WHITE PAPER

Infrastructure Simplification for Tiered Storage: IBM's TotalStorage Productivity Center

Sponsored by: IBM

Richard L. Villars   Rhoda Phillips
October 2004

EXECUTIVE SUMMARY

More than ever before, companies are dependent on the effective use of information to survive and grow. They also depend on easy access to and effective use of many different types of information. The growth in the volume of information required, however, is only part of the problem. The types of information that companies now need to create, manipulate, and store continue to proliferate, often leading to underutilization of current storage capacity as it is dispersed across multiple applications.

In addition, companies are more concerned than ever with ensuring both short-term and long-term access to all of this information. Accommodating the growing array of different information types while boosting overall business continuity goals poses significant challenges for today's budget-conscious IT executives. In this changed circumstance, boosting the effective utilization and integration of existing storage assets as well as improving the manageability of those assets are now primary goals.

This IDC white paper examines emerging storage technologies that allow companies to simplify their information infrastructure and reduce administrative costs through the deployment of a tiered storage infrastructure. It discusses how the evolution of tiered storage technologies will affect existing storage management products and processes.

IDC also assesses IBM's TotalStorage™ hardware and software products as components in companies' evolving storage deployment plans and evaluates IBM's recently introduced TotalStorage Productivity Center as part of a solution for boosting storage utilization and better managing future capacity expansion.

STORING AND PROTECTING INFORMATION IN TODAY'S ENTERPRISE

Large or small, public or private, companies today depend upon easy access to and effective use of many different types of information to find new customers, meet the needs of existing customers, and make sound decisions about future investments. The sheer growth in the volume of information required, however, is not the only concern for companies. The real challenge is the range of different types of information that companies now need to create, manipulate, and store on demand in order to succeed and grow.
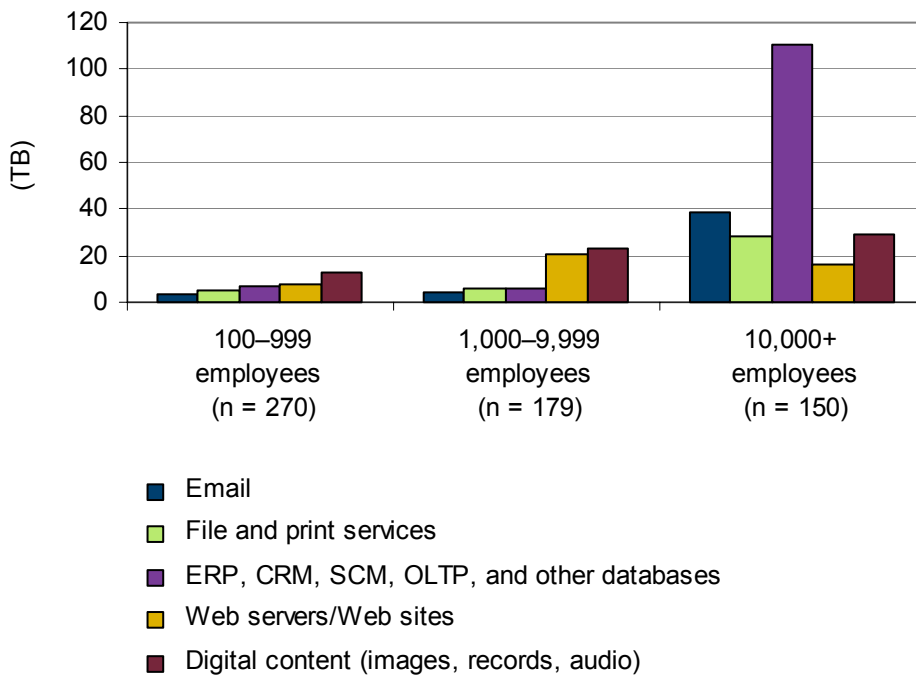
Expanding Range of Information Needs: New Applications and Data Types

In a recent survey of 833 *InfoWorld* subscribers who are IT and business managers at medium-sized and large U.S. businesses, IDC asked respondents to estimate the storage capacity currently allocated for different types of information. The size and range of data types that drive companies' storage investments are quite diverse, regardless of company size (see Figure 1).

