Enterprise SAN Performance

openBench Labs

A Division of Strategic Communications

Analysis: Raising the SAN Value Proposition By Automating Tiered Data Migration

compellent

Analysis: Raising the SAN Value Proposition By Automating Tiered Data Migration

Author: Jack Fegreus, Ph.D.
Chief Technology Officer
Strategic Communications
http://www.StratCommInc.com
July 6, 2006

Jack Fegreus is Chief Technology Officer at Strategic Communications, which consults with a number of independent publications. He currently serves as Editorial Director of Open magazine, Labs Director of HP World and contributes to InfoStor. He has served as Editor in Chief of Data Storage, BackOffice CTO, Client/Server Today, and Digital Review. Previously Jack served as a consultant to Demax Software and was IT Director at Riley Stoker Corp. Jack holds a Ph.D. in Mathematics and worked on the application of computers to symbolic logic.

Table of Contents

Executive Summary

> "To create a SAN environment that can provide the level of automated storage management traditionally associated with an ILM package, Compellent radically restructured the way storage is virtualized."

Neither Moore's Law—processing price/performance doubles every 18 months—nor Shugart's Law—magnetic storage price/bit halves every 18 months—shows any signs of being repealed in the foreseeable future. Nonetheless, resource planning remains a burning issue for IT as tight budget constraints, increased demands on administrator time, and unplanned increases in system workloads continue to plague IT planners. As a result, IT consolidation and simplification are securely ensconced within the mantra of many CIOs.

At the same time, CEOs, continue to pressure CIOs to provide an IT infrastructure that demonstrably improves business responsiveness and operational speed in order to meet rapidly changing business cycles. The global IT practice of McKinsey & Company, Inc. projects for 2006 that investments will be made in business intelligence software that can access data in Enterprise Resource Planning (ERP) systems for market analysis, as well as, software that can either improve the productivity of industry-specific processes or addresses industry-specific competitive issues.

More importantly, this IT expansion comes when overall IT budgets for capital and operational expenses are experiencing modest 3% growth. To fuel double-digit capital growth, CIOs are left with leveraging Moore's and Shugart's Laws to garner significant operational cost savings. A key resource targeted for significant cost-saving is storage, which is pegged by the Storage Networking Industry Association (SNIA) as having grown at a 79.6% compound annual rate for global companies.

To successfully wring out those savings, however, application workloads must be separated from infrastructure resources to minimize any risk of negatively impacting mission-critical applications, while optimizing and changing the underlying infrastructure. That separation process is termed "virtualization" and involves presenting an operating system (OS) with a logical representation of the device in place of the actual hardware. The logical or virtualized device masks any changes to the underlying hardware from the OS or any application running on the OS.

For storage, separating a logical resource representation from its physical implementation starts with the adoption of a storage area network (SAN). That makes a SAN a necessary condition to achieve the ultimate cost-cutting goal for storage: a self-managing resource that scales without disruption. While a SAN is necessary to achieve significant cost cutting, it is not sufficient to guarantee that those goals will be achieved. Garnering the level of storage cost savings needed to fund increased capital spending requires significant changes in the way storage is managed.

Gartner projects that three to five times more per GB is spent managing storage than acquiring it, which makes a significant reduction in the Total Cost of Ownership (TCO) dependent on driving down management costs. This is particularly important when dealing with structured and semi-structured data on which mission-critical applications, such as ERP and email, are built. These key production systems have performance limitations that require both optimally organized data and optimally configured storage. Automating an optimal storage environment for these applications has the potential to generate substantial management savings.

Until now, to approach that level of automated storage management required a sophisticated and costly Information Lifecycle Management (ILM) application running on top of a SAN implementation. To create a SAN environment that can provide the level of automated storage management traditionally associated with an ILM package, Compellent radically restructured the way storage is virtualized. While traditional SAN software virtualizes storage based on partitions of disk volumes, Compellent's Storage Center virtualizes storage based on disk blocks.

Fundamental Technologies

> "The Compellent Storage Center significantly reduces SAN TCO by obliterating an astonishing number of storage management issues and tasks."

