

# Critical Capabilities for Distributed File Systems and Object Storage

Published 25 October 2021 - ID G00738193 - 48 min read

By Analyst(s): Chandra Mukhyala, Julia Palmer, Jerry Rozeman, Jeff Vogel

Initiatives: [Data Center Infrastructure](#)

Vendors are increasingly investing in a single unstructured data storage platform for serving both file and object use cases. In this research, we evaluate 17 file and object storage products against eight critical capabilities in seven use cases important to I&O leaders.

**This Critical Capabilities is related to other research:**

[Magic Quadrant for Distributed File Systems and Object Storage](#)

[View All Magic Quadrants and Critical Capabilities](#)

## Overview

### Key Findings

- The feature sets of distributed file systems and object storage products are coming together, resulting in the merging of the use cases.
- Object storage products are being used for workloads beyond secondary storage use cases such as cloud-native applications and analytics archives.
- Unstructured data storage vendors are continuously improving their public cloud integration capabilities, simplifying the extended use for hybrid cloud storage use cases.

### Recommendations

Infrastructure and operations (I&O) leaders responsible for unstructured data infrastructure should:

- Avoid creating a file only or object only data storage silos by selecting a common unstructured data platform, if it meets all your application requirements.

- Improve data management and provide actionable insights to the business by selecting unstructured data platforms with strong data insights and life cycle management, including cyberprotection capabilities.
- Enable hybrid cloud storage to address use cases such as disaster recovery (DR), burst for capacity and burst for processing by shortlisting vendors with the ability to integrate with the public cloud.

## Strategic Planning Assumptions

By 2026, large enterprises will triple their unstructured data capacity stored as file or object storage on-premises, at the edge or in the public cloud, compared to 2021.

By 2025, 60% of the global unstructured data storage capacity will be deployed as software-defined storage (SDS), up from less than 25% in 2021.

## What You Need to Know

Starting with this Critical Capabilities research, Gartner is combining the evaluation of object storage and distributed file system products into a single report. The decision to combine the Critical Capabilities research for file and object storage products into a single research note reflects what Gartner is seeing from the vendors in this space and with adoption by I&O teams. Data centers, whether at the core, cloud or edge, are dealing with more amounts of unstructured data than structured, transaction-oriented data. Distributed file systems and object storage are the most common solutions deployed by infrastructure teams to address the challenges with unstructured data.

During the early years of object storage, there were distinct features object storage products offered compared to distributed file systems. Fast-forward to 2021, and the difference in capabilities to address unstructured data challenges between object storage and file systems have become minimal. Gartner inquiries suggest that many end users do not care whether the solution they deploy is based on a file system or an object storage, if they can access the data through their preferred access protocol. While not exactly same as the capabilities being evaluated in this report, the foundational capabilities that I&O teams look for when considering a solution for unstructured data include:

- Scalability in performance
- Scalability in capacity
- Availability and resilience

- Manageability
- Security
- Multitenancy
- Flexibility in deployment

All of the above capabilities can be addressed equally well with either file or object storage. The difference used to be in access protocols supported by either system, but today's products offer both file interfaces like Network File System (NFS), Server Message Block (SMB) and Hadoop Distributed File System (HDFS), and object interfaces like Amazon S3. Most vendors that used to offer separate products for file and object storage have decided to offer a single offering going forward to address both.

Beyond the foundational capabilities, a few additional capabilities have emerged recently to address the unique challenges of unstructured data. These include:

- Metadata analysis and tagging to enable business workflows
- Life cycle management of data from placement to tiering
- Ransomware detection and protection
- Hybrid cloud connectivity
- Support for cost-effective flash

Again, all the above capabilities can be equally addressed with file and object storage. The vast majority of use cases can be satisfied with either a file or object storage. As a result, vendors are offering a single product, and I&O teams are consolidating to a single product.

There are a small number of use cases where how the data is laid out — as a file system tree or a flat name space — does matter, and thus a stand-alone file or object storage would be a better fit. Specific workload characteristics that would benefit from a dedicated file or object storage as opposed to a unified storage include:

- Frequent data manipulation is best addressed by a file system because of its ability to modify data at a granular level. Additionally, objects in most object storage systems are immutable by default.

- Storing data as key-value pairs to allow access of objects using custom metadata is best addressed by native object storage systems, as opposed to file systems offering S3 interface.

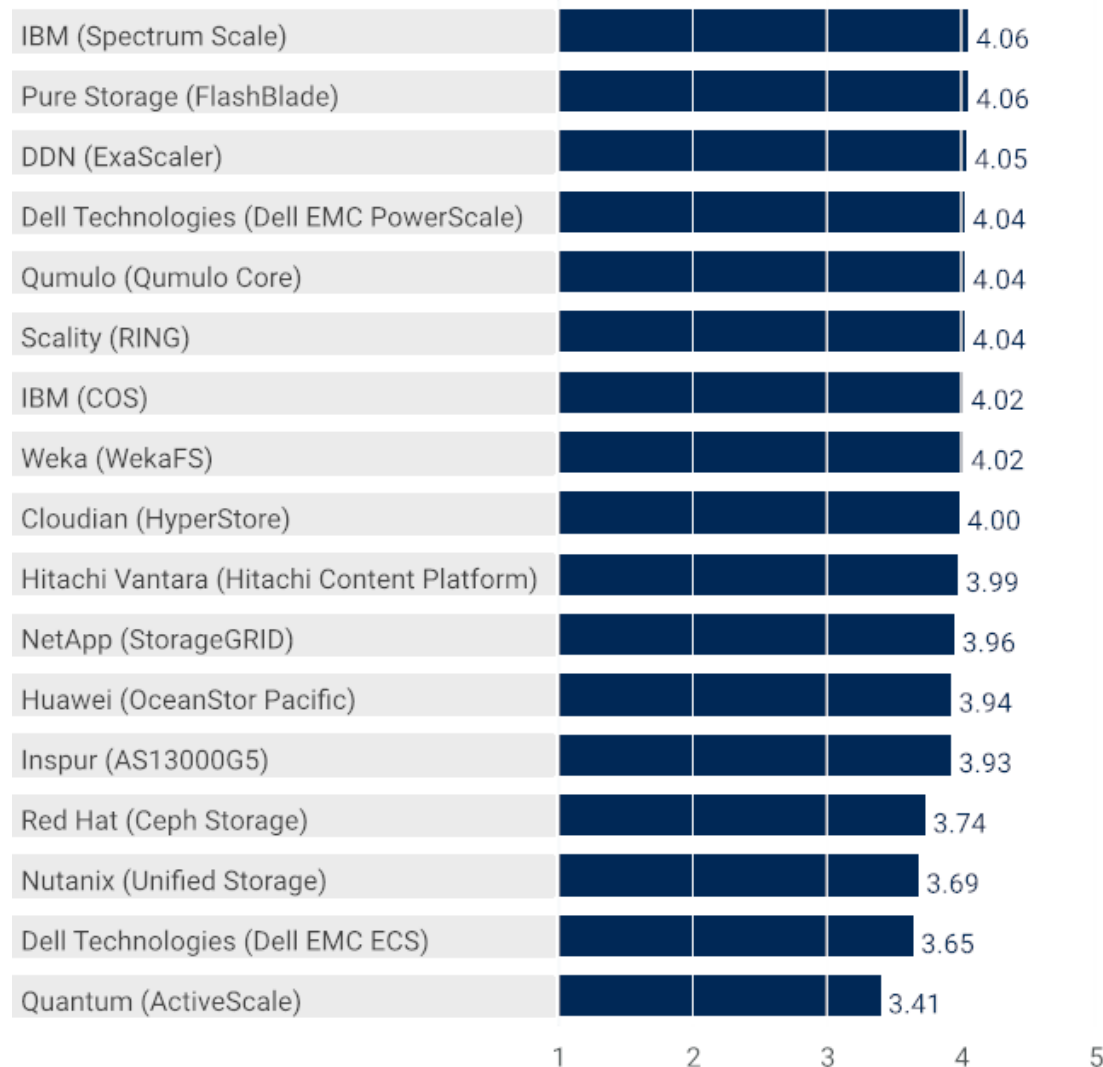
The research in this note evaluates products that are best-suited for file system-specific use cases, and products that are best-suited for object storage-specific use-cases, and other products that can address either use case. Not all use cases are applicable to all products. For example, commercial high-performance computing (HPC) is not applicable to products based on object storage, and the cloud-native applications use case is not applicable to most file system-based products.

## Analysis

### Critical Capabilities Use-Case Graphics

#### Vendors' Product Scores for the Analytics Use Case

Product or Service Scores for Analytics



As of 24 September 2021

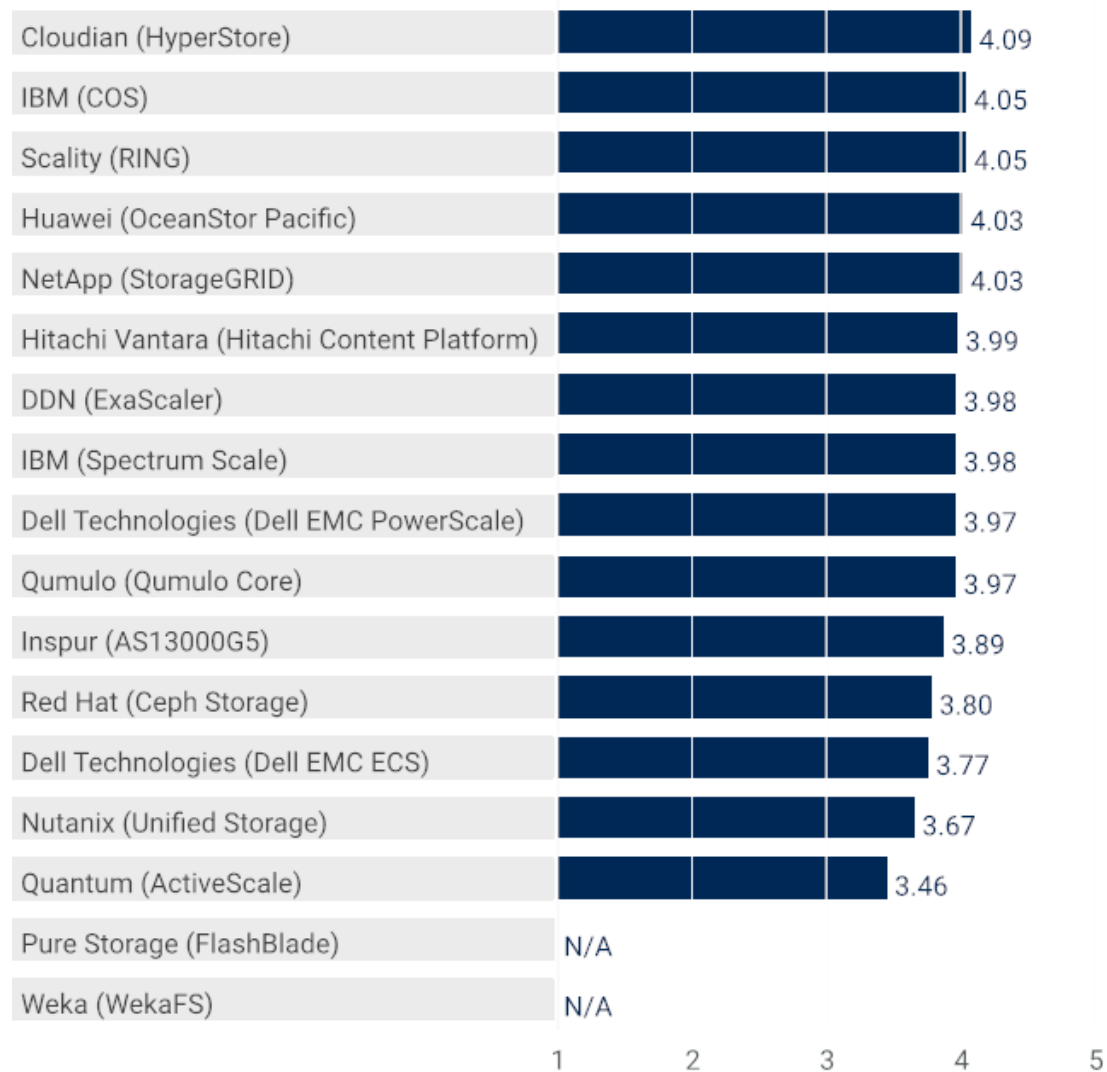
© Gartner, Inc

**Gartner**

Source: Gartner (October 2021)

