



WHITE PAPER

# Using HPE ProLiant Servers and Fusion ioMemory-based HPE PCIe Workload Accelerators from SanDisk®

**SanDisk®**  
a Western Digital brand

Western Digital Technologies, Inc.  
951 SanDisk Drive, Milpitas, CA 95035

[www.SanDisk.com](http://www.SanDisk.com)

## Table of Contents

<b>Executive Summary</b>	<b>3</b>
<b>Architecture</b>	<b>4</b>
<b>Performance</b>	<b>6</b>
<b>Capacity</b>	<b>8</b>
<b>Oracle Licensing Advantage</b>	<b>8</b>
<b>About the Fusion ioMemory Platform</b>	<b>9</b>
<b>About the HPE ProLiant DL380 Gen9</b>	<b>10</b>
<b>About Oracle Data Guard and Active Data Guard</b>	<b>11</b>
<b>Conclusions</b>	<b>14</b>

## Executive Summary

This reference architecture paper shows how to deploy Oracle 12c with extreme performance and high availability, using Fusion ioMemory™-based Hewlett Packard Enterprise PCIe Workload Accelerators from SanDisk, HPE ProLiant DL380 Gen9 Servers, and Oracle Data Guard physical standby database. Performance and system details are provided.

Each of the two low-profile HPE DL380 Gen9 servers in this reference architecture was powered by six internal 6.4 TB Fusion ioMemory flash memory storage devices (HPE PCIe Workload Accelerators), for a total usable capacity of 38.4 TB per server. System performance, as measured by the Oracle CALIBRATE\_IO utility, exceeded 1 million database IOPS with sub-millisecond latency in support of On-Line Transaction Processing (OLTP) workloads. Performance also hit 16 GB/s of streaming bandwidth critical to Decision Support Systems (DSS), Data Warehouses, and On-Line Analytic Processing (OLAP) systems. These metrics are per-server, and the reference model has two identically configured servers. By using Active Data Guard we were able to run the DSS workload on both servers, thus doubling the number of active users, and the total throughput hit 32 GB/s. This scalable model supports up to 31 servers per Data Guard cluster.

Oracle Data Guard 12c provides database high availability by creating one or more physical standby copies of the database, and maintaining all copies through the user's choice of synchronous or asynchronous redo apply. Data Guard supports graceful "switchovers" for planned events such as server maintenance, in which case the primary and a standby database simply switch roles. Data Guard also supports "failovers" for unplanned events such as a power outage or catastrophe, in which case the primary is presumed lost and one of the standbys is promoted to the role of primary. Failovers can be automated using Oracle Data Guard's Fast-Start Failover feature, which relies on an observer process to detect the unplanned outage of the primary database and signal the broker to transfer the primary role to one of the available physical standby databases.

Switchovers are planned events with graceful shutdowns and startups, so no recovery is required and the total transition period is quite brief. Failovers are unplanned and often require database recovery. Database recovery on legacy disk storage can take a very long time, but is significantly shortened by storing the database and recovery files on SanDisk flash storage. For example, at Oracle OpenWorld we conducted a live failover demonstration of a multi-terabyte database on Fusion ioMemory storage that was under a heavy workload. The entire process – including Data Guard Fast-Start Failover, database recovery, and transaction restarts – took less than 45 seconds. During that time, no users were disconnected, and no data was lost. After this brief "pause" all transactions were restarted automatically by Oracle Database.

Our reference architecture also leverages the Active Data Guard option to allow read-only transactions such as queries and backups to be offloaded from the primary database to a standby database. (See *About Oracle Data Guard and Active Data Guard* later in this paper for a discussion on the differences between the two types of Data Guard.) Standby databases are usually run in a constant state of recovery and are not usable by applications. Active Data Guard allows real-time redo apply, so each standby can be opened for read-only operations while continuing to apply redo and maintain synchronization with the primary.

