



WHITE PAPER

Meeting Backup and Archive Challenges – Today and Tomorrow

Sponsored by: Fujitsu

Nick Sundby
November 2014

IDC OPINION

IDC's end-user surveys show data integrity and availability remains a top priority for customers as they seek to improve business continuity while driving down costs. Backup and recovery and archiving and retrieval are consistently scoring highly as drivers for storage investment, and many companies also want to improve their disaster-recovery plans. These storage workloads pose a sizeable challenge as they stretch available storage and network resources, as well as IT teams. An increasing number of enterprises realize that solving these challenges requires leveraging best practices through the combination of new and also of well proven practices. As current ones cannot sustain service levels in the future.

Survey data indicates that spending growth for near-line disk backup target systems has exceeded that of primary disk arrays, and IDC estimates that companies already spend more on near-line storage overall. IDC's latest end-user survey of enterprises highlights a growing demand for backup, archive, and replication software in 2014, and possibly beyond, indicating increased investment levels into backup and archive.

This will drive demand for disk-based backup target appliances as they demonstrate a huge leap forward in comparison to old practices around tape only-backup. In comparison with legacy solutions, disk-based backup appliances achieve many of the operational and cost reduction objectives that customers face, including:

- A higher and more predictable data ingest rate than tape
- Data deduplication to increase effective storage capacity and allow more backup data to be kept online for faster restores
- Reduced physical footprint compared to tape library systems
- Integrated replication capability for quicker disaster recovery
- Reduced bandwidth requirement, since deduplication reduces the quantity of data sent during replication
- Reduction in tape-related costs including off-site transportation and storage, media costs, and operator handling

Tape, however, will continue to be present as it still offers a cost-effective storage option in the long term, which is decisive for long-term data retention in large scales. IDC research shows that many companies with 50 or more employees are using tape in their environment and only a small

proportion of them are planning to eliminate tape in the near future. IDC believes that tape will be around for the long term, primarily as an archive medium, but in some incidences for large scale backup, vaulting concepts and disaster-recovery operations, including the cloud.

The datacenter will continue to accommodate multiple systems for different functions in the future, leaving multiple management points for the IT team to look after. This overhead will need to be addressed to achieve the next level of storage economics by streamlining backup and archive data management. This will most likely be achieved by solutions that cover multiple if not all backup and archive data management functions on to a single platform.

METHODOLOGY

This IDC White Paper is based on a number of primary research sources, including IDC's *Quarterly Disk Storage Systems Tracker*, IDC's *Storage End-User Survey* and in-depth interviews with various end users.

Definitions

- **Backup** is the activity of copying files or databases so that they will be preserved in case of equipment failure or other catastrophe.
- **Data archiving** is the process of moving data that is no longer actively used to a separate data storage device for long-term retention.
- **Business continuity (BC)** is the activity performed by an organization to ensure that critical business functions will be available to customers, suppliers, regulators, and other entities that must have access to those functions.
- **Disaster recovery (DR)** is the processes, plan and technology needed to recover from an unforeseen incident at a datacenter.
- **Primary disk storage** is the main area in a computer in which data is stored for quick access by the computer's processor.
- **Second-tier storage** is primary storage that is optimized for lower cost and higher storage capacity.

IN THIS WHITE PAPER

This IDC White Paper focuses on the data management challenges of backup and archive data enterprise IT faces today and in the near future.

Backup and archive are particularly critical areas in which customers feel the need for better practices and improved economics for various reasons. These include enterprise data growth, budget cuts, and a more demanding business line.

This paper provides an overview of backup and archive trends, and looks at the main must-have features for a modern, sustainable storage solution. It also analyzes what Fujitsu has to offer to customers looking to enhance their backup and archive storage operations.

SITUATION OVERVIEW

IDC's storage end-user surveys show that backup and recovery, business continuity, archiving, and enhancements to disaster-recovery plans are all driving spending on storage solutions. This year, nearly half of the large enterprises interviewed mentioned backup and disaster recovery as one of the top three storage functions in which they had plans to invest, with over a third of medium-sized companies indicating the same. Archive was identified as an investment area by over a third of companies with over 1,000 employees.

Across almost all verticals, archiving initiatives, regulatory retention mandates, and compliance audits were major drivers for archival storage investments in 2014. eDiscovery needs and data mining requirements for generating operational and economic benefits continued to extend the scope of archival storage and software.

Archiving is not yet a fully established storage task on its own within many customer organizations. In a study conducted last year, only a fifth of customers claimed to use a dedicated solution for archiving, and almost a third of them simply used tape backups for long-term data retention. Half of the respondents simply kept data on disks. More sophisticated approaches to archiving, such as hierarchical tiered archive management, were minority cases.

At the same time, companies are looking to improve the economics of their storage operations, which is particularly pressing for larger enterprises, with four out of ten setting this as a priority. Many are also directly focused on reducing the number of systems and trying to ease management burden, while protecting virtual servers and ensuring data retention. Compliance also gained high response rates of over 20%.

