Technical White Paper
FUJITSU Storage ETERNUS AB and ETERNUS HB Feature Set

This white paper provides an overview of the main features supported by the FUJITSU Storage ETERNUS AB all-flash and ETERNUS HB hybrid series. It highlights their benefits and use cases and briefly describes each feature.

Contents

Introduction

As data is the most important asset in any company, storage system plays a vital role in the IT infrastructure of every enterprise. IT administrators need to ensure that data is stored on reliable, highly available platforms that can scale efficiently in order to handle ongoing business changes.

New IT trends are also imposing new challenges on storage systems. The increasing use of business analytics and data warehousing, the hype around big data, server and desktop virtualization, the enormous growth of unstructured data are just a few of the examples.

These trends not only require massive storage capacities, but also – and more importantly – storage performance. Parameters, such as IOPS (Input & Output Operations per Second), latency and bandwidth are gaining in significance. Many enterprises are purchasing additional storage systems, just to provide the required performance. The result is storage capacity overprovisioning as well as higher operational and capital expenditure.

Due to the diversification of storage requirements in recent years, it has become necessary to strengthen and expand storage portfolios. In order to support the latest technologies, Fujitsu has decided to collaborate with other companies to strengthen its product portfolio. In the HDS field, this will be achieved by strengthening collaboration with NetApp (a leader in Data Management Services).

The ETERNUS HB series is a hybrid storage that offers outstanding cost performance in a SAN environment with application-based data management. Versatile and reliable, the HB series handles huge amounts of data for uses such as media & entertainment, Hadoop and analysis of the earth's crust, in addition to applications such as databases and BI[1]. This storage system is suitable for high performance computing applications that require dedicated storage.

The ETERNUS AB series is an all-flash array that is suitable for performance-oriented environments that deliver outstanding IOPS performance and sub-millisecond latency. The AB Series, which combines excellent performance with enterprise-class reliability, can improve processing speed and responsiveness in OLTP[2], data warehouses, BI and virtualization to achieve business results quickly.

1 Abbreviation for Business Intelligence. With this method, the user analyzes and converts vast amounts of business data stored in their company's information system to make business and management decisions. Software and information systems that are used for this purpose are called BI tools or BI systems.
2 Abbreviation for On-Line Transaction Processing. This method instantly executes transaction processing that reliably executes multiple processes that are related to each other, based on requests from terminals.

Hardware Architecture

ETERNUS HB

ETERNUS HB portfolio is a set of hybrid storage systems which can be implemented to leverage and optimize storage load by combining traditional spinning drives together with flash technology. Redundant architecture ensures that in case of component failure storage system continues to work as it should and the part replacement does not require operation interruption.

Within ETERNUS HB storage portfolio Fujitsu offers NVMe host interfaces and SSD cache to meet performance hungry applications requirements. As automation is becoming a standard in IT world new hybrid family natively supports a RESTful web services API.

ETERNUS HB – Entry Models

The scalable entry models – ETERNUS HB1100/HB1200, ETERNUS HB2100/HB2200/HB2300 – are designed in a compact style to ensure an optimized footprint. These models are ideal for performing media streaming and data analyses in mid-sized companies and remote offices. They comprise the following hardware:

- Controller Shelf is always equipped with two Controller Modules (CM) and two Power Supply Units (PSU) in the rear as well as disk drives in the front
- Drive Shelf contains disk drives installed in the front, two PSUs and two IO modules (IOM) in the rear. The drive shelfs are always equipped in four Back-End (BE) ports. Various types of disk drives are described in detail in the following sections. Number and type of drive enclosures can be adjusted to the capacity requirements. Available variants of disk shelves are 2.5", 3.5" and 3.5" high-density
- The controller module (CM) runs the firmware-based operating system and contains host interface ports (HIC), management Ethernet ports, drive expansion ports, cache memory, SAS expander and BBU. All communication between two controllers goes through the midplane located in the shelf. SAS interfaces between the CMs are cross-connected to provide a redundant path to the disks installed on controller shelf. Each drive enclosure has four interface ports that are connected directly or through other shelfs to both CMs for path redundancy

ETERNUS HB – Midrange Models

The midrange models of the ETERNUS HB series are ETERNUS HB5100 and ETERNUS HB5200. They are designed to provide capacity, high performance thanks to NVMe ready interfaces and reliability. ETERNUS HB Midrange family is an ideal choice for enterprise customers who run data analyses and performance databases. They contain the following hardware:

- Controller Shelf base components are redundant Controller Modules (CM) and redundant Power Supplies (PSU) in the rear as well as disk drives in the front
- Drive shelves used in Midrange systems are available in 2.5" and 3.5" high-density variant. Their connection architecture is the same as in entry-level shelves
- The basic Midrange system consists of two interconnected controller modules, each with two host interface ports (HIC), SAS expander and internal cache which is mirrored between CMs to ensure data redundancy. Fujitsu has enriched Storage portfolio with high-speed NVMe compatible host interfaces available in ETERNUS HB5100 and ETERNUS HB5200. Options available to choose are: 100 Gbps InfiniBand, 100 Gbps NVMe over InfiniBand or RoCE and 32 Gbps NVMe over FC. Each system can be expanded with 2.5" or high-density drive shelves to meet capacity requirements.

ETERNUS AB

The ETERNUS AB series is a family of all-flash storage systems oriented on high performance data operations, low latency and availability. All these factors are available to achieve thanks to all-flash optimized storage systems. For the first time Fujitsu offers end-to-end NVMe array to provide system designed for the customers who desire high performance computing and AI workloads with ultra-fast response times. So as new hybrid series, ETERNUS AB supports RESTful API.

ETERNUS AB series combines excellent performance with enterprise-class reliability. It can improve processing speed and responsiveness in OLTP, data warehouses, BI and virtualization to achieve business results quickly.

ETERNUS AB – Entry Models

The entry model of ETERNUS AB – ETERNUS AB2100 – offers low cost and affordable performance. This model is ideal for performing media and data analyses for midsize businesses and remote offices.

Entry-level ETERNUS AB can grow up with additional 2.5" drive shelves and that allows reaching 1.4 PB of pure all-flash capacity. Different interface types can be mixed within controller modules. This makes ETERNUS AB series possible to adjust to any existing customer environment – based on Ethernet, FC or even SAS connections.

AB3100 is an entry-level end-to-end NVMe storage system ideal for midsize customers requiring consistent high-throughput for streaming media, analytics, AI/ML, and HPC workloads needing InfiniBand. It offers 367 TB of NVMe capacity within only 2U of rack space. It can deliver up to 670K IOPS and 24 GBps throughput with very low latency.

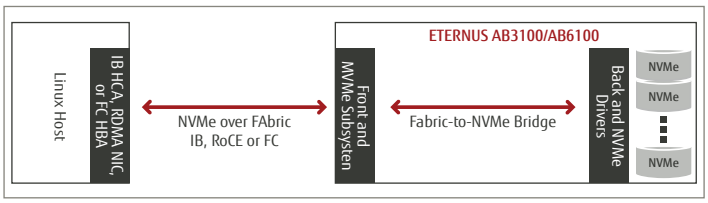
ETERNUS AB – Midrange Models

There are two ETERNUS AB midrange models – ETERNUS AB5100 and ETERNUS AB6100.

ETERNUS AB5100 is based on SAS architecture all-flash model designed for enterprise customers running data analyses and performance databases. It can scale up to 1.8 PB of all-flash capacity. With NVMe interfaces it enables to achieve sub-100 microsecond response time and up to 1 million IOPS.

ETERNUS AB6100 is the very first end-to-end NVMe storage array offered by Fujitsu with 2 million IOPS within only 2U of rack space. Fully populated storage system offers 367 TB of NVMe capacity.

Hardware architecture of this array differs from other traditional storage systems. Instead of most common SAS/SATA interface, disk drives are connected via PCIe interface and bridged to front-end NVMe interfaces. NVMe offers a more streamlined command set, to process a I/O request, than the SCSI and ATA command sets do. NVMe requires fewer than half the number of CPU instructions than the SCSI command set does with SAS devices and the ATA command set uses with SATA drives.



NVMe supports 64,000 commands in a single message queue and a maximum of 65,535 I/O queues. By contrast, a SAS device typically supports up to 256 commands, and a SATA drive supports up to 32 commands, in one queue.

Interfaces

The ETERNUS AB/HB provides different types of host interfaces, offering customers full flexibility in selecting the most appropriate data center infrastructure. Fibre Channel is the most widely used storage networking technology as it is highly reliable, efficient and secure. Up to 32 Gbps bandwidth is supported, offering NVMe over Fabric. 100 Gbps InfiniBand is now released to support such technologies as NVMe over InfiniBand, NVMe over RoCE, iSER or SRP. 1 Gbps, 10 Gbps & 25 Gbps iSCSI are also supported for connecting to IP networks. They are simple to operate and hence preferred by many customers. The host connections can also be realized by 12 Gbps SAS interfaces, except ETERNUS AB3100 and AB6100. Interfaces used as backend connectivity are 12 Gbps SAS and PCI Express in end-to-end NVMe storage system.

Types of Disk Drives

As the storage requirements differ according to the type of data and the frequency of its usage, various types of disk drives need to be supported in order to allocate the right disks for each type of data. Some data is mission-critical so it has to be accessed immediately in order to avoid loss of revenue or productivity degradation. This data must be stored on drives with very high performance, such as SSDs or even NVMe's. On the other hand, some types of data do not require very high performance, but need to be stored for longer periods. Then, it can be stored on more cost-efficient, high-capacity disks such as Nearline SAS disk drives, enabling the customer to balance speed, capacity and costs. In addition, 2.5" and 3.5" enclosures can be mixed in the same storage configuration of ETERNUS HB disk storage systems. SSDs and SAS disks can be mixed in the same 2.5" drive enclosures. 3.5" enclosures can host SSDs, SAS drives and Nearline disk drives, which provide up to 12 TB of capacity.

NVMe drives

Non-volatile memory express (NVMe) is a new, optimized, high-performance interface for PCIe based Solid State Drives (SSDs) that truly unleashes the power of Flash. NVMe capitalizes on the massively parallelized, low-latency paths to flash storage to offer faster storage response times and higher throughput, thereby powering application performance.

Solid State Drives (SSD)

Solid State Drives use semiconductor memory to store data. They contain no motors or moving parts, and thus have a much higher read/write access speed and reduced power consumption. They benefit those applications with high random access requirements, such as databases. In addition, with no motors or moving parts, they are more reliable than disk drives. The SSD used in ETERNUS AB/HB have enterprise-class performance and reliability. While maintaining compatibility with traditional disk drives, they support low-power consumption and high-speed operation.

Online SAS Disk Drives

For data volumes that are frequently accessed, but still do not require the very high performance of SSD, SAS disk drives are used, providing a balanced mix of performance and capacity, while keeping costs at a moderate level.

Nearline SAS Disk Drives

Storing infrequently accessed data on high-performance storage devices generates unnecessary costs. To meet the growing demand for cost-effective storage of less frequently accessed data, Fujitsu provides high-capacity, highly reliable, yet cost-optimized Nearline disk drives in its ETERNUS HB disk storage systems. This combination of online disk drives and Nearline disk drives in the same drive enclosure enables ETERNUS HB storage systems to support cost-effective operations, such as disk-to-disk backup and long-term storage of reference data.

High-Density Drive Enclosures

High-density drive enclosures are available for ETERNUS HB disk storage systems in order to meet the growing demand for more capacity. They house up to 60 3.5" disk drives in one DE, which can provide up to 720 TB in 4 rack units.

SANtricity operating system (OS)

ETERNUS AB/HB systems run on the SANtricity software platform. SANtricity OS and management software give storage administrators the ability to achieve maximum performance and utilisation of their flash arrays, perform online configuration and management tasks, and manage at scale.

