Storage Networking: An Essential Guide to Storage as a Key Component of Business Continuation

Abstract

According to the 2006 Global State of Information Security report, published by CIO Magazine in conjunction with PricewaterhouseCoopers in September of 2006, business continuity is the number four "to do" item for this year, down from number one last year. Data backup rose from number three in 2006 to number one in 2007. Why are these two important goals for 2007?

Information is a business's primary asset. Historically, data storage was a local issue, either on floppy discs or hard drives. In a data center, storage went to tape and was stored in libraries. Things are more complicated today, especially after September 11th, the natural disasters of 2005 (the Indonesian tsunami and Hurricanes Katrina, Rita, and Wilma), or even severe storms such as those that hit Greensburg, Kansas in 2007, as well as the constant threat of data loss due to security breaches. Companies are more concerned than ever about storage and business continuation in the event of catastrophes—large or small, natural or man-made. This article provides a high-level overview of the concepts of storage networking and business continuity, including the often-confusing terms associated with the industry. It concludes with a concise look at the storage marketplace and a plan for survival.

Hill Associates offers a course that expands on this paper. For more information visit our website at www.hill.com.

Introduction

The world is a much different place from what it was even two years ago. The year 2005 saw an unprecedented amount of challenges with regard to business. Whether it was the natural disasters of the year, or what appeared to be an increase lack of diligence with regard to transporting and managing tape backup, things were a mess. September 11 radically changed our view of safety and security in 2001. Telecommunication and data communication rules seem to be rewritten constantly. Chaos reigns. The pace of technological change appears to be increasing, and our ability to predict the future, or keep up with the present seems farther from our grasp. Yet, at the Gartner Security/IT Expo held in Tampa in October 2005, as Hurricane Wilma was stalled in the gulf, it was observed that "too many companies either do not practice plans to keep their operations going or have top leaders who are not committed to disaster plans." When the EF-5 storm hit Greensburg, Kansas in May 2007, it destroyed 95 percent of the town. (See http://www.usatoday.com/weather/storms/tornadoes/2007-05-05-kansas_N.htm.)

We have learned some valuable lessons, though. Businesses now recognize that they must better prepare for uncertainty. Business continuity planning (BCP)—a process that enables businesses to minimize their losses while ensuring availability of mission-critical application data in case of a catastrophe—is now at the top of the list of the "to do" list for 2007, according to the 2005 and 2006 Global State of Information Security Reports published jointly by *CIO Magazine* and PricewaterhouseCoopers. Data backup was number three. Financial institutions are clearly the model business to emulate.

(You can read the reports at *http://www.cio.com/archive/091505/global.html* and *http://www.cio.com/article/24979/The_Global_State_of_Information_Security_*.)

Companies are looking at business continuity management and recovery because a "hot standby" is no longer sufficient. Having a data center is no longer adequate. And backing up to tape does not provide a realistic ability to recover from a disaster, despite what is in the plan—assuming there is a plan.

Information is everything, and information comes from data. As a result, data is a critical asset. The manner in which data, as well as the information and the knowledge it represents, is stored, managed, moved, backed up, recovered, and used in a timely, cost-effective manner determines whether a corporation will be successful, especially in turbulent times. Part of a company's business continuity planning (BCP) process must include the appropriate storage and recovery of such valuable information.

Thus, BCP has now emerged as the leading driver for storage area networks (SAN) as well as network attached storage (NAS) and content addressable storage (CAS). IT personnel are considering all aspects of storage networking as a way to ensure business continuity in case of a disaster. Unfortunately, this is often a complicated topic, with many confusing concepts and terms. Figuring out how SANs, NAS, and CAS fit into a business continuity strategy can be difficult and time consuming. However, consider the price a company might pay if it did not have a plan. In this paper, we address some of the more common issues and concepts of BCP and storage area networking. We welcome your comments and experiences in a dialogue going forward.

Defining BCP

Data has value when you have it. That value goes up considerably when you do not have it.

I live in Colorado. We recently experienced a summer of fires unlike any other in recent memory. In 2005, the Indonesian tsunami and Hurricane Katrina caused destruction on a cataclysmic scale. Despite the material loss of homes and possessions, many families recognized that their lives are not just made up of their shelter in the forests. Planning for continuation of life and rebuilding was a critical factor in the ability to survive such a disaster. The same can be said for businesses.