FIGURE 1

Companies Use of Storage by Information/Application Type and Company Size

Q. What is the estimated total disk storage capacity, in terabytes, for the following kinds of data at your company, excluding disk drives on desktop systems?

- Email
- File and print services
- ERP, CRM, SCM, OLTP, and other databases
- Web servers/Web sites
- Digital content (images, records, audio)

Source: IDC's *Trends in Storage Survey*, 2004

It is not surprising to see significant allocations of storage for database applications and general file services (application types that have been around for decades). What is surprising for many business executives is the amount of storage allocated to recently introduced applications such as email, Web servers, and digital content (e.g., audio and image files).

This increasingly diverse array of applications poses a significant challenge for IT managers who must buy, maintain, and protect the storage assets used to store their companies' business information. Many installed storage systems are optimized to meet the performance and reliability needs of transaction-intensive applications. They are often too expensive and finely tuned to simultaneously support digital content or email applications where rapid capacity expansion and manageability are more critical concerns.

The consequence of this discontinuity is the proliferation of installed storage arrays with many different price, performance, and capacity characteristics. Each may support a particular class of business application, and each may require its own unique administration and data protection processes. As the information within these applications evolves and ages, the cost and capabilities of isolated storage systems do not match the value of the data being stored.

The effect of this discontinuity is further compounded by the relatively poor storage utilization rates (often under 40%) associated with the dispersed servers that support these applications. Many storage infrastructures lack the ability to narrow the gap between information value and storage system capabilities by using simple means.

Both of these problems translate into wasted capacity and reluctance to add additional capacity in times of tight budgets.

## Growing Concerns About Business Continuity: Information Availability and Reliability

Concurrent with the expansion in information types, companies are more concerned than ever with ensuring both short-term and long-term access to information. In the survey mentioned earlier, IDC also asked companies to identify their top two priorities relating to storage investments in the coming year (see Figure 2).

Companies Top Priorities for Storage Investments in the Next
12 Months

*Q. Which of the following are your company's two top priorities relating to storage for the next
12 months?*
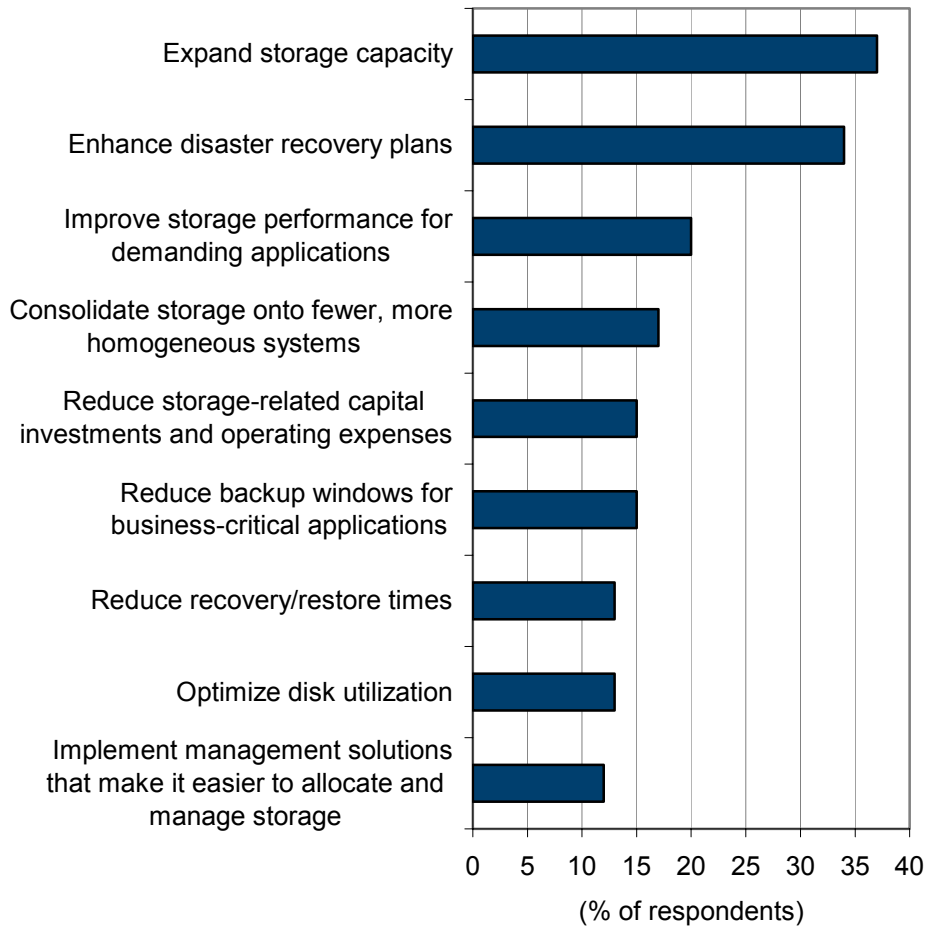
n = 492

