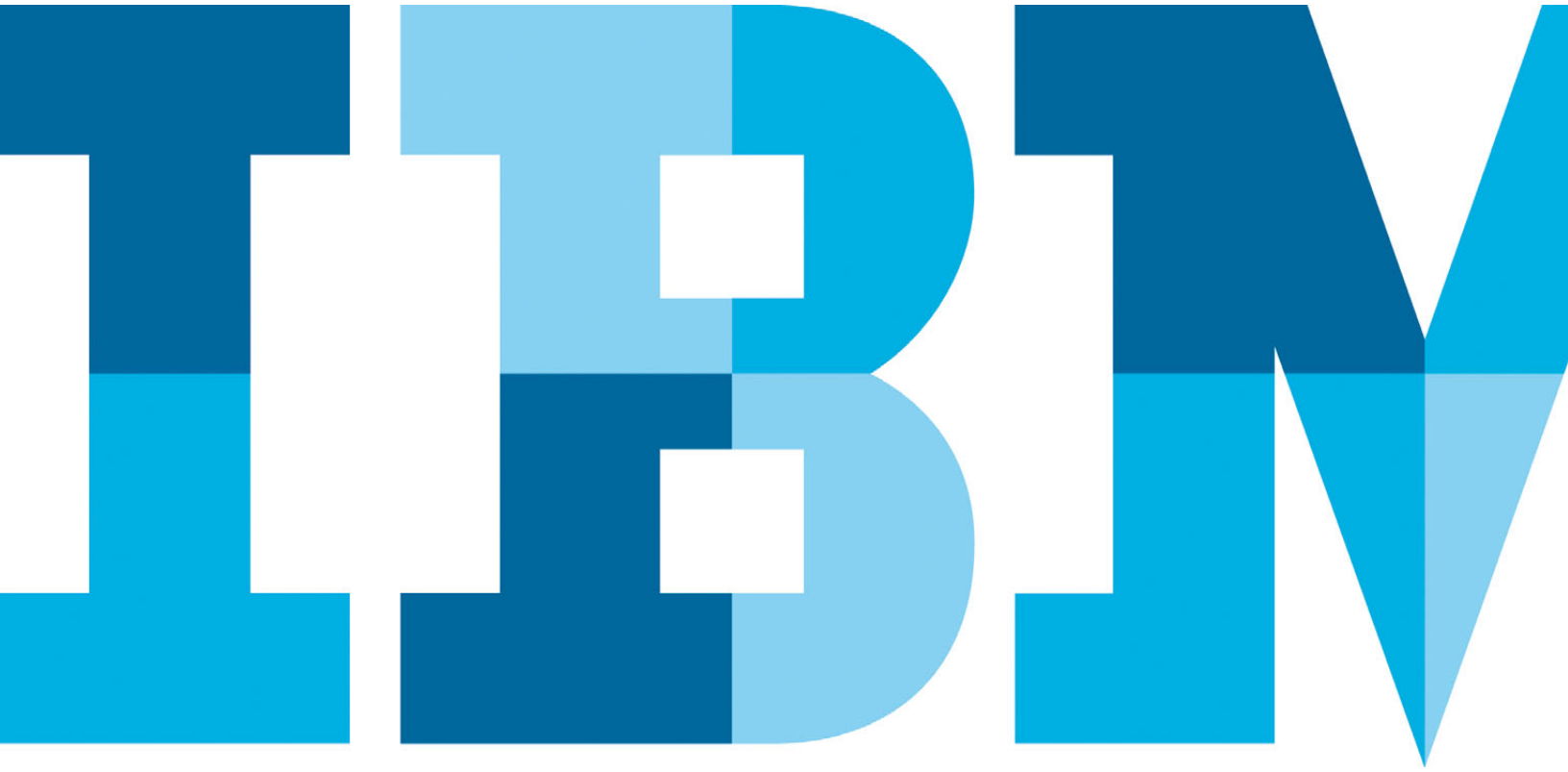
IBM Software
Thought Leadership White Paper

December 2011

Effective storage management and data protection for cloud computing

Protecting data in private, public and hybrid environments



IBM

Contents

Introduction

The U.S. National Institute for Standards and Technology (NIST) has defined cloud computing as "a model for enabling convenient, on-demand network access to a shared pool of configurable computing resources (for example, networks, servers, storage, applications and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction."

