Virtualized Oracle Database Study Using Cisco Unified Computing System and EMC Unified Storage

Technology Validation Report with Scale-Up, Scale-Out and Live Migration Tests

White Paper

January 2011

Contents

1. Executive Summary

Oracle database deployments can generate significant server sprawl due to the need to provision separate systems for development, quality assurance and production environments. As a result, Oracle database implementations can increase exponentially, and even small implementations can have a relatively large IT footprint. In traditional data center deployment, the databases are typically hosted on dedicated physical systems that are not fully utilized much of the time. This traditional "one workload, one box" approach to server provisioning inevitably leads to over-provisioning and underutilization of hardware assets. This is especially true with servers getting extremely powerful by improved compute capacity. This typically means that dedicated server deployments for departmental or branch office operations causes most servers operate at only about 5%-15% of their total load capacity. VMware vSphere enables the deployment of multiple virtual machines running different Oracle database solutions on the same physical server, resulting in increased CPU utilization and reduction in overall server requirements.

This paper highlights several scenarios of virtualized Oracle databases (version 11g Release 2) running on VMware ESXi 4.1 and Cisco Unified Computing system (UCS). These scenarios represent following situations that are commonly encountered in today's data centers.

- Scale up—Scaling up a virtual machine is the process of adding virtual CPUs (vCPUs) to increase a virtual machine's compute power. With VMware vSphere, a virtual machine can easily take advantage of multi-core processors with no special settings or configuration besides building the virtual machine with more vCPUs.
- Scale out—Scale out is the process of adding more virtual machines on the host which enables consolidation of multiple virtual machines, each running a single Oracle database, on a single ESXi host.
- Live migration via VMware vMotion™ - Migrate a busy virtual machine performing database transactions to another host.

The study used Swingbench kit for OLTP workloads; The Swingbench kit is comprised of scripts and load drivers that generate business transactions. These transactions simulate large scale order entry OLTP loads that are I/O-intensive. Test results conclude that virtualized Oracle databases on vSphere could handle intensive OLTP loads in single and consolidated multi-virtual machine scenarios. The key findings of this study are summarized below:

- A single instance Oracle database scaled effectively the virtual machine size increased in terms of vCPU and memory. The Transactions per Minute (TPM) increased in a near linear fashion as more vCPUs (2 vCPUs to 8 vCPUs) and memory (4 GB to 32 GB) were added.
- A single ESX server could effectively balance the Oracle database workload for each virtual machine. The results were *not* achieved by over-allocating any system resources (CPU and memory). The total number of vCPUs and total memory consumed by the virtual machines was less than or equal to the physical resources on the ESX host machine.
- When running an 8-vCPU virtual machine (which in the test environment here has more vCPUs than available cores on a NUMA node) ESX split the memory across the NUMA nodes. (See Appendix B.)
- A live migration of a virtual machine running an Oracle database OLTP load, using VMware vMotion™, experienced zero transaction errors and no data loss, though a temporary dip in throughput was observed during the migration. The minimal migration time of a very busy virtual machine reflected the superior architecture of Cisco UCS and underlying lossless and low latency 10GE network architecture.
- Dynamic Hot Add of vCPU helps in dynamically scaling the processing resources, improving overall performance and throughput of a saturated virtual machine running an Oracle Database Server processing an OLTP load and can be accomplished with zero down time.

  Note: This feature is available only on SuSE Linux Operating system.

2. Database Virtualization with VMware vSphere and Cisco Unified Computing System

This joint solution is architected with following key technologies from VMware, Cisco and EMC.

VMware vSphere

VMware vSphere is the industry's most trusted virtualization platform, transforming datacenters into a dramatically simplified cloud infrastructure and enabling the next generation of flexible, reliable IT services. VMware vSphere virtualizes and aggregates the underlying physical hardware resources across multiple systems and provides pools of virtual resources to the datacenter. Oracle databases are excellent candidates to take advantage of these features and they have been shown to run effectively on vSphere, providing significant scalability, availability and performance benefits. Virtualizing database workloads on vSphere can significantly reduce the number of physical systems your organization requires while achieving more effective utilization of datacenter resources. Customers realize tangible savings from this consolidation along with operational cost savings from reduced datacenter floor space, power, and cooling requirements.

VMware Virtual Symmetric Multi-Processing (SMP) enhances virtual machine performance by enabling a single virtual machine to use multiple physical processors simultaneously. Virtual SMP enables virtualization of processor and resource-intensive enterprise applications such as databases and ERP. It moves processing tasks between available processors to balance the workload and efficiently utilize all processing power.

VMware vMotion technology enables an entire running virtual machine to be instantaneously moved from one server to another. This migration is live, meaning that the guest OS and the Oracle database continue to run without disruption of any batch or online activity. The entire state of a virtual machine is encapsulated by a set of files stored on shared storage, and the VMware VMFS cluster file system allows both the source and the target ESX host to access these virtual machine files concurrently. The active memory and precise execution state of a virtual machine can then be rapidly transmitted over a high speed network. Because the network is also virtualized by ESX, the virtual machine retains its network identity and connections, facilitating a seamless migration process.

The VMware Hot-add vCPU feature enables CPU capacity to be added to virtual machines running Oracle databases when needed without disruption or downtime.

Figure 1. Cisco Unified Computing System

The processor, memory, and I/O are three of the most important subsystems in a server. All are critical to the workload performance and reliability of a server. At any given point, one of these subsystems tends to become a performance bottleneck. When an application is described as CPU, memory, or I/O bound, that subsystem is likely to create a bottleneck for that application. The Cisco® Unified Computing System is a next-generation data center platform that unites compute, network, storage access, and virtualization into a cohesive system designed to reduce total cost of ownership (TCO) and increase business agility. The system integrates a low latency, lossless 10 Gigabit Ethernet unified network fabric with enterprise-class, x86-architecture servers. The system is an integrated, scalable, multi-chassis platform in which all resources participate in a unified management domain. The Cisco UCS B250 M1 features Cisco's patented Extended Memory Technology. This Cisco technology provides more than twice as much industry-standard memory (384 GB) as traditional two-socket servers, increasing performance and capacity for demanding virtualization and large-data-set workloads. Alternatively, this technology offers a more cost-effective memory footprint for less-demanding workloads. In summary, the Cisco UCS B250 M2 Extended Memory Blade Server increases performance and capacity for demanding virtualization and large-data-set workloads. The server is a full-width, two-socket blade server with substantial throughput and more than twice the memory capacity of other Intel Xeon 5600 series–based two-socket servers.

The test cases in this paper used Cisco UCS M81KR Virtual Interface Card in the UCS. A Cisco innovation, the Cisco UCS M81KR Virtual Interface Card (VIC) is a virtualization-optimized Ethernet and Fibre Channel over Ethernet (FCoE) mezzanine adapter designed for use with Cisco UCS B-Series Blade Servers. The VIC is a dual-port 10 Gigabit Ethernet mezzanine card that supports standards-compliant virtual interfaces that can be dynamically configured so that both their interface type (network interface card [NIC] or HBA) and identity (MAC address and worldwide name [WWN]) are established using just-in-time provisioning. The Cisco M81KR VIC is a fully standards-compliant Fibre Channel adapter that delivers cutting-edge storage I/O operations per second (IOPS) and throughput performance.

EMC® Unified Storage

EMC Unified storage delivers industry-leading innovation in a midrange storage solution. The unique combination of flexible, scalable hardware design and advanced software capabilities enable CLARiiON storage systems to meet the diverse and growing needs of today’s enterprises with unique highly available data integrity features.

The EMC CLARiiON CX4 series (Figure 2) with UltraFlex™ technology is based on a new breakthrough architecture and extensive technological innovation, providing a midrange solution that is highly scalable, meeting the price points of most midrange customers. The unique modularity of the UltraFlex™ technology allows you to use a combination of protocols within a single storage system, providing online-expandable connectivity options. It also includes new levels of ease of use, making the CX4 easy to install, manage, and scale. The CX4 is the fourth-generation CX series, and continues EMC’s commitment to maximizing customer’s investments in CLARiiON technology by ensuring that existing resources and capital assets are optimally utilized as customers adopt new technology. The innovative technologies in the CX4 includes fully automated storage tiering and support for the latest generation of disk drive technologies, such as Flash Drives.

The CLARiiON CX4 series introduces thin LUN technology that builds on CLARiiON virtual LUN capabilities and seamlessly integrates with CLARiiON management and replication software. With CLARiiON Virtual Provisioning™, you can choose between traditional LUNs, metaLUNs, and thin LUNs. The ability to non-disruptively migrate data to different LUN and disk types allows you to deploy the best solution without incurring downtime. Virtual Provisioning enables organizations to reduce costs by increasing utilization without over provisioning of storage capacity, simplifying storage management, and reducing application downtime.

Figure 2. EMC CLARiiON CX4 Series

CX4-120

- Up to 120 drives
- 6 GB cache
- Standard 4 Fibre Channel/4 iSCSI
- Maximum 16 front-end Fibre Channel and/or iSCSI

CX4-240

- Up to 240 drives
- 8 GB cache
- Standard 4 Fibre Channel/4 iSCSI
- Maximum 20 front-end Fibre Channel and/or iSCSI

CX4-480

- Up to 480 drives
- 16 GB cache
- Standard 8 Fibre Channel/4 iSCSI
- Maximum 24 front-end Fibre Channel and/or iSCSI
- Flash drives

CX4-960

- Up to 960 drives
- 32 GB cache
- Standard 8 Fibre Channel/4 iSCSI
- Maximum 32 front-end Fibre Channel and/or iSCSI
- Flash drives

3. Test Methodology

Swingbench Order Entry Benchmark Kit

The SwingBench kit is comprised of scripts and load drivers that generate business transactions which simulate large scale order entry OLTP loads that are I/O-intensive. It is a TPC-C like workload generator that includes a data generator tool which was used to create larger schemas that generate much higher levels of I/O (larger index lookups). Data was loaded to grow the database size to 1TB. This workload used an approximate read/write ratio of 60/40.

