Recommendations

- Huawei Learning Website
  - http://learning.huawei.com/en
- Huawei e-Learning
  - http://support.huawei.com/learning/NavigationAction!createNavi?navId=MW000001_term1000025144&lang=en
- Huawei Certification
  - http://support.huawei.com/learning/NavigationAction!createNavi?navId=_31&lang=en
- Find Training
  - http://support.huawei.com/learning/NavigationAction!createNavi?navId=_trainingsearch&lang=en

More Information

- Huawei learning APP

Huawei Training & Certification

版权所有© 2018 华为技术有限公司

01 Latest Storage Technologies and Trends

www.huawei.com

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

Huawei Training & Certification
Huawei Training & Certification

Foreword

- This module briefly introduces the definition of storage, its development history and technological trends in storage. This module allows beginners to understand the how storage comes into existence and also know the direction of technological improvements in storage solutions.
- This module also briefly introduces the Huawei Storage Products and Solutions.

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

Huawei Training & Certification
Huawei Training & Certification

Objectives

- On completion of this course, you will be able to:
  - Understand the definition of storage;
  - Understand the development history of storage technologies;
  - Understand the evolution of storage technologies and its development trends;
  - Learn about Huawei Storage Products and Solutions.

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

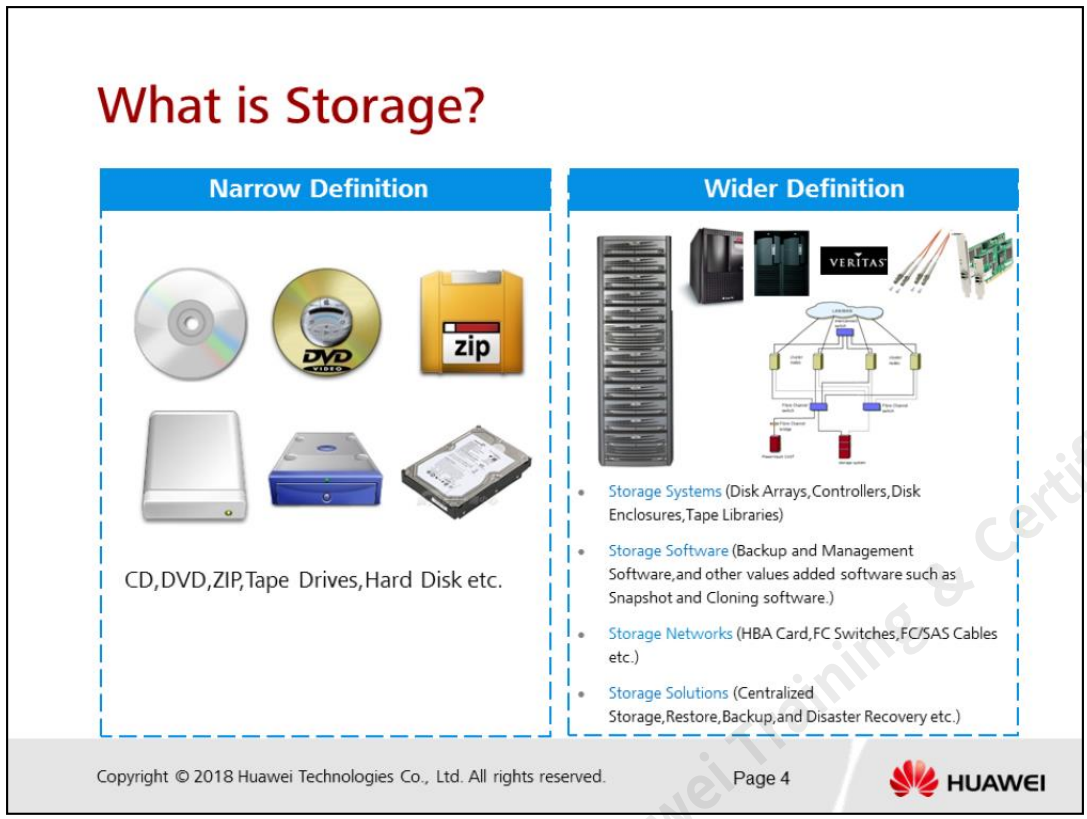
Huawei Training & Certification
Huawei Training & Certification

Contents

1. **Definition of Storage**
2. Development History of Storage
3. Evolution of Storage Technologies
4. Latest Storage Technologies and Trends
5. Storage Products and Solutions

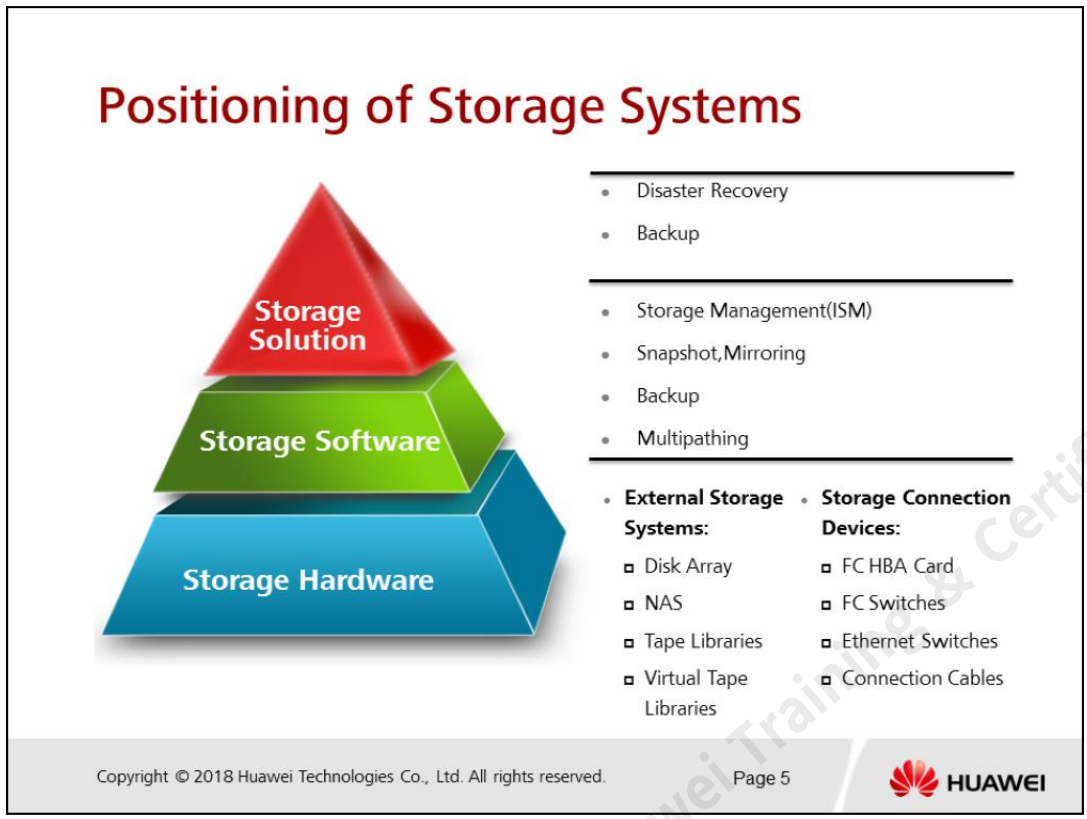
Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

Huawei Training & Certification
Huawei Training & Certification



- A narrow definition of storage can refer to the physical storage devices such as floppy disks, CD/DVD, hard drives or even tape drives that are still used in some enterprises. A wider definition of storage refers to the storage devices used in Datacenters such as storage systems, storage software, storage networks and storage solutions.
- In reality, multiple storage systems, hardware and software work together to form a solution that could satisfy the needs of business that requires high levels of data management such as data integration, backups and disaster recovery solutions. Etc. Servers accesses the data on storage hardware via a storage network and storage software manages the data for data integration, backup and disaster recovery purposes.
- This module shall introduce the wider definition of complex storage systems used for storing and managing critical business data.

Huawei Training & Certification



- Current storage technologies are no longer standalone systems, in reality, it consist of a series of components that makes it a complete storage system.
- Nowadays, storage systems are divided into 3 different components which are the storage hardware, storage software and storage solution. The storage hardware component is further divided into external storage hardware such as physical storage devices like disk arrays and tape libraries.
- Usability of storage systems are greatly increased due to the existence of storage software, features such as data mirroring, cloning, automatic backups and other data operational tasks can be managed and completed via storage software.
- A well designed storage solution ensures that data storage operations are conducted easier, however an excellently designed storage solution not only will allow easier storage system deployment, it also lowers the total cost of ownership (TCO) for the customer and ensures that the customer has the best value and protection on their investments.

Huawei Training & Certification

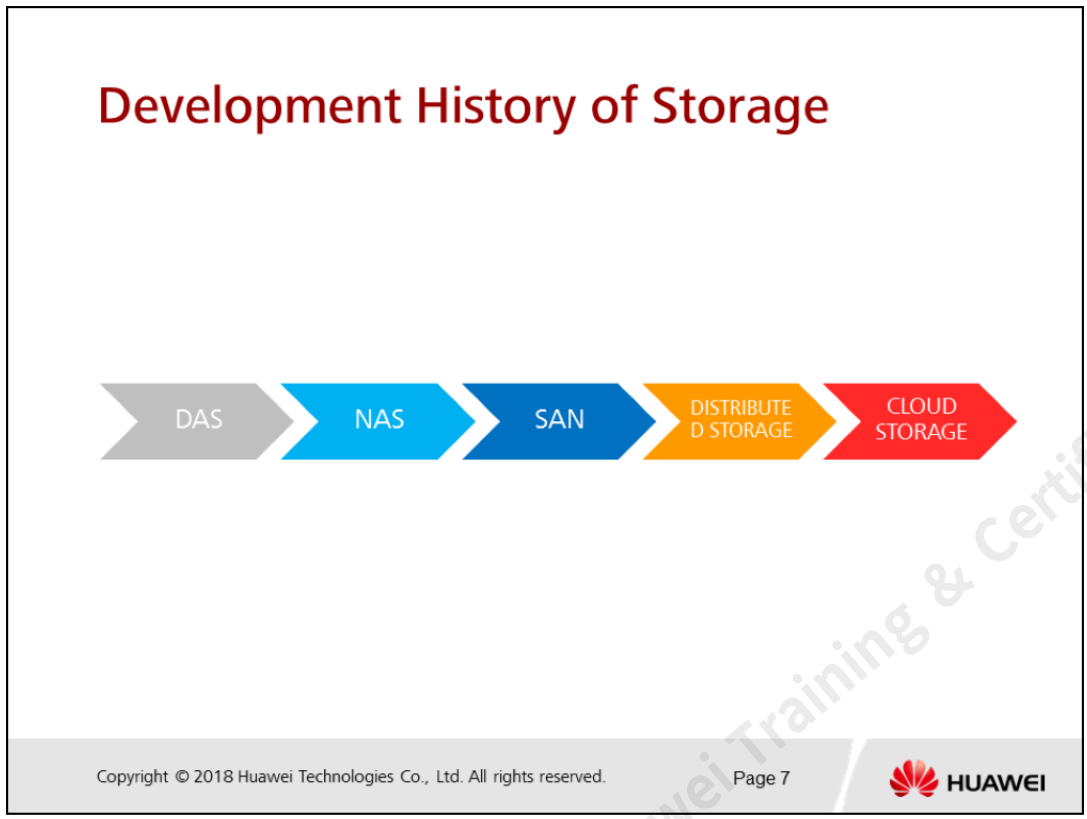
Contents

1. Definition of Storage
2. **Development History of Storage**
3. Evolution of Storage Technologies
4. Latest Storage Technologies and Trends
5. Storage Products and Solutions

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

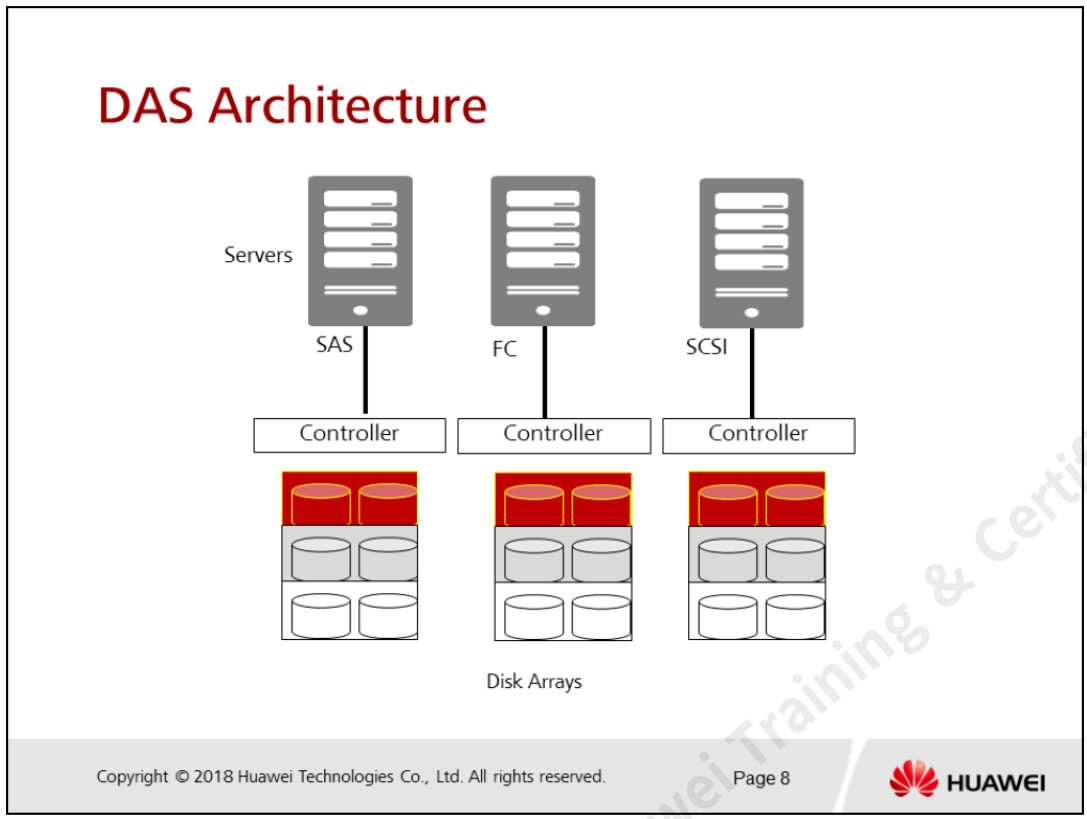
HUAWEI

Huawei Training & Certification
Huawei Training & Certification



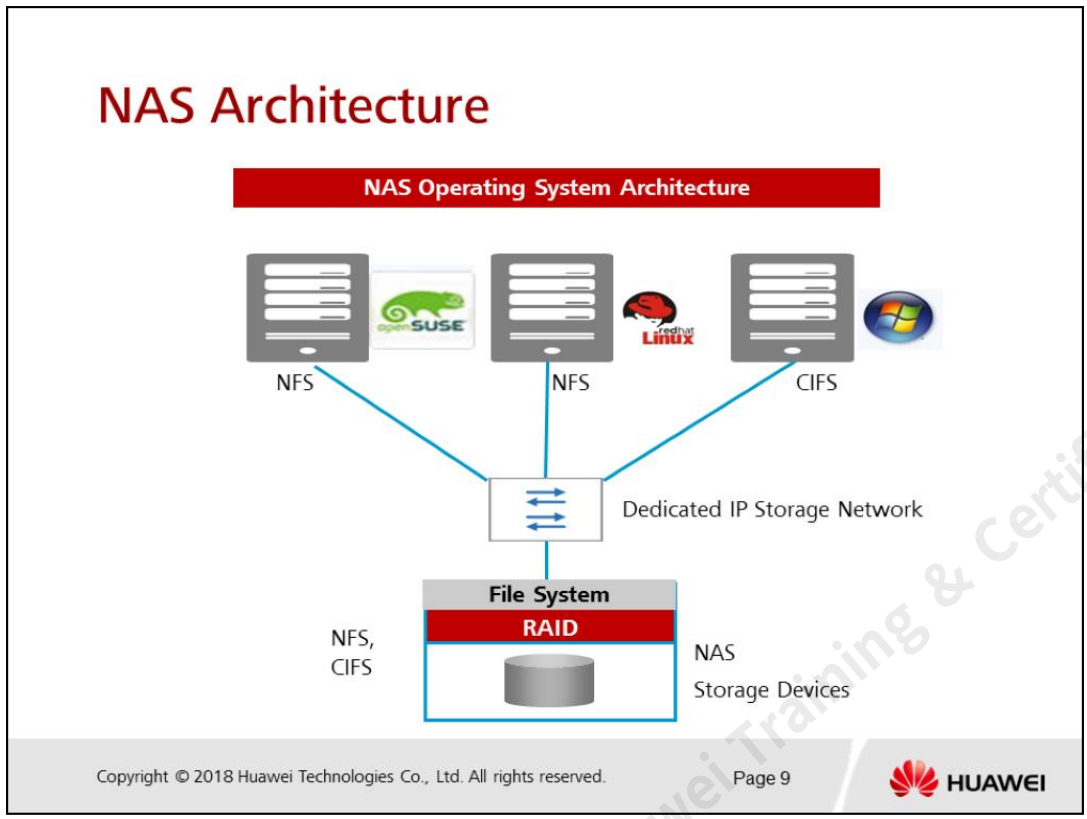
- Direct Attached Storage (DAS).
- Network Attached Storage (NAS).
- Storage Area Network (SAN).

Huawei Training & Certification
Huawei Training & Certification



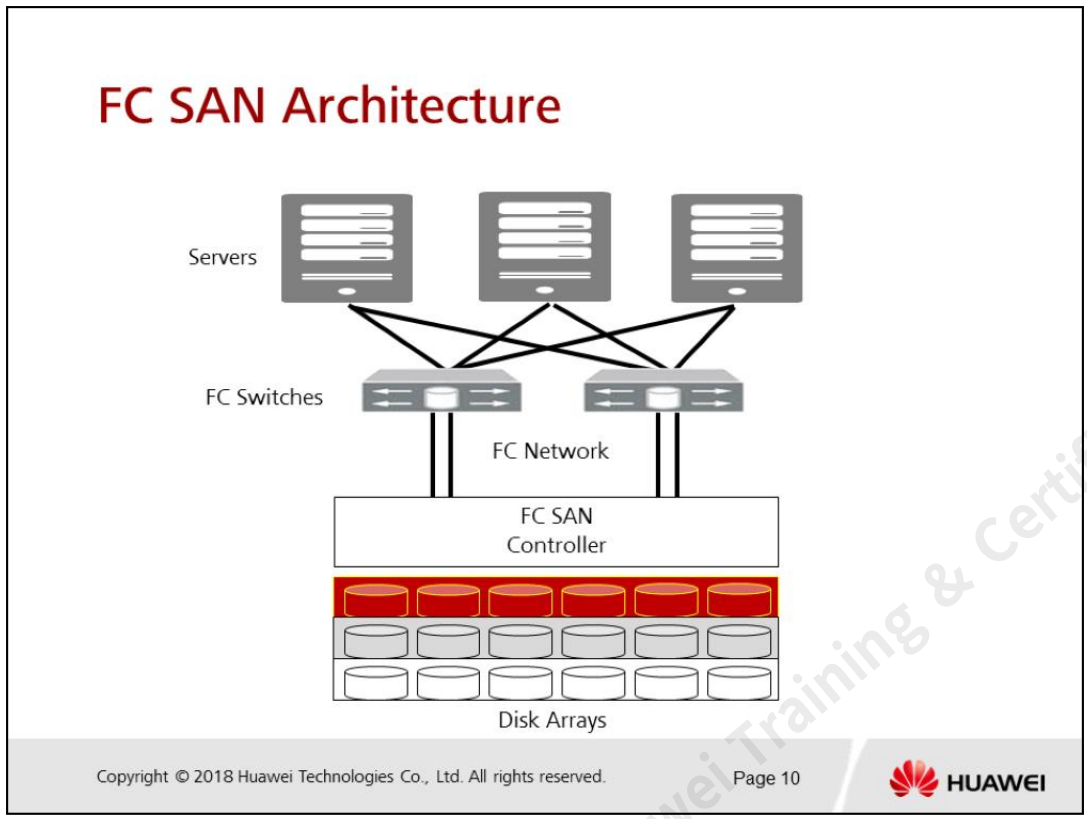
- Direct Attached Storage (DAS) : is a storage architecture that consist of storage devices are directly connected to the servers. DAS provides block-based storage for servers (Not File-based storage).
- Examples of DAS storage: Internal hard disk in servers, tape libraries that are directly connected to servers, and external hard drives that can be directly connected to servers. DAS can be differentiated into Internal DAS and External DAS based on the location of the storage in relation to the servers.
- Internal DAS:
  - In the architecture of the Internal DAS, storage devices are internally connected to the servers via serial or parallel connection to the internal bus, however, due to the physical limitation of the cables length, it only supports short distance high speed data transfers. Additionally, there are also limitation the number of devices that can be connected to the internal bus, and storage devices will also take up lots of space in the servers if they are placed within the servers, making maintenance of other parts within the servers difficult.
- External DAS:
  - In the architecture of External DAS, servers are directly connected to external storage devices. In most situations, they communicate to each other through FC or SCSI protocols. In comparison to internal DAS, external DAS has overcome the limitation of short distance and devices limit faced by internal DAS. Additionally, external DAS can even provide centralized management of the storage devices, making management of those storage devices easier.

Huawei Training & Certification



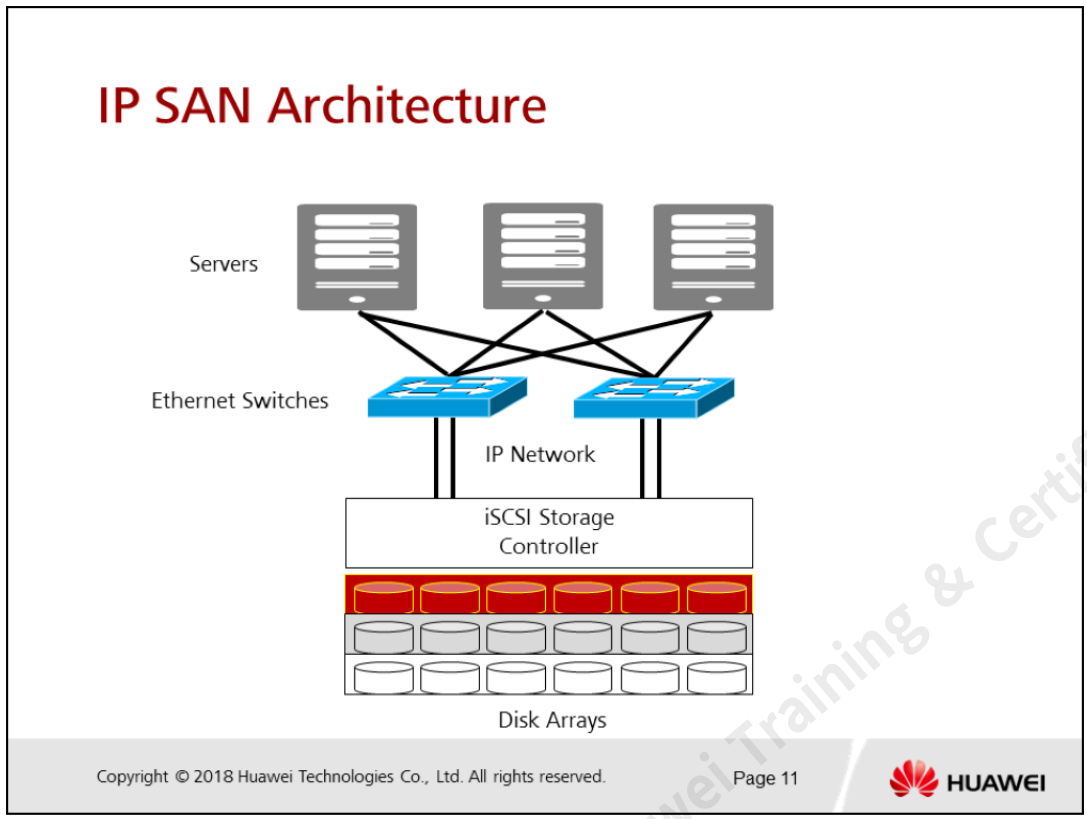
- Network Attached Storage (NAS) are IP-based file level storage devices that are connected to a local area network. NAS allows customers to share files quickly with lower storage management cost through file-based data access and sharing storage resources over a network. By utilizing NAS, there is no need to set up multiple file servers and it is one of the preferred file-sharing storage solutions. NAS also eliminates the issue of performance bottleneck faced when customers are accessing the file servers. NAS uses the network and file sharing protocols to implement data archiving and data storage. Among the protocols used by NAS also includes the TCP/IP protocols used for data transmission and CIFS and NFS protocols for remote file management.
- UNIX and Windows user can share data seamlessly through NAS and usually can do so via the sharing methods such as NAS and FTP. When NAS is shared, UNIX usually uses NFS protocol for file management while Windows uses CIFS protocol. With the advancement of network technologies, NAS has expanded to fulfill enterprises needs for high performance and high reliability data access. NAS devices are usually dedicated, high performance, and high speed devices with a single purpose of providing file service and storage. NAS devices uses their own operating system and integrated hardware and software to fulfill specific file service requirements. NAS optimizes its underlying operating system and file I/O (input/output) processes, making it perform file I/O better than common file servers. NAS devices can connect to more clients compared to traditional file servers which allows it to achieve the purpose of aggregating multiple traditional file servers with less NAS devices.

Huawei Training & Certification



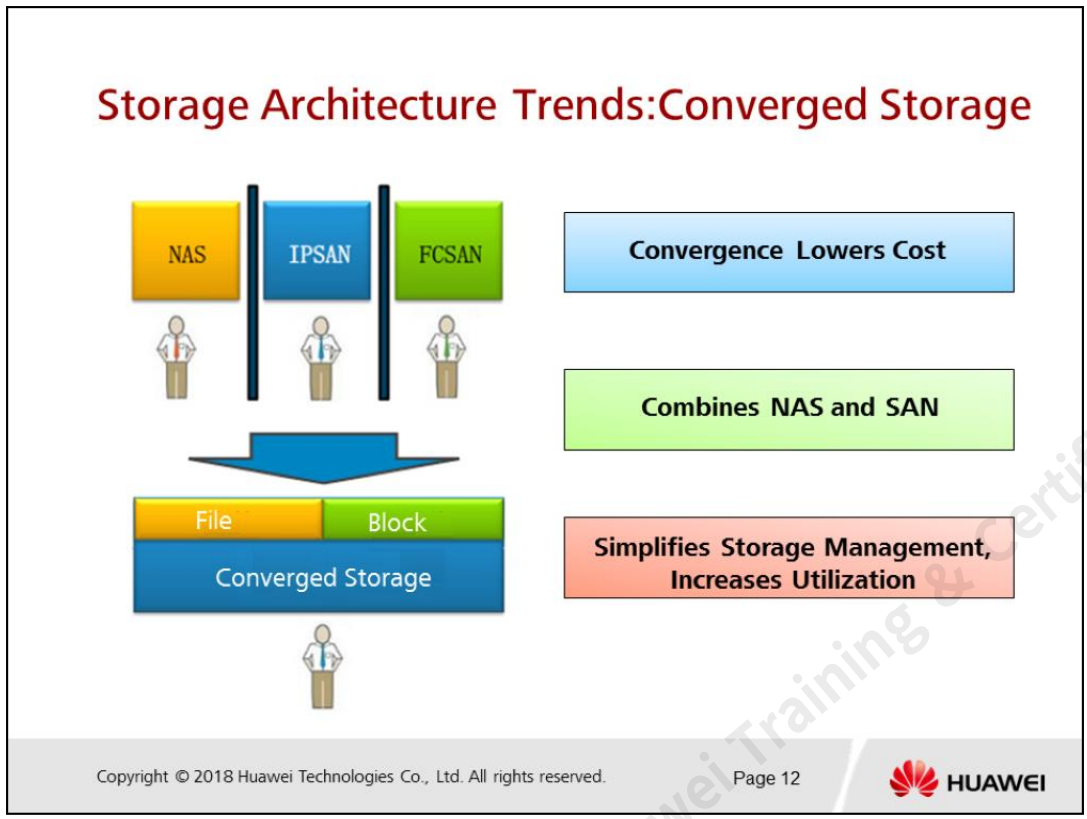
- Storage Area Network (SAN) : is a dedicated high performance network implemented between the server and storage resources. It has been optimized specifically for high amount of raw data transmission. Hence, FC SAN can be seen as an expanded form of SCSI protocols in terms of long distance application. The default protocols used by FC SAN is SCSI and Fiber Channel. Fiber Channel is especially suitable for SAN implementation due to the fact that it is capable of transmitting large blocks of data and is also capable of long distance transmission.
- The market for FC SAN is mostly centered at high end enterprise grade storage devices implementations. These storage implementations has high requirements for performance, redundancy and availability. Etc. storage arrays and backup devices.

Huawei Training & Certification



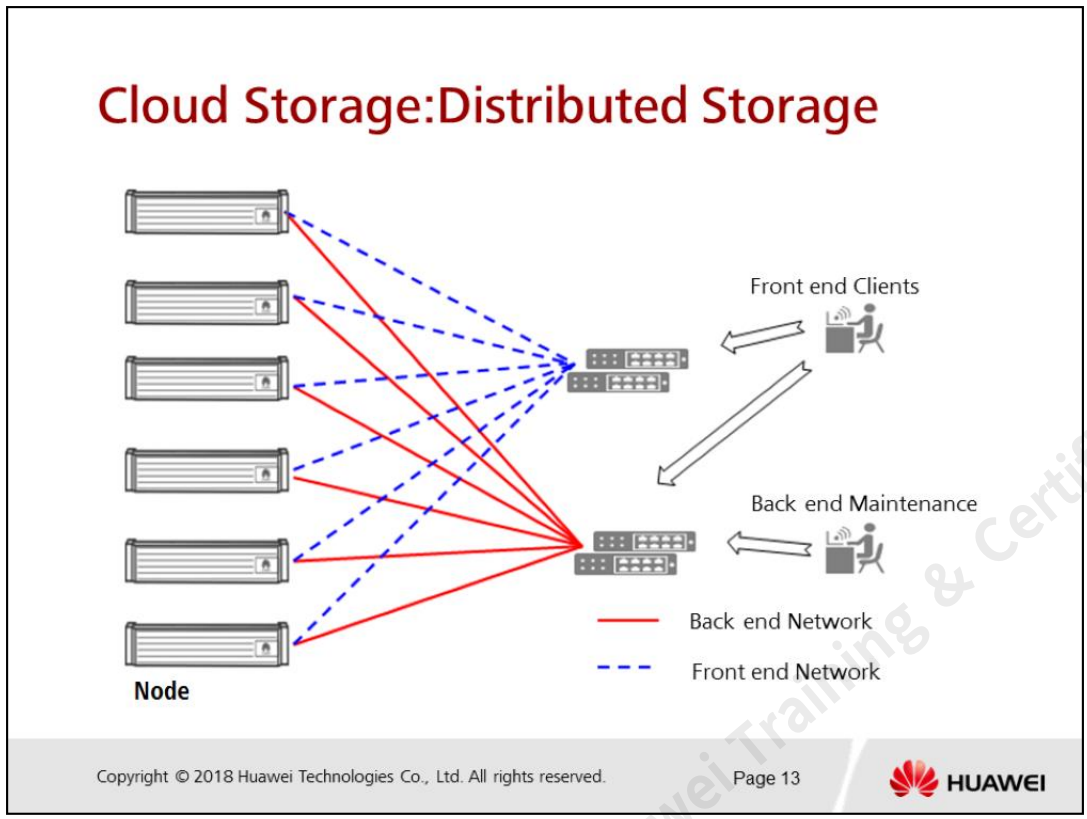
- IP SAN (Storage Area Network) : is a storage area network architecture based on TCP/IP protocols for data transmission using Ethernet as the medium.
- The default protocol for IP SAN implementation is iSCSI (Internet Small Computer System Interface), which is a protocol that encapsulates SCSI commands and transmit it over IP networks.

Huawei Training & Certification



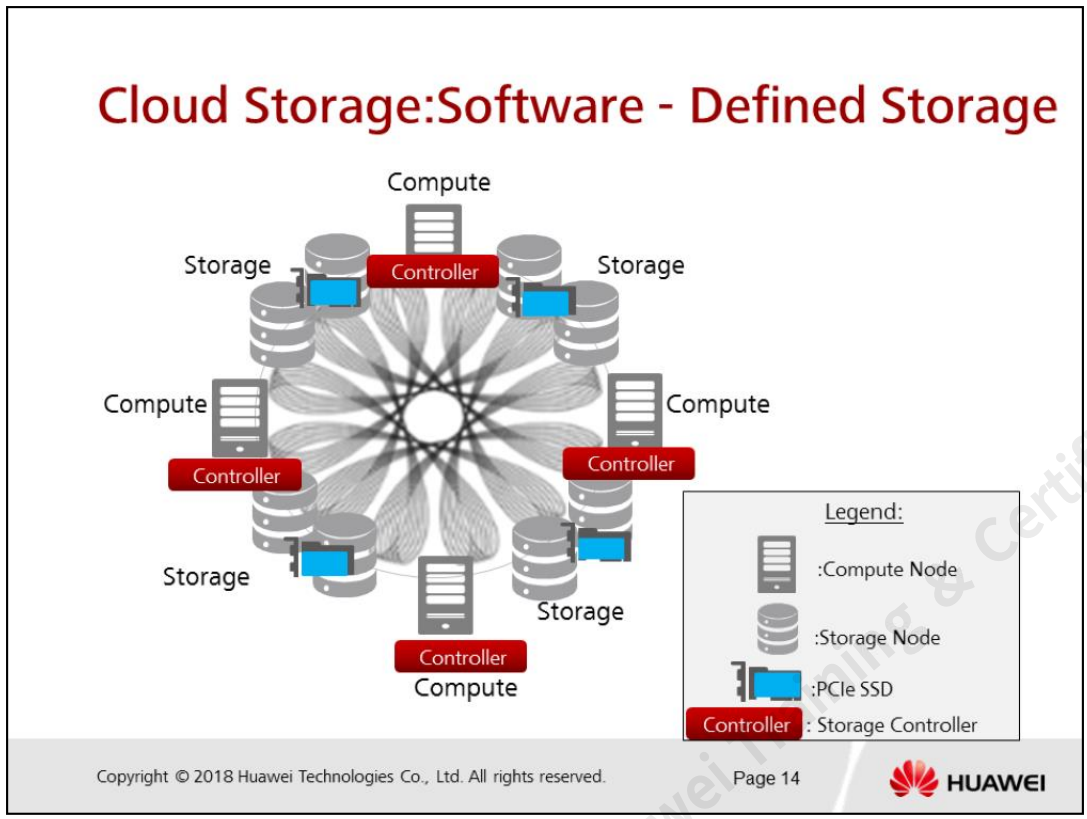
- Converged Storage combines both SAN and NAS storage, fulfills elastic business development, simplifies service deployment, increases storage resources utilization and effectively reduces total cost of ownership (TCO).

Huawei Training & Certification



- Distributed Storage in simpler terms can be defined as storing data in a distributed way across multiple data storage servers on a large scale storage resource pool built using standard x86 server's storage media such as internal HDD and SSD.
- Current Distributed storage systems are mostly built upon Google's experiences in building a distributed file system across multiple servers, and implementing data storage services on this distributed file system.

Huawei Training & Certification



- Software-defined storage uses distributed technology to build a large scale storage resource pool using storage media of X86 servers such as HDD and SSD, and provides industry standard SCSI and iSCSI ports to non-virtualized applications and virtual machines.

Huawei Training & Certification

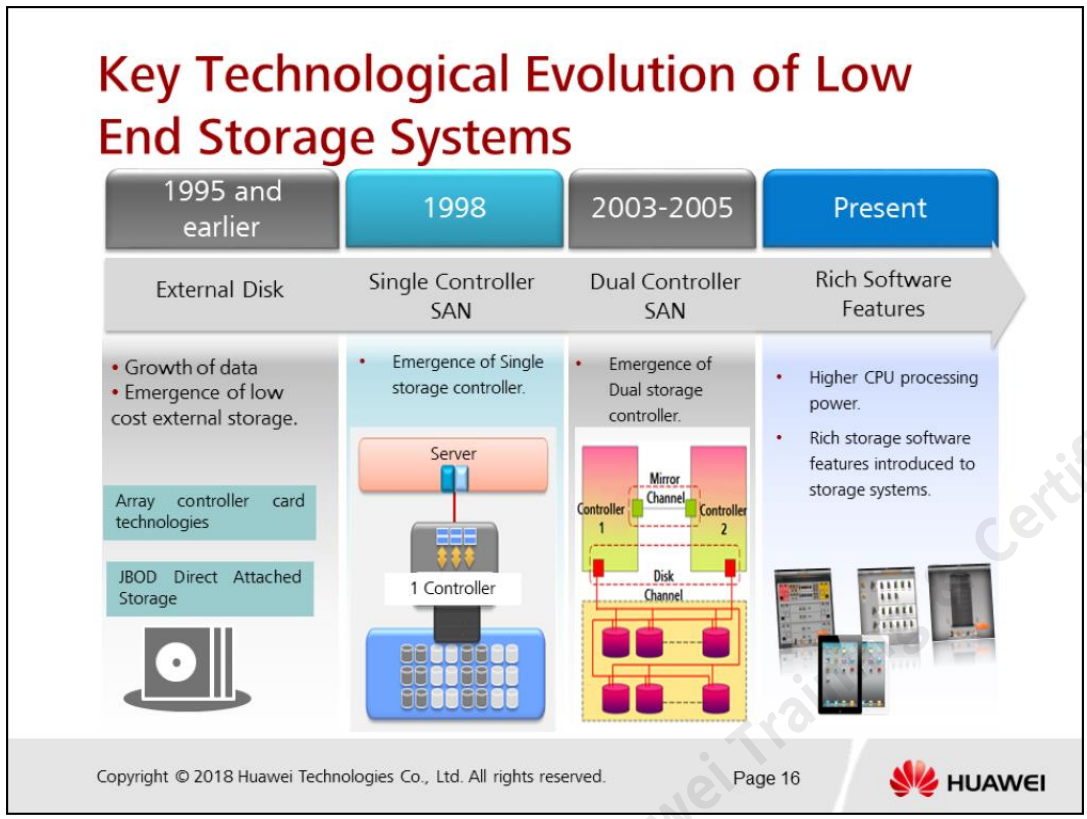
Contents

1. Definition of Storage
2. Development History of Storage
3. **Evolution of Storage Technologies**
4. Latest Storage Technologies and Trends
5. Storage Products and Solutions

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

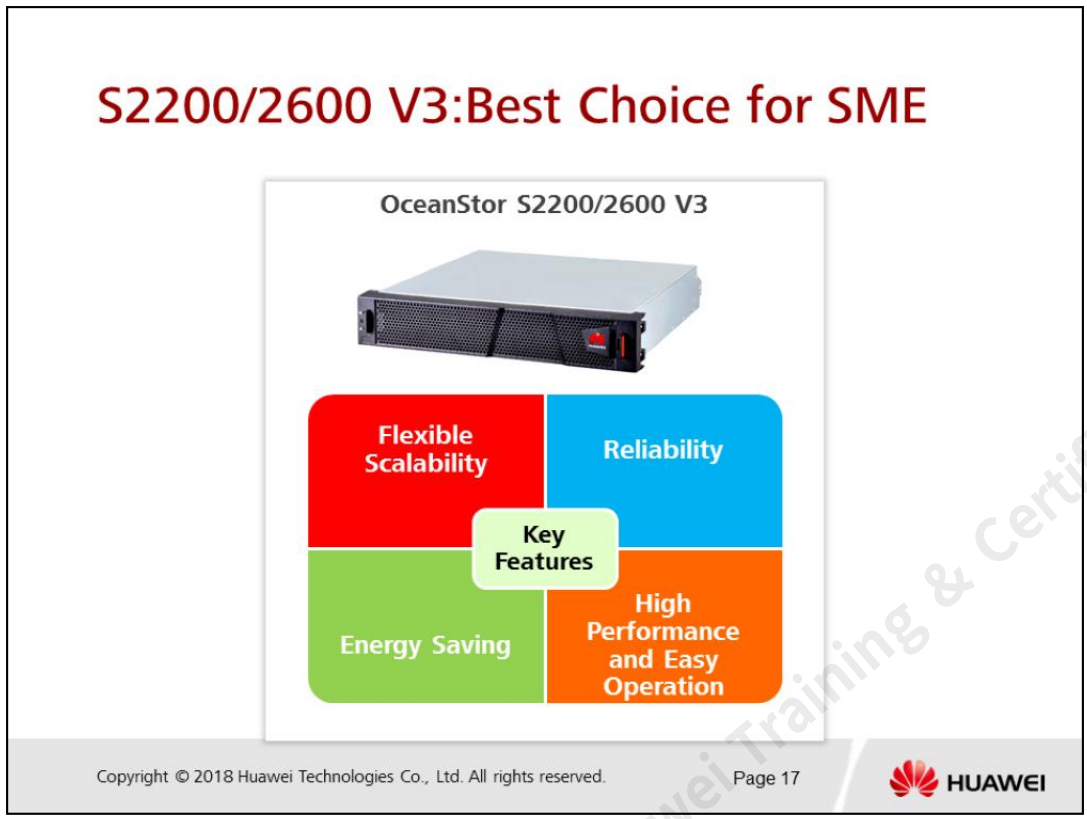
HUAWEI

Huawei Training & Certification
Huawei Training & Certification



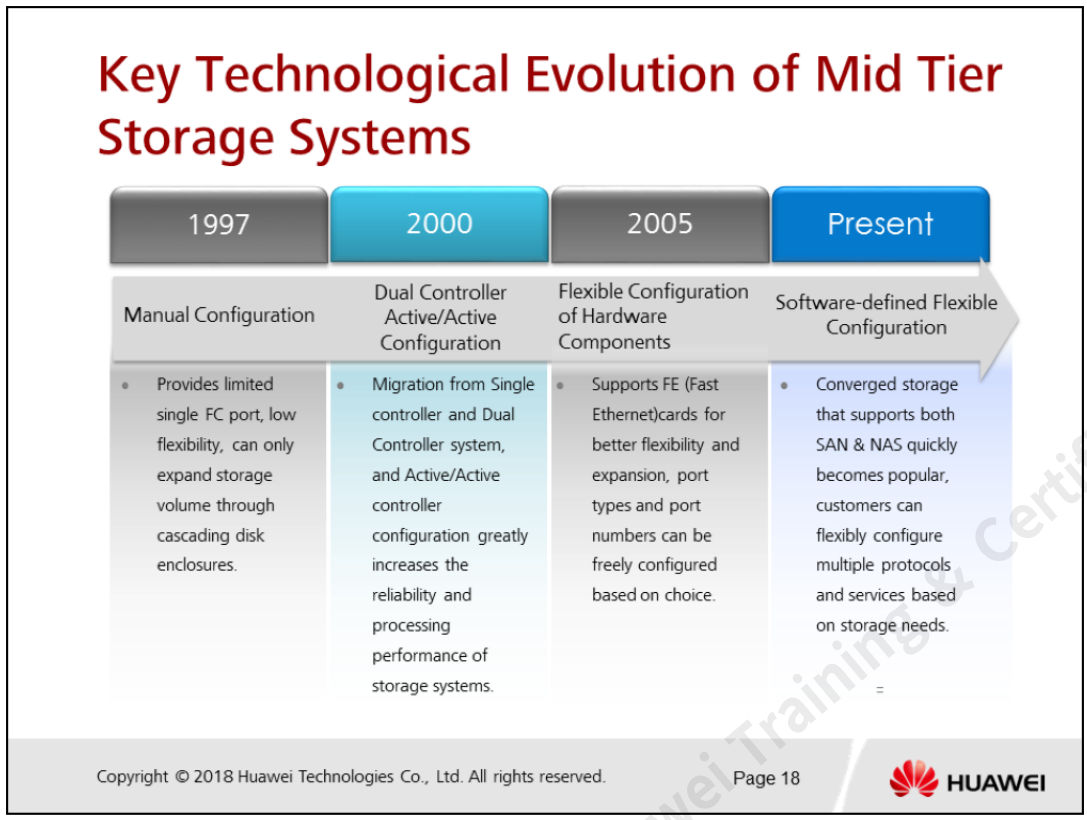
- Low end storage devices has characteristics such as simple, economical and easy to use, it is capable of running extendedly for 24x7 period, and it satisfies enterprises requirement for availability and data protection. It is considered as an entry level professional storage platform for enterprise datacenters.
- As of year 1995 and earlier, with the growth of data, local server storage is no longer able to fulfill enterprise requirements. Emergence of low cost external storage devices causes servers to be connected to multiple hard drives. Array controller card technologies was introduced which allows storage arrays to have more disks connected and managed. JBOD(Just a Bunch Of Disks) direct attached storage was also introduced in this period which increases the limit of storage hard disks that can be connected to servers.
- As of year 1998, single storage controllers was introduced into storage systems. Single Controller in SAN runs the RAID, Cache and Utility Software, which frees up the servers processing power.
- During the period of year 2003 to 2005, Single Controller SAN is slowly replaced by Dual Controller SAN due to decreasing cost, existence of single point of failure, low reliability, and poor scalability. Dual Controller SAN consist of 2 storage controllers which eliminates the single point of failure, has more reliability and much more scalable compared to single controller SAN.
- In the present, with the increase in CPU processing capabilities, low end storage devices can provide more and more service and features. This caused software features to move down from mid range to low range storage devices. Previously, only Mid tier storage system has the high CPU processing power to run multiple storage features, but as the low range storage systems is having more CPU power, these systems are also capable of running these features to a certain extent.

Huawei Training & Certification Huawei Training & Certification



- Flexible and Reliable:
  - Flexible scalability based on needs: Start small and scale up flexibly along with the growth of the enterprises and their needs for storage.
  - Fully Redundant System Design: Full redundancy with dual controllers architecture, ensuring the highest amount of system stability and data protection.
  - Easy operation: Unified management interface using the DeviceManager, 5-steps basic configurations, supports remote management using IPad, and one-click system upgrade for easier operation of your storage.
- Energy Saving and High Performance:
  - High performance and High Standard: Supports both FC and iSCSI , and satisfies complex storage requirements.
  - Multiple Energy Saving features and lowers the Total Cost of Ownership(TCO): Automatically hibernates the hard disk based on service loads and lowers the energy consumption by 40%, 16-speed smart fan speed control technology, and smart CPU frequency adjustment based on service workloads.

Huawei Training & Certification



- Mid tier storage systems has the key characteristics of convergence, flexibility, and simple system architecture which consists of multiple features. It satisfies the ever changing storage needs and requirement of customers, ensures 24*7 availability for enterprise systems, and it is a commonly used storage platform for enterprise datacenters.

Huawei Training & Certification

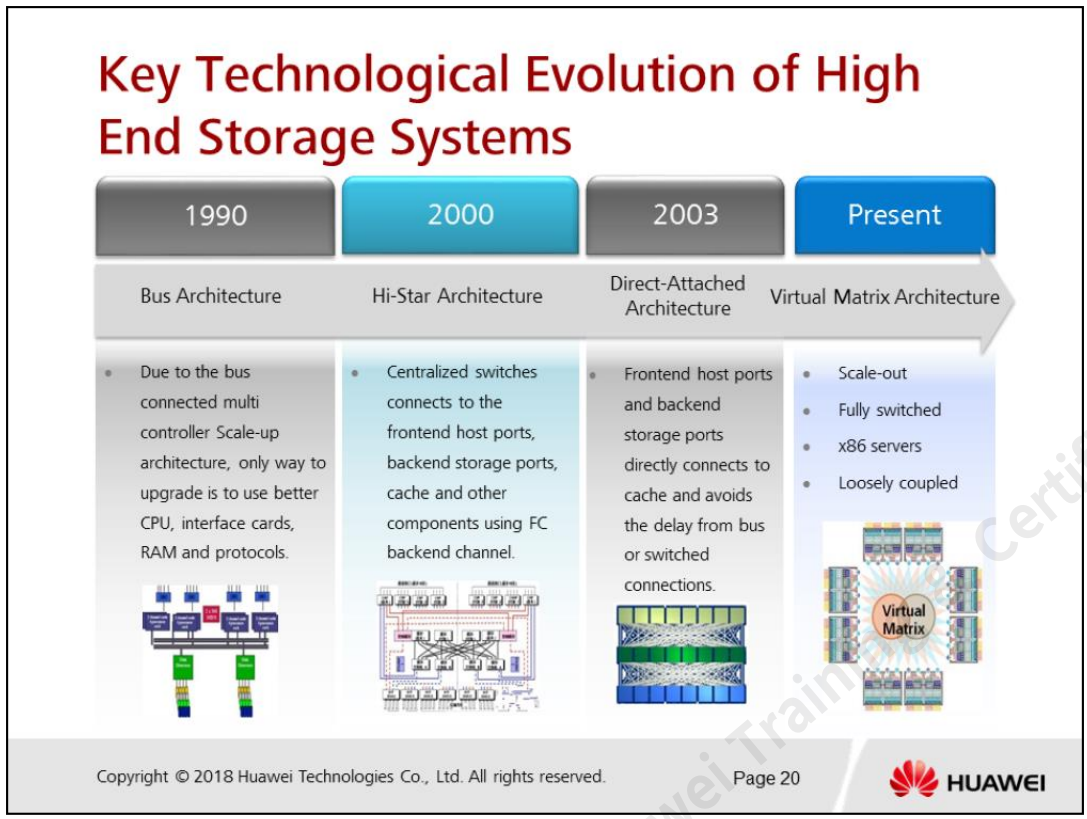
Converged Storage:Intelligent Multi Service Carrier



Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

- High Flexibility: Converges and supports FC/ISCSI/NFS/CIFS/FTP/FCoE/IB protocols, virtualization and all major applications of databases.
- High Performance: Frontend and Backend bandwidth exceeds the industry by 50%, and SPC-1 performance exceeds the industry by 30%.
- High Expansion: Cache and Storage Volume expansion capabilities exceeds 3 times of normal industry capabilities .
- High Fault-Tolerance: 30minutes/TB data reconstruction, zero hotdisks, auto repairs bad sectors.
- Fine Granularity: SmartTier/SmartCache/SmartMotion fine-grained hot data analysis and auto data migration enabling 3 times higher performance, SmartThin copy-on-write(COW) and disk reclamation technologies increases disk space utilization by 3 times.
- Service Adaption: Supports multiple standards for VAAI/VASA/SRM/SMI-S/Energy saving/security access/application platform, and adapts to the customer service environment and provides data protection.
- Auto Recovery: Proprietary RAID 2.0+ technology allows it to flexibly control data through fully virtualized mechanism which work seamlessly on virtualized platforms for data auto recovery and reconstructions.
- Interoperability with High End Storages: Industry's only storage device that is capable to interoperate and copy data between high end storage systems.
- Intelligent Management: Can complete deployment of 100 logical unit number(LUNs) in 1 minute allowing 10 times increased efficiency and an unified management system for managing all Huawei storage devices.

Huawei Training & Certification Huawei Training & Certification

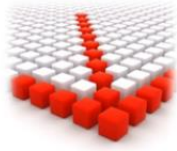
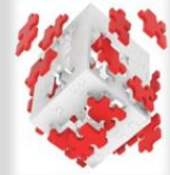


- High end storage devices has high reliability, high expansion system architecture, has rich data management and data management features. It ensures high level of service continuity and high level of service quality standards. It carries the critical business applications and is the core of the storage platform in enterprise grade datacenters.

Huawei Training & Certification

Current Features of High End Storage Systems

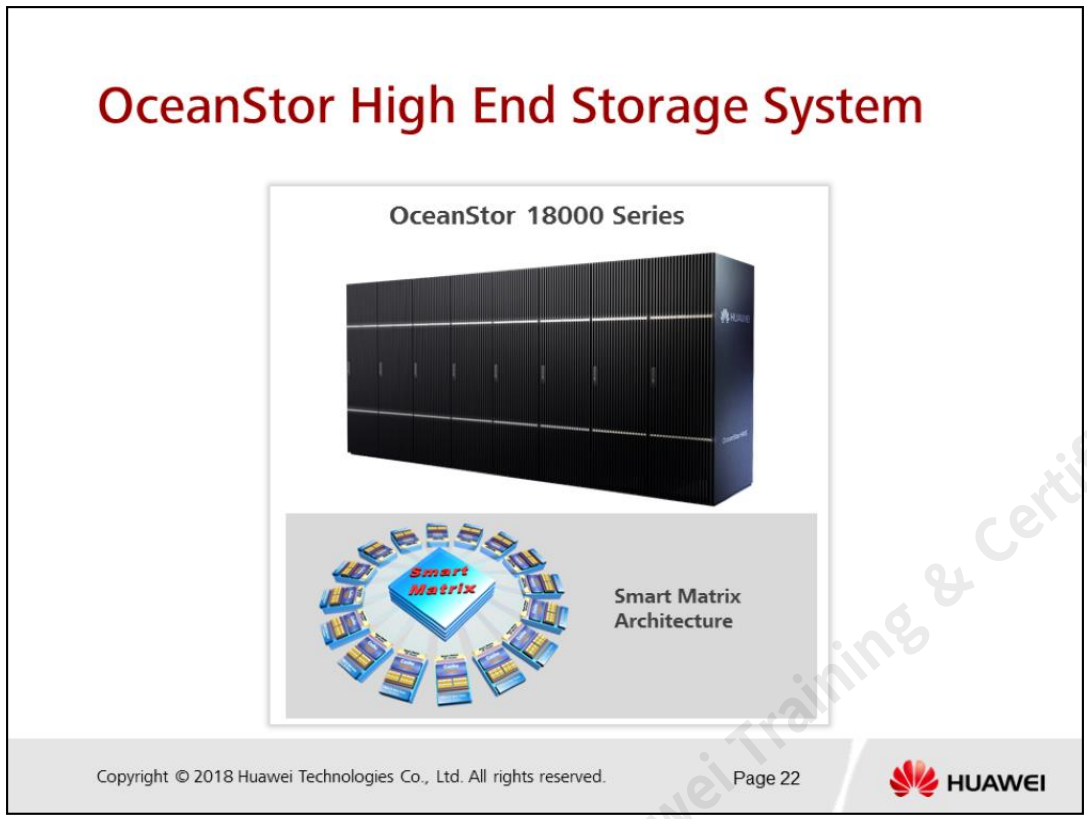
Core Storage Platform (Centralized)

System Architecture	Data Protection	Quick Response	Online Expansion
			
• Fully switched multi controller architecture • Global Cache • End to end verification • Limits number of disks in loop	• Local Snapshot Protection • Second-level copying	• Cache Partition, Global Cache • QoS • Auto data categorization • SSD Optimization	• Linear performance expansion • Online volume expansion • Virtualization

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

Huawei Training & Certification
Huawei Training & Certification



- Safe and Reliable:
  - Multi Controller Fully Redundant System Architecture: it eliminates single-point of failures (SPOF) Cache load balancing ensures stable operation of the system.
  - Based on RAID 2.0+ disk virtualization: Allows zero hotdisk, and 20 times data reconstruction speed which greatly enhances data reliability.
  - Complete Disaster Recovery Solution: Industry shortest 0~5 second Recovery Point Objective (RPO), with one-click disaster recovery management ensures highest degree of service continuity.
- Elastic and High Performance:
  - Elastic expansion of capabilities based on service need, which fulfills the storage needs for next 10 years.
  - Quickest Response: The only high end storage with millisecond level stable response, and increases the service processing speed by 10 times.
  - High Efficiency: Fine tuned for highest storage efficiency, and increases the storage utilization rate by 3 times.

Huawei Trainings & Certification
Huawei Training & Certification

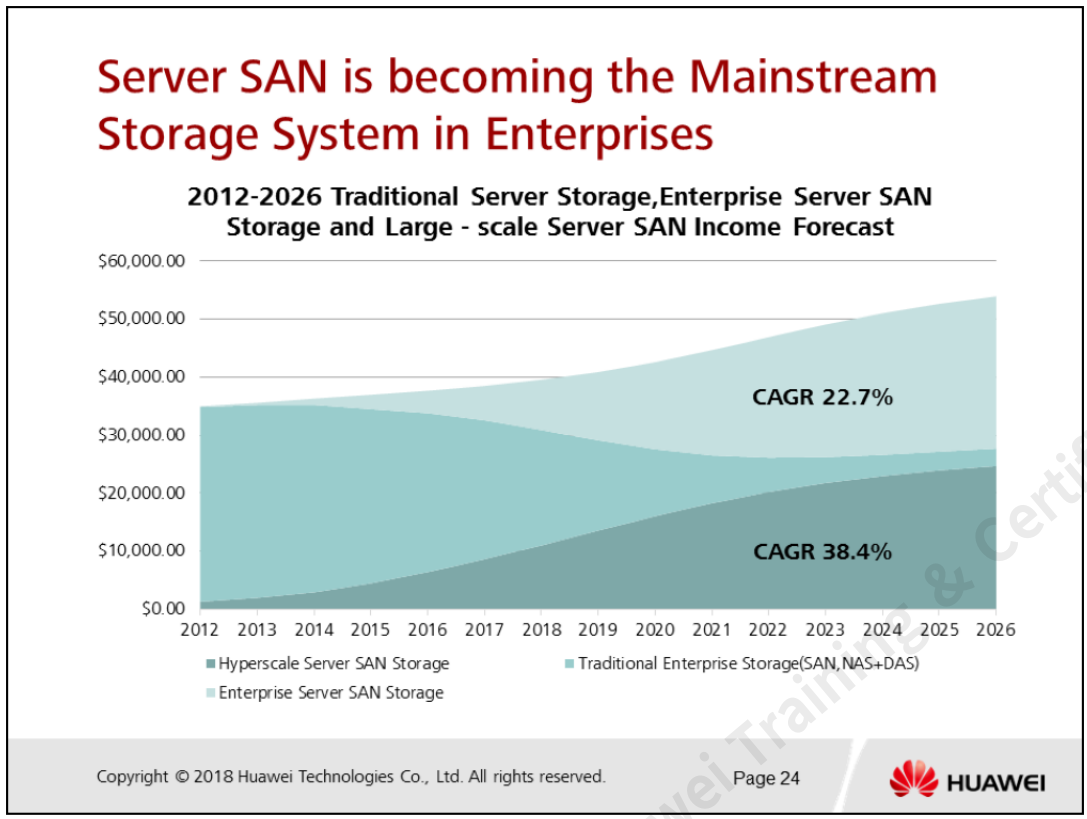
Contents

1. Definition of Storage
2. Development History of Storage
3. Evolution of Storage Technologies
4. **Latest Storage Technologies and Trends**
5. Storage Products and Solutions

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

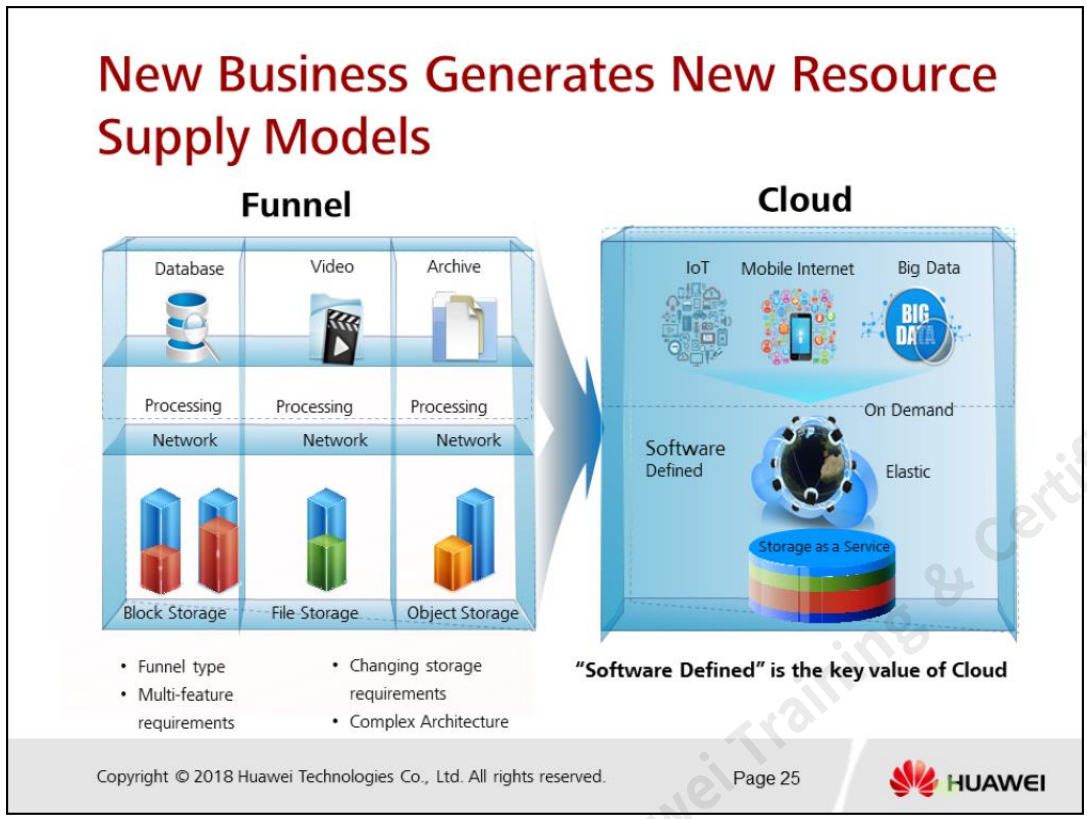
HUAWEI

Huawei Training & Certification
Huawei Training & Certification



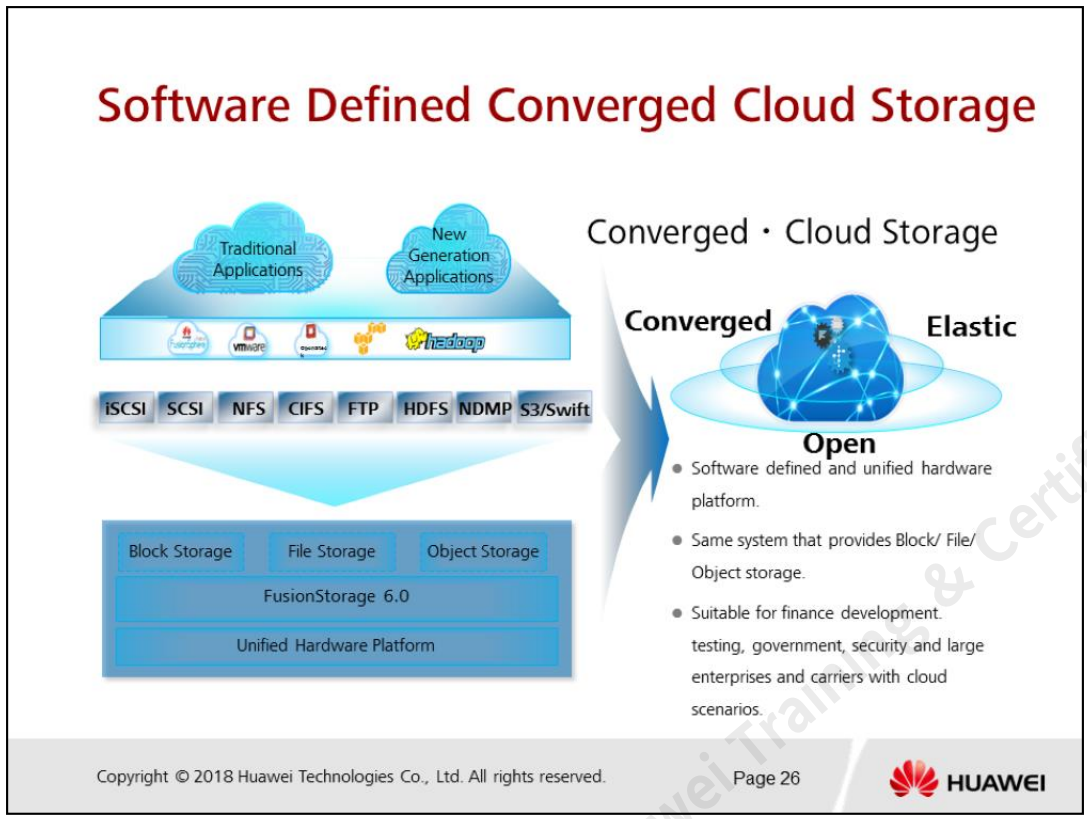
- Cloud Storage Market: Quickly growing and expected to reach $29B by 2020.
  - Benefited from the requirements of centralized storage resource pools, lower TTM (time to market) and TCO (total cost of ownership). Public cloud and private cloud is growing by 22.7% and 38.4% CAGR respectively.
- Public Cloud: Aims at the emerging trend of increasing application performance and reliability.
  - Standard Scenarios: Web、VDI、HPC、IOT、RDS、Wed Disk.
- Private Cloud: focuses increasing efficiency on large scale VM/Containers storage pools.
  - Standard Scenarios: VSI、VDI、Big data.
  - Key Requirement: Cloning/Active-Active、Deduplication/Compression、High bandwidth and Virtual machines compatibility.

Huawei Training & Certification



- Traditional storage are designed for single purpose or single scenario uses, and these funnels cannot satisfy the elastic expansion requirements. On the premise of cloudification, new business need new storage resources that supports elasticity, on demand and scalable expansion.

Huawei Training & Certification



- FusionStorage Software Defined Converged Cloud Storage is based on standard X86 hardware platform for high performance, high scalability and open compatibility with storage resource pools to provide file/block/object based storage for finance, government, and new businesses.
- The product has the high performance, elasticity and flexibility of distributed storage architecture for meeting agile business needs, and also has the reliability and usability of enterprise grade products to easily face the challenges of globalization, cloudification and big data applications.
- All new products are built based on demands and Cloud Storages are no exception. Enterprise Cloud datacenters has hundreds or even thousands of application that presents data in different service formats and data structures such as structures data and unstructured data. Traditional and New generation applications along with Big Data applications also coexist at the same time in the cloud datacenters.
- Huawei FusionStorage6.0 cloud storage system consist of multiple types of resource pools converged together to provide support for multi protocols data access. It also provides agile resource supply and file services such as block/file/object storage to higher level applications.
    - If you are building a Cloud Datacenter, choosing Huawei Cloud Storage not only allows you to fulfill current storage resource pool needs, it also helps in fulfilling future storage needs, volumes and TCO requirements.
    - If your Datacenters has large amount of databases applications, Huawei Cloud Storage resource pools can provide a highly efficient data storing and reading services.
    - If your Cloud Datacenter has applications based on Openstack or Hadoop, Huawei Cloud Storage can fit in seamlessly due to its open architecture.
    - Hence, software defined cloud storage can help protect your investments in datacenters and help realize the best match for resources and value.

Huawei Training & Certification

Contents

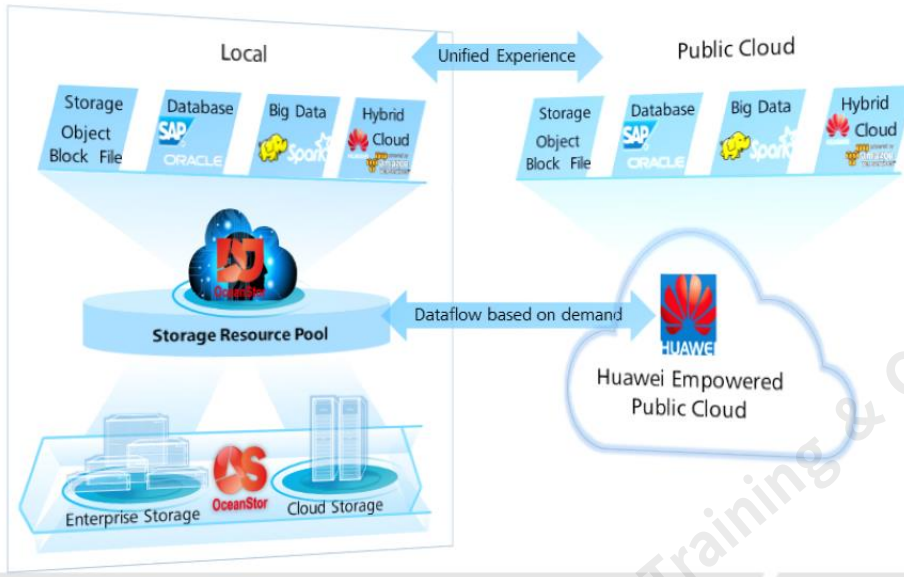
1. Definition of Storage
2. Development History of Storage
3. Evolution of Storage Technologies
4. Latest Storage Technologies and Trends
5. **Storage Products and Solutions**

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

Huawei Training & Certification
Huawei Training & Certification

STaaS:Storage as a Service Solution

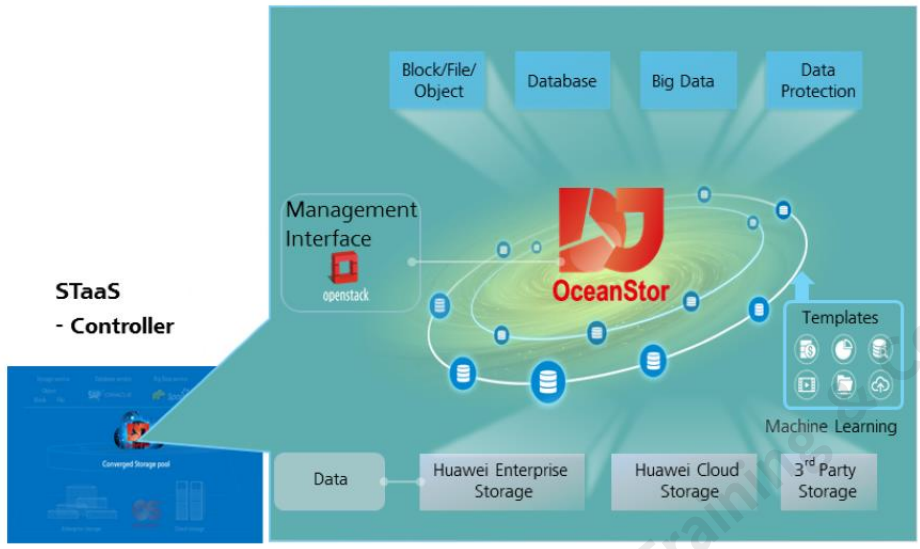


Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

Huawei Training & Certification
Huawei Training & Certification

OceanStor DJ:Easy,Intelligent,Open

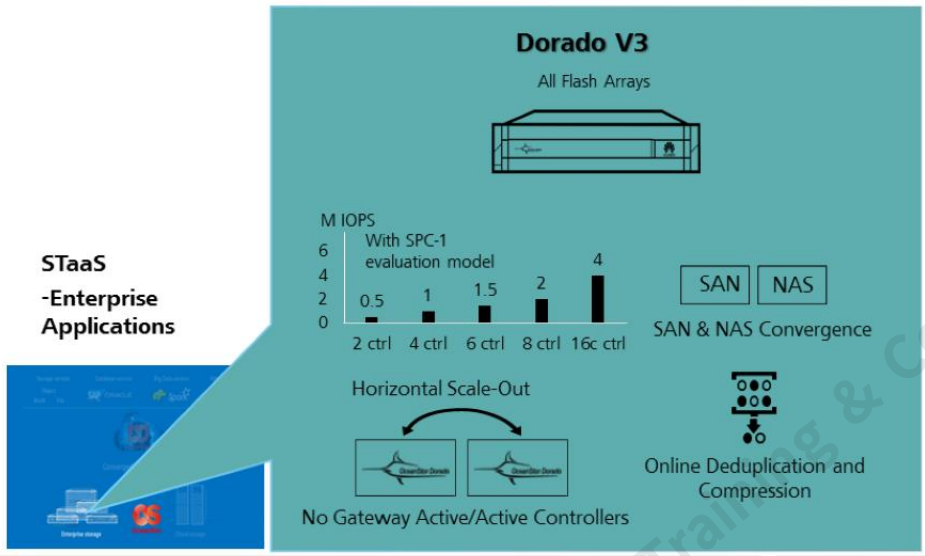


Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

Huawei Training & Certification

OceanStor Dorado V3:Fast & Reliable



Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

Huawei Training & Certification

OceanStor V3:Converged Storage for Agile Business

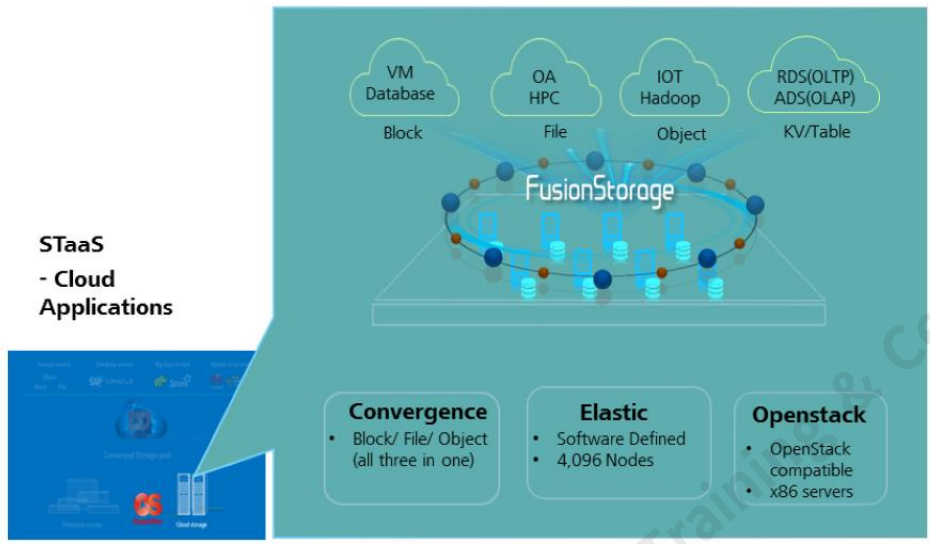


Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

Huawei Training & Certification

FusionStorage:Converged Cloud Storage

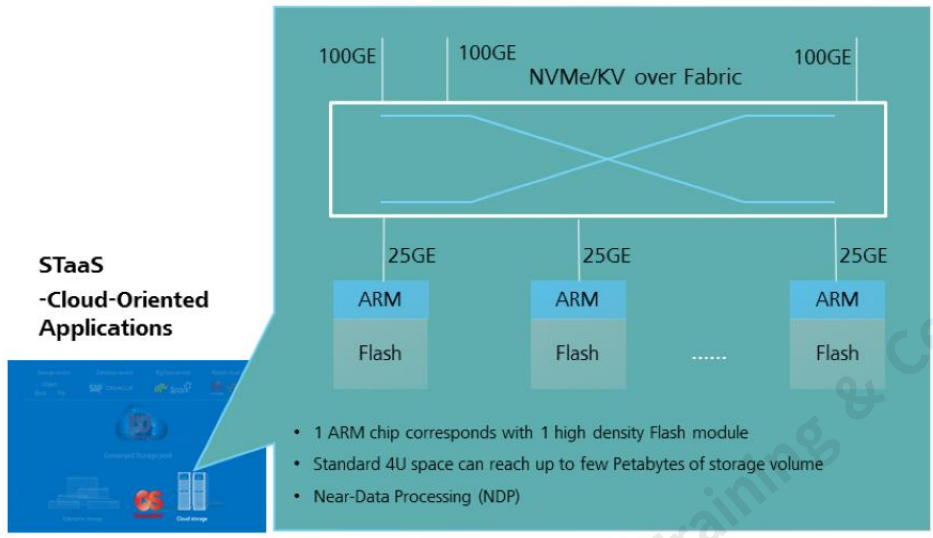


Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

Huawei Training & Certification
Huawei Training & Certification

Berreleye:Cloud - Oriented New Generation Hardware Architecture

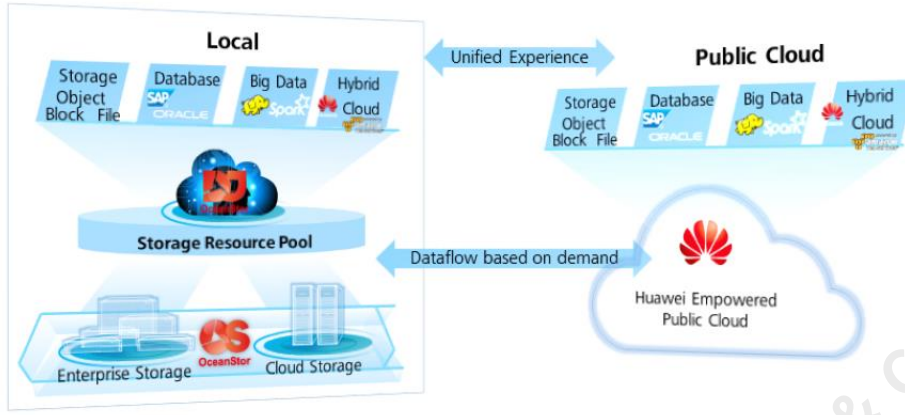


Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

Huawei Training & Certification
Huawei Training & Certification

STaaS:Hybrid Cloud Solution



Unified Management

- Unified Management of Local and Public Cloud Storage
- Storage as a Service

Dataflow based on Demand

- Incremental backup, supports CDP
- Logs replication, second-level RPO
- Online data replication

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

Huawei Training & Certification
Huawei Training & Certification

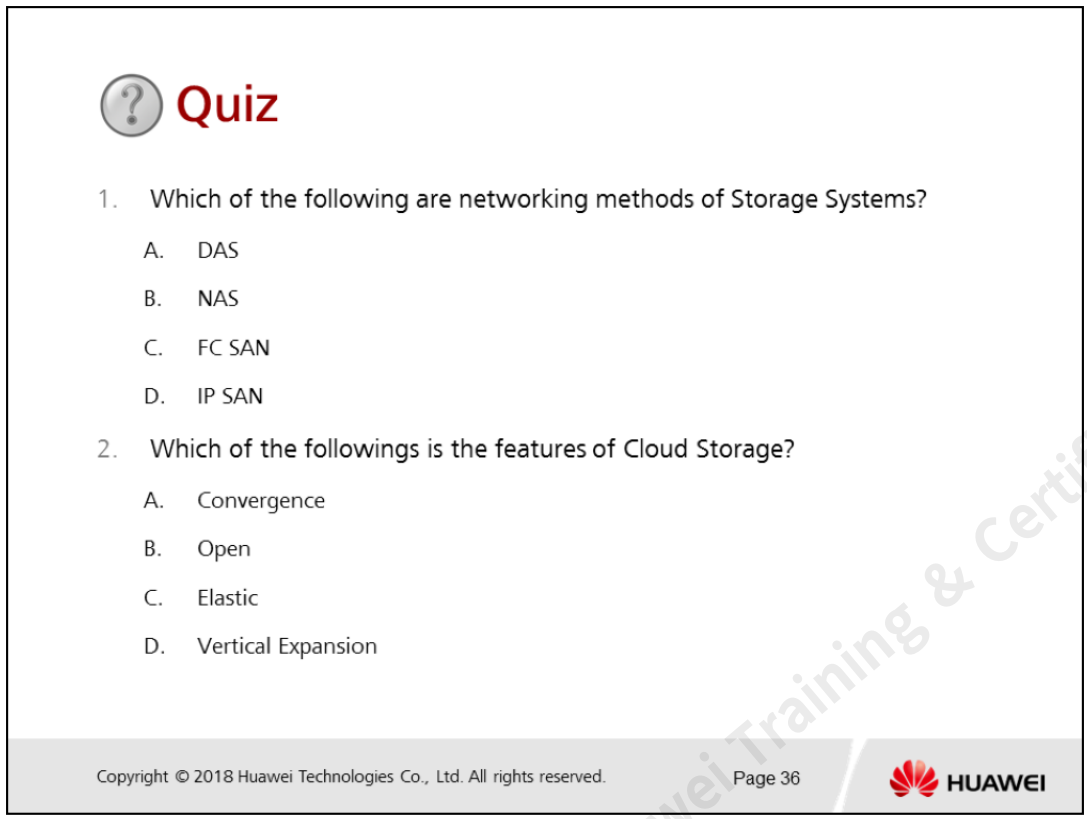
Section Summary

- This module mainly introduces the definitions, development history, technological evolution and technological trends of storage systems.
- This module also briefly introduces the different types of Huawei storage solutions.

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

Huawei Training & Certification
Huawei Training & Certification



- Answers:
  - ABCD
  - ABC

Huawei Training & Certification
Huawei Training & Certification

Thank You

www.huawei.com

Huawei Training & Certification
Huawei Training & Certification

02 Storage Technologies for AI, Big Data and the Cloud

www.huawei.com

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

Huawei Training & Certification
Huawei Training & Certification

Foreword

- This module mainly introduces:
  - Development trends of new ICT architectures.
  - Concept of the Cloud, the key technologies used and the application of storage in the Cloud.
  - Concept of Big Data, the key technologies used and the application of storage in Big Data.
  - Converged technologies and applications between the Cloud, Big Data and AI.

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

Huawei Training & Certification
Huawei Training & Certification

Objectives

- Upon completion of this module, you will be able to:
  - Understand the development trends of ICT.
  - Understand what is the Cloud, Big Data and AI.
  - Learn about storage technologies and its application in the Cloud.
  - Learn about storage technologies and its application in AI and Big Data.

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

Huawei Training & Certification
Huawei Training & Certification

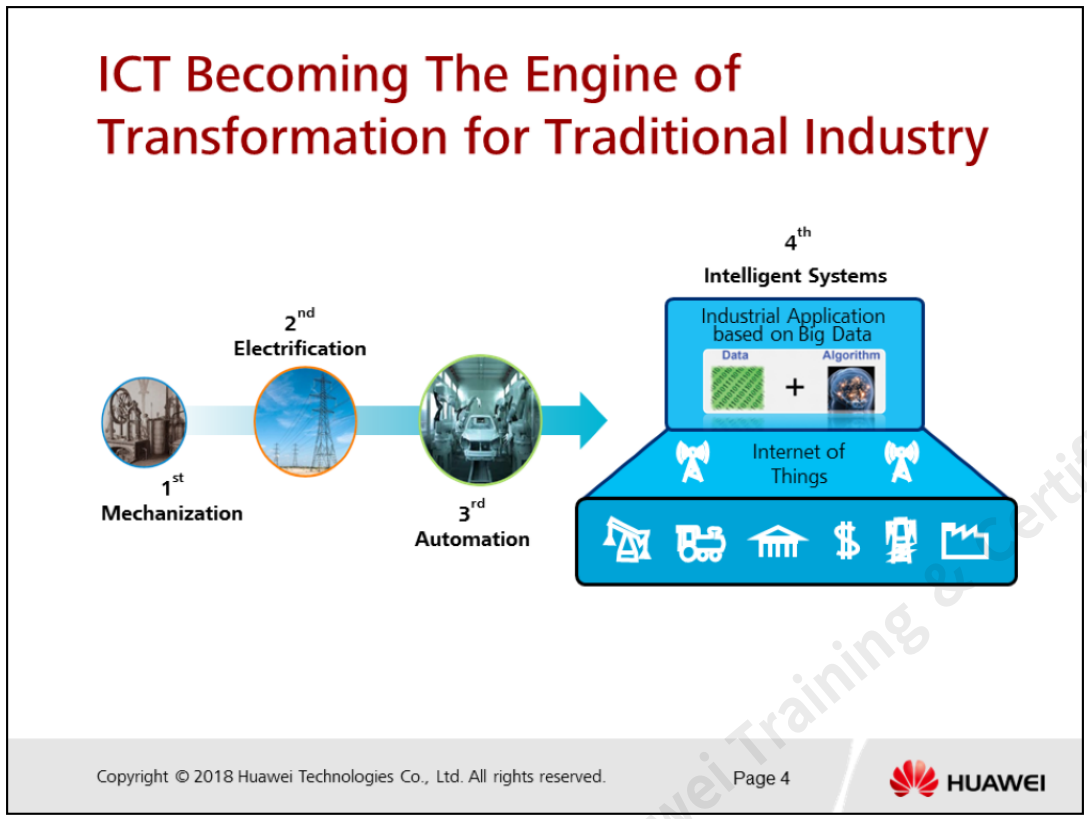
Contents

1. **ICT Technologies Development Trends.**
2. Storage Technologies and Its Application in the Cloud.
3. Storage Technologies and Its Application in AI and Big Data.

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

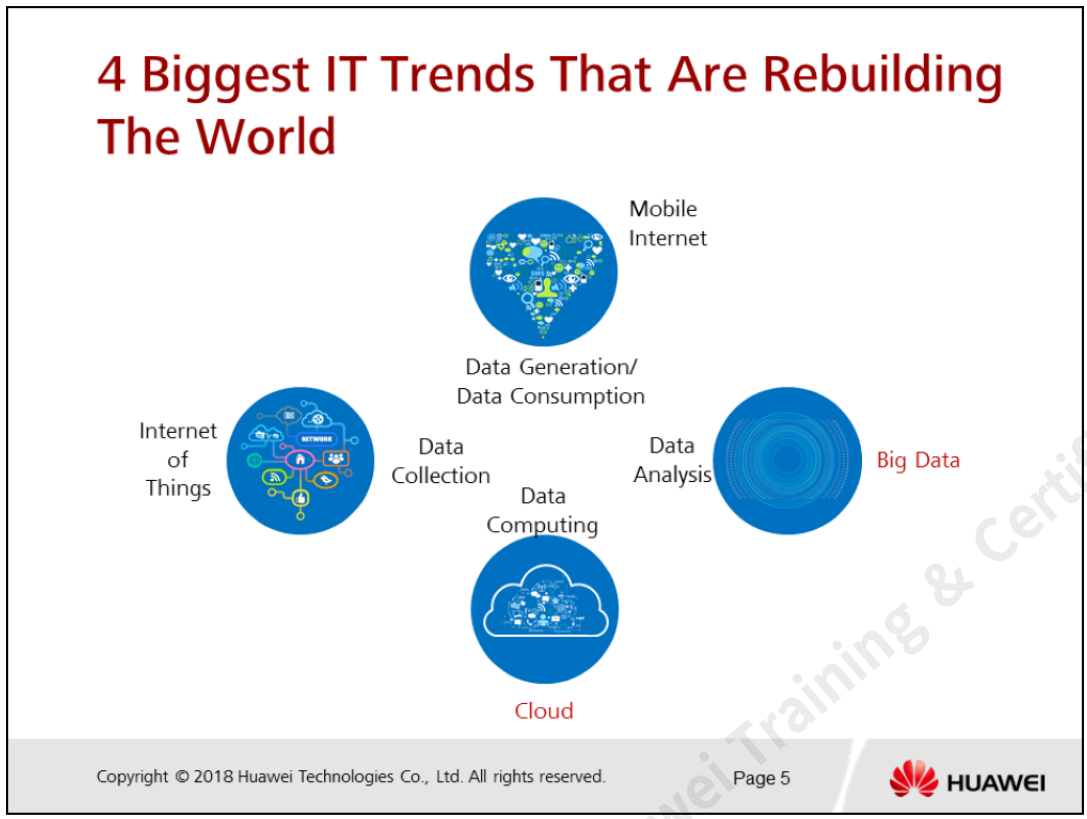
HUAWEI

Huawei Training & Certification
Huawei Training & Certification



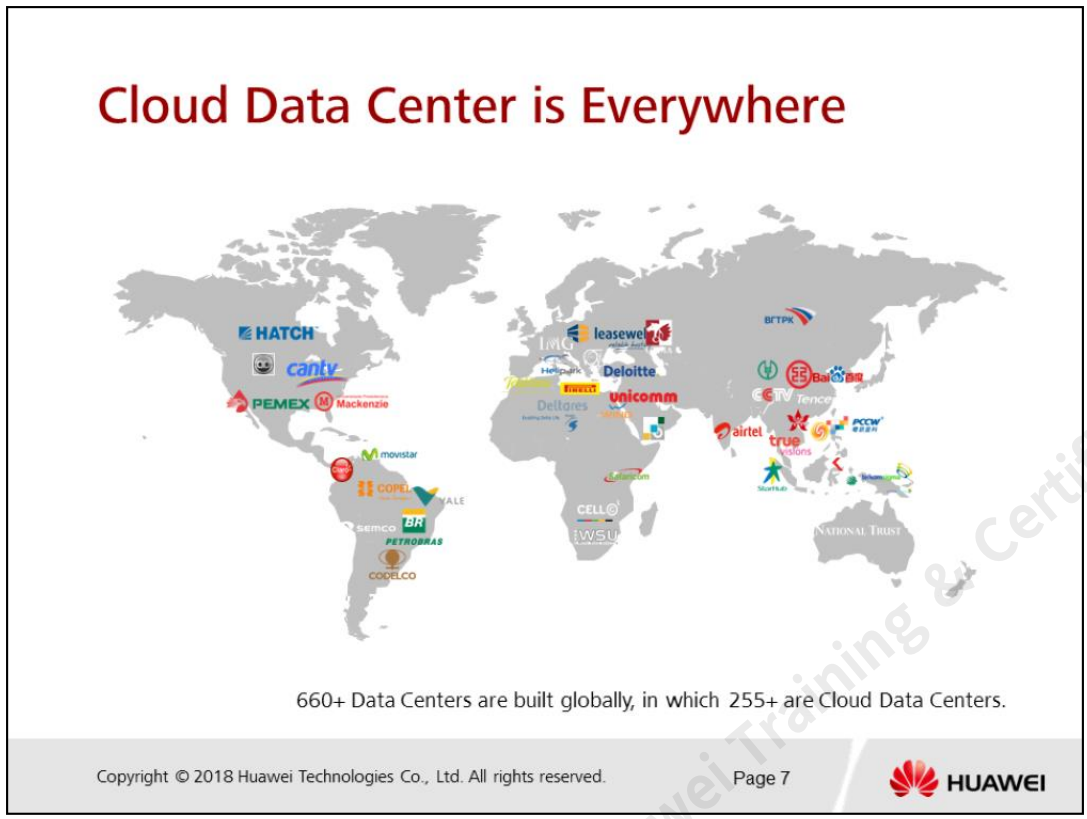
- All types of industries are facing the challenges of new IT technologies, and the new trends in IT are changing the world and rebuilding the rules that forms the basis of the digital world. As new technologies emerge, new industries are formed and there industries that are being replaced by these technologies, we can see the effect of technology on the world by looking at the timeline of the Industrial Revolution.
- The First Industrial Revolution used water and steam power to mechanize production. The Second Industrial Revolution used electric power to create mass production. The Third Industrial Revolution used electronics and information technology to automate production. Now a Fourth Industrial Revolution is building on the Third, the digital revolution that has been occurring since the middle of the last century. It is characterized by a fusion of technologies that is blurring the lines between the physical, digital, and biological spheres to create intelligent systems. Some of the intelligent systems such as industrial application based on Big Data, and Internet of Things are key elements that will bring forward the Fourth Industrial Revolution.

Huawei Training & Certification



- Mobile Internet: 7 billion mobile users, which is near to the total population on Earth. 78% CAGR of data growth and the total number of smart phones in the world (1.82 billion) has exceed the number of PC (1.78 billion). The compound annual growth rate (CAGR) is the mean annual growth rate of an investment over a specified period of time longer than one year.
- Internet of Things: Social Business and Social Media changed the people's way of life. 86% of enterprises are developing their business in social media sites.

Huawei Training & Certification



- Market trends has a huge change from prioritizing internal resources to external resources. Back in the days, enterprises are focused on building own data centers and server rooms, but slowly the market trend changed to using external services and resources, which promoted the growth of data centers worldwide. Data centers as a service are very common nowadays, and businesses no matter big or small can purchase resources from data center service providers worldwide.

Huawei Training & Certification

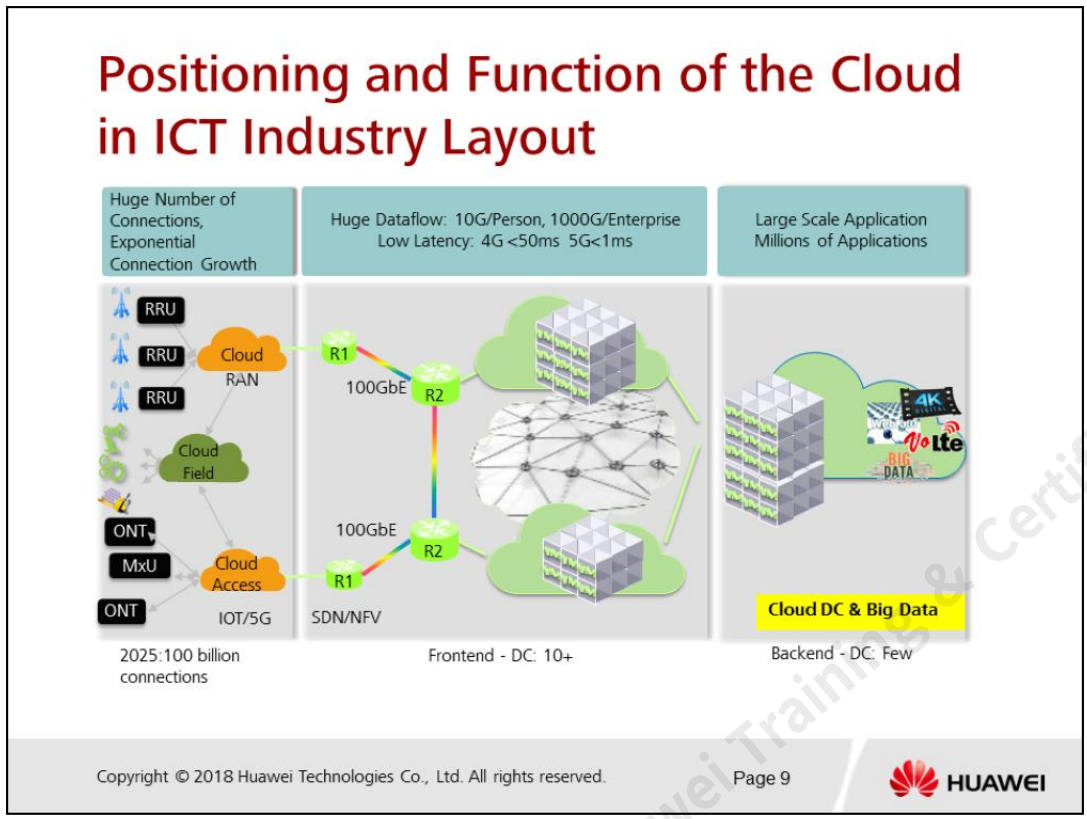
Contents

1. ICT Technologies Development Trends.
2. **Storage Technologies and Its Application in the Cloud.**
3. Storage Technologies and Its Application in AI and Big Data.

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

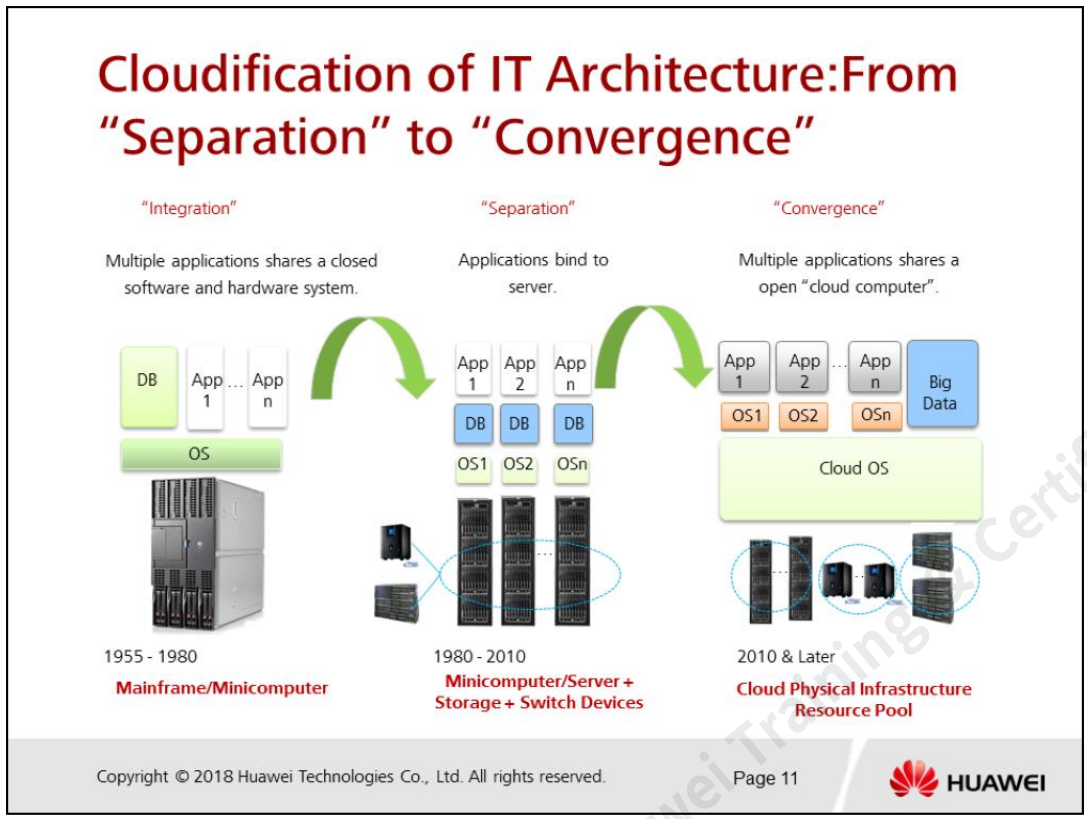
HUAWEI

Huawei Training & Certification
Huawei Training & Certification



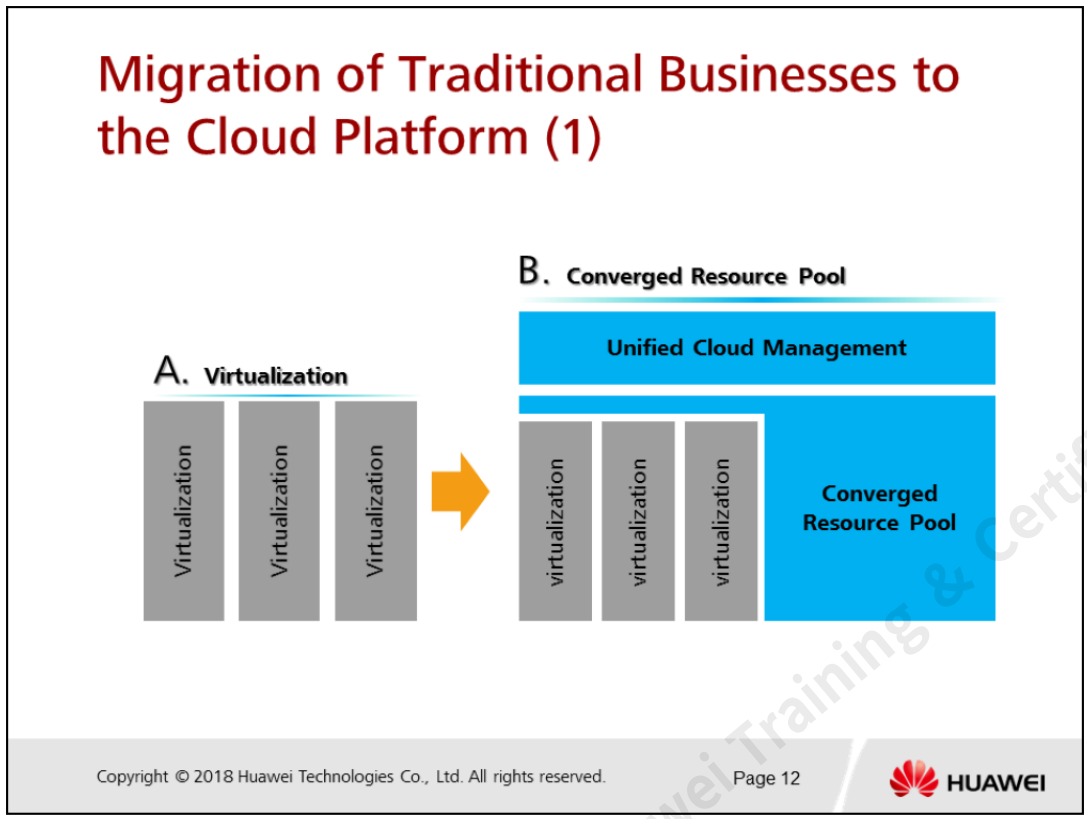
- The whole ICT industry layout has been set and just waiting for the maturity of the big modules and interconnection of those modules. Technologies within the modules such as 5G, NFV/SDN, Cloud Data Center, Big Data are growing in a very quick pace.
- Cloud Data Centers faces 3 big challenges: huge increase in connection flow, huge increase in data processing, and huge increase in application services.
- Internet of Things (IoT): It is the interconnection via the Internet of computing devices embedded in everyday objects, which enables them to send and receive data. The internet of things (IoT) is also a computing concept that describes the idea of everyday physical objects being connected to the internet and being able to identify themselves to other devices, and exchange data between them.
- SDN (Software Defined Network), is an emerging network architecture that decouples the network control and forwarding functions, which enables the network control to become directly programmable and the underlying infrastructure to be abstracted for applications and network services. In a software-defined network, a network administrator can shape traffic from a centralized control console without having to touch individual switches, and can deliver services to wherever they are needed in the network, without regard to what specific devices a server or other hardware components are connected to. The key technologies for SDN implementation are functional separation, network virtualization and automation through programmability.

Huawei Training & Certification



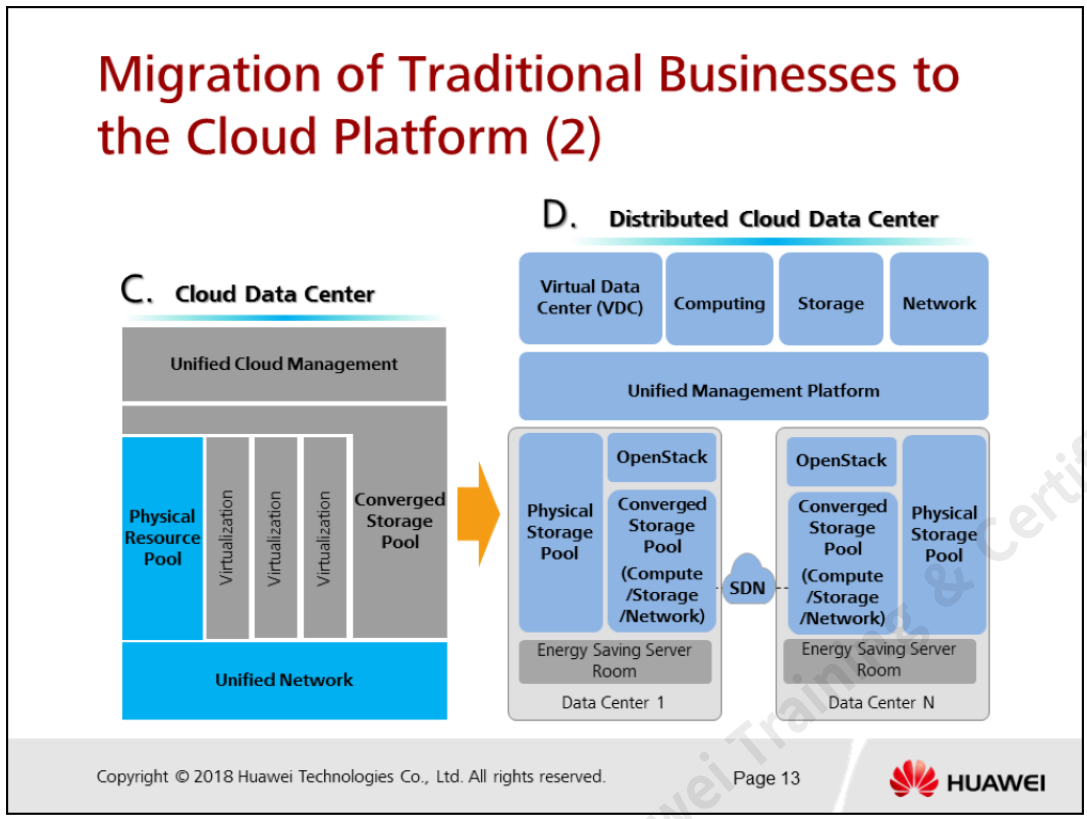
- Convergence of physical infrastructure: Heterogeneous computing, storage and network devices in the data center are unified and converged via cloud operating system with automated resource management.
- The driving force of value: Lower TCO(Total cost of ownership), higher efficiency in service deployment, more agile response towards client's requirements.

Huawei Training & Certification



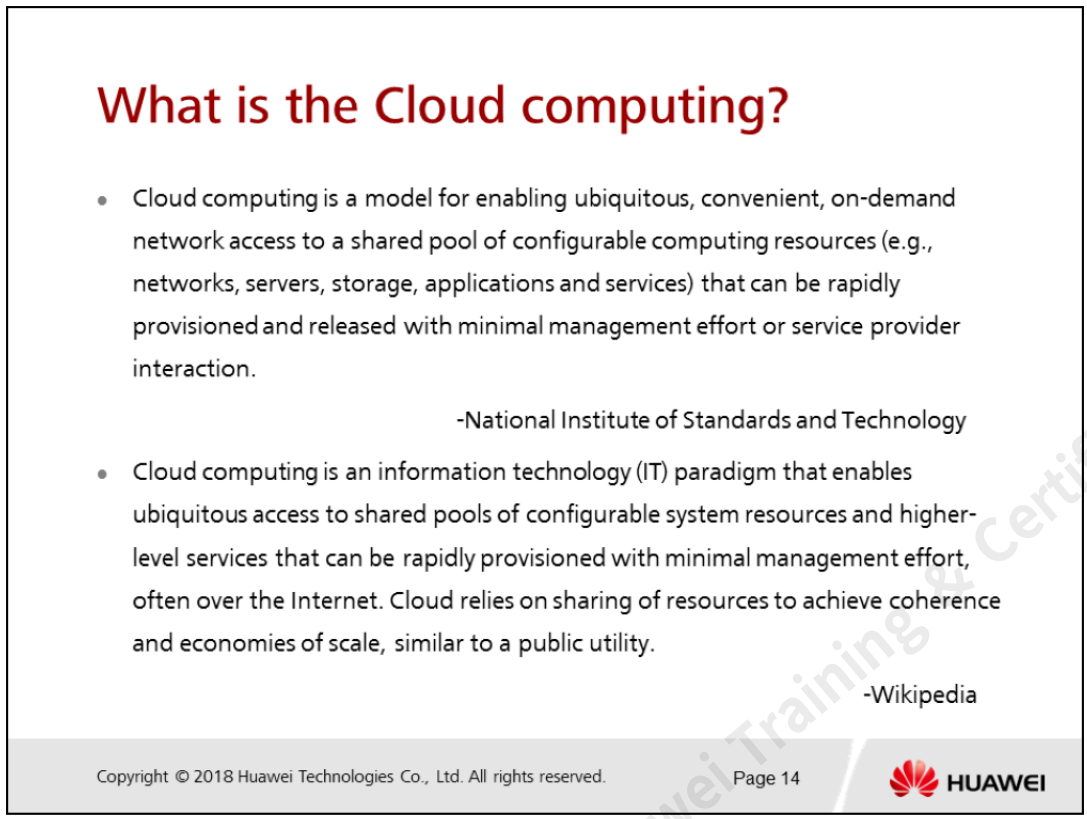
- Infrastructure:
  - Interconnection and evolution of existing networks.
  - Security isolation and security protection.
  - Building a differentiated resource pool that is able to satisfy different business needs.
- Management:
  - Management of existing resource pools.
  - Unified management of virtualized and physical resources.
  - Set resource pool management standard practices, standardized use of resources.
- Migration of services:
  - Smooth migration of old physical machines and data.
  - Smooth migration of virtual machines and data.

Huawei Training & Certification



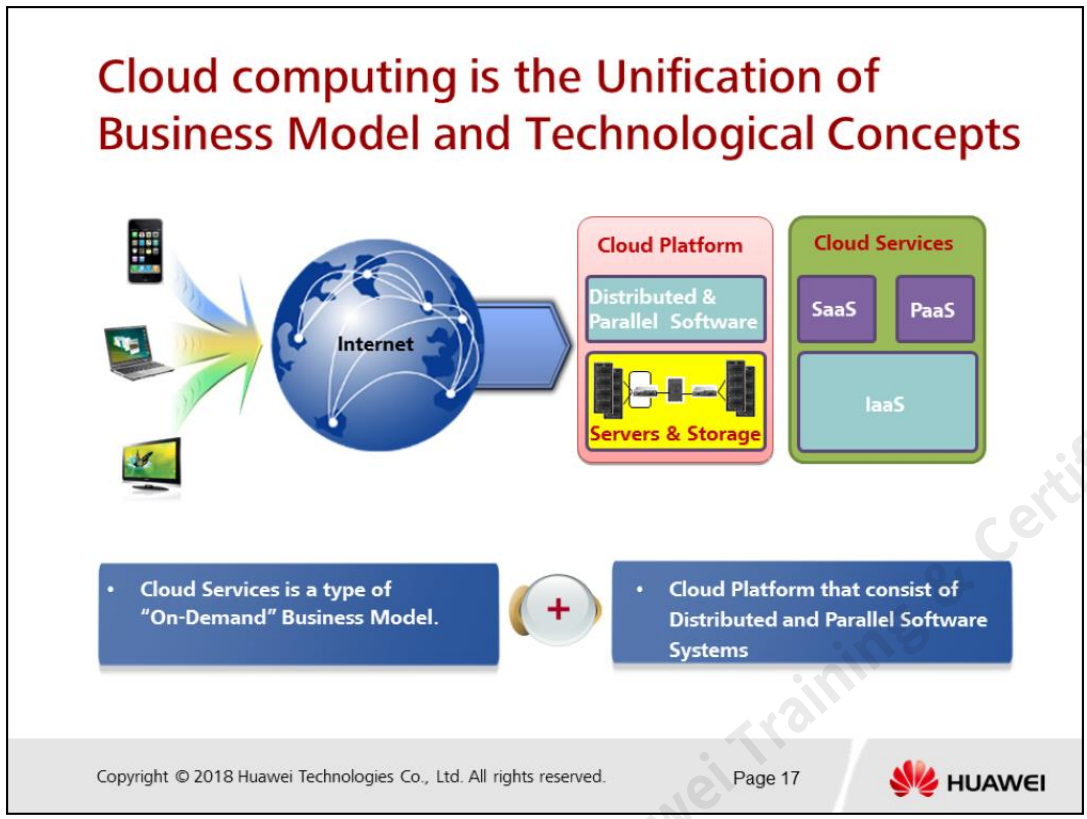
- Virtualization is the process of running a virtual instance of a computer system in a layer abstracted from the actual hardware. Most commonly, it refers to running multiple operating systems on a computer system simultaneously. To the applications running on top of the virtualized machine, it can appear as if they are on their own dedicated machine, where the operating system, libraries, and other programs are unique to the guest virtualized system and unconnected to the host operating system which sits below it.
- SDN encompasses multiple kinds of network technologies designed to make the network more flexible and agile to support the virtualized server and storage infrastructure of the modern data center. Software-defined networking can be defined as an approach to designing, building, and managing networks that separates the network's control or SDN network policy and forwarding planes thus enabling the network control to become directly programmable and the underlying infrastructure to be abstracted for applications and network services for applications as SDN, Cloud or mobile networks.
- OpenStack is a cloud operating system that controls large pools of compute, storage, and networking resources throughout a datacenter, which is all managed through a dashboard that gives administrators control while empowering their users to provision resources through a web interface.

Huawei Training & Certification



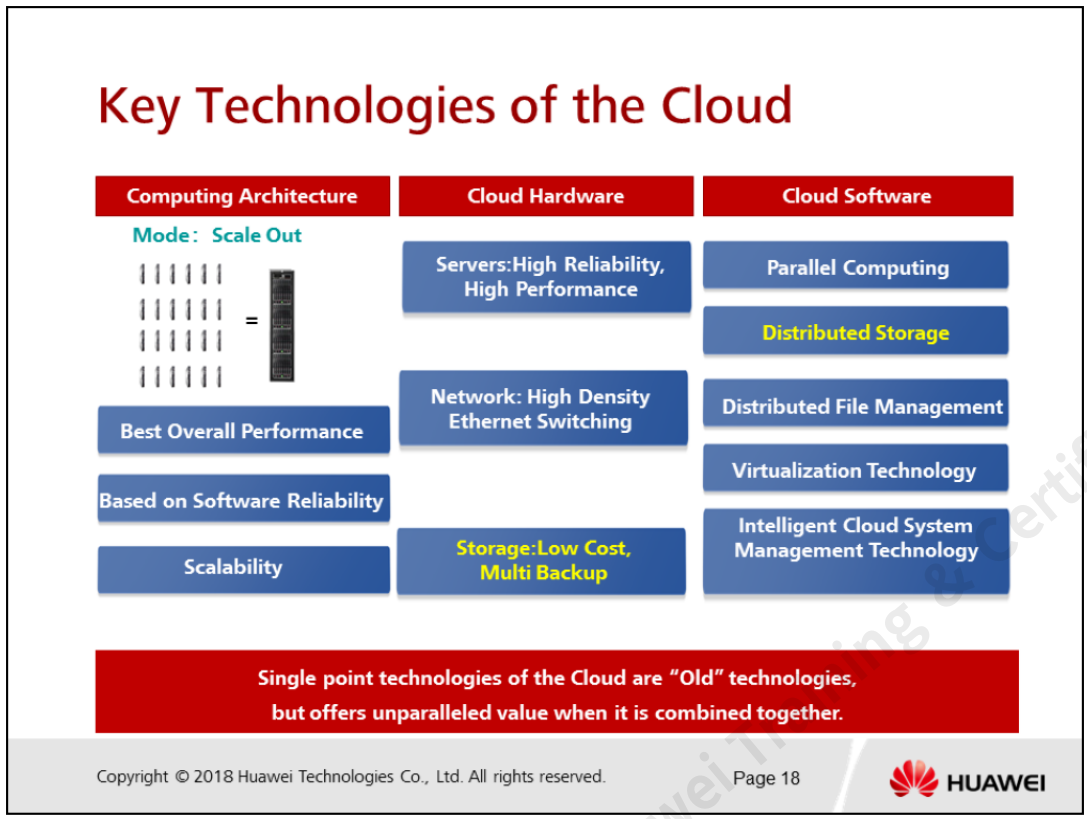
- Cloud refers to the addition, usage and delivery model of related internet-based services that typically involves the provision of dynamically scalable and often virtualized resources over the Internet. Cloud is a metaphor for the network and the Internet. In the previous diagrams, we often used a cloud to represent communication network, and even use it to represent the abstraction of the Internet and the underlying physical infrastructure. The narrow definition of the cloud refers to the delivery and usage model of IT infrastructure, in which resources that you require can be gained through the network based on demand in a easily scalable method. The wider definition of the cloud refers to the delivery and usage model of services, where the services you need can be obtained on demand and in a easily scalable way through the network. These services can be IT, software or internet related services or any other services. This in a sense also means that computing power can also be circulated over the internet as a form of product or commodity.
- Cloud resources are dynamically scalable, virtualized and provided over the Internet. End users do not need to know the details of the infrastructure in the cloud, and they do not have the expertise in cloud technologies, and they also do not need to directly control them, in fact they should just focus on what resources they really need and how they will be able to get the service over the network.
- The cloud provides a layer of abstraction between the end users and the physical infrastructure that is running the cloud. This makes it easier for end users to get the resources and services they need without the hassle of needing to learn or have the expertise in the multiple technologies required to run a cloud.

Huawei Training & Certification



- "On-Demand" business model is a business model where services are obtained based on customer's demand. The application and data required by the customer resides in the cloud, and the users use clients to access the required applications based on their demand. Cloud service providers provide the corresponding services to the customers based on their demand and collect corresponding payments for the services offered to the customers.
- From a business perspective, cloud services can be categorized as the following types:
  - IaaS (Infrastructure as a Service): IaaS provides the ICT infrastructure and resources that ensures normal operation of the services.
  - PaaS (Platform as a Service): PaaS provides a development platform for developers. In a traditional development environment, the developer needs to focus on the operating system and hardware while writing the codes for an application. However, with PaaS, the developer can focus on writing the codes for better applications while the infrastructure and platform is managed by PaaS vendors.
  - SaaS (Software as a Service): This is the earliest service model in the cloud. Users just require a simple device to connect to the application and operating system provided by the SaaS vendor. All the complex system maintenance tasks such as software and licensing upgrades are managed by the SaaS vendor.

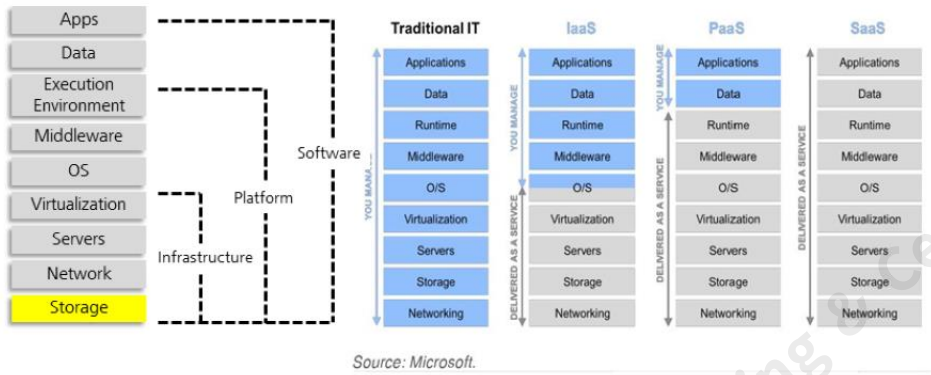
Huawei Training & Certification



- By summarizing multiple core technologies, it can be summed up in 3 main aspects which are: overall computing architecture, hosted hardware devices, and software system.
- The overall computing architecture requires high performance, high reliability and high scalability.
- Cloud hardware includes: High reliability and high performance computing servers that provides the computing resources. Low cost and data safe storage devices that provides data storage space. Supports data communication and exchange via high density switches in large layer 2 networks.
- Cloud software includes: parallel analysis computing technology used for Big Data. It also includes distributed storage technologies that provides dynamically scalable resource pool through converged storage resources, distributed file management used for data management and virtualization technologies for resource pooling of computing and storage. It also simplifies the operation and maintenance personnel's work through highly efficient and intelligent O&M system management technologies.

Huawei Training & Certification

Cloud Service Models

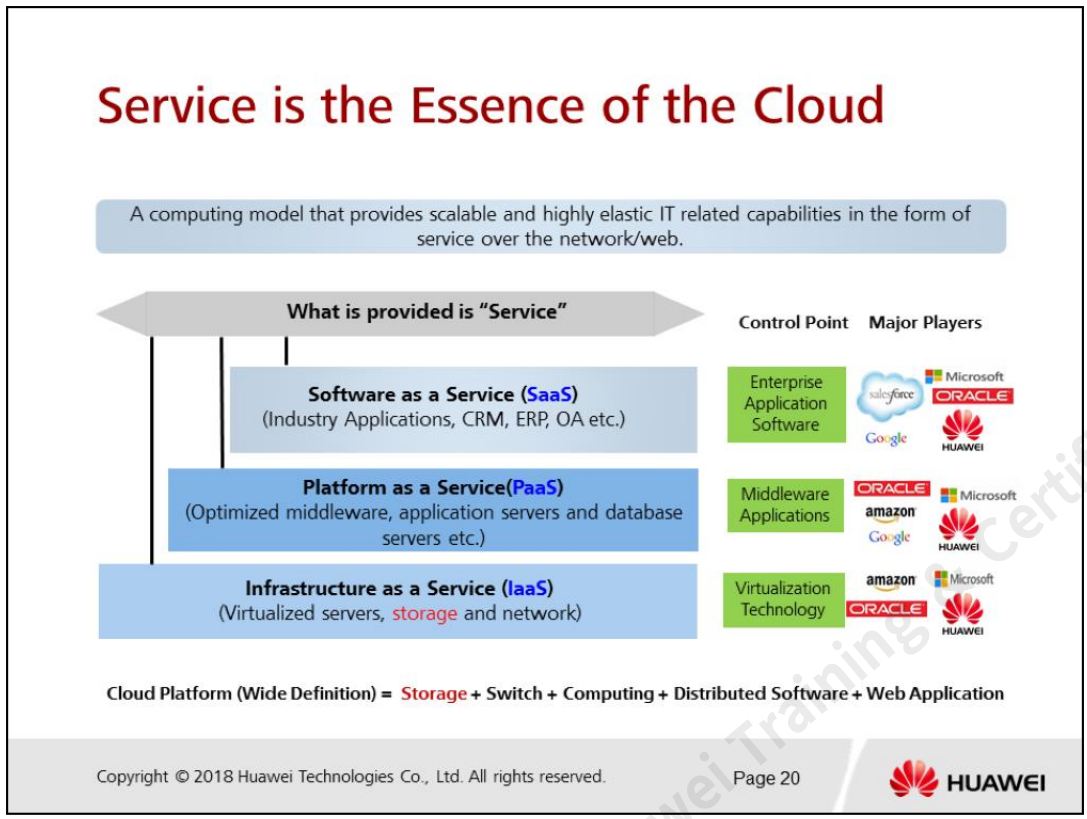


Source: Microsoft.

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

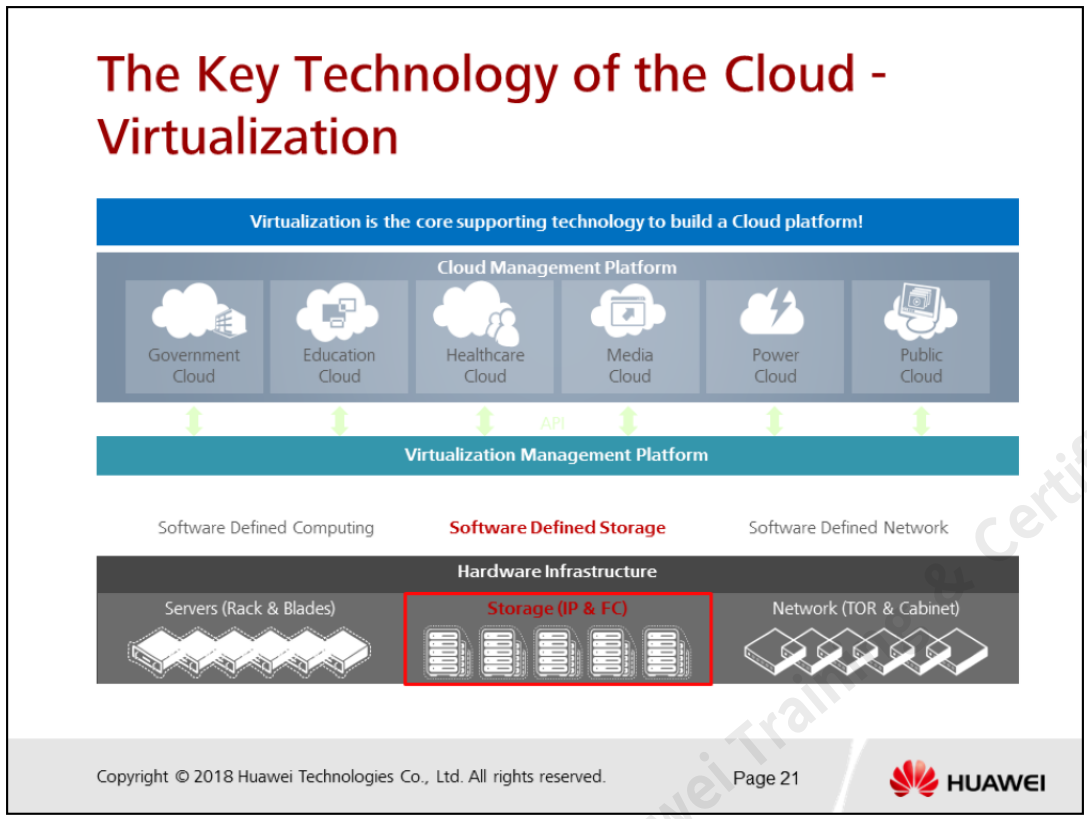
- The key characteristics of Cloud is:
  - On-demand Self-service
  - Ubiquitous Network Access
  - Location Independent Resource Pooling
  - Rapid Elastic
  - Pay Per Use

Huawei Training & Certification



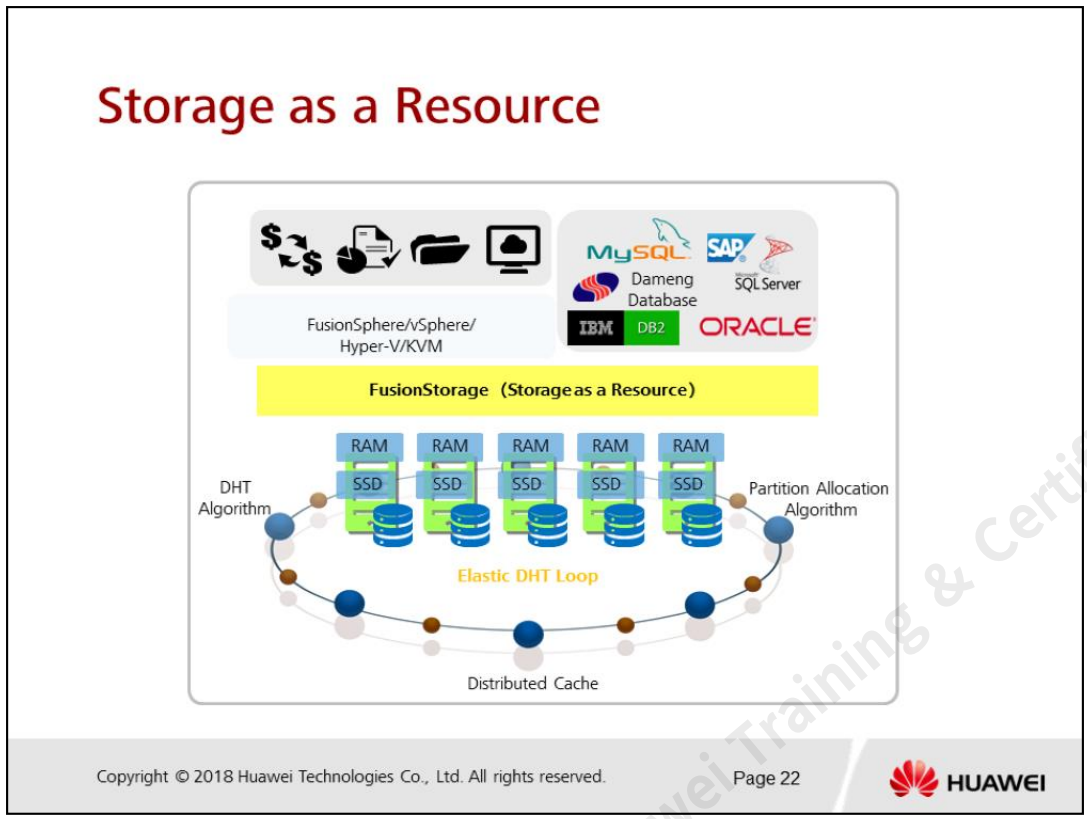
- The basic principle of a "Cloud Platform" is a distributed computing system, it provides On-Demand supercomputing and storage capabilities by distributing the computing load across a large number of computers. The terminal is also a part of the Cloud.
- The core part of the cloud platform is still the data center, but it is different in terms of technologies if compared to a traditional data center. Its grows in the direction that emphasizes standalone performance and clustering towards a "distributed, intelligent, and large capacity" data center.
- "Cloud" is just a metaphor, which refers to the networked distributed computing devices, and it also refers to the fact that data is hidden during the calculation process.

Huawei Training & Certification



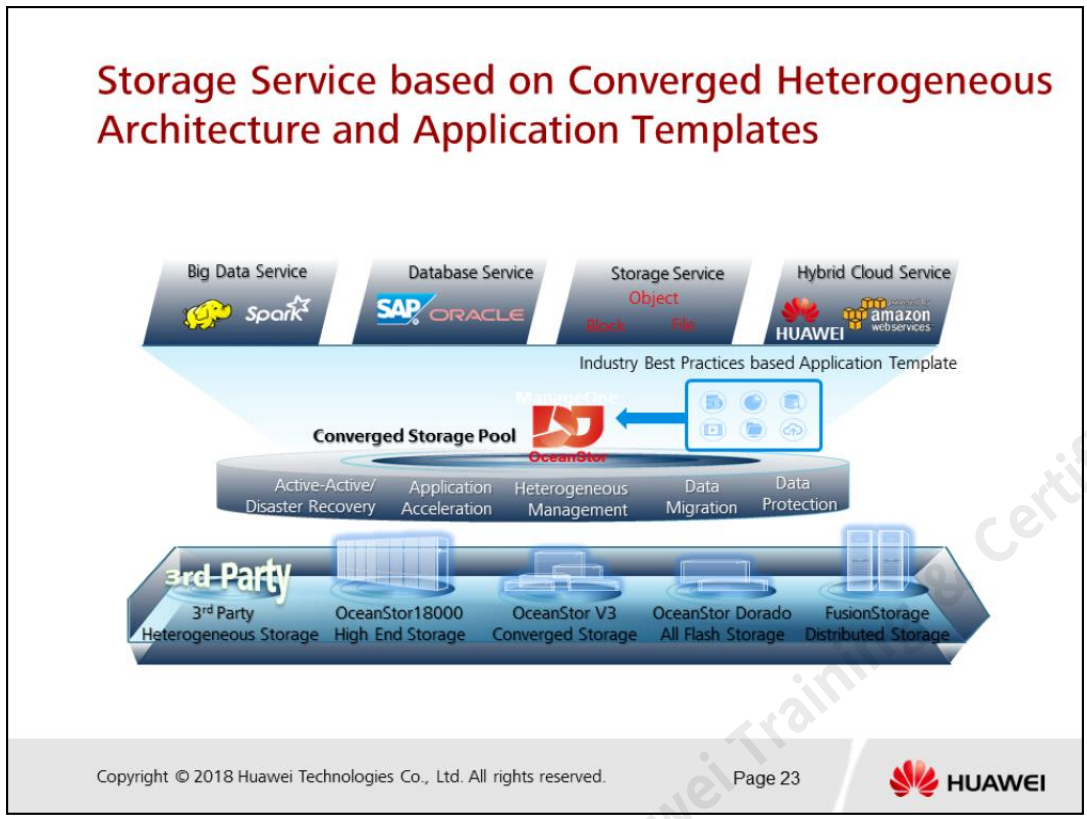
- Definition of Virtualization: Abstraction and isolation management technology towards physical resources in order to achieve the purpose of highest utilization rate of the physical resources.
- Types of Cloud deployment:
  - Private Cloud: Enterprise owned independent cloud infrastructure.
  - Public Cloud: Cloud infrastructure owned by cloud service providers that provides cloud service to the public or enterprises.
  - Hybrid Cloud: Combination of private and public cloud infrastructure through dedicated connection technologies to achieve data and application sharing.

Huawei Training & Certification



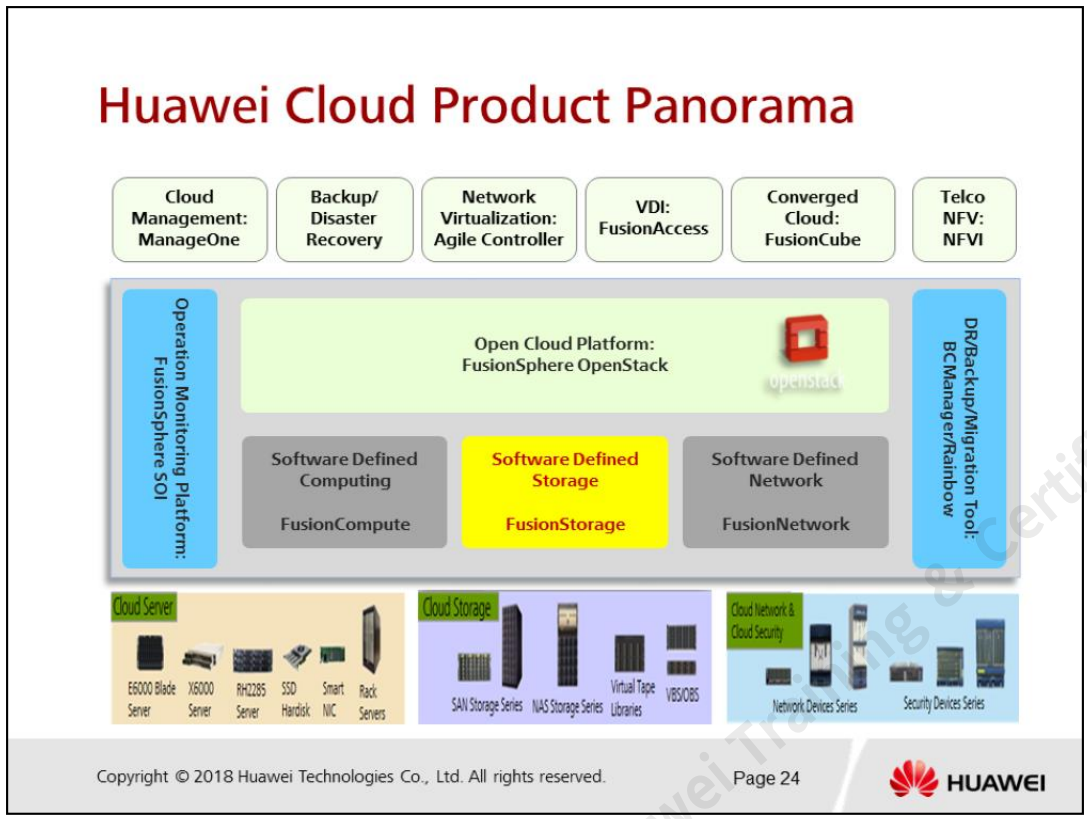
- Converged Storage:
  - Open: Supports deployment of multi vendor virtualized and physical database.
  - High Performance: Supports 4096 nodes, linear performance improvement, supports 10 million IOPS, and supports high performance databases and large capacity cloud resource pools.
  - High Reliability: Multiple backups, and supports up seven 9s reliability level (99.99999%), and has cabinet level reliability.
- Distributed Hash Table (DHT): It is designed for systems that has large number of nodes and the nodes are frequently added or removed.
- A distributed hash table (DHT) is a class of a decentralized distributed system that provides a lookup service similar to a hash table: (key, value) pairs are stored in a DHT, and any participating node can efficiently retrieve the value associated with a given key. Responsibility for maintaining the mapping from keys to values is distributed among the nodes, in such a way that a change in the set of participants causes a minimal amount of disruption. This allows a DHT to scale to extremely large numbers of nodes and to handle continual node arrivals, departures, and failures.
- DHTs form an infrastructure that can be used to build more complex services, such as anycast, cooperative Web caching, distributed file systems, domain name services, instant messaging, multicast, and also peer-to-peer file sharing and content distribution systems.

Huawei Training & Certification



- OceanStor DJ is a business-driven storage control software developed by Huawei for cloud data centers. It centrally manages data center storage resources, provides service-driven and automated storage services, improves storage resource utilization and effectiveness of storage service provisioning in cloud environments.
- The core of OceanStor DJ is based on the enhancement of OpenStack related services to realize the unified management of storage resources, on-demand distribution and data protection services. OceanStor DJ decouples applications from underlying storage and breaks the monopoly of legacy devices and application vendors. In cloud based scenarios, storage and data protection capabilities are provided in the form of XaaS, conforming to the storage value chain that is moving towards the direction of software and services.
- From the perspective of control plane and data plane, the value of OceanStor DJ is as following:
  - OceanStor DJ abstracts the storage features from the physical array on the control plane and logically combines multiple physical storage pools with the same or similar capabilities into a virtual storage pool. When a user requests storage resources, the user is allocated resources based on Resource Pool-based Service Level Agreement (SLA) without having to worry about which array on the backend provides storage services for their applications.
  - From the data plane perspective, OceanStor DJ has the ability to integrate all types of data services and the ability to support application access to block storage and file storage. At the same time, it also has the capability to use the unique features of the underlying storage arrays. Hence, OceanStor DJ retains the value added features of the storage arrays such as remote replication without adding to the purchasing cost of the user.

Huawei Training & Certification



- NFV (Network Function Virtualization) refers to the usage of common x86 hardware and virtualization technology to carry out software processing of multiple hardware features in order to lower the cost of purchasing expensive network devices. Through software and hardware decoupling along with the abstraction of features, network devices no longer depends on dedicated hardware. Thus, resources can be fully and flexibly shared, allowing quick development and deployment of new services, achieving automatic deployment based on actual business demands. It also allows elastic expansion and reduction, isolation of faults and self healing of services.

Huawei Training & Certification

Contents

1. ICT Technologies Development Trends.
2. Storage Technologies and Its Application in the Cloud.
3. **Storage Technologies and Its Application in AI and Big Data.**

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

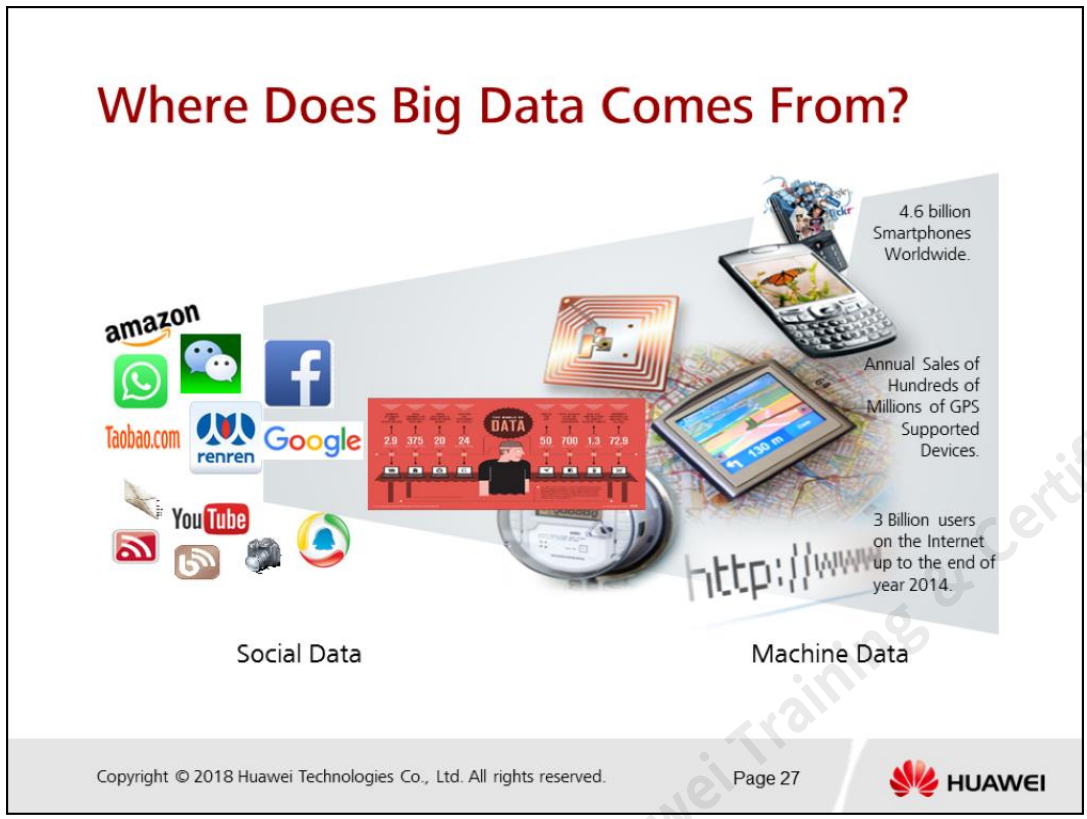
HUAWEI

Huawei Training & Certification
Huawei Training & Certification



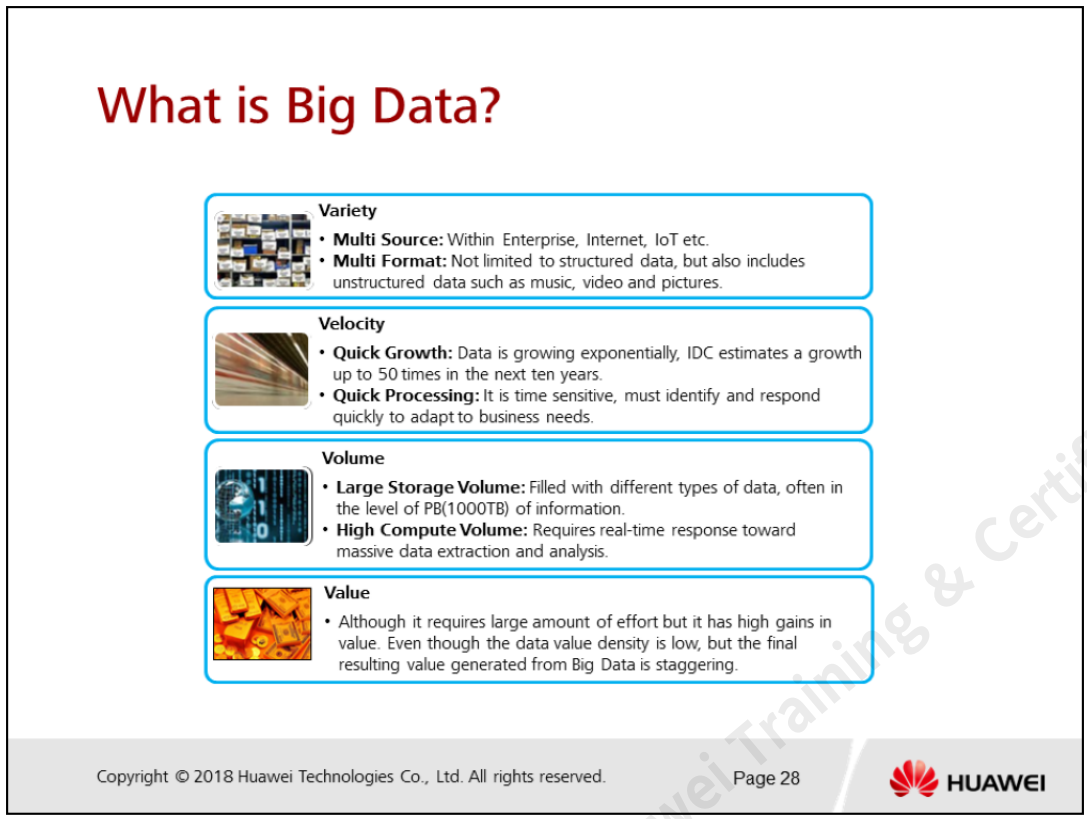
- Big Data Trends: It has overcome the initial stage, and become more and more mature, and it is leading towards the implementation phase.
- 50% of the enterprises has already invested and using Big Data, 33% of the enterprises are planning on how to utilize Big Data. As we see the trends of continual investment on Big Data technologies, Big Data is stepping towards the mature development phase.
- Cloud and Big Data are no longer "new" technologies, but have already become "mainstream" technologies.

Huawei Training & Certification



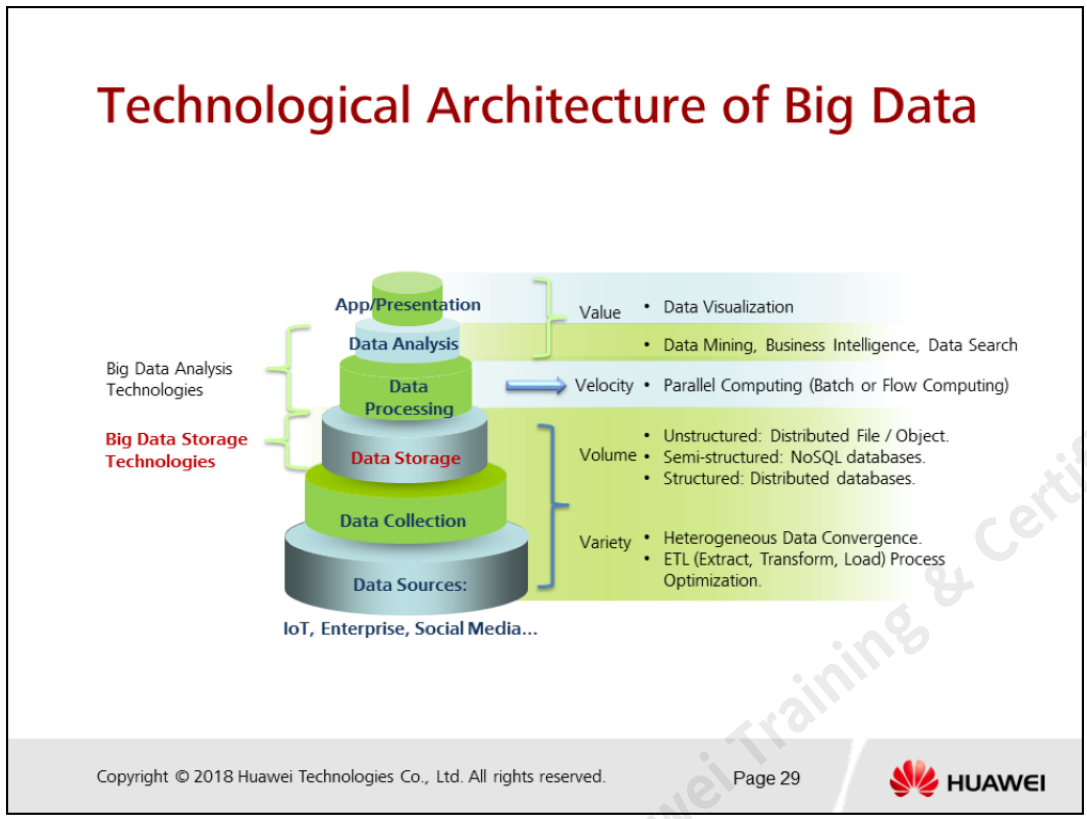
- Digital information created and shared globally increased nine-fold in five years to 3.8ZB (Zettabyte) in 2013.
- CERN (European Organization for Nuclear Research): Large Hadron Collider (LHC) generates 1PB/s worth of data. The LHC is the world's largest and most powerful particle collider, the most complex experimental facility ever built, and the largest single machine in the world.
- SKA(Square Kilometer Array): required storage space of 1EB (Exabyte) in 2015. The Square Kilometre Array (SKA) is a multi-billion dollar international project to build the world's largest radio telescope. Co-located primarily in South Africa and Western Australia, the SKA will be a collection of hundreds of thousands of radio antennas with a combined collecting area equivalent to approximately one million square metres, or one square kilometre.
- Construction of Cloud IDC also generated large concentration of data.
- Facebook: 50TB (Terabyte) of log data is generated daily, and up to 100TB of derived analytical data is generated daily.
- Manufacturing by machines and manufacturing by manual human labor generated massive amounts of data, and on the other hand, construction of centralized data centers also accelerates the centralization of data and data generation.

Huawei Training & Certification



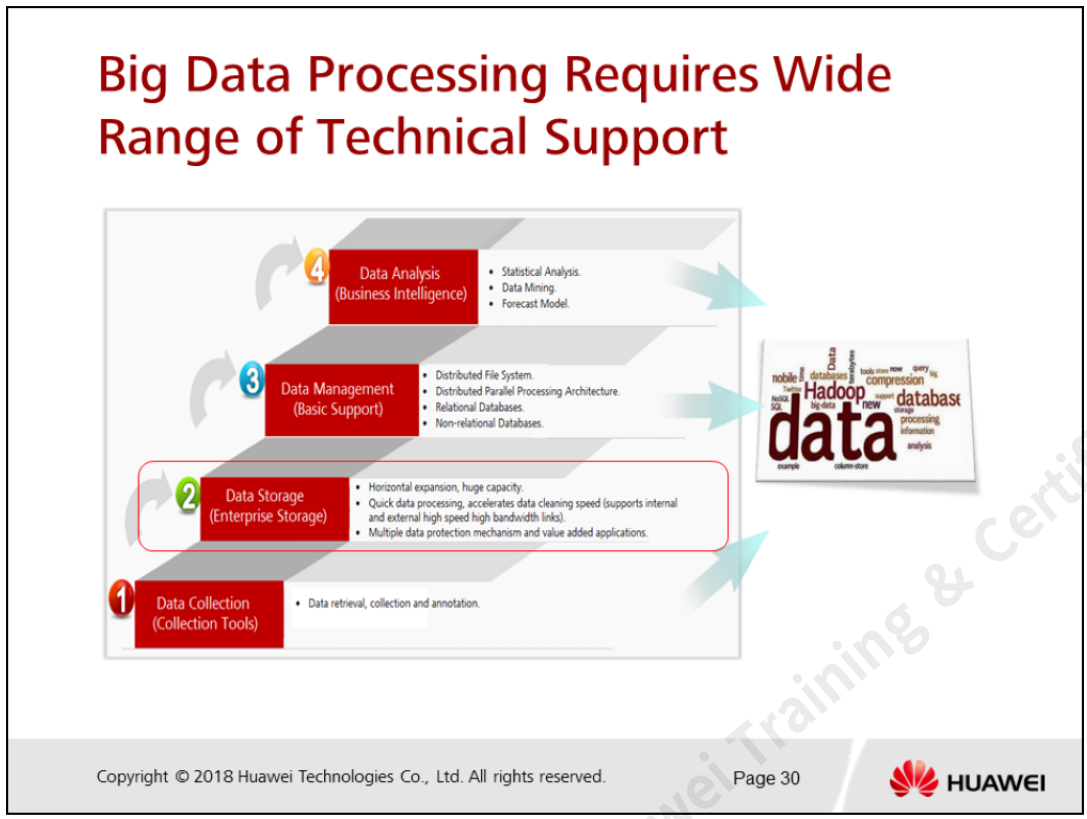
- Wikipedia: Big data is data sets that are so voluminous and complex that traditional data processing application software are inadequate to deal with them. Big data challenges include capturing data, data storage, data analysis, search, sharing, transfer, visualization, querying, updating and information privacy. The dimensions to big data is known as Volume, Variety, Velocity and Value.
- Gartner: Big Data is a massive, high-growth and diversified information asset that requires new processing models for greater decision-making, insight, and process optimization capabilities.
- The essence of Big Data: It is the reflection of the physical world in the digital world. For example, the flow of people in annual holidays will be recorded in the digital world.
- The logic of real world phenomena can also be found through big data analysis. For example, when a storm comes, we can see that the seagulls fly low. Through analysis, it is found that the seagulls fly low due to the appearance of many fishes on the surface of the sea which they could easily prey upon. So, why does the fishes swims near to the surface of the sea? Apparently, when the storm is coming, the water pressure in the sea increases, and fishes swim near to the surface of the sea where it is more convenient for them to breathe. These are all the relationship behind the phenomenon that can be figured out through big data analysis.
- Big data doesn't just refers to new data that are collected, in our society, the most valuable data still lies within the data that was accumulated by the enterprise over the years. Thus, data that was generated in the core systems of traditional data management is still relevant and valuable. It may not be the biggest amount of data or the latest or trendiest data, but it is the data that holds the highest commercial value to the enterprise itself.

Huawei Training & Certification



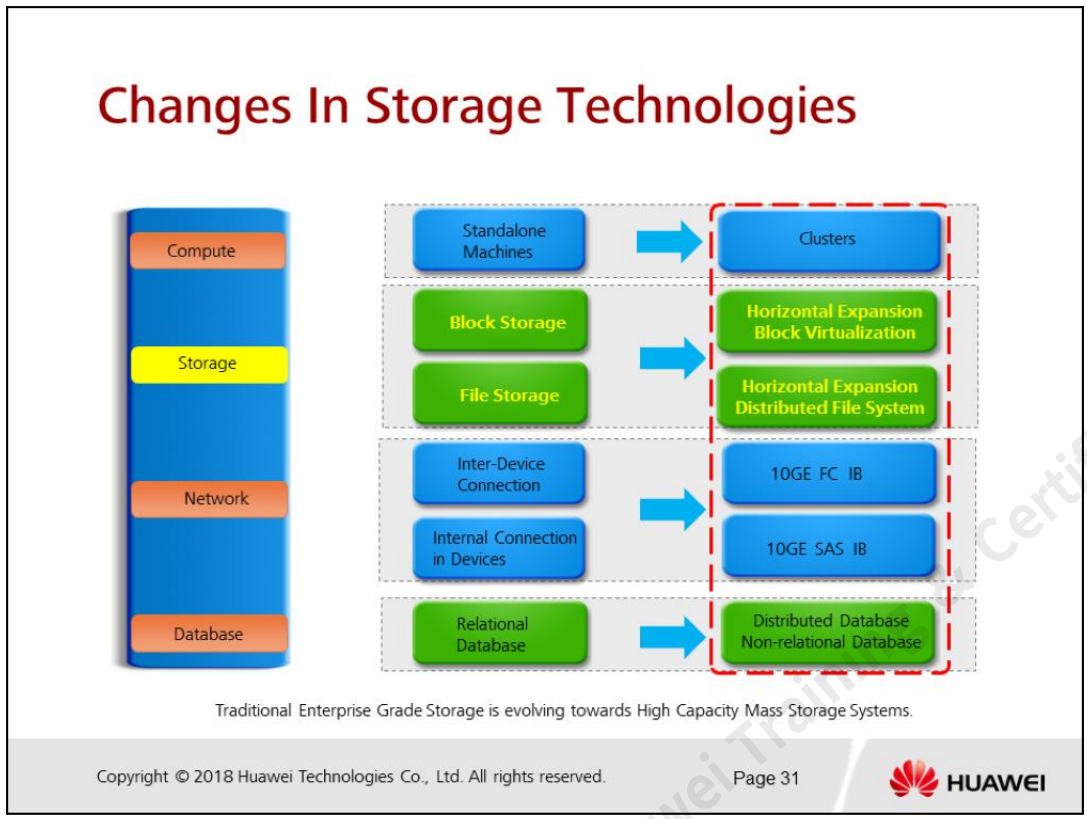
- From the perspective of the challenges faced by big data in terms of capacity, data diversity, processing speed and value mining, big data technologies cover a wide range of technologies ranging from mass storage and processing of data to applications that visualize those data, including convergence of heterogeneous data sources, massive distributed file system, NoSQL database, parallel computing framework, live stream data processing and data mining, business intelligence and data visualization.
- A typical big data processing system can be divided into 5 layers: Data Collection, Data Storage, Data Processing, Data Analysis and Application or Presentation of Data. The architecture of these layers can be seen as the diagram shown above.
- The data types and models of Big Data:
  - Unstructured Data: refers to data that not easily presented in the form of database or 2D logical charts, including all the formats of office files, text, pictures, XML, HTML, charts, images, music, and video.
  - Semi-structured Data: refers to the data that sits between structured data (such as relational database or object oriented database data) and unstructured data (such as sound and images). HTML files can be considered as a semi-structured data, as it is self descriptive, as the content and structure of the data are mixed together without clear distinction.
  - Structured Data: refers to data that can be stored within databases and can be represented or visualized in the form of 2D logical chart.

Huawei Training & Certification



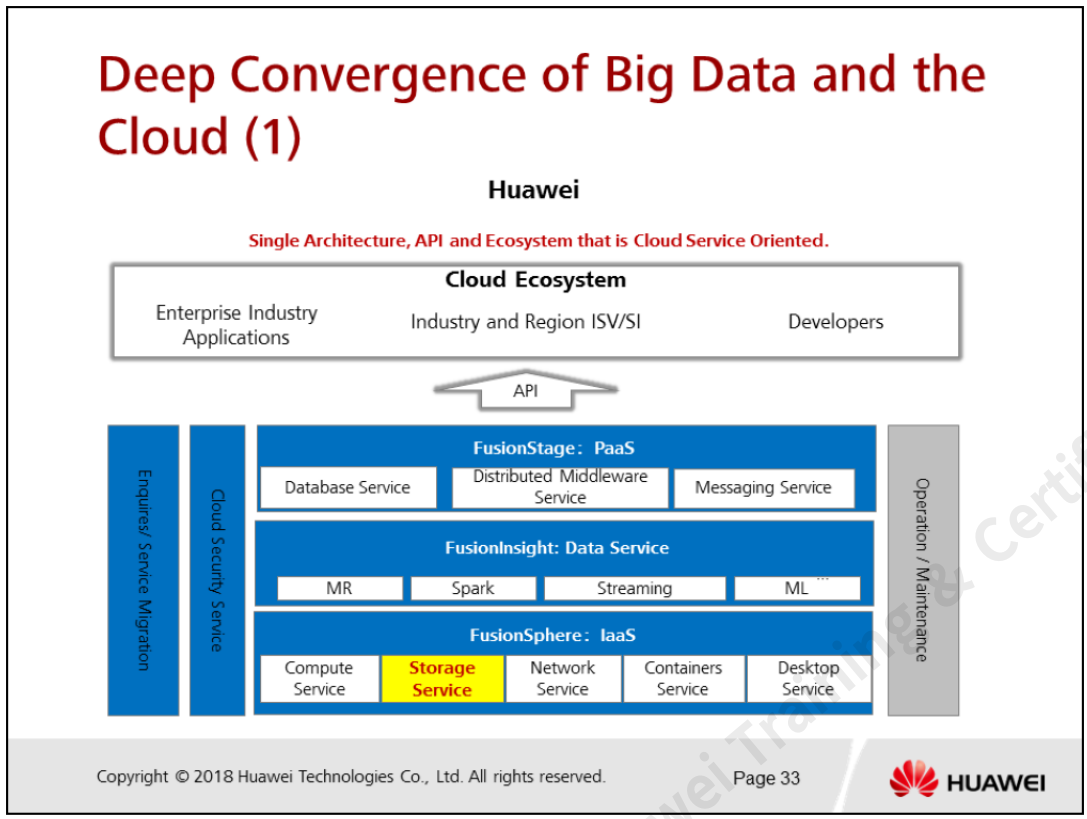
- The process of processing big data is basically the same as the process of general data processing, which includes the four process of data collection, data storage, data management and data analysis.
- The data collection process refers to collecting the required data for subsequent analysis through various sources, methods and tools. In Big Data scenarios, what kind of data needs to be collected? By what means should those data be collected? What methods should be used to collect those data? Which tool should be used for more efficient data collection? These are some of the questions that needs to be put under considerations during this process.
- The data storage process refers to the transmission and storage of data that was collected in the earlier process. With the exponential growth of data, traditional storage systems are no longer able to meet the requirements of Big Data storage. Thus, new technologies are needed to achieve the storage requirements of Big Data.
- The data management process can be considered as an extension to the data storage process. The data is further processed from the basis of the stored data, to realize much finer data segmentation and provides directly available metadata that improves the data analysis efficiency for subsequent data analysis process.
- The data analysis process refers mainly to the utilization of analysis methods, models, and tools to analyze the data and getting a final analysis result that satisfies the common analytical needs. Deeper mining based on the data that was queried and analyzed earlier, allows us to meet higher level of data analysis requirements.