Storage, as part of an overall BCP, is not just about managing disk arrays and capacity or making sure the server stays up. Storage is about managing one's risk. It's insurance. Loss of data or loss of access to that data puts a corporation at risk. Such risk can involve lost time, lost revenue, lost business opportunity, or perhaps even lost life if we consider a health care provider's risk.

A tightly integrated BCP that includes storage hardware, software, and networking services can mitigate risk of loss, while securely protecting data from even the most often cited reason for loss—human error. The U.S. Department of Labor estimates that 93 percent of businesses that experience a catastrophe do not reopen. Of those that do, most experience significant challenges.

The High Cost of Downtime

It is no longer adequate simply to back up data and store it off-site in a safe location, or even at an emergency hot site. IT disaster recovery plans must include much more. Corporations recognize they must preserve and continue key business practices, operations, and communications

infrastructure to ensure the availability of information systems and data. Studies by the Fibre Channel Association and Gartner Group show that there is a high cost for having downtime.

In the 1980s, mission critical meant ensuring that the phone lines were up. Data was almost an afterthought. Today, corporations have many applications and processes that they consider essential or critical. In the past, essential meant which applications and information could not, under any circumstances, be lost or destroyed. Now, mission critical is defined as what is required to bring a corporation back online and running immediately, as if nothing had happened. Ask employees what they consider mission critical, and they may say email or web access. When designing a BCP, corporations must define their time to data (TTD). For example, one corporation might include that an order entry process must restored in one hour, whereas email must be restored in eight hours. This concept underlies what is known in the industry as recovery time objective, or RTO. It is one of the most important aspects of any BCP.

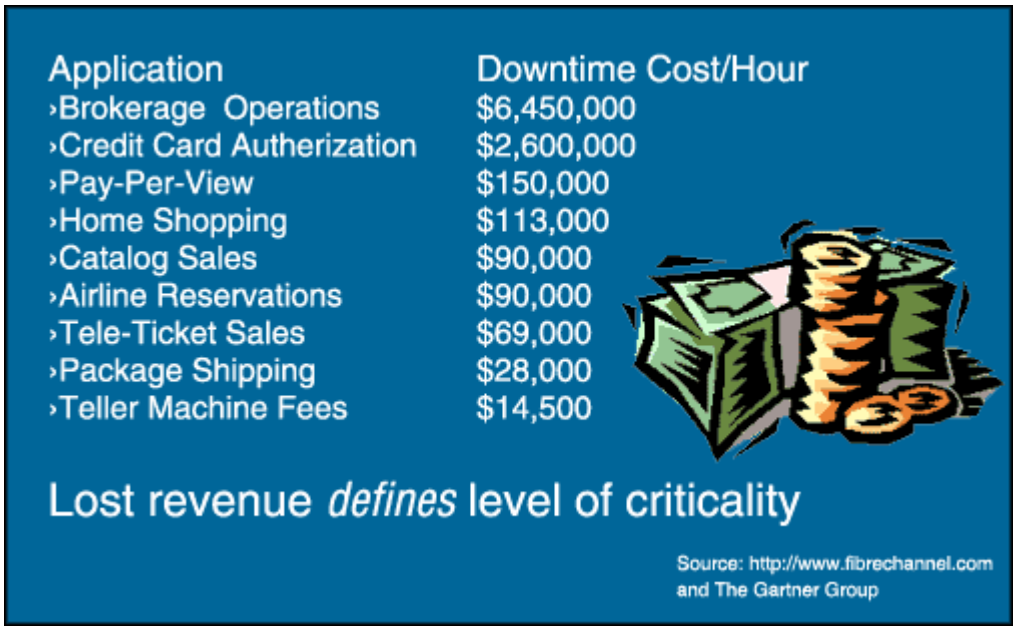

Figure 1: High Cost of Downtime

In the Beginning…

When computing power was introduced, there was a desire to have intelligence only in the data centers. Computing, processing power, and memory were all very expensive. The world of IT revolved around the mainframe. Then came the integrated circuit and Moore's Law. Computers evolved from enormous, large-scale centralized systems with dumb terminals hanging off to more nimble, portable, cheaper, and more efficient computing platforms. More information began to be created electronically, rather than on paper, and we needed a place to store it. In the 1980s, IT's focus was on the network and the data moving around it.

