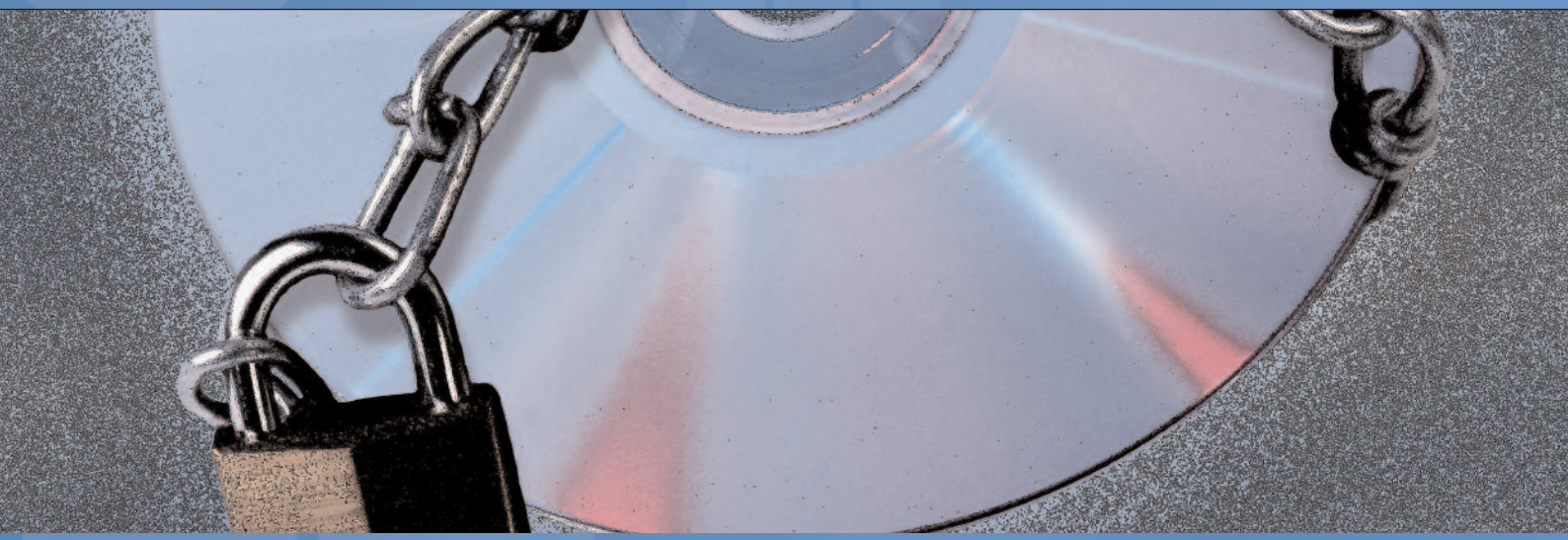
EMC² where information lives®



Affordable Data Protection: No Longer an Oxymoron

Leveraging midrange storage platforms for tiered data protection and business continuity

Table of Contents

Abstract

Storage technology developments have enabled tiered business continuity solutions on midrange storage platforms. As a result enterprise-class business continuity solutions are now more affordable and can be deployed in more areas in the enterprise. This EMC Perspective outlines the increasing emphasis enterprises are placing on protection of critical data assets and the growing list of cost-effective options that are available to address this requirement.

Enterprises are leveraging a growing range of affordable data protection solutions

Until recently, the comprehensive business continuity solutions used to protect critical information systems have been financially impractical to deploy throughout the enterprise. Except for the largest organizations that have already implemented comprehensive business continuity capabilities, the requirement to mitigate the risks resulting from a disruption to technology infrastructure has steadily grown from minimal to urgent over the past few years. Fortunately, as many of these companies have begun to address these requirements, technology developments have made a growing number of affordable solutions available.

Many factors beyond high-profile disaster events affect demand for business continuity solutions

Even as organizations continue to focus on the potential for high-profile disaster events like 9/11 and Katrina, many other forces have contributed to an increase in the demand for data protection and business continuity solutions for all data in the enterprise.

For some, it is the compressed business cycles brought about by online processes and increasingly demanding expectations of customers to be open and available for business at all times. As companies continue to put more processes online, the volume of data captured and stored continues to grow, which places more information and processes at risk.

