WHITE PAPER

A Performance Study of Hitachi Unified Compute Platform (UCP) Solution for Oracle Database

By Roger Bitar and Janos Horvath
Hitachi Data Systems

October 2016

Contents

Executive Summary

This white paper describes a performance study for Hitachi Unified Compute Platform (UCP) for Oracle Database with the Real Application Clusters option solution. The study, based on a proof of concept conducted for a large financial institution, uses Hitachi Unified Compute Platform 6000 for Oracle Real Application Clusters (UCP 6000 for Oracle RAC), with Hitachi Compute Blade 2500 and Hitachi Virtual Storage Platform G800 with Hitachi Accelerated Flash. It includes the performance and scalability benefits of this infrastructure for typical online transaction processing (OLTP) and data warehouse workloads.

This white paper includes the following:

- Hitachi Compute Blade 2500 with four CB520X B3 server blades. Each two CB520X B3 blades are connected via a two-node SMP connector, and represent one Oracle RAC node.
- Hitachi Virtual Storage Platform G800 with Hitachi Accelerated Flash.

Introduction

Oracle Databases are the foundation for transaction and analytics processing in many mission-critical applications that support today's businesses. Oracle Databases are used to hold the data for many business applications, including human capital management, supply chain management, customer relationship management, as well as financial, healthcare, media and entertainment applications. All electronic commerce operations rely on databases in the back end to hold customer information, such as creditworthiness and shopping history, to track inventory, as well as to maintain customer reviews. As of 2015, according to IDC, Oracle has owned about 48% of the relational database management system (RDBMS) market share for over a decade.

Converged Infrastructure

Over 60% of the Global Fortune 100 companies trust their Oracle Database deployment to Hitachi storage, which is known for its 99.999 (often called "five nines") uptime availability. This availability works out to 5.26 minutes of total downtime, planned or unplanned, in a given year and translates to 25.9 seconds per month, or 6.05 seconds per week of downtime.

Following Oracle's acquisition of Sun Microsystems in 2010, Oracle introduced version 2 of the Exadata engineered system: a database appliance that incorporates servers, networking and storage in a pre-configured rack that is optimized to run Oracle Databases. Oracle positioned the Exadata appliance as a database consolidation platform, and as a replacement for legacy IT resources that were deployed in a silo-like fashion. Oracle's Exadata appliance promised rapid deployment of a complete infrastructure, scalability, versatility and flexibility in the data center, in addition to screaming performance in terms of I/O operations per second or IOPS.

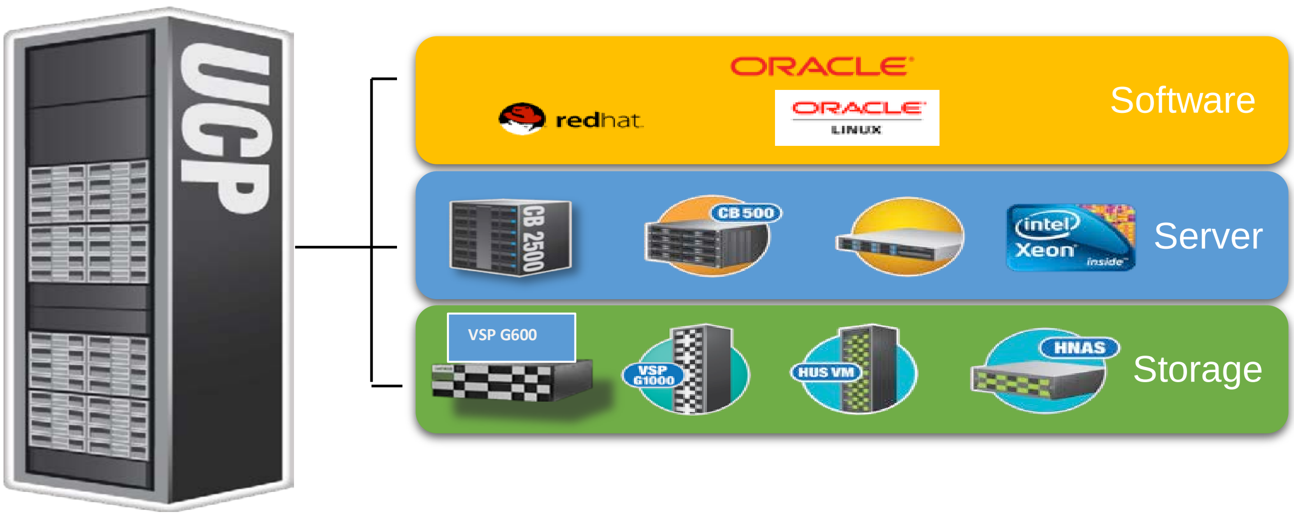
Following Oracle's successful introduction of the Exadata database appliance, many IT infrastructure vendors have followed suit, bringing their converged infrastructure to market. A converged infrastructure includes compute, networking and storage resources that were tested and certified to work together in a single rack. A converged infrastructure provides the following benefits:

- A centralized management infrastructure to manage all the components, freeing up IT administrators and resources to focus more on the need of the business.
- Lower costs, since consolidating all the components reduces the data center footprint and has a positive impact on maintenance costs, power consumption and cooling.
- Increased agility and faster provisioning, as the hardware resources can be easily repositioned to support any application or workload.
- Scalability and flexibility, as new computing or storage resources can be quickly added to support additional workloads and applications.
- Faster path to the private cloud, leveraging the many virtualization technologies.
- Faster time to value with the certified infrastructure, since all the components were tested and optimized to work together.
- Better support model, since all the components are usually provided by the same vendor, as customers are usually provided a single point of contact for support.

Hitachi Unified Compute Platform (UCP)

Hitachi Unified Compute Platform (UCP), as shown in Figure 1, consists of a compute layer using Hitachi Compute Blade chassis or Hitachi Compute Rack rackmount servers. UCP includes a network layer with redundant IP networking and optional SAN switches for high availability, and a choice of Hitachi storage arrays for the storage layer. Hitachi UCP supports Red Hat Linux RHEL, and Oracle Linux, starting with the v6.6 release, and Oracle UEK 4.1 release.

