WHITE PAPER

IBM Information Infrastructure Initiative Tames the Information Explosion

By Brian Babineau

April, 2009

Table of Contents

Executive Summary

Around the world, IT organizations are struggling to manage, protect, and maintain rapidly growing information assets. 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.

Over the past decade, corporations have routinely created, on average, 60% more data each year. Despite a global economic recession, including staggering unemployment rates, ESG estimates data will still grow by roughly 25% in 2009 as companies look to leverage Web 2.0 applications such as blogs and Wikis to share information inside and outside the organization. When businesses and economies recover, ESG believes annual worldwide information growth rates of 50% to 100% are expected to be the norm again—driven by new rich media content such as video surveillance and high definition television programming delivered over the web. In fact, if there is one constant businesses can count on, it is information growth.

This growth is compounded by the amount of data that companies save on a regular basis. Organizations around the world are starting to realize that information is a powerful business asset that can be leveraged to enter new markets, create new products and services, and improve operational efficiencies.

A perfect example of the intersection of information growth and information retention is the medical industry, where a current Computed Tomography (CT) study today consists of 64 image slices—which creates roughly 32 MB of data. This study will likely be retained for the life of the patient and be secured due to regulatory requirements. While health care IT departments may be able to deal with the current information generated by current imaging modalities, newer solutions such as the Magnetic Resonance Force Microscopy (MRFM) will quickly increase the size of images, which in turn, will impact data size. The MRFM, developed by IBM and Stanford University, offers the ability to study complex 3D biological structures at a nano-scale with volume resolution 100 million times finer than conventional MRI. As a result, physicians will have access to much more information to make an accurate diagnosis, but this will also force healthcare IT departments to be smarter about how they store, manage, protect, and secure information.

Many organizations will see information explode with the advent of a new business opportunity, similar to the health care industry. To get control of information growth rates—while reducing business risk and controlling storage management costs—IT decision makers must:

- Take a process- and services-oriented approach to managing their storage, information protection, and information security.
- Think creatively about how technologies, processes, and people can leverage the newly created and retained information, respond to external compliance and internal security requirements, and improve resource utilization.
- Select products and technologies by identifying the core capabilities and functional attributes required of their information infrastructure.

By defining requirements in terms of best practices, services, and processes, IT executives can make more effective investment decisions as their organizations' information management requirements become increasingly complex.

As they go about making these investments, customers may encounter IBM's Information Infrastructure initiative, which is a capability-oriented model to help customers identify, evaluate, and implement several IBM solutions—including recently acquired products and services. These capabilities—inclusive of information availability, information security, information retention, and information compliance—holistically consider the technologies, policies, and processes needed to deliver cost effective, high performance information and storage services

across the enterprise. Integrated solutions aligning with these capability areas include a significant number of the 70+ new product and service releases and major upgrades IBM announced over the past year, as well as a number of best practices that harvest lessons learned from custom engagements with government and Fortune 500 customers.

IBM's capability-based approach centers around mapping requirements to solutions—allowing 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 workshops, assessments, best practice templates, and technology recommendations to help master the specific capability or set of capabilities under consideration.

The IBM Information Infrastructure initiative is a new and vitally important pillar supporting the firm's Dynamic Infrastructure: Helping Build a Smarter Planet agenda as it enables organizations to consistently gather, store, and utilize information more intelligently—which, in turn, creates more business opportunities, lowers operating costs, and helps mitigate risk.