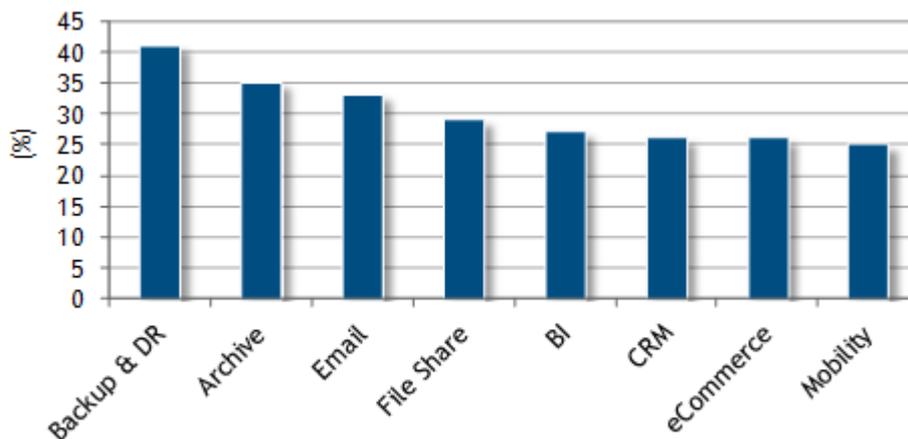
Not counting the cost of software licenses for backup and archive management, IDC data suggests that spending on backup and archive storage systems has exceeded that of primary storage. On average, the total volume of data stored grows at an annual rate of over 30%. This puts enterprises under enormous pressure to improve their existing storage management practices, which includes higher utilization of capacity, deployment of data reduction technologies, and increased operational efficiency. In most cases, this requires initial investments in new solutions that enable these best practices.

These findings confirm that while economic uncertainty suppresses growth and hampers overall storage market momentum, companies keep investing in their storage infrastructure at an increased level to achieve various objectives, including capacity and performance increases, enhanced backup and archive, and cost optimization.

FIGURE 1

Application Importance With Storage Investments

Q: Which applications are most important to you when considering future storage investments?



Note: Company size 1000 employees or higher; top 8 answers

Source: IDC, 2014

For most companies, management of backup and archive data may be the most complex IT task, accounting for a large portion of work-hours, software licenses, and professional services. Fragmented infrastructures and storage silos lead to higher management effort and difficulties in maintaining different/various SLAs in changing situations. As a result, storage carries a heavy bill associated with operation, and acquisition costs typically represent only a fraction of the total cost of ownership for storage systems over their lifetime. Empirical studies show that acquisition costs are dwarfed by operational expenditures over the course of three to five years.

As a result, reducing storage-related costs is among the top objectives for companies as they feel the pain of bad storage economics and want IT to take its fair share of cost cutting. This objective was expected to be reached by any new storage solution for a third of the companies surveyed, and 40% for large enterprises. IDC research shows that many companies have targeted legacy backup and archive infrastructure as a significant cause of operational burden, complexity, and expense. Mergers and acquisitions, autonomous IT departments and multiple office locations can all result in backup processes becoming complex and costly to administrate. When IT budget or staff resources are squeezed, a tipping point is reached and a new approach is required.

Archive Can Offset the Impact of Data Growth

Archive is now one of the most important applications driving investment in storage infrastructure. In the past, archive was often seen as "nice to have" but not essential. Many companies' archiving was often simply backup tapes that were moved to an off-site location, with little or no indexing or search capability. IDC research shows that investment in archive can have a rapid return, because of two factors. Firstly, the company benefits from converting aged data into a searchable, active resource. For example, email or files associated with a particular employee or project can be easily

retrieved. Secondly, the company can move unused data from the fast online storage into the archive. This saves money by mitigating rapid data growth and reducing the need for capacity upgrades. Storage management is simplified, significantly reducing management costs.

Figure 1 highlights the importance of improved disaster recovery in the respondent's investment plans. When IDC interviews end users, it is not uncommon to hear IT managers of significant companies admit that they have little or no real DR capability. For these companies, investment in an automated backup target with asynchronous replication to a second site is a cost-effective and efficient solution to improving business continuity of the datacenter.

The IT Efficiency Mantra

Many users talk to IDC about their storage challenges and the various initiatives they have taken to reduce cost or align better with the business units. As a general observation, the most successful projects tend to include the following three elements in some form.

- **Consolidate** – reduce the number of platforms or locations where possible
- **Standardize** – use a common set of hardware and software elements
- **Automate** – relieve IT staff of low-level manual tasks

A project that includes one or two of the above can certainly deliver benefits, but if all three can be combined, there is a chance for a step-change improvement that can benefit the company at a long-term strategic level.

The need is particularly acute in backup, archive, and replication, as many companies choose point solutions for each task that may be best of breed, but are managed as discrete functions. Many companies have a unique backup process for the datacenter, the remote offices, laptops, and key database platforms, further adding to complexity and cost. In IDC's view, an integrated approach to backup, archive, DR, and replication is highly desirable.

The Solution Landscape

The highlighted customer needs drive demand for solutions that can dramatically improve the economics of their backup and archive operation, including more labor efficient backup and recovery practices, fewer management points, and reduced long-term expenditures on capacity purchases. Many customers are still struggling with scattered tape-based backup infrastructures that require laborious field support and struggle to keep up with backup windows.

This has opened up the market for purpose-built backup appliances (PBBAs), a fast-growing segment of the storage market. PBBAs are specifically designed to serve as a target for backup and recovery. The disk-based solutions offer high ingest and recovery speeds, scalability, and the ability to integrate with existing backup software. Furthermore, most also offer integrated and bandwidth optimized remote replication capabilities for easy disaster-recovery planning.

Functions commonly supported by PBBAs include:

- **Deduplication of data to maximize the effective capacity and allow more data to be held online for faster recovery.** Deduplication typically leads to a reduction in backup storage requirement of 10-20x, depending on the mix of data types. Deduplication is effective for backup as subsequent backup typically carries very little change (new unique data)

relative to the size of the full data set. Instead of physically storing individual backups, PBBAs minimize the footprint by only storing changed blocks to the previous backups.

