FUJITSU storage ETERNUS CS800 data protection appliance

Technical white paper

The FUJITSU Storage ETERNUS CS800 backup appliance provides backup to disk with deduplication and radically reduced costs, through savings in the required storage capacity of up to 95 percent.

Contents

Introduction

When thinking about a backup strategy, companies have to consider the challenges they want to solve with such a strategy. The backup process has to be considered and also the question of data recovery speed and ease must be answered. The next question concerns the quality of the restored data: how old is the backup data? It's obvious that data is more accurate, the younger the backup – hopefully not too long ago. However, the latest backup is sometimes not good enough and there is a need to restore older files. But how many versions of backup data can be stored until capacity is exceeded? The backup window is shrinking just as the amount of data is constantly growing for most companies. Traditional backup-to-tape is no longer sufficient and many IT departments are looking for a faster backup, improved restores, more efficient disaster recovery and lower costs. The problem is selecting the correct products. Backup-to-disk products get full too quickly – that is the reason for the growing adoption of disk-based storage systems with deduplication as backup target systems. Finding the right deduplication solution is difficult. Appliances often have narrow scalability and interconnect options, forcing users to make decisions now that limit their future choices and hidden charges can make real costs unpredictable. Software-only options are often complex and limited in effectiveness. The ETERNUS CS800 – a disk-based backup appliance - provides deduplication without compromise, including hybrid distributed deduplication (SPEED mode), combining broad scalability and high performance to support the broadest range of IT environments today and into the future. Built on leading deduplication technology the ETERNUS CS800 appliance is a purpose-built solution that scales from 11 to 23 TB (CS800 ENTRY), 27 to 315 TB (CS800 SCALE) and 102 TB to 1020 TB (CS800 ENTERPRISE) of usable capacity, and provides fast backup performance – achieving up to 24 TB/hour in conventional inline target mode(CS800 ENTERPRISE). ETERNUS CS800 gives IT departments' optimal investment protection with a combination of maximum flexibility and value. With exception of CS800 ENTRY, which does not offer a VTL interface, all units provide the option to use simultaneously NAS, VTL, and OST interfaces, making it easy for users to add or change interconnect choices as their backup environments evolve. To keep acquisition costs low now and in the future, the appliance base price includes all the software licenses, including (depending on the chosen configuration) self-encrypting drives and AES 256-bit encrypted remote replication for automated DR protection, and direct tape creation for integrated long-term retention. ETERNUS CS800 provides IT departments with multi-site, multi-tier protection that is easy to manage – both for near-term restore and long-term retention in any backup environment. The end results are improved protection, streamlined management, and lower overall costs.

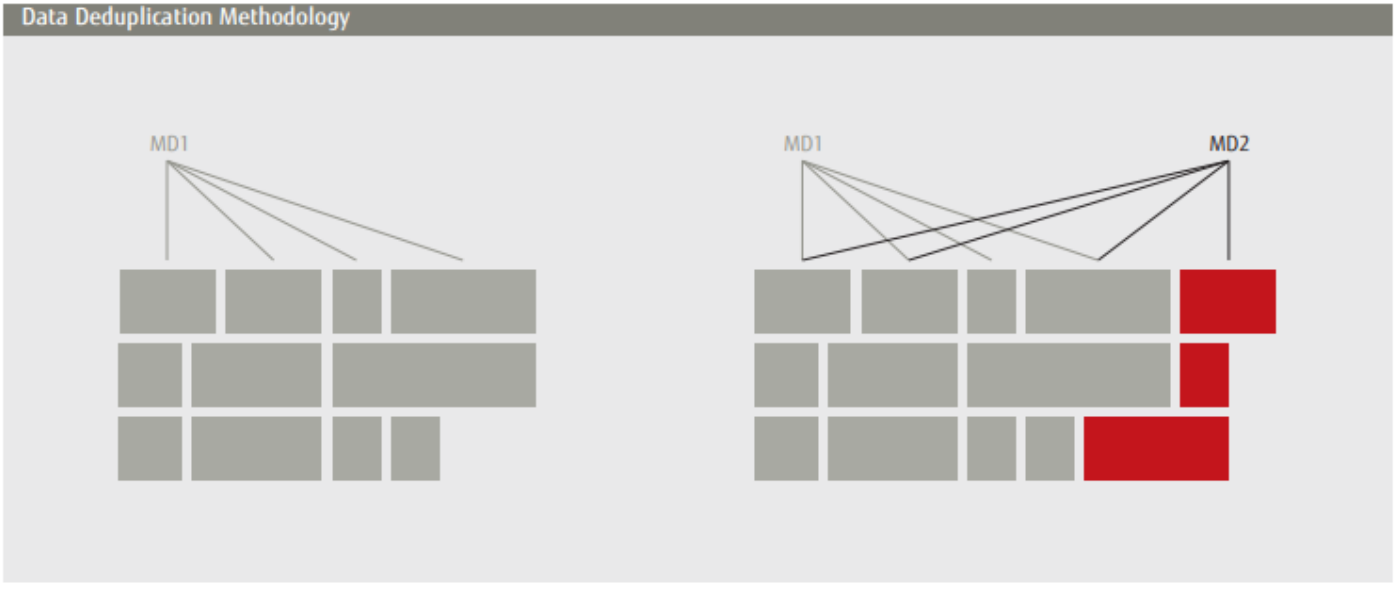
This white paper highlights the main features and concepts of an ETERNUS CS800.

Deduplication is the key

Data deduplication refers to a specific approach to reduce data and is built on a methodology that systematically substitutes reference pointers for redundant variable-length blocks (or data segments) in a specific dataset. The purpose of data deduplication is to increase the amount of information that can be stored on disk arrays and to increase the effective amount of data that can be transmitted over networks. Data deduplication can be combined with file-based data reduction systems to increase their effectiveness. Since the leverage of data deduplication is highest when there is a repeated attempt to write the same data, this technology is most frequently used today to store backup data.