SANtricity OS accelerates performance for high-IOPS and low-latency applications as well as high-bandwidth and high-throughput applications – all from the same enterprise-grade storage building block. It provides:

- Reliability and High Availability: SANtricity OS and ETERNUS AB/HB deliver optimal data access, integrity, and protection. Its automated I/O path failover, RAID and DDP protection, proactive monitoring, and online maintenance capabilities mean that data will always be available.
- Online Administration: All management tasks can be performed while the storage remains online with complete read and write data access. Storage administrators can use SANtricity management software to make configuration changes, perform maintenance, or expand storage capacity without disrupting I/O to attached hosts

- Advanced Data Protection: The SANtricity OS exceeds basic high-availability features to offer optimal data access, integrity, and protection. Its automated I/O path failover and extensive online configuration, reconfiguration, and maintenance capabilities enable data to be always available

SANtricity management software

SANtricity management software provides the building blocks that enable manageability advantages through configuration flexibility, custom performance tuning, scalable management, and complete control over data placement. It provides interfaces to manage in the way best for your environment and unlock the power of SANtricity OS.

Capacity Management

Dynamic Disk Pools (DDP)

In high-performance computing (HPC) and other big-data environments, success depends on the speed at which data is acquired, processed, and distributed. These environments use staggering amounts of data that can cripple most storage systems and break traditional architectures. For these sites, the ability to keep a system up and running at a consistently high level has replaced maximum raw performance as a critical need.

The Challenge

Big-data sites often have thousands of drives supporting a clustered file system; and with that many drives, failures are inevitable. As drive capacities continue to grow, traditional RAID technology struggles to keep up. Rebuild times on large-capacity drives can range from 18 hours for an idle system to multiple days or a week for an active system. Because idle time is rare, a drive failure and subsequent rebuild process can significantly affect a system's performance – in some cases up to 40%.

Dynamic Disk Pools

Dynamic Disk Pools (DDP) provides enormous value to sites with vast amounts of data supporting high-bandwidth programs and complex application processing. Its next generation technology minimizes the performance impact of a drive failure and can return the system to optimal condition up to eight times faster than traditional RAID. This powerful combination helps AB/HB storage systems deliver consistently high performance for maximum productivity.

Dynamic Disk Pools distributes data, parity information, and spare capacity across a pool of drives. Its intelligent algorithm (seven patents pending) defines which drives are used for segment placement; making sure data is fully protected. And its flexible disk pool sizing provides optimal utilization of any configuration for maximum performance, protection, and efficiency.

Consistent Performance

Large-scale compute clusters demand multiple gigabytes per second of bandwidth. For these sites, a drop in performance means that jobs run long or don't complete in their allotted window. Dynamic Disk Pools delivers and maintains exceptional performance under all conditions, whether optimal or under the stress of a drive failure.

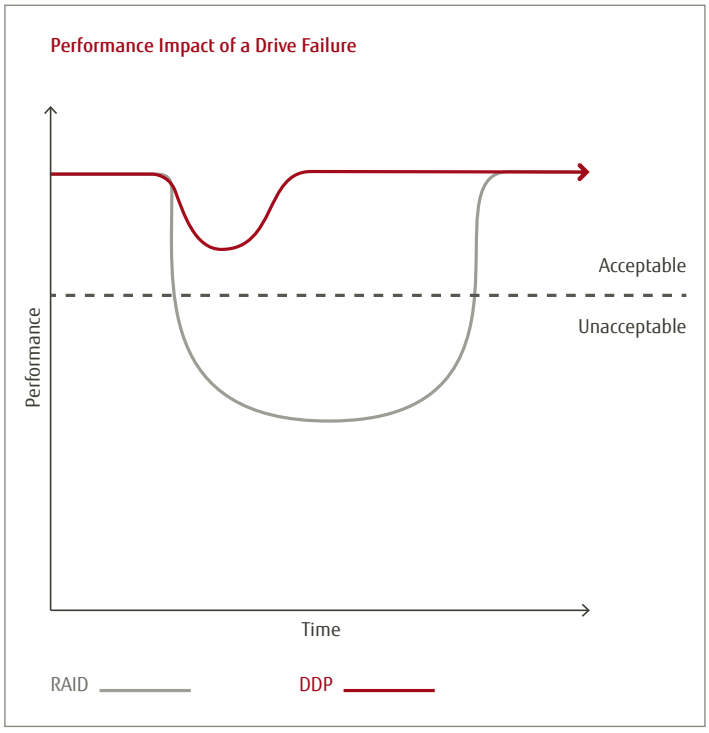
DDP minimizes the performance impact of a drive failure in multiple dimensions. By distributing parity information and spare capacity throughout the disk pool, DDP is able to use every drive in the pool for the intensive process of rebuilding a failed drive. This dynamic rebuild process is the reason that DDP can return the system to optimal condition up to eight times faster than traditional RAID. What's more, by distributing the rebuild workload across all drives, the overall impact of the process is greatly reduced. Rebuilds take less time and have less impact – a win-win for bigdata sites.

Improved Data Protection

Dynamic Disk Pools offers a level of data protection that simply can't be achieved with traditional RAID. Shorter rebuild times significantly reduce exposure to multiple cascading disk failures – a real concern as drive capacities get larger and larger. And thanks to its patented prioritize reconstruction technology, DDP is actually able to increase protection levels as the pool gets larger. Add to this the advanced protection features and extensive diagnostic capabilities that are standard with HB/AB series and you have a storage system that is optimized for excellent data protection.

Extreme Versatility

Maximum performance is often at odds with efficiency. High-performance applications require an optimized stripe size that typically doesn't align with the number of drives in the storage system. The result is either unused spindles, non-optimized stripe sizes, or the elimination of hot spare drives. Because of its extreme versatility, Dynamic Disk Pools is able to address wide-ranging application requirements without sacrificing efficiency. Drives can be configured into one large disk pool to maximize simplicity and protection, or into multiple smaller pools to maximize performance for clustered file systems. Different drive types can be used to create storage tiers, such as performance pools and capacity pools. Disk pools can also reside in the same system with traditional RAID groups. And flexible disk pool sizing – ranging from just 11 drives up to a full configuration – means an optimized fit for any configuration.



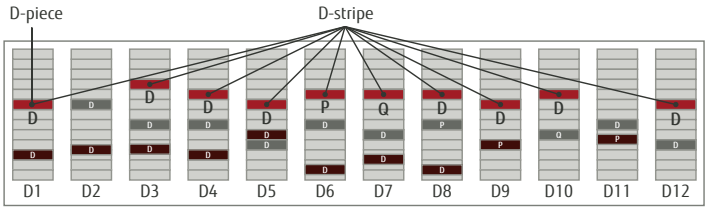
Dynamic Disk Pools is designed to maintain high performance even after drive failure

Technical overview

With DDP technology, ETERNUS AB/HB and management software allows you to create pools in addition to traditional volume groups (generally referred to as RAID groups). A pool can range in size from a minimum of 11 drives to as large as all the drives in a storage system. Pools can consist of either hard disk drives (HDDs) or solid-state drives (SSDs). In addition, pools and volume groups can coexist in the same system. For example, with a 24-drive storage system in which all drives have equal capacity, the following lists some possible combinations:

- One 8-drive RAID 10 (4+4) volume group, and one 16-drive pool
- One 24-drive pool
- Two 12-drive pools
- One 5-drive RAID 5 (4+1) volume group, one 4-drive RAID 10 (2+2) volume group, and one 15-drive pool

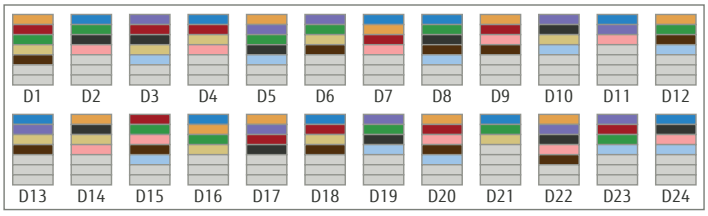
Distributed data allocation



D-piece and D-stripe example

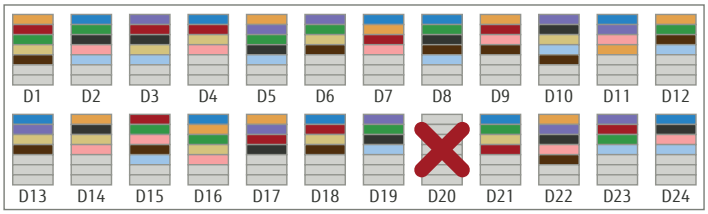
Data allocation distributed throughout the drives which consists the pool, keeping a redundancy equivalent to RAID6. Maximum installation number of DDP drives can be configured.

Data reallocation when a drive failure occurs



24-drive pool

Now suppose that one of the drivers in the pool fails, as below.



24-drive pool with one driver that has failed.

In the example in the figure, alternative spaces for the five D-pieces which exist in D20 when a drive failure occurs in the drive are prepared one each in the drives of D2, D10, D11, D16, and D21.

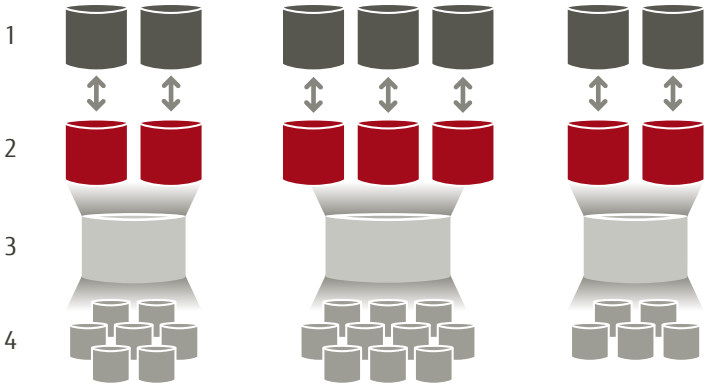
Then the data is written into the D-piece by using the RAID6 data reconfiguration method.

Pools and volume groups

To provision storage, you create either a pool or volume group that will contain the Hard Disk Drives (HDD) or Solid State Disk (SSD) drives that you want to use in your storage array. Physical hardware is provisioned into logical components so that data can be organized and easily retrieved. There are two types of groupings supported:

- Pools
- RAID volume groups

The pools and volume groups are the top-level units of storage in a storage array: they divide the capacity of drives into manageable divisions. Within these logical divisions are the individual volumes or LUNs where data is stored. The following figure illustrates this concept.



1 Host LUNs
2 Volumes
3 Volume groups or pools
4 HDD or SSD drives

When a storage system is deployed, the first step is to present the available drive capacity to the various hosts by:

- Creating pools or volume groups with sufficient capacity
- Adding the number of drives required to meet performance requirements to the pool or volume group
- Selecting the desired level of RAID protection (if using volume groups) to meet specific business requirements

You can have pools or volume groups on the same storage system, but a drive cannot be part of more than one pool or volume group. Volumes that are presented to hosts for I/O are then created, using the space on the pool or volume group.

Pools

Pools are designed to aggregate physical hard disk drives into a large storage space and to provide enhanced RAID protection for it. A pool creates many virtual RAID sets from the total number of drives assigned to the pool, and it spreads the data out evenly among all participating drives. If a drive is lost or added, System Manager dynamically re-balances the data across all the active drives.

Pools function as another RAID level, virtualizing the underlying RAID architecture to optimize performance and flexibility when performing tasks such as rebuilding, drive expansion, and handling drive loss. The RAID level is automatically set at 6 in an 8+2 configuration (eight data disks plus two parity disks).

Drive matching
You can choose from either HDD or SSDs for use in pools; however, as with volume groups, all drives in the pool must use the same technology. The controllers automatically select which drives to include, so you must make sure that you have a sufficient number of drives for the technology you choose.

Managing failed drives
Pools have a minimum capacity of 11 drives; however, one drive's worth of capacity is reserved for spare capacity in the event of a drive failure. This spare capacity is called "preservation capacity."

When pools are created, a certain amount of capacity is preserved for emergency use. This capacity is expressed in terms of a number of drives in System Manager, but the actual implementation is spread across the entire pool of drives. The default amount of capacity that is preserved is based on the number of drives in the pool.

After the pool is created, you can change the preservation capacity value to more or less capacity, or even set it to no preservation capacity (0 drive's worth). The maximum amount of capacity that can be preserved (expressed as a number of drives) is 10, but the capacity that is available might be less, based on the total number of drives in the pool.

Volume groups

Volume groups define how capacity is allotted in the storage system to volumes. Disk drives are organized into RAID groups and volumes reside across the drives in a RAID group. Therefore, volume group configuration settings identify which drives are part of the group and what RAID level is used.

