

White Paper  
Intel Information Technology  
Computer Manufacturing  
Server Virtualization

## Comparing Multi-Core Processors for Server Virtualization

Intel IT tested servers based on select Intel multi-core processors to analyze the potential role of each in our data center server virtualization strategy. Each server provided significant potential benefits in performance, power consumption per workload, and operating costs over older servers running non-virtualized workloads. The server based on the Quad-Core Intel® Xeon® processor 7300 series showed the greatest scalability and used the least power per job when running large numbers of virtual machines (VMs), and proved suitable for achieving high consolidation ratios in a busy enterprise data center. Servers based on the Quad-Core Intel Xeon processor 5300 series or the Dual-Core Intel Xeon processor 5100 series could effectively support more moderate consolidation ratios, with very low power consumption.

Robert E. Carpenter, Intel Corporation

August 2007

IT@Intel

## Executive Summary

Intel IT tested servers based on select Intel multi-core processors to analyze the potential role of each in data center server virtualization strategies. Each server provided significant potential benefits in performance, power consumption per workload, and operating costs over older servers running non-virtualized workloads.

Each server provided significant potential benefits in performance, power consumption per workload, and operating costs over older servers running non-virtualized workloads.

We compared a four-socket server based on the Quad-Core Intel® Xeon® processor X7350, with a total of 16 cores; a dual-socket server based on the Quad-Core Intel Xeon processor X5355, with eight cores; and a dual-socket server based on the Dual-Core Intel Xeon processor 5160, with four cores.

To test each server, we progressively added workloads, each consisting of a virtual machine (VM) with one copy of a CPU-intensive application.

- All three servers provided major improvements in performance and power consumption per workload over older servers running non-virtualized workloads.
- With each server, runtimes remained flat until the number of VMs reached the number of cores, then increased in a predictable, linear way.
- The Quad-Core Intel Xeon processor X7350-based server was about twice as scalable as the Quad-Core Intel Xeon processor X5355-based server. It consumed more power overall but used the least power per job when running large numbers of VMs.

Based on our tests, the servers could effectively support different consolidation scenarios. Expandable, scalable servers based on the Quad-Core Intel Xeon processor 7300 series enable high consolidation levels in a busy enterprise data center. The efficient performance of servers based on the Quad-Core Intel Xeon processor 5300 series and Dual-Core Intel Xeon processor 5100 series support more moderate consolidation ratios, with very low power consumption.

# Contents

<b>Executive Summary</b> .....	2
<b>Business Challenge</b> .....	3
Key Technologies and Architectures.....	4
<b>Server Virtualization Performance Tests</b> .....	4
Methodology.....	5
Comparing Performance.....	5
Results.....	6
<b>Analysis</b> .....	10
<b>Conclusion</b> .....	11
<b>Authors</b> .....	11
<b>Acronyms</b> .....	11

## Business Challenge

Like many IT organizations, Intel IT is under intense pressure to increase data center performance and capacity while driving down cost. Business computing needs are growing rapidly, and we continually look for ways to increase capacity in order to deliver the performance required to meet our business groups' needs. We must meet defined service level agreements (SLAs) while delivering performance at minimum cost.

Low server utilization has been a key obstacle hindering our efforts to reduce total cost of ownership (TCO). Like many organizations, Intel IT historically has dedicated a single application to each server in most cases. Most of our business servers have used less than 20 percent of CPU capacity.

Controlling data center energy consumption is also a growing challenge. Reducing energy consumption lowers TCO and can also help avoid the potentially even more serious problem of reaching the limit of a data center's available power supply.

## Key Technologies and Architectures

We have identified key technologies and architectural approaches that we are using to meet these challenges:

- **Virtualization.** This approach executes applications in standard containers—VMs. This enables us to consolidate multiple workloads

onto each server, increasing utilization and reducing power consumption per workload. This reduces server TCO. We can also configure a VM much more quickly than a new physical server—in minutes rather than hours—enabling us to respond rapidly to changes in business demand.

