CASE STUDY

Intel® Optane™ Solid State Drive
Intel® QLC 3D NAND Technology
Intel® Xeon® Scalable Processor
High Performance Storage System

intel®

Redefining Object Storage to Create a High-performance All-flash Solution for AI, Big Data, and HPC Applications

Baidu ABC Storage leverages Intel® Optane™ SSD and Intel® QLC 3D NAND SSD technology to drive performance and capacity for today's demanding application

"Baidu AI Cloud expects its high performance all-flash object storage solution to help private cloud users tackle the challenges posed by massive unstructured small files. The combination of Intel® Optane™ Solid State Drives (SSD) and Intel® SSD based on Intel® QLC 3D NAND technology has helped our solution yield optimum results in terms of stability and Input/Output Operations Per Second (IOPS)."

**- Baidu AI Cloud
ABC Storage Team**

Advanced technologies, such as Artificial Intelligence (AI) Training, Big Data Processing, and High Performance Computing (HPC), are driving the direction in development of private cloud storage services. Storage systems for massive data are also closely intertwined with enterprise needs, especially in the area of high-performance storage systems for massive quantities of unstructured small files. As a leading enterprise in IT and in the internet industry, Baidu AI Cloud* applied its years of experience in public cloud storage technologies to a private cloud storage solution as a crucial component in its ABC (AI, Big Data, Cloud) Strategy. Through its partnership with Intel, Baidu AI Cloud* employed a combination of SSDs with Intel® Optane™ technology and Intel® QLC technology for the core hardware of ABC Storage's all-flash object storage solution.

Data Growth—Opportunity and Challenge

The volume of worldwide data is expected to swell to 163 ZB (Zettabytes) by 2025.[1] Massive data, especially with the explosive growth of unstructured data, has become a driving force for the digitization of enterprise data, as well as the rapid and continued evolution of related IT technologies. This amount of data is expected to enable breakthroughs in technologies, such as computer vision, speech recognition, and financial risk control. Thus, effective management, processing, and utilization of massive data has become a key area of competitiveness for enterprises wishing to maintain an edge in their industries.

However, the storage of massive unstructured data creates challenges for traditional storage systems due to file size and quantity, indexing, accessing patterns, and legacy storage technologies (i.e. spinning drives). Additionally, block storage and file storage systems are not ideal for small file storage, while AI and other new applications demand higher requirements for storage systems in terms of read/write performance. These present interesting technology challenges.

File Size and Quantity—The performance of traditional file storage systems tends to be volatile and declines with the rapid increase of file quantities. In AI training scenarios, such as image recognition, the training datasets incorporate astounding file quantities, typically of small file size. Likewise, for popular internet applications, such as Media Asset Management, unmanned vehicles, and video services, the file quantities stored and processed in the system usually reach hundreds of millions. The rapid increase of file quantities results in the decline and volatility of IOPS performance in storage systems, especially in traditional file storage, such as Network Attached Storage (NAS) systems.

Indexing—In addition, file storage systems currently use Hash tree and B+ tree computing methods to manage and index directories. The algorithms used to manage and index directories tend to significantly decline in efficiency and performance when retrieving from directories containing over 100 million files.

Accessing—In certain application scenarios, "Read Once, Write Many" or "mixed read/write" access modes further exacerbate the challenges in terms of performance. Common file I/O processes comprise "open", "search", "read/write" and "close" operations. "Open" before "read" or "write" take up the most system time and resources. As such, when handling "mixed read/write" access modes, the system repeatedly executes "open" operations. When there are massive concurrent operations, a huge amount of the system's resources will be wasted and result in performance loss.

HDDs—The weaknesses of traditional HDDs in terms of IOPS and random read/write performance have hindered the performance upgrades of storage systems. Due to mechanical limitations, even the higher-performance HDDs only have IOPS figures in the hundreds for random read/write performance.[2] When processing small files, the efficiency is even lower, as the HDD is required to continuously search for and locate the files at different storage locations.

Baidu ABC Storage's High-Performance, All-Flash Storage Solution

Baidu has gained widespread recognition for its work in the area of search technologies. With over 100 billion pages, 2,000 Petabytes (PB) of data stored, and 100 PB of data processed per day,[3] Baidu is well-versed in the technological challenges brought about by the storage of massive unstructured small files.

