White Paper

Using VMware VMotion with Oracle Database and EMC CLARiiON Storage Systems

Applied Technology

Abstract

By migrating VMware virtual machines from one physical environment to another, VMware® VMotion can help eliminate application downtime resulting from hardware maintenance. This white paper demonstrates how VMotion can be used with Oracle® Database and EMC® CLARiiON® storage systems, and how much impact it can have on the database's service level.

February 2008

Part Number H4205

Table of Contents

Executive summary **4**
Introduction **4**
Audience 5
Terminology 5
VMotion testing with Oracle Database **6**
Test environment 6
Test methodology 7
Scenario 1: Validate live database migration 8
Scenario 2: Database performance impact 10
VMotion requirements and best practices **12**
Conclusion **12**
References **13**

Executive summary

Data center downtime is generally categorized as either "planned" or "unplanned." The ratio, and therefore cost, of planned downtime to unplanned downtime is usually estimated as four to one. Among the various reasons for planned downtime, hardware maintenance is the most common. Regardless of the cause of downtime, when mission-critical applications are not accessible, the business experiences significant negative impact including loss of productivity, revenues, and customer satisfaction. Any effective solution that can eliminate or significantly reduce planned downtime offers real operational benefits for system administrators.

In a physical environment, IT departments usually use one of three common methods to minimize the impact of planned downtime.

- Schedule maintenance for a time window after regular business hours to minimize the impact to end users. Scheduling a maintenance time window is challenging to global organizations.
- Keep redundant hardware ready for planned maintenance. This solution allows downtime to be relatively short, but it increases cost and complexity with its requirement for redundant systems.
- Cluster servers together so the application workloads can fail over to the remaining servers when the affected server is shut down for maintenance. Although clustering significantly reduces downtime, it is the least common approach due to the cost and complexity of setting up clustering for applications. Furthermore, not all applications are cluster-aware and so they cannot use this approach.

This white paper explains how VMware® VMotion can help eliminate most of the downtime that accompanies planned hardware maintenance. Also included are VMotion test scenarios that VMware and EMC used to validate live migration of Oracle® Database application virtual machines from one VMware ESX Server host to another. The paper also contains information on the impact of migration with VMotion on user transactions and response times.

Introduction

VMware VMotion enables the live migration of running virtual machines from one physical server to another with zero downtime, continuous service availability, and complete transaction integrity. This capability makes hardware maintenance possible at any time of the day and does not require clustering or redundant servers. VMotion makes it possible to move online workloads as required from one ESX Server host machine to another in order to maintain service levels and performance goals.

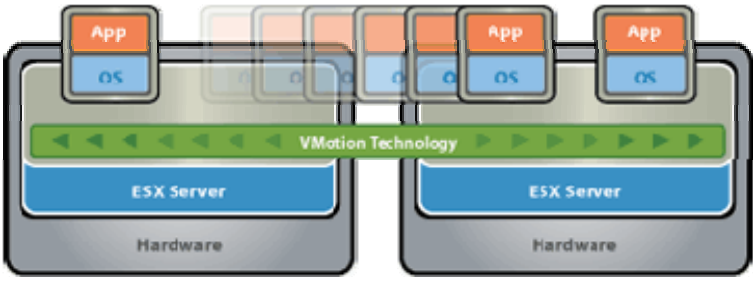

Figure 1. VMotion moving a live, running virtual machine from one host to another

VMotion technology's live migration process uses the hardware independence and encapsulation properties of virtual machines, leveraging shared storage to make it happen. The entire virtual machine is encapsulated by a set of files that reside on shared storage such as a storage area network (SAN).

The EMC® CLARiiON® CX3 UltraScale™ series of networked storage systems with its state-of-the-art I/O interconnect, based on native PCI Express technology within the new UltraScale architecture, delivers high bandwidth, low latency characteristics that enable the CX3 UltraScale series to deliver full 4 Gb/s

capabilities throughout the entire system. Customers benefit from advanced, industry-leading software to meet the most challenging availability and protection requirements. The UltraScale architecture enables seamless scalability from 365 GB to 353 TB, with the flexibility to mix and match high-performance and low-cost drives to meet different service levels in this tiered storage system. These EMC systems work in concert with VMware technology, providing the same high-performance shared storage availability to VMware virtual machines. For additional information on the CX3 UltraScale series, refer to the "References" section.