Data storage was a local issue. We first had floppy disk drives and then hard drives. Or, in the centralized data center, information was stored and backed up on tape or disk drives. There was a tight relationship between the host device and the storage media, with very little distance between these two. In data centers, tapes might have been stored in physical cabinets, or copies might have

been sent to a vault off-site. At the end of the 1980s, PCs began to have hard drives, and networking came into its own. Corporations deployed LANs with servers, and a need arose for specialized storage units.

Despite these advances, storage of any type was relatively expensive then. As mentioned, information is now most businesses' primary asset. Information, and the knowledge gained from that information, is driving today's Internet-based economy. We have seen increases in productivity and efficiency occur at unprecedented rates with this asset. It is therefore necessary to make this information always available, completely shareable by different computing platforms, anywhere in the world at any time, reliably. In addition, users want consistently fast access to this information, complete protection from loss, and a highly secure, cost-effective management structure. Today, storage has become the center of IT's world.

We are also in the midst of experiencing the perfect storm—three concepts coming together in time to create an opportunity unlike any other. The first is the continuation of Moore's Law, allowing for new data processing capabilities to create information that is used more proactively in business intelligence applications. Information Lifecycle Management has become critical to the success of businesses today.

Second, we recognize that memory, and thus storage, has become incredibly inexpensive relative to the media on which it is stored. Remember when getting one megabyte of RAM for under $100 was a good deal? As of January 2006, six gigabytes of memory on a micrcodrive costs $100. It is hard to find a new computer with a hard drive that is smaller than 80 Gigabytes. Memory is cheap today.

Tape is still many companies' preferred media for protecting data, and for creating a "backup copy." However, we now understand that tape suffers from many challenges. It is linear, and often backups are made in an incremental fashion, making for an RTO that is unacceptable for today's business. Tape backups are rarely encrypted. Therefore, when they are lost (as was reported almost monthly in 2005), the effort required to restore the trust can be astronomical. Finally, studies are now being published that demonstrate that roughly one-fourth of all successful tape backups cannot be restored. These are frightening statistics.

Finally, we must consider the death of distance. With bandwidth becoming a commodity today, companies are rapidly deploying technologies like Gigabit Ethernet to connect locations with astonishing speeds. Nothing is as cost-effective per megabit of transport as Ethernet. And with IP dominating everything as well, newer technologies like iSCSI and Fibre Channel over IP (FCIP) are making remote storage to disc very appealing.

Key Definitions

Before we discuss data storage and management in more detail, we should introduce some key definitions.

Direct Attached Storage

DAS requires that all storage access methods involve a CPU/server, with requests sent over an I/O bus like SCSI or IDE. As a result, DAS storage devices are tightly coupled with the host computer's operating system, and distance limitations come into play. DAS devices can be physically inside a server or external to it.

Network Attached Storage

NAS describes a specialized file server and storage device typically connected to an Ethernet LAN switch. Client machines "mount" volumes (really the disks of the server) to transfer data files to local machines. These servers typically have "skinny" operating systems that specialize in handling file read/writes. So, think of NAS as that updated file server down the hall.

Storage Area Network

A SAN is a dedicated, high performance network infrastructure optimized to move large amounts of data for applications that require optimal performance, redundancy, and availability. SANs are about having scalability, accessibility, manageability and intelligence to know what should go where and how. SANs often greatly extend the distance of a peripheral channel, whereas the NAS is simply another file server on the LAN.

We hear the SAN versus NAS, or SAN/NAS convergence question a lot these days. The answers are not always straightforward. NAS typically invokes thoughts of Ethernet and LANs in a small company or a department. SAN often invokes thoughts of Fibre Channel, data centers, and large Fortune 1000 companies. These two technologies are often linked, and in some cases, they complement each other. For instance, you can connect several NAS devices together with a Fibre Channel SAN topology. The main point is that most data will traverse an IP network, and the distinction between a SAN and NAS is becoming less important. SAN has become a generic term, like Kleenex or Xerox when we should say tissue and copy.

