

White Paper Fujitsu PRIMERGY Servers Performance Report PRIMERGY CX250 S2

This document contains a summary of the benchmarks executed for the PRIMERGY CX250 S2.

The PRIMERGY CX250 S2 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	2
Technical data	3
SPECcpu2006	
Disk I/O	10
vServCon	16
STREAM	
LINPACK	25
Literature	28
Contact	29

Document history

Version 1.0

New:

- Technical data
- SPECcpu2006

Measurements with Xeon E5-2600 processor series v2

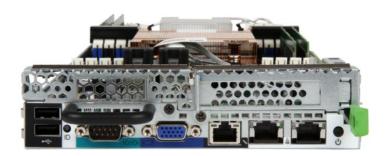
Disk I/O

Measurements with "LSI SW RAID on Intel C600 (Onboard SATA)" and "Raid SAS HBA Mezz Card 6Gb" controllers

- vServCon
 - Results for Xeon E5-2600 v2 processor series
- STREAM
 - Measurements with Xeon E5-2600 v2 processor series
- LINPACK
 - Measurements with Xeon E5-2600 v2 processor series

Technical data

PRIMERGY CX250 S2



Decimal prefixes according to the SI standard are used for measurement units in this white paper (e.g. 1 GB = 10^9 bytes). In contrast, these prefixes should be interpreted as binary prefixes (e.g. 1 GB = 2^{30} bytes) for the capacities of caches and storage modules. Separate reference will be made to any further exceptions where applicable.

Model	PRIMERGY CX250 S2
Form factor	Server node
Chipset	Intel C600 series
Number of sockets	2
Number of processors orderable	2
Processor type	Intel® Xeon® series E5-2600 v2
Number of memory slots	16 (8 per processor)
Maximum memory configuration	1024 GB
Onboard LAN controller	2 x 1 Gbit/s
Onboard HDD controller	SATA controller for up to 6 × 2.5" or 3 × 3.5" SATA HDDs
PCI slots	2 x PCI-Express 3.0 x16

The processor frequency specified in the following table is always at least achieved given full utilization. Processors with Turbo Boost Technology 2.0 additionally permit automatically regulated, dynamic overclocking. The overclocking rate depends on the utilization of the processor and its ambient conditions. As far as utilization is concerned, the number of cores subject to utilization as well as the type and strength of core utilization play a role. Added to these as influencing factors are the strength of the heating, the level of the ambient temperature and the heat dissipation options. As a result of overclocking it is even possible to exceed the thermal design power of the processor for short periods of time.

How much a processor benefits from the Turbo mode in an individual case depends on the respective application and can in some application scenarios even differ from processor example to processor example.

Processors (since			•	0.01		••			
Processor	Cores	Threads	Cache	QPI Speed	Processor Frequency	Max. Turbo Frequency at full load	Max. Turbo Frequency	Max. Memory Frequency	TDP
			[MB]	[GT/s]	[Ghz]	[Ghz]	[Ghz]	[MHz]	[Watt]
Xeon E5-2603 v2	4	4	10	6.40	1.80	n/a	n/a	1333	80
Xeon E5-2609 v2	4	4	10	6.40	2.50	n/a	n/a	1333	80
Xeon E5-2637 v2	4	8	15	8.00	3.50	3.60	3.80	1866	130
Xeon E5-2620 v2	6	12	15	7.20	2.10	2.40	2.60	1600	80
Xeon E5-2630Lv2	6	12	15	7.20	2.40	2.60	2.80	1600	60
Xeon E5-2630 v2	6	12	15	7.20	2.60	2.90	3.10	1600	80
Xeon E5-2643 v2	6	12	25	8.00	3.50	3.60	3.80	1866	130
Xeon E5-2640 v2	8	16	20	7.20	2.00	2.30	2.50	1600	95
Xeon E5-2650 v2	8	16	20	8.00	2.60	3.00	3.40	1866	95
Xeon E5-2667 v2	8	16	25	8.00	3.30	3.60	4.00	1866	130
Xeon E5-2650Lv2	10	20	25	7.20	1.70	1.90	2.10	1600	70
Xeon E5-2660 v2	10	20	25	8.00	2.20	2.60	3.00	1866	95
Xeon E5-2670 v2	10	20	25	8.00	2.50	2.90	3.30	1866	115
Xeon E5-2680 v2	10	20	25	8.00	2.80	3.10	3.60	1866	115
Xeon E5-2690 v2	10	20	25	8.00	3.00	3.30	3.60	1866	130
Xeon E5-2695 v2	12	24	30	8.00	2.40	2.80	3.20	1866	115
Xeon E5-2697 v2	12	24	30	8.00	2.70	3.00	3.50	1866	130

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

Detailed technical information is available in the <u>data sheet PRIMERGY CX250 S2</u>.

SPECcpu2006

Benchmark description

SPECcpu2006 is a benchmark which measures the system efficiency with integer and floating-point operations. It consists of an integer test suite (SPECint2006) containing 12 applications and a floating-point test suite (SPECfp2006) containing 17 applications. Both test suites are extremely computing-intensive and concentrate on the CPU and the memory. Other components, such as Disk I/O and network, are not measured by this benchmark.

SPECcpu2006 is not tied to a special operating system. The benchmark is available as source code and is compiled before the actual measurement. The used compiler version and their optimization settings also affect the measurement result.

SPECcpu2006 contains two different performance measurement methods: the first method (SPECint2006 or SPECfp2006) determines the time which is required to process single task. The second method (SPECint_rate2006 or SPECfp_rate2006) determines the throughput, i.e. the number of tasks that can be handled in parallel. Both methods are also divided into two measurement runs, "base" and "peak" which differ in the use of compiler optimization. When publishing the results the base values are always used; the peak values are optional.

Benchmark	Arithmetics	Туре	Compiler optimization	Measurement result	Application
SPECint2006	integer	peak	aggressive	Spood	aingle threeded
SPECint_base2006	integer	base	conservative	Speed	single-threaded
SPECint_rate2006	integer	peak	aggressive	Throughput	multi-threaded
SPECint_rate_base2006	integer	base	conservative	- Throughput	muiti-timeaded
SPECfp2006	floating point	peak	aggressive	Chood	aingle threeded
SPECfp_base2006	floating point	base	conservative	Speed	single-threaded
SPECfp_rate2006	floating point	peak	aggressive	Throughput	multi-threaded
SPECfp_rate_base2006	floating point	base	conservative	Throughput	multi-timeaded

The measurement results are the geometric average from normalized ratio values which have been determined for individual benchmarks. The geometric average - in contrast to the arithmetic average - means that there is a weighting in favour of the lower individual results. Normalized means that the measurement is how fast is the test system compared to a reference system. Value "1" was defined for the SPECint_base2006-, SPECint_rate_base2006, SPECfp_base2006 and SPECfp_rate_base2006 results of the reference system. For example, a SPECint_base2006 value of 2 means that the measuring system has handled this benchmark twice as fast as the reference system. A SPECfp_rate_base2006 value of 4 means that the measuring system has handled this benchmark some 4/[# base copies] times faster than the reference system. "# base copies" specify how many parallel instances of the benchmark have been executed.