Figure 1. Hitachi Unified Compute Platform (UCP)



UCP = Hitachi Unified Compute Platform, CB = Hitachi Compute Blade, VSP = Hitachi Virtual Storage Platform, HUS VM = Hitachi Unified Storage VM, HNAS = Hitachi NAS Platform

Hitachi Unified Compute Platform 6000 for Oracle Real Application Clusters Solutions

The Hitachi UCP 6000 for Oracle RAC solutions come in small, medium, large and extra-large configurations to meet varying customer needs. These Hitachi converged solutions for Oracle support Oracle Database versions 11 and 12c, in both RAC and single-instance deployments. Hitachi solutions for Oracle deployment offer:

- Lower total cost of ownership (TCO) with reduced infrastructure and license costs.
- High performance to meet business demands with 16Gbit Fibre Channel (FC) storage connections and Hitachi Accelerated Flash (HAF), also known as flash modules drives or FMD, for fast data access.
- Robust data protection with enterprise-class storage.
- Easy scalability from small to medium by adding an additional processor per blade (in a two-blade configuration for high availability).
- Easy scalability from medium to large by adding two more blades, and attaching an identical storage array.

- Extra-large configuration featuring SMP connector and our largest Hitachi Virtual Storage Platform G1000 (VSP G1000) storage array.
- Guaranteed business continuity with a high-availability, disaster recovery, backup and recovery, data protection solution.
- Highly reliable virtual environment with Hitachi logical partitioning technology (LPAR).
- Flexible configuration to meet customer needs.
- Latest CPUs from Intel, including large number of cores to support cloud deployment, and smaller core CPUs to mitigate the impact of Oracle Database licensing costs.
- Open architecture supporting new and legacy Oracle applications while avoiding vendor lock-in and proprietary high-cost platforms.
- Ease of management with tight integration with Oracle Enterprise Manager, and a single management software for server and storage.
- A turnkey solution, whereby all components are racked, stacked, installed, tested at the factory and shipped to the customer as an appliance.
- Faster time to value with a pre-integrated and pre-configured certified converged infrastructure optimized for Oracle database 12c.

The various components and technologies featured in the UCP 6000 for Oracle RAC are described below.

Hitachi Compute Blade 2500

The Hitachi Compute Blade 2500 (CB 2500) integrates network, I/O and server resources into a space-efficient, flexible solution. The rack-mountable, 12U chassis houses up to 14 CB520H server blade modules or up to eight full-width CB520X blades. Both half-width, and full-width blades can hold up to two sockets running with the latest Intel Xeon E5 and E7 family processors. For I/O versatility, there are two bays for internal network switches, and 28 PCIe Gen 3 expansion slots. The CB520X blades can be combined into two or four blades to create a single eight-socket SMP system with 192 memory DIMM slots by attaching an SMP connector, thus providing the scalability needed to run the most demanding applications.

With sophisticated, built-in reliability, availability and serviceability features, Hitachi Compute Blade 2500 is an ideal data center platform for the consolidation of mission-critical applications, virtualization and cloud computing applications. CB 2500 features Hitachi Logical Partitioning (LPAR) virtualization technology, which makes it possible to run multiple independent applications in completely isolated environments, and is certified and supported by Oracle. CB 2500 supports the latest available releases of the Linux operating system from Red Hat, and from Oracle.

Hitachi Compute Blade 2500 maintains high uptime levels through sophisticated failover mechanisms. The N+1 cold standby function enables multiple servers to share a standby server. It increases system availability while decreasing the need for multiple standby servers, or costly software-based high-availability solutions. It enables the system to detect a fault in a server blade and switch to the standby server, manually or automatically. The hardware switching is executed even in the absence of the administrator, enabling the system to return to normal operations within minutes.

Hitachi Storage Virtualization Operating System (SVOS)

Hitachi Storage Virtualization Operating System (SVOS) is the standard operating system for Hitachi Virtual Storage Platform G series and VSP F series. SVOS works with the virtualization capabilities of VSP and provides the foundation for global storage virtualization. It delivers software-defined storage by abstracting and managing heterogeneous storage to provide a unified virtual storage layer, resource pooling, and automation. SVOS also offers self-optimization, automation, and increased operational efficiency for improved performance and storage utilization.

Hitachi SVOS features and their capabilities include:

- **Global storage virtualization capability** provides an active-active clustering environment spanning multiple storage systems.
- **Hitachi Universal Volume Manager** enables virtualization of external heterogeneous storage.
- **Hitachi Dynamic Provisioning** provides thin provisioning for simplified provisioning operations, automatic performance optimization, and storage space savings.
- **Hitachi Device Manager** provides a single point of management for all Hitachi physical and virtualized storage, and it acts as the interface for integration with other management systems.
- **Hitachi Storage Advisor** is an alternative to Device Manager that provides a simpler configuration and management environment for less demanding storage environments where advanced replication or global storage virtualization is not required.
- **Hitachi Resource Partition Manager** supports secure administrative partitions for multitenancy requirements.
- **Cache partition manager** feature supports up to 32 cache partitions.
- **Hitachi Dynamic Link Manager Advanced** provides advanced SAN multipathing with centralized management.
- **Hitachi Storage Navigator** provides VSP element storage management functions.
- **Performance monitoring feature** provides an intuitive, graphical interface to assist with performance configuration planning, workload balancing and analyzing and optimizing storage system performance.
- **Storage-system-based utilities** include LUN manager, Hitachi Data Retention Utility, Hitachi Server Priority Manager, Audit Log and Volume Shredder.
- **Standard management interface support** includes SMI-S provider and SNMP agent.

Nondisruptive Migration

One of the biggest challenges during technology refresh cycles is to eliminate downtime and service disruption when the data used by the host is copied to a new volume on the new storage system and the host is reconfigured to access the new volume. Nondisruptive migration makes it possible to relocate data from existing storage systems to newer storage systems, without interrupting access by hosts. Data migration is accomplished using the global storage virtualization technology of VSP G series storage systems. Resources on the source storage system are virtualized on the target storage system. From the perspective of the host, I/O requests continue to be serviced by the source storage system during the migration process.

Hitachi Accelerated Flash

The second generation of Hitachi Accelerated Flash (HAF) flash module is built from the ground up to support concurrent, large-I/O enterprise workloads. HAF enables hyperscale efficiencies for applications such as large-scale transaction processing, online transaction processing (OLTP) databases and online analytical processes. At its core is an advanced embedded multicore flash controller that increases the performance of multilayer cell (MLC) flash to levels that exceed levels of costlier single-level cell (SLC) flash solid state disks (SSDs). Its inline compression offload engine and enhanced flash translation layer deliver up to 80% data reduction (typically 2:1) without slowing down operations.

Hitachi Accelerated Flash delivers outstanding value compared to enterprise SSD. Each HAF device delivers a sustained performance of:

- 150k IOPS for random reads: 8KB blocks.
- 80k IOPS for random writes: 8KB blocks.
- 2.0 GB/sec sequential reads (GB/sec).
- 1.0 GB/sec sequential reads (GB/sec).

