SecurityFirst™
Protecting Critical Data

WHITE PAPER

Why Object Storage is in Your Future

Most of us use it every day whether we realize it or not. The photos we take with our smartphones generally find their way to the cloud for safekeeping, where they're stored as objects. Some also get immediately posted to social media sites for sharing, which again uses an object storage repository rather than a more traditional hierarchical or block-based file system. The reason for this is simple: photos, videos, and other data we create consumes a lot of storage, and object storage systems are infinitely scalable.

What is Object Storage?

It's a relatively new technology primarily developed by Internet-based businesses including Amazon, Google, FaceBook, DropBox and others that required a more flexible form of storage to support their business models. The majority of this data is generally what's referred to as 'unstructured' data or information that doesn't fit within a fixed record length (database) format. Examples include documents, social media feeds, and digital pictures and videos.

An object storage system is organized as a flat plane, addressable as a single namespace. Each object has its own identifying details, comprised of flexible or custom metadata descriptors and a unique ID number that can be publicly accessible via Internet protocols (HTTP URLs) rather than storage commands. Data read and write operations are stateless, very simple and may be performed directly by a user's browser, via representational state transfer (RESTful web service), without having to go through the control of any server application.

Even though many are deployed within a cloud service, an object storage system can just as easily be an on-premises, cost-effective resource leveraging x86-based processing technology and commodity hard disk drives attached as a series of connected nodes. Data is protected using an erasure coding technique or a mathematical function to transform a set of data into a form that can be recreated from a subset of redundant pieces—providing additional benefits when using geographically dispersed datacenter locations.

But pure object storage is inadequate for some server workloads such as transactional data that changes frequently. The are no block-level, update-in-place capabilities. Instead, outdated information is deleted and new copies of the entire object are updated, then re-written in their entirety. Object storage is also not designed to replace network-attached storage (NAS) for shared file access because it doesn't have the locking and file-sharing facilities that ensure the single "truth" of a file.

Why Use It?

Beyond scalability, there are many reasons to still consider an object-based system to store your business and professional data compared to traditional file systems. Here are five more for starters:

1. Cheaper alternative for longer term storage
2. Customizable metadata for faster access
3. Built-in redundancy and resiliency without complicated fault-tolerant arrays
4. Easily accessible off-site storage
5. Reliability improvements over tape-based data backups

With today's exponential growth in data volume, it's most likely that only 10% of your data is used on a regular basis. And after ten or twenty days, that data may be rarely used again, but you can't delete it—at least not right away—and it probably has personal, sensitive, or confidential material that must be protected from unauthorized access. That's 90% of your data storage resources that could be better optimized and even outsourced to save building-out and managing multiple datacenters. Cloud Service Providers (CSPs) offer multiple pricing tiers where the bulk of the expense is incurred only when retrieving your data.

Each object has its own identifying details, comprised of metadata and an ID number, which the OS reads to retrieve data. Additional metadata descriptors can be generated and assigned to the objects providing faster searching using multiple indices. Also, without the need to trudge through file structures, retrieval is much faster, and users don't even need to know an object's exact location.

Traditional storage protections against physical device failures generally use replication, and some object storage systems do too, replicating at the object level across at least three locations. The main advantage of replication is its low computational overhead; the main disadvantage is its 200 percent storage overhead. To help reduce costs, better object storage systems use a parity-based protection method called erasure coding. Erasure coding systems divide objects into pieces called "shards," to which they add parity. The main disadvantage here is that it's more computationally expensive and can introduce latency during writes.

There are multiple cloud storage vendors from which to choose, and it makes a lot of sense to choose more than one. The ability to distribute your data will help you avoid any sort of cloud vendor lock-in situations and provide additional leverage against storage price increases. Vendors typically operate multiple locations distributed across multiple countries, helping address cross-border storage issues.

Magnetic tape-based backup usually offers the lowest cost per gigabyte of storage, but because tapes are a form of removable media, there can be challenges with achieving secure, off-site storage. It becomes less practical as data volumes increase, because of the sheer number of tapes required. Tapes are also a somewhat fragile media and have been known to fail during restoration attempts. Object storage backups leverage low-cost commodity disks and can offer significantly faster retrieval times.

But Can You Trust It?

Anytime a CIO considers outsourcing some of the organization's IT systems, the systems and suppliers must be thoroughly vetted. Overall cost savings and better, faster data access are enticing benefits for a new or replacement storage resource, but not at the expense of data security. If CIOs face challenges securing their own on-premises datacenters, why should anything be different for the CSP?

