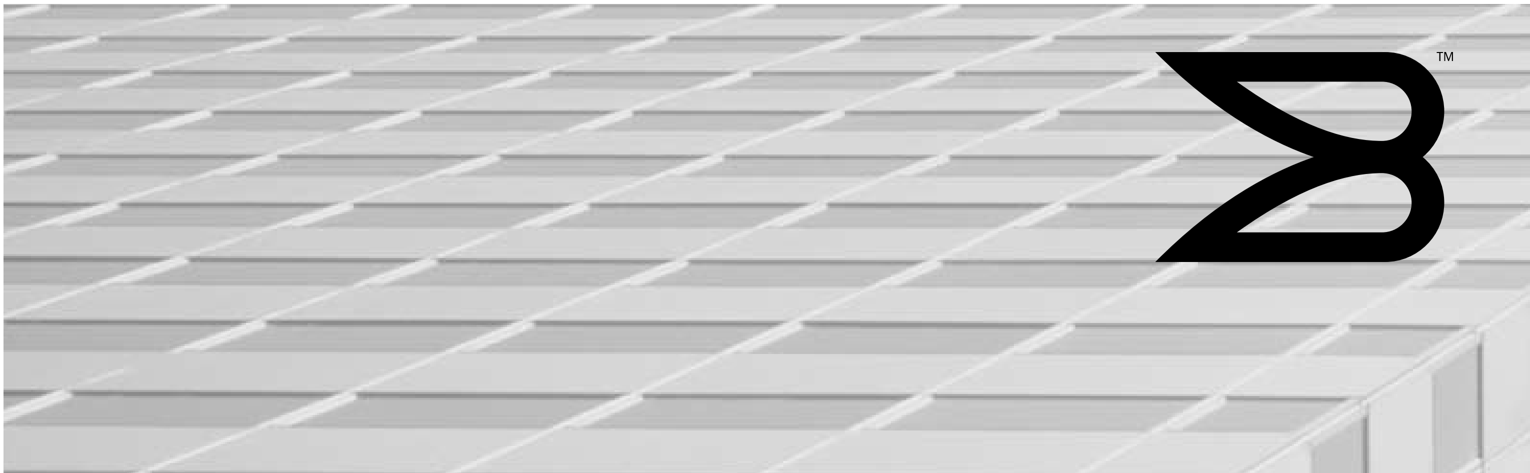
WHITE PAPER



DATA CENTER FABRIC

Benchmarking Microsoft SQL Server Using VMware ESX Server 3.5

The results of a benchmarking study performed in Brocade test labs demonstrate that SQL Server can be deployed on VMware ESX Server 3.5 for Online Transaction Processing (OLTP) applications in very favorable server consolidation ratios to meet corporate IT business requirements.

vmware®

BROCADE

The astounding success of the virtualization of applications on the VMware® ESX Server platform has changed the character of IT infrastructures worldwide. Organizations today are continuing to deploy applications using virtual machines (VMs) on ESX Servers at a record pace.

There are, however, classes of applications that organizations have been unwilling to move to virtualized environments. One example is, Online Transaction Processing (OLTP) applications running on Microsoft SQL Server database engines. These applications are characterized by high levels of Input/Output (I/O), in which the number of transactions is very high and the amount of data transferred per transaction is very low. Database Administrators (DBAs) suspect that these applications will put too much stress on ESX Server and, as a result, are somewhat hesitant to deploy I/O-intensive applications on virtualized platforms.

ABSTRACT

This paper presents the results of a benchmarking study, in which Brocade® used an industry-standard online transaction processing benchmark (TPC-C) in an ESX Server cluster to determine if ESX Server version 3.5 could run 10 simultaneously active VMs running SQL Server with a significant workload on a single ESX Server and 20 such VMs in a cluster of four hosts. The results demonstrate that ESX Server is more than capable of handling the task. Brocade is very confident that DBAs can deploy SQL Server on ESX Server 3.5 for OLTP applications and obtain similar server consolidation ratios, as we observed, which are very favorable and that make strong business sense.

Included in this paper is supporting performance data for ESX Server and the VMs running under ESX Server, which clearly shows that the workload performed successfully with platform memory, processor, and I/O resources to spare. In addition, the data shows that there is no evidence that ESX Server is a bottleneck in any way.

BACKGROUND

Deploying virtualized servers and applications in volume on VMware ESX platforms started more than two years ago. Even though the same application consolidation ratios apply to I/O-intensive applications as other applications virtualized today, end users have been reluctant to deploy I/O-intensive applications on ESX because of the perceived risk associated with those deployments. The general perception is that the ESX I/O abstraction layers cannot support the I/O throughput required to service I/O-intensive applications. Since many of these servers are considered much more critical to the enterprise, prevailing opinion is that it is better to avoid the perceived risk by continuing to deploy the applications on dedicated server platforms. If considered at all, consolidation of I/O-intensive applications was and for the most part still remains limited to increasing the number of application instances on more powerful dedicated platforms. This leaves a significant portion of end-user infrastructures, made up of Windows and Intel-based Linux and Sun Solaris servers, which are still not deployed in virtualized environments. And unfortunately, end users are missing the business benefits that would accrue from those deployments.