In addition, Hitachi Accelerated Flash modules provide superior data integrity while extending flash cell longevity. In addition, Hitachi Accelerated Flash modules feature workload priority access to ensure end-to-end application workload quality of service (QoS). This QoS extends from a virtual machine down to an individual NAND flash chip, over typical flash background tasks such as garbage collection and data refresh.

Hitachi Virtual Storage Platform Gx00

VSP Gx00 is a modular, rackmountable, storage system. The storage system has dual controllers that provide the interface to the disks. Each controller contains its own processor, dual in-line cache memory modules (DIMMs), cache flash memory (CFM), battery and fans. Each controller also has an Ethernet connection for out-of-band management using Hitachi Device Manager - Storage Navigator. The VSP Gx00 and all-flash VSP Fx00 storage systems feature complete system redundancy, hot-swappable parts, nondisruptive firmware upgrades, and hardware-based compression. If the data path through one controller fails, all hard drives remain available to connected hosts using a redundant data path through the other controller, until the defective controller is replaced. Self-contained, hardware-based RAID logical drives provide maximum performance in compact external enclosures.

Hitachi Virtual Storage Platform G600

The midrange VSP G600 system consists of a 4U enclosure that includes two clustered controllers per system. Each controller contains two Intel Xeon processors and up to 128GB of cache. Each controller supports a fixed number of slots, which can accept I/O modules that provide the connectivity to hosts via front-end modules, as well as the serial-attached SCSI (SAS) connectivity to the disks via back-end modules. Up to four front-end modules per controller are supported, which can connect to up to fifty-six 8GB Fibre Channel ports, or twenty-eight 16GB Fibre Channel ports, or twenty-eight 10Gbit/sec iSCSI ports. Each of the two back-end modules contain a single SAS protocol controller chip with optional encryption that provides two disk tray cable ports, each port having 4 x 12Gbit/sec SAS wide links. The VSP G600 can be attached to four types of optional disk boxes:

- DBS: 2U 24-slot small form factor (SFF) SAS box.
- DBL: 2U 12-slot large form factor (LFF) SAS box.
- DB60: 4U 60-slot dense LFF SAS drawer.
- DBF: 2U 12-slot FMD flash module HAF box.

The DBS box holds up to 24 of the 2.5" SAS drives. These include all the 10K hard disk drives (HDD), 15K HDD, and SSD options currently available. The DBL holds up to 12 of the 3.5" SAS drives. These are available as 4TB, or 6TB 7200 RPM drives. The DB60 drawer holds up to 60 of the 4TB or 6TB drives or, optionally, the 400GB SSD or 1.2TB 10K HDD with a special LFF conversion canister. The DBF box holds 12 of the proprietary flash module drives, which are available in 1.6TB, 3.2TB and 6.4TB capacities. The VSP G600 can connect to up to 24 DBL, DBS or DBF disk boxes, supporting a maximum of 288 LFF disks, or 576 SFF disks, or 288 FMDs, or up to 12 DB60 dense disk boxes, supporting a maximum of 720 LFF disks.

Hitachi Virtual Storage Platform G800

The midrange VSP G800 system is identical to VSP G600, but it provides higher performance and connectivity. VSP G800 supports up to 512GB of DRAM cache, sixty-four 8GB Fibre Channel ports, or thirty-two 16GB Fibre Channel ports, or thirty-two 10GbE iSCSI ports. VSP G800 can connect to up to 48 DBL, DBS or DBF disk boxes, supporting a maximum of 576 LFF disks, or 1152 SFF disks, or 576 FMDs, or up to 24 DB60 dense disk boxes, supporting a maximum of 1440 LFF disks.

Hitachi Dynamic Provisioning

Hitachi Dynamic Provisioning (HDP) is a volume management feature that allows storage managers and system administrators to efficiently plan and allocate storage to users or applications. It provides a platform for the array to dynamically manage data and physical capacity without frequent manual involvement. Dynamic Provisioning provides three important capabilities: thin provisioning of storage, enhanced volume performance and larger volume sizes.

Hitachi Dynamic Provisioning is more efficient than traditional provisioning strategies. It is implemented by creating one or more Dynamic Provisioning pools (DP pools) of physical storage space using multiple logical devices or LDEVs. Virtual DP volumes (DP-VOLs) can then be established and connected to the individual DP pools. DP-VOLs are of a user-specified logical size without any corresponding physical space. Actual physical space (in 42MB pool page units) is automatically assigned to a DP-VOL from the connected DP pool as that volume's logical space is written to over time. A new DP volume does not have any pool pages assigned to it. The pages are loaned out from its connected pool to that DP volume until the volume is reformatted or deleted. At that point, all of that volume's assigned pages are returned to the pool's free page list. In this way, capacity to support data can be randomly assigned on demand within the pool, rather than pre-allocating a huge amount of physical storage that may not be fully utilized, and potentially wasted. This handling of logical and physical capacity is called thin provisioning. In many cases, logical capacity will exceed physical capacity. Thin provisioning provides cost-effective capacity planning without disrupting operations.

Dynamic Provisioning software also combines many application I/O patterns and spreads the I/O activity across all available physical resources. This helps to alleviate disk "hot spots" and reduces disk contentions and performance bottlenecks.

Hitachi Dynamic Tiering

Hitachi Dynamic Tiering (HDT) simplifies storage administration by automatically optimizing data placement in one, two or three tiers of storage that can be defined and used within a single virtual volume. Tiers of storage can be made up of internal or external (virtualized) storage.

With HDT, data is written to the highest performance tier first. As data becomes less active, it migrates to lower-level tiers. If activity increases, data will be promoted back to a higher tier. All this is done automatically, and without administrators' involvement. Since 20% of data accounts for 80% of the activity, only the active part of a volume will reside on the higher performance tiers. By keeping hot data on flash storage for highest performance, and migrating cold data to lower cost high-capacity drives, HDT helps to improve performance and reduce operational costs by ensuring applications data are placed on the appropriate classes of storage.

