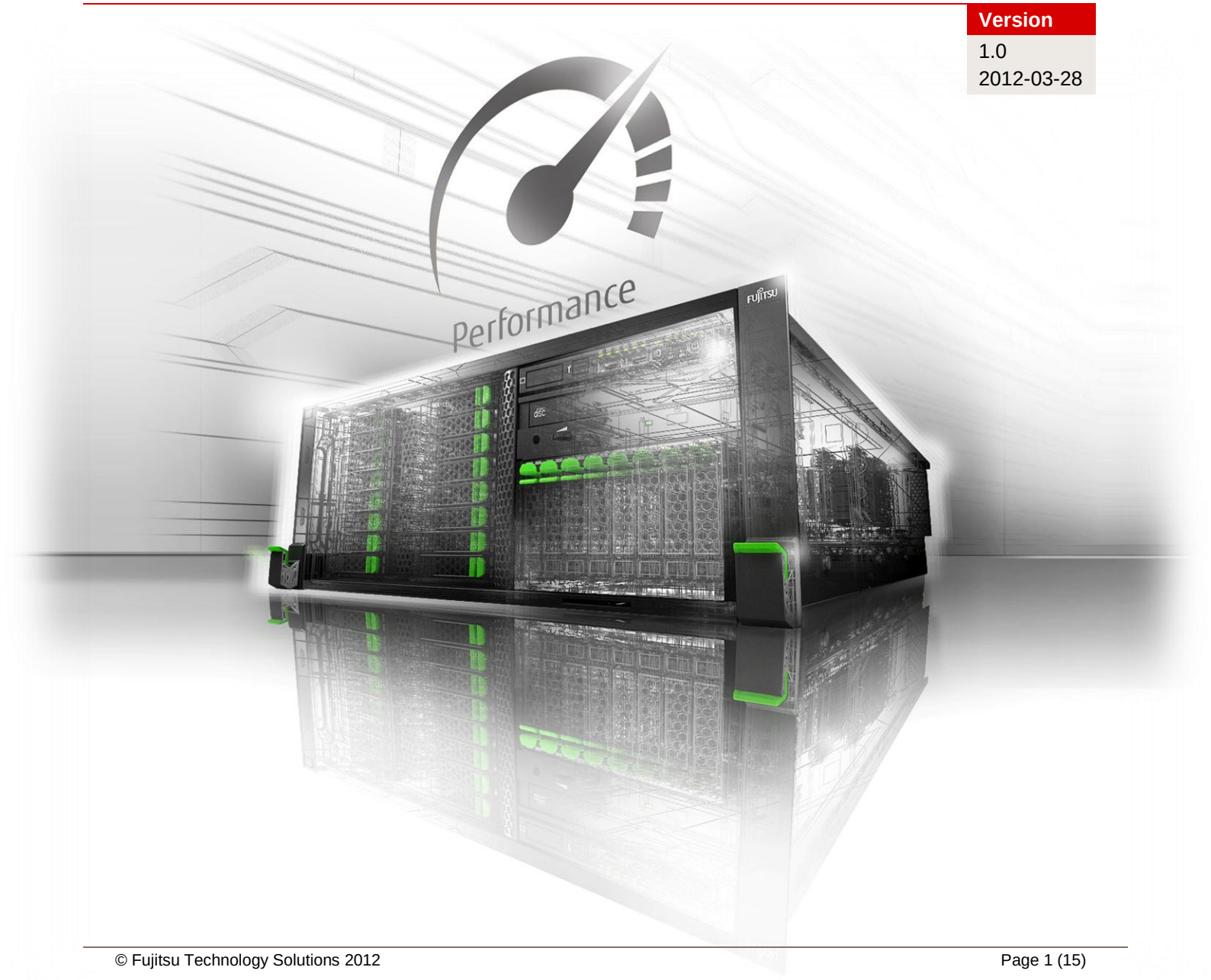
WHITE PAPER
FUJITSU PRIMERGY SERVERS
PERFORMANCE REPORT PRIMERGY SX980 S1

This document contains a summary of the benchmarks executed for the PRIMERGY SX980 S1.

In addition to the benchmark results, an explanation has been included for each benchmark and for the benchmark environment.

Version
1.0
2012-03-28



Contents

Document history

Version 1.0

New:

