WHITE PAPER
FUJITSU PRIMERGY SERVERS
Performance Report PCIe-SSDs ioDrive®2

This document looks closely at the ioDrive®2 PCIe-SSDs that are available for some PRIMERGY servers with regard to the disk I/O performance that can be achieved.

In addition to the measurement results, the measurement method and the environment, in which the measurements were carried out, are also briefly explained.

Version
1.0b
2013-07-04

Contents

Document history 2
Introduction and technical data 3
Measurement method 5
Measurement environment 6
Measurement results 7
  A single PCIe-SSD 7
  Two PCIe-SSDs 11
  Four PCIe-SSDs 13
  Influence of nominal capacity 14
Best practice 16
  Optimizing the current processor frequency 16
  Optimizing the processor allocation 18
  Select the server 18
Comparison with other storage media 19
Conclusion 20
Literature 21
Contact 21

Document history

Version 1.0

First report version

Version 1.0a

Minor corrections

Version 1.0b

Minor corrections

Introduction and technical data

The PCIe-SSD is an innovative, nonvolatile storage medium for PRIMERGY servers, which in certain cases can be used as a logical hard disk drive instead of “Hard Disk Drives” (HDDs) or “Solid State Drives” (SSDs) that have a SAS or SATA interface. The fact that the PCIe-SSD storage medium is directly coupled with the PCIe bus means that very high transaction rates and low latencies are possible. Such a storage medium is interesting for access acceleration to data that is stored on a nonvolatile basis, in cases such as:

- Databases
- Web2.0 Servers
- Page files (especially in virtual environments, because these files are an essential part of the memory management concepts there)

Technical data

The PCIe-SSDs on offer are flash memories that are based on a multi-level cell (MLC) NAND type of memory. There are currently two PCIe-SSD types, one model with a nominal capacity of 785 GB and one with 1.2 TB.

A powerful controller is integrated on the storage medium itself, thus making an additional RAID controller unnecessary. The bus interface to the server is of type PCIe 2.0, x4, which would allow a practical throughput of up to about 1720 MB/s. According to the data sheet “Fusion-io® ioDrive®2 Solid-State Storage devices“ the maximum throughput values of these PCIe-SSDs are:

Load profile	Maximum throughput ioDrive®2 PCIe-SSD	
	PCIe-SSD 785GB MLC	PCIe-SSD 1.2TB MLC
100% read, 1 MB block size	1.5 GB/s	1.5 GB/s
100% write, 1 MB block size	1.1 GB/s	1.3 GB/s

The ioDrive®2 PCIe-SSDs considered in this document are the second generation of this storage medium for PRIMERGY servers. The first generation (ioDrive®) was dealt with under performance aspects in the white paper “Performance Report PCIe-SSDs“. The essential maximum performance values for ioDrive®2 PCIe-SSDs have in comparison with the first generation been approximately doubled, as is shown in the following table of values from the data sheet “Fusion-io® ioDrive®2 Solid-State Storage devices“.

Load profile	Maximum throughput ioDrive® PCIe-SSD	
	PCIe-SSD 320GB MLC	PCIe-SSD 640GB MLC
100% read, 64 kB block size	735 MB/s	750 MB/s
100% write, 64 kB block size	510 MB/s	550 MB/s

The PCIe-SSDs cannot be used as drives for booting.

Administration software

Always part of a PCIe-SSD is the graphic administration software ioSphere®, which contains among other things the “Low-Level Format” menu with setting options for the formatting of the storage medium. A well-balanced setting consisting of a compromise of good write performance and good utilization of the storage capacity is used as standard. These formatting options can be used for special application scenarios to provide weighting either in the direction of higher storage capacity or higher write performance.

The formatting options on offer are:

- Factory Capacity: Standard setting for the formatting of the PCIe-SSD. The nominal capacity is available here as the real storage capacity.
- Maximum Capacity: Increases the real storage capacity of the medium. The write speed and stock of reserve blocks are reduced here for the blocks worn out by too many write cycles.
- Improved Performance: Increased write performance at the expense of about 10% of the storage capacity.
- High Performance: Maximum write performance at the expense of about 20% of the storage capacity.

Furthermore, the “Custom” selection also enables you to set an arbitrary nominal capacity percentage. Please note that the reformatting of a PCIe-SSD inevitably results in the deletion of its stored data, because the storage medium is restructured during this process.

Not least the storage capacity is interesting for the selection of a PCIe-SSD. Due to the formatting options the following actual storage capacities result from the nominal capacity:

Nominal capacity	Formatting option	Actual storage capacity
785 GB	High Performance	628 GB
	Improved Performance	706.5 GB
	Factory Capacity	785 GB
	Maximum Capacity	~845 GB
1.2 GB	High Performance	964 GB
	Improved Performance	1084.5 GB
	Factory Capacity	1205 GB
	Maximum Capacity	~1294 GB

If a PCIe-SSD is to be used for a page file, the standard setting “Disable Page File Support” must be reset for optimal performance to “Enable Page File Support” in the setting options of the “Page File Support” section. This setting has no impact on performance for other uses of the PCIe-SSD.

Measurement method

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“.

As standard, performance measurements of disk subsystems in PRIMERGY servers are carried out with a defined measurement method, which models the hard disk accesses of real application scenarios on the basis of specifications.

The essential specifications are:

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

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

Standard load 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]*

Measurement environment

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

System Under Test (SUT)	
Hardware	
Models	1 × PRIMERGY RX300 S7 1 × PRIMERGY RX900 S2
Processor	PRIMERGY RX300 S7: 2 × Xeon E5-2643 (3.3 GHz) 2 × Xeon E5-2603 (1.8 GHz) PRIMERGY RX900 S2: 4 × Xeon E7-8850 (2.0 GHz)
Data medium	4 × PCIe-SSD ioDrive®2 1.2 TB MLC 1 × PCIe-SSD ioDrive®2 785 GB MLC
Software	
BIOS	PRIMERGY RX300 S7: 1.12.0 PRIMERGY RX900 S2: 01.09
BIOS settings	PRIMERGY RX300 S7: For measurements of type “Performance”: Execute disable bit = Disabled; Frequency Floor Override = Enabled; Power Technology = Custom; Energy Performance = Performance; CPU C6 Report = Disabled; Package C State limit = C0 For measurements of type “Default”: Execute disable bit = Enabled; Frequency Floor Override = Disabled; Power Technology = Energy Efficient; Energy Performance = Balanced Performance; CPU C6 Report = Enabled; Package C State limit = No Limit PRIMERGY RX900 S2: Performance/Watt=Performance; Hyper-Threading=Disabled, NX Memory Protection=Disabled
Operating system	Microsoft Windows Server 2008 R2 Enterprise
Operating system settings	AFFINITY: The AFFINITY for the process that creates disk I/Os (=dynamo in the measurement method used) was set to the optimal CPU. Power Plan: For measurements of type “Performance”: Select a power plan = High performance; For measurements of type “Default”: Select a power plan = Balanced;
Firmware	For PCIe-SSD ioDrive®2: v7.0.2, rev 108609
Driver	For PCIe-SSD ioDrive®2: iomemory-vsl.sys 3.1.5 build 126
Administration software	For PCIe-SSD ioDrive®2: ioSphere® 3.2.2
Initialization of RAID arrays	RAID arrays are initialized before the measurement with an elementary block size of 64 kB (“stripe size”)
File system	NTFS
Measuring tool	Iometer 2006.07.27
Measurement data	Measurement file of 32 GB