Cloud-based systems have brought a new, scalable application delivery service model to the marketplace. Organizations are using cloud computing to save both capital and operational costs, shift capital expenses to operational expenses and gain increased flexibility. A "pay-as-you-go" usage model, in which cloud-based resources can be provisioned only as the user needs them, enables companies to better predict and manage expenses, reduce costs and simplify operations. The cost of adding new services and users in the cloud model, in fact, often can be almost completely eliminated.

Despite the advantages they gain from cloud-based application delivery, however, not all organizations are giving necessary attention to another key aspect of cloud deployment—storing and protecting the data that exists "in the cloud." Yet, ready access to cloud-based digital information is a foundational requirement of reliable cloud-based service delivery. Today, organizations that deal with large amounts of data

Cloud computing: Key points at a glance

Options for deploying cloud computing:

- **Build your own private cloud**
- **Subscribe to a private cloud**
- **Subscribe to a public cloud**
- **Subscribe to both (hybrid cloud)**

Data and storage management in the cloud are critical:

- **To provide a reliable, on-demand service experience**
- **To reduce costs and enable scalability**
- **To mitigate risks**

The keys to effective cloud storage management include:

- **Data protection and recovery**
- **Data security and lifecycle management**
- **Storage utilization and optimization**
- **Storage resource management**

(often termed "Big Data")—in various industries including media, banking and healthcare—are increasingly adopting cloud technology to deliver faster services, protect data in real-time, provide seamless communication between employees, partners and suppliers, enable business continuity, and address the pressure to become more energy efficient—to be a greener organization.

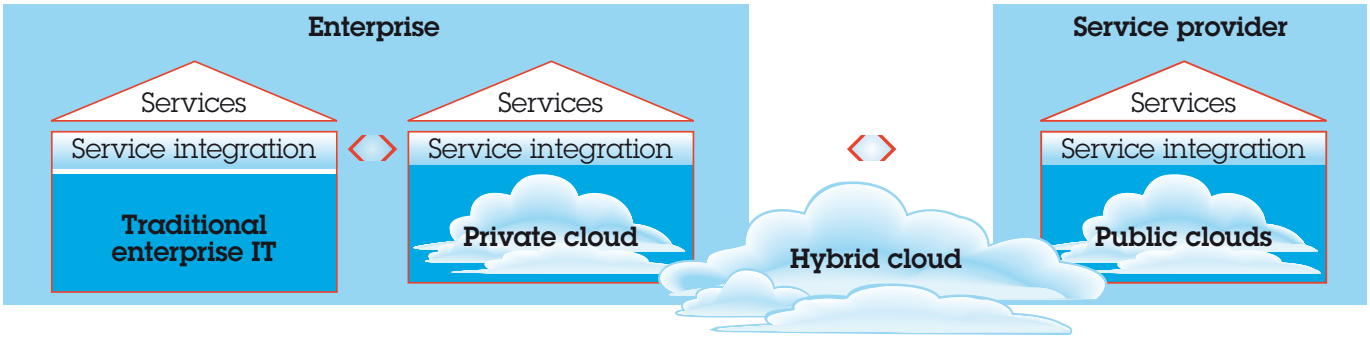
In cloud computing environments, applications reside in massively scalable data centers where compute resources can be dynamically provisioned and shared to achieve significant economies of scale. Storage and backup capacity needs to scale with compute resources—and both must be effectively managed to gain the maximum cloud benefits.

Cloud configurations

Several cloud computing models are currently available. These generally break down into a combination of the services being provided and the manner in which they are hosted. For example, a cloud may provide infrastructure (servers, storage or networking) as a service, or it may provide software platforms (desktop, server or web-based applications) as a service.

