Fujitsu Server PRIMERGY
Performance Report
PRIMERGY RX8770 M7

This document provides an overview of benchmarks executed on the Fujitsu Server PRIMERGY RX8770 M7.

Explaines PRIMERGY RX8770 M7 performance data in comparison to other PRIMERGY models. In addition to the benchmark results, the explanation for each benchmark and benchmark environment are also included.

Version
1.1
2023-10-20

Contents

Technical data

PRIMERGY RX8770 M7

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	PRIMERGY RX8770 M7
Form factor	Rack server
Chipset	Intel C741
Number of sockets	8
Number of configurable processors	8
Processor type	4th Generation Intel Xeon Scalable Processors Family
Number of memory slots	128 (16 per processor)
Maximum memory configuration	32,768 GB
Maximum number of internal storage disks	2.5 inch: 24
Maximum number of PCI slots	PCI-Express 5.0 (x16 lane): 24 (Low Profile)

Processor									
Processor model	Type	Number of cores	Number of threads	L3 Cache	UPI speed	Rated frequency	Maximum turbo frequency	Maximum memory frequency	TDP
				[MB]	[GT/s]	[GHz]	[GHz]	[MHz]	[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 8454H	XCC	32	64	82.5	16	2.10	3.40	4,800	270
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

All processors that can be ordered with PRIMERGY RX8770 M7 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
16GB (1x16GB) 1Rx8 DDR5-4800 R ECC	16	1	8	4,800			✓		✓
32GB (1x32GB) 2Rx8 DDR5-4800 R ECC	32	2	8	4,800			✓		✓
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	2,600W titanium PSU	6

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 PRIMERGY RX8770 M7.

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	PRIMERGY RX8770 M7
• Processor	8 x 4th Generation Intel Xeon Scalable Processors Family
• Memory	64 x 64GB (1x64GB) 2Rx4 DDR5-4800 R ECC
Software	
• BIOS settings	SPECspeed2017_int_base: • Adjacent Cache Line Prefetch = Disabled • LLC Dead Line Alloc = Disabled SPECSpeed2017_fp_base: • DCU IP Prefetcher = Disabled • Adjacent Cache Line Prefetch = Disabled • LLC Prefetch = Enabled • DBP-F = Enabled SPECrate2017_int_base: • DCU Streamer Prefetcher = Disabled • Adjacent Cache Line Prefetch = Disabled • CPU C1E Support = Disabled • LLC Dead Line Alloc = Disabled • SNC(Sub NUMA) =Enable SNC4 • FAN Control = Full SPECrate2017_fp_base: • Hyper Threading = Disabled • SNC (Sub NUMA) =Enable SNC4
• Operating system	SUSE Linux Enterprise Server 15 SP4 5.14.21-150400.22-default
• Operating system settings	Stack size set to unlimited using "ulimit -s unlimited"
• 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

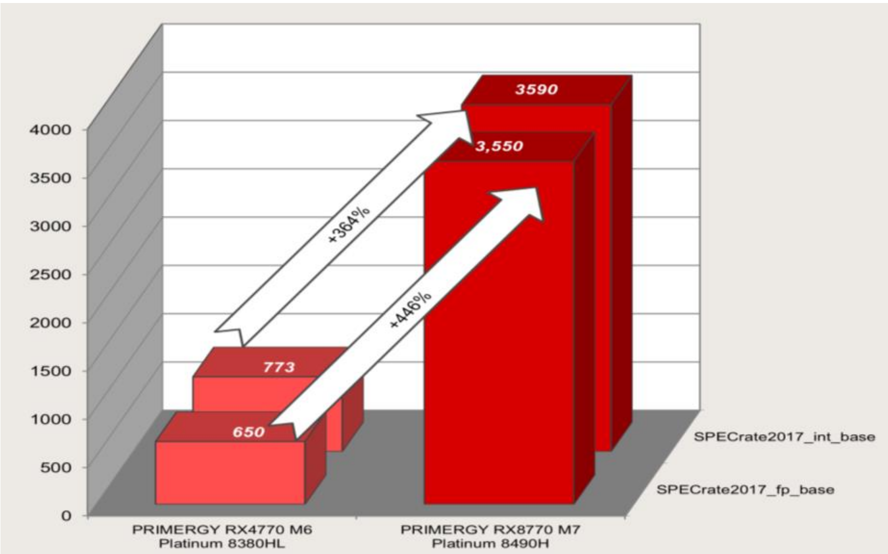
In terms of processors, the benchmark result depends primarily on the size of the processor cache, the support for Hyper-Threading, the number of processor cores, and the processor frequency. In the case of processors with Turbo mode, the number of cores, which are loaded by the benchmark, determines 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 processor	SPECrate2017_int_base	SPECrate2017_fp_base
Xeon Platinum 8490H	60	8	**3,590**	**3,550**
Xeon Platinum 8468H	48	8	**3,080**	**3,240**
Xeon Platinum 8460H	40	8	**2,820**	**3,050**
Xeon Platinum 8454H	32	8	**2,100**	**2,430**
Xeon Platinum 8450H	28	8	1,780 est.	2,250 est.
Xeon Platinum 8444H	16	8	1,230 est.	1,710 est.

Processor model	Number of cores	Number of processors	SPECspeed2017_int_base	SPECspeed2017_fp_base
Xeon Platinum 8490H	60	8	-	**343**
Xeon Platinum 8444H	16	8	**12.6**	-

The following graph compares the throughput of the PRIMERGY RX8770 M7 and the PRIMERGY RX4770 M6 with the highest number of sockets in the previous generation in a maximum performance configuration.
The PRIMERGY RX8770 M7 is 4.6 times better at SPECrate 2017_int_base and 5.5 times better at SPECrate 2017_fp_base than the PRIMERGY RX4770 M6 with half the number of sockets.



