

Intel Helps Kingsoft Cloud Create Hybrid Cloud Storage Products to Accelerate the Digital Transformation of Enterprises



“From the traditional point of view, the deployment and management of a company’s local data center storage and public cloud storage were separate from each other. Hybrid cloud storage has a stronger emphasis on the organic integration and centralized management of public cloud storage, data center storage, and edge storage in order to implement the centralized management of on-cloud and off-cloud storage and computation. This presents a lot of new challenges to the design of hybrid cloud storage architectures. Furthermore, storage systems of large and medium-sized enterprises will increasingly need to support AI and other performance-intensive loads. Traditional storage products are already unable to satisfy the requirements or support these services. Therefore, we hope to sufficiently utilize Intel’s innovative outcomes in areas such as processors and SSD and deploy a high-performance hybrid cloud storage system.”

– Liu Tao
Partner and Vice President of
Kingsoft Cloud

Overview

Digital transformation has become a key direction for enterprises to enhance competitiveness and promote business growth and innovation, while the importance of “getting on the cloud” is increasingly obvious. Thanks to its advantages in security and agility, etc., hybrid cloud has become the first choice for many enterprises in cloud construction of infrastructure. It has also become an important bridge for traditional enterprise customers who want to migrate to the cloud.

To help enterprises hasten the deployment of hybrid cloud storage, Kingsoft Cloud released the Kingsoft Cloud Hybrid Cloud Storage ES series of products, based on 2nd Generation Intel® Xeon® Scalable processors, Intel® SSD DC P4510 Series¹, and Intel® Ethernet Converged Network Adapters. This series of products have advantages, including high performance, high availability, high reliability, and scalability, that can help enterprise users meet the various data storage challenges they face while building hybrid cloud data centers. It also helps in unleashing the potential value of data.

Challenge: Deployment bottlenecks in hybrid cloud storage

The digital economy is currently being widely incorporated into many different industries, and there has been an explosive growth in data and a huge concentration of data. From this explosive growth, we can predict that in the coming years, enterprises will have to invest an increasing amount into data storage. Although current public cloud storage has already become an important choice for enterprises, there is still a certain degree of anxiety over public cloud storage due to factors such as security, compatibility, and business compliance. Against this backdrop, a new type of hybrid cloud storage that can implement the centralized management of public clouds as well as the enterprises’ own local and edge computing and storage resources emerges as a force to be reckoned with. By integrating data center and cloud computing resources, the hybrid cloud storage is able to combine commercial data centers, customer data centers, or even public and private cloud and local IT infrastructure in order to satisfy IT requirements of enterprises in the new era. Therefore, this will become a major trend in the development in digital transformation.

According to a prediction² by IDC, by 2021, over 90% of enterprises in China will rely on a combination of edge, local/dedicated private clouds and multiple public clouds, as well as offline data centers to satisfy their infrastructure requirements. By 2022, 50% of enterprises will have deployed centralized VMs, Kubernetes, and multi-cloud management processes and tools used to support the management and administration of multiple clouds deployed across local and public clouds. In addition, 85% of managers of large-sized enterprises have indicated that they plan to distribute data storage over multiple locations to improve data security and availability. They will use more hybrid cloud storage methods suitable for their businesses and use various data lifecycle management technologies to implement data storage and the free flow of data in hybrid cloud environments.

Enterprise storage architectures have continued to evolve, but in order to satisfy enterprise storage requirements that demand adaptivity to complex and variable scenarios, there are still many challenges in hybrid cloud storage:

- **Demand for high performance and elastic storage:** Hybrid cloud data centers must satisfy the high IOPS requirements of traditional databases and other applications while also supporting big data, deep learning, and other applications that have a large number of small files and require low latency data access. This presents a combination of demands on the overall hybrid cloud storage system including demand for high IOPS, high concurrency, high reliability, high flexibility, and scalability, which are all needed to solve the complexity and performance issues related to the rapid growth of data in hybrid cloud data center storage.
- **Multi-protocol support:** Various types of structured, semi-structured, and unstructured data needs to be stored on hybrid clouds, and different applications also require the use of block (iSCSI/ FC), file (POSIX/NFS/CIFS), object (S3), big data (HDFS), and other types of storage and interfaces. Therefore, planning and the meeting of different requirements in different types of applications present a major challenge when building a hybrid cloud storage system.
- **Data lifecycle management requirements:** Using typical AI training workflow as an example, the data collection, data cleaning and tagging processes require the processing of a huge amount of unstructured data such as image data. These types of data require a huge amount of storage space and high concurrency sequential

read and write access, which can become very costly. While at the same time, model training requires high concurrency, high IO, and storage support for random data access. If there wasn't any support from reasonable data lifecycle management policies, the entire storage system will face severe problems in cost control and performance bottlenecks.

