Fujitsu Server PRIMEQUEST
Performance Report
PRIMEQUEST 4400E

This document provides an overview of benchmarks executed on the Fujitsu Server PRIMEQUEST 4400E.

Explaines PRIMEQUEST 4400E performance data in comparison to other PRIMEQUEST models. In addition to the benchmark results, the explanation for each benchmark and benchmark environment are also included.

Version
1.0 2023-09-12

Contents

Technical data

PRIMEQUEST 4400E

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 4400E
Form factor	Rack server
Chipset	Intel C741
Number of sockets	2 per SB
Number of configurable processors	1 - 4
Processor type	4th Generation Intel Xeon Scalable Processors Family
Number of memory slots	64 (16 per processor)
Maximum memory configuration	16,384 GB
Maximum number of internal storage disks	2.5 inch: 32
Maximum number of PCI slots	PCI-Express 5.0, (x8 lane): 8, (x16 lane): 6 (Low Profile)

PRIMEQUEST 4400E processor

Processor model	Type	Number of cores	Number of threads	L3 Cache [MB]	UPI speed [GT/s]	Rated frequency [GHz]	Maximum turbo frequency [GHz]	Maximum memory frequency [MHz]	TDP [W]
Xeon Platinum 8490H	XCC	60	120	112.5	16	1.90	3.50	4,800	350
Xeon Platinum 8468H	XCC	48	96	105	16	2.10	3.80	4,800	330
Xeon Platinum 8460H	XCC	40	80	105	16	2.20	3.80	4,800	330
Xeon Platinum 8450H	XCC	28	56	75	16	2.00	3.50	4,800	250
Xeon Platinum 8444H	XCC	16	32	45	16	2.90	4.00	4,800	270
Xeon Gold 6448H	MCC	32	64	60	16	2.40	4.10	4,800	250
Xeon Gold 6434H	MCC	8	16	22.5	16	3.70	4.10	4,800	195
Xeon Gold 6418H	MCC	24	48	60	16	2.10	4.00	4,800	185
Xeon Gold 6416H	MCC	18	36	45	16	2.20	4.20	4,800	165

All processors that can be ordered with PRIMEQUEST 4400E support Intel Turbo Boost Technology 2.0. This technology allows you to operate the processor with higher frequencies than the rated frequency. The "maximum turbo frequency" listed in the processor list above is the theoretical maximum frequency when there is only one active core per processor. The maximum frequency that can actually be achieved depends on the number of active cores, current consumption, power consumption, and processor temperature.

As a general rule, Intel does not guarantee that maximum turbo frequencies will be achieved. This is related to manufacturing tolerances, and the performance of each individual processor model varies from each other. The range of difference covers the range including all of the rated frequency and the maximum turbo frequency.

The turbo function can be set in the BIOS option. Generally, Fujitsu always recommends leaving the [Turbo Mode] option set at the standard setting [Enabled], as performance is substantially increased by the higher frequencies. However, the Turbo Mode frequency depends on the operating conditions mentioned above and is not always guaranteed. The turbo frequency fluctuates in applications where AVX instructions are used intensively and the number of instructions per clock is large. If you need stable performance or want to reduce power consumption, it may be beneficial to set the [Turbo Mode] option to [Disabled] to disable the turbo function.

The processor with the suffix means it is optimized for the following feature.

Suffix	Workload
H	**DB/Analytics** Data analytics and big data usages

Please refer to the below URL for details.

https://www.intel.com/content/www/us/en/support/articles/000059657/processors/intel-xeon-processors.html

Memory modules									
Type	Capacity [GB]	Number of ranks	Bit width of the memory chips	Frequency [MHz]	3DS	Load Reduced	Registered	NVDIMM	ECC
32GB (1x32GB) 1Rx4 DDR5-4800 R ECC	32	1	4	4,800			✓		✓
64GB (1x64GB) 2Rx4 DDR5-4800 R ECC	64	2	4	4,800			✓		✓
128GB (1x128GB) 4Rx4 DDR5-4800 R 3DS ECC	128	4	4	4,800	✓		✓		✓
256GB (1x256GB) 8Rx4 DDR5-4800 R 3DS ECC	256	8	4	4,800	✓		✓		✓

Power supplies		Maximum number
Modular redundant PSU	2,200W platinum PSU	4
	2,600W platinum PSU	4

Includes components that will be supported after the system release. Also, some components may not be available in all countries or sales regions.