Building an Architecture on Virtual Disk Blocks

Unlike most intelligent SAN storage arrays, the Compellent Storage Center™ is sold as a complete modular storage area network (SAN)

OpenBench Labs Test Briefing:
Compellent Storage Center™ SAN

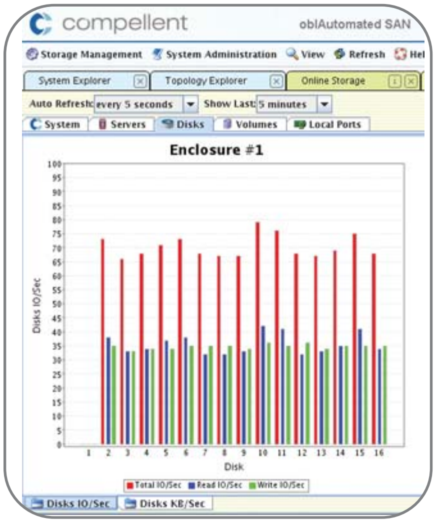
1) **Compellent virtualizes storage at the disk-block level versus traditional SAN virtualization at the array-partition level.**
2) **Dynamic Block Architecture™ tags logical blocks with metadata that includes last access time and RAID characteristics to emulate.**
3) **Logical blocks are served from tiered storage pools based on drive characteristics, such as interface—SATA or Fibre Channel—and rotational speed—15K or 10K rpm.**
4) **Compellent Storage Center manages logical blocks only when they are utilized and not when they are allocated.**
5) **Dynamic Capacity™ provides thin provisioning by automatically expanding disk pools to support allocated logical blocks.**
6) **Data Progression™ provides for automated tiered storage by migrating data blocks across pools based on access policies.**
7) **Using our oblDisk benchmark, I/O throughput was comparable to tests run on traditional SANs.**

solution and not as a single component. The reason for this is rooted in the product's remarkable value proposition. The Compellent Storage Center significantly reduces SAN TCO by obliterating an astonishing number of storage management issues and tasks. To achieve this goal, Compellent seized upon the notion of virtualizing the most fundamental element of storage: the data block.

All SAN virtualization software presents host systems with virtualized logical disks. Using traditional SAN software, that process starts with a SAN administrator combining physical disks into RAID volumes, partitioning those volumes, and then virtualizing the partitions into logical disks. Compellent software's rich functionality, however, is driven by a very different, sophisticated, and unique construct: Dynamic Block Architecture™. For an administrator using the Compellent Storage Center, virtualization does not begin at the level of a partition belonging to a disk array, but at the level of a logical disk block.

All of the disk blocks associated with all of the drives within a SAN are abstracted into a logical space of storage blocks, which can be larger than the physical space. Compellent accomplishes this extraordinary feat of legerdemain with the aid of a rich collection of metadata. Each logical disk block is associated with a collection of tags that represent notions that are normally associated with file-level and volume-level data constructs. File-oriented metadata includes such notions as data type and time stamps for events, including data block creation, the last access, and the last modification of the data. In addition to that file-oriented metadata, metadata that is typically associate with disk volumes includes constructs such as the type of disk drive, the associated disk tier, the underlying RAID level for data security, and the corresponding logical volume.

A very powerful and subtle aspect of Compellent's implementation of logical blocks and volume-oriented metadata is the rather remarkable



virtualization of the notion of RAID level. Within a Compellent Storage Center environment, RAID level is just a mathematical abstraction that relates to data security and availability. No longer does RAID level relate to a physical disk-formatting task.

With a Compellent SAN, the system manager no longer needs to make any decisions on how to physically format RAID disk volumes. Immediately, that takes an important task in storage management off of the table. What's more, the virtualization of RAID levels directly resolves another important storage management issue: the need to balance I/O requests. The Compellent Storage Center automatically spreads and balances I/O requests across all disks in a physical tier independently of the RAID meta-data classification. As a result, I/O performance scales optimally as disk drives are added.

Maximizing Performance and Functionality

Optimization of real-time performance is important to support virtualization and server I/O on any SAN controller. The fine-grained Dynamic Block Architecture and extended functionality of the Compellent Storage Center, which includes the ability to transparently migrate data across tiers of disk drives, only amplifies the importance of performance. On the other hand, the complexity associated with the implementation of such levels of device and data abstraction has the potential to introduce a significant amount of overhead.

