SQL Server Consolidation Using Cisco Unified Computing System and Microsoft Hyper-V

White Paper

July 2011

Contents

Executive Summary

This white paper details a study of Microsoft SQL Server consolidation on Cisco Unified Computing System (UCS). The primary objective of the study is to articulate the total cost of ownership (TCO) and return on investment (ROI) that can be achieved by companies who consolidate SQL Server on the Cisco Unified Computing System. Additionally, the study will also show that consolidated SQL Server implementations on the Cisco Unified Computing System can meet the scalability, availability and performance requirements mandated by today's high-volume database implementations.

Along with providing a framework for choosing among virtualization, multi-database, and multi-instance consolidation strategies for SQL Server Database Engine supporting OLTP applications by highlighting some of the key decision points based on technical analysis, this white paper also provides guidance and recommendations to help customers design a consolidation and virtualization strategy for deploying Microsoft® SQL Server® software on the Cisco Unified Computing System.

This experiment will show that the Cisco Unified Computing System can help companies realize a reduction in total cost of ownership (TCO) and achieve the level of infrastructure agility required to meet the challenges of today's fast paced and ever changing business requirements.

Introduction

In today's economic climate, enterprises are taking a closer look within their IT organizations to identify potential areas in which cost-saving strategies can be implemented to help reduce operating expenses. One of the challenges that IT organizations face today is how to develop an infrastructure that allows flexibility, redundancy, high-availability, ease of management, security, and access control while at the same time reducing costs, hardware footprints and complexity.

The growth in IT industry has realized tremendous growth in hardware computing capacity, database applications and physical server sprawl has resulted in costly and complex computing environments containing many over-provisioned and under-utilized database servers. Many of these servers implement a single instance of SQL Server realizing ten to fifteen percent CPU utilization on average which is not an optimal use of server resources. Additionally, due to the complexities introduced by all of this growth, many database administrators today are overburdened with redundant management and administrative tasks that must be implemented on each of the many servers they are responsible for. Also, during a catastrophic, physical database server failure, there is a significant impact on the sever administrator as they typically are the resources that are required to provision another physical server into the environment and prepare it for the application.

Database server consolidation is an area where companies can realize considerable cost savings with regards to total cost of ownership. Database server consolidation can also help companies to achieve the infrastructure agility they are seeking to stay competitive and provide the fastest time to market for their solutions.

Consolidation, in general terms, is the combining of various units into more efficient and stable larger units. When applied to an IT department, consolidation specifically translates into improved cost efficiency from higher utilization of resources, standardization and improved manageability of the IT environment, and recent focus on a "green" IT environment through reduced energy consumption. One of the important components in the IT environment is the database. Databases managements systems form the foundation of many business systems. Often each group or application create their own database to solve a specific problem, thus IT departments lose control of the number of databases and servers that need to be maintained. This leads to a proliferation of databases and machines running database instances also known as database sprawl. This makes the databases the prime candidates for consolidation.

Audience and Scope

This white paper is intended for customers, partners, solution architects, storage administrators, and database administrators who are evaluating, planning a consolidation and/or virtualization strategy. It provides an overview of various considerations and reference architectures for consolidated and virtualized SQL Server 2008 R2 deployments.

Today's Challenges

IT organizations face an enormous challenge to keep operations running around the clock with increasing demand and growing complexity, to the point that it becomes difficult for employees to request resources to fulfill their tasks. This has lead to circumventing the IT standards and procedures to get the job done faster, which in turn has lead to server sprawl especially at the database tier. Database applications are typically implemented in a three four-tier environment; web-tier, application-tier and database-tier. The tiered architecture allows compartmentalization of applications and separation of resources. Hardware at each tier provides a specific function and requires a specific set of tools to manage and maintain. All of the components are connected via a network, in most cases Ethernet. This separation of resources provides an easy way for employees to bypass and deploy database servers for test and development without the IT department's oversight and management.

The apparent benefit of three-tire architectures like ease of deployment and development can easily be overshadowed by potential challenges such as incompatibility issues that may arise during deployment to production, security vulnerabilities due to different security and patching requirements, and the operating costs associated with large number of servers including support, licenses, power and cooling. Hence servers that were deployed for a specific project that is no longer economical. As the reliability and performance aspects of industry standard x86 server platforms have improved, along with the maturity of virtualization technologies on x86, many organizations are considering virtualization and consolidation for their IT departments.

Why Consolidate?

Consolidation projects are typically started to achieve specific goals such as creating room for new applications while reducing operating expenditure. These goals can be broadly grouped into the following categories:

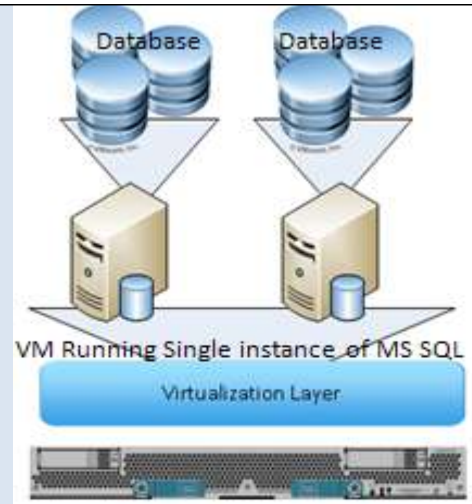
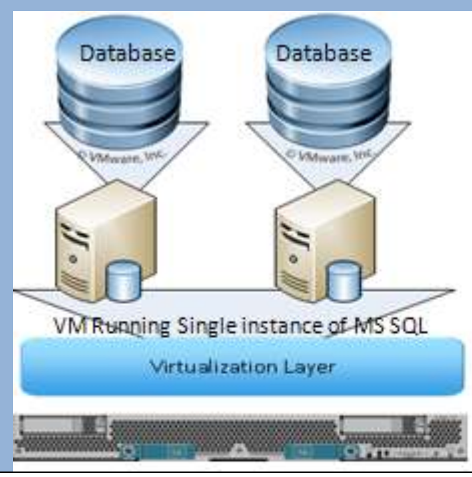
- Standardization and centralization
- Improve floor space and power efficiency
- Reduce the number of management domains IT agility

This paper focuses on consolidation strategies for typical online transaction processing (OLTP) applications based on SQL Server.

Some general traits that make an application a good candidate for consolidation are low machine resource utilization, moderate performance requirements, little active development, and low maintenance costs. Another factor to consider is the impact on the application's network and I/O latency, because both the network and storage resources become shared as part of consolidation.

