Technical White Paper
FUJITSU Storage ETERNUS DX8000 S3 Series – Hardware Architecture



Table of contents

Management Summary and General Remarks	2
System Architecture	3
Controller Module (CM)	5
Cache and Main Memory	5
Channel Adapter (CA)	6
Controller Enclosure (CE)	7
PCIe Flash Module (PFM) – Extreme Cache	7
Boot up and Utility Device (BUD)	7
Battery Backup Unit (BBU)	7
Front-end Router (FRT)	8
Front-end Enclosure (FE)	8
Drive Enclosure (DE)	9
Disk Drives	9
System Firmware	10
Remote Service	11
Environmental Sustainability	11
Summary	12
Appendix – Technical Specifications	12
Appendix – Online Resources	12

Management Summary and General Remarks

Based on a unique consistent system design, ETERNUS DX is a seamless family of hybrid disk storage systems ranging from economy, entry-level, and midrange systems up to the enterprise class. The major design principles of ETERNUS DX are compatibility and modularity. All models utilize the same disk types, disk shelves, racks, cables and other components, and use a uniform management software. Principally only the Controller Modules are model specific. The idea is to minimize operational, training and migration efforts and costs. Furthermore, this family concept allows upgrading a smaller model to a larger one by exchanging the control module. All other components can be reused. The uniform management system supports easy interaction between different model sizes: For example, ETERNUS DX100 S3 entry-level and DX500 S3 midrange models in remote locations can replicate data to any generation of ETERNUS DX enterprise models. The comprehensive system configuration flexibility of ETERNUS DX makes it easy to balance user requirements in terms of speed, capacity and costs.

The latest members of the family, the ETERNUS DX8000 S3 series of enterprise disk storage systems, are purpose-built for the requirements of behemoth data centers that demand outstanding storage peak performance of up to 4 million Input/Output Operations per Second (IOPS) and a huge capacity of up to 14 petabytes (PB), paired with enterprise-class resiliency and always-on availability. The systems incorporate a complete redesign of the system to achieve maximum modularity and configuration flexibility, while delivering highly scalable performance and operational robustness. This means that enterprises no longer need to precisely anticipate future capacity demand levels, nor select a specific high-end model with a fixed data-capacity range. The traditional, fixed-capacity, model-based approach typically led to a situation in which IT organizations invested in modules larger than they actually needed, just to be sure that future capacity could be covered and to avoid unplanned high migration costs. That's why Fujitsu based the ETERNUS DX8000 S3 generation on a completely new modular architecture that can be flexibly expanded according to business needs. With minimum initial cost, customers can start with a system having two controller modules and four drive enclosures holding two disks each. Maximum performance with up to 128/384 front-end ports for host connectivity can be achieved easily by scaling out to eight controller modules in ETERNUS DX8700 S3 and up to twenty-four controller modules in ETERNUS DX8900 S3. Simultaneous and flexible connection to different network types is possible. FC, FCoE and iSCSI host interfaces can be deployed in mixed configurations.

Fujitsu's Business-centric concept enables a higher level of automation for storage tiering, including support for SAS, Nearline SAS and SSD drive options and capacities that ensure an optimum service level in dynamic and fast-changing environments while also reducing the overall total cost of storage. The brand new system architecture of the enterprise systems can cope with multiple component failures without any interruption to service. In conjunction with the unique transparent failover feature – ETERNUS Storage Cluster – non-stop operation despite a complete system or site failure is guaranteed. The challenges posed by data protection are addressed using innovative Self-Encrypting Drives (SED) or system-inherent data encryption methods.

This white paper describes the architecture and hardware features of the ETERNUS DX8000 S3 models. Software functionalities and features are the subject of other publications.

System Architecture

Figure 1 shows the basic architecture of the systems – Controller Modules (CM) interconnected via PCI Express lines over Front-end Routers (FRT). The Controller Modules, as a unit, have one or more host-side interfaces Channel Adapter (CA) and one or more drive-side interfaces installed. To increase the availability of the device, each Controller Module is connected to each Front-end Router so the interconnection of the Controller Modules is quadruplicated – what we call the **Quad Star Architecture**. In addition, a Drive Enclosure (DE) for drive installation has two independent SAS-3 interface ports that are directly connected to two Controller Modules in the Controller Enclosure (CE) for path redundancy. This means that Drive Enclosures that are not directly connected to the Controller Enclosure, can be accessed from any other Controller Module via the Front-end Routers in the Front-end Enclosure (FE). Thus access to any data located in any drive is possible from any Controller Module via redundant paths.