Solution: Kingsoft Cloud's hybrid cloud storage based on Intel® architecture

Thanks to Kingsoft Cloud's wealth of experience operating with enterprise-level businesses, its many years of experience with public cloud storage services, as well as its in-depth knowledge of the storage industry and technology frameworks, Kingsoft Cloud has, based on 2nd Generation Intel Xeon Scalable processor, Intel SSD DC P4510 Series and Intel Ethernet Converged Network Adapter, released the ES series hybrid cloud storage products targeted at enterprise users to help them resolve the various types of data storage challenges they face while building hybrid cloud data centers.

The applications for Kingsoft Cloud's ES hybrid cloud storage series of products are shown in Figure 1. The ES product series currently include the ES3000 Distributed Block Storage System, the ES5000 Distributed File Storage System, and the ES10000 Distributed Object Storage System. Each has different features and can be used for different storage requirements. Enterprise users can select different combinations of the hybrid cloud storage products to satisfy their requirements in cloud computing, virtualization, broadcasting media, file sharing, file transfer, big data analysis, broadcasting media, file sharing, file transfer, big data analysis, backup and archiving, security, and other applications.

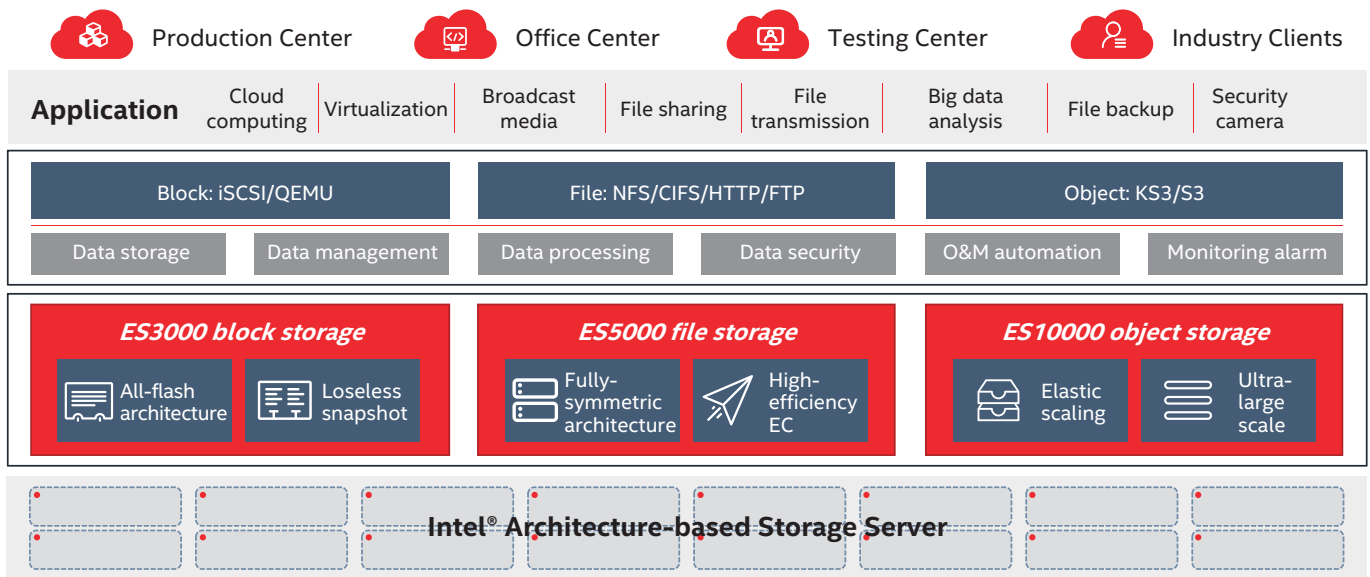


Figure 1. Kingsoft Cloud Hybrid Cloud Storage Product Family

ES3000 Distributed Block Storage System

The ES3000 Distributed Block Storage System uses the Kingsoft Cloud Elastic Block Storage infrastructure in its underlying layer and supports iSCSI standard ports, snapshots, replication and other enterprise-level features according to the business's requirements. It can provide enterprise users with high-performance, high-reliability, high-availability, and high-scalability elastic block storage services to satisfy the storage system requirements of private cloud platform storage, virtual platform storage, database, AI, big data storage, and other applications.