VMware Virtual Machine File System (VMFS) allows multiple ESX Server host machines to access the same virtual machine files on the storage system concurrently. As a result, the applications running in a virtual machine can remain active and uninterrupted as they are moved from one host to another. The active memory and precise execution states of the virtual machines rapidly transfer over a high-speed network connection, and VMotion technology ensures that the transfer period is imperceptible to users by recording ongoing memory transactions in a bitmap. The networks being used by the virtual machines are also virtualized by the underlying ESX Server software, ensuring that even after migration, virtual machine network identity and connections are preserved. In addition to shared storage for multiple hosts, a storage LUN can be divided into multiple datastores, to allow exclusive usage of storage space for individual virtual machines. Since the migration of a virtual machine with VMotion preserves active memory state, network identity, and access to the storage system, the result is zero downtime and no disruption to users.

Audience

This white paper is intended for any EMC customers, partners, or employees who want more information on how VMware Vmotion can be used with Oracle Database and the CLARiiON storage system

Terminology

Bandwidth: The average amount of read/write data in megabytes that is passed through the storage system per second.

Logical unit number (LUN): A unique identifier that is used to distinguish among logical storage objects in a storage system.

Online transaction processing (OLTP): A type of processing in which the computer responds immediately to user requests. Each request is considered to be a transaction.

Oracle Automated Stress Testing (OAST): An automated test suite designed to build OLTP type workloads for systems using Oracle Database. It creates tables, performs stress test runs, and outputs transaction-related performance data.

Oracle Automatic Storage Management (ASM): Storage management tool that provides cluster file system and volume management capabilities integrated into the Oracle Database. It performs automatic redistribution of storage capacity after the incremental addition or removal of storage disks.

Redundant Array of Independent Disks (RAID): RAID technology groups separate disks into one logical unit number (LUN). Data is distributed across these disks to improve reliability and performance. The CX3 series supports RAID level 0, 1, 1/0, 3, and 5.

Response time: The average time in milliseconds that it takes for one request to pass through the storage, including any waiting time.

Throughput: The average number of read/write requests in I/Os that are passed through storage per second.

Transactions per minute (TPM): The measured number of transactions completed by an OLTP database over a 1-minute interval.

VMotion: A VMware product that can move running virtual machines from one physical server to another with no impact to end users. With VMotion, users can perform live migrations of virtual machines with zero downtime.

VMotion testing with Oracle Database

This section describes the migration with VMotion of a live Oracle Database running in a virtual machine and includes data on the effect on response times and user transactions.