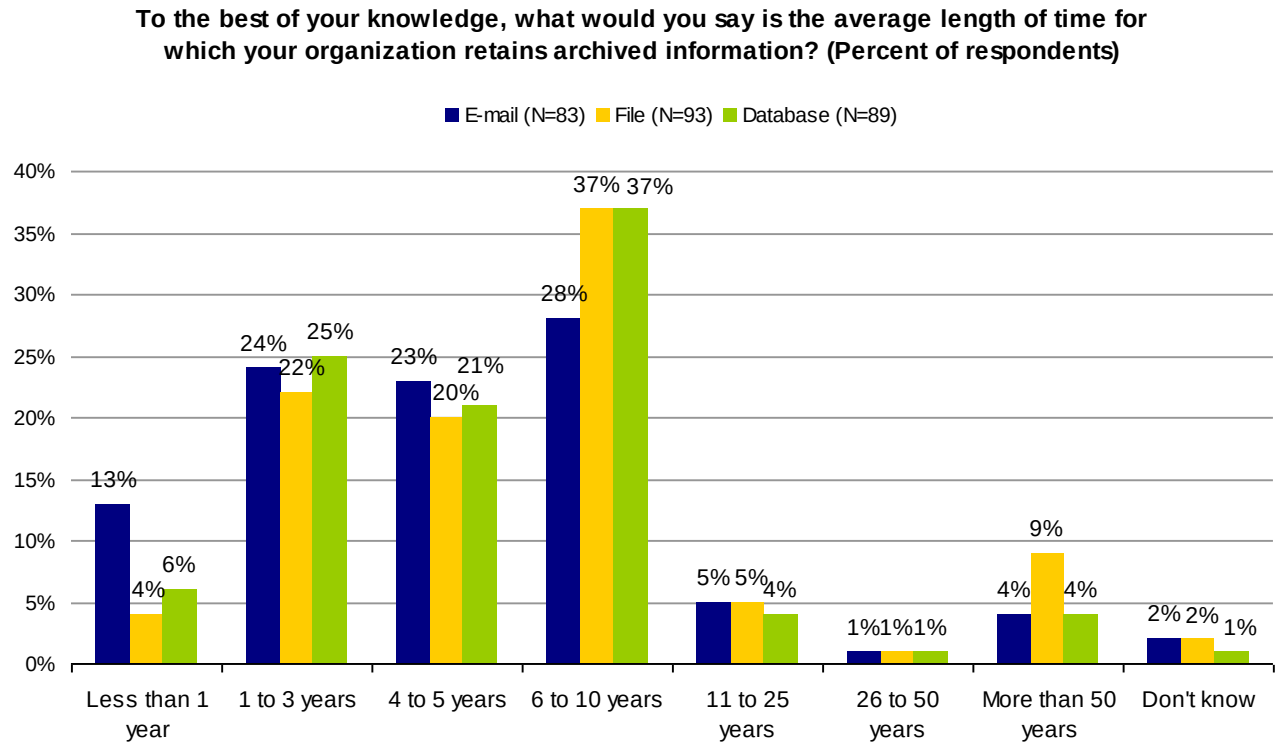
The Information Explosion

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. As a result, ESG believes that corporate data will increase 25% in 2009 after growing two to three times that rate per annum for much of this decade.[1]

Net new information is only part of the information explosion equation. Organizations are also retaining more information than ever before due to record retention regulations, electronic discovery requirements, and business intelligence initiatives that thrive on data. Organizations are saving e-mails, application files, and database records for several years and ESG expects this to continue as more and more companies realize that information is a powerful business asset that can be leveraged (see Figure 1).

[1] Source: ESG Research Report, *Enterprise Storage Survey*, November, 2008.

FIGURE 1. AVERAGE INFORMATION RETENTION PERIOD BY CONTENT TYPE



Source: ESG Research Report, Digital Archiving Survey, November 2007

The IT challenges created by this information explosion will not go away on their own, nor will they be successfully addressed with the incremental addition of several new terabytes of storage. As more data is generated and saved, businesses expand beyond traditional markets, extending use of mobile workforces and attempting to be more 'green' by implementing programs to reduce power consumption. 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 lower long-term capital expenses and promote operational efficiency.

A More Strategic Approach to Information Storage and Management

Old Methods Won't Work

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. For example, existing storage architectures may not have kept pace with data growth, forcing customers to constantly buy more and more of the same systems. In fact, 35% of organizations surveyed by ESG rank storage consolidation as an IT infrastructure initiative that will impact their storage spending over the next two years.[2] It is likely that this spending will include tiered storage and storage virtualization technologies to improve utilization of existing resources and promote data mobility over time.

[2] Source: ESG Research Report, *Enterprise Storage Survey*, November, 2008.

Likewise, organizations that could once comfortably accommodate their information protection requirements using tape backup solutions may find they now need to leverage disk in the backup process simply to complete once-a-day backups. These same companies may also realize that they need to introduce sophisticated archiving systems to comply with regulatory demands. Existing data retention policies that rely on saving everything "just in case" may be needlessly driving up storage capacity requirements.

The opening of remote offices or a major merger may constrain backup windows or create security challenges for centralized storage and data protection systems. Data reduction technologies, including data compression and deduplication, could help the organization save on hardware, power, cooling, facilities, and operating costs while enabling backup and archive operations to complete faster.