More information about the benchmark kit is available at http://www.dominicgiles.com/largesoe.html.

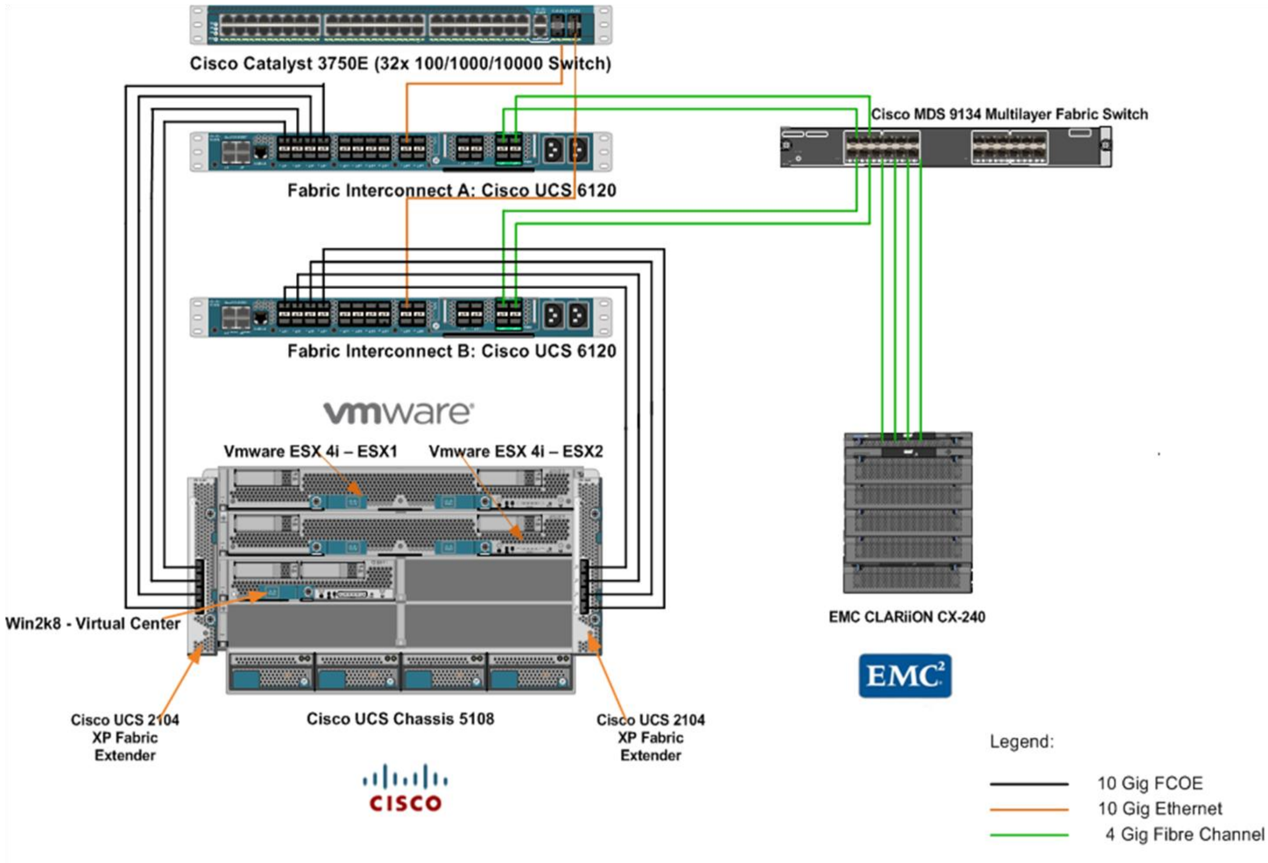
Test Setup and Infrastructure

Error! Reference source not found. summarizes the hardware and software used in the tests.

Table 1. Hardware and Software Used for Workload Characterization Tests

Hardware	2 x Cisco UCS B-250 blade server: 2 socket , 6 cores each, Intel Xeon X5670 @ 2.93GHz 192 GB RAM (running ESXi 4.1) One ESXi server is used host database server virtual machines, the other is used for the vMotion scenario and also runs Swingbench clients in virtual machines (one Swingbench client for each database server) 1 x B200 blade server running Windows 2008 and Virtual Center Network: 10GE Ethernet Shared storage: EMC CLARiiON CX4-240 (Flare 30) with 75 Fibre channel 15K RPM spindles
VMware Software	VMware Hypervisor ESXi 4.1 vCenter Server 4.1 vCenter Client 4.1 vMA – Virtual Management Appliance
Guest OS and Oracle Database Software in VM	Oracle Enterprise Linux 5.4 x86_64 Oracle 11gR2 Enterprise Edition Single Instance x86_64
Oracle OLTP Workload	Oracle Swingbench Load Generator – Large Scale Order Entry Benchmark

Figure 3. Physical Architecture



Summary of Test Scenarios

Table 2. Summary of Test Scenarios

Scenario	Virtual Machine (vCPU, RAM)	Test Description
Scaling up single Oracle database virtual machine (Hyper Threading OFF)	2vCPU/8GB 4vCPU/16GB 6vCPU/24GB 8vCPU/32GB	• Run Swingbench Large Scale Oracle Entry Benchmark user load against each virtual machine configuration • 4 separate test runs • 300 concurrent users • Min and max think time = 5 and 10ms respectively
Scaling out and consolidating multiple virtual machines running an Oracle Database (Hyper Threading ON)	4vCPU/16GB (1 – 5 virtual machines)	• Run Swingbench Large Scale Oracle Entry Benchmark user load against each virtual machine configuration • 5 separate test runs • 200 concurrent users per virtual machine • Min and max think time = 50 and 100ms respectively
Live migration of Oracle database in a virtual machine (vMotion)	4vCPU/16GB	• Run Swingbench Large Scale Oracle Entry Benchmark user load • 300 concurrent users • Min and max think time = 5 and 10ms respectively
VMware hot add CPU of Oracle database in a virtual machine	2vCPU/8GB (Hot Add 2 vCPU)	• Run Swingbench Large Scale Oracle Entry Benchmark user load • 100 concurrent users • Min and max think time = 5 and 10ms respectively

4. Test Scenarios and Results

These sections describe the setup and results for all the test scenarios. Performance results were primarily obtained from:

- Performance data captured by the VMware “esxtop” utility.
- Output report from the Oracle Swingbench generator. The report file summarizes the transaction throughput measured in Transactions per Minute (TPM) for the duration of the run.

Single Virtual Machine – Scale Up Test

The goal for the scale up test is validate the performance increase as the virtual machine size grows in terms of vCPU and memory. We identified and ran four independent test runs on virtual machines configured with 2, 4, 6 and 8 vCPUs along with memory configuration listed in the table below. The test details are described in **Error! Reference source not found.** and Single Virtual Machine–Scale Up Test below. (See Appendix A for virtual machine configuration details and properties).

Table 3. Single Virtual Machine - Test Configurations

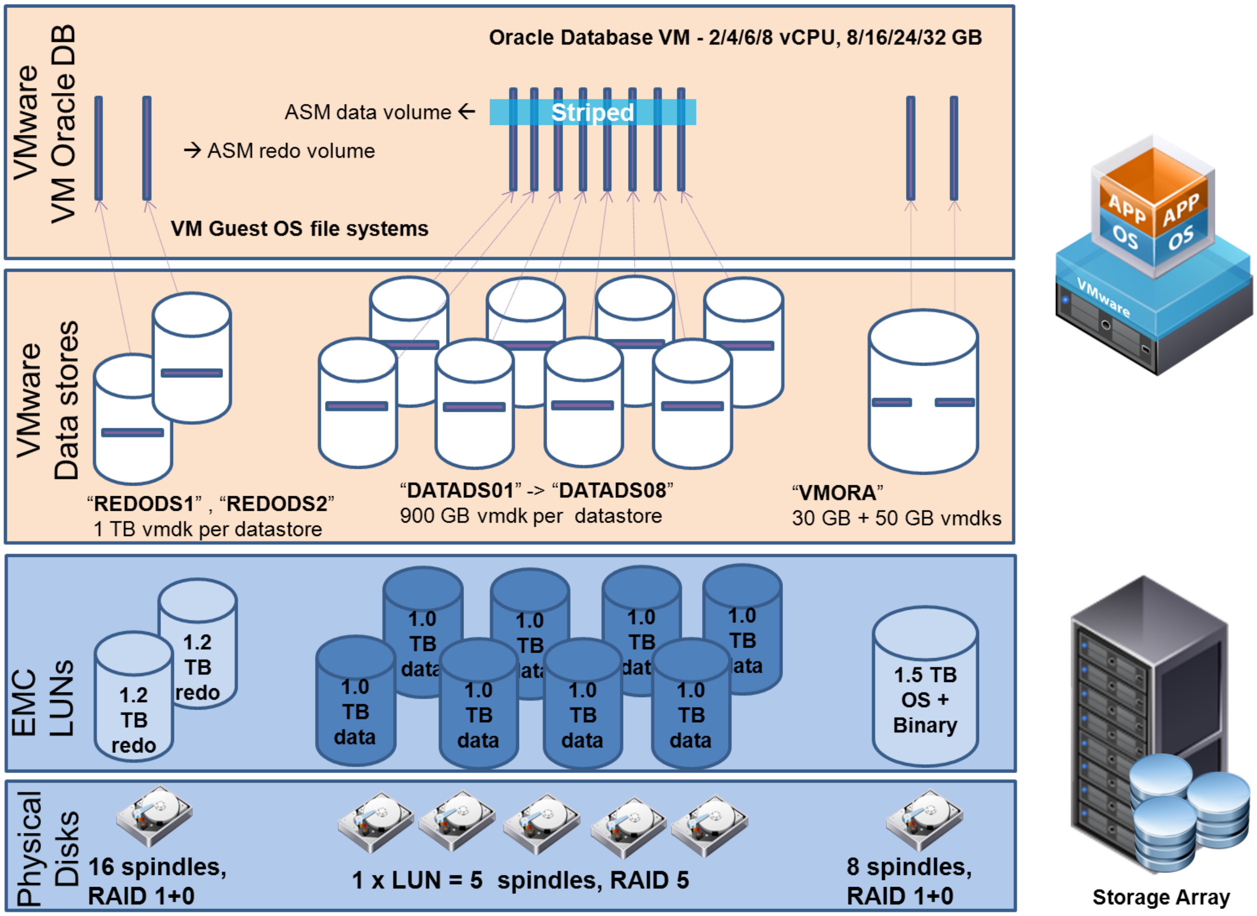
No. of vCPUs	Database Size (TB)	Concurrent Users (Min, Max think time - ms)	Virtual Machine Memory	Oracle Server SGA Size
2	1.1	300 (5,10)	8GB	6GB
4	1.1	300 (5,10)	16GB	10GB
6	1.1	300 (5,10)	24GB	16GB
8	1.1	300 (5,10)	32GB	20GB

Figure 4. Single Virtual Machine–Scale Up Test

Storage Layout for Single Virtual Machine

Figure 5 represents the storage layout of the Oracle database.

