HUAWEI

Huawei OceanStor
All-Flash Storage

Issue 2

Huawei Storage

Powers the Digital World

Contents

Leading Data Infrastructure to Advance the Digital Economy
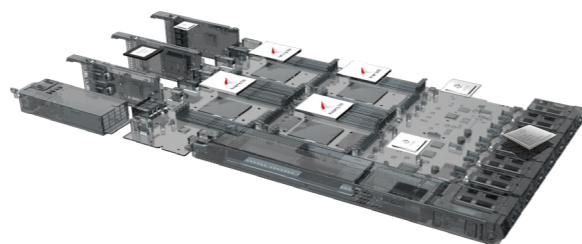
Leading Data Infrastructure to Advance the Digital Economy

By Peter Zhou, President of the Intelligent Data and Storage Domain of Huawei

"

Economies have developed and evolved throughout history. First, there were agricultural economies and then industrial economies. Today, we are entering the era of the digital economy. The global digital economy is growing 2.5 times faster than global GDP. Every dollar invested in digital technologies has a return 6.7 times higher than non-digital investments. China's digital economy reached CNY 31.3 trillion in 2018, accounting for 34.8% of the national GDP. The digital economy contributes over 50% of GDP growth. It is estimated that by 2025 the digital economy will be USD 23 trillion globally.

Data is the key factor to driving economic growth. Leaders in developing and innovating data infrastructure cannot only survive but thrive in the future. Predictions state that digital data will grow from 33 ZB in 2018 to 180 ZB by 2025. However, the drawbacks in our current technologies mean that less than 2% of the generated data is properly stored, and less than 10% of the stored data is fully utilized. Exponential data growth requires a large number of storage and computing resources. Limited enterprise storage resources cannot meet ever-increasing data volumes. This gives enterprises a hard time efficiently locating, fetching, and using data, seriously hindering

their development.

Enterprises in an intelligent world need a reliable and stable data infrastructure platform for full-lifecycle management of data storage, computing, management, and usage. This platform must integrate diverse IT hardware and software to help enterprises effectively store, compute, analyze, process, and use data. Most importantly, it must give enterprises a better understanding of their business and unlock data value for streamlined production management, efficient mission-critical applications, and stable operations services. Where there's data, there's value.

The goal of the data infrastructure industry is to achieve the best cost and value per bit through the entire data lifecycle. Huawei has launched its data infrastructure strategy for the Kunpeng computing industry, which integrates core database, big data, and storage capabilities to provide a converged, intelligent, and open data infrastructure. Huawei also aims to work with partners to build an open and prosperous Kunpeng ecosystem for the intelligent world by implementing its Platform+Ecosystem plan and achieving open hardware and open source software.

"

Peter Zhou

President of the Intelligent Data and Storage Domain of Huawei

HUAWEI

1 Setting Out

Luminaries' Insights into All-Flash Industry
The Time for Huawei Next-Gen OceanStor AFAs

Huawei Storage Ranks 3rd Globally
in Revenue with Growth of 34%

Finance: Innovation Pioneer in **Bank 4.0**

Carrier: The Optimum Data Infrastructure
for the **5G Era**

Behind Huawei OceanStor All-Flash Arrays' Success

Building Industry-Leading
Data Infrastructure for the Digital Economy

By Guo Tao, China Cloud News

Infrastructure is both the foundation and prerequisite for economic and social development. China is currently transitioning from a period of rapid economic growth to that of high-quality development across the board, in which such infrastructure as roads, railway lines, and bridges are being scaled up and upgraded. Meanwhile, a parallel process has been initiated in the digital economy; data infrastructure is also undergoing a major transformation, a phenomenon that marks the start of the "intelligence+" era.

The advances of the digital economy complement growth in the real economy. According to the White Paper on China's Digital Economy Development and Employment (2019) released by China Academy of Information and Communications Technology (CAICT), China's digital economy has been developing rapidly. In 2018, the volume of China's digital economy reached CNY 31.3 trillion (an increase of 20.9% year-on-year), accounting for 34.8% of GDP. Solid information infrastructure is required to ensure that such rapid and high-quality development within the digital economy is sustainable in the long-term. Amidst this period of digital transformation, Chinese enterprises also demand efficient, innovative, and secure information infrastructures, above all else.

! Support for the Digital Economy from the Ground Up

Information infrastructure construction is vital to both individual enterprises, and the industry at large. The real question is what type of infrastructure this should be. According to Peter Zhou, President of the Intelligent Data and Storage Domain of Huawei, data infrastructure will form the solid basis of the entire digital economy.

Data is at the root of data infrastructure, and a well-designed infrastructure enables enterprises to efficiently process, manage, and analyze data throughout the entire lifecycle. An infrastructure that possesses open and dynamic features, and continually integrates new technologies for evaluation, can offer better assistance to enterprises who are learning to cope with digitalization. In the "intelligent+" era, computing and intelligence are everywhere. Therefore, the data-based infrastructure should be improved, on both the inside and outside, by supporting intelligent data applications. When data is static or unused, it is unable to benefit enterprises. Many enterprise systems will be saturated by the sheer accumulation of data. Therefore, data infrastructure plays an essential role in extracting unforeseen value from the data.

Extending data infrastructure is no longer a simple matter of stacking the various elements, such as computing, storage, and network. Instead, it is a platform that integrates diverse resources from data centers, and gradually extends its function to data applications. The more value created by the platform, the better the infrastructure will be.

As a pioneer in the field of data infrastructure, Huawei has worked tirelessly to enhance its comprehensive infrastructure solutions, encompassing construction,

operation, and management, among other capabilities. Huawei aims to play a leading role in the data infrastructure industry, by providing users with intelligent and easy-to-use data solutions that grow the entire digital economy. Huawei has also extended its original storage service line to the field of intelligent data and storage, integrating new technologies and products, with the goal of providing innovative, end-to-end intelligent data solutions. In order to build this digital ecosystem, Huawei is committed to working with a wide range of partners around the world.
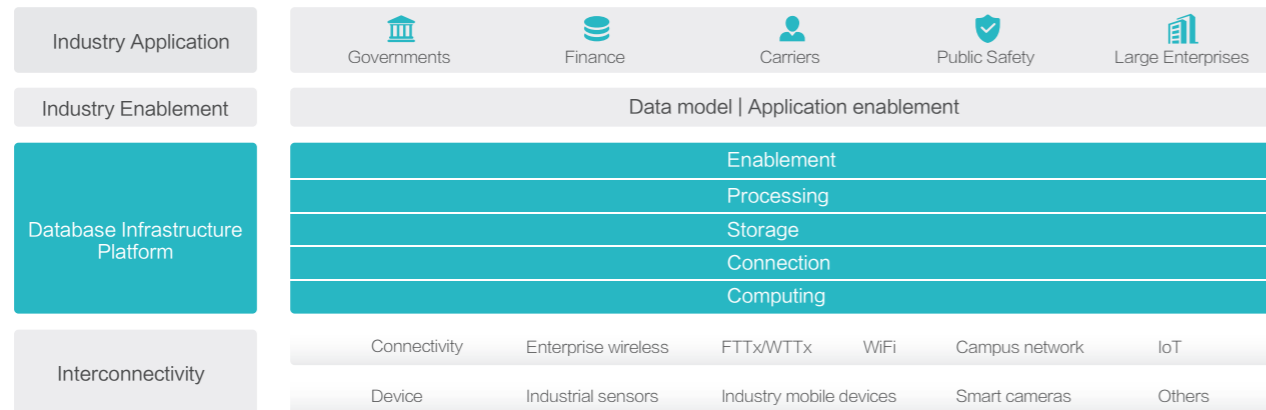
Industry Application	Governments	Finance	Carriers	Public Safety	Large Enterprises
Industry Enablement	Data model | Application enablement				
Database Infrastructure Platform	Enablement				
	Processing				
	Storage				
	Connection				
	Computing				
Interconnectivity	Connectivity	Enterprise wireless	FTTx/WTTx WiFi	Campus network	IoT
	Device	Industrial sensors	Industry mobile devices	Smart cameras	Others

Fig. 1 Data Infrastructure

Seamless Transformation

Huawei has transformed itself from a leader in the storage industry to a pioneer in the field of data infrastructure seamlessly. Storage is an indispensable carrier of data, and a key part of the data infrastructure. In the past, both vendors and users only paid attention to technology itself. Therefore, they believed that the process of purchasing, deploying and applying storage devices is complex. On the application side, Huawei has focused on improving data transfer, storage, analytics, and management throughout the entire lifecycle. Huawei is also capable of delivering data infrastructure as a comprehensive solution, which shortens the implementation process, and leads to greater customer satisfaction.

Huawei has redefined the nature of data infrastructure in the following five areas: computing, access, storage, processing, and data advancement, not only by leveraging its technical expertise and experience in computing and storage, but also in introducing new products, such as GaussDB, which provide customers with end-to-end intelligent data solutions. These capabilities make it easier to deploy and apply data infrastructure, supporting the wholesale digital transformation of industry.

This transformation not only made new products and solutions available; it has also changed the overriding perception that products are the only things of value. This shift in understanding has empowered vendors, who previously focused solely on products, to deliver services and solutions that support diversified applications as well. Instead of just purchasing products, users are now able to take service requirements into consideration, in building new data infrastructures, providing an impetus for the cloud-based transformation and digital services.

An increasing number of enterprises have undergone a major transformation in the field of information infrastructure, and users have benefited tremendously from the integration of Data+Intelligence. For example, in Beijing, Data+Intelligence enables one-stop government services, with the data doing all of the legwork, improving services efficiency by 50% in the process. In Shenzhen, Data+Intelligence turns the reactive "cars wait for the lights" (fixed signal cycle lengths) into the proactive "lights watch over cars" (flexible traffic signal timing), speeding up the traffic on the main road by 18% during morning peak hours. These are just a couple of the myriad of examples indicating how Data+Intelligence has helped vendors reinvent themselves, and how it has helped enterprises to accelerate their digital transformation.

The End of the HDD Era

Innovations always come with new challenges. Emerging technologies, such as cloud computing, big data, software-defined storage, and artificial intelligence (AI) have facilitated the evolution and transformation of storage, and thus, storage needs to be redefined. Previous transition periods were focused on only the technologies or products themselves, but in this instance, as more is known about data infrastructure, a shift in our common understanding of storage is needed as well. This means an end to HDD, and the full-scale implementation of SSDs.

Ever since the emergence of flash memory, it has been recognized as a game-changer for storage, with the ongoing trend of HDDs being replaced by flash memory. However, this transformation has not been as fast as initially expected. In light of continual technological innovation, evolving service scenarios, and optimized products, at long last, the time is ripe for "All-Flash for All".

Huawei has redefined the storage architecture to cope with challenges, such as low storage efficiency and complex management, changing it from a single type of storage to converged storage, forming the basis for converged processing. This indicates the enormous gains in efficiency achieved in applying all-flash technology.

Huawei has the leading market share for all-flash storage in China, and the leading growth rate for global all-flash storage. More than 3,000 enterprises around the world make use of Huawei's all-flash storage solutions. The launch of OceanStor Dorado intelligent all-flash storage represents a milestone in the realization of the "All-Flash for All" concept.

Huawei OceanStor Dorado all-flash storage comes equipped with high-performance Kunpeng 920 processors, a balanced and reliable SmartMatrix fully-interconnected architecture, and AI-based full-lifecycle intelligent management, to set a new storage benchmark for intelligent data solutions.

Huawei's long-term technical expertise in chip R&D is evident in the Kunpeng 920, network chips, management chips, SSD chips, and Ascend AI chip 310, which helps OceanStor Dorado all-flash storage deliver ultimate performance (20 million IOPS and 0.1 ms latency) for highly-responsive real-time operations.

Huawei's storage architecture is also second-to-none. OceanStor Dorado all-flash storage utilizes a SmartMatrix fully-interconnected architecture, which is capable of handling concurrent 7-controller failure without service interruption in any single system, serving as a model for high-accessibility.

AI-powered chips lead the way. OceanStor Dorado is the first-ever intelligent storage system with built-in Ascend AI chips, providing semantic association from machine learning to improve the read cache hit ratio by 50% for enhanced storage efficiency. In addition, edge-cloud AI synergy implements intelligent lifecycle management and further builds an automated data management system with a complete three-layered architecture, reducing OPEX.

Vendors engage in Flash Storms. However, discount promotions are not an effective way to attract users to an all-flash storage solution. Therefore, Huawei has not simply reduced the price of flash storage, but has also worked tirelessly to assist users in reducing the total cost of ownership (TCO) during the end-to-end construction, application, and O&M of intelligent data infrastructure. This emphasis on value is the essence of Huawei's "All-Flash for All" concept.

Jiang Libin, who works in the national customs information center of the General Administration of Customs, explained the impact of Huawei's solutions on the agency's work, "As the core application for customs, the clearance system needs to run 24/7. The system receives more than 300,000 customs declarations every day. Therefore, it has high requirements on data synchronization, disaster recovery, and backup. After using Huawei's OceanStor Dorado active-active solution to upgrade the original system, the reliability and performance have improved substantially, achieving second-level service recovery, and boosting data read and write speed by three times. Huawei has continuously innovated its chip and storage technology to launch the next-generation of OceanStor Dorado all-flash storage, which is faster, more stable, and more automated than its predecessors. It provides a new way to define the data infrastructure. We also look forward to further working with Huawei to build a reliable and smart customs data infrastructure platform, leading the way in customs technology innovation."

The "All-Flash for All" era is coming.

Huawei Storage Ranks 3rd Globally in Revenue with Growth of 34%

By Zhang Cha, IT We-Media

According to Gartner's latest market report, unlike the bleak x86 server market, the enterprise-class storage market is still growing, which is all the more notable given the sluggish growth of the global economy as a whole.
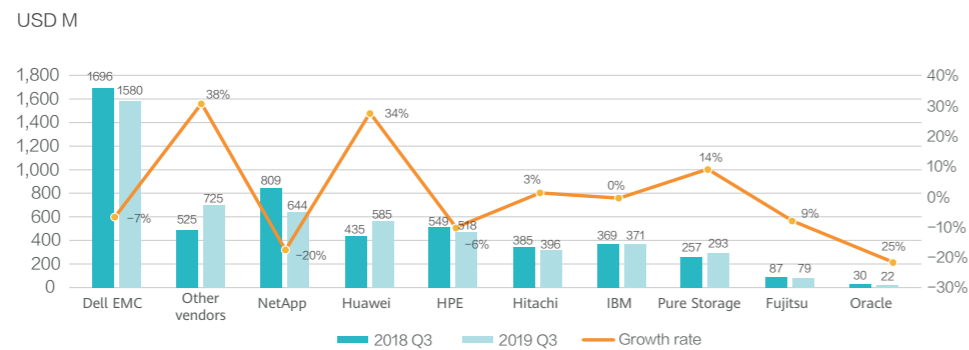
USD M

Source: Gartner's global storage market share report for Q3 2019

Fig. 1 Global storage vendor revenue ranking in Q3 2019

In this growing storage market, Huawei has gained prominence. **Its storage revenue has increased by 34% year on year, which is more than any other vendor. In Q3 2019, Huawei Storage ranked first globally in terms of growth rate and third in terms of storage revenue.**

Looking for a Fulcrum to Overtake at the Corner

For any organization looking to achieve leapfrog development, overtaking at the corner is a good strategy. However, accelerating at the corner under the influence of a huge centrifugal force can also cause us to spin off the track. Therefore, when we overtake at a corner, we need to find a solid fulcrum. Huawei is one of the best examples of this.
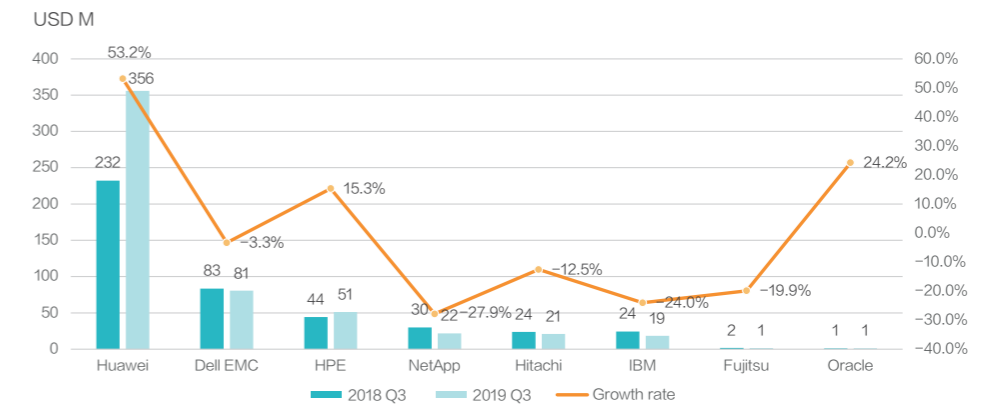
USD M

Fig. 2 Storage vendor revenue ranking in the Chinese market in Q3 2019

According to Gartner's report, in Q3 2019, Huawei's storage revenue grew by 53.2% year on year, accounting for a 43.7% market share. Furthermore, in the fast-growing all-flash field, Huawei is a pioneer, accounting for 57% of the total market. Chinese Mainland, Taiwan (China), Hong Kong (China), and Macao (China) together contribute 62% of Huawei Storage's revenue.

Source: Gartner's global storage market share report for Q3 2019

Fig. 3 All-flash vendor revenue ranking in the Chinese market in Q3 2019

From data, the conclusion is clear: Huawei's dominance of the Chinese storage market is secure, in both the larger storage market and the smaller but promising all-flash market. Furthermore, the success in the Chinese market will help Huawei attract more customers in the global storage market. In other words, Huawei has found a solid fulcrum it can use to overtake at the corner.

All-Flash as the Corner-Hitting Point

Now it has its fulcrum, Huawei Storage needs to consider the best way to approach the corner in order to quickly surpass its competitors.In Q3 2019, Huawei's all-flash revenue increased by 89.4% year on year. This is a remarkable improvement which underpins the company's overall storage market growth. The all-flash market is clearly the corner-hitting point Huawei needs to utilize in order to surpass its competitors.
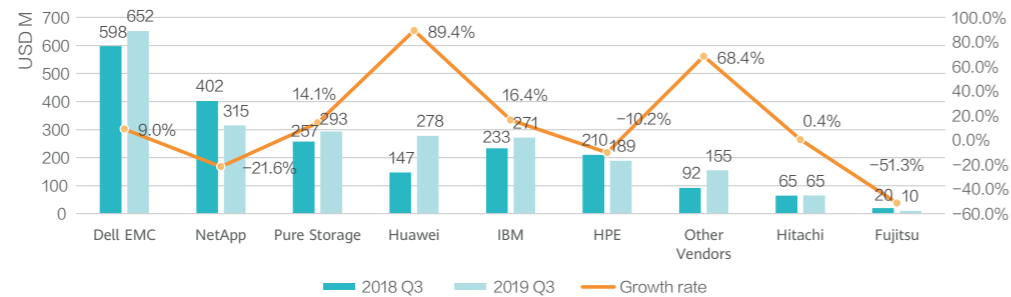
Source: Gartner's global storage market share report for Q3 2019

Fig. 4 Global all-flash vendor revenue ranking in Q3 2019

Currently, AFAs account for 50% of the storage market in North America and only 30% in China. With the economic growth and continuous rapid digital transformation in China, the all-flash market share will gradually approach and surpass that of North America. This means that the overall scale of the all-flash market in China will have a growth space of more than 20% in the next few years. With its strong market appeal and profound technical accumulation, Huawei Storage has a promising growth space. All-flash storage is the ideal "corner-hitting" point for Huawei to surpass competitors.
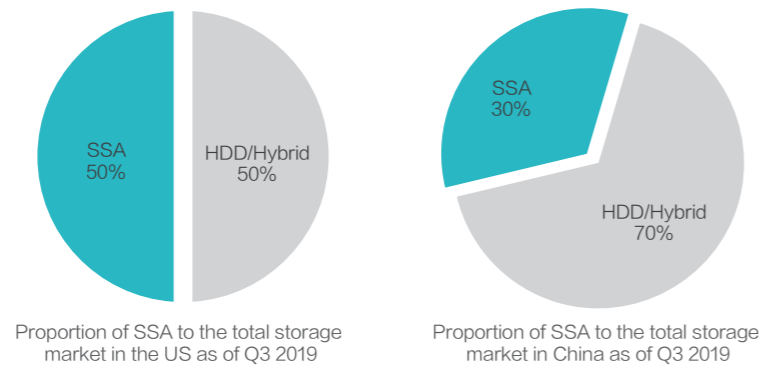

Proportion of SSA to the total storage market in the US as of Q3 2019

Proportion of SSA to the total storage market in China as of Q3 2019

Fig. 5 Storage market comparison: China and the US

Succeeding Together with Customers

Huawei's research into flash storage began in 2005. The company's long-term technical accumulation and massive investment not only enable it to fully understand SSD, but also establish it as an early mover in the R&D of main control boards and give it a deep understanding of protocols. Since 2011, Huawei has released multiple generations of all-flash products, set multiple performance records, and set a new benchmark for the storage industry.

> In the future, with the support of emerging technologies such as 5G and AI, data volume will increase exponentially. Rapid response will become the primary reason for users to choose all-flash. It is an exciting time in the development of the all-flash industry.

Finance:
Innovation Pioneer in Bank 4.0

By Li Bisheng, Huawei Senior Financial Architect

New technologies transform industries. The financial industry in China is gradually transforming from Bank 3.0 (mobile devices replacing physical branches) to Bank 4.0 (banking everywhere). Bank 4.0 aims to popularize modern FinTech services. It focuses on technological innovation to transform entire processes with technological applications, optimize service processes, and promote business innovation.

Next-generation IT technologies, such as cloud computing, big data, artificial intelligence, and blockchain continuously promote FinTech application, innovation, and transformation in the financial industry. However, the Internet era has placed new demands on the technical architecture of the financial industry:

1. Build energy-efficient data centers with automatic and intelligent O&M to improve the maturity of operation management capabilities.

2. Apply new technologies for infrastructure transformation and upgrade.

3. Foster cooperative innovation and resource sharing between banks and across industries with national strategies, such as Internet+ and big data.

The IT data center is a vital infrastructure for financial services. It carries key production core, front-end, payment, card exchange, mobile banking, e-banking, counter, and credit card systems. The IT data center must ensure 24/7 uninterrupted services, provide risk control plans in real time, and eliminate data silos between different data platforms.

IT infrastructure at a data center consists of servers, networks, and most crucial of all, storage devices. Bank 4.0 expands banks by integrating resources into one platform, generating revenue from retail finance, and prioritizing energy efficiency.

Integrating resources from multiple systems into one platform unifies services, promotes data sharing among departments, and greatly enhances operational efficiency. Integration requires shared storage to provide outstanding performance, low latency, and hundreds of thousands of IOPS to host applications on multiple systems, such as core production, online banking, front-end, and management support, meet highly concurrent system requests during peak hours.

The shared storage system must feature high scalability and use an elastic architecture to cope with the unpredictable explosive growth of services in this new era of million-level IOPS.

Such a large-scale platform must leverage cutting-edge technologies to stably provide 24/7 service operations even under heavy workloads.

Retail finance is the primary generator of revenue in Bank 4.0. Emerging financial services, such as Internet finance, credit loan, and third-party payment, have turned the traditional financial industry upside down. The overall profit growth of listed banks has been slowing down and single-digit growth of net profits is increasingly common. More banks are realizing that extensive operations are outdated and that new business models are necessary for sustainable development. Large state-owned and joint-stock banks are using retail finance to start their transformation. They concentrate their efforts on providing platform-based services with a focus on improving user experience. Mobile apps have become vital Internet financial products. Gaining mobile banking users and maintaining their loyalty are keys to determining market performance. As a result, banks require IT data centers to:

1.Quickly develop new Internet financial products and marketing platforms: Banks have IT teams to develop new services and carry out device O&M. New services must be thoroughly verified before rollout, with more than thirty verifications. Consequently, IT data centers must quickly build development and testing environments without wasting space or affecting production services.

2.Respond instantaneously to order placements: Helping users quickly complete transactions is crucial to financial retail profits. IT data centers continuously boost speeds and bank profits with millisecond-level storage responses by using leading edge technologies, such as NVMe over Fabric and SCM.

3.Use real-time user analytics and precise user profiling: Marketers often cast a wide net, but only a fraction of users will place orders and even a smaller fraction will become regular customers. Exploring prospective opportunities and turning first-time buyers into loyal customers require powerful capabilities for sales analytics. In the past, this required years of hard work and strenuous analytics. Today, technological advances and new IT configurations, especially in storage media performance, generate timely and accurate user profiles for commercial banks of all sizes. Banks can use real-time analytics of tens of millions of retail users and information sharing between banks and industries to explore retail opportunities from hundreds of millions of potential users.

Energy efficiency requires automatic and intelligent O&M, improved operational management for data centers, and reduced maintenance manpower. Additionally, environmentally friendly IT data centers must develop sustainable data infrastructure and economy and efficiently deal with the explosive data growth in Bank 4.0 with new technologies to improve storage space utilization and slash energy consumption in power and air conditioner cooling.

Flash media meets all of these requirements. Currently, all-flash arrays (AFAs) are used to accommodate core transaction services at financial data centers. In the future, AFAs will have a variety of applications, from core applications to common production services. They will bring important benefits to virtualization and cloud computing environments, such as low latency, quick responses, constant and scalable performance, high reliability, and low TCO.

Huawei Storage embodies independent innovation with a full series of storage systems using Kunpeng processors. One of the best examples is the next-generation OceanStor Dorado AFA released in Beijing on July 23, 2019. It provides 20 million IOPS, uses five intelligent chips developed by Huawei, and delivers unwavering reliability. It tolerates the failure of one out of two controller enclosures, setting a new benchmark for AFAs and driving the storage industry towards realizing the idea of "Ever Fast, Ever Solid, and AI-powered".

The differentiated competitive strengths of FinTech are crucial to success in the financial industry. Huawei provides fully developed and reliable storage products and solutions to help customers in finance achieve technological transformation, placing them at the forefront of financial digital transformation in China.

Carrier: The Optimum Data Infrastructure for the 5G Era

By Du Yumin and Zhao Zengzhu, Architects of Huawei Carrier Solution

5G is ON. 5G can be deployed twice faster than 4G. For example, a tier-1 carrier has 36,000 4G base stations. In 2019, the carrier purchased 5G spectrum through auction and plans to deploy 13,000 5G base stations in 2020. 5G, cloud, AI, and blockchain have been widely recognized in the industry as the new infrastructure to lead ICT development over the next 10 years. As an important part of the new architecture, data and storage are also ushering in great development opportunities.

Carrier Data Trends

5G has caused the centralized data layout of traditional infrastructure to evolve to a two-end model. Now, the business support system (BSS)/operations support system (OSS)/management support system (MSS) is deployed at the national/international data center (DC). Network functions virtualization infrastructure (NFVI), B2B, and B2C services are centralized for lower operations costs, while being moved downward to the edge for low latency.

To reduce duplicate data for cost effectiveness, carriers are beginning to deploy the BSS/OSS/MSS data scattered across the provincial, municipal, and county levels to the national/international DC. In addition, to further reduce capital expenditure (CAPEX), some low-value services will be directly deployed on the public cloud.

Rapid BSS Service Growth Due to Large Amounts of Billing Data

1.The average coverage radius of 5G base stations is 10–20% that of 4G base stations, meaning that up to 10 times more base stations are needed to cover the same area. Huge investment forces operators to build shared infrastructure. In addition to traffic- and time-based billing on 4G networks, 5G also supports QoS billing.

2.In the 5G era, the bandwidth increases by more than 10 times, and the number of Person to Machine (P2M) and Machine-to-Machine (M2M) terminals increases sharply. For example, in Germany, the number of M2M connections is estimated to increase from 500 per square kilometer to 4,500, which leads to the increase of terminals that require billing.
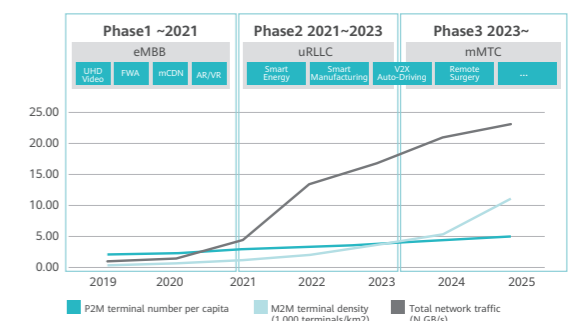
Fig. 1 5G terminal quantity change

Carriers have considered multiple simplified billing solutions, for example, traffic billing in the unit of 100 MB (instead of 10 MB) or complete machine-to-machine (M2M) monthly package. However, once 5G infrastructure and applications are well-interconnected, the increase in terminals and traffic will far exceed expectations. The performance, stability, and reliability of core systems needs to be improved for carriers.

Rapid OSS Service Growth Needs More Network Elements (NEs)

OSS services mainly grow from the increase in NEs. Take a tier-1 carrier subnet as an example. The subnet has four core network data halls, 900 transmission central office (CO) data halls, and 36,000 base stations, covering an area of 357,000 square kilometers. There are averagely 0.1 base stations per square kilometer.

1. In the first phase of 5G (eMBB), the carrier is expected to have 90 core network site addresses and 67,000 base stations — with an average of 0.2 base station per square kilometer.

2. In the second phase of 5G (uRLLC), the carrier is expected to have 900 core network site addresses and 148,000 base stations — with an average of 0.4 base station per square kilometer.

3. In the third phase of 5G (mMTC), the carrier is expected to have 10,000 core network site addresses and 336,000 base stations — with an average of 1 base station per square kilometer.
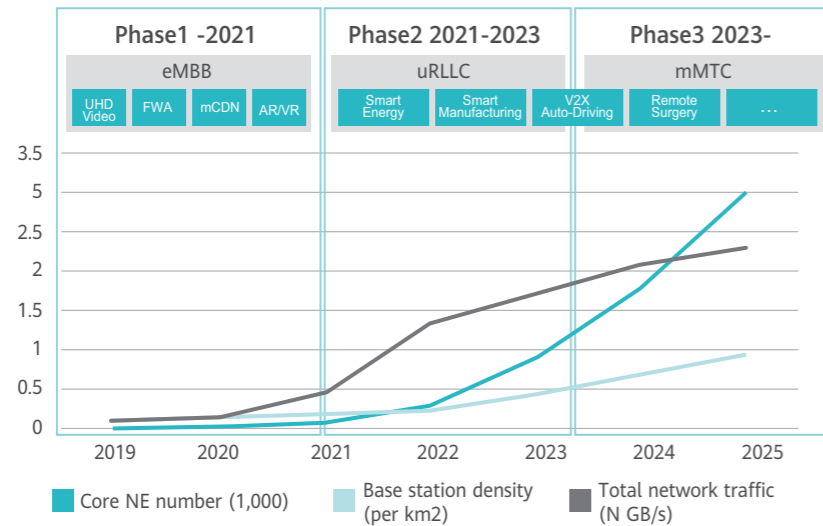

Fig. 2 Changes in 5G NE number

All-Cloud Mobile Office and IoT-based Asset Tracking for Rapid MSS Service Growth

1.5G makes cloud office possible, where data is stored on the cloud and accessible by employees using lightweight terminals anytime, anywhere. Centralized storage can greatly avoid duplicate data and save security software licenses, and a data disaster recovery solution effectively prevents data loss.

