THE POSSIBILITIES ARE INFINITE

WHITE PAPER

Performance Report PRIMERGY TX300 S5

Version 2.0
August 2009

Pages 52

Abstract

This document contains a summary of the benchmarks executed for the PRIMERGY TX300 S5.

The PRIMERGY TX300 S5 performance data are compared with the data of other PRIMERGY models and discussed. In addition to the benchmark results, an explanation has been included for each benchmark and for the benchmark environment.

Contents

Document history

Version 1.0

First report version including the benchmark chapters

- SPECcpu2006
  Measurements with Xeon E5502, E5504, E5506, E5520, E5540, X5550 and X5570
- SPECjbb2005
  Measurement with Xeon X5570
- OLTP-2
  Measurements with Xeon E5502, E5504, E5506, E5520, E5540, X5550 and X5570
- Terminal Server
  Measurements with Xeon E5504 and X5570

Version 1.1

New benchmark chapters:

- StorageBench
  Measurements with LSI MegaRAID SAS 1068 controller
  Measurements with LSI MegaRAID SAS 1078 controller

Version 1.2

Updated benchmark chapters:

- SPECcpu2006
  Measurements with Xeon L5506 and L5520
  Additional measurements with Xeon E5520, E5540 and X5570
- Technical Data, StorageBench, Literature (corrections)

New benchmark chapters:

- SPECsfs2008
  Measurement with Xeon X5570, 96 GB RAM and 3 x 6 data disks (146 GB, 10000 rpm) RAID 50
  Measurement with Xeon X5570, 96 GB RAM and 3 x 6 data disks (146 GB, 10000 rpm) RAID 0
- SAP SD
  Measurement with Xeon X5570 and SAP ECC Release 6.0
  Measurement with Xeon X5570 and SAP Business Suite software SAP enhancement package 4 for SAP ERP 6.0 (Unicode)
- vServCon
  Measurements with Xeon L5520, E5520, E5540, X5550 and X5570

Version 2.0

in all benchmark chapters: foot note regarding availability of components added

Updated benchmark chapters:

- SPECcpu2006
  Measurements with Xeon L5530, E5530, X5560 and W5590
  Additional measurements with Xeon E5520, E5540 and X5550
- SPECjbb2005
  Measurements with Xeon E5506, E5520 and W5590
- StorageBench (corrections)
- OLTP-2
  Corrections
  Measurements with Xeon L5530, E5530, X5570 and W5590
- vServCon (corrections, updated diagrams)

Technical Data

The PRIMERGY TX300 S5 is a dual socket tower server with an Intel 5520 chipset, two Intel Dual-Core or Quad-Core Xeon processors, up to 144 GB PC3-10600 or PC3-8500 registered ECC DDR3-SDRAM, a bus with 800, 1067 or 1333 MHz timing - depending on the processor used, a SAS controller with RAID 0 and RAID 1 for up to eight internal SATA or SAS hard disks or a SAS controller with RAID 5 and RAID 6 for up to 20 internal SATA or SAS hard disks, an onboard 2-port 1-GBit Ethernet controller and seven PCI-slots (2 PCIe-2 x8 and 5 PCIe-2 x4). The housing comprises up to eight 3.5" or up to 20 2.5" hard disks. As well as its predecessor PRIMERGY TX300 S4 the PRIMERGY TX300 S5 can be converted quickly and easily into a rack system with 4 height units for integration in 19-inch racks.




See [Data sheet PRIMERGY TX300 S5] for detailed technical information.

SPECcpu2006*

Benchmark description

SPECcpu2006 is a benchmark to measure system efficiency during integer and floating point operations. It consists of an integer test suite containing 12 applications and a floating point test suite containing 17 applications which are extremely computing-intensive and concentrate on the CPU and memory. Other components, such as disk I/O and network, are not measured by this benchmark.

SPECcpu2006 is not bound to a specific operating system. The benchmark is available as source code and is compiled before the actual benchmark. Therefore, the compiler version used and its optimization settings have an influence on the measurement result.

SPECcpu2006 contains two different methods of performance measurement: The first method (SPECint2006 and SPECfp2006) determines the time required to complete a single task. The second method (SPECint_rate2006 and SPECfp_rate2006) determines the throughput, i.e. how many tasks can be completed in parallel. Both methods are additionally subdivided into two measuring runs, "base" and "peak", which differ in the way the compiler optimization is used. The "base" values are always used when results are published, the "peak" values are optional.

Benchmark	Arithmetic	Type	Compiler optimization	Measuring result	Application
SPECint2006	integer	peak	aggressive	speed	single threaded
SPECint_base2006	integer	base	conservative		
SPECint_rate2006	integer	peak	aggressive	throughput	multithreaded
SPECint_rate_base2006	integer	base	conservative		
SPECfp2006	floating point	peak	aggressive	speed	single threaded
SPECfp_base2006	floating point	base	conservative		
SPECfp_rate2006	floating point	peak	aggressive	throughput	multithreaded
SPECfp_rate_base2006	floating point	base	conservative		

The results represent the geometric mean of normalized ratios determined for the individual benchmarks. Compared with the arithmetic mean, the geometric mean results in the event of differingly high single results in a weighting in favor of the lower single results. “Normalized“ means measuring how fast the test system runs in comparison to a reference system. The value of “1“ was determined for the SPECint_base2006, SPECint_rate_base2006, SPECfp_base2006 and SPECfp_rate_base2006 results of the reference system. Thus a SPECint_base2006 value of 2 means for example that the measuring system has executed this benchmark approximately twice as fast as the reference system. A SPECfp_rate_base2006 value of 4 means that the measuring system has executed this benchmark about 4/[# base copies] times as fast as the reference system. “# base copies“ here specifies how many parallel instances of the benchmark have been executed.

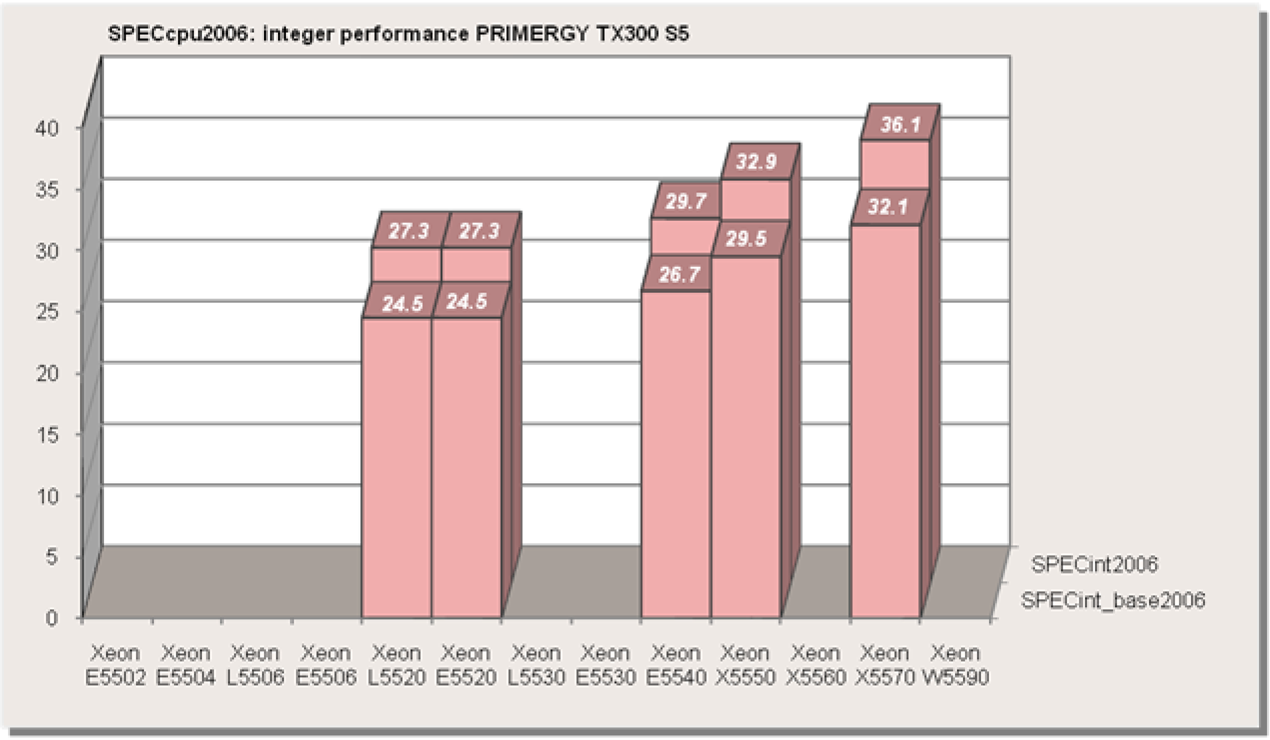
We do not submit all SPECcpu2006 measurements for publication at SPEC. So not all results appear on SPEC’s web sites. As we archive the log data for all measurements, we are able to prove the correct realization of the measurements any time.

Benchmark results

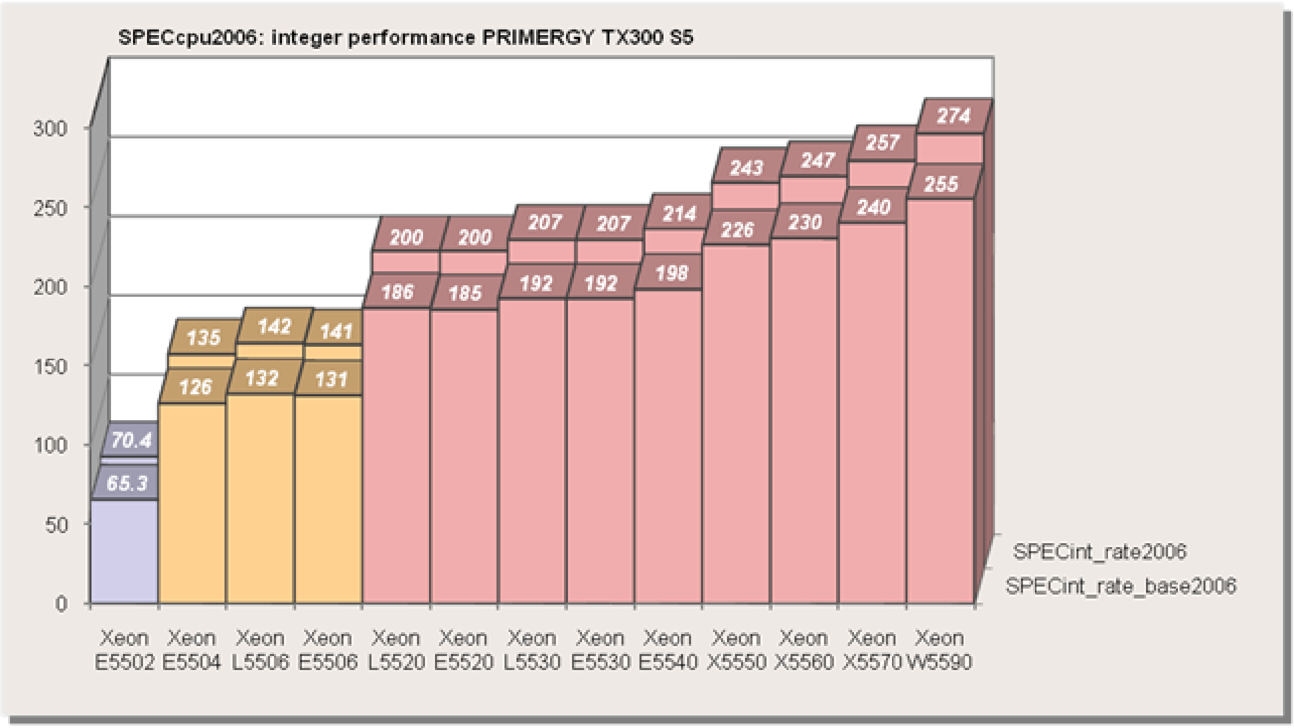
The PRIMERGY TX300 S5 was measured with the processors Xeon E5502, E5504, L5506, E5506, L5520, E5520, L5530, E5530, E5540, X5550, X5560, X5570 and W5590. The benchmark programs were compiled with the Intel C++/Fortran compiler 11.0 and run under SUSE Linux Enterprise Server 10 SP2 (64-bit). In the following tables bold values are published at http://www.spec.org. The values marked with „(est.)“ are estimated values.

* SPEC®, SPECint®, SPECfp® and the SPEC logo are registered trademarks of the Standard Performance Evaluation Corporation (SPEC).

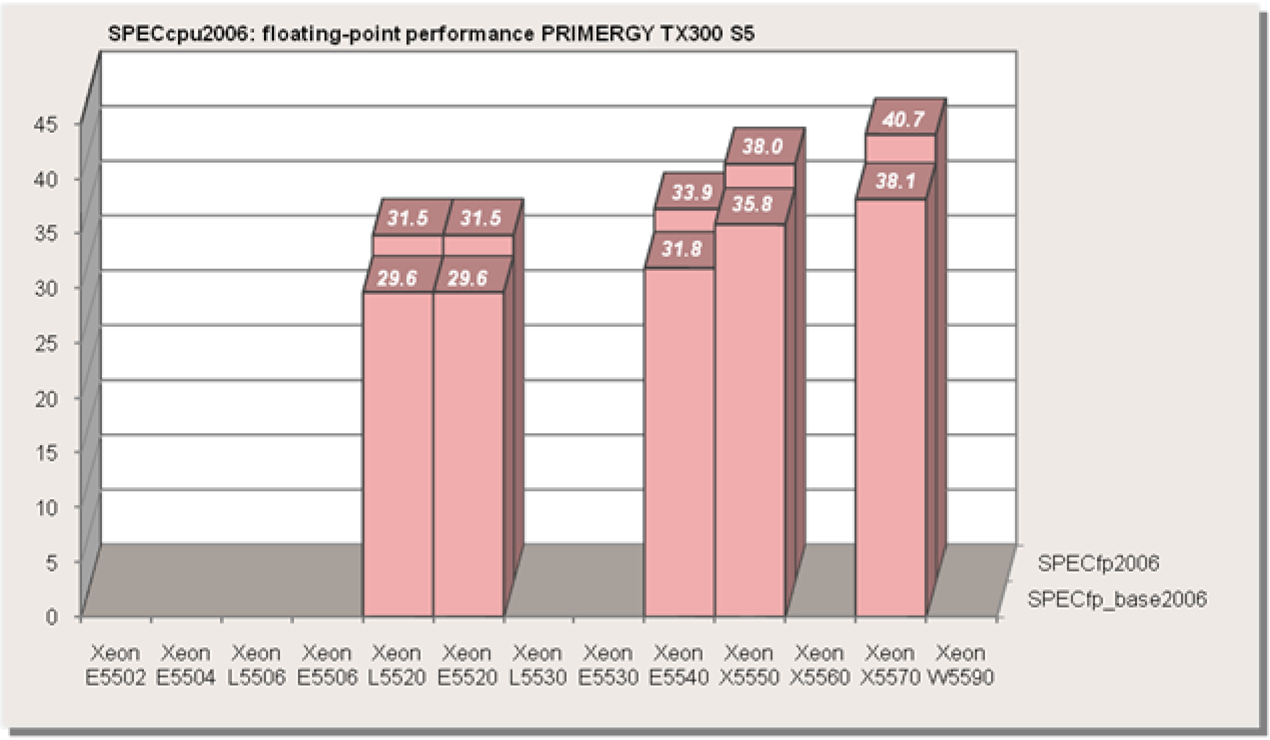
Processor	Cores	GHz	L3 cache	Bus	TDP	SPECint_base2006 2 chips	SPECint2006 2 chips
Xeon E5502	2	1.87	4 MB	800 MHz	80 watt		
Xeon E5504	4	2	4 MB	800 MHz	80 watt		
Xeon L5506	4	2.13	4 MB	800 MHz	60 watt		
Xeon E5506	4	2.13	4 MB	800 MHz	80 watt		
Xeon L5520	4	2.27	8 MB	1067 MHz	60 watt	24.5 (est.)	27.3 (est.)
Xeon E5520	4	2.27	8 MB	1067 MHz	80 watt	24.5	27.3
Xeon L5530	4	2.40	8 MB	1067 MHz	60 watt		
Xeon E5530	4	2.40	8 MB	1067 MHz	80 watt		
Xeon E5540	4	2.53	8 MB	1067 MHz	80 watt	26.7	29.7
Xeon X5550	4	2.67	8 MB	1333 MHz	95 watt	29.5	32.9
Xeon X5560	4	2.80	8 MB	1333 MHz	95 watt		
Xeon X5570	4	2.93	8 MB	1333 MHz	95 watt	32.1	36.1
Xeon W5590	4	3.33	8 MB	1333 MHz	130 watt		



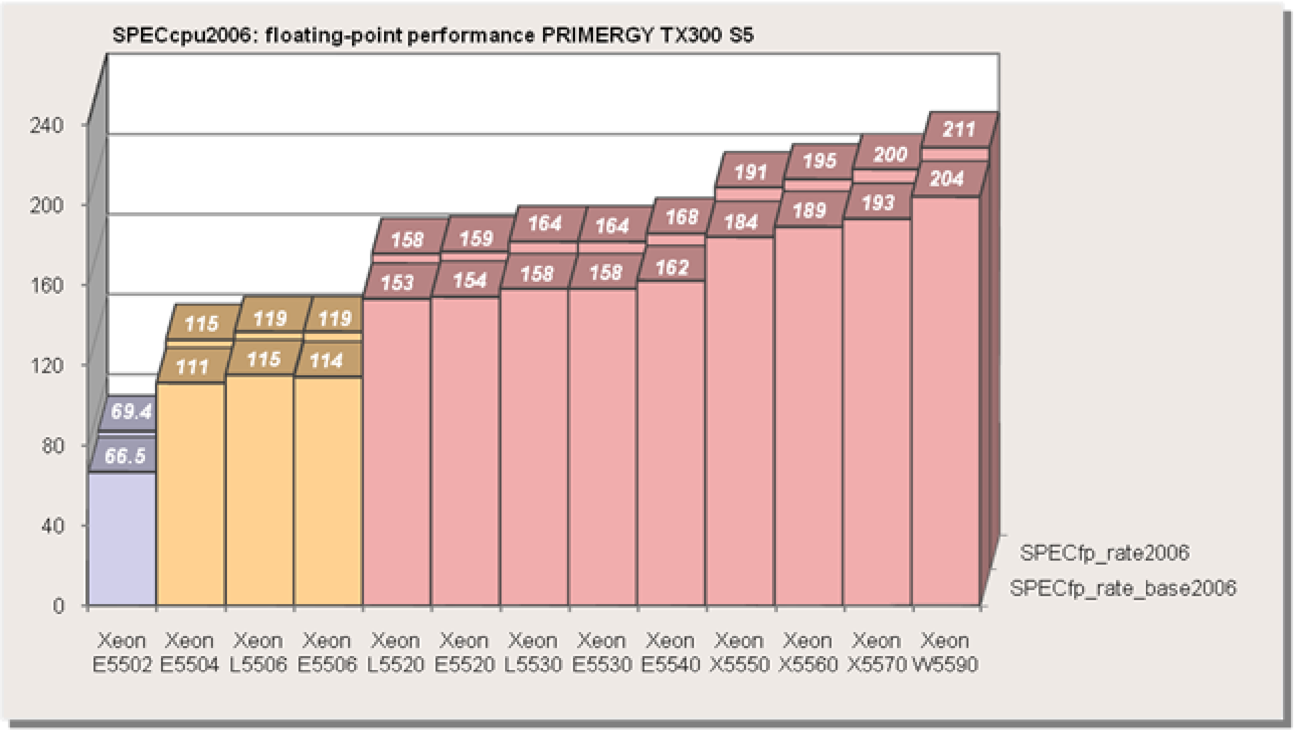
Processor	Cores	GHz	L3 cache	Bus	TDP	SPECint_rate_base2006		SPECint_rate2006	
						1 chip	2 chips	1 chip	2 chips
Xeon E5502	2	1.87	4 MB	800 MHz	80 watt		**65.3**		**70.4**
Xeon E5504	4	2	4 MB	800 MHz	80 watt		**126**		**135**
Xeon L5506	4	2.13	4 MB	800 MHz	60 watt		**132**		**142**
Xeon E5506	4	2.13	4 MB	800 MHz	80 watt		**131**		**141**
Xeon L5520	4	2.27	8 MB	1067 MHz	60 watt	95.7 (est.)	**186**	102 (est.)	**200**
Xeon E5520	4	2.27	8 MB	1067 MHz	80 watt	95.7	**185**	102	**200**
Xeon L5530	4	2.40	8 MB	1067 MHz	60 watt		**192**		**207**
Xeon E5530	4	2.40	8 MB	1067 MHz	80 watt		**192**		**207**
Xeon E5540	4	2.53	8 MB	1067 MHz	80 watt	102	**198**	110	**214**
Xeon X5550	4	2.67	8 MB	1333 MHz	95 watt	115	**226**	124	**243**
Xeon X5560	4	2.80	8 MB	1333 MHz	95 watt		**230**		**247**
Xeon X5570	4	2.93	8 MB	1333 MHz	95 watt	122	**240**	130	**257**
Xeon W5590	4	3.33	8 MB	1333 MHz	130 watt		**255**		**274**



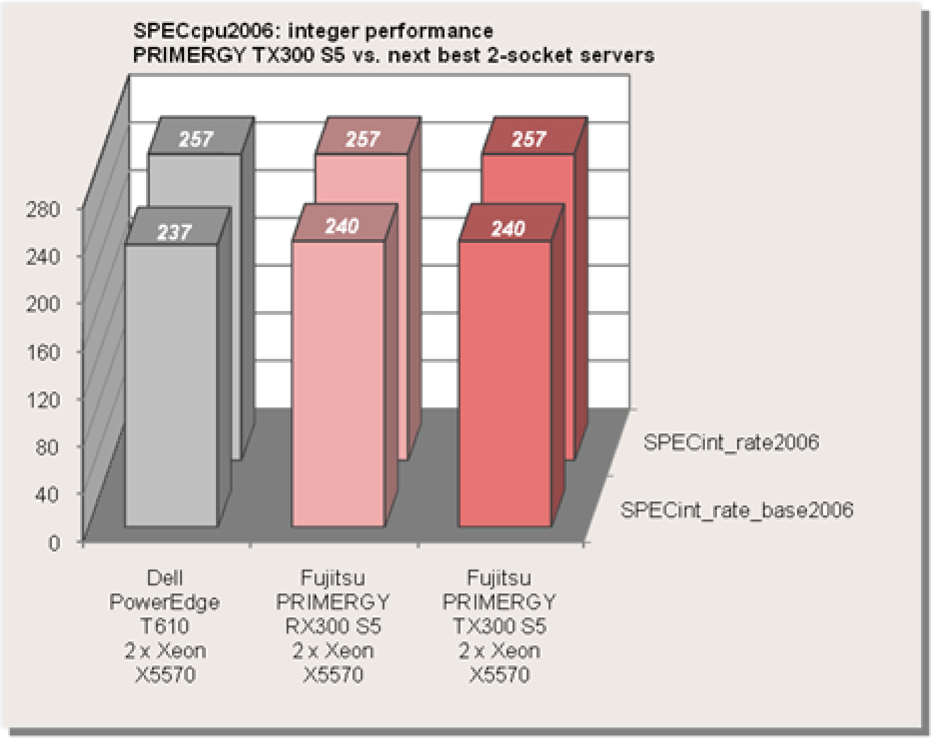
Processor	Cores	GHz	L3 cache	Bus	TDP	SPECfp_base2006 2 chips	SPECfp2006 2 chips
Xeon E5502	2	1.87	4 MB	800 MHz	80 watt		
Xeon E5504	4	2	4 MB	800 MHz	80 watt		
Xeon L5506	4	2.13	4 MB	800 MHz	60 watt		
Xeon E5506	4	2.13	4 MB	800 MHz	80 watt		
Xeon L5520	4	2.27	8 MB	1067 MHz	60 watt	29.6 (est.)	31.5 (est.)
Xeon E5520	4	2.27	8 MB	1067 MHz	80 watt	29.6	31.5
Xeon L5530	4	2.40	8 MB	1067 MHz	60 watt		
Xeon E5530	4	2.40	8 MB	1067 MHz	80 watt		
Xeon E5540	4	2.53	8 MB	1067 MHz	80 watt	31.8	33.9
Xeon X5550	4	2.67	8 MB	1333 MHz	95 watt	35.8	38.0
Xeon X5560	4	2.80	8 MB	1333 MHz	95 watt		
Xeon X5570	4	2.93	8 MB	1333 MHz	95 watt	38.1	40.7
Xeon W5590	4	3.33	8 MB	1333 MHz	130 watt		