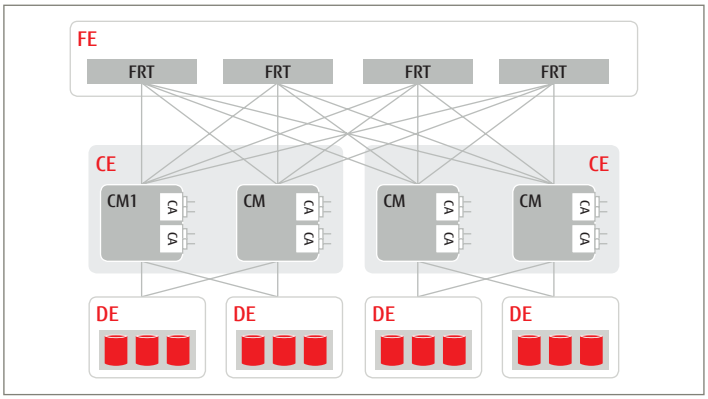


Figure 1

Figure 2 shows the schematic architecture of an ETERNUS DX8000 S3 Controller Module (CM). The basic functional blocks are the Intel Xeon processor which runs the firmware-based operating system, DRAM System Memory including the Data Cache and the Channel Adapters (CA) which realize the connection to the hosts. Each Controller Module connects to a maximum number of 16 Drive Enclosures (DE). The Drive Enclosures are connected via SAS-3 redundant paths to the SAS Expanders (EXP) in the Controller Modules. The four-ported PCIe 3.0 Interface (INF) realizes the high-bandwidth connectivity to the Front-end Routers.

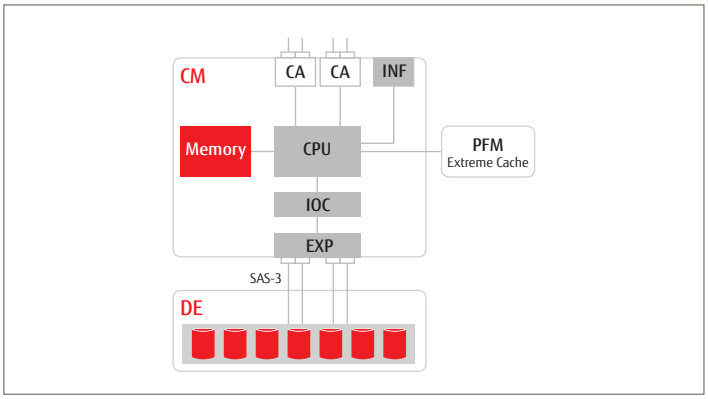


Figure 2

Each Controller Module can be optionally equipped with up to 2 PCI Express Flash Modules (PFM) which form the Extreme Cache.

The Controller Modules residing in pairs in a Controller Enclosure are interconnected via the Front-end Enclosures (FE) consisting of four redundant Front-end Routers (FRT) which enable PCIe 3.0 high-speed data interchange between the controllers (Figure 3).

The minimal configuration of an ETERNUS DX8000 S3 model consists of 2 controllers housed in one Controller Enclosure (CE) and can be enhanced by adding pairs of Controller Modules up to the maximum configuration of 24 controllers in 12 Controller Enclosures. Both the performance and the capacity of the system scale with the number of controllers.

This combination of Front-end Router PCIe 3.0 and the Back-end SAS-3 drive interface is key to enable outstanding data throughput in any ETERNUS DX8000 S3 system.

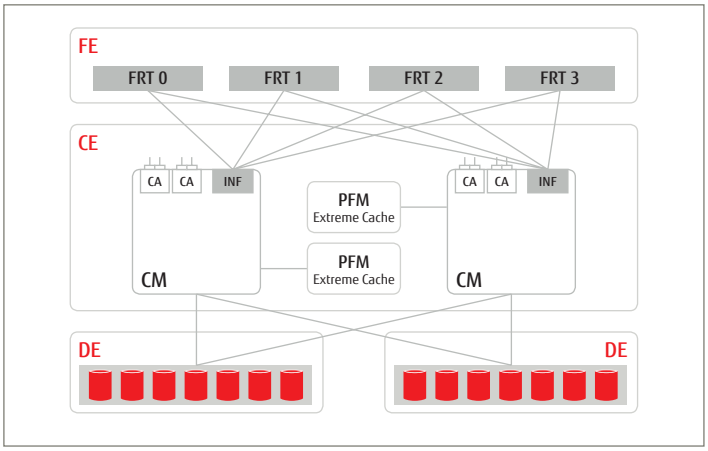


Figure 3

Legend

FE	Front-end Enclosure	CM	Controller Module
FRT	Front-end Router	CA	Channel Adapter
CE	Controller Enclosure	DE	Drive Enclosure

The schematic configuration of the highest possible configuration is shown in Figure 4.