Figure 5. Single Virtual Machine – Storage Layout



For each database virtual machine, we used 8 data LUNs, 2 Redo Disks that were managed by Oracle ASM. Starting with Physical disks on EMC CLARiiON storage, we carved out:

- 8 RAID-5 (4+1) LUNs with each LUN of 1 Terabyte in size and consisting of 5 individual spindles. These LUNs will used for Oracle data files.
- 2 RAID 1/0 LUNs with each LUN of 1.5 Terabyte and consisting of 8 spindles. These LUNs will be used for Oracle redo logs.
- 1 RAID 1/0 LUN consisting of 8 spindles. This LUN will have the OS (VM SAN Boot) and database binaries.
- Once these LUNs are visible to ESX host, VMFS partitions and vmdks (virtual disks) are created.
- Oracle ASM will view each vmdk as individual disks. We created 2 ASM disk groups as DATADG(8 vmdk) and REDODG(2 vmdk).

The following is the summary of storage configuration:

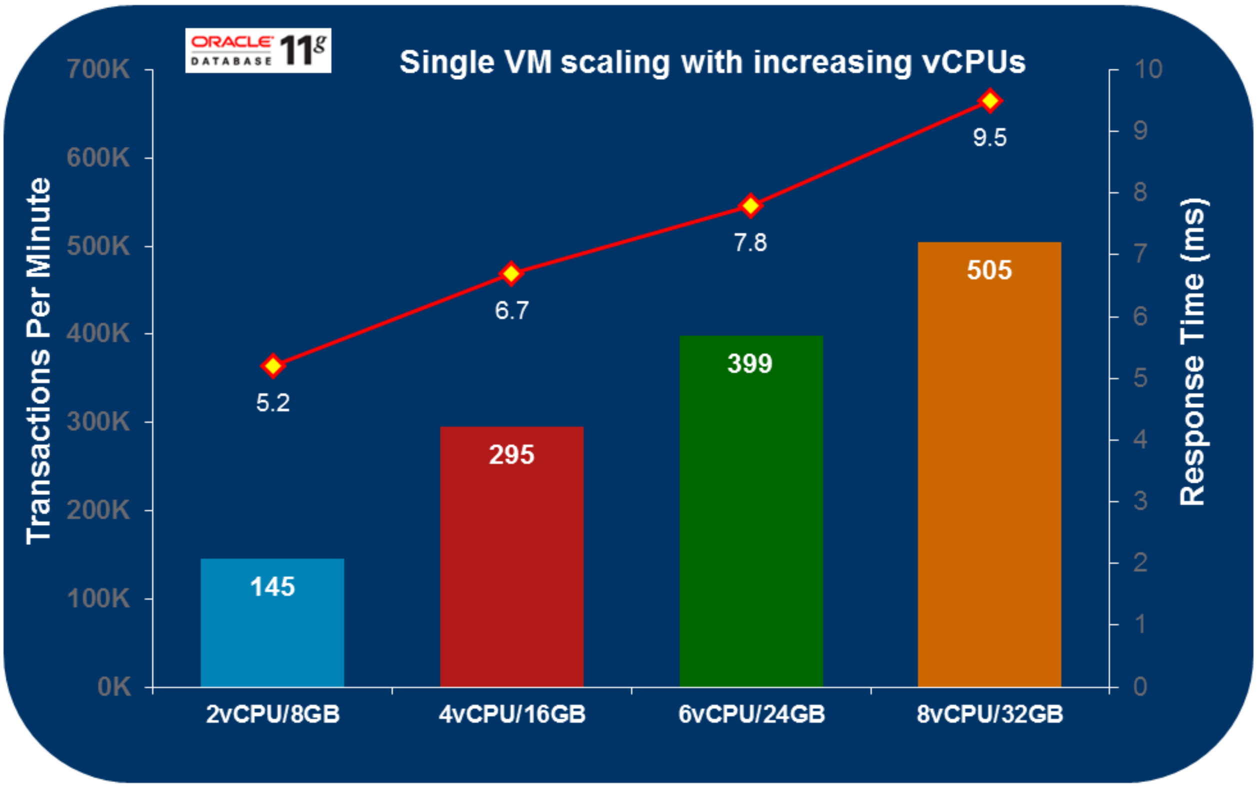
- LUN 1–8 (1TB each – RAID 5) -> DATA: datastore 1 – 8 VMFS -> 8 x VMDKs (900GB each) (Managed by Oracle ASM)
- LUN 9-10 (1.5TB each – RAID 10) ->REDO: datastore 9-10 VMFS -> 2 x VMDKs (1TB each) (Managed by Oracle ASM)

- LUN 11 (1.5TB – RAID 10) - > VMORA: datastore 11 VMFS -> 2 VMDK (30GB for OS, 50GB for Binary)

Single Virtual Machine Scale-Up Results

The results of the test runs are shown below. The chart shows the transaction throughput, measured in TPM, against the size of the virtual machine. In all runs, vCPU utilization within the virtual machine was maximized to nearly 100%.

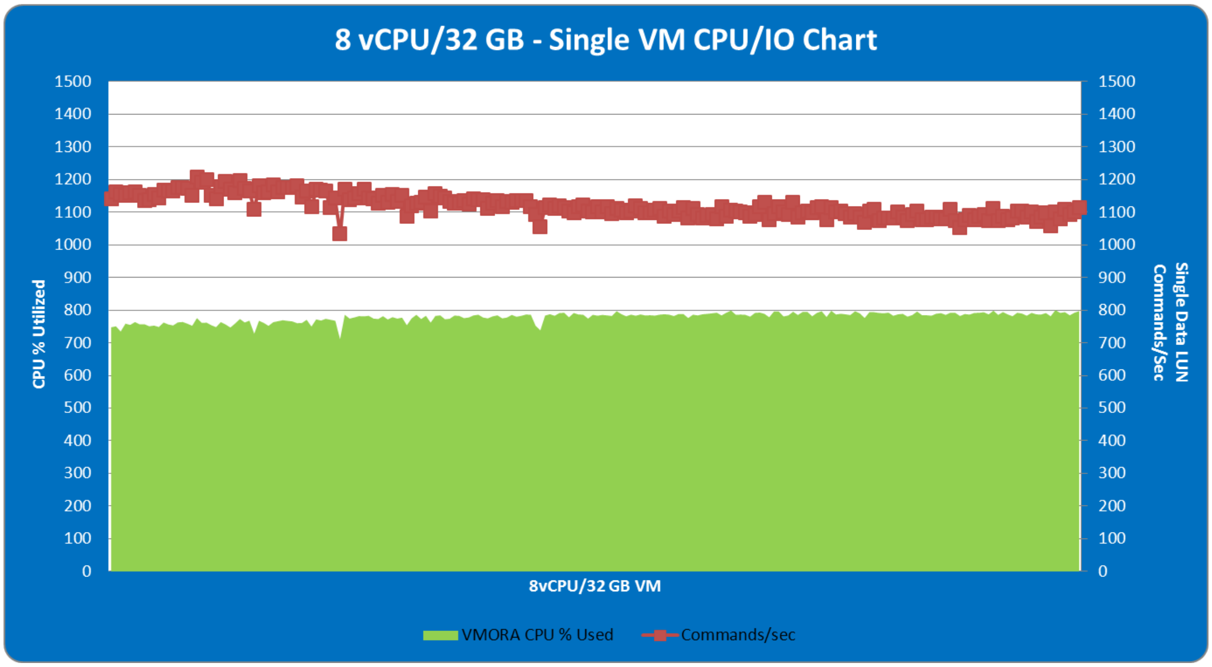
Figure 6. Single Virtual Machine – Scale-Up Chart



The results indicate a very reasonable scaling as the virtual machine size grew up in terms of vCPU and memory. The IO response time also increased as the virtual machine size grew up and the TPMs also increased. In other words, the cost of transaction (in terms of response time) increased as the virtual machine got busier and IO requests started queuing up at the storage. This is also reflected in the scaling factor for 8 vCPU virtual machine. We also observed that VMware vSphere effectively spanned the virtual machine vCPUs across 2 physical sockets (across NUMA nodes). This is explained further in Appendix B.

Figure 7 shows the CPU and IO statistics for the 8-way configuration.

Figure 7. Single Virtual Machine CPU/IO Chart – 8vCPU/32GB



This is a hard working virtual machine producing 505,000 TPM with cumulative vCPU utilization near 800% (as per standard "esxtop" reporting which corresponds to near 100% CPU utilization within the virtual machine). The disk statistics of one of the data LUNs reported 1120 IOPS (Commands/Sec) with latency of 9.5 ms/cmd.