NOTE: I/O-intensive applications are those that either produce a lot of storage traffic or generate a lot of I/Os per Second (IOPS). Microsoft SQL Server, Microsoft Exchange and Oracle database servers are popular examples of I/O-intensive applications.

Until now there has been very little third-party ESX benchmark data to support the deployment of I/O-intensive applications in virtual environments. What data was available was very limited; typically one guest running a load generator such as IOMETER in a single virtualized instance on a single ESX platform.

When VMware released ESX Server 3.5, Brocade felt it was time to revisit the possibility of virtualizing I/O-intensive applications on ESX Server. At the same time, more ISV vendors now support their applications in virtual environments. The time seemed right to generate data to convince end users to consider deploying more important applications in virtualized environments. Because of its wide deployment in IT environments, Brocade decided to generate benchmark data on Microsoft SQL Server

GOALS

The goal of this study was to provide meaningful data that would be relevant to end users to give them enough confidence to consider proliferating SQL Server in an ESX Server environment. Details were as follows:

- Use an application deployed in large numbers by end users on dedicated platforms today.
- Demonstrate that I/O-intensive applications are viable candidates for virtualization.
- Show an application consolidation ratio that makes business sense.
- Demonstrate that the risk is minimal when the application is virtualized.
- Highlight the probable bottleneck areas of such an implementation.
- Simulate as much as possible in a typical IT environment (resources, knowledge, and infrastructure). Note that the goal of the testing was not to break performance records or set high-water marks.
- Show that results apply equally in enterprise environments and Small-to-Medium Business (SMB) environments.
- Apply no application optimization that might skew results. In other words, a guest OS and ESX Server with an out-of-the-box configuration was used as much as possible.
- Limit the benchmarks to a specific workload type with clearly defined performance criteria.
- Perform benchmark testing that can be duplicated.

VMware has data that suggests that the workload for typical SQL Sever applications in most organizations is quite small: 90 percent are fewer than 20 Transactions per Second (TPS). For the purposes of this benchmark study, we established a performance goal of 50 TPS per virtualized SQL Server instance as measured by TPC-C. This corresponds to 500 concurrent users working at twice the normal speed on each VM running SQL Server, a workload that is very much greater than the average reported by VMware.

METHODOLOGY

An objective benchmark that is well known in the DBA community was used. This, combined with the decision to concentrate on an OLTP workload, led to the choice of Transaction Processing Council TPC-C benchmarks.

There were two phases—single-platform tests and a cluster test—as detailed below:

- A “bare metal” (BM) baseline was established by running TPC-C benchmark on one of the hosts in native Windows before ESX Server was installed. The SQL Server application was limited to one of the Central Processing Unit (CPU) cores and was allowed to use as much memory as it needed.
- ESX 3.5 was then installed on all the four Systems under Test (SUTs) designated to run the SQL Server application.
- In the single-platform phase: 10 separate tests were run on one of the platforms, and with each new test, an additional virtualized instance of SQL Server running the same workload was added. The tests were stopped at 10 simultaneous instances on a single platform due to time constraints.
- For the cluster test: Five instances of SQL Server were run per ESX Server, thus running a total 20 SQL Server instances simultaneously on all four platforms in the ESX cluster. A limit was reached at 5 instances per ESX platform (due to licensing constraints with the benchmark driver product used in testing).

TPC-C Overview

TPC-C is a standard benchmarking tool that emulates an OLTP processing environment, for example, an inventory management system. The tool reproduces new order entries, order status inquiries, and payment settlements and can be considered a true application with data entry. (Input field values, however, are not sanity-checked.)

Multiple warehouse inventories are scanned, stocking levels are checked, and deliveries are scheduled. Individual transactions can generate one or more additional transactions. No batch operations were performed; however, they were simulated by causing an application checkpoint to be performed once during each run.

Only new order transactions are counted in the TPS metric. The actual transaction score would have been higher if all transactions had been counted; TPS should be considered artificially low.

The data generated is representative of real-world demands and end users could validate it on their own if they wished. TPC-C has been available for a relatively long time and is credible as a benchmarking tool. Transactions per Second (TPS) or Transactions per Minute (TPM) are commonly accepted metrics. It is relatively easy to reproduce a given test. TPC-C generates a lot of IOPS—storage and network latency will be critical success factors.

Configuration Overview

The benchmarks were performed in the Brocade Solutions Center Labs at Brocade’s corporate headquarters in San Jose, California.