When consolidating database applications, we have considered the following two potential scenarios (Figure 1).

Figure 1. Graphical presentation of Physical machine, Virtual machine and databases

Multiple virtual machines per physical server hosting multiple databases in each VM	
Multiple virtual machines per physical server hosting single large database in each VM	

Option 1: Using a single physical machine to host multiple virtual machines and each virtual machine running a Microsoft SQL Server instances hosting multiple databases.

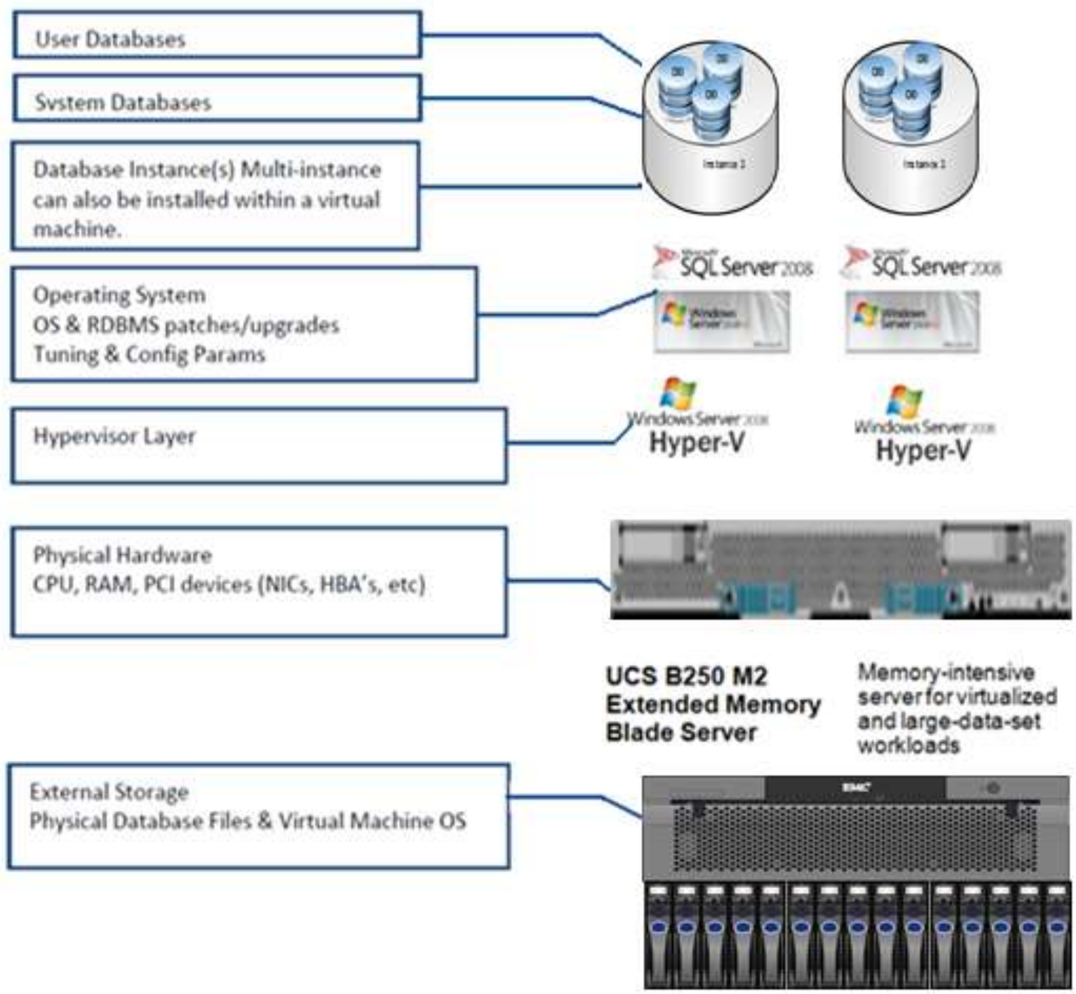
Option 2: Using a single physical machine to host multiple virtual machines and each virtual machine running Microsoft SQL Server database instance hosting a single database.

Each of these scenarios have their own security, compliance requirements, high availability, disaster recovery requirements, resource management benefits, level of consolidation density, and manageability characteristics.

Both these options have been tested in this study. But for option 1 we used a small database of 10 GB in size and for Option 2 we used a large database of 100 GB in size.

Figure 2 shows hosting of multiple virtual machines in a physical machine and each virtual machine has its own SQL Server Instance running on it.

Figure 2. SQL Server Consolidation using Virtualized Environment



Cisco UCS B250 M2 Blade Servers

The Cisco UCS B250 M2 blade servers with their unprecedented memory footprint will allow companies to consolidate and scale up their database implementations helping them to achieve higher densities and consume less datacenter real estate than ever before.

Additionally, the Cisco Unified Computing System Manager (UCSM) will empower database administrators and Server Administrators with an additional level of management and agility that will greatly simplify their management tasks especially in the areas of virtualization and high-availability.

The Cisco UCS B250 M2 Extended Memory Blade Server helps maximize performance and capacity for the most demanding virtualization and large dataset applications. The Cisco UCS B250 M2 server extends the capabilities of the Cisco Unified Computing System, using Intel's latest Xeon 5600 Series multicore processors to deliver even better performance and efficiency. Following are the most important features help to deploy database applications in a virtualized or non-virtualized environment using Cisco UCS B250 M2.

- Up to two Intel Xeon 5600 Series processors, which adjust server performance according to application needs
- Up to 384 GB memory technology, for demanding virtualization and large dataset applications or a more cost-effective memory footprint for less demanding workloads

- Two Small Form-Factor (SFF) Two dual-port mezzanine cards of up to 40 Gbps of I/O bandwidth. Mezzanine card options include Cisco UCS VIC M81KR Virtual Interface Card, converged network adapter (Emulex or QLogic compatible), or Intel 10GB Ethernet Adapter.

Because of large memory architecture, you can deploy as many as virtual machines depending on their business requirement.

Microsoft SQL Server 2008

Microsoft SQL Server 2008 is a comprehensive database platform that offers the security, reliability, and scalability to support mission-critical applications. Responding to exceptional database growth, Microsoft SQL Server 2008 features a comprehensive set of services and tools to support any data type or device. Microsoft SQL Server 2008 also provides improved resource utilization, enhanced locking, and optimized data storage, yielding substantial performance and scalability. An innovative policy-based infrastructure simplifies data platform management, with improved performance monitoring and reporting tools in Microsoft Performance Studio. Microsoft SQL Server 2008 also provides high availability with improved database mirroring and failover clustering, and it takes advantage of the inherent strengths of Microsoft Windows Server 2008.