## Vendors' Product Scores for the Archiving Use Case

Product or Service Scores for Archiving



As of 24 September 2021

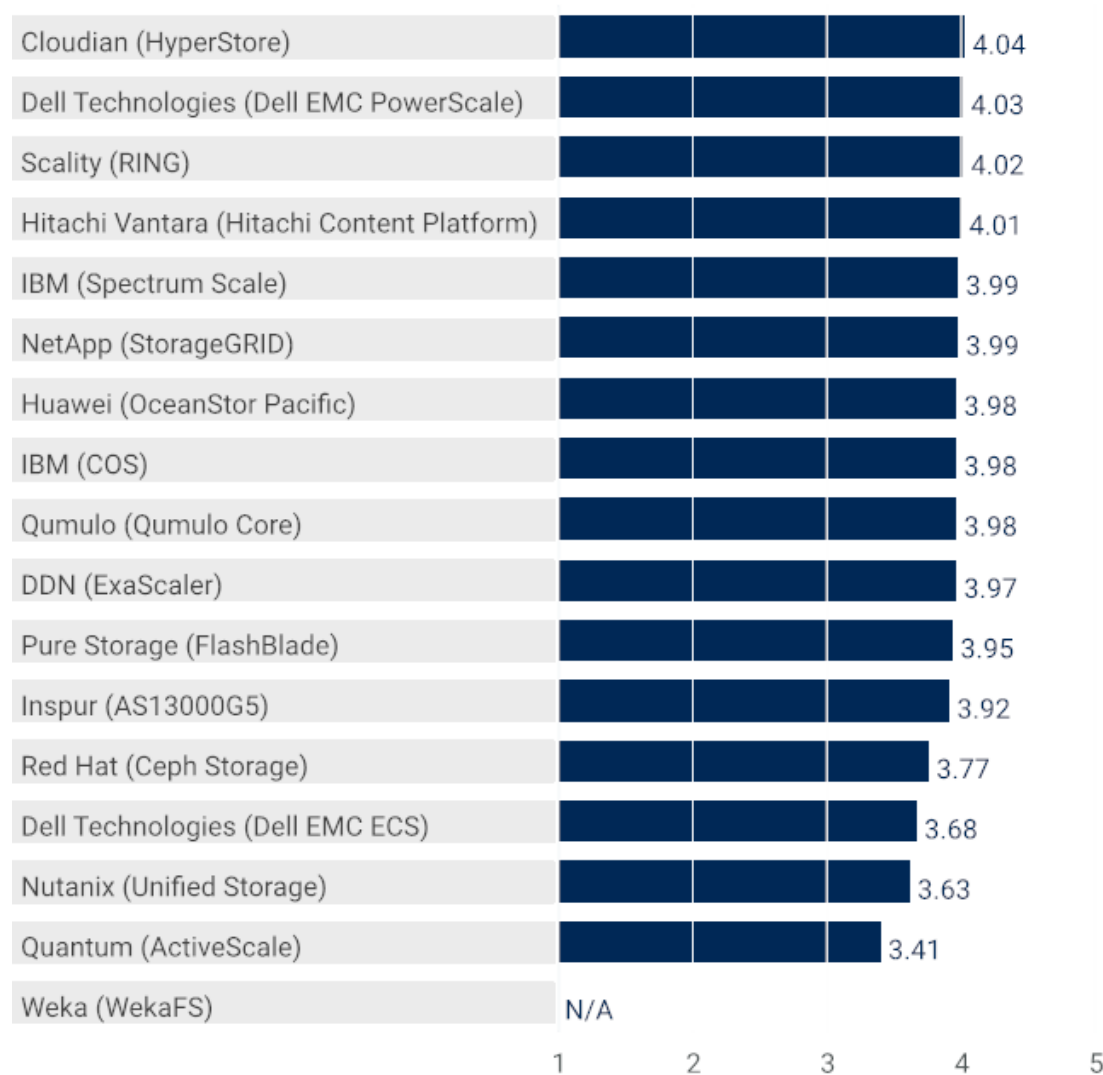
© Gartner, Inc

**Gartner**

Source: Gartner (October 2021)

## Vendors' Product Scores for the Backup Use Case

Product or Service Scores for Backup



As of 24 September 2021

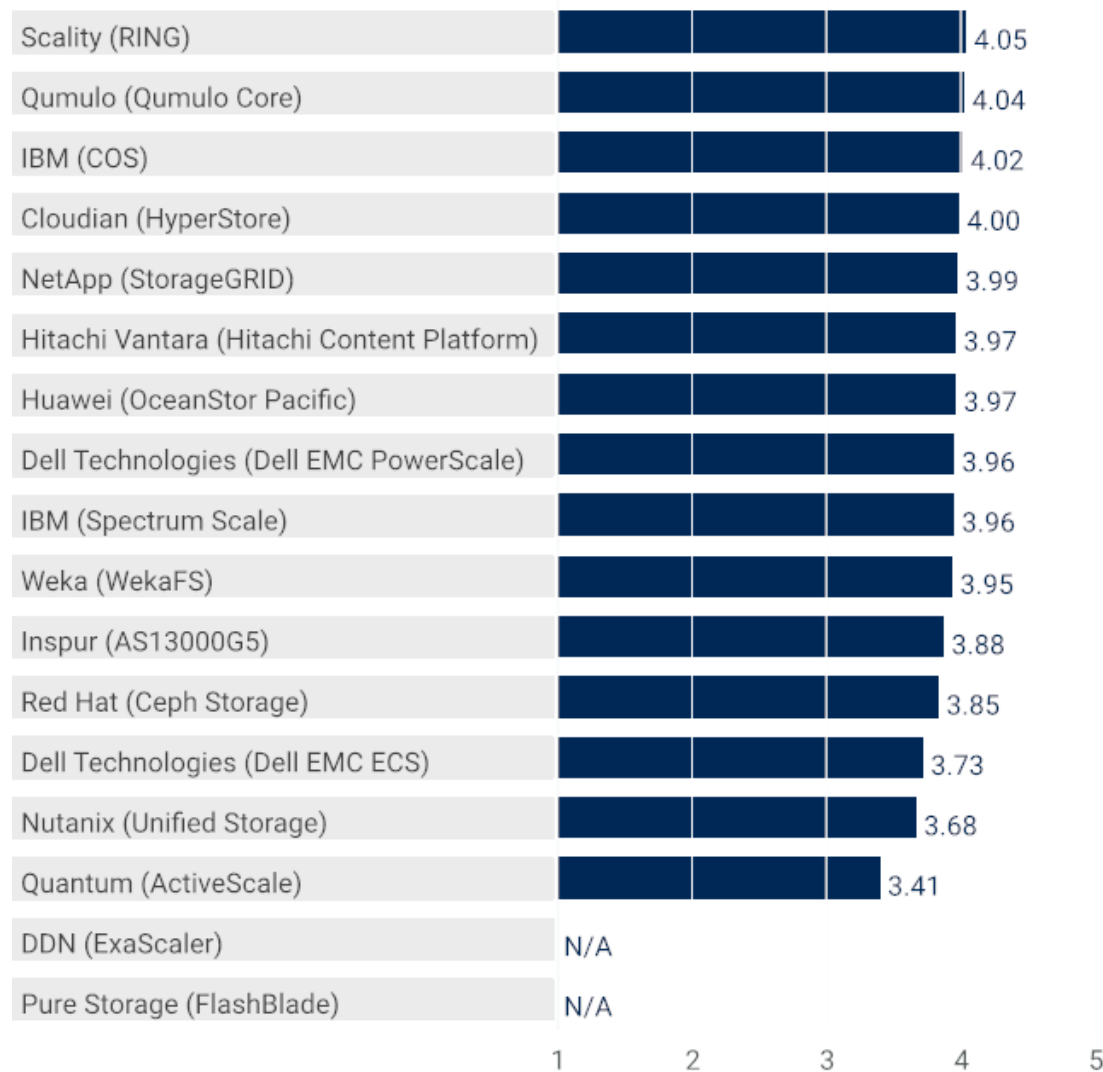
© Gartner, Inc

**Gartner**

Source: Gartner (October 2021)