For others, it is the growing requirement to meet regulatory demands that dictate the need for data protection. Over the past few years, as companies prepare themselves for compliance with Sarbanes-Oxley and other industry regulations, they take a renewed look at how their business processes and infrastructure ensure business continuity under any circumstance.

For many, it is simply the move to centralized storage networks that were initiated because of a need for greater efficiencies in managing overall storage resources. These companies realize that the risks of disruption of the centralized storage network are much greater with data concentrated on the network rather than spread out among many servers with direct-attached storage. So the move to centralized storage networks requires that business continuity considerations be addressed as part of the implementation.

These factors have forced enterprises to reassess data protection levels and begin to explore more flexible business continuity solutions in the face of continued IT budget pressures. According to a recent survey of medium-size organizations, nearly half are protecting data across buildings or across sites using replication technologies, while a majority of the remaining companies plan on doing so in the near future.

Successful companies evaluate business requirements first before considering technology

IT organizations that have experienced the most success with data protection/business continuity initiatives, work with the end users to establish the true business requirements first before making any technology decisions. Business users should be asked two questions regarding specific data and systems: how much data can we afford to lose if our applications went down; and what is the longest outage we could tolerate? The answers to these questions are referred to respectively as the recovery-point objective (RPO) and the recovery-time objective (RTO).

Before jumping to a technology solution however, the business user should carefully consider two factors in order to objectively determine the criticality of specific data and applications. These factors include the monetary value of a specific application and the associated data it holds for the business, and the costs that would be incurred should this system be unavailable. Once this business value is established, the conversation between the IT organization and the business unit regarding which level of protection the system should get becomes an objective assessment of value versus costs. With the emotional element removed, the technology decisions that are made using this approach tend to support the business requirements more cost-effectively.

Match the protection level to the business value

With an understanding of value and the maximum acceptable data loss and outage duration, the RTO and RPO requirements can be mapped to one of a range of data protection solutions by considering the costs for each, making sure the costs are consistent with the value. If the IT organization has already established a catalog of data protection service levels, users can quickly determine if they are willing to pay for the RPO and RTO they thought was needed.

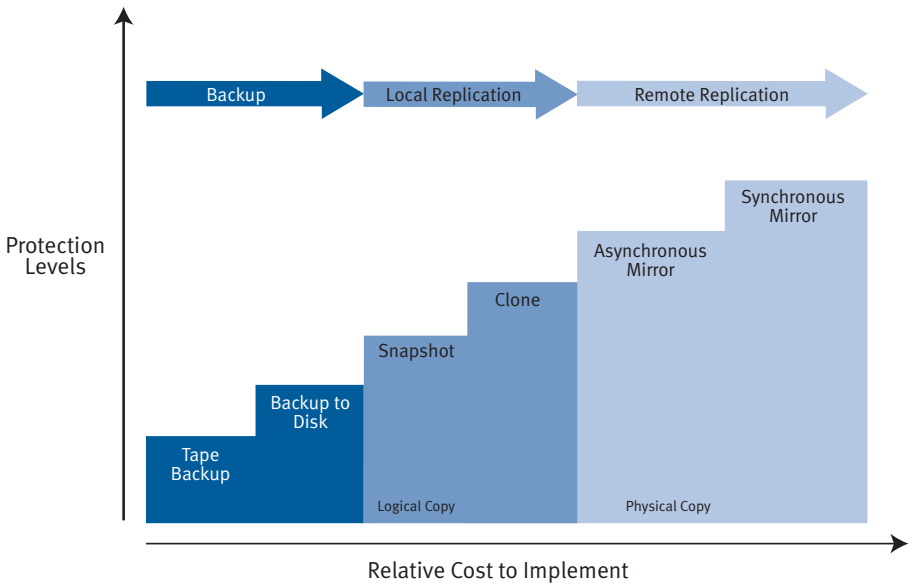


Figure 1: Tiered protection

At a very basic level data is protected by making copies of the production data. The method used to make these copies determines the level at which the data is protected as well as the costs required to reach that level. Figure 1 illustrates the range of methods available to deliver increasing levels of data protection. How quickly a copy can be made determines

how often the copies occur, which dictates the maximum potential data loss. More specifically, all transaction records that are captured after the last copy was made are potentially lost if a disruption occurs.

Tape backup at the left of the figure provides the lowest level of protection and results in the greatest potential data loss and longest time to restore in the event of a disruption. Many medium-size enterprises use tape backup as their entire data protection strategy, and most large enterprises use tape for some portion of their applications. Almost every organization is facing growing challenges in executing these backup procedures regardless of how large its storage environment.

