WHITE
PAPER

IBM Information Infrastructure Initiative Aims to Tame the Information Explosion

By Mary Johnston Turner

September, 2008

Table of Contents

Executive Summary

Around the world, IT organizations are struggling to manage, protect and maintain rapidly growing information assets. ESG's research indicates that, without intervention, the majority of organizations can expect their storage capacity requirements to increase at double digit rates annually for the foreseeable future. Getting control of these growth rates—while reducing business risk and optimizing storage management costs—requires IT decision makers to take a process- and services-oriented approach to managing their storage, information protection and information security requirements. Rather than dealing with these challenges by simply deploying more and more storage resources, IT managers need to think creatively about how to use technologies, processes and people to optimize capacity growth rates; respond to external compliance and internal security requirements; and reduce needless data storage, backup and archiving.

Each customer's specific information management solution will be determined by the nature of the organization's business requirements, existing and planned information assets, in-house IT skill sets, and existing storage hardware and information management software resources. In other words, the information infrastructure of the future must grow out of the information infrastructure of today.

In order to effectively respond to short term priorities in a way that will protect investments and improve efficiency of operations and availability over the long haul, customers need to identify the core capabilities and functional attributes required of their information infrastructure—before they select specific products or vendors. By defining requirements in terms of best practices, services and processes, IT decision makers can make more effective investment decisions as their organization's information management requirements become increasingly complex.

To better assist customers in tackling the information explosion, IBM has completed a number of product and service acquisitions to supplement internally developed IBM products and services. The result is a robust storage and information management portfolio that can help customers solve a wide range of current and emerging storage and information management challenges.

The September 2008 launch of the IBM Information Infrastructure initiative builds on many of these acquisitions. It features 40 new product and service releases and major upgrades, as well as a number of integrated solutions that harvest lessons learned from custom engagements with government and Fortune 500 customers. The centerpiece of the Information Infrastructure initiative is a capability-oriented model to help customers more effectively identify, evaluate and implement IBM solutions. These capabilities holistically consider the technologies, policies and processes needed to deliver cost effective, high performance information and storage services across the enterprise. They are:

- Information Availability
- Information Security
- Information Retention
- Information Compliance

IBM's capability-based approach to mapping requirements to solutions allows customers to concurrently address short term challenges such as storage capacity shortfalls and long term needs such as ensuring the ability to migrate information to new applications and platforms as needed. It provides customers with assessments, best practice templates and technology recommendations to help master the specific capability under consideration.