## Vendors' Product Scores for the Hybrid Cloud Storage Use Case

Product or Service Scores for Hybrid Cloud Storage



As of 24 September 2021

© Gartner, Inc

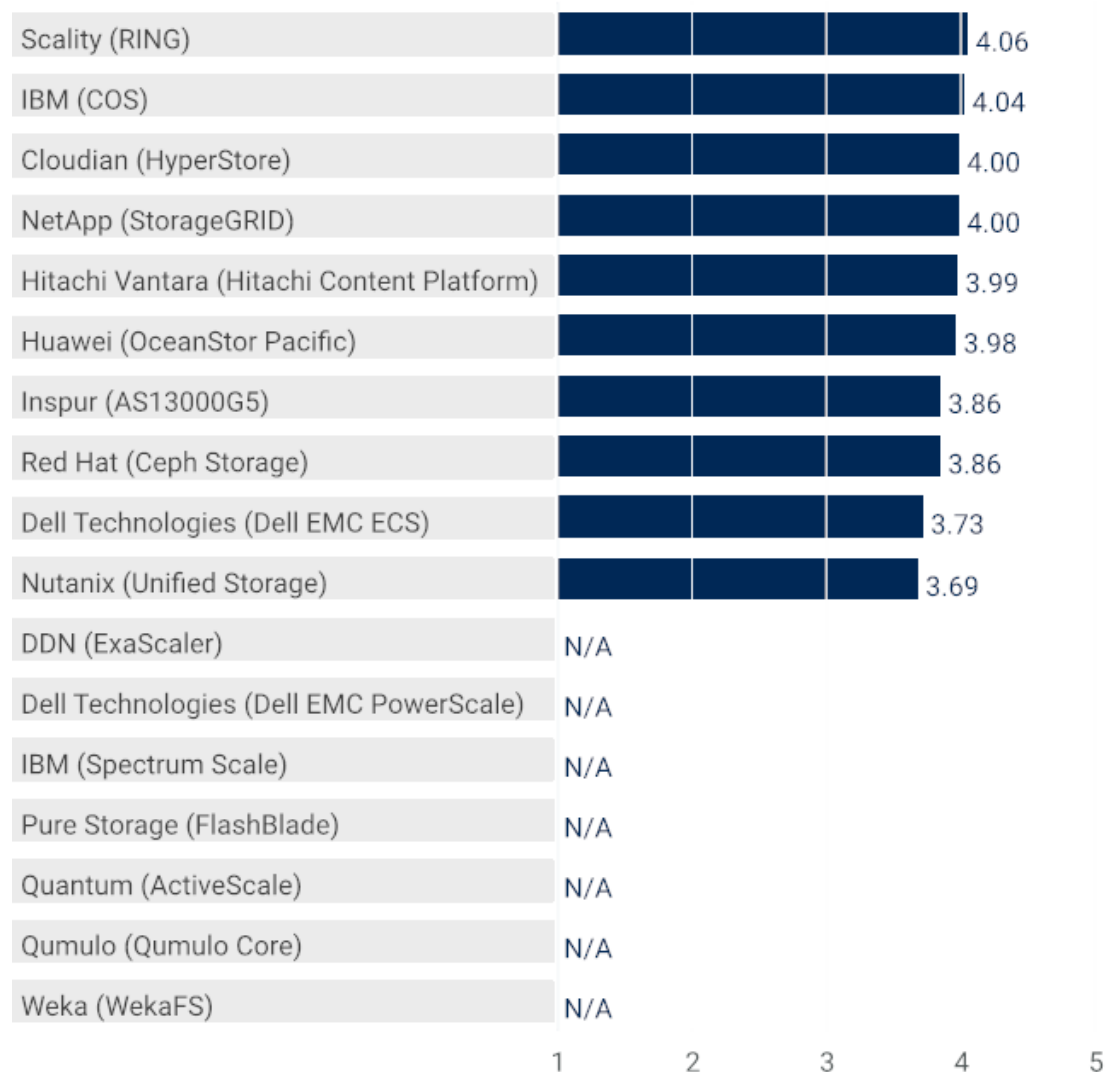
Gartner

Source: Gartner (October 2021)



## Vendors' Product Scores for Cloud-Native Applications Use Case

Product or Service Scores for Cloud-Native Applications



As of 24 September 2021

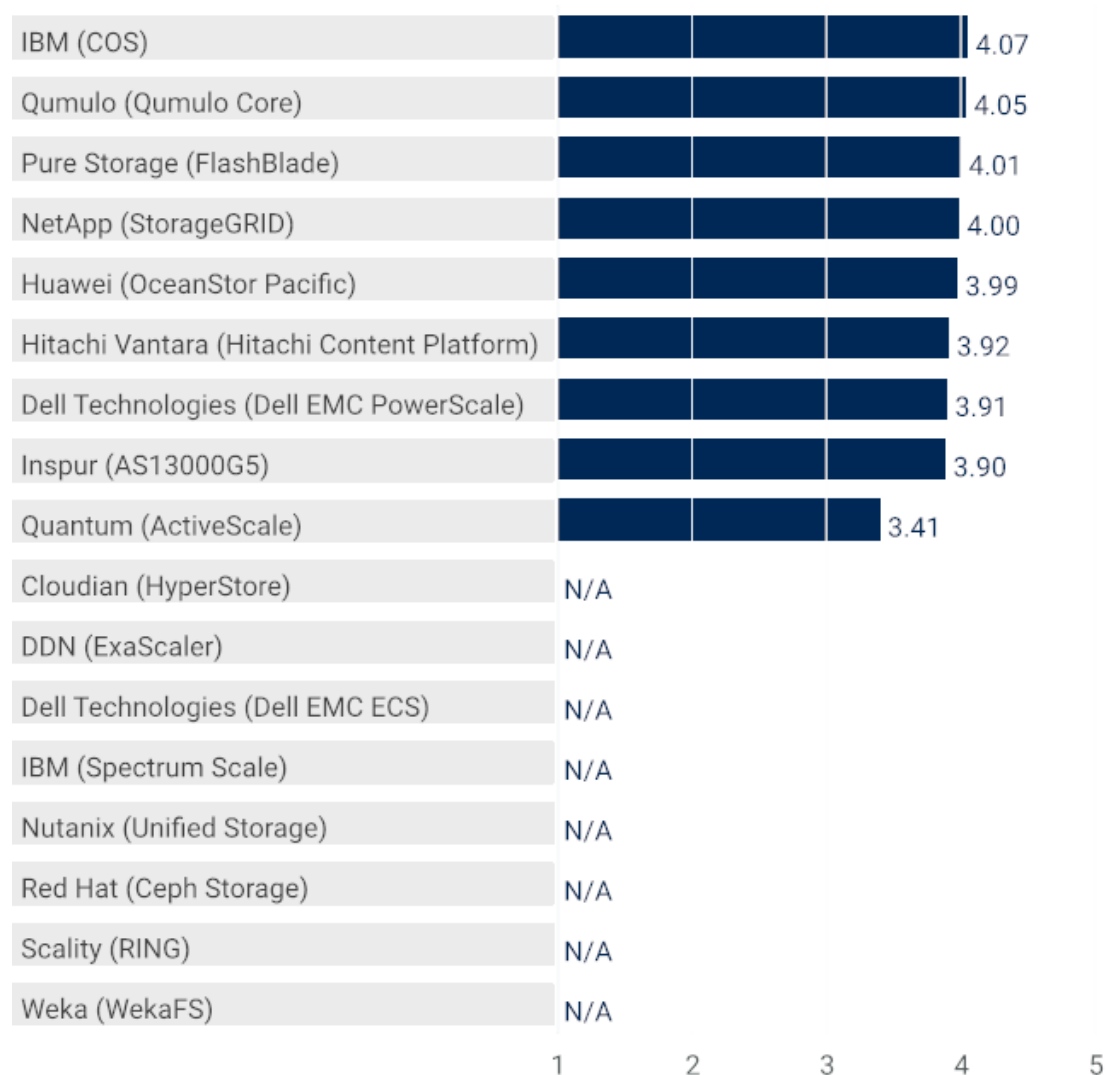
© Gartner, Inc

**Gartner**

Source: Gartner (October 2021)

## Vendors' Product Scores for the Cloud IT Operations Use Case

Product or Service Scores for Cloud IT Operations



As of 24 September 2021

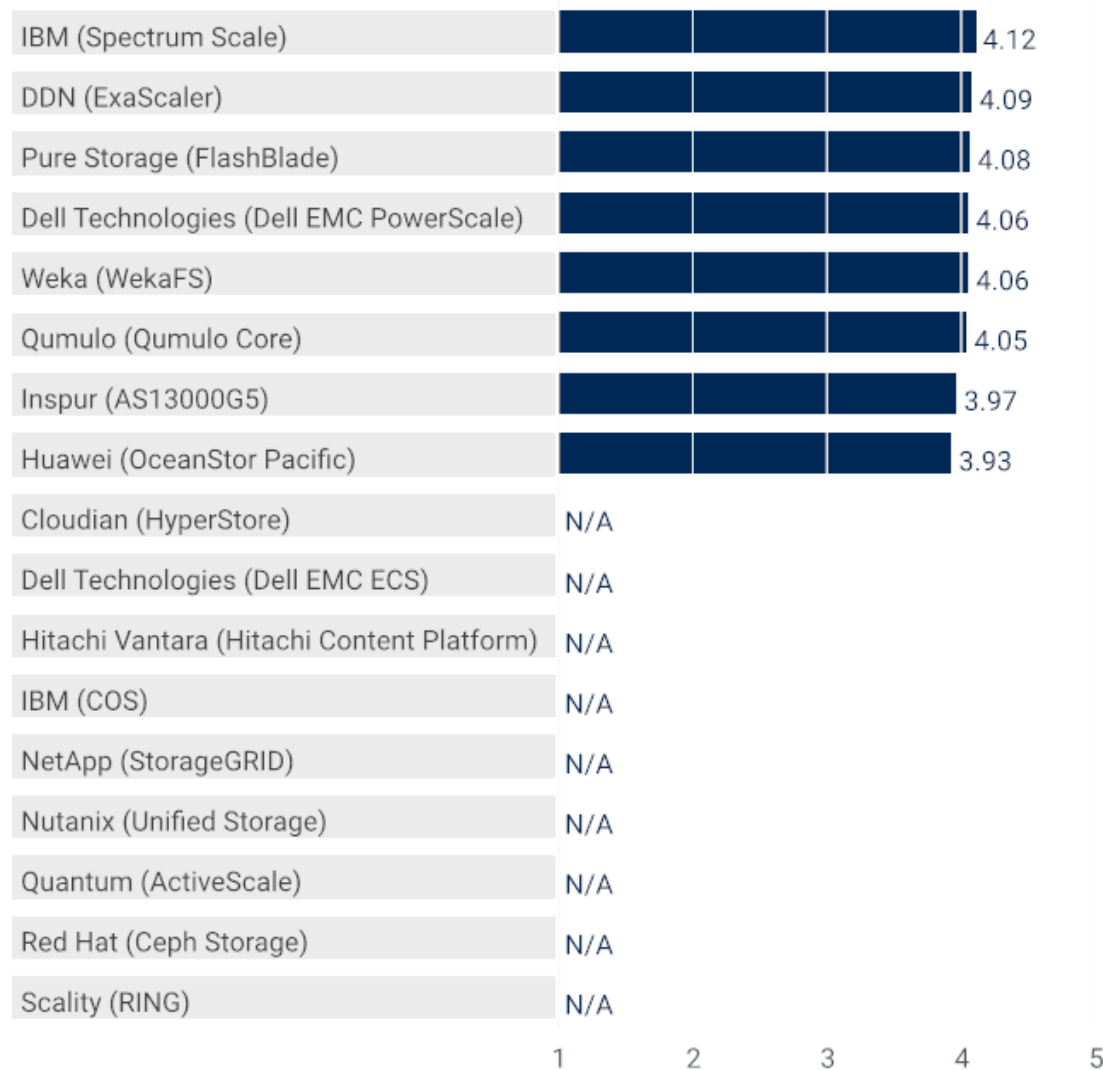
© Gartner, Inc

Gartner

Source: Gartner (October 2021)

## Vendors' Product Scores for Commercial HPC Use Case

Product or Service Scores for Commercial HPC



As of 24 September 2021

© Gartner, Inc

Gartner

Source: Gartner (October 2021)

## Vendors

### Cloudian (HyperStore)

Cloudian HyperStore is a native object storage platform in a scale-out architecture. HyperStore also supports NFS v.4 and SMB v.3 file services through an external HyperFile network-attached storage (NAS) gateway controller and operates on top of the scale-out object platform.