Detailed technical information is available in the data sheet of PRIMEQUEST 4400E.

SPEC CPU2017

Benchmark description

SPEC CPU2017 is a benchmark which measures the system efficiency with integer and floating-point operations. It consists of an integer test suite (SPECrate 2017 Integer, SPECspeed 2017 Integer) containing 10 applications and a floating-point test suite (SPECrate 2017 Floating Point, SPECspeed 2017 Floating Point) containing 14 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.

SPEC CPU2017 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.

SPEC CPU2017 contains two different performance measurement methods. The first method (SPECspeed 2017 Integer or SPECspeed 2017 Floating Point) determines the time which is required to process a single task. The second method (SPECrate 2017 Integer or SPECrate 2017 Floating Point) 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." They differ in the use of compiler optimization. When publishing the results, the base values are always used and the peak values are optional.

Benchmark	Number of single benchmarks	Arithmetics	Type	Compiler optimization	Measurement result
SPECspeed2017_int_peak	10	integer	peak	aggressive	Speed
SPECspeed2017_int_base	10	integer	base	conservative	
SPECrate2017_int_peak	10	integer	peak	aggressive	Throughput
SPECrate2017_int_base	10	integer	base	conservative	
SPECspeed2017_fp_peak	10	floating point	peak	aggressive	Speed
SPECspeed2017_fp_base	10	floating point	base	conservative	
SPECrate2017_fp_peak	13	floating point	peak	aggressive	Throughput
SPECrate2017_fp_base	13	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 favor of the lower individual results. "Normalized" means that the measurement is how fast is the test system compared to a reference system. For example, value "1" was defined for the SPECspeed2017_int_base, SPECrate2017_int_base, SPECspeed2017_fp_base, and SPECrate2017_fp_base results of the reference system. A SPECspeed2017_int_base value of 2 means that the measuring system has handled this benchmark twice as fast as the reference system. A SPECrate2017_fp_base value of 4 means that the measuring system has handled this benchmark about 4/[# base copies] times faster than the reference system. "# base copies" specifies how many parallel instances of the benchmark have been executed.

Not every SPEC CPU2017 measurement is submitted by Fujitsu for publication at SPEC. This is why the SPEC web pages do not have every result. As Fujitsu archives the log files for all measurements, it is possible to prove the correct implementation of the measurements at any time.

Benchmark environment

System Under Test (SUT)	
Hardware	
• Model	PRIMEQUEST 4400E
• Processor	4 x 4th Generation Intel Xeon Scalable Processors Family or 2 x 4th Generation Intel Xeon Scalable Processors Family
• Memory	32 x 64GB (1x64GB) 2Rx4 DDR5-4800 R ECC (4CPU configuration) or 16 x 64GB (1x64GB) 2Rx4 DDR5-4800 R ECC (2CPU configuration)
Software	
• BIOS settings	SPECspeed2017_int_base: • Override OS Energy Performance = Enabled • Energy Performance = Balanced Performance • CPU Performance Boost = Aggressive • SNC(Sub NUMA) = Enable SNC4 (Enable SNC2 when MCC are installed) • FAN Control = Full SPECSpeed2017_fp_base: • Hyper Threading = Disabled • Package C State limit = C0 • LLC Prefetch = Enabled • Homeless Prefetch = Enabled • DBP-F = Enabled • CPU Performance Boost = Aggressive • FAN Control = Full SPECrate2017_int_base: • DCU Streamer Prefetcher = Disabled • Package C State limit = C0 • LLC Dead Line Alloc = Disabled • CPU Performance Boost = Aggressive • SNC(Sub NUMA) =Enable SNC4(Enable SNC2 when MCC are installed) • FAN Control = Full SPECrate2017_fp_base: • Hyper Threading = Disabled (Enabled when MCC are installed) • Package C State limit = C0 • SNC (Sub NUMA) =Enable SNC4 (Enable SNC2 when MCC are installed) • CPU Performance Boost = Aggressive • FAN Control = Full
• Operating system	SUSE Linux Enterprise Server 15 SP4 5.14.21-150400.22-default
• Operating system settings	Default
• Compiler	C/C++: Version 2023.0 of Intel C/C++ Compiler for Linux Fortran: Version 2023.0 of Intel Fortran Compiler for Linux

