IDC

Moving Towards an All-Flash Data Center Era to Accelerate Digital Transformation

Sponsored by: HUAWEI

Table of Contents

Chapter I Data Is the Driving Force of Enterprise Development as the Digital Age Continues to Evolve 04

1.1 Global Digital Transformation: Move from Experimentation to Multiplied Innovation 04

1.2 Enterprise Digital Transformation Advances Steadily 06

1.3 Build Data-Driven Future Enterprises 08

Chapter 2 Demands Are the Catalyst for IT Infrastructure Development and Innovation 09

2.1 Data-Driven IT Infrastructure Upgrade 09

2.2 Five Challenges of IT Infrastructure Upgrade 11

Chapter 3 All-Flash Data Centers Help Enterprises Build a Green and Reliable Innovation Platform 15

3.1 All-Flash Data Centers Upgrade and Revamp IT Infrastructure 15

3.2 All-Flash Data Center Trends and Advantages 17

Chapter 4 Best Practices for Deploying All-Flash Data Centers 25

4.1 All-Flash Market Overview 25

4.2 Benefits of All-Flash Data Centers in Industries 26

Chapter 5 IDC Recommendations 40

IDC
Opinions

- According to the IDC's forecast, 65% of global GDP will be digitalized by 2022 as the digital economy grows; the digital economy will account for over 70% of China's GDP by 2025, boosted by the deployment of new infrastructure. IDC believes that digital transformation has entered the 2.0 era. The "future enterprises" are those organizations that firmly follow the digital path, where digital operations achieve scale, and the innovation speed is a magnitude greater than that of traditional business. How well organizations unlock their data value will determine how much they can digitally transform.

- As enterprises progress in digital transformation, data generation, collection, transmission, processing, storage, and management have undergone huge changes. These changes are fueled by things such as explosive data growth, diverse data types, higher data transmission and processing requirements, and the increasing complexity of data management. To address these challenges, enterprises need a modern, future-facing digital infrastructure. IDC believes that digital infrastructure is crucial to the future success of the digital economy, and is the foundation of enterprise business in the Digital Transformation (DX) 2.0 era. Such infrastructure is also intelligent, autonomous, and oriented towards multi-cloud architecture

- IDC's survey shows that more than 90% of enterprises plan to upgrade their digital IT infrastructure. In this process, the top 5 needs and challenges they face include reliability, cost reduction and efficiency improvement, data innovation, real-time agility, and intelligent management. Driven by these needs and challenges, the all-flash data center has become one of the most important development directions for future data centers. IDC believes that an all-flash data center refers to the data center that uses SSDs for at least 90% of its storage capacity (covering external storage systems and built-in storage of servers), while delivering high density and reliability, low latency, and energy efficiency. The all-flash data center supports the

development of businesses and applications, and helps enterprises maximize the value of data innovation.

- The major development trends of future all-flash data centers cover 5 points. First, all-flash storage is no longer restricted to certain applications and workloads. Now, it can be used to meet the demanding workloads of critical and value-added services. Second, all-flash data centers facilitate the green, energy-efficient and sustainable development of enterprises. They help reduce the footprint and number of data centers an enterprise needs, enabling them to better control energy expenses. Third, memory-driven infrastructure will promote the development of all-flash data centers and organizations moving to all-IP networks will have access to faster transmissions. Fourth, memory-driven infrastructure is going mainstream. With enough flexibility, it provides optimal latency for workloads and drives the development of all-flash data centers. Fifth, the realization of intelligent operations and maintenance in all-flash data centers will facilitate better management of cross-domain data over the full lifecycle.

- The construction of the all-flash data center can provide strong support for the development of enterprises, and satisfy their data-related needs. According to the IDC's global forecast, all-flash arrays (AFA) will grow at a compound annual growth rate (CAGR) of 7.5% from 2020 to 2024, and will account for more than half of the external storage market sales by 2024. In China, the market share of AFA has already reached 18.9% in 2020, with a growth rate of 24.0%. The construction of the all-flash data center will not only help enterprises upgrade storage media, but also achieve an all-round improvement in IT infrastructure architecture, data, network, and computing power. In China, the all-flash data center has already been applied in finance, government, healthcare, telecommunications, and manufacturing. In these industries, it helps meet the need for reliable, agile, stable, efficient, and intelligent infrastructure to support business innovation.

Chapter I

Data Is the Driving Force of Enterprise Development as the Digital Age Continues to Evolve

1.1 Global Digital Transformation: Move from Experimentation to Multiplied Innovation

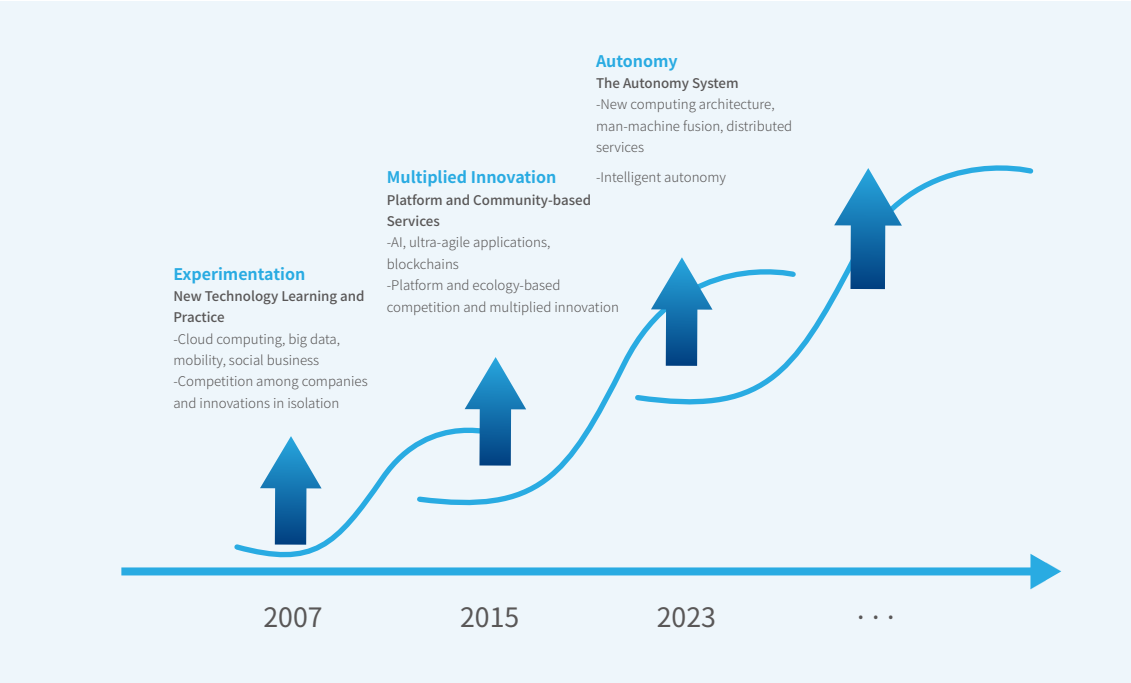
Scientific and technological development shapes our society and changes the ways of spurring economic growth. The emergence and progress of 3rd Platform technologies and services has inspired a range of innovations that generate new economic patterns, create new business forms, and bring new commercial value. The 3rd Platform technologies typically include cloud computing, big data and analytics (BDA), mobility, artificial intelligence (AI) technologies, social business, next-generation security, the Internet of Things (IoT), and other next-generation technologies such as blockchain and quantum computing. The digital economy is booming and becoming a new engine of economic growth. Based on IDC's prediction, as the global economy remains firmly on course for its digital destiny, 65% of the global GDP set to be digitalized by 2022; by 2025, the digital economy will account for over 70% of China's GDP, as a result of the new infrastructure policy.

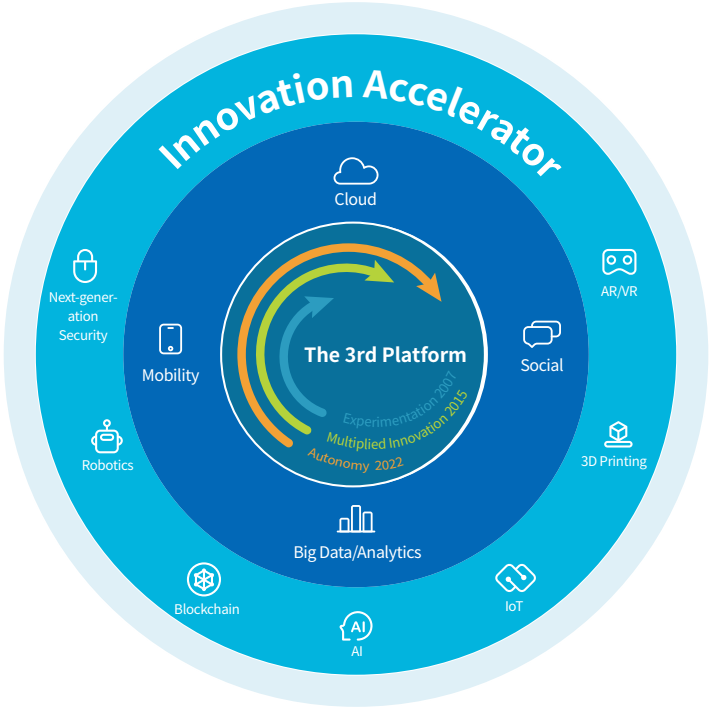
IDC believes that the digital economy has the following features:

· Digital knowledge and information as key production factors;
· Modern information network as a key carrier;
· Efficient utilization of ICT technologies as a key driving force behind the efficiency improvement and economic structural optimization.

According to IDC, as the 3rd platform technologies continuously deepen their support for business growth and innovation, the global digital transformation has leaped from the "experimentation" to the "multiplied innovation" phase. The new phase is based on the in-depth integration of various emerging technologies, and the five general trends will enable enterprises to achieve "multiplied innovation". This includes multiplied insights based on AI and big data, multiplied contact based on IoT and edge computing, multiplied development based on cloud computing and developers, multiplied expectations based on mobility and social business, and multiplied trust based on security and blockchains. In the future, as the digital economy continues to evolve, enterprises will gradually move towards the new phase of "autonomy", and more new business forms and patterns will emerge. This will bring more possibilities for product and service innovation and for the enterprise's intelligent autonomy.