HyperStore deployment models are highly flexible. The platform can be deployed on commodity hardware from Cisco; Hewlett Packard Enterprise (HPE); Lenovo; Seagate; Quanta Cloud Technology (QCT) and Supermicro. It can be deployed as a virtual machine (VM) on-premises, in the cloud, on near-edge cloud solutions such as Amazon Web Services (AWS) Outpost, or delivered as a Cloudian-branded preintegrated appliance. HyperStore is implemented by deploying individual nodes into a logical solution that supports both a local or geodistributed implementation that can span across on-premises and the public cloud. HyperStore also supports dissimilar hardware configurations while supporting very-high-density nodes. It supports geodistributed erasure coding, as well as data efficiency features, such as deduplication and compression, while supporting all-flash configurations together with performance management capabilities like quality of service (QoS).

Over the last 12 months, Cloudian developed its dual node file-based architecture into a scale-out architecture, attained AWS Outposts Ready certification, integrated with VMware Cloud Foundation with Tanzu, added a Kubernetes S3 Operator and added all-flash support.

Cloudian still lacks a native, integrated file system capability that doesn't require the deployment of the HyperFile NAS controllers on top of the HyperStore object storage platform.

HyperStore scores well across all capabilities and use cases, but it scores particularly well for the archiving and backup use cases due to its high scores for value (cost, services and support). HyperStore scores well on other critical capabilities as it supports a broad ecosystem of server vendors, maintains high compatibility with the S3 API and has a well-designed dashboard.

## **DDN (ExaScaler)**

DDN's EXAScaler product is based on the open-source parallel file system Lustre. EXAScaler uses a hardened version of Lustre and adds proprietary features for data management. EXAScaler includes open-source access protocol support for both file- and object-based access. Lustre, a parallel file system, is most frequently accessed through a parallel file system client, as opposed to NFS, SMB or S3 protocols. DDN's EXAScaler is primarily sold as purpose-built appliances with support for both all-flash nonvolatile memory express (NVMe) and hybrid storage media. EXAScaler is available as a software-only offering that can run on standard x86 servers and is also available in AWS, Google Cloud Platform (GCP) and Microsoft Azure marketplaces.

DDN continues to be a niche product and positioned mainly for use cases where performance is valued above everything else. It does not address object workloads adequately and has a very low number of customers using its S3 interface. In the file workload area, DDN is not used as general-purpose file storage, nor for collaboration or project-related file shares.

DDN EXAScaler excels at applications that require extreme throughput combined with parallel client access. Its ability to share data directly with GPU-based servers, while an emerging use case, makes DDN among the top choices for analytics. EXAScaler's Hot Pools feature allows intelligent use of the flash tier by moving hot files from the HDD tier to flash, and moving cooling files from flash to the HDD tier.

In the last 12 months, DDN enhanced its protocol support and cloud deployment capabilities, including addition of a managed storage service for EXAScaler. Additional features delivered include native synchronization capabilities for disaster recovery and data distribution.

Top use cases for DDN are analytics and commercial HPC because of its superior performance powered by a parallel file system. DDN's EXAScaler is typically used only in four of the seven use cases evaluated in this research. DDN scored high for analytics and commercial HPC use cases.

## **Dell Technologies (Dell EMC ECS)**

Dell Technologies' Dell EMC Elastic Cloud Storage (ECS) is a scale-out object storage offering that can be deployed on-premises on purpose-built, Dell EMC PowerEdge-based or third-party qualified servers. With VMware Cloud Director support, third-party service providers can offer ECS as a service for VMware customers. ECS supports both object and file access through NFS, but it is typically accessed through object interfaces such as S3. It is sold across all verticals, with customers ranging from very large enterprises to small and midsize enterprises. ECS is frequently positioned as a long-term storage repository to complement PowerScale high-performance file storage.

While ECS does not have architectural limits on scaling the number of nodes or objects, the current maximum capacity of a single ECS EX-Series rack is 11.5 PB raw. ECS uses erasure coding for protection from disk and other hardware failures, but it supports only two configurations of 12+4 and 10+2 data and parity chunks. ECS supports in-line compression, but has no native support for deduplication. It also supports in-line encryption, but not self-encrypting drives. ECS lacks the native ability to migrate data between different classes of storage within ECS or to tier data from ECS to public clouds. However, Dell EMC DataIQ, an optionally available data management tool, provides this functionality at no incremental cost when managing Dell EMC storage. ECS cannot be deployed in any of the major public clouds. ECS supports the Amazon S3 protocol for object access, but it does not currently support S3 Select or bucket notifications.

ECS lacks support for native SMB-based file access and depends on the GeoDrive gateway application. ECS requires a minimum of five nodes and can be expanded in increments of one node, if the capacity of the new node is the same as the previous node.

ECS ranked lower overall than most other vendors' offerings in this research due to lack of feature set and Dell Technologies' emphasis on Dell EMC PowerScale.

## **Dell Technologies (Dell EMC PowerScale)**

Dell Technologies' Dell EMC PowerScale family consists of the PowerScale nodes, Dell EMC Isilon nodes and the PowerScale OneFS file system. PowerScale nodes are based on Dell EMC PowerEdge servers and are all-flash or all-NVMe with 100Gb or 25Gb connectivity and always-on in-line data efficiency. The Isilon nodes are based on purpose-built appliances designed to offer high density and come in all-flash (F series), hybrid (H series) or archive (A series) configurations. All the PowerScale and Isilon nodes can be part of a single cluster and managed by a single PowerScale OneFS distributed file system.

PowerScale, being powered by a file system as opposed to an object store, is best-suited for file-based workloads. PowerScale does support S3 interfaces to accommodate applications that need simultaneous access to data over both file and object protocols, but is typically not used for exclusive object-only cloud-native workloads.

OneFS is a proven product that has been in the market for more than a decade, with increasing functionality delivered over the years. OneFS is typically deployed as an appliance in a cluster configuration, but is also available as a virtual machine that can run on VMware ESXi and is offered as a service for GCP. OneFS is used in almost all industry verticals and across medium to large organizations. Its primary use is for applications that require high bandwidth for file workloads. For storage efficiency, OneFS supports both compression and deduplication in-line. OneFS has broad support for independent software vendor (ISV) applications across all major industry verticals, and has strong native NFS, SMB and HDFS protocol support.

OneFS is not available to run on bare-metal industry standard x86 servers other than Dell's own servers. Dell EMC DataIQ, an optional no-cost data management software for Dell EMC Storage, adds insights into data usage for data stored on PowerScale, ECS, and third-party storage and cloud, including the ability for users to move data on-demand. Dell EMC CloudIQ provides centralized enterprise-grade storage management of PowerScale and ECS.

Notable enhancements during the last 12 months include support for NVIDIA GPUDirect leveraging NFS over remote direct memory access (RDMA) protocol, support for S3 protocol access, in-line compression and in-line data deduplication across the full PowerScale family, and parallel upgrades enabling smaller upgrade windows. Most significantly, Dell launched managed file storage services powered by Dell EMC PowerScale appliances, through its APEX Data Storage Services.

Dell EMC PowerScale ranked high for analytics, HPC and backup use cases because of its high performance and all-flash support, as well as its broad ecosystem of software partners.

## **Hitachi Vantara (Hitachi Content Platform)**

Hitachi Vantara's Hitachi Content Platform (HCP) cloud object storage solution is a software-defined scale-out offering for unstructured workloads. HCP is available in a variety of deployment models including a fully integrated appliance, software only, in a VM, in a container, in the cloud, on-premises, or with third-party servers and applications. As an appliance, HCP is available with preconfigured access nodes (HCP G series nodes) and dense storage nodes (HCP S series nodes). HCP is supported by a portfolio of optional software products that provide data management, governance, analytics, collaboration, remote office/branch office (ROBO) and file gateway capabilities. The HCP portfolio includes Hitachi Content Intelligence (HCI) for data quality and analytics; Content Monitor, which combines AI/ML techniques to provide advanced monitoring and visualization; and HCP Anywhere for user collaboration and cloud home directories. The HCP object system supports NFS and SMB file access and the S3 protocol, and tiers data to public cloud object stores.

Hitachi expanded its HCP solution to Hitachi Content Software for File (HCSF) to include a high-performance, all-flash, NVMe-native parallel file system through an OEM partnership with Weka and its WekaFS. The integrated solution is suited for use with artificial intelligence (AI), machine learning (ML), analytics and other GPU-accelerated workloads. Hitachi Vantara's integration of file and object storage provides independent scale of compute and storage capacity with metadata management automation to move data between on-premises and public cloud storage for cost, compliance and business continuity purposes. Performance scales linearly with node additions.

HCP and HCSF offer a menu of perpetual licenses based on functional needs, including basic, premium and extended use cases. HCP for Virtual Machines is available with software-only pricing and a variety of pricing structures for HCP appliances and HCP software. A new utility-style subscription license is delivered as a managed service with Hitachi Vantara's EverFlex license model for customers deploying in the cloud or in bare-metal environments, who are charged per-terabyte of data stored for one month.

HCP scores higher for hybrid cloud storage and cloud-native application use cases than other use cases because of high scores for interoperability and manageability.



## **Huawei (OceanStor Pacific)**

Huawei OceanStor Pacific is a recent rebrand of Huawei OceanStor 100D. OceanStor Pacific is now positioned and developed as a unified multiprotocol storage platform. While we evaluate only object and distributed file systems capabilities in this research, OceanStor Pacific also supports block and HDFS interfaces. OceanStor Pacific is deployed on-premises in enterprise data centers, but it also serves as the underlying storage layer for Huawei's hyperconverged infrastructure FusionCube platform and Huawei's public cloud infrastructure as a service (IaaS) offerings.

The product is most frequently deployed as a preintegrated Huawei appliance. It is available with traditional and pay-as-you-grow pricing. Subscription-based pricing is currently only available in cloud deployments. OceanStor Pacific is designed for large-scale implementations and is popular in the service provider, government and financial sectors. Each cluster can start with three nodes and scale up to 4,096 nodes in the cluster. The OceanStor Pacific administrator interface is a single pane of glass to manage all storage interfaces. Gartner clients express high satisfaction with scalability and performance of the product for large implementations and with data efficiency capabilities that outperform its competitors. OceanStor Pacific file services are still nascent and don't have much traction for large-scale distributed file use cases where legacy OceanStor 9000 products are still being supported. OceanStor Pacific integrates with Huawei Cloud, but does not have integration with top global IaaS public providers.

Over the last 12 months, the product has benefited from major R&D investment and added a lot of features and capabilities such as high-density performance and capacity nodes, data compression, asynchronous replication, encryption, and multitenancy support. The product now has a parallel file system and a distributed parallel client support to address commercial HPC and analytics workloads.

Huawei OceanStor Pacific addresses the hybrid cloud storage, archiving, backup and cloud-native applications use cases. Huawei OceanStor Pacific performed solidly across all file and object storage use cases.

