TechTarget

Special Report

Getting Started With Cloud Storage

Thinking about making a move to the cloud? Explore this guide to discover the key considerations you need to know before taking the plunge.

Cloud Storage Options: Public vs. Private vs. Hybrid Cloud Storage

The primary use of cloud storage today is for unstructured data, which is the fastest growing and most voluminous content, causing the most administrative pains. Cloud storage is less suitable for structured data, which continues to live on traditional enterprise data storage.

The benefits of cloud storage technology

The benefits of using cloud storage technology for unstructured data are compelling, starting with lower overall storage costs. Being service based, there's no storage hardware to buy, manage and maintain, and depending on the service, it can greatly reduce, if not eliminate, data center and storage administrator costs. Cloud storage eliminates expensive technology refreshes that usually kick in three years to five years after the initial purchase, needed to either get state-of-the-art technology or simply to get around purchasing expensive support contracts for older arrays.

The technology can provide close to 100% storage utilization by eliminating the massive amounts of unused storage that are needed with traditional data storage for anticipated growth and peak loads. Besides the overall cost savings, scalability of cloud storage and its ability to transparently support base and peak loads are its most appealing characteristics.

Public cloud storage

Public cloud storage services are a cloud storage option offered by a fast growing list of service providers: AT&T, Amazon, Iron Mountain Inc., Microsoft Corp., Nirvanix Inc., Rackspace Hosting Inc. and many others. Their storage infrastructure usually consists of low-cost storage nodes with directly attached commodity drives with an object-based storage stack that manages the distribution of content across nodes. Data in the cloud is typically accessed via Internet protocols, mostly Representational State Transfer (REST) and to a lesser degree Simple Object Access Protocol (SOAP). Resilience and redundancy is achieved by storing each object on at least two nodes. Usage is charged on a dollar-per-gigabyte-per-month basis and, depending on the service provider, there may be additional fees for the amount of data transferred and access charges.

Public cloud storage is designed for massive multi-tenancy that enables isolation of data, access and security for each client. The type of content stored on public clouds ranges from static non-core application data and archived content that needs to be

available, to backup and disaster recovery data. Public cloud storage isn't suited for active content that changes all the time. The primary concern of using public cloud storage in the enterprise is security and, to some extent, performance.

Internal or private cloud storage

Internal or private cloud storage runs on dedicated infrastructure in the data center and, as a result, address the two main concerns of security and performance, but otherwise offers the same benefits of public cloud storage. Internal storage clouds are usually for a single tenant, even though larger enterprises may use multi-tenancy features to segregate access by departments or office locations. Unlike their public cloud storage counterparts, scalability requirements are more modest, so internal cloud storage offerings are more likely to have traditional storage hardware under the hood. A case in point is Hewlett-Packard (HP) Co.'s CloudStart, which combines HP BladeSystem Matrix, an HP StorageWorks Enterprise Virtual Array (EVA) Family array and Cloud Service Automation (CSA) software into an internal cloud storage infrastructure. HP CloudStart by itself isn't a private cloud storage offering because it lacks the key element of being service based; instead, it's the enabling infrastructure that could be used by HP, one of its partners or even enterprises to offer it as a fully managed, pay-as-you-go cloud storage offering.

An example of a private cloud storage offering is the Hitachi Data Systems Cloud Service for Private File Tiering. Based on the Hitachi Content Platform (HCP), it resides in the customer's data center but is owned and managed by Hitachi. Besides an initial setup fee, the customer pays for it by usage. Similarly, Nirvanix hNode provides a fully managed, pay-as-you-go, internal cloud offering within the data center, based on the same technology that powers the Nirvanix Storage Delivery Network (SDN).

Hybrid cloud storage

Users who have a hybrid cloud storage environment manage resources both externally and in-house. Because hybrid cloud scenarios often provide an on-site appliance, they can provide local cache and memory, data deduplication and encryption for an IT shop's data.

However, a hybrid cloud solution must meet certain key requirements to make hybrid cloud storage work. They must behave like homogeneous storage, be virtually transparent and have mechanisms in place that keep active and frequently used data on-site while simultaneously moving inactive data to the cloud. These types of clouds also depend on policy engines to define when specific data gets moved into -- or pulled out -- of the cloud. For more on hybrid storage clouds, check out our tip on hybrid cloud implementation.

Public cloud vs. private cloud vs. hybrid cloud storage

The following chart provides a quick overview of available cloud storage options.