When you create a volume group, controllers automatically select the drives to include in the group. You must manually choose the RAID level for the group. The capacity of the volume group is the total of the number of drives that you select, multiplied by their capacity.

Drive matching
You must match the drives in the volume group for size and performance. If there are smaller and larger drives in the volume group, all drives are recognized as the smallest capacity size. If there are slower and faster drives in the volume group, all drives are recognized at the slowest speed. These factors affect the performance and overall capacity of the storage system.

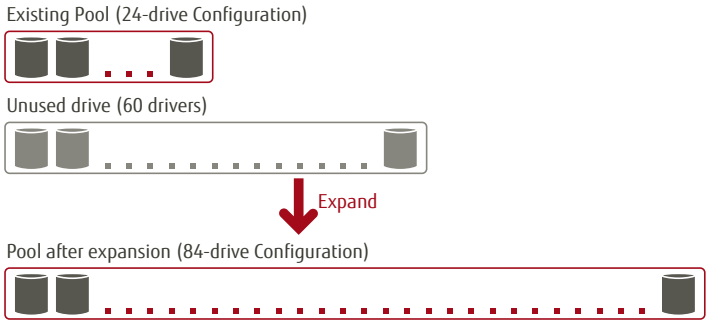
You cannot mix different drive technologies (HDD and SSD drives). RAID 3, 5, and 6 are limited to a maximum of 30 drives. RAID 1 and RAID 10 uses mirroring, so these volume groups must have an even number of disks.

Managing failed drives
Volume groups use hot spare drives as a standby in case a drive fails in RAID 1/10, RAID 3, RAID 5, or RAID 6 volumes contained in a volume group. A hot spare drive contains no data and adds another level of redundancy to your storage array.

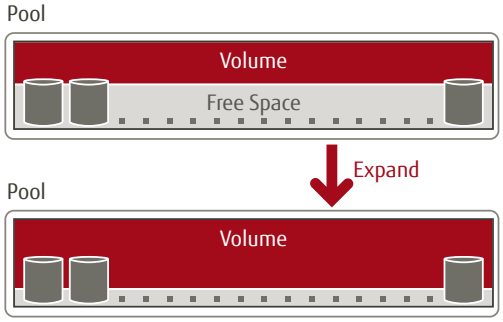
If a drive fails in the storage array, the hot spare drive is automatically substituted for the failed drive without requiring a physical swap. If the hot spare drive is available when a drive fails, the controller uses redundancy data to reconstruct the data from the failed drive to the hot spare drive.

Dynamic configuration

Capacity expansion of Pool and Volume Group is possible without stopping operation
For Pool, up to 60, and for Volume Group up 2, drives can be added at a time. By expanding Pool and Volume Group at a later time, operation management can be flexible.

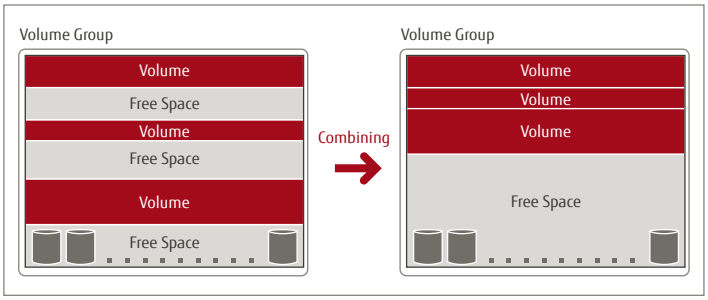


Capacity expansion of Volume is possible without stopping operation
The volume capacity can be expanded by using free space of Pool and Volume Group.



Combining free spaces in Volume Group without stopping operation

If creation and deletion of Volume are repeated in Volume Group, free spaces are divided into multiples, and the Volume of which the size is equal to the total free space may become not possible. In such case, multiple free spaces can be combined which have been created Volume Group. Even during combining free spaces, data in Volume can be accessed.

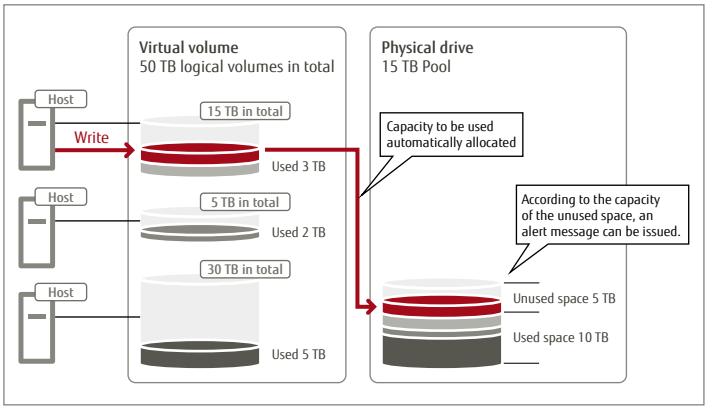


Reduction of initial investment, and operation cost with Thin Provisioning

Thin provisioning is a technology to virtualize the storage capacity, which enables the effective use of physical drives. When a volume is created, a virtual volume with the requested capacity is assigned to the host. At that time, the physical storage for it is not assigned yet. The physical storage is allocated when the data is written to the volume, according to the amount of the data.

In the figure, the 50 TB virtual volume is assigned according to the host request. On the disk array side, the 15 TB pool is assigned for operation. When a write is performed from the host, a space with the used capacity is automatically assigned from the unused space, and the data is written into it. By using the thin provisioning technology, logical volumes exceeding the physical drive capacity can be assigned to the host

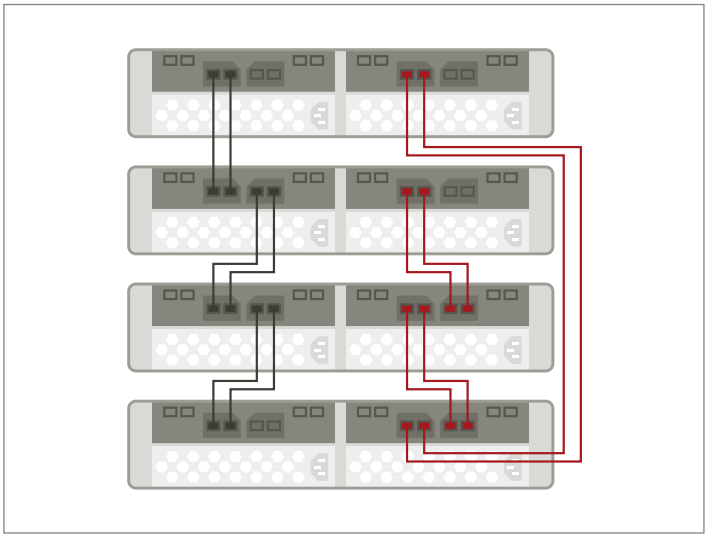
- Reduction of initial investment (small start)
- Reduction of unused installed drives allows reduction of operation cost



Availability Management

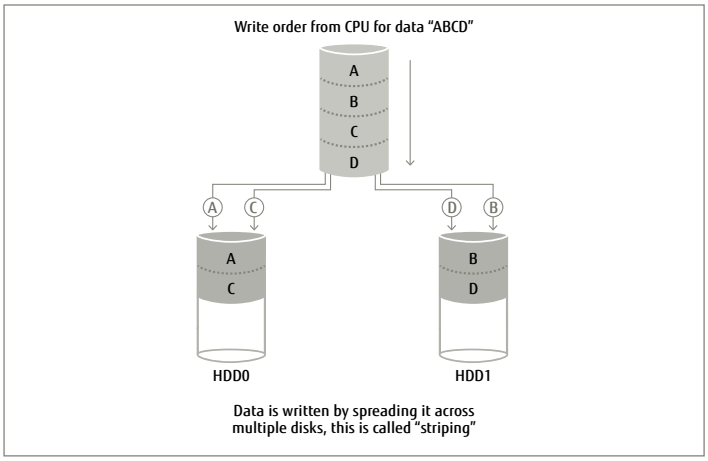
Reverse Cabling

In order to provide high availability, the controller enclosure is connected to the drive enclosure via reverse cabling. The connection of two paths is implemented in ascending order while the other two paths are connected in descending order, as shown in the figure. Because BE connections are doubled per CM it is called Quad Path Cabling. In the event of a drive shelf failure, only broken one is disconnected. All other drive enclosures remain accessible.

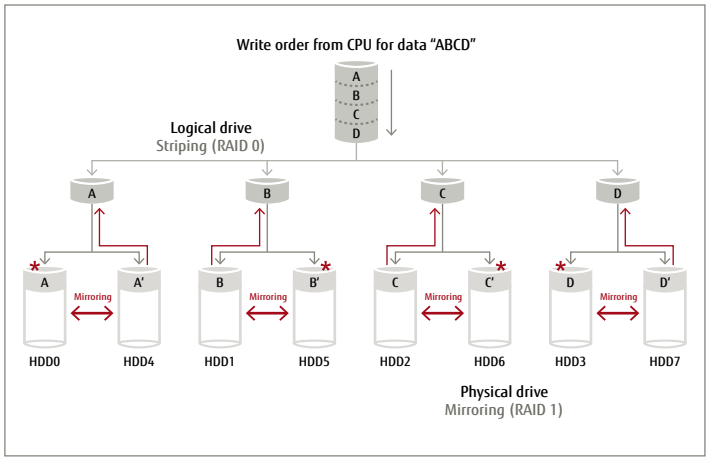


RAID: Improving Performance and Preventing Data Loss

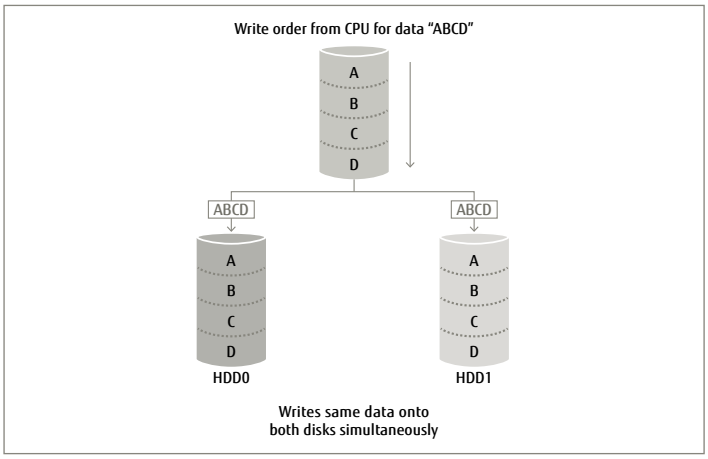
Damage to a company caused by disk failure is a steadily growing risk, as data volumes and disk capacities increase. Storage system downtime can result in companies failing to take full advantage of business opportunities, due to the management overheads involved in securing important data. RAID technology not only prevents such data loss, but also enhances business performance. RAID is the use of multiple disks to manage the data using a range of different techniques. These are divided into various levels. They all differ in terms of data deployment and the type of redundancy offered. It has also become popular to mix and match the various RAID level technologies in order to provide more specific cost reductions and performance enhancements. This document only concentrates on the most important RAID levels supported by ETERNUS AB/HB series.



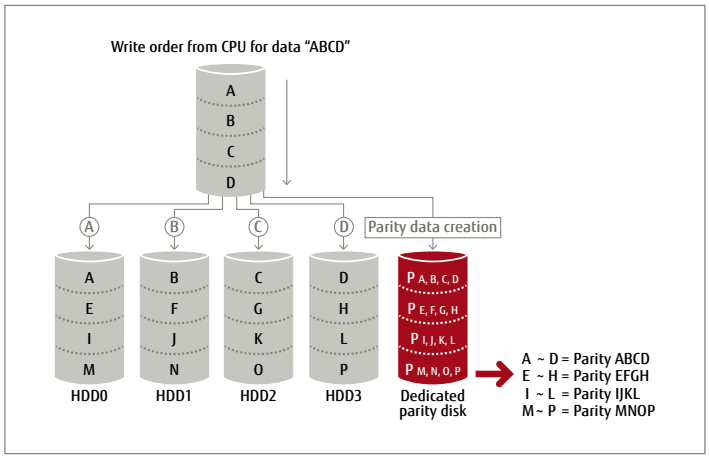
RAID 0



RAID 1+0



RAID 1



RAID 3

RAID 0

RAID 0 divides data into block units and writes them in a dispersed manner across multiple disks. As data is striped across every disk, this technique is also called "striping". This process enables high performance, as parallel access to the data on different disks improves the speed of retrieval. However, no recovery feature is provided if a disk failure occurs. If one disk fails, it affects both reads and writes. And as more disks are added to the array, the chance of a disk failure occurrence is higher.