- **Virtual tape library (VTL) functionality.** A VTL emulates the functions and characteristics of a physical tape library. It can integrate closely with the backup software, while adding additional management and backup and archive functions of its own. Since it can directly replace a tape library, it can utilize existing scripts and processes while offering significantly higher levels of performance. Some VTLs can be connected to a tape library and will manage the movement of data between the two devices either transparently or application-aware. High-end VTLs now offer some of the most sophisticated backup management capabilities for large enterprises, and are often used to consolidate a range of backup infrastructure onto a single platform – if they come with the sufficient capabilities and flexibility in capacity, performance, and functionality.
- **Backup to the cloud.** Adoption of cloud-based services is rising as companies seek to reap the benefits of flexible, lower cost IT services consumed on demand. Although backup and recovery services have been available via the cloud for some time, concerns about the security and reliability of services – and data residency issues – have served as a barrier to adoption for some organizations.
- **Mainframe connectivity.** A PBBA connected to both mainframe and open systems devices can become a consolidated backup platform serving the entire enterprise. This approach may be used to replace fragmented or dispersed backup/archive infrastructure with a single efficient platform using a common set of management practices and policies. Mainframe connectivity is typically restricted only to the most specialized high-end solutions on the market.
- **Replication.** Many PBBAs offer appliance-based replication capability, so that backup data is synchronously or asynchronously replicated to a remote location. This can be used between remote offices and a central datacenter, or between datacenters as a cost-effective way to enhance business continuity and is used by the majority of PBBA users.

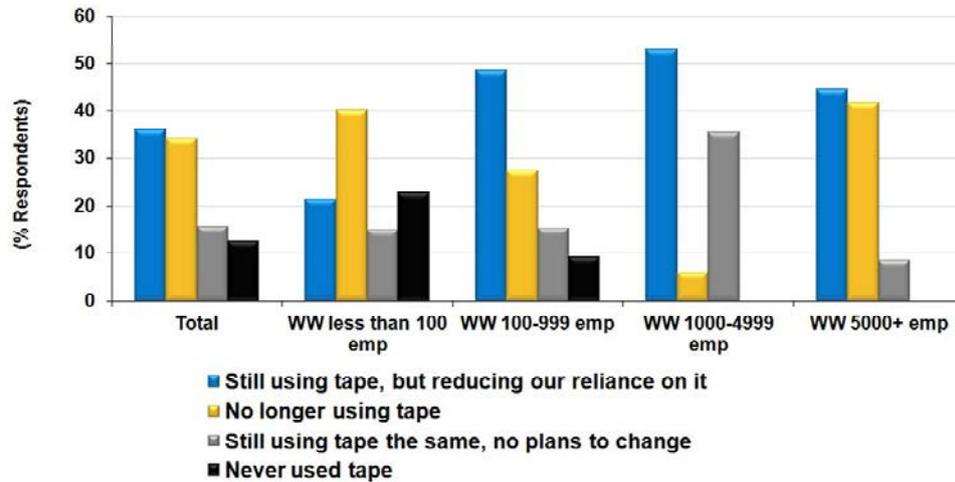
The strong combination of benefits has driven rapid growth in PBBA adoption. In 2011, the worldwide PBBA market value exceeded that of the worldwide branded tape market, which includes tape drives and automation products. At the close of 2012, the aggregate open system and mainframe worldwide PBBA market increased by 12.1%, totaling over \$2.75 billion. IDC expects total worldwide PBBA revenue will increase by a CAGR of 19.4% through 2016, totaling nearly \$5.9 billion.

This trend will continue for the foreseeable future as customers take advantage of the rich feature sets offered in PBBAs, particularly for virtual server protection, rapid recovery, and data deduplication.

FIGURE 2

Impact of Disk-Based Backups

Q: What is the impact of disk-based backups on the use of tape in your environment?



Source: IDC, 2013

Tape strongly complements disk-based targets as it offers long-term storage at the lowest cost per GB, and the removable media is ideal for secure offsite storage. Tape backup bottlenecks are now unlikely, as throughput performance has risen dramatically. IDC's end-user survey also shows that 49% of the companies interviewed with 1,000 or more employees will continue their usage of tape, but will reduce their reliance upon it in favor of disk-based alternatives.

Tape usage is therefore expected to continue, as many companies will use tape to protect and retain their data for disaster recovery, protection of IP, and regulatory compliance – and increasingly with the objective of offloading data from more expensive primary and backup disk systems. It's a common experience to find that only 10%-20% of a company's data is in active use, while the rest is accessed rarely or never at all. If the "cold" data is migrated off the primary storage and into the archive, the active data set is smaller and therefore easier to back up and faster to recover, provision, and manage. Data growth is also mitigated, saving money in capacity upgrades for the primary storage arrays.

Most modern PBBAs and archiving appliances, while more efficient than legacy approaches, are still point solutions and enterprises need to buy, deploy, and manage multiple devices for open networked, mainframe networked backup targets and archiving disk and tape platforms. Some of the best-selling solutions do not support mainframes or tape vaulting, limiting the use cases for the appliance and requiring additional devices for mainframe backups and archiving tasks.

Not only will this represent increased capital expenditures and multiple systems to integrate and manage, but it is likely to come from multiple vendor and management interfaces that are dissimilar. This increases the management overhead, especially when it comes to change management.

Deduplication – Strengths and Challenges

Deduplication is a function rather than a solution, so the benefits seen by the user will depend on the use case and environmental context. However, the most commonly seen strengths and challenges can be identified as follows.

