White Paper
FUJITSU Server PRIMEQUEST
Performance Report PRIMEQUEST 2800B2

This document contains a summary of the benchmarks executed for the FUJITSU Server PRIMEQUEST 2800B2.

The PRIMEQUEST 2800B2 performance data are compared with the data of other PRIMEQUEST models and discussed. In addition to the benchmark results, an explanation has been included for each benchmark and for the benchmark environment.

Version
1.1
2015-11-18

Contents

Document history

Version 1.0 (2015-09-15)

New:

- Technical data
- SPECcpu2006
  Measurements with Intel® Xeon® Processor E7-8800 v3 Product Family
- Disk I/O: Performance of RAID controllers
  Measurements with “PRAID EP420i” controller
- vServCon
  Results for Intel® Xeon® Processor E7 v3 Family
- VMmark V2
  “Performance Only“ Measurements with Xeon E7-8890 v3
- STREAM
  Measurements with Intel® Xeon® Processor E7-8800 v3 Product Family

Version 1.1 (2015-11-18)

New:

- OLTP-2
  Results for Intel® Xeon® Processor E7 v3 Family

Technical data

PRIMEQUEST 2800B2

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 memory modules. Separate reference will be made to any further exceptions where applicable.

Model	PRIMEQUEST 2800B2
Form factor	Rack server
Number of system boards orderable	1 – 4
Number of I/O units orderable	1 – 4
Number of disk units orderable	0 – 2
Per system board:	
Chipset	Intel® C602 Chipset
Number of sockets	2
Number of processors orderable	2
Processor type	Intel® Xeon® Processor E7-8800 v3 Product Family
Number of memory slots	48 (24 per processor)
Maximum memory configuration	3 TB
Max. number of internal hard disks	0
Per I/O unit:	
Onboard LAN controller	I/O Unit L (1GbE, 2xbaseTports) PQ2800E: 2 × 1 Gbit/s I/O Unit F (10GbE, 2xbaseTports) PQ2800E: 2 × 10 Gbit/s
PCI slots	I/O Unit L (1GbE, 2xbaseTports) PQ2800B: 4 × PCI-Express 3.0 x8 I/O Unit F (10GbE, 2xbaseTports) PQ2800B: 1 × PCI-Express 3.0 x8 2 × PCI-Express 3.0 x16
Per disk unit:	
Max. number of internal hard disks	4

Processors (since system release)									
Processor	Cores	Threads	Cache [MB]	QPI Speed [GT/s]	Rated Frequency [Ghz]	Max. Turbo Frequency [Ghz]	Max. Memory Frequency 1) [MHz]	TDP [Watt]
Xeon E7-8893 v3	4	8	45	9.60	3.20	3.50	1600	140
Xeon E7-8891 v3	10	20	45	9.60	2.80	3.50	1600	165
Xeon E7-8860 v3	16	32	40	9.60	2.20	3.20	1600	140
Xeon E7-8867 v3	16	32	45	9.60	2.50	3.30	1600	165
Xeon E7-8870 v3	18	36	45	9.60	2.10	2.90	1600	140
Xeon E7-8880 v3	18	36	45	9.60	2.30	3.10	1600	150
Xeon E7-8890 v3	18	36	45	9.60	2.50	3.30	1600	165

1) BIOS setting: *Memory Operation Mode = Performance Mode*

All the processors that can be ordered with the PRIMEQUEST 2800B2 support Intel® Turbo Boost Technology 2.0. This technology allows you to operate the processor with higher frequencies than the nominal frequency. Listed in the processor table is "Max. Turbo Frequency" for the theoretical frequency maximum with only one active core per processor. The maximum frequency that can actually be achieved depends on the number of active cores, the current consumption, electrical power consumption and the temperature of the processor.

As a matter of principle Intel does not guarantee that the maximum turbo frequency will be reached. This is related to manufacturing tolerances, which result in a variance regarding the performance of various examples of a processor model. The range of the variance covers the entire scope between the nominal frequency and the maximum turbo frequency.

The turbo functionality can be set via BIOS option. Fujitsu generally recommends leaving the "Turbo Mode" option set at the standard setting "Enabled", as performance is substantially increased by the higher frequencies. However, since the higher frequencies depend on general conditions and are not always guaranteed, it can be advantageous to disable the "Turbo Mode" option for application scenarios with intensive use of AVX instructions and a high number of instructions per clock unit, as well as for those that require constant performance or lower electrical power consumption.

Memory modules (since system release)								
Memory module	Capacity [GB]	Ranks	Bit width of the memory chips	Frequency [MHz]	Low voltage	Load reduced	Registered	ECC
16GB (2x8GB) 1Rx4 DDR4-2133 R ECC	16	1	4	2133			✓	✓
32GB (2x16GB) 2Rx4 DDR4-2133 R ECC	32	2	4	2133			✓	✓
64GB (2x32GB) 4Rx4 DDR4-2133 LR ECC	64	4	4	2133		✓	✓	✓
128GB (2x64GB) 4Rx4 DDR4-2133 LR ECC	128	4	4	2133		✓	✓	✓

Power supplies (since system release)	Max. number
Power supply 2.880W silver	6
Power Supply 2.880W platinum hp	6

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

Detailed technical information is available in the data sheet PRIMEQUEST 2800B2.

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	Type	Compiler optimization	Measurement result	Application
SPECint2006	integer	peak	aggressive	Speed	single-threaded
SPECint_base2006	integer	base	conservative		
SPECint_rate2006	integer	peak	aggressive	Throughput	multi-threaded
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	multi-threaded
SPECfp_rate_base2006	floating point	base	conservative		

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	
Model	PRIMEQUEST 2800B2
Processor	Intel® Xeon® Processor E7-8800 v3 Product Family
Memory	4 Sockets: 32 × 32GB (2x16GB) 2Rx4 DDR4-2133 R ECC 8 Sockets: 64 × 32GB (2x16GB) 2Rx4 DDR4-2133 R ECC
Software	
BIOS settings	Energy Performance = Performance
Operating system	SPECint_rate_base2006, SPECint_rate2006: Red Hat Enterprise Linux Server release 6.6 SPECfp_rate_base2006, SPECfp_rate2006: Red Hat Enterprise Linux Server release 7.1
Operating system settings	SPECint_rate_base2006, SPECint_rate2006: echo always > /sys/kernel/mm/redhat_transparent_hugepage/enabled
Compiler	SPECint_rate_base2006, SPECint_rate2006: Version 14.0.0.080 of Intel C++ Studio XE for Linux SPECfp_rate_base2006, SPECfp_rate2006: C/C++: Version 15.0.0.090 of Intel C++ Studio XE for Linux Fortran: Version 15.0.0.090 of Intel Fortran Studio XE for Linux