It's likely the case that a CSP will have better or at least additional resources to perform all the necessary system updates, scan for known vulnerabilities, and apply software patches in a more timely fashion because that's their whole business. But these vendors are not immune to zero-day exploits, employee mistakes and the potential for a malicious

insider attack. So how do you protect your data and protect your job when you move data into a cloud-based object store?

In most cases, CSPs will also take steps to secure the data within their environment, (server-side). However, this leaves big gaps in your security position: namely these gaps involve data in transition to the cloud (from on-premise to the cloud provider), and who is actually controlling the security (usually the CSP—not you).

The right move is to apply some client-side data protection to your data using encryption, pseudonymization, tokenization or other means and retain control of the decryption key or shared secret that returns clear text information to authorized users. Generate your own keys and store them on your systems away from the data. Then follow a best practices approach to rotating the keys on a periodic basis, and shredding them when the use for any data stored in the external object repository ends.

Maintaining this type of control ensures that only the original owners can access the data, and when you can finally delete the information, no spare or redundant copies are still floating around in somewhere waiting for some sort of space reclamation activity to overwrite them.

How Will You Use It?

Take your pick - just remember that it needs to be used securely:

1. Mass, scalable storage for 90% of your data
   a. Why? - because it is cheaper from a hardware, facilities, maintenance and technical resource perspective.
   b. Companies struggle with the on-going CAPEX needed to increase their storage capacity, and in most cases, the data-center infrastructure needed to house it. Turn that precious cash resource into a smaller OPEX with a cloud-based alternative.
   c. But be careful. Even with the 90% of the data you rarely need, it likely contains personally identifying, sensitive, or confidential data. Secure it or face the potential of significant fines if it' s ever stolen.

2. Backup/Recovery
   a. Why? – because it's cheaper, more efficient to restore, and more durable than tape archives.
   b. Provides faster access for restoring your data when disaster strikes.

3. Ransomware Mitigation
   a. Why? – because the U.S. Department of Homeland Security has stated that a backup/restore or disaster recovery solution is the only effective remediation or recovery from a ransomware attack.
   b. Best practice requires the backup to be secure and separated from the primary datacenter and network used for operations or production.

SecurityFirst DataKeep™ covers any potential gaps in your use of object storage repositories. It provides strong security, keeps your costs down, and puts you in control via its Object Store Agent:

- Starts with data protection on-premises, (client-side)
- Sends the data to object storage locally or in one or more clouds
- Splits the data to send shards (M:N recovery) into multiple clouds with resiliency, delivering the added benefit of avoiding CSP lock-in
- Supports BYOK (Bring Your Own Key) via KMIP compatible interface

In summary, you can finally leverage the low cost of object storage in the cloud with the trust needed to assure data privacy and resiliency. DataKeep and its Object Store Agent provide a superior means for storing infrequently used data within an off-site facility saving you the burden of planning for how and where your organization will retain exploding amounts of data in a broadly compliant manner.

About SecurityFirst

SecurityFirst™ delivers data-centric solutions that address the high-profile cyber threats facing organizations today, such as data breaches, ransomware and cloud security. We emphasize protection of the data itself to serve as your last line of defense. Data is always protected no matter where it resides and recoverable in the event of an unexpected failure or malicious attack such as malware or ransomware. As organizations and governments mandate stricter requirements for data privacy, SecurityFirst helps protect data from compromise and exposure.

About DataKeep

DataKeep™ secures critical data at its core to deliver unrivaled protection, control and resiliency. Customer-defined access policies, strong encryption and event logging combine with native secure backup/restore capabilities to address your data privacy, compliance and recovery needs.

Security teams can define and log access policies by job role and manage privileged and superuser access to block insider and external threats. DataKeep securely protects data at the source no matter where data resides, encrypting data at the volume or file level for attached storage or before sending to object storage. Native backup and restore commands can be leveraged to enable prompt recovery of archived data in the event of a ransomware attack.

DataKeep's ability to support M of N distributed shares allows companies to encrypt, split and distribute data across multiple object store locations or vendors for business continuity and operational efficiency. Organizations can utilize the backup and restore capabilities with object storage for secure cloud backup and archiving to improve resiliency. Secure your most valuable assets – your data, your brand and your reputation – with Datakeep.

SecurityFirst™
Protecting Critical Data

For a product demonstration or more information call 1-888-884-7152

securityfirstcorp.com