To highlight the benefits of Active Data Guard, we used HammerDB to simultaneously run a Decision Support System workload against the primary and standby databases. During this test each database delivered up to 16 GB/s of data from the in-server storage layer to the database engine, with an average combined data throughput of 30 GB/s. These results show linear scaling of queries: enabling Active Data

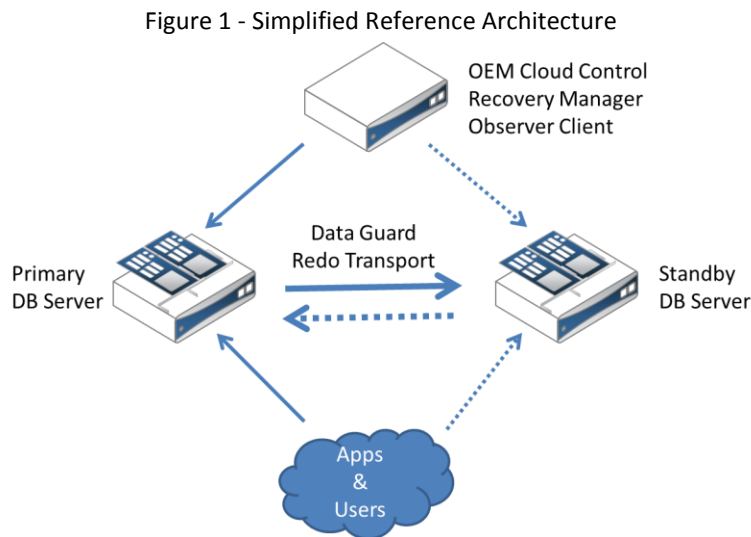
Guard on the standby database doubled the number of users and queries (including full table scans) completed per unit of time.

It is important to note that an Active Data Guard cluster is not limited to read-only operations; the primary database is fully operational. To illustrate this point we ran HammerDB's OLTP workload against the primary while simultaneously running its DSS workload against the standby. Results were outstanding: the primary maintained a very high transactional throughput, while the standby ran full table scans at speeds up to 16 GB/s.

Oracle Enterprise Manager Cloud Control 12c Release 5 was used to create the Data Guard standby database, and to monitor and maintain the entire Data Guard environment thereafter. Cloud Control provides a graphical user interface, and translates clicks into instructions that are passed to the Data Guard Broker. The same instructions could be entered directly into the Broker's command line interface when Cloud Control is not available.

## Architecture

We implemented a typical Oracle 12c Data Guard environment consisting of a primary database server, a standby database server, and an administrative server as shown in Figure 1. Performance and reliability were significantly enhanced by replacing legacy hard disk storage with SanDisk flash storage. SanDisk offers a wide range of flash storage products, including in-server and attached. In this reference architecture we used HPE PCIe Workload Accelerators from SanDisk (Fusion ioMemory PCIe Application Accelerator model SX350-6400).



Each database server in our reference architecture is an HPE ProLiant DL380 Gen9 dual-socket server with the following characteristics:

- A 2U rack-mount for maximum performance in a very small footprint
- Two Intel Xeon E5-2697 v3 processors with 14 cores each, TURBO enabled, and Hyper-Threading enabled

- From an Oracle licensing perspective, we used 56 processor cores throughout our cluster with a 0.5 core factor, for a total license of 28 cores.
- 256 GB DRAM
- For OLTP testing, 200 GB was allocated to the Oracle SGA and 8 GB for the PGA
- For DSS and analytics, 128 GB was allocated to the Oracle SGA
- All unallocated memory was shared by the operating system, drivers, and utilities
- 38.4 TB of 100% flash storage, using six 6.4 TB HPE PCIe Workload Accelerators (Fusion ioMemory PCIe Application Accelerators model SX350-6400)
- A single 10 Gb Ethernet port (on-board)
- Red Hat Enterprise Linux 7.1
- Oracle Database Server 12.1.0.2 Enterprise Edition with Grid Infrastructure (ASM and Oracle Restart)
- Three Oracle ASM diskgroups: GRID (60 GB), DATA (19170 GB), and RECO (19170 GB)
- HammerDB version 2.18 to populate the database with test data and drive a workload

The administrative server shown at the top of Figure 1 runs the following additional software components:

- Red Hat Enterprise Linux 6.5
- Oracle Database Server 12.1.0.2 Enterprise Edition
- Oracle Enterprise Manager (OEM) Cloud Control 12c
- A repository database for Oracle Management Server (OMS), the heart of OEM Cloud Control 12c
- A repository database for Oracle Recovery Manager (RMAN)
- Oracle Observer client application used for Data Guard Fast-Start Failover

The administrative server may be implemented as a new/dedicated server, or you may use an existing data center resource. It can even be implemented as a virtual machine. In our case, we leveraged an administrative server previously deployed in our data center that provides OEM Cloud Control and RMAN services to all of our Oracle databases. The only modification was to enable the Observer software, which is included with the Oracle 12c Enterprise Edition software and was therefore already installed on the server.