SPECrate2017: PRIMERGY RX4770 M6 and PRIMERGY RX8770 M7

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

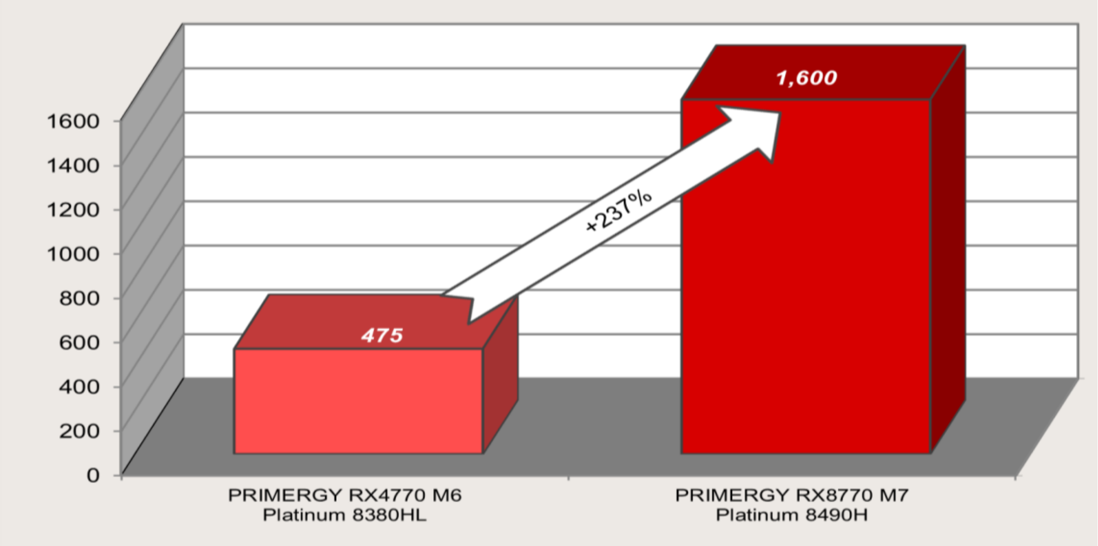
System Under Test (SUT)	
Hardware	
• Model	PRIMERGY RX8770 M7
• Processor	8 x 4th Generation Intel Xeon Scalable Processors Family
• Memory	64 x 64GB (1x64GB) 2Rx4 DDR5-4800 R ECC
Software	
• BIOS settings	• DCU Streamer Prefetcher = Disabled • SNC(Sub NUMA) = Enable SNC4 • Intel Virtualization Technology = 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]
Xeon Platinum 8490H	4,800	307	60	1.9	8	**1600**
Xeon Platinum 8468H	4,800	307	56	2.0	8	**1604**
Xeon Platinum 8460H	4,800	307	52	2.0	8	**1593**
Xeon Platinum 8454H	4,800	307	48	2.4	8	**1547**
Xeon Platinum 8450H	4,800	307	48	2.1	8	1513 .est
Xeon Platinum 8444H	4,800	307	32	2.8	8	1189 .est

The following graph compares the throughput of the PRIMERGY RX8770 M7 and the PRIMERGY RX4770 M6 with the highest number of sockets in the previous generation in a maximum performance configuration.
The stream performance of the PRIMERGY RX8770 M7 is 3.4 times better than that of the PRIMERGY RX4770 M6, which has half the number of sockets.



STREAM: PRIMERGY RX4770 M6 and PRIMERGY RX8770 M7

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.

https://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 https://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	
• Model	PRIMERGY RX8770 M7
• Processor	2 x 4th Generation Intel Xeon Scalable Processors Family
• Memory	64 x 64GB (1x64GB) 2Rx4 DDR5-4800 R ECC
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	Kernel Boot Parameter set with : nohz_full=1-X (X: logical core number -1)
• 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.

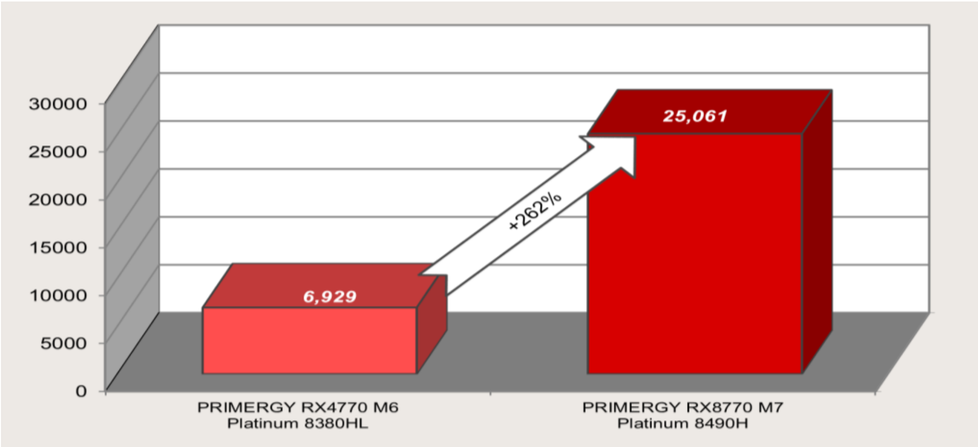
Processor	Number of cores	Rated Frequency [GHz]	Number of processors	Rpeak [GFlops]	Rmax [GFlops]		Efficiency
Xeon Platinum 8490H	60	1.90	8	29,184	**25,061**		86%
Xeon Platinum 8468H	56	2.00	8	25,805	**23,274**		90%
Xeon Platinum 8460H	52	2.00	8	22,528	**21,697**		96%
Xeon Platinum 8454H	48	2.40	8	17,203	**16,079**		93%
Xeon Platinum 8450H	48	2.10	8	14,336	13,112	est.	91%
Xeon Platinum 8444H	32	2.80	8	11,878	10,864	est.	91%

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 graph compares the throughput of the PRIMERGY RX8770 M7 and the PRIMERGY RX4770 M6 with the highest number of sockets in the previous generation in a maximum performance configuration.

