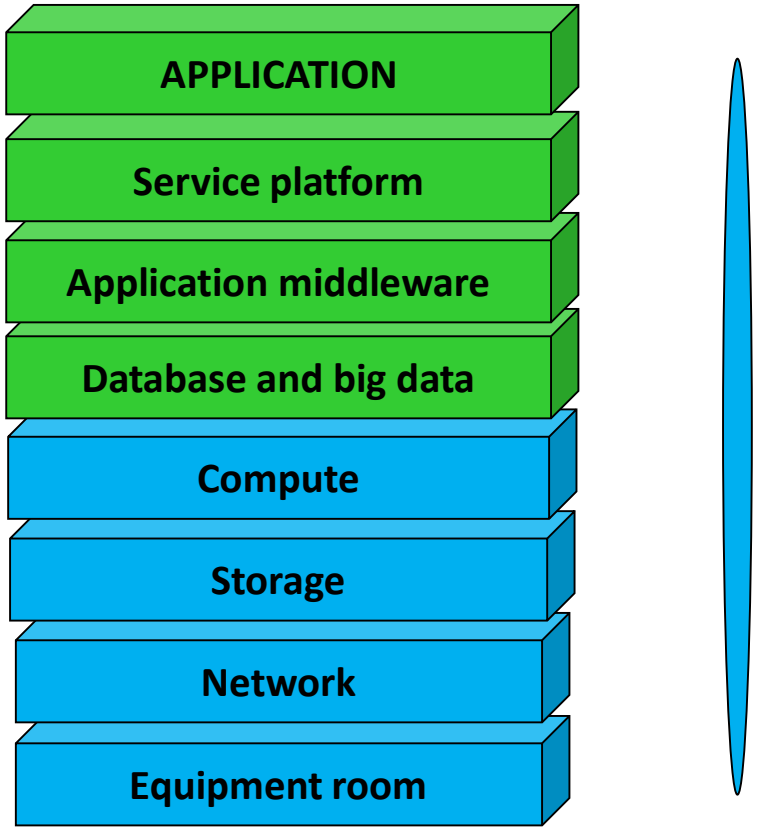
Basic Knowledge of Data Storage

Contents

New Data Center: New Compute, Storage, Network, and Energy



Layer 4 Application

- ERP
- PDM
- iSales
- iCare
- DW
- HRMS
- W3
- Email
- ...

Layer 3 Database, big data, and middleware

- WAS
- Oracle DB
- Integration via EIP
- Message middleware
- MySQL
- SQL Server
- LDAP
- DNS
- Big data

Layer 2 Compute, storage, and network

- Compute
- Storage
- Network

Layer 1 Energy

- Electric power system
- Cooling system
- Cabling system
- Fire extinguishing system
- Physical security
- ...

Reliable Storage Systems of High Performance and Large Capacity Dedicated to Data Centers

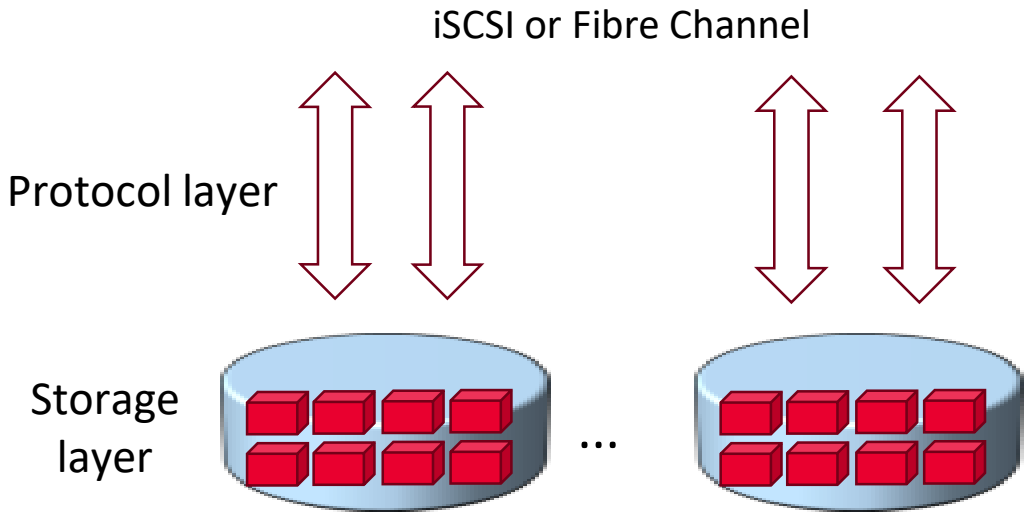
Storage Products for Individuals

- Requirements: cost-effectiveness and **fair and reasonable prices**
- Tolerance for disadvantages in response time, service life, and errors

Storage Products for Enterprise Data Centers

- Requirements: data integrity, **high reliability, high performance, and large capacity**
- Different from servers in high availability (HA) architecture, power failure protection, and quality control

Storage System Classification by Service – Block Storage

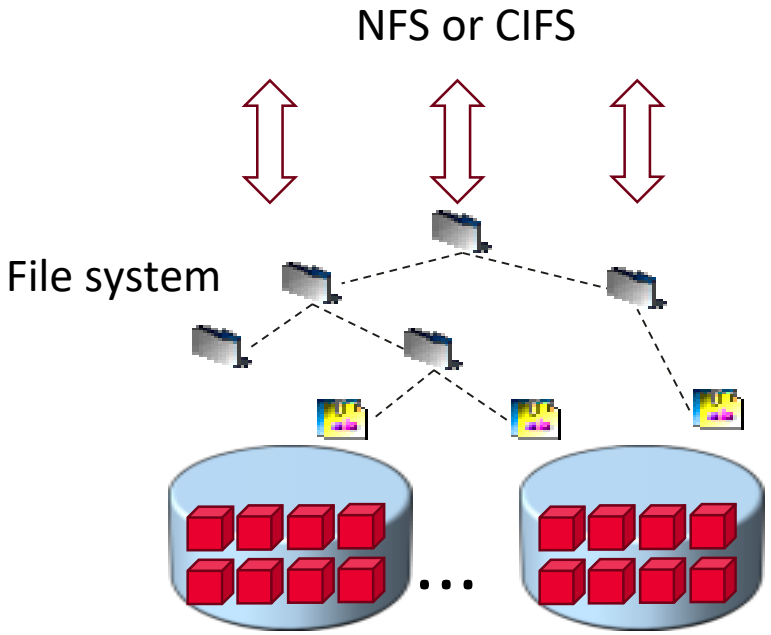


- Block storage is used to store structured data, that is, data is directly read and written by reading or writing one or more addresses from or into storage space.
- Generally, block storage is used to store important data of virtualization applications such as Oracle database and VMware.