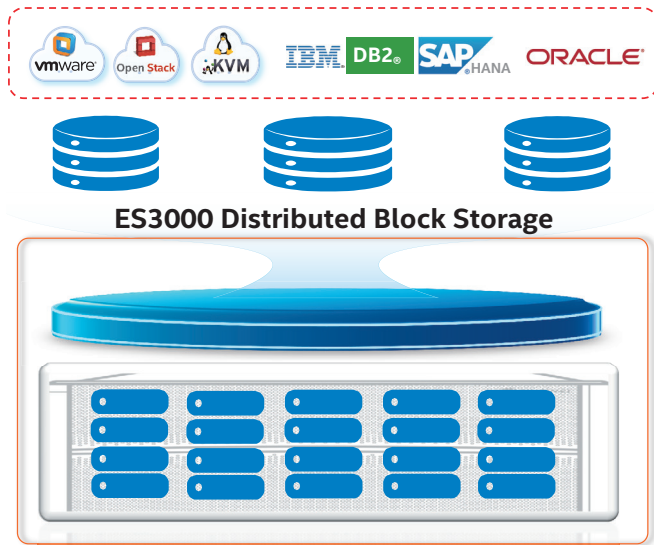


Figure 2. Architecture of the ES3000 Distributed Block Storage System

The ES3000 Distributed Block Storage solution can effectively resolve the following key challenges in hybrid cloud block storage:

- **High performance:** According to the actual configuration and business requirements, it can provide cloud hard disks that can be scaled to handle tens of thousands to millions of IOPS, and the performance can be improved linearly as the cluster expands.
- **High reliability:** After long-term and large-scale verification on the public cloud, the availability exceeds 99.95% and the reliability is as high as 99.999999999%.
- **Elastic scaling:** Elastic scaling allows storage to start at TB-level and be scaled to hundreds of PBs.
- **Enterprise storage product features:** Support cache acceleration, lossless snapshots, thin provisioning and other enterprise-level storage features.

ES5000 Distributed File Storage System

The ES5000 utilizes Kingsoft Cloud's Elastic Block Storage infrastructure in its underlying layer and supports standard file interfaces such as POSIX, CIFS, and NFS, as well as snapshots, replication, layering, WORM, and other enterprise-level storage product features. This provides high-availability, high-reliability, high-performance, and high-scalability file storage service to enterprise users and provides support for the requirements of broadcast media storage, HPC high performance storage, security video, storage, medical imaging storage, AI, big data, and other applications.

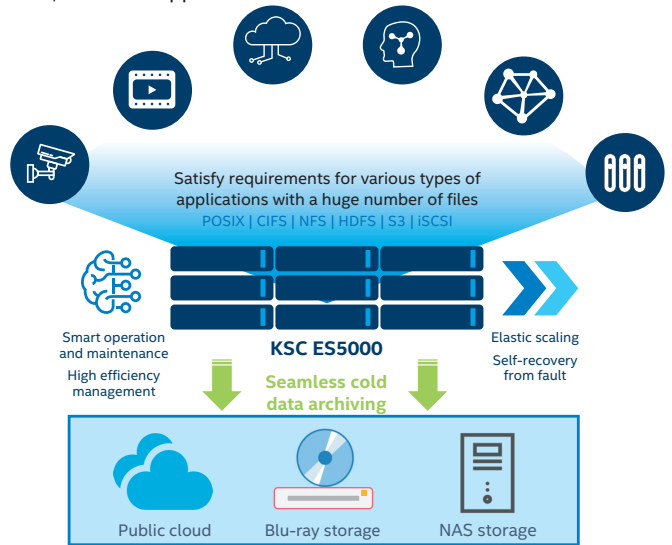


Figure 3. Architecture of the ES5000 Distributed File Storage System

The ES5000 Distributed Block Storage solution can effectively resolve the following key challenges in PB-level storage:

- **Unified namespace:** All storage servers within the system can replace the work of the I/O node, and the storage system comes with a load balancing mechanism which significantly improves the utilization of the compute nodes and the entire cluster system.
- **Software-defined storage architecture:** The general Intel® architecture design is used along with software-defined distributed file system, greatly reducing deployment and maintenance costs.
- **High performance:** ES5000 provides all-flash configuration, hybrid flash and HDD configuration to support millions of OPS and hundreds of GB/s bandwidth to satisfy the requirements of high-performance file storage applications.
- **Synchronization mechanism:** The system comes with its own data synchronization mechanism that can package small files, segment large files and perform differential synchronization, scheduled synchronization, and other functions to improve data synchronization efficiency.
- **Elastic scaling:** ES5000 uses a fully distributed architecture that supports horizontal scaling of hardware nodes, and can also guarantee system performance in large-scale applications via automatic load balancing policies.
- **Rich enterprise-level features:** ES5000 provides rich enterprise-level features and can help customers to easily build complete data protection mechanisms, to better plan and manage data resources and establish enterprise-level data centers.

ES10000 Distributed Object Storage System

ES10000 uses Kingsoft Cloud's object storage infrastructure (KS3) and object storage features in the underlying layer, and integrates the online functions of public clouds. Meanwhile, it is scaled to support iSCSI, CIFS, NFS and other interfaces according to the application requirements of private cloud clients, providing enterprise users with high-availability, high-reliability, and highly-scalable massive unstructured data storage service.

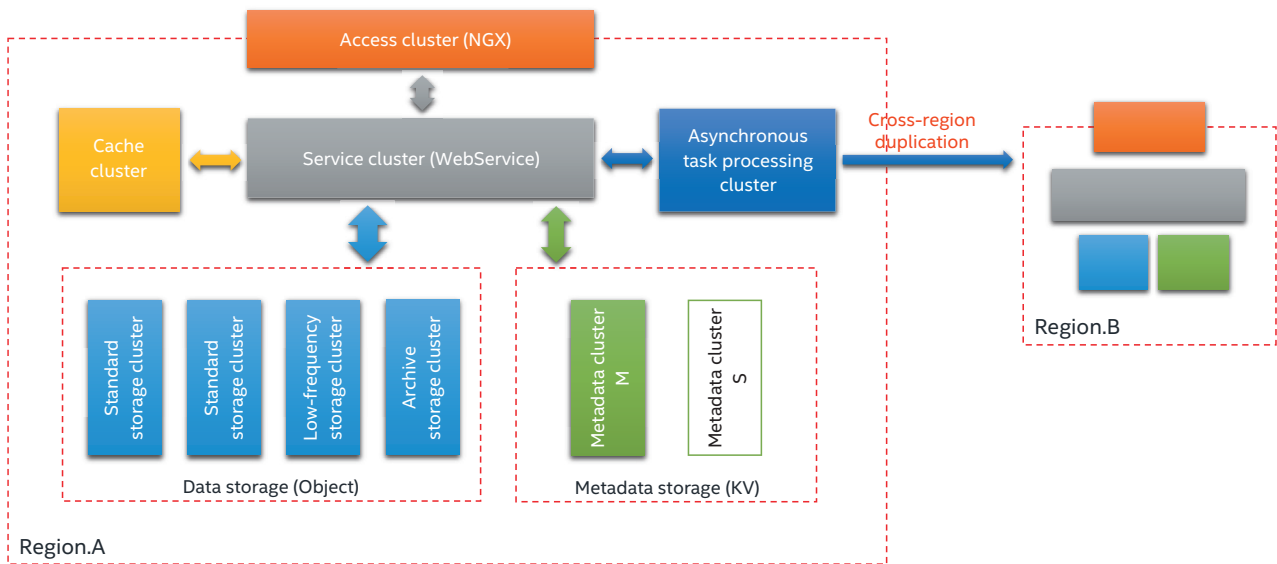


Figure 4. Architecture of the ES10000 Distributed Object Storage System

The ES10000 Distributed Block Storage solution can effectively resolve the following key challenges in object storage:

- **Massive single namespace:** ES10000 is able to support 1,000 nodes on a single cluster and 99 clusters in a single data center, as well as up to 10EB in capacity. EB-level storage on public clouds is also supported.
- **High concurrency access:** ES10000 supports bandwidths of hundreds of GB/s and millions of QPS and also supports higher access performance thanks to constant technological improvements.
- **No speed reduction even with faults:** ES10000 can better deal with the challenges of system operation under a single machine failure, and it contains four major hardware fault troubleshooting systems that can greatly reduce hardware failure damage and ensure that the system runs just as fast as before even when a failure occurs.
- **Quick performance and capacity scaling:** ES10000 supports TB-level capacity expansion completed in a week and PB-level capacity expansion in a month, and users will not even notice any disruptions during capacity expansion, so that enterprises will not have to pre-purchase a large number of machines thus avoiding high idle costs.
- **Elastic scaling:** ES10000 uses a fully distributed architecture and supports the linear increase in capacity and performance of the entire system via horizontal scaling of hardware nodes without the need for complicated resource requirement planning.
- **Rich enterprise-level features:** ES10000 provides rich enterprise-level features and supports enterprises to deploy complete data protection mechanisms and build reliable internal cloud storage systems.

"The entire product line of Kingsoft Cloud's ES hybrid storage supports business-oriented hybrid cloud storage lifecycle management, and users can select the optimal storage mode according to their actual requirements in performance, reliability, availability, and capacity. Data will flow between different storage resource pools as needed in order to reduce the total cost of ownership in a more targeted way."

– Liu Tao
Partner and Vice President of Kingsoft Cloud

Optimization: Intel helps Kingsoft Cloud's hybrid cloud storage to unleash the potential of data

Intel and Kingsoft Cloud have worked together closely to deploy high-performance hardware in a targeted manner according to the different performance features and requirements in the Kingsoft Cloud hybrid cloud storage product family. They also used software optimization to unleash the potential performance of hardware in different configurations.

After working with Intel's engineering team, and after actual measurements, evaluations and selection, Kingsoft Cloud chose to use 2nd Generation Intel Xeon Scalable processors to provide the computing power support and used the Intel SSD DC P4510 Series as the underlying medium for distributed storage. 2nd Generation Intel Xeon scalable processors have higher single core performance and are able to provide high performance and scalability to compute-intensive workloads in computing, storage and networking applications. Thanks to Intel® Ultra Path Interconnect (Intel® UPI), Intel® Infrastructure Management Technology (Intel® IMT), Intel® Advanced Vector Extensions 512 (Intel® AVX-512) and other leading technologies, Kingsoft Cloud meets the requirements of demanding I/O-intensive workloads, helps to create more powerful and agile services and provides more breakthrough functionalities in cloud data centers. The Intel® SSD DC P4510 series is based on Intel® 3D NAND technology, uses the NVMe interface, and can provide higher storage density, thus supporting a wider range of workloads applications. The P4510 comes with enhanced smart firmware algorithm, error correction, power failure protection and durability design, provides better QoS and can guarantee the optimal balance in read and write protection for data storage. Its annual failure rate (AFR) is significantly lower than HDDs, better satisfying the requirements of Kingsoft Cloud's hybrid cloud storage for availability and stability, etc. of storage system.

In addition to the Intel SSD DC P4510 Series, Kingsoft Cloud also uses the Storage Performance Development Kit (SPDK) for optimization³. SPDK provides a set of tools and libraries that can be used to write high-performance and scalable user-mode storage applications. It utilizes several key technologies to realize high performance such as moving some drivers to the user space to avoid system calls and allowing zero-copy access from applications. After using SPDK, user-mode drivers are implemented via polling hardware instead of relying on interrupts, helping to reduce total

latency and latency variance. Furthermore, when compared to using kernel drivers, there is a significant performance advantage in terms of IOPS for each CPU core. In addition, SPDK also has a lockless high performance I/O path mode that avoids all locks in key I/O paths but instead relies on message passing to share resources amongst multiple threads to achieve higher concurrency performance. Generally speaking, SPDK effectively integrates Intel's networking, data, and storage technologies and fully unleashes the performance potential of SSD storage mediums.