The Linpack performance of the PRIMERGY RX8770 M7 is 3.6 times better than that of the PRIMERGY RX4770 M6, which has half the number of sockets.



LINPACK: PRIMERGY RX4770 M6 and PRIMERGY RX8770 M7

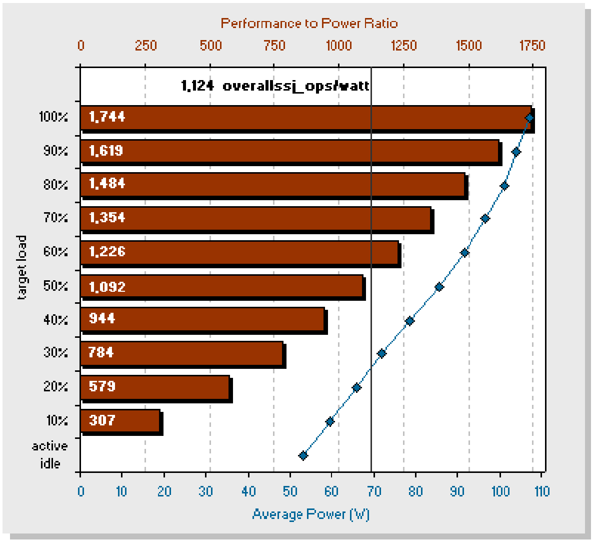
SPECpower_ssj2008

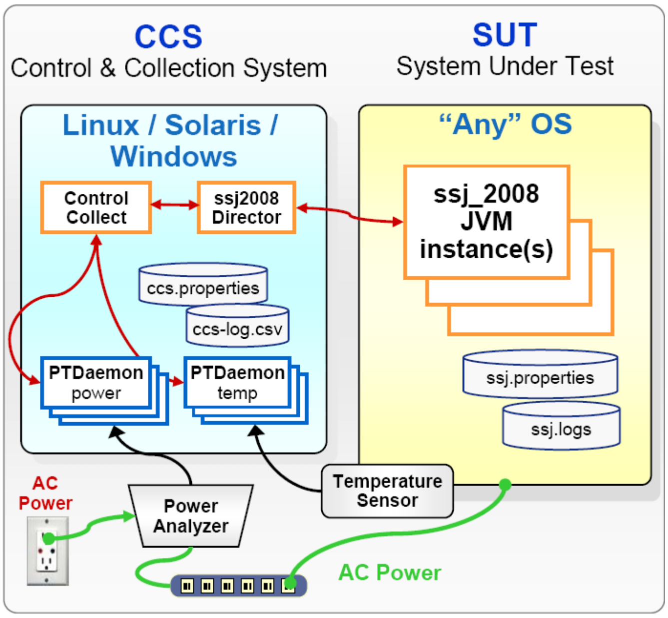
Benchmark description

SPECpower_ssj2008 is the first industry-standard SPEC benchmark that evaluates the power and performance characteristics of a server. With SPECpower_ssj2008 SPEC has defined standards for server power measurements in the same way they have done for performance.

The benchmark workload represents typical server-side Java business applications. The workload is scalable, multi-threaded, portable across a wide range of platforms, and easy to run. The benchmark tests CPUs, caches, the memory hierarchy, and scalability of symmetric multiprocessor systems (SMPs), as well as the implementation of Java Virtual Machine (JVM), Just In Time (JIT) compilers, garbage collection, threads, and some aspects of the operating system.

SPECpower_ssj2008 reports power consumption for servers at different performance levels — from 100% to “active idle” in 10% segments — over a set period of time. The graduated workload recognizes the fact that processing loads and power consumption on servers vary substantially over the course of days or weeks. To compute a power-performance metric across all levels, measured transaction throughputs for each segment are added together and then divided by the sum of the average power consumed for each segment. The result is a figure of merit called “overall ssj_ops/watt”. This ratio provides information about the energy efficiency of the measured server. The defined measurement standard enables customers to compare it with other configurations and servers measured with SPECpower_ssj2008. The diagram shows a typical graph of a SPECpower_ssj2008 result.





The benchmark runs on a wide variety of operating systems and hardware architectures and does not require extensive client or storage infrastructure. The minimum equipment for SPEC-compliant testing is two networked computers, plus a power analyzer and a temperature sensor. One computer is the System Under Test (SUT) which runs one of the supported operating systems and the JVM. The JVM provides the environment required to run the SPECpower_ssj2008 workload which is implemented in Java. The other computer is a “Control & Collection System” (CCS) which controls the operation of the benchmark and captures the power, performance, and temperature readings for reporting. The diagram provides an overview of the basic structure of the benchmark configuration and the various components.

Benchmark environment

System Under Test (SUT)

Hardware

• Model	PRIMERGY RX8770 M7
• Processor	8 x Intel Xeon Platinum 8490H 60C 1.90GHz 350W
• Memory	64 x 32GB (1x32GB) 2Rx8 DDR5-4800 R ECC
• Network interface	1Gbit/s (RJ45) on Motherboard
• Disk subsystem	1 x SSD SATA M.2 drive for booting, non hot-plug 240GB
• Power Supply Unit	4 x 2,600W titanium PSU