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

Measurement results

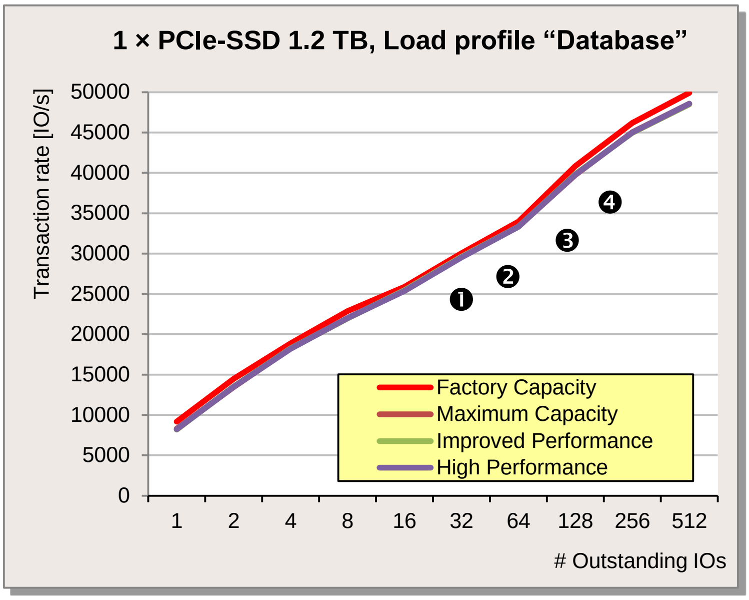
Here we use the five standard load profiles "Database", "File server", "File copy", "Streaming" and "Restore", which were described in the section "Measurement method", to examine the performance of the PCIe-SSDs. If various load intensities are considered in the diagrams for accesses to the storage medium, they are specified, as is usual for Fujitsu Technology Solutions, in "# of outstanding IOs". The applications with low load intensity are represented by one outstanding IO and the applications with a very high load intensity by 512 outstanding IOs.

A single PCIe-SSD

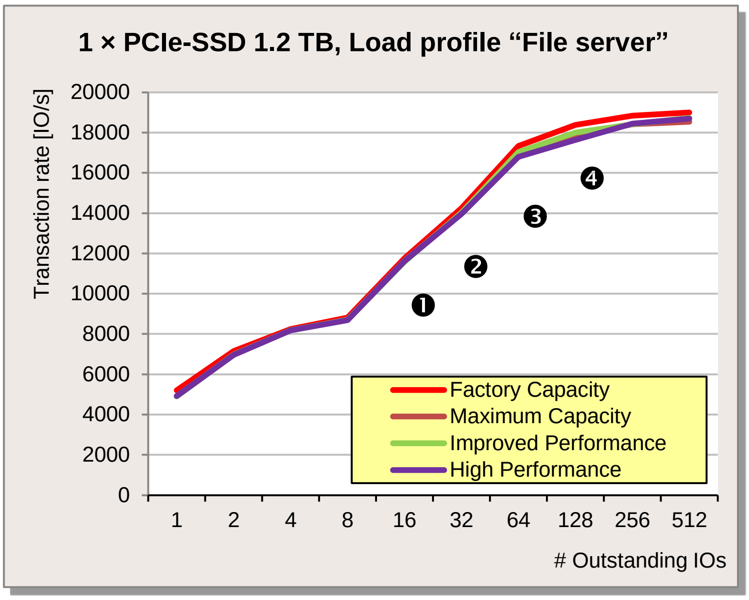
The fundamental considerations should first be made on the basis of a single PCIe-SSD with a nominal capacity of 1.2 TB. The load profiles with random access are dealt with first of all, and secondly the load profiles with sequential access.

Random accesses

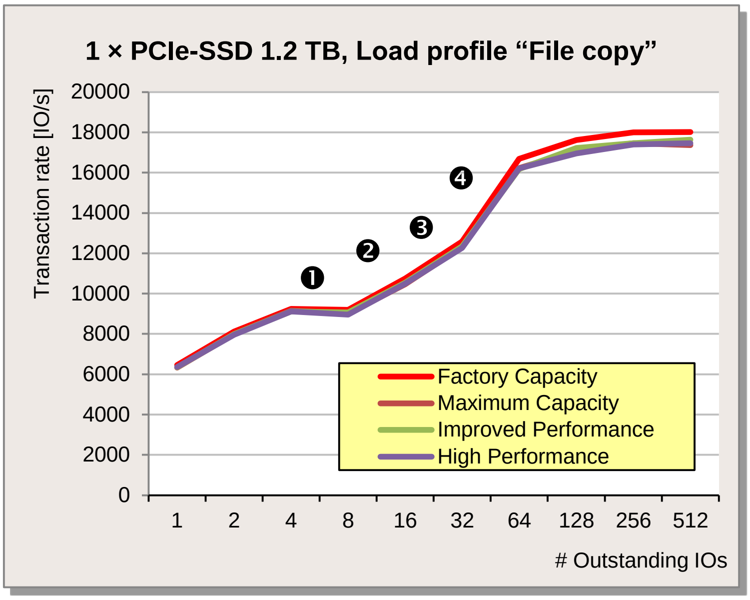
As is customary for random accesses, the transaction rate is specified below in IO/s as the measure for performance. The following diagram shows the transaction rates for the "Database" load profile (random access, 67% read, 8 kB block size). In the case of low load intensities the PCIe-SSD returns about 8000 to 9000 IO/s for all formatting options. As load intensity increases, the transaction rate increases evenly until it reaches just under 50000 IO/s with a very high load intensity. Performance here does not depend on the formatting option used.



The next diagram shows the transaction rates for the “File server” load profile (random access, 67% read, 64 kB block size). The transaction rates for low load intensities are about 5000 IO/s. As load intensity increases, so does the transaction rate. From about 64 outstanding IOs the growth in performance levels off and achieves a maximum of about 19000 IO/s with the highest load intensities. It is interesting to note that this maximum transaction rate is for the block size used equivalent to a data throughput of 1190 MB/s. This value already comes close to the maximum value for an ioDrive®2 PCIe-SSD, thus making the leveling off understandable. Performance for this load profile does not depend on the formatting option used, either.



The next diagram shows the transaction rates for the “File copy” load profile (random access, 50% read, 64 kB block size). The transaction rates for low load intensities are about 6500 IO/s. As load intensity increases, so does the transaction rate. From about 64 outstanding IOs the growth in performance levels off and achieves a maximum of about 18000 IO/s with the highest load intensities. In this case, performance does not depend on the formatting option used, either.



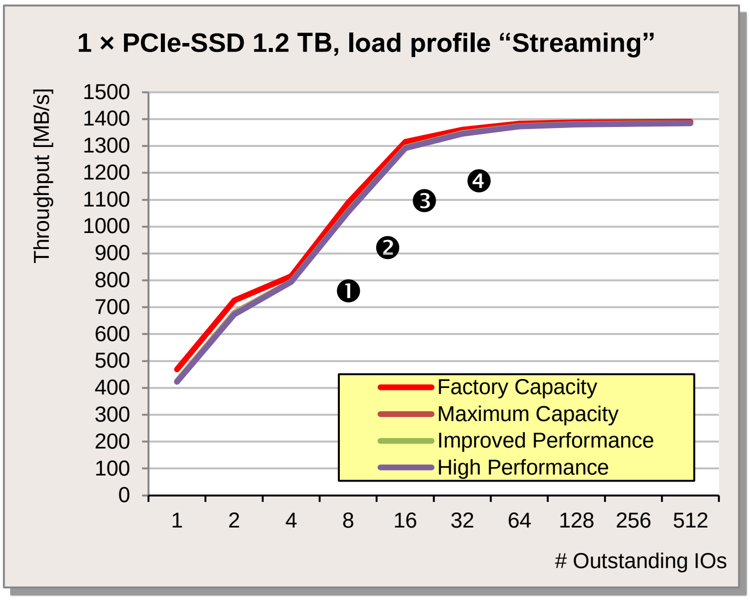
Sequential accesses

PCIe-SSDs are scarcely used for sequential accesses in productive applications, because conventional hard disks already achieve good performance values for this purpose. Nevertheless, such load profiles are for the sake of completeness also to be discussed here. As is customary for sequential accesses, the transaction rate is no longer specified below as the measure for performance, but throughput in MB/s. To understand the throughput values that can be achieved it is essential to know the maximum values of the storage medium. As already mentioned and according to the data sheet "Fusion-io® ioDrive®2 Solid-State Storage devices" the maximum values for a 1 MB block size are:

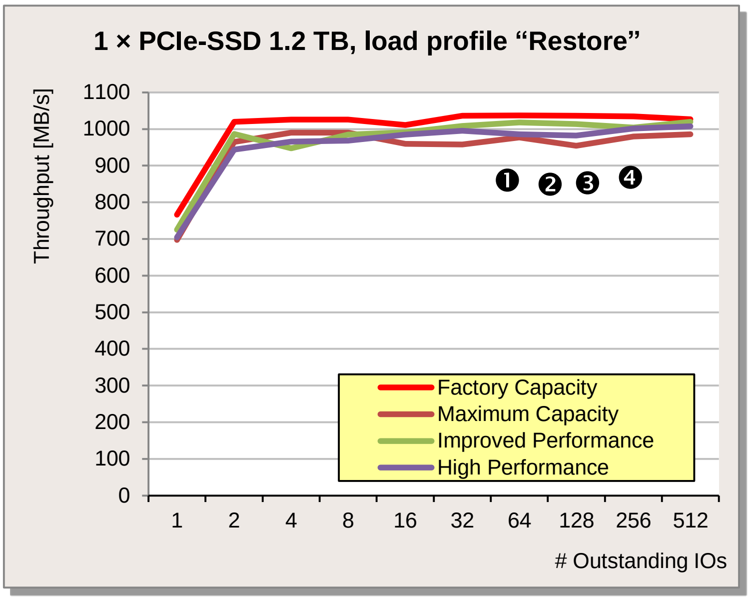
Load profile	Maximum throughput ioDrive®2 PCIe-SSD	
	PCIe-SSD 785GB MLC	PCIe-SSD 1.2TB MLC
100% read, 1 MB block size	1.5 GB/s	1.5 GB/s
100% write, 1 MB block size	1.1 GB/s	1.3 GB/s

These values are used as a frame of reference for the following analyses.

The following diagram shows the throughputs for the "Streaming" load profile (sequential access, 100% read, 64 kB block size). It is possible with the lowest load intensity to achieve a throughput of about 450 MB/s for all the formatting options. If load intensity increases to 16 outstanding IOs, the throughput quickly increases to 1300 MB/s. At this value the maximum is almost reached; up to the highest load intensities data throughput increases further up to about 1400 MB/s. In this case, performance does not depend on the formatting option used, either.



The next diagram shows the throughputs for the “Restore” load profile (sequential access, 100% write, 64 kB block size). Data throughput of about 700 to 770 MB/s is achieved with low load intensities. As soon as you increase load intensity to two outstanding IOs, the maximum data throughput of about 1000 MB/s is achieved. The various formatting options do not differ significantly with this load profile.



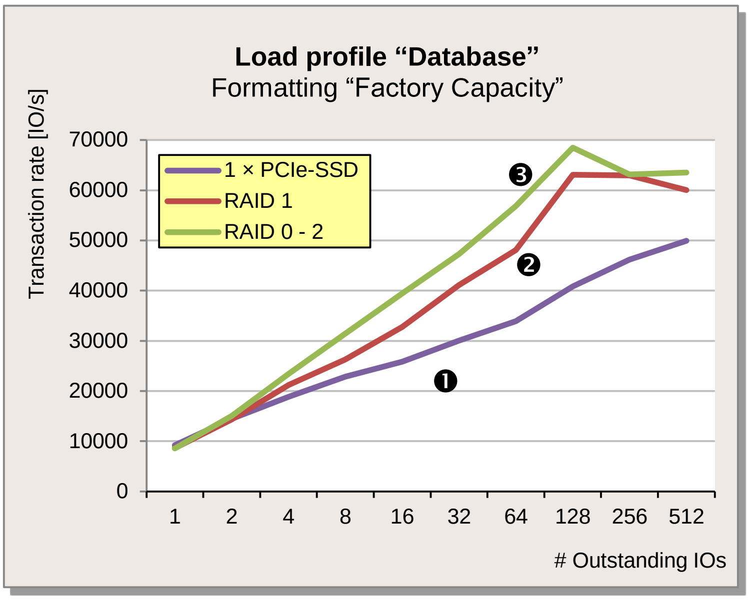
Overall comparison of the formatting options as regards performance

Unlike the predecessor generation, it is normally not necessary with the ioDrive®2 PCIe-SSDs to deviate from the standard formatting “Factory Capacity” for performance reasons. This is why the following sections should do without a discussion of any other formatting options than “Factory Capacity”.

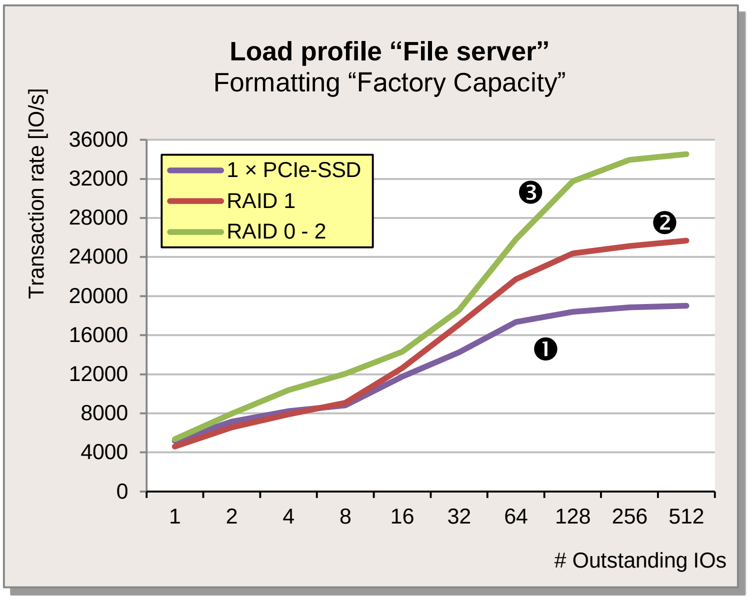
These deviating formatting options are only considered in special cases.

Two PCIe-SSDs

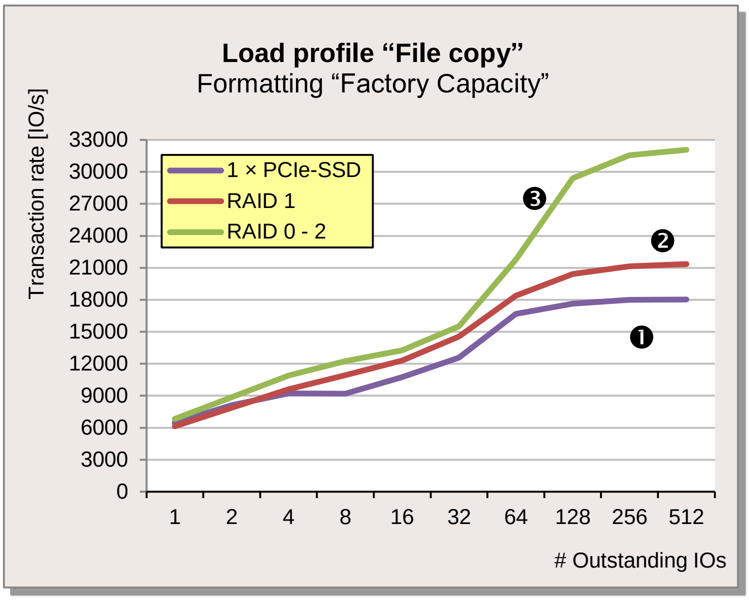
As with other logical drives, RAID arrays can also be formed from PCIe-SSDs at operating system level. To increase fail-safety you can use RAID 1. In the case of load profiles that are not exclusively write profiles this RAID level usually also results in an increase in performance for the read share of the accesses. In case of an application in which higher performance is more important than fail-safety, a RAID 0 array can also be formed from two PCIe-SSDs. The possible RAID configurations for various load intensities with a single PCIe-SSD (using the example of the 1.2 TB version) are to be compared below for each of the five standard load profiles.



