THE POSSIBILITIES ARE INFINITE

WHITE PAPER

Version 2.0
October 2009

Performance Report PRIMERGY BX620 S5

Pages 37

Abstract

This document contains a summary of the benchmarks executed for the PRIMERGY BX620 S5.

The PRIMERGY BX620 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, L5506, E5506, E5520, E5540, X5550, X5560 and X5570
- SPECjbb2005
  Measurement with Xeon X5570
- OLTP-2
  Measurements with Xeon E5502, E5504, E5506, E5520, E5530, E5540, X5550, X5560 and X5570
- StorageBench
  Measurements with LSI 1064E SAS IME Storage Module Controller
  and MegaRAID SAS PCI Express ™ ROMB Controller
- Terminal Server
  Measurements with Xeon E5504 and X5570
- 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
  Additional measurements with Xeon E5502, E5504, L5506, E5506, E5520, E5540, X5550, X5560 and X5570
  Measurements with Xeon L5520, L5530 and E5530
- StorageBench: typo
- OLTP-2: corrections, updated diagrams
- vServCon: corrections, updated diagrams

Technical Data

PRIMERGY BX600 S3 Blade Servers are highly scalable 19-inch rack systems occupying 7 height units (U) that accommodate a maximum of 10 compact dual server blades. They ideally satisfy the requirements of enterprise data centers and Internet or application service providers.

In addition to the hot-pluggable server blades, two redundant fan units with two fans each, up to four redundant power supply modules (each with three additional fans), one KVM switch, two redundant management blades and optionally up to four GBit Ethernet pass-thru blades or GBit Ethernet switch blades or GBit Ethernet intelligent blades as well as optionally up to two fibre channel pass-thru blades or fibre channel switch blades or fibre channel access gateways can be integrated in a PRIMERGY BX600 S3.

The PRIMERGY BX620 S5 dual server blades have an Intel 5520 chip set, two Dual-Core or Quad-Core processors Intel Xeon 5500 series, 12 DIMM slots for up to 96 GB PC3-10600 or PC3-8500 registered ECC DDR3-SDRAM, a bus with 800, 1067 or 1333 MHz timing - depending on the processor and memory used, two 1-GBit Ethernet controllers, pluggable storage modules for internal hard disks, two bays for 2.5'' SSD or SAS hard disks and a PCIe x8 slot.

The PRIMERGY server management solution ServerView (optionally with Deployment Manager and Remote Management) simplifies installation, administration and monitoring of the servers.




See Data sheet PRIMERGY BX620 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.

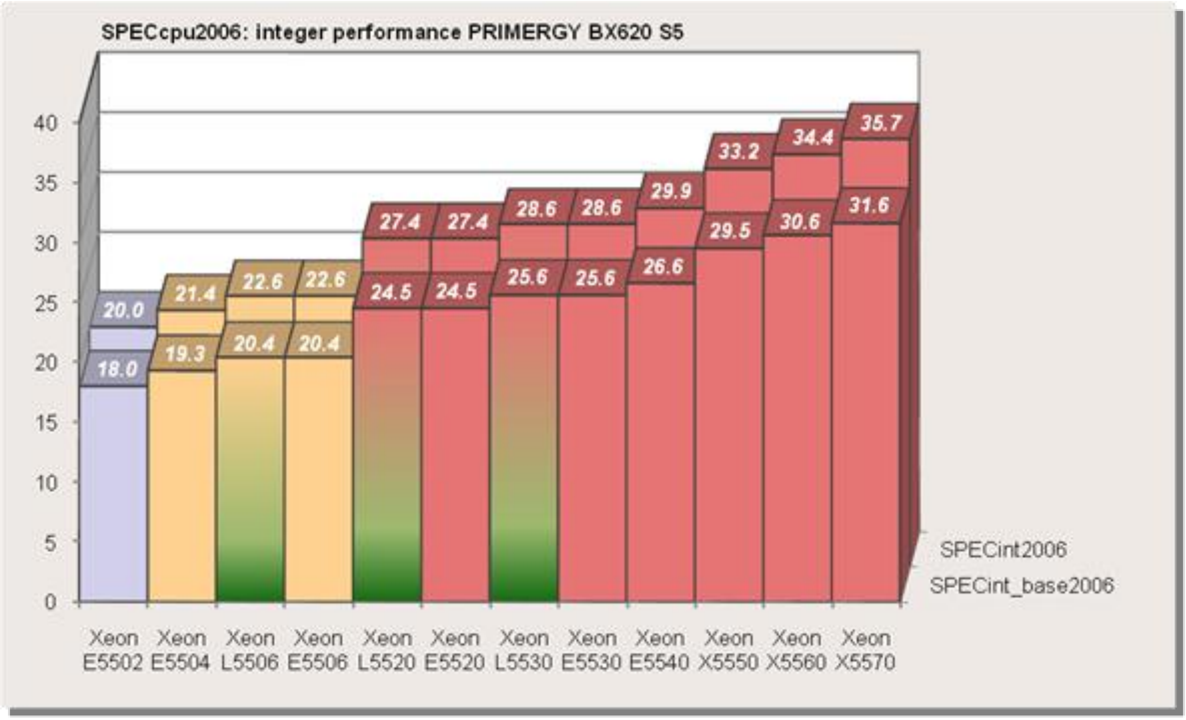
We do not submit all SPECcpu2006 measurements for publication at SPEC. So not all results appear on SPEC’s web sites. As we archive the log data for all measurements, we are able to prove the correct realization of the measurements any time.

Benchmark results

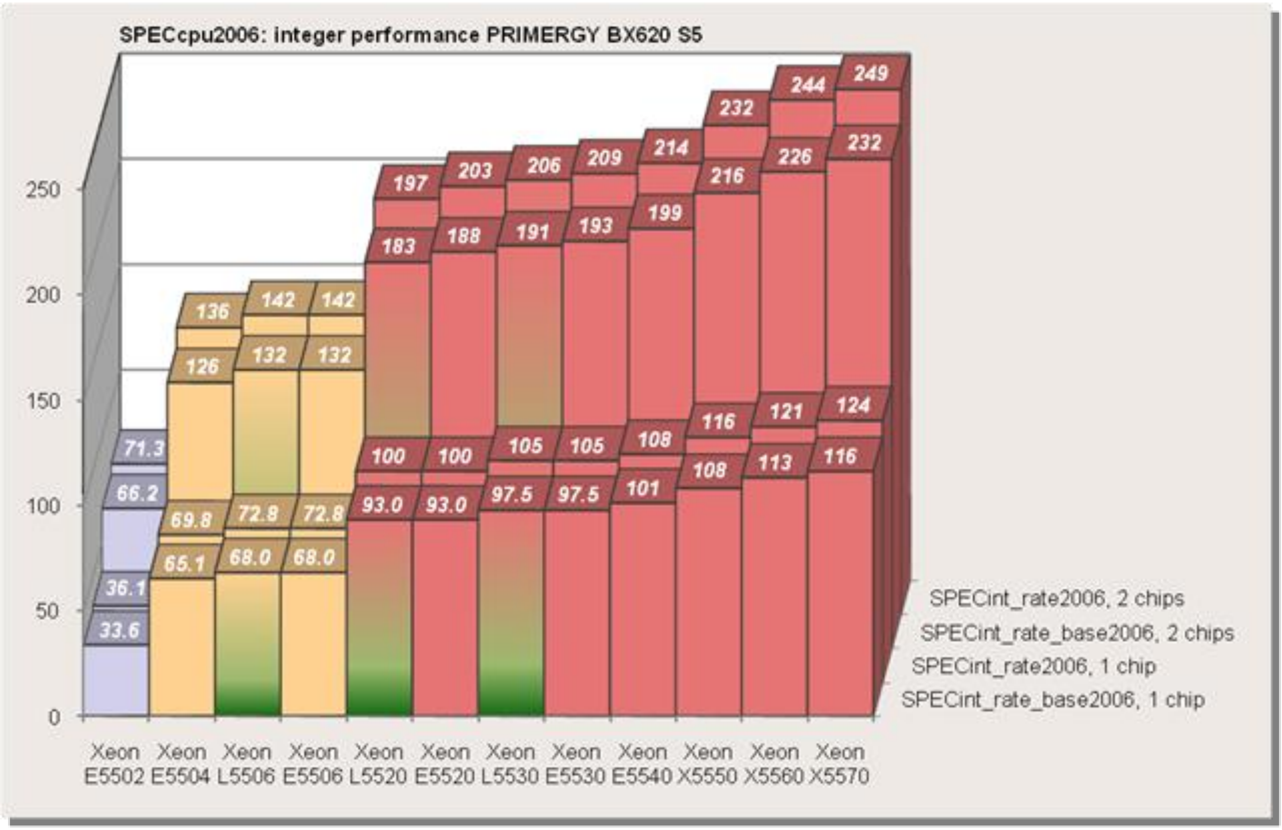
The PRIMERGY BX620 S5 was measured with the processors Xeon E5502, E5504, L5506, E5506, L5520, E5520, L5530, E5530, E5540, X5550, X5560 and X5570. The benchmark programs were compiled with the Intel C++/Fortran compiler 11.0 and run under SUSE Linux Enterprise Server 10 SP2 (64-bit). In the following 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).