Not every SPECcpu2006 measurement is submitted by us for publication at SPEC. This is why the SPEC web pages do not have every result. As we archive the log files for all measurements, we can prove the correct implementation of the measurements at any time.

Benchmark environment

System Under Test ((SUT)
Hardware	
Enclosure	PRIMERGY CX400 S2
Model	PRIMERGY CX250 S2
Processor	Xeon E5-2600 v2 processor series
Memory	Xeon E5-2620 v2, E5-2630 v2: 16 x 8GB (1x8GB) 2Rx8 DDR3-1866 R ECC All others: 8 x 16GB (1x16GB) 2Rx4 DDR3-1866 R ECC
Software	
BIOS settings	Energy Performance = Performance SPECint_base2006, SPECint2006, SPECfp_base2006, SPECfp2006: Utilization Profile = Unbalanced
Operating system	Red Hat Enterprise Linux Server release 6.4
Operating system settings	echo always > /sys/kernel/mm/redhat_transparent_hugepage/enabled
Compiler	Intel C++/Fortran Compiler 14.0

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

Benchmark results

In terms of processors the benchmark result depends primarily on the size of the processor cache, the support for Hyper-Threading, the number of processor cores and on the processor frequency. In the case of processors with Turbo mode the number of cores, which are loaded by the benchmark, determines the maximum processor frequency that can be achieved. In the case of single-threaded benchmarks, which largely load one core only, the maximum processor frequency that can be achieved is higher than with multi-threaded benchmarks (see the processor table in the section "Technical Data").

Processor	Number of processors	SPECint_base2006	SPECint2006	SPECint_rate_base2006	SPECint_rate2006
Xeon E5-2603 v2	2				
Xeon E5-2609 v2	2				
Xeon E5-2637 v2	2	56.2	59.8	409	426
Xeon E5-2620 v2	2			413	429
Xeon E5-2630Lv2	2				
Xeon E5-2630 v2	2	47.3	50.0	486	503
Xeon E5-2643 v2	2				
Xeon E5-2640 v2	2				
Xeon E5-2650 v2	2	52.6	57.0	651	675
Xeon E5-2667 v2	2				
Xeon E5-2650Lv2	2	34.7	37.0	540	561
Xeon E5-2660 v2	2	48.0	51.8	709	734
Xeon E5-2670 v2	2				
Xeon E5-2680 v2	2				
Xeon E5-2690 v2	2				
Xeon E5-2695 v2	2	50.7	54.9		
Xeon E5-2697 v2	2	55.3	60.0		

Processor	Number of processors	SPECfp_base2006	SPECfp2006	SPECfp_rate_base2006	SPECfp_rate2006
Xeon E5-2603 v2	2				
Xeon E5-2609 v2	2				
Xeon E5-2637 v2	2				
Xeon E5-2620 v2	2			378	386
Xeon E5-2630Lv2	2				
Xeon E5-2630 v2	2	83.0	86.0	423	433
Xeon E5-2643 v2	2				
Xeon E5-2640 v2	2				
Xeon E5-2650 v2	2				
Xeon E5-2667 v2	2				
Xeon E5-2650Lv2	2	66.6	69.3	454	464
Xeon E5-2660 v2	2			558	572
Xeon E5-2670 v2	2				
Xeon E5-2680 v2	2				
Xeon E5-2690 v2	2				
Xeon E5-2695 v2	2	91.1	95.5		
Xeon E5-2697 v2	2			654	673

Disk I/O

Benchmark description

Performance measurements of disk subsystems for PRIMERGY servers are used to assess their performance and enable a comparison of the different storage connections for PRIMERGY servers. As standard, these performance measurements are carried out with a defined measurement method, which models the hard disk accesses of real application scenarios on the basis of specifications.

The essential specifications are:

- Share of random accesses / sequential accesses
- Share of read / write access types
- Block size (kB)
- Number of parallel accesses (# of outstanding I/Os)

A given value combination of these specifications is known as "load profile". The following five standard load profiles can be allocated to typical application scenarios:

Standard load	Access	Type of access		Block size	Application	
profile		read	write	[kB]		
File copy	random	50%	50%	64	Copying of files	
File server	random	67%	33%	64	File server	
Database	random	67%	33%	8	Database (data transfer) Mail server	
Streaming	sequential	100%	0%	64	Database (log file), Data backup; Video streaming (partial)	
Restore	sequential	0%	100%	64	Restoring of files	

In order to model applications that access in parallel with a different load intensity, the "# of Outstanding I/Os" is increased, starting with 1, 3, 8 and going up to 512 (from 8 onwards in increments to the power of two).

The measurements of this document are based on these standard load profiles.

The main results of a measurement are:

■ Throughput [MB/s] Throughput in megabytes per second

Transactions [IO/s] Transaction rate in I/O operations per second

Latency [ms] Average response time in ms

The data throughput has established itself as the normal measurement variable for sequential load profiles, whereas the measurement variable "transaction rate" is mostly used for random load profiles with their small block sizes. Data throughput and transaction rate are directly proportional to each other and can be transferred to each other according to the formula

Data throughput [MB/s]	= Transaction rate [IO/s] × Block size [MB]
Transaction rate [IO/s]	= Data throughput [MB/s] / Block size [MB]

This section specifies hard disk capacities on a basis of 10 (1 TB = 10^{12} bytes) while all other capacities, file sizes, block sizes and throughputs are specified on a basis of 2 (1 MB/s = 2^{20} bytes/s).

All the details of the measurement method and the basics of disk I/O performance are described in the white paper "Basics of Disk I/O Performance".