Content Addressable Storage

Content addressable storage, or CAS, is the new new thing. In CAS, we create a baseline of objects linked with meta data, resulting in the assignment of an identifier to data objects for access in a fixed content environment.

Think of a customer order, and what happens if that customer orders something new. The bulk of the order, if entered, would be redundant to most everything already in a database: the name, address, phone number, primary contact, etc. The only thing that is new is the actual order. What if we could just link that order to a set of objects that already exist within the database? That is the essence of CAS. Another example of fixed content is a medical record, or financial transactions that result from money withdrawn from a bank, or an acceptable charge for merchandise in a store, or over the Web.

The benefits of CAS are numerous. The performance for readily accessed data of a fixed nature can be improved tremendously. CAS is also the most efficient way to protect and recover data. As a result, many organizations are looking at CAS as their preferred method for continuous data protection as well as satisfying other BC/DRP requirements.

Data Storage and Managing SANity

If we want to store data reliably and be able to access it fast, we need to ask the questions below.

- Where should the data be stored?
- How many copies are required?

- How often should the copies be made?
- How will the copies be restored and recreated?
- How should copies be retrieved if required?
- Under what circumstances should a plan of recovery be initiated?

When we consider these questions, it helps to view storage in three primary constructs.

1. **The data:** When we consider how the data is to be stored, we must examine alternatives such as storing the information simply as files, or in more traditional "block" fashion.

2. **Data management**: Not only must business continuity and storage policies be written and updated regularly, but the actual management of the storage platforms should be considered. Users are moving to SANs to upgrade their storage networking capabilities, but the problem is managing and maintaining them once they are in place. This is the number one concern of network administrators today when it comes to storage. Other than SNMP, there is no single standard yet for managing storage alternatives.

   Network managers are also concerned with interoperability. Vendors have created proprietary products to maintain a competitive advantage. Despite the promise of interoperability, there are now emerging standards for managing a heterogeneous storage environment. This leads to massive confusion. However, there is hope on the horizon. Leaders like IBM, HP, EMC, and Veritas (now owned by Symmantec), along with several other companies, are working to complete the Storage Management Initiative Specification (SMI-S). The industry is expecting that version 1.1 products will appear en mass in 2006.

3. **Security:** When storage was directly attached to mainframes and under IT's domain, controlling access was relatively easy. With the emergence of LAN technology and everyone wanting access to every device, things get a bit more complex. In addition, with multiple storage servers, it will become critical to prevent one server from overwriting another's information. Access control remains critical, but, as with management, each vendor implements its own version of secure access control differently. This proprietary nature of the industry might be good for locking in a customer, but it is frustrating for that customer as well.

In addition to these constructs of storage, there are three building blocks of storage area networking.

1. **Filing the data:** Where can we safely and reliably store the data? Should we use magnetic media such as tape, or should it be disc?
2. **Storing, organizing, and managing the files:** File systems are tightly linked to operating systems, but most people never worry about those things. If I buy an NT server, I get NT's File System (NTFS). With Solaris, I get the Unix File System. But this is not what the buzz is about today. It's about the logic. How will we fragment the data, if required? How will we provide fair access to the disc (i.e., manage access)? It is much easier to manage one thing than 25. The concept of virtualization comes into play here.

Virtualization manages the quantity of your storage. Storage virtualization software allows us to enable all of the discrete discs and storage devices connected to the storage network and view them as one virtual pool of storage capacity. And since any storage device can be virtualized, all devices can work together to maximize the utilization of the entire pool. Virtualization is considered the fastest growing segment of the SAN management software market and the specific point of the buzz. Veritas is the dominant player here, with an estimated 80 percent market share with their Volume Manager product.

3. **Providing connectivity to enable access:** Various users, systems, and devices need access to the information. This is a hot topic in the storage arena. Common terms include iSCSI, Fibre Channel, Gigabit Ethernet, and FCIP (in addition to NAS, SAN, and DAS).

The lines of demarcation among the above three building blocks of SANs are blurring, but there is a difference—at least for now (a difference that would take another white paper to explain). All of these serve the requirements for storing data/information. Storage networking has evolved to address the need to provide more open connectivity to data and information and the desire to break the bottlenecks that occurred with traditional DAS.