Block storage:

- Advantages: direct access, minimized overhead, and highest efficiency
- Disadvantages: highest cost and poor scalability
- Application scenarios: Oracle database and VMware virtualization

Storage System Classification by Service – File Storage

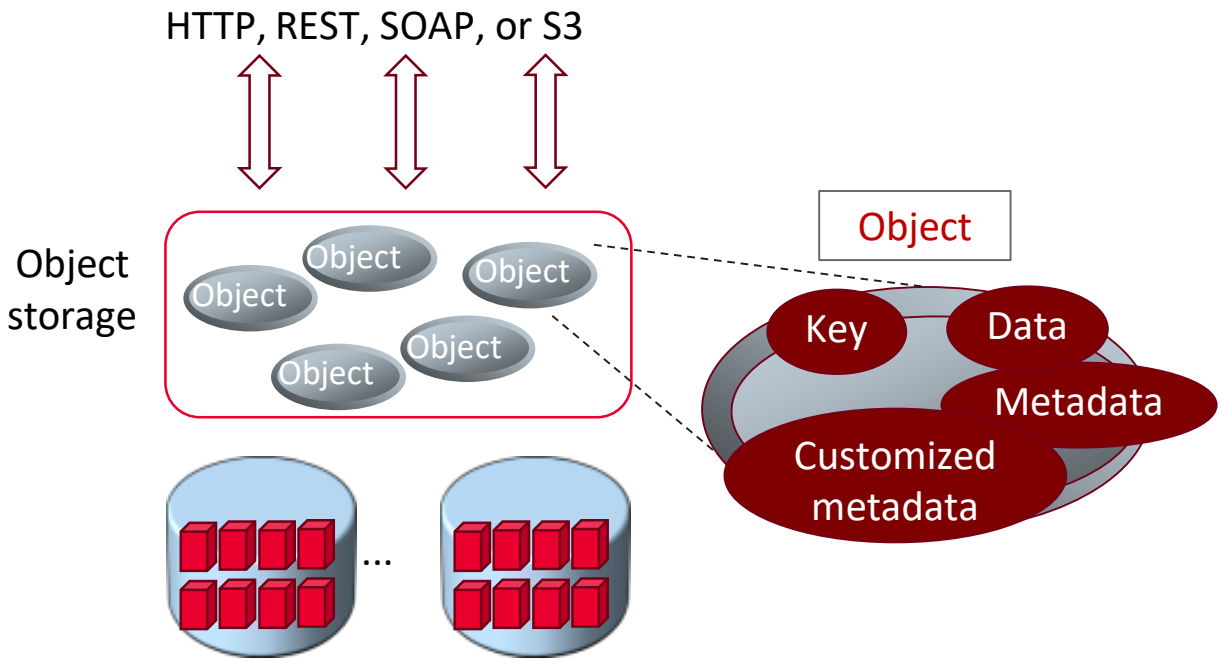


- File storage is mainly used to store unstructured data. Professional file systems are added to block storage devices to implement file sharing.
- Generally, file storage is mainly used to store file data, such as PACS (medical imaging) data, carrier CDR data, manufacturing (EDA) simulation data, oil exploration data, and HPC computing data.

File storage:

- Advantages: easy management and interconnection with applications
- Disadvantage: supports expansion, requiring ecosystem compatibility.
- Application scenarios: enterprises' internal application integration and file sharing