Data Guard offers three protection modes; we selected Maximum Availability. Please note that Data Guard standby databases are created in Maximum Performance mode by default. The three protection modes are briefly described here:

- **Maximum Performance** is the default and always operates in asynchronous mode. This mode allows users on the primary database to run without waiting for redo to be applied on a standby database.
- **Maximum Availability** guarantees zero data loss except in the case of a double failure. It provides a balance between the other two protection modes. It always attempts to operate in synchronous mode, but it will automatically switch to asynchronous mode when the standby is offline and automatically switch back to synchronous mode once the standby is back online.

- **Maximum Protection** guarantees zero data loss and always operates in synchronous mode. The primary is automatically shut down if Data Guard cannot write to a standby redo log of a standby database, which means two or more standby databases are generally required to use Maximum Protection mode. (If the environment had only one primary and one standby, and the standby was shut down to apply operating system patches, then synchronous writes would not be possible and Data Guard would automatically shut down the primary database).

The network link between the primary and standby database should be wide enough for the amount of redo being generated on the primary database. A high-performing Oracle database on SanDisk flash storage might generate between 200 and 300 MB per second of redo (not counting multiplexed files) to be transported across the public interface to the standby. This is in addition to routine network traffic. Therefore, it is recommended to use a network link of at least 10 gigabits per second. Redundant links are recommended for high availability.

When Data Guard is implemented “all at once” the three servers (primary, standby, and administrative) can be deployed and configured in parallel. Alternatively, existing Oracle 12c database environments can be upgraded to Data Guard by enabling the minimum requirements within an existing database and using OEM Cloud Control to create a physical standby copy of that database on another server.

Not shown in Figure 1 is a tape library system (or other media manager) for long-term storage of RMAN backups. It is recommended that RMAN backups be made to a local cache, and the local cache be backed up to a central location that is accessible to all standby databases. In the event of a catastrophic failure where the entire primary server is lost, including the RMAN cache, the standby that is promoted to the role of primary might require recovery using RMAN files in a central library.

## Performance

Performance was measured using four workload generators: the Flexible IO Tester, Oracle CALIBRATE\_IO, HammerDB On-Line Transaction Processing, and HammerDB Decision Support System.

Prior to installing the Oracle software we measured the performance of each storage device using the Flexible IO Tester (“fio”) by Jens Axboe. This test simply validates that the storage is operating within vendor specifications, but it makes no prediction as to database performance.

After installing the Oracle software and creating our primary database we re-evaluated the performance of our storage layer using Oracle CALIBRATE\_IO. This Oracle-supplied procedure runs inside the database. CALIBRATE\_IO was run nine times and the results averaged. Our primary server generated in excess of 1 million database IOPS using a typical 8 KB block size with sub-millisecond latency to support On-Line Transaction Processing (OLTP) workloads, and sustained 14 GB/s of streaming bandwidth critical to Decision Support Systems (DSS), Data Warehouses, and On-Line Analytic Processing (OLAP) systems.

The CALIBRATE\_IO tests were repeated after creating the Data Guard standby database. Running Oracle CALIBRATE\_IO on both the primary and standby databases simultaneously yielded double the results compared to a single server: we exceeded 2 million 8K IOPS with sub-millisecond latency, and 28 GB/s of large I/O bandwidth. Each server performed equally well and results scaled linearly. The reference architecture has two identically configured servers. Oracle Data Guard 12c supports up to 31 servers per Data Guard configuration, with one as the primary and 30 as standby. Ten can use synchronous replication,



and the remainder use asynchronous replication. In a reader farm configuration, the maximum theoretical performance would be 31 million IOPS and 434 GB/s bandwidth.

Real-world SQL performance was estimated using the HammerDB benchmark suite, which features data and workload generators for OLTP and DSS. We did not perform a true system benchmark: we simply used HammerDB to place the system under a heavy load and compare the performance of various configurations. The results are summarized in Figure 2. The unit of measure is New Order Transactions Per Minute (NOPM).

Figure 2 illustrates the relative performance of our Oracle database in various configurations. Phases 1, 2, and 3 are pre-Data Guard. Phases 4 & 5 are Data Guard configurations. The five phases are described below the chart. The Data Guard configuration with synchronous redo (Phase 5) was used to set the baseline at 1.000, because it is the most demanding of all configurations. The other configurations were compared to it, with their relative performance shown in the chart. Phase 3 perhaps best represents Oracle customers with standalone databases. Phase 4 shows that Data Guard offers higher availability in exchange for a nominal performance hit. Phase 5 implements synchronous redo transport, which as described earlier provides complete fault tolerance.