Fig. 1 Three Chapters of the Application of the 3rd Platform Technologies





Source: IDC, 2021

1.2 Enterprise Digital Transformation Advances Steadily

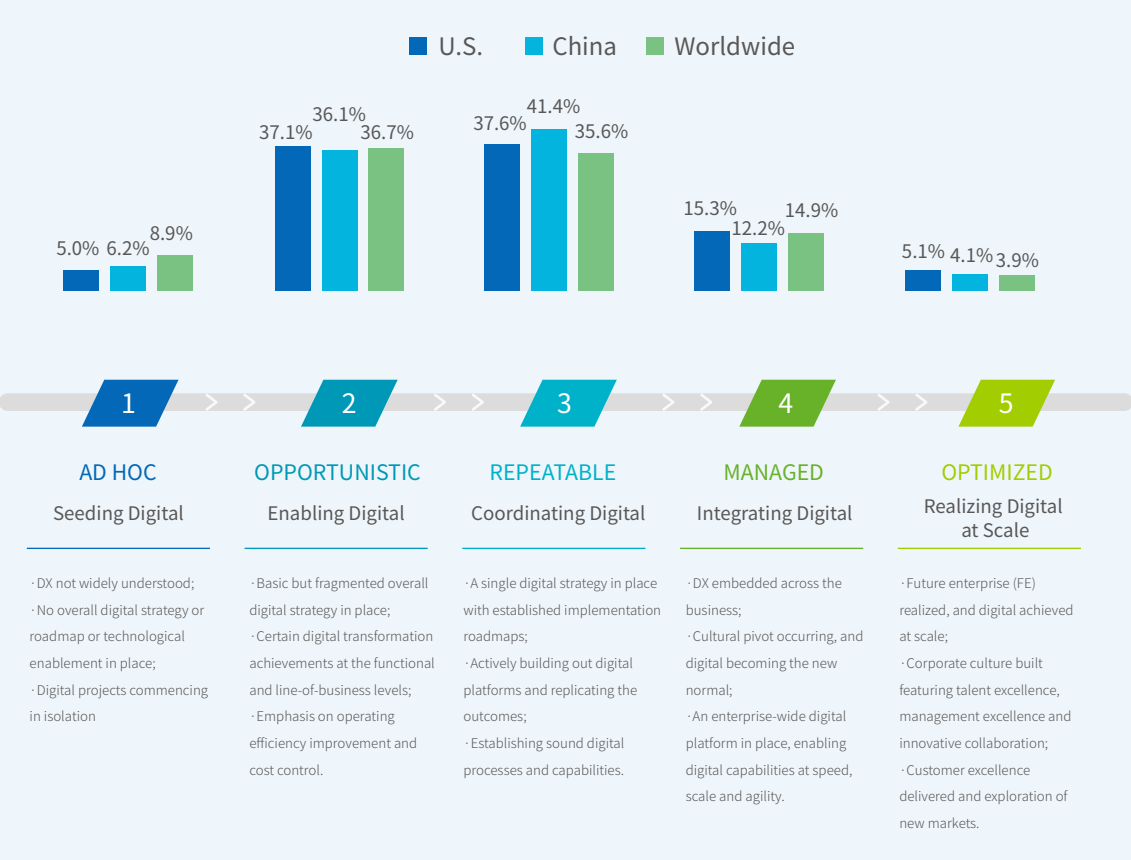
As the digital economy becomes increasingly important, enterprises must accelerate their digital transformation. According to IDC, digital transformation refers to the approach to using digital technologies to drive the innovation of an organization's business model and the reconstruction of its ecosystem. Technologies function as the driving force, and business growth and innovations are the core. The goal is to take advantage of digital technologies to reconstruct products, services, operating models and business models, improve user experience, optimize organization management, build the digital ecosystem, and promote sustainable innovations in the industry. IDC's data shows that in 2020, global spending related to enterprise digital transformation reached 1.3 trillion USD. That number is expected to grow to 2.4 trillion USD by 2024, while non-digital transformation related ICT investment will shrink at a CAGR of -1.6%.

Fig. 2 IDC Worldwide Enterprise Digital Transformation Spending Forecast
(Unit: Million USD)

Source: IDC, 2021

Enterprises are more conscious and capable of digital transformation than ever before. IDC divides the development of enterprise digital transformation into 5 stages based on maturity: ad hoc, opportunistic, repeatable, managed, and optimized. According to IDC's survey on the digital transformation of worldwide enterprises, nearly half of the world's enterprises moved to the 3rd or 4th stage in 2020. The overall trend in China is line with that of the world: 41% of enterprises have entered the 3rd stage, and digital transformation is still accelerating. According to IDC's forecast, 75% of enterprises will have an overall roadmap for the implementation of digital transformation by 2023, and by 2025, 75% business leaders will utilize digital platforms and ecosystems to adjust their value chains, so as to adapt to new markets, industries and ecosystems. More and more enterprises recognize the value of digital transformation and hope to make concrete actions to enhance the value of their data. IDC believes that digital transformation has entered the 2.0 era. The "future enterprises" are those organizations that firmly follow the digital path, where digital operations achieve scale, and the innovation speed is a magnitude greater than that of traditional business.

Fig. 3 Enterprise Digital Transformation Maturity: Worldwide, U.S., and China



Source: IDC, 2021

1.3 Build Data-Driven Future Enterprises

Data is the core resource in the digital economy era and a key driving force behind sustainable digital development. With the further development of the Internet, data becomes the carrier of the virtual world and highlights the real world. From a global perspective, all countries and regions have recognized that data is more than just a product of the digital era — it is the fuel that facilitates the development of the digital economy, and a fundamental strategic resource for the development of a country and its enterprises. For a long time, the U.S. has been formulating policies on data infrastructure construction, data management and transparency, data sharing and security, and e-government, continuously mining the value of data. The EU is also actively exploring better ways to use and share the value of data. In 2020, the European Commission released "A Whitepaper on Artificial Intelligence" and introduced the "European Data Strategy", aiming for Europe to adapt to the development trends of the digital era. The UK has released strategies such as "Digital UK", "Digital Charter", and "UK Digital Strategy", striving to regularize the Internet, create conditions for the development of the digital economy, achieve a leap from "digital" to "intelligent", and ensure the steady progress of the country. Lastly, China is committed to promoting the development of the digital economy. China considers data as a key production factor, emphasizes the importance of data on the national level, and puts forward three main points on accelerating the development of data: promoting the open exchange and sharing of government data, enhancing the value of social data resources, and strengthening the integration and protection of data resources.

For an enterprise, the value of data mainly lies in supporting decision-making and business goal setting, business development, product and service innovation, the all-round optimization of customer experience, and innovation at scale.

Support Corporate Decision-Making and Business Goal Setting	Support Business Developmen	Support Product and Service Innovation	Support All-Round Optimization of Customer Experience	Support Innovation at Scale
Data helps enterprises gain insights into social development rules and market trends to better make strategic deployment in key areas.	Data mining and analytics help to provide resources for the development and release of the next-generation business, accelerating the internal operation efficiency and the business implementation and iteration, and optimizing the capability to process business flows.	Data promotes product and service innovation and optimization while achieving fast product and service iteration. It also facilitates accurate delivery of customized products and services, and improves the chances of success for new products.	Data helps enterprises gain insights into customers to better provide the customers with customized services and achieve customer excellence.	Enterprises can form a data-centric insight mechanism, optimize the management process, integrate relevant resources and empower business innovations, so as to create differentiated values for customers and the market, and build the core competitiveness of enterprises.

As data plays an increasingly important role in future enterprise development, how well it utilizes the value of data will determine how far it can go in digital transformation. Enterprises need to build a data-driven culture centering on data management, Data as a Service (DaaS), and data analytics to spot and get data, and fully explore and achieve the value of data. Only in this way can they guarantee future competitiveness.

Chapter 2

Demands Are the Catalyst for IT Infrastructure Development and Innovation

2.1 Data-Driven IT Infrastructure Upgrade

The development of enterprise digital transformation has sparked many changes that affect data generation and collection, transmission and processing, storage and management, and data utilization. These changes include:

Massive data growth: The development of 3rd platform technologies, growth of products and services, innovation of business models, and acceleration of data production and process are driving rapid growth of data. According to IDC's Global DataSphere — a measure of how much new data is created, captured, or replicated in any given year — the total global data will grow from 64 zettabytes (ZB) to nearly 180 ZB by 2025, with a compound annual growth rate (CAGR) of 23% over the 2020-2025 forecast period. By 2025, China's data volume is expected to reach nearly 41 ZB, accounting for about one quarter of global data.

Diverse data types: As the data volume soars, data sources and structures become more complex. The IDC classifies data across over 70 categories of devices into 6 major types: entertainment data, IoT data, non-entertainment image/video, productivity/embedded data, social media data, as well as voice data. By 2025, the amount of data created by IoT devices will reach 26.6 ZB, accounting for 15% of the global data and making IoT the fastest growing data field, with a CAGR of 47% over the 2020-2025 period. Social media data follows closely behind with a projected CAGR of 35.4%. With the rise of machine vision technologies and AI video image analytics applications, intelligent camera systems will be adopted in more scenarios. The IDC predicts that the worldwide video surveillance camera market will grow to 44 billion USD by 2025, with a 5-year CAGR of nearly 13%. Intelligent video image systems will generate a large amount of data, a significant proportion of which will be unstructured. The IDC believes that unstructured data will grow quickly over the next 5 years, continuing to increase its proportion in the overall datasphere.

Increasing requirements on data transmission and processing: With the widespread implementation of technologies such as the IoT and AI in various industries, organizations have a growing need for real-time data. To ease the pressure on data transmission and processing, edge computing is emerging to reduce the latency of data perception and the response time. The IDC predicts that by 2023, 70% of enterprises will process data at the IoT edge, and the number of applications deployed at the edge will increase by 8 times. Furthermore, 50% of new infrastructure will be deployed at the edge. By 2025, the amount of data generated at the edge will reach 40 ZB, with a five-year CAGR of 33.4%.

Increasing amount of data to store and manage: As the data volume increases, storage capacity grows year by year. The IDC forecasts that the global installed storage capacity will grow at a CAGR of 19.2% over the 2020-2025 period, climbing from 6.7 ZB in 2020 to 16.1 ZB in 2025. In terms of global storage device types, enterprise-class storage became the storage device with the largest capacity in 2017, surpassing personal computer (PC) and tablets. It is estimated that the installed capacity of enterprise-class storage systems will account for 80% of the global installed capacity in 2025. In addition, the IDC forecasts that newly built storage capacity at the core will maintain a 5-year CAGR of 26%, while storage capacity at the edge will grow at a CAGR of 30.3%. In 2025, the newly built storage capacity at the core and the edge will account for 70.5% of all newly built storage capacity.

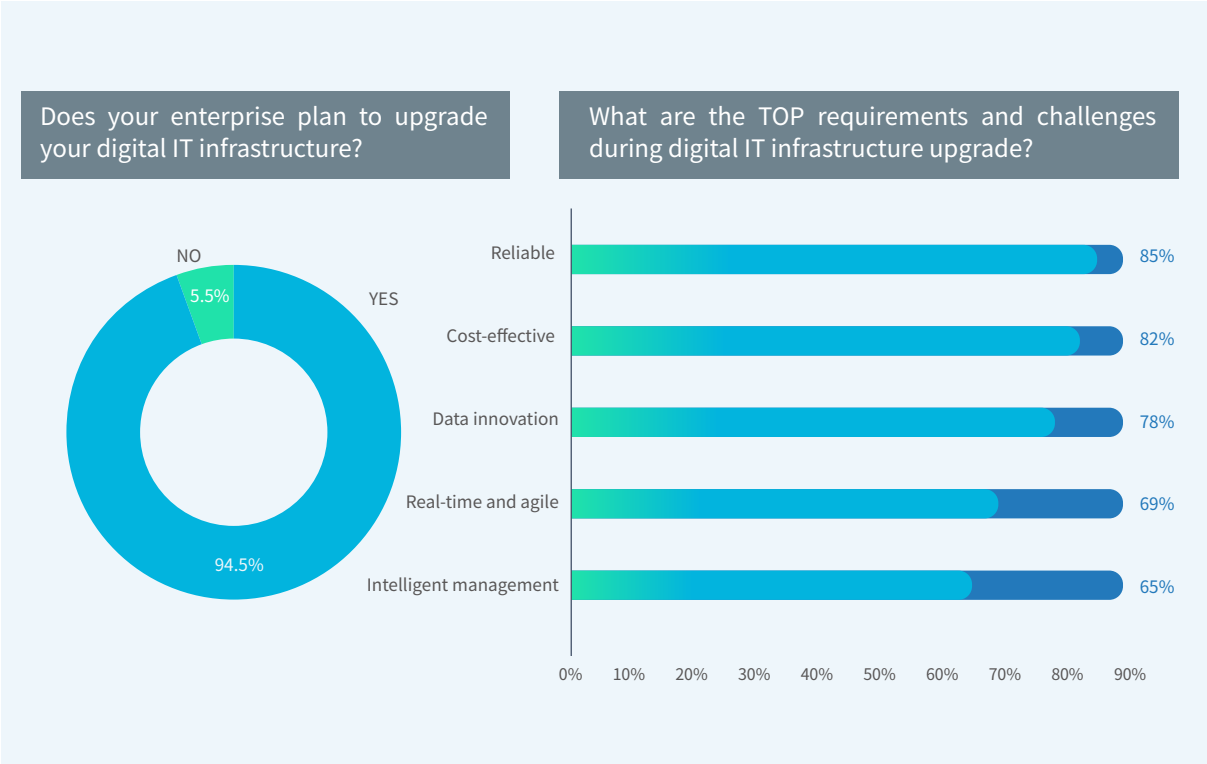
IDC believes that to successfully carry out digital transformation, companies need to build customer-centric, and data-driven digital services. Specifically, through the integration of digital and physical channels, data can be effectively collected, processed, and stored, and the value of data can be extracted to create a better experience for customers. IT infrastructure provides data-driven support for organizations and is closely related to business innovation and development. It is the key to improving production and management efficiency and optimizing user experience of products and services.