Storage System Classification by Service – Object Storage

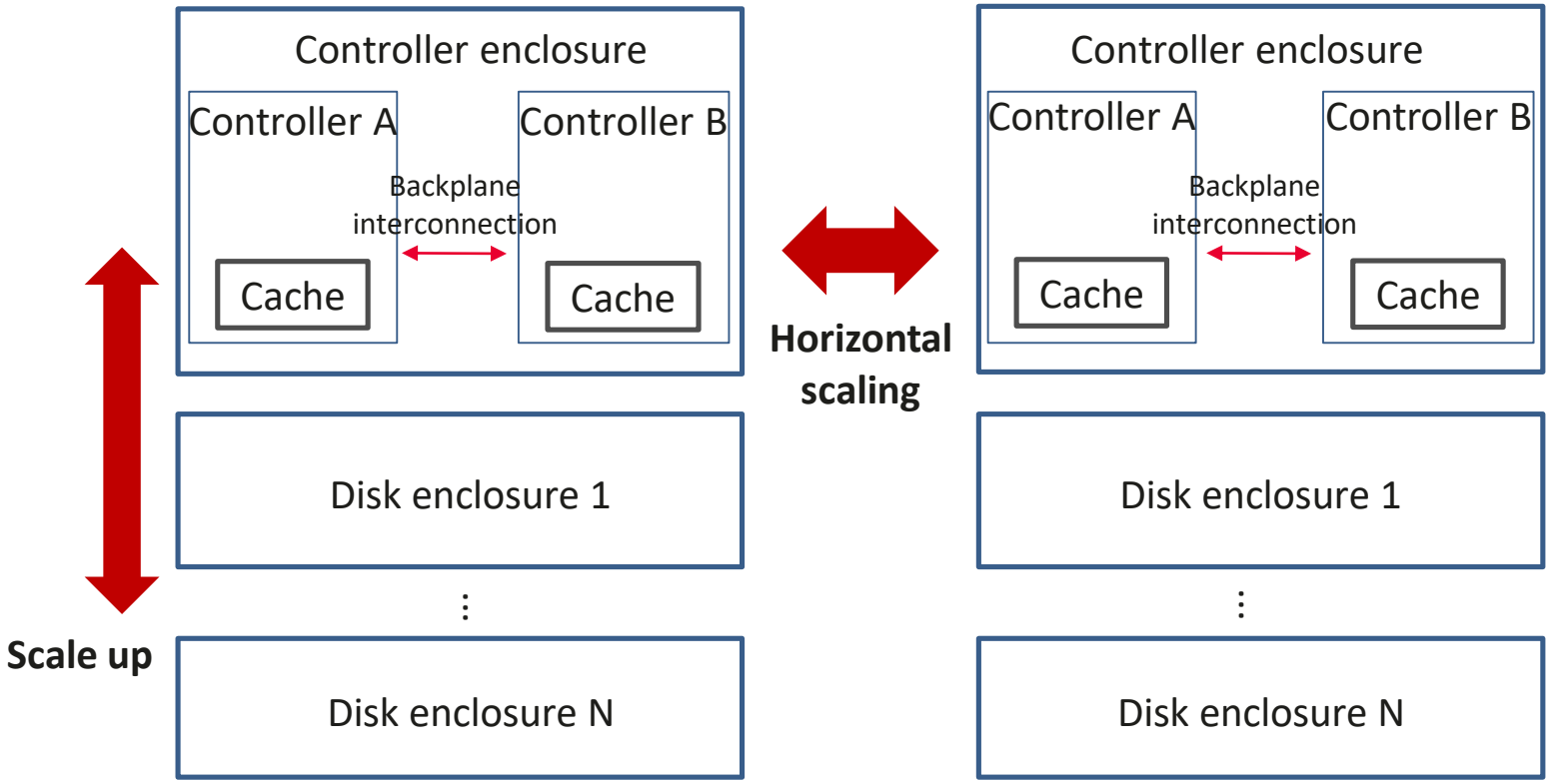


Features of object storage:

- Flat structure and nearly unlimited capacity expansion
- More intelligent self-management
- Standard Internet protocols and cross-region transmission capabilities
- Application scenarios: Internet-oriented storage, archiving, and backup

- Object storage, also known as object-based storage (OBS), is a network storage architecture. The differences between OBS and block storage or file storage lie in the interfaces (S3 interfaces) provided by OBS. OBS only generates an ID for the metadata of stored data and stores the ID, regardless of the data type.
- This storage architecture is mainly used in the application scenarios that have low requirements on performance but high requirements on capacity. It is mainly used to meet customers' requirements on large capacity and low price. The application scenarios include public cloud, Internet, and space leasing.

Storage System Classification by Architecture – Centralized Storage

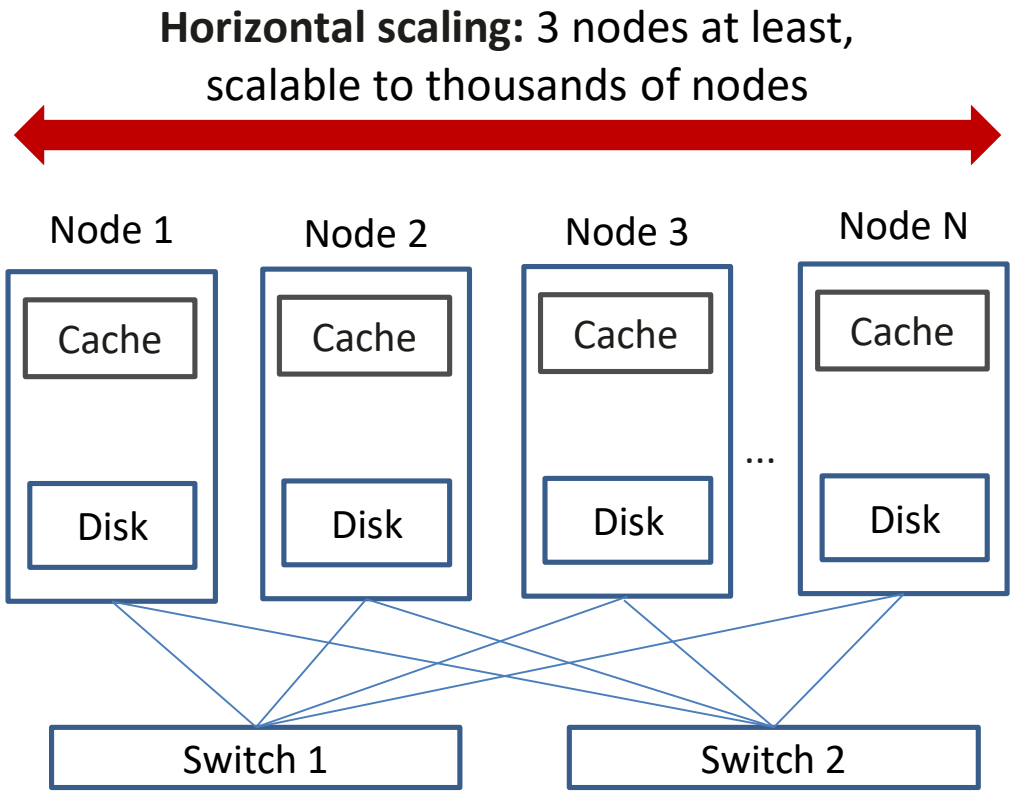


Centralized storage:

- Controller: manages all data. When a controller is faulty, another controller takes over services.
- Disk enclosure: only stores data. The controller determines how to store data.
- The capacity generally ranges from TB to dozens of PB.
- Controllers in a controller enclosure are directly connected through backplane circuits, and controller enclosures are interconnected through networks.

Note: Devices that support multiple controllers can be expanded horizontally. Generally, a maximum of 32 controllers are supported.