These findings show that a single database deployed in a virtual machine can be scaled out to handle larger transaction throughput. Customers can size virtual machines for Oracle databases to suit their capacity needs up to a maximum of 8 vCPUs, which is the current maximum for vSphere 4.1.

Multiple Virtual Machines – Scale Out and Consolidation Test

While consolidation saves significant time and money, it presents unique performance challenges as database users vie for shared resources. This can result in inconsistent, unreliable and overall disappointing database performance during expected peak times. As a rule, it is extremely important to understand the peak load requirements for the database consolidation candidates. In summary the consolidated database solution must:

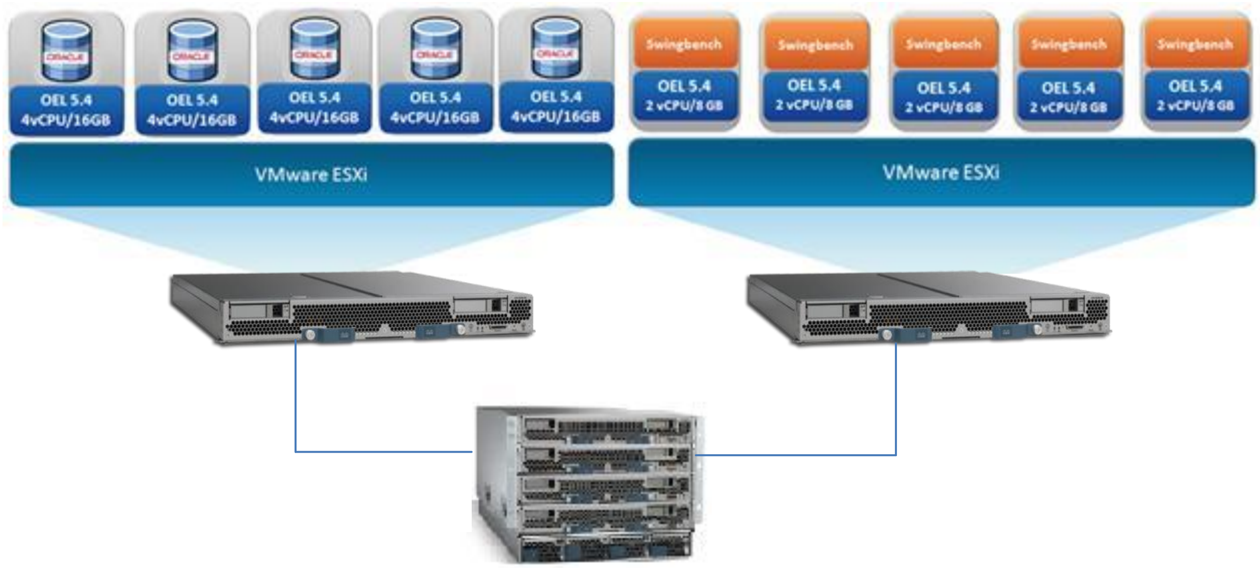
- Demonstrate fairness in resource sharing.
- Showcase sustained performance after consolidation

In the scale out test, we continued to stack virtual machines running independent database workloads on a single physical ESX server as long as they were able to sustain expected performance (~9500 TPMs) for a period of time. For our configuration, we were able to accommodate five independent virtual machines before storage IO started queuing up because of limited 75 spindles. The tests are described in **Error! Reference source not found.** and Figure 8 All five virtual machines shared the same LUNs/datastores.

Table 4. Multi-Virtual Machine – Test Configuration

vCPUs per VM	Database Size (TB)	Concurrent Users (Min, Max think time –ms)	Virtual Machine Memory	Oracle Server SGA Size
4 **(1 - 5 VMs)**	1.1TB	200 (50,100)	16GB	10GB

Figure 8. Multiple Virtual Machines – Scale-Out Test



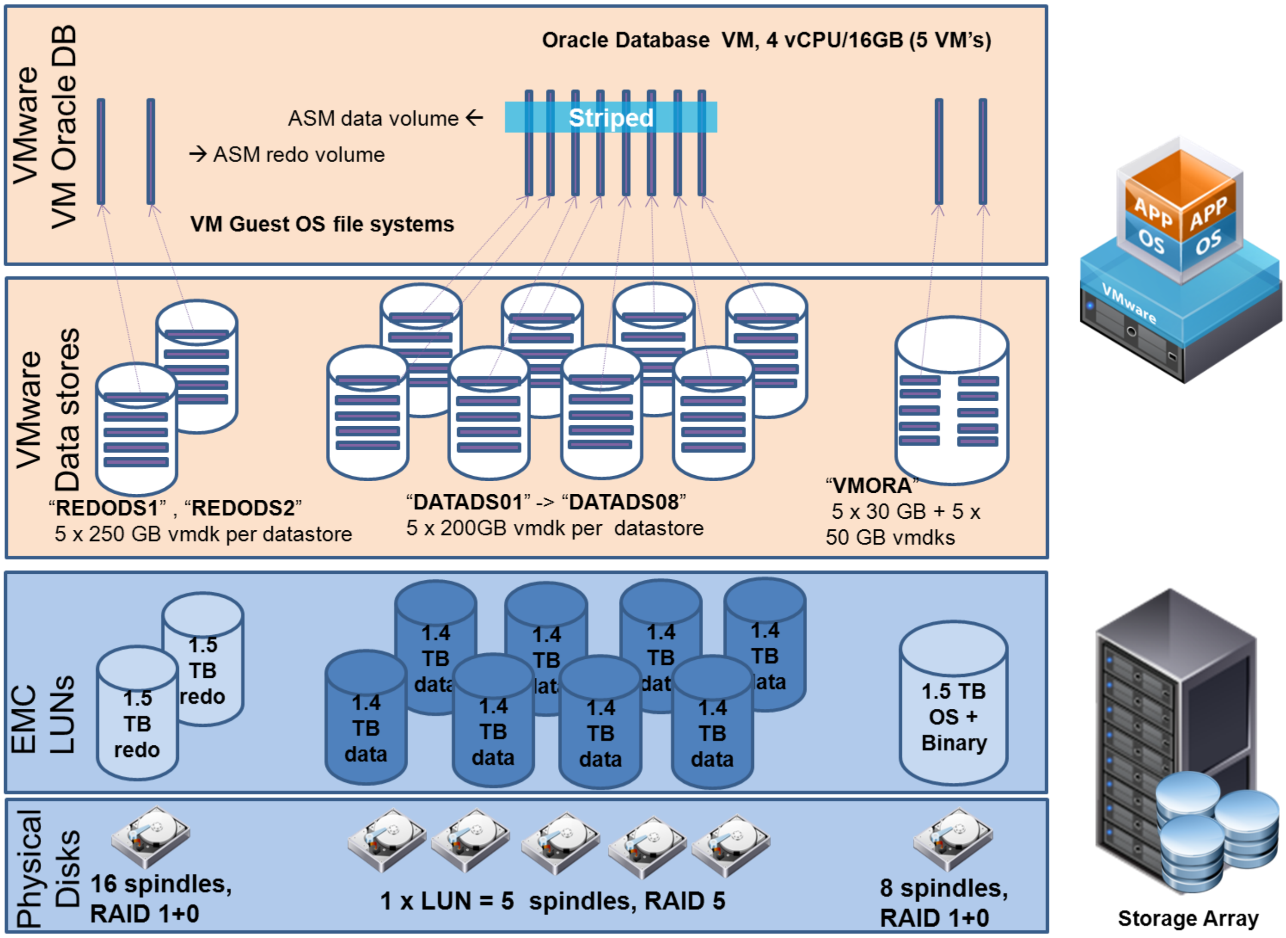
Storage Layout for Consolidation Tests

Figure 9 represents the storage layout for the Oracle databases.

- LUN 1–8 (1TB each – RAID 5) -> DATA: datastore 1 – 8 VMFS -> 8 x VMDKs (200GB each) for each Oracle database (Managed by Oracle ASM) – Each virtual machine has one VMDK from each datastore
- LUN 9-10 (1.5TB each – RAID 10) ->REDO: datastore 9-10 VMFS -> 2 x VMDKs (250TB each) for each Oracle database (Managed by Oracle ASM) – Each virtual machine has one VMDK from each datastore
- LUN 11 (1.5TB – RAID 10) - > VMORA: datastore 11 VMFS -> 2 VMDK (30GB for OS, 50GB for Binary) for each Oracle database – Each virtual machine has two VMDKs from the same datastore

  Note: Oracle ASM recommends only 1 partition per LUN as a best practice. In this consolidation case, we have storage system with 56 spindles for the database. So, we created 5 partitions for each LUN (one partition for each VM) to accommodate consolidated virtual machines. This way, each of the five VMs will have 8 Data LUNs and 2 Redo LUNs.