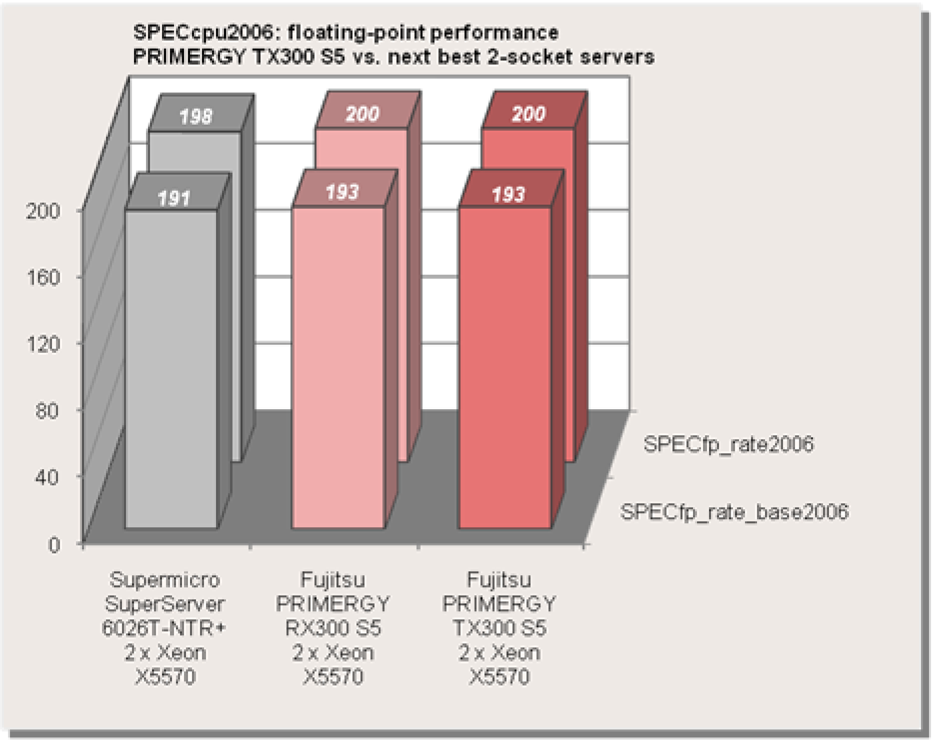


Processor	Cores	GHz	L3 cache	Bus	TDP	SPECfp_rate_base2006		SPECfp_rate2006	
						1 chip	2 chips	1 chip	2 chips
Xeon E5502	2	1.87	4 MB	800 MHz	80 watt		**66.5**		**69.4**
Xeon E5504	4	2	4 MB	800 MHz	80 watt		**111**		**115**
Xeon L5506	4	2.13	4 MB	800 MHz	60 watt		**115**		**119**
Xeon E5506	4	2.13	4 MB	800 MHz	80 watt		**114**		**119**
Xeon L5520	4	2.27	8 MB	1067 MHz	60 watt	79.1 (est.)	**153**	81.8 (est.)	**158**
Xeon E5520	4	2.27	8 MB	1067 MHz	80 watt	79.1	**154**	81.8	**159**
Xeon L5530	4	2.40	8 MB	1067 MHz	60 watt		**158**		**164**
Xeon E5530	4	2.40	8 MB	1067 MHz	80 watt		**158**		**164**
Xeon E5540	4	2.53	8 MB	1067 MHz	80 watt	83.5	**162**	86.4	**168**
Xeon X5550	4	2.67	8 MB	1333 MHz	95 watt	95.9	**184**	99.1	**191**
Xeon X5560	4	2.80	8 MB	1333 MHz	95 watt		**189**		**195**
Xeon X5570	4	2.93	8 MB	1333 MHz	95 watt	99.0	**193**	103	**200**
Xeon W5590	4	3.33	8 MB	1333 MHz	130 watt		**204**		**211**



On March 31, 2009 the PRIMERGY TX300 S5 achieved first place in the 2-socket server category in the benchmarks SPECint_rate_base2006, SPECint_rate2006, SPECfp_rate_base2006 and SPECfp_rate2006.[1] [2]

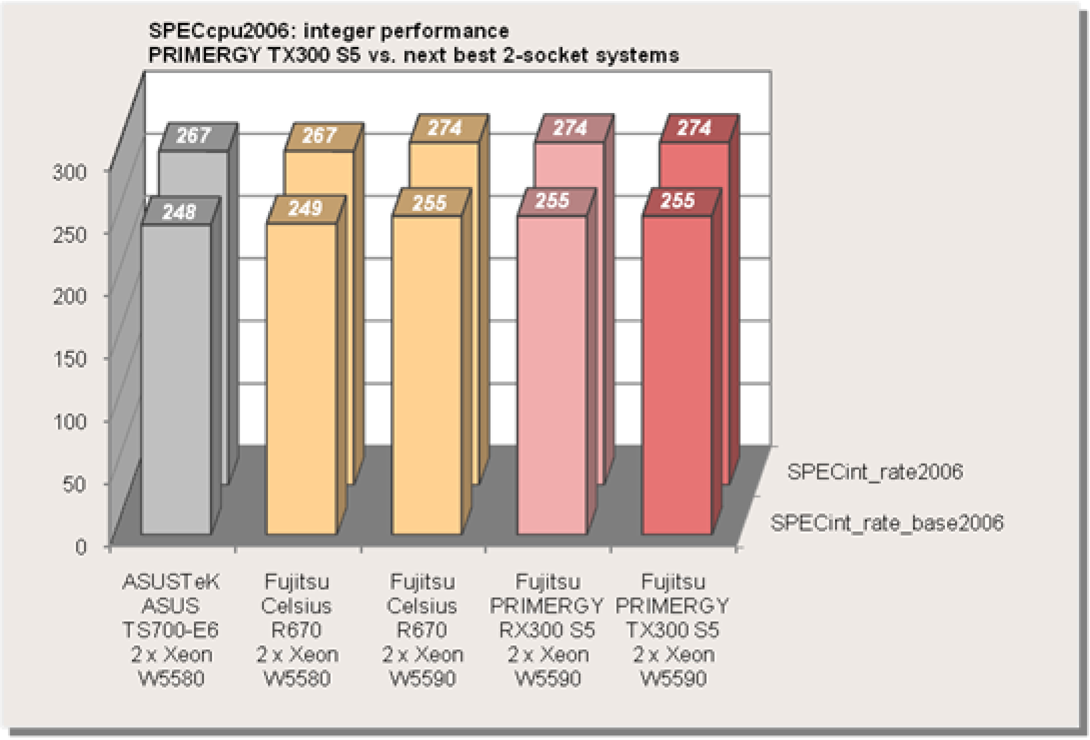


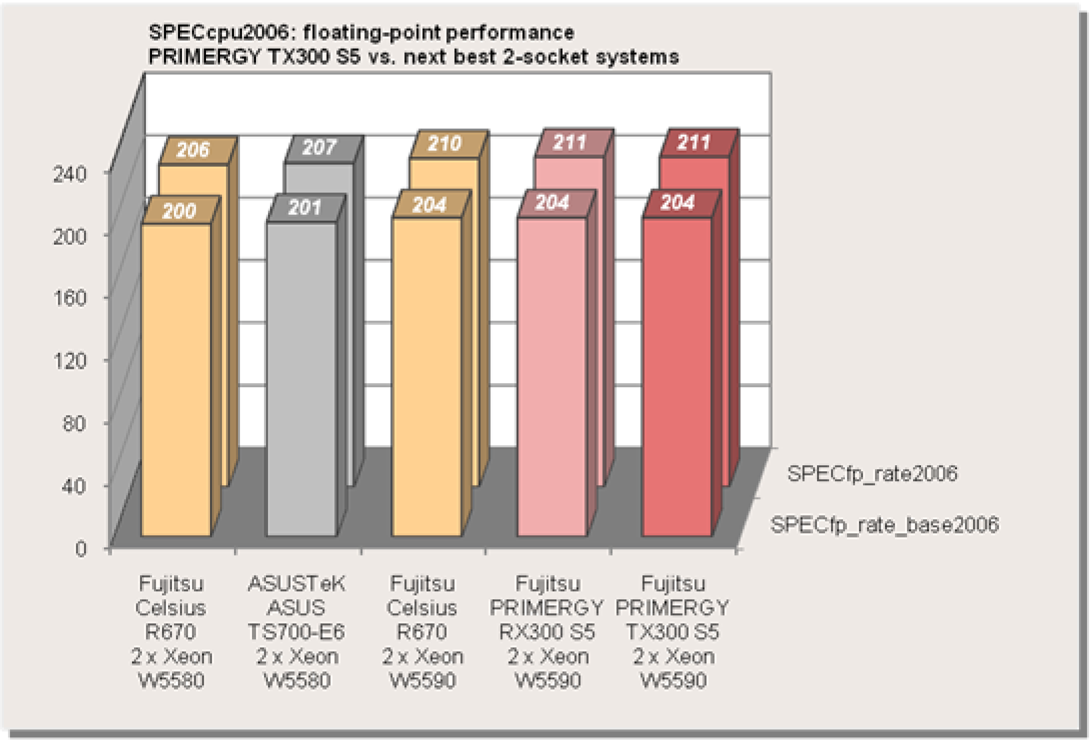


1 Competitive benchmark results stated above reflect results published as of March 31, 2009. The comparison presented above is based on the best performing 2-socket servers currently shipping by Dell and Fujitsu. For the latest SPECint_rate_base2006 and SPECint_rate2006 benchmark results, visit http://www.spec.org/cpu2006/results.

2 Competitive benchmark results stated above reflect results published as of March 31, 2009. The comparison presented above is based on the best performing 2-socket servers currently shipping by Supermicro and Fujitsu. For the latest SPECfp_rate_base2006 and SPECfp_rate2006 benchmark results, visit http://www.spec.org/cpu2006/results.

On August 18 2009 the PRIMERGY TX300 S5 came first in the rankings of the 2-socket system category in the benchmarks SPECint_rate_base2006, SPECint_rate2006, SPECfp_rate_base2006 and SPECfp_rate2006.[3 4]

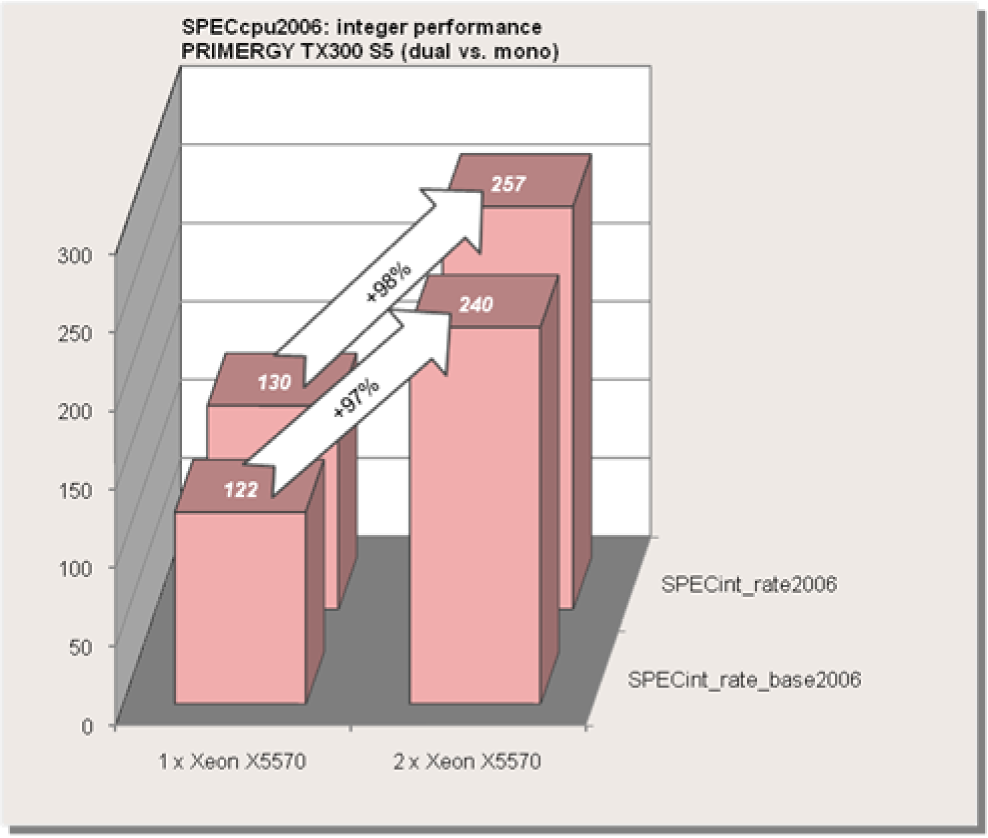


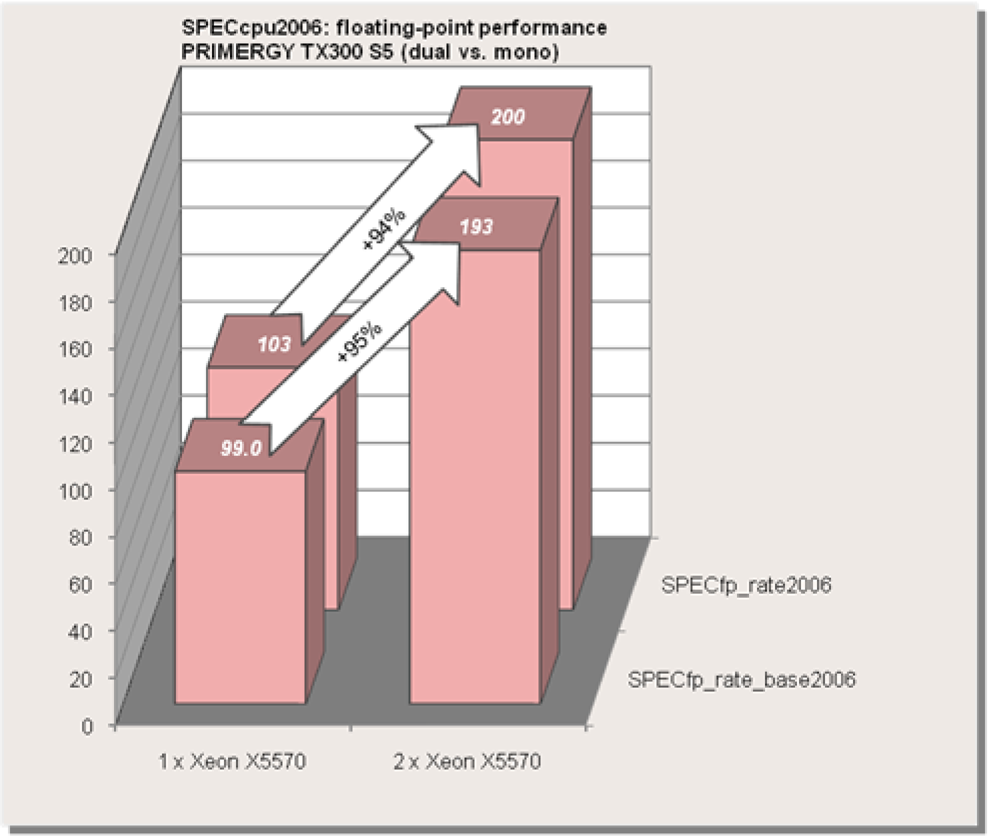


3 Competitive benchmark results stated above reflect results published as of August 18, 2009. The comparison presented above is based on the best performing 2-socket servers currently shipping by ASUSTek and Fujitsu. For the latest SPECint_rate_base2006 and SPECint_rate2006 benchmark results, visit http://www.spec.org/cpu2006/results.

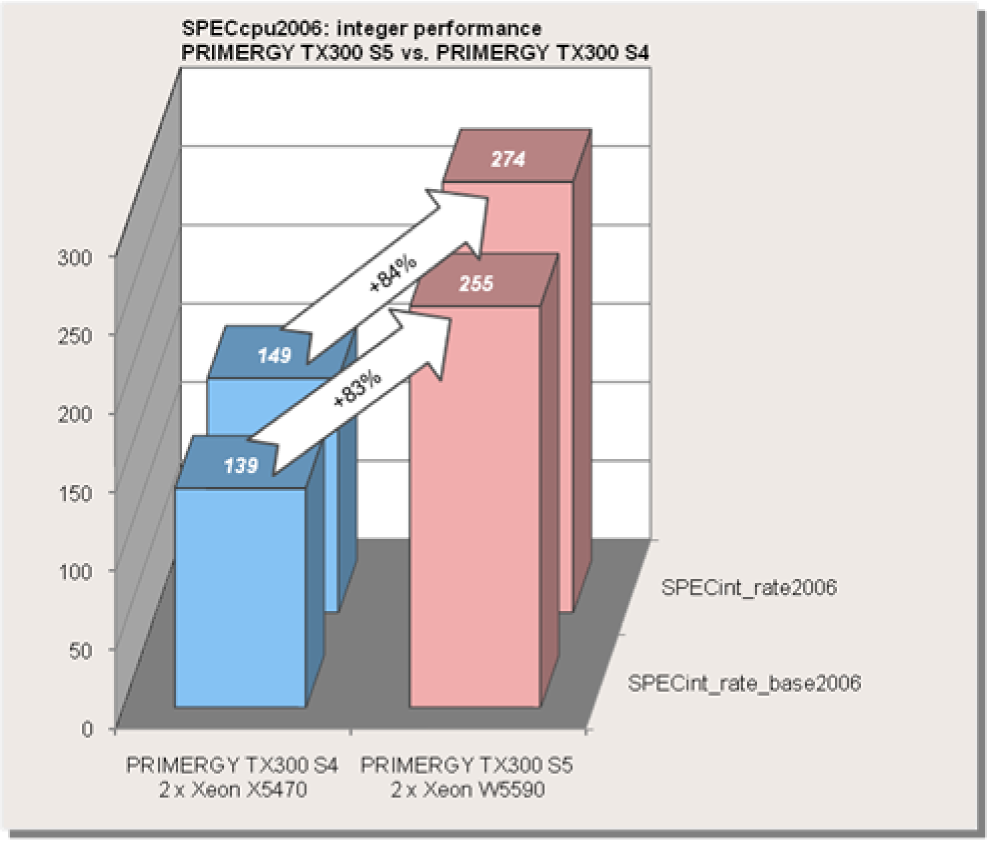
4 Competitive benchmark results stated above reflect results published as of August 18, 2009. The comparison presented above is based on the best performing 2-socket servers currently shipping by ASUSTek and Fujitsu. For the latest SPECfp_rate_base2006 and SPECfp_rate2006 benchmark results, visit http://www.spec.org/cpu2006/results.

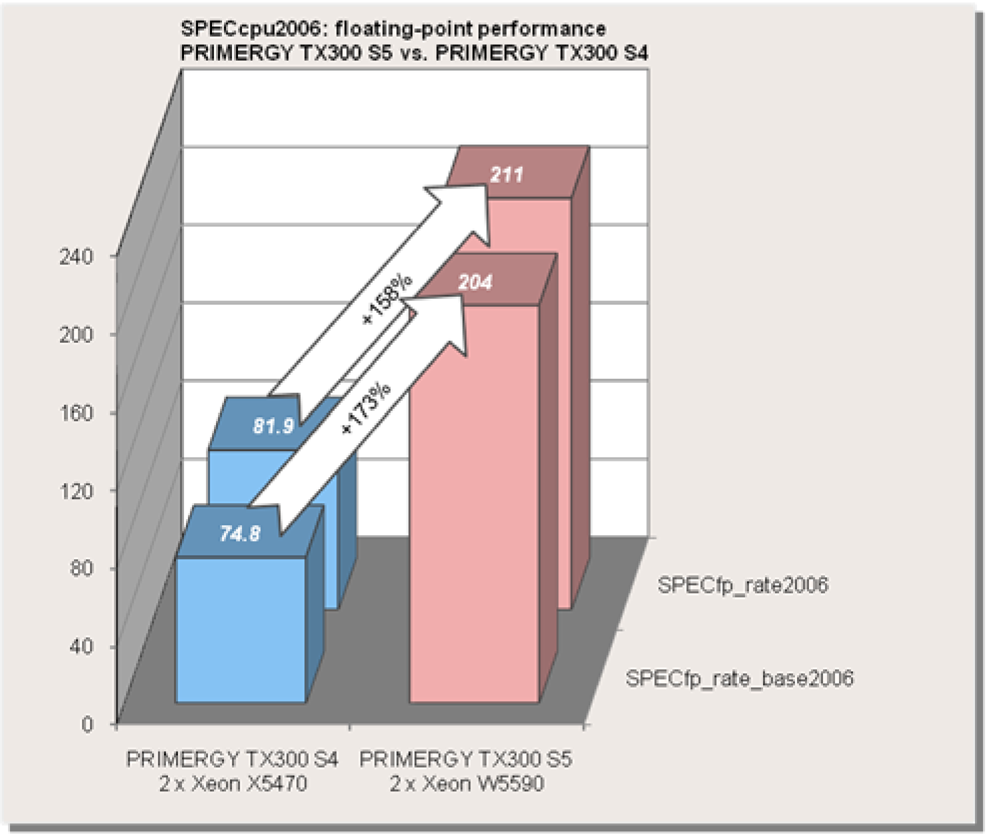
For both the integer and the floating-point test suite the throughput with two processors is almost twice as large as that with one processor.