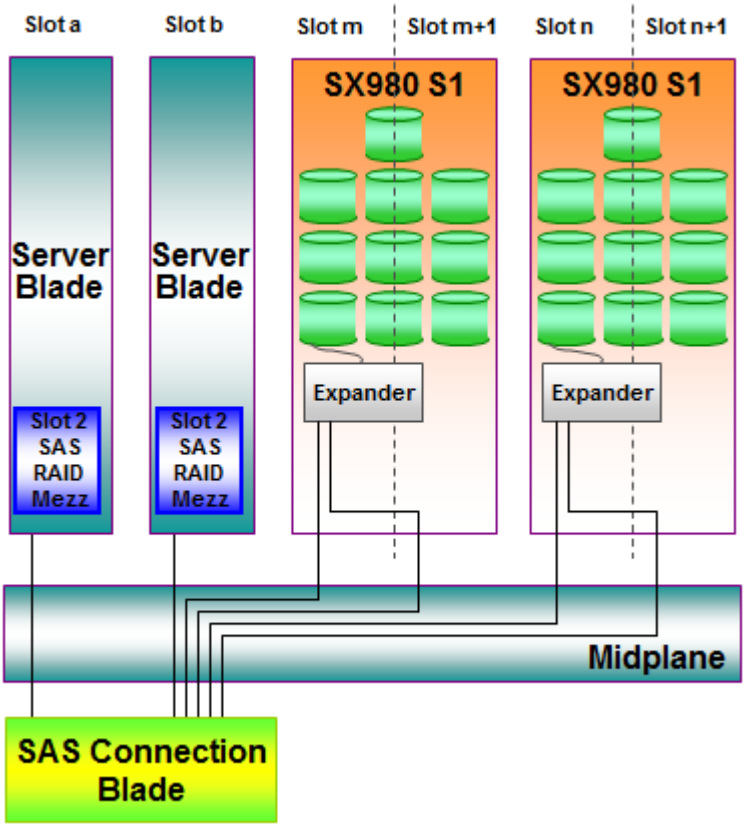
- Disk I/O
  Measurements with “PY SAS RAID Mezz Card 6Gb” Controller

Technical data

PRIMERGY SX980 S1 storage blades can be used to expand the internal disk capacity of a PRIMERGY server blade by up to 9 TB per storage blade. It has room for ten 2.5" hot-plug SAS-2.0 HDDs or SAS-2.0 SSDs, which can make their storage capacity available to all server blades via a SAS connection blade and the midplane of the blade server. The connection in the server blades is via a SAS RAID Mezzanine card.

The diagram opposite provides an overview of the components involved, their internal structure and the connections with each other.

Any required selection of hard disks from one or more storage blades can be allocated to any server blade. This is done via the graphic management interface of the SAS connection blade. Various configuration versions are possible within a blade server here.



Up to six PRIMERGY SX980 S1 can be plugged into the PRIMERGY BX900 S1 blade server. Up to three PRIMERGY SX980 S1 can be plugged into the PRIMERGY BX400 S1 blade server. Each one occupies two slots in the blade server.

Detailed technical information is available in the

- data sheet PRIMERGY SX980 S1
- data sheet PRIMERGY BX900 S1
- data sheet PRIMERGY BX400 S1

Disk I/O

Benchmark description

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

The essential specifications are:

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

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

Standard load 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]*

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 results were determined by way of example on a PRIMERGY BX920 S2.

System Under Test (SUT)	
Hardware (Shared)	
Enclosure	PRIMERGY BX400 S1
Switch	1 × Connection Blade SAS Switch 6 Gb 18/6
Hardware	
Model	4 × PRIMERGY BX920 S2
Disk subsystem	3 × PRIMERGY SX980 S1
Controller	1 × PY SAS RAID Mezz Card 6G (per server blade)
Drive	30 × EP HDD SAS 6 Gbit/s 2.5″ 15000 rpm 146 GB
Software	
Operating system	Microsoft Windows Server 2008 R2 Enterprise
Administration software	ServerView RAID Manager 5.0.2
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 27.07.2006
Measurement data	Measurement files of 32 GB with 1 – 8 hard disks; 64 GB with 9 – 16 hard disks; 128 GB with 17 or more hard disks

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

Benchmark results

The results presented here are designed to help you choose the right solution from the various configuration options of the PRIMERGY SX980 S1 in the light of disk-I/O performance. The determining factors here are the suitable components and their correct parameter settings. These two aspects should therefore be dealt with as preparation for the discussion of the performance values.

Components

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

- “Economic” (ECO): low-priced hard disks
- “Business Critical” (BC): very failsafe hard disks
- “Enterprise” (EP): very failsafe and very high-performance hard disks

The following table is a list of the hard disk types that have been available for the PRIMERGY SX980 S1 since system release.

Drive class	Data medium type	Interface	Form factor	krpm
Enterprise	HDD	SAS 6G	2.5″	10, 15
Enterprise	SSD	SAS 6G	2.5″	-

SAS-HDDs with a rotational speed of 15 krpm have better access times and throughputs than comparable HDDs with a rotational speed of 10 krpm. The 6G interface has in the meantime established itself as the standard among the SAS-HDDs.

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

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

After the hard disks the RAID controller is the second performance-determining key component.

On account of the architecture concept associated with the PRIMERGY SX980 S1 the RAID controller is located in the server blade. The following table is a list of the most important features of the RAID controllers that are available in the server blades in connection with the PRIMERGY SX980 S1. A short alias is specified here for each controller, which is used in the subsequent list of the performance values.