2.Operations involves the flow of a large number of materials and assets. Currently, most materials and assets are manually tracked, which will affect the reliability of data such as device locations, carried services, online time, and maintenance periods. As a result, some assets may be unavailable, and devices may exceed the maintenance period. This is a great waste of resources, and directly affects service rollout planning and adjustment policies. 5G makes mobile connection applications possible, where all assets can be tracked in real time using electronic labels, and the management databases can be updated in real time.

*BSS is used for billing and user management. The data volume is in direct proportion to the number of terminals and network bandwidth traffic.
*OSS is used for network management. The data volume is in direct proportion to the number of NEs.
*MSS is used for office management. The data volume is in direct proportion to the number of carrier employees and data consumption per employee.

Storage Architecture for 5G Data Centers

In the 5G era, the explosive growth of data volume increases the data service pressure on the central side. Data of the BOM, B2B, B2C, and NFVI domains all need to be moved to the cloud to quickly adapt to service requirements. Based on the importance of service systems, services in the data center are classified into five layers:

1. **Core applications:** production and transaction systems used by carriers themselves or rented by carriers to users. These include the billing and CRM functions of the BSS/OSS/MSS system, and the hot and warm layers of the B2B system (used for carrying core or high-performance applications of external users). NFV directly carries 5G core NEs. Its service is frequently accessed, meaning that in the event of unstable performance, user experience or internal process operation will be greatly affected. Therefore, performance stability and ultimate reliability are key requirements.

2. **Ultra-fast applications:** systems that require ultra-low latency, such as Internet of vehicles (IoV) and real-time decision-making

3. **Common applications:** OA, web service, AR, VR, and email

4. **High-capacity applications:** video and IoT. This type of service has lower requirements on latency and reliability, but the data volume increases sharply.

5. **Cold data applications:** media assets and archiving. These systems are not frequently accessed and do not have high performance requirements

All-Flash + Active-Active Solution for Ultra-fast, Stable Applications

SSDs yield a much more stable performance compared to HDDs. The average access latency of 10,000-rpm SAS HDDs is 5 ms, however, the performance difference between the inner and outer rings of the disk is more than three times. The latency of accessing different data blocks fluctuates between 3 ms and 10 ms. However, all-flash storage, covering host ports and disks, can use the E2E NVMe architecture to ensure that the latency of any data access can be controlled within 0.5 ms in a single system and within 1 ms in an active-active system. The delay per access is reduced from 5 ms to 1 ms, which means that the report output time for every 10 million devices is reduced from 10 hours to 3 hours.

Carriers need ultra-reliable core systems. They require 99.9999% service availability — always-on services during device upgrade, capacity expansion, and troubleshooting.
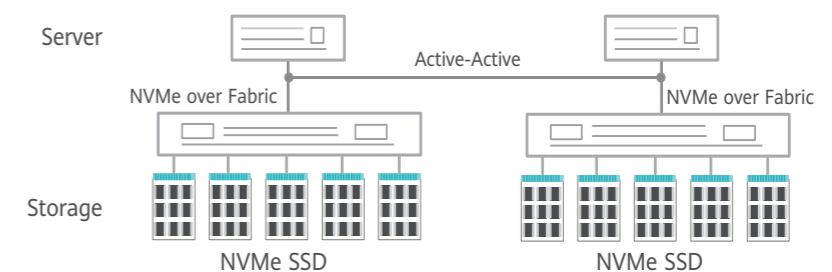

Fig. 3 Target architecture of core services

Therefore, the target architecture of the core system is the high-end all-flash, active-active solution.

1. E2E NVMe is supported from host ports to disks, ensuring stable and low latency.
2. The upgrade, capacity expansion, and migration are all online. Services are not interrupted.
3. The reliability of a single system is 99.999%, while that of an active-active system is 99.9999%.

Single Network: Building All-Ethernet Data Center

Currently, data centers have two types of networks: Fibre Channel and Ethernet. Core structured data is transmitted over the FC network, and other data is transmitted over the Ethernet network. The switches and optical modules differ between the two. In general, the FC network has two to three times more CAPEX per Gb/s compared to the Ethernet network. Deploying two types of network systems in a data center increases network complexity and doubles O&M costs.

RDMA over Converged Ethernet (RoCE) uses UDP instead of TCP to transmit data and supports direct memory access through converged Ethernet, ensuring stable and low latency for services. NVMe over Fabric abandons the SCSI protocol, and uses the flash-native interface to reduce the latency of access from the host to the storage by 50%. Therefore, NVMe over RoCE can provide performance and latency equivalent to FC, and ultimately can replace the FC network.
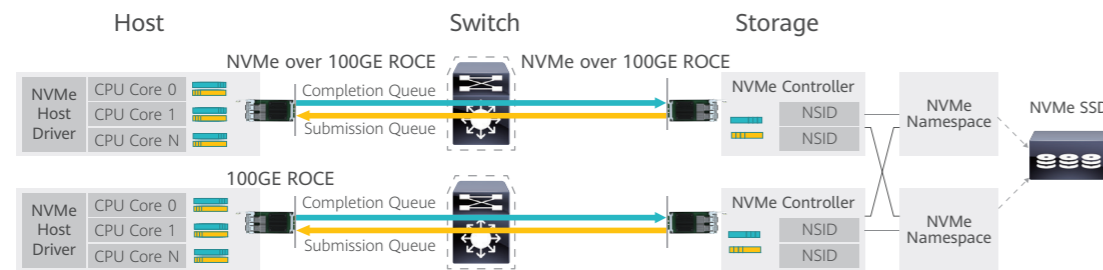

Fig. 4 NVMe over 100GE RoCE

To simplify the central IT architecture, Huawei aims to build an all-Ethernet data center. The evolution path of a core data network is FC > NVMe over FC > NVMe over 100GE RoCE. This evolution ensures a stable and low latency of the flash-native port for users, and a smooth transition from the FC network to the Ethernet network. With technology advancements, the block, file, object, and big data protocols evolve to iSCSI, NFS/CIFS, S3/Swift, and HDFS, respectively. In the future, NVMe over RDMA will be used to carry multi-form applications, providing ultimate performance and simplifying DC management.
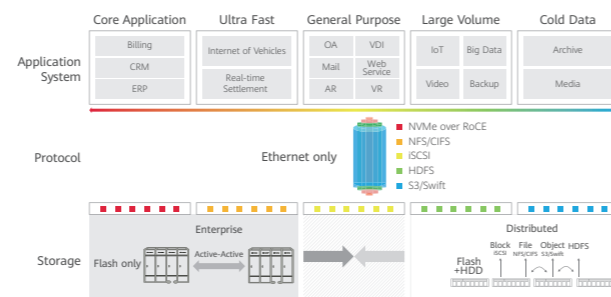

Fig. 5 Single DC, single network

5G is changing our lives in ways we never imagined. All-flash is the smart choice for enterprises deploying data infrastructure in the 5G era.

Behind Huawei OceanStor All-Flash Arrays' Success

By He Zhi, Huawei

The latest IDC report shows that Huawei's all-flash arrays (AFAs) ranked No. 1 in the global revenue growth rate in the first three quarters of 2019, with a growth rate of 234%. In China, Huawei Storage has ranked No.1 for four consecutive years by seizing 33.4% of the market share, while also occupying a leading share in carrier and financial markets.

Standing out of the competition for Huawei's AFAs is by no means pure luck. It comes from continuous innovation in product design and core technologies with more than 3,600 R&D engineers working in 12 storage R&D centers around the world producing innovative, premium products. Huawei can independently work on storage operating systems, controllers, and SSDs from end to end, and carry out end-to-end in-depth optimization specific for flash storage. "Huawei has accumulated over 14 years of know-how in flash, starting from as early as 2005," Lin Yu, chief architect of Huawei OceanStor Dorado AFAs, remarked, "Building from scratch, our flash storage products have gradually moved from common services to core applications in high-end storage for commercial use."
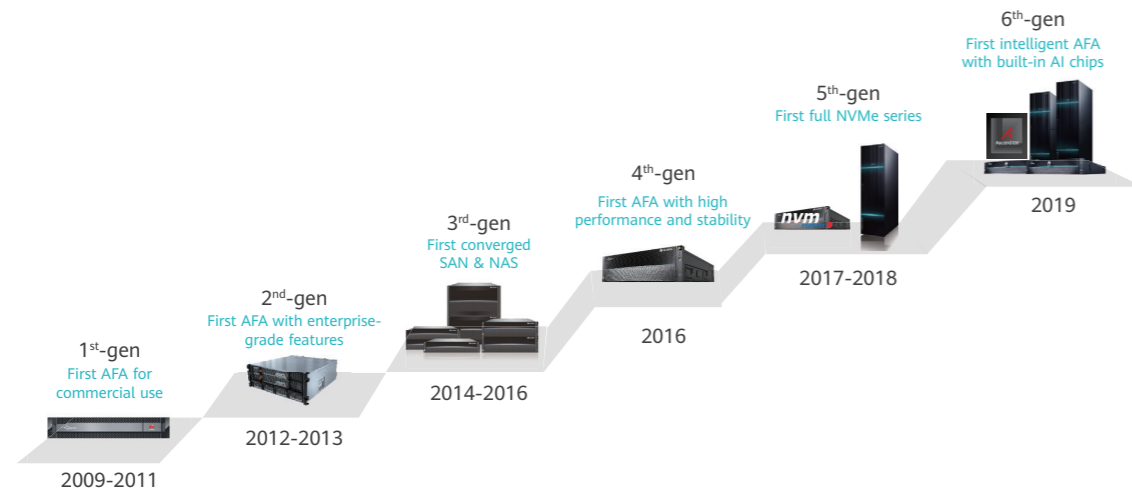
Overtaking on a Corner

"2008 was a critical year for Huawei's AFAs, and the success back then propelled us to develop all-flash technologies we use today." Lin Yu added. 2008 was the year Huawei Storage decided to prioritize flash storage to overtake competitors and break the HDD monopoly of traditional storage vendors from the mid-2000s. "In 2008, Silicon Valley was home to many startups that employed flash storage in enterprise-grade applications. We were amazed at the superior performance, power consumption, and reliability SSDs yield over HDDs. "Overtaking on a corner" is one concise yet powerful phrase summarized by one of our leaders. This phrase reflects Huawei's strategic storage thinking and important opportunities to build a competitive market dominated by traditional established vendors. It also inspired many R&D team members, myself included, to contribute in the flash field."

Recruiting talent became one of the most pressing tasks along with unified strategic thinking. Many storage experts, including Lin Yu, were tasked to set up R&D teams outside China, and also recruited top SSD experts from outside China to establish Huawei SSD R&D teams. "We were inspired by having the chance to work with the industry's top technical experts. Thanks to their insights, we've overcome many difficulties during the R&D process,"

she said. During this time, these R&D teams developed many high-value technical patents regarding SSD wear leveling, garbage collection, and ECC verification, and soon released the first SATA SSD boasting industry-leading performance and reliability. "Huawei's first self-developed SATA SSD is still fresh in my memory because of what it represented – a premium product that overtook legacy industry's benchmark — Intel products, in performance tests. It was a significant step forward for us, and helped strengthen our innovation in all-flash R&D."

6th-gen
First intelligent AFA with built-in AI chips

5th-gen
First full NVMe series

4th-gen
First AFA with high performance and stability

2019

3rd-gen
First converged SAN & NAS

2017-2018

2nd-gen
First AFA with enterprise-grade features

2016

1st-gen
First AFA for commercial use

2014-2016

2012-2013

2009-2011

Step One for Commercial Use of Flash Storage

Flash storage became more readily available in 2011 because many storage vendors launched their own flash storage solutions to accelerate tier-0 applications. Emerging storage vendors such as Violin had made a breakthrough in the mainstream database-storage acceleration field with ultra-high performance.

Lin Yu said, "By 2011 we had experience in commercial use of SSDs and storage arrays, and thanks to in-depth understanding of flash media and enterprise storage products and delivery, we removed many processes and features of the HDD architecture that cannot make the most of SSDs. Further, we were able to minimize storage system latency by combining storage arrays with SSDs and leveraging key technologies such as cache and scheduling optimized for flash features in storage arrays. This may not seem that impressive now, but it was groundbreaking at the time."

Thanks to the brand-new all-flash architecture designed by the storage teams, the very first-generation OceanStor Dorado AFA was launched and equipped with self-developed 6 Gbit/s SAS SSDs. This was the industry's first batch of all-flash products available for commercial use. The second-generation OceanStor Dorado AFA system was released in the following year. "Dorado is the fastest fish in the ocean. With this in mind, we used this name to stress the blazing-fast performance it can bring to users." OceanStor Dorado AFAs began to play its major role in the ocean of data for the very first time.

Converged Flash Storage for the Era of Fully Converged Data Architecture

In 2014, the business concept of convergence across five dimensions was successfully promoted and implemented. Among them, the convergence of SSDs and HDDs is an important feature for hybrid storage. Lin Yu remarked, "Because the price of SSDs was still high at that time, storage vendors gradually used hybrid storage instead of all-flash storage to accelerate tier-0 applications. Although we understood how SSDs work and knew which access modes yield optimal performance and SSD service life when it came to optimizing flash applications, we still didn't have much technical experience in tiered storage and SSD cache."

To keep up with competitors, the storage R&D teams set up a technical breakthrough research project focusing on hot and cold data identification and elimination algorithms. This was so successful that it delivered SSD cache and tiered storage

features within just one year. Huawei Storage can provide customers with an industry-leading solution that combines SSDs and HDDs, so customers can enjoy high performance and low latency of SSDs while massively reducing their OPEX.

In May 2014, Huawei released the third-generation OceanStor hybrid flash storage, unveiling an era of fully converged data architecture. At the same time the following year, Huawei released the OceanStor 18000 mission-critical converged storage, setting a benchmark for high-end storage with its superb stability, excellent performance, and future-proof innovation.

Technological Innovation, Setting a Benchmark in the Industry

In September 2016, the fourth-generation OceanStor Dorado AFA equipped with 12 Gbit/s SAS SSDs was launched, which features excellent performance that can power gateway-free active-active storage solutions. Lin Yu stated, "The fourth-generation OceanStor Dorado underwent a major architecture transformation for end-to-end I/O processing." This new generation allows I/Os to be written in ROW mode, uses large block sequential write, and ensures consistently stable performance even when snapshot and cloud features are enabled. The key to SSD reliability is to reduce write amplification, which is greatly different from the reliability requirement of HDDs. The new generation ensures the AFAs aggregate small I/O requests in log write mode to avoid RAID write amplification. This combined with wear leveling, patented anti-wear leveling, and innovative disk-controller collaboration algorithms makes the fourth-generation OceanStor Dorado AFA a top-notch product in the industry.

"Adjusting the write mode means that all features of the storage pool will be modified accordingly. During the entire delivery process, we experienced a hardship in optimizing garbage collection performance, constantly upgrading disk-controller collaboration algorithms, and ensuring stable system performance in the event of faults," Lin Yu recalled. The main functions of the fourth-generation had been developed in May 2016. After real-life tests on customers' existing network, performance fluctuation was controlled within 5%, a consistency that ranks top in the industry.

The fourth-generation OceanStor Dorado is the only AFA with outstanding performance and stability, setting a benchmark for the flash storage industry.

One Step Further with NVMe

Lin Yu continued, "From 2016 to 2017, all-flash technology matured at a steady rate, and now all-flash storage is the development trend of the entire storage industry. To make full use of all-flash capabilities, the matured NVMe technology has been adopted at unprecedented growth. Fortunately, we started our technical preparations three years ago, which enabled our technology team to use the stable architecture of our fourth-generation OceanStor Dorado, while maintaining our core strengths in NVMe and scalability. In March 2017, Huawei was the first vendor to release its NVMe-supported fifth-generation OceanStor Dorado AFA. It is industry's fastest AFA with a latency as short as 0.3 ms and can power the HyperMetro gateway-free active-active solution to ensure fast and stable operations of core applications."

Ever Fast, Ever Solid, AI-Powered Storage, a New Benchmark for AFAs

In July 2019, the sixth generation of OceanStor Dorado intelligent AFA was officially released. As a high-end storage product developed by Huawei's global R&D teams that are operating with over a decade of experience in the AFA field, this product has made many breakthroughs in the storage architecture. It is the first storage product to integrate five Huawei-developed chips that, combined with algorithms, can provide intelligent O&M throughout data lifecycle. "Our next-generation OceanStor Dorado embodies our ten years of continuous exploration and innovation in all-flash technology. I am proud to witness all of these," said Lin Yu.

"In terms of performance, we have built an end-to-end acceleration platform using five in-house developed chips, including Kunpeng 920 processor. This platform uses the end-to-end NVMe architecture and Huawei's unique FlashLink® intelligent algorithm to achieve 20 million IOPS and 0.1 ms latency, twice that of the next-best player, a benchmark for storage performance."

"We have designed the SmartMatrix fully interconnected architecture to deliver ultimate reliability to handle data growth in the 5G age. This architecture can tolerate the failure of 7 out of 8 controllers, the highest possible reliability in the industry while ensuring zero service interruption. Additionally, it can ensure ongoing services even in the event of failure of 1 out of 2 controller enclosures. This next-generation OceanStor Dorado delivers always-on core applications for enterprises and stands apart from peer products in terms of storage reliability."

"We have added many AI technologies to this next-generation AFA. For example, we have embedded the AI chip into storage devices to proactively analyze and understand the I/O rules of multiple application models, making storage faster than ever. Smart SSD enclosures are now able to process difficult workloads such as data reconstruction offloaded from storage controllers. General-purpose AI and customized edge AI work together for intelligent management throughout the product lifecycle. The success of Huawei's intelligent chips and edge synergy in the storage industry are testimony to the efforts made over the past decade."

The sixth generation of ever fast, ever solid, AI-powered OceanStor Dorado intelligent AFA outperforms other competitors in every aspect, delivering best-in-class user experience.

Conclusion

Huawei OceanStor AFAs have caused a stir in the industry. Now, Huawei ranks No.1 in the Chinese AFA market, and is also enjoying the fastest growth rate in the global market. This is the result of Huawei storage teams' hard work and dedication over the years.

These are the secret behind Huawei AFAs' success. For Huawei, it is a marathon, not a sprint, and as such will remain dedicated to continuously making efforts in flash storage (speed, stability, and AI) to provide users with more efficient, reliable, and easy-to-use storage. Quality speaks volumes in the storage industry, and it's no wonder why many new customers and partners are embracing the all-flash era with Huawei.

Building a fully connected, intelligent world

Huawei OceanStor All-Flash Storage

Ever Fast, Ever Solid, AI-Powered

2 Moving Forward

Huawei OceanStor AFAs
New Benchmark for Storage

Five-Chip Tech Accelerates End-to-End
Data Processing for Ever Fast Performance

The SmartMatrix Fully-Interconnected Architecture
A New Reliability Benchmark for High-End Storage

Huawei OceanStor Storage **Ecosystem with Openness**

Five-Chip Tech Accelerates End-to-End Data Processing for Ever Fast Performance

By Zheng Wei, Huawei

In the digital era, large data volumes are the norm. 5G and the internet of things (IoT) have driven an exponential increase in the amount of data processed. According to forecasts from IDC and Huawei, the volume of data will soar from 33 ZB in 2018 to 180 ZB in 2025. In order to face this flood of new data, enterprises will need to be armed with faster and more robust data processing and storage capabilities.

Huawei has accumulated substantial technical expertise, and continuously invested in pioneering chip technologies to develop the OceanStor Dorado, a next-generation all-flash storage system, and the most powerful SSD-based all-flash solution.

For the OceanStor Dorado all-flash storage system, Huawei has built an end-to-end chip platform by vertically integrating transmission, compute, intelligence, storage, and management. The platform incorporates five different types of Huawei-developed chips:
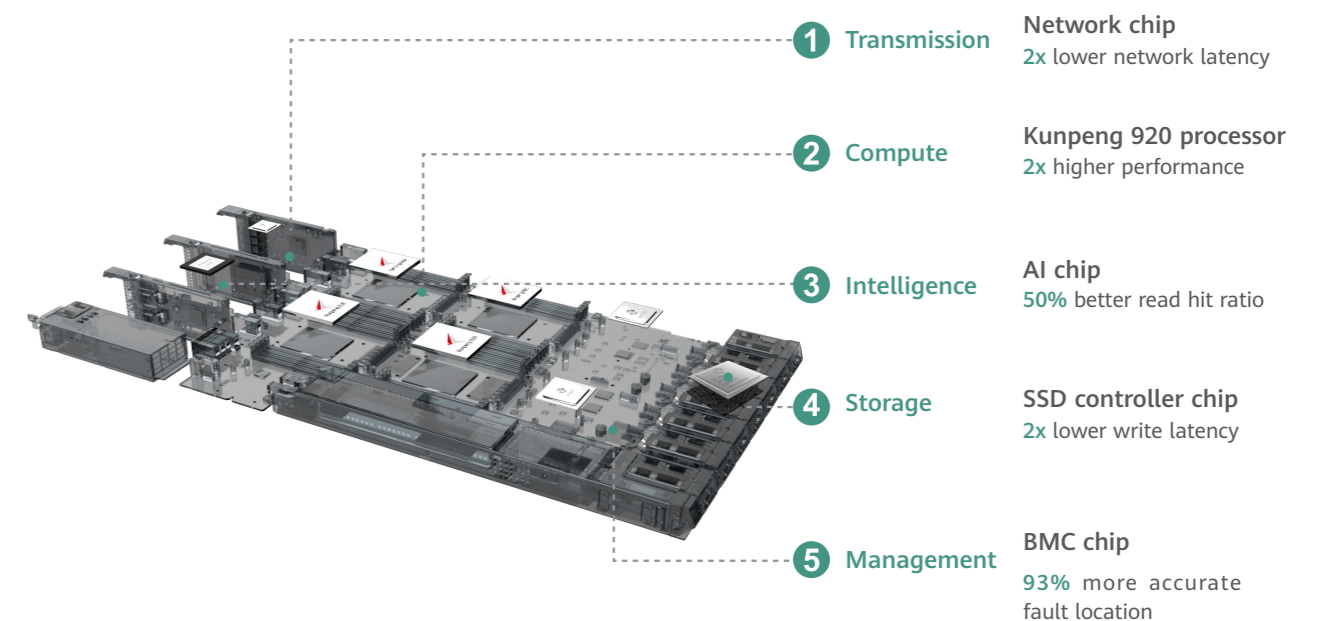

1 Transmission — **Network chip** / **2x** lower network latency

2 Compute — **Kunpeng 920 processor** / **2x** higher performance

3 Intelligence — **AI chip** / **50%** better read hit ratio

4 Storage — **SSD controller chip** / **2x** lower write latency

5 Management — **BMC chip** / **93%** more accurate fault location

Fig. 1 Overview of five chips

① Intelligent multi-protocol interface chip: hosts the protocol parsing previously performed by the general-purpose CPU, expediting front-end access performance by 2x.

② Intelligent processor chip, Kunpeng 920: utilizes the powerful CPU, delivering 25% more computing power than the industry average, and doubling the performance of a single controller.

③ Intelligent AI chip, Ascend 310: actively analyzes and understands the I/O rules of multiple application models, based on machine learning frameworks, to implement intelligent prefetch of memory space. This improves the read cache hit ratio by 50%.

④ Intelligent SSD controller chip: hosts the core flash translation layer (FTL) algorithm, accelerating data access within SSDs and providing 50% lower write latency.

⑤ Intelligent BMC management chip: contains a built-in Huawei storage fault library, which accelerates component fault diagnosis and location, with 93% accuracy.

Now, let's take a look at how OceanStor Dorado's next-generation all-flash storage is able to achieve such powerful performance by utilizing these five chips.

Intelligent Multi-Protocol Interface Chip Accelerates Front-End Read and Write

Data is transmitted from servers to storage arrays over the network. The first data station is the front-end interface module. A front-end interface module can be regarded as an air intake system; it separates data from the packet, and sends the data to the CPU for subsequent processing. The more powerful the front-end module, the more it can handle, which means that the smaller the portion of the CPU that needs to be allocated. This frees up the CPU for other important tasks, improving overall efficiency in the process.

In fact, the intelligent multi-protocol interface chip is the result of this pursuit. Huawei OceanStor Dorado all-flash storage adds a turbo to this air intake system for acceleration. It uses intelligent multi-protocol interface chips to host protocol stack functions, such as checksum and Fibre Channel (FC), which are originally implemented by the CPU based on software logic, meeting the application requirements for high concurrent I/Os, high throughput, and low latency.

According to recent test results, under the test models (7:3 read/write, 8 KB I/Os data blocks), Huawei OceanStor Dorado all-flash storage reduces data access latency from 160 μs to 80 μs, accelerating front-end access performance by two times, when compared with common front-end interface modules.

In addition, there are mainstream front-end modules, such as 8/16/32 Gbits/s FC and 1/10/25/40/100GE, in which each front-end module supports only a single protocol, preventing flexible conversion, and resulting in wasted resources. In order to use the front-end modules more efficiently, Huawei applies intelligent multi-protocol interface chips that integrate FC/IP protocol interfaces, and supports cutting-edge FC-NVMe and NVMe over RoCE.
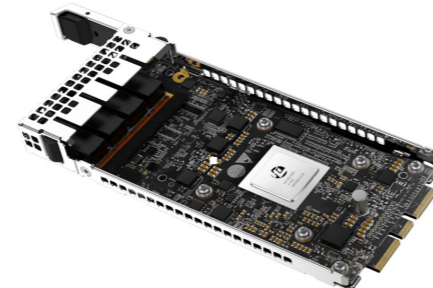
Fig. 2 Intelligent multi-protocol interface chip

Customers can integrate data hosted over IP, FC, and FCoE into one interface chip, greatly improving network flexibility, simplifying network management, and reducing network construction and maintenance costs. Better yet, Huawei is the first enterprise to support the front-end and back-end NVMe protocols in the entire series of OceanStor Dorado. Access latency is reduced to 100 μs, fully meeting the requirements of the most demanding transaction services, and equally significant, customers can select from a wide range of technologically advanced options, based on their budget, to host key enterprise services.

Intelligent Processor Chip Kunpeng 920 Doubles Processing via Many-Core Technology

The second station for data is the CPU. If we regard front-end interface modules as the turbo of a car engine, which greatly increases air flow speed, the CPU serves as the car engine, powering rapid data processing for the storage system.

How then can processing performance be optimized? First, the engine performance needs to be improved. The Kunpeng 920 processor scored 930 points in the SPECint benchmark, blowing past mainstream CPUs, for a record 25% increase in computing performance. Then, the number of engines needs to be increased. OceanStor Dorado provides a maximum of 768 CPU cores per controller enclosure, an industry-leading figure.

However, it is not an easy task to use such a high number of cores to maintain continuous and stable output performance. Typically, when the number of cores increases, it does not result in a linear increase in performance. This is because the CPU needs to do a lot of things, and different tasks have different processing flows and methods. The CPU frequently switches between different tasks, leading to significant waste, amidst the increased performance.

Huawei provides the intelligent many-core technology to resolve this longstanding issue. The intelligent multi-core technology works like the concept of an assembly line, popularized by automaker Henry Ford. Different "assembly lines" are established for each corresponding task. More cores are assigned for the more important tasks. For example, a read and write task is more important than a space reclamation task, therefore, the line that processes the read and write task will be assigned more cores.

What does intelligence refer to? First, it refers to an intelligent distribution algorithm that is configured to properly arrange tasks according to the progress of the job on each line. This algorithm tries to evenly distribute tasks to each line, in order to implement load balancing services. Second, it refers to the built-in intelligent scheduling algorithm of a CPU, which dynamically schedules core resources based on the core usage. This ensures that the 48 cores in a CPU can work in load balancing mode, and maximize the computing power. For example, there are 12 cores that are performing read and write I/O tasks. However,
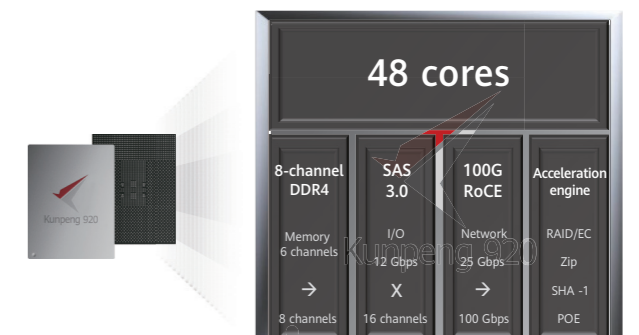

Fig. 3 Intelligent processor Kunpeng 920

a large number of read and write tasks queue up, while the load for the other eight cores, which perform the task of flushing cache to disks, remains relatively low. In this case, the intelligent scheduling algorithm temporarily schedules two out of the eight cores for read and write tasks, and reallocates the remaining six cores for load balancing. This maximizes the power of every core.

In summary, a single Huawei Kunpeng 920 processor provides for elite performance, and utilizes many-core technology to achieve linear performance growth as the number of controller nodes, CPUs, and cores is increased. This results in doubled performance for customers.