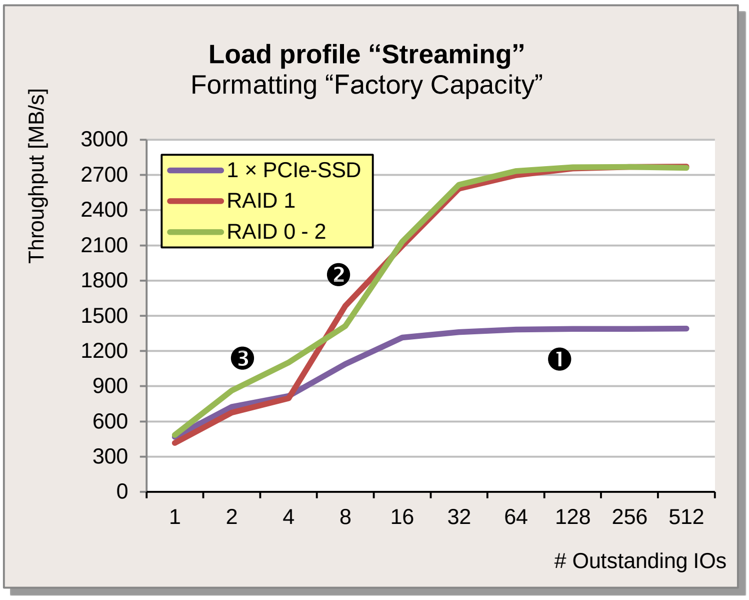
In the case of the “Database” load profile (random access, 67% read, 8 kB block size) the RAID 1 array (❷) and the RAID 0 array (❸) have the same transaction rate as a single PCIe-SSD (❶) for the lowest load intensity, namely about 8000 to 9000 IO/s. If load intensity increases up to 512 outstanding IOs, the transaction rates increasingly grow further apart: The single PCIe-SSD (❶) has up to approximately 50000 IO/s, the RAID 1 array (❷) achieves up to approximately 60000 IO/s, and the RAID 0 array (❸) has up to approximately 68000 IO/s.



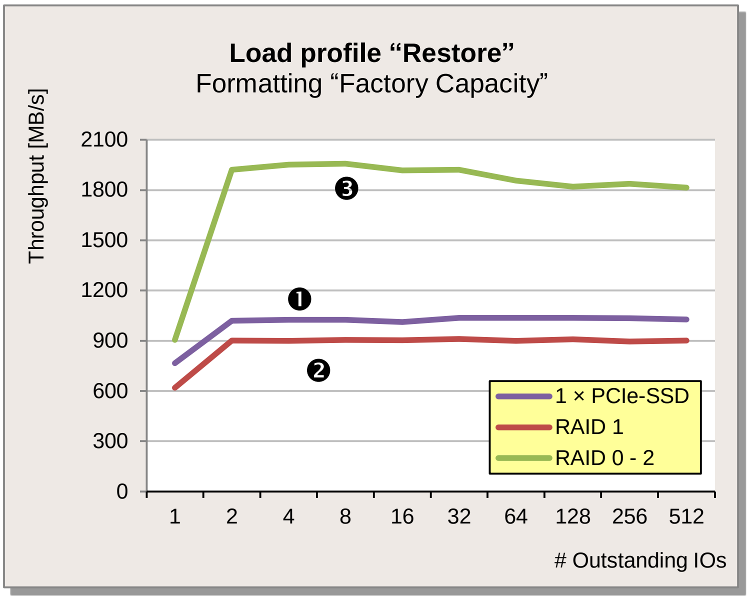
In the case of the “File server” load profile (random access, 67% read, 64 kB block size) both RAID arrays (❷ and ❸) have a transaction rate of about 5000 IO/s for the lowest load intensity, exactly the same as the single PCIe-SSD (❶). If you increase the load intensity, the transaction rates increase continuously both for the two RAID arrays and for the single PCIe-SSD (❶). In the case of the highest load intensity the RAID 0 array (❸) achieves about 35000 IO/s, the RAID 1 array (❷) about 26000 IO/s and the single PCIe-SSD (❶) about 19000 IO/s. The transaction rate for RAID 1 for all load intensities always lies between the one for a single PCIe-SSD and the one for RAID 0.



The case of the “File copy” load profile (random access, 50% read, 64 kB block size) is similar to the “File server” load profile. The two RAID arrays (❷ and ❸) achieve a transaction rate of about 6000 to 7000 IO/s for the lowest load intensity, exactly the same as the single PCIe-SSD (❶). If you increase the load intensity, the transaction rates increase continuously both for the two RAID arrays as well as for the single PCIe-SSD (❶). In the case of the highest load intensity the RAID 0 array (❸) achieves about 32000 IO/s, the RAID 1 array (❷) about 21000 IO/s and the single PCIe-SSD (❶) about 18000 IO/s. The transaction rate for RAID 1 for all load intensities always lies between the one for a single PCIe-SSD and the one for RAID 0.



In the case of the “Streaming” load profile (sequential access, 100% read, 64 kB block size) both the two RAID arrays (❷ and ❸) and the single PCIe-SSD (❶) have data throughputs ranging from 400 to 500 MB/s for the lowest load intensity. If you increase the load intensity up to 512 outstanding I/Os, RAID 0 achieves a maximum of 2770 MB/s and the single PCIe-SSD a maximum of 1390 MB/s. In the case of RAID 0 the maximum value is almost already achieved with 32 outstanding I/Os. The single PCIe-SSD already achieves the maximum value with 16 outstanding I/Os. Up to four outstanding I/Os the throughput for RAID 1 acts in the same way as with a single PCIe-SSD; from about eight outstanding I/Os the throughput then is almost exactly the same as with RAID 0.