## IBM (COS)

IBM Cloud Object Storage (COS) is based on a scalable object store and is positioned primarily for object storage use-cases. It is available as purpose-built appliances, and as software that can run on standard x86 servers from any of major server vendors. It is the underlying storage for IBM Public Cloud Object Storage service and is also available as a single tenant managed storage service either in IBM public cloud or other dedicated infrastructure. IBM COS has been in the market for more than a decade and has some of the largest deployments of object storage by size outside of public cloud. IBM COS supports file-based access, but primarily for archive workloads, through an external gateway called IBM Cloud Object Storage File Access. The file access capability supports access through NFS and SMB protocols and is based on a technology from CTERA that IBM resells.

IBM lacks support for tiering within the COS cluster and tiering external to the COS cluster. IBM COS also does not support compression or deduplication, and depends on such capabilities from external applications. IBM COS cannot tier data into any public clouds. End users have noted that the IBM COS user interface is complex and not intuitive.

IBM COS's main strength is in its proven resiliency and ease of management at very large-scale deployments. Notable updates from the last 12 months include object tagging, support of 18 TB SMR drives, and enhancements to object life cycle management and performance through Zone Slice Storage.

IBM COS is the only product in this evaluation that is also the underlying storage for a major public cloud: IBM Cloud. Powered by an object store, COS is best-suited for object-native workloads. IBM COS's high scalability and resiliency, powered by its geodispersed erasure coding, strong S3 support, low cost and high capacities, led to its high ratings for the cloud IT operations, cloud-native applications and archiving use cases.

## IBM (Spectrum Scale)

IBM's Spectrum Scale is a mature parallel file system, previously known as General Parallel File System (GPFS), that is the underlying file system for many of the top commercial enterprises and supercomputers in the world. Spectrum Scale is available as a file system that can be deployed on bare-metal servers, containers and virtual machines, including as a VM in AWS. Spectrum Scale is also available as an appliance from IBM under the Elastic Storage System (ESS) brand. Spectrum Scale's biggest strength is its highly scalable throughput, delivering 2.5TB per second sequential and 2.2TB per second random read/write bandwidth in production. Spectrum Scale provides a native POSIX interface along with NFS, SMB, S3 and HDFS. Spectrum Scale supports file access through open-source components: Ganesha for NFS and SAMBA for SMB. Spectrum Scale also provides the ability to write in one access protocol and read in a different access protocol. Even though IBM Spectrum Scale supports S3 access, it is rarely used as object storage; rather, S3 support exists to support multiprotocol access for file-based data. IBM's Spectrum Scale, when compared to other parallel file systems, has more enterprise features and, when compared to other commercial distributed file systems, has more scalability and performance.

Spectrum Scale lacks support for large-scale Microsoft applications where advanced SMB protocol support is required. Another area where Spectrum Scale continues to be challenging is ease of use and lack of antivirus support. It also does not support storage efficiency features like deduplication, but does provide compression.

Updates during the last 12 months include Active File Management (AFM) to cloud storage, enabling Spectrum Scale to cache data into the file system from S3-compliant object stores. Other updates include support for CSI Snapshot; immutability and append-only capabilities to AFM-DR filesets; enhanced TCP/IP networks utilization; QoS for projects at the fileset-directory level; a Spectrum Scale operator for Red Hat OpenShift; and a containerized Spectrum Scale Client.

Given its biggest strength in performance, commercial HPC and analytics are the top use cases where Spectrum Scale is deployed, followed by the backup, archiving and other file use cases.

## Inspur (AS13000G5)

Inspur AS13000G5 is a unified distributed file system and object storage product that supports block, file and object protocols in one cluster, but is mostly deployed for file workloads. The AS13000G5 series is offered in four appliance models targeting HPC, surveillance, backup and archiving use cases. The AS13000G5 models provide support for nondisruptive online storage capacity expansion by adding drives to existing nodes or adding additional nodes to an existing cluster up to a maximum capacity of 5,160PB. The maximum number of nodes per cluster is 5,120 or 5EBs, but the maximum cluster in production today is 239 or 70PBs. The minimum cluster size is 3 nodes or 12TB.

Over the last 12 months, Inspur launched updates to the AG13000G5 products that include a wide variety of enhancements to performance, manageability, security and resilience, and new features including object tagging, native HDFS support and S3 select. Object storage supports in-line data compression and in-line deduplication, but neither are supported for file storage. Customers highlight the AS13000G5 all-in-one multiprotocol support for SMB 1.0/2.0/3.0, NFS V3/V4, FTP, HDFS, HTTP, SWIFT and S3 API.

The AS13000G5 supports erasure coding, remote copy, snapshot, clone, thin provisioning and other storage features, and it provides hierarchical data protection and disaster recovery services. Tiering policy promotes and demotes data across tiers, and balances each tier based on access frequency and read/write usage profile. An Intel NVMe Optane storage Tier 0 is offered for write-intensive applications, and a quad-level cell solid-state drive (QLC SSD) tier is offered for read-intensive applications. Inspur supports SMB v.3.0 transparent failover, and in an SMB cluster with two or more nodes, when a node fails, the client automatically reinitiates a connection.

Inspur's multidevice storage management product, InView 2.0, provides a single pane of glass for basic real-time monitoring, including the use of cloud-based AIOps functions for use in performance and capacity trend analysis to predict performance issues.

The AS13000G5 file storage system is missing support for in-line storage data compression or deduplication as well as for the hybrid cloud storage ability to tier data to major public clouds. It lacks support for multisite replication and for the ability to run as a VM or software-defined instance on any of the top major public clouds.

The AG13000G5 is used primarily in commercial HPC, hybrid cloud storage and backup use cases in telecom, finance and education industries. The product scores above average for all file storage use cases evaluated.

## **NetApp (StorageGRID)**

NetApp StorageGRID is a mature object storage solution that is based on an underlying object store. StorageGRID is available as an appliance, as a VM that can be deployed on industry-standard servers, and as a managed storage service under NetApp Keystone. StorageGRID appliances are available at multiple capacity and performance points per node and can be deployed alongside software-only deployments in the same cluster. StorageGRID is primarily targeted at object storage workloads. StorageGRID does not have native file interface support. For file workloads, StorageGRID depends on NetApp FabricPool or third-party file gateways.

Where StorageGRID lags some of the leaders in the market is the absence of native file support to handle file heavy workloads. StorageGRID can tier data into AWS or Azure, but not into GCP.

StorageGRID provides a lot of flexibility in deployment choices to address both cost- and performance-centric use cases. It has good policy-based management of data placement and life cycle management. StorageGRID is the most popular storage tier behind NetApp's primary AFF and FAS storage due to easy policy-based tiering using NetApp FabricPool.

Over the last 12 months, StorageGRID added enhancements for AWS S3-compatible Object Lock, increased scale to support 2 billion objects per node, support for 18TB drives and fully managed storage as a service via NetApp Keystone in partnership with Equinix. Additional updates include enhancements to Key Management Interoperability Protocol (KMIP) with software-based node encryption and support for external key managers.

NetApp StorageGRID supports six of the seven use cases evaluated in this research. The only use-case StorageGRID does not support is the HPC use-case. StorageGRID ranked higher in Cloud Native Applications, Cloud IT Operations, closely followed by the remaining use-cases it supports. StorageGRID stands out for cloud-native applications and cloud IT operations because of its strong S3 API support and manageability capabilities.

## **Nutanix (Unified Storage)**

Nutanix provides scale-out distributed file (Nutanix Files) and object storage (Nutanix Objects) services supporting SMB, NFS and S3 access protocols. It is fully integrated with the Nutanix hyperconverged platform for a single architecture and management experience. Nutanix distributed storage solutions can be deployed across nodes, leveraging an existing hyperconverged infrastructure cluster infrastructure for smaller deployments, or on a dedicated-storage-only cluster for larger scale and better cost-efficiency.

Nutanix Files is widely used to consolidate legacy file storage systems across sites to provide a single architecture and a single point of management. It is used particularly for department shares, user profile, home directory data, imaging and video repository, and storage server consolidation. Nutanix Objects are often deployed for tiering big data applications, running cloud native applications and for large datasets in backup and archiving workloads.

Nutanix Unified Storage solutions leverage a subscription licensing model with terms from one to five years and a capacity-based metric based on used capacity at the TiB level. Nutanix Files and Nutanix Objects offer licensing for dedicated clusters or for mixed mode clusters, where the Nutanix Files and/or Objects storage system runs alongside the Nutanix hyperconverged infrastructure environment. Nutanix built a robust OEM route and its software-defined storage can be deployed on a wide variety of server platforms, including Dell, Fujitsu, HPE and Lenovo.

Over the last year Nutanix enhanced unified data management for all storage workloads on prem, at the edge or in the cloud, and added its deployment in AWS as part of Nutanix Clusters. Nutanix now has a data insight analytics engine with capabilities for data governance including data life cycle management, data compliance using audit trails, and data security by monitoring and detecting usage anomalies. Nutanix is investing in data protection and security capabilities and has recently introduced integrated ransomware protection for Nutanix Files.

Nutanix is a relative newcomer to the unstructured data market and does not address performance-sensitive workloads. Nutanix does not currently have support for file deduplication or tiering to the public cloud. While Files and Objects popularity is growing with existing Nutanix customers, its deployments tend to be smaller than the leaders in this market.

Nutanix Files and Objects is best suited for hybrid cloud storage, application data and analytics workloads.

## **Pure Storage (FlashBlade)**

Pure Storage FlashBlade is an all-flash array distributed file system that offers unified fast file and object store services in the same system for unstructured data applications. The majority of FlashBlade deployments are primarily file or object. FlashBlade was upgraded to Purity OS 3.2 with a wide range of enhanced features and support options, including native scale-out SMB v.2.1. FlashBlade is also available as an on-demand as-a-service offering and as a fully managed appliance that includes all hardware and software preintegrated to deliver simplified deployment and administration for performance-intensive workloads.

The FlashBlade product is based on a scale-out blade design that automatically distributes the metadata and data across the different blades in the system. Customers can nondisruptively scale capacity by adding additional blades to a FlashBlade chassis or by expanding to a multichassis configuration, without downtime. Each blade contains custom flash technology (DirectFlash) and storage processors. The DirectFlash technology is designed to fully exploit the parallelism of TLC NAND flash to eliminate the SSD drive controller as a bottleneck.

FlashBlade configuration starts with seven blades, or a minimum of 65TBs usable, and supports up to a maximum of 150 blades or 10 chassis with a raw capacity of 7.92PB in a single file or object namespace. Always-on compression yields an average data reduction of 1.4:1. The FlashBlade system internally optimizes and automatically chooses the erasure coding support, accounting for factors including the number of blades in the system and the distribution of blades across hardware chassis.

FlashBlade does not support either deduplication or native tiering to an external system. It has no meaningful QoS capabilities and lacks SMB v.3.0 support to provide continuous availability of SMB file shares. FlashBlade can't be deployed as software-defined instances on-premises or in the public cloud. However, starting with FB Purity 3.0, object data can be replicated to AWS. Pure has integrated the distributed Portworx control plane software as an orchestration and/or virtualization layer to the FlashBlade to address scaling issues within the container environment without going through a single CSI plug-in.