- **Utility computing.** Our goal is to create a more flexible, agile data center environment based on virtualization. We aim to be able to automatically allocate jobs to VMs on server resources anywhere within the environment. This maximizes the efficient use of data center resources, further reducing cost and improving responsiveness to business needs.
- **Intel multi-core processors.** Servers based on multi-core Intel Xeon processors underpin our strategy. They provide much greater performance, scalability, and power efficiency than previously possible. This enables us to implement server virtualization and consolidation, facilitating the development of a utility data center model.

# Server Virtualization Performance Tests

We conducted virtualization performance tests on servers based on different Intel multi-core processor series to evaluate their capabilities and potential roles in the data center.

We tested servers based on three configurations:

- Four-socket server based on the Quad-Core Intel Xeon processor 7300 series
- Dual-socket server based on the Quad-Core Intel Xeon processor 5300 series
- Dual-socket server based on the Dual-Core Intel Xeon processor 5100 series

In our tests, we focused on key measurements that would help determine potential roles for each server in a virtualized environment:

- Performance as we progressively increased the number of virtualized workloads; we focused on the time to complete each workload, since this reflected the server's ability to support SLAs in a virtualized environment
- Power consumption overall and per workload

We also estimated the operational cost benefits of using the servers to achieve different levels of consolidation.

Based on our analysis, we formulated potential roles for each of these server platforms in data center consolidation and virtualization strategies.

## Methodology

We conducted our tests using a synthetic CPU-intensive database application. This approach yields reproducible results and enables us to eliminate ambiguity related to the varying footprint of business applications on different platforms. Our application was representative of business applications such as business intelligence and financial forecasting, and also of technical applications such as seismic analysis.

To test each server, we progressively added workloads, each consisting of a VM with one copy of the application. We started with a minimum of two VMs, and we allocated each VM a capacity of one processor core. Each time we increased the number of VMs, we ran all workloads to completion. We recorded results including runtime for each workload and power consumption.

We continued adding VMs to each server until clear performance trends emerged. This resulted in different maximum numbers of VMs for each server. It also enabled us to identify relationships

between the number of VMs and the ability to deliver specific levels of performance, relative to the baseline performance with two VMs, in order to meet SLAs.

## Comparing Performance

To make it easier to compare the performance of the three servers, we created normalized measures of performance and workload.

We normalized runtime data for each server relative to the smallest workload. This enabled us to compare inherent behaviors of the different server platforms by eliminating performance differences due to clock speed.

We also calculated a normalized measure of workload: VMs per core. This enabled us to compare server behavior as the number of VMs approached and then exceeded the number of cores.

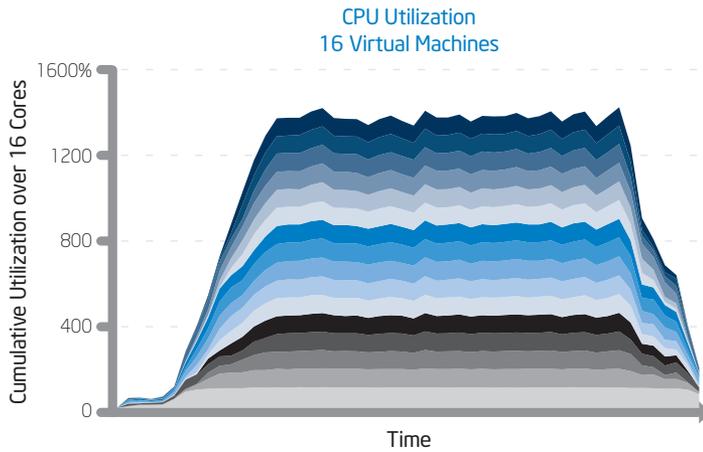
We also compared our test results for the three servers with previous performance data from a single, non-virtualized copy of the same application on an Intel® Pentium® III processor-based server, a traditional Intel IT server solution. Configurations of the four systems are shown in Table 1.