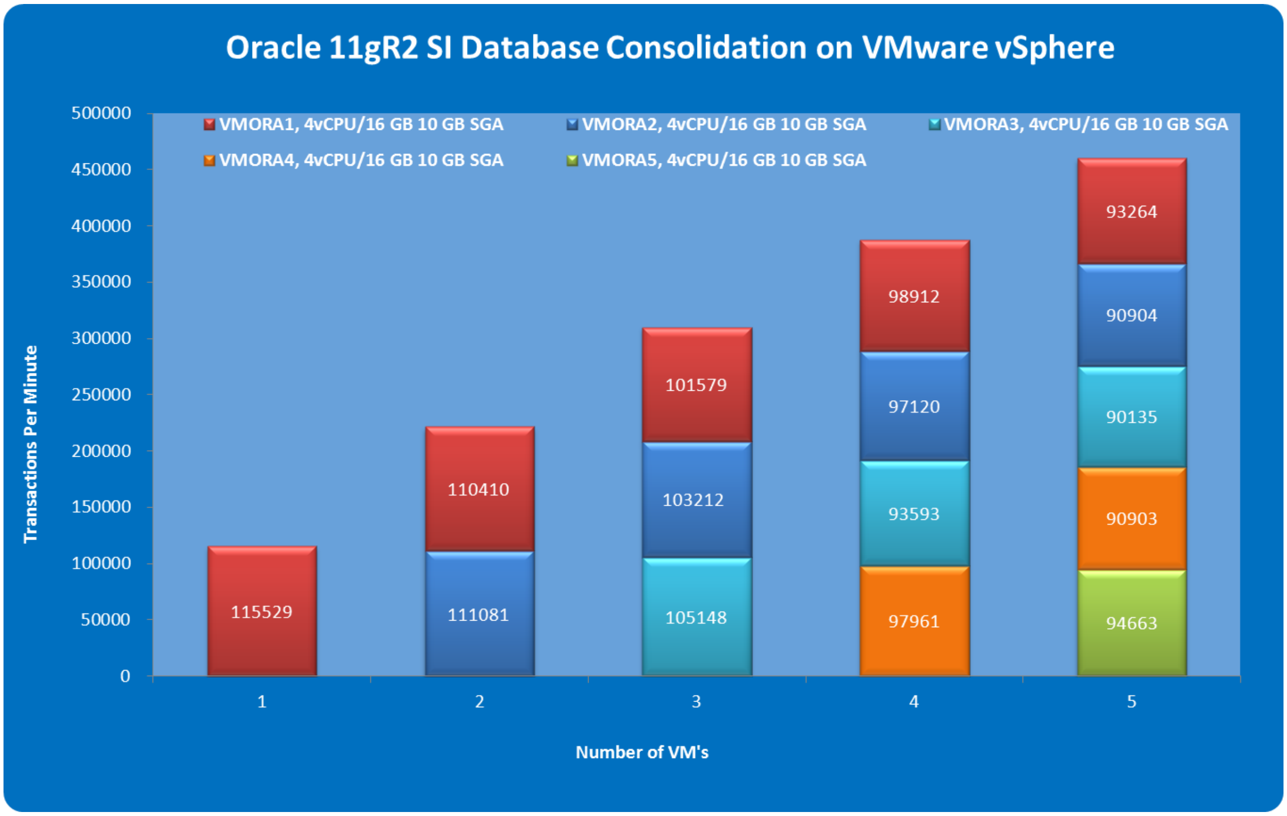
Figure 9. Multi-Virtual Machine Consolidation Study – Storage Layout



Multiple Virtual Machines Scale-Out Results

Figure 10 shows the results of the five test runs in the scale out test. The chart shows transaction throughput, measured in TPM, against the number of consolidated virtual machines. Each bar is broken down to show the TPM value of the individual virtual machine.

Figure 10. Multiple Virtual Machines – Consolidation



Note: We could not scale out beyond five virtual machines for our storage configuration. With five virtual machines running vCPU utilization within each virtual machine was about 45%.

Figure 9 demonstrates that the total throughput increases fairly linearly with additional virtual machines. The results show that ESX is able to fairly distribute CPU resources across the virtual machines. When all five virtual machines were running, CPU utilization of each virtual machine was fairly similar yielding similar performance is as demonstrated by the similar TPM values processed by each virtual machine. This can be attributed to the resource sharing and fairness properties of the vSphere scheduler. In summary, the ESX host was able to handle the workloads of the consolidated virtual machines. The general observations from this test are:

- With five virtual machines, the total number of vCPUs was more than the physical number of server cores. This CPU over-commitment was possible as the virtual machines were not fully saturated. This is common in scenarios where we typically see average CPU utilization of database servers well below 100% thus allowing more virtual machines to be consolidated on the same ESX host.
- The memory assigned to all the virtual machines was within the memory capacity of the ESX host so that the ESX balloon driver was not invoked during any of the scenarios.
- The multiple databases were able to share the same VMFS LUNs (Figure 8 shows the storage layout and consolidation of the multiple virtual machines in the same redo and data LUNs).

It should be noted that in cases of heavy I/O workloads, it is wise to separate such workloads out of the shared pool of storage to optimize the performance of those high transactional throughput applications. Since database workloads will vary, when looking at the number of virtual machines per LUN there is no exact rule to determine the limits of performance and scalability; this depends on the workload and an optimal storage design strategy between the application owner, VMware, database and storage administrators.

Live Migration of Virtual Machine (vMotion)

Typically, consolidation addresses the challenges around lower resource utilization. However, complex workloads and continuously changing workload demands make it very difficult to achieve an effective balance between server consolidation and resource availability. For example, how can someone grow the size of the virtual machine in the consolidated environment? VMware vMotion moves running virtual machines from one physical server to another with no impact to end users. When some of the consolidated virtual machines are moved, the freed up resources now can be made available.

In this vMotion test, we migrated a virtual machine running an Oracle database and performing OLTP transactions. The test execution sequence was:

- Virtual machine – VMORA1 running Oracle database server on ESXi Host 1.
- Execute 300-user Swingbench Large Scale Order Entry Benchmark for approximately one hour with minimum and maximum think time of 5ms and 10ms respectively.
- Live- migrate the virtual machine to target server ESXi Host 2.

  Figure 11 demonstrates the results. The chart shows transaction throughput against time. The first vertical line marks the start of live migration, the second the end of live migration.

Figure 11. Oracle Database Transactions per Minute During Migration with vMotion

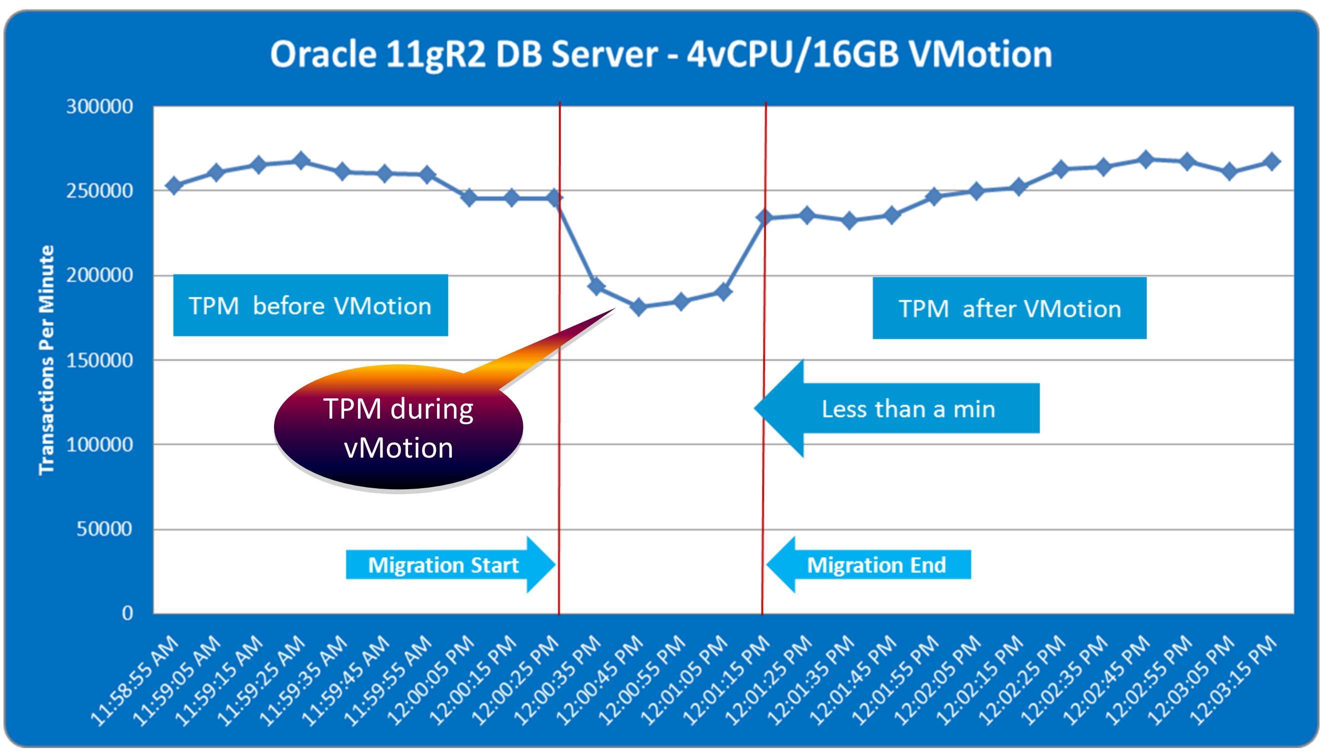


Figure 12 shows a slight impact on workload performance during migration, but transactions per minute immediately recover upon completion of the migration. No user errors or transaction failures are recorded and this busy 4 vCPU, 16GB RAM based virtual machine migrates to another server in less than a minute. The active memory and execution state transfer of virtual machine occurs over lossless, low latency and speedy 10GE network. This exercise validates that VMware VMotion allows datacenter administrators to optimize and allocate system resources to maximize capacity utilization with flexibility and availability.