Benchmark environment

All the measurement results discussed in this chapter were determined using the hardware and software components listed below:

System Under Test (S	SUT)
Hardware	
Controller	1 × "LSI SW RAID on Intel C600 (Onboard SATA)" 1 × "Raid SAS HBA Mezz Card 6Gb"
Drive	6 × EP HDD SAS 6 Gbit/s 2.5" 15000 rpm 146 GB 6 × EP SSD SAS 6 Gbit/s 2.5" 200 GB MLC 6 × EP SSD SATA 6 Gbit/s 2.5" 200 GB MLC 4 × BC HDD SATA 6 Gbit/s 2.5" 7200 rpm 1 TB 2 × BC HDD SATA 6 Gbit/s 3.5" 7200 rpm 3 TB
Software	
Operating system	Microsoft Windows Server 2008 Enterprise x64 Edition SP2 Microsoft Windows Server 2012 Standard
Administration software	ServerView RAID Manager 5.7.2
Initialization of RAID arrays	RAID arrays are initialized before the measurement with an elementary block size of 64 kB ("stripe size")
File system	NTFS
Measuring tool	Iometer 2006.07.27
Measurement data	Measurement files of 32 GB with 1 – 8 hard disks; 64 GB with 9 – 16 hard disks; 128 GB with 17 or more hard disks

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

Benchmark results

The results presented here are designed to help you choose the right solution from the various configuration options of the PRIMERGY CX250 S2 in the light of disk-I/O performance. The selection of suitable components and the right settings of their parameters is important here. These two aspects should therefore be dealt with as preparation for the discussion of the performance values.

Components

The hard disks are the first essential component. If there is a reference below to "hard disks", this is meant as the generic term for HDDs ("hard disk drives", in other words conventional hard disks) and SSDs ("solid state drives", i.e. non-volatile electronic storage media). When selecting the type of hard disk and number of hard disks you can move the weighting in the direction of storage capacity, performance, security or price. In order to enable a pre-selection of the hard disk types – depending on the required weighting – the hard disk types for PRIMERGY servers are divided into three classes:

"Economic" (ECO): low-priced hard disks"Business Critical" (BC): very failsafe hard disks

"Enterprise" (EP): very failsafe and very high-performance hard disks

The following table is a list of the hard disk types that have been available for the PRIMERGY CX250 S2 since system release.

Drive class	Data medium type	Interface	Form factor	krpm
Business Critical	HDD	SATA 6G	2.5"	7.2
Business Critical	HDD	SATA 6G	3.5"	7.2
Business Critical	HDD	SAS 6G	2.5"	7.2
Business Critical	HDD	SAS 6G	3.5"	7.2
Enterprise	HDD	SAS 6G	2.5"	10, 15
Enterprise	SSD	SATA 6G	2.5"	-
Enterprise	SSD	SAS 6G	2.5"	-

Mixed drive configurations of SAS and SATA hard disks are not possible.

The SATA-HDDs offer high capacities right up into the terabyte range at a very low cost. The SAS-HDDs have shorter access times and achieve higher throughputs due to the higher rotational speed of the SAS-HDDs (in comparison with the SATA-HDDs). SAS-HDDs with a rotational speed of 15 krpm have better access times and throughputs than comparable HDDs with a rotational speed of 10 krpm. The 6G interface has in the meantime established itself as the standard among the SAS-HDDs.

Of all the hard disk types SSDs offer on the one hand by far the highest transaction rates for random load profiles, and on the other hand the shortest access times. In return, however, the price per gigabyte of storage capacity is substantially higher.

More hard disks per system are possible as a result of using 2.5" hard disks instead of 3.5" hard disks. Consequently, the load that each individual hard disk has to overcome decreases and the maximum overall performance of the system increases.

More detailed performance statements about hard disk types are available in the white paper "Single Disk Performance".

The maximum number of hard disks in the system depends on the system configuration. The following table lists the essential cases.

Form factor	Interface	Connection type	Number of PCle controllers	Maximum number of hard disks
2.5"	SATA 3G	direct	0	4*)
3.5"	SATA 3G	direct	0	3
2.5"	SATA 6G, SAS 6G	direct	1	6
3.5"	SATA 6G, SAS 6G	direct	1	3

*) If no PRIMERGY RAID-Management is used, six hard disks are possible.

After the hard disks the RAID controller is the second performance-determining key component. In the case of these controllers the "modular RAID" concept of the PRIMERGY servers offers a plethora of options to meet the various requirements of a wide range of different application scenarios.

The following table summarizes the most important features of the available RAID controllers of the PRIMERGY CX250 S2. A short alias is specified here for each controller, which is used in the subsequent list of the performance values.

Controller name	Alias	Cache	Supported interfaces		Max. # disks in the system	RAID levels in the system	BBU/ FBU
LSI SW RAID on Intel C600 (Onboard SATA)	Patsburg A	-	SATA 3G	-	4 × 2.5" 3 × 3.5"	0, 1, 10	-/-
Raid SAS HBA Mezz Card 6Gb	LSI2108	512 MB	SATA 3G/6G SAS 3G/6G	PCIe 2.0 x8	6 × 2.5" 3 × 3.5"	0, 1, 5, 6, 10, 50, 60	√/-

The onboard RAID controller is implemented in the chipset Intel C600 on the motherboard of the server and uses the CPU of the server for the RAID functionality. This controller is a simple solution that does not require a PCIe slot.

System-specific interfaces

The interfaces of a controller to the motherboard and to the hard disks have in each case specific limits for data throughput. These limits are listed in the following table. The minimum of these two values is a definite limit, which cannot be exceeded. This value is highlighted in bold in the following table.

Controller alias	Effective in the	Connection				
	# Disk channels	Limit for throughput of disk interface	PCle version	PCIe width	Limit for throughput of PCIe interface	via expander
Patsburg A	4 × SATA 3G	973 MB/s	-	-	-	-
LSI2108	6 × SAS 6G	2918 MB/s	2.0	x8	3433 MB/s	-

More details about the RAID controllers of the PRIMERGY systems are available in the white paper "RAID Controller Performance".

Settings

In most cases, the cache of the hard disks has a great influence on disk-I/O performance. This is particular valid for HDDs. It is frequently regarded as a security problem in case of power failure and is thus switched off. On the other hand, it was integrated by hard disk manufacturers for the good reason of increasing the write performance. For performance reasons it is therefore advisable to enable the hard disk cache. This is particular valid for SATA-HDDs. The performance can as a result increase more than tenfold for specific access patterns and hard disk types. More information about the performance impact of the hard disk cache is available in the document "Single Disk Performance". To prevent data loss in case of power failure you are recommended to equip the system with a UPS.

