ORACLE MONSTER VIRTUAL MACHINE PERFORMANCE

VMware vSphere 6.5

vmware®

Table of Contents

Executive Summary

The ability to run extremely large virtual machines on vSphere 6.5 aligns closely with the needs of large databases in the enterprise today. Specifically, Oracle database instances in some cases require large amounts of RAM, many processors, and high-speed flash-based storage to achieve the performance needed. Tests were run in the VMware lab using monster virtual machines running Oracle databases to study their performance. Tests were run across a range of server generations and a flash storage array to illustrate how much performance has increased with each new generation of processor, the performance benefit of Hyper-Threading, and NUMA effects. The results of these tests show that the performance of Oracle databases on large virtual machines is excellent.

Introduction

In 2005, VMware ESX® 3.0 was released with support for virtual machines with 4 vCPU and 16 GB of memory. This capability was an early step toward the massive capacity of monster VMs in vSphere 6.5 that had a maximum of 128 vCPUs and 6 TB of RAM when it was released at the end of 2016. The capability of vSphere virtual machines has grown as servers have increased capacity and as large application needs have increased. The tests conducted in VMware labs shows how the performance of large Oracle databases running in monster VMs has scaled along with new server generations.

Combining the key elements of vSphere, Oracle database, IBM FlashSystem storage array, and Intel server processors creates a test environment for demonstrating what is possible within the enterprise today.

vSphere 6.5 is the latest version of the industry-leading virtualization and management platform. In addition to its ability to host large virtual machines, it also includes the tools to manage and effectively run both small and large virtualization environments.

Oracle provides one of the industry-leading database platforms for the enterprise. As a leader in this environment, its databases are used by many large organizations as the basis for information management at the heart of their operations.

The IBM FlashSystem® A9000 storage array used in the testing environment is an example of leading-edge storage performance. It integrates the extreme performance of IBM FlashCore® technology, highly parallel architecture, and comprehensive data reduction into one powerful solution. IBM FlashSystem A9000 is designed to accelerate workloads to help reduce costs, increase revenue, meet service level agreements, and improve customer satisfaction.

Performance testing with large virtual machines using databases on the latest server and storage technology was done with an open source database testing tool. Four generations of server hardware were used for the tests to illustrate how performance has improved over time relative to each new server processor.

Test Environment

Tests were conducted across four generations of servers. All servers were four-socket systems and configured with high performance Intel E7 Xeon processors.

SERVERS TESTED INTEL XEON MODEL	Total Cores/Threads	GHz	Codename	Release Date
E7-4870	40 / 80	2.40	Westmere-EX	2011
E7-4890 v2	60 / 120	2.80	Ivy-Bridge-EX	2014
E7-8890 v3	72 / 144	2.49	Haswell-EX	2015
E7-8890 v4	96 / 192	2.20	Broadwell-EX	2016

Table 1. Types of Intel Processors E7 Xeon tested

The servers were all connected via Fibre Channel to an IBM FlashSystem A9000 high performance, all flash array. All monster virtual machines were on this storage and simply moved between the different servers for each set of tests. Virtual machines were powered down and cold migrated between the hosts to allow them to take advantage of the features of each processor to the fullest.
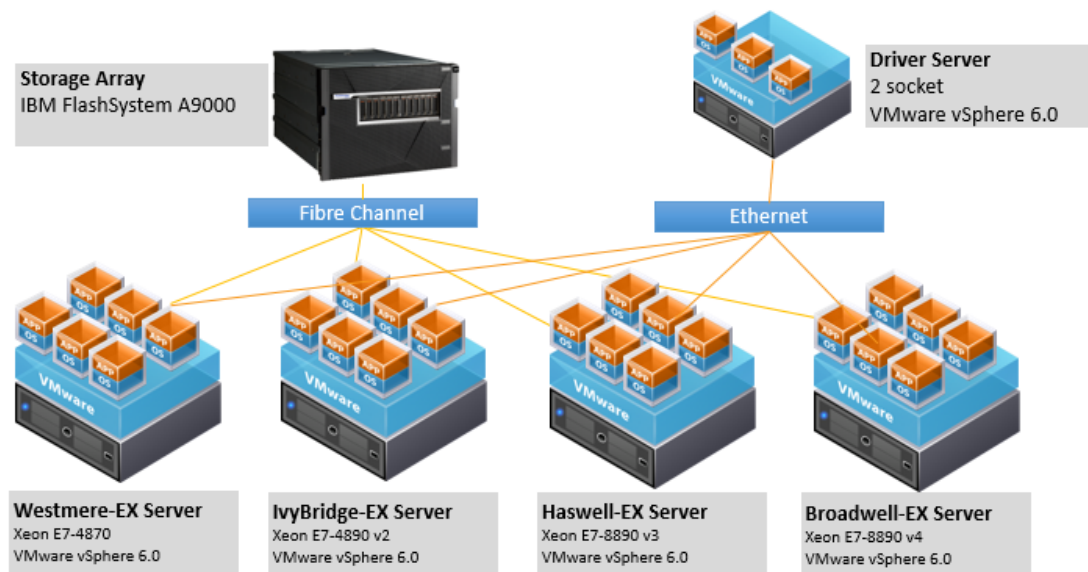