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 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.

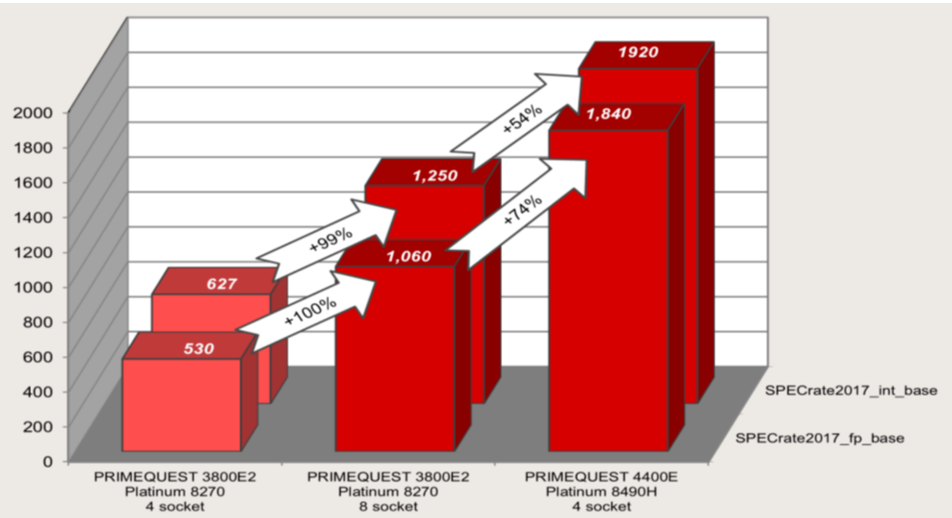
The results with "est." are the estimated values.

Processor model	Number of cores	Number of processors	SPECrate2017_int_base	SPECrate2017_fp_base
4CPU configuration				
Xeon Platinum 8490H	60	4	**1,920**	**1,840**
Xeon Platinum 8468H	48	4	**1,660**	**1,720**
Xeon Platinum 8460H	40	4	**1,480**	**1,640**
Xeon Platinum 8450H	28	4	**942**	**1,200**
Xeon Platinum 8444H	16	4	**650**	**907**
Xeon Gold 6448H	32	4	**1,230**	**1,400**
Xeon Gold 6434H	8	4	**397**	**567**
Xeon Gold 6418H	24	4	**881**	**1,110**
Xeon Gold 6416H	18	4	**665**	**892**

Processor model	Number of cores	Number of processors	SPECrate2017_int_base	SPECrate2017_fp_base
Xeon Platinum 8490H	60	4	-	**357**
Xeon Gold 6416H	16	4	**16.0**	-

The following graphs compare the throughputs of PRIMEQUEST 4400E and their older models, PRIMEQUEST 3800E2, with maximum performance configurations.

The performance of the PRIMEQUEST 3800E2 has been improved by more than 3 times compared to 4 sockets and by more than 1.6 times compared to 8 sockets.



SPECrate2017: Comparison of PRIMEQUEST 3800E2 and PRIMEQUEST 4400E

STREAM

Benchmark description

STREAM is a synthetic benchmark that has been used for many years to determine memory throughput and 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. This provides optimal load distribution for the available processor cores.

In the STREAM benchmark, a data area consisting of 8-byte elements is continuously copied to four operation types. Arithmetic operations are also performed on operation types other than COPY.