Source: IDC's *Trends in Storage Survey*, 2004

Only two items were identified as a top priority by more than 30% of the 492
respondents: expand storage capacity (37%) and enhance disaster recovery plans
(34%). Several of the less frequently cited priorities relate to these two issues as well.

Both the consolidation of storage onto fewer systems (17%) and optimizing disk
utilization (13%) are specific strategies for expanding storage capacity at less cost,
while reducing recovery/restore times (15%) and reducing backup windows for
business-critical applications are important elements in any disaster recovery plan.
These results indicate that boosting storage utilization and ensuring data integrity are
increasingly critical concerns that change the dynamics of companies' storage
investments.

The range of information that companies need to protect is expanding rapidly, and the pressure to ensure adequate protection for all corporate data continues to rise. In the past, concerns about business continuity centered on data replication for critical information systems including online transaction processing (OLTP) with a focus on rapid recovery at both local and remote sites. In today's enterprise, IT managers must deal with many different data types that require different data protection strategies.

☑ IT managers can't just focus on remote disaster recovery; they must coordinate multiple backups for interconnected applications while dealing with ever-shortening backup windows.

☑ IT managers can't just focus on retention of transaction data; they must implement verifiable retention for many types of data.

This more complex data protection environment is a source of complexity, exposing IT managers to greater chances for error, more inefficient operation/cost, and the risk of data loss.

For example, financial institutions are now required to retain all internal and external electronic correspondence (e.g., email and instant messaging) for a set period (typically two years). They must also be able to prove that the correspondence has been unaltered. In this environment, recovery times are important (hours to recovery are acceptable) but data loss at any point for the next two to seven years is unacceptable, subjecting the enterprise to financial and even criminal liability. In this environment, reducing backup increments, managing data retention policies, and staging data to a remote archive facility are the most important requirements.

IT managers are moving beyond the traditional solution of tape backup in an effort to address these increasingly complex business continuity needs. Today, IT managers employ a number of disk-to-disk replication and tape backup solutions to boost overall business continuity, but the diverse array of storage systems deployed to support different applications makes the creation and maintenance of consistent data protection/recovery policies difficult, if not impossible, for many companies.

Infrastructure Simplification: Maximizing the Value of Storage Investments

Accommodating the growing array of different information types while boosting overall business continuity goals poses a significant challenge for today's budget-conscious business executives.

Companies' expectations about the effectiveness of their IT environment has changed drastically in the past two years. Economic and business uncertainties have forced executives to closely monitor new IT expenditures and maximize returns on previous investments without always considering the long-term consequences of inefficient deployment. In this changed circumstance, boosting the effective utilization and integration of existing storage assets as well as improving the manageability of those assets are now primary goals.