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

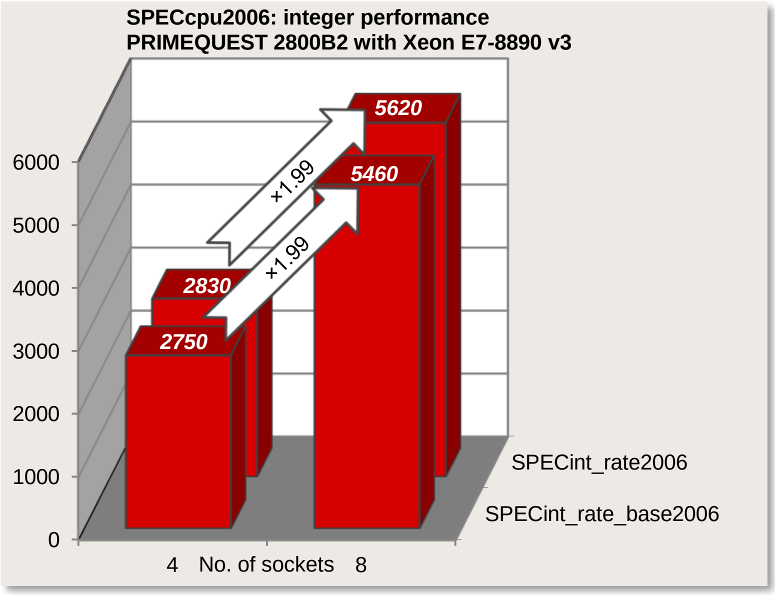
Benchmark results

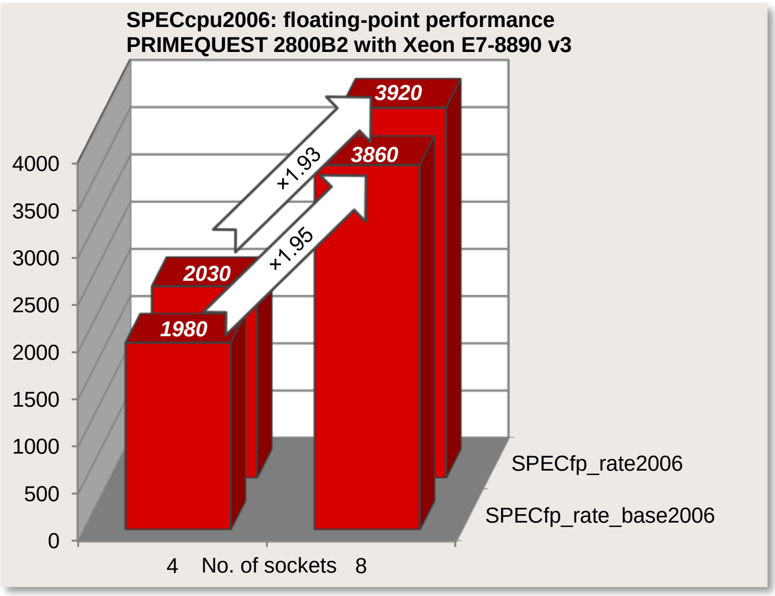
In terms of processors the benchmark result depends primarily on the size of the processor cache, the support for Hyper-Threading, the number of processor cores and on the processor frequency. The number of cores, which are loaded by the benchmark, determines the maximum processor frequency that can be achieved.

Processor	Number of processors	SPECint_rate_base2006	SPECint_rate2006	Number of processors	SPECint_rate_base2006	SPECint_rate2006
Xeon E7-8893 v3	4	850	881	8	1700	1760
Xeon E7-8891 v3	4			8	3780	3910
Xeon E7-8860 v3	4			8	4740	4870
Xeon E7-8867 v3	4			8	4920	5070
Xeon E7-8870 v3	4			8	5070	5220
Xeon E7-8880 v3	4			8	5250	5410
Xeon E7-8890 v3	4	2750	2830	8	5460	5620

Processor	Number of processors	SPECfp_rate_base2006	SPECfp_rate2006	Number of processors	SPECfp_rate_base2006	SPECfp_rate2006
Xeon E7-8893 v3	4	793	824	8	1580	1610
Xeon E7-8891 v3	4			8	3040	3090
Xeon E7-8860 v3	4			8	3520	3620
Xeon E7-8867 v3	4			8	3690	3750
Xeon E7-8870 v3	4			8	3670	3740
Xeon E7-8880 v3	4			8	3700	3770
Xeon E7-8890 v3	4	1980	2030	8	3860	3920

The two diagrams below reflect how the performance of the PRIMEQUEST 2800B2 scales from four to eight processors when using the Xeon E7-8890 v3.





Disk I/O: Performance of RAID controllers

Benchmark description

Performance measurements of disk subsystems for PRIMEQUEST servers are used to assess their performance and enable a comparison of the different storage connections for PRIMEQUEST servers. As standard, these performance measurements are carried out with a defined measurement method, which models the 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 profile	Access	Type of access		Block size [kB]	Application
		read	write		
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 capacities of storage media 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 (SUT)	
Hardware	
Controller	1 × "PRAID EP420i"
Drive	4 × 2.5" SAS SSD Toshiba PX02SMF040 4 × 2.5" SAS HDD HGST HUC156045CSS204
Software	
BIOS settings	Intel Virtualization Technology = Disabled VT-d = Disabled Energy Performance = Performance Utilization Profile = Unbalanced CPU C6 Report = Disabled
Operating system	Microsoft Windows Server 2012 Standard
Operating system settings	Choose or customize a power plan: High performance For the processes that create disk I/Os: set the AFFINITY to the CPU node to which the PCIe slot of the RAID controller is connected
Administration software	ServerView RAID Manager 6.1.4
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 PRIMEQUEST 2800B2 in the light of disk-I/O performance. Various combinations of RAID controllers and storage media will be analyzed below.

Hard disks

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).

Model versions

The maximum number of hard disks in the system depends on the system configuration. The PRIMEQUEST 2800B2 permits up to two disk units (DU). The disk units are also referred to below with the generic term "subunit".

The following table lists the essential cases. The two configuration versions of the disk unit are abbreviated as follows: "Disk Unit (1C)" is a disk unit with one controller and "Disk Unit (2C)" is a disk unit with two controllers.

Only the highest supported version is named for all the interfaces we have dealt with in this section.

Subunit	Form factor	Interface	Number of PCIe controllers	Maximum number of hard disks
Disk Unit (1C)	2.5"	SAS 12G	1	4
Disk Unit (2C)	2.5"	SAS 12G	2	2 × 2

Thanks to the modular architecture of the system it is sufficient to consider the disk-I/O performance for each controller. The possible overall performance of the system is the result of the sum of the performance maximums of all the controllers contained in the system.

RAID controller

In addition to the hard disks the RAID controller is the second performance-determining key component.

The following table summarizes the most important features of the available RAID controllers of the PRIMEQUEST 2800B2. A short alias is specified here for each controller, which is used in the subsequent list of the performance values.

Controller name / mounting position	Alias	Cache	Supported interfaces		Max. # disks in the subunit	RAID levels in the subunit	BBU/ FBU
PRAID EP420i in Disk Unit (1C)	PRAID EP420i (DU-1C)	2 GB	SAS 12G	PCIe 3.0 x8	4 × 2.5"	0, 1, 1E, 5, 6, 10	-/✓
PRAID EP420i in Disk Unit (2C)	PRAID EP420i (DU-2C)	2 GB	SAS 12G	PCIe 3.0 x8	2 × 2.5"	0, 1	-/✓

System-specific interfaces

The interfaces of a controller to the system board (also applies for the disk unit / system board interface) 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 configuration					Connection via expander
	# Disk channels	Limit for throughput of disk interface	PCIe-version	PCIe width	Limit for throughput of PCIe interface	
PRAID EP420i (DU-1C)	4 × SAS 12G	4120 MB/s	2.0	x4	**1716 MB/s**	-
PRAID EP420i (DU-2C)	2 × SAS 12G	2060 MB/s	2.0	x4	**1716 MB/s**	-

More details about these RAID controllers are available in the white paper "RAID Controller Performance".

Settings

In most cases, the cache of HDDs has a great influence on disk-I/O performance. 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. 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 the server. 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 system-specific interfaces are not exceeded. This is why all the performance statements of the document “RAID Controller Performance” also apply for the PRIMEQUEST 2800B2 if the configurations measured there are also supported by this system.