Kunpeng processors are deployed, not only on the next-generation OceanStor Dorado all-flash system, but on smart SSD enclosures (Huawei-innovated) as well, which enables SSD enclosures with computing capabilities. Therefore, reconstruction operations can be processed by the Kunpeng processors in smart SSD enclosures, achieving an ultimate reconstruction speed of up to 15 minutes/TB, 20 times higher than the industry average.
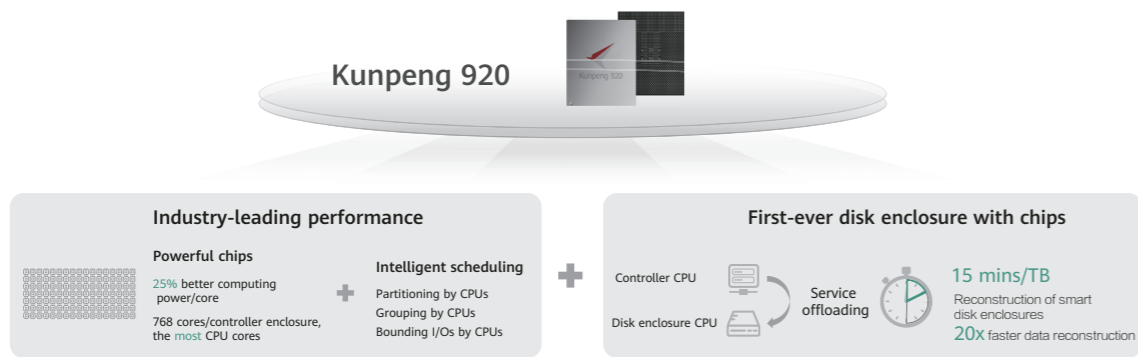
Kunpeng 920

| Industry-leading performance | First-ever disk enclosure with chips |

Powerful chips
25% better computing power/core
768 cores/controller enclosure, the most CPU cores

+

Intelligent scheduling
Partitioning by CPUs
Grouping by CPUs
Bounding I/Os by CPUs

+

Controller CPU
Disk enclosure CPU

Service offloading

15 mins/TB
Reconstruction of smart disk enclosures
20x faster data reconstruction

Fig. 4 Kunpeng 920 intelligent algorithm

Intelligent Ascend 310 Boosts Read Hit Ratio by 50%

Even with a powerful engine, an efficient and intelligent automatic gearbox is still needed. The intelligent AI chip Ascend 310 is developed to implement automatic acceleration, representing the first AI chip developed by Huawei specifically to boost computing efficiency. It is also capable of intelligently adjusting cache allocation based on service I/Os, improving the read hit ratio by 50%.

Ascend 310
FP16: 8 TeraFLOPS

Semantics-related I/Os with machine learning
10000001
1010101010

Read cache hit ratio
50%

Proactively analyzes and understands I/O rules of multiple application models based on machine learning frameworks
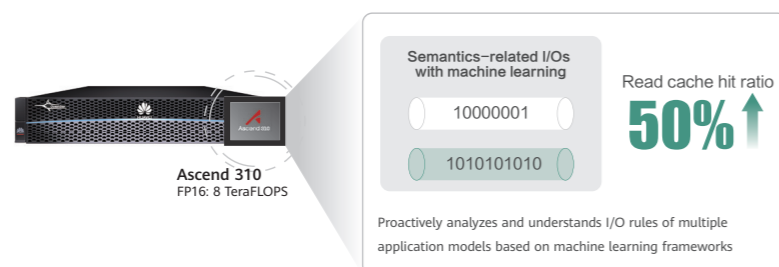
Fig. 5 Ascend 310 + intelligent algorithm

The Ascend 310 also helps improve forecast accuracy. The storage system receives many pieces of data, and each piece is called an I/O. The I/O can vary according to the corresponding service. However, there are also correlations between I/Os, including time correlation, space correlation and semantic correlation. Time correlation reveals the sequence of time, and space correlation refers to the relationships between spaces, such as Russia and Mongolia's location to the north of China. Semantic correlation finds associations, such as that between these two phrases: "The polar bear is the...", and "The biggest bear in the world". Such correlations are found with the goal of improving data prefetch accuracy.

The Ascend 310 offers a self-tuning deep learning algorithm, which is capable of quickly analyzing and deeply mining all of the I/O data of upper-layer services based on time, space, and semantic correlation. When an I/O arrives, the chip immediately identifies the data to be accessed, and instructs the CPU to quickly prefetch the data to the read cache. In addition, it continuously learns the existing data in the background to further improve accuracy. Lastly, the chip evaluates key performance indicators, such as the prefetch hit rate, waste rate, and latency, and then makes intelligent adjustments.

According to the test results, the read cache hit ratio of OceanStor Dorado increases from 19% to 69%, under the test model of 100% random reads and I/O block size within 64 KB.

2x Faster Write Speed with the Intelligent SSD Controller Chip

Data is stored on SSDs after passing through the front-end interface modules and CPUs. An SSD works like the fuel tank in a vehicle, storing all of the production data and power sources. Therefore, it requires high performance and stability. If a system can accurately control the fuel-injection quantity based on comprehensive factors, such as weather, distance, and vehicle condition, it can implement intelligently-informed acceleration, thereby prolonging the service life of the vehicle.

Fig. 6 Intelligent SSD controller chip

As data carriers, SSDs also require a highly-accurate control system. An SSD consists of a control unit (including SSD controllers and DRAM) and an NAND flash storage unit. The control unit is responsible for data read and write. The FTL algorithm stores the mapping between the logical block addressing (LBA) and SSD physical pages. As the core of an SSD, a control unit determines the response speed of SSD's data read and write. The NAND flash storage unit is the last place where data is stored.

In order to achieve blistering-fast performance, Huawei developed proprietary acceleration chips, in which the FTL algorithm was moved from the internal control software layer to the controller chip of an SSD. All FTL read and write operations are thus performed by chips, reducing software interactions and I/O response latency.

Testing data indicates that NVMe SSD write latency is reduced from 40 µs to 20 µs, and the SSD write speed is doubled in light-load scenarios (queue depth: 1, 4 KB random writes).

Intelligent BMC Management Chip for 93% Fault Location Accuracy

Extending the analogy, a good vehicle requires a high-quality center console that can adjust the driving status in real time, based on the road conditions. A higher-level console can intelligently predict potential future conditions, based on current road conditions and historical vehicle data, to prevent traffic accidents. Huawei OceanStor Dorado all-flash arrays use intelligent BMC management chips and built-in fault diagnosis and warning expert libraries to provide fast diagnosis and warning capabilities, ensuring the stable running of devices. A controller has many hardware units. In some cases, a hardware unit cannot detect failures in advance, and thus the built-in expert systems detect the sub-healthy hardware units in advance, and perform preliminary processing, such as by resetting the hardware units. If the processing fails, the specific fault information of the hardware unit will be reported to the upper-layer unit, a solution that boosts the fault location accuracy rate to 93%, and shortens the fault recovery time of Huawei storage from two hours to ten minutes.
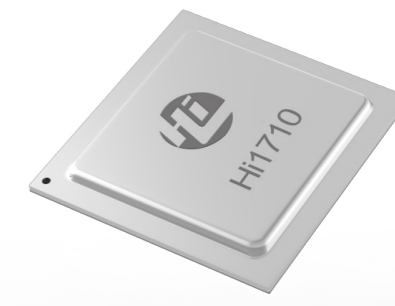

Fig. 7 Intelligent BMC management chip

! Conclusion

Huawei's five self-developed chips have achieved consistently ever fast performance, as indicated by the SPC-1 test, in which OceanStor Dorado all-flash arrays achieved an industry-leading performance of 200,000,000 IOPS, twice that of the next-best player.
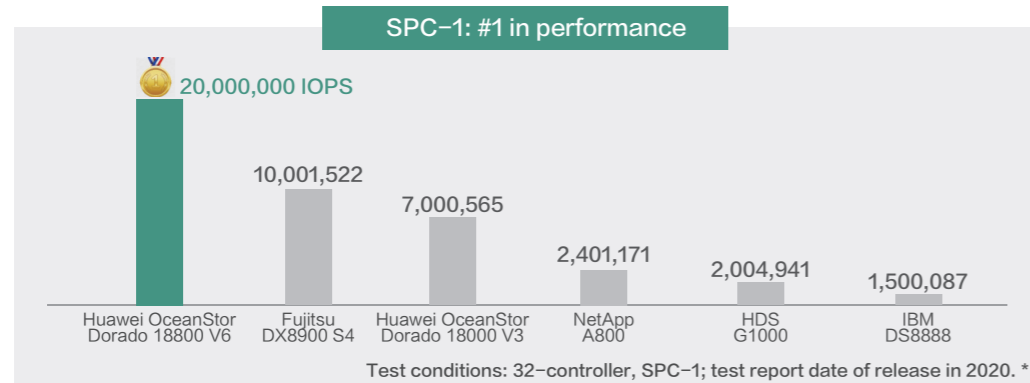

Fig. 8 IOPS in SPC-1

Tests performed in database scenarios reveal that Huawei OceanStor Dorado 18800 delivers 640,000 IOPS, also twice the standard figure from other vendors. (Test conditions: dual controllers, 100% random reads and writes, 7:3 mixed reads and writes, 80% space occupation, enabling 1 ms latency and effective capacity model.)
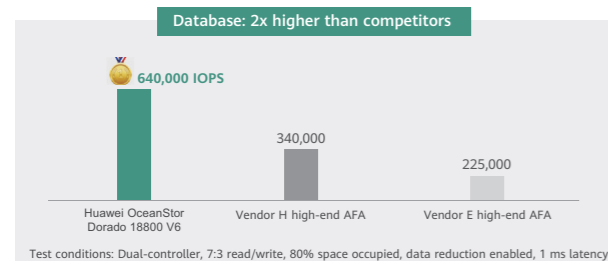
Word, PowerPoint, and Excel application tests performed in a virtual desktop scenario show that the response time for Huawei OceanStor Dorado 18000 is shortened by 80%. (Test conditions: Dual controllers, 100 x 3.84 TB SSDs, 8 TB/LUN, and 50 GB/VDI.)
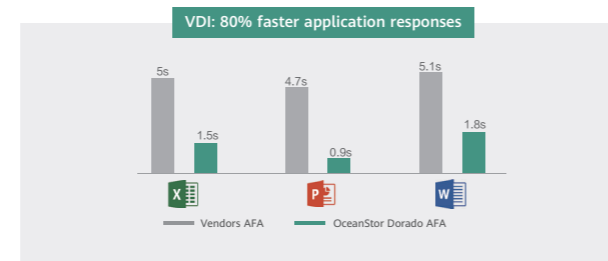

Fig. 9 IOPS in database

Fig. 10 Application response time in VDI

All of the above evidence demonstrates that in utilizing five self-developed chips, Huawei OceanStor Dorado all-flash arrays have eliminated performance bottlenecks caused by the unbalanced development of CPUs, media, and networks. In providing this solution, Huawei is equipped to help enterprises cope with the flood of data, and speed up the full-scale digital transformation process.

Lightning-Fast All-Flash Storage Embraces the E2E NVMe Era

By Wang Jiaxin, Huawei

Speed is the ultimate weapon to survive and thrive in fierce competition. Today, NVMe is maximizing speed to help users succeed.

Traditional SAS protocols are designed for HDDs and hinder SSD performance due to complex system architecture, excessive protocol parsing, and limited queue concurrencies. To solve these problems, NVM Express created NVMe protocol standards, and replaced complex protocol layers of the SAS system, such as I/O Scheduler and SCSI, with lightweight NVMe protocols. As a result, NVMe performs exceptionally well in all-flash arrays (AFAs).

Huawei embraced flash technologies in 2005 and consequently owns many core patents in flash memory developed over the past 15 years. Huawei has become one of the largest SSD-level patent holders with extensive research of NVMe protocol covering end-to-end (E2E) development of NVMe SSDs, flash controllers, and all-flash operating systems (OSs).

Huawei has remained a pioneer with its powerful capability matrix and E2E NVMe architecture for all OceanStor Dorado products for a 0.1 millisecond latency, making it the fastest in the industry.
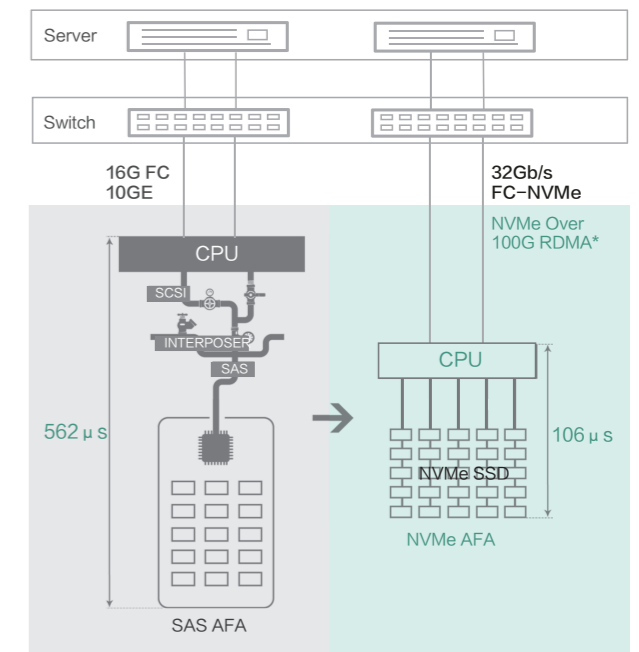

Fig. 1 Architecture: SAS vs. NVMe

But, how do Huawei NVMe AFAs outperform other SAS AFAs?

At the transmission path layer, I/Os of a SAS AFA are transmitted from: the front-end server to CPU → SAS chips through PCIe links and switches → SAS expander → SAS SSDs.
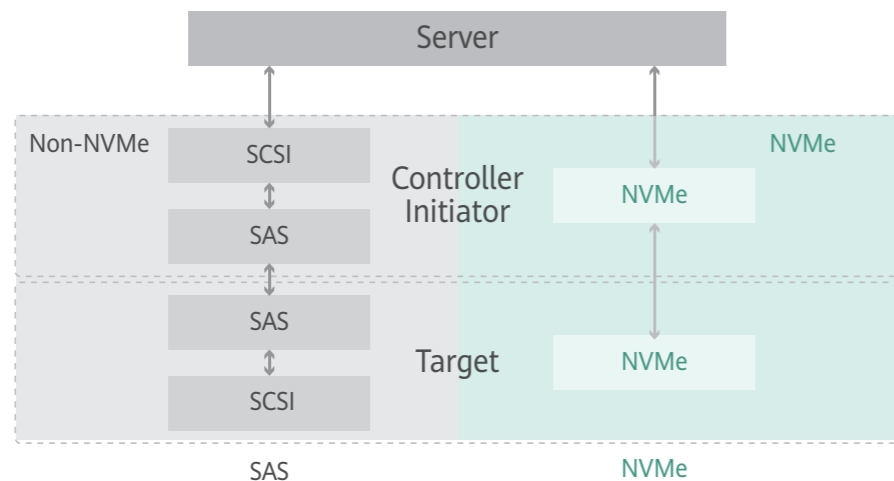
Fig. 2 AFA transmission path: Other SAS vs. Huawei NVMe

Huawei NVMe all-flash supports E2E NVMe. Data I/Os are sent from:

front-end servers → CPUs FC → NVMe/NVMe with RDMA front-end interface protocol of storage devices → NVMe SSDs through 100G RDMA at the back end.

The CPU of the NVMe AFAs directly communicates with NVMe SSDs for a shorter transmission path, higher transmission efficiency, and lower transmission latency.

Huawei is the first to use E2E NVMe across an entire all-flash storage series. Its OceanStor Dorado uses NVMe over Fabrics (NVMe-oF) to fully share SSD resources and provide networking designs for 32G FC-NVMe and NVMe over 100G RDMA for the connections of front-end networks and back-end SSD enclosures. Consequently, the controller interconnection for scale-out only needs one type of network protocol, providing three main advantages. Firstly, storage latency drops as low as 0.1 ms. Secondly, using one network system to control the whole data center simplifies storage network management. This avoids complex network protocol and planning, streamlines DC deployment, and reduces DC maintenance costs. Finally, users can customize with Huawei NVMe storage as needed while other vendors only provide E2E NVMe for high-end storage.
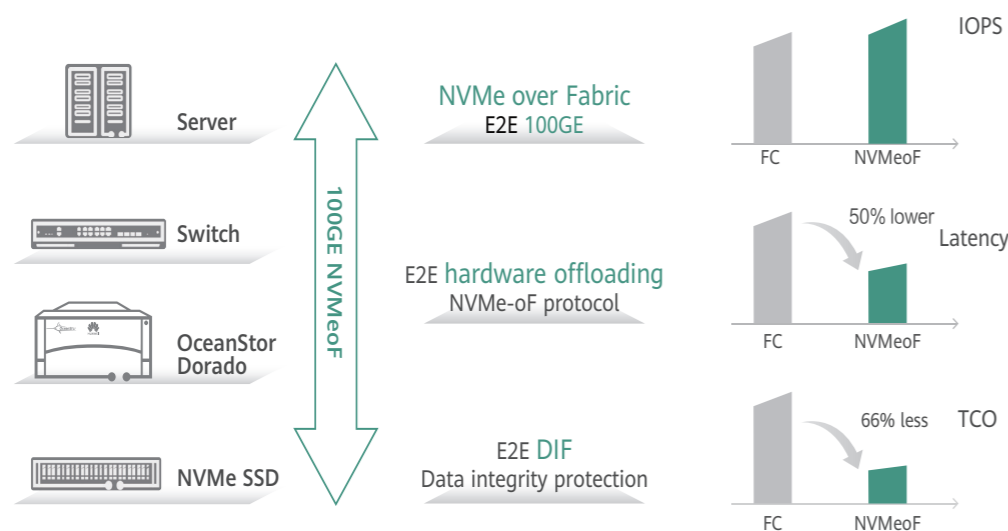

Fig. 3 Huawei's NVMe-oF E2E solution

At the layer of software protocol parsing, SAS and NVMe AFAs differ greatly in protocol interaction for data writes. A complete data write request of hosts requires 4 protocol interactions if using the SCSI (SAS back end) protocol. However, the NVMe protocol completes a write in only two interactions to double the write processing speed.
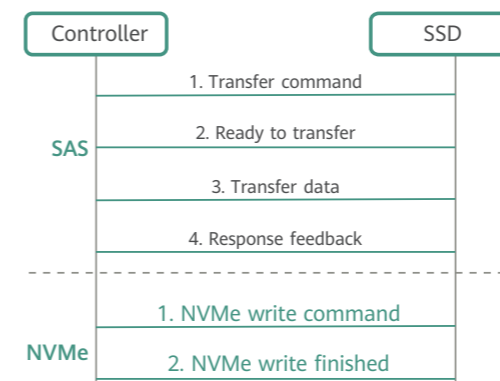

Fig. 4 Protocol parsing: SAS vs. NVMe

At the protocol encapsulation layer, SAS protocol stacks require I/O requests from block devices to reach SSDs through SAS links after the encapsulation of two layers (SCSI and SAS protocols). However, NVMe protocol stacks require only one encapsulation layer (NVMe protocol) for I/O requests. The simplified protocol stacks of Huawei NVMe AFAs lower encapsulation costs by 50%, reducing CPU consumption and I/O transfer delay caused by each encapsulation.

At the multi-queue concurrency layer, SAS protocol only supports a single queue, while NVMe protocol supports 64K queues. Each queue supports 64K concurrent commands to achieve higher concurrency processing and better cooperation with multiple SSD channels and dies.

At the layer of lock mechanism optimization, the SAS single-queue programming must be locked in the multi-core environment. Huawei has designed an I/O scheduling mechanism for NVMe to completely cancel the disk-level mutex on original I/O paths and avoid I/O processing conflicts. This allows concurrent processing of multi-core processors, reduces software overhead, and improves back-end processing.

At the layer of operating system optimization, Huawei's latest-gen all-flash storage has developed the innovative FlashLink® algorithm with intelligent chips and dedicated flash OSs. OceanStor Dorado has designed a many-core algorithm to capitalize on the potential of Kunpeng 920 chips for multi-core parallel computing. Huawei has developed an SSD enclosure with built-in Kunpeng 920 chips for improved computing capabilities, such as 15 min/TB reconstruction. This eases the controller enclosure load for a reconstruction speed that is 20 times faster than that of other vendors. For faster batch report processing, Huawei innovatively uses the Ascend 310 AI chip with the intelligent cache algorithm to learn and collect statistics on service rules from applications. As a result, the hit ratio of the read cache improves by 50%, and the batch processing is shortened by 40%. Huawei-developed SSDs and flash OSs minimize the impact of garbage collection on performance together with full-stripe writes, multi-streaming data separation, and E2E I/O priority adjustment. Storage controllers use optimized algorithms for internal software to understand and adjust the data layout in SSDs in real time. Controller data is written into the SSD in the required format to avoid migration and garbage collection after the data is flushed for high system performance and stability.

Innovative core technologies are driving enterprise development across industries. Not only has it benefited traditional industries like healthcare, manufacturing, and finance, but it has also driven growth in cutting-edge fields, such as AI, big data, autonomous vehicles, and block chain. Huawei is dedicated to building lightning-fast all-flash storage with its E2E NVME tech. Its continuous investment and innovation in the all-flash field aims to help enterprises deal with future challenges and accelerate their digital transformation. Ready to enjoy a more competitive all-flash storage in the flash era? Then think Huawei!

Unveiling the **5-Layer Reliability Design** of OceanStor Dorado

By Yang Jie, Huawei

> **Introduction: Data security and reliability are enterprise lifelines**
>
> Enterprises, governments, and other organizations rely on IT-based platforms because of their higher demands for business continuity and availability. Data loss and service suspension are fatal blows for organizations demanding informatization. Therefore, a reliable and effective reliability architecture that protects key data for 24/7 applications is essential.

❗ 5-Layer Reliability Design for Always-On Services

In July of 2019, Huawei publicly launched its next-generation OceanStor Dorado all-flash storage system. The system tops the industry with the highest IOPS at 20 million and the best latency at 0.1 milliseconds (ms). Its innovative SmartMatrix architecture propels the stability and reliability of high-end storage to new heights. Most importantly, the five-layer reliability design provides end-to-end (E2E) data protection with 99.9999% availability throughout the I/O process for always-on applications.
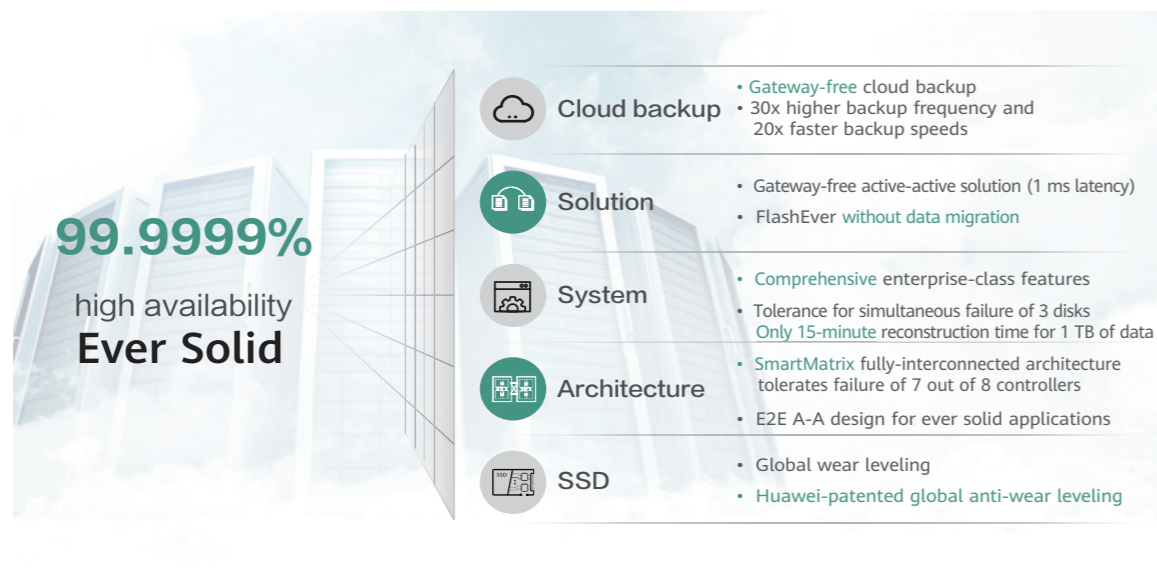
99.9999%
high availability
Ever Solid

Cloud backup
• Gateway-free cloud backup
• 30x higher backup frequency and 20x faster backup speeds

Solution
• Gateway-free active-active solution (1 ms latency)
• FlashEver without data migration

System
• Comprehensive enterprise-class features
• Tolerance for simultaneous failure of 3 disks Only 15-minute reconstruction time for 1 TB of data

Architecture
• SmartMatrix fully-interconnected architecture tolerates failure of 7 out of 8 controllers
• E2E A-A design for ever solid applications

SSD
• Global wear leveling
• Huawei-patented global anti-wear leveling

Fig. 1 Five-layer reliability structure of OceanStor Dorado

1.Always-On SSDs

Multiple data writes to SSDs decrease the SSD service life. Load imbalances may lead to overloading of certain SSDs, possibly resulting in unexpected failures. Huawei OceanStor Dorado addresses these issues by leveraging global wear leveling and anti-wear leveling. These technologies enhance SSD reliability and prevent multi-SSD failures for reliability throughout the entire lifecycle.

Global wear leveling balances loads among SSDs

OceanStor Dorado coordinates controller software and disk drives to regularly query the degree of SSD wear leveling, and uses that degree as a criteria for allocating space. In addition, RAID 2.0+ evenly distributes data among SSDs with Logical Block Address (LBA) and fingerprints for global load balancing.
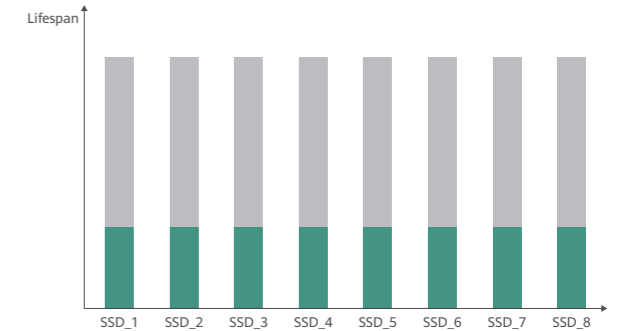
Global anti-wear leveling patented by Huawei prevents simultaneous failures of multiple SSDs

SSDs face challenges at the end of their lifecycle. For example, SSDs with wearing level exceeding 80% may have multiple SSDs fail simultaneously and lose data even if global wear leveling is still operational. The Huawei patent for anti-global wear leveling (patent No.: L201380000793.0) overcomes these challenges. The system selects the most severely worn SSD (such as SSD_1 in the right fig.) and writes data to it if it has idle space. This reduces the service life of the worn SSD faster than other SSDs, prompting replacement in advance and avoiding service disruption.
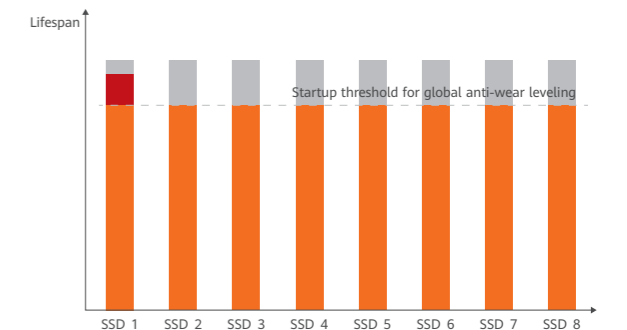
Fig. 2 Using global wear leveling tech in early SSD service life

Fig. 3 Using global anti-wear leveling tech in near-end SSD service life

2. Always-On Architectures

The fully-interconnected architecture of SmartMatrix tolerates failures for 7 out of 8 controllers

The high-end all-flash storage of OceanStor Dorado uses SmartMatrix architecture. This architecture uses shared front and back-end interface modules, shared SSD enclosures, and fully interconnected controllers. OceanStor Dorado uses three-copy cache to be the only system in the industry that tolerates failures for 7 out of 8 controllers and for one out of two controller enclosures with zero service interruption. The OceanStor Dorado sets the bar for storage reliability.

Active-active design balances loads

The high-end all-flash storage of OceanStor Dorado uses SmartMatrix architecture with Huawei-developed multipathing software, fully shared interface modules, global cache, and RAID 2.0+ technologies. It implements E2E load balancing for the host access, front-end, controller, and SSDs.
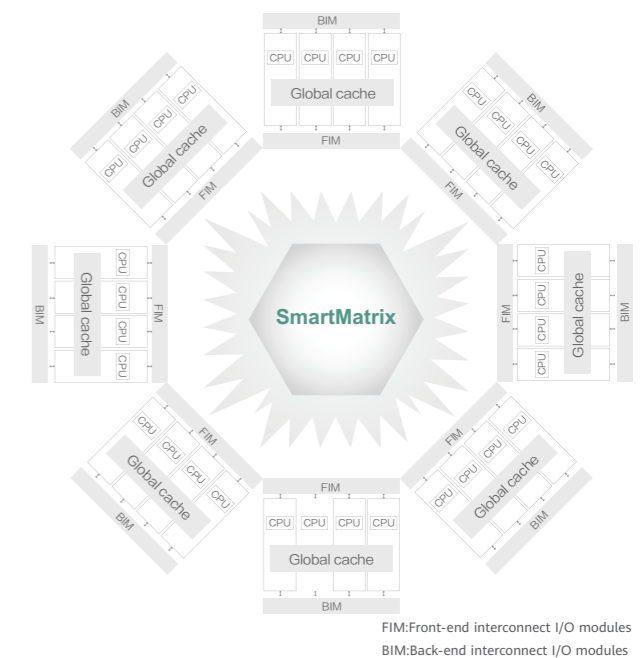
FIM:Front-end interconnect I/O modules
BIM:Back-end interconnect I/O modules

Fig. 4 SmartMatrix architecture

3.Always-On Systems

Comprehensive enterprise-class features for mainstream scenarios

OceanStor Dorado manages data throughout the lifecycle using Smart series software for improved efficiency, Hyper series software for data protection, and Cloud series software for local and cloud disaster recovery. It has full-lifecycle data management and builds a storage system with high reliability, performance, serviceability, and usability.