Going forward, organization will be built on a digital infrastructure platform centered on data. IDC believes that such future digital infrastructure is an inevitable product of the Digital Transformation 2.0 era, laying the foundation for future enterprise operations, with identified critical features of cloud-centric, autonomous operations and ubiquitous deployment.

2.2 Five Challenges of IT Infrastructure Upgrade

According to IDC research, organizations face the following five major challenges when upgrading their digital IT infrastructure and overhauling their data centers.

Fig. 4 Major Requirements and Challenges for Digital IT Infrastructure Upgrade



Source: IDC, 2021

Data Innovation

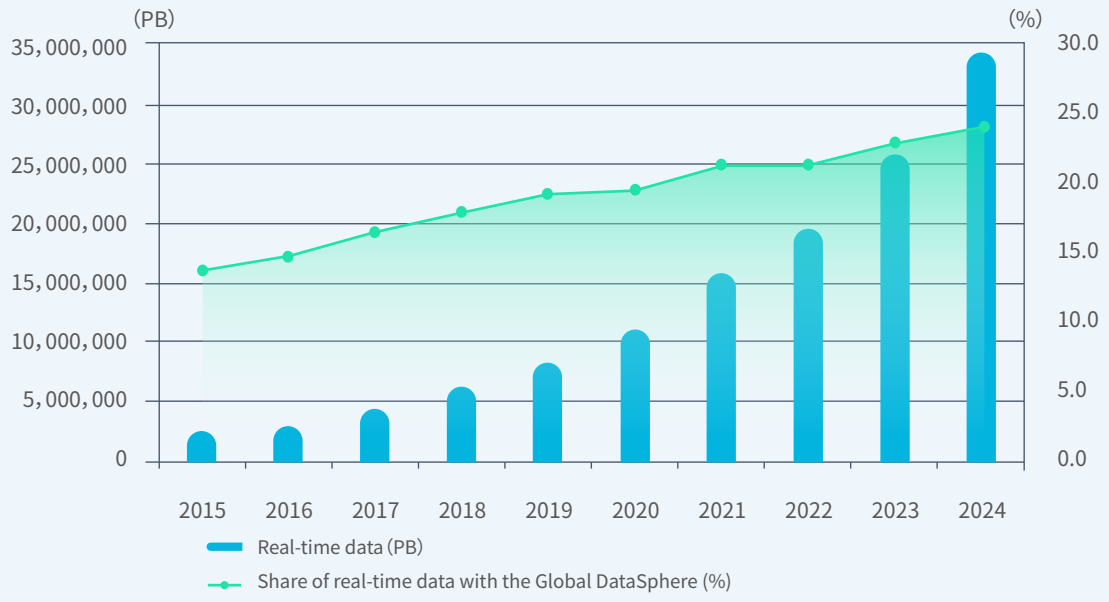
In the digital economy era, data is the new means of production and is regarded as an important asset by enterprises. Data innovation is the key for enterprises to improve their future competitiveness. More and more organizations are beginning to set key performance indicators (KPIs) around informatization construction and data innovation. For example, ensuring that at least 40% of the company's IT budget is allocated to digital transformation (DX) initiatives, that at least 40% of the DX initiatives are approved for implementation, and that the DX initiatives influence over 40% of enterprise revenue and/or profit growth. Organizations need to cope with the shift from cold to hot data and the challenges of data interconnection and interworking. They must also use intelligent platforms to enable data enablement to maximize data value and continuously innovate services and models. The rapid development of enterprise digital transformation has caused data to grow explosively, and data sources and structures have become more and more complex. The front-end data collection, storage, management, and application, as well as the entire process of data transfer inside and outside an enterprise require an IT infrastructure with all-round support, placing higher requirements on the existing IT infrastructure.

Real-Time and Agile

Enterprise services increasingly require a real-time data response. For example, high-frequency trading and real-time risk control systems in finance, billing systems in telecom, road monitoring, and autonomous driving. According to IDC forecasts, the share of real-time data will increase from 19.3% in 2019 to 24.4% in 2024, and this trend is getting obvious in new areas like big data analytics. In addition to rigid requirements for real-time data, future organizations will have increasingly complex business and application requirements, and the performance requirements for other production data will also rise. In the digital era, real-time response is a key factor for organizations to improve their core competitiveness. With the integration of emerging technologies, such as AI and IoT, more and more service scenarios require extremely low latency. Meeting this demand at the server, storage, and networking levels is one of the challenges organizations face when building the future digital infrastructures.

Real-time data: In today's interconnected world, data is everywhere and often customized, which requires it to be transmitted and processed in real time. According to IDC Global DataSphere, real-time data is defined as data that needs to be acted upon/reacted/recoded within 2 seconds, including data from systems such as on-site sensors, financial transactions, drones, and on-site controls.

Fig. 5 Worldwide Real-Time Data and Share, 2015-2024



Source: IDC, 2021

Green and Energy-Efficient

The COVID-19 pandemic has greatly impacted the global economy and production methods. More of the services we use for work, entertainment, and everyday life are moving online. With this, more and more organizations are producing better digital products and services for digital collaboration, aiming to improve their competitiveness. According to an IDC survey, among the benefits brought by digitalization, cost reduction and efficiency improvement have gradually become one of the most important priorities for organizations. For example, in the financial industry, banks are facing the challenge of increasing bad debt rates due to the pandemic. In the healthcare industry, cost reduction and efficiency improvement are two key benefits that will help them cope with the pressure brought by the pandemic. Building a green data center is a top priority for IT departments in many organizations.

Since the release of the Paris Agreement, countries around the world have taken proactive measures to reduce greenhouse gas emissions. For example, China has proposed plans to hit peak CO2 emissions by 2030 and become carbon neutral by 2060. As data volumes and applications surge, data centers support the development of the digital economy, but they also consume vast amounts of energy. The energy consumption of data centers has become an obstacle to sustainable digital development,

increasing the maintenance cost of data centers, and countering efforts of green development worldwide. Reducing data center energy consumption has become one of the biggest challenges facing most enterprises. Countries and regions have adopted low-carbon data center policies. The U.S., for example, has proposed the Data Center Optimization Initiative (DCOI) policy. European Commission has also taken measures to encourage the construction of climate-friendly, energy-efficient, and sustainable data centers.

Reliable

As enterprise digitalization continues, issues surrounding data security will become increasingly challenging. The IDC believes that future enterprises will build on digital infrastructure, which will manage the entire lifecycle of data in the cloud, edge and devices. As such, ensuring the data reliability is the key to building the digital infrastructure of the future. Digitalization makes data reliability risks more diverse. Data growth and architecture complexity pose new challenges to data reliability. Future services and applications will increasingly rely on sustainable services provided by infrastructure to support enterprise data operation services and provide operational support, information resource services, core systems, data storage, and backup, ensuring fast service response and continuity. Enterprises should start with infrastructure reliability, adopt highly reliable IT systems, and establish disaster recovery and backup systems to ensure maximum data reliability.

Intelligent Management

Looking forward, digital services will find their way into traditional data centers, clouds, edges, and devices where data is generated. Therefore, cross-domain management will become more critical, making management teams more flexible, and ridding of deployment location constraints. In the big data age, infrastructure management will depend on AI and machine learning technologies. Organizations need to make full use of big data analytics and multi-layer AI architecture in the cloud to build extensive system fault models via cloud training, in order to better predict faults, system performance, and capacity. Intelligent management can support the portability of higher-level workloads and highly dynamic applications. The digital infrastructure integrates with a variety of workflow automation tools by providing open application program interfaces (APIs), to enhance the tool ecosystem and achieve automatic O&M.

Chapter 3

All-Flash Data Centers Help Enterprises Build a Green and Reliable Innovation Platform

3.1 All-Flash Data Centers Upgrade and Revamp IT Infrastructure

The move to an all-flash data center will shape the future of infrastructure. IDC believes that at least 90% of storage capacity should be provided by high-density, high-reliability, low-latency, and energy-efficient SSDs in an all-flash data center. This includes external storage systems and built-in storage of servers. The all-flash data center helps enterprises harness the power of emerging technologies and applications and maximize the value of data innovation.

Enterprises not only care about device-level reliability, but also about business continuity, data recoverability, long-term compliance retention of records, and the agile management of the entire data infrastructure. Therefore, the construction of an all-flash data center involves more than just upgrading the media; it involves integrating data, networking, and compute resources to innovate the IT infrastructure architecture.

- **Storage:** Improving hardware capabilities has impacted storage. For example, magnetic media has given way to flash memory as storage media, and hard disk drives (HDDs) lost their place to solid state disks (SSDs). SSDs offer a data access capability of about 100 times faster, throughput 100 times higher, and single-disk IOPS more than 1000 times higher than traditional disks. The development of storage communication protocols takes the benefits of higher efficiency due to storage media upgrades and larger network bandwidth to the next level, which greatly enhances the processing power of computers.
- **Network:** Storage IP networks have become more vital as service volumes increase. To maximize the performance potential of SSDs, future storage networks should gradually adopt the NVMe protocol instead of the conventional SCSI protocol. Faster media and protocols require faster networks. The NVMe over Fabrics (NoF) storage network provides higher throughput and lower latency. It is a preferred solution for end-to-end NVMe storage networks.
- **Computing:** This is key to ensuring the stability and control of the entire lifecycle of data. The high-end all-flash solution offers full-lifecycle automated, intelligent management for data centers in the planning, provisioning, O&M, and optimization phases, reducing O&M challenges caused by data pooling, and implementing agile optimization.

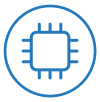
IDC believes that the future all-flash data center is developing towards the following trends:

- All-flash storage is not only applicable to specific applications and workloads. It is widely used to carry multiple service scenarios, such as critical and value-added services.
- The all-flash data center helps enterprises achieve energy efficiency and sustainable development by reducing footprint and energy consumption.
- All-flash technologies are being continuously optimized, promoting IP-based data center networks.
- Memory-driven infrastructure is flexible enough to ensure the lowest latency for workloads, driving all-flash data centers forward.
- The all-flash data center implements intelligent O&M, facilitating cross-domain deployment throughout the data lifecycle.

3.2 All-Flash Data Center Trends and Advantages

Opinion 1: All-Flash Storage Is Not Restricted to Certain Applications and Workloads

Future applications demand ever-increasing access speeds, meaning organizations must provide higher system performance to meet this growth. The all-flash system reduces the total cost of ownership (TCO) and expenditure per I/O, and simplifies management. It can be extensively adopted to fit a range of environments, and is not limited to specific applications and workloads. It perfectly meets the demanding workloads of critical and value-added services.