In the case of controllers with a cache there are several parameters that can be set. The optimal settings can depend on the RAID level, the application scenario and the type of data medium. In the case of RAID levels 5 and 6 in particular (and the more complex RAID level combinations 50 and 60) it is obligatory to enable the controller cache for application scenarios with write share. If the controller cache is enabled, the data temporarily stored in the cache should be safeguarded against loss in case of power failure. Suitable accessories are available for this purpose (e.g. a BBU or FBU).

For the purpose of easy and reliable handling of the settings for RAID controllers and hard disks it is advisable to use the RAID-Manager software "ServerView RAID" that is supplied for PRIMERGY servers. All the cache settings for controllers and hard disks can usually be made en bloc – specifically for the application – by using the pre-defined modi "Performance" or "Data Protection". The "Performance" mode ensures the best possible performance settings for the majority of the application scenarios.

More information about the setting options of the controller cache is available in the white paper "RAID Controller Performance".

Performance values

In general, disk-I/O performance of a RAID array depends on the type and number of hard disks, on the RAID level and on the RAID controller. If the limits of the <u>system-specific interfaces</u> are not exceeded, the statements on disk-I/O performance are therefore valid for all PRIMERGY systems. This is why all the performance statements of the document "<u>RAID Controller Performance</u>" also apply for the PRIMERGY CX250 S2 if the configurations measured there are also supported by this system.

The performance values of the PRIMERGY CX250 S2 are listed in table form below, specifically for different RAID levels, access types and block sizes. Substantially different configuration versions are dealt with separately.

The performance values in the following tables use the established measurement variables, as already mentioned in the subsection <u>Benchmark description</u>. Thus, transaction rate is specified for random accesses and data throughput for sequential accesses. To avoid any confusion among the measurement units the tables have been separated for the two access types.

The table cells contain the maximum achievable values. This has three implications: On the one hand hard disks with optimal performance were used (the components used are described in more detail in the subsection Benchmark environment). Furthermore, cache settings of controllers and hard disks, which are optimal for the respective access scenario and the RAID level, are used as a basis. And ultimately each value is the maximum value for the entire load intensity range (# of outstanding I/Os).

In order to also visualize the numerical values each table cell is highlighted with a horizontal bar, the length of which is proportional to the numerical value in the table cell. All bars shown in the same scale of length have the same color. In other words, a visual comparison only makes sense for table cells with the same colored bars.

Since the horizontal bars in the table cells depict the maximum achievable performance values, they are shown by the color getting lighter as you move from left to right. The light shade of color at the right end of the bar tells you that the value is a maximum value and can only be achieved under optimal prerequisites. The darker the shade becomes as you move to the left, the more frequently it will be possible to achieve the corresponding value in practice.

Random accesses (performance values in IO/s):

Configuration version				rel	dom ks id	dom Sks Id	lom ks ld	lom cks id	
RAID Controller	Hard disk type	Form factor	#Disks	RAID level	HDDs random 8 kB blocks 67% read [IO/s]	HDDs random 64 kB blocks 67% read [IO/s]	SSDs random 8 kB blocks 67% read [IO/s]	SSDs random 64 kB blocks 67% read [IO/s]	
	DC CATALIDD		2	1	550	447	17760	3951	
Patsburg A	BC SATA HDD EP SATA SSD	2.5"	4	0	1073	583	36497	8249	
	LI SAIASSD		4	10	828	446	28683	6665	
Patsburg A	BC SATA HDD	3.5"	2	1	500	448	N/A	N/A	
			2	1	862	731	19002	4400	
LSI2108	EP SAS HDD	2.5"	6	10	3190	1643	27432	15504	
	EP SAS SSD	2.5	6	0	4060	2125	68001	18495	
			6	5	2360	1245	18452	7573	

Sequential accesses (performance values in MB/s):

	Configuration version			el	ential :ks ad	ential :ks ite	ential :ks ad	ential :ks ite	
RAID Controller	Hard disk type	Form factor	#Disks	RAID level	HDDs sequential 64 kB blocks 100% read [MB/s]	HDDs sequential 64 kB blocks 100% write [MB/s]	SSDs sequential 64 kB blocks 100% read [MB/s]	SSDs sequential 64 kB blocks 100% write [MB/s]	
		BC SATA HDD		2	1	112	108	506	175
Patsburg A	EP SATA SSD	2.5"	4	0	422	419	946	718	
	LI GATAGOD		4	10	226	213	662	338	
Patsburg A	BC SATA HDD	3.5"	2	1	160	153	N/A	N/A	
			2	1	371	192	680	176	
LSI2108	EP SAS HDD EP SAS SSD	2.5"	6	10	788	577	1593	647	
			6	0	1170	1022	1858	1184	
			6	5	933	960	1651	1019	

At full configuration with powerful hard disks (configured as RAID 0) the PRIMERGY CX250 S2 achieves a throughput of up to 1858 MB/s for sequential load profiles and a transaction rate of up to 68001 IO/s for typical, random application scenarios.

vServCon

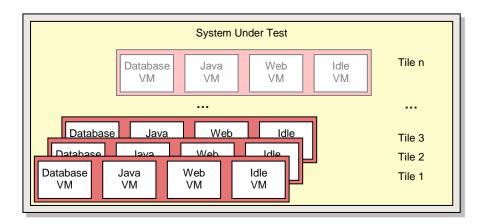
Benchmark description

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 not a new benchmark in the true sense of the word. It is more a framework that combines already established benchmarks (or in modified form) as workloads in order to reproduce the load of a consolidated and virtualized server environment. Three 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 application scenarios 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 the respective VM. In order to derive a normalized score, the individual benchmark results for one tile are put in relation to the respective results of a reference system. The resulting relative performance values are then suitably weighted and finally added up for all VMs and tiles. The outcome is a 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. This score thus reflects the maximum total throughput that can be achieved by running the mix defined in vServCon that consists of numerous VMs up to the possible full utilization of CPU resources. 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.

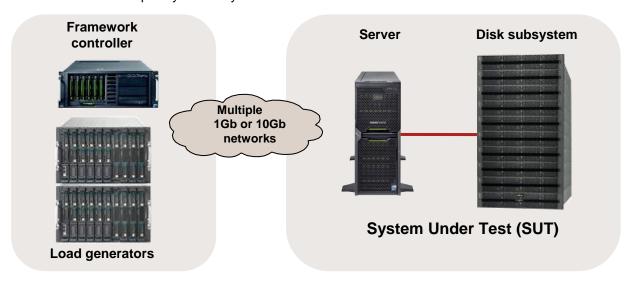
The progression of the vServCon scores for the tile numbers provides useful information about the scaling behavior of the "System under Test".