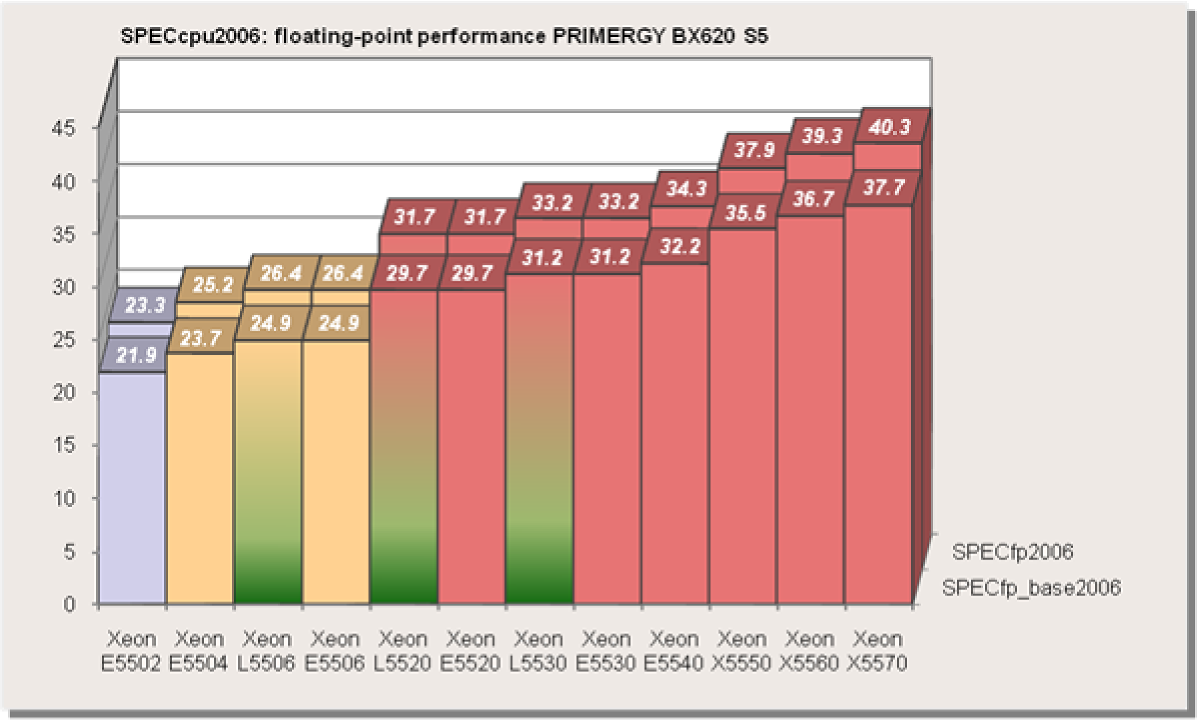
Processor	Cores	GHz	L3 cache	Bus	TDP	SPECint_base2006 2 chips	SPECint2006 2 chips
Xeon E5502	2	1.87	4 MB	800 MHz	80 watt	18.0	20.0
Xeon E5504	4	2	4 MB	800 MHz	80 watt	19.3	21.4
Xeon L5506	4	2.13	4 MB	800 MHz	60 watt	20.4 (est.)	22.6 (est.)
Xeon E5506	4	2.13	4 MB	800 MHz	80 watt	20.4	22.6
Xeon L5520	4	2.27	8 MB	1067 MHz	60 watt	24.5 (est.)	27.4 (est.)
Xeon E5520	4	2.27	8 MB	1067 MHz	80 watt	24.5	27.4
Xeon L5530	4	2.40	8 MB	1067 MHz	60 watt	25.6 (est.)	28.6 (est.)
Xeon E5530	4	2.40	8 MB	1067 MHz	80 watt	25.6	28.6
Xeon E5540	4	2.53	8 MB	1067 MHz	80 watt	26.6	29.9
Xeon X5550	4	2.67	8 MB	1333 MHz	95 watt	29.5	33.2
Xeon X5560	4	2.80	8 MB	1333 MHz	95 watt	30.6	34.4
Xeon X5570	4	2.93	8 MB	1333 MHz	95 watt	31.6	35.7



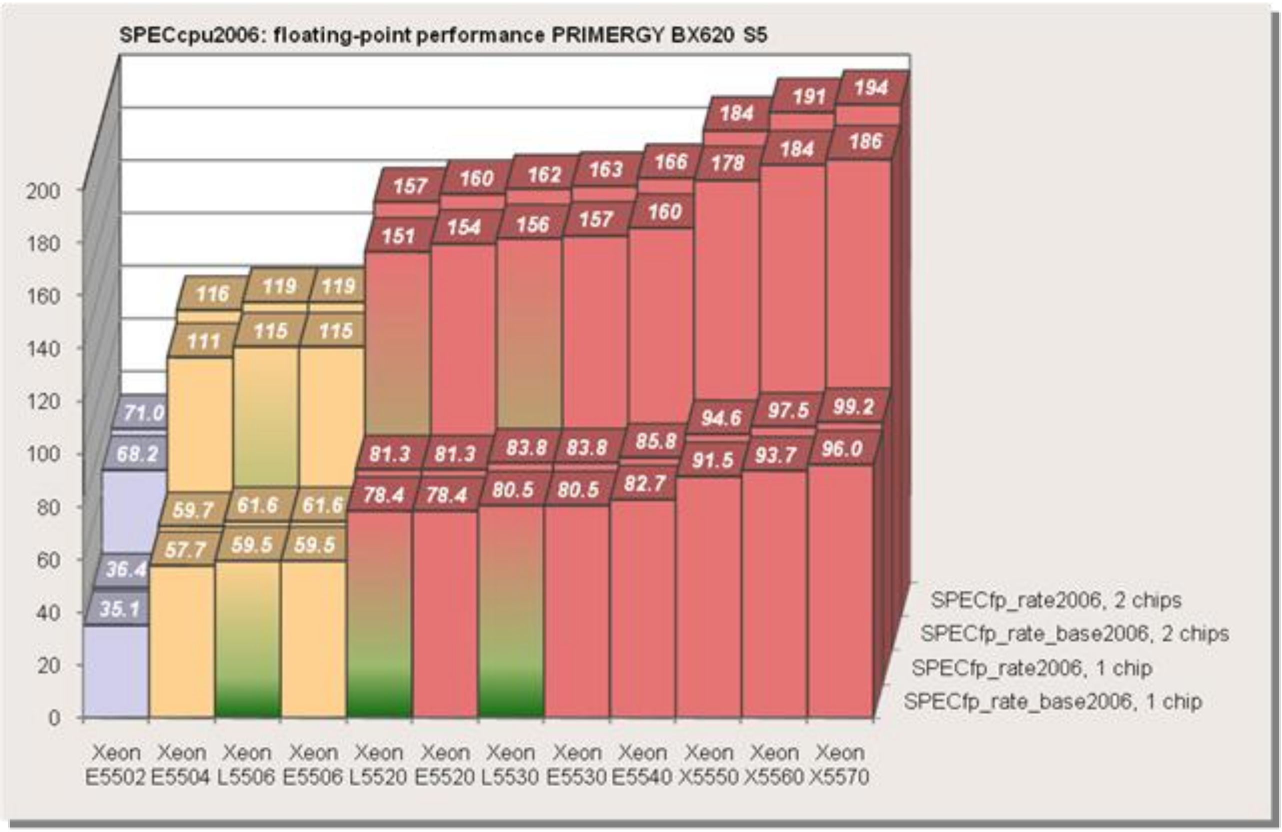
Processor	Cores	GHz	L3 cache	Bus	TDP	SPECint_rate_base2006		SPECint_rate2006	
						1 chip	2 chips	1 chip	2 chips
Xeon E5502	2	1.87	4 MB	800 MHz	80 watt	33.6	**66.2**	36.1	**71.3**
Xeon E5504	4	2	4 MB	800 MHz	80 watt	65.1	**126**	69.8	**136**
Xeon L5506	4	2.13	4 MB	800 MHz	60 watt	68.0 (est.)	132	72.8 (est.)	142
Xeon E5506	4	2.13	4 MB	800 MHz	80 watt	68.0	**132**	72.8	**142**
Xeon L5520	4	2.27	8 MB	1067 MHz	60 watt	93.0 (est.)	183	100 (est.)	197
Xeon E5520	4	2.27	8 MB	1067 MHz	80 watt	93.0	**188**	100	**203**
Xeon L5530	4	2.40	8 MB	1067 MHz	60 watt	97.5 (est.)	191	105 (est.)	206
Xeon E5530	4	2.40	8 MB	1067 MHz	80 watt	97.5	**193**	105	**209**
Xeon E5540	4	2.53	8 MB	1067 MHz	80 watt	101	**199**	108	**214**
Xeon X5550	4	2.67	8 MB	1333 MHz	95 watt	108	216	116	232
Xeon X5560	4	2.80	8 MB	1333 MHz	95 watt	113	226	121	244
Xeon X5570	4	2.93	8 MB	1333 MHz	95 watt	116	232	124	249



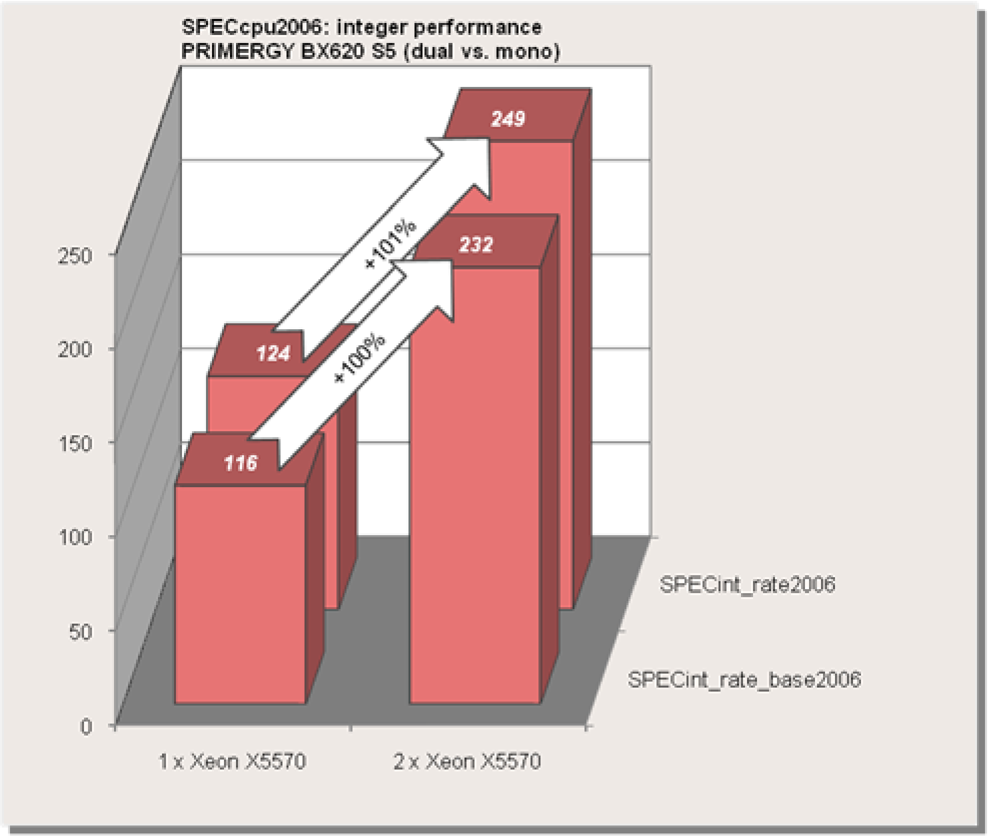
Processor	Cores	GHz	L3 cache	Bus	TDP	SPECfp_base2006 2 chips	SPECfp2006 2 chips
Xeon E5502	2	1.87	4 MB	800 MHz	80 watt	21.9	23.3
Xeon E5504	4	2	4 MB	800 MHz	80 watt	23.7	25.2
Xeon L5506	4	2.13	4 MB	800 MHz	60 watt	24.9 (est.)	26.4 (est.)
Xeon E5506	4	2.13	4 MB	800 MHz	80 watt	24.9	26.4
Xeon L5520	4	2.27	8 MB	1067 MHz	60 watt	29.7 (est.)	31.7 (est.)
Xeon E5520	4	2.27	8 MB	1067 MHz	80 watt	29.7	31.7
Xeon L5530	4	2.40	8 MB	1067 MHz	60 watt	31.2 (est.)	33.2 (est.)
Xeon E5530	4	2.40	8 MB	1067 MHz	80 watt	31.2	33.2
Xeon E5540	4	2.53	8 MB	1067 MHz	80 watt	32.2	34.3
Xeon X5550	4	2.67	8 MB	1333 MHz	95 watt	35.5	37.9
Xeon X5560	4	2.80	8 MB	1333 MHz	95 watt	36.7	39.3
Xeon X5570	4	2.93	8 MB	1333 MHz	95 watt	37.7	40.3