RAID 1

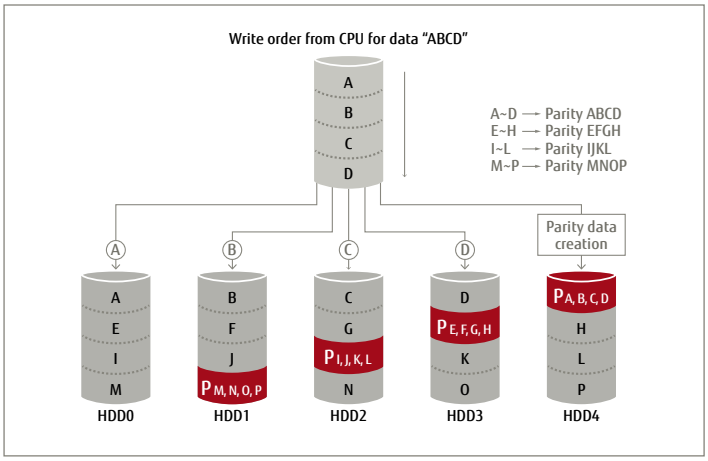
This level is called "mirroring" as it writes the same data to two disk drives simultaneously. Although there is no enhancement in access speeds, the automatic duplication of the data means there is less likelihood of data loss. RAID 1 provides failure tolerance. If one disk fails, the other automatically takes over and continuous operation is maintained. There could be some impact on the overall cost of the disk array because duplicating all the data within the array means that only half of the raw storage capacity is available for the hosts.

RAID 1+0

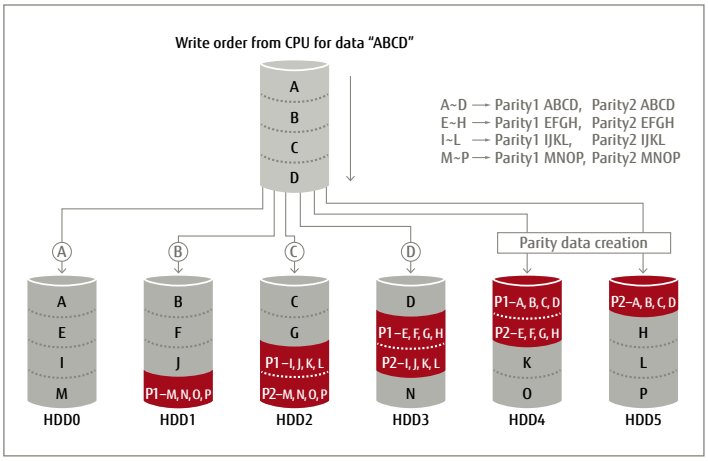
RAID 1+0 combines the benefits of RAID 0 and RAID 1. By configuring both technologies in a single array, both data duplication and improved access speed can be provided. Although this combination makes installation more expensive compared to other technologies. However both reliability and high I/O performance can be guaranteed. RAID 1+0 on Fujitsu ETERNUS Storage arrays also provides extra protection in those cases where a single drive failure can result in disruption of data access to users.

RAID 3

With RAID 3, data is divided into bit or byte units and written across multiple dedicated data disk drives. Parity information is created for each of separate data section and written to a dedicated parity drive. All disk drives can be accessed in parallel all the time and the data can be transferred in bulk, ensuring high-speed data transfer.



RAID 5



RAID 6

RAID 5

RAID 5 is the most commonly used RAID technology today. It is based on a technique that avoids concentration of I/O on a dedicated parity disk as with RAID 4. RAID 5 divides the data and creates parity information, but the parity data is written separately across multiple disks. It enables multiple write orders to be implemented concurrently because updated parity data is dispersed across the multiple disks. This feature ensures higher performance compared to RAID 4.

RAID 6

RAID 6 deploys two parity records to different disk drives (double parity), enabling two simultaneous disk drive failures in the same RAID group to be recovered. It is thus able to execute multiple write orders at the same time. This feature ensures higher performance. Especially for Nearline SAS, high capacity drives, RAID 6 (with double parity) should always be considered as first choice.

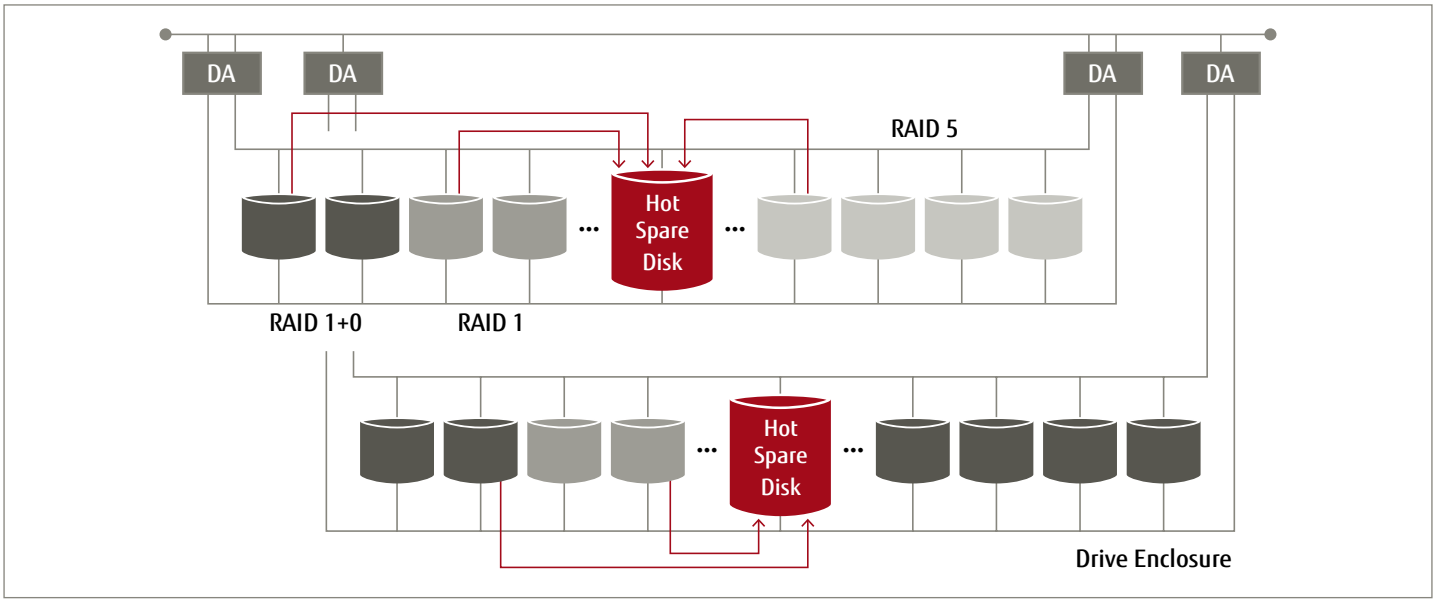
Proactive Drive Monitor / Data Evacuator

ETERNUS AB/HB series is constantly checking for a sign of a drive failure, and if it occurs, system is saving data in advance to ensure data continuity.

When a drive is nonresponsive storage system powers it OFF and ON to check if status had changed. If it is unclear, drive is marked as faulty. If a failure is expected to happen, data evacuator feature moves the data from the affected drive before a failure occurs. If a drive has failed, rebuild is started.

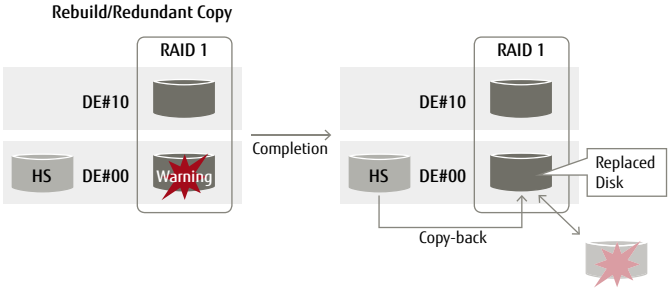
Global Hot Spare

Hot spare disks are preparatory disk drives that are kept on active standby for use when a disk drive fail. This global hot spare function enables hot spare disks to be used for any volume group. When a disk drive in a RAID group fails, data on the disk drive is automatically reconstructed on the hot spare disk in background.



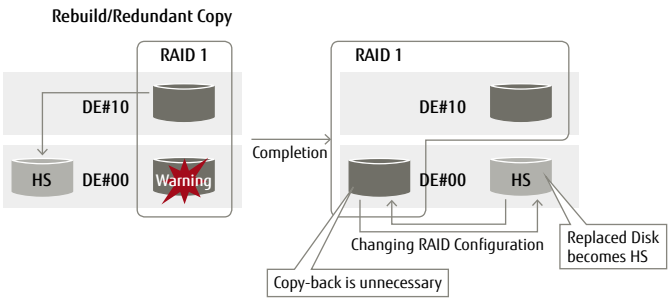
Copy-Back and Copy-Back-Less Operation

After the faulty disk has been replaced with a new disk, data on the hot spare disk is retrieved (copied back) to the new disk. The copy-back function restores the RAID group while maintaining redundancy after the rebuild has been performed for the hot spare in the RAID group.



Copy-back

Copy-back-less function is a feature which builds hot spare disks into a RAID configuration after the completion of rebuild or redundant copy and the internal RAID configuration of the failed disk is changed to hot spare disk. Immediately after maintenance and replacement, it can start working as a hot spare disk. This feature means that the copy-back process is no longer required.



Copy-back-less

Benefits:

- Disk replacement time is drastically reduced
- RAID availability is improved as the hot spare downtime is reduced
- I/O performance degradation during the copy-back process can be avoided

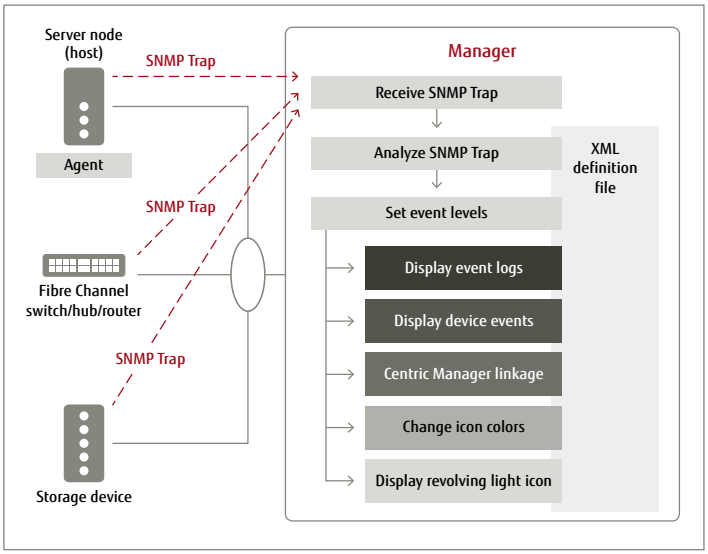
Event Management

Event notifications can be forwarded via e-mail. When Remote Support is configured, event notifications are also automatically sent to Fujitsu customer service.

Depending on the capabilities of the devices the event management function itself can be executed manually, via notification of the ETERNUS AB/HB or via external polling devices:

The event monitoring processes SNMP traps, decodes them, and displays them as an asynchronous event. The manager thus supports smooth operations because the displayed contents of the event are more detailed and easier to understand than decodes generated by a normal SNMP MIB compiler.

The customized content and the display format (show, hide) for events can be adapted in detail. Therefore, operation can be flexibly customized for specific operational environment requirements. Integration with various other management software products is also possible. Event monitoring using the device polling function regularly monitors the status of all devices connected via LAN by using SNMP, ping or unique protocol and can also be customized using definition files.