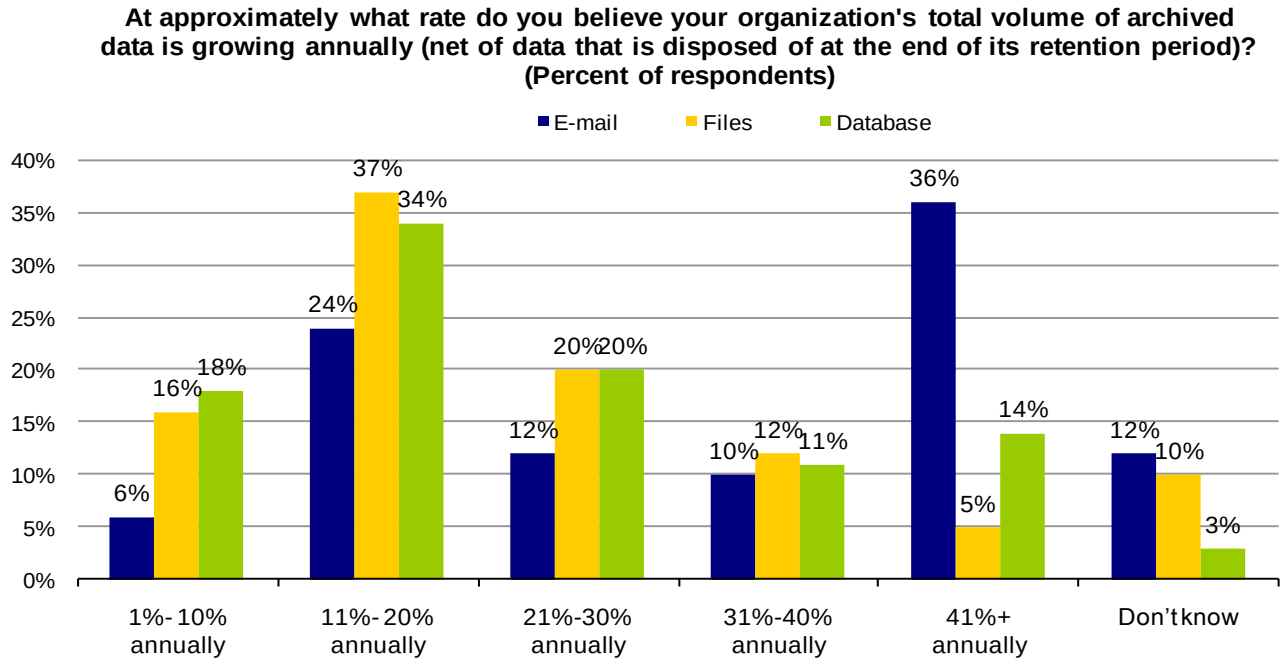
The IBM Information Infrastructure initiative is a new and vitally important pillar supporting the firm's New Enterprise Data Center program. By helping customers think holistically about storage, information service delivery, and automation of best practices, IBM is taking a major step towards enabling customers to cost effectively improve service availability and resource utilization in what is one of the most mission critical and fastest growing data center resource areas.

Information Explosion Demands New Data Center Strategies

Businesses of all types and sizes are generating extraordinary amounts of new information on a daily basis. Applications and authoring tools are becoming more and more user friendly, encouraging almost everyone to create not just presentations, documents and databases, but mash-ups, blogs, and wikis—all using rich audio, video, and image formats. Simultaneously, businesses are expanding beyond traditional markets; extending use of mobile workforces; attempting to reduce the costs of data center power, cooling and facilities; and finding themselves subject to more and more information compliance and legal discovery proceedings. To keep up with these changing business requirements while simultaneously taming the information explosion, IT decision makers need to deploy state-of-the-art storage and information management technologies and tools using best practice processes, policies and automation to promote operational efficiency and reduce business risk.

The business and IT challenges created by the information explosion will not go away on their own or be successfully addressed with the incremental addition of several new terabytes of storage. As can be seen in Figure 1, the vast majority of North American organizations polled by ESG anticipate double digit rates of storage capacity growth for most types of information for the foreseeable future.[1]

FIGURE 1. EXPECTED ANNUAL GROWTH RATES FOR FILE, E-MAIL AND DATABASE INFORMATION ARCHIVING REQUIREMENTS



Source: Enterprise Strategy Group, 2007

Specifically:

- More than one-third (36%) expect **e-mail** requirements to increase at an annual growth rate of **41%** or more
- Almost three-quarters (74%) expect their archiving and related storage requirements for **unstructured files** to increase by 11% or more annually

[1] Source: ESG Research, *2007 Digital Archiving Research Series*

- Fully 79% estimate **database** archiving and storage volumes will increase by 11% or more

Clearly, IT budgets and staff levels can't grow at the same rate as the information resources they support. Rather, IT leaders need to rethink their approaches to all aspects of storage and information management including hardware, software and process automation investments. Storage architectures may not have kept up with the pace of data growth. Tiered storage and storage virtualization technologies may be needed to improve utilization of existing resources and to promote data mobility over time.

Likewise, organizations that could once comfortably accommodate their information protection requirements using tape backup solutions may find they now need more sophisticated archiving and recovery systems to comply with regulatory demands. The opening of remote offices or a major merger may constrain backup windows or create security challenges for centralized storage and backup systems. Data retention policies that rely on saving everything "just in case" may be needlessly driving up storage capacity requirements. Data compression and de-duplication technologies could help the organization save on hardware, power, cooling, facilities and operational costs.