Focus on an Information Infrastructure's Capabilities

As organizations begin to re-evaluate their information infrastructures—the software, servers, storage, and networks that combine to deliver timely and secure information services throughout an organization and to its customers and partners—it is important to learn from the success of other IT strategies. For example, many companies have benefited from the introduction of service management best practices, such as ITIL. Use of a service management approach for data center operations allowed 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 enabled 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.

Applying the same principles to an information infrastructure, 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 centering specifically on individual products. Some of the most important information infrastructure capability areas worthy of consideration include:

- **Availability.** Data must be available when and where it is needed. Employees rely on it to do their jobs, business processes are designed around it, and now, more than ever, external constituents such as partners and customers need it to execute their own operations. Sixty percent of enterprise organization respondents to ESG research (1000 employees or more) indicated that if their mission critical applications were unavailable for more than four hours, they would experience significant revenue loss or other adverse impacts.[3]

- **Security.** With several governments and industries having information privacy laws and significant financial penalties, organizations have to worry about making information available to the right people. In the United States alone, 656 data breaches were reported during 2008—a 47% increase over 2007.[4]

- **Asset utilization.** In 2009, ESG expects storage budgets to increase a paltry 3% from 2008.[5] With information growths at eight times that rate, organizations must find ways to improve storage system utilization. In many cases, this is not possible due to legacy storage architectures and disaster recovery implementations where systems are replicated based on one-to-one relationships. Also, companies should look to extend the life of systems by moving data to other devices, including those that support different media types (disk, tape, and optical).

- **Data Protection.** Many organizations find that significant amounts of data are not consistently backed up due to a lack of standardized processes and policies, as well as the use of disparate 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.[6]

[3] Source: ESG Research Report, *Data Protection Market Trends*, January, 2008.
[4] Source: http://privacy.org/archives/002593.html
[5] Source: ESG Research Report, *Enterprise Storage Survey*, November, 2008.
[6] Source: ESG Research Report, *Data Protection Market Trends*, January, 2008.

- **Retention Management.** Companies save copies of information in the event of corruption or data loss. Oftentimes, copies are also created for disaster recovery reasons as well as for test and development, compliance, and many other purposes. In fact, 51% of organizations responding to an ESG survey create between two and ten copies of their primary database instances for the purpose of testing enterprise applications in user environments.[7] In data protection scenarios, companies have reasonable data expiration policies. However, copies of data saved for compliance, testing, and other purposes are rarely deleted. 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.

- **Compliance:** Information privacy rules are not the only things that impact information infrastructures. Organizations must also abide by record retention regulations that force them to save information for specified periods of time. Over half of organizations surveyed by ESG are saving e-mails to comply with such mandates.[8] Organizations may also be saving data to follow their own internal corporate governance mandates.

  What's more important than saving more information? Retrieving it to prove 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 quickly. It is also imperative that they are able to document who accessed and changed that information over its lifetime.

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 and technology requirements in terms of attributes. By selecting products and technologies that contain specific attributes, organizations will be assured that they can address the aforementioned capability areas, which in turn will help them meet their evolving business requirements over the long term. The attributes that organizations should consider looking for in products and technologies that will underpin their information infrastructure are:

- **Automated operations,** minimizing the need for IT staff to manually manage data over long periods of time.
- **Policy-driven compliance and security,** enabling organizations to swiftly adapt to any new mandates that impact information retention, availability, accessibility, and security.
- **Virtualized and optimized infrastructure,** preventing information from being bound by the underlying physical infrastructure.
- **Long-term platform independence,** ensuring data can be leveraged beyond the useful life of the hardware and software assets that currently store and manage it.

This capability-driven approach to requirement 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.

The Supporting Technologies

Once IT organizations begin to define and evaluate their storage and information management needs in terms of core capabilities and solution attributes, they will be better able to identify the key functional technology requirements needed to support operational flexibility and business growth over the long run. For example, recessionary times may drive requirements for a smarter approach to infrastructure management, rather than simply adding capacity to the existing infrastructure. For many organizations, technologies that merit consideration include:

[7] Source: Source: ESG Research Report, *2007 Database Archiving Survey*, December 2007.
[8] Source: Source: ESG Research Report, *2007 E-mail Archiving Survey*, December 2007.