Information Security Management

Data Confidentiality

Due to various data protection laws, enterprise information and the security involved has become much more important from a corporate social responsibility standpoint. Laws and internal guidelines require that access to relevant stored data is restricted only to authorized users and that sensitive information is protected against unauthorized or accidental access. ETERNUS AB/HB storage systems provide data encryption functions to address such requirements.

Role-Based Access Control (RBAC)

Safety management of device by limiting user access rights

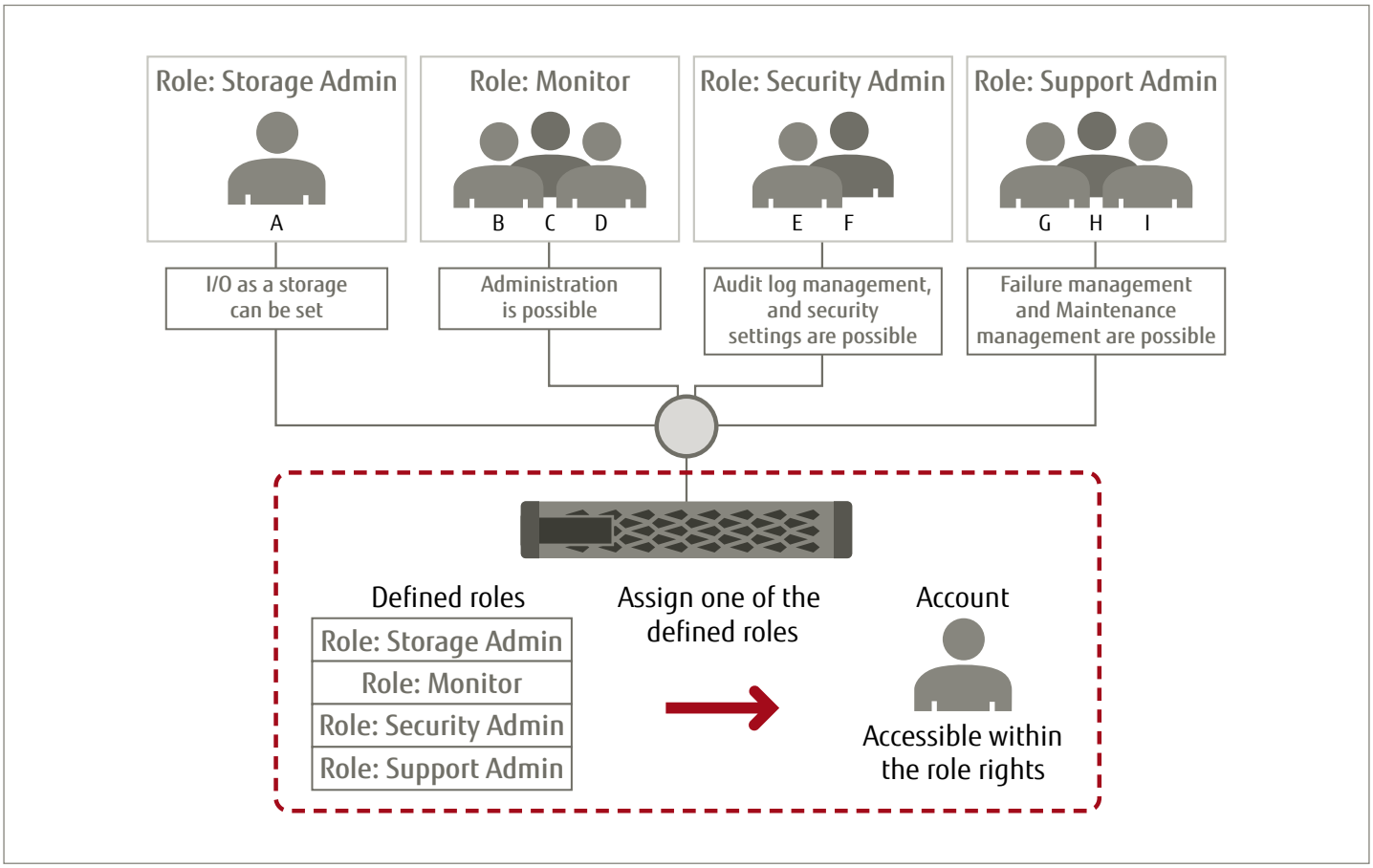
The RBAC function is a general function to "define rights as a role, and assign a role to an account".

User access restriction method by role:

- A role is a definition of rights
- By assigning a new account to a role, the rights can be easily set

Available roles:

- **Root Admin** – The special management right which enables access to all functions in the system. The manager password is set at the initial login, or later.
- **Storage Admin** – the full access right of read/write against the storage object (volumes and disk pools) is given. Access to the security setting is not possible.
- **Security Admin** – The access right to access management, certificate management, and audit log management is given. Enable/disable of the existing management interface (SYMbol) is also possible.
- **Support Admin** – The access right to all hardware resources of the storage array, failure data, MEL events, and controller firmware update is given. Access to the storage objects, and security settings are not possible.
- **Monitor** – The read-only access right to all storage objects is given. Access to the security setting is not possible.



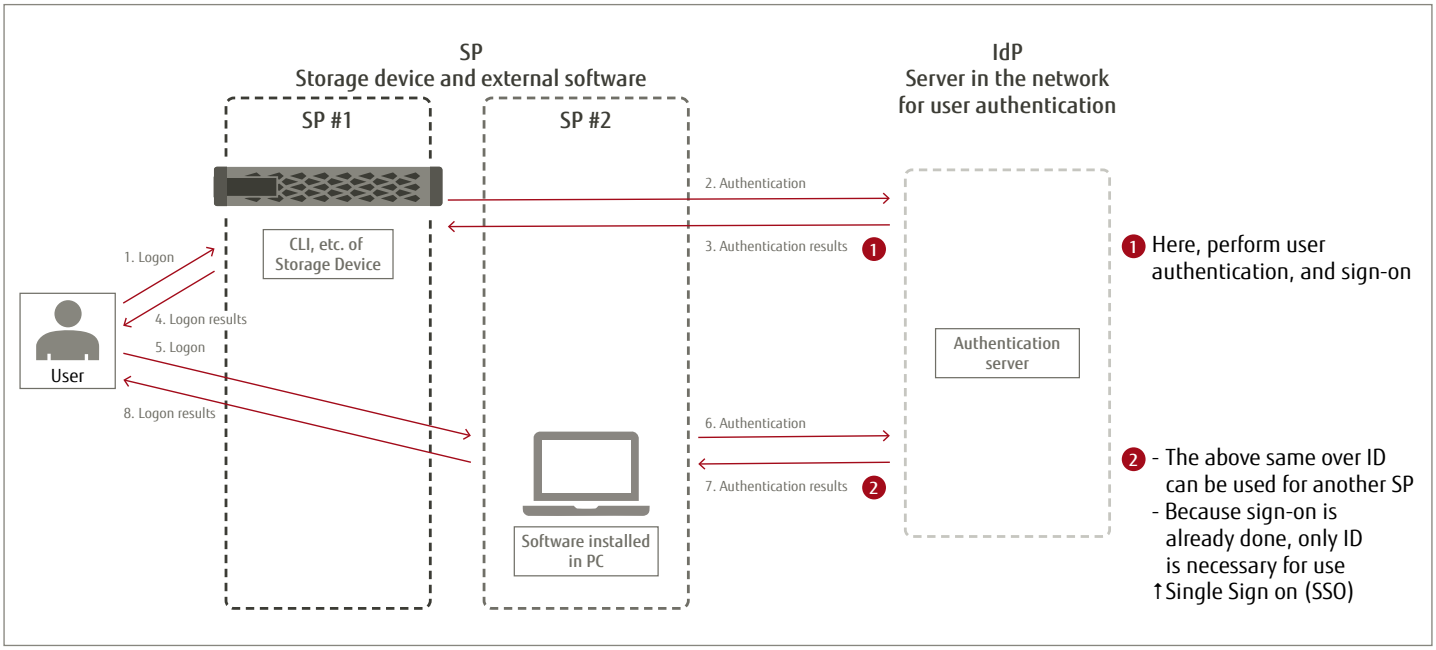
Security Assertion Markup Language (SAML)

Improved user-friendliness and security of user authentication with SAML

SAML is a standard of authentication information for user authentication between different internet domains. A protocol of security information such as user authentication information, attributes and approval of user rights described with the XML grammar. The service and user authentication is shared by Service Provider (SP) and ID Provider (IdP).

To use SAML, the Single Sign-on (SSO) function is used. SSO has the following two interpretations in general:

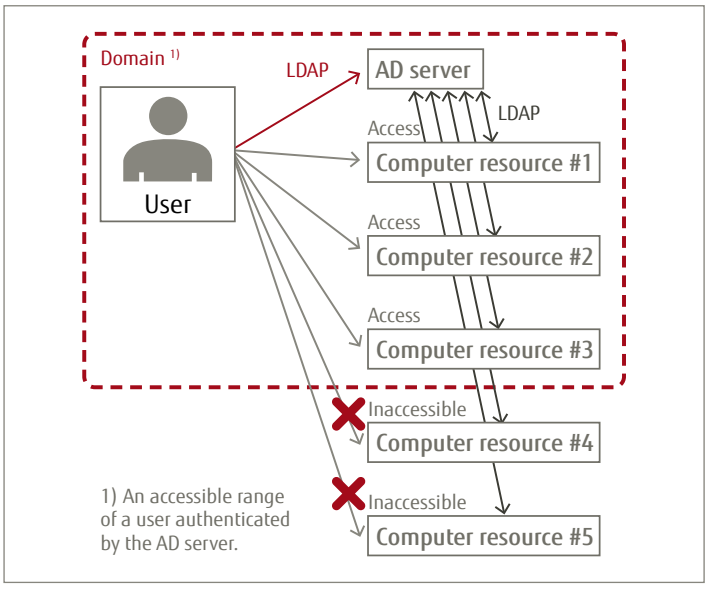
- A mechanism to arrow multiple services at one location.
- A mechanism to arrow multiple services in one authentication.



Active Directory (AD) and Lightweight Directory Access Protocol (LDAP)

Safety management by clarifying access range of user account

- AD: A generic name for the component group to manage users and computer resources
- LDAP: A protocol to connect to the directory service to manage users and computer resources



Encrypted Drives

Protection of stored data with encrypted drives

Combining encrypted drives and key management, data can be effectively protected from unauthorized access or modification due to theft, loss or diversion of disk drives.

Adopting encrypted drives ensures security without affecting performance. There are two types of encrypted drives:

1) FED (Full Disk Encryption)
- Full drive encryption at hardware level with AES-256 encryption standard.
- Hard drives contain ASIC chips that encrypt data when writing and decrypt data when reading.

2) FIPS (Federal Information Processing Standards)
- FIPS is a set of security requirements for cryptographic modules defined by the Information Technology Laboratory at the National Institute of Standards and Technology (NIST)
- FIPS 140-2 provides a set of rules that should be followed when a cryptographic module (here, a self-encrypting disk [SED]) is executed in FIPS compliant mode
- FIPS 140-2 defines security requirements in several levels. SEDs configured to run in FIPS compliant mode comply with the requirements listed in Level 2

Key Management Interoperability Protocol (KMIP)

Centralized and easy authentication key management

KMIP is a communication protocol that defines the communication method of the "client" and the "key management server" that manages encryption keys for data encryption.