Data Protection and Recovery

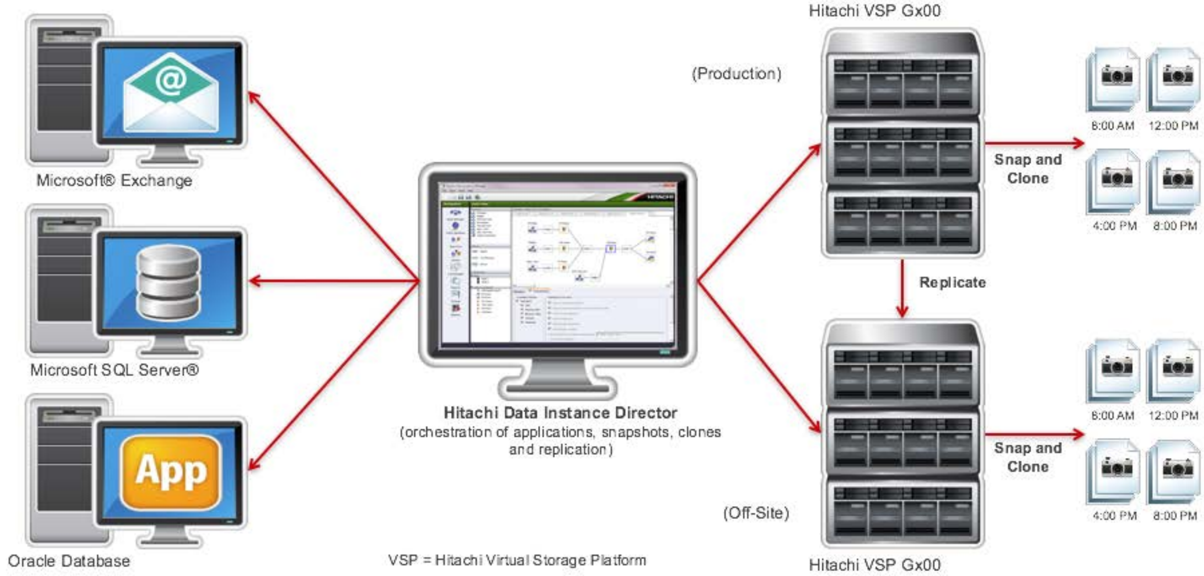
Hitachi Data Instance Director (HDID) unifies and simplifies data protection and recovery with copy data management. Through its unified, whiteboard-like, workflow-based policy engine, the data managed by HDID can be backed up, archived, used in operational and disaster recovery, as well as test and development, and many other purposes. HDID automates and orchestrates Hitachi storage-based copy technologies, including snapshots, clones, synchronous and asynchronous replication, all from a very user-friendly graphical user interface (GUI).

HDID is presented in an easy-to-use whiteboard style user interface that helps to map data protection processes to business priorities. HDID includes a wide range of fully integrated capabilities, including:

- Storage-based (block): Hitachi Thin Image snapshot, Hitachi ShadowImage clone, Hitachi TrueCopy synchronous replication, and Hitachi Universal Replicator asynchronous replication.
- Storage-based (file): Directory clone and file replication.
- Host-based: Continuous data protection (CDP), batch backup, archiving, software snapshots, repository replication, deduplication, bare metal recovery (BMR) and more.

These capabilities can be combined into complex workflows that automate and simplify copy data management, as shown in Figure 2.

Figure 2. Data Protection With HDID



HDID provides the following business benefits:

- Database-and application-aware snapshot and replication management software that automatically puts the database into a backup-ready quiesced state, executes the storage-based snapshot, and then releases the database or application to continue normal operations. This approach makes it possible to have fast, fully application-consistent operational recovery.
- Improved application and data availability to meet stringent service level agreements (SLAs), such as recovery point objective (RPO), recovery time objective (RTO) and retention.
- Faster, more frequent protection operations result in less data to recover following a disaster.
- Eliminates the need for siloed point solutions such as applications-based backup solutions, which can help to cut costs and risks.
- Dramatically reduces storage requirements and lowers capital and maintenance costs.

With HDID, you can orchestrate same or different snapshots schedules at multiple sites, have identical or different RPOs at each site, and choose same or different retention policies at each site. Storage-based snapshot and replication technologies makes it possible to:

Modern Data Protection With HDID
LEARN MORE

- Eliminate the impact of data protection operations from the database system. Production systems are not impacted by backup operations, which tend to be CPU intensive. You can be confident that your infrastructure will be able to continue supporting the heaviest load, without impacting your data protection plans.
- Eliminate the need of a backup window and associated downtime. This is crucial for critical, always-on applications, where any backup window may be unacceptable.
- Enable much more frequent backup operations, reducing the amount of data at risk by 90% or more.

