Data Mobility Group

**John Webster**
jwebster@datamobilitygroup.com
**November 12, 2004**

**Research Perspective**

# Storage Consolidation—Hidden Issues and Rewards

The value behind an IT consolidation project is conceptually easy to see: Fewer and larger servers as well as storage systems means easier management and potentially lower overhead, compared to a distributed architectural model.

In fact, one of the biggest pluses of a consolidation project has to do with software licensing costs and support fees, which IT administrators commonly identify as leading budget busters. Simply put: fewer servers and storage arrays means fewer software licenses to acquire and maintain, and that can translate into significant cost-savings. In this case, less is definitely better.

However, as enticing as consolidation projects may seem, there are some hidden risks and issues associated with them. For example, IT administrators need to consider the projected growth rate of data over the long term and whether the new architectural model will be able to meet corporate-dictated QoS levels as the IT infrastructure grows in size, capacity, and complexity.

Assume, for example, that the proposed consolidated architectural model is expected to support a number of critical applications over the next five years and that data processing volumes will grow by at least 60% per year over that period, the question IT administrators need to ask themselves is: Will the consolidation model scale to meet these demands? This will be a particularly critical question if data volume growth is well in excess of 60%, as is the case in many IT environments.

"From experience, we know that technology hurdles are often easier to overcome than people and organizational issues. The same holds true for consolidation projects."

We believe IT administrators contemplating consolidation projects should first determine the expected life cycle of the new architectural model over 3, 5, 7, and even 10 years and then determine whether or not the model will be able to evolve over time to meet the changing business and data environment.

This paper addresses some of the key issues that IT administrators need to consider before they consolidate their IT environments, as well as the major drivers of change—both business and IT-related—that will affect the new architectural model over its projected life cycle.

## Life-cycle Challenges

IT environments are never static; they must be flexible enough to respond to changes that are forced by business consolidations and restructuring, variations in the business climate, competitive pressures, and, most recently, regulations.

Also, it would be a mistake to assume that the only challenges IT administrators will encounter in implementing these plans will be technology-related. From experience, we know that technology hurdles are often easier to overcome than people and organizational issues. The same holds true for consolidation projects.

Database administrators, for example, may be happy about managing their own "silos" of server-attached storage, but blend these silos in with others in a networked-storage environment and you may have a problem—one that must be addressed before the project can move forward.

Application users must also be taken into account. A storage consolidation project will change—and challenge—the way IT administrators relate to user groups. Chargeback agreements, service level agreements (SLAs), policies regarding data protection and security will all be affected and must be dealt with at project inception as well as over the life cycle of the storage architecture.

Burgeoning data growth has become a fact of life for most storage administrators. Growth rates range of 50% per year are typical, and rates of 150% per year or more are not uncommon.

## SGI

If you are in the planning phases of a major consolidation project—or just contemplating one—enlisting the help of a preferred vendor is a good first step. Going it alone—as others before you have—is inadvisable.

Silicon Graphics, Inc. (SGI) has been supporting high-volume, data-intensive production processing environments for over a decade. Whether your consolidation project centers on NAS, SAN, or a combination of the two, SGI could be an invaluable source of information and technology.

SGI's consolidation offerings are anchored by their filesystem, XFS, which provides a single data format across NAS, SAN and data lifecycle management, and is architected to scale to 18 million TB. This approach allows SGI customers to scale capacity, performance and connectivity and even blend or move between storage architectures without massive disruptions like reformatting data, throwing out appliances and buying new ones, or simply running out of room. This single data format is extended under CXFS, the InfinteStorage Shared Filesystem, to IRIX, Linux, Solaris, AIX and MAC OSX systems. CXFS allows from 2 to 64 systems to instantly access the same data without requiring copies or network mounts—an approach that SGI customers appreciate for its significant productivity enhancement, disk utilization improvement and centralized management benefits.

SGI provides their consolidation storage hardware and software products separately, or packaged as in its InfiniteStorage solution Platforms such as the NAS 2000, 3000 and Gateway, and the SAN 2000, 3000 and Gateway models with an entry price as low as US $20,000. SGI's SAN Server comes with complete SRManagement capabilities including storage device and file system management, device configuration, performance monitoring, application performance visualization, provisioning, capacity planning, and reporting, both historical and real time.

Options are available for both NAS and SAN configuration for snapshot, mirroring, high-availability clustering, and data lifecycle management. For more information on SGI InfiniteStorage products and solutions, see ***www.sgi.com/storage.***

The present need to retain data under management for longer periods of time is one reason for these tremendous data growth rates, sometimes ten times longer than formerly required. However, we believe a number of emerging pervasive data technologies e.g. RFID tags and readers, multifunctional cell phones and wireless-enabled PDAs, digital video recording devices, and environmental sensors, will drive enterprise data storage requirements into the multi-petabyte range in the not-too-distant future.

In a word, architectural scalability is absolutely essential. In fact, "scalability" doesn't really adequately describe what is required. The new consolidated architectural model must allow administrators to grow storage capacity as needed *and* provide consistent, or better, levels of service to application users.

The problem is that while this type of model can minimize downtime, improve server levels, and eliminate unforeseen spikes in capacity usage, it can also leave storage administrators wondering if they've paid too much up-front. Over-provisioning is a very real concern.