VMware vCPU Hot Add (SuSE Linux Only)

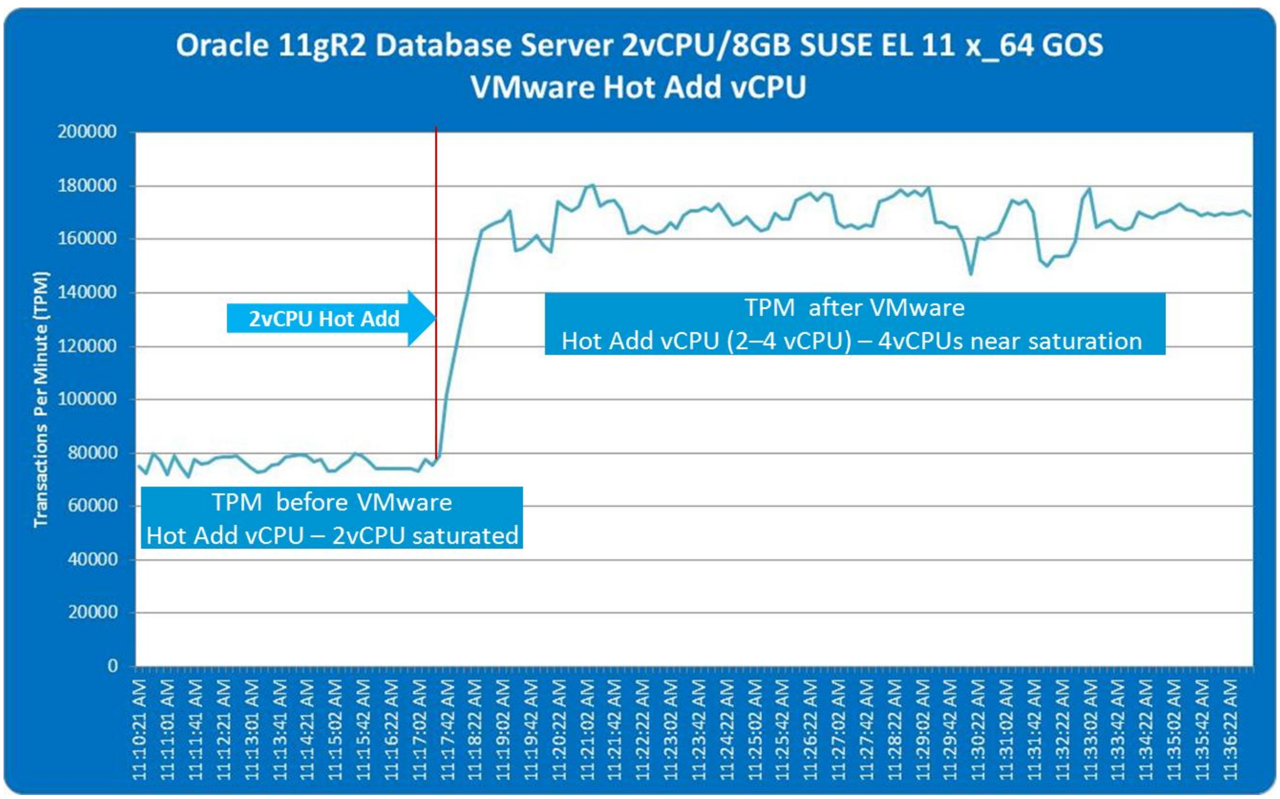
In the hot add scenario, we tested the vSphere hot plug feature for virtual devices which supports addition of virtual CPUs to a virtual machine without powering off the virtual machine. (Consult the [VMware Compatibility Guide] to determine support for hot add CPU for a specific guest OS.)

The test was executed as follows:

- Virtual machine – VMSUSE (Oracle database 11gR2 with SUSE Enterprise Linux 11 x_64 – VMware Hot Add vCPU supported Guest Operation System) running on ESXi Host 1 with CPU saturated near 100%.
- Execute 100 users Swingbench Large Scale Order Entry Benchmark load against 2 vCPU/8GB virtual machine for approximately one hour with minimum and maximum think time of 5 and 10ms respectively. Hot Add vCPU to VMSUSE (2 vCPU to 4 vCPU).

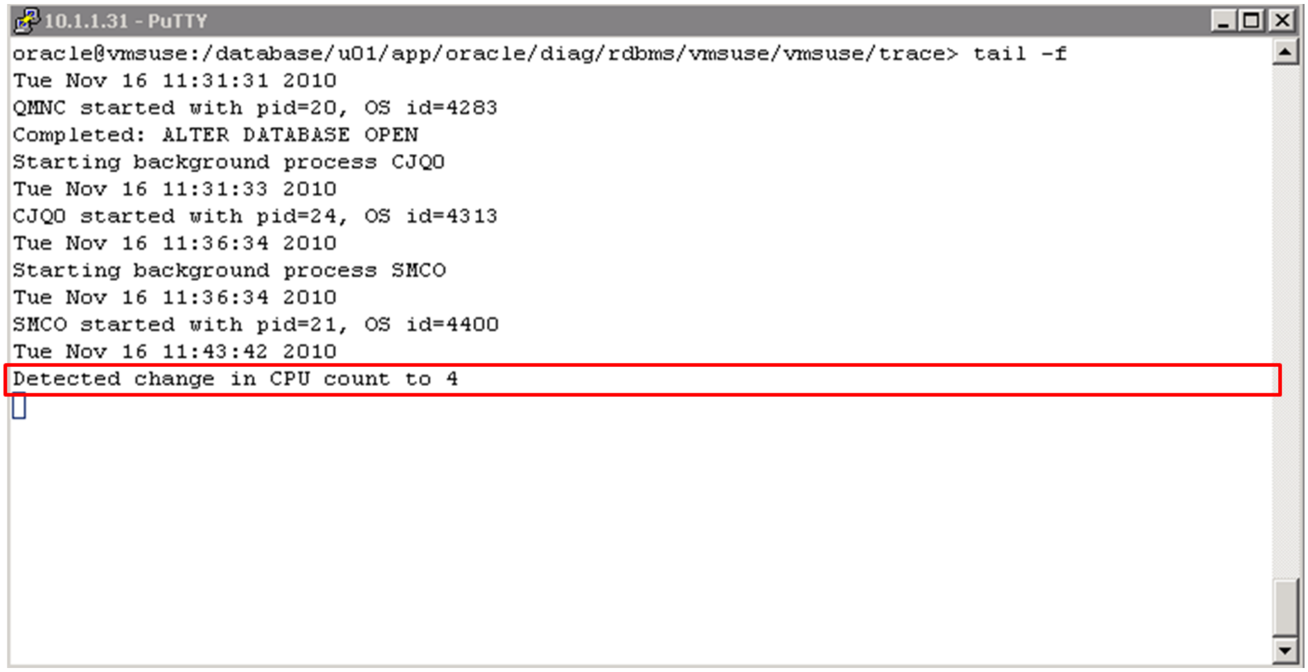
  Figure 11 shows transaction throughput, measured in TPM against time. It shows the TPM values of the virtual machine before and after the Hot Add vCPU operation. The vertical line marks the point in time the hot add operation was initiated in the vSphere client.

Figure 12. Oracle Database Transactions Per Minute with VMware Hot Add vCPU



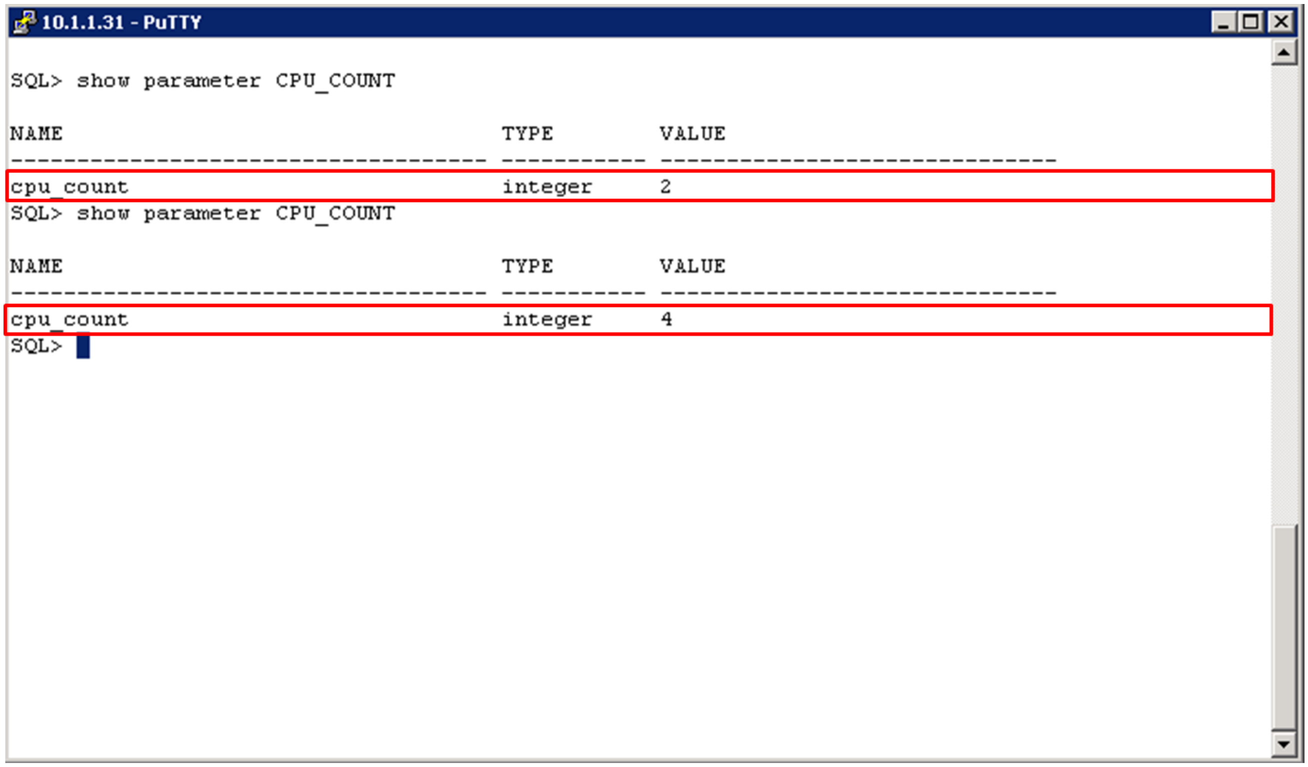
The chart shows the virtual machine initially delivered performance at just below 80000 TPM. After hot add of 2 vCPUs to the virtual machine, the transaction throughput (TPM) scaled appropriately to nearly double. As the number of vCPUs increased, the Oracle database recognized the extra vCPU availability. Figure 12 shows the detection of vCPU count in the Oracle alert log and Figure 13 shows that the Oracle database parameter CPU_COUNT changed from 2 to 4. This feature can be extremely useful for DBAs as they can add vCPUs on the fly without restarting the database or the virtual machine to meet the constantly changing workload demands. Databases can be sized in a smaller virtual machine, but as workload increases beyond initial estimates no downtime is required to add more processing capability.