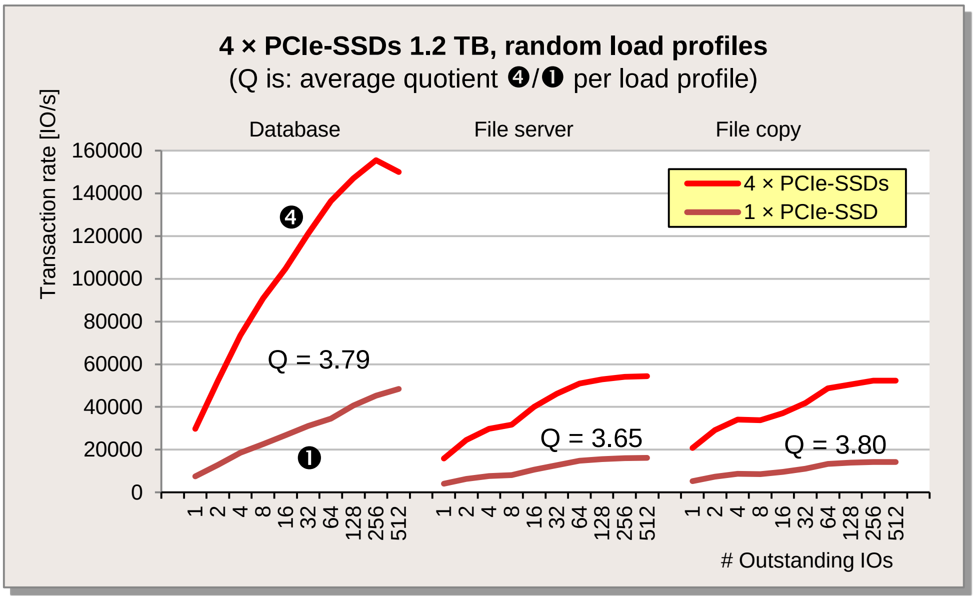


In the case of the “Restore” load profile (sequential access, 100% write, 64 kB block size) the single PCIe-SSDs (❶) achieves about 760 MB/s with one outstanding I/O. The maximum throughput, namely about 1030 MB/s, is achieved for all higher load intensities. The throughput for RAID 1 (❷) is about 160 MB/s lower than with a single PCIe-SSD for all load intensities. The highest values are achieved continuously with RAID 0 (❸): the data throughput for one outstanding I/O is about 900 MB/s and quickly increases to about 1900 MB/s for higher load intensities. If the load intensity increases further, the throughput then falls slightly to about 1800 MB/s.

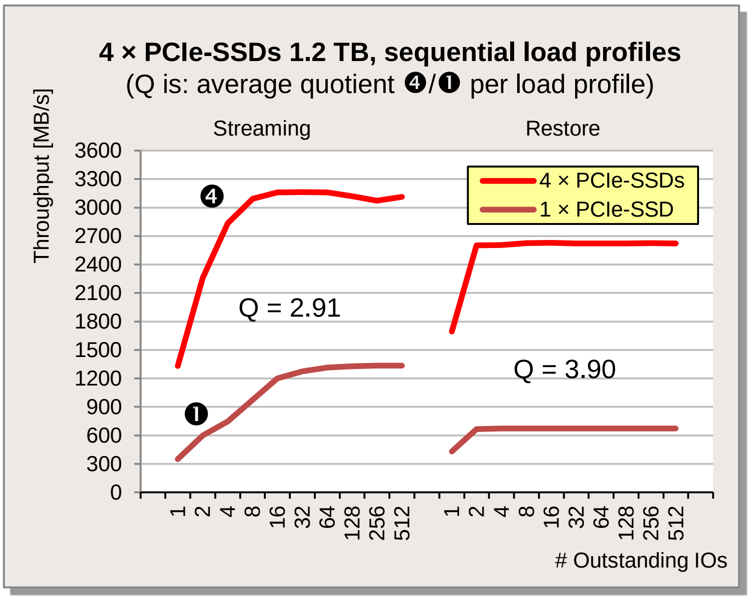
Four PCIe-SSDs

A number of PRIMERGY servers (e. g. PRIMERGY RX900 S2) can be equipped with up to four PCIe-SSDs. A RAID array is not always created from all the PCIe-SSDs in the case of such a number. You often have several processes or applications that create disk I/Os and which you prefer to have run on clearly separated PCIe-SSDs. It is interesting to know for such cases whether the available disk I/O performance scales with the number of PCIe-SSDs in the server. As an example the following two diagrams contain performance comparisons between a single PCIe-SSD and four PCIe-SSDs under parallel load (not in a RAID array), both of which are in a PRIMERGY RX900 S2. The load intensity created by the measurement method for each PCIe-SSD and the load profile are identical at every point of the curves. The average scaling factor (Q) for all load intensities (# outstanding I/Os) is entered in the diagrams for each load profile.

The first diagram deals with random accesses. The same standard load profiles as before are used: “Database” (random access, 67% read, 8 kB block size), “File server” (random access, 67% read, 64 kB block size) and “File copy” (random access, 50% read, 64 kB block size).



The second diagram deals with sequential accesses. The same standard load profiles as before are also used: “Streaming” (sequential access, 100% read, 64 kB block size) and “Restore” (sequential access, 100% write, 64 kB block size).

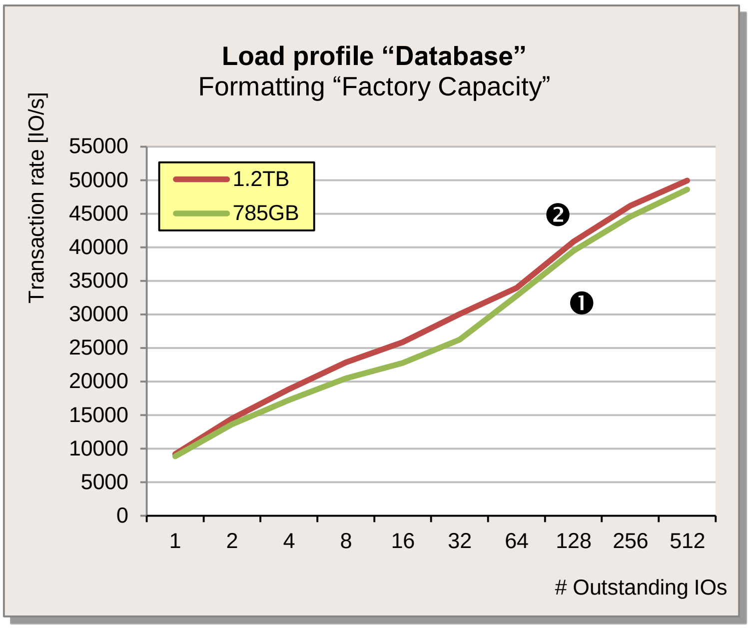


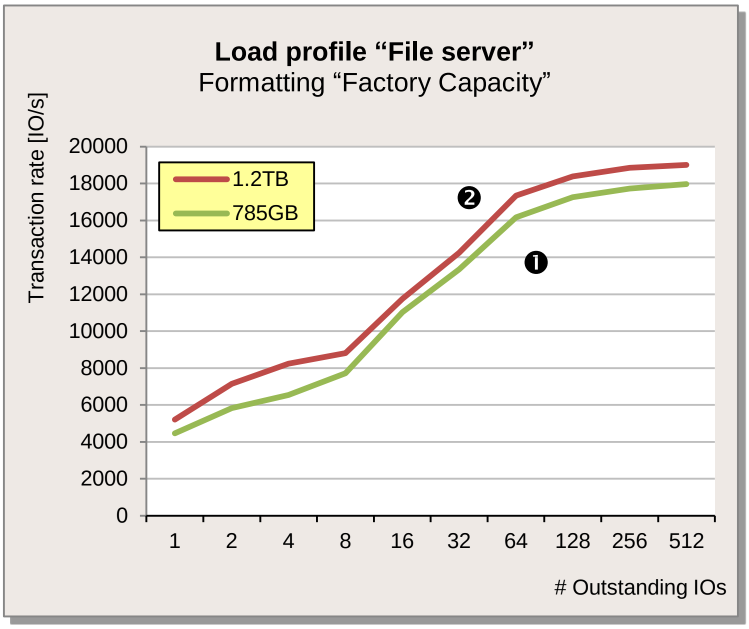
On the whole, the scaling factor Q from one to four PCIe-SSDs is in most cases just below 4.00. In a server with four PCIe-SSDs transaction rates of up to 155527 IO/s and data throughputs of up to 3165 MB/s are possible in total.