Tolerance for simultaneous 3-SSD failures and 15-minute reconstruction for 1 TB of data

OceanStor Dorado all-flash storage leverages RAID and smart SSD enclosure tech, supporting up to three simultaneous SSD failures in one RAID group with zero data loss. Tests show that it takes only 15 minutes, rather than the previous 30 minutes, to reconstruct 1 TB of data in the event of an SSD failure using smart SSD enclosure and controller synergy.
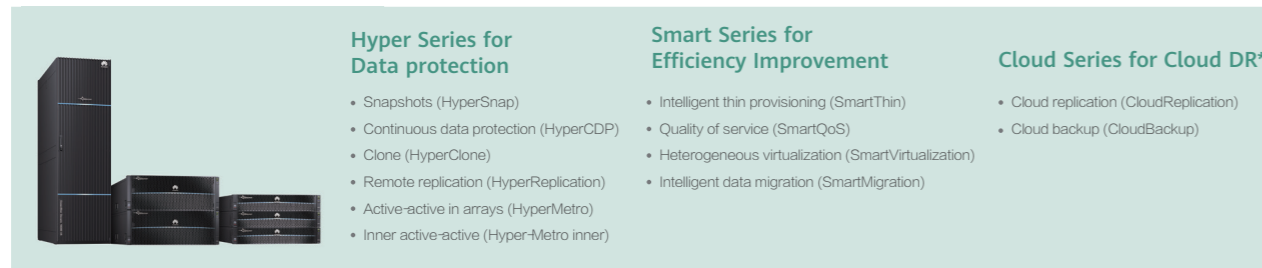
Hyper Series for Data protection	**Smart Series for Efficiency Improvement**	**Cloud Series for Cloud DR***
• Snapshots (HyperSnap)	• Intelligent thin provisioning (SmartThin)	• Cloud replication (CloudReplication)
• Continuous data protection (HyperCDP)	• Quality of service (SmartQoS)	• Cloud backup (CloudBackup)
• Clone (HyperClone)	• Heterogeneous virtualization (SmartVirtualization)	
• Remote replication (HyperReplication)	• Intelligent data migration (SmartMigration)	
• Active-active in arrays (HyperMetro)		
• Inner active-active (Hyper-Metro inner)		

Fig. 5 A full range of enterprise-class features

4.Always-On Solutions

HyperMetro gateway-free active-active model for always-on services

OceanStor Dorado uses the HyperMetro gateway-free active-active solution for zero service disruption upon any failure in a data center, meeting the most demanding reliability requirements of enterprises.
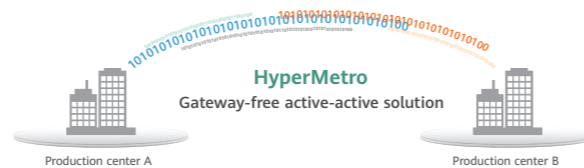
HyperMetro advantages:

*   Requires no additional storage gateways for I/O forwarding and directly groups two storage arrays into a cross-site cluster system for a quicker I/O response.
*   Active-active model. Data on two LUNs from two storage arrays synchronizes in real time, and can be accessed by the host simultaneously. RPO = 0 and RTO ≈ 0.
*   A variety of networking models are available: serial, parallel, or disaster recovery (DR) star networking.
*   Compatible with existing features of storage arrays with many data protection and BC&DR solutions.

HyperMetro
Gateway-free active-active solution

Production center A Production center B

Fig. 6 HyperMetro gateway-free active-active solution

5.Always-On Cloud Backup

OceanStor Dorado CloudBackup will allow the system to backup data to the public cloud or local NAS and object storage. Remote data backup and recovery will no longer require external backup servers, simplifying backup solutions and reducing costs.

❗ Conclusion

The reason that Huawei is always at the forefront of storage and data infrastructure transformation is simple. It continues to look ahead, gauge industry requirements, invest, and innovate. The all-flash storage of Huawei OceanStor Dorado is the result of a continual pursuit of perfection. Its innovative architecture design and five-layer reliability solution protect data throughout the I/O process for always-on services and a new benchmark in industry data reliability.

The SmartMatrix Architecture:
New Reliability Benchmark for High-End Storage

By He Yuan, Huawei

In recent years, all-flash storage is being used in more mission-critical service scenarios in fields that demand high reliability, such as finance and government fields. Research shows that one hour of unexpected downtime of mission-critical services causes losses of USD 6.48 million for the financial industry, USD 2.8 million for the energy industry, and USD 2 million for the telecommunications industry. On top of that, companies may also suffer client loss, brand image smear, as well as other immeasurable damage. In recent years, Huawei was determined to build an extremely stable storage architecture and ensure 24/7 business continuity for customers with its new OceanStor Dorado all-flash arrays (AFAs). In this article, we'll see how Huawei accomplished this mission.

A Full-Mesh, Balanced Architecture for a New Reliability Benchmark

The system architecture has long been considered as the core competency of a high-end storage system. For many years, no major breakthroughs have been made in mainstream high-end storage systems. Why do we think that Huawei OceanStor Dorado has set a new reliability benchmark? In the following figure, we can see that the SmartMatrix full-mesh balanced architecture adopted by OceanStor Dorado not only implements full interconnection of controllers, but also achieves full sharing of front-end and back-end resources. It features great tolerance toward hardware faults and, in the event of a fault, supports automatic fault tolerance within the system without causing service downtime.

Huawei: shared front end and no link switching upon controller failure | Vendor H: shared front end and no link switching upon controller failure | Vendor E: unshared front end and link switching upon controller failure

FC switch FC switch Service interrupted FC switch

Front end

Back end

Shared back end and zero service downtime upon failure of a controller enclosure | Unshared back end and service interruption upon failure of a controller enclosure | Unshared back end and service interruption upon failure of a controller enclosure
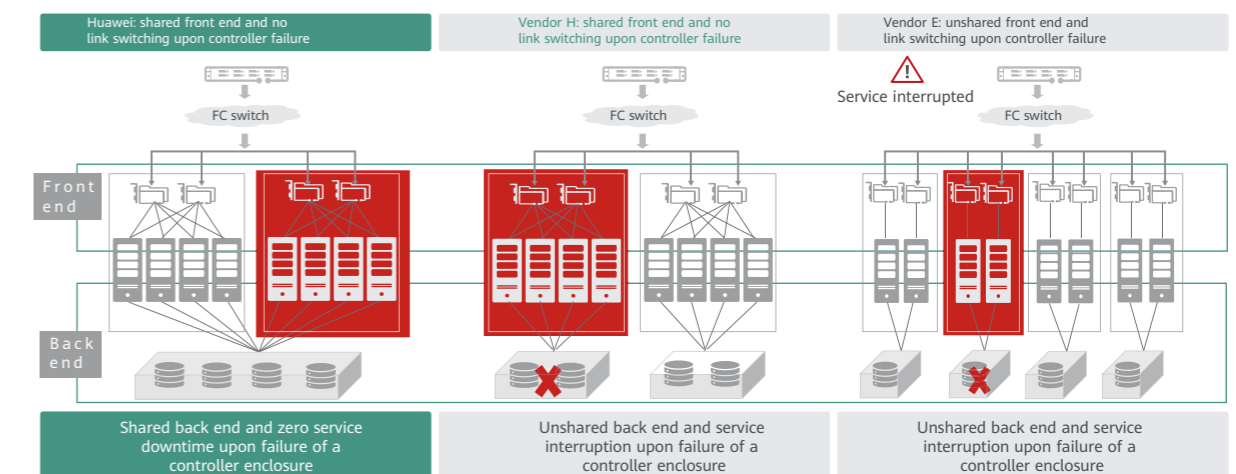
Fig. 1 Reliability architecture comparison

Compared with other mainstream high-end storage architectures, SmartMatrix has made breakthroughs and surpasses mainstream architectures in all aspects. As a result, Huawei OceanStor Dorado becomes the only storage series in the industry that tolerates failure of seven out of eight controllers, one out of two controller enclosures, and supports zero host service awareness upon a controller failure, setting a new benchmark in reliability.

	Huawei	Vendor H	Vendor E
Controller failure	Tolerance for failure of 7 out of 8 controllers	Tolerance for failure of 3 out of 4 controllers	Tolerance for failure of 2 out of 4 controllers
Controller enclosure failure	Tolerance for failure of 1 out of 2 enclosures	Failure of one enclosure is NOT tolerated	Failure of one enclosure is NOT tolerated
Link switching	No link switching upon controller failure	No link switching upon controller failure	Link switching upon controller failure

Table 1 Architecture reliability comparison

Now, let's see how OceanStor Dorado tolerates failure of seven out of eight controllers, one out of two controller enclosures, and supports zero service downtime in the event of a controller failure.

Breakthrough 1: No Link Switching and Zero Awareness of Host Services Upon a Controller Failure

The key to ensuring zero host service awareness when a controller fails is the switching time of front-end links. In traditional solutions, when a controller is faulty, front-end I/O path switching is triggered. Because the storage device cannot automatically switch I/O links, the host multipathing software must be used. The switching takes 5 to 30 seconds. OceanStor Dorado supports front-end interconnect I/O modules (FIMs), which intelligently identify and process host I/Os and distribute I/Os based on specific rules. When a controller is faulty, the port chip of the FIM detects that the PCIe link between the controller and the port chip is disconnected. Then in harmony with the service switchover within the controller, the FIM redistributes host requests to other controllers, implementing second-level controller failover and zero host service awareness.

Breakthrough 2: Tolerates Failure of Seven out of Eight Controllers and One out of Two Controller Enclosures Without Service Interruption

In a traditional high-end storage system, a disk enclosure cannot be accessed by all controllers, that is, it "belongs to" specific controllers. If a controller fails, the disk enclosure that belongs to it cannot be accessed and the system reliability is affected.

OceanStor Dorado leverages intelligent and fully shared back-end interface modules that connect disk enclosures with all eight controllers. Based on the three-copy cache technology, at least one valid write cache copy is provided to ensure zero service interruption in the event that two controllers fail at the same time or one controller enclosure fails. In addition, the continuous cache mirroring technology is provided so that when a controller fails, the mirror controller can be used to reconstruct cache mirror and enable the system to tolerate the failure of seven out of eight controllers without interrupting services.
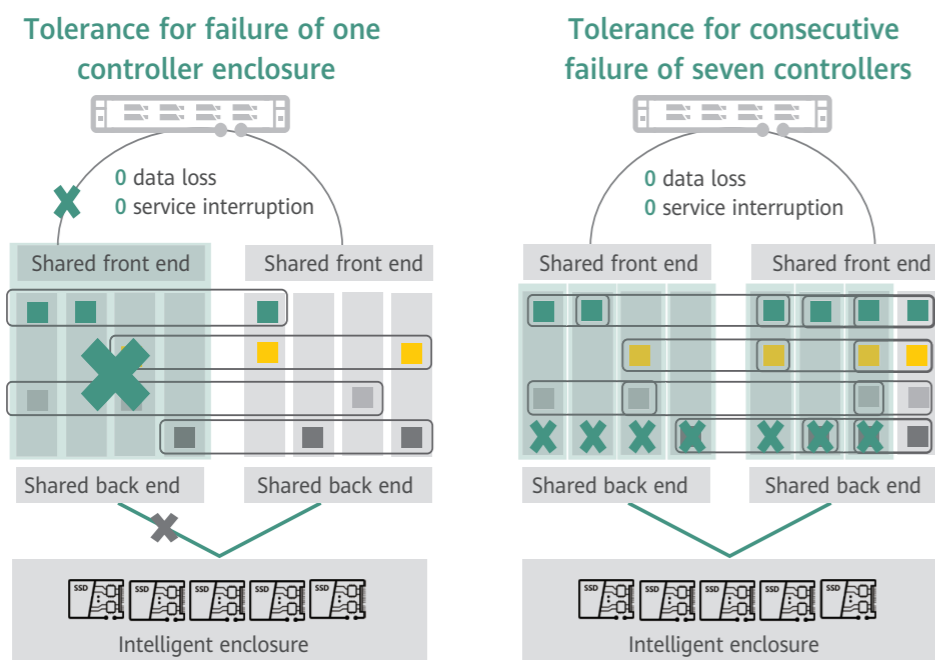

Fig. 2 Fault tolerance by SmartMatrix

End-to-End Global Load Balancing

The end-to-end global load balancing capability reflects the unique advantage of the SmartMatrix architecture. If we regard the hardware architecture as a skeleton, then the software architecture is the muscle and nerve. A tough skeleton works the best with flexible muscles and nerves. The OceanStor Dorado high-end AFAs are the only storage systems that implement end-to-end 4-layer global load balancing based on Huawei-developed multipathing software, FIMs, global cache, and RAID 2.0+ technologies. In comparison, other vendors only support one-layer load balancing.
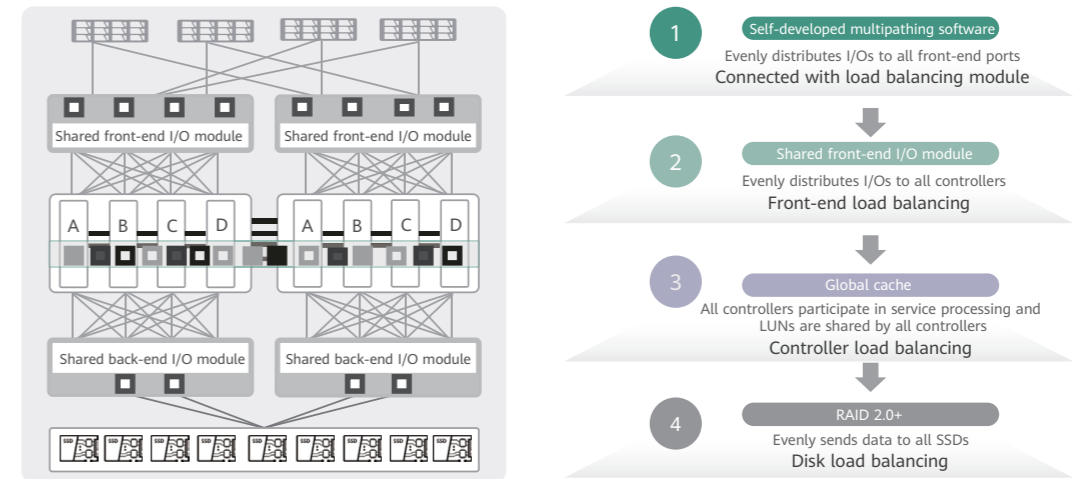

Fig. 3 End-to-end load balance

OceanStor Dorado uses the load balancing algorithm and innovative multipathing software to evenly distribute host I/Os to all front-end ports. Each front-end port evenly receives read and write requests from hosts, achieving load balance on the access side.

Front-end ports use the multi-controller interconnection design of the FIM and an optimized balancing algorithm to evenly distribute I/Os to all controllers. Each controller evenly receives read and write requests from hosts, implementing front-end load balance on the storage side.

The global cache technology is used to process read and write requests received by each controller in a unified manner globally. Because LUNs do not belong to any controller, I/Os are processed by the nearest controller. In this way, I/Os do not need to be forwarded to the controller to which a LUN "belongs." This ensures even load distribution and load balance among controllers.

RAID 2.0+ enables data to be evenly distributed to all disks in a storage pool, achieving load balance among SSDs.

Conclusion

Huawei OceanStor Dorado high-end AFAs employ the SmartMatrix fully-interconnected, balanced architecture. This powerful innovative architecture has set a new reliability benchmark for the industry and will lead the development of all-flash storage.

Huawei OceanStor Storage
Ecosystem with Openness

By Zhou Weigang, Senior Architect of Huawei Storage Ecosystem

The IT industry has its origins in closed systems. Mainframe systems such as those by IBM are a package of everything from computing, networking, and storage integrated hardware, to basic software of operating system (OS) and DR management, and applications such as databases. In the past, closed systems provided excellent IT services for financial institutions and large manufacturing enterprises; however, because they feature slow innovation, limited application scenarios, high costs, and low maintainability, they are insufficient for many diverse scenarios. For these reasons, closed systems are losing popularity.

In recent years, the rapid development of the IT industry is rooted in open systems. A number of outstanding IT enterprises have emerged, providing individual and enterprise users with comprehensive IT services. These enterprises would not be where they are today if they didn't leverage an openness concept, which has contributed to an increasingly open and powerful IT ecosystem.

Since its inception, Huawei Storage has adhered to the concept of openness, complied with open protocol standards and specifications, and is widely compatible with mainstream IT software and hardware products in the industry. OceanStor Dorado is the next generation of premium all-flash storage from Huawei. It inherits the 15-year ecosystem capabilities of Huawei storage, fully compatible with various third-party software and hardware, and actively involved in Huawei Kunpeng-powered data ecosystem.

! Openness, in the Genes of Huawei Storage

(1) Open Protocol Standards and Specifications

Today's IT systems are open because they run in accordance with a series of open protocol standards and specifications that are defined for data access, transmission, and management. Huawei storage products interact with third-party products in the IT ecosystem in strict compliance with related standards and specifications.

- Data access protocol standards: SCSI, SMB, NFS, and more
- Data transmission protocol standards: Fibre Channel (FC), iSCSI, FCoE, InfiniBand, and more
- Data and device management protocol standards and specifications: NDMP, SNMP, SMI-S, RESTful, OpenStack Cinder, OpenStack Manila, and more
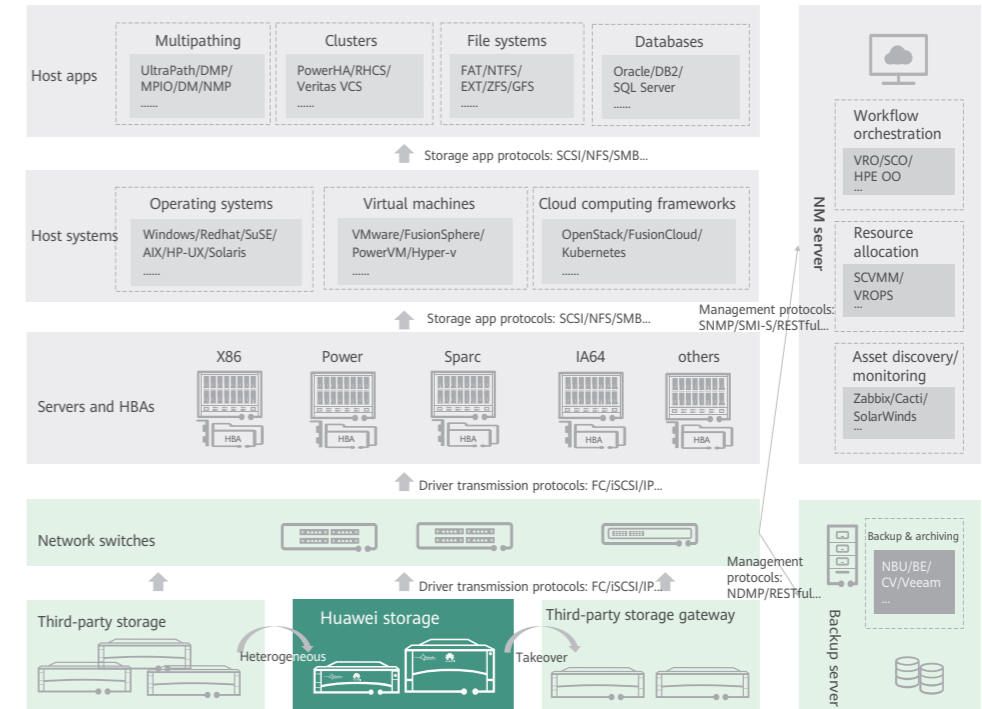

Fig 1 Overview of OceanStor ecosystem

With 15 years of experience in protocol understanding, design, and implementation, Huawei OceanStor Dorado is able to interoperate with mainstream IT software and hardware products, and continues to improve interoperability with other mainstream technologies over the product evolution.

(2) Integrating Data Management Software into SODA

SODA (formerly, OpenSDS) is the first branch of the Linux Foundation in the data and storage field. It is jointly established by Huawei, other industry-leading data and storage vendors and enterprises around the world. SODA was set up to address data management challenges of enterprise users' IT infrastructure architectures in cloud-native and multi-cloud scenarios.

With SODA's ecosystem capabilities, Huawei Storage opens abundant storage features to upper-layer application management software using SODA APIs. Third-party application software only needs to support SODA APIs. With the southbound ecosystem capabilities of SODA, third-party application software can be compatible with Huawei Storage.
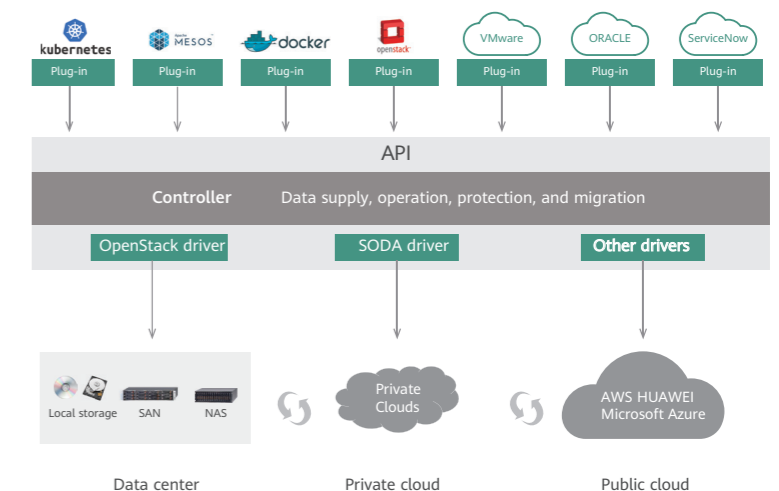

Fig 2 Open and unified data management with SODA

Openness, a Result from Huawei's Huge Investment

Building a Professional Interoperability Lab

Over the past 15 years, Huawei Storage has invested more than USD 200 million in building a world-leading professional storage interoperability lab.
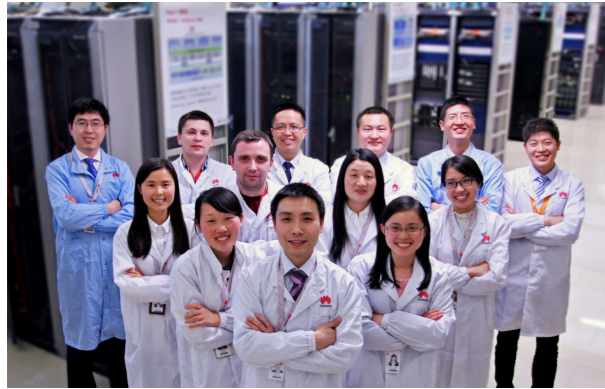

Fig 3 Huawei Storage Interoperability Lab

Cooperation with Authoritative Organizations

Huawei storage interoperability is strictly verified by using extensive software and hardware resources in Huawei Interoperability Lab. In addition, it is certified by authoritative organizations in and outside China. Third-party organizations evaluate Huawei storage interoperability more objectively, and enrich and supplement interoperability test scenarios. For example:

In China: Huawei Storage actively participates in the certification of China National Computer Products Quality Supervising Test Center (NCTC). This organization has China National Accreditation Service for Conformity Assessment (CNAS) qualification and the status of the multilateral, mutual recognition agreement party in International Laboratory Accreditation Cooperation (IALC). The certification result is authoritative around the world. Huawei Storage products have been certified by NCTC.

Outside China: Huawei Storage actively participates in European Advanced Networking Test Center (EANTC) certification. The EANTC is an internationally recognized objective test center, especially in Europe.

(1)Huawei Storage Interoperability Lab works on more than 4,000 third-party software and hardware products, including 800+ versions of OS software, 200+ server models, 200+ host HBA models, 100+ switch models, as well as a large quantity of backup, network management, cluster, database software, and hardware drivers.

(2)The storage interoperability lab features a workforce of more than 100 engineers, who are proficient in analyzing and verifying the interoperability of software and hardware products from various IT vendors. The lab provides interoperability consulting and services for more than 2,000 projects every year, and has received more than 1,000 visits from industry professionals. More than 10 million types of configurations are included in the interoperability matrix, allowing Huawei storage products to be widely compatible with mainstream IT products. Design and verification are performed to continuously integrate the interoperability capabilities into new products.

Fig 4 NCTC certificate for OceanStor

EANTC Independent Test Report
Huawei OceanStor Dorado and VMware Interoperability Test
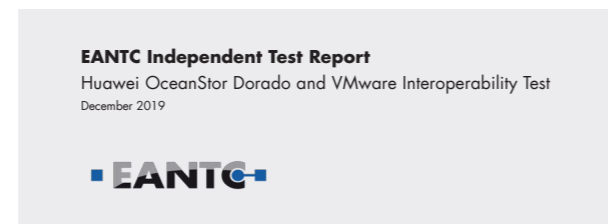December 2019

Fig 5 EANTC assessment certificate for OceanStor

Building an Open Cooperation Platform with Kunpeng Storage Ecosystem

While participating in the certification of authoritative organizations in the industry, Huawei Storage has also actively built Kunpeng storage ecosystem. Currently, Kunpeng certification platform has attracted multiple vendors of operating systems, databases, big data, and cloud platforms, and has released more than 100 Huawei certification reports.
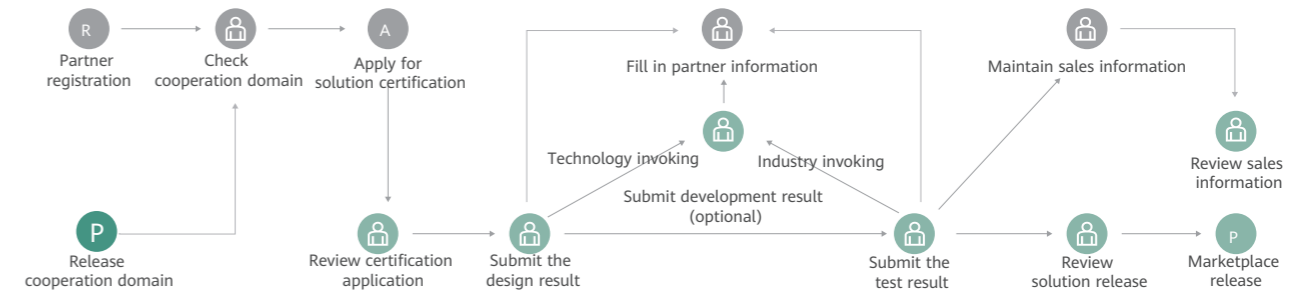

Fig 6 OceanStor storage ecosystem certification process

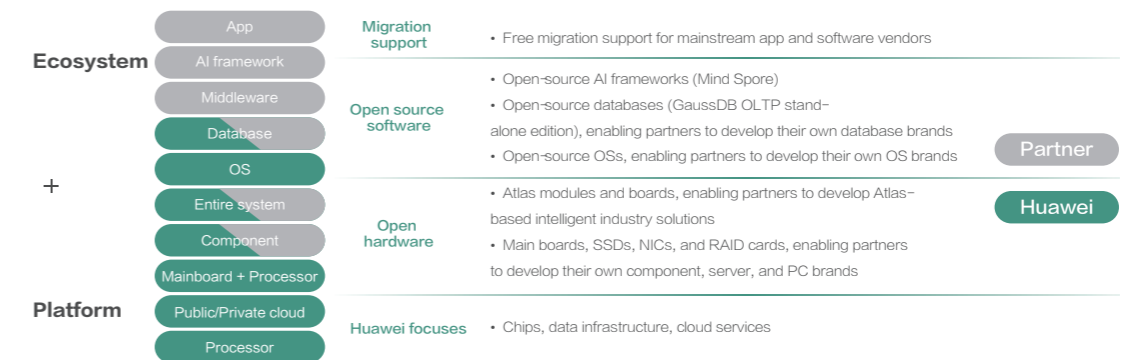# Openness, as in Huawei's Future Storage Strategy

Fig 7 Huawei storage ecosystem strategy

Since 2017, Huawei has been adhering to the Platform + Ecosystem strategy, and the three concepts of Technology Centered, Openness Enabled, and Partner First. Huawei has opened up hardware and software resources to build a fertile ecosystem, and worked with partners to help enterprises embrace a fully connected, intelligent world and achieve business success.

In 2019, Huawei OceanStor Storage actively adheres to the One Cloud, Two Wings, Two Engines strategy and builds core competitiveness with Kunpeng and Ascend processors. Currently, all series of centralized storage products are switched to the Kunpeng platform. Huawei also invests in the deployment of the Kunpeng-powered industry ecosystem and preferentially supports partners to develop data products based on Kunpeng processors.

Openness, as for Customer Trust

Huawei Storage has been established for 15 years, and has served over 10,000 users from Asia, Europe, Latin America, the Middle East, and Africa, covering sectors such as finance, carriers, large enterprises, and governments. Huawei storage products have been tested in various user application scenarios and have good interoperability with mainstream IT software and hardware products in the industry. Thanks to a back catalogue of successful projects, Huawei storage products have won the trust of users.

In the future, Huawei will look to improve the storage ecosystem and product interoperability by adhering to the concept of openness, helping users with more secure and reliable storage services.

