Building a Successful Strategy To Manage Data Growth

Abstract

In organizations that have requirements for a minimum of 30 tera-bytes to multiple petabytes of storage the "go to" technology for a successful data storage growth plan is automated tiered storage. Automated tiered storage is sometimes thought of as a solution for high-performance computing environments, but it is a technology that provides tremendous value for mid-size and large enterprise organizations that need to get their data storage infrastructure under control.

BLUE ARC®

There are few sure things in business these days, but here's one: If you are in business and expect to stay in business you can also expect that your need to store and manage data will be growing. In fact, for many businesses, the need to manage increased amounts of data will be expanding in epic proportions.

The research firm IDC estimates that data storage requirements in large enterprises are growing by an average of 60 percent a year. In some data-intensive industries—such as entertainment, healthcare and energy—those requirements are expanding by more than 100 percent a year.

Where Does It All Come From?

In many cases, the increase in demand for data storage and management is a result of fundamental changes in society and in the way we do business. Business is becoming more global and we are all becoming more and more reliant on digital communications and data-intensive applications for everything in our lives.

When you think about Web-based commerce; or medical imaging applications; or rendering for film or design work; or the sheer mass of social networks and ubiquitous digital devices, you realize that not only is the demand for data storage growing now, it will keep on growing for as far into the future as we can visualize.

That is why it is critical that businesses start building a strategy—now—for managing data growth. The key is not to try to limit growth: That would be virtually impossible. The key is to design a plan and build a technology infrastructure that limits the negative impact of data growth and, in fact, transforms data management into a vital strategic resource that can be used effectively for competitive advantage.

How is that possible? Well, if your business is spending less than your competitors on data storage, and is devoting less physical space to data storage, and is devoting less energy consumption and IT resources to data storage, and has a strategic plan in place for managing the future growth of data storage, and is not bogged down by slow IT performance, then the competitive advantages can keep on mounting.

Successfully Managing Data Growth: Getting Started

For many businesses, the first step in designing a successful strategy for managing data growth is to do an inventory of their existing data storage infrastructure combined with an analysis of future data requirements.

As part of this inventory and analysis it is critical to recognize that not all data is of equal importance: Some data needs to be accessible often and immediately; some may be equally critical but not accessed as often; some can be archived and accessed only in specific circumstances, such as for compliance, disaster recovery or archival purposes.

As organizations are going through this process, many are finding a few common themes:

- Much of the growth in data storage requirements is coming from unstructured data—such as audio or video files, CAD/CAM drawings, animation and graphic files—as opposed to the structured data that is found in the corporate database for HR files or business analysis, for example. IDC predicts that unstructured data will outgrow structured data in enterprises by as soon as 2011.

- Many organizations don't have a consolidated, strategic approach to handling the growth in data storage requirements they are managing now and in the future. Businesses often have a variety of different storage devices, such as solid state, Fibre Channel, SAS and SATA based arrays, but don't have an automated system by which to manage and classify the data that is stored on each of these devices.

BLUE ARC®

- Those businesses that do not have a strategic plan for managing data growth are finding that the costs of storage are increasing significantly and seemingly out of control. Even though the cost of storing data has fallen consistently through the years, enterprises are still spending as much as 15 to 20 percent of their IT budgets on data storage, according to IDC. Without a plan in place, that number goes up as businesses keep buying more and more storage devices to handle increased volume.

As businesses go through this assessment process, it is critical that they also go through a process of evaluating technology solutions to understand which technologies can help them solve their business challenges.

Turning To Automated Tiered Storage

In organizations that have requirements for a minimum of 30 terabytes to multiple petabytes of storage the "go to" technology for a successful data storage growth plan is automated tiered storage. Automated tiered storage is sometimes thought of as a solution for high-performance computing environments, but it is a technology that provides tremendous value for mid-size and large enterprise organizations that need to get their data storage infrastructure under control.

Among the advantages automated tiered storage offers to businesses are:

- Reduced actual costs and total cost of ownership
- Increase performance of accessing and storing data
- Spend less time managing data storage
- Scalability and the ability to "future proof" your storage infrastructure
- Consolidation, not only in storing data, but in the amount of physical space taken up by data storage equipment
- Improved energy efficiency across the organization and within the data center and IT resources to data storage, and has a strategic plan in place for managing the future growth of data storage, and is not bogged down by slow IT performance

Perhaps even more important than each of these is the sum of all the parts: By turning to an automated tiered storage solution, businesses are recognizing and sanctifying data storage as an asset that can provide competitive advantage and must be managed strategically.

Where many organizations may initially turn to automated tiered storage as a technology to get costs under control, they soon recognize that it is a key element in their overall strategy to manage their ever-expanding data growth.

Going through the data classification process inherent in building a tiered storage solution can be a defining business initiative that should be the foundation for a long-term strategic plan to manage data growth.

Also, in thinking of data storage as a key strategic resource, IT organizations should focus on functions and features that will support and help them manage data growth. For example, one of the key elements users should require of their automated tiered storage solution is the ability for the IT organization to set policies on common file attributes or access patterns so that the parameters are established by the organization and all data is automatically migrated to the appropriate tier of storage.