Storage System Classification by Architecture – Scale-Out Storage

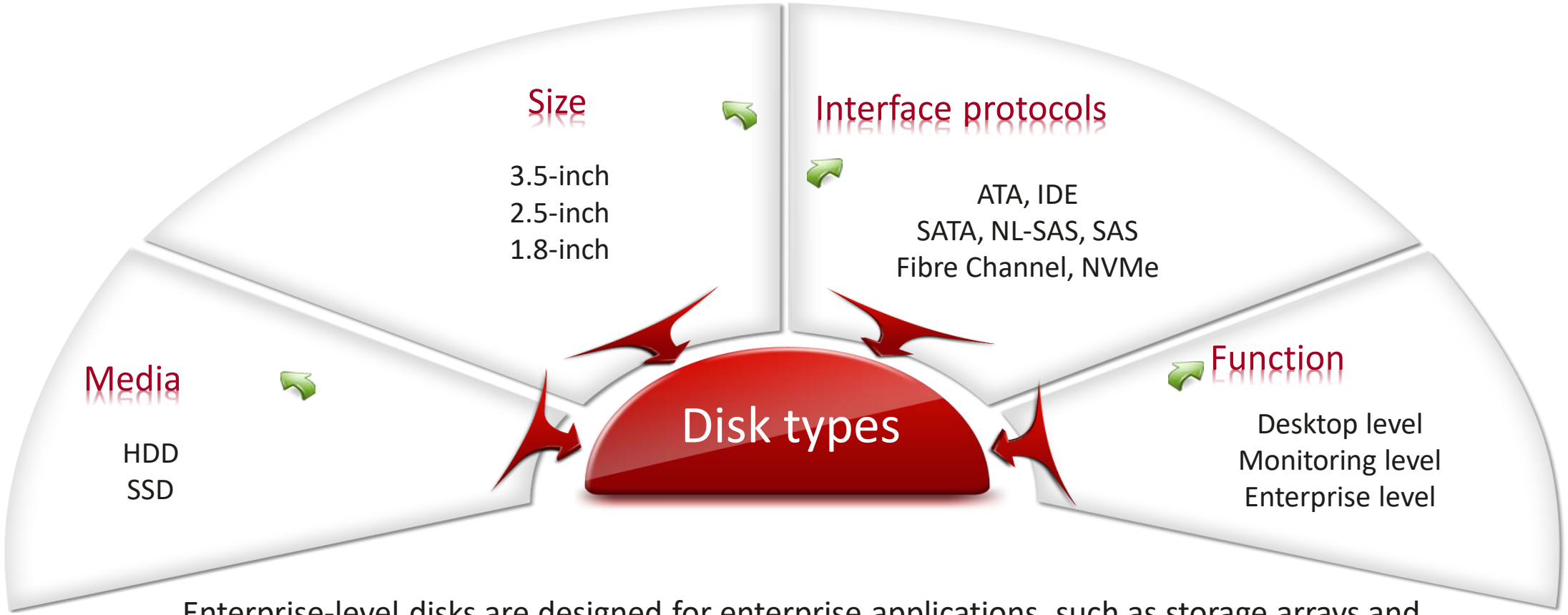


Scale-out storage:

- Each node participates in data management and storage. If one node is faulty, the remaining nodes continue to work.
- The capacity ranges from 100 TB to EB.
- Nodes are interconnected through the network.

Contents

Disk Types

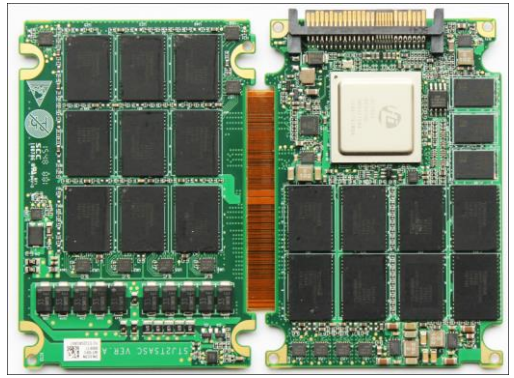


Enterprise-level disks are designed for enterprise applications, such as storage arrays and servers. Desktop-level disks are mainly used for home applications, such as desktop PCs and laptops.

Classification by Media

- **A hard disk drive (HDD)** is a traditional hard disk. It consists of a platter, a head, a platter rotating shaft, a control motor, a head controller, a data converter, interfaces, and cache.

- **A solid-state disk or solid-state drive (SSD)** is also called an electronic disk or a solid-state electronic disk. Different from HDDs that use mechanical parts such as disk bodies and heads, SSDs are composed of control chips and storage chips (flash or DRAM chips). In other words, SSDs are made of solid-state electronic storage chip arrays.



SSDs vs. HDDs

Item	SSD	HDD
Storage media	Flash chip	Magnetic disk
Shockproof and drop resistance	High	Low
Data storage speed	High: xxx, xxx IOPS, hundreds of MB/s	Low: 200 to 300 IOPS, dozens of MB/s
Power consumption	Low	High
Weight	Light	Heavy
Noise	None	Yes
Capacity	xxx GB to 32 TB	x TB
Service life	Long	Relatively long

By Protocol – Serial ATA (SATA)

Advantages:

1. SATA is a serial bus interface protocol. During data transmission, the data and signal cables are used independently, and the embedded clock frequency signal is used. Its rate can reach 30 times of the Parallel ATA (PATA).
2. Instead of simple PATA improvement, the bus structure is brand new.
3. Control information is scattered in data and transmitted by using a predefined bit.
4. One path is used to transmit data, and the other is used to return a response.
5. SATA has higher anti-interference capabilities and faster speed than PATA. The installation is easier and the number of cables used in the chassis is reduced.
6. The performance of single-thread tasks is good.

Disadvantages:

1. SATA, designed for entry-level applications, is not as powerful as SCSI in terms of big data throughput or multi-thread transmission.
2. When multiple threads are reading data, the head of a disk swings back and forth, causing the disk to be overheated.

By Protocol – Serial Attached SCSI (SAS)

1. SCSI is a storage interface protocol specially designed for small computer systems. 50-pin ports are used. The appearance of a 50-pin port is similar to that of a common disk port. The SCSI disks support connections to various devices. In addition, each SCSI disk has an independent chip for data processing. The CPU usage is low, the bandwidth can reach 320 MB/s, and the stability is good.
2. SAS is a disk connection technology that integrates the advantages of the parallel SCSI and serial connection technologies.
3. SAS is a point-to-point, full-duplex, and dual-port interface.
4. SAS is compatible with SATA, delivering high performance for enterprises, achieving interoperability with SATA, and bringing unprecedented flexibility and benefits to enterprises.
5. SAS features high performance, high reliability, and powerful scalability.

