White Paper

The Use of Object Storage for Modern Data-Intensive Workloads Requiring High Performance

Sponsored by: Nutanix

Carol Sliwa
October 2023

IDC OPINION

Object storage is no longer merely a cheap-and-deep repository for backups and archives of cold data. New and improved object stores now also target data-heavy workloads that require high performance, such as analytics, artificial intelligence (AI), machine learning (ML), and deep learning (DL). One key trend that helped expand the use cases for object storage was the introduction of flash drives to address applications with requirements for low latency and high input/output operations per second (IOPS). Start-ups demonstrated the potential for object storage built on container-based designs to deliver higher throughput than ever before, and many established vendors re-architected their object stores and added new capabilities to speed data access and metadata operations.

Amazon's Simple Storage Service (S3) validated the use of object-based storage for massive data sets, and its growing popularity fueled demand for other vendors to support the Amazon S3 API. Most enterprises now take a hybrid multicloud approach to infrastructure, and they stand to benefit from systems that support a common API. Software-defined storage built on a distributed architecture designed to span on-premises datacenters, edge sites, and public clouds can offer advantages over traditional SAN and NAS systems. Object stores can expand capacity and throughput with the addition of commodity server nodes to reduce costs. They use a global namespace to access potentially billions of objects across clusters that can span beyond hundreds of federated nodes, in contrast to hierarchical file systems with scaling limitations. Object stores also support more extensive, user-defined metadata than file or block storage to facilitate the search and analysis of enormous volumes of unstructured data across private and public clouds.

IDC forecasts that spending on object storage, inclusive of on-premises products and cloud-based services, will increase at a compound annual growth rate (CAGR) of 13.4% from 2021 to 2026. Shipped capacity will grow at a CAGR of 25.6% over that time frame, with most data in the multitenant shared/public cloud object stores of major providers such as Amazon, Google, and Microsoft. However, IDC predicts spending on single-tenant dedicated/private cloud object-based storage will increase at a higher CAGR of 22.9% than shared/public cloud (16.3%) through 2026, reflecting the ever-escalating volumes of unstructured data and growth of analytics, AI/ML, and Internet of Things (IoT) workloads that enterprises increasingly deploy as part of digital transformation initiatives. IDC research shows the fastest-growing workloads driving spending on enterprise compute and storage infrastructure from 2022 to 2027 will include AI life cycle (CAGR of 12%), unstructured databases, often used for AI (CAGR of 11.4%), and text and media analytics (CAGR of 10.3%).

Performance will continue to be one of the most important areas of focus for object storage suppliers that seek to expand to analytics and AI use cases. One such vendor, Nutanix, introduced object storage in 2019 as a complement to the block and distributed file storage that its popular hyperconverged infrastructure (HCI) platform already offered. S3-compatible Nutanix Objects Storage is now an integral part of the Nutanix Unified Storage (NUS) brand the vendor launched in 2022 to simplify licensing and packaging and spotlight the common management capabilities of its software-defined, multiprotocol storage. Nutanix places special emphasis on performance-oriented features, validations with prominent analytics vendors, and internal benchmark test results as it seeks to demonstrate that its object storage can meet the requirements of demanding data-intensive workloads. For primary object storage, Nutanix targets analytics platforms such as Vertica, Dremio, Snowflake, and Apache Spark and other ad hoc query engines, including Presto, Trino, and Spark SQL. Another primary storage use case is Hadoop Distributed File System (HDFS) replacement. Secondary storage targets for the Nutanix object store range from Splunk SmartStore, Confluent Platform for Apache Kafka, and Elasticsearch to backups and archives.

SITUATION OVERVIEW

The performance that scale-out object stores can now generate in clusters of tens, if not hundreds, of federated server nodes is significant. Start-ups began demonstrating several years ago that large object storage clusters could deliver triple-digit gigabytes per second of read and write throughput in test environments. Running a 32-node object storage cluster with low-latency, nonvolatile memory express (NVMe)–based solid state drives (SSDs) and 100GbE networking pushed performance to 350GBps for reads and 175GBps for writes in one benchmark test. Although "hero" numbers produced in ideal test conditions may not equate to real-world deployments, they suggest that object storage has the potential to push well beyond its historical limitations and handle demanding, data-intensive workloads, especially as clusters scale across tens or hundreds of nodes.