Test environment

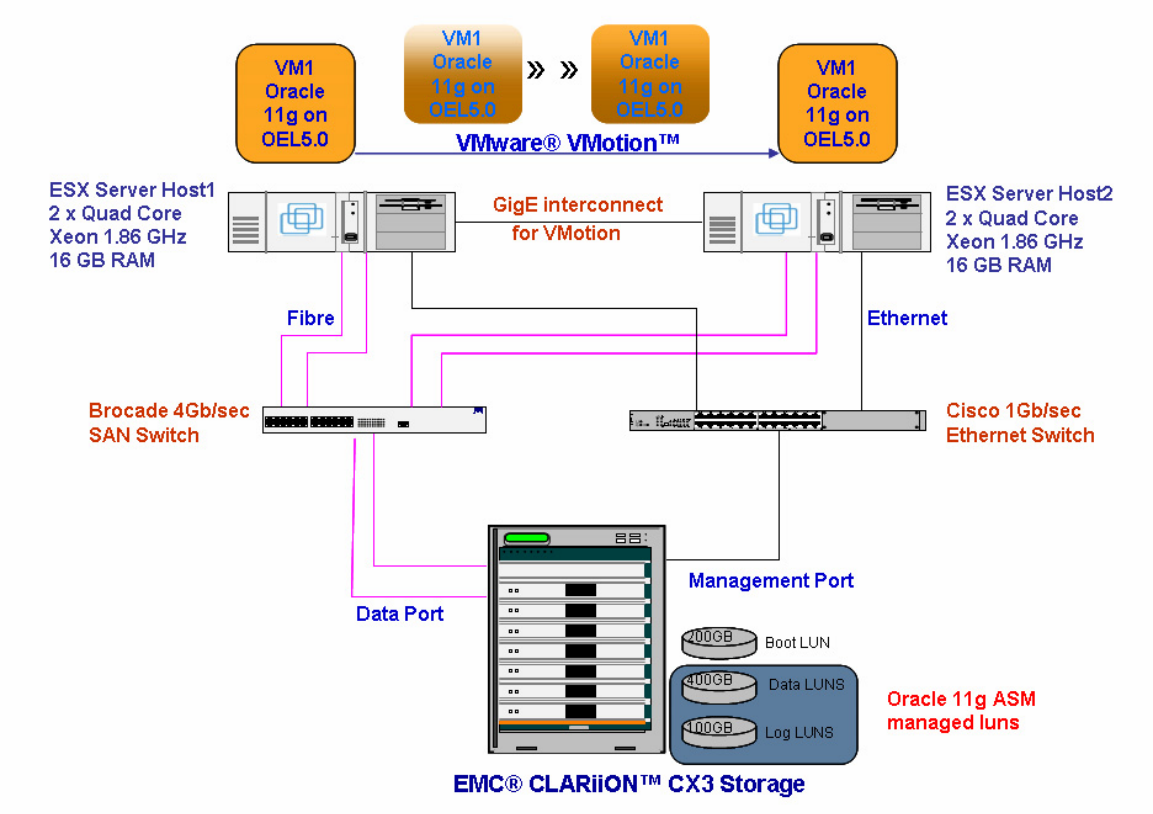

Figure 2. Test environment setup diagram

Server machines: DELL 2950

OS	VMware ESX Server 3.0.2
CPU	2 x Quad Core Xeon 1.86 GHz
Memory	16 GB

Storage system: EMC CLARiiON CX3-20f

Processors	2
Memory size	2 GB per SP
Number of disks	30 FC 68 GB @ 15k rpm
Base software	03.26.020.3.060
LUN configuration	200 GB RAID 5 MetaLUN for virtual machine OS (Boot)
	4 x 100 GB RAID 5 LUNs for Oracle Database datafiles (Data)
	2 x 50 GB RAID 5 LUNs for Oracle Database logfiles (Log)

Virtual Machine (VM1):

OS	Oracle Enterprise Linux (OEL) 5 (2.6.18-8.el5)
CPU	4 vCPU
Memory	4 GB

Database software: Oracle 11.1.0

OAST number of warehouses	100
OAST number of users	50
OAST runtime in seconds	3600

Test methodology

The exercise was developed to validate several critical points:

- Validate that Oracle Database functions as promised and that users can continue to perform transactions without interruption during migration.
- Understand what impact to database performance (transactions, response time, and so on) might occur during migration. Considering the complexities of database operations, migration with VMotion raises several interesting questions:
  - How much is the migration impact in terms of database transaction numbers and response times?
  - Will behavior be consistent if it's over frequent virtual machines migrations? In other words, would multiple migrations really affect resources differently over time or would the application (or virtual machine) performance degrade due to multiple migrations?
  - Finally, there are always concerns about resiliency of the application (database) and infrastructure when multiple migrations are performed in a short time. Would the application crash if moved across machines too many times?

To find the answers, testers created an OEL virtual machine with four virtual CPUs and 4 GB of virtual memory. Oracle Database was installed on ASM managed LUNs inside an EMC CLARiiON CX3 UltraScale storage system. The LUNs were exposed to both ESX Server hosts, as part of the Storage Group in the CLARiiON storage system. After database installation, an OAST stress testing workload was run with VMotion migration executed once, and was also scheduled every five minutes (12 migrations per hour). This test would address the questions above. Other considerations for the test approach included:

- It was very important that the workload put enough stress on virtual machine resources (> 70% vCPU utilization) so that VMotion would be validated against extreme real-world conditions.
- Oracle Database statspack reports were captured for runs with and without VMotion. The goal was not database performance tuning but to ensure that additional performance issues (and wait events) did not surface during migrations.

Most of the monitoring and operational tasks in this exercise were carried out through the VMware VirtualCenter Management Server that can run on any host system with a network connection to the ESX Server hosts and the CLARiiON storage system. VirtualCenter delivers centralized management, operational automation, resource optimization, and high availability.

To address the previously mentioned points, the exercise is divided into two scenarios:

- Validate live database migration with VMotion
- Study resource impact and validate behavior consistency and resiliency

Scenario 1: Validate live database migration

For the validation test, testers initiated the OAST run with 50 users to ensure that enough load was generated on the virtual machines. The screen capture from VirtualCenter shown in Figure 3 shows overall ESX Server host CPU utilization at 50 percent and virtual machine CPU utilization at approximately 98 percent.