Microsoft SQL Server 2008 provides a comprehensive data platform that supports the scalability, availability, security, and manageability needed for mission-critical OLTP applications:

- Scale and performance: With Microsoft SQL Server 2008, organizations can build database solutions with the performance and scalability capabilities that are required by the most demanding modern applications.
- High availability: Always-on technologies in Microsoft SQL Server 2008 provides high availability for databases, while reducing the management and performance overhead required for high-availability operation.
- Security: Microsoft SQL Server 2008 provides an enhanced secure data platform by encrypting valuable data, auditing changes to data and metadata, incorporating external cryptographic keys, and encrypting and signing data in backup files.
- Manageability: By providing innovative and automated policy-based administration, Microsoft SQL Server 2008 lets companies reduce the time and cost of managing their data infrastructure. Improved tools are provided for performance monitoring, troubleshooting, and tuning.

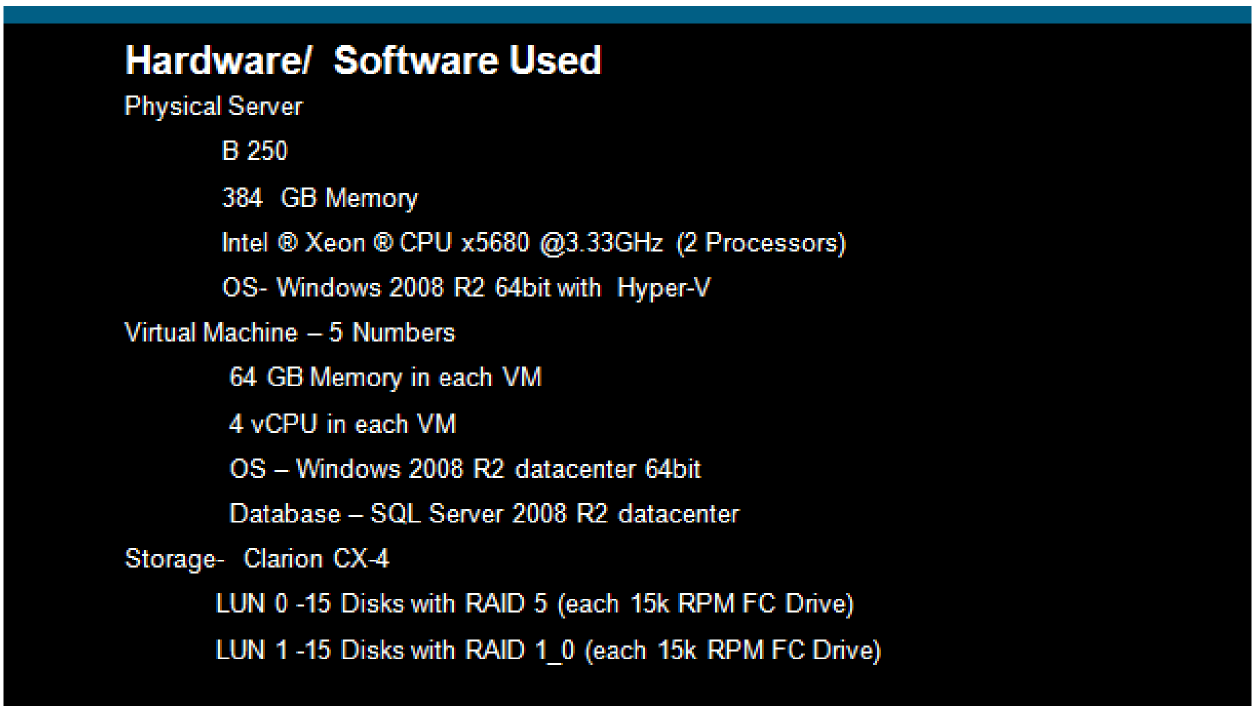
Microsoft SQL Server 2008 R2 is a complete set of enterprise-ready technologies and tools that help organizations derive the most value from information while lowering total cost of ownership (TCO). With Microsoft SQL Server 2008 R2, organizations can:

- Implement peer-to-peer replication quickly with the new visual designer and add nodes without affecting system availability
- Audit all actions across the enterprise and consolidate audit reporting
- Help protect sensitive data with automatic, transparent data encryption
- Use Microsoft Performance Studio to troubleshoot, tune, and monitor Microsoft SQL Server 2008 instances across the enterprise
- Take control of workload resource use with Microsoft Resource Governor
- Store all kinds of business data with native support for relational data, XML, file streams, and spatial data
- Reduce storage requirements and improve performance with data compression, backup compression, sparse columns, and filtered index and backup compression

- Optimize database mirroring performance and eliminate downtime with automatic recovery of suspect pages