The methods of hosting the cloud can be "private," in which the company owns and manages the cloud infrastructure; "public," in which a third-party owns and manages the infrastructure; or "hybrid," in which the public and private models are combined. In hybrid environments, the public cloud often acts as an overflow facility for the private cloud, or is used to satisfy other application needs, such as offsite information protection.

Because all of the data in a cloud resides in the same shared systems, effectively managing the storage resources and data in the cloud is of paramount importance to maintaining satisfactory service levels. Cloud services rely heavily on keeping the data and applications they are managing available at all times. They also must be able to restore operations quickly following any type of data disaster, such as database corruption, virus attack, hardware failure or local/regional disaster. Downtime, or failure to provide access to needed storage and data resources, constitutes a failure of the cloud service and should obviously be avoided.



Cloud computing typically exists as three types—private, public or hybrid. It may supplement or replace an organization's traditional enterprise IT operations.

One of the key challenges data center managers face is moving their existing heterogeneous storage infrastructures into cloud storage platforms. At the same time, effective utilization and optimization of the storage resources in the cloud infrastructure, along with the proper placement of data on different tiers of storage within the cloud, helps minimize the overall costs of hardware, software and administration.

Data protection in the cloud

For organizations embarking on cloud computing, storage management is extremely important. To avoid loss, the cloud system must provide data protection and resiliency. If loss does occur, the environment must be able to recover the data quickly in order to restore access to the cloud services.

This need is equally true whether the cloud environment is private, public or hybrid. It is possible to obtain data storage and protection services from companies that specialize in storage, and many companies that handle their own computing and applications choose this option for outsourcing storage. When outsourcing applications, however, a company should never assume that the service provider includes storage management, data protection or disaster recovery among its services—not all of them do. It is therefore important to ensure from the beginning that the provider delivers the necessary data storage and protection services, and to be familiar with the technologies and products used for storage management in the cloud.

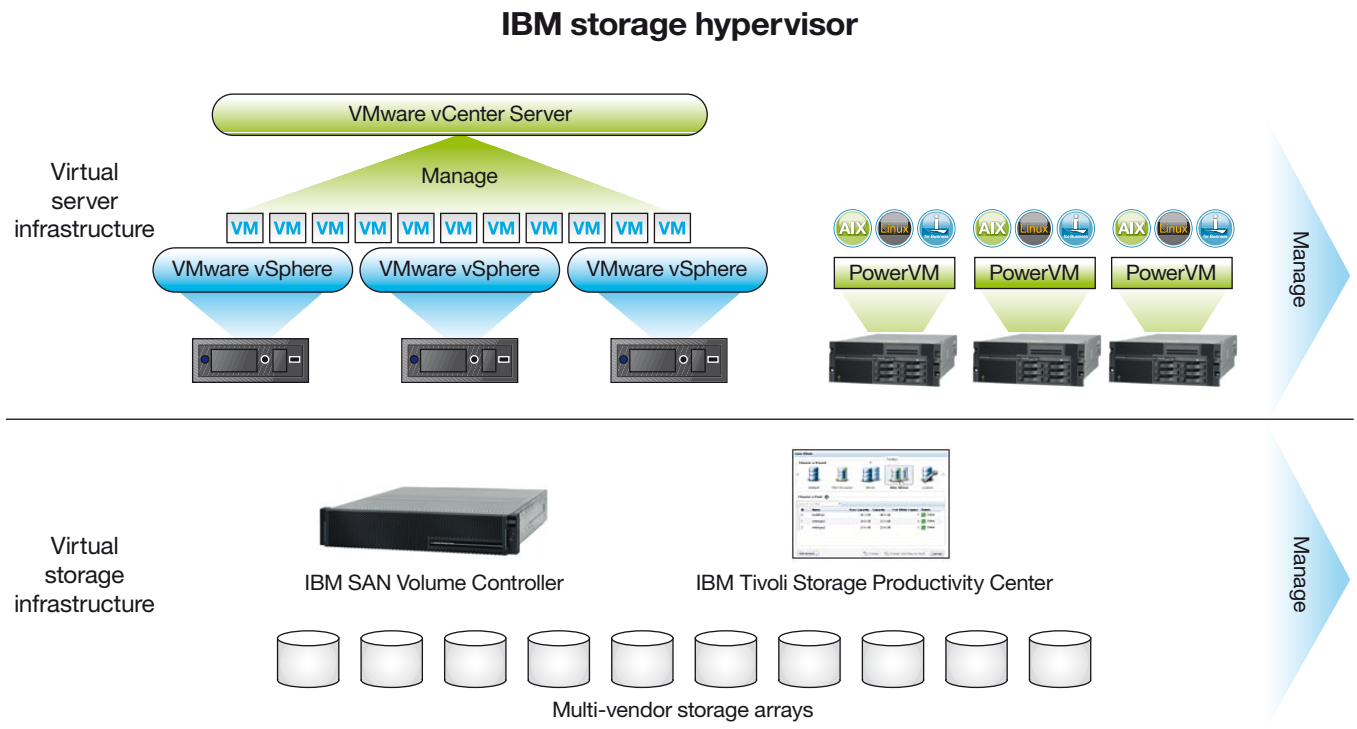
IBM storage management solutions for cloud computing