Controller name	Alias	Cache	Supported interfaces		Max. # Disks in the storage blade	RAID levels	BBU/ FBU
PY SAS RAID Mezz Card 6Gb	LSI2108	512 MB	SATA 3G/6G SAS 3G/6G	PCIe 2.0 x8	10 × 2.5”	0, 1, 5, 6, 10, 50, 60	✓/-

System-specific interfaces

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

Controller alias	Effective in the configuration					Connection via expander
	# Disk channels	Maximum throughput of disk interface	PCIe version	PCIe width	Maximum throughput of PCIe interface	
LSI2108	2 × SAS 6G	**973 MB/s**	2.0	x8	3433 MB/s	✓

Thus, the maximum data throughput that is possible in total from a server blade to all the hard disks connected via a SAS connection blade is about 973 MB/s. This also applies if these hard disks are in different storage blades.

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

In addition to the described maximum value from the viewpoint of a server blade, there is a maximum value from the viewpoint of a storage blade. On account of the four physical SAS 6G connections a maximum data throughput of about 1945 MB/s data is possible in total from all the accessing server blades to a storage blade.

Settings

In most cases, the cache of the hard disks 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. The performance can as a result increase more than tenfold for specific access patterns and hard disk types. More information about the performance impact of the hard disk cache is available in the document “Single Disk Performance”. To prevent data loss in case of power failure you are recommended to equip the system with a UPS.

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

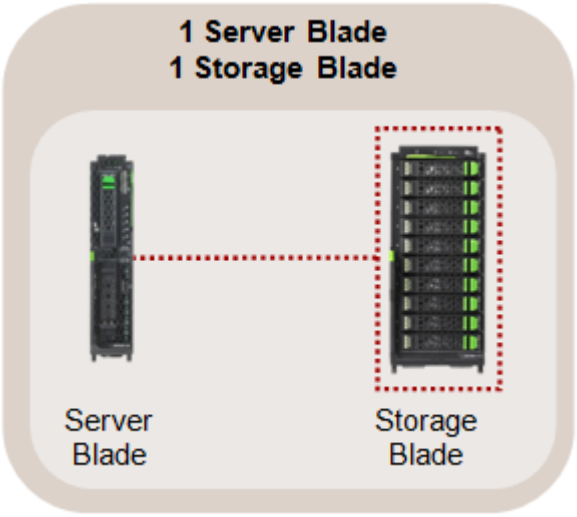
For the purpose of easy and reliable handling of the settings for RAID controllers and hard disks it is advisable to use the supplied RAID-Manager software “ServerView RAID”. All the cache settings for controllers and hard disks can usually be made at one go – specifically for the application – by using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

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

Maximum performance values for one server blade with one storage blade

In general, disk-I/O performance of a RAID array depends on the type and number of hard disks, on the RAID level and on the RAID controller. If the limits of the system-specific interfaces are not exceeded, the statements on disk-I/O performance are therefore valid for all PRIMERGY systems. This is why all the performance statements of the document "RAID Controller Performance" also apply for the PRIMERGY SX980 S1 if the configurations measured there are also supported by this system.

The following list now concentrates on the maximum achievable performance values of the system in connection with a single accessing server blade. The opposite diagram illustrates the basic structure.



The performance values are listed in table form, specifically for different RAID levels, access types and block sizes. The values can be achieved with hard disks (HDDs) that have been selected for optimal performance (the components used are described in more detail in the section Benchmark environment). Furthermore, cache settings of controllers and hard disks, which are made by "ServerView RAID" in "Performance" mode, are used as a basis for all the measured values of this document. More information about the selection of the RAID level and the cache settings for an access scenario is available in the document "RAID Performance".

Configuration version		RAID level	HDDs sequential 64 kB blocks 100% read [MB/s]	HDDs sequential 64 kB blocks 100% write [MB/s]	HDDs random 8 kB blocks 67% read [IO/s]	HDDs random 64 kB blocks 67% read [IO/s]
RAID Controller	#Disks					
1 × LSI2108	2	RAID 1	295	153	903	360
	10	RAID 5	978	868	3102	1701
	10	RAID 10	926	487	4172	2225
	10	RAID 0	981	970	5406	2891

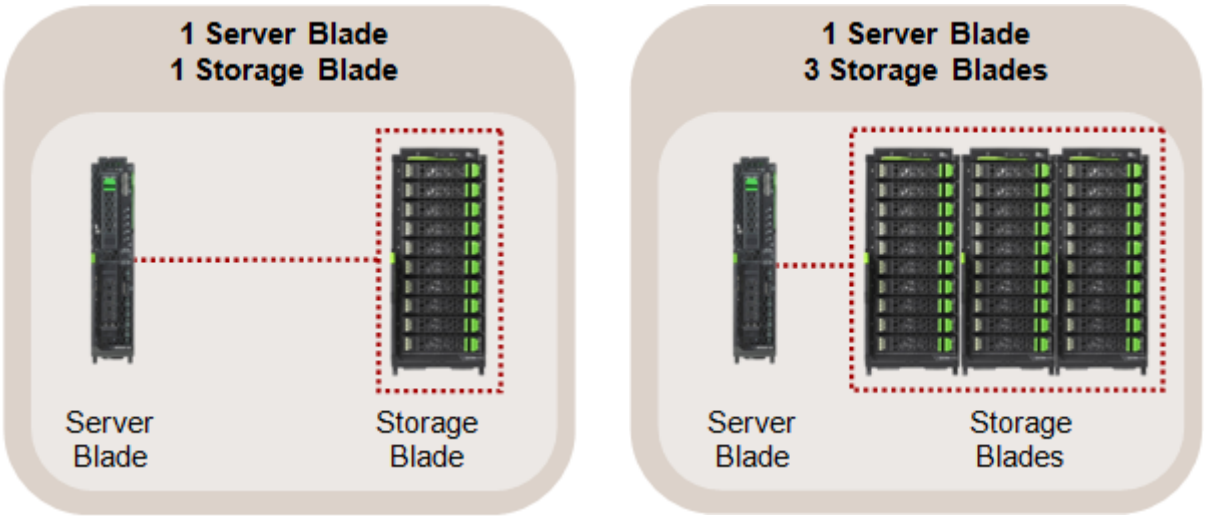
Thus, the PRIMERGY SX980 S1 achieves in connection with a single accessing server blade at full configuration with powerful HDDs (configured as RAID 0) a throughput of up to 978 MB/s for sequential load profiles and a transaction rate of up to 5406 IO/s for typical, random application scenarios.