The FlashBlade Evergreen premium and ultrasupport services provide guaranteed, nondisruptive elastic scaling of capacity and performance to ensure smooth transitions across platform generations. Rapid Restore leverages FlashBlade parallelism to back up and restore up to 270 TB/hour to address tight RTO windows. FlashRecover, in joint partnership with Cohesity Data Protect software, provides an integrated backup solution for large-scale recovery of VMs. FlashBlade's SafeMode snapshots provide immutability to protect data backups from ransomware attacks.

FlashBlade is used in — and scored solidly in — four use cases, including analytics, commercial HPC, backup and cloud IT operations. Its best showing was in analytics.

### **Quantum (ActiveScale)**

Quantum ActiveScale is a native object storage platform in a scale-out architecture. ActiveScale also supports NFS v.3, which is embedded into the architecture and supports separate volumes for file storage.

ActiveScale deployments are delivered as preintegrated appliances or as software on top of SuperMicro hardware. ActiveScale comes in three models: the low-capacity P100E3 that scales to 1PB, the midcapacity P100 that scales to 27PB and the high-capacity X200 that can scale to unlimited capacity. ActiveScale is implemented by deploying individual nodes into a clustered solution that supports both a local or maximum three-site implementation with geodistributed erasure coding. ActiveScale is S3-compatible, supports S3 versioning and object locking, and supports configurable erasure coding and replication to the public cloud.

Over the last 12 months, Quantum announced the smaller p100E3 appliance and the larger X200, which both offer unlimited capacity. It also released version 6.0 of its ActiveScale software, which offers increased capacity and performance. Quantum started its transition to offer a software-defined solution, and entered an agreement with Super Micro to deliver broader hardware options. Quantum also added subscription-based pricing to offer more pricing flexibility.

ActiveScale lacks support for file access through the SMB interface. ActiveScale scored low for storage efficiency, as it is missing features like deduplication, compression and data tiering. However, it can also improve on other capabilities, such as QoS and all-flash support. ActiveScale deployment scenarios have improved, but are still less flexible, as it cannot be deployed on servers of choice or within public cloud providers.



ActiveScale scores well for manageability, as it has a well-designed dashboard with useful at-a-glance metrics, such as the percentage of capacity used, data durability state and system performance characteristics. ActiveScale also scores well for value and for interoperability and capacity. Quantum ActiveScale scores average across the analytics, archiving, backup and hybrid cloud storage use cases. This is mostly due to a more limited feature set compared to other vendors in this research. ActiveScale is mostly deployed for the backup and archiving use case.

## **Qumulo (Qumulo Core)**

Qumulo's file data platform, Qumulo Core, is a software-defined distributed scale-out file system for managing unstructured data. Qumulo Core is available in two versions as Qumulo Server Q and Cloud Q. Qumulo Server Q is optimized for on-premises and preinstalled as an appliance through hardware vendors such as Arrow, Fujitsu and HPE. Cloud Q is available in Azure as a managed file service, achieving "well-architected" designation in the AWS and GCP marketplaces in a variety of instance types and configurable options. Qumulo Clusters can scale to petabytes of storage in a single namespace, with no enforced upper limit in the number of nodes.

Over the past 12 months, Qumulo has added a number of new enhancements and capabilities, including version 4.1 to support adding mixed nodes with different platform versions from the same hardware vendor or multiple vendors. Dynamic Scale, which allows expanding clusters with mixed node types, is an all-flash Server Q version for HPE. Other options include an all-NVMe Server Q Arrow P-368 and a new NVMe-hybrid line of Server Q Arrow C-432T and C-192T with up to double the write performance of SATA-based hybrid systems. Qumulo also offers Cloud Q support for AWS Outposts and a new native fully managed service on Azure. The product is still missing some enterprise features, such as NFS v.4 support (in beta), object protocols for file services (such as S3), compression, deduplication, write once read many (WORM), container support, and seamless, nondisruptive node removal.

In support of its primary media and entertainment vertical industry clients, Qumulo offers Studio Q as a cloud-based remote editing environment for video content editing. Qumulo offers a variety of data services for its file system. Qumulo Aware features monitoring with real-time analytics. Qumulo Integrate automates workflows. Qumulo Shift tiers data to AWS S3 and Glacier. Qumulo Protect offers backup and DR, and protects against external threats. Qumulo offers transparent, all-inclusive subscription pricing based on raw capacity.

Qumulo is broadly diversified in client use cases, but is best suited for unstructured data-intensive industries such as media and entertainment, life and earth sciences, telecommunications, automotive, HPC, manufacturing, oil and gas, and higher education. Qumulo Core received excellent scores in six of the seven use cases, with an emphasis on hybrid cloud storage and cloud IT operations.

## **Red Hat (Ceph Storage)**

Red Hat is the primary developer behind Ceph Storage, an open-source storage solution that delivers block, file and object storage capabilities. However, in this research, it is evaluated only on its object and file storage capabilities. Ceph's storage software development continuously benefits from the contribution of an active open-source community, with Red Hat as leading contributor. Ceph supports unlimited nodes and capacity, and it automatically rebalances and scales when adding additional servers to the pool. While Ceph's file system supports access over NFS protocol, it is not usually used in production deployments and lacks some enterprise features. Ceph does not support access over SMB protocol.

Ceph is sold as a software-only storage product and has reference architectures with OEM and original design manufacturer (ODM) server hardware vendors. Ceph receives significant R&D focus because Red Hat's container storage (OpenShift Data Foundation) is based on Ceph technology. Ceph is a highly tunable and scalable product, but performance depends on hardware selection, and thus the product does not provide guaranteed performance. Ceph is the best fit for large organizations where storage administrators are proficient with Linux system support and scripting.

Ceph has fewer ISV qualifications for backup, archiving, surveillance and gateway products compared with its competition. It is also missing some enterprise capabilities, such as deduplication and WORM. Some Ceph customers still cite a steep learning curve when onboarding the product. Ceph does not have much public cloud integration outside of Red Hat OpenShift. Over the last 12 months, Ceph introduced an object lock feature, added FIPS 140-2 cryptography support and expanded key management integration. It has also increased performance for block and object workloads, added NFS file system support, and improved efficiency with reduced resource consumption for small files and objects.

Red Hat Ceph is typically not deployed as a stand-alone storage product, but is often deployed as part of the Red Hat infrastructure portfolio, which includes Red Hat Enterprise Linux (RHEL), OpenShift and OpenStack. The primary use cases for Ceph are hybrid cloud storage, cloud-native applications, backup and analytics.

## Scality (RING)

Scality RING is a deeply integrated distributed scale-out object and file system solution that shares a common data pool between object and file system access. The Scality file system supports distributed NFS v.4 and SMB v.3 with backward compatibility.

Scality RING can be deployed on x86 commodity hardware in a local or geodistributed implementation. Scality goes to market with several hardware partners, including through an OEM arrangement with HPE. Scality is also on the price list for reselling with Cisco and Supermicro, and it meets in the channel with GIGABYTE, Lenovo and Western Digital.

Scality RING is scalable and S3-compatible, and is feature-rich with unique data-sharing capabilities. Scality RING can distribute files and objects across multiple sites and public cloud providers while providing uniform access to every file or object from every location by separating data and metadata across all sites.

Over the last 12 months, Scality has delivered NFS v.4 and SMB v.3, optimized its file system for shared access, and improved management and reporting. It also added object lock support with several other enhancements to improve its overall feature set, and added support for flash storage, including all NVMe support. Scality released version 8 of its RING software, which added several improvements like enhanced performance, improved tiering, SEC Rule 17a-4 certification and a new management interface for cloud management. Scality also delivered ARTESCA, a lightweight cloud-native object storage solution with a smaller capacity entry point.

Scality RING scored on the low end for storage efficiency, as it is lacking support for compression and deduplication. In addition, RING's architecture is a lesser fit for smaller-sized workloads. Scality could improve its overall ease of deployment and operations.

RING scores high on its resilience capabilities, as it delivers a 100% guaranteed availability program with compensation, broad data replication and availability capabilities such as stretch cluster implementations and/or synchronous replication. RING also scores high across the analytics, archiving, backup and hybrid cloud storage use cases, where it is often deployed for large datasets in backup and archiving, video and content distribution, hosted email, digital and medical imaging, and HPC archiving solutions.

## **Weka (WekaFS)**

Weka makes a parallel and distributed file system software product, WekaFS. First released in 2017, the software-only product was designed for I/O-intensive, low-latency file workloads. The software leverages industry-standard server infrastructure and NVMe SSD storage. The storage is virtualized into a logical pool and presented as a unified namespace file system to the host applications. The fully POSIX-compliant virtual file system supports the NAS NFS and SMB protocols. It can be deployed on dedicated storage servers (appliance model), in a converged solution (with compute and storage integrated in the same server) or natively in the cloud.

WekaFS is available on server platforms from major server OEMs as well as bundled solutions from Hitachi Vantara, HPE, Penguin Computing and Supermicro. WekaFS is designed for both high throughput and low latency workloads and optimized for performance by leveraging NVMe protocols and enterprise grade flash media. Weka delivers high performance through its journaled file system architecture that distributes metadata across the nodes and through accelerated writes to SSDs that are directly attached to the host. Weka namespace can be extended to take advantage of cost-effective and scalable S3 object storage on prem or in the public cloud. The cloud-based management system helps with system installation/configuration, day-to-day administration and analytics on performance (such as read/write IOPS, latency and throughput), system utilization and capacity planning.

Over the last 12 months, Weka released support for S3 API, NVIDIA GPUDirect Storage (GDS), autoscaling and tiering to multiple object storage buckets. WekaFS is not positioned as an object storage, but as a file system with the ability to communicate via an S3 object storage API. It does not have data compression or deduplication, and is typically not deployed as a general-purpose file system.

Weka is focusing on high-performance and low-latency use cases such as analytics, commercial HPC and cloud-native applications. Weka scored high in analytics and hybrid cloud storage.

## Context

Unstructured data storage products provide a set of foundation capabilities irrespective of the type of data that is stored in them: file or objects. Foundational capabilities include scalability, performance, resilience, availability and manageability. Distributed file systems, as the name suggests, are based on a file system with hierarchical structures, whereas object storage systems are based on a flat namespace consisting of key-value pairs. Where the value is the actual object, and the key is the metadata associated with that object. Distributed file systems have a long history, starting in high-performance computing for scientific research and now in commercial enterprises for various design, simulation and analysis use cases. Object storage, on the other hand, became a common form of storage only in the last decade due to the success of public cloud and the use cases around it.

Most of the enterprise application requirements can be satisfied with either file or object storage products; a true file system or object storage becomes a key requirement only in extreme cases. When data needs to be manipulated frequently, or when an application depends on the file path name, a file system is better suited. But for extreme scaling, as in several billions of files, the overhead of a file system causes inefficiencies, and an object storage is a better fit. Outside of these extremes, a common platform helps consolidate all unstructured data use cases into a single platform.

Today's products are increasingly catering to the issue of data management, and not just storing and protecting data. Data management broadly falls into three buckets. The first bucket is operational efficiency and other capabilities to simplify the job of the infrastructure administrator. The second is governance and security-related capabilities to simplify the job of security and regulatory teams. And the last bucket is about enabling data insights from the underlying data through the use of various analytics tools.