For critical real-time operations such as the reading and writing of disk data, such added overhead carries a high probability that it will turn response time to sludge. To resolve this issue, Compellent chose not to use a traditional real-time operating system (RTOS) on its SAN storage controllers. In place of a traditional RTOS, Compellent chose to use eCos (Embedded Configurable Operating System), an open source application specific operating system (ASOS).

Like a normal operating system, an RTOS runs independently under an application. In contrast, an ASOS is explicitly linked with the application—in this case the Compellent Storage Center—code to form a single executable image. This creates a precisely tuned environment for running the Compellent Storage Center software. What's more, it

When we ran a file I/O benchmark on a logical RAID volume located within a disk enclosure, I/O requests were distributed by the SAN software evenly across all of the disks in the enclosure.

From RTOS to ASOS

There are many real-time operating systems (RTOS) available, which are often based on the Linux kernel. While the Linux kernel can be customized for an application, the OS still requires a minimum set of system resources in order to run and never becomes application specific.

Using a traditional RTOS, applications run on top of the OS and are written using embedded APIs for that OS. On the other hand, with an application specific operating system (ASOS), such as eCos, POSIX APIs are embedded within the application. The application code and the ASOS are then linked to create a single executable module.

As a result, the application drives very fine-grained customization of the OS. This generates a minimal resource footprint for both the OS and application and facilitates optimal run-time performance.

embodies the SAN with the on-going ease-of-use that is usually reserved for devices that are dubbed "an appliance." In line with the notion of an IT appliance, the Compellent Storage Center solution provides IT with an infrastructure that is at once highly-modular and highly-customizable while exhibiting tight integration, real-time performance, and a high level of resource utilization .

Currently, the principal hardware components that can be utilized in building a comprehensive Compellent SAN solution are:

- Intel-based servers acting as Storage Center Controllers;
- Qlogic QLA®2342 Fibre Channel HBAs featuring automatic multipath failover support;
- Qlogic QLA®4050C ISCSI HBAs primarily intended to extend SAN connectivity to remote locations;
- McDATA/Brocade Fibre Channel Switches; and
- disk drive enclosures populated with either Fibre Channel or SATA drives.

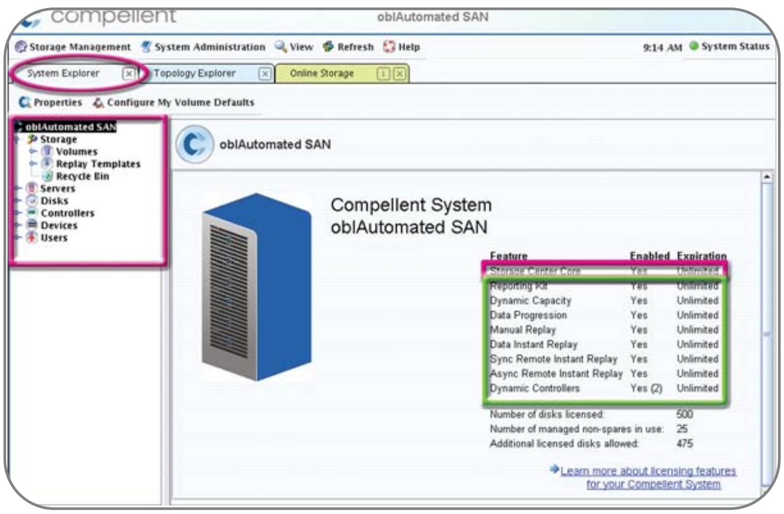
Moreover, through new software releases, this architecture can be readily expanded and scaled through the addition of new hardware within the Compellent Storage Center. Additions that are planned for near-term introduction include a new fast tier for Fibre Channel drives that connect via HBAs at 4Gb per second and a new middle tier for SATA-II drives that connect at 3Gb per second.

Using well-defined collections of hardware, a Compellent SAN is able to scale in performance, total storage capacity, and system availability with minimal impact on management overhead. Storage enclosures can be added in any increment to increase storage capacity and I/O will be automatically balanced over larger storage pools. For high availability, QLogic Fibre Channel HBAs provide multipath failover support between Storage Center controllers and disk enclosures. In addition, Storage Center controllers have on-board battery backup to protect data in the event of a power loss. For sites that need even more in high-availability support, Storage Center controllers can be clustered to provide for failover at the Compellent controller level.