Performance values for various numbers of storage blades

If several storage blades are used in connection with a server blade, the question arises as to whether the known scaling of the maximum disk-I/O performance also applies with the number of hard disks if the hard disks are housed in different storage blades.

In order to answer this question the number of storage blades associated with the server blade is to be varied in the blade server. This is done by comparing the performance values for one and three storage blades for random accesses. A current, powerful model, which is described in more detail in the section Benchmark environment, is used as the hard disk type.

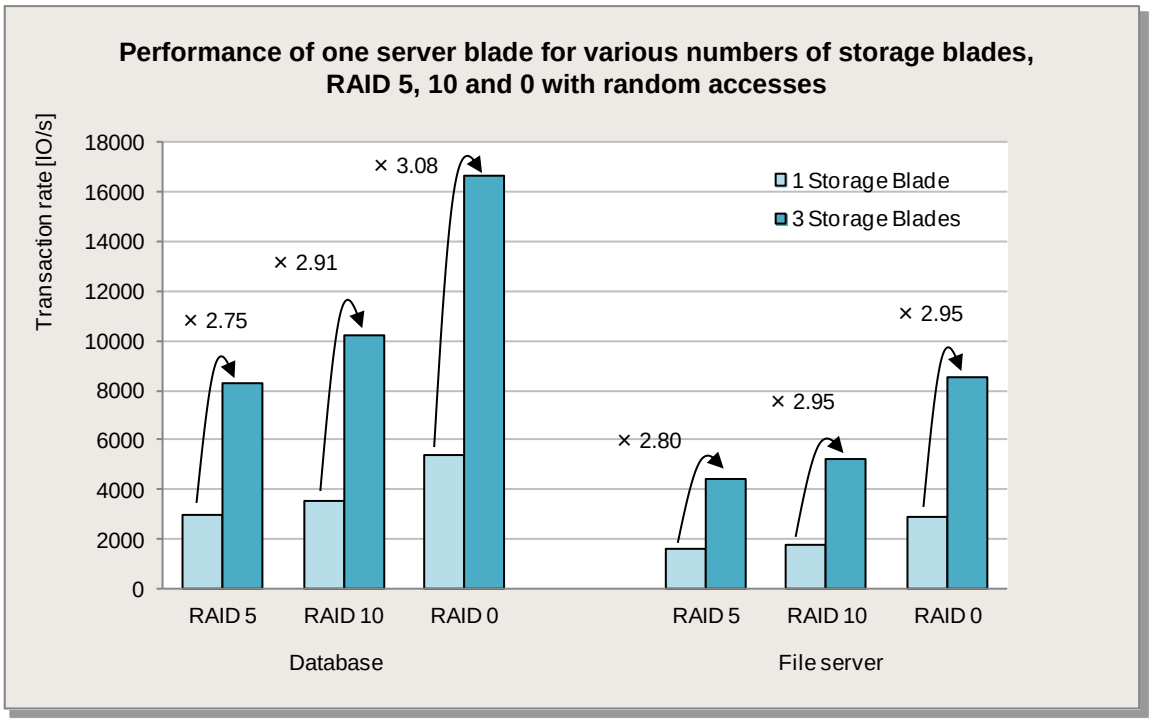
The next diagram illustrates the basic structure.



The cases dealt with here are the RAID levels 5, 10 and 0. The number of hard disks per storage blade is constant for each of the three RAID levels. The exact structure of the RAID arrays is described in the following table.

RAID level	Structure of the RAID arrays for the case of				
	1 storage blade		3 storage blades		
	Number of RAID arrays	Size of a RAID array	Number of RAID arrays	Size of a RAID array	Distribution
RAID 5	2	5 HDDs	6	5 HDDs	2 arrays per storage blade
RAID 10	1	8 HDDs	3	8 HDDs	1 array per storage blade
RAID 0	1	10 HDDs	1	30 HDDs	-

The maximum transaction rates for these cases are shown in the following diagram. The diagram shows the transaction rates for the standard load profile “Database” (random access, 67% read, 8 kB block size) in the left half and for the standard load profile “File server” (random access, 67% read, 64 kB block size) in the right half.



