White Paper

EMC COST-EFFICIENT INFRASTRUCTURE FOR ORACLE

EMC VNX5300, EMC FAST Suite, EMC VNX Snapshots, VMware vSphere, Oracle Standard Edition RAC

- Efficient–Fully virtualized Oracle environment using VMware vSphere 5
- Optimal–EMC FAST Suite with Flash, SAS, and NL_SAS drives
- Scalable–EMC VNX Snapshots enabling storage scalability
- Agile–Enhanced high availability with VMware HA and Oracle RAC

EMC Solutions Group

Abstract

This white paper describes configuration and deployment strategies for running increasingly larger virtualized Oracle OLTP workloads in an EMC® VNX®5300 block-based storage array infrastructure, creating a cost-effective platform for Oracle using EMC FAST Suite. It illustrates the latest EMC VNX technology, VNX Snapshots, to facilitate the simple and efficient provisioning and protection of Oracle databases.

October 2012

Part Number H10945.1

Table of contents

Executive summary

Business case

Small and midsize Oracle customers need a cost-efficient Oracle solution that balances the three characteristics of Oracle infrastructures:

- Agile–Ability to do tasks quickly and easily
- Efficient–Ability to accomplish more tasks with less or just enough resources and time
- Scalable–Ability to add increasing throughput and additional Oracle users on the existing infrastructure

This EMC solution describes the following EMC, VMware, and Oracle technologies that enable the needs of the small to midsize Oracle customers:

- **Efficient**–VMware® vSphere™ 5.0, including features such as VMware High Availability (HA), enables this solution to consolidate and run increasingly larger workloads to meet both availability and increased workloads on the VMware ESXi™ cluster.
- **Optimal**–EMC® FAST Suite enables optimization of the storage environment by automatically moving hot data slices from busy LUNs to a faster storage medium (for example, Flash drives) while keeping cold data on the slower drives (for example, Near-line Serial Attached SCSI (NL_SAS) drives). With FAST Suite, Oracle database administrators no longer need to worry about Oracle-to-storage tuning. It also avoids performance and operational issues as customers increase the number of Oracle databases in their environment.
- **Scalable**–EMC VNX® Snapshots technology enables storage scalability by quickly provisioning test development (test/dev) environments, quality assurance (QA) environments, or other non-production environments through using virtual copies of the production database.

In addition, the solution illustrates how to save money on Oracle licensing in a virtualized environment by deploying multiple virtual Oracle Standard Edition (SE) real application clusters (RAC) database servers with the current Oracle license. This enables the efficient use of owned licenses to the deployed use-case infrastructure.

Solution overview

The purpose of this solution is to illustrate how EMC, VMware, and Oracle technologies can simplify the daily tasks in an integrated and cost-efficient deployment.

The solution demonstrates how the following technologies create this cost-efficient solution:

- EMC VNX5300 arrays with VNX OE block release 5.32 support new features, such as VNX Snapshots, and provide an enterprise-class storage platform for the solution.
- FAST Suite, one of EMC's storage performance optimization features, combined with Flash, Serial Attached SCSI (SAS), and NL-SAS drives deployed on the VNX5300 array, reduces the cost and maximizes the performance.

- VMware virtualization technology, including VMware HA feature, demonstrates the ease of use and low cost features with which multiple Oracle RAC databases can be managed on a single SAN array.
- Multiple RAC databases, which utilize Oracle SE rather than Oracle Enterprise Edition (EE) to minimize Oracle software licensing cost, are used to illustrate a cost-efficient Oracle solution.

Key results

This solution demonstrates a cost-efficient deployment of the following:

- Hardware (server/network)
- Storage (array use)
- Oracle software (Oracle licenses)

Cost-efficiency is always a balance of cost to performance and the following are the technologies that produced an efficient solution.

EMC FAST Suite and EMC VNX5300 array storage tiers

EMC storage optimization technology, with EMC VNX5300 array storage tiers (SAS and NL_SAS drives) and EMC FAST Cache technology with Enterprise Flash drives, enables performance that scales as the workload increases.

Table 1 illustrates EMC FAST Suite and EMC VNX5300 array storage tiers, enabling the workload to increase, indicated in transactions per minute (TPM).

Table 1. Workload performance comparison

Items	Use case 1: One virtual RAC database	Use case 2: Two virtual RAC databases
Total TPM of workload	19,999 (TPM baseline)	37,298 (86 percent increase in TPM)

VNX Snapshots

VNX Snapshots enables the rapid provisioning of Oracle databases for test, development, patch, and other Oracle database needs. VNX Snapshots accelerates the fast recovery of databases (test/dev) to correct a database logical corruption. These are new snapshot features in VNX OE block release 5.32.

VMware vSphere 5.0

VMware vSphere 5.0 enables the efficient use of customers' hardware infrastructure by consolidating many legacy physical servers to many VMware virtual machines. VMware's HA feature enables the improved utilization of resources by running multiple SE Oracle RAC databases with the most efficient use of Oracle software licensing costs, enhancing the availability of Oracle RAC databases in the event of a physical ESXi server loss, with minimal downtime and performance impact.

For a detailed review on the cost-efficient use of Oracle and VMware, refer to the white paper: *Understanding Oracle Certification, Support and Licensing for VMware Environments*.

Introduction

Purpose

This white paper describes a cost-efficient solution built on EMC VNX5300 storage arrays. In this white paper, technologies such as EMC Fully Automated Storage Tiering for Virtual Pools (FAST VP), FAST Cache, VNX Snapshots, VMware HA, and Oracle Standard Edition RAC are deployed in a data center, illustrating the performance, efficiency, and agility of the Oracle OLTP use-case scenarios.

Scope

The scope of the white paper is to:

- Introduce the key technologies
- Describe the solution architecture and design
- Demonstrate how the key components are configured
- Show the configuration and performance of EMC FAST VP and FAST Cache technology
- Present cost-efficiency messages related to multiple Oracle RAC databases in a virtualized VMware environment
- Demonstrate the rapid provisioning and easy protection of Oracle testing and development systems with VNX Snapshots technology
- Show enhanced high availability with the combination of VMware HA and Oracle RAC

Audience

This white paper is intended for Oracle database administrators (DBAs), storage administrators, IT architects, and technical managers responsible for designing, creating, and managing mission-critical Oracle OLTP application, and especially for those who are looking for a cost-efficient storage solution for production or nonproduction environments.

Terminology

This white paper includes the items in Table 2.

Table 2. Terminology

Term	Definition
ASM	Automated Storage Management
CRS	Oracle Cluster Ready Service(CRS)
FAST™	Fully Automated Storage Tiering
FRA	Fast Recovery Area
NL_SAS	Near-line Serial Attached SCSI
OCR	Oracle Cluster Registry
OLTP	Online transaction processing
RAC	Real application cluster
RMAN	Oracle Recovery Manager
SAN	Storage area network
SAS	Serial Attached SCSI
SMP	Snapshot Mount Point
SGA	System Global Area
VMFS	VMware Virtual Machine File System
VMware HA	VMware High Availability
VNX OE	EMC VNX Operating Environment
VSI	EMC Virtual Storage Integrator

Technology overview

Introduction to the key components

This section provides an overview of key technologies used in this solution:

- EMC FAST Suite
- EMC PowerPath®/VE
- EMC Virtual Storage Integrator for VMware vSphere
- EMC VNX Snapshots
- EMC VNX5300 storage array
- Oracle Database 11*g* Release 2 SE
- VMware vSphere 5

EMC FAST Suite

EMC FAST Suite provides advanced data efficiency using:

- FAST Cache
- FAST VP

Both technologies reduce the number of drives needed to meet a given performance requirement, thereby significantly reducing the total cost of ownership (TCO) of the storage system and achieving the desired performance level.

FAST Cache

FAST Cache uses Flash drives to add an additional layer of cache between the dynamic random-access memory (DRAM) cache and rotating drives, thereby creating a faster medium to store frequently accessed data. It is an extendable, read/write cache. FAST Cache boosts application performance by ensuring that the most active data is served from high-performing Flash drives and can reside on this faster medium for as long as is needed.

FAST Cache tracks the data activity level at a 64 KB granularity and copies hot data to a collection of designated Flash drives. After the data has been promoted to FAST Cache, Flash drives handle subsequent access to the data, thereby providing extremely low latency for that data. As the temperature of the copied data decreases over time, the data is removed from FAST Cache, making room for new hot data.

FAST VP

EMC FAST VP is a policy-based, auto-tiering solution for enterprise applications. With FAST VP, you can create blended storage pools composed of various disk types, such as Flash, SAS, and NL_SAS. In a highly consolidated virtualized environment, this leads to the highest storage efficiency, both from a performance and capacity perspective, because all the drive spindles can be shared efficiently between applications.

FAST VP tracks the sub-LUN data activity level at a 1 GB granularity. In turn, these data slices are automatically migrated to the appropriate tiers within a pool, depending on the data activity level. The data migration is completely application-independent, nondisruptive, and controlled by simple user-defined policies.

EMC PowerPath/VE

EMC PowerPath/VE for VMware vSphere delivers PowerPath multipathing features to optimize vSphere virtualized environments. PowerPath/VE installs as a kernel module on the ESXi host and works as a multipathing plug-in (MPP) that provides enhanced path management capabilities to ESXi hosts. PowerPath/VE is transparent to the guest operating system running on the virtual machines.

EMC Virtual Storage Integrator for VMware vSphere

EMC Virtual Storage Integrator (VSI) for VMware vSphere is a plug-in to the VMware vSphere client, providing a single management interface for managing EMC storage within the vSphere environment. VSI provides a unified and flexible user experience that allows each feature to be updated independently, and new features to be introduced rapidly in response to changing customer requirements.

When PowerPath/VE is installed on an ESXi host, VSI presents important multipathing details for devices, such as the load-balancing policy, the number of active paths, and the number of dead paths.

EMC VNX Snapshots

VNX Snapshots were introduced in VNX OE block release 5.32.

Unlike copy-onfirst-write (COFW) snapshots, VNX Snapshots do not need to copy base-LUN data to the reserved LUN pool to preserve the point-in-time copy. Instead, new writes to the base LUN are directed to a different location in the storage pool.