Baidu AI Cloud has attempted to tackle the above challenges through software improvements and Intel-based hardware enhancements.
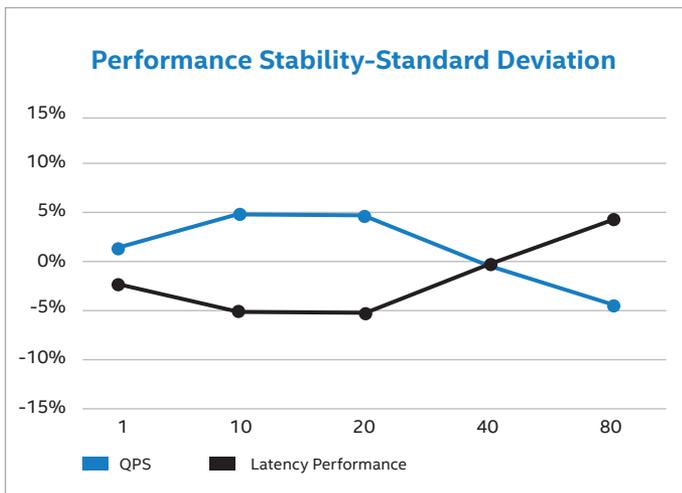
Figure 1. The performance stability test results of the ABC Storage object storage solution under Baidu AI Cloud

Software

Developers incorporated Baidu's high-performance object storage engine into the new solution, thereby enabling it to deliver great data life cycle management, data protection strategy, retrieval efficiency, InfiniBand* Architecture network and RDMA support, and flexible rights management mechanisms. Additionally, by leveraging flat deployment for object storage, high-efficiency retrieval and Exabyte scalability, the ABC Storage high-performance object storage engine is able to provide private cloud users with storage of massive unstructured small files.

An AI training process comprises data collection, cleaning and labeling, resizing, modeling, training, evaluation, and prediction. Each step requires the storage system to perform read, write, and retrieve operations. Throughout the training, the data will be subjected to high concurrency and iterative throughput, so as to provide sufficient data to train the system for full-load operations.

Baidu's object storage engine solves performance issues with massive files, enabling storage systems to achieve stable performance output and effectively boost the data utilization efficiency of AI applications. Meanwhile, for certain mixed read/write operations during training, the engine also performs further optimization to ensure that the system performance is unaffected under mixed read/write scenarios.

Testing results of various optimizations show that the software alone is able to maintain stable performance throughout with increasing file quantities. As shown in Figure 1, the Query Per Second (QPS) and latency performance fluctuated within a 5 percent[4] range as file quantities gradually increased from 100 million to 8 billion.

Hardware

As described above, HDDs present several challenges for high-performance storage solutions. SSDs have virtually no seek time or rotational latency, thereby resulting in high IOPS performance compared to HDDs. Baidu AI Cloud uses a combination of Intel® Optane™ SSD and Intel® QLC 3D NAND SSD technology to make up the core hardware for the ABC Storage all-flash object storage solution. Intel Optane SSDs feature innovative Intel® 3D XPoint™ storage media and incorporate advanced system memory controllers, interface hardware, and software technology, delivering low latency and high stability. The Baidu solution uses the following devices:

Intel® Optane™ SSD DC P4800X is deployed in core storage system areas, such as the cache, MDS, and log system. This device offers up to 550,000 IOPS of random read/write capacity and less than 10us of read/write latency,[5] enabling the solution to perform more effectively in multi-user and high-concurrency scenarios. Meanwhile, its drive writes per day (DWPD) performance also provides a longer lifespan and ensures better economic value.

Intel® SSD D5-P4320, based on QLC technology, offers large capacity data storage. Intel's 64-layer 3D NAND technology enables a single QLC SSD disk capacity of up to 7.68 TB in order to adequately fulfill the storage requirements of massive data. It also has a random read IOPS of up to

427,0007, and, when paired with the Intel® Xeon® Gold 6142 processor, it is especially suitable in terms of meeting "Write Once, Read Many" (WORM) performance requirements in application scenarios, such as AI training. The Intel SSD D5-P4320 used in the new solution effectively meets the requirements for large storage capacity.

In the ABC Storage solution, each storage server is deployed with four SSDs, which provide a total file storage quantity of up to 2 billion 15 KB files in 30 TB of capacity. More importantly, the price/performance ratio of the Intel QLC 3D NAND SSDs has enabled this combination of SSDs to ensure the high performance of this solution while effectively lowering the Total Cost of Ownership (TCO) for the system. Baidu testing has shown that the Baidu AI Cloud high performance all-flash solution could lower TCO by 60 percent.[6]

Results

With the support of Intel, the Baidu AI Cloud team carried out a detailed evaluation and measurement of the performance of the ABC Storage all-flash storage solution. Figure 2 shows the benchmark test framework, which includes a cluster made up of five servers with each server configured with two Intel® Xeon® Gold 6142 processors and 256 GB of memory. One 750 GB Intel Optane SSD DC P4800X and four 7.68 TB Intel SSD D5-P4320 drives were used. The system used a 40 GbE network to connect to the computing platform.