The performance values of the PRIMEQUEST 2800B2 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 established measurement variables, as already mentioned in the subsection Benchmark description, are used here. 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.

2.5" - Random accesses (maximum performance values in IO/s):

Base Unit PQ2800B2							
Configuration version			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]
RAID Controller	Hard disk type	#Disks					
PRAID EP420i (DU-1C)	HUC156045CSS204 SAS HDD PX02SMF040 SAS SSD	2	1	1949	1085	77312	12141
		4	10	3479	1445	106524	20473
		4	0	3939	1871	128697	32507
		4	5	2151	936	36808	13110
PRAID EP420i (DU-2C)	HUC156045CSS204 SAS HDD PX02SMF040 SAS SSD	2	1	1949	1085	77312	12141
		2	0	1950	970	105105	15537

2.5" - Sequential accesses (maximum performance values in MB/s):

Base Unit PQ2800B2							
Configuration version			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]
RAID Controller	Hard disk type	#Disks					
PRAID EP420i (DU-1C)	HUC156045CSS204 SAS HDD PX02SMF040 SAS SSD	2	1	424	272	1489	418
		4	10	586	467	1571	736
		4	0	968	901	1570	1469
		4	5	726	669	1556	1248
PRAID EP420i (DU-2C)	HUC156045CSS204 SAS HDD PX02SMF040 SAS SSD	2	1	424	272	1489	418
		2	0	504	492	1489	816

Conclusion

The use of one controller at its maximum configuration with powerful hard disks enables the PRIMEQUEST 2800B2 to achieve a throughput of up to 1571 MB/s for sequential load profiles and a transaction rate of up to 128697 IO/s for typical, random application scenarios.

In the maximum system configuration with two disk units with two controllers each, i.e. a total of four controllers, a maximum of eight hard disks could be run. If powerful hard disks are used in this maximum configuration, the system would mathematically achieve a total throughput of up to 5956 MB/s for sequential load profiles and a total transaction rate of up to 420420 IO/s for typical random application scenarios.

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 system.

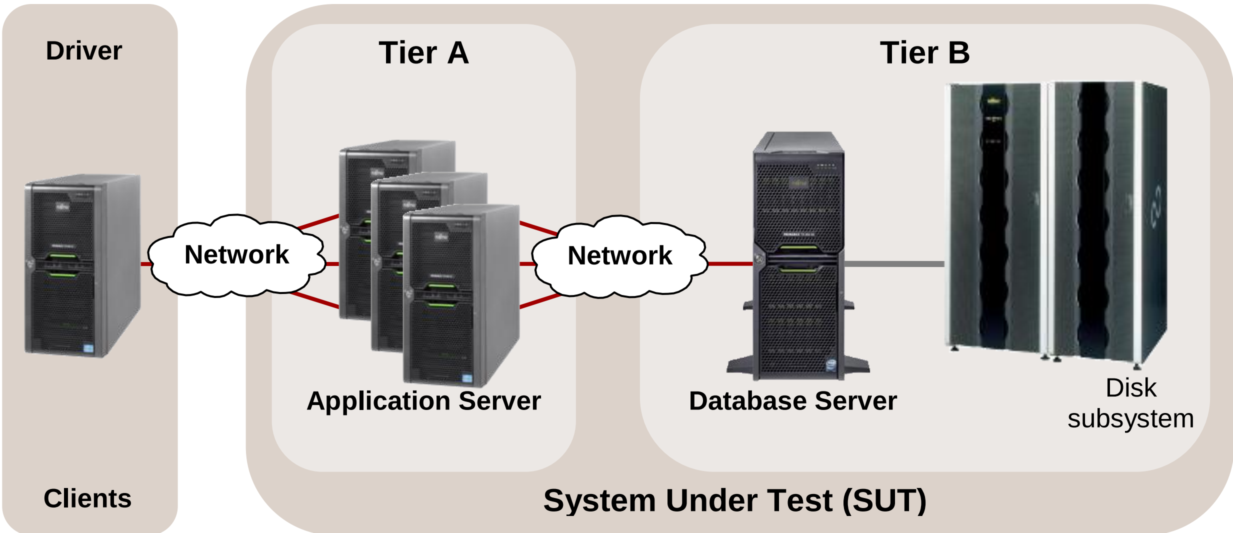
In contrast to benchmarks such as SPECint and TPC-E, which were standardized by independent bodies and for which adherence to the respective rules and regulations are monitored, OLTP-2 is an internal benchmark of Fujitsu. OLTP-2 is based on the well-known database benchmark TPC-E. OLTP-2 was designed in such a way that a wide range of configurations can be measured to present the scaling of a system with regard to the CPU and memory configuration.

Even if the two benchmarks OLTP-2 and TPC-E simulate similar application scenarios using the same load profiles, 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.

Further information can be found in the document Benchmark Overview OLTP-2.

Benchmark environment

The measurement set-up is symbolically illustrated below:



All results were determined by way of example on a PRIMEQUEST 2800E2.

Database Server (Tier B)	
Hardware	
Model	PRIMEQUEST 2800E2
Processor	Intel® Xeon® Processor E7- v3 Family
Memory	4096 GB: 64 × 64GB (2x32GB) 4Rx4 DDR4-2133 LR ECC 2048 GB: 32 × 64GB (2x32GB) 4Rx4 DDR4-2133 LR ECC 1024 GB: 16 × 64GB (2x32GB) 4Rx4 DDR4-2133 LR ECC
Network interface	2 × onboard LAN 10 Gb/s

Disk subsystem	PRIMEQUEST 2800E2: 1 × PRAID EP420i 2 × 300 GB 10k rpm SAS Drives, RAID1 (OS) 15 × PRAID EP420e 14 × JX40: Je 15 × 400 GB SSD Drive, RAID5 (data) 1 × JX40: 10 × 900 GB 10k rpm SAS Drives, RAID10 (LOG)
Software	
BIOS	Version BB15068
Operating system	Microsoft Windows Server 2012 R2 Standard
Database	Microsoft SQL Server 2014 Enterprise

Application Server (Tier A)	
Hardware	
Model	2 × PRIMERGY RX2530 M1
Processor	2 × Xeon E5-2697 v3
Memory	64 GB, 2133 MHz registered ECC DDR4
Network interface	2 × onboard LAN 10 Gb/s
Disk subsystem	2 × 300 GB 15k rpm SAS Drive
Software	
Operating system	Microsoft Windows Server 2012 Standard

Client	
Hardware	
Model	1 × PRIMERGY RX300 S8
Processor	2 × Xeon E5-2667 v2
Memory	64 GB, 1600 MHz registered ECC DDR3
Network interface	2 × onboard LAN 1 Gb/s 1 × Dual Port LAN 1Gb/s
Disk subsystem	1 × 300 GB 10k rpm SAS Drive
Software	
Operating system	Microsoft Windows Server 2012 R2 Standard
Benchmark	OLTP-2 Software EGen version 1.14.0