Note EMC still supports SnapView, which requires reserved LUN pool configuration.

A snapshot needs to be attached to the Snapshot Mount Point (SMP) to allow host access to it. SMP is a container that holds SCSI attributes, such as the World Wide Name (WWN), Name, and storage group LUN ID. SMP with an attached snapshot can be snapped, known as “snap of a snap”.

With “snap of a snap”, one of the key features of VNX Snapshots, users can create 256 read/writable snaps at a single point in time, which greatly increases the business value of read-only/writable checkpoints.

EMC VNX5300 storage array

The VNX5300 is a member of the VNX series storage platform, which delivers industry-leading innovation and enterprise capabilities for file and block storage in a scalable, affordable, and easy-to-use solution. Designed to meet the high-performance, high scalability requirements of midsize and large enterprises, the VNX series enables enterprises to dramatically grow, share, and cost-effectively manage multiprotocol environments.

The VNX series is powered by the Intel Xeon 5600 series processors, which help make it two to three times faster overall than its predecessors. The VNX quad-core processor supports the demands of advanced storage capabilities such as virtual provisioning, compression, and deduplication.

Oracle Database 11*g* Release 2 SE

Oracle Database 11*g* is available in a variety of editions tailored to meet the business and IT needs of all organizations. Of these editions, Oracle Database 11*g* SE is an affordable, full-featured data management solution that is ideal for small-sized to medium-sized companies. It is available on single or clustered servers with a maximum capacity of four sockets in total. It includes Oracle Real Application Clusters as a standard feature at no additional cost.

VMware vSphere 5

VMware vSphere 5 is the industry's most complete, scalable, and powerful virtualization platform, with infrastructure services that transform IT hardware into a high-performance, shared computing platform, and application services that help IT organizations deliver the highest levels of availability, security, and scalability.

VMware HA is one of the critical vSphere features that provides easy-to-use, cost-efficient high availability for applications running on virtual machines. In the event of a physical server failure, affected virtual machines are automatically restarted on other production servers with spare capacity.

The combination of VMware HA and other availability features of the vSphere platform enables organizations to select and easily deliver the level of availability required for all of their important applications.

Solution configuration

Introduction

This solution includes one VNX5300 storage array and two ESXi hosts running VMware ESX™ Server 5.0 in a VMware cluster environment.

The validation team created multiple virtualized 500 GB Oracle SE databases and connected them to the VNX5300 storage array using the FC protocol. To comply with the Oracle SE licensing restrictions, we limited the number of sockets on the ESXi cluster to four.

For this SAN configuration, each ESXi server contained a dual-port FC HBA to facilitate FC connectivity. We used EMC PowerPath/VE with the storage system to manage the I/O paths. PowerPath managed two active I/O paths and two passive I/O paths on each ESXi server.

In this solution, we enabled FAST Cache and FAST VP to optimize read and write performance. It is worth noting that FAST Cache and FAST VP technologies are under the same license, which means customers can purchase once for both technologies.

Solution architecture

Figure 1 depicts the overall architecture of the solution environment.

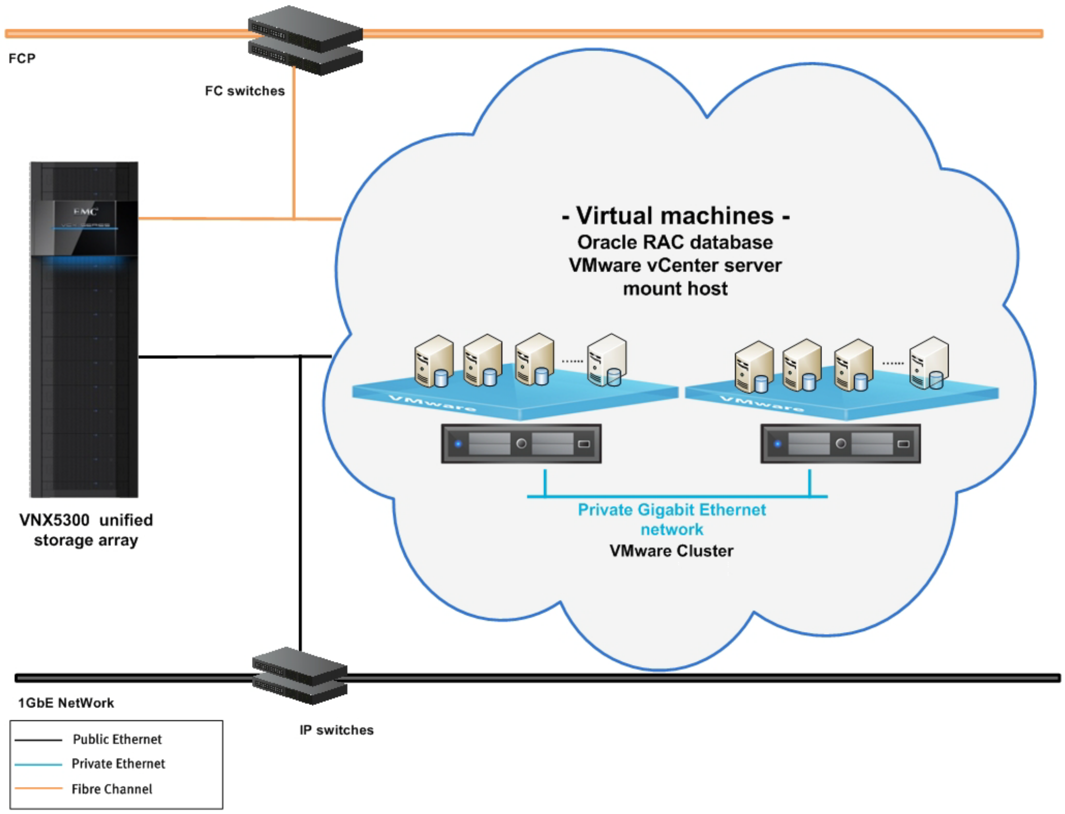

Figure 1. Solution architecture

Database and environment profile

Table 3 details the database and environment profile for the solution.

Table 3. Database and environment profile

Profile	Details
Database type	OLTP
Database size	500 GB
Oracle RAC	Two nodes
Oracle SGA	5 GB per node
Workload profile	TPC-C like benchmark, generated by Swingbench 2.4
Database read/write ratio	60/40
User scaling	200 to 800

Hardware resources

Table 4 details the hardware resources for the solution.

Table 4. Solution hardware environment

Hardware	Quantity	Configuration
EMC VNX5300 storage array	1	EMC VNX5300 with: • 17 x 2 TB 7.2k NL_SAS drives • 4 x 600 GB 15k SAS drives (vault) • 31 x 600 GB 15k SAS drives • 3 x 100 GB Flash drives
VMware ESXi servers for Oracle databases	2	Each server with: • 2 x 6-core CPUs • 48 GB RAM • 6 x 1 Gb NICs • 2 x 8 Gb HBA (dual-port HBA)
Ethernet switches	2	1 Gb/s Ethernet switches
FC switches	2	8 Gb/s FC switches

Software resources

Table 5 details the software resources for the solution.

Table 5. Solution software environment

Software	Version	Purpose
EMC VNX OE for block	05.32.000.5.006	VNX operating environment
EMC Unisphere™	1.2.0	VNX management software
EMC Powerpath/VE	5.7.b173	Multipathing and load balancing software
VMware vSphere	5.0 Update 1	Hypervisor hosting all virtual machines
VMware vCenter™	5.0 Update 1	Management of VMware vSphere
Oracle Database 11*g* Release 2	Standard Edition 11.2.0.3	Oracle database software
Oracle Enterprise Linux	6.2	Operating system for Database servers
Oracle Grid Infrastructure 11*g* Release 2	Standard Edition 11.2.0.3	Oracle cluster software

Virtualized Oracle server configuration

Table 6 lists the hardware and software components of the production RAC database server in the virtual environment.

Table 6. Production RAC database server components

Hardware component	Quantity
Processor	4 vCPUs
Memory	8 GB
Software component	**Version**
Oracle Enterprise Linux	6.2
Oracle Database 11*g* Release 2 SE RAC	11.2.0.3

Table 7 lists the hardware and software components of the test/dev single-instance database server in the virtual environment.

Table 7. Test/dev single instance database server components

Hardware component	Quantity
Processor	2 vCPUs
Memory	4 GB
Software component	**Version**
Oracle Enterprise Linux	6.2
Oracle Database 11*g* Release 2 SE	11.2.0.3

Oracle storage layout

In this solution, the validation team organized physical drives into the storage pools on the VNX5300 array, which is the recommended approach for the VNX storage system.

For more information about storage pools on VNX, refer to *Applied Best Practices Guide: EMC VNX Unified Best Practices for Performance.*

Table 8 lists the virtual pool design on VNX5300 and Figure 2 depicts the Oracle storage layout of the environment.