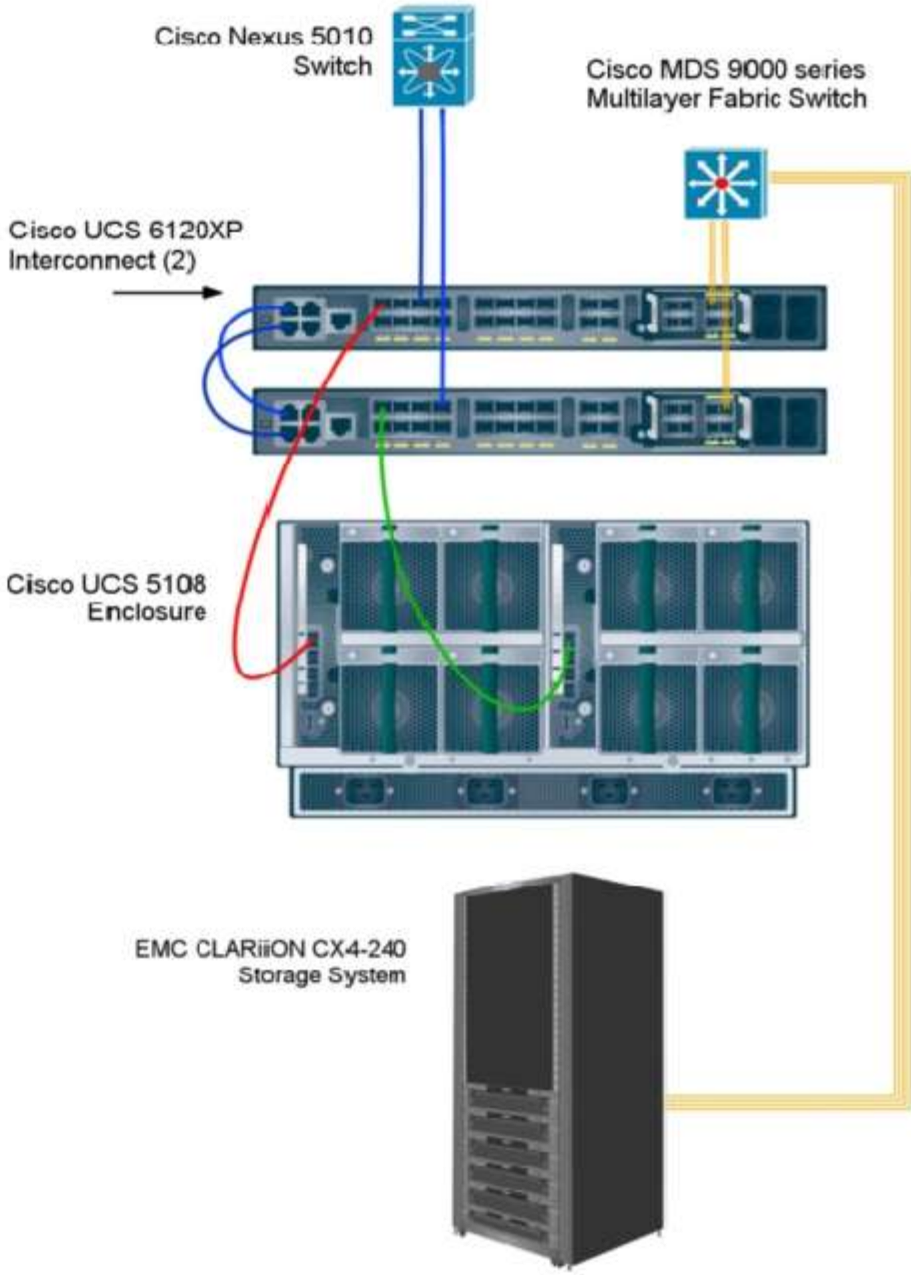
Tested Configurations and Methodology

Tests were conducted on a fully loaded blade server configuration consists of Cisco UCS B250 M2 Extended-Memory Blade Server equipped with two Intel® Xeon® Processor X5680 (3.333 GHz, 12MB L3 Cache, 130W), 384 GB (48 numbers of DIMM of 8 GB in each size) of memory and two QLogic HBA cards with 4Gbe bandwidth in each. Microsoft Windows 2008 R2 64bit operating system installed on the Server with Hyper-V enabled. Created multiple virtual machines and installed Windows 2008 R2 64bit operating system with SQL Server 2008 R2 datacenter edition.



Cisco UCS Chassis containing B250 M2 server connected to FIs (Fabric Interconnect), Nexus 6000 based Fabric Interconnect (FIs) which are then are connected to a specific zone of MDS where the storage is connected. Figure 3 depicts interconnect between different hardware components hosting the SQL Server databases.

Figure 3. Test bed configuration



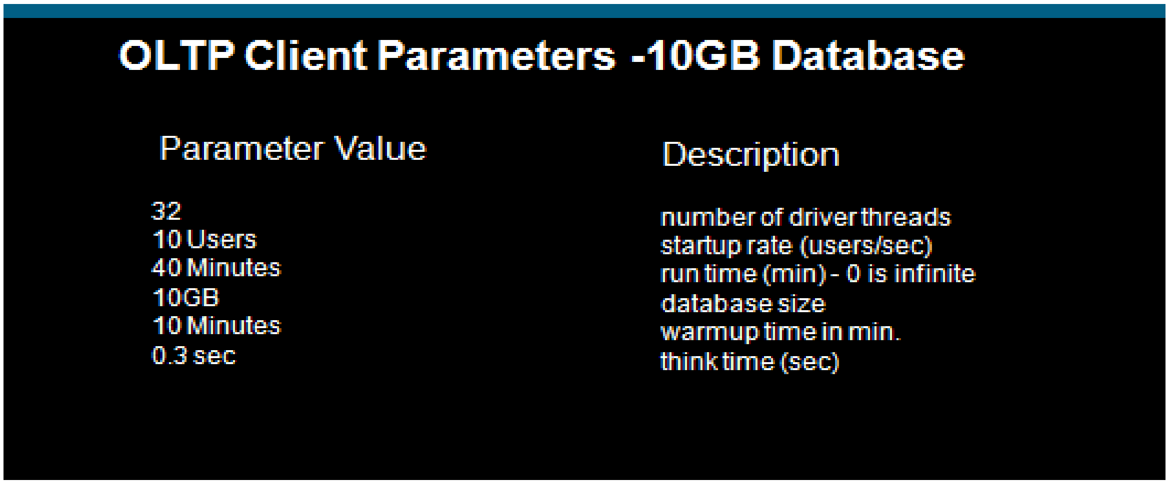
An OLTP workload was used to evaluate the performance and scalability of the server.

One client machine running with Windows 2008 Operating System was used to drive workload for the Database Server, used in this test configuration. From this client machine, tests were run to push the workload to the SQL Server database on virtual server.

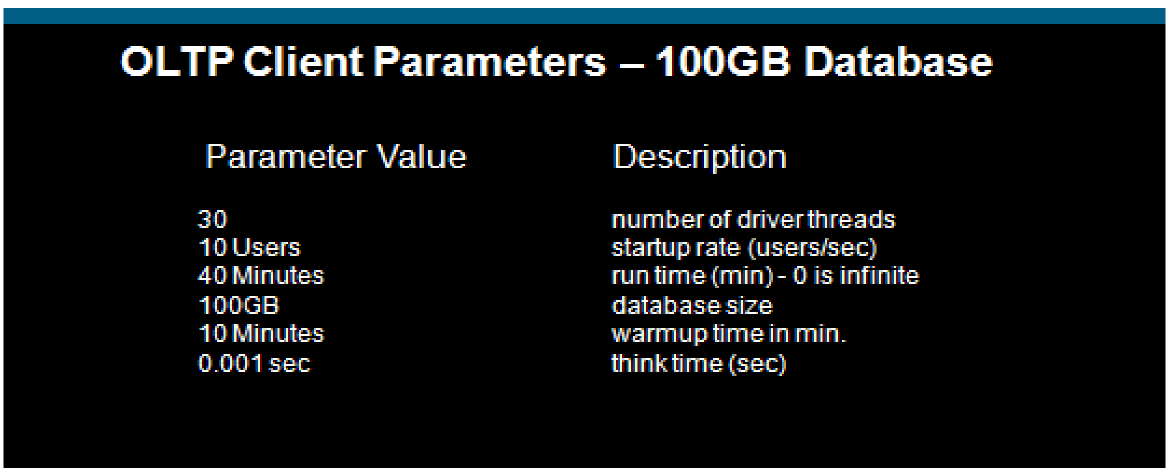
In each test case, processor utilization and memory utilization were captured along with number of operations per minute (OPM). Each test was run for 30 minutes with 15 minutes ramp-up time and the utilization of each virtual server as well as physical server were captured using windows perfmon tool.