Huawei Training & Certification



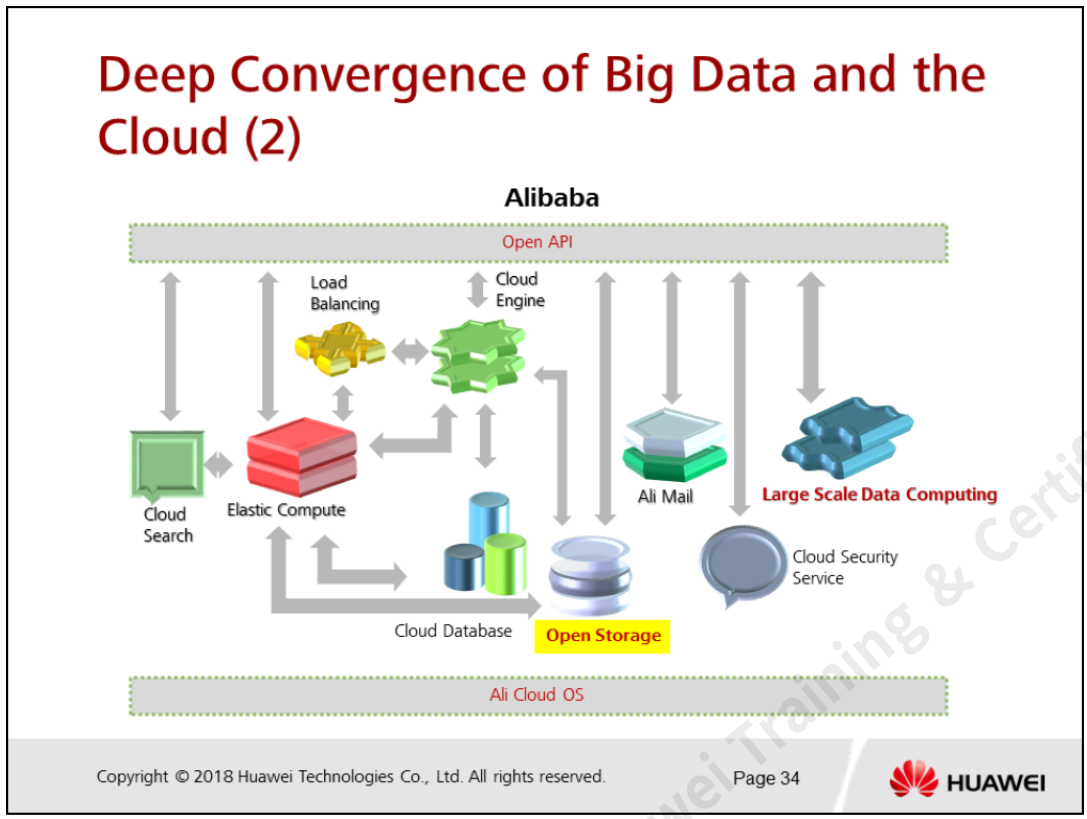
- Era of Big Data:
  - Compute: Developing towards the direction of clustering.
  - Storage: Both block and file storage are developing towards horizontal expansion, virtualization and rich software interfaces for external connections. However, file storage devices has higher horizontal expansion capabilities, and its hardware devices usually expands up to hundreds of nodes. File system also develops from local file system towards clustered file system and distributed file systems.
  - Network: No matter if it is the connection within the Internet or within internal networks, it is developing towards a direction of higher speeds, lower protocol overhead and more efficiency.
  - Database: The development trends moves from relational database towards distributed database and non-relational database (such as In Memory Database).
- The changes in these technologies brought upon new points of opportunities:
  - Database Revolution: Relational Database → Non-relational Database → Hybrid Database.
  - File System Revolution: Local File System → Clustered File System → Distributed File System.

Huawei Training & Certification



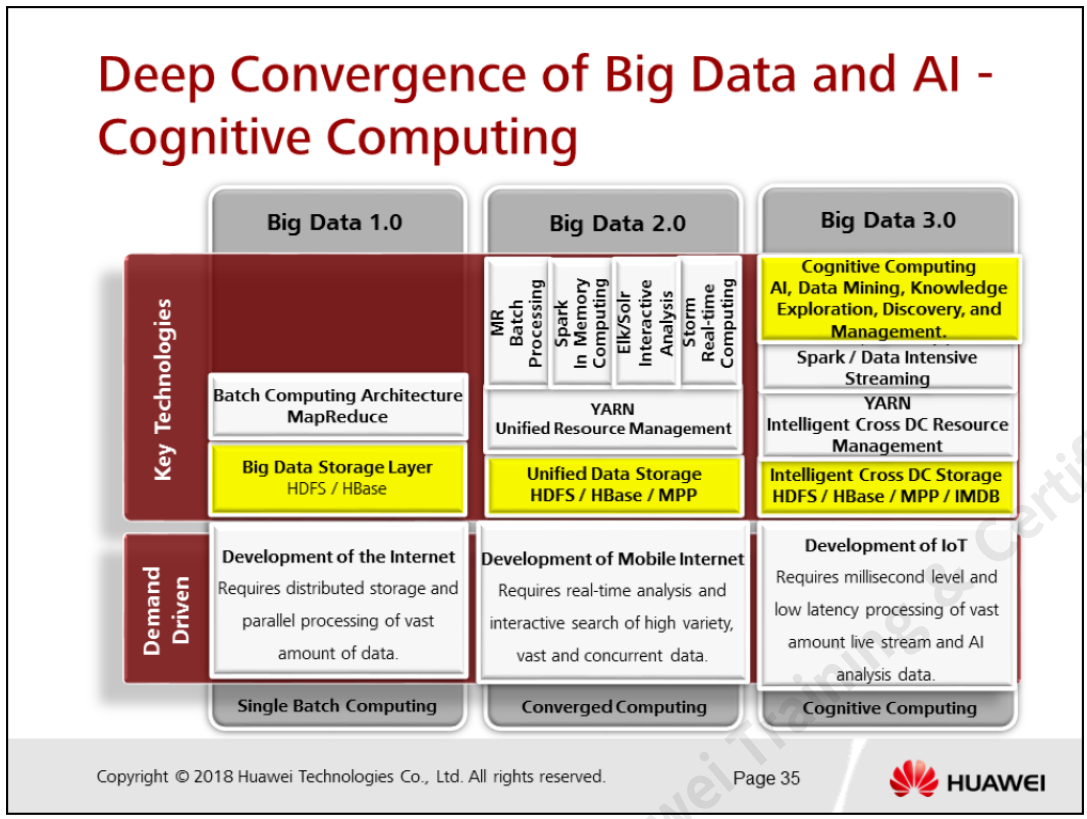
- Huawei focuses on its identity as the technology provider in the Big Data industry, where they provide the physical infrastructure, hardware devices, cloud platform and big data analysis platforms for enterprises in the market looking to build their own Big Data systems.
- The deep convergence of the cloud and big data can be further explained as follows:
- IaaS which is a part of the cloud provides the data storage and analysis services for Big Data, and this layer mainly corresponds and faces the personnel with the role of data administrators. Data administrators manages the vast amounts of data in terms of actual storage and data processing for Big Data.
- PaaS which is also a part of the cloud provides the development services for Big Data, where this layer mainly corresponds and faces the personnel with the role of developers. The developers create applications or software systems to utilize, process and visualize the data to form meaningful insights and value.
- Industry solutions, Big Data storage, Big Data analysis, and Big Data Applications are often represented in the form of SaaS, hence, Big Data can also be considered as one of the basic services offered in the cloud.

Huawei Training & Certification



- Big Data requires a certain technical service level to be fully utilized in a correct manner, but this issue is solved and covered by providing products and solution trainings, Huawei Big Data certifications, industry talent training programs and industry recognized certifications.
- Big data processing is inseparable from Cloud technology. Cloud provides a flexible and scalable infrastructure supporting environment for big data and is an efficient mode of data services. On the other hand, Big data provides new business value for Cloud. In overall, emerging computing trends such as Cloud, the Internet of Things, and the Mobile Internet are both areas that generate vast amounts of data which requires the analytical processing of Big Data.
- ISV refers to Independent Software Vendors, while SI refers to System Integrators where both of them are important partners that assist in building a cloud data center for enterprises and integrate all their existing business systems and even provide new systems or application that helps in the business operation of the enterprises.

Huawei Training & Certification



- AI refers to Artificial Intelligence.
- MapReduce (MR) is a programming model and an associated implementation for processing and generating big data sets with a parallel, distributed algorithm on a cluster.
- Spark is the leading platform for large-scale SQL, batch processing, stream processing, and machine learning.
- "ELK" is the acronym for three open source projects: Elasticsearch, Logstash, and Kibana. Elasticsearch is a search and analytics engine. Logstash is a server-side data processing pipeline that ingests data from multiple sources simultaneously, transforms it, and then sends it to a "stash" like Elasticsearch. Kibana lets users visualize data with charts and graphs in Elasticsearch.
- Solr is an open source enterprise search platform. Its major features include full-text search, hit highlighting, faceted search, real-time indexing, dynamic clustering, database integration, NoSQL features and rich document (e.g., Word, PDF) handling. Providing distributed search and index replication, Solr is designed for scalability and fault tolerance. Solr is widely used for enterprise search and analytics use cases and has an active development community and regular releases.

Huawei Training & Certification

Summary

- This module mainly introduced:
  - The Development Trends of ICT Architecture.
  - Concepts of the Cloud, Big Data and AI.
  - Storage Technologies and Its Application in the Cloud.
  - Storage Technologies and Its Application in AI and Big Data.

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

Huawei Training & Certification
Huawei Training & Certification

Quiz

1. Which of the following is the types of data in Big Data?
   A. Structured Data
   B. Unstructured Data
   C. Semi-structured Data
   D. Small Data
2. Which of the following is the feature of Storage as a Resource?
   A. High Performance
   B. High Reliability
   C. Openness

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

- Answers:
  - ABC.
  - ABC.

Huawei Training & Certification
Huawei Training & Certification

Thank You

www.huawei.com

Huawei Training & Certification
Huawei Training & Certification

03 ICT Architecture

www.huawei.com

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

Huawei Training & Certification

Huawei Training & Certification

Foreword

- This module introduces the technologies and components related to data storage in the ICT infrastructure. The main goal of enterprises is to provide services to customers, and data provides the support required for enterprises to achieve this goal.
- In this module, we shall first try to understand why enterprises need data, and then learn how these data are generated and stored.

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

Huawei Training & Certification
Huawei Training & Certification

Objectives

- Upon completion of this module, you will be able to:
  - Describe the differences between data and information.
  - Understand what is information lifecycle management.
  - Understand the definition of ICT, its architecture and knowledge of its components.

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

Huawei Training & Certification
Huawei Training & Certification

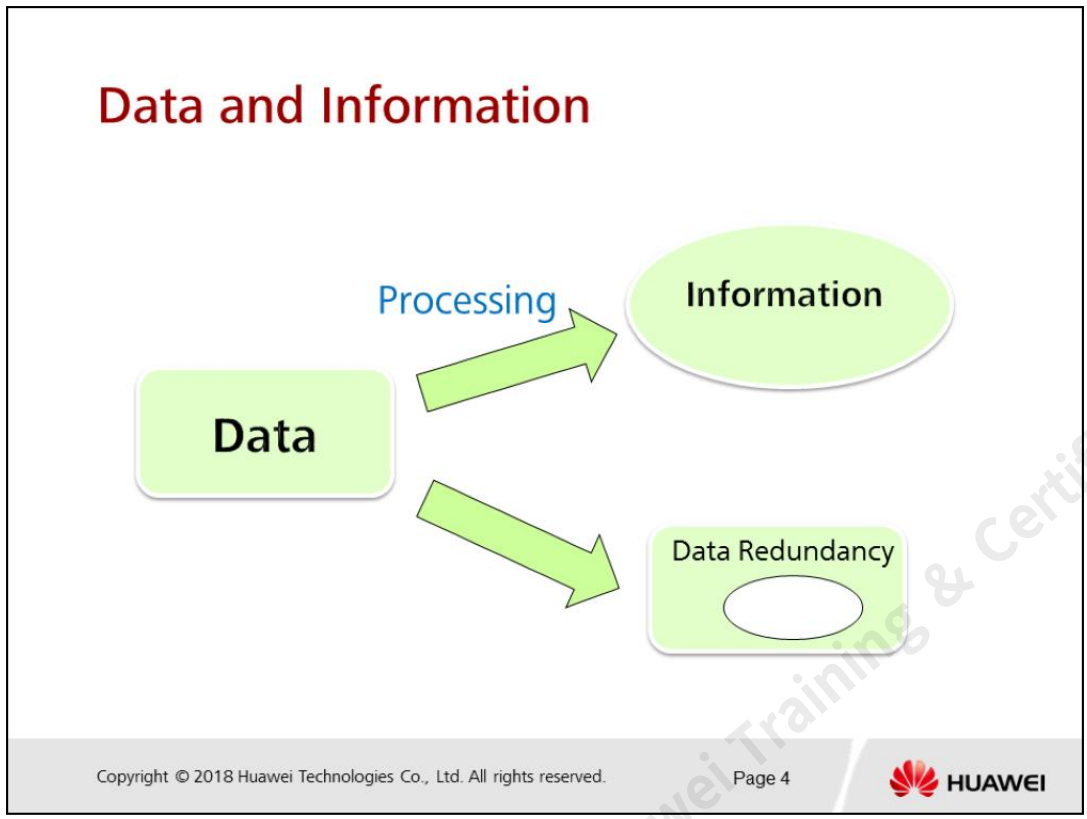
Contents

1. **Data and Information Management.**
2. The New ICT Era.
3. What Is ICT ?
4. Basic Architecture of ICT.
5. Components of ICT Infrastructure.

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

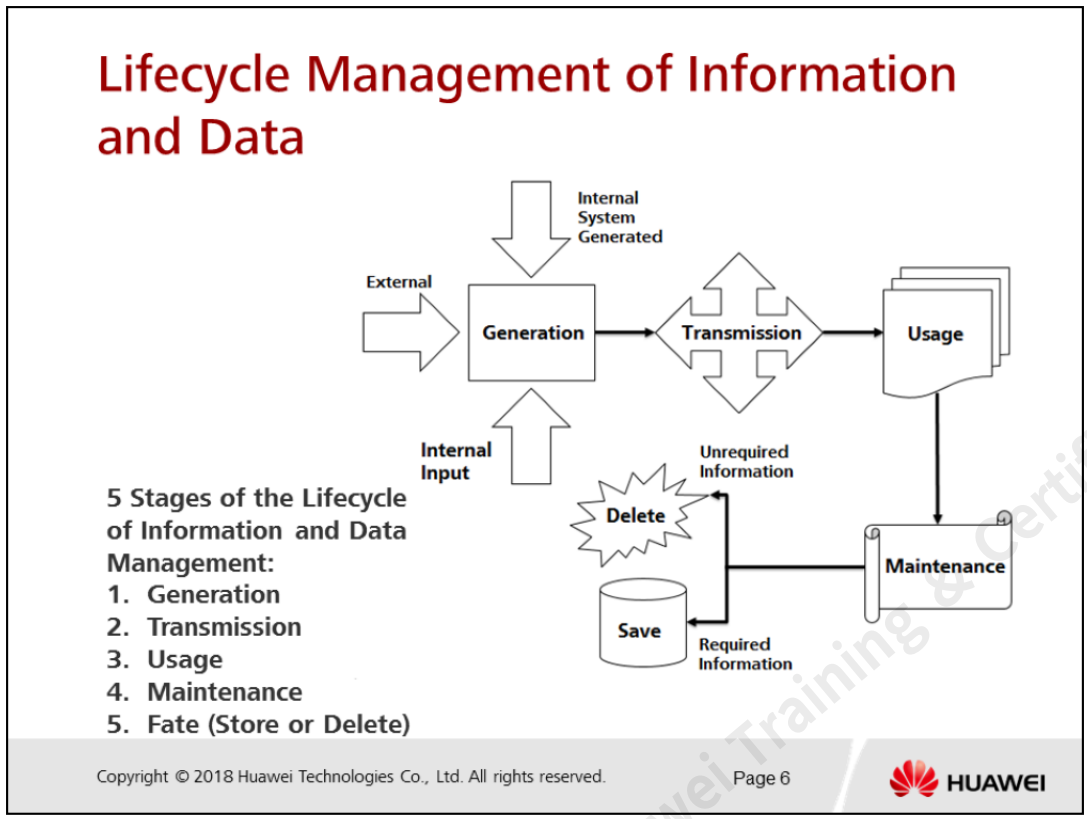
HUAWEI

Huawei Training & Certification
Huawei Training & Certification



- Data and information are interrelated. Data is the record that reflects the attributes of objective things and is the physical manifestation of information. Data becomes information after it has undergone processing, and information needs to be digitized into data before it can be stored or transmitted.
- So, does this mean that data refers to information that can be stored and transmitted? Not necessarily! Data and information has their differences. For the perspective of information theory, data that describes the source of information is the sum of the information and data redundancy, namely Data= Information + Data Redundancy.
- There is a subtle difference between data and information. Data are the facts or details from which information is derived. Individual pieces of data are rarely useful alone. For data to become information, data needs to be put into context. Data are simply facts or figures or bits of information, but not the information itself. When data are processed, interpreted, organized, structured or presented so as to make them meaningful or useful, they are called information. Information provides context for data.
- Data is obtained during the data collection process, information is the meaningful context that is derived from the collected data. For an example, a list of dates which is the data is meaningless without a context such as list of holiday dates which makes it into an information that you could understand.

Huawei Training & Certification



- Information lifecycle refers to the lifecycle where data and information is generated, used, undergone maintenance, saved and until it is finally deleted.
- In another meaning: Information lifecycle refers to the six stages of information collection, storage, processing, maintenance, usage and deletion.
- The information generation phase refers to the origin of information from scratch. Information can be created by one or more person within the organization, it can also be accepted from external sources or even generated by the systems that exists internally within the organization. For example, variety of business files and mails, input and output of computer systems, various reports, charts and statistics created by the staffs.
- Transmission phase refers to the stage where once the data is generated, it is transmitted in a certain way that reach the end user either internally within the organization or externally.
- The usage phase refers to the analysis of the data once it reach the end users to help in the commercial, political and ethical decision making.
- The maintenance phase refers to the management of the data, which includes storing, reading, transmit, copying and backup of those data.
- The fate phase refers to the final handling of the data that has already been used. This might be storing of those data such as in the scenario of legal compliance where data needs to be stored and archived for a certain point of time. This might also mean deletion of those data after it has been used and no longer required. For example, data such as personal emails after being read can be either stored or deleted. In fact, period of data storage also depends on the value of the information such as legal value, historical value, emotional value, commercial value, military value and political value. The higher the value of the information often leads to longer storage period.

Huawei Training & Certification

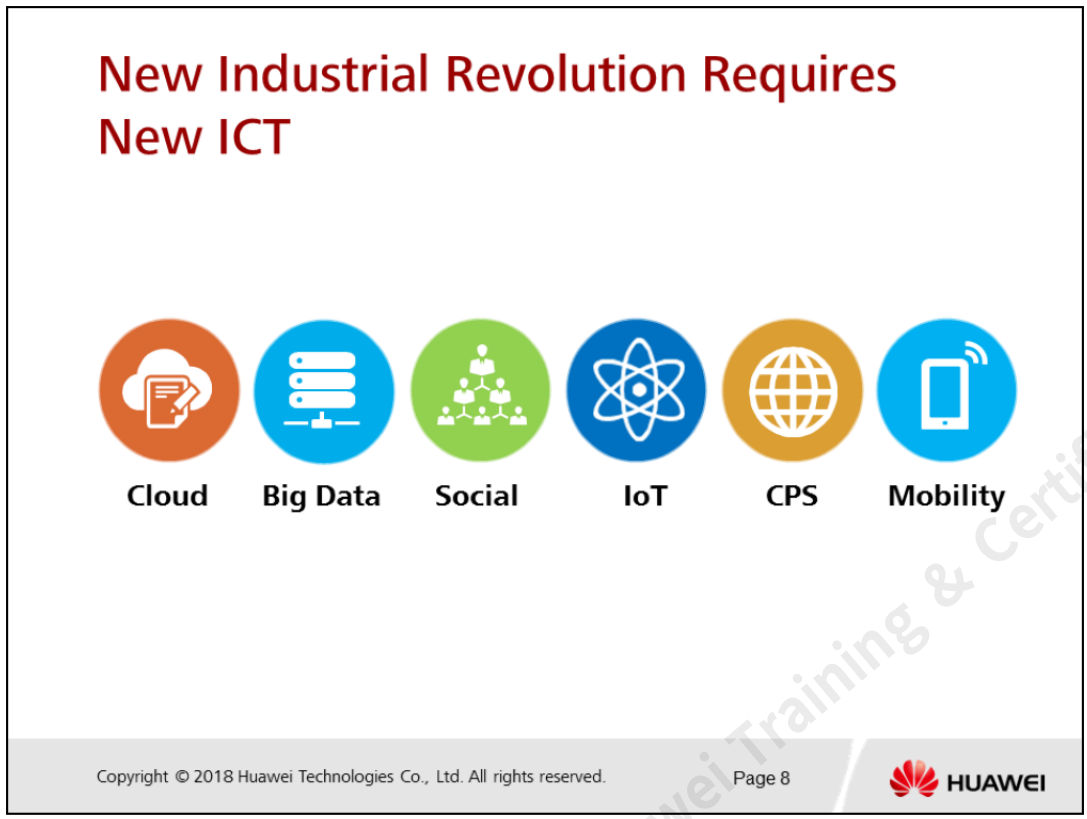
Contents

1. Data and Information Management.
2. **The New ICT Era.**
3. What Is ICT ?
4. Basic Architecture of ICT.
5. Components of ICT Infrastructure.

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

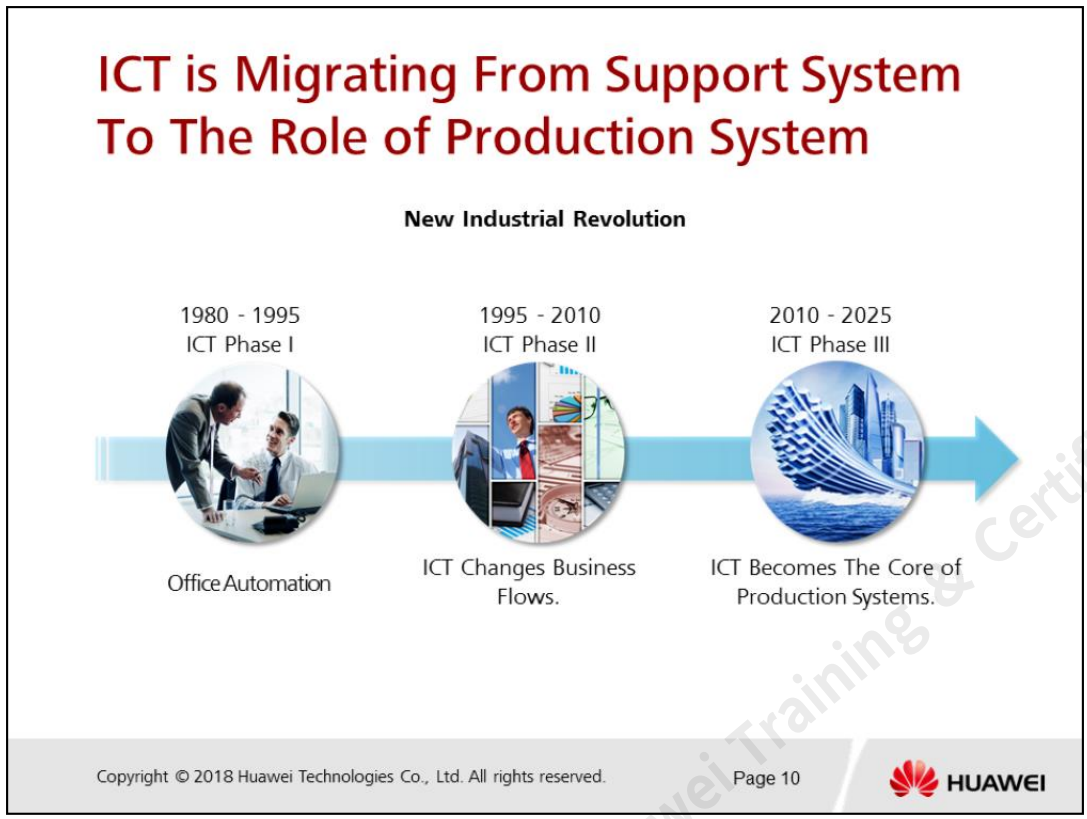
HUAWEI

Huawei Training & Certification
Huawei Training & Certification



- Internet of Things (IoT): It is the interconnection via the Internet of computing devices embedded in everyday objects, which enables them to send and receive data. The internet of things (IoT) is also a computing concept that describes the idea of everyday physical objects being connected to the internet and being able to identify themselves to other devices, and exchange data between them.
- Cloud refers to the addition, usage and delivery model of related internet-based services that typically involves the provision of dynamically scalable and often virtualized resources over the Internet. According to the official NIST definition, "cloud computing is a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications and services) that can be rapidly provisioned and released with minimal management
- Big Data refers to the collection of data that cannot be retrieved, managed and processed by conventional software and tools within a certain period of time. Amount of these data are too big for conventional software and tools to process within a short period of time that makes it useful. Big Data refers to technologies that are able to quickly obtain valuable information from a variety of data types. Among the technologies that are used in Big Data includes MPP(Massively Parallel Processing), data mining grids, distributed file systems, cloud computing platforms, Internet of Things(IoT), and scalable storage systems.

Huawei Training & Certification Huawei Training & Certification



- The above diagram shows the transition of the role of ICT during the new Industrial Revolution. In the first phase back in the 1980s to 1995, the role of ICT was more focused towards office automation and increasing productivity in the office. In the second phase, which is in the period of 1995 to 2010, ICT has gained importance in the daily operation of the businesses that it gradually changed the flow of businesses. It helped the transformation of traditional business to a much modern way of business operations with the application of ICT. Finally, at the third phase between 2010 to present and up to short period of future, ICT has become and will still be the core of production systems in enterprises. ICT has enabled many possibilities for increasing and managing the production process in terms of efficiency and speed, making ICT the core of modern day businesses.

Huawei Training & Certification

Contents

1. Data and Information Management.
2. The New ICT Era.
3. **What Is ICT ?**
4. Basic Architecture of ICT.
5. Components of ICT Infrastructure.

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

Huawei Training & Certification
Huawei Training & Certification

What is ICT ?

- ICT is the abbreviation of Information, Communication and Technology. It is a new concept and technological area formed from the combination of communication technology and information technology.

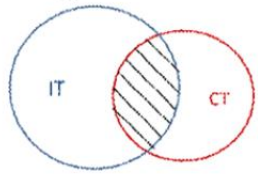


Figure 1: Narrow Definition Theory: ICT is the intersection of IT and CT (ICT=IT∩CT).

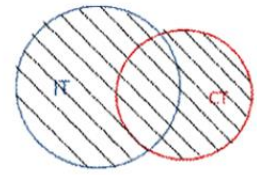


Figure 2: Combination Theory: ICT is the union or combination of IT and CT (ICT=IT∪CT).

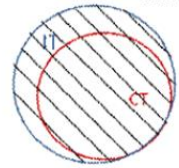


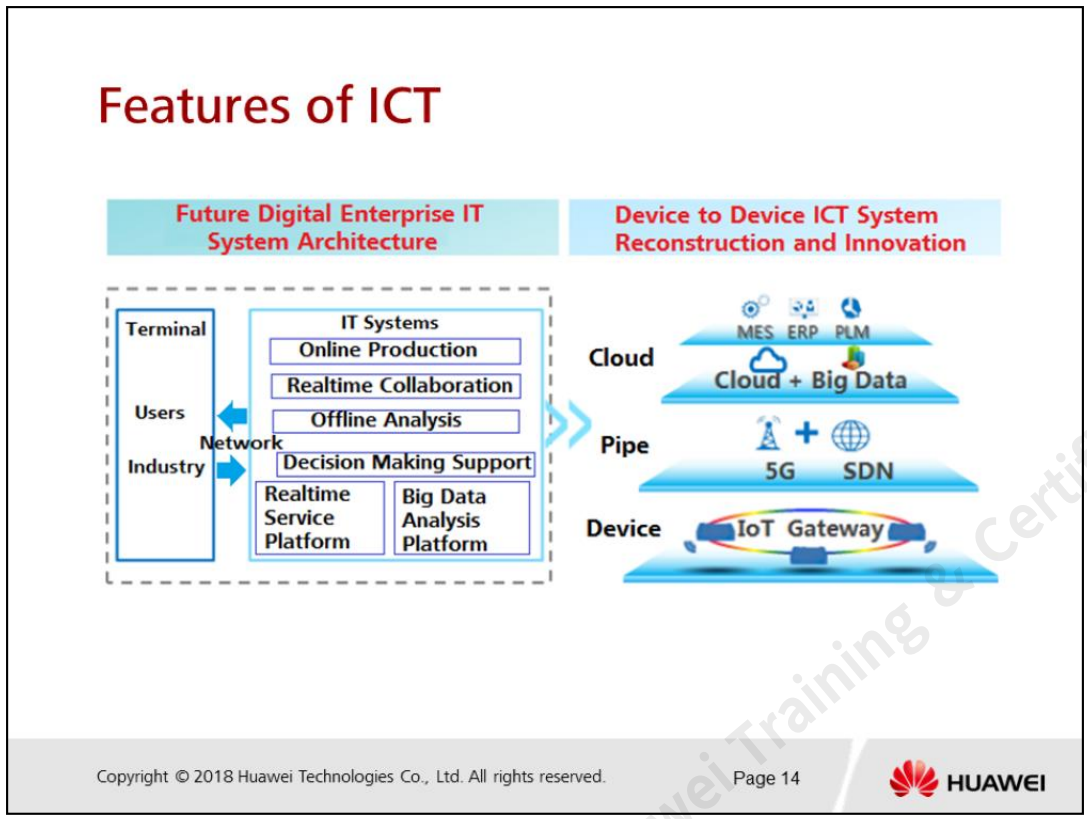
Figure 3: General IT Theory: CT is a part of the wider definition of IT(CT⊆IT,ICT=IT∪CT=IT).

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

- In order to have a better understanding of ICT, we must first look into the meaning of ICT(including all the nature-specific attributes of the concepts and thought models). Let us first define the concepts of IT and CT before we look into the definition of ICT.
- IT (Information Technology), refers to technologies that are related to information. (Huawei 3ms: mainly uses computer science and communication technology to design, develop, install and implement information systems and application systems.) There is varying explanation and definition of information technology based on different people and different books. But one viewpoint in common is that IT has 3 major components which is: sensor technology, communication technology and computing technology.
- CT(Communication Technology), refers to technologies that enables information exchange and information transmission through some kind of behavior or media between human and humans or between humans and nature. Communication technology's main purpose is to serve information, and its main goal is to transmit and exchange variety of information in a high speed, high quality, accurate, timely, secure and reliable manner.
- From the connotation of the concepts mentioned above, we can clearly see that IT technologies consist of basic information collection, information processing, information transmission(communication), and information application. The purpose of communication technologies is also to serve IT.

Huawei Training & Certification - Huawei Training & Certification



- One of the features of ICT is that it is cross domain and convergence. ICT uses CT as the basic infrastructure and presents the services provided by IT in a direct and relatable content to the customers such as remote learning, remote healthcare, IoT, Internet of Vehicles, and remote maintenance of elevator industries etc.
- Production systems of customers in the industry are deeply converging with ICT, Cloud, IoT technologies that are compatible with their industrial applications.
- The application of ICT has been the fundamental of business thinking, it focuses on the customers as the core of the business, bringing together new customer experiences in ways that was never possible before to help rebuild industries in education, logistics, and banking.
- ICT enterprises not only just focus on their own resources, business models, competitive advantage, but also focus on their positioning in the business ecosystem and their relationship with partners to complement each other in terms of resources, cross border innovation, and mutual development.
- ICT industry in a new industry that has continuous integration of technologies, which has become a catalyst and enabler for the transformation of businesses and industries.

Huawei Training & Certification

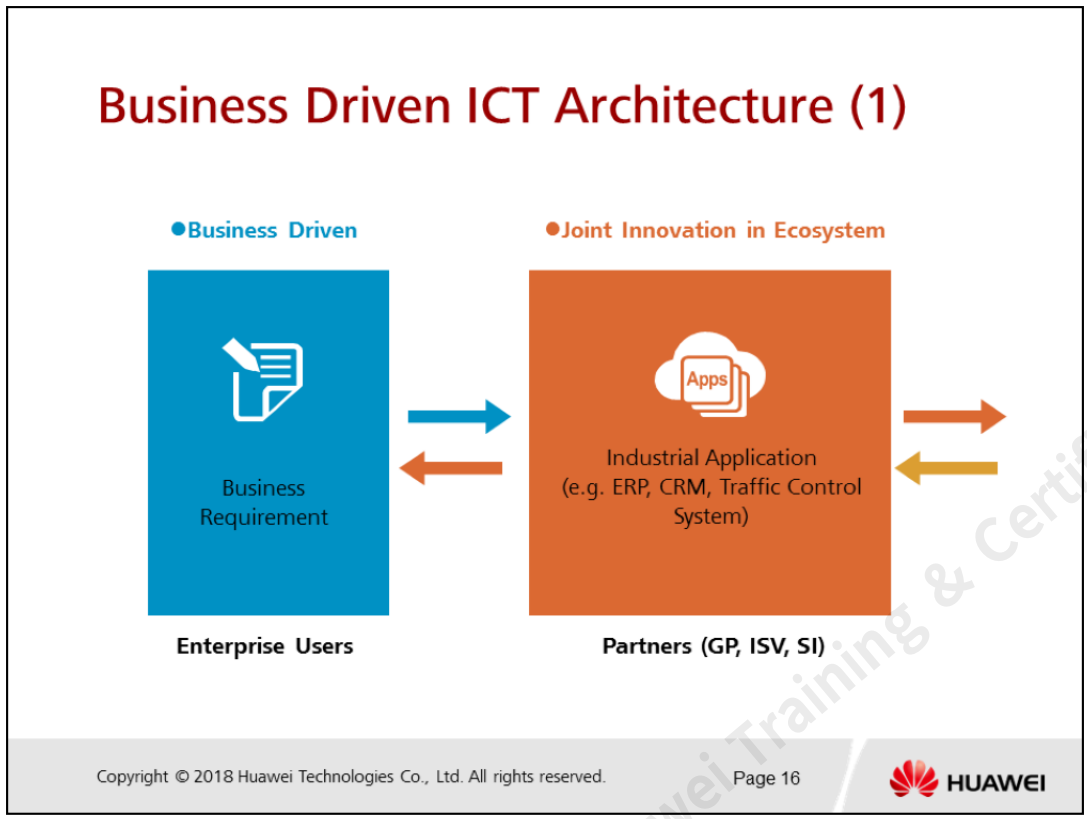
Contents

1. Data and Information Management.
2. The New ICT Era.
3. What Is ICT ?
4. **Basic Architecture of ICT.**
5. Components of ICT Infrastructure.

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

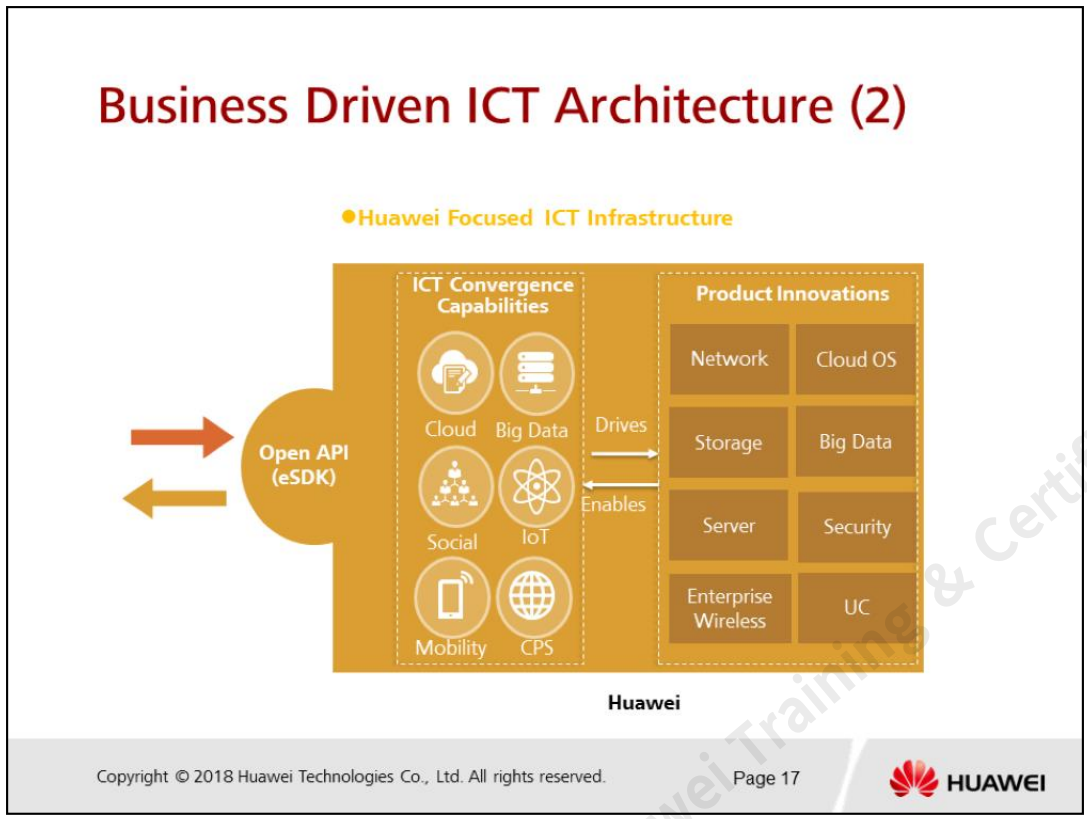
HUAWEI

Huawei Training & Certification
Huawei Training & Certification



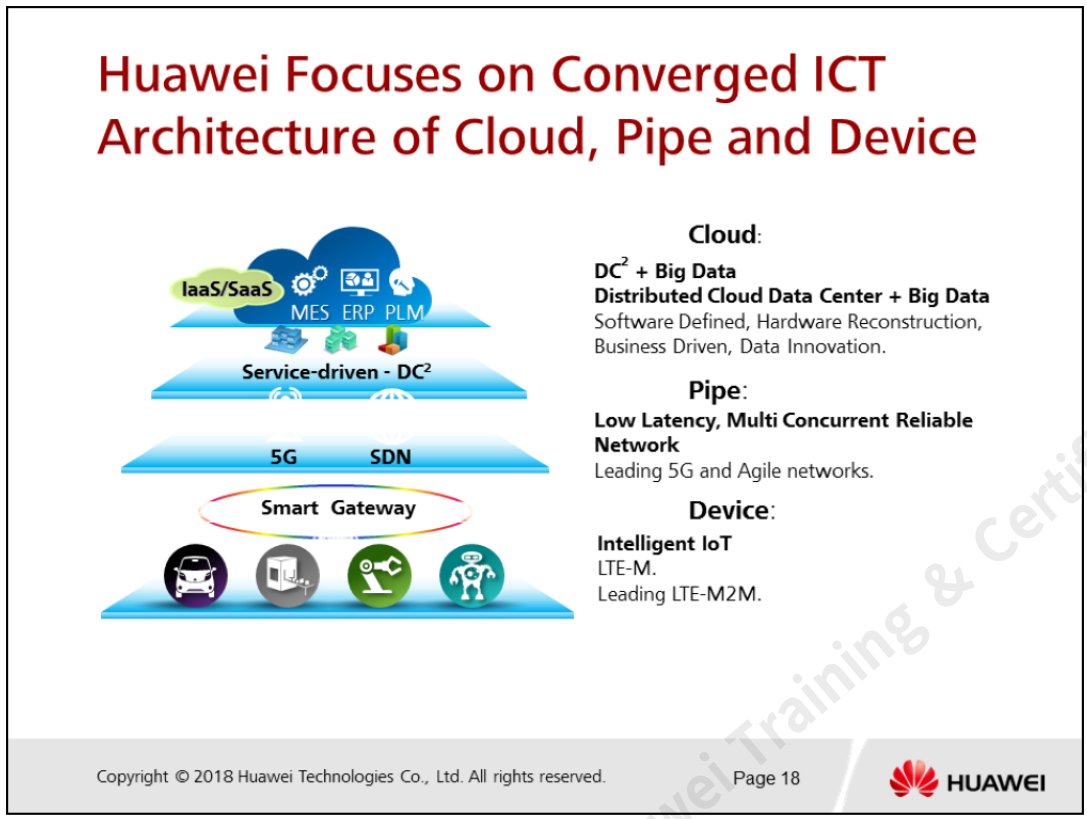
- BDII： Business-Driven ICT Infrastructure, which is an ICT infrastructure that is driven by the business demands and requirements.
- ERP： Enterprise Resource Planning, which was originally proposed by Gartner Group Inc. in the United States in 1990 that was first defined as an application software but quickly accepted by commercial enterprises around the world and has now become one of the modern theories of business management.
- CRM： Customer Relationship Management refers to technologies that enterprises use to manage and analyze customer interactions and data throughout the customer lifecycle, with the goal of improving customer service relationships and assisting in customer retention and driving sales growth.
- GP： Global Partner refers to global business partners.
- ISV： Independent Software Vendors is an individual or business that builds, develops and sells consumer or enterprise software.
- SI： Service Integrators or System Integrators is an individual or business that builds computing systems for clients by combining hardware, software, networking and storage products from multiple vendors.

Huawei Training & Certification



- Unified communications (UC) is a framework for integrating various asynchronous and real-time communication tools, with the goal of enhancing business communication, collaboration and productivity. Unified communications does not represent a singular technology; rather, it describes an interconnected system of enterprise communication devices and applications.
- Application program interface (API) is a set of routines, protocols, and tools for building software applications. Basically, an API specifies how software components should interact.
- Software Development Kit (SDK) is a programming package that enables a programmer to develop applications for a specific platform. Typically an SDK includes one or more APIs, programming tools, and documentation.

Huawei Training & Certification



- LTE-M (LTE for Machines), is a IoT technology that evolved from LTE. LTE-M is the abbreviation for LTE Cat-M1 or Long Term Evolution (4G) Category M1. This technology is for Internet of Things devices to connect directly to a 4G network, without a gateway and on batteries. LTE is the generic term for all of 4G.
- LTE-M2M (LTE Machine-to-Machine) refers to the IoT technology where IoT devices are connected and able to communicate from one machine to another machine using LTE network.
- SDN (Software Defined Network) is a concept of network design, or a revolutionary design thought. A network can be considered as SDN if the network devices can be centrally managed, programmable, and with control plane and data plane separation. SDN is not a physical technology or protocol, but it is more of a thought model and framework. The concept of software defined network also gave birth to other software defined technologies such as software defined security and software defined storage etc. In fact, SDN and software defined technologies are a trend that swept across the IT industry.

Huawei Training & Certification

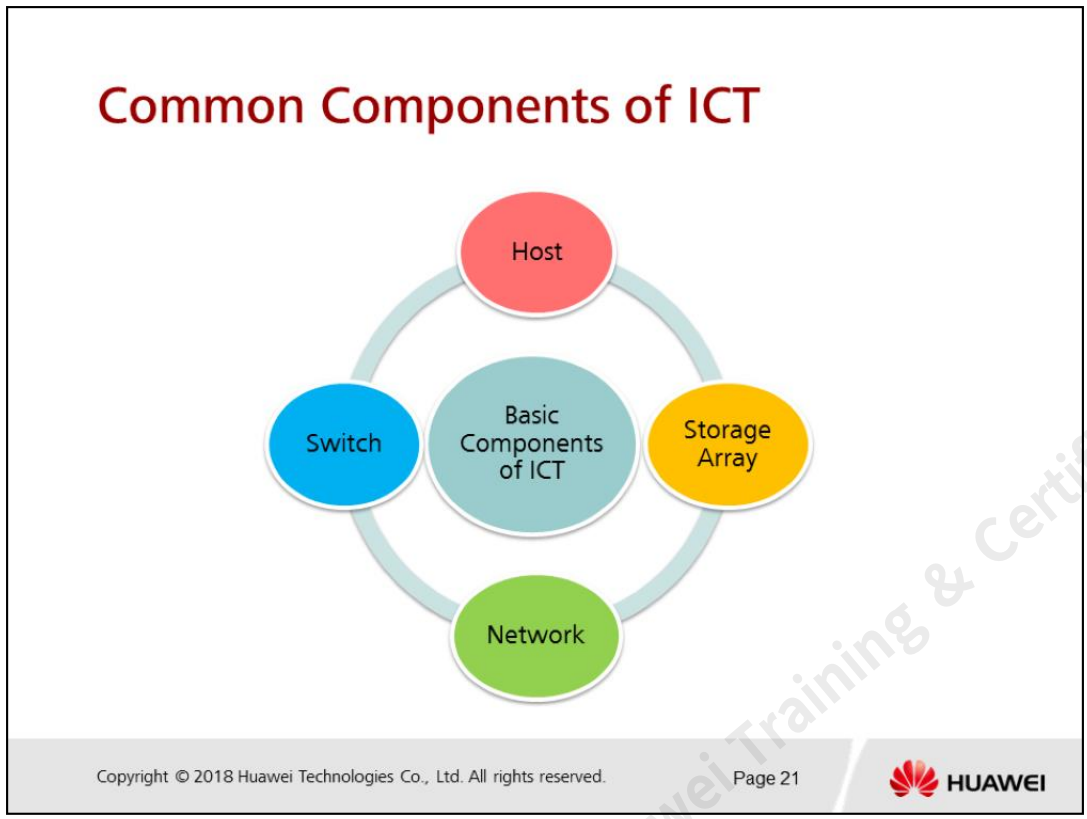
Contents

1. Data and Information Management.
2. The New ICT Era.
3. What Is ICT ?
4. Basic Architecture of ICT.
5. **Components of ICT Infrastructure.**

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

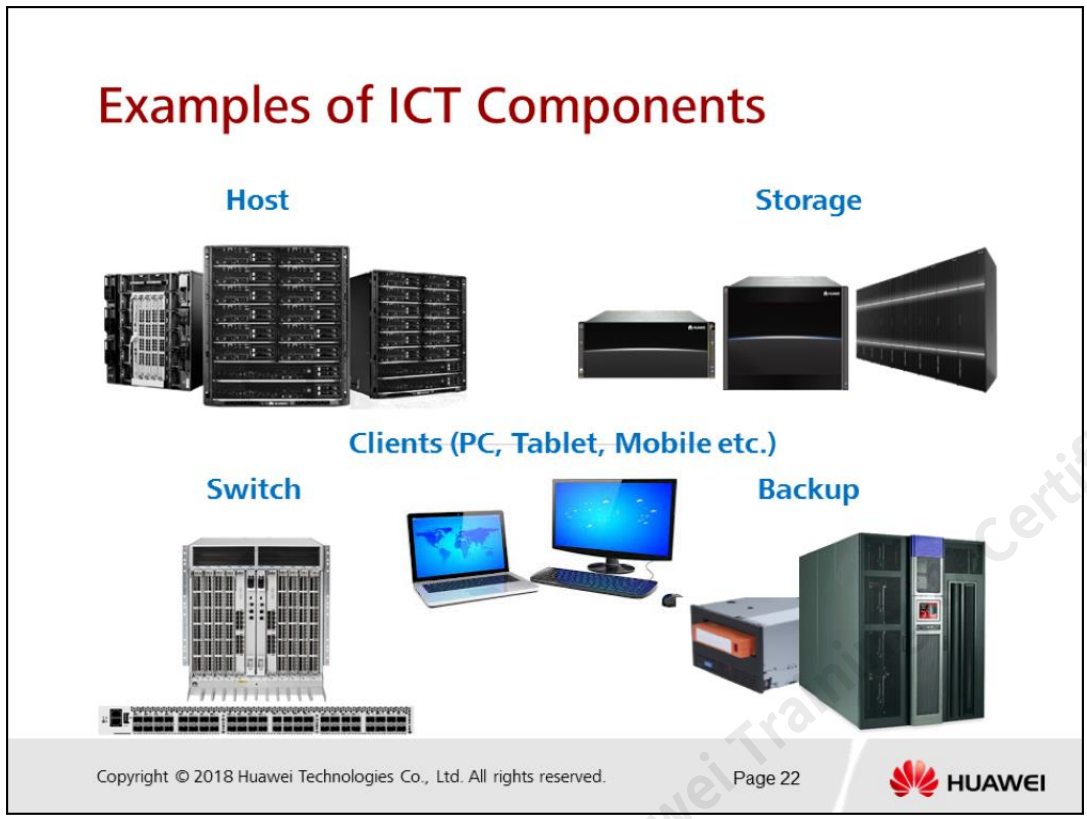
HUAWEI

Huawei Training & Certification
Huawei Training & Certification



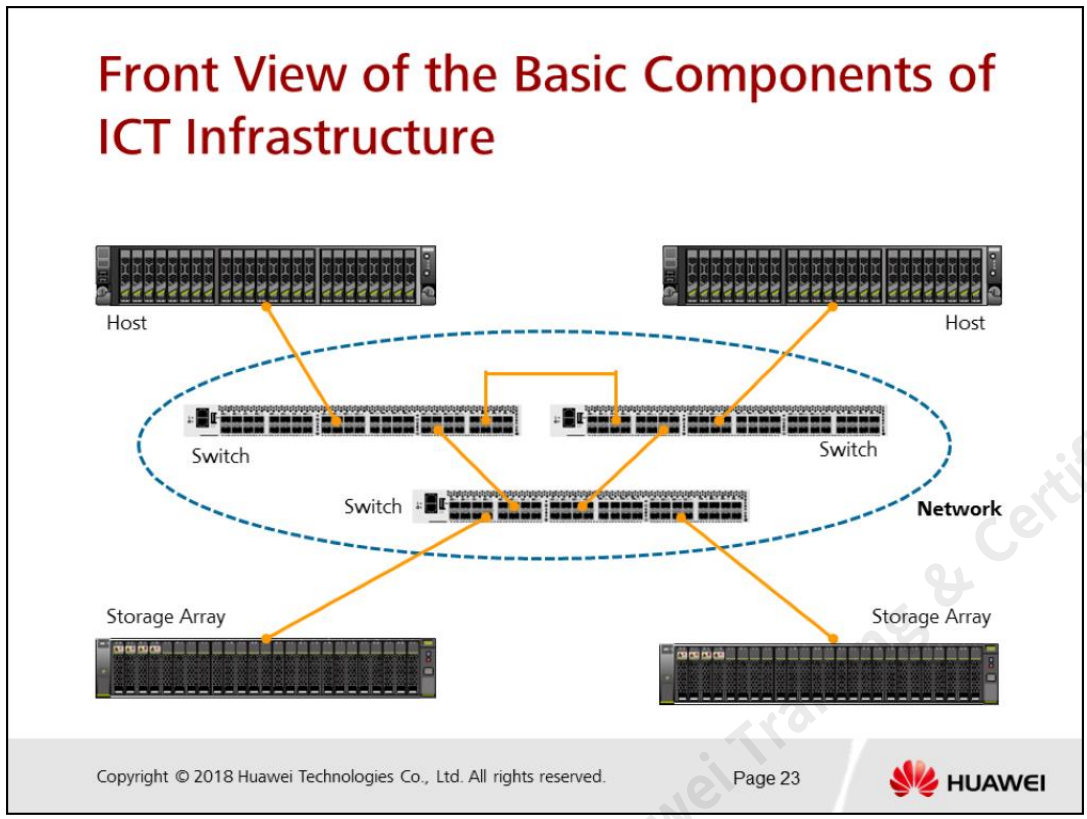
- The following are the common terminologies used in ICT infrastructure:
- Host: A computing system consisting of hard drives or file system that can provide data access and storage. Hosts typically do not include intermediary network devices like switches and routers, which are instead often categorized as nodes.
- Storage Array: A system that has the management software that can provide access to its sets of hard drives or tape libraries for the purpose of data access and storage. It is a hardware element that contains a large group of hard disk drives (HDDs). It may contain several disk drive trays and has an architecture which improves speed and increases data protection. The system is run via a storage controller, which coordinates activity within the unit.
- Network: It provides connection between sets of nodes and allows the connected devices(nodes) to communicate with each other.
- Switch: A network device used to switching network signals or data. It can provide a dedicated communication channel between any 2 nodes in the network that is connected to the switch. It is a networking device that connects devices together on a computer network by using packet switching to receive, process, and forward data to the destination device.

Huawei Training & Certification



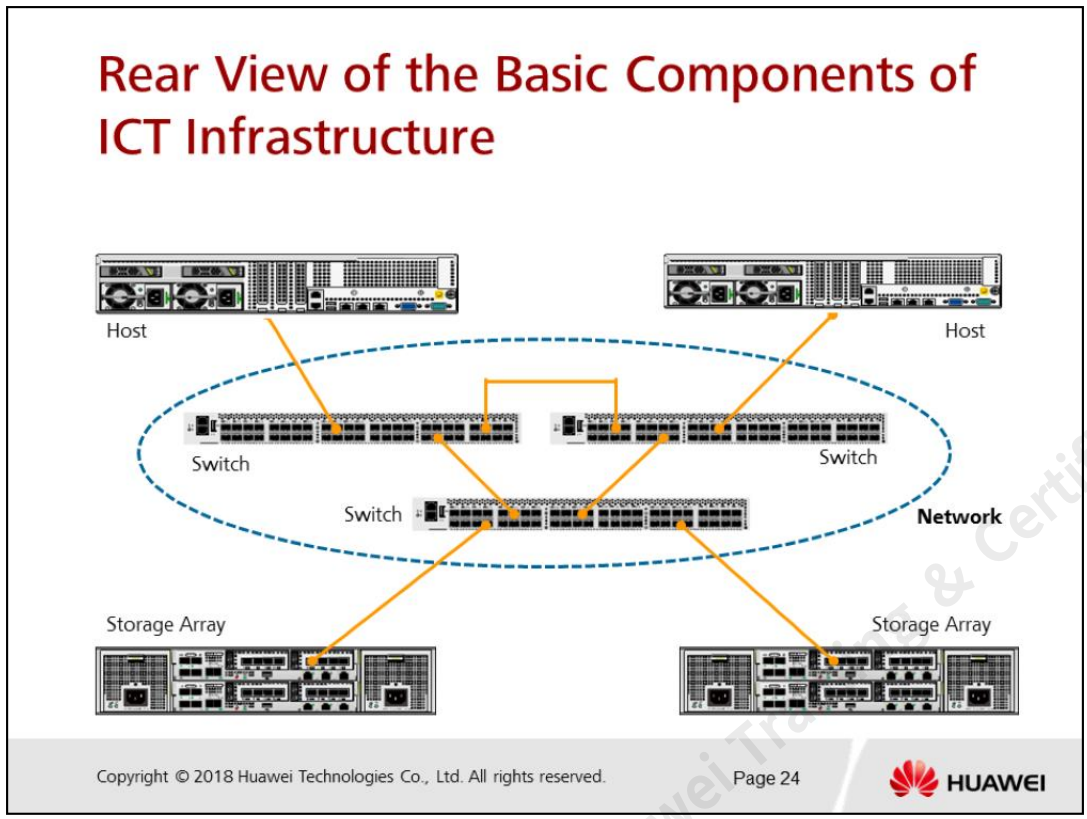
- The term host (sometimes also called servers) are commonly used to describe the computing device that has high specification within an enterprise, these devices usually have the important enterprise software installed within them, and are responsible for running important services. Desktop and laptops are simpler version of hosts. Hosts are commonly required to run 24x7 for many years before they reach their end of life(EOL) and replaced.
- Storage arrays commonly refers to the devices that handles data storage. Some of the storage arrays are quite small just like the size of a host, but some are much bigger and can be configured with thousands of hard drives.
- Switches are the network devices used for network interconnection between the devices in the network. Switches have many ports, and devices connects to each other by connecting through cables to these ports. The devices mentioned here can be hosts, storage arrays or even other switches.

Huawei Training & Certification



- The diagram above shows the front view of the common components of ICT such as hosts, switches and storage arrays connected to each other within the network.

Huawei Training & Certification



- The diagram above shows the rear view of the common components of ICT such as hosts, switches and storage arrays connected to each other within the network.

Huawei Training & Certification

Summary

- This module introduced:
  - Data and Information Management.
  - The New ICT Era.
  - What is ICT ?
  - The Basic Architecture of ICT.
  - Components of ICT Infrastructure.

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

Huawei Training & Certification
Huawei Training & Certification

Quiz

1. (True or False) Data is equal to Information.
2. (MCQ) Which of the following components are commonly used in ICT systems?
   A. Host
   B. Storage Array
   C. Network
   D. Switch

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

- Answers:
  - F.
  - ABCD.

Huawei Training & Certification
Huawei Training & Certification

Thank You

www.huawei.com

Huawei Training & Certification
Huawei Training & Certification

04 Architecture of Storage Systems

www.huawei.com

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.



Huawei Training & Certification
Huawei Training & Certification

Foreword

- This module mainly introduces the common system architecture of storage systems, the components of storage systems and also introduces the Huawei storage products.

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

Huawei Training & Certification
Huawei Training & Certification

Objectives

- Upon completing this module, you will be able to:
  - Understand the common storage system architectures.
  - Grasp the concepts of common storage components.
  - Understand the Huawei storage products.

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

Huawei Training & Certification
Huawei Training & Certification
Huawei Training & Certification

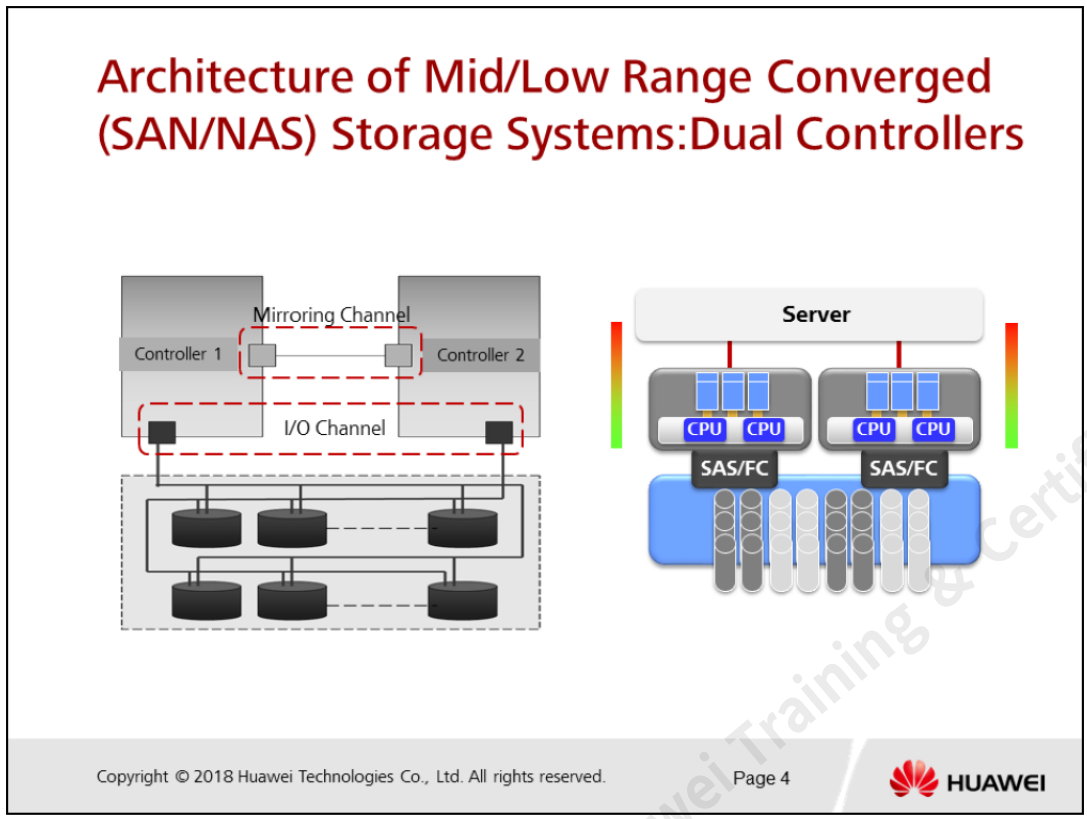
Contents

1. **Architecture of Storage Systems.**
2. Components of Storage Systems.
3. Introduction to Huawei Storage Products.

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

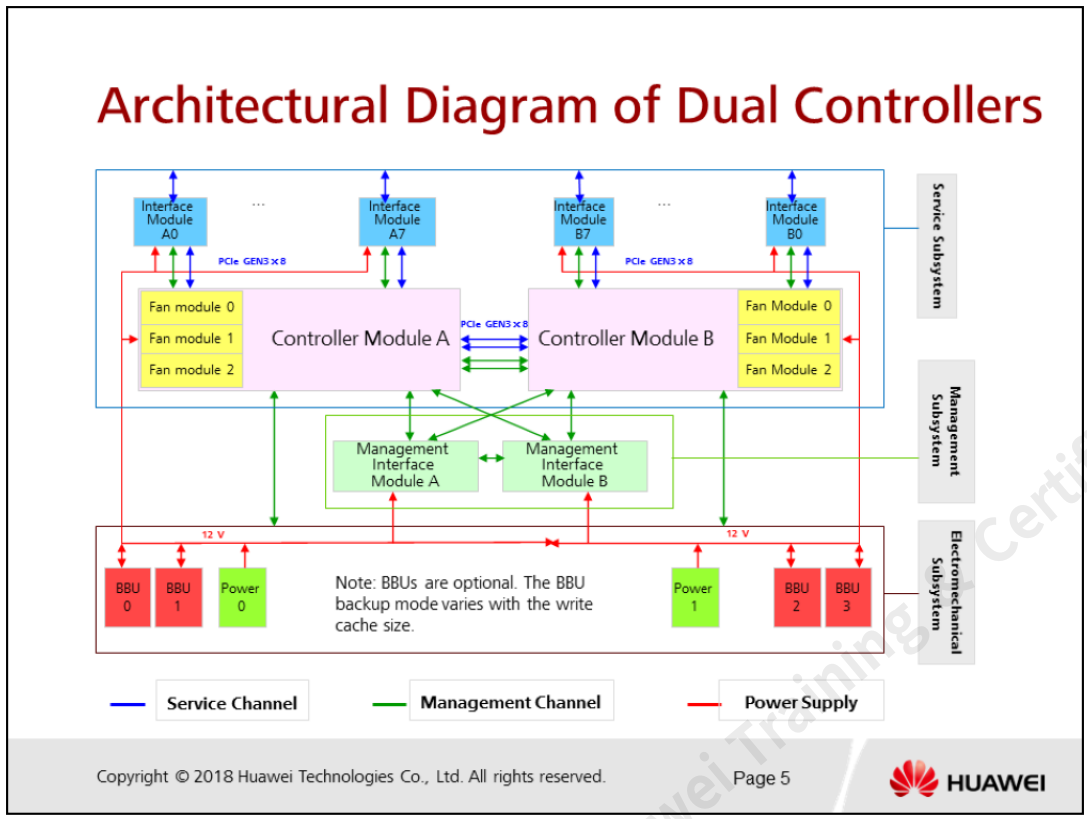
HUAWEI

Huawei Training & Certification
Huawei Training & Certification



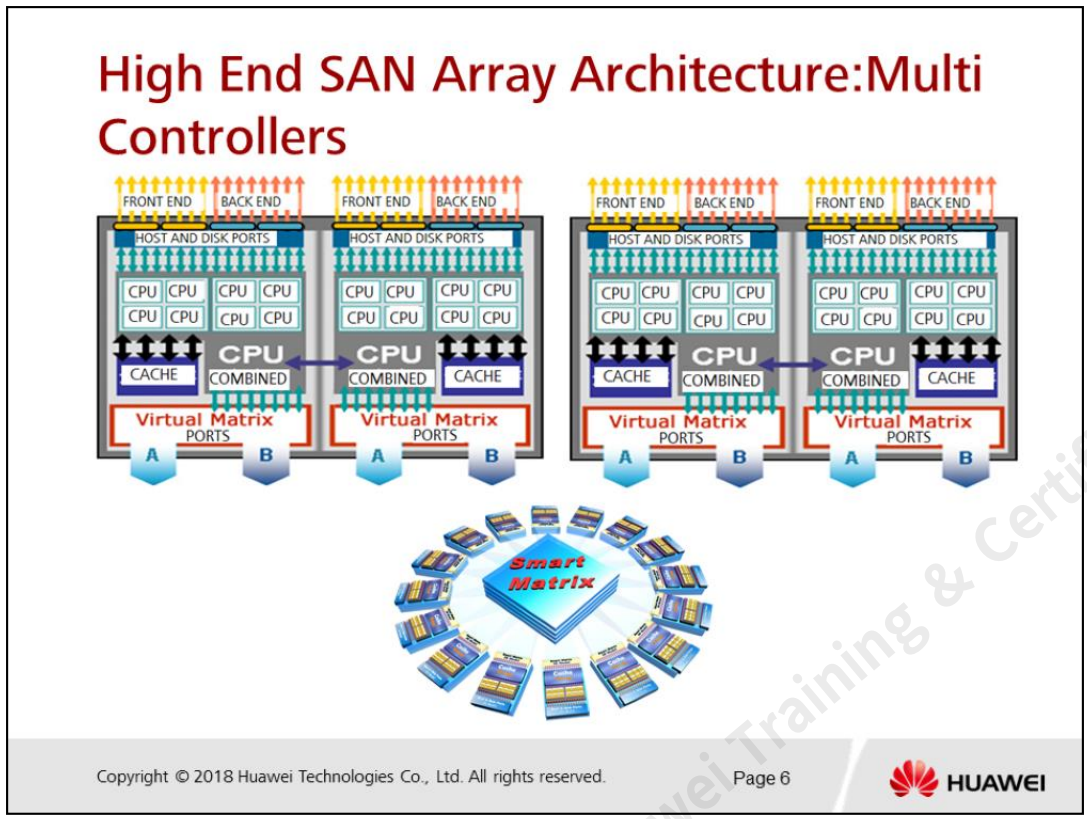
- Key features and benefits of Dual Storage controller system architecture:
  - A copy of all the data written during I/O process is kept in both storage controller's cache to avoid data loss when there is a single controller failure.
  - Each of the storage controller is connected via backend ports to the loop of physical disk enclosures for load balancing of the service workloads and increase the overall system performance.
  - Single or Multiple LUNs I/O service load is load balanced between the two storage controllers to avoid performance bottleneck.
  - Whenever a single controller is faulty, another storage controller can takeover the service workload with the help of the Multipathing software to ensure service continuity.

Huawei Trainings & Certification



- The diagram above shows the interconnection between the storage controllers and other hardware components of the storage systems such as the power and fan modules, management interface modules and the interface modules.
- The diagram also shows the redundant design of the storage controllers and management interface modules. Hence, by having Dual controllers, the single point of failure faced by single controller storage architecture are eliminated.

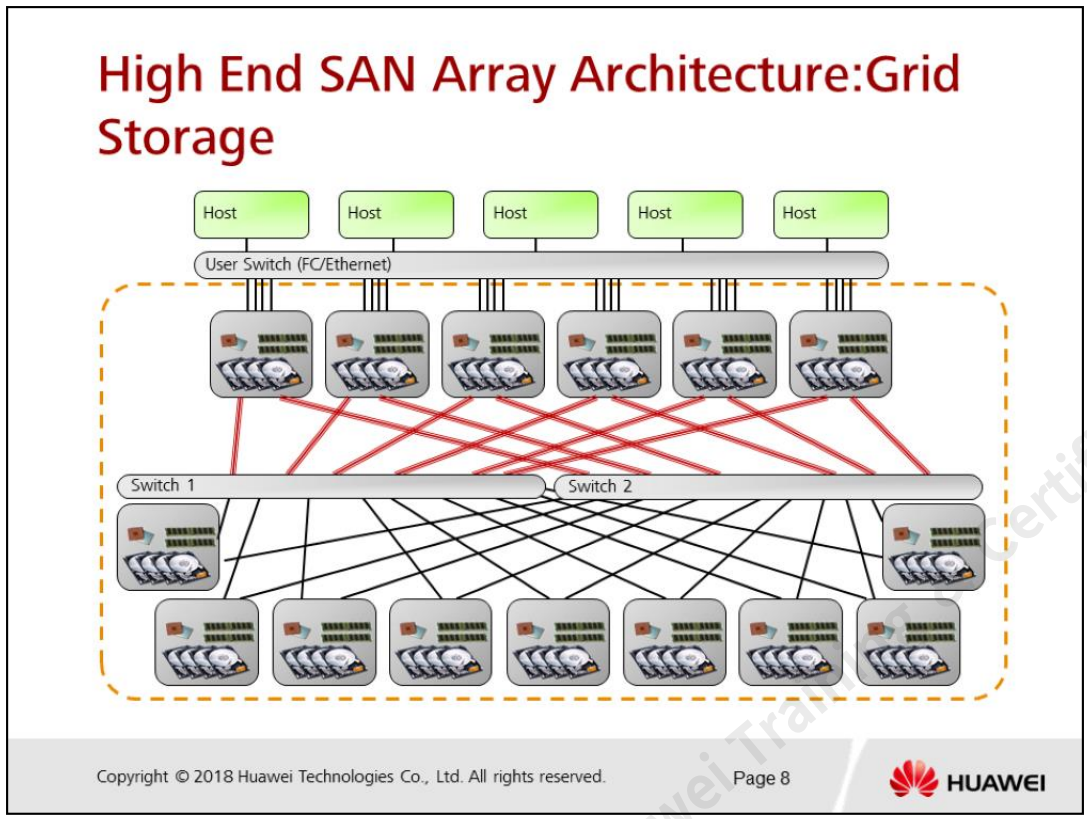
Huawei Training & Certification



Key features of High End Storage Systems:

- Data plane is interconnected using PCIe connections, and the 4 controllers within the storage engine are interconnected using PCIe backplane which allows high bandwidth rate up to 128G/s.
- Storage engines uses full duplex switched network with PCIe optical cables for inter-engine connection up to 128G/s bandwidth rate.
- The data transmission and mirroring uses the internal backplane interconnection channel as prioritized transmission route and reverts to use inter-engine switched network channel if unable to transmit the data between the storage controllers using the backplane.
- The total bandwidth for data transmission (including data transmission within same storage engine and data transmission between different storage engines) is up to 640G/s.
- Management and control plane uses Gigabit Ethernet(GE) technology and the 2 ports of the SVP(Service Processor) effectively creates 2 management planes for redundancy, and all the storage engines and switches can be connected to both of the management planes at the same time. When the GE network is faulty, the management control plane still can use the PCIe channel in the data plane to transmit management and control messages. This ensures that service continuity and management of storage system is not affected during the faulty period of the GE network of the management plane.

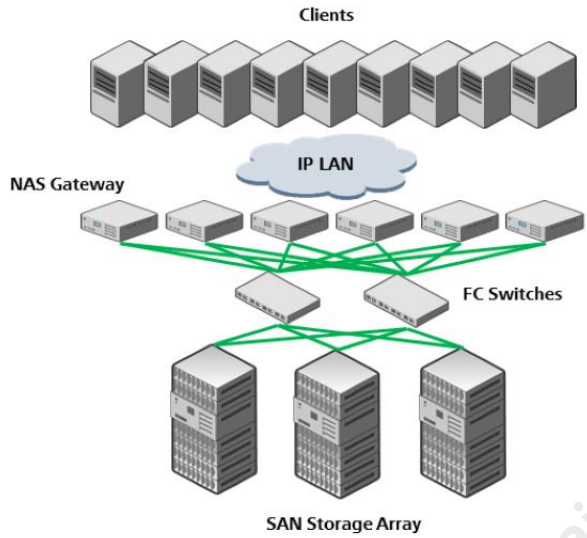
Huawei Training & Certification



- Grid storage is considered as the new generation storage architecture by the industry, it employs the combination of latest large scale parallel distributed grid storage technologies with the Scale-Out(Horizontal Expansion) storage architecture. This architecture allows it to use a multipath grid network between the storage nodes for load balancing and fine-grained data distribution algorithm to ensure that the data is evenly distributed between the storage nodes. This increases the overall system reliability, usability and IO performance, and makes it easily scalable. Once a new storage node is added, the rest of the nodes in the grid will be able to detect the new node and use it for storage making scalability much easier to perform.
- In grid storage, there is no centralized control module as the storage system consists of grids with self-contained storage nodes. These self-contained storage nodes are interconnected in the storage grid and able to communicate to each other without a centralized control module, introducing a new level of fault tolerance and redundancy into storage system architecture. Each grid is also known as data modules, and each independent data modules has their own loosely coupled CPU processing capabilities, caching and disk storage capabilities.

Huawei Trainings & Certification - Huawei Training & Certification

Clustered NAS Storage Architecture

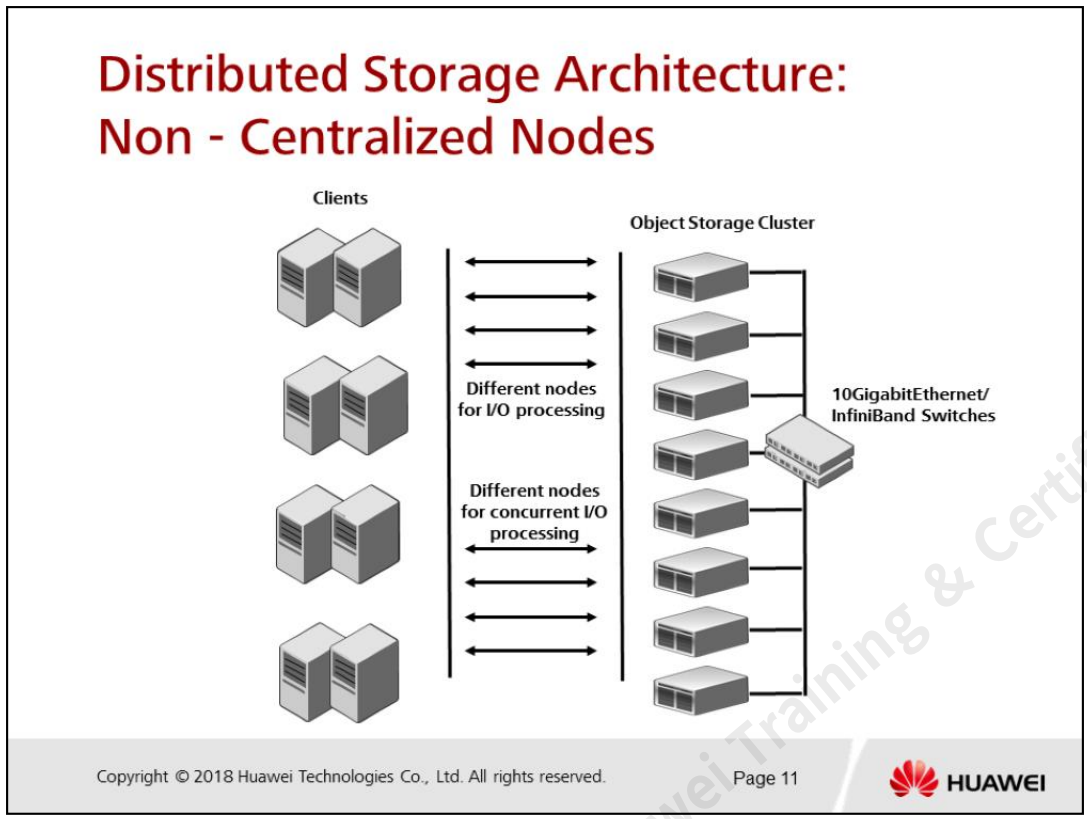


Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

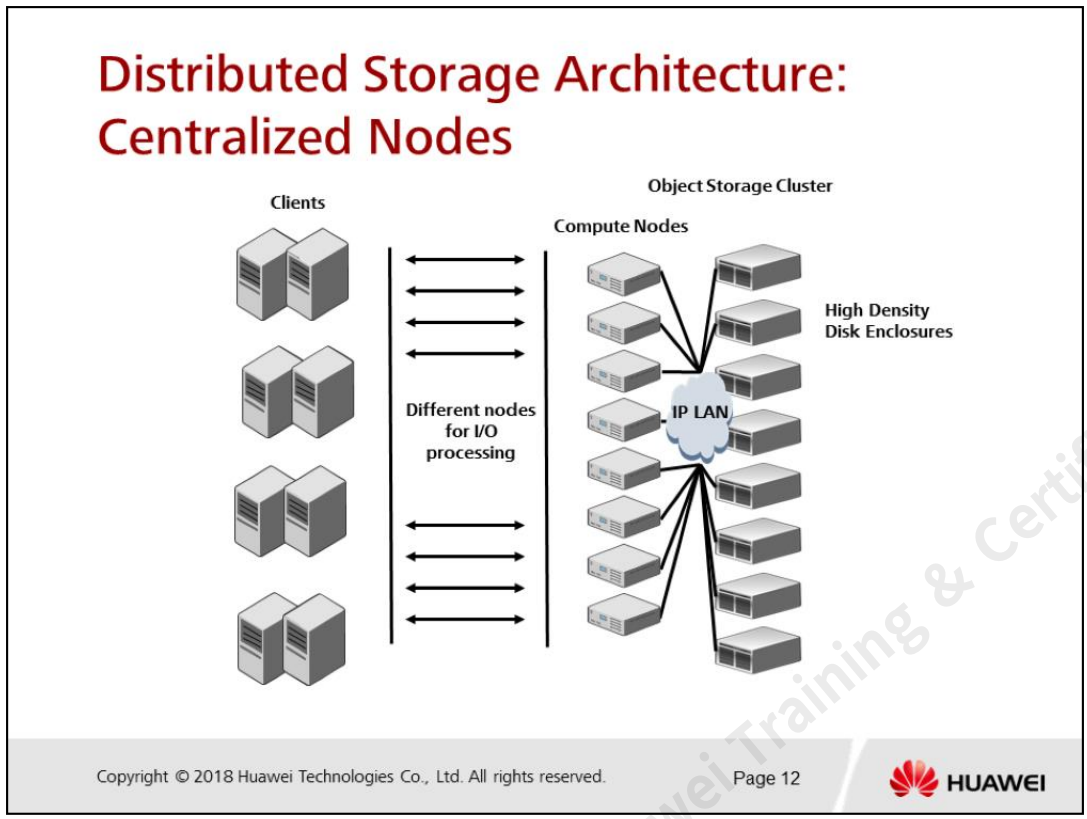
- Unified Global Namespace and Global Sharing.
  - A global namespace provides a consolidated view of multiple file systems or file servers in different location allowing data to be accessed without knowing where the data physically resides which is particularly beneficial for distributed implementation of storage. Without global namespace, each file system needs to be managed separately.
  - Global sharing allows the data to be shared globally and accessed from multiple locations making better data accessibility for the storage system.
- High Performance and Concurrent Multi Host IO Processing Array.
  - Clustered NAS allows IO process of multiple host to be processed in concurrent which in turn increases the overall performance of the storage systems.
- Online Expansion.
  - Expansion of the storage space and modules can be done without taking the whole storage system offline which eliminates downtime during storage expansions.

Huawei Training & Certification



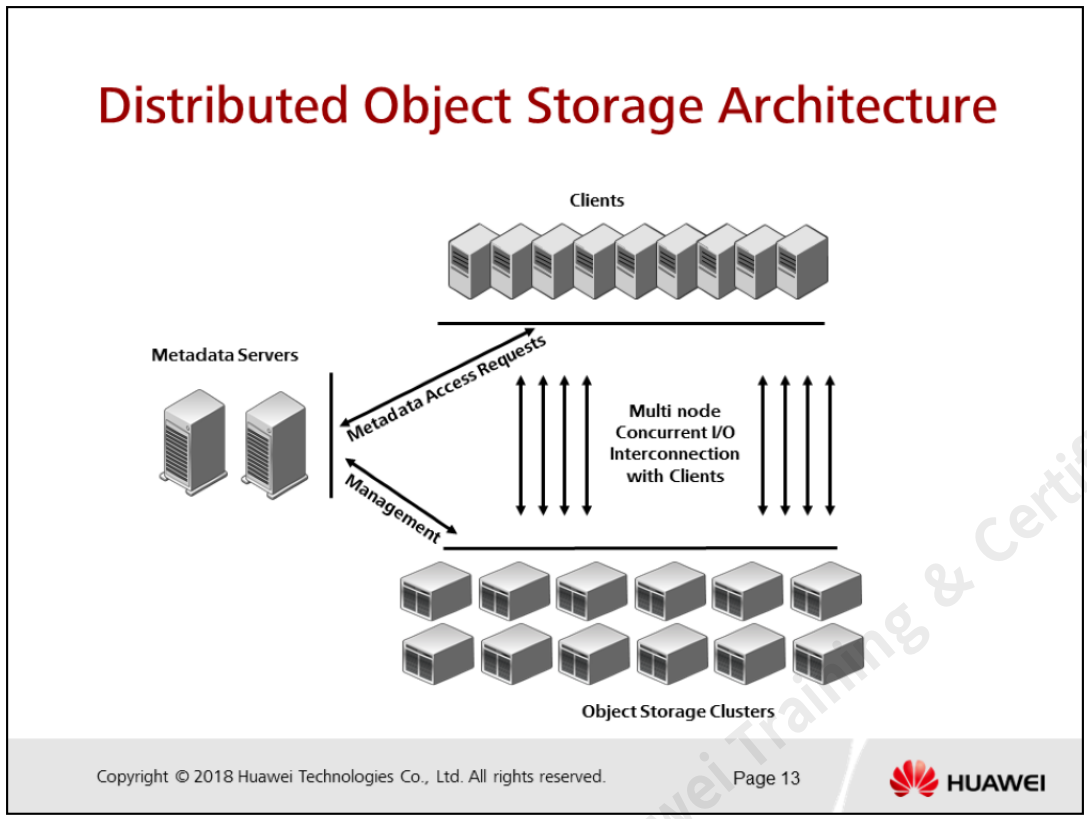
- Supports up to 10 billion files for storage.
- Global namespace and global sharing.
- High performance with concurrent I/O access to multiple nodes in the target storage cluster up to several GBs of combined I/O rate.
- Online expansion.
- Automatic load balancing.
- Low cost as it supports x86 servers.
- This distributed storage architecture has no need for centralized nodes for I/O processing as the I/O operation is sent directly from the client to the object storage cluster either concurrently or non-concurrently via different nodes.

Huawei Training & Certification



- Supports Exabyte (1000 Petabytes) level file storage
- Global namespace and global sharing.
- Online expansion.
- Low cost as it supports x86 servers.
- This distributed storage architecture has a layer of centralized compute nodes that will process the I/O operation from the clients and directs them to the targeted storage within the high density disk enclosures.

Huawei Training & Certification



- Supports up to ten billion files for storage.
- Global namespace and global sharing.
- High performance with concurrent I/O access to multiple nodes in the target storage cluster up to several GBs of combined I/O rate.
- Online expansion.
- Low cost as it supports x86 servers.
- Object storage is a computer data storage architecture that manages data as objects, as opposed to other storage architectures like file systems which manage data as a file hierarchy, and block storage which manages data as blocks within sectors and tracks. It uses metadata for each object for indexing, data management and optimizing storage.
- In this distributed object storage architecture, we can see from the diagram above that there are metadata servers that manages the access request and manages the data while the object storage clusters has a multi node concurrent I/O interconnection with clients. In a standard I/O operation, the client can read/write the data from the object storage clusters with the help of metadata servers that pinpoints the location of the targeted data by using metadata information.

Huawei Training & Certification

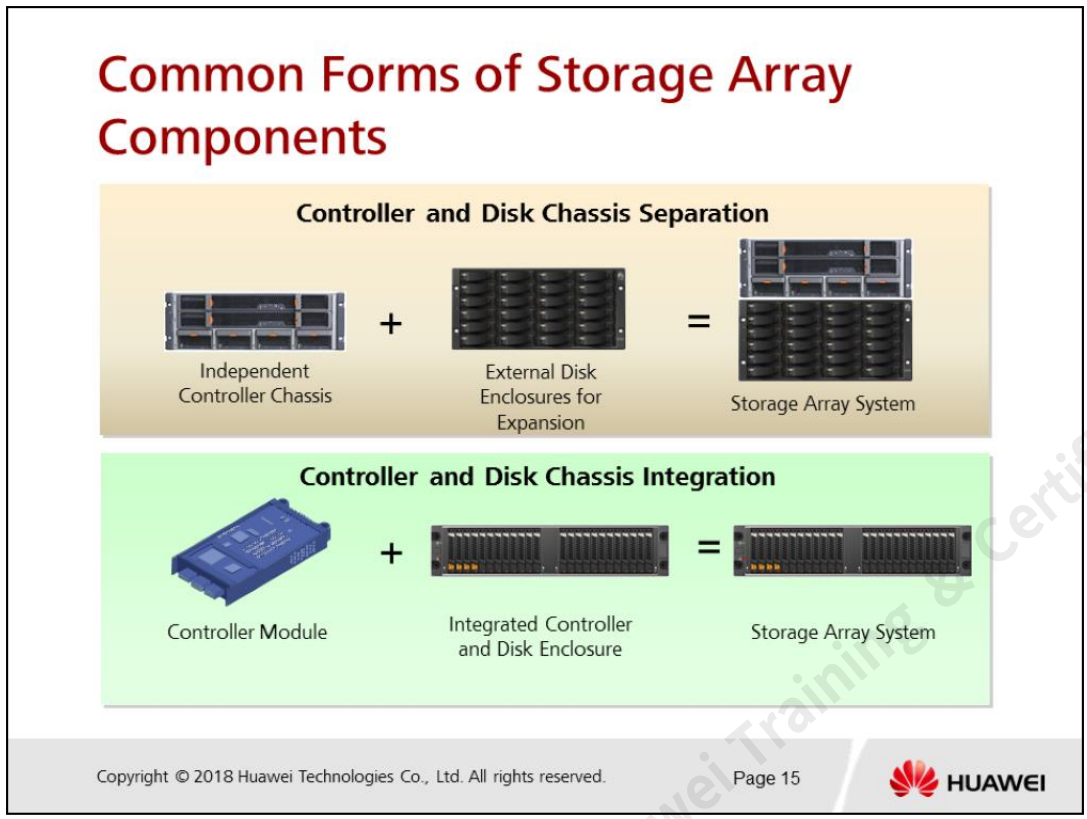
Contents

1. Architecture of Storage Systems.
2. **Components of Storage Systems.**
3. Introduction to Huawei Storage Products.

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

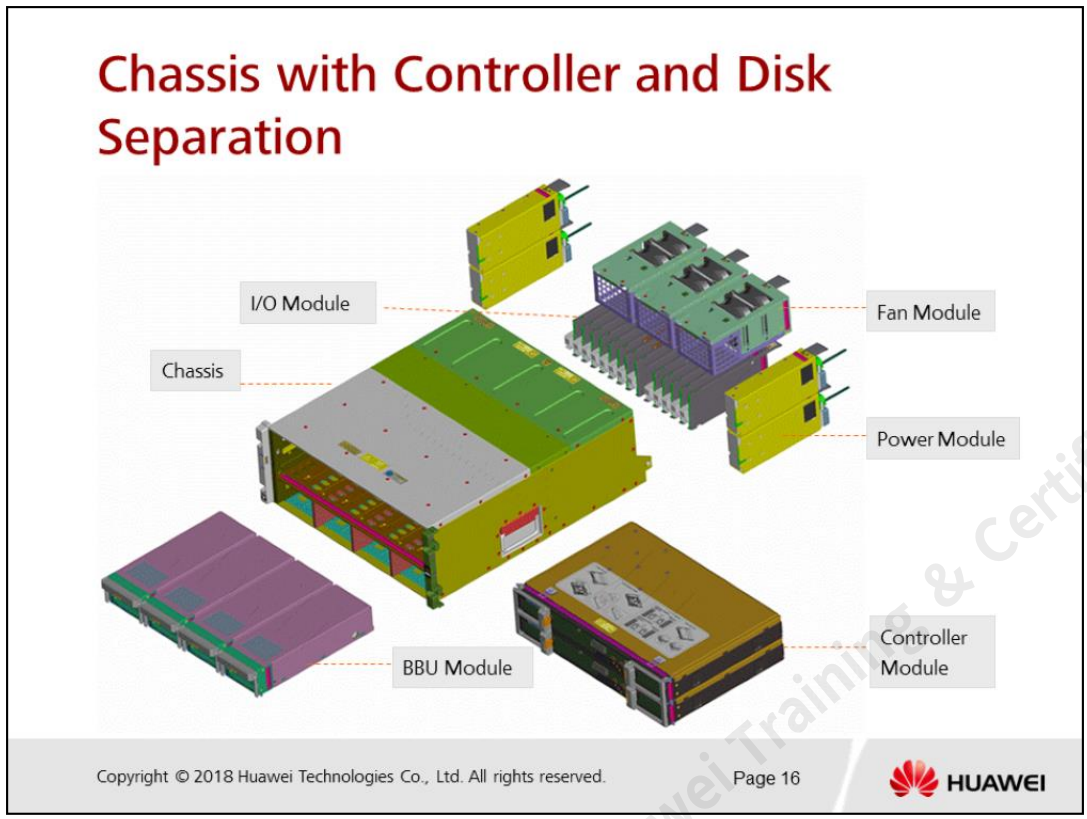
HUAWEI

Huawei Training & Certification
Huawei Training & Certification



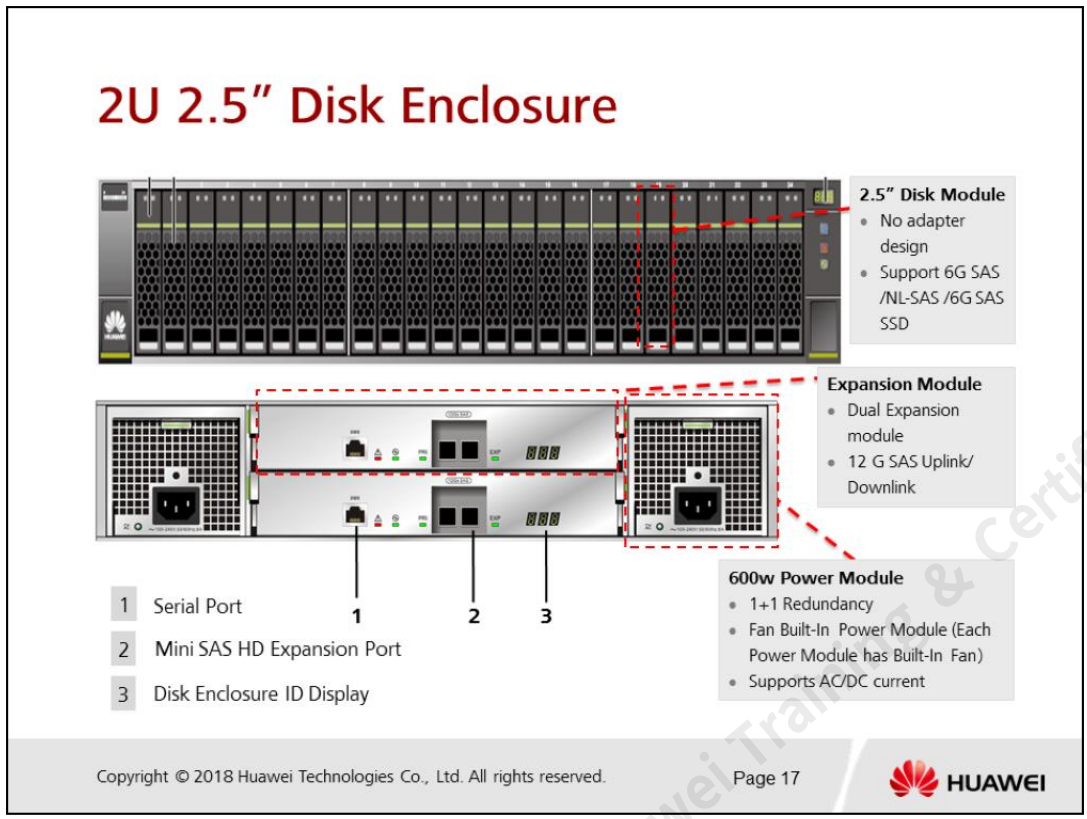
- The storage controller is the "brain" of the storage array, and it is the main component for processing and caching. It mainly manages the simple I/O operation and RAID management in the early days, but as technology progresses, it is now able to provide different kinds of advanced data management features such as snapshot, mirroring and cloning.

Huawei Training & Certification



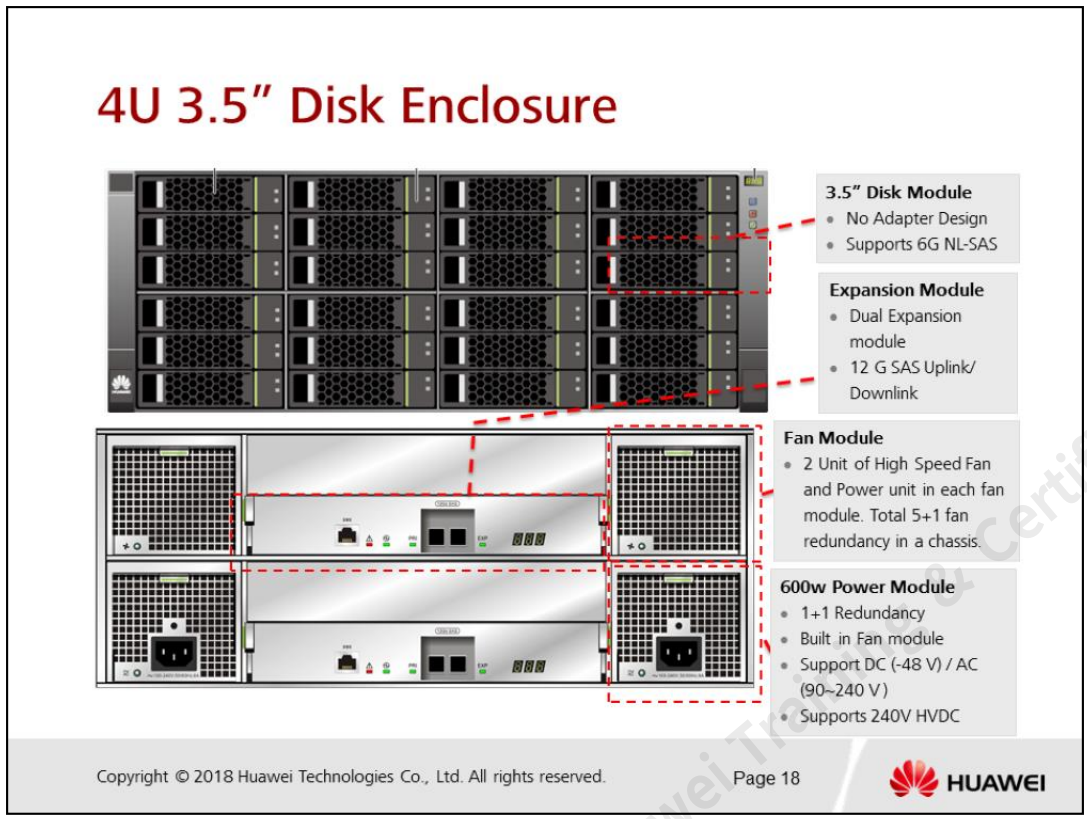
- The diagram above shows the different components that makes up the controller and disk separated chassis. The main components consist of the chassis, IO module, fan and power module and the controller and BBU (Backup Battery Unit) module.
- This chassis is connected to the external high density disk enclosures via the I/O module.
- BBU unit is the backup battery that provides power to the storage controller to save all the I/O cache into the coffer disks in the event of a power failure. Once the external power is restored, the cached data in the coffer disk will be flushed to the controller's cache and resume the I/O process to the physical disks in the external disk enclosures ensuring no data loss in the event of power interruptions.

Huawei Training & Certification



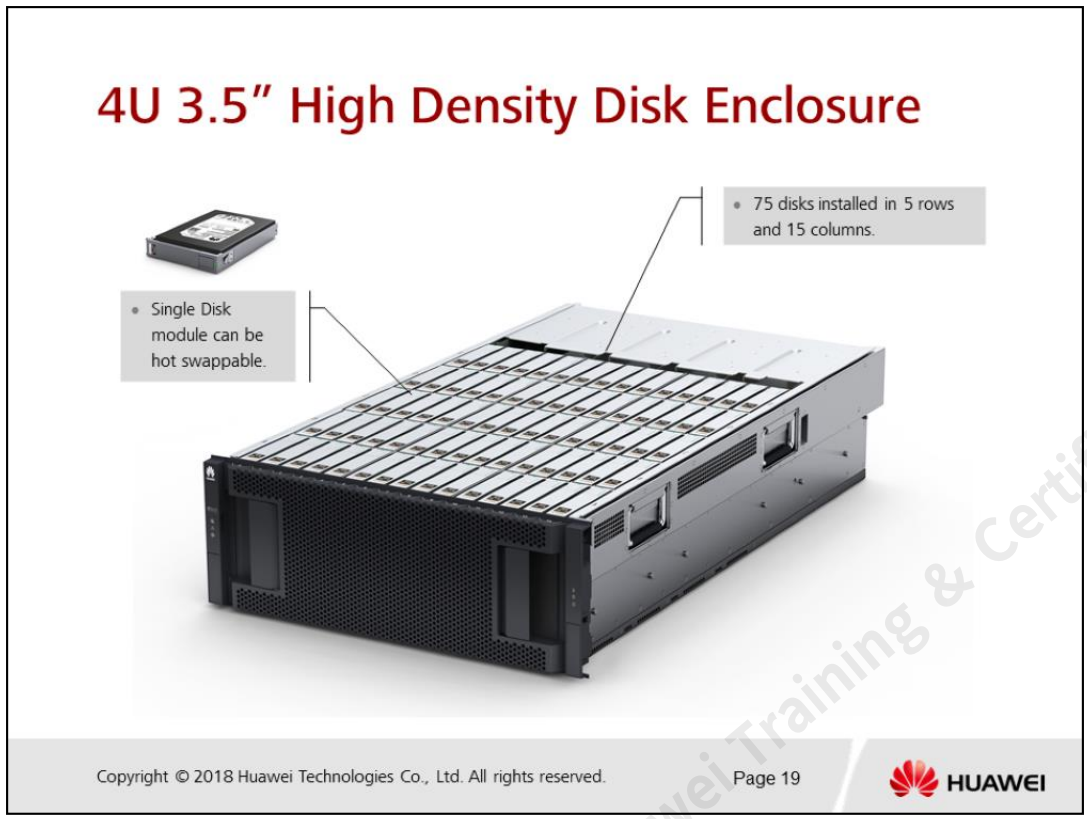
- The controller enclosure adopts a modular design and consists of a system sub rack(chassis), controllers, power-BBU modules, and disk modules.
- The disk module employs a no adapter design where the 2.5 inch hard disk are connected to the storage chassis directly via a Midplane without the need of an adapter board installed on each disks.
- This unit has a standard size of 2U and has a 1+1 redundant design for the power module and the controller module. The model shown above supports 25units of 2.5inch HDD in a single chassis.
- The controller module also acts as the expansion module to cascade multiple storage chassis together. The maximum number of storage controller in cascading mode differs if using either direct connection or switched network connection.
    - Cascading with Direct Connection between controllers(No switches) : Maximum number of cascaded controllers is 4.
    - Cascading in a switched network : Maximum number of cascaded controllers is 8.
    - The Mini SAS HD expansion port is used for expansion and cascading multiple storage chassis

Huawei Training & Certification - Huawei Training & Certification



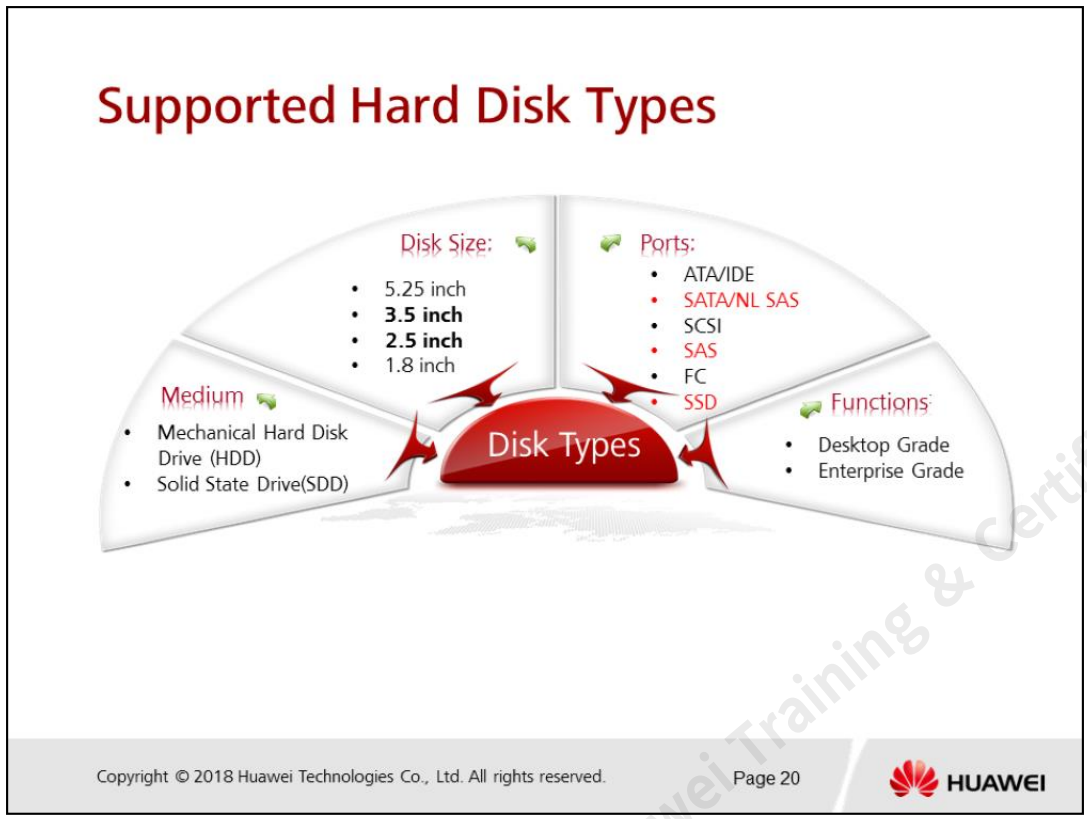
- The disk slots of a 4 U SAS disk enclosure are numbered 0 to 23 from left to right and from top to bottom.
- The expansion module has 2 Mini SAS HD expansion port (PRI/EXP) for cascading multiple chassis. There is also a Serial port for management console connection along with the disk enclosure ID display to show the ID for showing the ID of the storage chassis.
- The 600w power module supports both direct current (DC), alternating current (AC) and high voltage direct current (HVDC).

Huawei Training & Certification



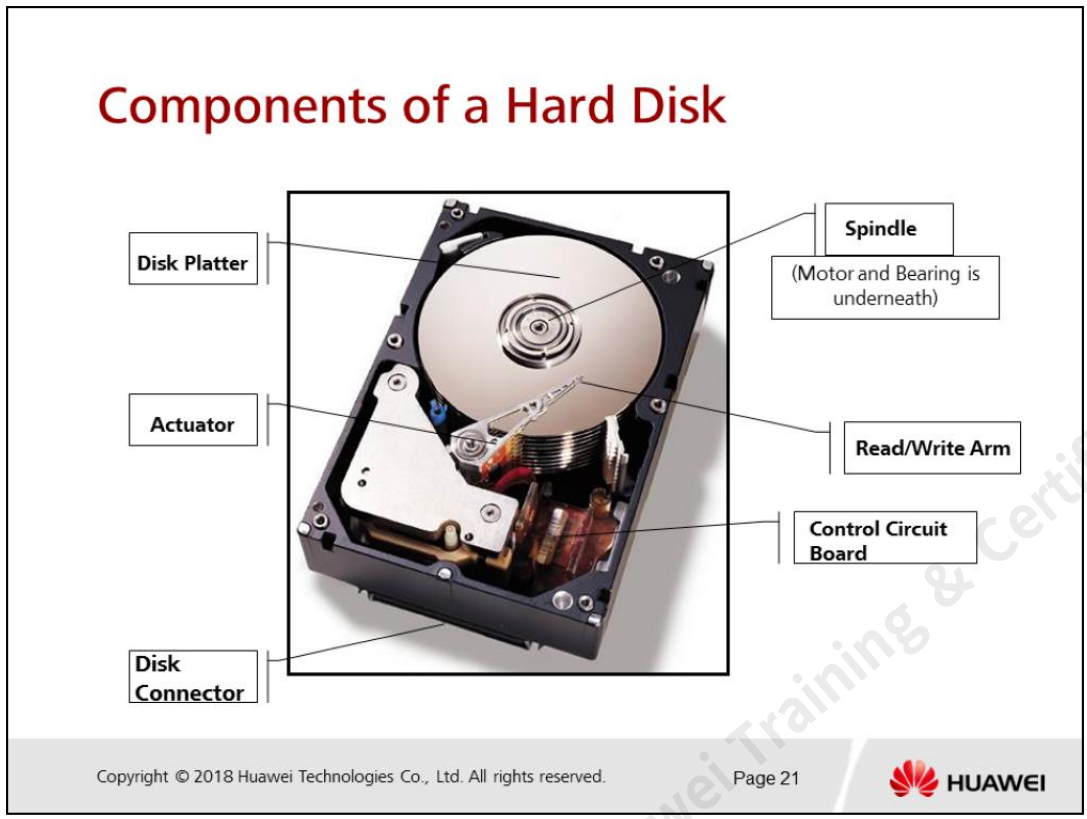
- The single disk module can be hot swappable without turning off the whole disk enclosure which minimizes the downtime during maintenance. However system performance might be impacted during disk replacement so it is suggested to do disk replacement during off-peak hours.
- The disk number of a high-density disk enclosure displayed on DeviceManager or CLI ranges from 0 to 74. These disks are numbered from left to right (15 columns) and from bottom to top (five rows).

Huawei Training & Certification



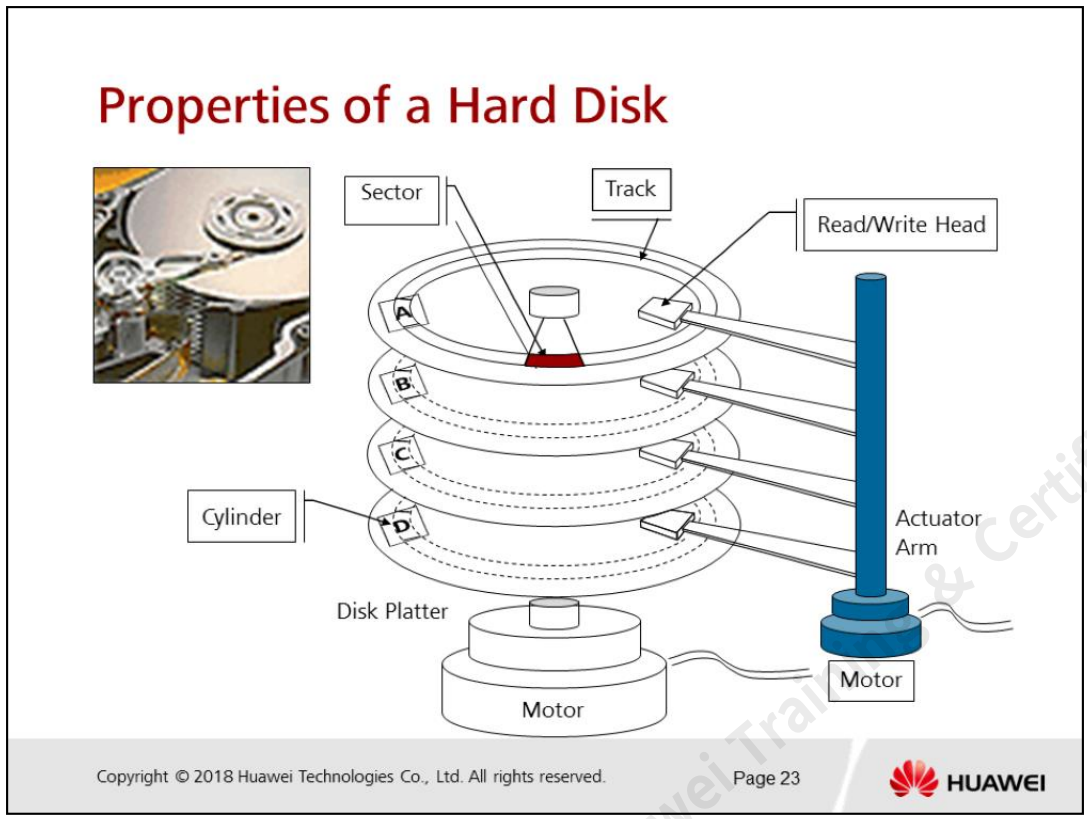
- The diagram above shows the different disk types that are supported by Huawei Storage Systems.
- Desktop grade hard disks are mainly for personal and home use and mostly used in desktop PC and laptops.
- Enterprise grade hard disks are mainly for enterprise level usage and mostly used in servers, storage arrays, and workstations.
- The difference between desktop and enterprise grade hard disks are as follows:
  - Capacity : Compared to desktop grade hard disks, enterprise grade hard disk have more storage capacity. Currently, single enterprise grade hard disk can have up to 4TB storage space or more.
  - Performance : Enterprise grade hard disk has higher performance mainly in rotation speed, caching and average seek time.
  - Reliability : Enterprise grade hard disk has higher mean time between failures (MTBF), commonly the desktop grade hard disk's MTBF is between 500, 000 hours whereas enterprise grade hard disks are above 1 million hours which makes enterprise grade hard disk much more reliable.

Huawei Training & Certification



- All the mechanical hard disks has the same working principles where the platters are covered with magnetic materials and the small sub micrometer region on the platter is polarized to represent a message or bit in binary form.
- The Read/Write Arm : Handles the data reading and writing :
  - Actuator: Moves the Read/Write arm to the specified locations on the platter.
  - Platter: Preserves or saves the written data.
  - Spindle: Rotates the disk platter and moves the disc's specified location to under the Read/Write arm.
  - Control Circuit Board: Controls the disk speed, movement of the actuator and send commands to the disk head.
- A hard disk uses a high speed moving read/write arm to read data from the magnetic particles on a flat disk platter. Data from the disk is read and transmitted to the CPU through the read head. The hard disk drive (HDD) consist of few platters, read/write heads, control circuit board, spindle and other mechanical components. The magnetic region on the disk allows unlimited time of data recording and data erasure but it is subject to the wear and tear of the disk drive. In this module, we will be introducing in detail the connectors on the hard drive, the components of the hard drive and the data saving mechanism and also discuss the factors that might affect the performance of the drives.
- A standard disk contains one or more flat round discs called disk platters. Data is recorded in the form of binary (0 & 1) onto these platters. These platters are hard and coated with magnetic materials on both sides, and data can be read/written to both sides of the platter. The number of platter and the storage volume of these platters determines the final storage capacity of the hard drive.

Huawei Training & Certification

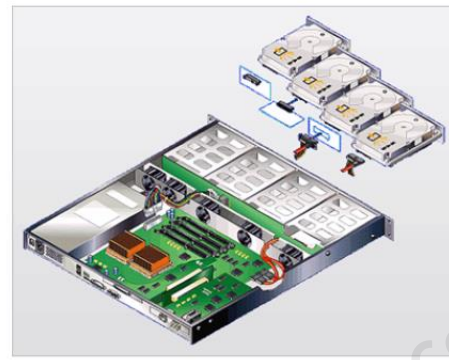


- A disk drive track is a circular path on the surface of a disk around the spindle where data in recorded magnetically. Track numbering starts from zero and the numbering starts from the edge of the platter to the innermost portion of the platter. We use the number of tracks per inch (TPI) to measure the density of the tracks alignment within a disk platter.
- Each track is divided into smaller units called sectors. Sectors are the smallest unit of storage that can be searched independently within a disk drive. The track and sector architecture is written to the disk by the disk manufacturers using a formatting tool on the disk platter. Different hard disk's track may contain different number of sectors. The first personal computer in the world had 17 sectors in each of its disk drive tracks, nowadays, our hard drives has much more sectors in each track compared to the past. The number of tracks within a disk drive differs based on the diameter and the TPI density of the disk platter, it may contain up to few thousand of tracks in a disk drive.
- Commonly, each sector can save 512 byte of user data, but some hard disk can be formatted to have a much bigger sector size such as 4KB sectors.
- Cylinders : All sets of the tracks that have the same track number on all the disk platter (Up and Bottom side) within a single disk drive forms a disk cylinder. The number of disk cylinder is equal to the number of disk tracks within the hard drive. However, the location of the magnetic head is determined by the disk cylinder number and not the track number which means that the disk controller moves the read/write head to any location on the disk platter based on disk cylinder number instead of track number. The addressing of individual sectors of the disk is traditionally done by referring to cylinders, heads and sectors (CHS).

Huawei Training & Certification

SAS Drives





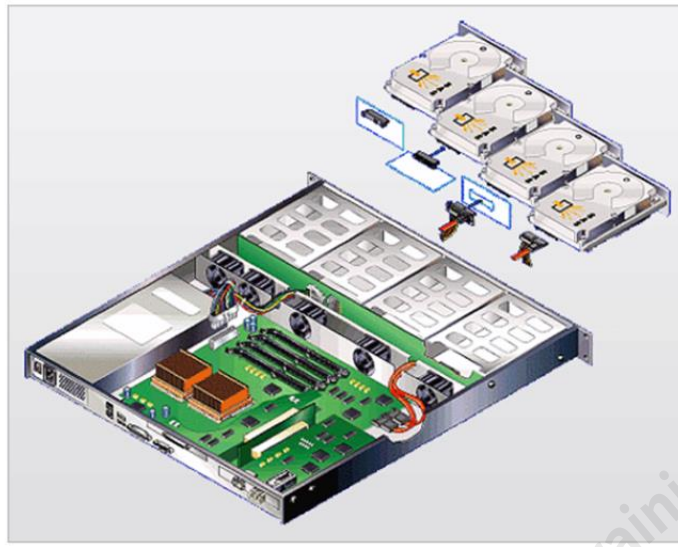
Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

- SAS stands for Serial Attached SCSI (SCSI Stands for Small Computer System Interface, typically pronounced as "scuzzy") and is a technology for transferring data from and to hard drives.
- SAS has provided a high performance, high reliability, scalable and easy operation solution for enterprise datacenters.
- SAS has provided unified support in terms of physical connector ports for ATA (Serial ATA) while is compatible with the parallel SCSI logics, providing more choices for servers and network storage for the disk types that can be used within their storage.
- SAS drives are faster and have more bandwidth throughput than SATA drives. SAS drives are also more reliable than SATA. However, SAS drives often comes with a higher cost range compared to SATA drives.

Huawei Training & Certification

NL - SAS Drives



Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

- Nearline-SAS (NL-SAS) are basically enterprise SATA drives with a SAS interface, R/W head and disk platters. In addition, they have the rotational speed (spin) of traditional enterprise-class SATA drives (7800 or slower RPM) with the fully capable SAS interface command sets. NL-SAS drives is a combination of SAS connectors and SATA drive platter which means that it has all the connectivity benefits of SAS while using a SATA disk platter that comes near to SAS drive level performance. Since, the internal disk structure uses SATA technologies, NL-SAS drive cannot reach full SAS drive level of performance and only come near to SAS level performance.
- NL refers to the abbreviation of Near Line, which means near line storage where it can be considered as the midpoint between the customer's online storage and offline storage. Here, online storage refers to data that has high access rate and frequently accessed by customer such as databases and offline storage where data access rate is very low and with very low access frequency such as customer data backup archives.

Huawei Training & Certification

Solid State Drives (SSD)

- As price lowers and storage capacity increases, SSD is becoming more and more popular in the market.
- SSD Principles:
  - Uses flash technology to store data.
  - No mechanical structure with lower energy consumption, low heat signature and low noise.
- Limited lifetime due to SSD usage frequency.
- 3 types of SSD:
  - SLC (Single Level Cell)
  - MLC (Multi Level Cell)
  - TLC (Triple Level Cell)

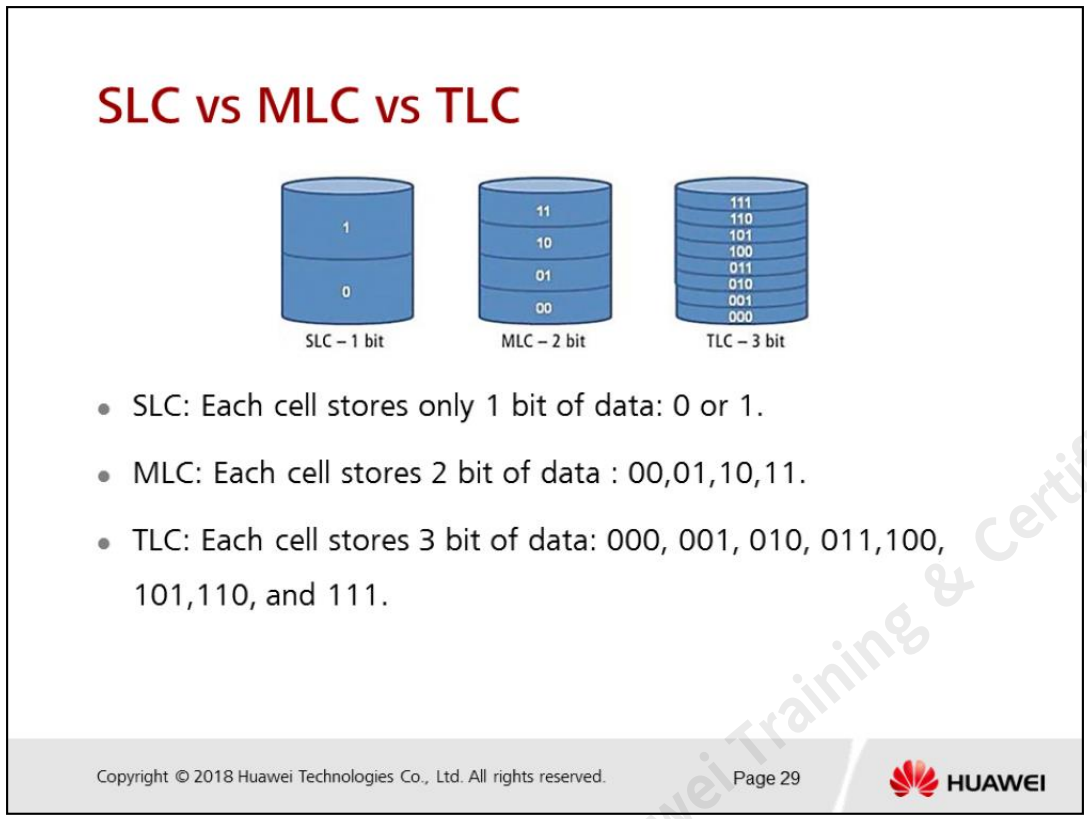


Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

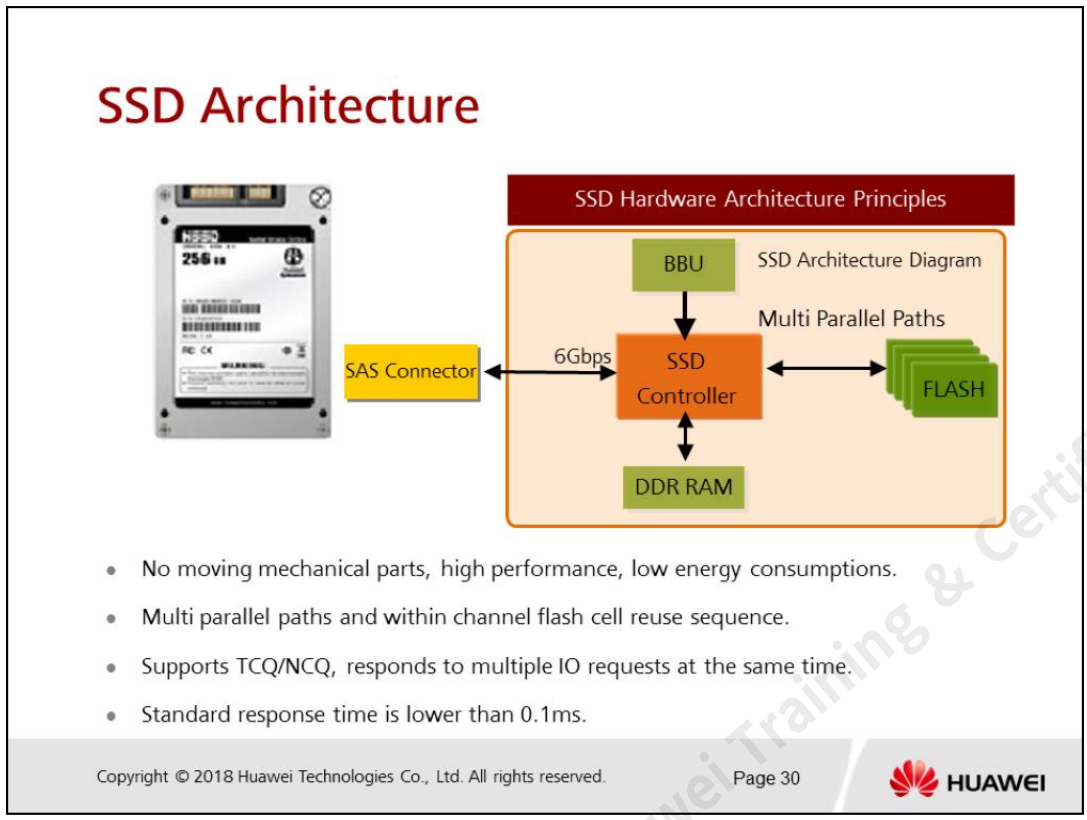
- Although traditional mechanical drives will not be disappearing from the market in a short period of time, but its replacement which is the Solid State Drive (SSD) has become more and more popular and widely used in the market. SSD does not use magnetic materials to store data but uses basic units of NAND Flash called "cell" as storage units to store data. NAND Flash is a non-volatile random access memory storing medium which has a unique feature of not losing any written data even if the power is cut off. This technology allows data to be quickly written to SSD drives much faster than compared to mechanical drives. Another benefit of SSD is that it does not produce loud noise or generate high heat signatures like traditional mechanical drives such as SATA/NL-SAS/SAS drives.
- SSD does not have moving mechanical parts which greatly lowers the risk of mechanical part failures, but this does not mean that its lifetime is unlimited. Due to NAND Flash is a non-volatile medium, each block of data needs to wiped before a new block of data can be written or else it may cause error data. This increases the write/erase cycles of the SSD drive. However, NAND Flash has a limited write/erase limit which means that each cell of the SSD has a certain limited number of times it can be erased and written into, once the limit is exceeded, that particular cell can no longer be used for data read/write. However, these kind of wear and tear can be easily monitored and predicted so that preparations can be made in advance for disk replacement in time. On the other hand, mechanical hard drives failures have no early symptoms which means that disk failure can happen at any time and replacement disk must be prepared at all times.

Huawei Training & Certification Huawei Training & Certification



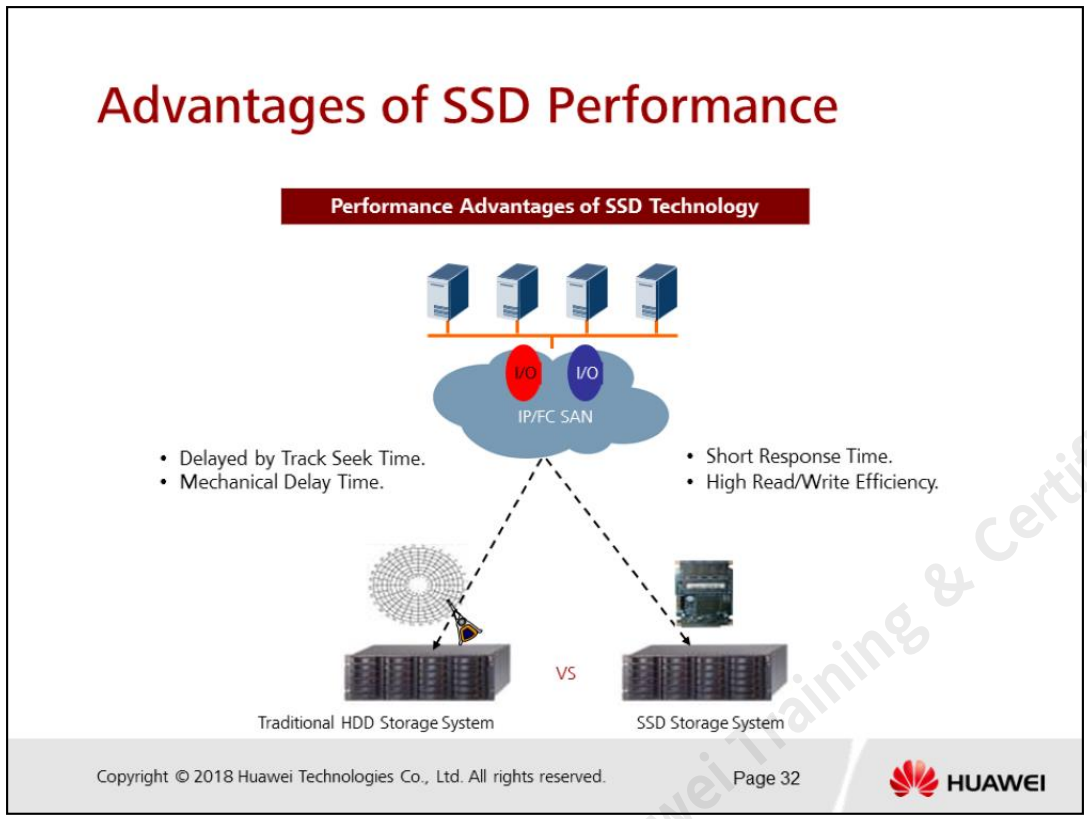
- Each cell in the NAND Flash of the SSD drive traditionally can store 1 bit of data, which is either 0 or 1. However, in newer generation of SSD drives, we are able to store more data in a cell using a special technologies which gave birth to the different types of SSD drives in the market today. The types of SSD drives can be categorized into Single Level Cell (SLC), Multi Level Cell (MLC) and Triple Level Cell (TLC).
- SLC only stores 1 bit of data which is in the binary form of 0 or 1. MLC on the other hand is able to store 2bits of binary data which means that it is able to store data in four states such as 00, 01, 10, 11. TLC is able to store 3 bits of data which means that it is able to store 8 different states of data which is shown on the diagram above.
- Three different types of cell although is able to store different amounts of data but their physical size remains the same, and this is also one of the reasons why SSD drive capacity are growing exponentially. Initially, SSD were only available in 64Gb or smaller capacity, but now the largest TLC SSD available on the market can store up to 2TB of data.
- However, different types of SSD has different level of wear resistance which results in different levels of drive reliability. The wear resistance of SSD disk is also an important factor in choosing a SSD disk.
- SLC has the highest reliability however comes at the highest cost, MLC has the middle level of reliability while TLC has the lowest reliability but with the lowest cost. Hence, choosing the right kind of SSD disk depends on the customer's requirement and the storage scenario and also the costing evaluations.

Huawei Training & Certification



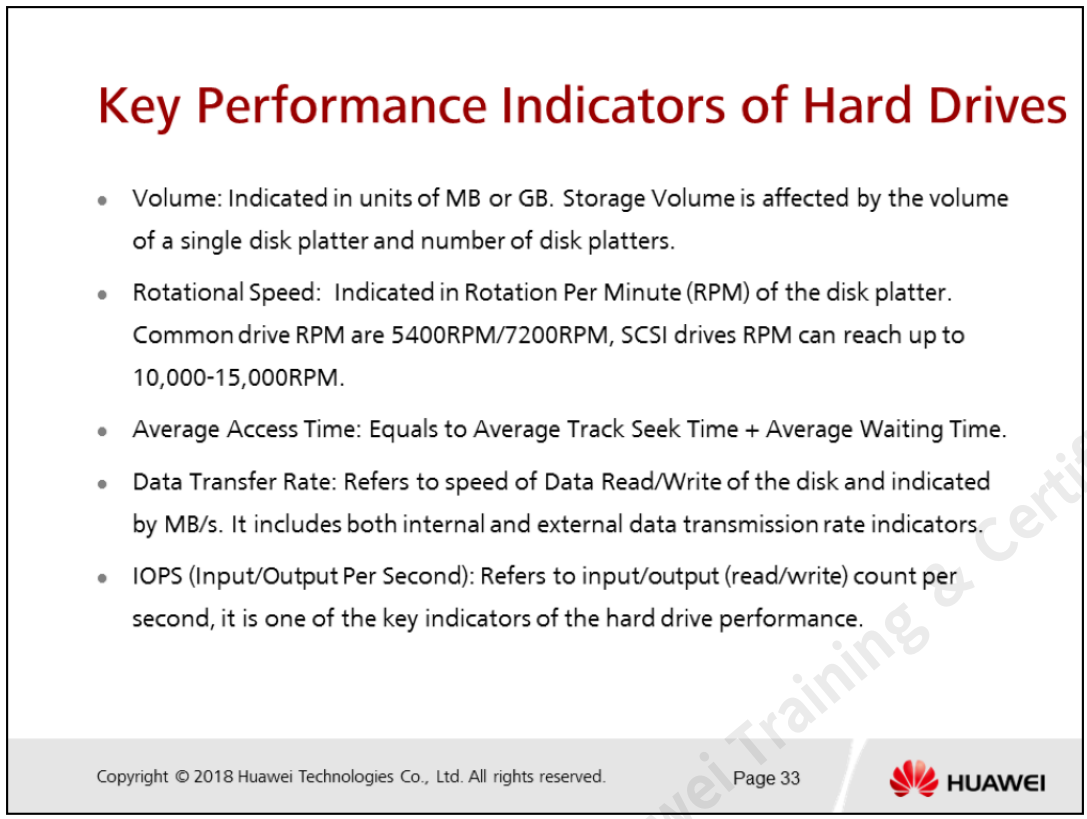
- The diagram above shows the architecture of the SSD and the technologies supported by it to optimize the read/write cycles of data for better performance.
- The SSD controller is able to process commands and IO requests in parallel due to the multi parallel path design that exist between the controller and the NAND Flash cells within the drive. The cell write/erase cycle is also optimized and allows the drive to reclaim and reuse any of the flash cells to store data efficiently.
- Native Command Queuing (NCQ) is an extension of the Serial ATA protocol allowing hard disk drives to internally optimize the order in which received read and write commands are executed. This can reduce the amount of unnecessary drive head movement, resulting in increased performance (and slightly decreased wear of the drive) for workloads where multiple simultaneous read/write requests are outstanding, most often occurring in server-type applications.

Huawei Training & Certification



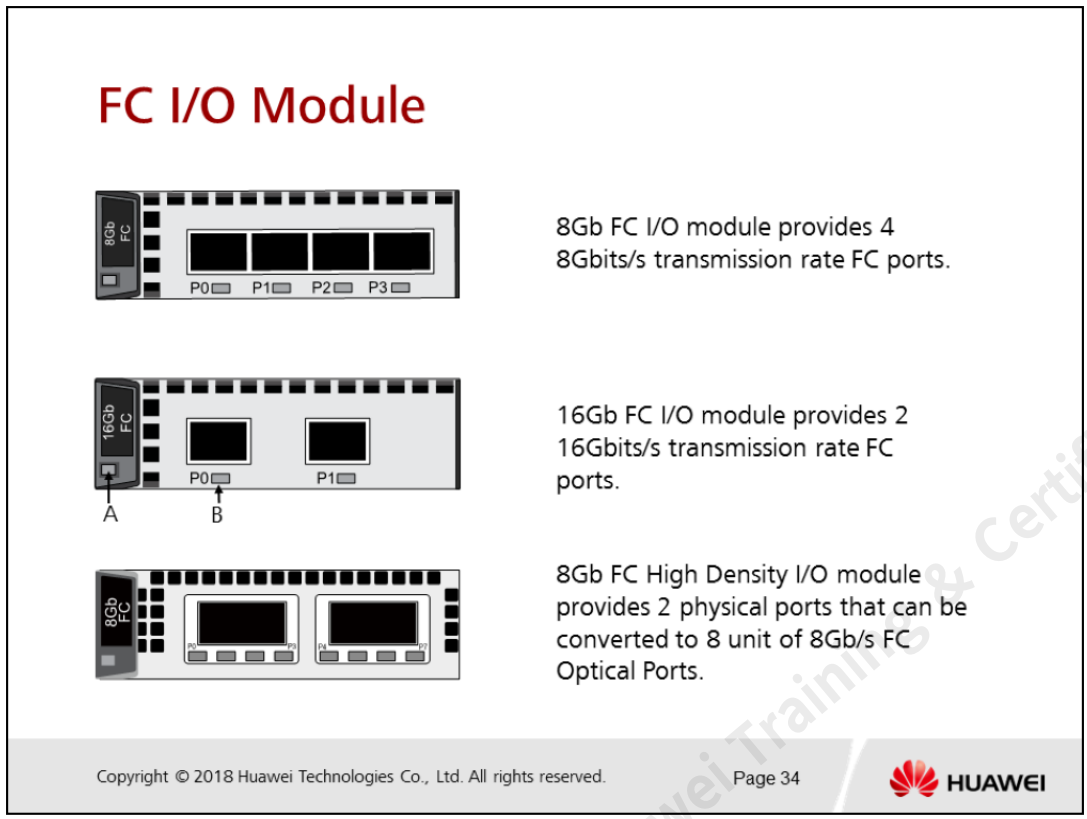
- Traditional HDD storage system has more delays compared to SSD storage system due to its mechanical parts. Traditional storage has to uses a portion of time for track seeking and moving the mechanical parts for data read/write operations. SSD storage has no moving mechanical parts and uses NAND Flash technologies without the need for track seeking which saves more time compared to traditional storages.
- Short Response Time:
  - Due to the nature of the mechanical parts in traditional storages, most of time is wasted in track seeking and mechanical delay, hence data transmission speed is heavily restricted by these parts. SSD storages has no mechanical parts which saves the time for track seeking and mechanical delay, which in turns allows a much faster IO request response rate.
- High Read/Write Efficiency:
  - Mechanical drives needs to move the heads constantly when it is performing random read or write operations which lowers the efficiency of the read/write process. SSD in the other hand, calculates the data storage locations internally and performs the read or write process, which saves a lot of mechanical operation time that in turns increases the read/write efficiency greatly.

Huawei Training & Certification



- IOPS is a key indicator for high frequency random read write applications such as OLTP (Online Transaction Processing). Another key indicator is **Throughput**, which refers the amount of data that can be successfully transferred in a single unit of time. For large amount of sequential data read write applications such as video editing of TV stations and video streaming, they are more focused on the indicator such as Throughput.

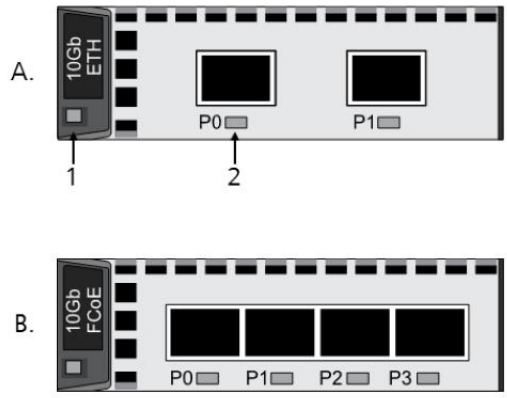
Huawei Training & Certification



- FibreChannel Input/Output Modules (FC I/O Module) provides connectivity and expansion to storage system by providing different number FC ports with varying transmission rate based on the models of the I/O module.
- Indicator lights on the modules above shows status of :
  - Refer to A in diagram above: Module Power Indicator/ Hot Swap Button. This shows the power status of the module and also serves as a button to initiate the hot swapping process of the module for replacement.
    - Steady green: The interface module is working correctly.
    - Blinking green: The interface module receives a hot swap request.
    - Steady red: The interface module is faulty.
    - Off: The interface module is powered off or hot swappable.
  - Refer to B in diagram above: Link/Speed Indicator Light. This shows the current running link speed of the I/O module FC port.
    - Steady blue: The data transfer rate between the storage system and the application server is 16 Gbit/s.
    - Blinking blue: Data is being transferred.
    - Steady green: The data transfer rate between the storage system and the application server is 4 Gbit/s or 8 Gbit/s.
    - Blinking green: Data is being transferred.
    - Steady red: The port is faulty.
    - Off: The link to the port is down.

Huawei Training & Certification

FCoE I/O Module



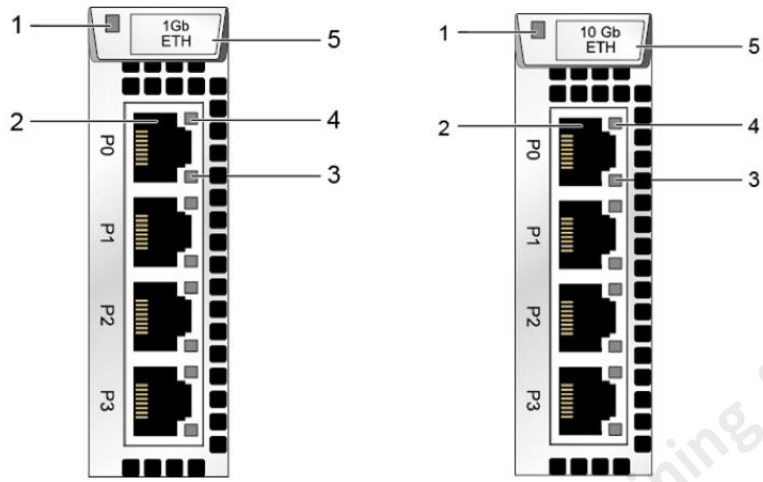
Note:
1 = Module Power Indicator Light / Hot Swap Button
2 = Port Link/Speed Indicator Light

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

- Refer to A in diagram above: 10Gb FCoE(FibreChannel Over Ethernet) I/O module provides 2 10Gbit/s data transmission rate FCoE port.
  - 10Gb FCoE I/O module (2 Ports) only supports Direct Network connection.
- Refer to B in diagram above: 10Gb FCoE(FibreChannel Over Ethernet) I/O module provides 4 10Gbit/s data transmission rate FCoE port.
  - 10Gb FCoE I/O module (4 Ports) only supports Switched Network connection.
- 10Gb FCoE I/O module is not recommended to run both iSCSI and FCoE protocols at the same time as it will lower the overall performance rate and have higher fluctuations in performance.

Huawei Training & Certification

Ethernet I/O Module

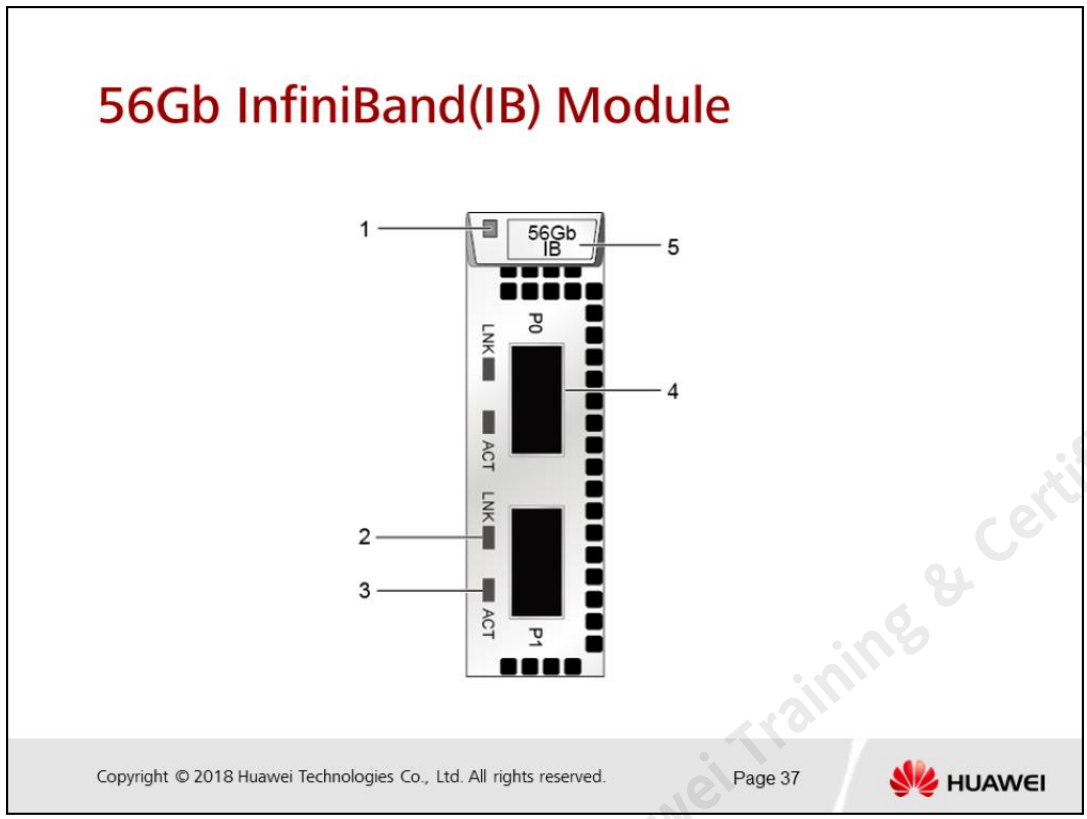


Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

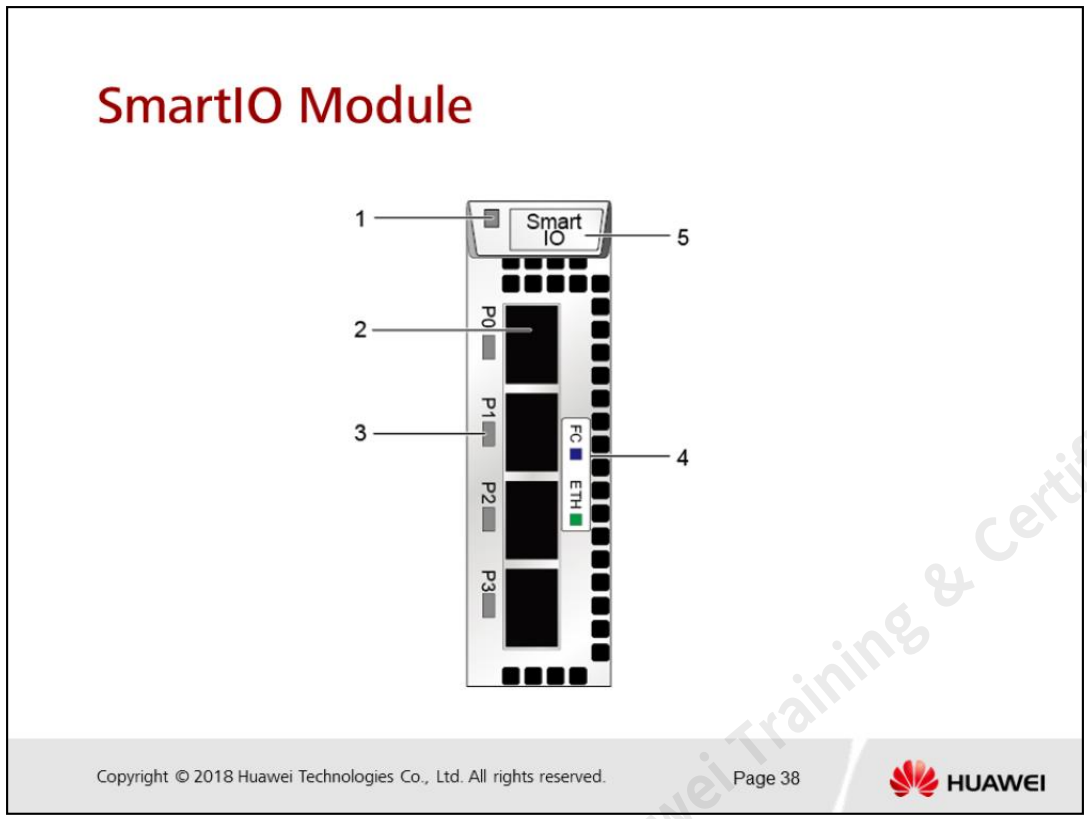
- GE (Gigabit Ethernet) Electrical I/O Module provides 4 1Gbits data transmission rate electrical Ethernet ports.
- 10GE Electrical I/O Module provides 4 10Gbit/s electrical Ethernet port, it supports auto speed negotiation between 10GE and GE for these ports.
- Indicator Lights:
  - Module Power Indicator Light/ Hot Swap Button.
  - 10GE Electrical Port.
  - 10GE Electrical Port Link/Active Indicator Light.
  - 10GE Electrical Port Speed Indicator Light.
  - Module Latch.

Huawei Training & Certification



- The 56Gb IB (InfiniBand) I/O module provides 2 (4x14Gbit/s) data transmission rate IB ports.
- Indicator Lights:
  - Module Power Indicator Light/ Hot Swap Button.
    - Steady green: The interface module is working correctly.
    - Blinking green: There is a hot swap request to the module.
    - Steady red: The module is faulty.
    - Off: The interface module is powered off or hot swappable.
  - 56Gb IB Port Link Indicator Light.
    - Steady green: The port is connected properly.
    - Off: The port link is down.
  - 56Gb IB Port Active Indicator Light.
    - Steady orange: Data is being transmitted.
    - Off: No data is being transmitted.
  - 56Gb IB Port.
  - Module Latch.

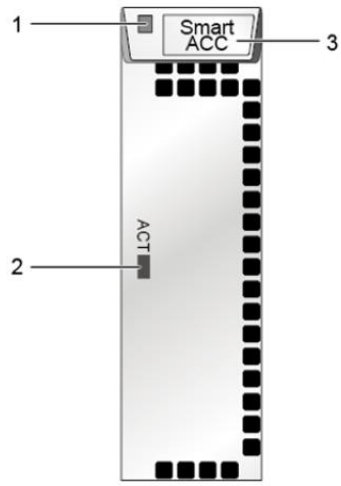
Huawei Training & Certification



- SmartIO Module supports three types of optical transceiver with data transmission rate of 16Gb, 8Gb, 10Gb.
- Indicator Lights:
  - Module Power Indicator Light/ Hot Swap Button.
  - SmartIO Port.
  - Port Link/Active/Mode Indicator Light.
  - Port Mode Silk Screen.
  - Module Latch.
- If the mode of the SmartIO port is set to FCoE/iSCSI or Cluster on the software interface, the port indicator is in ETH mode and a 10 Gbit/s optical module is required. 10 Gbit/s optical transceiver does not support GE/10GE auto negotiation.
- If the mode of the SmartIO port is set to FC on the software interface, the port indicator is in FC mode, and an 8 Gbit/s or 16 Gbit/s optical module is required.
- If the mode of the SmartIO interface module is set to FCoE/iSCSI and the host uses the FCoE protocol, the module must connect to FCoE switches for networking and a 10 Gbit/s optical module is required.
- If the mode of the SmartIO interface module is set to FCoE/iSCSI and the host uses the iSCSI protocol, the MTU value of the SmartIO port must be the same as that of the host.

Huawei Training & Certification

Smart Acceleration Module



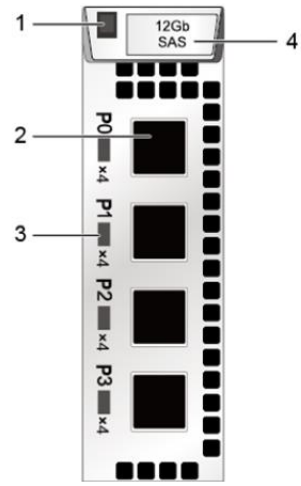
Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

- Smart Acceleration Module can concurrently provide hardware acceleration capabilities of data deduplication, fingerprint calculation, gzip decompression and compression, offloading the CPU load and increase the overall performance of the storage system in data deduplication and compression.
- Indicator Lights:
  - Module Power Indicator Light/ Hot Swap Button.
  - Active Indicator Light.
  - Module Latch/ Silkscreen.

Huawei Training & Certification

12Gb SAS Expansion Module

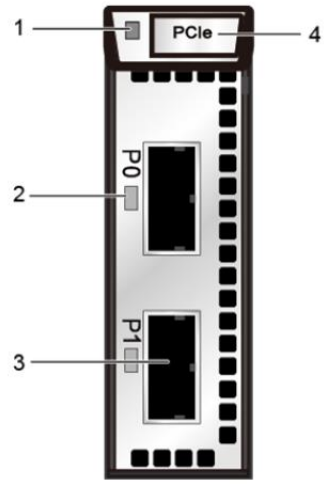


Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

- A SAS interface module provides four 4 x 12 Gbit/s mini SAS HD expansion ports that provide connectivity to disk enclosures.
- The SAS interface module connects to the back-end storage array of the storage system through a mini SAS HD cable.
- When the transfer rate of the connected device is less than that of the expansion port, the expansion port automatically adjusts the transfer rate to that of the connected device to ensure the connectivity of the data transfer channel.
- Indicator Lights:
  - Module Power Indicator Light/ Hot Swap Button.
  - Mini SAS HD Expansion Port.
  - Mini SAS HD Expansion Port Indicator Light.
  - Module Latch.

Huawei Training & Certification

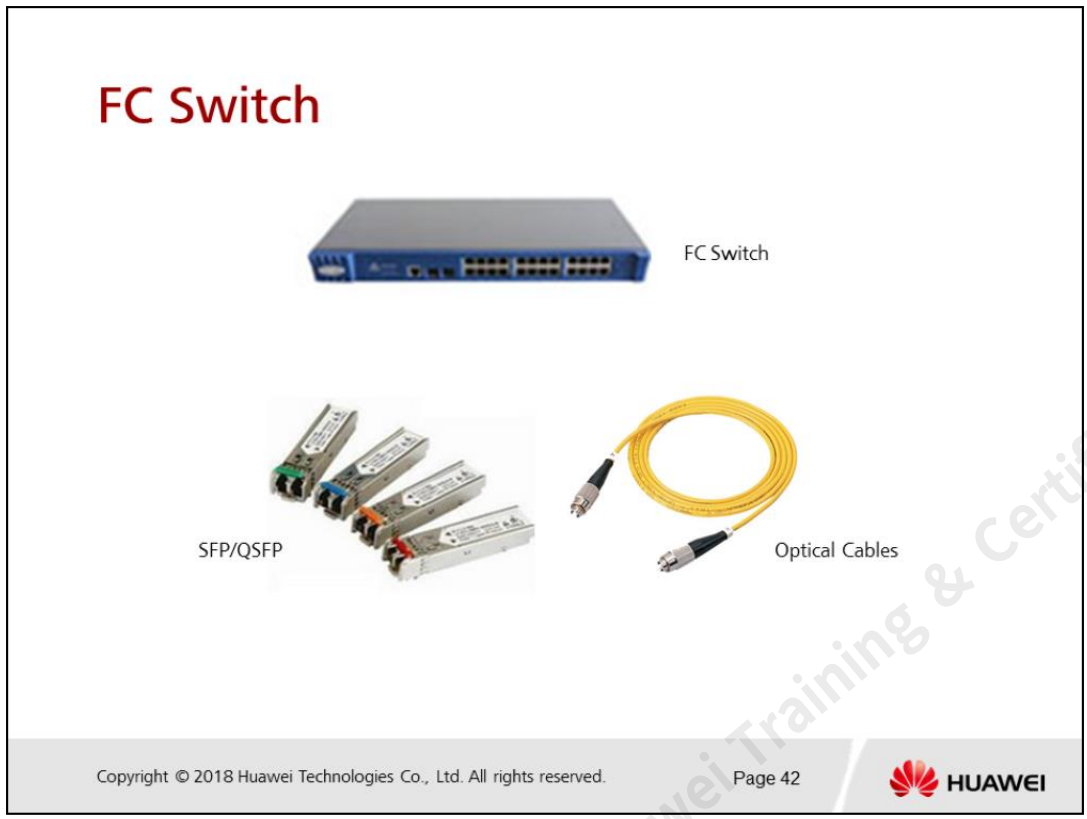
PCIe Module



Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

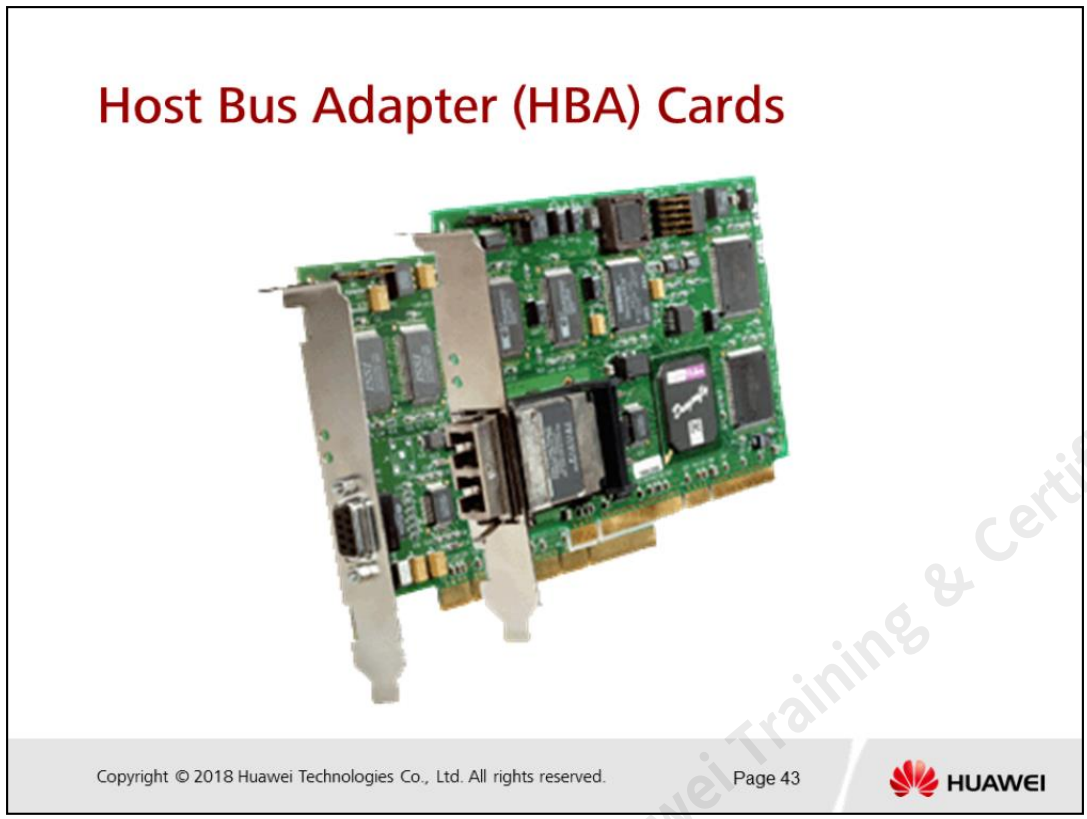
- PCIe Module provides 2 PCIe ports that are the main service port between Oceanstor 18000 V3 and the data switches, it is used to exchange dataflow and control flow messages between storage engines within the network.
- Indicator Lights:
  - Module Power Indicator Light/ Hot Swap Button.
  - PCIe Port Link/Speed Indicator Light.
  - PCIe Port.
  - Module Latch.