The diagram shows that in all these various cases a server blade with three storage blades achieves approximately three times the transaction rate than with a single storage blade.

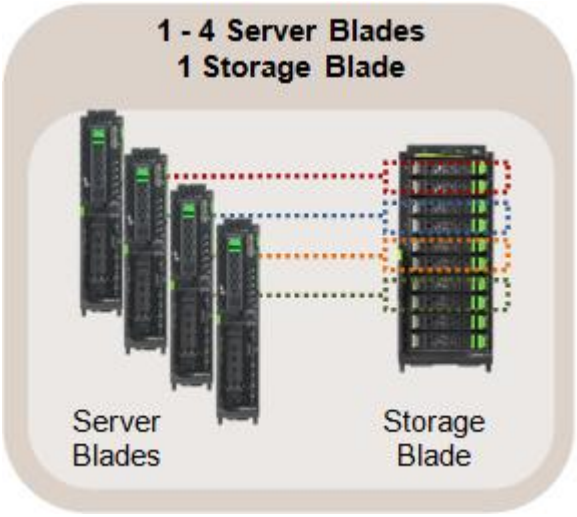
If fully configured with powerful hard disks the maximum possible data throughput for an accessing server blade (approx. 973 MB/s) can already be achieved for sequential load profiles in connection with a single PRIMERGY SX980 S1. In such cases, therefore, no further increase in performance can be achieved if several storage blades are allocated to a server blade. In all the other cases, however, the maximum disk-I/O performance that can be achieved with a server blade increases according to the number of allocated hard disks, regardless of the number of required storage blades. This was demonstrated in the analysis of the random accesses.

Performance values for various numbers of server blades

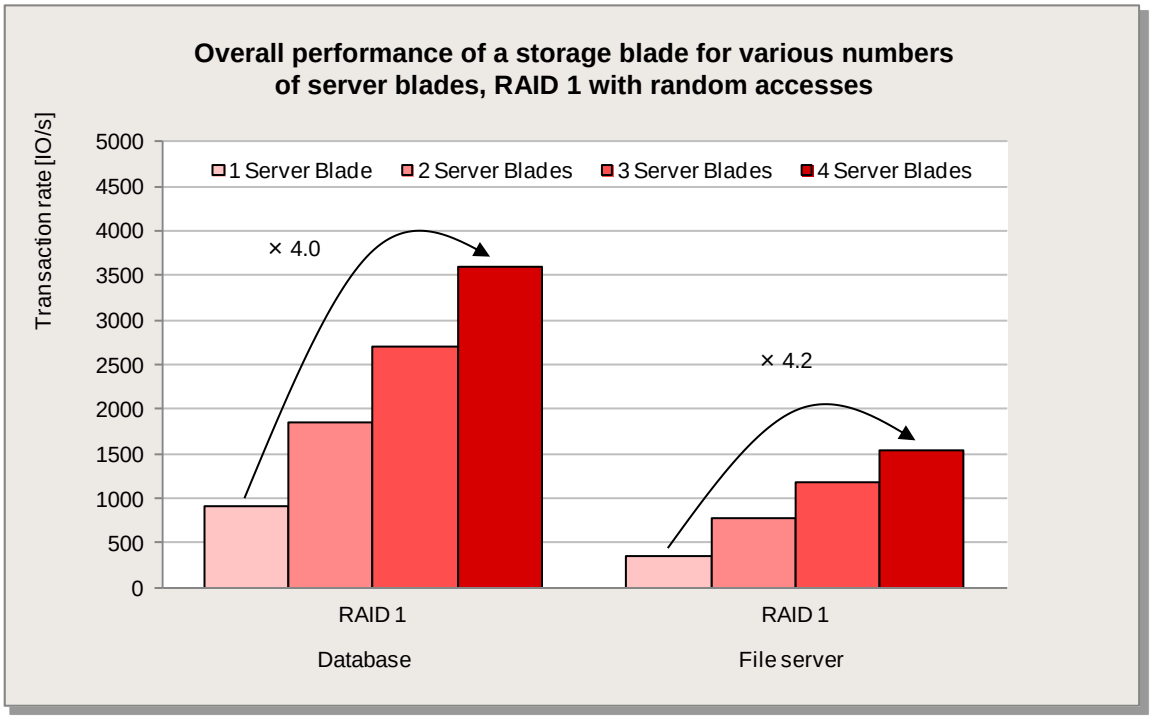
Next, the performance for various numbers of server blades is to be analyzed in connection with a single storage blade.

The first step in this conjunction is to show that the simultaneous access of several server blades to one PRIMERGY SX980 S1 storage blade does not have a significant effect on performance. The second step is to show when several accessing server blades can even be of an advantage.

The measuring setup consists of four server blades and a storage blade PRIMERGY SX980 S1. A RAID 1 array consisting of two HDDs in the storage blade is allocated to each server blade. The disk-I/O performance was analyzed for one, two, three and four server blades during access to their RAID 1 array. The opposite diagram illustrates the basic structure.