Storage Center Basics

"The Compellent SAN's single most powerful feature for reducing system management costs is automated Data Progression, which transforms the SAN into a virtual ILM appliance."

Core Software Features



The Compellent Storage Center uses an embedded Web server to provide a single interface that can support all of the SAN's underlying technology. The classic System Explorer view, uses an expandable tree to present the components that make up the Compellent Storage Center. Also, there is a Topology Explorer view that provides the SAN administrator with drag-and-drop graphical editing of the site's entire configuration. At the top of the System Explorer tree, information about the licensed SAN software components can be found. These components include what Compellent dubs the Storage Center Core and Storage Center applications, which are licensed separately.

Storage Center Core provides the foundation features for a Compellent SAN: Dynamic Block Architecture sets up block-level data management; advanced virtualization enables managing physical disks as a single pool; data caching supports multi-threaded read-ahead operations and mirrored writes; servers can boot from the SAN and eliminate all internal drives; and logical volumes can be copied, mirrored and migrated without impacting users.

While all of the necessary components to set up a functional SAN are contained within the Storage Center Core, the advanced features needed to change the storage management paradigm are licensed as optional

In addition to running the Storage Center embedded Web interface on IE with Windows XP Pro, we had no problems running the interface on a laptop with SUSE Linux 10 using the Epiphany Web browser. Epiphany is part of the GNOME desktop and uses the Mozilla layout engine to display web pages.

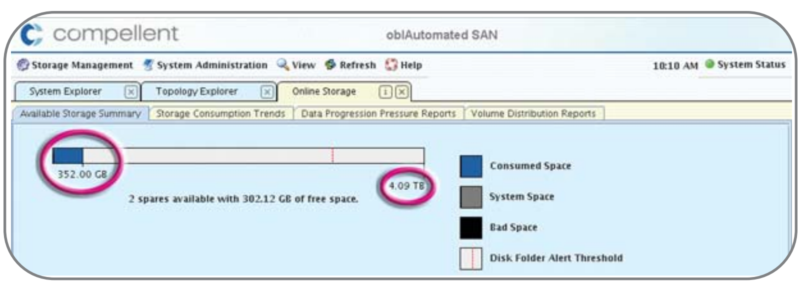
Going Ahead with Management

Compellent builds its management interface using the GoAhead open source Web server, which has a very small footprint and is designed to be embedded in a wide range of small devices. GoAhead supports JavaScript, in-memory CGI processing, and Active Server Pages (ASP)—Windows CE devices are among Web server's targets. The Compellent Storage Center uses embedded JavaScript to create dynamic data in ASP pages which GoAhead delivers to client systems performing storage management tasks.

applications. Of these Storage Center options, three applications are critical in order to dramatically reduce operational overhead costs that relate to system and storage management: Dynamic Capacity, Data Progression, and Data Instant Replay.

Compellent Storage Center Applications	
APPLICATION	**FEATURE**
Dynamic Capacity™	**Dynamic expansion of storage capacity for thin provisioning**
Data Progression™	**Automated tiered storage puts data on the most cost effective media vis à vis throughput and capacity**
Data Instant Replay™	**Immediate recovery from any data hazards via local data replication**
Remote Instant Replay™	**Immediate recovery from any data hazards via remote data replication**
Reporting Kit	**Detailed storage trend reporting**