Arithmetics type	Arithmetics	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.

In this chapter, throughputs are indicated as a power of 10. (1 GB/s = 10^9 Byte/s)

Benchmark environment

System Under Test (SUT)	
Hardware	
• Model	PRIMEQUEST 4400E
• Processor	4 x 4th Generation Intel Xeon Scalable Processors Family or 2 x 4th Generation Intel Xeon Scalable Processors Family
• Memory	32 x 64GB (1x64GB) 2Rx4 DDR5-4800 R ECC (4CPU configuration) or 16 x 64GB (1x64GB) 2Rx4 DDR5-4800 R ECC (2CPU configuration)
Software	
• BIOS settings	• DCU Streamer Prefetcher = Disabled • SNC(Sub NUMA) = Enable SNC4 (Enable SNC2 when MCC type installed)) • Intel Virtualization Technology = Disabled • LLC Dead Line Alloc = Disabled • Stale Atos = Enabled
• Operating system	SUSE Linux Enterprise Server 15 SP4 5.14.21-150400.22-default
• Operating system settings	Default
• Compiler	C/C++: Version 2023.0 of Intel C/C++ Compiler for Linux
• Benchmark	STREAM Version 5.10

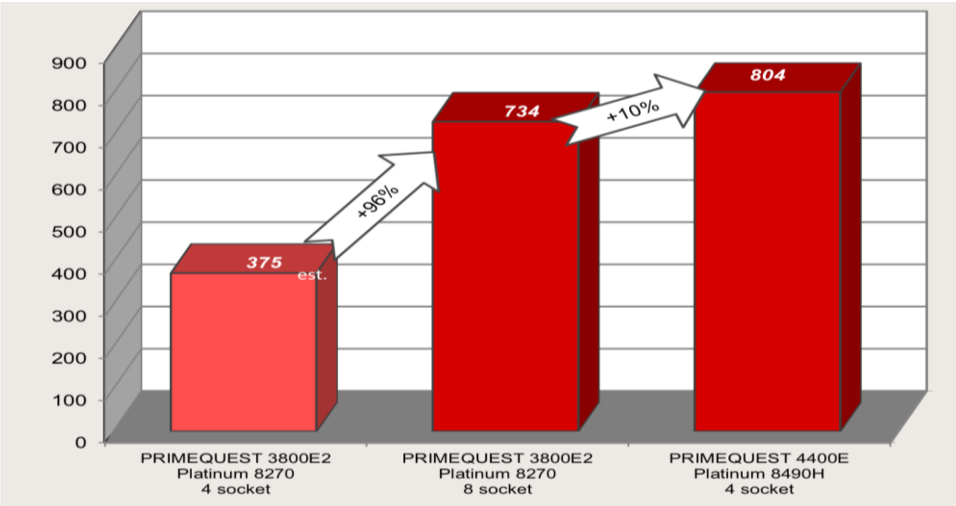
Benchmark results

The results with "est." are the estimated values.

Processor	Memory frequency [MHz]	Maximum memory bandwidth [GB/s]	Number of cores	Rated frequency [GHz]	Number of processors	TRIAD [GB/s]
4CPU configuration						
Xeon Platinum 8490H	4,800	307	60	1.90	4	**805**
Xeon Platinum 8468H	4,800	307	48	2.00	4	**808**
Xeon Platinum 8460H	4,800	307	40	2.00	4	**817**
Xeon Platinum 8450H	4,800	307	28	2.10	4	**766**
Xeon Platinum 8444H	4,800	307	16	2.80	4	**643**
Xeon Gold 6448H	4,800	307	32	2.40	4	**701**
Xeon Gold 6434H	4,800	307	8	3.70	4	**434**
Xeon Gold 6418H	4,800	307	24	2.10	4	**679**
Xeon Gold 6416H	4,800	307	18	2.20	4	**579**

The following graphs compare the throughputs of PRIMEQUEST 4400E and their older models, PRIMEQUEST 3800E2, with maximum performance configurations.