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

- **Data compression and data deduplication** technologies that reduce the overall volume of data to be stored in the first place.

- **Tiered storage environments** that include networked storage and a mix of disk and tape platforms to make efficient use of resources. A multi-tiered, virtualized infrastructure also enables organizations to build and offer their own 'cloud computing' infrastructures.

- **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.

- **Policy-based, compliance management** tools to reliably apply information retention 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.

These technologies may come in the form of individual or integrated products. Regardless, each can deliver substantial benefits through cost savings and operational improvements. These opportunities are magnified as organizations implement more of them in conjunction with each other to meet business requirements.

IBM's Information Infrastructure Initiative

A Comprehensive Portfolio

IBM has long offered its customers a significant number of storage-related products and services. However, 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,** which offered storage systems with 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,** a leading storage resource management software and services provider with deep expertise in large scale storage management optimization, workflow standardization, and best practices enforcement.
- **Diligent,** a company that successfully built data deduplication software capable of being integrated with server and storage infrastructures to help organizations reduce the amount and cost of physical storage required in data centers.
- **FilesX,** which provided continuous data protection and recovery software for enterprises and remote/branch offices.
- **Softek Storage Solutions**, which offered local and remote data protection and data migration software and services.
- **Arsenal Digital,** an online storage backup and information protection services provider.
- **Internet Security Systems,** a leading vendor of security products and services protecting information from known and unknown threats in the data center, host systems, network, and endpoints.
- **FileNet,** a long-time marketing leader providing a family of content management software tools that help organizations streamline and automate business processes, including records management.
- **Princeton Softech (Optim),** which developed enterprise application lifecycle management software optimized for major ERP and CRM suites, providing archiving and test data depersonalization capabilities.

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.

Alignment with Business Requirements

The challenge for organizations, however, is to find the right IBM technologies and services to address their specific sets of requirements and then implement and operate the chosen solution effectively. The IBM Information Infrastructure initiative aims to help organizations of all sizes to understand and take advantage of 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. To address specific challenges that are the direct result of ever-changing business requirements, IBM organized the Information Infrastructure integrated solutions into four groups based on the capability areas they enable.

- **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 across diverse platforms. This capability area includes end-to-end storage virtualization solutions, including virtualization for disk, tape, SAN Directors, networks, and file systems. 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, and data security. IBM's research in cryptography, disk, tape, and data center security anchors this capability area. Important solutions include IBM encryption, which is built directly into storage devices in order to minimize overhead, and 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 deduplication and compression technologies. IBM's Scale-Out File Services product is a featured offering in this capability area. This solution can search very large directory structures quickly, leveraging inexpensive servers and disk in a grid configuration which provides affordable scale-out capability.

- **Information Compliance:** Reduces compliance and audit risk via the use of data discovery tools, protected storage, and information management professional services to provide best practices and solutions covering information retention, access, and destruction. An example of an integrated solution in this area is the IBM Compliance Warehouse for Legal Control, which includes agents to collect business records and discoverable information from e-mail and other repositories, policy-based retention to simplify management of legal hold orders, and analytic and reporting tools 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 organization 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.

Measuring Success

Numerous IBM customer success stories demonstrate 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 their 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, which enables the company to easily capture all of the data

being generated and make it quickly available for business analytics. Additionally, the company saved server images on networked storage. Information availability quickly improved as the firm automated web server deployment, thereby shrinking the implementation process from hours to minutes with more consistent results. By taking a capability-based approach, Zappos.com was able to improve information and web server availability by deploying automated management and high performance storage solutions.

Conclusion

Businesses must adapt to their own information explosions while dealing with tighter budgets and finding opportunities in new markets, products, and services to become more operationally efficient. While decision makers might be tempted to stick with the status quo by adding more capacity, today's economic climate demands a smarter approach.

Companies need to think about an information infrastructure and what capabilities are required as part of that. Then, they should look at specific technologies to build that information infrastructure—ones that will improve utilization, ensure information availability, and facilitate compliance via the appropriate retention and security policies. IT must also look for best practice recommendations to become operationally efficient, further maximizing the investment in an information infrastructure.

IBM, with its Information Infrastructure initiative, makes it much easier for customers to take a smarter approach and make technology investments based on capabilities rather than point-product features. By focusing on capabilities, customers will make more educated purchases that will help them take advantage of the information explosion rather than simply dealing with it.

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

www.enterprisestrategygroup.com