**Table 1. Test System Configurations**

Processor	Clock Speed	Cores per Processor	Number of Processors	Total Cores	Cache per Processor	Bus	Memory Type	Amount of Memory	Memory per VM
Quad-Core Intel® Xeon® Processor X7350	2.93 GHz	4	4	16	8 MB Level 2	1066 MHz DHSI	667 MHz FB DIMM RAM	64 GB RAM	2 GB
Quad-Core Intel Xeon Processor X5355	2.66 GHz	4	2	8	8 MB Level 2	1333 MHz FSB	667 MHz FB DIMM RAM	32 GB RAM	2 GB
Dual-Core Intel Xeon Processor 5160	3 GHz	2	2	4	4 MB Level 2	1333 MHz FSB	667 MHz FB DIMM RAM	32 GB RAM	2 GB
Intel® Pentium® III Processor	733 MHz	1	1	1	256 K	133 MHz	SDRAM	2.5 GB	N/A <sup>1</sup>

<sup>1</sup> Not applicable: This workload was not virtualized.

**DHSI** dedicated high speed interconnects  
**FB DIMM** fully buffered dual inline memory module  
**FSB** front side bus  
**VM** virtual machine



**Figure 1. Our test application intensively utilized all the allocated cores on each tested server.** With the four-socket server based on the Quad-Core Intel® Xeon® processor X7350, core utilization approached an average of 90 percent across all 16 cores when running 16 virtual machines. Intel internal measurements, July 2007.

## Results

The application effectively utilized the available cores. For all the servers tested, the application consumed 85 to 90 percent of all allotted processor cores over the range of VMs run on each server. For example, on the server based on the Quad-Core Intel Xeon processor X7350, average utilization across all 16 cores approached 90 percent when running 16 VMs, as shown in Figure 1.

We observed clear trends differentiating the three servers in scalability and power consumption. Results are shown in Table 2.

### Performance and Scalability

The three servers showed markedly different levels of scalability. We did, however, observe a similar pattern of performance with increasing load on each server.

As we increased the number of VMs, average runtimes remained approximately constant until the number of VMs approached the total number of cores. As the number of VMs began to exceed

**Table 2. Runtimes and Power Consumption**

Number of Virtual Machines	Quad-Core Intel® Xeon® Processor X7350					Quad-Core Intel Xeon Processor X5355					Dual-Core Intel Xeon Processor 5160				
	Average Time	Normalized Time	Average Watts	Watt-Minutes (W-M)	W-M/job	Average Time	Normalized Time	Average Watts	W-M	W-M/job	Average Time	Normalized Time	Average Watts	W-M	W-M/job
2	6.23	1.00	495	3083	1542	4.92	1.00	478.00	2350	1175	4.38	1.00	330.45	1448	724
4	6.40	1.03	523	3346	837	4.95	1.01	458.90	2272	568	4.67	1.06	366.40	1710	427
6	5.87	0.94	536	3148	525	5.01	1.02	479.30	2401	400	5.84	1.33	372.82	2177	363
8	5.83	0.94	557	3246	406	5.61	1.14	489.43	2747	343	6.64	1.51	370.56	2461	308
10	5.70	0.91	589	3358	336	6.64	1.35	481.08	3196	320					
12	5.58	0.90	625	3489	291	7.61	1.55	494.54	3763	314					
14	5.39	0.86	639	3443	246	8.42	1.71	489.29	4121	294					
16	5.49	0.88	675	3705	232										
18	7.04	1.13	850	5982	332										
20	7.46	1.20	842	6282	314										
22	7.81	1.25	864.68	6752	307										
24	8.08	1.30	865	6982	291										
26	8.63	1.38	876	7558	291										
28	8.90	1.43	902	8023	287										
30	9.13	1.47	911	8318	277										
32	10.34	1.66	916	9479	296										

the number of cores, average runtimes began to increase due to the total time required to run all the jobs and the fact that the VM hypervisor spent more time managing VMs among the available cores. We observed predictable, approximately linear increases in runtime as the number of VMs increased.