First, as data volumes grow, backup procedures take increasingly longer to execute. Traditionally, this meant taking production systems offline sometime in the middle of the night. With the continued proliferation of online processes and 24x7x365 operations, this is no longer acceptable. The window to shut down production systems and back up data has disappeared.

Today, many companies use low-cost ATA-based storage arrays to do backup-to-disk (B2D). B2D is a faster and more reliable backup process so it alleviates some of these challenges and provides protection levels that are higher than tape backup at a slightly higher price point.

Another way customers are getting higher protection levels at a lower cost point is by deploying point-in-time snapshot technology. A snapshot is not an actual copy of data but a pointer to the original data as the data was at the time the snapshot was taken. Snapshot technology allows a logical copy of data to be made with much greater frequency since a snapshot is instantly accessible, unlike a full copy which needs to be fully synchronized before accessing. A snapshot also consumes less disk space than an actual physical copy. One thing that is important to consider is that with a full physical copy of the data, if there is a failure of the original dataset, you can restore the data set from that volume without significant data loss. Snapshot technology does not enable that level of protection.

EMC SnapView and SAN Copy ensure Shawnee County's critical records are protected across its CLARiiON SAN

Shawnee County is a local government organization serving 12 townships in Kansas, including the city of Topeka. Shawnee County's information technology department was serving more of its 170,000 residents' online information needs, but finding it difficult to manage the growing pool of direct-attached storage for the 45 servers scattered across its facilities. "Whether it was over-provisioned or under-provisioned, we just couldn't seem to get storage capacities right," explained Network Administrator Pat Oblander. "We simply couldn't hit the mark anymore with direct-attached storage."

As a result, Shawnee County made a decision to implement a storage area network (SAN) based on EMC CLARiiON CX series systems to consolidate the most critical information and begin to address its storage provisioning challenges. Since the SAN would result in data being concentrated in a single location, Oblander knew it was important to ensure the storage environment was hardened against failure. Prior to the SAN, the risk of data loss was spread out among the storage attached to the 45 servers so the risk levels were tolerable.

Fortunately, two EMC CLARiiON systems deliver the advanced data protection and business continuity capabilities needed. Shawnee County uses EMC SnapView to create a local clone of the SQL Server databases residing on a CLARiiON CX series array at its North Annex facility. These databases hold registered deeds and GIS system records supporting any application having to do with mapping, including E911, dispatch, appraisals, elections, etc. EMC's SAN Copy software is then used to push copies of the clones across the SAN to another CLARiiON CX series system located three miles away at the courthouse facility. These remote copies enable Shawnee County to test upgrades, fixes, and configuration changes without disrupting the production system. The result has significantly improved the reliability and accuracy of these processes.

"EMC's CLARiiON platform has provided the bullet-proof platform we needed," said Oblander. "We not only get a more manageable, straightforward storage environment that eliminates the provisioning headaches, but also get data that is more protected than ever. And we can recover from server failures much more quickly as well. If a server fails we can restart by simply booting a second server from the SAN. We're just beginning to tap into the operational capabilities afforded by EMC's replication technologies, like non-disruptive backup-to-disk."

With critical data records expected to grow 50 percent over the next year, and possible new mapping applications easily doubling that growth rate, CLARiiON's data protection and business continuity capabilities will play an increasingly important role in Shawnee County's information infrastructure.

Because of this, snapshots are usually implemented in combination with some form of actual physical copy or replica. Once again the decision should be based on the service level required for the environment.

At a high level, physical copies are either clones or mirrors. A clone is a separate physical copy of the production dataset at a specific point in time. A mirror is a separate physical copy that continually tracks or mirrors the changes made to a production dataset.

A clone can be made by either applying to the cloned dataset only the incremental changes that have occurred in the production data since the last time a clone was made, or by "splitting off" or disconnecting one of the mirrored copies from the mirroring process. This is often referred to as a business continuance volume or BCV.

A mirror can be either an asynchronous or a synchronous mirror. An asynchronous mirror accumulates all changes made to the production dataset and then applies these changes at a specified interval. Any transactions made between the time the accumulated changes are applied may be lost if the production dataset is disrupted.