The following sections describe IBM solutions that support storage services in cloud computing environments.

Storage hypervisor

IBM pioneered the concept of the storage hypervisor—an appliance solution that can help storage managers virtualize heterogeneous storage infrastructures with ease and control, and manage these resources through comprehensive management software. The IBM storage hypervisor solution includes IBM® System Storage® SAN Volume Controller, IBM Tivoli® Storage Productivity Center and IBM Tivoli Usage and Accounting Manager.

Apart from improved utilization, the storage hypervisor helps create a smart storage infrastructure that can support IT's cloud ambitions. Necessarily, the storage hypervisor enables thin provisioning for cost advantage, data compression for storage efficiency, insights for storage tiering decisions and comprehensive support for heterogeneity of storage systems. With more than 7,000 deployments, IBM offers the world's most widely deployed storage hypervisor.



A virtualized storage environment from IBM can operate seamlessly with virtualized server environments built on VMware or other solutions.

Data protection and recovery

Any cloud environment should include solutions for protecting the data that "lives" in the cloud, and for recovering lost data in the event of an error or failure. Cloud-based storage is also often used as a repository for backup and archive data.

IBM Tivoli Storage Manager provides unified data protection and recovery management across the entire infrastructure, including the cloud, to provide reliable and predictable access to data. Tivoli Storage Manager includes automated data lifecycle management, built-in data reduction technologies, advanced application protection and more.

Data managed by cloud service providers must be protected appropriately, according to its value. For business data, this typically means managing multiple copies and multiple versions of every file. Tivoli Storage Manager provides flexible levels of data protection, and can meet a variety of recovery point objectives and recovery time objectives, offering multiple levels of service options.

More importantly, Tivoli Storage Manager can be configured to segregate data by client. Cloud providers can assure their clients that onsite and offsite backup media will not get into the hands of other users, even in a disaster recovery scenario.

IBM Tivoli Storage Manager FastBack® is an advanced, continuous data protection and near-instant recovery software solution for business-critical Microsoft Windows and Linux servers. Tivoli Storage Manager FastBack helps organizations reduce the amount of data at risk between backups to almost zero, and it reduces the recovery time for almost any data loss to mere seconds. Within cloud environments, Tivoli Storage Manager FastBack can be deployed to provide enhanced service level protection for the most critical applications, whether on physical or virtual servers.