The serial technology is also used, which is better than SCSI in transmission rate and anti-interference. However, the price is higher.

By Protocol – Near-Line SAS (NL-SAS)

1. NL-SAS disks integrate SAS interfaces and SATA disks. In other words, NL-SAS disks refer to SATA disks with SAS interfaces and near-SAS performance.
2. Near-line storage is oriented to applications between online storage and offline storage. Data that is not frequently used or seldom accessed is stored on the storage devices of which performance is relatively low. However, these devices must provide fast addressing capabilities and a high transmission rate.

SAS SSD

Size: standard 2.5-inch SAS SSDs accessed based on the SAS protocol

Short access time: The xx µs-level average read/write response time meets the requirements of latency-sensitive service applications and minimizes the response time to customers' requirements.

High throughput: Provides hundreds of thousands of random read/write IOPS, meeting the requirements of high-performance storage system.

High read/write speed: GB-level read/write bandwidth and stable performance.

Low power consumption: The energy consumption per unit capacity is much lower than that of SAS HDDs.

NVMe SSD

Size: 2.5-inch and 1.8-inch NVMe SSDs accessed based on the NVMe protocol
Short access time: The xx µs-level average read/write response time, shorter than that of SAS SSDs, meets the requirements of latency-sensitive service applications and minimizes the response time to customers' requirements.
High throughput: Provides hundreds of thousands of random read/write IOPS, meeting the requirements of high-performance storage system.
High read/write speed: GB-level read/write bandwidth and stable performance.
Low power consumption: The energy consumption per unit capacity is much lower than that of SAS HDDs.

Comparison of Mainstream Disk Types

Indicator	SATA HDD	SAS HDD	NL-SAS HDD	SAS SSD	NVMe SSD
Rotational speed (rpm)	7200	15,000 or 10,000	7200	N/A	N/A
Capacity (TB)	4, 6, 8, 10, or 14	1.2, 1.8, or 2.4	4, 6, 8, 10, or 14	1.92, 3.84, 7.68, 15.36, etc.	1.92, 3.84, 7.68, 15.36, etc.
MTBF (h)	1,200,000	1,600,000	1,200,000	2,000,000	2,000,000
Remarks	Being developed from ATA disks, SATA 2.0 supports 300 MB/s data transfer, and SATA 3.0 supports up to 600 MB/s data transfer. The annual failure rate of SATA disks is about 2%.	SAS disks are designed to meet enterprises' high performance requirements and are compatible with SATA disks. The transmission rate ranges from 3.0 Gbit/s to 6.0 Gbit/s, and will be increased to 12.0 Gbit/s in the future. The annual failure rate of SAS disks is less than 2%.	NL-SAS disks are enterprise-class SATA drives with SAS interfaces. They are applicable to tiered storage in a disk array, which simplifies the design of the disk array. The annual failure rate of NL-SAS disks is about 2%.	An SSD is made up of a solid-state electronic storage chip array. An SSD consists of a control unit and a storage unit (flash and DRAM chips). SSD is the same as the common disks in the regulations and definition of interfaces, functions, usage, as well as the exterior and size. The annual failure rate is about 0.5% to 0.8%.	An SSD is made up of a solid-state electronic storage chip array. An SSD consists of a control unit and a storage unit (flash and DRAM chips). SSD is the same as the common disks in the regulations and definition of interfaces, functions, and usage. The latency is lower than that of SAS SSDs, and the size can be thinner than that of standard SAS SSDs. The annual failure rate is about 0.5% to 0.8%.

SCM – Next-Generation Storage Media

Storage Class Memory (SCM) is a new storage medium popular in the industry. SCM is similar to storage in its persistence and to memory in its byte-level access.

The SCM SSD, which uses NVMe block interface and is compatible with the native architecture, is the primary application form of SCM. Optane P4800X series launched by Intel is an example. This product has little impact on the system architecture but provides better performance than that of flash SSDs. In addition, SCM SSDs do not require garbage collection, which prevents performance deterioration similar to that of NAND SSD after long-time running, and curbs latency at an appropriate level. This type of storage products is springing up in the industry, and Huawei is also working on the R&D of relevant products.

High-performance SSDs provide the following application forms in a storage system:

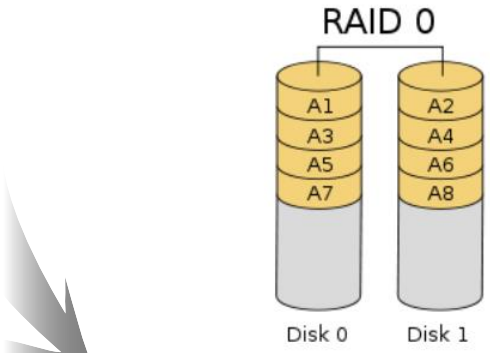
- Metadata cache: As metadata cache of AFA, SCM SSDs work with the DRAM to build a memory + SCM SSD two-tier cache, which avoids bottlenecks in memory capacity and supports larger user capacity with stable performance.
- Data cache: SCM SSDs serve as the acceleration layer of user data and improves performance in typical application scenarios.
- Main storage: SCM SSDs serve as the storage layer of user data and provides a high-performance storage system to meet the performance requirements of some scenarios.

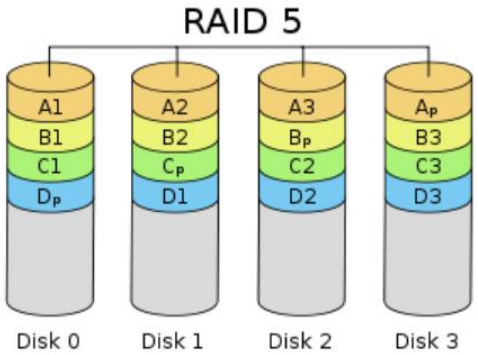
Limited by the price of SCM, its high requirements on the storage architecture, and unclear application scenarios, it is difficult to put SCM into large-scale commercial use in a short term.