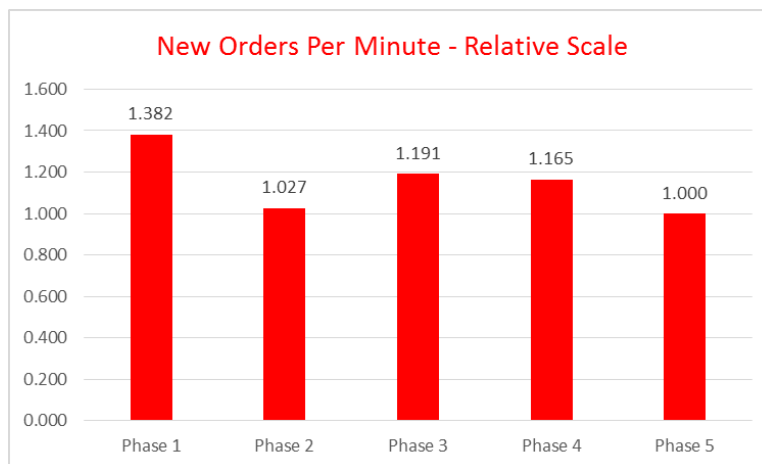


Figure 2 – HammerDB Average Score Per Configuration

- Phase 1 is a standalone Oracle 12c database in ARCHIVELOG mode.
- Phase 2 adds Force Logging, Flashback Logging, Block Change Tracking, and Broker.
- Phase 3 adds two hidden `init.ora` parameters to account for the heavy logging added in Phase 2.
- Phase 4 adds Data Guard in the default Maximum Performance Mode (asynchronous).
- Phase 5 changes Data Guard to Maximum Availability (synchronous).

Notice in Figure 2 that the New Orders Per Minute score did not increase when using Data Guard, despite the fact we doubled the number of servers. This is because write workloads cannot run on a physical standby database. We therefore instrumented HammerDB's Decision Support System (DSS) benchmark on the standby. The DSS test can execute purely read-only operations. While the OLTP benchmark was running on the primary, the DSS benchmark was running in parallel on the standby. We observed no degradation in the primary's performance while the standby was under heavy load. This shows the standby may be used to offload read-only operations, scale the number of users, and improve total system throughput without impacting users on the primary database.

DSS workload performance was also assessed in all five of the configurations noted above. The ability to perform long sequential reads was unaffected by each configuration. In all cases, each database was able to run DSS queries at speeds up to 16 GB/s. By running the DSS workload on both the primary and standby in parallel we observed sustained aggregate throughput of 28 GB/s.

Data Guard allows up to 31 databases per cluster. If results were to continue scaling linearly, we could expect aggregate throughput in the range of 434 to 496 GB/s.

## Capacity

Each of our Oracle Data Guard servers has a raw, usable 100% flash capacity of 38.4 TB, uncompressed. The effective storage space could be tripled by leveraging the Oracle Advanced Compression Option, which features row compression and LOB de-duplication. The Oracle Advanced Compression Option was not used in this paper.

Although the reference architecture has two identical servers, the usable capacity remains 38.4 TB, as a physical standby's storage is essentially a mirror of the primary's storage.

Oracle ASM with Normal Redundancy (RAID 0) was configured on each server to maximize performance and usable capacity. Oracle Data Guard synchronous replication effectively maintained RAID 0+1 protection of all data across the two servers.

The server used in this reference architecture has a small footprint, just 2 rack units high, and has a total of six PCIe slots for HPE PCIe Workload Accelerators (Fusion ioMemory PCIe Application Accelerators). Fusion ioMemory storage devices are available from HPE as PCIe Workload Accelerators in a range of sizes from 1.25 TB to 6.4 TB per device and compatible with all HPE Gen8 and Gen9 ProLiant servers. There are also HPE PCIe Workload Accelerators available in 1.2 TB and 1.6 TB capacities as mezzanine cards for use with HPE BladeSystem Gen8 and Gen9 server blades.

## Oracle Licensing Advantage

Core-based licensing is very common with Oracle Database 12c. Oracle customers are very sensitive to the number of server cores, their utilization, and their impact on the total system cost. Our reference architecture uses a relatively small number of processor cores, using them wisely to control costs. Rather than implementing a larger server with more cores (and correspondingly higher licensing costs) that mostly spin while waiting for data to be delivered from slow legacy storage, our reference architecture uses a different, more modern approach. It takes advantage of SanDisk flash storage connected directly to the system board, which provides data to the processors at memory speeds. This changes processor cycles from spin waits to data processing events, allowing Oracle to fully utilize all resources. By fully leveraging the existing server processors, additional processors and additional servers were unnecessary, so cost was contained without sacrificing performance.