Figure 1. Testbed configuration

Test Workload

The test workload used was the open-source DVD Store 3 [1]. DVD Store simulates an online store that allows customers to log on, search for DVDs, read customer reviews, rate the helpfulness of reviews, and purchase DVDs. Many common database features are used to support these operations including tables, indexes, triggers, stored procedures, transactions, foreign keys, simple queries, and complex multi-join queries. A client program is included that generates load against the database by simulating the actions of users on the site. Each simulated user thread does the full set of steps for a user from logon through purchase. The workload reports throughput in orders per minute (OPM).

For each test configuration, the number of user threads was increased with each iteration until the throughput reached its maximum level. It was found in these tests that the maximum throughput corresponded to host CPU utilization of between 90 and 99 percent.

DVD Store 3 allows for any size database to be created, and, for these tests, the on-disk size was approximately 200 GB.

Virtual Machine Configuration

The configuration of the virtual machines was kept constant except for the number of vCPUs, which were varied depending on the specific test. VMXNET3 virtual network adapter and paravirtual SCSI (PVSCSI) adapters were used. Data and log virtual disks were assigned to separate paravirtual adapters. Memory was configured to be 256 GB.

Red Hat Enterprise Linux (RHEL) 7.1 was used as the operating system, and Oracle Database 12.0.1.2 was installed. The Oracle-provided installer RPM was used and the Oracle best practices for installing were followed according to their documentation. Huge pages were configured for RHEL and enabled in Oracle. The Oracle sga_target parameter was set to 100 GB. This ensured that a large amount of the database would in the buffer cache, but not all so that some disk accesses would occur during the tests.

Generational Monster VM Comparison

The graph below shows the results of running a monster VM with vCPUs equal to physical cores across 4 generations of Intel-based, four-socket servers. The earliest generation server was based on the Intel E7-4870 processor that was originally released in 2011 and was codenamed Westmere-EX. Each processor has 10 cores, which means that the server has a total of 40 cores. The virtual machine, when run on this server, was configured with 40 vCPUs to match the physical core count. The virtual NUMA configuration was set to match the underlying physical server with 10 vCPUs per socket. The same approach was used for the other servers and virtual machines in the test, up to the latest generation server with 96 cores and a virtual machine with 96 vCPUs.

The results show that performance increased over 3x from the Westmere- to Broadwell-based servers. This shows that the performance of monster virtual machines scales with more vCPUs and on newer servers with higher core counts.
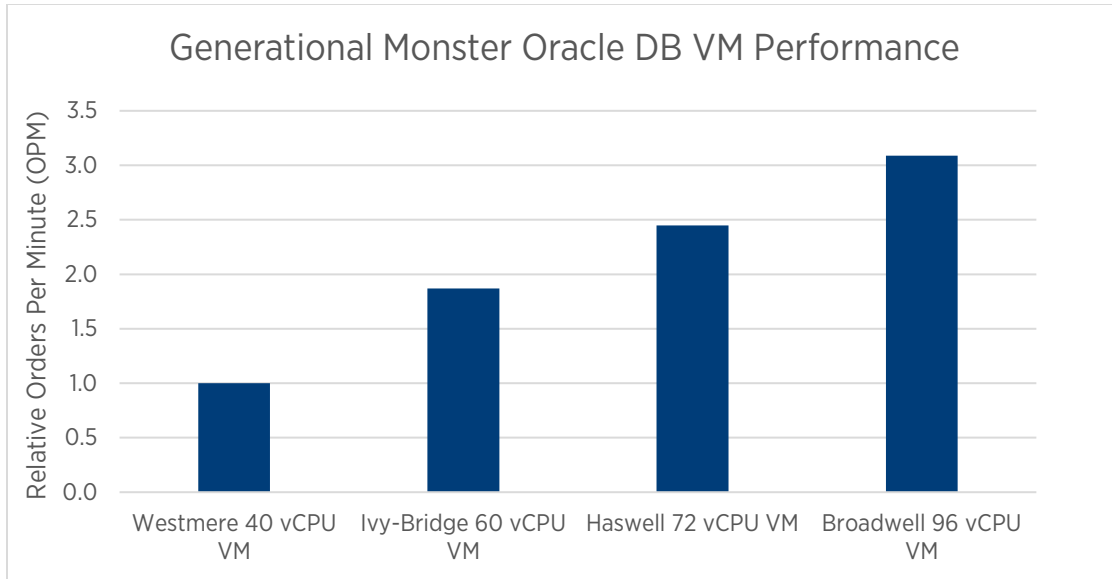
Figure 2. Oracle DB VM performance increased 3x from Westmere to Broadwell servers