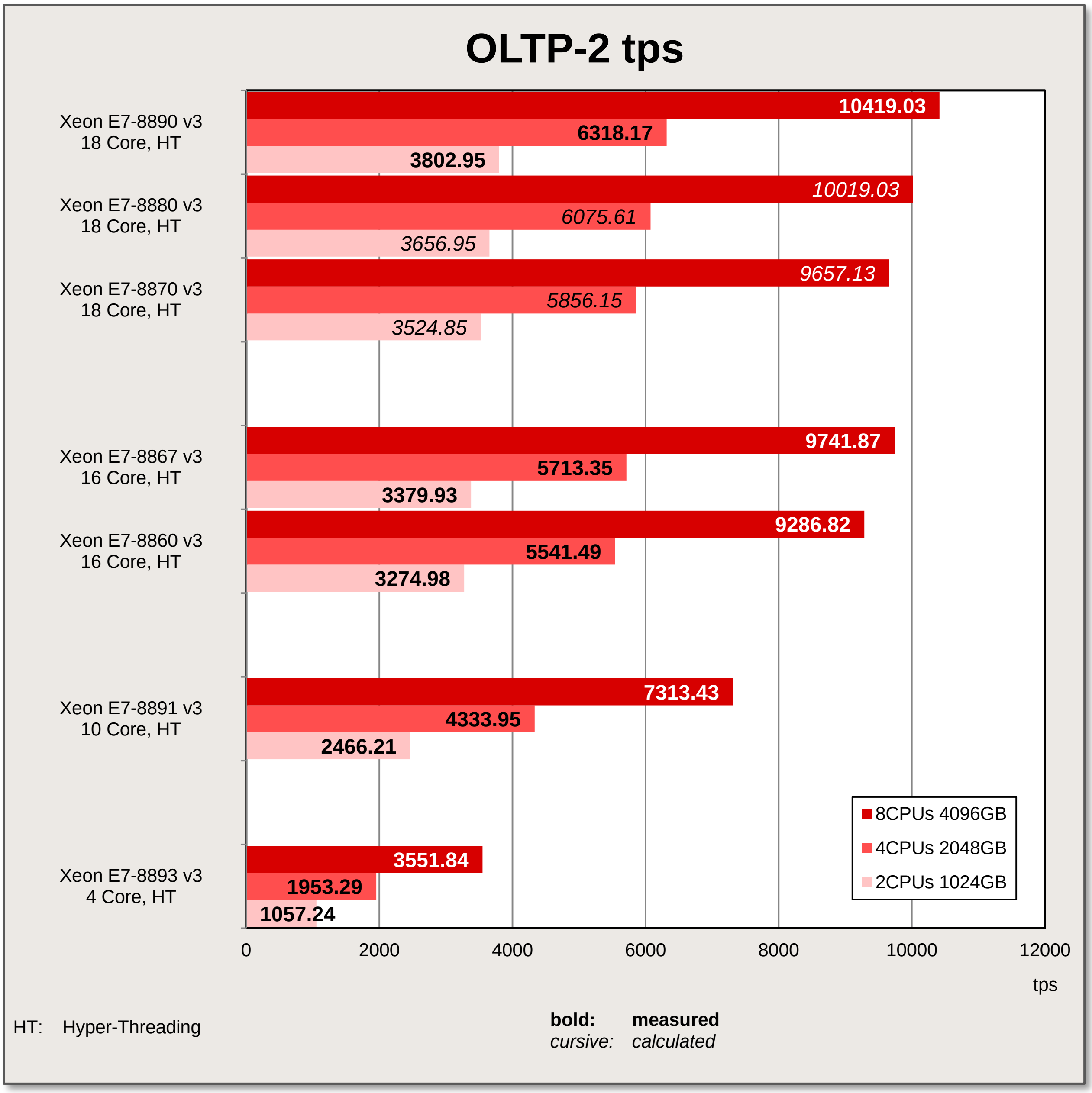
Some components may not be available in all countries / sales regions.

Benchmark results

Database performance greatly depends on the configuration options with CPU, memory and on the connectivity of an adequate disk subsystem for the database. In the following scaling considerations for the processors we assume that both the memory and the disk subsystem has been adequately chosen and is not a bottleneck.

A guideline in the database environment for selecting main memory is that sufficient quantity is more important than the speed of the memory accesses. This why a configuration with a total memory of 1024 GB was considered for the measurements with two processors, a configuration with a total memory of 2048 GB for the measurements with four processors and a configuration with a total memory of 4096 GB for the measurements with eight processors. The memory configurations had memory access of 1600 MHz. Further information about memory performance can be found in the White Paper Memory performance of Xeon E7 v3 (Haswell-EX)-based systems.

The following diagram shows the OLTP-2 transaction rates that can be achieved with two, four and eight processors of the Intel® Xeon® Processor E7 v3 Family.



It is evident that a wide performance range is covered by the variety of released processors. If you compare the OLTP-2 value of the processor with the lowest performance (Xeon E7-8893 v3) with the value of the processor with the highest performance (Xeon E7-8890 v3), the result is a 2.9-fold increase in performance.

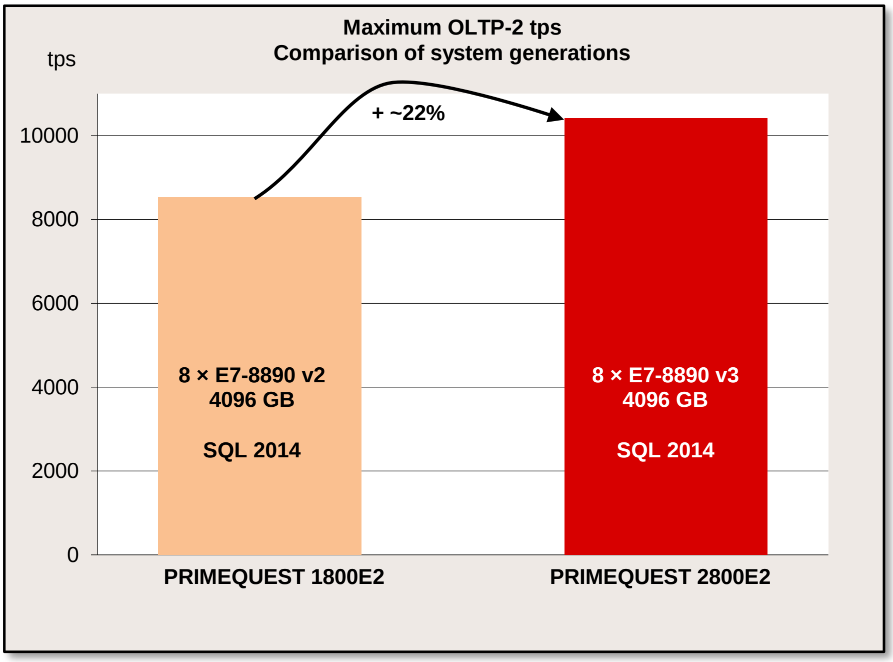
Based on the number of cores the processors can be divided into different performance groups:

The start is made with Xeon E7-8893 v3 as processor with four cores and Hyper-Threading.

The groups of 10- and 16-core processors offer in this processor series a medium-range OLTP-2 performance. Due to the various technical features of the processors in these groups (see. “Technical data”) it is possible to choose the right CPU depending on the usage scenario.

The group of processors with 18 cores is to be found at the upper end of the performance scale. Due to the graduated CPU clock frequencies an OLTP performance of between 9657.13 tps (8 × Xeon E7-8870 v3) and 10419.03 tps (8 × Xeon E7-8890 v3) is achieved.