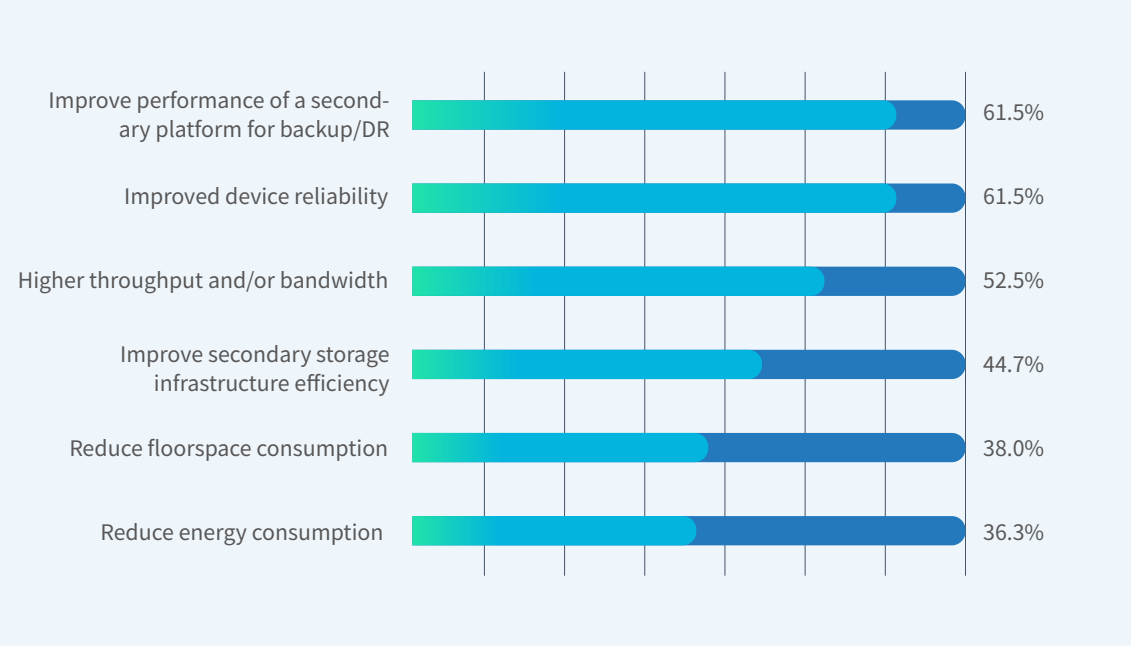


All-flash storage for critical services: Core enterprise services usually have high system performance, reliability, and stability requirements. The IT system needs to be very responsive to ensure the stability of business applications; even a little latency could result in the loss of a large number of users. All-flash storage could provide million-level IOPS and millisecond-level latency, and it is definitely more reliable and stable. In the past, enterprises have often been concerned about the service life of SSDs—after all, data security and reliability are what enterprises care about the most. Different from HDDs, an SSD's service life is determined by the number of writing and erasing of the NAND flash storage. Currently, TLC NAND has become the mainstream of enterprise-level SSDs, and QLC NAND will be extensively adopted in cloud, CDN, and tiered storage solutions. Thanks to the continuous development of writing algorithms, the service life of an SSD is greatly extended at the

software level, so it won't become a bottleneck in terms of the SSD's actual use and replacement cycles. Besides, since there are no mechanical moving parts in an SSD, mechanical failures can be avoided. Compared with HDDs, enterprise-level SSDs are more resistant to bumps and collisions, therefore, all-flash data centers have higher-level system stability and are less likely to encounter accidental data loss. All-flash storage systems are an ideal choice for critical and high-concurrency application scenarios. It is increasingly popular for core enterprise business. SSDs equipped with new ports, storage-class memory (SCM), and the end-to-end NVMe network architecture supercharge system performance. More importantly, all-flash isn't limited to the primary storage; it can also be used in DR solutions to accelerate backup speeds and shorten recovery time, better meeting the ever-increasing needs of applications in the digital age.

All-flash storage for value-added services: Organizations can expect digital infrastructure to be the catalyst for more value-added services. For example, real-time translation in e-commerce promotes sales and improves customer stickiness, whereas those working in social network can quickly analyze user behaviors for precision marketing. Innovative data applications are a key driving force for enterprises' digital transformation. Distributed storage is an ideal solution for storing massive amounts of unstructured data and achieving fast data response. It is applicable to a variety of scenarios such as databases and virtualization platforms that feature high concurrency. Additionally, a distributed architecture that decouples storage from computing improves resource utilization and adopts all-flash storage for higher reliability. The price erosion of SSD and improvements in storage capacity, density, and performance have upended industries across all fields. Distributed storage is starting to adopt more SSDs to reduce performance and bandwidth costs per GB. Such deployment also uses intelligent tiering to place data at the memory and SSDs according to access frequency, greatly increasing data read speeds and maximizing hardware performance. According to IDC's research, a large number of organizations prefer to use all-flash systems as secondary storage thanks to the enhanced backup and data recovery performance, and high storage system stability.

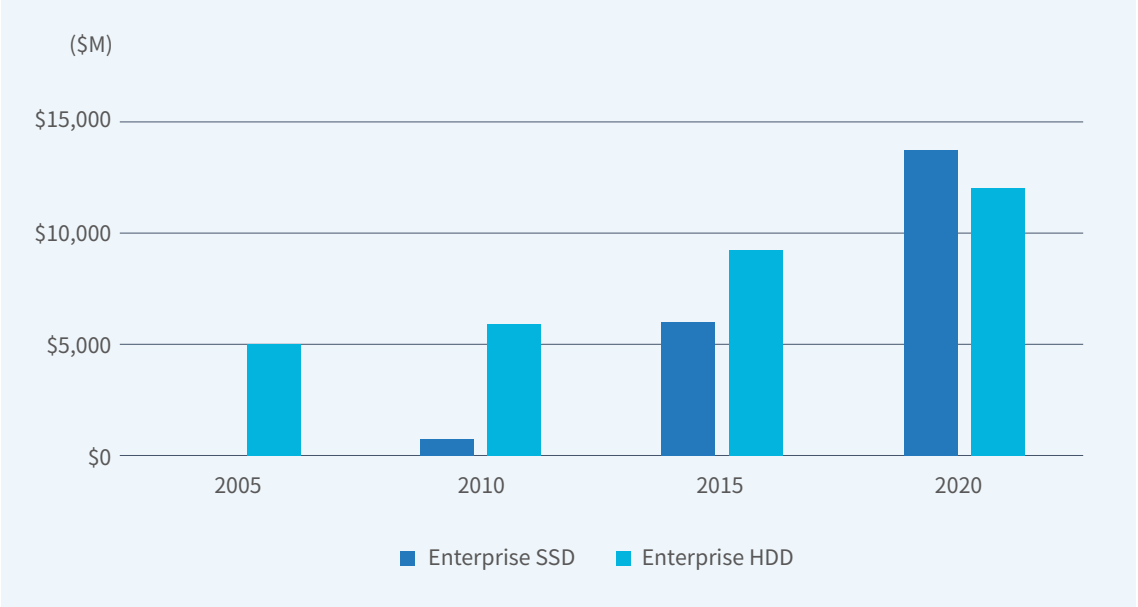
Fig. 6 Driving Factors for Enterprises/Organizations to Use All-Flash Systems as Secondary Storage



Source: IDC, 2021

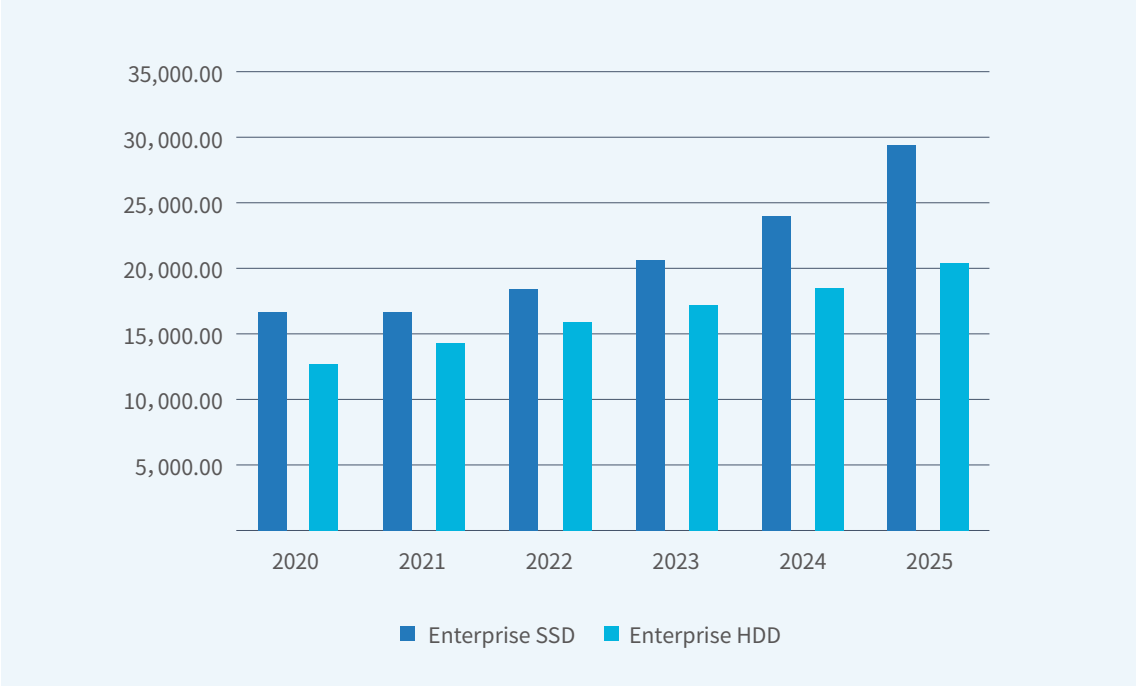
According to IDC research, the global total of spending on enterprise SSDs surpassed that of traditional HDDs in 2020, and this trend is expected to continue.

Fig. 7 Worldwide Enterprise Spending on HDD and SSD, 2005-2020



Source: IDC, 2021

Fig. 8 Worldwide Enterprise HDD and SSD Spending Forecast ($M), 2020-2025



Source: IDC, 2021

Accelerated data protection deployment in multiple services and application scenarios:

As an increasing number of industry customers go digital, business development and innovation heavily rely on data stability, availability, and reliability. Data becomes an important enterprise asset. This makes it subject to multiple threats and risks. Natural disasters, virus attacks, unauthorized intrusions, and system software and hardware faults all pose challenges to data security. Effective data protection has become an urgent need. Therefore, it is imperative to accelerate data protection deployment in multiple services and scenarios.

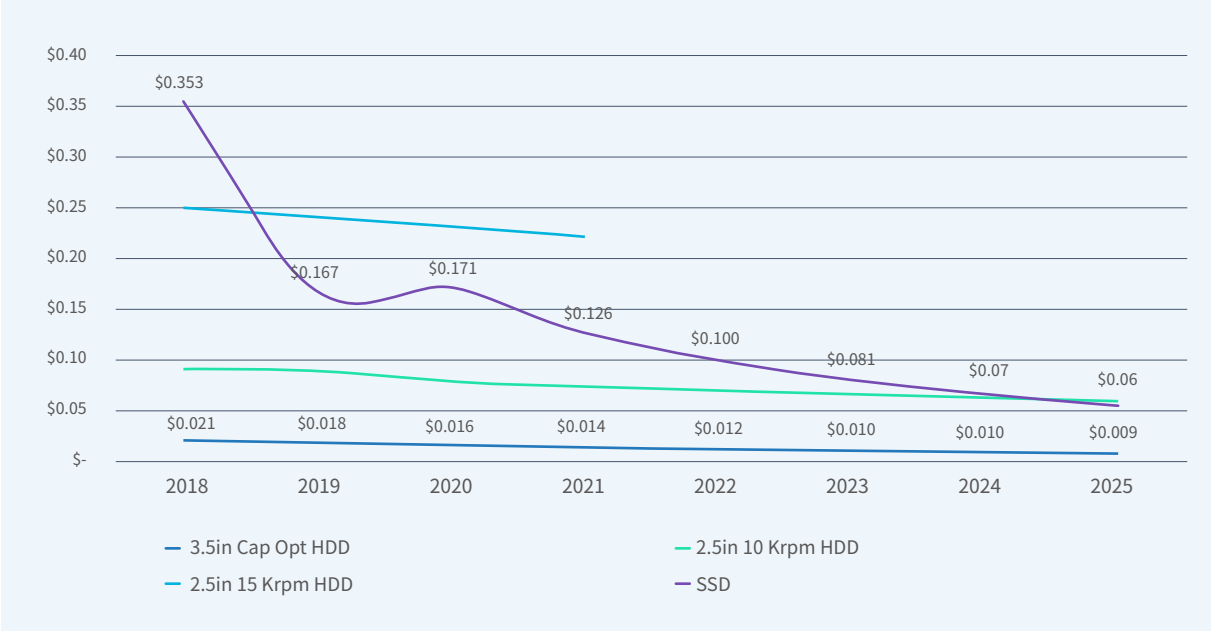
Increasing data volumes, diverse service types, and frequent invocations exponentially drive the amount of data to be stored and used. Mobility and cloud-based deployment increase the proportion of unstructured data of various types. What's more, data is getting hotter and hotter due to frequent access. In response to the preceding changes, data protection faces increasingly complex challenges. During planning and design, enterprises need a data protection solution that can store more data, support diverse types, provide better performance, and apply to a growing number of emerging application scenarios.