Table 8. Virtual pool design on VNX5300

Storage pool	Types of disks	Protection type	Disks in the pool	Number of LUNs	Size of LUNs	File type	FAST Cache enabled?
Data pool (FAST VP)	Tier 1: SAS drive 10k rpm	RAID 5	(4+1)x3	6	1 TB	Datafiles, control files	Yes
	Tier 2: NL_SAS drive 7k rpm	RAID 6	6+2				
Redo pool	SAS drive 10k rpm	RAID 10	4+4	10	10 GB	Redo log files	No
Temp pool	SAS drive 10k rpm	RAID 10	2+2	6	200 GB	Temp files	No
Fast Recovery Area (FRA) pool	NL_SAS drive 7k rpm	RAID 6	6+2	6	1 TB	FRA files	No
Cluster Ready Service (CRS) pool	SAS drive 10k rpm	RAID 10	2+2	10	5 GB	CRS files	No
Virtual machine pool	SAS drive 10k rpm	RAID 1	1+1	2	60 GB	Virtual machine hosts	No

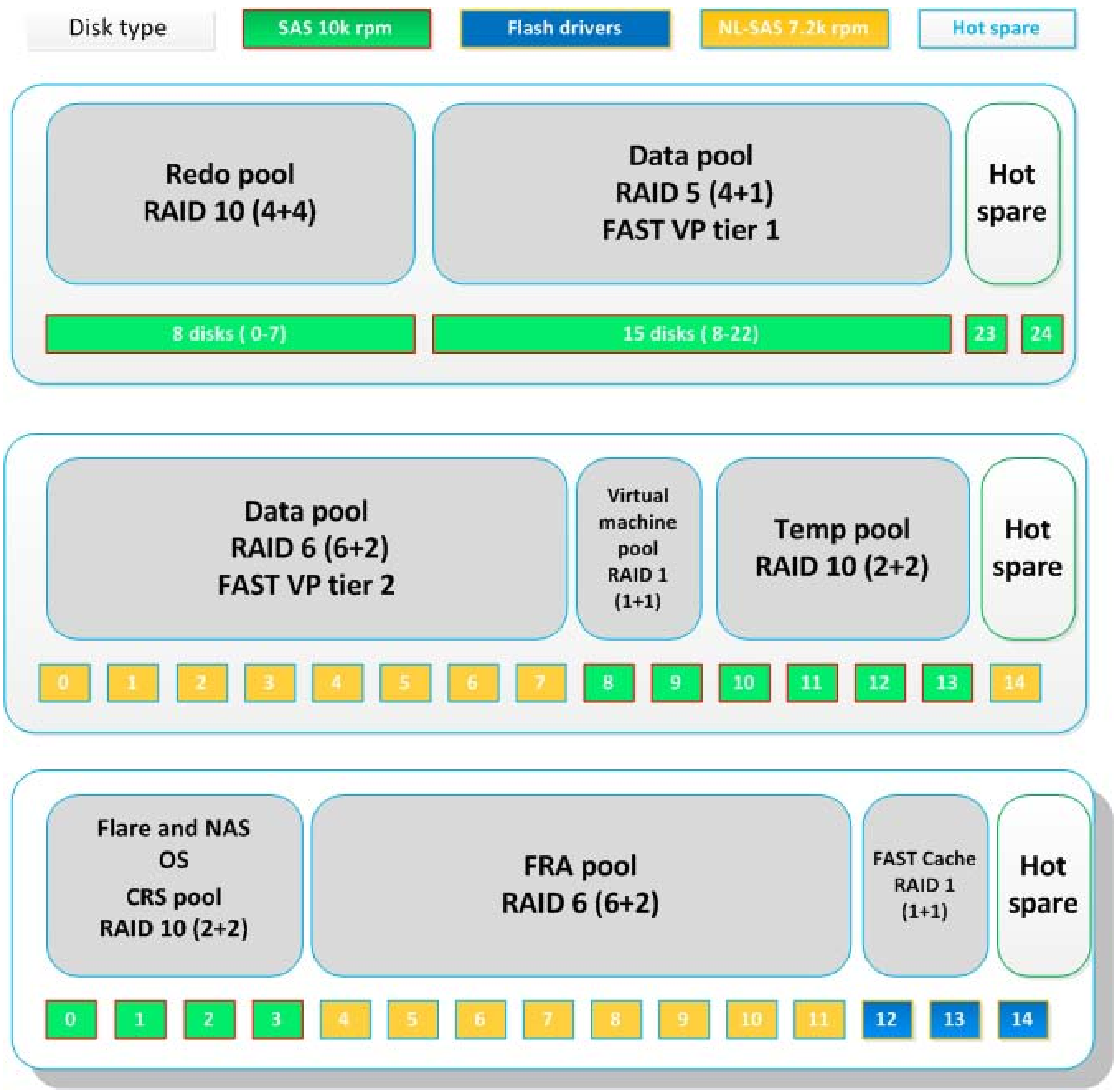

Figure 2. EMC VNX storage layout

All virtual machines in this configuration use virtual disks (VMDK) from VMware Virtual Machine File System (VMFS) data store volumes. Those virtual disks are shared by multiple virtual machines by enabling the multi-writer option. Each VMFS datastore hosts a single VMDK disk, ensuring high performance and zero contention.

For more information about using the multi-writer flag, refer to the article *Disabling simultaneous write protection provided by VMFS using the multi-writer flag* on the VMware Knowledge Base.

On the VNX storage array, this configuration allowed for the placement of the Oracle datafiles, temp files, control files, and log files on different types of disks. The datafiles and control files resided on the data pool composed of NL_SAS drives (with RAID 6 protection) and SAS drives (with RAID 5 protection). Redo logs and temp files were on the SAS drive pool (with RAID 10 protection), and FRA files were on the NL_SAS drive pool (with RAID 6 protection).

FAST Cache and FAST VP configuration

FAST Cache configuration

Overview

FAST Cache provides read and write caching using a private RAID 1 group consisting of Flash disks.

To use FAST Cache, make sure you meet the following criteria:

- The application workload must first be analyzed to determine whether the application will benefit from FAST Cache and to determine the optimal size of the FAST Cache.
- The storage system must have the FAST Cache enabler installed.
- The storage system must have Flash disks that are not already in a storage pool.
- FAST Cache must be configured on the storage system.
- FAST Cache must be enabled for the RAID group LUNs and/or the storage pools that are to use the FAST Cache.

FAST Cache is most effective when application workloads exhibit data activity skew where a small subset of data is responsible for most of the data set's activity. For example, for a LUN that has less than 5 percent data accounts for more than 80 percent IOPS, if we promote the 5 percent data to FAST Cache, we can achieve a millisecond response time at the lowest possible cost. For those applications whose data sets exhibit a high degree of skew, FAST Cache can be assigned to concentrate a high percentage of application IOPs on the Flash capacity.

For further information on configuring FAST Cache, refer to the white paper *Leveraging EMC FAST Cache with Oracle OLTP Database Applications*.

Analyzing the application workload

The decision to implement FAST Cache should only be made after the application workload characteristics are measured and analyzed. Array-level tools are available to EMC field and support personnel for determining both the suitability of FAST Cache for a particular environment and the right size FAST Cache to configure. Contact your EMC sales teams for guidance.

The size of the application's active working set, the IOPS requirement, the RAID type, and the read/write ratio determine whether a particular application will benefit from using FAST Cache, and what the optimal cache size should be.

Preferred application workloads for FAST Cache include:

- Small-block random I/O applications with high locality
- High frequency of access to the same data

The workload characteristics of OLTP databases make them especially suitable for using FAST Cache.

As shown in Figure 3, we used an array-level tool, EMC Unified Block Locality Analyzer, to determine the active data set size and suitability of FAST Cache. In this solution, it was confirmed that the Oracle database workload would benefit from using FAST Cache and indicated that two 100 GB Flash drives would be the optimal configuration.

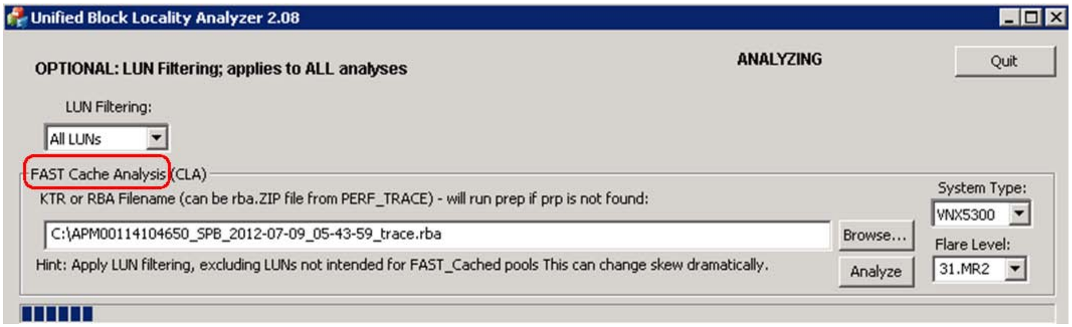

Figure 3. FAST Cache analysis by Unified Block Locality Analyzer tool

For further information, see the following white papers: *EMC CLARiiON, Celerra Unified, and VNX FAST Cache* and *Deploying Oracle Database on EMC VNX Unified Storage*.

In line with EMC recommendations, FAST Cache was enabled for the Oracle datafiles only. Oracle archive files and redo log files have a predictable workload composed mainly of sequential writes. These can be efficiently handled by the array's write cache and assigned hard disk drives (HDDs). Enabling FAST Cache on these files is neither beneficial nor cost efficient.

Figure 4 shows the FAST Cache configuration for the solution:

- **Number of Disks** was set to 2, as recommended by the workload analysis.
- **RAID Type** was set to RAID 1.
- Per EMC best practices, we manually selected two Flash disks in Bus 0 Enclosure 0 from the list of available disks presented by Unisphere.

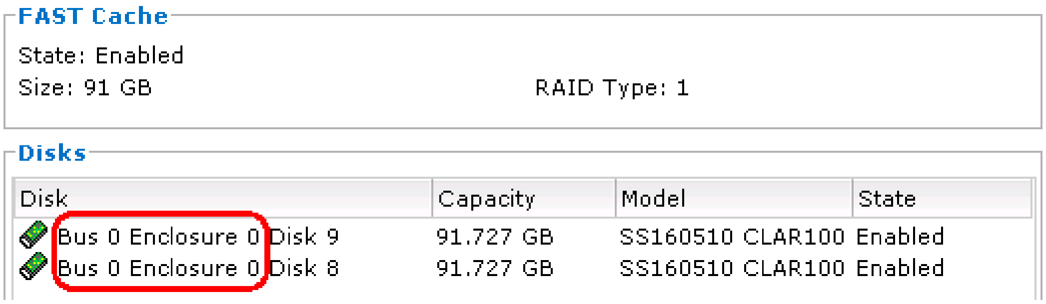

Figure 4. FAST Cache configuration

FAST VP configuration

Overview

FAST VP is a simple and elegant solution for dynamically matching storage requirements with changes in the frequency of data access. Generally, FAST VP segregates disk drives into the following three tiers:

- Extreme performance tier—Flash drives
- Performance tier—SAS drives
- Capacity tier—NL-SAS drives

You can use FAST VP to aggressively reduce TCO and/or to increase performance. A target workload that requires a large number of performance tier drives can be serviced with a mix of tiers, and a much lower drive count.