Physical Infrastructure

The test environment, shown in Figure 1, included the following:

- Quest Benchmark Factory was used to generate the TPC-C workload with 500 concurrent end users per guest Operating System (OS).
- The Systems under Test were dual-socket, quad-core ESX Servers containing 32 gigabytes of memory.
- Storage was configured with a shared Virtual Machine File System (VMFS) volume for the guest OS files on the Storage Area Network (SAN) and Raw Device Mapping (RDM) LUNS for the database and log files.
- The SAN consists of two Brocade 200E Switches and a Hitachi AMS1000 storage array.
- No optimization was performed for the ESX Server, Windows, and SQL Server. Detailed configurations for these components are available separately (http://www.brocade.com/resources/whitepapers.jsp).
- One CPU was allocated to each SQL Server instance, but no limits were set on how much CPU core the instance could use. 1.5 GB of memory was reserved for each VM running a SQL Server instance. The memory configuration was determined by running the benchmark monitoring paging function. The memory allocation was increased until paging stopped, and then a small amount of memory was added.

NOTE: See "Detailed Configuration Information" toward the end of the paper to find out how to access detailed information on all the components. If you wish to duplicate this infrastructure, refer to the Brocade Web site, from which you can download the configurations for the components used in the benchmark.

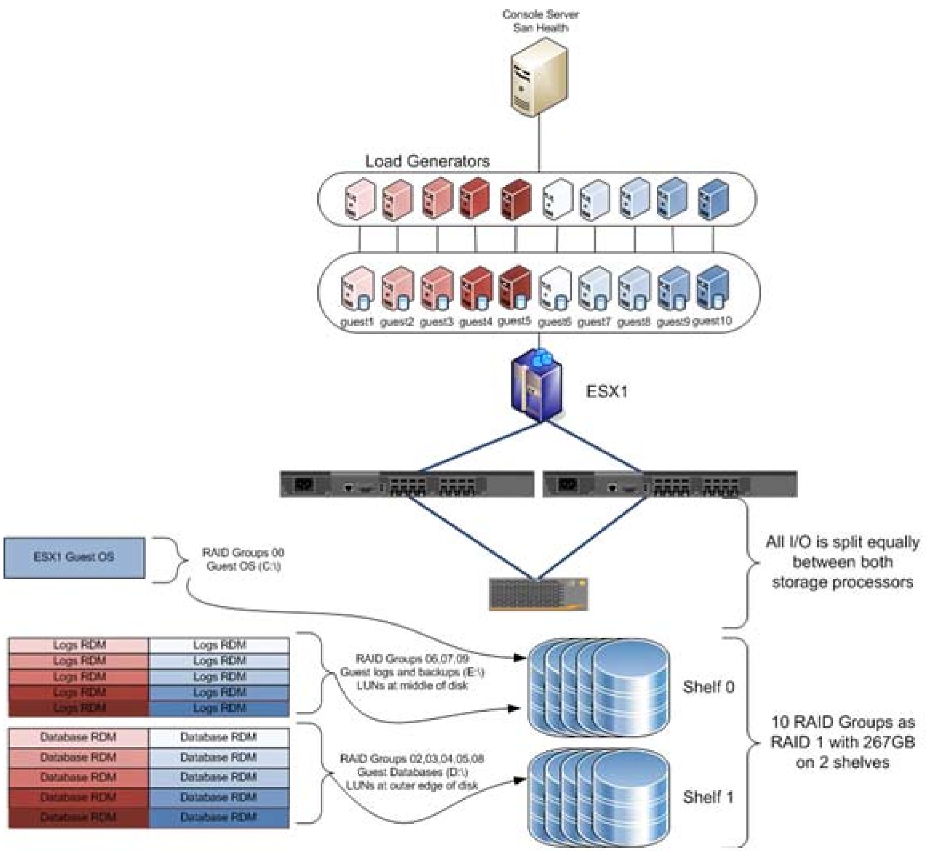


Figure 1.
Benchmark infrastructure for the single-platform test.

Further Details

The following provides additional details about the test infrastructure illustrated in Figure 1:

- The console is the Benchmark Factory control point, which is connected to the driver systems shown below it to run the agents driving the SUT running from 1 to 10 VMs configured with SQL Server.
- The SUT is shown in blue near the bottom.
- The ESX 3.5 host is dual-connected to the Fibre Channel SAN composed of two Brocade 200E 4 Gbit /sec switches. The switches in turn are connected to a Hitachi storage array.
- The SAN provides redundant paths to the storage, but only one side is actually active at a time, which means that all the traffic to and from the storage array flows through the same switch. In the case of a failure in any Fibre Channel component in the path, it is possible to fail over to the alternate switch and continue operations. This normally occurs without interruption to the applications.

Emulating the End-User Experience

The benchmark study was conducted by testers who represent typical IT Windows System Administrators. For example, they had no prior Fibre Channel (FC) experience and no previous knowledge of TPC-C, Benchmark Factory, or storage optimization. (Lack of optimization in the SUTs was discussed earlier.)

RESULTS

The results of the benchmark study and the degree of performance and stability experienced with ESX Server 3.5 were satisfying. Workload goals were met with no adverse effects to the hosts, SAN, or storage environments. The results are presented below in two parts: single-platform tests and the ESX Server cluster tests. Table 1 summarizes the TPC-C scores obtained for the “bare-metal” (BM) run and the 10 virtualized instance runs.

