Storage Education Series

Best Practices: Optimizing Oracle on VMware vSphere 5.1

Authors:

Leah Schoeb, Sr. Partner
Evaluator Group

Ed Whalen, Chief Technologist
Performance Tuning Corporation

February 1, 2013

Evaluator Group

Enabling you to make the best technology decisions

Table of Contents

Introduction

Virtualization has become widely accepted in today's enterprise datacenter, from consolidating thousands of desktops to mission critical Tier 1 applications in server virtualization. One rapidly growing area for server virtualization is running critical Oracle database applications on VMware® vSphere® virtual machines. VMware vSphere virtual machines can now be the same size as some of the largest physical x86 servers running large Oracle databases. In turn, these virtual machines have the same capabilities, performance and resources to run database environments.

Determining the size and demands of a virtual machine running an Oracle application will depend on how many other virtual machines and their applications can share resources without impacted performance. Even though VMware supports much larger virtual machines, they are still subject to the same issues all other applications run into in a shared virtualized infrastructure. VMware vSphere 5.1 now supports 64 vCPUs and 1TB of main memory, which has been measured to support 1 million IOPS performance and over 36 Gb/sec in throughput.

Database engines and their infrastructures provide mission-critical support to most large IT departments worldwide. VMware vSphere simplifies IT infrastructures enabling IT administrators to leverage compute, network, and storage resources more efficiently and with greater agility. Virtualized infrastructures also bring rapid deployment and TCO improvements, which allow IT administrators to quickly support customer requests and easily manage to keep up SLAs.

The Oracle Database is one of the most popular applications being virtualized today, so VMware has put resources into optimizing performance for Oracle Database applications. Building large Oracle databases on VMware vSphere 5.1 is not only practical but may be preferred. We will present best practices for deploying Oracle databases on vSphere 5.1.

Oracle Database is a relational database management system whose primary job is to structure, manage and maintain data. Oracle databases are used to manage and serve business critical data where the loss of service and loss of data can be catastrophic to business. Oracle databases range in size from a few Gigabytes to hundreds of Terabytes and even Exabytes. Oracle servers are typically resources intensive and configured with a large number of CPUs and memory.

Because the Oracle database server maintains enormous amounts of data, its performance is very sensitive to the performance of the I/O subsystem. In order to optimize database performance, the database server uses large amounts of main memory and solid state technology for caching data.

VMware has provided multiple features for optimizing I/O, network, memory, and CPU. In addition to performance features that can be used to optimize the Oracle database server on VMware, high availability and management features are available as well. This paper will discuss VMware features and enhancements and how they might be used to provide a stable, robust, high performing Oracle database environment.

VMware vSphere 5.1 Benefits

There are many benefits for virtualizing a system with VMware vSphere running an Oracle database. Recent performance results from VMware have demonstrated that Oracle performance in a virtualized infrastructure performs very near native physical environments with the added benefits:

- **Manageability and Agility** – VMware's vMotion® moves virtual machines from server to server with no downtime to the Oracle database and can simplify operations, such as hardware maintenance and upgrades.
- **Consolidation** – Multiple application servers can be consolidated on a single physical host. This allows for maximum use of physical resources (CPU, memory, network, storage).
- **Provisioning** – vSphere can encapsulate and application into an image that can be moved or duplicated. This helps immensely in needs for rapid deployment.
- **High Availability (HA)** – For unexpected or unplanned downtime VMware vSphere will restart virtual machines on another host in a cluster. This feature is transparent to the Oracle Database while minimizing downtime and offers higher levels of service for Oracle Databases and other applications.
- **Fault Tolerance (FT)** – The fault tolerance offered by VMware vSphere is architected to support no downtime or data loss and provide continuous availability for Oracle Databases and other applications.
- **Dynamic Resource Scheduling (DRS)** – VMware vSphere DRS® and Storage DRS® provides load balancing for both virtual machines and VMDKs across a cluster. This allows for virtual machines and VMDKs to automatically be relocated in a cluster to keep the cluster load balanced based on any reservations (guaranties), limits (hard caps) and shares (contention based on prioritization) configured for Processor, memory, and storage resources. This means virtual machines with Oracle Databases will automatically get moved to servers and storage with more resources in order to keep up SLAs. In addition, DRS will also allow configuration of affinity and anti-affinity rules. These rules will govern which virtual machines live on the same host or never live on the same host.
- **Power Management** – VMware's Dynamic Power Management (DPM) allows for conservation of power when resources aren't needed.
- **Rapid refresh** – Virtualized environments can be quickly refreshed for test/dev, training, and QA/UAT environments. This saves a lot of time in development, testing, and training cycles.
- **Network I/O Control -** NetIOC allows for assignment of bandwidth or relative weighing between VM's for sharing the network pipes and even more so for 10Gb interfaces.
- **Storage I/O Control –** SIOC allows a level of storage predictably and fairness to ensure tier 1 workloads, like many Oracle Databases, get what they need, or what they are entitled too. It is

not a guarantee of CPU, memory, and network resources, but 'fair' access. This is turned on in vSphere by default.