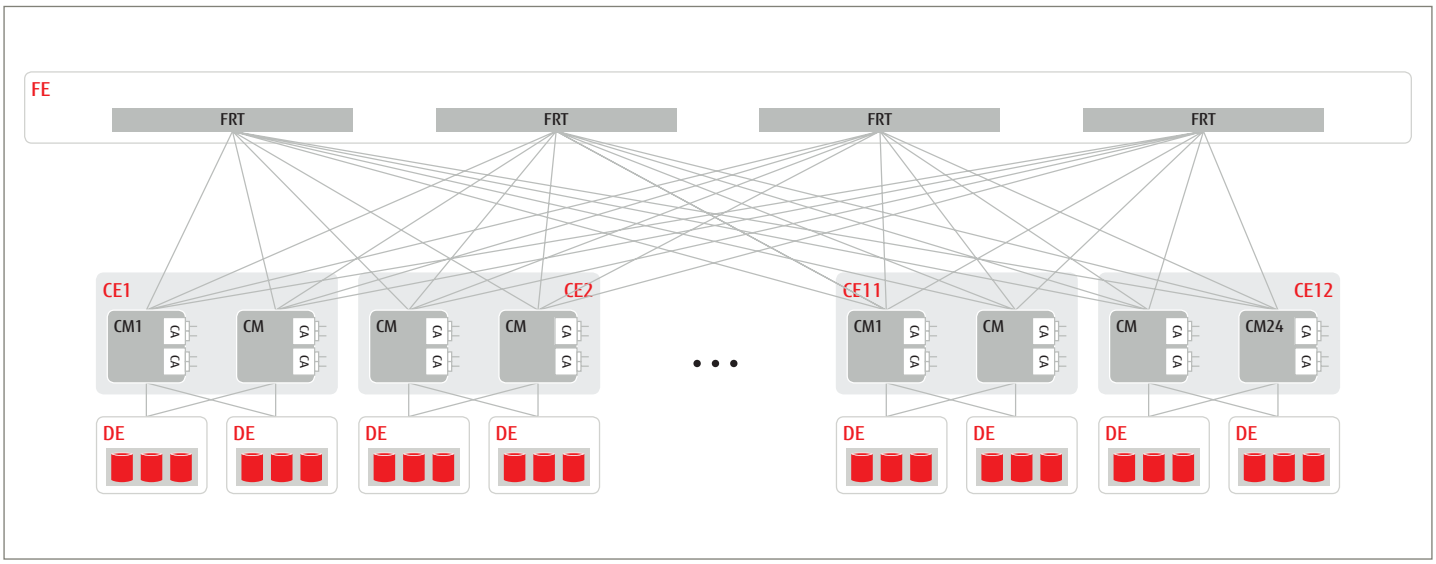


Figure 4

The maximum configuration of ETERNUS DX8700 S3 consists of eight Controller Modules with 32 Channel Adapters and 64 Drive Enclosures holding up to 1,536 disk drives.

The maximum configuration of ETERNUS DX8900 S3 consists of 24 Controller Modules with 96 Channel Adapters and 192 Drive Enclosures holding up to 4,608 disk drives.

Mechanically the system populates one 19", 42U rack (minimum); ETERNUS DX8900 S3 can be expanded up to 12 fully populated racks (Figure 5).

This architecture is the basis for the extremely flexible configuration options of Fujitsu's high-end system. In addition, 3.5" and 2.5" disk drive shelves can be freely mixed in the system.

The number of controllers is related to the number of disks and vice versa. Thus the new architecture ensures that performance always scales out when higher capacity is needed.

Furthermore, the number of Solid State Drives (SSD) in the system is not limited. This means that the system performance can be tailored in many ways.

The Quad Star Architecture delivers big performance and high-availability advantages as the internal data paths are virtual point-to-point connections, thus avoiding latencies which are typical of switched or matrix-like back-end concepts.

In addition, it offers wide-ranging scaling capacity by simultaneously determining (without limits) which disk tier classes are utilized in line with the needs of specific applications.

The same wide choice applies to the host connectors, which are just modules within the controller and can be easily exchanged as fabric requirements change.

Minimum configuration: 1 rack, 4DE, 8 drives



Maximum configuration: 12 racks, 192DE, 4,608 drives

Figure 5 – ETERNUS DX8900 S3 can be expanded up to 12 fully populated racks

Controller Module (CM)

The Controller Module is the basic component of the ETERNUS DX8000 S3 system.

This block diagram shows the components which are incorporated in the Controller Module. Up to four Channel Adapters (CA) per Controller Module realize the host connections.

The CPU is an Intel Xeon 2.6GHz 6 core processor for DX8700 S3, while DX8900 S3 uses an Intel Xeon 3.0GHz 10 core processor.

All components within the Controller Module are connected as indicated in this table:

Connection	Technology
CPU-Memory	DDR3
CPU-PFM	PCIe 2.0
CPU-CA	PCIe 3.0
CPU-INF	PCIe 3.0
CPU-IOC	PCIe 3.0
CPU-BUD	SATA 6Gb