A synchronous mirror continually applies each change made to the production dataset to the copy as well before committing the transaction back to the server. Since the copy is always in synch with the original, no transactions are lost in the event of a disruption to the original. Since the server has to wait for both the production and mirrored datasets to change, there is a slight performance penalty that occurs that increases as the physical distance separating the production and mirrored datasets increases.

Mirrors are usually made remotely from the production data in order to decrease the chances of an event that disrupts both the production and mirrored copies.

Replications can be made using the processing power of the application server or the storage processor inside the storage array itself. Server-based replication may be a lower-cost solution, but reduces the number of valuable server cycles that can be dedicated to production workloads. Executing replication procedures using the processing power of the storage arrays helps ensure the availability of critical production applications.

Willis-Knighton protects critical medical records using EMC MirrorView/S

Willis-Knighton is one of the fastest growing health systems in the United States with an aggressive plan to eliminate paper throughout its network of four hospitals in the Shreveport, Louisiana. area. As paper disappears, however, digital records grow and must be properly protected—not only to ensure effective patient care in the event of data loss, but also to comply with strict federal regulations, such as HIPAA, that govern information security.

Among the most critical data at Willis-Knighton Health System are Picture Archiving and Communication System (PACS) images, each of which can be several gigabytes in size. Willis-Knighton stores these images on a high-performance EMC CLARiiON Fibre Channel storage area network (SAN) in a primary data center at the Willis-Knighton Medical Center. To further protect these vital medical records, Willis-Knighton uses EMC MirrorView/Synchronous (MirrorView/S) software to continuously replicate the PACS images to a second CLARiiON SAN at a disaster recovery site seven miles away.

Jonathan Lee, Network Coordinator at Willis-Knighton Health System, said, "Loss of our PACS images could have a very serious impact on patient care because our physicians reference them frequently in the course of treating a patient. CLARiiON and MirrorView provide us with rapid access to actively used images, and the peace of mind knowing that we could recover these images quickly in the event of a disaster. Using MirrorView is important because we need to verify that the data is intact when it reaches the disaster recovery site.

"Previously, our PACS images were stored on optical disk, which would have required hours to restore. If we ran into mechanical problems, it could take longer—possibly even days. Today, we can fail over to our disaster recovery site in a matter of minutes. So our physicians can continue to access the PACS images with virtually no disruption to patient care."

Sophisticated software extends protection levels for complex application environments

New software technology is available that takes advantage of these replication technologies to automate the process of restarting an application when a primary server or storage array is disrupted. This technology couples business data and its associated mirrors with the application servers to manage the application restart process. When a primary site fails, the software seamlessly transfers control from the failed production storage to the remotely mirrored storage resource and restarts the application on the remote secondary server.

With the growth of Web-based applications and Web services, more environments involve multiple databases and interrelated applications. For instance, sales, manufacturing, e-commerce, and customer service records all may share common databases and therefore need to be consistent with each other at the transaction level. Therefore it is important that a copy being made of one database must represent the same point in time as the copy being made of another interrelated database before any attempt is made to restart a portfolio of applications that span these databases. If the two copies are not consistent and a restart is attempted, significant data integrity issues may result. Advanced consistency software technologies ensure that replicas made in these federated application environments are consistent at the transaction level. This protects the integrity of all data in the event of a disruption.

Compromise is no longer necessary

No organization can cost-effectively protect all of its data assets with just one of these technologies. Less critical data may require a simple backup while the most critical data likely requires a synchronous mirror coupled with an automated application restart capability. Using one or the other technology to meet the protection needs of both types of data will result in either excessive exposure to data loss risks or excessive costs. The most effective approach combines these technologies—from backup to snapshots, clones and mirrors—into a tiered protection infrastructure that delivers the most appropriate levels of protection to data based on its value to the organization. It is no longer necessary for organizations to compromise—either accepting greater risks than they need to, or implementing a solution that overprotects data and costs more to implement than is justified by the value of the data.

Taken together, these technologies allow an organization to not only protect its critical data assets, but also use the copies of production data to support parallel processing activities such as backup, application development and testing, and data warehouse refreshes while increasing the availability of production systems. Decision-support activities are more effective because the data warehouse can be refreshed with current data more frequently without bringing user query activity to a halt. Backups can be taken more frequently which reduces the potential for data loss. System upgrades, maintenance fixes, and configuration changes can be thoroughly tested offline using the most current copies of production data, which reduces the chances of failure or data integrity issues when the changes go live.