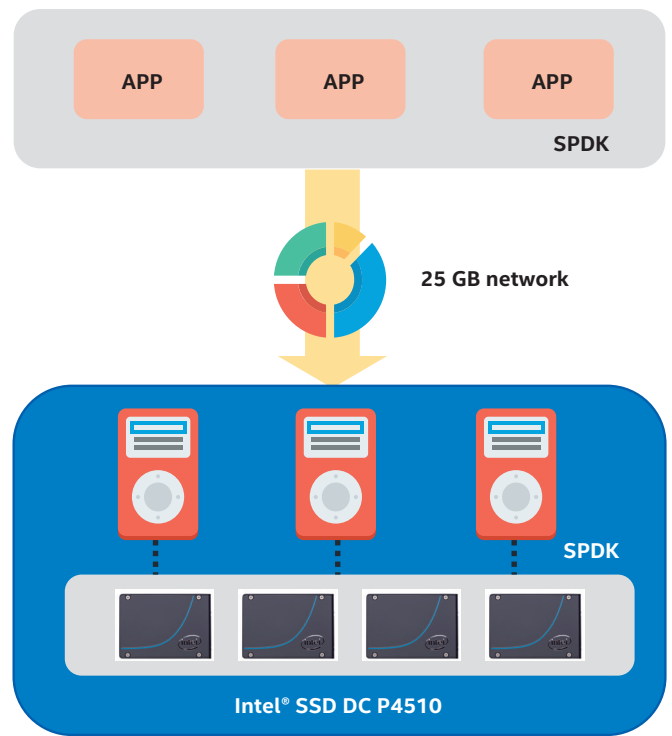


Figure 5. SPDK Application in Kingsoft Cloud's Hybrid Cloud Storage

The 2nd Generation Intel Xeon Scalable processor also helps to further improve the performance of Kingsoft Cloud's hybrid cloud storage system. This processor provides more and newer enterprise-level functions and can serve the requirements in hybrid clouds and data-driven applications. It also helps in improving daily operations. This multi-functional server platform brings a breakthrough in computing and I/O performance for compute-intensive and latency-sensitive applications.

By combining with the Intel® Optane™ SSD, TLC and QLC 3D NAND SSD product family, the storage platform built on the Intel® Xeon® Scalable platform can perform efficient management of massive data flowing through storage, cache, and memory to satisfy the various requirements in the data and cloud era.

Kingsoft Cloud and Intel worked closely together in creating Kingsoft Cloud's hybrid cloud storage products. Besides using Intel® processors, SSDs, other hardware products, as well as SPDK and other software technologies, they also collaborated to optimize the underlying layer of the hybrid cloud storage system

for actual storage application scenarios, further improving the performance of the SSD storage medium to satisfy users' performance requirements of the hybrid cloud storage system while also reducing TCO at the same time.

To verify the performance of Kingsoft Cloud's Intel® architecture-based hybrid cloud storage system, Kingsoft Cloud built test platforms based on the original configuration and the latest configuration to test the two platforms separately (the configurations for the two platforms are shown in Table 1):

	Configuration Details	Hosts
Original configuration	2 * Intel® Xeon® Silver 4110 processor, 128 GB RAM, dual-port 10 GB network card, 3.84T SATA SSD*8	5
New configuration	2 * Intel® Xeon® Silver 4210 processor, 128 GB RAM, dual-port 25 GB network card, Intel® SSD DC P4510 7.68 TB*8	5

Table 1. Test platform configurations including the original configuration and the upgraded new configuration

The results of the 4KB random read and write IOPS test (results shown in Figure 6) and the 4KB random read and write latency test (results shown in Figure 7) show that the performance of the overall storage system has improved by over 300% at this stage, and it is expected that it will eventually achieve up to 500% improvement in performance. The storage cost for the same storage capacity and the same performance has reduced significantly, providing enterprise users with higher cost

effectiveness as well as better storage performance and experience. Furthermore, Intel SSDs have outstanding performance consistency with very little IOPS variation due to differences in application environments. Such outstanding performance ensures that hybrid cloud storage systems are able to provide stable and high-performance storage services in various application scenarios.

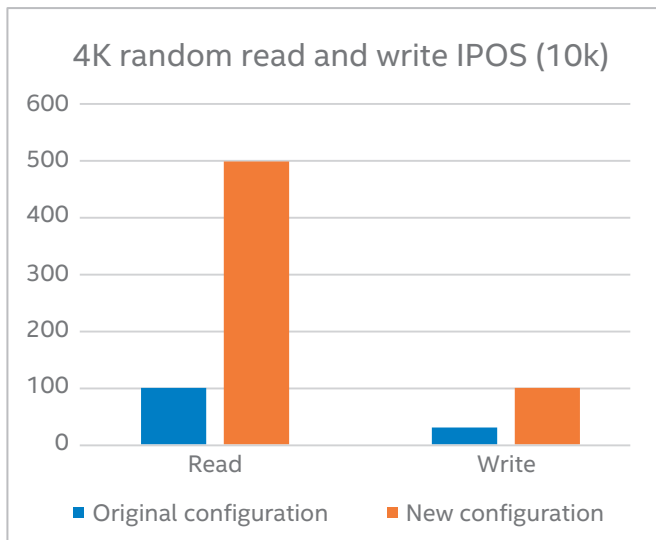


Figure 6. 4KB Random Read and Write IOPS Test (the higher the better)