The diagrams below illustrate the performance of the PRIMERGY TX300 S5 compared with its predecessor PRIMERGY TX300 S4, both in their highest performance configurations





Benchmark environment*

All SPECcpu2006 measurements were performed on a PRIMERGY TX300 S5 with the following hardware and software configuration:

Hardware	
Model	PRIMERGY TX300 S5
CPU	Xeon E5502, E5504, L5506, E5506, L5520, E5520, L5530, E5530, E5540, X5550, X5560, X5570 and W5590
Number of CPUs	1 chip: Xeon E5520, E5540, X5550 and X5570: 4 cores, 1 chip, 4 cores/chip 2 chips: Xeon E5502: 4 cores, 2 chips, 2 cores/chip others: 8 cores, 2 chips, 4 cores/chip
Primary Cache	32 KB instruction + 32 KB data on chip, per core
Secondary Cache	256 kB on chip, per core
Other Cache	Xeon E5502, E5504, L5506 and E5506: 4 MB (I+D) on chip, per chip others: 8 MB (I+D) on chip, per chip
Memory	Xeon L5506: PC3-8500R Xeon E5502, E5504, E5506, L5520, E5520, E5540, X5550, X5570: PC3-10600R
Software	
Operating System	SUSE Linux Enterprise Server 10 SP2 (64-bit)
Compiler	Intel C++/Fortran Compiler 11.0

* Some components may not be available in all countries / sales regions.

SPECjbb2005*

Benchmark description

SPECjbb2005 is a Java business benchmark that focuses on the performance of Java server platforms. It is essentially a modernized version of SPECjbb2000 with the main differences being:

- The transactions have become more complex in order to cover a greater functional scope.
- The working set of the benchmark has been enlarged to the extent that the total system load has increased.
- SPECjbb2000 allows only one active Java Virtual Machine instance (JVM), whereas SPECjbb2005 permits several instances, which in turn achieves greater closeness to reality, particularly with large systems.

On the software side SPECjbb2005 measures the implementations of the JVM, JIT (Just-In-Time) compiler, garbage collection, threads and some aspects of the operating system. As far as hardware is concerned, it measures the efficiency of the CPUs and caches, the memory subsystem and the scalability of shared memory systems (SMP). Disk and network I/O are irrelevant.

SPECjbb2005 emulates a 3-tier client/server system that is typical for modern business process applications with emphasis on the middle tier system:

- Clients generate the load, consisting of driver threads, which on the basis of the TPC-C benchmark generate OLTP accesses to a database without thinking times.
- The middle-tier system implements the business processes and the updating of the database.
- The database takes on the data management and is emulated by Java objects that are in the memory. Transaction logging is implemented on an XML basis.

The major advantage of this benchmark is that it includes all three tiers that run together on a single host. The performance of the middle tier is measured, thus avoiding large-scale hardware installations and making direct comparisons possible between SPECjbb2005 results of different systems. Client and database emulation are also written in Java.

SPECjbb2005 only needs the operating system as well as a Java Virtual Machine with J2SE 5.0 features.

The scaling unit is a warehouse with approx. 25 MB Java objects. Precisely one Java thread per warehouse executes the operations on these objects. The business operations are assumed by TPC-C:

- New Order Entry
- Payment
- Order Status Inquiry
- Delivery
- Stock Level Supervision
- Customer Report

However, these are the only features SPECjbb2005 and TPC-C have in common. The results of the two benchmarks are not comparable.

SPECjbb2005 has 2 performance metrics:

- bops (business operations per second) is the overall rate of all business operations performed per second.
- bops/JVM is the ratio of the first metrics and the number of active JVM instances.

In comparisons of various SPECjbb2005 results it is necessary to state both metrics.

The following rules, according to which a compliant benchmark run has to be performed, are the basis for these metrics:

* SPEC®, SPECjbb® and the SPEC logo are registered trademarks of the Standard Performance Evaluation Corporation (SPEC).

A compliant benchmark run consists of a sequence of measuring points with an increasing number of warehouses (and thus of threads) with the number in each case being increased by one warehouse. The run is started at one warehouse up through 2*MaxWhm but not less than 8 warehouses. MaxWhm is the number of warehouses with the highest operation rate per second the benchmark expects. Per default the benchmark equates MaxWH with the number of CPUs visible by the operating system.

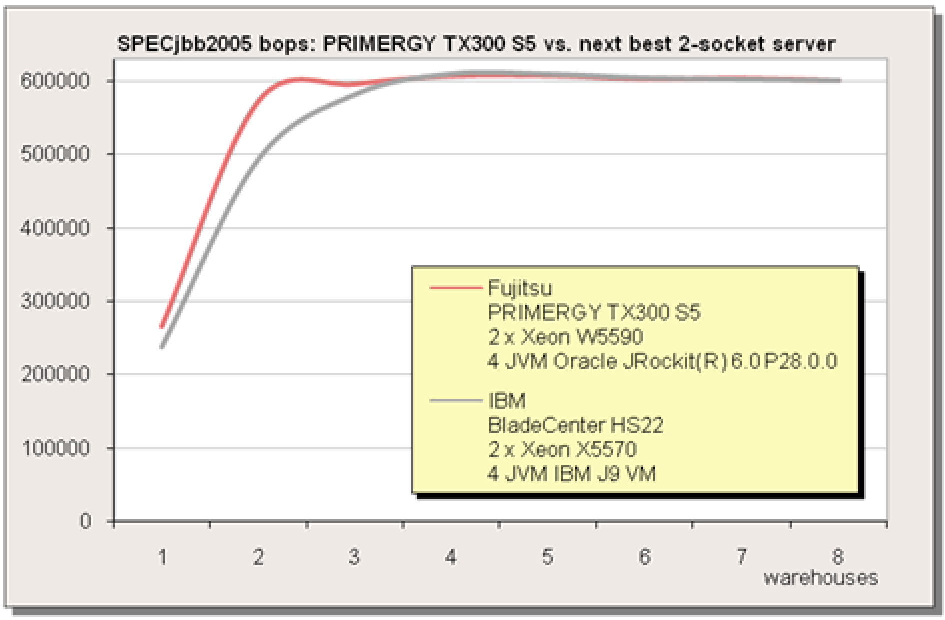
The metrics bops is the arithmetic average of all measured operation rates with between MaxWhm warehouses and 2*MaxWhm warehouses.

Benchmark results

In March 2009 the PRIMERGY TX300 S5 was measured with two Xeon X5570 processors and a memory of 24 GB PC3-10600R DDR3-SDRAM. The measurement was taken under Windows Server 2008 Enterprise x64 Edition. Two instances of JRockit(R) 6 R28.0.0 (build P28.0.0-14-111048-1.6.0_05-20090303-1104-windows-x86_64) from Oracle were used as JVM for the measurement. The benchmark result includes all the measuring values from 8 to 16 warehouses.

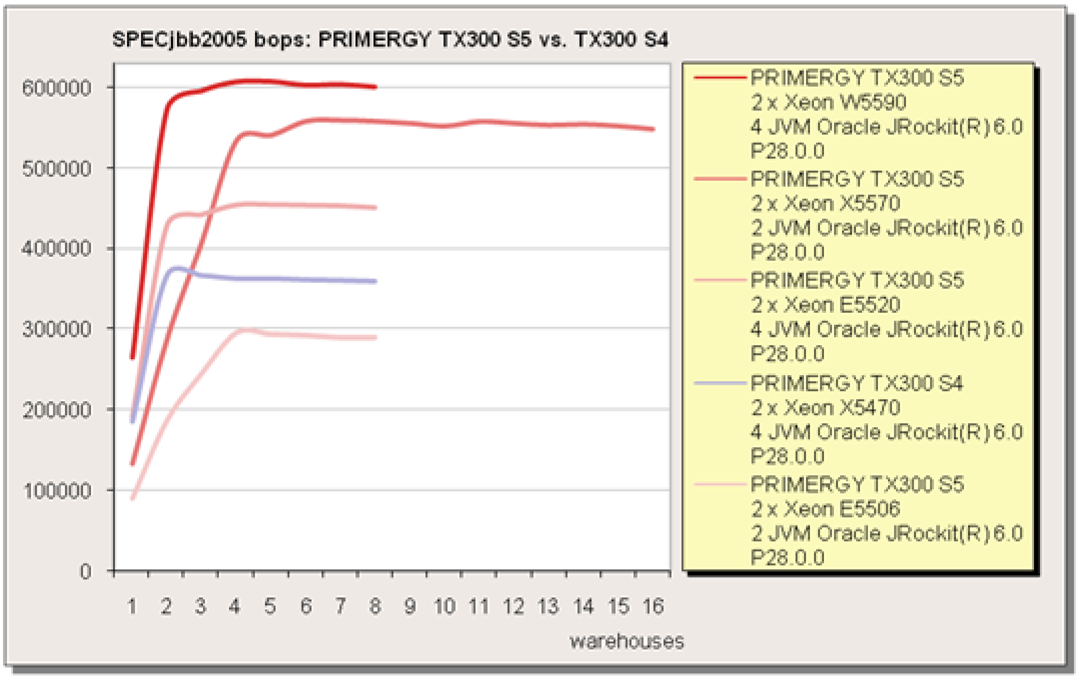
In May 2009 the PRIMERGY TX300 S5 was measured with two processors - Xeon E5506 and E5520. In comparison with the measurement of March 2009, Windows Server 2008 Enterprise x64 Edition SP2 and a more recent JVM version were used. In the case of the Xeon E5506 two instances of JRockit(R) 6 P28.0.0 (build P28.0.0-29-114096-1.6.0_11-20090427-1759-windows-x86_64) from Oracle were used and in the case of the Xeon E5520 four instances were used. All the measurement values of 4 to 8 warehouses were incorporated in the benchmark result.

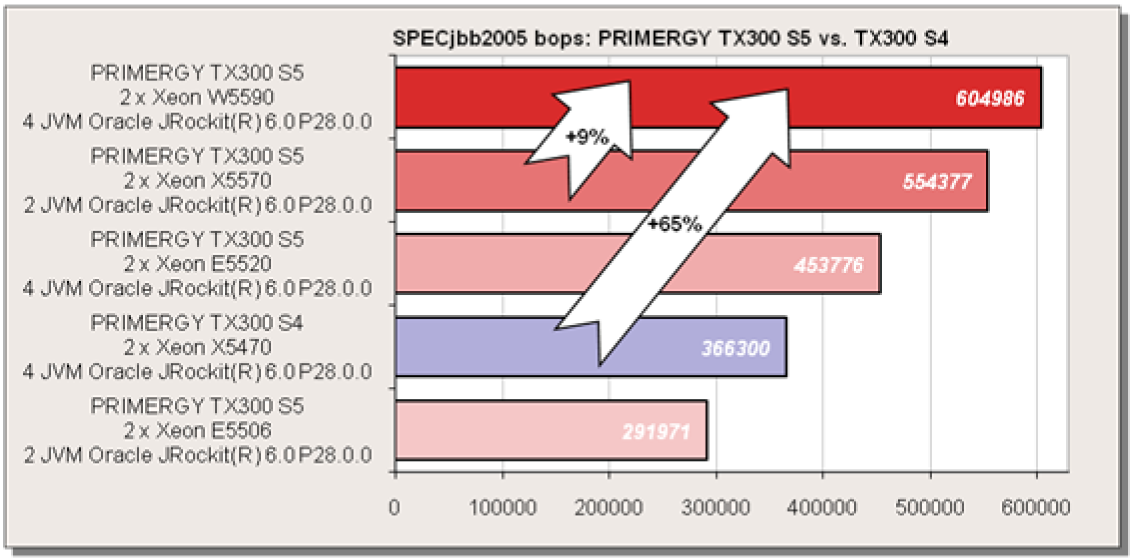
In August 2009 the PRIMERGY TX300 S5 was measured with two Xeon W5590 processors and a memory configuration of 48 GB PC3-10600R DDR3-SDRAM. The configuration otherwise corresponded to the measurement of May 2009 with the Xeon E5520. Here the PRIMERGY TX300 S5 and the PRIMERGY RX300 S5 achieved the best result of all 2-socket servers.*



* Competitive benchmark results stated above reflect results published as of August 26, 2009. The comparison presented above is based on the best performing servers with 2 processors currently shipping by IBM and Fujitsu. For the latest SPECjbb2005 benchmark results, visit http://www.spec.org/jbb2005/results.

When the PRIMERGY TX300 S5 is compared to its predecessor the PRIMERGY TX300 S4, a throughput increase of +65% exists in the respective top performance configurations.





Benchmark environment*

The SPECjbb2005 measurement was performed on a PRIMERGY TX300 S5 with the following hardware and software configuration:

Measurement of March 2009:

Hardware	
Model	PRIMERGY TX300 S5
CPU	Xeon X5570
Number of chips	2 chips, 8 cores, 4 cores per chip
Primary Cache	32 kB instruction + 32 kB data on chip, per core
Secondary Cache	¼ MB (I+D) on chip, per core
Other Cache	8 MB (I+D) on chip, per chip
Memory	6 x 4 GB PC3-10600R DDR3-SDRAM
Software	
Operating System	Windows Server 2008 Enterprise x64 Edition
JVM Version	Oracle JRockit(R) 6 P28.0.0 (build P28.0.0-14-111048-1.6.0_05-20090303-1104-windows-x86_64)

Measurement of May 2009:

Hardware	
Model	PRIMERGY TX300 S5
CPU	Xeon E5506, E5520
Number of chips	2 chips, 8 cores, 4 cores per chip
Primary Cache	32 kB instruction + 32 kB data on chip, per core
Secondary Cache	¼ MB (I+D) on chip, per core
Other Cache	Xeon E5506: 4 MB (I+D) on chip, per chip Xeon E5520: 8 MB (I+D) on chip, per chip
Memory	6 x 4 GB PC3-10600R DDR3-SDRAM
Software	
Operating System	Windows Server 2008 Enterprise x64 Edition SP2
JVM Version	Oracle JRockit(R) 6 P28.0.0 (build P28.0.0-14-111048-1.6.0_05-20090303-1104-windows-x86_64)

Measurement of August 2009:

Hardware	
Model	PRIMERGY TX300 S5
CPU	Xeon W5590
Number of chips	2 chips, 8 cores, 4 cores per chip
Primary Cache	32 kB instruction + 32 kB data on chip, per core
Secondary Cache	¼ MB (I+D) on chip, per core
Other Cache	8 MB (I+D) on chip, per chip
Memory	12 x 4 GB PC3-10600R DDR3-SDRAM
Software	
Operating System	Windows Server 2008 Enterprise x64 Edition SP2
JVM Version	Oracle JRockit(R) 6 P28.0.0 (build P28.0.0-14-111048-1.6.0_05-20090303-1104-windows-x86_64)

* Some components may not be available in all countries / sales regions.

StorageBench

Benchmark description

To estimate the capability of disk subsystems Fujitsu Technology Solutions defined a benchmark called StorageBench to compare the different storage systems connected to a system. To do this StorageBench makes use of the Iometer measuring tool developed by Intel combined with a defined set of load profiles that occur in real customer applications and a defined measuring scenario.

Measuring tool

Since the end of 2001 Iometer has been a project at http://SourceForge.net and is ported to various platforms and enhanced by a group of international developers. Iometer consists of a user interface for Windows systems and the so-called "dynamo" which is available for various platforms. For some years now it has been possible to download these two components under "Intel Open Source License" from http://www.iometer.org/ or http://sourceforge.net/projects/iometer.

Iometer gives you the opportunity to reproduce the behavior of real applications as far as accesses to IO subsystems are concerned. For this purpose, you can among other things configure the block sizes to be used, the type of access, such as sequential read or write, random read or write and also combinations of these. You can also configure the number of simultaneous accesses ("Outstanding IOs"). As a result Iometer provides a text file with comma separated values (.csv) containing basic parameters, such as throughput per second, transactions per second and average response time for the respective access pattern. This method permits the efficiency of various subsystems with certain access patterns to be compared. Iometer is in a position to access not only subsystems with a file system, but also so-called raw devices.

With Iometer it is possible to simulate and measure the access patterns of various applications, but the file cache of the operating system remains disregarded and operation is in blocks on a single test file.

Load profile

The manner in which applications access the mass storage system considerably influences the performance of a storage system. Examples of various access patterns of a number of applications:

Application	Access pattern
Database (data transfer)	random, 67% read, 33% write, 8 KB (SQL Server)
Database (log file)	sequential, 100% write, 64 KB blocks
Backup	sequential, 100% read, 64 KB blocks
Restore	sequential, 100% write, 64 KB blocks
Video streaming	sequential, 100% read, blocks ≥ 64 KB
File server	random, 67% read, 33% write, 64 KB blocks
Web server	random, 100% read, 64 KB blocks
Operating system	random, 40% read, 60% write, blocks ≥ 4 KB
File copy	random, 50% read, 50% write, 64 KB blocks

From this four distinctive profiles were derived:

Load profile	Access	Access pattern		Block size	Outstanding IOs	Load tool
		read	write			
Streaming	sequential	100%		64 KB	3	Iometer
Restore	sequential		100%	64 KB	3	Iometer
Database	random	67%	33%	8 KB	3	Iometer
File server	random	67%	33%	64 KB	3	Iometer

All four profiles were generated with Iometer.

Measurement scenario

In order to obtain comparable measurement results it is important to perform all the measurements in identical, reproducible environments. This is why StorageBench is based, in addition to the load profile described above, on the following regulations:

- Since real-life customer configurations work only in exceptional situations with raw devices, performance measurements of internal disks are always conducted on disks containing file systems. NTFS is used for Windows and ext3 for Linux, even if higher performance could possibly be achieved with other file systems or raw devices.
- Hard disks are among the most error-prone components of a computer system. This is why RAID controllers are used in server systems in order to prevent data loss through hard disk failure. Here several hard disks are put together to form a “Redundant Array of Independent Disks”, known as RAID in short – with the data being spread over several hard disks in such a way that all the data is retained even if one hard disk fails – except with RAID 0. The most usual methods of organizing hard disks in arrays are the RAID levels RAID 0, RAID 1, RAID 5, RAID 6, RAID 10, RAID 50 and RAID 60. Information about the basics of various RAID arrays is to be found in the paper Performance Report - Modular RAID for PRIMERGY.

  Depending on the number of disks and the installed controller, the possible RAID configurations are used for the StorageBench analyses of the PRIMERGY servers. For systems with two hard disks we use RAID 1 and RAID 0, for three and more hard disks we also use RAID 1E and RAID 5 and, where applicable, further RAID levels – provided that the controller supports these RAID levels.
- Regardless of the size of the hard disk, a measurement file with the size of 8 GB is always used for the measurement.
- In the evaluation of the efficiency of I/O subsystems, processor performance and memory configuration do not play a significant role in today’s systems - a possible bottleneck usually affects the hard disks and the RAID controller, and not CPU and memory. Therefore, various configuration alternatives with CPU and memory need not be analyzed under StorageBench.

Measurement results

For each load profile StorageBench provides various key indicators: e.g. “data throughput” in megabytes per second, in short MB/s, “transaction rate” in I/O operations per second, in short IO/s, and “latency time” or also “mean access time” in ms. For sequential load profiles data throughput is the normal indicator, whereas for random load profiles with their small block sizes the transaction rate is normally used. Throughput and transaction rate are directly proportional to each other and can be calculated according to the formula

Data throughput [MB/s]	*= Transaction rate [Disk-I/O s^{-1}] × Block size [MB]*
Transaction rate [Disk-I/O s^{-1}]	*= Data throughput [MB/s] / Block size [MB]*

Benchmark results

The PRIMERGY TX300 S5 is equipped with controllers from the “Modular RAID” family. The variety of the RAID solutions enables the user to choose the right controller for his application scenario.

The PRIMERGY TX300 S5 has the following RAID solutions to offer:

1. RAID Controller LSI MegaRAID SAS 1068

   The controller is supplied as a PCI Express card. The maximum number of SATA and SAS hard disks that can be connected to the controller is eight. Support is provided for RAID levels 0, 1 and 1E. The controller does not have a cache.

2. RAID Controller LSI MegaRAID SAS 1078

   The controller is supplied as a PCI Express card and offers the user a complete RAID solution. Both SATA and SAS hard disks can be connected. Support is provided for RAID levels 0, 1, 5, 6, 10, 50 and 60. Two different versions of this controller are on offer with either a 256 MB or 512 MB cache. The controller cache can be protected against power failure by an optional battery backup unit (BBU). The controller supports up to 240 hard disks.

Various SATA and SAS hard disks can be connected to these controllers. Depending on the performance required, it is possible to select the appropriate disk subsystem. And depending on the model version the PRIMERGY TX300 S5 offers six hot-plug bays for 3½" hard disks or twelve hot-plug bays for 2½" hard disks. Optionally, an extension box with a further 8 × 2½" hot-plug bays is available for the 2½" version or an extension box with a further 2 × 3½" hot-plug bays for the 3½" version.

The following hard disks can be chosen for the PRIMERGY TX300 S5:

- 2½" SAS hard disks with a capacity of 73 GB and 146 GB (10 krpm)
- 2½" SAS hard disks with a capacity of 36 GB and 73 GB (15 krpm)
- 3½" SAS hard disks with a capacity of 73 GB, 146 GB, 300 GB and 450 GB (15 krpm)
- 3½" SATA hard disks with a capacity of 250 GB, 500 GB, 750 GB and 1 TB (7.2 krpm)

LSI MegaRAID SAS 1068

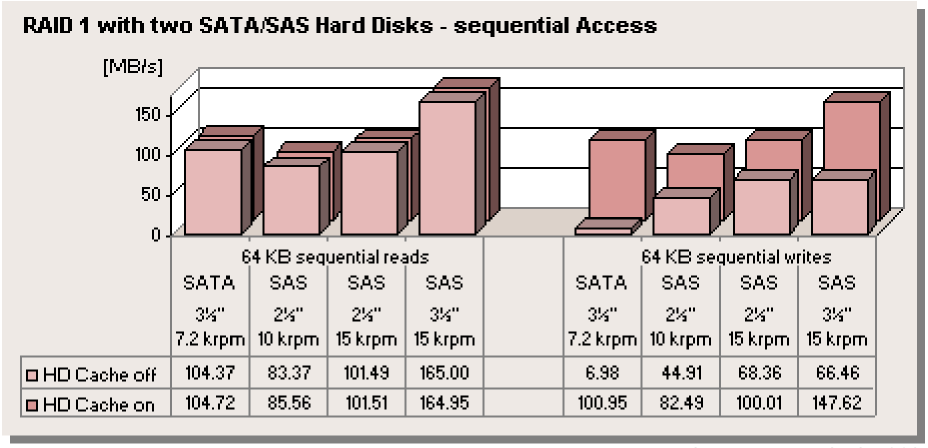
The performance of the available hard disk types on the LSI MegaRAID SAS 1068 controller is compared below.

This controller does not have a controller cache. This is why only the impact of the disk cache parameters was examined in the measurements and the measurements for the hard disk comparison were in each case performed with and without a disk cache.