The key to meeting these goals is infrastructure simplification. Companies need to deploy solutions that reduce the complexity of managing existing and future assets while boosting their effective utilization. Simplification is more than just rearranging storage topologies; it's also about implementing management practices that are enabled by a simpler infrastructure. It means:

☑ Automating repetitive tasks

☑ Eliminating application downtime when migrating data from one class of storage to another

☑ Having more flexibility to allocate storage capacity on demand without precluding future upgrades

Based on our ongoing conversations with senior executives in industries ranging from financial services to life sciences to retail, IDC believes these changes in evaluating overall storage solutions reflect a permanent adjustment in attitude. IT executives' approach toward evaluating the value of all information assets and the IT solutions that support them must reflect new concerns for integration, effective utilization, manageability, and availability. These new priorities will remain even as the global economy recovers and companies' overall investment mind-set becomes more liberal.

EXPANDING STORAGE OPTIONS FOR ENTERPRISES

In an effort to meet the sometime conflicting application and cost containment demands of business executives, IT managers began to look for storage solutions that addressed specific IT requirements (e.g., better performance, higher reliability, greater capacity, or lower cost). In response, the industry has seen a proliferation of storage systems and supporting technologies that address specific customer requirements.

Expanding Storage System Options

IT managers can now choose from an array of different storage systems, including:

☑ High-end monolithic storage arrays that provide high performance and reliability in a single, centrally managed system

☑ Midrange storage arrays that allow IT managers to add storage capacity for applications in smaller increments

☑ Low-end arrays that enable the sharing of storage capacity across multiple servers in small businesses and remote sites

In addition, the past year has seen the emergence of a new class of storage systems that leverage lower-cost and larger-capacity disk technologies (e.g., Serial ATA) to significantly boost array capacities while reducing capacity costs. While not appropriate for all application workloads, these capacity-oriented technologies provide IT managers with greater control over costs when allocating storage for specific needs.

Today, incorporating capacity-oriented storage systems into existing environments is a high priority for companies. IT managers want to leverage these new systems for specific requirements:

☑ More cost-effective storage of proliferating fixed content assets such as digital images, large reference files, and email archives (content-aware storage)

☑ More rapid and frequent backup/recovery of business-critical information at both local and remote datacenters (disk to disk to tape)

The Rise of Networked Storage

Concurrent with the expansion in storage array options, IT managers are connecting application servers and storage systems via storage area networks (SANs). SANs help accelerate the consolidation of server, storage, and tape assets by detaching storage devices from individual servers and applications while simultaneously providing a platform for centralized control of those assets.

The combination of SANs and multiple storage systems is also spurring the next step in infrastructure simplification: the development of solutions that allow IT managers to pool storage resources across multiple systems and applications while shielding server and application administrators from the complexities of configuring storage systems and managing tape backups.

These new solutions introduce layers of logical abstraction between the physical blocks of data on storage devices and the volumes/files that servers and applications access. Continued improvements in performance, cost, and reliability of storage networking technologies now make it possible to move functions traditionally handled by general-purpose servers or array-based controllers into network-based systems. These functions include LUN/volume management, data replication/backup services, and file/database management services.

Building a Tiered Storage Infrastructure

Companies of all sizes that employ these new storage solutions are consciously and unconsciously constructing a tiered storage infrastructure.

They continue to deploy storage systems that deliver many different classes of storage ranging from high performance/high cost to high capacity/low cost. Through the deployment of SANs, many of these storage assets are now physically connected to servers that run many different types of applications and create many kinds of information. Finally, with the arrival of network-resident storage services for distributed management of volume, files, and data replication, IT managers can intelligently provision, reallocate, and protect storage assets to meet the needs of many different applications across the network instead of device by device.

Results from IDC's survey of IT and business managers reinforce the growing interest in purchasing solutions that support the development of a tiered storage infrastructure. Adding support for low-cost/high-capacity storage arrays and common data replication services were both options that respondents wanted their storage array suppliers to deliver in the next 12 months (see Figure 3).