Results for Single-Platform Tests

Table 1.
Transaction TPC-C scores for 1 to 10 concurrent instances.

Run	Average Guest TPS	Total TPS	Users	Total TPM
BM	52.50	53	500	3,150
1	52.72	53	500	3,163
2	52.68	105	1,000	6,322
3	52.57	158	1,500	9,463
4	52.41	210	2,000	12,578
5	52.11	261	2,500	15,632
6	52.26	314	3,000	18,814
7	52.94	371	3,500	22,235
8	52.28	418	4,000	25,094
9	52.16	469	4,500	28,166
10	52.10	521	5,000	31,260

While there are minor variations, the results are very consistent for each run and VM instance. A single ESX Server demonstrated an overall rate of 31,260 TPM servicing 5,000 users in 10 VMs while using only about 17 percent of the CPU resources available on the SUT.

Running 10 VMs, each with one virtual CPU, on an 8-core host means that the ESX Server was over-committed in terms of CPU resources. The results show that ESX Server is more than capable of handling such over-commitment.

CPU Utilization

Table 2 summarizes average ESX Server and virtual machine CPU utilization in concurrently running VMs from 1 through 10 concurrent SQL Server instances. The "Average VM CPU Utilization" column shows the CPU utilized by each VM as a percent of the amount they were allocated. (The ESX Server information is also graphed in Figure 2.)

VMs	Average ESX CPU Utilization	Change from Previous Run	Average VM CPU Utilization
BM	-	-	9.7%
1	2.9%	0.0%	15.8%
2	3.7%	0.9%	13.3%
3	2.9%	- 0.9%	12.7%
4	4.9%	2.0%	12.7%
5	7.4%	2.5%	12.1%
6	9.6%	2.2%	10.3%
7	10.9%	1.3%	10.5%
8	12.8%	1.9%	10.6%
9	14.4%	1.6%	10.3%
10	16.6%	2.2%	15.0%

Table 2.
ESX Server and virtual machine CPU Utilization.

The first thing to notice is that the ESX Server CPU utilization does not increase much as VMs are added. The data shows that the incremental ESX processor utilization is less than 2.5 percent across the range of concurrent instances tested. The incremental overhead for running the SQL Server instances using the TPC-C workload in these tests over the range of VMs tested is minimal and very predictable.

Virtual machine CPU utilization is also very consistent, which is to be expected because every SQL Server instance is performing the same work with the same data.

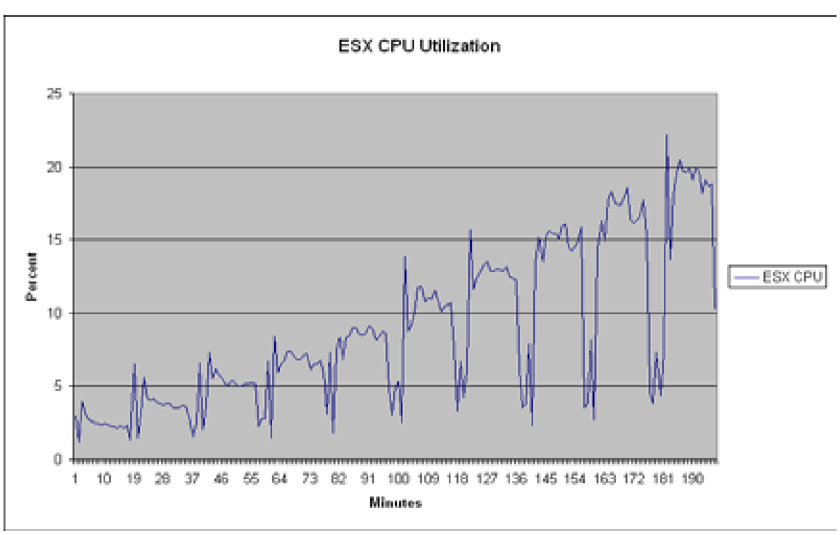


Figure 2.
ESX Server CPU utilization for 1 to 10 concurrent instances.

NOTE: Virtual machine accounting data is subject to the overall functioning of the virtualizing platform. VMware has produced a detailed white paper on VM time keeping (*Timekeeping in VMware Virtual Machines*, http://www.vmware.com/pdf/vmware_timekeeping.pdf).

Memory Utilization

Table 3 summarizes average ESX Server and virtual machine memory usage from 1 through 10 concurrent SQL Server instances. The "Average VM Memory Utilization" column shows the memory utilized by each VM as a percent of the amount they were allocated.

Table 3. ESX Server and VM memory utilization.

VMs	Average ESX Memory Utilization	Change from Previous Run	Average VM Memory Utilization
BM	-	-	37.0%
1	9.0%	0.0%	49.9%
2	14.2%	5.3%	46.9%
3	19.6%	5.4%	41.9%
4	24.3%	4.7%	40.3%
5	27.8%	3.5%	37.6%
6	30.7%	2.9%	33.8%
7	33.8%	3.1%	33.4%
8	38.1%	4.2%	33.1%
9	41.8%	3.8%	32.5%
10	46.5%	4.7%	32.4%

Two trends are immediately apparent.