The tests were conducted in two different scenarios as described in the previous sections.

Scenario 1: One SQL Server Instance was running on each virtual machine, hosting six numbers of 10 GB size of databases. We used following parameters from the Client machine to push the OLTP workload to each databases.



Scenario 2: One SQL Server Instance was running on each virtual machine, hosting one 100 GB size of database. We used following parameters from the OLTP Client machine to push the OLTP workload to each database.



The following table shows virtual machine configuration details.

Table 2. Virtual Machine Configurations

PHYSICAL MACHINE	VIRTUAL MACHINE	DATABASES
B250 M2 (two 6 core Intel Xeon X 5680 CPU of 3.33 GHz speed each, 384 GB memory of 1067 MHz speed each DIMMs).	**Each Hyper-V VM with 4 vCPU and 64 GB Memory.**	**6 Databases in one virtual machine with each database of size 10 GB .**
B250 M2 (two 6 core Intel Xeon X 5680 CPU of 3.33 GHz speed each, 384 GB memory of 1067 MHz speed each DIMMs).	**Each Hyper-V VM with 4 vCPU and 64 GB Memory.**	**One Databases in one virtual machine with each database of size 100 GB .**

Workload

To build the workload, we used a typical OLTP transaction involving ordering, searching and delivery. We used different parameters such as think time, thread, etc. for two different sizes of database to saturate each virtual machine. We used 32 numbers of concurrent thread and 0.3 seconds of think time for the small database (10GB) and 30 numbers of concurrent threads and 0.001 sec think time for large database (100GB).

Cisco UCS B250 M2 blade servers were used for this testing running Microsoft Windows 2008 R2 datacenter with Hyper-V enabled. In this consolidation process we created up to five virtual machine with four vCPU and 64Gb memory in each virtual machine. Microsoft SQL Server 2008 datacenter version was installed on each virtual machine.

We tested with two different sizes of databases in two different scenarios, in this consolidation process. In the first scenario we created a small-sized database, for small and medium scale businesses, and in the second scenario, we created a large database for large scale businesses. In both these scenarios, each virtual machine had one SQL instance running on it. Up to six databases of small-sized databases in SQL instance running on each virtual machine was tested in the first scenario. In second scenario, we created one copy of large-sized database in one SQL instance running on each virtual machine. All data was contained in a storage system and was accessed through FC (Fiber Channel). The client machine was running windows 2008 to initiate the workload for each database.

In the first scenario we started the test with one virtual machine and one database. One SQL server instance captured the benchmark score and resource utilization. Later, we added one database at a time to the same SQL server instance on the same virtual machine and simultaneously started the workload on all the databases on that virtual machine. We tested up to six databases on the same virtual machine and captured the benchmark score at every instance of adding new to database to the virtual machine. The result was that throughput scaled linearly up to six databases in one single SQL Server instance on one virtual machine. We then added more virtual machines to the configuration with each virtual machine now supporting 6 Databases under one SQL server instance. Benchmark scores and resource usages were collected with this configuration as well.

In the second scenario we started the test with one database in one SQL instance on one virtual machine and captured the benchmark score and resource utilization. Later we started adding one virtual machine at a time to the same physical machine and simultaneously started the workload on all the virtual machines. We tested up to five virtual machines on the same physical machine and captured the benchmark score at every instance of adding new virtual machine. The throughput scaled linearly up to five virtual machines.

We initiated the workload on both the scenarios. Performance of the blade server was measured, especially as it relates to an active OLTP environment.

The primary metric is OPM, which the driver program calculates and reports in a test file on the client machines.

We ran this workload on all virtual machines simultaneously for 30 minutes with 15 minutes of warm-up time. We considered a blade server to have delivered acceptable performance since the total OPM for the server did not decrease and since the average OPM results across all clients at the end of the test was above the baseline.

Test Results

Figure 4 shows the results of scenario 1. It reports number of recorded from each database. This result was captured for one virtual machine after adding one DB at a time to the same SQL server instance on one virtual machine and simultaneously started the workload on all the DBs on that virtual machine. As it shows in graph, OPM linearly scaled up to six databases after adding one after another database into the same SQL server Instance running on one virtual machine .