Performance Gain from Hyper-Threading

With Hyper-Threading enabled, there are two execution threads sharing a single physical core. Because they are sharing a single physical core, the performance gain is not doubled but is usually in the 10 to 20 percent range. The 96-core Broadwell-based server was used to measure the performance gain from Hyper-Threading in these tests. The performance of a single 96-vCPU virtual machine was compared against the combined performance of two 96-vCPU virtual machines. Using two 96-vCPU virtual machines means that there are 192 vCPUs active and all 192 execution threads on the server are active. The gain in performance due to hyper threading was 11%, as shown in the graph below.
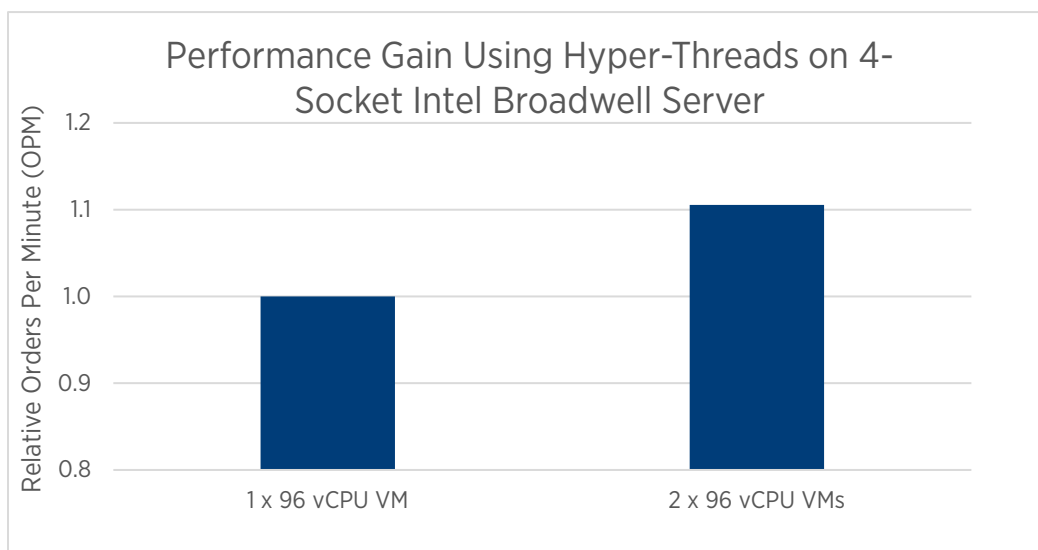

Figure 3. Performance gain was 11% using Hyper-Threads

NUMA Performance

Today's servers use Non-Uniform Memory Access (NUMA) in their architecture. This means that each processor has memory that is directly connected to it. The combination of a processor and its memory is known as a NUMA node. Performance can be optimized by reducing the number of remote memory accesses for a process from the local NUMA node it is running on.

A best practice for virtual machines is to right-size them to the amount of CPU and memory that is needed, and to keep the NUMA node boundaries in mind when doing this sizing. In other words, assign the amount of vCPUs and memory that are needed for the VM, but limit this to what is really needed. If the size of the VM can be kept to equal or less than the size of the NUMA node, then performance can be optimized, as remote NUMA node memory access is limited or eliminated.