The hard disk cache has influence on disk I/O performance. This is frequently seen as a safety problem in the event of a power failure and is therefore disabled. Nevertheless, it was for a good reason integrated by the hard disk manufacturers to increase write performance. For performance reasons it is advisable to enable the disk cache for the SATA hard disks in particular, which in comparison with the SAS hard disks rotate slowly. The by far larger cache for I/O accesses and thus a potential safety risk (data loss) in the event of a power failure is situated in the main memory and is administered by the operating system. To prevent data losses it is advisable to equip the system with an uninterruptible power supply (UPS).

In the test setup two hard disks were connected to the controller and configured as a RAID 1. In the measurements all hard disk types currently available for the PRIMERGY TX300 S5 were analyzed. The throughputs of the individual hard disk types in RAID 1 are compared below with different access patterns.

The diagram shows that as the rotational speed increases, the throughput for sequential reads and writes with a 64 KB block size rises.



	64 KB sequential reads					64 KB sequential writes			
	SATA 3½" 7.2 krpm	SAS 2½" 10 krpm	SAS 2½" 15 krpm	SAS 3½" 15 krpm		SATA 3½" 7.2 krpm	SAS 2½" 10 krpm	SAS 2½" 15 krpm	SAS 3½" 15 krpm
HD Cache off	104.37	83.37	101.49	165.00		6.98	44.91	68.36	66.46
HD Cache on	104.72	85.56	101.51	164.95		100.95	82.49	100.01	147.62

LSI MegaRAID SAS 1068

If for sequential read a hard disk with a rotational speed of 15 krpm is used instead of one with a speed of 10 krpm, the result for the 2½" hard disk is an increase in throughput of about 19% If you compare the throughputs of the 2½" and 3½" hard disks both with a rotational speed of 15 krpm, you can see that the throughput for the 3½" hard disk is about 63% higher than for the 2½" hard disk.

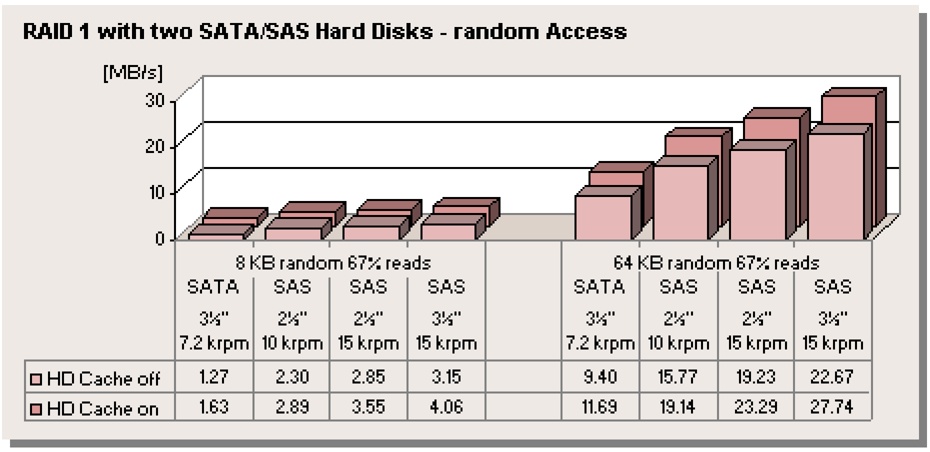
If you compare the 3½" SAS hard disk with the 3½" SATA hard disk, you can then see that the throughput of the SAS hard disk with 15 krpm is about 58% higher than the SATA hard disk with 7.2 krpm for sequential read.

If for sequential write with enabled disk cache a hard disk with a rotational speed of 15 krpm is used instead of one with a speed of 10 krpm, the result for the 2½" hard disk is an increase in throughput of about 21%. If you compare the throughputs of the 2½" and 3½" hard disks both with a rotational speed of 15 krpm, you can see that the throughput for the 3½" hard disk is about 48% higher than for the 2½" hard disk.

If you compare the 3½" SAS hard disk with the 3½" SATA hard disk, you can then see that the throughput of the SAS hard disk with 15 krpm is about 46% higher than the SATA hard disk with 7.2 krpm for sequential write and with an enabled disk cache.

A special increase in throughput for sequential write, up to 14-fold, can be achieved with the SATA hard disk by enabling the disk cache. The increase in throughput gained with SAS hard disks through enabling the disk cache is not so pronounced as with the SATA hard disks, but it is still significant. For the 2½" hard disks with 10 krpm throughput increases by 84% and by about 46% for the 2½" hard disks with 15 krpm. For the 3½" hard disks with 15 krpm the throughput increases by a factor 2.2.

The following diagram shows that for random access with 67% reads the disk cache also plays an important role in improving throughput.



	8 KB random 67% reads					64 KB random 67% reads			
	SATA 3½" 7.2 krpm	SAS 2½" 10 krpm	SAS 2½" 15 krpm	SAS 3½" 15 krpm		SATA 3½" 7.2 krpm	SAS 2½" 10 krpm	SAS 2½" 15 krpm	SAS 3½" 15 krpm
HD Cache off	1.27	2.30	2.85	3.15		9.40	15.77	19.23	22.67
HD Cache on	1.63	2.89	3.55	4.06		11.69	19.14	23.29	27.74

LSI MegaRAID SAS 1068

Through enabling the disk cache an increase in throughput of about 29% was achieved for random access with 8 KB blocks in RAID 1 with two 3½" 7.2 krpm SATA hard disks or 3½" 15 krpm SAS hard disks.

If you compare the throughput of the 2½" and 3½" 15 krpm SAS hard disks, it is then evident that the throughput of the 3½" hard disk for random access with 8 KB blocks and enabled disk cache is about 23% higher than that of the 2½" hard disk.

If you compare the throughput of the 3½" SAS hard disk with the 3½" SATA hard disk, you can then see that the throughput of the SAS hard disk with 15 krpm is about 2.5 times higher for random access with 8 KB blocks and enabled disk cache than in the SATA hard disk with 7.2 krpm.

Through enabling the disk cache an increase in throughput of up to 24% was achieved in the SATA hard disk with 64 KB blocks. In the SAS hard disks the improvement in throughput for 64 KB blocks is somewhat lower and lies a little above 20%.

If you compare the throughput of the 2½" and 3½" 15 krpm SAS hard disks, it is then evident that the throughput of the 3½" hard disk for random access with 64 KB blocks and enabled disk cache is about 19% higher than that of the 2½" hard disk.

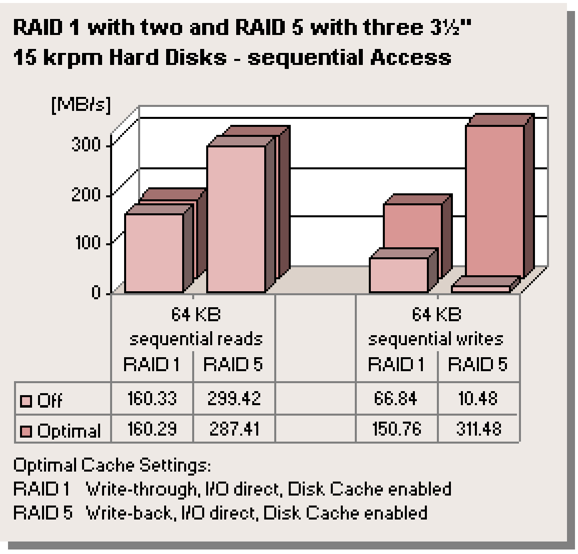
If you compare the throughput of the 3½" SAS hard disk with the 3½" SATA hard disk, you can then see that the throughput of the SAS hard disk with 15 krpm for random access with 64 KB blocks and enabled disk cache is about 2.4-fold higher than in the SATA hard disk with 7.2 krpm.

LSI MegaRAID SAS 1078

The RAID array defines the way in which data is treated as regards availability. How quickly the data is transferred in the respective RAID array context depends largely on the data throughput of the hard disks. The number of hard disks configured for the measurements in a RAID array was defined depending on the RAID level. Between two and three hard disks were used. To ensure that the hard disks do not represent a bottleneck when determining the efficiency of the controller under various cache settings, the measurements were performed with hard disks with a rotational speed of 15 krpm.

The throughput can in certain cases be considerably increased through the cache settings. However, these increases in throughput differ – depending on the data structure and access pattern. For the measurements the controller cache option “Read-Mode” is always set to “No Read-ahead” and the option ”I/O cache” is always set to “I/O direct”. The options “Write-Mode” and “Disk cache” were varied.

The following diagram shows the throughputs for sequential read and write with 64 KB blocks and for different cache settings in RAID 1 with two and in RAID 5 with three 3½" hard disks.



LSI MegaRAID SAS 1078 with 512 MB Cache

The sequential read throughput achieves very good values and does not depend on the cache settings.

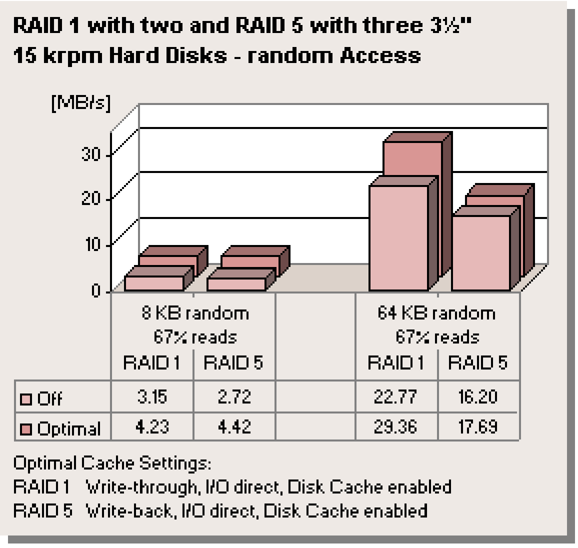
In contrast, the write throughput depends on the cache settings. In order to achieve optimal performance with RAID 1, it is necessary to use the “Disk cache enabled” option as the optimal cache setting. In our case the throughput was improved by a factor 2.3 using sequential write.

The importance of optimal cache settings for a good performance can be seen particularly clearly with RAID 5. The diagram shows that sequential write throughput increases considerably, by a factor of 30, as a result of enabling the controller cache with the option “Write-back” and the disk cache with the option "enabled".

To achieve optimal throughput for random access with RAID 1 it is important to set the write mode option of the controller cache to "Write-through" and to enable the disk cache of the hard disk. As a result of these optimal cache settings, improvements in throughput of 34% and 29% are achieved depending on whether blocks of 8 KB or 64 KB are used for random access.

To achieve optimal throughput for random access with RAID 5 it is important to set the write mode option of the controller cache to "Write-back" and to enable the disk cache of the hard disk. Due to these optimal cache settings, improvements in throughput of 63% and 9% are achieved depending on block size.

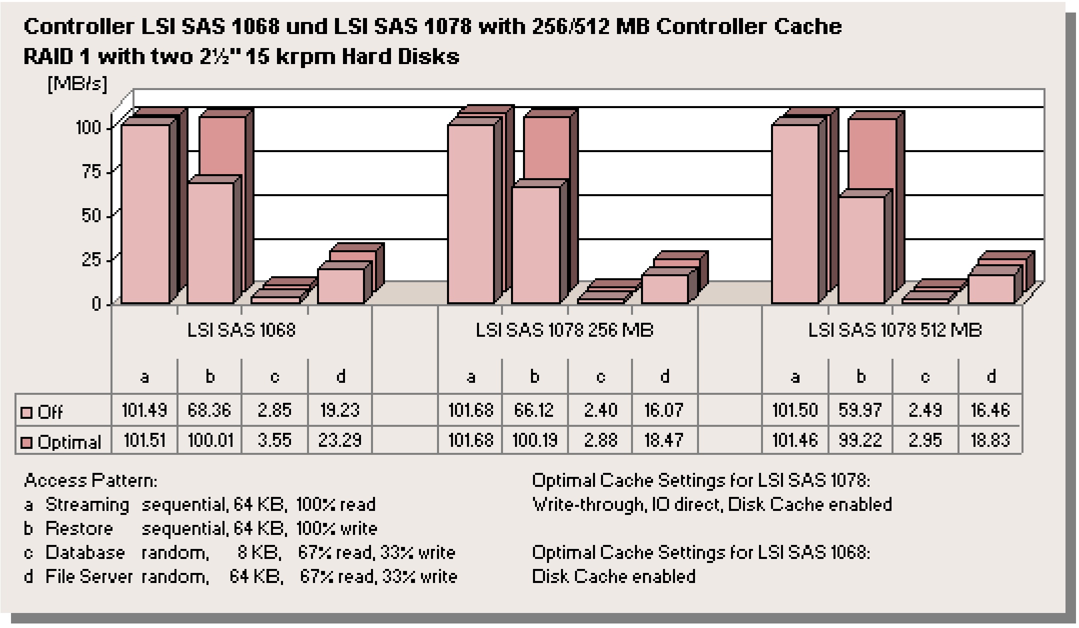
More detailed information about this topic is available in the paper [Performance Report - Modular RAID for PRIMERGY](#).



LSI MegaRAID SAS 1078 with 512 MB Cache

Controller comparison

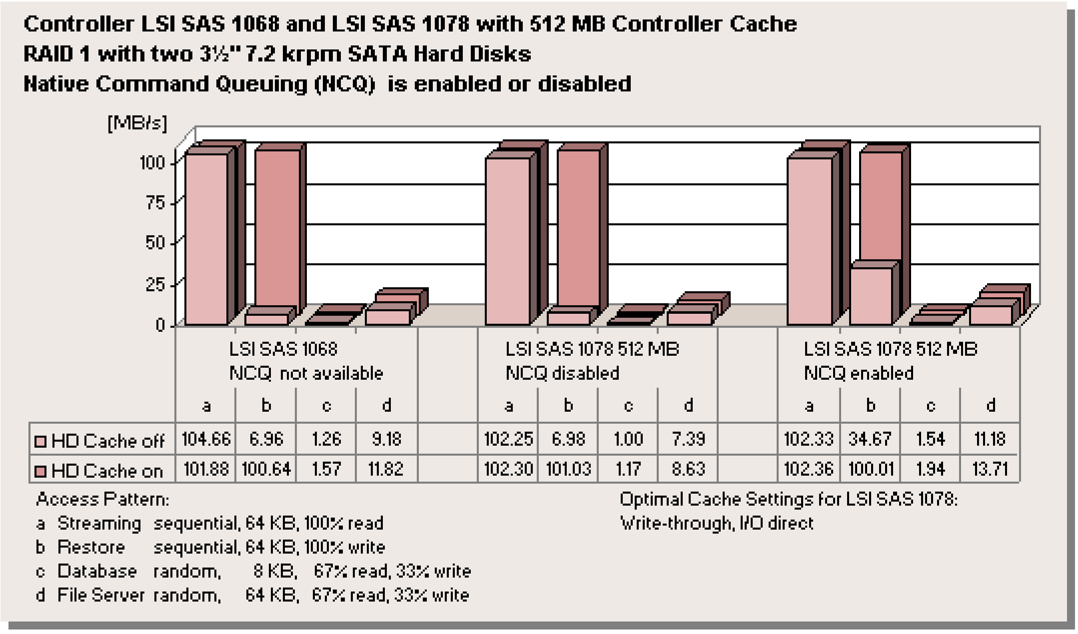
The following comparison depicts the throughputs of the various controllers. The measurements were made with the same hard disk types in the same RAID 1 array. The diagram shows the throughputs achieved with disabled caches (Off) and with optimal cache settings (Optimal).



The differences in performance of the controllers used are minimal with pure sequential access and optimal cache settings. With sequential read all the controllers achieve the maximum throughput rate regardless of the cache settings. Even for sequential write all the controllers are within the same performance range and data throughput can be increased by up to 65% through optimal cache settings.

With read access in RAID 1 the entry-level controller, LSI MegaRAID SAS 1068, shows a somewhat higher data throughput for this load profile than the LSI MegaRAID SAS 1078 controller, which is optimally equipped with its controller cache and the extended functionality for the higher RAID levels and also performs well in RAID 1.

When SATA hard disks are used, the LSI MegaRAID SAS 1078 controller has an advantage over the LSI MegaRAID SAS 1068 controller, because it supports Native Command Queueing (NCQ). NCQ collects the incidental I/O requests and optimizes the sequence in which they are processed. This avoids any unnecessary movements in the hard disk head, which is reflected in shorter latency times and higher throughput rates. This procedure can only be used with SATA hard disks. As can be seen in the diagram, NCQ results in a significant improvement in performance, particularly with sequential write without disk cache and for random access.



The LSI MegaRAID SAS 1078 controller with enabled NCQ, the default setting at present, therefore shows an approximately 5-fold throughput with sequential write without enabled disk cache. For random access, regardless of the disk cache, NCQ makes an approximately 22% higher throughput possible.

It is advisable to always leave NCQ enabled on the LSI MegaRAID SAS 1078 controller when running SATA hard disks so that the controller can also show its capability.

Conclusion

With the “Modular RAID” concept, the PRIMERGY TX300 S5 offers a plethora of opportunities to meet the various requirements of different application scenarios.

The entry-level controller, represented by the LSI MegaRAID SAS 1068 controller, offers the basic RAID solutions RAID 0, RAID 1 and RAID 1E and supports these RAID levels with a very good performance.

The “high-end” controller, represented by the LSI MegaRAID SAS 1078 controller, offers all today’s current RAID solutions; for the PRIMERGY TX300 S5, which can be expanded with up to twenty internal hard disks, this can be RAID levels 0, 1, 5, 6, 10, 50 and 60. This controller is supplied with a 256 MB or 512 MB controller cache and can as an optional extra be secured with a BBU. Various options for setting the use of the cache enable controller performance to be flexibly adapted to suit the RAID levels used.

Use of RAID 5 or RAID 6 enables the existing hard disk capacity to be utilized economically for a good performance. However, we recommend a RAID 10 for optimal performance and security.

The PRIMERGY TX300 S5 offers a choice between SATA and SAS, and for SAS hard disks between 2½" hard disks and 3½" hard disks and also different rotational speeds of 10 krpm or 15 krpm. Depending on the performance required, a decision must be taken as to which hard disk type with which rotational speed is to be used. Hard disks with 15 krpm offer an up to 23% better performance. As a result of using 2½" hard disks it is possible – depending on the RAID level – to achieve higher parallelism through the use of more hard disks in the RAID array.

For maximum performance it is advisable, particularly with SATA hard disks or when using a controller without a controller cache, to enable the hard disk cache. Depending on the disk type used, the increase in performance is 14-fold. When the hard disk cache is enabled we recommend the use of a UPS.

Benchmark environment*

All the measurements presented here were performed with the hardware and software components listed below.

Component	Details
Server	PRIMERGY TX300 S5
Operating system	Windows Server 2008, Enterprise Edition Version: 6.0.6001 Service Pack 1 Build 6001
File system	NTFS
Measuring tool	Iometer 27.07.2006
Measurement data	Measurement file of 8 GB
Controller LSI MegaRAID SAS 1068	Product: LSI RAID 0/1 SAS 1068 Driver Name: lsi_sas.sys, Driver Version: 1.29.03.00 Firmware version: 1.29.03.00 BIOS version: 06.26.00.00
Controller LSI MegaRAID SAS 1078	Product: LSI RAID 5/6 SAS 1078 Driver name: megasys.sys, Driver version: 3.9.0.64 Firmware package: 11.0.1-0008 Firmware version: 1.40.32-0580 BIOS version: 2.06.00 Controller cache: 256 MB or 512 MB
Hard Disk SATA, 3½", 7.2 krpm	Seagate ST3500320NS, 500 GB
Hard Disk SAS, 2½", 10 krpm	Seagate ST973402SS, 73 GB
Hard Disk SAS, 2½", 15 krpm	Seagate ST973451SS, 73 GB
Hard Disk SAS, 3½", 15 krpm	Seagate ST3146356SS, 146 GB

* Some components may not be available in all countries / sales regions.

SPECsfs2008*

Benchmark description

SPECsfs2008 is a benchmark suite developed by the Standard Performance Evaluation Corporation (SPEC) measuring file server throughput and response time. It provides a standardized method for comparing performance across different vendor platforms.

SPECsfs2008 results summarize the server's capabilities with respect to

- the number of operations that can be handled per second
- the response time (time required to complete individual operations)

In relation to the predecessor version the new code in SPECsfs2008 causes performance changes. Therefore SPECsfs2008 results are not comparable to results from earlier versions of the SPECsfs benchmark.

The SPECsfs2008 benchmark is used to measure the performance of NFS or CIFS fileservers. It creates a mixed workload that simulates a typical server environment. The NFS workload, following NFS protocol version 3, and the CIFS workload are based on data collected by SFS committee members from thousands of real NFS and CIFS servers operating at customer sites.

While there are some similarities, especially with respect to the file sets each workload operates on, the NFS and CIFS workloads are not comparable. For example, if the CIFS results for a SUT (System Under Test) are 20% higher than the NFS results for the same SUT, it should not be inferred that the SUT is 'better' at delivering CIFS operations than NFS operations. The workloads are very different and no attempt was made to normalize the NFS and CIFS workloads. The only valid comparisons that can be made are between published results for different SUTs operating against the same SPECsfs2008 workload, either NFS or CIFS.

The total file set size created for a given load point amounts to 120 MB per ops/sec. During the measurement 30 percent of the files in the file set are accessed. The maximum file size is 32 MB. In the NFS workload the relation between the commands READ and WRITE is 9:5. In the CIFS workload the relation between the commands READ_ANDX and WRITE_ANDX is about 7:3.

The benchmark requires that the user has one or more client systems that can be used to generate load on the server. One of these clients, referred to as the 'prime' client, is used to coordinate all the load-generating clients. Client operating systems supported are AIX, FreeBSD, Linux, Mac OS X, Solaris 10 and Windows.

In a typical SPECsfs2008 test configuration, a series of load generating clients are directed through a network at file systems shared or exported from the SUT. Clients are directed by the manager program to execute a series of ten or more tests at increasing, equally spaced 'load points'. A load point represents a throughput level (number of operations) for the workload the clients will present to the server.

At the start of each load point each client will start a specified number of load generating processes. Each process will mount the exported or shared file systems, create a directory structure, and fill it with a series of files of various sizes. The number of files the load generating process creates is determined by and increases proportionally with the number of operations per second specified for the given load point.

When all load generating processes have completed the initialization of the file set (known as the 'INIT' phase of the benchmark), they will begin to request from the server a series of NFS or CIFS operations based on the workload parameters established by the benchmark. The load generators will spend 300 seconds running in WARMUP mode, where no official measurements are recorded. After the WARMUP phase is completed, the benchmark will start a 300 second measurement phase, where the actual measurement of throughput and response time for the load point is collected.

Upon successful completion of the 300 second measurement period, statistics from all load generators on all clients are collected and centralized by the manager program, which will report and record the overall results of the test. If the test completes without violating any benchmark rules, the summary report produced by the manager may be used to produce an SFS submission which can be sent to SPEC for review and publication. For this purpose the benchmark reporting tools are used to present the peak achieved throughput as well as calculate an 'overall response time' metric, which is meant to reflect the results of average response time measurements across the series of load points.