Figure 4. OPM and database scaling

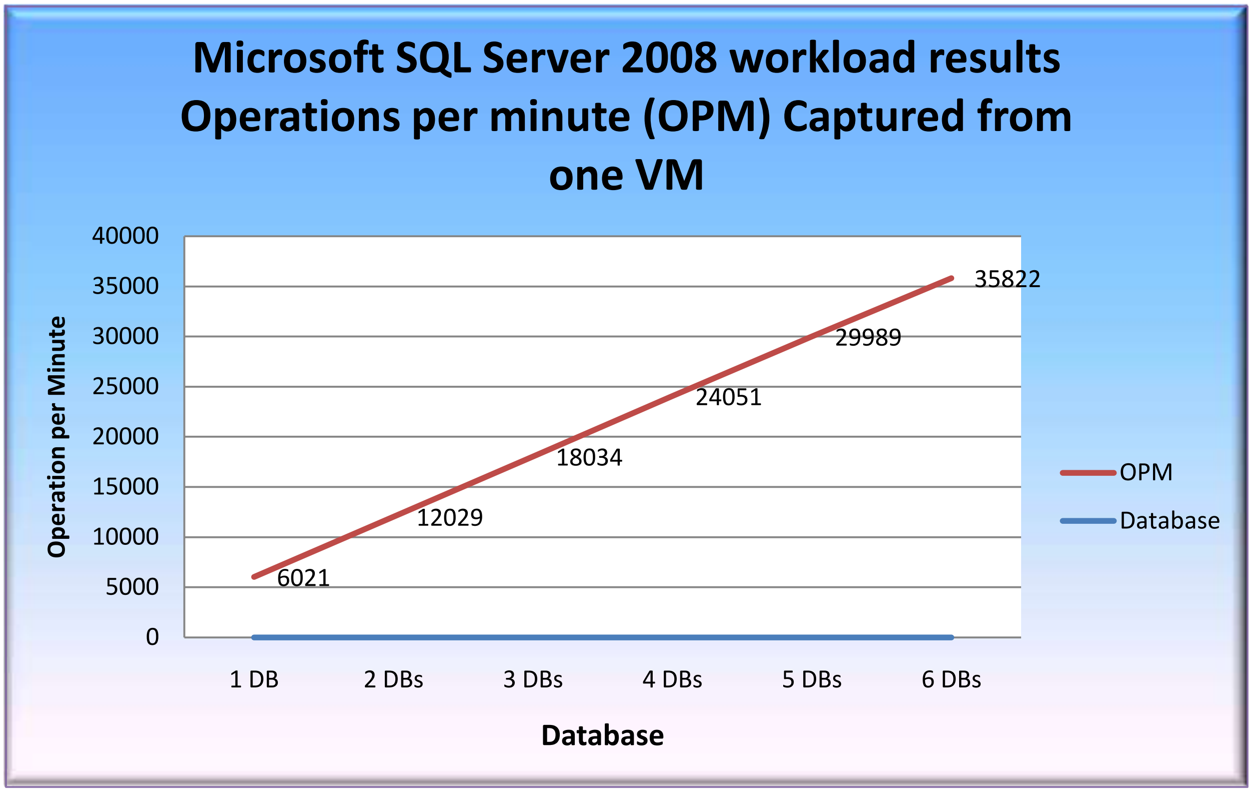


Figure 5 shows the resource utilization, especially CPU and Memory of virtual machine and physical machine after adding each database to the same virtual machine mentioned in scenario 1. We observed up to 85 percent of memory utilization and 75 percent of CPU utilization of the virtual machine, which is the datacenter standard for maximum resource utilizing.

Figure 5. CPU and Memory utilization of virtual machine after adding each database

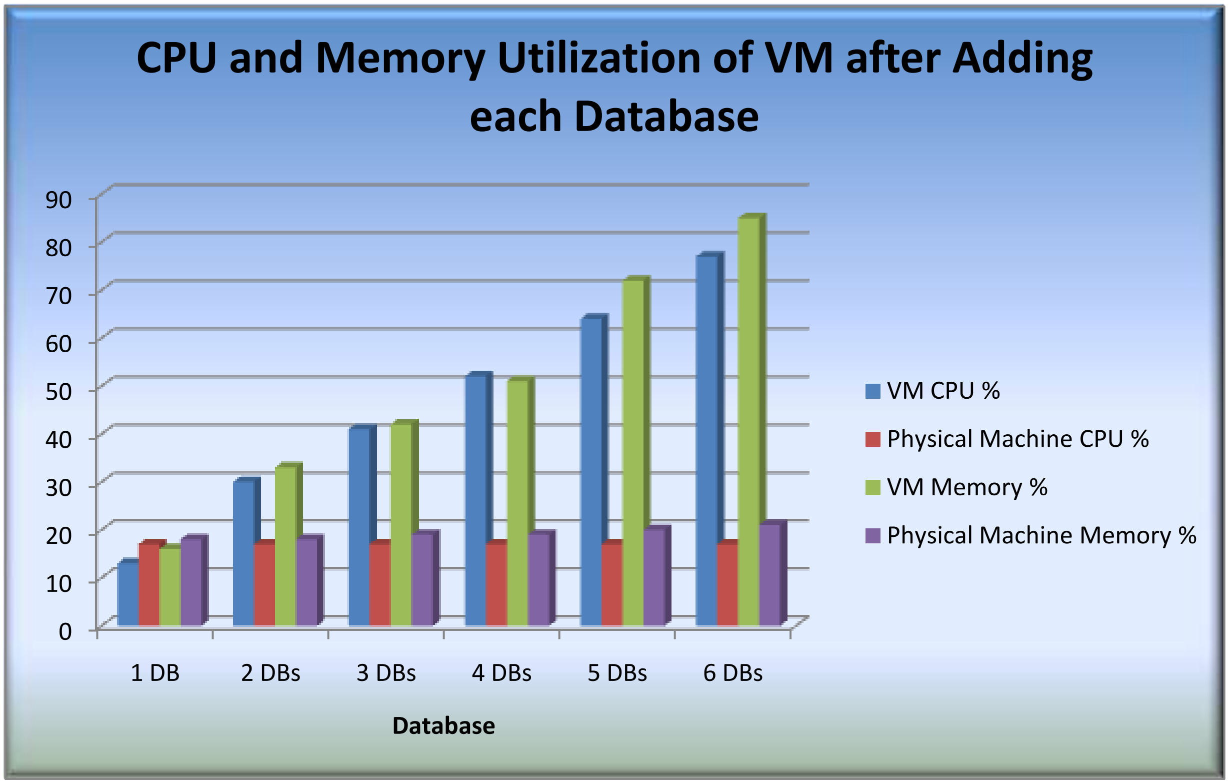


Figure 6 shows OPM recorded in scenario 1 after adding each virtual machine into the same physical machine. In this scenario we added six databases to each virtual machine and there were five virtual machine created on the same physical machine. We found linear scaling of OPM as well as the linear scaling of resource utilization.