How Deduplication works

In short, data deduplication operates by segmenting a dataset in a backup environment – normally a stream of backup data – into blocks and writing those blocks to a disk target. To identify blocks in a transmitted stream, the data deduplication engine creates a digital signature like a fingerprint for each data segment and an index of the signatures for a given repository. The index, which can be recreated from the stored data segments, provides the reference list to determine whether blocks already exist in a repository. The index is used to determine which data segments need to be stored. When data deduplication software sees a block that has been processed before, it inserts a pointer to the original block in the dataset's metadata instead of storing the block again. If the same block shows up multiple times, multiple pointers to it are generated. Data deduplication technology stores multiple sets of discrete metadata images, each of which represents a different dataset but all of which reference blocks contained in a common storage pool. Since the leverage of the data deduplication technology is highest when there are repeated data segments, the technology is most frequently used today to store backup data. The methodology allows the disk to support retention of backup data sets over an extended length of time, and it can be used to recover files or whole data sets from any of multiple backup events. ETERNUS CS800 is built around the deduplication methodology to provide cost-efficient disk storage capacity for backup data and fast recovery times.



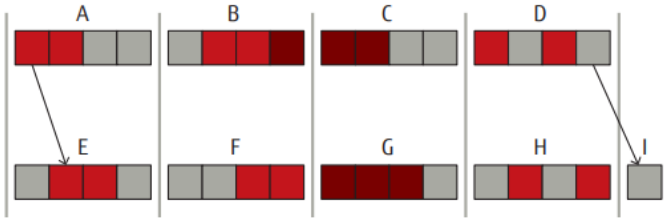
Variable-length data segments vs. fixed-length blocks

The deduplication technology used within ETERNUS CS800 divides the data stream into variable-length data segments using a methodology that can find the same block boundaries in different locations and contexts. This block-creation process allows the boundaries to "float" within the data stream so that changes in one part of the dataset have little or no impact on the boundaries in other locations of the dataset. Through this method, duplicate data segments can be found

- at different locations inside a file
- inside different files
- inside files created by different applications and
- inside files created at different times.

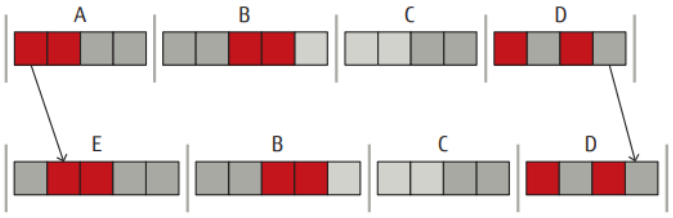
Another method to look for repeated blocks in transmitted data is using fixed-length blocks, an approach that is currently being used by several backup software suppliers to include deduplication as a feature of the software. However, the fixed block approach achieves substantially less effective reduction than a variable-block approach. The reason is that the primary opportunity for data reduction in a backup environment is in finding duplicate blocks in two transmitted data sets that are made up mostly but not completely of the same segments. If we divide a backup data stream into fixed-length blocks, any change in size to one part of the dataset creates changes in all the downstream blocks the next time the data set is transmitted. Therefore, two data sets with a small amount of difference are likely to have very few identical blocks.

Here is an example of applying fixed block lengths to a data sequence:



The upper line shows the original block division – the lower line shows the blocks after making a single change to Block A (an insertion). In spite of the fact that the shaded sequence of information is identical in the upper and lower lines, all of the blocks have changed content and no duplication is detected. If we stored both sequences, we would have 8 unique blocks.

And this is what the same example looks like when applying variable-length segmentation:



Data deduplication utilizes variable-length blocks or data segmentswhen looking at a data sequence. In this case, Block A changes when the new data is added (it is now E), but none of the other blocks is affected. Blocks B, C, and D are all recognized as identical to the same blocks in the first line. If we stored both sequences, we would have only 5 unique blocks.

The deduplication ratio

When a backup to an ETERNUS CS800 is processed for the first time, the number of repeated data segments varies widely depending on the nature of the data. The effect can range from negligible benefit to a gain of 50 percent or more in storage efficiency. With the following sequence of backups from the same specific disk volume, the benefit will typically become very significant, as similar datasets are written to the deduplication pool and each new write operation only increase the size of the total pool by the number of new data segments. Normally in data sets representing conventional business operations, the change rate of data between two backup events is only 1 percent or 2 percent. Higher change rates are sometimes seen. The total number of data segments stored over multiple backup events also depends to a very great extent on the retention policies set by the user – the number of backup jobs and length of time they are held on disk. The difference between the amount of space that would be required to store the total number of backup datasets in a conventional disk storage system and the capacity used by an ETERNUS CS800 system is referred to as the deduplication ratio and is represented by the following formula:

$$\text{Deduplication} = \frac{\text{Total Data Before Reduction}}{\text{Total Data After Reduction}}$$

To learn more about the deduplication ratio and when deduplication becomes really powerful, please read the white paper “ETERNUS CS800 – data deduplication background”.

.

Replication and disaster protection

Disaster protection is a persistent IT problem. The minimum disaster protection required from every IT organization is ensuring that backup data is safe from site loss or damage. Equipment and applications can be replaced eventually, but digital assets are often irreplaceable. No matter how resilient or redundant a given storage or backup system may be or how many layers of redundancy it might have, when all copies of data are located at a single site and in a single hardware system, they are vulnerable to site-specific damage, including natural disasters, fire, theft, and malicious or accidental equipment damage. Managing removable media, especially across multiple remote sites, is expensive, time-consuming and error prone. Traditional disk-to-disk systems do not offer an effective solution since backup volumes are too large to allow replication for most users.