IBM Tivoli Storage FlashCopy® Manager enables fast and frequent backup of critical applications to limit data loss without disruption to operations. It leverages advanced snapshot technologies in IBM System Storage solutions. Tivoli Storage FlashCopy Manager can work as a standalone solution to provide near-instant, application-aware restore capabilities in physical and virtual server environments, as well as integrating with Tivoli Storage Manager to provide longer-term data management.

Data lifecycle management

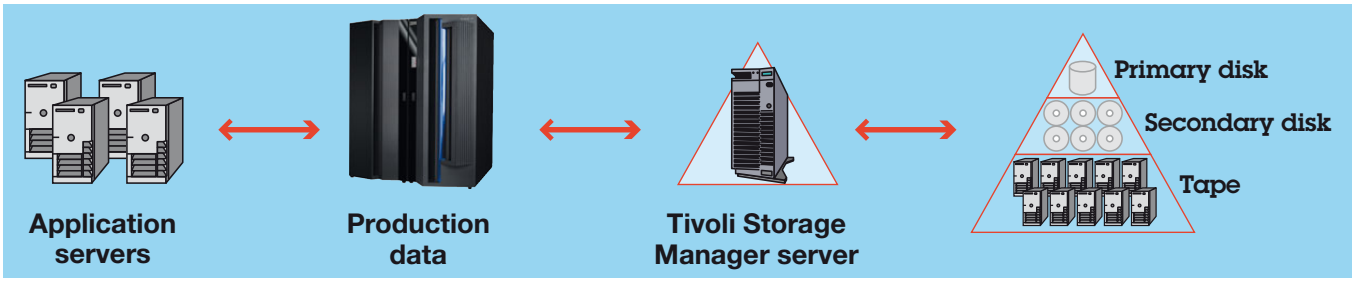
An important consideration in planning and managing storage in cloud environments is to make the most efficient use of storage resources by placing data on the most appropriate and cost-effective tier of storage that meets service delivery requirements—then to eliminate data that is no longer needed. The process of automatically moving and deleting data based on business rules is referred to as *data lifecycle management*.

IBM Tivoli Storage Productivity Center for Data is a sophisticated data reporting tool that helps set policies and identify data that can be safely removed from the cloud. With Tivoli Storage Productivity Center for Data, you can get reports on where your data is, sorted by access or saved dates, who owns it, the application that created it and numerous other filters. From the intelligence you gain from Tivoli Storage Productivity Center for Data reports, you can set meaningful policies in your data management software to automatically take the appropriate action on data that shouldn't be clogging your cloud systems and running up usage charges. It can also help you identify and eliminate duplicate data, orphan data, temporary data and non-business data.

The next step is to automate these data management policies through three distinct processes: migration, archival and expiration.

The IBM Tivoli Storage Manager family includes the capabilities that automate data migration between multiple storage tiers based on policies. The net result of these processes is to remove unneeded data from your primary cloud storage systems, reducing or postponing the need to acquire more expensive hardware and helping to reduce administrative costs—all without impacting key operational processes.

Tivoli Storage Manager data migration solutions not only help you clean up your primary storage systems to help them run more efficiently, they can also be used to easily move data to new storage technologies as they are deployed within the cloud.



Implementing a hierarchical storage management model that migrates data from application and production systems to secondary storage tiers is a key part of an efficient and effective cloud environment.

These solutions work transparently in the background, automatically selecting and moving files from primary to secondary tiers of storage—or back again—based on the policy criteria that you set, such as file size or length of time since a file has been opened. All operations are transparent to the users and applications.

Data migration within the cloud can help you get control of, and efficiently manage, data growth and its associated storage costs by providing automated space management. It provides the following key features:

- Storage pool virtualization that helps maximize utilization of the managed storage resources
- Restore management capabilities that are optimized based on data location
- Migration that is transparent to the users and to applications
- Migrations that are scheduled to minimize network traffic during peak hours
- Automatic migrations that occur outside the backup window
- Threshold limit settings that can eliminate "out of disk space" messages

In a hybrid cloud model, data lifecycle policies can automate the transparent movement of data from the private cloud to the public cloud, and vice versa. That movement can be greatly accelerated with the use of a cloud gateway solution such as Riverbed Whitewater or the Netex HyperIP.