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

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

Benchmark environment

The measurement set-up is symbolically illustrated below:



System Under Test (SUT)					
Hardware					
Processor	Xeon E5-2600 v2 processor series				
Memory	1 processor: 8 x 8GB (1x8GB) 2Rx4 L DDR3-1600 R ECC 2 processors: 16 x 8GB (1x8GB) 2Rx4 L DDR3-1600 R ECC				
Network interface	1 × dual port 1GbE adapter 1 × dual port 10GbE server adapter				
Disk subsystem	1 × dual-channel FC controller Emulex LPe12002 ETERNUS DX80 storage systems: Each Tile: 50 GB LUN Each LUN: RAID 0 with 2 × Seagate ST3300657SS-Disks (15 krpm)				
Software					
Operating system	VMware ESX 5.1.0 U1 Build 1065491				

Load generator (incl. Framework controller)						
Hardware (Shared)	Hardware (Shared)					
Enclosure	PRIMERGY BX900					
Hardware						
Model	18 x PRIMERGY BX920 S1 server blades					
Processor	2 × Xeon X5570					
Memory	12 GB					
Network interface	3 x 1 Gbit/s LAN					
Software						
Operating system	Microsoft Windows Server 2003 R2 Enterprise with Hyper-V					

Load generator VM (per tile 3 load generator VMs on various server blades)					
Hardware					
Processor	1 x logical CPU				
Memory	512 MB				
Network interface	2 x 1 Gbit/s LAN				
Software					
Operating system	Microsoft Windows Server 2003 R2 Enterprise Edition				

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

Benchmark results

The PRIMERGY dual-socket systems dealt with here are based on Intel Xeon series E5-2600 v2 processors. The features of the processors are summarized in the section "Technical data".

The available processors of these systems with their results can be seen in the following table.

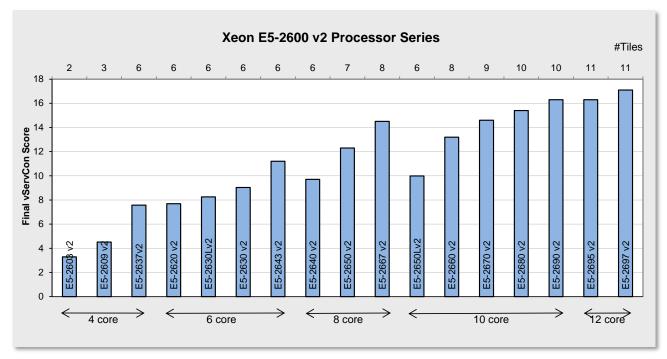
	Process	or	RX200 S8	RX300 S8	RX350 S8	TX300 S8	BX924 S4	CX250 S2	CX270 S2	#Tiles	Score
	4 Cores	E5-2603 v2	✓	✓	✓	✓	✓	✓	✓	2	3.29
	4 00100	E5-2609 v2	✓	✓	✓	✓	✓	✓	✓	3	4.52
	4 Cores, HT, TM	E5-2637 v2	✓	✓	✓	✓	✓	✓	✓	6	7.57
		E5-2620 v2	✓	✓	✓	✓	✓	✓	✓	6	7.69
	6 Cores	E5-2630Lv2	✓	✓	✓	✓	✓	✓	✓	6	8.25
es	HT, TM	E5-2630 v2	✓	✓	✓	✓	✓	✓	✓	6	9.03
Series		E5-2643 v2	✓	✓	✓	✓	✓	✓	✓	6	11.2
٧2		E5-2640 v2	✓	✓	✓	✓	✓	✓	✓	6	9.70
5600	8 Cores HT, TM	E5-2650 v2	✓	✓	✓	✓	✓	✓	✓	7	12.3
E5-2600	111, 1111	E5-2667 v2	✓	✓	✓	✓	✓	✓	✓	8	14.5
Xeon		E5-2650Lv2	✓	✓	✓	✓	✓	✓	✓	6	9.99
		E5-2660 v2	✓	✓	✓	✓	✓	✓	✓	8	13.2
	10 Cores HT, TM	E5-2670 v2	✓	✓	✓	✓	✓	✓	✓	9	14.6
		E5-2680 v2	✓	✓	✓	✓	✓	✓	✓	10	15.4
		E5-2690 v2	✓	✓	✓	✓	✓	✓	✓	10	16.3
	12 Cores	E5-2695 v2	✓	✓	✓	✓	✓	✓	✓	11	16.3
	HT, TM	E5-2697 v2	✓	✓	✓	✓	✓	✓	✓	11	17.1

HT = Hyper-Threading, TM = Turbo Mode

bold: measured, cursive: calculated

These PRIMERGY dual-socket systems are very suitable for application virtualization thanks to the progress made in processor technology. Compared with a system based on the previous processor generation an approximate 26% higher virtualization performance can be achieved (measured in vServCon score in their maximum configuration).

The first diagram compares the virtualization performance values that can be achieved with the processors reviewed here.

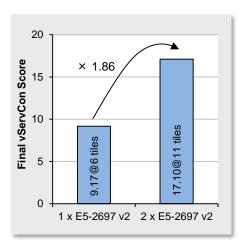


The relatively large performance differences between the processors can be explained by their features. The values scale on the basis of the number of cores, the size of the L3 cache and the CPU clock frequency and as a result of the features of Hyper-Threading and turbo mode, which are available in most processor types. Furthermore, the data transfer rate between processors ("QPI Speed") also determines performance.

A low performance can be seen in the Xeon E5-2603 v2 and E5-2609 v2 processors, as they have to manage without Hyper-Threading (HT) and turbo mode (TM). In principle, these weakest processors are only to a limited extent suitable for the virtualization environment.

Within a group of processors with the same number of cores scaling can be seen via the CPU clock frequency.

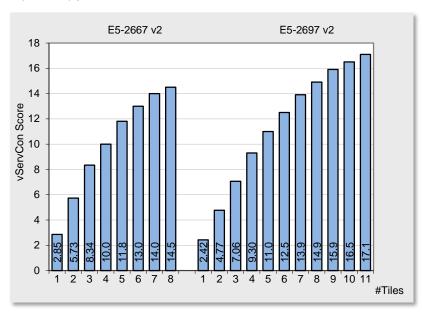
As a matter of principle, the memory access speed also influences performance. A guideline in the virtualization environment for selecting main memory is that sufficient quantity is more important than the speed of the memory accesses. The vServCon scaling measurements presented here were all performed with a memory access speed - depending on the processor type - of at most 1600 MHz. More information about the topic "Memory Performance" and QPI architecture can be found in the White Paper Memory performance of Xeon E5-2600 v2 (Ivy Bridge-EP)-based systems.