FAST VP can be used in combination with other performance optimization software, such as FAST Cache. A common strategy is to use FAST VP to gain TCO benefits while using FAST Cache to boost overall system performance. There are other scenarios where it makes sense to use FAST VP for both purposes.

In this case, we used two tiers (SAS and NL_SAS) for FAST VP configuration. From cost-efficiency perspective, FAST VP can optimize disk utilization and efficiency, while FAST Cache with a small number of Flash drives serves the data that is accessed most frequently.

Tiering policies

FAST VP includes the following tiering policies:

- Start high then auto-tier (new default policy)
- Auto-tier
- Highest available tier
- Lowest available tier
- No data movement

Start high then auto-tier (new default policy)

Start high then auto-tier is the default setting for all pool LUNs on creation. Initial data placement is on the highest available tier and data movement is subsequently based on the activity level of the data. This tiering policy maximizes initial performance and takes full advantage of the most expensive and fastest drives first, while providing subsequent TCO by allowing less active data to be tiered down, making room for more active data in the highest tier.

When there is a pool with multiple tiers, the start high then auto-tier design is capable of relocating data to the highest available tier, regardless of the drive type combination. Also, when adding a new tier to a pool, the tiering policy remains the same and there is no need to manually change it.

Auto-tier

FAST VP relocates slices of these LUNs based on their activity level. Slices that belong to LUNs with the auto-tier policy have second priority for capacity in the highest tier of the pool, after the LUNs set to the highest tier.

Highest available tier

The highest available tier setting should be selected for those LUNs that, although not always the most active, require high levels of performance whenever they are accessed. FAST VP prioritizes slices of a LUN with the highest available tier selected above all other settings.

Slices of LUNs set to the highest tier are rank ordered with each other according to the activity. Therefore, in cases where the sum total of LUN capacity set to the highest tier is greater than the capacity of the pool's highest tier, the busiest slices occupy that capacity.

Lowest available tier

The lowest available tier setting should be selected for LUNs that are not performance-sensitive or response time-sensitive. FAST VP maintains slices of these LUNs on the lowest storage tier available, regardless of the activity level.

No data movement

The no data movement policy may only be selected after a LUN has been created. FAST VP does not move slices from their current positions when the no data movement selection has been made. Statistics are still collected on these slices for use when the tiering policy is changed.

In this solution, the **Start High then Auto-Tier** policy was set to **Scheduled**.

The **Data Relocation Schedule** setting was configured as **Monday** to **Sunday**, starting from **00:00** to **23:45**, which determines the time window when FAST VP moves data between tiers.

Note The **Data Relocation Rate** and **Data Relocation Schedule** are highly dependent on the real workload in a customer environment. Usually, setting the **Data Relocation Rate** to **Low** has less impact on the current running workload.

For details about FAST VP configuration, refer to *EMC FAST VP for Unified Storage Systems*.

EMC VSI overview

Overview

EMC VSI provides enhanced visibility into VNX storage directly from the vCenter GUI.

Datastore and storage pool information from VSI

Figure 5 shows the datastore and storage pool information, which provides information about storage pool usage for the Data5A datastore.

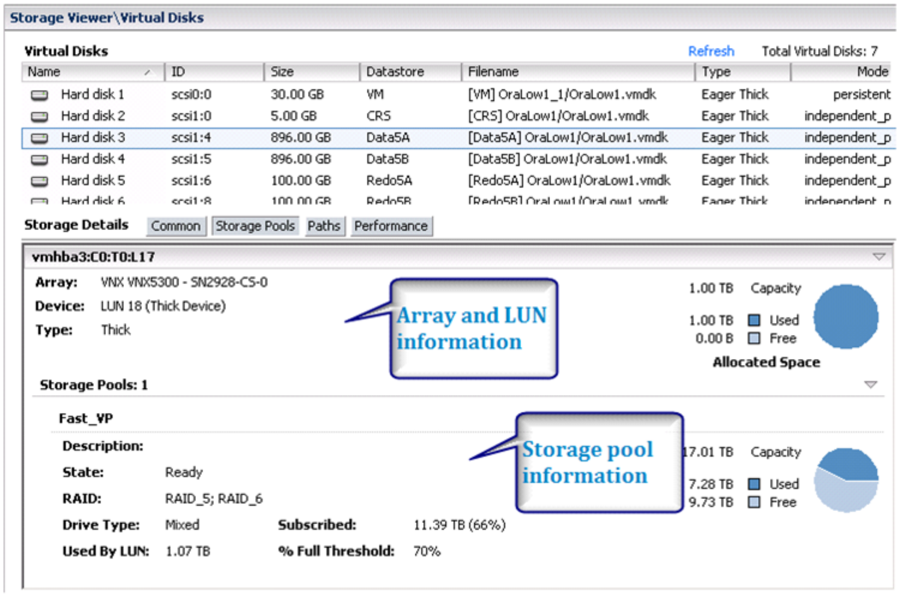

Figure 5. Datastore and storage pool information viewed from VSI

Figure 5 shows the ESXi server and the VNX5300 storage mapping with details about VMDK, LUNs, and storage pools of the array. The Storage Viewer identifies the details about VMDK, such as the VNX LUNs that host VMDK, the paths to the physical storage, the storage pool usage information, and the virtual disk performance statistics.

Testing and validation

Test overview

This solution summarizes a series of implementation procedures and optimal configurations that we observed during the validation of a cost-efficient solution for Oracle Database 11*g* Release 2 SE using FAST VP and FAST Cache, VNX Snapshots, and VMware HA on the VNX5300 storage array.

Notes

- Results are highly dependent on workload, specific application requirements, and system design and implementation. Relative system performance will vary as a result of these and other factors. Therefore, it is suggested not to use this workload as a substitute for a specific customer application benchmark when contemplating critical capacity planning and product evaluation decisions.
- All performance data contained in this report was obtained in a rigorously controlled environment. Results obtained in other operating environments may vary significantly.
- EMC Corporation does not warrant or represent that a user can or will achieve similar performance expressed in transactions per minute.

Test objectives

The objectives of this solution are to demonstrate:

- How EMC FAST VP and FAST Cache work together to achieve more transactional throughput by deploying more Oracle 11*g* Release 2 SE databases on the same physical devices.
- The use of EMC VNX Snapshots to:
  - Create multiple Oracle test/dev environments by using the latest VNX Snapshots "writable snap of a writable snap" feature to replicate a running Oracle production database, with minimal performance impact and no downtime.
  - Carry out the point-in-time backup of the Oracle test/dev environment by using the VNX Snapshots "read-only snap of a writable snap" feature to restore the test/dev environment in case of logical errors in the test/dev database.
- The enhanced high availability of Oracle applications by showing VMware HA's ability to automatically restart RAC nodes on the surviving ESXi host and achieve higher levels of protection against the failure of the physical server.

Test scenarios

This section contains the test scenarios shown in Table 9.

Table 9. Test scenarios

Test scenarios
Scenario A: Storage resource optimization
Scenario B: Virtual test/dev environment rapid provisioning and protection
Scenario C: Business continuity and disaster recovery

Scenario A: Storage resource optimization

Overview

This scenario was designed to demonstrate how FAST VP and FAST Cache allow customers to optimize the storage array for deploying one or more Oracle databases.

This scenario shows how:

- EMC FAST VP technology creates a storage pool out of mixed drive types, monitors the data activity level, and automatically moves data at sub-LUN granularity to the corresponding tiers available within a pool.
- EMC FAST Cache technology, as a means of introducing an extra layer of the cache using Flash drives between the DRAM and rotating spindles, improves the performance of the most latency-sensitive data.

Test configuration

To begin with, we configured the FAST VP two-tier storage pool for datafiles using high-performance SAS drives and high-capacity NL_SAS drives, to migrate the hot data from the NL_SAS tier to the SAS tier in the pool. We identified the hot data by periodically analyzing the characteristics of the workload. After FAST VP relocated the data for optimal performance, we enabled FAST Cache to further optimize the storage performance by caching the hot data on the Flash drives.

Note In line with EMC recommendations, we enabled the FAST VP pool and FAST Cache for the Oracle datafiles only. For more details, refer to *EMC VNX unified best practices for performance.*

To minimize the Oracle software licensing cost, we used Oracle SE rather than Oracle Enterprise Edition (EE). Two 500 GB OLTP databases (single-instance and RAC database) were deployed on this environment.

Oracle SE is available on a single server or clustered servers with a maximum of four sockets, including Oracle RAC as a standard feature with no additional cost.

VMware vSphere provided a virtualization platform for this scenario, with VMware virtual machines hosting the Oracle single-instance or RAC database nodes. VMware vCenter provides a centralized management platform for the vSphere environments, enabling control and visibility in every level of the virtualized infrastructure.

Test procedure

We used the following procedure for this scenario:

1. Create a FAST VP pool composed of two drive types (SAS and NL_SAS).
2. Deploy a virtualized single-instance Oracle 11*g* Release2 SE database (500 GB).
3. Start the Swingbench workload using the TPC-C like benchmark with the 200-user load against the single-instance database.
4. Enable FAST Cache.
5. Convert the single-instance database to a two-node RAC database.

6. Start the Swingbench workload using the TPC-C like benchmark with the 400-user load against the RAC database.
7. Add the second RAC database.
8. Start the Swingbench workload using the TPC-C like benchmark with the 800-user load against both RAC databases.

Testing results

Single instance database

Figure 6 compares the Swingbench transactions per minute (TPM) and the response time before and after FAST VP and FAST Cache were enabled.