Figure 13. Oracle Database Server Alert Log



```
oracle@vmsuse:/database/u01/app/oracle/diag/rdbms/vmsuse/vmsuse/trace> tail -f
Tue Nov 16 11:31:31 2010
QMNC started with pid=20, OS id=4283
Completed: ALTER DATABASE OPEN
Starting background process CJQ0
Tue Nov 16 11:31:33 2010
CJQ0 started with pid=24, OS id=4313
Tue Nov 16 11:36:34 2010
Starting background process SMCO
Tue Nov 16 11:36:34 2010
SMCO started with pid=21, OS id=4400
Tue Nov 16 11:43:42 2010
Detected change in CPU count to 4
```

Figure 14. Oracle Database Server CPU Count



```
SQL> show parameter CPU_COUNT

NAME                                 TYPE        VALUE
------------------------------------ ----------- ------------------------------
cpu_count                            integer     2
SQL> show parameter CPU_COUNT

NAME                                 TYPE        VALUE
------------------------------------ ----------- ------------------------------
cpu_count                            integer     4
SQL>
```


5. Conclusion

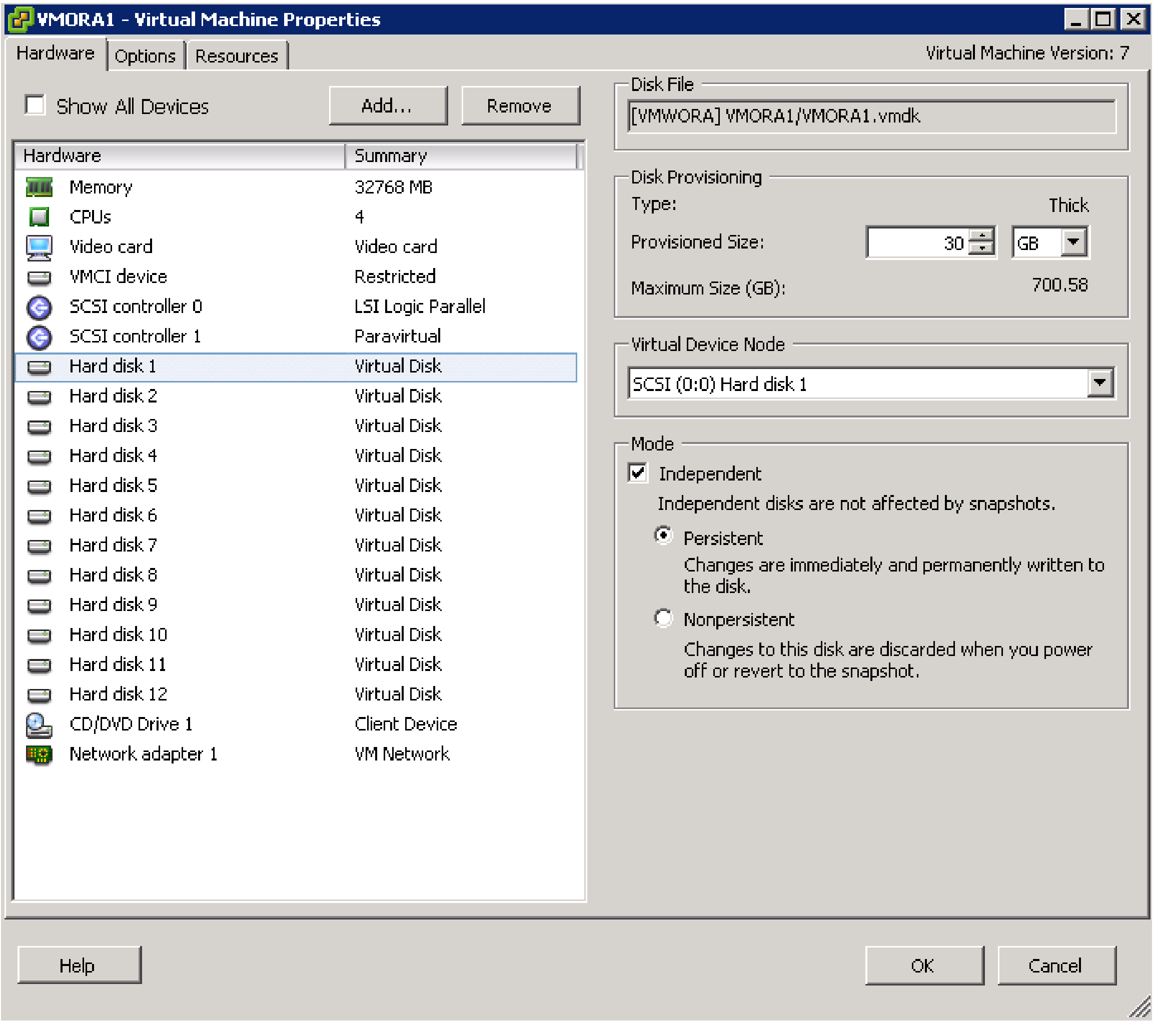
Demand for virtualized data centers is high due to the complexity of managing and provisioning physical resources, securing that environment, maximizing utilization of assets, numerous network connections, and the rising costs of facilities and energy usage. Cisco, EMC and VMware have worked closely over the past year on a shared vision for the future of enterprise data centers. Using the test cases that represent frequent real world challenges, we are able to address the common concerns that DBAs share for database virtualization.

- **Effective Database Performance and scalability:** Oracle databases run at performance similar to that of a physical system. We tested both scale up and scale out scenarios that highlighted CPU and memory based scalability along with consolidation.
- **Minimal Database I/O Overhead:** We did not see any significant IO overhead or queuing during consolidation test case even when running for a longer duration. The virtual machines were able to sustain consistent performance over the period of time.
- **Multi-core and CPU frequency based scaling:** We observed that VMware vSphere scalability beyond a single processor for each virtual machine. The performance boost for increased CPU frequency was also visible in the large memory test cases.
- **Capacity on demand with flexibility:** With vMotion and CPU hot add feature(SuSE Linux only) tests, we validated that system administrators gain substantial flexibility in being able to balance virtual machines freely across ESX hosts to improve or maintain performance or when they need to perform server maintenance with no impact to end-users. Today’s constantly changing workload demands can now also be addressed via scale up or scale out mechanism that empowers the DBAs to achieve the optimum performance for their database virtual machines and while providing more flexibility to right-size system deployment.

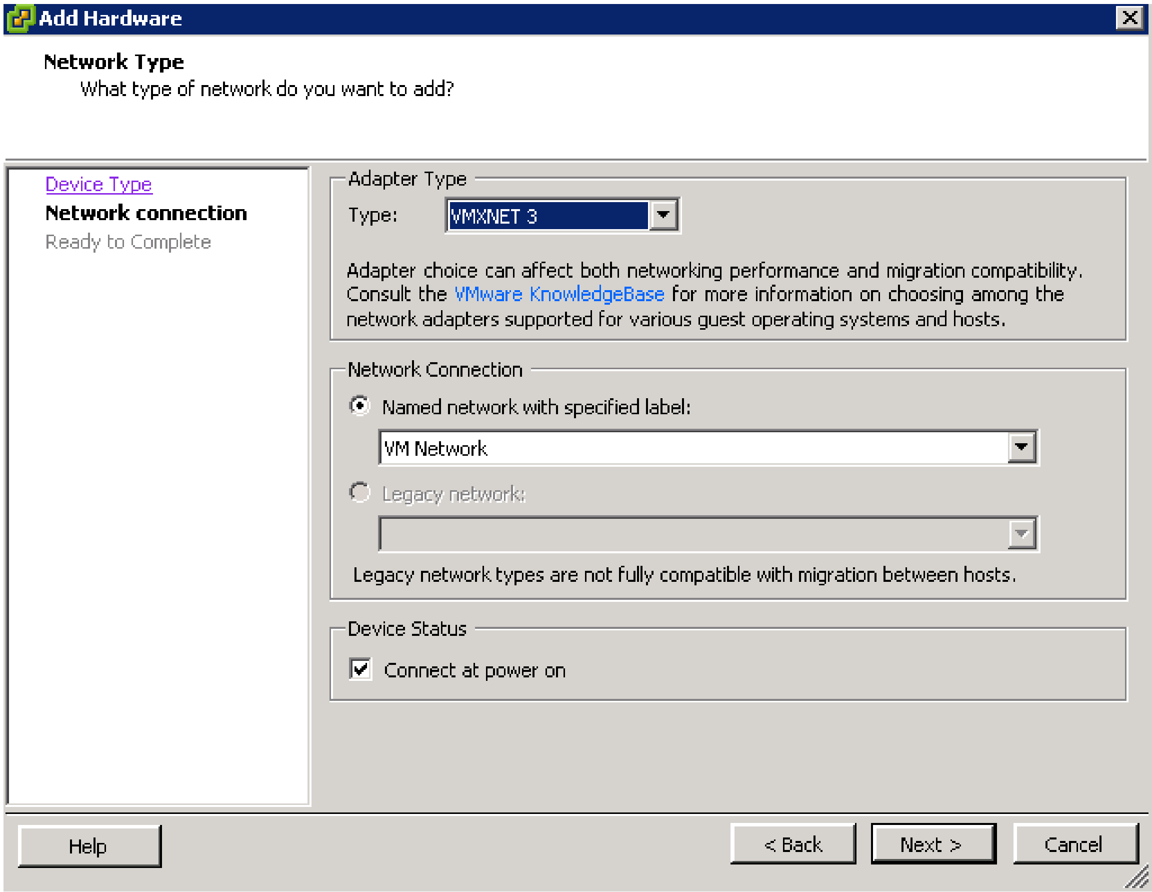
Appendix A: Virtual Machine Configuration

This Appendix describes some of the virtual machine configuration steps and properties.