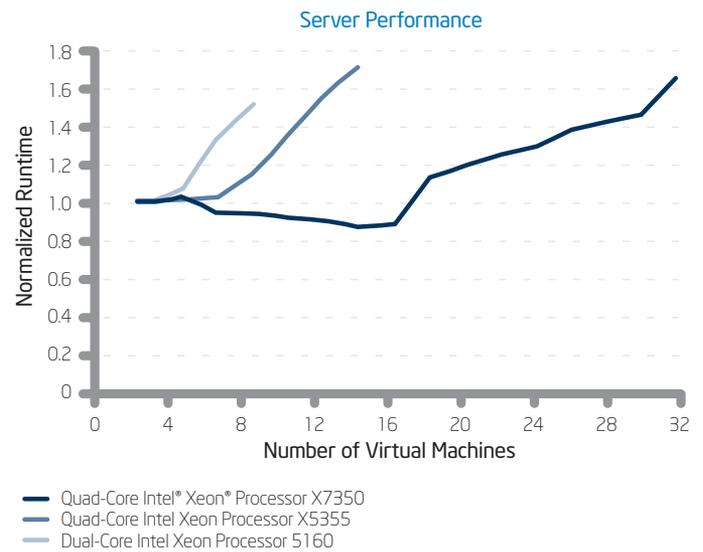
This relationship resulted in substantial differences in scalability, as shown in Figure 2. With the Quad-Core Intel Xeon processor X7350-based server, average runtime remained approximately flat until the number of VMs reached 16, then gradually and predictably increased. With 24 VMs, average runtime increased by 40 percent; with 32 VMs, it increased by about 70 percent. With the servers based on the Dual-Core Intel Xeon processor 5160 and the Quad-Core Intel Xeon processor X5355, average runtime started to increase at about four and eight VMs respectively.

These trends were also apparent in the non-normalized runtimes, though we observed minor variations.

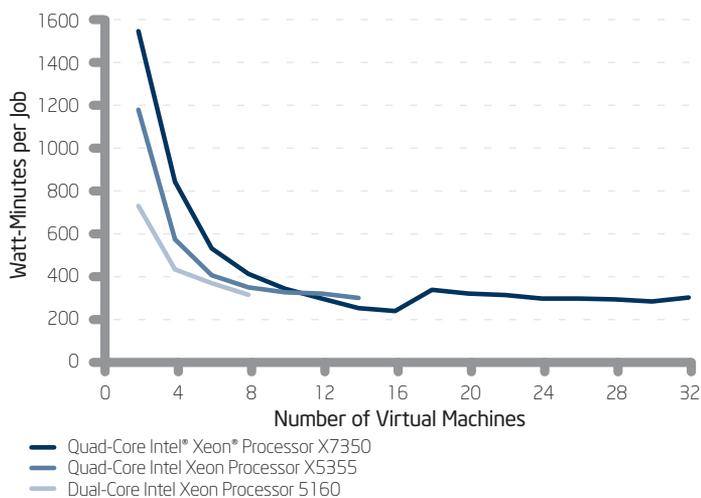
When running few VMs, the server based on the Dual-Core Intel Xeon processor 5160 completed each job more quickly on average than the other servers. This was due to factors related to VM hypervisor efficiency. On the Quad-Core Intel processor X7350-based server, the VM software utilized all 16 cores to some extent, even when running only two VMs. This indicates that on a lightly loaded machine, the hypervisor spends a significant amount of time checking cores and moving VMs between them relative to the time spent executing workloads. We did not customize the hypervisor to pin VMs to specific cores, because to achieve the flexibility needed for the virtual utility data center, we cannot decide in advance where an individual job will run. The effect was greatest on the Quad-Core Intel Xeon processor X7350-based server because it has the most cores.

However, this effect became insignificant as the number of VMs increased; performance was increasingly determined by the availability of CPU resources. As hypervisor efficiency increased, runtimes on the Quad-Core Intel Xeon processor X7350-based server actually decreased slightly with six to 16 VMs.

As a result of its larger number of cores, with eight or more VMs, the Quad-Core Intel Xeon processor X7350-based server completed each job faster than the Dual-Core Intel Xeon processor 5160-based server. With 10 or more VMs, it also outpaced the Quad-Core Intel Xeon processor X5355-based server.



**Figure 2. In tests, the three servers based on Intel multi-core processors showed differing levels of scalability.**  
Intel internal measurements, July 2007.



**Figure 3. Power consumption per job with increasing numbers of workloads.** Intel internal measurements, July 2007.

### Power Consumption

We also saw a crossover in power consumption per job as we increased VMs, due to the scalability of the Quad-Core Intel Xeon processor X7350-based server.

This server was a larger system overall than the others; besides having twice as many cores, it also was configured with twice as much memory, more fans, and redundant power supplies.