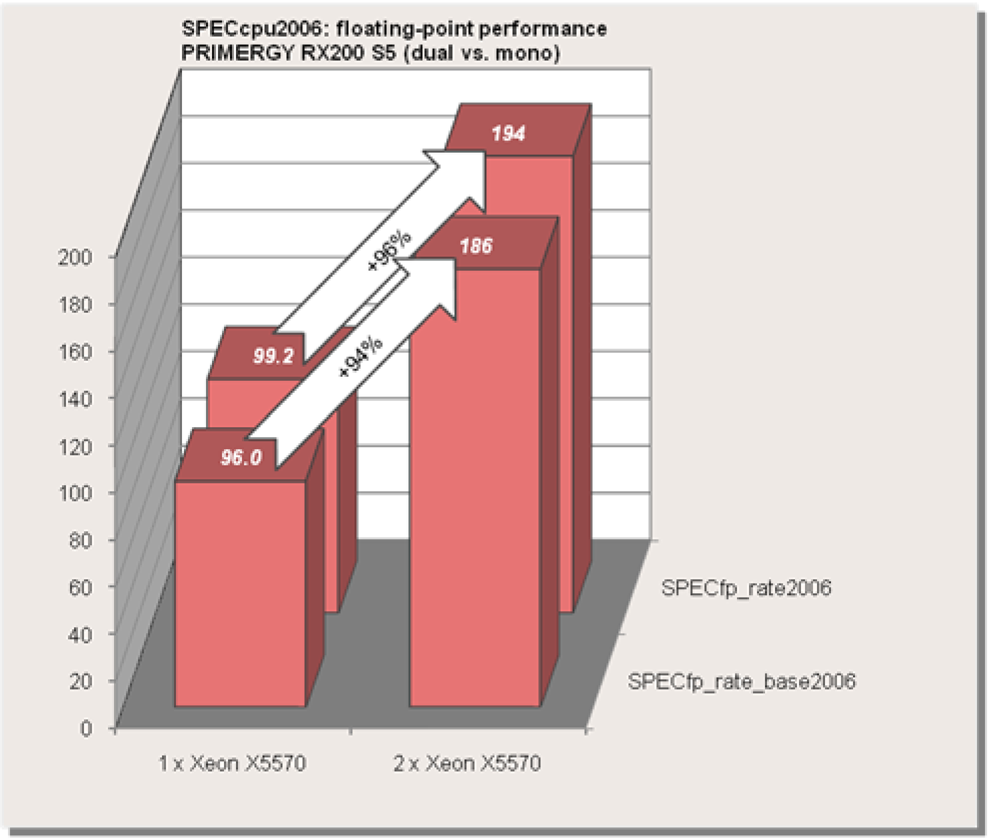


Processor	Cores	GHz	L3 cache	Bus	TDP	SPECfp_rate_base2006		SPECfp_rate2006	
						1 chip	2 chips	1 chip	2 chips
Xeon E5502	2	1.87	4 MB	800 MHz	80 watt	35.1	**68.2**	36.4	**71.0**
Xeon E5504	4	2	4 MB	800 MHz	80 watt	57.7	**111**	59.7	**116**
Xeon L5506	4	2.13	4 MB	800 MHz	60 watt	59.5 (est.)	115	61.6 (est.)	119
Xeon E5506	4	2.13	4 MB	800 MHz	80 watt	59.5	**115**	61.6	**119**
Xeon L5520	4	2.27	8 MB	1067 MHz	60 watt	78.4 (est.)	151	81.3 (est.)	157
Xeon E5520	4	2.27	8 MB	1067 MHz	80 watt	78.4	**154**	81.3	**160**
Xeon L5530	4	2.40	8 MB	1067 MHz	60 watt	80.5 (est.)	156	83.8 (est.)	162
Xeon E5530	4	2.40	8 MB	1067 MHz	80 watt	80.5	**157**	83.8	**163**
Xeon E5540	4	2.53	8 MB	1067 MHz	80 watt	82.7	**160**	85.8	**166**
Xeon X5550	4	2.67	8 MB	1333 MHz	95 watt	91.5	178	94.6	184
Xeon X5560	4	2.80	8 MB	1333 MHz	95 watt	93.7	184	97.5	191
Xeon X5570	4	2.93	8 MB	1333 MHz	95 watt	96.0	186	99.2	194

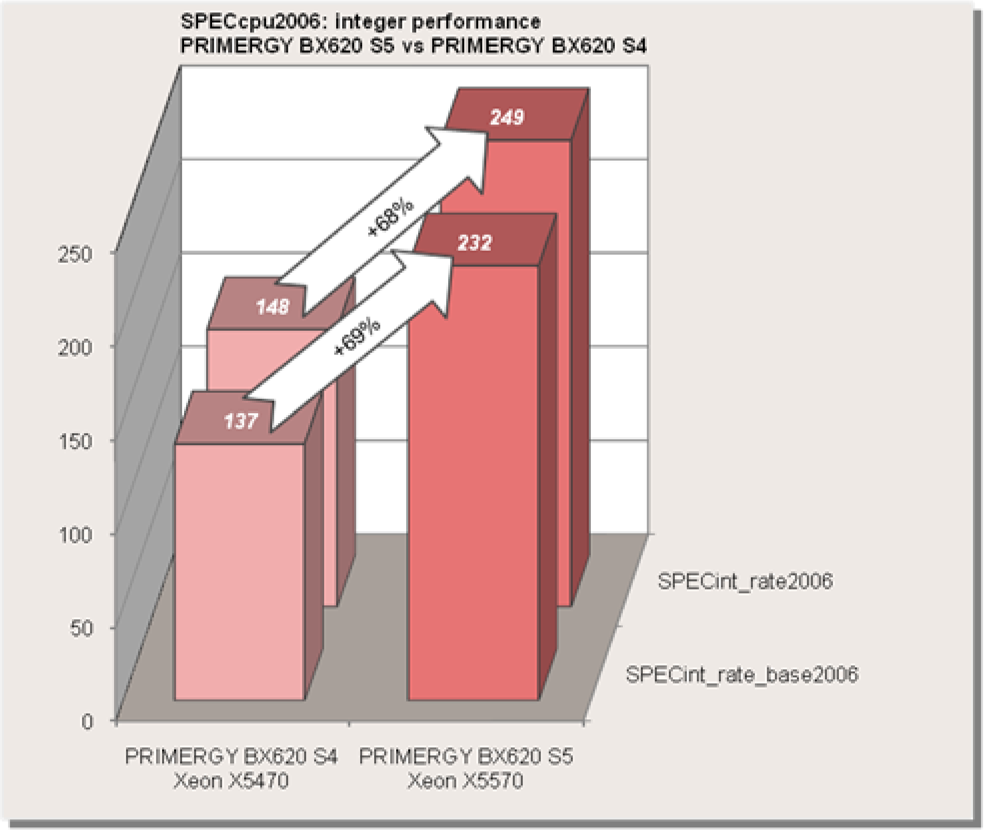


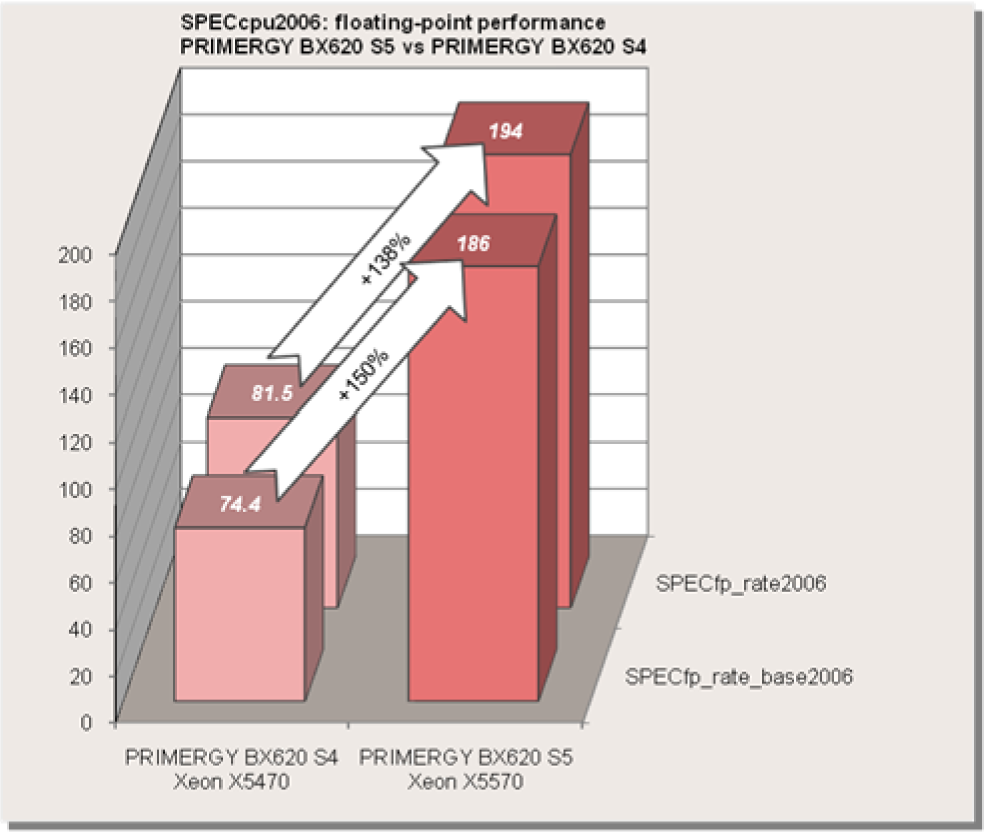
For both the integer and the floating-point test suite the throughput with two processors is about twice as large as that with one processor.





The following two diagrams show the performance of the PRIMERGY BX620 S5 in relation to its predecessor PRIMERGY BX620 S4 both in their highest performance configurations.