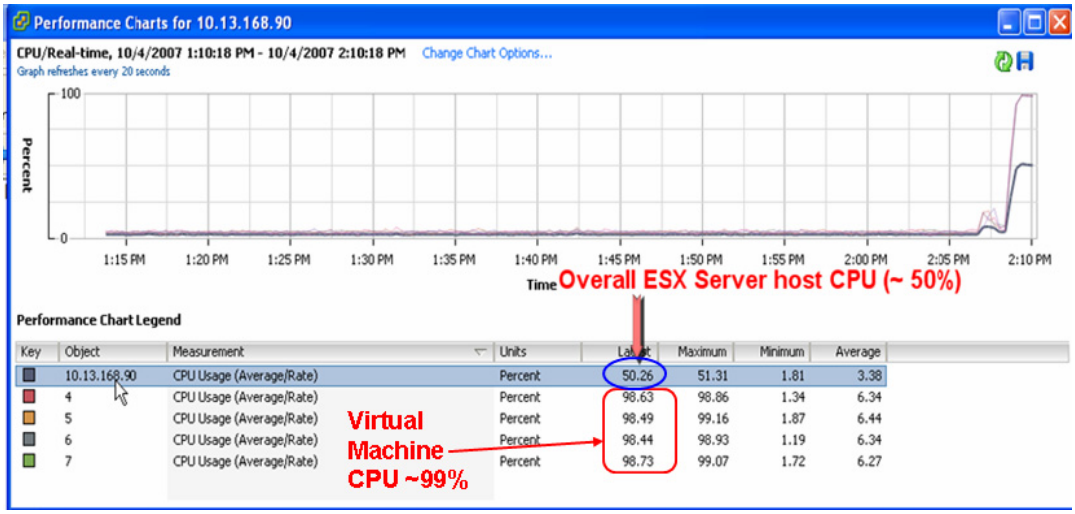

Figure 3. ESX Server and virtual machine CPU usage rates during OAST run

Once the users ramped up, testers manually initiated the database virtual machine migration from VirtualCenter. Figure 4 shows the interface of the VirtualCenter management console from which the VMotion migration test was initiated.