## Product/Service Class Definition

Object storage refers to storage hardware and software infrastructure that house data in structures called "objects" and serve hosts via protocols (such as HTTP) and APIs (such as Amazon S3). Conceptually, objects are similar to files in that they are composed of content and metadata. In general, objects support richer metadata than file storage by enabling users or applications to assign attributes to objects that can be used for administrative purposes, data mining and information management.

A distributed file system clusters multiple storage nodes together, presenting a single namespace and storage pool to provide high bandwidth for multiple hosts in parallel. Data is distributed over multiple nodes in the cluster to handle availability and data protection in a self-healing manner and to provide high throughput and linear capacity expansion.

The products in this category must have “distributed architecture.” This refers to a distributed computing architecture in which each node is independent and self-sufficient, and there is no single point of contention across the system. More specifically, none of the nodes share memory or disk storage. People typically contrast distributed architecture with systems that keep a large amount of centrally stored state information, whether in a database, an application server or any other similar single point of contention.

## Critical Capabilities Definition

Enterprises should consider the following eight critical capabilities when deploying unstructured data storage platforms. Enterprises can work toward these goals by evaluating object storage products in all capability areas.

### **Storage Efficiency**

A product’s ability to support storage efficiency technologies, such as compression, single-instance storage/deduplication, tiering and configurable erasure coding to reduce total cost of ownership (TCO).

### **Security and Multitenancy**

The native security features embedded in the platform that provide granular access control, enable enterprises to encrypt information, provide robust multitenancy and QoS, offer data immutability, and ensure compliance with regulatory requirements.

### **Capacity**

The ability of the product to support growth in capacity in a nearly linear manner. We looked at object storage capacity scalability limitations in theoretical and real-world configurations, such as maximum theoretical capacity, object size and production deployment.

### **Interoperability**

The ability of the product to support third-party ISV applications across horizontal and vertical industry-specific use cases, public cloud APIs, and various deployment models.

## **Manageability**

The automation, management, monitoring and reporting tools and programs supported by the product. In addition, ease of setup and configuration, as well as metadata management capabilities, are considered.

These tools and programs can include single-pane management consoles, monitoring systems and reporting tools. They are designed to help personnel seamlessly manage systems, monitor system usage and efficiencies, and anticipate and correct system alarms and fault conditions before or soon after they occur.

## **Performance**

The per-node and aggregated throughput for reads and writes that can be delivered by the cluster in real-world configurations. It also includes support for flash for metadata and data for latency-sensitive applications.

## **Resilience**

The platform capabilities for providing high system availability and uptime. Options include high tolerance for simultaneous disk and/or node failures, fault isolation techniques, built-in protection against data corruption, and data protection techniques such as erasure coding and replication.

Features are designed to meet users' recovery point objectives (RPOs) and recovery time objectives (RTOs). There are several methods for data protection in today's object storage products. Redundant array of independent disks (RAID) is becoming less popular due to huge capacity overheads and long rebuild times.

The simplest way to protect data is replication, which stores multiple copies of the data locally or in a distributed manner. A more innovative data protection scheme is erasure coding, which breaks up data into "n" fragments and "m" additional fragments across n+m nodes, offering clients configurable choices depending on their cost and data protection requirements. Enterprises often combine erasure coding and replication because the former performs well with large files, whereas the latter works well with large numbers of small files. WAN costs and performance considerations in distributed environments are also factors.

## **Value**

This is the price of the product relative to the capabilities an enterprise stands to experience.

## Use Cases

### **Analytics**

This applies to storage consumed by big data analytics applications and packaged business intelligence (BI) applications for domain or business problems.

### **Archiving**

This applies to a product's ability to provide a cost-effective, scalable and resilient long-term data storage repository.

### **Backup**

The storage system acts as the destination for backup data. Ingest and restore performance, security, storage efficiency, and value are treated as critical capabilities for this use case.

This use case is not about the backup capabilities of the storage system.

### **Hybrid Cloud Storage**

In this use case, enterprises leverage public cloud storage to facilitate a hybrid cloud workflow between on-premises and public cloud IaaS platforms.

### **Cloud-Native Applications**

Applications designed specifically to run in a cloud environment where storage is provisioned, created, consumed and deleted all through object interfaces.

In this use case, the storage system provides the underlying storage for these applications.

### **Cloud IT Operations**

In this use case, the storage system is used to provide storage as a service through cloudlike consumption, delivery and management platforms.

### **Commercial HPC**

In this use case, an enterprise uses distributed file system storage to provide high throughput and parallel read-and-write access to large volumes of data.



## Vendors Added and Dropped

### Added

- Nutanix
- Weka

### Dropped

- Caringo (Acquired by DataCore, which does not meet inclusion criteria for revenue)

## Inclusion Criteria

To qualify for inclusion, vendors must meet *all* of the following requirements:

- Above \$10 million of recognized product revenue over the last four quarters (as of May 2021) for the distributed file systems and/or object storage solutions between 1 May 2020 and 30 April 2021. In addition, the vendor must have at least 75 production customers each consuming more than 500TB raw capacity through either distributed file or object storage protocols only. The product must be installed in at least three (out of four) major geographies, with a minimum of 20 production customers brought to revenue in at least three of the four geographies (North America, EMEA, Asia/Pacific and South America).
- The product should be deployed across at least five out of the seven use cases
- The product must be designed for primarily on-premises workloads and not as a pass-through solution where data will be permanently stored elsewhere.
- Products should not be offered exclusively as a service.
- The vendor should own the storage software intellectual property and be a product developer. If a product is built on top of open-source software, the vendor must be one of the top 10 active contributors to the community (in terms of code contribution over the past 12 months).
- The vendor must have a product including features and capabilities generally available before 5 April 2021 that meet the packaging and product capability criteria listed below.

### Packaging:

- The product must be sold as either an appliance or software-based storage solution.
- The product must be available for purchase and consumed as a stand-alone file-and/or object-storage-only product and not as part of an integrated, converged or hyperconverged system with a compute and hypervisor bundle.

## Product capabilities:

- The product must have file and/or object access to the common namespace/file system.
- The product must have a *fully distributed architecture* where data and metadata are distributed, replicated or erasure-coded over the network across multiple nodes in the cluster. The product must have the ability to handle disk, enclosure or node failures in a graceful manner without impacting availability.
- The product must be a single file system capable of expanding beyond 500TB.
- The product must have a global namespace capable of 2PB expansion.
- The cluster must span more than four nodes.
- The product must offer support for horizontal scaling of capacity and throughput in a cluster mode or in independent node additions with a global namespace/file system.

**Table 1: Weighting for Critical Capabilities in Use Cases**

(Enlarged table in Appendix)

<i>Critical Capabilities</i>	<i>Analytics</i>	<i>Archiving</i>	<i>Backup</i>	<i>Hybrid Cloud Storage</i>	<i>Cloud-Native Applications</i>	<i>Cloud IT Operations</i>	<i>Commercial HPC</i>
Storage Efficiency	10%	12%	12%	10%	10%	10%	10%
Security and Multitenancy	12%	10%	15%	15%	14%	12%	10%
Capacity	11%	14%	12%	11%	12%	10%	12%
Interoperability	13%	11%	11%	13%	17%	13%	10%
Manageability	12%	11%	10%	20%	15%	20%	10%
Performance	20%	8%	14%	11%	10%	10%	23%
Resilience	10%	12%	10%	10%	12%	15%	10%
Value	12%	22%	16%	10%	10%	10%	15%
As of 24 September 2021							

Source: Gartner (October 2021)

This methodology requires analysts to identify the critical capabilities for a class of products/services. Each capability is then weighed in terms of its relative importance for specific product/service use cases.

## Critical Capabilities Rating

Each of the products/services that meet our inclusion criteria has been evaluated on the critical capabilities on a scale from 1.0 to 5.0.

**Table 2: Product/Service Rating on Critical Capabilities**

(Enlarged table in Appendix)

<b>Critical Capabilities</b>	<b>Quantum (ActiveScale)</b>	<b>Inspur (AS13000G5)</b>	<b>Red Hat (Ceph Storage)</b>	<b>Cloudian (HyperStore)</b>	<b>DDN (ExaScaler)</b>	<b>Dell Technologies (Dell EMC ECS)</b>	<b>Dell Technologies (Dell EMC PowerScale)</b>	<b>Pure Storage (FlashBlade)</b>	<b>Hitachi Vantara (Hitachi Content Platform)</b>	<b>IBM (COS)</b>	<b>IBM (Spectrum Scale)</b>	<b>Nutanix (Unified Storage)</b>	<b>Huawei (OceanStor Pacific)</b>	<b>Qumulo (Qumulo Core)</b>	<b>Scality (RING)</b>	<b>NetApp (StorageGRID)</b>	<b>Weka (Weka FS)</b>
Storage Efficiency	2.0	4.0	4.0	4.2	2.5	3.0	4.5	3.5	4.0	2.5	3.7	3.0	4.5	3.5	3.0	4.0	3.0
Security and Multitenancy	3.4	3.9	3.9	3.9	3.5	3.6	4.1	3.5	4.3	3.4	3.4	3.0	3.9	3.8	4.0	4.0	3.6
Capacity	3.5	4.0	4.0	4.0	4.5	3.5	4.0	3.8	4.0	4.0	4.5	3.5	3.5	3.5	4.5	3.8	4.5
Interoperability	3.1	3.1	4.3	4.0	4.2	4.3	4.1	3.8	4.4	4.5	3.4	4.2	4.1	3.8	4.2	4.2	3.8
Manageability	3.8	3.8	4.2	4.0	4.2	4.2	3.6	4.2	3.6	4.3	4.0	4.0	3.9	4.7	4.2	4.0	4.0
Performance	3.5	4.4	3.0	3.8	4.8	3.0	4.5	5.0	4.0	4.0	4.9	3.8	3.5	4.5	4.0	3.6	4.7
Resilience	3.4	4.3	3.5	3.5	3.5	3.3	3.1	4.5	3.3	4.8	4.0	4.0	4.2	4.0	4.1	4.0	4.0
Value	4.3	3.8	3.5	4.7	4.4	4.5	4.0	3.5	4.2	4.5	4.0	3.8	4.3	4.1	4.2	4.3	4.0
As of 24 September 2021																	

Source: Gartner (October 2021)

Table 3 shows the product/service scores for each use case. The scores, which are generated by multiplying the use-case weightings by the product/service ratings, summarize how well the critical capabilities are met for each use case.