Huawei Training & Certification



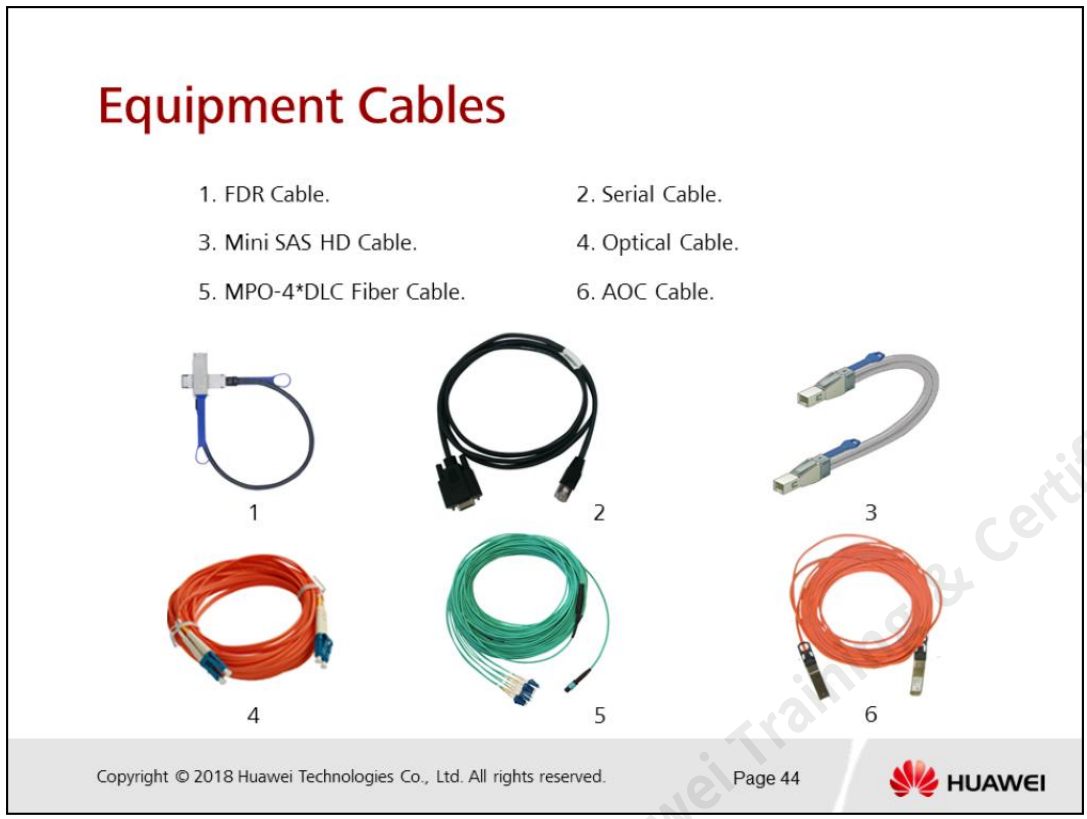
- Fibre Channel switch is a network switch compatible with the Fibre Channel (FC) protocol. It allows the creation of a Fibre Channel fabric, that is the core component of a storage area network (SAN). The fabric is a network of Fibre Channel devices which allows many-to-many communication, device name lookup, security, and redundancy. FC switches implement zoning, a mechanism that disables unwanted traffic between certain fabric nodes.
- The small form-factor pluggable (SFP) is a compact, hot-pluggable optical module transceiver used for both telecommunication and data communications applications. An SFP interface on networking hardware provides the device with a modular interface that the user can easily adapt to various fiber optic and copper networking standards such as FibreChannel and Ethernet.
- The Quad Small Form-factor Pluggable (QSFP) is a compact, hot-pluggable transceiver used for data communications applications similar to SFP. Certain technologies such as InfiniBand uses QSFP transceivers for the connection to networking hardware.
- A fiber optic cable is a network cable that contains strands of glass fibers inside an insulated casing. They're designed for long distance, very high performance data networking and telecommunications. Compared to copper cables, fiber optic cables provide higher bandwidth and can transmit data over longer distances.

Huawei Training & Certification



- A host bus adapter (HBA) is a circuit board and/or integrated circuit adapter that provides input/output (I/O) processing and physical connectivity between a host system, or server, and a storage and/or network device. Because an HBA typically relieves the host microprocessor of both data storage and retrieval tasks, it can improve the server's performance time.
- FC HBA Card: Commonly known as Fibre Network Cards. The transmission protocol used by FC HBA cards is FibreChannel protocol and usually is connected to the FC Switches using fibre optic cables.
- iSCSI HBA Card: Uses iSCSI protocol for data transmission and has same ports with Ethernet network cards. By integrating the hardware iSCSI initiator on the circuit board and using TCP/IP offload engine technology to do data processing on the HBA cards, we are able to lessen the host CPU's workload and provide a high usability hardware multi channel feature along with server remote booting feature. Booting from SAN is also possible if the iSCSI card supports remote booting or boot from LAN features.

Huawei Training & Certification



- FDR (Fourteen Data Rate) cable is custom used for 56Gb InfiniBand I/O Module.
- Serial Cable is used for connecting equipment's serial port to the maintenance terminal. One end of the serial cable is RJ-45 connector used to the equipment and another end is the DB-9 connector used to connect to the maintenance terminal.
- Mini SAS HD cables is used for connecting disk enclosures with high density disk enclosures or connecting between 2 disk enclosures/storage chassis.
- Devices communicate between each other through the optical cables and FC switches. One end of the optical cables connect to the FC HBA card, while the other end connects to the FC switch or Application server. Both end of the optical cables are LC (Lucent Connector) type.
- MPO-4*DLC Fiber Cables is custom used by the 8Gb High Density I/O Module.
- AOC (Active Optical) Cable is used for connecting storage system controller module's PCIe port to the data switches.

Huawei Trainings & Certification

Contents

1. Architecture of Storage Systems
2. Components of Storage Systems
3. **Introduction to Huawei Storage Products**

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

Huawei Training & Certification
Huawei Training & Certification

Huawei Converged Storage Products

- Entry Level Storage:
  - OceanStor 2200/2600 V3 Series
- Middle Range Storage:
  - OceanStor 5300/5500 V3 Series
  - OceanStor 5600/5800 V3 Series
- High End Storage:
  - OceanStor 6800 V3 Series
  - OceanStor 18000 V3 Series

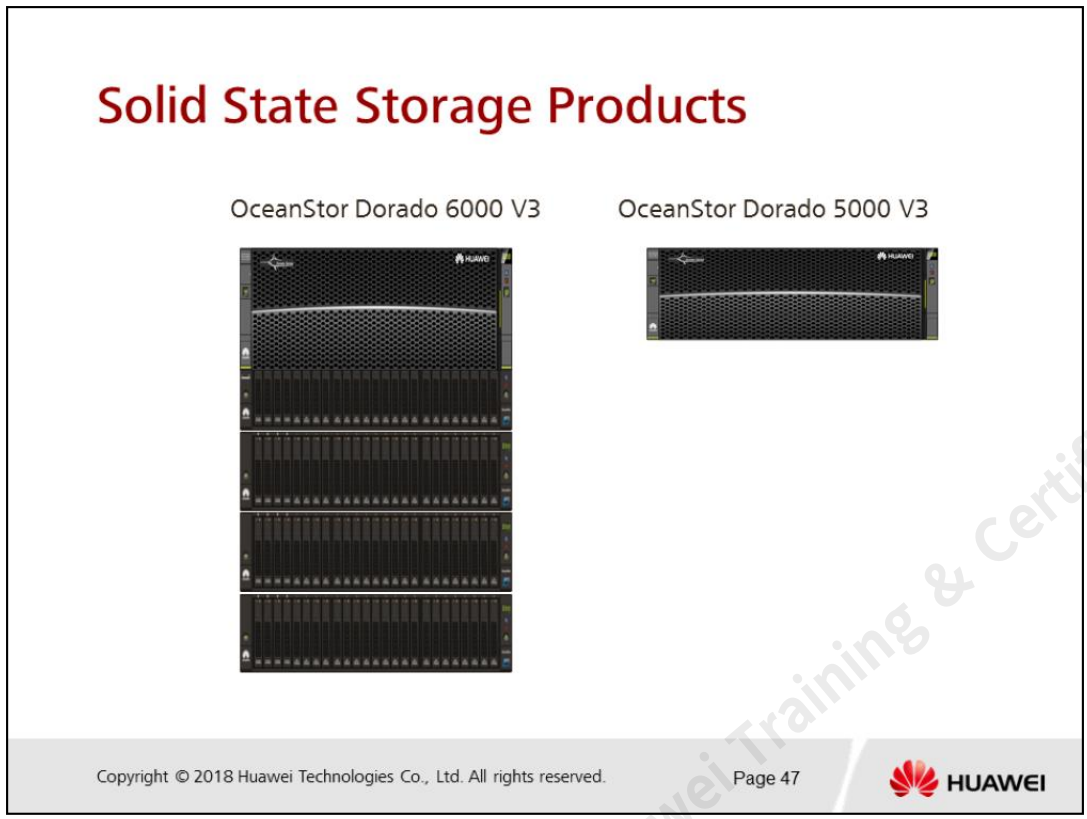


Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

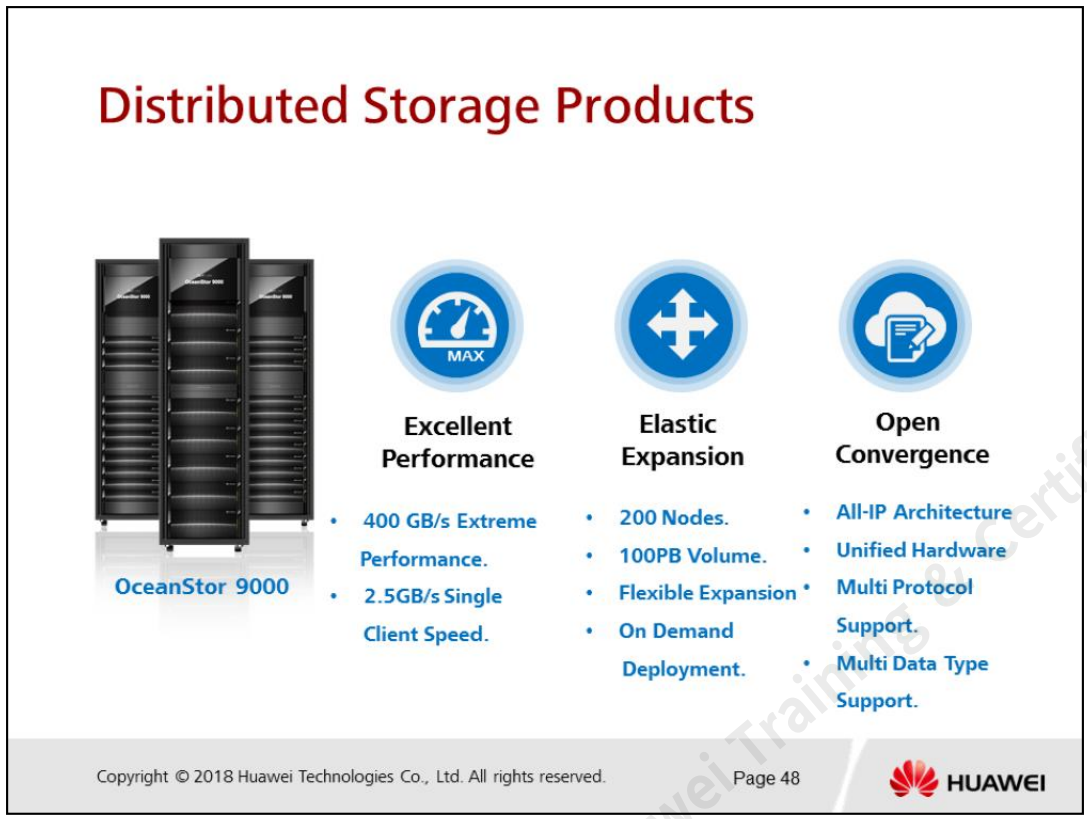
- OceanStor V3 series storage systems has high standard in terms of hardware architecture, unified block and file storage, combines multiple different advanced data application and data protection technologies, ensuring that Huawei storage systems has high performance, high scalability, high reliability and high usability to satisfy large and medium sized enterprise's different kinds of needs for storage products.
- OceanStor 2200 V3 only provides block storage services.

Huawei Training & Certification



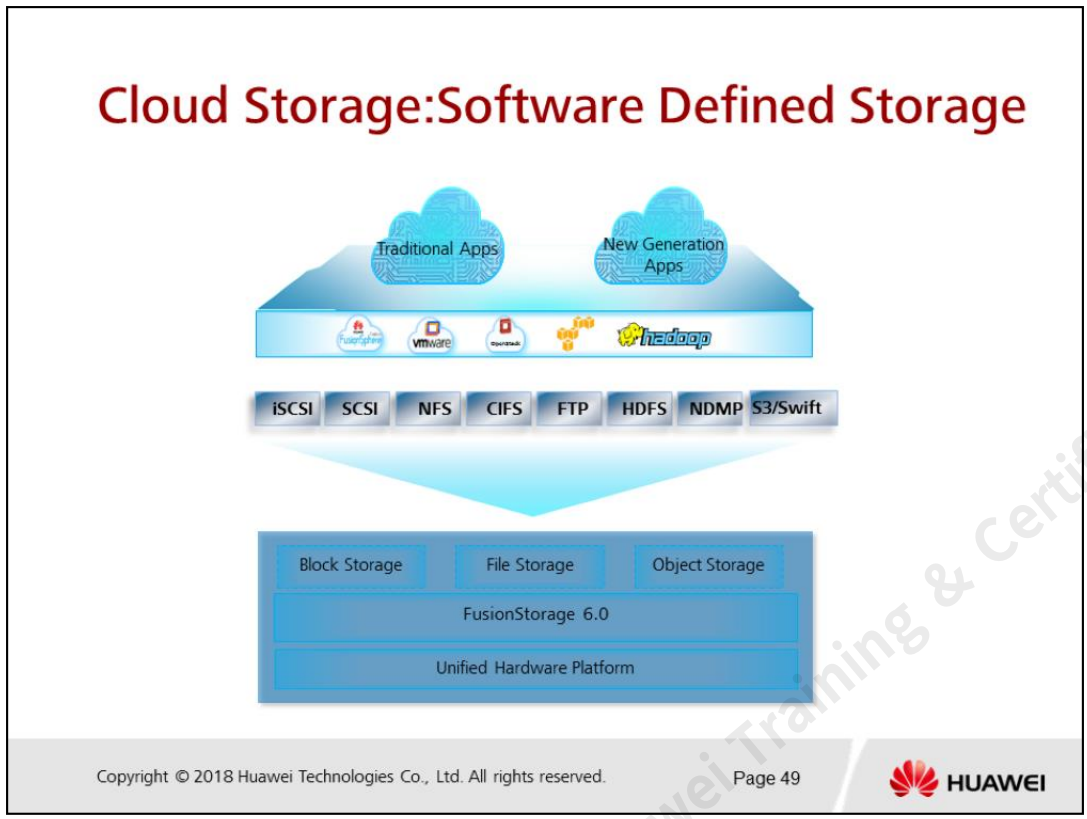
- Highest Speed: 500us stable time delay at millions IOPS, making database faster by 20 times.
- Highest Reliability: Multi Controller Fully Redundant Architecture ensuring no single point of failures in the storage system, all of the components supports online maintenance.
- Highest Savings: Converged storage, online deduplication and decompression saves up to 3 times the storage space and lowers total cost of ownership (TCO) by 70%.
- OceanStor Dorado V3 all-flash storage is the ideal choice for enterprises' mission-critical business. It is the industry's first commercial use of NVMe all-flash storage, and it delivers high-performing, reliable, and efficient storage services.

Huawei Training & Certification



- As a big data storage product, the OceanStor 9000 provides advantages such as large capacity, high performance, and flexible scalability as well as a range of value-added features. It can be applied to various fields such as broadcasting, media and entertainment (M&E), high-performance computing (HPC), Internet operation, data centers, and large-sized enterprises.
- OceanStor 9000 is a Big Data storage platform that uses the general x86 architecture hardware and a scale-out NAS system. OceanStor 9000 adopts a fully symmetric architecture and supports seamless expansion of 3 to 288 nodes.
- The OceanStor 9000 employs a networking mechanism that features full mesh and redundancy, allowing any nodes to concurrently access any data. Multiple nodes can concurrently access different areas of the same data, achieving high-performance data read and write. The OceanStor 9000 provides standard NFS/CIFS/FTP/HDFS/object storage service compatible with Amazon S3 and OpenStack Swift APIs.

Huawei Training & Certification



- FusionStorage Software Defined Converged Cloud Storage is based on standard X86 hardware platform for high performance, high scalability and open compatibility with storage resource pools to provide file/block/object based storage for finance, government, and new businesses.
- The product has the high performance, elasticity and flexibility of distributed storage architecture for meeting agile business needs, and also has the reliability and usability of enterprise grade products to easily face the challenges of globalization, cloudification and big data applications.
- All new products are built based on demands and Cloud Storages are no exception. Enterprise Cloud datacenters has hundreds or even thousands of application that presents data in different service formats and data structures such as structures data and unstructured data. Traditional and New generation applications along with Big Data applications also coexist at the same time in the cloud datacenters.

Huawei Training & Certification

Summary

- This module mainly introduces:
  - The different kinds of common storage system architectures and its differences.
  - The different kinds of storage components that exist within a storage systems and its functions.
  - The different ranges of Huawei Storage Products and briefly highlights the product's key features.

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

Huawei Training & Certification
Huawei Training & Certification

Quiz

1. Which of the following are types of SSD Drive?
   A. SLC
   B. MLC
   C. TLC
   D. NLC
2. What are the key indicators for assessing mechanical hard drive's performance?
   A. Disk Volume
   B. Disk Rotational Speed
   C. Data Transmission Rate
   D. Average Access Time

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

- Answers:
  - ABC.
  - BCD.

Huawei Training & Certification

Thank You

www.huawei.com

Huawei Training & Certification
Huawei Training & Certification

05 Introduction to Common Storage Protocols

www.huawei.com

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

Huawei Training & Certification
Huawei Training & Certification

Foreword

- This module mainly describes the common storage protocols:
  - Definition, Technical Principles, and Applications of SCSI\iSCSI Protocols.
  - Definition, Technical Principles, and Applications of SAS Protocol.
  - Definition, Technical Principles, and Applications of PCIe Protocol.
  - Definition, Technical Principles, and Applications of IB Protocol.
  - Definition, Technical Principles, and Applications of CIFS\NFS Protocols.
  - Definition, Technical Principles, and Applications of FTP\HTTP Protocols.

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

Huawei Training & Certification
Huawei Training & Certification

Objectives

- Upon completion of this module, you will be able to:
  - Describe the definitions of common storage protocols.
  - Understand the technical principles of common storage protocols.
  - Understand the application scenarios of common storage protocols.

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

Huawei Training & Certification
Huawei Training & Certification

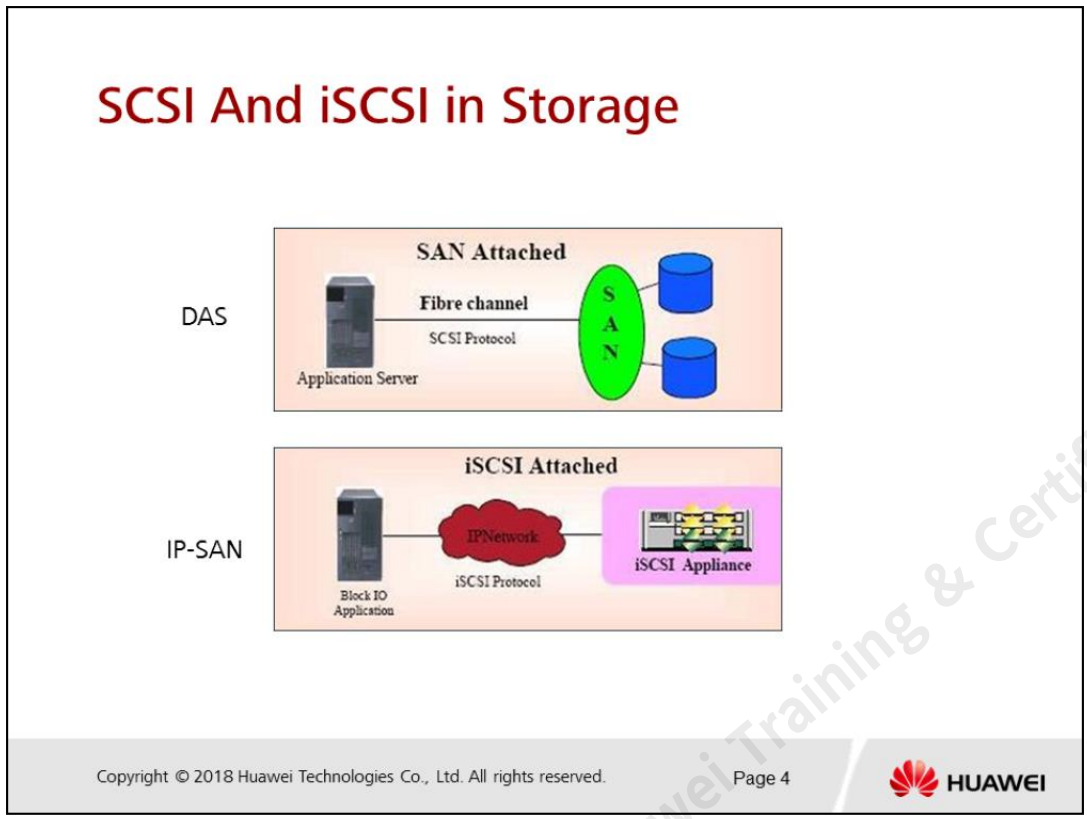
Contents

1. **SCSI/iSCSI**
2. SAS
3. FC/FCOE
4. PCIe
5. IB
6. CIFS/NFS
7. FTP/HTTP

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

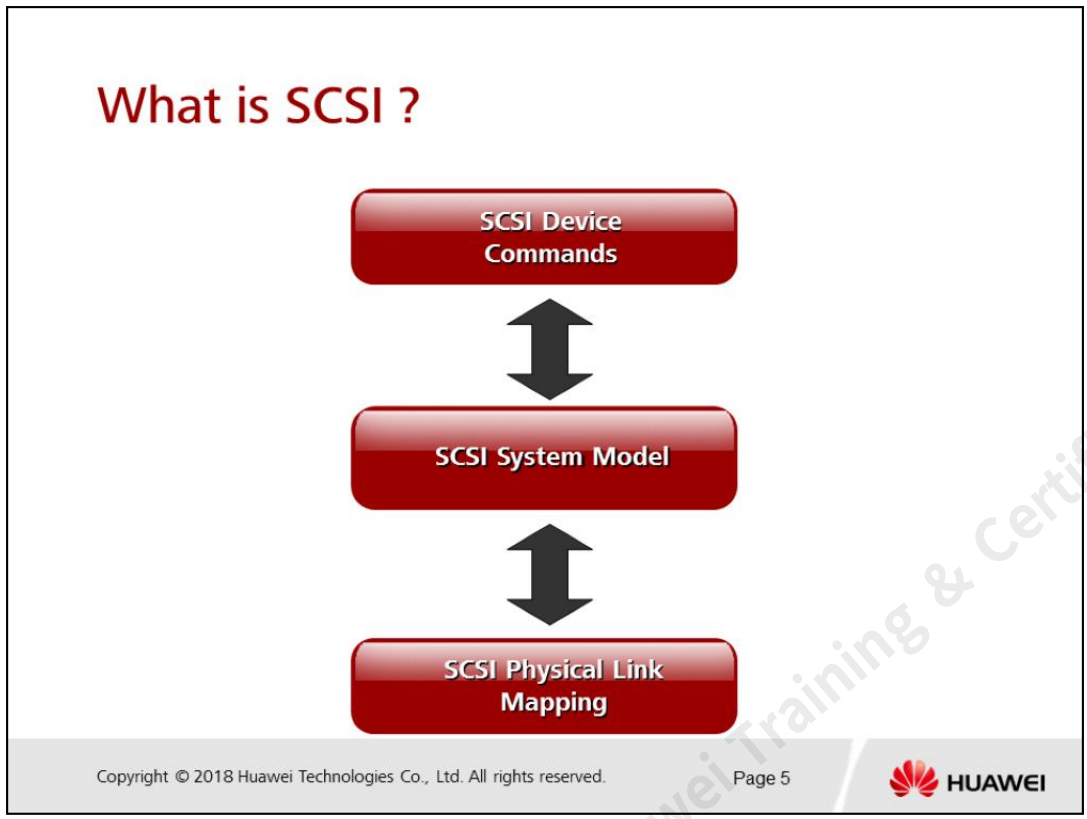
HUAWEI

Huawei Training & Certification
Huawei Training & Certification



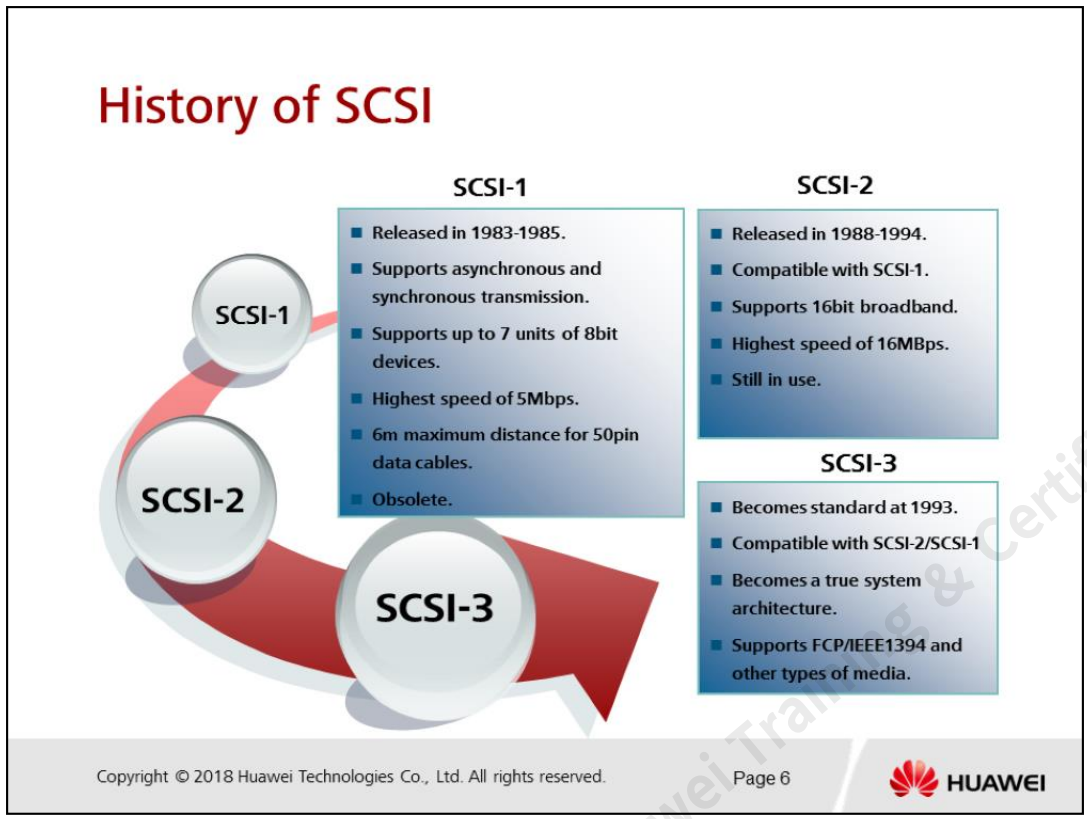
- SAN (Storage Area Network): It is a storage architecture that connects storage device and application servers through network, and this network is solely used for the data access between the host and the storage. When there is a data retrieval request, the data can be transmitted in high speed between the servers and the back end storage devices through the SAN.
- IP SAN: It uses Fast Ethernet/Gigabit Ethernet/ 10 Gigabit Ethernet network to connect the application servers and the storage systems. IP SAN transmits the SCSI commands and data blocks through high speed Ethernet network. It also inherits all the benefits of Ethernet network, and achieves the purpose of building an open, high performance, high reliability, highly scalable storage resource platform.

Huawei Training & Certification



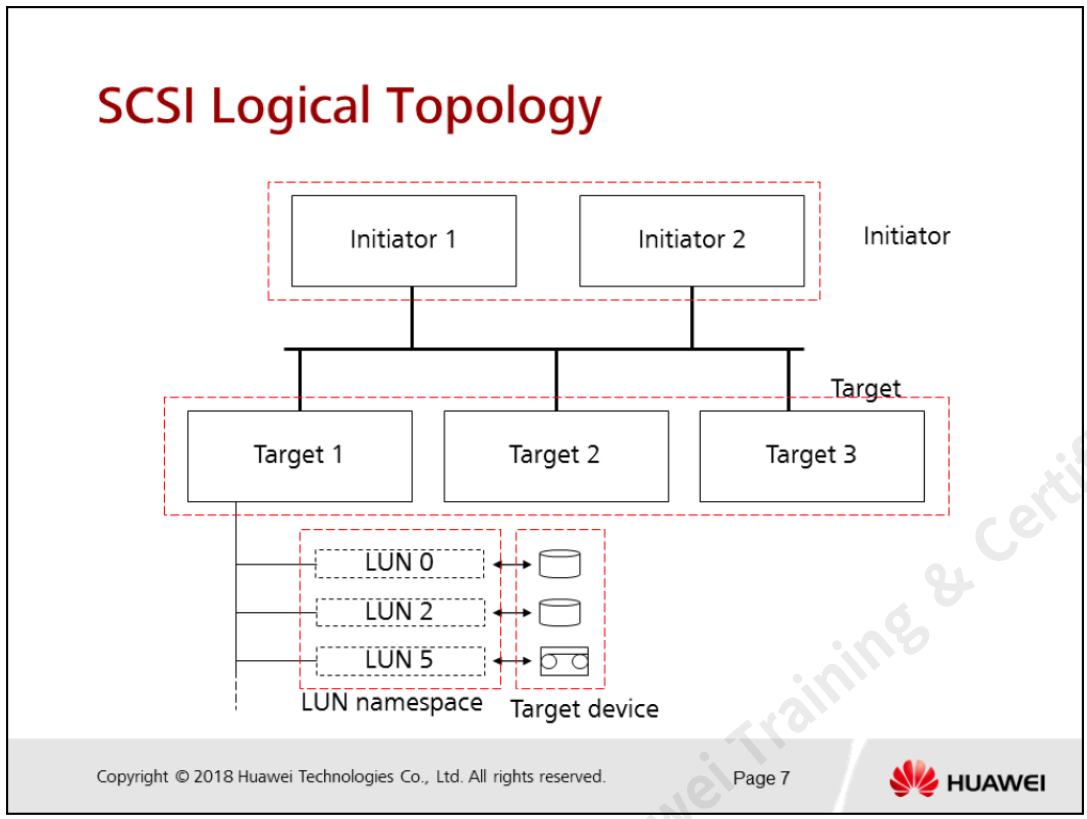
- SCSI, which is Small Computer System Interface, refers to a large protocol system, and has undergone evolution of SCSI-1/ SCSI-2/ SCSI-3 up till now.
- SCSI protocol defines the commands list and the system model for information communication within the architecture for different devices (disks, tapes, processors, optical devices, network devices etc.). This means that all these different devices can communication with each other using the SCSI protocol to work together in order complete tasks such as data read/write or data transmission.
- SCSI protocol in its essence is not related to the transmission media, and SCSI protocol can run on may different types of transmission medium or even virtual medium. For example, iSCSI protocol that is based on IP protocol or FCP (Fibre Channel Protocol) that is based on iSCSI protocol using underlying Fibre Channel connection. This means that SCSI protocol can run in different types of medium such as copper cables, optical cables or even in virtual connections in virtualization environments.

Huawei Training & Certification



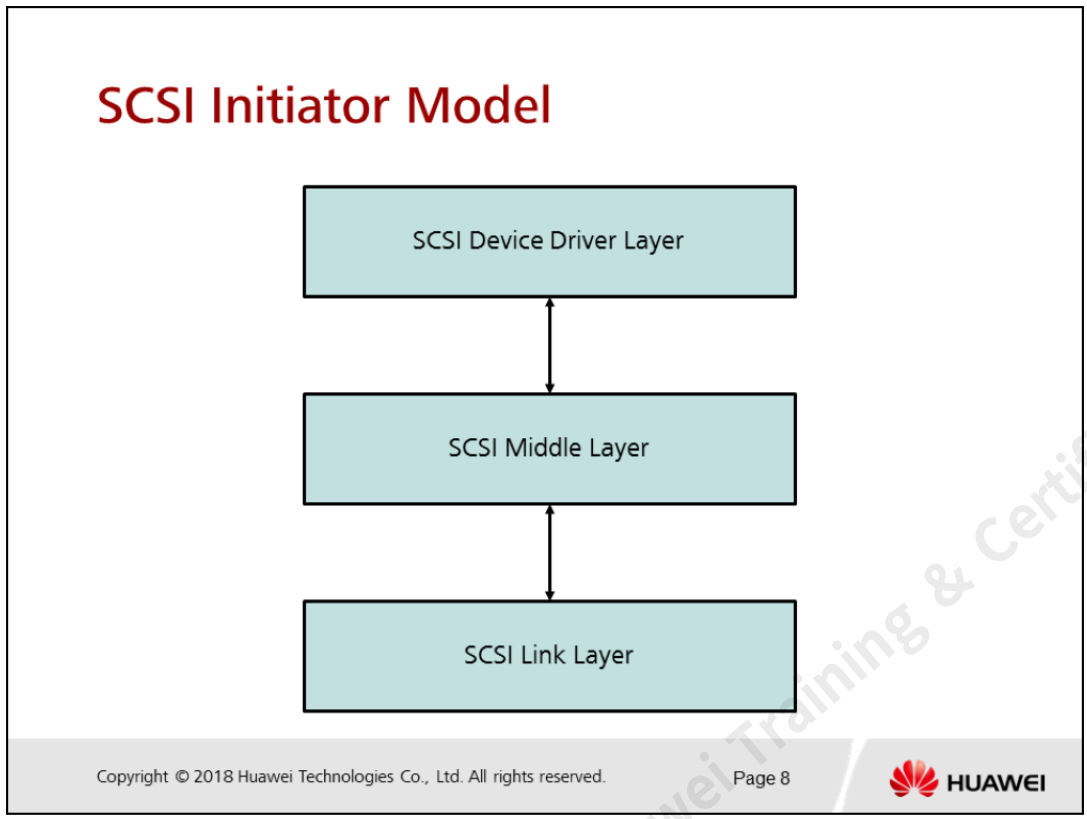
- The diagram above shows the history of SCSI technology ranging from its first version introduced in 1983 to the latest version introduced in 1993. The comparison above shows the changes that SCSI technology has gone through in order to keep up with the advancement of technologies and user requirements.

Huawei Training & Certification



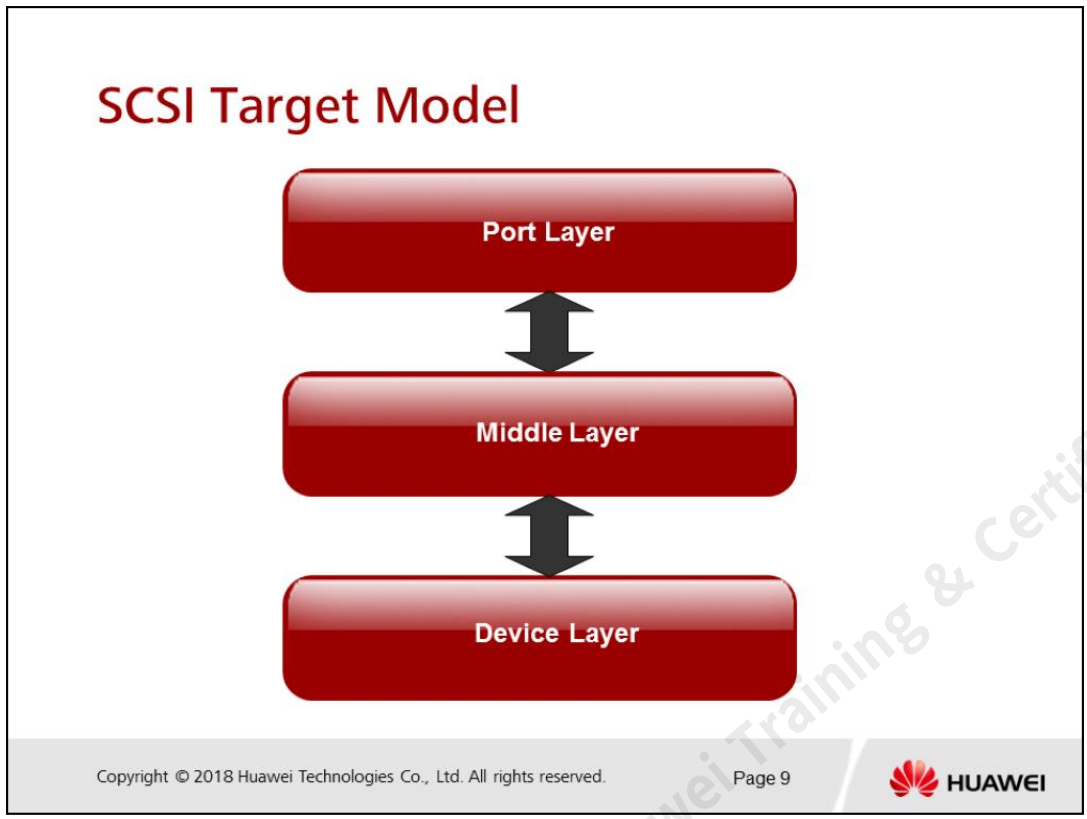
- LUN (Logical Unit Number): LUN refers to the logical storage resource connected to the target. An target can connect to multiple LUN and each LUN can have different properties. LUN#0 can be a disk drive, and LUN#1 can be other storage devices.
- Initiator: In reality, SCSI is a client/server(C/S) architecture, the clients are the initiator which is responsible to send commands to the SCSI target. Commonly, hosts plays the role of the initiators.
- Target: Target is the server end that processes the SCSI commands. It receives the commands from the host(initiator) and then analyze and processes those commands. An example of the target is the role of the storage array.
- SCSI initiator and target forms a standard C/S model, where each of the SCSI commands are run in "request/response" model.
    - Main functions of Initiator: Sends SCSI requests.
    - Main functions of Target: Respond to SCSI requests, and provide storage services through the LUN, and provide task management capabilities through the task processor.

Huawei Training & Certification



- The SCSI system in the host usually operates under the "Initiator" mode. From the perspective of SCSI system architecture, it is divided into "Middle Layer", "Devices Layer", and "Transmission Layer". Thus common operating systems including Windows, Linux, AIX, Solaris, BSD also divides SCSI into 3 layers.
- The initiator architecture under Windows: Windows separates SCSI initiator into 3 logical layers, in which SCSI Port is responsible for the basic SCSI processing such as device discovery, namespace scanning etc. To understand more details on the Windows drivers, you may refer to the DDK (Driver Development Kit) or WDM (Windows Driver Model) documentations.
- The initiator architecture under Linux: Linux also divides the SCSI initiator into 3 logical layers, in which scsi mod middleware layer handles the complex processing of process that doesn't involve SCSI devices and adapters such as abnormalities and namespace maintenance etc. On the other hand, HBA driver provides the encapsulation and decapsulation and transmission of SCSI commands. Device driver contains certain SCSI devices drivers such as some of the well known drivers for SD (SCSI Disk), ST (SCSI Tape) and SR (SCSI CDROM) etc.
- The initiator architecture under Solaris: The architecture for Solaris is similar to Linux/Windows, so we will not explain it further here.
- The initiator architecture under AIX: The AIX architecture is also divided into 3 layers which are SCSI Device Driver Layer, SCSI Middle Layer and SCSI Adapter Driver Layer.

Huawei Training & Certification Huawei Training & Certification

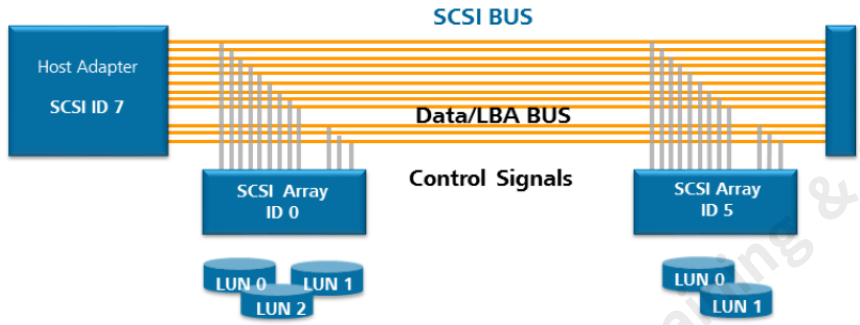


- Target is also divided into 3 layers which are Port Layer, Middle Layer and Device Layer based on SCSI system architecture. The most important layer is the Middle Layer which is designed using SAM(SCSI Architecture Model)/SPC (SCSI Primary Commands) specifications. The middle layer manages and maintains the LUN namespace, link ports, target device, tasks, commands and communication. Port layer drivers are loaded in registered mode whereas the device layer drivers are dynamically loaded.
- PORT Model in Target: PORT drivers are dynamically loaded, and the main functions of PORT is to complete the encapsulation and decapsulation of the SCSI commands carried over the link/connection. For example, encapsulating the SCSI commands into FPC or iSCSI or SAS packets, or decapsulate the FCP/iSCSI/SAS packets into SCSI commands. The target model drivers for iSCSI/FCP/SAS hardware belongs to the PORT category. The functions provided by the PORT includes transmitting responses (xmit_response), transmitting data (xfer_data), notification on completed management tasks, ending the processing task (cmd_done), port control (reset control) etc.

Huawei Training & Certification Huawei Training & Certification

SCSI Protocol and Storage System

- SCSI protocol is the fundamental protocol between hosts and storage disks.
- DAS(Direct Attached Storage) uses SCSI protocol for host and storage device interconnection.

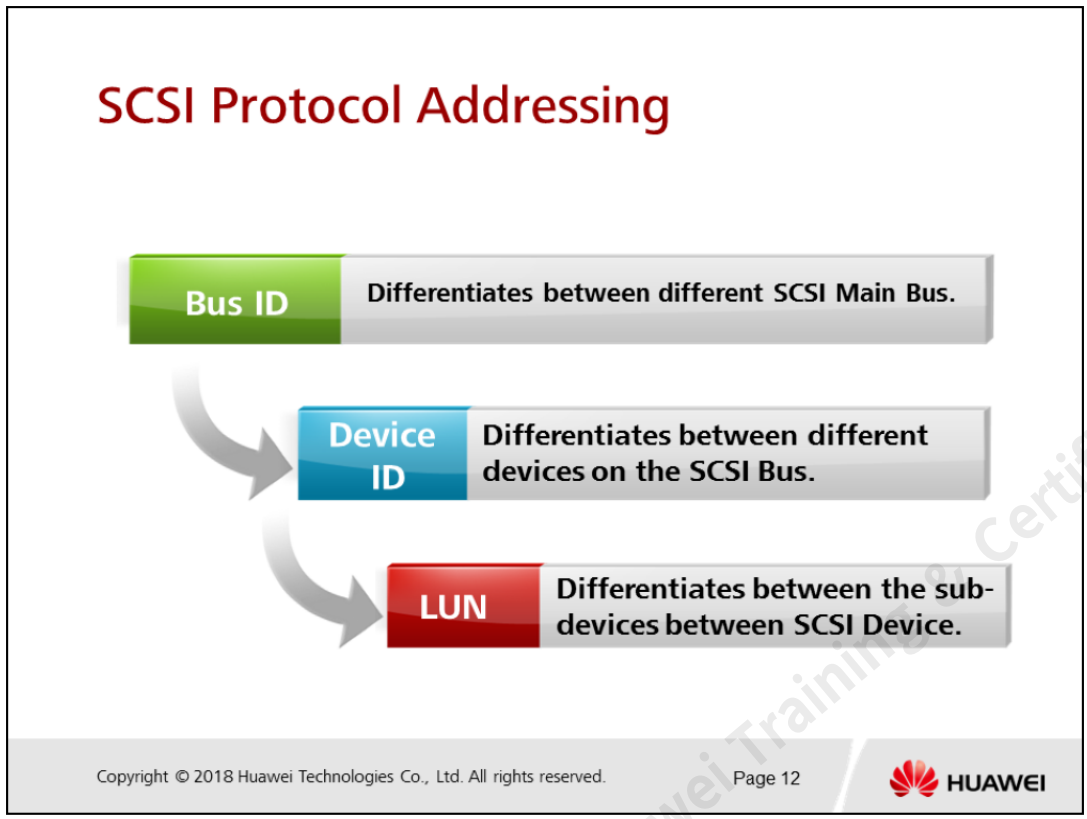


Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

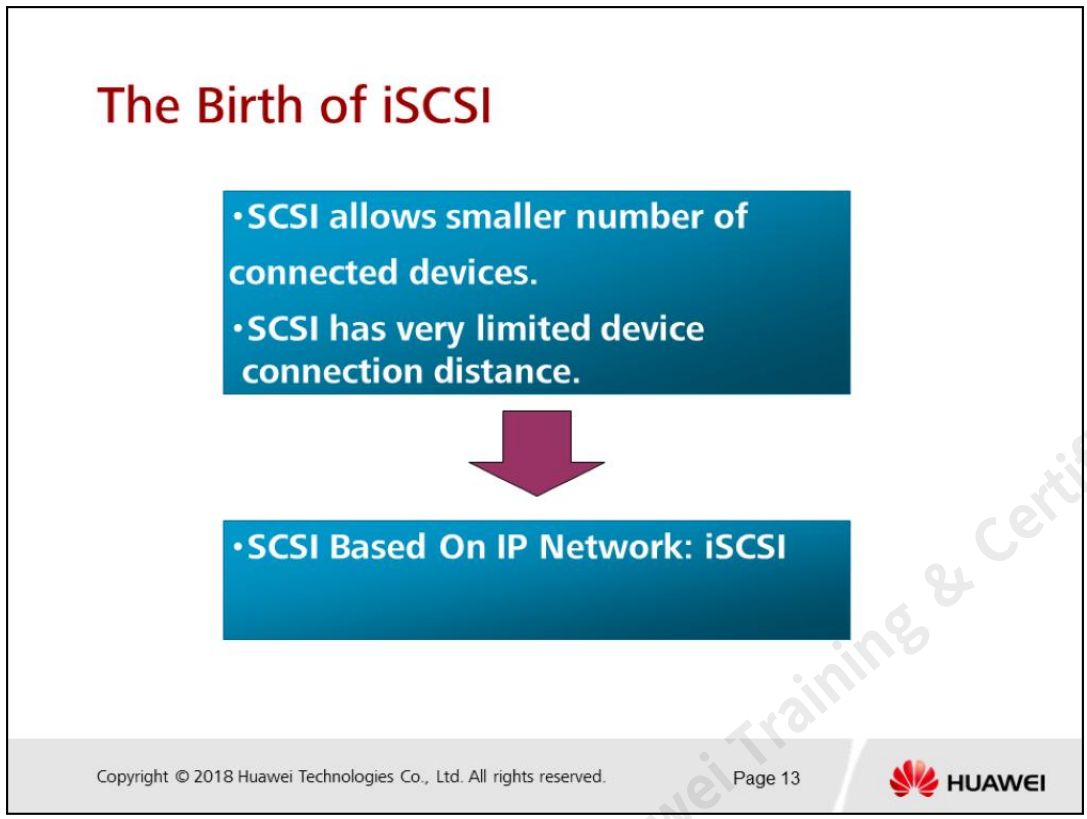
- The storage controller send a request signal to the Bus processor to use the Main Bus. When the request is accepted, the storage controller cache will begin to perform the send operation. In this process, the controller utilizes the main bus, and the other devices cannot use the main bus during that period. However, if the main bus has the Interrupt function, the main bus controller can stop the sending or transmission process at any time and pass the main bus control to other devices for a much higher priority operation.
- SCSI controller is similar to a small CPU, and has its own command sets and cache. SCSI is a special bus architecture, and can perform dynamic task allocation and task operation for multiple devices in the computer at the same time. This means that it can dynamically allocate and dynamically complete multiple tasks required by the system at the same time.

Huawei Training & Certification



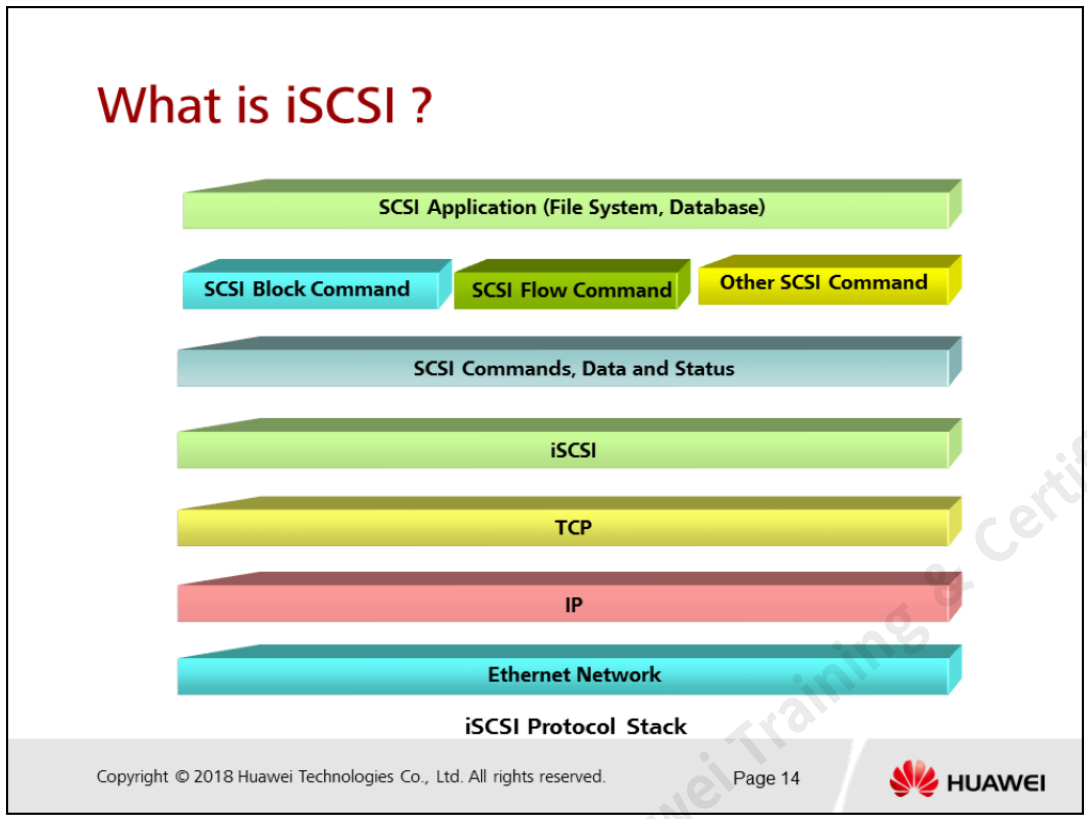
- In order to address the devices connected to the SCSI main bus, SCSI protocol introduced SCSI Device ID and LUN(Logical Unit Number). In the SCSI main bus, each device must have a unique device ID, and the HBA(Host Bus Adapter) in the servers also have their own device ID which is fixed at 7 by default. Each bus allows a maximum of 8 or 16 device IDs including the HBA. Device ID is not only used for addressing, but also used to identify the device priority on the bus. Additionally, all the different devices connected to the main bus must have different device ID to avoid addressing and priority conflicts.
- Each storage device includes multiple sub devices such as disks, tape drives etc. Hence, SCSI protocol introduced the concept of LUN ID for the addressing of all these sub devices within the storage devices.
- Traditional SCSI controller connects to a single bus, and only has a single bus ID. An enterprise level server may be configured with multiple SCSI controller, and thus have multiple SCSI bus. When SAN is introduced, each FC HBA or iSCSI network cards are connected to a bus, hence each bus needs to be assigned with a bus ID for differentiation between different buses. We can use a three element description to identify or label a SCSI target in the form of : Bus ID/Target Device ID/LUN ID.

Huawei Training & Certification



- iSCSI protocol was initially introduced by IBM, CISCO and HP. It became the IETF (Internet Engineering Taskforce) standard by 2004, and the current iSCSI protocols are based on SAM2(SCSI Architecture Model 2).
- iSCSI (Internet SCSI) encapsulates the SCSI commands and data blocks into TCP packets and transmit them over the IP network.
- iSCSI became the SCSI transmission level protocol, and its initial starting point is to use mature IP network technologies to implement and extend Storage Area Network (SAN) capabilities.

Huawei Training & Certification

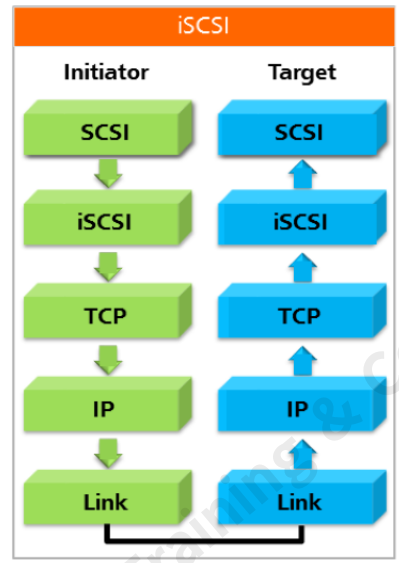


- iSCSI is the abbreviation of Internet Small Computer System Interface, and it is a standard in data blocks transmission over TCP/IP. It can be considered as SCSI over IP.
- iSCSI can be used to build IP based SAN to provide high speed, low cost, long distance storage solutions.
- iSCSI encapsulates the SCSI commands into TCP/IP packets which allows I/O data blocks to be transmitted over IP network.
- iSCSI is the mapping of SCSI remote procedure call to the TCP/IP protocol. SCSI layer is responsible to generate the CDB (Command Descriptor Block) and sends it to the iSCSI protocol later where it is further encapsulated into PDU(Protocol Data Unit) and transmitted over the IP network.

Huawei Training & Certification

iSCSI Initiator - Target Model

- Initiator
  - SCSI layer is responsible to generate the CDB (Command Descriptor Block) and sends it to iSCSI.
  - iSCSI is responsible to generate iSCSI PDU(Protocol Data Unit), and send it to the target over IP Network.
- Target
  - iSCSI layer receives the PDU and sends the CDB to the SCSI layer.
  - SCSI layer is responsible for interpreting the meaning of the CDB and send responses if required.

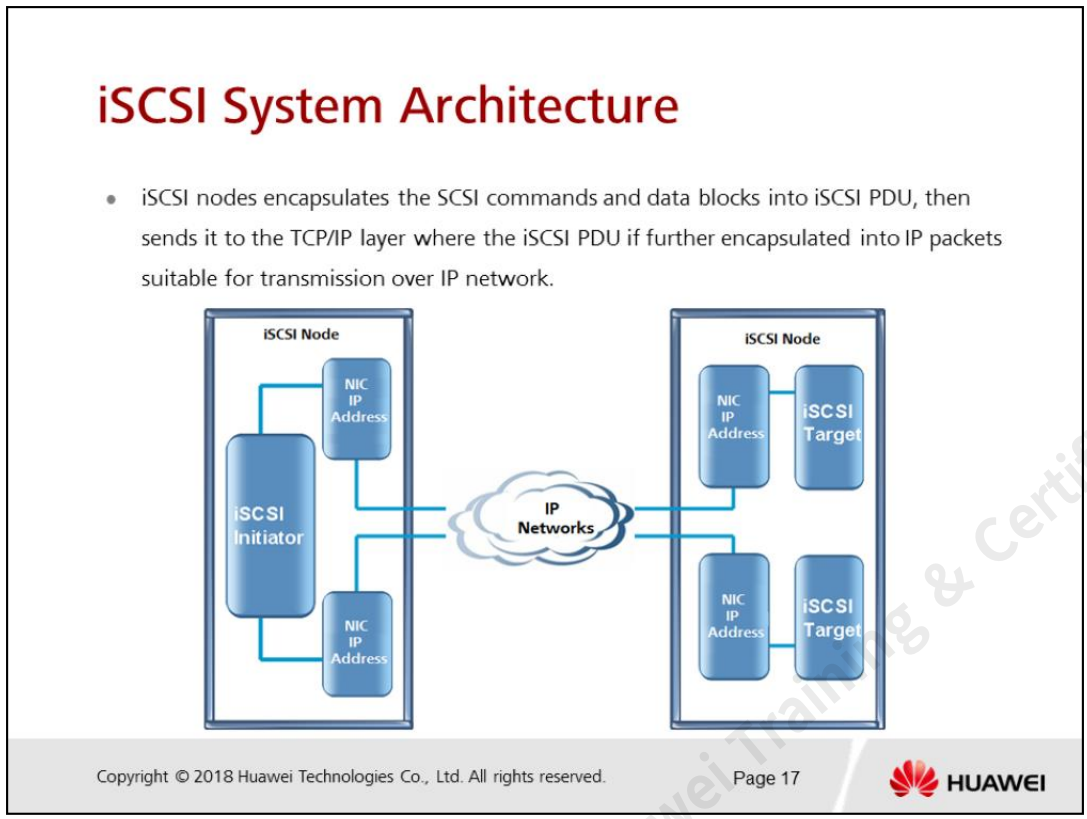


Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

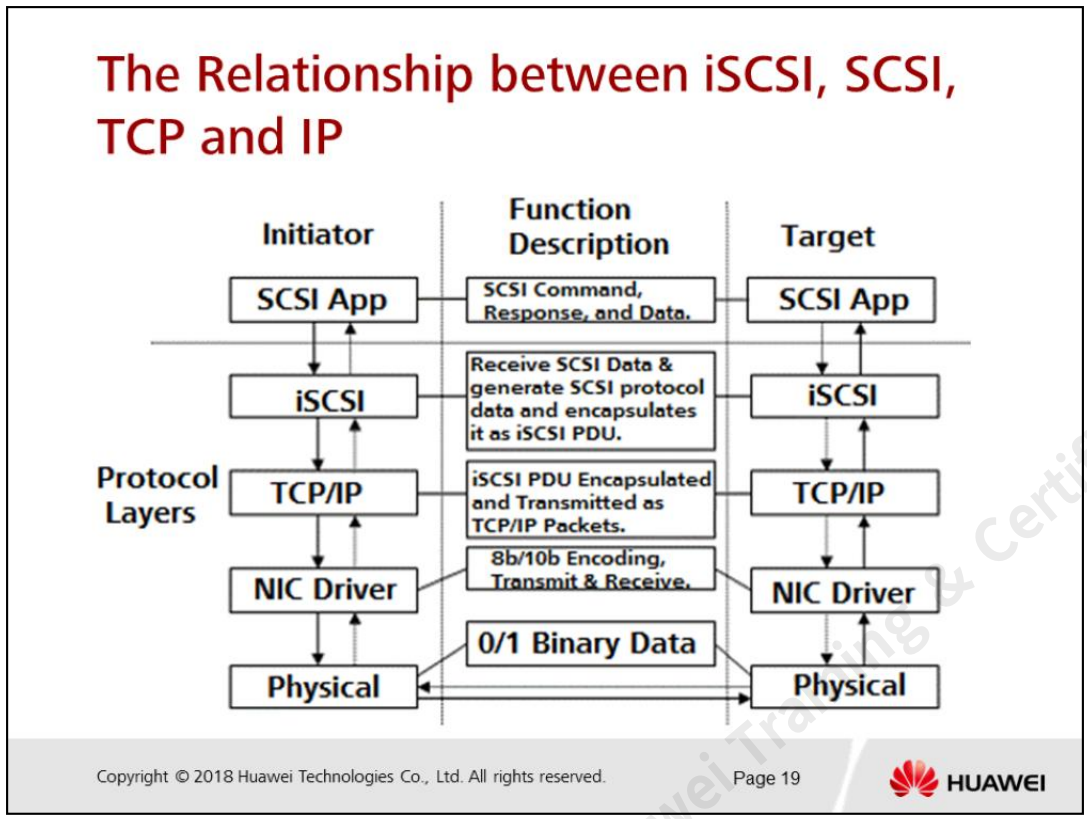
- iSCSI communication system still inherits certain characteristics of SCSI. In iSCSI communication, there is a Initiator device that sends the I/O request, and a Target device that responds and execute the actual I/O operations based on the I/O requests received. Once the Initiator and Target establish a connection, the Target will be the main device that controls the whole working process.
- iSCSI Initiator: It can be divided into 3 types, which are software Initiator driven program, hardware TOE (TCP Offload Engine) card and the iSCSI HBA card. Judging from performance, the software initiator has the lowest performance, TOE card has medium performance and iSCSI HBA cards has the best performance.
- iSCSI Target: The target are usually the iSCSI disk arrays or iSCSI tape libraries.
- iSCSI protocol has defined a set of naming and addressing method for Initiator and Target. All the iSCSI nodes are identified by iSCSI names. This naming convention avoids iSCSI names and host names to be confused with each other.

Huawei Training & Certification



- In a system that supports iSCSI, the users sends out instructions to send or fetch data from a SCSI storage device, the operating system processes that request and converts it into one or many SCSI commands which are then send to the target SCSI controller card. iSCSI node will do the encapsulation process on the SCSI commands and data blocks which forms a iSCSI PDU, which is sent to the TCP/IP layer. The TCP/IP layer will encapsulate the iSCSI PDU into the IP packets suitable for transmission over the IP network. It can also do secure encryption on the encapsulated SCSI commands, then send it over unsecure networks.
- Packets can be transmitted over the local area network or Internet. On the receiving storage controller, the data will be recombined and the storage controller will interpret the SCSI commands and data contained within those packets and send them to the corresponding disk drives. The disk drives will then execute the functions initially requested by the host or application. If the received request is for data retrieval, then the data will be fetch from the disks, encapsulated and transmitted to the requesting computer. This whole process is transparent to the user.
- Although the SCSI commands execution and data preparation can be completed by standard TCP/IP or current network card software, there are some caveats where if the encapsulation and decapsulation of the packets process are done by software, it will take a toll on the CPU and impact performance. This is due to the fact that host CPU will require a lot of CPU computation cycles to process all the data and the SCSI commands.

Huawei Training & Certification



- PDU: Protocol Data Units. A protocol data unit is information delivered as a unit among peer entities of networks containing control information, address information or data. In layered systems, PDU represents a unit of data specified in the protocol of a given layer, which consists of protocol control information and user data. PDU is a significant term related to the initial four layers of the OSI model. In Layer 1, PDU is a bit, in Layer 2 it is a frame, in Layer 3 it is a packet and in Layer 4 it is a segment. In Layer 5 and above, PDU is referred to as data.
- TCP/IP: Transmission Control Protocol/ Internet Protocol. In simple words, it is the language a computer uses to access the internet. It consists of a suite of protocols designed to establish a network of networks to provide a host with access to the internet. TCP/IP is responsible for full-fledged data connectivity and transmitting the data end to end by providing other functions, including addressing, mapping and acknowledgment.
- 8b/10b encoding takes a group of continuous 8 bit of data and then splits it into 2 sets of data, the first set with 3 bits and the second set of 5 bits of data. After encoding, it forms a set of 10 bit transmission character with a group of 4bit character and another group of 6 bits of character, which are transmitted over the network.
- 8b/10b encoding: It is an algorithm for encoding data for transmission in which each eight-bit data byte is converted to a 10-bit transmission character. Invented and patented by IBM Corporation. 8b/10b encoding is used in transmitting data on Fibre Channel, ESCON, and Gigabit Ethernet. 8b/10b encoding supports continuous transmission with a balanced number of ones and zeroes in the code stream and detects single bit transmission errors.

Huawei Training & Certification

Contents

1. SCSI/iSCSI
2. **SAS**
3. FC/FCOE
4. PCIe
5. IB
6. CIFS/NFS
7. FTP/HTTP

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

Huawei Training & Certification
Huawei Training & Certification

SAS in Storage System

- In enterprise storage systems, SAS (Serial Attached SCSI) interface has already replaced SCSI and SATA interface.
- SAS uses point to point architecture, and its performance can go up to 300MB/s, 600MB/s or higher.



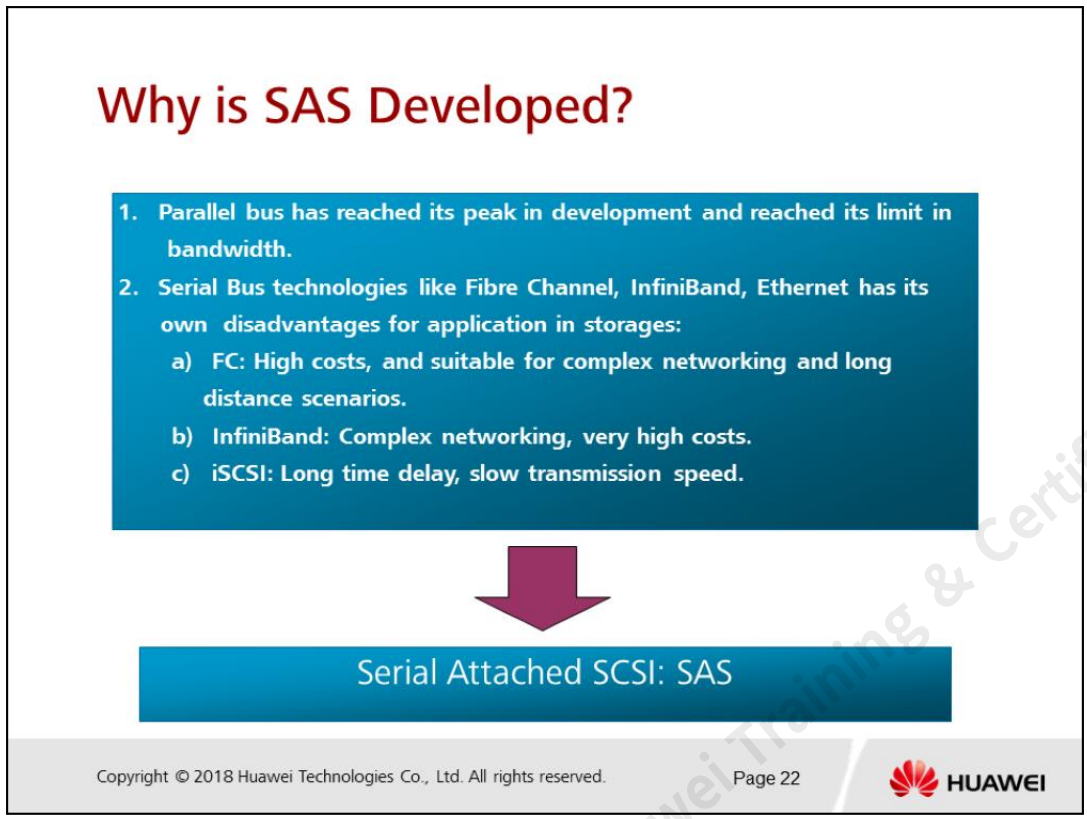


Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

- SAS uses a point-to-point connection design to enable a dedicated link to be established between two devices in the communication. The multi point bus design in Parallel SCSI (often called as SCSI) allows multiple devices share the same bus, which means that the more devices are connected, the performance gets worse. With point-to-point connections in SAS, the communication speed is much faster because it is not necessary to detect whether the two devices communicating are allowed to use the connection link before communication. Each device is connected to the specified data path to increase the bandwidth.
- The serial interface has a simple structure, supports hot swap, has high transmission speed, and high operational efficiency. Commonly, large parallel cables can cause electronic interference but the SAS cable structure can solve this problem. The SAS cable structure saves space and improves the heat dissipation and ventilation capabilities of SAS hard disk based servers.

Huawei Training & Certification

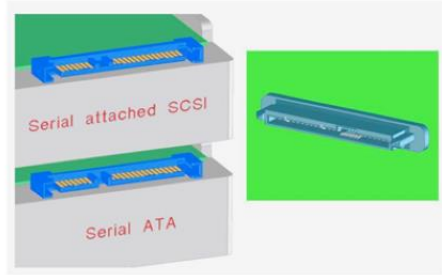


- SAS is the abbreviation of Serial Attached SCSI, which is a serial connection form of SCSI.
- SAS is the serial standard for SCSI bus protocol.

Huawei Training & Certification

What is SAS?

- SAS: Serial Attached SCSI, is the serial standard for SCSI bus protocol.
- SAS uses serial technology to obtain higher transmission rate and better expansion capability, and it is compatible with SATA hard disks.
- SAS current transmission rates are up to 3Gbps, 6Gbps, 12Gbps or higher, and supports full-duplex mode connection.

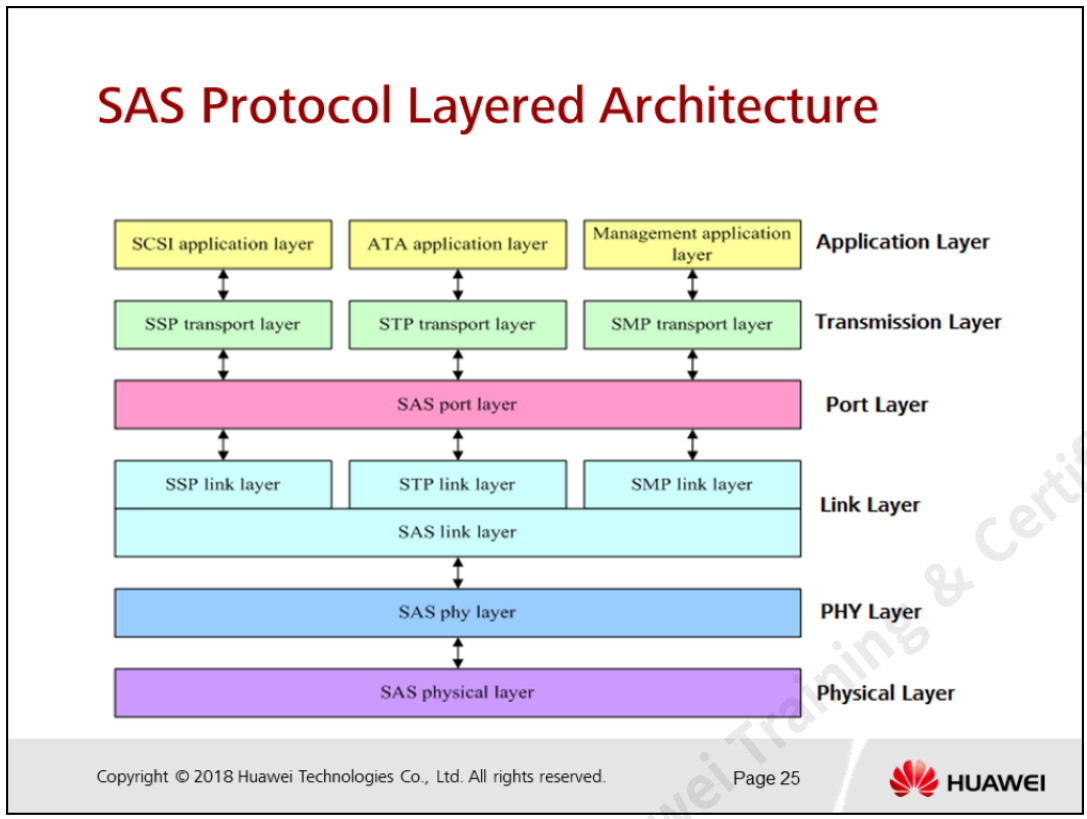


Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

- Lower Costs:
  - SAS backplane is compatible with both SAS and SATA hard disks, which means that there is no reinvestment required when using different types of hard disks and compatible with the existing SAS and SATA hard disks within the enterprise.
  - Lowers the cost of designing different products for SCSI and SATA standards, and at the same time lowers the complexity in cabling and layers in PCB(Printed Circuit Board) which saves cost in manufacturing.
  - System Integrators do not need to purchase different backplanes and cables to configure the different types of hard disks in the client environment.
- Connects More Devices:
  - SAS technology introduces SAS expanders that allow SAS systems to connect more devices, each of these allows multiple ports to be connected, and each port can connect to SAS devices, hosts, or other SAS expanders.

Huawei Training & Certification

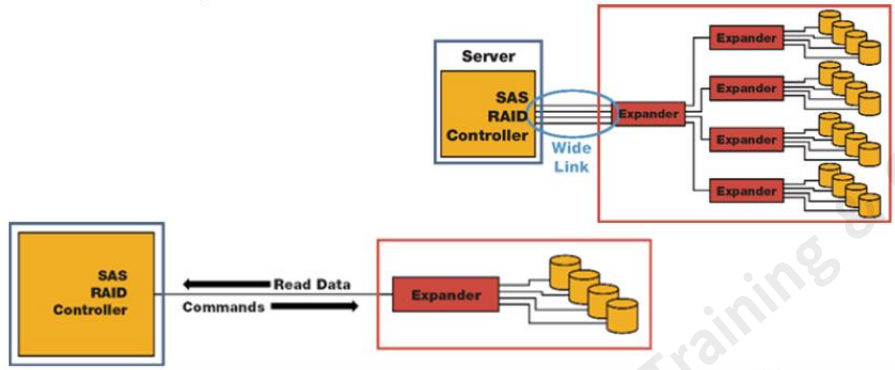


- SAS standard has divided the SAS architecture into 6 layers, which are in the order of lowest layer to the highest layer : physical layer, sas phy layer, link layer, port layer, transmission layer and application layer. Each layer is responsible for certain functions:
  - Physical Layer: Defines the hardware such as cables, interfaces, and transmitters.
  - SAS PHY Layer: Includes the low level protocols, data encoding (8b10b), signaling (Out of Band, OOB) and connection reset sequences.
  - Link Layer: Describes how to control phy layer connection management, as well as primitives, CRC checksum and descrambling, and transmission rate matching & processing.
  - Port Layer: Describes the interface between the link layer and the transmission layer, including how to request, break and building the connection.
  - Transmission Layer: Defines how to transmit all the commands, status and data encapsulation in the SAS frames and also how to perform decapsulation of the SAS frames.
  - Application Layer: Describes the details on how to uses SAS for different types of applications.

Huawei Training & Certification

SAS Features

- SAS serial communication method allows multiple devices in multiple data paths to communicate at full speed.
- SAS supports wide link that is formed through the binding of multiple narrow links.
- SAS architecture uses expander for interface expansion, and has very good expansion capabilities.
- SAS uses Full-Duplex Mode in data communication.



Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

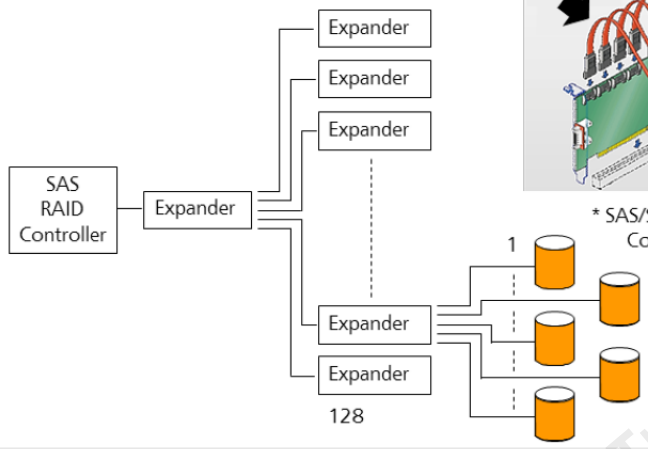
HUAWEI

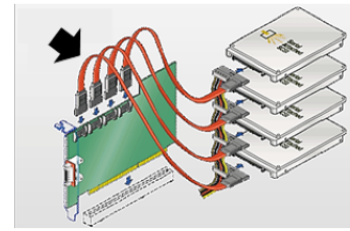
- Compared to SCSI, the main improvements of SAS are:
  - Provides higher throughput due to the usage of serial communication, and has possibility for higher performance in the future.
  - 4 narrow links can be bonded together to form a Wide Link and provide higher throughput.
- SAS uses a full duplex (two-way) communication mode instead of one-way communication. Traditional parallel SCSI can only communicate in one direction. In parallel SCSI scenario, when the device receives a packet, if the device wants to respond to the packet, it needs to reestablish a new SCSI communication link after the previous link is disconnected. With SAS, two-way communication is possible. Each SAS cable has 4 cables which is used for 2 inputs and 2 outputs. Thus, SAS can read and write data at the same time, and full-duplex data operations improve data throughput efficiency.

Huawei Training & Certification

SAS Expansion Capabilities

- SAS architecture uses expander to perform interface expansion, and has very good expansion capabilities. 1 SAS domain can connect up to the maximum of 16384 units of storage devices.





* SAS/SATA Standard Single Port Connectors And Cables.

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

- SAS Expander: An interconnection device in the SAS domain similar to Ethernet switches. Through the cascading of Expanders, it can greatly increase the number of connected terminal devices, in which HBA(Host Bus Adapter) costs can be saved. Each expander can connect up to a maximum of 128 terminal devices or 128 expanders. A SAS domain consists of the following components which are: SAS Expanders, Terminal Devices and Connection Devices (SAS connection cables).
  - SAS Expander is equipped with address routing table for tracking and recording of the addresses of all the SAS drives connected.
  - Terminal Devices includes initiators (commonly known as SAS HBA cards) and targets (SAS/SATA disks or HBA cards in the Target mode).
- Loops cannot be formed with the expanders within the SAS domain to ensure the normal operation of the device discovery process.
- In actual use, due to the bandwidth factor, the connected terminal devices in the expanders are often lesser than 128 devices.
-

Huawei Training & Certification

Principles of SAS Cable

- SAS cables commonly has 4 paths, and the speed of each path commonly used is 12GB/s.
- SAS devices are connected in the form of loops (also known as chains).
- The bandwidth of the cable is 4X12GB/s, which limits the number of disk devices in the SAS loop.
- The current best practice for the maximum number of hard disks in a SAS loop is 168 disks, which means a maximum of 7 hard disk enclosures with 24 disk slots each to form a loop.



Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

Huawei Training & Certification

- Most current storage device vendors use SAS cables to connect the disk enclosures to the controller enclosures or use it for interconnection between disk enclosures. SAS cables usually bundle four separate channels (narrow ports) into one wide port to provide more bandwidth. Each of the 4 independent channels can operate at 12 Gb/s, so the entire wide port can provide 48 Gb/s of bandwidth. To ensure that the amount of data on the SAS cable does not exceed the maximum bandwidth of the entire SAS cable, we need to limit the total number of hard disks connected to a SAS loop.
- For Huawei devices, the maximum number of hard disks in a SAS loop is 168 disks, which means that up to seven disk enclosures with 24 disk slots can form a SAS loop. However, the prerequisites condition is that all the hard disks in the loop are traditional SAS hard disks. Now that SSDs are more commonly used, we must realize that SSDs are much faster than SAS disks in terms of speed. Therefore, for the SSD disks, the best practice for the maximum number disks in a loop is 96 disks, which is a loop consisting of 4 disk enclosures with 24 disk slots.
- The SAS cable interface is called a Mini SAS cable when the SAS cable is a single-channel 6Gb/s cable type. Now that the single-channel speed is increased to 12Gb/s, the corresponding SAS cable is now called a high-density Mini SAS cable.

Comparison Between SAS And Other Transmission Technologies

Types of Technology	Main Advantages	Main Disadvantages	Fields Of Application
ATA	Low Cost	Low Performance	PC
SCSI	Medium Performance	Drawbacks of Parallel Technologies	Enterprise Grade Storages
FC	High Performance, High Reliability	High Cost	High End Storages
SATA	Low Cost, High Capacity	Low Performance and Low Reliability	Mid and Low End Storages
SAS	High Performance, High Reliability	Medium Cost	Mid and Low End Storages

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

- The table above shows the comparison different types of transmission technology with SAS with the main advantages and disadvantages highlighted.
- SAS has high performance and reliability and comes with a medium cost range which is suitable for mid tier and low tier storage systems.
- Although FC has the highest performance, but the higher costs of implementation makes it not the cost effective choice for small and medium enterprises.

Huawei Training & Certification

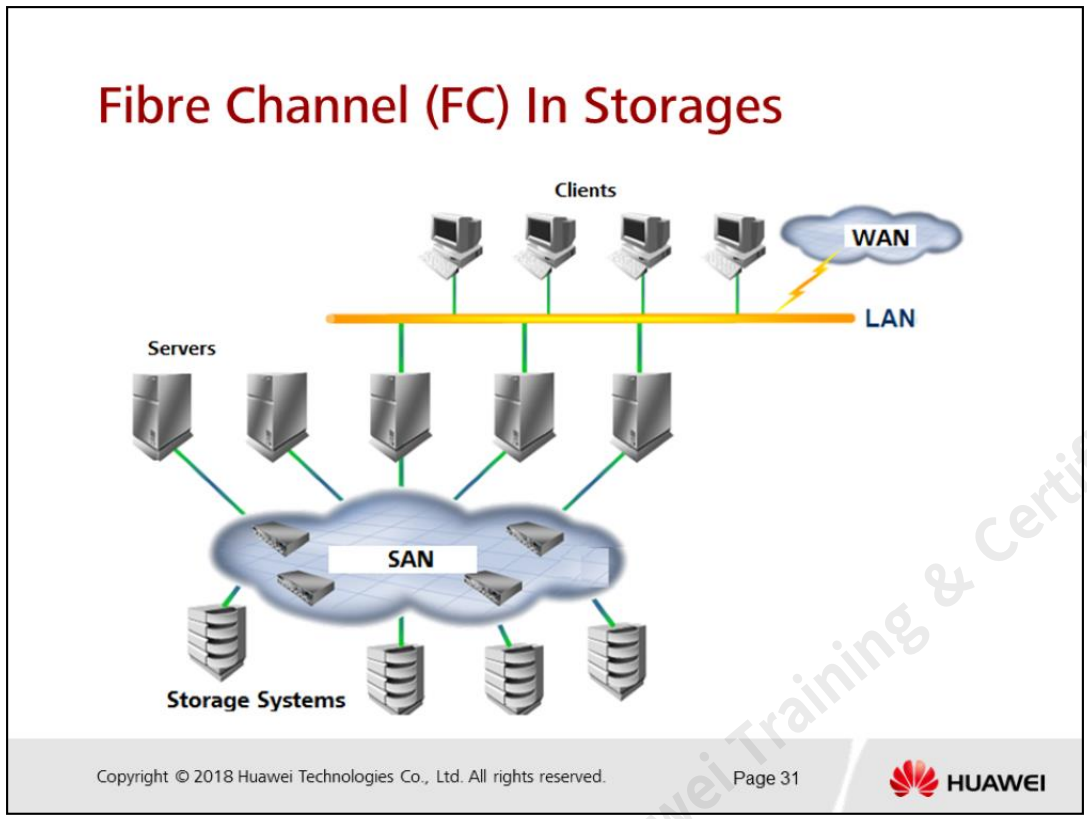
Contents

1. SCSI/iSCSI
2. SAS
3. **FC/FCOE**
4. PCIe
5. IB
6. CIFS/NFS
7. FTP/HTTP

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

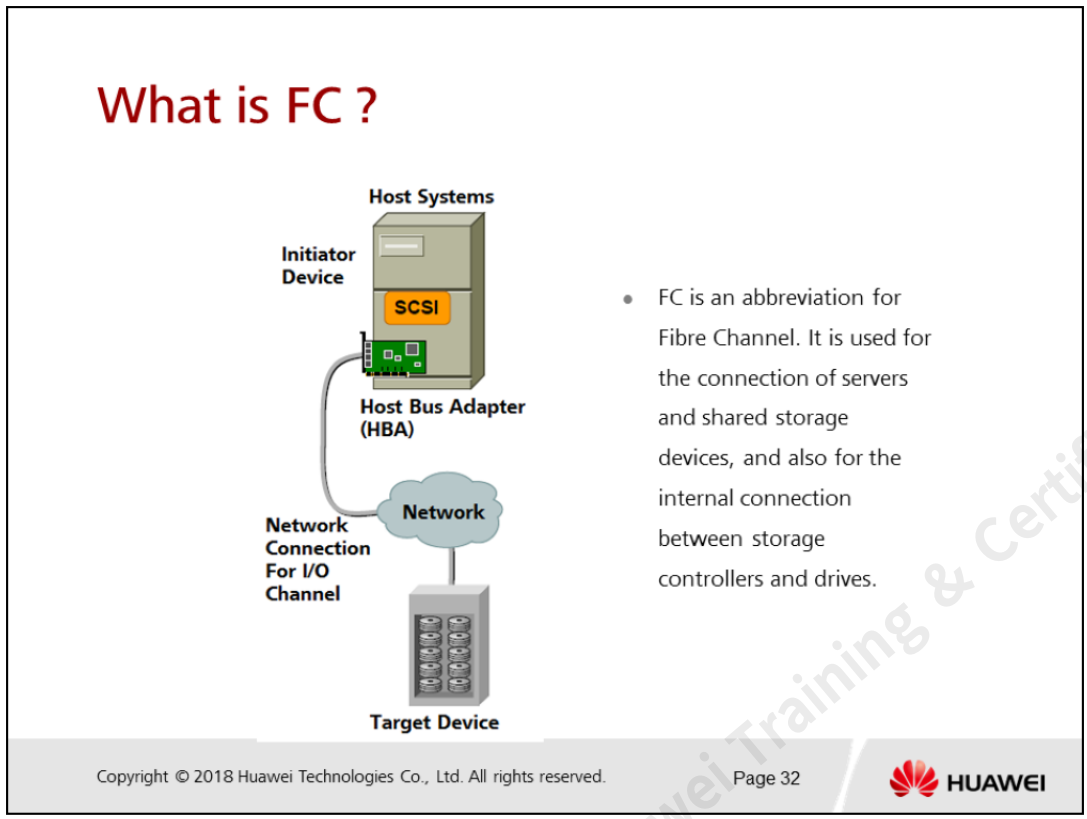
HUAWEI

Huawei Training & Certification
Huawei Training & Certification



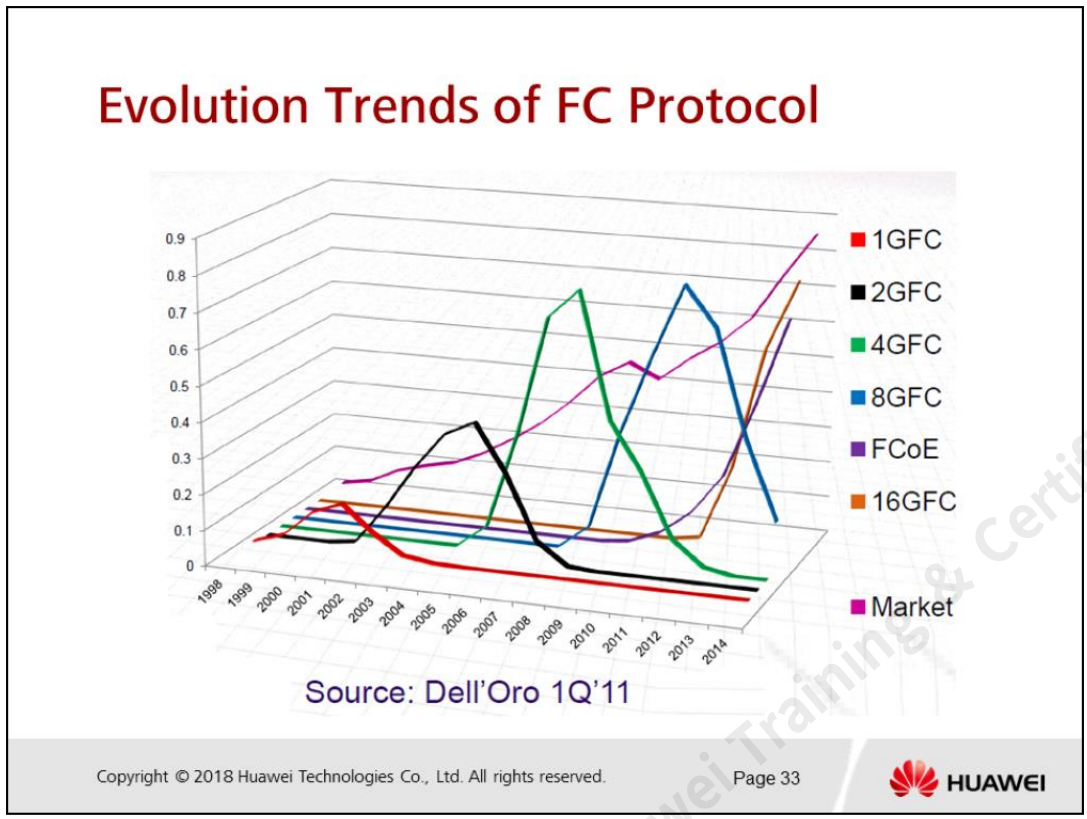
- Fabric is a term used to refer to an intelligent network composed of intelligent Fibre Channel switches with good system design. This network can provide enterprise-level performance, scalability, manageability, reliability, and availability.

Huawei Training & Certification



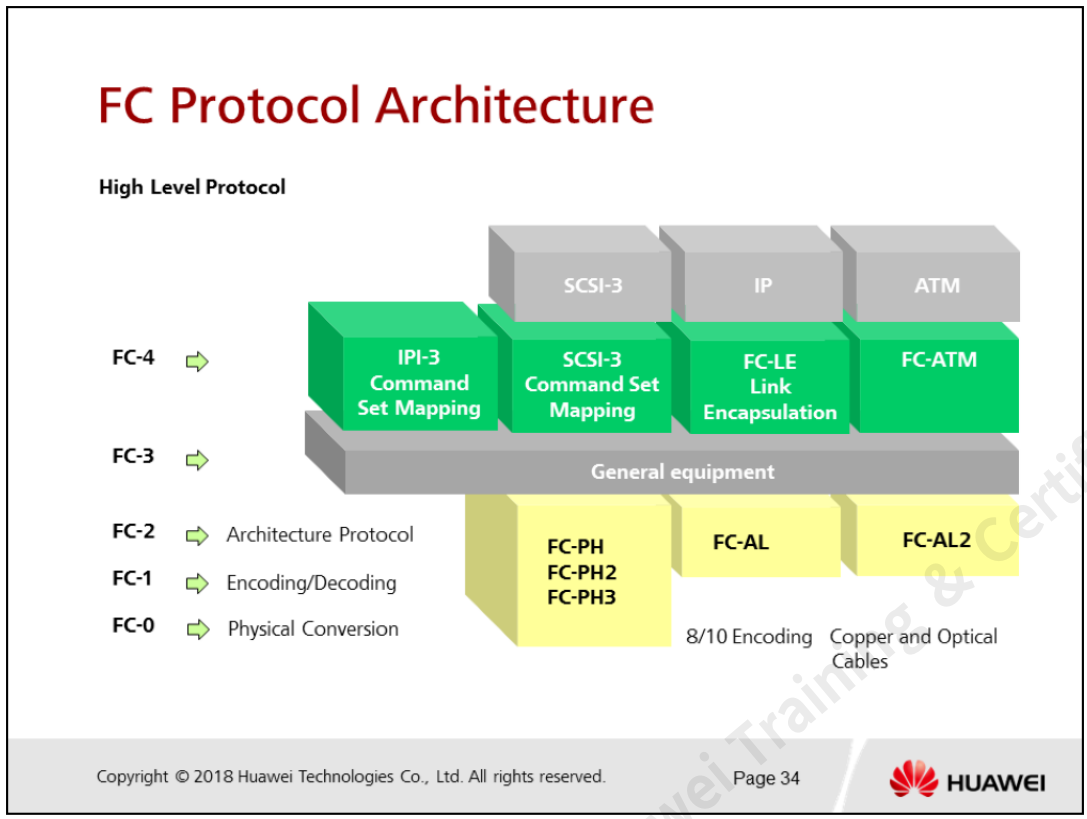
- FC is a standard for high-performance serial connection. Its interface transmission rate is 4Gbps, 8Gbps, 16Gbps or higher. FC transmission medium can be chosen from the option of copper cable or optical fiber. It has long transmission distance, and supports multiple types of network topologies.
- The definition of Fibre Channel: FC refers to the meaning of "channel in the form of fibre", and it is often misunderstood with Fiber Channel due to the similarities of the word Fibre and Fiber. Fiber Channel can be understood as "optical channel". Both refers to different types technological context. Compared to TCP/IP, FC protocol stack contains a lot of similar concepts. For example FC switches, FC routers, SPF(shortest-path-first) algorithm etc. We can consider FC protocol as the TCP/IP protocol within SAN because both of them follow the OSI model in their architecture design.
- FC protocol cannot be considered as Fiber protocol because the FC connection medium can be optical fiber, twisted pair cables or coaxial cable. Due to FC protocols commonly adopt optical fiber as the transmission medium, thus a lot of people tend to consider FC as fiber channel protocol instead of the actual Fibre Channel Protocol.
- FC Protocol Advantages: It has high bandwidth, high reliability, high stability, low latency, and resistance to electromagnetic interference (EMI)etc. It also can provide very stable and reliable fiber connections, its easy to build large-scale data transmission and communication networks, and currently supports 1x, 2x, 4x and 8x bandwidth connection rate. As with the continuous development of technology, the bandwidth for FC is still expanding to meet the technical performance requirements of higher bandwidth data transmission.

Huawei Training & Certification



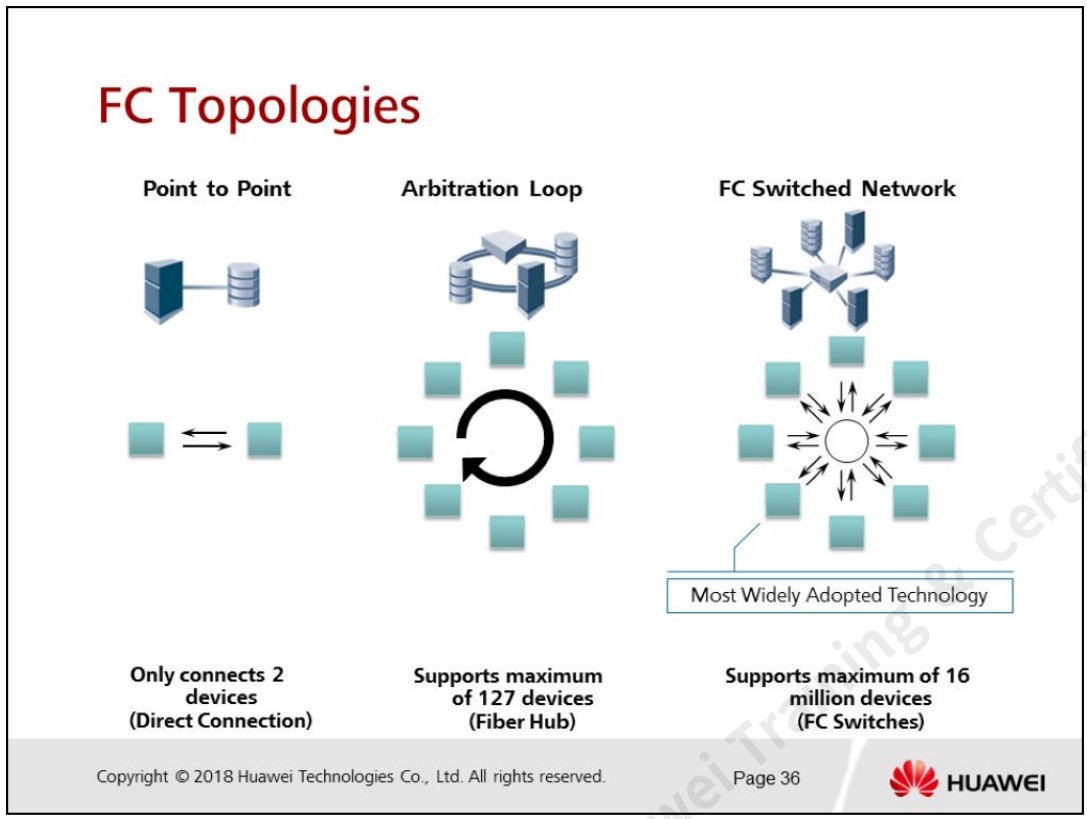
- FC originated in the year 1988, during that time, FC's main purpose was to simplify the connection and increase the data transmission distance instead of increasing the data transmission rate. As time progresses, it is used to increase the transmission bandwidth of hard disk protocols, with more inclination towards high speed, high efficiency and reliable data transmission. FC SAN started to gain wide application and adoption starting from the end of 1990s.
- Future Analysis:
  - From the perspective of trend analysis of research institutions, 16G FC has become an inevitable trend, and with the advent of PCIe3.0 and the increase in the density of virtual machines, higher I/O throughput is required, thus 32G FC will also gain more traction and has very promising future.
  - Standards update and commercialization pattern: The FC standard will be updated once every 36 to 48 months. After the new standard comes out, it will be commercially available within 12 to 24 months. Subsequently, the mainstream manufacturers will launch relevant products within 2 years. The standard life cycle of each generation is about 6 years, but from the point of view of 8G and 16G FC, the life cycle is obviously shortened, and the replacement cycle is obviously accelerated. Thus, FC technology has become an inevitable trend in the storage and SAN industry.

Huawei Training & Certification



- The main part of the Fibre Channel protocol architecture is actually FC-2. Among the layers, FC-0 to FC-2 is called FC-PH, which is also called "physical layer". Fibre Channel mainly transmit data through FC-2 layer. Therefore, Fibre Channel is also often referred to as "Layer 2 Protocol" which is similar to the Ethernet Protocol.
- Fibre Channel data units are called frames. Even though the fibre channel itself consists of five layers, but most of the time Fibre Channel is referred as a Layer 2 protocol. A Fibre Channel frame has a maximum of 2148 bytes, and the head of the Fibre Channel frame is different from the Ethernet packet. FC uses only one frame format and performs various tasks on multiple layers.
- The function of the frame determines its format. The Fibre Channel frame starts with the start of frame (SOF) flag, followed by the frame header. We will discuss more details on the frame header later. Then it is followed by the data, or Fibre Channel content, and finally the end of frame (EOF) flag. In simpler words, the frame structure is SOF + Data + EOF. The purpose of this encapsulation is to allow Fibre Channel frames to be carried by other protocols like TCP when its needed.

Huawei Training & Certification



- FC protocol used a long period of time since its initial introduction in 1988 to become a mature technology. Nowadays, SAN has a few methods to interconnect the physical components:
  - Point to Point: 2 devices are directly connected to each other. This is the most simple topology and has limited connection capabilities.
  - Arbitration Loop: In this type of connection, all the devices are connected into a loop similar to a Token Ring. Any addition or removal of the device within this loop will cause all the operation in the loop to stop. A faulty device within the loop will cause the whole loop to malfunction. By using a Fiber Hub, multiple devices can be connected together forming a logical loop, and bypass the faulty nodes meaning that a faulty device will not affect the communication of the whole loop. Arbitration loop was used in small SAN environments, but are obsolete nowadays. The main reason for it to be obsolete is because an Arbitration Loop can only contain a maximum of 127 devices, and the amount of devices used in SAN environment nowadays are way more than 127 devices.
  - Switched Network: This is the method that builds modern FC SAN networks. It uses FC switches to interconnect host and storage devices. In the modern SAN, it is the recommended best practice to use 2 switches to connect the host and storage devices, in order to form a redundant link which increases the reliability of SAN. Switches are an intelligent devices, it not only can perform the interconnection between devices but can do many more functions related to the network.

Huawei Training & Certification

Types of Ports on FC Switches

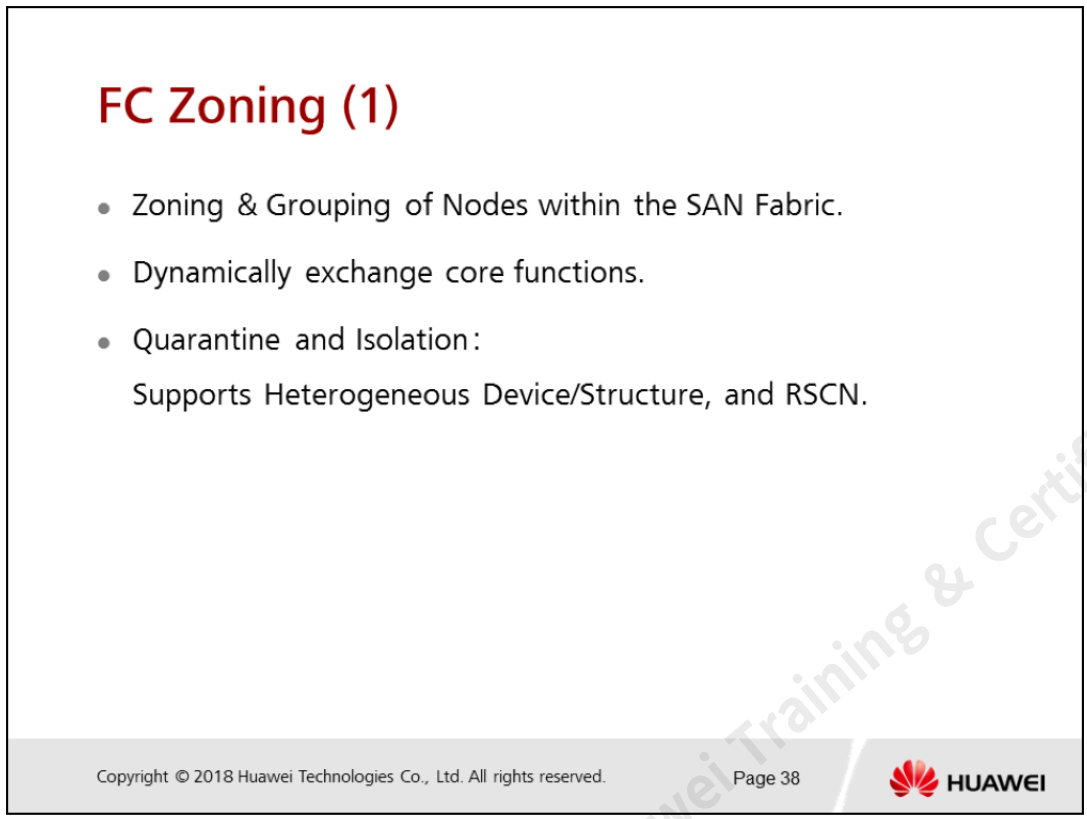
- D_Port (Diagnose Port): It is used to isolate the ISL (Inter-Switch Link) to diagnose the link level faults. It can only be used for diagnostics and testing, and does not carry any Fabric traffic or data flow.
- E_Port (Expansion Port): It is used for ISL with other switches to achieve Fabric expansion.
- EX_Port: It is a type of E_Port that is used to connect the FC Router with the outer fabric. The EX_Port is terminated by the Router and cannot be used to combine fabrics like done by the E_Ports of switches.
- F_Port (Fabric Port): The port which is connected to the Node Port (N_Port) of the FC devices. For example, when connecting storage devices, the switch port will be displayed as the F_Port.
- FL_Port(Fabric Loop Port): It is a type of L_Port that is able to perform the function of an F_Port, attached via a link to one or more NL_Ports in an Arbitrated Loop topology.
- G_Port(Generic Fabric Port): A generic switch port that may function either as an E_Port, A_Port, or as an F_Port.
- M_Port(Mirror Port): Used to replicate all the traffic flow between the source port to the destination port.
- U_Port / GL_Port (Universal/Generic Fabric Loop Port): The most basic FC port type, all unidentified or uninitialized ports belongs to U_Port or GL_Port.

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

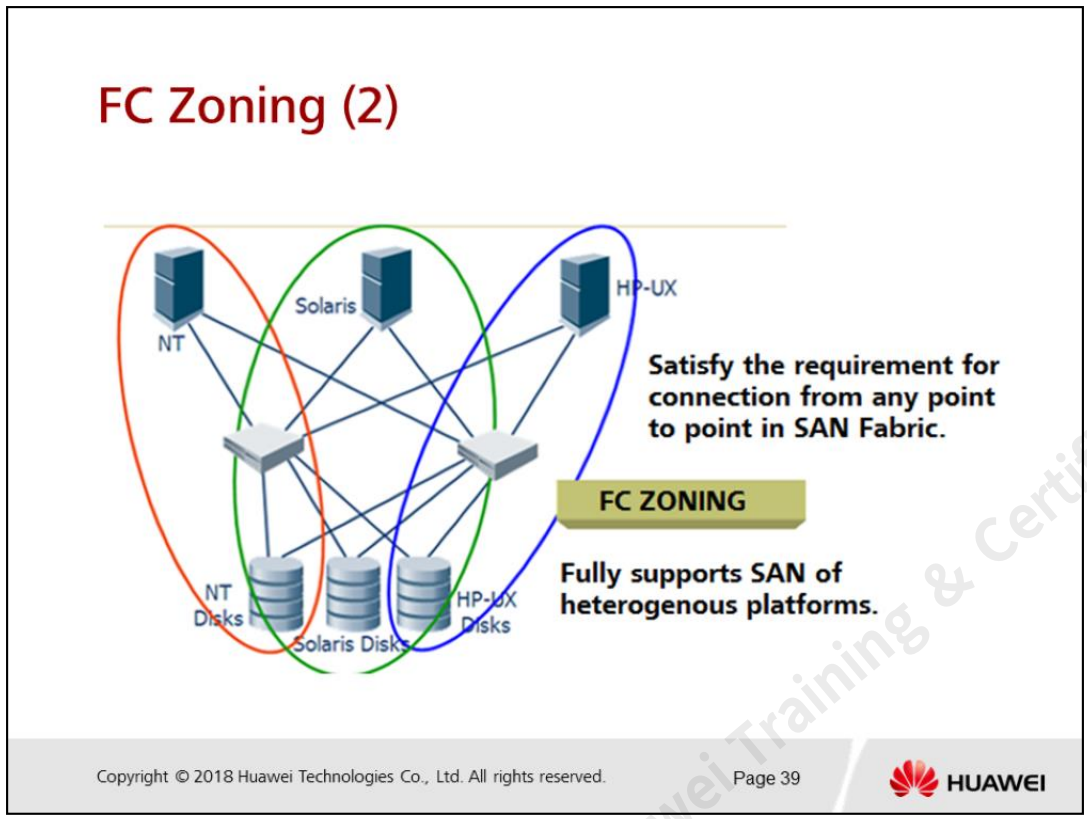
- The list above shows the different types of Ports in the FC fabric and their respective functions. The usage of correct port type can help prevent issues in accessing the fabric from the nodes.

Huawei Training & Certification



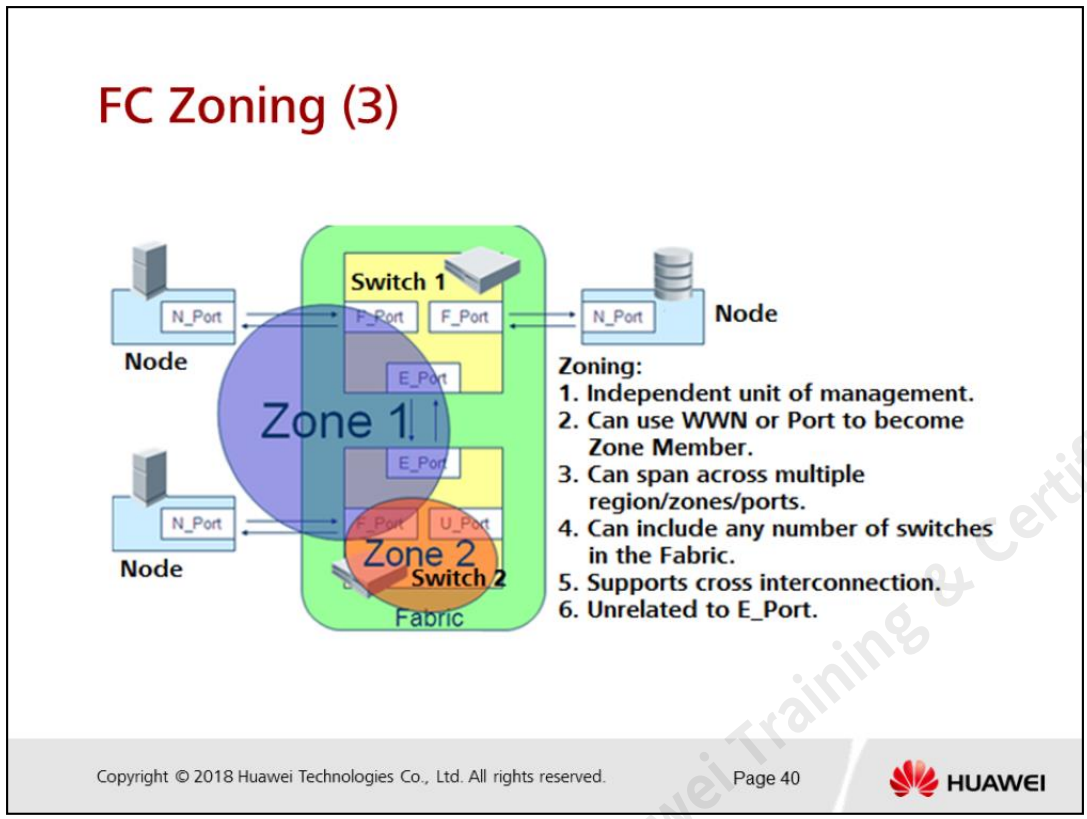
- FC switches Zoning function is similar to the VLAN functions of Ethernet networks, which is used to achieve interconnection of specific devices in the network and avoiding flooding of broadcast packets.
- Zoning function allows access control of the members of the same VSAN(Virtual SAN), providing further division of regions within the VSAN. Different N_Port members can be added into a Zone based on different purposes, and this also allows the isolation between the N_Port members of different zones.
- Registered State Change Notification (RSCN) : A feature in the FC switches that is responsible to notify the nodes of the changes within the network fabric or changes within the architecture itself.
- In Fibre Channel protocol, a registered state change notification (RSCN) is a Fibre Channel fabric's notification sent to all specified nodes in case of any major fabric changes. This allows nodes to immediately gain knowledge about the fabric and react accordingly.

Huawei Training & Certification



- The diagram above shows the application of FC zoning in a heterogeneous SAN platform.
- FC zoning works across storage platforms in the fabric and can logically group the host nodes and storage nodes within the SAN fabric.

Huawei Training & Certification

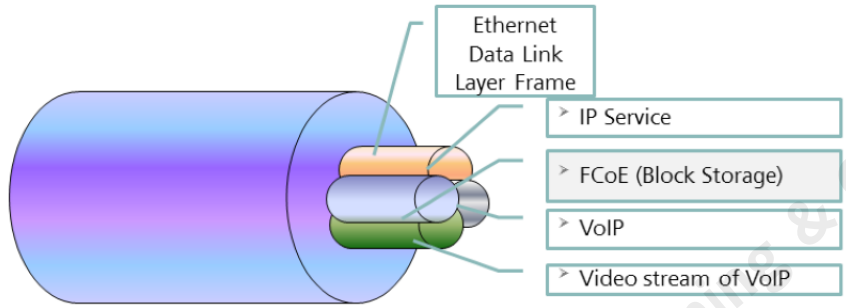


- The diagram above shows a scenario where FC zoning is applied by spanning across multiple switches and interconnection of two FC zones.
- The nodes are logically isolated in the fabric by the zones in the SAN fabric.

Huawei Training & Certification

FCoE Protocol

- It is a protocol that directly transmits FC signals over the enhanced lossless Ethernet network.
- FCoE encapsulates the FC frame within an Ethernet Frame, allowing LAN and SAN service traffic to be transmitted through the same Ethernet network

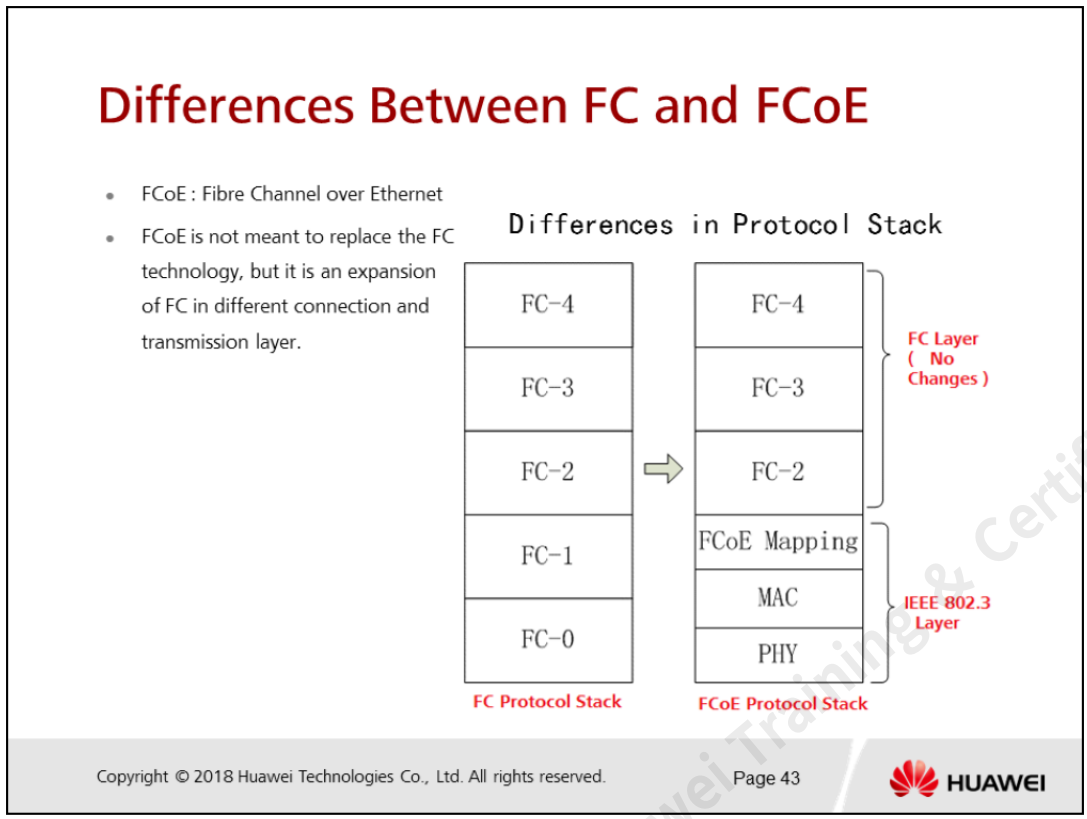


Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

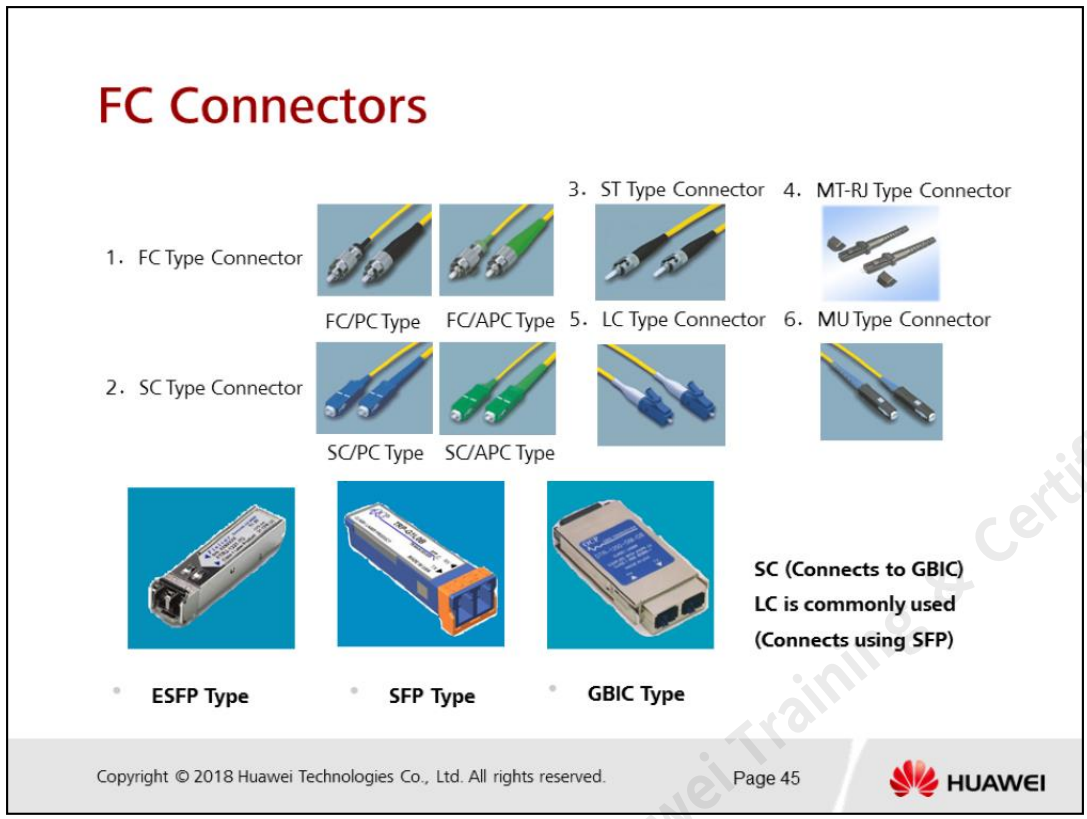
- FCoE (Fibre Channel Over Ethernet) can provide standard FC services in Ethernet networks such as device discovery, global namespace, zoning and these services can operate according to the actual original standards, retaining the low latency, high performance capabilities of FC.
- From the perspective of FC protocols, FCoE is just carrying FC on a new type of link which is the Ethernet link. One thing to be noted is that, this Ethernet network must be Enhanced Lossless Ethernet Network, in order to satisfy the transmission requirements of FC protocols towards the link layer.
- Features of FCoE:
  - Standards Organization: The protocol was submitted to ANSI (American National Standards Institute)T11 committee for approval in 2008, and it needs to work closely with IEEE (Institute of Electrical and Electronics Engineers). Thus, this protocol is verified and accepted globally.

Huawei Training & Certification - Huawei Training & Certification



- FCoE retains the protocol stack above the FC-2 layer, and replaces the FC-0 and FC-1 layers with the Physical and Data Link Layer in the Ethernet protocol.
- Due to the fact that FC-0 main purpose is to define the types of transmission medium, and FC-1 was responsible for the encoding and decoding of the frames, both of these layers was required in the FC SAN network during data transmission. However, since FCoE runs on the Ethernet network, these 2 layers are no longer required and can be replaced with the Physical and Data Link Layer of the Ethernet protocol.
- The value of FCoE lies in the fact that the users have the choice of either having the whole logical network to function as the dedicated network for storage data transmission and signaling, or to have a shared hybrid network that has mixed traffic of storage data transmission, VoIP, video streaming and other data transmission. The objective of FCoE is to integrate storage data transmission with the Ethernet network in the preconditions that it continues to support users requirement towards high performance and features extensiveness of FC SAN.
- Different Operating Environment: FC protocol runs on tradition FC SAN network. Meanwhile, FCoE runs as the storage protocol in the Ethernet network.

Huawei Training & Certification



- FC type optical fiber connector: FC is the abbreviation of the Ferrule Connector, which indicates that the external reinforcement method uses a metal sleeve, and the fastening method is a turnbuckle. Such connectors has a simple structure, it is easy to operate, and it is easy to manufacture, but the fiber end is more sensitive to micro dust, and Fresnel reflection is easy to occur in these connectors, and it is difficult to improve the return loss performance of these connectors.
- SC type optical fiber connector: SC is the abbreviation of Square Connector, its named such because the connector housing is square in shape. The pin size and the coupling sleeve used in these connectors are the same as the FC type, in which the fiber end of the pin is mostly PC or APC type. The fastening method is a plug-and-pin latch type, which does not require rotation. These connectors are inexpensive, easy to insert and withdraw, have small fluctuations in insertion loss, high compressive strength, and high installation density.

Huawei Training & Certification

Contents

1. SCSI/iSCSI
2. SAS
3. FC/FCOE
4. **PCIe**
5. IB
6. CIFS/NFS
7. FTP/HTTP

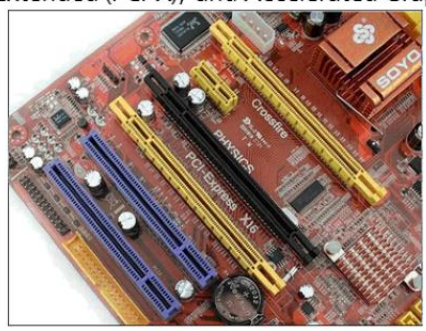
Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

Huawei Training & Certification
Huawei Training & Certification

What Is PCIe ?

- PCI Express (also known as PCIe) is a high-performance, high-bandwidth serial communications interconnection standard first introduced by Intel and later developed by the Peripheral Component Interconnect Special Interest Group (PCI-SIG) to replace bus-based communications Architectures such as: PCI, PCI Extended (PCI-X), and Accelerated Graphics Port (AGP).

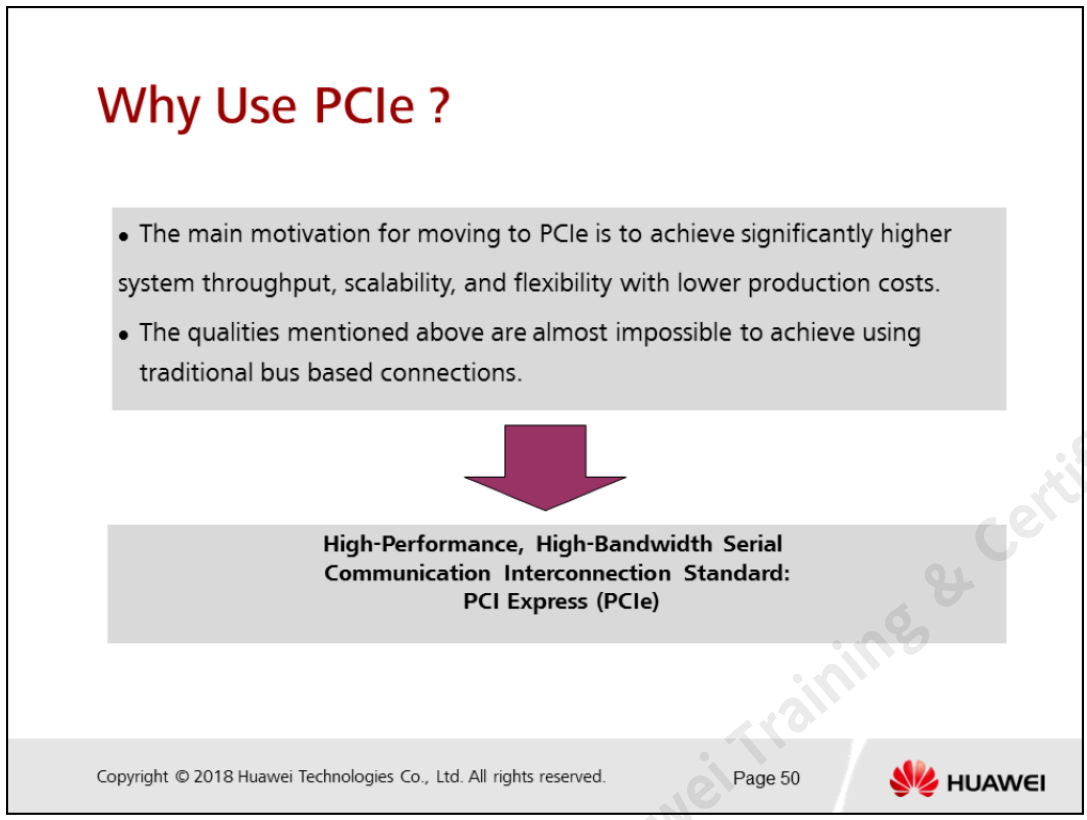




Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

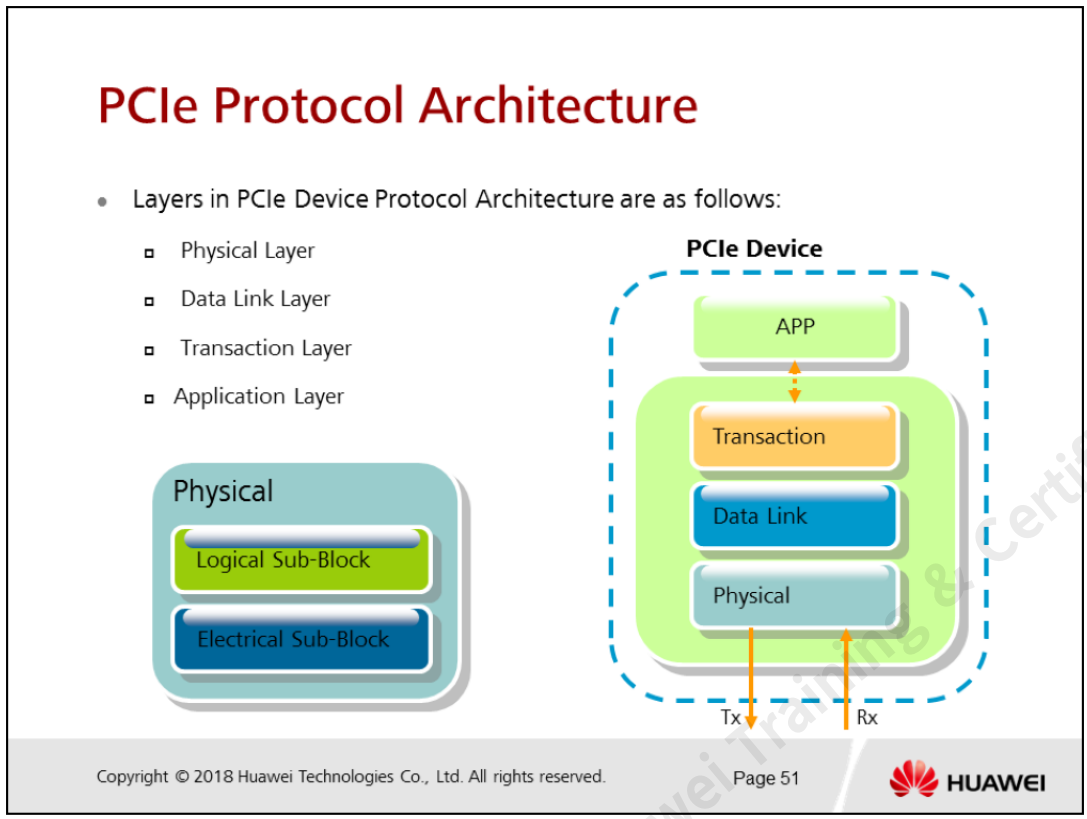
- In the year 1991, Intel Corporation introduced the concept of PCI. With the development of modern processor technology, the replacement of the parallel bus with the high-speed differential bus is a general trend in the field of interconnection. Additionally, when compared with single-ended parallel signals, the high-speed differential signals can be used for higher clock frequencies, and thus induced the emergence of PCIe bus.
- PCI Express (aka PCIe) is a high-performance, high-bandwidth serial communications interconnection standard first introduced by Intel and later developed by the Peripheral Component Interconnect Special Interest Group (PCI-SIG) to replace bus-based communications. Architectures such as: PCI, PCI Extended (PCI-X), and Accelerated Graphics Port (AGP).
- PCI Express has the following advantages over the traditional PCI bus:
  - Dual-channel, high-bandwidth, fast transfer rate: Achieves a similar full-duplex transmission mode (RX and TX are separated). It has high transmission rate. The bandwidth of the first generation PCIe X1 is 2.5 Gigabits per second (Gbps), the second generation has up to 5.0 Gbps, and the recently released PCIe 3.0 standard is capable of supporting 8.0 Gbps. For larger bandwidths, it can be achieved by expanding the number of links, and the resulting bandwidth is the number of paths(N) multiplied by the bandwidth of each paths.

Huawei Training & Certification



- The formulation of the PCI Express standard is based on the concerns for the future and it continues to evolve to provide higher throughput for the system. The throughput of the first-generation PCIe protocol was 2.5 gigabits per second (Gbps), while the second-generation PCIe protocol reached 5.0 Gbps, and the recently released PCIe 3.0 standard can support 8.0 Gbps. While the PCIe standard continues to utilize the latest technology to provide increasing throughput, the transition from PCI to PCIe will be simplified using a layered protocol by keeping the driver software compatible with existing PCI applications.
- The PCIe protocol features include:
  - Point-to-point connection
  - High reliability
  - Tree network
  - Full duplex
  - Frame-based transmission

Huawei Training & Certification



- Physical Layer:
  - The physical layer in the PCI Express bus architecture mainly defines the physical characteristics of the bus. In future development, the performance of the PCI Express bus can be further improved by speeding up or changing the codec mode. These changes will only affect the physical layer and will not affect other structures, thus facilitating the upgrade process.
- Data Link Layer:
  - The important role of the data link layer is to ensure the correctness and reliability of data packets transmitted over the PCI Express bus. It will check if the data packet encapsulation is complete and correct, add the serial number and cyclic redundancy check code (CRC) to the data for detection and error correction, and use ack/nack handshake protocol for detection and correction.
- Transaction Layer:
  - There are 2 main role of the transaction layer: The first role is to accept read and write requests sent from the software layer, or to create a request encapsulated packet itself to the data link layer, this data packet is called "Transaction Layer Packet". The second role of this later is to accept the response packet (Data Layer Packet, DLP) sent from the data link layer and associate it with the relevant software request and send it to the software layer for processing.
- Application Layer:
  - This layer mainly consists of applications that sends requests and receive responses from PCIe devices.

Huawei Training & Certification

Bandwidth Of PCIe Links

Specification	Bandwidth/IO Count	Work Frequency	Transmission Rate	Encoding
PCI 2.3	32 bit	33/66 MHz	133/266 MB/s	--
PCI-X 1.0	64 bit	66/100/133 MHz	533/800/1066 MB/s	--
PCI-X 2.0 (DDR)	64 bit	133 MHz	2.1 GB/s	--
PCI-X 2.0 (QDR)	64 bit	133 MHz	4.2 GB/s	--
AGP 2X	32 bit	66 MHz	532 MB/s	--
AGP 4X	32 bit	66 MHz	1.0 GB/s	--
AGP 8X	32 bit	66 MHz	2.1 GB/s	--
PCI-E x1	4 (Differential)	2.5 GHz	500 MB/s (Full Duplex)	8b/10b
PCI-E x2	8 (Differential)	2.5 GHz	1.0 GB/s (Full Duplex)	8b/10b
PCI-E x4	16 (Differential)	2.5 GHz	2.0 GB/s (Full Duplex)	8b/10b
PCI-E x8	32 (Differential)	2.5 GHz	4.0 GB/s (Full Duplex)	8b/10b
PCI-E x16	64 (Differential)	2.5 GHz	8.0 GB/s (Full Duplex)	8b/10b
PCI-E Gen2 x1	4 (Differential)	5.0 GHz	1.0 GB/s (Full Duplex)	8b/10b
PCI-E Gen2 x2	8 (Differential)	5.0 GHz	2.0 GB/s (Full Duplex)	8b/10b
PCI-E Gen2 x4	16 (Differential)	5.0 GHz	4.0 GB/s (Full Duplex)	8b/10b
PCI-E Gen2 x8	32 (Differential)	5.0 GHz	8.0 GB/s (Full Duplex)	8b/10b
PCI-E Gen2 x16	64 (Differential)	5.0 GHz	16.0 GB/s (Full Duplex)	8b/10b
PCI-E Gen3 x1	4 (Differential)	8.0 GHz	~2.0 GB/s (Full Duplex)	128b/130b
PCI-E Gen3 x2	8 (Differential)	8.0 GHz	~4.0 GB/s (Full Duplex)	128b/130b
PCI-E Gen3 x4	16 (Differential)	8.0 GHz	~8.0 GB/s (Full Duplex)	128b/130b
PCI-E Gen3 x8	32 (Differential)	8.0 GHz	~16.0 GB/s (Full Duplex)	128b/130b
PCI-E Gen3 x16	64 (Differential)	8.0 GHz	~32.0 GB/s (Full Duplex)	128b/130b

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

- The table above shows the different bandwidth, working frequency, encoding method and the transmission rate for the specifications of PCI an d PCIe technology of different generations.

Huawei Training & Certification

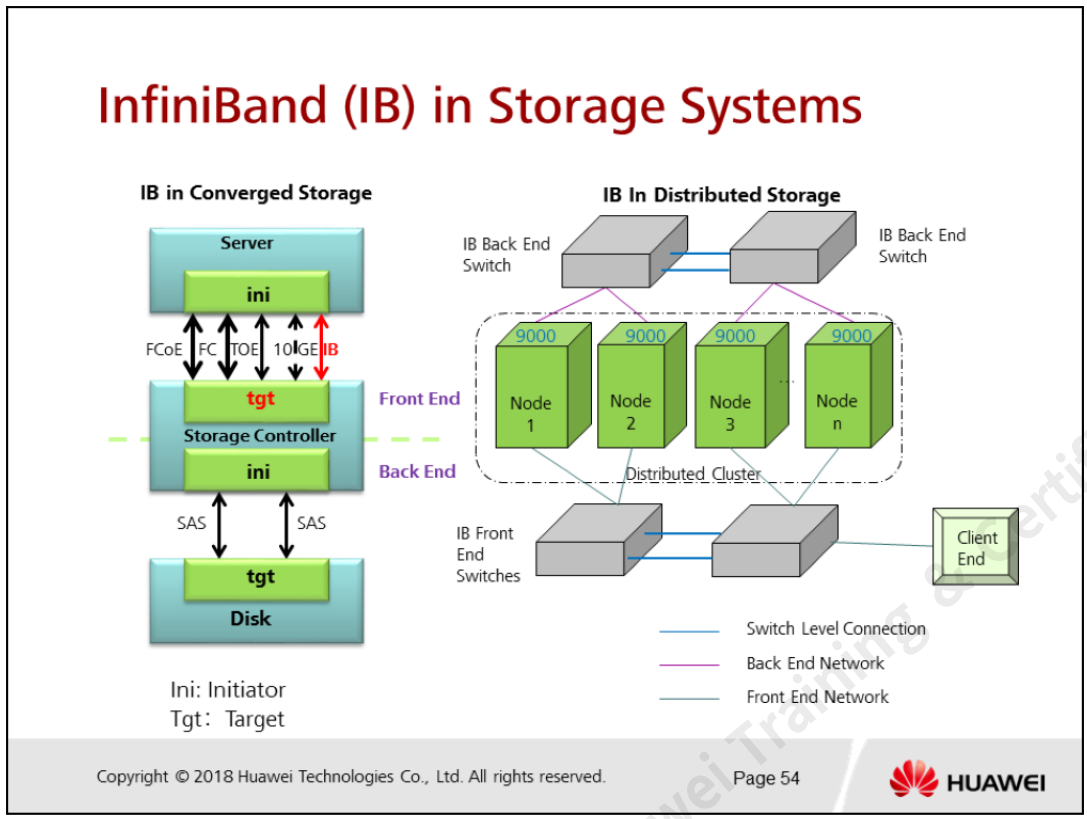
Contents

1. SCSI/iSCSI
2. SAS
3. FC/FCOE
4. PCIe
5. **IB**
6. CIFS/NFS
7. FTP/HTTP

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

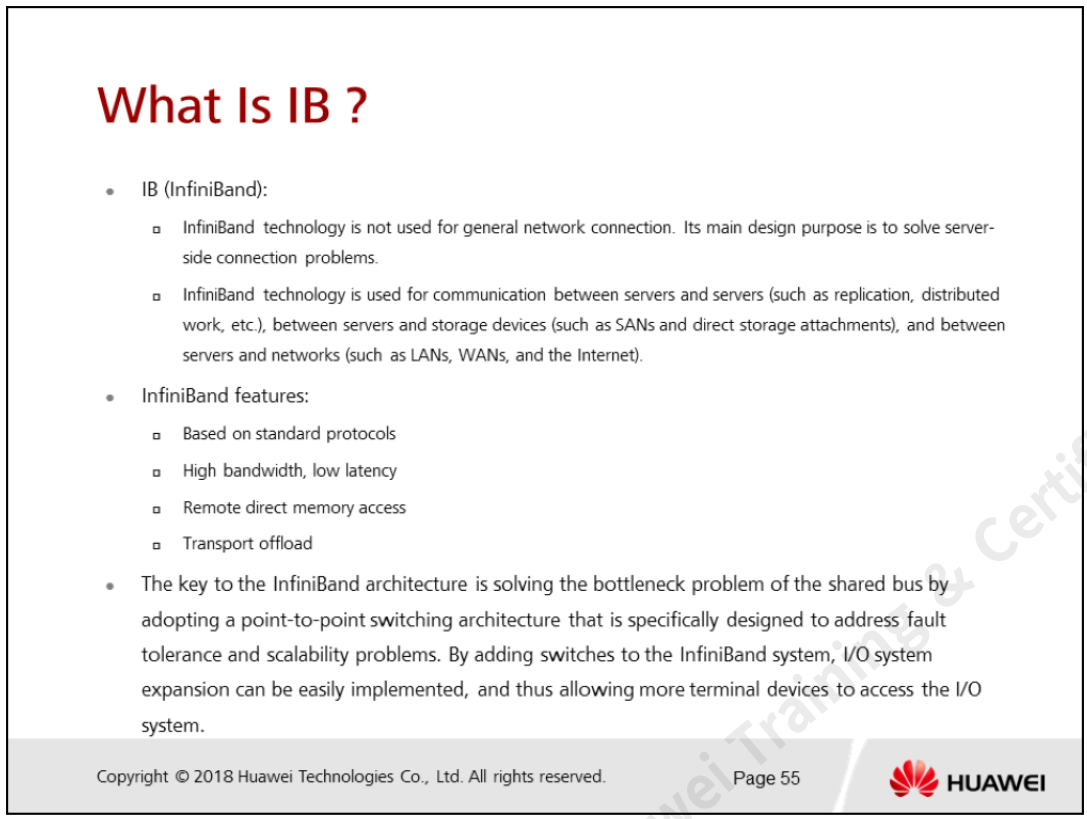
HUAWEI

Huawei Training & Certification
Huawei Training & Certification



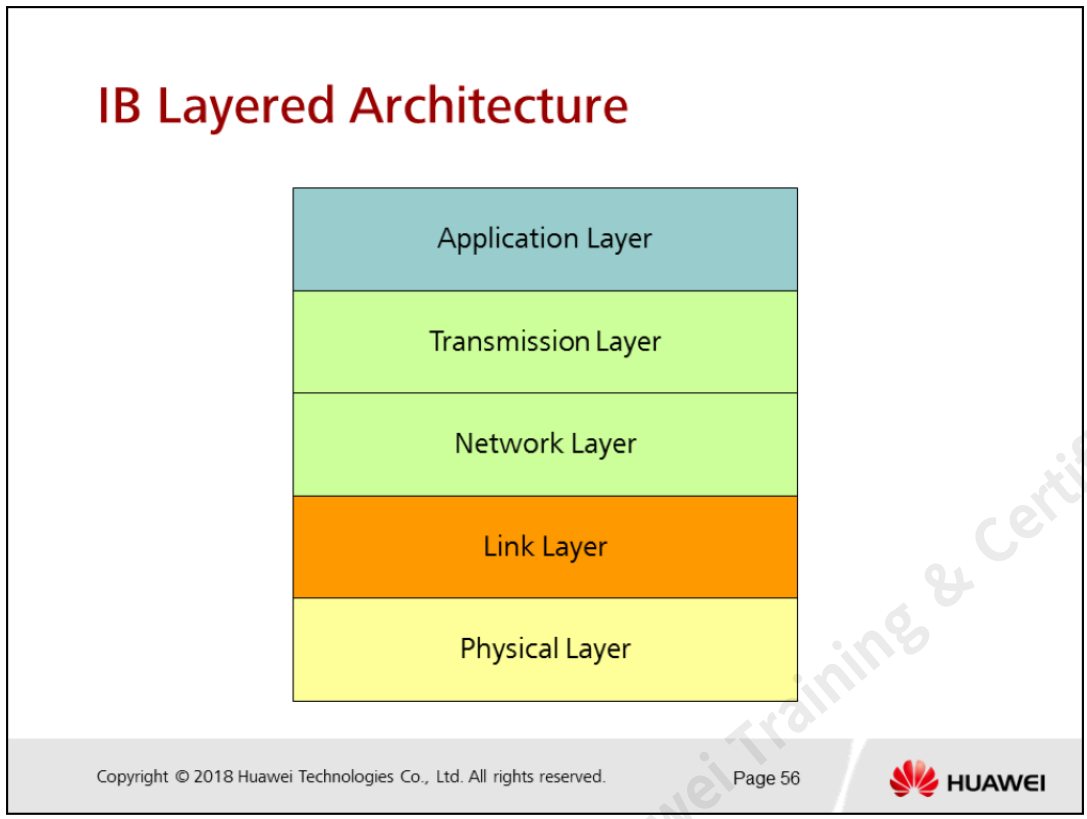
- IPoIB (IP over IB), which is the adaptation layer of Linux kernel and IB driver, is responsible for constructing, destroying IP header, sending and receiving IP message.
- The IB front-end network is used to communicate with customers for data exchange networks, and transmits data based on the IPoIB protocol.
- The IB back-end network is used to store the data exchange network between the nodes inside the device. The RPC(Remote Procedure Call) module uses RDMA (Remote Direct Memory Access) to complete the data synchronization between nodes.

Huawei Training & Certification



- The InfiniBand standard defines a set communication system used by multiple devices, which includes channel adapters, switches, and router channel adapters for connection to other devices, including host channel adapters (HCA) and target channel adapters (TCA).
- The main features of the InfiniBand protocol:
  - Protocol Based On Standards: The InfiniBand Trade Association, which was established in 1999, consists of 225 companies that have jointly designed this open standard. The members that mainly control the association include: Agilent, Dell, HP, IBM, InfiniSwitch, Intel, Mellanox, Network Appliance and Sun Microsystems. More than 100 other members assisted in the development and promotion of the IB standard.
  - Speed: InfiniBand's performance of 10 gigabytes per second significantly exceeds the speed of 4 gigabits per second for existing Fibre Channel technology and also exceeds 1 gigabit per second speed for Ethernet technology.
  - Memory: InfiniBand-enabled servers use host channel adapters to convert protocols to the server's internal PCI-X or PCI-Xpress bus. HCA has RDMA functionality, sometimes also referred to as Kernel Bypass. RDMA is a good fit for a cluster because it can use a virtual addressing scheme to let the server know and use some of the other server's memory without having to refer to the operating system's kernel.
  - Transport Offload: RDMA can help offload transfers, which shifts packet routing from the OS to the chip level, saving the processor's processing workload. An 80 GHz processor is required to process data at 10 Gbps in the OS.

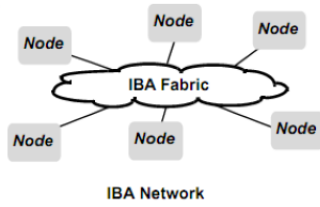
Huawei Training & Certification



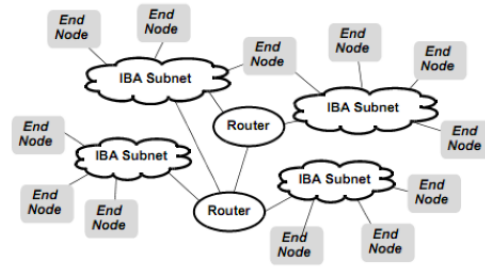
- Physical layer: Defines connections of three different rate of transmission which are, 1X, 4X, and 12X respectively, and their signal transmission rates are 2.5, 10, and 30 Gb/s respectively. This means that the IBA (InfiniBand Architecture) allows multiple connections until a 30Gbps connection speed is achieved. Due to the use of full-duplex serial communication, a single-speed two-way connection requires only four cables, and in the case of a 12-speed mode, it requires only 48 cables, which is very attractive in terms of cable usage.
- Link layer: The link layer provides functions such as packet design, point-to-point connection operation, and packet switching in the local subsystem. At the packet communication level, two special packet types are specified, which are data transmission and network management packets. The network management package provides functions such as device enumeration operation control, subnet indication, fault tolerance, etc. On the other hand, the data transmission package is used to transmit actual data information. The maximum length of each packet is 4 KB. Within each specific device subnet, the direction and switching of each packet is completed by the subnet manager through the usage of the local 16-bit identification address.
- Network layer: This layer provides a routing mechanism for information packets from one sub-structure to another sub-structure. Each of the source and destination nodes has a Global Routing Header (GRH) and a 128-bit IPv6 address. The network layer also embeds a standard global 64-bit identifier that is unique across all subnets. Through intricate exchanges between these identification values, data is allowed to travel across multiple subnets.

Huawei Training & Certification

IB Architecture (IBA)



IBA Network



IBA Network Components

- Components of InfiniBand Architecture(IBA):
- Node: Host Channel Adapter (HCA), Host Target Adapter(HTA).
- Network: Switch, Routers.
- Physical: Links(Fiber or Cables), Repeaters.

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

- IBA originally targeted to become the standard interface for internal and external connections for computing, communication and storage. However, IBA lost the opportunity for internal system operation in communication devices due to the implementation of Rapid I/O by multiple communication equipment manufacturers. So, IBA can only set its target towards computing and storage devices for its adoption. But coincidently the rapid growth of SAN, made FC technology to be widely adopted which marks that IBA lost its ground for adoption in storages. Afterwards, IBA sets its sights at computing devices, aiming to replace PCI standard, but PCI-X took the place in the replacement of PCI. After PCIe was proposed by Intel to replace PCI-X, it made it impossible for IBA become the standard for internal connections in computing devices. Even the main competitors of Intel such as Sun and AMD used HyperTransport/HTX and not IBA which made the chances for IBA adoption in internal system connections for computing devices much smaller.

Huawei Training & Certification

IB Interface

- **Channel Adapter (CA)** is divided into 2 different types as below:
  - **Host Channel Adapter (HCA):** It is used at the hosts side such as Mellanox HCA products.
  - **Target Channel Adapter (TCA):** It is used mainly on the IB switches side or serve as the interface on the storage systems.

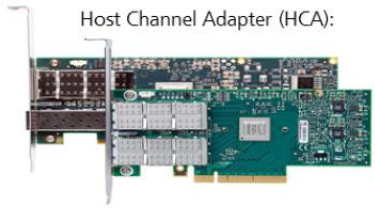
IB Switches:



IB Connectors:



Host Channel Adapter (HCA):



Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

- Mellanox Technologies was founded in 1999 and has headquarters in Santa Clara, California, and Yokneam, Israel. Mellanox is a leading provider of server and storage end-to-end connectivity solutions. In the end of 2010, Mellanox completed the acquisition of the famous Infiniband switch manufacturer Voltaire, which enabled Mellanox to gain more comprehensive capabilities in the HPC, cloud computing, data center, enterprise computing and storage markets.
- Channel adapters are divided into:
  - Host Channel Adapter (HCA), which is used to control the NODE.
  - Target Channel Adapter(TCA), which is used for peripheral NODE, so that the IO of the device is offloaded from the host and transferred directly in the network.
- Channel adapters implement the physical layer, link layer, network layer, and transport layer functions.
- The channel adapter is an important part of the IB network interface. It is a programmable DMA device with specific protection features allowing local and remote DMA operations.

Huawei Training & Certification

IB Signaling Mode

Transmission Rate (Gbit/s)	SDR	DDR	QDR	FDR	EDR
1X	2.5	5	10	14	26
4X	10	20	40	56	104
8X	20	40	80	112	208
12X	30	60	120	168	312

- SDR: single data rate
- DDR: double data rate
- QDR: quad data rate
- FDR: fourteen data rate
  EDR: enhanced data rate
- HDR : High Data Rate
  NDR : Next Data Rate

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

- 8b/10b encoding for SDR, DDR, FDR-10, and QDR: This encoding method has 8 bits data per 10 bits transmitted.
- 64b/66b encoding for FDR and EDR: This encoding method has 64 bits of data per 66 bits transmitted.
- The remaining 2 bits in these encoding methods are at the start and end of frame.
- The effective one-way theoretical throughput (actual data rate, and non-signaling rate) will be lower than this value.

Huawei Training & Certification

What Makes IB Performance So Fast?

The channel-based end-to-end switched interconnection structure does not share the bus and there are no related electronic restrictions, arbitration conflicts, and memory consistency issues.

Simple and efficient protocol, low overhead, protocol supports hardware offload.

QoS: 16 level VL(Virtual Lanes) and 16 level SL(Service Level), which helps in achieving efficient service quality management and credit-based dual-layer flow control mechanism.

RDMA + Zero Copying technology for data.

Supports multiple concurrent links: Theoretically more number of concurrent links equals to more speed, such as QDR: X1 10Gbps, X4 40Gbps, X8 80Gbps, X12 120Gbps.

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

Huawei Training & Certification
Huawei Training & Certification

Contents

1. SCSI/iSCSI
2. SAS
3. FC/FCOE
4. PCIe
5. IB
6. **CIFS/NFS**
7. FTP/HTTP

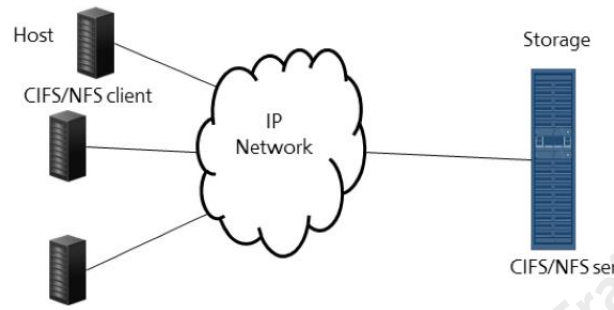
Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

Huawei Training & Certification
Huawei Training & Certification

CIFS/NFS In Storage Systems

- The two most common network sharing protocols for NAS are: CIFS and NFS.
  - **CIFS (Common Internet File System):** CIFS refers to the collective name of SMB (Server Message Block). Network file sharing between Windows hosts is achieved by using Microsoft's own CIFS service.
  - **NFS (Network File System)** is a network file system that uses NFS extensively for cloud computing and databases. UNIX-like operating systems such as Linux/UNIX/AIX/HP-UX/Mac OS X uses NFS to provide network file system storage services.

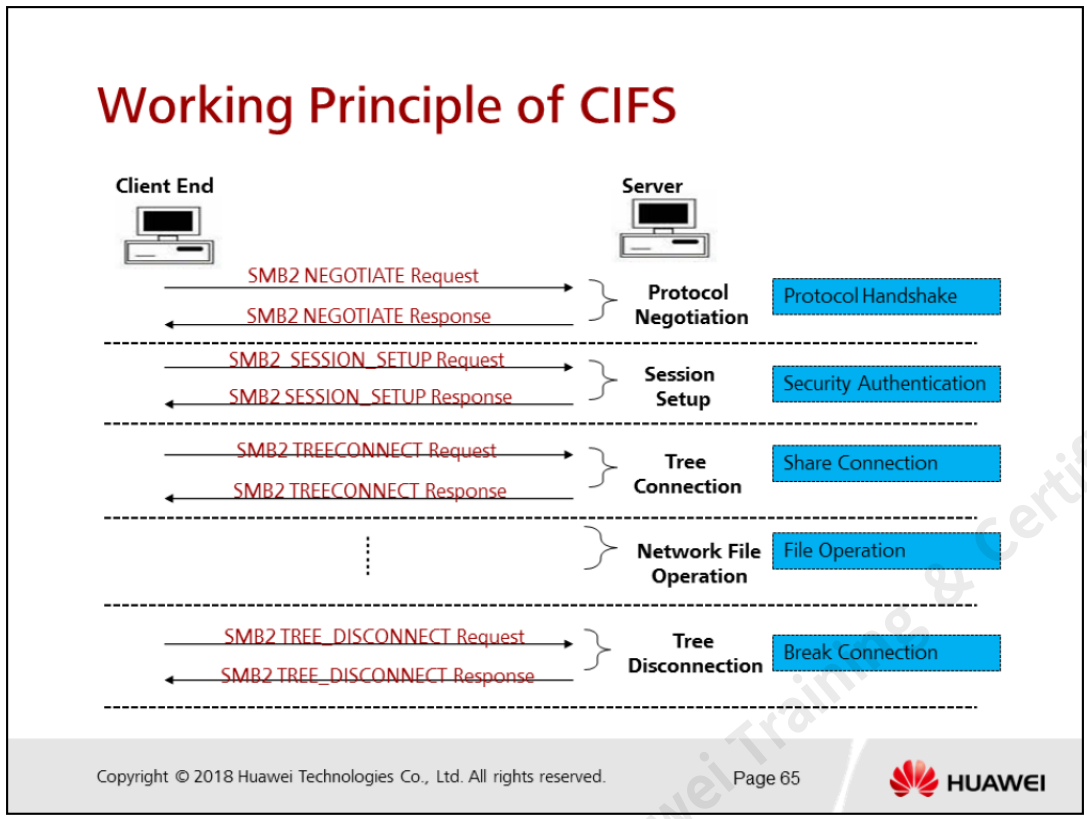


Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

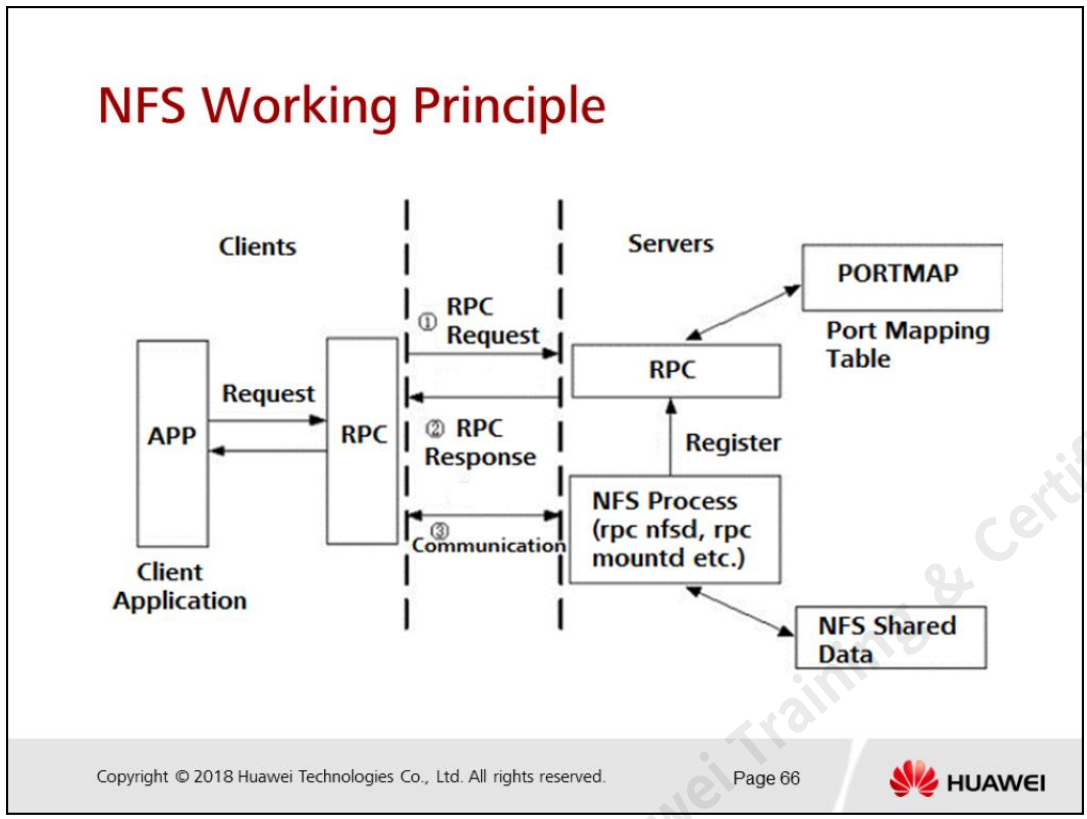
- CIFS:
  - In 1996, Microsoft proposed to rename the SMB as CIFS (Common Internet File System) and added many new features. Now CIFS refers to the general name of SMB, specifically each version is SMB1, SMB2, SMB3.0. SMB is a client/server, request/response mode protocol.
  - The Server Message Block (SMB) was originally developed by IBM's Barry Feigenbaum. Its purpose is to transform the local file interface "Interrupt 13" in the DOS operating system into a network file system. The SMB protocol is mainly used to share files, printers, serial ports, etc. between computers.
  - The SMB has been continuously improved since 1988 and has evolved from SMB to SMB2 (2007) and SMB3 (2012).
  - CIFS (Common Internet File System) is a file sharing protocol developed by Microsoft for connecting Windows clients and servers.
  - CIFS is a standard formed after the publication of the Server Message Block (SMB) developed and used by Microsoft. SMB is mainly used by computers on the network to share files, printers, and serial ports.
  - After redevelopment by UNIX server vendors, SMB can be used to connect UNIX servers and Windows clients and perform tasks such as printing and file sharing.