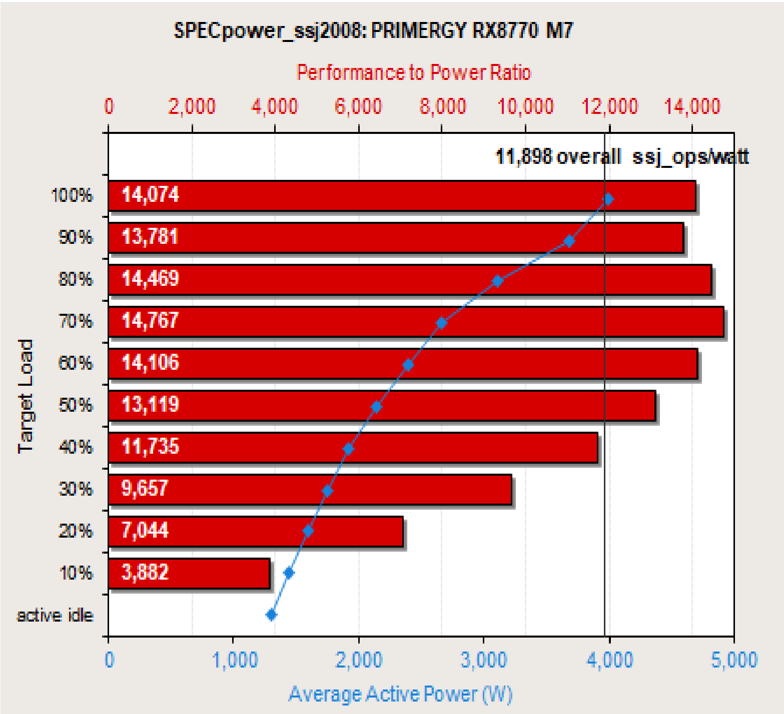
Software

• BIOS settings	ASPM Support = Auto Hardware Prefetcher = Disabled Adjacent Cache Line Prefetch = Disabled DCU Streamer Prefetcher = Disabled Intel(R) VT-d = Disabled Energy Performance = Energy Efficient Package C State limit = C2 Uncore Frequency Scaling = Power balanced SNC(Sub NUMA) = Enable SNC2 USB Port Control = Disable all ports Network Stack = Disabled
• Operating system	Windows Server 2022 Standard
• Operating system settings	Turn off hard disk after = 1 Minute PCI Express Link State Power Management = Maximum power savings Minimum processor state = 0% Maximum processor state = 100% Turn off display after = 1 Minute POWERCFG /SETACVALUEINDEX SCHEME_CURRENT SUB_PROCESSOR PERFBOOSTMODE 4 POWERCFG /SETACVALUEINDEX SCHEME_CURRENT SUB_PROCESSOR PERFINCTHRESHOLD 95 POWERCFG /SETACVALUEINDEX SCHEME_CURRENT SUB_PROCESSOR PERFDECTHRESHOLD 93 POWERCFG /SETACVALUEINDEX SCHEME_CURRENT SUB_PROCESSOR PERFDECTIME 1 POWERCFG /SETACVALUEINDEX SCHEME_CURRENT SUB_PROCESSOR IDLESCALING 1 POWERCFG /S SCHEME_CURRENT Using the local security settings console, "lock pages in memory" was enabled for the user running the benchmark. Benchmark was started via Windows Remote Desktop Connection.
• JVM	Oracle Java HotSpot(TM) 64-Bit Server VM 18.9 (build 11.0.16.1+1-LTS, mixed mode)
• JVM settings	-server -Xmn1500m -Xms1625m -Xmx1625m -XX:+UseLargePages -XX:AllocatePrefetchDistance=256 -XX:AllocatePrefetchLines=4 -XX:InlineSmallCode=3900 -XX:MaxInlineSize=270 -XX:MaxTenuringThreshold=15 -XX:ParallelGCThreads=2 -XX:SurvivorRatio=1 -XX:TargetSurvivorRatio=99 -XX:-UseAdaptiveSizePolicy -XX:+UseParallelOldGC -XX:FreqInlineSize=2500 -XX:LoopUnrollLimit=45 -XX:InitialTenuringThreshold=12 -XX:-ThreadLocalHandshakes -XX:UseAVX=0

Benchmark results

The PRIMERGY RX8770 M7 in Microsoft Windows Server 2022 Standard achieved the following result:

SPECpower_ssj2008 = 11,898 overall ssj_ops/watt



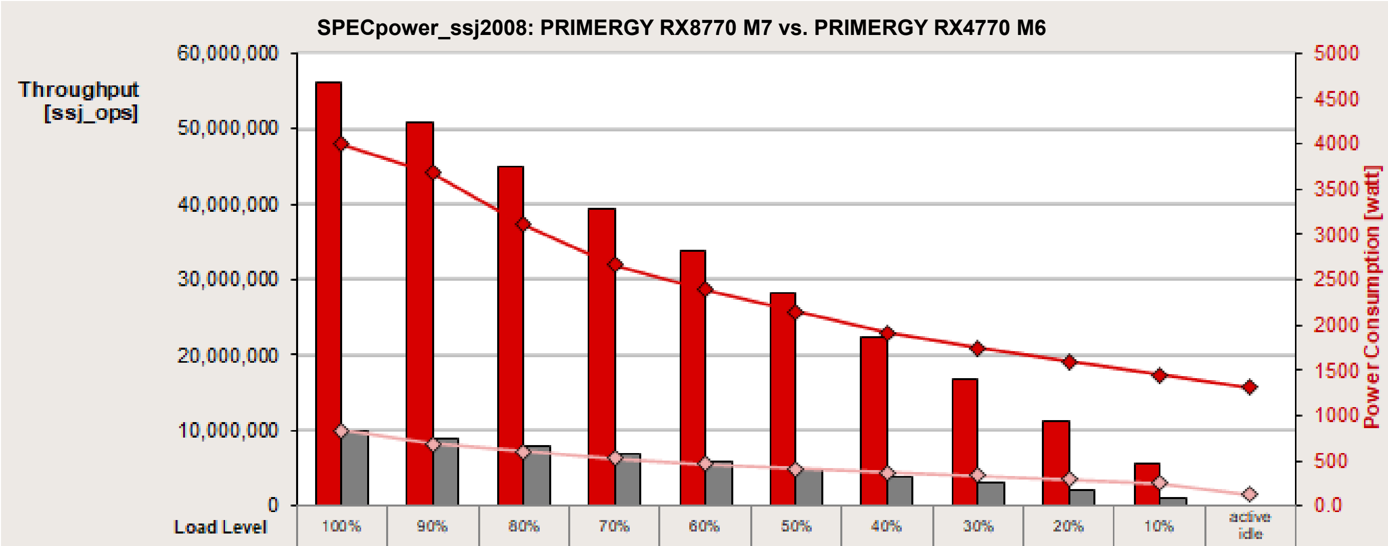
The adjoining diagram shows the result of the configuration described above. The red horizontal bars show the performance to power ratio in ssj_ops/watt (upper x-axis) for each target load level tagged on the y-axis of the diagram. The blue line shows the run of the curve for the average power consumption (bottom x-axis) at each target load level marked with a small rhomb. The black vertical line shows the benchmark result of 11,898 overall ssj_ops/watt for the PRIMERGY RX8770 M7. This is the quotient of the sum of the transaction throughputs for each load level and the sum of the average power consumed for each measurement interval.