Benchmark environment*

All SPECcpu2006 measurements were performed on a PRIMERGY BX620 S5 with the following hardware and software configuration:

Hardware	
Model	PRIMERGY BX620 S5
CPU	Xeon E5502, E5504, L5506, E5506, L5520, E5520, L5530, E5530, E5540, X5550, X5560 and X5570
Number of CPUs	1 chip: Xeon E5502: 2 cores, 2 cores/chip others: 4 cores, 4 cores/chip 2 chips: Xeon E5502: 4 cores, 2 cores/chip others: 8 cores, 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
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:

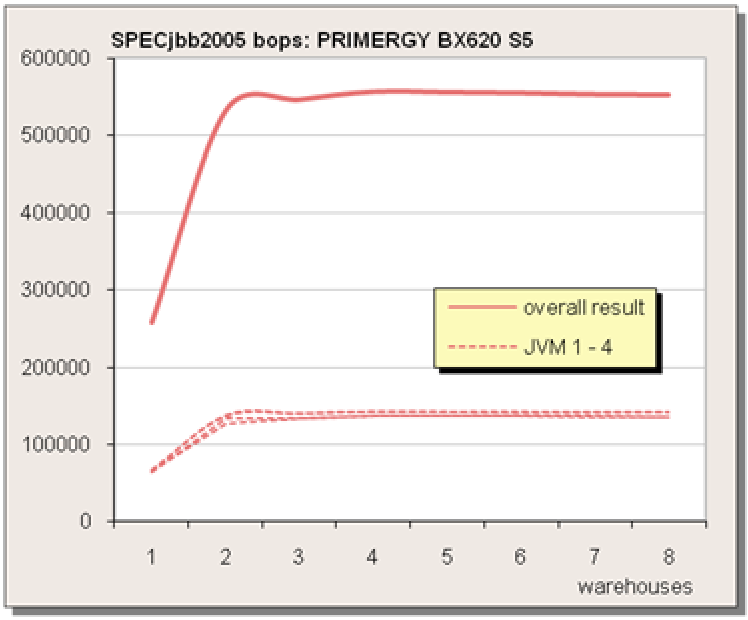
* SPEC®, SPECjbb® and the SPEC logo are registered trademarks of the Standard Performance Evaluation Corporation (SPEC).

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 May 2009 the PRIMERGY BX620 S5 was measured with a memory configuration with 24 GB PC3-10600R DDR3-SDRAM with two Xeon X5570 processors. The measurement was taken under Windows Server 2008 Enterprise x64 Edition SP2. Four instances of JRockit(R) 6 R28.0.0 from Oracle were used as JVM for the measurement. The benchmark result includes all the measuring values from 4 to 8 warehouses.



The following result was achieved:

SPECjbb2005 bops = 554740

SPECjbb2005 bops/JVM = 138685

Benchmark environment*

The SPECjbb2005 measurement was performed on a PRIMERGY BX620 S5 with the following hardware and software configuration:

Hardware	
Model	PRIMERGY BX620 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 SP2
JVM Version	Oracle JRockit(R) 6 P28.0.0 (build P28.0.0-29-114096-1.6.0_11-20090427-1759-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 BX620 S5 server has the following controllers to offer:

1. LSI 1064E SAS IME Storage Module Controller

   The controller does not have a controller cache and support is provided for RAID levels 0 and 1. The two offered version of the controller, with or without a rise option for an additional PCI Express card, do not differ in performance.

   MegaRAID SAS PCI Express ™ ROMB Controller

   The controller has a 256 MB controller cache and support is provided for RAID levels 0 and 1.This controller cache can be protected against power failure by an optional battery backup unit (BBU).

Various SAS hard disks can be connected to these controllers. Depending on the performance required, it is possible to select the appropriate disk subsystem. The PRIMERGY BX620 S5 server offers two hot-plug bays for 2½" hard disks. These two hard disks are adequate for the installation of an operating system. If greater storage capacity is needed, the server can be optionally extended with a PRIMERGY SX650 storage blade.

The following hard disks can be chosen for the PRIMERGY BX620 S5:

- 2½" SAS hard disks with a capacity of 36 GB, 73 GB and 146 GB (10 krpm)
- 2½" SAS hard disks with a capacity of 36 GB and 73 GB (15 krpm)

The 2½" drives have great advantages to offer, because they can reduce both power consumption and heat development and thus cut the costs for device cooling. The better space utilization and consequently the more compact design of the PRIMERGY BX620 S5 server only became at all possible through the use of 2½" hard disks.

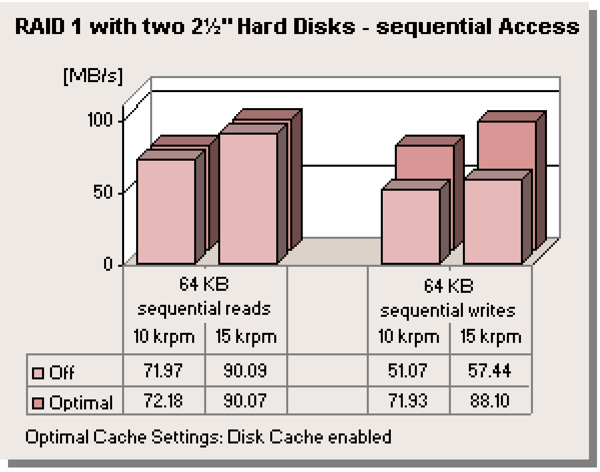
The RAID array defines the way in which data is treated as regards availability. How quickly the data is transferred in the respective RAID array context depends largely on the data throughput of the hard disks. The number of hard disks configured for the measurements in a RAID array was defined depending on the RAID level.

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. 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).

LSI 1064E SAS IME Storage Module

The throughputs of the hard disks in RAID 1 are compared below with different access patterns. In the measurements all hard disk types currently available for the PRIMERGY BX620 S5 server were analyzed. 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 diagram shows that as the rotational speed increases, the throughput for sequential reads and writes with a 64 KB block size rises. 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 is an increase in throughput of about 25%. The sequential read throughput achieves very good values and does not depend on the cache settings.



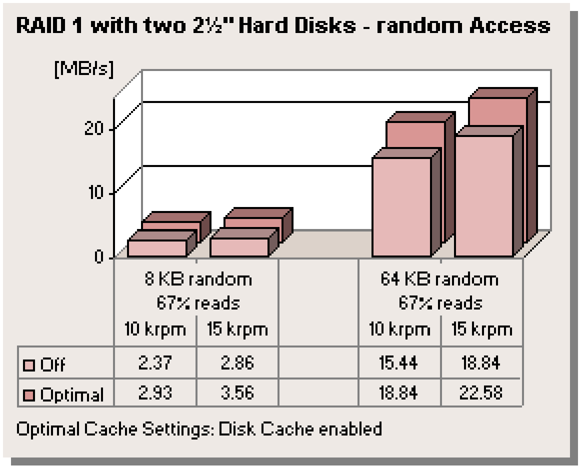
LSI 1064E SAS IME Storage Module

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 is an increase in throughput of about 22%.

As can be seen in the diagram, a relevant increase in throughput with sequential write can be achieved by enabling the disk cache: for hard disks with 10 krpm the throughput increases by about 41% and for hard disks with 15 krpm the throughput increases by about 53%.

The following diagram shows that the disk cache plays an important role in throughput improvement even for random access with 67% read share. For access with 8 KB and 64 KB blocks the throughput improves through the enabling of the disk cache by an average of 22%, regardless of whether a 10 krpm or a 15 krpm hard disk is used.

If you compare the throughput of the 10 krpm and 15 krpm hard disks, you can see that the throughput of the 15 krpm hard disks for random access with 8 KB and 64 KB blocks and enabled disk cache is about 21% higher than with the 10 krpm hard disks.

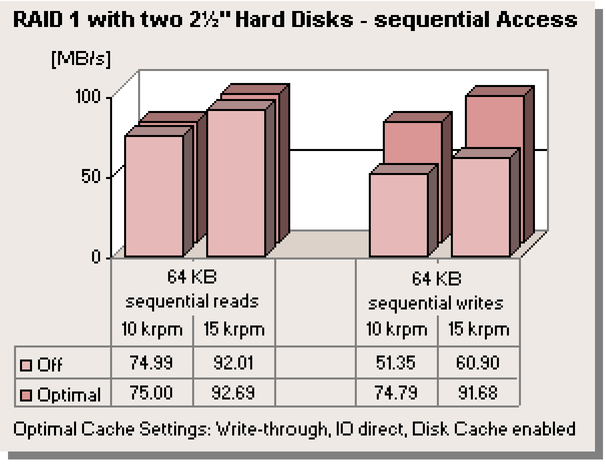


LSI 1064E SAS IME Storage Module

MegaRAID SAS PCI Express ™ ROMB

The throughputs of the hard disks in RAID 1 are compared below with different access patterns. In the measurements all hard disk types currently available for the PRIMERGY BX620 S5 server were analyzed.

The diagram shows that as the rotational speed increases, the throughput for sequential reads and writes with a 64 KB block size rises. 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 is an increase in throughput of about 23%. The sequential read throughput achieves very good values and does not depend on the cache settings.



MegaRAID SAS PCI Express ™ ROMB

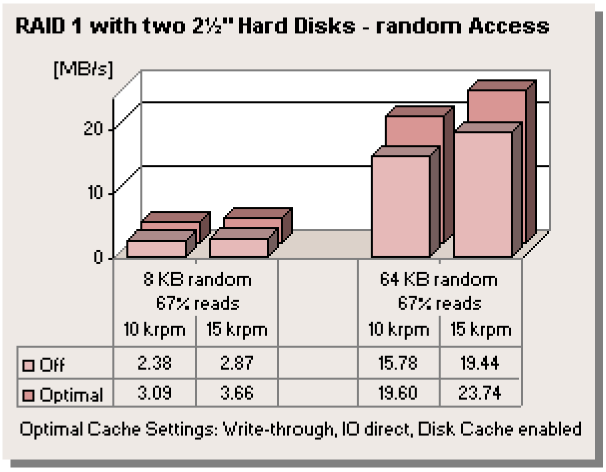
If for sequential write with enabled disk cache a hard disk with a rotational speed of 15 krpm is used instead of one with a speed of 10 krpm, the result is an increase in throughput of about 23%.

As can be seen in the diagram, a relevant increase in throughput with sequential write can be achieved by enabling the disk cache: for hard disks with 10 krpm the throughput increases by about 46% and for hard disks with 15 krpm the throughput increases by about 51%.

The following diagram shows that the disk cache plays an important role in throughput improvement even for random access with 67% read share.

For access with 8 KB blocks the throughput improves through the enabling of the disk cache by an average of 30% and 27%, depending on whether a 10 krpm or a 15 krpm hard disk is used. For access with 64 KB blocks the throughput improves through the enabling of the disk cache by an average of 24% and 22%, depending on whether a 10 krpm or a 15 krpm hard disk is used.

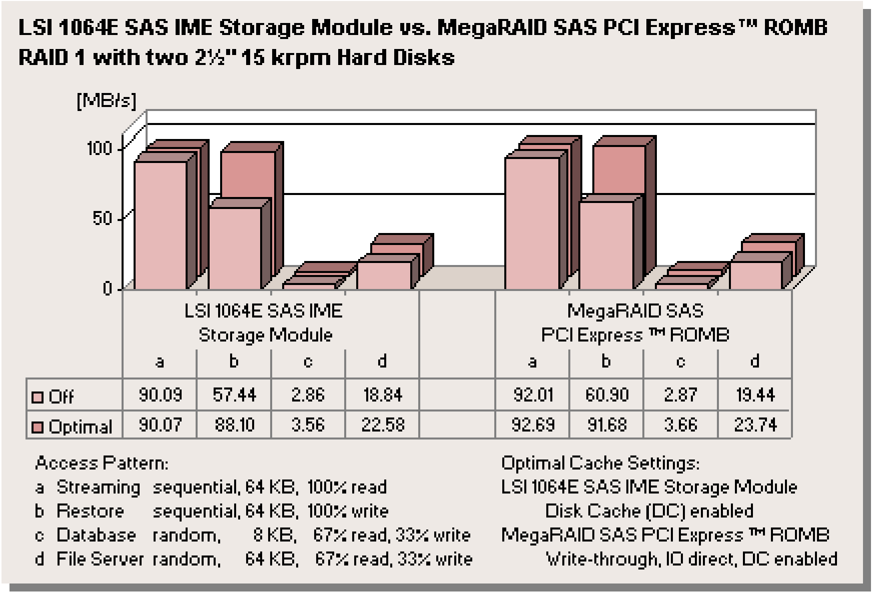
If you compare the throughput of the 10 krpm and 15 krpm hard disks, you can see that the throughput of the 15 krpm hard disks for random access with 8 KB and 64 KB blocks and enabled disk cache is about 20% higher than with the 10 krpm hard disks.