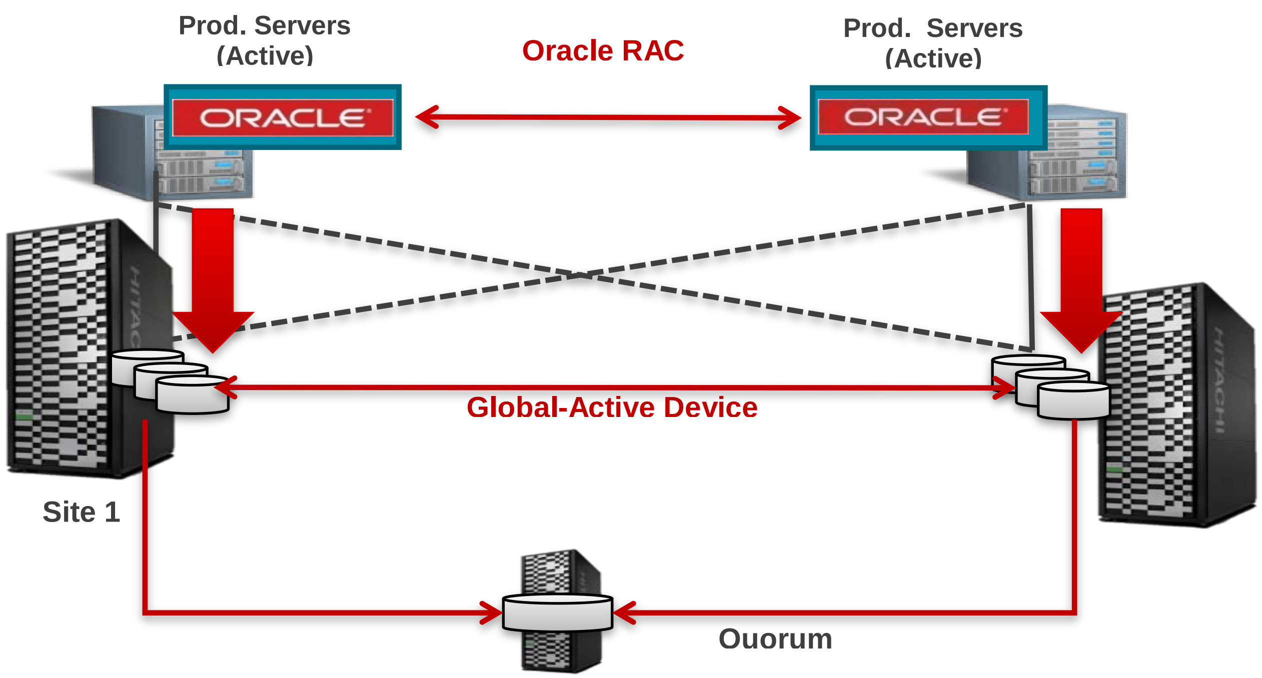
Global-Active Device

Global-active device is a feature of Hitachi Storage Virtualization Operating System. It simplifies and automates high availability to ensure continuous operations for your most mission-critical data and applications. Global-active device enables active-active continuous operations for production workloads on systems in different sites, while maintaining full data consistency and protection. This is achieved by continuous mirroring of storage volumes between two VSP systems that accept read/write I/O operations on both sides and are continuously updated, as shown in Figure 3. If a disk controller becomes unavailable at one site due to technical or catastrophic failure, the controller at the other site automatically takes over and accepts read/write I/O operations.

Where continuous computing is required, global-active device ensures that an up-to-date storage volume is always available. The feature lets you take full advantage of all resources, and also greatly simplifies fault-tolerance design and administrative operations. As a native solution to Hitachi storage, global-active device is the ideal solution for nonstop operations, eliminating the cost and complexity of adding additional appliances and their network connections. Global-active device provides the following benefits:

- Active-active configuration that provides simultaneous access to both sites.
- High availability across data centers for distances up to 100 kilometers or 60 miles.
- Zero RPO and RTO.
- Avoid complex disaster recovery configurations and procedures.
- Bidirectional access across sites.
- Better return on investment (ROI) and return on assets (ROA) with better utilization of disaster recovery site resources.
- Minimal to no impact on business critical workloads when the following occur:
  - Storage array component failure
  - Storage network failure
  - Wide area network failure
- Ability to nondisruptively migrate application workloads to another physical location within metro distance during site maintenance.

Figure 3. Global-Active Device



Benchmark Description

We ran a proof of concept for a large financial institution on our Hitachi UCP. In this POC, we used Benchware benchmark, recently renamed as "peakmarks," in our tests. The key performance metrics measured by this benchmark are listed in the data sheet.

Configuration Details

This section describes the configuration of the system under test. We configured two Oracle RAC cluster nodes, with each RAC node consisting of two CB520X B3 blades connected via a two-node SMP connector, for a total of four Intel Xeon E7-8880 v4 Broadwell processors, or seventy-two cores, and 1TB of total memory per node. The configuration used in this benchmark consists of the following components:

Hardware Configuration

The hardware configuration consisted of the following components:

- 1 x CB 2500.
- 2 x 10 GB LAN switches..
- 4 x CB520X B3 blades configured as two Oracle RAC nodes, by using a two-node SMP connector to connect every two blades together. Each node included 4 x E7-8880v3 2,3GHz – 72 cores CPUs, with 1 TB RAM.

- 1 VSP G800 storage array configured with:
  - 2 Back-end modules with encryption.
  - 32 x 16GB Fibre Channel ports.
  - 6 RAID groups of 1,6TB FMD in RAID-6 (6 FMDs for data + 2 FMDs for parity).
  - 6 RAID groups of 600GB 10K SAS HDDs in RAID-6 (6 HDDs for data + 2 HDDs for parity).

For our tests, we used FMDs to store the database files.

Software Configuration

We used the following software releases in our tests:

- Oracle Database Release: 11.2.0.4.0.
- Oracle Enterprise Manager 12c.
- Hitachi Command Suite 8.4.
- Benchware Release: 8.6 Build 160523.

Database Configuration

The database under test was 11 - 12TB in size, with the following defined parameters:

Tablespace	Capacity [GB]
BENCHWARE	11263.410
SYSAUX	2.031
SYSTEM	0.742
UNDOTBS1	96.000
UNDOTBS2	51.000
USERS	1.520

Tablespace	Capacity [GB]
TEMP	120.000

Tablespace	Used [GB]
BENCHWARE	8404.922
SYSAUX	1.925
SYSTEM	0.739
UNDOTBS1	70.253
UNDOTBS2	37.147
USERS	1.412

Database I/O Calibration

Calibrate I/O is a tool provided by Oracle as of Oracle Database 11g, as part of the DBMS_RESOURCE_MANAGER.CALIBRATE_IO package. From the Oracle performance tuning manual: "*The I/O calibration feature of Oracle Database enables you to assess the performance of the storage subsystem, and determine whether I/O performance problems are caused by the database or the storage subsystem*".

From Oracle Doc ID 727062.1: *"When Calibrate I/O is invoked it will generate I/O intensive read-only random I/O (db_block_size) and large-block (1MByte) sequential I/O workloads. Unlike various external I/O calibration tools, this tool uses the Oracle code stack and runs in the database, issuing I/O against blocks stored in the database. The results, therefore, much more closely match the actual database performance".*

The random block-sized reads (8KB by default) are used for testing the maximum IOPS, and for predicting maximum OLTP database throughput. The sequential large 1MB reads are used to test the maximum data warehouse throughput that can be sustained by the storage subsystem. For an Oracle Real Application Clusters (RAC) database, the workload is simultaneously generated from all instances.

Calibrate I/O was invoked using the following command:

exec dbms_resource_manager.calibrate_io (32, 10, :p_iops, :p_mbps, :p_latency);

- The first input parameter num_physical_disks. From the documentation this parameter should be set to the "approximate number of physical disks in the database storage."
- The second input parameter is max_latency: "Maximum tolerable latency in milliseconds for database-block-sized IO requests.
- The remaining three parameters are output parameters which are the results of the measurements: max_iops, max_mbps, and latency.

The I/O calibration package produced the following result in Figure 4:

Figure 4. Calibrate I/O Test Results

```
                                                    Latency          Max MBps
Day              Begin End   #Disks    Max IOPS     [ms] Max MBps  Process
---------------- ----- ----- ------ ----------- ------- -------- --------
Fri 17-Jun-2016 00:01 00:17      32   7.994E+05       0    12415     1500
```

The results obtained by the Calibrate I/O test show a maximum of 799,400 IOPS for random I/O, and 12,415 MB/sec for sequential I/Os.

Benchmark Results

This section describes the tests performed and the results of the tests.