The following table shows the benchmark results for the throughput in ssj_ops, the power consumption in watts and the resulting energy efficiency for each load level.

Performance		Power	Energy Efficiency
Target Load	ssj_ops	Average Power (W)	ssj_ops/watt
100%	56,228,882	3,995	14,074
90%	50,717,458	3,680	13,781
80%	44,977,955	3,109	14,469
70%	39,367,763	2,666	14,767
60%	33,730,309	2,391	14,106
50%	28,102,680	2,142	13,119
40%	22,493,026	1,917	11,735
30%	16,884,058	1,748	9,657
20%	11,245,927	1,596	7,044
10%	5,626,680	1,449	3,882
Active Idle	0	1,308	0
			Σ ssj_ops / Σ power = 11,898

Comparison with the predecessor

The following diagram shows for each load level (on the x-axis) the throughput (on the left y-axis) and the power consumption (on the right y-axis) of the PRIMERGY RX8770 M7 compared to the PRIMERGY RX4770 M6 with the highest number of sockets in the previous generation.

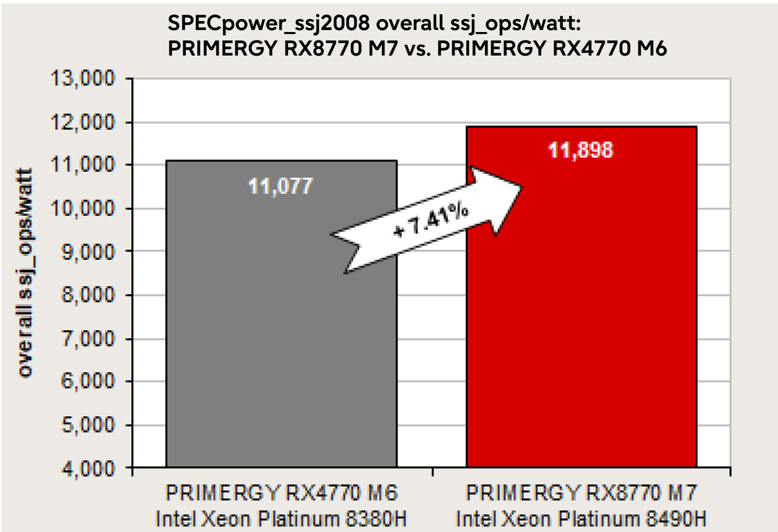


Load Level	100%	90%	80%	70%	60%	50%	40%	30%	20%	10%	active idle
RX8770 M7 ssj_ops	56,228,882	50,717,458	44,977,955	39,367,763	33,730,309	28,102,680	22,493,026	16,884,058	11,245,927	5,626,680	0
RX4770 M6 ssj_ops	9,904,110	8,914,839	7,933,444	6,944,695	5,948,299	4,956,147	3,966,282	2,973,866	1,980,442	990,995	0
RX8770 M7 watt	3,995	3,680	3,109	2,666	2,391	2,142	1,917	1,748	1,598	1,449	1,308
RX4770 M6 watt	834	688	603	532	467	413	370	333	296	257	129

100% throughput of the PRIMERGY RX8770 M7 is 56,228,882 ssj_ops, which is 5.7 times higher than 9,904,110 ssj_ops of PRIMERGY RX4770 M6 with half the number of sockets.

On the other hand, the average power consumption of the PRIMERGY RX8770 M7 during benchmark is 2,364 W, which is 5.3 times higher than 447 W of the PRIMERGY RX4770 M6.

The overall energy efficiency of the PRIMERGY RX8770 M7 has improved by 7.41 % compared with RX4770 M6.



Measurement results of SPECpower_ssj2008 (June 7, 2023)

11,898 SPECpower_ssj2008

On June 7, 2023, PRIMERGY RX8770 M7 with eight Xeon Platinum 8490H processors achieved a performance value of 11,898 on the Windows Server 2022 Standard in the SPECpower_ssj2008 benchmark, in the Windows division of the 4th Generation Intel Xeon Processor Scalable Family category and won first place in 4-socket server SPECpower_ssj2008 performance.

For the latest results of the SPECpower_ssj2008, see https://www.spec.org/power_ssj2008/results/.

SAP BW Edition for SAP HANA Standard Application Benchmark

Description of the benchmark

With the increasing importance of SAP HANA and in particular SAP Business Warehouse (SAP BW) on HANA, a new benchmark was introduced in July 2016: the SAP BW Edition for SAP HANA Standard Application Benchmark, referred to as SAP BWH Benchmark in the following.

The benchmark represents a typical mid-size customer scenario and volumes and utilizes the new capabilities of SAP HANA which enable customers to enhance their BW processes.

Since its first edition in 2016, the SAP BWH Benchmark has been further developed and adapted to customer requirements. In the meantime, SAP BWH Benchmark version 3 is available. Benchmarks with the older versions won't be certified anymore. The results of different versions must not be compared with each other.

The SAP BWH Benchmark consists of 3 phases:

- Data load phase
- Query throughput phase
- Query runtime phase

Data load phase

The data flow starts with a data load from the source object into the corporate memory layer. The source object is shipped with the backup.

The source object contains 1.3 billion records (= 1 data set). It is possible to load this data set of 1.3 billion records multiple times.

The data set stored in the source is fetched and propagated through the different layers in 25 load cycles. In other words, 1 load cycle processes 1/25 of the data set.