Contents

RAID

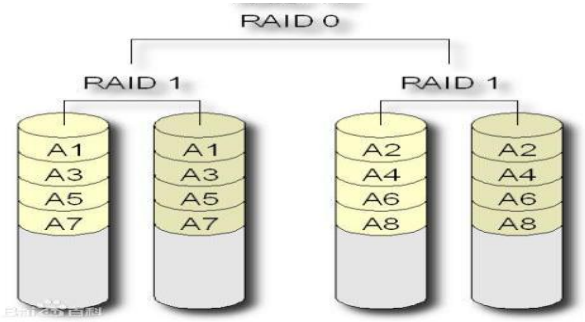
The term RAID was invented by David Patterson, Garth A. Gibson, and Randy Katz at the University of California, Berkeley in 1987. It combines multiple independent physical disks into a virtual logical disk using related algorithms to provide larger capacity, higher performance, and better error tolerance capabilities.



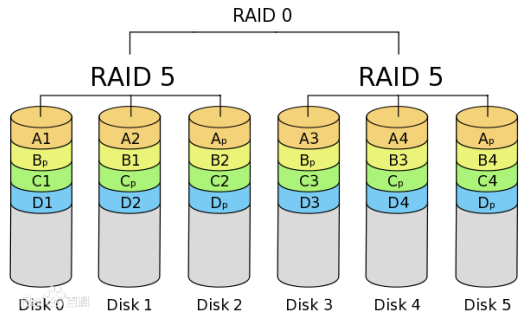


- RAID is classified into different RAID levels based on the combination methods, such as RAID0, RAID1, RAID3, RAID5, and RAID6.
- Two different RAID levels can be combined to form a new RAID level.

RAID10



RAID50



HUAWEI

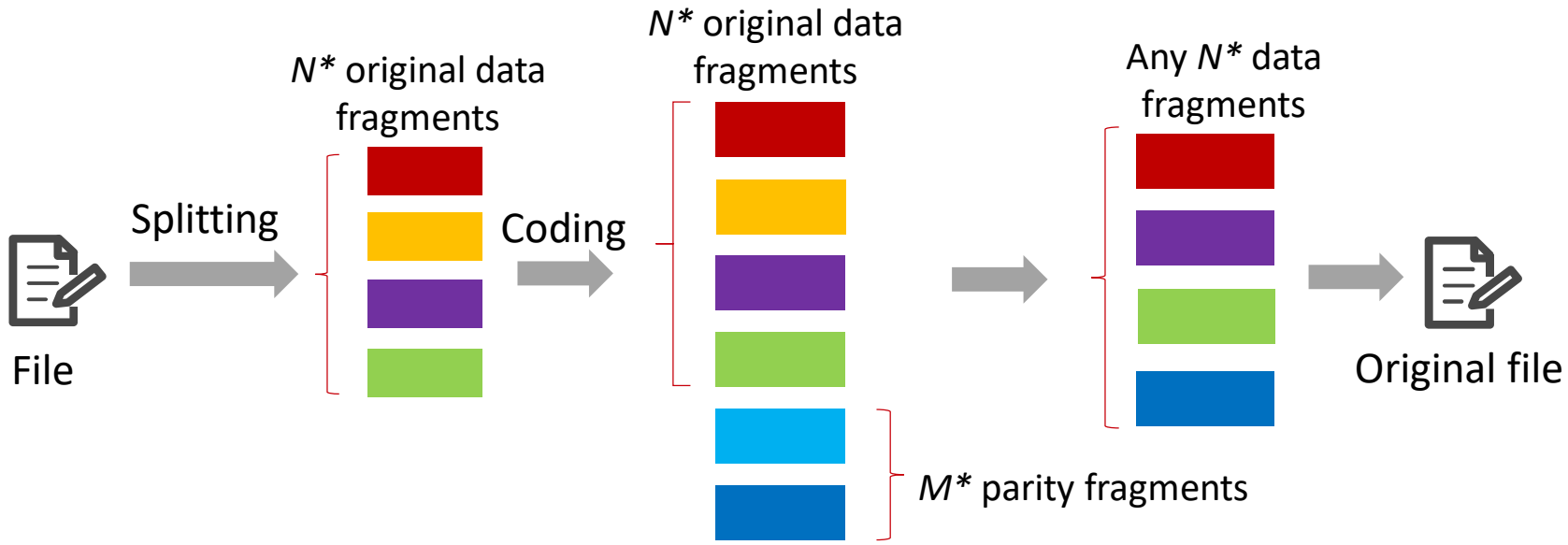
Comparison between Mainstream RAID Levels

Level	RAID0	RAID1	RAID3	RAID5	RAID6	RAID10
Data arrangement characteristics	Striping	Mirroring	Dedicated parity stripe	Distributed parity stripe	Double parity stripe	Mirroring and striping
Fault tolerance	None	High	Medium	Medium	High	High
Hot backup	None	Yes	Yes	Yes	Yes	Yes
Read performance	High	Low	High	High	High	Medium and high
Random write performance	High	Low	Medium	Medium and high	Medium and high	Medium
Sequential write performance	High	Low	Medium	Medium and high	Medium and high	Medium and high
Number of disks	≥ 1	≥ 2	≥ 3	≥ 3	≥ 4	mn ≥ 4 (m indicates the number of RAID 0 groups.)
Available capacity	All	1/n	(n-1)/n	(n-1)/n	(n-2)/n	1/m

Erasure coding (EC)

Erasure coding (EC) is a data redundancy protection mechanism widely used in scale-out storage. Data written into the scale-out storage system is **divided into N data fragments, and M parity fragments are generated for the N data fragments using EC**. If M fragments are damaged in an EC group, the system implements data recovery from the N fragments.