Strengths

- A 10-25x increase in effective storage capacity that allows more backup data to be kept online for faster restores. Virtual server image files often contain large amounts of redundant data and may allow 50x reduction or more
- Reduced physical footprint compared to tape library systems
- Integrated replication capability for quicker disaster recovery. Reduced WAN bandwidth requirement, since deduplication reduces the quantity of data sent during replication
- Reduction in tape-related costs including off-site transportation and storage, tape hardware upgrades, media costs, service contracts, and operator handling

Challenges

- Generally PBBAs exhibit a restore throughput performance that is lower than the ingest rate, extending the time for major restore operations
- Data that is compressed (video, audio, imagery, PDF files) or encrypted cannot normally be de-duplicated. Since there is a cost associated with deduplication, users with a significant portion of such data will gain little benefit and costs will be increased
- PBBAs can be a single point of failure within the backup infrastructure
- PBBAs generally require internal 'housekeeping' time in order to reclaim released capacity and to defragment the data. This function runs on a daily or weekly basis and may take several hours to complete. During this time, backups are either not possible or run at a slower rate. This can become an issue in a large environment (say 100TB or more) where the PBBAs are heavily used. The "housekeeping" process means that the unit cannot be treated as a 24 x 7, continuously available device. IDC has conducted several cost-benefit studies to identify and quantify the business value of deduplication. In general, the user benefits far outweigh the challenges, and users typically see a financial return after 5-8 months of use. The largest savings typically come from lower operational overheads, productivity savings for IT users, elimination of tape hardware/service/media investment and reduced exposure to outages due to improved disaster recovery capability.

FUJITSU'S BACKUP AND ARCHIVE STRATEGY

Fujitsu has taken a different approach to most in the industry. Instead of just trying to solve one problem, the company has developed a long-term vision for backup and archive. This not only embraces backup and recovery but also archive to create a single platform for all backup and archive needs, be it mainframe, Unix, or x86 environments. Also, it provides an affordable storage repository for less-used files. In the spirit of this concept, Fujitsu chose to integrate advanced disaster-recovery capabilities across all functions as well as tape vaulting support into the hierarchical storage management.

Fujitsu has chosen to create a service-level-oriented view of the data-protection operation supported by underlying automation of tasks. This includes the integrated management of disk,

deduplication and tape, automated migration of data across media types, and systems based on policies and features and functions for high data availability.

Automated storage tiering delivers financial benefits and is commonly used on primary storage arrays. Fujitsu believes that backup and archive can be approached in a similar way, with a range of backup and archive service levels depending on the RTO/RPO objectives for the data in question. Fujitsu's vision is that multiple automated service levels should be available from a single unified backup and archive platform. Users should be able to choose the backup service level that suits the data or workload, so that resources are allocated efficiently. This vision also aligns with a self-service internal cloud, in which users can choose the backup service from a service catalog with known costs and service levels.

Rather than position a single storage system against a competing storage system, Fujitsu positions a consolidated solution against point solutions. This strategy allows Fujitsu to integrate with existing environments and improve operations even for datacenters with the latest PBBAs and tape libraries. The picture is completed by branch-office PBBAs for centralization of backup and archive and standard workflows.

The Fujitsu appliances are based on standard components such as robust and high-performance Fujitsu servers and storage subsystems glued and complemented by a respective software stack. Thanks to standardization, these appliances allow easy and step-by-step integration in existing environments, and the ready-to-use solutions enable easy administration and management.

FUJITSU ETERNUS CS PORTFOLIO OVERVIEW

Fujitsu's vision is seen in three product lines, the ETERNUS CS8000 series (formerly known as CentricStor and ETERNUS CS High End), the ETERNUS CS800 and the ETERNUS CS200c. With the ETERNUS CS8000, Fujitsu has created a highly scalable, high-performance, and unified backup and archive platform for midsized to enterprise datacenters. For SMB environments, Fujitsu offers the ETERNUS CS800, a purpose-built deduplicating backup appliance with a remote/branch-office concept. ETERNUS CS200c integrates correctly-sized hardware, backup software, and licenses into a complete backup solution enabling small and midsized environments a fast and hassle-free setup of a comprehensive backup environment.

Fujitsu ETERNUS CS8000 Overview

The ETERNUS CS8000 fulfils Fujitsu's vision of providing customers with a singular management point for all backup and archive needs. ETERNUS CS8000 is an appliance that acts as a backup target for hosts and media servers, but unlike other targets it provides considerable flexibility about how the data is handled.

The first step is that it caches incoming data. Depending on the chosen service level the data will stay on disk only, deduplication only, or it will be moved to tape, or it will stay on disk and on tape. It creates a logical representation of a backup target and acts as a virtual layer between hosts and external disk and tape systems. For archive, the disk storage is the primary storage, and automated HSM is also available to move the archive data to tape.