Automatic Storage Expansion

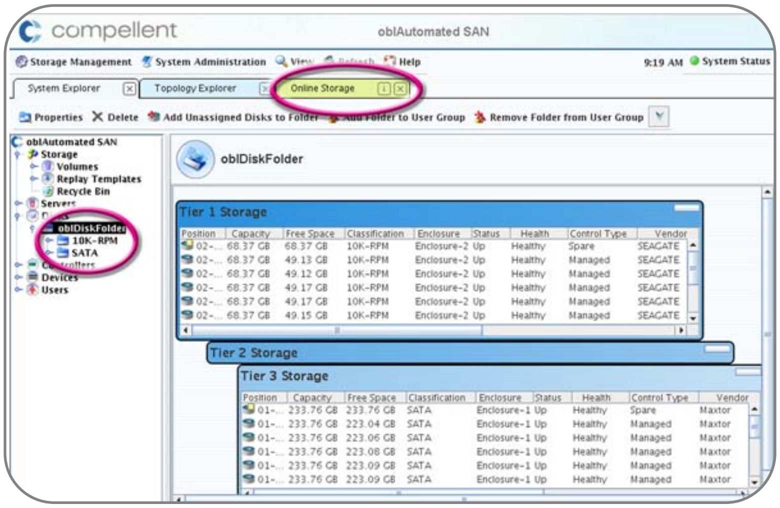


Storage Center advanced virtualization assigns logical disk blocks to logical drives only when data bits are written to the drive. The Dynamic Capacity application builds upon block-level virtualization and expands the real capacity of a disk folder automatically as drives are added. This allows SAN administrators to define logical volumes that have larger capacities than the real physical capacity available in a folder. This is especially useful whenever a host operating system does not readily support the expansion of a disk volume's capacity.

Our oblDiskFolder had a physical capacity of 4TB. Storage Center, however, deals only with utilized data blocks to minimize system overhead, which in our case was 352GB.

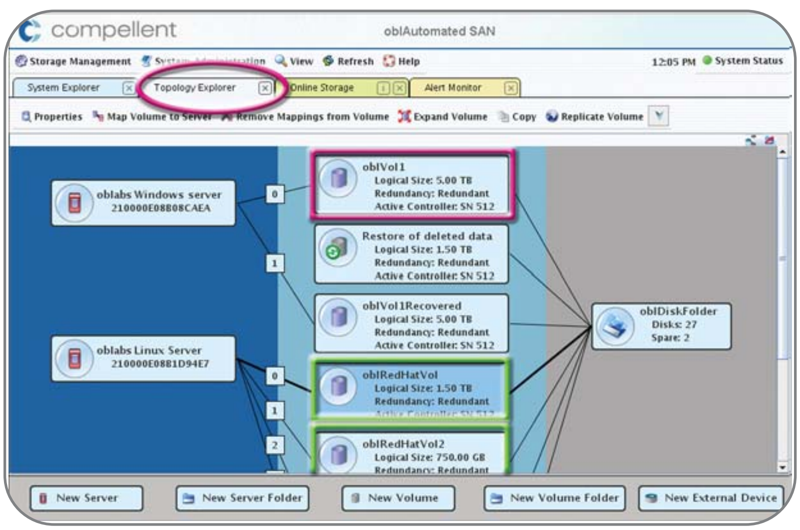
By providing for the allocation of more storage than is physically installed—dubbed thin provisioning—and consuming physical disk resources only when data bits are written, Dynamic Capacity takes a number of the issues complicating capacity planning off of the table. One of the most important of these issues involves a trade off between current capital expenses and future administrative expenses.

A number of applications and operating systems require significant



management intervention when their underlying storage volumes must be modified. As a result, a decision must be made before implementation as to whether it is more cost effective to immediately acquire all of the disk capacity that will likely be needed in the future in order to avoid all of the additional management tasks that will be needed to modify the storage architecture in the future. Complicating this calculated tradeoff is the fact that disks are a rapidly deflating commodity—Shugart's Law equates an 18-month delay with a net cost savings of 50%—while storage management is an inflating labor cost.

Setting up physical disks in a folder is astonishingly trivial: administrators put disks in a managed pool. There are no RAID arrays to create or manage. Storage Center groups drives in tiers based on drive type—SATA or Fibre Channel—and rotational speed—15K or 10K rpm. In our tests, the oblDiskFolder had two tiers: Tier 1 with 10K rpm Fibre Channel drives and Tier 3 with SATA drives.