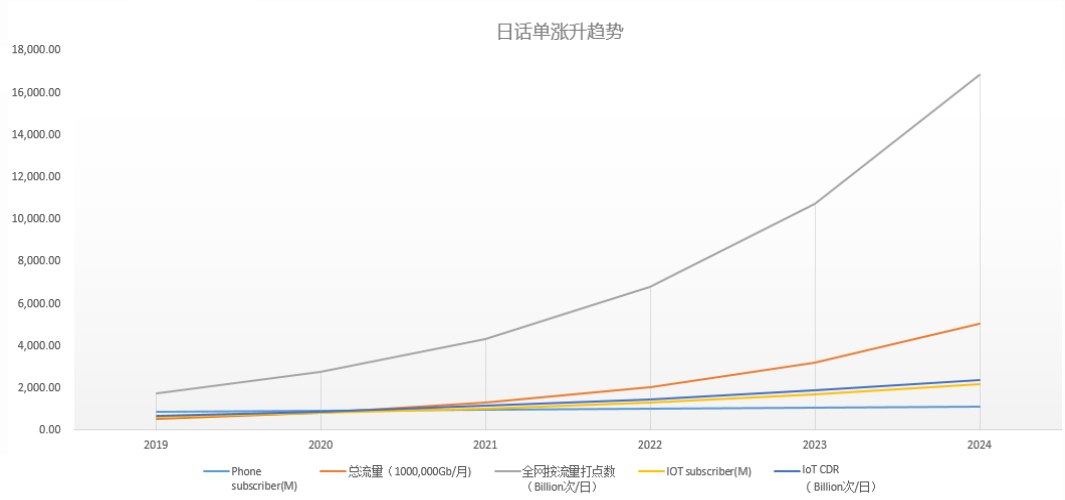
Compared with the multi-copy mechanism, EC improves both disk utilization and storage reliability, thereby cutting costs. If N+M EC protection is used, the utilization rate of storage space is N/(N+M).



Contents

Stable Services, Represented by Production Transactions, Require Powerful and Reliable Storage

Carrier: Core billing systems carry a heavier load in the 5G era

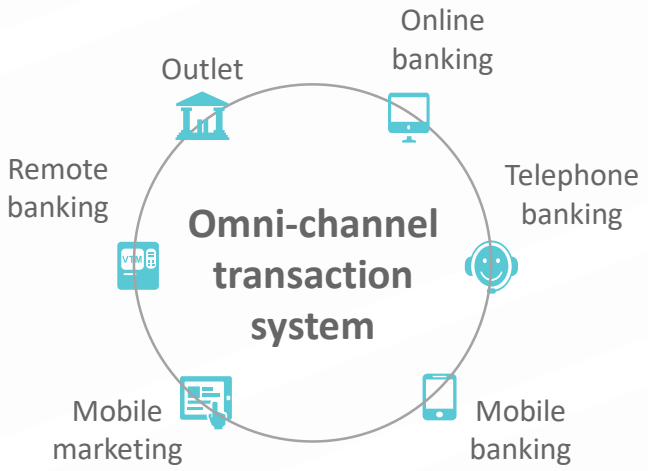


Mobile phone | Network-wide traffic | Mobile phone CDR | IoT devices | IoT CDR

User device growth: **3x**

CDR growth: **10x**

Finance: Transaction channel diversity brings sharp transaction volume increase



[The People's Bank of China Report] Q3, 2019, banks

- Total number of accounts: 11.017 billion with **11.6%** YoY **growth**
- Banks processed 59.464 billion mobile payments with **31.5%** YoY **growth**

Emerging micropayment transactions increase **over 10x**

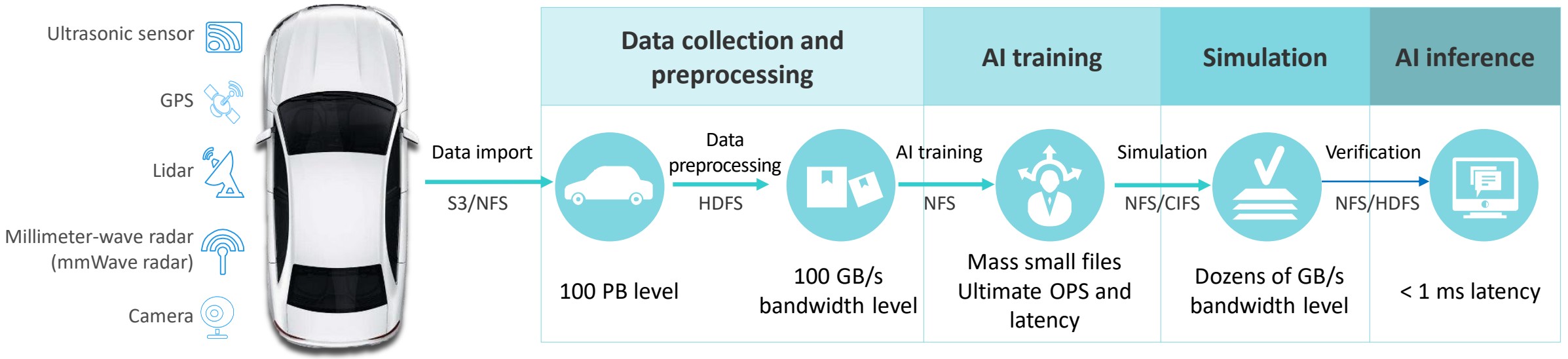
24/7 service availability

Latency **5 ms -> 0.5 ms**

99.999% -> 99.9999% Reliability

HUAWEI

Agile Services, Represented by HPDA, Demand Efficient, Large-Capacity Storage



L3 to L4

50x more drive test mileage generates 50x more data.

Each vehicle generates 64 TB of data daily.

100 PB data level **requires more storage space.**

S3, NFS, and HDFS interfaces

Massive data copy **demands efficient analysis.**

Requires high bandwidth and high OPS at once.

Ultra-low latency **requires high performance.**

HUAWEI

Intelligent Management: Rapid Data Growth and Various Devices Complicate Management

Digital transformation faces increasing challenges, including diversified applications, media, and devices.

Enterprise storage (structured data)

CRM

ERP

Billing

Manufacturing

New mass data storage (unstructured data)

EDA

Container

HPC

Big data

Video

AI

Stream processing

Media

NVM /SCM

Blu-ray Disc

Network

CXL

CCIX

OpenCAPI

NVMe over Fabrics

GEN Z

Architecture

Centralized architecture

Distributed architecture

Composable architecture

Hyper-converged architecture

Data volume managed per capita increases by several times.