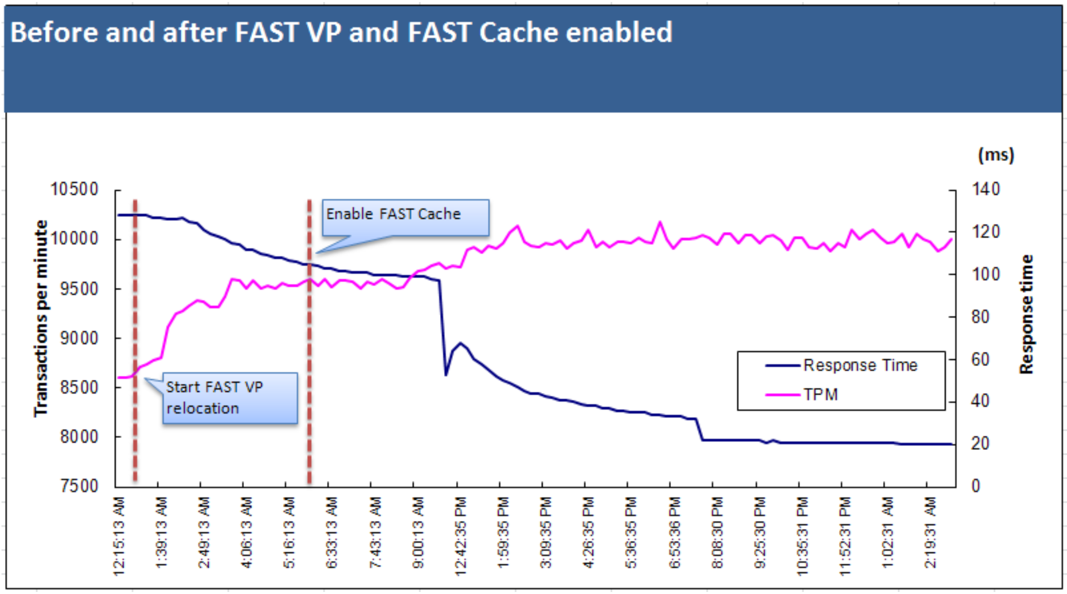

Figure 6. Before and after FAST VP and FAST Cache enabled (single instance)

In this case, FAST VP took eight hours to relocate the data for optimal performance. As more hot data was promoted from the NL_SAS tier to the SAS tier during the data relocation period, as shown in Figure 7 and Figure 8, the proportion of SAS tier increased from 21.2 percent to 93.59 percent. As shown in Figure 6, during data relocation, the number of transactions processed per minute increased by 11 percent, peaking at 9,600. The response times decreased steadily from 130 ms to 100 ms.

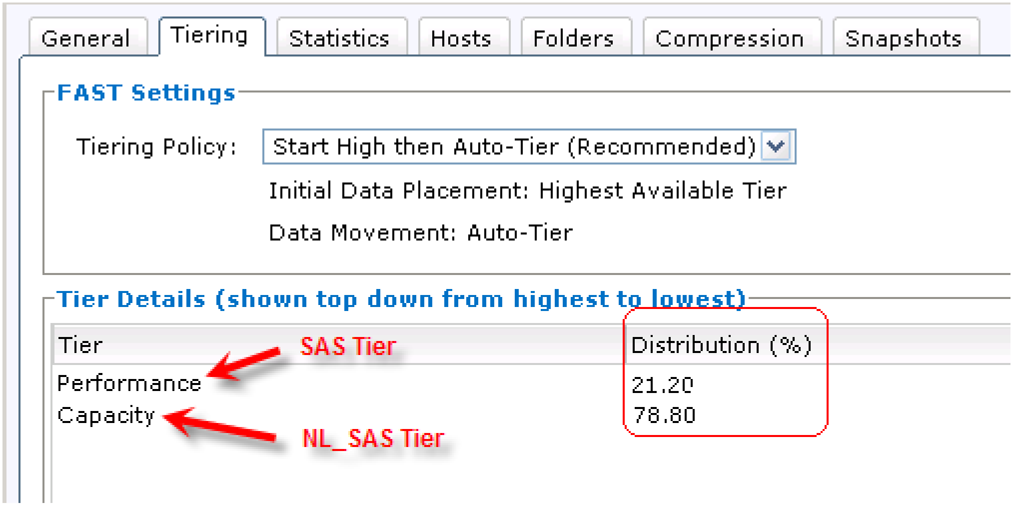

Figure 7. FAST settings and tiering details before FAST VP relocated data

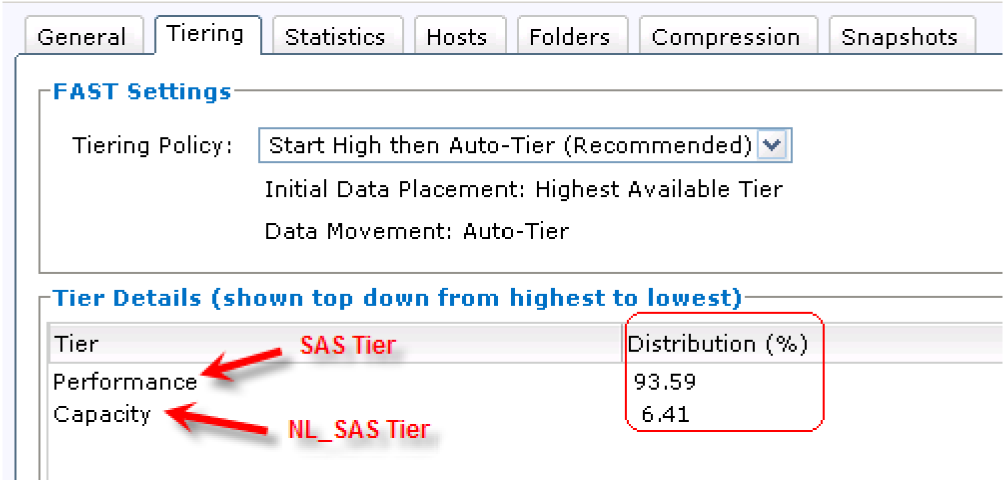

Figure 8. FAST settings and tiering details after FAST VP relocated data

After we enabled FAST Cache for the thin pool where the datafiles were located, the hot data was continually promoted to the Flash drives. The number of TPM increased constantly from 9,600 to approximately 10,200. The response time decreased from 100 ms to 20 ms.

Figure 9 and Figure 10 were captured from the single-instance Oracle database Automatic Workload Repository (AWR) report before and after we enabled FAST VP and FAST Cache.

Top 5 Timed Foreground Events

Event	Waits	Time(s)	Avg wait (ms)	% DB time	Wait Class
db file sequential read	1,896,046	35,892	19	95.27	User I/O
DB CPU		1,450		3.85	
log file sync	269,477	408	2	1.08	Commit
read by other session	131	4	33	0.01	User I/O
library cache: mutex X	1,123	2	1	0.00	Concurrency

Figure 9. TOP 5 timed foreground events before FAST VP and FAST Cache were enabled

Top 5 Timed Foreground Events

Event	Waits	Time(s)	Avg wait (ms)	% DB time	Wait Class
db file sequential read	932,886	3,205	3	82.68	User I/O
DB CPU		579		14.93	
log file sync	130,717	132	1	3.40	Commit
read by other session	88	2	24	0.05	User I/O
library cache: mutex X	375	1	2	0.02	Concurrency

Figure 10. TOP 5 timed foreground events after FAST VP and FAST Cache were enabled

There was a significant improvement on the I/O subsystem before and after FAST VP relocated the data, and when we enabled FAST Cache. With FAST VP and FAST Cache, the bottleneck from the **db file sequential read** events was reduced by more than 80 percent, from 19 ms to 3 ms.

During the test, we made no changes to the database. All performance improvements directed resulted from enabling FAST VP and FAST Cache.

RAC databases

We continued the test by converting the single-instance database to two-node RAC database. We then added the second two-node RAC database to get more workload under the same hardware/software configuration.

Figure 11 shows that the number of TPM reached 37,298 in total when two RAC databases were running on the VMware vSphere virtualized platform. It also demonstrates that, with FAST Cache and FAST VP monitoring and optimizing the storage to meet the changing demands of the workload, there was little performance impact on any running RAC databases when we scaled out the number of RAC database.

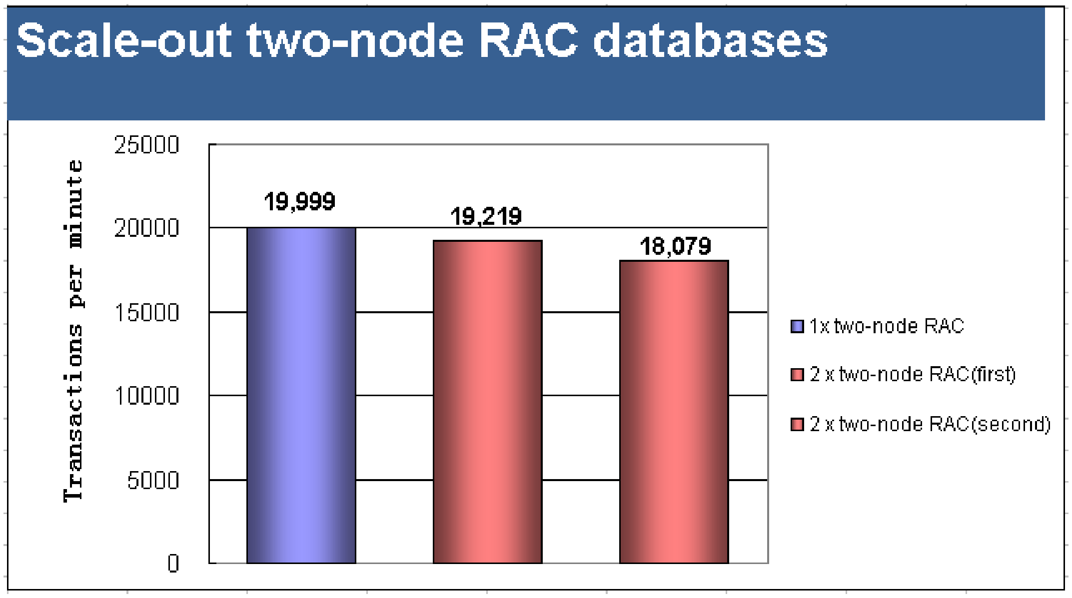

Figure 11. Scale-out two-node RAC databases

Scenario B: Virtual test/dev environment rapid provisioning and protection

Overview

This scenario provides a rapid, low-impact solution for provisioning Oracle test/dev environments by replicating the production database with EMC VNX Snapshots. This scenario also covers the use of VNX Snapshots to restore a logically corrupted test/dev database.