Characteristic	Public cloud storage	Private cloud storage	Hybrid cloud storage
Scalability	Very high	Limited	Very high
Security	Good, but depends on the security measures of the service provider	Most secure, as all storage is on-premises	Very secure; integration options add an additional layer of security
Performance	Low to medium	Very good	Good, as active content is cached on-premises
Reliability	Medium; depends on Internet connectivity and service provider availability	High, as all equipment is on-premises	Medium to high, as cached content is kept on-premises, but also depends on connectivity and service provider availability
Cost	Very good; pay-as-you-go model and no need for on-premises storage infrastructure	Good, but requires on-premises resources, such as data center space, electricity and cooling	Improved, since it allows moving some of storage resources to a pay-as-you-go model

Each cloud storage option discussed here has its pros and cons. Public clouds have high scalability, but often lag in performance. Private clouds generally have high reliability, but limited scalability. And hybrid clouds might offer the in-house control that some companies are looking for, but also tend to cost more. Depending on your specific needs, the size of your environment, and your budget, one of these cloud storage options is bound to be a good fit for your organization.

Eight Questions to Ask Before Using Cloud Storage

Before using cloud storage, review these decision points from Arun Taneja to avoid issues with existing infrastructures or new cloud providers.

Why should you consider using cloud storage for at least some of your data and applications? Well, unless you've been living under a rock for the past six years, you probably have a good idea of the potential advantages that cloud-enabled storage can offer.

Public cloud storage can provide significant economies, as you pay only for the capacity you need and use, and thereby save on the Capex costs of new storage systems as well as the space, power and cooling costs to house and run them. Most providers offer virtually unlimited scalability, allowing your data storage to cost effectively expand to meet the needs of your growing business. Depending on the cloud storage service (not true with commodity cloud storage like Amazon Simple Storage Service, but certainly true with some higher-level services), you may also be able to offload data protection tasks such as backups and replication along with other administrative tasks to the provider, leading to reductions in associated hardware, software and management costs.

But we frequently hear IT managers asking a more difficult question: How can I best decide between cloud versus traditional data center storage when it comes to new initiatives, major growth and expansion, or periodic technology refreshes? Should I stick with the networked (or sometimes, direct-attached) storage approaches my organization has been deploying and managing internally, or should I plan to store at least some of my data in a public cloud?

While there's no single correct answer to this question, it's useful to consider some of the tradeoffs that may lead you to take one approach or the other. Here are eight major decision points to take into account before your organization uses cloud storage:

1. What is the primary use case? Is it backup, disaster recovery (DR), collaboration or primary storage? This is the first and most pivotal question, since the answer will narrow your focus and streamline your decision-making process. For example, the requirements a provider will need to meet for cloud-enabled backups will likely be far less stringent than those for primary storage.

2. What type of information will I be storing? Is it user and application data? Will my applications also run in the cloud? If it's the latter, you may be looking at a storage

service packaged with cloud compute, such as Amazon Elastic Block Store (EBS), which is used to support Amazon Elastic Compute Cloud.

3. Am I looking at object storage for cloud-architected applications, or block or file storage for traditional applications? The former will likely be better suited to storage in a public cloud with access via REST-like or SOAP APIs, while the latter will require access via block or file storage protocols.

4. Will public cloud storage meet my needs for data availability? This means knowing the provider's service-level agreement (SLA) terms, and what recourse you have if the terms are not met.

5. What about my data security and privacy requirements? You'll want to review the provider's encryption capabilities, both for data in transit and data at rest. What types of access controls and authorization methods are in place?

6. Will using public cloud storage allow me to comply with relevant industry regulations such as HIPAA or Sarbanes-Oxley? Does your organization have internal IT policies that you also must satisfy?

7. Can a provider meet my recovery times? If I'm using the cloud for backup and recovery, will the provider's SLA allow me to meet my recovery time and recovery point objectives?

8. What about primary storage? If at least some of the data will be primary storage, does the provider offer a gateway appliance or service to provide the feel and accessibility of local storage? Can you meet peak requirements for IOPS and response time?

If you're not satisfied with the answers to these questions for particular data sets and use cases, then you should go with in-house storage. More than likely, you'll find that some use cases and data sets are suitable for cloud data storage, while others are not. For example, many users are already using cloud storage for backup, especially for data sets they consider to be "non-critical." Remember that cloud storage gateways can serve as a "bridge" between cloud and in-house storage, providing advantages such as secure access, a local cache to boost performance, and support for traditional protocols such as Network File System (NFS) for files and iSCSI for block storage. Some gateways are targeted at specific use cases, such as backup or DR, while others are designed to accommodate a range of storage needs.

While this article has focused primarily on *public* clouds, you may also want to consider *private* cloud storage. There are not a lot of options on the market, but you can find

them if you look. For example, Nirvanix offers CloudComplete, a managed private cloud storage offering. Most private cloud products today are provider-managed, with the storage nodes and capacity usually residing in your own facilities. You gain the scalability, ease of management and most of the cost benefits of public cloud storage, including equipment Capex savings. However, you will likely end up paying operating costs.