- The extra memory required by ESX Server for each new instance is in the range of 3 to 5 percent for environments in which memory is not constrained. This illustrates very clearly the effects of ESX Server VM page sharing. ESX Server provides an additional optimization to memory management by sharing pages among VMs. In this instance, page sharing decreases memory requirements for each VM running SQL Server even further.
- Each virtual machine consumes an average of less than half of the amount of memory with which it was originally configured, and the average memory utilization declines as more VMs are added.

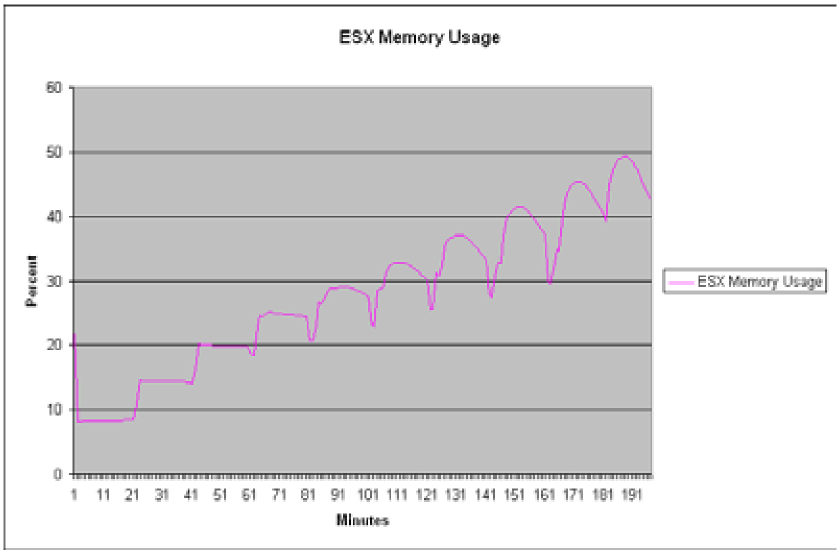


Figure 3.
ESX memory utilization for 1 to 10 concurent instances.

I/O Utilization

I/O generated by TPC-C workload can be characterized by a very large number of short block transfers. This section presents the results for the ESX Server only, to show how much I/O actually gets through the ESX Server abstraction layers and out the Fibre Channel Host Bus Adapter (HBA) ports. Table 4 shows the total platform I/O throughput obtained in each iteration of the test in megabytes per second.

The first column is the number of SQL Server VMs; the second column shows the average transfer level in megabytes per second; and the third column shows the peak reached

Number of VMs	ESX Average Host I/O in MB/sec	ESX Peak I/O in MB/sec
BM	4.3	-
1	4.0	9.3
2	6.0	18.9
3	7.8	28.7
4	9.6	29.2
5	10.7	26.1
6	14.7	38.3
7	15.7	51.2
8	18.4	70.6
9	20.9	71.6
10	21.9	62.8

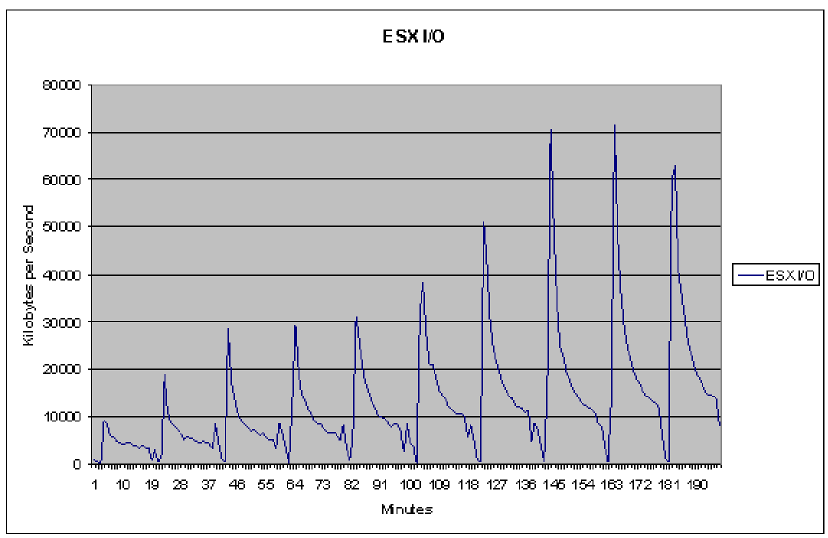
Table 4.
ESX platform total I/O utilization.

during each run. Notice that the traffic levels are not particularly high by FC standards. This is due to the nature of the test, in which the number of IOPS is high, but the amount of data transferred is low.