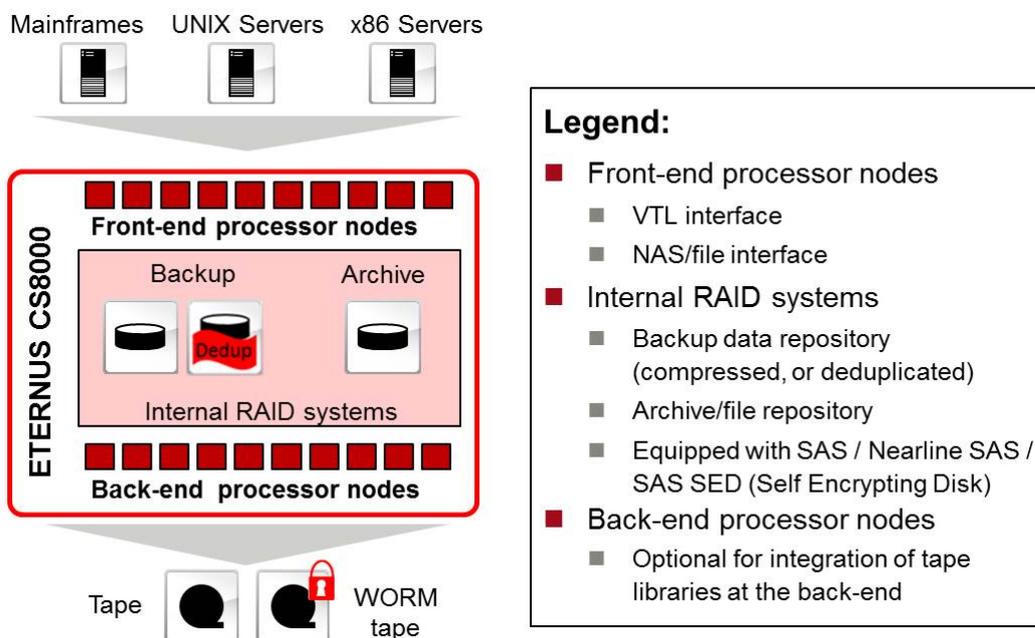
This results in a single platform for all backup and archiving jobs to commit against, simplifying the configuration of the network, hosts, and storage. The ETERNUS CS8000 provides open networked

and mainframe VTL interfaces, as well as CIFS/NFS host connectivity, while supporting a mix of internal disk and deduplication volumes, and up to ten tape libraries in the back-end, managing data volumes into the Exabyte magnitude. Via the NAS/file interface up to 200x500 million files are supported.

The grid-based system architecture is highly scalable and modular, offering dedicated front-end resources facing the hosts, and a dedicated back-end for managing data flow between the appliance and the storage targets. Front-end or back-end nodes can be added to increase the performance as required, up to a total of 20. The disk cache, managed as internal storage by the ETERNUS CS8000, can be scaled up to 22.2 Petabytes raw capacity on up to 16 RAID systems.

FIGURE 3

ETERNUS CS8000



Source: Fujitsu, 2014

Many solutions offer "scale up" (to grow in capacity), but ETERNUS CS8000 also offers flexibility for "scale out." The grid of processor nodes can be enlarged or reconfigured to meet future performance demands, avoiding the need for an over-specified, under-utilized system. This is a very important part of ETERNUS CS8000, and a big benefit not only for large enterprise customers but also for midsize environments. ETERNUS CS8000 works with the same architecture, independent of the level of capacity, performance, and SLA (replication, HA, disk/tape combination, internal management of several copies with dual/triple/multiple safe) being used.

This scale-out characteristic differentiates ETERNUS CS8000 from other point solutions and can be used in different environments. For example, an entry ETERNUS CS8000 configuration starting with 7 TB internal disk storage with only backup functions can be flexibly expanded to the full capabilities of the largest ETERNUS CS8000 configuration with full capacity (22.2 Petabyte internal disk, managing data on the backup-end in the exabyte magnitude) and performance (up to 150 TB/hour and beyond) as well as full functions for backup, archive, replication, and so on. The

system can be scaled in throughput or capacity as needed; the user buying only what is required for the current needs.

Additionally ETERNUS CS8000 manages multiple copy images on tape targets on the back end, also with asynchronous replication (cascading) to other systems thousands of kilometers away. There is no effort from the backup application necessary. This also reduces the management efforts on the SW side. The backup software only maintains a single backup image – ETERNUS CS8000 ensures that there is the right level of availability within the backup storage, with multiple copies of this image. And in the case that one copy is not available, the self-healing functionalities of the ETERNUS CS8000 ensure that the copy will be recreated.

This allows the ETERNUS CS8000 to offer uncompromised ingest speeds and the option for near-instantaneous recovery, as well as centralized disaster-recovery operations. All functions are policy based SLA-driven and the freedom of using disk and tape targets of various types allows customers to balance protection level and speed for cost. For example, some hosts may require both fast recovery times and high levels of backup performance.

Transparent Split-Site Capability

The core element of the most disaster resilient architecture is one logical ETERNUS CS8000 system which is spread over two geographically separated sites, the so called "split-site configuration" with cache mirror. The internal infrastructure is thereby extended to a second site which can be 100 km away. Still the connected applications from the production environment see one logical setup without being aware of the geographical location. The advantage is obvious: although distributed across two sites, there is still only one single system. Data can be written to or read from both sites.

The even more valuable benefit appears when it comes to a failure at one site, for example in the case of a disaster. Traditional systems require a classical failover which is difficult to set up. Even more difficult and error-prone are the manual actions that have to be done in a stressful situation, perhaps while sprinklers are running.

The split-site configuration of ETERNUS CS8000 provides the solution. If one site goes down, the workload is automatically routed to the other site. The applications are not aware of this process. Backup and archive processes continue as before without any action necessary by the staff. This is called automated continuation. It dramatically increases the availability of data and thereby secures business processes – all of it at the same time with a dramatic reduction of administration costs.

Automated Tape Management and Migration

To reduce complex, time-consuming and error prone tape management, ETERNUS CS8000 takes over the management of all target systems, providing a highly available virtualized tape robot. The redundant architecture enables reception of the backup data, even if a component should fail. For example, if the physical tape fails, then the data will automatically be directed to another tape drive or tape library to ensure that the data is safe.