Storage utilization and optimization

IBM software solutions for managing storage in cloud environments are designed to deliver optimized data availability and performance, and to improve utilization of storage assets. They can help organizations reduce the effort required to manage the storage infrastructure stack in the cloud by streamlining resources to support data protection and management, and by being able to virtualize the entire storage infrastructure and offer it as a single resource to the cloud.

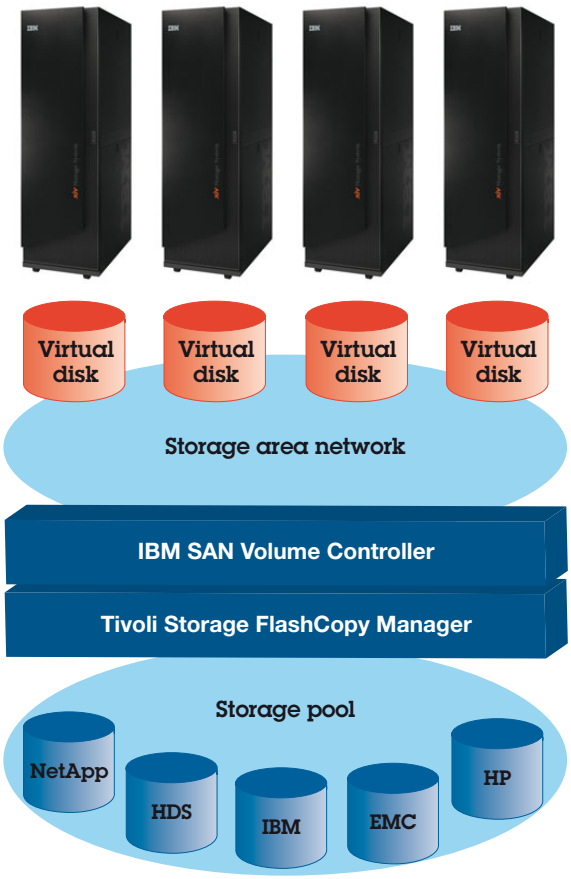
IBM System Storage SAN Volume Controller virtualizes IBM and non-IBM storage to enable resource pooling, thin provisioning and simplified management. SAN Volume Controller eliminates storage silos by enabling the sharing of storage resources among applications. It auto-assigns and releases capacity on demand, addressing a critical requirement of all clouds—the redeployment of resources as utilization needs change.

IBM Scale Out Network Attached Storage (SONAS) is a multipetabyte offering for information storage and distribution. It is built on the IBM General Parallel File System (GPFS™) to store and manage billions of files in a single namespace—which is critical to larger compute and storage clouds.

SONAS meets today's storage challenges with quick and cost-effective IT-enabled business enhancements, combining massive scalability, high-availability and automated data management to efficiently deliver information around the world.

Storage resource management

The absence of sophisticated storage management systems in a cloud often results in a lack of visibility into the storage environment. Cloud administrators have difficulty in understanding how much capacity is available, where it is located, which applications are accessing it, how it is performing, how secure the storage is and whether it is possible to meet stringent recovery time objectives. Gaps that may exist between the unpredictable demand for data availability and the ability of business to meet that demand in an efficient way can result in unmet service levels, additional downtime, new hardware and operational costs and lower customer satisfaction.



System Storage SAN Volume Controller combines heterogeneous storage devices into a single pool of capacity. It manages the placement of data within the virtual storage pool based on business policies.

IBM Tivoli Storage Productivity Center is an advanced storage resource management solution that manages virtualized storage environments with tools that can help companies reduce complexity by centralizing, simplifying and automating storage tasks associated with cloud infrastructures. This integrated solution can help improve storage availability within cloud environments by combining the assets, capacity, performance, analytics, configuration, operations and data replication in a single management platform. Designed to manage both IBM and non-IBM storage systems, Tivoli Storage Productivity Center is easy to deploy and manage, and its components are tightly integrated with Tivoli service management software to deliver automated management of data and resources.