Storage latency plays a critical role in the ability of an application such as SQL Server to maintain the level of performance for the response times required by OLTP applications. Given that the host platform has no bottlenecks, both the storage array and the storage network must respond quickly to keep response times low. Proper storage array configuration minimizes device delays and the Brocade FC SAN ensures that there is no latency in the network. Latency is not a friend to OLTP environments, because it can cause application performance issues and higher response times. At worst, it triggers resource budgeting activities in the database servers, which can cause longer-term performance issues. In extreme cases, the database engine may decide that the device is no longer functional and shut down.

Figure 4 graphs the level of I/O traffic for the ten benchmark runs. Since each run activates an additional instance of SQL Server running the same workload, the regular increase in workload is expected. The start of each run is clearly shown by a spike in the traffic, caused by end users logging in to the application and some SQL Server startup activity such as building up the buffer cache. The OLTP transaction portion of the benchmark follows. This accurately simulates the use case associated with start-of-day activities, when end users come into work and start logging on to their applications.

Figure 4.
ESX I/O levels in kilobytes per second for 1 to 10 concurrent instances.



In Figure 4, the 10 runs show a decrease in I/O activity to 0 (zero) between each run. The result curves for all 10 runs have a similar shape with a peak of I/O activity at the start of the run followed by a tapering off of activity as the run progresses. A benchmark run starts by logging on the 500 users (the initial spike of resource usage including some SQL Server buffer caching), followed by the running of the TPC-C transactions. Once the transactions are complete, the I/O activity goes to 0 (zero) and the test ends.

Figures 5 and 6 show the I/O performance data for the run of 10 simultaneous instances as measured by ESX Server and independently by the Brocade SAN Health monitoring tool. It is clear from the minor differences in the data that ESX Server is producing performance data consistent with, although somewhat lower than, SAN Health. (The differences are most likely due to differences in sampling intervals.) Overall the data is very consistent between ESX Server and SAN Health monitoring.

Figure 5.
Brocade SAN Health I/O data
for 1 to 10 concurrent instances.

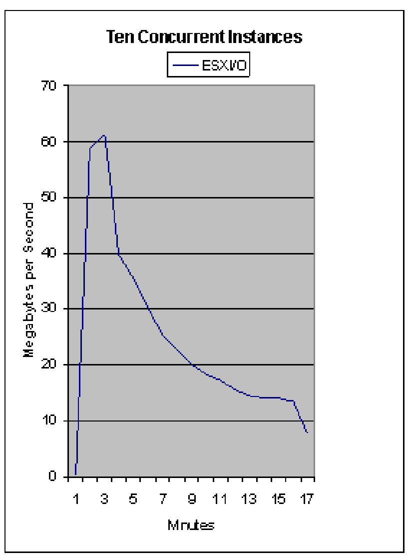


Figure 6.
ESX I/O data for
10 concurrent instances.

Other Resources
ESX Server also produces information on IP network activity. These results were not included in this study, because the levels were too low to influence overall performance in any way. The absence of storage-related latency and the low level of the IP activity, coupled with no resource bottlenecks in the virtual SQL Server instances, allowed the IP data to be safely ignored.

Results for Cluster

In the cluster test, the four ESX Servers were run together with each host running 5 instances of SQL Server. The benchmarking tool used was licensed for 20 concurrent clients, and the clients were distributed across all four ESX Servers. The CPU and Memory data for each ESX Server was consistent with the 5-VM data point in the single platform tests discussed earlier in this paper. The most interesting data was the combined traffic levels from the cluster, so this number of instances per host was judged acceptable. Measured results from a single host were used to estimate the overall throughput for the fabric, as measured from one of the hosts.

The only change to the overall benchmarking configuration was the activation of the three extra SUTs. Each SUT runs five VMs. No changes were made to the SAN fabric other than the connection of the extra SUT hosts.

Figure 7 shows the configuration for the cluster test.