FIGURE 3

Capabilities That Companies Want from Their Storage Array Vendors

Q. Which two of the following do you want your current storage array vendor to deliver in the next 12 months?

■ 100–999 employees (n = 270)
□ 1,000–9,999 employees (n = 179)
■ 10,000+ employees (n = 150)
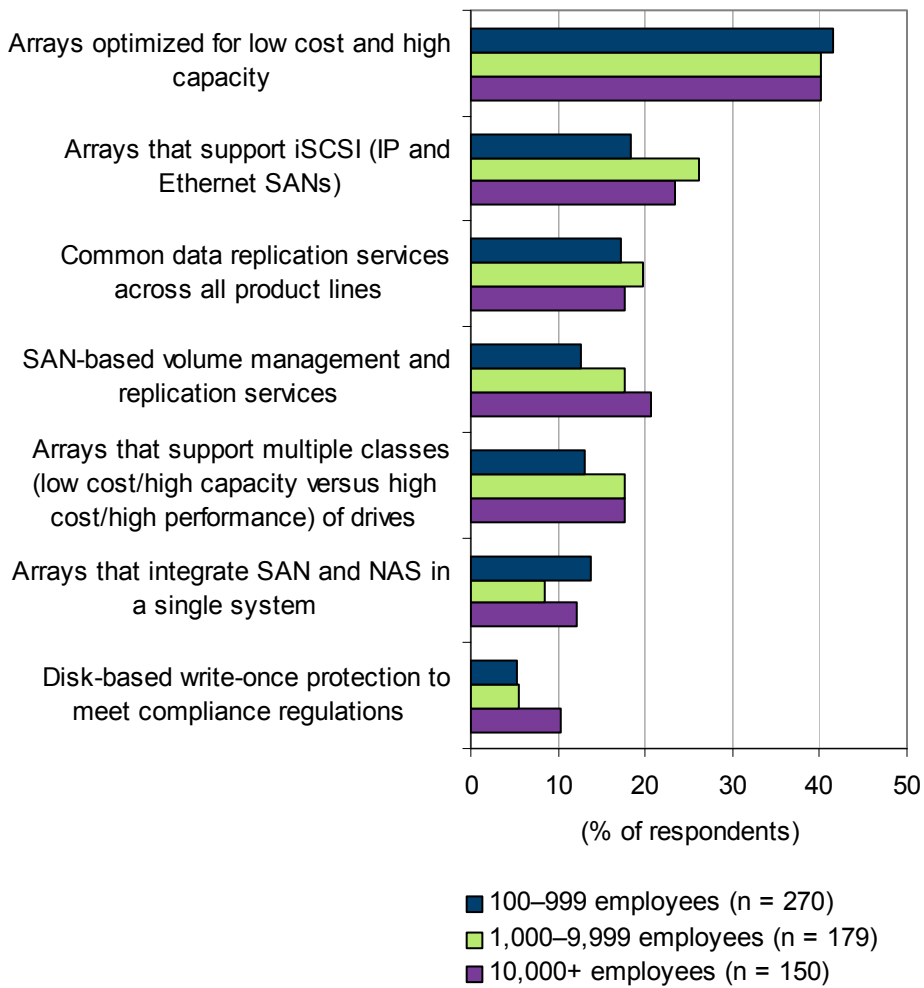
Source: IDC's *Trends in Storage Survey*, 2004

MANAGING IN A WORLD OF TIERED STORAGE

Over the next two years, turning the idea of tiered storage into a usable reality for IT managers will be an important goal for the storage hardware and software industries. For IDC, the most important idea that IT managers should focus on as they evaluate and deploy solutions within a tiered storage infrastructure is that these solutions only work when used together. In today's storage world, that remains a daunting goal.

The historic sensitivity of the connection between storage systems, SAN switches, and storage software has always translated into customer demands for extensive testing of different combinations of arrays, tape systems, servers, operating systems, and applications. In today's world of diverse server platforms, operating systems, SAN infrastructure, and storage products, the testing required to confirm all combinations is growing exponentially. IT managers face the daunting task of managing a growing array of storage-related devices, network controllers, and storage applications.

If not managed properly, this diversity translates into growing administrative overhead and reduced interoperability.

Laying a Management Foundation for Tiered Storage

The Storage Networking Industry Association (SNIA) has long championed the development of standard processes and protocols. The SNIA storage model helped define the logic behind many of today's new technologies, including volume management and data replication.

SNIA's most important near-term effort is to drive the deployment of the Storage Management Initiative Specification (SMI-S) for monitoring and managing all storage-related systems. Storage vendors of all stripes have already begun to deliver SMI-S within their storage systems and SAN infrastructure products, and more will follow in the coming year.