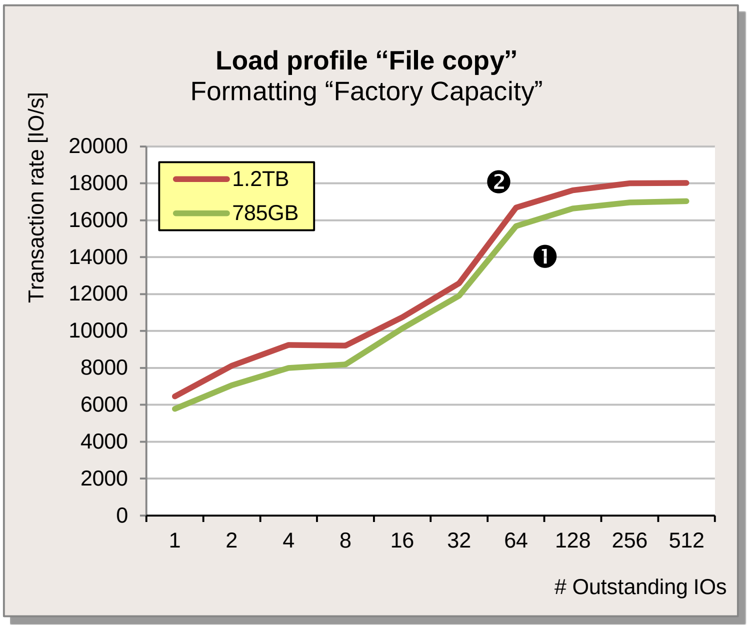
Influence of nominal capacity

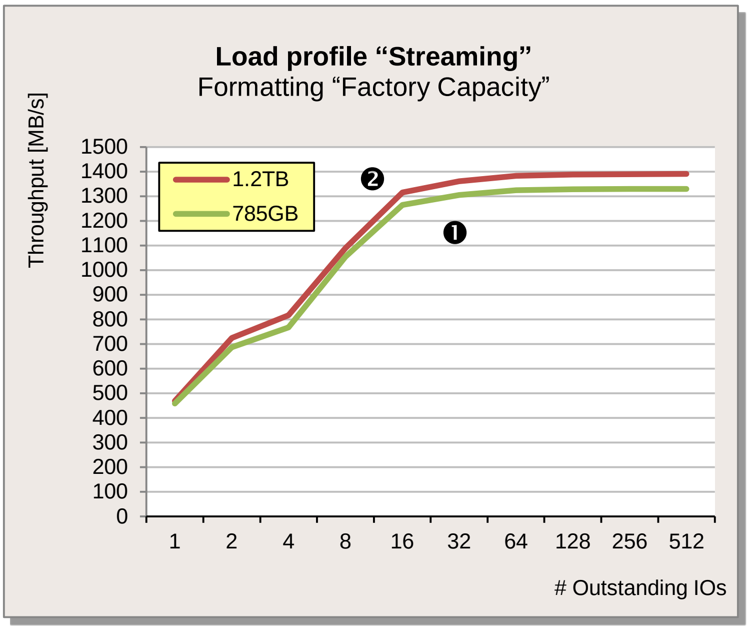
All else being identical, there is also a slight difference in performance between the capacity versions 785 GB and 1.2 TB. The following diagrams show this for a single PCIe-SSD in a PRIMERGY RX300 S7.

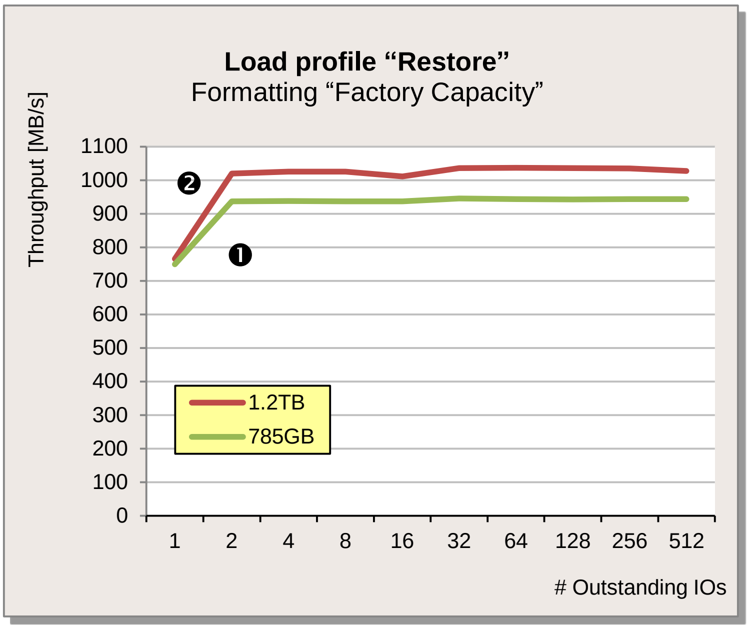
The advantage of the larger capacity version tends to be greater for load profiles with a write share. In these cases the advantage usually ranges between 4% and 10%.






Load profile "File copy"
Formatting "Factory Capacity"
Transaction rate [IO/s]
20000
18000
16000
14000
12000
10000
8000
6000
4000
2000
0
1.2TB
785GB
❷
❶
1
2
4
8
16
32
64
128
256
512
Outstanding IOs


Load profile "Streaming"
Formatting "Factory Capacity"
Throughput [MB/s]
1500
1400
1300
1200
1100
1000
900
800
700
600
500
400
300
200
100
0
1.2TB
785GB
❷
❶
1
2
4
8
16
32
64
128
256
512
Outstanding IOs


Load profile "Restore"
Formatting "Factory Capacity"
Throughput [MB/s]
1100
1000
900
800
700
600
500
400
300
200
100
0
❷
❶
1.2TB
785GB
1
2
4
8
16
32
64
128
256
512
Outstanding IOs

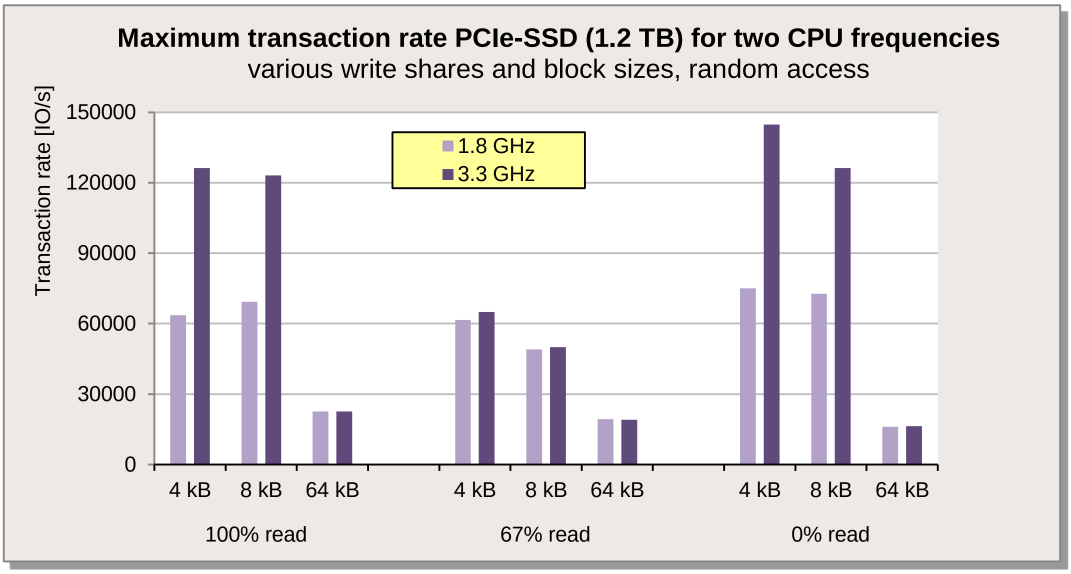
Best practice

Like every I/O component of a server, PCIe-SSDs also depend on the hardware conditions of the server. The processor, main memory and PCIe bus in particular play a role here. The performance of PCIe-SSDs can be positively influenced by the right selection, configuration and setting of the server.