Opinion 2: All-Flash Data Centers Facilitate Green, Energy-Effective, and Sustainable Development of Enterprises

Driven by national policies and the demand for improved efficiency at lower costs, energy saving practices have been implemented in every aspect of organizations' production and operation processes. At the data center level, high-density all-flash systems are an ideal solution. In the future, as more organizations will look to transform digitally, data centers will become big energy consumers. Organizations can deploy all-flash data centers to greatly reduce footprint and energy expenses, thereby achieving sustainable digital development. And the resulting higher resource utilization and cost efficiency will become key factors to drive growth in all-flash system shipments.

Currently, price is the key factor that makes organizations reluctant to adopt all-flash systems. However, in the long run, the price gap between SSDs and HDDs is gradually narrowing. According to IDC, statistics show the price per GB of SSDs decreased by 25% annually on average from 2015 to 2020, and suggest this trend will continue over the next five years. IDC expects that the price per GB of enterprise-class, NAND SSDs will decline at a CAGR of –19% globally from 2020 to 2025. From the price trends about enterprise-class SSDs and 2.5-inch 10Krpm HDDs, it can be inferred that the price per GB of SSDs was 2.2 times that of HDDs in 2020. It is expected that by 2025, the price of enterprise-class SSDs will be lower than that of 2.5-inch 10Krpm HDDs. In the long run, with the decline in SSD prices, all-flash data centers will not only be energy-efficient but give you more bang for your buck.

Fig. 9 Worldwide Enterprise HDD and NAND SSD $/GB Trend Comparison, 2018-2025

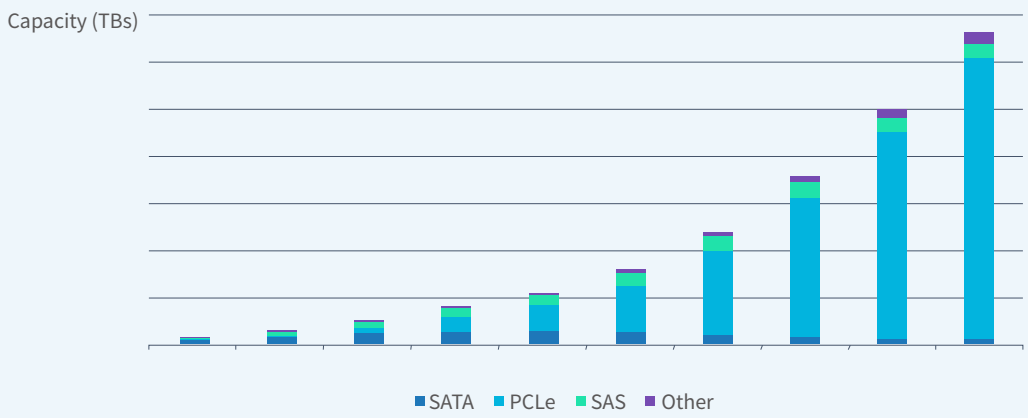


Source: IDC, 2021

Opinion 3: Flash Technologies Are Getting Better and Driving IP Networking for Data Centers

All-flash storage features better manageability and maintainability. In addition, SSDs deliver more flexibility in size and are available at multiple lengths, widths, and heights. In terms of interfaces, NVMe is gradually becoming mainstream due to its excellent performance in reducing storage network latency, and SSD interfaces will transition from traditional SATA/SAS to PCIe (NVMe). PCIe SSDs can further release performance and make hardware directly communicate with CPUs to improve the response speed, while NVMe unifies interface protocols to allow interconnection between multiple OSs.

Fig. 10 Worldwide Enterprise SSD Capacity by Interface Forecast, 2015-2024



Source: IDC, 2021

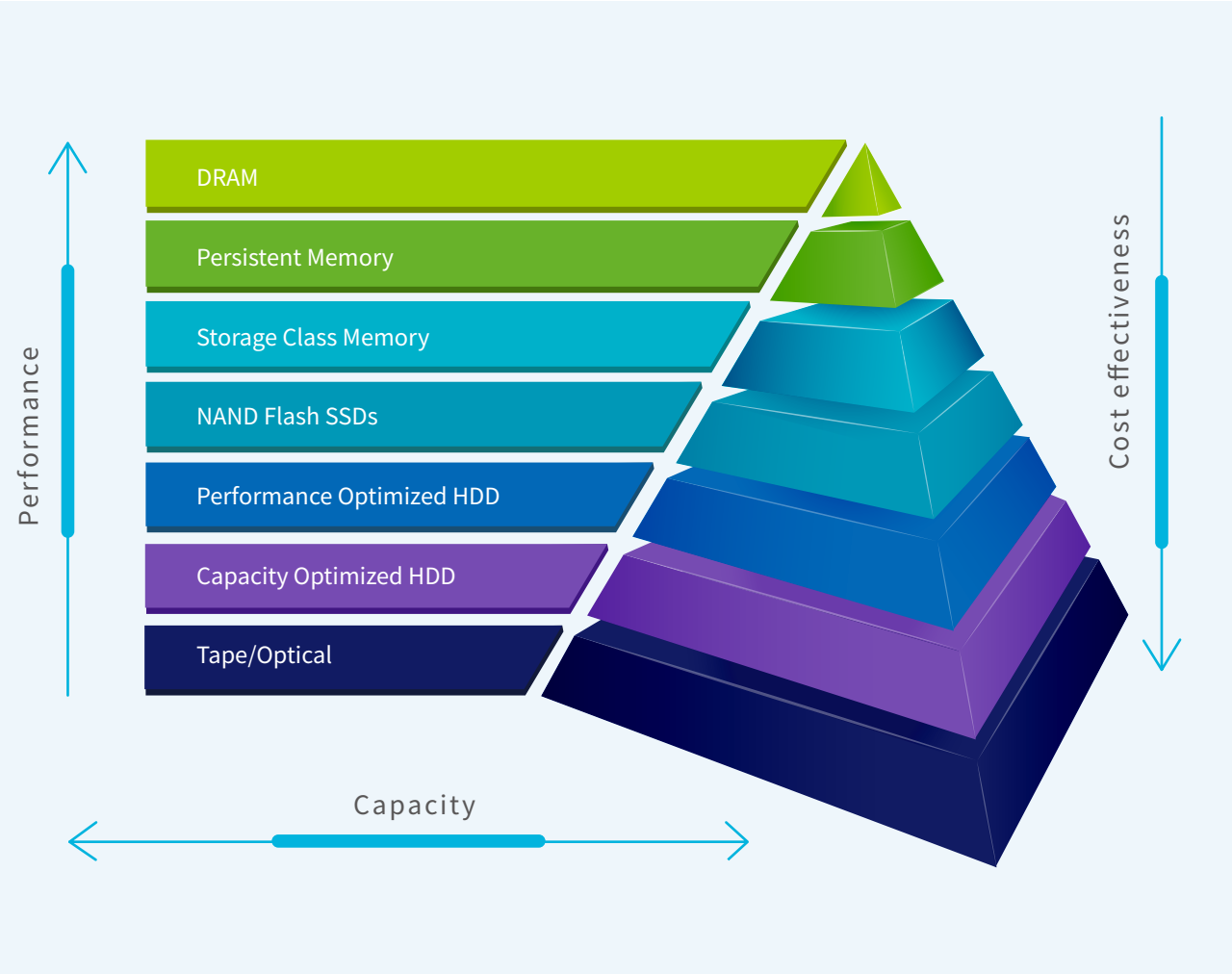
To further extend the NVMe protocol, NVMe over Fabrics is introduced. Currently, there are two mainstream forms of NVMe over Fabrics: FC-NVMe and NVMe over RoCE. The former is adopted in many mainstream storage networks because FC networks deliver certain advantages in stability. NVMe over RoCE is considered to be a more promising option with advantages in more open Ethernet and rapid improvement of bandwidth speeds. It has been five years since NVMe over Fabrics was released. This technology has gradually matured, laying a solid foundation for its large-scale implementation. As the evolution trend, NVMe over Fabrics will leverage its performance and cost advantages to drive architecture evolution, especially in the storage-compute decoupling scenarios. Organizations who move to all-IP networks will have access to higher transmission performance, wider connectivity, and optimal manageability, as well as simpler O&M.

Opinion 4: Memory-Driven Infrastructure Will Promote the Development of All-Flash Data Centers

With the advance of digital transformation, organizations face new challenges, including those presented by real-time data. Facing new service requirements, the memory-driven infrastructure is going mainstream. With enough flexibility, it provides optimal latency for workloads and uses cutting-edge technologies such as AI, machine learning, and deep learning to provide high data concurrency at low latency for workload expansion.

The memory-driven infrastructure uses emerging persistent memory technologies to reduce data access latency while delivering superior performance and enterprise-class storage management capabilities. It leverages technologies such as SCM to meet the real-time, availability, and functionality needs of next-generation applications. SCM provides many benefits. Although SSDs deliver higher performance than HDDs, there are still latency bottlenecks between memory and SSDs. SCM provides DRAM-like performance and larger storage capacity at a lower cost per GB than traditional memory. Moreover, SCM's performance density increases efficiency in other aspects. For example, less back-end storage capacity is needed, further reducing the cost of all-flash data center deployment. As a technical trend, SCM still faces some challenges, such as high costs. Despite costing less per GB than DRAM, using SCM in a memory-driven infrastructure causes the highest cost per GB among all types of persistent storage products. In addition, some applications that do not perform partial overwriting cannot fully utilize the performance of emerging persistent memory technologies. To solve this utilization issue, some software programs supporting memory-driven infrastructures provide APIs that allow existing applications to leverage these technologies without application modification. Reliability verification and risks caused by architecture reconstruction also hinder enterprises' adoption of SCM. SCM will be first used in specific scenarios, such as real-time workloads, tiered data solutions, and caching applications. More and more hardware architectures are also being developed for a better use of SCM.

Fig. 11 Breaking Traditional Memory/Storage Architectures with Persistent Memory Solutions



Source: IDC, 2021

Opinion 5: All-Flash Data Centers Enable Cross-Domain Deployment of Data over the Full Lifecycle, and Intelligent O&M Is Essential

Future digital infrastructure is likely to interact more with front-end business, and gradually evolve to a ubiquitous infrastructure that spans the entire data lifecycle. Diversified deployments can both meet business requirements but also present challenges to IT infrastructure O&M. All-flash data centers allow organizations to consolidate workloads onto fewer platforms (with the increased infrastructure density of an all-flash system), reducing management costs and complexity while delivering higher unit performance and reliability.

In the future, digital infrastructure O&M will be more automated with the help of AI. AIOps, or AI for IT operations, is a practice of applying big data and predictive AI/machine learning technologies to various IT operations and management processes. This enables IT systems to detect anomalies, predict failures, and recommend remediation, as well as perform self-tuning, autonomous IT provisioning, load balancing, data migration, patching, and more. AIOps will be frequently used to detect correlations, anomalies, and patterns, and proactively predict infrastructure health, capacity requirements, and performance changes. It is increasingly the enabler of many cloud and digital infrastructure management tools and services. To improve efficiency, most AIOps-enabled tools require large amounts of relevant data and training to identify normal and abnormal patterns. Thanks to innovations in data flow processing technologies for continuous data analysis, higher levels of automated learning are possible given enough time. IDC believes that future digital infrastructure will largely depend on the ability of complex, interconnected cloud infrastructures to dynamically self-regulate and self-optimize against real-time changes in resource requirements, application performance, and end-user experience. AIOps is expected to evolve into an important technical foundation for this vision.