Until now we have looked at the virtualization performance of a fully configured system. However, with a server with two sockets the question also arises as to how good performance scaling is from one to two processors. The better the scaling, the lower the overhead usually caused by the shared use of resources within a server. The scaling factor also depends on the application. If the server is used as a virtualization platform for server consolidation, the system scales with a factor of 1.86 or better. When operated with two processors, the system thus achieves a significantly better performance than with one processor, as is illustrated in the diagram opposite using the processor version Xeon E5-2697 v2 as an example. In this case, the scaling of the processor versions with a lower overall performance is somewhat better than for the CPU reviewed here with the largest number of cores.

The next diagram illustrates the virtualization performance for increasing numbers of VMs based on the Xeon E5-2667 v2 (8 core) and E5-2697 v2 (12 core) processors.

In addition to the increased number of physical cores, Hyper-Threading, which is supported by almost all Xeon processors of the E5-2600 v2 series, is an additional reason for the high number of VMs that can be operated. As is known, a physical processor core is consequently divided into two logical cores so that the number of cores available for the hypervisor is doubled. This standard feature thus generally increases the virtualization performance of a system.

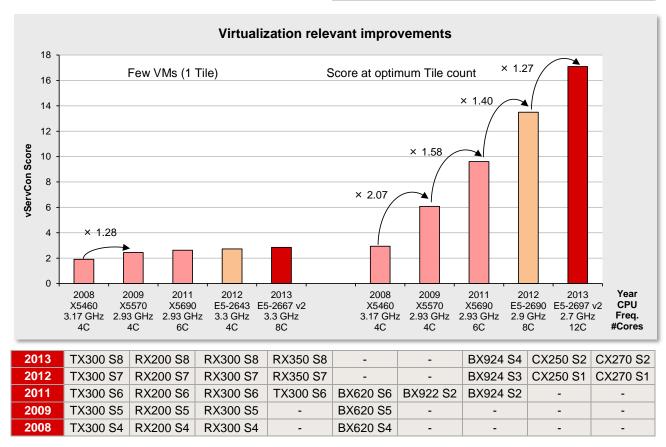


The previous diagram examined the total performance of all application VMs of a host. However, studying the performance from an individual application VM viewpoint is also interesting. This information is in the previous diagram. For example, the total optimum is reached in the above Xeon E5-2667 v2 situation with 24 application VMs (eight tiles, not including the idle VMs); the low load case is represented by three application VMs (one tile, not including the idle VM). Remember: the vServCon score for one tile is an average value across the three application scenarios in vServCon. This average performance of one tile drops when changing from the low load case to the total optimum of the vServCon score - from 2.85 to 14.5/8=1.81, i.e. to 64%. The individual types of application VMs can react very differently in the high load situation. It is thus clear that in a specific situation the performance requirements of an individual application must be balanced against the overall requirements regarding the numbers of VMs on a virtualization host.

The virtualization-relevant progress in processor technology since 2008 has an effect on the one hand on an individual VM and, on the other hand, on the possible maximum number of VMs up to CPU full utilization.

The following comparison shows the proportions for both types of improvements. Five systems are compared: a system from 2008, a system from 2009, a system from 2011, a system from 2012 and a current system with the best processors each (see table opposite) for few VMs and for highest maximum performance.

	Best Performance Few VMs	∨ServCon Score 1 Tile	Best Maximum Performance	vServCon Score max.
2008	X5460	1.91	X5460	2.94
2009	X5570	2.45	X5570	6.08
2011	X5690	2.63	X5690	9.61
2012	E5-2643	2.73	E5-2690	13.5
2013	E5-2667 v2	2.85	E5-2697 v2	17.1



The clearest performance improvements arose from 2008 to 2009 with the introduction of the Xeon 5500 processor generation (e. g. via the feature "Extended Page Tables" (EPT)¹). One sees an increase of the vServCon score by a factor of 1.28 with a few VMs (one tile).

With full utilization of the systems with VMs there was an increase by a factor of 2.07. The one reason was the performance increase that could be achieved for an individual VM (see score for a few VMs). The other reason was that more VMs were possible with total optimum (via Hyper-Threading). However, it can be seen that the optimum was "bought" with a triple number of VMs with a reduced performance of the individual VM.

Where exactly is the technology progress between 2009 and 2013?

The performance for an individual VM in low-load situations has only slightly increased for the processors compared here with the highest clock frequency per core. We must explicitly point out that the increased virtualization performance as seen in the score cannot be completely deemed as an improvement for one individual VM.

The decisive progress is in the higher number of physical cores and – associated with it – in the increased values of maximum performance (factor 1.58, 1.40 and 1.27 in the diagram).

Up to and including 2011 the best processor type of a processor generation had both the highest clock frequency and the highest number of cores. From 2012 there have been differently optimized processors on

4

¹ EPT accelerates memory virtualization via hardware support for the mapping between host and guest memory addresses.

offer: Versions with a high clock frequency per core for few cores and versions with a high number of cores, but with a lower clock frequency per core. The features of the processors are summarized in the section "Technical data".

Performance increases in the virtualization environment since 2009 are mainly achieved by increased VM numbers due to the increased number of available logical or physical cores. However, since 2012 it has been possible - depending on the application scenario in the virtualization environment – to also select a CPU with an optimized clock frequency if a few or individual VMs require maximum computing power.

STREAM

Benchmark description

STREAM is a synthetic benchmark that has been used for many years to determine memory throughput and which was developed by John McCalpin during his professorship at the University of Delaware. Today STREAM is supported at the University of Virginia, where the source code can be downloaded in either Fortran or C. STREAM continues to play an important role in the HPC environment in particular. It is for example an integral part of the HPC Challenge benchmark suite.

The benchmark is designed in such a way that it can be used both on PCs and on server systems. The unit of measurement of the benchmark is GB/s, i.e. the number of gigabytes that can be read and written per second.

STREAM measures the memory throughput for sequential accesses. These can generally be performed more efficiently than accesses that are randomly distributed on the memory, because the CPU caches are used for sequential access.