This scenario enables the use of "writable snap of a writable snap" and "read-only snap of a writable snap", both of which are new features of VNX Snapshots.

This solution demonstrates how to create:

- A "writable snap of a writable snap" for dev/test environment provisioning purposes.
- A "read-only snap of a writable snap" by using it as a backup source.

Test configuration

The storage configuration in this scenario was identical to the storage solution in Scenario A: Storage resource optimization.

We used one RAC database as the source database running with 400-user load when VNX Snapshots-related operations were performed on the source database.

Test procedure

We used the following two procedures to validate the dev/test environment fast provision and recovery by using VNX Snapshots.

Procedure 1: Writable snap of a writable snap

In this procedure, by using the following VNX Snapshots technology, "writable snap of a writable snap," we provisioned the dev/test instances in less time and used less storage.

Figure 12 shows the procedure for creating multiple writable snaps of a writable snap.

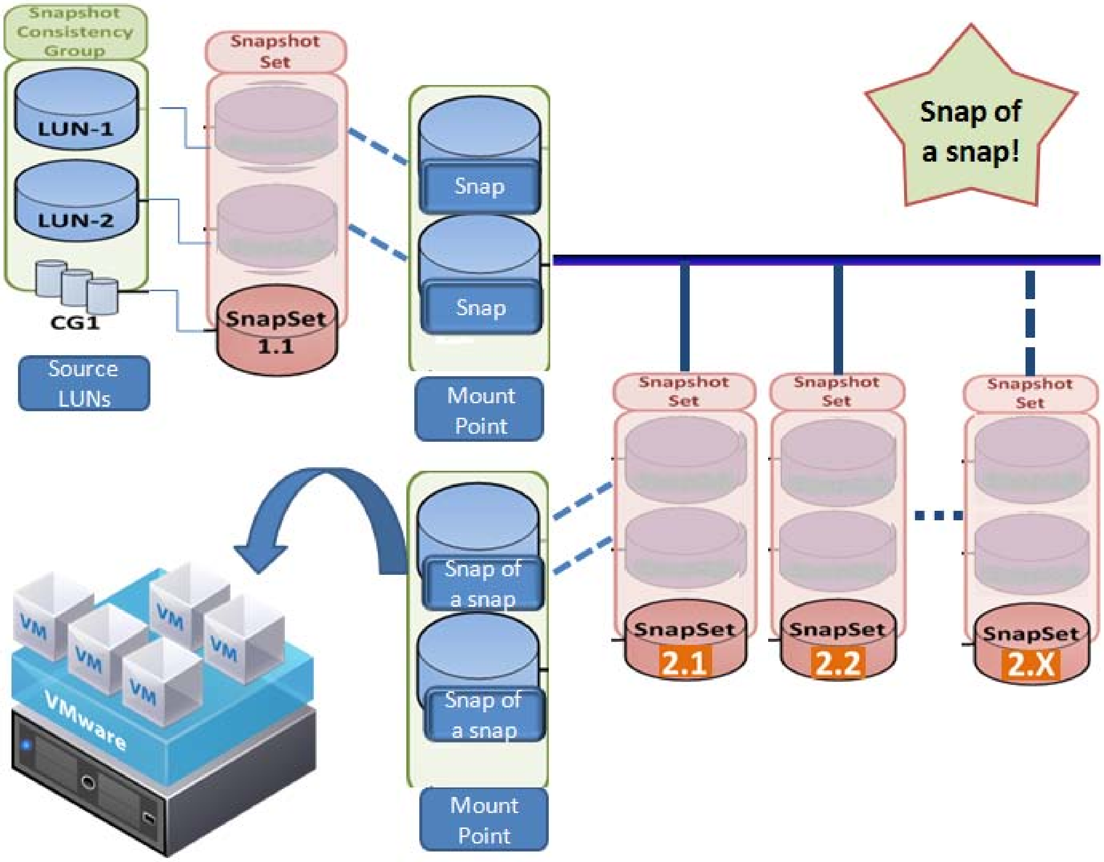

Figure 12. Example of writable snaps of a writable snap

Follow these steps to create multiple writable snaps of a writable snap:

1. Create one snapshot consistency group that includes the individual datafile and log file LUNs of the production RAC database.
2. Create the snap (SnapSet 1.1) of the LUNs from the consistency group.
3. Create SMPs for the consistency group and attach the snap (SnapSet 1.1) to the SMPs in VNX Unisphere as shown in Figure 13.

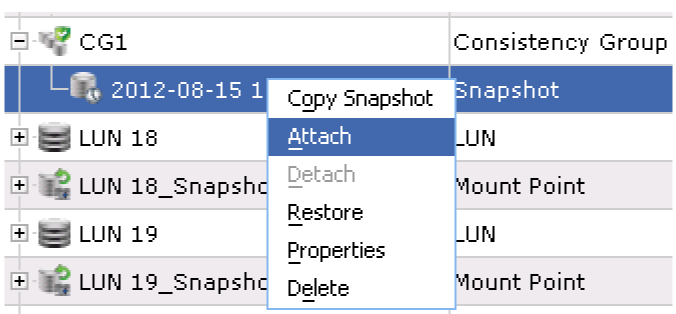

Figure 13. Attaching snap to the mount point

As shown in Figure 14, the properties of the snapshot become writable after attaching the snap.

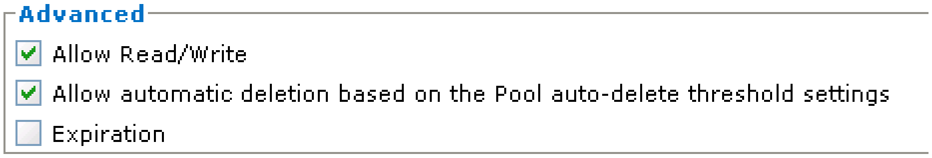

Figure 14. Properties of the snapshot

4. Create another snapshot consistency group that includes the SMPs created in Step 3, as shown in Figure 15.

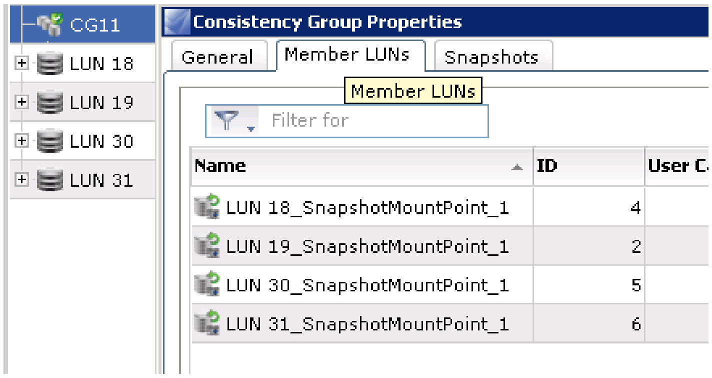

Figure 15. Consistency group composed of SMPs properties

5. Create the snap (SnapSet 2.1) of the SMPs created in step 3, which is a snap of snap.

6. Create another eight snaps (SnapSet 2.2 to 2.9) of snap by copying the snapshot function in VNX Unisphere, as shown in Figure 16.

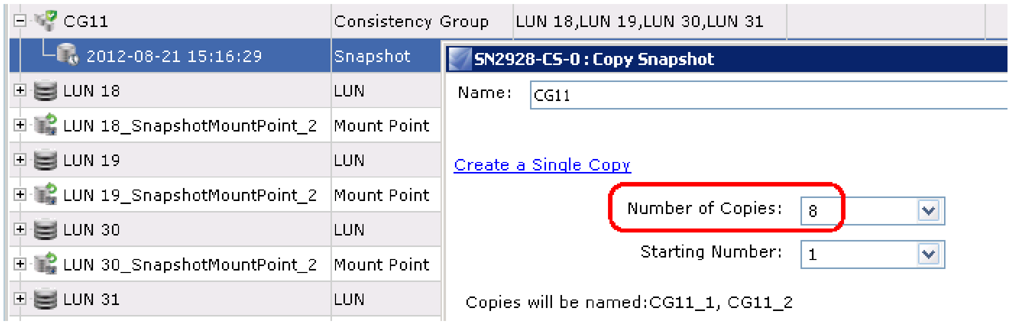

Figure 16. Rapid Oracle-instance provisioning with one click on VNX Unisphere

Procedure 2: Read-only snap of a writable snap

In this procedure, by using the following VNX Snapshots technology, "read-only snap of a writable snap," we snapped the dev/test database as a backup source and recovered the database in minutes after the database logical corruption took place.

Figure 17 shows an example of the read-only snap of a writable snap.