Core Capabilities and Enablers of Effective Information Infrastructure Environments

The information infrastructure—e.g., the people, processes and technologies that combine to deliver timely and secure information services to an organization—is a critical component of today's broader data center environment. Many enterprise data centers have benefited from the introduction of service management best practices such as ITIL and enabling automation tools to support processes such as provisioning, change control and software distribution. Use of a service management approach for data center operations allows IT teams to look above technology building blocks and focus on what it takes to deliver consistent and reliable end-to-end service levels to the business. It also enables IT staff to identify the core capabilities and enabling tools needed to consistently deliver these end-to-end services, even as underlying technologies evolve and change over time.

As organizations begin to re-evaluate their information infrastructure requirements, it is important to apply these same types of principles to storage environments. Specifically, organizations need to think in terms of how to best deliver information services and then implement an integrated, coordinated set of policies, processes and technologies to meet business requirements. This approach needs to be organized around operational priorities and capability areas rather than centered specifically on individual product attributes. Some of the most important information infrastructure capability areas that need to be considered include:

- **Availability and Utilization**: As storage requirements escalate, many organizations fail to consider technologies that could make more efficient use of existing resources. Frequently, storage capacity is not easily shared across applications and users, and storage architectures make it costly and difficult to move information from one media type to another. Dedicated resources and one-to-one disaster recovery capacity allocations also drive up costs and make it difficult for IT staff to operate efficiently. By implementing tools and processes to better share resources across heterogeneous platforms, IT organizations can improve overall information availability cost effectively. They can also ensure that as hardware platforms evolve over time, data can be easily migrated to newer environments as needed.

- **Security**: Many organizations find that significant amounts of data are not consistently protected or secured due to a lack of standardized processes, policies and automated enabling tools. Among enterprise and medium size IT decision makers, ESG's research indicates only 28% are completely confident that their organization's information is fully protected.[2] Mastering information security requires organizations to implement and enforce security and threat assessments and mitigation programs using a wide range of technologies including encryption, access control and identity management technologies—both inside the data center and over the web. Organizations need to be able to track and report compliance and violations with these programs and policies.

- **Retention, Backup and Recovery**: Many organizations create and retain too many copies of individual files and databases in too many places. Copies are made and replicated for disaster recovery reasons