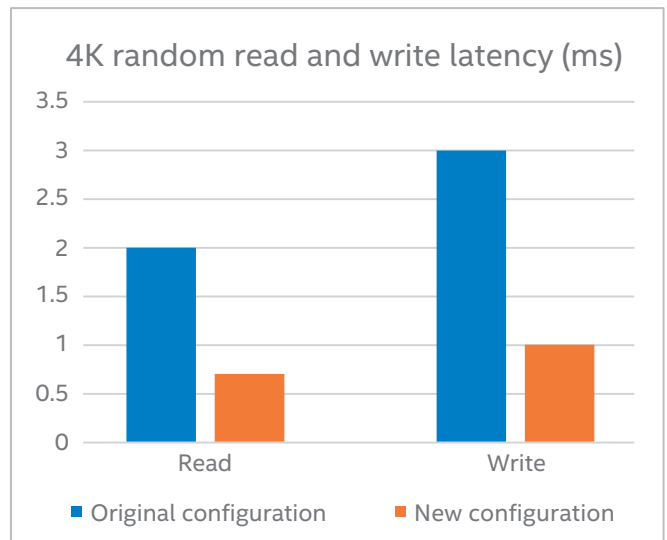


Figure 7. 4KB Random Read and Write Latency Test (the lower the better)

Kingsoft Cloud's Intel® architecture-based hybrid cloud storage system accelerates the digital transformation of enterprises

Thanks to the introduction of Intel® processors, SSDs and other products as well as software optimization technologies, Kingsoft Cloud's hybrid cloud storage products have reached the intended objectives and can help large and medium-sized enterprises build future-oriented hybrid cloud storage systems and support their digital transformation strategy. Specifically speaking, the deployment of Kingsoft Cloud's hybrid cloud storage system can help enterprises realize the following benefits:

- **High performance storage with support for traditional and new storage loads:** Kingsoft Cloud's hybrid cloud storage system is able to satisfy the data throughput performance and low latency requirements of traditional database applications as well as big data, deep learning and other applications. Furthermore, the variation in IOPS in different application environments is very small. This ensures that the enterprise's hybrid cloud platform can still provide high-speed and stable storage services even in high load conditions.
- **Elastic storage that helps enterprises satisfy service development requirements and control TCO:** With the capability for linear scaling, Kingsoft Cloud's hybrid cloud storage system is able to support enterprises in planning their storage system according to their needs so as to give play to the value of the original data and avoid the waste of resources caused by excessive construction investment. At the same time, the Kingsoft Cloud's hybrid cloud storage nodes applying the latest Intel products and technologies can be introduced in the future expansion, so that the latest storage technologies and developments can be utilized at any time.
- **Satisfy the data storage requirements of various types of businesses:** Kingsoft Cloud's hybrid cloud storage system product line includes block storage, file storage, object storage, and other products. It can satisfy different types of business requirements and is able to meet even the demanding requirements of big data, AI training tasks and other complex storage workloads via reasonable data lifecycle management policies.
- **High availability for the stable continuity of businesses:** Intel SSDs can guarantee a very low failure rate even under high load conditions. Combined with Kingsoft Cloud's troubleshooting, multi-data duplication, erasure code, and other such protection technologies, the availability of the system can be raised to a very high level. This not only ensures the continued operations of businesses, but also effectively reduces operation and maintenance costs.

At present, Kingsoft Cloud's enterprise hybrid cloud storage, based on Intel® hardware architecture, has been widely used in government, finance, healthcare, media, video, and the Internet, and has helped enterprises build hybrid clouds and accelerate digital transformation.

Taking a media company in the radio and television industry as an example. The company has accumulated an increasing number of video clips with higher and higher resolution during the development process of its new media business for film and television. The rapid growth of system data has presented demanding requirements in terms of massive storage, elastic scaling, clip distribution, storage interface, etc. Furthermore, concurrent access to massive files generated across dozens of workstations is required in the content gathering and content creation processes. A solution is urgently needed to solve the slow read and write speeds and lagging caused by subpar performance.