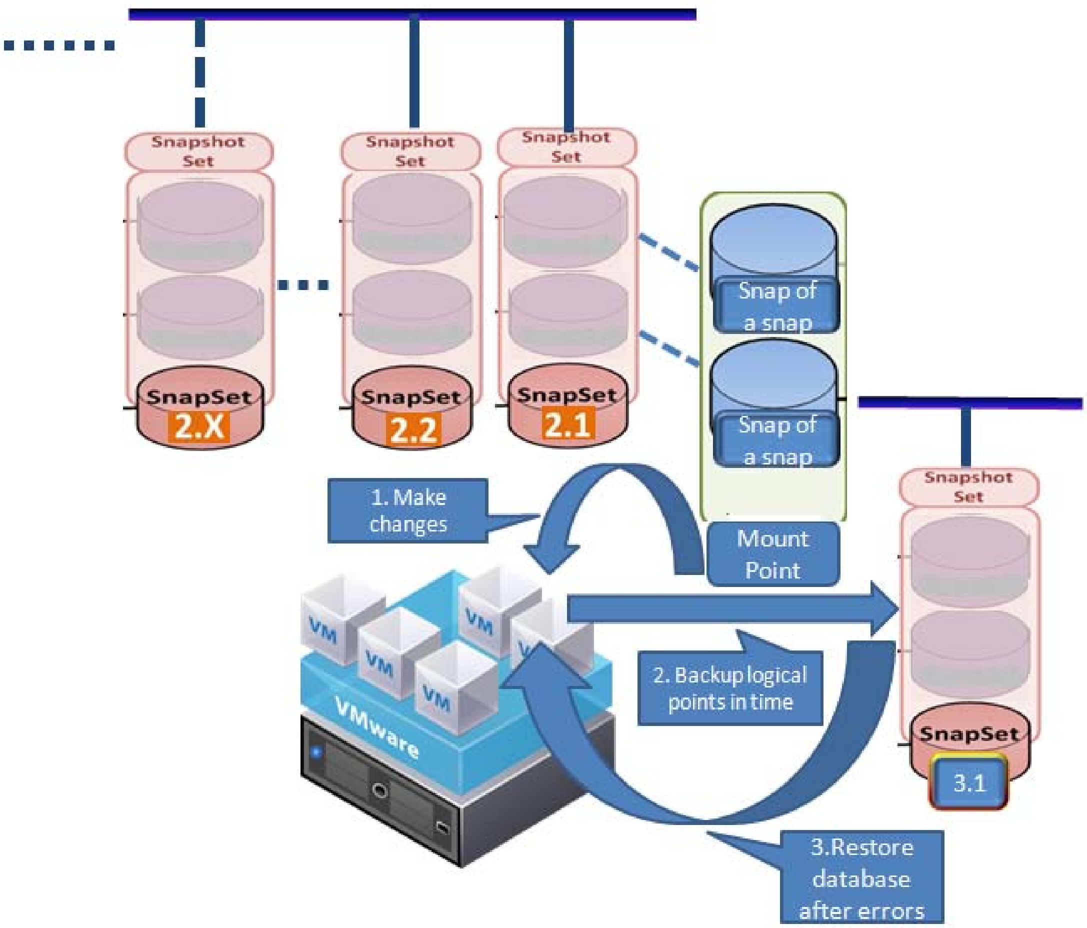

Figure 17. Example of read-only snap of a writable snap

Follow these steps to test a read-only snap of a writable snap:

1. Present the snapshot LUNs to the ESXi server and the virtual machine to be mounted, mount one of the writable snaps of a writable snap as a new test/dev database, and open the database.
2. Insert records into the single-instance dev/test database. Figure 18 shows the record count and the timestamps of every record. This information was later used to validate the restore of the database.

```
SQL> select * from test;

         X Y
---------- --------------------
         3 15-AUG-2012 17:30:24
         4 15-AUG-2012 17:30:24
         5 15-AUG-2012 17:30:24
         6 15-AUG-2012 17:30:24
         7 15-AUG-2012 17:30:24
         8 15-AUG-2012 17:30:24
         9 15-AUG-2012 17:30:24
        10 15-AUG-2012 17:30:24
         1 15-AUG-2012 17:30:24
         2 15-AUG-2012 17:30:24

10 rows selected.
```

Figure 18. Record count and timestamp in the dev/test database before data deletion

3. Create a read-only snap of the writable snap of the dev/test database, as shown in Figure 19.

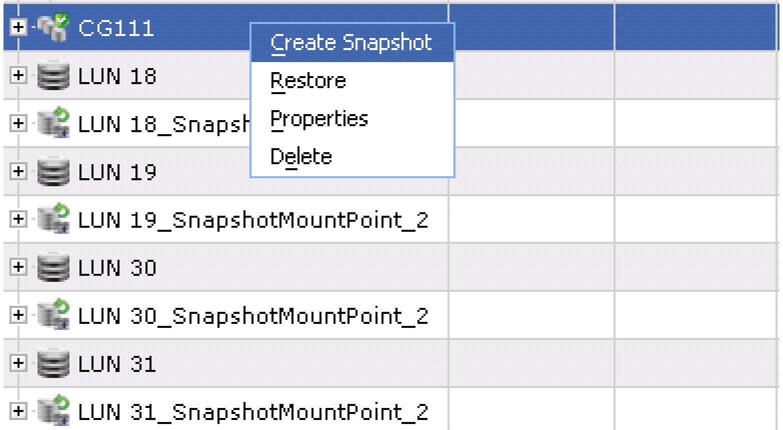

Figure 19. Read-only snap of the writable snap creation

4. Simulate “accidental” data deletion and close the test/dev database, as shown in Figure 20.

```
SQL> drop table test purge;

Table dropped.

SQL> desc test
ERROR:
ORA-04043: object test does not exist


SQL>
```

Figure 20. “Accidental” data deletion

5. Restore the database by selecting **Restore** from the list, as shown in Figure 21.

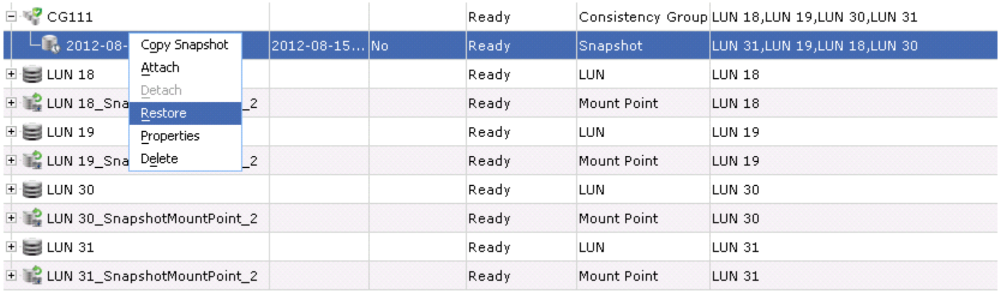

Figure 21. Restore the database using "read-only snap of a writable snap" (Part I)

6. The wizard in Figure 22 is displayed. Click **Yes** to continue.

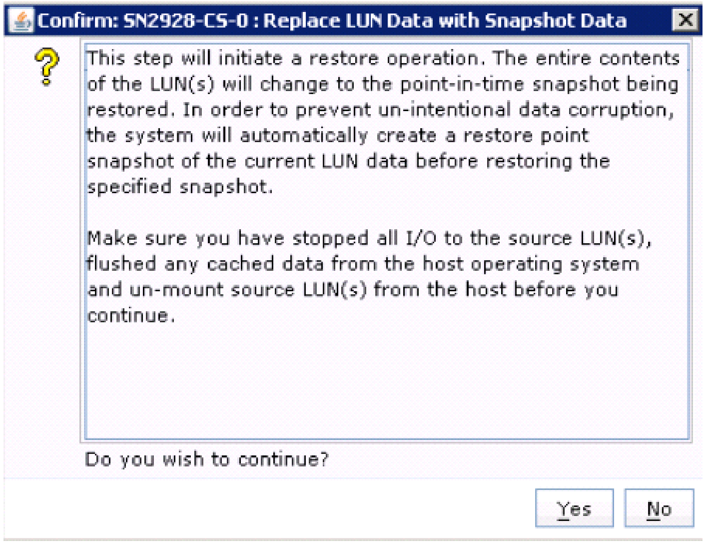

Figure 22. Restore the database using "read-only snap of a writable snap" (Part II)

A restore operation is initiated and the entire contents of the LUNs are changed to the point in time that the snapshot was restored.

7. To verify the data was restored, display the record count and timestamp from the restored dev/test database. The output of the query verified that the database was restored to the point before the logical corruption, as shown in Figure 23.

```
SQL> select * from test;

         X Y
---------- --------------------
         3 15-AUG-2012 17:30:24
         4 15-AUG-2012 17:30:24
         5 15-AUG-2012 17:30:24
         6 15-AUG-2012 17:30:24
         7 15-AUG-2012 17:30:24
         8 15-AUG-2012 17:30:24
         9 15-AUG-2012 17:30:24
        10 15-AUG-2012 17:30:24
         1 15-AUG-2012 17:30:24
         2 15-AUG-2012 17:30:24

10 rows selected.
```

Figure 23. Record count and timestamp in the dev/test database after database restore

Testing results

Writable snap of a writable snap

The production database was snapped and then nine other point-in-time snaps were created, which are also called “snap of a snap”. During the creation, in order to validate whether there was any performance impact on the production database, we ran a Swingbench workload on the production database when the snaps were created.

Figure 24 shows a slight decrease in TPM during the first snap creation. During the creation of the other nine snaps (snap of a snap), TPM was reduced by 4 percent, from 19,900 to 19,150.

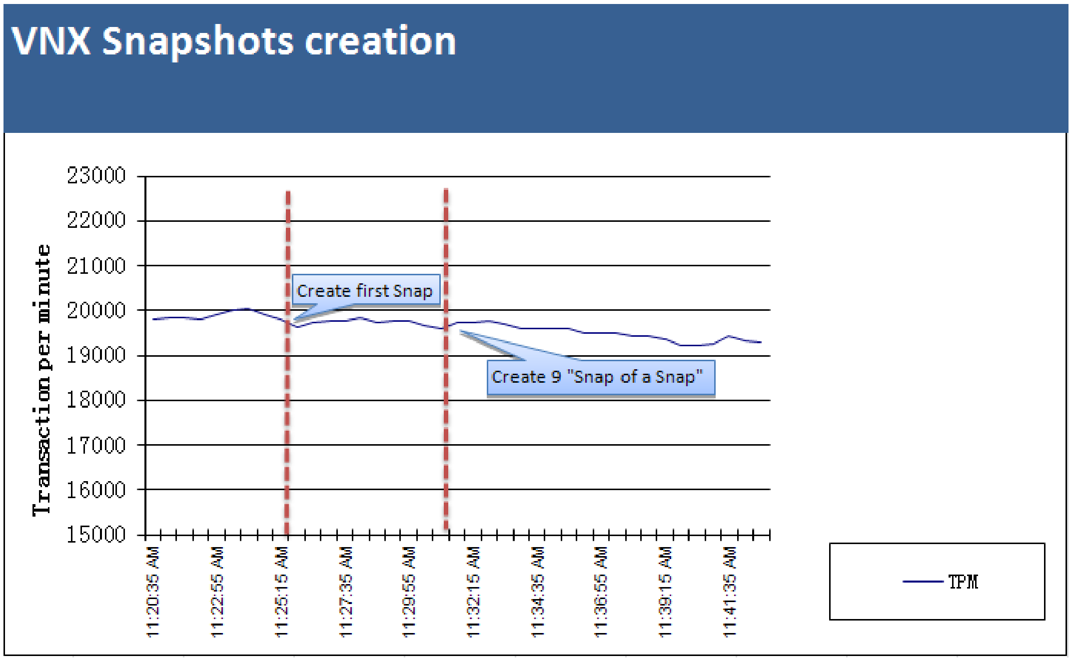

Figure 24. VNX Snapshots creation

This test scenario shows that by using VNX Snapshots, the production database can be provisioned quickly for test, development, and other purposes, while all the involved operations having minimal performance impact on the source production database.