Applying data deduplication to replication

Deduplication does not only provide storage benefits but the technology also dramatically reduces the bandwidth needed to copy data over networks. The result gives disk backup a practical way to provide WAN-based disaster recovery protection and to reduce requirements for removable media.

ETERNUS CS800 uses these advantages provided by its deduplication technology and offers a replication feature that is available with every appliance. Replication uses the Ethernet protocol, to efficiently transport a complete copy of user data residing on one ETERNUS CS800 (“the source”) to another ETERNUS CS800 (“the target”). High efficiency is achieved by transporting only the unique data blocks plus metadata from source to target.

There are two generally accepted models for replication: synchronous and asynchronous

- Synchronous replication often referred to as mirroring, continuously maintains two primary, active datasets in the same state by transferring blocks between two storage systems at each I/O cycle. Synchronous replication is normally designed to provide very rapid failover to the replica if the primary dataset is corrupt.
- Asynchronous replication can also be applied to non-dynamic, pointin-time images, including backup images, to provide site loss and disaster recovery protection. The technique is much less complex to implement than mirroring techniques. Backup data is a good replication candidate for DR purposes – it is a point-in-time copy of the primary data, and it is isolated from the primary applications by the backup process.

Data deduplication makes the process of replicating backup data practical by reducing the bandwidth and cost needed to create and maintain duplicate datasets over networks. At a basic level, deduplication-enabled replication is similar to deduplication-enabled data stores. Once two images of a backup data store are created, all that is required to keep the replica or target identical to the source is the periodic copying and movement of the new data segments added during each backup event, along with its metadata image, or namespace.

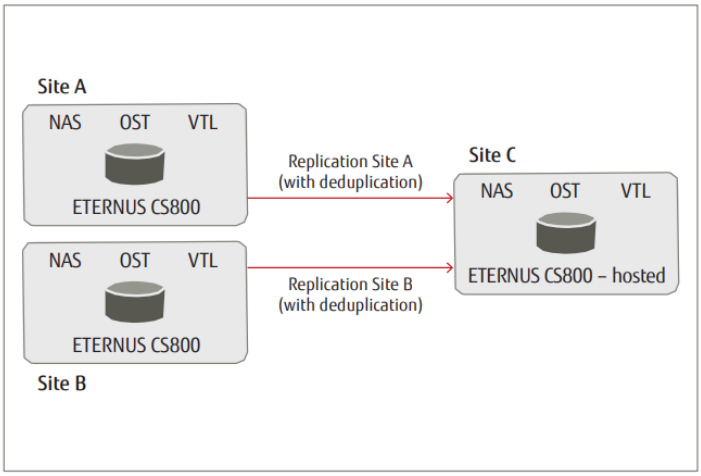
ETERNUS CS800 uses asynchronous replication to create and maintain duplicate images of backup datasets on different devices using transmission over WAN connections. This replication is bidirectional and can be performed in parallel to more than one location. The replication process begins by copying all the data segments in one division of a source appliance to an equivalent division in a second, target appliance.

After the source and target are synchronized, the replication process only sends the new data segments for each new backup event written to

the source. If the new backup event has changed by 1 percent, the expected bandwidth requirement to create the replica will be 1/100 of the bandwidth that would have been needed to replicate the entire backup dataset written to the source.

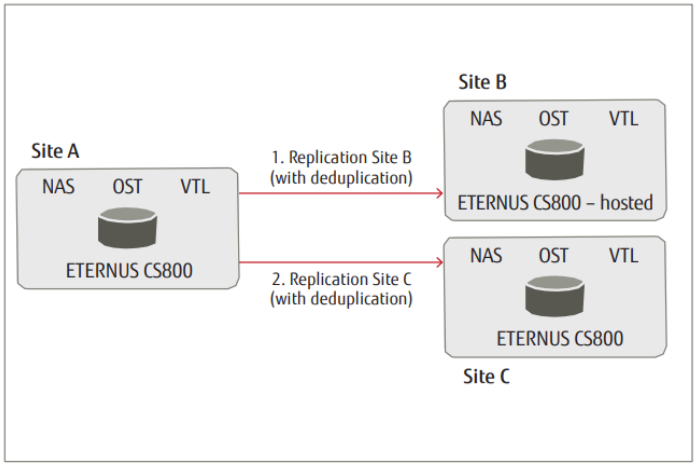
The bandwidth requirement might be reduced further because of a two-stage, pre-transmission process as part of the replication software implemented in an ETERNUS CS800. In this system, before any data is sent to the target device, the ETERNUS CS800 replication software sends a list of the blocks available for replication to the target device. The target ETERNUS CS800 checks the list of data segments against the index of data segments it has already stored, and it returns a list of elements that are not already locally available and that needs to be sent from the source appliance. The source then sends copies of only the new data segments over the network. The data segments are sent in the background, the process begins as soon as the backup job has started to be written to the source, and the replication is completed when the metadata for the new backup image is transmitted. At that point, the backup image is available for recovery at the target. Once the replication process is set up in the ETERNUS CS800 it runs automatically without further human interaction.

The ETERNUS CS800 replication software allows multiple sources appliances to point to the same target device and replication normally takes place on a partition-to-partition basis, i.e. each source consists of a specific appliance partition that replicates data to a similar image on a source device, either a NAS share or a virtual library partition. All replication images at the target are supported by a common deduplication pool which deduplicates data segments across all backup images sent. That means that deduplication takes place between different source sites. The picture on the right shows that if the same blocks are backed up at source sites A and B, they only have to be stored once on a common site C when both A and B are replicating data to the same target appliance.