Best Practices for Oracle Database in a VMware vSphere Infrastructure

An Oracle deployment in a VMware vSphere virtualized infrastructure is not much difference than deploying Oracle in a physical configuration. This gives Oracle DBAs and opportunity to use their current skill set for deploying and managing the Oracle Database. The economic benefits lower CAPEX and OPEX are popular reasons why IT departments are virtualizing their database applications at a rapid rate.

In order to provide for an optimally performing Oracle database system, proper sizing is essential. Sizing the virtual machine for an Oracle database server requires proper sizing of CPU, Memory and I/O (both capacity and performance). Sizing the Oracle server system for VMware is not unlike sizing an Oracle server system for a standalone server.

There are a number of configuration options that will enhance the performance, stability and usability of your Oracle virtual machine. These configuration options are fairly easy to implement and provide an optimal platform for your Oracle database server. Careful sizing and capacity planning for any given workload is still highly recommended when taking advantage of the configuration options mentioned in this section.

VMware Tools

It is recommended that VMware Tools be installed on all virtual machines that will be running the Oracle database. VMware tools provide optimized disk, network and video drivers that enhance the performance and usability of the virtual machine.

CPU and vCPU Allocation

Unlike physical machines, CPUs in a virtualized environment (vCPUs) share the same resources. vCPUs can also be over allocated, causing vCPU starvation. Therefore it is important to properly allocate vCPU resources for each virtual machine. Additionally, it is important to properly size the host so that physical CPUs are not overused.

CPU resources can be over allocated for efficiency. If you are over allocating CPU resources, you should consider CPU reservations, in order to guarantee a minimum number of CPU cycles. A CPU reservation is used to guarantee resources to the virtual machine.

Bigger is not always better - Depending on the application, adding more vCPUs may not offer more performance. There are many reasons why adding more vCPUs does not yield more performance and they may be fixed through tuning or adding other resources. In any case, each vCPU imposes scheduling

constraints and any excess vCPUs can cause degradation in performance with unnecessary additional constraints.

Non-Uniform Memory Architecture (NUMA) - Since the virtual CPU is abstracted from the underlying physical CPUs and VMware vSphere supports NUMA aware, we recommend that NUMA be disabled at the virtual machine layer. In order to disable NUMA, add *numa=off* to the boot string in the grub.conf file. This will disable NUMA at boot time. Oracle NUMA support is disabled by default for Oracle 11g and above. Details can be found in the Oracle support documentation. VMware does recommend at the physical host BIOS and guest OS levels that NUMA stay enabled. There are performance advantages of taking advantage of hyperthreading.

Virtual Symmetric Multiprocessing (Virtual SMP) – Since Oracle architecture is multithreaded, it is a very good candidate for Virtual SMP. Only applications that support running multiple processes or threads in parallel can really benefit from SMP. Single threaded applications do not qualify to take advantage of this architecture.

Memory

Host Physical Main Memory - Using a physical host that supports a Hardware-Assisted Memory Management Unit (MMU) is very important. Hardware-Assisted MMU is a chipset features that specifically supports a memory management unit for virtualization. This x86 chipset support from Intel and AMD is referred to as EPT for Intel and RVI for AMD. These processor manufacturers have put in an additional level of hardware-implemented page tables. These tables contain address translations from virtual machine to physical machine translation. By default in vSphere, this is set to Automatic in the CPU/MMU Virtualization option.

Memory, like CPU, can also be over allocated. VMware created a memory management architecture designed for efficiency in over-commitment of memory.

1. Transparent Page Sharing (TPS) allows for duplicate pages from different virtual machines that are identical to be shared. This both reduces memory utilization and is more efficient. Memory can safely be over allocated, but memory reservations should be used for important virtual machines.
2. Ballooning – This overcommit technology allows a virtual machine to use more memory than has been allocated.
3. Memory Compression – This feature optimizes the use of memory by compressing a virtual machine's memory 2:1 on the fly, freeing up space as needed.
4. Swap to Solid State Storage – If there are solid state devices (SSDs) or PCIe flash cards present in the physical host an option will be made available to create a swap file to solid state storage. If there is no solid state storage present this option will not be offered.