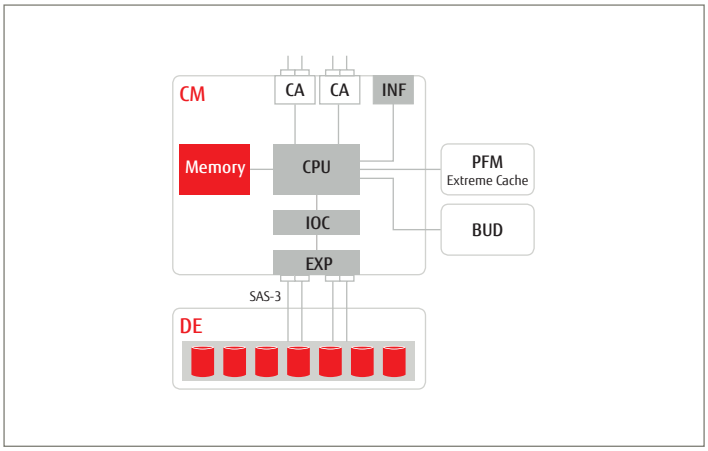


Figure 6

Controller Module (CM)
Cache and Main Memory

The main memory or cache memory consists of eight DIMM slots that can be populated with 4, 8, 16 or 32 gigabyte (GB) DDR3 DIMMs. The minimum main memory for one Controller Module is 16 GB and can be expanded to 128 GB for a DX8700 S3 Controller Module and to 256 GB for a single DX8900 S3 Controller Module, so that a fully extended ETERNUS DX8900 S3 can offer up to 6,144 GB of cache memory.

System	ETERNUS DX8700 S3	ETERNUS DX8900 S3
min. Cache/CM	16 GB	16 GB
min. Cache/CE	32 GB	32 GB
max. Cache/CE	256 GB	512 GB
max. Cache/System	1,024 GB	6,144 GB

The combination of highly dynamic cache allocation mechanisms with a very lean operating system kernel supports very efficient and performance-optimized use of the main memory for caching purposes and not for system overhead.

Controller Module (CM)
Channel Adapter (CA)

The Channel Adapters (CA) realize the connectivity to the hosts via the storage network as well as the array-to-array connectivity for remote mirroring. Up to four Channel Adapters can be mounted on each Controller Module.

Currently five different types of Channel Adapters are available
- 2-port FC 16 Gb/s
- 4-port FC 16 Gb/s
- 2-port iSCSI 10 Gb/s
- 2-port iSCSI 1 Gb/s
- 2-port FCoE 10 Gb/s

Remote replication between arrays can be realized with both FC and iSCSI.

All types of Channel Adapters can be mixed on a Controller Module or in the complete system without limitations. This offers the highest flexibility in terms of connectivity and disaster recovery options during the entire lifecycle of the system. Even if new connectivity technologies or protocols come to market, the flexible CM/CA design enables quick adoption in existing systems, and thus offers additional investment protection.