When using Data Guard, Oracle requires Oracle Database 12c Enterprise Edition be licensed on both the primary and standby servers. If you implement the optional Real-Time Apply component of the reference architecture, which allows off-loading read-only operations to the standby server, then Active Data Guard must also be licensed on both servers. ASM and Data Guard are both included with Enterprise Edition and require no additional licenses. There are no additional licensing fees associated with the Oracle Enterprise



Manager (OEM) Cloud Control, Oracle Management Server (OMS) repository, Recovery Manager (RMAN) software, the RMAN repository, the Data Guard Broker, or the Fast-Start Failover Observer.

To summarize, this reference architecture requires 56 core licenses of Oracle Database Server 12c Enterprise Edition (28 per server), and 56 core licenses of Oracle Active Guard (again, 28 per server). Our Intel processors have an Oracle licensing core factor of 0.5, thus reducing all core licenses by 50%. The total cost of our Data Guard environment represents a significant savings compared to an Oracle RAC deployment with identical hardware, and RAC does not include database redundancy. Savings have two primary reasons: Active Data Guard costs less than half of Oracle RAC, and a flash-based system requires fewer processor cores to achieve equal or better performance compared to a typical Oracle RAC database stored on legacy disk arrays.

Several Oracle software add-ons are available to enhance the reference architecture including the Oracle Partitioning Option, Advanced Security Option (i.e., encryption), and the Advanced Compression Option (ACO). ACO has many benefits in a Data Guard environment, including compression of RMAN backups, Data Pump Exports, and redo transport.

Our reference architecture also leveraged the Oracle Enterprise Manager Cloud Control 12c R5 to monitor and maintain the overall environment, including the creation of standby databases. The basic deployment and use of OEM Cloud Control does not incur Oracle license fees, as noted in Chapter 10 of the Enterprise Manager Licensing Information guide. Cloud Control is most effective when the optional Diagnostic and Tuning packs are licensed separately.

This flash-enabled, shared-nothing architecture is highly scalable and benefits users with a very efficient platform delivering dramatically higher levels of work and performance per server while also reducing total core-based license costs.

Licensing information for Enterprise Edition including databases and ASM can be found here: <http://docs.oracle.com/database/121/DBLIC/toc.htm>

Enterprise Manager licensing information can be found here: [http://docs.oracle.com/cd/E24628\\_01/doc.121/e24474/toc.htm](http://docs.oracle.com/cd/E24628_01/doc.121/e24474/toc.htm)

## About the Fusion ioMemory Platform

HPE PCIe Workload Accelerators are based on the Fusion ioMemory platform from SanDisk. It combines Fusion ioMemory VSL™ (Virtual Storage Layer) software with enterprise-grade Fusion ioMemory hardware to take enterprise applications and databases to the next level. The Fusion ioMemory platform provides consistent microsecond data access latency for mixed workloads, multiple gigabytes per second access, and hundreds of thousands of IOPS from a single product. With industry leading reliability (e.g., unrecoverable bit error ratio of  $10^{-20}$ ), the sophisticated Fusion ioMemory architecture allows for nearly symmetrical read and write performance with best-in-class low queue depth performance, making the Fusion ioMemory platform ideal across a wide variety of real world, high-performance enterprise environments.

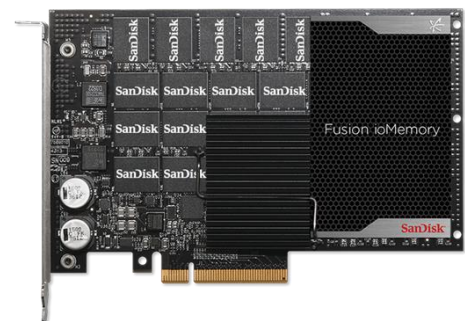


Figure 3 – HPE 6.4 TB PCIe Workload Accelerator (Fusion ioMemory SX350-6400)

This paper features the HPE 6.4 TB PCIe Workload Accelerator (Fusion ioMemory SX350-6400) storage device, with 6.4 TB raw usable capacity of SanDisk NAND flash memory. The Fusion ioMemory SX350 line is available in a standard PCIe form factor for HPE Gen8 and Gen9 ProLiant rack-mount servers, such as the DL380, with per-device capacities ranging from 1.25 TB to 6.4 TB. The Fusion ioMemory SX350-6400's endurance rating is 22 petabytes written. These are also available in a mezzanine form factor for use with HPE BladeSystem Gen8 and Gen9 blade servers in 1.2 TB and 1.6TB capacities.)