Sequential Write Tests

For the sequential write tests, we tried to determine the saturation point for one node before moving on to a two-node cluster tests. The test results as presented in Figure 5, show that we hit a saturation point of 64 parallel jobs with one node and reached a maximum throughput of 5,892MB/sec. For the two-node test, we hit a saturation point of 64 parallel jobs and reached a maximum throughput of 7,386 MB/sec.

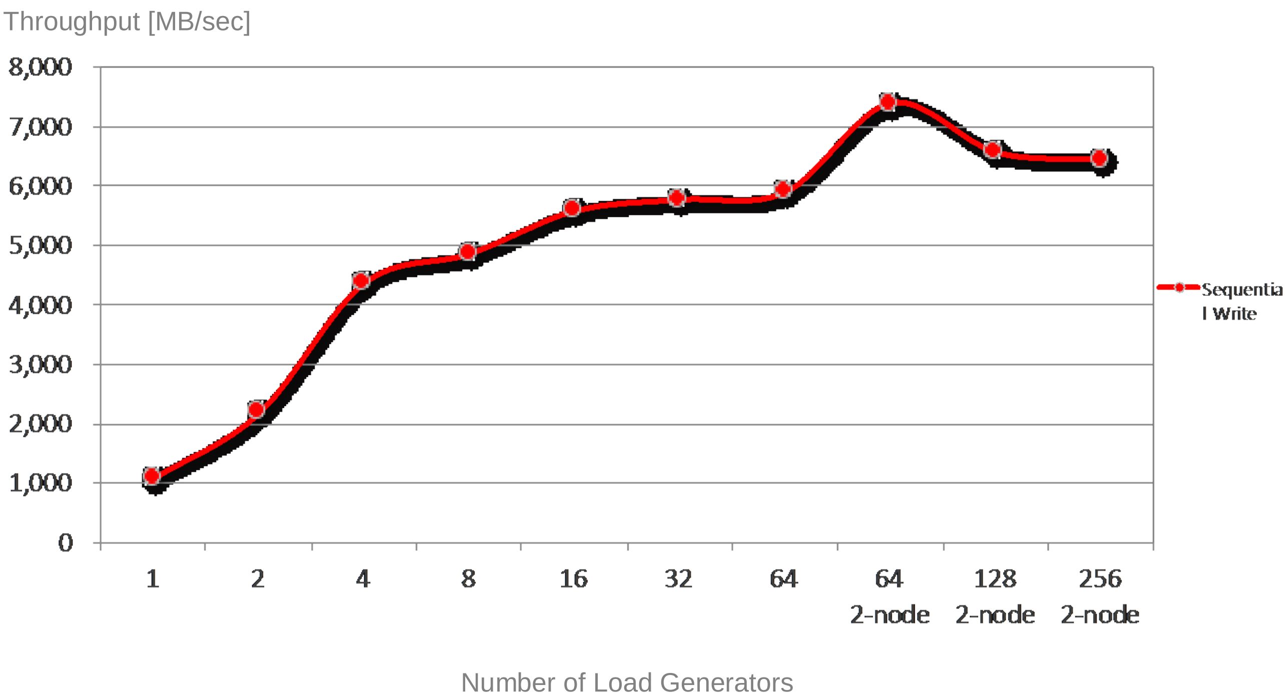
Figure 5. Sequential Write Test Results

Run	Tst	Code	#N	#J	#T	CPU busy [%]	CPU sys [%]	Physical read [iops]	Physical read [dbps]	Physical read [MBps]	Physical write [iops]	Physical write [dbps]	Physical write [MBps]	REDO write [iops]	Hitrate db flash [%]	Hitrate exa flash [%]	Elap time [s]
9	1	STO-22	1	1	1	0	0	14	1	0	1166	34	1093	35	0	0	15
	2	STO-22	1	2	1	1	0	20	0	0	2326	68	2185	64	0	0	15
	3	STO-22	1	4	1	1	0	34	0	1	4618	136	4370	85	0	0	15
	4	STO-22	1	8	1	1	0	44	0	1	5174	201	4856	118	0	0	27
	5	STO-22	1	16	1	1	0	44	0	1	5896	197	5579	98	0	0	47
	6	STO-22	1	32	1	1	0	44	0	1	6021	179	5763	39	0	0	91
	7	STO-22	1	64	1	1	0	47	1	1	6159	195	**5892**	37	0	0	178
	8	STO-22	1	128	1	2	0	46	2	1	5873	184	5624	30	0	0	373
	9	STO-22	2	64	1	2	0	67	7	1	7726	235	**7386**	57	0	0	142
	10	STO-22	2	128	1	2	0	56	2	1	6861	216	6555	51	0	0	320
	11	STO-22	2	256	1	2	0	61	5	1	6738	207	6435	60	0	0	652

- Run = benchmark run id
- Tst = benchmark test id
- Code = benchmark test code
- #N = number of RAC nodes
- #J = number of load generators (jobs)
- [iops] = i/o operations per second
- [dbps] = database blocks per second
- [MBps] = megabyte per second
- [s] = time in seconds

Since we have four 16 Gb/sec Fibre Channel connections to storage, the theoretical maximum data transfer that can be reached is 6.4GB/sec, our results of 5,892MB/sec shows we are very close to reaching the theoretical maximum for throughput. As for the two-nodes test, the theoretical maximum that can be reached is 7.8GB/sec, and we got very close results with 7,386MB/sec. The chart in Figure 6 shows the scalability of the sequential write operations as we increase the number of jobs that were executed in parallel on the x-axis.