The permissible data volumes are a multiple of 1.3 billion initial data records. The minimum number of data sets to be loaded is dependent on the size of the main memory.

The data load phase takes several hours and is a combination of CPU- and IO-intensive load. When several HANA nodes are used (see "SAP HANA Scale-up and Scale-out Configuration Architecture" below), significant network load is generated.

Query throughput phase

The queries for the throughput phase must be executed via an ABAP program with a variant containing 380 queries. Users execute the set of navigation steps in random order (via asynchronous RFCs). The queries contain typical query patterns which can be found in BW productive systems of customers.

The query throughput phase runs one hour and is CPU bound. In a HANA multi-node environment, also significant network load is generated.

Query runtime phase

For the query runtime phase the same ABAP program as for the throughput phase is used with a different variant. The variant contains 10 queries which are executed sequentially. These queries are used to measure the runtime. They contain complex query patterns which are executed in BW productive systems of customers, but which are typically not executed by many users in parallel but selectively by some power users. Therefore, they are executed sequentially.

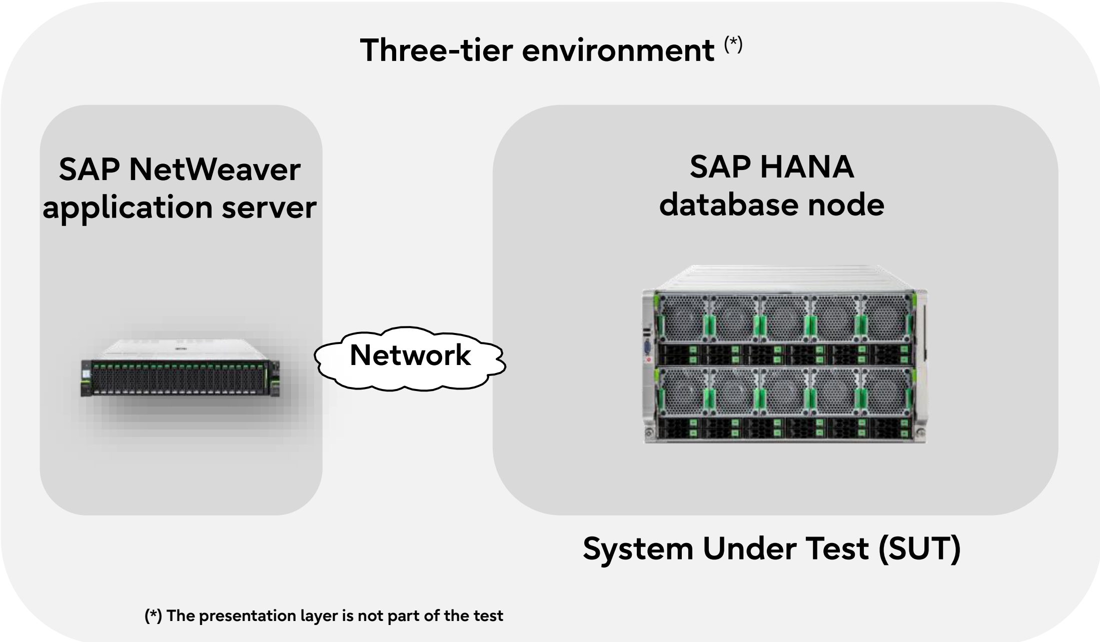
The query runtime phase takes a short time and generates a small load. Only a few processors cores are used, single thread performance is important for short runtimes.

Certified and published SAP BWH Benchmarks are published on SAP's benchmark site here.

Benchmark environment

In general, a single database node or multiple database nodes can be used for SAP benchmarks to scale the workload. In the context of SAP HANA and particularly the SAP BW Edition for SAP HANA Standard Application Benchmark it is referred to as a scale-up configuration in the case of a single database node and a scale-out configuration in the case of multi database nodes.

The SAP BWH Benchmarks for PRIMERGY RX8770 M7 were performed on a scale-up configuration.



Although an application server is involved in the benchmark, neither performance metrics are measured nor does the server appear on the benchmark certificate.

System Under Test (SUT)	
Hardware	
• Model	PRIMERGY RX8770 M7
• Processor	8 x Intel Xeon Platinum 8490H processor 60C 1.9GHz 350W
• Memory	128 x 64 GB (1x64GB) 2Rx4 DDR5-4800 R ECC
• Network interface	1 Gbit/s (RJ45) on Motherboard
• Storage subsystem	1 × PRAID EP640i LP internal RAID controller 1 x SSD SAS 2.5" Mixed Use 1.6TB 2 x PRAID EP680e LP external RAID controller 3 x ETERNUS JX40 S2 Enclosure 25 × JX40 S2 TLC SSD 1.6TB 3DWPD
Software	
• Operating system	SUSE Linux Enterprise Server 15 SP4
• Database	SAP HANA 2.0

Application Server	
Hardware	
• Model	PRIMERGY RX2540 M5
• Processor	2 x Intel Xeon Platinum 8280L 28C 2.7GHz 205W
• Memory	12 x 64GB (1x64GB) 2Rx4 DDR4-2933 R ECC
• Network interface	1 Gbit/s (RJ45) on Motherboard
• Storage subsystem	1 x PRAID EP420i RAID Controller 2 x HDD SAS 2.5" 15K 600GB 3 x SSD 1.6TB 1 x PACC EP P4800X AIC PCIe-SSD 750GB
Software	
• Operating System	SUSE Linux Enterprise Server 15 SP2
• Technology platform release	SAP Netweaver 7.50

Benchmark results

On August 14, 2023, the following SAP BW edition for SAP HANA Standard Application Benchmark Version 3 was certified:

Certification number 2023035	
• Benchmark Phase 1 Number of initial records Runtime of last Data Set (seconds)	 15,600,000,000 11,508
• Benchmark Phase 2 Query Executions per Hour CPU utilization of database server	 9,457 97%
• Benchmark Phase 3 Total Runtime of complex query phase (seconds)	 128
• Operating system	SUSE Linux Enterprise Server 15
• Database	SAP HANA 2.0
• Technology platform release	SAP Netweaver 7.50
• Configuration Database Server	Fujitsu Server PRIMERGY RX8770 M7, 8 processors / 480 cores / 960 threads, Intel Xeon Platinum 8490H processor, 1.90 GHz, 80 KB L1 cache and 2,048 KB L2 cache per core, 112.5 MB L3 cache per processor, 8,192 GB DRAM

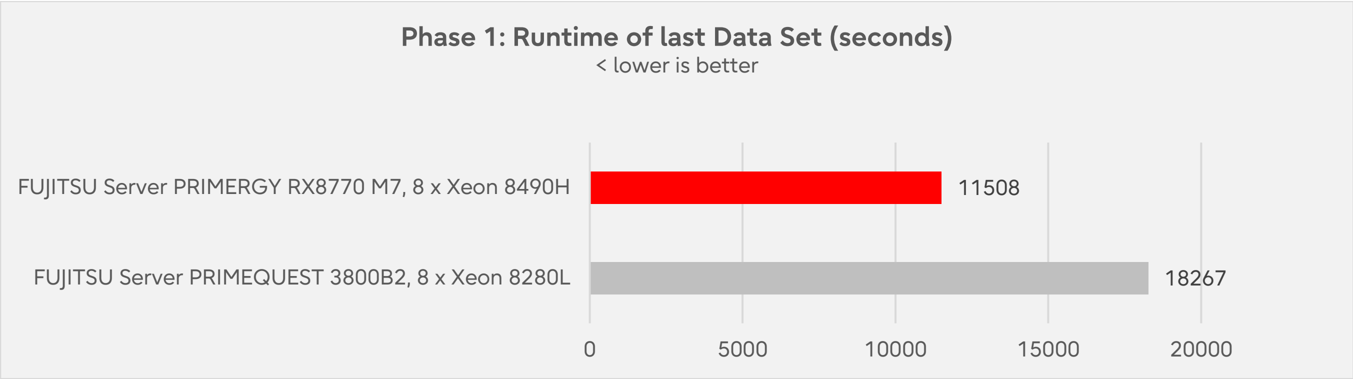
The SAP BWH Benchmark certificate can be found here: Certification 2023035.

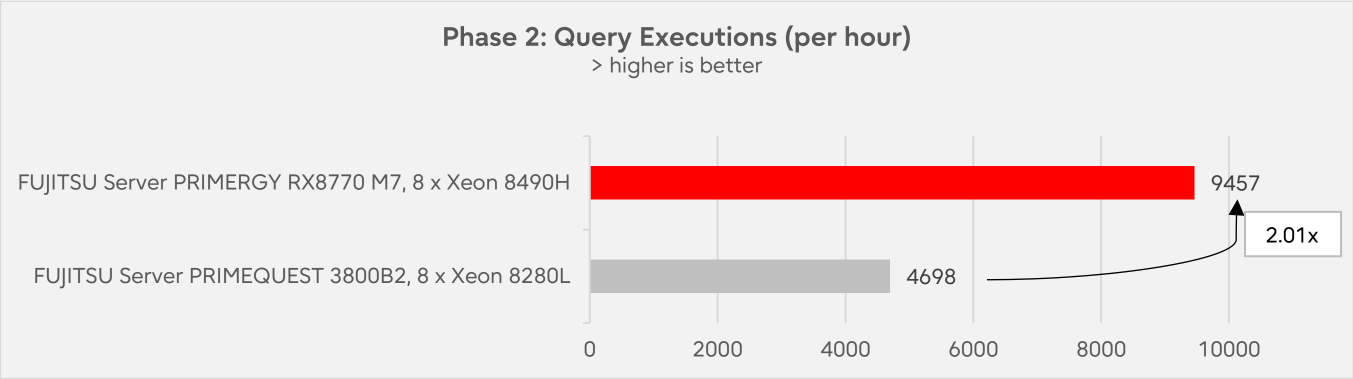
The following charts compare the results of all three benchmark phases of the SAP BW Edition for SAP HANA Standard Application Benchmark for PRIMERGY RX8770 M7 and its predecessor PRIMEQUEST 3800B2.

On PRIMEQUEST 3800B2 the benchmark was also conducted with 15.6 billion initial records. Here are the details of the benchmark:

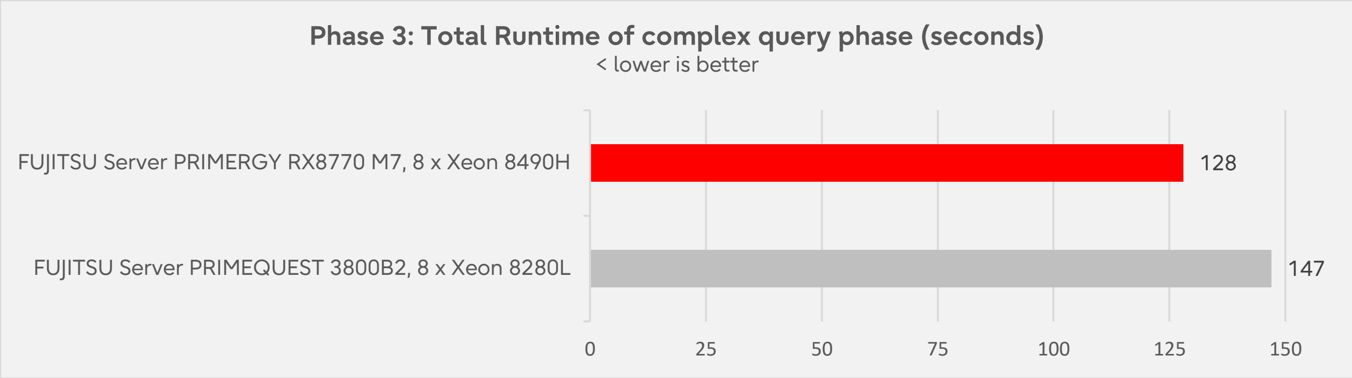
FUJITSU Server PRIMEQUEST 3800B2, 8 processors / 224 cores / 448 threads Intel Xeon Platinum 8280L processor, 3072 GB DRAM, 6144 GB Persistent Memory, SAP BW edition for SAP HANA Standard Application Benchmark Version 3, 3-tier scale-up configuration, Bare-Metal environment, SAP HANA 2.0, SAP NetWeaver 7.50, SUSE Linux Enterprise Server 12, 15.6 billion initial records, 18267 seconds runtime of last data set, 4698 query executions/h, 147 seconds total runtime of complex query phase. Certification 2019058.

