

White Paper

FUJITSU Server PRIMERGY

HDD or SSD or NVMe drives for servers – what is more suitable?

SATA, SAS, SSD and NVMe: numerous selections for just one drive? This white paper explains how to select the correct storage drives for servers - to help administrators



CONTENT	
HDD or SSD or NVMe drives or solid state disk drives for servers – what is more suitable?	2
Function of a hard disk	2
Technical details	4
Quality factors	7
Provision in servers	7
Drive classes	9
Selecting the configuration	12
Why hard disks from Fujitsu?	12
Conclusion	12

HDD or SSD or NVMe drives for servers – what is more suitable?

When considering how to increase server performance, most people think about swapping the processor, expanding memory capacity or increasing I/O throughput and the useable bandwidth. Storage media are often not considered as a major performance factor when it comes to servers even though server technology - like the rest of the market - has recently undergone tremendous progress. The developments made in old and new interfaces, the new types of memory access technology and the hardware involved are the main reasons why administrators and IT managers have to select whether a hard disk or SSD or NVMe drive is more suitable for their servers. This White Paper is designed to help administrators and IT managers understand the detailed requirements involved and to select the correct memory configuration for their Fujitsu server systems.

Storage media determine how well a server can "serve" an application or network in which it is incorporated. Storage media must not only provide suitable speeds and performance but much more. They must be reliable and have low latency and low power consumption rates; they must be easily adaptable to the various customer applications. All these parameters are involved in ensuring an optimal configuration.

Main component

Since the beginning of computer technology hard disks have been used not only to store operating systems, programs and data, but also as an extension of a computer's RAM. Before we look at the details, let us understand that the term "hard disk" is often used for saving data not only on hard disks but on magnetic disks as well as on drives using flash memory and which do not have any moving parts. The latter are thus known as "solid state" drives (abbreviated to SSD). Today we have a lot of traditional workloads storing their data on slower HDDs (providing maximum storage consolidation) or the much faster SSDs (which offer the best price/performance), both of which reuse traditional interface protocols. With the ever-growing demand for additional data streams to the storage system, unlocking parallel access to storage became a challenge. This led to the entire storage world coming together to come up with a new interface protocol that will revolutionize storage – and NVMe was the result. This White Paper will make clear distinctions regarding these storage drive categories.

Function of a hard disk

The function of a hard disk in a server is different to the functions used in desktops, notebooks or other computers. Its function is mainly determined by the task of the server. The system is also integrated in a network. For this reason, hard disks in servers must handle a higher work load and serve a more users than in standard PCs. Data transfer requirements can be issued at any time and have to be processed with as little latency as possible because any unforeseen delays would result in the decentral user having to wait and that is something he cannot tolerate. This means that the hard disk of a server must be functional and active at all times whereas the hard disk of a desktop can be switched to "idle" when it is not required. And it is not just "one" hard disk: servers should never be equipped with only one hard disk drive (HDD or SSD or NVMe) but at least two which are bundled in a RAID array in order to provide greater performance and reliability.

Typical applications and their requirements

When selecting hard disks for a server, there are three main factors to consider:

- What are the applications on the system?
- Which data is to be saved?
- What is the significance of the application and data for the company?

From a technical viewpoint, the emphasis should be on the performance of a drive which has three main elements:

- The first is **speed** or "how fast can a hard disk provide and save information". Decisive factors in this conjunction are the number of I/O processing transactions per second (IOPS), which a hard disk can process, as well as the quantities of raw data (payload without header), which are transferred within a specific time frame; this is referred to as data throughput.
- Another factor regarding performance is **latency**, i.e. the time that elapses between requesting the data and the moment when the data reaches the user.

- The third factor is: How **reliable** is the drive? The lifespan can become a decisive factor if the data is to be saved for a long period of time and access is required after or outside of the production cycle.