The performance of the PRIMEQUEST 3800E2 has been improved by more than 2 times compared to 4 sockets and by more than 1.1 times compared to 8 sockets.



STREAM: Comparison of PRIMEQUEST 3800E2 and PRIMEQUEST 4400E

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. The description can be found in the following 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 x 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 x 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. In other words, if n is doubled, the measurement time will be approximately eight times longer. The size of n also has an influence on the measurement result itself. As n increases, the measured value asymptotically approaches its 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. he 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: 1 billion floating point operations/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 Rated 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/. This requires using an HPL-based LINPACK version (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

Benchmark environment

System Under Test (SUT)	
Hardware	
• Model	PRIMEQUEST 4400E
• Processor	4 x 4th Generation Intel Xeon Scalable Processors Family or 2 x 4th Generation Intel Xeon Scalable Processors Family
• Memory	32 x 64GB (1x64GB) 2Rx4 DDR5-4800 R ECC (4CPU configuration) or 16 x 64GB (1x64GB) 2Rx4 DDR5-4800 R ECC (2CPU configuration)
Software	
• BIOS settings	• HyperThreading = Disabled • CPU Performance Boost = Agressive • Fan Control = Full
• Operating system	SUSE Linux Enterprise Server 15 SP4 5.14.21-150400.22-default
• Operating system settings	Default
• Compiler	C/C++: Version 2023.0 of Intel C/C++ Compiler for Linux
• Benchmark	Intel Optimized MP LINPACK Benchmark for Clusters

Benchmark results

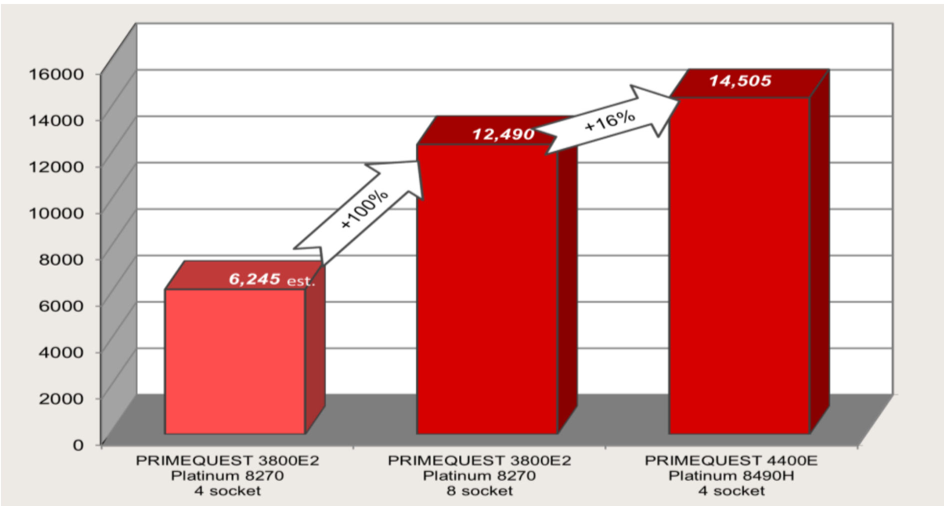
The results with "est." are the estimated values.

Processor	Number of cores	Rated frequency [GHz]	Number of processors	Rpeak [GFlops]	Rmax [GFlops]	Effic.
4CPU configuration						
Xeon Platinum 8490H	60	1.90	4	14,592	**14,505**	99%
Xeon Platinum 8468H	48	2.10	4	12,902	**12,656**	98%
Xeon Platinum 8460H	40	2.20	4	11,264	**11,872**	105%
Xeon Platinum 8450H	28	2.00	4	7,168	**7,844**	109%
Xeon Platinum 8444H	16	2.90	4	5,939	**6,042**	102%
Xeon Gold 6448H	32	2.40	4	9,830	**10,101**	103%
Xeon Gold 6434H	8	3.70	4	3,789	**3,743**	99%
Xeon Gold 6418H	24	2.10	4	6,451	**7,090**	110%
Xeon Gold 6416H	18	2.20	4	5,069	**5,586**	110%

Rpeak values in the table above were calculated by the base frequency of each processor. Since we enabled Turbo mode in the measurements, the average Turbo frequency exceeded the base frequency for some processors.