Figure 6. OPM scaling after adding each virtual machine

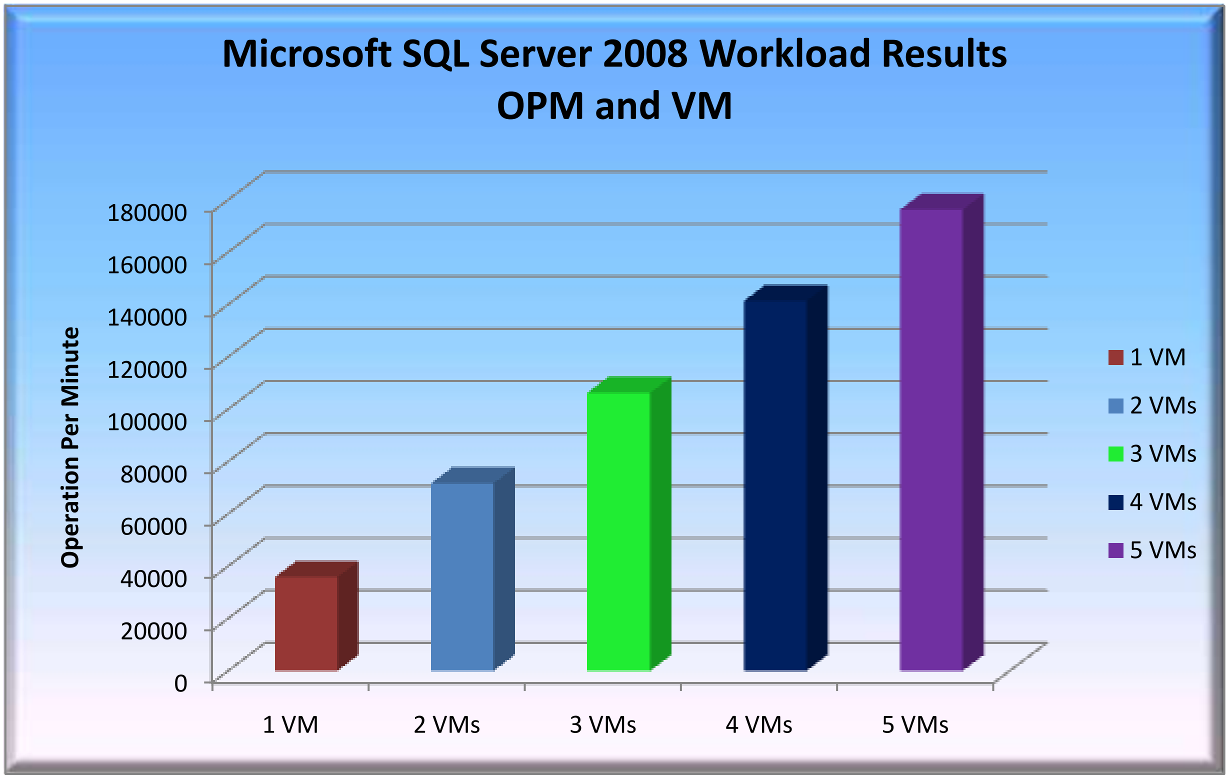
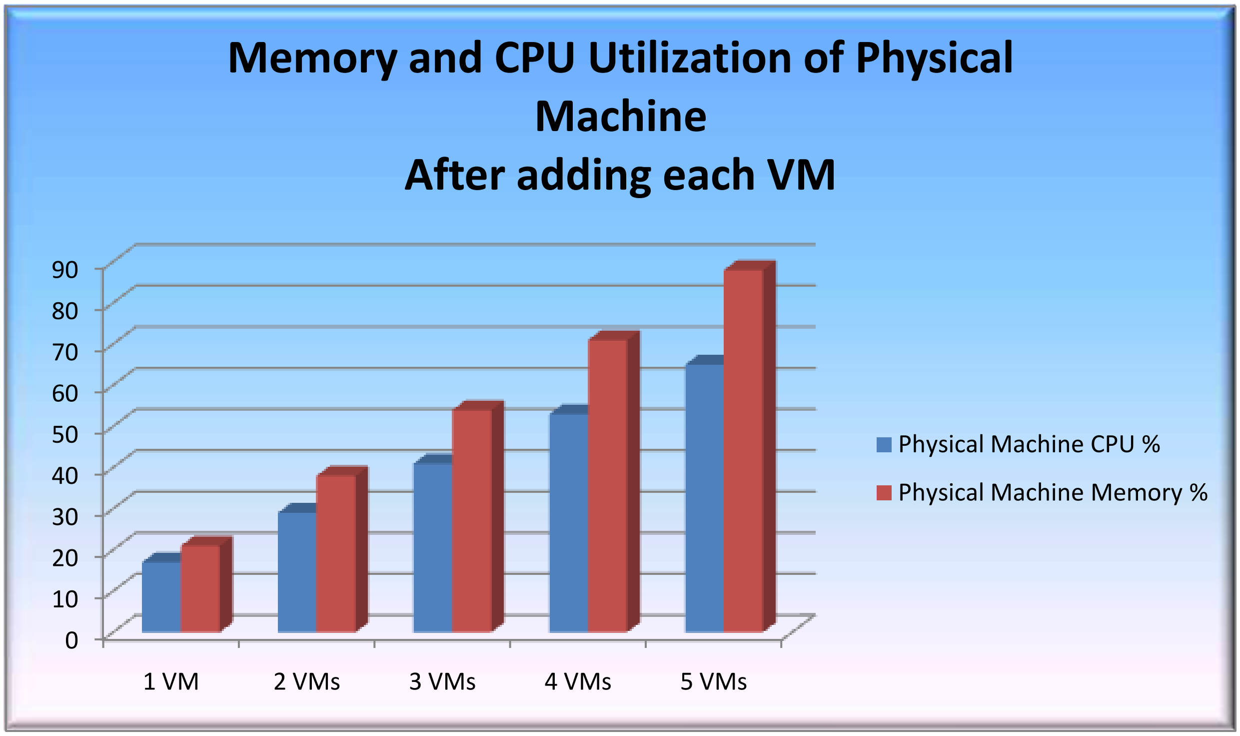


Figure 7 shows the resource utilization, especially CPU and Memory of Physical machine after adding each virtual machine one after another with six Databases in each virtual machine and concurrently running OLTP workload one each virtual machine. Workload was capped to make sure that resources, memory and CPU, were not saturated. In this testing process we observed 85 percent of memory utilization. As it is shown in the graph, physical machine still had enough CPU to handle any unexpected spikes in the workloads.

Figure 7. CPU and Memory utilization of Physical Server after adding each virtual machine



Scenario 2

Scenario 2 as described in workload section, we started test with one large database in the SQL Server Instance running on the virtual machine. We captured the test results at the end of test. Later we keep on adding one virtual machine at a time with the same memory and database instance. At every virtual machine count we captured metrics of the workload. It was observed that the best utilization of CPU and Memory of each virtual machine as well as Physical machine was sustained until five virtual machines were deployed on the same physical machine.

Figure 8 shows the linear scaling of CPU and memory after adding each virtual machine to the same physical machine.