While many of today's object stores can linearly scale performance and capacity to address a broad range of workloads, from data lakes for demanding analytics and AI applications to colder backups and archives, that was not always the case. Object storage traces its history to the content-addressable storage (CAS) market that emerged in the early 2000s. Like modern object stores, CAS appliances aimed to store large amounts of fixed content across a cluster of servers more cost effectively than block-based SAN arrays, scale more easily than NAS systems built on hierarchical directories and SANs, and speed access to data over tape-based archives. Products convert content into immutable objects, using hashing algorithms to assign a unique identifier and address to each one, and add metadata to help make the content searchable. The main downside of the old disk-based CAS appliances was slower performance than block and file alternatives, largely due to the processing overhead, and the systems were generally relegated to secondary storage.

Although the CAS market eventually dissolved, the underpinning architecture and technology lived on in the object storage market that emerged in its aftermath among a group of established storage vendors and start-ups. Performance remained a drawback, but object storage got a major boost with the 2006 launch of the cloud-based Amazon Simple Storage Service, providing storage on demand for users anxious to minimize time-consuming infrastructure planning, provisioning, and management. Amazon S3 grew so widely popular that demand grew for other vendors to support the S3 APIs, as developers increasingly started to build applications to them. On-premises storage vendors enabled data offloads to S3-based object storage for backups, archives, and disaster recovery. Those integrating S3 support into products began promoting hybrid cloud and multicloud strategies. Several major storage vendors acquired start-ups to get into the burgeoning object storage business. More recently, some have taken the extra step of launching on-premises storage-as-a-service options designed to mimic the public cloud model and address some of its drawbacks, including the transaction and data egress fees that have led some organizations to repatriate workloads to onsite datacenters.

On-premises and public cloud object storage providers now share a common objective of expanding target use cases to data lakes, analytics, AI, ML, DL, and other data-intensive applications that require higher performance than backups and archives. Most AI/ML platforms support S3 API-based object storage. Amazon S3 performance guidelines advise customers to examine network throughput, CPU, and DRAM requirements and use HTTP analysis tools to look at DNS lookup time, latency, and data transfer speed. Other tips include scaling storage connections horizontally to maximize accessible bandwidth, combining Amazon S3 storage and Amazon Elastic Compute Cloud (EC2) instances in the same AWS region, and using Amazon S3 Transfer Acceleration to reduce latency over long geographic distances. On-premises object storage can offer a potentially compelling alternative for enterprises that hope to minimize latency, have greater control over performance tuning, and eliminate transaction and egress costs, as well as organizations that need to keep data in their own datacenters for regulatory or compliance purposes.

On-premises object stores that support NAND flash storage have grown in importance with the influx of modern workloads that require high throughput, high IOPS, and low latency to accelerate data queries and memory operations. Standard SAS/SATA solid state drives can provide a substantial performance boost over hard disk drives (HDDs). Newer nonvolatile memory express–based PCI Express (PCIe) SSDs can reduce latency and increase throughput even more than SAS/SATA SSDs. Flash benefits also include lowering power consumption and condensing storage footprints since SSD-based systems can generally meet performance requirements with fewer drives. SSDs also tend to have lower failure rates than hard disk drives to ease system maintenance, and they can reduce rebuild and recovery times if a drive fails, helping ensure data durability as SSD densities increase to 30TB and beyond. On the downside, SSDs cost more on a price-per-gigabyte basis than spinning disk, although new SSD options with dense quad-level cell (QLC) flash that stores 4 bits per cell can lessen the cost differential with HDDs.

Since many workloads do not demand the highest possible levels of performance, organizations must determine the balance of flash and spinning disk that will most cost effectively meet their requirements. Enterprises in need of tens or hundreds of petabytes, if not exabytes, of capacity often opt for HDDs for cool and cold data. Cheap spinning disk may suffice with batch-oriented analytics applications that do not require instant response time or with workloads that have moderate throughput needs. Reserving SSDs for metadata operations may be an option to improve performance in an HDD-based object storage cluster.

Historically, enterprises used block-based SANs and file-based NAS systems for applications requiring the highest levels of performance. Given the changing nature of modern workloads, organizations would be wise to consider software-defined storage. Analytics and AI/ML/DL applications often require massive data sets to produce insightful results, and storage capacity tends to grow significantly over time. Scale-out software-defined systems, such as parallel file systems and object storage, are designed to expand to store petabytes, if not exabytes, of data. Object storage vendors have introduced new systems, re-architected existing products, and/or added new features to reduce the overhead associated with content and metadata processing and improve performance. Enterprises have a choice of single-protocol object storage, products that support file and object, and systems that support data access via block, file, and object protocols. Nutanix Unified Storage, for instance, supports block, file, and object storage services on its popular HCI platform.