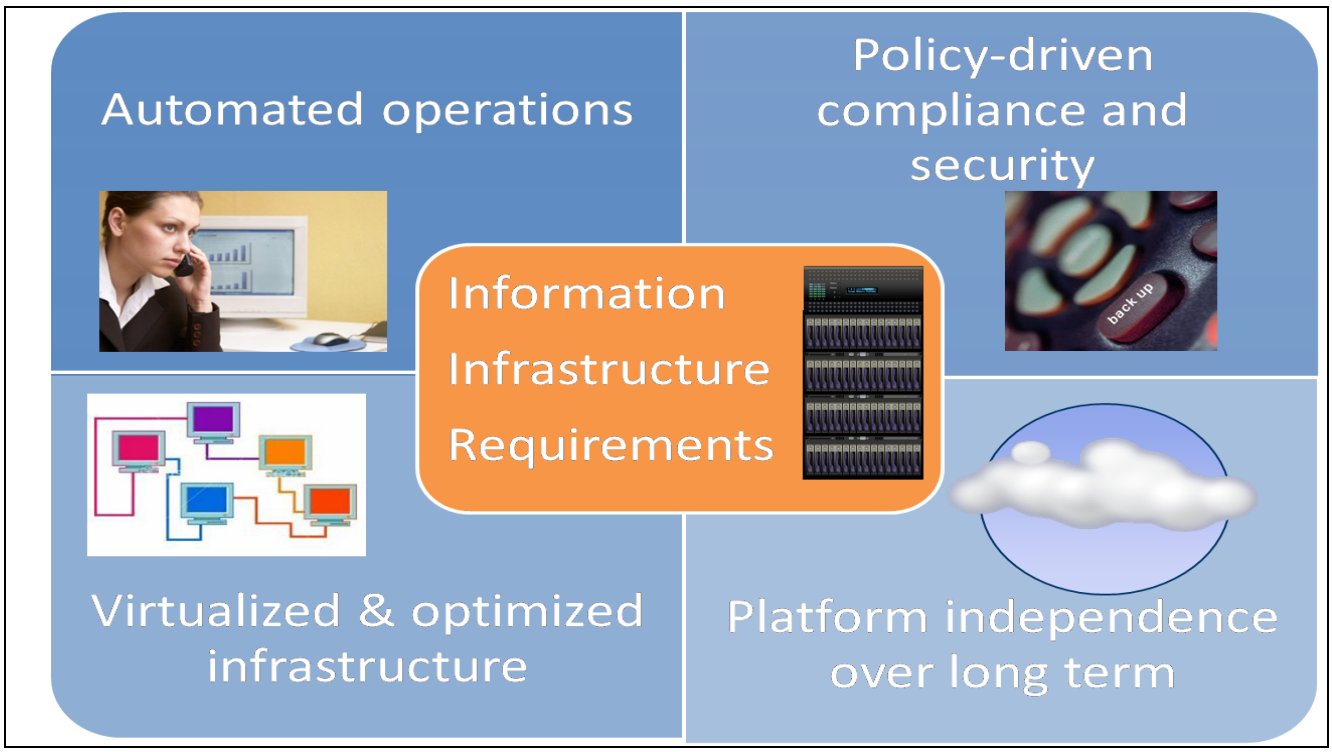
[2] Source: ESG Research Report, *Data Protection Trends*, 2008

as well as for test and development, archiving, compliance and many other purposes. Many files are attached to e-mails and backed up again as part of e-mail system archiving and data protection programs. A lack of consistent information disposal processes and policies often means that large numbers of files are retained much longer than required by business needs or compliance and audit mandates. The net result is excessive use of costly storage resources. Organizations addressing this capability area need to define and implement consistent retention policies and support automated data migration across tiered storage resources. They need to integrate archiving systems and take full advantage of compression and data de-duplication technologies.

- **Compliance:** Requests to quickly respond to regulatory agency information management edicts, legal e-discovery demands and internal audit inquiries are becoming daily occurrences in many organizations. In these types of situations, organizations need to know they can find the information they need quickly and they need to be able to document who touched and changed that information across its lifecycle. While consistent policies—supported by automated tools—can help standardize and enforce business requirements, they need to be defined and implemented in such a way that allows the day-to-day business to remain productive. Achieving success in this capability area requires effective communications with end-users, use of well defined best practices and processes, and reliance on automated, policy-driven tools.

Clearly, the business requirements expressed in each of these capability areas can be supported by a number of technologies, depending on the specific needs of the organization. After beginning the assessment process by defining requirements in terms of information infrastructure capabilities and best practices, IT managers can move on to specifying product requirements in terms of product neutral attributes (see Figure 2)—then implement the best technologies and products available to support these business requirements.

FIGURE 2. ATTRIBUTES OF EFFECTIVE INFORMATION INFRASTRUCTURE ENVIRONMENTS



Source: Enterprise Strategy Group, 2008

This capability-driven approach to requirements definition is significantly different from the way many storage and information management concerns have been addressed in the past. Historically, specific technologies

have been evaluated and selected to solve specific tactical problems, often locking customers into solutions that either failed to meet all their strategic needs or fell short over the long term.

Once IT organizations begin to define and evaluate their storage and information management needs in terms of core capabilities and solution attributes, they are better able to identify the key functional technology requirements needed to support operational flexibility and business growth over the long run. For many organizations, technologies that may merit consideration include:

- **Storage virtualization** to enable seamless resource sharing and data mobility across applications and user groups. By increasing the flexibility of resource allocation, virtualization enables IT staff to more efficiently manage resources while improving overall utilization, performance and resource availability.