The Fusion ioMemory SX350-6400 storage device uses a PCIe 2.0 x8 slot, making it compatible with nearly all enterprise-class servers. Most servers support multiple devices, somewhere between two and twelve depending on the server. All Fusion ioMemory storage devices keep data center costs to a minimum by consuming than 25 watts of power or less.

Fusion ioMemory PCIe Application Accelerators are unique in their ability to sustain writes as well as or better than reads. Most Oracle databases perform more read operations than write operations, but writes can still be a bottleneck for Oracle databases. Consider that many OLTP and Operational Data Stores have a workload consisting of 40% writes. These writes include inserts, updates, and deletes to row data and indexes, and corresponding Undo, Redo, and Archive Log writes generated by these operations. Even Decision Support Systems (DSS) and databases used for On-Line Analytic Processing (OLAP) may experience slowness on checkpoints and logging. When selecting a storage product it is imperative to consider the database's dependency on write operations, not just read acceleration.

To learn more about HPE PCIe Workload Accelerators, please refer to the HPE website: <http://www8.hp.com/us/en/products/accelerators/product-detail.html?oid=8699245>. You can also visit the SanDisk website to learn more about the Fusion ioMemory SX350 line of PCIe Application Accelerators featured in this paper by referring to the video at <https://www.youtube.com/watch?t=50&v=qweog75HTL8> and the data sheet at <http://www.sandisk.com/assets/data-sheets/fusion-iomemory-sx350-pcie-application-accelerators-datasheet.pdf>.

## About the HPE ProLiant DL380 Gen9

The HPE ProLiant DL380 Gen9 Server delivers the latest performance and expandability in HPE's 2-Processor portfolio of rack mount servers. It is purpose-built for flexibility, efficiency, and manageability and designed to adapt to the needs of any environment, from large enterprise to remote office/branch office. The DL380 Gen9 Server offers enhanced reliability, serviceability, and continuous availability, backed by comprehensive warranty.

Figure 4 – HPE ProLiant DL380 Gen9 Server



Designed to reduce costs and complexity, the DL380 Gen9 Server leverages Intel's latest E5-2600 v3 processors with up to 70%<sup>1</sup> performance gain, plus the latest HPE DDR4 SmartMemory supporting 1.5 TB and up to 14%<sup>2</sup> performance increase.

Manage the DL380 Gen9 Server in any IT environment by automating the most essential server lifecycle management tasks: deploy, update, monitor, and maintain with ease.

The HPE ProLiant DL380 Gen9 Server can run every application from the most basic to mission critical, and can be deployed with confidence. With the HPE ProLiant DL380 Gen9 Server, you can deploy a single platform to handle a wide variety of enterprise workloads:

- Big Data—Manage exponential growth in your data volumes—structured, unstructured, and semi-structured
- Storage-centric applications—Remove bottlenecks and improve performance
- Data analytics—Find the information you need, when you need it, to enable better business decisions
- Virtualization—Consolidate your server footprint by running multiple virtual machines on a single DL380
- Customer relationship management (CRM)—Gain a 360-degree view of your data to improve customer satisfaction and loyalty
- Enterprise resource planning (ERP)—Trust the DL380 Gen9 to help you run your business in near real time
- Virtual desktop infrastructure (VDI)—Deploy remote desktop services to provide your workers with the flexibility they need to work anywhere, at any time, using almost any device

To support your heterogeneous IT environment, the DL380 Gen9 Server supports Microsoft Windows® and Linux® operating systems, as well as VMware® and Citrix® virtualization environments.

## About Oracle Data Guard and Active Data Guard

This section briefly describes Data Guard as it pertains to our reference architecture. For a complete understanding of all Data Guard capabilities and configuration options, refer to Oracle's on-line documentation located at [http://docs.oracle.com/database/121/nav/portal\\_booklist.htm](http://docs.oracle.com/database/121/nav/portal_booklist.htm)

Data Guard is a free, high availability feature of the Oracle Database Server 12c Enterprise Edition that enables customers to instantiate multiple copies of a database. One copy of the database supports live interaction with users and is called the "primary database". All other copies of the same database are called "standby databases" because they are standing by to assume the role of primary should the need arise. Changes made in the primary database are forwarded from the primary database to the standby databases.