Chapter 4

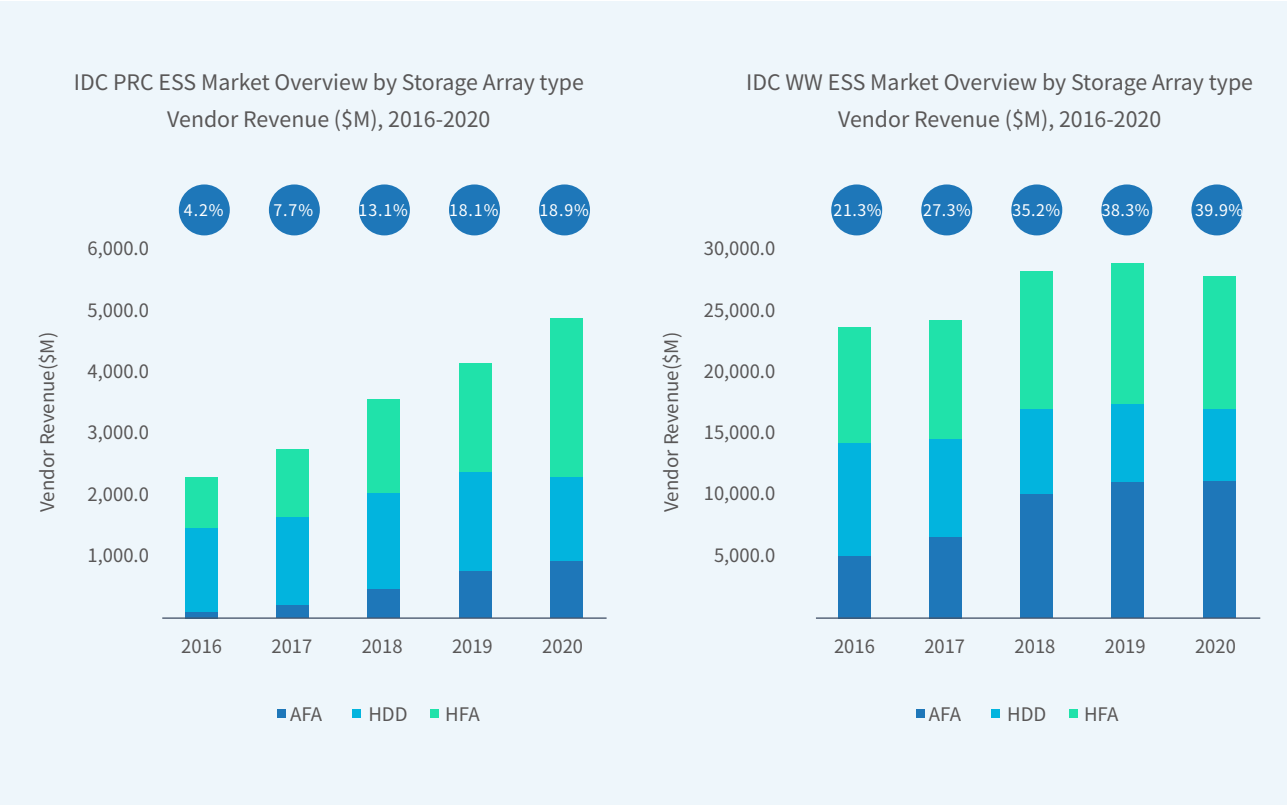
Best Practices for Deploying All-Flash Data Centers

4.1 All-Flash Market Overview

Digital transformation opens a gateway to new opportunities for organizations, driving R&D investments, platform innovation, and new practices such as digital industrialization, industry digitization, smart services, and industry ecosystems, which will foster new consumption patterns. This process involves organizations to emphasize environmental protection and develop energy-efficient practices for a green economy. In terms of constructing data center infrastructure, data centers based on servers, storage, and networks are transforming toward goals that fit into various industries — better performance at lower costs and better ROI, offering more value. Because data centers handle centralized storage, compute, and switching of data resources, organizations wanting to transform digitally must invest in a modern data center that fits their business needs.

All-flash arrays are becoming a rigid demand due to their outstanding performance and energy saving advantages. According to IDC China's external storage system report, the market share of all-flash arrays reached 18.9% in 2020, with a growth rate of 24.0% and higher than the ESS market average growth rate of 17.5%. There is plenty room for AFA market to grow in China. Globally, all-flash arrays took up nearly 40% of the external storage market, of which there was a 4.0% overall growth decline due to the 2020 pandemic. IDC forecasts that globally all-flash arrays will likely grow at a 7.5% CAGR from 2020 to 2024, accounting for more than half of the external storage market sales by 2024.

Fig. 12 PRC and Worldwide External Storage Market Overview by Storage Array Type
Vendor Revenue ($M), 2016-2020

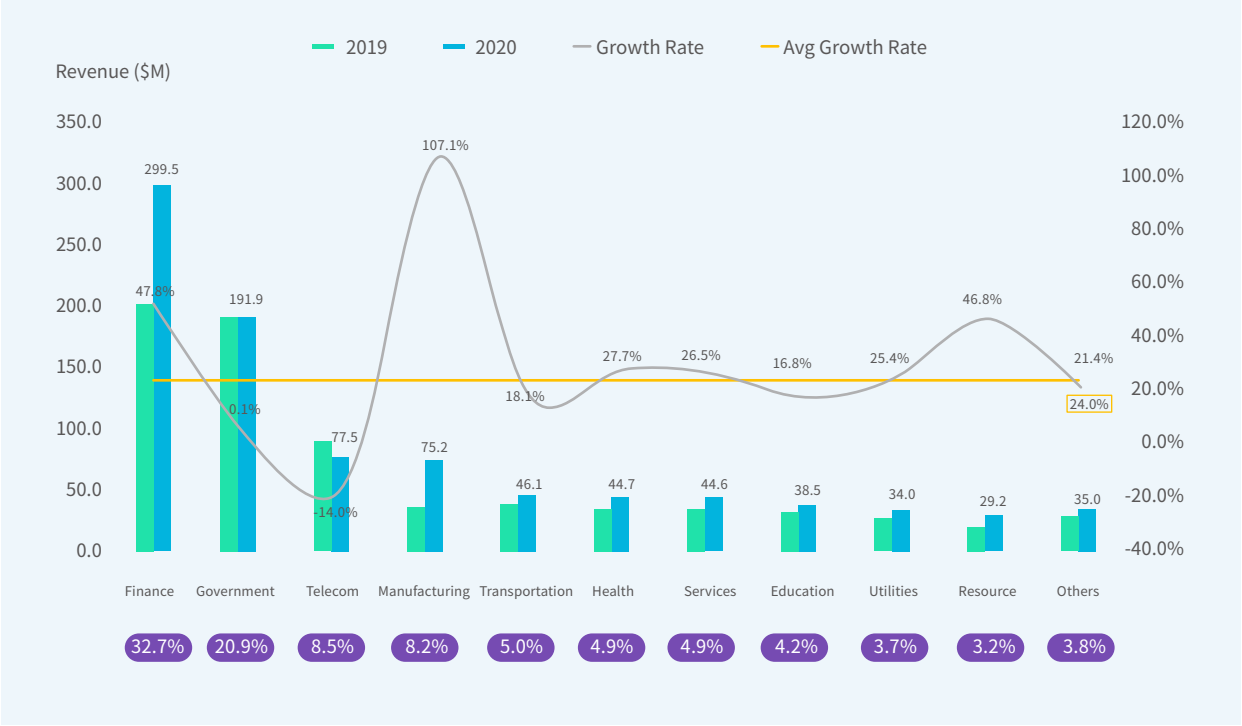


Source: IDC, 2021

4.2 Benefits of All-Flash Data Centers in Industries

All-flash storage has been widely used in industries that require high performance and stability, such as finance, government, healthcare, telecommunications, and manufacturing. Leveraging on the advantages of all-flash technologies, the all-flash data center will help more enterprises unlock the value and potential of data.

Fig. 13 PRC AFA Market by Vertical, Vendor Revenue ($M), 2020/2019



Source: IDC, 2021

Finance: All-Flash Data Centers Ensure Data Security and Support Financial Service Innovation and Development

-Industry Features and Deployment of the All-Flash Data Center

As mobile becomes more prevalent, digital financial services (including banking, insurance, and securities) are developing rapidly, giving rise to new online, intelligent, and contactless features. New financial services, such as digital banking, personal wealth management, digital lending, and omni-channel payment, are emerging one after another, creating massive amounts of data, which used to mainly be structured, but now are becoming increasingly unstructured and diversified. With the acceleration of digitalization and informatization, the financial industry has turned its focus to ensuring the high availability of infrastructure. Financial institutions and organizations whose businesses depend on data have high requirements on the security, business continuity, disaster recovery, and backup of critical and sensitive data. In addition, the industry as a whole has high requirements for service response. As mobile services increase and their scale expands, high concurrency and ultra-high peak rate scenarios are becoming more frequent. More and more organizations rely on stable, efficient, reliable, and elastic IT infrastructure to achieve business goals and plan strategically. According to IDC research, data security and compliance are the top priorities of digital financial construction, followed by data interconnection and cross-department sharing.

The all-flash data center provides robust support for scientific and technological innovation in the financial industry by upgrading media, integrating resources, and reorganizing and building an architecture that meets the digital development needs of enterprises. IDC China's external storage report shows that in 2020, the demand for all-flash arrays in the financial industry accounted for 32.7% of the total all-flash array market, and this continues to grow. Considering the financial industry's needs for high storage security and a high data access rate, organizations usually prefer mid-range and high-end all-flash storage arrays. About 70% of their expenditure goes on all-flash products with a unit price of 100,000 USD or above.

The all-flash array offers high performance and a low failure rate. To meet the different data processing requirements of the financial industry, the all-flash data center provides solutions with high performance and stability to improve business agility. These solutions deliver high input/output operations per second (IOPS), low latency, easy scalability, and high security for key applications and databases, and effectively support data disaster recovery, backup, and archiving of financial institutions.

-Major Workloads in the All-Flash Data Center

Take the application of a database system in an all-flash data center for example. A stable and high-performing financial service database system is critical to the performance of the application system, and a fast, reliable storage system is integral to the database.

- For online transaction processing (OLTP) databases, a large number of data updates cause high concurrency, so storage needs to handle high IOPS at a low latency. At the same time, due to the direct connection between the transaction system and the user, it is necessary to maintain high stability in the transaction process.
- For online analytical processing (OLAP) databases, data is analyzed and queried, and reports are generated based on multiple dimensions to aid decision-making. A large number of query operations have certain requirements on the overall processing time of application tasks. Disk systems are evaluated mainly based on their throughput capability.
- For hybrid transaction/analytical processing (HTAP) databases, low I/O latency and high data throughput are important requirements for storage.

For OLTP database applications running in all-flash data centers, local NVMe SSDs or high-end NVMe all-flash storage are suitable choices. Compared with traditional technologies with the same space and power consumption, the flash technology provides several times the IOPS and greatly shortens the response time. Database memory is more efficient, and operations are concurrent to ensure business continuity of the transaction system. End-to-end NVMe all-flash storage is required for OLAP database applications, due to capacity reasons. Under the same IOPS requirement, flash technology can significantly reduce footprint and power consumption. Similar to the competition between disks and tapes at the component level, all-flash data centers using end-to-end NVMe flash technology can achieve higher concurrency, higher throughput, and lower I/O response latency while reducing the storage capacity cost per gigabyte, which will be the dominant technology adopted in the future all-flash data center.

Case Studies

ITAU——The All-Flash Data Center Facilitates Collaboration and Business Development during the Pandemic

ITAU is the largest bank in Latin America and one of the largest in the world. It has about 96,000 employees and is present in more than 20 countries. Following the outbreak of COVID-19, ITAU closed its physical branches. As a result, the number of employees working remotely increased from 3,000 to 22,000 and the volume of online and mobile banking businesses increased sharply. ITAU has two data centers that were built in 2014. Increasing workloads caused sub-par system efficiency and service system performance. In particular, the statistical analysis system (SAS) was affected. This limited the performance of the core debit card service system and made overcoming the storage bottleneck the bank's top priority.

Through Proof of Concept (PoC) tests, ITAU finalized and deployed an all-flash solution that provides stable storage performance. The all-flash solution features NVMe, innovative hardware, and intelligent algorithms, fully unleashing the potential of all-flash. In terms of data reliability, the all-flash solution supports RAID-TP so that the system can tolerate the failure of three disks at the same time with zero service disruption, ensuring banking and financial business continuity and mission-critical business stability.