Figure 7.
Benchmark infrastructure for the cluster test

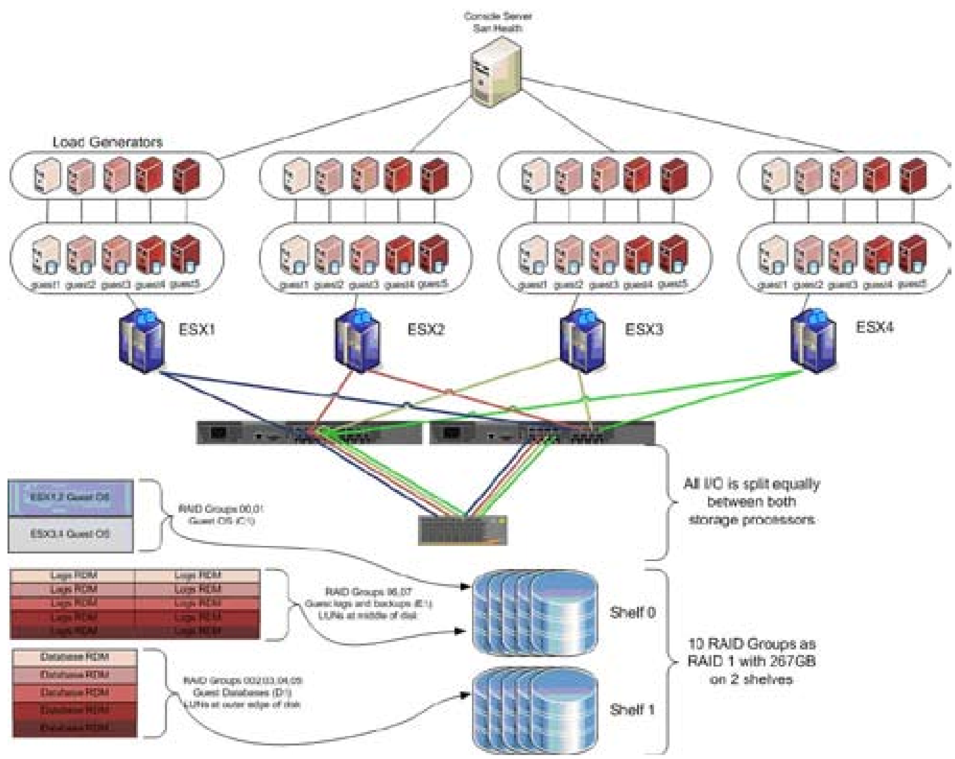


Figure 8 shows the ESX Server data. The total platform I/O is represented by the top line, and the average I/O per ESX Server is shown by the bottom line. Expected sustained rates of approximately 40 MB/sec are shown with peaks in the range of 140 to 150 MB/sec. Measured transfers from Brocade SAN Health are shown in Figure 9, a stacked graph showing the Average I/O levels from each of the four ESX Servers in which the total bandwidth is represented by the top (blue) line. The measured data mirrors what is expected from ESX Server in Figure 8.

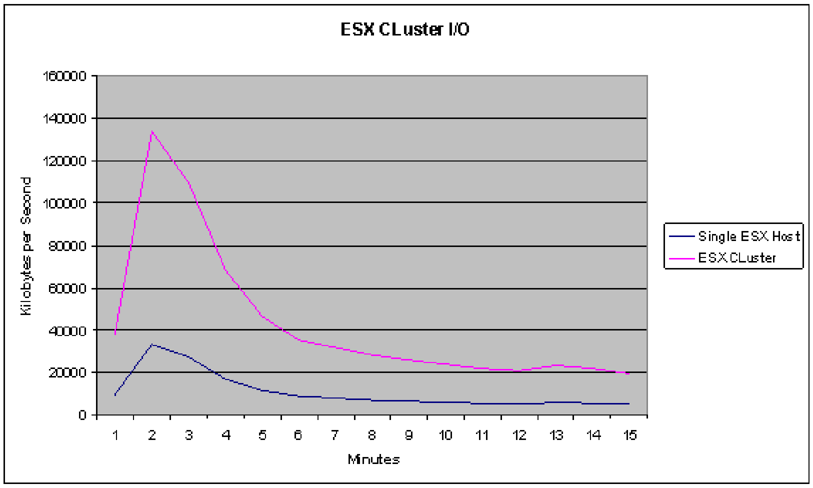


Figure 8.
Estimated disk throughput based on ESX Server data.

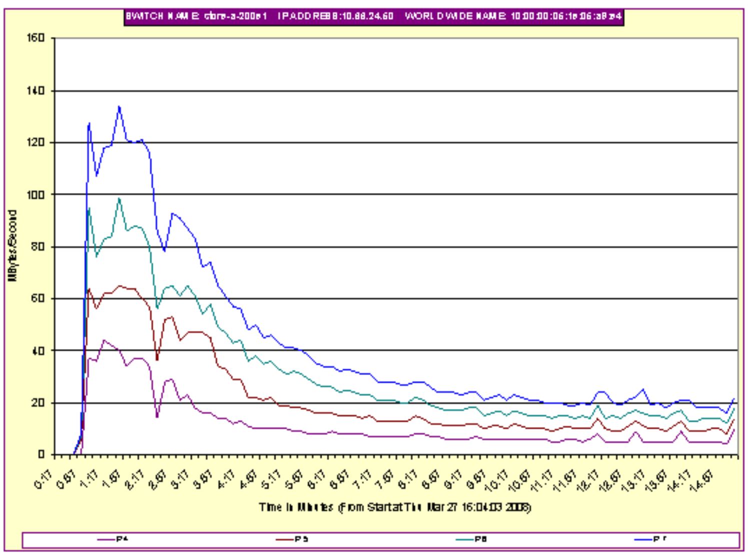


Figure 9.
Actual I/O levels measured by Brocade SAN Health.

Table 5 summarizes the TPS scores for the four ESX Servers, each running with 5 concurrent instances of SQL Server. Aggregate TPC-C scores for the cluster were approximately 1,050 TPS and 63,000 TPM for 10,000 users.

Table 5.
TPS results for the cluster run of 5 VMs per ESX Server.