HUAWEI

Transforming Transaction Latency into Transaction Efficiency

Huawei OceanStor Dorado improves online transations during peak hours by **5x**

7 of the world's top 20 banks process their data with the world's fastest all-flash storage
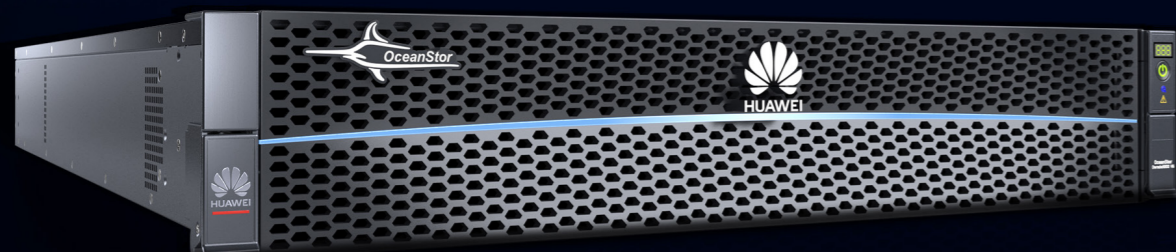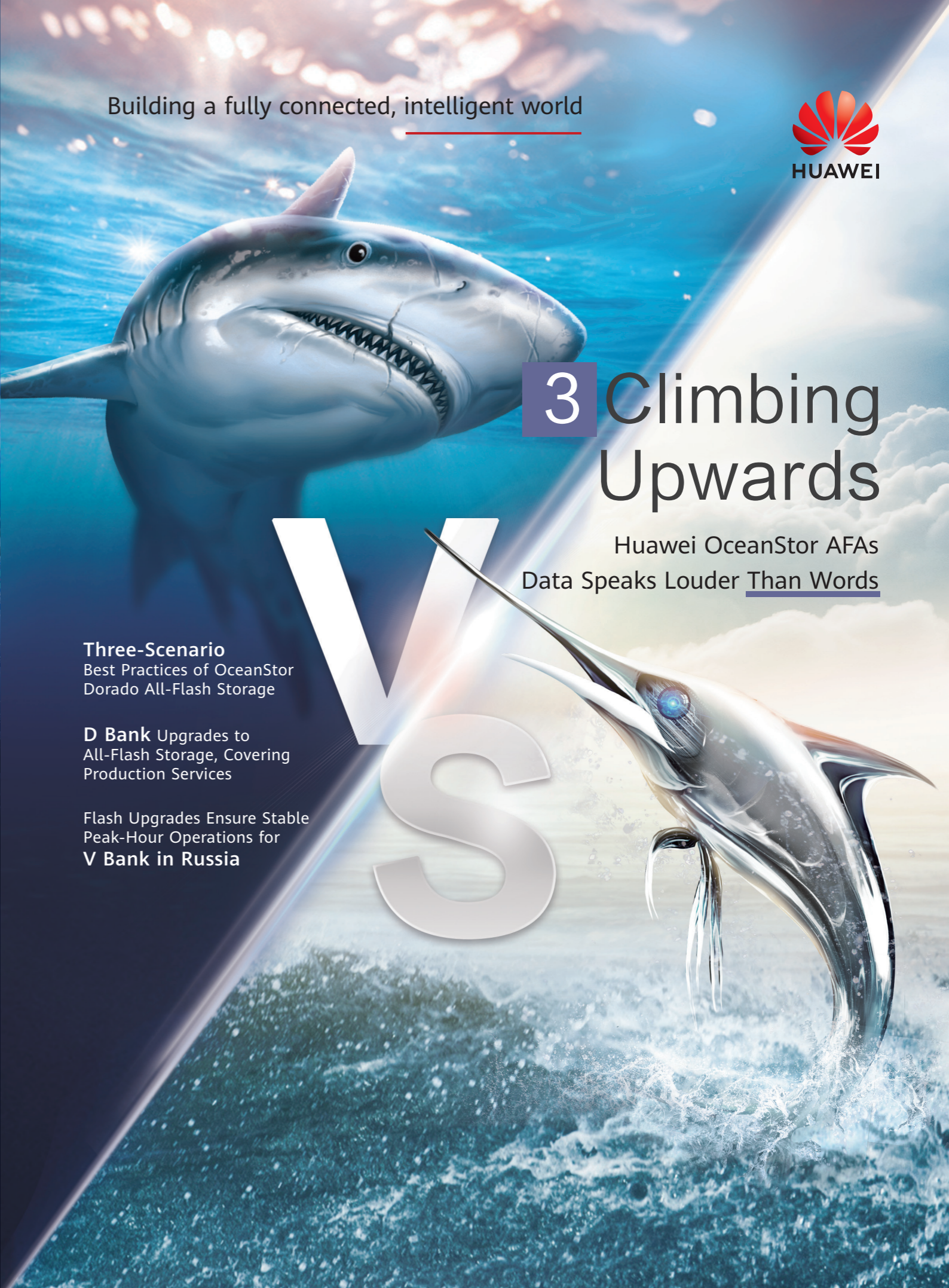
3 Climbing Upwards

Huawei OceanStor AFAs
Data Speaks Louder Than Words

VS

Three-Scenario
Best Practices of OceanStor Dorado All-Flash Storage

D Bank Upgrades to All-Flash Storage, Covering Production Services

Flash Upgrades Ensure Stable Peak-Hour Operations for **V Bank in Russia**

Three-Scenario Best Practices of
OceanStor Dorado All-Flash Storage

By Tang Jialong, Huawei Test Expert

OceanStor Dorado is the next generation of all-flash array (AFA) from Huawei. The OceanStor Dorado uses five intelligent chips and a multitude of efficient algorithms to deliver 20 million IOPS, 100-microsecond latency, and hundreds of gigabytes of bandwidth. It more than satisfies the requirements of online transaction processing (OLTP) and online analytical processing (OLAP) in mainstream databases for the finance and carrier industries. In addition, OceanStor Dorado has adapted and enhanced algorithms for virtualization and supports ecosystem plug-ins for a variety of virtualization platforms to greatly improve VM efficiency and advance the all-flash era across industries.
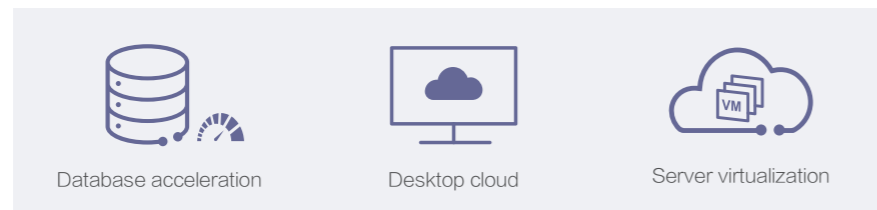
Database acceleration   Desktop cloud   Server virtualization

Fig. 1 Mainstream app scenarios of OceanStor Dorado

❙ Scenario 1: Database Acceleration

Oracle is the most common database for large enterprises. It carries key data of the ERP system. The data explosion in the 5G/AI era has greatly challenged Oracle performance and increased its operating cost.

Oracle Database Challenges

Storage performance from building-block expansion

With the development of informatization, banks and large e-commerce companies now process over 100 million daily transactions. The IOPS of core service systems exceeds 200,000 and requires a latency of around 1 ms. Traditional storage systems cannot satisfy these new requirements through building-block expansion.

Business continuity

Service continuity is vital for the Oracle OLTP, which manages mission-critical businesses for enterprises.

Oracle license costs increase annually

Oracle license fees have always consumed a large proportion of the IT O&M budget. For example, the Oracle license fees of a 10 TB database account for 50% of the annual O&M costs and increase as the data volume grows. How to cut Oracle license fees is a critical challenge for many companies.

All-Flash Solution for Oracle Databases

The Oracle database all-flash solution uses the powerful performance of Huawei OceanStor Dorado for a 5x service scale for online transactions. In addition, a low latency of 0.5-ms reduces the I/O waiting time, improves storage and server resource utilization, and significantly reduces Oracle license fees. The innovative HyperCDP high-density snapshot provides the highest data protection frequency in the industry and lossless performance during snapshot creation.

A verification test for the OceanStor Dorado performance for Oracle databases was set up. 12 hosts delivered I/O to the storage array for a stable 500-µs latency. Fig. 2 shows the test network and configuration.
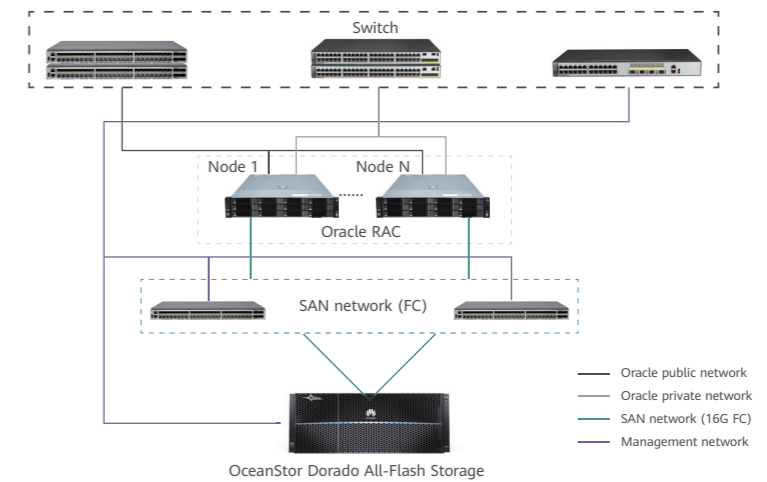
Fig.2 Ultimate performance test networking (Oracle)

Category	Type	Quantity/Size	Release
Storage	OceanStor Dorado 18000	40*960 GB SSD	
Database servers	RH2288H V6	Intel(R) Xeon(R) CPU E5-2660 v4 @ 2.00GHz (total28 cores), 256 GB RAM	
Operating system (OS)		RHEL 7.5	
Host bus adapter (HBA)	Qlogic	2 x 16 Gb dual-port HBAs	QLE2672
Oracle database	Oracle RAC		Oracle Database 18c Enterprise Edition Release 18.0.0.0.0
Benchmark tools	OLTP	SLOB	v2.4
		SwingBench	v2.5
	OLAP	TPC-H	v2_17_3
Host number		12	

Table 1 Test configuration of Oracle performance

6x faster than the industry benchmark Dell EMC VMAX-F

The SwingBench tool simulated an online transaction system. The performance of the 4-controller OceanStor Dorado 18000 reached **840,000 IOPS at 0.5-ms latency, 50x that of traditional storage systems, and 6x that of Dell EMC VMAX250F. In addition, OceanStor Dorado supports scale-out and linear performance growth.**

Storage Device	Test Tool	Workload Model	IOPS
OceanStor Dorado 18000	Swingbench(OE2)	7:3, 8 KB, random	840,000@500µ s
Dell EMC VMAX250F	Swingbench(OE2)	7:3, 8 KB, random	133,000@1ms

Table 2 Result comparison of Oracle test

Array · Avg.I/O Response Time(μs)

Array · Block Bandwidth(GB/s)

Array · Total IOPS

Array−Total IOPS
2019−11−30 11:20:05
● System   846,091 IOPS

Fig.3 OceanStor Dorado performance test (840,000@500 μs)

100-μs latency, 10x faster than the industry average

Huawei OceanStor Dorado not only provides high IOPS, but can also reduce the latency to 100 μs in scenarios with high cache hit ratio. A test was performed with the SLOB tool. The 4-controller OceanStor Dorado 18000 achieved a 75% read cache hit ratio, 220,000 IOPS, and less than 100-μs latency in the 25% update and 8:2 read/write ratio scenario (the most typical database I/O model). Even if the IOPS of the core service system reaches 200,000, OceanStor Dorado can still keep the latency at about 100 μs. This is 10x the 1 ms latency of the all-flash storage of other vendors and far exceeds the requirements of mission-critical applications.

Array · Avg.I/O Response Time(μs)

Array · Block Bandwidth(GB/s)

Array · Total IOPS

Array−Total IOPS
2019−11−30 10:42:50
● System   220,528 IOPS

Fig.4 OceanStor Dorado performance test (220,000@100 μs)

HyperCDP without compromising performance

Service continuity and data reliability are critical for Oracle databases because they manage mission-critical applications for enterprises. OceanStor Dorado safeguards mission-critical applications using high-density snapshots with the highest protection frequency in the industry for a shortened snapshot interval of 3 seconds and lossless performance during snapshot creation.

In a snapshot test for four-controller OceanStor Dorado 18000 and HDS F1500 all-flash storage, Huawei OceanStor Dorado far exceeded other vendor products in snapshot density and activation impact.

As shown in Fig. 5, OceanStor Dorado supports a minimum snapshot interval of 3 seconds. In contrast, HDS snapshots are only configurable for 15-minute intervals through APIs.
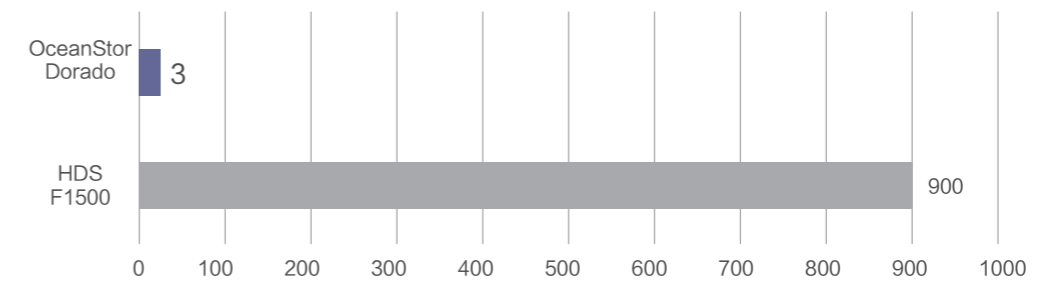

Fig. 5 Snapshot frequency comparison

Fig. 6 shows the impact on OceanStor Dorado and HDS F1500 when snapshots are created, activated, and then mapped to hosts. OceanStor Dorado performance is uncompromised since a snapshot is created and activated within several seconds. Conversely, HDS F1500 took 75 seconds to create and activate a snapshot, greatly increasing latency.
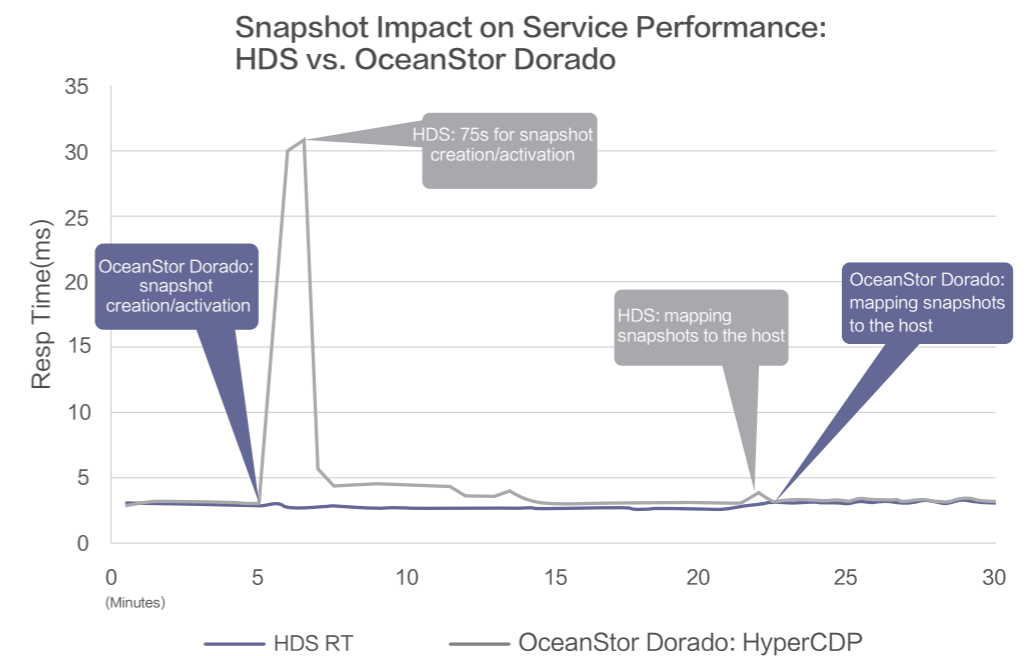

Fig. 6 Snapshot performance impact

Oracle license fees slashed by 30% or more

The ultra-low latency of OceanStor Dorado reduces the I/O waiting time and host CPU usage. Fewer CPU resources are needed to handle the same service pressure, greatly reducing costs for Oracle license procurement.

Fig. 7 shows the test results of traditional storage and OceanStor Dorado when SwingBench OE2 was used to simulate online transactions. Under the same service pressure (800 TPS), the host CPU usage of OceanStor Dorado was 38.7%, while traditional storage usage was 60% or more. Storage arrays are no longer a service bottleneck. Fewer host CPU resources carry the same services, reducing Oracle license and host purchasing costs. At 800 TPS, the Oracle license fees are reduced by 31.3%.
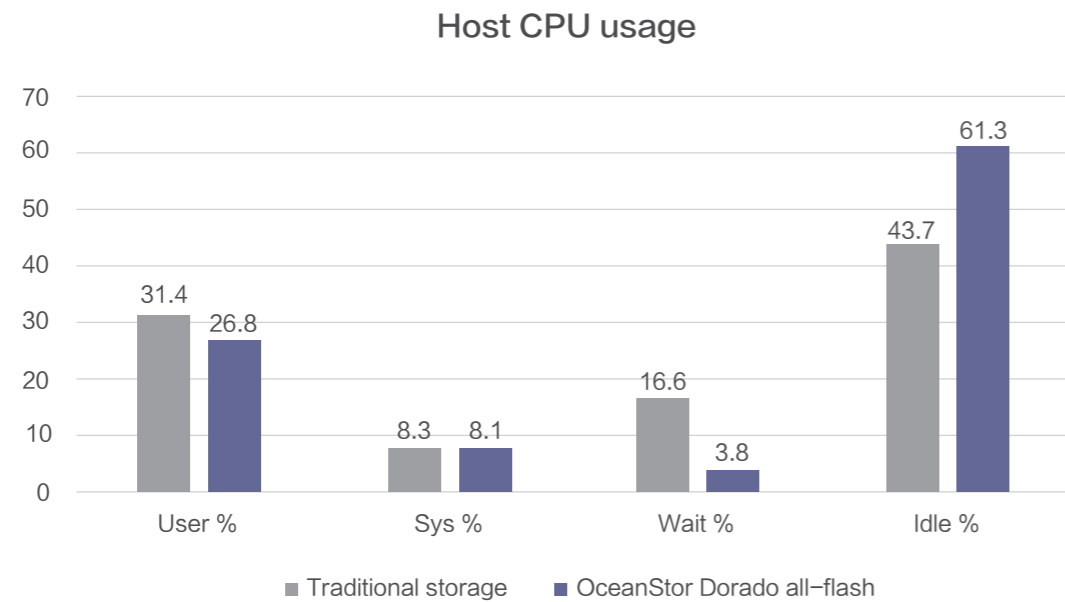
Host CPU usage

Fig. 7 Comparison of host CPU usage

Scenario 2: Better User Experience with Virtual Desktop Infrastructure (VDI)

Traditional PC-centric tools have lengthy and costly management and delivery services. VDI is widely used in enterprise IT systems so that end users can work on any device anytime and anywhere. However, the centralized VDI system has much higher requirements for availability, performance, and maintenance than the traditional PC environment.

VMware VDI Challenges

Time-consuming O&M: It takes five hours to deploy 500 virtual desktops with traditional storage because of the protracted waiting time for capacity expansion. This has adverse effects during emergencies. For example, it takes more than two hours for patch installation and restart of 500 desktops, practically stopping daily office work.

Poor user experience: There is a 10-second response time to open or save a PPT, Word, or Excel file with a traditional VDI solution.

VDI Storage Solution

The industry-leading performance and low latency of OceanStor Dorado fosters quick maintenance of VDI systems and an enjoyable user experience.
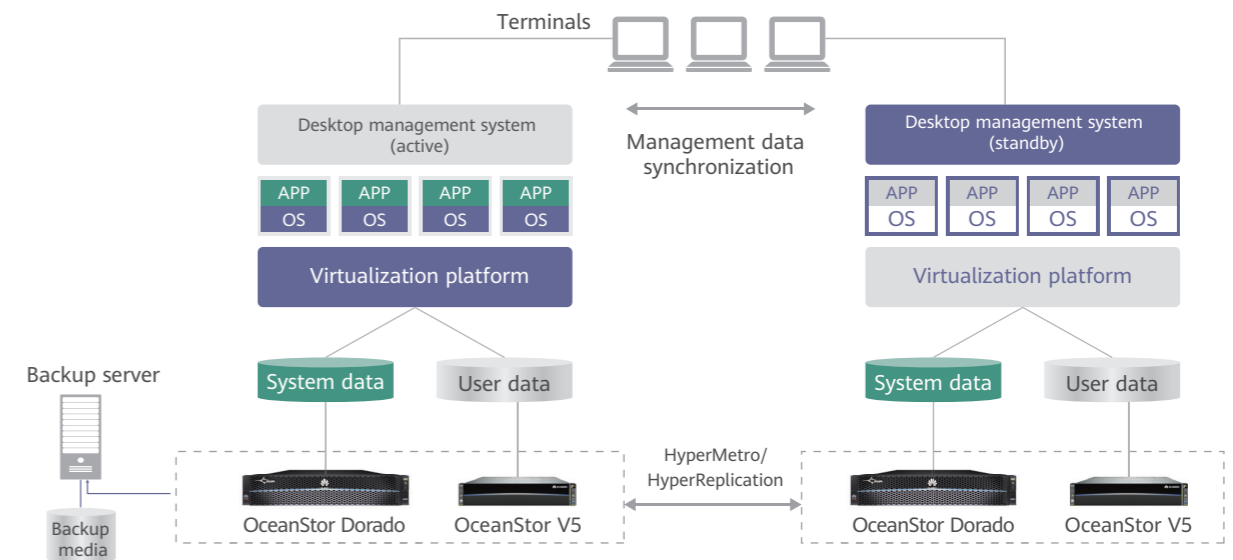

Fig. 8 Reference architecture of OceanStor Dorado for deploying VMware VDIs

The following test shows the differences between OceanStor Dorado and traditional storage arrays in common VDI scenarios.

Hardware	Configuration	Quantity
Server	E9000	8
Storage	OceanStor Dorado 5000 2 x Controllers 1 x 2 U 25-slot SSD enclosure 25 x 960 GB SSDs	1
Fibre Channel (FC) switch	16 Gbit/s Huawei FC switch	2
1GE switch	1GE Huawei Ethernet switch	1

Table 3 VDI hardware configuration

Test Item	Software Version
VDI host operating system	VMware ESXi 6.7.0 Update 2
Multipathing software	Huawei UltraPath 6.2.0
Virtualization software	VMware vSphere 6.7.0 U2
Virtual desktop operating system	Windows 7 Enterprise 64-bit
VDI management software	VMware View 6.1
Test tool	View Planner3.6
Virtual machine (VM) specifications	4 vCPUs, 8 GB memory, and 50 GB disks

Table 4 VDI software configuration

80% acceleration of batch VDI deployment and startup for better O&M

A verification test for VDI deployment with a set of two-controller OceanStor Dorado 5000 was set up. It took five hours to deploy 500 virtual desktops with traditional storage but only 9 minutes with OceanStor Dorado 5000. **The deployment speed was over 30 times greater.**

A verification test for VDI startup with a set of two-controller OceanStor Dorado 5000 was set up. It took 15 minutes to start up 500 virtual desktops with traditional storage but only 176 seconds with OceanStor Dorado 5000. **The boot storm time was shortened by 80%.**
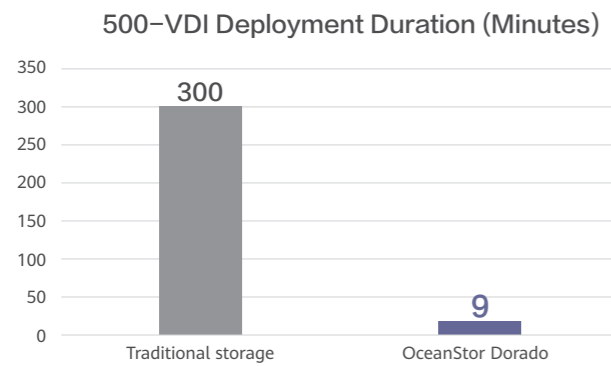
500-VDI Deployment Duration (Minutes)

Fig. 9 500-VDI deployment duration

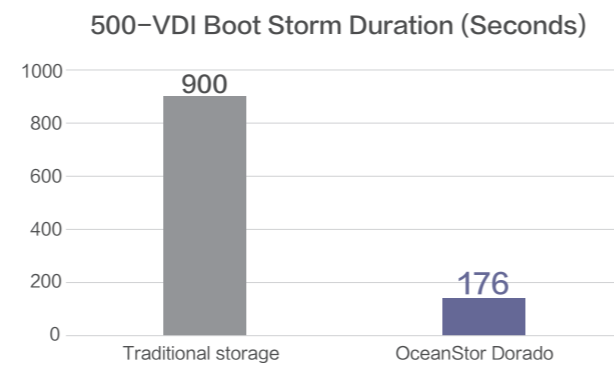### 500-VDI Boot Storm Duration (Seconds)

Fig. 10 500-VDI boot storm

In VDI scenarios, storage arrays have daily multiple I/O peaks. Boot and login storms occur every morning. Virus scanning is periodically started and patches are installed on desktops at night. A high volume of I/O negatively affects storage arrays. The extremely high performance of OceanStor Dorado maintains stability and low latency through I/O peaks.

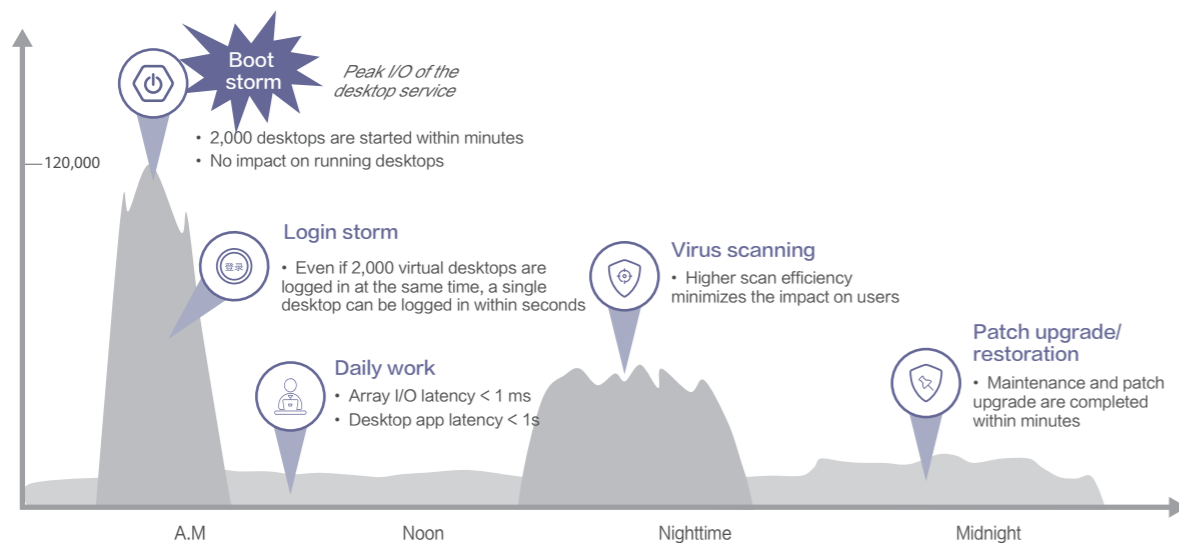## 24/7 stability through boot storms, maintenance, and emergencies

Fig. 11 VDI daily operation

80% acceleration of batch VDI deployment and startup for a better user experience

Microsoft Office software is most frequently used in virtual desktops. View Planner simulated 500 VDI users using Microsoft Office software and OA users accessing a wide range of Office products. Table 5 lists the configuration details.

Office load scenario

Load	Percentage
Word	30%
Excel Sort	30%
PowerPoint Presentation	30%
Outlook	10%

Hybrid load office scenario

Load	Percentage
Office software, Internet Explorer, and Adobe Reader	70%
Windows Media Player	20%
7-Zip compression and decompression	10%

Table 5 VDI desktop test load configuration

Two rounds of tests with View Planner produced the results as provided in Table 6:

Test Item	Duration (s)	Avg. Latency (ms)	Avg. Bandwidth (MB/s)	Max. Latency (ms)	Max. IOPS
500 desktops – Office scenario	2,040	0.156	125	0.276	17,742
500 desktops – Hybrid load office scenario	2,400	0.197	466	0.288	24,889

Table 6 Load test results for 500 desktops

The maximum latency of OceanStor Dorado was less than 300 μs, and the average latency was only 100 μs. These results show that OceanStor Dorado delivers excellent performance in the VDI solution for a superior experience that surpasses traditional PCs. Fig. 12 compares traditional storage and OceanStor Dorado in the time required to open these applications in Microsoft PowerPoint and Word. Compared with traditional storage, OceanStor Dorado takes less time to open, close, and save Word, PPT, XLS, and Excel files. The response time is shortened by 50% for a greatly improved user experience.
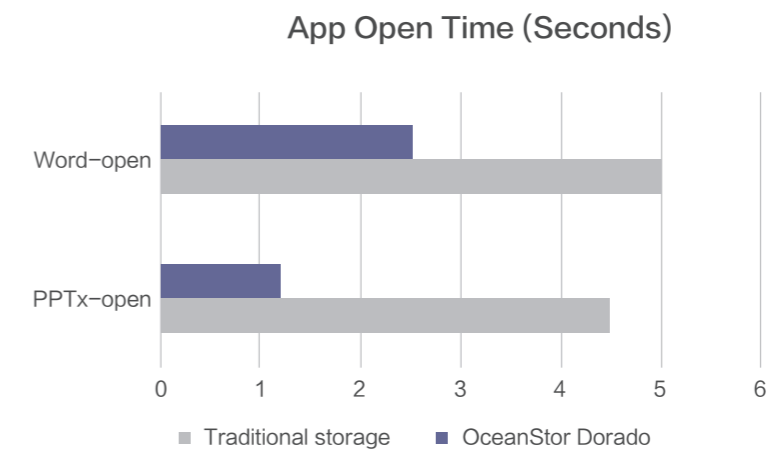
App Open Time (Seconds)

Fig. 12 Open comparison of Office apps

Scenario 3: Server Virtualization (VMware) for Higher IT Efficiency

Server virtualization can dramatically improve enterprise IT efficiency and elasticity. However, storage is the biggest bottleneck for a system to achieve centralized business deployment and maximum resource utilization. OceanStor Dorado all-flash storage can cope with high-performance challenges in server virtualization scenarios, providing enterprises with unparalleled service convenience.

Challenges of VMware Server Virtualization

Service consolidation

The continuous expansion of businesses (application types and unit capacity performance) requires extensive integration and scalability from the virtualization platform. Common services of server virtualization include Microsoft services, databases, web applications, and OA applications. Mainstream servers with high loads require about 2,000 to 5,000 IOPS/TB, and medium-load applications require 500 to 1,000 IOPS/TB. However, the performance density of traditional storage is unable to meet these requirements due to its low 300 to 500 IOPS/TB.

Lengthy development and testing decrease performance

The development and testing system of the virtualization platform needs to quickly obtain production data copies with no impact on production businesses. Traditional replication at the application layer is time-consuming, taking about 10 days to clone 140-TB data. This can seriously stall development and testing. If a problem occurs in the production environment and needs to be reproduced in the development and testing environment, 10 days is an unacceptable delay. In addition, data replication at the application layer consumes production system resources, adversely affecting production performance and user services.

Intensive time and labor investment for O&M

During data center virtualization, almost all enterprise applications run on VMs and containers. The internal data center of a large enterprise usually has hundreds of VMs, but VM O&M requires a huge investment of time and labor. Consequently, in many scenarios VM O&M has become a bottleneck for enterprise service development.

Server Virtualization Solution

The all-flash storage of Huawei OceanStor Dorado provides powerful performance, large-scale service integration, and ultra-low latency. It effectively manages the high-performance and density requirements of server virtualization and the workloads of multi-business scenarios.