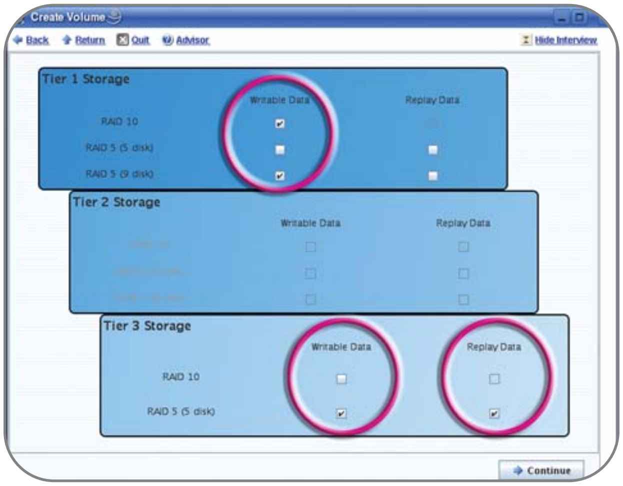


Dynamic Capacity obviates the need to make that trade off: Leveraging this feature, IT can cut up-front expenditures, avoid future management costs, and incur no penalty. The key to these benefits is the ability of Compellent's Storage Center to work with utilized rather than allocated space. Since any block that is not utilized is ignored, an extended number of logical disk blocks can be allocated to a logical volume without incurring any loss in performance. In our testing, we allocated a logical volume to a server running Windows 2003 Server that was larger than all of our physical disk space and measured no impact on SAN performance.

For our test site we set up two servers: one ran Windows® Server 2003 and the other Red Hat Linux®. Logical disks were created from a single folder of 29 physical disks. We assigned one logical disk, oblVol1, a capacity of 5TB, which was 1TB greater than the total physical capacity of the disk pool, and exported that disk to our Windows server.

Automatic Data Relocation

The Compellent SAN's single most powerful feature for generating



With Data Progression running, we were able to utilize multiple RAID levels and multiple disk tiers for both "Writable" and "Replay" data, which are Compellent's terms for normal and snapshot data. Without Data Progression, we would have been limited to one RAID level and one disk tier for the underlying disk structure as in a traditional SAN.

savings by reducing storage management overhead is Data Progression, which transforms the SAN into a virtual ILM appliance. Data Progression provides the support needed for automated tiered storage. With Data Progression, the SAN administrator defines policies about the frequency with which data is accessed. Then the Compellent controller tracks data access patterns on a real-time basis and transparently migrates logical data blocks between storage tiers according to those policies. Data Progression elevates the Compellent SAN from a storage appliance to an ILM appliance.

All of the storage growth projections that put the compound annual growth rate (CAGR) as high as 125% make Data Progression very interesting for IT. Even the more modest projection of the SNIA, a 79.6% CAGR, pegs a typical site's storage volume as growing by an order of magnitude by 2010. That will put an extraordinary storage management burden on any site trying to balance growing demands for data with the need to optimize their IT infrastructure.

A key strategy to maintain a cost-effective storage infrastructure calls for administrators to optimize the placement of data on devices based on the frequency with which that data is accessed and the performance characteristics—hence cost—of the storage device. In this scenario, only the most frequently accessed data files are retained on the highest performing devices. The goal is to purchase fewer high-performance storage devices in order to reduce overall storage expenditures.

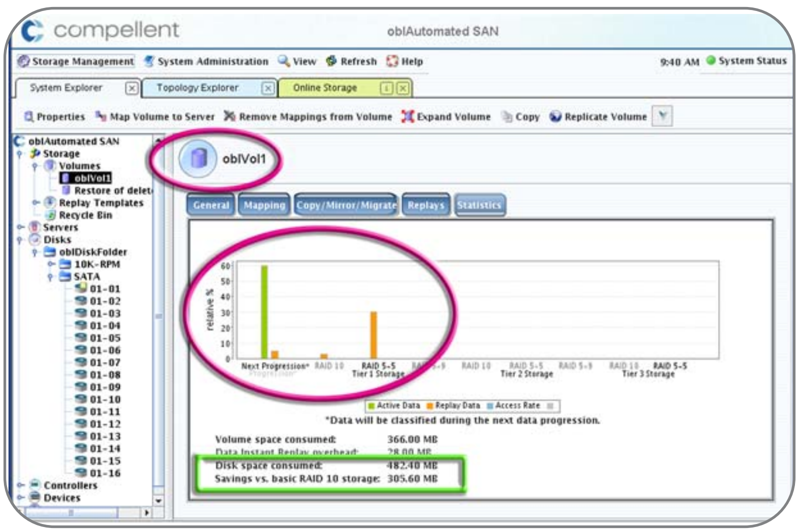
Data Optimization: Location, Location, Location

For mission-critical applications that use either structured or semi-structured data, that scenario can become an exceedingly complex task. Thanks to data retention regulations, such as the Sarbanes-Oxley Act, mission-critical production databases must now retain significantly greater amounts of historical data, which creates a significant overhead for both storage and database administrators. The lack of storage granularity within database software compounds the impact of the growing use of production databases as historic data repositories.