5. Swap to Disk – When all other options in the hierarchy have been exhausted data will be swapped to disk.

Memory should be sized based on Oracle database needs. It is rare when over-commitment is not recommended for an application and Oracle Databases is one of those rare exceptions for its SGA. The memory should be able to support the SGA and PGA memory that is required by the database instances without over commitment.

Guest OS Memory

Use Large Memory Pages - For Oracle SGA sizes over 2 GB in size we recommend using Linux Hugepages. Linux Hugepages allow for Oracle SGA memory pages to be allocated with a 2 MB page size rather than the traditional 4 KB page size. This will greatly reduce the number of pages that are stored in the pagetable. This will result in a smaller pagetable, a better TLB cache hit ratio and less kswapd utilization. **Do not use Automatic Memory Management**. Oracle 11g's new feature Automatic Memory Management (AMM) does not support hugepages.

Oracle SGA size – Make sure to set memory reservations to the same size as the Oracle SGA. This is to avoid performance degradation with kernel swapping between ESX and the guest OS activity. There is a trade off in setting reservations that can limit the use of vMotion for that virtual machine. Virtual machines can vMotion to another host if that host has the same amount of free physical memory or greater availability to take on the set reservation. If the Oracle database is being used in a non-productive environment, for example test/dev, then overcommitting and taking advantage of VMware's memory management architecture is acceptable.

Balloon Driver – VMware recommends not to disable the balloon driver as it is very valuable to the Oracle database's up time. Even if the balloon driver gets enabled, for whatever reason, it will save the application from a severe drop off in performance by swapping to disk.

Storage Subsystem

The I/O subsystem is critical to the performance of the Oracle database server. Sizing of the I/O subsystem is no different than from a stand-alone server if dedicated storage is used. When using shared storage, I/O performance should be closely monitored. I/O performance in Oracle database systems is one of the key causes of performance issues.

Dedicated Datastores - With Oracle database servers, I/O performance (IOPS) and latency is more important than throughput (MB/sec). Just like a physical implementation of a mission-critical workload with dedicated physical LUNs, creating dedicated datastores for Oracle databases is key.

Using VMFS vs. RDM - To balance performance and manageability VMware does recommend using VMware vSphere's VMFS; but for extreme cases of needed performance to ensure SLAs, using Raw Device Mapping (RDM) may be necessary. Performance data from VMware has shown that VMFS and RDM demonstrate similar performance. Raw Device Mapping allows the virtual machine to directly

access a disk partition. Since Oracle is very sensitive to I/O latencies, the use of RDM is recommended for large partitions and heavy I/O use databases.

Alignment - Just like any other file system, VMFS will have degraded performance if a partition is not properly aligned. Make sure to create VMFS partitions in VMware vCenter™ for automatic alignment during the creation of a VMFS partition. Partitions with high activity *diskpart* may be used. For further alignment needs consult your storage vendor.

Paravirtualized Drivers – Paravirtualized SCSI adapters have shown to have low overhead and give good performance for highly demanding Oracle data files. Combining the latest Paravirtualized SCSI driver (PVSCSI) and ESXi kernel optimizations demonstrates good performance for Oracle data files. For Oracle databases with light loads (<2000 IOPS) and had more than four outstanding I/Os, PVSCSI is not recommended.

Network Array Storage (NAS) – If a decision has been made to use networked storage with iSCSI or NFS file protocols for the price/performance advantages, enabling jumbo frames is a must. This allows for a larger payload of traffic and will yield better performance.

Storage Array Network (SAN) – Fibre Channel (FC) is usually the first choice in high speed online disk storage to handle the demands of mission critical Oracle databases. Recently with the introduction of solid state technology, solid state storage has been a popular choice for indexes, redo logs and frequently used data to help either accelerate or meet the demands of the database I/O load.

Planning - Proper storage sizing requires two dimensions, performance and capacity. Work with your storage vendor's Oracle best practice documentation for capacity planning and performance management.

Virtual Machine Storage Subsystem

There are several things that can be done to the virtual machine storage subsystem to improve performance and improve usability. They include the use of VMware Tools, Oracle ASM and by assuring proper partition alignment. Using RDM is also recommended.

Virtualization Driver - VMware Tools includes a virtualization driver that improves the performance of the I/O subsystem by reducing the code path necessary for performing I/O operations. The virtualization driver is an important part of optimizing the virtual machine for Oracle.