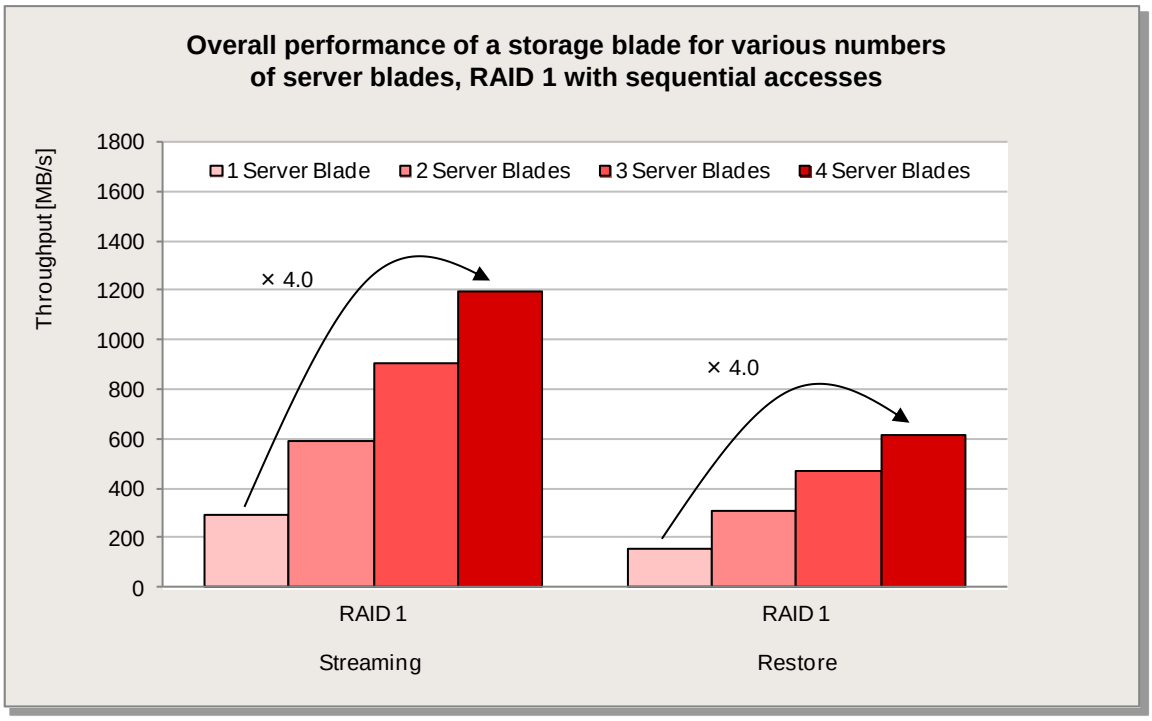


The first diagram is intended to deal with these four cases for the random load profiles. The two groups of columns in the diagram show the maximum transaction rates for the standard load profiles “Database” (random access, 67% read, 8 kB block size) and “File server” (random access, 67% read, 64 kB block size).



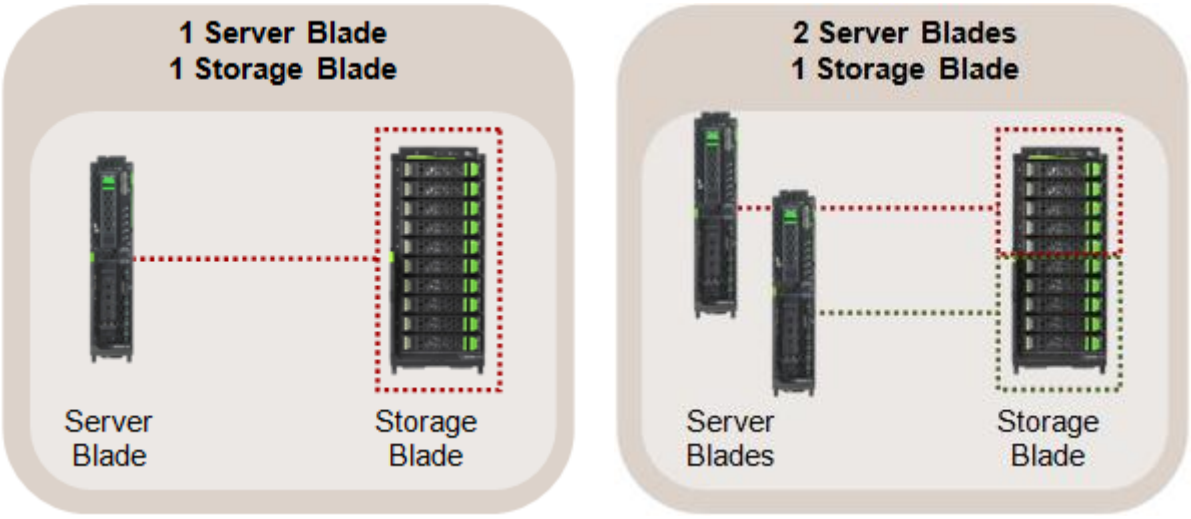
In the diagram you can clearly see that the transaction rates created by several accessing server blades virtually do not impair each other.