Multi-source Replication to one Target

The pre-transmission process that checks to see which data segments are already present at the target site is an important feature of the ETERNUS CS800 replication process. It makes sure that identical data segments are not stored twice even if they come from different source systems. For example, a data segment was backed up yesterday at source site A and today at source site B, both sites replicate to site C where this data segment is only stored once and due to the pre-transmission process even only sent once over the network. Only the metadata needs to be sent and stored. This pre-transmission deduplication of the data segments can significantly reduce the bandwidth needed for replication in environments where users in distributed sites work on similar file sets.



1:2 Replication

To increase the level of disaster protection ETERNUS CS800 has an option of replicating its backup data to a second remote site. The process of replication is the same as described before. After the pre-transmission process that checks the existence of identical data segments on the target systems site B and site C, the decision is made whether the whole data block has to be sent to the target system or only the metadata. To guarantee the highest level of synchronism on all involved systems, both replication processes from source site A can start in parallel.

Applying data deduplication to replication

As many organizations use public data exchanges to supply WAN services between distributed sites and the data transmitted between sites can take multiple paths from source to target, deduplication appliances need to offer encryption capabilities in order to ensure the security of data transmissions. In the case of ETERNUS CS800, all replicated data (metadata and actual blocks of data) can be encrypted at source level using SHA-AES 128-bit encryption and decrypted at the target appliance.

Encryption is available with every ETERNUS CS800 system and free of extra charge.

Integrated path-to-tape

Quite a lot of organizations still need to store some of their backup data on tape, even as they move to disk-based backup. As a result of Fujitsu's expertise in both tape and disk-based backup, Fujitsu can provide an integrated solution that supports both.

ETERNUS CS800 offers an application-specific direct path-to-tape (PTT) as a standard feature on SCALE and ENTERPRISE models to provide optimal integration of short-term backup/restore with economical long-term retention. This feature writes backup data directly from the appliance to an attached tape library over an FC link, without sending the data back through a media server. Direct tape creation reduces loading on media servers, and it makes tape creation an automated process that takes place outside the backup window.

ETERNUS CS800 supports two kinds of PTT:

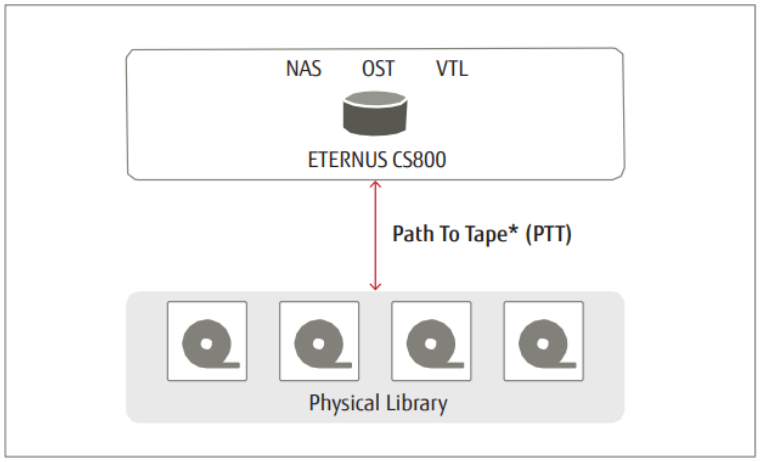
- Application-specific PTT and
- OST PTT.

Both methods use tape drives connected to ETERNUS CS800 via dedicated back-end FC interfaces, but differ in the front-end presentation used to bring data into ETERNUS CS800:

- at different locations inside a file Application-specific PTT is used with an ETERNUS CS800 FC VTL partition, with command and control traffic from the backup server occurring via NDMP.
- OST PTT uses Veritas' proprietary OST API, and both – commands and data – flow via Ethernet to an OST storage server defined on ETERNUS CS800.

In all cases, tapes are written in the application format, exactly as written by the application. This means recovery from tape does not require an ETERNUS CS800, keeping DR easy. All it takes to read the tapes is a server running the appropriate software and a tape drive.

Data flows out of ETERNUS CS800 FC PTT interfaces in native, uncompressed form. When it is received at the tape drive, the drive compresses it and writes the data to the physical tape medium. This provides a capacity benefit based on the compressibility of each data set. Tape drive compression also provides a potential throughput benefit, realized in cases where ETERNUS CS800 PTT throughput exceeds the native write throughput rating of the physical tape drive being used. This benefit is at its greatest when using older tape technologies, such as SDLT or early-generation LTO. Current generation drives are fast enough to keep up with ETERNUS CS800 PTT output stream.



The changing role of tapes

In traditional backup environments, tapes play an important role as backup target. Backups have been performed during idle times of the servers in the night, using the power of these servers for running the backup application. There was typically one 'backup window' during which this work had to be completed.

Today, several things have changed that challenge this traditional model. The adoption of disk-based backup means data is protected on-site more quickly and tape creation becomes a secondary process supporting longterm retention or DR objectives. At the same time, the prevalence of 24×7 operations mean that servers

that were formerly idle at night are now busy all the time. Furthermore, the trend toward virtualized server infrastructures allows much higher utilization of physical server resources, eliminating the resource that was so handy at tape creation time.

PTT allows the tape creation load to be moved onto ETERNUS CS800, minimizing the impact of backup on server loading and contention for valuable computing resources. Since the application initiates the PTT activity, it is aware of all the data copies. A side effect is that manual tape handling, which can become error-prone and needs special expertise, is substituted by an automatic creation of tapes.

ETERNUS CS800 perfectly combines the need to have a local copy of the backup data in addition to the off-site copy. ETERNUS CS800 fulfils this need locally with the PTT option, while ETERNUS CS800 replication may be used to provide off-site DR protection. This leaves tape serving for long-term data retention and moves tape creation out of the critical nightly data protection window.

Like any function or feature PTT has limits, and it fits some customer requirements better than others.

To learn more about PTT and where it fits best please read “ETERNUS CS800 understanding Integrated Path to Tape”.