As explained in the section "Technical Data," Intel generally does not guarantee that the maximum turbo frequency can be reached in the processor models due to manufacturing tolerances. A further restriction applies for workloads, such as those generated by LINPACK, with intensive use of AVX instructions and a high number of instructions per clock unit. Here the frequency of a core can also be limited if the upper limits of the processor for power consumption and temperature are reached before the upper limit for the current consumption. This can result in the achievement of a lower performance with turbo mode than without turbo mode. In such a case, disable the turbo function in the BIOS option.

The following graphs compare the throughputs of PRIMEQUEST 4400E and their older models, PRIMEQUEST 3800E2, with maximum performance configurations.
The performance of the PRIMEQUEST 3800E2 has been improved by more than 2 times compared to 4 sockets and by more than 1.1 times compared to 8 sockets.



LINPACK: Comparison of PRIMEQUEST 3800E2 and PRIMEQUEST 4400E

Disk I/O: Performance of storage media

Benchmark description

Performance measurements of disk subsystems for PRIMEQUEST servers 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 as follows.

- Random access / sequential access ratio
- Read / write access ratio
- Block size (kiB)
- Queue Depth (number of IO requests to issue at one time)

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 [kiB]	Application
		read	write		
Filecopy	Random	50%	50%	64	Copying files
Fileserver	Random	67%	33%	64	Fileserver
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 files

In order to model applications that access in parallel with a different load intensity the Queue Depth is increased from 1 to 512 (in steps to the power of two).

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

The main measurement items are as follows.

- Throughput [MiB/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 following formula.

Data throughput [MiB/s]	= Transaction rate [IO/s] x Block size [MiB]
Transaction rate [IO/s]	= Data throughput [MiB/s] / Block size [MiB]

In this section, a power of 10 (1 TB = 10^{12} bytes) is used to indicate the capacity of the hard storage medium, and a power of 2 (1 MiB / s = 2^{20} bytes) is used to indicate the capacity of other media, file size, block size, and throughput.

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 section apply for the hardware and software components listed below.

System Under Test (SUT)

Hardware

Controller: PRAID EP540i

Storage media	Category	Drive name
HDD	SAS HDD (SAS 12Gbps, 15k rpm) [512n]	ST300MP0006 ST600MP0006
	SAS HDD (SAS 12Gbps, 10k rpm) [512e]	AL15SEB18EQ AL15SEB24EQ
	SAS HDD (SAS 12Gbps, 10k rpm) [512n]	AL15SEB060N AL15SEB120N
SSD	SAS SSD (SAS 12Gbps, Write Intensive)	XS400ME70084 XS800ME70084 XS1600ME70084
	SAS SSD (SAS 12Gbps, Mixed Use)	XS800LE70084 XS1600LE70084 XS3200LE70084
	SAS SSD (SAS 12Gbps, Read Intensive)	XS960SE70084 XS1920SE70084 XS3840SE70084 XS7680SE70084

Controller: PRAID EP680i

Storage media	Category	Drive name
SSD	PCIe SSD (Write intensive)	SSDPF21Q400GB SSDPF21Q800GB SSDPF21Q016TB
	PCIe SSD (Mixed Use)	KCM61VUL1T60 KCM61VUL3T20 KCM61VUL6T40
	PCIe SSD (Read intensive)	KCM61RUL960G KCM61RUL1T92 KCM61RUL3T84 KCM61RUL7T68

Controller:	Intel C741 Standard SATA AHCI controller		
	Storage media	Category	Drive name
	SSD	M.2 Flash module	MTFDDAV240TGA
			MTFDDAV480TGA
			MTFDDAV960TGA