The next diagram deals with the four numbers of server blades with RAID 1 for the sequential load profiles. The two groups of columns in the diagram show the throughputs for the standard load profiles “Streaming” (sequential access, 100% read, 64 kB block size) and “Restore” (sequential access, 100% write, 64 kB block size).



You can also see very clearly here that the overall performance of the PRIMERGY SX980 S1 multiplies at approximately the rate at which the number of RAID arrays increases. Thus, the accesses of several server blades to the same PRIMERGY SX980 S1 storage blade do not appreciably impair each other.

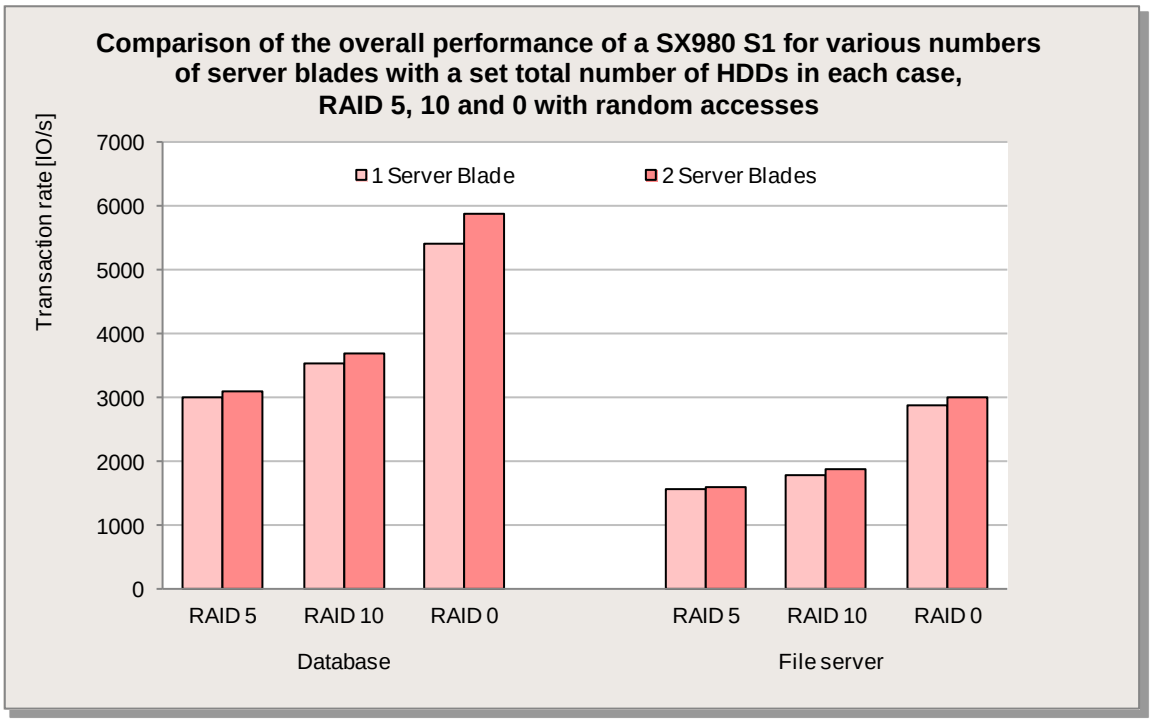
In the first step we kept the number of allocated hard disks per server blade constant. The other question that arises in the light of the best possible utilization of disk-I/O performance of the given hard disks in a PRIMERGY SX980 S1 storage blade is as follows: Is it possible to achieve advantages in performance by not only allocating a set total number of hard disks to one server blade, but to two different server blades - with half to each one? For this purpose, a single storage blade is considered below, the performance values of which are compared for one and two allocated server blades with an identical total number of HDDs in each case. The next diagram illustrates the basic structure.



The RAID levels analyzed are 5, 10 and 0. The size and number of RAID arrays for the individual cases are described in the following table.

RAID level	Structure of the RAID arrays (in 1 storage blade) for the case of			
	1 server blade		2 server blades	
	Number of RAID arrays	Size of a RAID array	Number of RAID arrays	Size of a RAID array
RAID 5	2	5 HDDs	2	5 HDDs
RAID 10	1	8 HDDs	2	4 HDDs
RAID 0	1	10 HDDs	2	5 HDDs