IBM SmartCloud Resilience Services offerings

For those companies looking for a managed services approach to cloud-based data protection and recovery, IBM SmartCloud Resilience Services offer a suite of solutions designed to help your business leverage cloud services technologies to quickly and cost effectively recover in the event of a disruption, data loss or disaster. Our solutions, which include server recovery and data protection and backup services, can help you efficiently manage risk, reduce costs and meet regulatory compliance mandates.

IBM SmartCloud Managed Backup is a cloud-based service that enables security-rich, managed protection of critical data. Our service provides on- or off-site data backup to help you reduce operational risk and total cost of ownership. It can help you keep your critical business data available in the event of an outage. By offering private or public cloud-based backup services, it enables you to choose and implement a plan for backup, retention and retrieval so that you can gain security-rich, around-the-clock access to information.

IBM SmartCloud Archive provides a cloud-based data archiving solution that can index, search, retrieve and store archived content more cost effectively—helping you leverage information for enhanced decision-making and better regulatory compliance management.

Integrated service management provides an effective way to manage storage in a cloud computing environment—so information assets are safe and available to help achieve business goals.

Summary

IBM software and service solutions for smarter storage management and information protection in cloud environments help ensure that business and IT operations are fully aligned and supported by integrated service management. They help deliver a workload-optimized approach, and they offer a choice of implementation options for superior service delivery with agility and speed.

IBM now offers second-generation storage management technology for cloud environments, delivering faster return on investment. IBM's first-generation technology was developed several years ago, deployed internally within IBM, and used successfully by more than 300,000 IBM employees worldwide. The best practices and lessons learned then formed the base for IBM's second-generation software technology, built to support the needs of a smarter planet—more scalable, more reliable, more efficient and more flexible.

Unlike other cloud computing solutions that use hardware and software combinations that are not tested for scalable results, IBM solutions for storage management in cloud environments are tested and proven reliable—and they are backed by a strong commitment to further research and development of IBM cloud technology. Across the entire portfolio, IBM storage software products have several things in common—they are fueled by business expertise, built to scale and can integrate and adapt to a customer's changing needs. And they have many years of proven reliability in real-world environments.

IBM reduces the complexity of managing cloud environments by offering a complete portfolio of automated solutions for managing data and storage infrastructure, enabling better efficiency for business resiliency and helping to reduce costs and improve security, all while increasing visibility, control and automation of the cloud storage infrastructure.

For more information

IBM is uniquely positioned to provide a wide range of storage management solutions, even for mixed-vendor environments. IBM can help you implement a dynamic storage infrastructure that scales to meet your changing business requirements.

To learn more about how IBM Tivoli storage management solutions can help you address data management concerns in your environment, contact your IBM sales representative or IBM Business Partner, or visit: **ibm.com**/tivoli/solutions/storage

Additional information can be found at:
ibm.com/tivoli/solutions/storage-cloud

About Tivoli software from IBM

Tivoli software from IBM helps organizations efficiently and effectively manage IT resources, tasks and processes to meet every-shifting business requirements and deliver flexible and responsive IT service management, while helping to reduce cost. The Tivoli portfolio spans software for security, compliance, storage, performance, availability, configuration, operations and IT lifecycle management, and is backed by world-class IBM services, support and research. For more information on Tivoli software from IBM, visit **ibm.com**/tivoli

Additionally, financing solutions from IBM Global Financing can enable effective cash management, protection from technology obsolescence, improved total cost of ownership and return on investment. Also, our Global Asset Recovery Services help address environmental concerns with new, more energy-efficient solutions. For more information on IBM Global Financing, visit: **ibm.com**/financing

IBM Corporation Software Group
Route 100 Somers, NY 10589
U.S.A.

Produced in the United States of America
December 2011

Please Recycle

TIW14056-USEN-01