Whether block, file, object, or multiprotocol, most storage systems now support configurations tailored for data-intensive analytics and AI/ML/DL, and some are better suited than others to handle the differing aspects of workloads. An AI data pipeline may include stages for data ingestion, labeling, training, validation, inference, and retention. Each stage may have different storage requirements. For instance, the ingest stage requires massive concurrency and write throughput, but the training stage needs massive read throughput, and inference demands low-latency access and read performance. A 2021 IDC survey of 2,000 IT and line-of-business professionals showed 53% use block storage, 53% use file, and 45% use object, in varying combinations, for AI training. A unified data repository for structured and unstructured data was the most popular choice for AI inference, at 50%, followed by file (45%), object (43%), and block (21%) storage. Flash-based object storage with support for multiple protocols and interfaces may be a consideration for most, if not all, AI stages, while disk-based object systems might be limited to workloads such as model training or data lakes. Some organizations combine a fast file system on the front end with a more scalable, cost-efficient, metadata-rich object store on the back end for long-term data retention and protection. Factors to take into consideration when weighing storage decisions include the supported data access protocols of the applications, performance and latency requirements, the size of the files and objects, and the projected capacity over time.

One of the fastest-growing workloads for disk-based object storage is Splunk SmartStore. Launched in 2018, Splunk SmartStore enables the decoupling of compute and storage resources and offloading of colder data to a separate object-based tier. SmartStore software analyzes the age, access frequency, and priority of the data to determine the content that needs to remain cached close to the compute resources for real-time analytics and the colder data that can move to less expensive S3-based object storage on premises or in the cloud. Splunk SmartStore tiers can grow quite large. The disk-based Splunk SmartStore object storage for one financial institution's fraud detection data lake now exceeds 80PB.

Software-defined object storage merits consideration for modern applications with petabyte-level, if not exabyte-level, capacity needs, given its potential to ease scaling over many block- and file-based alternatives. Enterprises need to take stock of the I/O and latency requirements of all stages of their AI data pipelines to make the optimal storage decisions for each one. Cost-effective scale-out object stores can help customers save money with warm and cold unstructured data that may constitute 80% to 90% of their storage needs. Public cloud may also be an important part of the unstructured data storage strategy for many enterprises, underscoring the need for software-defined infrastructure with strong support for hybrid, multicloud integration and management.

NUTANIX OBJECTS STORAGE PROPOSITION FOR DATA-INTENSIVE WORKLOADS

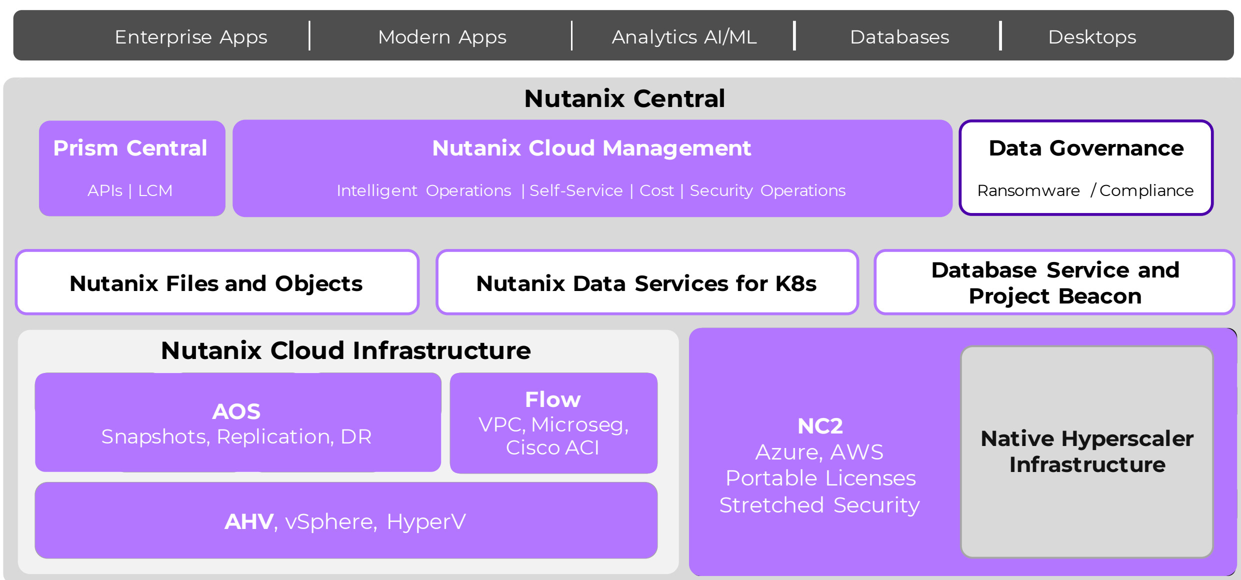
Founded in 2009, Nutanix is a $1.5 billion provider of software-defined infrastructure that can run on commodity hardware and span hybrid multicloud environments. Nutanix grew its customer base to about 25,000, based largely on its pioneering work with hyperconverged infrastructure designed to ease deployment, operations, and management by combining compute, storage, networking, and virtualization resources in a single system. IDC research has shown that most organizations now take a hybrid cloud or hybrid multicloud approach, and HCI users often replace traditional three-tier architectures with consolidated hyperconverged systems as they refresh their aging IT infrastructure.