As a result, it consumed more power than either of the other two servers: an average of 495 watts (W) when running two VMs, compared with 478 W for the Quad-Core Intel Xeon processor X5355-based server, which also had redundant power supplies, and 330 W for the Dual-Core Intel Xeon processor 5160-based server.

As we added more VMs, all the servers became more power-efficient, in terms of power consumption per job, as shown in Figure 3.

However, this trend was strongest with the Quad-Core Intel Xeon processor X7350-based server. As a result, with increasing numbers of VMs, power consumption differences between the servers at first narrowed and then were reversed in many cases, as shown in Table 3 and Figure 3. For instance, when loaded with 24 VMs—a VM-to-core ratio of 1.5 to 1—the server based on the Quad-Core Intel Xeon processor X7350 consumed slightly less power per job than the other servers when they were loaded to the same VM-to-core ratio.

All the multi-core processors demonstrated dramatic power savings compared with running non-virtualized workloads on the original Intel Pentium III processor-based servers, as shown in Table 3.

**Table 3. Power Needed to Run 24 Workloads**

	Runtime	Average Watts	Watt-Minutes	Virtual Machines per Server	Number of Servers	Total Energy Budget in Watt-Minutes
Quad-Core Intel Xeon Processor X7350	8.96	823	7374	24	1	7374
Quad-Core Intel Xeon Processor X5355	7.61	495	3767	12	2	7534
Dual-Core Intel Xeon Processor 5160	5.84	373	2178	6	4	8713
Intel Pentium III Processor	8.28	157.5	1304	N/A	24	31298

### Operational Costs

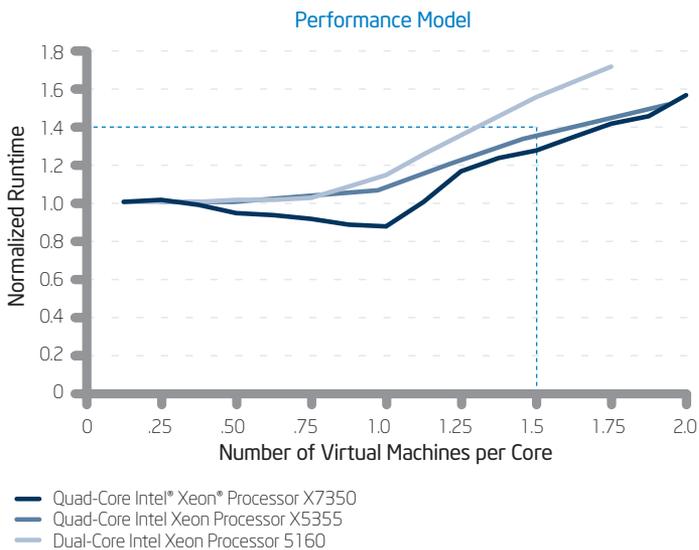
Higher consolidation levels create potential operating efficiencies. To estimate these, we built an operational cost model based on an intensively used data center environment and a consolidation ratio of 1.5 VMs per core.

We chose this ratio of 1.5 VMs per core because, for our CPU-intensive test application, it corresponded to an example SLA constraint that all jobs within the data center running on fully loaded machines should complete within 1.4x of the time required to run in a single VM on an otherwise empty machine. We selected this performance range as a theoretical but reasonable example, based on input from an Intel operations team. With our test application, each server achieved this performance with approximately 1.5 VMs per core, as shown in Figure 4. Different service levels will require different consolidation ratios.