Optimizing the current processor frequency

The actual processor frequency has an influence on the performance of PCIe-SSDs. In the ideal case, the PRIMERGY system is equipped with the processor type with the highest possible nominal frequency. However, this nominal frequency is in itself not yet significant, because modern processors can both temporarily increase and decrease the actual frequency compared with the nominal frequency by using technologies such as Turbo mode and energy-saving functions. The reduction can e. g. occur in applications with smaller computing requirements. If performance is more important than energy saving in an application, this lowering of frequency can be prevented by changing the settings. The following examples illustrate the performance effect you achieve through optimizing the CPU frequency.

The first issue to be examined is the impact of choosing a processor with higher nominal frequency. For this purpose, the following diagram considers a PCIe-SSD 1.2 TB in a PRIMERGY RX300 S7, which by means of suitable BIOS and operating system settings (listed in the section "Measurement environment") is tuned for maximum performance. The diagram compares the data throughputs for a CPU with 1.8 GHz and a CPU with 3.3 GHz for selected load profiles.



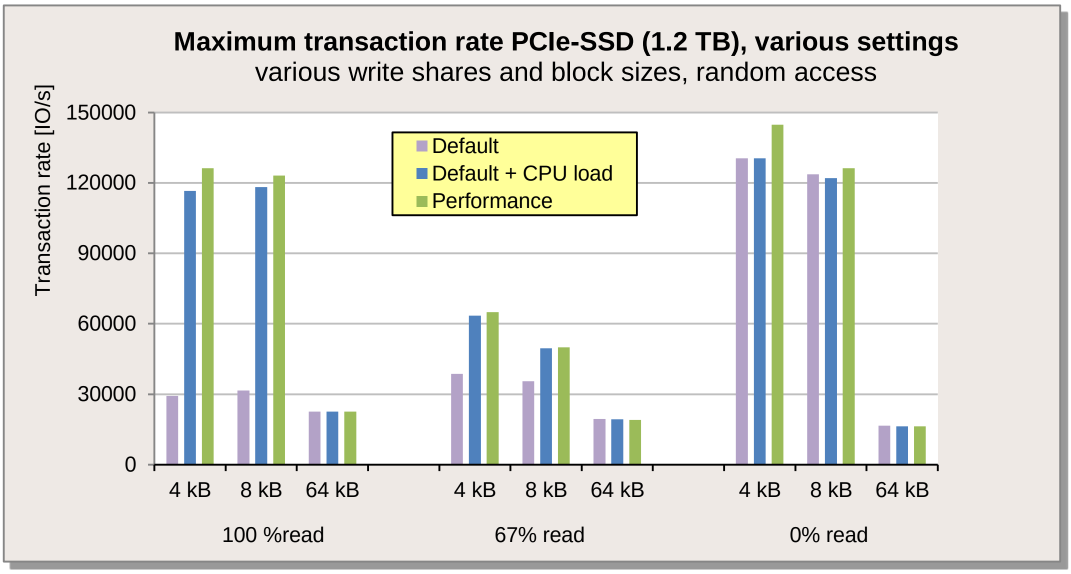
The diagram demonstrates that a significant advantage can be gained for block sizes up to 8 kB through a higher nominal frequency. A lower nominal CPU frequency has no impact for load profiles with mixed read and write accesses as well as for larger blocks.

The random accesses were merely chosen as an example in the diagram; the information the diagram provides also applies for sequential accesses.

Secondly, the following diagram examines the impact of temporarily lowering the frequency of a CPU because of low processor utilization. The data series shown represent:

- Default settings, low CPU utilization (default)
- Default settings, average CPU utilization (default + CPU load)
- Performance settings, low CPU utilization (performance)

"Default" and "Performance" refer to the sets of BIOS and operating system settings that are described in the section "Measurement environment".



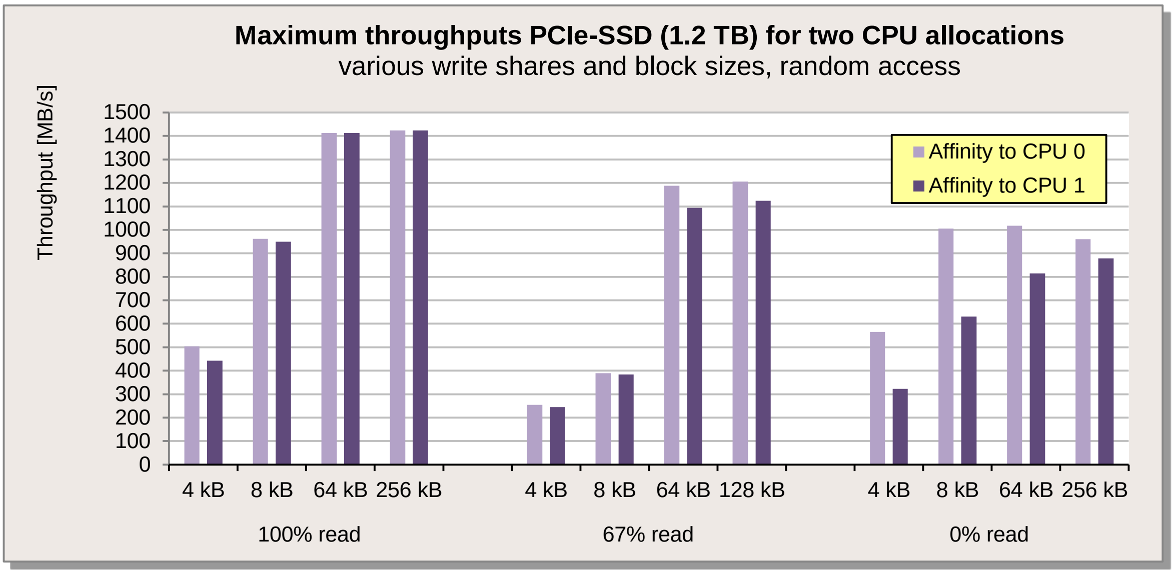
The diagram illustrates the increase in performance that can be achieved by inducing maximum CPU frequency in situations with low CPU utilization. It can be seen that a more than three-fold increase is possible for read-dominated load profiles with small block sizes (≤ 8 kB). The maximum CPU frequency can be induced in two ways: you either select optimal performance settings ("Performance"), or you assign a suitable process to the CPU that processes the interrupts of the PCIe-SSD ("Default + CPU load"). This assigned process should have a certain minimum computing requirement. However, it should not subject its CPU to a peak load situation. In the "Performance" case the BIOS parameter "Frequency Floor Override" ensures that the CPU always works with the maximum frequency; this is ensured in the "Default + CPU load" case by the process.

The effect is only minor for larger block sizes and write-dominated accesses.

The random accesses were merely chosen as an example in this diagram as well; the information the diagram provides also applies for sequential accesses.

Optimizing the processor allocation

The allocation of a process that creates disk I/Os to the most favorable CPU has a major influence on the performance of a PCIe-SSD. This allocation is possible for Windows operating systems, e. g. by starting the process via “start /affinity” or the “Task Manager”; and e. g. the utility “numactl” is used for Linux for this purpose. The most favorable CPU is the one that processes the interrupts of the PCIe-SSD. The following diagram compares the situations “favorable CPU” (in this case “CPU 0”) and “unfavorable CPU” (in this case “CPU 1”) for a PCIe-SSD 1.2 TB in a PRIMERGY RX300 S7.