The SPECsfs2008 Run and Reporting Rules have been established by SPEC to ensure that results generated with this suite are meaningful, comparable to other generated results, and are repeatable. Per the

* SPEC®, SPECsfs® und das SPEC-Logo sind eingetragene Warenzeichen der Standard Performance Evaluation Corporation (SPEC).

SPEC license agreement, all results publicly disclosed must adhere to these Run and Reporting Rules. Furthermore, SPEC requires that any public use of results from this benchmark follow the SPEC OSG Fair Use Policy. In the case where it appears that these guidelines have not been adhered to, SPEC may investigate and request that the published material be corrected.

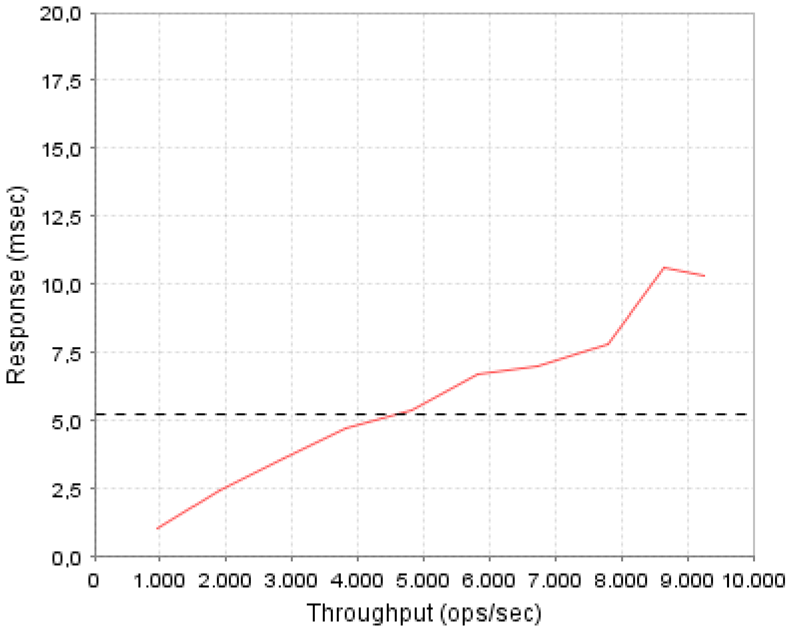
Benchmark results

In March 2009 the PRIMERGY TX300 S5 was measured with the CIFS workload of the SPECsfs2008 benchmark. The server was equipped with two Xeon X5570 processors, 12 x 8 GB PC3-8500R DDR3-SDRAM, two 8-port SAS RAID controllers with 512 MB cache and BBU and 20 SAS hard disks. The operating system was on two hard disks (36 GB, 15000 rpm), which were put together to form a RAID 1. The data area of the file server was to be found on 18 hard disks (146 GB, 10000 rpm), which were put together to form three RAID 50s. The PRIMERGY TX300 S5 was connected to the load generators via the onboard 2-port 1-Gigabit Ethernet controller. The measurement was made under Windows Server 2003 R2 Enterprise x64 Edition SP2.

A target value of 960 ops/sec was stipulated for the first measuring interval. In accordance with this stipulation the benchmark generated a data area with a size of about 112 GB. The target value was increased by a further 960 ops/sec for each of the following measuring intervals and thus also the size of the data area by about a further 112 GB. A result of SPECsfs2008_cifs = 9251 Ops/Sec (Overall Response Time = 5.22 msec) was achieved.

The table below shows the throughput achieved in the ten measuring intervals in operations/second as well as the average response times in milliseconds. This is reflected in the graph, which also shows the overall response time as an interrupted line, that is the average response time calculated for all measuring intervals as well as the maximum permitted average response time of 20 msec for an orderly benchmark run at the end of the y-axis. In the last measuring interval the target value of 9600 ops/sec is no longer achieved. The red curve ends at x = 9251. At this point in time the server is at full capacity.

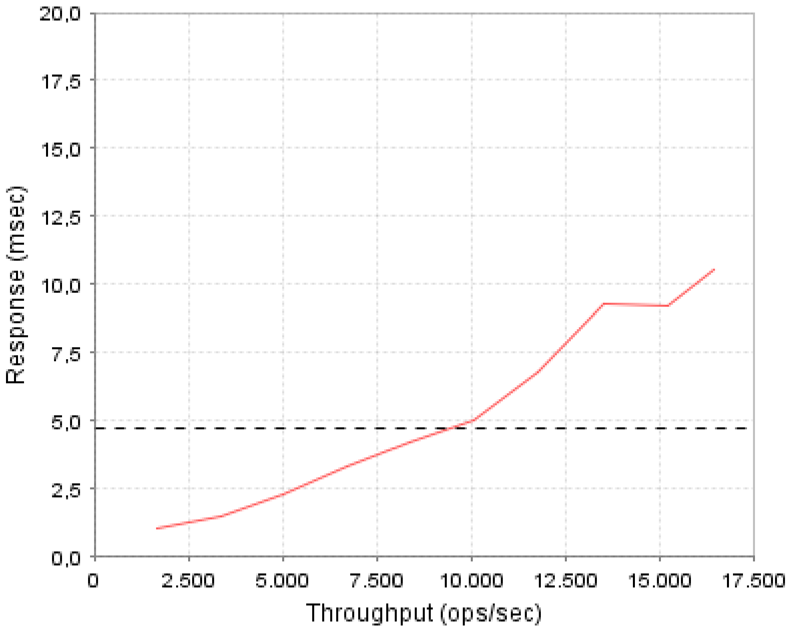
Throughput (ops/sec)	Response (msec)
958	1.0
1913	2.4
2875	3.6
3826	4.7
4808	5.4
5820	6.7
6747	7.0
7803	7.8
8657	10.6
9251	10.3



A further measurement was made in April 2009, in which the hard disks for the data area of the file server were put together to form three RAID 0s.

A target value of 1680 ops/sec was stipulated for the first measuring interval. In accordance with this stipulation the benchmark generated a data area with a size of about 192 GB. The target value was increased by a further 1680 ops/sec for each of the following measuring intervals and thus also the size of the data area by about a further 192 GB. A result of SPECsfs2008_cifs = 16488 Ops/Sec (Overall Response Time = 4.68 msec) was achieved.

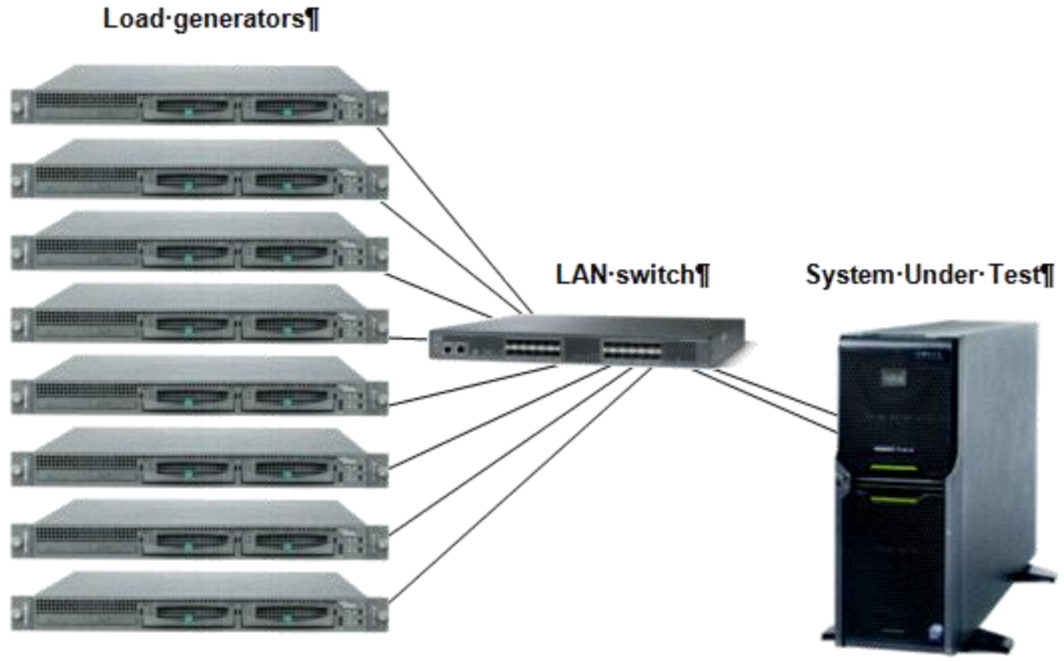
Throughput (ops/sec)	Response (msec)
1679	1.0
3357	1.5
5033	2.3
6712	3.3
8403	4.2
10064	5.0
11805	6.8
13512	9.3
15222	9.2
16488	10.5



Both measurement results have been published at http://www.spec.org.

The measurements show what throughput performance the PRIMERGY TX300 S5 achieves in the respective RAID configurations with its internal hard disk capacity. During the measurements the processors were by all means not at full capacity. The throughput was especially limited by the number of hard disks.

Benchmark environment*



System Under Test (SUT)	
Hardware	
Server	PRIMERGY TX300 S5
Processor	2 x Xeon X5570
Memory	12 x 8 GB PC3-8500R DDR3-SDRAM
Network Interface	1 x onboard 2-port 1-Gigabit Ethernet controller
Disk Subsystem	2 x 8-port SAS RAID controller with 512 MB cache and BBU 2 x SAS disk, 36 GB, 15000 rpm, RAID 1, for OS 18 x SAS disk, 146 GB, 10000 rpm, for data measurement 1: 3 x RAID 50 measurement 2: 3 x RAID 0
Software	
Operating System	Windows Server 2003 R2 Enterprise x64 Edition SP2

Load Generators	
Hardware	
Model	8 x PRIMERGY RX100 S3
Processor	2 x Pentium D 950
Memory	2 x 1 GB DDR2 SDRAM PC2-4200
Network Interface	2 x onboard 1-Gigabit Ethernet Broadcom BCM5721 (1 used)
Software	
Operating System	Windows Server 2003

* Some components may not be available in all countries / sales regions.

OLTP-2

Benchmark description

OLTP stands for Online Transaction Processing. The OLTP-2 benchmark is based on the typical application scenario of a database solution. In OLTP-2 database access is simulated and the number of transactions achieved per second (tps) determined as the unit of measurement for the performance of the system measured.

In contrast to benchmarks such as SPECint and TPC-E, which were standardized by independent bodies and for which adherence to the respective rules and regulations are monitored, OLTP-2 is an internal benchmark of Fujitsu Technology Solutions. The partially enormous hardware and time expenditure for standardized benchmarks has been reduced to a reasonable degree in OLTP-2 so that a variety of configurations can be measured within an acceptable period of time.

Even if the two benchmarks OLTP-2 and TPC-E simulate similar application scenarios using the same workload, the results cannot be compared or even treated as equal, as the two benchmarks use different methods to simulate user load. OLTP-2 values are typically similar to TPC-E values. A direct comparison, or even referring to the OLTP-2 result as TPC-E, is not permitted, especially because there is no price-performance calculation.

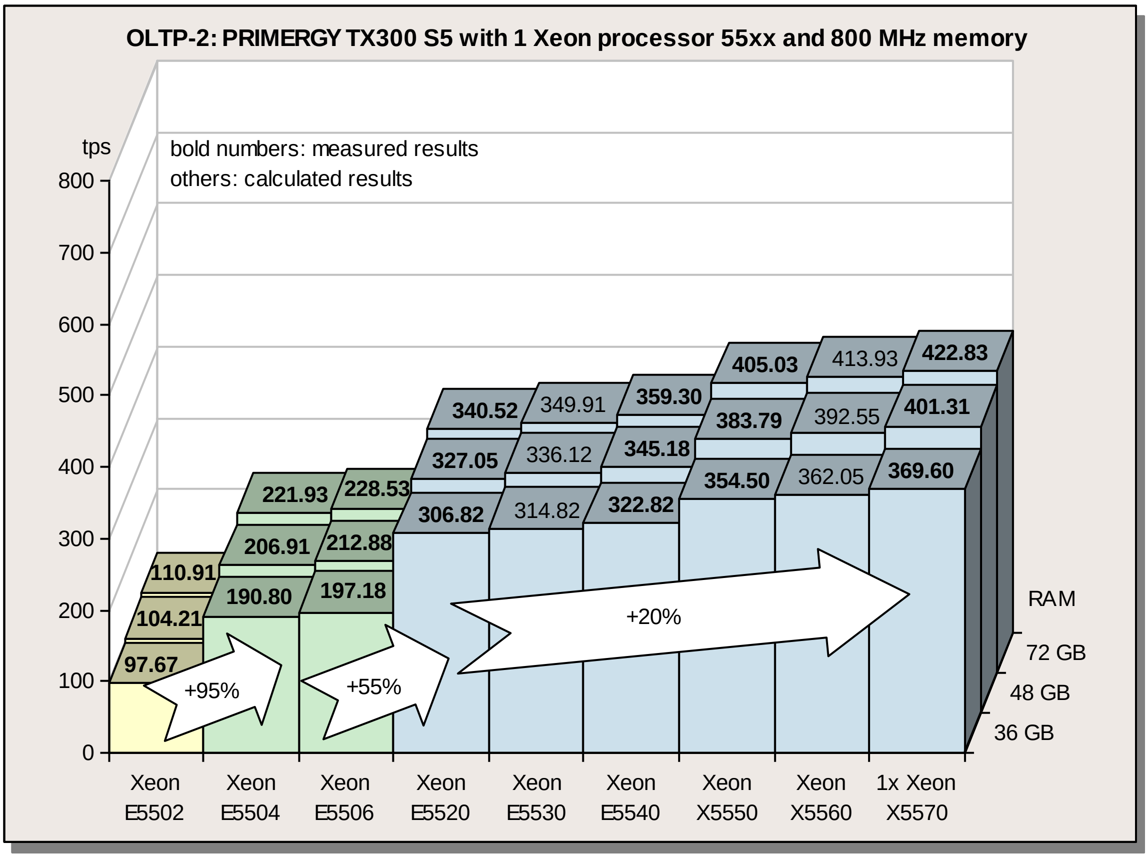
Benchmark results

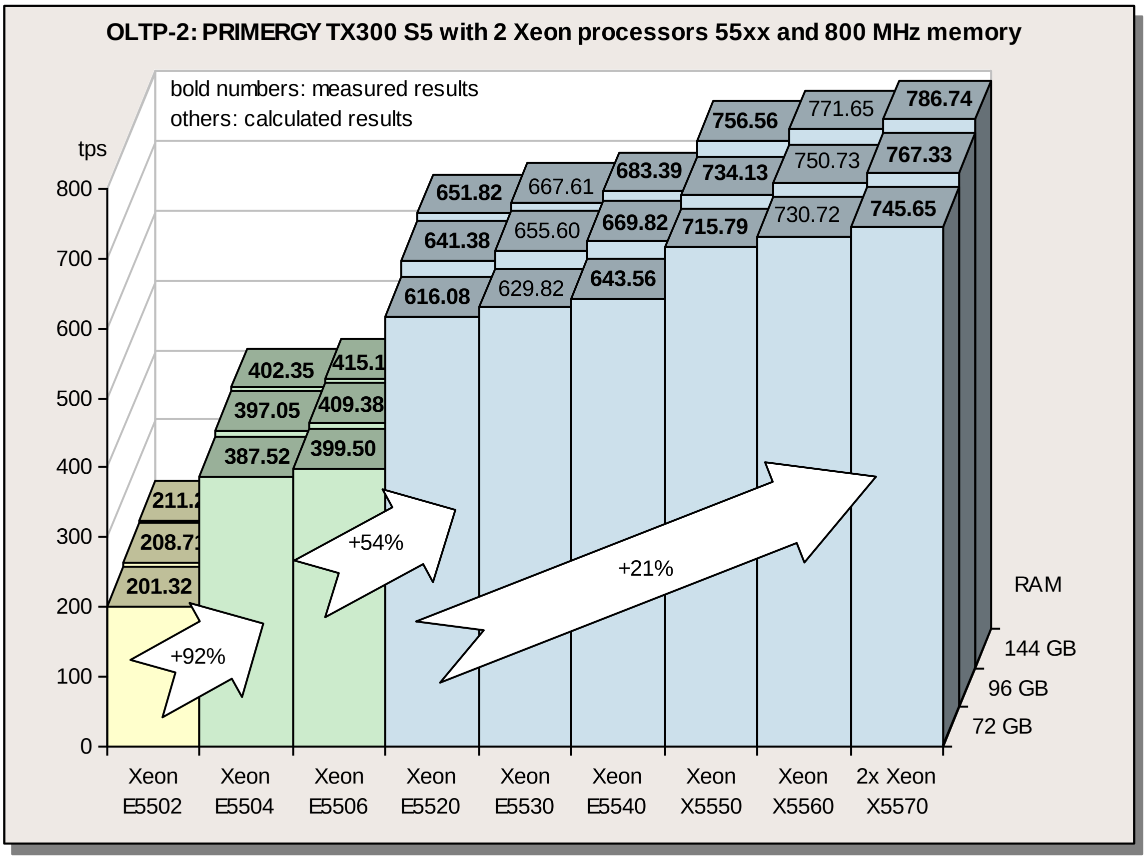
The PRIMERGY TX300 S5 has been measured with Intel Xeon Processors series 55xx at a memory size of 36 GB, 48 GB, 72 GB 96 and 144 GB. All results were determined on the basis of the operating system Microsoft Windows Server 2008 Enterprise x64 Edition and the database SQL Server 2008 Enterprise x64 Edition. OLTP-2 benchmark results depend to a great degree on the configuration options of a system with hard disks and their controllers. Therefore, the system was equipped with five LSI SAS MegaRAID 8880EM2 controllers that were connected to 30 PRIMERGY SX40 and a total of 360 SAS-hard disks. The disk subsystem was dimensioned to be no bottleneck within the measurements. Comparable results may also be achievable with other disk subsystems being no bottleneck. See the Benchmark environment section for further information on the system configuration.

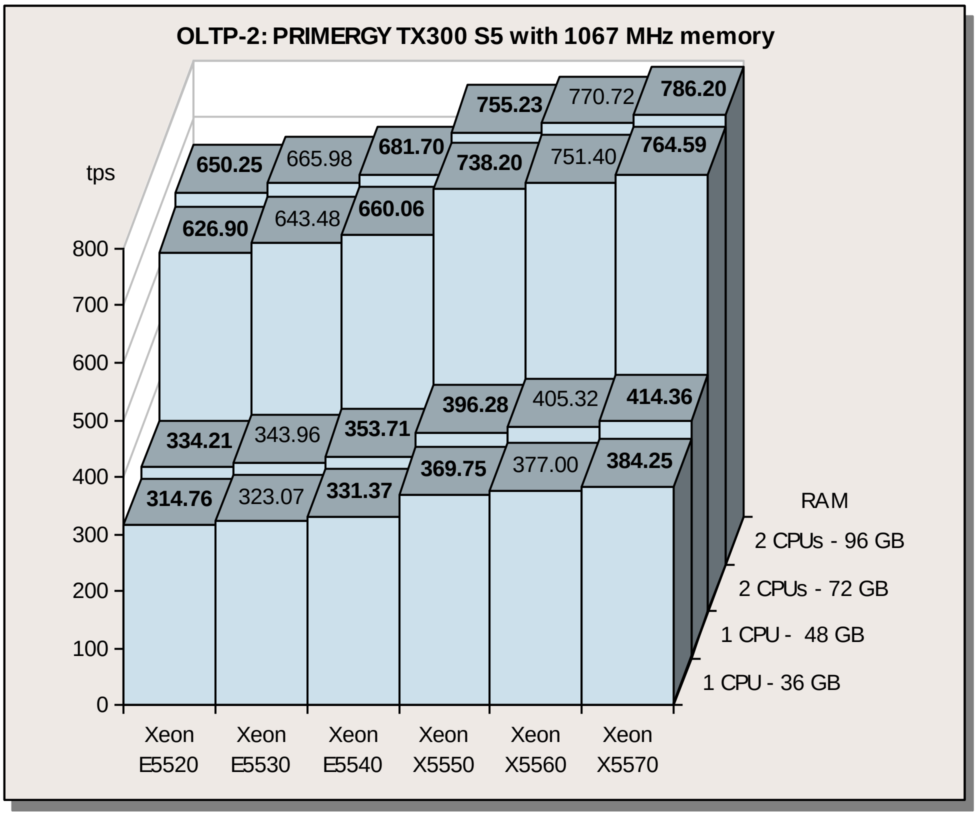
In the maximum memory configuration of PRIMERGY TX300 S5 with 9 memory modules at one processor and 18 memory modules at two processors memory access speed is 800 MHz. If only 6 memory modules are used per processor, memory access speed is 1067 MHz with Intel Xeon E5520, E5530, E5540, X5550, X5560 and X5570. There are measurements for all types of processors using 800 MHz memory access. In addition the six largest processor types have been measured with 1067 MHz memory access using less memory than possible.

The next diagrams show the OLTP-2 performance data for the PRIMERGY TX300 S5 separated in two groups with one and two processors Intel Xeon series 55xx (E5502, E5504, E5506, E5520, E5530, E5540, X5550, X5560 and X5570) and a memory access speed of 800 MHz. Performance values for Xeon L5530 are equivalent to Xeon E5530, values for Xeon L5520 equivalent to Xeon E5520 and values for Xeon L5506 equivalent to Xeon E5506.

The largest scaling over all processor types with +92% to +95% is at E5502 to E5504. In this case the number of processor cores is doubled from two to four. There is also a large increase of +54% to +55% at E5506 to E5520 by doubling the processor cache from 4 MB to 8 MB and the use of Hyper-Threading. Finally the scaling across E5520 to X5570 is +20 to +21%. The memory scaling from 36 GB to 48 GB is about +8%, from 72 GB to 96 GB about +3%. This depends on the workload of the OLTP-2 benchmark and is not typical for all database applications.


