



SWIFTSTACK WHITEPAPER

An IT Expert Guide:

Understanding Object Storage and How to Use It

November 2014

The explosion of unstructured data is creating a groundswell of interest in object storage, certainly among cloud services providers, but also within enterprise IT departments. Object storage is infinitely more scalable than traditional file and block storage solutions, and is far more cost efficient, flexible and manageable in addressing the challenges of dealing with large data volumes in multi-use cloud environments.

Object storage has been around for a while but began gaining widespread attention in recent years as the storage platform for some of the largest-scale compute environments in the world, most notably Amazon S3 and Google Cloud Storage. The challenges faced by Amazon and Google in cost effectively delivering massive scale for unstructured data are now being encountered by enterprise IT organizations, particularly with the growth of private and hybrid clouds.

For many IT professionals, object storage is like a shiny new object that sounds and looks pretty but doesn't come with a clear set of instructions on how, when and why it should be used. That's why we're here. This white paper provides an overview of object storage and describes the use cases where it can be most effective in addressing the challenges facing enterprise IT. In addition, we explain how object storage fits into a software-defined architecture and what to look for in a solution.

Part 1: Understanding Object Storage

To understand the difference between object storage and traditional file/block storage, it is instructive to look at the history of files. File storage evolved from physical filing cabinets and desktops, and those ways of organizing data defined computing storage—built on “file systems” that used “folders” to organize data. As technology has evolved, however, we now expect that our data can be shared and accessible across all of our devices. The old network-attached storage or storage area network approach of data as documents stored in file folders breaks down when there is a need to have simultaneous access to data across many devices, as is the case in cloud environments and other modern architectures.

Object storage was designed to replace existing data organization paradigms, such as putting one folder in another or allowing only a single person to edit a file at a time. It treats everything as a piece of data, equal to every other piece of data. This allows object storage systems to store and share data across millions of computers and devices. With object storage, applications don't have to spend time with tasks around file locking, hardware failures or capacity management. Instead, applications can focus on their own functionality and treat the storage as a utility resource.¹

Two other important points about object storage:

- Software-defined object storage solutions can be deployed on commodity server hardware and also use the same spinning disks and solid-state disks as in legacy storage environments. You don't have to rip and replace everything in order to deploy an object storage platform and, in fact, with specific file system gateway solutions, you can expose object-stored data as files and integrate them with file-based applications. Also, with the right solution in place you can utilize hardware from multiple vendors while mixing storage node density and storage device size.

¹ “Object Storage vs. File Storage,” Enterprise Storage Forum, Sept. 26, 2014, interview with John Dickinson, director of technology, SwiftStack

- Object storage is a different way of organizing data on a disk. For many use cases—and particularly for unstructured data—it is a more efficient way of storing data. Like files, objects contain data; but, unlike files, objects are not organized in a hierarchy. Every object exists at the same level in a flat address space called a storage pool. Objects are characterized by their metadata; each object is assigned a unique identifier, which allows a server or end user to retrieve the object without needing to know the physical location of the data.²

Part 2: Object Storage Use Cases

One of the most important benefits of object storage is that it enables IT to easily and effectively deliver Storage as a Service. You can manage a larger pool of storage that is inherently multi-tenant, with massive scale that traditional storage architectures either can't deliver at all or can deliver only at an exorbitantly expensive cost. In addition to massive scale, object storage delivers several other major advantages versus traditional file and block approaches, including:

- The ability to distribute data more widely in either multiple data centers or across the far corners of a single data center.
- Better availability of the data under all sorts of different failure scenarios because data is automatically replicated between nodes, clusters and sites.
- The ability to manage all of your storage resources as a single resource pool so that you never have to overprovision resources, and you always deliver the right storage to the right workload at the right time.
- The ability to serve multiple use cases with the same storage footprint.

Fundamentally, with object storage you are not presenting a consistent block or file view of the data. If the system tried to present a consistent view, it would have to wait for all of the redundant copies to make all of the changes. With an object storage platform, there are strategies within the system to automatically reconcile changes.