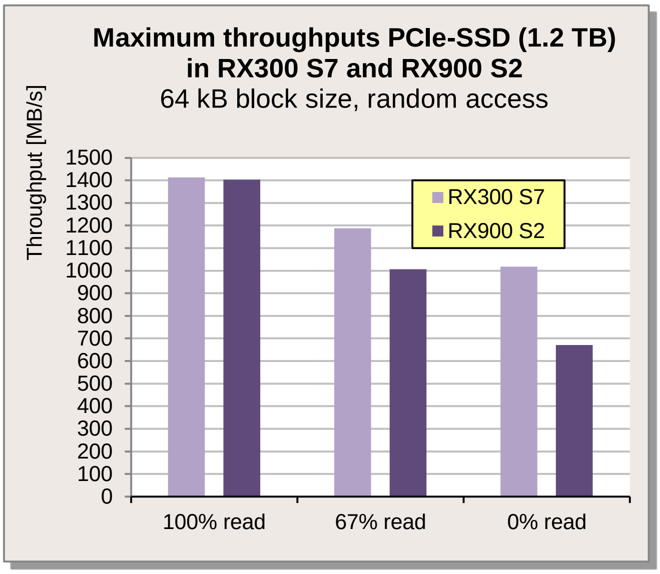


It becomes clear that correct processor allocation is becoming relevant for strongly write-dominated load profiles. The effect is the most distinct for block size 8 kB. Performance increases of up to 80% are possible for this case. The effect is still significant even for larger block sizes.

The random accesses were merely chosen as an example in this diagram as well; the information the diagram provides also applies for sequential accesses.

Select the server

The selection of the server also plays a role for the performance of PCIe-SSDs. The following diagram compares a PRIMERGY RX300 S7 (based on the Intel Xeon E5-2600/4600 processor family) and a PRIMERGY RX900 S2 (based on the Intel Xeon E7-4800/8800 processor family), which are equipped with a PCIe-SSD 1.2 TB in each case. The diagram compares by way of an example the maximum values of random load profiles with read shares of 100%, 67% and 0%. In the case of sequential load profiles the picture would be almost identical. The selected block size (64 kB) ensures that any differences in processor frequencies are not taken into account; this would only be done with smaller block sizes.



Due to their architecture the servers based on the E5-2600/4600 processor family have a clear advantage with strongly write-dominated load profiles in particular. On the other hand the servers based on the E7-4800/8800 processor family enable more performance per server, as support is provided in them for up to four PCIe-SSDs. Thus, for example up to 155000 IO/s are possible with the “Database” load profile and up to 3170 MB/s with the “Streaming” load profile in a PRIMERGY RX900 S2, whereas in the PRIMERGY RX300 S7 the corresponding maximum values are 68000 IO/s and 2770 MB/s (described in the subsections “Two PCIe-SSDs” and “Four PCIe-SSDs”).

Comparison with other storage media

The ioDrive®2 PCIe-SSD is to be compared here with its own predecessor as well as the storage media that are currently available for conventional drive bays: a conventional SAS-2.0 hard disk (SAS-2.0-HDD) and a SAS-2.0-SSD.

The following table compares the performance values of the four storage media for the five standard load profiles that have already been used.

Maximum performance for 1 – 512 outstanding IOs							
Load profile	**SAS-2.0-HDD 146 GB, 15 krpm, 2.5″ Disk cache enabled**	**SAS-2.0-SSD MLC 200 GB 2.5″ Disk Cache enabled**	**PCIe-SSD ioDrive® 640 GB Formatting "Advertised Capacity"**	**PCIe-SSD ioDrive®2 1.2 TB Formatting "Factory Capacity"**	**Quotient PCIe-SSD ioDrive®2 / SAS-2.0-HDD**	**Quotient PCIe-SSD ioDrive®2 / SAS-2.0-SSD**	**Quotient PCIe-SSD ioDrive®2 / PCIe-SSD ioDrive®**
Database	706 IO/s	14184 IO/s	28653 IO/s	49919 IO/s	70.71	3.52	1.74
File server	591 IO/s	3006 IO/s	8003 IO/s	19005 IO/s	32.16	6.32	2.37
File copy	571 IO/s	2985 IO/s	7693 IO/s	18018 IO/s	31.55	6.04	2.34
Streaming	192 MB/s	377 MB/s	617 MB/s	1390 MB/s	7.24	3.69	2.25
Restore	191 MB/s	196 MB/s	541 MB/s	1037 MB/s	5.43	5.29	1.92

As the table shows, the performance values of the ioDrive®2 PCIe-SSD achieve – depending on the load profile – a maximum of more than 2.3 times that of the ioDrive® predecessor generation, approximately 3.5 times that of the SAS-2.0-SSD, and more than 70 times that of the SAS-2.0-HDD.

The following table compares the minimum latency times when reading and writing small blocks.

Load profile	**SAS-2.0-HDD 146 GB, 15 krpm, 2.5″ Disk Cache enabled**	**SAS-2.0-SSD MLC 200 GB 2.5″ Disk Cache enabled**	**PCIe-SSD ioDrive® 640 GB Formatting "Advertised Capacity"**	**PCIe-SSD ioDrive®2 1.2 TB Formatting "Factory Capacity"**
Read Latency (1 kB sequential)	0.089 ms	0.090 ms	0.029 ms	0.066 ms
Write Latency (1 kB sequential)	0.098 ms	0.128 ms	0.033 ms	0.018 ms

On the whole, the minimum write latency (0.018 ms) is clearly lower with the ioDrive®2 PCIe-SSD than with other storage media in the comparison.

Conclusion

The ioDrive®2 PCIe-SSDs are an advanced non-volatile storage medium, which offers very high performance in a small space, above all for I/O accesses with a large number of transactions. The performance of the predecessor generation has with regard to both maximum transaction rate and maximum data throughput been approximately doubled. An ioDrive®2 PCIe-SSD achieves e. g. transaction rates of up to about 50000 IO/s for accesses that are typical of databases, and data throughputs of up to 1390 MB/s for sequential read accesses - as with video streaming.

The disk I/O performance per server can be further increased by forming RAID arrays with operating system means or by operating these storage media independently. It is possible to operate up to four of these PCIe-SSDs in a number of PRIMERGY servers.

Tuning operations for processor frequency make sense for load profiles with small blocks in order to optimize the disk I/O performance of this storage medium; especially if the accesses are exclusively read-dominated or exclusively write-dominated.

Under certain conditions, tuning operations for processor allocation of the application that creates disk I/Os also make sense for write-dominated accesses. The latter can become necessary in particular for access profiles with small blocks, thus ensuring that the application always runs on the most favorable processor.

Taking the processor architecture into consideration also plays a role when selecting a system.

Literature

PRIMERGY Systems

http://primergy.com/

PRIMERGY Performance

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

Basics of Disk I/O Performance

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

Fusion-io® ioDrive®2 Solid-State Storage devices

Data sheet
http://docs.ts.fujitsu.com/dl.aspx?id=cb40337c-4292-44fe-ac8f-50a96bc653e1

Fusion-io® ioDrive® Solid-State Storage devices (predecessor generation)

Data sheet
http://docs.ts.fujitsu.com/dl.aspx?id=fbbf6a0c-1210-4c23-8df7-081401e9444d

Performance Report PCIe-SSDs
http://docs.ts.fujitsu.com/dl.aspx?id=1d1cc40f-880e-45e8-bfa8-c1426dcb5afa

Information about Iometer

http://www.iometer.org/

Contact

FUJITSU

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

PRIMERGY Product Marketing

mailto:Primergy-PM@ts.fujitsu.com

PRIMERGY Performance and Benchmarks

mailto:primergy.benchmark@ts.fujitsu.com

ioDrive, ioDrive2, ioMemory and ioSphere are registered trademarks of Fusion-io.