SwiftStack

SWIFTSTACK WHITEPAPER

Build the Business Case:

Object Storage for Backup, Archiving and Recovery

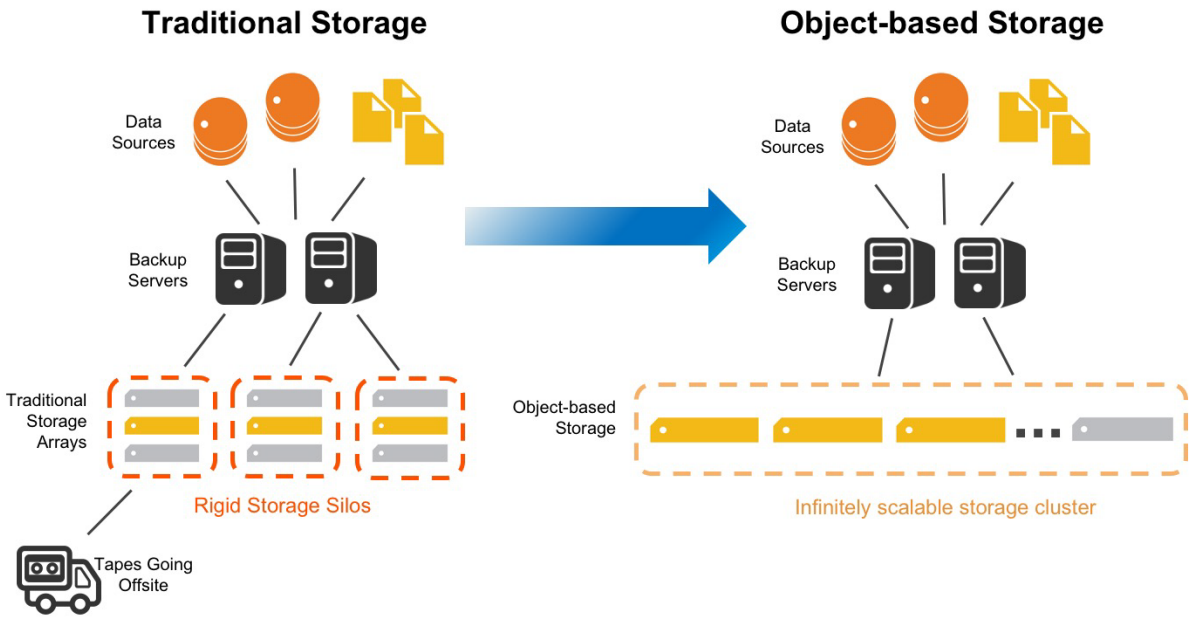
The dramatic and relentless growth of unstructured data is causing headaches for IT organizations and storage administrators. Not only are they looking to keep pace with the need to store more data than ever—without busting their budgets—but they are also facing a variety of operational, legal, regulatory and security pressures to retain, organize, manage and produce that data on a moment's notice. Data storage requirements are growing more than 50% annually, according to many research estimates, with unstructured data taking up the lion's share of that growth. This is making it harder than ever to back up, archive and recover data coming from many sources, physical locations, data formats, workloads, storage topologies, and compute architectures. While there are numerous ways of dealing with runaway storage growth for backup, archiving and recovery workloads, many organizations are constrained by such factors as cost, manpower, management complexity and performance. Among the challenges:

- Tape remains attractive for its low cost and ability to protect from a major disaster by transport ing it offsite, but it takes time to get tapes back onsite and restore attempts fail way too often.
- Hard disk-based arrays offer capacity and performance, but are traditionally challenging to scale and they can become very expensive. In an attempt to reduce costs, adopting proprietary deduplication appliances will just lock you into another hardware platform.
- Backing up to the public cloud gives you the disaster recovery ability that offsite tapes provide, it scales indefinitely, and it easy to use, but you lose fast local restores and your bandwidth out may not be sufficient to meet your defined data protection policies.

Considering Object Storage for Backup, Archiving and Recovery

For storage administrators looking for a more cost-efficient, reliable, scalable, and easier to use solutions for backup, archiving and recovery, object storage is an excellent destination target. Traditionally, backup applications needed disk storage to emulate tape or be locally mounted to the media server. Now, most leading backup applications support the Swift and S3 APIs, enabling them to utilize object storage. This type of storage offers several attractive capabilities for these workloads, including:

- Nearly unlimited scalability of capacity without reducing performance. No more running into hard capacity limitations of a storage silo, forcing you to recycle your backup data and/or manage it across many namespaces.
- Automatic disaster recovery protection by replicating data offsite. Cluster nodes can exist in multiple physical locations, so you do not need to worry about hauling media to the mountain anymore.
- Cost-efficient hardware infrastructure based on standard servers, Ethernet-based networking, and affordable disk drives rather than on more expensive proprietary solutions. That's the advantage of software-defined storage.