The typical usage scenarios for servers are:

- **Mail servers** are responsible for all types of communication. This includes email traffic and other types of messaging. Modern applications are also responsible for the connection to mobile devices or home office where employees need to access company resources via the Internet. Mail servers are basically "hard disk servers" as the CPU workload is not a bottleneck. Mail servers require reliable hard disk drives. Speed is also a prerequisite, but is not so significant here as in other situations, except for those extremely comprehensive mail databases where low latency has a very high priority.
- **Application servers** run programs for individual users with office tasks, or programs for a company or for numerous users in a transaction service or even for several million users in the Internet. This sector requires the fastest and most reliable of drives - an absolute must.
- **Storage servers** are mainly used to store files. They not only contain their own hard disks but are also connected with external arrays connected. Reliability is also a top priority in this situation. Storage servers are usually full of data during the various phases of their lifespan. In other words: this data can be essential for the company in its production line or other business processes. It must be kept online or nearline or it is stored in a manner which conforms to archiving regulations, i.e. off-line. This does not mean that archived information is not so important; it is simply not required as often and consequently does not have to be saved on very fast or very powerful drives. Storage servers are usually linked with backup solutions away from their location, for example with tape libraries, optical storage media or other equipment, such as online storage services. However, if access to archived contents should be necessary on a more frequent basis, a company will possibly require fast drives or a combination of reliable and fast drives.
- **Database servers** store a database which is accessed either via the corporation network or via the Internet. Furthermore, the database must sometimes be connected to special application servers. Database servers have to simultaneously handle numerous parallel requests from various users which make low latency and high reliability decisive factors. Multiple tiers of HDD and SSD can be effective for database servers to acquire better performance. In particular, PCIe SSD with NVMe Interface is the right choice to have high throughput rates and low latency. NVMe deploys flash storage on a PCIe bus and offers up to 64,000 parallel pathways from the CPU, thus overcoming serial limitations in storage I/O processing. The protocol capitalizes on the multiple parallel, low latency paths to flash storage to offer faster storage response times and higher throughput for speedier application performance.
- **Streaming servers** provide multimedia data to a company's employees or its customers. The users can access such data either within a limited period or have 24hr access (if systems are accessible via the World Wide Web). This scenario requires fast drives so that large files can be accessed; they must have the necessary speed and performance.
- **Virtual servers** run a virtual operating system and one or more virtual machines where a different application runs on each machine. These servers have an important role to play - just like application servers - or are even more important as each system - and not just one - can be used to host several critical business applications.
- Such a **combined use** is customary when a company only has one single extremely large server as a host or when several machines work together as if they are only one. In this situation, hard disks should be assigned in various chassis for the various scenarios or the hard disk most suitable should be chosen depending on the task with the highest resources requirement should a user prefer to work with just one type.

Technical details

As already stated, the same type of hard disks is not always used in servers. The first technical difference has already been explained: HDDs or SSDs or NVMe drives a combination of all three can be used in servers as storage media. The further differences lie in the interface technology, drive size and internal components. These factors determine the performance, reliability and energy-efficiency of a server.

HDD, SSD and NVMe technologies

The rise and convergence of new age technologies like Artificial Intelligence, Big Data, Analytics and Internet of Things have led to the advent of next-generation data-intensive workloads and ever-proliferating applications which are exponentially increasing the demand for accelerated data storage.

Today we have a lot of traditional workloads storing their data on slower HDDs (providing maximum storage consolidation) or the much faster SSDs (which offer the best price/performance), both of which reuse traditional interface protocols. With the ever-growing demand for additional data streams to the storage system, unlocking parallel access to storage became a challenge.

This led to the entire storage world coming together to come up with a new interface protocol that will revolutionize storage – and NVMe was the result. NVMe deploys flash storage on a PCIe bus and offers up to 64,000 parallel pathways from the CPU, thus overcoming serial limitations in storage I/O processing. The protocol capitalizes on the multiple parallel, low latency paths to flash storage to offer faster storage response times and higher throughput for speedier application performance.

Here are main technical features of HDD, SSD and NVMe drives:

- From the outset hard disk drives (HDDs) have consisted of several disks equipped with a magnetic layer. A read/write mechanism accesses the data on the individual disks. Depending on the size and type of hard disk, up to 10 terabyte (TB) of information can be saved on HDDs. This conventional type of data storage offers a satisfactory degree of reliability whereby the performance depends on various factors (number of revolutions, interface, cache size) all of which will be explained later in this document
- Solid state drives (SSDs) are based on flash memory, basically the same type of memory as used in SD cards or non-volatile memory in mobile devices. However, there are certain differences between an SSD and an SD card: SD cards have a different type of controller and are regarded by the operating system as a replaceable medium. However, the main difference is in the reliability. SD cards are only suitable for short-term data exchange whereas SSDs have much greater reliability and are thus ideal for storing data on a long-term basis. SSDs have a data throughput level that is much greater than that of HDDs but their performance varies which greatly depends on the type of access. A direct comparison of HDDs and SSDs in benchmarks is usually a difficult process. Traditional HDD benchmarks usually concentrate on areas where HDDs have difficulty, namely rotation latency and search times. As SSDs do not have to rotate nor look for local data, they seem to do much better in these tests compared to HDDs. SSDs thus offer much better performance values than HDDs in most of the application scenarios. The earlier very short life expectancy of SSDs has since been constantly increased in recent years. The SSDs supplied by Fujitsu are ideal for long-term use in servers.
- NVMe is the promising new optimized, high-performance interface that is designed for systems that use PCIe based SSDs. NVMe provides a standard access protocol for flash to leverage the PCIe bus to directly connect CPUs to attached storage, thereby reducing latency, increasing throughput and most importantly offering massive parallelization capabilities. So while mainstream Flash SSDs will continue to be available with SAS or SATA for price/performance scenarios, NVMe SSDs will be very attractive for usage scenarios that demand extreme performance. In fact, Gartner statistics project an attach rate of 69% for PCIe NVMe SSDs in servers and 26% for PCIe NVMe SSDs in storage arrays in the next two years. Although still very nascent, NVMe over Fabric extends the latency, parallelism and performance advantages to the entire fabric (over Fibre Channel, RDMA and TCP, etc.). With NVMe-oF, the potential impact of NVMe would not just accelerate SSDs – it will also power computing, fabrics, storage, system management and more.

Interface

From the numerous interfaces available on the market, only three types are significant for servers.

- **SATA (Serial Advanced Technology Attachment) is the successor to the parallel ATA (PATA) system.** Due to its use for serial data transfer, SATA is much faster than the parallel connection. Whereas PATA offered a maximum speed of 133 MB per second (MB/s), the current SATA standard is 600 MB which equals a raw bandwidth of 6 Gbit/s per hard disk. SATA hard disks are ideal when the focus is on low-price, dependable read/write speeds.

- **SAS is the abbreviation for "Serial Attached SCSI".** This term refers to SCSI, the founder of modern server interfaces (Small Computer System Interface). SAS uses the same command sequence as SCSI and adapts it to a fast serial connection. The interface offers a bandwidth of 6 Gbit/s and a throughput of 600 MB/s per hard disk is possible. The SAS protocol thus offers the advantage of a more stable and faster protocol which ensures a higher rate of user data, improved interference resistance and the option of a dual-port connection of storage media for cluster operation. SAS data media are thus always required when the focus is on high throughput rates and high-level reliability.
- **PCIe SSDs are not connected via the normal SATA or SAS controllers, but directly to the PCIe bus of the server system.** Consequently, this by far exceeds the limited transfer rates provided by the SAS/SATA interface. PCIe has a number of advantages, including lower latencies, scalable performance/bandwidth, better I/O performance and low power consumption, thanks to its direct connection to the CPU. Currently, PCIe 3.0 can support up to 16 lanes, with each lane supporting 1Gb/s of throughput per lane. Besides the noticeable performance improvement, this enables NVMe to support massive concurrence of parallel requests. For example: NVMe can support up to 64,000 queues and 64,000 commands per queue, while SATA can queue only 32 commands in a single queue, and SAS can queue about 256 commands in a single queue.

An interesting fact for administrators is that SAS is "downwards compatible" with SATA, i.e. SATA and SAS drives can be connected to an SAS controller. However, this does not function in the reverse direction - SAS drives cannot be connected to SATA host adapters. In order to obtain the best results, future-oriented administrators should equip all servers with SAS controllers, including those that only have SATA drives. You can thus convert to SAS drives without any problems should the performance requirements increase. However, this is not compulsory as controllers can also be upgraded in a running system without any problems.

Size / Form Factor

The size of the drives installed in a server determines both its storage capacity as well as its energy efficiency. Three different types of form factor are currently available for customers, of which only two, 3.5" and 2.5", are relevant for servers.