Fig. 13 shows typical VSI networking. The upper-layer application is a VMware HA cluster with dual switches and Huawei OceanStor Dorado.
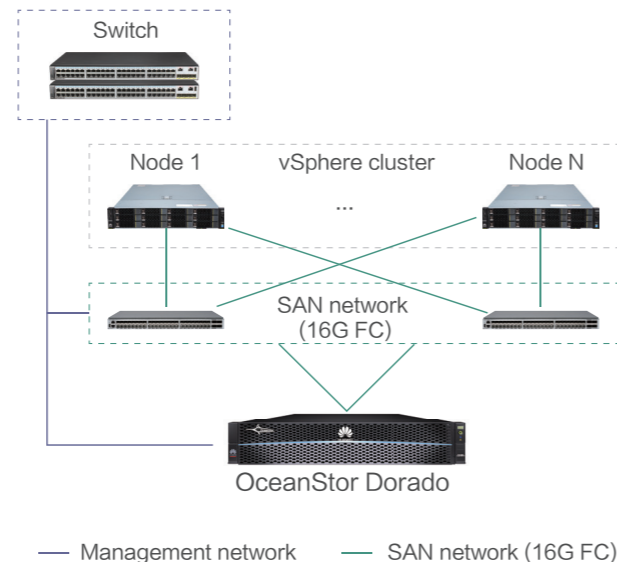
Fig. 13 VSI networking

Hardware	Model	Quantity
Rack server	RH 2288 V3	2
FC switch	SNS3096	2
Ethernet switch	S5700	1
Storage	OceanStor Dorado 6000	1

Table 7 Hardware configuration

Software	Product	Version
Virtualization	VMware vSphere	6.7
Client OS	Windows Server	2012 R2 x 64

Table 8 Software configuration

VM Component	Specifications
vCPU	4
Memory	8 GB
Drive	50 GB

Table 9 VM configuration

Better business integration capability for VMs of mission-critical businesses

Compared with traditional storage, Huawei OceanStor Dorado provides 10,000 IOPS at 100 GB/VM, which can carry 800 VMs at 100 GB. As a result, the VM benefits from a six-fold improvement in bearing capability for service integration tailored for heavy-load virtualization.
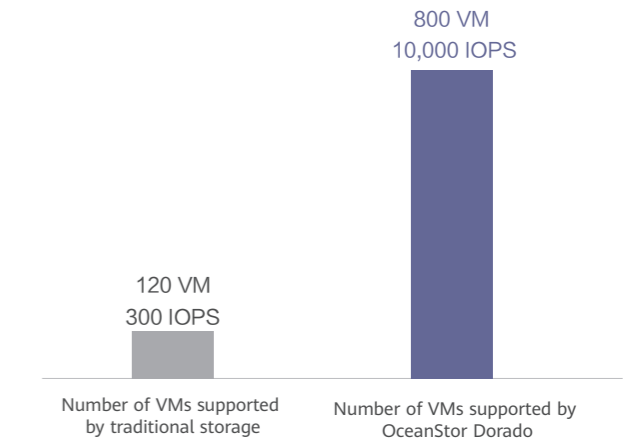
Fig. 14 Performance density comparison

20x efficiency with development and testing environment deployment in 10 minutes

The high-density snapshot of Huawei OceanStor Dorado deploys a development and testing environment within minutes. In addition, development and testing data can be updated at any time from production data. Table 10 compares Huawei OceanStor Dorado with traditional storage from vendors in the time required for deploying a development and testing environment on the existing network of user A.

System	Production Data	Traditional Backup Deployment	HyperCDP Deployment
ERP	140 TB	10 days	10 minutes
HRMS	180 TB	12 days	10 minutes

Table 10 Development and testing deployment comparison

User A generates a data copy on the backup storage by copying data at the application layer to avoid impacting the production environment. This solution requires two sets of storage devices, one for production and the other for backup. Asynchronous remote replication is implemented between the two sets of storage devices. In addition, LUN clone or host application layer copy is used on the backup storage to generate the data copies required by the production development and testing environment. It takes more than 10 days to copy data at the application layer. If a LUN clone is used, it takes several hours. If HyperCDP and data protection consistency groups are used, it takes **only 10 minutes** to deploy multiple development environments. In addition, HyperCDP can be performed **every 3 seconds** for data upgrades in the development and testing environments at any time. In contrast, updates in the traditional solution repeat all processes, requiring intensive time and labor investments. **The all-flash storage of Huawei OceanStor Dorado has demonstrated its leading role in development and testing environment deployment.**

30 VMs deployed within 14 seconds for 40x O&M efficiency

The most vital responsibility of virtualization administrators is to deploy VMs for new services, and the OceanStor Dorado streamlines this process. OceanStor Dorado optimizes vStorage APIs for Array Integration (VAAI) and deploys VMs in just seconds with XCopy, reducing the rollout time for the new service. As shown in Fig. 15, it takes only 3 seconds to deploy a VM and 14 seconds to deploy 30 VMs. If XCopy is disabled, 810 seconds are required to deploy 30 VMs. OceanStor Dorado is 40x more efficient in VM deployment.
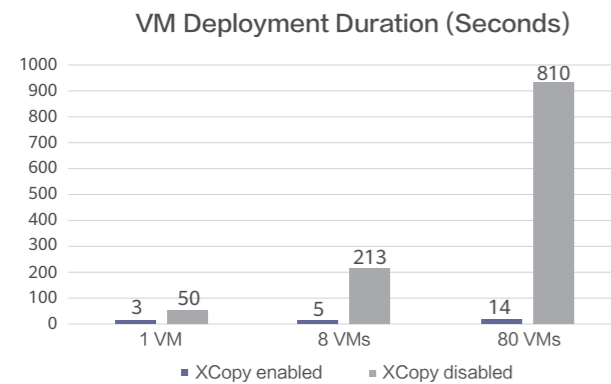
VM Deployment Duration (Seconds)

	1 VM	8 VMs	80 VMs
XCopy enabled	3	5	14
XCopy disabled	50	213	810

Fig. 15 VM deployment using a template

Enabled XCopy takes only 3 seconds to migrate one VM and 11 seconds to migrate 30 VMs. When XCopy is disabled, it takes 946 seconds to migrate 30 VMs. Thanks to the powerful performance of OceanStor Dorado and the customized optimization for virtualization O&M, VM migration is nearly finished in the time it takes users to switch between pages. There is a 90x improvement in migration efficiency. As a result, the overall O&M efficiency of VMs is greatly improved.

The all-flash storage of Huawei OceanStor Dorado features Huawei-developed chips and excellent software and hardware optimization capabilities. It is a perfect choice for deploying mission-critical applications and virtualization at enterprise data centers. OceanStor Dorado perfectly resolves the pain points in storage performance and O&M of data centers, and lays the foundation for rapid enterprise growth.

30 VMs migrated within 11 seconds for 90x the O&M efficiency

The responsibilities of virtualization administrators also include VM O&M. Extended running of services may result in an insufficient datastore capacity or unbalanced usage of datastore resources. Storage vMotion can avert these potential problems. OceanStor Dorado implements second-level migration in VM storage vMotion scenarios to greatly shorten the service migration time. Fig. 16 compares online VM migration storage when XCopy is enabled and disabled.
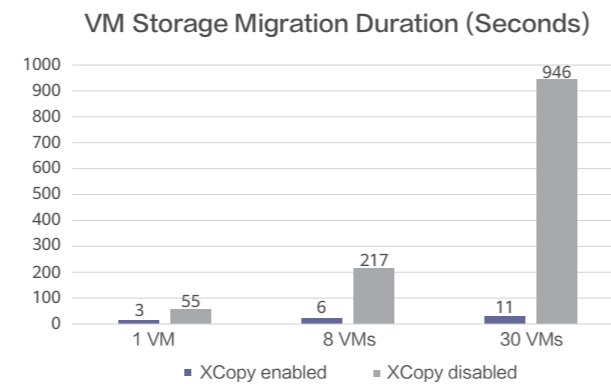
VM Storage Migration Duration (Seconds)

	1 VM	8 VMs	30 VMs
XCopy enabled	3	6	11
XCopy disabled	55	217	946

Fig. 16 VM storage migration

D Bank Upgrades to All-Flash Storage, Covering Production Services

By Ma Yan, Huawei

D Bank, headquartered in Singapore, is one of the world's leading commercial banks, and has garnered numerous awards, such as World's Best Digital Bank 2018 and World's Best Bank 2019. To stay ahead of the curve, D Bank designs future-oriented solutions with its advanced digital technologies and innovations. Its goal is to be the best digital bank in the world for years to come.

Due to the increasing service pressure in the production environment, storage on the existing network cannot meet the real-time requirements of the production system. During peak hours, low-priority services may compete for storage resources with high-priority services. In addition, the high concurrent access to the production system requires a robust storage system that can balance workload and ensure 24/7 stable running.

For the development and testing environment, because D Bank launches a large quantity of innovative marketing services every year to continuously increase profits. The development period of new services must be short and new services need to be quickly launched. This poses high requirements on the construction and reuse efficiency of the development and testing environment. The customer wants a platform that can create services in minutes and use one copy of data for multiple systems, to ensure smooth service rollout at a lower total cost of ownership (TCO).

The bank also requires a storage system with superb performance to cope with the pressure of rapid service growth and service development, and improve the return on investment (ROI). Huawei's OceanStor Dorado high-end all-flash storage can deliver up to 20,000,000 IOPS and 0.1-ms latency, far outperforming peer products. At the end of October, 2019, OceanStor Dorado was tested at the POC site. The test networking and configuration are as follows.
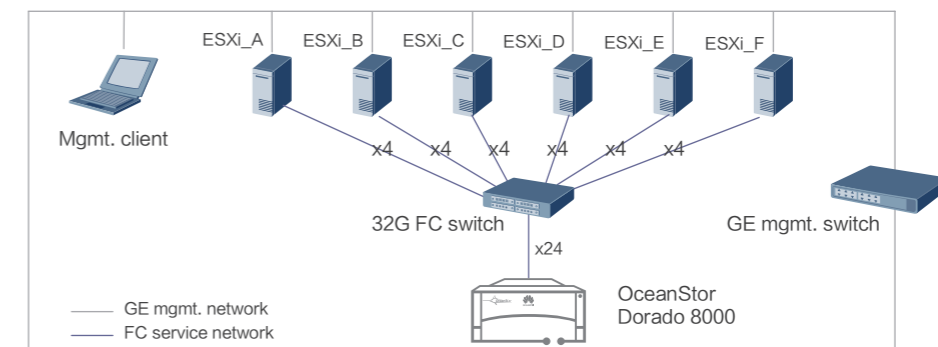

Fig.1 Performance test networking

Name	Description	Quantity
OceanStor Dorado 18000	OceanStor Dorado 18000 all-flash storage with four controllers	3
2 U SSD enclosure	2 U SSD enclosures on OceanStor Dorado 18000 all-flash storage	3
32 Gbit/s FC front-end interface module	4 x 32 Gbit/s FC front-end interface modules	16
100GE back-end module	2 x 100Gb RDMA back-end interface modules	6
GE front-end interface module	4 x 1 GB front-end GE interface modules	4
SSD	Huawei 2.5-inch 7.68 TB SSDs	75

Table 1 Test configuration

Test: The Vdbench software simulated production services (8 KB data blocks, 100% random, 7:3 mixed read/write, and raw capacity). The performance of four controllers reached 2,139,636 IOPS, and the latency was stable at 0.996 ms. Under the same workload, the available-capacity service model reached 1,660,790 IOPS, over twice that of competitors with the same configuration.
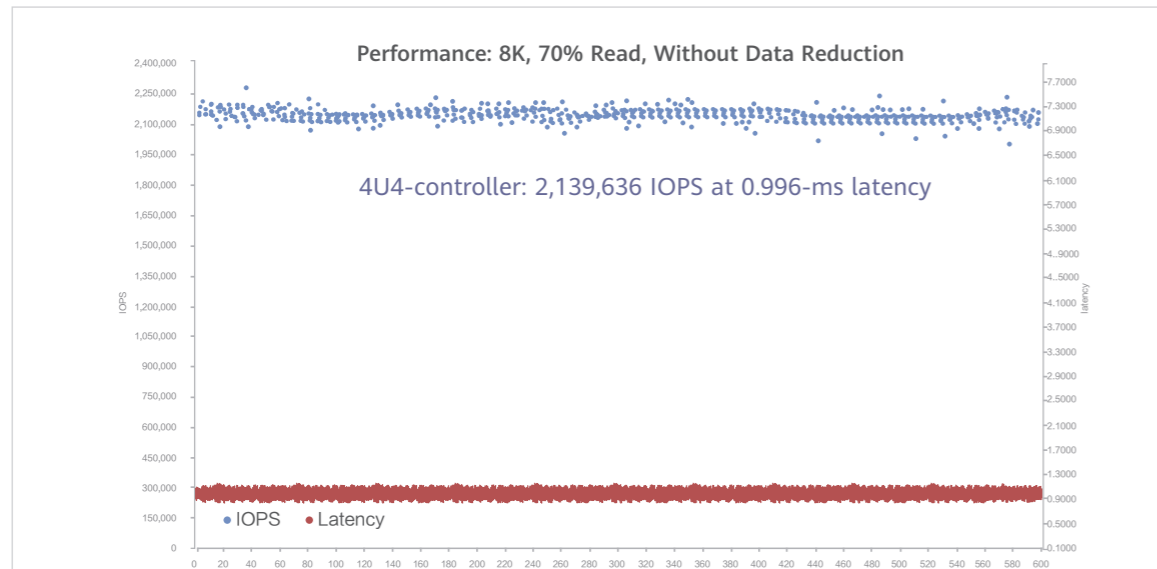

Performance: 8K, 70% Read, Without Data Reduction

4U4-controller: 2,139,636 IOPS at 0.996-ms latency

● IOPS   ● Latency

Fig. 2 Performance and latency curves of OceanStor Dorado 8000 (raw capacity) service model

Performance: 8K, 70% Read, with Data Reduction

4U4-controller: 1,660,790 IOPS, twice better than that of other vendors
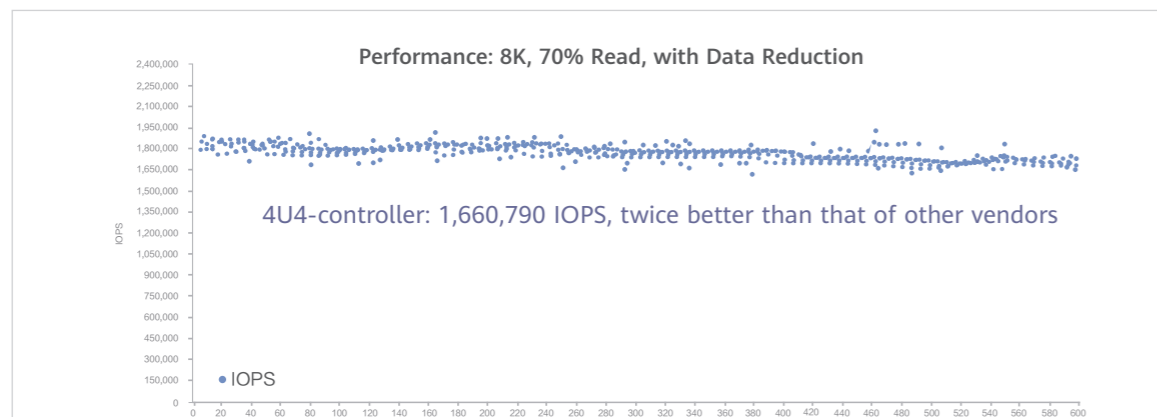
● IOPS

Fig. 3 Performance curve of OceanStor Dorado 8000 (available capacity) service model

The results astounded the customer, who raised the following questions about its excellent performance:

- Can the performance support key services during peak hours?
- Can the storage system balance loads under heavy service pressure?
- Can services keep running even if multiple storage controllers are faulty?

Test data 1: Vdbench simulated production services (8 KB data blocks, 100% random, 7:3 mixed read/write). As shown in Fig. 4, the initial IOPS of a single controller was 28,000. However when SmartQoS was enabled, the IOPS of a single controller decreased to 18,000. As shown in Fig. 5, SmartQoS set the burst IOPS to 97,500 and maximum IOPS to 78,000. The burst IOPS was kept for a period of time, evened out, and then became stable at the maximum IOPS. The SmartQoS feature met the performance requirements of key services at peak hours.

Per-Controller-IOPS

30,000 IOPS before limitation

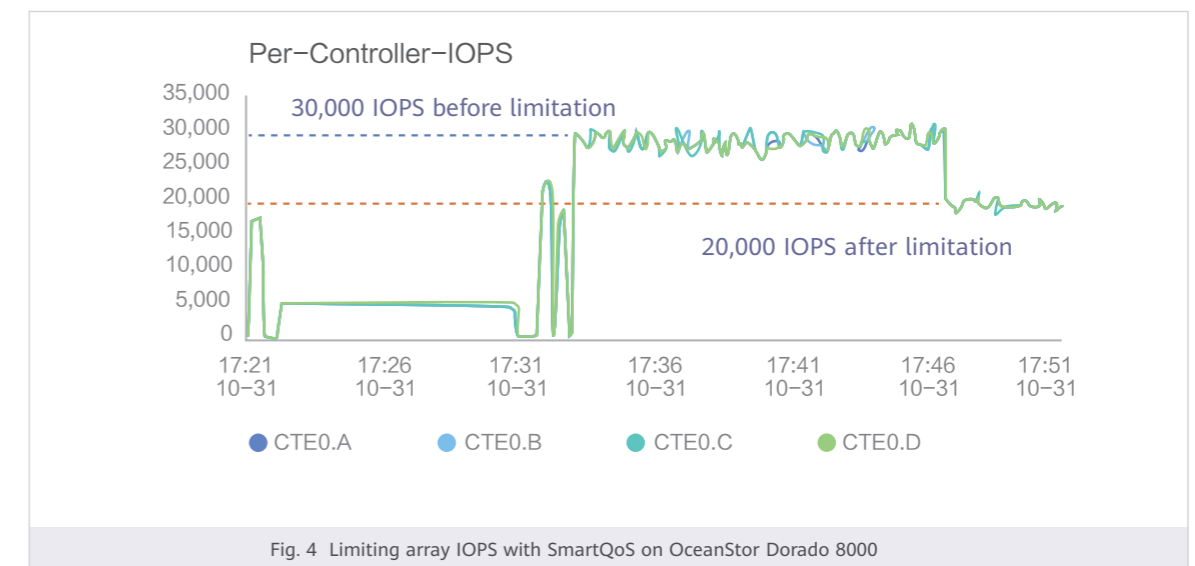20,000 IOPS after limitation

● CTE0.A   ● CTE0.B   ● CTE0.C   ● CTE0.D

Fig. 4 Limiting array IOPS with SmartQoS on OceanStor Dorado 8000

Normalized IOPS

● Max   ● Burst   ● Normalized IOPS   Avg:53,167,232 IOPS

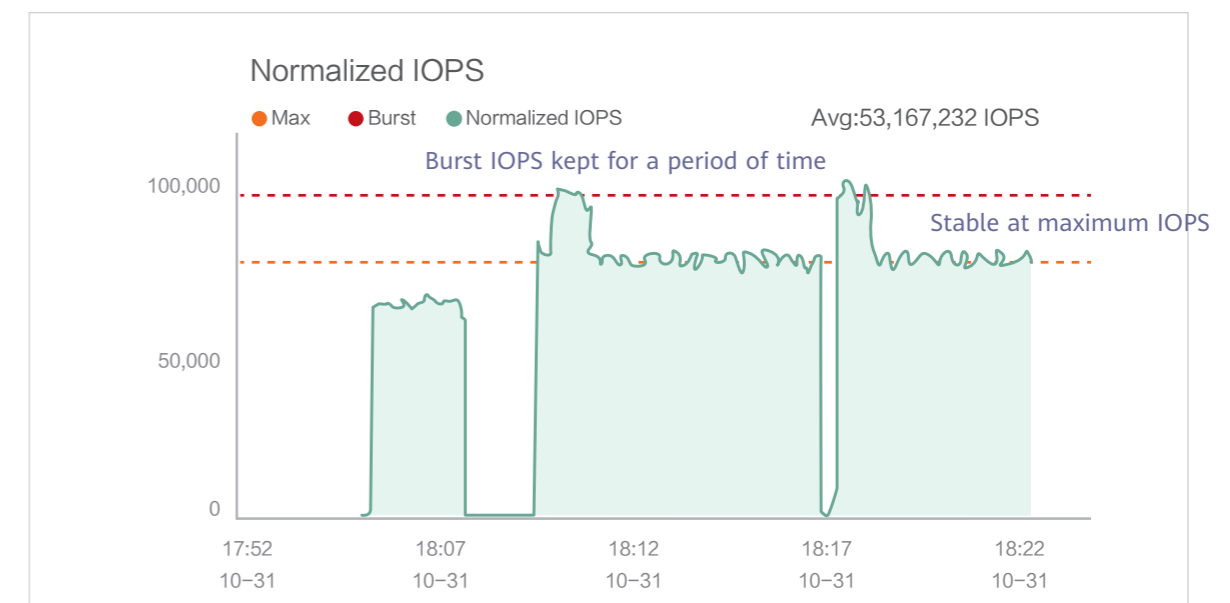Burst IOPS kept for a period of time

Stable at maximum IOPS

Fig. 5 Limiting burst IOPS and maximum IOPS with SmartQoS on OceanStor Dorado 8000

Test data 2: Vdbench simulated production services (8 KB data blocks, 100% random, 7:3 mixed read/write). As shown in Fig. 6 and Fig. 7, the IOPS and CPU usage of each controller were the same, that is, 25,000, and 15%, respectively. The end-to-end active-active architecture achieves global load balancing. The CPU usage and IOPS load balancing rate of the controllers were both greater than 98%.
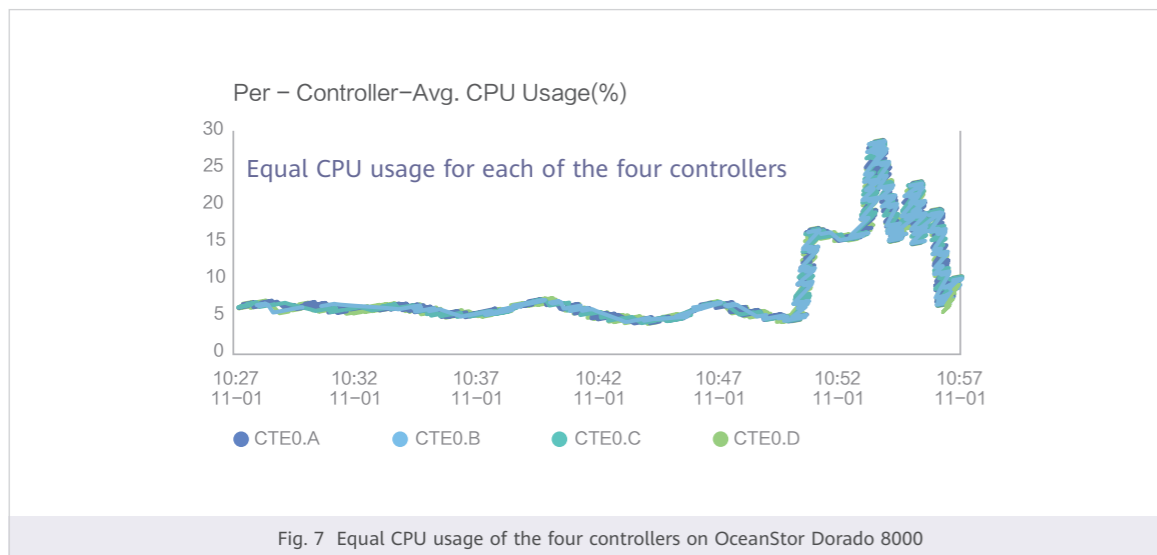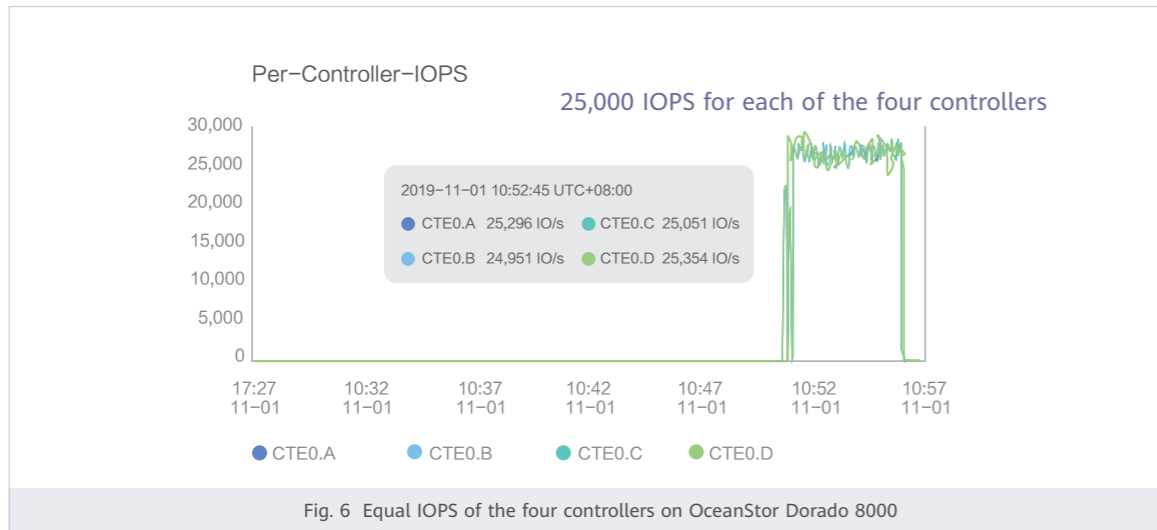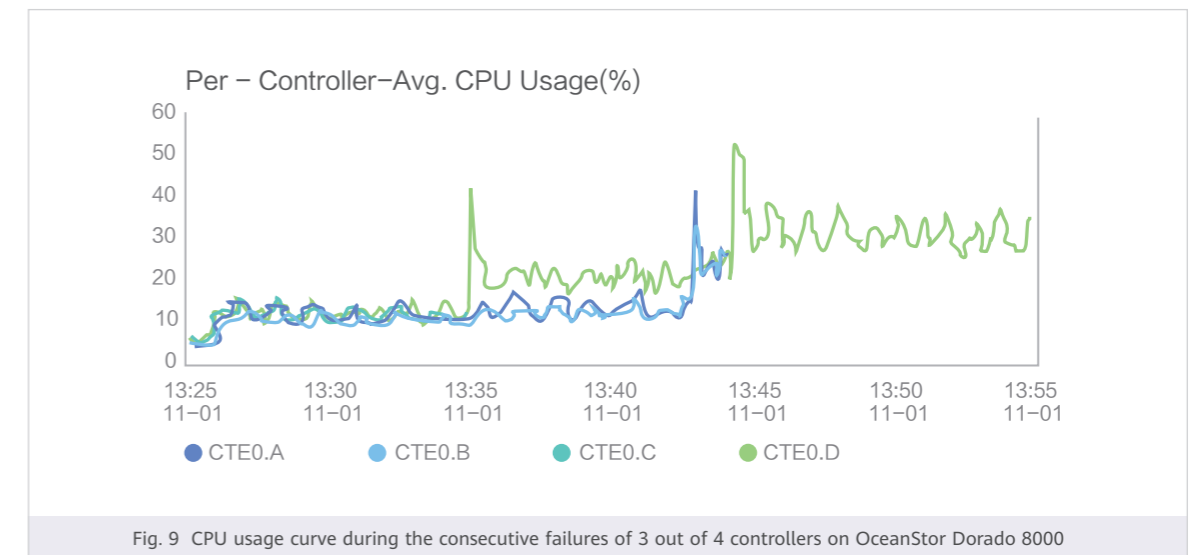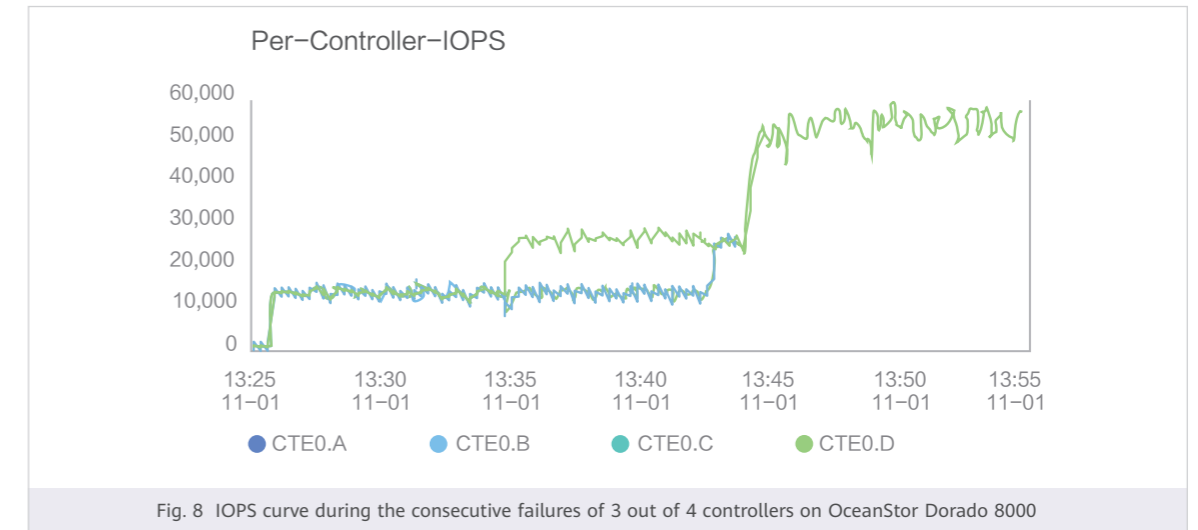
Test data 3: Vdbench simulated production services (8 KB data blocks, 100% random, 7:3 mixed read/write). As shown in Fig. 8 and Fig. 9, four controllers operated properly between 13:25 and 13:35, and the IOPS of a single controller was 12,000.

After controller C was removed at 13:35, controller D took over services from controller C. The performance of controller D increased to 24,000 IOPS, and that of controllers A and B remained 12,000 IOPS.

After controller A was removed at 13:43, controller B took over services from controller A. The performance of controller B increased to 24,000 IOPS.

After controller B was removed at 13:45, controller D took over services from controller B. The performance of controller D increased to 48,000 IOPS.

During the period, the IOPS and CPU usage of each controller were stable.

Fig. 6 Equal IOPS of the four controllers on OceanStor Dorado 8000

Fig. 8 IOPS curve during the consecutive failures of 3 out of 4 controllers on OceanStor Dorado 8000

Fig. 7 Equal CPU usage of the four controllers on OceanStor Dorado 8000

Fig. 9 CPU usage curve during the consecutive failures of 3 out of 4 controllers on OceanStor Dorado 8000

Test data 4: Vdbench simulated production services (8 KB data blocks, 100% random, 7:3 mixed read/write). As shown in Fig. 10 and Fig. 11, services operated properly after controllers A and C were removed at 14:25. The IOPS of a single controller was 25,000.