Huawei Training & Certification



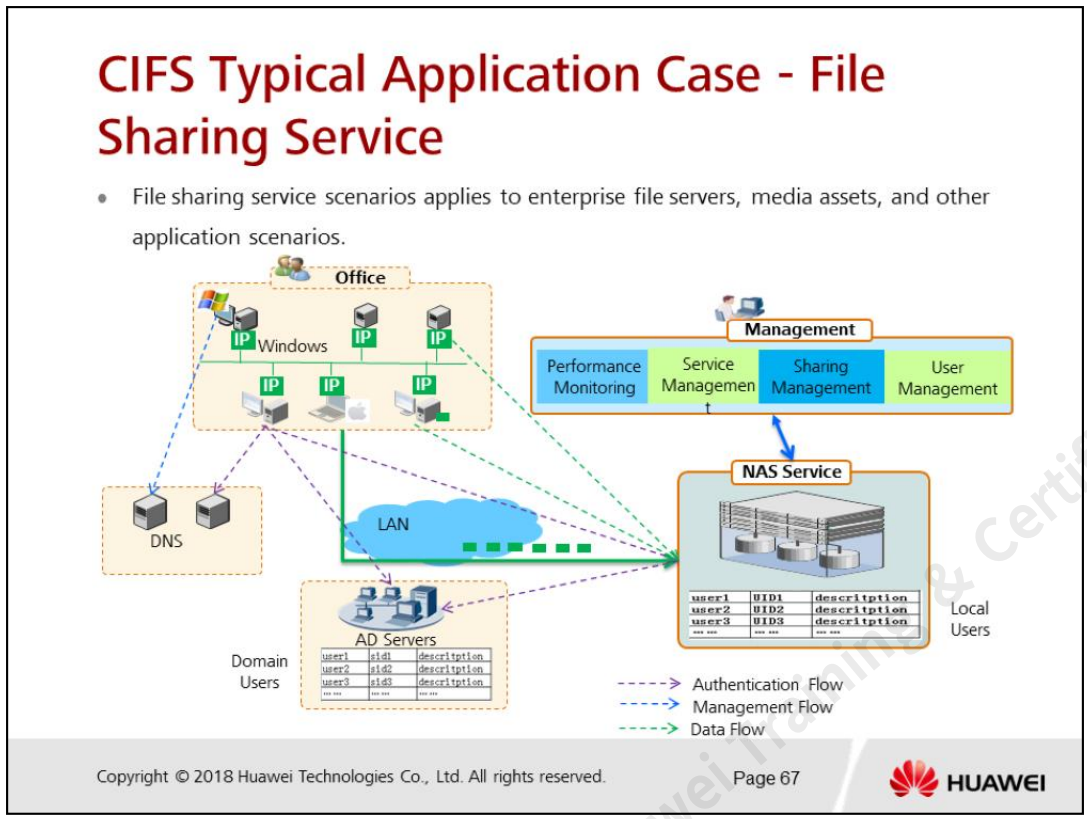
- Authentication method: NTLM, Kerberos etc.
- Network traffic can be carried over TCP or other ways such as: RDMA etc.
- Network file operations are processed and transferred to the actual file system.
- NTLM security supports provider services. NTLM means NT LanManger, one of the authentication methods provided under NT which uses 64-bit encryption.
- Kerberos network authentication. Kerberos is a network authentication protocol designed to provide powerful authentication services for client/server applications through a key system.

Huawei Training & Certification



- RPC (Remote Procedure Call): A remote procedure call, which is a protocol for requesting services from a remote computer program over a network without having to understand the underlying network technology. The RPC protocol assumes the existence of certain transport protocols, such as TCP or UDP, to carry information data between communications programs. In the OSI network communication model, RPC crosses the transport layer and the application layer. RPC makes it easier to develop applications that include network distributed programming.
- RPC uses client/server mode. The requestor is a client, and the service provider is a server. First, the client that invokes a RPC process sends an invocation of the process parameter to the service process and then waits for a reply. On the server side, the process stays asleep until the RPC message arrives. When a RPC message arrives, the server obtains the process parameters, calculates the result, sends the reply message, and then waits for the next RPC message. Finally, the client-side RPC process receives the reply message, obtains the process result, and then invokes execution to proceed.

Huawei Training & Certification

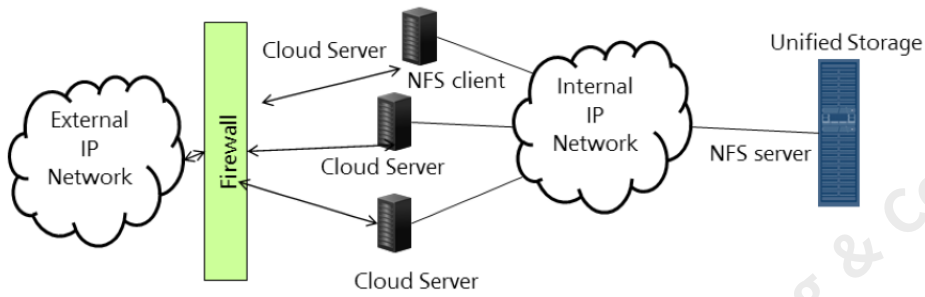


- DNS: Refers to domain name service, which is used to implement the mapping of domain names and IP.
- AD: Refers to Active Directory. An AD domain is used for providing directory services.
- The CIFS protocol is mainly used for file sharing. The following are two of its typical application scenarios:
  - File sharing service scenario
    - File sharing service is the most typical application scenario. It is mainly applied to enterprise file servers, media assets, etc., and provides users with file sharing services.
  - Hyper-v virtual machine application scenario
    - Microsoft promotes Hyper-V virtual machines and uses SMB to share virtual machine images. In this scenario, the failover feature of SMB 3.0 needs to be relied on. This feature ensures that the node switchover occurs in the event of a node failure and the service is not interrupted, thus ensuring the reliability of the virtual machine operation.

Huawei Training & Certification

NFS Typical Application Case:Shared Storage In Cloud Computing

- Cloud Computing uses NFS servers as its internal shared storage:



Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

- Cloud virtualization software (such as VMware) is optimized specifically for NFS clients to create virtual machine storage space on the NFS server's shared space.
- This cloud-optimized NFS client provides better performance and reliability.

Huawei Training & Certification

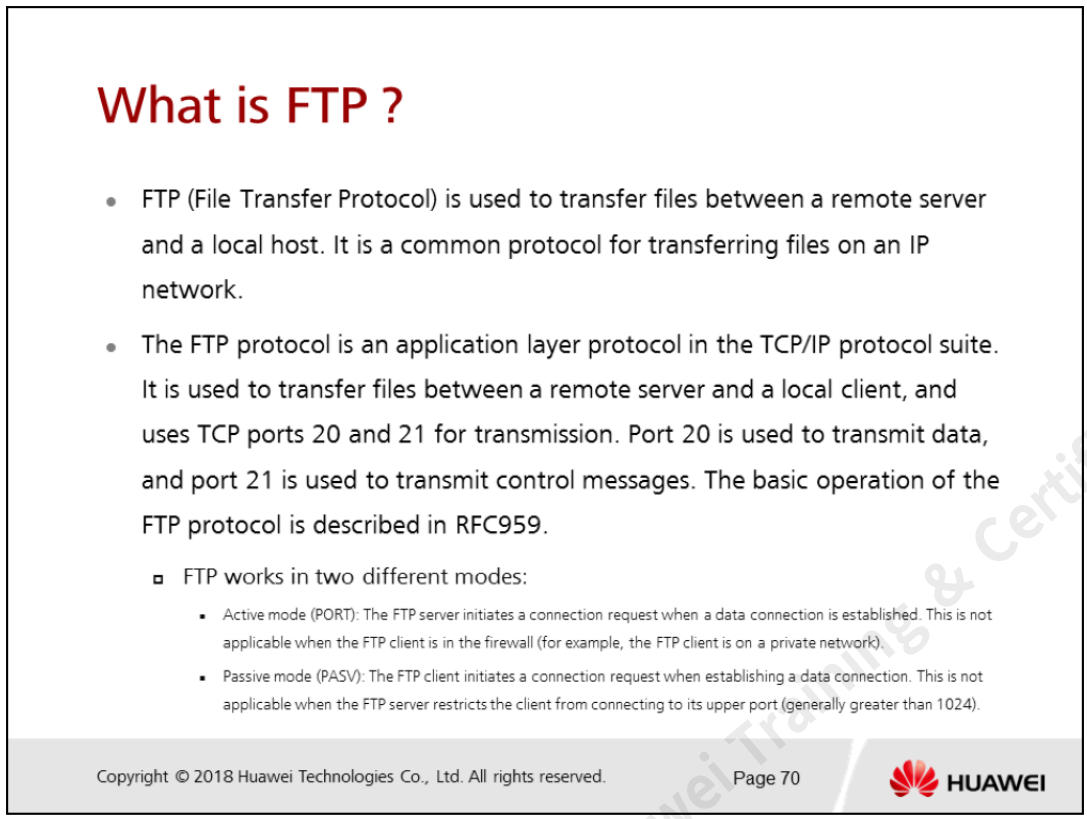
Contents

1. SCSI/iSCSI
2. SAS
3. FC/FCOE
4. PCIe
5. IB
6. CIFS/NFS
7. **FTP/HTTP**

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

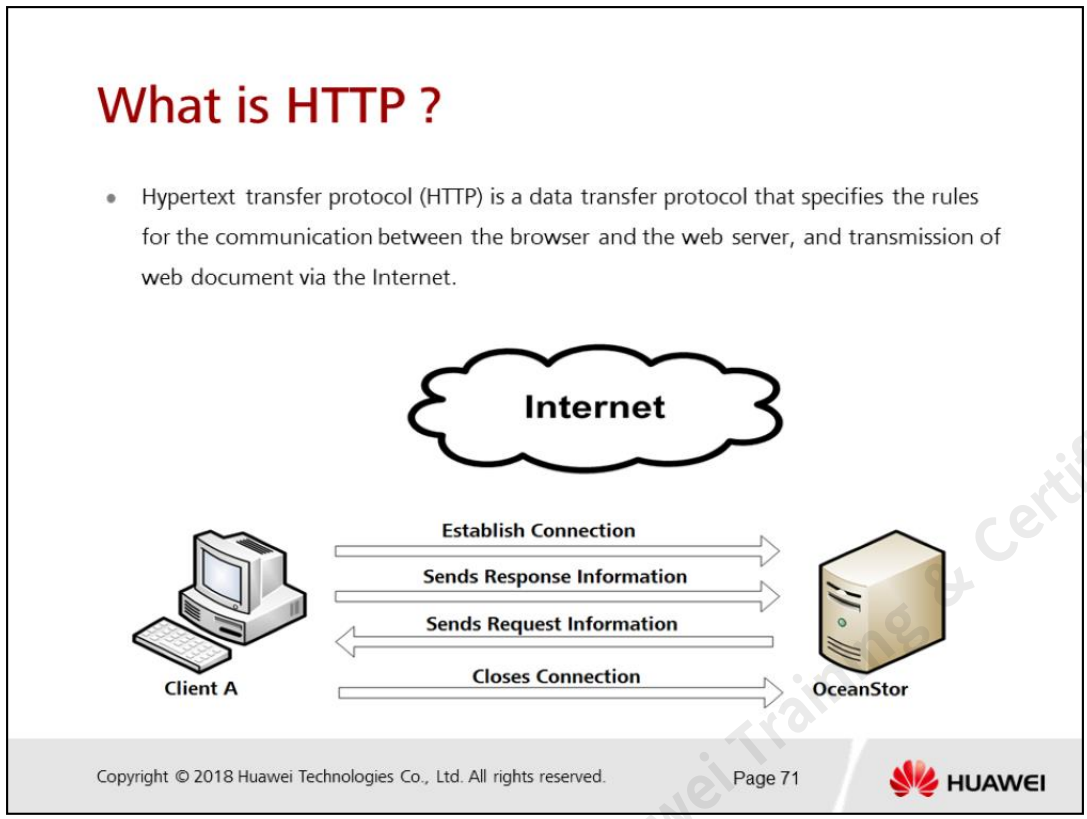
HUAWEI

Huawei Training & Certification
Huawei Training & Certification



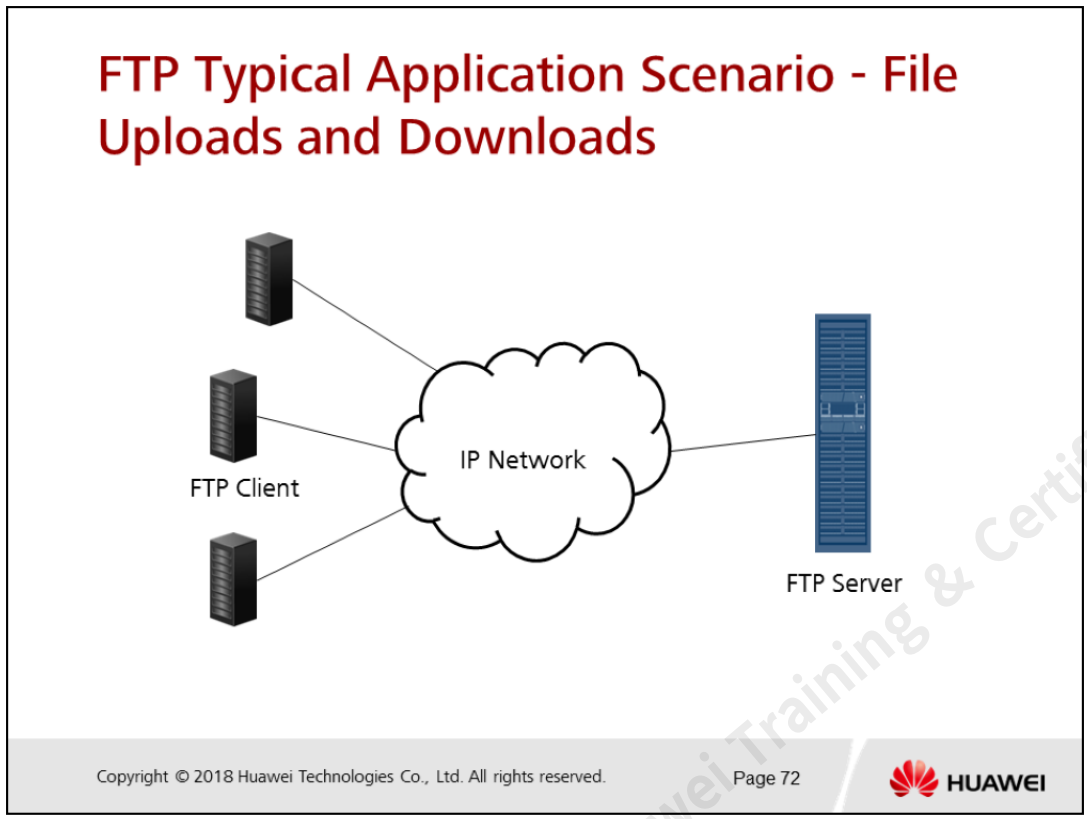
- FTP is the abbreviation of File Transfer Protocol. It is used for controlling the two-way transmission of files on the Internet. At the same time, it is also an application. There are different FTP applications based on different operating systems, and all these applications follow the same protocol to transfer files.
- In the use of FTP, users often encounter two concepts which are: "Download" and "Upload". "Downloading" a file means copying a file from a remote host to your computer and “Uploading" a file means copying the file from your computer to a remote host. In popular Internet terms, users can upload (download) files to (from) remote hosts through client programs.
- FTP has two file transfer modes:
  - Binary mode for transferring program files (e.g. files with the suffixes .app, .bin, and .btm).
  - ASCII mode for transferring text files (such as files with the suffixes .txt, .bat, and .cfg).

Huawei Training & Certification



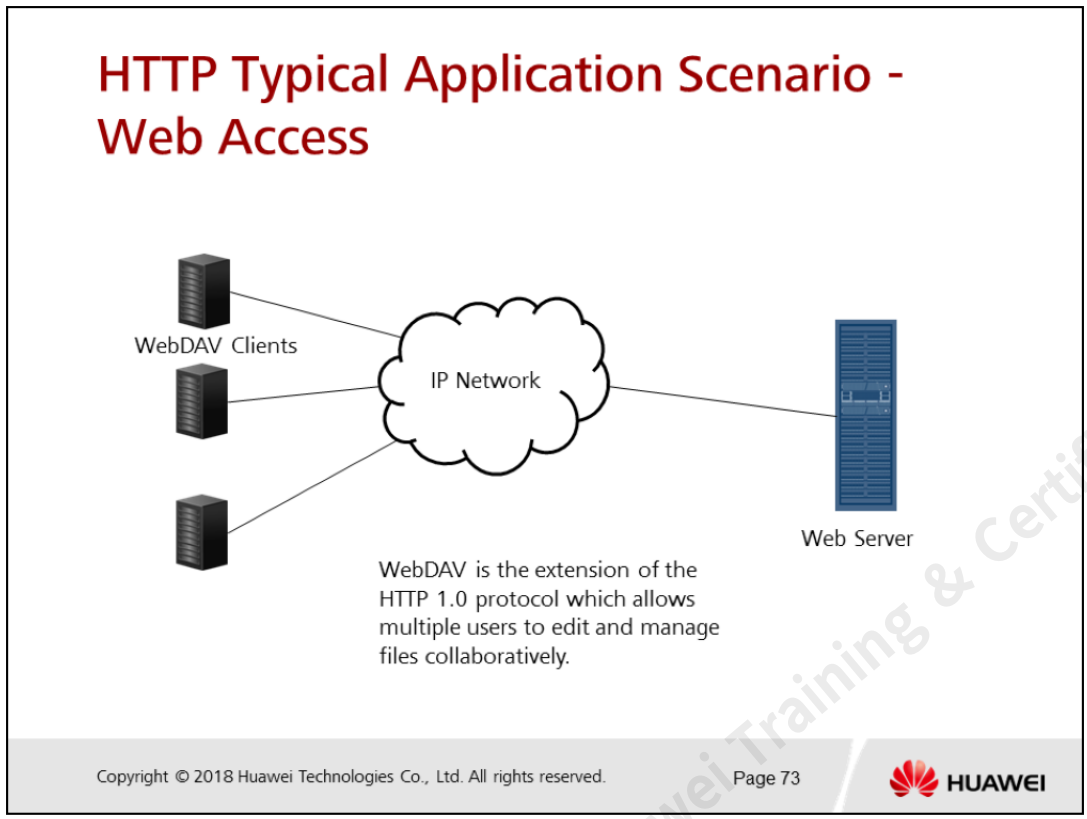
- The HTTP protocol defines how a Web client requests a Web page from a Web server and how the server delivers the Web page to the client.
- HTTP uses a short connection for packet transmission and the connection is interrupted/disconnected after each transmission.
- Different from the CIFS and NFS protocols, through the use of WebDav extension protocol and the mounted file system, it enables HTTP share creation, and the upload, download, modification, and locking of resource files.
- HTTPS is a secure version of the HTTP protocol and is based on SSL/TLS security encryption. SSL (Secure Sockets Layer) and its successor Transport Layer Security (TLS) are security protocols that provide security and data integrity for network communications. TLS and SSL encrypt the network connection at the transport layer.

Huawei Training & Certification



- FTP is used to transfer files between the remote server and the local host.
- The File Transfer Protocol (FTP) is a standard network protocol used for the transfer of computer files between a client and server on a computer network.
- FTP is built on a client-server model architecture and uses separate control and data connections between the client and the server. FTP users may authenticate themselves with a clear-text sign-in protocol, normally in the form of a username and password, but can connect anonymously if the server is configured to allow it. For secure transmission that protects the username and password, and encrypts the content, FTP is often secured with SSL/TLS (FTPS). SSH File Transfer Protocol (SFTP) is sometimes also used instead but it is technologically different.

Huawei Training & Certification



- The WEB server is also called a WWW (World WIDE WEB) server and provides users with online information browsing services through the HTTP protocol. The user sends a request to the server through the browser to browse the resource information on the server.
- Since the Web has become the basis of the Internet, HTTP 1.1 (Hypertext Transfer Protocol) has proven to be a very flexible and versatile protocol for transmitting data.
- However, there are some obvious disadvantages of HTTP, which restricts it from being adopted as a comprehensive Internet communication protocol: it is well-suited for viewing static documents, but it cannot provide enough complexity (to provide rich authoring capabilities to clients) The way to handle documents. For example, when two authors make changes to a document at the same time without communicating, there will be a "lost update" problem. Only changes made by the last author and re-uploaded to the server will remain, and changes made by another author will be lost.
- WebDAV (Web Distributed Authoring and Versioning) extends the HTTP/1.1 protocol and allows clients to publish, lock, and manage resources on the Web collaboratively.

Huawei Training & Certification

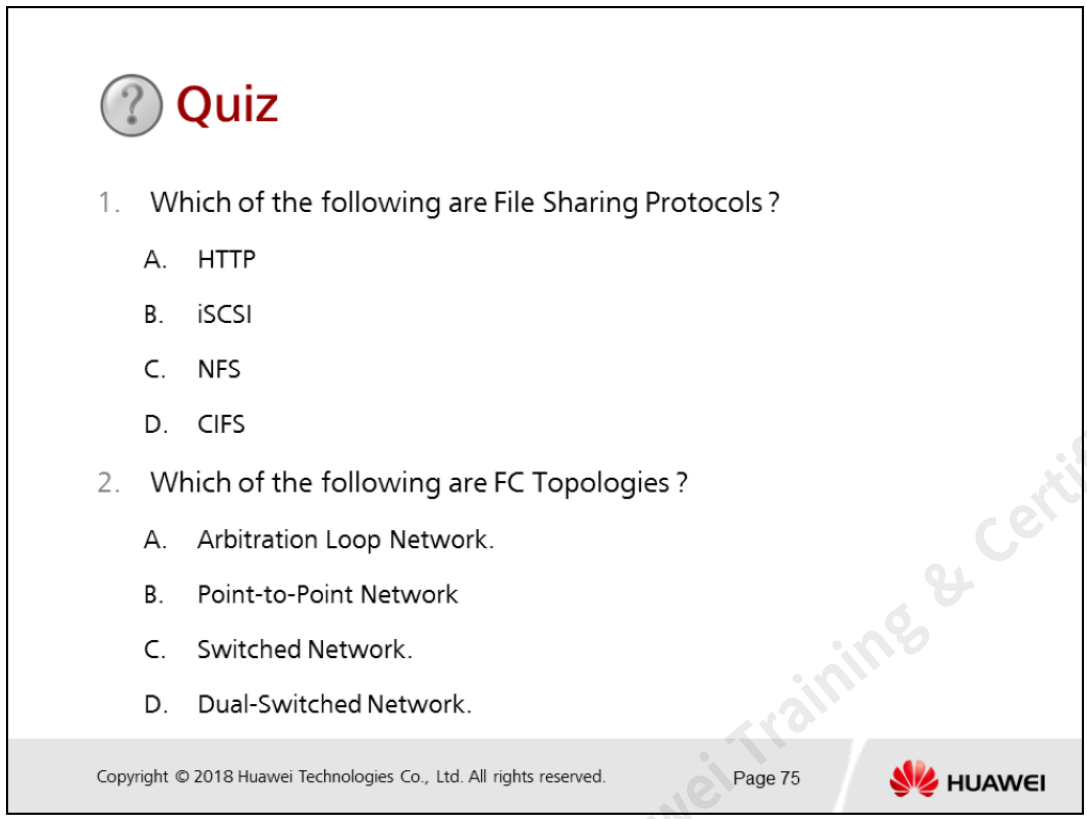
Summary

- This module mainly introduced:
  - The brief definitions of common storage protocols.
  - In depth introduction of the technical principles of common storage technologies.
  - The brief introduction of application scenarios of common storage protocols.

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

Huawei Training & Certification
Huawei Training & Certification



- Answers:
  - ACD.
  - ABC.

Huawei Training & Certification
Huawei Training & Certification

Thank You

www.huawei.com

Huawei Training & Certification
Huawei Training & Certification

06 Storage Network Technologies

www.huawei.com

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

Huawei Training & Certification
Huawei Training & Certification

Foreword

- This module introduces the types of storage network and the comparison of the advantages and disadvantages between different types of storage networks.
- This module also introduces the definition of DAS, NAS, and SAN.
- This module also describes in detail on the architecture of DAS, NAS and SAN storage network, and its advantages and disadvantages.

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

Huawei Training & Certification
Huawei Training & Certification

Objectives

- Upon completion of this module, you will be able to:
  - Describe the characteristics of DAS, NAS and SAN storage network technologies.
  - Differentiate between the advantages and disadvantages of DAS, NAS, and SAN storage network technologies.

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

Huawei Training & Certification
Huawei Training & Certification

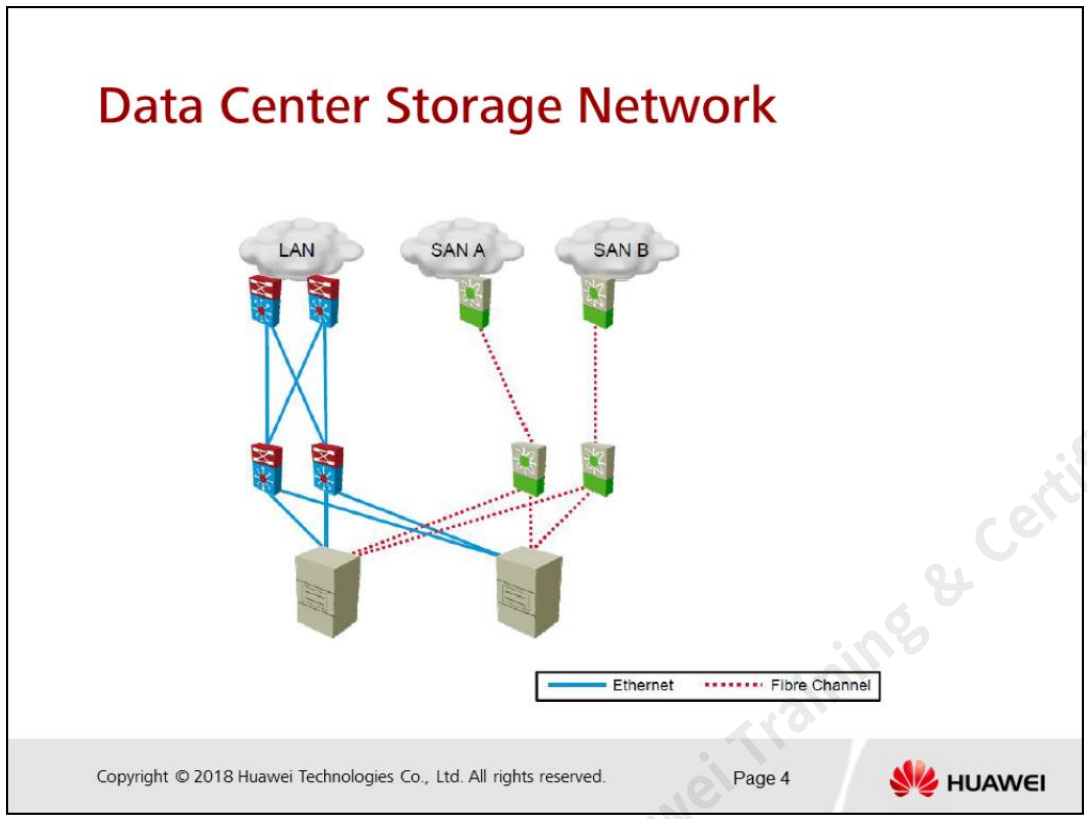
Contents

1. **Classification of Network Storages.**
2. DAS - Traditional Storage Technology.
3. SAN - Storage Area Network.
4. NAS - Network Attached Storage.

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

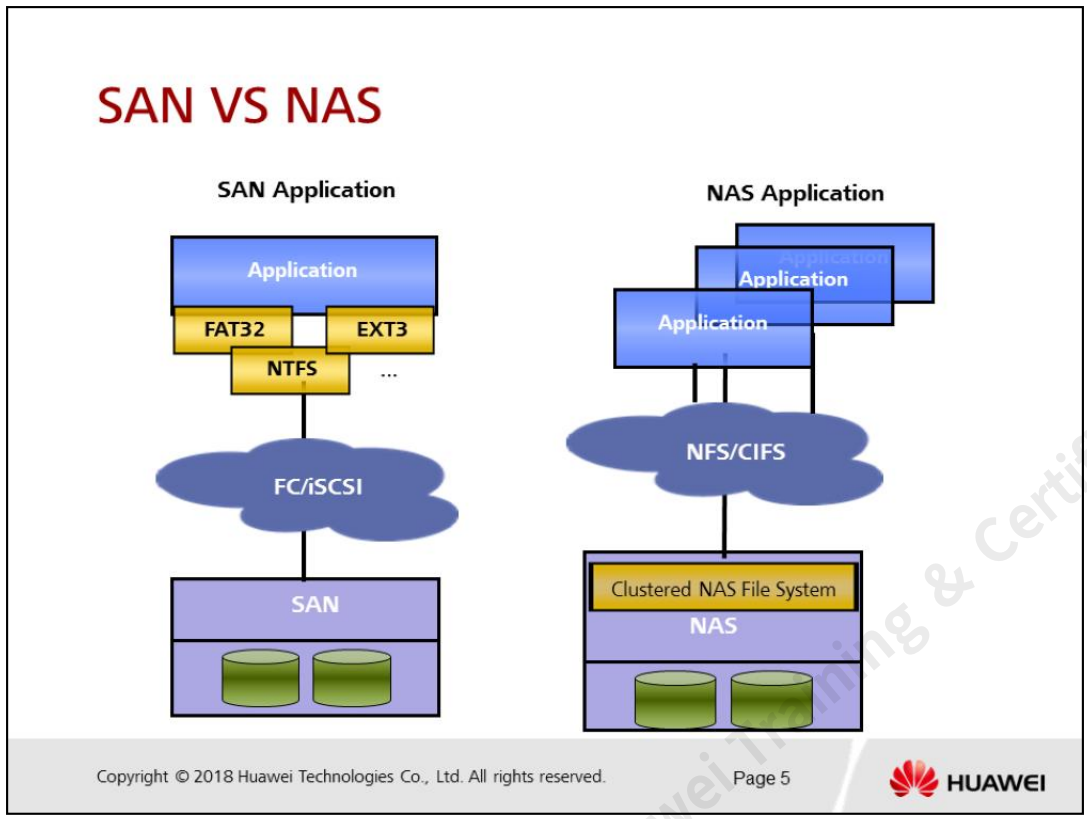
HUAWEI

Huawei Training & Certification
Huawei Training & Certification



- Data storage has become one of the popular technologies nowadays and created a new technological trend after the emergence of the Internet. It brings the web into a data centric era.
- Data storage went through 3 stages of development: Direct Attached Storage (DAS), Network Attached Storage (NAS), and Storage Area Network (SAN). DAS is a product of CPU centric era and adapted to the initial development of the computing industry. The biggest technological differences between SAN and NAS is the way they use dedicated protocol or existing IP technology and data sharing. The advantage of SAN is that it solved the initial problem of network bandwidth for storage issue, while the advantage of NAS is its versatility and data sharing capabilities.

Huawei Training & Certification



- NAS (Network-Attached Storage)
  - Protocols: NFS/CIFS (Based on TCP/IP).
  - Provides file access, and suitable for file storage requirements.
  - Suitable for file sharing and file access application, supports heterogeneous file sharing platform.
  - File data migration is much simpler and reliable than raw device.
- SAN (Storage Area Network)
  - Protocols: FC/iSCSI
  - Supports raw device access, suitable for traditional database access.
  - Relies on application host to provide file access. Data sharing access requires the support of the cluster software, high overhead on processing access conflicts, lower performance, and hard to support data sharing between heterogeneous platforms.
  - Raw device data migration is difficult.

Huawei Training & Certification

Comparison of DAS, NAS and SAN (1)

Storage System Architecture	DAS	NAS	SAN
Data Transmission Protocol	SCSI/FC/ATA	TCP/IP	FC
Data Transfer Object	Block	File	Block
Uses Standard File Sharing Protocols	No	Yes (NFS/CIFS...)	No
Centralized Management	Not Exactly	Yes	Requires Management Tools
Increases Server Efficiency	No	Yes	Yes
Disaster Tolerance	Low	High	High, Dedicated Solutions

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

- The table above shows the comparison of DAS, NAS and SAN storage architecture.

Huawei Training & Certification

Comparison of DAS, NAS and SAN (2)

Storage System Architecture	DAS	NAS	SAN
Suitable Clients	SME Servers, JBOD	SME, Monitoring, Broadcasting	Large Enterprises, Data Centers
Application Environment	Local Area Network(LAN). Low file sharing level, independent operating platform, and less server management	Local Area Network(LAN). High file sharing level, high demand in heterogeneous storage format.	Fibre Channel Storage Network. Complex network environment, high file sharing level, heterogeneous OS platform, high number of servers.
Capacity Expansion Capabilities	Low	Medium	High

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

Huawei Training & Certification
Huawei Training & Certification

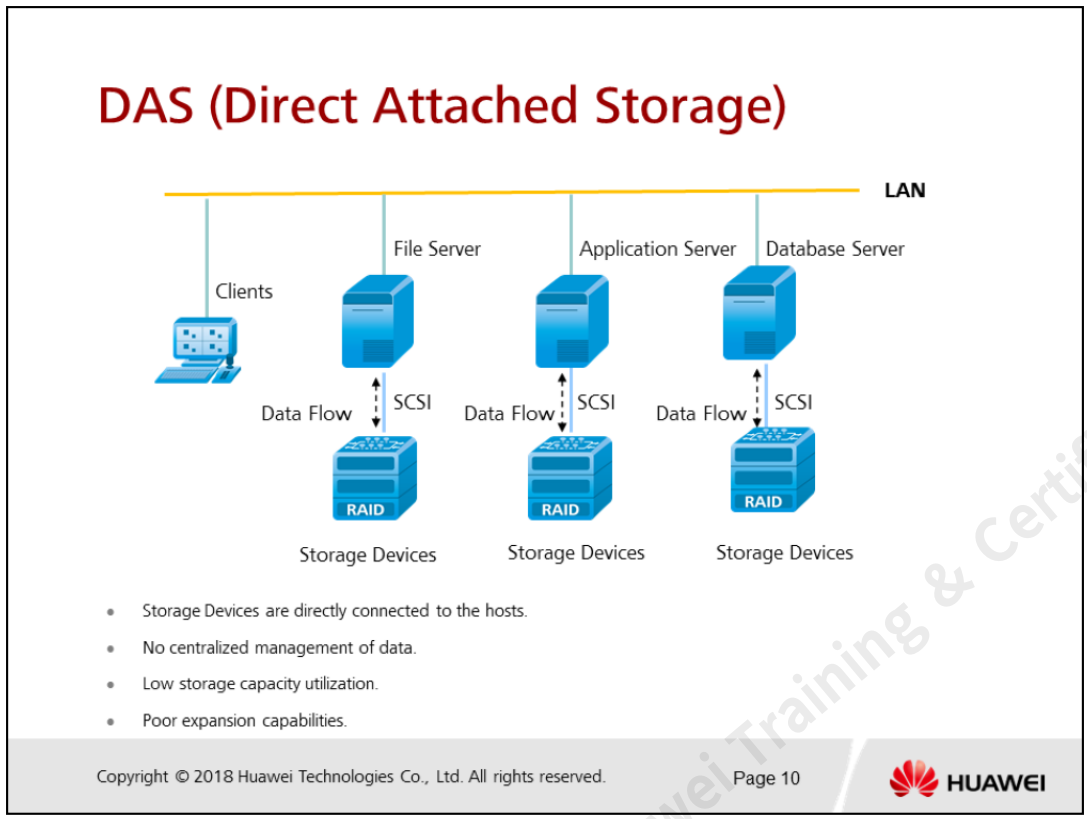
Contents

1. Classification of Network Storages.
2. **DAS - Traditional Storage Technology.**
3. SAN - Storage Area Network.
4. NAS - Network Attached Storage.

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

Huawei Training & Certification
Huawei Training & Certification

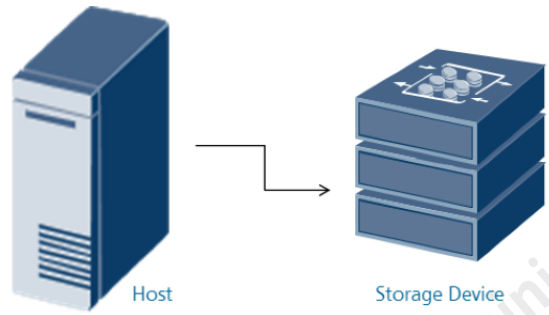


- DAS is very suitable for small or medium sized local area networks that has low storage capacity requirements and fewer number of servers. Its main advantage is easy and quick implementation of storage capacity expansion with small cost of investment.
- A typical DAS system is made of a data storage device (for example enclosures holding a number of hard disk drives) connected directly to a computer through a host bus adapter (HBA). The main characteristics of DAS is that between those two points (server and storage device), there is no network device (like hub, switch, or router).

Huawei Training & Certification

Definition of DAS

- Definition of DAS: DAS refers to 1 or more storage devices that are directly connected to the servers which provides block storage services to the servers.
- Classification of DAS: Based on the location of those storage devices relative to the servers, DAS is divided into Internal DAS and External DAS.

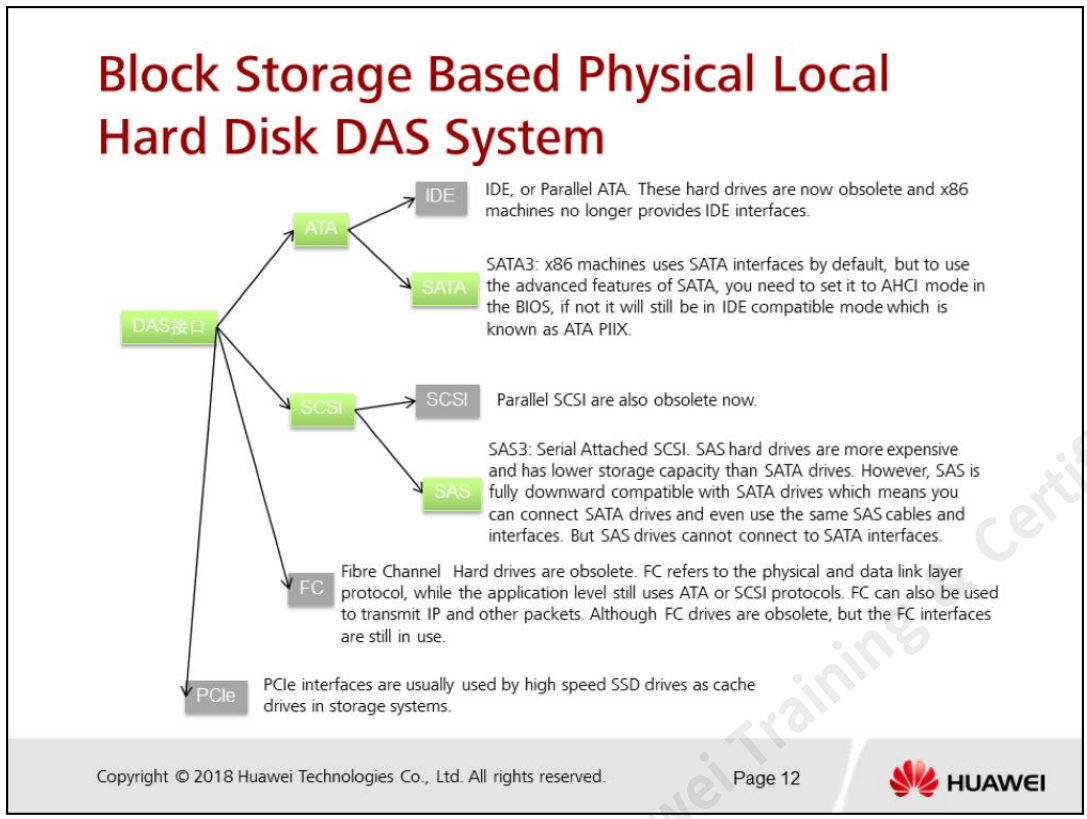


Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

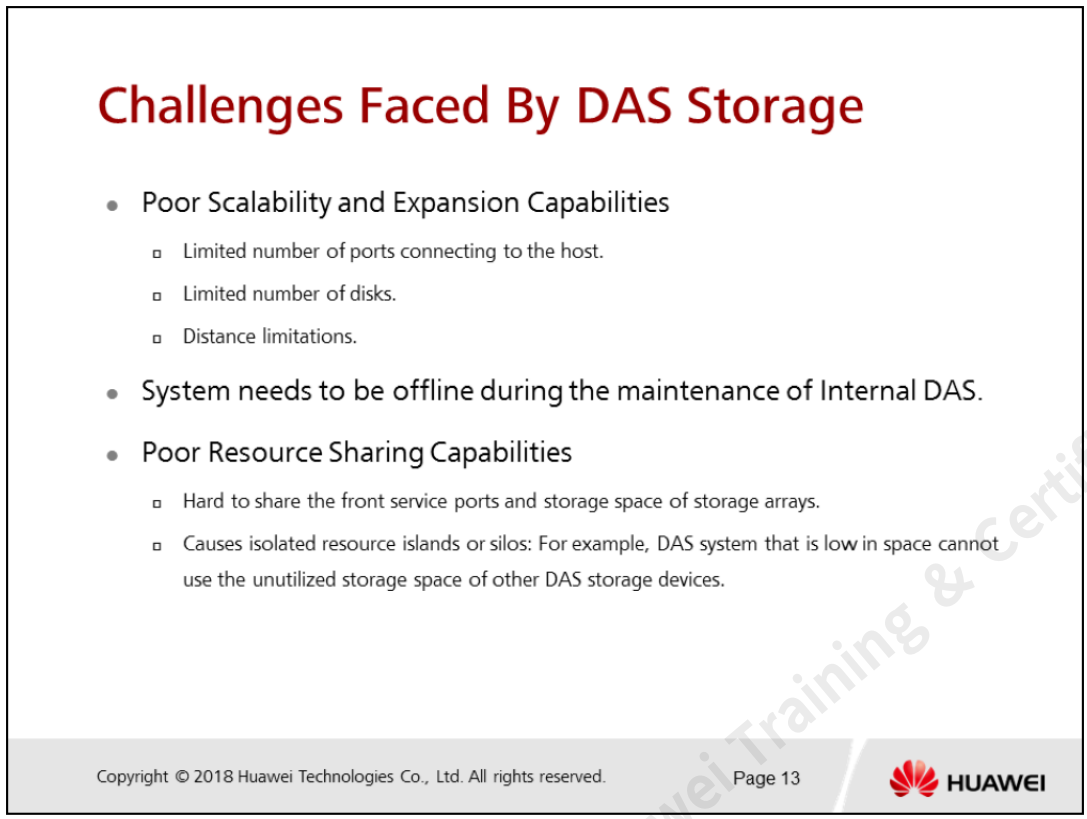
Huawei Training & Certification

- Directly Attached Storage (DAS): is a storage architecture where the storage devices are directly connected to the servers. DAS provides block storage (not file storage) services to the servers. Example of DAS are: internal hard drive of servers, directly connected tape libraries, directly connected external disk enclosures. Based on the location of the storage device relative to the server, DAS can be divided into Internal DAS and External DAS.
- Internal DAS:
  - In the architecture of Internal DAS, storage devices connects to the server internally via parallel or serial connections. However, due to the physical distance limitation of the system bus, it could only support short distance high speed data transmission. Additionally, there is also limitation on the number of devices that can be connected internally. It takes up a lot space when these storage devices are placed within the server's chassis, causing difficulties for maintenance of other parts within the server.
- External DAS:
  - In the architecture of External DAS, external storage devices are connected directly with the servers, and not located within the servers itself. In most scenarios, these storage devices communicate between each other using FibreChannel (FC) protocols or SCSI protocols. In comparison with Internal DAS, External DAS has overcame the limitation of distance and limitation of number of connected devices. Other than that, External DAS even provides centralized management of storage devices making management and maintenance of these devices much easier.



- Direct-Attached Storage(DAS), can be explained simply as using local hard drives for storage and these drives are connected using local interface. X86 architecture usually uses ATA (IDE or SATA) interfaces, minicomputers and high end x86 servers uses SCSI (SCSI or SAS) interface. In high performance scenarios, FC optical interface are used. However, with the developments of serial interface technologies, even the slower SATA interface speed also have exceeded the speed of hard drives read/write speed, which means that SATA interface speed is sufficient for even SSD drives. As the serial interface technology such as SATA becomes much faster, the FC hard drives becomes obsolete and SAS usage are gradually decreasing. High speed SSD cache drives on the other hand uses PCIe interface directly for much higher performance.

Huawei Training & Certification, Huawei Training & Certification



- DAS has poor scalability and expansion capabilities. Limited number of usable host ports caused a limitation on number of host that can be connected to the storage. At the same time, limited bandwidth of DAS devices caused it to have limited IO processing capabilities. When the IO processing requirement from the connected host is high and reaching the upper limit of the DAS devices IO processing capabilities, the service on the host will be affected and even will cause a chain reaction affecting the performances of all the connected host of the DAS storage.
- Due to the limited sharing capabilities of the front end host ports, it caused DAS to have low resource utilization rate. In the DAS system, to reallocate and configure the remaining unutilized storage space is difficult causing the formation of resource islands or silos. As some of the DAS devices are under utilized with remaining lots of space and some are over utilized with critically low storage space, the fact that DAS devices cannot do resource sharing between devices does not able to help or solve this situation. Additionally, disk utilization rate, throughput, internal memory of DAS devices also affects the performance of DAS storage systems. Additional factors such as RAID group level configuration, host storage controller protocol type and host bus efficiency also affects the performance of DAS system. The main advantage of DAS system compared to other storage architecture lies only on the fact that it does not have the complexity for devices interconnection and network latency between storage devices.

Huawei Training & Certification

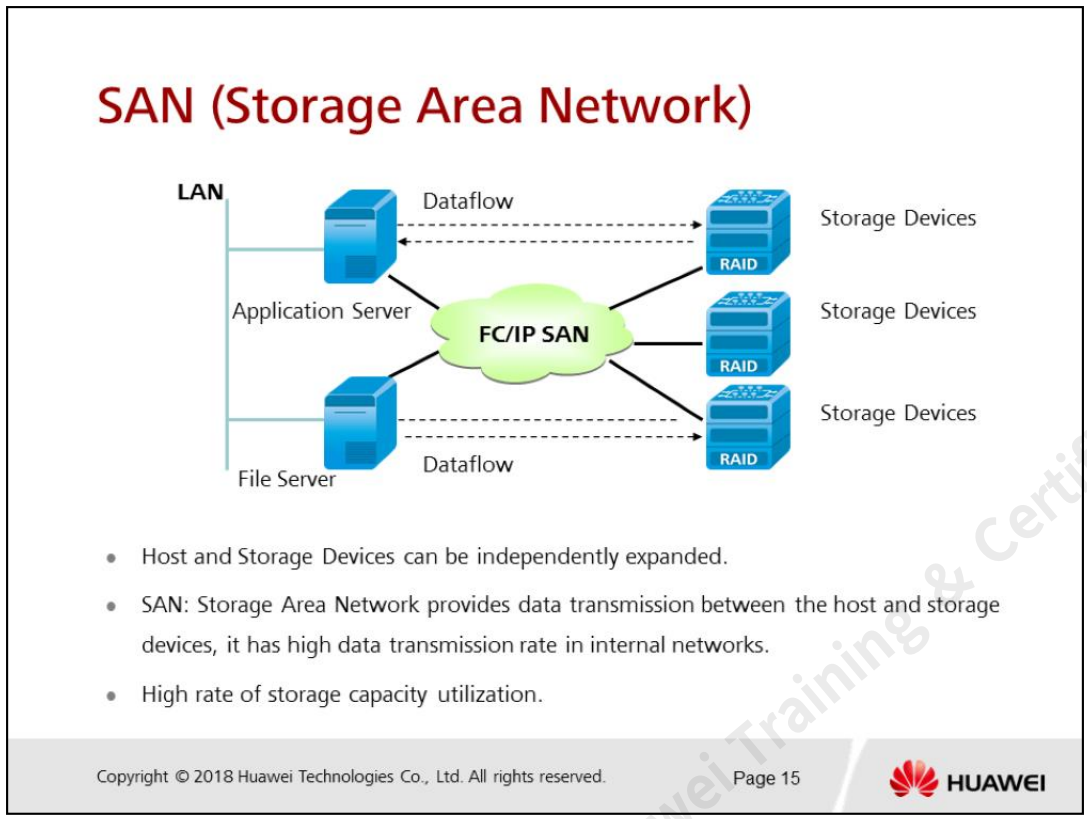
Contents

1. Classification of Network Storages.
2. DAS - Traditional Storage Technology.
3. **SAN - Storage Area Network.**
4. NAS - Network Attached Storage.

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

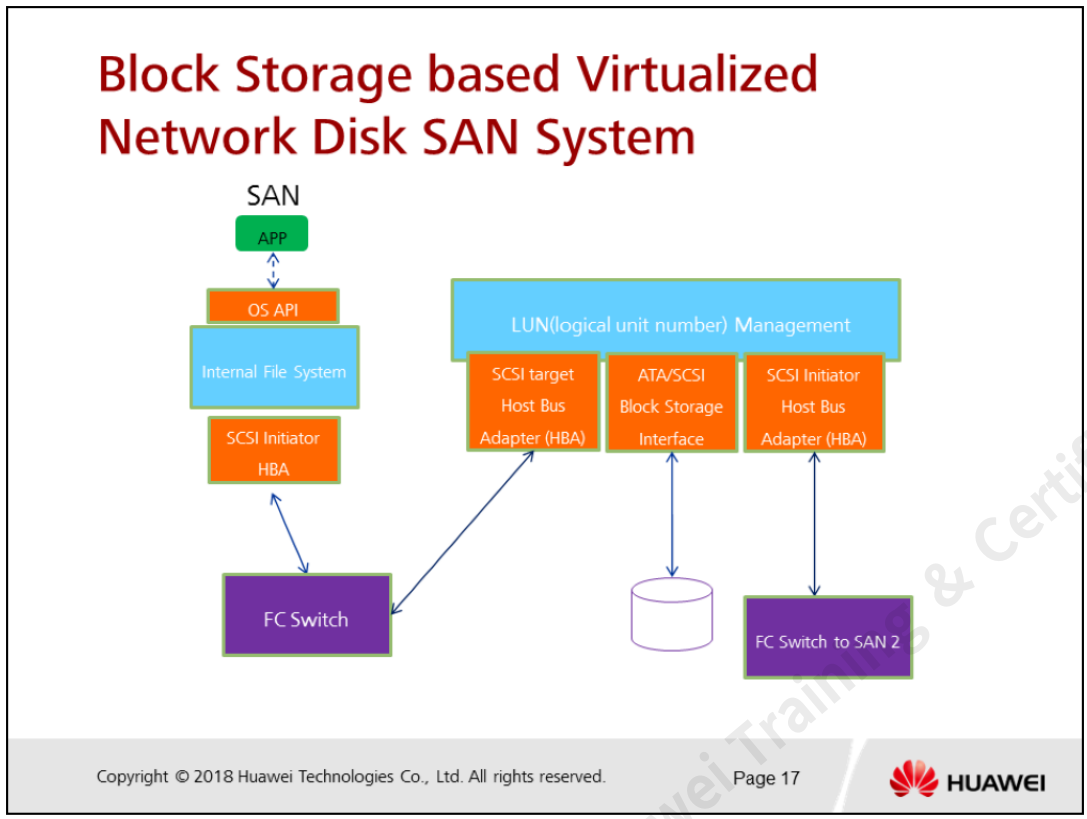
HUAWEI

Huawei Training & Certification
Huawei Training & Certification



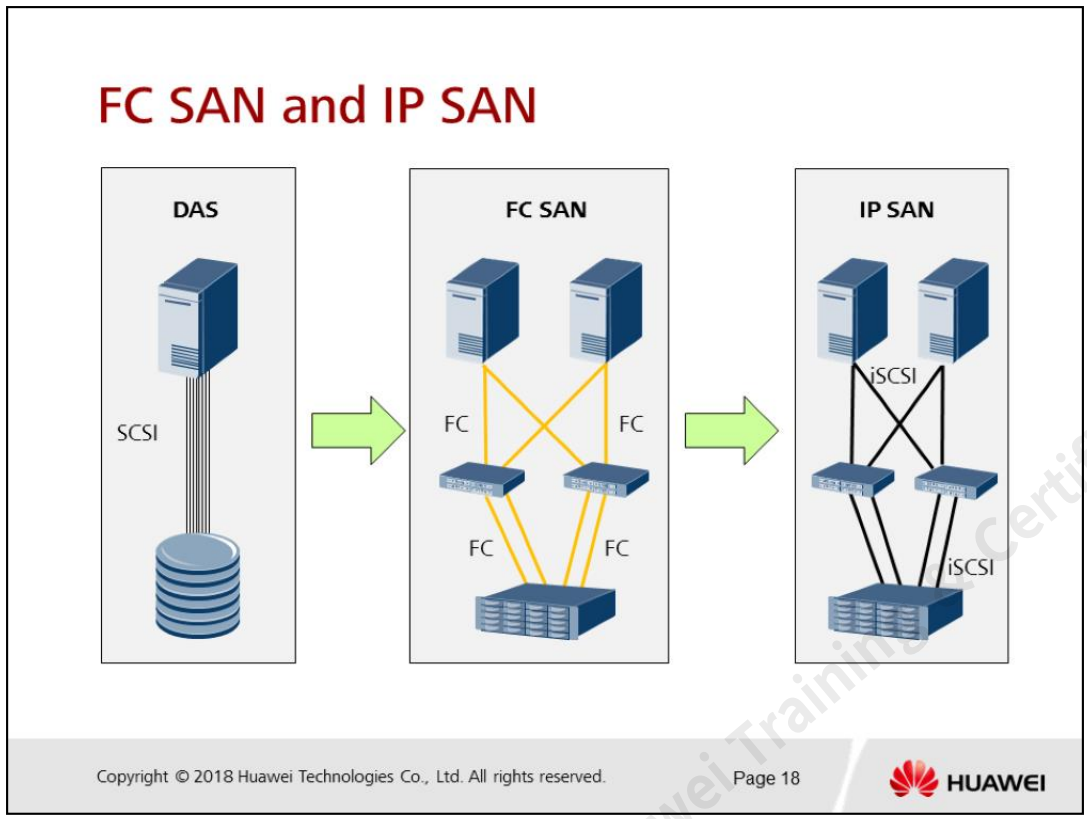
- SAN is the abbreviation of Storage Area Network, it is a type of dedicated network for high speed reliable data access and transmission between host and external storage devices or between independent storage resources.
- SAN implements an expandable and scalable network topology connecting the servers and storage devices, and each storage devices does not belong to a dedicated server, as it allows equal level of resource sharing between all of the storage devices and hosts within the network.
- A complete SAN includes: host devices that support SAN, storage devices that support SAN, connectivity devices used to connect all the devices in the SAN, management software that supports SAN, and finally services that supports SAN.
- Storage Area Network (SAN) : is a dedicated high performance network implemented between the server and storage resources. It has been optimized specifically for high amount of raw data transmission. Hence, FC SAN can be seen as an expanded form of SCSI protocols in terms of long distance application. The default protocols used by FC SAN is SCSI and Fiber Channel. Fiber Channel is especially suitable for SAN implementation due to the fact that it is capable of transmitting large blocks of data and is also capable of long distance transmission. The market for FC SAN is mostly centered at high end enterprise grade storage devices implementations. These storage implementations has high requirements for performance, redundancy and availability. For an example, components such as storage arrays and backup devices are storage devices.

Huawei Training & Certification



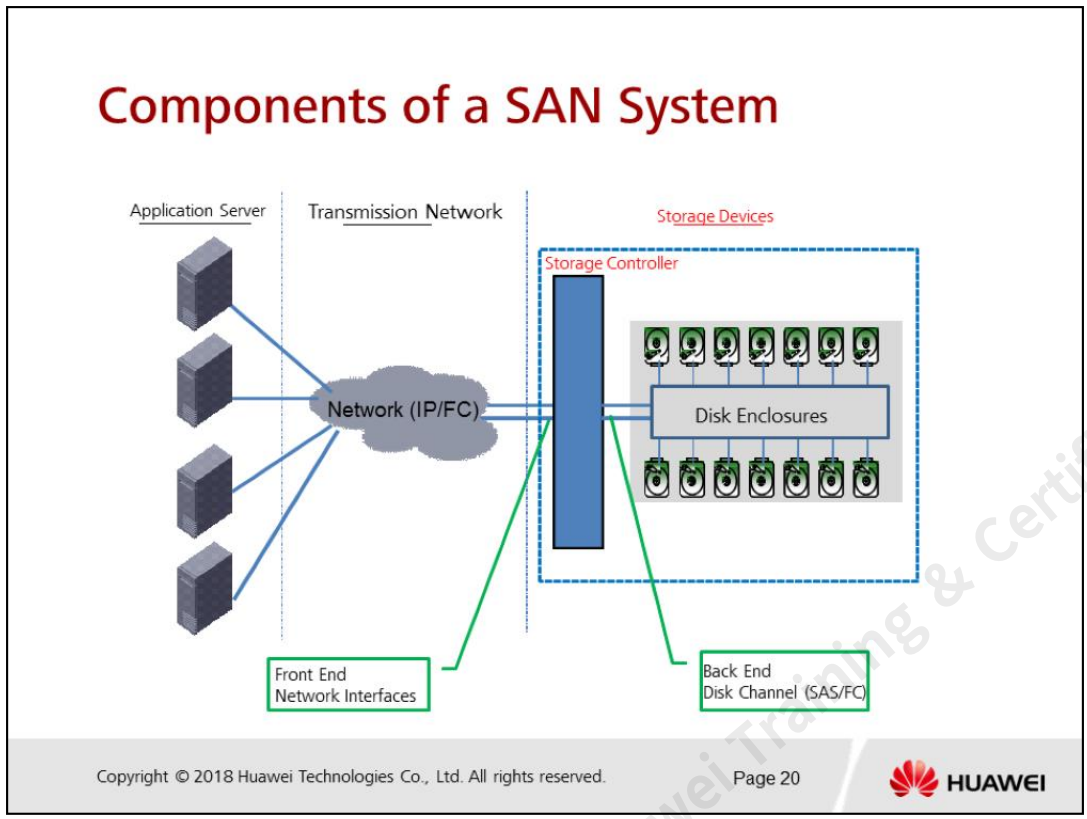
- SAN provides block storage similar to DAS, but also has the advantages of remote connection using network similar to NAS.
- In computer storage, a logical unit number (LUN), is a number used to identify a logical unit, which is a device addressed by the SCSI protocol or Storage Area Network protocols which encapsulate SCSI, such as Fibre Channel or iSCSI.
- SCSI initiator is the endpoint that initiates a SCSI session, that is, sends a SCSI command. The initiator usually does not provide any Logical Unit Numbers (LUNs).
- On the other hand, a SCSI target is the endpoint that does not initiate sessions, but instead waits for initiator's commands and provides required input/output data transfers. The target usually provides to the initiators one or more LUNs, because otherwise no read or write command would be possible.

Huawei Training & Certification



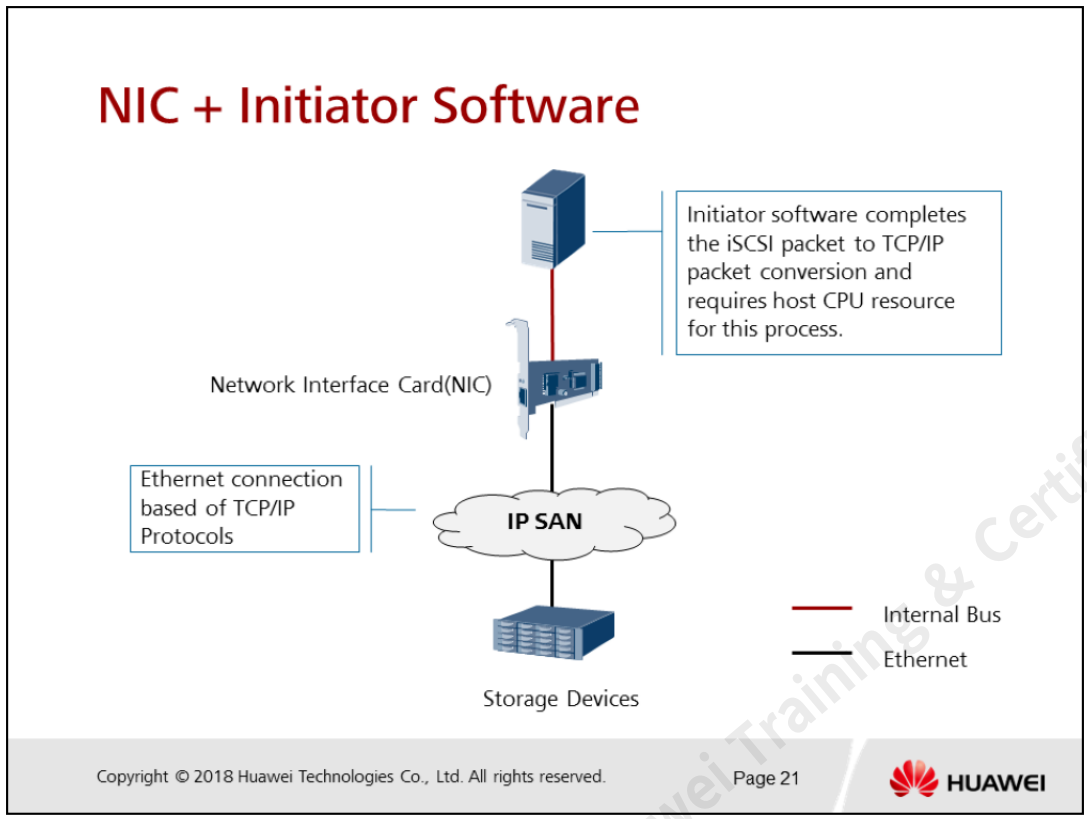
- What is IP SAN?
  - It uses TCP/IP as the base level transmission protocol, and is a storage area network architecture that uses Ethernet as the transmission medium.
  - Default protocol for IP SAN is iSCSI, which defines the encapsulation of SCSI commands in packets and the transmission of those packets over the IP network.
- IP-SAN default networking methods are:
  - Direct Connection: The host and storage devices are directly connected via Ethernet network cards, TOE (TCP Offload Engine) cards, or iSCSI HBA cards. These types of connection are simple, economical but will be difficult to share storage resources in scenarios where there a large number of hosts.
  - Single Switching: Hosts and storage devices are connected through a single Ethernet switch, where the connection on the ends of the host and storages are still using Ethernet network cards, TOE cards or iSCSI HBA cards. This type of connection allows multiple host to share the same storage device, has good scalability, but has a single point of failure at the network switch. If the network switch is faulty, all the host are unable to access the storage devices.

Huawei Training & Certification



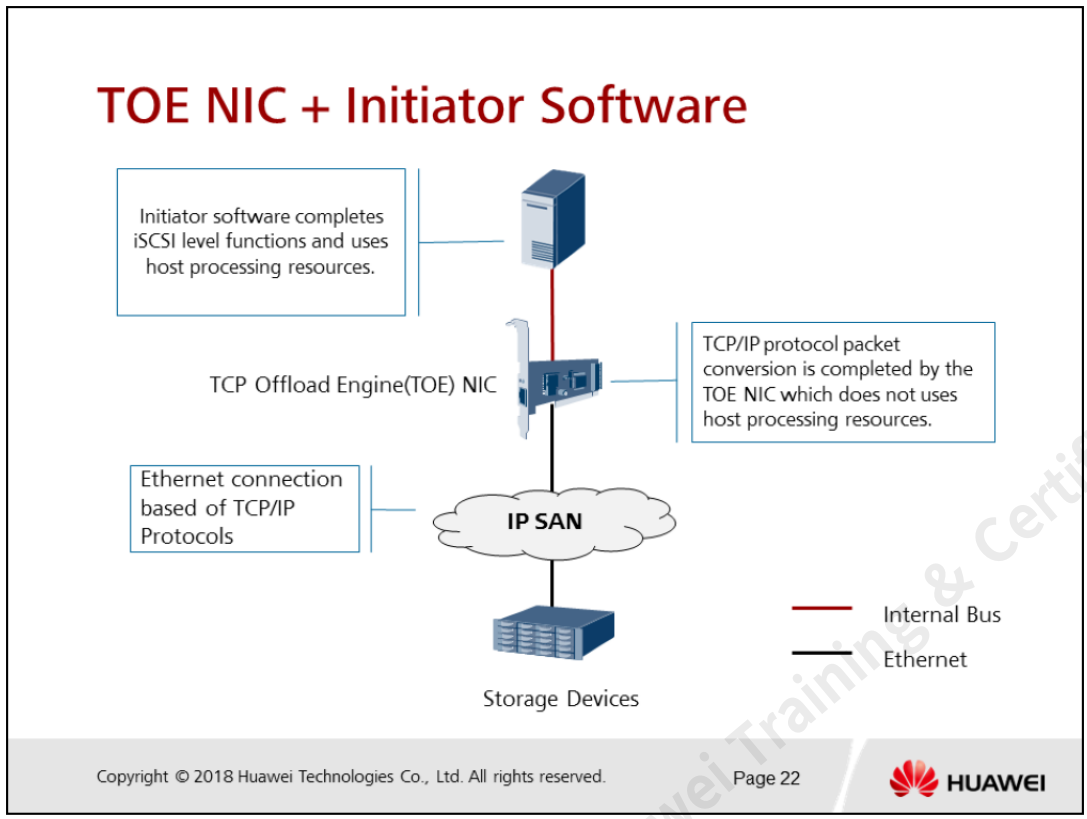
- As shown on the diagram above, a complete SAN system commonly includes the front end application server, transmission network, and storage devices. Servers refers to x86 servers, minicomputers or clustered servers that are running operating systems such as Windows, Linux or Unix. Transmission network can be FC network or IP network, and the types of network and interface cards used within the SAN depends on the type of network that the key storage equipment supports.
- Storage devices consists of the front end network interfaces storage controllers, back end disk channel, disk enclosures and hard drives. Storage controller is the core of the storage device. It is a manifestation of processing performance, and can have corresponding optimization measures for various protocol architectures. There is no obvious difference among various storage devices.
- Hence, the difference in storage devices lies at the front end network interfaces (IP or FC), and the back end disk channel (SAS or FC). Based on the different combination of the front end and back end technologies, it formed 4 main storage device architecture: which are IP+SAS, IP+FC, FC+SAS, FC+FC. Commonly, FC SAN refers to storage devices that connects to server using FC protocols through the FC network interface, while IP SAN refers to the storage devices that connects to the server using iSCSI protocol through IP network interface.

Huawei Training & Certification



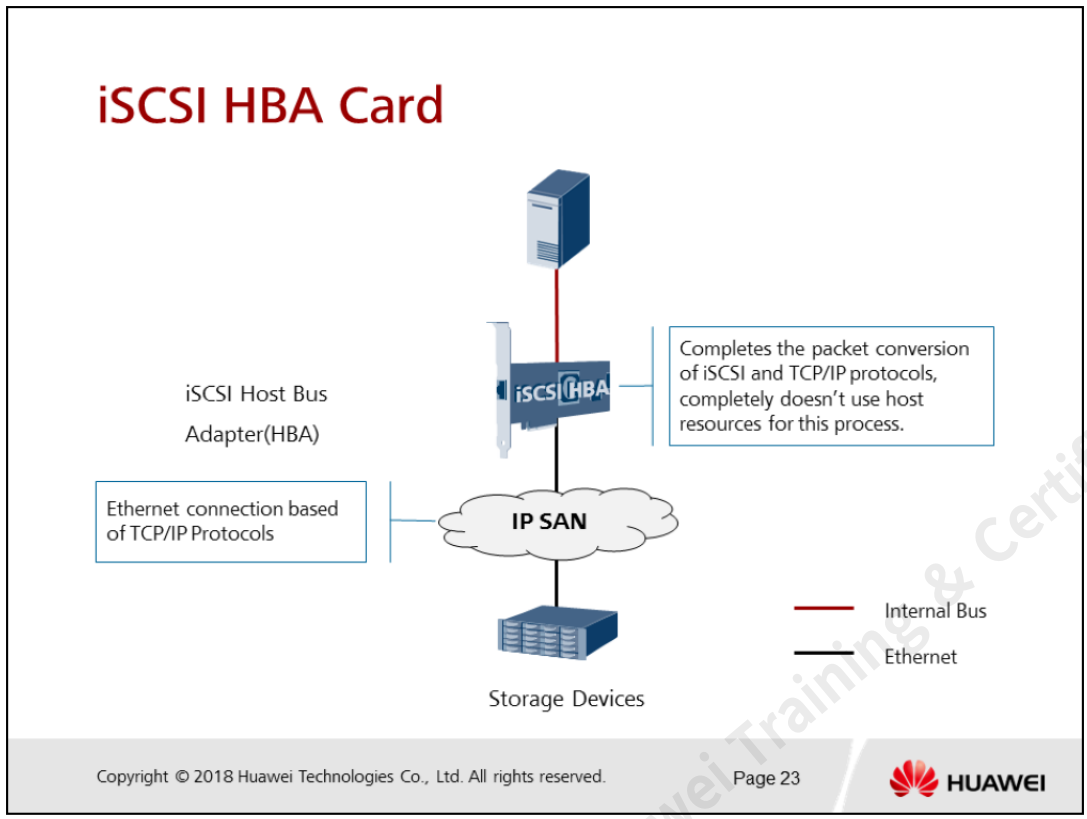
- Host devices such as servers and workstations, connects to Ethernet switches using standard NIC cards, and iSCSI storage devices are also connected to Ethernet switches or directly to the host NIC card. Initiator software is installed on the host to virtualize the Ethernet card as an iSCSI card which is used to send and receive iSCSI packets and makes both iSCSI and TCP/IP protocol data transmission possible between the host and the iSCSI storage device. This implementation has the lowest hardware cost due to the use of standard Ethernet cards and Ethernet switches, eliminating the need for additional adapters. However, this mode of implementation needs to occupy host processing resources when performing iSCSI and TCP / IP data packet conversion, which in turn increases host operating overhead and reduces system performance. However, the basic requirements for data access are met in applications or scenarios with low I/O and bandwidth requirements.

Huawei Training & Certification



- TOE NIC cards handles the TCP/IP protocol level functions, while the host processes the iSCSI protocol level functions. Hence, by using TOE NIC allows a high increase in data transmission rate. Compared to the method where only software is used, this method greatly reduces the host operation overhead and does not increase the cost of building the network greatly which serves as a configuration solutions that offers good compromise for performance and costing.
- The TCP/IP offload engine is a technology that is gaining popularity in high-speed Ethernet systems for the purpose of optimizing throughput. TOE components are incorporated into one of the printed circuit boards, such as the network interface card (NIC) or the host bus adapter (HBA).
- In recent years, the communication speed in Ethernet systems has increased faster than computer processor speed. This produces an input/output (I/O) bottleneck. The processor, which is designed primarily for computing and not for I/O, cannot keep up with the data flowing through the network. As a result, the TCP/IP flow is processed at a rate slower than the speed of the network. TOE solves this problem by removing the burden (offloading) from the microprocessor and I/O subsystem.

Huawei Training & Certification



- iSCSI HBA cards are installed on the hosts to achieve highly efficient data transmission between host and switches and between host and storage devices. iSCSI and TCP/IP protocol level functions are both completed by the HBA card installed on host, which bring the host CPU resource consumption during data transmission to the minimum. This method provides the best data transmission performance, however it comes with the highest cost to build this network compared to other methods.
- The iSCSI communication system inherits some of the features of SCSI protocol. In iSCSI communication, there is an Initiator device that initiates I/O requests and a Target device that responds to requests and performs actual I/O operations. After a connection is established between the Initiator and Target, the Target acts as master in the data transmission operation and controls the entire process. Examples of the target includes: iSCSI disk arrays and iSCSI tape libraries.

Huawei Training & Certification

Comparison between FC SAN and IP SAN

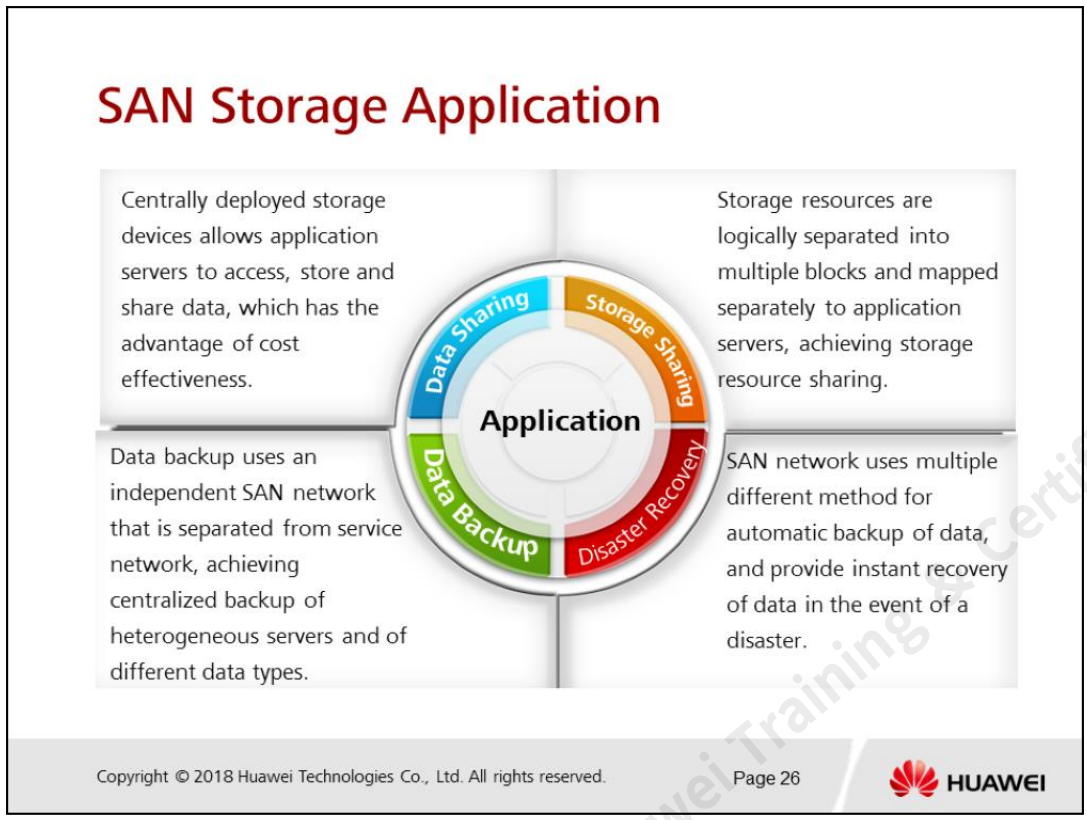
Description	FC SAN	IP SAN
Network Speed	4Gb、8Gb、16Gb	1Gb、10Gb、40Gb
Network Architecture	Custom built optical network and HBA cards.	Uses existing IP networks.
Transmission Distance	Limited by optical transmission distance.	Theoretically no distance limitation.
Management & Maintenance	Complex management and technologies.	Easy operation like normal IP devices.
Compatibility	Poor compatibility.	Compatible with all IP network devices.
Performance	Very high transmission and read/write performance.	Mainstream of 1Gb/s, and 10 Gb/s speed is still growing.
Cost	Purchase (FC Switches, HBA cards, FC Disk Arrays),and Maintenance (Train Personnel, System Configuration and Monitoring) costs are high.	Compared to FC SAN, the purchase and maintenance cost are lower with higher ROI.
Disaster Recovery	Disaster recovery hardware and software cost are high.	Can achieve local and remote disaster recovery with lower cost.
Security	Higher security.	Lower security.

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

- IP SAN only requires very less hardware configurations, and these hardware are widely in use, so this makes the cost of implementing IP SAN much lower than FC SAN. Most of the host are already equipped with suitable network interface cards (NIC) and connected to suitable switches, and these NIC and switches are also suitable (although not perfect) for iSCSI protocol based data transmission. High performance IP SAN on the other hand often requires custom iSCSI HBA cards to be installed and connected to high end switches in order to achieve high performance data access and transmission.

Huawei Training & Certification



- A SAN storage solution has a high total cost of ownership (hardware and experienced IT personnel), so SAN is mostly used in medium or large enterprises. Data requirement of these companies are usually:
  - Has high requirement regarding the response time, reliability, scalability for critical database applications.
  - Centralized backups and high storage performance, and also data integrity and data reliability.

Huawei Training & Certification

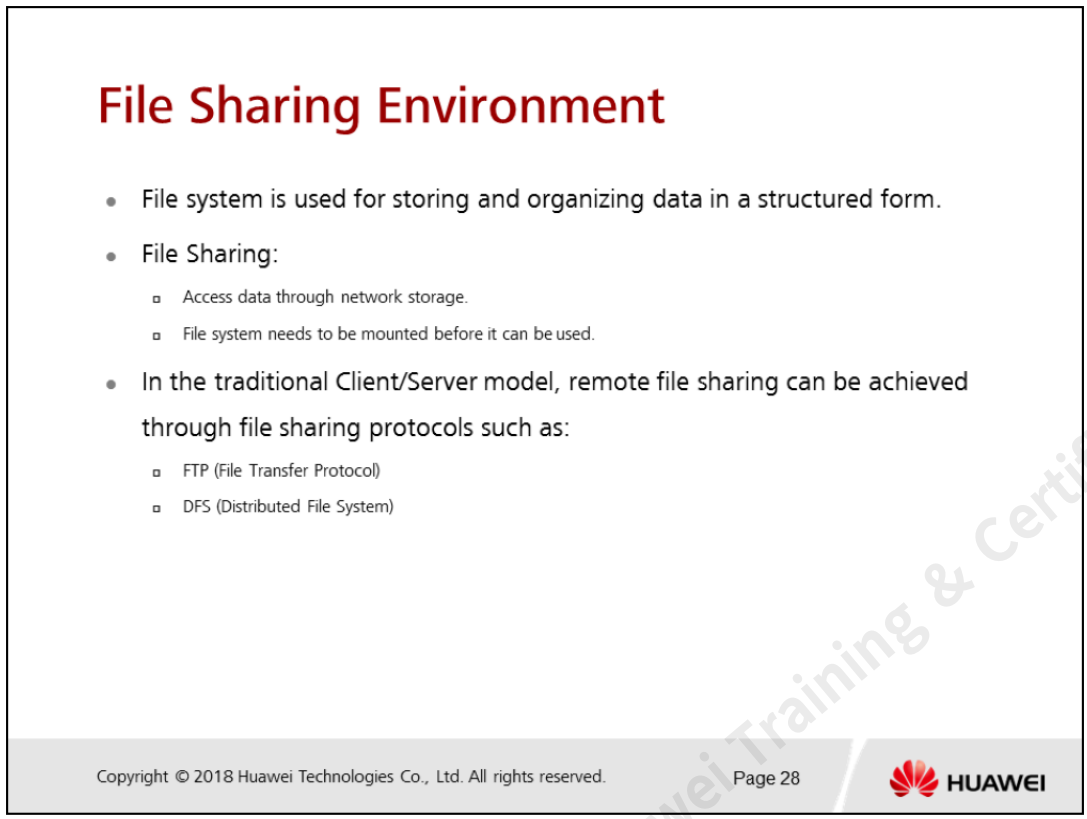
Contents

1. Classification of Network Storages.
2. DAS - Traditional Storage Technology.
3. SAN - Storage Area Network.
4. **NAS - Network Attached Storage.**

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

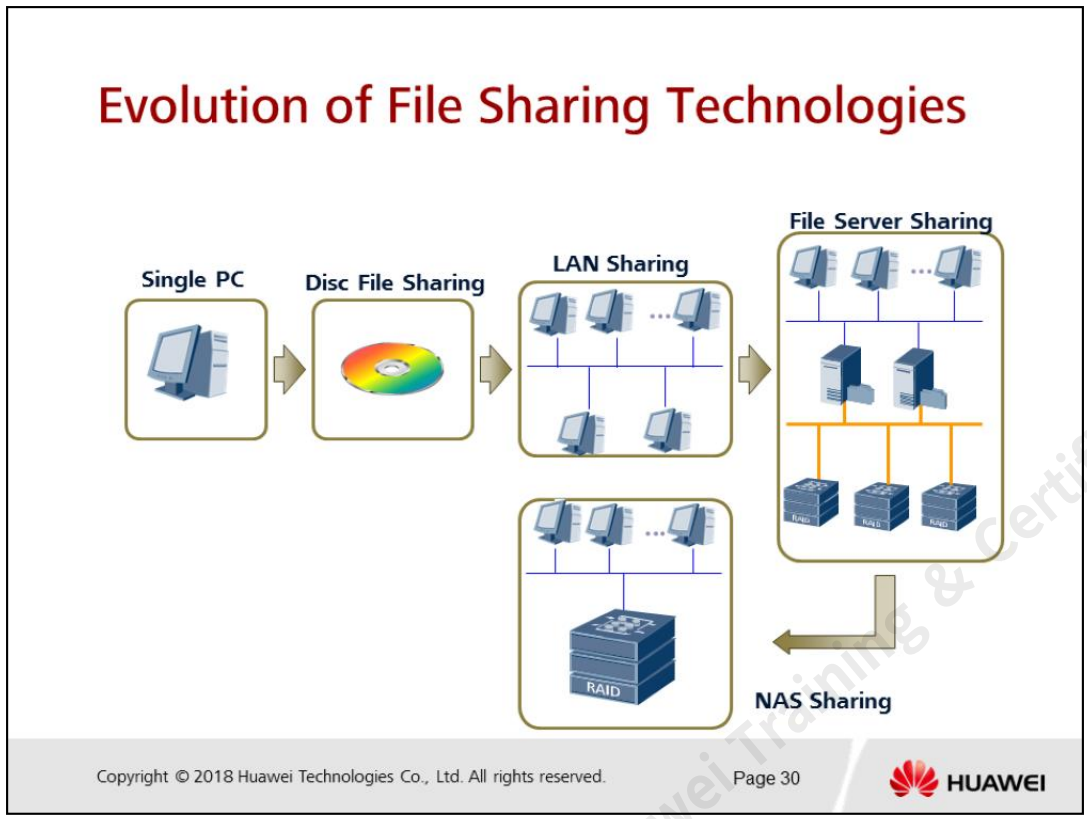
HUAWEI

Huawei Training & Certification
Huawei Training & Certification



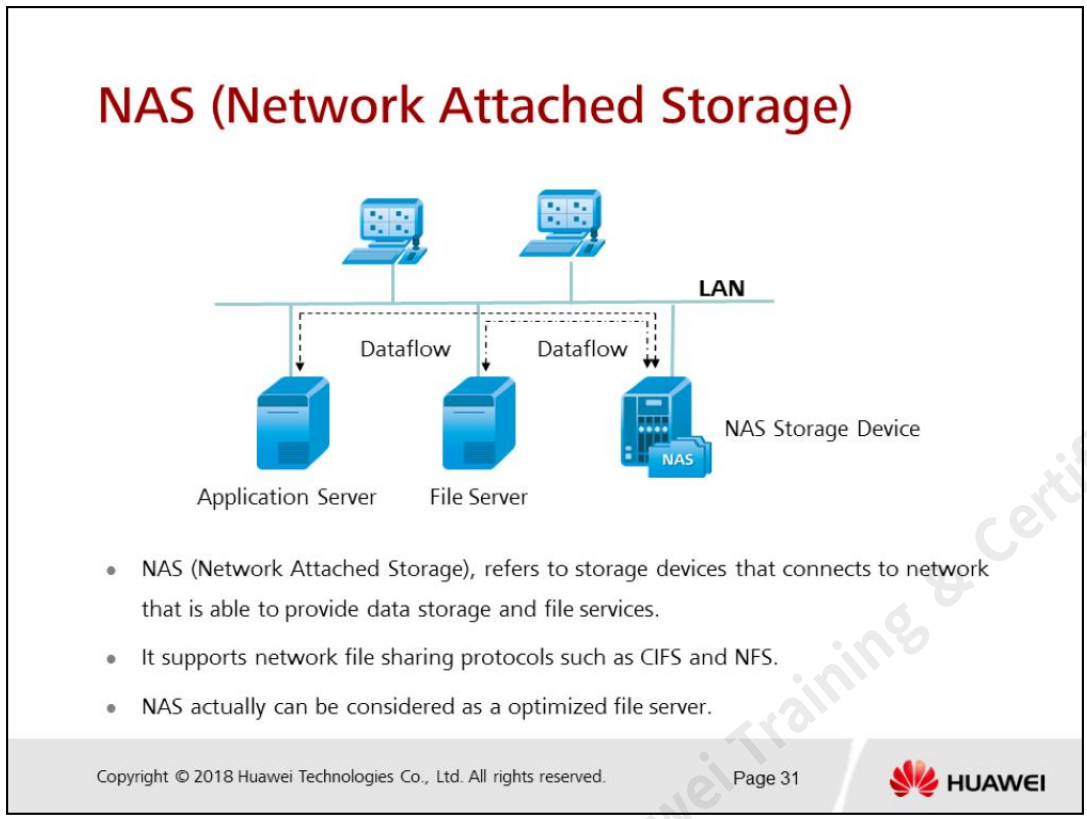
- File sharing involves data storage and access. In a file sharing environment, users who create files can specify file usage permissions (read, write, execute, append, delete, list) for other users and can control file changes. When files are shared, if more than one user accesses the same file at the same time, a method of protection is needed to maintain the integrity of the data. The method used in the traditional Client/Server (C/S) model is by implementing file sharing protocols and using Distributed File System (DFS). The following are some examples of these implementations:
    - File Transfer Protocol (FTP) allows the transmission of data in networks. It is a standard file transmission protocol, and uses TCP/IP protocol in the server end and client end of the data transmission process. FTP data are not encapsulated during transmission, which means that there is no security protection on the data transmission. FTP over Secure Shell (SSH) added the security protection features on top of the FTP protocol that allows data to be transmitted over the network securely without data security concerns compared to FTP.

Huawei Training & Certification



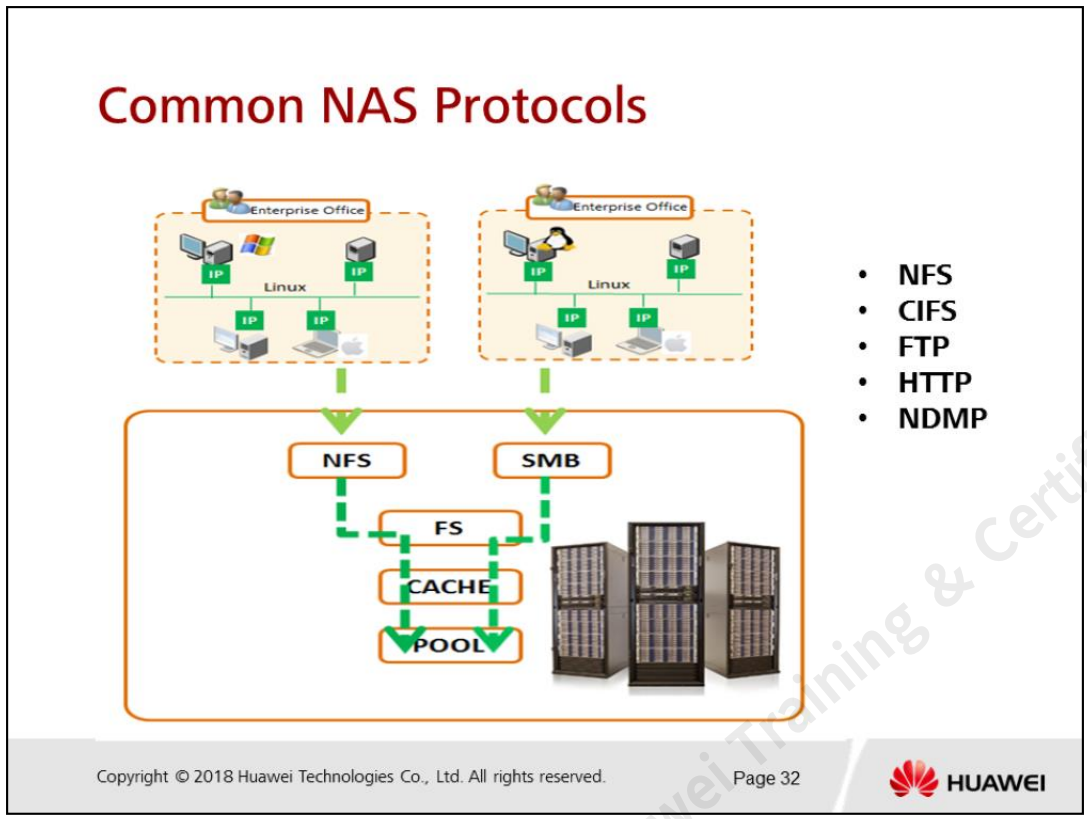
- Back in the past, Kilobyte(GB) level file sharing floppy disk was widely used. However, as time progresses, the need for larger share volumes across organization increases, which lead to the inventions of removable storage media such as flash drive that is able to store data on Gigabyte(GB) level, which in turn fully replaced floppy disks over time.
- Enterprises needs to not only store large amounts of data, but also needs to share those data over networks, where NAS undoubtedly becomes a good choice. For servers and hosts, NAS is a external device that has great flexibility for over the network deployment. NAS also provides file based storage and not block based storage, which makes it easy for clients to access those data over the network.

Huawei Training & Certification



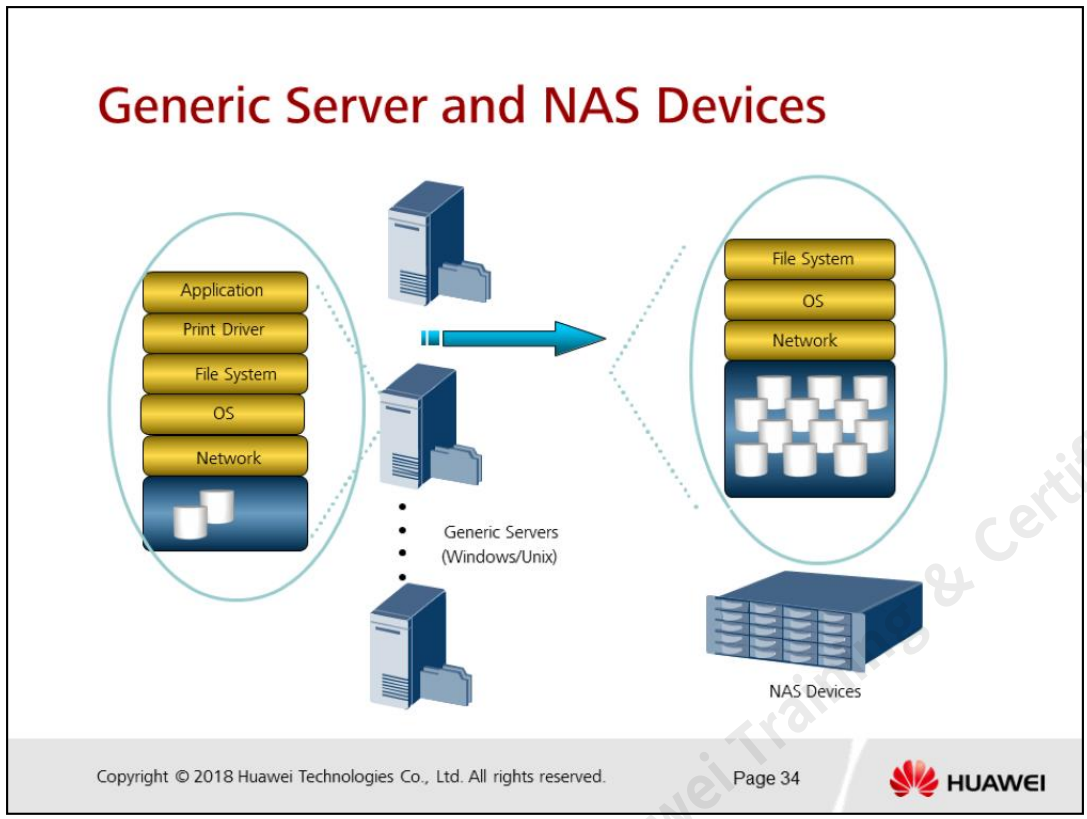
- NAS (Network Attached Storage) connects storage devices to an existing network to provide data and file services. NAS generally consists of several components such as storage hardware, operating system and a file system. It is based on the TCP / IP protocol to achieve file-level data access services.
- Instead of relying on a common operating system, NAS connects storage devices through a standard network topology, eliminating the need for a dedicated server to access the network. NAS uses a simplified user-oriented operating system designed for data storage with all the required protocols built in, therefore making the management and setup of the entire system relatively simple.
- NAS is mainly used for highly efficient file sharing services, which is suitable for scenarios where large capacity file transmission or sharing over the network is required.

Huawei Training & Certification



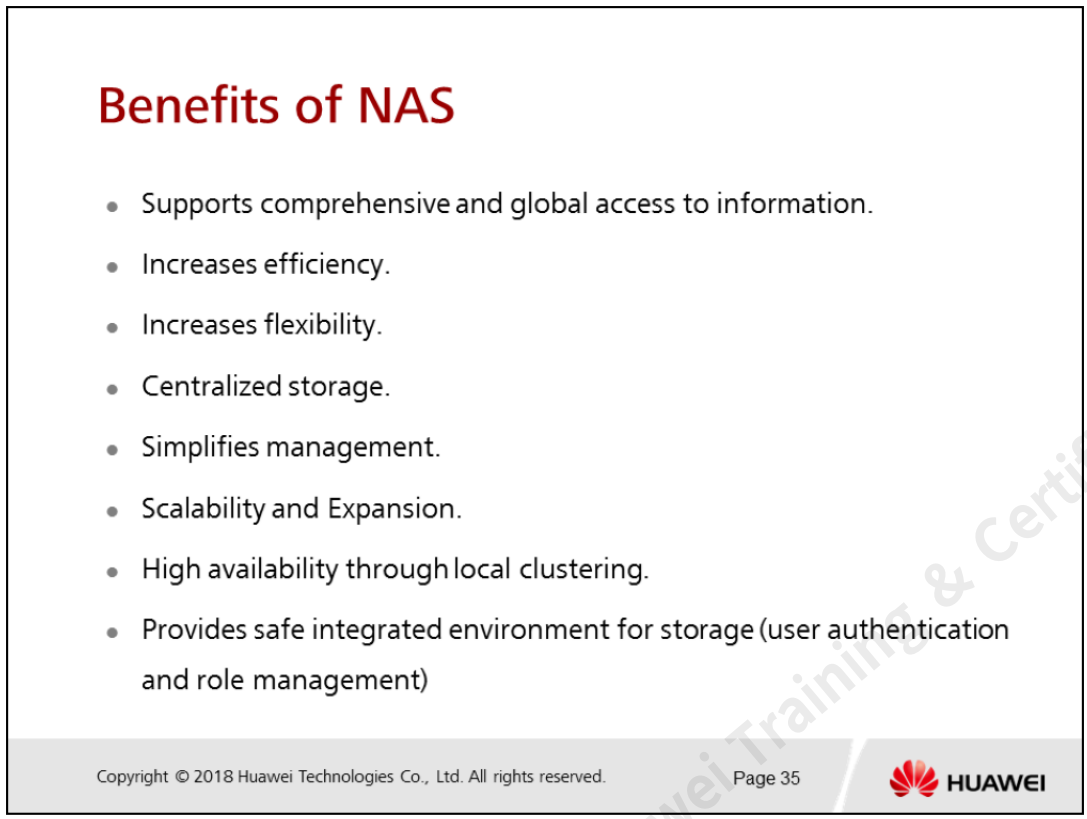
- CIFS (Common Internet File System), is a file sharing protocols that works under the Microsoft Windows environment, it is based on Server Message Block(SMB) protocol.
- NFS (Network File System) protocol is the traditional file sharing protocol under the UNIX operating system environment.
- Most NAS devices support multiple file service protocols to handle remote file system I/O requests. As mentioned earlier, NFS and CIFS are common file-sharing protocols. NFS is primarily used in UNIX-based operating environments; while CIFS is used in Microsoft Windows-based operating environments. These file sharing protocols allow users to share file data across different operating environments and provide users with transparent migration of files across operating systems.
- NFS is a client/server application that uses Remote Procedure Calls (RPCs) to communicate between computers. Users updates the files on the remote NAS as if they are updating files on their local storage.
- The user's system needs an NFS client to connect to the NFS server. Since, NFS servers and clients uses TCP/IP protocol to transmit files, the clients and server system must install the TCP/IP protocol stack before NFS can work properly.

Huawei Training & Certification



- Generic server refers to all the different types of servers such as web server, application server, web server, mail server and virtual servers. In this context, the diagram above shows that NAS devices are compatible and able to provide storage and file services universally to any type of servers as long as they support the same file system.
- NAS devices are optimized on the basis of common servers, hence they have the features such as file service, storage, search and retrieval, and access to client application files.
- As shown on the diagram, a generic server can be used to install any types of application, and it runs on a common operating system(OS). NAS differs from generic servers, it is solely dedicated for file services, and it provides file sharing to other operating system through standard open protocols. In order to increase the high availability performance of NAS devices, some NAS vendors even provides NAS clustering feature.
- NAS devices consist of the following components:
  - NAS controller (CPU and Memory Etc.)
  - One or more network interface cards (NIC) which provides network connectivity. For example, Ethernet and Gigabit Ethernet network cards.
  - An optimized operating system for NAS feature management.
  - File sharing protocols such as NFS and CIFS.
  - Disk resources that supports industry standard storage protocols such as ATA, SCSI, FC.
- NAS environment includes the clients that access NAS devices through IP network using standard protocols.

Huawei Training & Certification



- NAS provides the following benefits:
  - Supports comprehensive global access to information: NAS provides high efficiency file sharing, and supports many to one or one to many configuration. Many to one configuration allows the NAS to be accessed by multiple clients simultaneously, while one to many configuration allows a single client to connect to multiple NAS devices.
  - Increases efficiency: NAS devices provides dedicated optimized OS for file services, which can lower file service operation through file servers, increases the utilization ratio of generic servers, and eliminates the bottleneck faced by generic servers when accessing files.
  - Increases flexibility: NAS uses industry standard protocols, and has great compatibility with Windows and Unix clients. It can flexibly provide file services to different types of clients that accesses the same resource.
  - Centralized Storage: NAS allows storages to be centralized and minimize the data duplication on client workstations, simplifies data management and provides higher data protection.

Huawei Training & Certification

Summary

- This module mainly introduces:
  - Data Center Storage Network.
  - Network Architecture of DAS, SAN and NFS.
  - Advantages and Disadvantages of DAS, SAN and NFS Network Technologies.

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

Huawei Training & Certification
Huawei Training & Certification

Quiz

1. What is the data transmission rate of FC protocol? ( )
   A. 4Gb
   B. 8Gb
   C. 10Gb
   D. 16Gb
2. What are the common protocols in SAN? ( )
   A. FC
   B. iSCSI
   C. FCoE
   D. IB

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

- Answers:
  - ABD.
  - ABC.

Huawei Training & Certification

Thank You

www.huawei.com

Huawei Training & Certification
Huawei Training & Certification

07 Storage Reliability Technologies

www.huawei.com

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

Huawei Training & Certification
Huawei Training & Certification

Foreword

- This module mainly describes the traditional RAID technologies, RAID 2.0+ technologies, Multipathing technologies and disk reliability technologies.
- This module also describes the functions of these technologies towards the aspect of data protection.

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

Huawei Training & Certification
Huawei Training & Certification

Objectives

- Upon completion of this module, you will be able to:
  - Describe the concepts, principles and types of RAID technologies.
  - Describe the principles of RAID 2.0+ technologies.
  - Understand the Multipathing technologies of hosts.
  - Understand the technologies implemented for ensuring the reliability of hard disk.

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

Huawei Training & Certification
Huawei Training & Certification

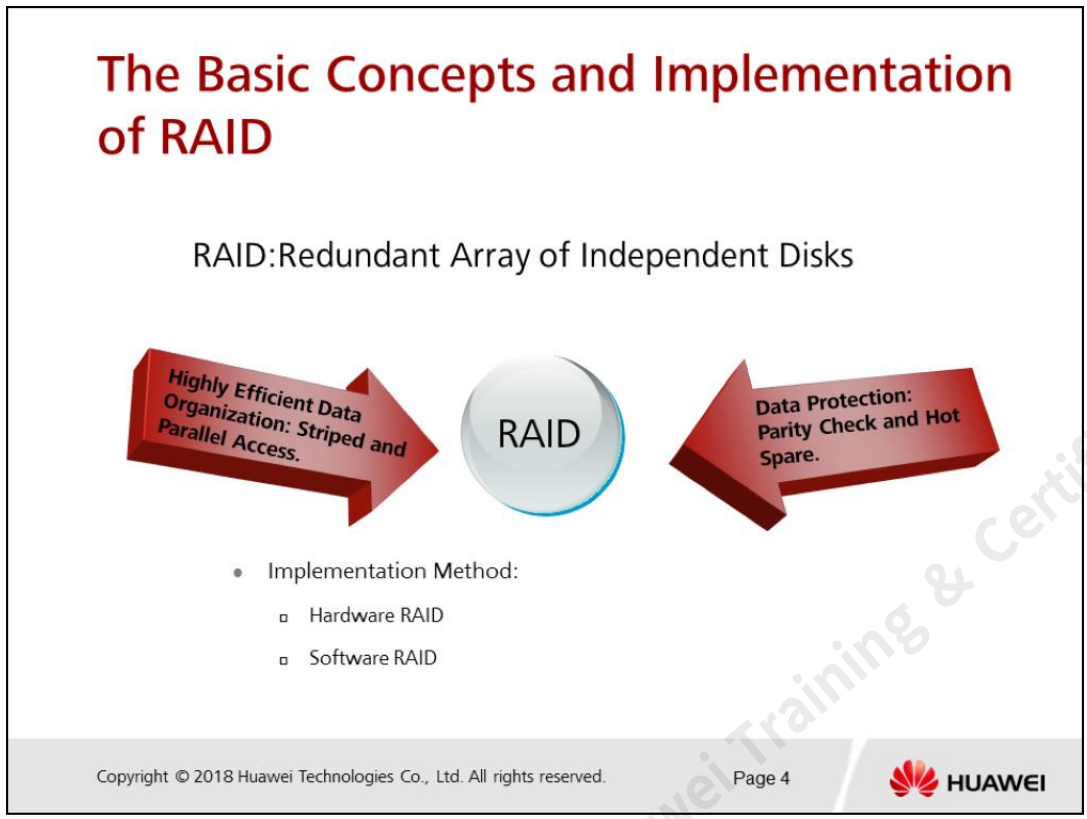
Contents

1. **Traditional RAID Technologies.**
2. RAID 2.0+ Technologies.
3. Host Multipathing Technologies.
4. Disk Reliability Technologies.

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

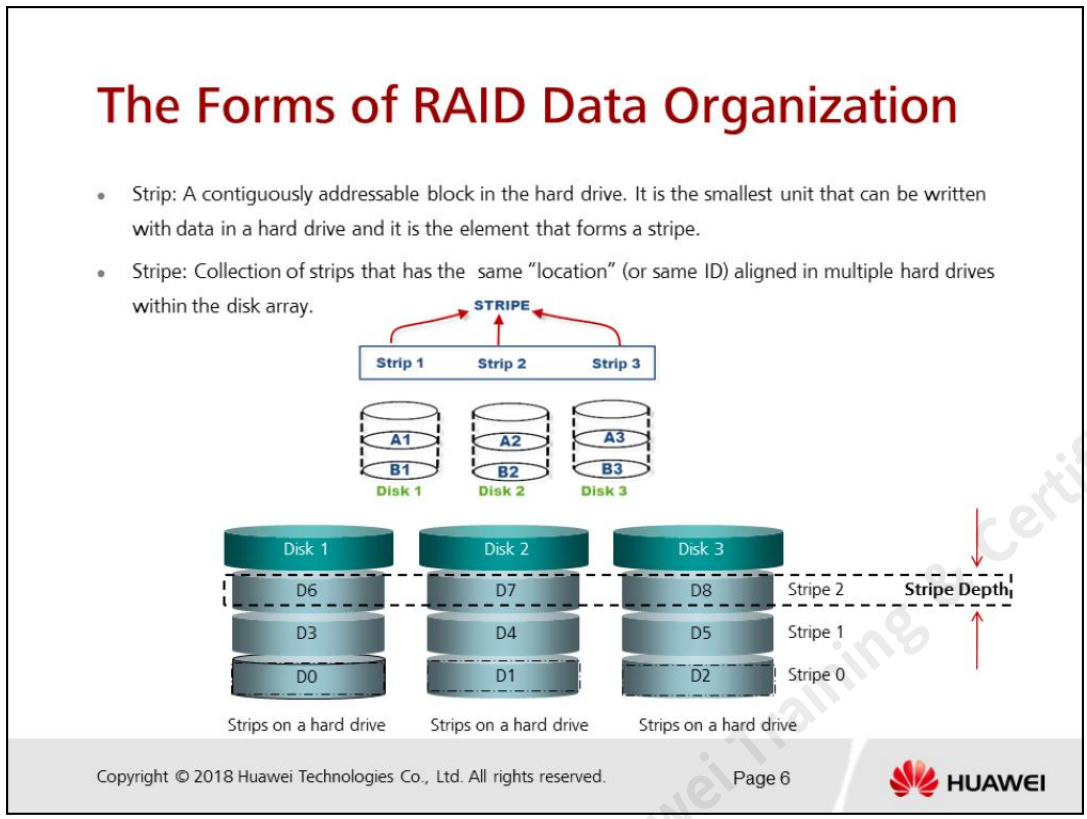
HUAWEI

Huawei Training & Certification
Huawei Training & Certification



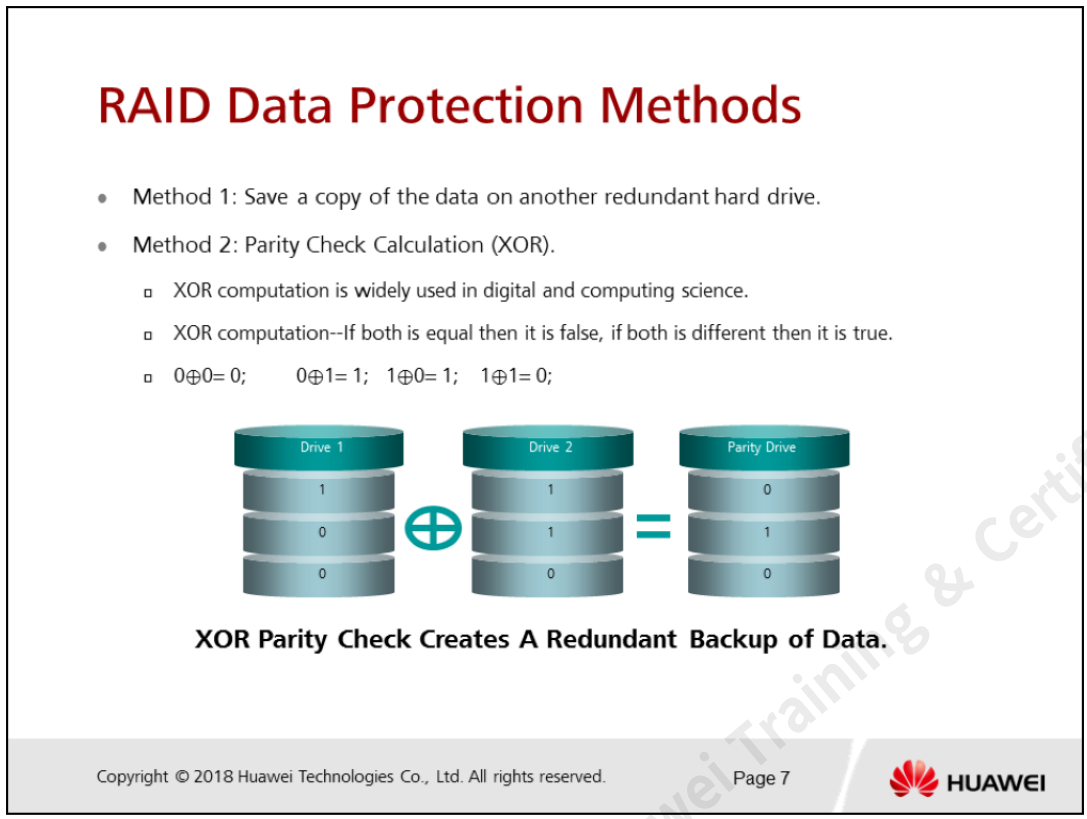
- The initial purpose of RAID technology was to combine multiple hard disks with small capacities in order to obtain a larger storage space. The RAID technologies that we will be mentioning in this module is more related towards data protection. In other words, when the physical devices fails, RAID is able to prevent the loss of data.
- The main functions of RAID technologies:
  - By striping the data on the hard disk to achieve block access of data, it reduces the mechanical seek time of the drive and improves the data access speed.
  - By reading the data simultaneously from few drives in the array (Parallel Access), it reduces the total mechanical seek time of the drives and improves the data access speed.
  - By mirroring or parity check methods, it achieves redundant data protection in the drives.
- With the development of array technologies, it produced a lot of different types of RAID. However, only a few types of RAID are still in use actively. In this module, we will be discussing the most common types of RAID and learn the other related functions of RAID such as data protection. At the same time, we will learn on why choosing different types of RAID means that there will be a difference in the performance or cost.

Huawei Training & Certification



- Stripe Width:
  - Stripe width is the count of the number of member disk in a stripe. In simple words, stripe width refers to the number of parallel stripes that can be written to or read from simultaneously. This is equal to the number of disks in the array, so a four-disk striped array would have a stripe width of four.
- Stripe Depth:
  - Stripe depth refers to the storage capacity of a stripe in each individual disk in the array. For an example, the stripe depth a stripe size of 128KB across two hard disk in RAID 0 configuration will be 128KB/2=64KB.
  - In other forms of RAID that has parity involved, please note that you need to exclude the parity disk and only include the data disks before calculating the stripe depth. For example, a RAID 5 configuration with 3 hard disks with a stripe size of 128KB, will be 128KB/2=64KB. It is divided by 2 because the sum of storage capacity equivalent to 1 disk is used for parity (For RAID 5 with 3 disks configuration every 2 blocks of Data, there is 1 block of Parity Data, D0+D1+P0).

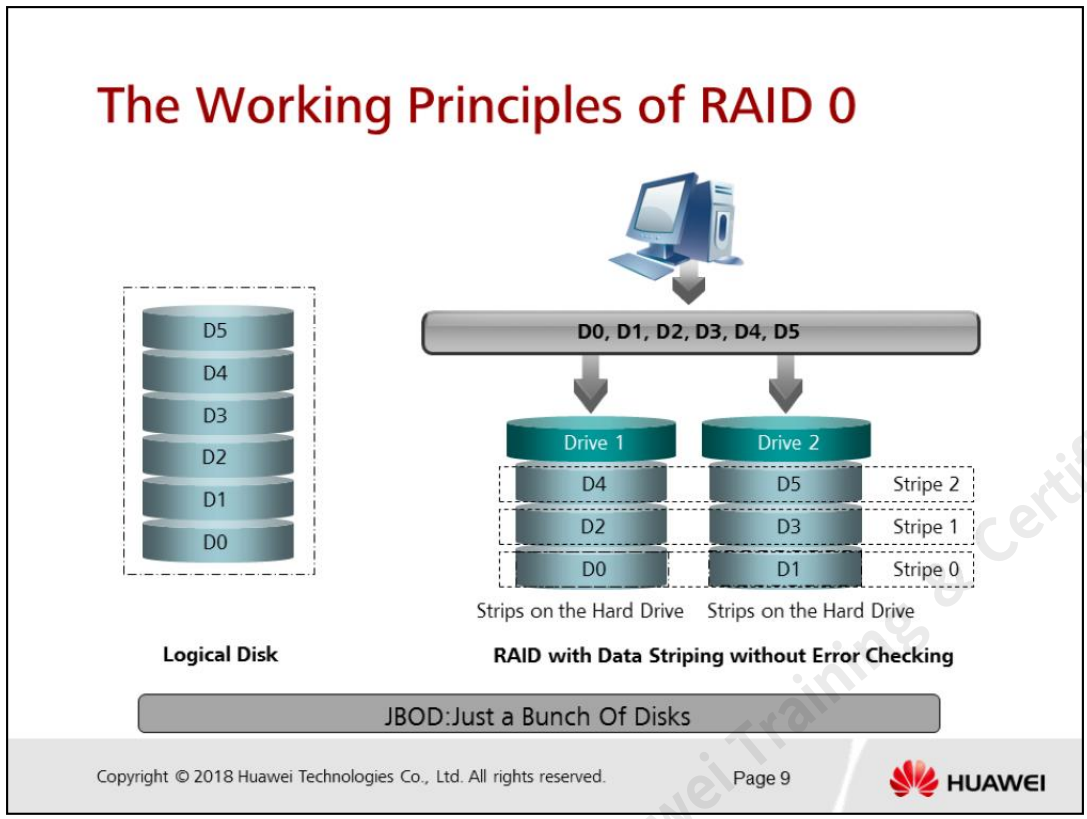
Huawei Trainings & Certification
Huawei Training & Certification



- RAID technologies usually have 2 types of methods for data protection. One of the method is to store a backup of the data on another redundant hard drive(mirroring). The second method is by using XOR computation to calculate the parity of the data. The XOR computation will use the user data to calculate the additional data which is the parity that can be used to restore data in the event if one of the drives in the array fails. For RAID types that uses XOR computation for parity check, it means that there is a need for additional parity drives.
- A very simple Boolean operation is used at the binary level to create RAID parity. This operation is the Exclusive Disjunction operation also known as Exclusive OR (XOR). By using XOR, the raw binary data (1 or 0) is passed through an operation that results in a final binary result, which can be used for redundancy and error correction.
- The table for the operation of XOR on raw binary data is shown on the diagram above.

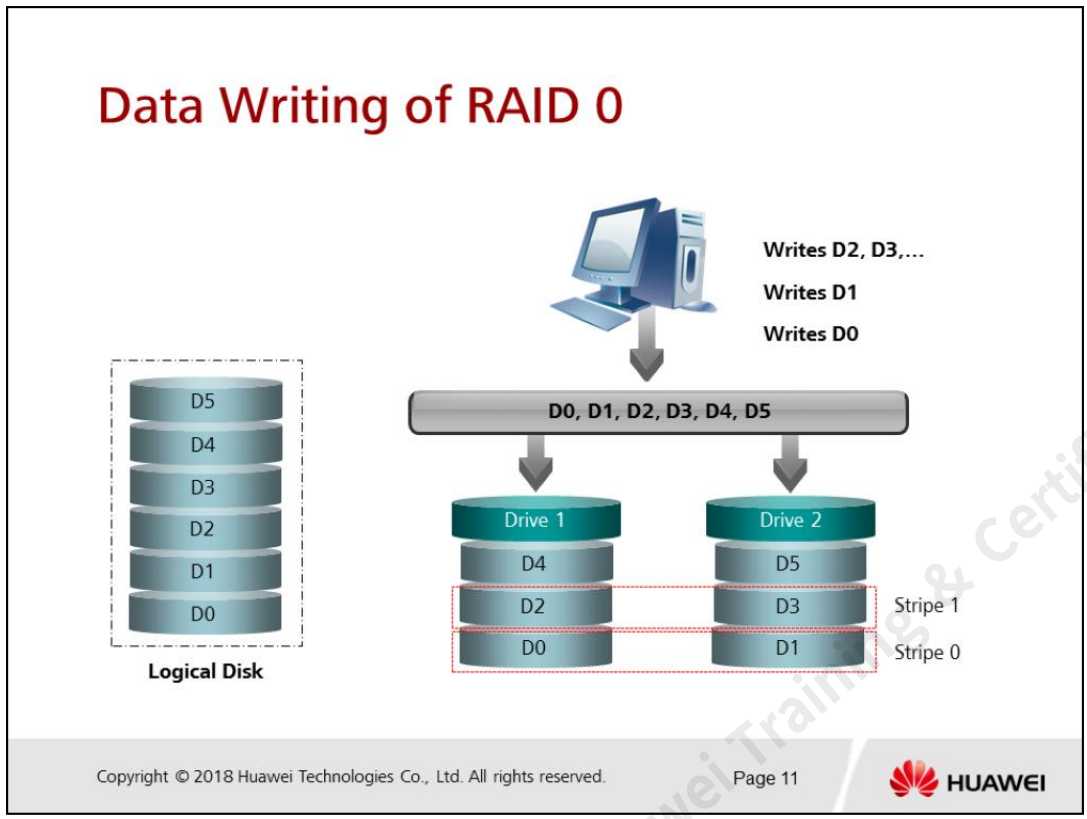
Huawei Training & Certification

Huawei Training & Certification
Huawei Training & Certification
Huawei Training & Certification



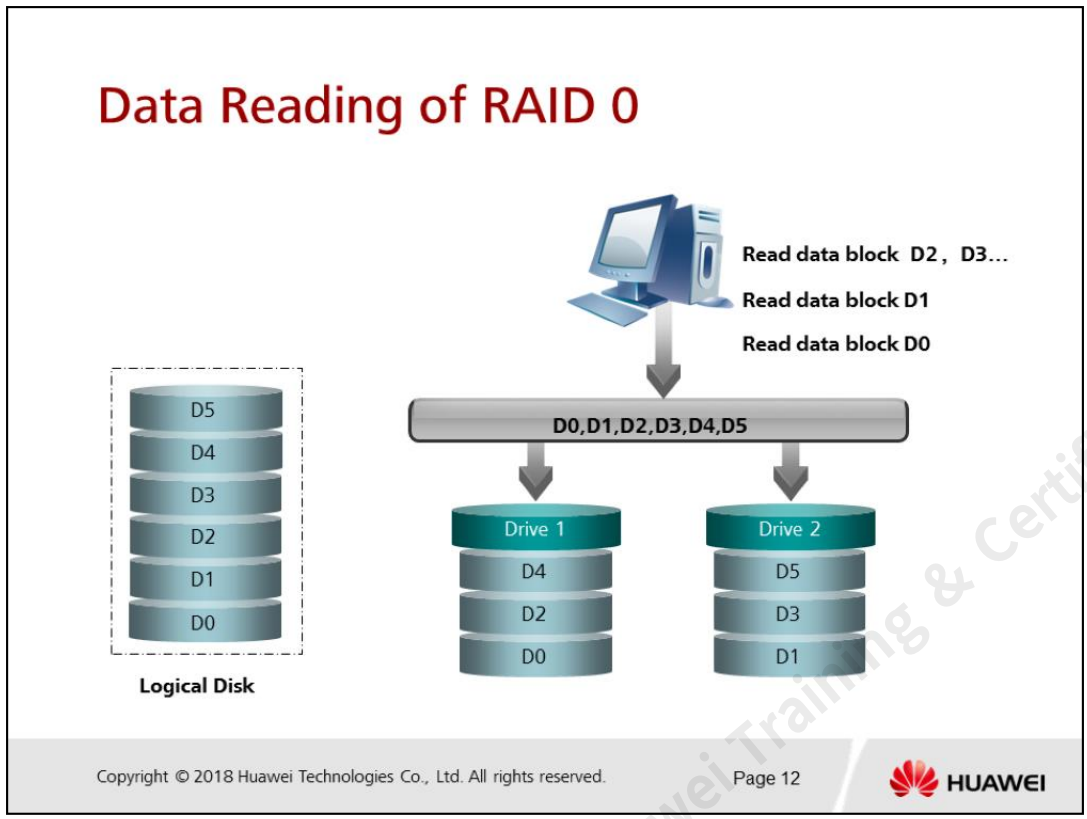
- Among all the RAID types, RAID 0 (also known as Striped RAID) has the highest storage performance. RAID 0 uses striping technologies to distribute the data across all the disks in the RAID group within the array.
- A RAID 0 implementation has a minimum of 2 disks. RAID 0 groups strips data of into data blocks of varying sizes ranging from 512 bytes to megabytes (usually it is the multiple of 512 bytes) and writes them into different drives. As shown on the two disk that forms RAID 0 at the diagram above: The first 2 data blocks are written into stripe 0, in which, the first data block is written in the stripe 0 of disk 1, and the second data block is written in parallel at the stripe 0 of disk 2. The next data block is written at stripe 1 of disk 1 and so on. Using this method, the workload of I/O is distributed and balanced across all the disks. Due to the higher speed of the data transmission bus compared to the disk read/write speeds, it can be considered that the drives within the RAID group can read/write simultaneously in parallel.

Huawei Training & Certification



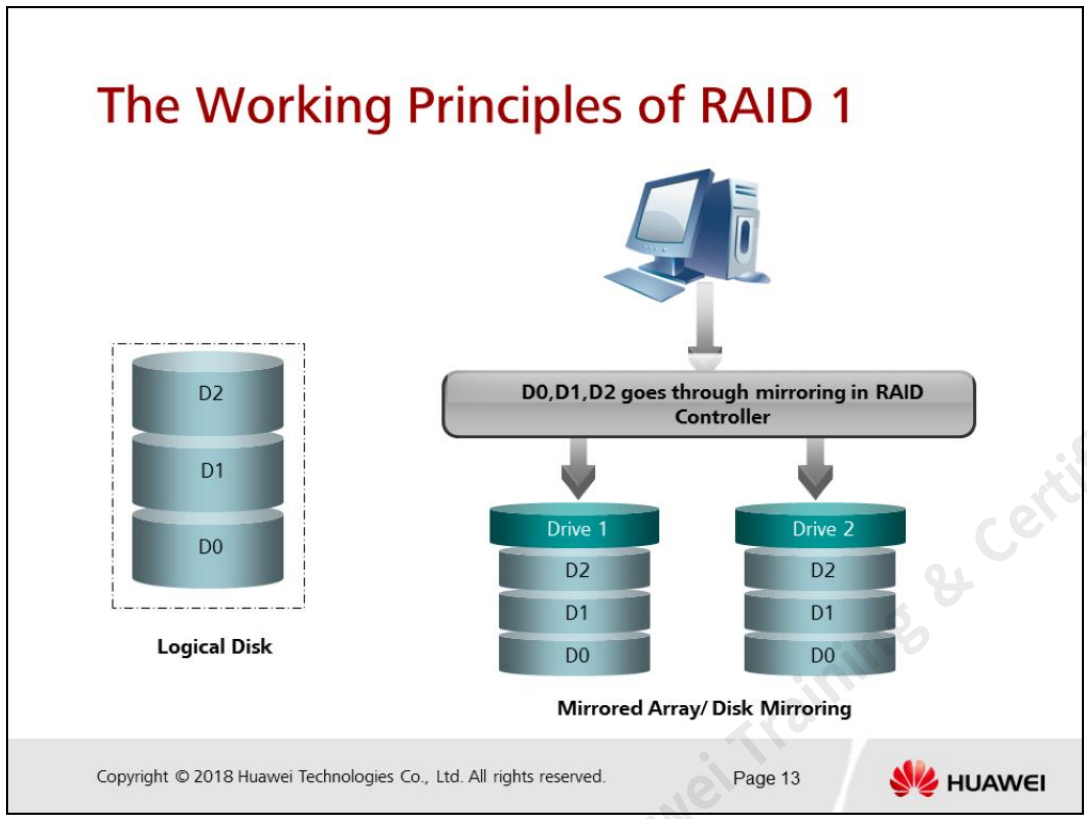
- RAID 0 uses data striping technologies to write data into the disk groups, it divides the data into data blocks, and evenly distribute them across all the hard disks within the RAID group. Only when the previous stripe is fully written by data, then the data is written to another stripe. As shown in the diagram above, data blocks D0, D1, D2, D3, D4, D5 is waiting to be written into a RAID 0 disk group. D0 is written into the first stripe (Stripe 0) into the first drive (Drive 1), and D2 is written to the first stripe in drive 2 and so on. All the remaining data is written to the disk group using the data striping method in a RAID 0 disk group.
- The data write performance of RAID 0 is directly proportional to the number of disks within the disk group. In simpler words, the more number of disk are in the RAID 0 disk group, the higher the data write performance.

Huawei Training & Certification



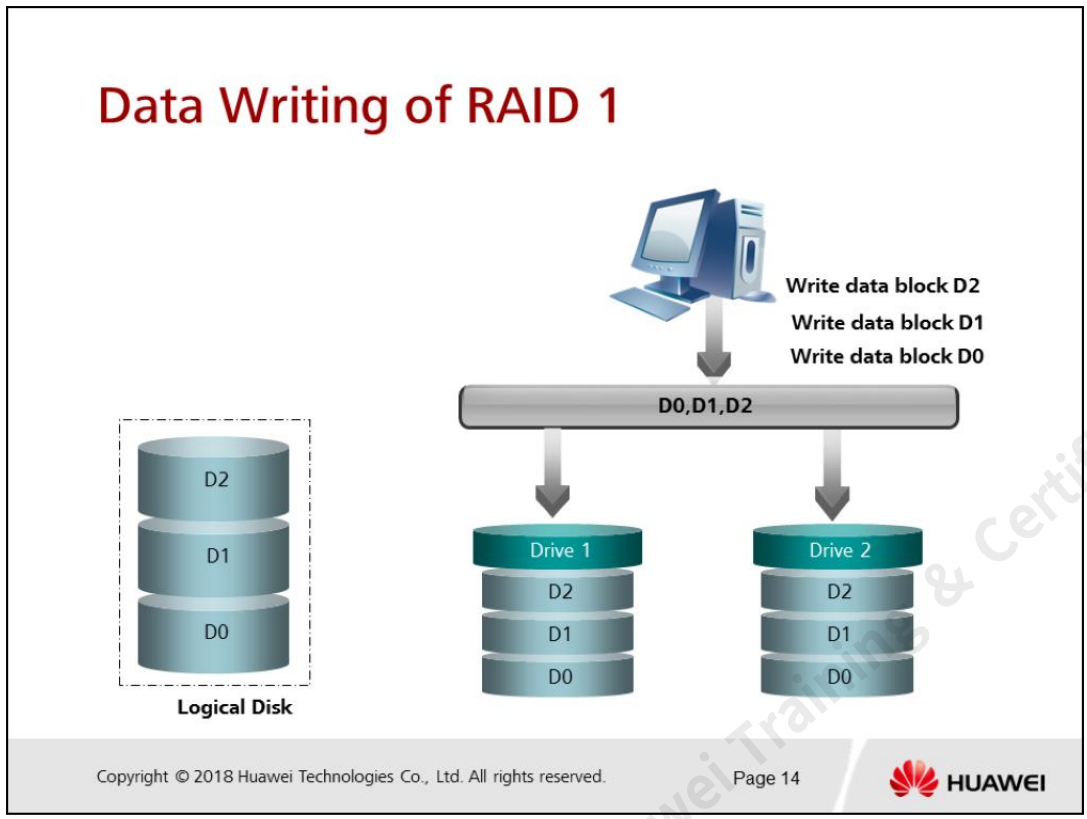
- When RAID 0 receives a data read request, it will search and retrieve the targeted data from all the disks within the RAID group. In the diagram above, we can see the whole data read process.
- Firstly, the array receives the read request for data blocks D0, D1, D2, D3, D4, D5. Subsequently, the array reads the data in parallel, which is reading D0 from drive 1 and D1 from drive 2, and so on for all the other data. When all the data blocks are read from RAID, they are integrated by the RAID controllers to form a complete data and then sent to the host. Since, RAID 0 is using data striping technologies where data is split across drives, the data blocks must be recombined to form the complete unstriped data before it is sent to the host.
- Similar to data writing, RAID 0 data read performance is directly proportional to the number of disks within the RAID group.

Huawei Training & Certification Huawei Training & Certification



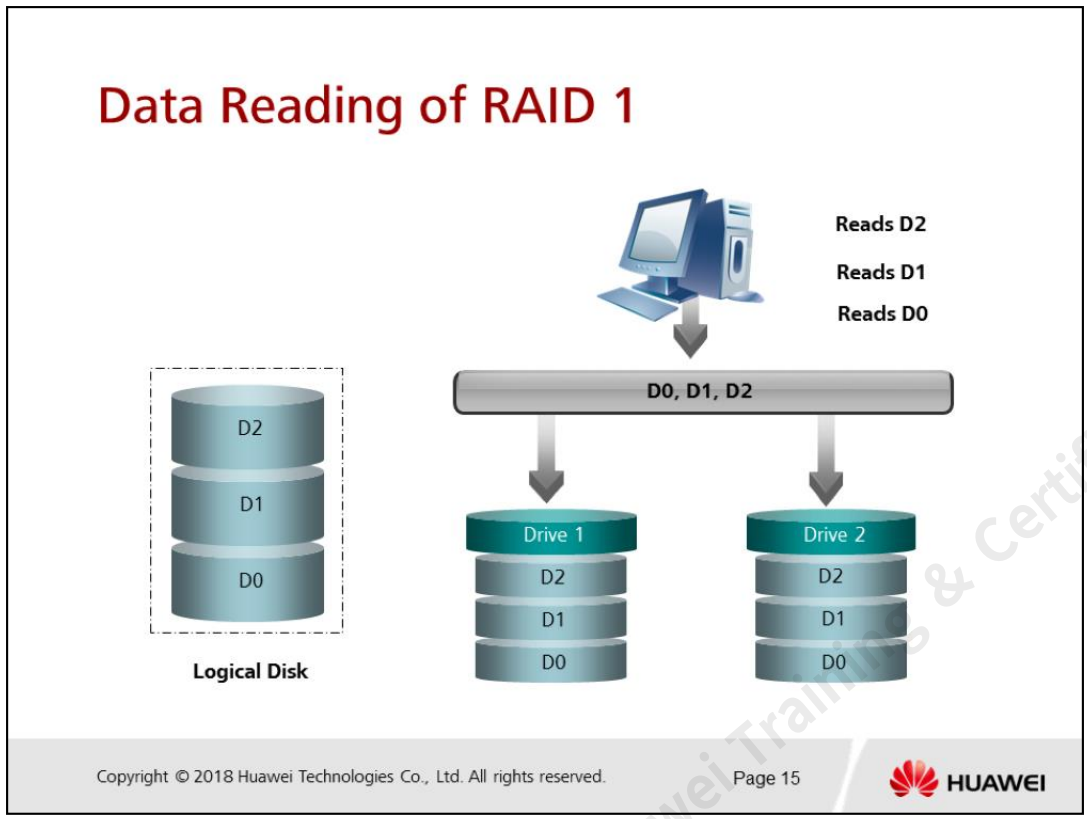
- RAID 1 (also known as mirrored disk array) is a type of RAID that is built on top of the principles of high safety. RAID 1 system uses 2 identical hard drives that is configured with mirroring. When data is written to one of the drives, the data copy will be stored simultaneously on the mirrored drive. When the source drive (physical drive) is faulty, the mirrored drive will takeover services to ensure service continuity. The mirrored drive serves as backup, and provides high data reliability.
- A RAID 1 disk group data storage capacity is only equal to the single disk capacity, the other disk stores the copies of the data, which means that for every 1 Gigabyte of data, it consumes 2 Gigabytes of storage space. Hence, the RAID 1 system that consist of 2 drives only has the space utilization rate of 50%.
- The two disks used for RAID 1 need to be identical in storage size. If the storage capacity of the drives are different, the usable storage capacity will be based on the smaller drive.

Huawei Training & Certification



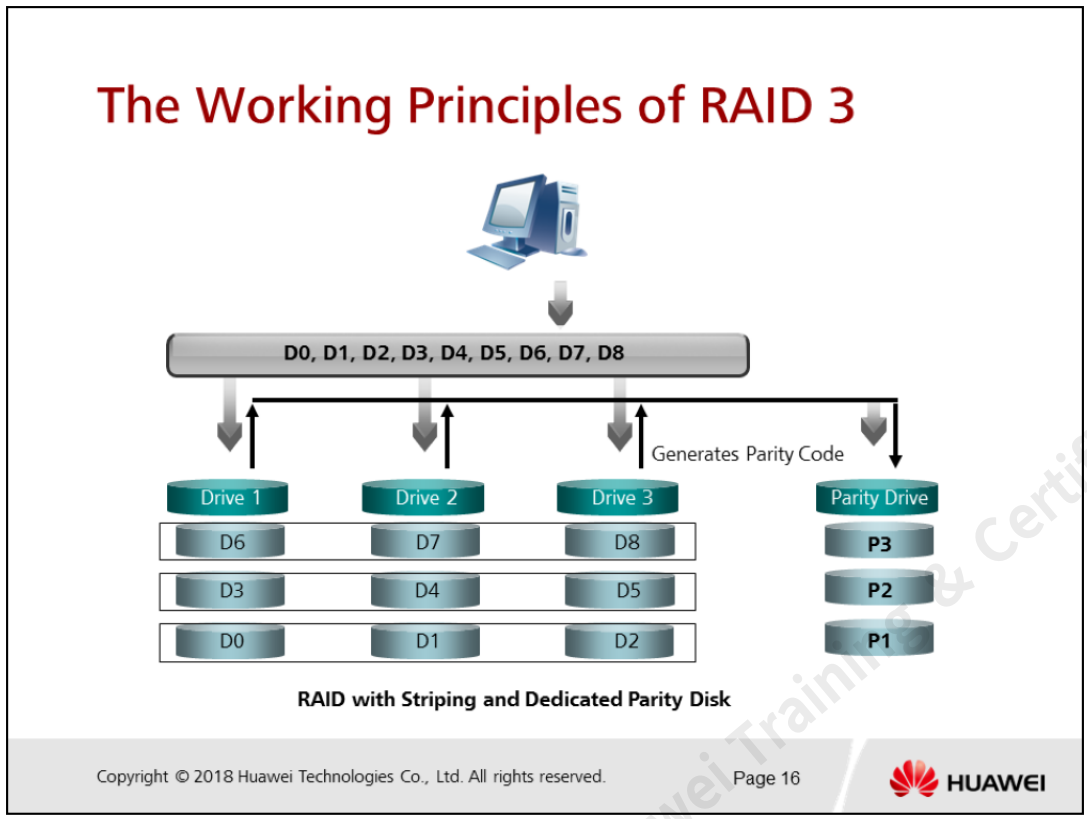
- RAID 0 uses data striping technologies to write different data blocks in parallel into multiple drives, but in contrast, RAID 1 writes the same data blocks to each disks within the RAID array. Data on all the member disks are the same. As shown on diagram above, data blocks D0, D1, D2 are waiting to be written to the drives. D0 and copy of D0 is written to both of the hard drives (Drive 1 and Drive 2) at the same time, and the other data uses the same method (mirroring) to be written to drives within the RAID 1 disk groups.
- Generally, a RAID 1 write performance is equal to the write performance of a single disk.

Huawei Training & Certification



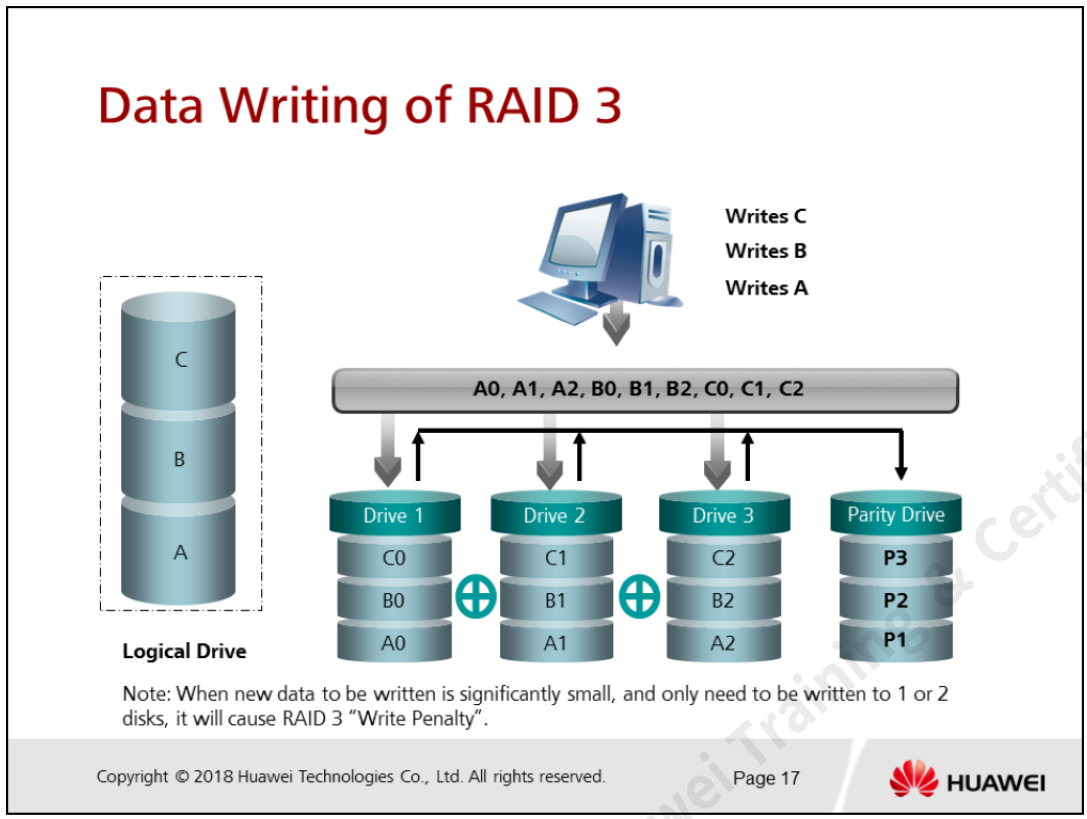
- When RAID 1 reads data, it will read from the data disk and mirrored disk simultaneously, which increases the reading performance. If one of the drives fails, the data can be read from the other drive.
- RAID 1 read performance is equal to the sum of the performance of both disk. But the write performance is only 50% since each data need to be written twice.

Huawei Training & Certification



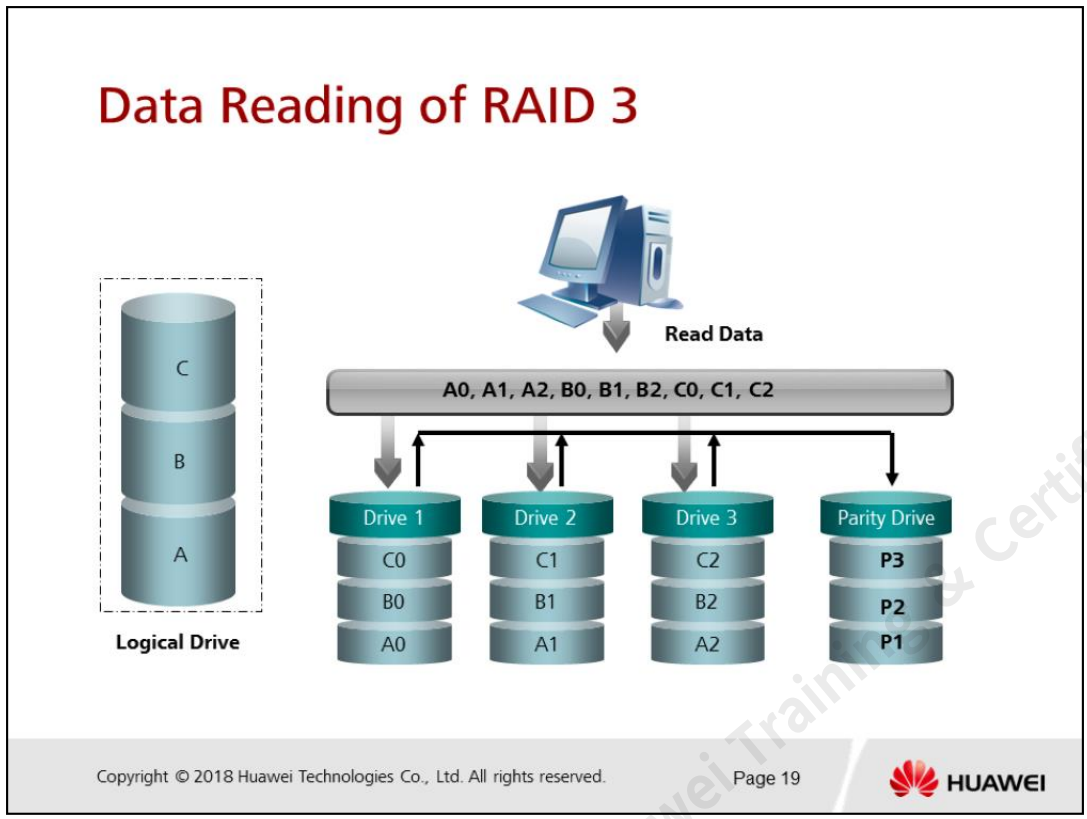
- RAID 3 is similar to RAID 0, but the difference lies with the fact that RAID 3 has a dedicated parity disk that stores parity information of the data. The dedicated parity disk stores the parity information of the data stripe. If any data error is detected or the drive is faulty, we can use the parity information to restore the data on the faulty drive. RAID 3 is suitable for data with high density or single user environments, that require long sequential read or write access to data such as uncompressed video editing or streaming media. RAID 3 assigns the data write operations to its data disk members of the RAID group. However, when there is new data to be written, no matter if it is written to any of the data disks, RAID 3 needs to recalculate and rewrite the parity information. Hence, when some applications requires large amount of data writing, the parity disk in RAID 3 will have high workload. Due to the fact that there is a wait time for the parity check, the RAID 3 read and write performance is affected to a certain level. Additionally, due to the high workload of the parity disk, it is commonly the disk that is most likely to fail within a RAID 3 system. This is why the parity disk is often considered as the bottleneck of RAID 3 systems.
- Assuming the number of disks in a RAID 3 system as N, the usable storage capacity will be N-1 (deducting the storage capacity of the 1 drive used as parity disk). Similar to the other RAID types, the member disks within the RAID group need to have the same storage capacity and rotational speed for optimum performance.

Huawei Training & Certification



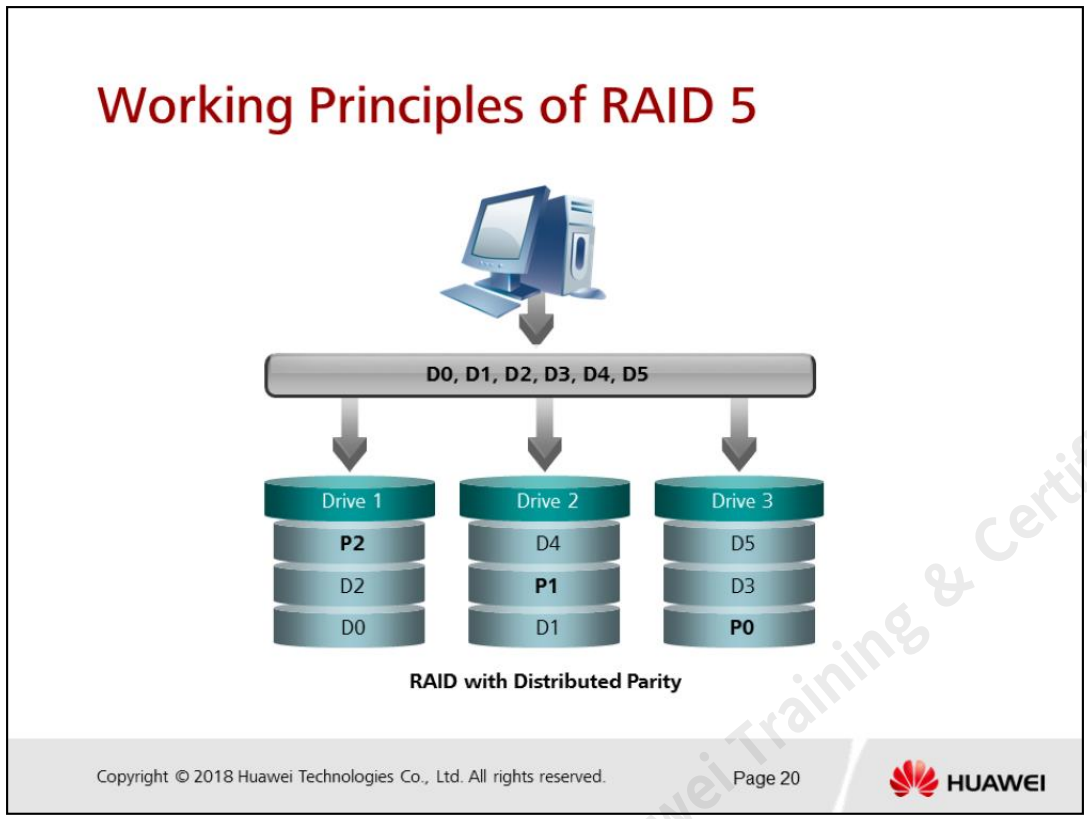
- RAID 3 implements single disk fault tolerance and parallel data transmission. In other words, RAID 3 uses data striping technologies to split up data into data blocks, runs XOR calculations to generate the parity information that is written into the dedicated parity drive. When a drive is faulty, the data is written to the drives that are not faulty and the parity check continues.
- RAID 3 performance is not fixed. In principle, RAID 3 uses N+1 data protection method. This means that when N number of drives containing user data that needs to be protected, you will need a extra 1 drive to store parity information. In this scenario, when a new data is written to the disks, parity information will be calculated and written to the parity disk.

Huawei Training & Certification



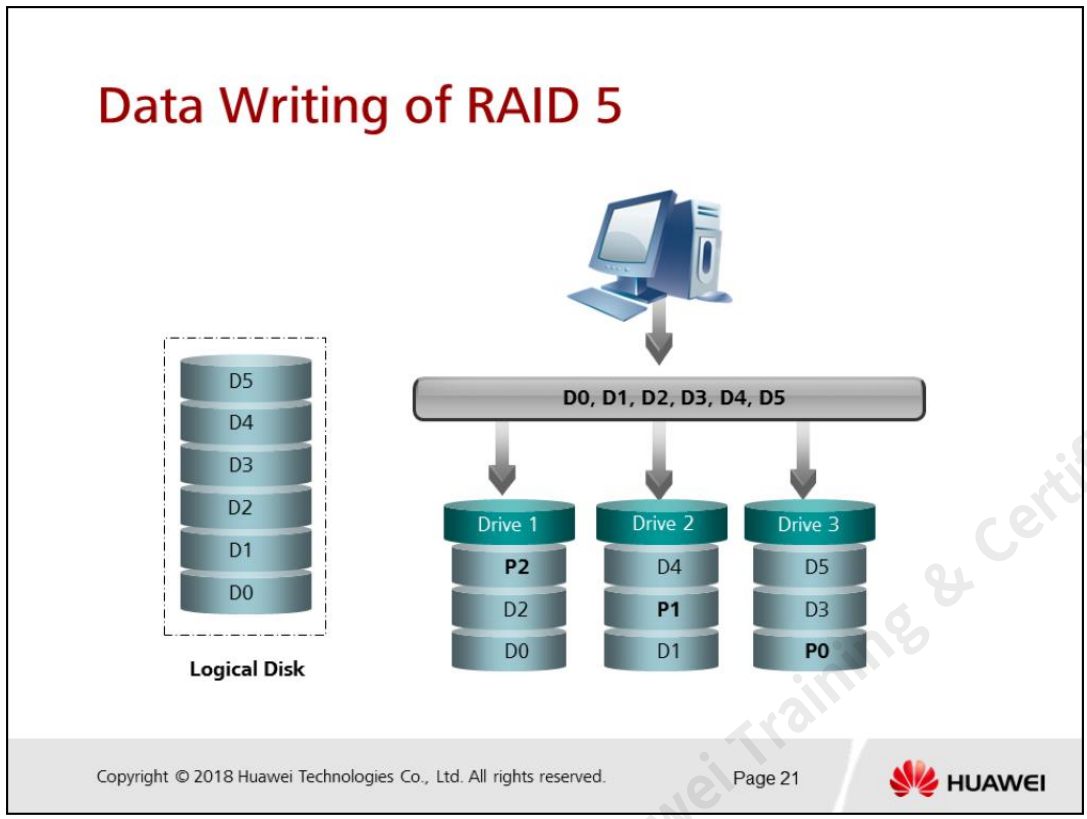
- In RAID 3, data are read in stripe method. Each drive in the RAID is controlled by the disk driver, so the data blocks in the same stripe can be read in parallel. Hence, each of the drives in the RAID 3 system is utilized during the data reading operation, which increases the red performance.
- RAID 3 uses parallel data read(write) mode.
- RAID 3 data read performance depends on the amount of data to be read and the number of disk drives within the RAID 3 array.

Huawei Training & Certification



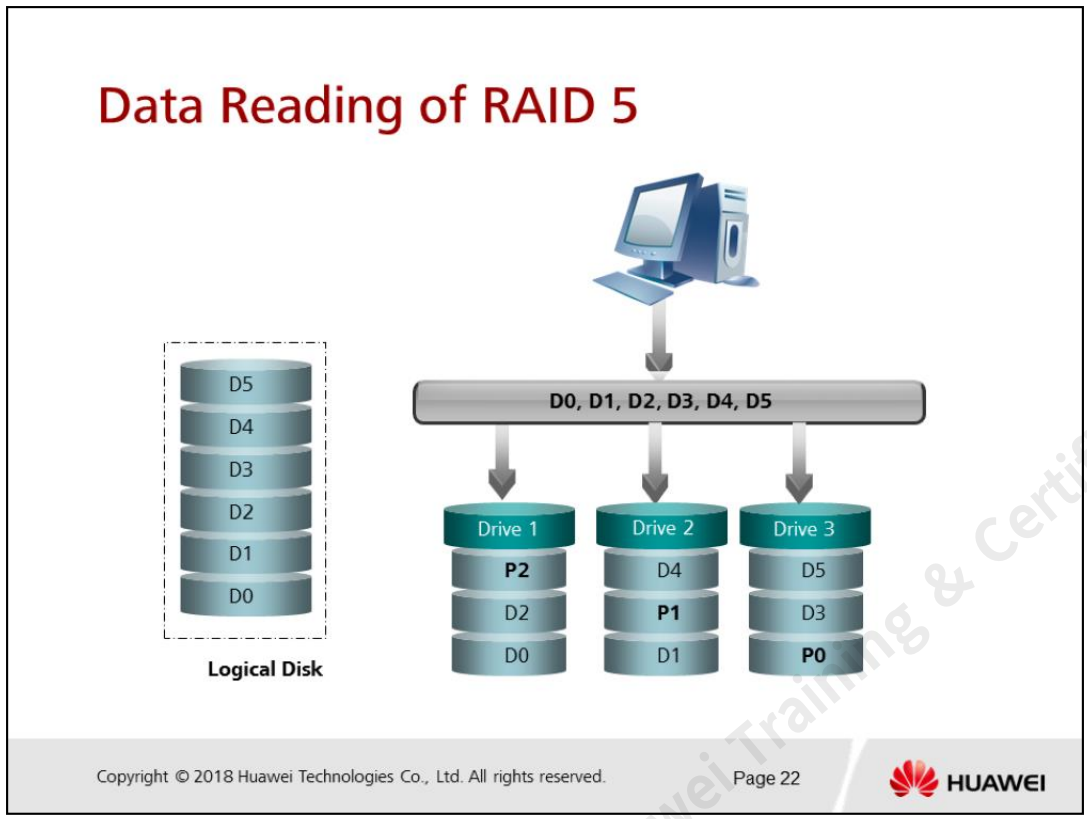
- RAID 5 is the improved version of RAID3, and it uses data striping, calculates and store parity information. RAID has a dedicated parity drive where all the parity information is read and written, this also causes the performance bottleneck of RAID 3 that we have mentioned previously. RAID 5 uses distributed parity, where each member disk in the RAID group will store data and parity information. So, RAID 5 do not have a bottleneck or a hot spot like RAID 3.
- Assuming the number of drives in a RAID 5 group is N, the usable data storage capacity is equal to N-1. Similar to the other RAID, all the member disks in the RAID 5 array need to have the same disk capacity and rotational speed.
- In the disk array of RAID 3 and RAID 5 types, if a disk fails, the RAID group will change from Online state to Failed state, until the data on the faulty disk is fully reconstructed. If another disks in the RAID failed at the same time, then all the data in the RAID array will be lost.

Huawei Training & Certification



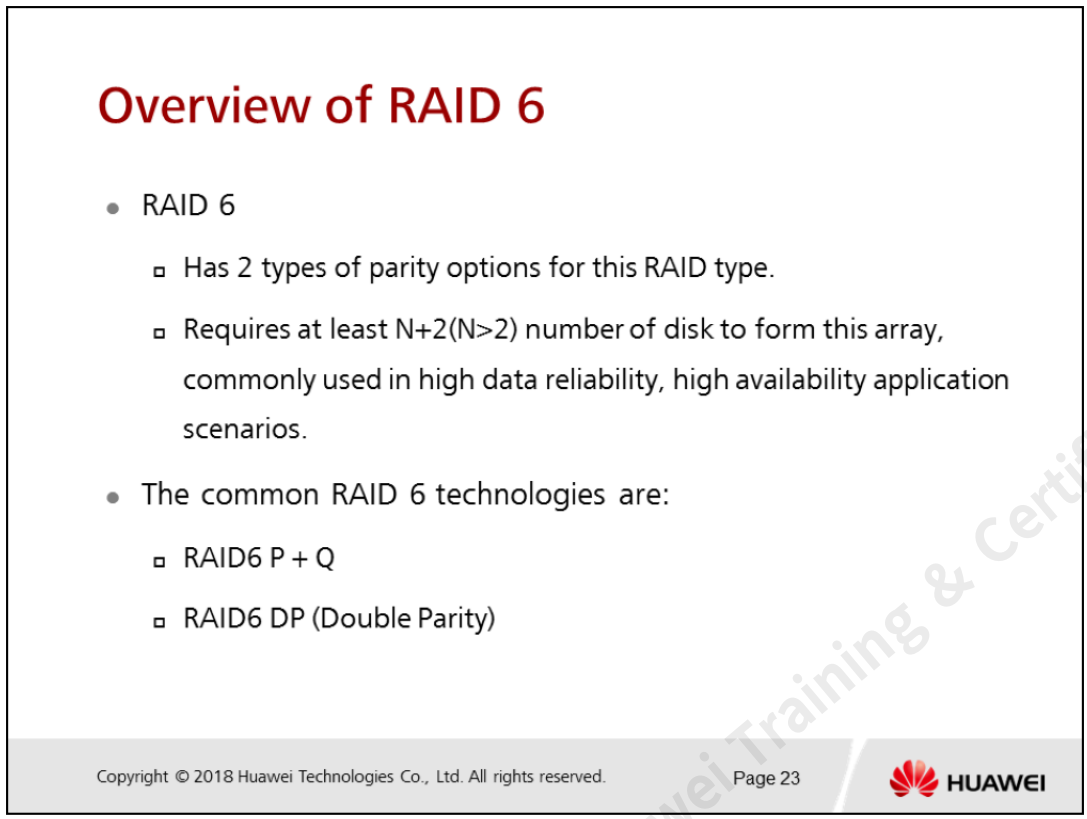
- In RAID 5, data is written in the form of stripes into the disk groups. Each member disk of the RAID 5 group stores data and parity information, when data is written in stripes, the parity information will be written in the corresponding disk. When RAID 5 is continuously writing data, different stripes uses different disks to store parity information. Hence, the parity information in RAID 5 is not saved in a fixed dedicated parity disk, but it is distributed and stored across the disks based on certain logics.
- When writing small amount of data, RAID 5 has write penalty similar to RAID 3.
- RAID 5 write performance depends on the amount of data written and the amount of disks in the RAID 5 array. Assuming the number of disks in a RAID 5 array is N, in the scenario where all the member disks has same rotational speed and without considering write penalty, the RAID 5 sequential IO write performance in theory will be slightly smaller than N-1 times the performance of a single drive (calculation of redundant check requires additional computing time).

Huawei Training & Certification



- Data within a RAID 5 array is stored in the form of stripes on the disks. Only N-1 number of data disks is required to fully restore all the data. It has the fault tolerance of a single drive, meaning that one of the hard disk can fail without losing data, but if more than 1 fails at the same time then the data can not be reconstructed.
- The data read performance of RAID 5 depends on the amount of data and the number of disks within the RAID group.