- **3.5" is the most widely used size for HDDs.** Only drives of this size can accommodate the maximum data quantity of 16 TB (as of July 2019). 3.5" HDDs are therefore the generally recommended solution for high storage requirements. However, they consume more power than the smaller 2.5" drives. SSDs are not available in this size. The features of 3.5" hard disks are high maximum capacity and a more favorable price. They thus offer the best price per GB.
- **2.5" is the usual size for HDDs and SSDs.** Although these drives are only one inch smaller, they use much less power than 3.5" hard disk drives. A maximum of 5 TB data can be saved per drive in 2.5" drives. Compared to their larger companions, 2.5" hard disks are not only power-savers, but also offer the best read/write speeds when operated in a network of several data media. They are thus ideal for use in "green" implementation scenarios and for maximum performance.
- **SATA DOM (DOM=Disk on Module)** is a kind of Form Factors based on SSDs. It is an extremely space and energy-saving flash memory that is designed and released as a cost efficient boot drive in servers. For a range of PRIMERGY servers, Fujitsu offers a DOM with a SATA interface, which can be inserted directly into the SATA port of the system board. Installing OS on a DOM makes it possible to maximize the data storage capacity of the server using the other HDDs or SSDs.

Hard disk manufacturers are constantly able to increase the maximum capacity of drives. Fujitsu will offer any new hard disks as an option as soon as they are available for delivery.

Internal components

When looking more closely at this topic, we see that HDDs and SSDs differ in the way they save information.

- Data on an HDD is saved on "disks", actually glass disks with a magnetizable material layer. The higher the density of the respective disk and the more disks a drive contains, the higher its capacity. All disks at a specific speed that is measured in revolutions per minute (rpm) which defines its bandwidth, access time and also its power consumption. Servers with high performance requirements are usually equipped with hard disks with speeds of 7,200, 10,000 or 15,000 rpm. Faster drives not only consume more power, but also require suitably advanced cooling mechanisms. Another problem is caused by vibrations. When the hard disks in a chassis rotate at different speeds, they can impair operations as they disrupt each other's write cycles. All drives that work in the same area should thus always have the same rotation speed (i.e. the same interface). If drives with different speeds or interfaces are combined in a server with different application areas (see the table at the end of this White Paper), they should always be accommodated in different chassis units.

- **SSDs have no moveable parts, they save information in flash-memory cells.** The tests implemented by Fujitsu reveal improved I/O performance with a factor of 100 compared to that with hard disk drives. And at the same time they only use a fifth of the power consumption as there is no need for a motor to run when the drive is inactive. Due to the absence of moveable parts, SSDs are not subject to any mechanical wear and tear and are not sensitive to heat and vibrations. However, one disadvantage of SSDs is their limited lifespan. Whereas HDDs age due to their mechanical wear and tear, SSD wear and tear is due to an electrical effect. Although read operations from intact flash cells can be performed on an unlimited basis, a flash cell can only handle 3,000 to 100,000 write operations depending on its quality. Reading information is then no longer reliable. This fast wear and tear caused problems for the SSDs particularly at the very beginning. This has been counteracted for several years now by the "Wear-Leveling" procedure. The controller in the flash drive distributes write operations to all the memory cells so that each cell - where possible - is written equally. This distribution procedure leads to a lifespan similar to or even exceeding that of conventional hard disks.

Quality factors

When selecting the correct drives for a server, the following important features should be taken into consideration:

- The lifespan of a hard disk is difficult to estimate. Manufacturers use a value known as MTBF. This abbreviation stands for "mean time between failures" (average number of operating hours between downtimes) and refers to a time after which the drive will probably fail. MTBF is measured in hours which means that the value is always enormous. However, please note that this is merely a static value which does not guarantee that the drive will not fail before this time is reached.