But with storage on the premises, you will enjoy greater control and data security compared to a public cloud deployment, along with an easier path to compliance with industry regulations. Nirvanix also offers the ability to scale out your private cloud storage by adding nodes in other locations, with all the nodes federated under a single namespace. One additional advantage of private cloud storage is that it's a stepping stone to a hybrid model, in which data can be stored and easily moved as needed between private and public storage clouds. Private cloud storage can be implemented with software from the likes of Caringo, EMC Atmos, Mezeo Software or Scality (all allow the use of commodity hardware). While using private cloud storage is still in an early stage of evolution, it is something to investigate if some of the answers to the above eight questions don't add up to an obvious choice.

Cloud Storage Gateway Eases Transition to Cloud Storage Services

The most popular applications for cloud storage are long-term retention of archives and as a lower-cost alternative for storing lower value data. Public and private cloud storage systems are just now becoming a standard part of the IT arsenal. However, conventional applications may not be able to access the unique interfaces common to many cloud storage services and products. This has given rise to a new class of cloud storage gateway applications that ease the transition by presenting cloud storage in a way that applications can use.

What's wrong with cloud storage systems?

An important differentiator between modern cloud storage and the managed storage of years past is the API-based interface used by many cloud systems. Rather than using a conventional SAN or NAS protocol, most public cloud providers rely on Internet protocols, using a REST API over HTTP. This is wonderfully useful for today's programmers, but totally incompatible with legacy applications.

Even if a conventional storage protocol is used, most cloud storage systems are engineered more for scalability than absolute performance. Off-site services must contend with network latency in addition to any architectural choices made in favor of flexibility. This means that, regardless of the interface used, a cloud storage gateway can be beneficial or even a downright requirement.

Another consideration is the unique security model associated with most cloud storage services. Access to block storage is usually controlled by region, for example using LUN masking to restrict access to a single host bus adapter (HBA) on a single server. Above this lies the file system, which controls access based on users, groups and directories in a tree. Cloud storage systems use an entirely different security model, one based on tenants, buckets and the objects. Although potentially more flexible, the cloud security model is entirely incompatible with conventional systems.

Enter cloud storage gateway vendors

A number of cloud storage gateway products have recently appeared to ease the transition between conventional and cloud storage interfaces. Each product focuses on a unique touch point between applications and their storage. Cloud gateways can make cloud storage appear to be a NAS filer, a block storage array, a backup target or even an extension of the application itself.

- Asigra Inc. produces an integrated backup solution for public and private cloud storage services with integration both on-site and at the service provider location.
- Cirtas Systems produces the Bluejet cloud storage controller, which allows public cloud storage services to be accessed as if they were on-site block storage arrays.
- Nasuni Corp. produces a virtual appliance that acts as a file server, complete with snapshots, caching and encryption.
- StorSimple Inc. produces a storage appliance with caching and provisioning, deduplication, encryption and WAN optimization targeted at SharePoint implementations.
- The TwinStrata Inc. CloudArray is a data protection and disaster recovery appliance integrated with public cloud storage.
- Other companies, like CommVault Technologies Inc. and Symantec Corp., have integrated cloud storage as a backup target on par with tape and disk.

The common theme for all these products, beyond bridging local applications to remote cloud storage, is the value-added features they introduce. Most include encryption technology to safeguard data stored off-site, as well as compression, deduplication and WAN optimization technology to accelerate performance. Many also take advantage of

the scalability of cloud storage for features like snapshots, version control and data protection. And in nearly every case, local storage is used as a cache for improved performance.

Is cloud data storage a tier?

Although it's tempting to see the integrated storage provided by these gateways as part of the overall trend toward tiered storage, cloud data storage can be much more than a low-cost tier. Indeed, each of these applications also demonstrates that the unique capabilities of cloud storage can provide new functionality previously unknown in conventional storage systems. Each cloud storage gateway vendor is blazing a new trail, bringing new capabilities rather than merely lowering the cost of storing data.

Free resources for technology professionals

TechTarget publishes targeted technology media that address your need for information and resources for researching products, developing strategy and making cost-effective purchase decisions. Our network of technology-specific Web sites gives you access to industry experts, independent content and analysis and the Web's largest library of vendor-provided white papers, webcasts, podcasts, videos, virtual trade shows, research reports and more —drawing on the rich R&D resources of technology providers to address market trends, challenges and solutions. Our live events and virtual seminars give you access to vendor neutral, expert commentary and advice on the issues and challenges you face daily. Our social community IT Knowledge Exchange allows you to share real world information in real time with peers and experts.

What makes TechTarget unique?

TechTarget is squarely focused on the enterprise IT space. Our team of editors and network of industry experts provide the richest, most relevant content to IT professionals and management. We leverage the immediacy of the Web, the networking and face-to-face opportunities of events and virtual events, and the ability to interact with peers—all to create compelling and actionable information for enterprise IT professionals across all industries and markets.

Related TechTarget Websites

- SearchDataBackup
- SearchCloudStorage
- SearchDisasterRecovery
- SearchVirtualStorage
- SearchSMBStorage
- SearchStorage