Today, many IT managers don't list SMI-S support as a high priority for their storage systems suppliers, but many of the capabilities listed in Figure 3 depend upon the use of this standard. For the first time, IT managers will be able to collect and access common information for storage systems throughout the enterprise. Such information is a powerful resource for better managing storage capacity, ensuring consistent data replication, and limiting the impact of inevitable failures.

IT suppliers must deliver solutions that employ the SMI-S data provided by multiple devices (e.g., SAN switches and network storage controllers) to provide common discovery, performance, and configuration information to IT planners and administrators.

Centralizing Control of Storage Services and Processes

Delivering effective management for tiered storage requires more than common control of the physical devices. The true value of tiered storage is that it allows IT managers to provision and define storage at a logical level as well. IT managers can create classes of storage based on performance, capacity, cost, and level of data protection across multiple physical systems and then assign application resources based upon their unique requirements.

This is also the level at which IT managers can apply consistent rules related to regulatory compliance requirements, such as:

☑ **Service availability:** ensuring applications meet set availability requirements through judicious use of data replication and data protection solutions

☑ **Information access control:** ensuring that access to information is controlled at both the user and application levels

☑ **Data retention:** ensuring the accurate retention of specific information sets for set time periods

At this logical level, the administrative concerns for IT managers shift more to areas such as process automation, capacity planning, policy management, and application-level performance monitoring. Effective tiered storage management solutions must deliver capabilities that enable:

☑ Centralized volume and file service creation and reconfiguration

☑ Cross-system data replication and protection management

☑ Workflow systems for identifying and automating repetitive administrative tasks

Keys to Delivering Effective Management Solutions for Tiered Storage

The advantages of reusing existing storage assets along with new storage systems within a tiered storage infrastructure hold great promise for IT managers struggling to support new types of information while improving information availability and holding down storage expenditures. Delivering on the promise, however, requires the providers of such solutions to address several important customer concerns.

Emerging solutions based on these new architectures must do the following:

☑ Deliver robust physical device management capabilities that work across heterogeneous, multivendor storage systems and complementary SAN switches

☑ Ensure flexibility in administrative options (centralized or distributed) without sacrificing consistency and commonality of functions

- ☑ Enable the limited deployment of specific management functions cost-effectively without sacrificing the ability to quickly add more functionality

- ☑ Work effectively with existing system and network management systems

Equally important, these solutions must deliver on these promises without:

- ☑ Requiring wholesale replacement of existing management tools and processes

- ☑ Reducing the reliability of existing data protection processes

- ☑ Introducing additional management complexity that overwhelms existing staff resources

IBM'S TOTALSTORAGE PRODUCT FAMILY

One company that is heavily involved in the delivery of storage hardware and software solutions that support a tiered storage infrastructure is IBM. The company provides solutions through its IBM TotalStorage family of hardware and software products and through a number of management solutions from IBM Tivoli.

TotalStorage hardware products provide companies with storage options ranging from highly scalable and reliable ESS (Shark) arrays for high-end datacenters to the modular FAStT family of arrays. The FAStT products also include support for capacity-oriented Serial ATA drives that are of much interest to IT managers.

IBM also sells and supports SAN fabric switches and Directors to connect its storage systems to servers ranging from the largest mainframes to low-end Intel servers. These include products from Brocade, Cisco, CNT, and McDATA.

IBM's TotalStorage Open Software Family addresses storage management issues across a number of different areas, including:

- ☑ Hierarchical storage management (IBM Tivoli Storage Manager for Space Management)

- ☑ Archive management (IBM Tivoli Storage Manager for Data Retention)

- ☑ Recovery management (IBM Tivoli Storage Manager)

In the past year, IBM also introduced several solutions that focus on storage and file virtualization as well as heterogeneous data replication. These network-based storage services solutions are:

- ☑ The TotalStorage SAN Volume Controller is a complete software and hardware solution that combines the storage capacity from attached storage arrays into a single reservoir of storage through the use of logical volumes. The SAN Volume Controller allows IT managers to consolidate diverse existing storage assets into an easily shared and consistently managed pool while providing common data replication services across storage devices.