Apart from raw capacity and performance values, customers should also note that most server drives have to be run continuously or to be more precise: they must be able to run continuously without any interruption. Fujitsu only sells drives that fulfill these conditions. In contrast, the MTBF for desktop HDDs is specified on the basis of an 8-hour operational day. SSDs contain flash memory cells that are wearing parts, so SSDs can only tolerate a limited number of write jobs and has a built-in Wear-Out indicator. In contrast to HDDs, not only is there therefore a general, average lifespan, but also one related to write jobs, which is called write endurance. Depending on usage, there is a possibility to reach its write endurance limit within the system warranty period. In this case the warranty for such a component, as an exception to the system warranty, is restricted to the time period until the indicator reaches the exhaust level. The DWPD (device writes per day) can be considered in order to assess the lifespan expectancy of the SSDs, i.e. how often the storage medium can be written to per day. Management software "ServerView RAID Manager" is available for monitoring the write endurance for SSDs. The write endurance should be checked regularly so that the replacement of SSDs can be initiated before reaching the limit. The write endurance of the SSDs offered by Fujitsu is specified in the corresponding product data sheet.

- A special feature of SSDs is their restricted data retention capability when switched off. If an SSD is removed from a server and, for example, put in a cabinet as backup, the information stored will remain available for ten years as best-case. Factors, for example flash technology (SLC/MLC), the previous usage intensity (DWPD) or the environmental temperature will shorten the storage period. The minimum storage period in relation to the respective forecast lifespan is six months for SLC SSDs and three months for MLC SSDs. In order to ensure that all HDDs and SSDs fulfill these conditions, they must be subjected to strict tests and certified accordingly. This is the case with all drives installed by Fujitsu in its PRIMERGY servers. This is why each drive is the best solution for the planned specialized area of usage.

Provision in servers

The features and the quality of each drive only accounts for half their performance in a server. The other half depends on the quality of interaction between the drives. As stated earlier, hard disks should never be used individually in servers in contrast to desktop and mobile systems. Instead, they should be linked together via a host adapter in order to assure the required redundancy via RAID configurations. In order to attain optimal throughput, these controllers should also be equipped with a sufficient amount of cache memory.

RAID configurations

RAID stands for "Redundant Array of Independent Disks" and contains a logical system for mirroring and for striping data across several drives in order to attain shorter access times and higher data protection. There are currently twelve so-called RAID levels that offer different redundancy and security levels; however, all have to accept certain compromises regarding capacity and speed. The configurations used most frequently are RAID 0, 1 (including derivative), 5 and 6.

- **RAID 0 (striping)** distributes data across several hard disks so that their capacity can be fully used and their overall read/write speed can be improved. The disadvantage of RAID 0 is that it does not provide any data protection: all the information on all hard disks is lost if a hard disk fails.
- **RAID 1 (mirroring)** duplicates all the information across all the individual hard disks in the array, thus offering complete redundancy. Two or more hard disks have completely identical data at all times. No data is lost as long as one hard disk remains functional. The total capacity of the array thus corresponds to that of a hard disk. Each hard disk in the array is always completely identical with any other hard disk in the array at any given moment in time. Disadvantage: Only 50% of the "pure" (nominal) storage capacity can be used; the other half is used for mirroring, i.e. storage arrays must always be "twice as large" as the amount of data saved.
- **RAID 10 (mirroring with striping)** combines RAID 1 and RAID 0 and thus requires four hard disks.

- **RAID 1E** mirrors individual data blocks on the following hard disks. RAID1E requires an odd number of hard disks. Data loss only occurs when two sequentially connected hard disks fail (or the first and last hard disk in a chain of hard disks). The usable capacity corresponds to 50% of nominal capacity.
- **RAID 5 (striped-hard disks with parity)** combines at least three hard disks so that there is no data loss if any of the hard disks should fail. The storage capacity of the array is thus reduced by one hard disk; in other words: one hard disk per RAID array ensures redundancy. Data loss cannot be prevented if more than one hard disk fails.
- **RAID 6 (striped hard disks with double parity)** can compensate the failure of two hard disks.

Cache

One important means of increasing speed is to provide an adequately sized controller cache. The cache acts as a buffer for the large quantities of read/write operations which hard disks have to perform. A cache should be provided even when the number of operations is low. For security reasons, an emergency power supply (BBU, Battery Backup Unit) is normally included in the controller so that any data in the cache that has not been saved is not lost during a power failure. In contrast, a write cache in hard disks should only be used if it is protected by an emergency power supply. On the other hand, SSDs have an internal emergency power supply which supplies sufficient power to an internal write cache, thus ensuring that all write operations are completed even when there is a power failure.