For a database, the smallest discrete addressable storage component is a table, which ties tables to logical drives. Without ILM functionality, record access activity is a relatively meaningless statistic in terms of where data can be directly located. For storage optimization, DBAs are limited to placing tables on logical drives.

The only way to optimize the location of historical records is to restructure the tables within the database. Typically, that involves creating new instances of tables for storing historic data that are different from the production-instance tables. With that restructuring, tables can be placed on different logical disks with different underlying physical characteristics.

None of that manual intervention is necessary on a Compellent SAN with Data Progression. Since disk blocks are as virtual as the logical drives that contain them, the Compellent controller can freely place infrequently accessed data blocks—and the records that those blocks represent—on the most cost-effect storage devices, without changing the way logical drives are presented to the host OS. More importantly, block migration is completely transparent to both the OS and any applications. ILM software requires application-specific modules that embed stubs in the application's data files to redirect that application to any new data location.



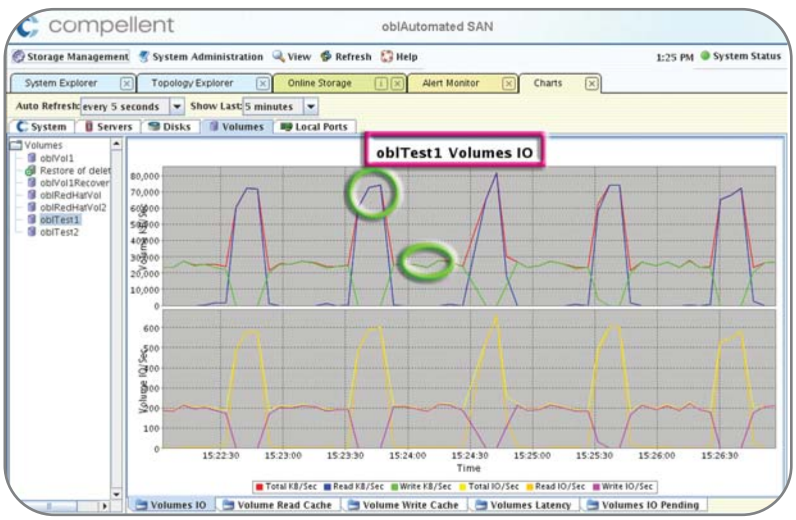
Shortly after creating our oblVol1 logical drive, the automated Data Progression application had spread its logical blocks over both RAID-10 and RAID-5 logical blocks in our Tier 1 storage. Moving blocks from RAID-5 to RAID-10 saved over 300MB. Over time, logical blocks would also be moved to Tier 3 (SATA) storage.

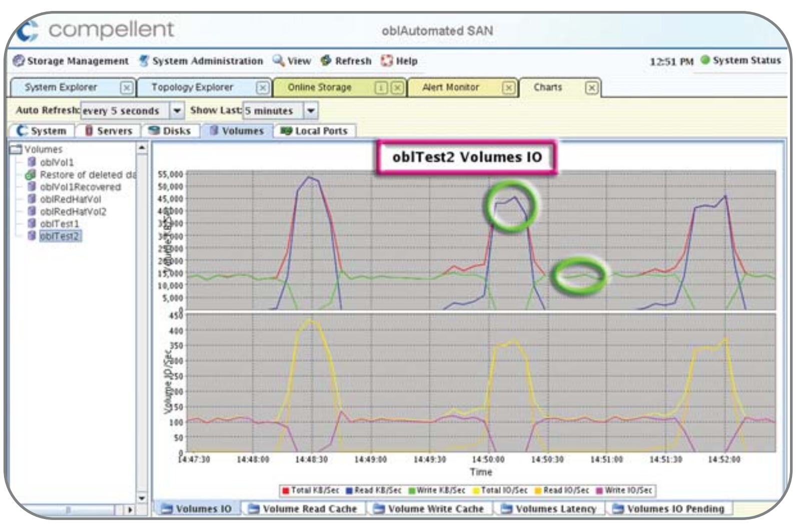
Compellent's analog to traditional snapshots is dubbed Data Instant Replay. The read-only copies of data are called replays and consume minimal storage space. Replays can be scheduled for automatic creation at specific intervals via intuitive templates or created on demand. Replays provide for extremely fast recovery from business interruptions including viruses, file deletions, and file corruption. Storage Center's architecture allows for the creation of an unlimited number of replays. Without Data Progression licensed, all replay data resides within only one RAID level and disk tier.