The tape reorganization function maximizes utilization of the capacity available. Furthermore, regular reorganization of tape media refreshes the magnetization. The appliance automatically makes quality checks of duplicate copies. If an error occurs, the self-healing functionality recovers

the inaccessible information from the other copy to ensure the availability of data even over the very long-term. According to Fujitsu, average utilization of tape capacity is around 90% using ETERNUS.CS8000 management, compared to an average 50% usage with enterprise backup applications. The result is reduced requirement for tape media and floor space.

Self-healing also plays an important role if a copy is requested for a restore case and not available for some reason. The recall request is automatically redirected to the second copy, the data is provided and the job for the backup software is done. In addition the automated self-healing recovers the failed copy. Thus the level of redundancy is restored.

Tape libraries are decoupled from the backup software. Therefore ETERNUS CS8000 more effectively utilizes the streaming performance of tape media. A wide range of libraries and tape drives is supported. Even if the libraries and drives are different, they can be connected in parallel. This makes it very easy to introduce a new generation of tape libraries or tape drives, since ETERNUS CS8000 takes care of the data migration from the old library to the new one.

Improved Deduplication

The ETERNUS CS8000 has improved deduplication capability compared to the previous generation. Deduplication is fully integrated into the architecture so that chunking, hashing, and compression functions are handled as in-line processes by the front-end processors. Ingest performance can be increased as required by adding additional nodes. This delivers greater performance, resilience and scalability than the post-process approach of the previous generation. Parallel backups are supported, so that deduplicated disk and standard tape backups can be handled concurrently. Due to its architecture and flexibility, the constitutional limitation of deduplication appliances, the housekeeping process, can be avoided and the platform allows fully available 24 x 7 operations.

Multitenancy is supported on several levels. For different users dedicated virtual tape drives, or dedicated front-end processors can be deployed. Furthermore, the data may be stored on specific volume groups with dedicated user access only. For the highest level of user separation up to ten independent deduplication services with an individual deduplication storage pool are supported, all managed within one ETERNUS CS8000 system.

Automated Tape Creation

The ETERNUS CS8000 allows a single stream performance beyond 270MB/s between ETERNUS CS8000 back-end and the tape library. Fujitsu claims that in typical environments without ETERNUS CS8000 just 20% of tape drive performance is used on average.

The optimized streaming can reduce the numbers of tape drives required and related service costs, and reduce the SAN ports required.

Archive Capability

The appliance can also be used for archiving by flexibly adding the related components and functions. The archiving function is also flexible and highly automated, and integrates well with archiving software solutions. It guarantees write protection with WORM functionality, automated migration of data for cost optimization, and the creation of redundant copies for long-term data integrity.

This appliance is already certified for the storing and retrieval of electronic documents based on the principles of proper accounting (<http://www.kpmg.de/bescheinigungen/requestreport.aspx?28388>).

NAS Capability

Not only can the appliance act as a target for the archiving software, but the ETERNUS CS8000 can also act as an online NAS repository for files that are rarely accessed, thus saving valuable capacity on higher-performance online storage systems. It's common to find less than 20% of the data on primary storage being actively used. If the older or unused data is moved to a more cost-effective storage tier, then more capacity is freed up for the production systems and backups and backup/replication can be managed more easily. Later, data can be automatically migrated onto tape for further cost savings while still representing them as accessible for end users.

Protecting the Archive Data

The archive software identifies inactive files using the archive policy and moves them to the archive storage. Since the files are no longer in the production environment, they are no longer protected by the standard backup process. Therefore it is important to protect the archive data, which may need to be restored in a full DR recovery.

The ETERNUS CS8000 offers a number of ways to protect archive data.

- Archive data can be written to a local tape library and also to a remote tape library over a long distance. This provides an easy way to introduce disaster resiliency with a remote copy on tape, which will be available even if the complete datacenter is affected by a catastrophic event like fire or flooding.
- Asynchronous replication to a remote ETERNUS CS8000 system is supported. No extra replication software is necessary. Data protection can be implemented especially for long distance replication in the range of thousands of kilometers.

The replication functionalities mentioned above are a disaster recovery copy, and not a real backup. The replicated copy helps to recover data if the primary storage system is gone, for example if the datacenter is affected by a catastrophic event like fire or flooding. But the disaster copy only allows recovery of the latest status. For example, if the user deletes data accidentally, than this deletion process is also replicated to the disaster copy. There is no option to go back to a previous version. To provide a true backup of the archive an integrated and automated snapshot-based backup solution for the ETERNUS CS8000 NAS subsystem is provided.

The backup copies can be replicated. For example, the most recent backup may be stored at the local system, while the complete set of required backups like all daily backups of the last week can be stored at the second system.

The NAS functionality is also based on the ETERNUS CS8000 scale-out architecture using a distributed file system to handle up to billions of files. Fujitsu's ETERNUS CS8000 therefore consolidates backup, archiving, disaster-recovery, and second-tier storage functions onto a single platform.

The consolidation also leads to higher utilization of backup and archive storage systems such as tape libraries; i.e., investment protection through optimized usage of existing infrastructures. On the other hand, the tape handling is decoupled from the backup applications and managed by

ETERNUS CS8000, so it is easy to introduce new tape libraries or tape technologies without changing anything at the backup application level.

Cloud Integration

Another option which may play a more important role in the future is the option to store backup data in the cloud. The adoption of cloud-based services is rising as companies seek to reap the benefits of flexible, lower-cost IT services consumed on demand. Although backup and recovery services have been available via the cloud for some time, concerns about the security and reliability of services – and data residency issues – have served as a barrier to adoption for some organizations. As a future enhancement, the ETERNUS CS8000 will be capable of using the cloud as an additional target via a cloud gateway. This gateway will allow customers to store data optionally into the cloud.