If you compare the maximum achievable OLTP-2 values of the current system generation with the values that were achieved on the predecessor systems, the result is an increase of about 22%.



vServCon

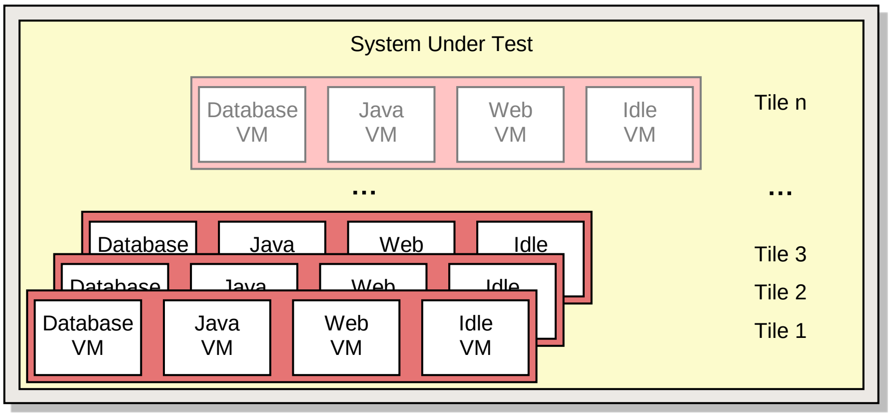
Benchmark description

vServCon is a benchmark used by Fujitsu 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.

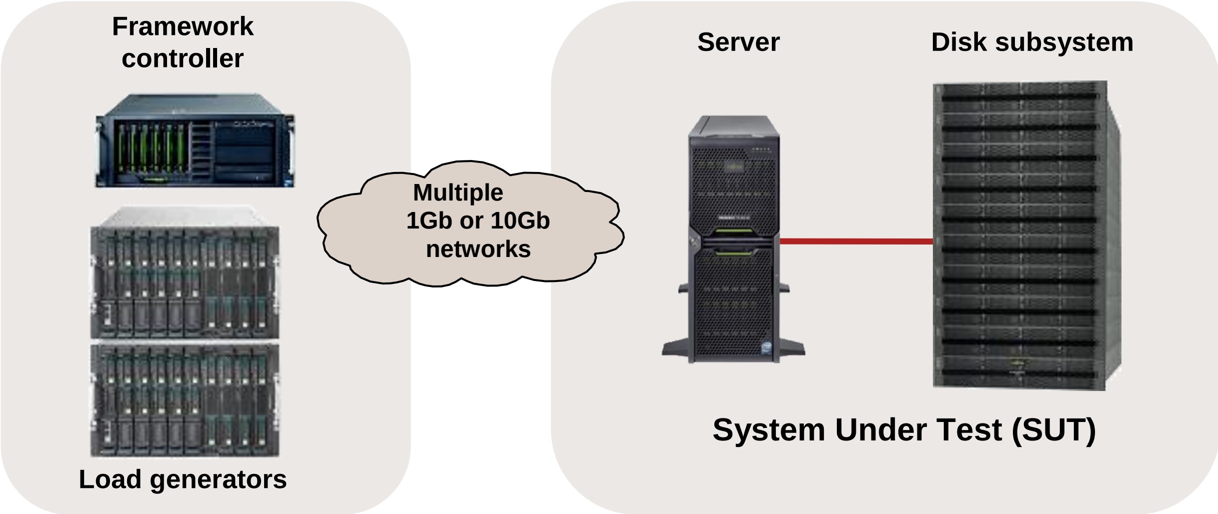
The progression of the vServCon scores for the tile numbers provides useful information about the scaling behavior of the “System under Test”.

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	Intel® Xeon® Processor E7 v3 Family
Memory	2 TB: 64 × 32GB (2x16GB) 2Rx4 DDR4-2133 R ECC
Network interface	1 × dual port 1GbE adapter 1 × dual port 10GbE server adapter
Disk subsystem	1 × dual-channel FC-Controller Emulex LPe16002 LINUX/LIO basiertes Flash Storage System
Software	
Operating system	VMware ESX 6.0.0 Build 2724185

Load generator (incl. Framework controller)	
Hardware (Shared)	
Enclosure	PRIMERGY BX900
Hardware	
Model	18 × PRIMERGY BX920 S1 server blades
Processor	2 × Xeon X5570
Memory	12 GB
Network interface	3 × 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 × logical CPU
Memory	512 MB
Network interface	2 × 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 PRIMEQUEST eight-socket systems dealt with here are based on processors of the Intel® Xeon® Processor E7 v3 Family. 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.

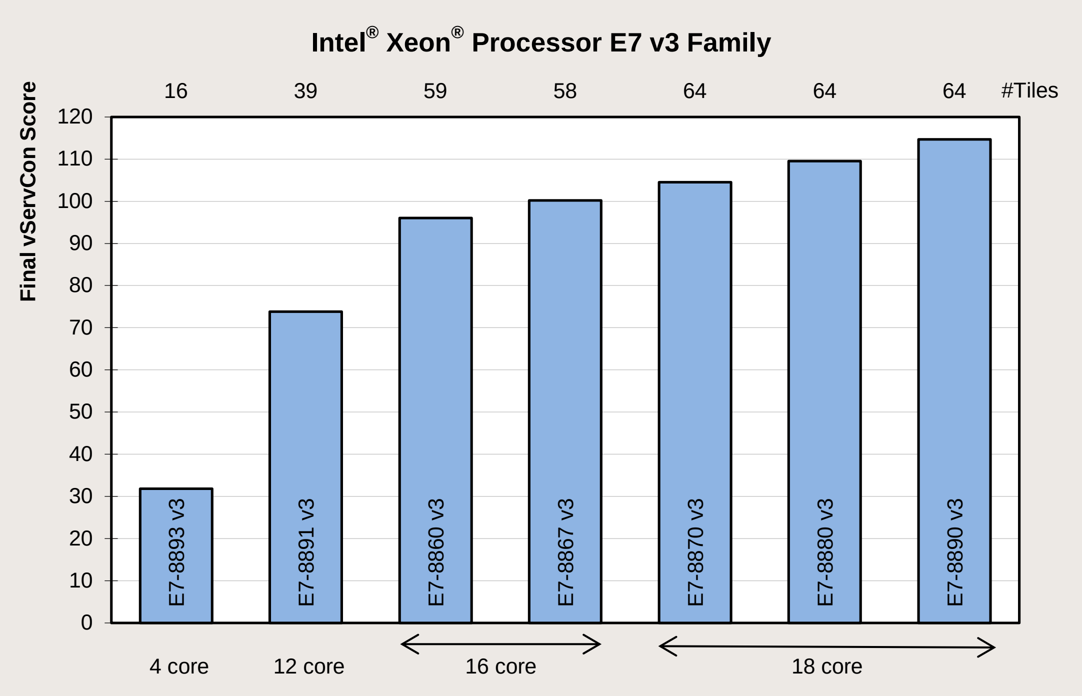
Processor			Score	#Tiles
Intel® Xeon® Processor E7 v3 Family	4 Cores Hyper-Threading, Turbo-Mode	E7-8893 v3	31.8	16
	10 Cores Hyper-Threading, Turbo-Mode	E7-8891 v3	73.8	39
	16 Cores Hyper-Threading, Turbo-Mode	E7-8860 v3	96.0	59
		E7-8867 v3	100.2	58
	18 Cores Hyper-Threading, Turbo-Mode	E7-8870 v3	104.5	64
		E7-8880 v3	109.5	64
		E7-8890 v3	114.7	64

These PRIMEQUEST eight-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 40.5% higher virtualization performance can be achieved (measured in vServCon score in their maximum configuration).

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. 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.

More information about the topic "Memory Performance" can be found in the White Paper Memory performance of Xeon E7 v3 (Haswell-EX)-based systems.