**Table 3: Product Score in Use Cases**

(Enlarged table in Appendix)

Use Cases	Quantum (ActiveScale)	Inspur (AS13000G5)	Red Hat (Ceph Storage)	Cloudian (HyperStore)	DDN (ExaScaler)	Dell Technologies (Dell EMC ECS)	Dell Technologies (Dell EMC PowerScale)	Pure Storage (FlashBlade)	Hitachi Vantara (Hitachi Content Platform)	IBM (COS)	IBM (Spectrum Scale)	Nutanix (Unified Storage)	Huawei (OceanStor Pacific)	Qumulo (Qumulo Core)	Scality (RING)	NetApp (StorageGRID)	Weka (WekaFS)
Analytics	3.41	3.93	3.74	4.00	4.05	3.65	4.04	4.06	3.99	4.02	4.06	3.69	3.94	4.04	4.04	3.96	4.02
Archiving	3.46	3.89	3.80	4.09	3.98	3.77	3.97	-1.00	3.99	4.05	3.98	3.67	4.03	3.97	4.05	4.03	-1.00
Backup	3.41	3.92	3.77	4.04	3.97	3.68	4.03	3.95	4.01	3.98	3.99	3.63	3.98	3.98	4.02	3.99	-1.00
Hybrid Cloud Storage	3.41	3.88	3.85	4.00	-1.00	3.73	3.96	-1.00	3.97	4.02	3.96	3.68	3.97	4.04	4.05	3.99	3.95
Cloud-Native Applications	-1.00	3.86	3.86	4.00	-1.00	3.73	-1.00	-1.00	3.99	4.04	-1.00	3.69	3.98	-1.00	4.06	4.00	-1.00
Cloud IT Operations	3.41	3.90	-1.00	-1.00	-1.00	-1.00	3.91	4.01	3.92	4.07	-1.00	-1.00	3.99	4.05	-1.00	4.00	-1.00
Commercial HPC	-1.00	3.97	-1.00	-1.00	4.09	-1.00	4.06	4.08	-1.00	-1.00	4.12	-1.00	3.93	4.05	-1.00	-1.00	4.06
As of 24 September 2021																	

Source: Gartner (October 2021)

To determine an overall score for each product/service in the use cases, multiply the ratings in Table 2 by the weightings shown in Table 1.

## Evidence

The critical capabilities scoring is primarily supported by the comprehensive responses from vendors to Gartner's Magic Quadrant survey on this topic. The weights associated with each critical capability for the use cases evaluated are based on the analyst team's understanding of the use cases, which is influenced by inquiries and one-on-one meetings with Gartner clients regarding object storage and distributed file system solutions, conducted since the publication of the last iteration of this Critical Capabilities research.

## Critical Capabilities Methodology

This methodology requires analysts to identify the critical capabilities for a class of products or services. Each capability is then weighted in terms of its relative importance for specific product or service use cases. Next, products/services are rated in terms of how well they achieve each of the critical capabilities. A score that summarizes how well they meet the critical capabilities for each use case is then calculated for each product/service.

"Critical capabilities" are attributes that differentiate products/services in a class in terms of their quality and performance. Gartner recommends that users consider the set of critical capabilities as some of the most important criteria for acquisition decisions.

In defining the product/service category for evaluation, the analyst first identifies the leading uses for the products/services in this market. What needs are end-users looking to fulfill, when considering products/services in this market? Use cases should match common client deployment scenarios. These distinct client scenarios define the Use Cases.

The analyst then identifies the critical capabilities. These capabilities are generalized groups of features commonly required by this class of products/services. Each capability is assigned a level of importance in fulfilling that particular need; some sets of features are more important than others, depending on the use case being evaluated.

Each vendor's product or service is evaluated in terms of how well it delivers each capability, on a five-point scale. These ratings are displayed side-by-side for all vendors, allowing easy comparisons between the different sets of features.

Ratings and summary scores range from 1.0 to 5.0:

1 = Poor or Absent: most or all defined requirements for a capability are not achieved

2 = Fair: some requirements are not achieved

3 = Good: meets requirements

4 = Excellent: meets or exceeds some requirements

5 = Outstanding: significantly exceeds requirements

To determine an overall score for each product in the use cases, the product ratings are multiplied by the weightings to come up with the product score in use cases.

The critical capabilities Gartner has selected do not represent all capabilities for any product; therefore, may not represent those most important for a specific use situation or business objective. Clients should use a critical capabilities analysis as one of several sources of input about a product before making a product/service decision.

## Document Revision History

[Critical Capabilities for Object Storage - 21 October 2020](#)

[Critical Capabilities for Object Storage - 25 November 2019](#)

[Critical Capabilities for Object Storage - 30 January 2019](#)

[Critical Capabilities for Object Storage - 25 January 2018](#)

[Critical Capabilities for Object Storage - 31 March 2016](#)

[Critical Capabilities for Object Storage - 11 February 2014](#)

---

## Recommended by the Authors

Some documents may not be available as part of your current Gartner subscription.

[How Products and Services Are Evaluated in Gartner Critical Capabilities](#)

[Magic Quadrant for Distributed File Systems and Object Storage](#)

[Product Manager Insight: Top 5 Trends in Unstructured Data Storage, 2021](#)

[Market Trends: External Object Storage Offerings Are Ready to Replace Additional File System Cases](#)

[Hype Cycle for Storage and Data Protection Technologies, 2021](#)

© 2021 Gartner, Inc. and/or its affiliates. All rights reserved. Gartner is a registered trademark of Gartner, Inc. and its affiliates. This publication may not be reproduced or distributed in any form without Gartner's prior written permission. It consists of the opinions of Gartner's research organization, which should not be construed as statements of fact. While the information contained in this publication has been obtained from sources believed to be reliable, Gartner disclaims all warranties as to the accuracy, completeness or adequacy of such information. Although Gartner research may address legal and financial issues, Gartner does not provide legal or investment advice and its research should not be construed or used as such. Your access and use of this publication are governed by [Gartner's Usage Policy](#). Gartner prides itself on its reputation for independence and objectivity. Its research is produced independently by its research organization without input or influence from any third party. For further information, see "[Guiding Principles on Independence and Objectivity](#)."



Table 1: Weighting for Critical Capabilities in Use Cases

<b>Critical Capabilities</b> ↓	<b>Analytics</b> ↓	<b>Archiving</b> ↓	<b>Backup</b> ↓	<b>Hybrid Cloud Storage</b>	↓	<b>Cloud-Native Applications</b> ↓	<b>Cloud IT Operations</b> ↓	<b>Commercial HPC</b> ↓
Storage Efficiency	10%	12%	12%	10%		10%	10%	10%
Security and Multitenancy	12%	10%	15%	15%		14%	12%	10%
Capacity	11%	14%	12%	11%		12%	10%	12%
Interoperability	13%	11%	11%	13%		17%	13%	10%
Manageability	12%	11%	10%	20%		15%	20%	10%
Performance	20%	8%	14%	11%		10%	10%	23%
Resilience	10%	12%	10%	10%		12%	15%	10%
Value	12%	22%	16%	10%		10%	10%	15%
As of 24 September 2021								

Source: Gartner (October 2021)

Table 2: Product/Service Rating on Critical Capabilities

<b>Critical Capabilities</b>	<b>Quantum (ActiveScale)</b>	<b>Inspur (AS13000G5)</b>	<b>Red Hat (Ceph Storage)</b>	<b>Cloudian (HyperStore)</b>	<b>DDN (ExaScaler)</b>	<b>Dell Technologies (Dell EMC ECS)</b>	<b>Dell Technologies (Dell EMC PowerScale)</b>	<b>Pure Storage (FlashBlade)</b>	<b>Hitachi Vantara (Hitachi Content Platform)</b>	<b>IBM (COS)</b>	<b>IBM (Spectrum Scale)</b>	<b>Nutanix (Unified Storage)</b>	<b>Huawei (OceanStor Pacific)</b>	<b>Qumulo (Qumulo Core)</b>	<b>Scality (RING)</b>	<b>NetApp (StorageGRID)</b>	<b>Weka (WekaFS)</b>
Storage Efficiency	2.0	4.0	4.0	4.2	2.5	3.0	4.5	3.5	4.0	2.5	3.7	3.0	4.5	3.5	3.0	4.0	3.0
Security and Multitenancy	3.4	3.9	3.9	3.9	3.5	3.6	4.1	3.5	4.3	3.4	3.4	3.0	3.9	3.8	4.0	4.0	3.6
Capacity	3.5	4.0	4.0	4.0	4.5	3.5	4.0	3.8	4.0	4.0	4.5	3.5	3.5	3.5	4.5	3.8	4.5
Interoperability	3.1	3.1	4.3	4.0	4.2	4.3	4.1	3.8	4.4	4.5	3.4	4.2	4.1	3.8	4.2	4.2	3.8

Manageability	3.8	3.8	4.2	4.0	4.2	4.2	3.6	4.2	3.6	4.3	4.0	4.0	3.9	4.7	4.2	4.0	4.0
Performance	3.5	4.4	3.0	3.8	4.8	3.0	4.5	5.0	4.0	4.0	4.9	3.8	3.5	4.5	4.0	3.6	4.7
Resilience	3.4	4.3	3.5	3.5	3.5	3.3	3.1	4.5	3.3	4.8	4.0	4.0	4.2	4.0	4.1	4.0	4.0
Value	4.3	3.8	3.5	4.7	4.4	4.5	4.0	3.5	4.2	4.5	4.0	3.8	4.3	4.1	4.2	4.3	4.0
As of 24 September 2021																	

Source: Gartner (October 2021)

Table 3: Product Score in Use Cases

Use Cases	Quantum (ActiveScale)	Inspur (AS13000G5)	Red Hat (Ceph Storage)	Cloudian (HyperStore)	DDN (ExaScaler)	Dell Technologies (Dell EMC ECS)	Dell Technologies (Dell EMC PowerScale)	Pure Storage (FlashBlade)	Hitachi Vantara (Hitachi Content Platform)	IBM (COS)	IBM (Spectrum Scale)	Nutanix (Unified Storage)	Huawei (OceanStor Pacific)	Qumulo (Qumulo Core)	Scality (RING)	NetApp (StorageGRID)	Weka (WekaFS)
Analytics	3.41	3.93	3.74	4.00	4.05	3.65	4.04	4.06	3.99	4.02	4.06	3.69	3.94	4.04	4.04	3.96	4.02
Archiving	3.46	3.89	3.80	4.09	3.98	3.77	3.97	-1.00	3.99	4.05	3.98	3.67	4.03	3.97	4.05	4.03	-1.00
Backup	3.41	3.92	3.77	4.04	3.97	3.68	4.03	3.95	4.01	3.98	3.99	3.63	3.98	3.98	4.02	3.99	-1.00
Hybrid Cloud Storage	3.41	3.88	3.85	4.00	-1.00	3.73	3.96	-1.00	3.97	4.02	3.96	3.68	3.97	4.04	4.05	3.99	3.95

Cloud-Native Applications	-1.00	3.86	3.86	4.00	-1.00	3.73	-1.00	-1.00	3.99	4.04	-1.00	3.69	3.98	-1.00	4.06	4.00	-1.00
Cloud IT Operations	3.41	3.90	-1.00	-1.00	-1.00	-1.00	3.91	4.01	3.92	4.07	-1.00	-1.00	3.99	4.05	-1.00	4.00	-1.00
Commercial HPC	-1.00	3.97	-1.00	-1.00	4.09	-1.00	4.06	4.08	-1.00	-1.00	4.12	-1.00	3.93	4.05	-1.00	-1.00	4.06
As of 24 September 2021																	

Source: Gartner (October 2021)