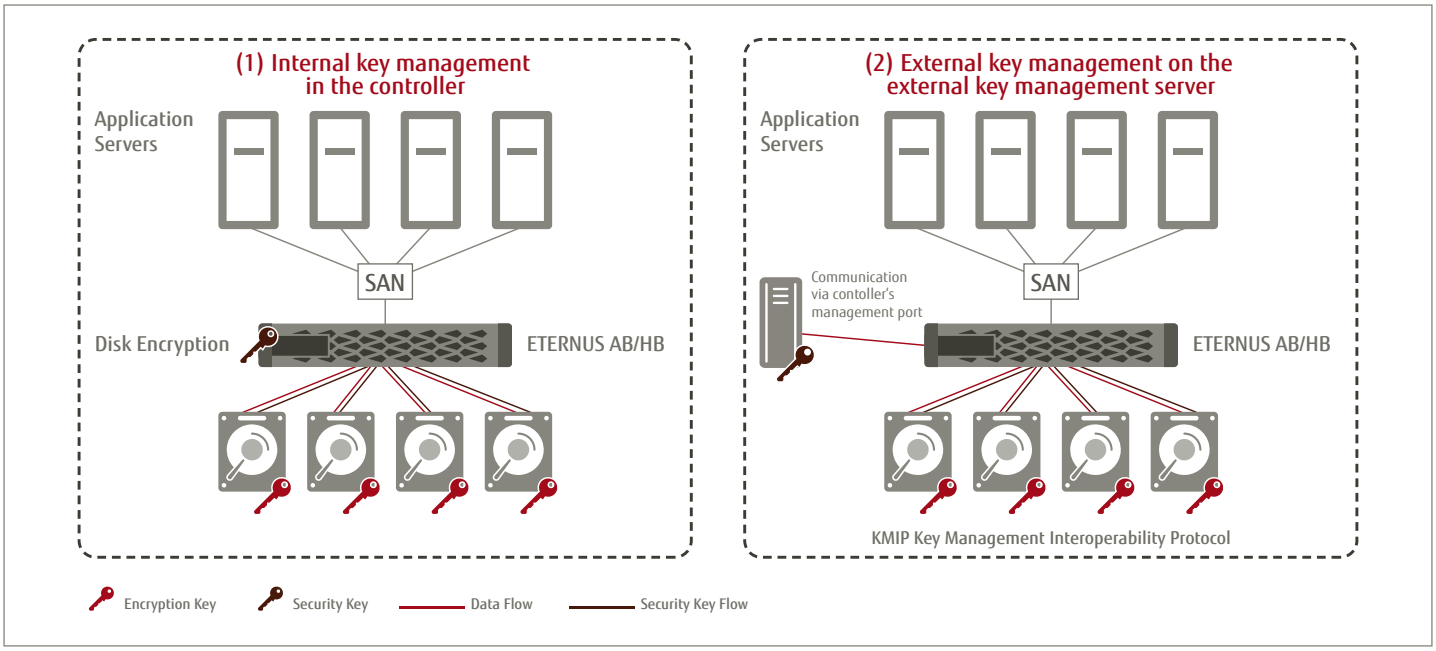
A template for encryption key generation and usage is defined, and it is a technology that enables easy management of multiple keys.

Combining local key management with drive level encryption provides comprehensive protection for stored data without impairing performance or ease of use.

Even if all drives are carried away from a data center due to reintroduction, removal, maintenance, etc., there is no risk of leaking confidential data.

The ETERNUS AB/HB series manages security keys using internal or external key management.

- With the internal key management, security keys are managed on Array
- With the external key management, security keys are managed on the external KMIP server

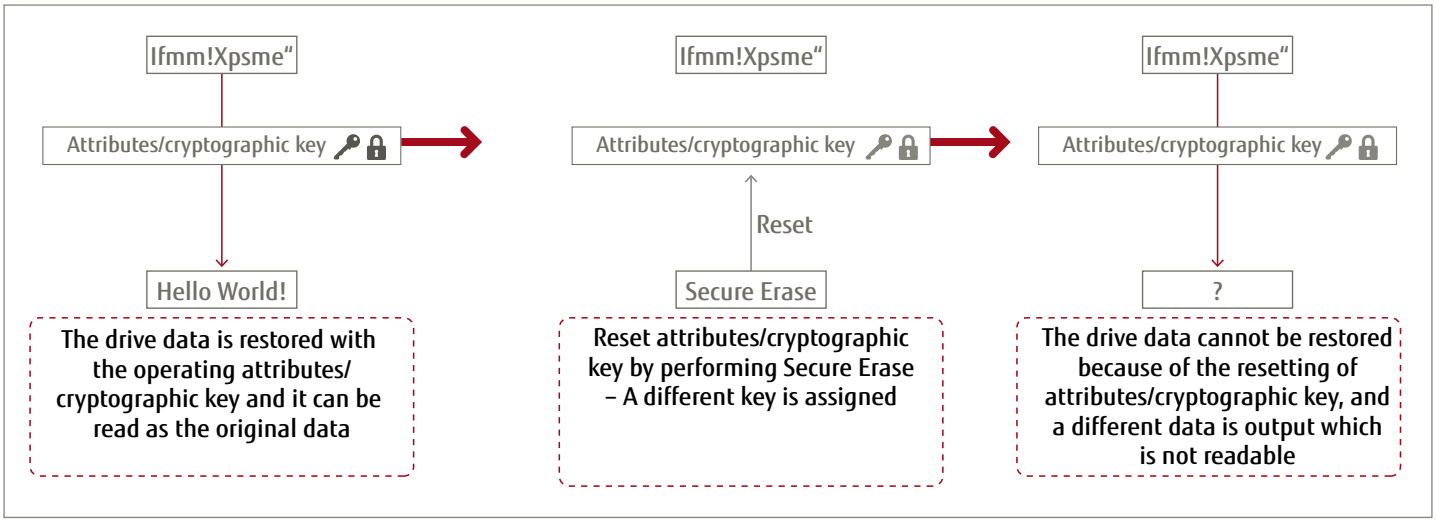


Data protection

Drive Data Erase Function: Secure Erase

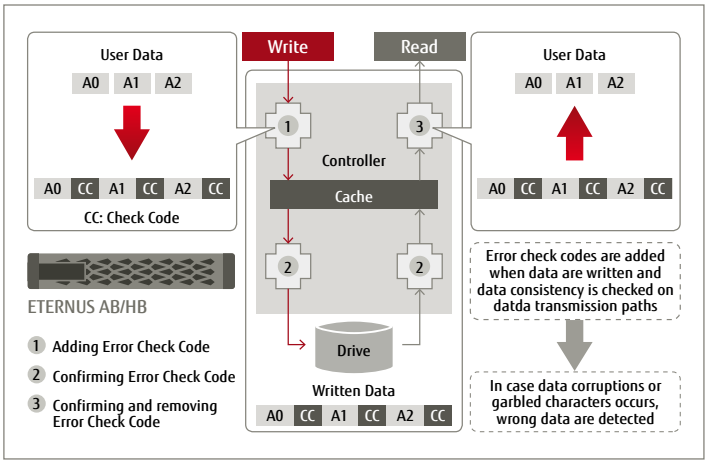
An active defense measure against data leakage caused by unlocking discarded drives

Only security ready drives (FIPS/FDE) can be processed. By resetting the drive attributes, the data becomes unrecoverable. Here, at least, cryptographic key is included in the attributes.



Data Protection with T10 Protected Information (PI) or DIF (Data Integrity Field)

If the function is enabled, the storage system adds Error Check Code (CRC) to each data block in the volume. After the data block is transferred, possible errors during transfer are checked by using the CRC code. Potentially corrupted data will not be written into the drive, nor returned to the host.



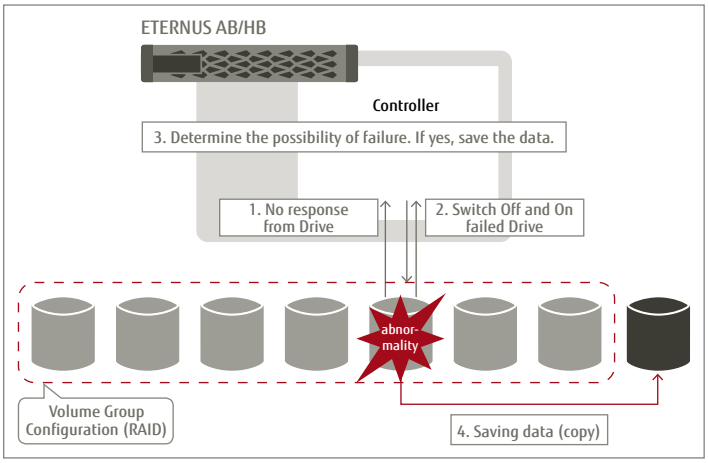
Data Block Guard

Proactive Drive Monitor / Data Evacuator

For a drive with no response, automatically power OFF and ON to check the Fault status. If it is unclear, it is regarded as Fault.

If a failure is suspected, the evacuator feature moves the data from the affected drives before a failure occurs. If a drive has failed, Rebuild is called to reduce the time until Rebuild ends.

The data move is equivalent to the activity of Redundant Copy which is explained in a different chapter.



Features and management

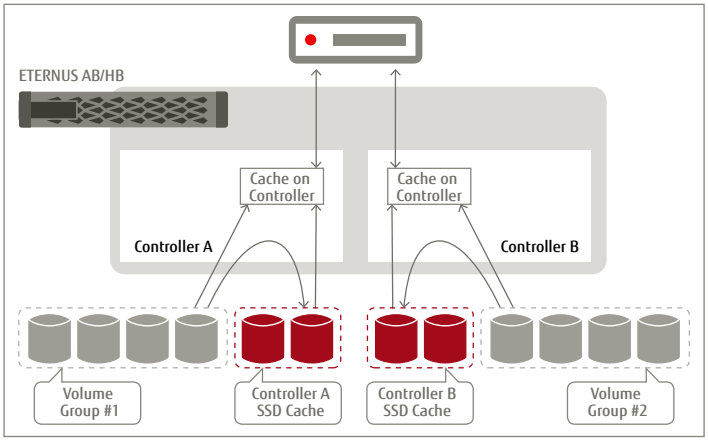
SSD Cache

For workloads that requires large amounts of Read processing SSD Cache allows us to reach desired performance results. SSD Cache use SAS-3 SSD disks installed in front-side drive slots of the array. This feature can be used on both – ETERNUS AB and HB arrays.

By copying frequently accessed data (hot data) from an HDD volumes to an SSD cache it allows us to realize efficiency of hard disk operation, reduced latency and fast reading speed. There is no need to create schedules for data movement between tiers, because data is permanently monitored by system and moved to SSD cache if it is required. SSD cache capacity can be expanded up to 5 TB and is shared between both CMs – one RAID 0 volume for each controller. Second cache improves only Read operations – data Writes are handled by primary cache of the storage system.

This feature is best suited and most beneficial for environments where random Read workload exceeds 85%. If workload type is mostly sequential or Write ratio exceeds 15% then use of SSD cache may not be cost-effective.

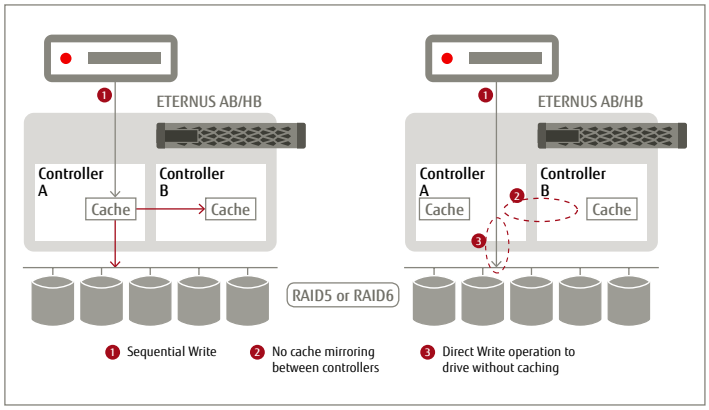
SSD Cache can also be used for volumes created on SED drives but then drive security function of the volume and the SSD cache needs to be the same.



Full Stripe Write Acceleration

Full Stripe Write Acceleration (FSWA) is a function to improve the throughput of the host sequential Writes when a specific conditions are satisfied. With this feature enabled data written to storage device goes directly to disk drives skipping cache write which requires replication to second controller cache. This is done to avoid replication which extends completion time of write process to the array.

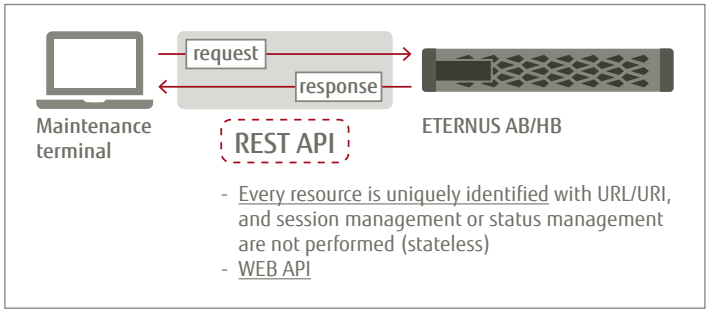
Note that in the basic operation of ETERNUS AB/HB, caching is performed when the drive is HDD. When the drive is SSD, the data is directly written into the drive without caching.