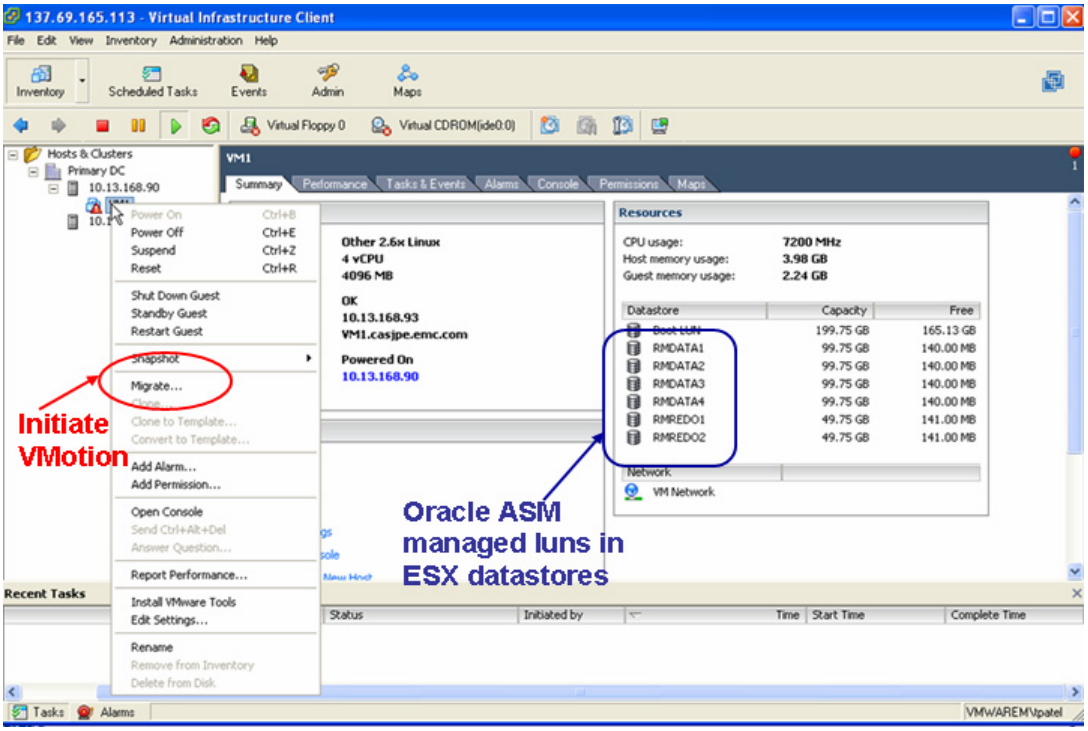

Figure 4. VMotion initiation from the VirtualCenter management console

The VirtualCenter management console facilitates centralized management of ESX Server hosts and virtual machines. Along with standard management and provisioning functions, VirtualCenter also monitors system availability and performance. Figure 4 shows virtual machine resources details, virtual machine state, and the datastores where Oracle ASM-managed LUNs reside.

Within about two minutes, the database virtual machine migration process completed successfully. Figure 5 provides a view of both ESX Server hosts involved in migration exercise.

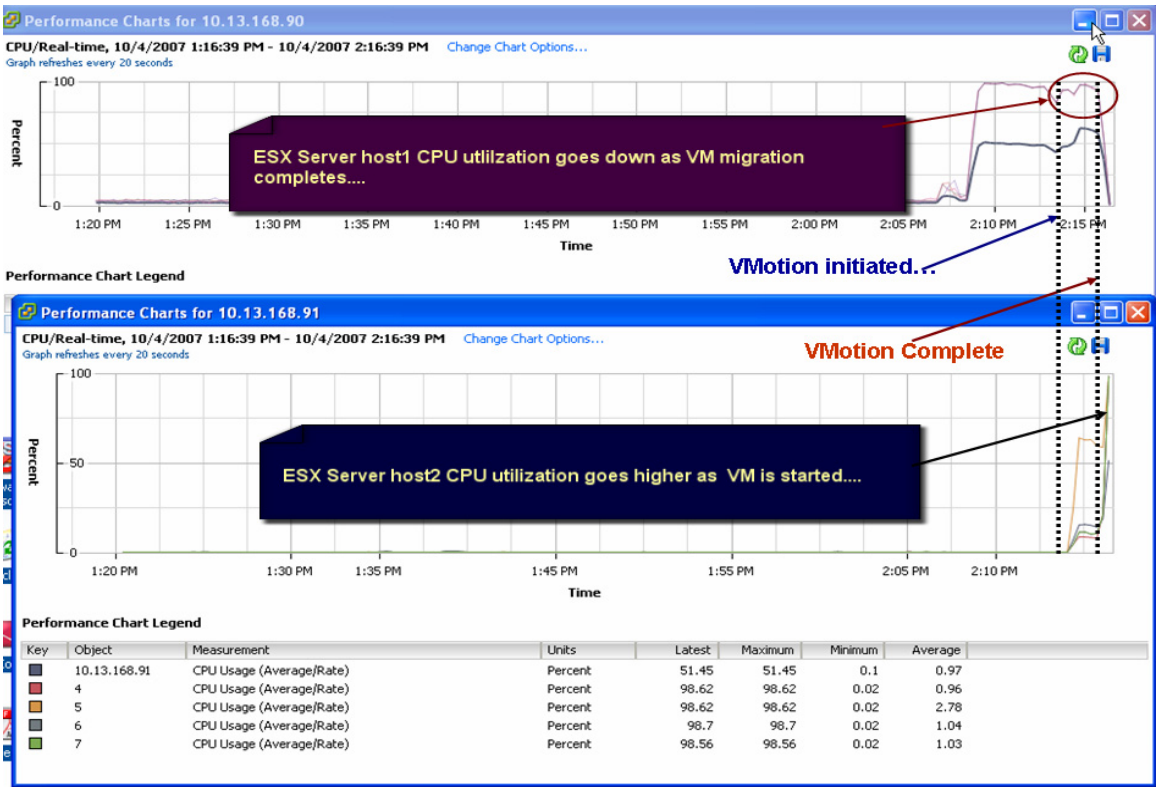

Figure 5. CPU usage rates of ESX Server host1 (top) and ESX Server host2 (bottom) during VMotion migration

A few observations to note:

- The CPU utilization on ESX Server host1 went marginally higher as the VMotion process initiated memory transfer across the ESX Server hosts.
- A 4 GB RAM VMotion process completed in a couple of minutes and user transactions continued on ESX Server host2 as the virtual machine migrated from ESX Server host1 to ESX Server host2.
- The VMotion migration time depends upon "VMotion interconnect" configuration, speed, and memory changes to some extent.
- At the end of the run, the OAST workload generated a detailed statistics report. The report file indicated that the transactions per minute (TPM) were reduced and response time went higher during VMotion. However, the testers did not notice any timeouts or freeze in transactions. Figure 6 shows the OAST report on the VMotion migration exercise.