Huawei Training & Certification

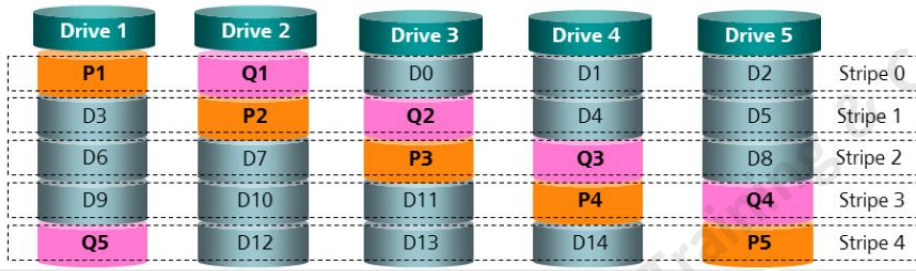


Huawei Training & Certification

- The RAID group that we have mentioned earlier in this module only considered the scenario where only a single disk has failed (excluding RAID 0). Nowadays, the storage capacity has greatly increased, and the disk reconstruction has also increased. A RAID 5 system that is formed with a lot of large capacity drives may need a few days and not few hours to reconstruct the data of the failed drives. During reconstruction, the system is in the downgraded state, which means that in this situation, any additional drive failures will cause the array to fail and the data is lost. This is why some organization or enterprises requires a dual redundancy system. In other words: A RAID group should allow 2 drives to fail and at the same time all the data is still accessible. This type of dual redundancy data protection methods usually can be implemented in the following ways:
  - The first method is multi mirroring. Multi mirroring refers to the data stored in the main disk is stored in multiple copies to additional redundant hard drives. This method involves a very large expenditure and cost.
  - The second method is RAID 6 array. RAID 6 array provides data protection for up to 2 disk failures at the same time.

Working Principles of RAID 6 P+Q

- RAID 6 P+Q calculates 2 different sets of parity data P and Q, when data is lost, the data can be recovered based on P and Q. The parity information of P and Q is calculated based on the formula below:
  - P = D0 ⊕ D1 ⊕ D2…
  - Q = (α ⊕ D0) ⊕ (β ⊕ D1) ⊕ (γ ⊕ D2)…

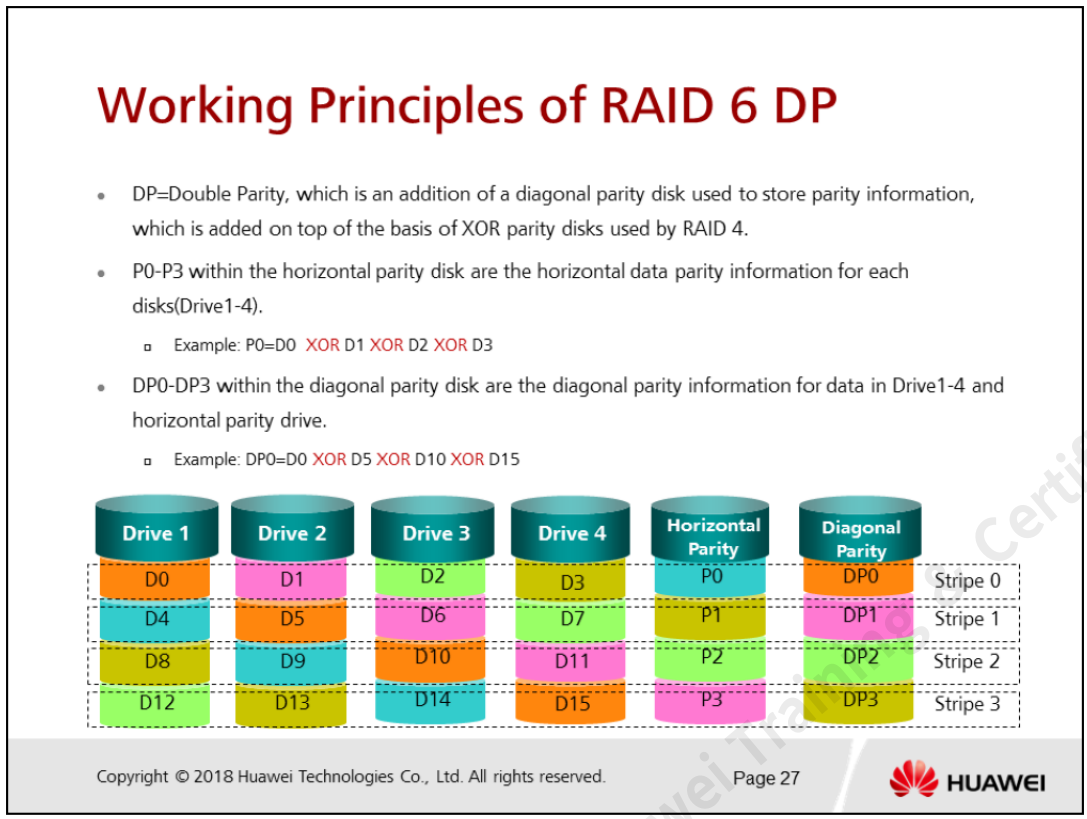


Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

- When RAID 6 implements P+Q parity, P and Q are 2 independent set of parity value. They use different calculations, the user data and the parity information of the same stripe is distributed across the disks.
- P is calculated using simple XOR operation. Q is calculated by GF (Galois Field) conversion then with XOR operation. α (alpha) , β (beta) and γ (gamma)are constants used in the GF formula, and the resulting value will produce a "Reed Solomon Code". This calculation involves the GF conversion and XOR operation of the data in the same stripes within the data disks. Reed Solomon codes is widely used for error correction due to its capability to convert the raw data distributed across multiple disks into parity information, and it is able to restore the lost data by running calculations based on available parity information.
- As shown on the diagram above, P1 is derived by running XOR operation on the stripe 0 containing D0, D1 and D2. Similarly, P2 and P3 are derived using the same method on stripe 1 and stripe 2 respectively.

Huawei Training & Certification



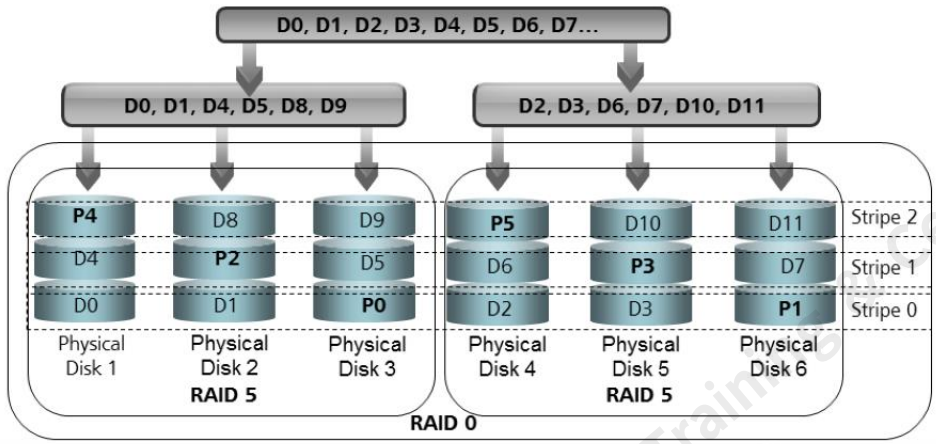
- Another type of RAID 6 implementation is RAID 6 DP(Double Parity). RAID 6 DP has two independent parity data blocks. The first parity information is the same as P parity of RAID 6 P+Q. The second parity information differs from RAID 6 P+Q, and uses diagonal XOR computation to calculate the parity information for diagonal data blocks across the disks in the array. Horizontal parity information is calculated using XOR computation on the data blocks in the same stripe. As shown on the diagram above, P0 is derived from XOR computation of the data blocks in stripe 0 which are D0, D1, D2 and D3. P1 is derived from XOR computation on data blocks D4, D5, D6, D7 on stripe 1 etc. Hence, we can conclude that P0 = D0 ⊕D1⊕ D2⊕D3， P1 = D4⊕D5⊕D6⊕D7 and so on for the other data on the remaining stripes in the disk array.
- The second parity data block is calculated using XOR computation on diagonal data blocks within the disk array. The process of choosing the data blocks for this computation is quite complex. DP0 is derived from XOR computation on data blocks: D0 on stripe 0 on drive 1, D5 on stripe 1 on drive 2, D10 on stripe 2 on drive 3 and D15 on stripe 3 on drive 4. Similarly, DP1 is derived from data blocks: D1 on stripe 0 on drive 2, D6 on stripe 1 on drive 3, D11 on stripe 2 on drive 4, and the horizontal parity data block P3 on stripe 3 on the horizontal parity disk. Hence, we can conclude that DP0 = D0⊕D5⊕D10⊕D15， DP1 = D1⊕D6⊕D11⊕P3, and so on for the other remaining stripes on the disk array.

Huawei Training & Certification

Huawei Training & Certification
Huawei Training & Certification
Huawei Training & Certification

Hybrid RAID - RAID 50

- RAID 50 is the RAID type that combines RAID 5 and RAID 0. The first level is RAID 5 and the second level is RAID 0.



Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

- RAID 50 is the combination of RAID 0 and RAID 5. Both subgroups are configured as RAID 5, and these 2 subgroups in turn forms a RAID 0. Each RAID 5 subgroup is independent from each other. RAID 50 requires at least 6 disks because a RAID 5 subgroup requires a minimum of 3 disks.
- As shown on the diagram above, physical disk 1, 2 and 3 forms a RAID 5 and physical disk 3, 4, and 5 forms another RAID 5 group. Both of theses RAID 5 subgroups forms another RAID 0.
- In RAID 50 array, it can tolerate multiple drive failures at the same time. However, in the event that 2 disks from the same RIAD 5 subgroup fails at the same time, the data on RAID 50 array will be lost.

Huawei Training & Certification

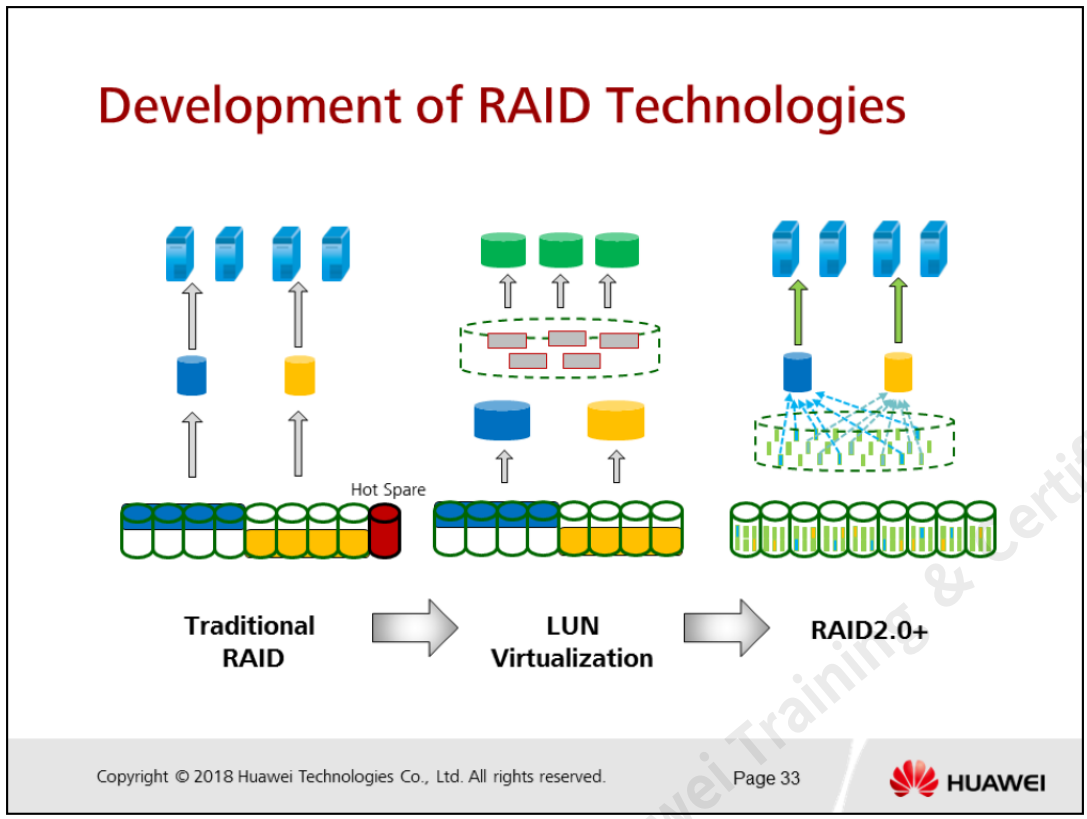
Contents

1. Traditional RAID Technologies.
2. **RAID 2.0+ Technologies.**
3. Host Multipathing Technologies.
4. Disk Reliability Technologies.

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

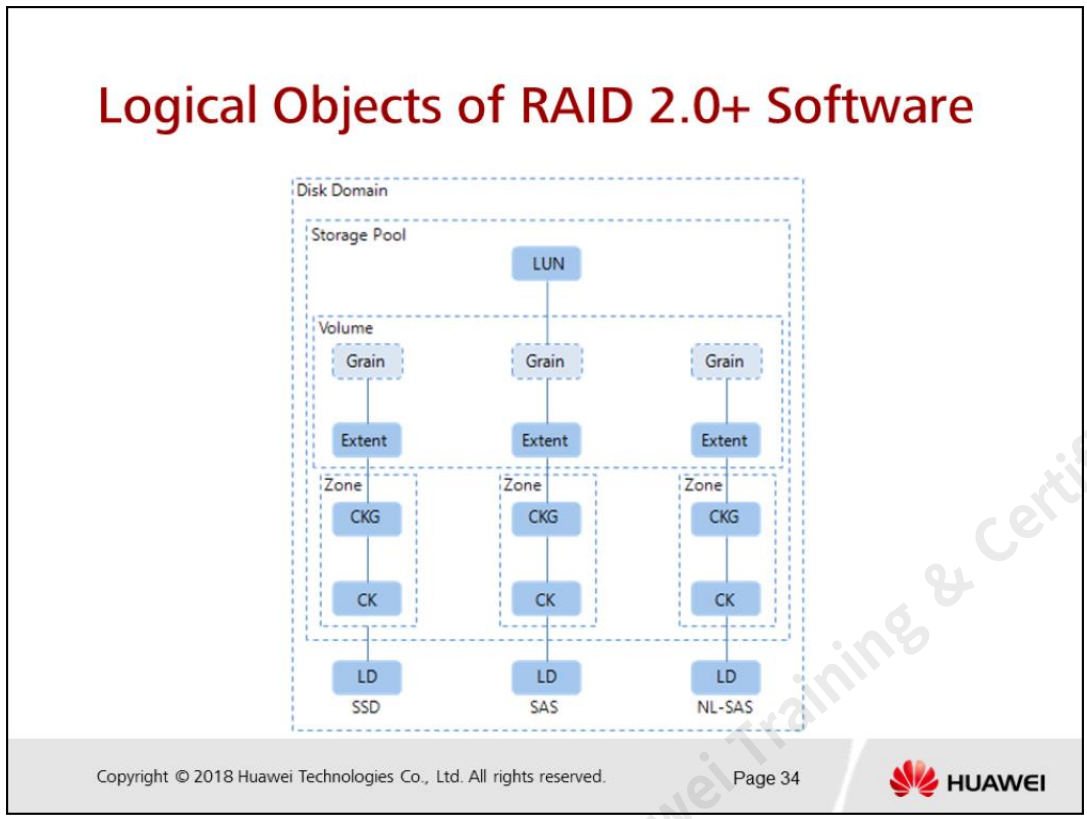
HUAWEI

Huawei Training & Certification
Huawei Training & Certification



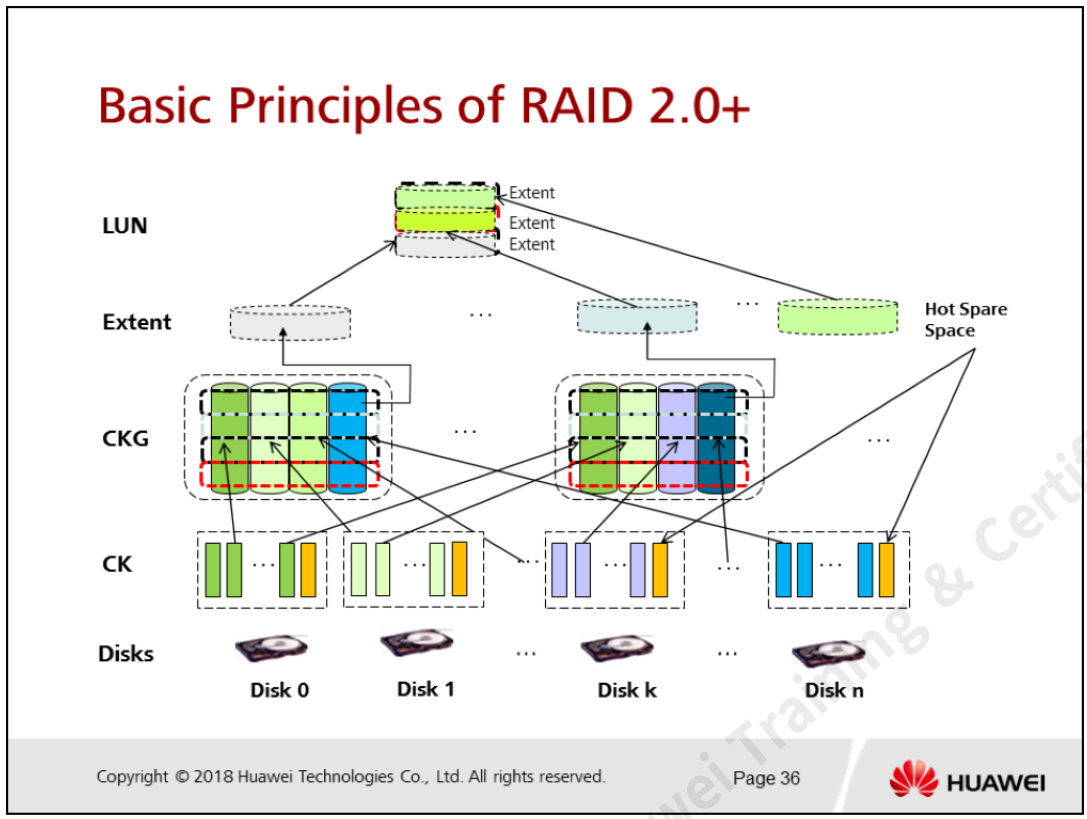
- In the earliest form of RAID technologies, it was just combination of couple smaller capacity inexpensive disks into a larger capacity logical disk for mainframe usage. Eventually, as the capacity of hard drives gets larger, the initial purpose of building RAID is longer towards creating a large capacity disks, but to utilize the RAID technologies to achieve data reliability and data security, and increase storage performance. As each individual disk capacity has become larger, and the storage capacity of digital disks in RAID are much larger, smaller logical unit number (LUN) are allocated from the RAID and used by servers. Following the development of disk technologies, the single disk capacity has reached few terabytes, and the time needed for disk reconstruction gets much longer which increases the risk of other disk failure during the disk reconstruction period that will cause data loss. In order to solve this issues, block virtualization technologies was introduced. RAID 2.0+ is a Huawei proprietary block virtualization technology, it distributes the physical and digital storage space into separate smaller blocks across multiple disks, and can utilize the data read and write performance of the storage system to its fullest extent. It is easier for expansion, allocation of storage space based on demand, distribution of hot and cold data, and data migration. It is the fundamental for all of the Huawei Smart storage software technologies. At the same time, since the hot spare space is also distributed across multiple disks, the data reconstruction can be done simultaneously across all disks, and it avoids the performance bottleneck of writing to a single hot spare disk, which in turn greatly reduces the data reconstruction time.

Huawei Trainee & Certification Huawei Training & Certification



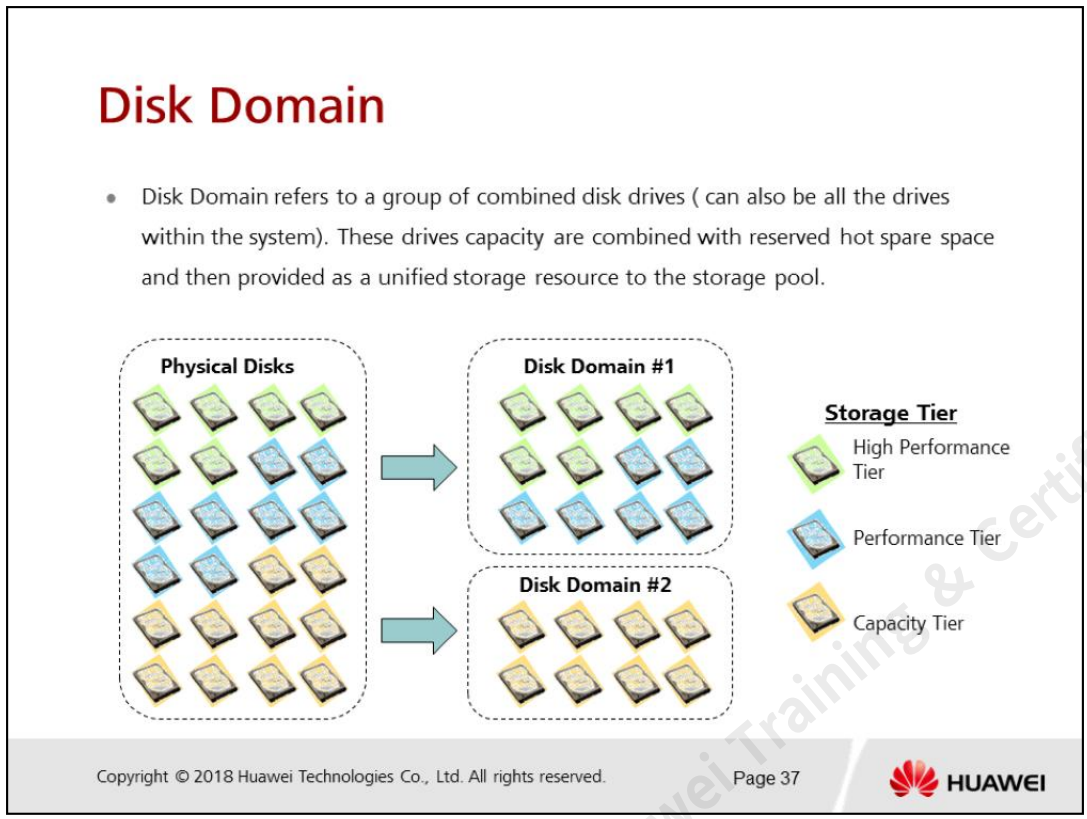
- Huawei RAID 2.0+ implements 2 layer virtualization management model that consist of low level disk management and high level resource management. Inside the system, the storage space of each disk is divided into smaller chunks, and these chunks are used as the fundamentals to build the RAID group, which allows data to be evenly distributed to all the disk in the storage pool. At the same time, chunks are used as the unit for resource management, which greatly increases the efficiency of resource management.
  - OceanStor storage system supports different types of disks (SSD, SAS, NL-SAS). In theory, SATA disks can also be used, but because of the low performance, it is less used in enterprise grade storage. These disks in the storage system forms a Disk Domain. In a disk domain, same type of disks can form a storage tier, and disks within the storage tier can be further allocated into different Disk Groups.
  - The disks in each storage tier is divided into Chunk(CK) of fixed size, in fact the CK size of SSD layer and SAS layer is 64MB and the CK size of NL-SAS layer is 256M by default.
  - OceanStor storage system uses random algorithm to form a Chunk Group (CKG) using the Chunk (CK) in each storage tier based on the user defined "RAID Policy". Users can set a RAID Policy for each of the storage tier within the Storage Pool.

Huawei Training & Certification



- In simpler terms, the overall principles of RAID 2.0+ can be represented as shown on the diagram above.

Huawei Training & Certification



- In traditional RAID, the first step before providing the storage space to the host is to create a RAID group. But there are limitations and requirement for this step: The disk type within the RAID group must be the same, the storage capacity and the rotational speed of the disks must be identical, and the maximum number of disks are recommended to not exceed the count of 12 disks.
- Huawei RAID 2.0+ technology implements another way to provide storage space to hosts. In this method, the first step is no longer creating a RAID group but to create a Disk Domain. A disk domain refers to a group of disk drives. 1 disk drive can only belong to 1 disk domain. OceanStor storage system can create one or more disk domains.
- At a glance, the concept of disk domain and RAID groups seems similar, which are groups of disk drives, but in reality, there are big differences between the two. In the creation of RAID groups, the set of drives used to form the RAID group have already been configured to a fixed RAID type, and the drive type, drive capacity and drive rotational speed must be the same. However, in a disk domain, the number of drives can be huge such as hundreds of drives, and can have a maximum of 3 different drive types, where each type of the drive can be assigned to a storage tier.

Huawei Training & Certification

Storage Pool & Tier

- Storage Pool is the container for storing storage resources. All the storage space used by servers are provided by the storage pool.
- Storage Tier is the integration of storage media with similar performance within the storage pool. It is used to manage the storage media with different performances and makes it easier to provide different storage space based on performance requirements.

Storage Tier	Tier Name	Supported Disk Type	Application
Tier0	High Performance Tier	SSD	Highest performance and cost, suitable for high frequency data access and data storage.
Tier1	Performance Tier	SAS	Medium performance and cost, suitable for medium frequency data access and data storage.
Tier2	Capacity Tier	NL-SAS	Low performance, low cost and high capacity. Suitable for large sized data and low frequency data access and data storage.

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

- A Storage Pool is created based on a disk domain, it can dynamically allocate the Chunk(CK) resources, form Chunk Group (CKG) based on the Raid Policy of each Storage Tier, to provide RAID protected storage resource to applications and users.
- A storage pool can be divided into different tiers, and the storage tiers and types of disk drives supported by the OceanStor storage system is shown on the table above.
- When you create a Storage Pool from a disk domain, you can specify the storage tier types and the "RAID Policy" and "Capacity" of each tier.
- The Capacity tier mostly consists of high capacity SATA or NL-SAS drives, and the recommended RAID policy is to use double parity RAID 6. (SATA drives are less commonly used nowadays, and part of the products no longer supports them in terms of specifications.)

Huawei Training & Certification

Storage RAID Type & RAID Policy

RAID Type	RAID Policy
RAID1	1D+1D，1D+1D+1D+1D
RAID10	System automatically chooses 2D +2D or 4D+4D
RAID3	2D+1P，4D+1P，8D+1P
RAID5	2D+1P，4D+1P，8D+1P
RAID50	(2D+1P)*2，(4D+1P)*2，(8D+1P)*2
RAID6	2D+2P，4D+2P，8D+2P，16D+2P

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

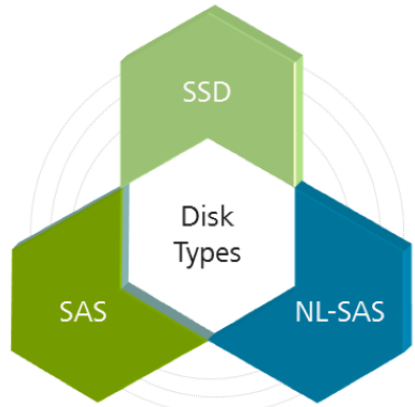
HUAWEI

- OceanStor storage system supports RAID 1, RAID 10, RAID 3, RAID 5, RAID 50, and RAID 6. The supported RAID policy and configurations are shown as the table above.

Huawei Training & Certification

Disk Group (DG)

- Disk Group consist of the combination or integration of multiple similar types of hard drives within a Disk Domain.
  The types of hard drives includes SSD, SAS and NL-SAS.

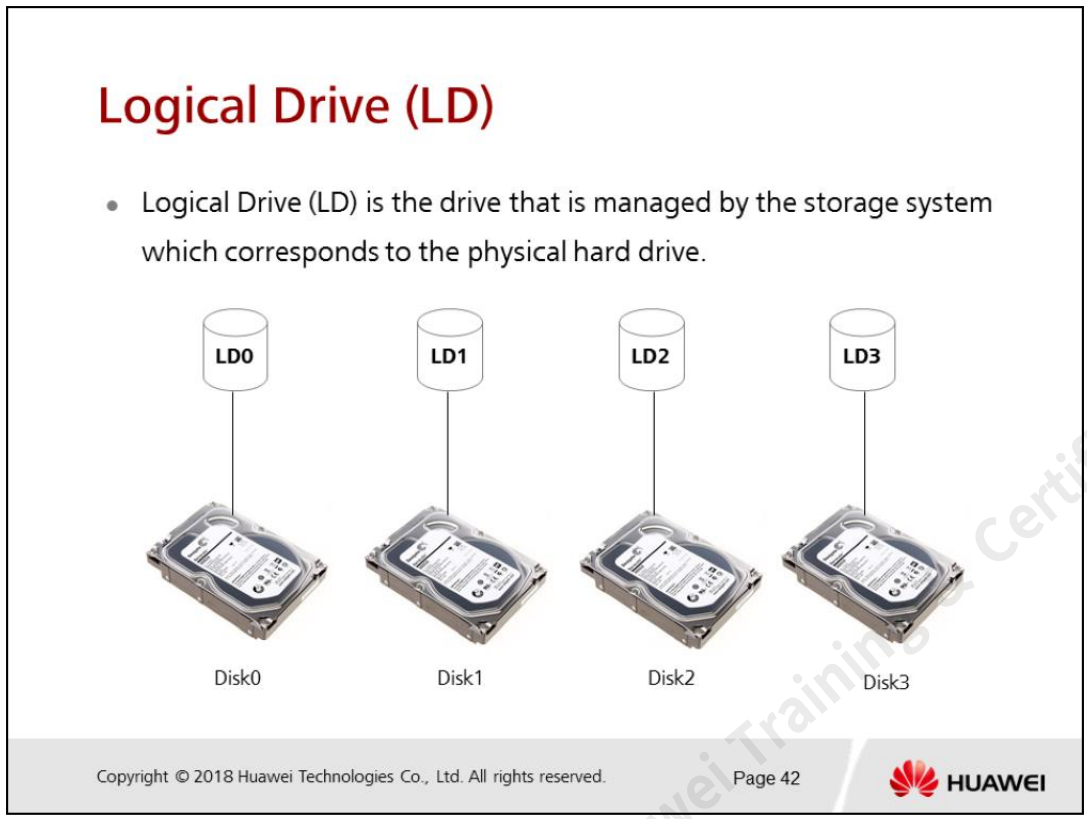


Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

- OceanStor storage system will automatically divide the disks within the disk domain into one or more Disk Groups (DG) based on the number of disks within each disk type.
- 1 Disk Group (DG) can only contain 1 type of hard disk drives.
- Anyone of the many Chunks (CK) in a CKG (Chunk Group) originates from the different hard drives within the same Disk group.
- Disk Group (DG) belongs as the internal system object, and its main purpose is for fault isolation, it is automatically configured by the OceanStor storage system and not represented for external visualization for users.

Huawei Training & Certification

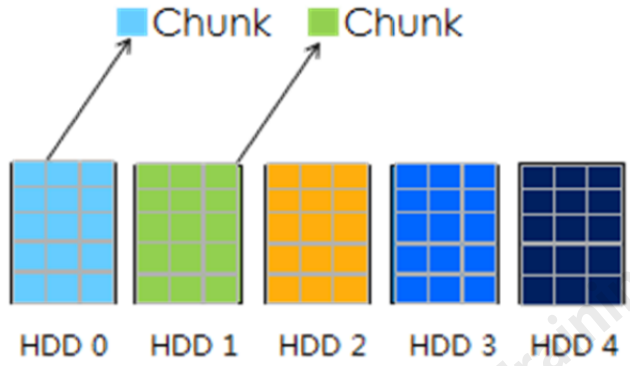


- A logical drive is a drive space that is logically created on top of a physical hard disk drive. A logical drive is a separate partition with its own parameters and functions, and it operates independently. A logical drive can also be called a logical drive partition or logical disk partition.
- A physical drive refers to the hardware unit within a computer, laptop, or server. It is a data storage device that can store and retrieve digital information using one or more platters coated with a thin magnetic layer. Physical drives are usually attached to a computer or laptop. In some cases, they can be external, such as a USB pen drive or an external memory. A physical drive can be partitioned into two or more logical disks which can operate independently of each other.

Huawei Training & Certification

Chunk (CK)

- Chunk or abbreviated as CK, refers to the fixed size physical storage space that is divided from the hard drive capacity within the storage pool, and it is the basic unit for forming RAID.



Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

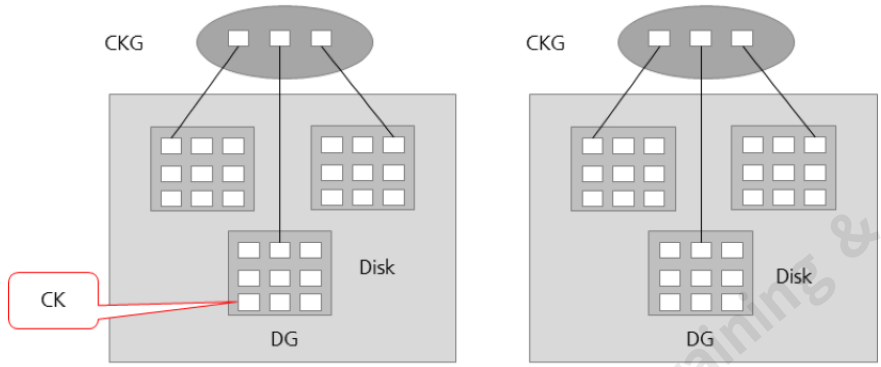
HUAWEI

- The Chunk size is the set of consecutive physical storage space of fixed size divided by the storage system. The Chunk size cannot be changed.
- In simpler terms, a Chunk is the unit for dividing the disk space in a storage system. A chunk is 64 MB in size. It is the fundamental unit in the Huawei RAID2.0+ technology.

Huawei Training & Certification

Chunk Group (CKG)

- Chunk Group (CKG), is formed from Chunks (CK) of multiple disks from the same Disk Group (DG). CK is the logical storage unit based on RAID, and it is the smallest unit for resource allocation in the Storage Pool.



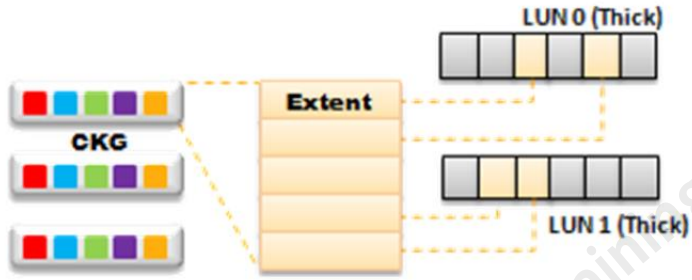
Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

- The CK in the CKG are obtained from the hard disks within the same DG. CKG has RAID properties (RAID properties are actually configured on the Storage Tier). Both CK and CKG belong to internal system objects, and are automatically configured by the storage system which are not represented visually to external application and users.
- In a storage pool, a specified number of chunks make up a chunk group based on a specified RAID group.
- Chunks in a chunk group must come from disks of the same type.

Huawei Training & Certification

Extent

- Extent is the logical storage space of fixed size that is formed using the CKG as the fundamental unit. The size of extent can be changed, and it is the smallest unit for hot data analysis and data migration (data migration granularity).



Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

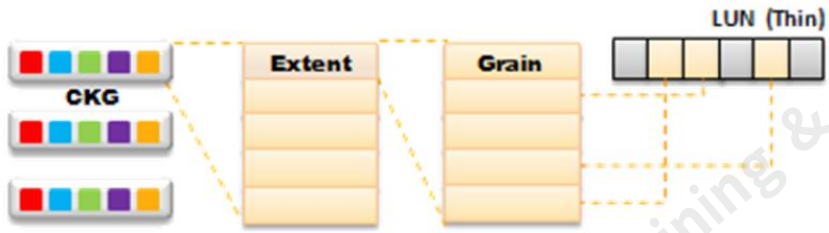
HUAWEI

- An Extent belongs to a Volume or a LUN, the size of the Extent can be configured when creating the storage pool. Once it is configured, the size of Extent can no longer be changed. The size of Extent in different storage pool can be different, but the size of the extent in the same storage pool is unified and identical.
- An extent is fixed-sized logical space by which space in a storage pool is divided. Extent is the smallest unit for applying for and freeing up space and migrating data. An extent is usually 512 KB to 64 MB in size. The default size is 4 MB.

Huawei Training & Certification

Grain

- In Thin LUN mode, the Extent is further divided into finer and smaller blocks. These blocks are called Grain. Thin LUN uses Grain for space allocation. The LBA within the Grain is continuous.



Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

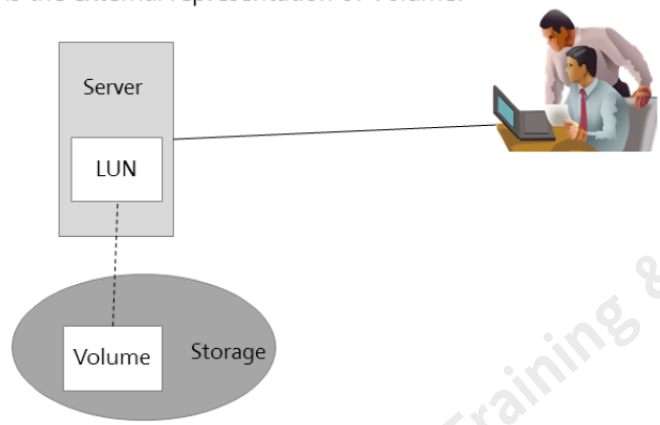
HUAWEI

- Thin LUN is mapped to the LUN using Grain as the unit. This means that multiple Grains forms a LUN which is used to store user data. Grain is only used to form Thin LUNs and there is no such object used to form the Thick LUN. Thick LUN is formed using Extent instead of Grain.
- When a user creates a thin LUN, OceanStor storage system divide the extents into grains and map the grains to the thin LUN. In thin LUN mode, extents are divided into 64 KB grains. In this way, fine-grained management of storage capacity is implemented.
- Logical block addressing (LBA) is a common scheme used for specifying the location of blocks of data stored on computer storage devices, generally secondary storage systems such as hard disk drives. LBA is a particularly simple linear addressing scheme where blocks are located by an integer index, with the first block being LBA 0, the second LBA 1, and so on.

Huawei Training & Certification

Volume & LUN

- Volume is the internal management object within a storage system.
- LUN is the storage unit that can be directly mapped to host for data read/write. It is the external representation of Volume.



Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

- A Volume object is used to organize all the logical storage units such as Extent and Grain in a LUN. We can dynamically apply or release the amount of Extents to increase or decrease the actual storage space of a Volume.
- A LUN is made up with multiple extents. A LUN can contain multiple extents that are from different storage media, so more disks participate in data read and write, which ensures fast response, high efficiency, and improved performance.

Huawei Training & Certification

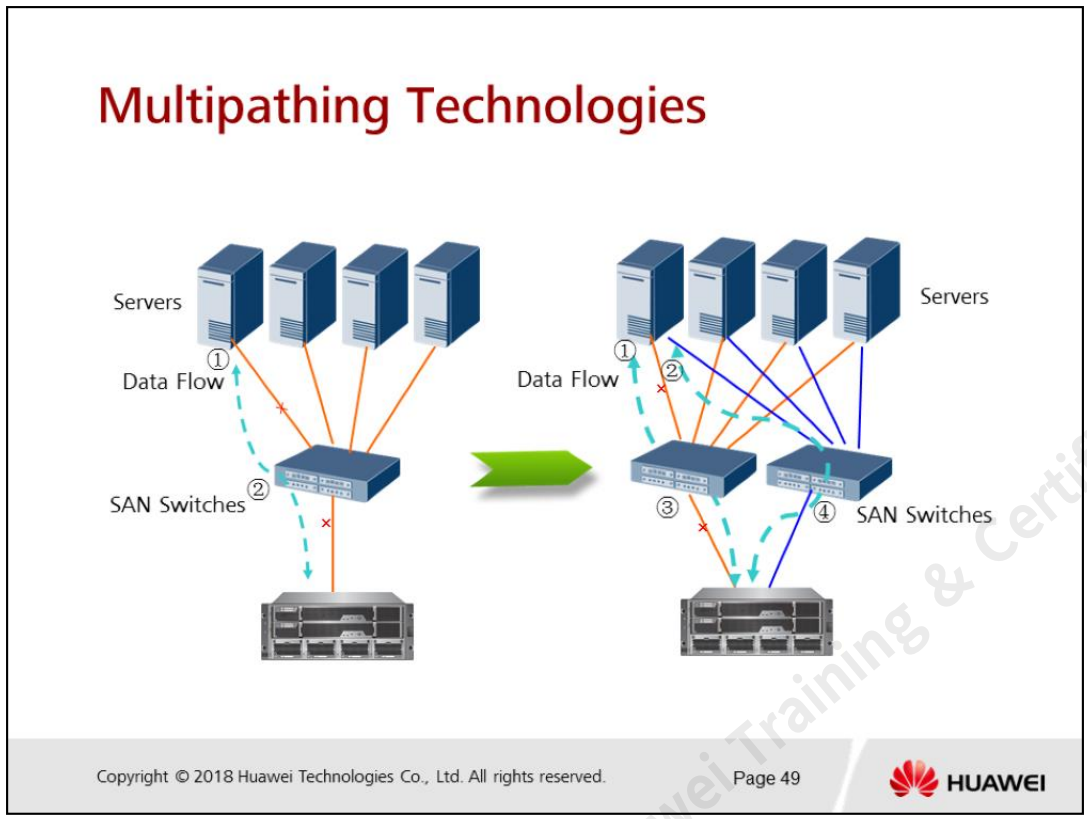
Contents

1. Traditional RAID Technologies.
2. RAID 2.0+ Technologies.
3. **Host Multipathing Technologies.**
4. Disk Reliability Technologies.

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

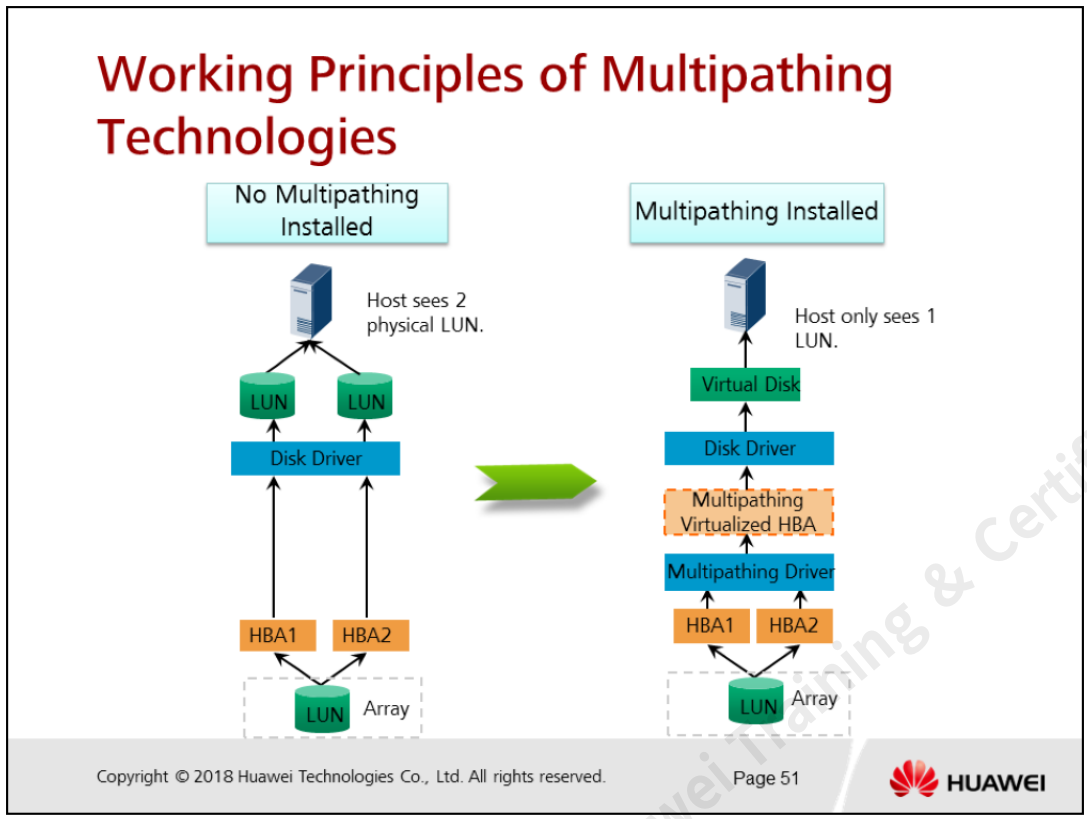
HUAWEI

Huawei Training & Certification
Huawei Training & Certification



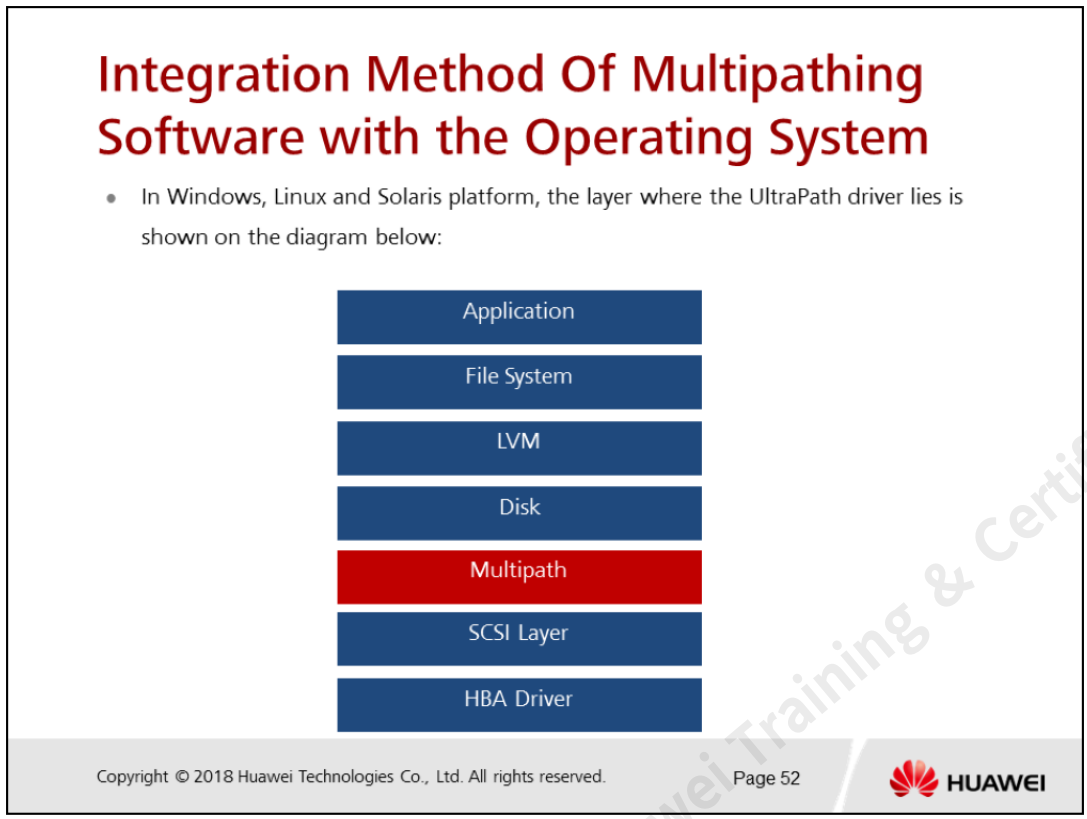
- The common path for user data to reach the disk array: Host → SAN Network → Storage System Controller → Disks in Storage System.
- Multipathing technologies refers to implementation of multiple connections and paths between the host and the storage array, which allows data to be transferred in more than 1 path between the host and storage array across multiple switches to avoid single point of failures. As shown on the right diagram above, there are 2 possible paths between the host and the storage array which are (1, 3) & (2, 4), and there are 2 independent switches along these paths. In this mode, when path (1, 3) is down, data flow from the host will be redirected by the Multipathing software to choose the path (2, 4) to reach the storage array. At the same time, if one of the switch is faulty, it will redirect the data flow towards the other path with the working network switch. When the 1st path has been recovered, the IO flow will automatically switch back to the normal route. The whole path switching and recovery process is transparent towards the host applications, which fully avoids the IO interruption between the host and storage array due to path failures.

Huawei Training & Certification



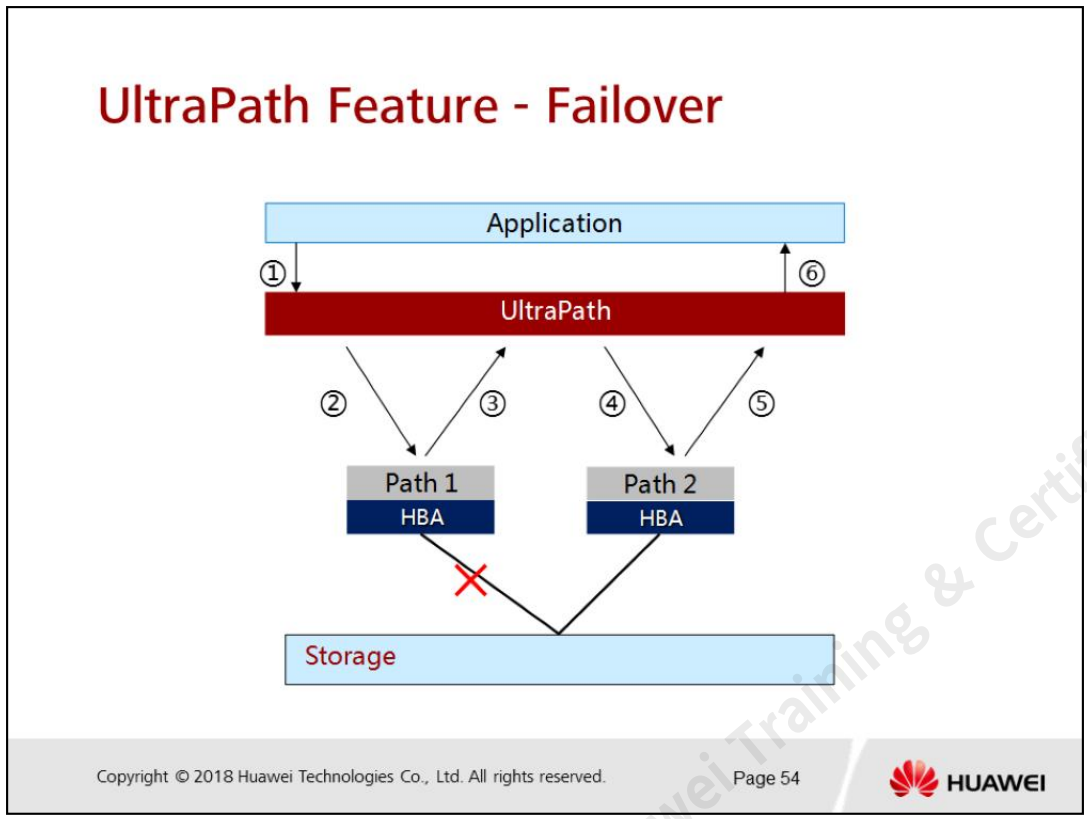
- Functions of Multipathing software:
  - Avoids logical error in operating systems caused by multiple paths directing to the same LUN.
  - Increases links and connection reliability, and avoids system failure due to single path or link failures.

Huawei Training & Certification



- Generally, Multipathing software works underneath the disk drivers to provide virtualized disk to upper level application access. But it can also work above the disk drivers and below the LVM, like most mainstream Multipathing software in the Linux platform.
- Why is there a need to mask the LUNs when using Multipathing software?
  - In a redundant storage network, an application server with no Multipathing software detects a LUN on each path. Therefore, a LUN mapped through multiple paths is mistaken for two or more different LUNs. Redundant LUNs exists because each path reports a LUN directly to the application server. Due to the identification errors of the application server, different applications on the application server repeatedly write different data to the same location of the LUN, resulting in data corruption.
  - To resolve this problem, the application server must identify which is the real and available LUN. As UltraPath is able to acquire configuration information of the storage system, it clearly knows which LUN has been mapped to the application server.
  - UltraPath installed on the application server masks redundant LUNs on the operating system driver layer to provide the application server with only one available LUN, the virtual LUN. In this case, the application server only needs to deliver data read and write operations to UltraPath that masks the redundant LUNs, and properly writes data into LUNs without damaging other data.

Huawei Training & Certification Huawei Training & Certification



- When the same storage controller has multiple paths connecting to the host, and when the LUN belong to that particular storage controller, if a main path has failures, it will preferably choose the path on the other ports on the same storage controller. If no such path exists, then it shall choose the path that is connected to the other storage controller to send or receive I/O from host.
- UltraPath will automatically transfer the I/O using other paths in the event of a path failure. The process for this failover is shown on the diagram above:
    1. Application sends I/O towards the virtual disk generated by the UltraPath Multipathing software.
    2. UltraPath forwards the I/O onto Path 1 (SCSI device).
    3. Faulty path causes the I/O on the Path 1 fails.
    4. UltraPath retransmit the I/O on the other alternative Path 2.
    5. Path 2 I/O process is successful.
    6. UltraPath sends the I/O process success message to the upper layer application.

Huawei Training & Certification

UltraPath Feature - Failback

- UltraPath will retransmit the I/O back to the normal route after detecting that the normal path has recovered from failure.
- There are two methods of path recovery:
  - For systems that supports hot plugging technologies (such as Windows), the broken path between the host and storage will cause the SCSI device to be removed, correspondingly when the link has recovered it will trigger the automatic connection of the SCSI device. In this scenario, UltraPath will detect the path recovery instantly.
  - For systems that doesn't support hot plugging technologies (such as AIX, earlier version of Linux), UltraPath detects path recovery in periodic interval cycle. This means that UltraPath will check the paths in the configured specific interval of time to see if the path is recovered or not.

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

Huawei Training & Certification
Huawei Training & Certification

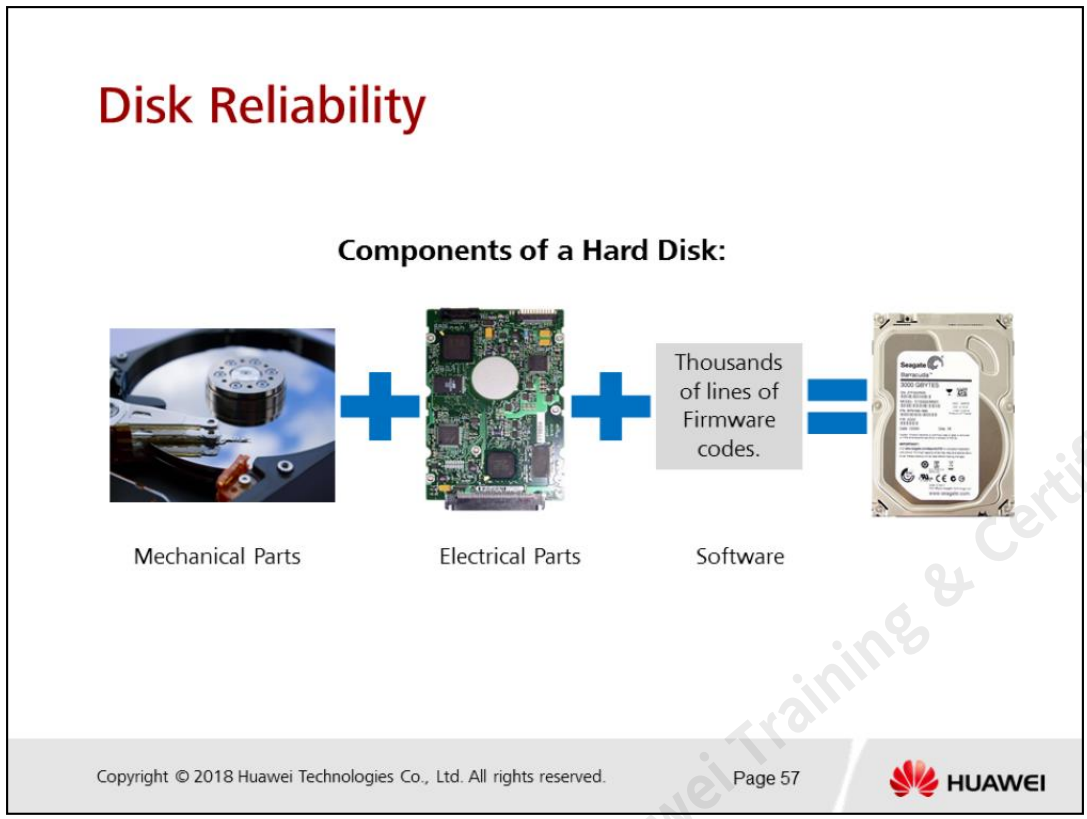
Contents

1. Traditional RAID Technologies.
2. RAID 2.0+ Technologies.
3. Host Multipathing Technologies.
4. **Disk Reliability Technologies.**

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

Huawei Training & Certification
Huawei Training & Certification

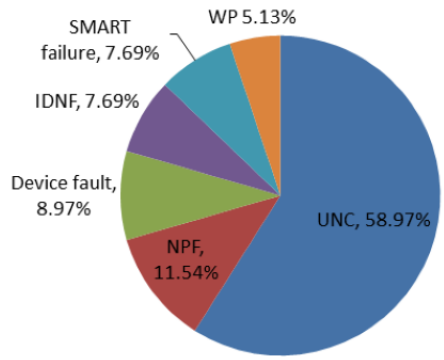


- Electrical parts and software's function is to move the mechanical parts to complete the data access, retrieval and data storing.
- Hard disks are the component that store the user data in the storage systems, it is a piece of hardware that is frequently accessed for data reading and data writing. Core applications and important business data are stored in these devices which is crucial for it to be reliable when the data within it is needed.
- As disk capacity grows, more and more data is stored on a single drive. This makes it important to ensure that the hard disks used are reliable because a faulty drive may mean hours of data reconstruction or data loss.
- As the hard disks are made from mechanical, electrical and software parts, it means that there are many ways and factors that can cause a drive failure.
- In the following segment, we will look through the different types of disk failures, how to recover from those failures and the types of technologies used to increase the disk reliability to ensure higher data protection.

Huawei Training & Certification

Types of Disk Failures

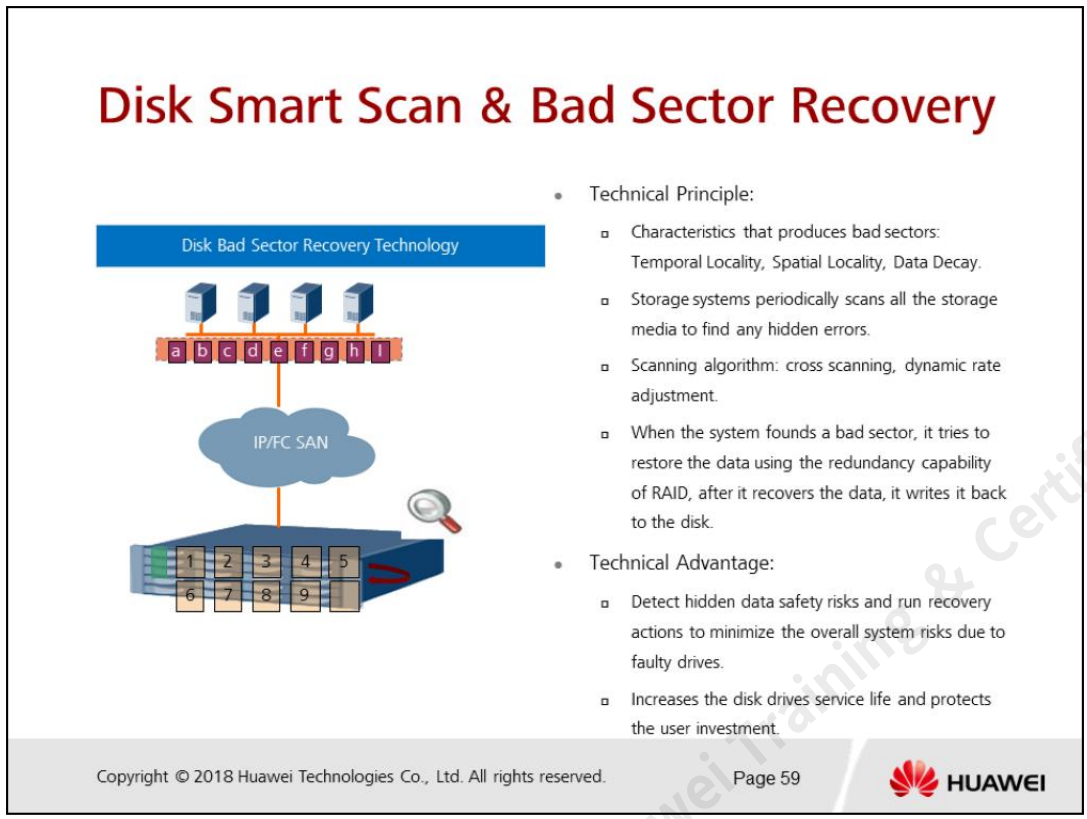
- TOP 1: UNC (Uncorrectable) about 60% of the disk errors are caused by bad sectors.
- TOP 2: NPF (No Problem Found) about 12% of the disk errors found no problems after the disk is returned to factory for analysis.
- TOP 3: Disks will predict failures before it happens, but the failure detection rate is just 7.69%, most of the failures are not able to be predicted in advance.



Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

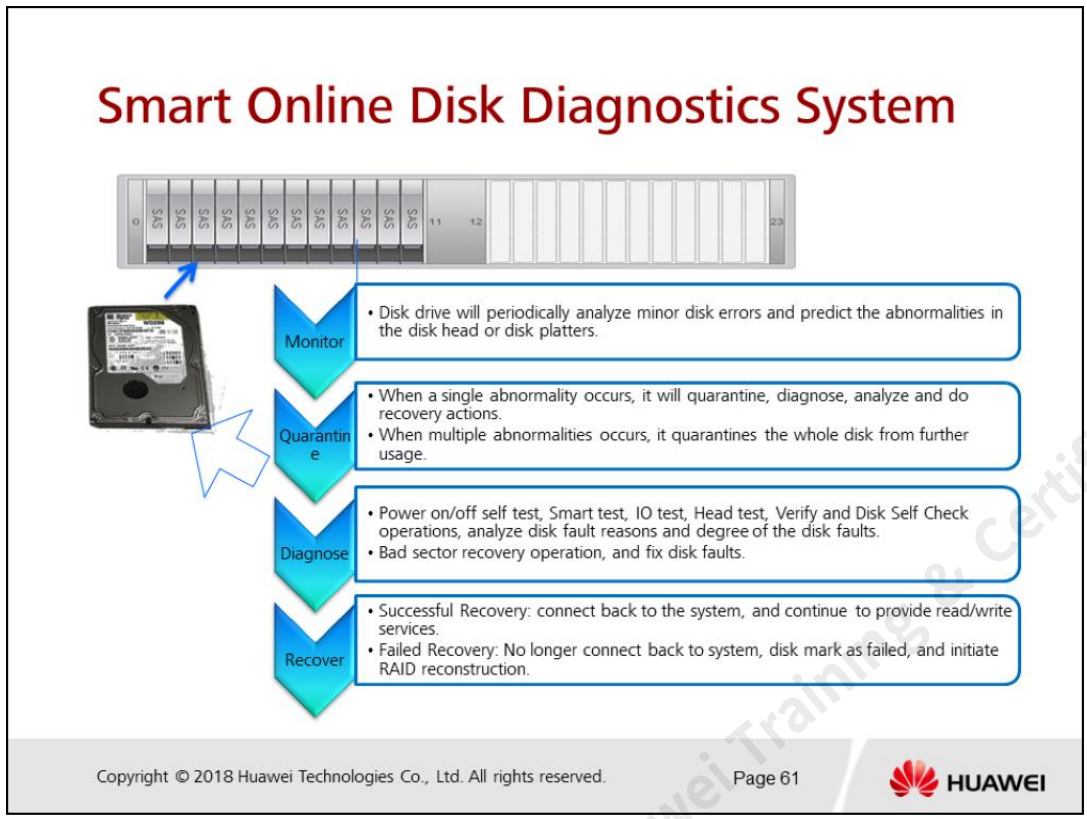
- UNC: Uncorrectable error which is an ECC error in the data field could not be corrected (a media error or read instability).
- IDNF: ID Not Found error where the required cylinder, head, and sector could not be found, or an ECC error occurred in the ID field.
- WP: Write Protected error where data could not be written to the disk.
- Factors that cause bad sectors:
  - Logical bad sectors: The read/write disk head has stains, or dust particles within the hard drive, external disturbances causing instable movement of the head, accidental interruption when data writing(e.g. power issues), vibration.
  - Physical bad sectors: Flaws on the disk platter, metal particles within the hard drive, external shock causing head crash etc.
  - Degeneration of disk head: Electrostatic discharge (ESD) or electrostatic overstress (EOS), material issues of the disk head, bad working environment for the disks accelerates the degeneration (high temperature, instable voltage) etc.
- EOS is a term used to describe the thermal damage that may occur when an IC is subjected to a current or voltage that is beyond the datasheet specification limits of the device. An EOS event can degrade the IC or cause permanent functional failure. EOS is a much slower phenomenon than ESD, but the associated energy is very high.
- Electrostatic discharge is the sudden flow of electricity between two electrically charged objects caused by contact.

Huawei Training & Certification



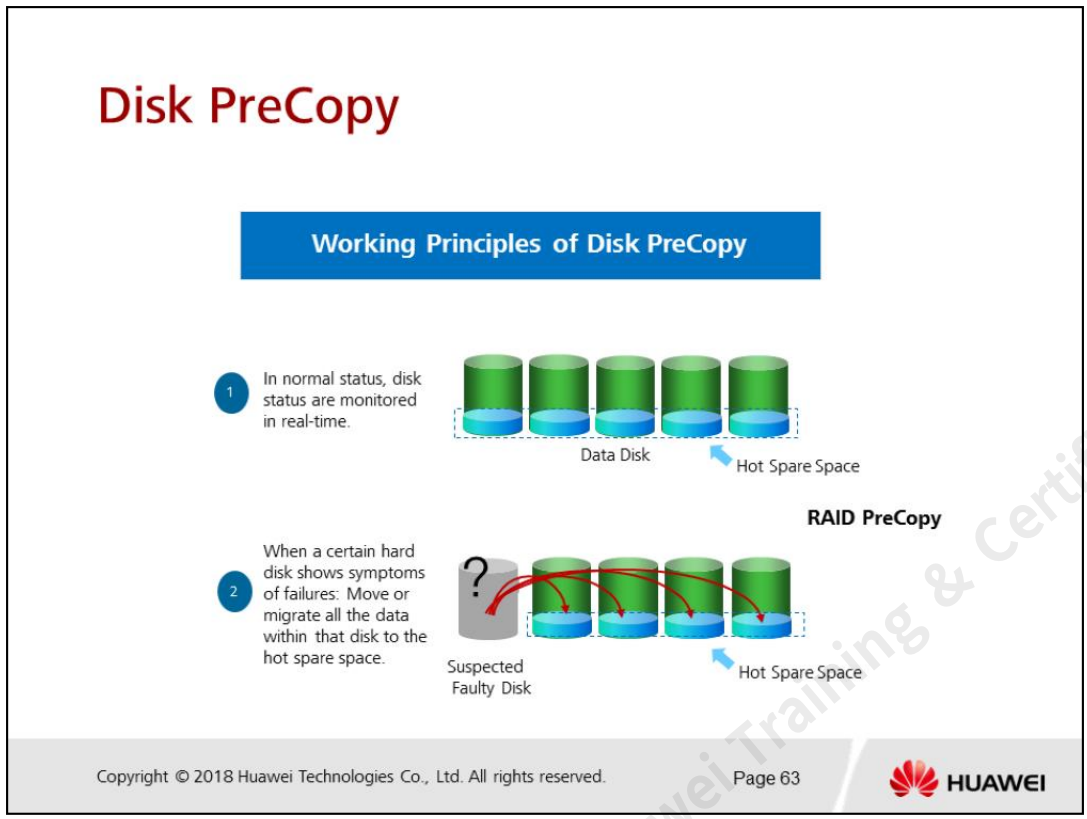
- Bad sectors are often caused during write operations or scratches, and these bad sectors are not found instantly and are only found during data reading process.
- Hence, the background scanning for bad sectors are very important, if not when we read the data, there might be a chance that both data on the RAID stripe is damaged which will cause data loss because it can no longer be recovered. Example: In a RAID 5 with 3 disks scenario, both 1 block of data and 1 block of parity data is corrupted at the same time, which makes the data recovery not possible to recover the lost data.
- As the disk capacities are getting larger and larger, simple sector by sector scanning will take 1 month or even a few months to complete 1 cycle of scanning in large storage systems. This in reality loses the effect of finding the bad sectors quickly in time to recover before further damage is done.
- Huawei has discovered some patterns in the generation of bad sectors during research:
  - If one sector is bad or damaged, the other nearby sectors are likely to be damaged.
  - If a bad sector is found at a certain point of time, at that same point of time there is a high possibility that other sectors might be damaged.
- Based on those pattern found, Huawei designed a scanning algorithm: It will periodically leap through and examine certain sectors within the disk, once it found a bad sector, it will check the nearby sectors and increase the speed of scanning, once it found that the bad sectors is reducing then the speed of scanning is lowered.

Huawei Training & Certification



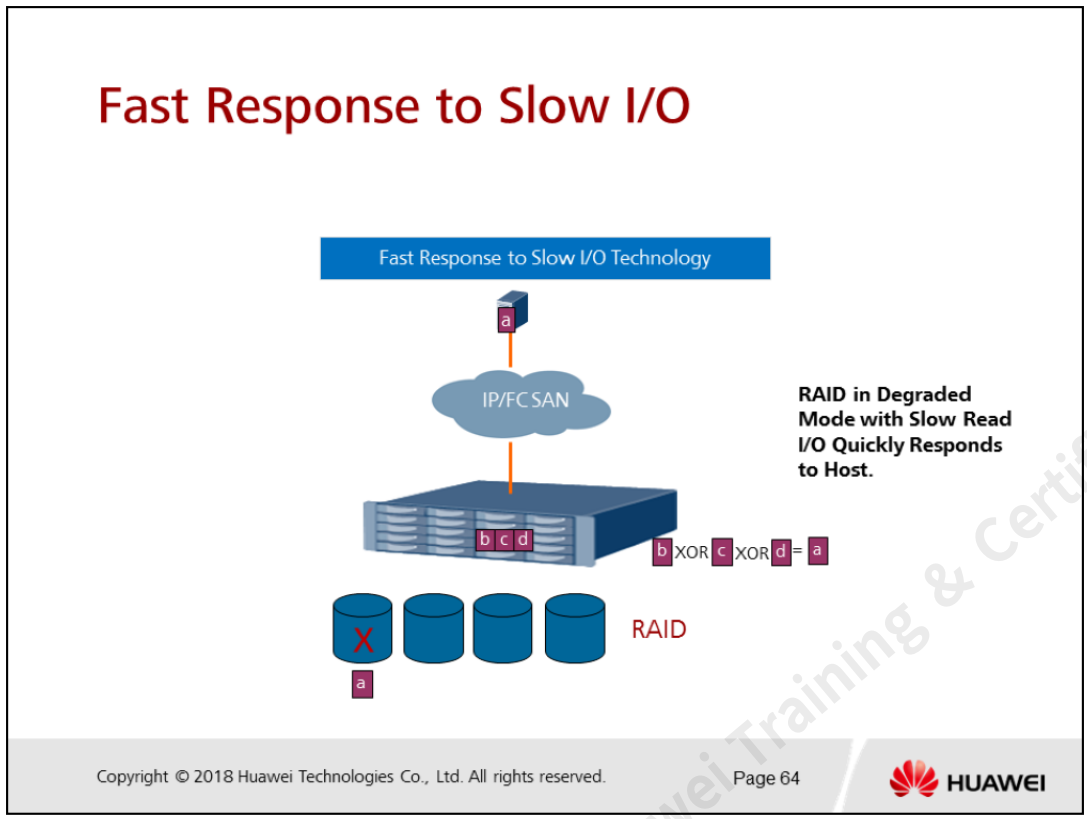
- Advantage: Solves the NPF(No Problem Found) issue (due to abnormality on disk firmware, short external vibration shock or other factors causing disk failures), minimizes the rate of RAID failure, and maximizes the business continuity and data reliability to the greatest extent.
- Based on analysis, annual SATA drives failures amounts to 2.5%, 70% of it due to repairable bad sectors. Periodic scanning on the disks in a 30 days cycle, means that bad sectors on the disks can be found and repaired every 30 days. Hence, we can calculate the annual disk failure rate after implementing the periodic disk scanning to be $\lambda=2.5\%\times0.3+2.5\%\times0.7\times30/365=0.89\%$. The RAID failure rate can also be lowered by 1 level.
- The are two methods to detect the bad sectors on the disks
  - Automatic Analysis of Read/Write Failures: There are many reasons for read/write failures, such as spoilt disk head, disk interface failure, connection cable failure, unplugged hard disks etc. When a disk read/write failure occurs, the storage system will analyze whether the cause of failure is due to bad sectors based on checking the current state of the storage system, current state of hard disk, I/O failure information.

Huawei Training & Certification



- Disk drives has built in mechanism to predict failures, but those mechanisms are not accurate to fully rely on.
- Huawei has design its own failure prediction mechanism based on the fundamentals of the built in failure detection mechanism of disk drives, which increased the accuracy of the failure detection by a few times.
- Technical Principle:
  - The storage system runs health monitoring on the disks, when it finds that the amount disk errors has went over the preconfigured threshold, it will immediately mark that disk as faulty.
  - RAID group uses the copy method to move the data in the affected disk to the hot spare space, and send an alarm to the admin personnel at the same time to remind them to change the suspected faulty disks.
- Technical Advantage: Greatly lowers the rate of RAID reconstruction, and increase system reliability.

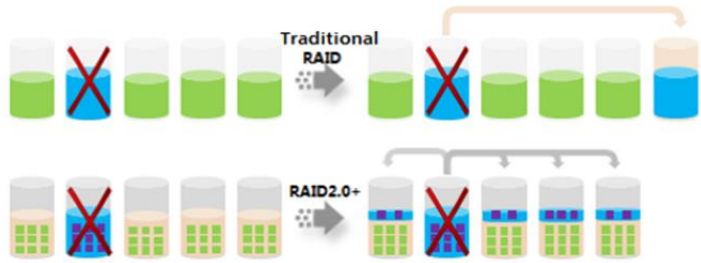
Huawei Trainings & Certification



- Disk drives have mechanical parts, so some slow I/O at times is quite normal, and normal application may not even noticed the difference. But for customer core applications, it is quite sensitive towards access delay, the slow I/O may even affect the application or even cause interruption in service.
- Huawei storage fully utilizes the RAID technologies in this kind of scenario, when a certain disk I/O is quite slow, it quickly gives up on waiting for that disk and uses the data in other disks and parity information to quickly generate the data that is required by the core application. This effectively ensures the business continuity of the core applications.
- Technical Principle:
  - Disks causes service to hang or disrupted due to the long I/O response time (seconds level) because of physical bad sectors, head problems, vibration shock, multiple retries etc.
  - Storage system monitor the disk I/O in real-time, if it is slower than the preconfigured threshold, it will use RAID degrade mode to quickly generate the data using parity information and respond to host.
- Technical Advantage: Quickly respond to business services, truly support the core enterprise applications that are very sensitive towards time delay.

Huawei Training & Certification

Overview of RAID Reconstruction Technologies



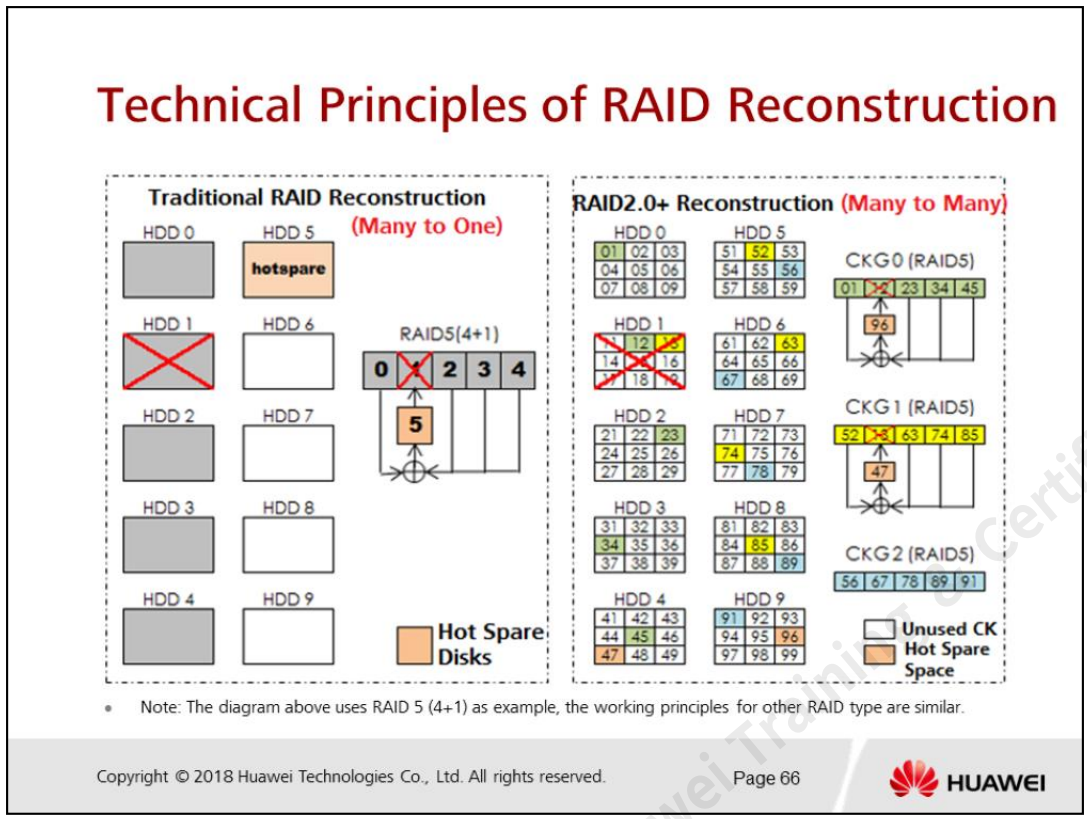
Traditional RAID	RAID2.0+
Need to manually configure the independent global or local hot spare disks.	Distributed hot spare space, no need for manual configuration.
Many to one reconstruction, where the reconstructed data are written in sequence to a single hot spare disk.	Many to many reconstruction, where reconstructed data in written in parallel to many disk drives.
Hot spot exists and longer reconstruction time.	Load balancing and short reconstruction time.

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

- OceanStor storage system has implemented multiple fault tolerant designs towards the disks, which includes reliable protection features such as online disk diagnostics, DHA (Disk Health Analyzer) for disk diagnostics and failure prediction, background bad sector scanning. RAID 2.0+ technology will automatically reserve a certain amount of hot spare space in the disk domain based on the hot spare policies configured, and the users no longer have to manually configure them independently.
- When the system detects a region on the disk that in unrepairable or the whole disk has failed, the system will automatically begin data reconstruction, which will quickly reconstruct the affected data to the hot spare space of other hard disks to achieve fast self healing of the storage system.

Huawei Training & Certification



- The growth of disk capacity makes Traditional RAID to face a serious problem: 10 years ago, reconstructing a disk only takes minutes, but now a single disk reconstruction could take few hours or dozens of hours. Longer disk reconstruction makes the enterprise storage to stay in a Degraded RAID mode without any fault tolerance for long hours, this poses a huge risk for data loss. Real life scenarios such as data loss due the pressure of data reconstruction process and business operations during the reconstruction process are quite common.
- RAID2.0+ technology based on low level block virtualization solves the problem of the bottleneck target disk (hot spare disk) in traditional RAID reconstruction, making the write bandwidth of the reconstructed data flow no longer the bottleneck in data reconstruction speeds. This effectively increased the reconstruction speed greatly and lowered the rate of dual disk failures, and at the same time increased the overall system reliability.

Huawei Training & Certification

Summary

- This module introduced:
  - Concepts, Working Principles and Types of RAID Technologies.
  - Technical Principles of RAID2.0+ Technologies.
  - Functions of Host Multipathing Technologies.
  - Disk Reliability Technologies.

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

Huawei Training & Certification
Huawei Training & Certification

Quiz

1. (True or False) Is RAID 5 more higher level than RAID 0 ? ( )
2. Which of the followings are common RAID types ? ( )
   A. RAID 10
   B. RAID 5
   C. RAID 0
   D. RAID 4

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

- Answers:
  - F.
  - ABC.

Huawei Training & Certification
Huawei Training & Certification

Thank You

www.huawei.com

Huawei Training & Certification
Huawei Training & Certification

08 Common Advanced Storage Technologies

www.huawei.com

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

Huawei Training & Certification
Huawei Training & Certification

Foreword

- This module introduces the features of the common advanced storage technologies such as:
  - SmartThin
  - SmartTier
  - Smart QoS
  - SmartPartition
  - Snapshot
  - SmartQuota

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

Huawei Training & Certification
Huawei Training & Certification

Objectives

- Upon completion of this module, you will be able to:
  - Understand the principles, configuration process and application scenarios of SmartThin technology.
  - Understand the principles, configuration process and application scenarios of SmartTier technology.
  - Understand the principles, configuration process and application scenarios of SmartQoS technology.
  - Understand the principles, configuration process and application scenarios of SmartPartition technology.
  - Understand the principles, configuration process and application scenarios of SmartQuota technology.

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

Huawei Training & Certification
Huawei Training & Certification

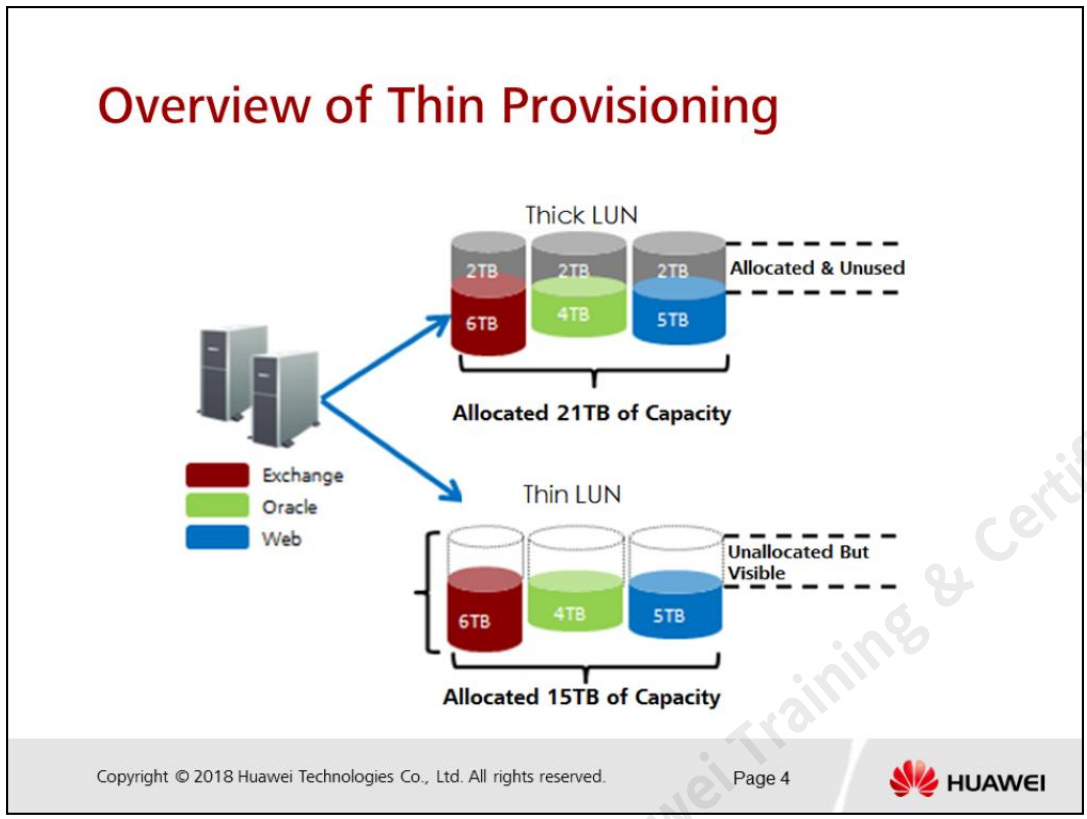
Contents

1. **SmartThin**
2. SmartTier
3. SmartQoS
4. SmartPartition
5. Snapshot
6. SmartQuota

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

Huawei Training & Certification
Huawei Training & Certification

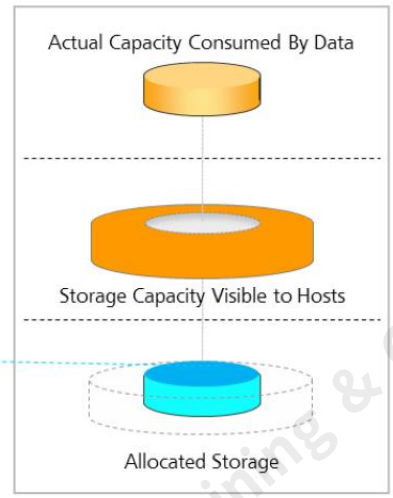


- The diagram above shows the differences in storage space allocation between the Thin LUN provided by SmartThin and the Traditional LUN.
- SmartThin technology has improvements as the following on top of traditional LUN provisioning:
  - Improvement 1: The storage space of LUNs that uses SmartThin technology is not allocated upon creation, but only allocated when it actually needs that storage space.
  - Improvement 2: Based on the fundamentals of improvement 1, the storage space of LUNs that uses SmartThin technology can be dynamically adjusted in size.
- When data capacity has exceeded the initial prediction, we can dynamically adjust the storage space of that LUN. The unused space can be used as public storage space that can be allocated to any LUNs that require the storage space. Thus, there won't be any private unutilized space in LUNs, which increases the utilization rate and efficiency ratio. At the same time, dynamic storage space adjustment provides the capability of online configuration of storage space size, which allows online expansion of storage that doesn't affect the business in operation.

Huawei Training & Certification

Thin LUN

- It is a logical disk accessible by hosts.
- It is referred to as Thin LUN and its storage space are dynamically allocated from the storage pool based on actual user storage capacity in use.

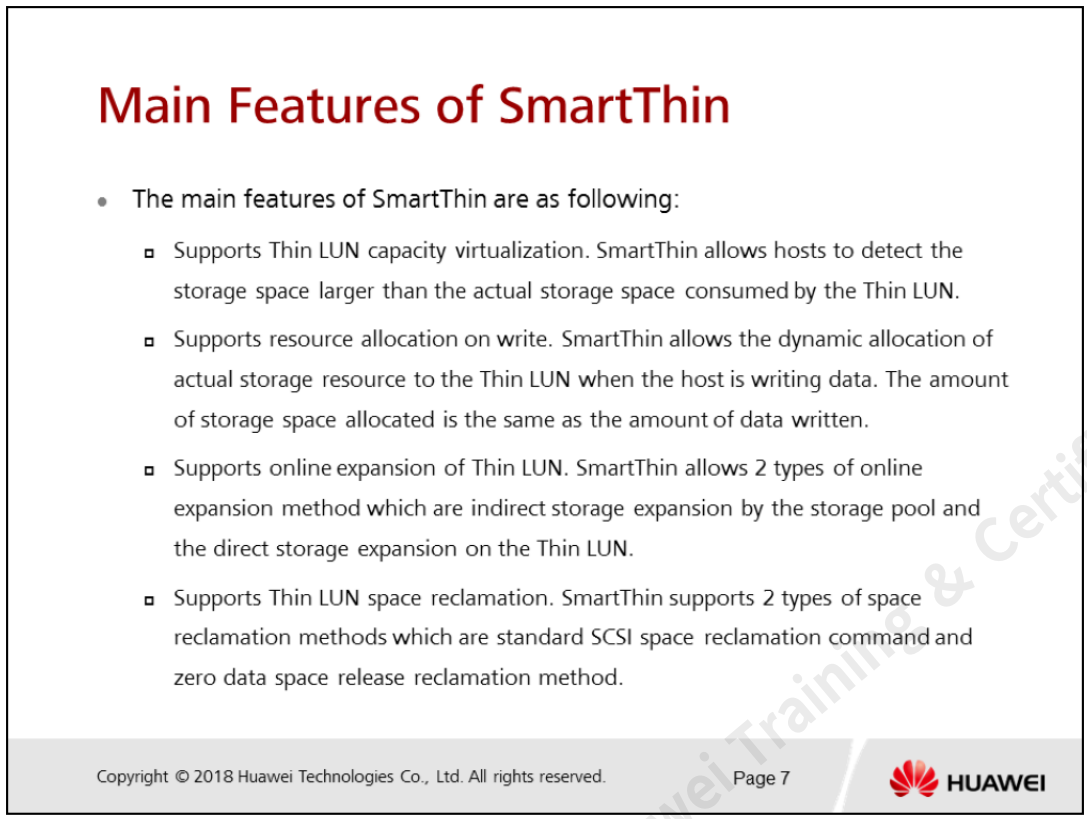


Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

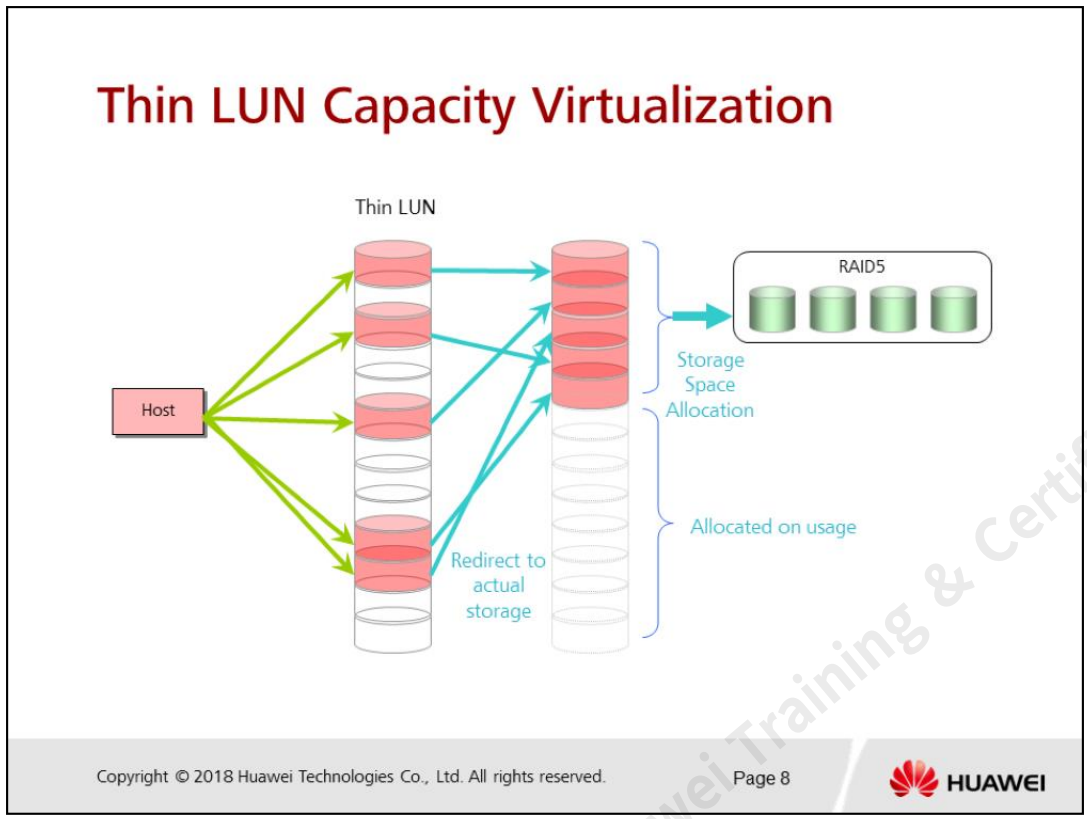
- Data Integration: It is a LUN that can be mapped to host from the perspective of storage systems.
- Fully Usable: Normal Read/Write on the Thin LUNs.
- Dynamic Allocation: Storage resource are allocated on write.
- Two different types of LUNs may be created on Storage Pools – Thick LUNs and Thin LUNs. While both allocate space on demand, there are significant differences between them in terms of both operation and performance.
- When a Thick LUN is created, the entire space that will be used for the LUN is reserved. If there is insufficient space in the Storage Pool, the Thick LUN will not be created. For example, if a Thick LUN of 100GB is created, 100GB worth of storage space from the storage pool is allocated upon creation to the Thick LUN.

Huawei Training & Certification



- SmartThin provides a storage management mode that supports on-demand allocation of storage space.
- SmartThin applies to the following scenarios:
  - Core system services that have high requirements for service continuity use SmartThin for online expansion, ensuring ongoing services. For example, SmartThin applies to financial systems.
  - Services whose data growth is hard to predict use SmartThin to allocate physical storage space based on requirements, avoiding wasting storage space. For example, SmartThin applies to email and web disk services.
  - Mixed services that have diverse storage requirements use SmartThin to contend for physical storage space and achieve optimal configuration of physical storage space. For example, SmartThin applies to carrier's services.

Huawei Training & Certification

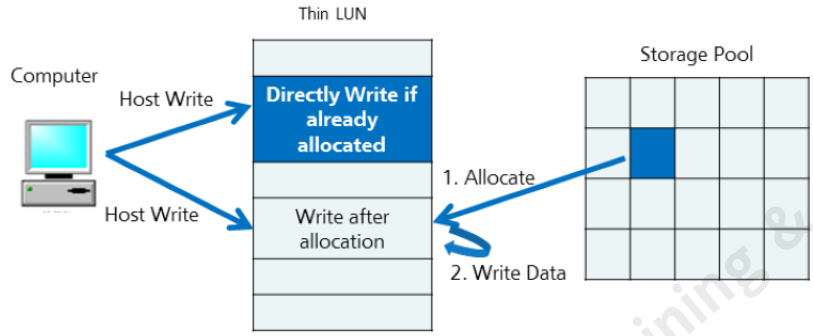


- SmartThin creates Thin LUN based on RAID2.0+ virtualized storage pools, which means that the Thin LUN and the traditional LUN coexists in the same storage pool.
- Thin LUN is a logical unit that can be created in the storage pool and mapped to host for direct access.
- The capacity of Thin LUN is not the actual physical storage space, but it is a virtualized value. The physical storage space is only applied through on write allocation policy from the storage pool when there is actual I/O process on the Thin LUN.
- In the RAID2.0+ environment, storage system will divide the storage capacity of the storage pool into individual small data blocks called CHUNK. These CHUNKs are used to form the CHUNK Groups (CKG) for RAID. This allows the data to be evenly distributed across all the hard disks in the storage pool, and resource management is based on CHUNK as the unit. SmartThin uses CKG and divides it further into smaller Extent (the smallest storage unit that can be allocated) as the unit for storage space organization.
- Thus, Thin LUN and Thick LUN exists within the same storage pool, and can utilize the physical storage capacity within the storage pool making storage planning much easier and flexible, and avoid the issues of providing separate storage pool space for Thin LUN and Thick LUN.

Huawei Training & Certification

Capacity - on - write

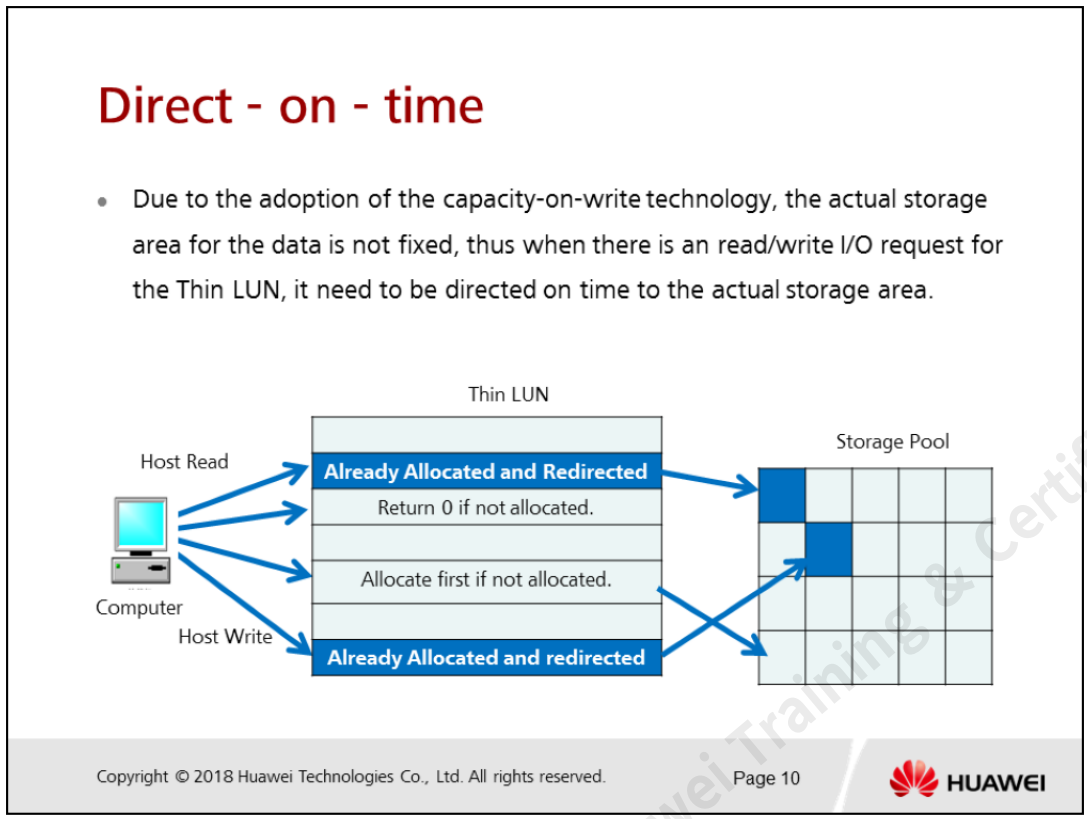
- Write I/O request to the Thin LUN will trigger the storage capacity allocation.
- When the currently allocated physical storage is lower that the threshold, it will apply for new storage space from the storage pool.



Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

- Capacity-on-write: When Thin LUN receives a data write request from the host, it will firstly use the direct-on-time technology to determine whether the logical storage area for the write request has already been allocated with physical storage space. If it is not allocated, then it will trigger the space allocation, and the smallest unit for the space allocation for Thin LUN is grain, where the size of the Grain is 64Kb. Once the space is allocated, the data will be written to the newly allocated actual storage area.

Huawei Training & Certification



- Due to the adoption of capacity-on-write technology, the relationship between actual storage area of data and the logical storage area of the data is not fixed and not constant, thus can no longer be calculated based on a exact formula, but calculated during write time based on the principle of direct-on-time. In simpler words, since the storage space is allocated on write, it requires a calculation during the write process to map the logical storage area on the LUN to the actual storage area in the storage systems. Similarly, when there is a read request, a check is required to determine the mapped actual storage area to access the data. Both these calculations are run on-time when the request is initiated and thus are referred to direct-on-time.
- Thus when there is a read/write request to the Thin LUN, it needs to be redirected based on the relationship of the logical and actual storage area, and these mappings relies on the mapping table.
- The mapping table's main function is to record the mapping relationship between the actual storage area and the logical storage area. During the write process the mapping table is dynamically updated, and the same mapping table is checked during the read process. Hence, the Direct-on-time operation can also be divided into Read Redirection and Write Redirection.

Huawei Trainings & Certification Huawei Training & Certification

Application Scenario

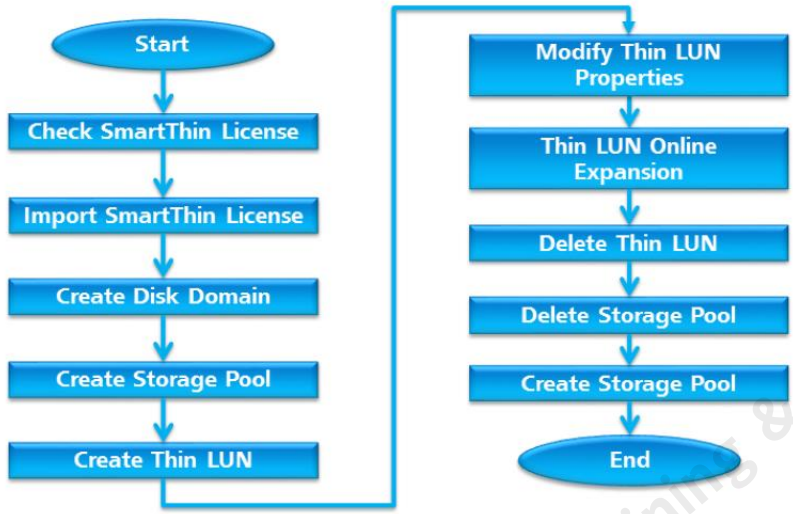
- For core business services that has high requirements for business continuity, using SmartThin configuration allows online expansion of the system without interruption to business service. For example: bank transaction systems.
- For businesses that has high application data growth speed that is unable to be accurately evaluated, using SmartThin configuration allows on demand allocation of physical storage space and avoid wastage in utilization rate. For example: Email services and Web Disk services.
- For business with complex systems and has different requirements for storage for different services and applications, using SmartThin configuration allows different business services to compete for the physical storage space, allowing optimum configuration for physical storage space. For example: Carrier Service Providers.

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

Huawei Training & Certification
Huawei Training & Certification

Configuration Process



Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

Huawei Training & Certification
Huawei Training & Certification

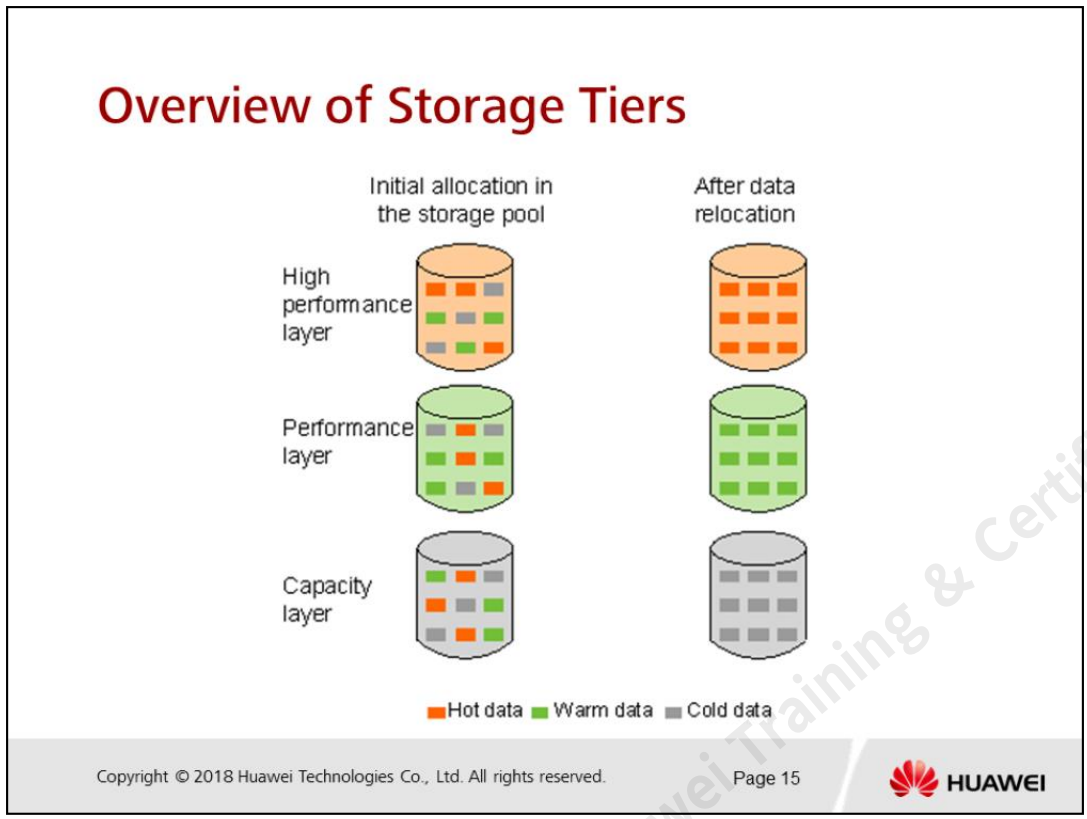
Contents

1. SmartThin
2. **SmartTier**
3. SmartQoS
4. SmartPartition
5. Snapshot
6. SmartQuota

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

Huawei Training & Certification
Huawei Training & Certification



- SmartTier is a proprietary storage tier software developed by Huawei based on the RAID2.0+ technology. It automatically matches the data of different activity level with the different characteristics of the storage media, in order to increase the storage system performance and lower the user costs.
- Tier storage technology completes tiered storage of data through data migration. Currently, tier storage technology is divided into manual migration and automatic migration.
- Manual migration is performed by system maintenance personnel based on the current state of operation of the storage system, and the pressure on each business systems. They manually migrate the data from one LUN to another LUN.
- Automatic migration is based on the access frequency of the files or blocks of data. The access frequency helps to differentiate between the hot data which are frequently accessed and cold data that are less accessed. Different types of storage media is matched by their characteristics to the hot or cold data. Frequently accessed hot data is automatically migrated to the high performance storage media, and cold data are automatically migrated to the high capacity and low cost storage media.

Huawei Training & Certification

Levels of Storage Tier

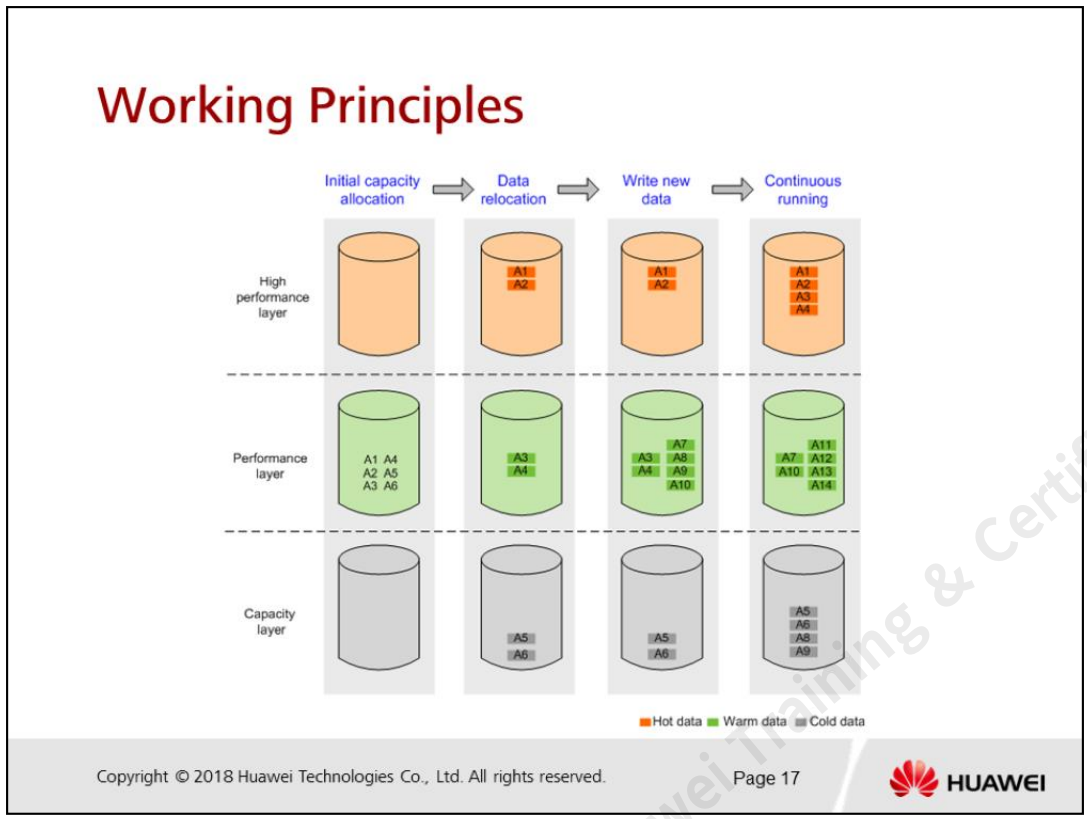
Storage Tier	Disk Type	Disk Characteristics	Data Characteristics
High Performance	SSD	SSD has very short response time, each unit of storage capacity has high costs.	High Performance Tier is suitable for data with high frequency of access.
Performance	SAS	SAS disks has short response time and each unit of storage capacity has medium costs.	Performance Tier is suitable for data with medium frequency of access.
Capacity	NL-SAS	NL-SAS disks has relatively longer response time, and each unit of storage capacity has low cost.	Capacity Tier is suitable for data with low frequency of access.

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

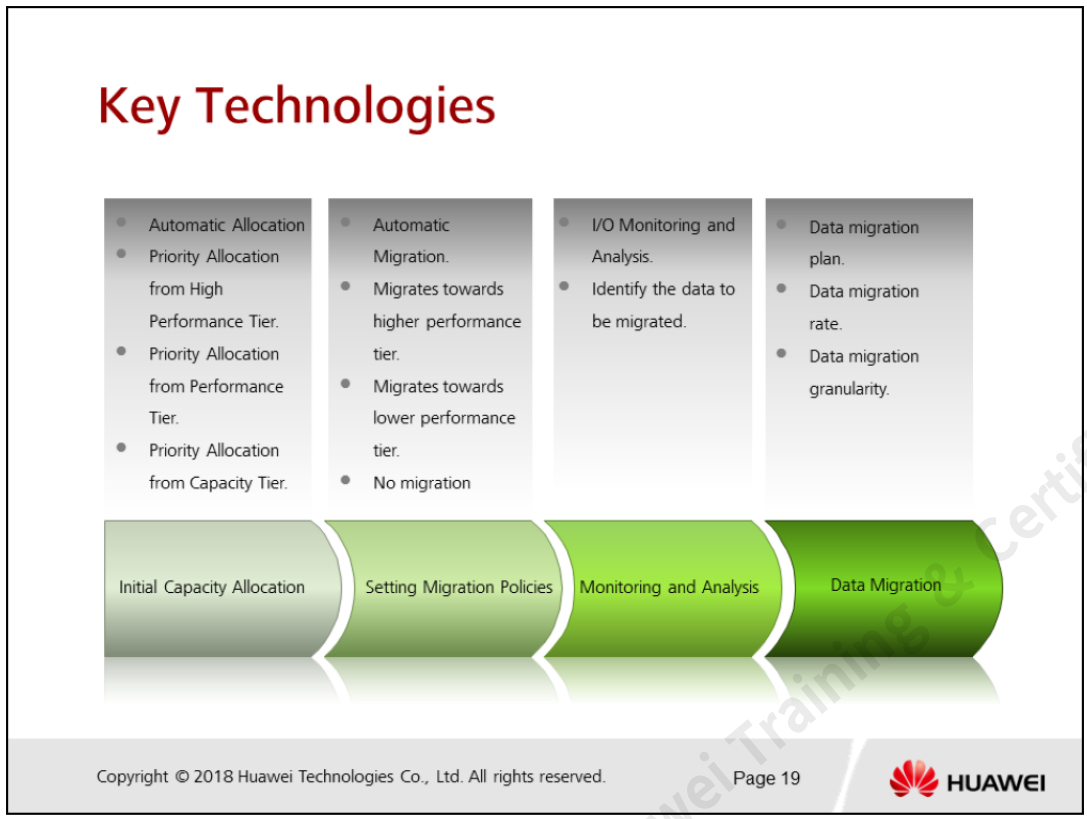
- SmartTier divides the different types of storage media into 3 levels of storage tiers based on the performance. SSD drives with the highest performance is categorized into the High Performance Tier, SAS drives forms the Performance Tier and NL-SAS drives forms the Capacity Tier.
- Each storage tier independently uses the same disk type and RAID policies.

Huawei Training & Certification



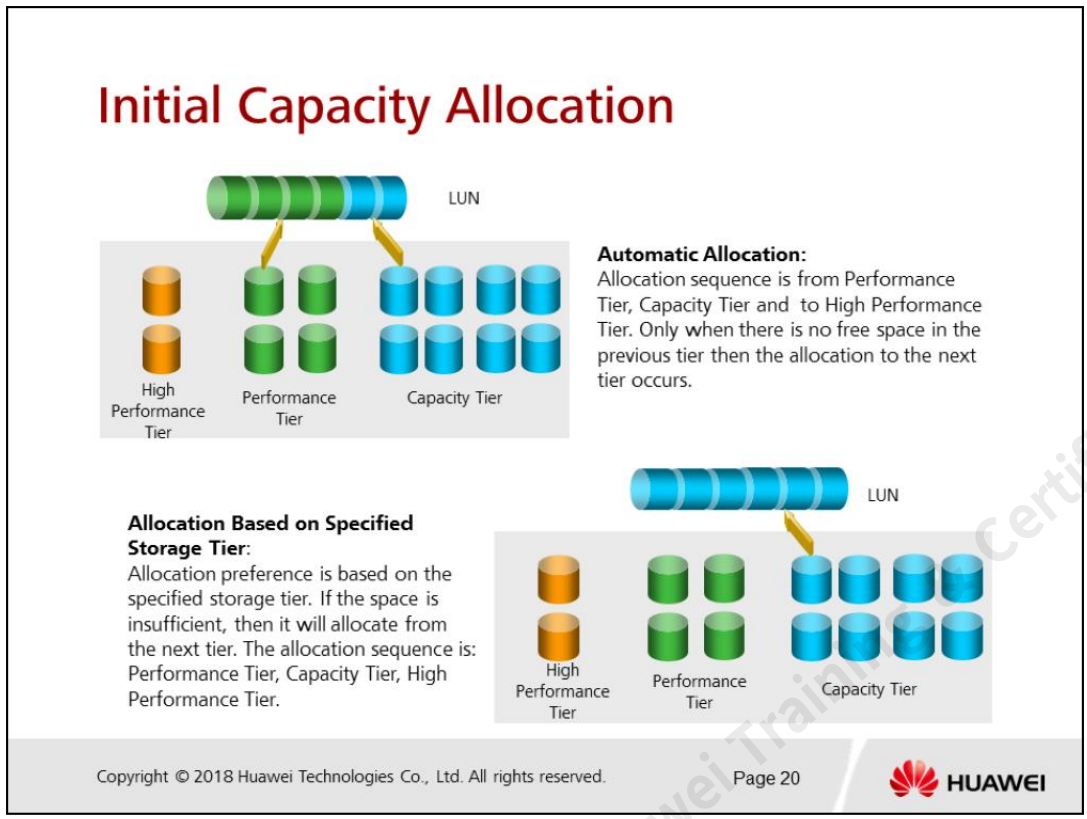
- Storage Pool is the logical combination of one or more storage tiers, and can support up to 3 types of storage tiers. The disks contained within the storage pool determines what kind of storage tiers can be created in the storage pool. A storage pool that consist of a single disk type cannot create different storage tiers, and thus cannot use SmartTier for intelligent data storage management.
- LUNs are created in the storage pool, it distributes the data in the LUN to the different storage tiers within the storage pool before SmartTier functions are applied on those data.
- The SmartTier partitions LUN data by certain granularity. Such granularity is called Data Migration granularity or data block. The data migration granularity is defined when creating a storage pool and the value cannot be changed after being set.
- If a storage pool contains more than one type of disks, the SmartTier can be used to fully utilize the performance of each storage tier. Data migration is an action that promotes data blocks with higher activity and demotes data blocks with lower activity. In this process, storage pool uses blocks as the unit to differentiate the data activity level and then migrates the whole block of data to the other storage tier.

Huawei Training & Certification



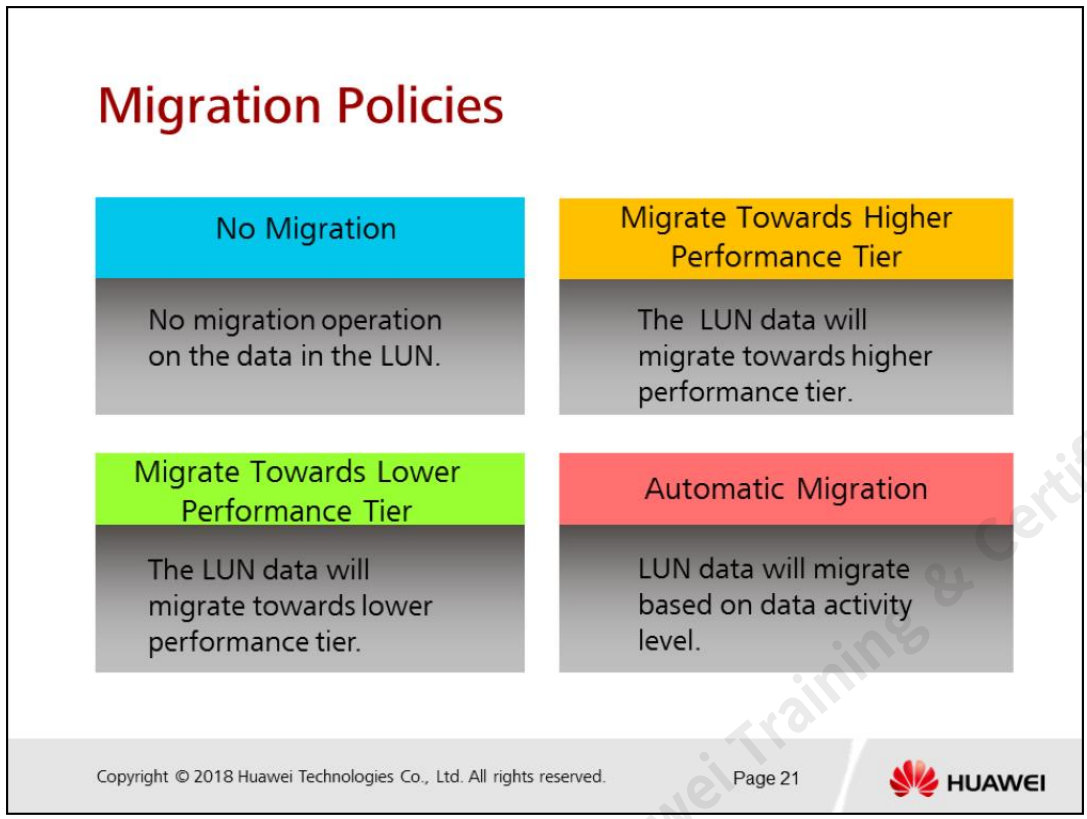
- The initial capacity allocation policy determines the tier to which a new data is written. For example, if the allocate from the performance layer first policy is chosen, new data blocks written into the storage system are placed into the performance tier. There are four initial capacity allocation policies available: automatic allocation, allocate from the high performance layer first, allocate from the performance layer first, and allocate from the capacity layer first. Automatic allocation means that the storage system distributes new data to the performance tier, the capacity tier, and then the high performance tier in sequence.

Huawei Training & Certification



- There are 4 types of initial capacity allocation policy which are: Automatic allocation, Allocate from the high performance layer first, Allocate from the performance layer first, and Allocate from the capacity layer first.

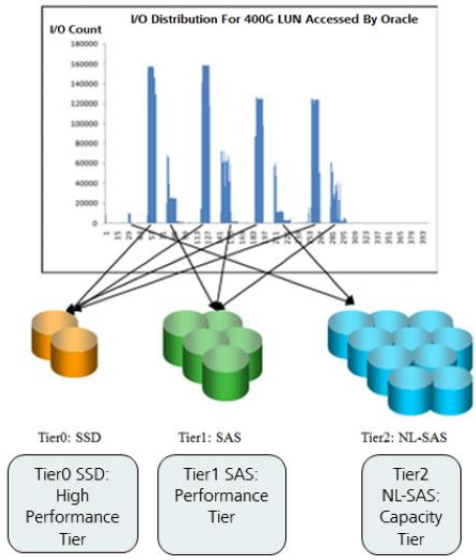
Huawei Training & Certification



- Automatic Migration: When the LUN is configured with Automatic Migration policies, the LUN will do data migration based on data access frequency. The data with higher access frequency migrates towards higher performance tier, and data with lower access frequency migrates towards lower performance tier. If there is no special requirements, automatic migration is recommended for better overall performance.
- Migration Towards Higher Performance Tier: When the LUN is configured with this setting, no matter if the data block access frequency is high or low, the data blocks in the LUN will migrate towards higher performance tier, and consume more high performance tier storage resources. Configure this setting only if you have special requirements for the LUN.
- Migration Towards Lower Performance Tier: When the LUN is configured with this setting, no matter if the data block access frequency is high or low, the data blocks in the LUN will migrate towards lower performance tier. When the LUN is used for archive services, or business services with low performance requirements, this setting can be configured for that LUN.

Huawei Training & Certification

I/O Monitoring and Analysis

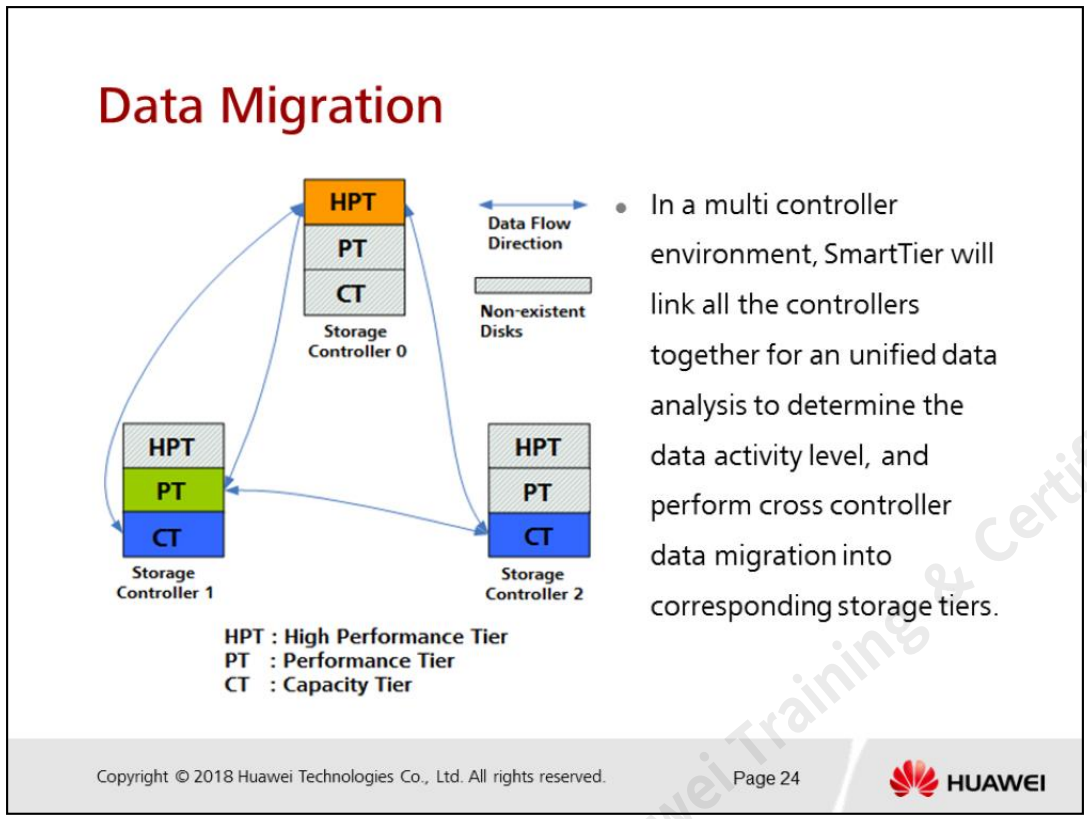


- Based on the I/O statistics information during the monitoring period of the monitored storage pool, such as Read/Write I/O, Average I/O size, the I/O distribution in the storage pool can be determined periodically. Thus, it can be used to determine the data activity level (hot/cold) and then figure out the data blocks to be migrated through analysis.

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

- The is no absolute value for differentiating between hot and cold data. The high performance storage tier has the capacity to store A number of data blocks, while the performance tier has the capacity to store B number of data blocks. During migration, all the LUNs within the storage pool that has automatic migration policies configured will move A number of data blocks with the highest access frequency to the high performance tier and move B number of data blocks to the performance tier. Other data blocks that has lower access frequency will be moved to the capacity tier.
- The main principle is to allows data with the highest access frequency to use the storage media with the best performance level.

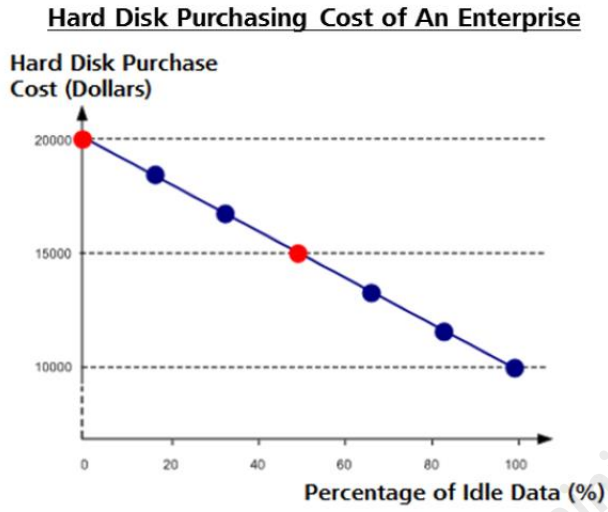
Huawei Training & Certification



- The data migration speed can be configured to High, Medium and Low which correspond to the actual data migration speed of a single storage controller which are 100M/s, 20M/s and 10M/s. The default migration speed is set to Low.
- To lower the impact of SmartTier data migration towards the performance of host business services, the configured data migration speed is the upper limit of the migration speed allowed in the storage system. Data migration always gives priority to the host business services, and uses the idle time of the storage devices to perform data migration, and it adjusts the migration speed based on the real-time pressure on the host business services. When the host business services has lower pressure, the migration speed is adjusted higher whereas when the pressure on the business systems is high, it will lower the data migration speed. No matter how the speed of the migration is adjusted automatically, the actual data migration speed will not exceed the configured value of data migration speed.
    - Low: Suitable for data migration period where there is high pressure on the business systems.
    - Medium: Suitable for data migration period where there is medium pressure on the business systems.
    - High: Suitable for data migration period where there is no business services are running, or when the host and business services are not sensitive towards the performance requirements.

Huawei Training & Certification

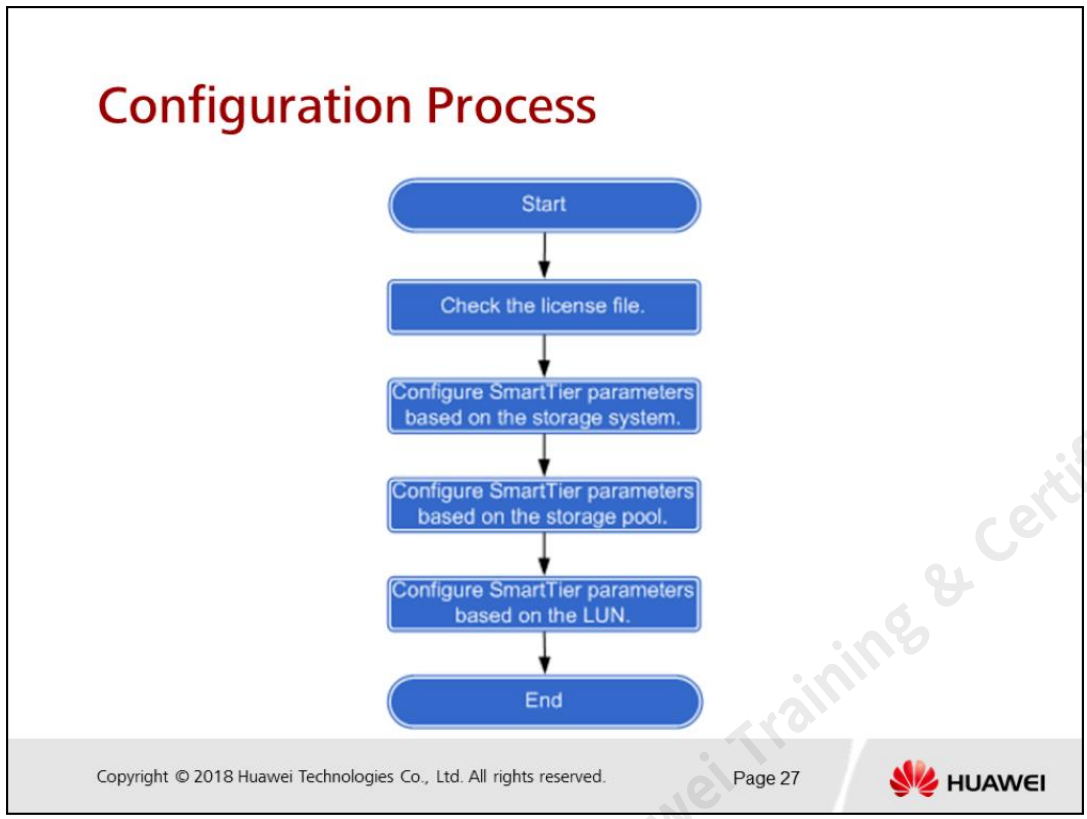
Application Scenario - Lowers TCO



Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

- Due to the free data (Not highly accessed) are stored in the NL-SAS drives, high performance SSD drives could release their storage space and allows more of the highly accessed data to be stored in the SSD drives. SSD drives provides shorter response time and higher IOPS for highly accessed data which in turn increases the performance of the storage system.
- Better utilization rate of the hard disks allows savings on the amount of hard disks purchased, which effectively lowers the TCO(Total Cost of Ownership) of the storage system.
- SmartTier allows better usage of hard disks of different performance levels, which not only increases the overall performance of the storage system but also lowers the cost required for hard disk purchase as it allows better utilization of hard disks with different performance levels with data of different activity levels.

Huawei Training & Certification



- The flowchart above shows the process of configuring the SmartTier feature in Huawei OceanStor V3 storage systems.
- The process for configuring SmartTier within a storage system consists of checking the license file and configuring SmartTier parameters based on storage system levels, storage pool levels, and LUN levels.

Huawei Training & Certification
Huawei Training & Certification
Huawei Training & Certification

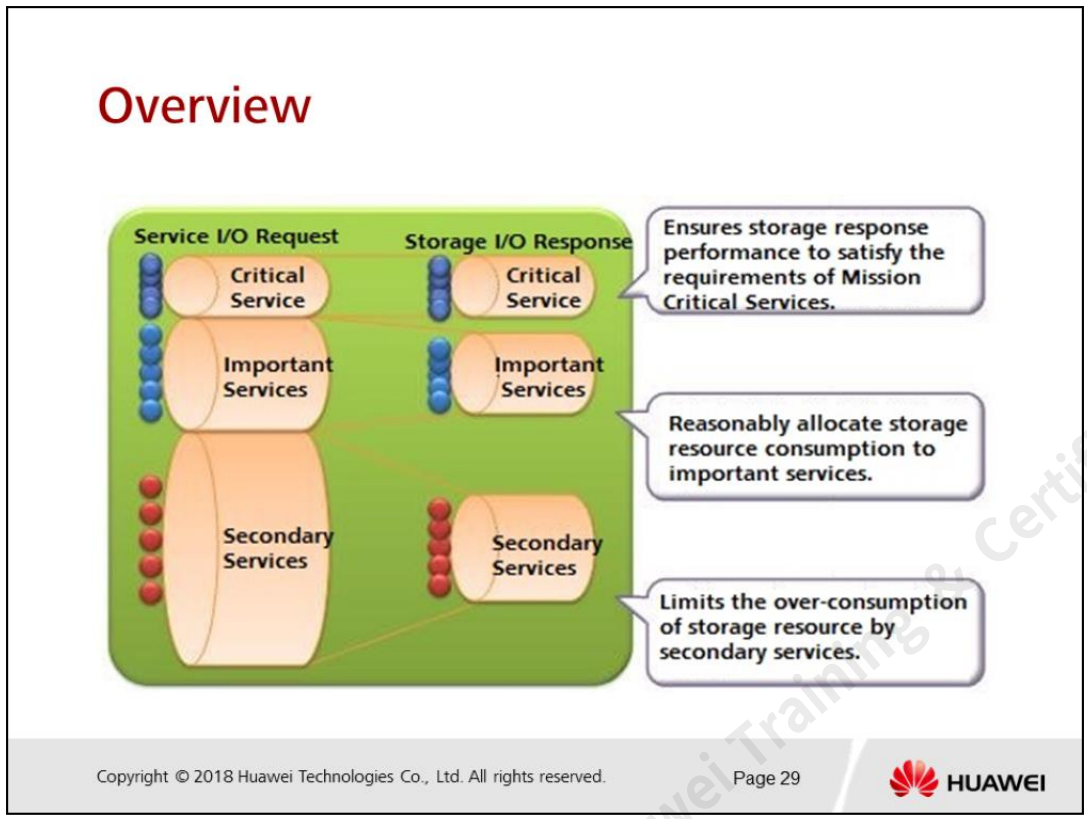
Contents

1. SmartThin
2. SmartTier
3. **SmartQoS**
4. SmartPartition
5. Snapshot
6. SmartQuota

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

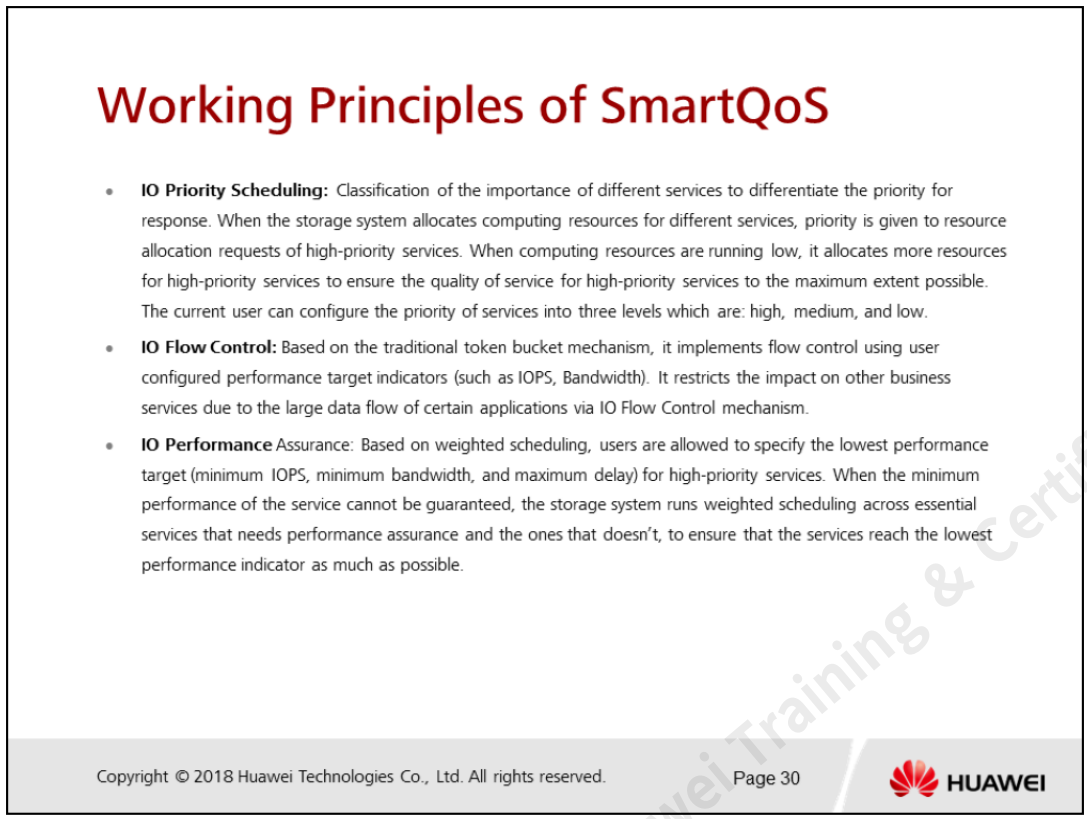
HUAWEI

Huawei Training & Certification
Huawei Training & Certification



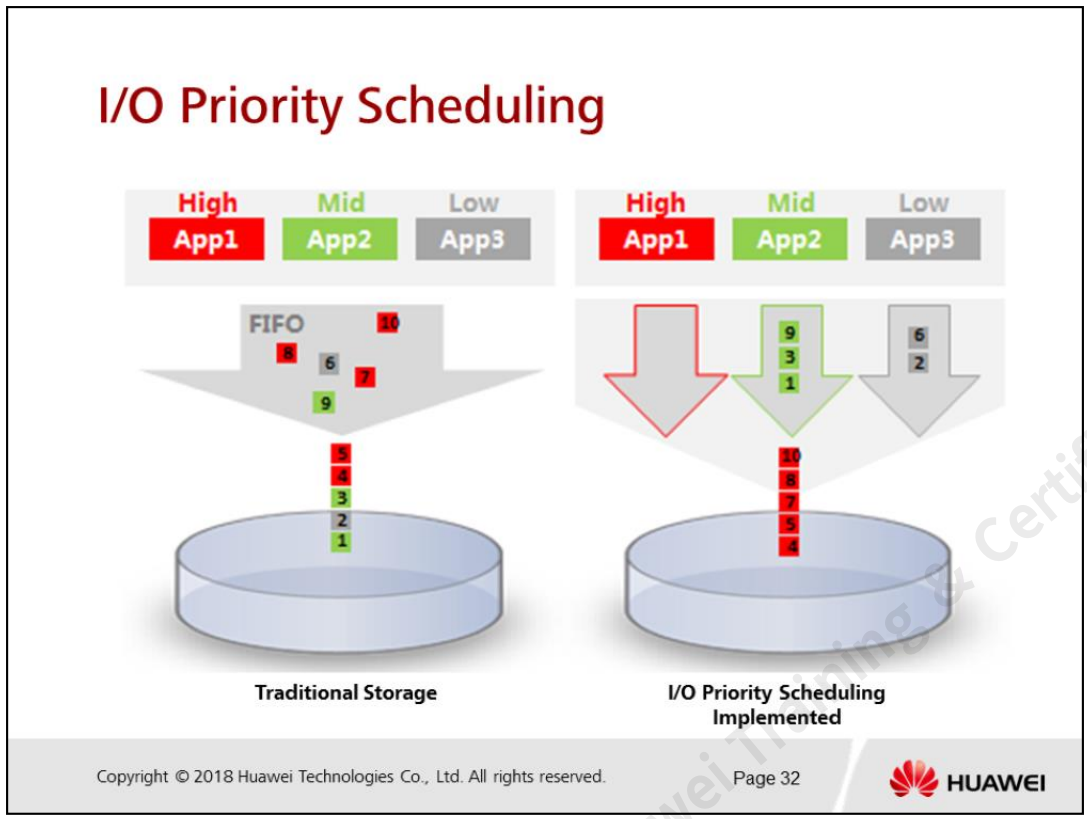
- SmartQoS is a storage Quality of Service(QoS) feature provided in the OceanStor V3 converged storage systems.
- SmartQoS is an intelligent QoS control feature developed by Huawei that dynamically allocates a storage system's resources to meet specific performance goals of certain applications.
- As shows on the diagram above, it can dynamically allocate and adjust the resources within the storage system, and has fine granular control from one end to another end along the data I/O path of the storage system, to ensure and fulfill different QoS requirements in different applications of various importance on the same storage device.

Huawei Training & Certification



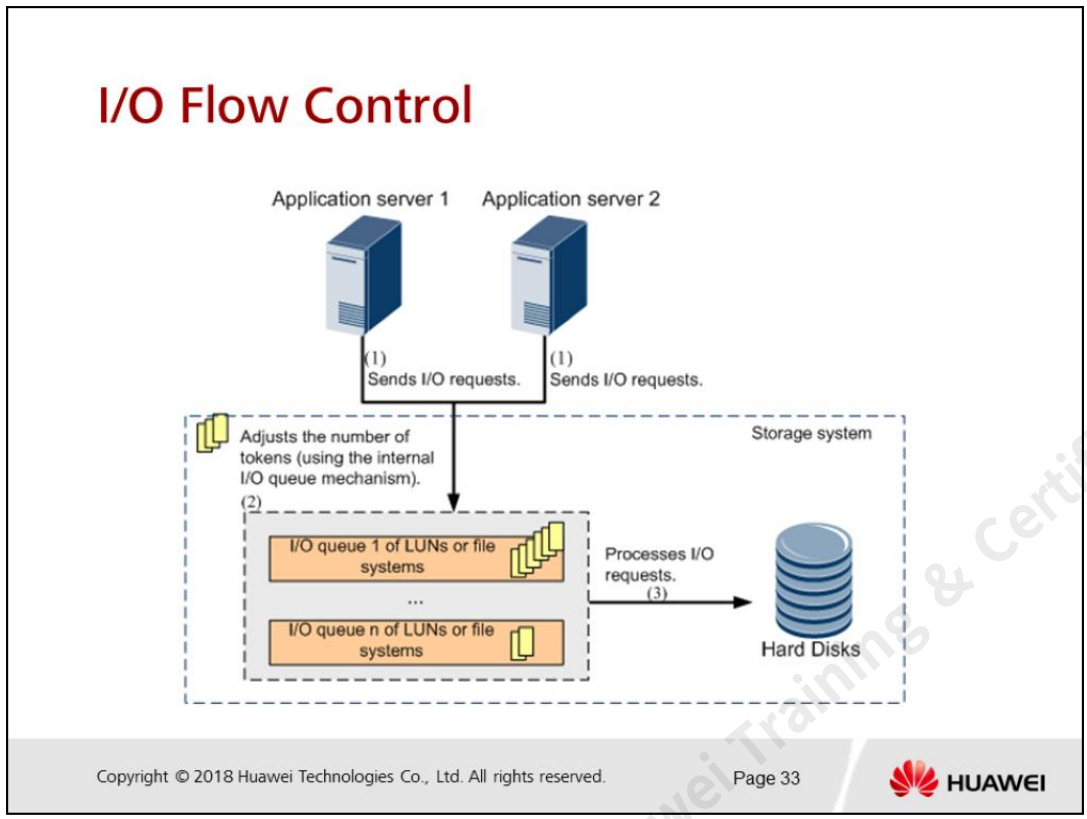
- SmartQoS Features:
  - Full IO Path Priority Scheduling:
    - SmartQoS provides priority scheduling on the whole IO path through computing resource scheduling and allocation control. Compared to the other vendor's priority scheduling policies, the control granularity is much finer.
  - Different QoS policies for different application scenarios:
    - SmartQoS has finer classification of application scenarios, and provides different QoS policies for different application scenarios. When the user has specific performance indicator target requirement, it can use flow control or lowest performance assurance policies to satisfy user requirements. When the user doesn't have specific performance indicator target requirement, it can use the priority mechanism for applications to ensure best performance for high priority applications.

Huawei Training & Certification



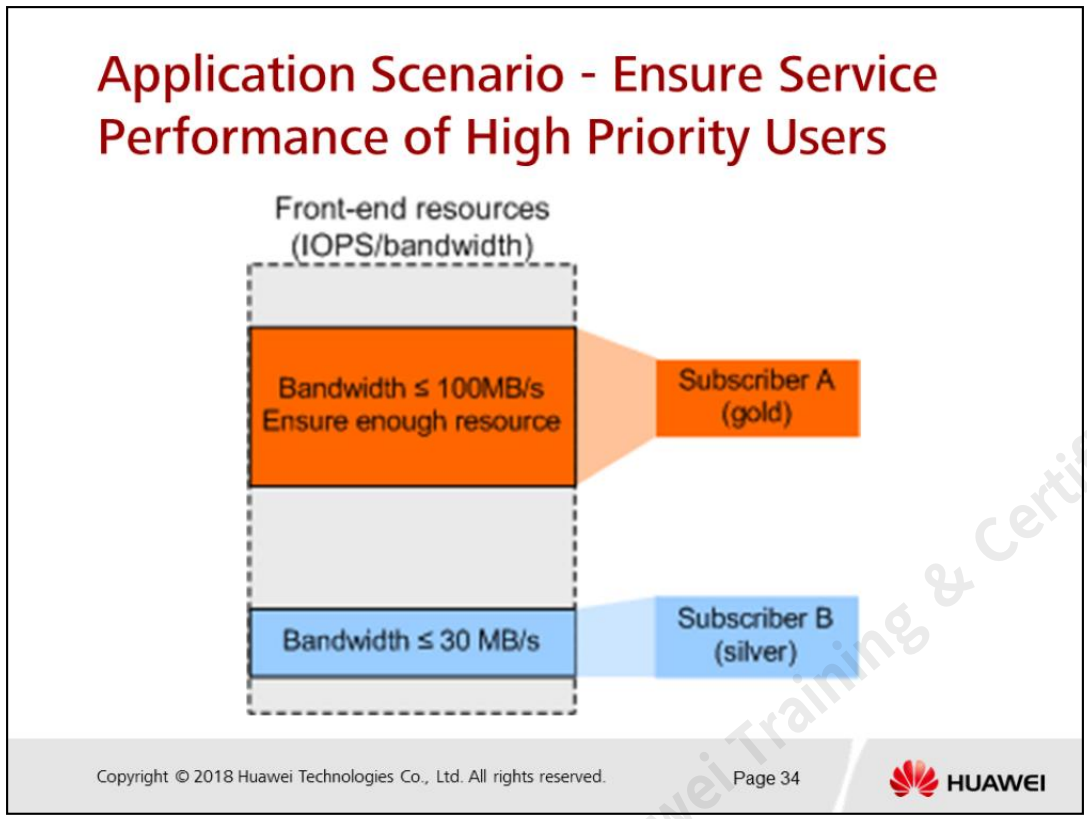
- SmartQoS feature's priority scheduling technology is based on LUN and File System priority levels.
- Hence, each LUN and File System has a priority level properties, and these properties are user configured and saved in the database. When an I/O request is sent from host to storage array, this I/O will obtain its priority level property information from the LUN or File System it belongs to and carries this priority information along the whole I/O path until it reach its destination.
- Users need to configure the priority level when creating the LUN or File System. If it is not configured, the LUN or File System created by have the lowest priority level by default.
- After the LUN or File System is created, its priority level can be manually changed based on user requirements.

Huawei Training & Certification



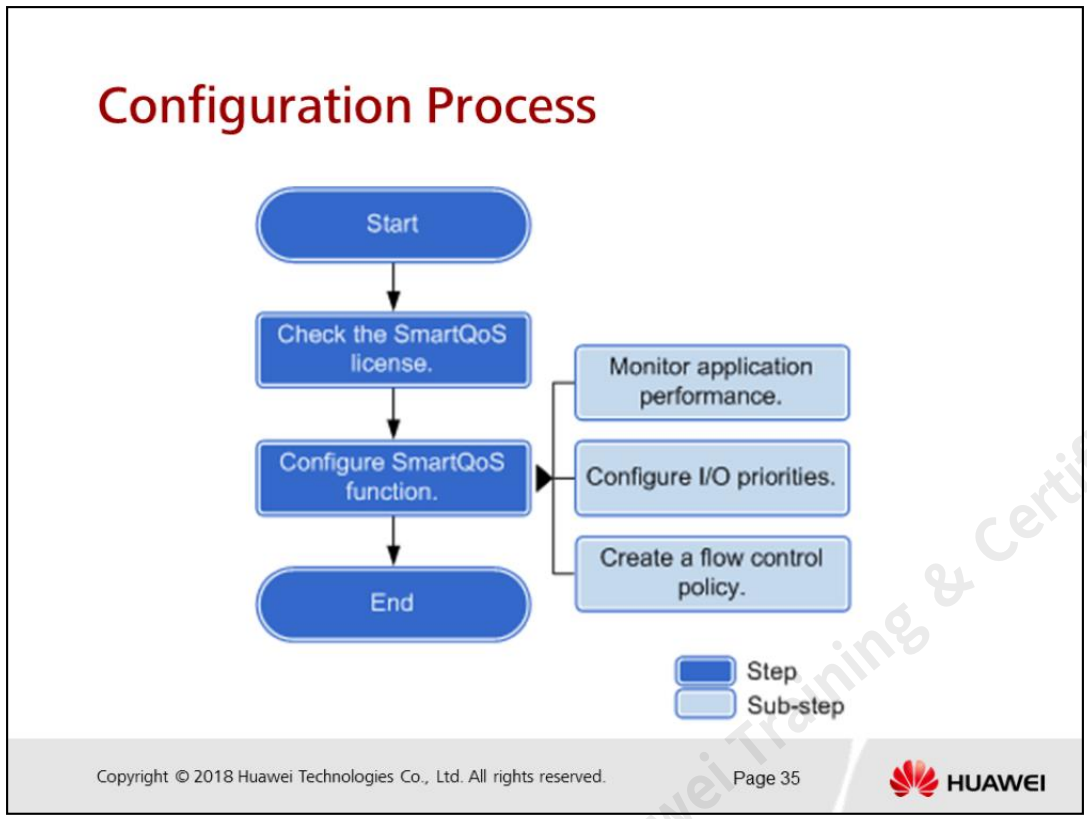
- SmartQoS I/O Flow Control's performance indicator target are achieved through a mechanism of token distribution and control. When users configured the performance upper limit for a certain flow control group, this upper limit will be converted into corresponding tokens. In the storage system, if the users wants to restrict the flow type of IOPS, then each IO corresponds to a token. If the performance indicator is set to Bandwidth, then each sector corresponds to a token.
- Each flow queue has a token bucket, and SmartQoS will periodically put in a certain amount of tokens within the bucket of each flow queue. The amount of tokens depends on the user configured performance upper limit of this flow queue. Thus, if the user configure the performance indicator upper limit as the IOPS=10000, then the token distribution algorithm will set the amount of tokens in the token bucket of the flow queue to be 10000 tokens per second.
- When the flow queue is managed, it will check whether there is sufficient token within the bucket of the flow queue. If there is tokens in the bucket, then it will retrieve an I/O to be processed, and consume the tokens. If there is insufficient tokens in the bucket, then the flow queue must wait till the bucket has enough tokens to be used.

Huawei Training & Certification



- To satisfy the high priority users, we can configure the I/O priority level on the LUN to ensure the operations of the service.
  - Configure user A with higher priority, to ensure that the services for user A will operate as normal.
  - Configure user B with lower priority, to ensure that services for user B will not impact the normal operations of user A's services.

Huawei Training & Certification



- The flowchart above shows the process of configuring the SmartQoS feature in Huawei OceanStor V3 storage systems.
- The SmartQoS configuration process includes checking the SmartQoS license and configuring the SmartQoS function.
- If performance is a major consideration for the services in a storage system, make sure you understand the true performance requirement of each application. The storage system provides a service monitoring function to help you find the performance bottlenecks of each application.

Huawei Training & Certification

Contents

1. SmartThin
2. SmartTier
3. SmartQoS
4. **SmartPartition**
5. Snapshot
6. SmartQuota

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

Huawei Training & Certification
Huawei Training & Certification

Overview of SmartPartition

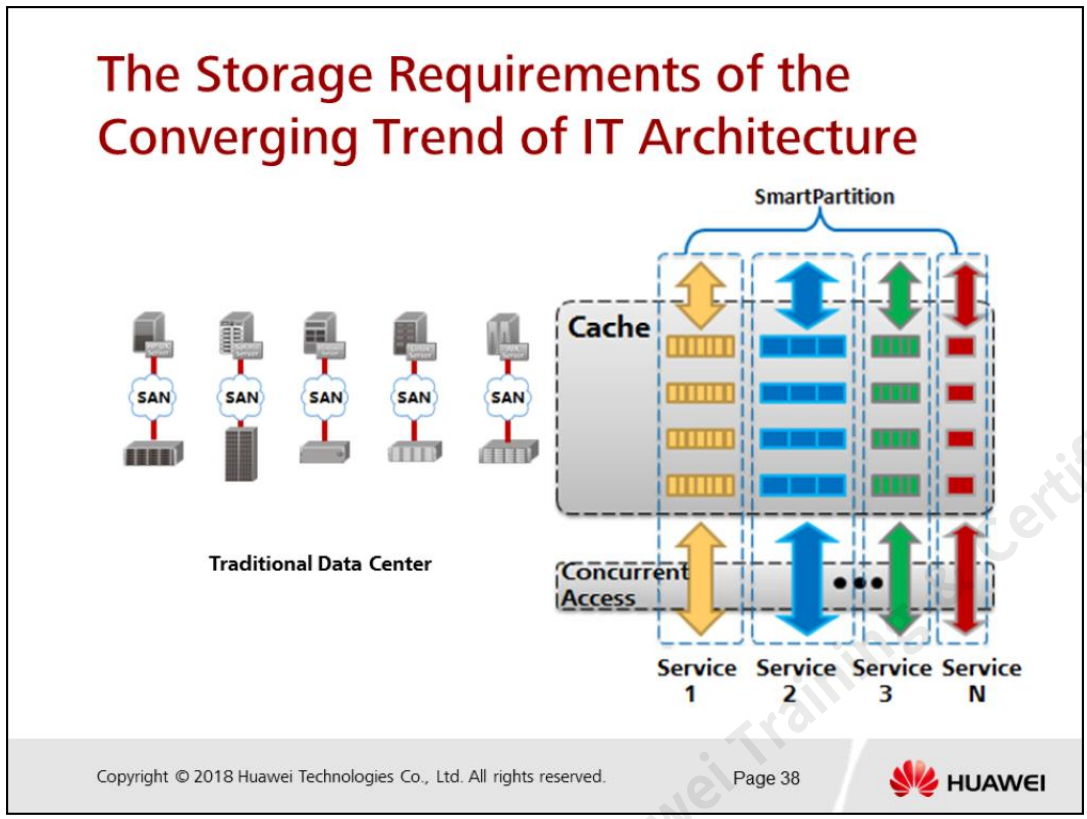
- **SmartPartition** is a smart cache partitioning technology in OceanStor V3 converged storage systems designed to meet the challenge of QoS under the trend of storage convergence. Its core idea is to ensure the performance of key applications by partitioning system core resources. The administrator can configure different sizes of cache partitions for different applications. The system will ensure that the cache resources in the partition are exclusively occupied by the application, and dynamically allocate other resources in different partitions in real time according to the actual conditions of the application, thereby ensuring the applications performance of the applications located in the cache partition.
- **SmartPartition** is essentially a Cache partitioning technology. Cache partitioning is a relatively mature technology in the industry, and it has been relatively long after the time it was initially launched. Current mainstream storage vendors, such as EMC and HDS, have introduced their own cache partitioning features. From a technical point of view, all the major vendors have basically the same approach for cache partitioning where they all divide the limited cache resources into multiple logical areas for management. Thus, the main difference lies in the specific algorithms and adjustment strategies inside the Cache partition.

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

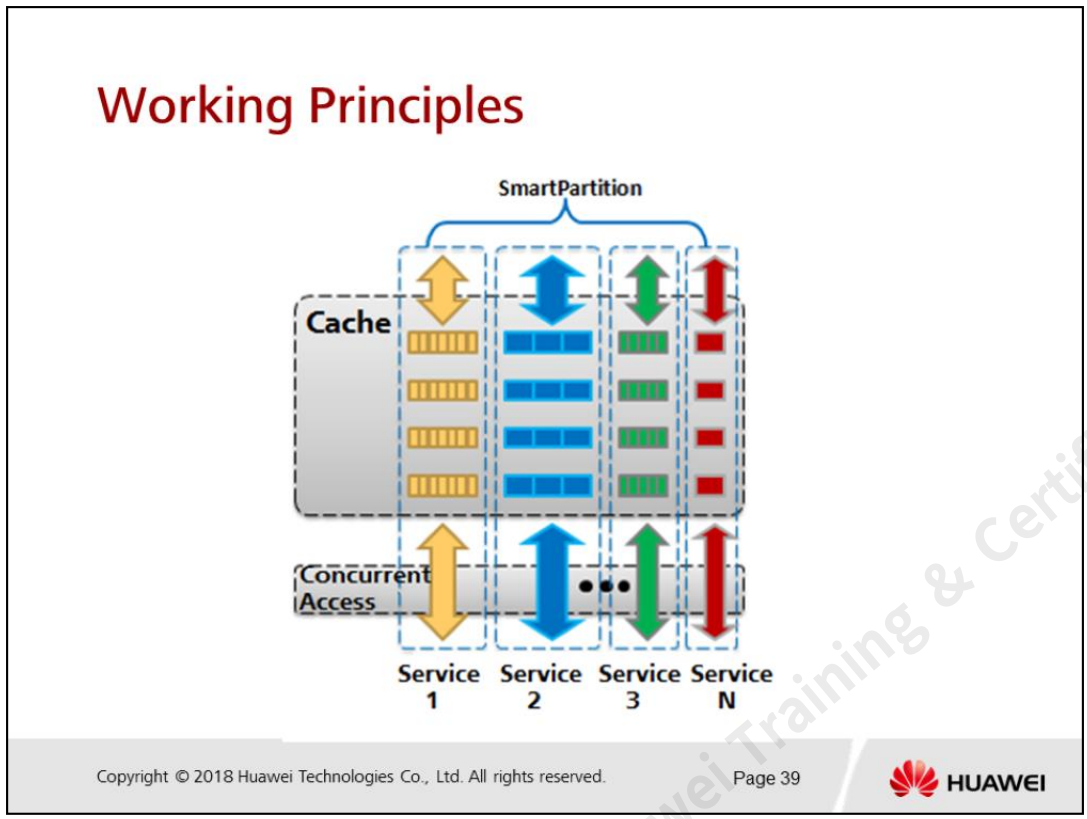
- SmartPartition can even cooperate with the other QoS technologies such as SmartQoS within the OceanStor V3 converged storage systems, in order to achieve better service quality assurance effect.