OLTP-2: PRIMERGY TX300 S5 with 1 Xeon processor 55xx and 800 MHz memory
tps
bold numbers: measured results
others: calculated results
800
700
600
500
400
300
200
100
0
97.67
104.21
110.91
190.80
206.91
221.93
197.18
212.88
228.53
306.82
327.05
340.52
314.82
336.12
349.91
322.82
345.18
359.30
354.50
383.79
405.03
362.05
392.55
413.93
369.60
401.31
422.83
+95%
+55%
+20%
RAM
72 GB
48 GB
36 GB
Xeon E5502
Xeon E5504
Xeon E5506
Xeon E5520
Xeon E5530
Xeon E5540
Xeon X5550
Xeon X5560
1x Xeon X5570


OLTP-2: PRIMERGY TX300 S5 with 2 Xeon processors 55xx and 800 MHz memory
bold numbers: measured results
others: calculated results
tps
800
700
600
500
400
300
200
100
0
201.32
387.52
397.05
402.35
399.50
409.38
616.08
641.38
651.82
629.82
655.60
667.61
643.56
669.82
683.39
715.79
734.13
756.56
730.72
750.73
771.65
745.65
767.33
786.74
+92%
+54%
+21%
RAM
144 GB
96 GB
72 GB
Xeon E5502
Xeon E5504
Xeon E5506
Xeon E5520
Xeon E5530
Xeon E5540
Xeon X5550
Xeon X5560
2x Xeon X5570



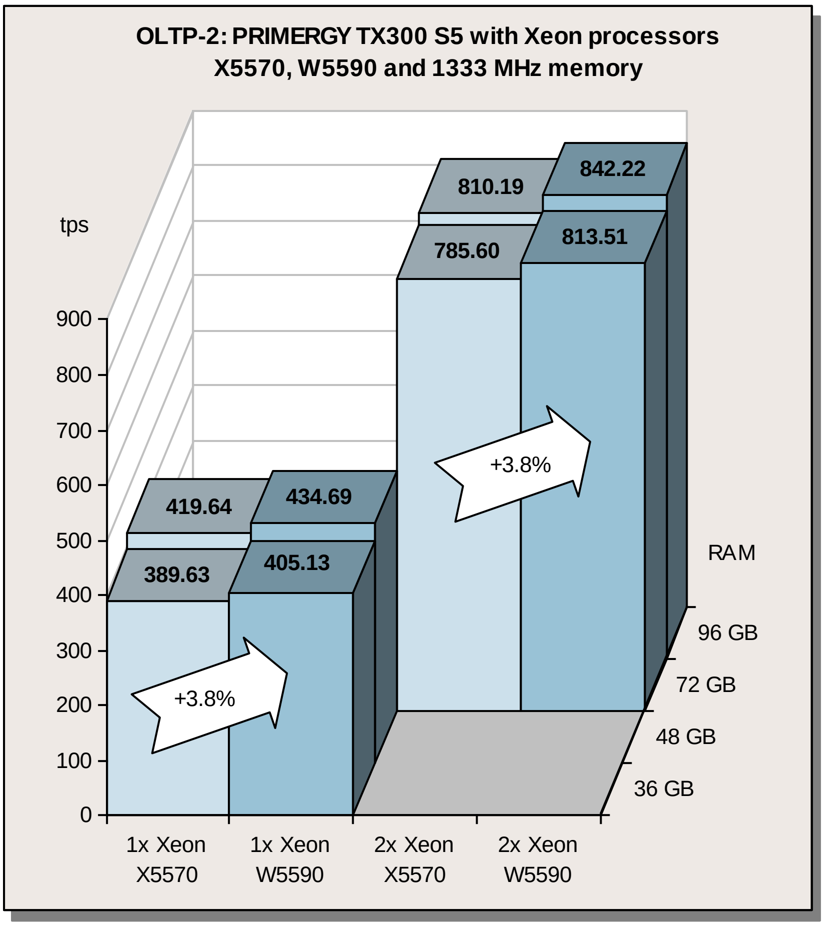
The following graph shows the OLTP-2 performance data for 1067 MHz memory access speed and processor types Xeon E5520, E5530, E5540, X5550, X5560 und X5570. Memory size at one processor is 36 GB and 48 GB. Memory size at two processors is 72 GB and 96 GB. Smaller memory latencies with higher memory frequency result in overall higher values.

The later release of the processor type W5590 required another measurement, because SQL Server 2008 Enterprise x64 Edition SP1 increased in performance and results are not comparable to the first measurement. In addition memory access frequency increases from 1067 MHz to 1333 MHz at six memory modules per processor. This measurement compares the processors Xeon X5570 and X5590. See the Benchmark environment section for further information on the new system configuration.

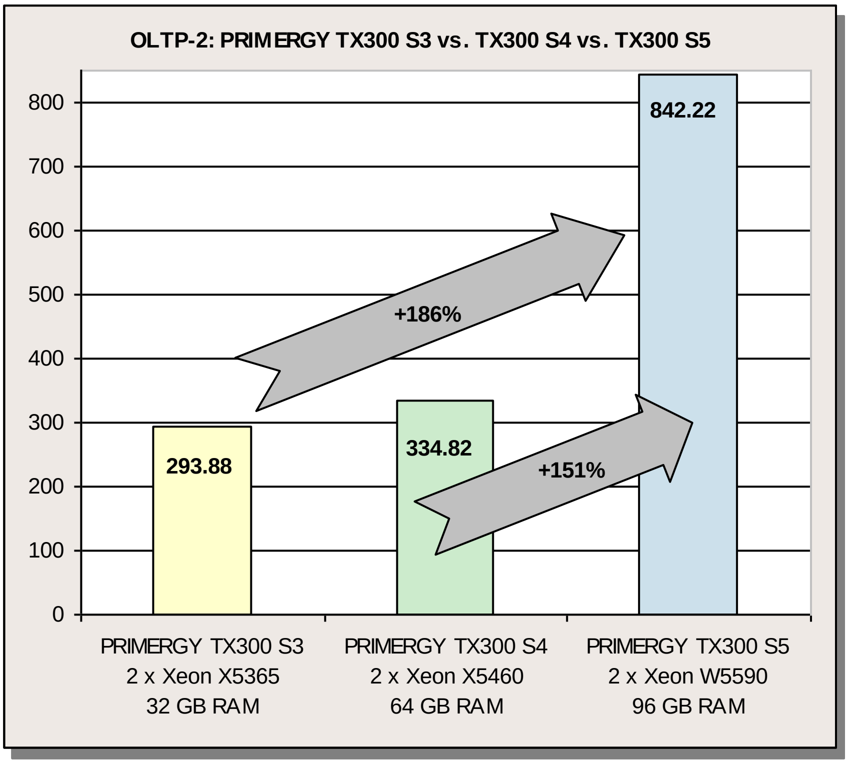
Beside graph show the OLTP-2 performance values of Xeon X5570 and W5590 with SQL Server 2008 Enterprise x64 Edition SP1. Six memory modules are used per processor and memory access is 1333 MHz. Maximum memory is 48 GB and 96 GB.

The performance increase of processor W5590 compared to Xeon X5570 is 3.8%.

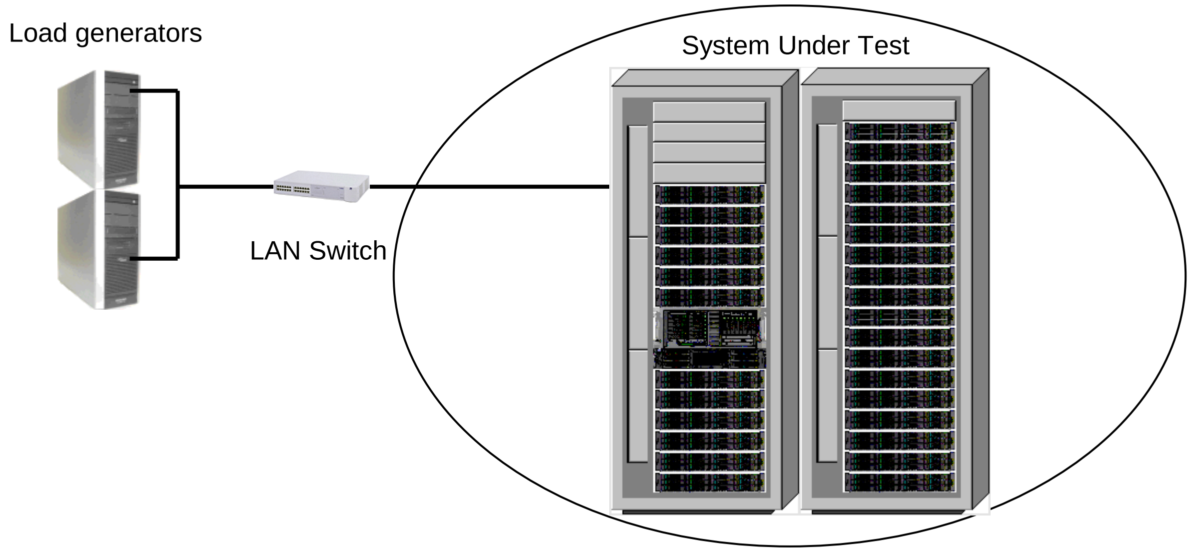
Software and hardware improvements for the second measurement increase about +2.1% compared to the first one.



When comparing the PRIMERGY TX300 S5 and its predecessors the PRIMERGY TX300 S3 and TX300 S4 all in their highest performance configurations, a throughput increase of +186% and +151% is noted.



Benchmark environment*



System Under Test (Tier B)	
Hardware	
Server	PRIMERGY TX300 S5
Processor	Intel Xeon E5502, E5504, E5506, E5520, E5530, E5540, X5550, X5560, X5570 Second test: X5570, W5590
Memory	Up to 18 x 8 GB DDR3 PC3-8500R
Settings (default)	Turbo Mode enabled, NUMA Support enabled, Hyper-Threading enabled
Network Interface	2 x 1-GBit LAN Intel (onboard)
Disk Subsystem	PRIMERGY TX300 S5: 1 x LSI SAS PCI Express™ ROMB 2x 3.5" 36GB 15K Fujitsu MAY2073RC RAID-1, OS 6x 3.5" 146GB 15K Fujitsu MAY2073RC RAID-0, log 5x LSI SAS MegaRAID 8880EM2 30 x FibreCAT SX40: 192 x Seagate 73 GB 15 krpm, RAID-0, data 168 x Seagate 146 GB 15 krpm, RAID-0, data
Software	
Operating System	Windows Server 2008 Enterprise x64 Edition
Database	First test: SQL Server 2008 Enterprise x64 Edition Second test: SQL Server 2008 Enterprise x64 Edition SP1

Front End (Tier A)	
Hardware	
Model	1 x PRIMERGY RX300 S4
Processor	2 x Xeon E5420 2.50 GHz, 2 MB L2 cache
Memory	4 GB FBD667 PC2-5300F
Network Interface	2 x 1-GBit LAN (onboard), dual port LAN 1-GBit
Software	
Operating System	Windows Server 2003 R2 Standard x64 Edition

Load Generators

* Some components may not be available in all countries / sales regions.

Hardware	
Model	2 x PRIMERGY Econel 200
Processor	2 x Xeon 3.40 GHz, 2 MB L2 cache
Memory	2 GB DDR-SDRAM PC2700
Network Interface	1 x 1-GBit LAN (onboard)
Software	
Operating System	Windows Server 2003 Standard Edition SP1 (x86)
OLTP-2 Software	First test: EGen version 1.6.0-1011 Second test: EGen version 1.8.0-1015

SAP SD

Benchmark description

The SAP application software consists of modules to manage standard business processes. There are modules for ERP (Enterprise Resource Planning) like Assemble-to-Order (ATO), Financial Accounting (FI), Human Resources (HR), Materials Management (MM), Production Planning (PP) und Sales and Distribution (SD), as well as for SCM (Supply Chain Management), Retail, Banking, Utilities, BI (Business Intelligence), CRM (Customer Relation Management) or PLM (Product Lifecycle Management).

This application software always runs on top of a database system, so a complete SAP configuration consists (in addition to the supporting hardware) of the software components operating system, database and the SAP software itself.

To assess the performance, stability and scalability of a SAP application system, SAP AG has developed the SAP Standard Application Benchmarks. These benchmarks (the SD benchmark being the most important) analyze the performance of the entire system and thus deliver a measure for the quality of the integration of the single components.

The benchmark differentiates between a two-tier and a three-tier configuration. With the two-tier configuration, the SAP application and the database are installed on one server. With a three-tier configuration, the individual components of the SAP application can be distributed over several servers and another server takes over the database.

A complete specification of the benchmark developed by SAP AG, Walldorf, Germany is available at http://www.sap.com/benchmark.

Benchmark results

With the certification number 2008072, SAP certifies that the PRIMERGY TX300 S5, equipped with 2 Xeon X5570 processors (with SAP ECC release 6.0 and SQL Server 2005 x64), attained the following results on December 5, 2008 under Windows Server 2003 Enterprise x64 Edition SP2:

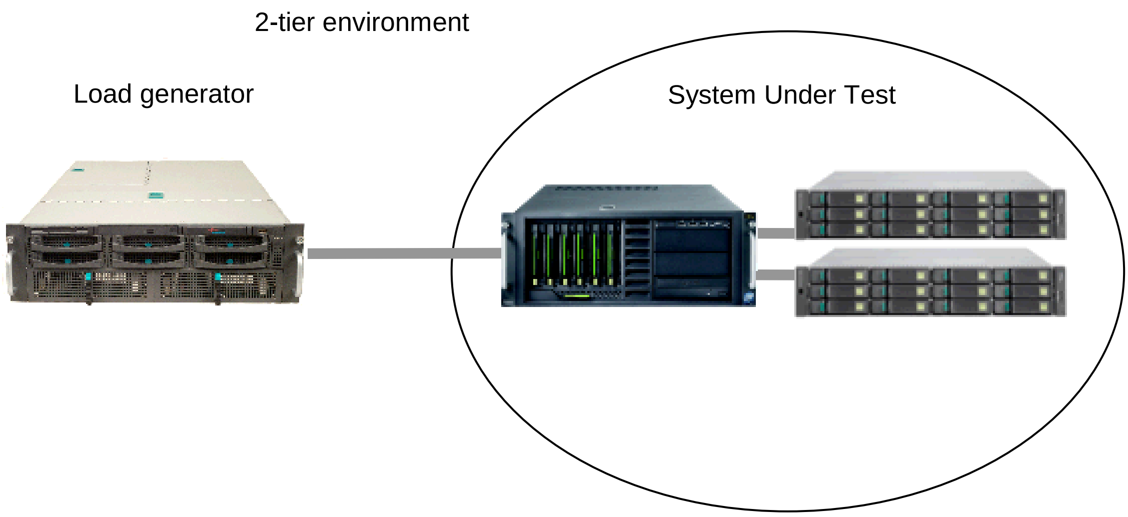
Number of benchmark users	4715 SD (Sales & Distribution)
Average dialog response time	1.96 seconds
Throughput	
Fully Processed Order Line items / hour	473000
Dialog steps / hour	1419000
SAPS	23650
Average DB request time (dia/upd)	0.011 sec / 0.040 sec
CPU utilization central server	99%
Operating System central server	Windows Server 2003 Enterprise x64 Edition
RDBMS	SQL Server 2005 x64
SAP ECC Release	6.0
Configuration Central Server	PRIMERGY TX300 S5 2 × Xeon X5570, 2.93 GHz, ½ MB L2 cache per core, 8 MB L3 cache per chip, 48 GB RAM

With the certification number 2009014, SAP certifies that the PRIMERGY TX300 S5, equipped with 2 Xeon X5570 processors (with SAP enhancement package 4 for SAP ERP 6.0 (Unicode) and SQL Server 2008 x64), attained the following results on April 27, 2009 under Windows Server 2008 Enterprise x64 Edition:

Number of benchmark users	3328 SD (Sales & Distribution)
Average dialog response time	0.99 seconds
Throughput	
Fully Processed Order Line items / hour	363330
Dialog steps / hour	1090000
SAPS	18170
Average DB request time (dia/upd)	0.025 sec / 0.014 sec
CPU utilization central server	99%
Operating System central server	Windows Server 2008 Enterprise x64 Edition
RDBMS	SQL Server 2008 x64
SAP Business Suite software	SAP enhancement package 4 for SAP ERP 6.0 (Unicode)
Configuration Central Server	PRIMERGY TX300 S5 2 × Xeon X5570, 2.93 GHz, ½ MB L2 cache per core, 8 MB L3 cache per chip, 48 GB RAM

Benchmark environment*

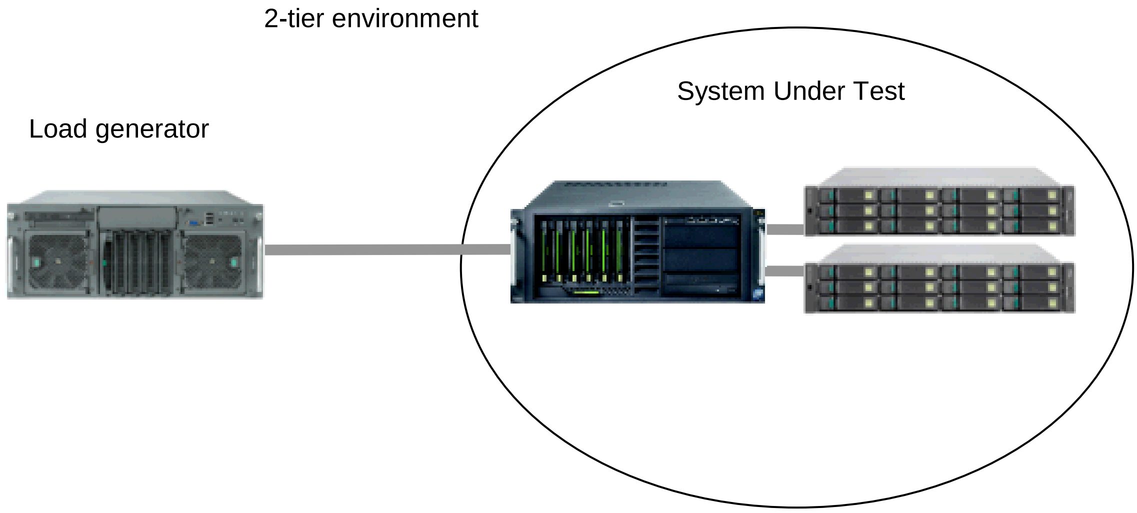
Certification number 2008072



System Under Test (SUT)	
Hardware	
Server	PRIMERGY TX300 S5
Processor	2 x Xeon X5570
Memory	12 x 4 GB PC3-8500R DDR3-SDRAM
Disk Subsystem	PRIMERGY TX300 S5: 1 x LSI MegaRAID SAS 1068 controller 3 x 2.5" SAS disks, 73 GB, 15 krpm 1 x LSI MegaRAID SAS 1078 controller with 512 MB cache 4 x 2.5" SAS disks, 73 GB, 15 krpm 1 x LSI MegaRAID SAS 8880E controller with 512 MB cache and BBU 2 x FibreCAT SX40: 14 x 2.5" SAS disks, 73 GB, 15 krpm
Software	
Operating System	Windows Server 2003 Enterprise x64 Edition SP2
Database	SQL Server 2005 x64
SAP ECC Release	6.0

Load Generator	
Hardware	
Model	PRIMERGY RX600
Processor	4 Xeon MP 2.50 GHz, 512 KB L2 cache, 1 MB L3 cache
Memory	8 GB PC2100 DDR-SDRAM
Software	
Operating System	Linux 2.6

Certification number 2009014



System Under Test (SUT)	
Hardware	
Server	PRIMERGY TX300 S5
Processor	2 x Xeon X5570
Memory	12 x 4 GB PC3-8500R DDR3-SDRAM
Disk Subsystem	PRIMERGY TX300 S5: 1 x LSI MegaRAID SAS 1068 controller 3 x 2.5" SAS disks, 73 GB, 15 krpm 1 x LSI MegaRAID SAS 1078 controller with 512 MB cache 4 x 2.5" SAS disks, 73 GB, 15 krpm 1 x LSI MegaRAID SAS 8880E controller with 512 MB cache and BBU 2 x FibreCAT SX40: 14 x 2.5" SAS disks, 73 GB, 15 krpm
Software	
Operating System	Windows Server 2008 Enterprise x64 Edition
Database	SQL Server 2008 x64
SAP Business Suite software	SAP enhancement package 4 for SAP ERP 6.0 (Unicode)

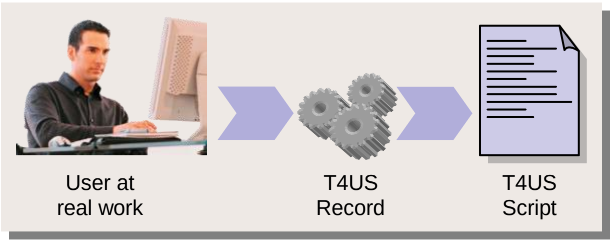
Load Generator	
Hardware	
Model	PRIMERGY RX600 S2
Processor	4 x Xeon 7040, 3 GHz, 4 MB L2 cache
Memory	8 GB PC2-3200 DDR2-SDRAM
Software	
Operating System	SUSE Linux Enterprise Server 11

* Some components may not be available in all countries / sales regions.

Terminal Server

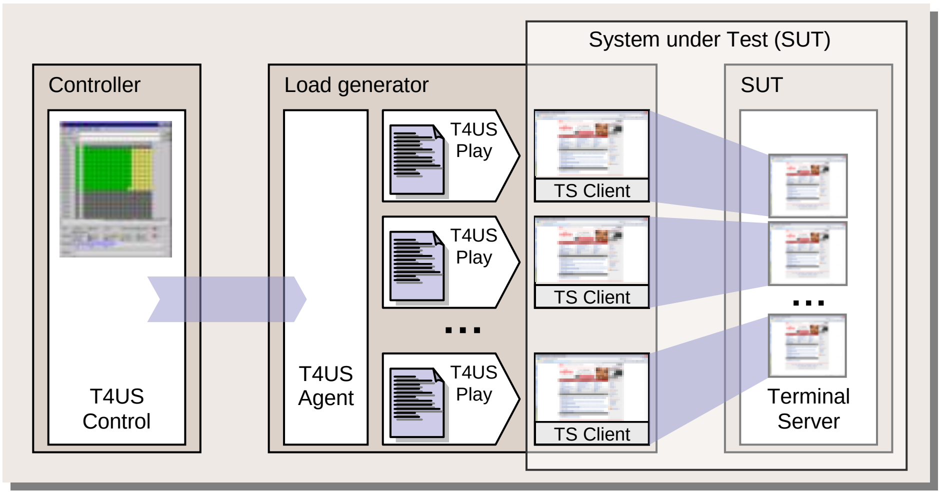
Benchmark description

For Terminal Server measurements there are a number of load simulation tools, whose results cannot be compared with each other and which are not a standard benchmark. The existing load simulators are not in a position to measure Microsoft Terminal Services and Citrix Presentation Server under the same conditions or have other limitations. Fujitsu Technology Solutions therefore uses a self-developed program named **T4US** (Tool for User Simulation). This is a flexible tool that can simulate any terminal-server-based scenario – independent of the operating system or application software used – and that carries out an in-depth measuring of response times and utilization of all the different system components.



The **T4US Record** tool records user input as keyboard and mouse activities in real time as well as display outputs and stores it in a **T4US Script**. T4US Scripts are the load profiles used during the measurement.

The T4US load simulator has three components.



T4US Control centrally controls and monitors the entire simulation process and evaluates measurement data during the measurement. Several instances of **T4US Playback** run on the load generator. Each T4US Playback "feeds" keyboard and mouse inputs in real time to a terminal server client on the basis of T4US Scripts recorded with T4US Record, and monitors the display content of the terminal server client. Thus, the response time of the terminal server is determined by means of high-resolution timers. A **T4US Agent** runs on every load generator. The T4US Agent is responsible for handling communication with the controller, controls and monitors the instances of T4US Playback and transfers the measured response times to the controller.