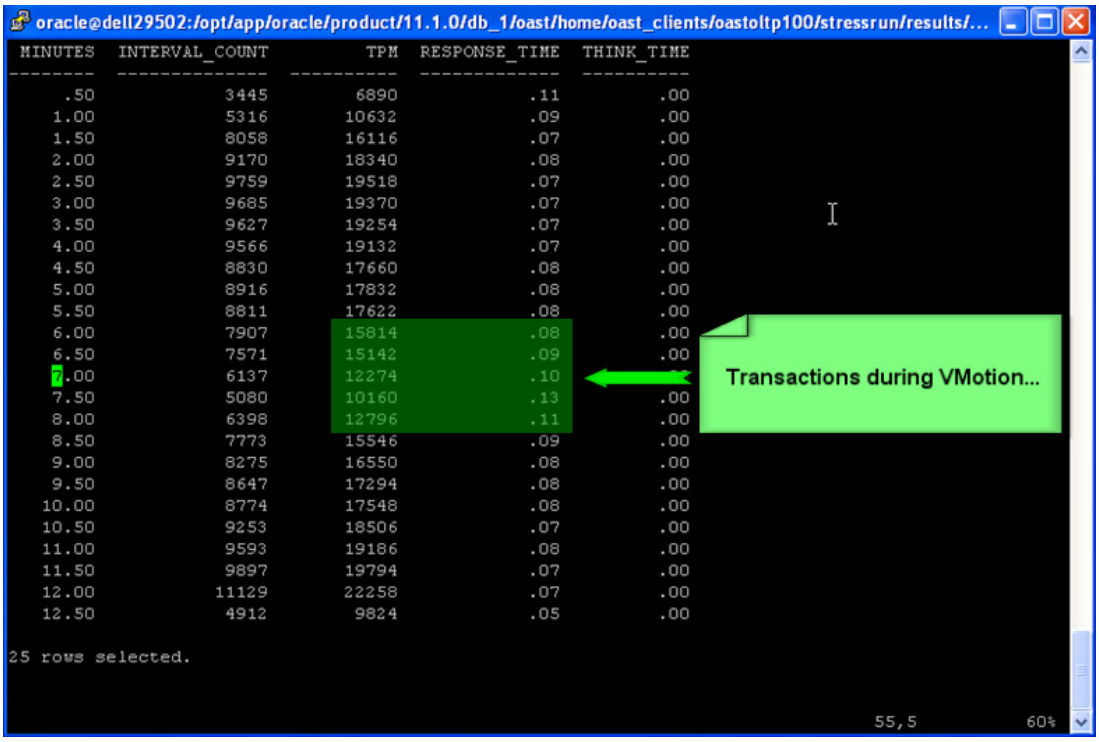

Figure 6. OAST report from the Oracle Database client during VMotion

Scenario 2: Database performance impact

This scenario addresses VMotion resource impact, behavior consistency (or degradation over time) in relation to multiple migrations, and resiliency. As the first step, testers completed a baseline one-hour OAST run with no VMotion migrations. Next was a one-hour OAST run with VMotion migrations initiated every five minutes. These test runs provide a basis for comparison to address the issues mentioned in the Test Methodology.

The following figures compare transactions per minute and response times for the two runs: baseline (no migrations) and migrations with VMotion every five minutes.