As we move toward a more global networking model with any-to-any connectivity, we will move toward a general concept of storage networking. A physical connectivity to the actual storage device will no longer matter. What will matter is storage device management, information availability through its lifecycle, automatic backup and recovery of that information, and the quality of the network service. We are not there yet, but there is certainly more awareness that storage networking needs to be more widely implemented.

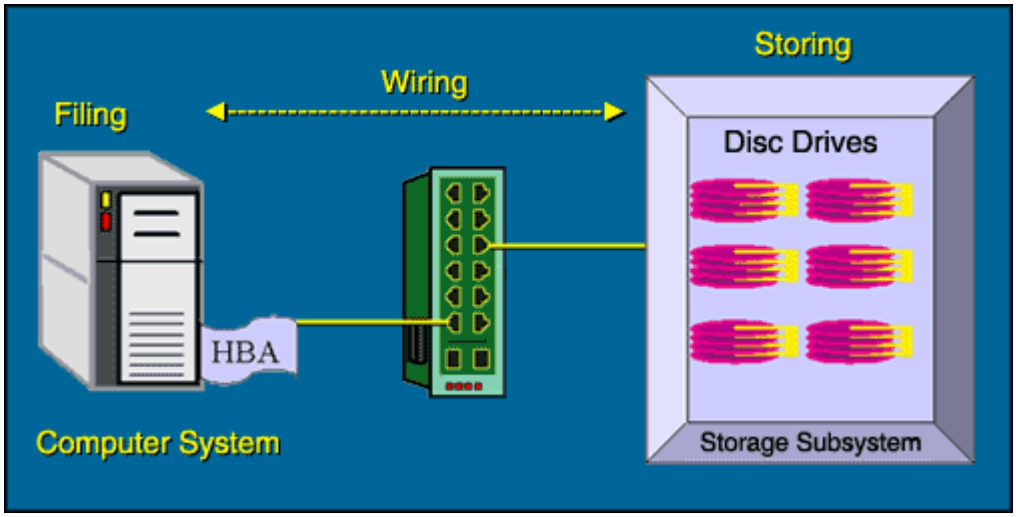

Figure 2: Simple Storage Model

The Storage Market

The storage market is growing. However, it seems that there are various ways to play in this game—as a switch provider, a storage device manufacturer, provider of RAID products, a managed or hosted storage provider, or some combination of all of these. HP and EMC are the clear leaders

today, but other names come up that are not represented on the chart below are Brocade (with 90 percent of the fabric switch market), McDATA (a major player in the Director switch market and acquired by Brocade in January 2007), Veritas (a player in the storage management space and owned by Symantec), and Cisco and IBM (they see it as a huge opportunity).

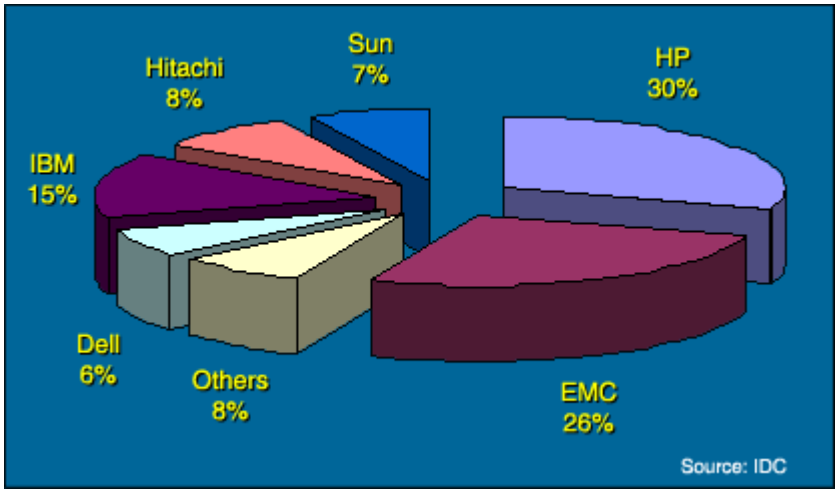

Figure 3: Breakdown of the Storage Services Market

IDC predicts that the market for US storage management services will grow by 5% CAGR through 2007 with worldwide growth for SAN revenue estimated at 11% CAGR through 2007. By some estimates, storage technology, a subset of the overall IT budget, accounts for up to 60% of overall IT spending by U.S. companies. While it might be difficult to pinpoint who is responsible for storage administration and BCP, storage seems to be one of the hottest markets today, despite the downturn in overall telecom capex.