Reducing complexity can help IT departments stay agile and meet fast-changing business and application needs as they pursue digital transformation projects to gain greater value from their data assets.

Nutanix HCI appliances originally supported block-based storage, but the company added distributed scale-out file storage in 2017 and object storage in 2019 with its shift from an appliance focus to a software-defined infrastructure approach. The company introduced the Nutanix Unified Storage brand in February 2022 as part of a portfoliowide effort to simplify packaging, pricing, and management for customers deploying the products across core datacenters, edge locations, and public cloud sites. Built on the Nutanix Cloud Platform, NUS offers file, object, and block storage services through a single capacity- and consumption-based licensing model designed to let customers pay as they grow (see Figure 1). Customers can independently scale computing and storage resources and use the object store in dedicated/standalone mode, with all VMs in a node running storage services, or in HCI mode, with the node's VMs also able to run applications.

FIGURE 1

Nutanix Cloud Platform Architecture



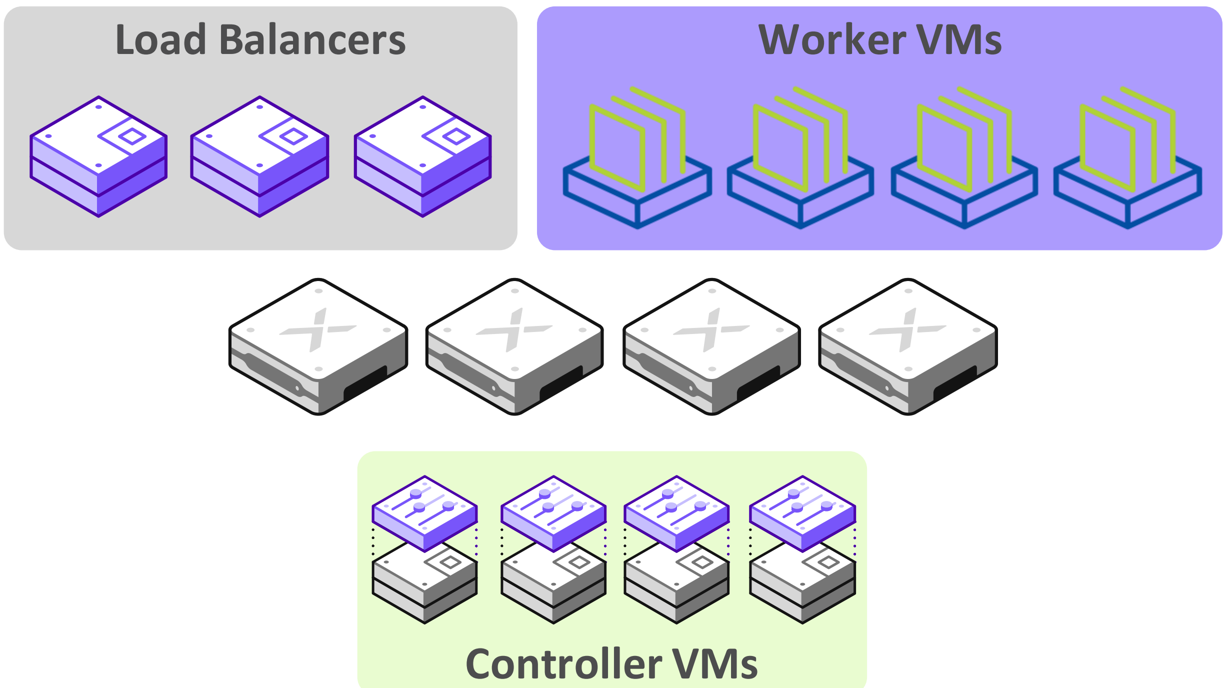
Source: Nutanix, 2023

Software-defined Nutanix Objects Storage is a containerized deployment that offers customers a choice of running on dedicated, standalone hardware or on servers that also run Nutanix HCI, which can manage Files, Objects, or Volumes data. NUS leverages the Nutanix Acropolis Operating System (AOS) for capabilities such as compression, deduplication, encryption, and erasure coding. Each physical node runs a controller virtual machine (CVM) to build out the virtualization layer and AOS cluster. Nutanix designed the object store to communicate directly to the CVM using its own native APIs for improved performance. Worker node VMs run the Nutanix microservices platform with the containerized, Kubernetes-orchestrated component services that make up Nutanix Objects Storage. Each worker node contains an object controller that contributes to the available read/write bandwidth.

Built-in load balancers receive and distribute S3 client requests to the worker VMs, which provide the S3 interface, store data and metadata, and integrate with the Prism Central unified management tool (see Figure 2). The Objects architecture is designed to support multiple instances of the main services to prevent performance bottlenecks.

FIGURE 2

Nutanix Objects Storage Architecture



Source: Nutanix, 2023