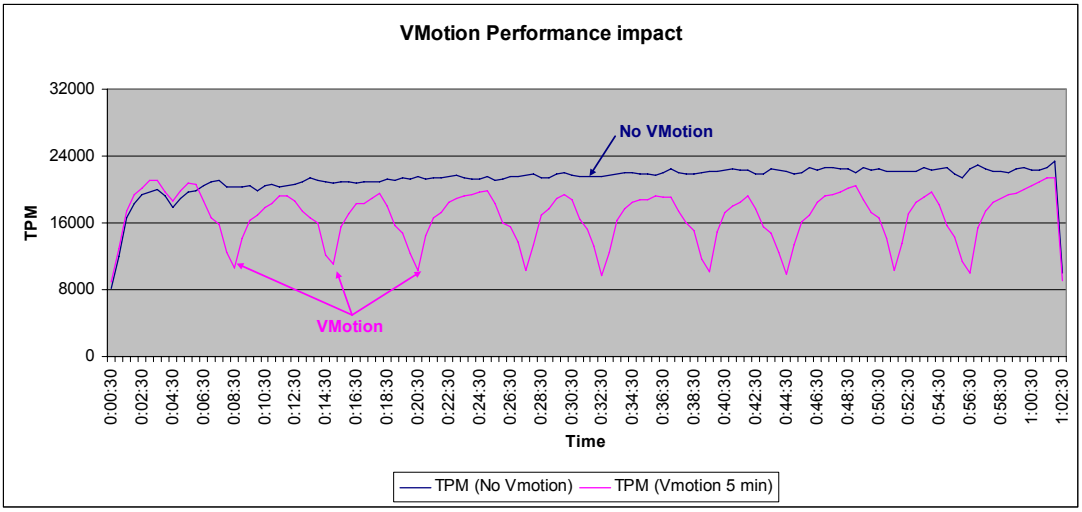

Figure 7. TPM comparison with no VMotion and with VMotion occurring every five minutes

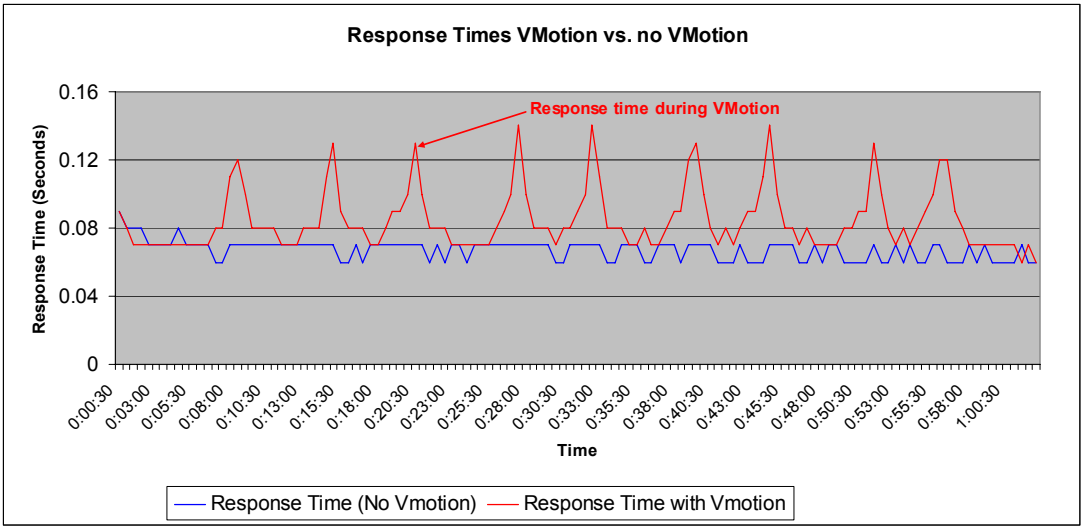

Figure 8. Response time comparison with no VMotion and with VMotion occuring every five minutes

TPMs and response times were affected for a brief period while migration was in progress. TPM numbers went down and response times became longer during VMotion migration. Multiple virtual machine (database) migrations were handled extremely well and did not have any adverse effects on database behavior.

Figure 9 shows how the storage LUNs used by the virtual machine performed during the test.