RESTful API

The SANtricity REST API is an application programming interface which doesn't require user interaction after the set of commands has been issued. This interface is dedicated for storage management, process automation and monitoring purposes. REST API documents can be accessed with the help function of SANtricity System Manager.

Thanks to client libraries available for Python and Java SDK it is possible to customize API to customer environment.

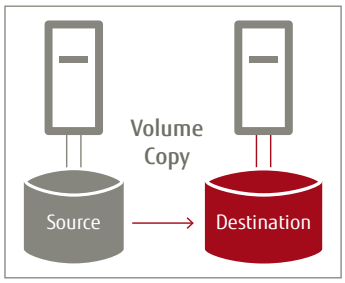


Copy functions

While data is growing exponentially, its importance is also growing from regulatory compliance and business continuity standpoints. Thus efficient easy replication and backup of large data volumes is becoming a critical issue.

The embedded copy functions of ETERNUS AB/HB supports the copying of data from a business volume to another volume, both within the same ETERNUS AB/HB storage system and across multiple ETERNUS AB/HB storage systems in the storage network. There is also a possibility of creating volume snapshot which is especially useful for backup systems. Complete set of copy functions is described in detail in the next few sections.

Volume copy

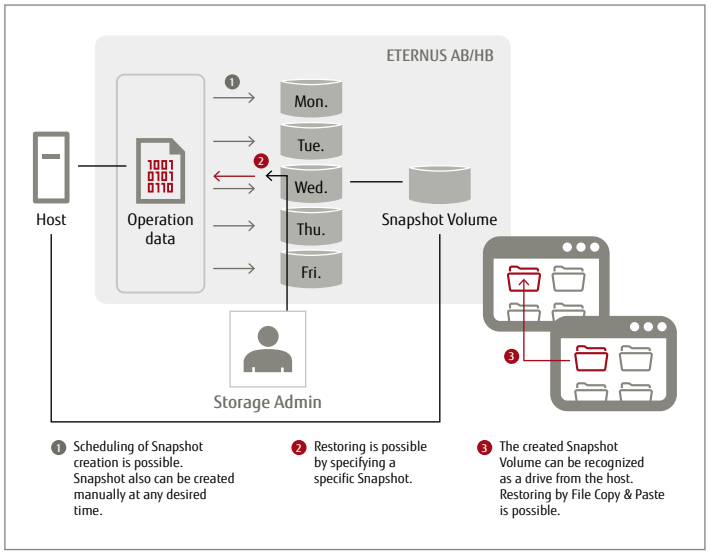


Volume copy is a complete and independent physical copy of the source volume. This operation creates online or offline clone of the prime volume within the same ETERNUS AB/HB system. Destination volume must be at least the same size as source but it can be protected with different RAID level underneath. This protection function is the right choice for creating a replica or a backup of the data as it is a complete copy.

Snapshot

A snapshot function provides instant and capacity efficient solution to create the backup image of the source volume. Technology used in this process is copy-on-write, which use very little space to keep the snapshot. Volume restore can be performed by specifying a particular Snapshot Image. If there is a need to retrieve deleted or changed files, then Snapshot Volume needs to be created from Snapshot Image and mapped to the host. Snapshot copies of multiple volumes created at the same time can be collected in Consistency Group.

It is possible to set up snapshot schedules and warning/alert notifications for data update space, depending on used capacity.

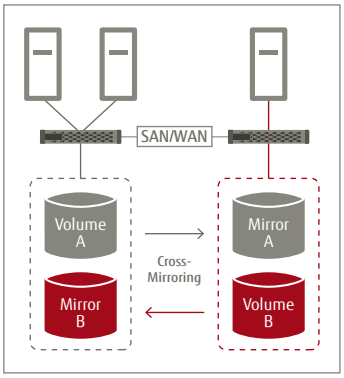


Synchronous Mirroring

Consistent and mirrored data in both company locations can be assured with Synchronous Mirroring. This feature creates full copy of the source data in remote location and keeps mirroring session active all the time so the data is validated with each write request on both sides. Synchronous mirroring require low RTT (Round Trip Time) between sites to keep satisfying performance.

Asynchronous Mirroring

Data volumes can be mirrored asynchronously between two ETERNUS AB/HB systems to ensure data availability in case of any disaster in primary datacenter. It is a great solution for datacenters connected via links with slow bandwidth because data is sent periodically – it guarantees that bandwidth will not be saturated by mirroring data.



Remote Mirroring Concept

SANtricity Management

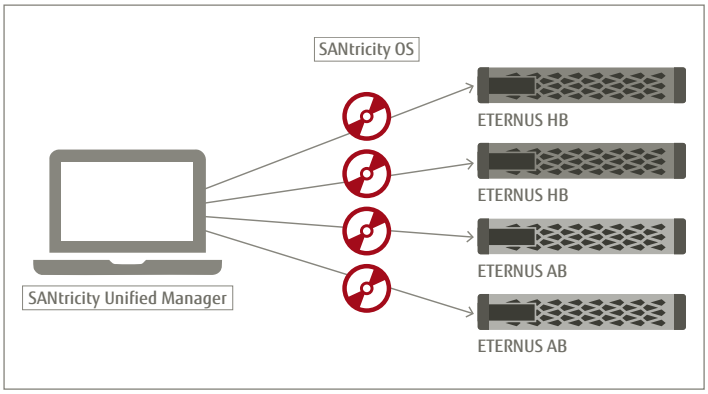
SANtricity System Manager

SANtricity System Manager is the Web-based management GUI embedded in ETERNUS AB/HB with following features:

- Simple and intuitive Web GUI
- The overview of storage system data such as performance, capacity and storage layer is provided in a single view with a dashboard
- Easy access to common storage operations with tab menu
- Real-time graphic display of the storage performance information such as IOPS, throughput and CPU utilization rate
- Visualization of storage capacity and storage configuration
- Intuitive maintenance and management operations are possible by integration of hardware location display with relevant hardware operations

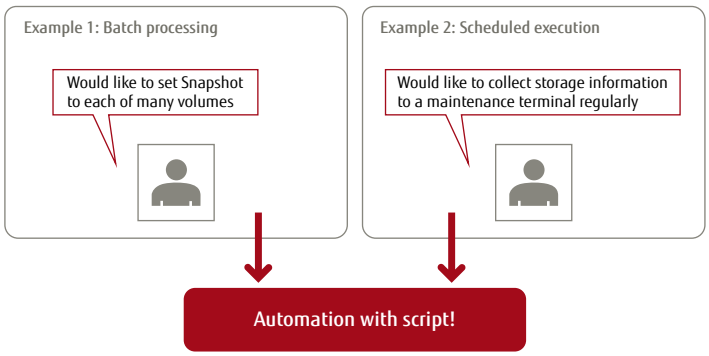
SANtricity Unified Manager

Fujitsu offers unified management software to monitor and manage whole ETERNUS AB/HB infrastructure and it is called SANtricity Unified Manger. It is possible to do maintenance tasks, such as upgrade, OS installation or log collection, across all available storage systems in Unified Manager.



SMcli

SMcli is a storage management tool installed in a terminal in Windows and Linux. Main purpose of this tool is to automate and script most common tasks executed on ETERNUS AB/HB storage systems. All commands are executed from command line level and their results are equivalent to GUI available operations.



Active IQ ConfigAdvisor

Software for verifying configurations and checking the health of ETERNUS AB/HB storage systems. ConfigAdvisor is a tool installed on client device, accessible via web interface and used during trouble-shooting and maintenance procedures. According to the collected information such as hardware configuration or health check, various analysis can be displayed, also in form of report i.e. generated in pdf file format.

Active IQ OneCollect

Active IQ OneCollect so as ConfigAdvisor is an application installed on client device so there is no need to provision server resources to run it. OneCollect is an integrated information collection tool for storage environments including not only ETERNUS AB/HB systems but also servers and switches.

SANtricity software

The latest version of SANtricity software (v11.7x) offers significant feature enhancements:

1. High Availability

- I/O Path Protection: Availability and load balancing are maintained even if a path fails.
- Mirroring Volumes that use consistency groups are mirrored synchronously/asynchronously.

2. Data Protection

- Dynamic Disk Pool: To enhance data protection, rebalancing is dynamically performed when the number of drives changes.
- Various RAID Levels: Multiple RAID levels are used flexibly to optimize workloads. The level can be changed without stopping the system.
- Data Assurance: Complies with the T10-PI standard.
- Proactive Drive Health Monitor: Detects drive problems before a failure occurs.
- Background Media Scan: Proactively checks for media failure and fixes problems in advance.
- Snapshot: Creates point-in-time images of data and related data volume sets.
- Volume Copy: Creates a complete physical copy (clone) of the volume in the storage system.

3. Diagnosis

- Recovery Guru: Assists with troubleshooting and suggests the appropriate procedures.

4.Performance

- SSD Cache used as the primary data cache and automatically provides frequently accessed data.

Coordination with Host

VMware Linkage

With VMware vSphere Client, displays integrated information of infrastructure in virtual environment

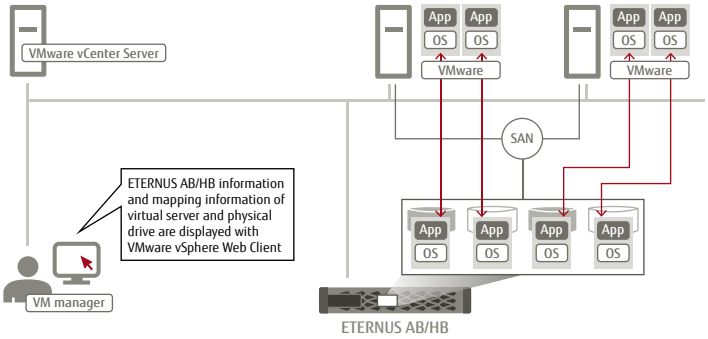
VMware vSphere Web Client is a browser base management tool for the vSphere environment, which enables operation and monitoring of the vSphere infrastructure.

Server Manager can perform various operations such as storage volume settings and backup from the VMware management screen.

ETERNUS AB/HB provides the embedded volume tag function for the administrator to be able to organize volumes in the array by the workload type.

By using the volume creation wizard, the recommended configuration associated with Vmware or the volume segment size setting can be done.

With this setting, the time needed for the provisioning of volumes by Vmware vSphere can be shortened.



vStorage API for Array Integration (VAAI)

ETERNUS AB/HB performs the VMware server processing to reduce the server side load

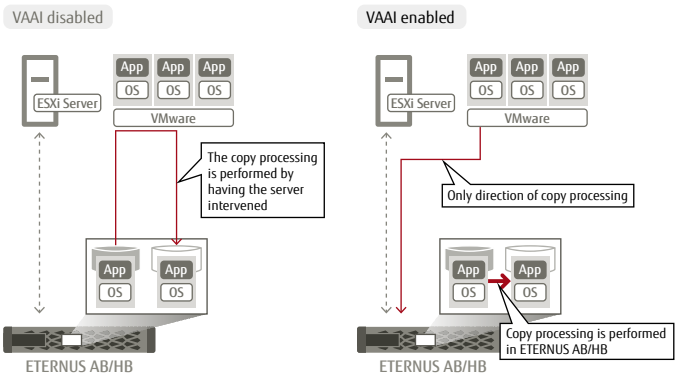
The VMware vSphere software realizes the virtualization of a platform. The "VAAI" API is provided to improve the system performance and expandability, utilizing the storage device resources of VMware vSphere.

ETERNUS AB/HB Series cooperates with VAAI to integrate various functions, improving the operational efficiency in the large server virtualization environment

Copy Offload is a function to perform the copy processing by not having the server intervened. Instead, it performs the copy processing in the storage device. Conventionally the copy processing such as duplication of a virtual machine and migration needs to be performed by having the server intervened, which requires quite a few resource consumption of the server.