The correct controller

When selecting controllers, current and future requirements must be taken into account. In addition to an SAS interface and sufficient cache, you must also have a suitable number of ports so that an increasing number of drives can be connected in order to suit the ever-increasing requirements of a company. Fujitsu always ensures that the controllers in a system have a sufficient number of ports, or that a suitable SAS expander is installed on the backplane. Several PRIMERGY servers have an onboard SATA RAID controller with a maximum transfer speed of 3/6 Gbit/s. If higher speeds are desired, a high-performance SAS controller can be added as required. Non-RAID controllers should be used for connecting tape drives or external storage units as they have their own RAID controllers.

Drive classes

In order to summarize all these aspects and simplify the customer's decision, Fujitsu has defined several hard disk classes:

- **Economic (ECO).** The drives in this class have a low unit price. However, their performance and reliability levels mean that they are only suitable for entry-level applications. They should be used in non-critical areas with low I/O traffic and moderate speed requirements as higher workloads can impair their reliability. ECO drives provide speeds of up to 5.4/7.2K rpm and have a SATA interface.
- **Business Critical (BC) or Nearline.** The drives in this class offer the highest capacity with the lowest cost per GB; these drives are designed to provide good performance and suitable reliability. Depending on the server implementation, BC drives can be equipped with an SAS or and SATA interface and offer a speed of 7.2K rpm. If top I/O throughput rates are required, then Enterprise HDDs or SSDs should be used.
- **Enterprise (EP).** Drives in this class provide maximum performance and reliability. They are designed to cope with maximum workloads and the highest throughputs. This class uses the SAS interface and the rotation speeds are 10/15 K.
- **SSD.** Similar to HDDs, SSDs have several capacity points and interfaces like SAS, SATA and PCIe. Another important and SSD-specific criterion is their write endurance. In this case, SSDs can be roughly divided into three categories, which meet completely different requirements:

Endurance Class	DWPD	Description/Suitability
Read-Intensive	<= approx. 1	SSDs of the lower price category. They are suited to load situations that are characterized by a low write intensity. DOM is in this class and best suited for Boot drive. Examples: Boot drive, streaming services
Mixed-Use	> approx. 1 and < approx. 10 usually approx. 3	SSDs of the medium price category. They are suited to load situations that are characterized by a moderate write intensity. Examples: File servers, web servers, mail servers
Write-Intensive	>= approx. 10	SSDs of the upper price category. They are suited to load situations that are characterized by a high write intensity. Examples: Virtualization servers, databases

Selecting the configuration

The correct decision regarding the hard disk configuration is a fundamental factor for long-term server performance. The correct drives must enable the server to function in a suitable manner. The following table illustrates which drives and interfaces should be implemented in the various usage scenarios.

Use of server	Importance of data	Decisive performance factors	Drive class and type	Interface/controller	Comments
E-mail	High	Average speed, low latency	BC HDD/SSD	SATA	Combined use of HDD and SSD with an extremely comprehensive mail database
Application	High	High speed	EP HDD/SSD	SAS	SATA for SSD, which is linked with an SAS controller; A combination with NAS or DAS strongly recommended.
Database	High	Low latency	EP SSD or HDD	SAS (SATA from an SSD drive viewpoint)	SATA SSD, with an SAS controller connected; SAS SSD where applicable or an SSD/HDD combination for low latency
Storage	From low to high	Average speed, low latency, depending on the frequency of the required access to database and user	ECO, SSD or HDD for online or nearline, HDD for offline maintenance	SAS 2 for online or nearline, SAS 1 or SATA for offline maintenance	Combined use of SAS and SATA, however not in a shared RAID array; a backup or separate medium/location is strongly recommended.
Streaming	Medium	High speed	BC HDD	SATA or SAS	High performance SATA drives could possibly suffice as the speed is high, but access is mainly sequential.
Virtual	High	High speed	EP HDD	SAS	SSDs are not recommended, due to the high number of random write operations required.
Mail and applications combined	High	High speed, low latency	BC or BC and EP, SSD and HDD	SAS	A RAID array is recommended for each implementation area
Mail, applications, storage combined	High and low	Depending on the use in each area, see above	BC or EP SSD or HDD for mail and applications, HDD for data storage	SAS 2 for mail and applications, SAS or SATA for data storage	One RAID array is recommended for mail and applications, another one for data storage