Nutanix Objects Storage users can start with a single node and scale performance and capacity in linear fashion to 48 nodes and 16.8PB per cluster, with options to federate multiple clusters under a single shared namespace. The product supports flash-based NVMe and SAS/SATA SSDs for high performance and dense HDDs as a lower-cost alternative with hybrid or disk-only configurations. Nutanix says the largest all-flash deployment is about 6PB. Nutanix Objects Storage features include streaming-style replication, tiering to other S3-based object stores on premises or in public clouds, write once, read many (WORM) buckets, and object locking for built-in ransomware protection. Although its focus is now software, Nutanix continues to partner with vendors such as Commvault and Veeam on turnkey appliances that bundle storage and backup software.

Nutanix estimates that backup and archive represent about 45% of the object storage use cases, and the remainder includes data-intensive workloads such as analytics and AI/ML. For primary use cases, Nutanix Objects Storage is validated for use with the Dremio, Vertica, and Snowflake analytics platforms. Nutanix Objects Storage is designed to support any query engine that works with HDFS and S3 buckets, including open source Apache Spark SQL, Presto, and Trino. For primary storage batch jobs, Nutanix also focuses on open source software, with support for Apache Spark and Hadoop MapReduce. In the realm of secondary storage, Nutanix validated its Objects storage with Confluent Kafka and Splunk SmartStore. Another key secondary storage use case, Elasticsearch, does not require validation because Nutanix supports the open source version. The August 2023 announcement of Nutanix GPT-in-a-Box adds a software-defined "AI ready" platform and services designed to help customers configure hardware and software infrastructure to run generative pretrained transformers (GPT), including large language models at the edge or in datacenters, using open source AI and MLOps frameworks on the Nutanix Cloud Platform with Nutanix Objects Storage and Nutanix Files Storage.

The 2023 partnership with Snowflake represents one of Nutanix's most recent collaborative efforts to offer customers greater flexibility. Two years earlier, Snowflake announced the general availability of an External Tables feature that lets customers query external data as if it were stored inside a Snowflake table. Customers are now able to run ad hoc queries directly on data stored in external systems, such as Nutanix Objects Storage, without having to first copy the potentially large data set to the Snowflake Data Cloud. Nutanix supports queries on select portions of the data sets to help speed response time.