Value Proposition

“The Compellent Storage Center SAN generates storage environment savings that are entirely transparent and involve neither the explicit nor implicit manipulation of any file structures for any applications.”

With the extensive functionality and pervasive virtualization, which makes disk blocks logical rather than physical entities, the issue of I/O performance should hold top-of-mind attention for any IT decision maker who is assessing a Compellent SAN. To get a handle on performance, we ran our streaming I/O benchmark, oblDisk v3.0, on two logical RAID-5 disks.





The first logical disk, oblTest1, was created from a pool of Fibre Channel drives spinning at 10,000 rpm. The second logical disk, oblTest2, was created from a pool of SATA drives. Next, we exported both logical drives to an Intel-based server that was running Red Hat Linux.

Our oblDisk I/O benchmark repeatedly provided throughput results consistent with previous tests of 10K Fibre Chanel and SATA logical drives that were created using logical partitions of physical arrays with traditional SAN software. On reads, I/O consistently peaked around 72MB per second using the volume created in Tier 1 with 10K

We created two logical RAID-5 drives at opposite ends of the performance spectrum and exported them to a server running Red Hat Linux. These volumes were formatted on the Red Hat 8.0 server with the ext3 file system. The Tier 1 drive, oblTest1, utilized 10K Fibre Channel drives, while the Tier 3 drive, oblTest2, utilized SATA drives. We ran our oblDisk v3.0 benchmark on both logical drives and monitored progress with the Compellent Storage Center’s performance reporting tools. The results were entirely consistent with logical drives created on RAID-5 partitions created with traditional SAN software.

Fibre Channel drives and 28MB per second using the Tier 3 volume on SATA drives. Write throughput was 42MB per second and 14MB per second respectively.

In our throughput test, we found no significant overhead penalty on system performance using logical drives created from virtualized disk blocks rather than from a virtualized physical array partition, which allowed us to freely benefit from the ILM capabilities of the Compellent SAN. In particular, we were able to set policies for logical disk blocks that associated their location in storage tiers based on the frequency of access and the frequency at which Compellent would automatically realign disk blocks within the storage tiers.

Compellent's Cost-Savings Fundamentals

DRIVE	DRIVE TIER	CAPACITY	COST/GB	RAID	READ I/O RATE	WRITE I/O RATE
Seagate Cheetah	Fibre (10K)	147GB	$2.99	5	72MB/second	42MB/second
Seagate Barracuda	SATA	250GB	$0.36	5	28MB/second	14MB/second

The savings in hard storage costs alone with automated tiered progression can be prodigious. While Fibre Channel drives offer a three to one advantage in throughput performance, SATA provides better than an eight to one advantage in capacity costs. This can be particularly attractive in areas such as high-performance technical computing, where IDC projects that storage requirements will grow from a median level of 27TB today to 500TB in the next five years.

1TB Cost-Savings Scenarios

SCENARIO	ACTIVE DATA	TIER1 STORAGE COST	TIER3 STORAGE COST	COST SAVINGS
Email Server	15%	$448	$270	$2,220
ERP Database	70%	$2,093	$108	$789

The relative ease of these savings is particularly compelling for mission-critical applications that are built on either structured or semi-structured data files. The Compellent Storage Center SAN generates storage environment savings that are entirely transparent and involve neither the explicit nor implicit manipulation of any file structures for any applications. In many cases, the cost of a DBA manually restructuring the appropriate database tables would exceed the savings in hardware costs and could potentially lead to problems with the application induced by those changes.