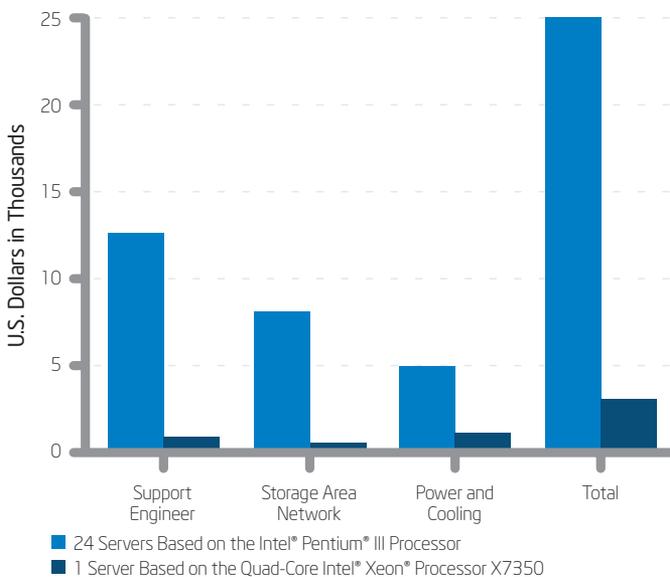
Based on the model, there are potential differences in annual operating costs depending on consolidation level, as shown in Figure 5. All three servers offered considerable potential savings compared to the original Intel Pentium III processor-based servers with non-virtualized workloads.

We project reduced support, storage area network, and power and cooling costs. This calculation is conservative and does not include many other “per-system” savings, including those due to reduced server room square footage, racks, network switches and routers, network drop costs and power run costs, backup power supplies, uninterruptible power supplies (UPSs), and additional administrative and management costs within the operations center.

These potential savings increase with the higher consolidation ratios achievable with servers based on the Quad-Core Intel Xeon processor 7300 series, compared with using a larger number of single-socket or dual-socket servers to deliver equivalent processing power.



**Figure 4. Each server could support a theoretical service level agreement set at 1.4x baseline performance when running about 1.5 virtual machines per core. Intel internal measurements, July 2007.**



**Figure 5. Cost comparison of 24 to 1 server consolidation.**

**Assumptions:**

- **System Engineering Resources.** One system engineer per 250 physical servers, with a full-time employee burden rate of USD 130,000 per year.
- **Storage Area Network.** One SAN card per device at USD 1,000 per SAN card. Cards are depreciated over three years.
- **Electricity.** USD 0.08 per kilowatt-hour. Server power consumption is multiplied by 1.75 to include estimated data center cooling.
- **Capacity.** Runs 22 hours per day with a full load (24 workloads), and 2 hours per day with a light load (2 workloads).

# Analysis

In a virtualized environment, jobs may be automatically allocated to any server with available capacity. To meet SLAs, servers have to be able to complete each job within a predictable time while running a variable mix of other workloads.

Each of the three multi-core processor-based servers continued to deliver predictable runtimes as the number of CPU-intensive workloads increased, making them good candidates for a virtualized environment. However, the differences in scalability suggest different potential roles for each server.

We used a straightforward method to analyze these potential roles. To meet an SLA, we need to complete workloads within an acceptable time period, as agreed with a business group. To determine this acceptable runtime, we could first establish a baseline based on the time it takes the server to run the virtualized workload when running only a single VM. Then we could agree on an acceptable range of variation, such as 1.4x this baseline runtime.

We can use this goal to determine the server capacity that we need. With our test application, each server achieved this performance when running an average of approximately 1.5 VMs per core, as was shown in Figure 4.

This translates into consolidation ratios of approximately 24 to 1 for the Quad-Core Intel Xeon processor X7350-based server, 12 to 1 for the Quad-Core Intel Xeon processor X5355-based server, and 6 to 1 for the Dual-Core Intel Xeon processor 5160-based server.

Data center managers can then select the most suitable processors and servers based on factors

such as the desired consolidation ratio and the type of data center environment. Our results and analysis suggest the following roles:

- **Quad-Core Intel Xeon processor 7300 series.** The scalable server based on the Quad-Core Intel Xeon processor 7300 series was suitable for achieving the highest consolidation ratios in a busy enterprise data center environment. It delivered predictable runtimes over the widest range of workload levels, meeting our theoretical 1.4x SLA performance level when running between one and 24 or more workloads. When heavily loaded, this server delivered the lowest power consumption per workload, with the potential for the greatest operational savings due to high consolidation levels.
- **Quad-Core Intel Xeon processor 5300 series and Dual-Core Intel Xeon processor 5100 series.** Servers based on the Quad-Core Intel Xeon processor 5300 series or the Dual-Core Intel Xeon processor 5100 series effectively supported more moderate levels of consolidation. They also used less power overall and less power per workload when running moderate numbers of workloads. The consolidation levels we observed in tests with these servers could result in large operational cost savings over our older, non-virtualized server solutions.