One performance-related capability that can be helpful with Hadoop MapReduce batch jobs is a "fast copy" feature that enables Nutanix Objects Storage to speed the deduplication of Objects data to minimize performance overhead and reduce the need for additional storage capacity. To accelerate data crunching, MapReduce breaks up a task into chunks to enable individual compute nodes to address a portion of the problem. The process is designed to work with HDFS and generally involves a significant amount of file renames. Renaming a file may be a quick, inexpensive operation in a POSIX file system, but copy operations can be taxing on an object storage system that must create a new unique object to make any changes or updates. Nutanix's fast copy feature enables the object's metadata to serve as a new logical copy, eliminating the need to store another full copy of the object.

IDC interviewed two users of Nutanix Objects Storage about the performance they observed with analytics-related workloads. One, Nutanix's internal IT department, deploys Nutanix Objects Storage with Dremio's analytics engine to power a self-service platform designed to give about 250 employees access to more than 40 on-premises and cloud data sources. Nutanix launched the data-as-a-service platform in 2019 on HDFS-based storage and replaced it with Nutanix Objects in 2021 to lower costs and simplify the architecture and management. Proof-of-concept (POC) testing done before the switch from Hadoop-based storage to Nutanix Objects Storage showed query response time improved, at a range from 2.4% to 16.0%, in selected test results supplied by Nutanix IT. With its next platform iteration, Nutanix IT plans to replace its VM-based Nutanix Objects Storage implementation with a Kubernetes-based Nutanix cluster to reduce the infrastructure footprint, simplify cluster management, enable rolling upgrades, and improve linear scalability. Nutanix IT viewed the early POC results from the scaled-down Kubernetes test cluster as close enough to the VM-based Nutanix production cluster in most cases to merit the switch. POC testing showed that query runtimes were higher in nine cases (ranging from 0.5% to 25.0%, or an average of 13.4%, excluding one aberrant query result with an increase of more than 400%), lower in three instances (ranging from 5.9% to 25.0%), and equivalent in three. The data-as-a-service infrastructure using Kubernetes remained in testing at the time this document was written.

One longtime Nutanix customer, a government agency based in Asia, viewed Objects as a natural addition, for a new Splunk SmartStorage archive, to maintain a single IT ecosystem. The agency deploys roughly 10 Nutanix Objects Storage clusters, each consisting of 4 to 8 hybrid nodes, with

SSDs for caching warm data and HDDs for colder data. The Splunk infrastructure manager noted that the agency performed simple file-transfer tests to ensure adequate throughput prior to deploying Nutanix Objects in late 2019 or early 2020, but it has done no testing since the system went into production. He was unable to offer performance comparisons to prior storage because Splunk SmartStore was a new deployment. His main qualitative observation on the Nutanix Objects Storage production environment was that the system could handle a "massive" amount of simultaneous PUT and GET requests without issue, with the hotter data stored on SSDs returned faster than the colder data on HDDs.

Nutanix continues to make performance-focused improvements to its Objects storage software. One noteworthy update, in Objects 3.5.1, increased the maximum capacity for the product's metadata block cache. The update can be especially helpful in lowering latency for the repeated list operations that an analytics application must perform to identify target objects for the execution of data operations. Nutanix engineers optimized the Objects architecture to make available extra memory for the cache. Another important update, with the Objects 3.2 release, was a rewrite of the S3 front-end adapter that forms part of the object controller. Nutanix replaced the open source code that it used for front-end S3 functionality with a new "TurboS3" adapter that it claims improved the operations per second of small-object GETs/reads by up to 95% and small-object PUTs/writes by up to 32% in an internal test environment.

Nutanix's testing demonstrated the level of performance — 49,503 operations per second (ops/s) for GETs and 20,469ops/s for PUTs — a 4-node all-flash Objects cluster could achieve on small 8KB objects. The highest throughput Nutanix reported on larger 8MB objects in the 4-node all-flash environment was 14.3GBps for reads and 7.2GBps for writes. Nutanix also provided results to show that performance scales linearly in an Objects storage cluster. Company tests indicated the throughput of an 8-node hybrid cluster was roughly double the throughput of a 4-node hybrid cluster under the same test conditions. Nutanix tested another key performance metric, latency, by measuring the elapsed time between the submission of a GET request and the receipt of the first byte of data. Time to first byte (TTFB) for a 4-node all-flash cluster was less than 3 microseconds, according to the internal Nutanix tests.