Summary

Business continuity planning is a process whose time has come. The biggest recent change has nothing to do with the technology, but rather a lack of a sense of urgency in the need to invest in deploying a business continuity plan. There has been a radical change in many enterprises following the devastation caused by the natural disasters of 2005 and 2006, along with the declining prices associated with bandwidth and remote storage on something other than tape. With the economy starting to pick up, these strategic topics are on the top of customers' wish lists.

This is all a significant shift in emphasis for IT organizations. There is still much confusion in the market regarding terminology. However, we need to remember the same is true for all new technologies. Over time the industry settles to provide standard terminology and standards for products and services. A good reference is the Storage Networking Industry Association's site at http://www.snia.org. Another good reference is Webtorials and the *Network World* Storage Resource Center.

The following is a high-level view of a simple business continuity plan.

A Plan for Survival

- Assume a total disaster

- Identify critical applications and equipment. (What would be the impact if they were not available?)
- Understand your needs (How fast do your mission-critical applications and data need to be up and running? How current must that data be?)
- Devise a disaster recovery team and determine how to mitigate risks of loss
- Recognize that there are many alternatives and approaches, with no perfect solution
- Test the plan before you commit to everything on that strategy
- Deploy the plan in phases
- Train and educate staff
- Monitor the process
- Test the results over time
- Reevaluate and have regular updates to the plan and process
- Rinse and repeat

Information management is an integral part of business continuity planning as information has become the lifeblood of today's enterprise. Providing secure access to the information from anywhere, at anytime has forced companies to raise the importance of storage considerations from important to mission critical. The information and the access networks are now among the most important corporate assets. Application and information must be available in a 24/7 environment. To meet these challenges the storage concept has been developed, and now expanded. While the concept is relatively new, the storage network market is one of the fastest growth areas in IT. Competition in the area is intense. HP and EMC, the market leaders in enterprise storage systems, are seeing increased competition from IBM, IronMountain, StorageTek, and AT&T. Brocade and McDATA, established switch manufacturers, are expected to see more competition among themselves and from Cisco and Dell, the latest entrants into this market. This is an important area to watch.

This white paper provides a high-level perspective of these exciting areas to help you understand the core elements of business continuity planning and SANs. If you have interesting experiences to share on the subject please forward them to the author at m.steinberg@hill.com.

About the Author

Mark Steinberg is the Director of Business Development and a Senior Member of Technical Staff for Hill Associates. Mark has spent more than 25 years in the business of technology, providing consultative services regarding the strategic implications of technological change.

Business leaders rely on Mark's insights as they determine their business and technology strategies. His knowledge of convergence, both voice and data and wireline and wireless, has allowed him to create a variety of programs that address these business strategies. His engagements have been with companies such as CANTV, Telstra, AT&T, Qwest, Global Crossing, BellSouth, Verizon, Cingular, Ernst & Young, and Sprint—in the U.S., South America, Europe, Asia, and Australia.

A Senior Member of the Technical Staff, Mark has been with Hill Associates since 1994. He is a Certified Information Systems Security Professional (CISSP #49990) and holds an MBA in International Finance from the University of Santa Clara, CA, a BA in Pre-med/Biology from Hartwick College in Oneonta, NY, and has done graduate work in Astronomical Physics at the University of Colorado, Colorado Springs, CO.

About Hill Associates

Hill Associates is a premier provider of Education and Marketing Services. For over 25 years, we have demystified the complex world of telephony, network, security, and data communications for the most recognized players in the industry. Today we offer a variety of programs: leader-led courses, hands-on workshops, self-directed e-learning, and virtual classroom e-learning—all focused on enabling success in this rapidly evolving industry. Our Marketing Services bring objective industry and subject matter expertise to sales events, customer meetings, white papers, and Webinars.

At Hill Associates, our collaborative approach addresses each client's unique needs and challenges. Whether it's strategic executive briefings, programs for field teams or corporate IT, or marketing initiatives designed to bring knowledge to prospects, Hill Associates delivers your solution.

Visit our website (www.hill.com) for more information. While there, check out our newsletter, blog, and podcasts.