Software		
Operating system		Microsoft Windows Server 2019 Standard
Benchmark version		3.0
RAID type		Type RAID 0 logical drive consisting of 1 hard disk
Stripe size		HDD: 256KiB, SSD: 64 KiB
Measuring tool		Iometer 1.1.0
Measurement area	HDD, SSD (Except M.2)	RAW file system is used. The first 32GiB of available LBA space is used for sequential access. The following 64GiB is used for random access.
	SSD (M.2)	NTFS file system is used. The first 32GiB of available LBA space is used for sequential access. The following 64GiB is used for random access.
Total number of Iometer worker		1
Alignment of Iometer accesses		Aligned to access block size

Benchmark results

The results shown here are intended to help you select the appropriate storage media under the aspect of disk-I/O performance. For this purpose, a single storage medium was measured in the configuration specified in the subsection "Benchmark environment."

Controller

The measurements were made using controllers in the table below.

Storage media	Controller name	Cache	Supported interfaces		RAID levels
			host	drive	
SSD/HDD	PRAID EP540i	-	PCIe 3.0 x8	SATA 6G SAS 12G PCIe x16	0, 1, 1E, 10, 5, 50
PCIe SSD 2.5”	PRAID EP680i	-	PCIe 4.0 x8	SATA 6G SAS 12G PCIe x16	0, 1, 1E, 10, 5, 50
M.2 Flash	C741 Standard SATA AHCI controller	-	DMI 3.0 x4	SATA 6G	-

Storage media

When selecting the type and number of storage media you can move the weighting in the direction of storage capacity, performance, security or price. The following types of HDD and SSD storage media can be used for PRIMEQUEST servers.

Model	Storage media type	interface	Form factor
2.5 inch model	HDD	SAS 12G	2.5 inch
	SSD	SAS 12G	2.5 inch
		SATA 6G	M.2
		PCIe 4.0	2.5 inch

HDDs and SSDs are operated via host bus adapters, usually RAID controllers, with a SATA or SAS interface. The interface of the RAID controller to the chipset of the system board is typically PCIe or, in the case of the integrated onboard controllers, an internal bus interface of the system board.

Of all the storage medium types SSDs offer by far the highest transaction rates for random load profiles as well as the shortest access times. In return, however, the price per gigabyte of storage capacity is substantially higher.

Cache 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.

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 PRIMEQUEST 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 mode "Performance" or "Data Protection."

The "Performance" mode ensures the best possible performance settings for the majority of the application scenarios.

Performance values

The performance values are summarized in the following tables. In each case specifically for a single storage medium and with various access types and block sizes. 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.

The table cells contain the maximum achievable values. This means that each value is the maximum achievable value of the whole range of load intensities (number 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.

Storage media performance

HDDs

Capacity [GB]	Storage device	Interface	Transactions [IO/s] Database	Fileserver	filecopy	Throughput [MiB/s] Streaming	Restore
- SAS HDD 15krpm [512n]							
300	ST300MP0006	SAS 12G	790	696	666	304	304
600	ST600MP0006	SAS 12G	736	651	601	301	300
- SAS HDD 10krpm [512e]							
1,800	AL15SEB18EQ	SAS 12G	767	631	624	255	249
2,400	AL15SEB24EQ	SAS 12G	754	620	617	264	260
- SAS HDD 10krpm [512n]							
600	AL15SEB060N	SAS 12G	698	586	600	232	232
1,200	AL15SEB120N	SAS 12G	732	604	615	230	226

SSDs

* PRIMEQUEST 4400L does not support SSDs.