# Conclusion

Based on our testing, each of these servers can deliver significant benefits in data center performance, power consumption, and operational costs over older non-virtualized servers. The selection of the most suitable server may depend on the specific data center environment and consolidation strategy.

In a busy enterprise data center environment, servers based on the Quad-Core Intel Xeon processor 7300 series potentially can achieve the greatest scalability, the highest consolidation ratios, and the best performance per watt, based on our tests. These servers are the most expandable and typically are configured with more memory and other features designed to enable continuous intensive use in an enterprise data center.

It is important to note that our workload was designed to create a “worst case” scenario; we drove CPU utilization to nearly 100 percent and maintained that level in a steady state. Actual business applications differ, in that demand is typically lower and varies over time. This lower, varying demand leaves additional headroom that can be used to manage peak loads or to increase the number of VMs while maintaining service levels.

In practice, we expect to see higher consolidation ratios without service level degradation.

Servers based on the Quad-Core Intel Xeon processor 5300 series and the Dual-Core Intel Xeon processor 5100 series effectively supported more moderate levels of consolidation with low power consumption. Servers based on the Quad-Core Intel Xeon processor 5300 series have more potential to scale and achieve higher consolidation levels than servers based on the Dual-Core Intel Xeon processor 5100 series.

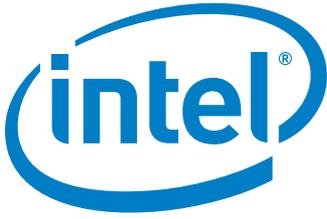
Other factors may also help determine the most suitable class of servers for a specific data center environment, such as the type of applications to be consolidated, system administration considerations, and licensing costs of virtualization and application software.

## Authors

**Robert E. Carpenter** is an enterprise architect with Intel Information Technology.

## Acronyms

<b>DHSI</b>	dedicated high speed interconnects	<b>TCO</b>	total cost of ownership
<b>FB DIMM</b>	fully buffered dual inline memory module	<b>UPS</b>	uninterruptible power supply
<b>FSB</b>	front side bus	<b>VM</b>	virtual machine
<b>SLA</b>	service level agreement	<b>W</b>	watts



[www.intel.com/IT](http://www.intel.com/IT)

Test systems included an HP server based on the Quad-Core Intel® Xeon® processor X7350.

Performance tests and ratings are measured using specific computer systems and/or components and reflect the approximate performance of Intel products as measured by those tests. Any difference in system hardware or software design or configuration may affect actual performance. Buyers should consult other sources of information to evaluate the performance of systems or components they are considering purchasing. For more information on performance tests and on the performance of Intel products, visit Intel Performance Benchmark Limitations at [www.intel.com/performance/resources/limits.htm](http://www.intel.com/performance/resources/limits.htm).

Intel® Virtualization Technology requires a computer system with an enabled Intel® processor, BIOS, virtual machine monitor (VMM) and, for some uses, certain computer system software enabled for it. Functionality, performance or other benefits will vary depending on hardware and software configurations and may require a BIOS update. Software applications may not be compatible with all operating systems. Please check with your application vendor.

This paper is for informational purposes only. THIS DOCUMENT IS PROVIDED "AS IS" WITH NO WARRANTIES WHATSOEVER, INCLUDING ANY WARRANTY OF MERCHANTABILITY, NON-INFRINGEMENT, FITNESS FOR ANY PARTICULAR PURPOSE, OR ANY WARRANTY OTHERWISE ARISING OUT OF ANY PROPOSAL, SPECIFICATION OR SAMPLE. Intel disclaims all liability, including liability for infringement of any proprietary rights, relating to use of information in this specification. No license, express or implied, by estoppel or otherwise, to any intellectual property rights is granted herein.

Intel, the Intel logo, Intel. Leap ahead. and Intel. Leap ahead. logo, Pentium, and Xeon are trademarks of Intel Corporation in the U.S. and other countries.

\* Other names and brands may be claimed as the property of others.

Copyright © 2007, Intel Corporation. All rights reserved.

Printed in USA  
0807/SEP/RDA/PDF

Please Recycle  
ITAI Number: 07-2601w