After controllers A and C were recovered at 14:50, the four controllers shared the service pressure of the original two controllers. The IOPS of a single controller decreased to 12,500. During the period, the IOPS and CPU usage of each controller were stable.
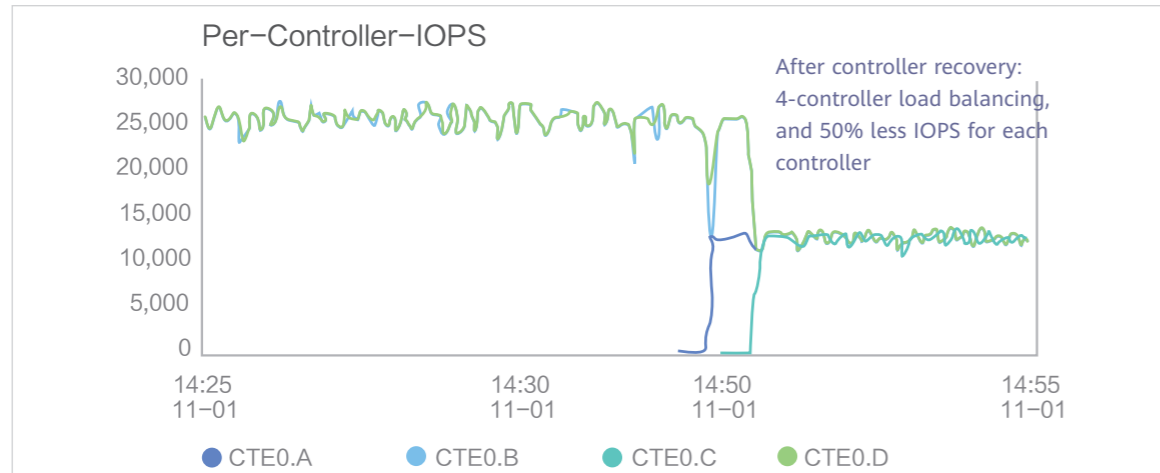

Fig. 10 IOPS curve during the simultaneous failures and recovery of 2 controllers on OceanStor Dorado 8000
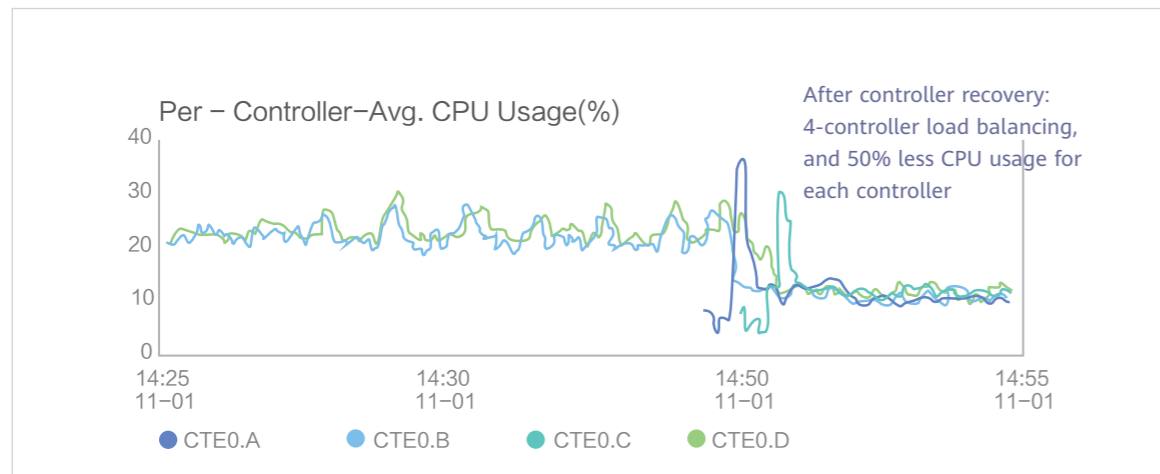

Fig. 11 CPU usage curve during the simultaneous failures and recovery of 2 controllers on OceanStor Dorado 8000

The POC test successfully verified that Huawei's OceanStor Dorado high-end all-flash storage provides ultimate read/write performance, service quality control, system load balancing, and tolerance for simultaneous failures of dual controllers/consecutive failures of three out of four controllers in the simulated production scenario. These results made it an easy decision for D Bank to choose Huawei as its strategic partner to run the premium OceanStor Dorado 8000 for its production systems. This will set up D Bank as a next-gen digital bank for years to come.

Flash Upgrades **Ensure Stable Peak-Hour Operations** for V Bank in Russia

By Guo Kaizhen, Huawei

Despite the speed of growth slowing down for the global economy, Russia's economic growth remains dynamic. That's the view of Russian President Vladimir Putin, who was speaking at the Russia Calling Investment Forum. All of this is closely related to the continued business growth of V Bank. As a leading universal bank in Russia, V Bank provides extensive banking services and products in Russia, Commonwealth of Independent States (CIS), Europe, Asia, Africa, and the United States.

However, it was no all plain sailing for V Bank. While business grew steadily, the processing capability of the legacy IT infrastructure in data centers (DCs) was unable to keep pace with the service growth due to rapid large-scale expansion. The legacy infrastructure had many issues, such as lack of data sharing due to information silos, high total cost of ownership (TCO) for storing redundant data, and insufficient workload processing during peak hours. All resulted in poor user experience. Digital services and infrastructure caused unprecedented challenges to the bank. It lacked a reliable disaster recovery (DR) solution as the legacy option lacked backup models among multiple DCs, and was unable to efficiently purchase a single IT device without impacting business continuity. All these factors meant the long term strategy of the bank needed reviewing, and it was decided that digital transformation would be a top priority.

"Both China and Chinese companies are strategic partners for Russia. I've met Chinese companies many times from the transportation and finance fields. We are digitalizing banks. For this project and the next, Huawei is our natural digital partner."

—From President of V Bank Group

After several rounds of in-depth meetings, the OceanStor Dorado 6000 all-flash storage was determined to be a very appropriate product for the bank's digital goals. Based on the ever-evolving all-flash architecture, the OceanStor Dorado with dual controllers delivers ultra-high performance of hundreds of thousands of IOPS. It is well regarded in many industries, and is known as a star product in the storage industry thanks to its 0.1 milliseconds (ms) latency and 99.9999% reliability. Leading software features meet the requirements of existing applications, and the scale-out and scale-up capabilities are sufficient to cope with the service growth in the next 5 to 10 years.

"Does your storage really double the speed of similar products from other vendors?" V Bank's CTO doubted.

"The results speak for themselves. Let's take the test!" Then Huawei invited the technical team of V Bank for the proof of concept (POC) test and compete with peer products from other vendors. The test covered:

· The system latency and IOPS, and if these could meet the upper-layer service requirements under the normal workload pressure of the database.

· The system's stable running during peak hours (service burst in a short time).
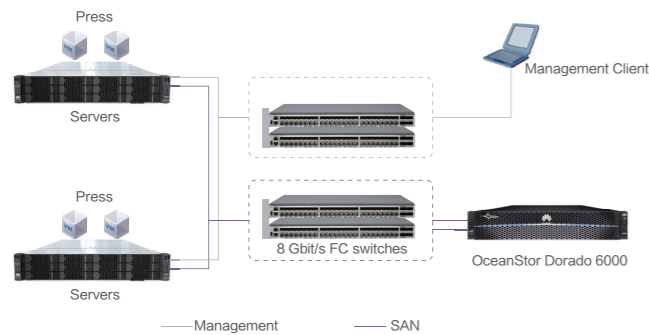

Fig. 1 POC test networking

Type	Name	Quantity
Test server	RH5885 V2 servers	4
Test storage	OceanStor Dorado 6000	1
Test software	Vdbench	20
	Redhat Linux 6.5 x64	10
	Windows Server 2012	10

Table 1 POC test configuration

A foreign trade bank, unlike a common bank, has high requirements of stable system throughput when processing loads of transactions due to time difference. During the test, the software Vdbench is used to simulate the database OLTP workloads. Conditions include: 8 KB data blocks, 100% random, and 7:3 read/write, and 16 LUNs. After running for 20 minutes, OceanStor Dorado 6000 with two controllers delivered 500,000 IOPS and a stable latency of 0.3 ms. The services run for three days without any disruption fluctuation, and all tasks were completed successfully. The system stays well ahead of other products, which impressed the banks representatives.
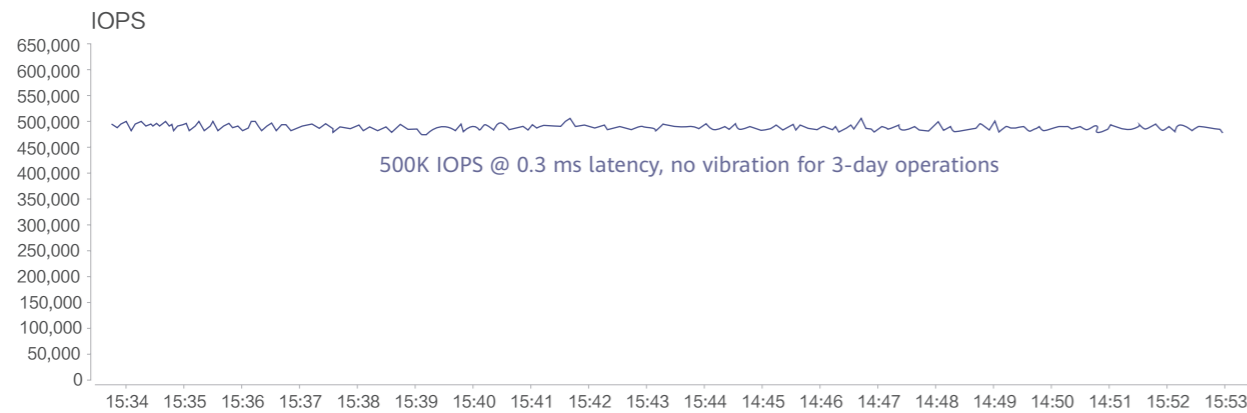

Fig. 2 Performance test - IOPS

The second part tested the system capabilities of coping with data bursts. One of the biggest risks for a foreign trade bank is the exchange rate risk, which refers to the risk that changes in the relative value of certain currencies will make the settlement of international credit and debt out of control. Foreign exchange involves the credit and debt process in the international financial market by which, if the foreign exchange rate increases, the debtor will suffer great losses. The sharp changes of exchange rate can make or break large enterprises. The fluctuation also directly affects the value of the foreign exchange reserves for a country, which poses huge difficulties to any central banks.

For banks, time is money. If global exchange rates fluctuate abnormally, accounting for the existing funds and deriving all profit and loss scenarios are required to help executives analyze accordingly. If the IT system cannot predict and analyze the risks as and when needed, the impact can be catastrophic. Latency is the lifelines of storage systems that host critical banking services. It is an indicator that will never be perfect for vendors. During the tests, Vdbench simulated a burst access scenario of the database (8 KB data blocks, 70% random, and 7:3 mixed read/write). Despite the demanding environment, OceanStor Dorado 6000 still reached 600,000 IOPS with 0.5 ms latency, and could still host more than twice of workloads on the legacy network.
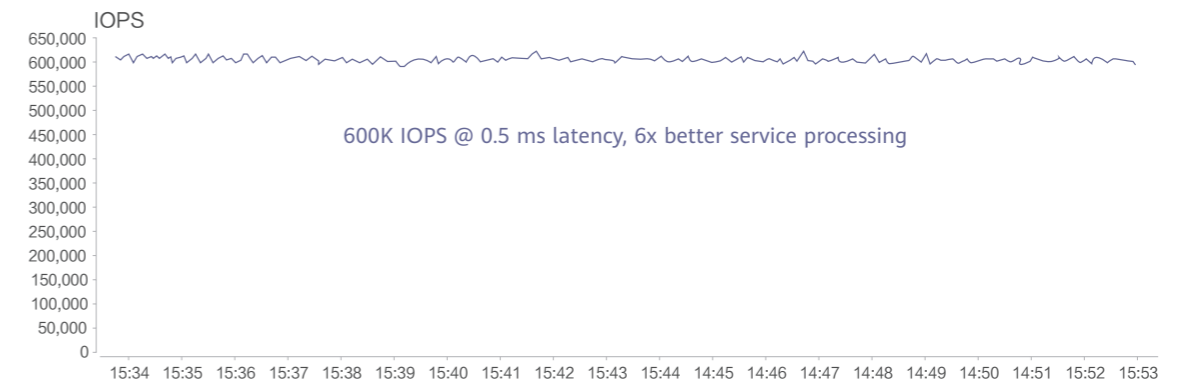
Fig. 3 High pressure scenario

Huawei OceanStor Dorado 6000 ranked first in the POC test with proven performance and stability, far exceeding the user expectations. At the beginning of the first phase for the DC reconstruction, the bank purchased OceanStor Dorado 6000 with a total capacity of 900 TB. The overall service processing capability is expected to improve six fold compared to the equipment on the legacy network, which put many of V Bank concerns at rest. Since then, Huawei OceanStor Dorado has been running stably, and was given an endorsement by business director of V Bank, who stated how satisfied the bank has been with Huawei storage. In the second phase, the bank plans to keep purchasing Huawei OceanStor all-flash storage while doubling its business scale.

HUAWEI

4 Enjoying Success

Huawei OceanStor AFAs
Story Sharing

Transforming Citizen Wait into Citizen Satisfaction

Huawei OceanStor Dorado improves taxation system performance by **3x**

1145 government customers from 114 countries (non-china) choose Huawei all-flash storage to accelerate digital transformation and innovation

Bank of Jiangsu Deepens Cooperation with **Huawei**, Leading Transformation of Urban Commercial Banks

Guotai Junan securities Improves IT Infrastructure Service with Huawei OceanStor AFAs

Carrier MTS Supercharges Efficiency by 5x and Reduces TCO by 35% with OceanStor AFAs

Bank of Jiangsu Deepens Cooperation with Huawei, Leading Transformation of Urban Commercial Banks

By Wang Zhen, Huawei

Following groundbreaking scientific and technological progress in recent years, China has climbed to one of the largest economic entities in the world. This rapid development is in part thanks to the country's digital economy, which serves as a new momentum for banks and other financial services.

Bank of Jiangsu was officially listed and opened on January 24, 2007. Adhering to the mission of "creating a better life" and the core values of "integration and innovation, pragmatic responsibility and lean growth", Bank of Jiangsu is committed to building a featured and comprehensive best-in-class commercial bank capable of delivering a smarter lifestyle with FinTech. As the largest legal-entity commercial bank in Jiangsu province, the bank ranked 91st in the 2018 global bank ranking of Banker (a British English-language monthly international financial affairs publication). It is one of the highest ranking banks on the Forbes list and has been rated as the "Most Competitive Small and Medium Bank" and "Most Innovative Bank" by Financial Times, as well as the "Best Urban Commercial Bank in China" by Global Finance.

Rapid development and maturity of technologies, such as artificial intelligence (AI), big data, and all flash, have expedited digitalization in the financial industry. Interactive real-time banking services are transforming IT technologies in banking production systems, and banks like Bank of Jiangsu are having to adapt just to keep up.

Core Infrastructure Faces Severe Challenges

To further facilitate the transformation and upgrade of the real economy, better meet the business development requirements in the digital transformation era, and accelerate the construction of intelligent, characteristic, internationalized, and integrated banking services, Bank of Jiangsu was in urgent need of a next-generation data infrastructure architecture. However, the bank's legacy IT system faced the following challenges:

1. How to enhance reliability of user data

It is a top priority for the bank to build a reliable IT infrastructure providing 24/7 secure, trustworthy, and sustainable services.

2. How to achieve optimal user experience

In the mobile Internet era, demanding data volumes and requirements on processing efficiency of core systems have caused the bank to review its processes. Specifically, maintaining stable IT systems' services in massive concurrent transactions, ensuring quick response to transaction operations of each end user, and providing better service experience are all crucial to development.

3. How to improve the ROI

The bank needs to reduce the overall investment cost and its storage architecture must be future-proof and scalable to meet service requirements over the next five years. New services, such as smart retail and vehicle life platform, require a new scalable architecture that enables new services to be rolled out quickly, making the traditional IT architecture of the bank outdated.

Best-fit Solution Targets to Fast, Stable, and Accurate Planning and Implementation

Serving as the lifelines of a bank, IT systems running core banking systems, integrated front-end systems, debit cards, and payment systems have always been the top priority.

In 2015, Bank of Jiangsu chose Huawei to play an important role in its construction of the next-generation IT system. Huawei storage covers O&M and peripheral production systems at the earliest stage to the full deployment of core systems now. This long-term partnership displays the cooperation between both parties and is testimony to the innovation and success of Huawei storage.

At the end of 2018, Bank of Jiangsu started core infrastructure reconstruction to improve the reliability and transaction processing performance of its core system. In January 2019, the IT department of Bank of Jiangsu set up a core infrastructure reconstruction team who worked on analyzing and determining reconstruction principles, device and solution suppliers, and technical architectures until the core system was officially launched at the end of October 2019. The entire process faced tight schedules and heavy tasks.

Large-scale projects, such as core system rollout, usually take at least one to two years to complete. However, due to the unremitting efforts of Bank of Jiangsu, Huawei, and other partners, the project was completed within just 10 months. Data of dozens of key databases and application systems, including core standard operating procedures (SOPs), e-banking, mobile banking, and gen-2 payment systems, was migrated to a secure and reliable storage system. This ensures top-class data storage and management to effectively support service development over the next three to five years. In practice, Bank of Jiangsu has set an example of how to plan, construct, and deploy an innovative storage system and service platforms and become the first bank to apply innovative storage technologies to its core system among trillions of urban commercial banks.

Bank of Jiangsu Cooperates with Huawei to Build a Secure and Reliable Mission-Critical Service System

Core transaction service continuity and reliability of IT systems are the top priorities of banks. Huawei high-end all-flash storage system leverages the industry-leading SmartMatrix architecture to effectively ensure the reliability of a single system. In addition, the gateway-free active-active and geo-redundant data protection solutions ensure service continuity from the system architecture to the DR solution, laying a foundation for the first attempt to apply innovative storage technologies to the core system of a bank.

1. Rock-solid security and reliability

Banks are in urgent need of lightning-fast and rock-solid storage for their core service systems. Huawei OceanStor storage systems use all-flash and multi-controller technologies to eliminate performance bottlenecks and deliver superb performance. After Bank of Jiangsu uses Huawei OceanStor high-end all-flash storage system, the service processing performance of its core system is accelerated by over 40%. During peak hours, the storage performance of AWR for mobile banking is remarkably accelerated, and the latency is reduced more than one-fold compared with the original architecture.

2. Rapid service rollout and superior experience

Thanks to the great improvement in data processing efficiency, the core system effectively supports the rapid growth of new Internet financial services of the bank. After the new system is launched, customers on the Che Sheng Huo (vehicle-related services for easier life) platform increase exponentially. In addition, the Ai Xue Xi (iLearning), Ai Jian Kang (iHealth), and E Rong Zhi Fu (E-Loan Payment) platforms also witness quick increases on contracted institutions and registered customers. The Shui E Rong (Tax E-Loan) service has been implemented and upgraded. The next-generation core storage system effectively accelerates process optimization, service rollout, and service migration as well as quickly responds to massive concurrent transactions, delivering better user experience.

3. Effective investment protection

The performance and capacity of the storage system can be expanded as required, effectively supporting direct banking services for users (30 million as of 2018) to hundreds of millions of users in the future and protecting long-term investments of banks.

FinTech has become a bridge closely connecting banks and customers. Bank of Jiangsu will continue cooperating with Huawei to promote service innovation, supply-side reform, and digital transformation, as well as build a featured and comprehensive best-in-class commercial bank capable of delivering a smarter lifestyle with FinTech.

Guotai Junan Securities Improves IT Infrastructure Service with Huawei OceanStor AFAs

By Gao Xun, Guotai Junan IT Technical Director

In August 2019, an article entitled Guotai Junan Securities Introduces Huawei OceanStor Dorado AFAs to Remove Data Storage Bottlenecks was a resounding success with positive feedback from both the global market commentators and competitor executives. This article was just a reflection of how technology developed domestically effectively builds IT infrastructure service.

Guotai Junan Securities' IT development strategy is to independently innovate, deliver personalized products and services, and build service for the entire IT line. The IT technical platform domestically constructed with innovative technologies matches both the national development strategy and our own IT development strategy. Building service for our IT infrastructure platform is integral to our development. At an IT conference held at the beginning of 2018, the Guotai Junan CIO Yu Feng stated that IT infrastructure must lead the development of IT hardware technologies and an IT infrastructure platform must consist of network communication, storage resource pools, cloud computing, and IT security. Over the past two years, Guotai Junan has greatly improved IT infrastructure service by adopting innovative technologies.

Huawei OceanStor AFAs Improve Database Service for Critical Application Systems

Centralized storage is an important IT infrastructure platform for Guotai Junan Securities. Data of important application systems is stored on centralized storage. The read and write performance of centralized storage is one of the vital factors that directly affects data processing for application systems. This is especially true when the volume of application data is large.

After running strict proof of concept (POC) tests and rigorous technical analysis, we ultimately chose Huawei OceanStor Dorado AFAs to replace the traditional HDD storage devices of IBM, Dell EMC, and other brands. Huawei OceanStor Dorado AFAs store data of core application systems, such as centralized transactional and comprehensive

wealth management systems.

Huawei AFA benefits:

The centralized transactional system only takes 3 hours to batch process historical data for a 40% improvement in efficiency.

The response time of production databases for comprehensive wealth management is shortened from 5 ms to less than 1 ms.

OceanStor Dorado AFAs ensure stable and prompt responses even if the service volume of Junhong App increases 10-fold or high concurrency scenarios occur dealing with 500,000 transactions per second (TPS).

OceanStor Dorado AFAs reduce the equipment footprint by 85% with one AFA only requiring 2 U of cabinet space, compared to the 13 U needed by an HDD storage system.

OceanStor Dorado AFAs significantly drive down cooling costs and reduce power consumption by 71%.

They are also projected to reduce total maintenance costs by more than 70% in five years.

OceanStor Dorado AFAs have been operating without problems, further enhancing our confidence in Huawei storage devices for core applications systems and accelerating our adoption of AFAs to host all critical applications. We also use the latest OceanStor Dorado AFAs from Huawei to accommodate the next-generation clearing system and IQ databases for big data. Real-time test results confirm that the performance of the AFAs of Huawei OceanStor Dorado is more than double that of AFAs from international vendors such as Dell EMC and HDS.

Consequently, more than one third of Guotai Junan's data centers are using Huawei OceanStor Dorado AFAs.

HyperMetro Gateway-Free Active-Active Solution for 24/7 Stable Operations of the Financial Cloud Platform

Cloud-ready data centers are the wave of the future. They can implement standard construction of IT infrastructure resource pools, quick resource delivery, self-help cloud services, and centralized management and control, as well as transparent use of heterogeneous resource pools. They greatly improve the utilization of data center infrastructure and infrastructure resources and reduce the total cost of ownership (TCO). The accelerated construction for the financial cloud over the past two years has resulted in the migration of more than 80 production application systems to the cloud.

But how do we ensure high availability of cloud resource pools? We use the fully developed commercial virtualization platform and the HyperMetro gateway-free active-active solution powered by OceanStor Dorado AFAs in the production service cloud. This allows us to build cloud resource pools and self-healing for resource pools. In addition, optimistic lock technology reduces I/O interactions between two sites, improving the overall performance. At the data layer, two sets of active-active storage devices are deployed between virtual clusters to provide read and write services for each cluster. If a storage system at one site fails, the other system can seamlessly take over services for zero data loss, zero service interruption, and 24/7 stable operations of the financial cloud platform.

The benefits of Huawei's innovative technologies don't end there. They also help us improve IT infrastructure service capabilities, such as remote management, unified O&M, operational audit management, and unified data protection management. Compared to similar products from other large international companies, Huawei OceanStor Dorado AFAs are much more cost-effective. Most importantly, OceanStor Dorado AFAs fit our overall IT development strategy.

In the future, we will continue to capitalize on the advanced achievements of Chinese innovation, and continuously improve the service of IT infrastructure for unified resource pool construction, quick delivery, and security control.

Union Hospital Achieves 3x More Responsive HIS with OceanStor All-Flash Active-Active Solution

By Wang Jiaxin, Huawei

"Hospital information system (HIS), at the heart of healthcare IT systems, is crucial to any treatment, and is used as early as registration and diagnosis. Because we need it to service requirements instantaneously and be available 24/7, the obvious choice for changing our infrastructure was Huawei's OceanStor Dorado all-flash active-active solution. We adopted it to host our HIS. And we saw results instantly, with speed improving by over three times, greatly helping healthcare experience. The gateway-free active-active solution ensures 24/7 stable running of core services, strengthening the hospital's healthcare efficiency and capabilities. The system has been running stably for half a year since its rollout, improving the overall efficiency."

— Zhou Bin, Computer Management Center of Union Hospital

HIS Upgrades with OceanStor All-Flash Active-Active Solution

Huawei proposed its OceanStor Dorado with the HyperMetro gateway-free active-active solution for the Union Hospital after analyzing the existing network. The solution required no changes to the hospital's architecture, databases, medical software, and O&M models. What's more, when implemented, the solution expedites I/O reads and writes speed in an end-to-end (E2E) manner to address the hospital's slow HIS responses during peak hours. Two OceanStor Dorado systems mirror with each other, meaning data can be synchronized in real time and the hosts will not be affected in the event of a failure of a single device, helping ensure the continuity of critical healthcare services.

HIS Management

| Outpatient | Inpatient | Medical supply storehouse | Pharmacy | Healthcare | Nursing |

HyperMetro gateway-free A-A

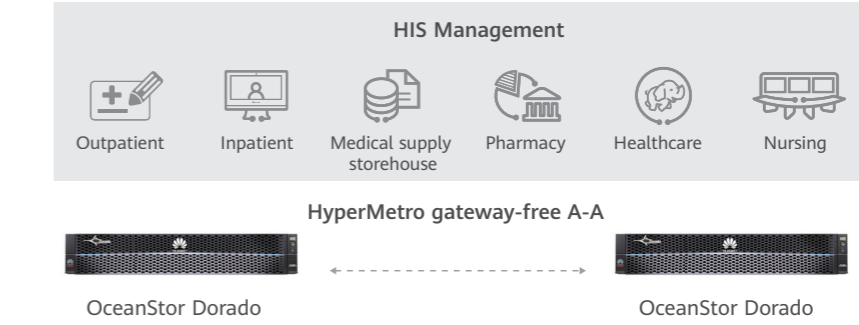OceanStor Dorado   OceanStor Dorado

Fig. 1 HIS upgrades with OceanStor Dorado active-active solution

3x Faster HIS for Improved Healthcare Experience

During peak hours, departments such as diagnosis and treatment, payment, and pharmacy complained about the slow or even unresponsive systems. The response time for each operation was even more than 30 seconds, severely affecting the individual and overall healthcare experience. After analysis, the cause of the system performance bottleneck was due to the legacy storage on the network, which could not respond to the high-concurrency requests.

The OceanStor Dorado all-flash solution greatly improves the response speed of the HIS system during peak hours, where the response speed improved by more than three times. This prevents service interruption during peak hours and allows all departments of the hospital to effectively handle the processing and treatment of patients.
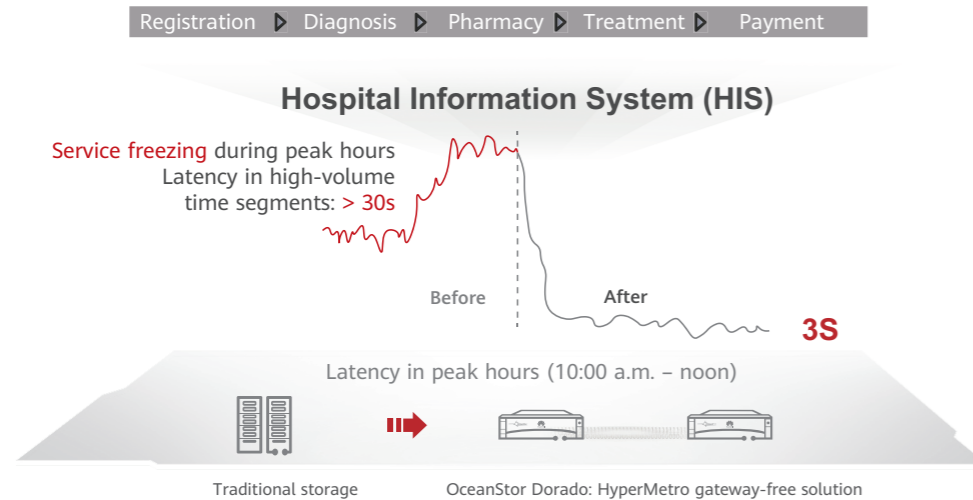
Registration ▷ Diagnosis ▷ Pharmacy ▷ Treatment ▷ Payment

Hospital Information System (HIS)

Service freezing during peak hours
Latency in high-volume
time segments: > 30s

Before   After

3S

Latency in peak hours (10:00 a.m. – noon)

Traditional storage   OceanStor Dorado: HyperMetro gateway-free solution

Fig. 2 HIS upgrade effect

Gateway-Free Active-Active Solution for Fast, Stable Operations

Rapid HIS development and service concentration have posed many operational risks for the hospital IT center. In the event the production center is hit by an upgrade/maintenance operation, power outage, fire, or natural disaster, there is a risk of halting business processes or losing customer data, which can cost hospitals dearly in terms of money and reputation. It is imperative to build a reliable disaster recovery (DR) system. Tertiary hospitals constructing DR must now adhere to the regulations of state regulatory institutions, requiring the outpatient system be recoverable within 5 to 10 minutes, and the information system available 24/7.

An active-active data center (DC) solution is ideal for maintaining high reliability of the HIS as its designed to ensure two data centers are running concurrently, and the two centers host services back up each other. If a data center fails, services are automatically switched to the other center, during which services are not interrupted with zero data loss. Although the system reliability of the previous HIS was improved thanks to the active-active solution using gateways, one additional gateway leads to another node failure risk, complicating networking and increasing the probability of manual errors. This is because an additional gateway means one more data transmission channel, which prolongs the latency while bringing down the performance of the entire IT system. Huawei's gateway-free active-active solution is an ideal choice to solve this problem. It supports load-balanced active-active mirroring and non-disruptive cross-site takeover, ensuring a zero recovery point objective (RPO) and a close-to-zero recovery time objective (RTO). The gateway-free design reduces the risks of faulty nodes, simplifying networking and management. The active-active layout can be smoothly upgraded to a geo-redundant one and to a converged data management solution without using additional gateway devices. Further, it provides solution-level availability of 99.9999% and fast, stable critical applications, ensuring that core applications are not likely to experience system breakdown.

In June 2018, Union Hospital migrated its HIS data to Huawei OceanStor Dorado. Their system has run stably since service rollout, avoiding downtime during peak hours and greatly improving the quality of healthcare.