Data integrity architecture

Let's just recall what deduplication on ETERNUS CS800 means:

- data is segmented into variable length blocks
- a digital signature is generated for each segment
- the blocks are written to the system's RAID storage array
- and the signatures are stored in an index.

As data is written to an ETERNUS CS800, each unique block only needs to be stored once. If a block is sent to the system again, a pointer to the original instance of the data segment is stored instead of storing the block again. Since the pointer uses a small fraction of the space used by the original block, data deduplication technology dramatically increases the capacity of the disk to retain data. The same logic is also applied to replication of deduplicated backup data between different ETERNUS CS800 systems. The replication source checks the target before sending blocks to see if the target already holds copies of them, and it only sends new blocks during replication. The effect is that replication of deduplicated backup sets uses much less bandwidth than a conventional backup.

In this highly automated and optimized world of an ETERNUS CS800 it is essential that the integrity of the backup data is ensured at all times. ETERNUS CS800 configuration options thus provide multiple levels of automated processes which monitor, protect, and maintain the integrity of the system and the data stored within.

Solution level protection

- Built-in system checks
- Auto signature verify
- Auto index recovery
- Auto-data re-acquition
- Direct media creation

File system level protection

- Full-segment writes
- Append-only writes
- Resilient, high capacity industry standard file system

Hardware level protection

- Dedicated memory allocation
- Power loss protection
- Data block guard
- Industry-Standard RAID
- Enterprise redundancy
- SMART monitoring
- Dynamic sector repair

Solution level integrity processes

- **Built-in system integrity checks:** These routines protect ETERNUS CS800 systems by continually checking the state of the system's hardware and software conditions using a built-in, automated test process during normal operations. If anomalies are detected, full data and index verification/correction processes are performed in the background.
- **Automated signature verification:** During all read, access, and replication operations, the system checks the validity of the index signatures in real time. If a fault were to occur, the system would invoke correction activities, automatically recovering the index or re-acquiring the segment. In replication deployments, recovery can be made from any replicating location.

- **Auto-index recovery:**
  If any faults are identified in the signature index, the index is automatically corrected by regenerating the index signatures from the original blocks.
- **Auto-data re-acquisition:**
  If faults were to occur in the data blocks, an automated process is initiated to correct blocks by re-acquiring the segments during subsequent backup events.

Direct creation of removable media:

The configuration options of ETERNUS CS800 can offload data directly from internal disks to tape. This is particularly valuable for long term backup data. The path-to-tape (PTT) feature is available via the OST or VTL protocol for selected backup applications.

File system level integrity processes

- **Full-segment write policies:**
  Data is collected into a full-segment before being written to disk to both optimize performance and increase the integrity of the write process.
- **Append-only write policies:**
  New data is only appended to existing data pool. Data blocks are never deleted or made available for overwrites until the data status and all references are verified.
- **Resilient, high capacity file system:**
  ETERNUS CS800 leverages a high performance and secure file system technology, which can handle millions of files.
- **Enterprise-level redundancy:**
  All ETERNUS CS800 configuration options include redundant, hot swappable components (fans, power supplies) as a standard feature. ETERNUS CS800 ENTERPRISE systems also include dual redundant RAID controllers, redundant interconnect with path failover and dynamic disk pools. With DDP, customers avoid high density drawbacks with faster rebuilds and lower performance impact during rebuilds, which means less impact on the backup window.
- **SMART predictive monitoring:**
  ETERNUS CS800 systems continually monitor system-level trend data using the industry-standard system monitoring and reporting technology (SMART) protocol, including recoverable error rates and re-tries. When specific thresholds are reached, the system issues alerts recommending preventive maintenance.

Hardware system level integrity processes

- **Dedicated, internal memory allocation:**
  Isolates all data deduplication processes and indexes from other operations to ensure integrity.
- **Power loss protection:**
  All data written to disk is maintained in a protected RAM that utilizes capacitors to provide unlimited time protection for cached data (as opposed to using battery packs). Using this method data can be maintained and ultimately recorded correctly even in the event of a power failure or accidental drive removal/insertion.
- **Dynamic sector repair:**
  In the event of a read failure from any drive sector, the system automatically re-directs the request to a redundant copy of the data maintained on disk. It then isolates the faulty sector and creates a new sector to maintain full internal redundancy

Support of Veritas OST

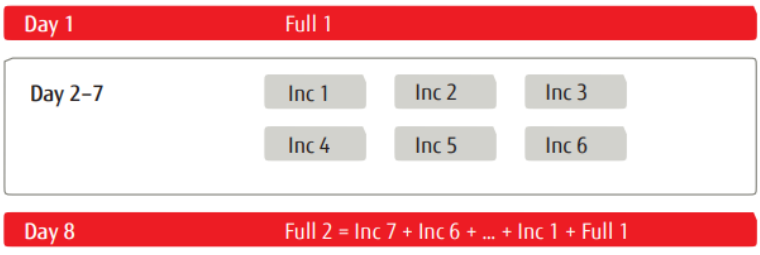
The Veritas OST (open storage) API allows storage product vendors to create an interface that enables NetBackup and BackupExec to manage the creation, retention, duplication (multiple copies), and capacity quotas through a common interface and completely managed by the Veritas application. In NetBackup, control is provided by storage lifecycle policies. The basic elements in any OST environment are a "storage server" – which is an interface element set up on the appliance – and a "Logical Storage Unit" (LSU) – which is a specific location for data within the Storage Server. Backup data in an OST environment is written by a media server to an LSU on the storage server, and data can be read back to the media server from the LSU. It is the LSU and its contents that are managed by the application. Veritas OST allows NetBackup and Backup Exec to seamlessly integrate with an ETERNUS CS800 disk backup system. Once installed and configured, NetBackup or Backup Exec can manage the backups through the ETERNUS CS800 and take advantage of the system's capabilities such as data deduplication, replication or direct tape creation. OST support in the ETERNUS CS800 means that users can create Storage Server on an ETERNUS CS800 system, and create Logical Storage Units in it for writing and reading backup data. It also means that Fujitsu supplies plug-in software that can be installed on Veritas media servers to allow them to communicate with the Storage Servers on the ETERNUS CS800 system. For users who have standardized on NetBackup or BackupExec, OST provides the potential to increase ingest performance in some scenarios and to streamline management of backup data across multiple sites and multiple tiers. To operate via OST interface the media server has to be set up with an ETERNUS CS800 specific OST Plug-in. After configuration of the ETERNUS CS800 the following processes are available via OST to increase backup performance and disaster protection:

- Optimized synthetic full backups
- OST optimized duplication - optimized duplication - auto Image replication - concurrent deduplication
- Hybrid deduplication – SPEED mode

Optimized Synthetic Full Backups

A synthetic backup is a backup in which content from a series of backups, starting with a full backup and followed by incremental backups, is combined into a new full backup.

For a synthetic backup to be successful, all the backups from the series must be intact and internally consistent. Intact means that each backup created still exists and no backup in the series has been expired. Internally consistent means that the individual backup catalogs are all consistent. During a synthetic backup, the backup application first queries and analyzes the backups for the given policy and client and then attempts to use their catalog data to combine the content into a new logical backup.

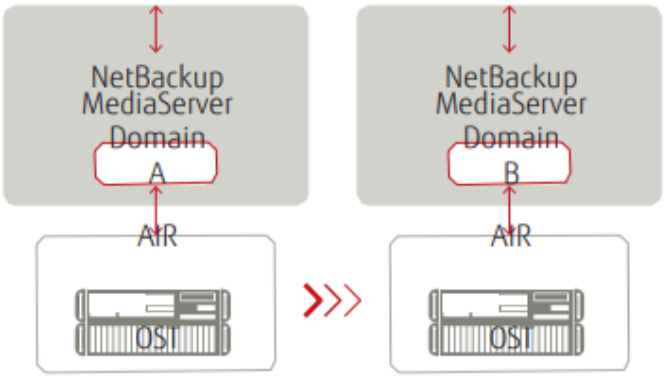


OST optimized duplication

ETERNUS CS800 has the capability to perform replication. NetBackup or Backup Exec uses this capability to initiate an optimized duplication of backup images between these appliances. The duplication operation of NetBackup or Backup Exec triggers the replication function in the OST disk appliance if both the source and destination volumes for the copy are OST LSUs.

OST optimized duplication reduces the workload on the NetBackup or BackupExec media server because the replication is performed by the ETERNUS CS800. Duplication is done in the background, and it is faster because it uses data deduplication capabilities to reduce the copy bandwidth. Duplication is still initiated, managed, and controlled by the NetBackup or BackupExec media server while the actual data movement process is offloaded to gain the maximum benefits from the appliance's replication capabilities. The ETERNUS CS800 can replicate OST data to another ETERNUS CS800 using the following methods:

- Optimized Duplication - with optimized duplication, backup images on a storage server can be replicated to another storage server on an ETERNUS CS800 that resides in the same NetBackup or BackupExec domain. The duplication occurs when it is initiated in NetBackup.
- AutomaticImage Replication (AIR) - by using Veritas NetBackup 7.1 or higher, an LSU for Automatic Image Replication can be configured. If enabled, data on an LSU is automatically replicated to a remote LSU that resides on an ETERNUS CS800 in a different NetBackup domain. The timing of the replication, as well as the backup images that are replicated, are determined by the storage lifecycle policies (SLPs) configured in NetBackup. The primary purpose of Auto Image Replication is to create off-site copies of mission critical backups to protect against site loss. The duplicate copy is available at the disaster recovery site as soon as the duplication has completed. This offers increased data protection for the most critical data.



- Concurrent Optimized Duplication – For both optimized duplication and Automatic Image Replication, Concurrent Optimized Duplication can optionally be enabled. If enabled, as data is written to the storage server, it is simultaneously replicated to the target ETERNUS CS800. When optimized duplication or Automatic Image Replication subsequently occurs, the operation is more efficient because a portion of the required data has already been replicated to the target storage server.

It is important to remember that, with automatic image replication, the local and remote LSUs reside in different NetBackup domains. This differs from optimized duplication, which occurs between two LSUs residing within the same NetBackup domain.

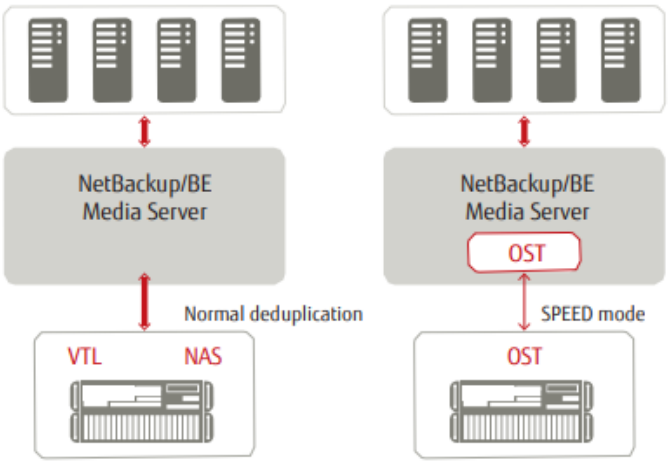
Hybrid deduplication – SPEED mode

Hybrid deduplication accelerates backups and reduces network bandwidth requirements by distributing deduplication between the backup server and the ETERNUS CS800.