While data requirements can be accommodated more easily if the resources needed to non-disruptively absorb that growth are in place, IT administrators feel that additional capacity should be added on an as-needed basis rather than bought up-front since the cost of a unit of hardware performance (expressed as "cost per GB" or "cost per switch port") typically declines over time. But a scalable architecture must be able to do more than just meet future capacity requirements; it must also be able to integrate and interoperate seamlessly with other IT resources, provide adequate performance, and be adequately protected. And, it must do these things at a price that can be justified to senior executive managers who are highly cost conscious.

Corporate mergers and acquisitions have become commonplace today, and the success of these mergers hinges partly on the ability of management teams to integrate "unlike" systems. For this reason, open standards—i.e., standard that promote interoperability and parallel operation between unlike devices—should be built into the new architecture right from the beginning.

Similarly, storage administrators must address performance and data protection issues from the get-go. As the data store grows, the stress on the existing data delivery and access network also increases. This stress can show up in the form of network performance variability, which can ultimately impact user productivity. Storage administrators must factor in additional users and applications over the life cycle.

Meanwhile, if the architecture's backup and restore functions fail to support the growing data management environment, data protection can become an issue. While it is undoubtedly easier—and even cheaper—to manage data in consolidated rather than distributed environments, storage administrators must make sure that the part of the new infrastructure tasked with data protection can scale upward in terms of backup performance and responsiveness disasters as data volumes grow. Failure to do so puts data at risk.

## Data Protection vs. Archive—DLM Done Right

Backup rules are changing, and IT administrators should update their data protection strategies to meet these evolving requirements. New federal and state regulations, for example, dictate extended retention periods for a variety of data types, such as Word documents, email messages, voicemail messages, and digitized images. The result is a complicated data life-cycle management (DLM) environment—one that requires IT administrators to think outside of the traditional data protection and archive box.

Many IT administrators mistakenly think of backup and archival as interchangeable processes when they are actually distinctly different. Moreover, the two can be implemented separately or combination as part of a larger DLM strategy in a consolidation project.



76 NorthEastern Blvd. Suite 29A Nashua, NH 03062 Phone: 603.563.8384 Fax: 603.563.8282 datamobilitygroup.com

With DLM, the purpose of backup does not change: to make a copy of a volume, LUN, or file at a particular point-in-time (PIT) for future use in the event the original volume, LUN, or file is corrupted or deleted. (In a DLM environment, the backup copy used to restore data in the event of a failure or disaster of some sort can be kept on disk, tape, or optical media.)

The purpose of an archive is not to restore data in the event the data is accidentally deleted or corrupted but to help IT administrators logically group and manage growing volumes of "inactive" data that need to be kept for potentially long periods of time. Keeping inactive data on high-performance storage resources for extended periods of time can significantly degrade application performance.

The goal is to design a consolidated storage architecture that extracts inactive data (and associated table definitions, indexes, and relationships) and moves it to higher-capacity, lower-cost storage devices. Data from these devices can be restored by manual or automated means, when needed.

Keeping backup and archive data separate from one another on tiered storage devices can have significant benefits: Not only can it lower overall storage costs (because data is stored in appropriate storage tiers), but it can also streamline recovery processes in regulatory compliance or legal situations.

The bottom lines is that the new consolidated architecture must respond to IT's need to protect data as volumes grow as well as business and regulatory compliance requirements.

## Managing the Managers

Outside the realm of IT, departmental managers generally like storage that is captive to their servers. So, for many, a consolidation project built around networked storage can upset the data order of the universe, if you will, and introduce new uncertainties. Database administrators, in particular, may strenuously resist changes out of fear that their jobs will become harder to do. Consolidating enterprise storage also often means consolidating enterprise storage user groups. Any way you slice it, it means a brand-new management paradigm!

To make the project seem less threatening, IT administrators must make the new storage environment as transparent as possible to these user groups. Being able to see the new architecture, understand how it works, and, most importantly, grasp its value makes the process easier to accept.

To do this, storage administrators must be armed with the facts! The Computer Measurement Group's motto is "You can't manage what you can't measure." The same applies to consolidated storage environments. Enter management applications software or storage resource management software (SRM). This type of software can be used to ensure that IT organizations are meeting critical SLAs.

Essentially, SLAs define a quality of service that must be maintained to justify charging particular user groups for services. SLAs demand constant monitoring and statistical data gathering of key storage services, including capacity usage, performance, and uptime.

The tools can also be used to calculate total cost of ownership (TCO) and return on investment (ROI), as related to the enterprise storage domain, and justify storage-related expenses to upper management in a predictive fashion or, in the case of ROI modeling, to support an acquisition. We believe both models should be used on an ongoing basis.

## Conclusion

The decision to consolidate is far from a no-brainer. While the case for consolidation is built on many factors, the most compelling reason for this type of data environment is centered on doing more with less. IT administrators can deliver more and better services to users with proportionately less IT staff time devoted to storage and data management with consolidated environments.

As for the inherent risks of this type of project, with a little extra effort and a little due diligence, IT administrators can expose potential problems and, ultimately, implement storage models that will have them reaping benefits over the long term.