Capacity [GB]	Storage device	Interface	Transactions [IO/s] Database	Fileserver	filecopy	Throughput [MiB/s] Streaming	Restore
- SAS SSD (10DWPD)							
400	XS400ME70084	SAS 12G	122,956	22,969	19,438	1,052	872
800	XS800ME70084	SAS 12G	123,848	23,784	19,435	1,052	874
1,600	XS1600ME70084	SAS 12G	123,277	23,725	19,270	1,051	884
- SAS SSD (3DWPD)							
800	XS800LE70084	SAS 12G	121,914	23,707	19,257	1,052	871
1,600	XS1600LE70084	SAS 12G	122,949	23,771	19,455	1,052	874
3,200	XS3200LE70084	SAS 12G	123,090	22,816	19,418	1,051	872
- SAS SSD (1DWPD)							
960	XS960SE70084	SAS 12G	123,014	23,678	19,424	1,052	870
1,920	XS1920SE70084	SAS 12G	123,093	23,760	19,423	1,052	874
3,840	XS3840SE70084	SAS 12G	122,810	22,949	19,406	1,051	871
7,680	XS7680SE70084	SAS 12G	123,461	22,899	19,516	1,051	880
- PCIe SSD							
400	SSDPF21Q400GB	PCIe4 x4	303,783	91,576	84,727	6,693	4,562
800	SSDPF21Q800GB	PCIe4 x4	290,266	99,852	94,882	6,738	4,512
1,600	SSDPF21Q016TB	PCIe4 x4	304,687	108,995	110,292	6,682	4,382
- PCIe SSD (3DWPD)							
1,600	KCM61VUL1T60	PCIe4 x4	272,211	49,350	47,236	6,649	2,740
3,200	KCM61VUL3T20	PCIe4 x4	314,143	72,898	75,032	6,649	4,062
6,400	KCM61VUL6T40	PCIe4 x4	305,271	67,808	71,273	6,649	3,853
- PCIe SSD (1DWPD)							
960	KCM61RUL960G	PCIe4 x4	77,623	9,719	6,428	6,633	1,400
1,920	KCM61RUL1T92	PCIe4 x4	180,706	19,204	12,678	6,649	2,730
3,840	KCM61RUL3T84	PCIe4 x4	315,657	72,526	75,132	6,649	4,048
7,680	KCM61RUL7T68	PCIe4 x4	311,548	68,020	71,191	6,649	3,853

Capacity [GB]	Storage device	Interface	Transactions [IO/s]			Throughput [MiB/s]	
			Database	Fileserver	filecopy	Streaming	Restore
- M.2 SATA SSD							
240	MTFDDAV240TGA	SATA 6G	34,363	5,680	5,730	500	353
480	MTFDDAV480TGA	SATA 6G	43,056	6,473	6,540	503	490
960	MTFDDAV960TGA	SATA 6G	50,096	6,984	7,049	505	494

Literature

PRIMEQUEST Servers

https://www.fujitsu.com/global/products/computing/servers/mission-critical/index.html

PRIMEQUEST 4400E

This Whitepaper

- https://docs.ts.fujitsu.com/dl.aspx?id=894abdf0-c011-49a2-882b-d8ac1cf512f7
- https://docs.ts.fujitsu.com/dl.aspx?id=f176128b-c037-4383-9594-fffe38a4c694

SPEC CPU2017

https://www.spec.org/osg/cpu2017

Benchmark Overview SPECcpu2017

https://docs.ts.fujitsu.com/dl.aspx?id=20f1f4e2-5b3c-454a-947f-c169fca51eb1

STREAM

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

LINPACK

The LINPACK Benchmark: Past, Present, and Future
https://www.netlib.org/utk/people/JackDongarra/PAPERS/hplpaper.pdf
TOP500
https://www.top500.org/
HPL - A Portable Implementation of the High-Performance Linpack Benchmark for Distributed-Memory Computers
https://www.netlib.org/benchmark/hpl/
Intel Math Kernel Library – LINPACK Download
https://www.intel.com/content/www/us/en/developer/articles/technical/onemkl-benchmarks-suite.html

Document change history

Version	Date	Description
1.0	2023-09-12	New: • Technical data • SPEC CPU2017, STREAM, LINPACK Measured and calculated with 4th Generation Intel Xeon Processor Scalable Family • Disk I/O Measured with 2.5 inch model

Contact

Fujitsu

Web site: https://www.fujitsu.com

PRIMEQUEST Performance and Benchmarks

mailto:fj-benchmark@dl.jp.fujitsu.com