Case Studies

The All-Flash Data Center NVMe over RoCE Enables a Large State-owned Bank to Achieve Industry-Leading Innovation

In 2018, a large state-owned bank made "technology-led digitalization" one of its top priorities in the new round of strategy-making. As the digitalization of banking speeds up, requirements for information systems also get increasingly higher. Since the existing FC network develops relatively slowly, the commercial network technology still hasn't broken through the bottleneck of a 32G-bandwidth, so it might be hard for the current bandwidth to deal with the future business development. With the continuous expansion of business scale, it is expected that storage will face problems due to the bottleneck of bandwidth. Second, the two sets of networks – for customers and storage – adopt the IP network and FC network, respectively. Because the two technology stacks are different, two teams are needed for operation and maintenance, meaning that when a failure occurs, a lot of time will be spent on checking the network connection, and the identification of the problem and clarification of responsibilities are delayed.

After comprehensively reviewing its IT infrastructure architecture, the bank took the lead in building the next-generation storage network. It selected the NVMe over RoCE solution, completed the technical test of all-flash storage in the NVMe over RoCE environment, and deployed the system successfully. At present, the all-flash data center could ensure the uninterrupted service of its mainframe when failures occur at the controller, the front-end interface, or the front-end business lines. The network switch will enable LLDP, greatly reducing the time required for service switching. In the meantime, RoCE-based networking can provide high performance, with a 7:3 read/write ratio and an 8 KB I/O size. The overall performance is greatly improved compared with FC in the full-random I/O model. In the scenario of network congestion, compared with the common RoCE, the performance in bandwidth, IOPS, latency, and unexpected high traffic scenarios is significantly improved.

Government: Digital Service Capability Improvement for Reliable and Available Data

-Industry Features and Deployment of the All-Flash Data Center

Emergencies such as the COVID-19 make the governments around the world realize the importance of digital transformation to the effective prevention and control of the situation and the rapid recovery of the economy. Digital government transformation serves as a good example for various industries to follow. In China, "speeding up the construction of a digital economy, society, and government, and driving the reform of production modes, lifestyles, and governance modes through digital transformation" is an important topic in the government plan.

According to the IDC's prediction and analysis on global digital government transformation, by 2021, 45% of global governments will employ AI and process automation to provide civil services; by 2022, 60% of organizations will deploy AI technologies to recruit, train, and retain employees, and enhance their future job skills; and to ensure trust, by 2025, 65% of government agencies will pay more attention to protecting the reliability and privacy of their digital assets, and apply predictive analysis to identify, control, measure, and address reliability risks.

The Chinese government is also conducting its own digital transformation and has made stable progress in implementing the network power, digital China, and smart society strategies. In addition to embracing new technologies, such as big data, AI, and block chain, the Chinese government is also updating its service flow modes and widening the coverage of government services. The Chinese government will keep increasing its input in building its information capacity, introduce a more sophisticated flow automation capability to handle civil services in real time, and make investment in building a digital operation model in order to provide more regulated, intelligent, digital work solutions.

According to IDC, the government sector accounts for 20% of the total demand for external all-flash storage arrays in the Chinese market. In the future, China will further accelerate the construction of a digital government, achieving remarkable improvements in social informatization level and fulfilling public requirements for government services, including healthcare, education, and elderly care.

-Major Workloads in the All-Flash Data Center

In diverse digital government transformation scenarios, governments are promoting digital government services and optimizing governance capabilities. In China, many cities are promoting utilization of facial recognition systems to provide better services. Taking "intelligent taxation" as an example. By using facial recognition technology, a huge number of remote taxation terminals can upload mass files and materials by using a high-speed portable scanner. An all-flash data center provides stable support to transfer and store mass files and enables a fast response to taxation services, creating a "face-to-face" service experience in online taxation. The "Smart City" is another highlight of global digital government transformation. The IoT connectivity can support environment monitoring, public utilities, parking management, and elderly care. Efforts must be made to promote deployment of IoT applications with low energy consumption, multiple connections, and wide coverage. Along the IoT application path, the all-flash data center provides a higher transfer speed, more stable performance, and greater cost-effectiveness in terms of network, compute, and storage, across multiple dimensions, including data acquisition, transfer, analysis, and application.

Case Studies

Ministry of Finance of the Republic of Indonesia — Infrastructure Upgrade for Higher Service Level of Government Affairs

Background

The Ministry of Finance of the Republic of Indonesia (Ministry of Finance) is an important government agency that is in charge of economic development and protecting the financial assets and finances of Indonesia. The Ministry of Finance has adapted well to managing state financial affairs by leveraging information and communication technologies. To improve the service level, the Ministry of Finance must upgrade its infrastructure, such as servers, storage systems, and network, from time to time. However, in view of the characteristics of financial affairs information, storage migration is faced with stringent data reliability requirements. Meanwhile, the Ministry of Finance's data is growing at PB scale each year and therefore, scalable storage systems must be deployed to meet its capacity requirements. The growing amount of data has also posed a challenge to the performance of technology infrastructure. Traditional storage was subject to performance bottleneck and data transfer latency increased to more than 10 milliseconds.

Solution

The Ministry of Finance deployed an all-flash data center whose architecture is fully compatible with its existing architecture. Data migration can be implemented either at the management program level or at the storage level, with zero impact on upper layer services. Supported by the active-active data centers solution and the fully interconnected, reliable architecture, it tolerates the failure of up to seven out of eight controllers and full load balancing among controllers, further ensuring business continuity and always-on services. In this way, the infrastructure provides highly available data services as well as an open, efficient, and reliable data management environment.

Results

The all-flash data center improves data synchronization performance, reducing report processing time from 5 to 2 hours. With the highest level DR assurance, no data will be lost in the event of power outages or natural disasters. The all-flash data center also improves system reliability and simplifies networks and O&M operations. Furthermore, it has a small footprint. A chassis can contain up to 36 NVMe SSDs, increasing disk density by more than 40% while saving energy and physical space in equipment rooms.

Telecom: All-Flash Data Centers Improve Internal Management and Enable Service Innovation

-Industry Features and Deployment of the All-Flash Data Center

The telecommunication industry is active and performs well in digital transformation. With next-gen IT technologies, it also helps other industries achieve digital development. With the implementation of technologies such as 5G and cloud computing, telecom operators are facing the pressure of higher workloads internally. Optimized data capabilities can better support the BSS system, which has the potential to increase customer loyalty and optimize the OSS system to improve O&M efficiency. Externally, telecom carriers provide value-added services such as smart transportation, smart retail, IoV, gaming and entertainment, and AR/VR applications. A huge amount of new data and an increasing number of data innovation scenarios require fast data access. The data centers of telecom operators must provide high-concurrency, low-latency data transmission, and service continuity. This places high requirements on the performance and stability of the infrastructure.

According to the IDC's research, many telecom carriers' devices are aging and out of warranty. As a result, the device failure rate is high and the performance is insufficient. In addition, traditional storage devices occupy a large amount of space, consume a lot of energy, and cost a lot to maintain, resulting in high operating expenditure (OPEX). In light of the rapid increase of concurrent key services, such as billing and CRM services, telecom carriers are more willing to purchase all-flash storage arrays. They expect to use powerful all-flash arrays to accelerate the consolidation and replacement of old devices, modernize storage systems, and maximize their storage investments. This enables the telecom industry to cope with service growth and ensure service efficiency.

-Major Workloads in the All-Flash Data Center

Improving customer loyalty is one of the key KPIs in emerging value-added service scenarios in the information era, no matter the industry. In many cases, end-user experience is what forces enterprises to undergo digital transformation. Again, this highlights the value of data, which drives the comprehensive optimization of customer experience. Take the distributed database scenario for example. In the past, telecom operators mainly deployed databases in silos. Compute and storage resources were bundled, which made the IT infrastructure complex and the utilization rate of IT resources low. In addition, unexpected service peaks could have very high resource requirements. This made it difficult to adapt to service changes. Only computing resources could be rapidly expanded, which led to high management costs.

All-flash data centers can solve the preceding problems. All-flash data centers support decoupled storage-compute deployment to resolve the reliability weaknesses of a coupled storage-compute solution. In addition, resources are optimized to reduce resource redundancy and implement elastic expansion of compute and storage resources. Backup and recovery also use mature shared storage to ensure data security and stability and promote data innovation. Flash technologies are now continuously optimized; emerging persistent memory technologies reduce data access latency and provide excellent performance and enterprise-class storage management functions. In addition, NVMe interfaces are becoming mainstream, making all-flash data centers more manageable and improving data security. In addition, the NVMe transmission standard allows SSDs to directly connect to the CPU through the PCI-E channel, effectively reducing data latency. Registers do not need to be read when executing commands, lowering the latency of the controller and software interface, improving the data transfer speed, and enabling data innovation. The physical footprint is also reduced, the disk density is improved, and the power consumption is lowered. In conclusion, the enhancement of the all-flash technology helps users reduce costs and increase efficiency while achieving the ultimate goals of being green and saving energy.

Case Studies

A Telecom Carrier's Provincial Branch in China — The All-Flash Data Center Ensures Stable, Sustainable Core Services and Improves Customer Satisfaction

As its service volume increases, a telecom carrier's provincial branch (referred to as "the company") uses Internet channels to handle service requests from subscribers. As a result, the service workload increases significantly compared with the counter service period, and the workload peak keeps changing instead of being fixed in the past. This gives rise to challenges for the company's IT system performance and stability. To ensure customer service efficiency and quality, the company developed the TeleDB database based on MySQL. It reconstructed the IT system architecture in stages to improve database reliability and automate O&M, during which the company faced challenges from three aspects: 1. High reliability is required to ensure smooth running of core businesses covering a wide range of end-users; 2. Agility is required to meet the development of new business; 3. Cost-effectiveness is always the key task by choosing optimized architectures to say cost as much as possible.

Based on the requirements above, the company adopted a decoupled storage-compute architecture for its new all-flash data center, so as to ensure the reliability and stable operation of its core services. It also used a large number of NVMe SSDs to improve service performance. With the help of the all-flash data center, failure recovery was shortened from hours to minutes; comprehensive energy consumption was reduced, and the floor space was significantly saved. In addition, such a future-ready architecture facilitates the company's development of new businesses in the era of "Internet+" and mobility.

Manufacturing: All-Flash Data Centers Facilitate Efficient and Reliable Storage Resource Utilization in Intelligent Manufacturing

-Industry Features and Deployment of the All-Flash Data Center

In the digital transformation process, the manufacturing industry must not only support the operation of large-scale ERP systems, but also take into account the application of new technologies such as the Internet of Things (IoT) and sensors. The key to successful transformation lies in properly managing the entire data lifecycle, from collection, storage, and transmission, to presentation, analysis, and optimization. In the past, manufacturers have built large-scale application systems such as ERP, CRM, PLM, MES and EDA, but information silos exist between different systems. In production and manufacturing, more and more sensors are used in automated production, and more data needs to be collected, stored, computed, managed, and utilized, requiring strong support from new data infrastructure.

To achieve intelligent manufacturing, the manufacturing industry is undergoing a digital reform. All-flash data centers can provide reliable and agile real-time data services for high-end manufacturing. According to IDC's report on China's external storage, the manufacturing industry's demand for all-flash storage arrays has increased significantly in 2020, with a doubled year-on-year growth. All-flash products are being procured not just for core production systems, but for edge service scenarios as well. More companies join digital transformation, from high-end electronics in the past to process manufacturing enterprises such as car and footwear manufacturers now.

-Major Workloads in the All-Flash Data Center

Taking the application of electronic design automation (EDA) tools for example. Chip product manufacturers have handed over the whole process of circuit design, performance analysis, and IC layout design to computers. Logical and physical verification in the design is very important, and manufacturers need to ensure that products meet all performance requirements before production is started.