Manpower Required

1 PB

5 PB

Monthly deployment time required: 15 days

Monthly deployment time required: 3 days

Data Volume

O&M management complexity increases.

Planning	Building	O&M	Optimization
Survey and analysis 2 months	5 persons, 2 nights Manual deployment	Segment-by-segment check 2 hours	Labor-consuming 3 hours

HUAWEI

Data Security: With Digital Transformation in Full Swing, Data Protection Becomes More Important than Ever

Natural disasters

Fire in an ISP data center in Europe

>3.6 million websites down
Huge amounts of customer data unable to be restored

Datacom (Australia) hit by a flood

A large number of devices damaged, and user data permanently lost

Software and hardware failures

Failure of core databases at a bank

A large number of services interrupted for 37 hours

UPS overload at an IDC service provider

Power failure in the equipment room and system breakdown
Services of multiple financial institutions interrupted for 7 hours

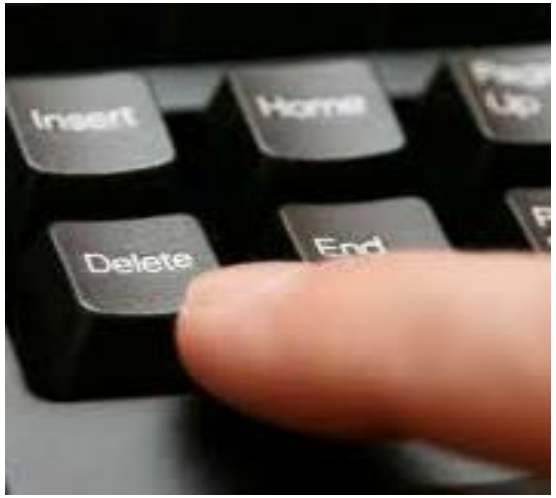
Virus/Unauthorized access

Ransomware in 2020

> 150% increase in ransomware attacks
18 days of downtime on average caused by an attack

Hacker attack on a bank

Personal information of over 100 million users leaked
Plummeting stock price and potential legal claims



Human errors

Database of a listed company deleted

Market value slashed by more than HK$3 billion
CNY 120 million in compensation paid to its clients

Code deleted by mistake

Service downtime: 12 hours
Direct loss: ~CNY 80 million

HUAWEI

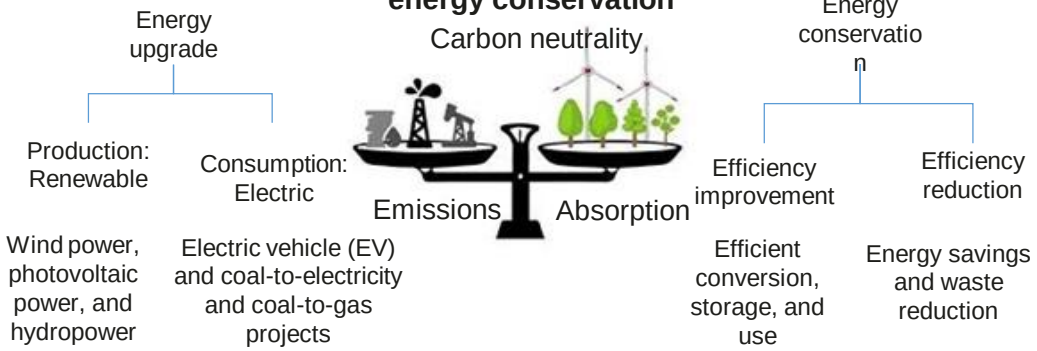
Green Storage Is a Must to Achieve Its Carbon Peak and Carbon Neutrality Goals

New energy methods to meet carbon neutral targets

- Carbon neutral targets of major economies

China	EU	US	Japan
Carbon peak by 2030 Carbon neutrality by 2060	Delivers the Green Deal Carbon neutrality by 2050	Rejoined the Paris Agreement Carbon neutrality by 2050	Announced the Green Growth Strategy Carbon neutrality by 2050

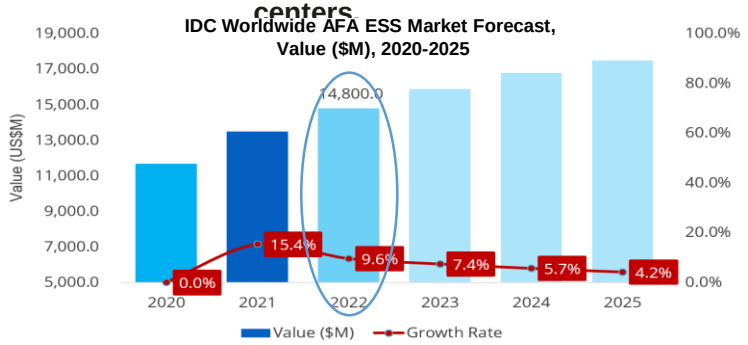
- **The key to carbon neutrality: New energy methods in production and consumption areas, efficient energy use, and energy conservation**



All-flash data centers to save energy and reduce emissions

- **All-flash storage devices are becoming more ubiquitous in data centers**

According to projections, the all-flash market will grow by 9.6% to $14.8 billion by 2022.



- **Increased storage power consumption makes all-flash devices essential to achieving energy goals.**

According to IDC, China's data center storage power consumption will grow rapidly in the next five years.

21% CAGR in power consumption (in current mode)

Example: DC with 10 PB capacity

Status quo		After all-flash upgrade
140,000 kWh/year HDDs 20%: 2.4 TB SAS 80%: 10 TB SATA	**2.4x lower**	**60,000 kWh**/year SSDs 100%: 7.68 TB SSD

Equal to the amount of CO_2 absorbed by **16,000** trees a year

Report from Huadian Group:
Carbon emissions from **1 kWh electricity** = 0.96 kg of CO_2 = The amount absorbed by **0.2 trees** a year

HUAWEI

Thank you.

把数字世界带入每个人、每个家庭、
每个组织，构建万物互联的智能世界。

Bring digital to every person, home, and organization for a fully connected, intelligent world.