This aspect of NUMA performance was measured on the 4-socket (or 4 NUMA node) Broadwell server. Four 24-vCPU virtual machines were tested and compared against the performance of a single 96-vCPU VM. The results in the chart below show that the total throughput of the four VMs was 15% higher than the throughput of the single VM.
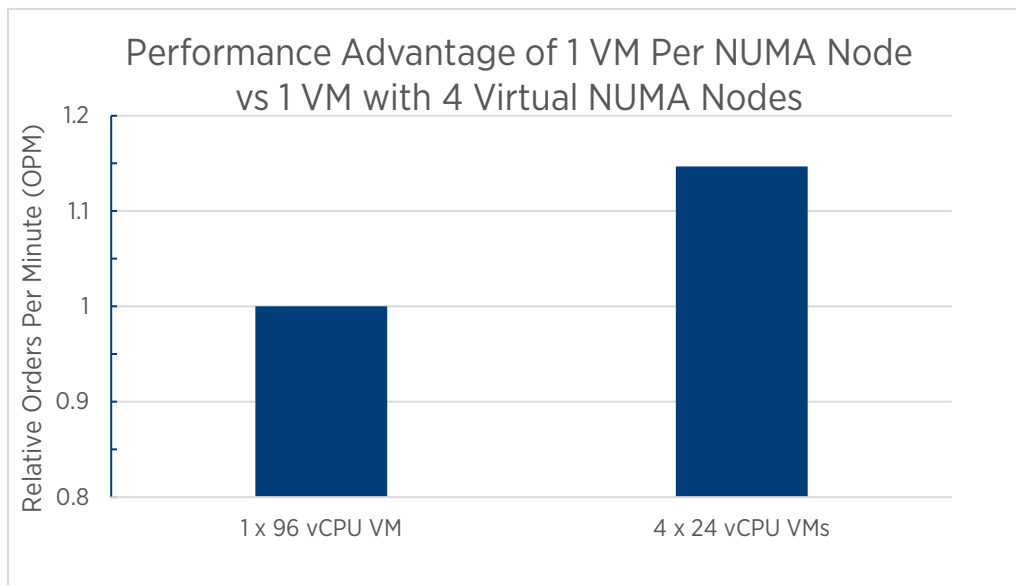
Figure 4. Total throughput of four VMs is 15% higher than the throughput of one VM, where the logical total of vCPUs in each case is 96

Best Practices

The VMware-published Oracle Databases on VMware Best Practices Guide [1] is the best source for complete details on best practices. In this section, the key best practices for large virtual machines and Oracle databases are summarized, but for more details the full best practices guide should be consulted.

CPU

The latest generation server CPUs provide the best performance. Whenever possible, use the latest generation processors to get the best performance for database virtual machines.

Right-sizing of vCPUs and memory for the needed load is important to achieve best performance. Monster virtual machines can provide great performance, so assign the database virtual machine all the vCPUs that it needs. However, overprovisioning virtual machines is counterproductive to achieving best overall performance. Keeping the physical NUMA node sizes in mind when sizing virtual machines is a good guide. Limit the number of physical NUMA nodes to a minimum that a virtual machine will span to get the best performance.

Do not use CPU affinity. While it is possible to achieve some small gains in performance with CPU affinity in some cases, it is more likely to cause problems and performance issues than it is to provide significant benefit.

Enable Hyper-Threading on the host. It has been found that Hyper-Threading can improve host performance by 10 to 20 percent with many database workloads.

Memory

Size memory based on NUMA node size. Overallocation of memory to a virtual machine beyond what is needed can result in crossing physical NUMA node boundaries unnecessarily. Assign the needed memory, but don't over allocate.

Set a memory reservation for the virtual machine equal to the Oracle SGA. This will ensure that the memory needed for the buffer cache of the database is reserved and will ensure good performance.

Use large memory pages. Large memory pages are more efficient and improve the performance of databases due to their heavy usage of memory. In Linux, large pages are referred to as huge pages.

Storage

Use dedicated LUNs or datastores for data and log virtual disks. Ensure that the storage being used for the data and log virtual disks is not shared with other virtual machines or applications.

Configure all database virtual machines with only paravirtual SCSI adapters.

Use VMFS for all vSphere datastores. Tests have shown that VMFS and RDM provide essentially the same performance, but VMFS is much easier to manage and maintain.

High performance all-flash storage should be used when possible. Oracle databases greatly benefit from the low latency performance of high-end storage arrays. The IBM FlashSystem A9000 used in these tests maintained low latency performance during all tests, enabling excellent performance for each VM configuration.

Oracle Database

Follow Oracle best practices within the virtual machine. Oracle inside a properly configured virtual machine behaves like Oracle on a bare metal system.

Evaluate the use of Oracle's NUMA support before using it. The default value for Oracle's NUMA support is disabled, and in testing with the DVD Store it was found that this default value of disabled provided the best performance. It is important to note here that NUMA can be enabled or disabled also at the virtual machine and physical server levels. The best combination for performance is to leave it enabled for virtual machine and physical server levels, while leaving it disabled for Oracle. These are the default settings.

Conclusion

Testing shows that running Oracle databases on monster virtual machines provides excellent performance. By using established best practices and using the latest server and storage technology, virtual databases provide the needed performance capacity for enterprises of any size.

References

[1] Todd Muirhead. (2017, April) DVD Store 3.
https://github.com/dvdstore/ds3

[2] VMware, Inc. (2016, May) Oracle Databases on VMware Best Practices.
https://www.vmware.com/content/dam/digitalmarketing/vmware/en/pdf/solutions/vmware-oracle-databases-on-vmware-best-practices-guide.pdf

About the Author

Todd Muirhead is a performance engineer at VMware and has worked with many different databases, servers, and storage systems to develop best practices and performance guides. He is also the co-creator and maintainer of the DVD Store open source benchmark.

Acknowledgements