To understand the full benefit of hybrid deduplication let us compare the different ways backup data are sent to ETERNUS CS800:

- In normal environments all backup data is sent over the network to the appliance where the blocks are identified, signatures created and checked against the index. Only unique blocks are then compressed and stored.
- Working in a hybrid deduplication environment the workload is distributed between the media server and the appliance. The signatures are already created on the media server and sent to the ETERNUS CS800, where they are checked against the index. The appliance reports to the media server which blocks are unique and only these blocks were compressed and sent over the network to the ETERNUS CS800 to be stored.

This means that hybrid deduplication results in increased backup performance and minimizes the workload on the network. To use hybrid deduplication the media server has to be configured with an OST Plug-in provided by ETERNUS CS800.



OST + PTT

Application specific PTT is used to create physical tapes locally at the site that hosts the ETERNUS CS800 to which the data was originally backed up, namely the source site. Because OST allows the backup server to manage optimized duplication (replication) as well as tape creation, OST PTT enables the creation of tapes at a target site without requiring the presence of a server at that site. All you need is the target ETERNUS CS800 with an attached tape library. The backup application remains aware of all copies of the data and they may all be managed independently whether on the source ETERNUS CS800, the target ETERNUS CS800, or on tape.

NetBackup accelerator

The shrinking backup windows, older network infrastructures and the growing use of virtual environment can be addresses with the NetBackup accelerator feature. Compared to the classic Backup were all data transferred from the client, to the media server and to the backup device (Figure A) the new NetBackup accelerator enables the client to send only the changed data on a file level directly to the ETERNUS CS800 (Figure B). This reduces the network load drastically.

By placing the Synthetic Engine to summarize all the incremental to a new full backup in the ETERNUS CS800, significant hardware in the rest of the environment can be saved or used for other tasks. By that the backup window shrinks and the SLA fulfillment is easier to achieve. This does not only apply for the regular files and folders, it also works for VMware's environments and saves license costs duo the reduced Hardware utilization.

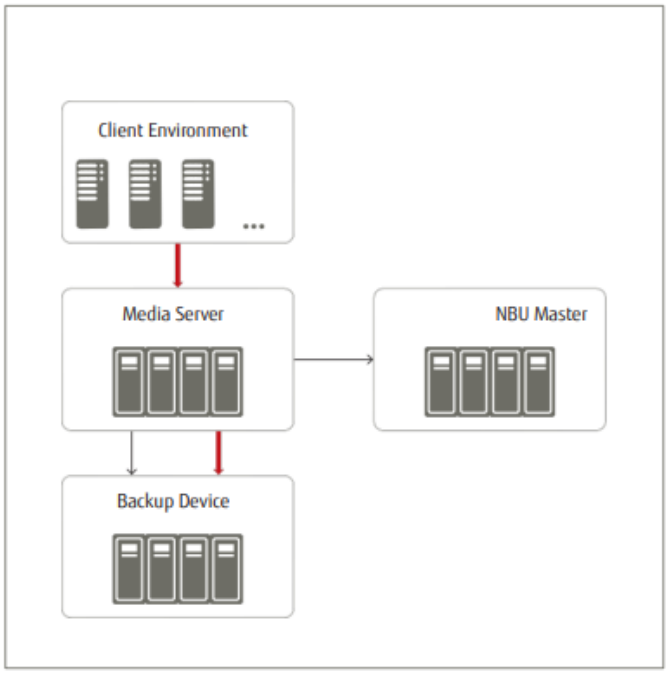


Figure A

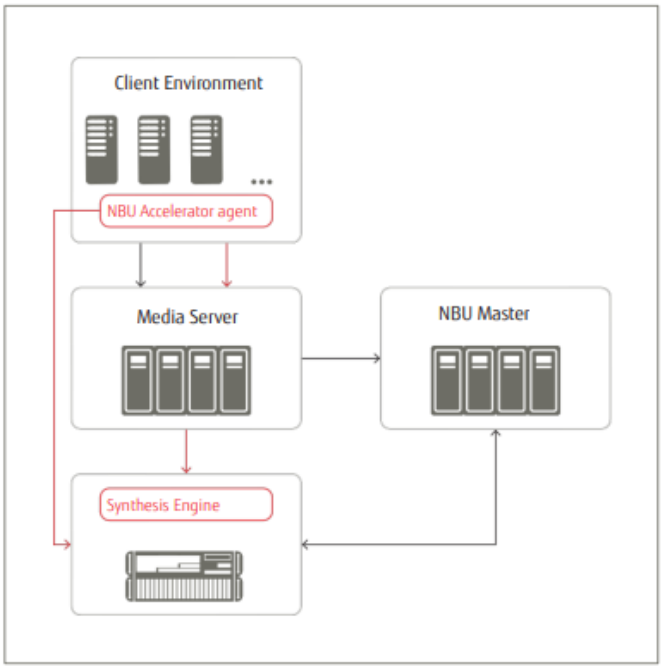


Figure B

Scalability and models

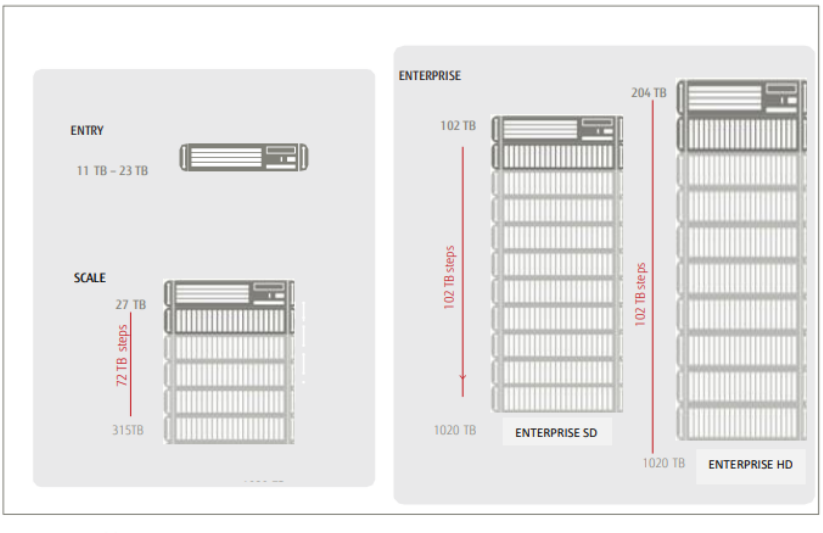
ETERNUS CS800 offers outstanding scalability and makes sure that no investment in unused storage capacity has to be made in advance. Based on three models – ENTRY, SCALE and ENTERPRISE – the appliance makes it easy to start with just the capacity that fits to the current backup environment. If needed, more storage capacity is easily added with- out additional software license costs.

