795-2027-
```

```

9cb6549b0608/OLTP2/OLTP2-73d3a1c2.vswp"
unity.wasCapable = "FALSE"
sched.scsi1:1.throughputCap = "off"
sched.scsi1:2.throughputCap = "off"
sched.scsi1:3.throughputCap = "off"
sched.scsi1:4.throughputCap = "off"
sched.scsi1:5.throughputCap = "off"
sched.scsi2:0.throughputCap = "off"
sched.scsi2:1.throughputCap = "off"
sched.scsi2:2.throughputCap = "off"
sched.scsi2:3.throughputCap = "off"
sched.scsi2:4.throughputCap = "off"
sched.scsi2:5.throughputCap = "off"
sched.scsi3:0.throughputCap = "off"
sched.scsi3:1.throughputCap = "off"
sched.scsi3:2.throughputCap = "off"
sched.scsi3:3.throughputCap = "off"
sched.scsi3:4.throughputCap = "off"
sched.scsi1:0.throughputCap = "off"
sched.scsi0:1.throughputCap = "off"
numa.vcpu.preferHT = "TRUE"
monitor_control.disable_flexpriority = "FALSE"
replay.filename = ""
scsi0:0.redo = ""
scsi0:1.redo = ""
scsi1:0.redo = ""
scsi1:1.redo = ""
scsi1:2.redo = ""
scsi1:3.redo = ""
scsi1:4.redo = ""
scsi1:5.redo = ""
scsi2:0.redo = ""
scsi2:1.redo = ""
scsi2:2.redo = ""
scsi2:3.redo = ""
scsi2:4.redo = ""
scsi2:5.redo = ""
scsi3:1.redo = ""
scsi3:0.redo = ""
scsi3:2.redo = ""
scsi3:3.redo = ""
scsi3:4.redo = ""
featMask.vm.cpuid.SMEP = "Val:0"
mem.hotadd = "FALSE"
sched.cpu.affinity = "all"
vmci0.present = "FALSE"
numa.nodeAffinity = "2, 3"

```

## Appendix J: Auxiliary Worldlet Pinning Script for Virtual Machine

A vsish command was used in a script as below to bind auxiliary worldlets to physical CPUs:

```
export PCPU0=17
export PCPU1=16
export CCPU=$PCPU0
export pcount=0

for worldlet in `vsish -e ls /worldlets/objs/ | awk 'BEGIN { FS = "." };
{ print $2, $1}' | sort | awk '{print $2 "." $1 }'`;
do
    vsish -e set /worldlets/objs/${worldlet}cpuAffinity $CCPU
    echo vsish -e set /worldlets/objs/${worldlet}cpuAffinity $CCPU

    pcount=`expr $pcount + 1`

    if [ $pcount -eq 1 ] ;
    then
        pcount=0
        if [ $CCPU -eq $PCPU0 ]
        then
            CCPU=$PCPU1
        else
            CCPU=$PCPU0
        fi
    fi
done
```

## For more information

[Universal Database Solution for Mission-Critical x86](#)

[HP ProLiant DL580 Technology](#)

[HP ProLiant DL580 Gen8 Server](#)

[HP 3PAR StoreServ 7450 All-flash storage](#)

[HP 3PAR Remote Copy Software](#)

[HP Serviceguard Solutions for Linux](#)

[HP Networking](#)

To help us improve our documents, please provide feedback at [hp.com/solutions/feedback](http://hp.com/solutions/feedback)

### Sign up for updates

[\*\*hp.com/go/getupdated\*\*](http://hp.com/go/getupdated)



Share with colleagues



Rate this document

---

© Copyright 2015 Hewlett-Packard Development Company, L.P. The information contained herein is subject to change without notice. The only warranties for HP products and services are set forth in the express warranty statements accompanying such products and services. Nothing herein should be construed as constituting an additional warranty. HP shall not be liable for technical or editorial errors or omissions contained herein.

Intel and Intel Xeon are trademarks of Intel Corporation in the U.S. and other countries. Microsoft and Windows are trademarks of the Microsoft group of companies. Oracle is a registered trademark of Oracle and/or its affiliates. Red Hat is a registered trademark of Red Hat, Inc. in the United States and other countries. Linux is the registered trademark of Linus Torvalds in the U.S. and other countries. VMware is a registered trademark or trademark of VMware, Inc. in the United States and/or other jurisdictions.