Before execution the source code is adapted to the environment to be measured. Therefore, the size of the data area must be at least four times larger than the total of all CPU caches so that these have as little influence as possible on the result. The OpenMP program library is used to enable selected parts of the program to be executed in parallel during the runtime of the benchmark, consequently achieving optimal load distribution to the available processor cores.

During implementation the defined data area, consisting of 8-byte elements, is successively copied to four types, and arithmetic calculations are also performed to some extent.

Туре	Execution	Bytes per step	Floating-point calculation per step
COPY	a(i) = b(i)	16	0
SCALE	$a(i) = q \times b(i)$	16	1
SUM	a(i) = b(i) + c(i)	24	1
TRIAD	$a(i) = b(i) + q \times c(i)$	24	2

The throughput is output in GB/s for each type of calculation. The differences between the various values are usually only minor on modern systems. In general, only the determined TRIAD value is used as a comparison.

The measured results primarily depend on the clock frequency of the memory modules; the CPUs influence the arithmetic calculations. The accuracy of the results is approximately 5%.

This chapter specifies throughputs on a basis of 10 (1 GB/s = 10^9 Byte/s).

Benchmark environment

System Under Test	System Under Test (SUT)				
Hardware					
Enclosure	PRIMERGY CX400 S2				
Model	PRIMERGY CX250 S2				
Processor	2 processors of Xeon E5-2600 v2 processor series				
Memory	8 × 16GB (1x16GB) 2Rx4 DDR3-1866 R ECC				
Software					
BIOS settings	Processors other than Xeon E5-2603 v2, E5-2609 v2: Hyper-Threading = Disabled				
Operating system	Red Hat Enterprise Linux Server release 6.4				
Operating system settings	echo never > /sys/kernel/mm/redhat_transparent_hugepage/enabled				
Compiler	Intel C Compiler 12.1				
Benchmark	Stream.c Version 5.9				

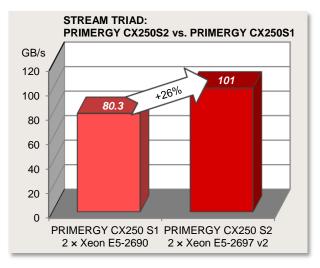
Some components may not be available in all countries or sales regions.

Benchmark results

Processor	Cores	Processor Frequency [Ghz]	Max. Memory Frequency [MHz]	TRIAD [GB/s]
2 x Xeon E5-2603 v2	4	1.80	1333	48.4
2 x Xeon E5-2609 v2	4	2.50	1333	59.4
2 x Xeon E5-2637 v2	4	3.50	1866	82.2
2 x Xeon E5-2620 v2	6	2.10	1600	78.9
2 x Xeon E5-2630Lv2	6	2.40	1600	80.5
2 x Xeon E5-2630 v2	6	2.60	1600	81.9
2 x Xeon E5-2643 v2	6	3.50	1866	
2 x Xeon E5-2640 v2	8	2.00	1600	83.4
2 x Xeon E5-2650 v2	8	2.60	1866	97.0
2 x Xeon E5-2667 v2	8	3.30	1866	
2 x Xeon E5-2650Lv2	10	1.70	1600	81.9
2 x Xeon E5-2660 v2	10	2.20	1866	96.1
2 x Xeon E5-2670 v2	10	2.50	1866	97.3
2 x Xeon E5-2680 v2	10	2.80	1866	97.9
2 x Xeon E5-2690 v2	10	3.00	1866	98.5
2 x Xeon E5-2695 v2	12	2.40	1866	101
2 x Xeon E5-2697 v2	12	2.70	1866	101

The results depend primarily on the maximum memory frequency. The processors with only 4 cores, which do not fully utilize their memory controller, are an exception. The smaller differences with processors with the same maximum memory frequency are a result in arithmetic calculation of the different processor frequencies.

The following diagram illustrates the throughput of the PRIMERGY CX250 S2 in comparison to its predecessor, the PRIMERGY CX250 S1, in their most performant configuration.



LINPACK

Benchmark description

LINPACK was developed in the 1970s by Jack Dongarra and some other people to show the performance of supercomputers. The benchmark consists of a collection of library functions for the analysis and solution of linear system of equations. A description can be found in the document http://www.netlib.org/utk/people/JackDongarra/PAPERS/hplpaper.pdf.

LINPACK can be used to measure the speed of computers when solving a linear equation system. For this purpose, an $n \times n$ matrix is set up and filled with random numbers between -2 and +2. The calculation is then performed via LU decomposition with partial pivoting.

A memory of $8n^2$ bytes is required for the matrix. In case of an n × n matrix the number of arithmetic operations required for the solution is $^2/_3n^3 + 2n^2$. Thus, the choice of n determines the duration of the measurement: a doubling of n results in an approximately eight-fold increase in the duration of the measurement. The size of n also has an influence on the measurement result itself: as n increases, the measured value asymptotically approaches a limit. The size of the matrix is therefore usually adapted to the amount of memory available. Furthermore, the memory bandwidth of the system only plays a minor role for the measurement result, but a role that cannot be fully ignored. The processor performance is the decisive factor for the measurement result. Since the algorithm used permits parallel processing, in particular the number of processors used and their processor cores are - in addition to the clock rate - of outstanding significance.

LINPACK is used to measure how many floating point operations were carried out per second. The result is referred to as **Rmax** and specified in GFlops (Giga Floating Point Operations per Second).

An upper limit, referred to as **Rpeak**, for the speed of a computer can be calculated from the maximum number of floating point operations that its processor cores could theoretically carry out in one clock cycle:

Rpeak = Maximum number of floating point operations per clock cycle

- x Number of processor cores of the computer
- x Maximum processor frequency[GHz]

LINPACK is classed as one of the leading benchmarks in the field of high performance computing (HPC). LINPACK is one of the seven benchmarks currently included in the HPC Challenge benchmark suite, which takes other performance aspects in the HPC environment into account.

Manufacturer-independent publication of LINPACK results is possible at http://www.top500.org/. The use of a LINPACK version based on HPL is prerequisite for this (see: http://www.netlib.org/benchmark/hpl).

Intel offers a highly optimized LINPACK version (shared memory version) for individual systems with Intel processors. Parallel processes communicate here via "shared memory", i.e. jointly used memory. Another version provided by Intel is based on HPL (High Performance Linpack). Intercommunication of the LINPACK processes here takes place via OpenMP and MPI (Message Passing Interface). This enables communication between the parallel processes - also from one computer to another. Both versions can be downloaded from http://software.intel.com/en-us/articles/intel-math-kernel-library-linpack-download/.