- The core component of the ETERNUS CS800 ENTRY model is based on a rack mountable Industry Standard Server. ENTRY model starts with 11 TB usable capacity and can be upgraded to 23 TB with 4TB incremental upgrades.
- The core components of the ETERNUS CS800 SCALE model are based on a rack mountable industry standard server and up to 4 JBOD storage extensions. ETERNUS CS800 SCALE starts with 27 TB usable capacity without JBOD and can be upgraded in steps of 72 TB until a maximum of 315 TB.
- The core components of the ETERNUS CS800 ENTERPRISE SD model are based on a rack mountable lindustry standard server plus one RBOD storage extension and up to nine EBOD storage extensions, which can be added to the system in steps of 102 TB, starting with 102 TB up to a maximum capacity of 1020 TB.
- The core components of the ETERNUS CS800 ENTERPRISE HD model are based on a rack mountable Industry standard server plus one high-density RBOD storage extension and up to one additional highdensity EBOD storage extension. Capacity upgrades can be added to the RBOD/EBOD in steps of 102 TB, starting with 204 TB up to a maximum capacity of 1020 TB.

ENTERPRISE models also feature dynamic disk pooling (DDP), to avoid high density drawbacks with faster rebuilds and lower performance impact during rebuilds, which means less impact on the backup window.

All models are ready to be configured as unified data protection platform. That means backup and restore data in NAS and SAN environments can be managed within one ETERNUS CS800 and Ethernet and Fibre Channel controllers can be combined flexibly.

These standard building blocks and the high flexibility in network components enable each system to be configured and seamless scaled up according individual needs. Each configured system is manufactured in Fujitsu's OEM partner factories and has to pass strongest quality assurance processes before delivery.



ETERNUS CS800 Models

Administration and reporting

Effective management is critical for obtaining full value from deduplication deployments. ETERNUS CS800 provides an advanced set of management tools for disk appliances and the entire backup infrastructure.

The ETERNUS CS800 system utilizes a Web-based interface which allows configuring and managing the system via Web pages from a remote workstation on the same network. Alternatively a command line interface is also available for advanced configuration and management of an ETERNUS CS800.

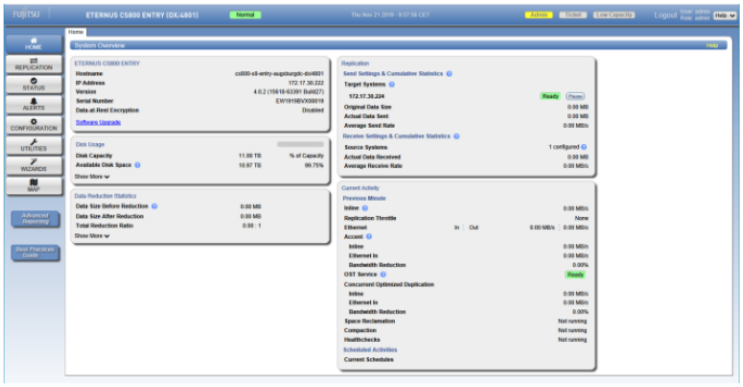
Advanced Reporting is included on every ETERNUS CS800 appliance and combines comprehensive performance data logging with powerful visual reporting and analysis tools to help identifying potential problems and optimizing system operation. As an extension to the ETERNUS CS800 remote management interface it provides an array of performance statistics for the system and a history of change in those data. This helps to identify trends or determine when a problem began. By showing how various operations affect performance,

Advanced reporting also helps to optimize the network ecosystem and business procedures for backup, recovery and replication. Advanced Reporting runs on an ETERNUS CS800 system and continually works in the background to log performance data. To view logged data, graphic reports are available on demand through a web-based interface. Data can be displayed for the current system status or for any time period since data logging began.

Advanced reporting reports provide system and performance statistics, such as:

- Capacity utilization
- Ethernet and fibre channel activity
- CPU load
- System load
- Data deduplication
- Space reclamation activities
- OST and SPEED backups
- Per share information (Chargeback)

Each report includes one or more graphs that show current data but also data over a longer time period to a maximum of six years. This enables comparisons of current and past performance. It also displays the effect of any network configuration or business processes as they affect the ETERNUS CS800 compared to earlier configurations or processes.



Conclusion

Fujitsu Storage ETERNUS CS800 offers an easy-to-deploy and easy-to-use high-performance and scalable platform for effective backup and disaster recovery. The ETERNUS CS800 allows customers to consolidate the whole backup environment as it easily ingests backup streams from multiple sources and achieves the highest possible deduplication ratio.

As a backup target it can be used by all market leading backup applications and in heterogeneous environments. By consolidating backup data storage onto a single scalable and high-performance appliance, higher capacity utilization, more effective deduplication, and easier DR setups can be achieved.

FUJITSU storage ETERNUS CS800 data protection appliance

Technical white paper

For more information on:

www.fujitsu.com/eternus_cs800

Contact

Fujitsu Technology Solutions

Mies-van-der-Rohe-Straße 8

80807 Munich

Germany

www.fujitsu.com/eternus_cs800