NVMe Storage Adoption: Recommendations

- NVMe memory technologies will become available in storage system by several vendors this year. To truly benefit from NVMe technology, a storage system needs to be designed from the ground-up and be optimized end-to-end. It stands to reason that some early designs may blend current with new technologies, thereby compromising the advantages of NVMe. Go only for fully developed NVMe systems and conduct technical evaluations to ensure that the desired performance improvements are realized.
- There will be a lot of news about NVMe over Fabrics as the next big thing, but as this standard has only been released recently, do not expect mature designs in the next two years.
- This year will be a year of early adopters, current SAS-based all-flash storage will become mainstream. Current SAS-based all flash systems are fully mature and increasingly improve the price per performance and price per capacity ratio. Their performance characteristics will meet the demands of most application scenarios. Consequently, their adoption will further grow in the range of 30% p.a. NVMe will not replace these designs but will be added as an additional storage tier for usage scenarios with highest demands on storage response time. Look for vendors who can offer SAS- and NVMe based all-flash storage with full management compatibility to avoid operational silos.

- Drivers and optimizations for operating systems, hypervisors and adapters are currently being developed and will take more time for mainstream acceptance of NVMe storage. The performance and latency benefits will appeal initially to very high-end users who demand extreme throughput and IOPS for future workloads along with lowest possible latency. Expect a long-term coexistence of SAS and NVMe-based all-flash storage systems until NVMe will become the new mainstream.

Why hard disks from Fujitsu?

You could be asking why you cannot use generic drives from accessories trade instead of hard disks and SSDs from Fujitsu. They all look alike. However, the decisive factor for the security of your data and the performance of your system is in those more "invisible" differences.

The only difference immediately obvious regarding the required hot-plug function is the Fujitsu drive chassis. This ensures that hard disks and SSDs can be easily removed and inserted in the event of an error and when making upgrades. The status of the data medium can be easily seen via the control lamps on the front. Read/write activity, idle or errors are indicated by flashing lights.

The main differences are not visible in the drives. New drives in the target systems are tested in great detail, checked and adapted to the PRIMERGY servers.

The first step is to adapt the drive firmware - together with the manufacturer - to the requirements of PRIMERGY servers. This includes setting parameters and activating special firmware that has been specially developed for Fujitsu. All the problems determined during later tests are corrected and the corrections are integrated into the modified firmware. Incidentally, in a best-case scenario generic firmware will only include these modifications by the drive supplier after a certain delay.

The next step is to subject the drives to various shock, vibration and temperature tests in order to discover any mechanical weak points. Constantly running endurance tests simulate many years of data media usage. Additional checks ensure that the defined read/write speeds are actually attained. When the drives have passed such tests, they are subsequently tested in the target system and then certified.

Drives are not just tested in these early stages. Once they have been successfully certified, the components are continually subjected to goods entry tests (random tests in accordance with AQL values). 100% of the delivered data media is initially tested. If no faults are encountered during these initial tests, the random sample lots are later reduced to 10% of the delivered components. If errors start to occur during later tests, the number of lots tested is increased again.

And even after the drives have been delivered, Fujitsu still checks the reasons and frequency of errors reported to Fujitsu Service. Problems can thus be avoided at an early stage, for example by introducing modified firmware.

If you decide to increase your memory capacities after a while, you will be able to obtain compatible hard disks and SSDs from Fujitsu for at least three years. This is because new drives are also tested and certified for old systems.

As you can see: hard disks and SSDs from Fujitsu provide you with tested quality as well as the additional benefits from the knowledge gained by our experts over many years regarding storage technology and system integration. Starting with the initial certification of new drives, the fine adjustment of the firmware and up to the quality control of individual deliveries and supplies, Fujitsu does everything possible to guarantee you the best performance and highest fail-safety.

Please note that components not yet released by Fujitsu can affect your guarantee.

➔ Warranty conditions: <http://support.ts.fujitsu.com/warranty/Index.asp?LNG=COM>

Conclusion

Selecting the correct HDD or SSD or NVMe drives for the servers in your company is not a simple task. As well as technical requirements, administrators and IT managers must also consider other aspects, such as low investment costs and low total cost of ownership, when it comes to finding the best solution for their network. There are indeed a lot of options available and our Fujitsu consultants will be glad to provide you with support.