Read-only snap of a writable snap

In this scenario, "read-only snap of a writable snap" feature took only a few minutes to restore the test/dev database to a point in time before the logical corruption.

It is validated that with this point-in-time versioning of a writable snap, customers can use writable snaps of production databases in nonproduction scenarios such as patch testing or reporting.

Scenario C: Business continuity and disaster recovery

Overview

In this scenario, we used both VMware vSphere HA and Oracle RAC to maximize the availability of the Oracle RAC databases hosted on the VMware virtual machines.

The scenario demonstrates that, in the case of an ESXi host failure, VMware HA can automatically migrate the virtual RAC nodes from the failed ESXi host to the surviving ESXi host. Thus the RAC databases were restored to full capacity as soon as possible, while keeping the Oracle RAC databases available during the whole process.

Test configuration

The storage configuration in this scenario was identical to the storage solution in Scenario A: Storage resource optimization.

This solution included a single vSphere cluster and two ESXi hosts within the cluster. We used multiple two-node RAC databases as the production databases with each virtual RAC node running on different ESXi hosts.

Test procedure

We used the following procedure to validate the high availability solution using vSphere HA and Oracle RAC:

1. Set up vSphere HA.

   The following are the required components for setting up vSphere HA:

   - VMware Infrastructure Suite Standard or VMware Infrastructure Enterprise
   - Two ESXi hosts (at a minimum)
   - SAN storage shared between the ESXi servers
   - CPUs compatible between the hosts. For more details, see vSphere High Availability Deployment Best Practices.

2. Configure vSphere HA to create a VMware HA cluster and move the ESXi hosts into the cluster.

3. Deploy two-node RAC databases on the HA cluster, with each node running on different ESXi hosts.

4. Shut down one of the ESXi hosts to simulate the host failure.

5. Check if the RAC nodes were migrated from the failed ESXi host to the surviving ESXi host and the Oracle RAC environment was returned to full capacity.

Testing results

To test this scenario, we restarted one of the ESXi hosts to simulate an ESXi host failure. The system of one RAC database responded to this failure, as shown in Figure 25.

```
[oracle@oralow2 ~]$ srvctl status database -d costa
Instance costa1 is running on node oralow1
Instance costa2 is running on node oralow2
[oracle@oralow2 ~]$ srvctl status database -d costa
Instance costa1 is not running on node oralow1
Instance costa2 is running on node oralow2
```

Figure 25. RAC node went offline after the ESXi server failure

In a physical environment, customers need to spend a lot of time to repair or rebuild a node or to add new node to a RAC database. However, in a virtualized environment, VMware HA detects the RAC node failure and automatically restarts the virtual RAC nodes on the surviving ESXi host in minutes. Figure 26 and Figure 27 show that the failed Oracle RAC nodes were automatically restarted on the surviving ESXi host.

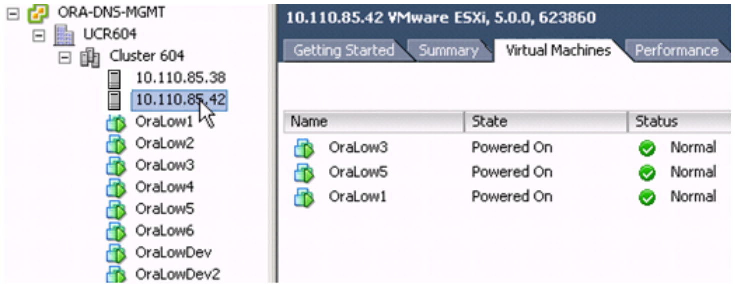

Figure 26. Before the ESXi server failure

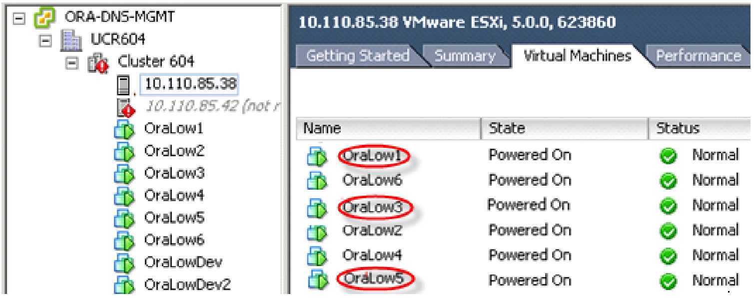

Figure 27. After the ESXi server failure

Figure 28 shows that the RAC database system responded to the restarted RAC node.

```
[oracle@oralow2 ~]$ srvctl status database -d costa
Instance costa1 is not running on node oralow1
Instance costa2 is running on node oralow2
[oracle@oralow2 ~]$ srvctl status database -d costa
Instance costa1 is running on node oralow1
Instance costa2 is running on node oralow2
```

Figure 28. RAC node restarted on another ESXi server

By leveraging VMware HA, a virtualized Oracle RAC environment can provide even higher availability than a physical Oracle RAC environment. A multinode Oracle RAC deployment is highly available. However, VMware HA coexists with and complements a virtualized Oracle RAC installation, which requires no administrator intervention during the server crash, in comparison with deploying Oracle RAC databases in a hardware environment.

Conclusion

Summary

This cost-efficient solution demonstrates how to optimize the Oracle environment to meet today's business needs. It shows how EMC, VMware, and Oracle technologies simplify a customer's daily tasks and processes in an integrated and cost-efficient deployment.

Along with VNX OE File release 7.1 and block release 5.32, EMC VNX5300 creates one of the most efficient storage platforms for Oracle. Deploying multiple virtualized Oracle databases on two-tier, dynamic FAST VP and FAST Cache can greatly reduce cost while improving performance.

By deploying the following technologies, the solution provides a balance between cost and performance, and at the same time, improves efficiency and resiliency.

EMC FAST VP and FAST Cache technology:

- Reduces the amount of time required for Oracle tuning and administration while delivering performance benefits for Oracle OLTP workloads
- Reduces the spindle count of the physical drives necessary to support the required performance

EMC VNX Snapshots technology:

- Enables Oracle DBAs to snap writable snaps (up to 256) and to create a new branch for general data reuse and testing purposes
- Provides point-in-time versioning of a writable snap that is consistent with the source database so that it can be used for effective and fast recovery

VMware virtualization technology:

- Shares resource across database servers by improving resource utilization and increasing infrastructure efficiency
- Reduces the cost of Oracle licensing by physically consolidating virtual machines running Oracle software to their own dedicated VMware environment

Findings

Solution key technologies

The key technologies of this solution are as follows:

- **VMware vSphere 5.0**–Enhances high availability and reduces Oracle licensing cost by isolating Oracle virtual machines physically to their own dedicated VMware HA cluster:
  - Oracle 11*g* SE databases are available on clustered servers within a maximum capacity of four sockets in total, including Oracle RAC feature with no additional cost.
  - Oracle EE is priced on a core basis, while Oracle SE is priced on a processor socket basis. This is very beneficial for customers who are keen on cost savings on Oracle software license.
- **FAST VP and FAST Cache technologies**–Provide automatic Oracle-to-Storage optimization.

  The performance of transactions per minute was improved by 86 percent by running more workloads in a multiple RAC databases environment.
- **VNX Snapshots**–Enable Oracle DBAs in an easy provisioning and protection process:
  - By using "Writable snap of a writable snap," multiple snaps of an Oracle production database are created in seconds for testing, development, quality assurance, or other purposes.
  - "Read-only snap of a writable snap" enables fast recovery of the test/dev database in minutes to correct the database logical corruption.
- **VMware HA**–A virtualized Oracle RAC environment that provides a higher availability and faster return-to-service ability than a physical Oracle RAC environment.

  An ESXi host failure in a virtual deployment results in an automatic restart of the failed Oracle RAC node on another available ESXi host.

Cost efficiency

This solution demonstrates that the capital expenditure (capex) can be reduced by using:

- FAST Suite to maximize the storage resources available, using less disk drives in a targeted manner. This ensures that the correct data is on the correct tier at all times.
- Oracle SE rather than Oracle EE with Oracle RAC option to reduce licensing costs.

The reduced operational expenditure (opex) is also demonstrated:

- Reduced power and cooling costs
- Reduced maintenance costs

Table 10 compares the cost-efficiency messages of one virtual RAC database and two virtual RAC databases.

Table 10. Comparison of cost-efficiency messages

Configuration scenario	Total users of workload	Total TPM of workload	Cost-efficiency messages
One virtual RAC database	400	19,999	Baseline
Two virtual RAC databases	800	37,298	• 86 percent increase in TPM because of the database added • Better utilization of hardware resources and increased operational efficiency, because of the consolidation of workloads • No extra software and hardware cost

For a detailed review on the cost-efficient use of Oracle and VMware, refer to the white paper: *Understanding Oracle Certification, Support and Licensing for VMware Environments*.

References

White papers

For additional information, see the white papers listed below.

- *Deploying Oracle Database Applications on EMC VNX Unified Storage*
- *Deploying Oracle Database on EMC VNX Unified Storage*
- *EMC CLARiiON, Celerra Unified, and VNX FAST Cache*
- *EMC VNX Unified Best Practices for Performance*
- *Leveraging EMC FAST Cache with Oracle OLTP Database Applications*
- *Virtual test/dev environment provisioning for Oracle RAC 11g with EMC unified storage*
- *Understanding Oracle Certification, Support and Licensing for VMware Environments* (VMware Website)

Other documentation

For additional information, see the following documents.

- *Oracle Grid Infrastructure Installation Guide 11g Release 2 (11.2) for Linux*
- *Oracle Database Installation Guide 11g Release 2 (11.2) for Linux*
- *Disabling simultaneous write protection provided by VMFS using the multi-writer flag* (VMware Knowledge Base)