Carrier MTS Supercharges Efficiency by 5x and Reduces TCO by 35% with OceanStor AFAs

By Zhang Di, Huawei

MTS is the largest mobile network carrier in Russia in terms of revenue and subscribers with its business in the country contributing 93% towards its total revenue. The carrier has an international customer base with approximately 78 million subscribers across Russia, Ukraine, Armenia, and Turkmenistan. In order to further increase its revenue and keep up with the crowd in the 5G era, MTS is dipping its toes into developing digital services such as those on the cloud. However, these emerging services present a fresh set of challenges to the existing IT and storage systems. In the face of these challenges, MTS needed to digitally transform its infrastructure and business plan.
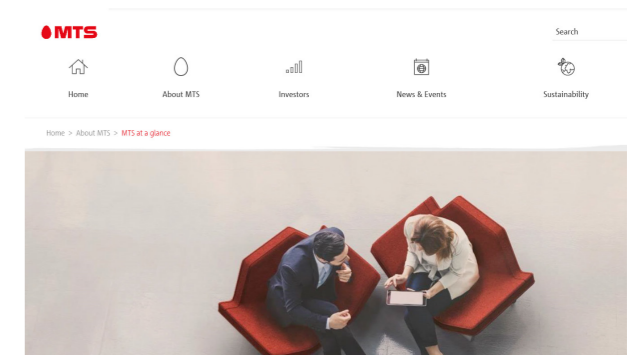
Stable and High Performance for 5x Higher Service Efficiency

As one of the leading carriers in Europe, MTS has preemptively deployed 5G network management systems (NMS) and data storage systems to cope with the flood of data from new services. We found that the existing mission-critical services and cloud systems caused performance bottlenecks and were insufficient to meet the requirements for more high-quality services.

After evaluating products from mainstream storage vendors, MTS turned its attention to Huawei's data acceleration solutions powered by OceanStor Dorado AFAs. The industry-leading AFAs implement end-to-end optimizations in flash-oriented OSs, chips, and SSDs. With five intelligent chips and FlashLink® intelligent algorithm, the AFAs reach up to 20 million IOPS and 0.1 ms ultra-low latency.

In addition to attaining the highest storage performance, MTS also paid significant attention to stability during peak hours. In particular, MTS required at least 98% I/O ratio of the products and less than 4 ms latency. Therefore, the customer will not choose products with a high performance but low stability. Based on the results of existing network tests, the I/O ratio of OceanStor Dorado reached up to 99.91% and less than 4 ms latency in Oracle database scenarios (the most important scenarios), ranking first among all storage products involved in the test. This result surpassed MTS's expectations and met the requirements for its mission-critical services.

< 4 ms latency:
98%

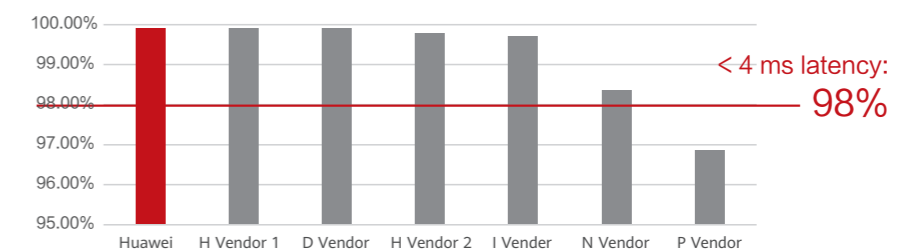|						
100.00%						
99.00%						
98.00%						
97.00%						
96.00%						
95.00%	Huawei	H Vendor 1	D Vendor	H Vendor 2	I Vender	N Vendor	P Vendor

Fig. 1 POC test results of AFAs

Thanks to these successes in the test conditions, MTS chose Huawei OceanStor Dorado on its existing network, hosting its important services, such as billing, CRM, and cloud services. The CTO of MTS, who is responsible for cloud services said, "With the help of Huawei, MTS is going to strengthen its presence in the field of cloud services. The problem with previous storage devices of vendor D to carry VM services was that it took 15 minutes to boot 1000 VMs in batches and led to a significant amount of resource waste. Fortunately, now we are using the next-generation OceanStor Dorado AFAs, our service efficiency has improved fivefold. It only takes three minutes for the boot of 1,000 VMs during peak hours".

Ultimate Reliability for Service Continuity

Service stability and continuity are critical for carriers. For a carrier with tens of millions of subscribers, a one-minute service interruption can lead to a substantial financial loss (up to USD2 million). Therefore, MTS made the reliability of the live network running on Huawei storage devices a priority.

After a series of tests, MTS was pleased to choose the HyperMetro gateway-free active-active solution powered by OceanStor Dorado AFAs to deploy its mission-critical services with 99.9999% availability. In addition, the gateway-free solution reduces the number of fault nodes and simplifies management, providing an industry-leading stable performance with 530,000 IOPS at 1 ms, perfect to deliver the high availability and stable performance needed for MTS.

Batch SSD failure was also a headache for the O&M team. OceanStor Dorado AFAs implement industry-leading RAID-TP technology to tolerate a concurrent three-SSD failure in a RAID group without interrupting services, greatly improving O&M experience.

35% Lower TCO for Soaring Service Requirements

MTS aims to increase its average annual revenue by 30% to 50% in the next three years to improve competitiveness, with an envisaged extension of its cloud services. As part of its aim to become the largest cloud provider in Russia, MTS has acquired several cloud providers, including IT-Grad 1 Cloud LLC. In 2018, MTS's cloud service revenue increased nearly sixfold, however, costs also increased sharply due to rapid data growth. It is estimated that the data volumes will exceed 30 PB in five years, leading to a high storage cost for MTS.

To tackle this problem, MTS wants to reduce the quantity of suppliers and device models to have a streamlined architecture. Then, they want to integrate or replace storage devices on the existing network within three years, which will optimize the use of resources and reduce O&M costs. The final part of the plan will be to replace the traditional centralized storage with AFAs, transforming into an all-flash-powered model.

Huawei OceanStor Dorado AFAs adopt large-capacity SSDs and data reduction technologies with fewer SSD slots and enclosures in the same user capacity. MTS's upgraded IT data center on the existing network provides 30% lower space occupation and 45% less power consumption, reducing cooling costs and maintenance costs by 45% and 70% respectively. It is estimated that the three-year TCO will be reduced by 35%. This decrease of TCO will lead to a decrease of MTS's investment of services.

MTS chose Huawei OceanStor Dorado AFAs to integrate services on the existing network in three phases. The first phase has been completed, helping MTS make significant success in its mission-critical and cloud services. Over the next two phases, MTS will further work with Huawei to realize its short- and long-term goals.

Make money online

Top 5 Banks in Brazil
Accelerate FinTech Development
with Huawei OceanStor AFAs

By Zhu Leshan, Huawei

The top 5 banks in Brazil process nearly 80% of loans and deposits in the country. These banks have increased their investments in the digital transformation of financial services in response to the growing ubiquity of e-finance and smart devices. Their goals are to adapt to rapidly changing emerging services, win the trust and support of more customers, and maximize profitability through IT infrastructure upgrades. These five banks have each chosen Huawei as their partner in digital transformation and are using Huawei OceanStor AFAs to accelerate the evolution of their FinTech.

In 2017, C Bank Decreased Costs and Increased Efficiency with OceanStor AFAs for Market Breakthroughs

C Bank is the second largest public bank in Brazil with financial services for thousands of urban users. Retail services are the most stable growth point for profit of C Bank. They include channel access and operation analysis with an extensive customer base, frequent random concurrent accesses, and demanding requirements for real-time services.

However, the performance bottlenecks of the storage system on its live network restricted random concurrent accesses to retail services. Inadequate supplies of power in Brazil have increased the cost of power over the last two years to an almost 90% increase of BRL 420.28/MWh. As result, C Bank had to reduce the power consumption of data centers as part of its KPIs.

C Bank conducted in-depth comparative analysis on solutions provided by mainstream storage vendors and finally chose Huawei OceanStor AFAs to replace 20 storage devices from Dell EMC on their live network. This replacement provided a more than 5x increase in IOPS processing efficiency, increased the number of concurrent online transactions by 210%, and fully met retail service requirements for high concurrency.

Huawei minimized power consumption with its best-fit OceanStor AFAs storage solution. 2.5-inch SSDs replaced 3.5-inch 15,000 rpm SAS HDDs to reduce the power consumption of a single disk. A single HDD consumes 15 W, but an SSD consumes only 5 W for 66% less power consumption than an HDD. Finally, 10 sets of OceanStor high-end AFAs replaced 20 storage devices from Dell EMC on the live network. This not only reduced the equipment room footprint by 50%, but also reduced overall power consumption by 55%. Estimates predict the OPEX of storage systems will decrease by 33% over the next five years.

20 sets of EMC VMAX high-end or VNX mid-range storage devices

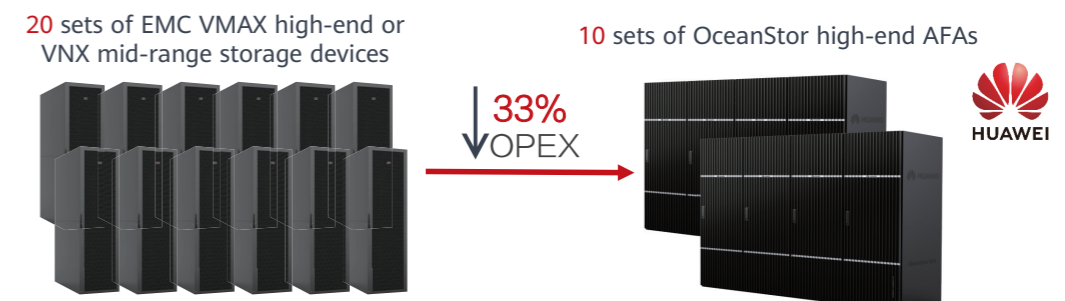**10** sets of OceanStor high-end AFAs

↓**33%** OPEX

HUAWEI

Fig. 1 Replacement of storage devices on the live network of C Bank in Brazil

In 2018, OceanStor AFAs Impressed B Bank with Ultimate Performance

B Bank is one of the largest companies for financial services in Brazil with one out of every three Brazilians as subscribers. After comprehensive technical exchanges with Huawei, B Bank decided to upgrade its outdated IT system to adapt to service needs. Specifically:

1. The SAS grid computing and analysis system supports banking service analysis and investment decision-making. However, decisions were hindered by the slow performance of the system which delayed report generation up to four hours. In the high-speed world of modern finance, the slow response of SAS storage was a huge disadvantage.

2. The IT system of B Bank generates 1.1 million transaction records every month with 85% of the transactions being processed with e-banking. The changes in modes of access demand a faster response speed from storage systems.

In 2018, B Bank decided to collaborate with Huawei to replace 10 PB HPE high-end storage systems on their live network with 12 sets of OceanStor AFAs. These AFAs were distributed to different parts of the network: two sets for the background SAS grid computing and analysis system, three sets for the public service system that processes channel and product services, and seven sets for cloud access on the channel access system.

OceanStor AFAs significantly improved the performance of the IT system. The latency of the SAS grid computing and analysis system access to a storage system dropped from 200 ms to 0.3 ms (a 99% decrease). The storage solution shortened analysis report generation from 4 hours to 15 minutes to remarkably improve service processing efficiency for the SAS grid computing and analysis system.
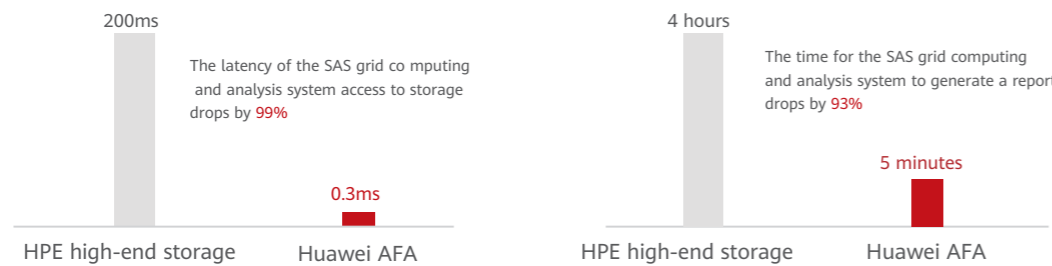

200ms
The latency of the SAS grid co mputing and analysis system access to storage drops by 99%
0.3ms
HPE high-end storage Huawei AFA

4 hours
The time for the SAS grid computing and analysis system to generate a report drops by 93%
5 minutes
HPE high-end storage Huawei AFA

Fig. 2 B Bank accelerated performance of its IT system with Huawei OceanStor AFAs

In 2019, POC Tests of Huawei OceanStor AFAs Galvanized I Bank

I Bank is the largest bank in Latin America. It operates in more than 20 countries in North America, South America, Asia, and Europe, and invests USD 2 billion in technology R&D every year.

In 2014, I Bank built two large-scale data centers. However, their storage devices were unable to cope with the requirements of their continuous service expansion in recent years.

The top priority of IT teams is the stable and reliable running of service systems. The existing storage architecture of I Bank was outdated, and its system availability and DR capabilities were inadequate in meeting the reliability requirements of its service systems. The old devices with expired warranties on the live network and a high component fault rate further complicated the task of the IT team.

Therefore, I Bank paid special attention to the architecture design and reliability of the storage products

that they purchased. In the POC tests conducted for alternative products, the excellent performance of Huawei AFAs surpassed the expectations of the IT team. On an active-active DC (HyperMetro) network, the storage performance was stable, and the loads between the two data centers were fully balanced. On a 3DC network, services seamlessly switched between three data centers for always-on services. RAID-TP technology from Huawei tolerated concurrent failures of three disks without interrupting services. After obtaining the POC test results, the IT team of I Bank was impressed by the powerful disaster recovery capabilities and system availability of the Huawei storage solution.

It was imperative for I Bank to overcome its storage bottlenecks. System efficiency suffered from performance allocation that did not match the service system and a SAS production system that restricted the core service system for personal cards. Despite this, the resources of the storage

---

system running the Splunk application system were not fully utilized.

Huawei proactively assisted in sorting out the problems of the IT service system. Four standard performance models (Small, Medium, Large, and X-Large) and special models were customized for different types of services. Each model corresponded to a storage system for an optimal performance level. For example:

The new performance models fully utilized storage resources and eliminated performance bottlenecks.

Model	Capacity (TB)	IOPS	Bandwidth (GB/s)	Read/Write Ratio	Block Size	Cache Read Hit	Latency (ms)
Small	256 TB	192,000	6 GB/s	70%	32 K	50%	0.8 ms
Medium	512 TB	384,000	12 GB/s	70%	32 K	50%	0.8 ms
Large	1,024 TB	512,000	24 GB/s	70%	32 K	50%	0.8 ms
X-Large	1,536 TB	512,000	24 GB/s	70%	32 K	50%	0.8 ms

Table 1 New performance models

The POC test results revealed the advantages of OceanStor AFAs in QoS customization for different I Bank applications regarding IOPS, bandwidth, and response latency. I Bank fully utilized the performance advantages of OceanStor AFAs in different service scenarios.

Before Huawei, the data centers of I Bank used 154 sets of network storage devices. These devices occupied a lot of cabinet space in the three data centers but provided only 36.9 PB of available capacity. The extremely high power consumption of the devices and cooling systems also created a high TCO.

Only 13 OceanStor AFAs provided 10 PB of effective capacity for I Bank with converged resource pools and excellent data reduction capabilities.

OceanStor AFAs dropped:
- The cabinet space footprint by 60%.
- The power consumption for devices and cooling systems each by 45%.
- The maintenance cost by 65%.
- The three-year TCO by 55%.

Huawei OceanStor AFAs Push FinTech Development for Top 5 Banks in Brazil

In 2017, C Bank decreased costs while increased efficiency with OceanStor AFA for market breakthroughs.

In 2018, OceanStor AFAs impressed B Bank with ultimate performance. Intense collaboration with Huawei built the confidence of B Bank in Huawei technical capabilities and strengths for IT. After deploying OceanStor AFAs on the live network, B Bank greatly accelerated its efficiency in IT system processing by significantly reducing the response time of the SAS grid computing and analysis and channel transaction systems.
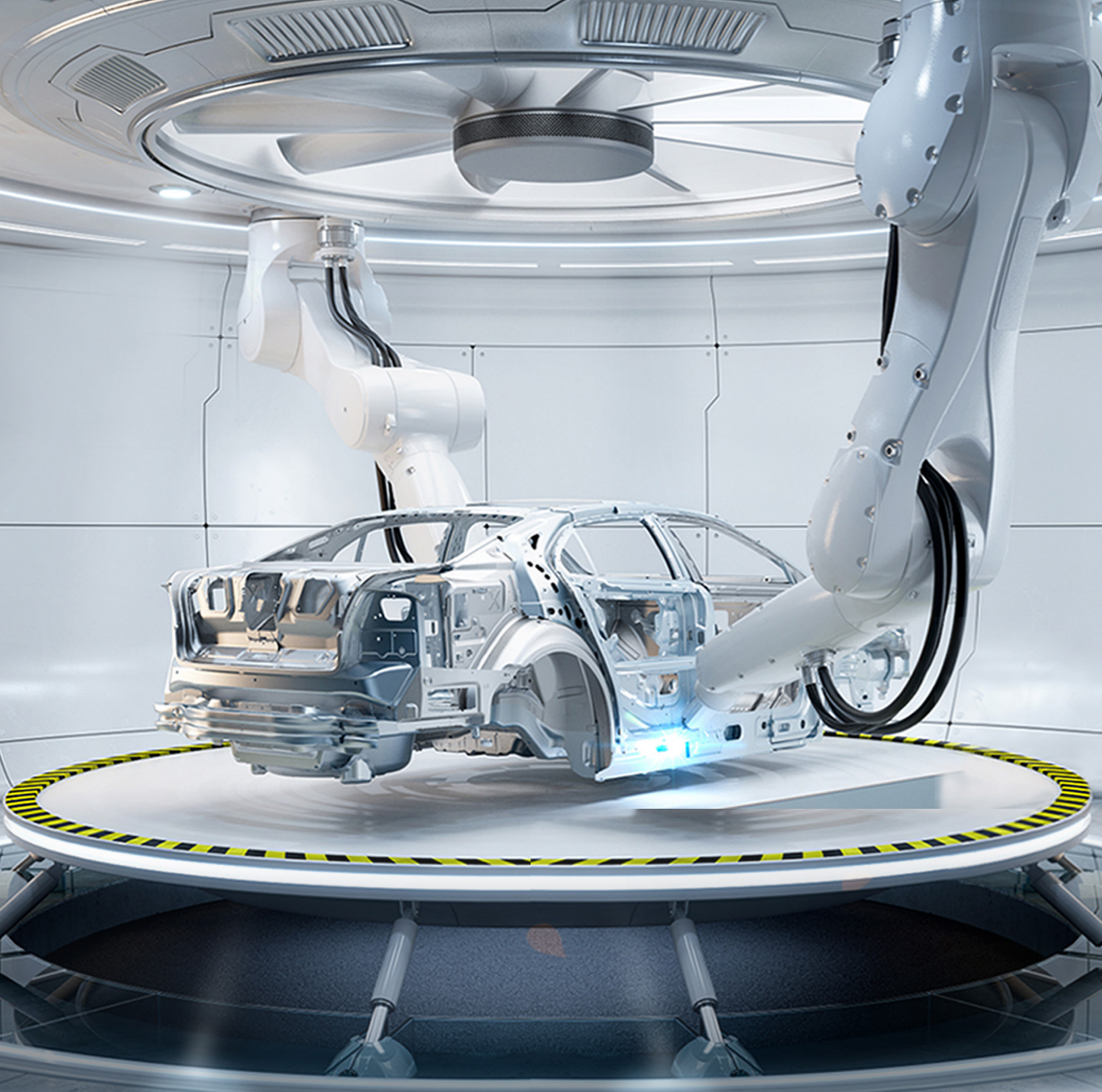
In October 2018, BC Bank purchased two sets of OceanStor AFAs to consolidate its storage systems on the live network, eliminating data silos, fully utilizing storage resources, simplifying O&M, and reducing TCO by 20%.

In 2019, OceanStor AFAs won the I Bank bid through their excellent performance in POC tests. I Bank partnered with Huawei to explore new solutions for digital transformation.

In August 2019, BB Bank signed a storage framework agreement with Huawei and used four sets of OceanStor AFAs to double its effective capacity, save space by 30%, and reduce power costs by 64%.

The top 5 banks in Brazil know that future success lies in the continuous digital transformation of today. They have become leaders in the development of science and technology finance.

Building a fully connected, intelligent world

5 Looking Deeper

Huawei OceanStor AFAs
Full Sail Ahead

Transforming More Data into More Profits

Huawei OceanStor Dorado improves ERP BW system perfomance by **3x**

Huawei OceanStor All-Flash Storage
— **Full Sail Ahead**

By Yuan Shaolong, IT We-Media

It was freezing cold in Beijing in early 2017. Zhang Peng, Chief Architect of Huawei OceanStor Dorado all-flash arrays (AFAs), was waiting to board his overseas flight with mixed feelings. While he was happy and excited that the next-gen OceanStor Dorado development project had just got approved, he was still a little uncertain. Could he and his team help Huawei enter the world of high-end storage?

In July 2019, Huawei launched a brand new generation of OceanStor Dorado AFAs. This high-end product was globally developed by different Huawei teams and features the highest performance of 20 million IOPS and lowest latency of 0.1 ms in the entire industry. Its innovative SmartMatrix architecture redefines stability and reliability for high-end storage. Its built-in AI chip is the next step in the evolution of high-end storage.

Zhang Peng, Chief Architect of Huawei All-Flash Arrays

Looking back over the past two years, Zhang Peng's doubt was gone, "The new-gen OceanStor Dorado is innovative, it's very unique in the industry. I'm very proud."

Ten-Year Quest for Success

"Huawei is not a novice in the all-flash field. It launched the industry's first AFAs in 2009," Zhang Peng said. "Huawei has developed an expertise in many all-flash technologies and is ready to seize any market opportunity."

That opportunity has presented itself. The growing popularity of SSDs is pushing HDDs out of the market. Replacement of traditional HDD-based systems creates a huge market for AFAs, or as Zhang Peng put it, "Speed is the name of the game in the market and technology revolution, and being the first to launch new products is vital to success."

It takes long-term technological accumulation of the team to successfully develop a storage product. "A few years ago, I was appointed to build a research team in a foreign country. I had no clue how to start," Zhang Peng recalled. He delved into the pool of industry talent and called prospective hires one by one on the phone. "Huawei Storage has grown rapidly in recent years with a significant market presence. But step one was not easy, especially recruiting. We eventually developed a great team, and I think the main reason why they joined us is the big stage here at Huawei and the bright future of flash."

For many years, the "Big Three" dominated the high-end storage market. The extremely high technical requirements of high-end storage alone were enough to exclude most challengers. Technological innovation was subsequently slowed to a crawl. But Huawei persevered. "We share the same dream at Huawei, and OceanStor Dorado is the product of the hard work of the world's top storage talents."

Smashing the Limits of Traditional Architecture

Scale-up architecture was standard for high-end storage throughout the traditional HDD era. Exponential data growth and a drive for performance then gave rise to scale-out architecture in recent years. Though it's easy to expand, scale-out architecture lacked the performance and efficiency of the old scale-up architecture. Breaking limits of traditional architecture became the biggest challenge facing the Huawei storage team. The research objective was clear from day one: something brand new combining the advantages of both types of architecture. "Some members of our team used to work for established big names in the industry and were very excited about the challenge of a brand new architecture. Giving them free reign to explore their ideas proved key in redesigning our architecture," said Zhang Peng.

The team encountered another challenge during architecture design. Because the architecture was completely new, many components did not yet exist. "We had extremely high requirements on components, such as interface modules. After many rounds of testing, suppliers gave up because our requirements were just too high." Development slowed to a crawl.

Then someone proposed HiSilicon. After careful deliberation, they decided to give it a shot and tried to achieve architecture breakthrough with it.

Chip verification and testing was both a difficult and urgent task. It was complicated by the use of five chips for controller processing, SSD control, AI, protocol processing, and BMC management. Consequently, communication between the storage team and the chip team was paramount to resolve any adaptation issue. "We had a colleague whose only job for two years was communicating our requirements to the chip team. It sounds simple, but his work was actually integral to our success." Zhang Peng noted.

Their constant efforts paid off in the end. The SmartMatrix architecture of OceanStor Dorado AFAs innovatively separated storage controllers from disk enclosures and connected the front end with the back end.

The separation of controller enclosures from disk enclosures allowed independent upgrade and expansion for either. The SmartMatrix helps OceanStor Dorado support up to 32 controllers in a single system and tolerate the failure of seven out of eight controllers for 99.9999% availability.

Looking back on the sometimes torturous development process, Zhang Peng could not help reflecting, "Good things never come easy." Conforming to their predecessors and peer vendors would have limited their creativity and produced a mediocre storage system. "OceanStor Dorado proved to be worth it. The fast data growth and changing business workloads of today require redefined storage architecture. The innovative OceanStor Dorado provides that much needed innovation as a pioneer in storage technology."

Achieving Ultimate Performance

Stabilizing the storage latency to a low level is a long process that requires repetitive adjustment and optimization. Once steady progress was made in architecture design, the research team started to optimize the software layer. Because the Kunpeng 920 processor had 48 cores, optimization mainly focused on multi-CPU and multi-core utilization. "Our goal was 20 million IOPS. But Rome wasn't built in a day. I/Os were frequently interrupted for varying reasons and we had to analyze each interruption one by one. Then we had to try to plan around them and test again and again," Zhang Peng recollected.

He then added, "Storage system processors mainly process data. Multi-core processors like Kunpeng 920 excel at this because they can dedicate resources to data processing jobs, like deduplication and compression, whereas general-purpose x86 processors are better at compute-intensive applications." Tests showed that OceanStor Dorado surpassed peer products by 50% in performance even with features like snapshot and clone enabled.

Customers sometimes play an unexpectedly important role in product development. "They are our best product managers. The success of OceanStor Dorado would have been impossible without their input," Zhang Peng said. He recalled how a European banking customer gave the team a valuable piece of advice when they showed off their OceanStor Dorado prototype. "We showed him our prototype and were convinced we had thought of everything." The customer was impressed, but still had one suggestion. Although large I/O latency occurred only once per ten thousand, it was a large problem for him that he hoped they could solve.

"This led to some spirited debate among the team." Zhang Peng admitted frankly. Addressing the issue required a system design modification, which would require a huge investment of time and money. Zhang Peng stated the problem succinctly, "The long-tail I/O latency problem adversely affected the original design as tasks needed separation. It involved multiple layers, such as the hardware and software layers." The team decided to follow the Huawei customer-oriented philosophy and resolve the issue despite the increased investment.

Rounds of discussion, modification, and testing finally delivered an OceanStor Dorado with a low latency of just 0.1 ms.

Achieving Long-Term Stability

Another reason why high-end storage is regarded as the crown jewel of storage is its extremely high reliability and stability. This is also why customers in many industries use it for mission-critical workloads.

"Conversations with customers in recent years have communicated an increasing need for storage reliability," Zhang Peng said. He went on to discuss how different industries were worried about multiple points of failure (MPOF) in data centers, "More digital applications means bigger data centers with a lot more equipment. Inevitably, MPOF becomes an issue."

The new-generation OceanStor Dorado has reinforced its reliability with five layers: component, architecture, system, solution, and cloud. These reinforcements set a new benchmark for reliability of high-end storage with 99.9999% service availability.

OceanStor Dorado AFAs use wear-leveling technology to balance service loads among all SSDs and prevent collective failure of multiple disks. The innovative SmartMatrix architecture adopts a fully-interconnected front-end and back-end design and an intelligent multi-protocol interface chip. A fully symmetric active-active controller design allows hosts to access LUNs from any controller. In the event that a controller is faulty, services can be switched over to functioning controllers within one second. OceanStor Dorado AFAs tolerate the failure of seven out of eight controllers. In addition, an all-in-one backup and disaster recovery design provides active-active storage capabilities without requiring gateways. This reduces the number of possible faulty nodes and simplifies system deployment. OceanStor Dorado systems are also able to work seamlessly with the public cloud for data backup and service recovery on the cloud within minutes.

"With zero service downtime from hardware or software errors, the new OceanStor Dorado truly has always-on services," Zhang Peng said proudly.

Making Storage Smart

Despite appearances, the high-end storage market is full of opportunities. Artificial intelligence (AI) is reinventing high-end storage and making intelligence the wave of the future.

"Reinvigorating high-end storage requires something strong, something innovative," said Zhang Peng. "The fully interconnected architecture of OceanStor Dorado is ideal for AI technologies, such as machine learning. Being able to learn and analyze the global I/O patterns and supporting intelligent storage operations will definitely improve system performance and efficiency." The new OceanStor Dorado is equipped with many technologies for advanced intelligence. A multi-protocol interface chip efficiently parses protocols.

An AI chip proactively analyzes and learns the I/O patterns of many application models, improving read cache hit by 50%. Smart disk enclosures take over data reconstruction from array controllers to speed up the process to 15 minutes/TB.

Additionally, the new AFAs support full-lifecycle intelligent O&M with their AI chips and algorithms to provide resource planning, service provisioning, system tuning, risk prediction, and fault location. They support 60-day prior performance and capacity prediction, 14-day prior disk failure prediction, and can provide immediate solutions for 93% of identified problems.

Conclusion

Zhang Peng was honest about the intense pressure the team worked under for the last two years. They worked overtime almost every day to tackle all the technical challenges of the new hardware, software, and OS designs. "I'm glad that our hard work has finally paid off. The new OceanStor Dorado is a huge success."

OceanStor Dorado has been a resounding success in the all-flash market that was fueled by the dedication and hard work of the Huawei storage team. They have not only made Huawei the leading vendor in the Chinese market, but have also made it the fastest growing vendor world-wide.

A brand new logo was released on the same day as the new AFAs: a swordfish leaping out of the sea. The swordfish, one of the fastest fishes in the ocean, symbolizes the speed of the AFAs and their ability to help customers race forward into the ever-growing ocean of data.

Huawei OceanStor All-Flash Storage Issue 2
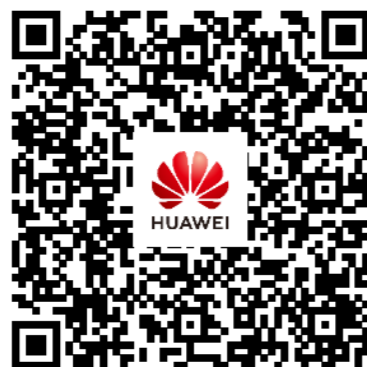
Huawei OceanStor Storage Product Portfolio

OceanStor V5
(Extreme Convergence)

OceanStor Dorado
(Superior Performance)

OceanStor 18500/18800 V5

OceanStor Dorado 18000 V6

OceanStor 5300/5500/5600/5800 V5

OceanStor Dorado 8000 V6

OceanStor 6800 V5

OceanStor Dorado 3000/5000/6000 V6

Ever Fast, Ever Solid, AI-Powered