EDA requires high performance from the entire data center. The whole process of computer aided design, manufacturing, testing, and computer aided engineering includes two parts: logic simulation at the front end and physical simulation at the back end. Front-end logic simulation involves RTL encoding and comprehensive optimization simulation. The I/O features are OPS services with KB-level small-file reads and writes and large numbers of metadata operations, which place high requirements on file system I/O performance. Back-end physical simulation involves conversion of netlists to layouts and layout optimization simulation. The workload features a lot of GB-level large file writes which require high bandwidth. According to the service model analysis, the file system performance is the top 1 requirement of EDA simulation tools.

The manufacturing industry has constantly raised its requirements for data center performance. For example, simulation is a crucial step in the manufacturing process. Simulation must be completed as quickly as possible without faults to save time for R&D and design, which poses challenges to the performance of infrastructure, especially storage systems. Storage system performance must not become the bottleneck. In addition, the simulation program requires parallel multi-tasking and storage resource sharing. As different tasks have different requirements for storage performance, they tend to contend for NAS resources, so how to avoid NAS resource preemption is an urgent problem.

High-end all-flash products typically use streamlined and reliable active-active NAS solutions to ensure stable online simulation in the manufacturing industry. The solutions feature automatic failover, recovery, smooth upgrade and expansion, and five-fold switchover speed. They also provide higher IOPS and lower latency for EDA simulation. To address NAS resource preemption, an all-flash data center network can use various basic technologies to implement QoS, provide better service capabilities for specific network links, limit the OPS and bandwidth of project teams, and avoid resource preemption and mutual interference between teams. This solves network latency and congestion problems and effectively utilizes all-flash data center resources.

Case Studies

HiSilicon: – Building the All-Flash Data Center to Ensure Business Stability and Increase Customer Satisfaction

HiSilicon Semiconductor Co., Ltd is a leading provider of chips and solutions, and its products include chips and solutions for wireless network, fixed network, digital media, etc. HiSilicon boasts leading IC design and verification technology, and has an advanced EDA design platform with a sound development process and specifications. The EDA data of HiSilicon can be divided into four categories, including public data, project result-related data, project process-related data, and archived data. The overall simulation business accounts for a large proportion and thus has higher requirements for performance. The specific requirements for different types of data are as follows:

- Public data with more reading and less writing requires high IOPS and high reliability;
- Project result-related data with frequent reading and writing requires high reliability;
- Project process-related data of the front-end design phase is mainly small files and involves frequent reading and writing; it has high requirements for OPS, with the peak OPS of a single project easily exceeding 300,000. In contrast, project process-related data of the back-end design phase mainly consists of big files and also involves frequent reading and writing; it has high requirements for bandwidth, and the peak bandwidth requirement can be as high as 1.5 GB/s.
- Archived data marking the conclusion of a project needs to provide remote backup capabilities.

In order to meet the above requirements, HiSilicon uses all-flash NAS storage for its EDA business. Its speed and capacity greatly reduce the time required for simulation. It also deploys a full SSD-virtualized resource pool, a reliable system, a fully-connected reliable architecture, and a gateway-free active-active solution to ensure the 24/7 operation of the simulation business. In addition, the intelligent QoS service can manage multi-user OPS and bandwidth and ensure the priority of key businesses.

Healthcare: All-Flash Data Centers for Free Data Movement and a Solid Foundation for Smart Healthcare

The world's population is aging, giving rise to increasing health requirements. With this, the global healthcare industry is projected to maintain long-term growth with increasing investment in biotechnology, pharmaceuticals, and medical IT construction as well as increasing procurement of ICT infrastructures. Until now, China's healthcare industry has made remarkable achievements in digital development. A comprehensive information system that integrates computing network platforms, medical service software, and digital medical devices has been initially formed, and a system of data interconnection, regional collaboration, hierarchical diagnosis and treatment, and health services has been basically formed within hospitals.

According to the IDC's external storage market report for China in 2020, procurement demand for external storage systems in the healthcare industry has temporarily slowed down due to the pandemic, but is forecasted to increase again, driven by regional medical construction and the rise of remote medical services. The healthcare industry's demand for all-flash storage arrays is expected to grow at an estimated CAGR of 7.9% over the next five years.

Case Studies

Cliniques Universitaires Saint-Luc — All-Flash Reconstruction Enhances the Capabilities of a Unified Data Platform

As one of the largest hospitals in Brussels, Belgium, Cliniques Universitaires Saint-Luc built a unified information platform to manage all clinical, medical, and administrative information to improve medical services for better patient experience. The subsystem includes a series of applications such as databases, virtualization platforms, file sharing, email, and office applications. However, the rapid growth of medical services is bottlenecked by slow system response and insufficient storage space. The existing infrastructure had performance limitations and was difficult to scale, and storage was insufficient to meet the growing storage requirements of data centers.

The hospital decided to reconstruct its data center into an all-flash model. Now, the end-to-end data storage solution enables clinical research systems and databases running on the unified information platform to obtain high computing power and storage performance, and the high-end SAN storage provides excellent scalability.

Case Studies

Beijing Tongren Hospital — Data Silos Eliminated for Optimal Digital Services

About Beijing Tongren Hospital

Beijing Tongren Hospital, founded in 1886, is a large-scale comprehensive hospital with an annual number of outpatient and emergency visits reaching 2.876 million. The hospital's main goal is to transform from digital to intelligent, while meeting regional construction requirements of medical facilities.

Background

The existing storage arrays at the Tongren Hospital were outdated, with scattered HDDs in a siloed architecture. As the hospital's outpatient service volume grows and new services such as online appointment and telemedicine are introduced, the storage performance and core service systems stalled, resulting in prolonged response time. In addition, the hospital lacked a unified deployment platform, with a large number of devices in multiple service silos, causing complex and costly O&M management. Therefore, the hospital wanted to streamline the existing data architecture and eliminate data silos to maximize data value. One key aspect for Tongren Hospital was the need for integrated HIS and PACS core systems, which can enable analytics, diagnostics, and research of medical images in real time, and provide auxiliary medical support. However, the integration of HIS and PACS has the following challenges:

- The HIS system covers the hospital's full service process and needs to handle all patient information. When the HIS system is integrated with the PACS system, even more data will be generated. One major challenge for the hospital is finding an effective way to intelligently manage daily data. This includes: recording accurate and complete medical information, simplifying patient treatment and visitation procedures, reducing transactional workloads, and improving the quality and efficiency of hospital services.
- The PACS system generates the largest amount of operational data and has the strictest data precision and transmission requirements. During the collection, transmission, and storage of medical images, the hospital must address the issues of how to efficiently and accurately store massive volumes of unstructured data, and quickly invoke these data when necessary.
- New information systems (electronic medical records) and the enhanced clarity of medical images places more importance of reliable storage and stable transmission. Now, hospitals are expected to handle and process massive volumes of medical and 4K image data.
- Integrated HIS and PACS data has high requirements on the accessibility and accuracy of data. Between the payment desk and a doctor's office, the medical payment information in the HIS system needs to be updated instantly. Medical images generated by inspection should be reflected in PACS system in time. It is vital that PACS systems provide timely and accurate consultation information for doctors.

Solution

Tongren Hospital builds an all-flash data center and deploys the all-scenario flash acceleration solution, all-scenario data protection solution, and disaster recovery and backup solution. Together, this environment serves as an upgrade on the data center, improving data reliability and availability.

- All-scenario flash acceleration enables quick and easy data migration from existing production application systems (such as HIS and PACS) to the all-flash data center. For new service systems (e.g. Internet access service, online registration service, etc.), cloud resource pools and databases are directly deployed in all-flash data centers to meet agile deployment and capacity expansion requirements in the future.
- All-scenario data protection adopts the modern all-flash array architecture to ensure uninterrupted and stable service running, and uses hard disk wear and anti-wear leveling technologies to minimize service failures, meeting data protection requirements in large-capacity storage scenarios.
- In addition, Tongren Hospital adopts the solution of full disaster recovery of core data and hot backup of important data. The active-active configuration is enabled for key service storage, ensuring that core services are online in real time. It also uses the intelligent O&M management tool to implement end-to-end disaster recovery management and one-click drill, reducing O&M workload and improving O&M efficiency.

Results

With all-scenario flash acceleration and global topology visualized management, Tongren Hospital integrated HIS and PACS systems, improved storage performance, and tripled outpatient registration efficiency. In terms of service stability, core services run 24/7 without interruption, and active-active service access is established to eliminate single points of failure. This project reconstructs of the legacy siloed architecture for lower O&M costs, and the global topology enables centralized management of multiple physical devices, improving resource utilization.

To provide superb system stability, the all-flash data center offers decoupled storage-compute capabilities to facilitate real-time processing and expansion of massive amounts of unstructured medical image data, covering archives, recordings, videos, pathology, and PACS. In addition, high-end all-flash storage arrays with AI accelerator cards provide powerful computing capabilities. The system now uses deep learning algorithm to learn service load rules and predict service behaviors, injecting intelligence into read cache and QoS control processes. Other services include intelligent fault reporting, capacity and disk health prediction, performance tidal and bottleneck analysis, and anomaly detection. Together, the all-flash system can meet the future IT requirements of new policies, regulations, and services. The all-flash data center not only intelligently manages services and improves process efficiency, but also reduces O&M costs and accelerates data openness and sharing, helping hospitals maximize data value in digital transformation.

Chapter 5

IDC Recommendations

Organizations developing their storage infrastructure strategies must focus on five key aspects: agility, efficiency, availability, reliability, and cloud adaptability. To adopt the best strategy for digital transformation, organizations must understand business and technical issues, as well as be aware of the latest storage technologies and architectures. IDC lists the following technologies and architectures that deserve attention:

All-flash systems

All-flash data centers offer a smart choice to adapt to the digital future. Whether it is to take advantage of the high performance and capacity densities, to simplify infrastructure at lower energy and space needs, or to improve management efficiency, IT decision makers must determine an ideal strategy for building future storage infrastructures. Flash media costs are declining in recent years, which is why more and more organizations are using all-flash systems for secondary storage workloads such as backup, disaster recovery, archiving, and big data analytics. Considering the efficiency and the total cost of ownership of these solutions over the organization's expected storage lifecycle, these all-flash devices offer better savings than conventional alternatives.

Storage consolidation:

For organizations looking to update their storage systems, IDC recommends considering the option to consolidate workloads (that is, leveraging the increased infrastructure density of newer systems). The fewer platforms you buy, use, and manage, the more cost-effective it will be. By consolidating your storage, the new systems can consolidate more intensive workloads while delivering higher per-unit performance and reliability.

All-IP networks

In order to fully realize the potential of SSDs in all-flash data centers, IDC recommends considering innovations in storage networks besides the existing FC networks. NVMe over Fabrics is the trend, and NVMe over RoCE is considered to be a more promising option that can improve performance while reducing costs and simplifying management.

Data reliability and availability

For any organization working with critical workloads, service disruptions must be avoided. That means investing in key infrastructure offering cutting-edge technical specifications such as high data reliability and availability. For CIOs looking to refresh their storage infrastructure with non-disruptive upgrades, expansions, and technologies, using all-flash systems help speed up recovery and move data faster than HDDs during rebuilds. In addition, data protection is included as a key task in routine management. Organizations are recommended to enhance investment in data protection resources and raise requirements for data protection levels, so that data can be easily utilized under comprehensive protection.

Automation

Data management becomes more difficult as data grows and deployment mode advances, making automation essential to improve responsiveness and operational reliability. Automation is already a mature concept, and better tools can extend automation capabilities to hybrid cloud environments. In a broad sense, automation is not just a booster of efficiency, it can also help organizations meet a variety of agility and data governance needs.

About IDC

International Data Corporation (IDC) is the premier global provider of market intelligence, advisory services, and events for the information technology, telecommunications and consumer technology markets. IDC helps IT professionals, business executives, and the investment community make fact-based decisions on technology purchases and business strategy. More than 1,100 IDC analysts provide global, regional, and local expertise on technology and industry opportunities and trends in over 110 countries worldwide. For 50 years, IDC has provided strategic insights to help our clients achieve their key business objectives. IDC is a subsidiary of IDG, the world's leading technology media, research, and events company.

IDC China

No.36 3rd Ring Road

China 100013

+86.10.5889.1666

Twitter: @IDC

idc-community.com

www.idc.com