Midrange platforms are available with a full range of data protection capabilities

Midrange storage platforms like EMC® CLARiiON® are helping large and medium-size enterprises implement cost-effective tiered data protection that only the largest organizations could afford previously. CLARiiON offers a full range of end-to-end business continuity solutions that, when coupled with EMC Consulting services, are designed to meet every business continuity need.

EMC CLARiiON storage provides the ability to deploy a full range of tiered storage options within a single platform. High-performance disk drive technologies support mission-critical production volumes. Less-critical environments, like development, test, or reporting can utilize high-capacity/low-cost drives, and tape-emulating disk libraries can be used for backup and recovery operations.

In addition to working transparently with host-based replication software like EMC's RepliStor®, CLARiiON also works with a full range of array-based replication software that delivers snapshots, clones, or mirrors without consuming valuable server cycles or LAN bandwidth.

EMC SnapView™ software runs within the CLARiiON array and delivers either snapshots or full-volume clones that can be used as the source data for fast, frequent, and non-disruptive backups or for development and testing, or data warehouse procedures.

EMC SAN Copy™ also runs within the CLARiiON array and is designed to pull or push full or incremental copies between CLARiiON and other storage platforms including EMC Symmetrix®, Hitachi, HP, IBM, and Sun storage arrays. By integrating SAN Copy with SnapView, snapshots and BCVs can be used as the source volume from which SAN Copy pulls data for faster remote recovery operations.

EMC MirrorView™ is advanced, array-based replication software that provides synchronous or asynchronous mirrors between two or more CLARiiON systems. By integrating MirrorView with SnapView snapshots, copies of production data can be used for parallel operational procedures from a secondary location.

In addition, MirrorView replicas can be used with advanced high-availability software for near-instantaneous restart of failed servers at a remote location.

Both SnapView and MirrorView provide application consistency technology to ensure that copies made in environments with multiple databases and interrelated applications represent a consistent point in time, which ensures the integrity of any process that uses those copies.

CDW relies on EMC MirrorView/S for superior data protection

CDW® (NASDAQ: CDWC) is a Fortune 500 provider of information technology (IT) solutions and advice for business, government, and education. The $6.3 billion company built its IT infrastructure for the most rapid possible responsiveness to its customers' unique needs and therefore recognizes the importance of reliable access to core business information.

As a result, CDW protects its key customer transaction and financial information by replicating it from its main production center to a disaster preparedness facility 30 miles away. The data resides on EMC CLARiiON CX series Fibre Channel systems, which are installed at each site and connected by EMC MirrorView/Synchronous (MirrorView/S) replication software.

Steve Staines, CDW's Manager, Data Center Services, said, "We have a transaction-heavy business that places significant demands on our IT infrastructure requiring high performance, high availability, and business continuity. The combination of EMC's MirrorView/S replication solution and CLARiiON storage systems typically enables us to recover in five to 15 minutes in the unlikely event of a site failure and still ensure zero transaction loss."

More companies are now taking advantage of affordable data protection technologies

Midrange platforms like EMC CLARiiON are helping organizations like Shawnee County, Willis Knighton, and CDW implement cost-effective, tiered data protection solutions that only large organizations could previously afford.

To be sure, many IT executives are still under the impression that an investment in data protection is essentially an insurance policy that sits idly until a disaster strikes. Initial deployment costs along with the lack of business benefits to justify the costs are cited most often in industry surveys as the reasons for not implementing business continuity solutions to protect data.

In reality, the benefits go well beyond an insurance policy. Technology that was originally intended solely for disaster recovery purposes has evolved to the point where it is now being used to make operational procedures, such as development, testing, and data warehouse refreshes much more efficient. As a result, both IT staff and even end users utilizing decision-support systems are much more productive.

These ongoing productivity benefits make it much easier for medium-size enterprises to justify the investments required.

EMC has helped thousands of similar companies successfully implement comprehensive data protection and business continuity solutions on its CLARiiON storage platforms. Its experienced team has extensive knowledge of replication technologies and storage deployment best practices along with the proven methodologies to help medium-size organizations quickly assess both the risks and consequences of data loss, develop optimal data protection implementation plans, and maintain the availability of critical applications.

For more information, please contact your EMC account manager or visit http://www.EMC.com/continuity.

EMC²
where information lives®

EMC Corporation
Hopkinton
Massachusetts
01748-9103
1-508-435-1000
In North America 1-866-464-7381

EMC Perspective
H2216