It is important as well that the movement of the data be transparent to users and applications, i.e that users continue to access data from the original location where the user stored it, and applications do not have to be "reconnected" after data has been moved. This way, as new data is generated it is automatically moved to the appropriate tier based on its value to the organization, as defined by the organization. It is impossible to understate the importance of data transparency: It is the key to any strategic plan to manage data growth because it means that organizations are really managing data as opposed to just managing devices.

Choosing A State-Of-The-Art Network Attached Storage Solution

In the world of automated tiered storage, however, not all solutions are created equal. In choosing a solutions approach, it is important that the IT organization establishes priorities for data—based on the needs of the businesses—and maps out a strategy that addresses data growth for both current requirements as well as future needs. In addition to data transparency, the key considerations should be:

Support for multiple drives and protocols: Most IT environments are neither heterogeneous nor permanent. In fact, most businesses have a storage infrastructure that has multiple types of drives and multiple protocols, so the ideal automated tiered storage solution would support them all and bring them all together.

User-defined policies: As noted, this is critical in establishing a long-term solution to deal with growing data storage requirements. In a recent survey by Storage Magazine, 42 percent of IT professionals said classifying data is their biggest pain point related to their tiered storage system. Having an automated tiered storage solution that provides IT with a large palette of options can provide huge benefits to the organization and significant cost savings.

Performance: Data should be moved seamlessly across the different tiers and users should be able to access migrated files at any time. The overall performance of the storage infrastructure should not be degraded, even though the organization may be consolidating the number of storage devices. In the world of automated tiered storage, performance is an integral part of the solution to ensure that SLAs are met.

Manageability: You should be able to manage and deploy different types of storage within the same network-attached storage (NAS) environment. You should also be able to manage all the diversity—from immediate mission-critical data to infrequently needed archival data—all from a single NAS platform.

There are other key criteria for choosing the right automated tiered storage solution as well as the right vendor, based on the needs of the business and the priorities of the IT organization. In the right environment, with the right solution, automated tiered storage can address run-away storage costs, physical space shortages in the data center and energy consumption issues.

The key is to develop a plan and a strategy, recognizing that the growth you're seeing now in unstructured data is really just the beginning of a long-term trend. By acting now and putting a plan in place, organizations can build the foundation of an infrastructure that strategically manages data growth and treats data storage as a true corporate asset.

Choosing A State-Of-The-Art Vendor

Just as not all solutions are created equal, not all vendors are equal as well. Within the field of automated tiered data, BlueArc has emerged as a market leader and pioneer. Because of the company's vast experience in providing high-performance NAS solutions, it has been able to develop a tiered data architecture that delivers exceptional performance without compromising scalability, bandwidth or capacity.

BlueArc's file system SiliconFS™ is an embedded product that serves as the common point of integration for all elements of a storage solution, including storage servers, information services and the storage ecosystem. BlueArc's approach provides maximum flexibility because it enables businesses to mix and match storage to achieve an optimal blend of performance, capacity and cost.

BLUE ARC®

In addition, BlueArc's approach to tiered storage delivered through its state-of-the-art BlueArc Data Migrator™ is focused on the data itself and not on the physical media. This means that users do not have to be notified of the new data location and applications don't have to be "reconnected" to the new location. The business benefits from this approach because it gains broad flexibility in establishing user-defined policies and in making those policies transparent to users and applications within the organization.

The ability to provide data transparency is a key differentiator and one of the factors that make the BlueArc solution both more customer-friendly and more "future proof." For business customers thinking about creating a long-term plan to manage data growth it is important to understand the key points of difference between solutions and to ask vendors about these important features such as data transparency, performance and scalability.

As you move toward building a strategy for managing data growth and you define the need for automated tiered storage, you want to make sure that any solution you implement will help maximize the efficiency of your current infrastructure. In addition, you want a solution that will provide you with the robustness and flexibility to manage your existing and future data in accordance with its value to the organization at any given point in time.

About BlueArc

BlueArc is a leading provider of high performance unified network storage systems to enterprise markets, as well as data intensive markets, such as electronic discovery, entertainment, federal government, higher education, Internet services, oil and gas and life sciences. Our products support both network attached storage, or NAS, and storage area network, or SAN, services on a converged network storage platform.

We enable companies to expand the ways they explore, discover, research, create, process and innovate in data-intensive environments. Our products replace complex and performance-limited products with high performance, scalable and easy to use systems capable of handling the most data intensive applications and environments. Further, we believe that our energy efficient design and our products' ability to consolidate legacy storage infrastructures, dramatically increases storage utilization rates and reduces our customers' total cost of ownership.

BLUE ARC®

BlueArc Corporation
Corporate Headquarters
50 Rio Robles Drive
San Jose, CA 95134
t 408 576 6600
f 408 576 6601
www.bluearc.com

BlueArc UK Ltd.
European Headquarters
Queensgate House
Cookham Road
Bracknell RG12 1RB, United Kingdom
t +44 (0) 1344 408 200
f +44 (0) 1344 408 202