MegaRAID SAS PCI Express ™ ROMB

Controller comparison

The following comparison depicts the throughputs of the both 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 MegaRAID SAS PCI Express ™ ROMB controller shows a slightly higher performance than the LSI 1064E SAS IME Storage Module controller, as the throughput rates reached with sequential access are up to 6% more than with the LSI 1064E SAS IME Storage Module controller.

Almost the same performance differences exist with option-free access. The throughput rates reached with the MegaRAID SAS PCI Express ™ ROMB controller are up to 5% greater than with the LSI 1064E SAS IME Storage Module controller.

Both controllers achieve substantial performance increases with write operations as a result of optimal cache settings.

Conclusion

In conjunction with the PRIMERGY BX620 S5 server the both controllers support the basic RAID solutions RAID 0 and RAID 1. A RAID 1 ensures full data redundancy on the existing hard disks. Optionally, the server can be extended with a PRIMERGY SX650 storage blade.

The PRIMERGY BX620 S5 server offers 2½ SAS hard disks with rotational speeds of 10 krpm or 15 krpm. Depending on the performance required, a decision must be taken as to which rotational speed is to be used. The higher rotational speed of the 15 krpm hard disks is reflected in a higher throughput.

For maximum performance it is advisable, particularly when using a controller without a controller cache to enable the hard disk cache. Depending on the disk type used and access pattern, this leads to a performance increase of up to 53%. 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 BX620 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
SAS Storage Module Controller	Product: LSI 1064E SAS IME Driver name: lsi_sas.sys, Driver version: 1.29.03.00 Firmware version: 1.27.00.00 BIOS version: 06.26.00.00
SAS/RAID Storage Module Controller	Product: MegaRAID SAS PCI Express ™ ROMB Driver name: megasas.sys, Driver version: 2.13.0.64 Firmware package: 7.0.1-0066 Firmware version: 1.12.230-0066 BIOS version: MT33 Controller cache: 256 MB
Hard Disk SAS, 2½", 10 krpm	Seagate ST973402SS, 73 GB
Hard Disk SAS, 2½", 15 krpm	Seagate ST973451SS, 73 GB

* 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.

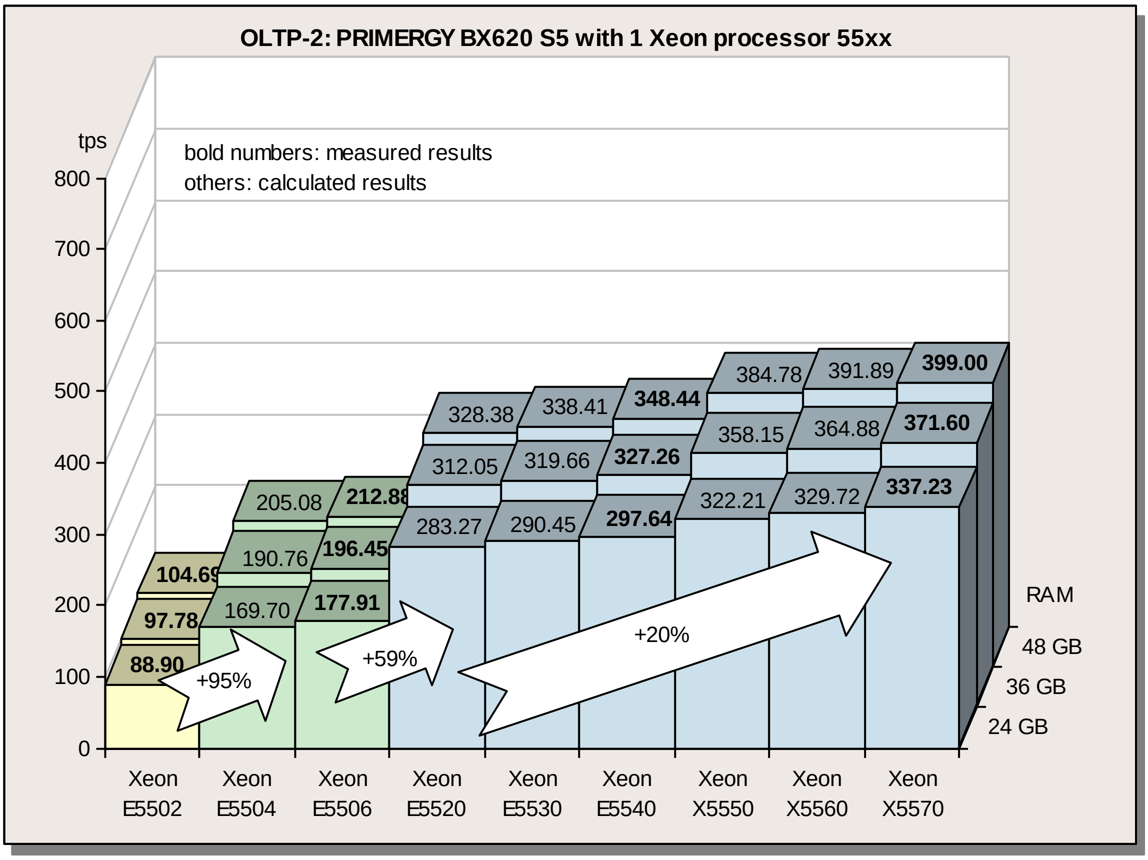
Benchmark results

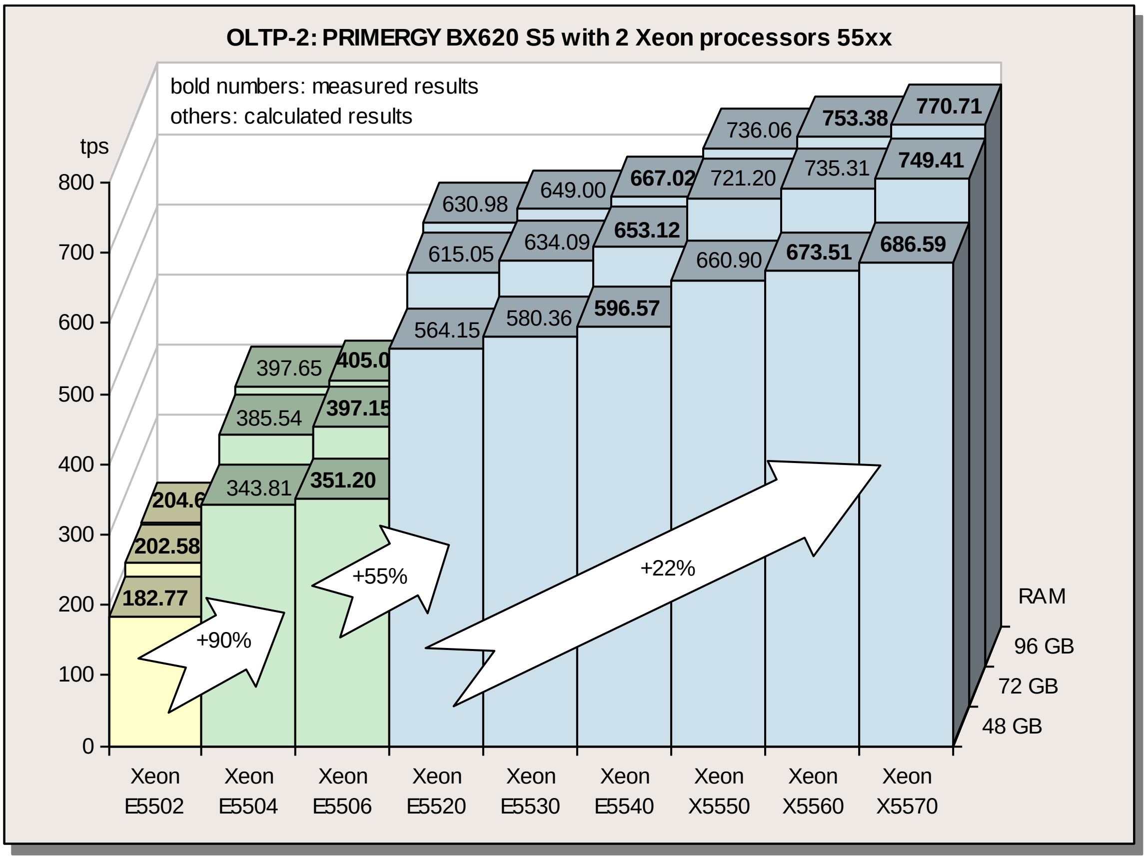
The PRIMERGY BX620 S5 Blade Server has been measured with Xeon Processors series 55xx at a memory size of 24 GB, 36 GB, 48 GB, 72 GB and 96 GB. All results were determined on the basis of the operating system Microsoft Windows Server 2008 Enterprise x64 Edition and the database SQL Server 2008 Enterprise 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 two dual-channel Fibre Channel controllers that were connected to a total of 465 hard disks within five FibreCAT CX500. 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 BX620 S5 with 6 memory modules at one processor and 12 memory modules at two processors memory access speed is 1067 MHz for processor types Xeon E5520, E5530, E5540, X5550, X5560 and X5570. With the use of Xeon E5502, E5504 und E5506 memory access speed is 800 MHz.