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

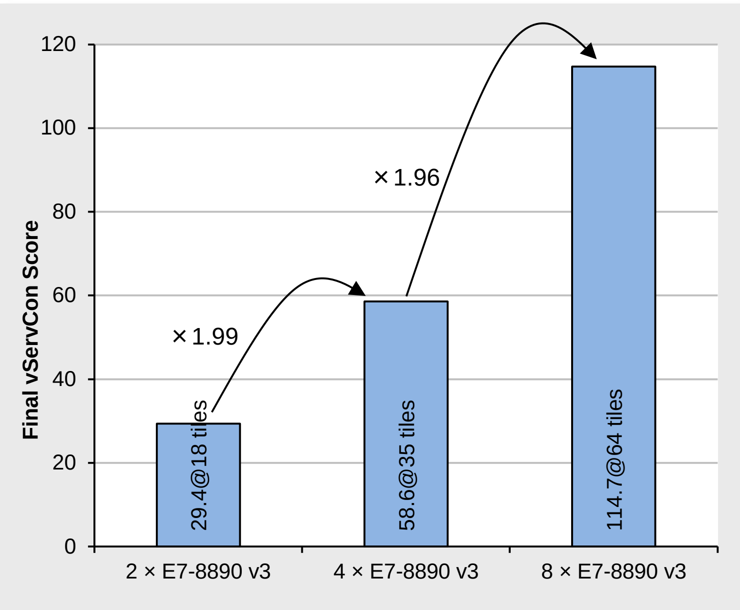


The Xeon E7-8893 v3 as the processor with four cores only makes the start.

An increase in performance is achieved by the processor with twelve cores (Xeon E7-8891 v3).

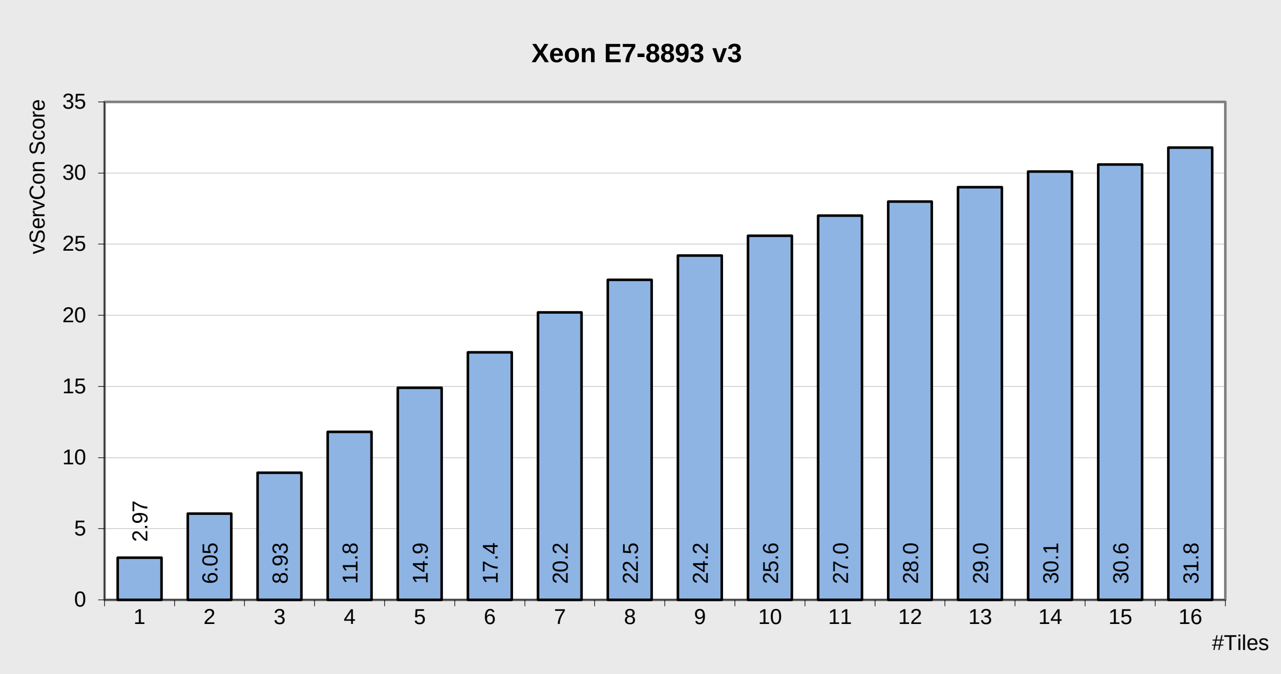
The group of processors with 18 cores, which achieves a higher performance than the 16-core processors, is to be found at the upper end of the performance scale.

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



Until now we have looked at the virtualization performance of a fully configured system. However, with a server with eight sockets the question also arises as to how good performance scaling is from two to four or eight 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 1.99 or 1.96. When operated with four or eight processors, the system thus achieves almost twice the performance as with two or four processors, as is illustrated in the diagram opposite using the processor version Xeon E7-8890 v3 as an example.

The next diagram illustrates the virtualization performance for increasing numbers of VMs based on the Xeon E7-8893 v3 (4-Core) processor.



In addition to the increased number of physical cores, Hyper-Threading, which is supported by all Xeon E7 processors, 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 scaling curves for the number of tiles as seen in the previous diagram are specifically for systems with Hyper-Threading. 32 physical and thus 64 logical cores are available with the Xeon E7-8893 v3 processors; approximately four of them are used per tile (see Benchmark description). This means that a parallel use of the same physical cores by several VMs is avoided up to a maximum of about eight tiles. That is why the performance curve in this range scales almost ideal. For the quantities above the growth is flatter up to CPU full utilization.

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 E7-8893 v3 situation with 48 application VMs (16 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 2.97 to 32.8/16=2.05, i.e. to 45%. 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.

VMmark V2

Benchmark description

VMmark V2 is a benchmark developed by VMware to compare server configurations with hypervisor solutions from VMware regarding their suitability for server consolidation. In addition to the software for load generation, the benchmark consists of a defined load profile and binding regulations. The benchmark results can be submitted to VMware and are published on their Internet site after a successful review process. After the discontinuation of the proven benchmark “VMmark V1” in October 2010, it has been succeeded by “VMmark V2”, which requires a cluster of at least two servers and covers data center functions, like Cloning and Deployment of virtual machines (VMs), Load Balancing, as well as the moving of VMs with vMotion and also Storage vMotion.

In addition to the “Performance Only” result, it is also possible from version 2.5 of VMmark to alternatively measure the electrical power consumption and publish it as a “Performance with Server Power” result (power consumption of server systems only) and/or “Performance with Server and Storage Power” result (power consumption of server systems and all storage components).

VMmark V2 is not a new benchmark in the actual sense. It is in fact a framework that consolidates already established benchmarks, as workloads in order to simulate the load of a virtualized consolidated server environment. Three proven benchmarks, which cover the application scenarios mail server, Web 2.0, and e-commerce were integrated in VMmark V2.

Application scenario	Load tool	# VMs
Mail server	LoadGen	1
Web 2.0	Olio client	2
E-commerce	DVD Store 2 client	4
Standby server	(IdleVMTest)	1

Each of the three application scenarios is assigned to a total of seven dedicated virtual machines. Then add to these an eighth VM called the “standby server”. These eight VMs form a “tile”. Because of the performance capability of the underlying server hardware, it is usually necessary to have started several identical tiles in parallel as part of a measurement in order to achieve a maximum overall performance.

A new feature of VMmark V2 is an infrastructure component, which is present once for every two hosts. It measures the efficiency levels of data center consolidation through VM Cloning and Deployment, vMotion and Storage vMotion. The Load Balancing capacity of the data center is also used (DRS, Distributed Resource Scheduler).