- **Tiered storage environments** that include networked storage and a mix of disk and tape platforms to make efficient use of resources.

- **Data migration and mobility tools** to efficiently move information across platforms to support archiving and storage policies as well as to enable seamless migration to new platforms in the future.

- **Data compression and data de-duplication** technologies that reduce the overall volume of data to be stored in the first place. Policies and tools that can efficiently migrate information to lower cost physical resources as information ages also contribute to optimal resource utilization.

- **Policy-based, automated compliance management** tools to reliably apply information compliance archiving policies and facilitate tracking, audit and reporting processes. Compliance management tools also identify records for deletion, which reduces both storage management costs and business risk.

- **Encryption, access management, identity management and other security** tools to protect information from unauthorized access or changes.

- **Automated management and operations** tools to maximize IT staff productivity and improve service levels by reducing human error and downtime. The costs associated with IT staff on ongoing operations are generally much higher than the cost of system acquisition and deployment. Using policy-based management automation helps reduce these costs significantly.

The groundwork put in place by taking a capability-driven approach to defining technology requirements enables IT decision makers to evaluate alternative technologies and vendor offerings in the context of business requirements rather than on the basis of isolated point-solution capabilities and architectures. The process of evaluating competing products becomes much more effective and efficient when decision makers have achieved solid agreements about what matters most to the success of the business.

IBM Information Infrastructure Initiative Helps Customers Navigate a Comprehensive Portfolio

IBM has long offered its customers a significant number of storage-related products and services. However, over the last two years, the firm has made a renewed and expanded commitment to assembling a comprehensive set of products and services to address the full spectrum of emerging information infrastructure requirements. To do this in such a short span of time, IBM has aggressively supplemented its organically developed products with a number of important acquisitions including:

- **XIV:** highly scalable thin provisioning, snapshot and remote mirroring software that supports iSCSI and Fibre Channel, and is optimized for emerging workloads such as Web 2.0 applications and digital media

- **NovusCG:** storage resource management software and services with deep expertise in large scale storage management optimization, workflow standardization and best practices enforcement

- **Diligent:** data de-duplication software that can be integrated with server and storage infrastructures to help organizations reduce the amount and cost of physical storage required in data centers

- **Softek Storage Solutions:** local and remote data protection and data migration software and services
- **Arsenal:** online storage backup and information protection services
- **Internet Security Systems:** security products and services protecting information from known and unknown threats in the datacenter, host systems, network, and endpoints
- **FileNet:** a family of content management software tools that help organizations streamline and automate business processes and records management
- **Princeton Softech (Optim):** enterprise application lifecycle management software optimized for major ERP and CRM suites
- **FilesX:** continuous data protection and recovery software for enterprises and remote/branch offices

Along with Tivoli storage resource management and power optimization software, IBM's internally developed storage-specific professional services, remote managed infrastructure services capabilities, and a wide range of storage hardware platforms, IBM can now address many complex storage-related challenges facing CIOs in a wide range of industries. The challenge for customers, however, is to find the right technologies and services to address their specific sets of requirements and then to implement and operate the chosen solution effectively.

The recently introduced IBM Information Infrastructure initiative aims to help customers efficiently navigate the full breadth and depth of the IBM product and services portfolio in a way that addresses business needs rather than pure technology criteria. As part of this program, IBM is showcasing a number of integrated solutions, as well as the full spectrum of IBM's individual storage and information management product and service offerings. The integrated solutions draw on best practices validated during the delivery of complex engagements with some of IBM's largest customers. Solution templates, configurations and deployment strategies have been optimized for broader use.