Given these benefits, there are specific use cases where object storage offers significant advantages over other solutions, with the two most prominent being:

1. **Storing and managing unstructured data:** Object storage makes it much easier to manage, store and scale unstructured data because it provides structure to this data by treating it as objects, just as it treats structured data. This is a huge breakthrough in the management of unstructured data at a particularly critical time: The amount of data is doubling every two years,³ and it has been estimated that approximately 90% of that data is of the unstructured variety. In any application where there are significant amounts of unstructured data—think documents, images, video, backups and archives—object storage will be absolutely critical to keeping costs under control, enabling scalability, supporting high availability and reducing management complexity.

² "Object storage," SearchStorage.com

³ "The Digital Universe of Opportunities: Rich Data and the Increasing Value of the Internet of Things," IDC, April 2014

- 2. Applications support in the building out of IT as a Service:** Given the inherent multi-tenant nature of object storage, as well as its single shared storage footprint and massive scalability, it is clear why leading cloud vendors such as Amazon and Google have deployed it to enable their cloud services. Another significant advantage of an object storage platform in the delivery of IT as a Service is that it is programmable, so that developers can actually build applications around the storage environment.

Part 3: Why to Use a Software-Defined Approach

To take full advantage of the benefits of object storage, it is necessary to utilize a software-defined approach. With software-defined storage, there are basically three layers:

- The access layer: how you get data in and out, including, in the case of object storage, a file system gateway.
- The actual storage itself, i.e., the disk.
- A control plane to ensure availability, performance, capacity, etc.,

In an object storage platform, the control plane is critical in enabling you to abstract and manage all of your storage resources so that they appear as—and function as—a single resource pool. This approach allows for the use of commodity hardware because the nature of software-defined object storage is much more forgiving in the failure of components. A software-defined object storage system can sustain significant hardware failures without data loss.

An important factor in choosing the right solution for object storage is to deploy a management platform to ensure that you can reduce the total cost of ownership of running, managing and maintaining the infrastructure. With the right platform in place you can leverage the lower costs of commodity hardware to, in effect, create a new paradigm in how storage is deployed, managed and scaled.

Part 4: What to Look for in a Solution

Given the relentless growth of unstructured data and the shift to cloud computing models, it is almost a given that most enterprises—and certainly any organization dealing with multi-petabytes of unstructured data—will be deploying object storage in the near future, if they aren't doing so already.

As you evaluate what to look for in an object storage platform, there are several important characteristics to consider. One is the opportunity to deploy an open-source solution. There are many benefits of going the open-source route, such as access to a wide community of developers and the ability to avoid vendor lock-in. In object storage, open source is a particularly attractive option because there is a huge base of support for OpenStack Swift, which already powers some of the largest cloud deployments in the world.

In deploying an object storage platform based on Swift, you will also need to work with a company that has expertise in deploying Swift in enterprise environments and delivers a full set of solutions to maximize its benefits. These solutions should include:

- **A software-defined controller** that provides improved visibility and control into a storage system even when it spans multiple data centers. The controller should enable operators to manage and scale their storage environment from a single pane of glass.
- **A file system gateway** that enables you to integrate your object storage platform with existing file-based applications.
- **Enterprise authentication integration** to provide easy plug-and-play integration with enterprise management systems such as LDAP and Active Directory.
- **Global data distribution** so that data can be distributed across multiple regions in a single cluster, thus reducing latency for data access and ingest while providing high availability.
- **Flexible storage policies** to help simplify management and lower the overall costs of the storage infrastructure.
- **Heterogeneous hardware support** to leverage existing investments and allow you to turn a multitude of disparate hardware into a cohesive storage system.

Part 5: Taking the Next Step

Every enterprise should be at least looking at object storage, and any enterprise dealing with massive amounts of unstructured data or deploying IT as a Service through the cloud should strongly consider adding object storage to its storage environment.

In taking the next step, it is important to work with a vendor that has the highest levels of expertise in object storage so that you can really understand how to leverage total cost of ownership and take the best advantage of this major innovation in data center infrastructure. You also want to consider a vendor that supports OpenStack Swift and delivers all of the major features and functions described above.

When you look at the field of vendors in object storage, you will see very clearly why SwiftStack is considered one of the leaders in object storage software and why it is enabling customers to derive unprecedented value and innovation for their storage environments. Are you ready to take the next step in object storage? **Here's where to start.**