TABLE 1

Fujitsu ETERNUS CS8000: Key Functions and Associated Business Value

Fujitsu CS8000: Key Functional Characteristic	Description	Business Value Benefit
Unified backup and archive platform	<p>Open systems and mainframe connectivity (parallel support of: Windows, SUSE, AIX, Linux, Red Hat Linux, Solaris, IBM z/OS, System i, Fujitsu BS2000/OSD).</p> <p>One appliance for backup and archiving purposes.</p>	<p>Allows consolidation of fragmented and disparate backup and archive infrastructure into a coherent platform with consistent management processes, saving operational costs.</p> <p>Increases utilization of target disk and deduplication disk arrays and tape libraries.</p>
Modular grid architecture	<p>Scale-out architecture based on modular building blocks allows the adding of performance (processor nodes) and capacity (storage arrays) without disrupting existing workloads.</p> <p>16 Gb FC / 10 GbE / FICON 8 host interface support.</p>	<p>Investment protection.</p> <p>Platform can start small and grow over time, reducing over-provisioning.</p> <p>Internal storage can grow from 7TB to over 22.2PB (physical capacity before compression and deduplication).</p> <p>Increase performance with the same system up to 150 TB/hour and beyond (no need to invest in a second system and buy the related capacity just for performance reasons).</p> <p>Existing storage assets can remain part of the storage pool, extending their useful life.</p> <p>Array investments can typically be depreciated over five years rather than three.</p> <p>Speed the ability to deploy new backup and archive capabilities to enable key business initiatives or to support changing workload profiles.</p>

TABLE 1

Fujitsu ETERNUS CS8000: Key Functions and Associated Business Value

Fujitsu CS8000: Key Functional Characteristic	Description	Business Value Benefit
Simplified and integrated management of disk, deduplication and tape technology	<p>Completely virtualized storage separates management of storage services from hardware configuration.</p> <p>Automated ILM (Information Lifecycle Management) between disk, and tape media.</p>	<p>Flexibility in using the optimized combination of disk, deduplication and tape technology.</p> <p>Reduced complexity of IT operations and lower cost of operations by reducing the need for specialized expertise and minimizing time-consuming, error-prone manual tasks.</p> <p>Staff time is freed up to allow increased focus on understanding and supporting business requirements.</p> <p>Reduced management overhead by automatic relocation of data to most cost-effective storage tier while cost/performance is continually optimized.</p> <p>This results in more consistent performance in order to meet SLA requirements. IDC research shows that manually re-positioning data to address performance problems can take days to execute since data must be selected, moved and then monitored to test that the problem has been solved. Automation of the process allows the system to respond correctly to surges in demand without operator intervention.</p>
Automated availability and disaster recovery functionalities	<p>Automated copy management (several copies, local and remote, on disk and tape) incl. self-testing and self-healing.</p> <p>Split system with cache mirror to a local or remote system (still representing one single system for automated continuation).</p>	<p>Easy use of automated high availability and disaster recovery options for backup and archive data.</p> <p>Automation eliminates human errors in critical situations.</p> <p>Ensures an always available copy, even on the very long term.</p> <p>Continuous operation, even in the case of whole-site outage – no classic failover necessary.</p> <p>No single point of failure platform, allowing 24 x 7 operations.</p>
Protection of archive data for DR and backup	<p>Archive data can be replicated to remote sites for DR</p> <p>Archive data can be protected using an integrated snapshot-based backup capability</p>	<p>Inactive data that is stored in the archive usually falls outside the standard backup process for the production environment. Replicating the archive is good for DR, but is not the same as a backup. For example, if a user deletes a file in the archive tier, the deletion will also be replicated to the DR site.</p> <p>ETERNUS CS8000 has an integrated snapshot-based backup of the archive data, allowing full version control of archive data.</p> <p>The archive tier then therefore mitigates data growth on the primary storage, without compromising on data protection or resilience.</p>

TABLE 1

Fujitsu ETERNUS CS8000: Key Functions and Associated Business Value

Fujitsu CS8000: Key Functional Characteristic	Description	Business Value Benefit
Leveraging backup system for long-term storage, as a second-tier NAS storage for inactive file data and for compliant archiving	<p>Flexible connectivity (for SAN, VTL, TCP/IP & NAS) and support of all major backup and archive applications.</p> <p>Parallel support of all major tape libraries with tape automation functions, tape refresh, self-healing and tape migration.</p>	<p>Viable long-term storage capability, reducing the cost of meeting compliance and data governance obligations.</p> <p>Allows mitigation of the cost of data growth by removing inactive data from expensive primary storage to ETERNUS CS8000, reduces the burden of managing and protecting active data and allows backups and recovery jobs to complete within the available time window.</p> <p>Automatic migration between legacy and new generations of tape drive, with minimal operator intervention.</p> <p>Single platform for backup, archive, and secondary storage purposes to consolidate backup and archive infrastructure, utilize storage capacity of disk and tape systems and reduce management efforts.</p>

FUTURE OUTLOOK

Disk-based backup has enabled many companies to reinvent their backup and disaster-recovery capabilities, so that files are recovered in minutes rather than hours, and rapid failover is possible in the event of a major outage. However, data volumes continue to grow exponentially and IT staff resources remain under constant pressure.