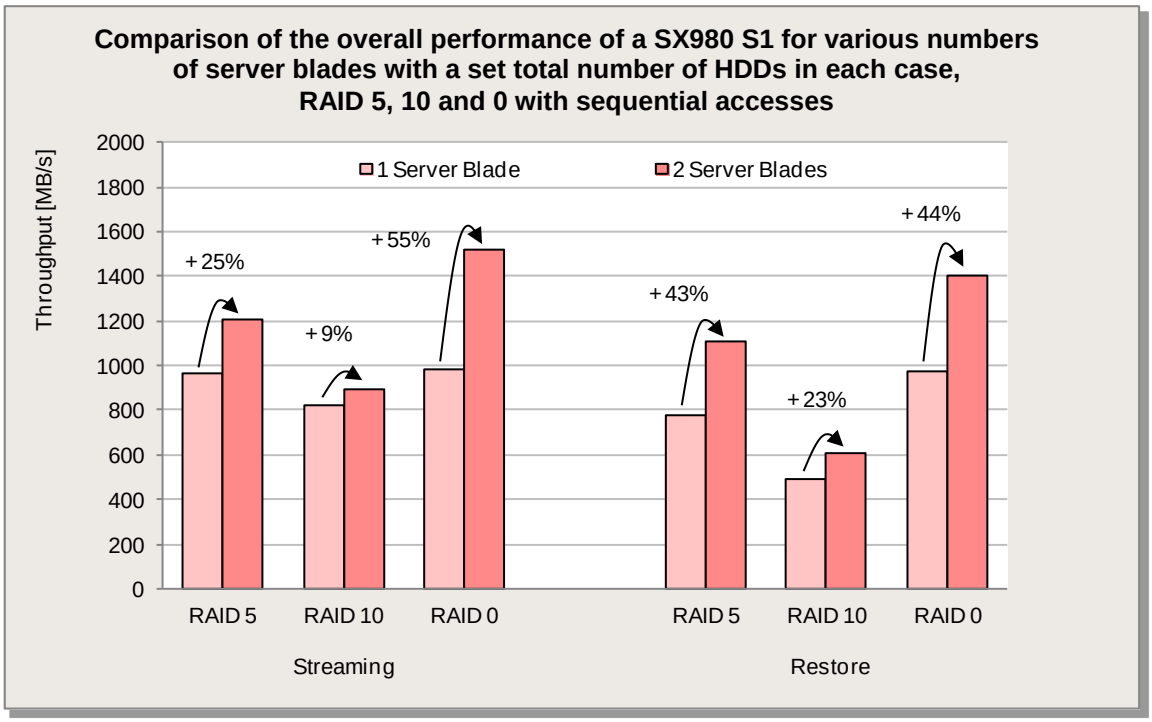
The first diagram deals with random accesses. It shows the maximum transaction rates for the standard load profile “Database” (random access, 67% read, 8 kB block size) in the left half and for the standard load profile “File server” (random access, 67% read, 64 kB block size) in the right half.



For the random accesses dealt with here it can thus be seen for all three RAID levels that the overall performance increases slightly due to the allocation of the load-generating applications to two server blades instead of to one.

In contrast to random accesses, experience has shown that sequential accesses make other requirements of the components involved, because sequential accesses can generate high data throughputs. That is why it is interesting to also make the corresponding performance comparisons for sequential load profiles. As regards its maximum possible data throughput the PRIMERGY SX980 S1 is designed in such a way that a storage blade can also work together efficiently with two server blades, which both generate their maximum disk-I/O load. Consequently, it is possible in specific configurations when the hard disks of a PRIMERGY SX980 S1 are allocated to two different server blades to achieve a significantly higher total throughput than would be possible with only one server blade.

The following diagram illustrates this for the same system configurations and the same RAID configurations that were also previously used for the random load profiles. The diagram shows the maximum data throughputs for the standard load profile ”Streaming” (sequential access, 100% read, 64 kB block size) in the left half and for the standard load profile “Restore” (sequential access, 100% read, 64 kB block size) in the right half.



The diagram shows that in the three RAID levels analyzed a PRIMERGY SX980 S1 achieves significantly higher total throughputs for sequential accesses if two server blades, instead of one server blade, access the same total number of hard disks in a storage blade

Conclusion

The PRIMERGY SX980 S1 enables very versatile and flexible allocations between server blades and hard disks in storage blades. As a result, it is possible to allocate for instance very many hard disks (e.g. 30) to a single server blade, which makes accordingly high performance (up to approx. 973 MB/s per server blade and approx. 1945 MB/s per storage blade) and also storage capacity available. RAID arrays within a storage blade, which belong to different server blades, do not appreciably impair each other from a performance point of view. Such a distribution of the hard disks of a storage blade over several accessing server blades is even advantageous when it comes to making the best possible use of the overall performance of a storage blade.

Literature

PRIMERGY Systems

http://primergy.com/

PRIMERGY SX980 S1

Data sheet
http://docs.ts.fujitsu.com/dl.aspx?id=b4de3482-a07c-4cda-8f9c-de4e3869291b

PRIMERGY BX900 S1

Data sheet
http://docs.ts.fujitsu.com/dl.aspx?id=0a5dcae5-f5a2-42dc-9039-7f887182bc5e

PRIMERGY BX400 S1

Data sheet
http://docs.ts.fujitsu.com/dl.aspx?id=fab1fbb5-5d63-4b2d-816d-3def6c7cb6da

PRIMERGY Performance

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

Disk I/O

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

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

RAID Controller Performance
http://docs.ts.fujitsu.com/dl.aspx?id=ada7c1bf-74e3-4953-b783-839cdeec790b

RAID Performance
http://docs.ts.fujitsu.com/dl.aspx?id=c55404e9-69c6-4d1e-a556-6a322b27da5b

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