Figure 6. Sequential Write Scalability



Sequential Read Tests

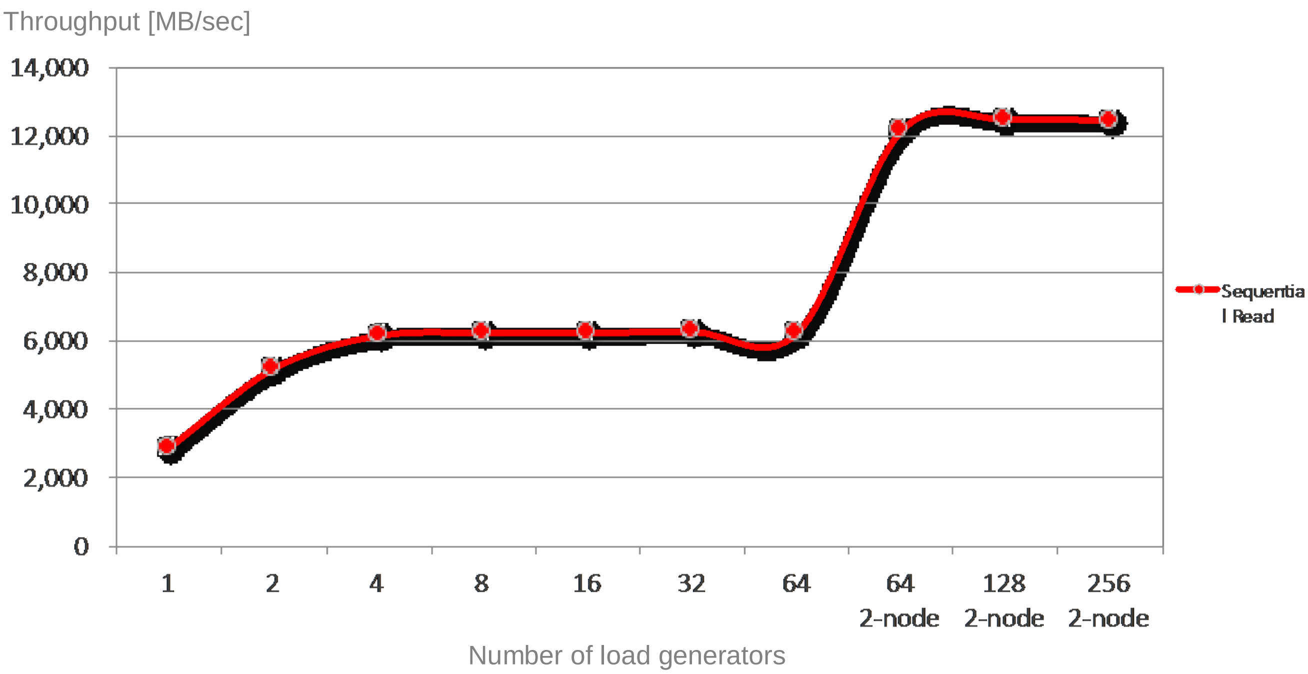
The sequential read tests, as presented in Figure 7, show that we hit a saturation point of 32 parallel jobs and reached a maximum throughput of 6,271MB/sec, very close to the theoretical maximum of 6.4GB/sec. For the two-node test, we hit a saturation point of 128 parallel jobs and reached a maximum throughput of 12,462MB/sec, which is very close to the theoretical maximum of 12.8GB/sec.

Figure 7. Sequential Read Test Results

Run	Tst	Code	#N	#J	#T	CPU busy [%]	CPU sys [%]	Physical read [iops]	Physical read [dbps]	Physical read [MBps]	Physical write [iops]	Physical write [dbps]	Physical write [MBps]	REDO write [iops]	Hitrate db flash [%]	Hitrate exa flash [%]	Elap time [s]
9	21	STO-12	1	1	4	1	0	2915	367939	2875	7	5	0	1	0	0	299
	22	STO-12	1	2	4	1	0	5251	664502	5192	8	6	0	1	0	0	298
	23	STO-12	1	4	4	1	0	6218	788157	6158	7	6	0	1	0	0	302
	24	STO-12	1	8	4	1	0	6302	798758	6240	8	7	0	2	0	0	303
	25	STO-12	1	16	4	1	0	6304	798781	6241	7	5	0	1	0	0	308
	26	STO-12	1	32	4	1	0	6333	802626	6271	6	4	0	1	0	0	314
	27	STO-12	1	64	4	1	0	6294	797586	6231	7	5	0	2	0	0	321
	28	STO-12	2	64	4	1	0	12256	1555018	12149	9	7	0	2	0	0	328
	29	STO-12	2	128	4	2	0	12571	1595154	12462	9	6	0	3	0	0	321
	30	STO-12	2	256	4	2	0	12533	1590206	12424	8	6	0	4	0	0	322

The chart in Figure 8 shows the scalability of the sequential read operations as we increase the number of jobs that were executed in parallel on the x-axis.

Figure 8. Sequential Read Scalability



In the above test, we reached a maximum throughput of 12,462MB/sec. This result matches the result obtained by the Calibrate I/O test of 12,415MB/sec in Figure 4.

Random Read Tests

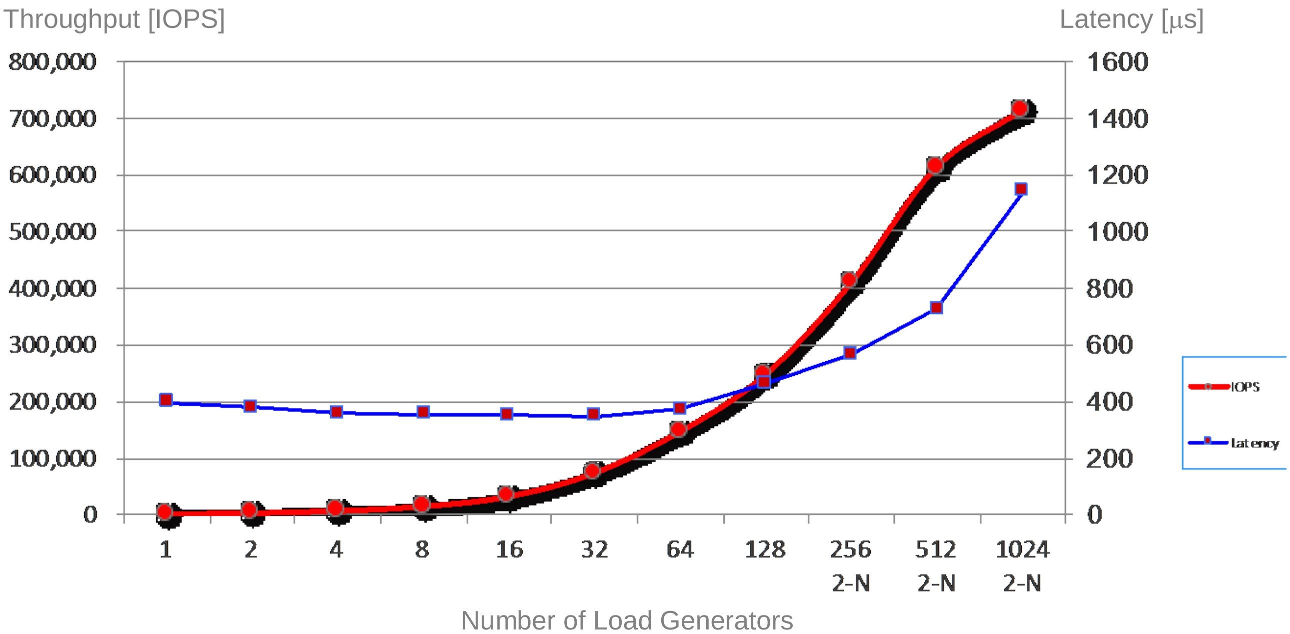
The random read tests were executed using an 8KB block size, as shown in Figure 9. The tests show that we hit a saturation point of 512 parallel read operations, and reached a maximum of 416,991 IOPS, with a latency of 0.5 msec as shown in Figure 10. For the two-node test, we hit a saturation point of 1024 parallel read operations, and reached a maximum of 713,848 IOPS with very low latency 1.1 msec, as shown in Figure 10.