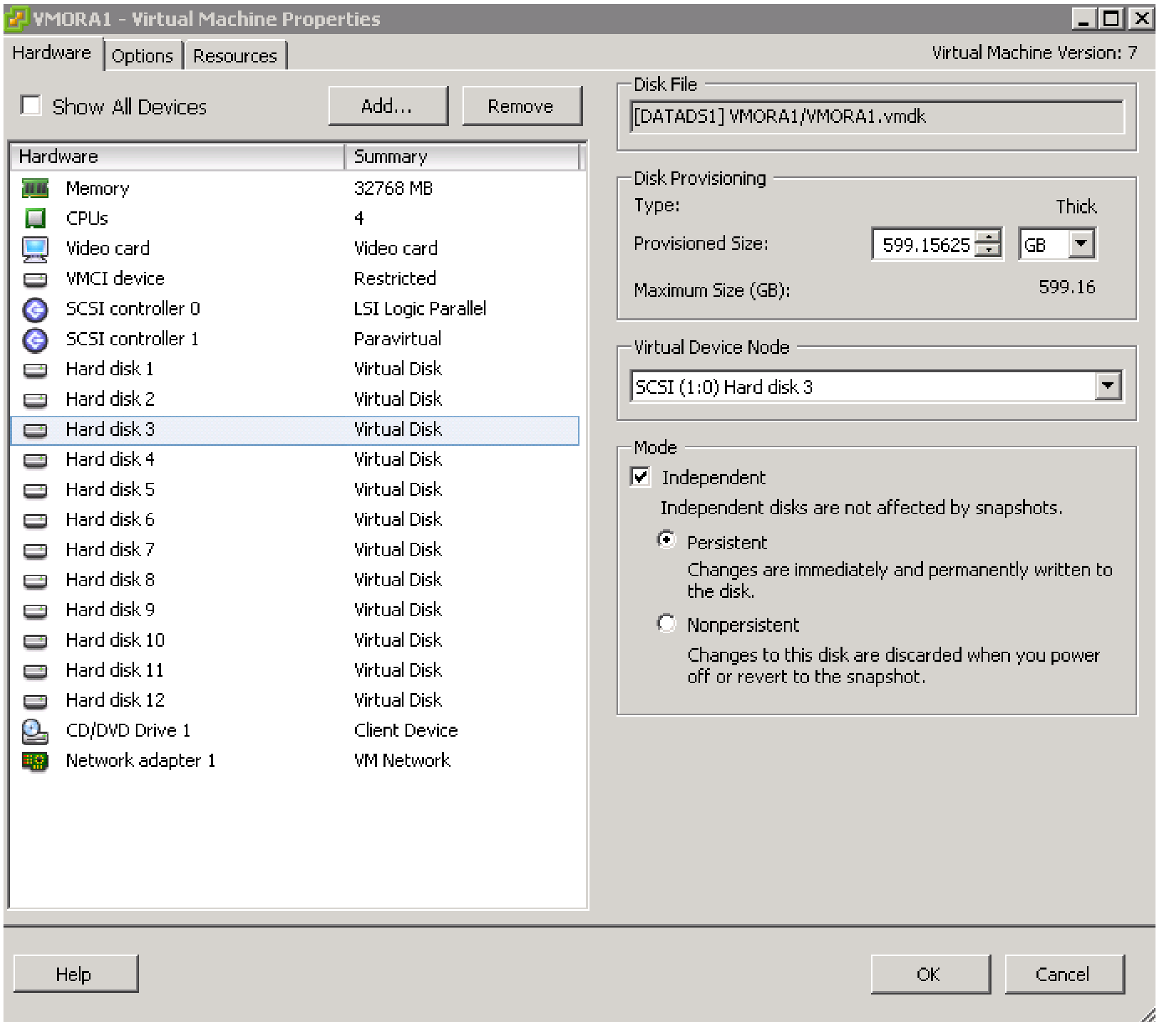
1. Create the Oracle database virtual machine with the following properties.



2. Network: Create or add network adapter and select **VMXNET3** for better performance. VMware Tools must be installed for the VMXNET3 adapter type to be available.



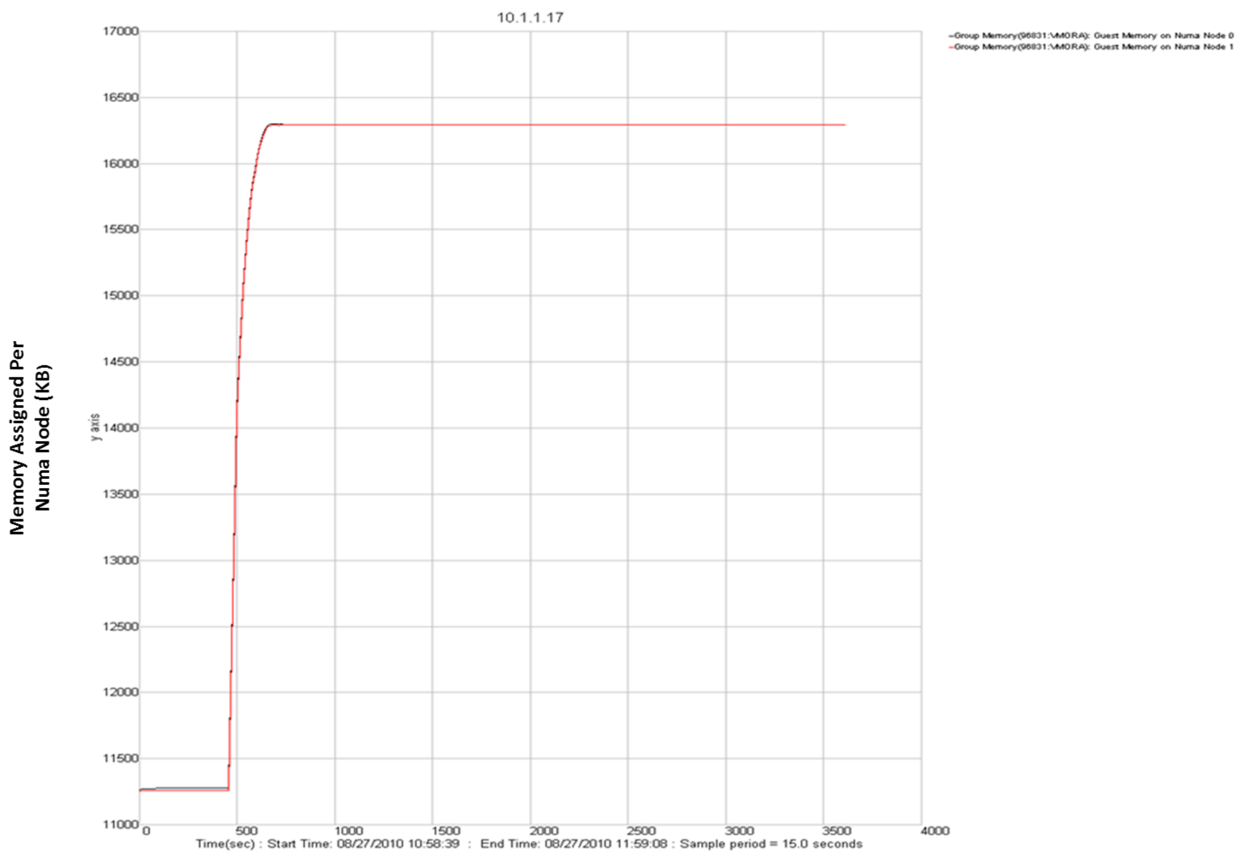
3. Hard Disk: Create Hard Disk 1 and 2 with default SCSI (LSI Logic Parallel) adapter for both VM files and Oracle Binary.
4. Create Hard disk 3 – 10 for Oracle Data and Hard disk 11 – 12 for Oracle Redo logs with a VMware paravirtual adapter for better performance.
5. Select SCSI 1:0 for virtual device node to create a new SCSI controller and change the controller type to **VMware Paravirtual**. VMware Tools must be installed for the paravirtual option to be available. Follow VMware KB 1010398 to configuring disks to use VMware Paravirtual SCSI (PVSCSI) adapters



6. Install Oracle Enterprise Linux 5.4 64-bit and install and configure Oracle Database 11gr2, ASM with database size 1TB. Database disk layouts are described in each tested scenario.

Appendix B: 8 vCPU/32GB Virtual Machine Memory Split

This Appendix shows the NUMA memory statistics captured during the single virtual machine scale up run where the virtual machine was configured with 8 vCPU and 32 GB RAM. The following chart was obtained from "esxtop" data and extracted using "esxplot". It shows that the virtual machine memory of 32 GB was evenly split across two NUMA nodes.



The above behavior is expected in vSphere 4.1. A Wide-VM is defined as a virtual machine that has more vCPUs than the available cores on a NUMA node – this is the case in the above example where the ESX server consisted of two 6-core sockets, ESX 4.1 allows wide-VMs to take advantage of NUMA management. NUMA management means that a virtual machine is assigned a home node where memory is allocated and vCPUs are scheduled. By scheduling vCPUs on a NUMA node where memory is allocated, the memory accesses become local, which is faster than remote accesses.

Wide-VM NUMA support is accomplished by splitting a wide-VM into smaller NUMA clients whose vCPU count does not exceed the number of cores per NUMA node, and assigning a home node to each client. The 8-vCPU virtual machine on the ESX server has two 4-vCPU NUMA clients. The wide-VM has multiple home nodes because it consists of multiple clients, each with its own home node. This is implemented in a manner to improve performance of memory access – for details see the whitepaper VMware vSphere: The CPU Scheduler in VMware ESX 4.1 .

Cisco Systems, Inc.
170 West Tasman Drive
San Jose, CA 95134-1706
USA
www.cisco.com
Tel: 408 526-4000
800 553-NETS (6387)
Fax: 408 527-0883

Document number: UCS-TR100026