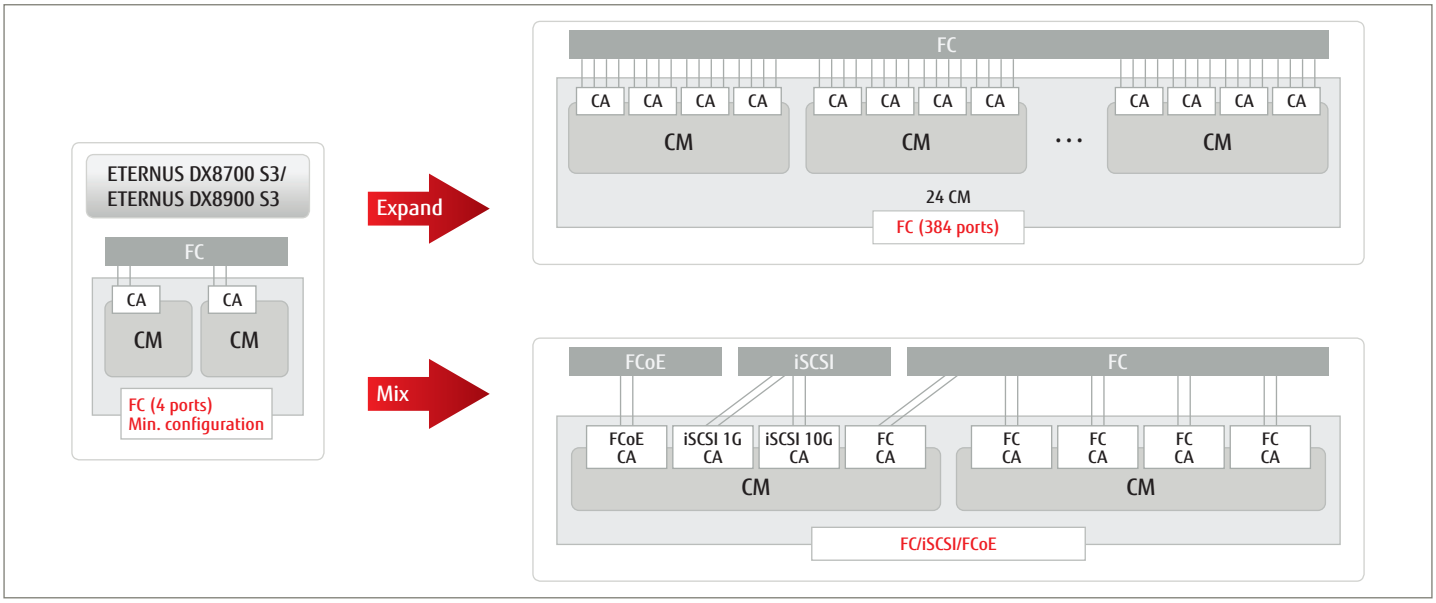


Figure 7

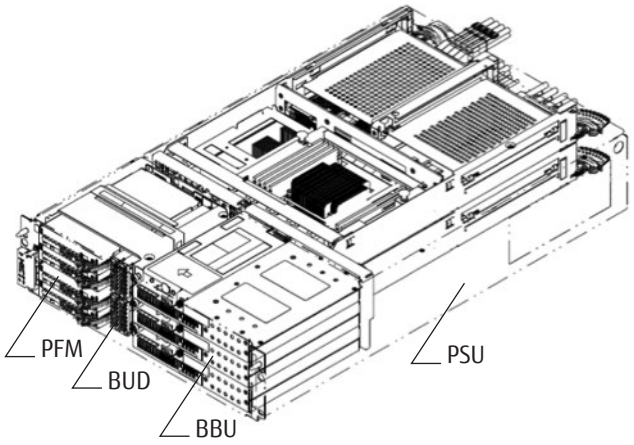
Controller Enclosure (CE)

A Controller Enclosure is a 3U, standard 19" frame and always holds as main components two Controller Modules, the optional PCI Express Flash Modules and the Boot up and Utility Devices (BUD). It also incorporates the Service Controller, Power Supplies, the Battery Backup Units and the SAS-3 drive interfaces. Up to 16 Drive Enclosures in a mixed configuration, Small Form Factor 2.5" (SFF) and Large Form Factor 3.5" (LFF), can be connected per Controller Enclosure.

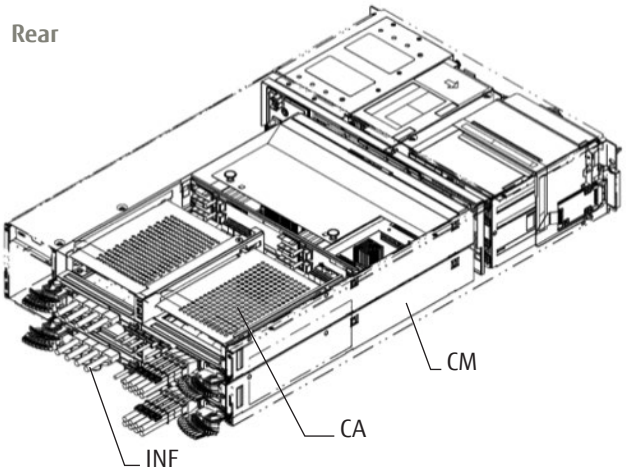
Its architecture was designed under the stipulation that all components must be non-disruptively upgradable or replaceable in the running system. Thus all components in the Controller Enclosures are hot-pluggable.

Starting with one Controller Enclosure, ETERNUS DX8700 S3 can scale to 4 Controller Enclosures/8 Controllers and ETERNUS DX8900 S3 scales up to 12 Controller Enclosures/24 Controllers.

Front



Rear



Legend

PFM	PCI Express Flash Modules	PSU	Power Supply Unit
BUD	Boot up and Utility Device	INF	PCIe 3.0 Interface
BBU	Battery Backup Unit	CA	Channel Adapter
		CM	Controller Module

PCI Express Flash Module (PFM) – Extreme Cache

The system cache can be expanded with flash storage located directly in the controller enclosure. These high-speed PCI Express Flash Modules (PFM) are used as large capacity read cache – the Extreme Cache – which provides higher performance while using fewer disk drives, optimizing system-wide costs and power consumption. One Controller Enclosure can host up to two modules per Controller Module, providing up to 5.6 terabyte (TB) cache capacity for read operations. A fully equipped ETERNUS DX8900 S3 can thus be equipped with 67.2 TB of second-level cache.

Boot up and Utility Device (BUD)

The Boot up and Utility Devices are very fast non-volatile SSD drives which have two major functions:

1. They contain the operating system software and the system configuration data
2. They provide sufficient storage capacity for flushing cached data, which has not yet been written to disk from cache, to the SSD in case of power faults.

A BUD consists of two such SSDs in a mirrored configuration.

Battery Backup Unit (BBU)

During a power outage or power failure, the Battery Backup Unit provides sufficient power to destage the data residing in the cache to the non-volatile vault drives realized by the Boot up and Utility Device. This protects the data in the cache memory without any limitations to the retention period, no matter how long the outage lasts.

Front-end Router (FRT)

The Front-end Routers are the central components of the Quad Star Architecture. They are high-performance PCIe switches and realize the any-to-any connections between the controllers. Four units which are hot-swappable in case of failures are always installed in the Front-end Enclosure.