<sup>1</sup> Source as of 3 April 2014: Intel internal measurements on platform with two E5-2697 v2 (12C, 2.7GHz), 8x8GB DDR3-1866, RHEL6.3. Platform with two E5-2697 v3 (14C, 2.6GHz, 145W), 8x8GB DDR4-2133, RHEL6.3.

<sup>2</sup> Based on similar capacity DIMM comparing HPE server vs. non HPE server with DDR4, July 2014.

The method by which changes are sent differs for physical and logical standby databases. In this reference architecture we used a physical standby database, which means all changes made to the primary database were sent to the standby in the form of redo entries. This will be illustrated later.

The illustration below provides a high-level overview of an Oracle Data Guard environment.

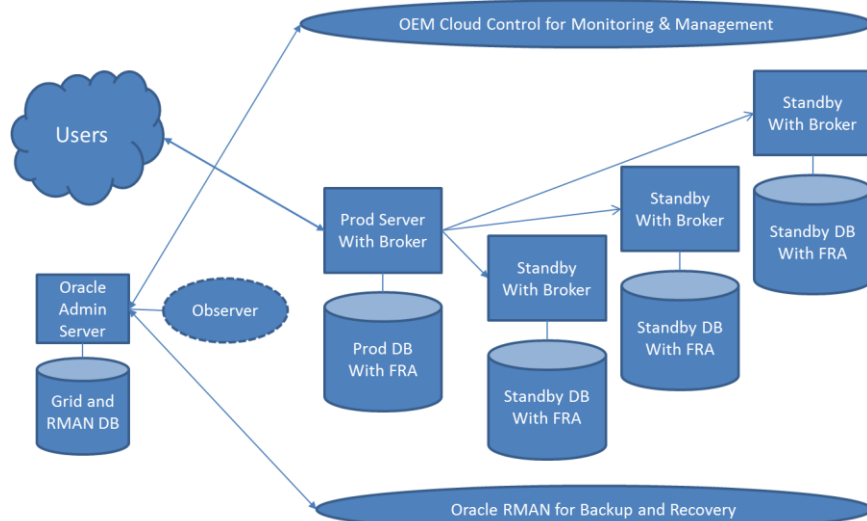


Figure 5 - Oracle 12c Data Guard Big Picture

This illustration shows users interacting with the primary database, and changes being forwarded to one or more standby databases. Each copy of the database has its own storage, which can be in-server or centrally managed. An administrative server (at the left) runs three optional components: the Oracle Enterprise Manager Cloud Control software to monitor and manage the entire environment; the Oracle Recovery Manager repository; and the Oracle Observer client software used by the Data Guard Fast-Start Failover capability.

Notice in the illustration that each database has a Broker. The internal interface to the Broker is an Oracle background process named DMON (Data Guard Monitor). DMON is included with all Oracle databases based on the Enterprise Edition. Once the Broker has been enabled, DMON coordinates instructions between the primary and all standby databases across a standard Oracle Net connection. The Broker is required by the DGMGRL command line utility, the Observer, and by OEM Cloud Control when used to manage the Data Guard environment. Commands from OEM Cloud Control, DGMGRL, and the Observer are automatically passed through the Broker for dissemination throughout the cluster across the standard Oracle Net topology.

Data Guard supports high availability through graceful “switchovers” for planned outages and “failovers” for unplanned outages. In both cases, the role of primary is transferred to the first available standby database.

- **Switchovers** are manual events that are planned and graceful. A switchover transfers the role of primary to one of the standby databases and then converts the old primary into a standby, so that the two database copies have literally switched roles. For example, prior to patching the operating system on the

primary server, a switchover should be initiated to transfer the role of primary to another database. Once the original primary's role has been downgraded to standby, it can be shut down so that the host can be upgraded and the Oracle binaries relinked, with zero impact to the production database. Because switchovers are graceful, no recovery is required, and there is zero chance of data loss. Optionally, a second switchover event transfers the role of primary back to the original database. Assuming each copy of the database runs on equally powerful hardware, it does not matter where the primary runs.

- **Failovers** occur when there is a problem with the primary database or its host that prevents normal operations. Failovers can be manual or automated events, but either way they are not planned and are not graceful. Automated database recovery is typically required after a failover event. Many Oracle Data Guard users prefer manual failover, because if the issue is minor it can be resolved by the database administrator and/or system administrator without failing over to a standby and recovering the database. However, issues often arise in the middle of the night when no administrators are on hand, in which case a fully automated Fast-Start Failover is preferred. A failover requires the old primary be rebuilt or restored from a backup. To avoid the post-failover case of running a primary with no available standby it is recommended to have at least three copies of the database.

