File, Block and Object: Storage Fundamentals in the Cloud Era

TechTarget®

In this handbook:

File, block and object:
Storage fundamentals in the
cloud era

While on-premises hardware block, file, and object storage products are available, the increase in cloud adoption has also led to hybrid and distributed forms of these storage options.

What's more, although file, object, and block are long-running fundamentals of storage, the ways they are being deployed in the cloud era are changing.

To ensure you maintain a pulse on cloud-based file, object, and block storage developments, browse this custom TechTarget e-guide, which shares the benefits, fundamental principles, and deployment considerations for each approach.

Sponsored by:

TechTarget | PARSEC LABS

File, block and object: Storage fundamentals in the cloud era

ANTONY ADSHEAD, STORAGE EDITOR

Despite the many changes in data storage over the decades, some fundamentals remain. One of these is that storage is accessed by one of three methods – block, file and object.

This article will define and expand on the characteristics of these three, while also looking at the on-prem and cloud products you will typically find that use file, block and object storage.

What we see is that while on-prem (usually) hardware form factor block, file and object storage products are available, these types of access to storage are also offered in the cloud to serve the workloads there that require them.

The rise of the cloud has also led to hybrid – datacentre and cloud – and distributed forms of file and object storage.

So, although file, object and block are long-running fundamentals of storage, the ways they are being deployed in the cloud era are changing.

Sponsored by:

TechTarget | PARSEC LABS

FILE AND BLOCK: WHOLE AND PART

The file system has always been a mainstay of storage technology. Block and file access storage offer two ways to interact with the file system.

File access storage is when you access entire files via the file system. Usually that is via network-attached storage (NAS) or a linked grid of scale-out NAS nodes. Such products come with their own file system on board and storage is presented to applications and users in the drive letter format.

In block access, the storage product – usually deployed on-prem in storage-area network (SAN) systems, for example – only addresses blocks of storage within files, databases, etc. In other words, the file system that applications talk through resides higher in the stack.

File systems give all sorts of advantages. Among the most prominent is that this is how most enterprise applications are written – and that won't go away too soon.

A key characteristic of file system-based methods is that there are methods – such as those found within the Posix command set – to lock files to ensure they cannot be simultaneously over-written, at least not in ways that corrupt the file or the processes around it.

Sponsored by:

TechTarget   PARSEC LABS

File storage accesses entire files, so it gets used for general file storage, as well as more specialised workloads that require file access, such as in media and entertainment. And, in its scale-out NAS form, it is a mainstay of large-scale repositories for analytics and high-performance computing (HPC) workloads.

Block storage provides application access to the blocks that files comprise. This might be database access where many users work on the same file simultaneously and from possibly the same application – email, enterprise applications such as enterprise resource planning (ERP), for example – but with locking at the sub-file level.

Block storage has the great benefit of high performance, and not having to deal with metadata and file system information, etc.

FILE AND BLOCK: CLOUD AND DISTRIBUTED

File storage still exists in standalone NAS format, especially at the entry level, and scale-out NAS, intended for on-prem deployment, is commonplace.

But the advent of the cloud, and its tendency to globalise operations, has affected things has had a twofold effect.

On the one hand, there are a number of suppliers that offer global file systems that combine a file system distributed across public cloud and local network

Sponsored by:

TechTarget | PARSEC LABS

hardware, with all data in a single namespace. Providers here include Ctera, Nasuni, Panzura, Hammerspace and Peer Software.

On the other hand, all the key cloud providers – Amazon Web Services, Google Cloud Platform and Microsoft Azure – offer their own file access storage services, and also those of NetApp, in the case of AWS. IBM also offers file storage though its cloud offering.

BLOCK IN THE CLOUD

Some storage suppliers, such as IBM and Pure, offer instances of their block storage in the cloud. And the big three all offer cloud block storage services, aimed at applications that require the lowest latency, such as databases and analytics caching, as well as virtual machine (VM) work.

Probably because of the nature of block storage and its performance requirements, no distributed block storage seems to have emerged in the way it has with file.

OBJECT STORAGE: A WORLD APART

Object storage is based on a "flat" structure with access to objects via unique IDs, similar to the domain name system (DNS) method of accessing websites.

In this handbook:

File, block and object:
Storage fundamentals in the
cloud era

For that reason, object storage is quite unlike the hierarchical, tree-like file system structure, and that can be an advantage when datasets grow very large. Some NAS systems feel the strain when they get to billions of files.

Object storage accesses data at the equivalent of file level, but without file locking, and often more than one user can access the object at the same time. Object storage is not strongly consistent. In other words, it is eventually consistent between mirrored copies that exist.

Most legacy applications are not written for object storage. But far from that necessarily being a disadvantage, historically speaking, object storage is in fact the storage access method of choice for the cloud era. That is because the cloud is generally far more of a stateless proposition than the legacy enterprise environment, and also comprises probably the bulk of storage offered by the big cloud providers.

Also, objects in object storage offer a richer set of metadata than in a traditional file system. That makes data in object storage well-suited to analytics, too.

OBJECT IN THE CLOUD – AND ON-PREM WITH FILE

The cloud has been object storage's natural home. Most storage services offered by cloud providers are based on object storage, and it is here that new de facto standards, such as S3, have emerged.

TechTarget   Sponsored by: PARSEC LABS

File, block and object:
Storage fundamentals in the
cloud era

With its easy access to data that that can happily exist as largely stateless and eventually consistent, object is the bulk storage of the cloud era.

You can get object storage for on-prem deployment, such as Dell EMC's Elastic Cloud Storage, which is solely for datacentre deployment. Meanwhile, Hitachi Vantara's Hitachi Content Platform, IBM's Cloud Object Storage and NetApp's StorageGrid can operate in hybrid- and multicloud scenarios.

Some specialist object storage suppliers, such as Cloudian and Scality, offer on-prem and hybrid deployments.

And in the case of Scality, along with Pure Storage (and NetApp, to an extent), converged file and object storage is possible, with the rationale here being that customers increasingly want to access large amounts of unstructured data that may be in file or object storage formats.