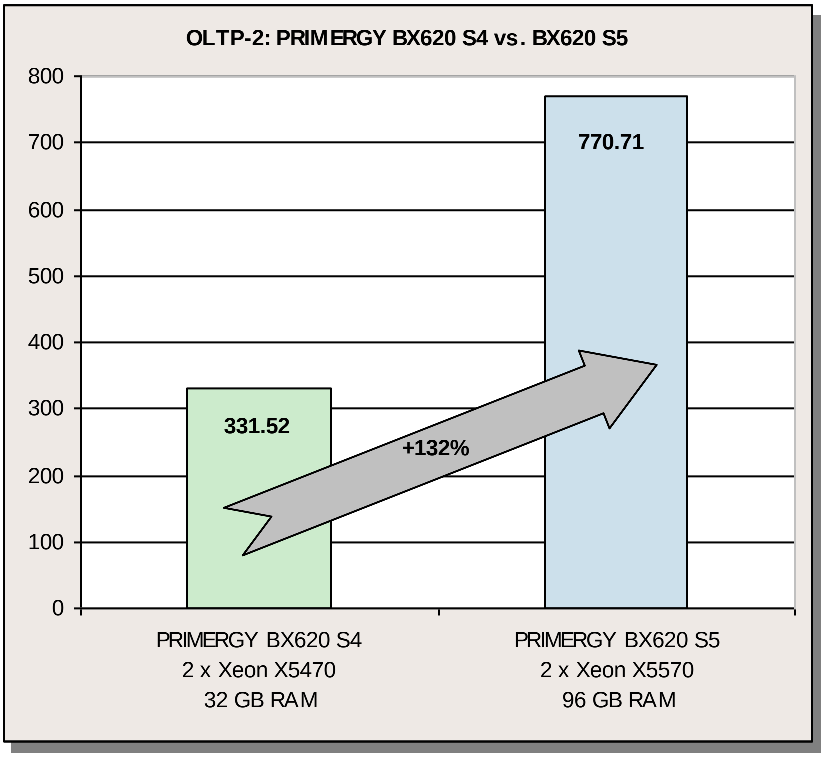
The next diagrams show the OLTP-2 performance data for PRIMERGY BX620 S5 separated in two groups with one and two processors Xeon series 55xx (E5502, E5504, E5506, E5520, E5530, E5540, X5550, X5560 and X5570). Performance values for Xeon L5530 are equivalent to Xeon E5530, values for Xeon L5520 are equivalent to Xeon E5520 and values for Xeon L5506 equivalent to Xeon E5506.

The largest scaling over all processor types with +90% to +95% is at Xeon E5502 to Xeon E5504. In this case number of processor-cores is doubled from two to 4. There is also a large increase of +55% to +59% at Xeon E5506 to Xeon E5520 by doubling the processor-cache from 4 MB to 8 MB and the use of Hyper-Threading. Finally the scaling across Xeon E5520 to Xeon X5570 is +20 to +22%. The memory scaling from 24 GB to 36 GB is about +10%, from 36 GB to 48 GB about +7%, from 48 GB to 72 GB about +10% and from 72 GB to 96 GB about +2.5%. This depends on the workload of the OLTP-2 benchmark and is not typical for all database applications.

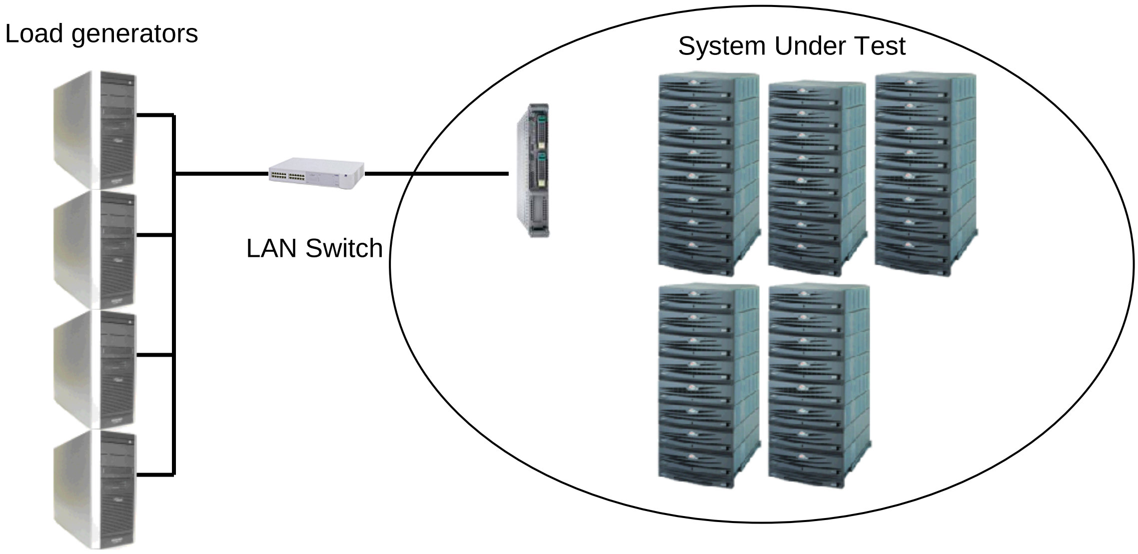

OLTP-2: PRIMERGY BX620 S5 with 1 Xeon processor 55xx
tps
800
700
600
500
400
300
200
100
0
bold numbers: measured results
others: calculated results
104.6
97.78
88.90
205.08
190.76
169.70
212.8
196.45
177.91
328.38
312.05
283.27
338.41
319.66
290.45
348.44
327.26
297.64
384.78
358.15
322.21
391.89
364.88
329.72
399.00
371.60
337.23
+95%
+59%
+20%
RAM
48 GB
36 GB
24 GB
Xeon E5502
Xeon E5504
Xeon E5506
Xeon E5520
Xeon E5530
Xeon E5540
Xeon X5550
Xeon X5560
Xeon X5570


OLTP-2: PRIMERGY BX620 S5 with 2 Xeon processors 55xx
bold numbers: measured results
others: calculated results
tps
800
700
600
500
400
300
200
100
0
204.6
202.58
182.77
397.65
385.54
343.81
405.0
397.15
351.20
630.98
615.05
564.15
649.00
634.09
580.36
667.02
653.12
596.57
736.06
721.20
660.90
753.38
735.31
673.51
770.71
749.41
686.59
+90%
+55%
+22%
RAM
96 GB
72 GB
48 GB
Xeon E5502
Xeon E5504
Xeon E5506
Xeon E5520
Xeon E5530
Xeon E5540
Xeon X5550
Xeon X5560
Xeon X5570

When comparing the PRIMERGY BX620 S5 and its predecessor the PRIMERGY BX620 S4 all in their highest performance configurations, a throughput increase of +132% is noted.



Benchmark environment*



System Under Test (SUT)	
Hardware	
Server Blade	PRIMERGY BX620 S5
Processor	Xeon E5502, E5504, E5506, E5520, E5530, E5540, X5550, X5560, X5570
Memory	Up to 12 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 BX620 S5: 1 x LSI SAS with 1064E 1x 2.5" 73GB 15K Fujitsu MBC2073RC RAID-0, OS 1 x dual-channel FC mezzanine card Emulex 1 x dual-channel FC controller Emulex LPe11002-M4 5 x FibreCAT CX500: 315 x Seagate 36 GB 15 krpm, RAID-0, data 135 x Seagate 73 GB 15 krpm, RAID-0, data 15 x Seagate 36 GB 15 krpm, RAID-0, log
Software	
Operating System	Windows Server 2008 Enterprise x64 Edition
Database	SQL Server 2008 Enterprise x64 Edition

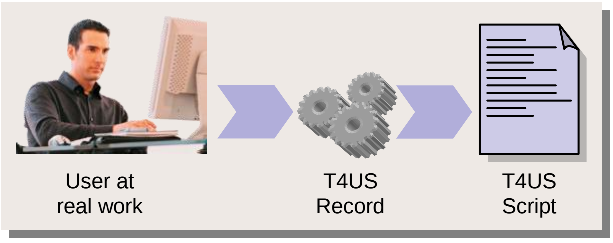
Load Generators	
Hardware	
Model	4 x PRIMERGY Econel 200
Processor	2 x Xeon 3.40 GHz, 2 MB L2 cache
Memory	2 GB DDR-SDRAM PC2700
Network Interface	1 x 1-GBit LAN (onboard)
Software	
Operating System	Windows Server 2003 Standard Edition SP1 (x86)
OLTP-2 Software	EGen version 1.6.0-1011

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

Terminal Server

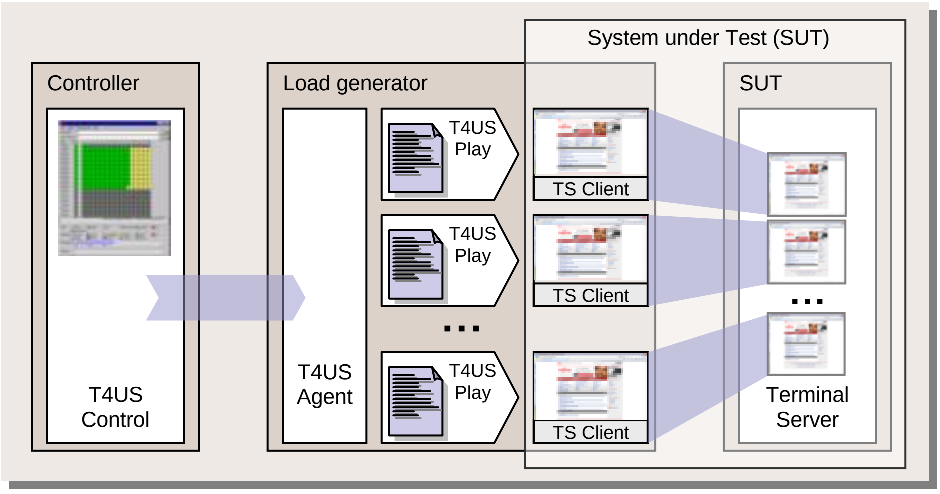
Benchmark description

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.

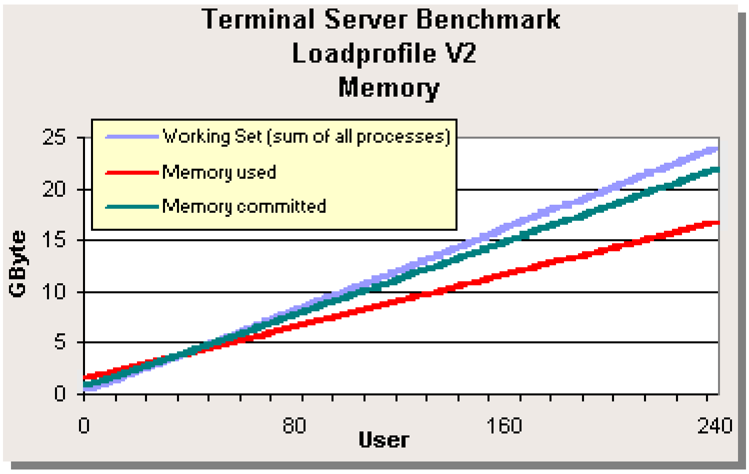
During the measurement the number of users working with Terminal Server is continuously increased. The Terminal Server response times are monitored by the T4US controller and compared with stored reference values which were determined from a previous reference measurement with only few users. If the 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.

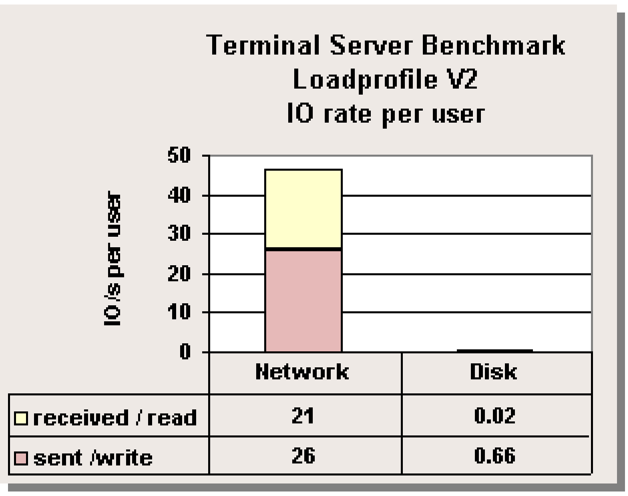
The new load profile V2 is characterized by the fact that a user that is to be simulated works with various Microsoft Office applications. In addition to the creation of a Microsoft Word document, a PowerPoint presentation is also designed. Calculations are also made on a new Excel spreadsheet. The number 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.