☑ The TotalStorage SAN File System provides an enterprise-class architecture for delivering a SAN-based global name space and file system across multiple distributed storage systems that is usable by all application servers. The SAN File System provides IT managers with improved access and sharing of data, automated policy-based file provisioning and improved utilization, and enhanced backup and recovery.

IBM TotalStorage Productivity Center

As the provider of a broad spectrum of storage hardware and software products that will play key roles in companies' deployment of tiered storage solutions, IBM is also enhancing and expanding its storage management infrastructure solutions to better match the needs of its customers.

In March 2004, IBM announced the IBM TotalStorage Productivity Center as part of an initiative to improve storage utilization availability and administrator productivity through tight integration of diverse storage management applications. The goal of this effort was to deliver a more robust but simplified set of storage management solutions for managing IBM's own storage hardware and software products as well as SMI-S–enabled products from other storage vendors.

Today, IBM TotalStorage Productivity Center components include:

☑ IBM TotalStorage Productivity Center for **Data** (formerly IBM Tivoli Storage Resource Manager)

☑ IBM TotalStorage Productivity Center for **Disk** (formerly TotalStorage Multiple Device Manager — Performance Manager)

☑ IBM TotalStorage Productivity Center for **Fabric** (formerly the IBM Tivoli SAN Manager)

☑ IBM TotalStorage Productivity Center for **Replication** (formerly TotalStorage Multiple Device Manager — Replication Manager)

These products that make up the IBM TotalStorage Productivity Center leverage common database administrative functions and share a centralized suite install. IBM offers flexible licensing options so customers can pick and choose the functions that they want to deploy at any given time without sacrificing the ability to make future enhancements.

IBM TotalStorage Productivity Center for Data

This component focuses on data management and optimization for disks, file systems, and databases that reside within a tiered storage environment. It provides tools for monitoring and tracking data growth and performance so that IT managers can better manage file system extensions and disk provisioning. IBM TotalStorage Productivity Center for Data also includes support for the SAN Volume Controller and SAN File System.

IBM TotalStorage Productivity Center for Disk

In a tiered storage environment, IT managers will need to monitor and configure many different storage devices and applications from a single interface. IBM TotalStorage Productivity Center for Disk provides a common and integrated set of management functions for all IBM storage hardware and software products (including the SAN Volume Controller and SAN File System) as well as other storage products that support the SMI-S management standard. IBM TotalStorage Productivity Center for Disk addresses specific storage management tasks and provides services that help IT managers:

- Provision, allocate, and assign physical and virtual LUNs on storage arrays and the SAN Volume Controller

- Perform best LUN analysis on aggregate pools of storage

- Monitor volume, file capacity, and cache thresholds to boost availability

IBM TotalStorage Productivity Center for Fabric

This component provides IT managers with the tools to more effectively provision, optimize, and diagnose problems in the underlying SAN infrastructure (e.g., SAN switches) that provides the connection between the diverse disk arrays, tape systems, and servers that are part of a tiered storage infrastructure. By integrating Productivity Center for Fabric, Productivity Center for Data, Productivity Center for Disk, and Productivity Center for Replication, IBM is providing IT managers with a more seamless and timely solution for diagnosing performance and availability problems across the entire environment.

IBM TotalStorage Productivity Center for Replication

This component provides services that help IT managers:

- Create and monitor data replication services across multiple arrays and large storage pools

- Develop and manage consistency groups to ensure reliable and consistent data replication across interdependent data sources located on different systems

- Monitor the status of data replication activities to ensure reliable recovery and overall application availability

Challenges for IBM

IBM's new capabilities and the integration of its storage management software products provide several benefits for IT managers. As companies seek to more effectively leverage existing storage investments and take advantage of new technologies within a tiered storage infrastructure, they will need storage management solutions that simplify demonstrative tasks and provide a more coherent view of the overall environment.

The slow evolution of both network and system management solutions in the 1990s reveals, however, that achieving true integration from an IT manager's perspective is not an easy task.