Huawei Training & Certification



- Typical IT architecture consist of the combination of Compute, Network and Storage. In the traditional "Silo" architecture, different application systems are isolated, and a single storage only needs to face a low amount of application (commonly 5 or below) that consume the storage resources.
- The diagram on the right above represents different kinds of services using a different color. We can see that the I/O mode of these applications are different and has different requirements for cache and concurrent access.
- Different kinds if services are mixed together in the storage system and they compete for concurrent access resources and cache resources, causing unguaranteed quality of service. For example, concurrent access of service N doesn't match the concurrency of its host, causing the data for service N to become more and more in the cache and impact the performance of other services. If service 3 is a key business service, then it means that the business service performance cannot be guaranteed.

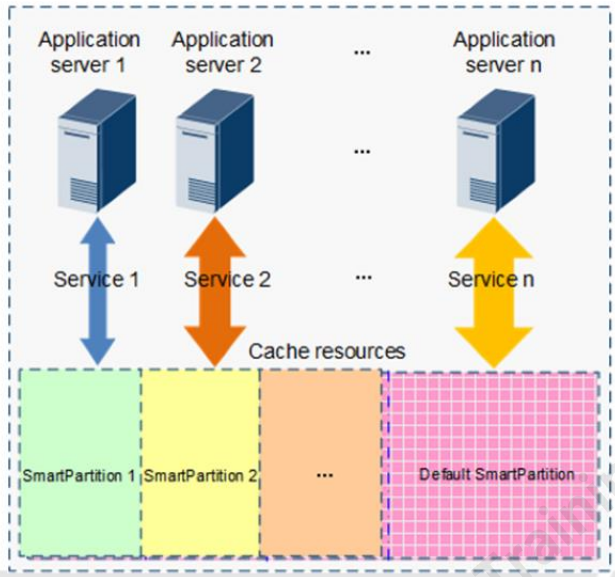
Huawei Training & Certification



- Cache partitioning technologies ensure the service quality of key services through isolating and partitioning the required cache resources for the different services.
- In storage systems, the amount of cache that can be used by one service is the main factor of whether it can impact the service quality of other services.
- The size of the cache that a service can occupy on the storage system is the most important factor that affects the performance of the storage system. A large cache capacity translates into:
  - Higher write combination rate, higher write hit ratio, and better disk access sequence for an application that delivers write I/Os.
  - Higher read hit ratio for an application that delivers read I/Os.
  - For an application that delivers sequential I/Os, its cache capacity does not need to be large but must meet I/O requirements.
  - Higher access rate and better performance for an application that delivers random I/Os.

Huawei Training & Certification

Scheduling of Multiple Applications

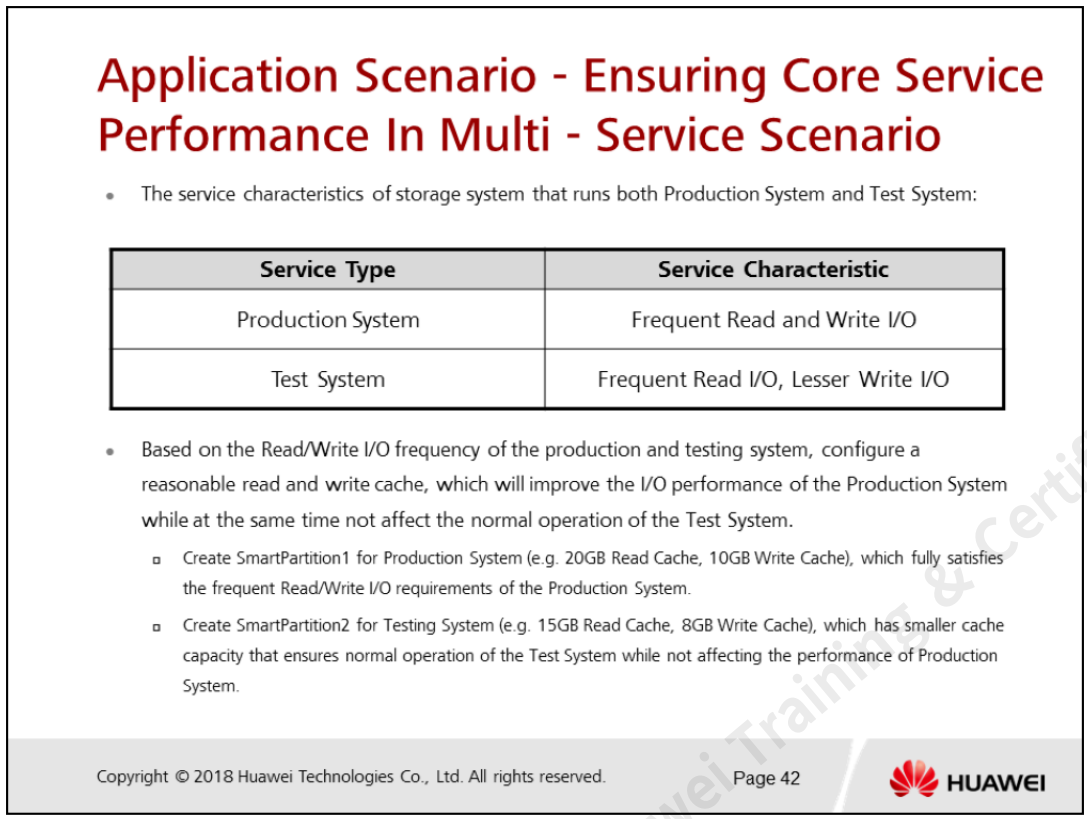


Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

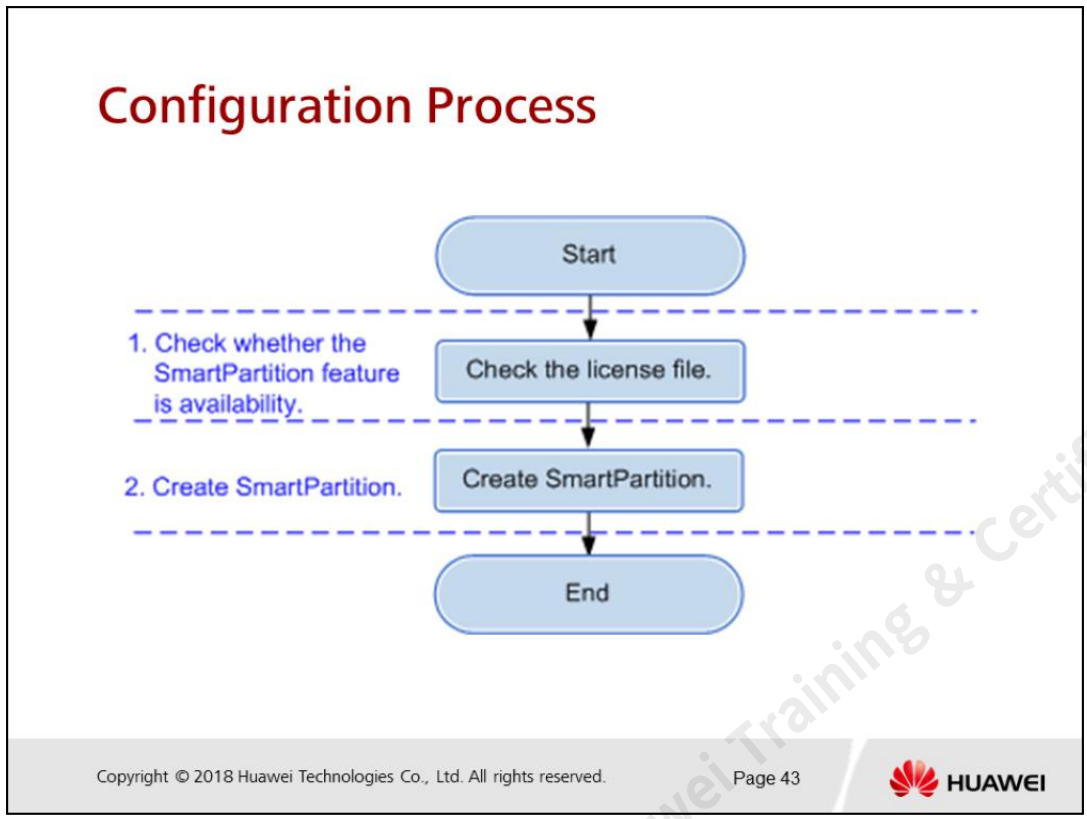
- SmartPartition partitions are created based on service LUNs or file systems. Each SmartPartition partition is exclusively accessed and does not interfere with other partitions.
- After you manually set cache capacity for a partition, SmartPartition will periodically analyze the number of I/Os processed by each partition to achieve an optimal configuration and ensure the QoS of mission-critical services.

Huawei Training & Certification



- As the performance and capacity of a storage system keep growing, multiple applications are typically deployed in the same storage system to simplify the storage architecture and reduce configuration and management costs. However, those applications contend for storage resources, seriously affecting the performance of each service.
- Based on service characteristics, SmartPartition can allocate different sub partitions to different services, so that cache resources in a sub partition are exclusive to the corresponding service, thereby meeting the performance needs of different services and ensuring that mission-critical services run smoothly.
- At the same time, according to service read and write pressure, the corresponding read partition and write partition size can be set separately. If there is a large amount of read I/O traffic, you can allocate more read partition capacity; if more traffic on write I/O services, you can configure a smaller read partition capacity.

Huawei Training & Certification



- The flowchart above shows the process of configuring the SmartPartition feature in Huawei OceanStor V3 storage systems.
- The SmartPartition configuration process includes checking the SmartPartition license and configuring SmartPartition functions.

Huawei Training & Certification

Contents

1. SmartThin
2. SmartTier
3. SmartQoS
4. SmartPartition
5. **Snapshot**
6. SmartQuota

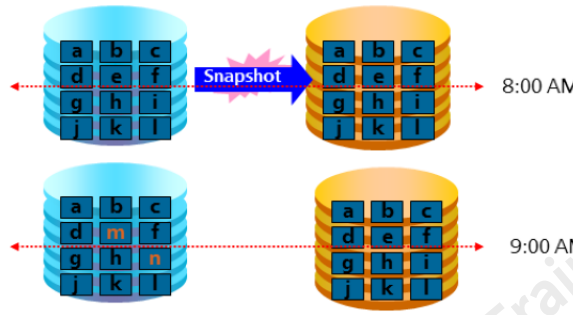
Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

Huawei Training & Certification
Huawei Training & Certification

LUN Snapshot Overview

- Definition: Snapshot refers to consistent data copies from the source data at a certain point of time. After the snapshot is generated, it can be accessed by the host or used as a backup of data at a certain point of time.
- The main features of Snapshot includes:
  - Instant Generation: Storage system can generate a snapshot in seconds, and obtain a consistent copy of the source data.
  - Takes up less storage space: The generated snapshot is not a full complete copy of the physical data, it does not takes up much of the storage space. Thus, even if the source data is huge, snapshots only takes up very less amount of storage space.



Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

- Snapshot not only can quickly generate a consistent copy of the source volume at a certain point of time, it also provide the mechanism to restore the data in the source volume.
- When the data within the source LUN was intentionally deleted, destroyed or infiltrated by virus, we can quickly restore the source LUN using snapshot rollback to restore the data in the source LUN to the time when the snapshot was activated, in order to reduce the loss of data.

Huawei Training & Certification

Related Concepts

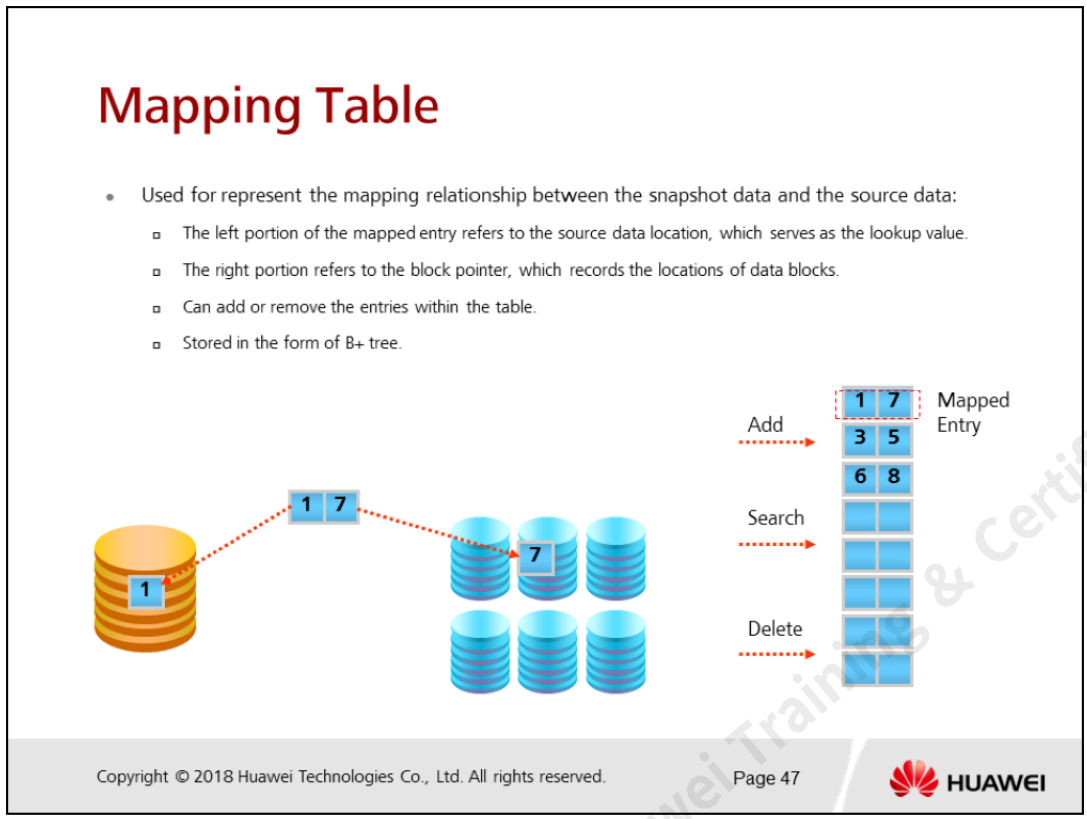
Terms	Definition/Description
source volume	The volume where the source data resides for snapshot operation, it is represented to the host as LUN. Source Volume includes Meta Volume and Data Volume. -Meta Volume: Records the location of source data. -Data Volume: Records the data stored in the source volume.
COW data space	COW data space is dynamic storage allocated from the storage pool where the source LUN resides. This storage is used to store COW data following snapshot generation and activation. All snapshot LUNs of a source LUN share a single COW data space.
snapshot volume	It is the logical data copy generated after the snapshot is created from the source volume. It is represented as a LUN to the users.
snapshot rollback	Replicate the data from the snapshot LUN to the source LUN, which restores the source LUN data to the point where the snapshot was made.
mapping table	It is used to record the changes of data in the source and snapshot LUN at a certain point of time and the changed storage location of data. It is divided into shared and private mapping tables.
inactive	It is a status of the snapshot. In this status, snapshot are unable to be used, and requires activation operation before snapshots can be used as normal.

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

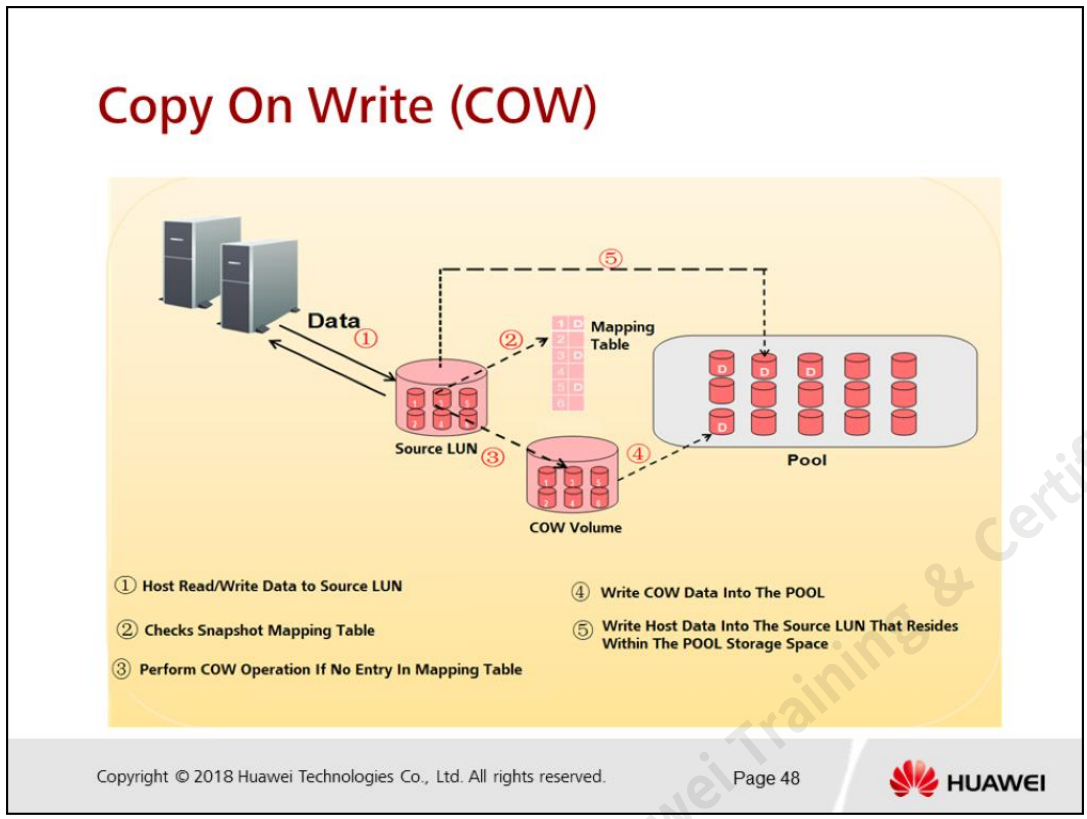
- The technical terms explained on the table above are quite commonly used to explain the different steps and status of Snapshot feature.

Huawei Training & Certification



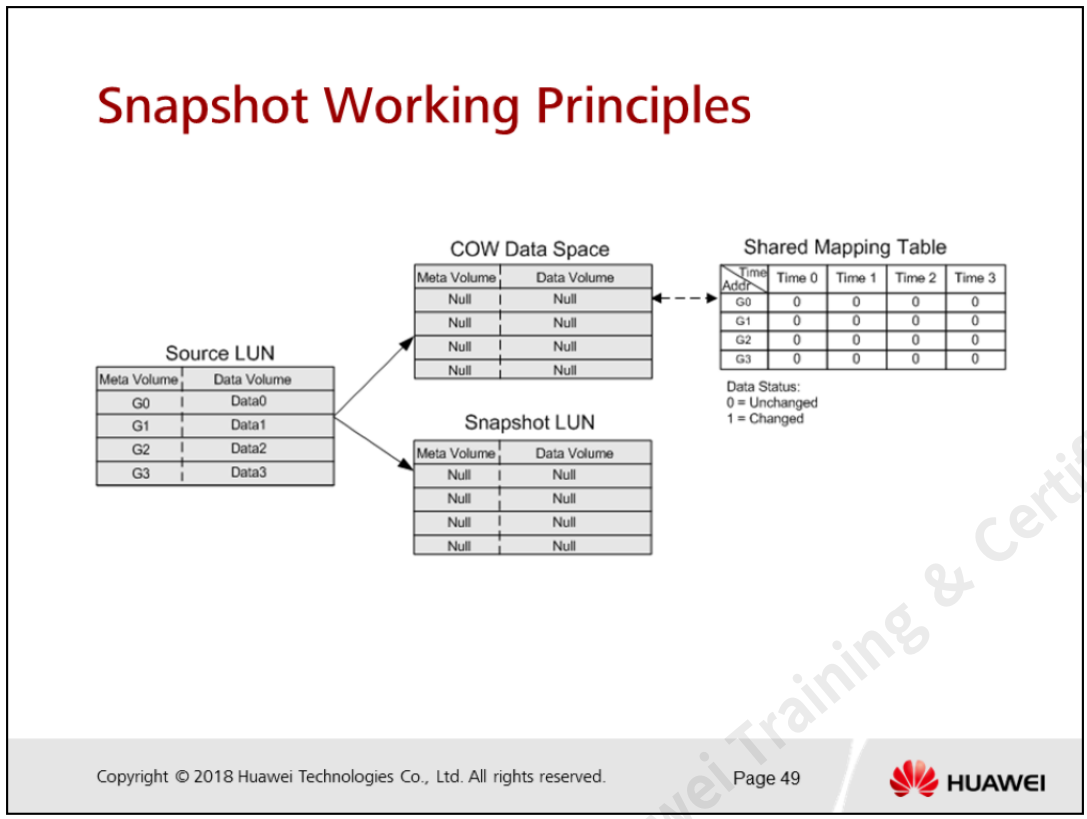
- Mapping table is used to represent the relationship between the snapshot data and actual data.
- Mapping table is divided into 2 types: Private mapping table and Shared mapping table. Both have the same working principles, but the main difference lies where the private mapping table records the changes of data in the snapshot LUN while the shared mapping table records the changes of data in the source LUN.

Huawei Training & Certification



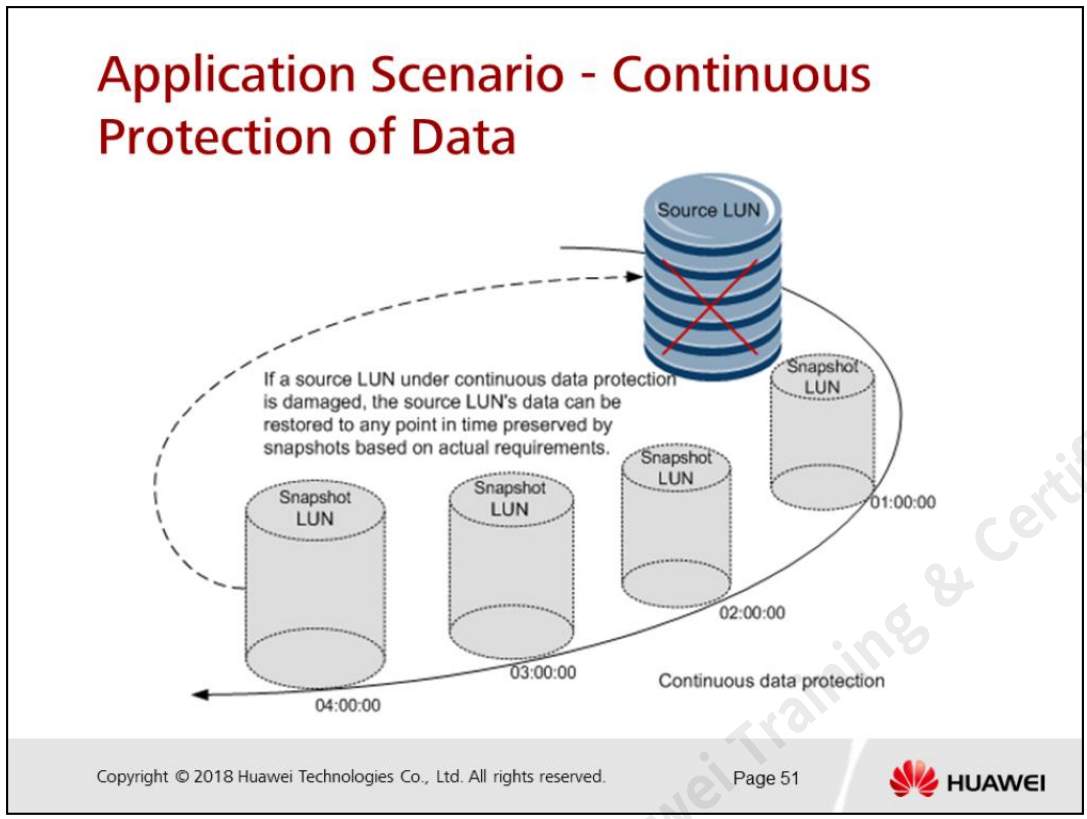
- When the snapshot is activated (snapshot point-in-time), the read/write operation of the hosts are as following:
  - After snapshot is activated, data is written to the source LUN.
  - It will first check the mapping table, if there is no mapping entry for the corresponding location, then it will perform copy-on-write(COW). Afterwards, it will record a backup copy of the source LUN data in the mapping table. If the mapping entry already exists, then it will directly overwrite the location on the source LUN.
  - COW reads the data in the corresponding location and copies it to the COW data space.
  - COW data space and the source LUN data space are distributed in the same POOL, which means that writing to the COW data space is the same as writing to the POOL data space.
  - After COW is completed, write the data from the host to the POOL space where the source LUN resides.
- Copy on write (COW) is a technology that saves changed data to a source LUN. If snapshots are active, when an application server writes data to the source LUN, the storage system performs the following steps:
  - Copies the to-be-replaced original data (COW data) to a COW data space.
  - Changes the mapping relationship of the COW data.
  - Records the new location of the COW data in the COW data space.
  - Writes the new data to the source LUN.

Huawei Training & Certification



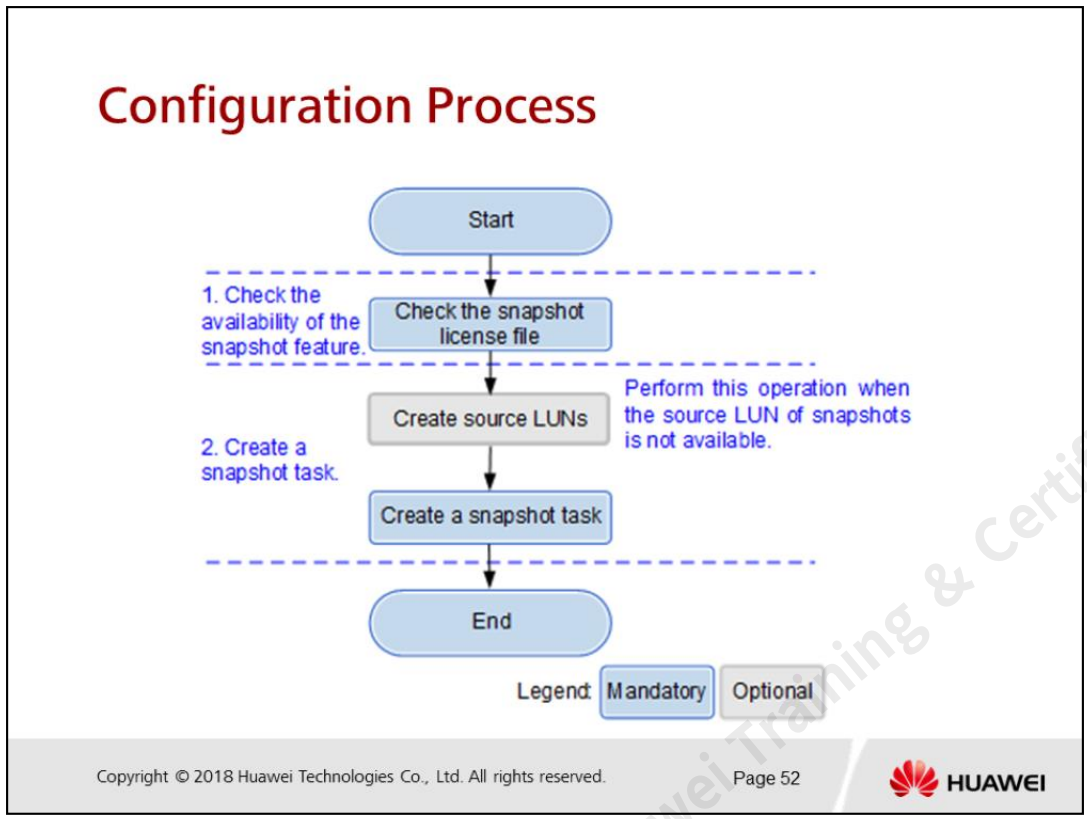
- Following the snapshot creation and activation, a data duplicate of the source LUN is generated. The storage system dynamically allocates COW data space from the storage pool where the source LUN resides, and generates a snapshot LUN.
- Due to the fact that no data is written to snapshot or source LUNs, thus no data is recorded to the COW meta and data areas or to the snapshot meta and data volumes. Diagram above shows the initial state of a snapshot after creation and generation.
- If an application server attempts to write data to the source LUN of an activated snapshot, the storage system performs the following before processing the write request:
  - Uses the COW mechanism to copy COW data to the COW data space.
  - Changes the mapping relationships in the mapping table.
  - Writes new data to the source LUN
- In a snapshot period, COW is performed only once for data of each location. Following that, newly written data overwrites existing data. For example, if the corresponding value of DataX is recorded as 1 in the mapping table, COW has been performed for DataX. For data write requests made in the future, the storage system will write data directly to the source LUN without copying DataX to COW data space.

Huawei Training & Certification



- Snapshots can be used directly as point in time backup copies of data. The usage of snapshot allows quick recovery of data in the following scenarios:
  - Virus infection.
  - Human error in operation.
  - Malicious tampering.
  - Data corruption due to system downtime.
  - Data corruption due to BUG in applications.
  - Data corruption due to BUG in storage systems.
  - Damaged storage media (Only snapshots based on split mirror technology can recover data in this case).

Huawei Training & Certification



- The flowchart above shows the configuration process for Snapshot feature in Huawei OceanStor V3 storage systems
- Configuring a snapshot task includes checking the license file for the snapshot feature, creating LUNs, and creating snapshots.
- Steps of configuring a snapshot task:
  - Check the availability of the snapshot feature by checking the snapshot license file.
  - Create a snapshot task by either optionally create a source LUN for snapshot activation or create a snapshot task following the wizard instruction on a existing source LUN.

Huawei Training & Certification

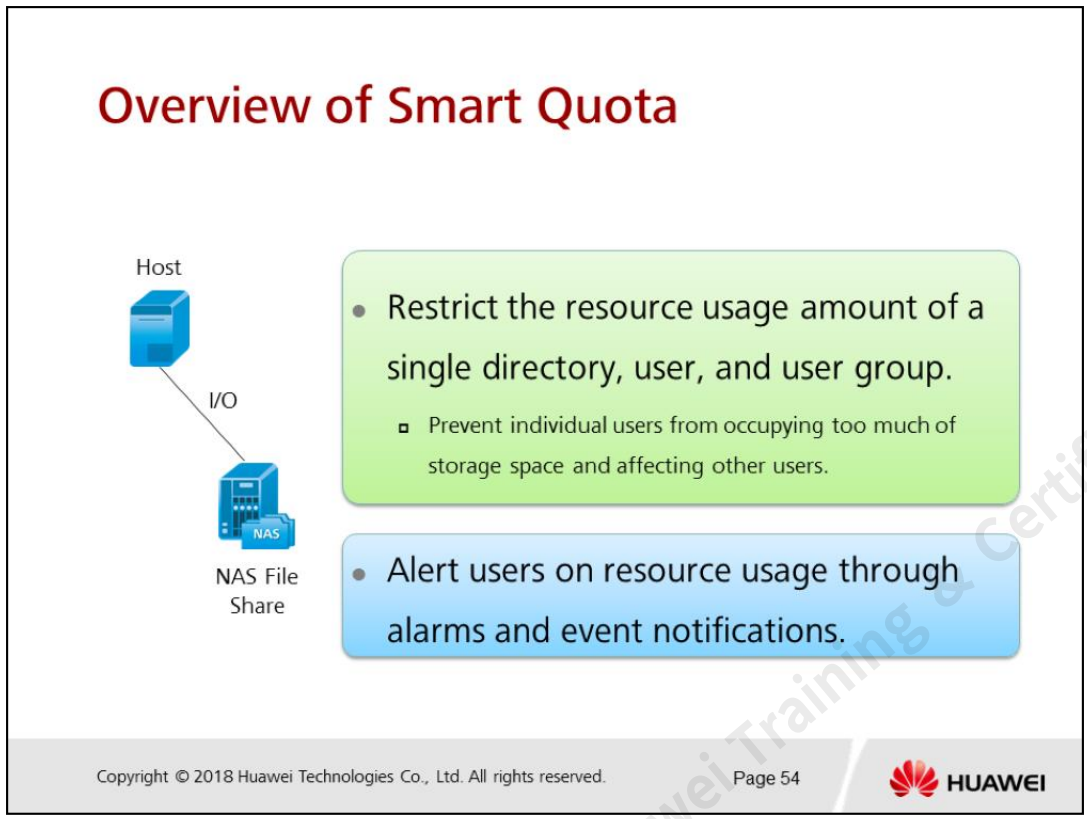
Contents

1. SmartThin
2. SmartTier
3. SmartQoS
4. SmartPartition
5. Snapshot
6. **SmartQuota**

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

Huawei Training & Certification
Huawei Training & Certification



- With the development of virtualization and cloud computing technologies, IT systems faces the challenges of resource utilization efficiency and how to effectively manage them. In a typical IT storage system, as long as there are usable storage resource (disk space), then it will be utilized by the users until it has depleted. From an economical point of view, we need a method to control the usage of storage resources and its growth.
- In the NAS file service environment, resources are usually provided to departments, organizations or users in the form of shared directories, and each of them has unique storage requirements and limitations. Thus, systems need to allocate resources and restrictions to each user based on shared directory or local conditions. SmartQuota is the technology developed to solve this requirement. It can be used to perform resource allocation and restrictions based on directories, users, and user groups.
- The main role of SmartQuota is to make things easier for system administrators to control the usage of the storage resources by the users (including directories, users and user groups) and to limit the disk space usage of specific users to avoid the problem of over-consumption of storage resources by some users.

Huawei Training & Certification

Related Terms

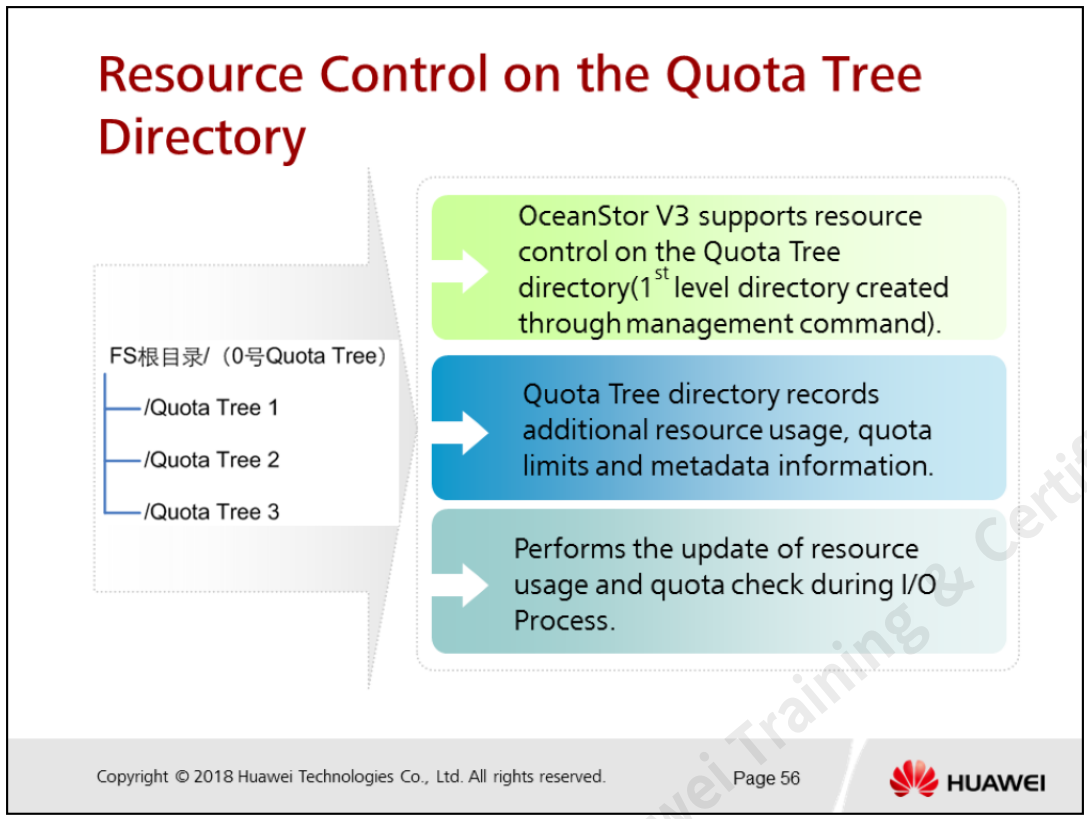
Terms	Descriptions
Quota Tree	Quota tree is a special directory of a file system. In a quota tree, you can set directory quotas, file limit quotas and manage storage space occupied by files in the directory.
Root Quota Tree	The root quota tree is actually the root directory of the file system. User quotas and group quotas can be set on the Root Quota Tree to limit the resources available to users.
Soft Quota	Soft quotas refers to a quota value that can be configured. Once the resources used by users exceed this value, the system generates alarms. Alternately, if the used resources are changed from higher to lower than the soft quotas, the alarms are eliminated.
Hard Quota	Hard quota refers to the maximum value of the resources that can be consumed by the users. The amount of resource used by users cannot exceed this value.

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

- The table above explains the key terms involved in the SmartQuota feature. It is important to differentiate the key differences between the soft and hard quota when configuring the resource limitation for users using the SmartQuota feature.

Huawei Training & Certification



- In the SmartQuota feature, the Quota Tree is an important concept. In simpler words, Quota Tree is the first level directory of the file system. From the management viewpoint, Quota Tree is not just a directory, but also a configuration entity. This means that, Quota Tree can only be created, deleted or modified through the management terminal (Command Line or GUI Management Interface), and it cannot be modified through the client terminal or hosts. Additionally, it serves as the entity to configure the directory quota, user quota, and user group quota, and these quotas can only be configured on the Quota Tree itself. In the following, lets summarize the differences between the Quota Tree and normal directories:
- Quota tree can only be created, deleted, renamed through command line or GUI management interface by the system administrator. System administrators can only delete empty quota trees.
- Quota tree can be shared using protocols, but no rename or deletion is allowed when the quota tree is shared.

Huawei Training & Certification

Resource Usage in a Quota Tree

Quota Tree 1

| ---- confFile.conf (2MB, usr 3, grp 5)

| ---- run.dat (1MB, usr 3, grp 8)

| ---- doc (0B, usr 4, grp 8)

| | ---- study.doc (5MB, usr 7, grp 9)

|

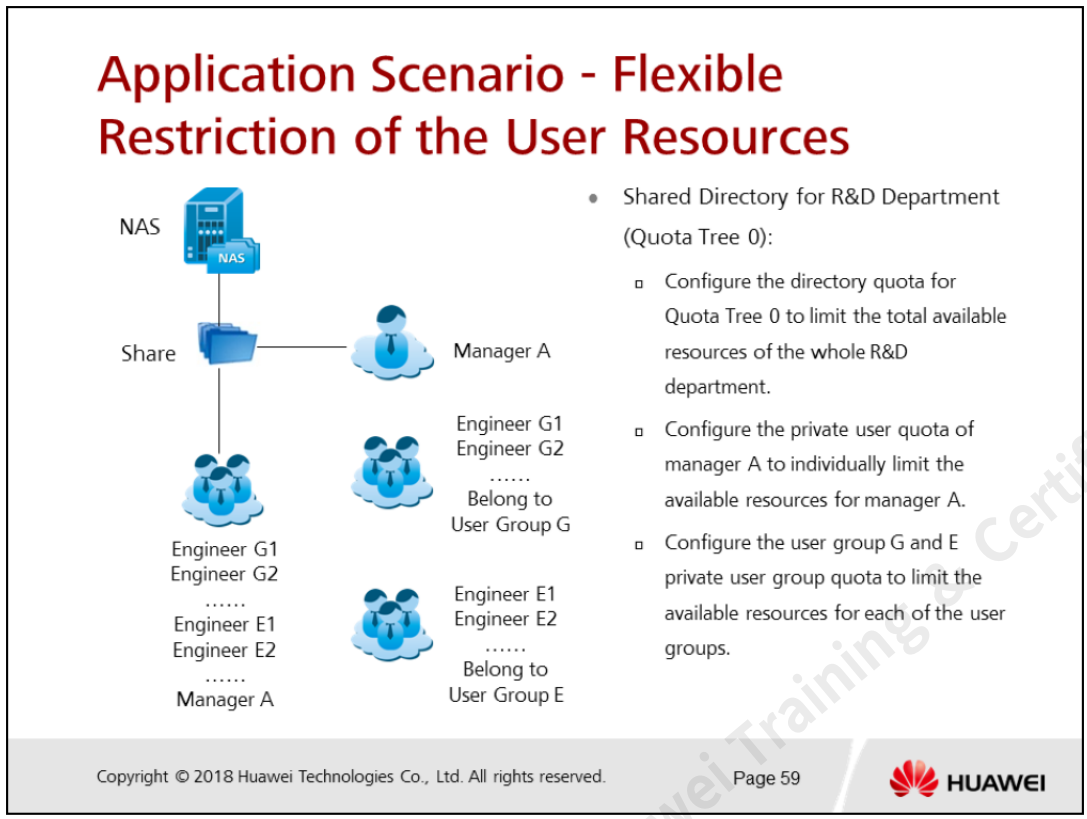
Quota Tree 1	Capacity	Number of Files
Directory	8MB	4
User		
3	3MB	2
4	0	1
7	5MB	1
User Group		
5	2MB	1
8	1MB	2
9	5MB	1

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

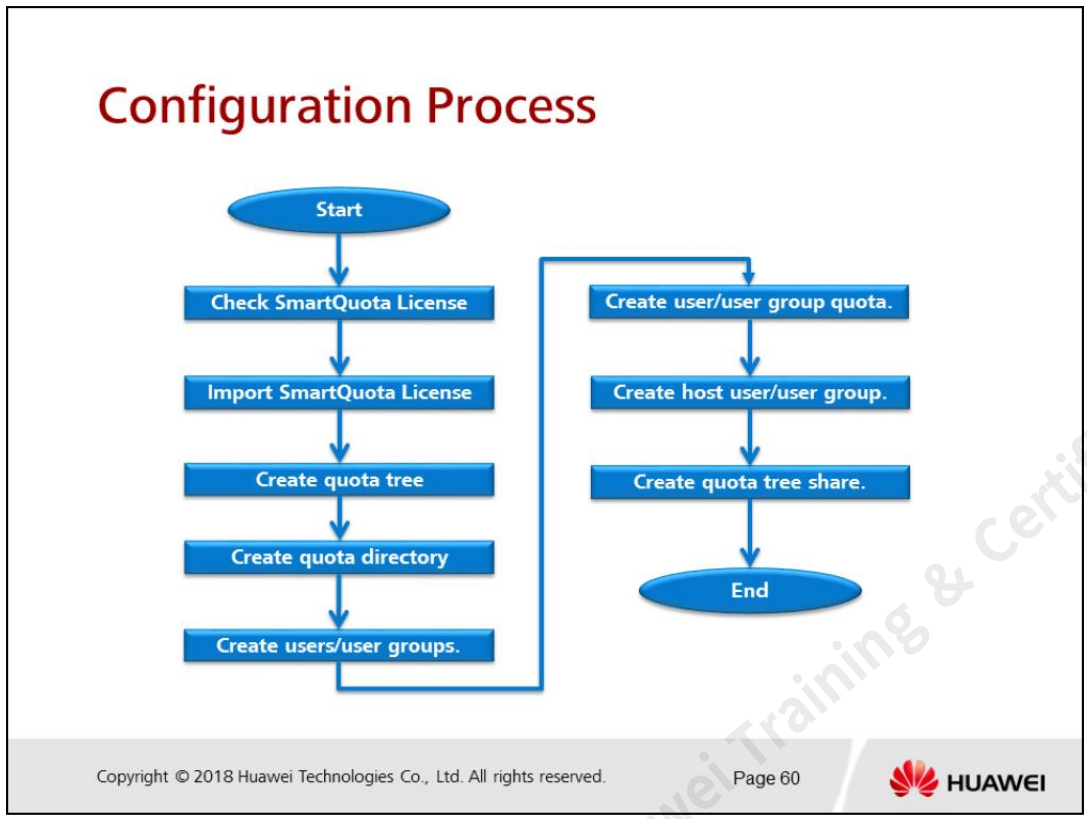
- Resource usage of directories (Statistics of directory quotas)
  - The total amount of the storage capacity and number of files for all the files within the directories of Quota Tree.
- Resource usage of user/user groups (Statistics of user/user group quotas)
  - The total amount of the storage capacity and number of files for all the files in the directory created for the specific user/user group within the directories of Quota Tree.

Huawei Training & Certification



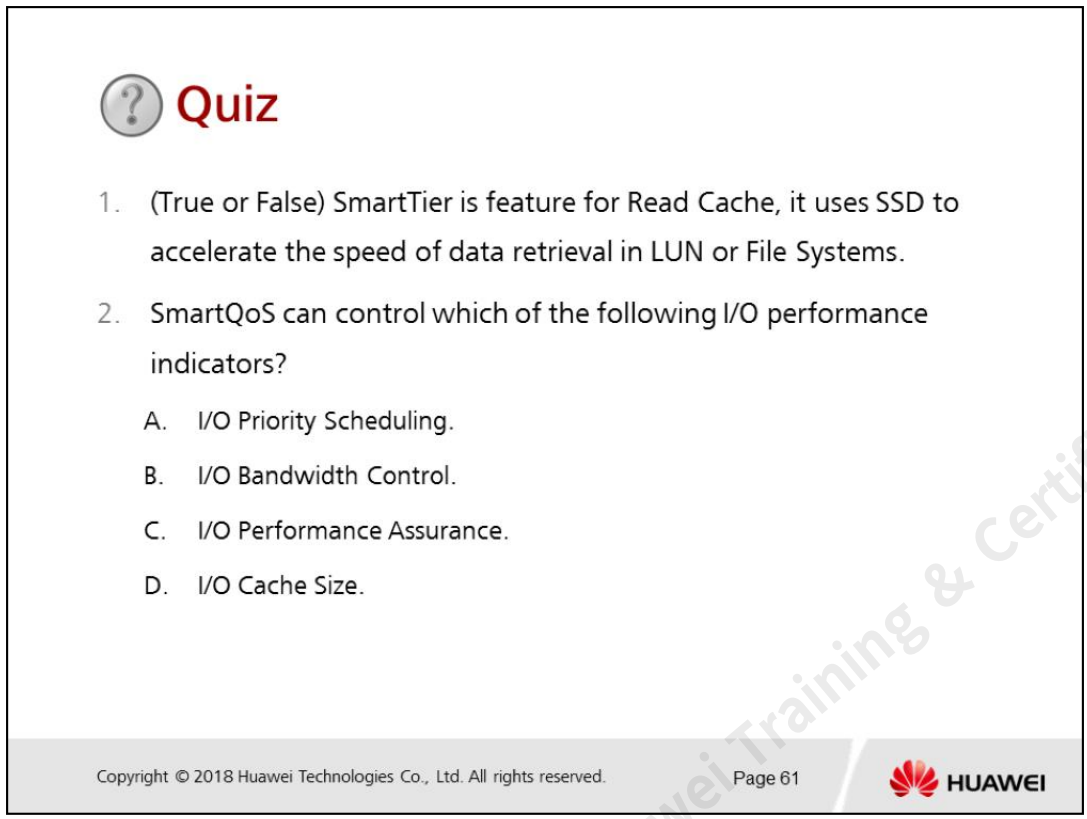
- SmartQuota feature allows flexible restriction or limitation on the user resources within department by implementing 3 different types of quota which are :
  - Directory Quota.
  - User Quota.
  - User Group Quota.
- Within a single Quota Tree, the system administrators can configure the directory quota to limit the total amount of resources available for the corresponding department, then configure the user quota and user group quota for each user and user groups in the department to flexibly limit the amount of resource usage of each users and groups.

Huawei Training & Certification



- The flowchart above shows the configuration process of the SmartQuota feature in Huawei OceanStor V3 storage system.
- The configuration process involves checking the license for the feature availability, if it is not available then the license must be imported to activate this feature.
- Subsequently, create the quota tree and the quota directories and configure the users/user groups and their quotas respectively.
- Finally, create the hosts user/user groups and link them together with the quota directories through sharing protocols.

Huawei Training & Certification



- Answers:
  - False.
  - ABC.

Huawei Training & Certification
Huawei Training & Certification

Summary

- This module introduced the common advanced features used in storage systems which are:
  - SmartThin
  - SmartTier
  - SmartQoS
  - SmartPartition
  - Snapshot
  - SmartQuota

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

Huawei Training & Certification
Huawei Training & Certification

Thank You

www.huawei.com

Huawei Training & Certification
Huawei Training & Certification

09 Overview of Business Continuity Solutions

www.huawei.com

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

Huawei Training & Certification
Huawei Training & Certification

Foreword

- This module introduces:
  - The importance and challenges faced by business continuity.
  - The definition and measuring standards of business continuity.
  - The types of business continuity solutions.
  - The panorama of Huawei Disaster Recovery solutions.

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

Huawei Training & Certification
Huawei Training & Certification

Objectives

- Upon completion of this module, you will be able to:
  - Understand the importance and the challenges faced by business continuity.
  - Understand the definition of business continuity, standards and the relationship between the costs and risks in ensuring business continuity.
  - Understand the common solutions for business continuity.
  - Understand the panorama of Huawei business continuity solutions.

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

Huawei Training & Certification
Huawei Training & Certification

Contents

1. **Challenges of Business Continuity.**
2. Definition of Business Continuity.
3. Overview of Business Continuity Solutions.
4. Product Panorama of Huawei Business Continuity Solutions.

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

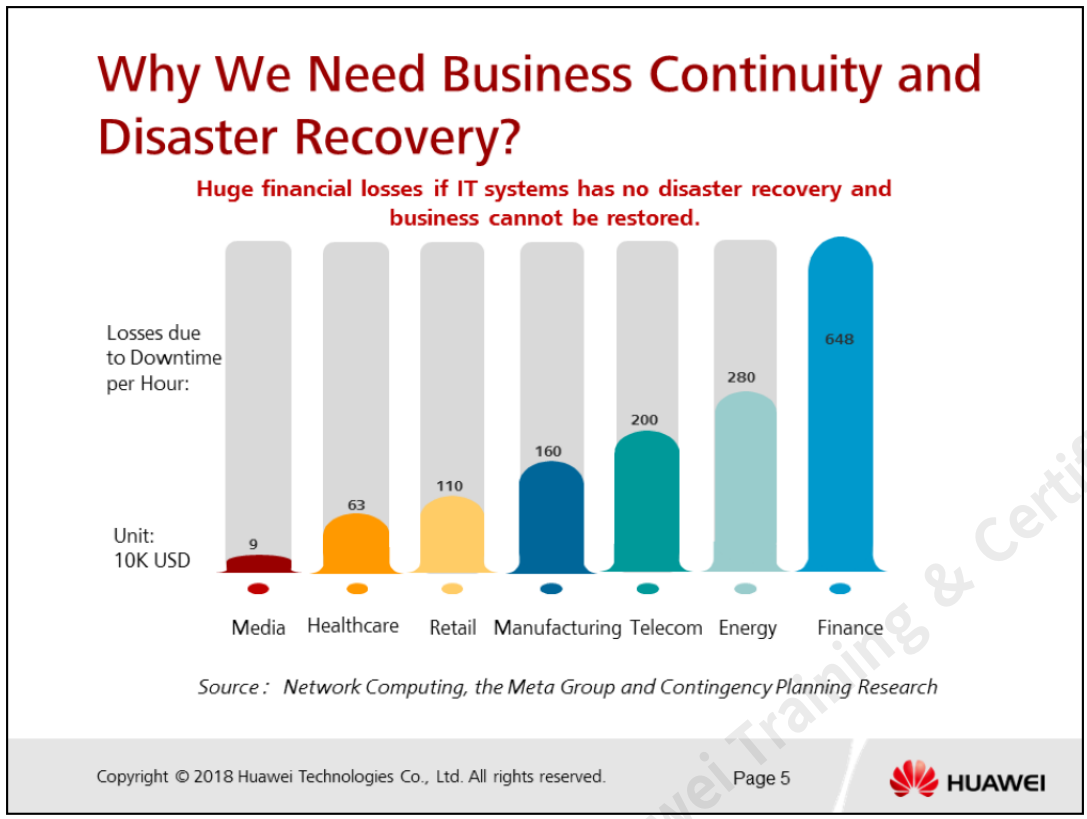
HUAWEI

Huawei Training & Certification
Huawei Training & Certification



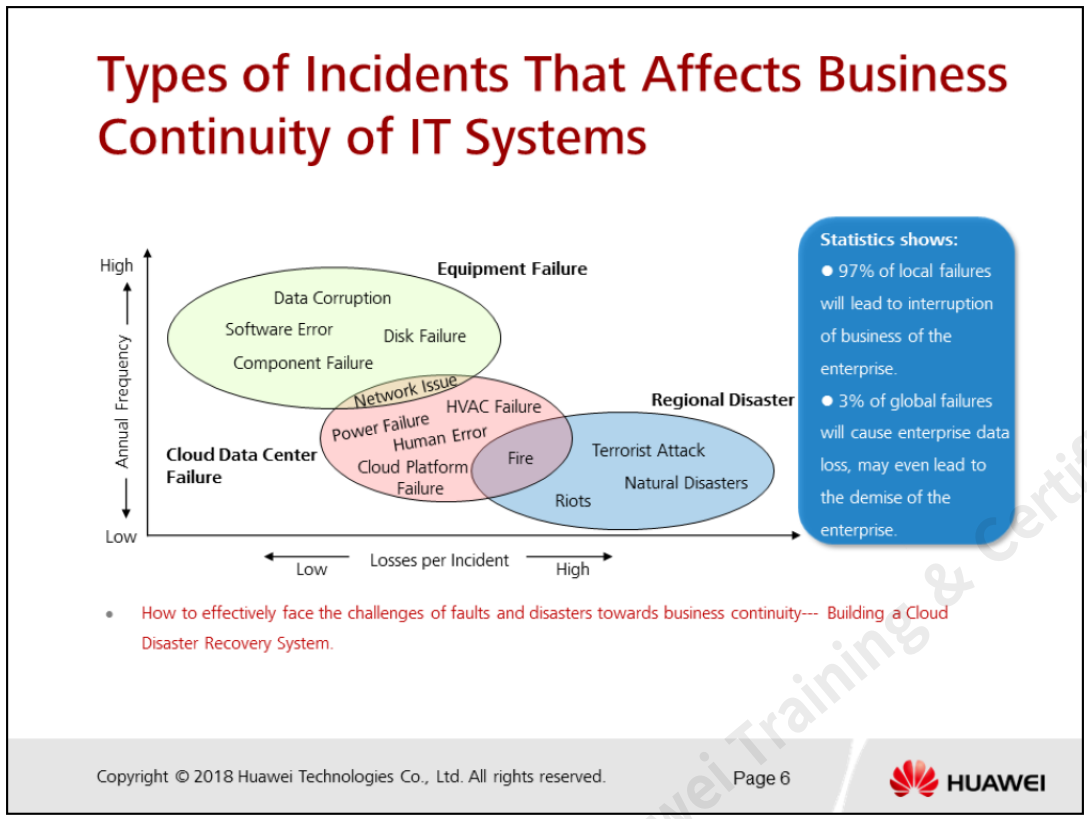
- Disasters are frequent, and able to destroy the important assets such as physical infrastructure, key personnel, information system, and key business data, which directly threatens the continuity and operation of businesses. Do you already have a plan and ready to face disasters now?
- Enterprises need to prepare for everything from natural disasters to cyber-attacks with disaster recovery plans that detail a process to resume mission-critical functions quickly and without major losses in revenues or business operations.

Huawei Training & Certification



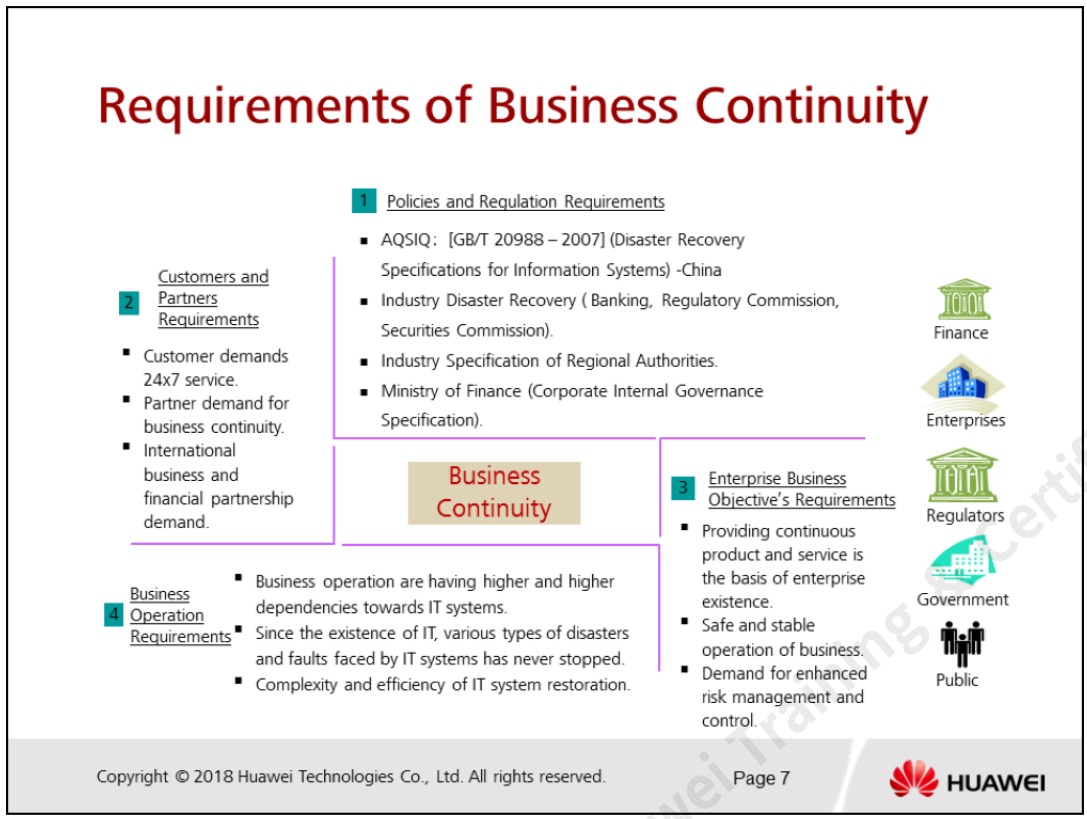
- From traditional data centers to cloud environments, business downtime will bring great financial losses and affect the reputation of the company. Customers has the demand for business continuity and data protection at any given time.
- Some government has regulated laws to build IT disaster recovery systems, such as:
  - Government regulations to require Telecom service providers to have compulsory disaster recovery systems to protect user data, and the RTO (Recovery Time Objective) must be lower than 1 hour. -Columbia
  - Government has published official documents on 2 occasions to emphasize the importance of disaster recovery and backup, and also a guide on disaster recovery, which has stressed the importance of disaster recovery, backup and restoration in the aspects of laws and regulations. -China
- Enterprises without a proper disaster recovery plan face the risks of potential financial losses and affect their reputation in the event of a disaster. No business is invulnerable to IT disasters, but speedy recovery due to a well-crafted IT disaster recovery plan is expected by today's ever-demanding customers. Too many businesses fail because they were ill prepared for an IT disaster.

Huawei Training & Certification



- Equipment level failures refers to disk failure, failure of components in storage device or breakdown of the whole storage system.
- Data center level failure refers to incidents such as long power outage of data center, HVAC (Heat, Ventilation, and Air Conditioning) failure in the data center, or a fire that caused the collapse of the whole business system.
- Regional disasters refers to floods, earthquakes, or any major disasters that happen in the region, that cause the collapse of the whole IT system within the region which may cause permanent data loss within the data center.
- In order to prevent different types of errors, faults, failures and disasters, we need to implement different types of disaster recovery solutions. For example, to face equipment level failures, we could implement high availability disaster recovery solution locally. To face data center level disasters, we can implement a same city disaster recovery center. Additionally, to face regional disasters, we can build a remote region disaster recovery center.
- As the disaster level is different, there is no one size fit all solution for disaster recovery, we must first understand the customer requirements in terms of the expectation on the disaster recovery solution and build a disaster recovery solution that meets all the needs and expectations of the customer.

Huawei Training & Certification



- The diagram above shows the requirement of business continuity from 4 different aspects which are from policies and regulations, customers and partners, business operation and business objectives.
- AQSIQ refers to The General Administration of Quality Supervision, Inspection and Quarantine of the People's Republic of China, which is a ministerial department of China, that serves as a regulator for quality and standards. In many countries, there are similar departments or regulatory bodies that specifies the standards and set the regulation for enterprises to implement a certain level of business continuity plan based on their sectors or industries.
- The public, government, regulators, enterprises and financial sector also plays a part in demanding business continuity plan to ensure smooth operation of the industries and businesses that are critical for the society. For example, government enforce policies so that key services to the public such as hospitals need implement a certain level of business continuity or disaster recovery plan, so that the public will not face the risk of not having proper healthcare in the event of a disaster.

Huawei Training & Certification



- Business continuity management (BCM) is a holistic process to ensure uninterrupted availability of all key business resources required to support critical business activities, whether manual or IT-enabled, in the event of business disruption. Business continuity planning (BCP) involves planning and procedural aspects, encompassing emergency response, crisis communications, business continuity and disaster recovery. Disaster recovery planning (DRP) is the technical component of BCP and focuses on the continuity of information and communication technology systems that support business functions.
- During the process of planning for business continuity or even managing a current business continuity plan, there are a lot of considerations to be made as shown on the diagram above. The key idea is to build and manage a business continuity solution that balances the investment cost and meets all the requirement expected. During BCP, it is best to consider all the factors involved and implement a solution that fits your current needs with some rooms for expansion as the business grows.

Huawei Training & Certification

Challenges Faced By Business Continuity and Disaster Recovery (2)



How to simplify the DR system building, wide management areas, and high difficulties of DR ?

- How to effectively complete DR operations with limited manpower to ensure enterprise business continuity?
- How to simplify, and complete the building of DR system that meet the requirements quickly when it consist of a series of steps such as business analysis, risk analysis, solution design, implementation, walkthroughs and process flows.

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

Huawei Training & Certification
Huawei Training & Certification

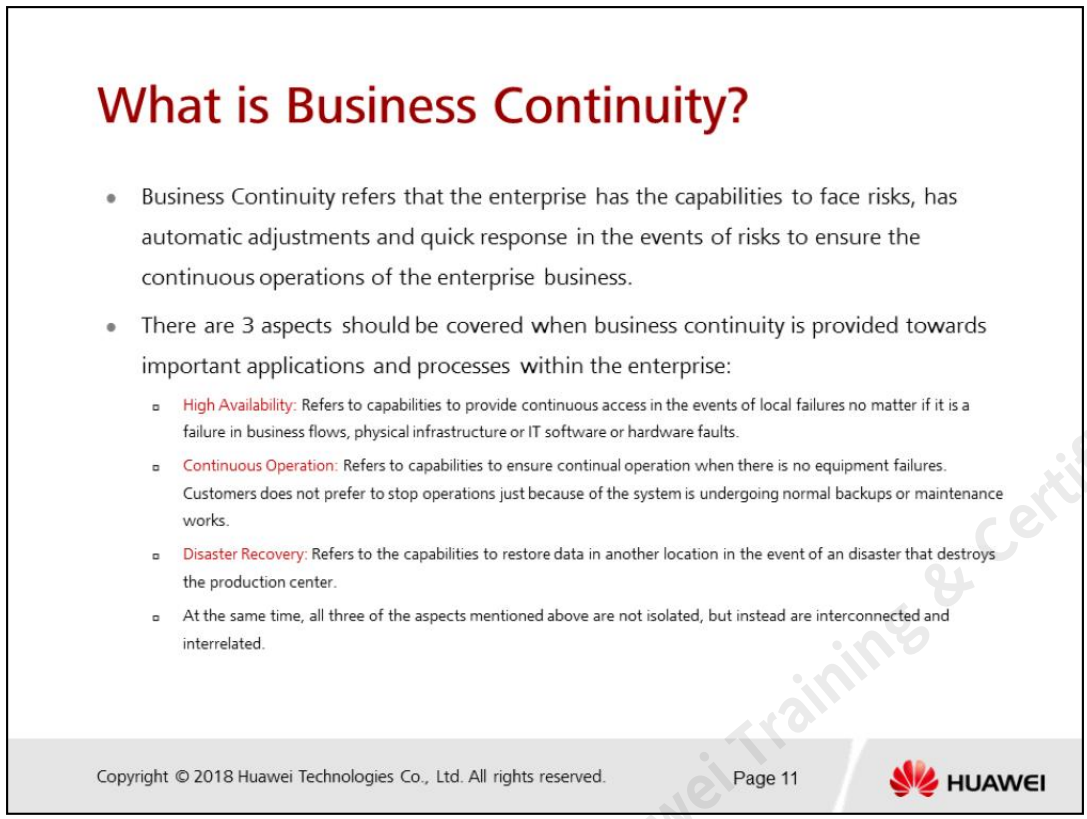
Contents

1. Challenges of Business Continuity.
2. **Definition of Business Continuity.**
3. Overview of Business Continuity Solutions.
4. Product Panorama of Huawei Business Continuity Solutions.

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

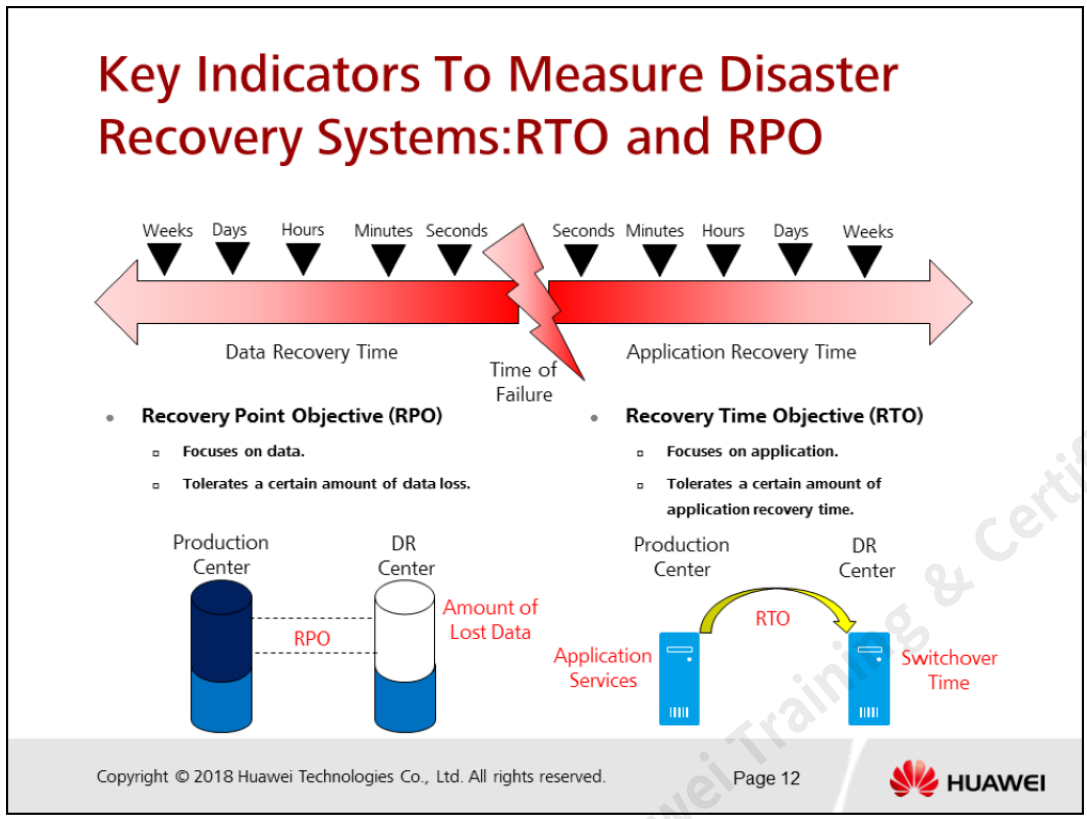
HUAWEI

Huawei Training & Certification
Huawei Training & Certification



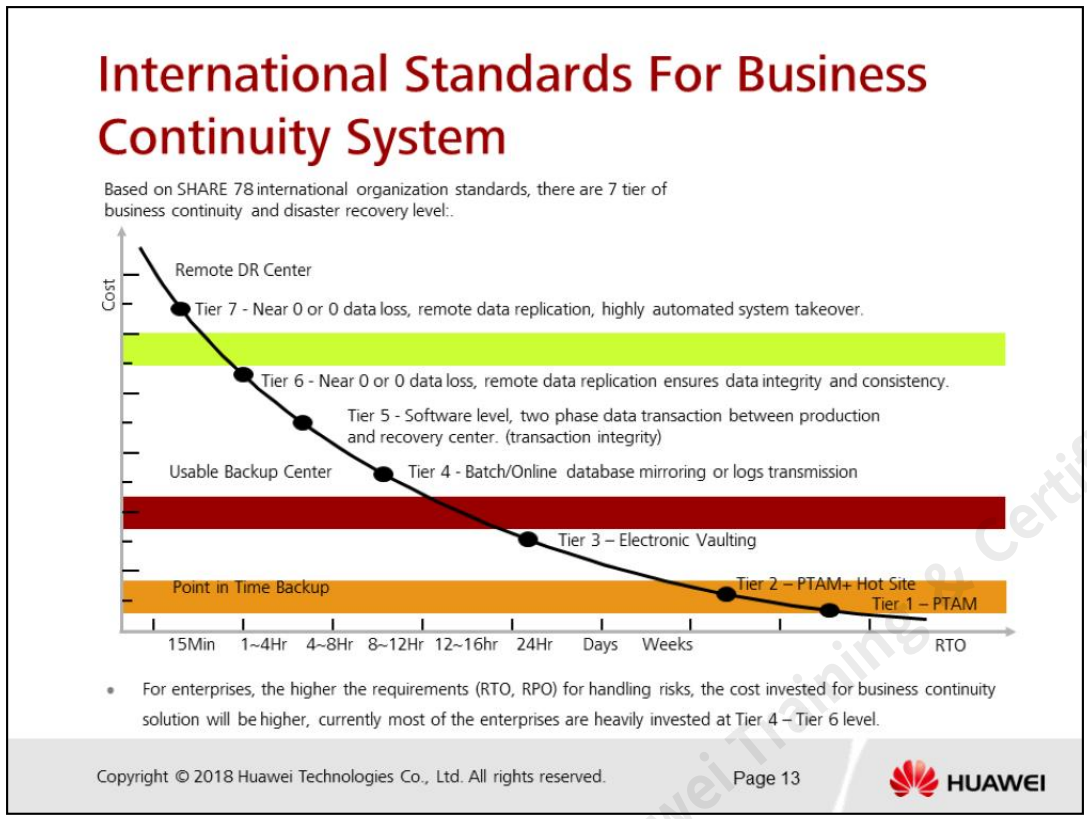
- Business continuity management originated in the disaster recovery planning of 1970s. At that era, disaster recovery activities are managed by data processing managers. During that time, if there is a major fault or crisis, downtime is calculated by days and not hours. Financial organization, such as banks and insurance companies largely choose to store backup tapes in another location away from the main data center. Recovery activities are usually caused by fire, flood, typhoon or other physical damages.
- By the 1980s, there are already a lot of commercial disaster recovery center that provides computing service on shared devices but they mainly focus on IT disaster recovery. By the 1990s, IT industry had a huge revolution, and disaster recovery plan developed into business continuity plan.
- Disaster recovery (DR) involves a set of policies, tools and procedures to enable the recovery or continuation of vital technology infrastructure and systems following a natural or human-induced disaster. Disaster recovery focuses on the IT or technology systems supporting critical business functions, as opposed to business continuity, which involves keeping all essential aspects of a business functioning despite significant disruptive events. Disaster recovery is therefore a subset of business continuity.

Huawei Training & Certification



- RTO (Recovery Time Objective) refers to the time required for the information system or business function to recover from a breakdown after a disaster occurred. The smaller the value, means that the smaller the service interruption time.
- RPO (Recovery Point Objective) refers to a point in time where systems and data are restored after a disaster has occurred. The smaller the value means that the lesser the amount of data is lost.
- How to set the RTO indicator for active-standby system switchover ?
    - RTO is the requirement for the disaster recovery switching time for business systems, it is mainly derived from the business impact analysis. From the perspective of DR system design, the switchover time of the current DR system need to be considered from the overall application system as a whole including storage, database, and application in the aspect of time required for them to takeover services, and not only considered from the aspect of storage layer.
- In another sense, RPO limits how far to roll back in time, and defines the maximum allowable amount of lost data measured in time from a failure occurrence to the last valid backup. RTO is related to downtime and represents how long it takes to restore from the incident until normal operations are available to users.

Huawei Training & Certification



- PTAM (Pickup Truck Access Method) refers to the method of transporting backups between production and off site storage facility using pickup trucks.
- The 7 tiers of Disaster Recovery:
  - Tier 1: Data Backup with No Hot Site. Regular backups are transported to offsite storage facilities. No backup system or infrastructure to restore system in the event of disaster.
  - Tier 2: Data Backup with Hot Site. Regular backup combined with offsite facility and infrastructure to restore system in the event of a disaster.
  - Tier 3: Electronic Vaulting. It is build on top of Tier 2 functionality. Data is electronically vaulted which is more current than via PTAM. As a result, less data recreation or loss when a disaster occurs. The facilities for providing Electronic Remote Vaulting consists of high-speed communication circuits, some form of channel extension equipment and either physical or a virtual tape library and an automated tape library at the remote site.

Huawei Trainings & Certification

Contents

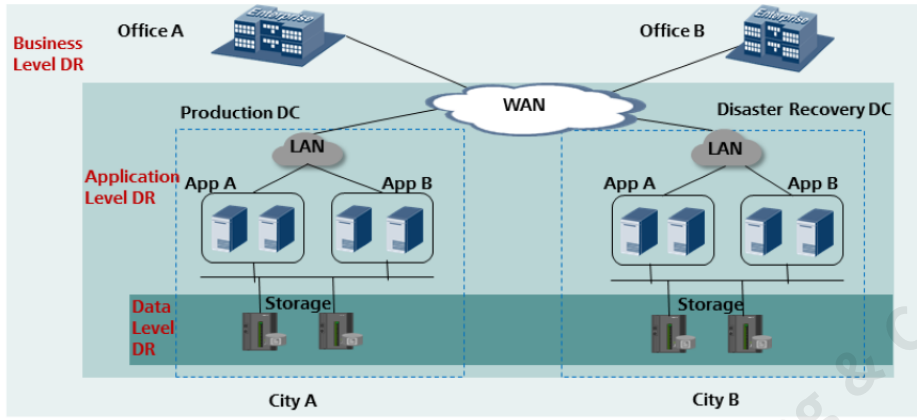
1. Challenges of Business Continuity.
2. Definition of Business Continuity.
3. **Overview of Business Continuity Solutions.**
4. Product Panorama of Huawei Business Continuity Solutions.

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

Huawei Training & Certification
Huawei Training & Certification

DR Classification Based on DR Effect

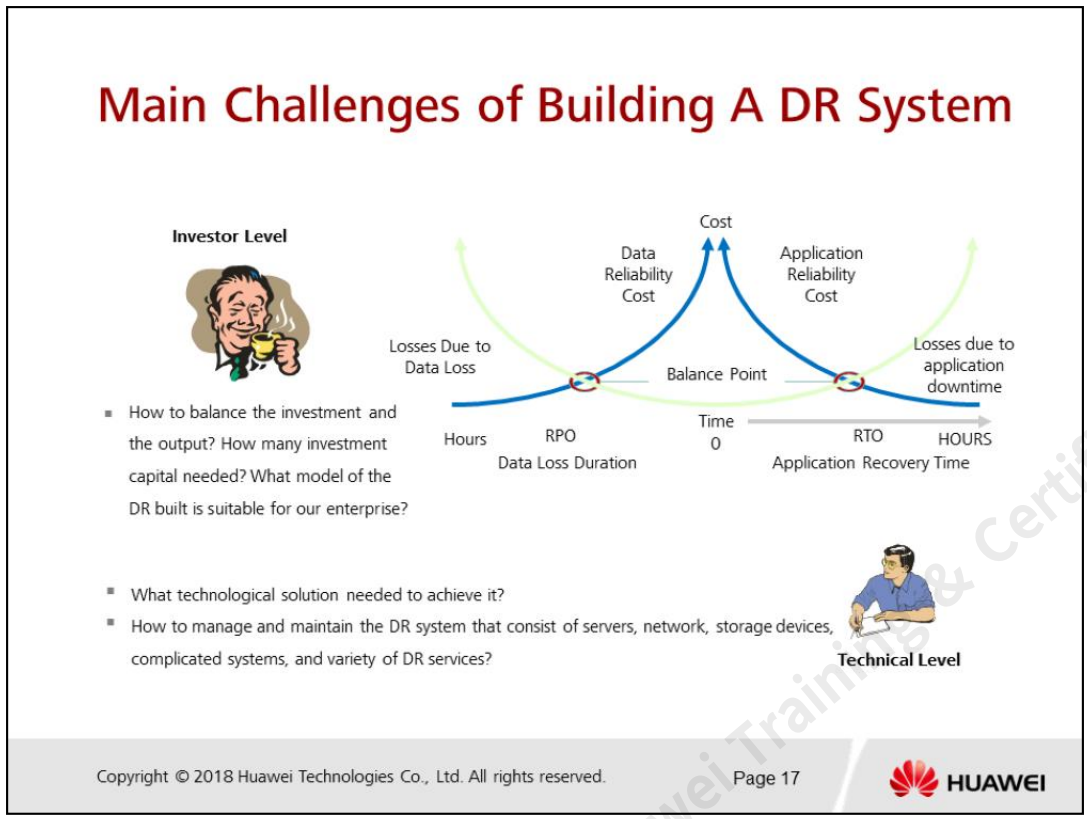


Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

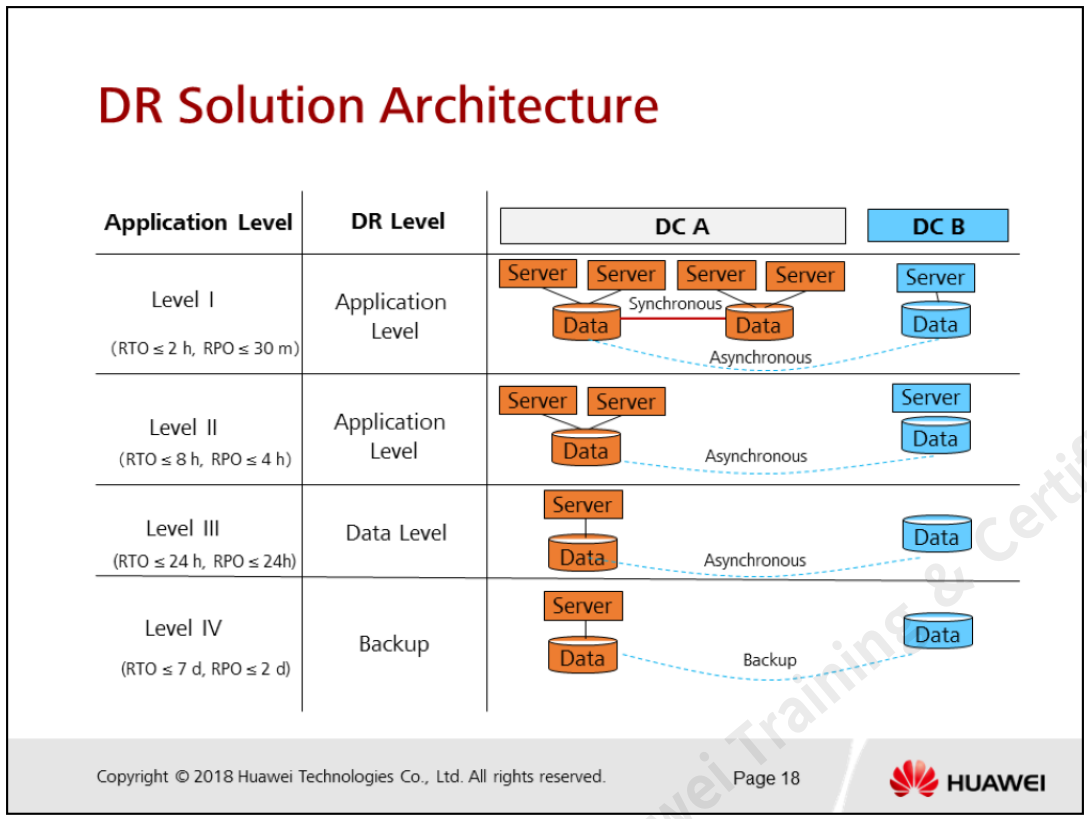
- Based on DR effect, the DR system is classified into data level DR, application level DR and business level DR.
- Data Level DR: Remote DR system data is a copy of the local key application data, when a disaster occurs at the local system, the system at least has an usable remote copy of key business data.
- Application Level DR: Building on top of the basis of data level DR, a set of backup environment equivalent to the local production system is set up remotely, including the configuration of host, network, application, and IP resources. When the local system encounters a disaster, the remote system can provide a fully usable production environment to takeover the services.
- Business Level DR: It is the DR plan for the whole business that requires all the physical infrastructures including the non IT systems such as phones and offices. When a huge disaster occurs, the original office will be destroyed, besides the data and the application restoration, we need to have a backup working environment that is able operate the business as normal.

Huawei Training & Certification



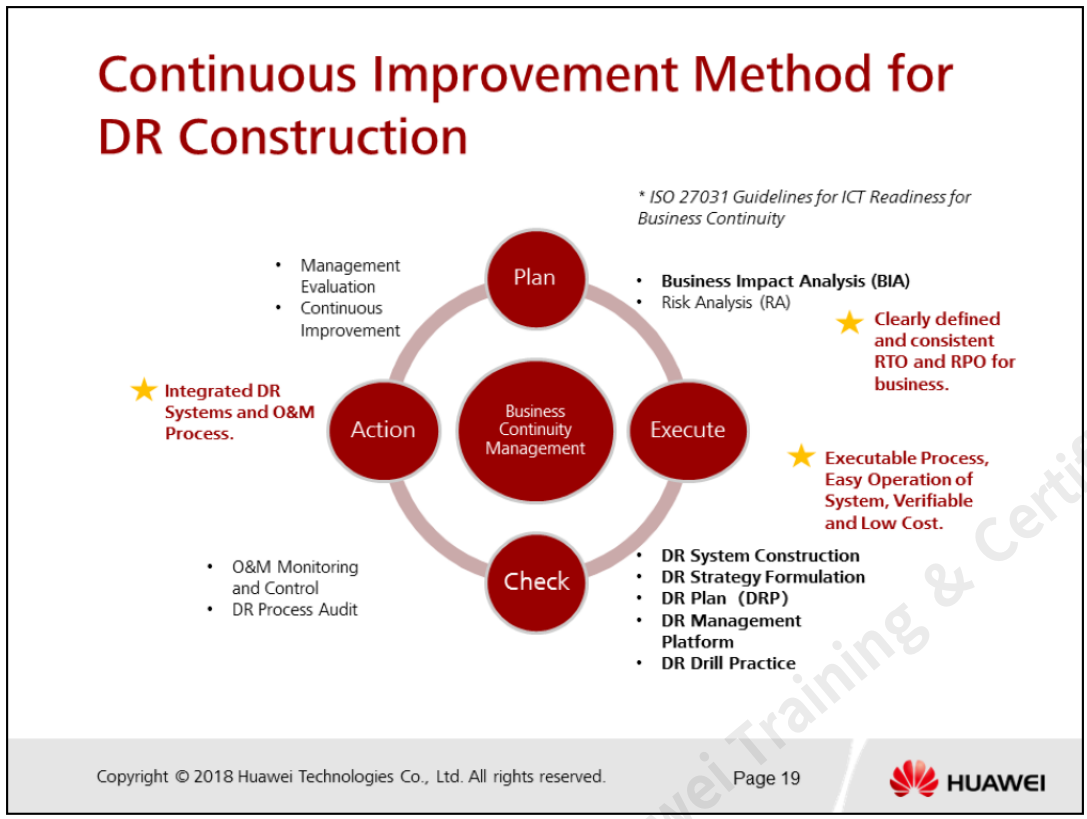
- The main challenges of building a DR system lies within the investor and technical level.
- From the investor's perspective, the amount of investments needs to be viable and has a positive rate of returns. In short, the amount invested for DR system needs to be practical and it is not too high in budget or too low that it doesn't meet the requirements of the DR plan itself. It should have a practical ROI (Return of Investment) and also scalable when the business grows.
- From the technical perspective, the main point of consideration would be which technologies should be implemented to achieve the requirement of the DR plan as there are abundant choices out there in the market for the components of a DR system. The choice of technologies must be reliable and cost effective. Other than that, the technologies implemented need to have easy operation and maintenance process, so that the O&M cost is lower and the process is more efficient with limited number of personnel.

Huawei Training & Certification



- BIA (Business Impact Analysis) analyzes the business continuity requirements from the business level. For the following example, lets consider BIA analysis for E-government service scenario. First of all, there is a need to sort out the services of the e-government before running the BIA analysis.
  - Consider the impact of downtime towards the openness, authority, social stability, government credibility, laws violations, and business scope. Impact ratings can be classified into 3 levels such as "low", "medium", and "high". By calculating the sum of impact value in various businesses and sectors, we can derive the overall impact value for the downtime. The threshold for implementing DR is set at impact rating level 2 (impact value of 5-8).
  - By evaluating the impact of business downtime on 8 points of time (1h, 4h, 8h, 24h, 2d, 3d, 5d, 10d), we can derive the MAO (Maximum Acceptable Outage) for variety of businesses.
  - By analyzing the relationship between applications (including business support systems, integrated relationship between applications), we can analyze and derive the requirements of DR (RPO & RTO).
  - Based on RTO and RPO, we can classify the DR level for the e-government system into 4 different levels as shown on the table above, formulate the strategies for DR, and formulate the basis of DR solutions.
- Some countries like China published official documentations for DR standards and specifications. An example would be the document numbered (GB/T 20988-2007) titled "Information Security Technology - Disaster Recovery Specifications For Information System".
- Social Security Center: Basic data must have a complete backup daily, and the core system RTO must meet the requirement of 4 hours.

Huawei Training & Certification - Huawei Training & Certification



- The diagram above shows the continuous improvement method for the construction of DR systems based on the best practices in the industry and ISO guidelines.
- The process above serves as a guideline to construct, evaluate and manage the DR system to ensure that it is up to date and able to respond and recover from disasters efficiently.
- As technologies changes, more and more option is available for better DR system construction or management. As business changes, the requirements for DR may not be the same as the initial planning or setup. Hence, there is a need to continuously improve the DR system to ensure that it meets the current business requirements. It is also crucial to test the DR system or do simulated drill practices to ensure that the DR process is valid, suitable and all the personnel involved in the process are aware of their roles. This ensures that in the event of the disaster, the DR process can be carried out smoothly with lower amount of time to resume the business operations as usual.

Huawei Training & Certification

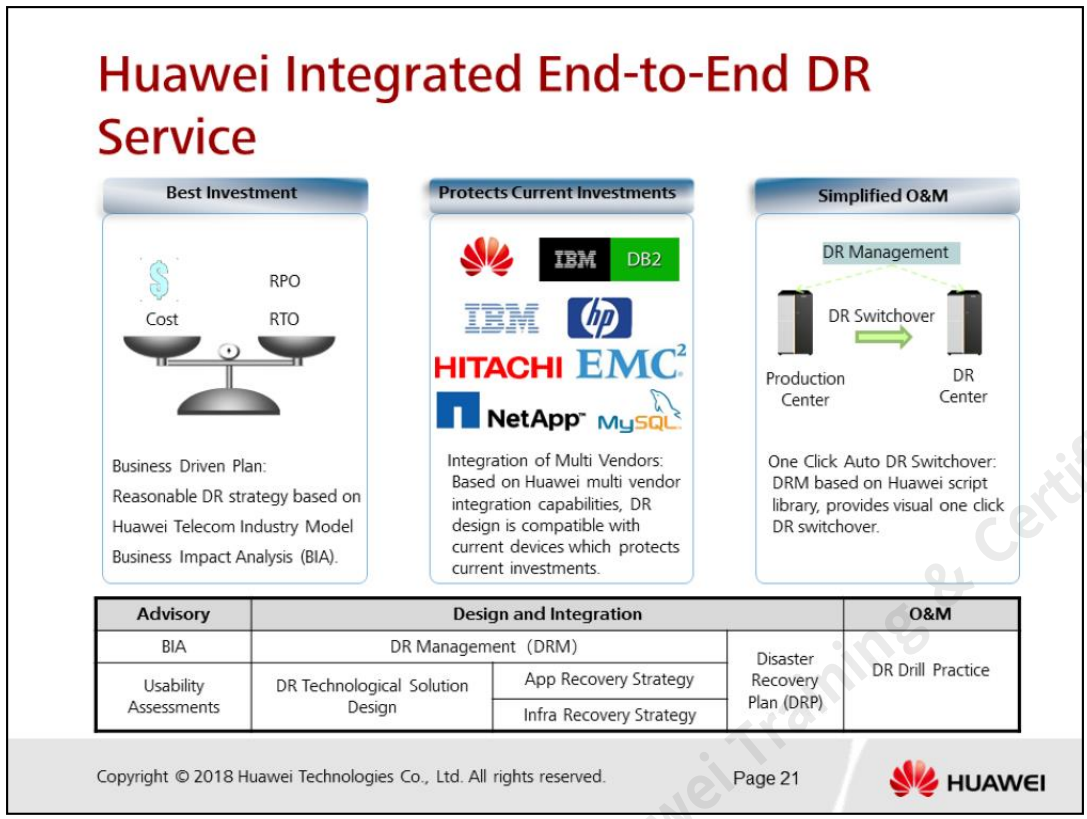
Contents

1. Challenges of Business Continuity.
2. Definition of Business Continuity.
3. Overview of Business Continuity Solutions.
4. **Product Panorama of Huawei Business Continuity Solutions.**

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

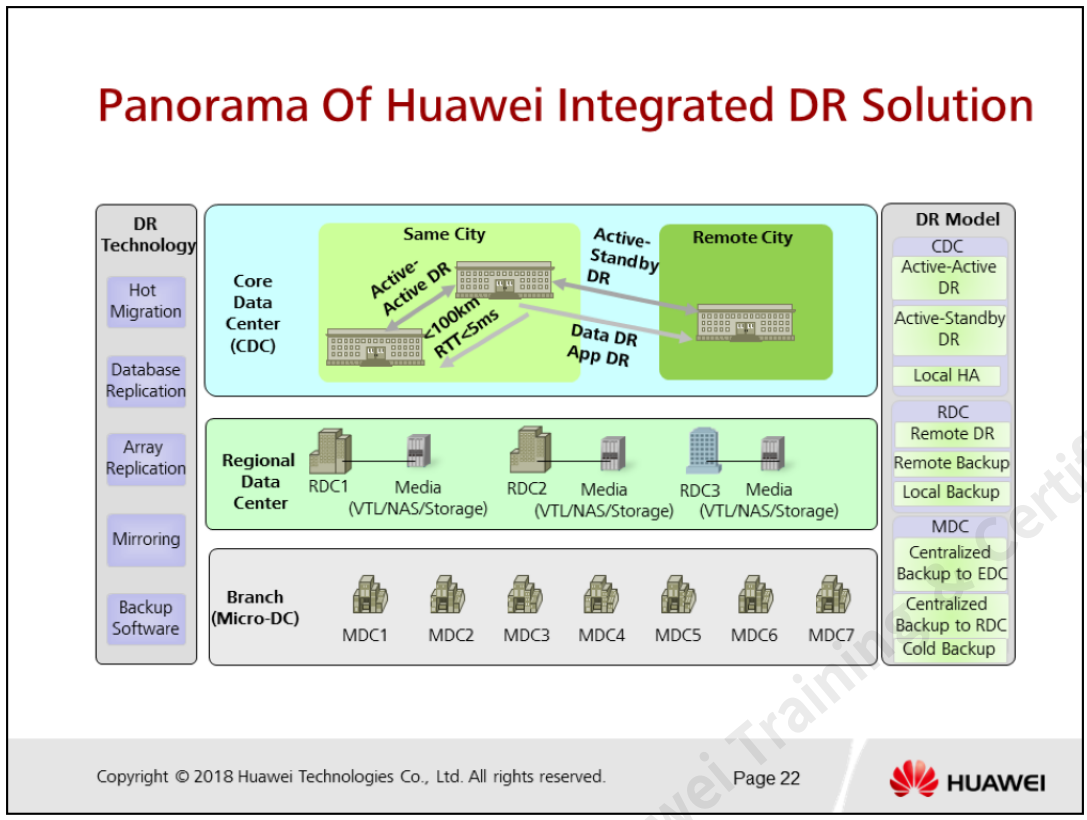
HUAWEI

Huawei Training & Certification
Huawei Training & Certification



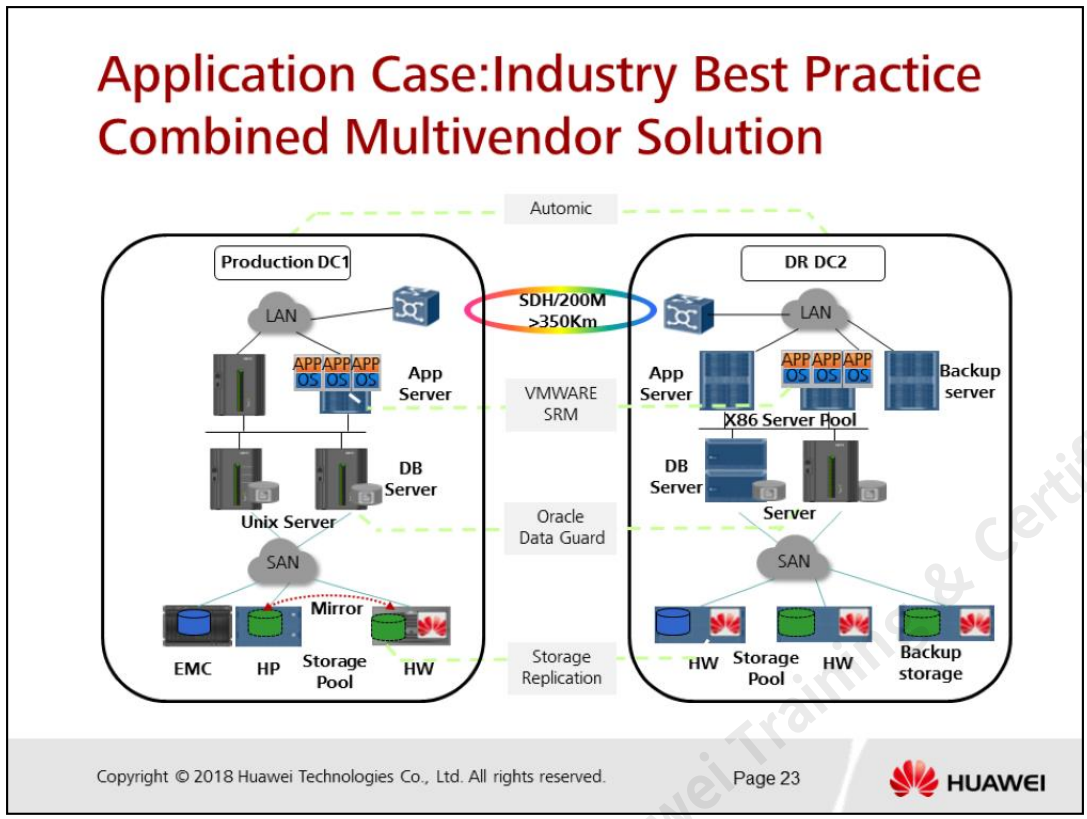
- Business driven plan can be defined as the most reasonable DR target to achieve the best investment. It includes: sorting out the business systems, formulate the BIA and provide the most suitable DR strategy plan.
- Huawei's capabilities for multi vendor integration with solution design that is compatible with current devices, effectively protects current investments. Huawei provides complete DR solutions from data level to application level that can be configured to meet the different requirements of variety of businesses.
- One click auto switchover lowers the complexity of DR operation and maintenance. Other enhancements of Huawei DR solution are:
  - Visual daily management and switchover management: Through the use of DRM management platform, the daily O&M and the switchover process can be visualized and the progress of the switchover can be monitored in real time.
  - Switchover Drill Practice: The DR switchover process can be simulated through the DRM platform, which lowers the cost of O&M drill practice.
  - One click auto switchover: It is based on Huawei script libraries which allows servers, databases and DNS switchover to be scripted, and all the action script are run with an one click operation. Hence, in the event of a disaster, the switchover process from production to DR center can be completed with ease.

Huawei Training & Certification



- The diagram above shows the panorama of the Huawei DR systems. Huawei's solution for disaster recovery and business continuity is comprehensive and covers all aspects of data protection, application recovery and data center level disaster recovery either within the city or from a remote region.
- Huawei's DR solution design is compatible with lots of different and latest DR technologies within the market, and can be built in various DR models to meet customer requirements.
- Different configurations in the levels of branches, regional data center and core data center allows a flexible and scalable DR system to be implemented which satisfies the different DR needs from each levels.

Huawei Training & Certification



- Key Technologies of the solution:
  - Implements Active-Standby DC model over a distance of 350 km.
  - Main DC uses Oceanstor 6800 V3/18K with heterogeneous architecture of EMC/IBM/HP storage.
  - VMWARE SRM manages the virtual machine synchronization.
  - Oracle Data Guard manages the database synchronization.
  - Implements Huawei storage replication for data synchronization.
  - Implements Automic DRM software for automated DR management.

- Benefits to the customer:
  - Help the customer to solve the problems of constructing a one-stop IT disaster recovery center, and eliminate potential safety problems.
  - Dual insurance on the data restoration, increases data protection level.
  - Core application RTO is lower than 2 hours, RPO lower than 10 minutes.
  - Storage online dynamic expansion, which saves maintenance cost by 15%.

Huawei Training & Certification

Summary

- This module mainly introduced:
  - The Challenges Of Business Continuity.
  - The Definition Of Business Continuity.
  - The Overview Of Business Continuity Solutions.
  - Panorama Of Huawei Business Continuity Solutions.

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

Huawei Training & Certification
Huawei Training & Certification

Quiz

1. What are the key indicators for measuring DR systems?
   A. RPO
   B. CIO
   C. RTO
   D. RTT
2. Which of the followings is a business continuity solution?
   A. Active-Active DC Solution.
   B. Active-Standby DR Solution.
   C. 3DC Solution.
   D. Local HA Solution.

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

- Answers:
  - AC.
  - ABCD.

Huawei Training & Certification
Huawei Training & Certification

Thank You

www.huawei.com

Huawei Training & Certification
Huawei Training & Certification

10 Backup Solutions And Its Applications

www.huawei.com

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

Huawei Training & Certification
Huawei Training & Certification

Foreword

- This module mainly describes about the related knowledge of backup solutions, including:
  - The Concepts of Backup Solutions.
  - The System Architecture of Backup Solutions.
  - The Components of Backup Solutions.
  - The Related Technologies of Backup Solutions.
  - The Typical Application Cases of Backup Solutions.

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

Huawei Training & Certification
Huawei Training & Certification

Objectives

- Upon completion of this module, you will be able to:
  - Describe the related technologies of Storage Backup Solutions within the Datacenter.
  - Understand the technical differences of Backup and Disaster Recovery solutions.
  - Learn how to deploy a Backup Solution through application cases.

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

Huawei Training & Certification
Huawei Training & Certification

Contents

1. **Overview Of Backup Solutions.**
2. Architecture Of Backup Solutions.
3. Common Backup Networks.
4. Common Backup Technologies.
5. Application Cases Of Backup Solutions.

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

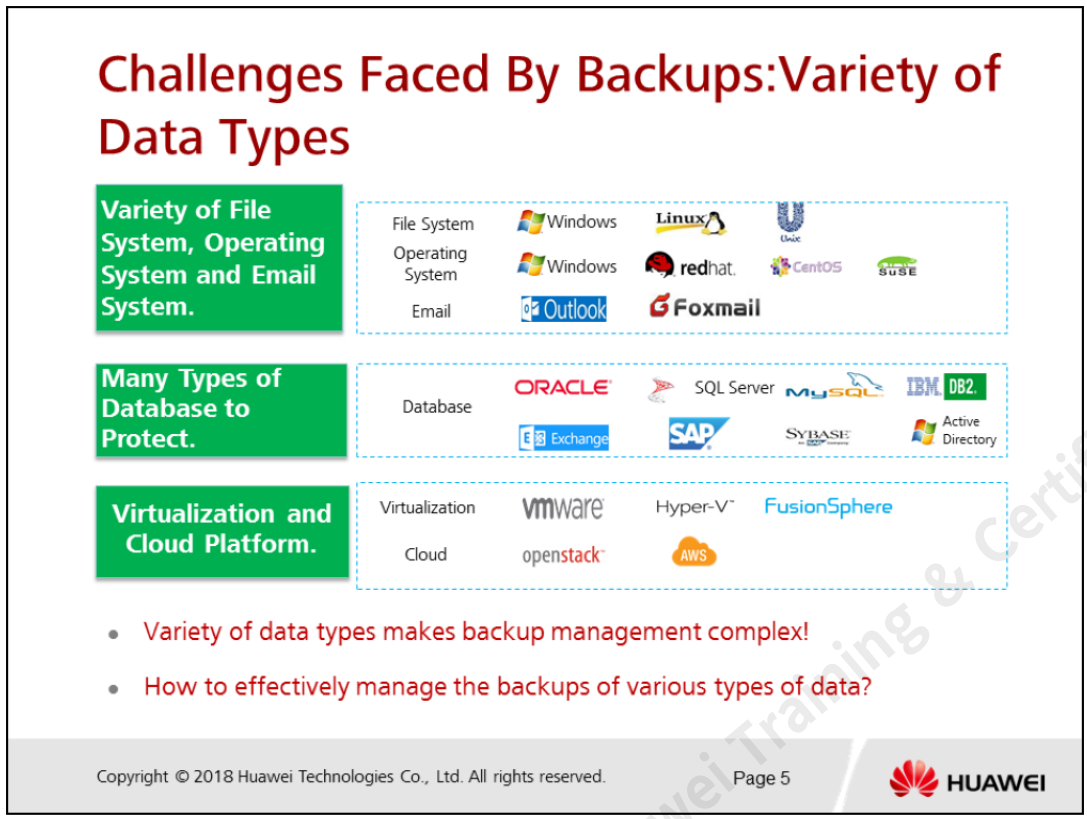
HUAWEI

Huawei Training & Certification
Huawei Training & Certification



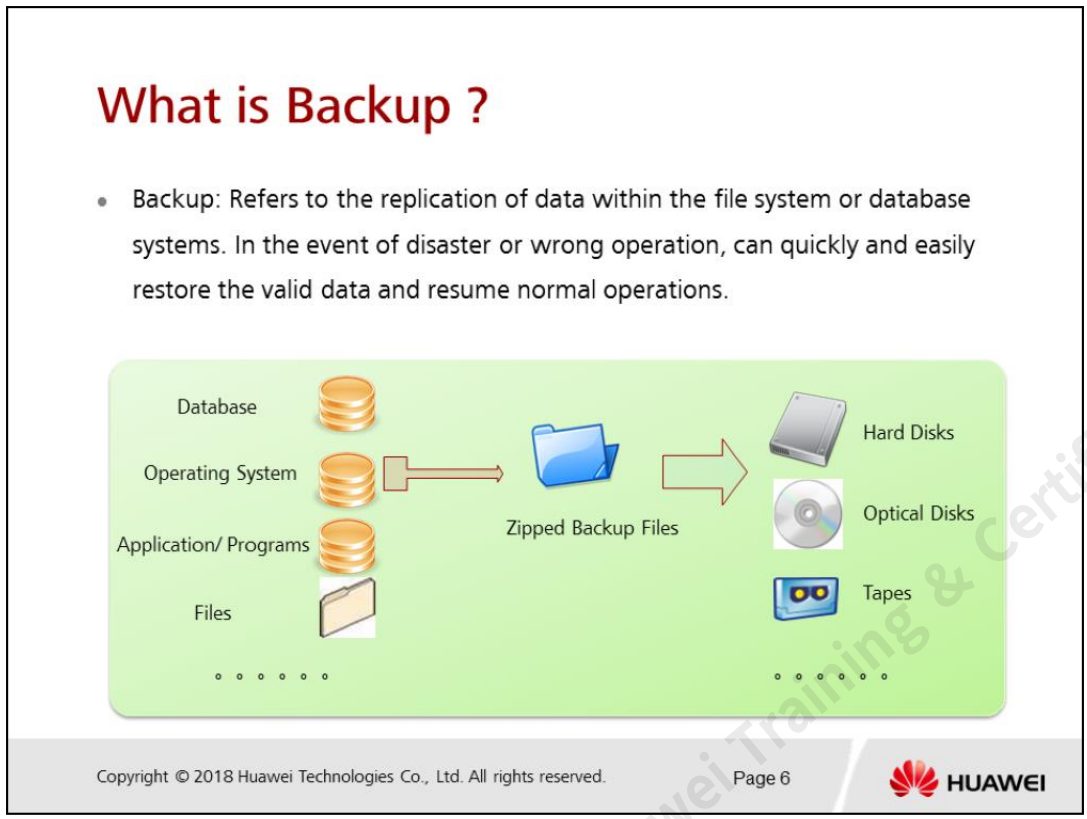
- Data backup is the last line of defense in data protection.
- Compared to the developments of IT environments, the growth and development of backup software are lagging behind in pace.
- Regular backups are an insurance policy for your business. With so much of the modern business supported by computer networks, a loss of your network data can have a crippling financial effect on your business.
- Backups help ensure that regardless of your current security strategy, you have secure and clean data to keep your business running in the case of data loss, a hack/virus attack or even a natural disaster.

Huawei Trainings & Certification



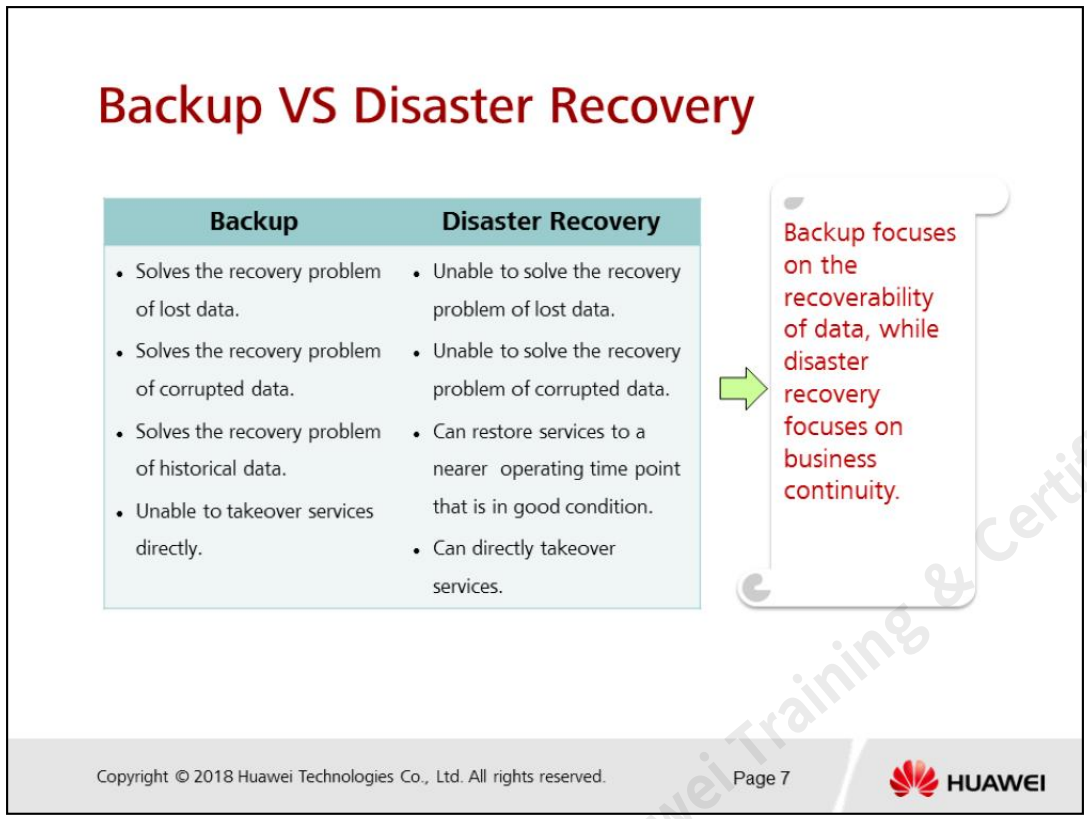
- The variety of data types made the backup process complicated. We are no longer dealing with single format or type of data, but are facing with myriads of data types and applications running on physical, virtualized or cloud environments.
- For backup administrators, this serves as a challenge to manage the different types of data format, and ensuring that all backup tasks are completed and valid.
- Normal backup solutions may work for small and medium enterprises, but for large scale enterprises, backup becomes very complex and require solutions that provides unified management of backup tasks for many types of data and application.

Huawei Training & Certification



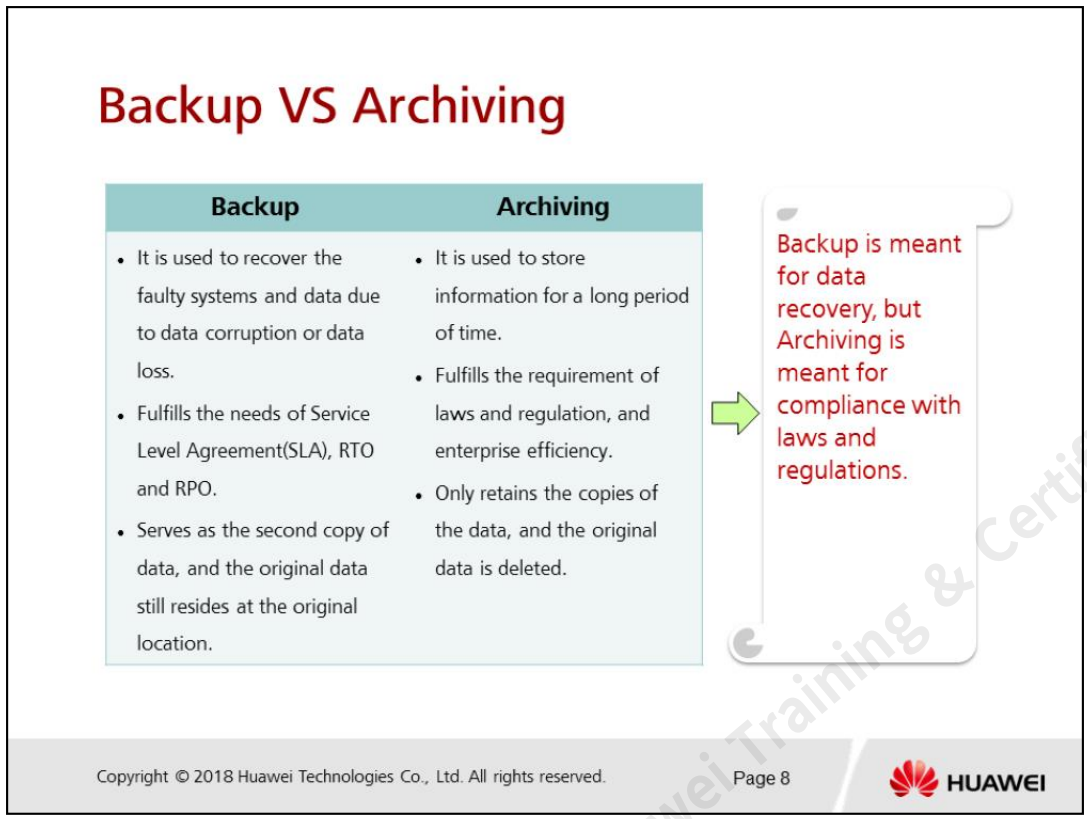
- Definition of Backup:
  - Replication of data within the storage device to another mass storage device such as tapes to cope with the incidents of files or data loss.
  - Backup in reality is just a type of "replication" operation, and data backups usually are stored in a another format.
  - The purpose of backup is for recovery of data in the incident of disasters.
- Commonly, the following components are within a set of backup systems:
  - Backup Servers: It is the host for the backup software, and generally are PC servers or minicomputers.
  - Backup Software: Its the core of backup system. It copies the production data to the storage media, and manages the data backups. Examples of backups software are, Symantec Backup Exec and NetBackup etc.
  - Storage Device: The device where the backup data is stored, commonly are disk arrays, physical tape libraries or virtualized tape libraries.

Huawei Training & Certification



- Backup is the basis of disaster recovery.
- Many tends to think that backup and disaster recovery are the same, but each of them has their own merits and differences.
- Backup in simpler terms refers to the process of copying the data to another location for recovery in the event of data loss or corruption.
- Disaster recovery refers to the process of maintaining business continuity in the event of a disaster.
- Disaster recovery can be considered as a set of policies and procedures which focus on protecting an organization from any significant effects in case of a negative event, which may include cyber attacks, natural disasters or building or device failures. Disaster recovery helps in designing strategies that can restore hardware, applications and data quickly for business continuity.

Huawei Training & Certification



- Data archiving is the process of moving data that is no longer actively used to a separate storage device for long-term retention. Archive data consists of older data that is still important to the organization and may be needed for future reference, as well as data that must be retained for regulatory compliance. Data archives are indexed and have search capabilities so files and parts of files can be easily located and retrieved.
- Data archives are often confused with data backups, which are copies of data. Data backups are used as a data recovery mechanism that can be used to restore data in the event it is corrupted or destroyed. In contrast, data archives protect older information that is not needed for everyday operations but may have to be accessed occasionally. The data archives serve as a way of reducing primary storage consumption and related costs, rather than acting as a data recovery mechanism. Some data archives treat archive data as read-only to protect it from modification, while other data archiving products treat data as read / write. Data archiving is most suitable for data that must be retained due to operational or regulatory requirements, such as document files, email messages and possibly old database records.

Huawei Training & Certification

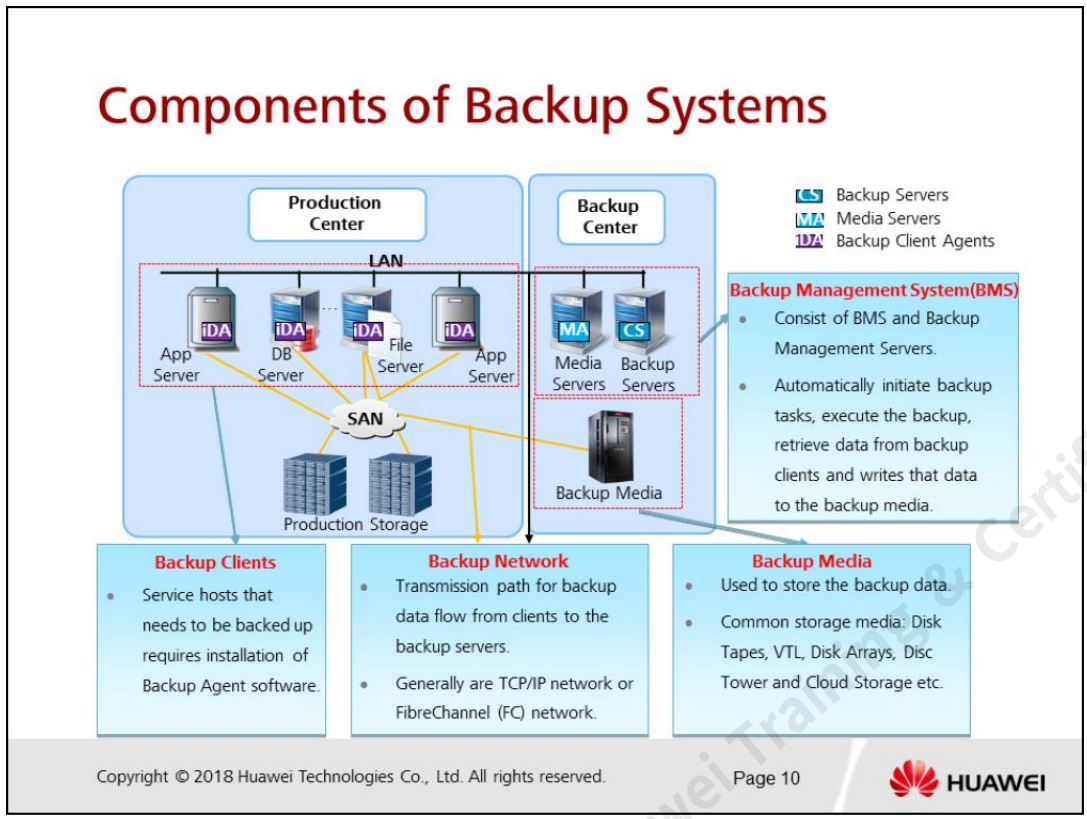
Contents

1. Overview Of Backup Solutions.
2. **Architecture Of Backup Solutions.**
3. Common Backup Networks.
4. Common Backup Technologies.
5. Application Cases Of Backup Solutions.

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

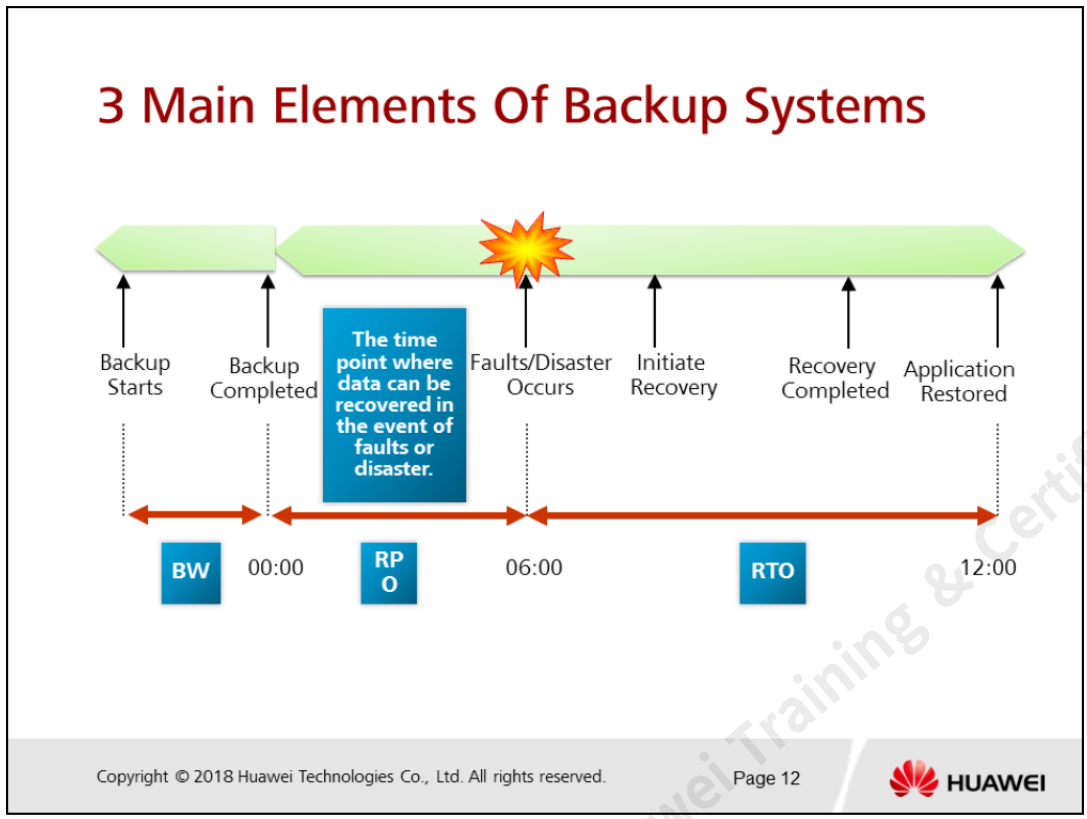
HUAWEI

Huawei Training & Certification
Huawei Training & Certification



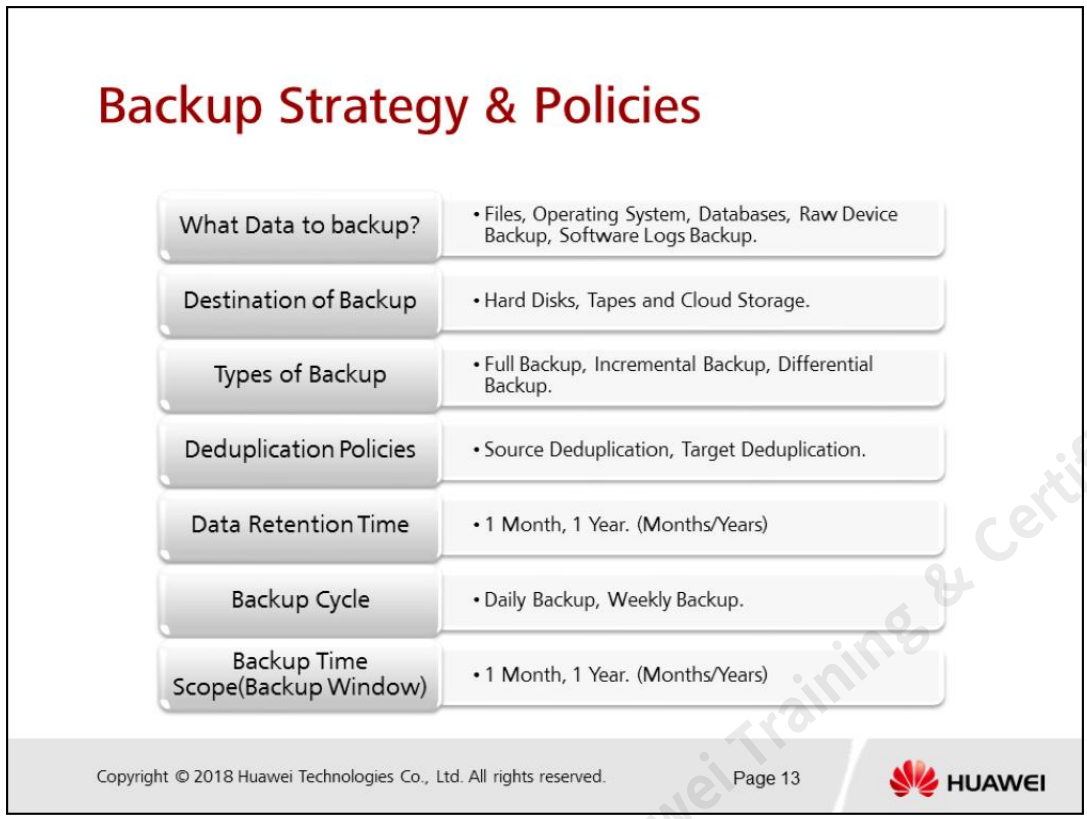
- Backup server: Maintains all the configuration information (such as information on clients, media agents, backup devices, etc.), and manages all the job and historical task records, and provides GUI/WEB and other management interfaces.
- Media server: Responsible for managing the data transmission between the physical storage device and its corresponding clients and agents.
- Backup media:
  - Disk array: It is applicable and useful in occasions where the backup data volume is large and the backup window is relatively small, which requires high performance and reliability of the backup device, such as within data centers. Examples of suitable products are Huawei OceanStor V3 series and 18000 series.
  - Tape library: It is suitable for scenarios where the backup data volume is small, the time window is large, or in the case where you want to save the data offsite for a long period of time. Examples of such products are HPMSL6060, IBM TS3310 etc.
  - Virtual tape library: It is applicable in the scenario where it is necessary to inherit the enterprise's original tape backup architecture and policies, but at the same time also need to improve the backup speed. Products such as Huawei VTL6900 are suitable in these scenarios. (VTL: Refers to Virtual Tape Library. It can be integrated into existing tape backup environments and used along with tape products to form a solution that combines the advantages of both tape and disk technology.)

Huawei Training & Certification



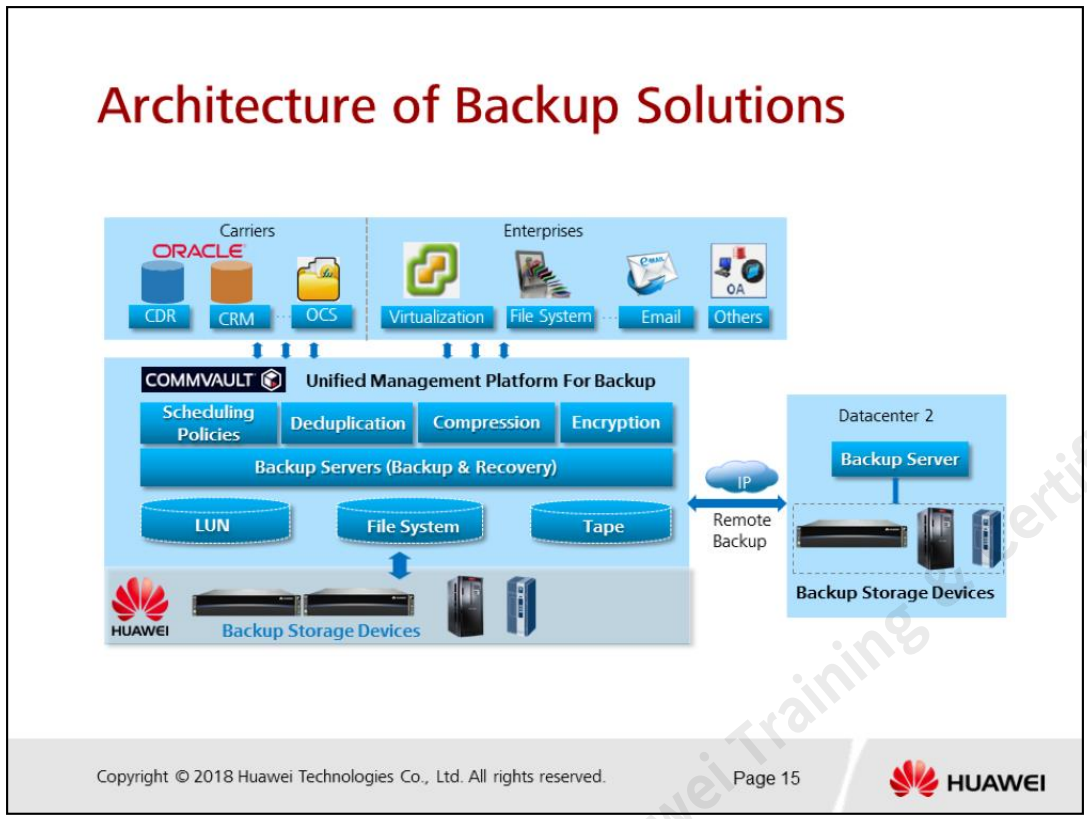
- BW (Backup Window): The length of time for a backup to be completely backed up in a work cycle.
- RPO (Recovery Point Objective): The maximum amount of data loss that can be tolerated.
- RTO (Recovery Time Objective): The longest downtime that can be tolerated.

Huawei Training & Certification



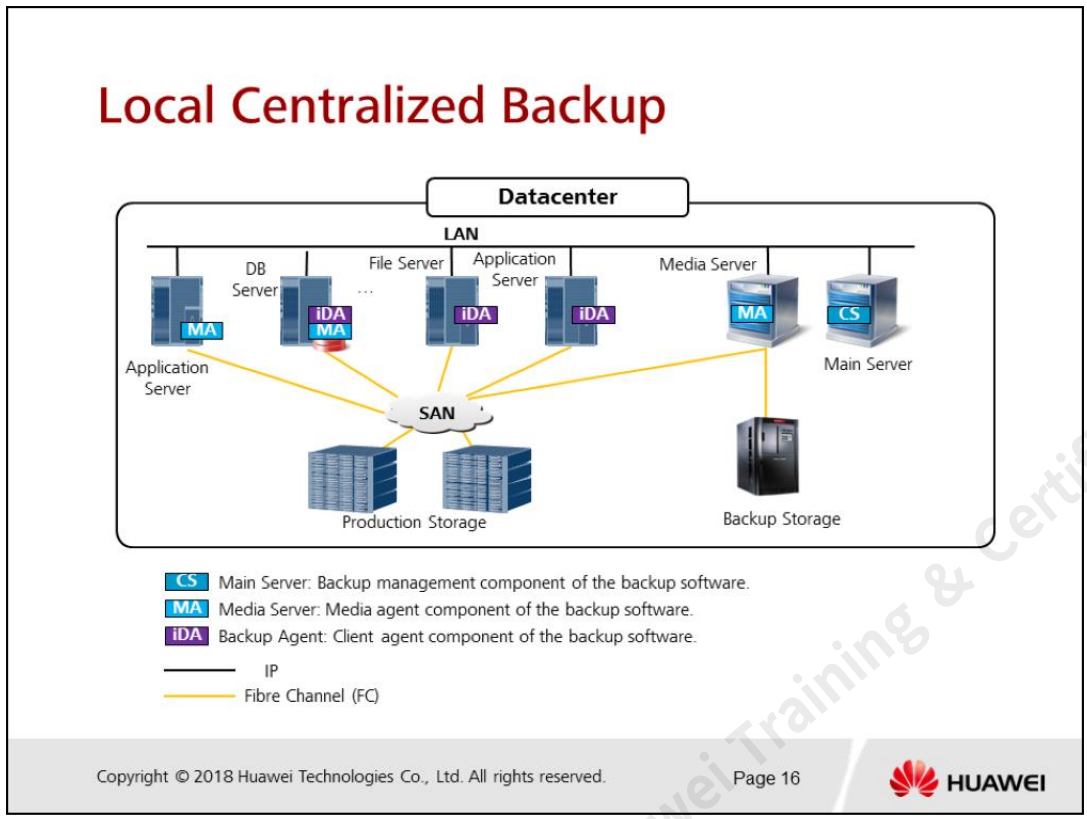
- Backup strategy and policies option:
  - More inclined towards the operating system and application software, which performs a full backup after each system update or installation of new software.
  - For critical application data, which involves a large numbers of daily data updates, but the total amount of data is not very large. Thus, we can make a full daily backup when users just uses a relatively small amount of data in this scenario.
  - For critical applications, where the total amount of data has only a small amount of data updates every day. Thus, we can do a full backup on a monthly or weekly basis, and on an incremental basis, we can do incremental backups at a series of shorter intervals of time.
- Typical backup type:
  - Full backup: full daily backup.
    - Advantages: Ability to quickly restore data based on the last full backup, with a small recovery window.
    - Disadvantages: It takes up a lot of storage space. Each backup takes a long time and the backup window is large.

Huawei Training & Certification



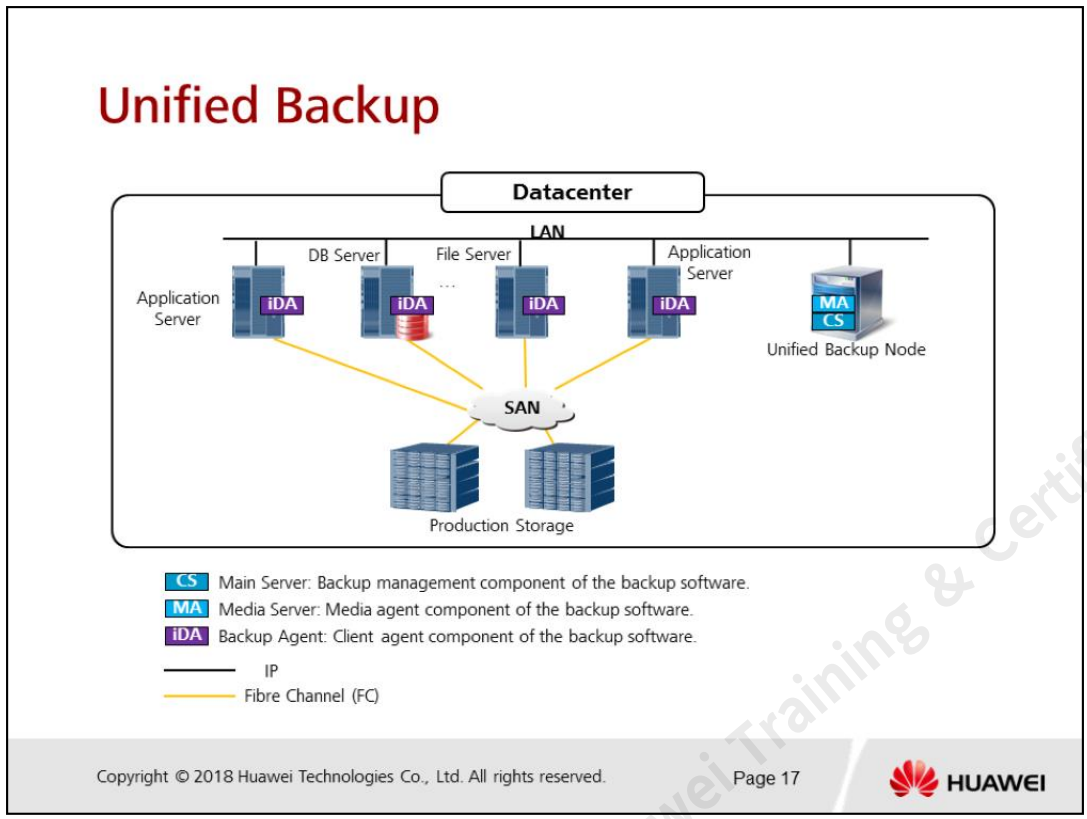
- The diagram above shows the overview of the architecture of backup solutions.
- Backup solutions actually consists of one or more components in design to provide the best backup performance to meet customer requirements. The components can be software or hardware or external platforms such as the cloud.
- Backup solutions need to be designed with the specified requirements in mind, and flexible in expansion to cater with the growth of data. Thus, the components chosen in the backup solutions needs to be compatible with each other to gain maximum performance and maximize the value of investment.

Huawei Training & Certification



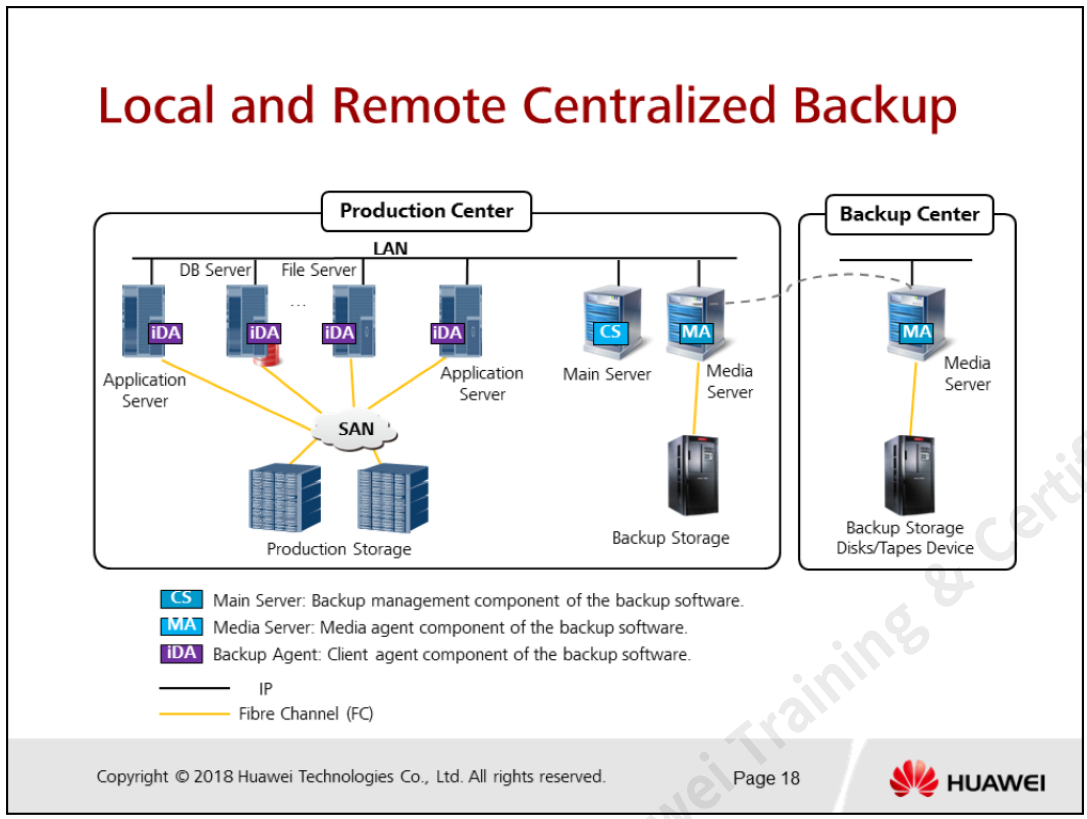
- Scenario: Used for Local DC Centralized Backup, Carrier, and Enterprise, etc.
- Supported platforms:
  - Supports file and application backups for Windows, Linux, and UNIX platforms.
  - Support backups for VMware, Hyper-V, and FusionCompute virtualization platform.
- Highlights of the solution:
  - Supports parallel deduplication, saving more backup storage space.
  - Supports hardware based snapshots of Huawei storage arrays for data backup, zero service impact, and second-level recovery.
  - Full graphical GUI unified management, there is no need to write any script, and has simple operation and maintenance.

Huawei Training & Certification



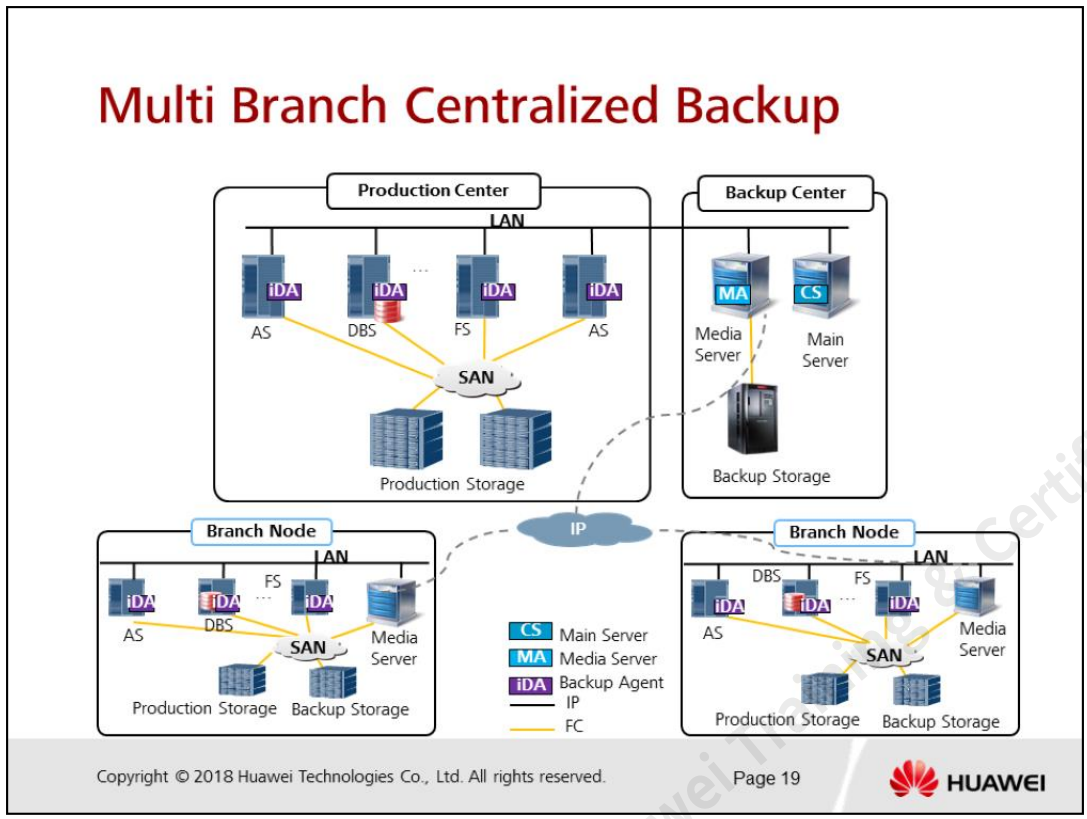
- Scenario: Usage for Local backup of SMB and branch offices that small amount of data.
- Supported platforms:
  - Supports file and application backups for Windows, Linux, and UNIX platforms.
  - Support backups for VMware, Hyper-V, and FusionCompute virtualization platform.
- Highlights of the solution:
  - Backup server and backup storage integration, allows equipment and maintenance costs to be reduced by 50%.
  - Full graphical GUI unified management, there is no need to write any script, and has simple operation and maintenance.

Huawei Training & Certification



- Scenario:
  - Used for Local DC Centralized Backup, Carrier, and Enterprise, etc.
  - Remote Backup.
- Supported platforms:
  - Supports file and application backups for Windows, Linux, and UNIX platforms.
  - Support backups for VMware, Hyper-V, and FusionCompute virtualization platform.
- Highlights of the solution:
  - Supports parallel deduplication, saving more backup storage space.
  - Supports hardware based snapshots of Huawei storage arrays for data backup, zero service impact, and second-level recovery.
  - Full graphical GUI unified management, there is no need to write any script, and has simple operation and maintenance.

Huawei Training & Certification



- Scenario:
  - Local DC Centralized Backup, Carrier and Enterprises.
  - Unified Management.
- Supported platforms:
  - Supports file and application backups for Windows, Linux, and UNIX platforms.
  - Support backups for VMware, Hyper-V, and FusionCompute virtualization platform.

Huawei Training & Certification

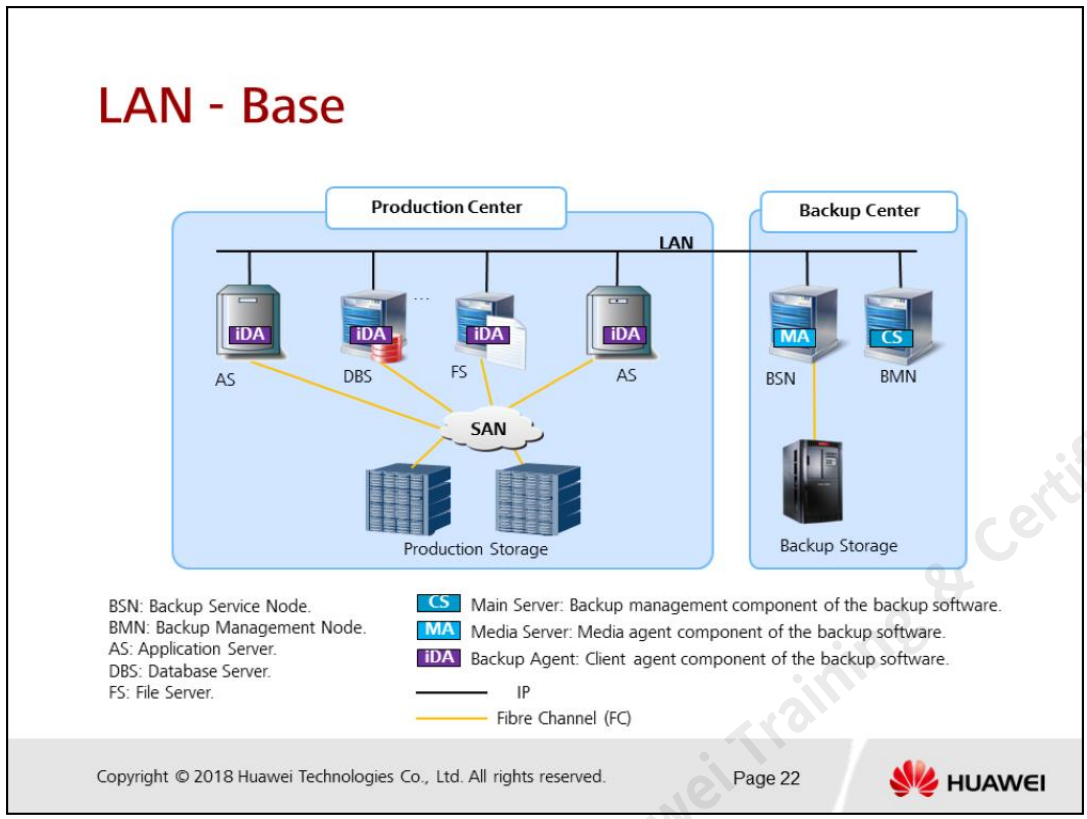
Contents

1. Overview Of Backup Solutions.
2. Architecture Of Backup Solutions.
3. **Common Backup Networks.**
4. Common Backup Technologies.
5. Application Cases Of Backup Solutions.

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

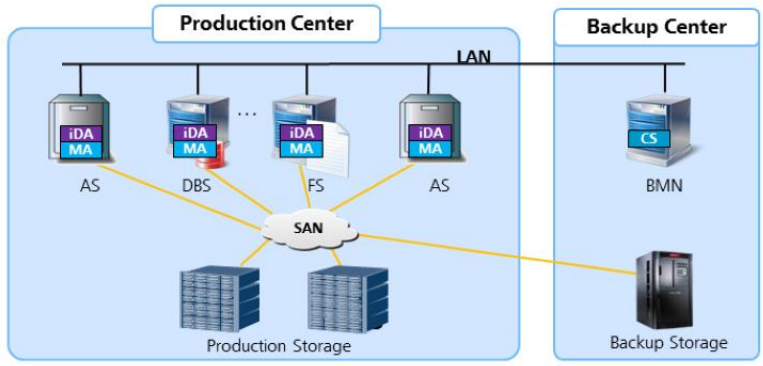
Huawei Training & Certification
Huawei Training & Certification



- LAN-Base: The backup data stream is transmitted over the existing LAN network.
- Advantages:
  - Make full use of existing networks and save investment.
  - Low equipment requirements.
- Disadvantages:
  - Takes up existing network bandwidth.
  - Backup performance is limited.
  - Has a certain influence on the host application.

Huawei Training & Certification

LAN - Free



BSN: Backup Service Node.
BMN: Backup Management Node.
AS: Application Server.
DBS: Database Server.
FS: File Server.

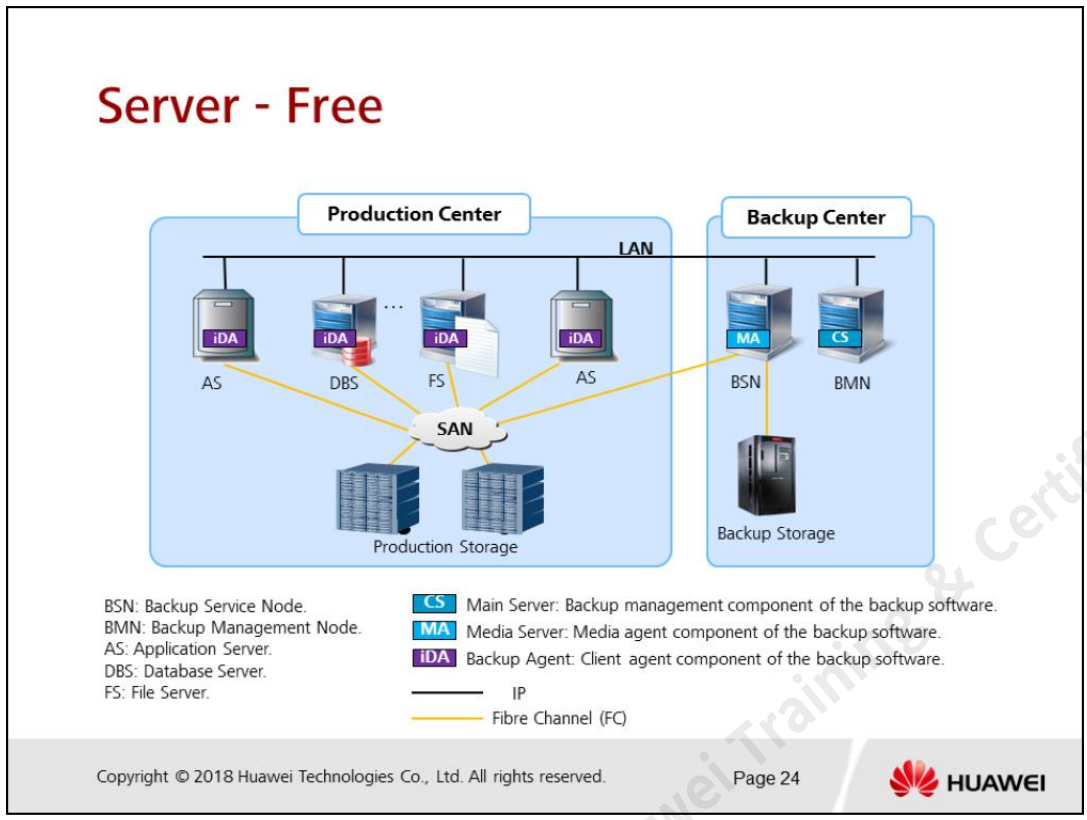
CS Main Server: Backup management component of the backup software.
MA Media Server: Media agent component of the backup software.
iDA Backup Agent: Client agent component of the backup software.
IP
Fibre Channel (FC)

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

- LAN-Free: The backup data stream is transmitted over the SAN network.
- Advantages:
  - Small impact on existing business networks.
  - Good backup performance.
- Disadvantages:
  - Larger investment in networks.
  - High requirements for equipment.

Huawei Training & Certification



- Server-Free: The backup data stream does not go through the production server but is transmitted over a separate network.
- Advantages:
  - Little impact on business hosts.
  - Little impact on existing business service networks.
  - Good backup performance.
- Disadvantages:
  - Larger investment in networks.
  - High demand for equipment.

Huawei Training & Certification

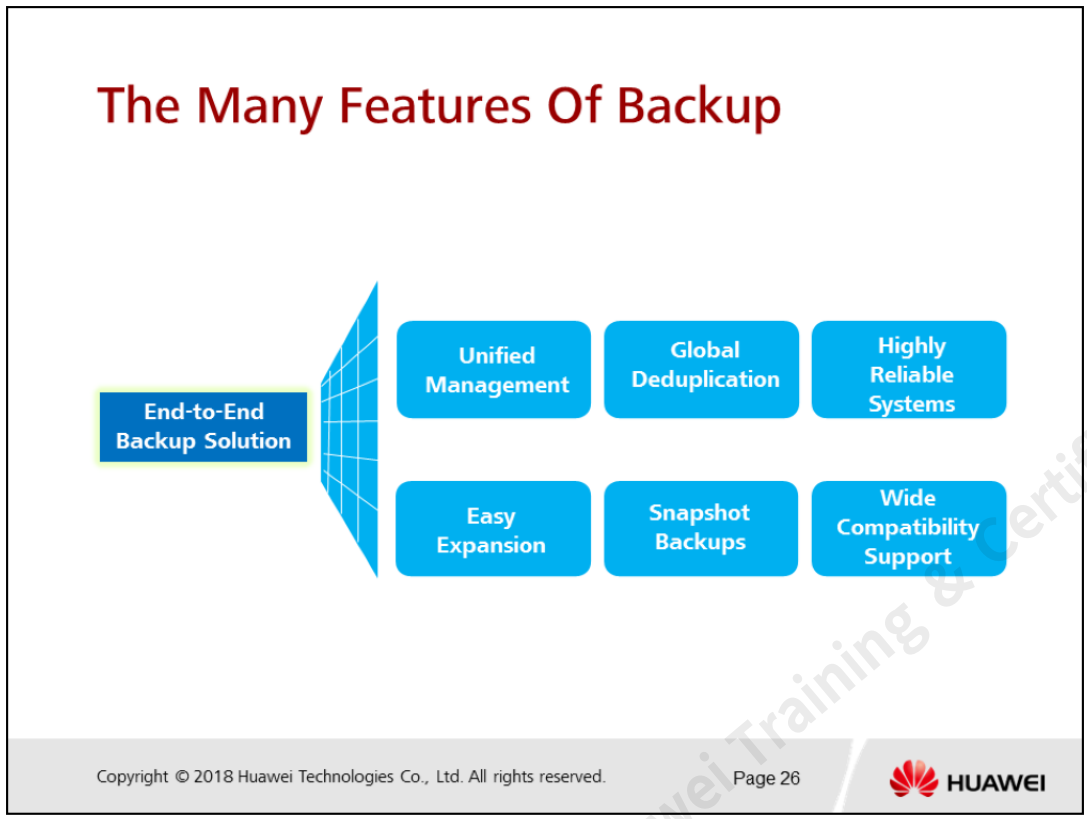
Contents

1. Overview Of Backup Solutions.
2. Architecture Of Backup Solutions.
3. Common Backup Networks.
4. **Common Backup Technologies.**
5. Application Cases Of Backup Solutions.

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

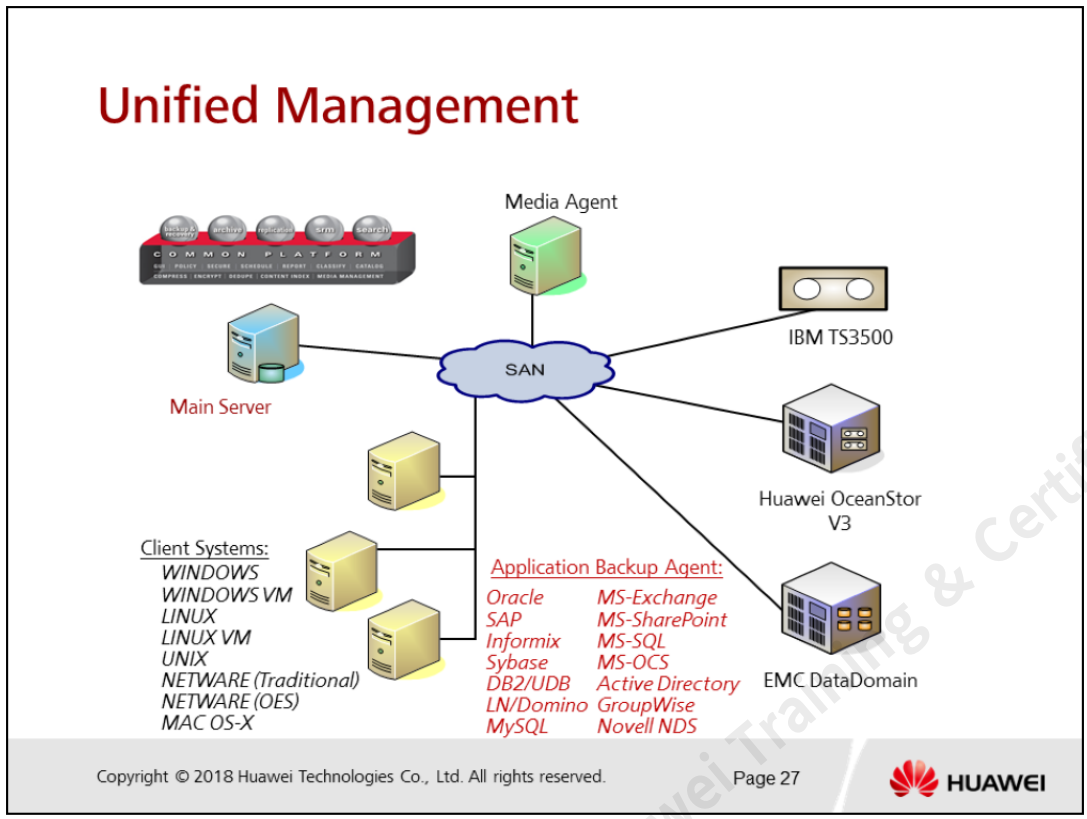
HUAWEI

Huawei Training & Certification
Huawei Training & Certification



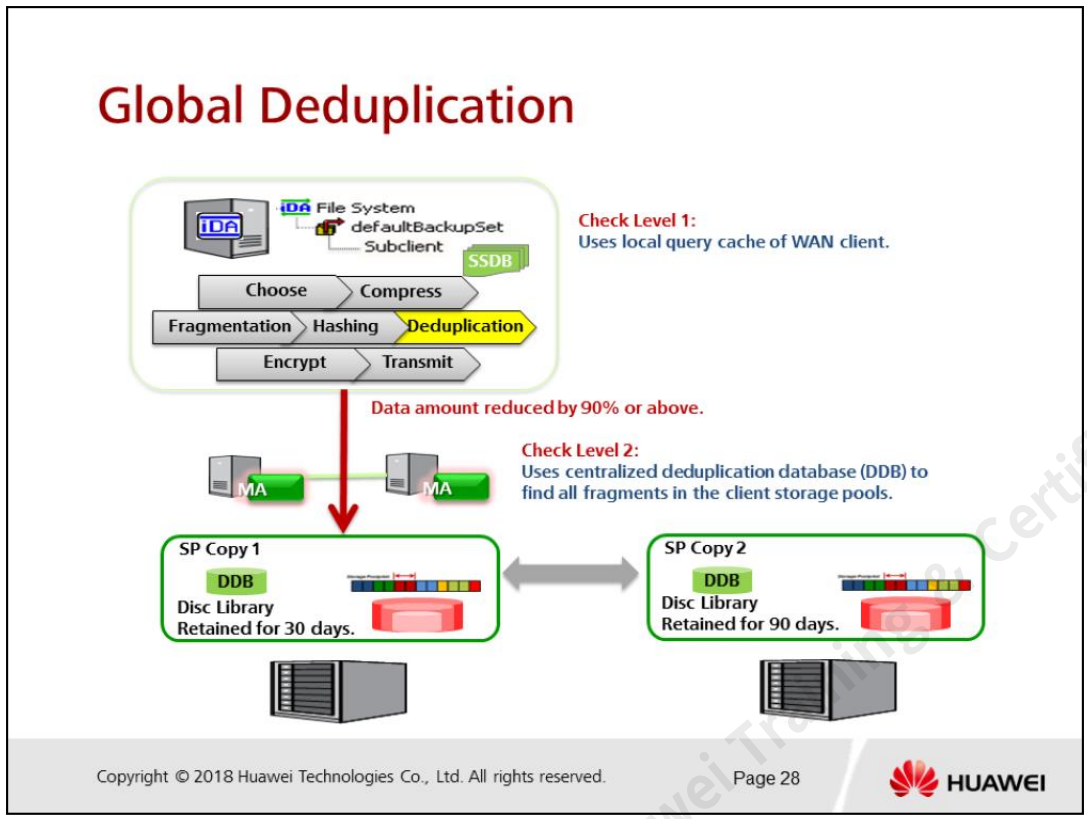
- A good backup solution has many desirable features that not only maximize the backup efficiency but also provides easy operation and maintenance.
- An end to end backup solution must have a unified management interface to easier management of multiple components and backup tasks, easily expandable, highly reliable and has wide compatibility support.
- Features such as global deduplication and snapshot backups allows the increase in backup efficiency and saves on the consumption of storage space for backups.