Figure 9. Random Read Test Results

CPU Run	CPU Tst	Physical Code	#N	Physical #J	#T	Physical busy [%]	sys [%]	Physical read [iops]	Physical read [dbps]	Physical read [MBps]	REDO write [iops]	Hitrate write [dbps]	Hitrate write [MBps]	Elap write [iops]	db flash [%]	exa flash [%]	time [s]
9	31	STO-62	1	1	1	0	0	1614	1608	13	9	6	0	1	0	0	305
	32	STO-62	1	2	1	1	0	3450	3443	27	7	5	0	1	0	0	300
	33	STO-62	1	4	1	1	0	7162	7155	56	5	4	0	0	0	0	303
	34	STO-62	1	8	1	1	0	15173	15167	119	10	7	0	2	0	0	307
	35	STO-62	1	16	1	2	1	32814	32809	256	8	7	0	1	0	0	309
	36	STO-62	1	32	1	3	1	74211	74204	580	13	11	0	1	0	0	308
	37	STO-62	1	64	1	4	1	147179	147172	1150	15	13	0	4	0	0	306
	38	STO-62	1	128	1	5	2	247656	247649	1935	21	18	0	3	0	0	317
	39	STO-62	1	256	1	7	3	373460	373454	2918	26	20	0	6	0	0	332
	40	STO-62	1	512	1	9	4	416991	416984	3258	34	26	0	10	0	0	352
	42	STO-62	2	64	1	5	2	149625	149618	1169	15	16	0	2	0	0	308
	43	STO-62	2	128	1	7	2	261644	261637	2044	24	21	0	4	0	0	309
	44	STO-62	2	256	1	8	3	409601	409594	3200	31	24	0	8	0	0	309
	45	STO-62	2	512	1	12	4	614358	614352	4800	45	35	0	11	0	0	324
	46	STO-62	2	1024	1	15	6	713848	713841	5[illegible]77	56	41	0	20	0	0	363

The chart in Figure 10 shows the scalability of the sequential read operations as we increase the number of jobs that were executed in parallel on the x-axis. It also shows the measured latency on the y-axis to the right.

Figure 10. Sequential Read Scalability



In the above test, we reached a maximum of 713,848 IOPS. This result matches the result obtained by the Calibrate I/O test of 799,400 IOPS in Figure 4 above.

Conclusion

The above performance tests show that with a two-node Oracle RAC clusters, we were able to achieve a maximum throughput of 7,386MB/sec for 100% sequential write tests, a maximum throughput of 12,462MB/sec for 100% sequential read tests, and 713,848 random IOPS with very low latency of 1.1 msec. These results closely match the results that were obtained using the Calibrate I/O tool provided by Oracle, demonstrating that there were no I/O performance problems generated by the storage subsystem. In addition, the graphs show linear scalability as we added the number of jobs executing in parallel in one node until we reached saturation, and further scalability as we added an additional Oracle RAC node.

The above tests show that Hitachi Unified Compute Platform is the ideal platform to deploy your Oracle databases and mission-critical enterprise applications. It provides a balanced server and storage converged solution with sufficient performance to meet your business requirements. It scales linearly as your workload increases, whether horizontally by adding more blade servers as RAC nodes, or vertically by using an SMP connector to combine the resources of up to four full-widths blades. The solution comes in three sizes: small, medium and large that can be upgraded easily, simply by adding CPUs, blades or a storage array, thus protecting your investments for the future as your business grows.

In addition to deploying Oracle applications, the UCP can be used to deploy Oracle and non-Oracle workloads on Linux, or Microsoft Windows®. It uses an open architecture without a vendor lock-in or proprietary, high-cost components thereby providing a lower TCO, and better ROI on your infrastructure spending.

UCP from Hitachi provides a faster time to value, being a pre-integrated, pre-configured, and pre-tested converged infrastructure that is optimized for Oracle Database deployments. With database and application-consistent data protection as provided by Hitachi Data Instance Director, Hitachi helps to improve application and data availability to meet stringent SLAs, such as RPO, RTO and retention schedules. HDID allows for more frequent protection operations, resulting in less data to recover following a disaster, without impacting your production environments. In addition, the built-in reliability, availability, serviceability and sophisticated failover mechanisms, such as the N+1 cold standby in the server, or the global-active device in the storage, provide end-to-end data protection, high availability and disaster recovery options for your databases and applications. These features make it possible to position Hitachi's UCP as an always available, dependable platform to deploy your mission-critical enterprise applications.

References

1. Hitachi Unified Compute Platform:
https://www.hds.com/en-us/products-solutions/converged-infrastructure.html
2. Hitachi Unified Compute Platform for Oracle Database:
https://www.hds.com/en-us/products-solutions/converged-infrastructure/unified-compute-platform-for-oracle-databases.html
3. Hitachi Compute Blade 2500:
https://www.hds.com/en-us/products-solutions/servers/compute-blade-2500.html
4. Hitachi Virtual Storage Platform:
https://www.hds.com/en-us/products-solutions/storage/virtual-storage-platform-g-series.html
5. Hitachi Data Instance Director:
https://www.hds.com/en-us/products-solutions/data-protection/data-instance-director.html
6. Global Active Device:
https://www.hds.com/en-us/products-solutions/data-protection/global-active-device.html
7. Benchware benchmark:
http://www.benchware.ch/

WHITE PAPER

Corporate Headquarters
2845 Lafayette Street
Santa Clara, CA 95050-2639 USA
www.HDS.com **community.HDS.com**

Regional Contact Information
Americas: 866 374 5822 or info@hds.com
Europe, Middle East and Africa: +44 (0) 1753 618000 or info.emea@hds.com
Asia Pacific: +852 3189 7900 or hds.marketing.apac@hds.com

HITACHI is a trademark or registered trademark of Hitachi, Ltd. VSP, ShadowImage and TrueCopy are trademarks or registered trademarks of Hitachi Data Systems Corporation. Microsoft, Windows and SQL Server are trademarks or registered trademarks of Microsoft Corporation. All other trademarks, service marks, and company names are properties of their respective owners.
WP-555-A R. Bitbar October 2016