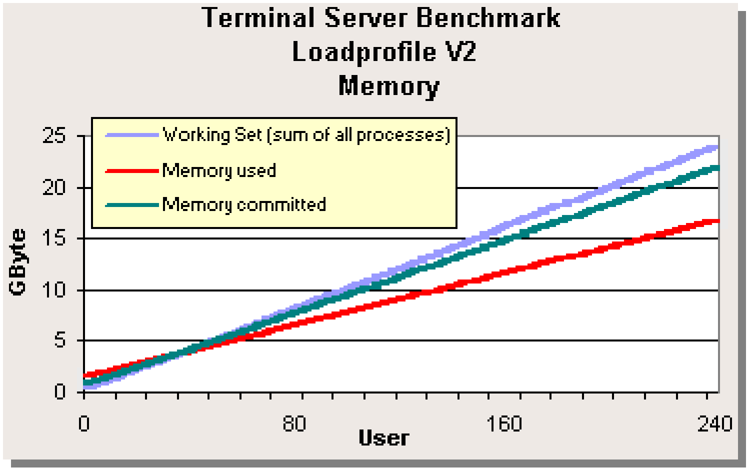
During the measurement the number of users working with Terminal Server is continuously increased. The Terminal Server response times are monitored by the T4US controller and compared with stored reference values which were determined from a previous reference measurement with only few users. If the response time of the application has deteriorated to such a degree that it no longer complies with the predefined rules, the measurement is terminated and the number of users is the result of this measurement. However, this number should not be seen as absolute, because the number of users who can support a system always depends on the actual user profile. Instead the results should primarily be regarded as relative, that is, "a PRIMERGY System A is twice as efficient as a PRIMERGY System B" or "the doubling of the main memory results in a x% increase in performance."

Load profile V2

The load profile V1, which has until now been used in Terminal Server measurements, and with which each user logs on to Terminal Server on a cyclical basis, creates a text with images and then logs off again, can no longer be used. Due to improved performance in the systems to be measured the benchmark reaches a scale at which the number of users is achieved due to the logon/logoff processes to be performed and no longer through the processor performance of the system, in other words there are restrictions in the operating system. This means that the benchmark already reaches its limits without using the processor to capacity. An improvement in processor performance cannot be measured by the benchmark. This is why a new load profile V2 is used with the measurements performed here.

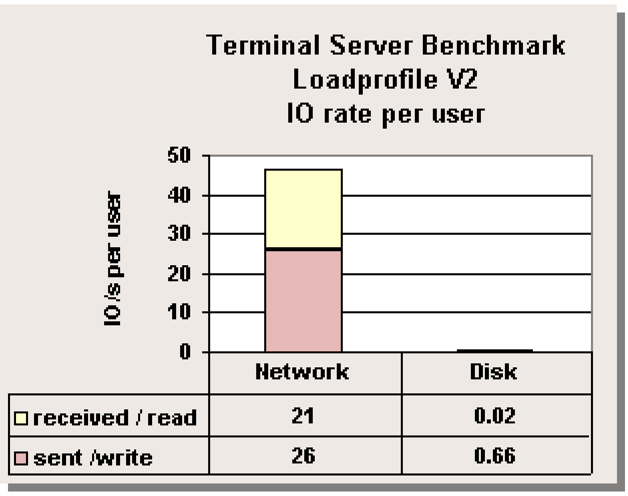
The new load profile V2 is characterized by the fact that a user that is to be simulated works with various Microsoft Office applications. In addition to the creation of a Microsoft Word document, a PowerPoint presentation is also designed. Calculations are also made on a new Excel spreadsheet. The number of logon/logoff processes is reduced in comparison with the old load profile. On average only every sixth user logs on to and off from Terminal Server on a cyclical basis. Also every sixth user prints on average a Word document. Additional CPU load is achieved by packing and unpacking files in the memory. The typing speed of the simulated user is between 330 and 440 characters per minute.

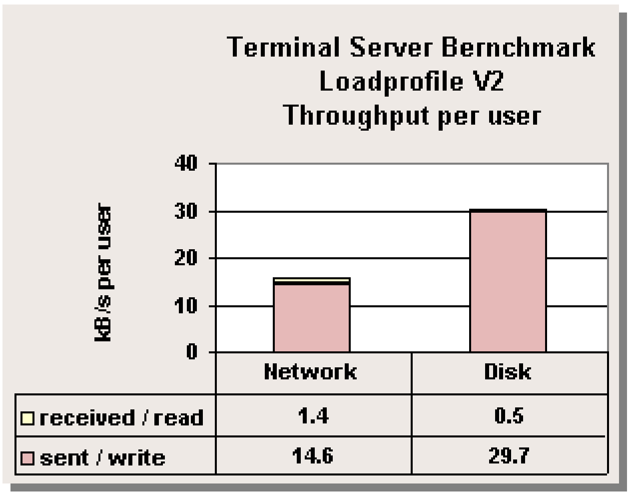
The memory requirements of the Terminal Server benchmark increase in proportion with the number of users and depend on the underlying operating system, there are in particular differences in the 32-bit and 64-bit operating systems. This aspect is handled in depth in the document Terminal Server Sizing Guide - 64-bit Technology (see Literature).



In the diagram opposite the memory requirements of the benchmark with load profile V2 are presented on a 64-bit Windows Server 2008 system. Due to the fact that the users now work with various applications, the memory used for load profile V2 is higher than with the original load profile V1.

The other diagrams show the average IO rates of the disks and the network as well as the relevant data throughputs that this load profile V2 generates on a Windows Server 2008 x64 system.





Benchmark results

In all the measurements performed the Terminal Server system was equipped with the operating system "Windows Server 2008 x64 Enterprise Edition without Hyper-V SP1". Measurements with a 32-bit operating system were omitted because of the restriction of virtual address space and kernel structures and, as a consequence thereof, the limitation of supported users.

All installations for which no optimizations were performed on the server or client are standard. The only settings that are changed to subject all PRIMERGYs to the same test conditions are the following ones:

- The page file of the operating system was set to a fixed size of 28 GB.

The following performance-relevant factors are critical for a terminal server system:

- Computing performance
- Main memory
- Disk subsystem
- Network

Network

A Terminal Server-based infrastructure is substantially influenced by the underlying network infrastructure. Because we are discussing the performance of an individual Terminal Server in this case, the network has been dimensioned in such a way that it does not represent a bottleneck.

Disk subsystem

The disk subsystem is a further performance-relevant component. In the measurement environment used here the operating system is saved on one partition of a RAID 0 array of two hard disks, the data of the users and the pagefile are saved on partitions of a RAID 0 array of a further two hard disks of the Terminal Server. This configuration is used to ensure that the measurement results between the various PRIMERGY systems are comparable and that the disk subsystem does not become a bottleneck during measuring. However, this does not mandatory correspond to the real customer configuration, because there the user data is typically placed on appropriate disk subsystems or external file servers and not on local hard disks of a terminal server. To achieve maximum throughput, all caches, including the write caches, have been activated. Hard-disk write caches make a considerable contribution toward increasing performance and it is recommended - also in productive use - to make use of this functionality, which is available on all hard disks. In this regard, it is advisable to use a UPS to protect against power failures and the data loss that these entail.

Main memory

The main memory has the greatest influence on the performance of the terminal server. This is particularly reflected in the response time. As and when required, Windows acquires further virtual memory by relocating (swapping) data currently not needed from the main memory (RAM) to the swap file on the hard disk. However, since disk accesses are about a thousand times slower than memory accesses, this results directly in a breakdown in performance and a rapid increase in response times.

Since a Terminal Server works with numerous users and various applications, it is primarily important to have equipped the system with adequate memory. Memory access speed is then a subordinate factor. With a maximum memory configuration of up to 144 GB the PRIMERGY TX300 S5 provides a good platform for Terminal Server.

The memory access speed of the PRIMERGY TX300 S5 not only depends on the processor but also on memory configuration. The best access speed is achieved when the memory DIMMs are only inserted on one bank, distributed across the channels allocated to the CPU.

In the measurements performed here the measured Terminal Server systems were equipped with sufficient memory. With 6 × 4 GB memory, distributed over three channels per CPU, the PRIMERGY TX300 S5 was optimally configured for the number of simulated users and also optimally configured for smaller memory access times. Doubling the memory to 48 GB did not bring about any improvement in the benchmark results.

Computing performance

Depending on requirements, the PRIMERGY TX300 S5 can be equipped with various processors, which differ in respect to clock frequency, cache, transfer rate of the Quick Path Interconnect (giga transfer, GT) and number of cores.

The system was measured for the Terminal Server benchmark with both the smallest quad-core processor, the Xeon E5504, and the currently most powerful quad-core processor for this system, the Xeon X5570. In contrast to the Xeon E5504, the Xeon X5570 has both Hyper-Threading and Turbo Boost Technology, which - depending on the application – automatically overclocks the processor when working below maximum energy consumption (Thermal Design Power (TDP)).

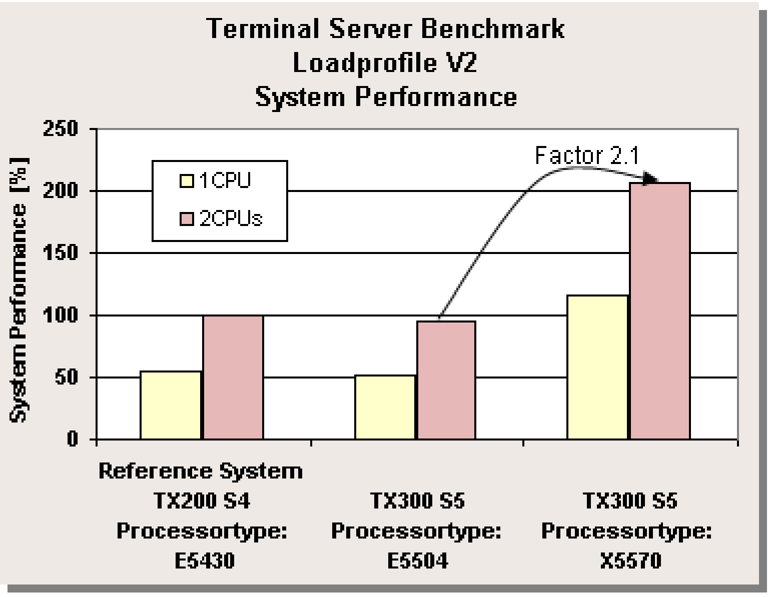
Other features of the processors measured

- Xeon E5504, 2.00 GHz, 4.8 GT, max. 800 MHz DDR3 bus speed, 4 MB L3 cache, 80 watt
- Xeon X5570, 2.93 GHz, 6.4 GT, max. 1333 MHz DDR3 bus speed, 8 MB L3 cache, 95 watt

The maximum user numbers achieved per system with the new load profile V2 cannot be compared with the user numbers that were achieved through the former load profile V1. To avoid confusion the benchmark results are therefore no longer presented in absolute user numbers, but only in relation to a previously measured reference system, here a PRIMERGY TX200 S4 equipped with up to two Xeon E5430 processors that have neither Hyper-Threading nor Turbo Boost Technology.

- Xeon E5430, 2.67 GHz, 1333 MHz front-side bus, 2 × 6 MB L2 cache, 80 watt

The Terminal Server benchmark with the new load profile V2 scales well. As a result of doubling the number of processors, that is to say from four to eight cores, a 1.8-fold increase in system performance is achieved both in the reference system and in the PRIMERGY TX300 S5 with the Xeon E5504.



When Hyper-Threading is enabled, the Xeon X5570 has eight logical CPU cores. In other words, the adding of a second CPU means increasing the number of logical cores from eight to 16. Even in these measurements the benchmark still achieves good scaling of system performance - namely of more than 1.7-fold.

As regards system performance, the PRIMERGY TX300 S5 system with the Xeon E5504 is of the same magnitude as the reference system.

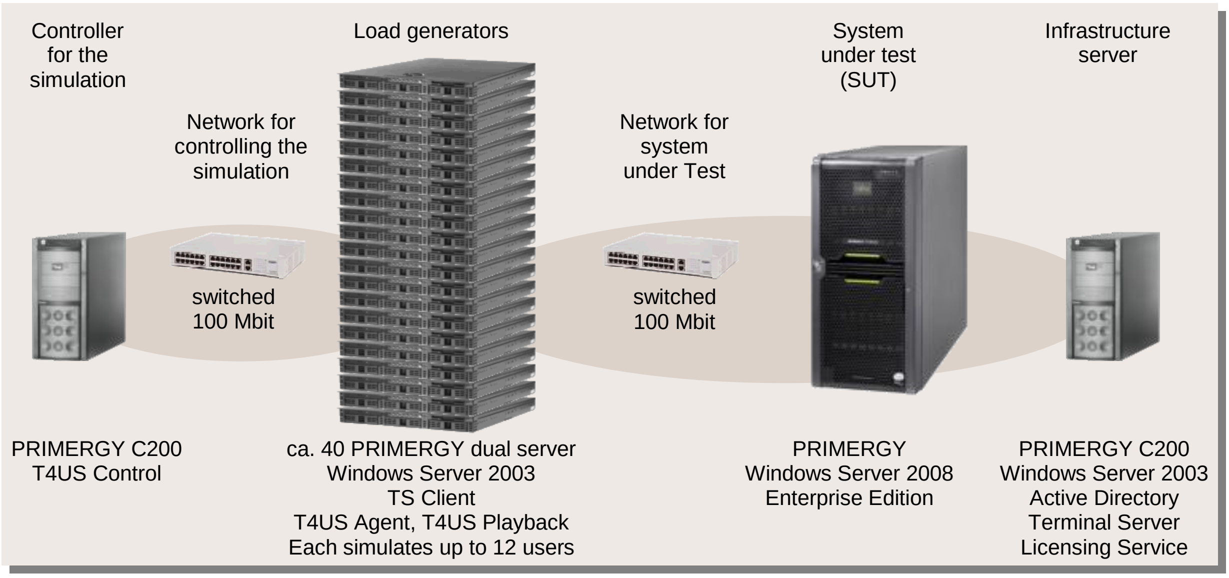
If the PRIMERGY TX300 S5 system is equipped with the more powerful Xeon X5570 processors, more than a doubling of the system performance is achieved with the Terminal Server benchmark. In addition to the higher clock frequency and the larger second-level cache, memory access is also faster. In both measurements the system was equipped with 6 × 4 GB memory. The Terminal Server benchmark benefits greatly from the additional logical cores. The Turbo Boost Technology of the Xeon X5570 also has a performance-enhancing impact on the load peaks of the Terminal Server benchmark.

On the whole, the PRIMERGY TX300 S5 is well suited for Terminal Server applications. Technologies such as Hyper-Threading and Turbo Boost result in a strong processor performance, which in connection with a large memory configuration permits a high number of Terminal Server users in practice. However, the actual number of users is always based on the current customer load profile.

Benchmark environment*

The figure below shows the environment in which the terminal server performance measurements are implemented.

A load-generator can simulate a great number of users because the applications run on the server. With the terminal server protocols, only keyboard input and mouse clicks are transferred to the server and changes to the screen content to the client. Thus, a large network bandwidth is not needed. The connection of the load simulators to the terminal server (also called "system under test" (SUT)) was established by means of a 100-Mbit Ethernet network where the terminal server was connected through the gigabit uplink. The user profiles were stored on the terminal server. The users' files to be read and written during the measurement were also maintained locally on the terminal server. The infrastructure server also located in the SUT network provides basic services such as Active Directory, DNS, and Terminal Services Licensing. Log-in of the simulated users was always effected to the Active Directory.



System Under Test (SUT):
The terminal server runs the Microsoft Terminal Services, that are included in the operating system. Apart from the applications listed in the table, no other software was installed on the Terminal Server.

T4US Measurement Environment:
The load generators simulate different users working with the terminal server. One T4US controller centrally controls and monitors the entire simulation process. The infrastructure server provides basic services.

Hardware	
Model	PRIMERGY TX300S5 // PRIMERGY TX200S4
Processor	1-2 Xeon E5504 // 1-2 Xeon E5430 1-2 Xeon X5570
Memory	24 GB // 12 GB
Network Interface	1×1-GBit LAN Intel (onboard) // Broadcom
Disk Subsystem	2 × SAS Controller: LSI 1078 Modular RAID 4 × SAS disks, 15 krpm, RAID 0
Software	
Operating System	Windows Server 2008 x64 Enterprise Edition without Hyper-V
Version	Service Pack 1
Network Protocol	TCP/IP
Disk Organization	1 Volume: OS 1 Volume each: data and pagefile
Terminal Server Software	Microsoft Terminal Services
Application	Microsoft Office 2003 (32-bit), 7-Zip 4.57

Load Generator-Hardware	
Model	PRIMERGY RX100 S3 // PRIMERGY BX300
# of Load Generators	20 // 20
Processor	Pentium D 940 // 2 × Pentium III 933 MHz
Memory	2 GB // 1 GB
Network Interface	2 × 1 GBit LAN
T4US Controller and Infrastructure Server Hardware	
Model	PRIMERGY C200
Processor	2 × Pentium III 1.40 GHz
Memory	1.5 GB
Network Interface	2 × 100 MBit LAN
Software	
Operating System	Windows Server 2003 Standard Edition SP2
Network Protocol	TCP/IP
RDP Client	6.0.6000.16459
T4US Version	3.3
T4US Load Profile	T4US Load profile V2

* Some components may not be available in all countries / sales regions.

vServCon

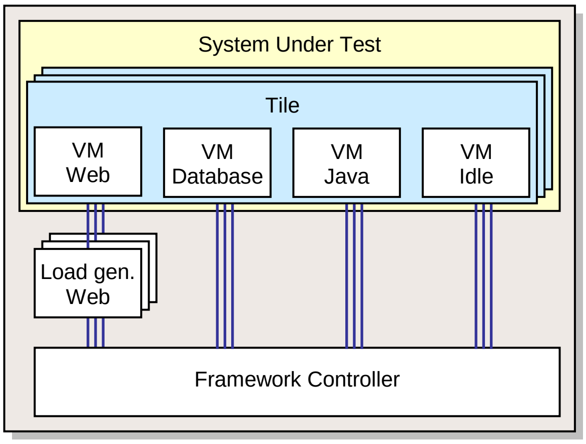
Benchmark description

vServCon is a benchmark used by Fujitsu Technology Solutions to compare server configurations with hypervisor with regard to their suitability for server consolidation. This allows both the comparison of systems, processors and I/O technologies as well as the comparison of hypervisors, virtualization forms and additional drivers for virtual machines.

vServCon is a framework that summarizes already established benchmarks in order to reproduce the load of a consolidated and virtualized server environment. Four proven benchmarks are used, which cover the application scenarios database, application server and web server.

Application scenario	Benchmark	No. of logical CPU cores	Memory
Database	Sysbench (adapted)	2	1.5 GB
Java application server	SPECjbb (adapted, with 50% - 60% load)	2	2 GB
Web server	WebBench	1	1.5 GB

Each of the three standard benchmarks is allocated to a dedicated virtual machine (VM). Add to these a fourth machine, the so-called idle VM. These four VMs make up a »tile«. Depending on the performance capability of the underlying server hardware, you may as part of a measurement also have to start several identical tiles in parallel in order to achieve a maximum performance score.



Each of the three vServCon application scenarios provides a specific benchmark result in the form of application-specific transaction rates for each VM. In order to derive a score for a specified number of tiles the individual benchmark results are put in relation to the respective results of a reference system. A PRIMERGY RX300 S3 has been defined as reference system for vServCon. The resulting dimensionless performance values are then weighted allowing for the number of virtual CPUs and memory size and are added up for all VMs and tiles. The outcome is the vServCon score for this tile number.

Starting as a rule with one tile, this procedure is performed for an increasing number of tiles until no further significant increase in this vServCon score occurs. The final vServCon score is then the maximum of the vServCon scores for all tile numbers, and reflects the maximum total consolidation benefit of all VMs for a server configuration with hypervisor.

vServCon also documents the total CPU load of the host (VMs and all other CPU activities) and where possible the electrical power consumption.

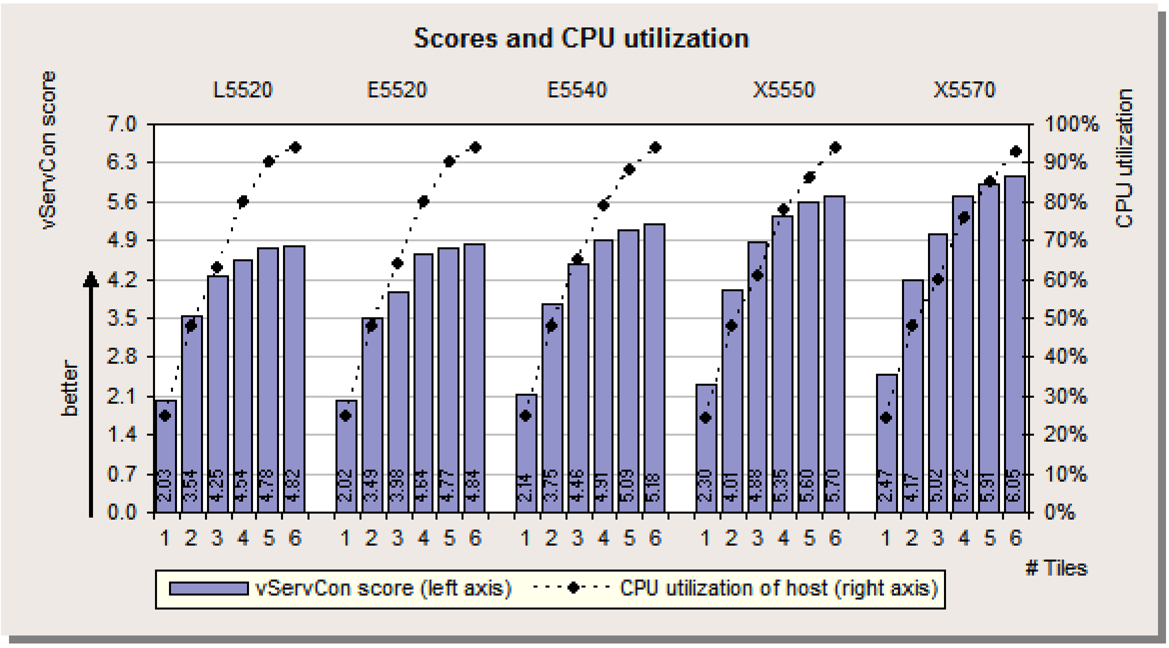
The score is intended to express a virtualization-specific system performance that can be achieved - right through to maximum utilization of the CPU resources - with a many VMs. In other words, the score would not be significant if a limitation were to occur during a vServCon measurement for an unnecessarily small number of tiles, e.g. as a result of an inadequately sized disk connection or insufficient main memory. This is why the measurement environment for vServCon measurements is designed in such a way that only the CPU is the limiting factor and that no limitations occur as a result of other resources. For this purpose and for purposes of comparability an exactly defined profile is used for the virtual hardware resources, the operating system and the applications for all the VMs used in vServCon.

A detailed description of vServCon is available in the document: vServCon - Benchmark Overview.