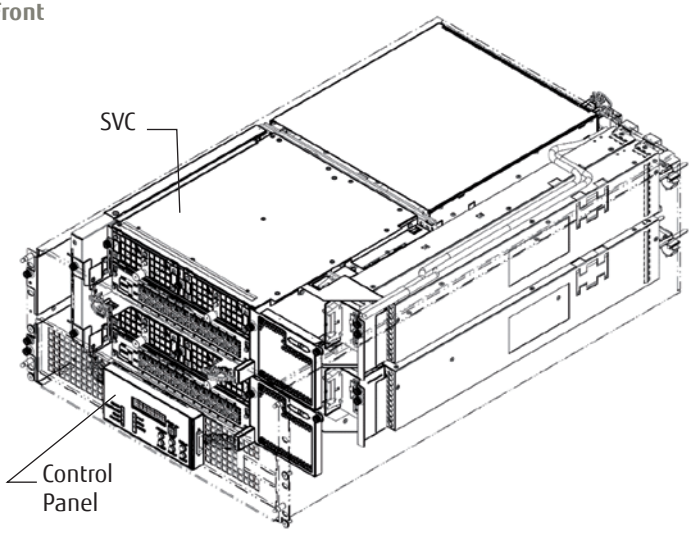
One Controller Enclosure and the Drive Enclosures connected directly to that Controller Enclosure, collectively forms a "node." The Quad Star Architecture realizes the high-performance access of Drive Enclosures on a different node, accessing the Drive Enclosure indirectly through the Front-end Router using ultra-fast PCIe 3.0 connections.

Front-end Enclosure (FE)

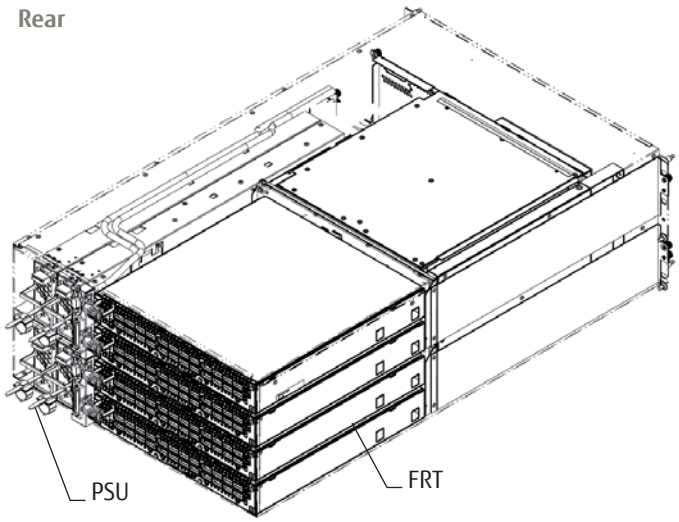
The Front-end Enclosure is the central component of the ETERNUS DX8000 S3 storage systems. It realizes the controller-controller connection and communication. It is a 5U, standard 19" frame and holds as main components the 4 Front-end Routers (FRT) at the rear. In addition, it contains fans and Service Controllers (SVC) at the front. The Service Controllers control the monitoring function for the entire ETERNUS DX system.

The integrated System Operation Panel provides basic information about the system status using LED indicators and messages on the LCD screen.

Front



Rear



Legend

SVC	Service Controllers
PSU	Power Supply Unit
FRT	Front-end Routers

Drive Enclosure (DE)

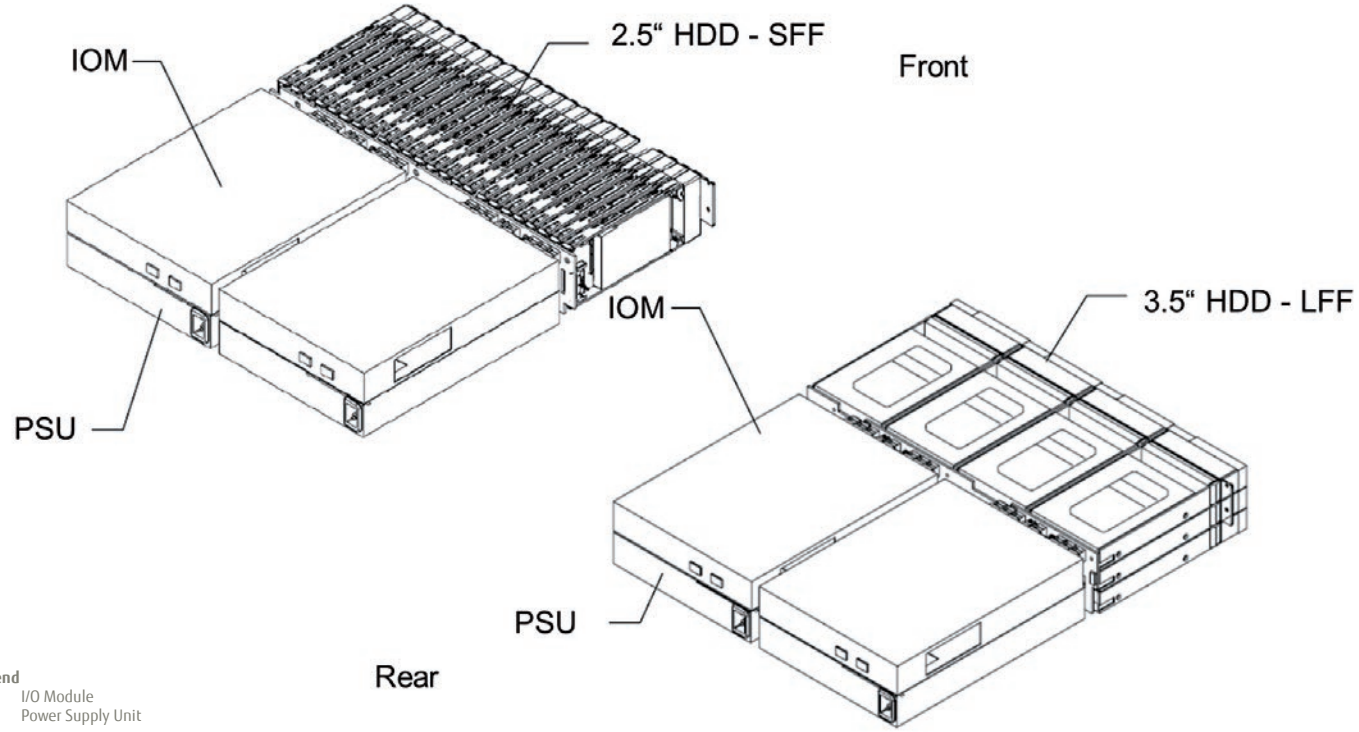
The utilized Drive Enclosure types are identical from ETERNUS DX60 S3 up to ETERNUS DX8900 S3. Drive Enclosures are 2U, standard 19" enclosures, and ETERNUS DX8900 S3 can be equipped with a maximum of 192 Drive Enclosures (DE) in a total of 12 cabinets.