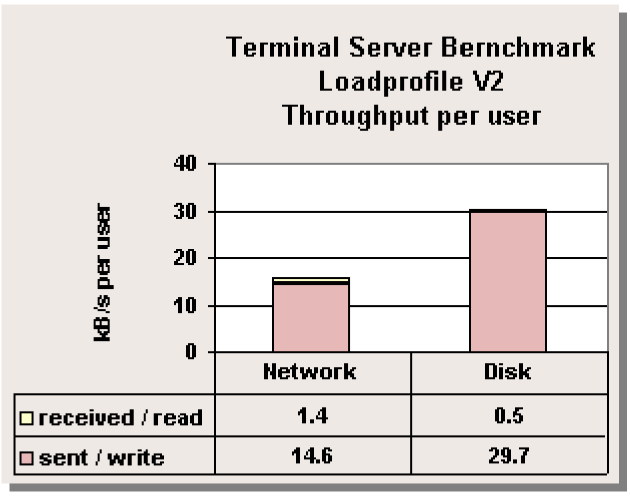
The memory requirements of the Terminal Server benchmark increase in proportion with the number of users and depend on the underlying operating system, there are in particular differences in the 32-bit and 64-bit operating systems. This aspect is handled in depth in the document Terminal Server Sizing 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 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 36 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, the data of the users and the pagefile are each stored on one partition of the Terminal Server – with the partitions being on a RAID-0 array of two hard disks. 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 96 GB the PRIMERGY BX620 S5 server blade provides a good platform for Terminal Server.

The memory access speed of the PRIMERGY BX620 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 BX620 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 BX620 S5 server blade 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)).

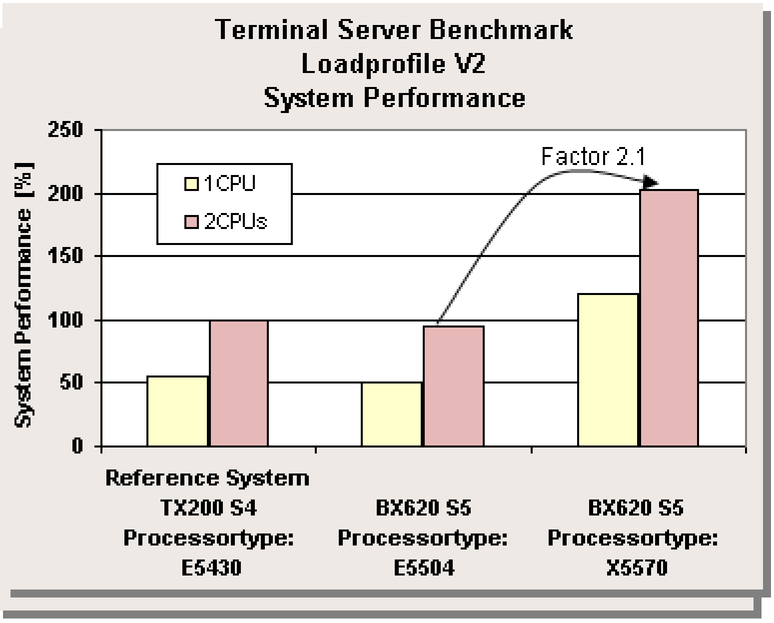
Other features of the processors measured

- 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 BX620 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 1.7-fold.

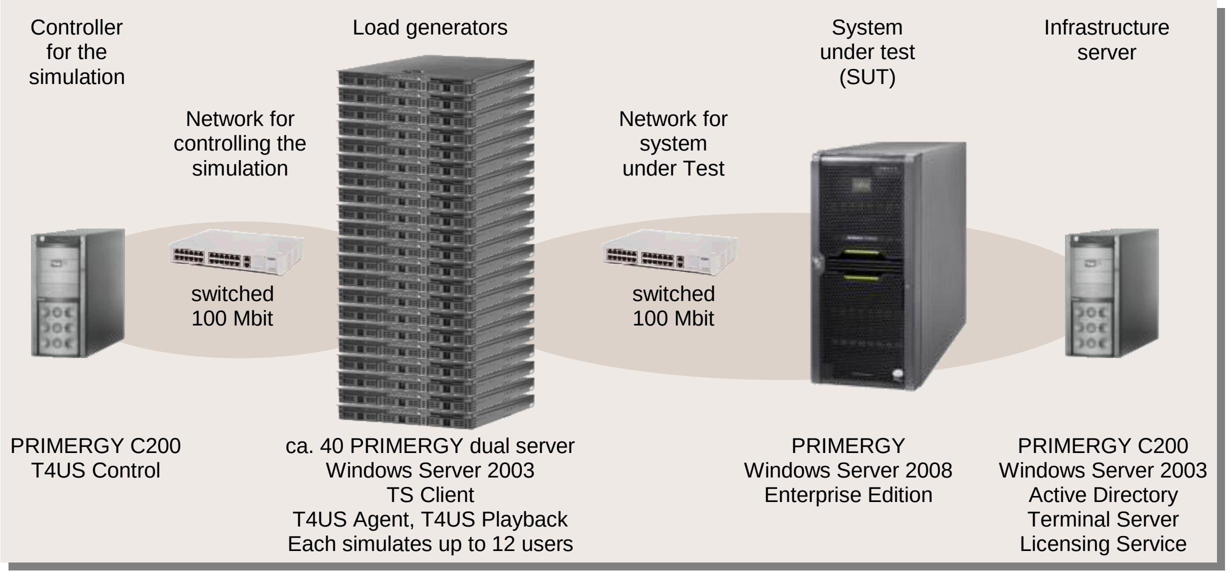
As regards system performance, the PRIMERGY BX620 S5 server blade with the Xeon E5504 is of the same magnitude as the reference system.

If the PRIMERGY BX620 S5 server blade 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 BX620 S5 server blade 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.

Hardware	
Model	PRIMERGY BX620S5 // PRIMERGY TX200S4
Processor	1-2 Xeon E5504 // 1-2 Xeon E5430 1-2 Xeon X5570
Memory	24 GB // 12 GB
Network Interface	2×1-GBit LAN Intel (onboard) // Broadcom
Disk Subsystem	SAS/RAID Storage Module 2 × SAS disks, 15 krpm, RAID 0
Software	
Operating System	Windows Server 2008 x64 Enterprise Edition
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

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.

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

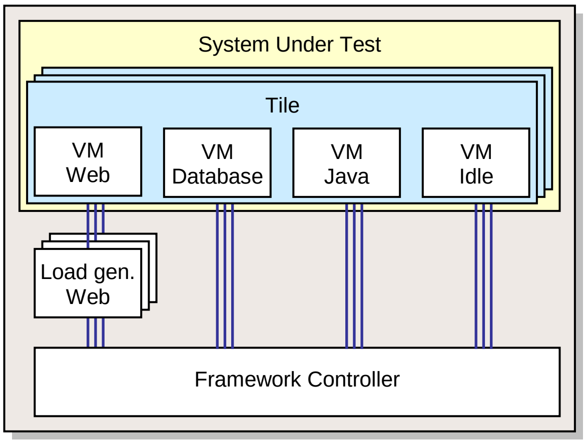
Benchmark description

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.

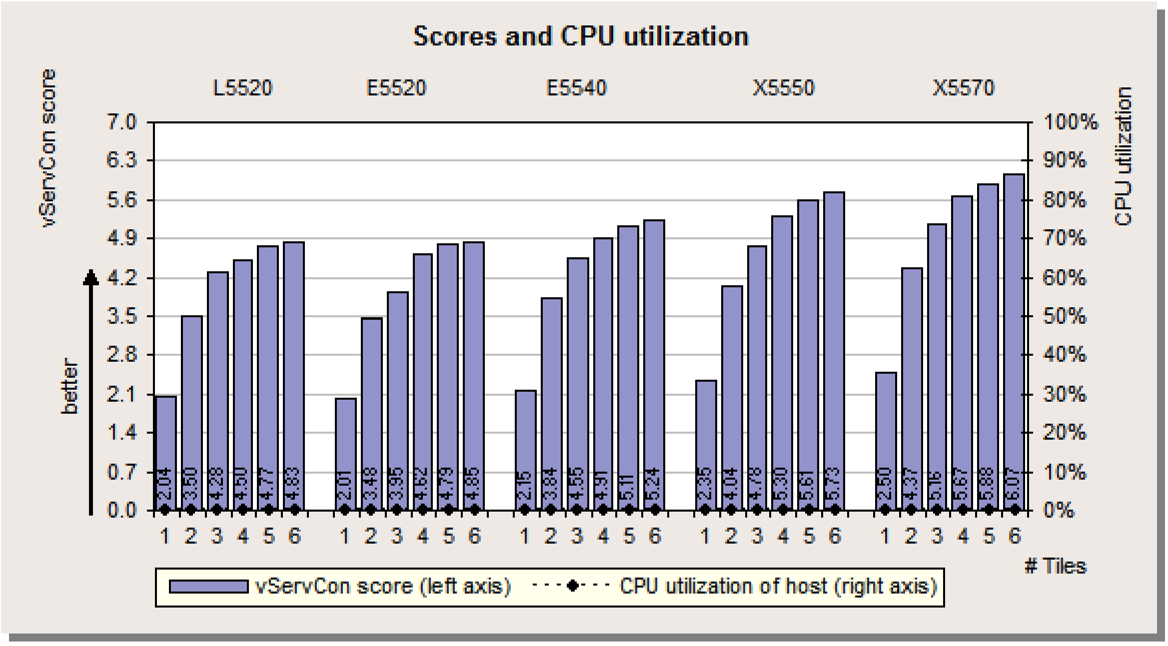
The score is intended to express a virtualization-specific system performance that can be achieved - right through to maximum utilization of the CPU resources - with a many VMs. In other words, the score would not be significant if a limitation were to occur during a vServCon measurement for an unnecessarily 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 BX620 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 BX620 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.