Nutanix has conducted extensive tests to validate the performance of its object storage, in some cases with popular analytics applications and query engines. IDC reviewed the results, but since it was not involved in the Nutanix test process, it cannot verify the test conditions or the results. However, IDC recognizes that performance is a top concern for customers with modern data-intensive workloads and includes selected Nutanix-provided results with the caveat that test conditions often do not mirror real-world environments. Performance comparisons are difficult, but the test results serve as an indicator of the potential performance a system can achieve under optimal conditions or with a similar configuration. They also demonstrate the ongoing commitment Nutanix has made to continue to improve the performance of its object storage product for modern data-intensive workloads.

CHALLENGES/OPPORTUNITIES

The competitive landscape in the object storage market consists largely of mature standalone products that well-established vendors updated significantly over the years and newly designed systems that start-ups and other vendors built on lightweight container-based designs with modern S3- and Kubernetes-based cloud-native and data-intensive applications, such as data lakes for AI/ML and analytics, in mind. As part of a popular hyperconverged infrastructure platform, Nutanix Objects

Storage offers a different sort of value proposition than standalone object stores. The Nutanix Cloud Platform combines computing, storage, networking, and virtualization resources under common management, and its Nutanix Unified Storage supports block, file, and object interfaces. Nutanix Objects Storage has a considerable opportunity for growth, especially among existing HCI users with data-intensive workloads and enterprises seeking to condense their IT infrastructure to ease deployment, operations, and management.

One ongoing challenge all HCI vendors face is convincing potentially skeptical customers that HCI-based storage can handle the performance and scalability requirements of modern data-intensive workloads. IDC's January 2023 *IT Infrastructure for Storage and Data Management Survey* showed that HCI users encounter challenges getting sufficient storage performance without excessive cost and find it difficult to scale storage separately from compute resources. Nutanix has done extensive work to improve the performance and efficiency of its Objects storage, and it must continue to enhance the product and find effective ways to relay the information to customers if it hopes to expand the use of its object storage to additional workloads.

CONCLUSION

Use cases for object storage have expanded beyond backups and archives to modern data-intensive workloads that require higher performance, including analytics, artificial intelligence, and machine learning. The transition makes sense from a storage capacity standpoint, given an object system's ability to cost effectively scale without theoretical limit through the addition of commodity server nodes to clusters. Performance has always been the main concern, due to the overhead associated with the way object storage processes and handles content and metadata. Object stores can linearly scale throughput in the same way they scale storage capacity, but they face challenges with workloads that require high IOPS and/or low latency. The advent of flash-based object storage started to address the performance questions, as did an influx of new and rearchitected systems with improved capabilities to help streamline processing. Object stores still may not match the highest-performing block- and file-based storage systems, but they warrant consideration in particular for workloads that need to scale throughput with massive data sets.

Nutanix Objects Storage is one of the many object-based storage systems to prioritize performance-related enhancements to better target modern data-intensive workloads. Nutanix rewrote its S3 front-end adapter and added features, such as fast copy for MapReduce, to help improve the performance of its Objects storage with analytics platforms and query engines. The company has also expended considerable effort to validate Nutanix Objects Storage with popular big data applications such as Vertica, Snowflake, Splunk SmartStore, and Confluent Kafka and recently added a Nutanix GPT-in-a-Box option to ease infrastructure deployment for generative AI workloads. Nutanix conducted extensive internal tests to try to demonstrate that its object storage can meet the performance requirements of demanding workloads. Performance metrics can be difficult to assess and compare since test conditions and configurations do not equate to the real-world environments of customers. However, viewing the test results in the context of the performance-related updates the company has made, Nutanix can make a strong case that its object storage merits consideration for modern data-intensive workloads, especially among customers that already use Nutanix HCI products or hope to consolidate their IT infrastructure.

About IDC

International Data Corporation (IDC) is the premier global provider of market intelligence, advisory services, and events for the information technology, telecommunications, and consumer technology markets. With more than 1,300 analysts worldwide, IDC offers global, regional, and local expertise on technology, IT benchmarking and sourcing, and industry opportunities and trends in over 110 countries. IDC's analysis and insight helps IT professionals, business executives, and the investment community to make fact-based technology decisions and to achieve their key business objectives. Founded in 1964, IDC is a wholly owned subsidiary of International Data Group (IDG, Inc.).

Global Headquarters

140 Kendrick Street
Building B
Needham, MA 02494
USA
508.872.8200
Twitter: @IDC
blogs.idc.com
www.idc.com