So what does the future hold for backup and archive? In IDC's view, most companies no longer have the time or resources to spend managing backup processes. As far as possible, backup should be an automatic and self-managing process that requires little or no manual intervention. It must happen without disruption to the application workloads, and recovery performance (for single files or complete systems) must comfortably exceed SLA requirements.

Archive has an important role to play in this story. By migrating cold data from the primary storage into an archive, data growth issues are mitigated and backup is more manageable. However, many companies still have a confused view of archive, often using old backups as a substitute. Arguably, current archive solutions are not delivering a sufficiently compelling proposition to the market.

Part of the problem may be that backup, archive, and replication are often implemented as standalone functions. Yet efficient IT is often based on consolidation, standardization, and automation. If backup, archive, and replication can be brought together into a single platform, the management process would be simplified.

Achieving an integrated, automated, scalable, and efficient backup infrastructure is a challenge, as most solutions are built from a collection of best-of-breed but disparate components. Fujitsu's approach with its ETERNUS CS8000 allows the consolidation of multiple systems onto a single automated disk and tape backup and archive platform, and this is a significant step toward meeting this goal.

Cloud storage providers are increasingly used by large companies as backup and archive targets, so the cloud gateway function of the ETERNUS CS8000 is a useful capability. Users should consider how they will recover large quantities of data rapidly. Some cloud providers will ship the array to the customer site, or will host the user workloads until the customer site is back online. As cloud providers refine their offerings and provide better business value justification, use of the cloud storage tier as a complement to on-premises storage will become increasingly mainstream.

CHALLENGES/OPPORTUNITIES

IDC research identifies huge cost pressure on storage operations as economies stagnate. Customers have been growing increasingly conscious of the hard and soft costs associated with data storage and management, including an out of proportion labor overhead, a complex and expensive software licensing portfolio, and technical maintenance for legacy systems. This creates demand for solutions that improve the economics of the operation while solving capacity, performance, and backup and archive issues. Much of the capacity needs are driven by the accumulation of old and copied data, for few organizations have data life-cycle and retention mechanisms in place that would be intelligent enough to delete redundant or obsolete data safely. IDC estimates that spending on backup and archive storage has already surpassed primary this year.

Current practices around data backup, disaster recovery, and archive are a challenge for many organizations due to a lack of adequate strategic planning and implementation. Transformational projects are typically triggered by unsustainable operations and breach of SLAs, IDC observes. Also, service providers need to hit very low cost levels in order to make their business viable, and preferably in predictable, pay-as-you-grow financial models.

CONCLUSION

Fujitsu's vision of a unified backup and archive platform has been in place since the advent of the ETERNUS CS in 1999. The ETERNUS CS products have undergone a continuous and sustained development process over 15 years, often based on feedback from the active and loyal user community. IDC research shows that IT managers in large companies are increasingly looking to converge their backup, archive and replication infrastructure. In our view, Fujitsu's strategy was ahead of its time in 1999 but is now aligned closely with backup and archive objectives in the enterprise.

Fujitsu's ETERNUS CS8000 offers enterprise customers some highly beneficial features. The unified backup and archive data management platform across open and mainframe networks, backup and archive functions, and the simultaneous and automated support for disk, dedup disk and tape media are highly differentiated against the competition, most of which offer a set of different products to address all these needs. Moreover, not only can it take over the management of backup and archive and centralize disaster recovery, but it allows customers to keep using any

of their existing tape systems, as ETERNUS CS8000 can act as a front end for those secondary storage systems. Investments into tape libraries are protected and the return improved. IDC believes that the combination of these features and benefits is unique.

The ETERNUS CS800 completes this picture as a smart PBBA for SMB environments and branch offices. The appliance offers an easy to deploy and use high-performance and scalable platform for effective backup, archive, and disaster recovery. For small and medium-sized organizations, the appliance can deploy in multiple locations for bidirectional replication. This enables Fujitsu to tap into the fast-growing PBBA midmarket.

The ETERNUS CS portfolio now also comprises an integrated backup appliance, the ETERNUS CS200c. The appliance enables a fast and straightforward setup of a comprehensive backup and archiving environment. Leading software technology from CommVault Simpana is highly integrated with Fujitsu system technology in order to deliver the right performance for the selected capacity range. Backup capacity can be expanded during the lifecycle, thus supporting data growth and delivering investment protection. Each capacity extension is fully licensed, making commercial procedures simple.

About IDC

International Data Corporation (IDC) is the premier global provider of market intelligence, advisory services, and events for the information technology, telecommunications and consumer technology markets. IDC helps IT professionals, business executives, and the investment community make fact-based decisions on technology purchases and business strategy. More than 1,100 IDC analysts provide global, regional, and local expertise on technology and industry opportunities and trends in over 110 countries worldwide. For 50 years, IDC has provided strategic insights to help our clients achieve their key business objectives. IDC is a subsidiary of IDG, the world's leading technology media, research, and events company.

IDC U.K.

Chiswick Tower
389 Chiswick High Road
London W4 4AE, United Kingdom
44.208.987.7100
Twitter: @IDC
idc-insights-community.com
www.idc.com

Copyright Notice

This IDC research document was published as part of an IDC continuous intelligence service, providing written research, analyst interactions, telebriefings, and conferences. Visit www.idc.com to learn more about IDC subscription and consulting services. To view a list of IDC offices worldwide, visit www.idc.com/offices. Please contact the IDC Hotline at 800.343.4952, ext. 7988 (or +1.508.988.7988) or sales@idc.com for information on applying the price of this document toward the purchase of an IDC service or for information on additional copies or Web rights.

Copyright 2014 IDC. Reproduction is forbidden unless authorized. All rights reserved.