Testing showed that the combination of the Intel Optane SSD and Intel 3D NAND QLC SSD technology adequately meets the storage system performance requirements for AI training application scenarios. Table 1 shows the performance results of the basic ABC Storage version.
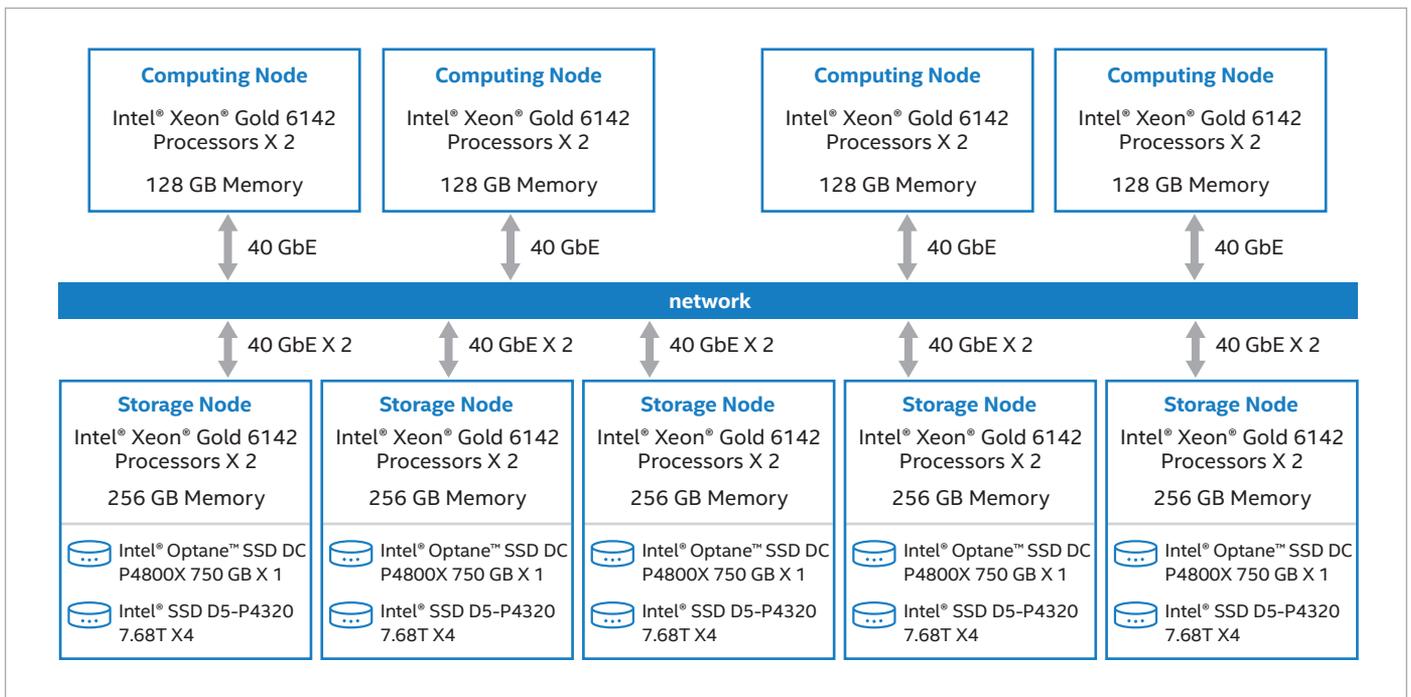
Figure 2. The Benchmark Test Framework for ABC Storage's All-flash Storage Solution

Packet Size (KB)	Data Operation	Throughput Rate (packet/second)	Throughput Rate for Every Storage Node	Average Response Time (ms)
4	Write	33929.8	6785.96	14.6
8	Write	33150.9	6630.18	14.944
16	Write	34158.5	6831.7	14.504
4	Read	210760	42151.9	2.366
8	Read	199679	39935.8	2.496
16	Read	187384	37476.9	2.66

Table 1. Benchmark performance test results for ABC Storage's all-flash storage solution[4]

Future Prospects

As one of the crucial practical outcomes of the Baidu AI Cloud ABC strategy, the ABC Storage high-performance all-flash object storage solution has provided strong and reliable support for private cloud application scenarios, such as AI training, big data analysis, and high-performance computing, with its improved storage performance and storage size.

Intel's products and technologies are crucial factors in the success of the solution. In the future, both parties plan to embark on more partnerships to optimize the performance of the existing solutions, while incorporating more of Intel's products and technologies. Meanwhile, both parties also plan to extend the all-flash high-performance object storage solution to more application scenarios to truly convert massive data into a driving force that will propel the transformation of the development of IT technologies and the digitization of enterprises.

The Advantages of the Baidu AI Cloud Solution

- The ABC Storage high-performance object storage engine provides an integrated object storage interface for application scenarios, such as AI training and high performance computing, thereby providing stable performance output even with a rapid increase in file quantities.

- With targeted optimization processes, the ABC Storage high-performance object storage engine helps storage systems maintain good performance, whereby "read/write", WORM and "mixed read/write" scenarios are required for massive data.

- The combination of the Intel® Optane™ SSD and the Intel® SSD based on Intel® QLC 3D NAND technology enables the ABC Storage all-flash object storage solution to maintain high performance, while drastically reducing TCO.

(intel)