The result of VMmark V2 for test type „Performance Only“ is a number, known as a “score”, which provides information about the performance of the measured virtualization solution. The score reflects the maximum total consolidation benefit of all VMs for a server configuration with hypervisor and is used as a comparison criterion of various hardware platforms.

This score is determined from the individual results of the VMs and an infrastructure result. Each of the five VMmark V2 application or front-end VMs provides a specific benchmark result in the form of application-specific transaction rates for each 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 dimensionless performance values are then averaged geometrically and finally added up for all VMs. This value is included in the overall score with a weighting of 80%. The infrastructure workload is only present in the benchmark once for every two hosts; it determines 20% of the result. The number of transactions per hour and the average duration in seconds respectively are determined for the score of the infrastructure workload components.

In addition to the actual score, the number of VMmark V2 tiles is always specified with each VMmark V2 score. The result is thus as follows: “Score@Number of Tiles”, for example “4.20@5 tiles”.

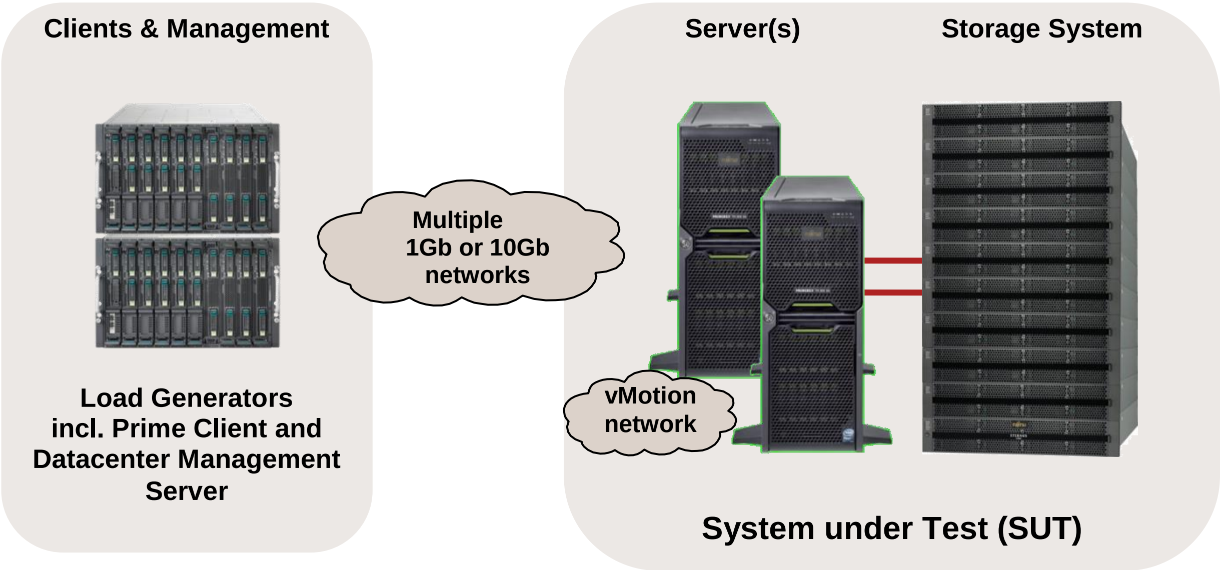
In the case of the two test types “Performance with Server Power” and “Performance with Server and Storage Power” a so-called “Server PPKW Score” and “Server and Storage PPKW Score” is determined, which is the performance score divided by the average power consumption in kilowatts (PPKW = performance per kilowatt (KW)).

The results of the three test types should not be compared with each other.

A detailed description of VMmark V2 is available in the document Benchmark Overview VMmark V2.

Benchmark environment

The measurement set-up is symbolically illustrated below:



System Under Test (SUT)	
Hardware	
Number of servers	2
Model	PRIMEQUEST 2800B2
Processor	8 × Xeon E7-8890 v3
Memory	2048 GB: 64 × 32GB (2x16GB) 2Rx4 DDR4-2133 R ECC
Network interface	2 × Dual port Emulex OCe14102 10GbE Adapter
Disk subsystem	Dual port PFC EP LPe16002 4 × PRIMERGY RX300 S8 configured as Fibre Channel target Details see disclosures
Software	
BIOS	Version 1.22
BIOS settings	See details
Operating system	VMware ESXi 6.0.0b Build 2809209
Operating system settings	ESXi settings: see details

Datacenter Management Server (DMS)	
Hardware (Shared)	
Enclosure	PRIMERGY BX600
Network Switch	1 × PRIMERGY BX600 GbE Switch Blade 30/12
Hardware	
Model	1 × server blade PRIMERGY BX620 S5
Processor	2 × Xeon X5570
Memory	24 GB
Network interface	6 × 1 Gbit/s LAN
Software	
Operating system	VMware ESXi 5.1.0 Build 799733

Datacenter Management Server (DMS)VM	
Hardware	
Processor	4 × logical CPU
Memory	10 GB
Network interface	2 × 1 Gbit/s LAN
Software	
Operating system	Microsoft Windows Server 2008 R2 Enterprise x64 Edition

Prime Client	
Hardware (Shared)	
Enclosure	PRIMERGY BX600
Network Switch	1 × PRIMERGY BX600 GbE Switch Blade 30/12
Hardware	
Model	1 × server blade PRIMERGY BX620 S5
Processor	2 × Xeon X5570
Memory	12 GB
Network interface	6 × 1 Gbit/s LAN
Software	
Operating system	Microsoft Windows Server 2008 Enterprise x64 Edition SP2

Load generator	
Hardware	
Model	3 × PRIMERGY RX600 S6 1 × PRIMERGY RX500 S7
Processor	4 × Xeon E7-4870 (PRIMERGY RX600 S6) 4 × Xeon E5-4650 (PRIMERGY RX500 S7)
Memory	PRIMERGY RX600 S6: 512 GB PRIMERGY RX500 S7: 256 GB
Network interface	PRIMERGY RX600 S6: 2 × 10 Gbit/s LAN PRIMERGY RX500 S7: 2 × 10 Gbit/s LAN
Software	
Operating system	VMware ESX 4.1.0 U2 Build 502767
Load generator VM (per tile 1 load generator VM)	
Hardware	
Processor	4 × logical CPU
Memory	4 GB
Network interface	1 × 10 Gbit/s LAN
Software	
Operating system	Microsoft Windows Server 2008 Enterprise x64 Edition SP2

Details	
See disclosure	http://www.vmware.com/a/assets/vmmark/pdf/2015-05-05-Fujitsu-PRIMEQUEST2800E2-40.pdf