To meet the aforementioned challenges, this media company deployed the Kingsoft Cloud ES10000 Hybrid Cloud Storage solution. The solution provided 30PB raw capacity and no less than 10PB available storage space, which can sufficiently satisfy the storage needs of customers for various video clips, and can reliably support customers' future elastic scaling needs in storage capacity and performance. In addition, this solution also provided functions and user experience that are completely consistent with Kingsoft Cloud's self-developed KS3 object storage on the public cloud as well as professional service guarantees, which strongly promoted the company's business development.

Future Prospects

Intel and Kingsoft Cloud have, through powerful combination, created a more competitive hybrid cloud storage product. The two parties will, relying on continuous product and technology innovation, continue to engage in deeper cooperation on even more levels, and further unleash the performance potential of Intel's computing, storage, and networking products. The two parties are currently still continuing to cooperate based on the new-generation Intel SSDs and other products, and software optimizations to continue to improve the performance, availability, and other advantages of the hybrid cloud storage system, and to help Kingsoft Cloud accelerate its business in the field of hybrid cloud storage.

Not only does Kingsoft Cloud's hybrid cloud storage system have features such as low latency, high bandwidth, high IOPS, and high availability in its infrastructure layer, it also realizes centralized management of public clouds and enterprises' own local and edge

storage resources thanks to innovations in cloud architecture. It is also capable of meeting the storage requirements of AI, machine learning, and other new types of applications. We believe that as Intel and Kingsoft Cloud continues to work together more closely, the two parties will bring customers even more innovative products and solutions and hasten the unleashing of the power of data.

As data has become an enterprise's most valuable asset, Intel, Kingsoft Cloud, and other partners have joined together and will continue to strive to promote innovation and ecosystem partnership in storage technology, and create more optimized storage solutions to satisfy the data processing requirements of different levels, and drive, with storage innovation, the digital transformation of the enterprise into a new stage of development.

About Kingsoft Cloud

Founded in 2012, Kingsoft Cloud is a well-known independent cloud service provider in China. Its business scope covers multiple countries and regions around the world. Over the past eight years since its establishment, Kingsoft Cloud has always adhered to the customer-centric service concept, providing secure, reliable, stable, and high-quality cloud computing services. Kingsoft Cloud has already built a complete cloud computing infrastructure and operating system and, through the organic combination with artificial intelligence, big data, IoT, blockchain, edge computing, AR/VR, and other advantageous technologies, provided solutions that are applicable in areas of government, finance, AIoT, healthcare, industrial engineering, media, video, game, education, the Internet, and other industries.

About Intel

Intel (NASDAQ: INTC), is an industry leader, creating world-changing technology that enables global progress and enriches lives. Inspired by Moore's Law, we continuously work to advance the design and manufacturing of semiconductors to help address our customer's greatest challenges. By embedding intelligence in the cloud, network, edge and every kind of computing device, we unleash the potential of data to transform business and society for the better. To learn more about Intel's innovations, go to newsroom.intel.cn and official website intel.cn.



¹ <https://ark.intel.com/content/www/cn/zh/ark/products/series/122570/intel-ssd-dc-p4510-series.html>

² https://www.tmtpost.com/nictation/4254570.html?ulu-rcmd=0_061df_art_3_eddc3a41493d46498f4f2eac9eaa03e

³ <https://spdk.io/>

Intel does not control or audit third-party data. Please review the content, consult with other sources, and confirm whether the data mentioned are correct.

Intel technologies may require enabled hardware, software or services activation. For more information, inquire with the original equipment manufacturer or retailer.

In specific tests of component performance in a particular system, any difference in hardware, software, or configuration may affect actual performance. When considering purchase, please refer to other information for performance evaluation. For more complete information on performance and benchmark results, please visit: www.intel.com/benchmarks

Intel technologies' features and benefits depend on system configuration and may require enabled hardware, software or service activation. Actual performance may vary depending on system configuration. No computer system can be absolutely secure. Check with your system manufacturer or retailer, or learn more at intel.com

No computer system can be absolutely secure.

Intel makes no warranties, express or implied, including, but not limited to, the implied warranties of merchantability, fitness for a particular purpose, and non-infringement, and any warranties arising out of the performance process, the trading process, or trade practices.

For more information about performance and benchmark results, please visit: www.intel.com/benchmarks

Intel, the Intel logo, and other Intel trademarks are registered trademarks of Intel Corporation or its subsidiaries in the United States and/or other countries.

Copyright © Intel Corporation. All rights reserved