Data Guard offers three protection modes (discussed earlier in this paper):

- Maximum performance – the default, always operating in asynchronous mode
- Maximum protection – guarantees zero data loss and always operates in synchronous mode
- Maximum availability – provides a balance between the two by operating in synchronous mode when possible and automatically switching between synchronous and asynchronous modes when performance objects are not being met. Standby databases are created in maximum performance mode, but it is generally recommended to use maximum availability mode to minimize any risk of data loss.

The flow of data from primary to standby is shown in the figure below. Data entered by users into the primary database is always recorded in the redo buffer. If Data Guard is configured to use synchronous transport, then the primary database's Log Writer process will duplex the redo data to both the redo log file and the LNS process. In asynchronous transport mode the Log Writer is only used by the local database and is not used to ship redo. The LNS process will try to read the data from the redo buffer, but if it misses the data then LNS will retry to read it from the online redo log file. Furthermore, in asynchronous mode, if the redo log is archived before LNS reads the data, then the Archiver process will perform "gap resolution" by reading from the archived log file and sending data to the remote server's RFS process.

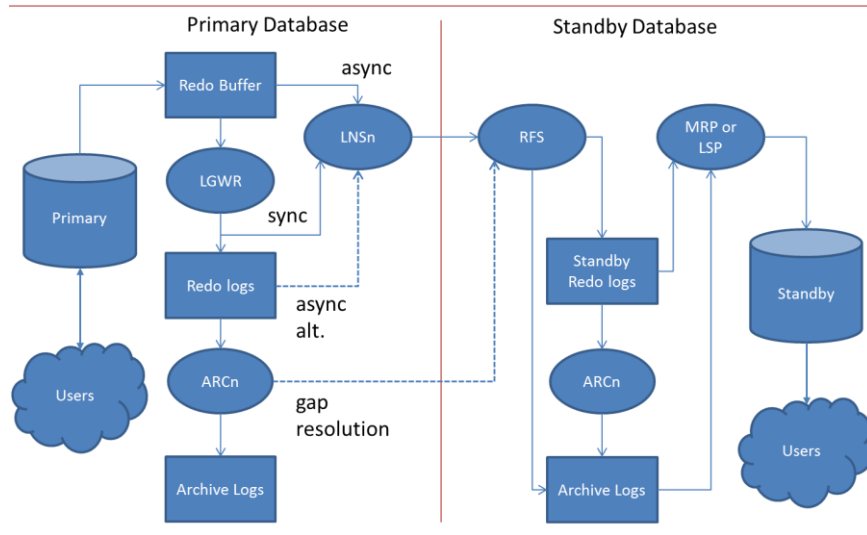


Figure 6 - Oracle Data Guard Redo Flow

Active Data Guard is a separately licensed product that can be used to enhance Data Guard. Active Data Guard allows standby databases to be used actively for read-only operations while still applying redo: changes made to the primary will continue to be applied to the standby while open for read-only operations. This allows the standby to be used for queries, reports, and backups. Without Active Data Guard, redo cannot be applied while the standby is being used: the backlog of redo can quickly fill the system and shutdown Oracle, so standby databases without Active Data Guard should only be opened read-only for very short periods of time. With Active Data Guard, the standby can be opened as read-only indefinitely while still applying changes.

## Conclusion

Today, there is a cost-effective solution to optimize your Oracle Database 12c Data Guard infrastructure with HPE ProLiant Gen9 servers powered by HPE Workload Accelerators from SanDisk - providing the perfect balance of performance, capacity, and high availability. Eliminating I/O wait times by storing and running the database on high-performance flash allows a smaller number of system processors to be used much more efficiently, while also driving down software licensing fees and operating expenses, as well as minimizing the data center footprint.

For more information about the products described in this reference architecture, contact your SanDisk or HPE sales representatives or visit us online at <http://hpe.sandisk.com/>.

Specifications are subject to change. ©2016 Western Digital Corporation or its affiliates. All rights reserved. SanDisk and the SanDisk logo are trademarks of Western Digital Corporation or its affiliates, registered in the U.S. and other countries. Fusion ioMemory is a trademark of Western Digital Corporation or its affiliates. Other brand names mentioned herein are for identification purposes only and may be the trademarks of their respective holder(s). 5090EN 20160615

Western Digital Technologies, Inc. is the seller of record and licensee in the Americas of SanDisk® products.