Scale Without Limits

Traditional storage is bound by the size of the disk pool and increasing the pool size requires significant processing time or is sometimes destructive. Also, if using protection methods like RAID5, the pool can only grow so large. Object storage is inherently designed to scale incrementally and can grow to many petabytes in size. Instead of your storage system dictating how you should manage your backup data, eliminate this significant boundary so you can manage your data higher up the stack.

For example, when the storage system tells you it's 80% full, start thinking about scaling up the storage capacity. To add storage, simply add another node to the cluster, which involves just a few simple software configuration steps. One step to install the software package and another couple of steps to include it in the specified cluster. When complete, the backup application automatically has more capacity to utilize.

Also with object storage, performance and capacity can be scaled independently based on the role of the node. If more performance is required, additional proxy nodes can be added. If more capacity is needed, storage nodes can be added. For smaller or starter environments, nodes can perform both tasks just fine. This is a set of tools that backup administrators did not have at their disposal before, allowing the environment to be managed over time with no storage disruption.

Automatic Disaster Recovery

Even with multiple data copies, if they all reside in a single physical location, you are not protected from a major disaster like fire, flood, and theft. Instead of having a service haul your tapes offsite or using supplemental software to replicate data to another data center, let the object-based storage cluster do it for you.

As you're thinking of ways to better protect all of your data, you most likely do not want to consider wholly replacing on-premises backup with public cloud backup because you would eliminate fast restore times. Have the best of both with fast local backup and restore along with automatic offsite protection.

This functionality exists at the core of object-based storage, where nodes of a cluster can exist in multiple physical

locations. Data will be redundantly written to local nodes quickly, allowing the application to move on. Then the storage cluster copies the data over the slower pipe to the other site. This is referred to as eventual consistency. Since the offsite node is a member of the cluster, if a node local to the backup server fails, it will just access the data from the offsite node. It's a single namespace, so the backup server does not know the difference. This significantly reduces restore times from offsite media by eliminating the need to first replicate the data back so the backup server can access it.

Significantly Lower Cost Hardware

Since the beginning of time, it seems like there's always a promise to lower cost, when in reality, one vendor just wants you to transfer your purchases from another vendor to them. With a truly hardware-independent software-defined storage solution, the savings actually are real.

To start, use your server of choice. It's most likely the case that you're already purchasing a few server models that meet the needs of your scaling applications. Why not just use those servers for your storage solution as well? Also, run the storage software on top of the operating system that you already know how to manage and have most likely optimized for your needs. And last but not least, eliminate the drive tax by procuring drives from the best source that are fit for your application at hand. No more having to choose drives that don't best meet your needs and cost significantly more because they come from your array vendor.

Considering SwiftStack

As object-based storage continues to gain interest and acceptance by IT professionals and storage administrators, organizations are seeking out and evaluating potential suppliers of object storage to meet the growing capacity, management and economic requirements for backup workloads.

SwiftStack had been the lead contributor to the widely used Swift project in OpenStack and is led by some of the best in the application, cloud, and storage spaces. While SwiftStack still contributes significantly to the core, it has recently released its 3rd major version of its enterprise-grade software-based storage solution, called the SwiftStack Storage System. This solution is not hardware-centric, allowing you to choose the servers, drives, and networking that best meets your needs without the extra-cost built-in to traditional storage.

The SwiftStack Storage System is comprised of the out-of-band controller, storage nodes, and an optional filesystem gateway. This solution is built on the OpenStack Swift Object Storage Engine and contains all of the deployment, management, and analytics functionality you'd expect from modern, enterprise-grade storage.

The new version of SwiftStack adds some great features for data protection solutions:

AWS S3 API emulation - if your app can write to S3, it can now write to SwiftStack.

Erasure codes - increases storage efficiency by up to 50%, which means you can back up more data per node.

Flexible controller model - the storage controller can be hosted by SwiftStack or deployed behind your firewall, providing options so you have visibility to your globally-distributed storage nodes from anywhere.

The SwiftStack Controller supports a browser-based dashboard that enables multiple operators to have role-based access control, while providing single-pane-of-glass management of a distributed storage architecture for backup, archiving and recovery. It enables consistent configuration management across multiple nodes, while supporting reports for capacity planning and storage utilization for chargebacks.

Conclusion

As backup, archiving and recovery workloads become larger, more complex, and more important, IT organizations have become severely challenged. They must keep up with fast-mounting data backup capacity requirements, as well as the need for intelligent and highly automated archiving solutions and reliable and instantaneous recovery. Of course, all this must be accomplished against the reality of tighter storage infrastructure budgets.

Object-based storage has emerged as an excellent option for backup, archiving and recovery because of its infinite capacity scalability, automatic protection from a major disaster, and cost-reducing attributes, all while being simple to deploy and manage. This approach delivers the advantages of public cloud storage with the custody of your data on-premises. SwiftStack is the leading provider software-based object storage built on the power of open source.

Learn about SwiftStack's backup solution partners and how others are leveraging object-based storage in their data protection environments at SwiftStack.com/backup.

COMMVAULT

VERITAS

trilioData
Business Assurance Platform

IBM

rubrik

CloudBerry Lab
Cloud Storage Tools