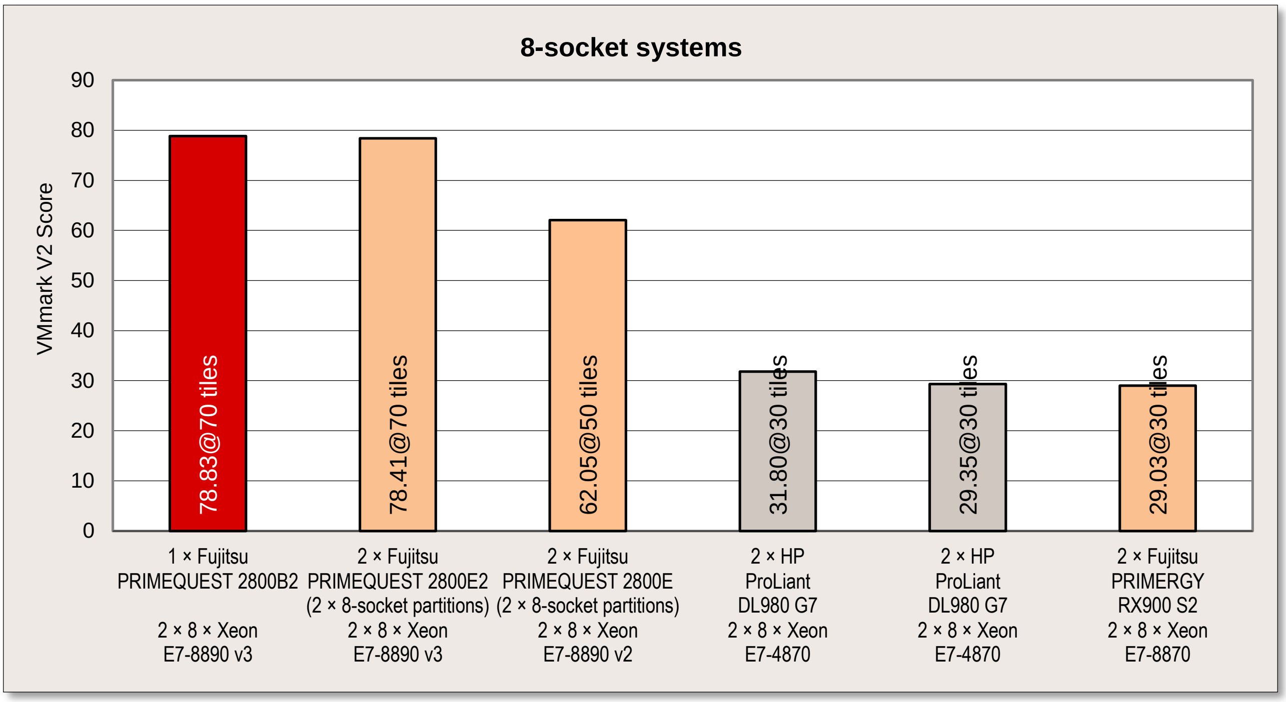
Some components may not be available in all countries or sales regions.

Benchmark results

On August 4, 2015 Fujitsu achieved with two PRIMEQUEST 2800B2 systems with Xeon E7-8890 v3 processors and VMware ESXi 6.0.0 a VMmark V2 score of “78.83@70 tiles” in a system configuration with a total of 2 × 144 processor cores and when using two identical servers/partitions in the “System under Test” (SUT). With this result the PRIMEQUEST 2800B2 is in the official VMmark V2 ranking the most powerful 8-socket server worldwide as well as the most powerful server worldwide in a “matched pair” configuration consisting of two identical hosts (valid as of benchmark results publication date).

All comparisons for the competitor products reflect the status of 4th August 2015. The current VMmark V2 results as well as the detailed results and configuration data are available at http://www.vmware.com/a/vmmark/.

The diagram shows the result of the PRIMEQUEST 2800B2 in comparison with all 8-socket systems.



The table opposite shows the difference in the score (in %) between the Fujitsu system and the other 8-socket systems.

8-socket systems	VMmark V2 score	Difference
Fujitsu PRIMEQUEST 2800B2	78.83@70 tiles	
Fujitsu PRIMEQUEST 2800E2	78.41@70 tiles	0.54%
Fujitsu PRIMEQUEST 2800E	62.05@50 tiles	27.04%
HP ProLiant DL980 G7	31.80@30 tiles	147.89%
HP ProLiant DL980 G7	29.35@30 tiles	168.59%
Fujitsu PRIMERGY RX900 S2	29.03@30 tiles	171.55%

The processors used, which with a good hypervisor setting could make optimal use of their processor features, were the essential prerequisites for achieving all PRIMEQUEST 2800B2 result. These features include Hyper-Threading. All this has a particularly positive effect during virtualization.

All VMs, their application data, the host operating system as well as additionally required data were on a powerful Fibre Channel disk subsystem. As far as possible, the configuration of the disk subsystem takes the specific requirements of the benchmark into account. The use of flash technology in the form of SAS SSDs and PCIe-SSDs in the powerful Fibre Channel disk subsystem resulted in further advantages in response times of the storage medium used.

The network connection to the load generators was implemented via 10Gb LAN ports. The infrastructure-workload connection between the hosts was by means of 1Gb LAN ports.

All the components used were optimally attuned to each other.

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 processor 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 12 times larger than the total of all last-level processor 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.

Type	Execution	Bytes per step	Floating-point calculation per step
COPY	a(i) = b(i)	16	0
SCALE	a(i) = q × b(i)	16	1
SUM	a(i) = b(i) + c(i)	24	1
TRIAD	a(i) = b(i) + q × 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 processors influence the arithmetic calculations.

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

Benchmark environment

System Under Test (SUT)	
Hardware	
Model	PRIMEQUEST 2800B2
Processor	8 processors of Intel® Xeon® Processor E7-8800 v3 Product Family
Memory	64 × 32GB (2x16GB) 2Rx4 DDR4-2133 R ECC
Software	
BIOS settings	EnergyPerformance = Performance
Operating system	Red Hat Enterprise Linux Server release 6.6
Operating system settings	Transparent Huge Pages inactivated
Compiler	Intel C++ Composer XE 2015 for Linux
Benchmark	STREAM version 5.10

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

Benchmark results

Processor	Memory Frequency [MHz]	Max. Memory Bandwidth [GB/s]	Cores	Processor Frequency [GHz]	Number of Processors	TRIAD [GB/s]
Xeon E7-8893 v3	1600	102	4	3.20	8	397
Xeon E7-8891 v3	1600	102	10	2.80	8	436
Xeon E7-8860 v3	1600	102	16	2.20	8	427
Xeon E7-8867 v3	1600	102	16	2.50	8	442
Xeon E7-8870 v3	1600	102	18	2.10	8	443
Xeon E7-8880 v3	1600	102	18	2.30	8	442
Xeon E7-8890 v3	1600	102	18	2.50	8	442

Literature

PRIMEQUEST Servers

http://ts.fujitsu.com/primequest

PRIMEQUEST 2800B2

This White Paper:

- http://docs.ts.fujitsu.com/dl.aspx?id=e23ba4d1-936d-4e2d-98bd-d63ed173ec4a
- http://docs.ts.fujitsu.com/dl.aspx?id=135a7dc2-78c8-48ae-badc-96110c7bb12a

Data sheet
http://docs.ts.fujitsu.com/dl.aspx?id=bb8a1f75-7d1b-4bb8-bd51-73edd336d9d5

PRIMEQUEST Performance

http://www.fujitsu.com/fts/x86-server-benchmarks

Performance of Server Components

http://www.fujitsu.com/fts/products/computing/servers/mission-critical/benchmarks/x86-components.html

Memory performance of Xeon E7 v3 (Haswell-EX)-based systems
http://docs.ts.fujitsu.com/dl.aspx?id=324913b1-3a67-4ee7-a809-c01bc9a6d00b

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

Disk I/O: Performance of storage media and RAID controllers

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

Information about Iometer
http://www.iometer.org

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

STREAM

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

VMmark V2

Benchmark Overview VMmark V2
http://docs.ts.fujitsu.com/dl.aspx?id=2b61a08f-52f4-4067-bbbf-dc0b58bee1bd

VMmark V2
http://www.vmmark.com

vServCon

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

Contact

FUJITSU

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

PRIMEQUEST Product Marketing

mailto:Primergy-PM@ts.fujitsu.com

PRIMERGY Performance and Benchmarks

mailto:primergy.benchmark@ts.fujitsu.com

2015-11-18 WW EN