Figure 8. CPU and Memory utilization of Physical Machine after adding each virtual machine

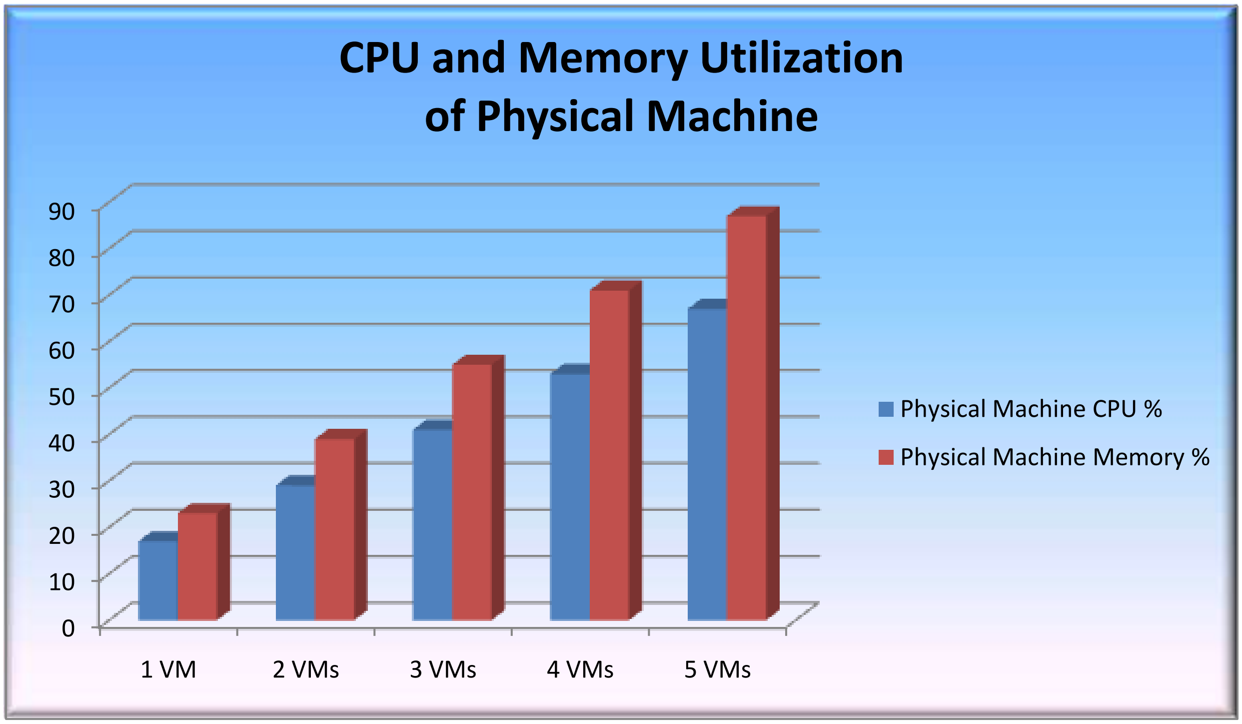
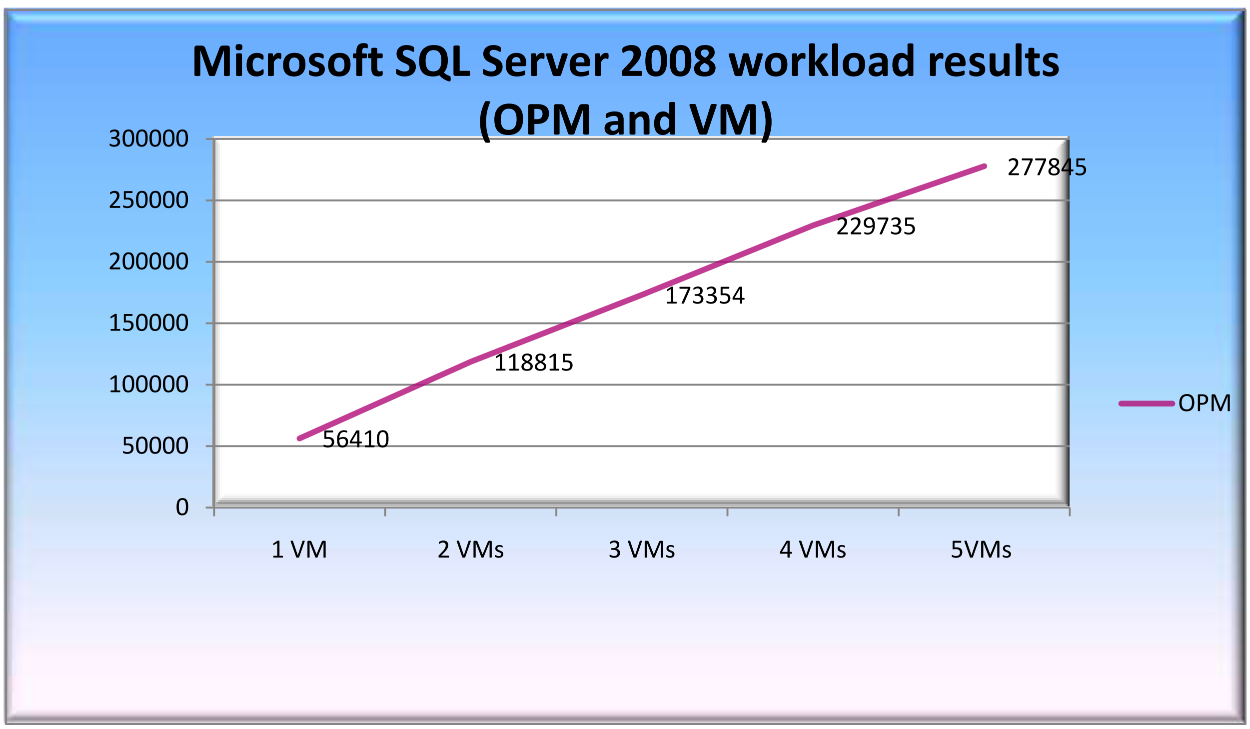


Figure 9 shows the OPM captured in scenario 2 after adding each virtual machine.

Figure 9. OPM scaling after adding each virtual machine



Conclusion

This study provides an overview of considerations to develop a strategy to maximize hardware utilization and reducing costs associated with database sprawl. It describes the strategies and reference architectures as a starting point to consolidate SQL Server using a Cisco blade server implementing a building block approach to design, configure, and deploy using best practice recommendations to simplify IT.

To simplify the design and deployment of a virtualized infrastructure, Cisco offers solution architectures bundles for blade servers, Microsoft Hyper-V. The bundles provides configuration and best practices to achieve full redundancy—with no single point of failure, scalability, and ease of management.

The tests performed in Cisco labs showed that significant gains can be achieved by developing a strategy to maximize hardware utilization, reducing sprawl, power and cooling costs by consolidating and virtualizing SQL Server on the latest Cisco blade servers, while providing the performance to meet the most demanding customers' workloads.

For More Information

For more information about the Cisco and Microsoft alliance, visit http://www.cisco.com/go/microsoftalliance or e-mail cisco_microsoft_info@cisco.com.

Printed in USA

Document number: UCS-TR1000xx 08/10