The four major Information Infrastructure capability areas are:

- **Information Availability:** Combines heterogeneous storage virtualization technologies, such as the IBM SAN Volume Controller, with data migration tools and a unified management console to ensure data mobility and migration as needed across diverse platforms throughout the lifecycle of the information. This capability area includes end-to-end storage virtualization solutions, including virtualization for disk, tape, large SAN Directors, file systems, and files. It also incorporates automated local and remote site failover capabilities, which define, automate and apply best practices to deliver near-continuous remote site failover across both mainframe and distributed environments.
- **Information Security:** Protects information, maintains privacy, and enables secure information sharing via the use of data and media encryption technologies, access management, identity management, threat mitigation, data security, and web services integration. IBM's research in cryptography, disk, tape, and data center security anchors this capability area. Important solutions include IBM encryption that is built directly into storage devices in order to minimize overhead, and key management software that simplifies encryption key management over the life of the media while enabling secure delegation of authority.
- **Information Retention:** Supports an organization's information retention policies, even when information will be kept longer than the life of the media, by combining tiered storage and integrated archiving products and services with cost effective data de-duplication and compression technologies. IBM's Scale-Out File Services product is a featured offering in this capability area. This solution can quickly search very large directory structures and includes integrated backup. It uses inexpensive servers and disk in a grid configuration, providing affordable scale-out capability.
- **Information Compliance:** Reduces compliance and audit risk via the use of data discovery and policy-based retention tools, protected storage, and information management professional services to provide

best practices and tools covering design and implementation of programs for information retention, destruction, change control and long term access. An example of an integrated solution in this area is the IBM Compliance Warehouse for Legal Control, which includes agents to collect compliance information from e-mail and other repositories, policy-based retention to simplify management of legal hold orders, eDiscovery to automate reporting, and analytics to help investigate suspicious activities. Storage options include policy-based protection to prevent premature deletion and remote site replication to maintain resiliency.

Given that each customer begins at a different point with regard to their existing storage and information management assets—as well as business priorities—IBM and its business partners will work with customers to craft a solution specifically designed for the customer's specific needs.

During the Information Infrastructure launch, IBM offered a number of customer examples for how this capability-based approach to requirements definition and solution development is already paying dividends. The experience of one of these customers, online shoe and clothing retailer Zappos.com, exemplifies the benefits provided by using a capability model to drive definition of requirements and selection of products. The Zappos.com data center team was truly drowning in information as the number of customer and inventory records, online images, and business transactions under management increased daily. Working with IBM business partner Sycomp, the Zappos.com team zeroed in on improving its Information Availability capabilities.

Using IBM-developed assessment and design methodologies, the team determined it needed a significant data center refresh including deployment of automated management systems to keep up with rapidly growing business requirements. Zappos.com selected the N series Gateway with Snapshot software and IBM DS8300 disk storage to provide 25 TB of storage resources. Information availability quickly improved as the firm automated web server deployment, thereby shrinking the deployment process from hours to minutes with more consistent results. By taking a capability-based approach, Zappos.com was able to improve availability by deploying automated management solutions and high performance storage solutions. Without taking a capability-based approach to the problem, the firm might have overlooked opportunities to use automation to improve availability and service levels across the data center.

Waiting Won't Solve the Problem - Time to Take Action

Many customers may feel overwhelmed by the IT costs, risks, and operational impacts resulting from their own organization's information explosion. While decision makers might be tempted to add more capacity to the system and call it a day, they really cannot afford to rely on band-aid solutions. The information explosion will not slow down on its own. ESG's research indicates that information volumes will only continue to increase at double digit rates—unless customers take steps to better control and manage their resources.

Managing the information explosion may require customers to coordinate use of multiple information infrastructure technologies including tiered storage, storage virtualization, information archiving, security, data reduction software products, and automation. It will also require IT decision makers to take advantage of best practices for deployments and ongoing operational processes. Using a capability-based approach to define requirements and sort through the options will help customers make better decisions faster. The sooner IT decision makers take a hard look at their information infrastructure requirements, the sooner they can begin to save money, reduce risk and improve operational efficiencies.

20 Asylum Street
Milford, MA 01757
Tel: 508-482-0188
Fax: 508-482-0218

www.enterprisestrategygroup.com