Huawei Training & Certification



- Centralized management targets includes the enterprise's main server, all media agents, storage devices, and client agents.
- The unified management includes:
  - Unified performance management, such as historical performance statistics and performance data export etc.
  - Unified monitoring and management, such as real-time monitoring of system status and query of resource service status etc.
  - Unified device management, such as physical device management, simple configuration wizards etc.
  - Unified resource management, such as unified storage resource management, unified management of physical disks and shared storage space etc.

Huawei Training & Certification

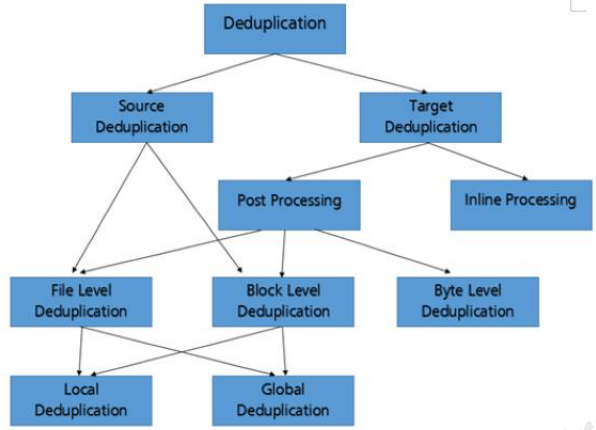


- Deduplication technology, in simple terms, refers to the technology that eliminates duplicate data. It uses software or hardware to process stored data to eliminate redundant data, thereby reducing the storage space occupied by data. .
- Implementation process of deduplication: The data that needs to be stored is compared with the hash unit by block, and the data block that is already stored previously is no longer stored again, and only the index is used to record the data block. The new data blocks that are not stored previously are physically stored and then indexed so that the same data blocks are physically stored only once. Through the index, you can see the complete logical view of the data, but in reality there is just little physical storage of data.
    - Raw data is stored on the media.
    - In the case of deduplication, comparisons are performed in units of a certain block size.
    - Only the unique data block is stored in the storage space after deduplication. The following data is compared with the unique data. Duplicate data is then deleted, and the Index and Metadata are stored in the specified space.
    - Data without duplicates are kept as the unique data in the storage space after deduplication, and the Index and Metadata are written at the same time.

Huawei Training & Certification

Classification of Deduplication Technologies

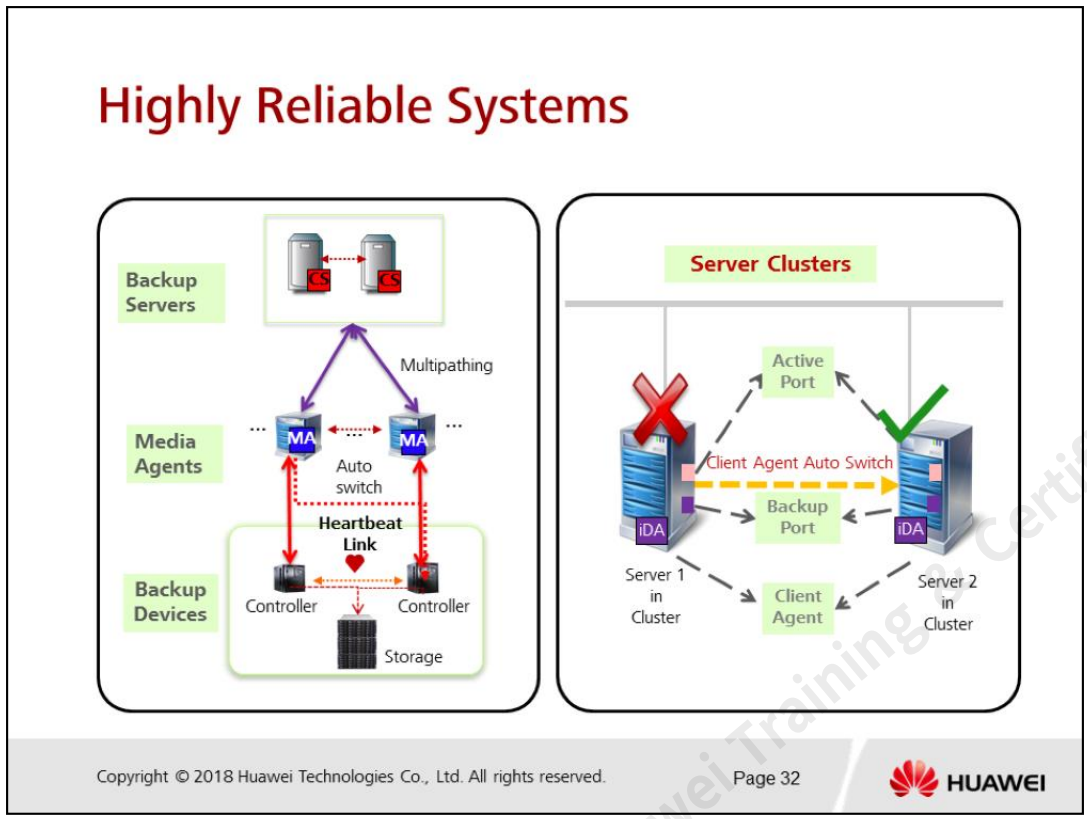
- The deduplication technology can be classified according to multiple dimensions such as location, time, granularity, and extent of deletion.



Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

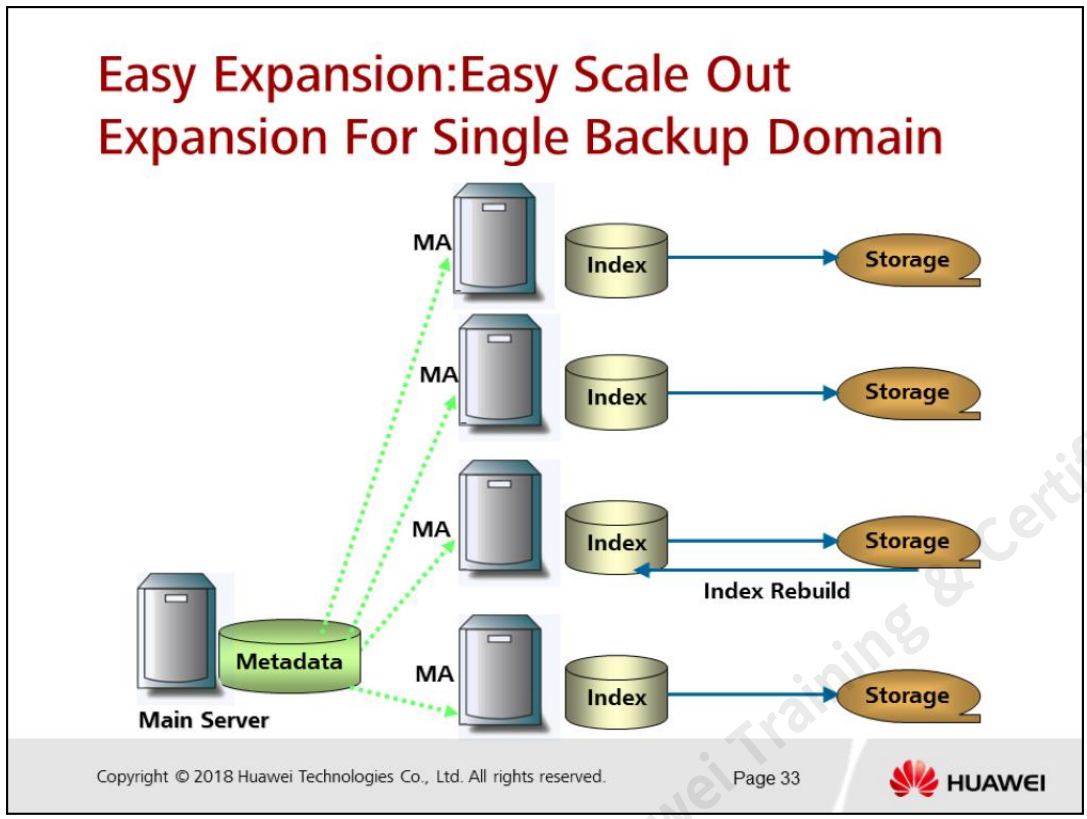
- The deduplication technology can be classified according to multiple dimensions such as location, time, granularity, and extent of deletion.
  - Source Deduplication: Delete duplicate data first and then sends the data to the backup device. Which means that, when the data is transmitted from the "source" to the "target", the transmitted data block is first hashed at the source end. If the data block has been previously transmitted, only the transfer hash is needed to be transmitted. If the data block has not been previously transmitted, then the data block is transmitted and the hash value of the data block is recorded. Advantage: It can save transmission bandwidth. Disadvantages: Occupy source resources for deduplication.

Huawei Training & Certification



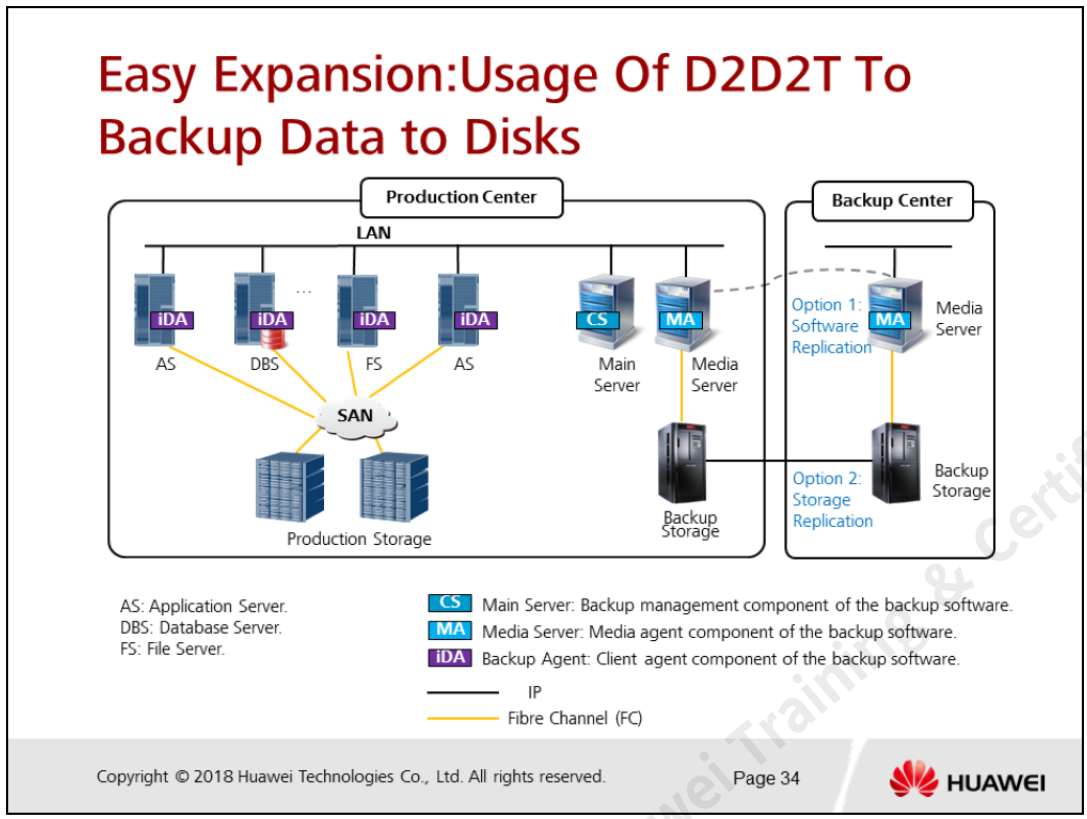
- All three layers of high availability design in the backup solution design ensures that the daily backup tasks can be completed without any system damage.
  - Backup server HA: Two backup servers can be deployed in HA mode to improve reliability.
  - Multipath in transmission media: All storage media agents work in Active-Active (A-A) mode and have multipath support for load balancing and switching.
  - Dual controllers: Dual controllers for storage devices make storage devices more stable and reliable.
- The backup client automatically switches to another server in the cluster through server switching.
  - Automatic switching: Automatically detect the status of the cluster and switches automatically.
  - Fast switching: Quickly restore backup jobs after the server switches.
  - Less maintenance work: No need to constantly monitor status of backup jobs, allows reliable backup, and less maintenance work.

Huawei Training & Certification



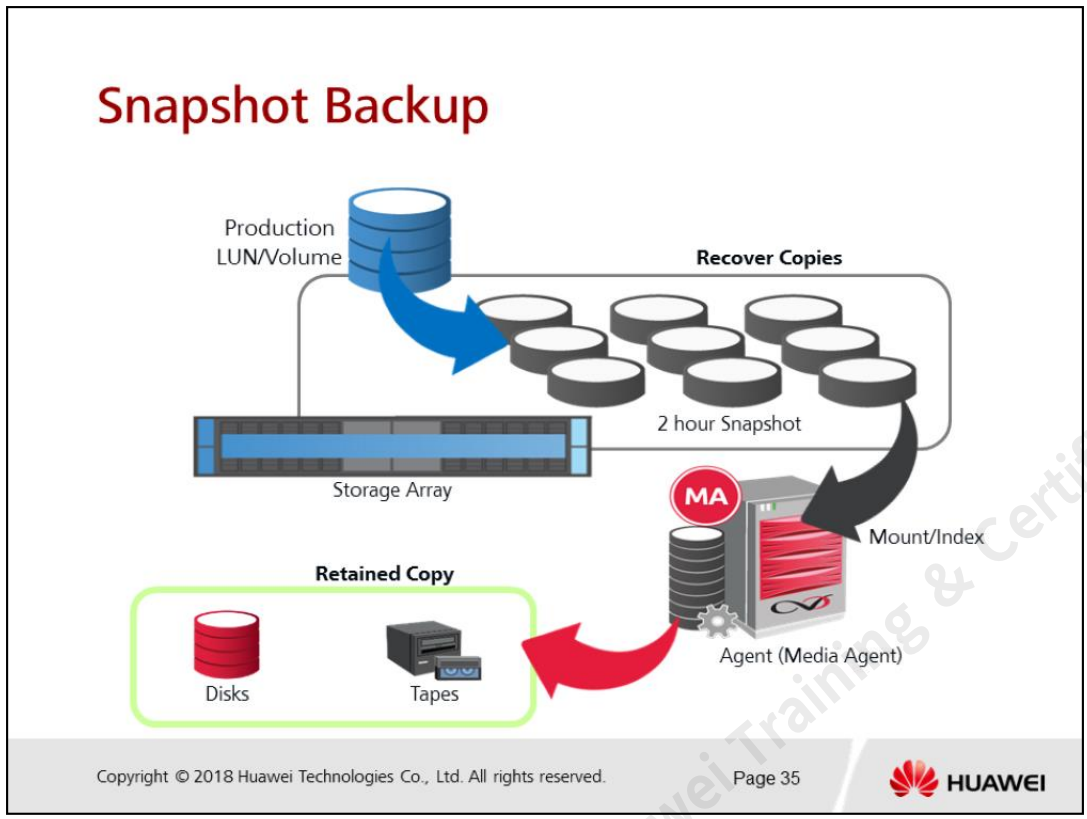
- Manage the entire data protection of the enterprise in a single domain. It has good scale-out capabilities. Media agents maintain granular index information to support millions of objects.
    - Centralized management: Uses split indexes to centrally manage all information. There is no extra maintenance cost for this process.
    - High flexibility: Media agents are easily expanded to support more data.
    - Large-scale: Supports 10,000 clients, about 100 times more than a centralized index.
    - Granular recovery: Granular recovery of individual objects which is suitable for VMs, files, Exchange, SharePoint, and Lotus Notes.

Huawei Training & Certification



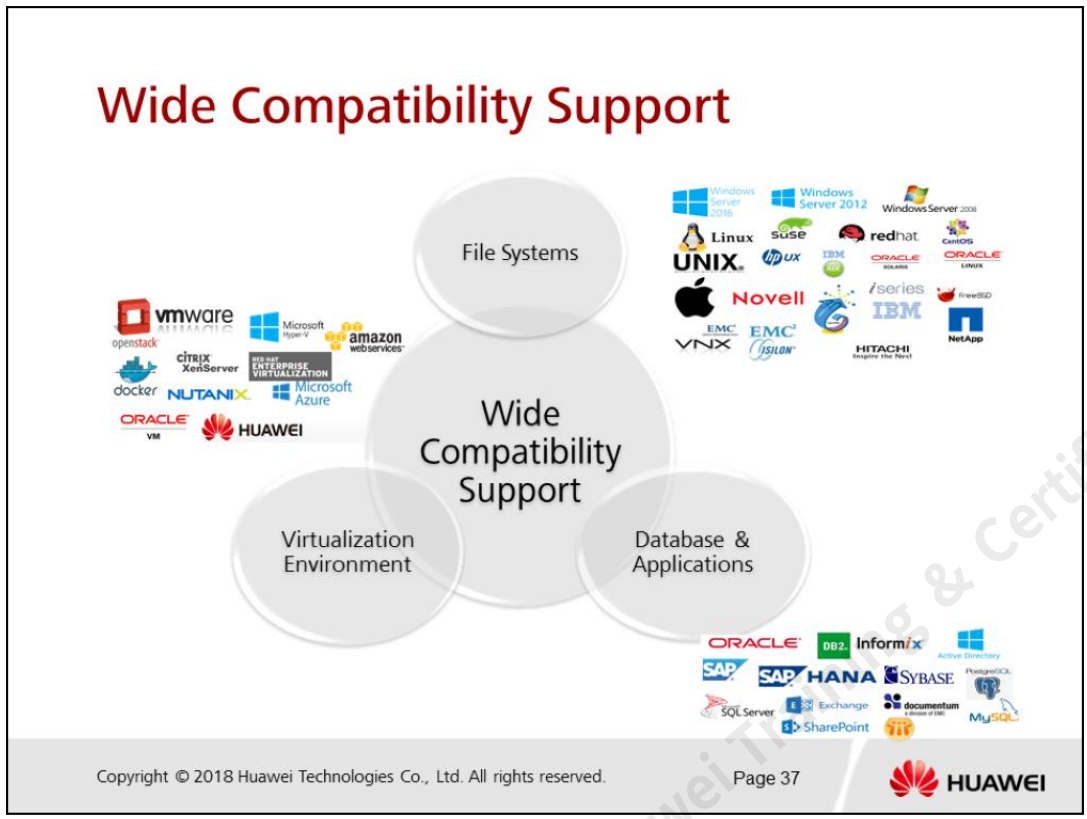
- Software Replication:
  - Advantages: Easy to manage, and supports recovery/restoration of backup data directly from backup center.
  - Disadvantages: Software replication requires a license.
- Storage replication:
  - Advantages: No software license required.
  - Disadvantages: Complex management.
- Disk-to-disk-to-tape (D2D2T) is an approach to computer storage backup and archiving, in which data is initially copied to backup storage on a disk storage system and then periodically copied again to a tape storage system.

Huawei Training & Certification



- Supports storage arrays with snapshot integration feature.
- Supports off-host agent-based backups.
- Equipped with fast recoverability:
  - Browse the contents of the backup to quickly restore the selected object.
  - Combine incremental backups into full backups in the background to quickly restore data.
  - Restoring data from hardware snapshots, and can also use snapshots for fine-grained recovery.
- Recovery copy:
  - Storage array protection: Supports native format, storage auto-discovery, its fully integrated (no scripts), and with snapshot support.
  - Storage array recovery: Restore volume, clone/mount, copy data back from backups.

Huawei Training & Certification



- For more information, please refer to : http://documentation.commvault.com/commvault/v11/article?p=landing_pages/c_backup_agents.htm.
- Wide File System compatibility support:
  - Windows: 2016, 2012, 2008, 2003, 10, 8, 7, Vista, XP.
  - Linux:
    - Asianux, Debian, Fedora, Gentoo, Mandriva, NeoKylin.
    - OpenSuSE, Red Flag, Red Hat, Scientific, Slackware.
    - Source Mage, SuSE , Ubuntu, Z-Linux, Oracle.
  - Unix: IBM AIX, HP-UX, Oracle Solaris, FreeBSD
  - Other OS: Mac, IBM iSeries, OpenVMS, Novell OES
  - Storage:
    - Dell, EMC, Hitachi, Huawei, IBM N-Series
    - NetApp, NextaStor, Nexsan, Oracle ZFS, Stratus

Huawei Training & Certification

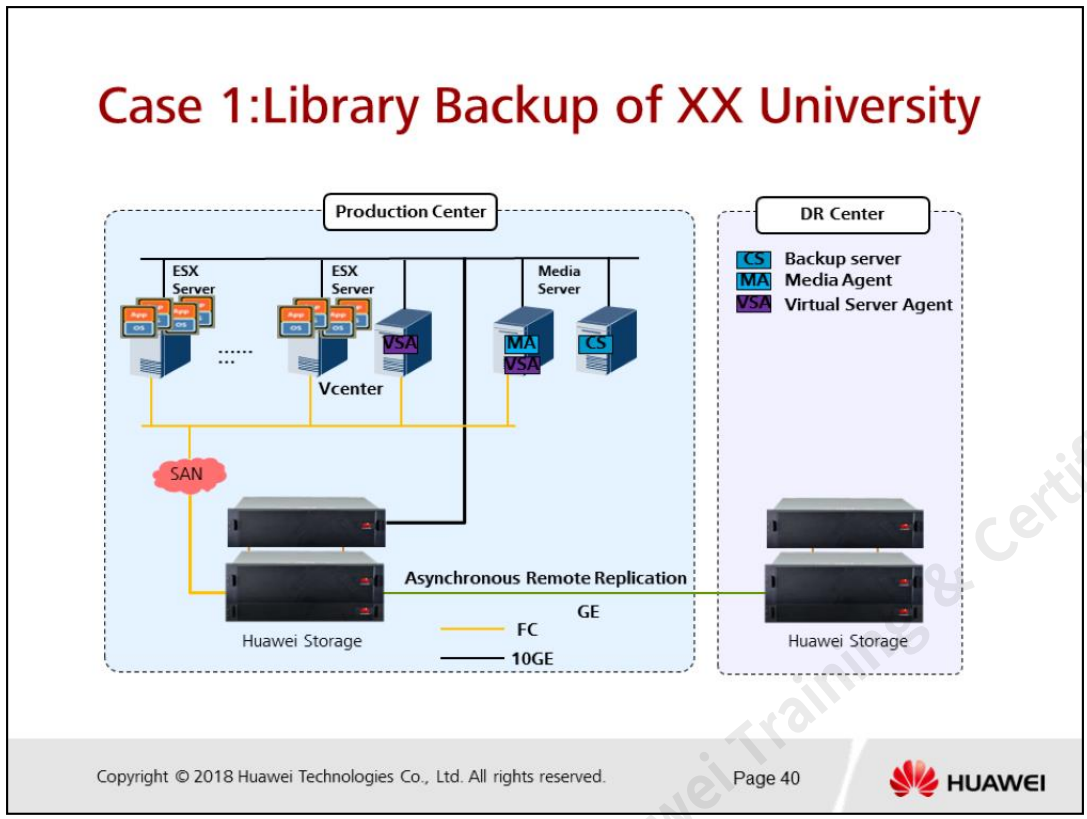
Contents

1. Overview Of Backup Solutions.
2. Architecture Of Backup Solutions.
3. Common Backup Networks.
4. Common Backup Technologies.
5. **Application Cases Of Backup Solutions.**

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

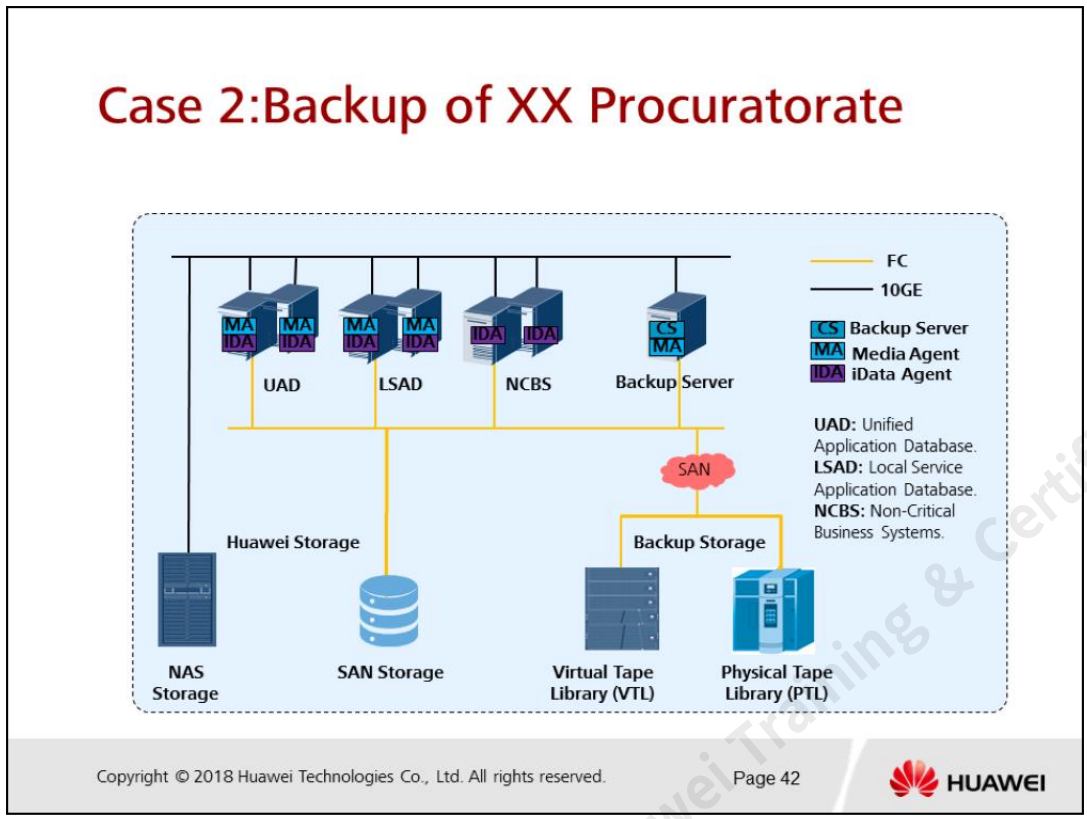
HUAWEI

Huawei Training & Certification
Huawei Training & Certification



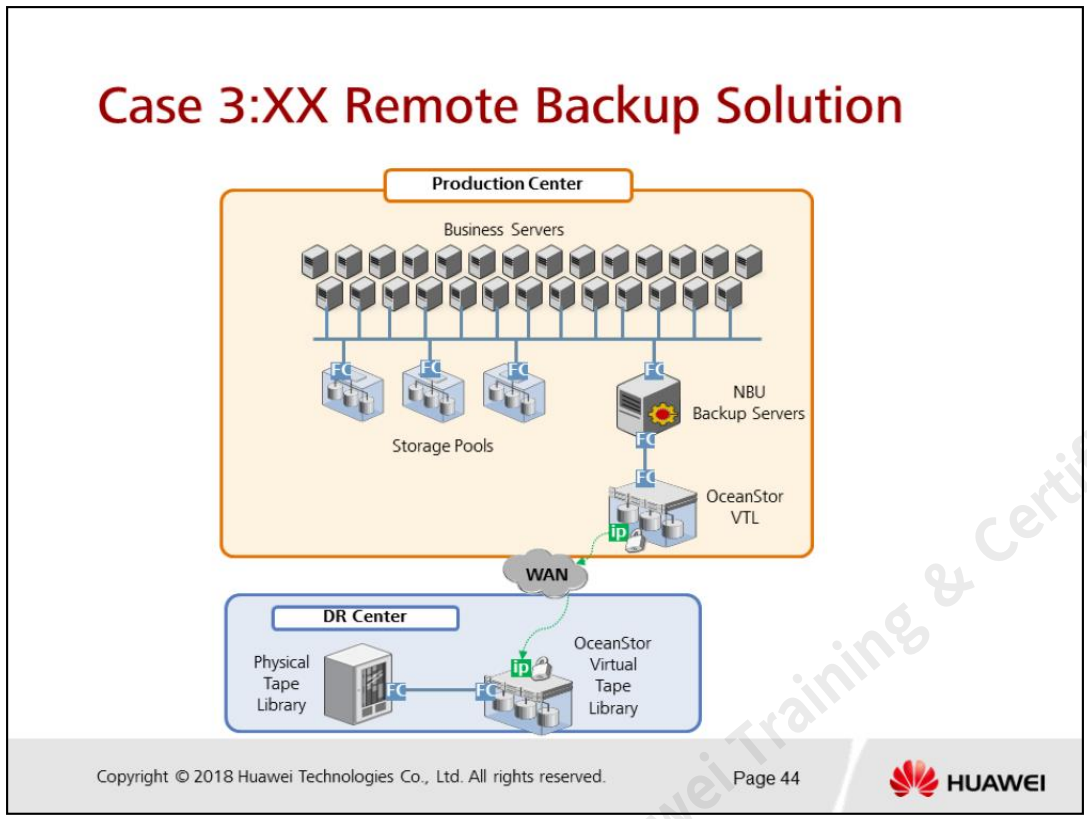
- Business Challenges:
  - The existing network applications are based on virtual machine platforms and have low backup performance.
  - There are a large number of library books and there is a lack of a unified backup platform.
  - An off-site disaster recovery center needs to be established to ensure the security of backup data.
- Solution:
  - Use the LAN_Base networking mode to deploy and use the Enterprise centralized backup solution.
  - Simpana is used as the backup software. Huawei storage is used as a backup medium for local and offsite disaster recovery centers.

Huawei Training & Certification



- Business Challenges:
  - Prosecution service data has grown rapidly and management work has been difficult.
  - Lack of unified service application platform for service backup.
  - It is hoped that a backup system will be built to enable the service backup platform to exchange data with the superior and the subordinate units.
- Solution:
  - Build a unified backup platform, use Simpana as backup software, and VTL6900 as a backup medium.
  - Deploy LAN_Base and LAN_Free hybrid networking to maximize the usage of existing network resources.
  - Use deduplication technology to reduce storage space consumption.

Huawei Training & Certification



- Business Challenges:
  - Existing DAS storage architecture, storage expansion difficulties, and high management costs.
  - No centralized backup system, requires manual operation, and low reliability.
- Huawei Solution:
  - Production Center deploys a NBU + VTL Backup System.
  - Disaster Recovery Center deploys a VTL(Virtual Tape Libraries).
  - Remote backup and data replication for disaster recovery between the VTL of production center and disaster recovery center through IP network.

Huawei Training & Certification

Summary

- This module mainly introduces the contents of the following topics:
  - Concepts, Policies and Key Technologies of Backup Solutions.
  - Architecture and Components of Backup Solutions and the Structure of Backup Systems.
  - The Common Backup Networking in Backup Solutions.
  - The Typical Application Cases of Backup Solutions.

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

Huawei Training & Certification
Huawei Training & Certification

Quiz

1. (True or False) Backup is the same as Archive. ( )
2. (Multiple Choice) Which of the following are components of Backup Systems? ( )

   A. iDA B. MA C. BA D. CS
3. (Single Choice) Which of the following does not belong to the networking of backup solutions? ( )

   A. LAN-Free B. SAN-Base C. LAN-Base D. Server-Free

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

- Answers:
  - False.
  - ABD.
  - B.

Huawei Training & Certification
Huawei Training & Certification

Thank You

www.huawei.com

Huawei Training & Certification
Huawei Training & Certification

11 Disaster Recovery Technology and Its Applications

www.huawei.com

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

Huawei Training & Certification
Huawei Training & Certification

Foreword

- This module mainly introduces:
  - The definition of Disaster Recovery system and its importance towards Enterprise Business Continuity.
  - Common Disaster Recovery Solutions.
  - Common Disaster Recovery Technologies.
  - Success Story of Disaster Recovery Solutions.

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

Huawei Training & Certification
Huawei Training & Certification

Objectives

- Upon completion of this module, you will be able to:
  - Describe the concepts and importance of Disaster Recovery.
  - Differentiate the Advantages and Disadvantages of common Disaster Recovery solutions.
  - Understand the technical principles of Disaster Recovery solutions.
  - Learn how to deploy a Disaster Recovery solution through case study of typical application of Disaster Recovery Solution.

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

Huawei Training & Certification
Huawei Training & Certification

Contents

1. **Overview of Disaster Recovery Solution.**
2. Architecture of Disaster Recovery Solution.
3. Common Disaster Recovery Technologies.
4. Case Study of Disaster Recovery Application.

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

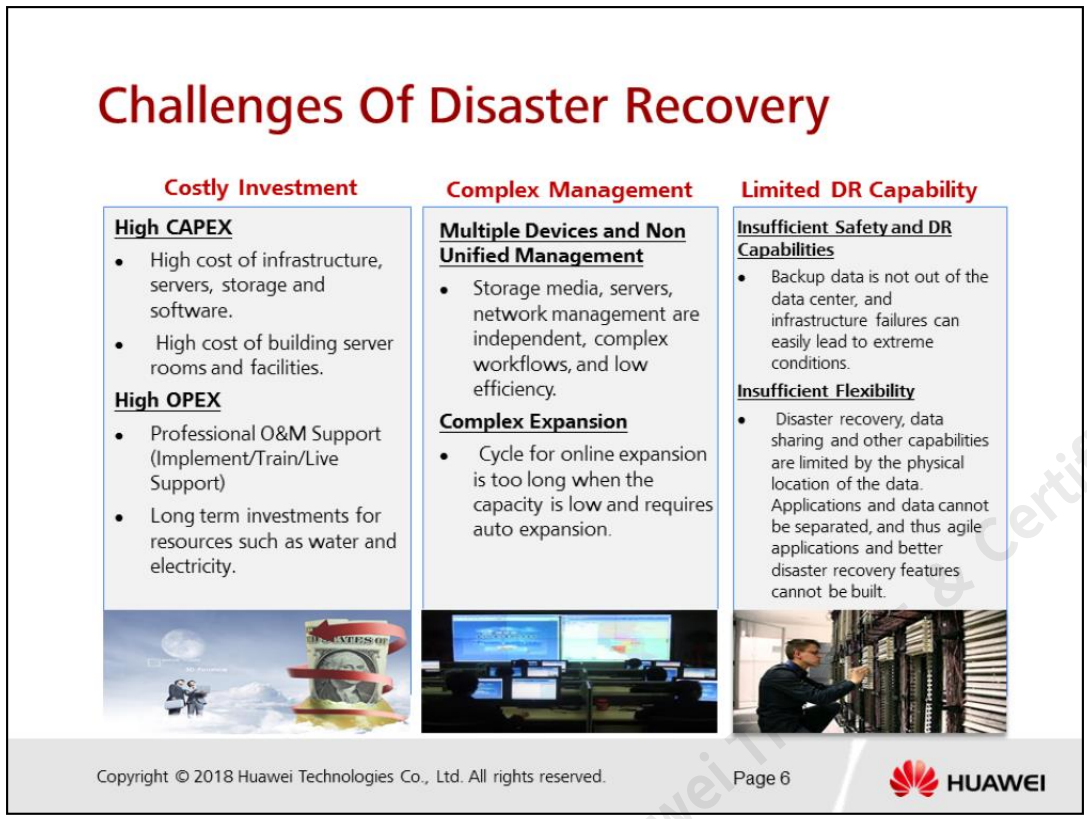
HUAWEI

Huawei Training & Certification
Huawei Training & Certification



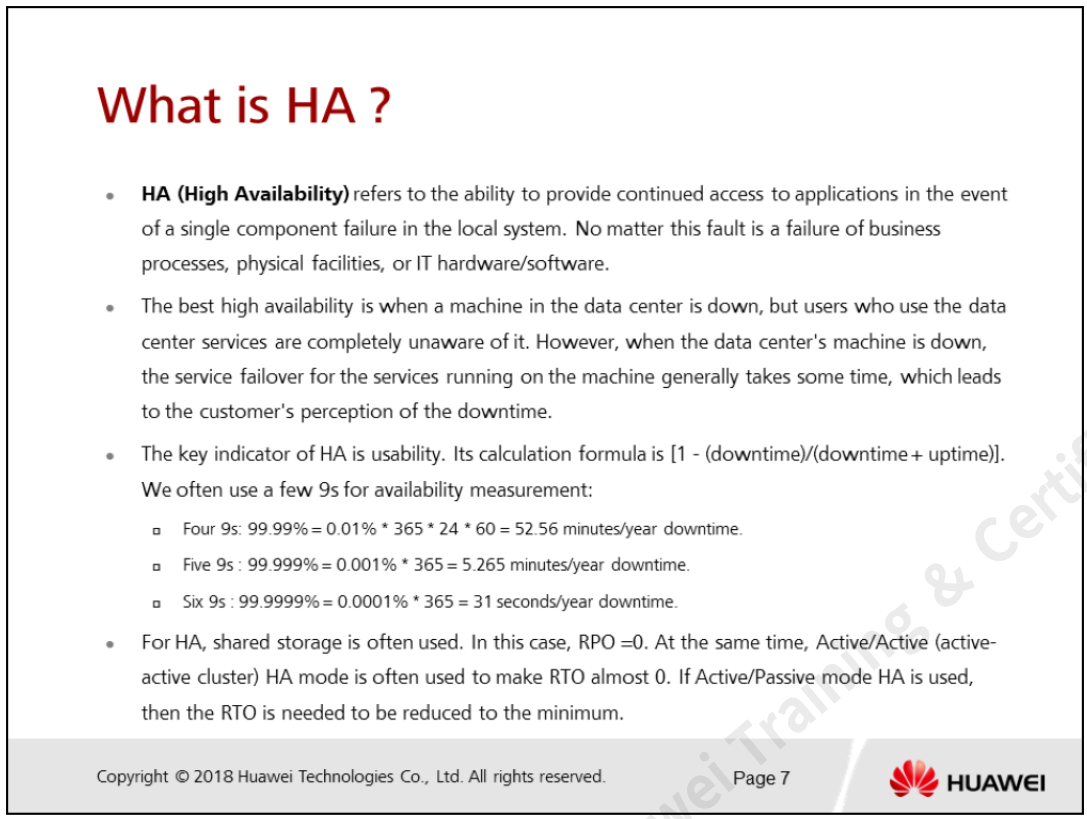
- The necessity of disaster recovery system construction: Disasters are everywhere and can happen at anytime.
- From the analysis of statistics on insurance from a respected organization in Switzerland:
  - Global direct losses due to natural disasters and human accidents reached US$123 billion in 2004/
  - About 400 catastrophes occurred in the world in 2005, with losses exceeding US$230 billion.
  - In 2006, the losses caused by natural disasters and man-made disasters were lower than the long-term trend, but still with the impact of a direct loss of 48 billion U.S. dollars.
  - Globally, compared with the 1960s, the number of catastrophic natural disasters in the world has increased threefold and the economic loss has increased by 9 times up till the 1990s.

Huawei Training & Certification



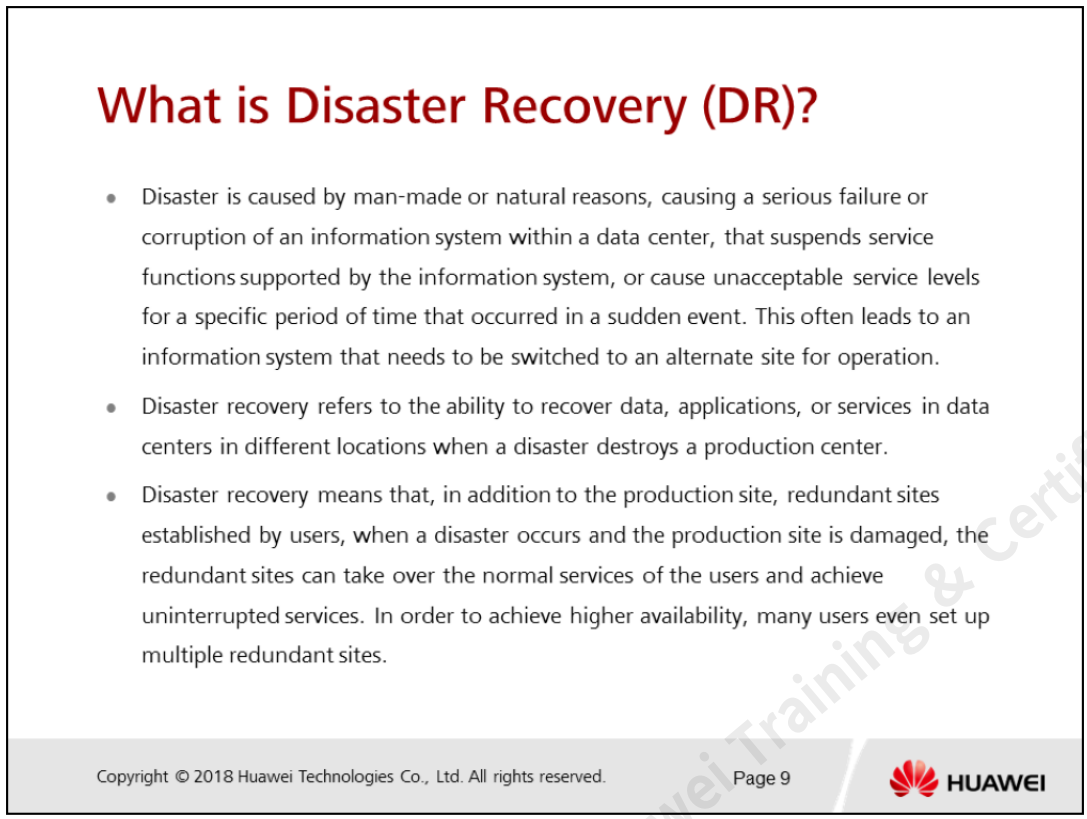
- Various number of applications and inconvenient management: More and more business systems are running in enterprise IT systems, and increasing number of applications are required for disaster recovery protection as they are part of the key business services. Common applications include Oracle, DB2, SQL Server, Exchange, etc. The cloudification trend of IT systems is becoming more obvious. Many virtual machines also need to be protected by the DR plan. Additionally, there is a lack of a unified management system.
- DR process is complicated, time-consuming and error-prone: Different application have different configurations, and their recovery processes are also different making overall DR configuration difficult. On the other hand, service failover and recovery require operations by trained professionals, which is time-consuming and error-prone. It also lacks of automated creation and deployment processes.
- Black box operation is difficult to understand: Traditional service failovers, drills and other operational processes are black box operations that occurs within the equipment, and with the lack of visual effects making it difficult to understand.

Huawei Training & Certification



- HA needs to use redundant servers to form a cluster to run the workloads, which includes applications and services. This type of redundancy also divides HA into two categories which are:
- Active/Passive HA:
  - The cluster only includes two nodes, these nodes are referred in short as master and backup. In this configuration, the system uses primary and standby machines to provide services, and the system only provides services on the primary device.
  - When the master device fails, the service on the standby device is started to replace the service provided by the master device.
  - Typically, CRM software such as Pacemaker can be used to control the switching between active and standby devices and provides a virtual IP to provide services.

Huawei Training & Certification



- Disaster recovery (DR) is an area of security planning that aims to protect an organization from the effects of significant negative events. DR allows an organization to maintain or quickly resume mission-critical functions following a disaster.
- A disaster can be anything that puts an organization's operations at risk, from a cyberattack to equipment failures to natural disasters. The goal with DR is for a business to continue operating as close to normal as possible. The disaster recovery process includes planning and testing, and may involve a separate physical site for restoring operations.

Huawei Training & Certification

The Relationship between HA and DR (1)

- The two are related to each other, complement each other, and intercross with each other, but at the same time have with significant differences between them:

Dimension	HA(High Availability)	DR (Disaster Recovery)
Scenarios	HA refers to the local high availability system, which means that in the case of multiple servers running one or more applications, it should ensure that when any of the server fails, the application it runs won't be interrupted, and the application and system should be able to quickly switch over to operate on another server, which resides within the local system cluster and hot backup.	DR refers to highly available systems in different places (in the same city or different places), indicating the ability to recover data, applications, and services in the event of a disaster.

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

- High Availability(HA) is the process in which a system is supposed to take over when another system stops working efficiently or at all. For IT, this must occur with as minimal downtime as possible—so minimal that most users don't even know there was a problem. Data loss must also be negligible; if there is a service issue, HA takes over so that whatever employees are working on isn't lost. No system is perfect—as each has occasional hiccups, scheduled maintenance, peak usage times, upgrades, minor outages, and so on. High availability accounts for these events and guarantees they don't cut into standard operations and productivity.
- Although disaster recovery also is concerned with downtime and data security, it focuses on bigger threats, surprise crises, and potentially longer outages: beyond the normal recovery point objective (RPO) and recovery time objective (RTO) of high availability. DR deals with the unexpected—events that weren't scheduled, were apart from the norm, and often carry longer-lasting consequences. Good disaster recovery doesn't mean companies are completely caught off guard, but rather, that they are trained and prepared to respond to an event, according to set plans, in order to restore normal operations quickly, whether the systems they turn to are in house, at another facility, or in the cloud (or a combination thereof).

Huawei Training & Certification

The Relationship between HA and DR (2)

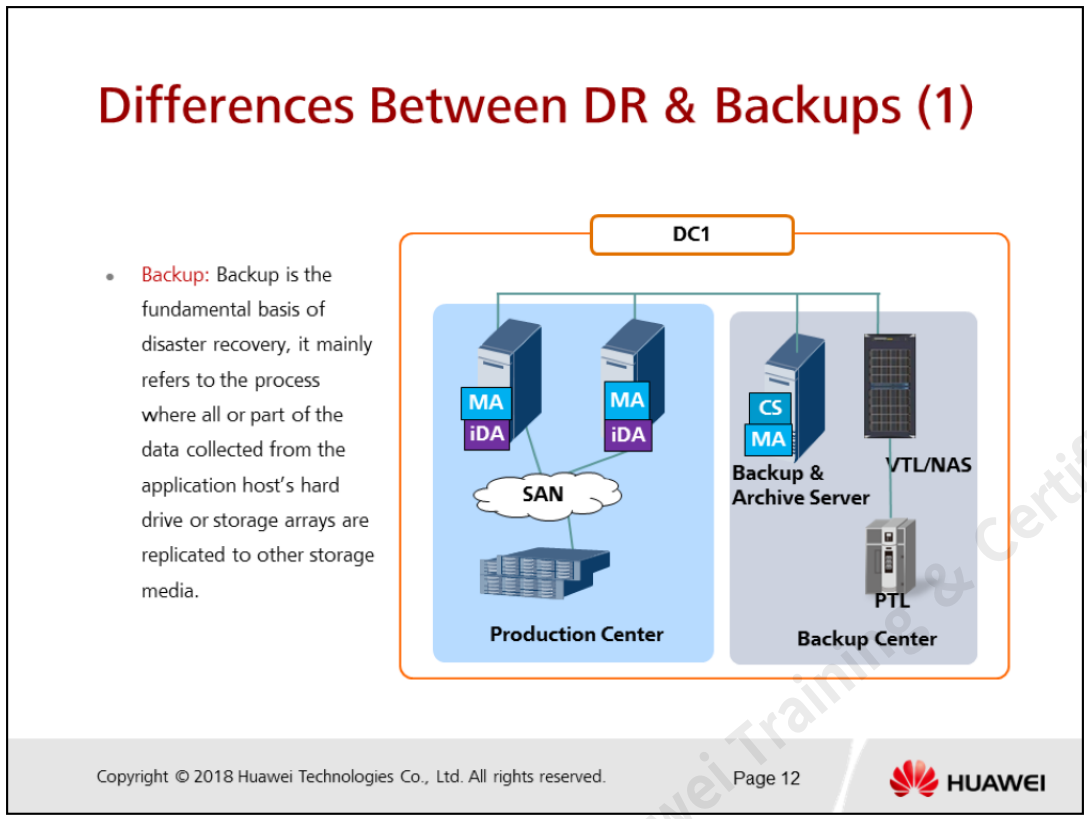
Dimension	HA(High Availability)	DR(Disaster Recovery)
Storage	HA often uses shared storage, which means there is no lost of data (RPO=0), more considerations are on the RTO due to the time period taken for service switching/failover.	Off-site disaster recovery data recovery uses data replication. According to the different data replication technologies used (synchronous and asynchronous), data loss often results in RPO > 0; while off-site application switching often takes longer, so RTO > 0.
Faults	Mainly handles single-component failures causing the load to switch between servers in the cluster.	Copes with large-scale failures causing the load to switch between data centers.
Network	LAN scaling tasks are within the category of HA.	WAN scaling tasks are within the category of DR.
Cloud	HA is a mechanism for ensuring business continuity within a cloud environment.	DR is a mechanism for ensuring business continuity among multiple cloud environments.
Objective	HA's main purpose is to ensure that the business/service is highly available.	DR is a business continuity solution that is based on reliable data protection.

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

- RPO is the maximum age of files that an organization must recover from backup storage for normal operations to resume after a disaster. The recovery point objective determines the minimum frequency of backups. For example, if an organization has an RPO of four hours, the system must back up at least every four hours.
- RTO is the maximum amount of time, following a disaster, for an organization to recover files from backup storage and resume normal operations. In other words, the recovery time objective is the maximum amount of downtime an organization can handle. If an organization has an RTO of two hours, it cannot be down for longer than that.
- The RPO and RTO helps administrators choose optimal disaster recovery strategies, technologies and procedures.

Huawei Training & Certification

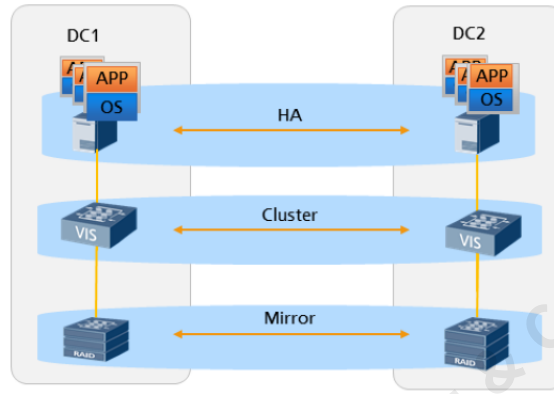


- Backup solutions are designed to provide quick data access in the event of a sudden, small-scale computing outage, such as the loss of a laptop, accidental file deletion or crashed hard drive. At their most basic form, backup could be just the process of simply copying critical files to a second storage environment. This could be a tape backup, secondary in-office workstation, or hosted cloud system. Backups excels at providing immediate access to critical files and quickly restoring data. Thus, reducing the need to rely on time-consuming OS re-installations or manually retrieving and copying files from long-term storage solutions.
- Disaster recovery, meanwhile, is designed to provide a large-scale action plan for restoring critical data and systems in the event of a serious computing outage. While the term "disaster" often calls up images of massive storms, earthquakes or tornados, a disaster covers anything that causes widespread network outages or prevents your employee from completing necessary tasks. Disaster recovery solutions typically cover the entire process necessary to restore system function — from re-establishing network connections to repopulating data to establishing recovery time objectives (RTOs).
- While backups form a critical part of your disaster recovery plan, they're not enough on their own and need to work with a proper DR plan to ensure business continuity in the event of a disaster.

Huawei Training & Certification

Differences Between DR & Backups (2)

- **Disaster recovery:** Disaster recovery system refers to the establishment of two or more sets of IT systems with the same function in remote locations separated from each other. Health status monitoring and function switching can be performed between them. When a system stops working due to the event of disaster (such as fire, earthquake), the entire application system can be switched to another place so that the system function can continue to work normally.



- Generally, disaster recovery refers to backup of data or application systems that are not in the same server room or location. Backup refers to local data or system backups.
- In general, DR combines disaster recovery with backup, which is, local backup combined with remote data replication to achieve perfect data protection.

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

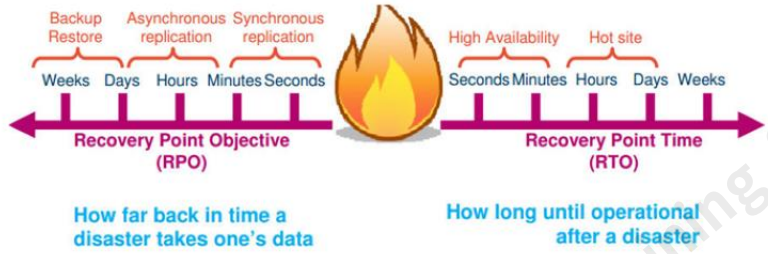
HUAWEI

- Generally, backup is implemented using backup software technology, and disaster recovery is implemented through replication or mirroring software. The fundamental difference between the two is:
  - The format of the data processed by the backup software is inconsistent and must be restored before the data can be used. The data format processed by the replication or mirroring software does not change and thus can be directly mounted to the host.
  - The data protection cycle of the two is inconsistent, and the time period for copying or mirroring is shorter.
  - Generally, backup is the last line of defense for data protection and it is more biased towards archiving data.

Huawei Training & Certification

Key Indicators For Measuring DR System Performance

- RPO (Recovery Point Objective) represents the amount of data that is allowed to be lost when a disaster occurs.
- RTO (Recovery Time Object) represents the system recovery time.
- The smaller the RPO and RTO, the higher the availability of the system and the greater the investment required by the user.



Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

- Developing a good continuity plan means deciding upon how much data your business needs to protect, and how much time it can afford to be without it.
- These distinct volume sets are referred to as your Recovery Point Objective (RPO) and Recovery Time Objective (RTO), respectively. Creating a comprehensive business continuity plan requires setting tolerances for each of them.
- Learning the difference between RPO and RTO, and setting realistic tolerances for each, is one of the most important steps you can take to ensure business continuity in the event of a data disaster.

Huawei Training & Certification

Levels Of DR Systems (1)

Level	Definition	RTO	TCO
Data Level	Through the establishment of remote disaster recovery centers, remote backup of data will be performed to ensure that the original data will not be lost or destroyed after the disaster. In the data-level disaster recovery mode, the established disaster recovery center can be simply considered as a remote data backup center. At the data-level disaster recovery level, applications are interrupted when a disaster occurs. Data-level disaster recovery has a long recovery time, but its cost is lower than other disaster recovery levels, and its implementation is relatively simple. Data sources are the source of life for all critical business systems, so data-level disaster recovery is essential.	The longest RTO (several days) because after the disaster, the equipment needs to be redeployed to recover the business services using backup data.	Lowest

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

- The table above shows the detailed description of data level disaster recovery and its expected RTO and TCO.
- Different level of DR has their own method of implementation and they cover different aspects of DR which are data, application and service.
- It is important to formulate the DR plan of your business based on the DR requirement on each level of DR, investment cost and expected RTO/RPO.

Huawei Training & Certification

Levels Of DR Systems (2)

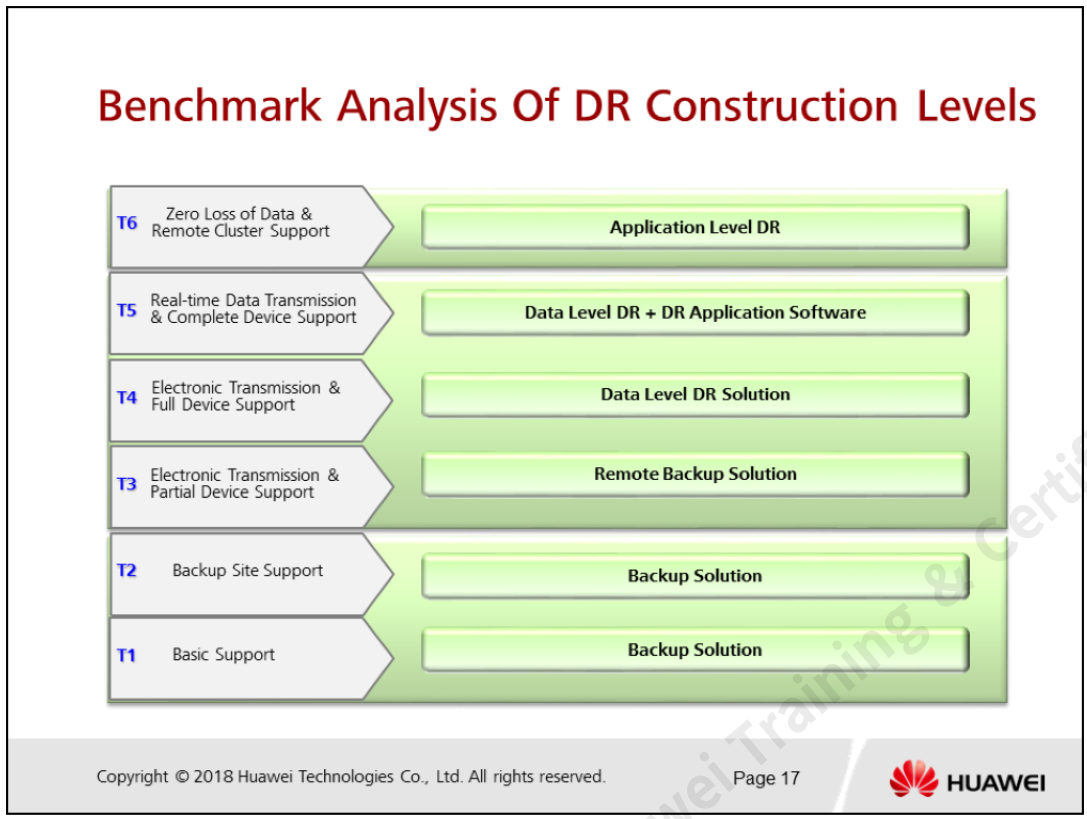
Level	Definition	RTO	TCO
Application Level	On the basis of data-level disaster recovery, similar sets of application systems are also built at the backup site. Through synchronous or asynchronous replication technology, key applications can be guaranteed to recover within the allowable time range, minimizing the loss caused by disasters, and allows users to be unaware of the disaster. Thus, the services provided by the system are complete, reliable and safe.	Medium RTO (Few Hours)	Medium. Same or similar system or subsystems is built offsite.
Service Level	In addition to the necessary IT-related technologies, full-service disaster recovery requires full infrastructure. Most of its components are non-IT systems (such as telephones, office locations, etc.). When a catastrophe occurs, the original office space will be destroyed. In addition to the restoration of data and applications, a back-up workplace needs to be provided so business can be conducted as usual. .	Smallest RTO (Few Minutes or Seconds)	Highest

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

- The table above shows the detailed description of application level and service level disaster recovery and its expected RTO and TCO.
- Different DR level has different performance in terms of service failovers in the event of the disaster. Although higher performance (low RTO, high RPO) DR level ensures greatest amount of business continuity protection to the enterprise, it is important to consider the amount of investment required for implementation.
- It is best to choose the DR plan that fits the current scale of business with some rooms for growth and improve your DR plan along with the growth of your business.

Huawei Training & Certification



- The diagram above shows the different level for the benchmarking of the DR construction levels.

Huawei Training & Certification

Panorama Of DR Solution

Local High Availability Solution

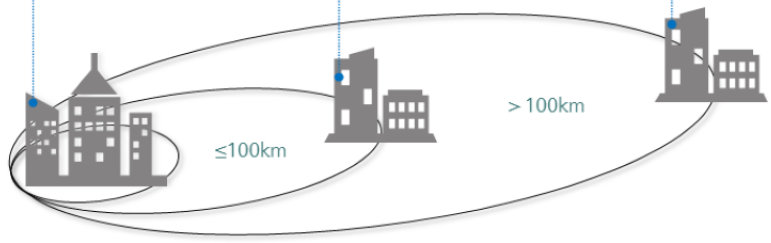
- Local HA solution.

Same City DR Solution

- Active-Active DC Solution.
- Active-Standby DR Solution.

Remote Region DR Solution

- 3DC DR Solution.
- Active-Standby DR Solution.



Local Production Center

Same City DR Center

Remote Region DR Center

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

Huawei Training & Certification
Huawei Training & Certification

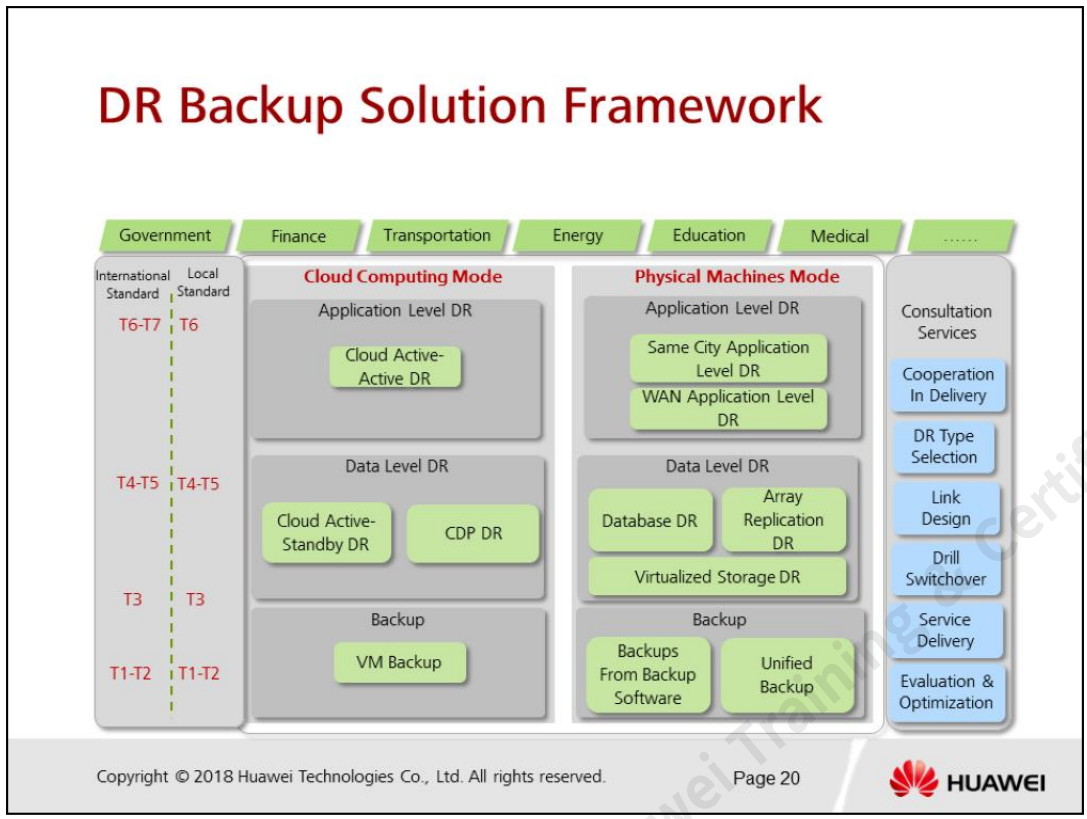
Contents

1. Overview of Disaster Recovery Solution.
2. **Architecture of Disaster Recovery Solution.**
3. Common Disaster Recovery Technologies.
4. Case Study of Disaster Recovery Application.

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

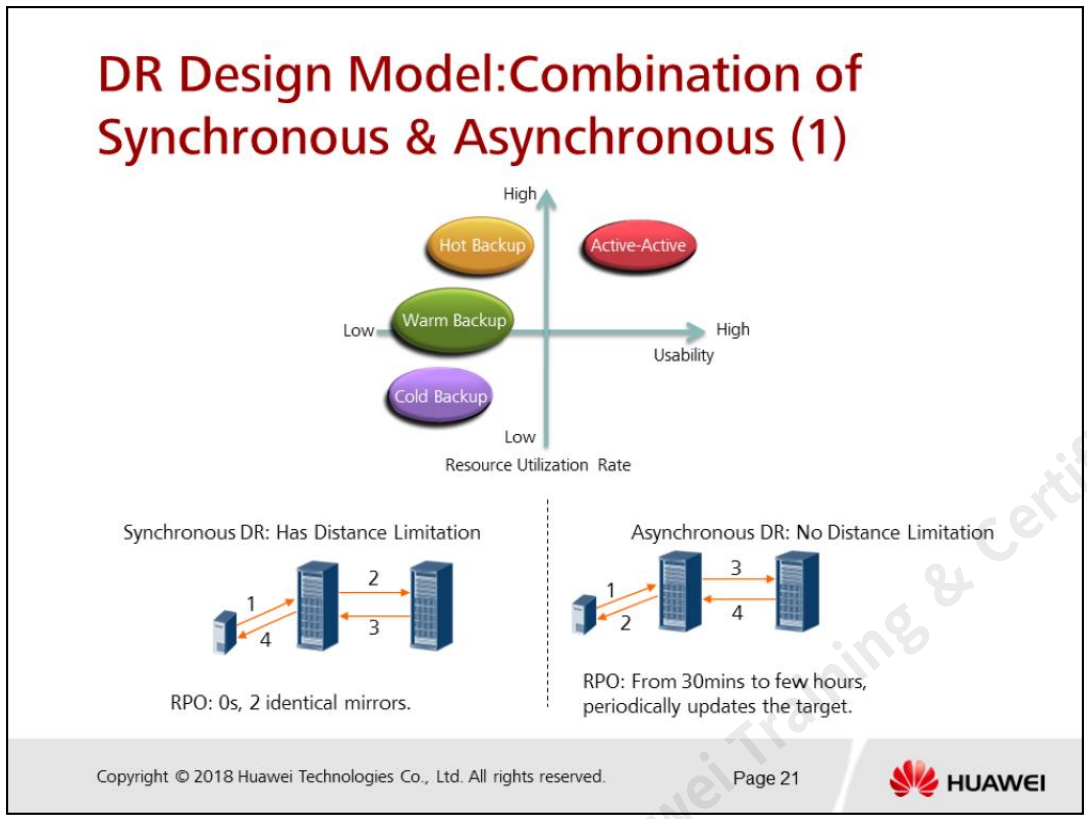
HUAWEI

Huawei Training & Certification
Huawei Training & Certification



- Match customer business and development strategies, and provide professional services from strategic consulting, disaster recovery planning, and service implementation for continuous operation and management.

Huawei Training & Certification



- One of the important factors to consider when constructing DR solution is to choose where and when to implement the synchronous and asynchronous DR technologies.
- Both synchronous and asynchronous DR has its challenges and limitations in implementation and requires careful planning and commissioning.
- Both these technologies can be implemented on its own for DR, but it can also be combined to provide better results and protection of the business data.

Huawei Training & Certification

DR Design Model:Combination of Synchronous & Asynchronous (2)

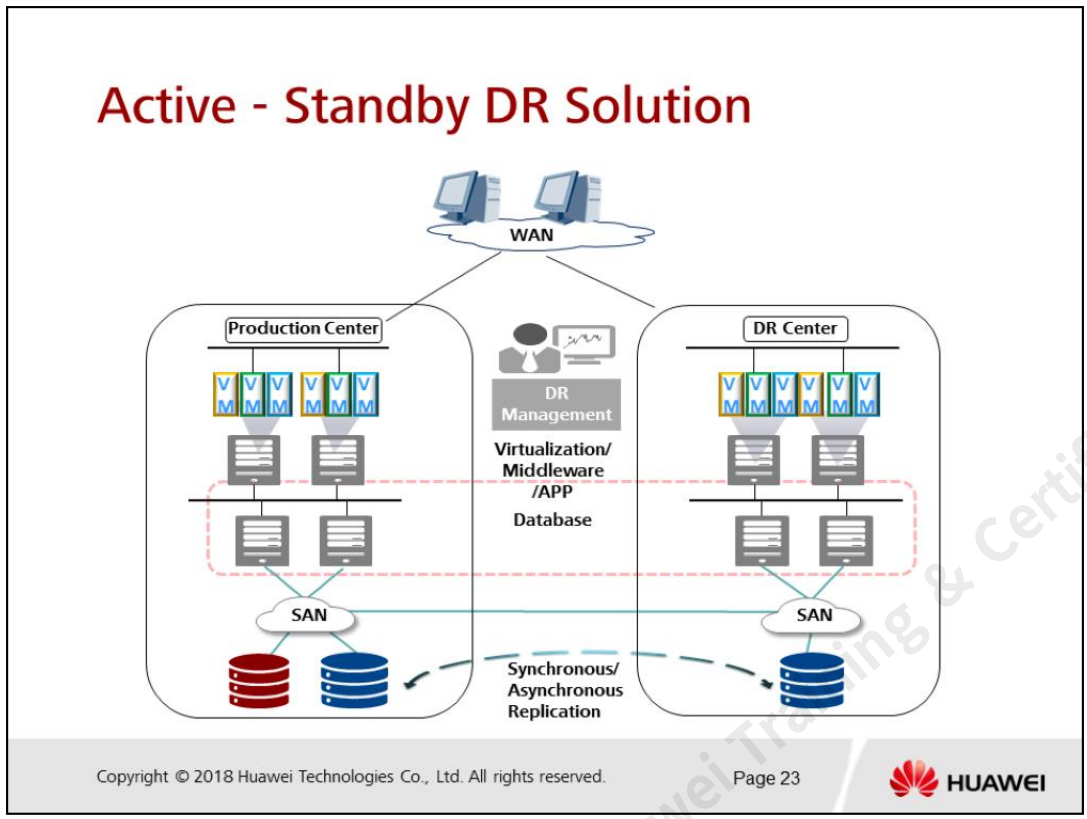
DR Mode	Reliability Solution	DR Recovery	Data Backup Requirements
Active-Active	Cluster + Load Balancing	Automatic	Real-time synchronous replication (<100KM)
Hot Backup	Cluster	Automatic	Real-time synchronous replication (<100KM)
Warm Backup	Manual Intervention	Manual	Asynchronous replication (>100KM)
Cold Backup	Strong Manual Intervention	Manual	Same as above

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

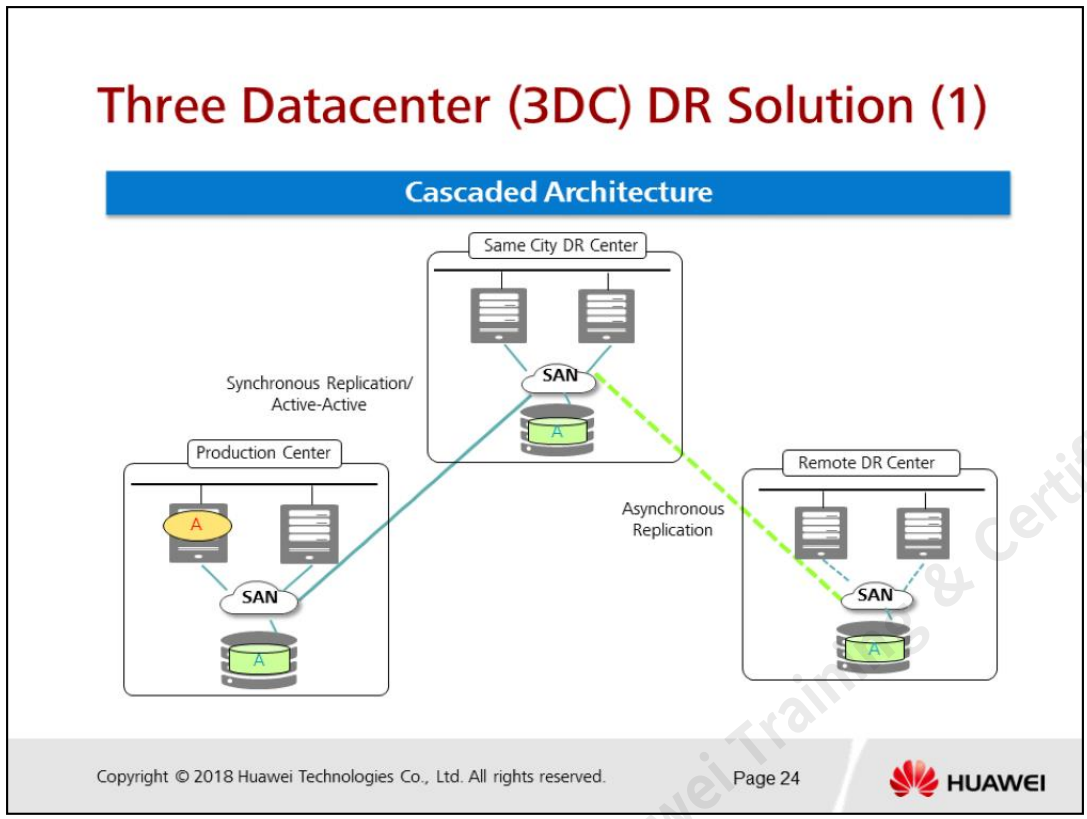
- The table above shows the different DR mode that can be implemented with comparisons on its reliability, recovery process and data backup requirements.

Huawei Training & Certification



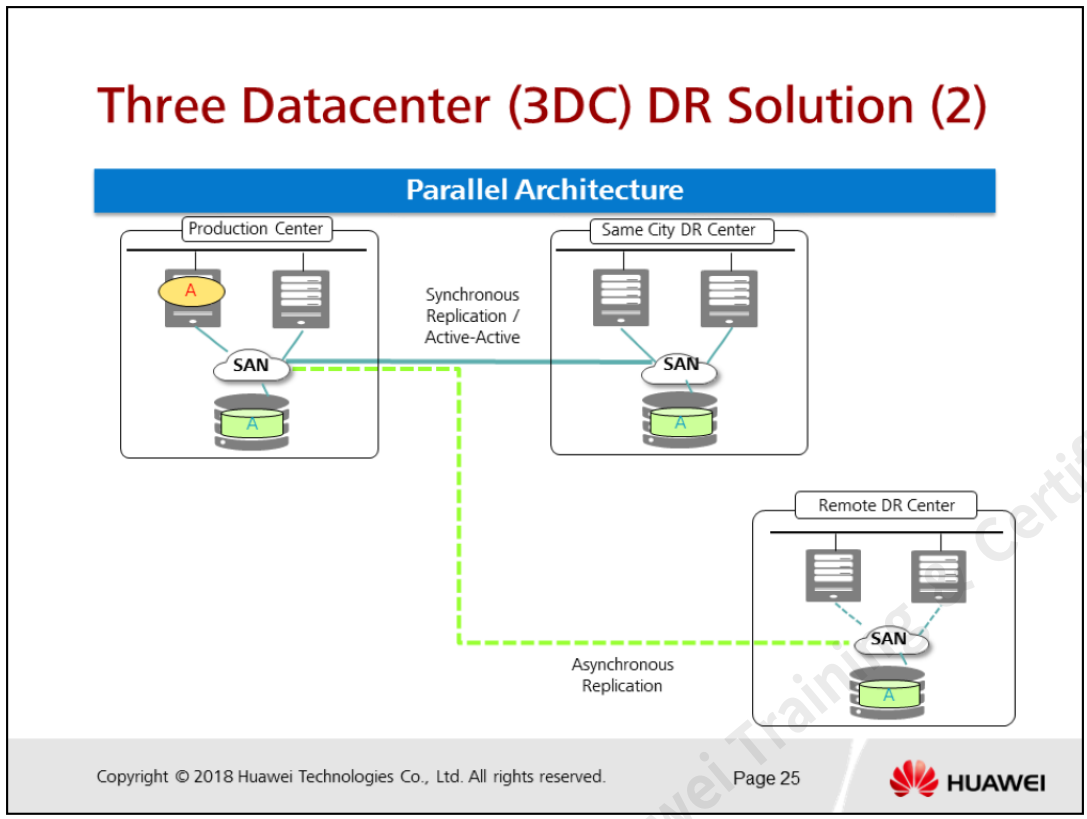
- Disaster Recovery System Management Visualization:
  - Support disaster recovery management software deployment and commissioning.
  - Supports disaster recovery drills and one-click switching, and has capabilities to assists in the completion of custom scripting tools, and assists you in recovering the standby business system in one click.
- Disaster recovery service is mature and efficient:
  - Disaster recovery system has one-stop analysis, design, delivery, and drill service.
  - Specialized disaster recovery drill process and technical plan, with 20+ disaster recovery drill project implementation experiences including finance, government, and medical.
  - Support the implementation of the disaster recovery plan for old non-Huawei equipment and reduce the cost of customers by 40%.

Huawei Training & Certification



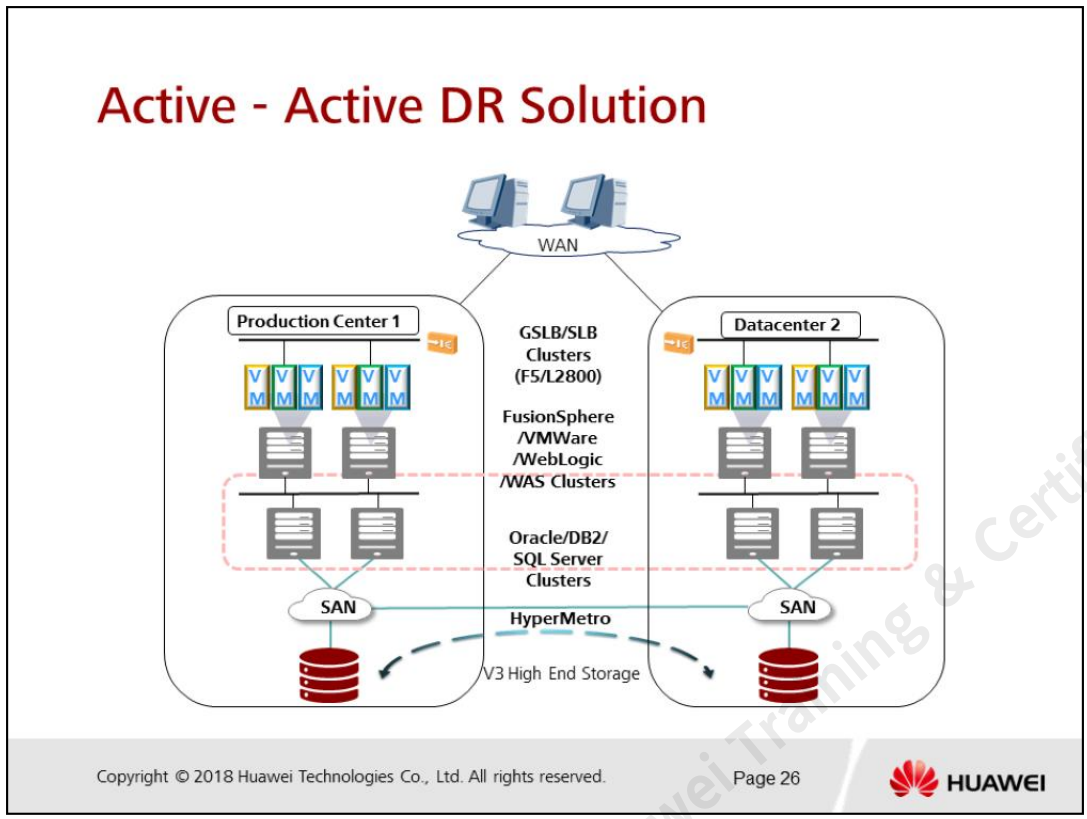
- Disaster recovery construction cycle is short and delivery risk is low:
  - Disaster recovery construction period shortened from 10 months to 7 months, which is shortened by 30%.
  - Multi-vendor cooperation with effective management, and thus shortens the project delivery cycle.
  - Effective assessment and analysis of multiple services and multiple applications to ensure rapid construction of disaster recovery systems.
  - Has Asia's largest integrated verification center to ensure effective validation of disaster recovery design and reduces the project implementation risk.

Huawei Training & Certification



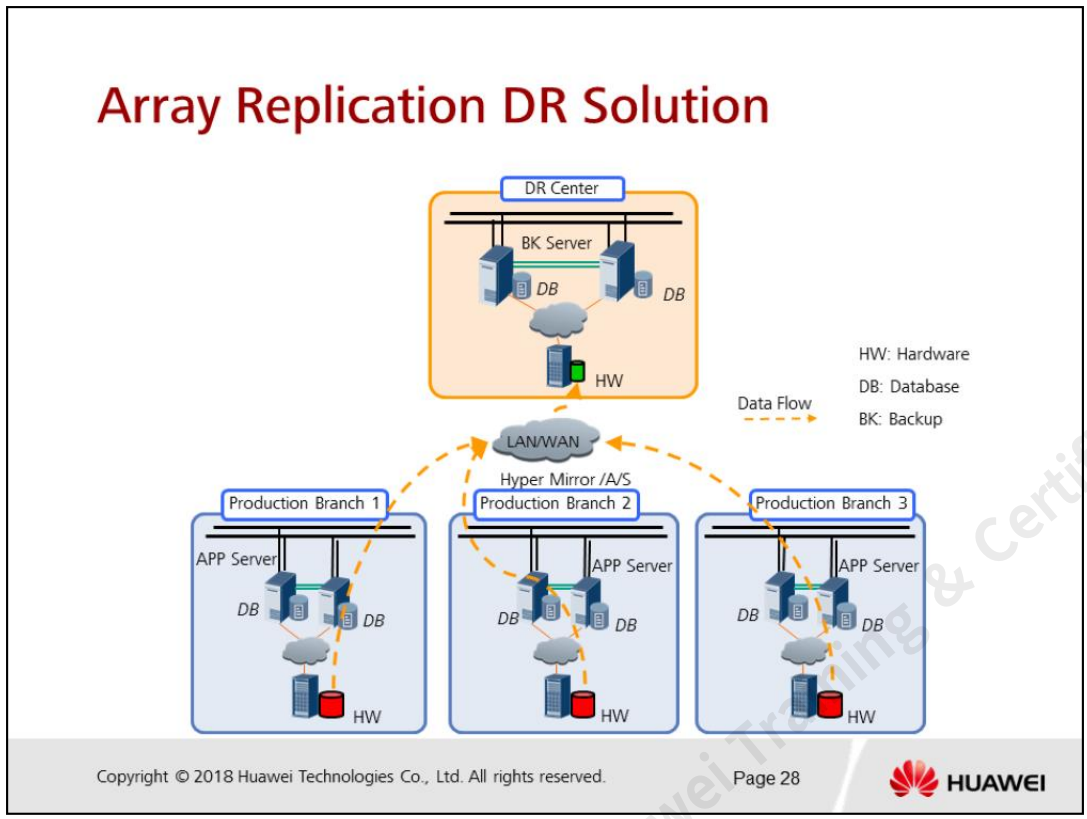
- Remote disaster recovery system management visualization:
  - Support one-click visual deployment and commissioning of disaster recovery management software.
  - Support unified management and monitoring of production centers, disaster recovery facilities in the same city, and remote disaster recovery center equipment, and thus simplifies the equipment maintenance process.
  - Support one-button disaster recovery drill and switching, and customer-customized script recovery of standby service system, and simplified management and maintenance of disaster recovery system.

Huawei Training & Certification



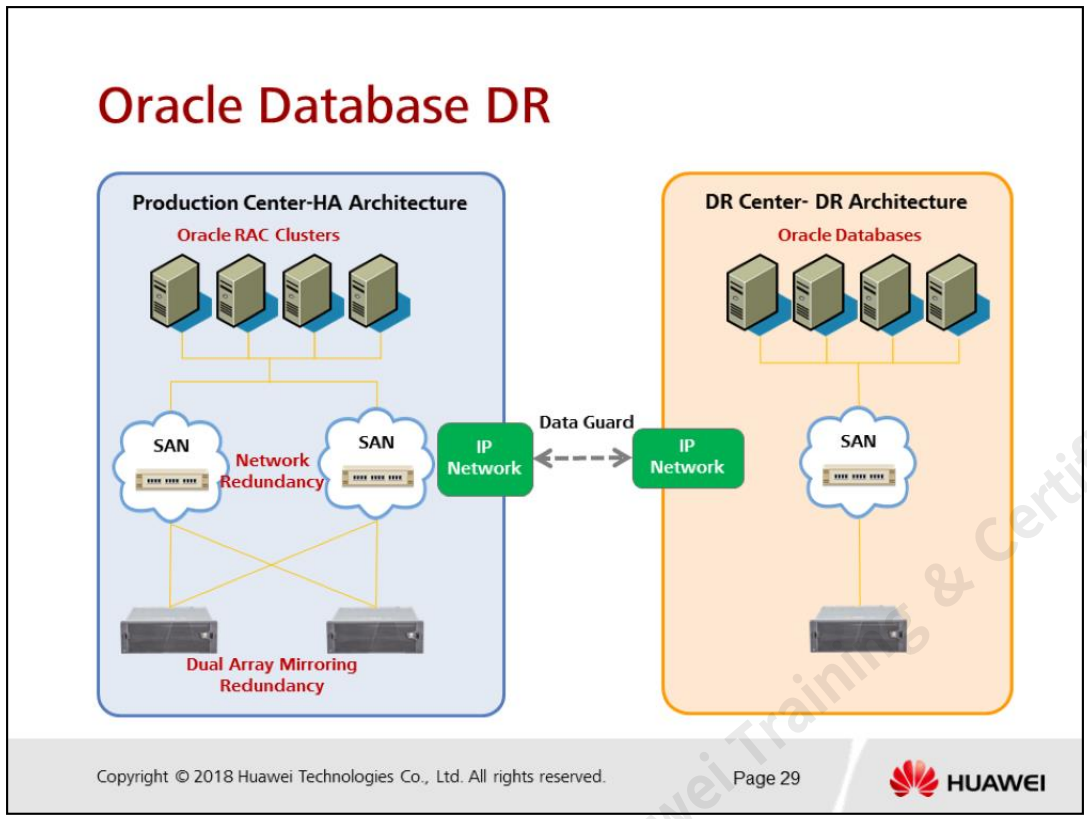
- Active-Active gateway free disaster recovery, has high efficiency in implementation:
  - Enterprise grade "Active-Active" high reliability, no service downtime even when there is data center level of fault occurs, 7*24hours online.
  - Active-Active storage layer with no virtualized gateways, reduces points of failures, and simplifies implementation and commissioning.
- Rich experiences in Active-Active DR services:
  - Rich experience in the implementation of active-active disaster recovery and implementation of more than 80+ active-active DR projects in the fields of finance, government, and medical care.
  - Completes the design and implementation of the 3DC DR plan from the existing active-active DR plan without stopping the running services.
  - With the largest integrated verification center in Asia, extensive range of IT equipment allows comprehensive test verification and remote demonstration of the disaster recovery plan to be conducted to ensure the perfect design of the plan.
  - Extensive disaster recovery implementation experience, strong professional service tools support, improve disaster recovery program implementation efficiency, and reduce TCO by up to 30%.

Huawei Training & Certification



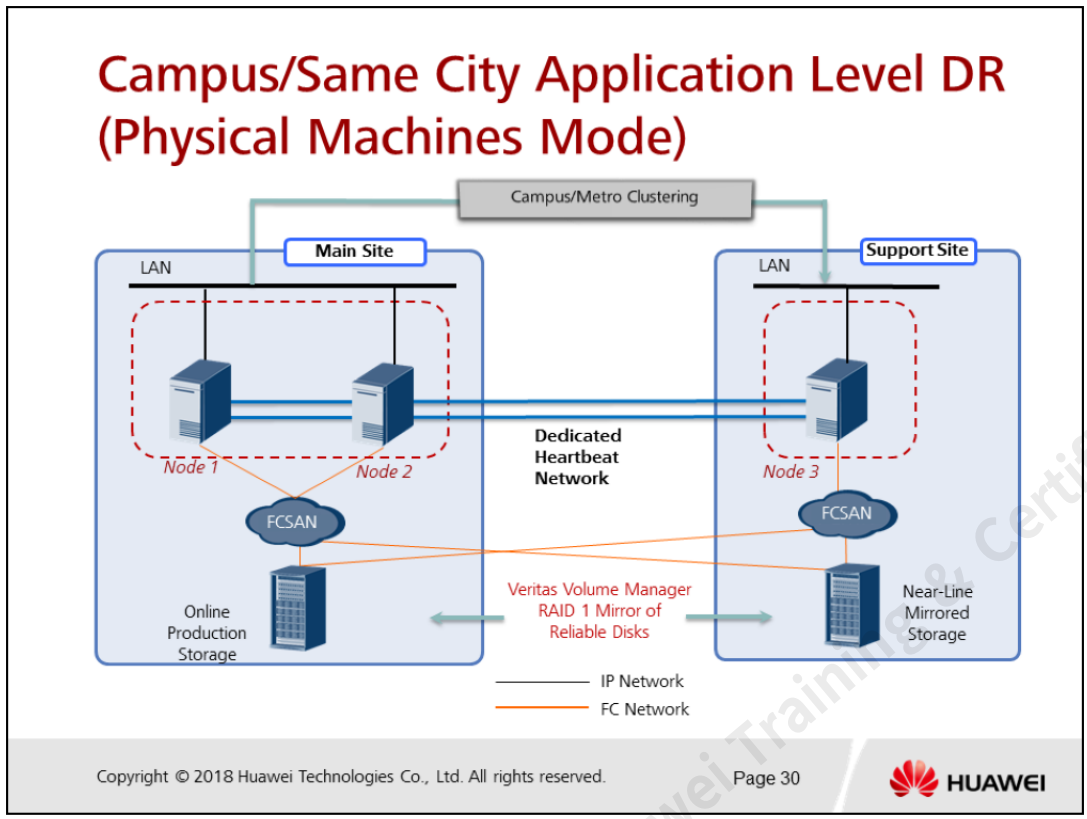
- Application scenario:
  - Support one-to-one disaster recovery or multi-to-one disaster recovery.
  - Multi-branch's data is stored independently. A unified disaster recovery system needs to be established to provide unified disaster recovery protection for each branch.
- DR Solution Features:
  - Supports up to 31 branches.
  - Support synchronous/asynchronous replication mode, flexible selection according to the characteristics of each branch.
  - Unified management of disaster recovery storage systems in each branch.
  - Different levels of arrays can replicate with each other, thus reducing TCO.
  - Supports flexible online addition/removal of branch sites.
  - With the snapshot function, continuous data protection can be achieved.
- Value to Customer:
  - Disaster recovery resource sharing, and saving investment.
  - Disaster Recovery Center maintains and manages resource sharing, and also saving maintenance and management costs.
- HyperMirror is a continuous data protection technology that creates two physical mirror copies of a LUN for redundant backup protection against host service interruption.

Huawei Training & Certification



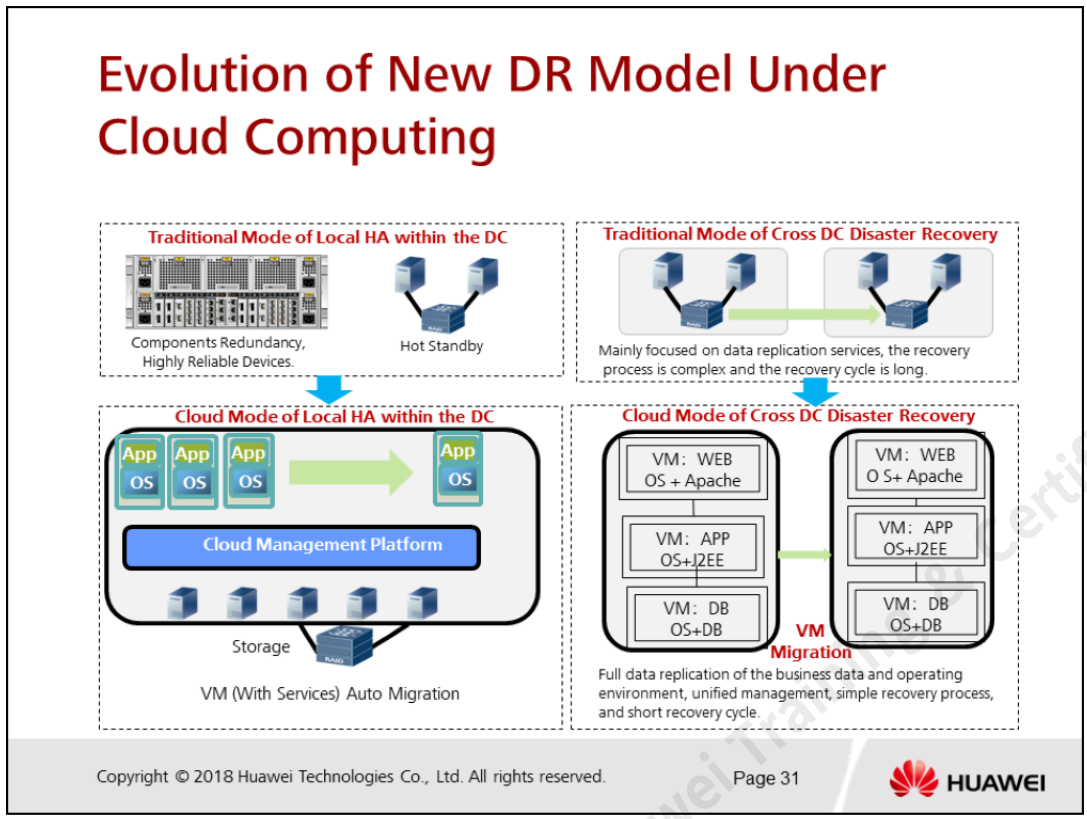
- Application scenario:
  - Adopts Oracle database.
  - Need to establish remote disaster recovery.
- Solution Features:
  - Local high availability, and remote disaster recovery.
  - Fully redundant architecture design.
  - Heterogeneous hardware compatibility.
- Value to Customer:
  - Little changes to the existing network environment.
  - Simple maintenance.
  - Redo-based replication with low bandwidth requirements.
- Oracle RAC(Real Application Cluster) is a cluster database with a shared cache architecture that overcomes the limitations of traditional shared-nothing and shared-disk approaches to provide highly scalable and available database solutions for all business applications.
- Oracle Data Guard ensures high availability, data protection, and disaster recovery for enterprise data. Data Guard provides a comprehensive set of services that create, maintain, manage, and monitor one or more standby databases to enable production Oracle databases to survive disasters and data corruptions.

Huawei Training & Certification



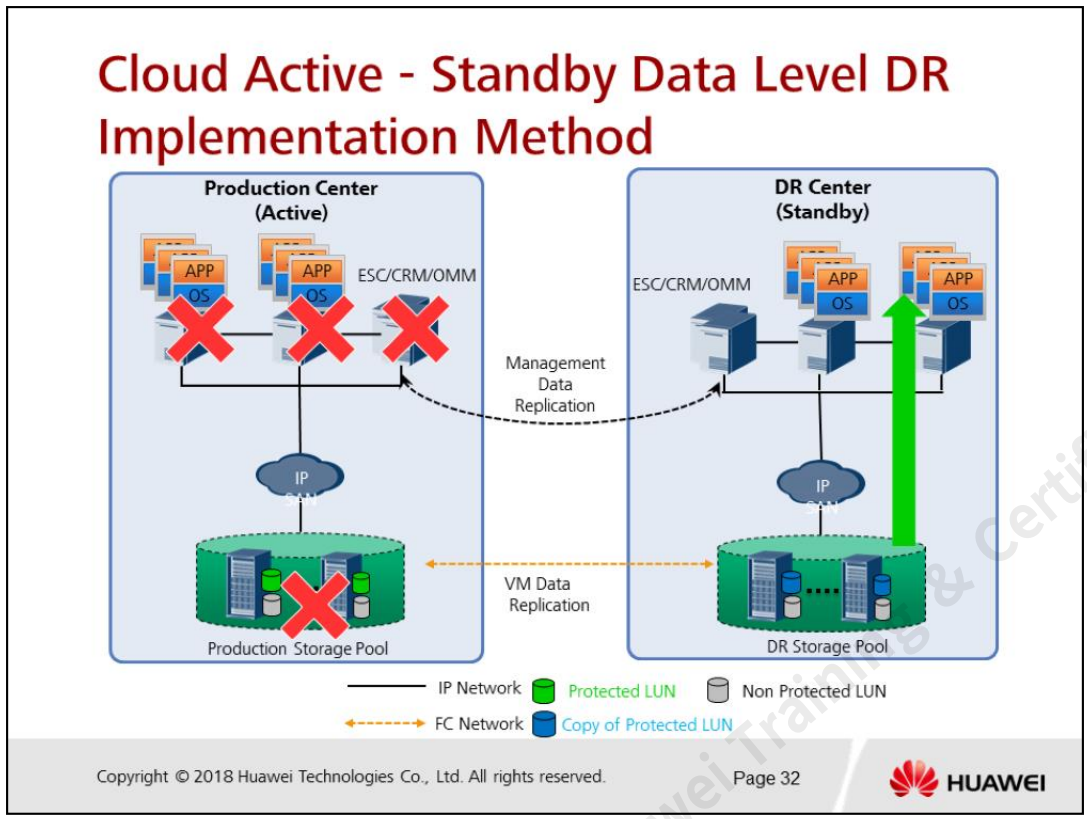
- Application scenario:
  - Physical machine mode (minicomputer, X86 server).
  - Two data center's distances must be <80KM.
- Value to Customer
  - Application automatic switching to quickly recover services.
  - Supports a wide range of application types and can be tailored to customer needs.
  - Develop a customizable switching script for each application.
- Program Features
  - Achieve local high availability and remote application level disaster recovery.
  - Equipped with openness for storages, and supports heterogeneous storage.
  - Cluster nodes can be arbitrarily combined (the total number of hosts in the cluster must be <= 64).
- Veritas Volume Manager: A Veritas software product installed on storage clients that enables management of physical disks as logical devices. Volume Manager enhances data storage management by controlling space allocation, performance, data availability, device installation, and system monitoring of private and shared systems.

Huawei Training & Certification



- After IT is centralized and cloudified, higher requirements are placed on business continuity, including on network requirements, data security requirements, and business reliability requirements.

Huawei Training & Certification



- Production Center and Disaster Recovery Center independently deploy cloud management platform (two cloud platforms deployed).
- Set up a synchronization strategy and regularly perform remote copying of the cloud management data and business data (VM) from the production center to the disaster recovery center
- In the service planning, users can divide two types of LUNs according to actual needs: protection LUNs and unprotected LUNs, create VMs that need disaster recovery on the protection LUNs, and configure array replication only for protection LUNs to save on the requirements of storage capacity within the DR center.
- When a fault occurs in the production center, the disaster recovery center uses the disaster recovery management software to restore the virtual machines in one click.

Huawei Training & Certification

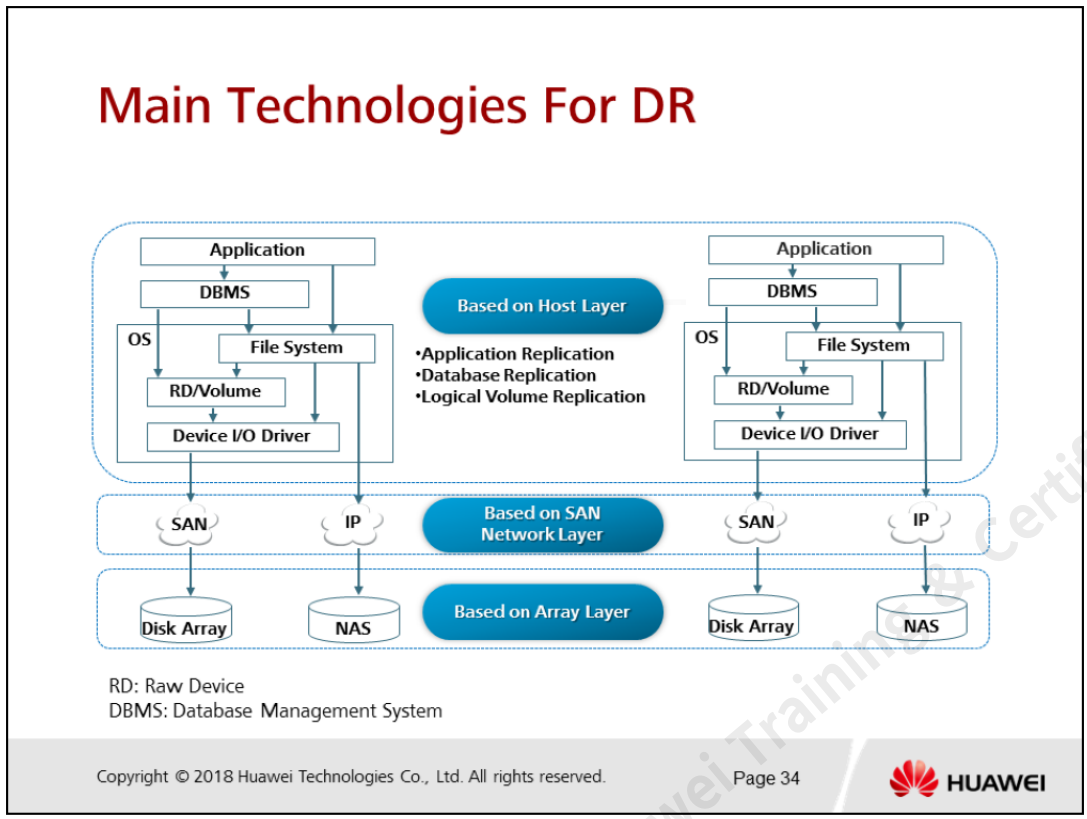
Contents

1. Overview of Disaster Recovery Solution.
2. Architecture of Disaster Recovery Solution.
3. **Common Disaster Recovery Technologies.**
4. Case Study of Disaster Recovery Application.

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

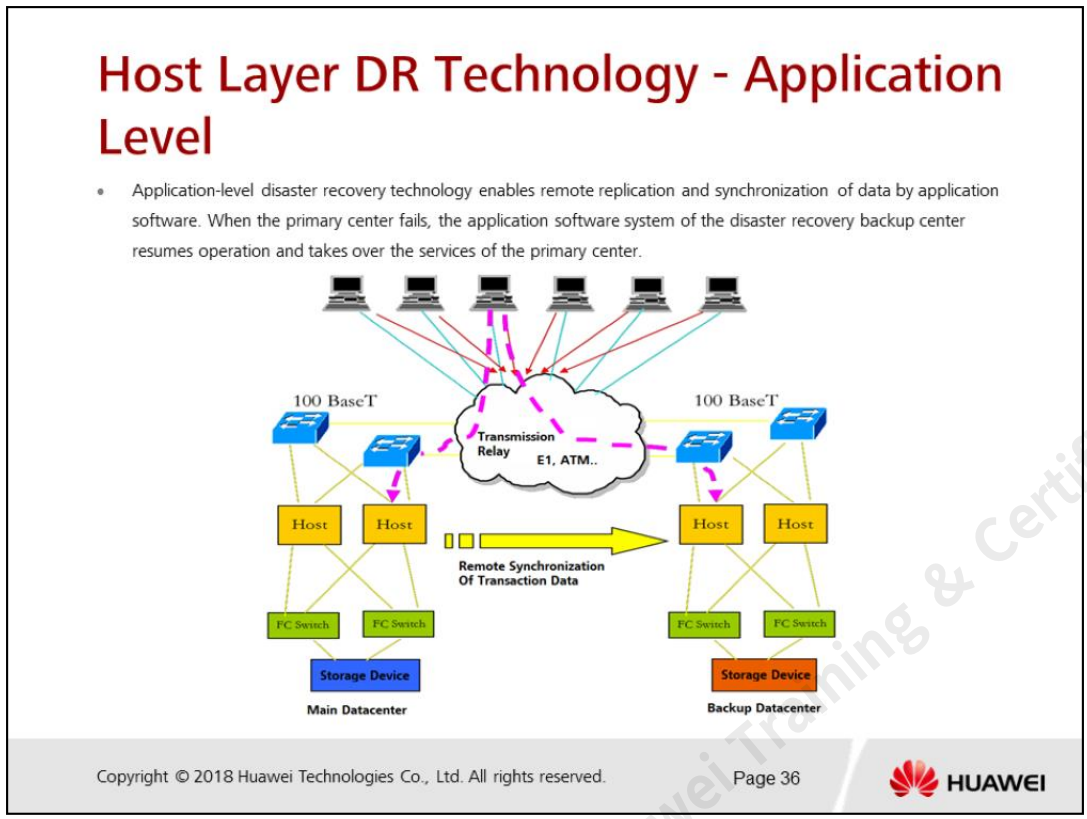
HUAWEI

Huawei Training & Certification
Huawei Training & Certification



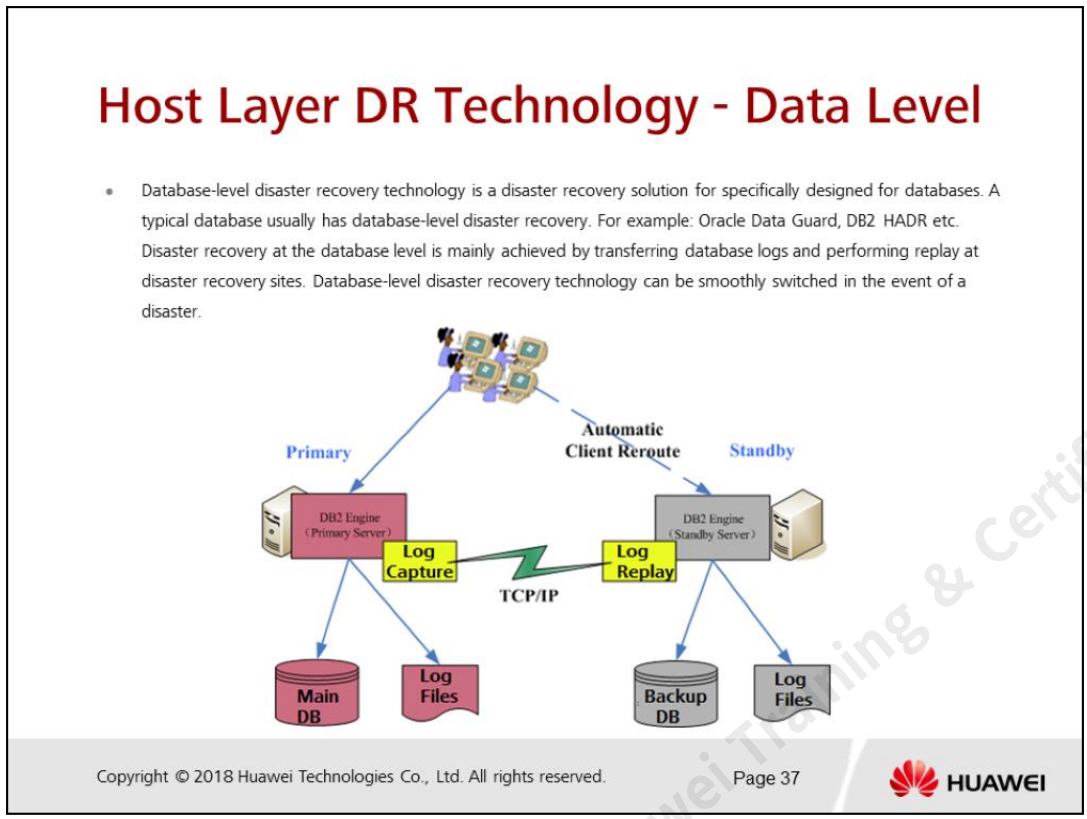
- Disaster recovery technology based on host layer :
  - Install dedicated data replication software, such as volume replication software, on servers in production centers and disaster recovery centers to enable remote replication. There must be a network connection between the two centers as data channels. The remote service switching function/service failover software can be added at the server layer to form a complete application-level disaster recovery solution.
  - This method of data replication has relatively less investment, mainly in the cost of software procurement. Its compatibility is better, and can be compatible with different brands of servers and storage devices, which is more suitable for users with complex hardware composition. However, this method requires synchronization of software operations on the server, and occupies a large amount of host resources and network resources.
- Disaster recovery technology based on network layer :
  - The SAN-based data replication technology consists of a storage area network (SAN) between the front-end application server and the back-end storage system. It joins/connected to the storage gateway, and connects to the front-end server host, and also connects to back-end storage devices.

Huawei Training & Certification



- Working principle: Two remote databases are connected through the application software. Each time the business services process data and separately stores them in the database of the main center and the backup center.
- Advantages and disadvantages:
  - Supports wide area networks, and no separate hardware or software support is required. Data logic replication prevents proliferation of human errors and these processes are transparent to disk subsystems.
  - Consistency checks need to be performed regularly. Backup data from the backup center cannot be quickly restored back to the main center and major modifications to the application are required.

Huawei Training & Certification

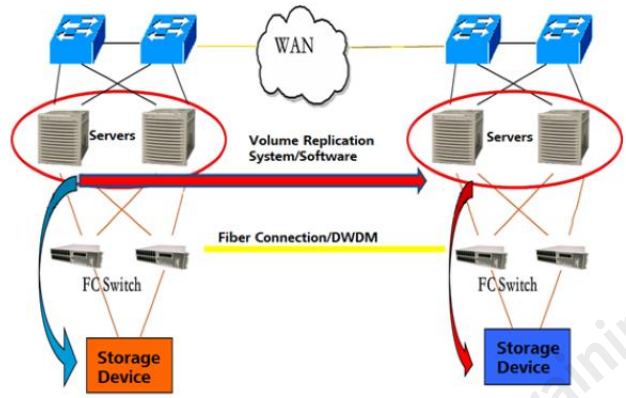


- Working principle:
  - Configure primary and standby database servers.
  - Once the primary database has transaction operations, the log files are simultaneously transferred to the standby database, and then the standby database replays the received log files to maintain consistency with the primary database.
  - When the primary database fails, the standby database server can take over the transactions of the primary database server.
- Advantages and disadvantages:
  - Support for wide area networks and no separate hardware support required. It is transparent to disk subsystems. Implementation of logical replication to reduce the risk of proliferation of human error. There is no need to modify the application. The data in the main center/disaster recovery center can be accessed at the same time.
  - The backup center's backup data cannot be quickly restored back to the main center. Non-database data can't be copied remotely. The synchronization system has a great impact on the production system. In the asynchronous mode, more data will be lost and at least one log file will be lost. Complex service switching process and production transformation is complicated.

Huawei Training & Certification

Host Layer DR Technology - Logical Volume Level

- Logical disk volume-based remote data replication refers to the remote synchronous (or asynchronous) replication of one or more volumes based on demand. The implementation of this solution is usually done through software.



Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

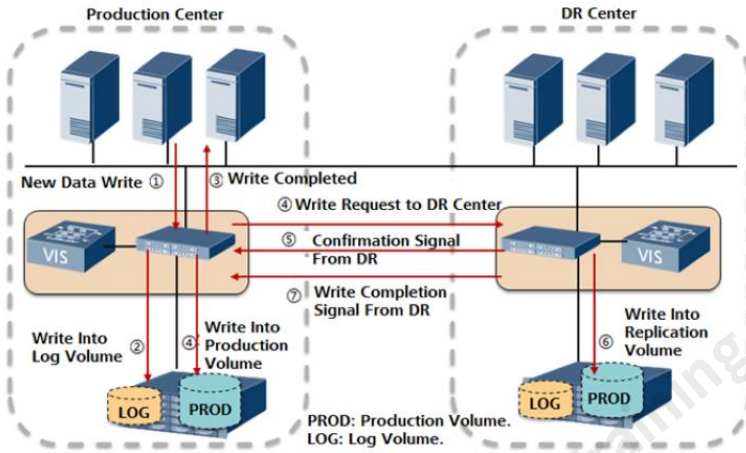
HUAWEI

- Working principle: The remote replication control management software copies the operation data of each I/O on the volume of the active node system to the corresponding volume of the remote node, in real time (or with delay) so as to achieve the data synchronization between the two remote volumes whether if it is synchronous or asynchronous.
- Advantages and disadvantages:
  - Ensures data integrity and consistency. The structure is relatively simple and transparent to the disk subsystem.
  - The write performance of the host is affected by the distance. If there is no host at the disaster recovery center, data-level disaster recovery cannot be performed and thus logical level disaster cannot be prevented.

Huawei Training & Certification

Network Layer DR Technology

- The SAN-based data replication technology is a storage area network (SAN) between the front-end application server and the back-end storage system. A layer of intelligent switches is added, where the front-end is connected to server and hosts, and the back-end is connected to the storage devices.

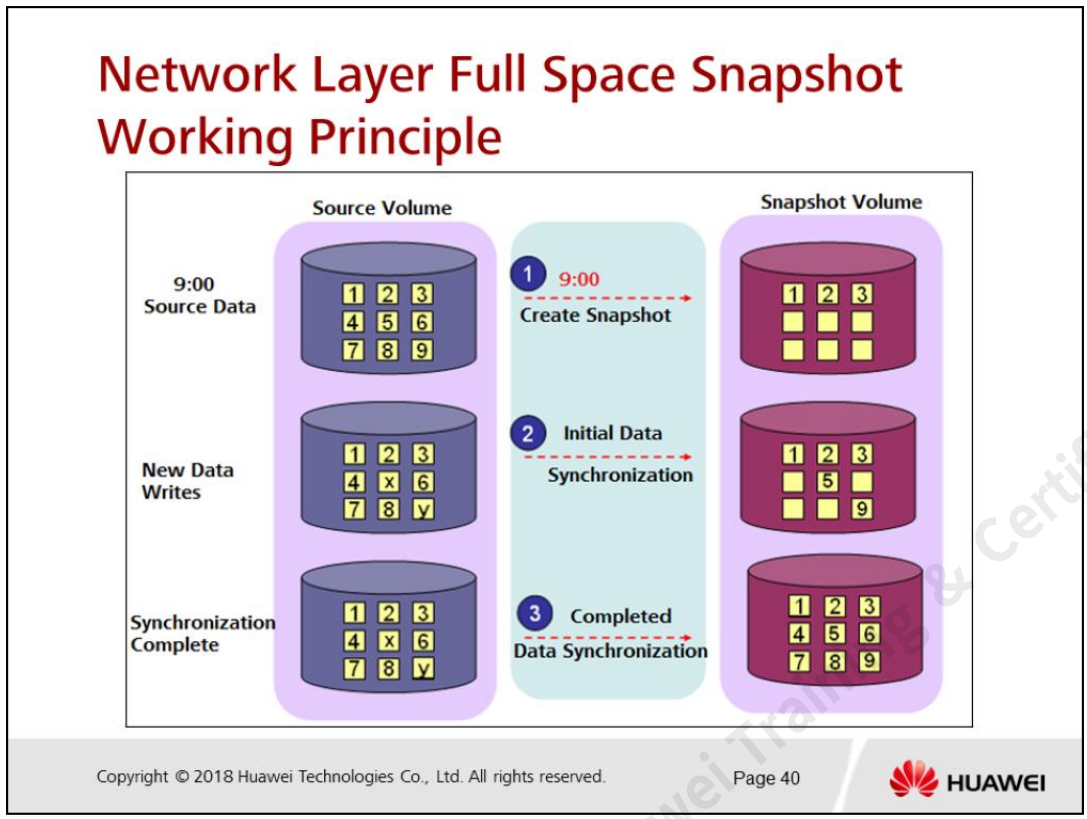


Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

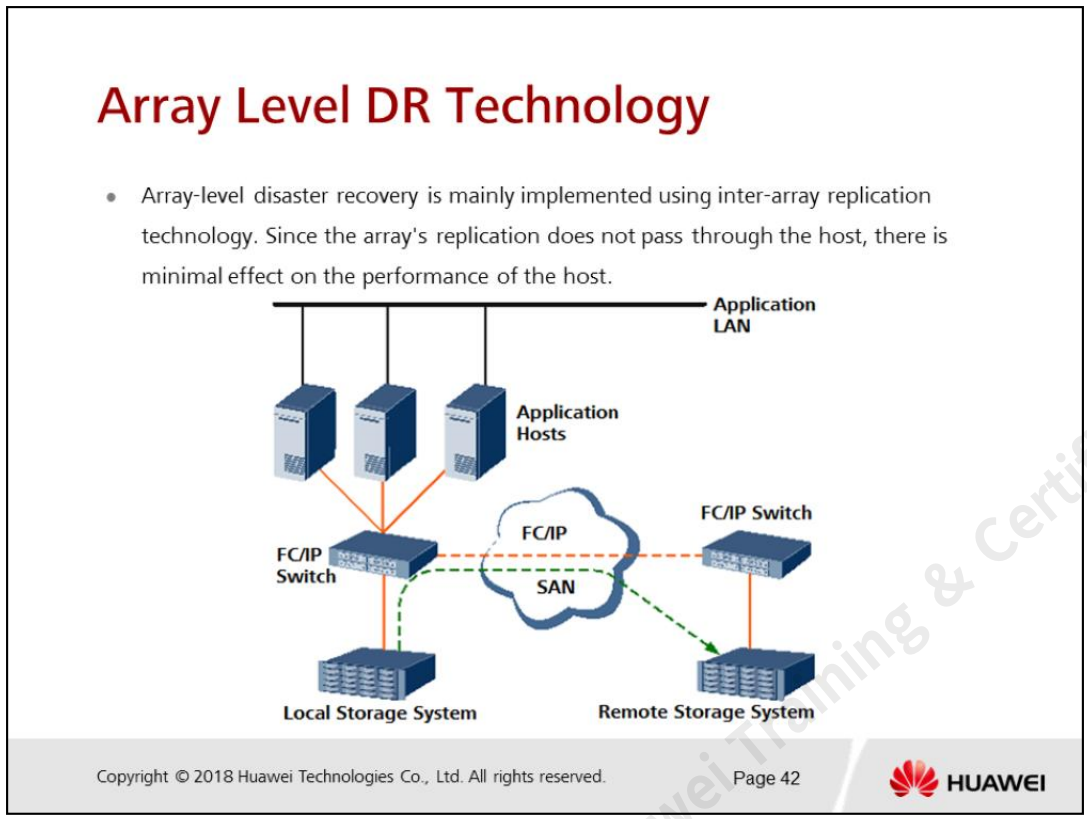
- Working principle:
  - The production center host writes data to the local virtual gateway.
  - The production-side virtual gateway writes data to the local log volume.
  - After the data is successfully written to the log volume, the virtual gateway of the production center returns the "confirmation" to the local host.
  - When the production-side virtual gateway writes data to the local production volume, it sends a data write request to the disaster recovery-side virtual gateway.
  - After the disaster recovery virtual gateway receives the write request, it returns a "confirmation" to the production virtual gateway.
  - The disaster recovery-side virtual gateway then writes data to the disaster recovery-side replication volume.
  - After the data is successfully written to the replication volume in the disaster recovery center, the virtual gateway in the disaster recovery center returns a “completed” signal to the virtual gateway in the production center.
- Advantages and disadvantages:
  - Support heterogeneous storage devices, allows virtualization integration, achieves unified management, and improve storage utilization.
  - Need to transform the SAN network.

Huawei Training & Certification



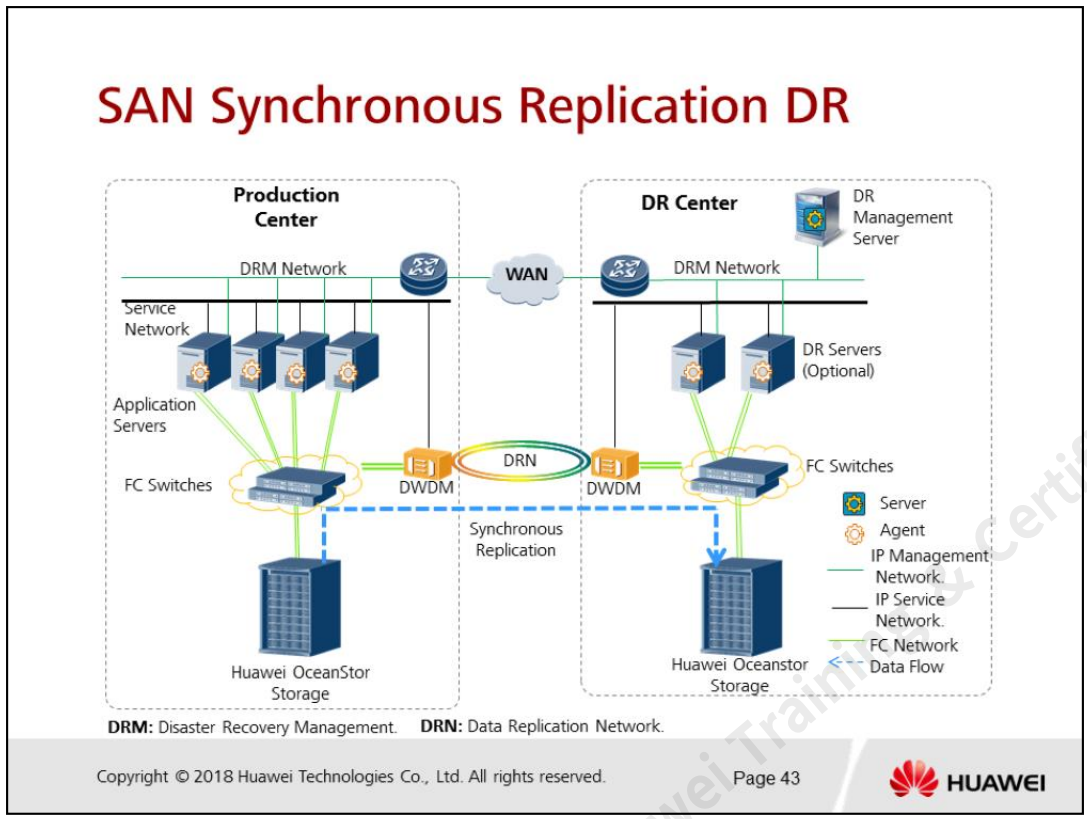
- Technical principle:
  - The working principle of the full-space snapshot technology: When the snapshot time is reached, the system allocates a physical space with the same size as the snapshot volume for the source data volume, and starts background data synchronization. After the synchronization data is completed, the snapshot is created at this point in time is considered as a success.
  - A full-sized snapshot is a point in time physical copy of the data at the source volume's snapshot.
- Steps:
  - Create a volume with the same size as the source volume as the snapshot volume and start background data synchronization.
  - If new data is written to the source volume in the data synchronization process and the data is written to the content that has not been copied to snapshot volume yet, the original data is written to the snapshot volume. Subsequently, the new data is written to the source volume to retain its most up to date status. The data is up-to-date if the location of the newly written data is part of the already synchronized copy of contents in the snapshot volume, thus only the new data is written to the source volume and the contents of the snapshot volume data are unchanged.

Huawei Training & Certification



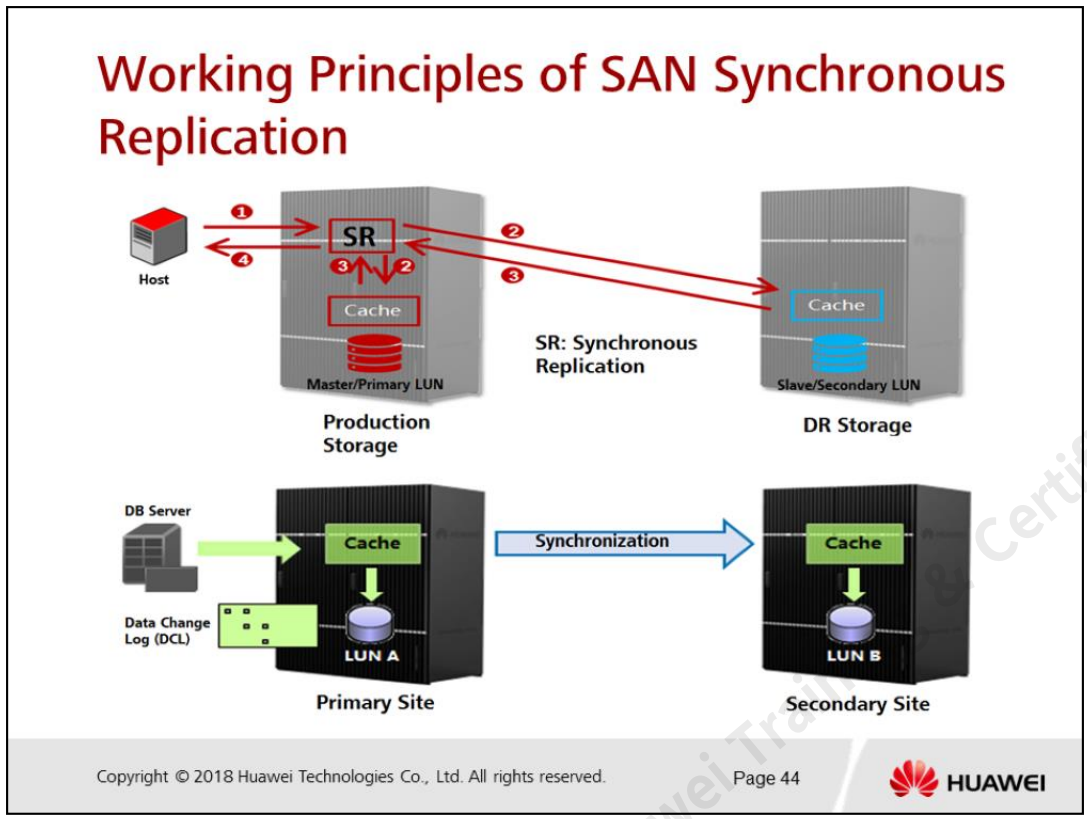
- Array Level DR technology are one of the simple DR methods used to protect data by utilizing 2 similar storage arrays to replicate data. Although it has the benefits of low impact to host performance and simple implementation, there are a few downsides.
- The Downsides to Array Based Replication
  - The first problem with array based replication, in almost all cases, is that it requires a near identical unit from the same manufacturer at the DR site. The DR site though may not need the same level of performance or redundancy as the primary, and a less expensive system may be a more appropriate fit.
  - This inability to mix storage hardware also means that adding new hardware from a new vendor must be considered carefully, since the introduction of a new vendor will mean learning another replication module and buying a second storage system for the DR site. It may also mean that a newer more innovative storage vendor may not be considered because they lack a replication capability.
  - The third problem is that array based replication software is not application aware. These modules just replicate blocks of data as it changes on the array with no understanding of the application that is changing that data or how other servers may be related to that application. Software based replication software that is application aware can ensure that clean copies of data are captured. They can also make sure that groups of servers that are dependent on each other can be kept in sync

Huawei Training & Certification



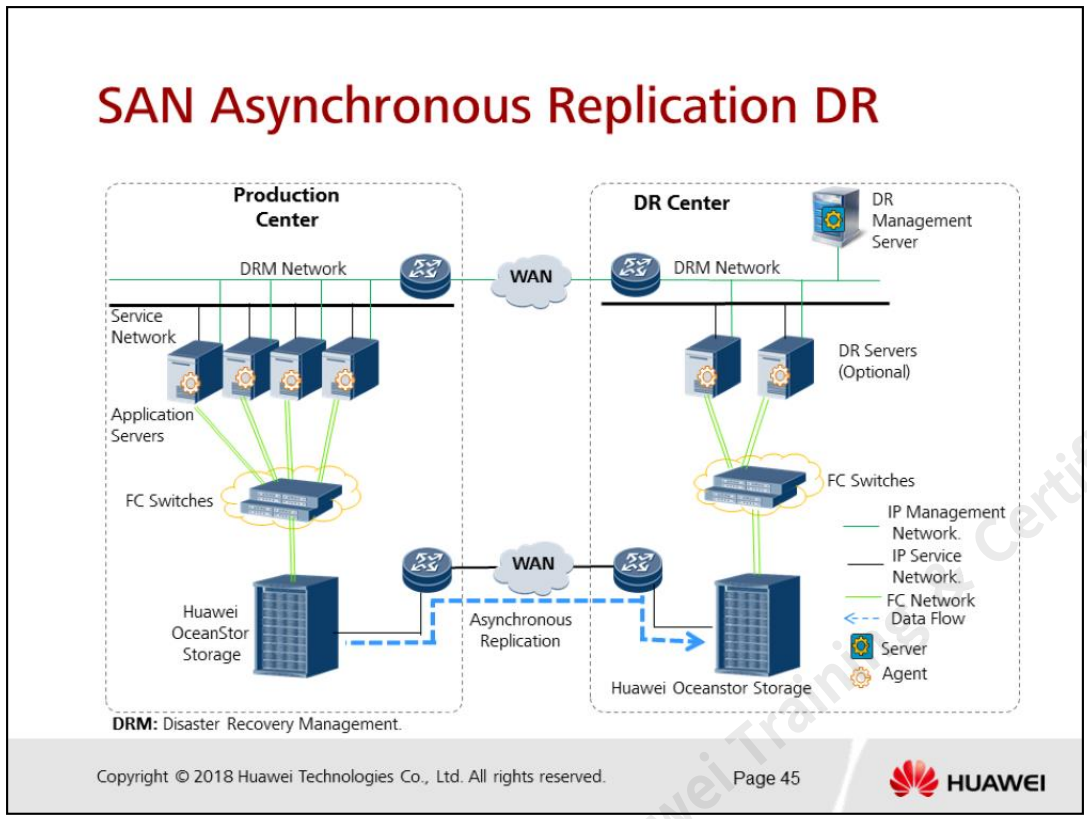
- The deployment method is shown in the figure above has the target RPO=0 and level of RTO within minutes.
- SAN-based disaster recovery replication supports synchronous replication, which is recommended to be within the distance of 100km.
- Remote Datacenter (RD) mainly provides disaster recovery management functions, including topology, disaster recovery testing, drills and disaster recovery.
- When performing application management and disaster recovery application recovery, the Agent needs to be installed on the server
- The RD management network needs to communicate with the host and storage.
- It supports FC/iSCSI links, and it is recommended to use FC links for synchronous replication.
- Dense wavelength division multiplexing (DWDM) is a technology that puts data from different sources together on an optical fiber, with each signal carried at the same time on its own separate light wavelength. Using DWDM, up to 80 (and theoretically more) separate wavelengths or channels of data can be multiplexed into a light stream transmitted on a single optical fiber.

Huawei Trainings & Certification



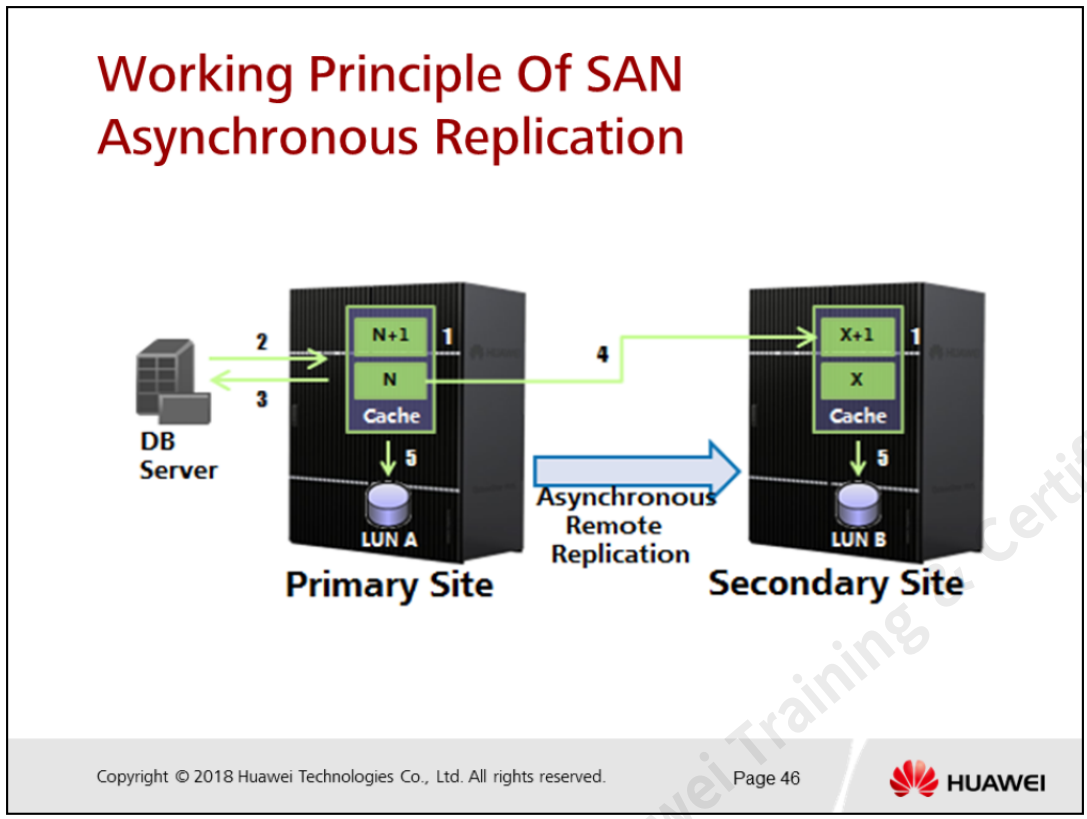
- Synchronization steps:
  - The production storage system receives a write request from host. HyperReplication logs the request. Only the address information is recorded in the log, and the data content is not recorded.
  - It write this request to the master and slave LUNs. Normally, LUNs are in write-back mode and the data is written to Cache.
  - HyperReplication returns both write processing results for the master and slave LUNs. If all the writes are successful, the log is cleared. Otherwise, the log is retained and it enters an abnormal disconnection state. When the synchronization starts back, the data block corresponding to the log address is copied again.
  - It returns the processing result of the host write request, which is based on the processing result of the data write to the master LUN.
- Split:
  - In the split state, the write request of the production host is only written to the master LUN, and the difference log records the difference between the master and slave LUN data. When you want to re-maintain data consistency between the primary and secondary LUNs, you can manually start the synchronization operation. The synchronization process is the process of incrementally copying the data block marked as "differenced" in the differential log from the primary LUN to the secondary LUN. The principle of I/O processing here is similar to the principle of initial synchronization mentioned above.

Huawei Training & Certification



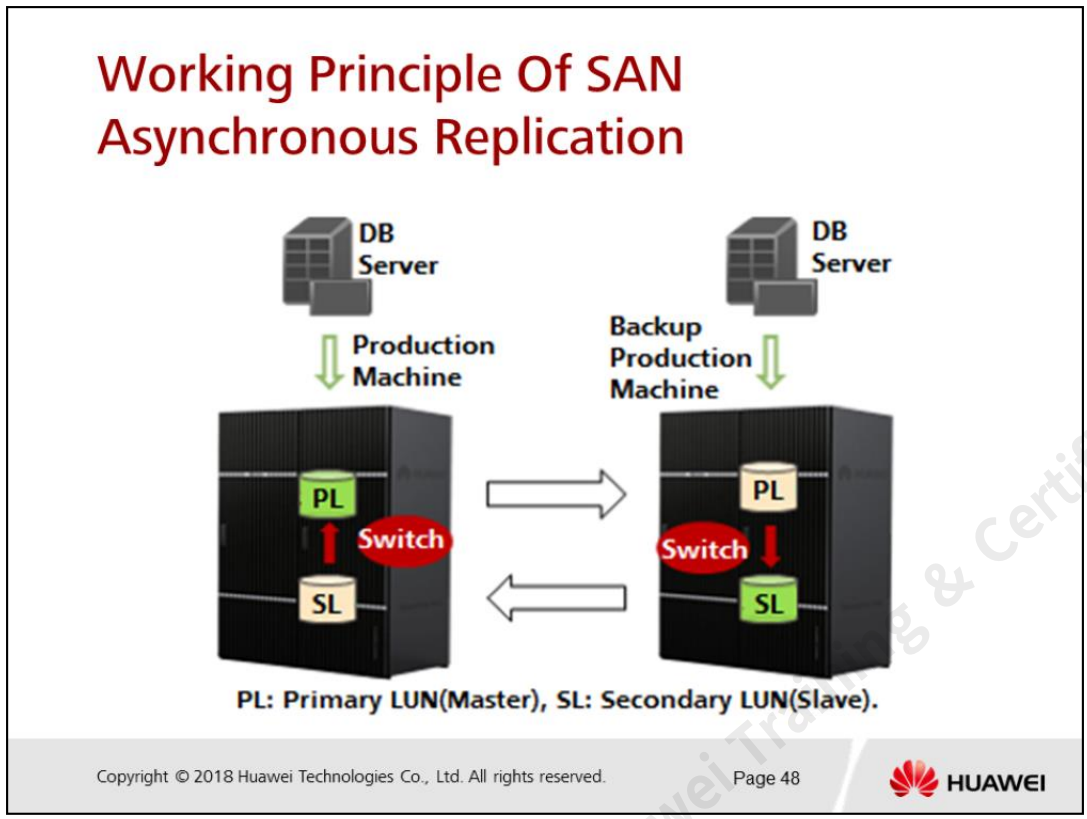
- The deployment method is shown in the figure above. The target RPO is >3s and RTO level is in minutes.
- Different from synchronous replication, asynchronous replication has time-interval replication strategies, and theoretically no limitations in distance.
- RD mainly provides disaster recovery management functions, including replication strategy, topology, disaster recovery testing, drills, and disaster recovery.
- When performing application management and disaster recovery application recovery, the Agent needs to be installed on the server
- The RD management network needs to communicate with the host and storage.
- Seconds-level replication is triggered on production storage, meanwhile replications that can tolerate for more than 15 minutes of delay can be triggered on RD.

Huawei Trainings & Certification



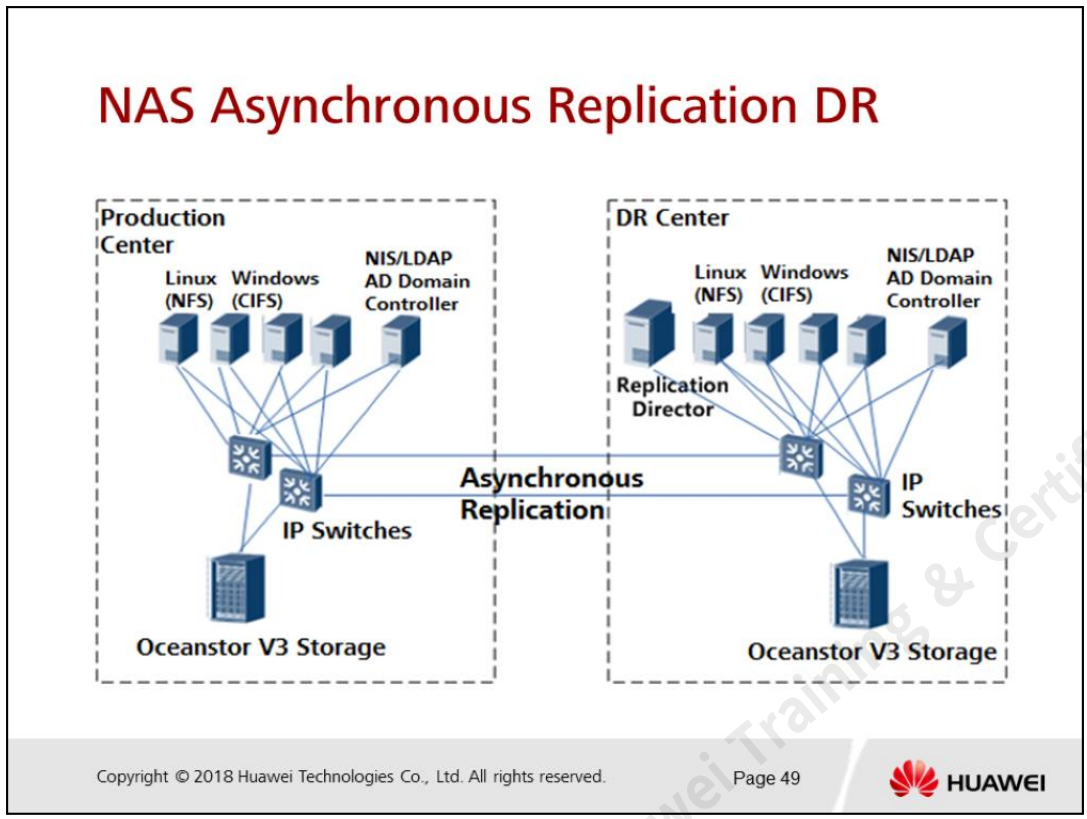
- Time slice: Logical space for writing data in the cache for a period of time (data size is not limited).
- In a low-RPO scenario, the asynchronous remote replication cycle is short. OceanStor storage system caches all data in multiple time slices. If the host service bandwidth or disaster recovery bandwidth is abnormal or fails, the replication cycle becomes longer or interrupted, at this time the data in the Cache will be automatically flushed to provide consistent protection according to the disk flushing policies. It will then read back the flushed data from the disk when the replication resumes.
- Whenever a synchronization period is set (Its set by the user, and the range is from 3s to 1440min), the system will automatically initiate a synchronization process to synchronize the master site data incrementally to the secondary site (if the synchronization type is manual, the user is required to trigger synchronization.). New time slices (TPN+1 and TPX+1) are generated in the cache of the primary LUN (LUN A) and the secondary LUN (LUN B) at the start of each replication cycle.
- The main site receives the production host's write request.

Huawei Training & Certification



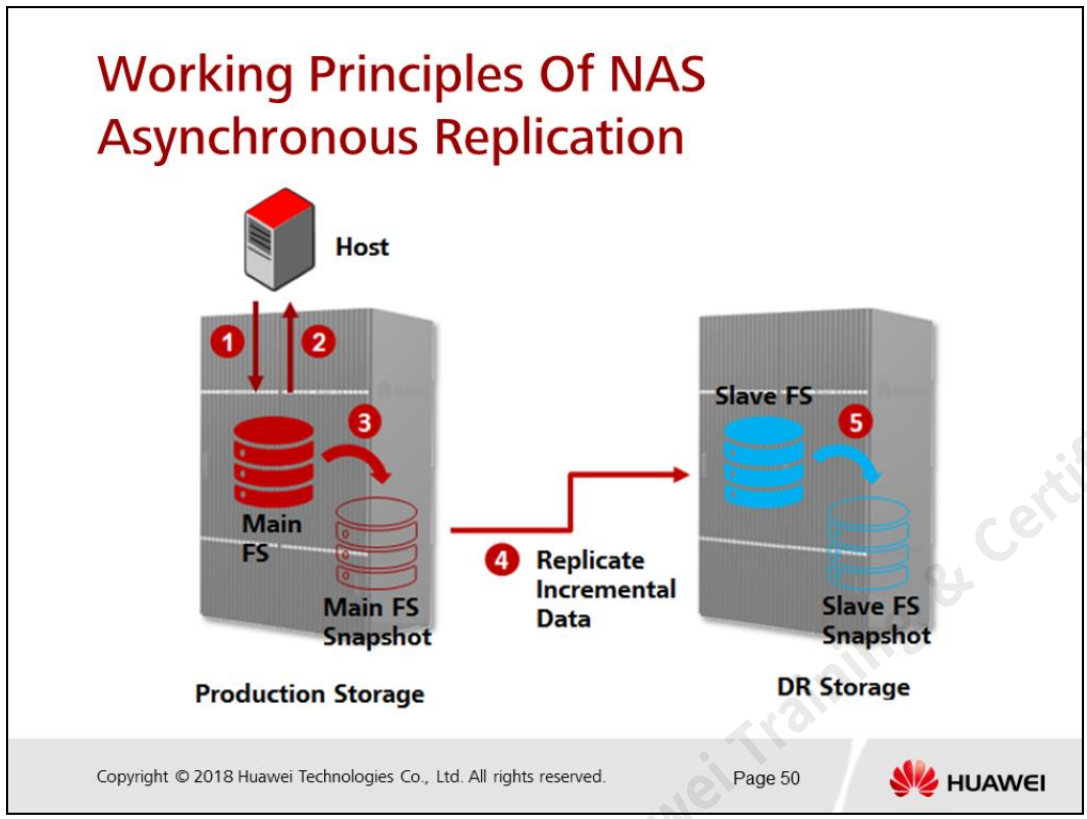
- Switching/Service Failover:
  - Synchronous remote replication can perform master-slave switching under normal conditions.
  - In the split state, you need to set the slave LUN for master-slave switchover.
  - Asynchronous remote replication is in a split state.
  - In the split state, you need to set the slave LUN to be writable.

Huawei Training & Certification



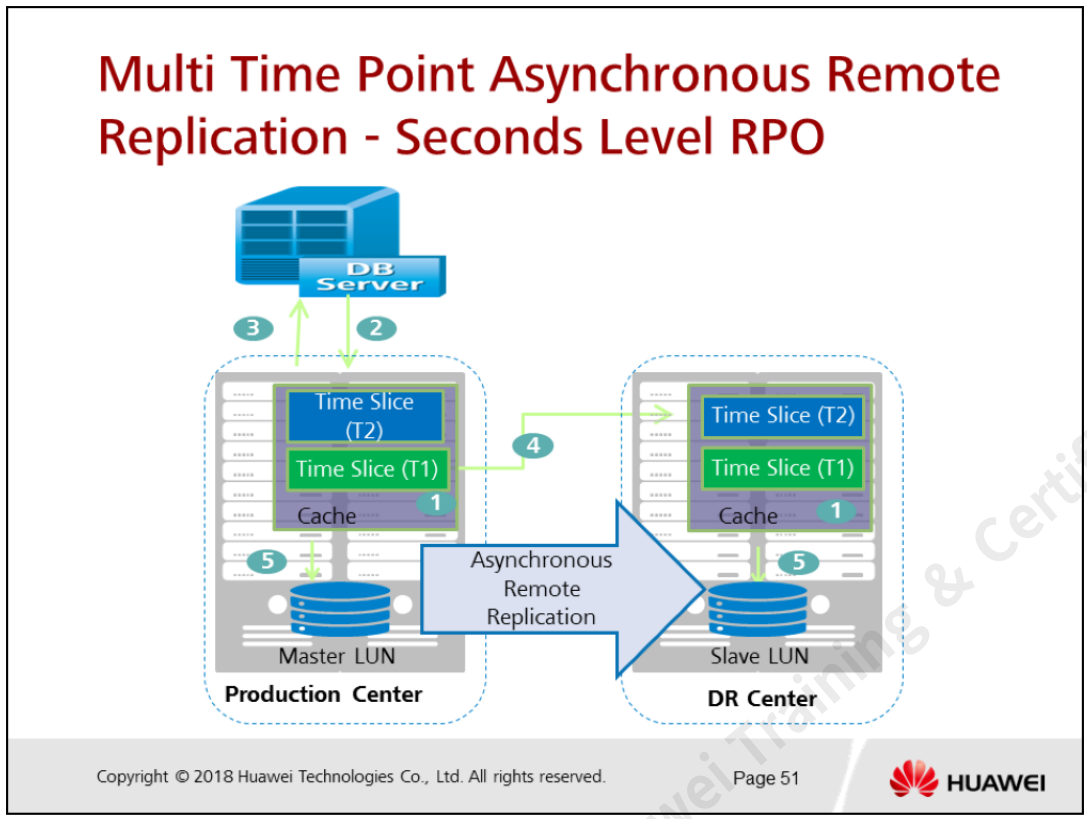
- NAS file system replication is currently available only in the V3R2C10 version and is implemented using ROW.
- The NAS's NAS disaster recovery management does not deploy the Agent on Linux or Windows. It only manages the replication policies and disaster recovery of V3 storage.
- The file system currently mainly supports NFS/CIFS. Disaster recovery management currently manages only the FS replication section, file system, and permission control section. When the system is created, it needs to be configured.
- File system replication is similar to SAN and supports FC/iSCSI links.

Huawei Training & Certification



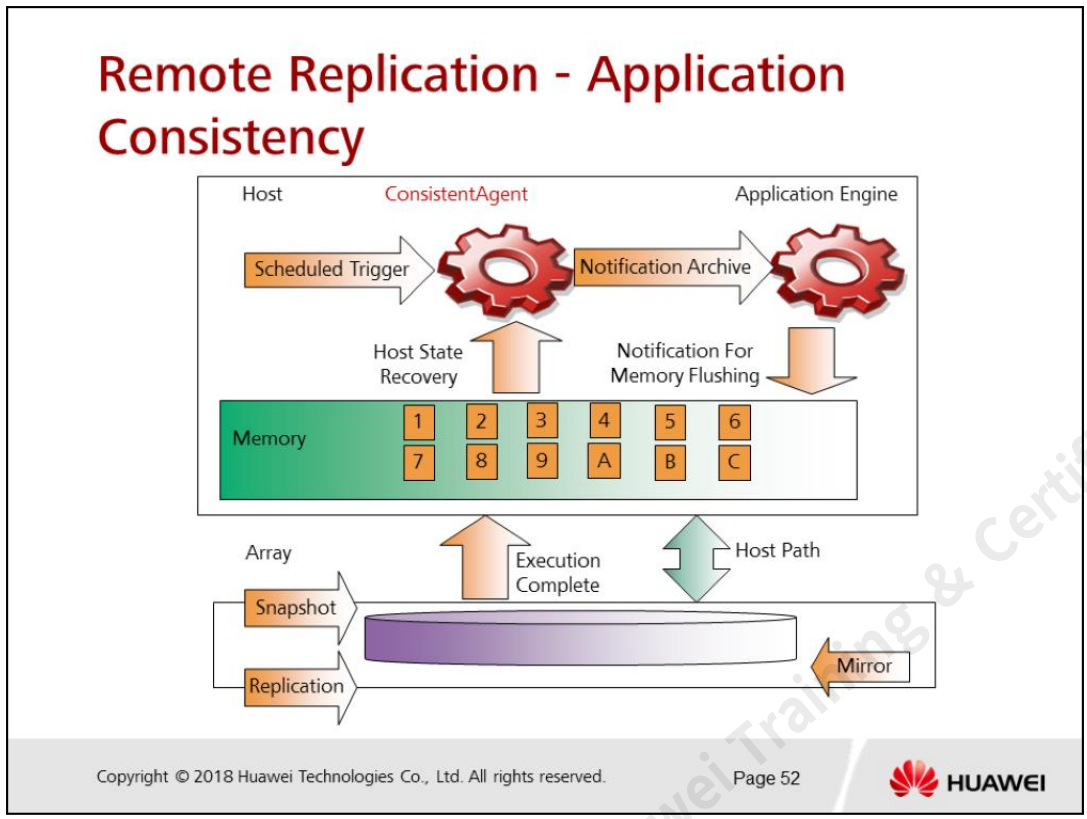
- At the beginning of each cycle, the file system asynchronous remote replication creates a snapshot of the primary FS (primary file system). Based on the delta information during the period from the completion of the previous cycle to the start of the current cycle, the snapshot data is copied to the secondary FS. After the incremental copy is completed, the content of the FS is the same as the content of the snapshot of the primary FS, and the data consistency point is formed on the FS.
- It supports long distance remote replication from one file system to another, but it does not support directory-to-directory or file-to-file remote replication.
- A single file system can only be included in one replication task, but a replication task can include many different file systems.
- File system only supports one-to-one replication, and the same file system cannot be set as the source and destination of the remote replication process. It does not support cascaded replication and also doesn't support 3DC.
- The smallest unit of incremental replication is the file system block size of (4K-64K). The shortest asynchronous replication synchronization cycle period is 5 minutes.
- Supports replication resuming if the connection was interrupted or broken during the replication process.

Huawei Training & Certification



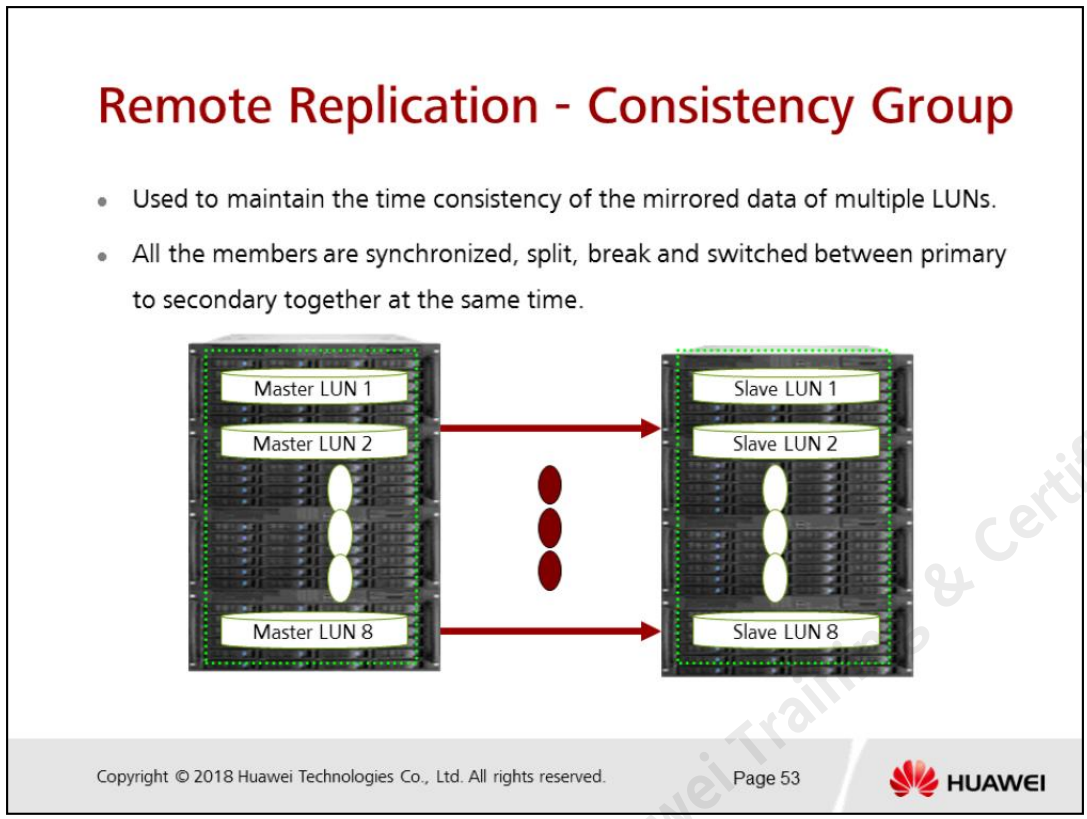
- A minimum of 3 seconds for a consistency point:
  - A new time slot*(T2, P2) is generated in the cache of the master LUN and the slave LUN at the start of each replication cycle.
  - The newly written data of the host is cached in the time slot T2 of the main LUN cache.
  - Response is sent to host on write completion status.
  - The data of the time slice T1 is directly copied to the slave LUN and written to the time slot P2 of the slave LUN.
  - The data received by the primary and secondary LUNs will be written to disks.
    - The copy reads data directly from the cache with a small delay.
    - Snapshots do not require real-time COW(Copy On Write) updates. Synchronization has a small impact on performance and the replication cycles can be reduced to a 3-second replication cycle. This means that the smallest amount of time to wait before the next cycle starts can be set to a minimum 3 seconds.