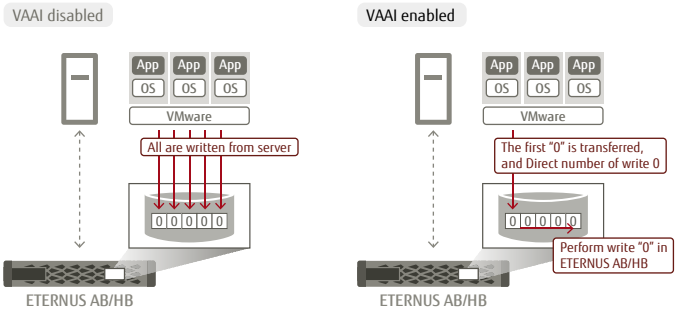
By utilizing this function, the copy processing can be off-loaded to the storage device side, reducing the server load, and improving the system performance.

Copy Offload: Offload copying of virtual machine, and copy processing at migration to storage

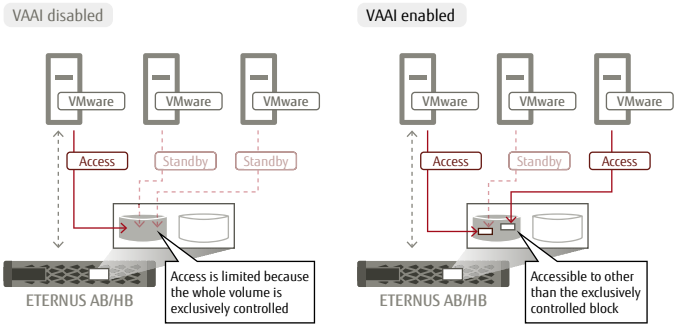


When storage spaces are assigned to create a new virtual machine, etc., the storage spaces need to be initially filled with zero data. Write Same is a function to perform this processing on the storage device side. By performing the processing on the storage device side instead of in the server, the server load is reduced, and the capacity assignment of a virtual machine is accelerated. For the exclusive control of a specific storage area, ATS (Atomic Test & Set) realizes the exclusive control by using the block unit control function of the storage device, instead of the control in the unit of LUN (logical volume) implemented in "VMware vSphere". With this function, the storage space of which the access is limited in the exclusive control can be minimized, improving the operational efficiency of a virtual machine.

Write Same – zero disk blocks: Zero write processing of storage space allocation for new virtual machine



ATS (Atomic Test & Set): Exclusive control in the unit of logical volume by the connection server is minimized to the unit of block

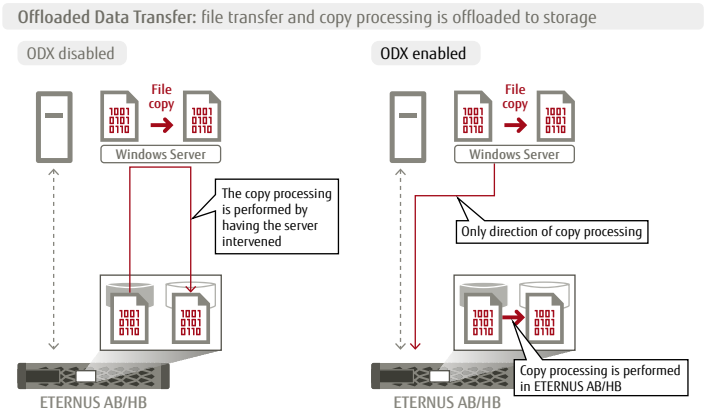


Windows Linkage with Offloaded Data Transfer (ODX)

A file is copied or moved with server read/write process up to now. The server CPU load would be high in a process.

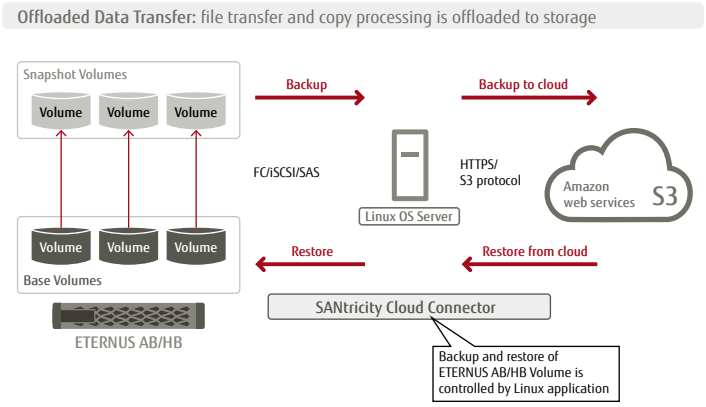
ETERNUS AB/HB series supports "Offloaded Data Transfer (ODX)" function which is a Windows Server function.

This function makes the server CPU load accompanied with file copying or migration offload to a storage device, and the server CPU load can be reduced.



Cloud Linkage

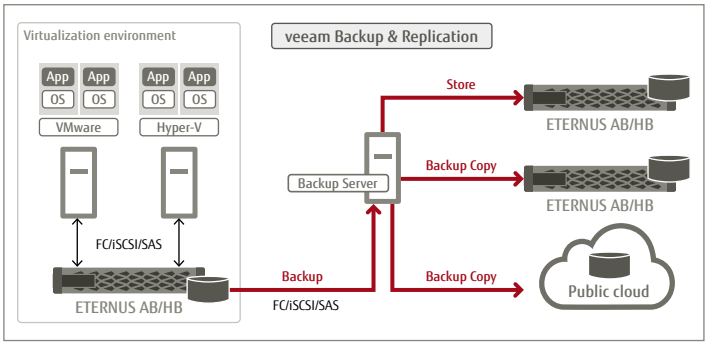
Cloud Connector is a software to backup data in ETERNUS AB/HB series to Cloud, and restore by using the backup data. By directly connecting the system to AWS S3 and performing backup, a high cost performance backup is realized.



Veeam Linkage

For the virtual environment of VMware vSphere and Microsoft Hyper-V, Veeam Backup & Replication integrates backup and replication of the advanced VM machine into a single solution. It is software to realize virtualized applications, and a high speed and reliable recovery of data.

By using Veeam Backup & Replication and ETERNUS AB/HB series, virtualized applications, and a high speed and reliable recovery of data become possible, and business continuation and reduced risk are realized.



Summary and Conclusion

The FUJITSU Storage ETERNUS AB/HB series is designed for small and medium sized enterprises and businesses that require a dedicated solution, and provides an industry-leading, simple and effective platform. The FUJITSU Storage ETERNUS AB/HB series has a high affinity for VMware and is the best choice for accelerating business operations.

With the FUJITSU Storage ETERNUS AB/HB series, Fujitsu has created a SAN array that is future-ready, usable today; and easily implemented within your current operational processes and procedures.

The FUJITSU Storage ETERNUS AB/HB series delivers high performance, consistent low latency and advanced HA features which are easy to provision with the embedded System Manager or REST API.

The FUJITSU Storage ETERNUS AB/HB series also has robust built-in monitoring capabilities that enable you to troubleshoot performance issues at the logical level, physical level, and application level, as shown in the Performance Monitor.

In addition to solving the extreme latency requirements, the ETERNUS AB/HB series also resolves Oracle database challenges in the following ways:

- It dramatically boosts the performance of existing applications and lowers the cost per IOPS ratio without requiring that you rearchitect the application
- It increases Oracle performance with RAID10 and DDP
- Better response times increase user productivity, improving business efficiency
- Oracle ASM provides an alternative solution for high availability and disaster recovery

Technical specifications ETERNUS AB and ETERNUS HB

ETERNUS AB ALL-FLASH STORAGE	AB2100	AB3100	AB5100	AB6100
				
Number of controllers	2	2	2	2
Max. system memory	64 GB	32 GB	128 GB	256 GB
Max. storage capacity	1,468 TB	367 TB	1,836 TB	367 TB
Max. No. of SSDs installable	96	24	120	24
Max. No. of host interfaces	12 Port [FC (4, 8, 16Gbit/s)] 8 Port [FC (32Gbit/s)] 12 Port [iSCSI (10Gbit/s)] 8 Port [iSCSI (10GBase-T)] 8 Port [iSCSI (25Gbit/s)] 8 Port [SAS (12Gbit/s)]	8 Port [FC (8, 16, 32Gbit/s) 8 Port [NVMe over FC (32Gbit/s)] 4 Port [NVMe over InfiniBand (100Gbit/s)] 4 Port [NVMe over RoCE (100Gbit/s)] 8 Port [iSCSI (25, 10Gbit/s, SFP+)] 4 Port [IB iSER (100Gbit/s)] 4 Port [IB SRP (100Gbit/s)]	12 Port [FC (4, 8, 16Gbit/s)] 8 Port [FC (32Gbit/s)] 8 Port [NVMe over FC (32Gbit/s)] 12 Port [iSCSI (10Gbit/s)] 8 Port [iSCSI (25Gbit/s)] 4 Port [IB iSER (100Gbit/s)] 4 Port [IB SRP (100Gbit/s)] 8 Port [NVMe over InfiniBand (100Gbit/s)] 4 Port [NVMe over Ethernet (100Gbit/s)] 8 Port [SAS (12Gbit/s)]	16 Port [FC (8,16,32Gbit/s)] 16 Port [NVMe over FC (32Gbit/s)] 8 Port [NVMe over InfiniBand (100Gbit/s)] 8 Port [NVMe over RoCE (100Gbit/s)] 16 Port [iSCSI (10, 25Gbit/s)]
Drive interface	SAS 12 Gbps	NVMe	SAS 12 Gbps	NVMe
Supported RAID	0, 1, 1+0, 3, 5, 6, DDP (Dynamic Disk Pools)			
No. of drive shelves (DE) connected	3	0	4	0

For information on the management software, please see the data sheet.

ETERNUS HB HYBRID STORAGE	HB1100/HB1200	HB2100/HB2200/HB2300	HB5100/HB5200
			
Number of controllers	2		
Max. system memory	16 GB	64 GB	128 GB
Max. storage capacity	HB1100: 432 TB HB1200: 194 TB	HB2100: 3,456 TB HB2200: 2,894.4 TB HB2300: 3,456 TB	HB5100: 7,927.2 TB HB5200: 8,640 TB
Max. No. of HDDs installable	48 (HB1200), 24 (HB1100)	192 (HB2300), 168 (HB2200), 192 (HB2100)	480 (HB5200) 444 (HB5100)
Max. No. of SSDs installable	8	120	120
Max. No. of host interfaces	4 Port [FC (4, 8, 16Gbit/s)] 4 Port [iSCSI (10Gbit/s)] 4 Port [iSCSI (10GBase-T)]	12 Port [FC (4, 8, 16Gbit/s)] 8 Port [FC (32Gbit/s)] 12 Port [iSCSI (10Gbit/s)] 8 Port [iSCSI (10GBase-T)] 8 Port [iSCSI (25Gbit/s)] 8 Port [SAS (12Gbit/s)]	12 Port [FC (4, 8, 16Gbit/s)] 8 Port [FC (32Gbit/s)] 8 Port [NVMe over FC (32Gbit/s)] 12 Port [iSCSI (10Gbit/s)] 8 Port [iSCSI (25Gbit/s)] 4 Port [IB iSER (100Gbit/s)] 4 Port [IB SRP (100Gbit/s)] 4 Port [NVMe over InfiniBand (100Gbit/s)] 4 Port [NVMe over RoCE(100Gbit/s)] 8 Port [SAS (12Gbit/s)]
Drive interface	SAS 12 Gbps		
Supported RAID	0, 1, 1+0, 3, 5, 6, DDP (Dynamic Disk Pools)		
No. of drive shelves (DE) connected	1	3	7

For information on the management software, please see the data sheet.

Published by
FUJITSU LIMITED
Copyright 2021 FUJITSU LIMITED
www.fujitsu.com/eternus

All rights reserved, including intellectual property rights. Technical data subject to modifications and delivery subject to availability. Any liability that the data and illustrations are complete, actual or correct is excluded. Designations may be trademarks and/or copyrights of the respective manufacturer, the use of which by third parties for their own purposes may infringe the rights of such owner.
For further information see www.fujitsu.com/eternus