Benchmark results

The PRIMERGY TX300 S5 is very suitable for application virtualization thanks to the considerable progress made in processor technology. Compared with the previous system, a practically double virtualization performance can be achieved (measured in vServCon score). On the basis of the previously described vServCon profile almost optimal utilization of the CPU system resources is possible with 18 real application VMs (equivalent to six tiles) if the system is fully assembled with two Xeon processors.

For the PRIMERGY TX300 S5 this is illustrated in the first diagram by the vServCon scores in relation to processor and number of tiles. The respective CPU loads of the host have also been entered. The number of tiles with optimal CPU load is typically at about 90%; beyond that you have overload, which is where virtualization performance no longer increases, and sinks again respectively.



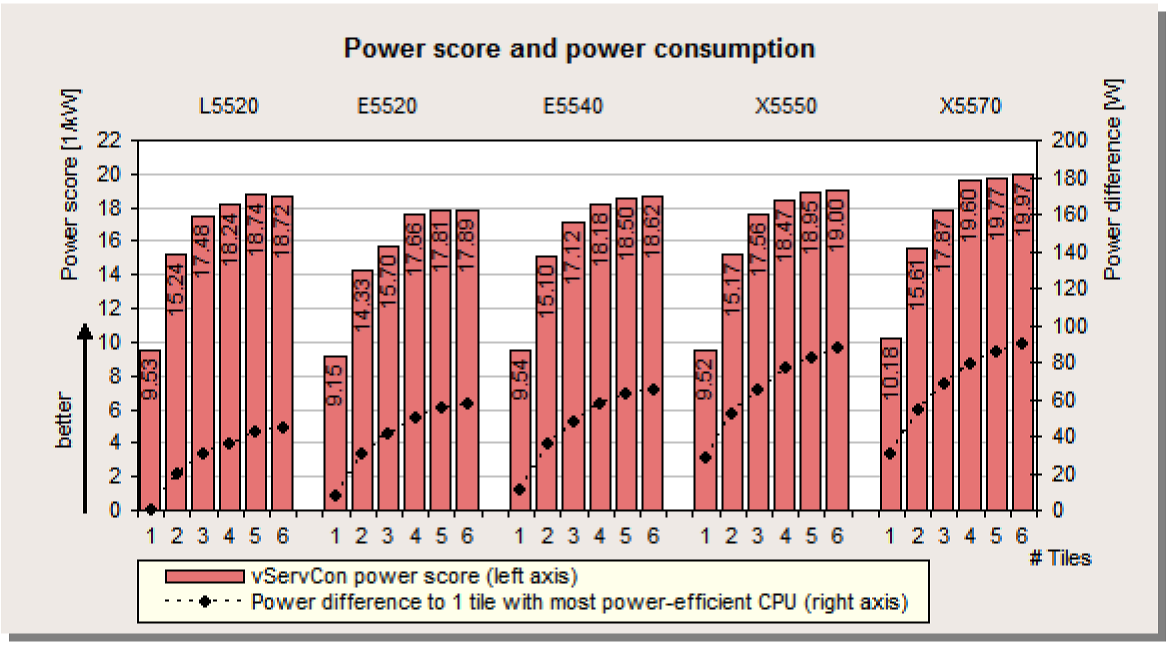
The performance capability of these current Xeon processors is seen in the increase in overall performance as far six tiles. The vServCon scores also noticeably increase with the processor frequency.

The high number of tiles is enabled particularly via Hyper-Threading whereby a physical processor core is divided into two logical cores so that 16 logical cores are available for the hypervisor. This standard feature thus generally increases the virtualization performance of a system.

The scaling curves for the number of tiles are specifically for systems with Hyper-Threading. As a result of the approximate four logical CPUs used per tile (see the benchmark description), a parallel use of the same physical cores by several VMs is avoided up to and including two tiles. That is why the performance curve scales almost ideally in this sector. For tile quantities above two tiles the growth is flatter up to CPU full utilization.

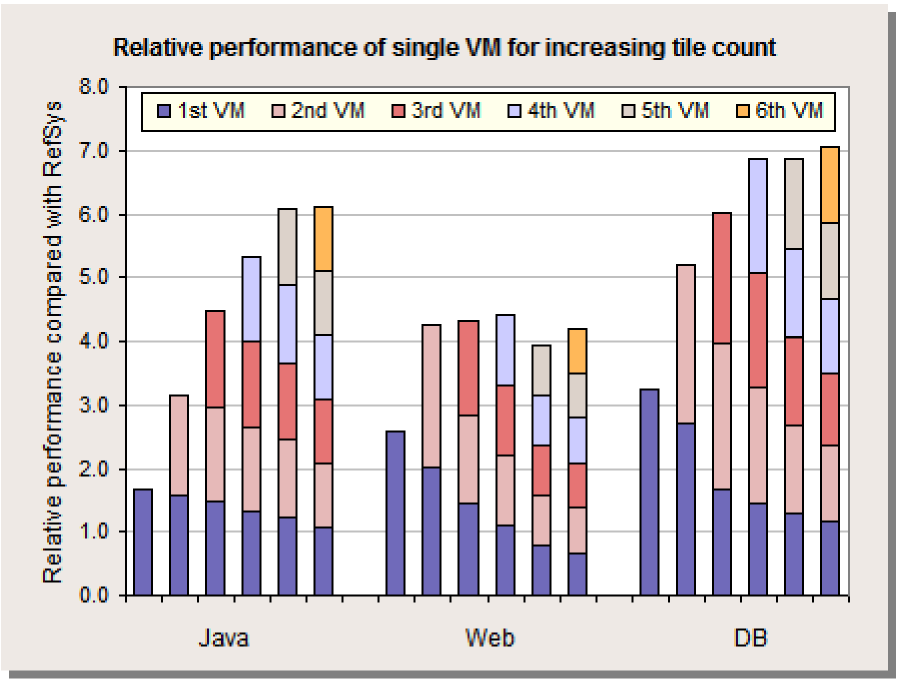
A guideline in the virtualization environment for selecting main memory is that a sufficient quantity is more important than the speed of the memory accesses.

Saving electrical energy is an important aspect of server consolidation. With the Xeon E5540 processor it is e.g. possible to increase the virtualization performance by 75% merely by doubling the number of real application VMs from 3 to 6, while at the same time electrical power consumption only increases by about 11%. The power aspects for the processors depicted above are illustrated in the following diagram. It shows on the one hand the absolute differences in power consumption and on the other hand the ratio of the vServCon score to power consumption in kW, denoted in the diagram in short as »vServCon power score«.



Previously, the virtualization performance of the system was analyzed as a whole. Below, performance is also to be discussed from the viewpoint of an individual application VM in the described virtualized environment. As an example, the system is analyzed for this purpose with the processor Xeon X5570.

If the number of application VMs is optimal as far as the overall performance is concerned, the performance of an individual VM is already notably lower than in operational low-load situations. This is illustrated in the diagram opposite through the relative performance in ratio to the reference system with an individual application VM of each of the three types for increasing VM numbers. The first column of a group views one VM in the array of a total of three application VMs (1 tile), the second one is for the array of 6 application VMs (2 tiles), etc. The values are presented - both individually and in total for all VMs of the respective type - through the height of the stacked columns.



With regard to the numbers of VMs on a virtualization host it is necessary in a specific case to weigh up the performance requirements of an individual application against the overall requirements.

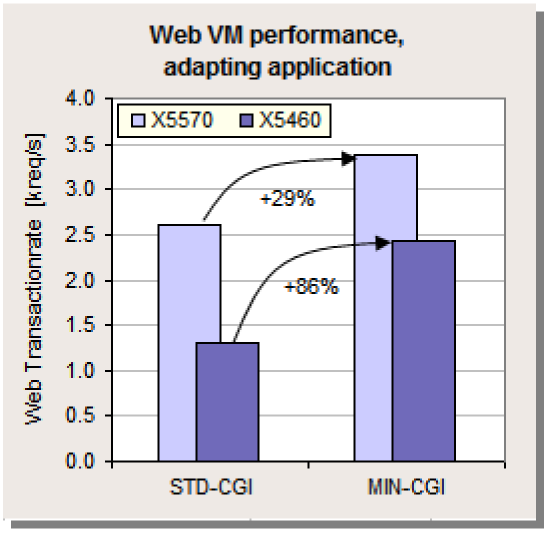
If you want to run applications in virtual machines at maximum performance, it is worth looking at the application profiles that make higher demands of a virtualization solution more closely. These include application scenarios like web server that are a great drain on memory management.

The first method of optimization is applied to the application scenario. The influence of the implementation of dynamic contents on performance can be impressively seen in the example of a web server with dynamic pages. Dynamic contents are frequently implemented as CGI programs (or scripts). Each time they are selected, these CGI programs generate a new process, which is rather complex for the hypervisor. Alternatively, dynamic contents can be implemented by using PHP, ASP or similar methods, which result in no overheads through newly generated processes. This can be simulated in vServCon by varying the share of HTTP requests, which start such CGI programs, in the load profile of the web server VMs. The diagram below illustrates the impact on performance of an unmodified Linux kernel in the VM. The two load profiles compared are:

Load profiles for web server	
STD-CGI	This defines that 16% of all HTTP requests and 2% of all HTTP-SSL requests on the web server start a CGI program. Makes great demands of a virtualization solution.
MIN-CGI	STD-CGI profile, but without the 16% CGI-HTTP requests. The load on a web server is decreased by this reduction in the number of CGI processes; but this reduces the costs within the virtualization solution a great deal more. Both effects together make so much additional CPU performance available that the web transaction rate for VMs is significantly increased.

All the previously described measurements use the STD-CGI profile as standard.

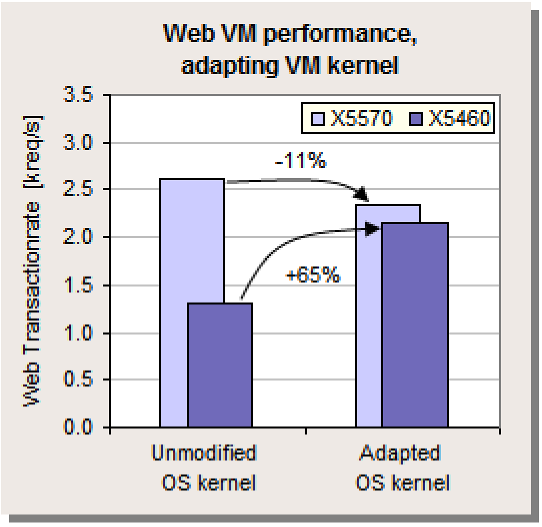
The diagram clearly shows the advance of processor technology in the virtualization sector. A PRIMERGY TX300 S5 (with Xeon X5570) is compared with the previous system. Whereas a Web-VM with the previous system could achieve an increase in performance of 86% as a result of optimization in the application scenario, this is only 29% with the current processor generation. Due to the »Extended Page Tables« (EPT) the system is so good with the more demanding load profile »STD-CGI« that the optimization reserve based on the application scenario is reduced. The performance growth when we look at »MIN-CGI« instead of »STD-CGI« is a similar value to the value for a non-virtualized system.



Conclusion about optimization for the application scenario: This is of interest but the benefit is the same level as in a non-virtualized system.

The second method of optimization is applied below the application level in the VM. Increases in performance are in principle possible both through appropriate processor functions and through a suitable hypervisor or also through an operating system or driver in the VM that has been specially adapted to the hypervisor. Such an adapted VM actively supports the hypervisor in its work, and as a consequence the virtualization overhead can in part be significantly reduced.

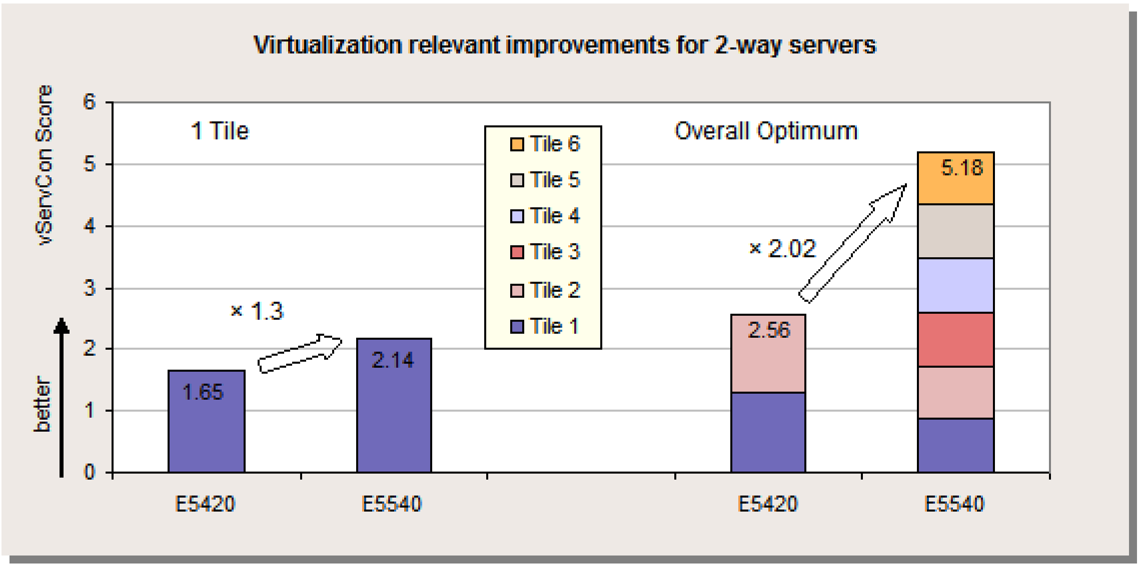
With the previous generations of the PRIMERGY servers the use of a kernel modified for virtualization for the web server VM resulted in a large increase in performance. The diagram compares a PRIMERGY TX300 S5 (with Xeon X5570) with the previous system for two kernels under the load profile STD-CGI. The one kernel is the unmodified LINUX kernel, and the other is the adapted kernel.



The diagram shows that the quality of the virtualization support with a current processor from the Xeon 5500 series is so good in the meantime that such a kernel modification can no longer be recommended. With the previous system with a Xeon X5460 processor an increase in performance of 65% was possible.

The non-modified LINUX kernel (SMP) is the standard for the measurements described above unless otherwise specified.

The virtualization-relevant progress in processor technology compared to the previous system has an effect on the one hand on an individual VM (e.g. via EPT) and, on the other hand, on the possible maximum number of VMs up to CPU-full utilization (via Hyper-Threading). The following comparison shows the percentages for both types of improvements more clearly. The comparison is made with a previous system with 2 × Xeon E5420 (2.5 GHz) and a PRIMERGY TX300 S5 with 2 × Xeon E5540 (2.53 GHz); both have the same number of physical cores.



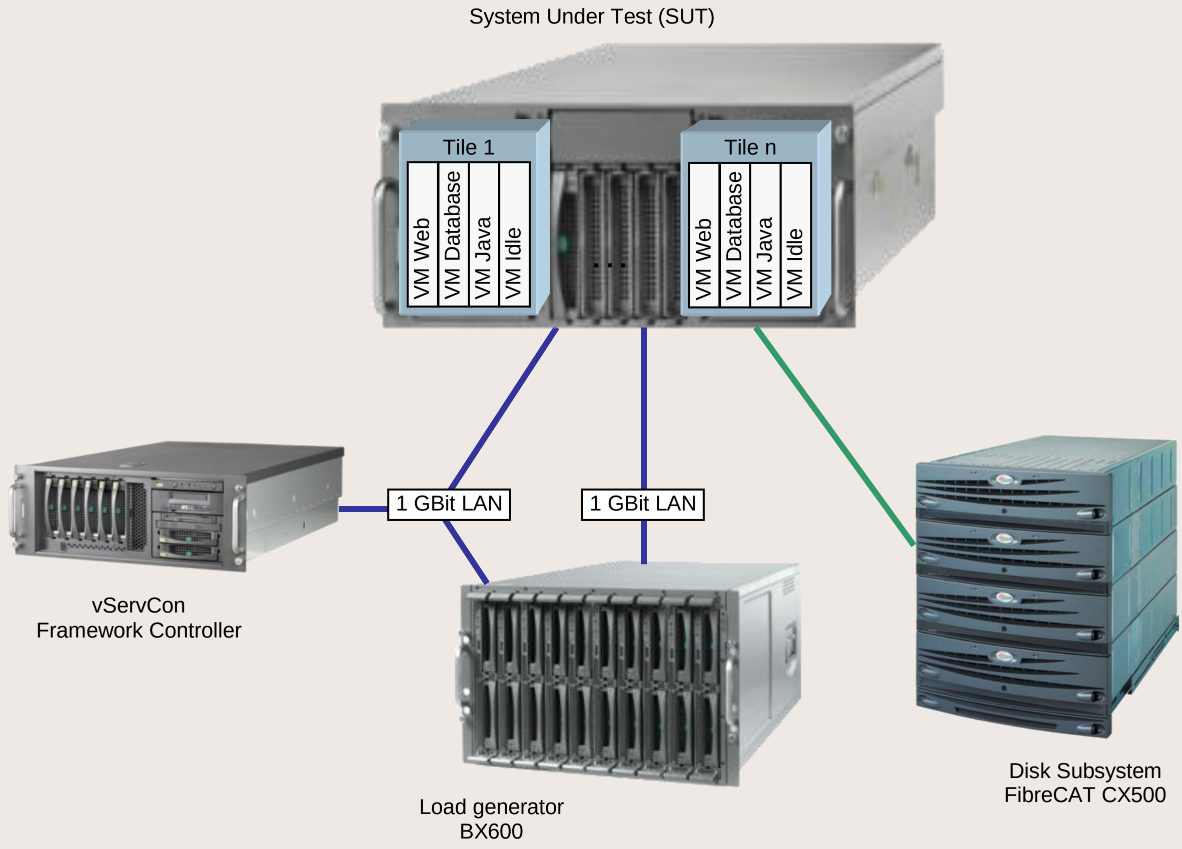
One sees an increase of the vServCon score by a factor of 2.02. The one reason is the performance increase that can be achieved for an individual VM (factor 1.30, see score with one tile). The other reason is that more tiles are possible. However, this is not a true-value tripling (from two to six tiles) as the performance growth flattens out the higher the number of tiles.

We thus explicitly warn against regarding the improved virtualization performance – shown in the score – as a complete improvement for an individual VM. More than some 30% - 50% more throughput in contrast to a processor with the same clock speed in the previous generation is not possible.

Benchmark environment*

The measurements were made with the environment described below:

vServCon Benchmark Environment



SUT Hardware	
Model	PRIMERGY TX300 S5
Processor	2 × Xeon L5520 (2.27 GHz) 2 × Xeon E5520 (2.27 GHz) 2 × Xeon E5540 (2.53 GHz) 2 × Xeon X5550 (2.67 GHz) 2 × Xeon X5570 (2.93 GHz)
Memory	48 GB (a PC3-10600R each, 8 GB, in DIMM-1A until DIMM-1F)
Network interface	2 × 1-GBit LAN (onboard); one for load, one for control
Disk Subsystem	No internal hard disks were used, solely one storage system FibreCAT CX500. One 50 GB LUN per tile for the »virtual disk files« of the VMs. Each LUN is a RAID 0 array consisting of 6 Seagate ST373454 disks (15 krpm)
Storage connection	Via FC controller Qlogic QLE 2460
SUT Software	
Operating system	Hypervisor VMware ESX Server.
Version	Version 4.0.0 build 157368
BIOS	Version 6.00 R1.02a3.2619; Default settings

SUT: Virtualization-specific details	
Web server VM kernel, original	SLES10 SP2, 32-bit, 2.6.16.60-0.23-smp
Web server VM Kernel, adapted	SLES10 SP2, 32-bit, 2.6.16.60-0.23-vmi (kernel with VMware VMI interface)
General details	Described in Benchmark Overview vServCon

Load Generator Hardware	
Model	3 server blades per tile in PRIMERGY BX600 S2 chassis
Processor	X86 Family 15, Model 4, Stepping 1, Genuine Intel 3000 MHz
Memory	1 – 2 GB
Network interface	2 × 1 GBit LAN each
Operating system	W2K3 EE

* Some components may not be available in all countries / sales regions.

Literature

PRIMERGY Systems

http://ts.fujitsu.com/primergy

PRIMERGY TX300 S5

Data sheet
http://docs.ts.fujitsu.com/dl.aspx?id=e09f066d-690f-48ba-aec6-fa9fbd07b049

Memory performance of Xeon 5500 (Nehalem EP) based PRIMERGY servers
http://docs.ts.fujitsu.com/dl.aspx?id=1a55e793-4231-490d-8941-42c89f5670ad

PRIMERGY Performance

http://ts.fujitsu.com/products/standard_servers/primergy_bov.html

OLTP-2

Benchmark Overview OLTP-2
http://docs.ts.fujitsu.com/dl.aspx?id=e6f7a4c9-aff6-4598-b199-836053214d3f

SAP SD

http://www.sap.com/benchmark

Benchmark Überblick SAP SD
http://docs.ts.fujitsu.com/dl.aspx?id=0a1e69a6-e366-4fd1-a1a6-0dd93148ea10

SPECcpu2006

http://www.spec.org/osg/cpu2006

Benchmark Overview SPECcpu2006
http://docs.ts.fujitsu.com/dl.aspx?id=1a427c16-12bf-41b0-9ca3-4cc360ef14ce

SPECjbb2005

http://www.spec.org/jbb2005

Benchmark Overview SPECjbb2005
http://docs.ts.fujitsu.com/dl.aspx?id=5411e8f9-8c56-4ee9-9b3b-98981ab3e820

SPECsfs2008

http://www.spec.org/sfs2008

StorageBench

Performance Report - Modular RAID for PRIMERGY
http://docs.ts.fujitsu.com/dl.aspx?id=8f6d5779-2405-4cdd-8268-1f948ba050e6

iometer
http://www.iometer.org

Terminal Server

Terminal Server Sizing Guide (EN)
http://docs.ts.fujitsu.com/dl.aspx?id=377f3a1c-54ce-4224-8322-3008ac213e41

Terminal Server Sizing Guide - 64-bit Technology (EN)
http://docs.ts.fujitsu.com/dl.aspx?id=d9f67e26-900d-4745-8abe-f6fd677f829a

Microsoft Windows 2008 and Terminal Services
http://www.microsoft.com/windowsserver2008/en/us/ts-product-home.aspx

vServCon

Benchmark Overview vServCon
http://docs.ts.fujitsu.com/dl.aspx?id=b953d1f3-6f98-4b93-95f5-8c8ba3db4e59

Contact

PRIMERGY Hardware

PRIMERGY Product Marketing
mailto:Primergy-PM@ts.fujitsu.com

PRIMERGY Performance and Benchmarks

mailto:primergy.benchmark@ts.fujitsu.com

Published by department:

Enterprise Products
PRIMERGY Server
PRIMERGY Performance Lab
mailto:primergy.benchmark@ts.fujitsu.com

Internet:
http://ts.fujitsu.com/primergy

Extranet:
http://partners.ts.fujitsu.com/com/products/servers/primergy