Oracle ASM - Oracle's Automatic Storage Management is a high performance I/O subsystem that uses a small Oracle instance for managing I/O operations. Introduced in version 10.1 it has been used for a long time, is stable and performs very well. Oracle ASM works well in a virtualized environment and is recommended. There are a couple of considerations when using Oracle ASM:

- ASM disk groups need to be created with equal disk types. The performance of an ASM disk group will only perform as fast as it's slower member of the group.

- Create ASM disk groups based on I/O behavior characteristics (random, sequential, log files, etc…).

For NAS environments do not use Oracle ASM failure groups because of unpredictable behavior; and, it may consume additional resources during a disk failure. Use external redundancy to make disk failures transparent to the database.

Virtual Controllers – Use multiple virtual SCSI controllers to increase parallelism for I/O traffic inside the guest OS. Just like in physical database environments, it is recommended to separate Redo Log traffic from data files through separate virtual controllers. Guest OS and swap should be on a separate controller as well.

Partition Alignment - It is very important to make sure that partitions are aligned within the virtual machine. VMware itself maintains partition alignment for virtual machine storage at the host level, however, partition alignment when using RDM is the responsibility of the virtual machine.

Network

The number of network adapters allocated for the Oracle virtual machine will be determined by application requirements. Jumbo frames should be used when required, for things such as RAC interconnects.

Network performance is important to the overall performance of the Oracle database server. There are several things that can be done within VMware to improve network performance.

Paravirtualized network Adapters – Paravirtualized network adapters minimize overhead by offering an optimized interface between the virtual machine and the physical network.

Security – Separate infrastructure and virtual machine traffic. Sharing these types of traffic is not only a security violation for most production datacenters but this can also cause negative impacts on performance.

Load Balancing – NIC teaming is a great way to load balance network traffic and can also provide passive failover in cases of failure.

Jumbo Frames - Jumbo Frames is supported within VMware, but only recommended when Jumbo Frames are appropriate. Jumbo Frames are typically recommended for the following cases: Interconnect in Oracle RAC environments; storage networks when NAS is used; and, between Oracle Data Guard Systems.

Guest OS

It is also important to optimize the OS for Oracle stability, functionality and performance. There are a few things that are very important for Oracle database servers.

Time Keeping for Virtual Machines - It is important for Oracle database servers to have the correct time. The VMware Tools time synchronization option is only recommended for desktop systems. We recommend disabling this feature and enabling NTP (Network Time Protocol) for time synchronization.

OS Configuration for Oracle - For consistency and accuracy we recommend setting up Linux packages and configuration by using the Oracle pre-install packages. For Oracle Enterprise Linux (OEL) or Red Hat Linux 5.x use the oracle-validated RPM. For Oracle Enterprise Linux (OEL) or Red Hat Linux 6.x use the oracle pre-install RPM.

Disabling unnecessary processes – VMware recommends that all unnecessary foreground and background processes within the guest OS be disabled.

Other considerations

Use at least VMware vSphere 5.0 – VMware has shown in many Oracle Database environments there has been as much as a 20% gain in performance using at least vSphere 5.0. Each ESXi host can support up to 2TB main memory, 512 virtual machines and 2048 CPUs. vSphere 5.0 supports 32 vCPUs and 1TB of main memory, while vSphere 5.1 supports 64 vCPUs and 1TB of main memory per virtual machine.

Optimize the physical computing environment – There are certain BIOS settings on the physical host that can maximize compute resources for vSphere environments with Oracle Databases. An example is disabling unnecessary process and peripherals. The VMware website has the details on BIOS settings and their recommendations.

Choosing the best virtual machine monitor – Make sure for the CPU/MMU Virtualization option that Automatic is chosen. This allows for vSphere to choose the best virtual machine monitor, based on the CPU and guest OS.

Summary

VMware provides an efficient and optimal environment for Oracle database servers. The benefits of using Oracle on VMware will vary based on your environment and needs. Oracle databases have been in use on VMware for many years with great success. Since VMware supports efficient I/O and memory access, the overhead incurred by VMware is minimal. By following the best practices and tips provided in this paper, your Oracle environment can be configured optimally with maximum performance capabilities.

About Evaluator Group

*Evaluator Group Inc. is dedicated to helping **IT professionals** and vendors create and implement strategies that make the most of the value of their storage and digital information. Evaluator Group services deliver **in-depth, unbiased analysis** on storage architectures, infrastructures and management for IT professionals. Since 1997 Evaluator Group has provided services for thousands of end users and vendor professionals through product and market evaluations, competitive analysis and **education**. **www.evaluatorgroup.com** Follow us on Twitter @evaluator_group*