With this announcement IBM is taking the necessary first steps. It is leveraging an evolving management standard (SMI-S) to enable consistent communication with all storage systems. This decision, however, does mean that IBM is dependent upon other storage system and software suppliers to expand their support for SMI-S. Otherwise, IBM's solutions will be significantly less useful in heterogeneous environments. At present, IDC expects SMI-S adoption to expand rapidly in 2004, but it will still take time to become widely deployed in most customer environments.

Beyond leveraging standards, IBM must ensure that its own applications provide more than simple "integration on the glass." Deeper integration includes:

☑ Employing consistent processes and workflows for completing tasks that cross devices and applications (promised by IBM later in 2004)

☑ Leveraging common databases for configuration and performance information

☑ Employing standard tools for job scheduling, reporting, and security

IBM's decision to employ the Multiple Device Manager as the common data collection repository for management information lays a solid foundation for further enhancements in administrator usability and automation.

FINAL THOUGHTS: PROCESSES AND IMPLEMENTATION ARE KEYS TO SUCCESS

IT managers need to maximize their return on investment in all types of information and the storage assets that ensure accessibility to and reliability of that information. In response, providers of storage solutions are delivering capacity at lower prices and in more modular packages. They are also taking advantage of improved networking to simplify companies' storage infrastructure and boost reliability through the use of a tiered storage infrastructure.

IBM offers a comprehensive set of storage hardware and software products to help companies move to a tiered storage infrastructure. It is also enhancing its storage management solutions to enable more effective deployment and use of these products. The IBM TotalStorage Productivity Center products are excellent representatives of this new approach to managing storage assets as a single resource and should be on your list for consideration in any storage consolidation or tiered storage effort.

Successful deployment and use of a tiered storage infrastructure, however, requires more than new hardware and management products. IT managers must shift from today's fragmented environment, in which many different administrators perform

similar tasks on different systems. As companies begin to deploy storage systems as part of a tiered storage solution, they must address a number of questions. IT managers must:

☑ Rethink data ownership

☑ Centralize the setting of data protection and retention policies

☑ Evaluate the reallocation of administrative resources

Implementation Is Key

The most advanced and integrated management solution is of little use if IT departments do not change existing policies and procedures to automate repetitive tasks. Effective implementation, which is always a major concern, is an especially critical issue when considering these emerging solutions. IT managers need to establish rules and procedures for naming conventions, application upgrades, as well as provisioning and replication policies. Choosing the right partner to assist in implementation is, therefore, a high priority.

What should you look for in an implementation partner (e.g., major system/storage vendor or value-added reseller)? The following features are important:

☑ **Understand benefits and pitfalls of tiered storage.** Although tiered storage solutions hold great promise for improving capacity utilization and reducing management costs in the long run, today's economic environment requires a near-term return on investment. Your implementation partner must have well-formed and documented storage management practices that don't limit options while enabling more targeted returns on IT infrastructure investments.

☑ **Support of heterogeneous server and storage environments.** Although pilot deployments may be in relatively homogeneous environments, the real benefit of a tiered storage infrastructure is its ability to consolidate and migrate information across heterogeneous server and storage systems. Your implementation partner must have both knowledge and testing facilities to identify and quickly resolve any interoperability problems.

☑ **Experience in deploying storage networks.** Tiered storage solutions are only as good as the networks on which they are built. Your implementation partner must demonstrate that it can design and deploy a SAN infrastructure that supports current targeted needs without jeopardizing the future expansion of services across the enterprise.

☑ **Incorporation of management into the deployment process.** Ongoing administration/management is the most overlooked and misunderstood issue in most IT deployment plans. Your implementation partner must help you develop a management and provisioning process that both works with your existing management systems and takes advantage of the solution's new capabilities to improve management efficiency.

Companies such as IBM must ensure that IT managers can quickly deploy components of its TotalStorage Productivity Center without inhibiting future efforts to modify storage management processes. To achieve this goal, IBM and its partners must ensure that their professional service organizations are able to help customers use these solutions to boost utilization, improve application reliability, and streamline existing management policies and procedures.