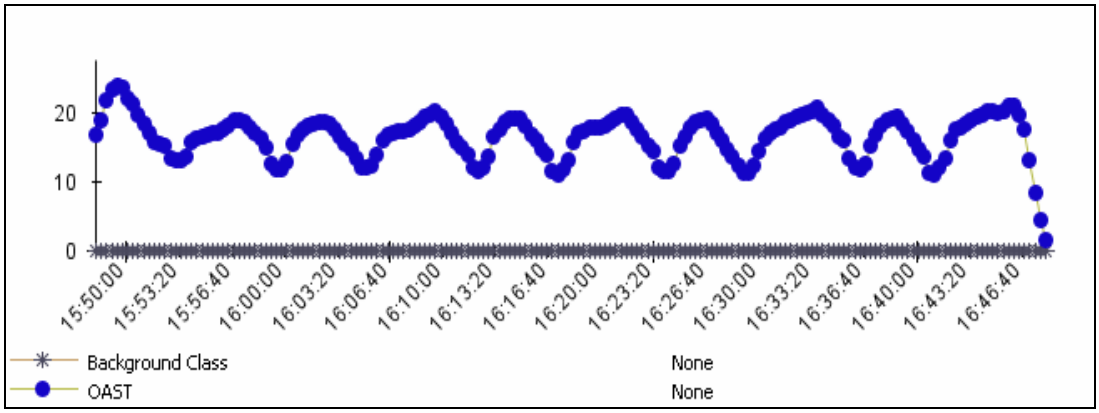

Figure 9. Bandwidth (MB/s) of LUNs used by the virtual machine

The bandwidth received from the virtual machine's database into storage indicates the same kind of behavior as the TPM graph (Figure 7) during VMotion migration. Although the bandwidth decreased when migration was in progress, the storage was continuously utilized as the virtual machine migrated back and forth between the two ESX Server hosts.

The storage utilization, work completed (TPM), and response times affected by migration returned to normal (state prior to migration) after each migration. In other words, we did not observe any overall database performance (or virtual machine) degradation due to multiple migrations. This result should encourage IT administrators facing challenges to minimize downtime.

VMotion requirements and best practices

VMware recommends that users of VMotion adhere to these requirements and best practice recommendations.

- All ESX Server host hardware, in particular the CPU, must be compatible.
- All virtual switches must be configured in the same way for both ESX Server hosts.
- Use a separate, non-routable subnet for all VMotion traffic. Unless management traffic and hot migration events are infrequent, have dedicated NICs for the Service Console and VMotion.
- Run a private Gigabit Ethernet migration network between all VMotion-enabled managed hosts.
- In order to connect VMware ESX Server hosts to the CLARiiON storage system, install the NaviAgentCLI program on the host and give firewall access before starting the service.
- In case a LUN has to be shared between multiple VMs, set DisallowSnapshotLUN value to 0 in VirtualCenter

Conclusion

VMware VMotion helps administrators keep their IT environments running by giving them unprecedented flexibility and availability to meet the increasing demands of today's businesses and end users. VMotion allows users to:

- Perform hardware maintenance without scheduled downtime.
- Automatically optimize and allocate entire pools of resources for maximum capacity utilization with flexibility and availability.
- Proactively move virtual machines away from failing or underperforming servers.

VMotion does have an impact on host system resources. Therefore, careful planning is necessary to minimize the downtime impact. In this study, testers built an extreme scenario by using a real workload to generate heavy demand on the database running inside a virtual machine. The results demonstrate that a very busy virtual machine performing more than 300 database transactions per second can be moved from one physical server to another in a couple of minutes without stopping the database or losing any transactions. The tests revealed that the transaction response times and transactions per minute did suffer a slight impact but there were no service interruptions or timeouts. Multiple migrations of a very busy virtual machine in a short time verified application (database) behavior consistency as well as storage and application resiliency.

The tests also proved that CLARiiON storage systems can provide secure and stable access of VMFS formatted storage disks required by VMware. There was no overhead at the storage level when I/Os were requested by virtual machines, and the storage system was instrumental in providing storage access to multiple hosts, so that VMotion can be done without any storage reconfiguration.

In summary, infrastructure downtime is inevitable in today's complex IT environments because it is impossible to plan for every scenario. However, planned downtime can be significantly reduced with effective planning and use of technology such as VMotion. This study clearly demonstrates that VMware VMotion coupled with the CLARiiON storage system provides a powerful way to eliminate or greatly reduce planned downtime at a lower cost than traditional approaches.

References

On VMware.com:

- VMware VMotion product data sheet

http://www.vmware.com/pdf/vmotion_datasheet.pdf

- VMware and Oracle database solutions page

http://www.vmware.com/partners/alliances/technology/oracle.html

- VMware Server Virtualization Infrastructure page

http://www.vmware.com/products/data_center.html

- VMware DRS product page

http://www.vmware.com/products/vi/vc/drs.html

On EMC.com:

- Solutions for Oracle page

http://www.emc.com/solutions/application-environment/oracle/index.htm

- EMC CLARiiON Family page

http://www.emc.com/products/family/clariion-family.htm

- EMC CLARiiON CX3 UltraScale Series data sheet

http://www.emc.com/collateral/hardware/data-sheet/c1148-clariion-cx3-ds.pdf