Manufacturer-specific LINPACK versions also come into play when graphics cards for General Purpose Computation on Graphics Processing Unit (GPGPU) are used. These are based on HPL and include extensions which are needed for communication with the graphics cards.

Benchmark environment

System Under Test (SUT)						
Hardware						
Enclosure	PRIMERGY CX400 S2					
Model	PRIMERGY CX250 S2					
Processor	2 processors of Xeon E5-2600 v2 processor series					
Memory	16 x 8GB (1x8GB) 2Rx8 DDR3-1866 R ECC					
Software						
BIOS settings	All processors apart from Xeon E5-2603 v2, E5-2609 v2: Hyper Threading = Disabled All processors apart from Xeon E5-2603 v2, E5-2609 v2: Turbo Mode = Enabled (default) = Disabled					
Operating system	Red Hat Enterprise Linux Server release 6.4					
Benchmark	HPL version: Intel Optimized MP LINPACK Benchmark for Clusters 11.0 Update 5 for Linux OS					

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

Benchmark results

Processor	Cores	Processor frequency [Ghz]	Maximum turbo frequency at full load [Ghz]	Number of processors	Without Turbo Mode		With Turbo Mode	
					Rpeak [GFlops]	Rmax [GFlops]	Rpeak [GFlops]	Rmax [GFlops]
Xeon E5-2603 v2	4	1.80	n/a	2	115	110		
Xeon E5-2609 v2	4	2.50	n/a	2	160	152		
Xeon E5-2637 v2	4	3.50	3.60	2	224	213	230	220
Xeon E5-2620 v2	6	2.10	2.40	2	202	192	230	219
Xeon E5-2630Lv2	6	2.40	2.60	2	230	219	250	238
Xeon E5-2630 v2	6	2.60	2.90	2	250	238	278	265
Xeon E5-2643 v2	6	3.50	3.60	2	336		346	
Xeon E5-2640 v2	8	2.00	2.30	2	256	244	294	280
Xeon E5-2650 v2	8	2.60	3.00	2	333		384	
Xeon E5-2667 v2	8	3.30	3.60	2	422		461	
Xeon E5-2650Lv2	10	1.70	1.90	2	272	259	304	289
Xeon E5-2660 v2	10	2.20	2.60	2	352	335	416	396
Xeon E5-2670 v2	10	2.50	2.90	2	400		464	
Xeon E5-2680 v2	10	2.80	3.10	2	448	426	496	459
Xeon E5-2690 v2	10	3.00	3.30	2	480	456	528	487
Xeon E5-2695 v2	12	2.40	2.80	2	461		538	
Xeon E5-2697 v2	12	2.70	3.00	2	518		576	

Rmax = Measurement result

Rpeak = Maximum number of floating point operations per clock cycle

x Number of processor cores of the computer

x Maximum processor frequency[GHz]

The following applies for processors without Turbo mode and for those with Turbo mode disabled:

Maximum processor frequency[GHz] = Nominal processor frequency[GHz]

Processors with Turbo mode enabled are not limited by the nominal processor frequency and therefore do not provide a constant processor frequency. Instead the actual processor frequency swings - depending on temperature and power consumption - between the nominal processor frequency and maximum turbo frequency at full load. Therefore, the following applies for these processors:

Maximum processor frequency[GHz] = Maximum turbo frequency at full load[GHz]

Literature

PRIMERGY Systems

http://primergy.com/

PRIMERGY CX250 S2

This White Paper:

http://docs.ts.fujitsu.com/dl.aspx?id=fa24cd13-260c-4ab1-b8b4-7372ae5b3bd3

http://docs.ts.fujitsu.com/dl.aspx?id=93b172a3-94fa-4718-98ad-a65892476d34

http://docs.ts.fujitsu.com/dl.aspx?id=a9e347a9-a481-4e96-b963-11a9dd6599cd

Data sheet

http://docs.ts.fujitsu.com/dl.aspx?id=5344296a-14c7-477d-837f-17fb2e035a89

Memory performance of Xeon E5-2600 v2 (Ivy Bridge-EP)-based systems http://docs.ts.fujitsu.com/dl.aspx?id=a344b05e-2e9d-481b-8c9b-c6542defd839

PRIMERGY Performance

http://www.fujitsu.com/fts/products/computing/servers/primergy/benchmarks/

Disk I/O

Basics of Disk I/O Performance

http://docs.ts.fujitsu.com/dl.aspx?id=65781a00-556f-4a98-90a7-7022feacc602

Single Disk Performance

http://docs.ts.fujitsu.com/dl.aspx?id=0e30cb69-44db-4cd5-92a7-d38bacec6a99

RAID Controller Performance

http://docs.ts.fujitsu.com/dl.aspx?id=e2489893-cab7-44f6-bff2-7aeea97c5aef

Information about Iometer

http://www.iometer.org

LINPACK

The LINPACK Benchmark: Past, Present, and Future

http://www.netlib.org/utk/people/JackDongarra/PAPERS/hplpaper.pdf.

TOP500

http://www.top500.org/

HPL - A Portable Implementation of the High-Performance Linpack Benchmark for Distributed-Memory Computers

http://www.netlib.org/linpack/

Intel Math Kernel Library - LINPACK Download

http://software.intel.com/en-us/articles/intel-math-kernel-library-linpack-download/

SPECcpu2006

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

Benchmark overview SPECcpu2006

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

STREAM

http://www.cs.virginia.edu/stream/

vServCon

Benchmark Overview vServCon

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

Contact

FUJITSU

Website: http://www.fujitsu.com/

PRIMERGY Product Marketing

mailto:Primergy-PM@ts.fujitsu.com

PRIMERGY Performance and Benchmarks

mailto:primergy.benchmark@ts.fujitsu.com

© Copyright 2013 Fujitsu Technology Solutions. Fujitsu and the Fujitsu logo are trademarks or registered trademarks of Fujitsu Limited in Japan and other countries. Other company, product and service names may be trademarks or registered trademarks of their respective owners. Technical data subject to modification and delivery subject to availability. Any liability that the data and illustrations are complete, actual or correct is excluded. Designations may be trademarks and/or copyrights of the respective manufacturer, the use of which by third parties for their own purposes may infringe the rights of such owner. For further information see http://www.fujitsu.com/fts/resources/navigation/terms-of-use.html
2013-10-02 WW EN

http://ts.fujitsu.com/primergy Page 29 (29)