A Drive Enclosure has two independent drive interface ports. Path redundancy is maintained by connecting the drive enclosure directly to two controllers. This configuration allows operation to continue even if one of the connection paths fails.

The ETERNUS DX8000 S3 series also uses reverse cabling connections for the data transfer paths between the controllers and the Drive Enclosures. This ensures continued access even if a failure occurs in a Drive Enclosure.

Two different Drive Enclosure types are available:

- a Small Form Factor (SFF) Drive Enclosure holds up to 24 x 2.5" disk drives
- a Large Form Factor (LFF) Drive Enclosure holds up to 12 x 3 .5" disk drives



Legend
IOM I/O Module
PSU Power Supply Unit

Different disk types having the same form factor can be mixed within one Drive Enclosure without restriction. Up to 16 Drive Enclosures in mixed configurations (SFF and LFF) can be connected per Controller Enclosure.

As the operating system and the system configuration data resides on the backup disks in the controller enclosure, all disks in the drive enclosures can be used for the storage of customer data.

Disk Drives

All disk drives in ETERNUS DX S3 systems are equipped with an SAS-3 12 Gb/s interface to realize the connection to the drive enclosures. Three different main types are available and realize different storage tiers in terms of performance and reliability:

- Very fast Solid State Drives (SSD) for the highest IOPS and response time requirements.
- High-performance SAS drives with both 15k and 10k RPM
- Low spinning Nearline-SAS drives with 7.2 RPM for backup, archiving and low-access data

All these types are available with different capacities in both LFF and SFF, and optionally as Self-Encrypting Disks (SED).

ETERNUS DX8000 S3 thus offers flexibility with nearly unlimited disk combination options to perfectly match any kind of application or customer requirement.

System Firmware

ETERNUS DX8000 S3 series offers multifold firmware-based features and functionalities which complement the hardware design to improve the reliability and performance of the system.

Advanced RAID organization ensures data access even during a drive enclosure failure. Each disk drive in a RAID group is located in a different Drive Enclosure, and attached to different Controller Enclosures. This keeps the RAID group accessible even if one entire Drive Enclosure, Controller or Controller Enclosure becomes unusable. Available RAID levels are RAID 1, 1+0, 5, 5+0 and 6.

Fast Recovery rebuilds data very quickly by simultaneously relocating data in the failed drive to space reserved on the remaining drives of the RAID.

Data block guard guarantees very high data integrity and improved system reliability. This function adds check codes for identification of all data written. It can also confirm/guarantee data integrity at multiple checkpoints along the data transmission path.

Disk drive patrol is a background process that constantly checks the disk drives to detect errors and write failures. Erroneous data is reconstructed and written to a new valid area on the disk drive. Thus disk write failures are avoided in a preemptive manner.

Controller based encryption prevents unauthorized data access LUN-wise, whereas Self-Encrypting Disks do not offer this kind of granularity. The selected encryption technology can either be the world-standard 128-bit AES technology or Fujitsu's unique encryption with high-process performance.

Encrypted communication between ETERNUS DX and user terminals via SSL/SSH guarantees secure transfer of data through a network and prevents unauthorized access and information theft.

Thin provisioning helps lower initial investments in disk drives while offering virtual pools of disk capacity.

Quality of Service options help administrators improve or maintain system performance for the most demanding and business-critical applications.