ESX Host 1		ESX Host 2		ESX Host 3		ESX Host 4		Cluster	
VM	TPS	VM	TPS	VM	TPS	VM	TPS	TPM	TPS
ESX11	52.38	ESX21	52.38	ESX31	52.53	ESX41	52.64	-	-
ESX12	52.5	ESX22	52.87	ESX32	52.32	ESX42	52.82	-	-
ESX13	52.63	ESX23	52.8	ESX33	52.61	ESX43	52.61	-	-
ESX14	52.33	ESX24	52.88	ESX34	52.04	ESX44	52.77	-	-
ESX15	51.51	ESX25	52.71	ESX35	50.84	ESX45	52.62	-	-
Total TPS	261.35	-	263.64	-	260.34	-	263.46		1,049-
Total TPM	15,681	-	15,818	-	15,620	-	15,808	62,927	

NOTE: The tools used made it difficult to synchronize the start and execution of VM instances over several ESX Servers. Clusters will be studied in greater detail in future projects.

> The results of this benchmarking exercise show that running Microsoft SQL Server in a virtualized environment on ESX Server 3.5 is both achievable and practical.

Lessons Learned

TPC-C benchmarks look very good when you are trying to set high-water performance records on highly optimized, single-platform architectures, often with direct-attached, highly optimized storage. The goals of this study were exactly the opposite, and it took longer than expected to automate the tests.

ESX Server 3.5 aggregates its performance data in the repository frequently, so be sure to download detailed data as soon as possible after the test.

RAID sets and the number of physical I/O ports are important in this environment. Ensure that:

- The RAID sets are configured properly.
- You have enough performance to handle the number of connections from the hosts and the number of IOPS they generate.

WHY FIBRE CHANNEL?

Brocade considers consolidating virtualized SQL Server platforms as a Fibre Channel opportunity, because FC meets high I/O utilization with low latency and deterministic behavior of the storage, which are difficult requirements. In such an environment, even very small FC switches can deal with very high levels of virtualized I/O, especially when latency is a taken into account.

I/O-intensive applications can be highly sensitive to latency, which causes application response times to climb or, even worse, to seriously impact database performance budgets. This presents a use case for Brocade advanced fabric services such as:

- Flow-based Quality of Service (QoS). In this example, SQL Server instance I/O would receive priority over other virtualized applications, for which I/O performance is not among the critical performance criteria.
- Congestion detection and management. Even within a given level of QoS, it will be possible to give better service to any application that can be identified by a World Wide Name (WWN).
- Problem determination tools, which can provide further information in the event that manual intervention is required.

These platforms seem ideal for bladed server environments. The Access Gateway capability of Brocade Fabric OS® (FOS) offers a scalable storage connectivity solution, which can be installed directly into bladed server enclosures to provide seamless access to storage.

APPLICATIONS

Once it is established that it is reasonable to add IO-intensive applications into a virtualization environment, a great many opportunities open up for an infrastructure architect. An architect could, for example, define a workload based on different application profiles, such as I/O-intensive, CPU-intensive, and low-resource-consuming applications. Those profiles could be used to define standard platforms to be deployed as a single architectural building block.

The benefits of such an approach are clear:

- A known workload can be verified on a platform before deployment to reduce deployment risk and performance issues over time.
- The building blocks can be structured based on application workload requirements instead of infrastructure constraints. The result would be more predictable performance from the architecture over time.
- Larger building blocks facilitate capacity planning because the larger units are easier to work with.
- Clear application profiles enable simpler chargeback strategies, which are more relevant to the actual application. A good example of this is the traditional approach of charging SAN usage based on storage space allocation. This approach typically penalizes the end user who has a lot of storage that is not accessed very often. The database applications typically consume less storage, but data is accessed intensively. As a result these end users are usually charged too little for their SAN usage compared to the first end user.

Blade server systems have emerged as a very popular server deployment vehicle. This platform allows for straightforward application profile definitions at the server blade, blade enclosure, and rack level to further facilitate planning and workload deployments over time.

DETAILED CONFIGURATION INFORMATION

This document is available on the Brocade Web site (http://www.brocade.com/resources/whitepapers.jsp) along with the Zip file described below:

- PDF document
- Zip file with a Read Me file spreadsheets of detailed configuration information, including:
  - Host platform configurations
  - ESX Server configuration
  - Windows configuration
  - Storage array configuration
  - SQL Server configuration information
  - SAN zone configurations

CONCLUSION

Based on these results, Brocade believes that it is possible to deploy SQL Server in virtualized environments with consolidation ratios of up to 10:1 when deploying ESX Server with configurations similar to those described in this paper. ESX Server combined with a Brocade SAN and properly configured storage can easily handle OLTP workloads. In addition, we believe that storage latency, always the bane of performance metrics, can be avoided by deploying a well architected Fibre Channel SAN with properly configured storage arrays.

Corporate Headquarters
San Jose, CA USA
T: (408) 333-8000
info@brocade.com

European Headquarters
Geneva, Switzerland
T: +41 22 799 56 40
emea-info@brocade.com

Asia Pacific Headquarters
Singapore
T: +65-6538-4700
apac-info@brocade.com