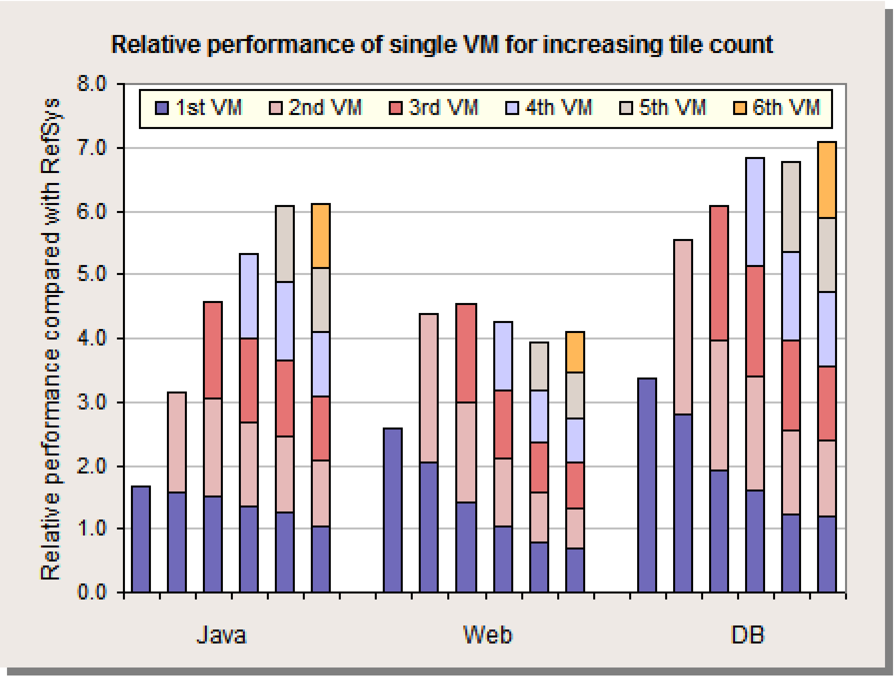
The high number of tiles is enabled particularly via Hyper-Threading whereby a physical processor core is divided into two logical cores so that 16 logical cores are available for the hypervisor. This standard feature thus generally increases the virtualization performance of a system.

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.

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.

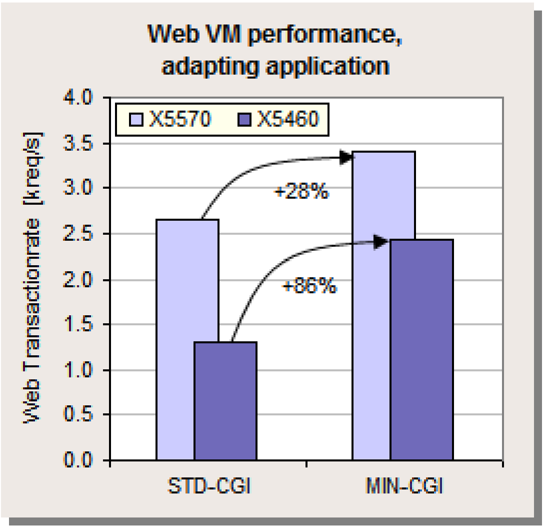
If you want to run applications in virtual machines at maximum performance, it is worth looking at the application profiles that make higher demands of a virtualization solution more closely. These include application scenarios like web server that are a great drain on memory management.

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.

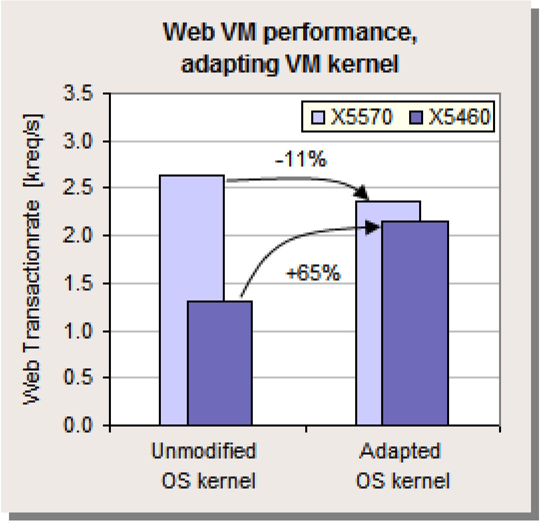
The diagram clearly shows the advance of processor technology in the virtualization sector. A PRIMERGY BX620 S5 (with Xeon X5570) is compared with the previous system. Whereas a Web-VM with the previous system could achieve an increase in performance of 86% as a result of optimization in the application scenario, this is only 28% 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.

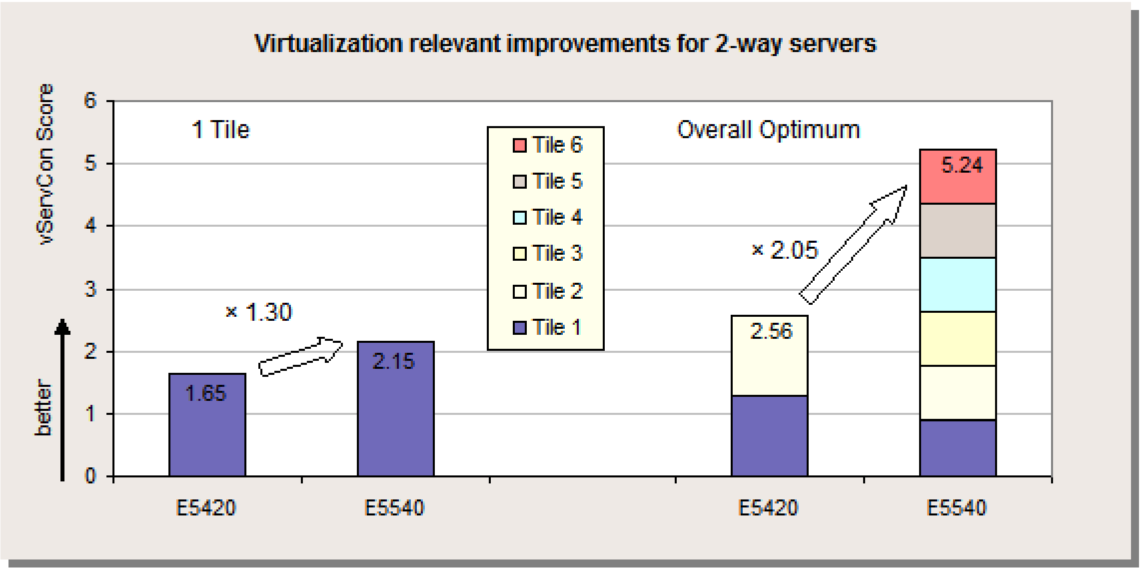
With the previous generations of the PRIMERGY servers the use of a kernel modified for virtualization for the web server VM resulted in a large increase in performance. The diagram compares a PRIMERGY BX620 S5 (with Xeon X5570) with the previous system for two kernels under the load profile STD-CGI. 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 BX620 S5 with 2 × Xeon E5540 (2.53 GHz); both have the same number of physical cores.

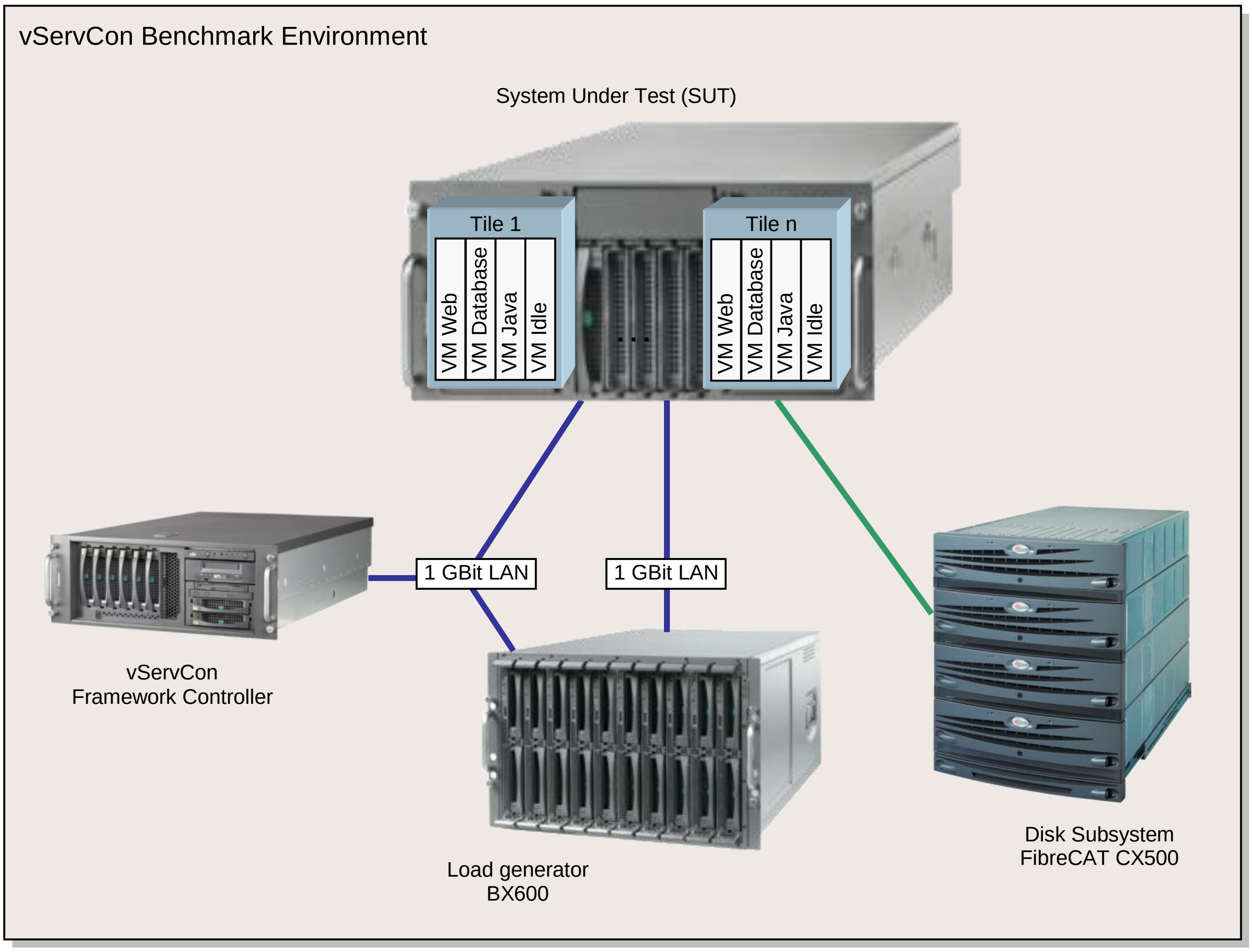


One sees an increase of the vServCon score by a factor of 2.05. The one reason is the performance increase that can be achieved for an individual VM (factor 1.30, see score with one tile). The other reason is that more tiles are possible. However, this is not a true-value tripling (from two to six tiles) 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:



SUT Hardware	
Model	PRIMERGY BX620 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 LPe12002
SUT Software	
Operating system	Hypervisor VMware ESX Server.
Version	Version 4.0.0 build 157368
BIOS	Version 3A28; 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 BX620 S5

Data sheet
http://docs.ts.fujitsu.com/dl.aspx?id=ef5b1c88-78b4-4c00-a923-488be5866e1a

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

PRIMERGY SX650

Data sheet
http://docs.ts.fujitsu.com/dl.aspx?id=095dbb1b-5aa2-4064-bacc-1afff784c066

Performance Report PRIMERGY SX650
http://docs.ts.fujitsu.com/dl.aspx?id=9384e31a-c21b-44b8-a3a2-5c7a5902491b

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

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

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