A more detailed description of all DX8000 S3 features and the even more sophisticated functionalities of the ETERNUS SF management suite (which complements and enhances ETERNUS DX systems) is available in the ETERNUS DX software white paper.



Remote Service

Fujitsu's AIS Connect is an optional software feature that offers call-home functionality in case of alerts, incidents and other events. In addition, the solution allows server-less remote access to ETERNUS storage systems for remote services in customer environments.



It can either be leveraged as a software functionality of the firmware or with a separate device – the AIS Connect Remote Support Gateway. The device consists of a rack-mountable 1U, standard 19" enclosure.

Implementation of this real-time service reduces the potential downtimes of ETERNUS storage systems to a minimum. The call-home feature can alert the maintenance provider to issues long before a potential system failure becomes imminent. System specialists can remotely access the system for preemptive maintenance, fast analysis and troubleshooting to prevent incidents.

Environmental Sustainability

The ETERNUS DX8000 S3 series was engineered to reduce the environmental load for both the manufacturing process and operations.

The main target in development and design was low power consumption, which has been achieved by the deployment of power-saving SFF and SSD hard drives, along with highly efficient power supplies and intelligent control of fan rotation.

In addition, power and cooling can be reduced by the ECO-mode feature which controls drive spindle rotation and disk "spin down" for idling disk drives.

Real-time monitoring of power consumption and temperature with ETERNUS SF management suite helps raise storage administrators' awareness for power savings.

The required battery capacity was significantly reduced thanks to the power-saving backup method realized with the Backup Disk implementation on the Controller Module. This was also a prerequisite for the deployment of NiMH batteries which save 50 percent in weight and 80 percent in volume while almost doubling the battery lifetime when compared to the lead batteries formerly used.

The avoidance of hazardous chemical substances is a key prerequisite throughout the entire manufacturing process and is complemented by the 3R design of systems and manufacturing – Recycle, Reuse, Reduce.

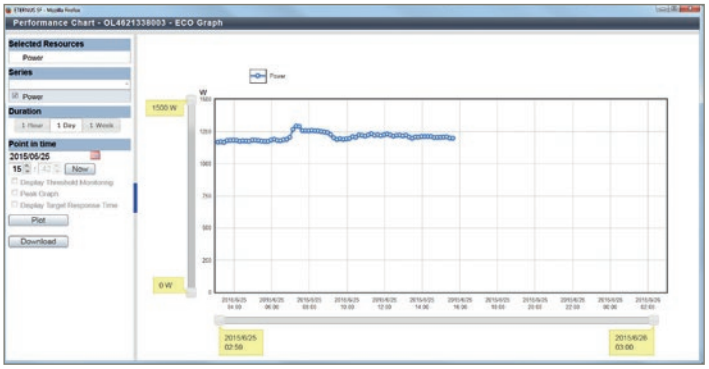


Figure 8:
Real-time monitoring of power consumption and temperature with ETERNUS SF helps raise storage administrators' awareness for saving energy.

Summary

The Fujitsu Storage ETERNUS DX8000 S3 series provides non-disruptive capacity upgrades up to 4,608 disk drives, linearly scaling performance and capacity by adding Controller Modules and Drive Enclosures. Maximum performance is provided by configurations based on from 2 to 24 controller units, with maximal 384 front-end ports for host connectivity. For simultaneous connection to different network types, FC, FCoE and iSCSI host interfaces can operate in mixed configurations.

Tiered storage optimization using a range of SAS, Nearline SAS and SSD drive options lowers the overall total cost of storage. The challenges of data protection are addressed by Self- Encrypting Drives (SED). Flexible management across different quality of service profiles, using different drive types for each application scenario, ensures efficient assignment of system resources matched to the performance priorities of different applications.

System robustness with redundant interconnections to the storage array, internal redundant components and local data copy options ensure high availability for today's non-stop business. Additional extensive high-availability and disaster recovery features including remote replication and its native transparent fail-over capability make ETERNUS DX enterprise models the ideal storage systems for all business-critical data – Business-centric Storage.

Appendix – Technical Specifications

	ETERNUS DX8700 S3	ETERNUS DX8900 S3
Architecture	Scale-out	Scale-out
Storage controllers	2-8	2-24
Maximum storage capacity	4,608 TB	13,824 TB
Maximum disk drives	1,536	4,608
Maximum cache memory (DRAM)	1 TB	6 TB
Maximum second-level cache (SSD)	22.4 TB	67.2 TB
Maximum connectable hosts	8,192	8,192

Appendix – Online Resources

- Fujitsu Global Portal: ETERNUS DX8700 S3 product page
- Fujitsu Global Portal: ETERNUS DX8900 S3 product page
- Datasheet: ETERNUS DX8700 S3
- Datasheet: ETERNUS DX8900 S3
- Brochure: ETERNUS DX S3 Family
- Flyer: ETERNUS DX S3
- Whitepaper: ETERNUS DX Feature Set
- Whitepaper: ETERNUS SF

Published by
Fujitsu Limited
Copyright © 2015 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