Benchmark phase 2, the CPU bound query throughput phase, is best to compare the performance of the 4th Generation Intel Xeon Scalable Processor aka Sapphire Rapids against the 2nd Generation Xeon Scalable Processor aka Cascades Lake.





PRIMERGY RX8770 M7 achieves more than factor 2 query executions per hour, compared to its predecessor PRIMEQUEST 3800B2 with Intel Xeon 8280L (8 processors, 224 cores, 448 threads, 2nd Generation Xeon Scalable Processor aka Cascade Lake).


Phase 3: Total Runtime of complex query phase (seconds)
< lower is better
FUJITSU Server PRIMERGY RX8770 M7, 8 x Xeon 8490H
128
FUJITSU Server PRIMEQUEST 3800B2, 8 x Xeon 8280L
147
0
25
50
75
100
125
150

Disk I/O: Performance of storage media

Benchmark description

Performance measurements of disk subsystems for PRIMERGY 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
SSD	SAS SSD (SAS 12Gbps, Mixed Use)	XS800LE70084 XS1600LE70084 XS3200LE70084
	SATA SSD (SATA 6Gbps, Mixed Use)	MTFDDAK480TDT MTFDDAK960TDT MTFDDAK1T9TDT MTFDDAK3T8TDT

Controller: PRAID EP680i

Storage media	Category	Drive name
SSD	PCIe SSD (Mixed Use)	KCM61VUL1T60 KCM61VUL3T20 KCM61VUL6T40

Controller: Intel C741 Standard SATA AHCI controller

Storage media	Category	Drive name
SSD	M.2 Flash module	MTFDDAV240TDS MTFDDAV480TDS

Controller: Intel C741 Standard NVM Express controller

Storage media	Category	Drive name
SSD	M.2 Flash module (NVMe)	MTFDKBA480TFR MTFDKBA960TFR

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 3.0x16	0, 1, 1E, 10, 5, 50
PCIe SSD 2.5"	PRAID EP680i	-	PCIe 4.0 x8	SATA 6G SAS 12G PCIe 4.0 x16	0, 1, 1E, 10, 5, 50
M.2 Flash	C741 Standard SATA AHCI controller	-	DMI 3.0 x4	SATA 6G	-
M.2 Flash (NVMe)	C741 Standard NVM Express controller	-	DMI 3.0 x4	PCIe 3.0 x2	

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 PRIMERGY servers.

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

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 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 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

SSDs

Capacity [GB]	Storage device	Interface	Transactions [IO/s]			Throughput [MiB/s]	
			Database	Fileserver	Filecopy	Streaming	Restore
- SAS SSD (MU)							
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
- SATA SSD (MU)							
480	MTFDDAK480TDT	SATA 6G	49,138	6,383	6,600	508	437
960	MTFDDAK960TDT	SATA 6G	50,488	6,970	7,136	508	486
1,920	MTFDDAK1T9TDT	SATA 6G	50,669	7,183	7,336	508	487
3,840	MTFDDAK3T8TDT	SATA 6G	49,490	7,115	7,208	493	474
- PCIe SSD (MU)							
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
- M.2 SATA SSD							
240	MTFDDAV240TDS	SATA 6G	31,923	5,489	5,512	504	299
480	MTFDDAV480TDS	SATA 6G	39,553	6,331	6,516	501	394
- M.2 NVMe SSD							
480	MTFDKBA480TFR	PCIe3 x2	74,947	15,849	12,564	1,644	685
960	MTFDKBA960TFR	PCIe3 x2	147,206	31,459	25,928	1,644	1,381

Literature

PRIMERGY Servers

https://www.fujitsu.com/global/products/computing/servers/primergy/

PRIMERGY RX8770 M7

This Whitepaper

https://docs.ts.fujitsu.com/dl.aspx?id=68375d84-4c05-4d98-a9af-bc7f3f5e5088

Data sheet

https://docs.ts.fujitsu.com/dl.aspx?id=2fd409a6-beaa-422e-bb6d-285635560514

PRIMERGY Performance

https://www.fujitsu.com/global/products/computing/servers/primergy/benchmarks/

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

SPECpower_ssj2008

https://www.spec.org/power_ssj2008

Benchmark Overview SPECpower_ssj2008

https://docs.ts.fujitsu.com/dl.aspx?id=166f8497-4bf0-4190-91a1-884b90850ee0

SAP BWH

https://www.sap.com/benchmark

Benchmark results

https://www.sap.com/dmc/exp/2018-benchmark-directory/#/bwh

Benchmark overview

http://docs.ts.fujitsu.com/dl.aspx?id=70a4c869-586c-49f3-a6a4-47f188dd72b3

Document change history

Version	Date	Description
1.1	2023-10-20	Update • SAP BWH Standard Application Benchmark Measured with Intel Xeon Platinum 8490H • Minor corrections
1.0	2023-07-26	New: • Technical data • SPEC CPU2017, STREAM, LINPACK Measured and calculated with 4th Generation Intel Xeon Processor Scalable Family • SPECpower_ssj2008 Measured with Intel Xeon Platinum 8490H • Disk I/O Measured with 2.5 inch model

Contact

Fujitsu

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

PRIMERGY Performance and Benchmarks

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