Huawei Training & Certification



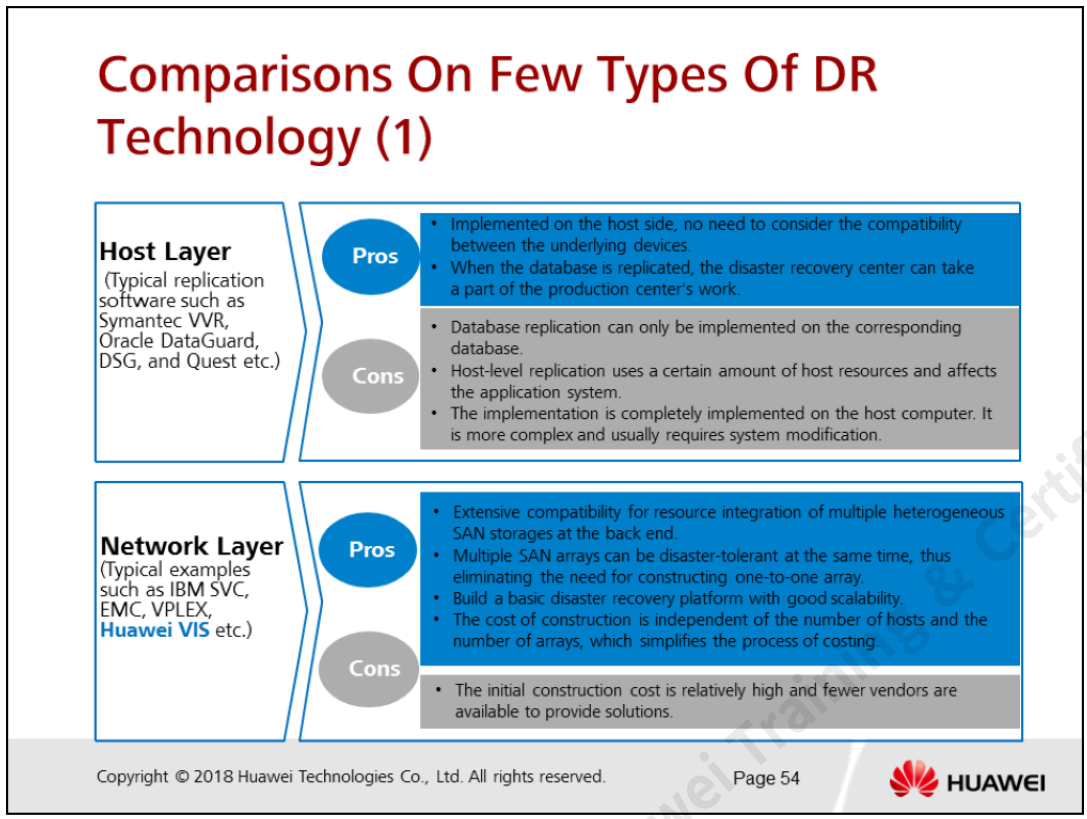
- Application consistency:
  - The consistency agent is installed on the host to implement the linking between the snapshot of the array and the database.
- When the snapshot task executes:
  - First, it puts the database in backup mode, perform checkpoints, and write all the dirty data in memory to the storage system.
  - Then, it informs the array to take a snapshot.
  - Finally, the database is taken out of backup mode.
- Advantages:
  - The disaster recovery end can pull up the data and uses it directly without the need of Roll Forward and Rollback process.

Huawei Training & Certification



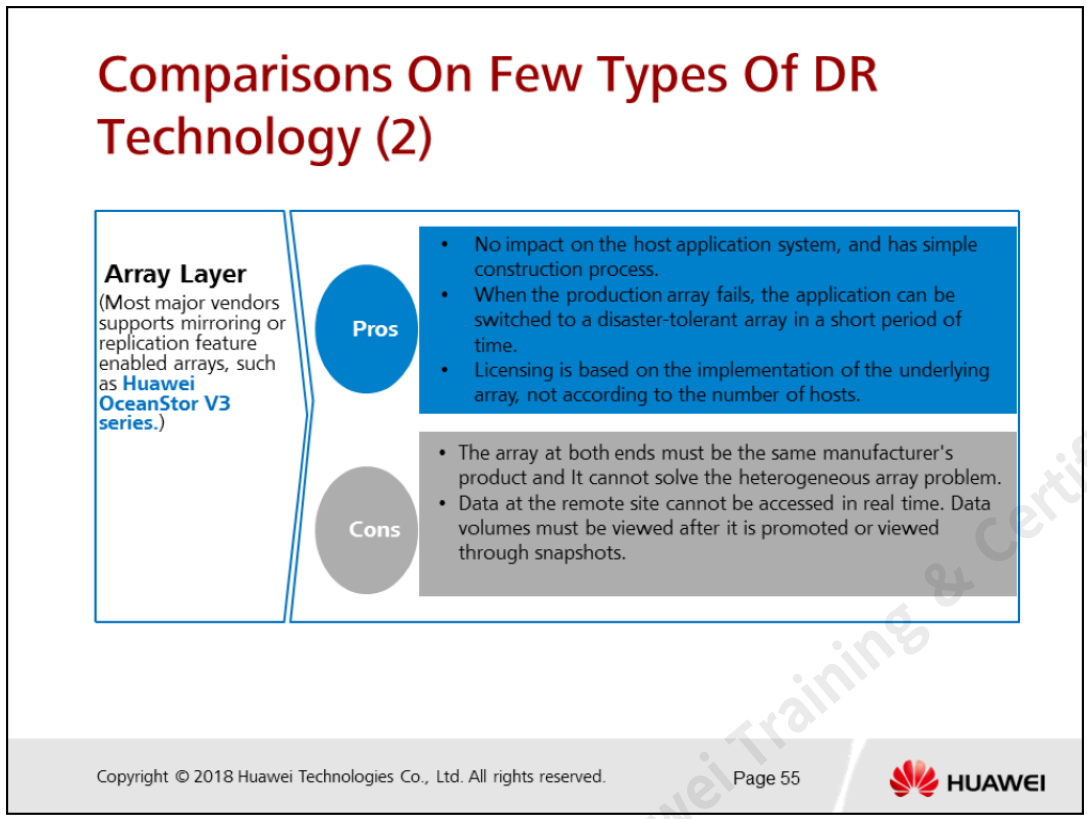
- In large and medium-sized database applications, data, logs, and modification information are stored in different LUNs of the array. Generally, the associated LUN is a non-stand-alone LUN. If the data of one LUN is missing, the data in other LUNs will be invalid. Thus these LUNs must remain consistent in time for the data to be valid and useful in the event of a disaster.
- We hope to simultaneously perform data synchronization or split operations on these LUNs to ensure that the data association between multiple secondary LUNs remains unchanged, thus ensuring the integrity and availability of disaster recovery backup data. The technology introduced to solve this issue is remote replication consistency group technology.
- The maximum number of remote replication pairs in the remote replication consistency group of the Huawei array is 8, and cross-array consistency groups are not supported.
- Note: Remote replication of associated LUNs should be placed in a consistency group. LUNs without an association should not be placed in a consistency group. In addition, synchronous remote replication and asynchronous remote replication cannot be placed in the same consistency group. All slave LUNs for remote replication must be on the same remote storage system.

Huawei Training & Certification



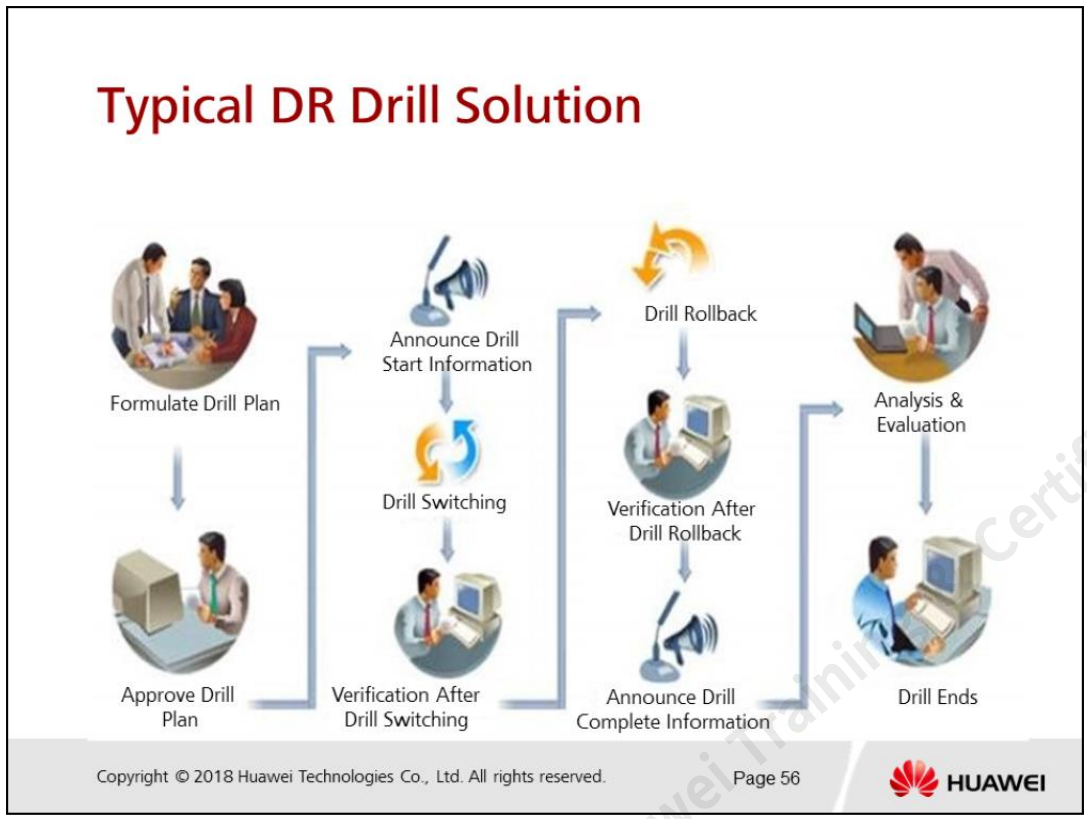
- The table above shows the comparison of DR technologies based on the host and network layer implementation.

Huawei Training & Certification



- The diagram above shows the pros and cons of array layer DR technology compared to the earlier host layer and network layer DR technologies.

Huawei Training & Certification



- Features and Benefits of Disaster Recovery Drill Exercise Process:
  - Real Disaster Recovery steps drill exercise.
  - Through Disaster Recovery Drill Exercise, familiarize with the Disaster Recovery Process and improve the Service Recovery Ability in the event of disaster.
  - Through disaster recovery drill exercises, check the business service and data integrity to properly prepare for handling the actual event of disasters.

Huawei Training & Certification

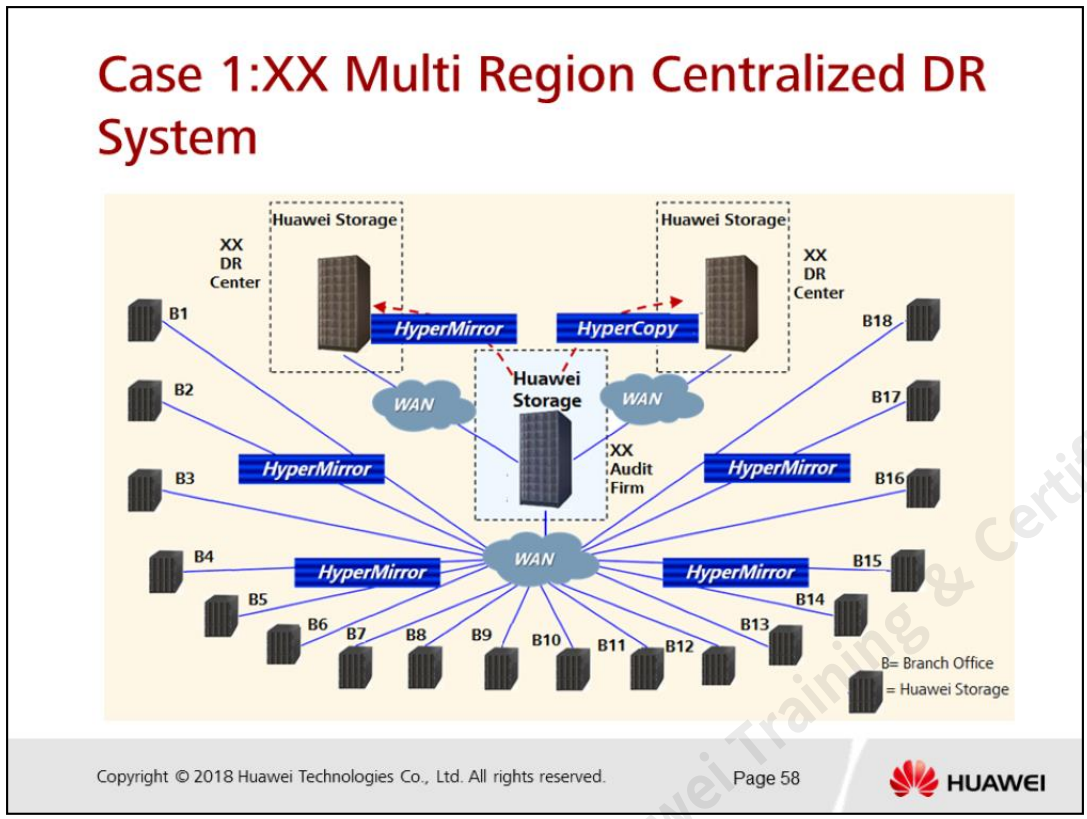
Contents

1. Overview of Disaster Recovery Solution.
2. Architecture of Disaster Recovery Solution.
3. Common Disaster Recovery Technologies.
4. **Case Study of Disaster Recovery Application.**

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

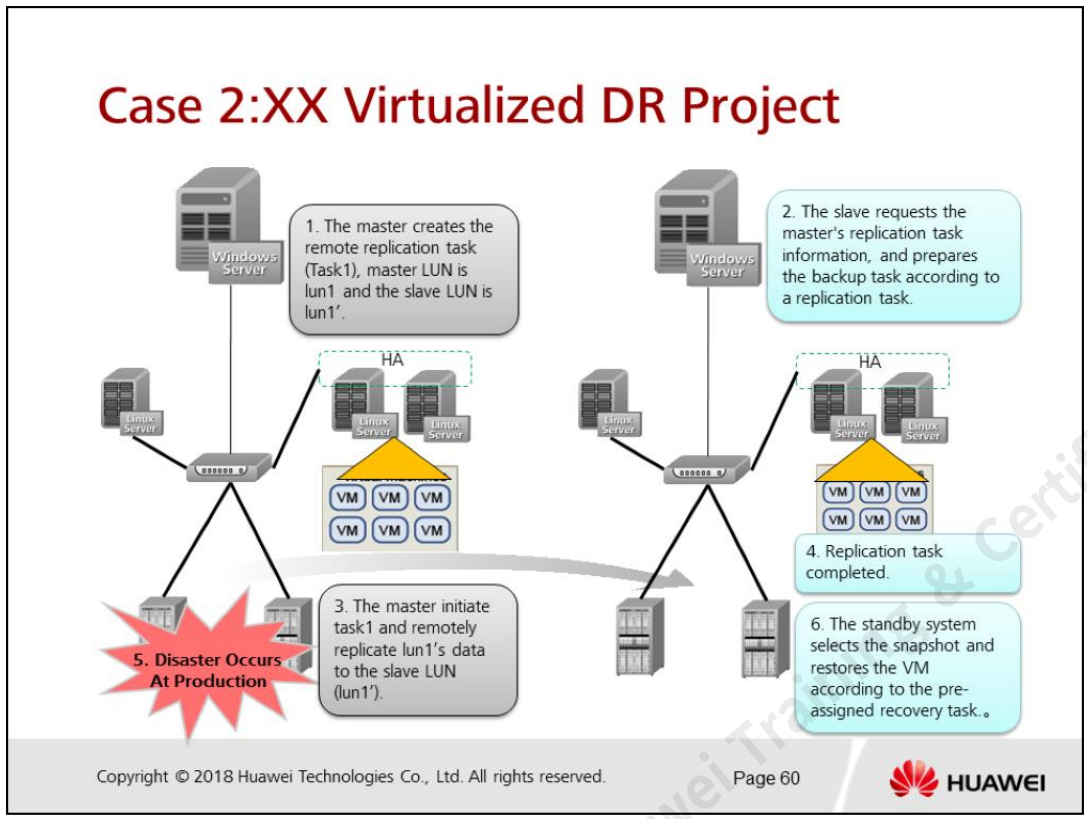
HUAWEI

Huawei Training & Certification
Huawei Training & Certification



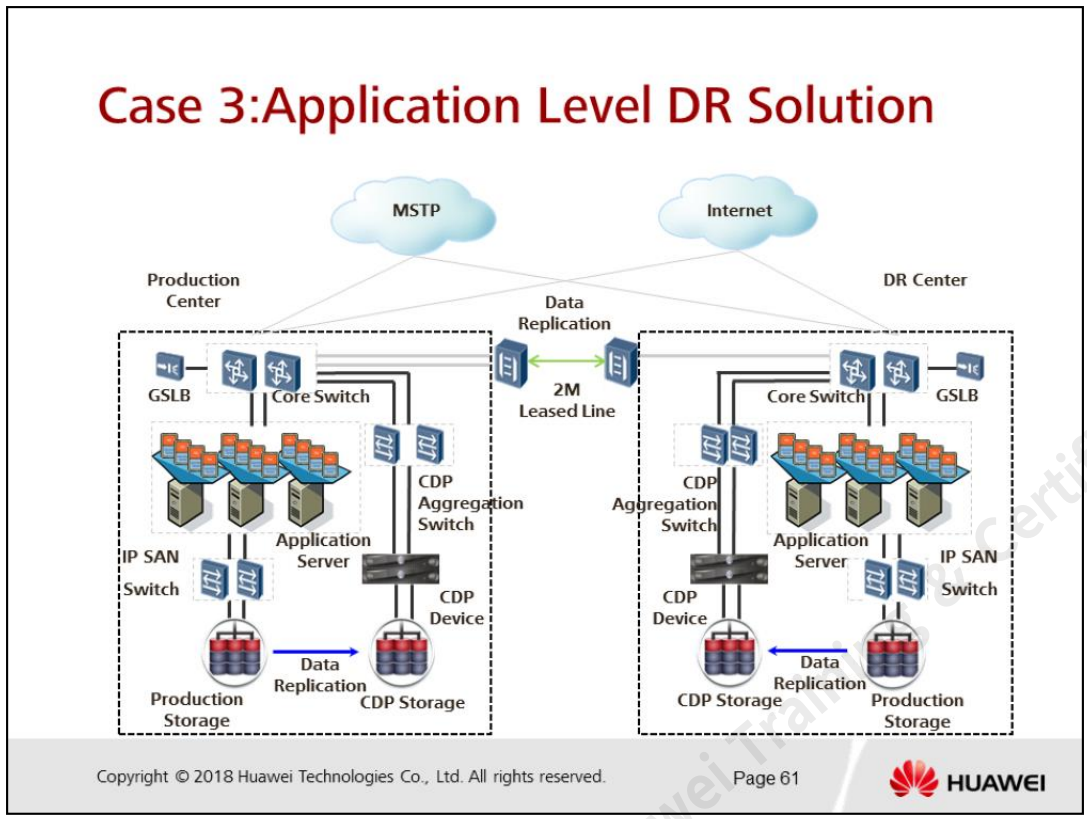
- The challenges faced:
  - The 18 special offices of the XX Audit Firm are distributed in various provinces and have a wide geographical distribution. The investment in disaster recovery and construction is extremely demanding and the implementation of the later projects is difficult.
  - The technical expertise of the various agencies within the firm is weak, making it difficult to maintain and perform DR Drill exercises of the disaster recovery system.
  - The firm has large amount of data, and high demand for disaster recovery network.
- Huawei Solution:
  - All the 18 branches were deployed with a storage system, and the disaster recovery is centralized to the disaster recovery center in Changsha. The whole DR system deployed 21 sets of medium-to-high-end arrays with a total deployment capacity of more than 1PB.
  - Utilized Huawei's 32:1 remote replication technology to achieve centralized disaster recovery on the existing network.
  - The DR Drill software is customized at the disaster recovery center in Chang XX City to meet the requirements of users for the maintenance and drill of the disaster recovery system.

Huawei Training & Certification



- The challenges faced:
  - The customer already has a vSphere virtual data center and needs to build a new data center for disaster recovery.
  - Customer requires Low TCO, and high return on investment.
- Huawei Solution:
  - Deploy a set of IT systems in disaster recovery centers, including storage, services, networking, and virtualization platforms.
  - Huawei UltraVR disaster recovery components is installed in production and disaster recovery centers.
  - The host of the virtual machine, which is installed with the consistency agent called "ConsistentAgent", that implements application-level protection of the virtual machines.
- Value to Customer:
  - There is no need to change the existing network architecture.
  - Flexible setting of disaster recovery strategy and policies, with one-click recovery of the systems.
  - Support disaster recovery drill exercise and disaster recovery switching/failovers.

Huawei Training & Certification



- The challenges faced:
  - The current state of IT is difficult to support business development needs, and it is difficult to ensure the continuous operation of online business.
  - IT operation and maintenance complexity, high energy consumption, and low resource utilization.
- Huawei Solution:
  - Move the business systems to Huawei Cloud Platform.
  - Deploy CDP(Continuous Data Protection) storage and CDP software in two data centers and use CDP technology to implement application-level disaster recovery for both data centers in the same city.
- Value to Customer:
  - Resilient resources, reuse of resources, improve resource utilization, and reduce operation and maintenance costs.
  - The key business RTO and RPO are 0. When the production center fails, services and data are automatically switched to the disaster recovery center.
  - Guaranteed business continuity.

Huawei Training & Certification

Summary

- This module mainly let us learn about:
  - The concepts of DR, and the importance of constructing the DR systems.
  - The common DR solutions and the advantage and disadvantages of these solutions.
  - The different technologies used to implement these DR solutions and the typical DR application cases.

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

Huawei Training & Certification
Huawei Training & Certification

Quiz

1. (True or False) Synchronous Remote Replication can achieve RTO=0 and RPO=0. ( )
2. (MCQ) The core of DR technology is the data replication technology but there are many types of these technologies, thus which of the following 3 layers can be deployed with devices that implements data replication functionality? ( )
   A. Application Layer.
   B. Host Layer.
   C. Network Layer.
   D. Storage Layer.

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

- Answers:
  - False.
  - BCD.

Huawei Training & Certification
Huawei Training & Certification

Thank You

www.huawei.com

Huawei Training & Certification
Huawei Training & Certification

12 Introduction to Data Center

www.huawei.com

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

Huawei Training & Certification
Huawei Training & Certification

Foreword

- This module mainly introduces the development history of Data Centers, and the basic modules and components of Data Centers. This module also introduces the evolution and trends of Cloud Data Centers.

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

Huawei Training & Certification
Huawei Training & Certification

Objectives

- Upon completing this module, you will be able to:
  - Understand the history of Data Centers;
  - Familiar with the basic components and modules of the Data Center;
  - Understand the evolution and current trends of cloud Data Centers.

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

Huawei Training & Certification
Huawei Training & Certification

Contents

1. **History of Data Centers.**
2. Basic Components and Modules of Data Centers.
3. Challenges of Cloud Data Centers and its Development Trends.

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

Huawei Training & Certification
Huawei Training & Certification

Evolution of Data Centers



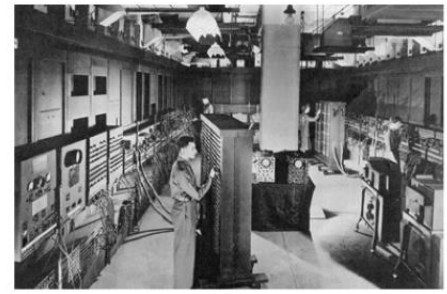
Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

Huawei Training & Certification
Huawei Training & Certification

Earliest Development of Data Centers - ENIAC

- ENIAC (Electronic Numerical Integrator And Computer) was developed in 1946 for calculating the artillery firing tables for the United States Army's Ballistics Research Laboratory.
- At that moment, there was no other computer that can accomplish the task required for those calculations. ENIAC had 17468 vacuum tubes, 7200 crystal diodes, 70000 resistors, 10000 capacitors, 1500 relays, and more than 6000 switches. It can perform 5000 simple addition and subtraction operations or near to 400 multiplication operations in a second. It was 1000 times faster than a calculator and 20000 times faster than manual hand calculations.



Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

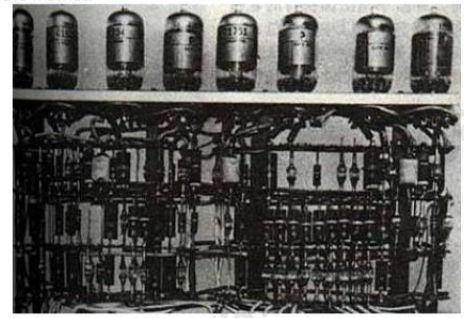
HUAWEI

- The idea of developing an electronic computer was generated during the period of World War II. During those times, the war was going on actively and the military equipment of all the countries was still very primitive, and the main strategic weapons used at the time was planes and artillery. Hence, it was crucial to develop and invent new types of artillery and missiles to aid the front lines of the war, so the US army established the Ballistics Research Laboratory at Aberdeen Proving Ground, Maryland.
- The US Army wanted the laboratory to provide 6 ballistic firing tables each day to the Army's Artillery Unit for missiles research purposes . In reality, each of the firing tables is generated by calculating hundreds of missile trajectories and each trajectory's mathematical model is a very complex set of non linear equations. There is no exact solutions can be calculated for these complex equations and can only be numerically approximated, which means that a lot of numeric needs to be calculated to find the approximate values for these equations.

Huawei Training & Certification - Huawei Training & Certification

Commercialization of Virtualization Technologies in Data Centers - TRADIC

- Unlike vacuum tube systems, the first transistor based computer (TRADIC) was developed in 1954, but more advanced commercial systems appeared in the 1960s and lead to a breakthrough in mainframes technology such as the IBM System Mainframe. Mainframes are primarily developed for government and military use at the time and have unique requirements for environmental sites and security systems. With the increasing demand for system performance, people are hoping to provide support for additional features and implement resource sharing on the same system. As a result, virtualization technologies gradually came into the attention and viewpoint of the market.
- The concept of virtualization quickly gained traction and was further refined by the multitasking processor mechanism in the mainframes. The first commercially available virtualization technology was implemented on the IBM VM/370 Operating System in 1972.
- As one of the most important technologies of the data center, the development of virtualization technology also becomes an important part in the entire Data Center development process.



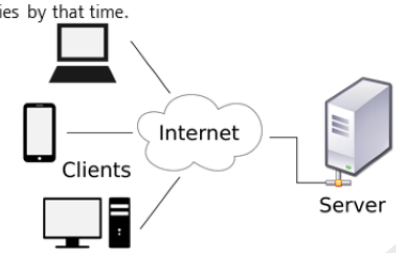
Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

- Transistor computer is the second generation of the electronic computer. In 1954, Bell Laboratories in the United States successfully developed the first computer using a transistor circuit called "TRADIC" with 800 transistors.
- In the first generation of computers which is before the 1950s, computers used vacuum tubes as components. Vacuum tubes or also known as electron tubes is a device that controls electric current between electrodes in an evacuated container. However, during operation, these tubes generates too much heat, has poor reliability, has slow operation speed, expensive and are bulky in size. These factors caused the development of computers to be limited and as a result, transistors started to be used as a component for computers. Transistors not only able to realize the same functions of a vacuum tube, but they also have the advantages of being smaller in size, light in weight, has longer lifespan, higher efficiency, generates less heat and has lower energy consumption. After the transistor is used as the component of the computer, the electronic circuit structure of the computer is greatly improved, making high speed electronic computers even easier to be developed and implemented.

Huawei Training & Certification

C/S Architecture Computing Model and the Internet

- 20 years ago, the microcomputer industry flourished, older generation of PC lost its charms and slowly faded from the market, and was slowly replaced by connected network devices, especially when Client/Server technology model emerged, which led to gradual emergence of external hosting and external Data Centers.
- In the mid of 1990s, the emergence of the Internet brought a huge impact on the market and provided more choices for the deployment of Data Centers in the next decade or so. With the company's needs for support for Internet business applications, network connectivity and collaboration services becomes the must-have IT services to deploy by businesses. Network providers and colocation vendors have been widely adopted in the creation of hundreds of data centers, and the data center as a service model is already accepted by most companies by that time.



Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

- C/S Architecture:
  - Client/Server architecture is a well known software system architecture which reduces the communication overhead of the system by reasonably allocating tasks to the clients and servers. It requires that clients to be installed for running management operations.
  - Client and server side programs are different in nature, user programs are mainly installed at Clients and the Server's main function is to provide data management, data sharing, data and system maintenance and concurrency control for the client program to complete user's specific service or business needs. For example, servers can provide databases services to multiple clients while client program can access and manipulate the data in the databases via a client side program interface.
  - Development is made easier with this architecture, with greater ease of operation. However, application upgrade and client programs maintenance is a bit difficult in this architecture. For example, during a new upgrade of the system, all the clients need to upgraded and patched for every new features which increases the maintenance workload.

Huawei Training & Certification

Emergence of Modular Data Centers

- In the recent years, Data Center has evolved into a new form which is the modular data center. It deploys Data Center servers and devices typically in containers, hence it is also known as Container Data Center. One of the best known modular Data Center is the Sun Blackbox, where its 280 servers are deployed in standard 20 feet containers and shipped around the world.
- These types of modular Data Center lacks in outer appearance compared to the well designed traditional Data Center, but it takes up only 1% of the cost of building a traditional Data Center. It has flexible mobility and dramatically reduces the time in deployment cycles, making deploying Data Centers in remote regions faster and more efficient.



Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

- Modular data center is a new generation of data center deployment based on cloud computing. In order to cope with the trends of such as cloud computing, virtualization, centralized and high-density computing developments in server technologies, modular data center adopts a modular design concept to minimize the impact of infrastructure coupling with equipment room in modern Data Centers. Essentially, modular Data Center allows Data Center equipment and services to be quickly deployed without the need of building Data Center infrastructure beforehand. It integrates subsystems such as power distribution, cooling, cabinets, airflow containment, integrated wiring and monitoring equipment together in a container form, allowing Data Center services can be deployed quickly without building all those systems first in a traditional data center. This design improves the overall efficiency, enables fast deployment, allows flexible expansion and energy savings. Modular Data Center design can be found in the smaller form factors of Micro Data Center (MDC) and larger form factor of Container Data Center (CDC).
- A micro Data Center (MDC) is a smaller, modular Data Center system that is designed to solve different sets of problems or to take on different types of workload that cannot be handled by traditional facilities or even large modular Data Centers.

Huawei Training & Certification

Cloud Data Centers

- Software as a Service (SaaS) marked the shift of the demand for computing resources brought by infrastructure to the on-demand ordering model. This business model leverages the network infrastructure and data center operators to jointly deliver large-scale expandable data bandwidth resources to users. They are able to provide a variety of IT services through these resources in the cloud data centers.
- Although in the beginning, everyone did not realize that this kind of development will be so rapid, but cloud service providers such as Amazon and several other infrastructure as a service providers which are based on the cloud data center platform turned out to have mass number of subscribers for their services.



Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

- Cloud computing is an information technology (IT) paradigm that enables ubiquitous access to shared pools of configurable system resources and higher-level services that can be rapidly provisioned with minimal management effort, often over the Internet.
- The cloud refers to software and services that run on the Internet, instead of locally on your computer. The advantage of the cloud is that you can access your information on any device with an Internet connection. Popular examples of cloud services includes Software as a Service (SaaS) where your software is fully installed and managed on the cloud without the need of installing it at your local server or local Data Center.
- Cloud Data Centers are optimized and built for the purposes of running the Cloud.

Huawei Training & Certification

Contents

1. History of Data Centers.
2. **Basic Components and Modules of Data Centers.**
3. Challenges of Cloud Data Centers and its Development Trends.

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

Huawei Training & Certification
Huawei Training & Certification

What is a Data Center ?

- Data Center usually refers to a physical space that centrally process, store, transmit, exchange and manage information. Computers, servers, network and storage equipment are generally considered as key Data Center equipment. The environmental factors that are required for the operation of these key Data Center equipment are the power distribution systems, cooling systems, cabinet systems, fire protection systems, and monitoring systems which are generally considered as key physical infrastructures of Data Centers.

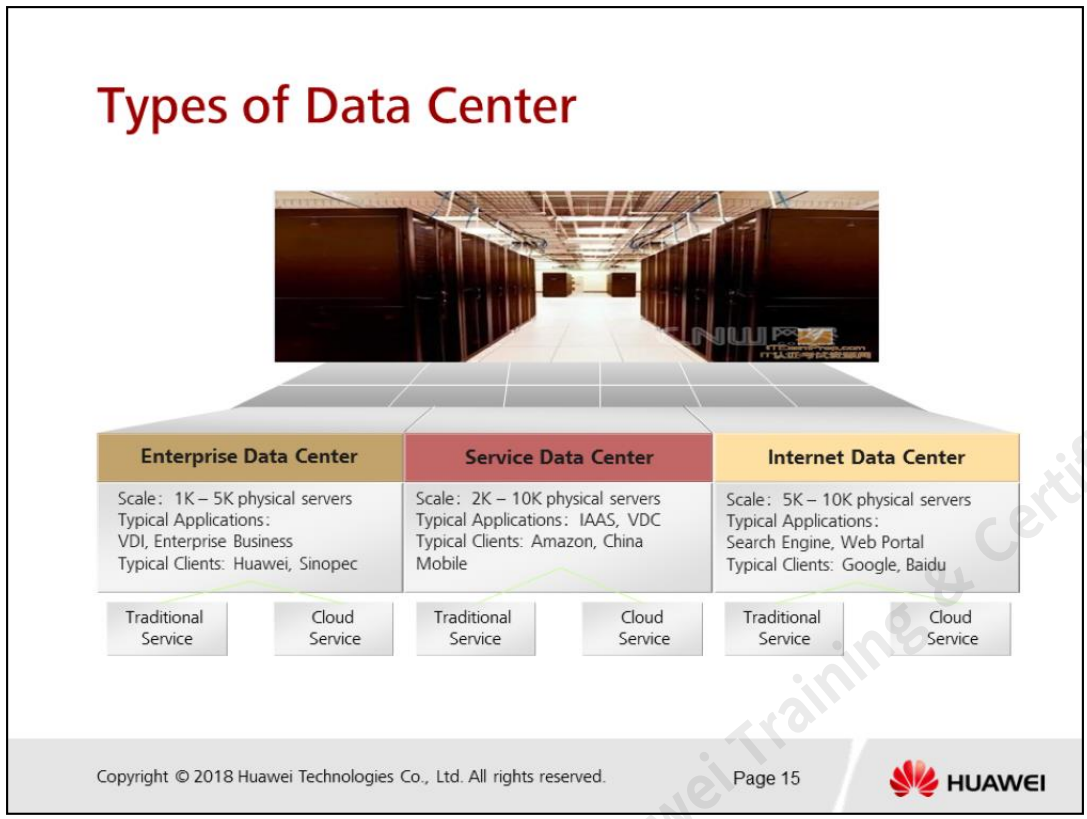


Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

- Wikipedia defines a data center as : "A facility used to house computer systems and associated components, such as telecommunications and storage systems. It generally includes redundant or backup power supplies, redundant data communications connections, environmental controls (e.g. air conditioning, fire suppression) and various security devices."
- Data Centers are not a single thing, but rather, a combination of elements. At a minimum, Data Centers serve as the principal repositories for all manner of IT equipment, including servers, storage subsystems, networking switches, routers and firewalls, as well as the cabling and physical racks used to organize and interconnect the IT equipment. A data center must also contain an adequate infrastructure, such as power distribution and supplemental power subsystems, including electrical switching; uninterruptable power supplies; backup generators and so on; ventilation and data center cooling systems, such as computer room air conditioners; and adequate provisioning for network carrier (Telco) connectivity. All of this demands a physical facility with physical security and sufficient physical space to house the entire collection of infrastructure and equipment.

Huawei Training & Certification



- Data Centers are facilities built to host and access the server's services, and mainly it can be categorized into a few types such as: Enterprise Data Center, Service Data Center and Internet Data Centers.
- Each type of these data centers have different needs that they cater to, and the scale of these data center varies by the types of services, application scenarios and client requirements.
- Enterprise Data Center usually caters to the enterprise business needs such as virtual desktop environment(VDI) and hosting enterprise business applications. The scale of the enterprise data center depends on the business requirements and can be expanded as per needed. It is usually owned and maintained by enterprises internally.
- Service Data Center usually caters to the needs of service providers that provides services such as Infrastructure as a Service (IAAS) and Virtual Data Center (vDC) to the public. It usually have similar scale or higher scale compared to enterprise data centers. Clients that subscribes to these services do not have to maintain the data center themselves and could just consume the service and resources subscribed. All the data center equipment is maintained by the service provider and resources are sold as a package to the clients.
- Internet Data Center usually caters to the needs of the Internet. It has the largest scale compared to the other types of data center as it need to serve a large number of users on the Internet. Large search engines and web portals such as Google and Baidu are usually hosted on the Internet Data Centers.

Huawei Training & Certification

Components of Data Center

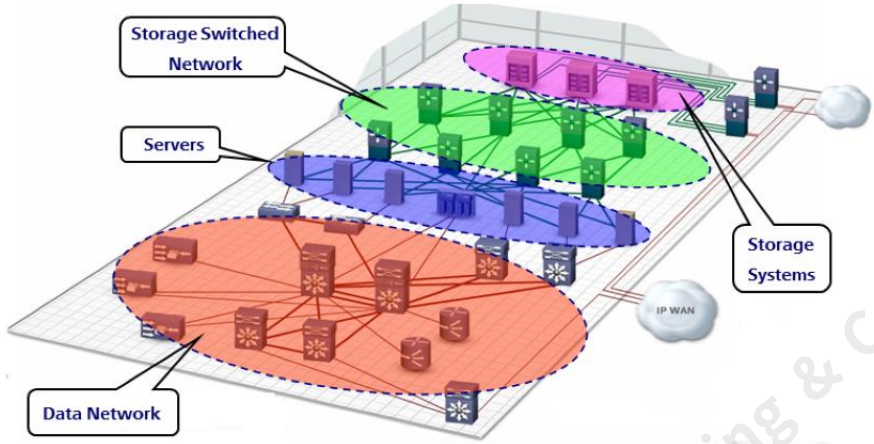
- Data center is a business-oriented infrastructure that is the core to supporting enterprise business operations and future growth, it mainly includes the following components:
- Safe Web Architecture.
- Reliable Supporting Equipment (Server Room, Generator, UPS, Air Conditioning etc.)
- Unified Server/Application Platform.
- Centralized Storage and Backups.
- Unified System Management Platform.
- Service-oriented Operation & Maintenance (O&M) Organization and Process.

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

Huawei Training & Certification
Huawei Training & Certification

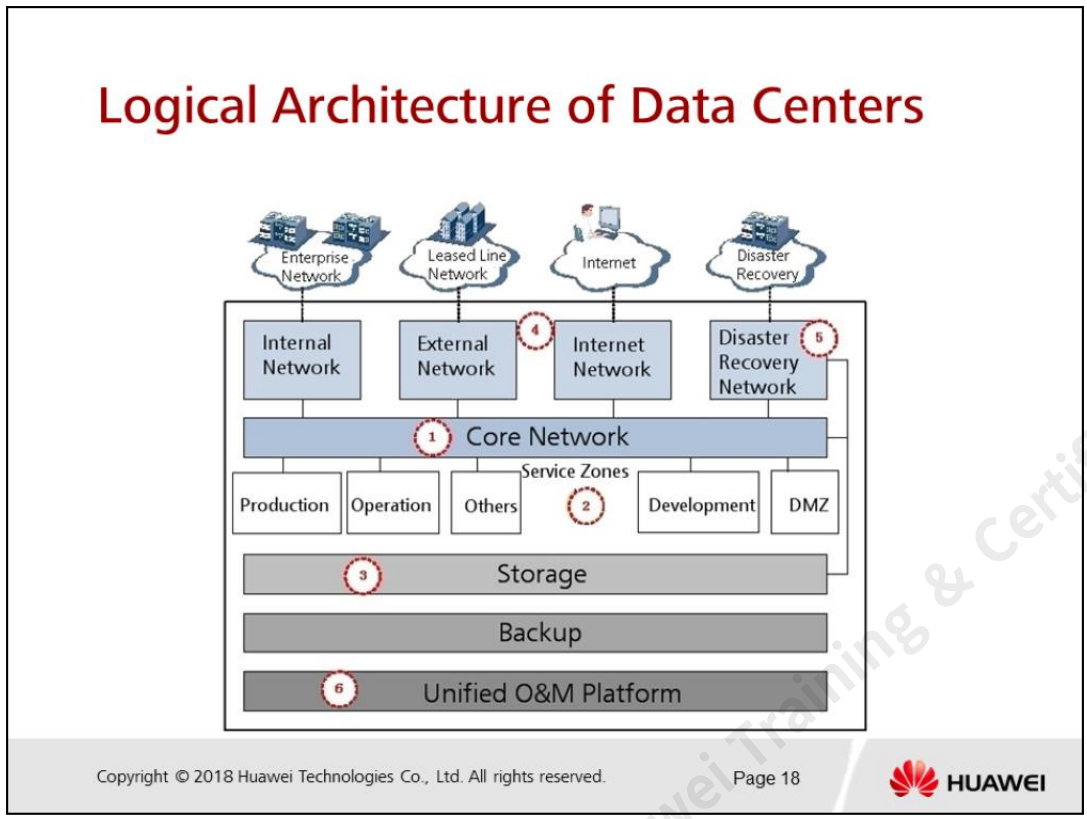
Standard Data Center Architecture



Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

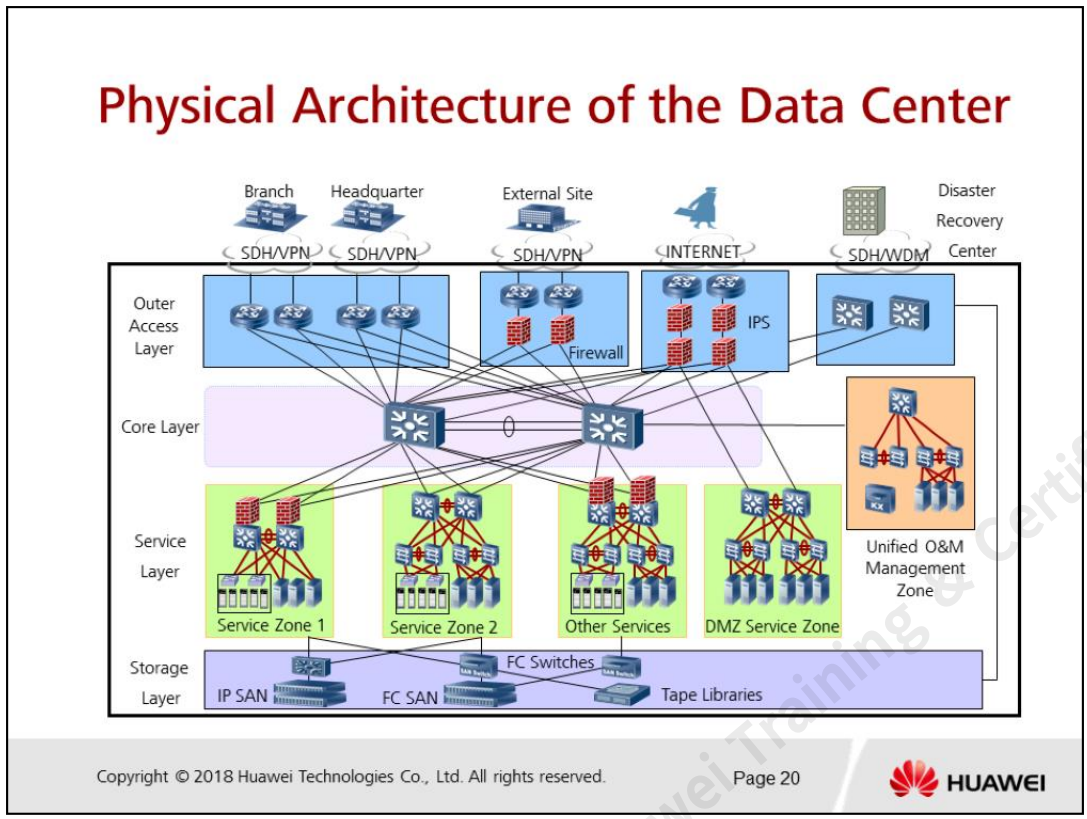
HUAWEI

Huawei Training & Certification
Huawei Training & Certification



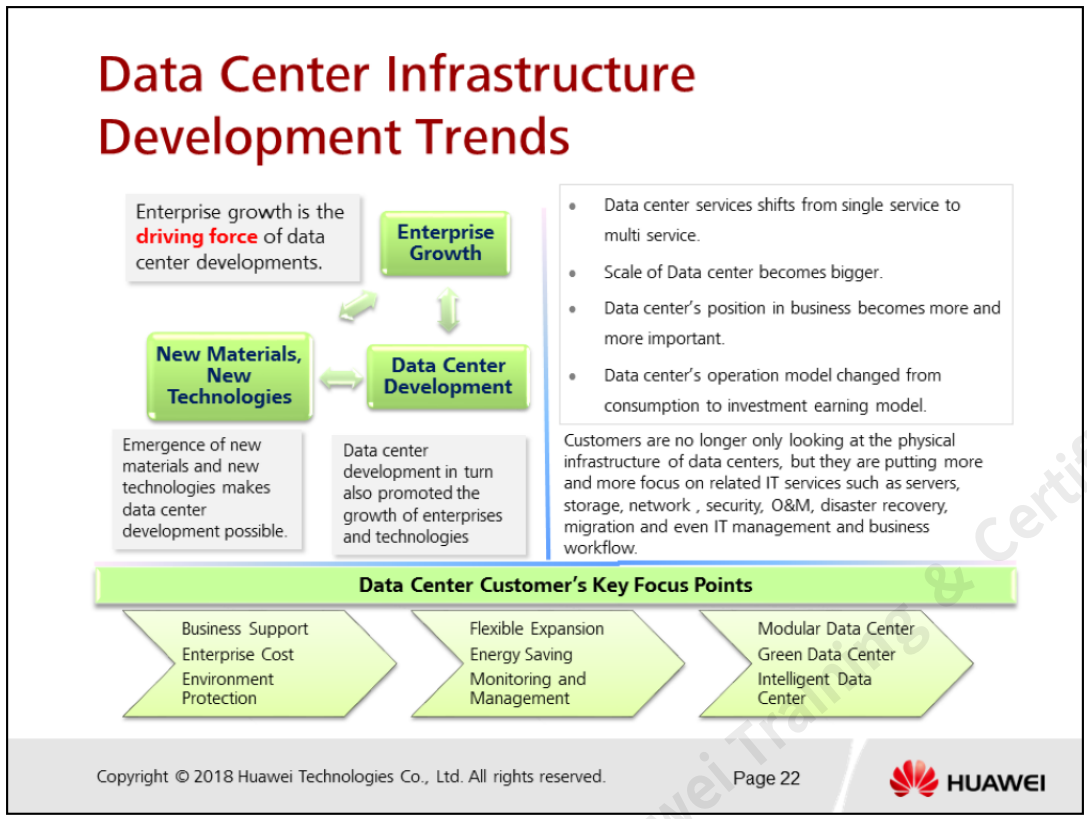
- Core Network Zone: It is the core of the data center network. It connects the internal server areas, enterprise internal network, partner's network, disaster recovery center and the external user access network.
- Server Zone: This the area for the deployment of servers and application systems. Based on security and scalability considerations, it can be further divided into: Production area, Business Operation area, Testing area, and DMZ (Demilitarized Zone) area.
- Storage Zone: This area includes the storage and network devices for the FC SAN and IP SAN(Storage Area Network).

Huawei Training & Certification



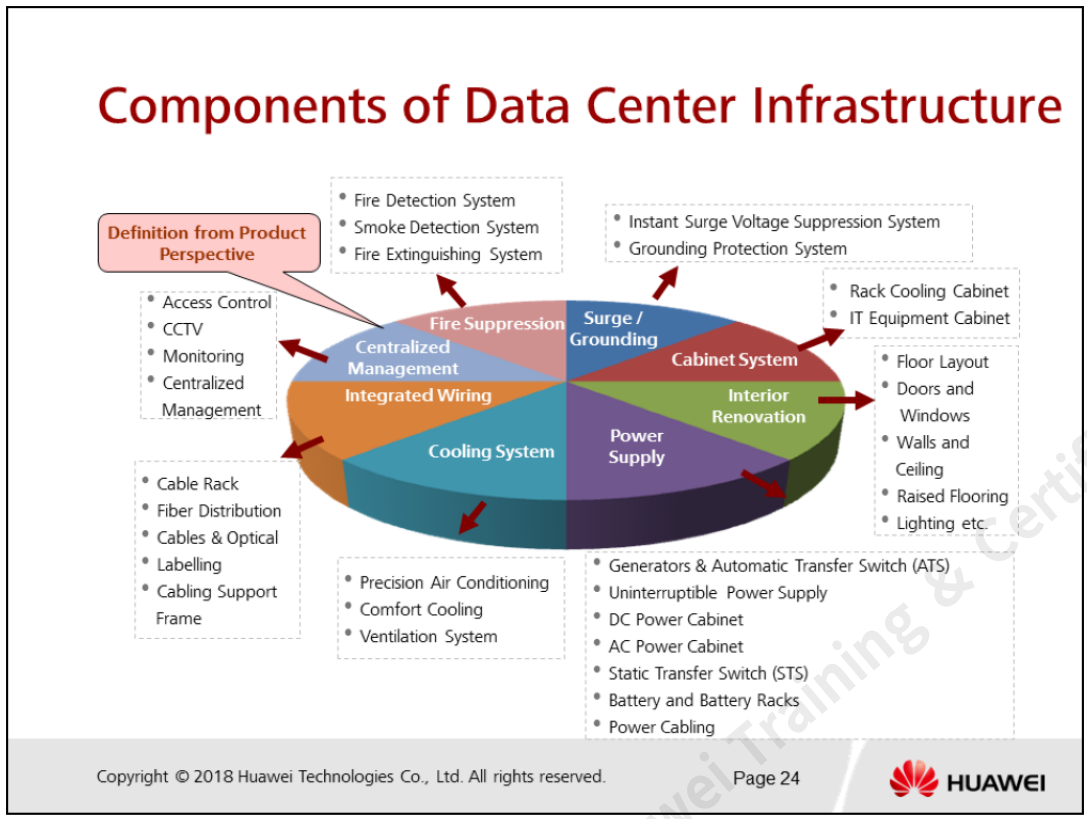
- The diagram above shows the physical architecture of the data center. Generally, the data center is segregated into different layers that that has different functions for better scalability and security.
- The Access layer provides the connectivity for the data center equipment to the devices and networks and sites outside of the data center. It provides connection to the headquarter and branches using Synchronous Digital Hierarchy (SDH) and Virtual Private Network (VPN) technologies. Data center also connects to the Internet via the Access layer, and it connects to the Disaster Recovery Center via SDH or Wavelength Division Multiplexing (WDM) technologies.
    - SDH is a standard technology for synchronous data transmission on optical media allowing real-time data transmission using optical cables .
    - VPN essentially extends a private network across a public network, and enables users to send and receive data across shared or public networks as if their computing devices were directly connected to the private network. Both these technologies are frequently used in data centers worldwide for external connection to private enterprise network.

Huawei Training & Certification



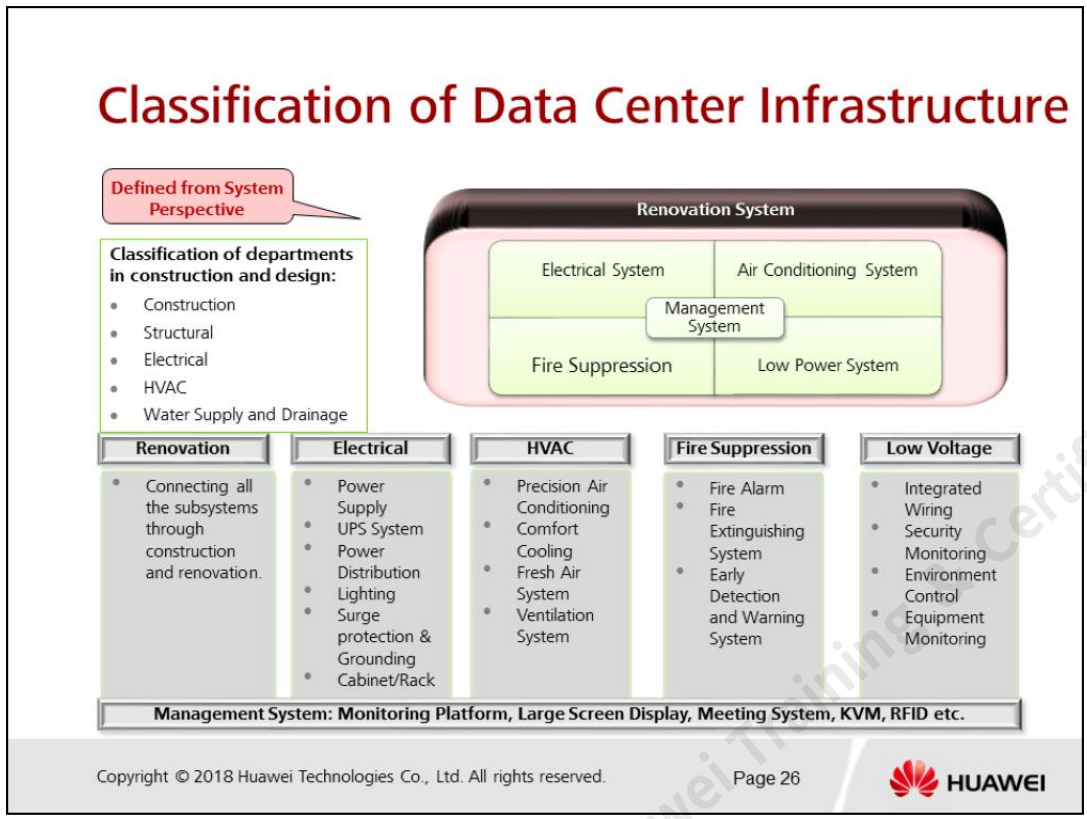
- IT systems and information construction have become the driving force of enterprise growth and development, which means that enterprise growth or more specifically its business growth has generated continuous demand on IT systems and in turn became the primary driving force on data center development.
- Information construction refers to the process of increasing business performance in productivity, management, operation and efficiency through the use of modern information technologies such as computing, network to extensively use information resource in depth. In context, this process can refer to the digitization of business information in aspects of sales, management, products and making use of these information in business decision making process for better business performance.
- On the other hand, data center developments such as the emergence of multiple features and services in data centers also contributed towards the growth of enterprises. At the same time, the development of data center provided new demands and requests towards materials and technologies. The appearance of new materials and technologies makes the demands of data centers possible and promotes the growth of both data centers and enterprises. All three factors stated here is interrelated to each other and creates a cycle for growth and development between each other.

Huawei Training & Certification



- Previously, we have introduced the components, features and classification of the data center infrastructure from the perspective of physical space. In the following segment, we shall look at the components of the data center infrastructure from the perspective of systems, including what type of systems that makes up the physical infrastructure of the data center.
- As we have known, the data center physical infrastructure consist of multiple subsystems. The picture above shows definition of multiple systems in data center infrastructure based on Huawei Level 1 (L1) data center design. Based on this definition, the physical infrastructure of data center is classified into 8 different systems.
- Power Supply System: It includes diesel generators, Automatic Transfer Switch (ATS), Uninterruptible Power Supply (UPS), DC power cabinet, AC power cabinet, Static Transfer Switch (STS), backup battery and battery racks, and power cabling.
- Cooling System: It includes precision air conditioning, comfort cooling and ventilation systems.
- Other than that we have the interior renovation, cabinet system, lightning surge protection and grounding system, fire suppression system, integrated wiring system, and centralized management system that makes up the main 8 systems of the data center infrastructure. Each system has multiple features and functionalities, but we won't go into much detail on each of the systems here.

Huawei Training & Certification



- For the following segment, lets look at the content of the Data Center L1 Physical Infrastructure Design from a system perspective without focusing on products or equipment.
- Building a data center involves the expertise of multiple different fields in order to build and interconnect all the different subsystems in the data center infrastructure. Generally, the design and construction institute provides a systematic classification on the departments of specialization in the fields such as construction, structural, electrical, HVAC(Heat, Ventilation & Air Conditioning), water supply and drainage. However, due to the special nature of the data center, the industry commonly classify the different fields of expertise required to build a data center into different systems such as renovation, electrical system, air conditioning system, fire suppression system, low power system and management system. Some of the data center building or design companies even classify the management system into low voltage distribution system).

Huawei Training & Certification

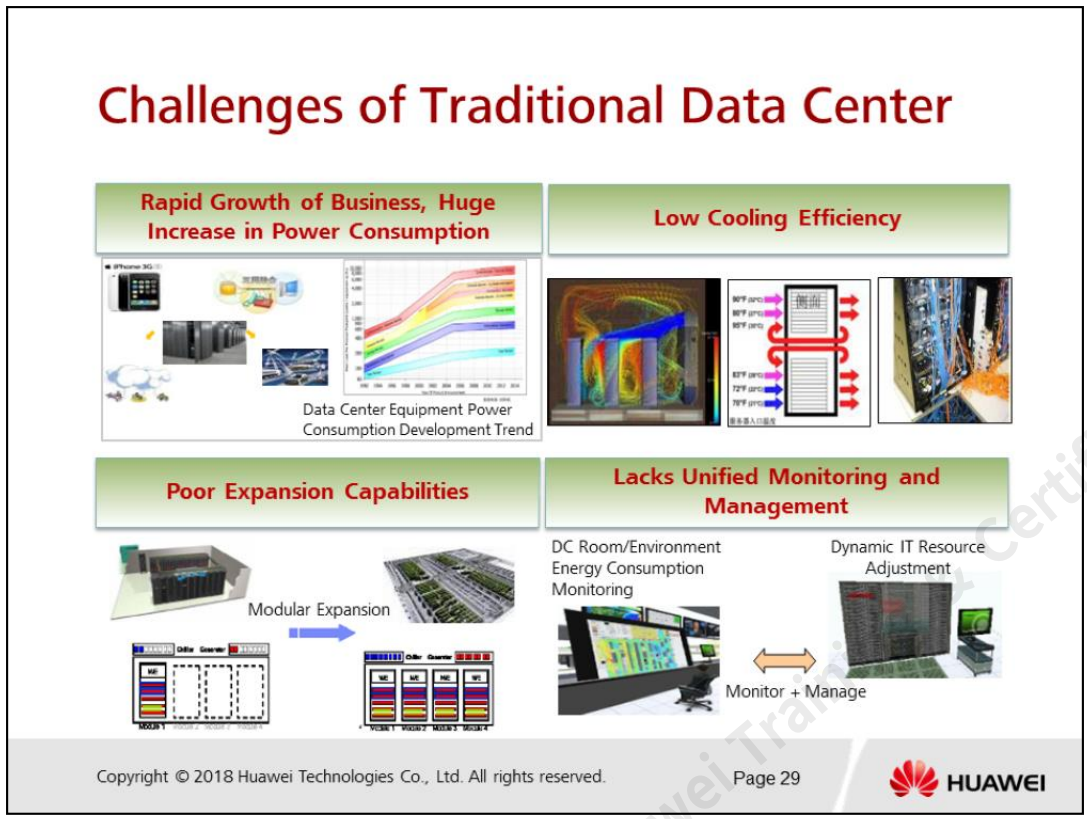
Contents

1. History of Data Centers.
2. Basic Components and Modules of Data Centers.
3. **Challenges of Cloud Data Centers and its Development Trends.**

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

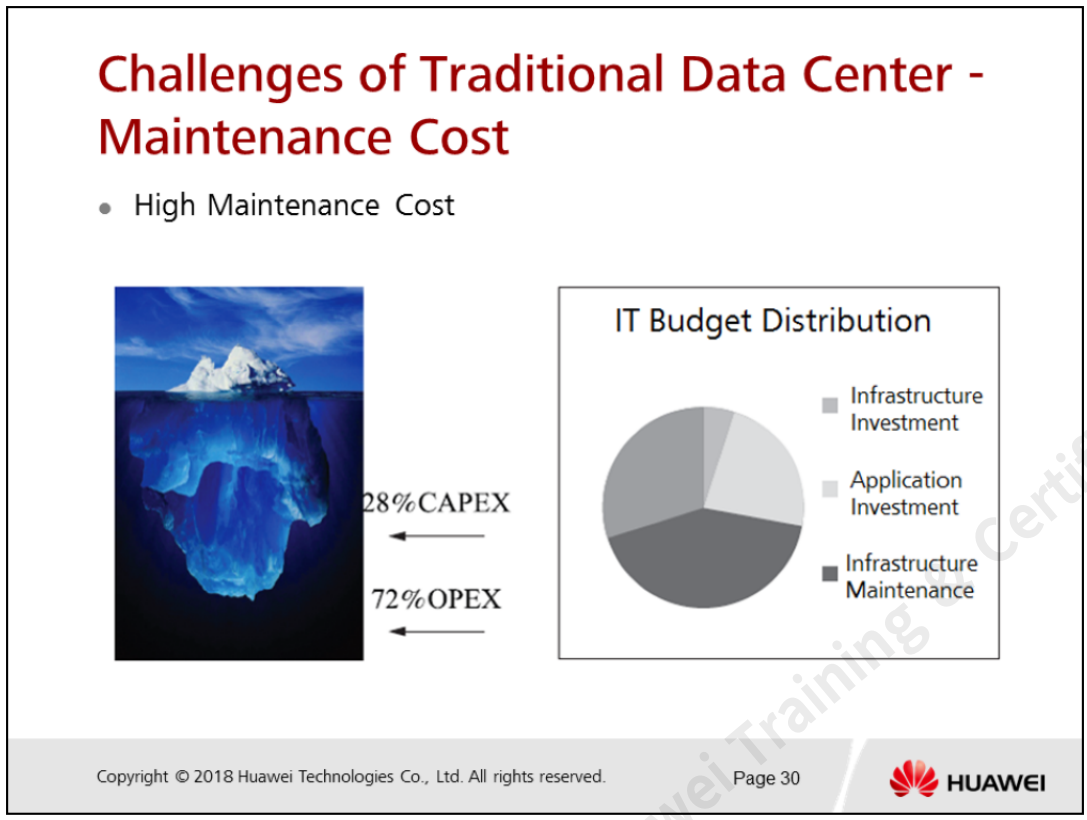
HUAWEI

Huawei Training & Certification
Huawei Training & Certification



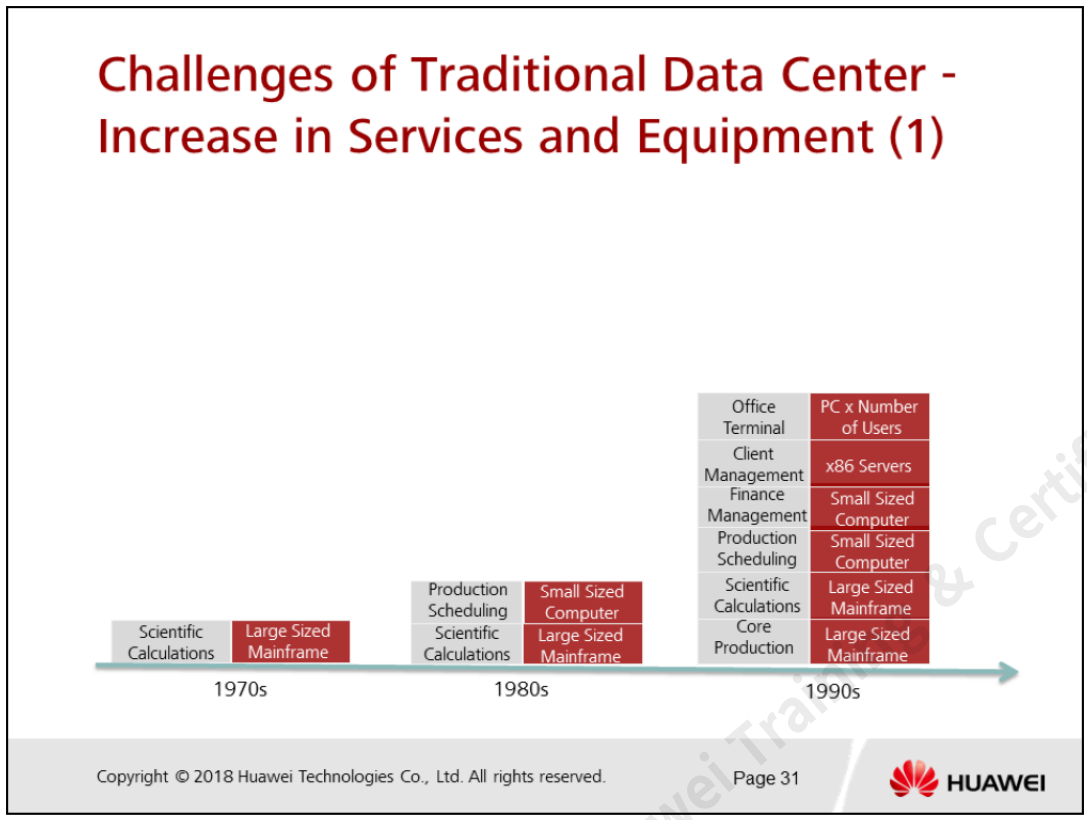
- With the growth of 3G, Cloud Computing, and Internet of Things, the power consumption of data center dramatically increased. Traditional data center power density increased from 3-5kW/cabinet to 10kw+ and overheating of certain areas in the data center became the biggest risk of safe operation of data center equipment.
- Unidirectional placement of cabinets, cascaded heating, upper air ventilation, serious mixture of hot and cold airflow, no airflow separation boards within cabinet, return of hot airflow back to the cabinet, and wrong cabling arrangement causes the airflow supply and extraction to be hampered which affects the maintenance and lowers the reliability.
- The data center building cycle ranges from 1 to 2 years, in the early stages we are unable to predict accurately the growth of the business, so the data center equipment needs to be deployed on demand and expanded modularly so that resources are highly utilized.
- Traditional data center involves only monitoring and not management, newer generation of data centers involves both monitoring and management where IT equipment and physical infrastructure of the data center can intelligently interact with each other in order to increase management efficiency.

Huawei Training & Certification



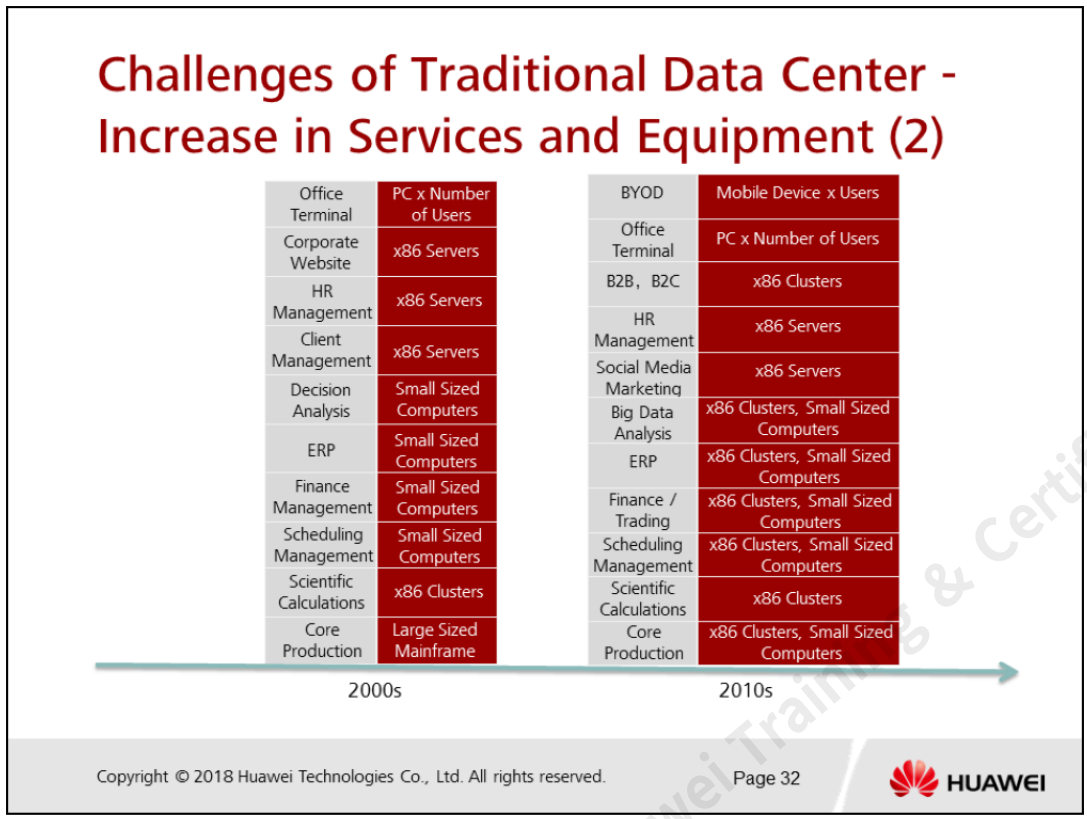
- Data centers are now becoming the synonym to "Cost Center" and "Energy Consumption Center". The increasing number of servers and other IT equipment not only increased the capital expenditure (CAPEX) of the enterprise but also gradually increasing the operating expenditure(OPEX). Based on Accenture's survey on the market, operating cost has taken up to 72% of the total IT budget. The cost involved of running a data center can be best expressed as the tip of the iceberg as shown on the picture above.

Huawei Training & Certification



- What the chart above is trying to convey is that IT is getting more and more complex with the mixed deployment of computers and x86 servers, and it not trying to convey the trend that computers are taking over more and more roles in the industry. The increased number of equipment and complexity of deployment and management poses a great challenge to the traditional data center.
- The growth of enterprise IT is a process of where "human labor" and "human brain" is being replaced by IT equipment. Enterprise IT is a very complex system similar to the human body. When Enterprise IT replaces more and more tasks that previously required manual human effort, the complexity increases. Hence, as the type of business services that need to be supported by IT increases, so does the number of equipment used increases. If we factor in the fact that scale of enterprises that keeps getting bigger, Enterprise IT has become a very huge and complex system.

Huawei Training & Certification



- For an example, consider the scale of a large scale enterprise such as Telefonica, they have:
  - 85+ Data Centers
  - 50,000+ Servers
  - 35+ Programming Languages
  - 5+ Operating Systems, with 20+ different version
  - 4000+ Applications
  - 10+ Databases
  - 25+ Middleware
  - 10+ PB Storage
- Another large enterprise such as Huawei has 150,000 employees, spans across 140+ countries, with 700 offices, 1.92 billion emails annually, 120 million of orders annually, with the total storage volume reaching up to 5.69PB. Their IT system computing power and storage capacity increases annually by 39% and 32% respectively!

Huawei Training & Certification

Challenges of Traditional Data Center - Increase in Data Traffic (1)



- Enterprises requires stronger processing capabilities, higher connection bandwidth, and much more flexible service interaction. High density 10GE/40GE connection to storage servers, and 40GE/100GE network interconnection is the evolution direction that cloud computing network are heading towards in the future.
- High density and high performance connected devices will become mainstream in the future.

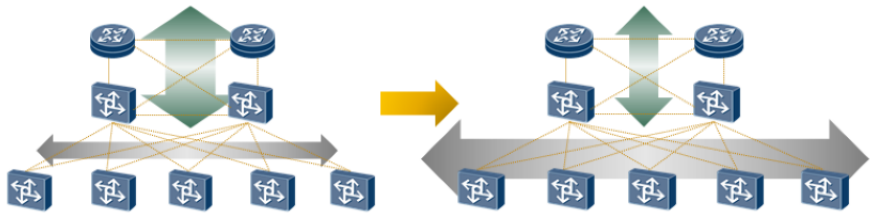
Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

Huawei Training & Certification

Challenges of Traditional Data Center - Increase in Data Traffic (2)

Changes in Data Traffic: Increased Horizontal Flow

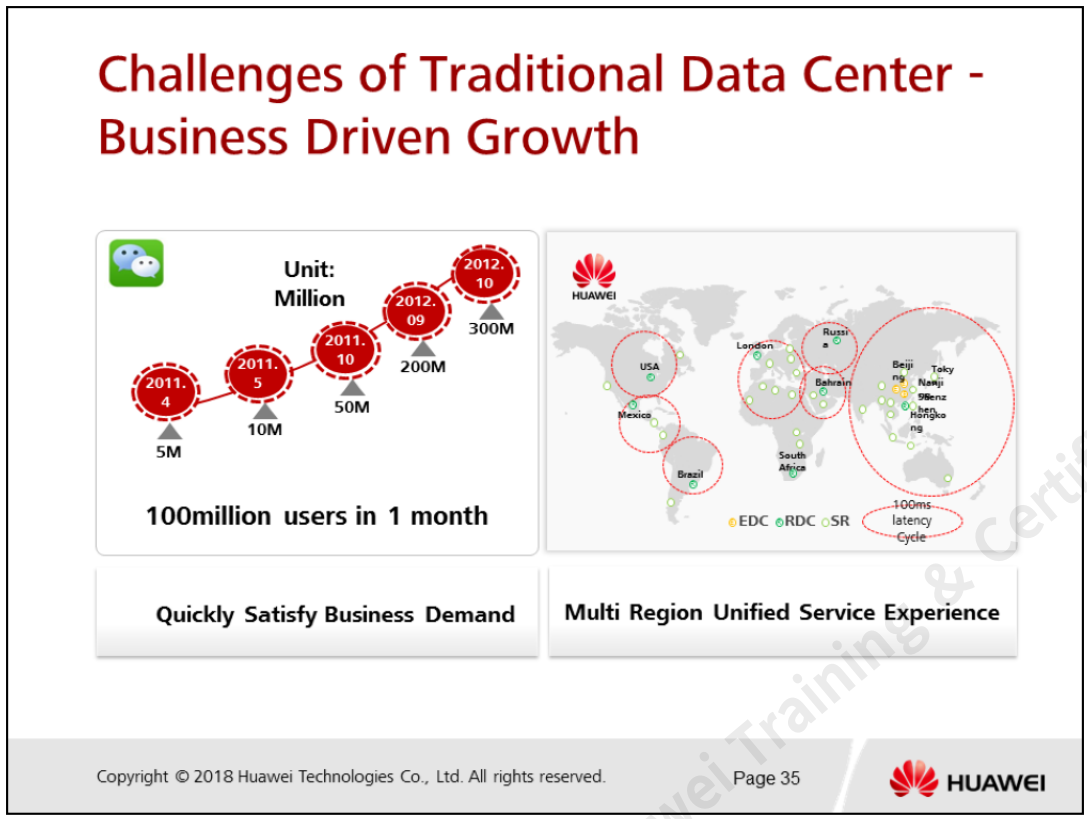


- Big Data, Parallel Processing, 3D Rendering, and Search Engine services requires the combined processing of clustered servers and generates high amount of horizontal flow of data within the network.
- Elastic scaling and online migration of virtual machines, and high amount of data transferred between these virtual machines also contributes to the increased horizontal flow of data traffic within the network.

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

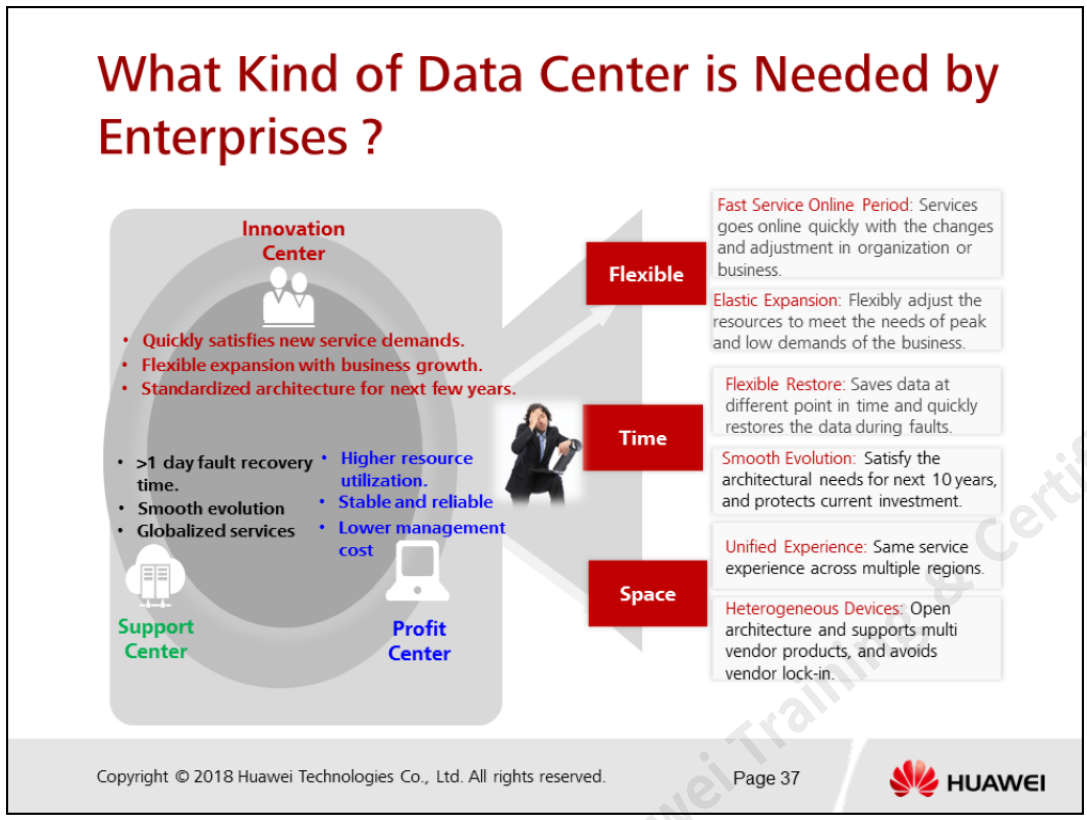
HUAWEI

Huawei Training & Certification
Huawei Training & Certification



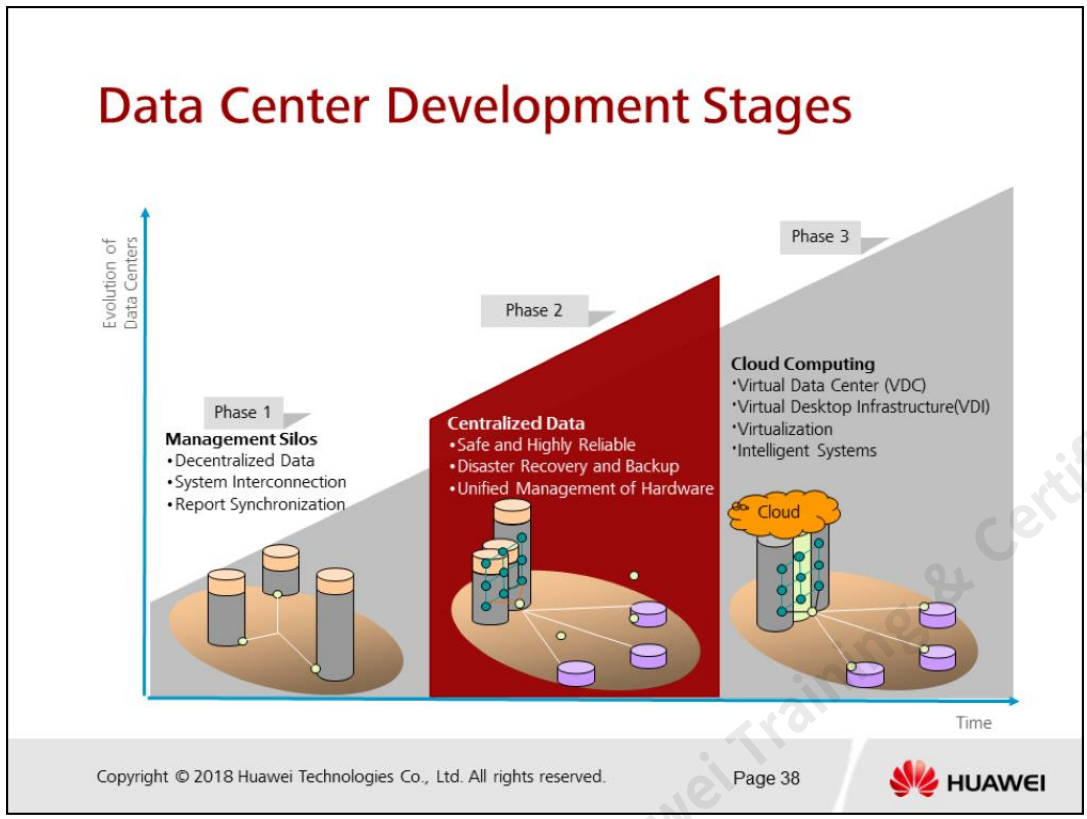
- Wechat launched its first version in 2011, and its user base increased from 5 million to 50 million in 6 months, however it only used 1 month to expand from 200 million users to 300 million users. This is an example that shows businesses are growing in a much faster pace, and how to quickly respond to the growing business resource demand have become one of the greatest challenges of the traditional data center. For another example, Yuebao taken 200 days to reach 250 billion worth of funds, but to reach 400 billion from 250 billion, it only took 1 month. Wechat is a social messaging application whereas Yuebao is a money market fund application that is very popular in China.
- Nowadays, many enterprises especially multinational group of companies, are facing a huge challenge in how to satisfy the different business needs of the region and providing the best service experience when the scale of the business is getting bigger and spans across the globe in multiple countries. As shown on the picture above is Huawei's multiple data center across the globe.

Huawei Training & Certification



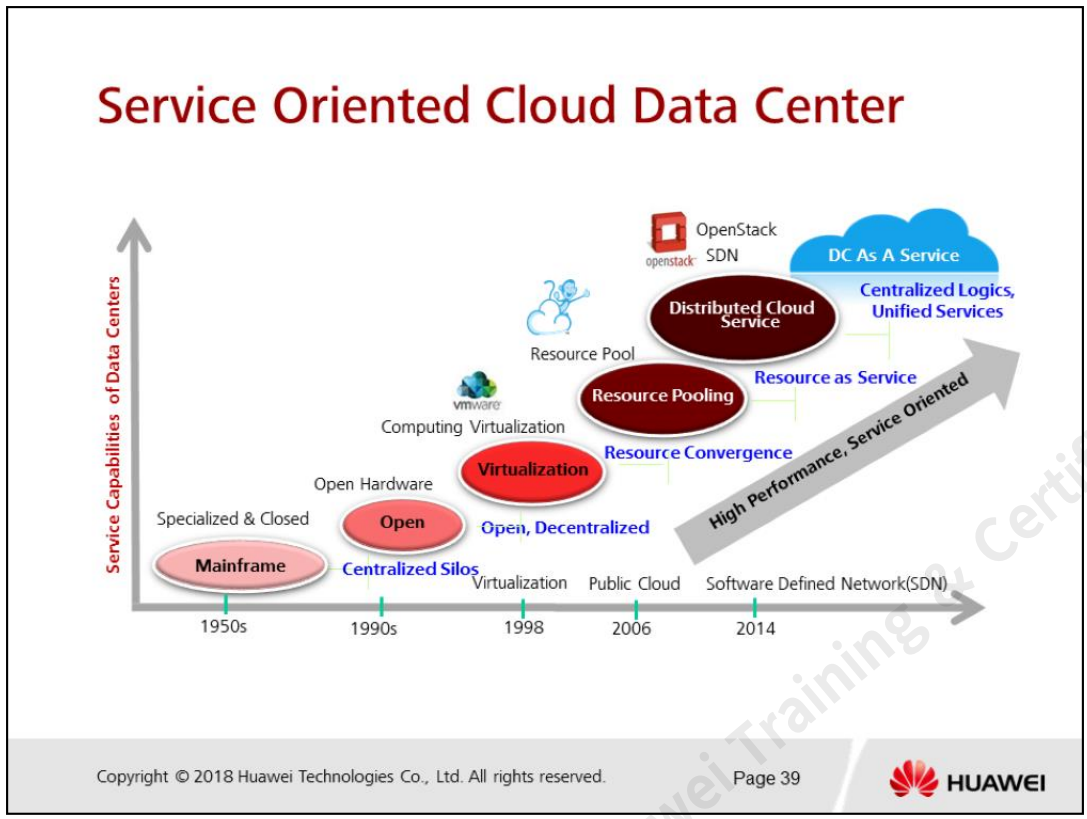
- The 2 key point in this page are :
  - Data center architecture is getting more and more open. Open architecture of data center protects the current investment and makes it much easier to be compatible with 3rd parties in the future.
  - Data center is evolving towards service oriented, and distributed data center which further promotes technologies such as Virtual Data Center (VDC) and distributed computing.

Huawei Training & Certification



- As information are getting highly centralized, development of network technologies, and 100G network bandwidth platform emergence, data centers are evolving towards cloud computing services, following the trends of virtualization, automation and intelligent systems.
- Cloud computing is the next generation data center platform that has the following characteristics:
  - New commercial service model: IaaS(Infrastructure as a Service), PaaS(Platform as a Service), SaaS(Software as a Service).
  - Unlimited expansion of computing resources: IT resources can be obtained at any time based on demand and it is easily scalable.
  - Revolutionary IT and CT technologies: Virtualization, Distributed Computing, Automation, and Intelligent systems.
- A virtual data center(VDC) is a pool or collection of cloud infrastructure resources specifically designed for enterprise business needs.
- Virtual desktop infrastructure (VDI) is virtualization technology that hosts a desktop operating system on a centralized server in a data center.

Huawei Training & Certification



- The trend in data center technologies moves toward open architecture, and flexible service oriented data centers.
- Centralized computing is computing done at a central location, using terminals that are attached to a central computer. The computer itself may control all the peripherals directly (if they are physically connected to the central computer), or they may be attached via a terminal server. Alternatively, if the terminals have the capability, they may be able to connect to the central computer over the network.
- Decentralized computing is the allocation of resources, both hardware and software, to each individual workstation, or office location. In contrast, centralized computing exists when the majority of functions are carried out, or obtained from a remote centralized location. Decentralized computing is a trend in modern-day business environments

Huawei Training & Certification

Unique Features of Cloud Data Centers

- Cloud Data Centers has the following advantages:
  - Increases the utilization rate of IT devices and hardware.
  - Simplifies management.
  - Fast deployment of services, more agile support on enterprise business growth.
  - Energy saving.

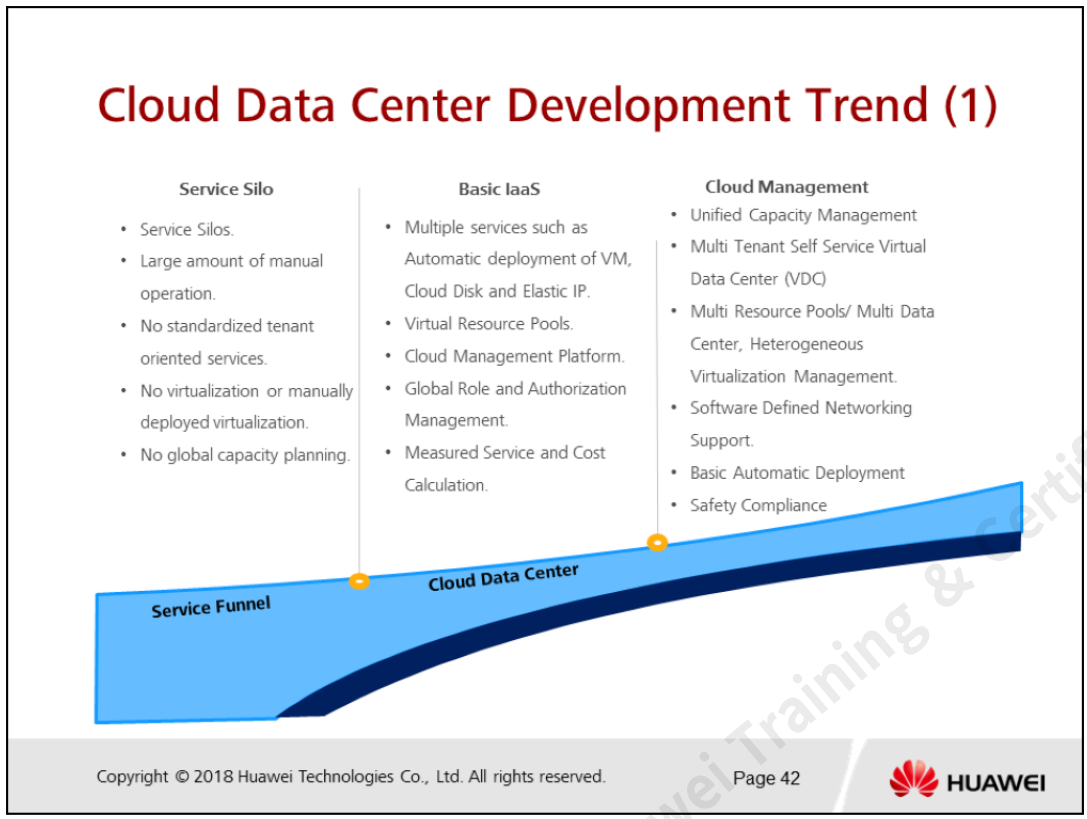


Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

- Increasing Utilization Rate of IT Equipment: Cloud computing enables the sharing of IT equipment such as servers, storage and networks through virtualization so that multiple applications can run on the same physical server. By sharing these resources through virtualization allows the CPU utilization rate to be increased from 15% to 60% or even more.
- Simplifies Management: In cloud data centers, management personnel are facing virtual machines instead of different types of physical servers. Unified management and configuration of these virtual machines can be achieved through cloud management software without the need of worrying about the hardware differences of the servers running those virtual machines.

Huawei Training & Certification



- A silo describes any system that is unable to operate with any other system, meaning it's closed off from other systems. Silos create an environment of individual and disparate systems within an organization. They are systems that have specialized functions that are unable to share its data or resources with other systems.
- Communication within a silo is always vertical, making it difficult or impossible for the system to work with unrelated systems.
- An Elastic IP address is a static, public IPv4 address designed for dynamic cloud computing. You can associate an Elastic IP address with any instance or network interface for any virtual machines(VM).

Huawei Training & Certification

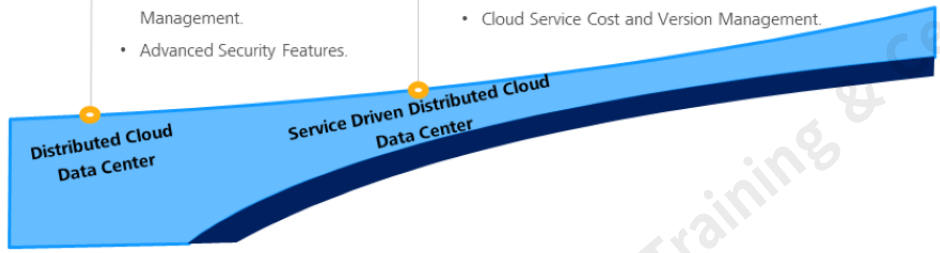
Cloud Data Center Development Trend (2)

Cloud as a Service

- Automated Virtual Network Service.
- Storage as a Service (STaaS).
- Disaster Recovery as a Service (DRaaS).
- Platform as a Service(PAAS).
- Hybrid Cloud Service.
- Complete Cloud Service Catalog.
- Service Scheduling and SLA Management.
- Advanced Security Features.

Service Driven Cloud Service

- Supports Cloudification of Key Services in Industry.
- Supports professional clouds such as Enterprise Cloud, including Power Cloud, Finance Cloud, Police Cloud, Government Cloud, and Carrier Cloud etc.
- Formation of Cloud Service Industry Chain, Independent Software Vendors (ISV) develops industry applications based on Distributed Cloud Data Center (DC^2).
- Intelligent Cloud Service Operation and Maintenance
- Cloud Service Cost and Version Management.

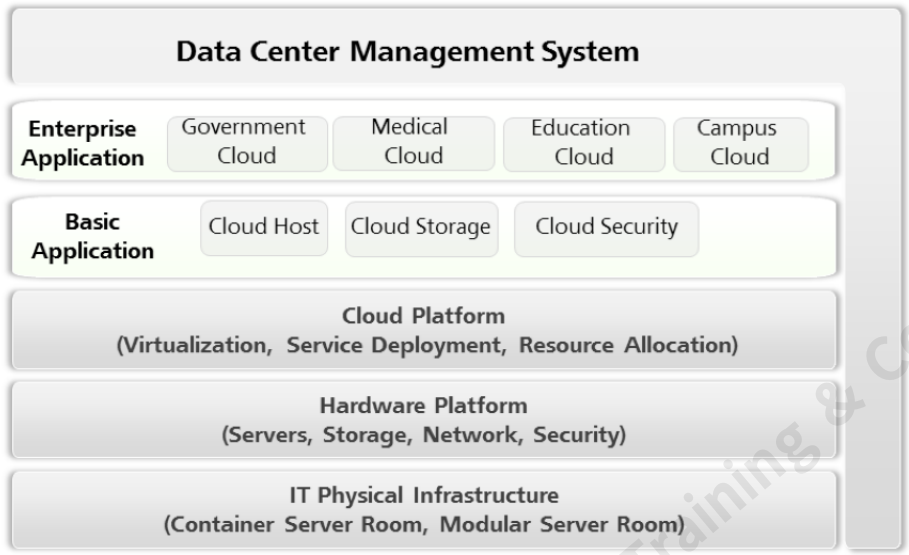


Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

Huawei Training & Certification
Huawei Training & Certification

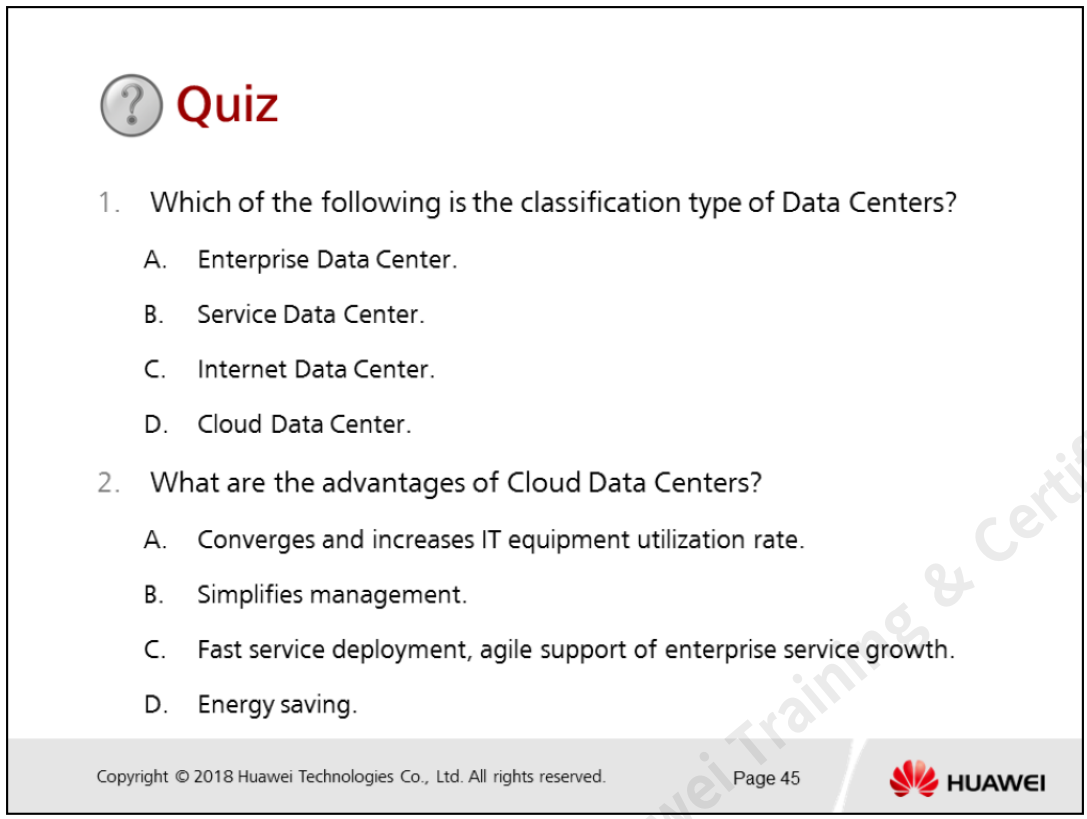
Next Generation Data Center Architecture



Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

Huawei Training & Certification
Huawei Training & Certification



- Answers:
  - ABC. (Refer to page 13)
  - ABCD. (Refer to page 32)

Huawei Training & Certification
Huawei Training & Certification

Summary

- This module mainly introduces about:
  - The Development History of the Data Center.
  - Basic Components of the Data Center.
  - Evolution Trends of the Cloud Data Center.

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

Huawei Training & Certification
Huawei Training & Certification
Huawei Training & Certification

Thank You

www.huawei.com

Huawei Training & Certification
Huawei Training & Certification

13 Data Center Storage Management

www.huawei.com

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

Huawei Training & Certification
Huawei Training & Certification

Foreword

- This module mainly introduces:
  - Initial Configuration Process And Configuration Methods Of Storage Systems In the Data Center.
  - Basic Block Storage Services Application Scenario and Its Configuration Methods In The Data Center.
  - File Services Application Scenario and Its Configuration Methods In The Data Center.

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

Huawei Training & Certification
Huawei Training & Certification

Objectives

- Upon completion of this module, you will be able to:
  - Describe the initial operations and configuration of the storage system in the data center.
  - Describe the configuration process of the block storage services, its application scenario including the disk domain, storage pool allocation configuration and how the host and application servers accesses the storage volumes.
  - Describe the configuration process and application scenario of file services, including the application scenario and configuration methods of CIFS and NFS.

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

Huawei Training & Certification
Huawei Training & Certification

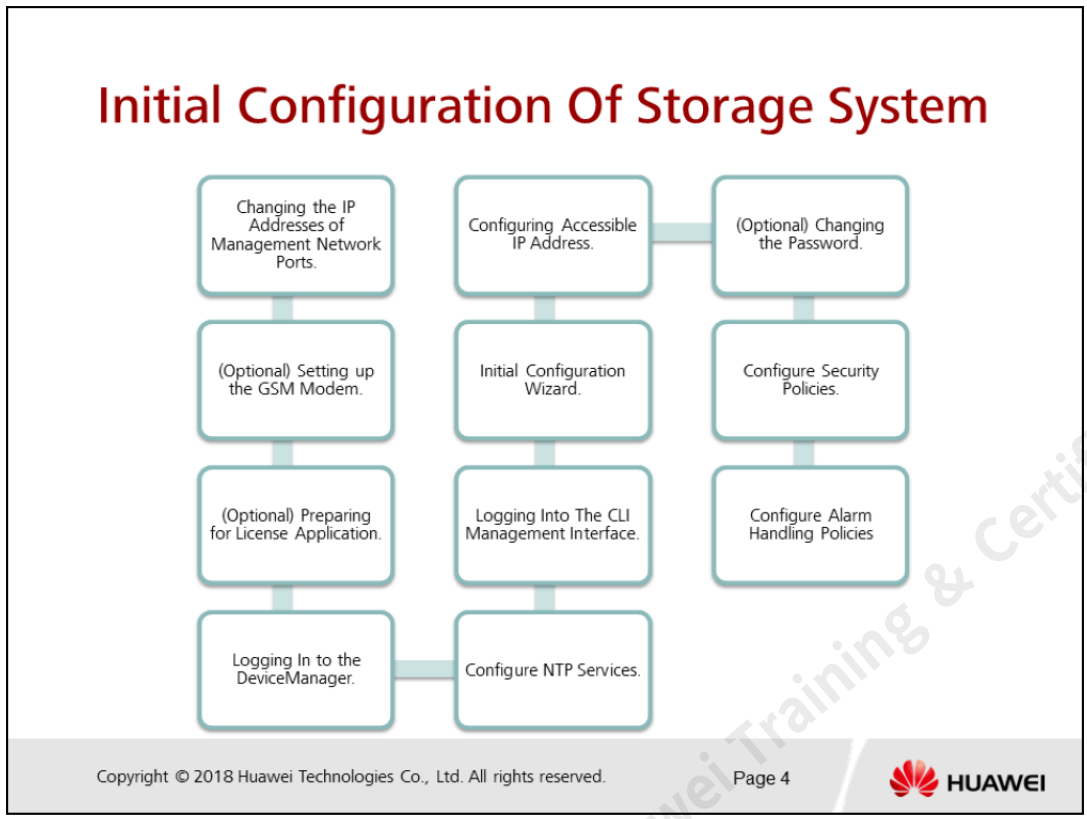
Contents

1. **Initial System Configuration.**
2. Configuration of SAN Storage Service.
3. Configuration of NAS Storage Service.

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

Huawei Training & Certification
Huawei Training & Certification



- After checking that the storage is correctly powered on, we need to initialize the storage system. The steps for the initial configurations are as follows:
- Changing the IP Addresses of Management Network Ports:
  - The storage system can be accessed through the IP addresses of the management network ports to configure and manage the storage system. During system initialization, you need to modify the IP addresses of the management network ports to connect those ports to user networks. After the management network ports communicate with the maintenance terminal properly, you can use the maintenance terminal to configure and manage the storage system.
- (Optional) Setting up a GSM Modem:
  - After setting the GSM modem, you can configure short message notification in management software. The storage system will send alarm information to a specified mobile phone. Thus, you can learn about storage system exceptions and solve them in a timely manner.

Huawei Training & Certification

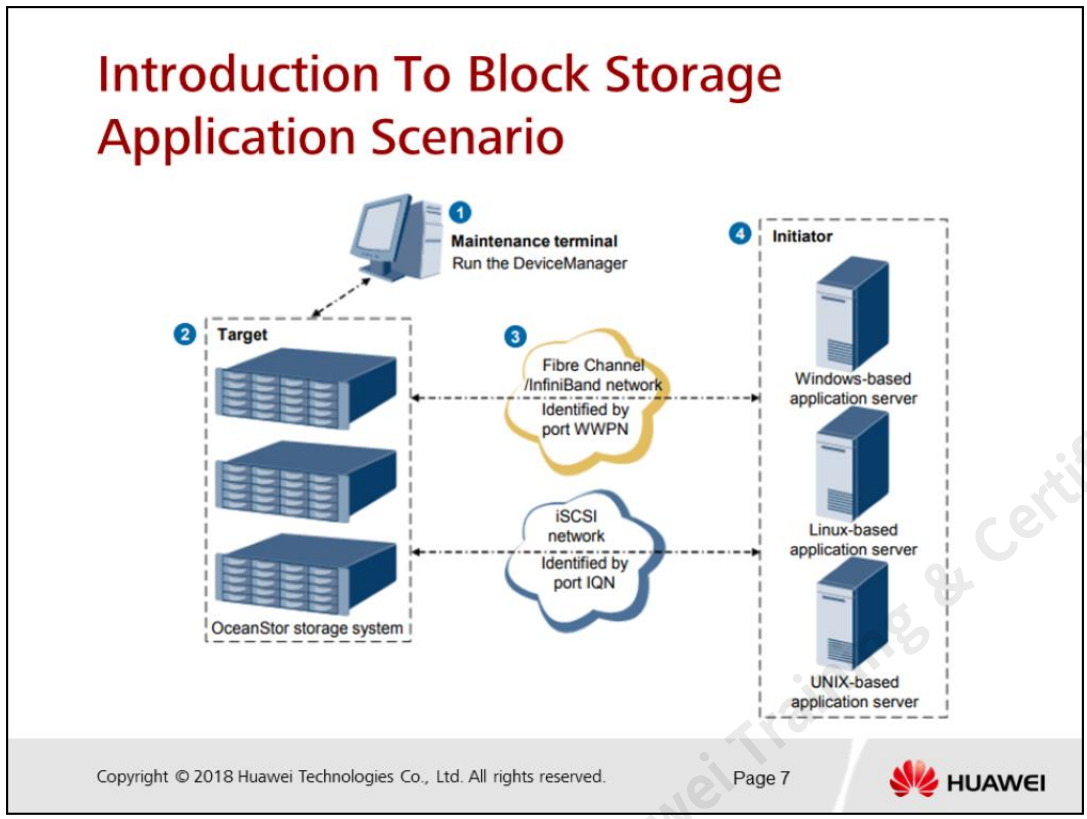
Contents

1. Initial System Configuration.
2. **Configuration of SAN Storage Service.**
3. Configuration of NAS Storage Service.

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

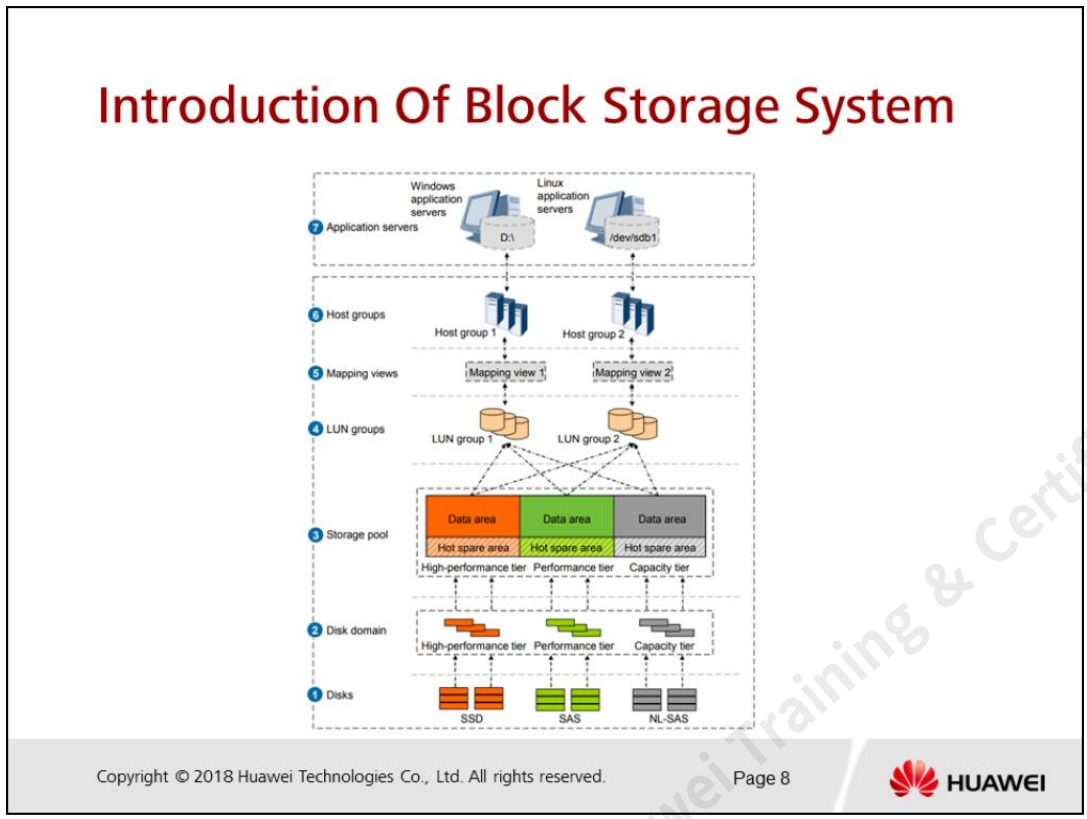
HUAWEI

Huawei Training & Certification
Huawei Training & Certification



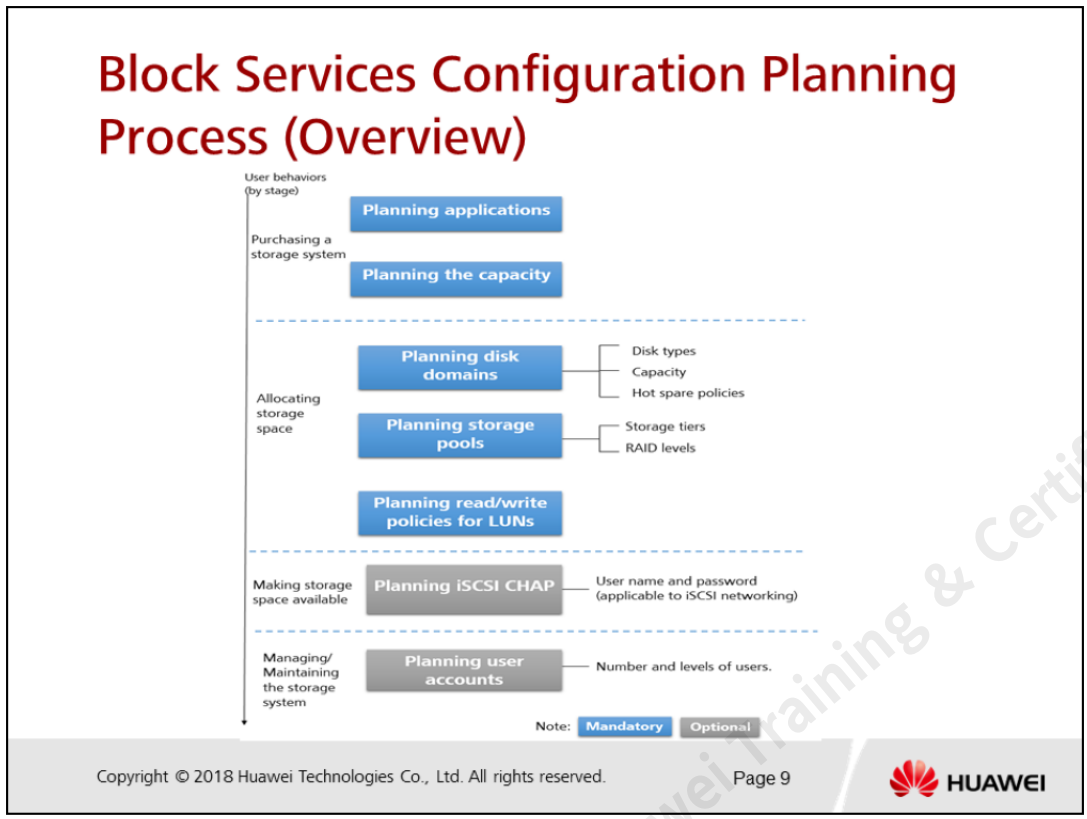
- Refer to 1: Users can manage and maintain the storage systems through the DeviceManager software from a maintenance terminal that is connected to the storage management network ports or SVP(Service Processor in 18000 series or 18000F series). DeviceManager is the integrated management software developed by Huawei.
- Refer to 2: Storage systems provide storage space for application servers.
- Refer to 3: The OceanStor enterprise storage systems can be connected to application servers running different operating systems including Windows, Linux, and UNIX over an Internet Small Computer Systems Interface (iSCSI), Fibre Channel, and InfiniBand network. According to data transmission protocols, an application server functions as the initiator for data transmission, and a storage system serves as the target for the information. The initiator sends data read and write requests to the target. The target receives, processes, and responds to the requests.
- Refer to 4: Application servers runs client programs. The storage system can connect to application servers running different operating systems including Windows, Linux (SUSE, and Red Hat), and UNIX (Solaris, AIX, and HP-UX).

Huawei Training & Certification



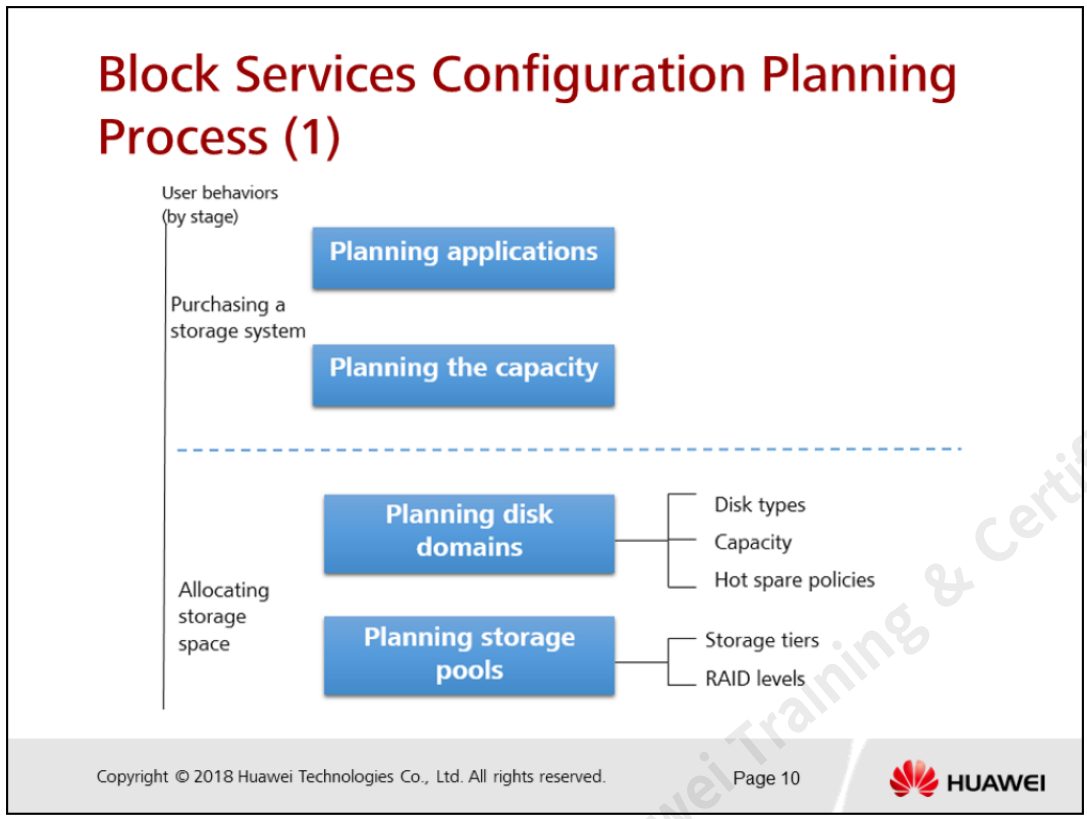
- Refer to 1: The storage system automatically identifies all disks.
- Refer to 2: Disk domains are comprised of different types of disks. Services of different disk domains are isolated from each other.
- Refer to 3: Storage pools are created in disk domains and comprised of RAID groups formed by disks of different performance. Storage pools provide logical storage space.
- Refer to 4: A LUN obtains storage space from the storage pool. LUNs are the minimum logical storage units that can be identified by application servers. A LUN group may contain one or multiple LUNs.
- Refer to 5: After mappings between host groups and LUN groups are established, related application servers can access the LUNs.
- Refer to 6: After initiators are added to hosts, one-to-one logical mappings between hosts and application servers are established. Then application servers can use storage space provided by the storage system. A host group may contain one or multiple hosts.
- Refer to 7: The application server identifies LUNs as logical disks. Then it can access the detected logical disks in the same way it would access local disks.

Huawei Training & Certification



- Before using the storage systems, please plan the logical sequence of the steps that need to be followed in order to efficiently configure the basic storage services.
- The planning process is shown as the diagram above, with the phases set according to user behaviors when they are using the storage system.
- Planning applications: The storage system provides application-based wizards to create storage resources. The applications include Microsoft Exchange, VMware, Hyper-V, Oracle, and SQL Server. Therefore, you can use the corresponding wizard to create storage resources for the preceding five applications. To create storage resources for other applications, follow the configuration procedure for basic storage services as shown at the diagram above.
- Planning the capacity: Storage capacity is the user's most basic requirement and must be planned first.
- Planning disk domains: A disk domain provides storage space for storage pools, whose storage tiers and available capacities depend on the disk types, capacity, and hot spare policy of the disk domain. Therefore, the disk types, capacity, and hot spare policy must be properly planned for a disk domain.
    - Disk type: Disk types of a disk domain determine the storage tiers of storage pools. Plan disk types based on your requirements.

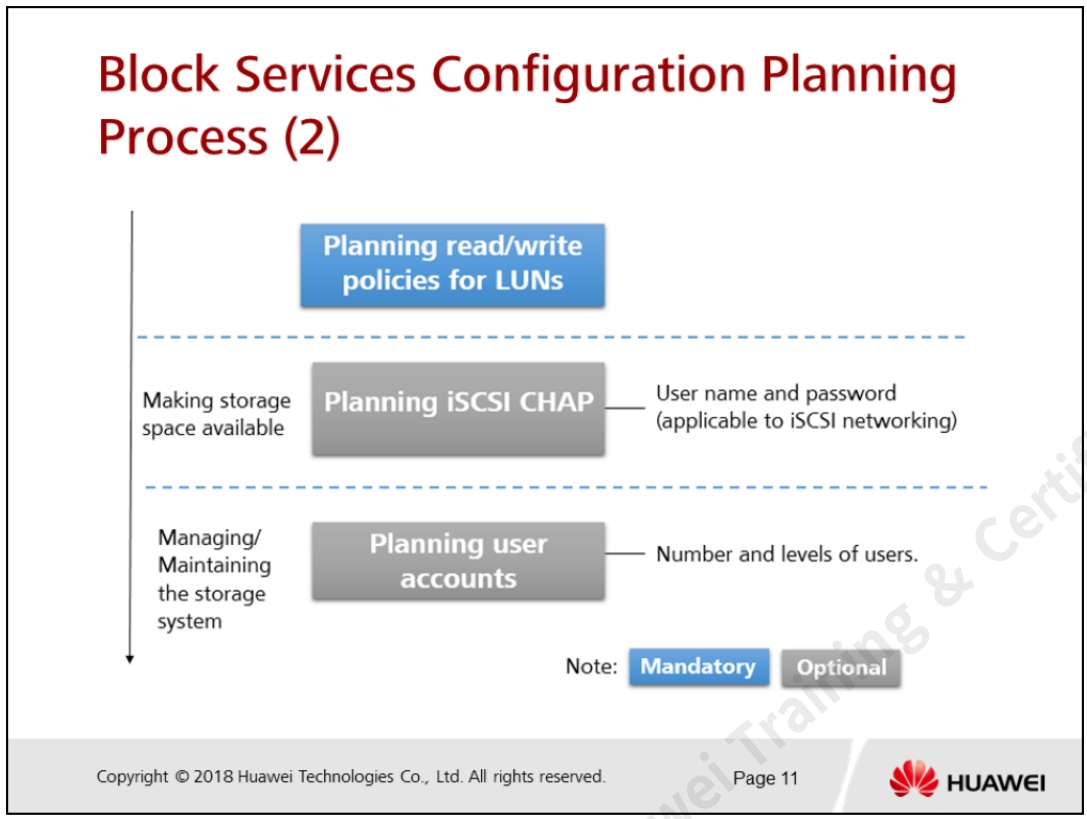
Huawei Training & Certification



- Capacity: The capacity of a disk domain determines the available capacities of storage pools.
- Hot spare policy: Plan hot spare policies and hot spare space so that the hot spare space can take over data from failed member disks.

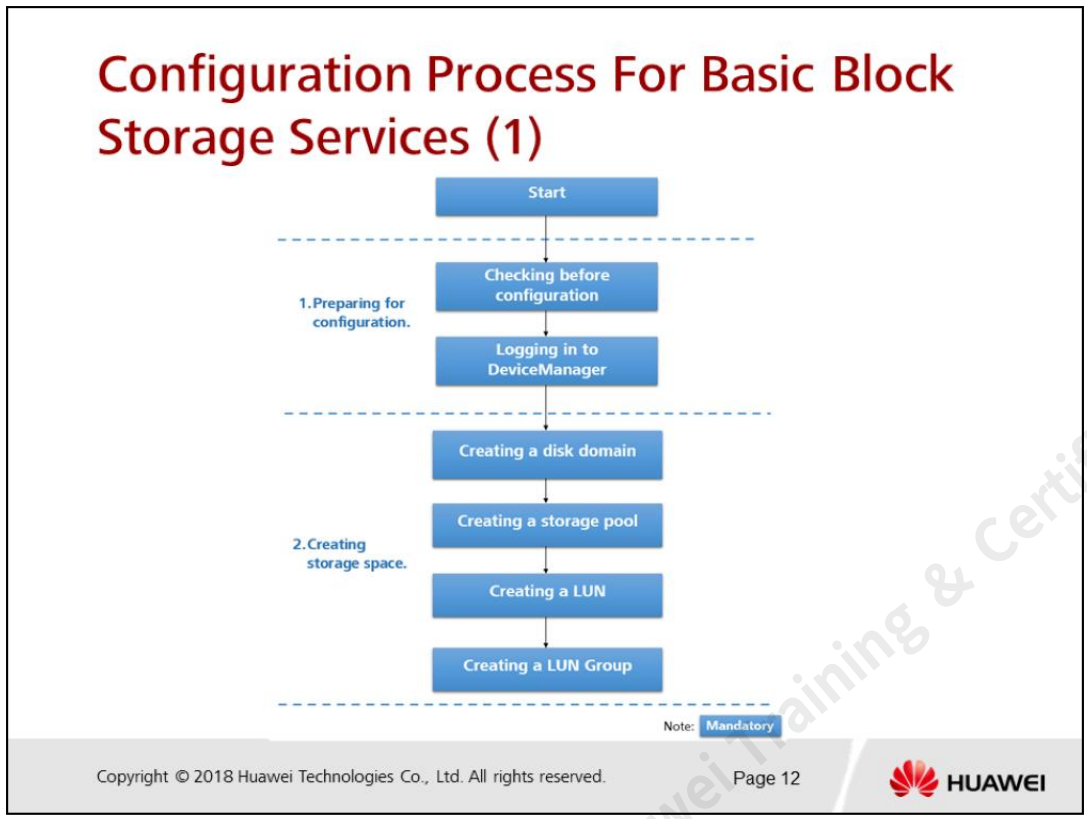
- Planning storage pools: A storage system provides storage space for application servers in the form of storage pools.
  - The storage tiers and RAID levels of the storage pools must be properly planned in advance for better storage utilization.
  - Storage tier: Plan storage tiers to meet the need for optimal distribution of hot data and cold data.
  - RAID level: Plan RAID levels based on actual needs because RAID levels vary in performance, cost, and reliability.
- Planning read/write policies for LUNs: Different read/write policies configured for LUNs affect the response time of the storage system. Properly planned write policies and prefetching policies can help improve the response speed of the storage system.

Huawei Training & Certification



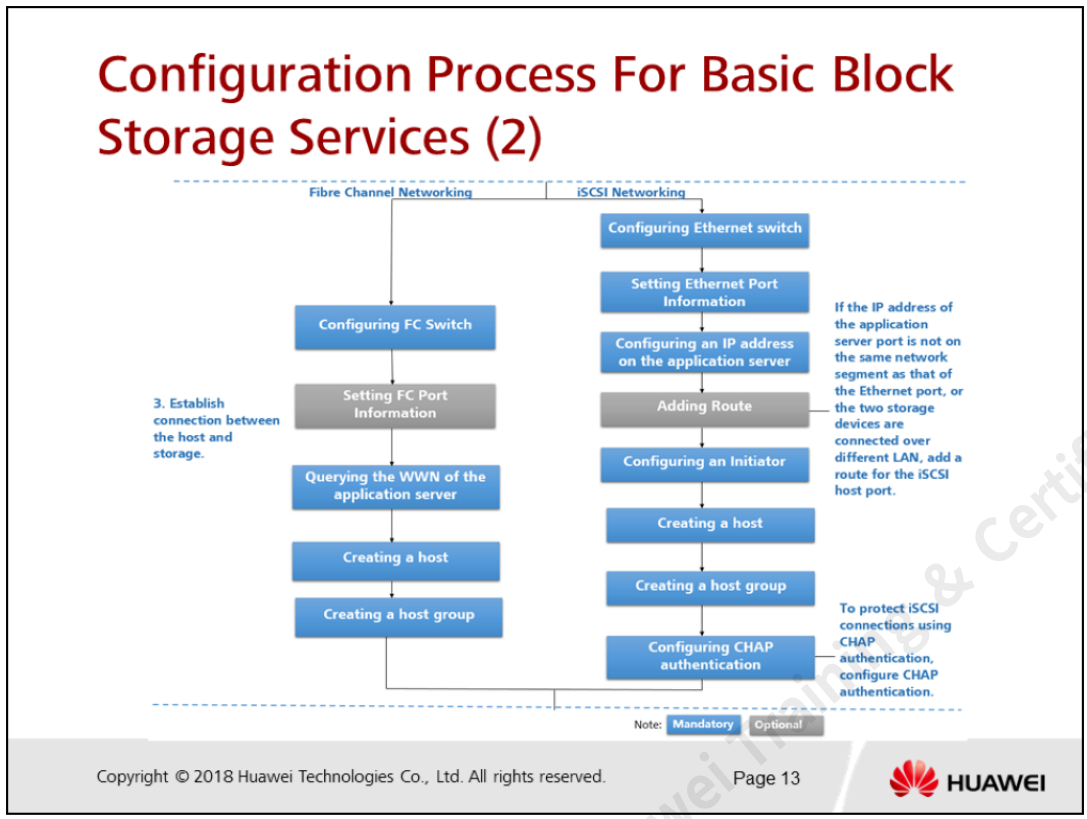
- (Optional) Planning iSCSI challenge handshake authentication protocol (CHAP): Considering the storage system access security when application servers connect to the storage system through iSCSI and use the storage space provided by the storage system, you should plan the CHAP user name and password to control the access to the storage system.
- Planning user accounts: Plan the number of users and their privileges carefully for subsequent management and maintenance. Users with different responsibilities should have different permissions.

Huawei Training & Certification



- The configuration process shown above introduces the main configuration procedures for basic block storage services. You can understand the configuration logic of the storage space configuration procedure through the flowchart shown above.
- Preparing for configuration:
  - Checking before configuration: Before configuring the storage services, check whether the software installation and initial configuration meet the requirements for storage space configuration.
  - Logging in to the DeviceManager: The DeviceManager is a device management platform program developed by Huawei. You can log in to the DeviceManager to manage and maintain the storage system. Make sure that you have all the login credentials prepared for logging into DeviceManager.
- Creating storage space:
  - Creating a disk domain: A disk domain provides storage space for storage pools, whose storage tiers and available capacities depend on the disk types, capacity, and hot spare policy of the disk domain. The storage system automatically allocates hot spare space with different capacities based on hot spare policies to take over data from failed member disks.

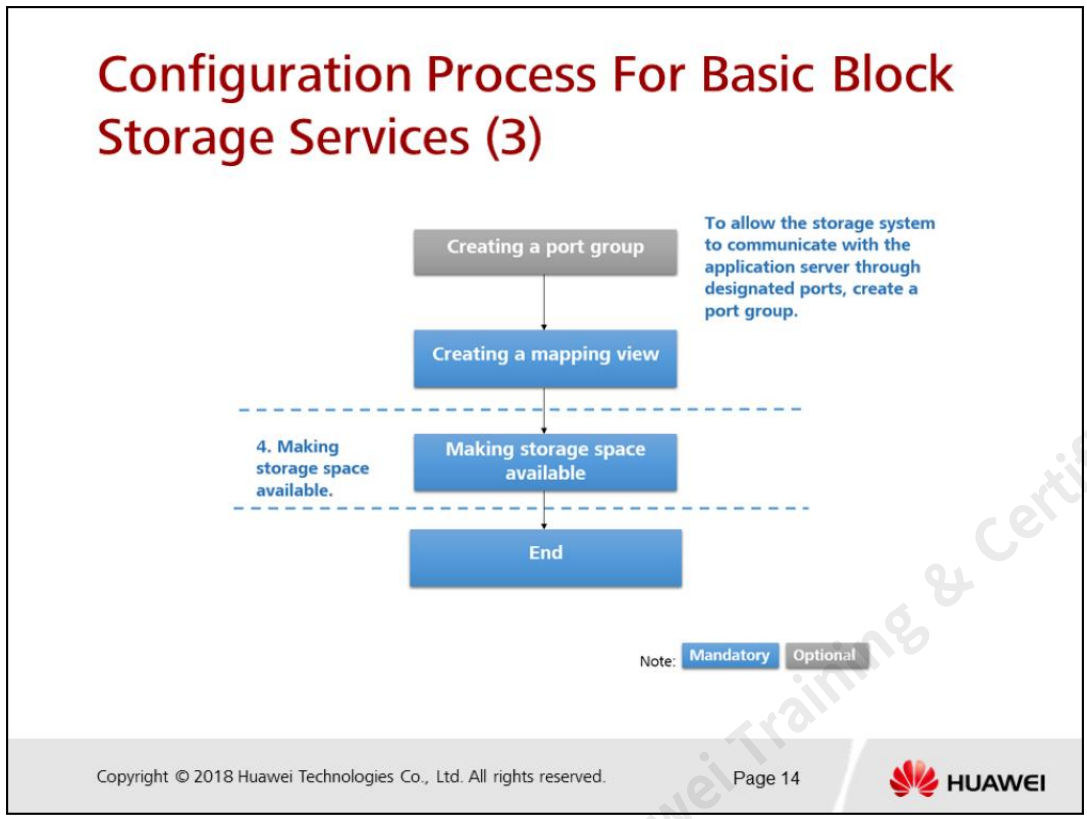
Huawei Training & Certification



- Creating a storage pool: The storage space used by application servers is provided by the storage pools on the storage system.
- Creating a LUN: A LUN indicates a logical unit in storage space.
- Creating a LUN group: A LUN group is a set of LUNs. Only LUN groups can be added to a mapping view. Single LUN cannot be added to the mapping view. Mapping view is the logical relationship and connection between the storage space allocated (such as LUNs in block storage) and the host that accesses the storage.

- Configuring Host Connectivity and Setting Up A Connection:
  - Configuring Host Connectivity:
    - IP SAN: Need to configure Ethernet switch, set Ethernet port information, add route and configure the IP address of the application server. Need to configure an initiator on the application server.
    - FC SAN: Need to configure FC switch and set FC port information of the storage system. Need to query the WWN of the application server.
  - Configure the connectivity between host and storage system.

Huawei Training & Certification



- Creating a host: A host is a virtual concept of an application server. Each host has its own application server. The host can be a physical machine or a virtual host like a virtual machine. An initiator is added to a virtual host to establish a logical connection between the application server, the storage system and the virtual host.
- Creating a host group: A host group is a set of hosts. Only host groups can be added to a mapping view. Single host cannot be added to the mapping view. A mapping view allows LUN Group to connect to Host Group, in which all the LUNs in the LUN Group is accessible by all the hosts in the Host Group.
- (Optional) Configuring CHAP authentication (iSCSI connection): CHAP authentication is configured to restrict the access to the storage system from application servers to increase the security of the storage system.
- (Optional) Creating a port group: Multiple ports can be logically added to a port group to facilitate port management.

Huawei Training & Certification

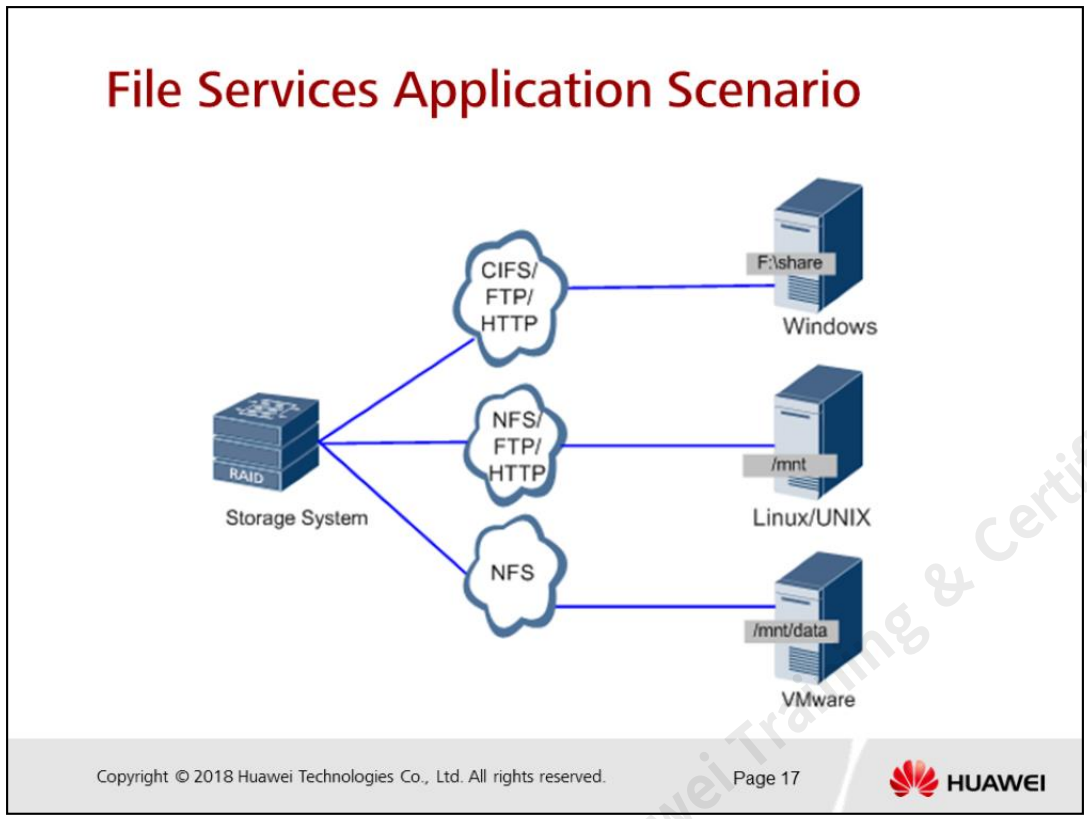
Contents

1. Initial System Configuration.
2. Configuration of SAN Storage Service.
3. **Configuration of NAS Storage Service.**

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

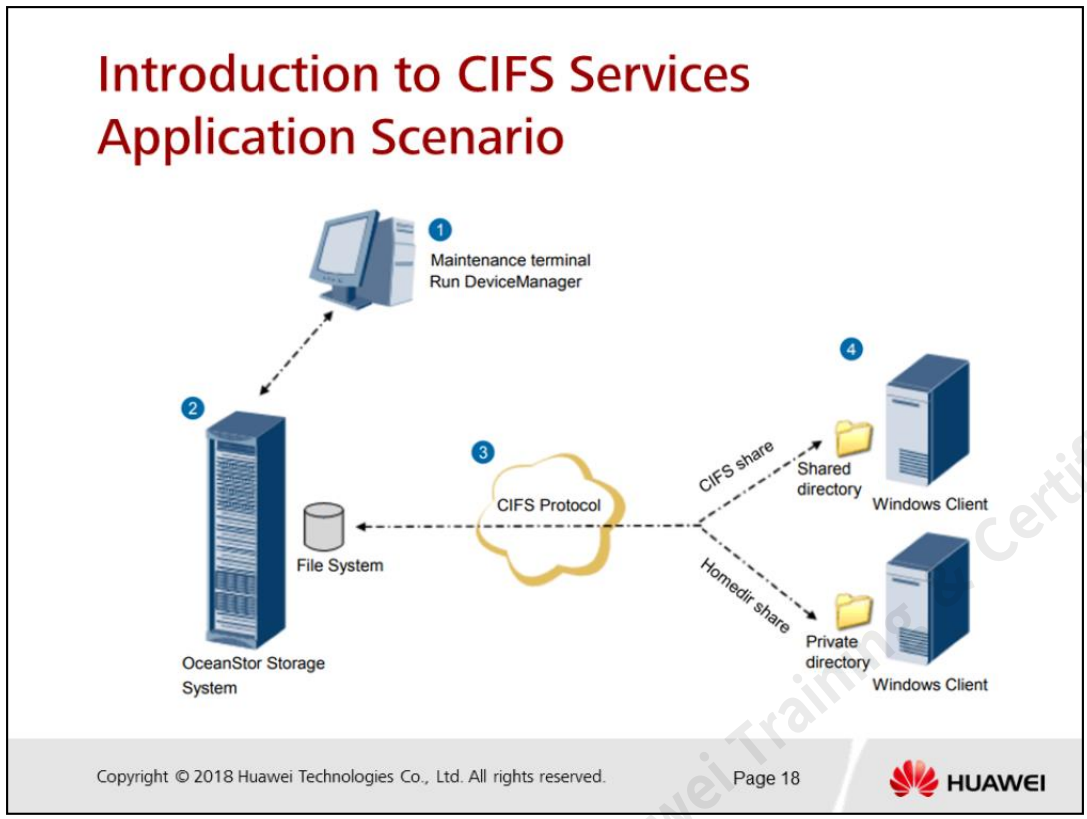
HUAWEI

Huawei Training & Certification
Huawei Training & Certification



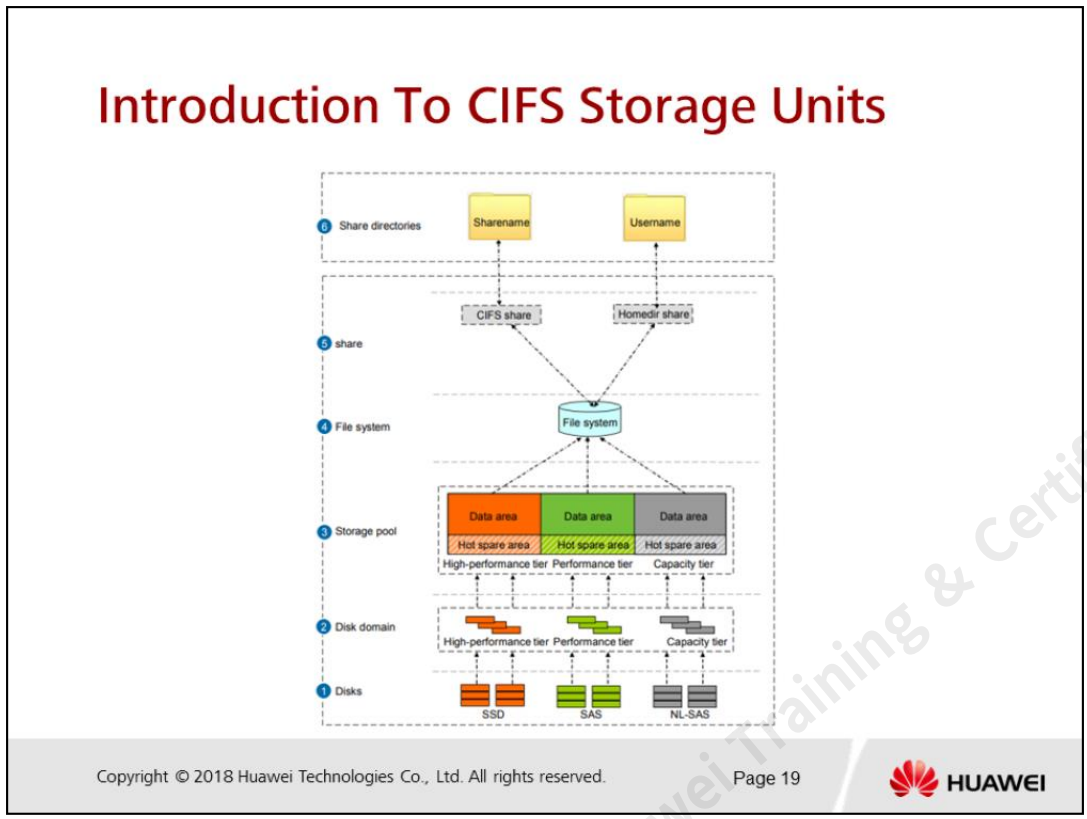
- Storage systems supports application servers to access shared files via NFS, CIFS, FTP or HTTP protocols. Application servers include Windows application servers, Linux application servers, UNIX application servers (such as AIX application servers, HP-UX application servers and Solaris application servers) and VMware application servers.
- The scenario where application accesses the storage system is shown on the diagram above.
- The way that each operating system(OS) accesses the storage space in the storage system using file storage services is different. In Windows, the OS access the storage space through shared directories with specified share names. For example, the shared directory such as F:\Share. In Unix, the OS can mount the shared directory to the /mnt directory. After mounting, the users can access the remote files as if they are accessing files on the local system. Similarly, virtual machines created under VMware can also mount shared file directories and use it to store and access files.

Huawei Training & Certification



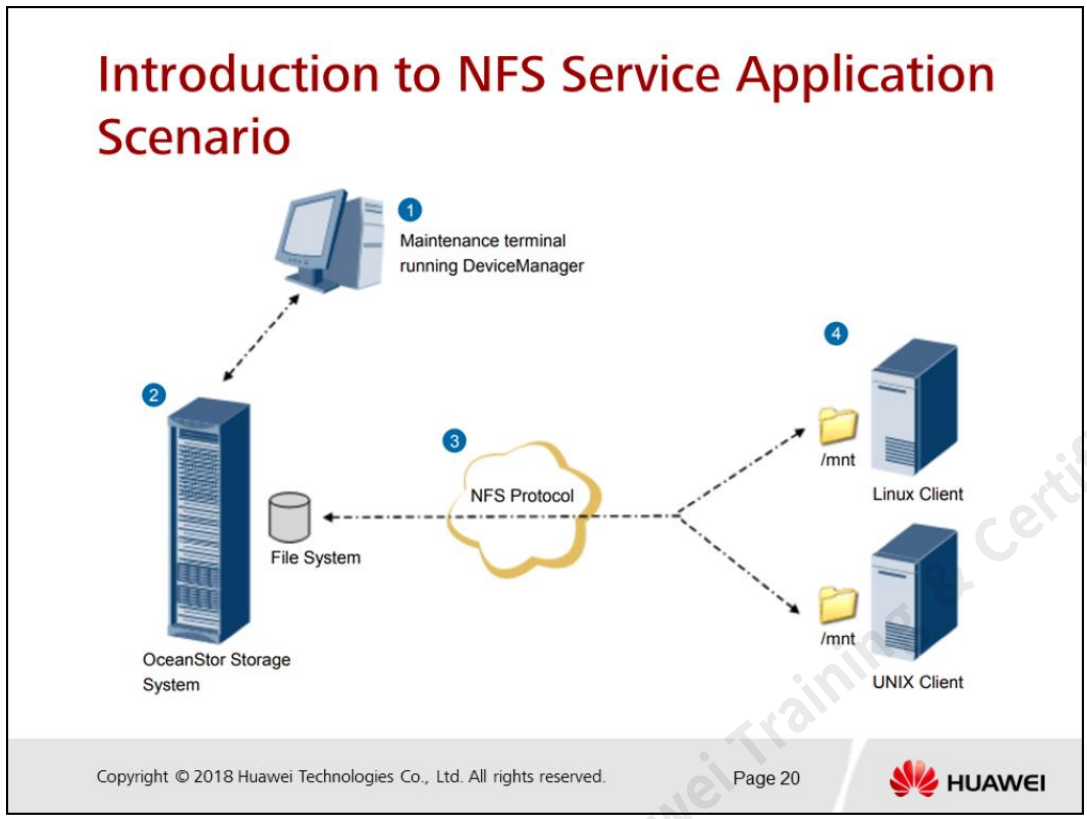
- Refer to 1: Users can manage and maintain the storage systems through the DeviceManager software from a maintenance terminal that is connected to the storage management network ports or SVP(Service Processor in 18000 series or 18000F series). DeviceManager is the integrated management software developed by Huawei.
- Refer to 2: File system of storage system provide file level data storage services featuring high performance and enhanced security.
- Refer to 3: CIFS is a file system sharing protocol developed by Microsoft and it is intended for Windows operating system. Through CIFS, files can be transferred and shared between Windows clients.
- Refer to 4: Generally, CIFS shares provided by storage systems are public directories and can be accessed by all authorized users. If you want to set a private directory, enable Homedir so that each user can only access the directory that share the same user name with the user.

Huawei Training & Certification



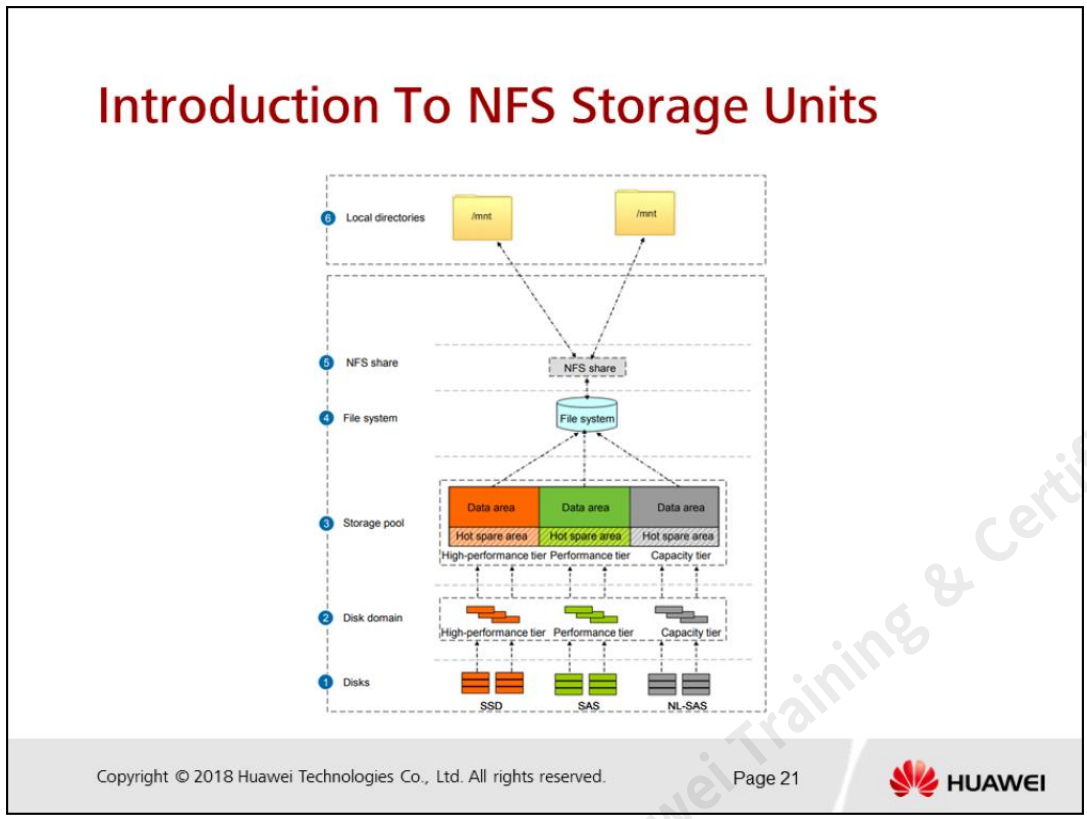
- Refer to 1: The storage system automatically identifies all disks.
- Refer to 2: Disk domains are comprised of different types of disks. Services of different disk domains are isolated from each other.
- Refer to 3: Storage pools are created in disk domains and comprised of RAID groups formed by disks of different performance. Storage pools provide logical storage space.
- Refer to 4: A file system can be created in a storage pool to provide storage space for a CIFS share.
- Refer to 5: After creating a CIFS share, you can set different access permissions for users.
- Refer to 6: A shared file system appears as a directory on a Windows-based client when the client accesses the CIFS share provided by a storage system. A storage system shares its file system with a user in CIFS or Homedir mode under the Windows operating platform. If the share mode is CIFS, then the shared file system appears as a directory with any specified sharename. On contrast, if the share mode is Homedir, then the directory name is the same as the user name. The Homedir directory can be accessed only by this user.

Huawei Training & Certification



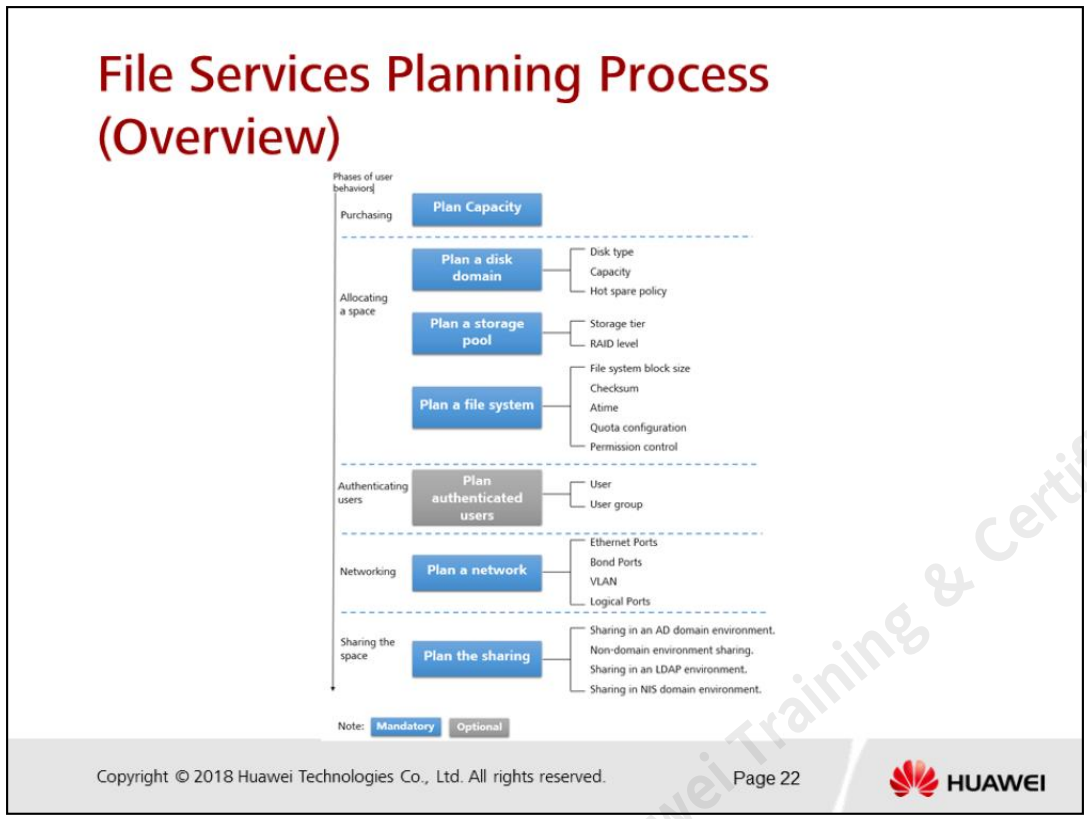
- Refer to 1: Users can manage and maintain the storage systems through the DeviceManager software from a maintenance terminal that is connected to the storage management network ports or SVP(Service Processor in 18000 series or 18000F series). DeviceManager is the integrated management software developed by Huawei.
- Refer to 2: File system of storage system provide file level data storage services featuring high performance and enhanced security.
- Refer to 3: NFS is a file system sharing protocol developed by Sun Microsystems and it is intended for operating systems such as Linux and UNIX. Through NFS, files can be transferred and shared between clients that are running Linux (including SUSE and Red Hat) or UNIX (including Solaris, AIX, and HP-UX).
- Refer to 4: By using NFS, a storage system functions as a server and allows clients to access a shared file system. The clients mount the directory where the shared file system resides to their local directories, and then users can access the files on the remote server like they are accessing local files.

Huawei Training & Certification



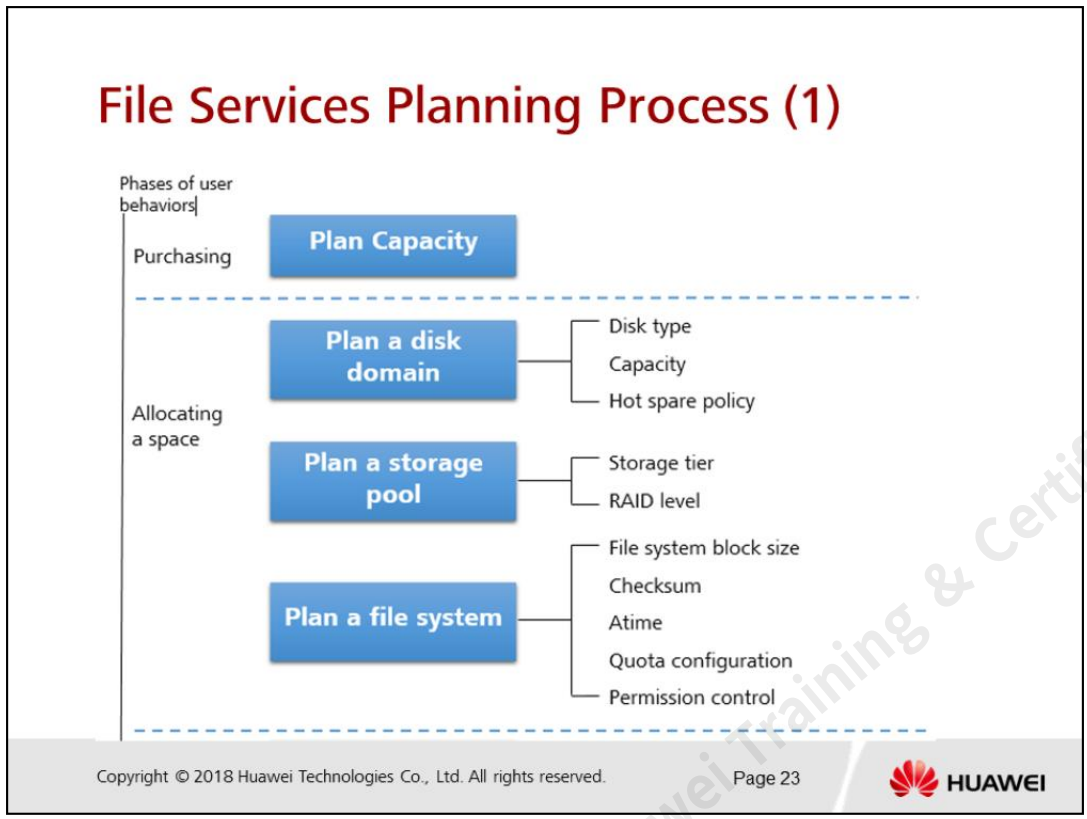
- Refer to 1: The storage system automatically identifies all disks.
- Refer to 2: Disk domains are comprised of different types of disks. Services of different disk domains are isolated from each other.
- Refer to 3: Storage pools are created in disk domains and comprised of RAID groups formed by disks of different performance. Storage pools provide logical storage space.
- Refer to 4: A file system can be created in a storage pool to provide storage space for an NFS share.
- Refer to 5: After creating an NFS share, you can set different access permissions for clients.
- Refer to 6: A Linux or UNIX-based client can access an NFS share provided by a storage system. A shared file system appears as a directory. The NFS share is mounted to the mnt directory on the Linux or Unix-based clients.

Huawei Training & Certification



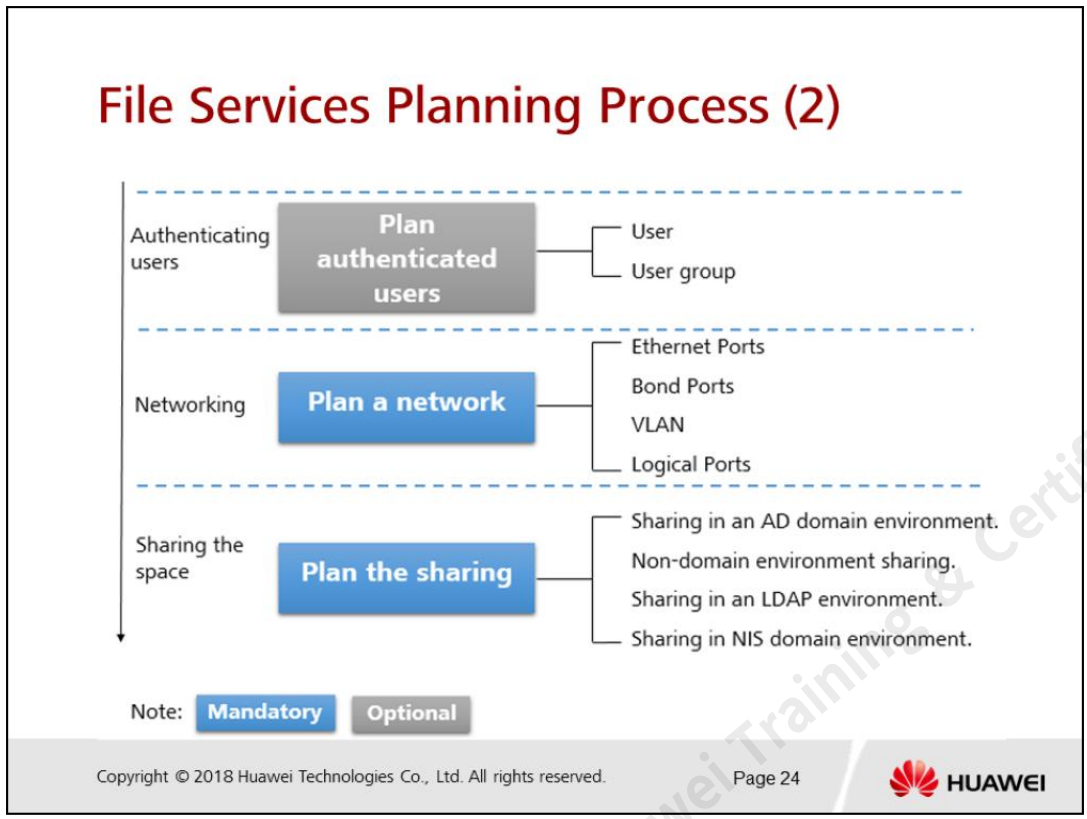
- Planning the capacity:
  - The storage system capacity is used for service data and system data storage. To ensure sufficient capacity for service data, the storage system capacity must be properly planned.
- Planning disk domains: A disk domain provides storage space for storage pools, whose storage tiers and available capacities depend on the disk types, capacity, and hot spare policy of the disk domain. Therefore, the disk types, capacity, and hot spare policy must be properly planned for a disk domain.
  - Disk type: Disk types of a disk domain determine the storage tiers of storage pools. Plan disk types based on your requirements.
  - Capacity: The capacity of a disk domain determines the available capacities of storage pools.
  - Hot spare policy: Plan hot spare policies and hot spare space so that the hot spare space can take over data from failed member disks.

Huawei Training & Certification



- Planning storage pools: A storage system provides storage space for application servers in the form of storage pools. The storage tiers and RAID levels of the storage pools must be properly planned in advance for better storage utilization.
    - Storage tier: Plan storage tiers to meet the need for optimal distribution of hot data and cold data.
    - RAID level: Plan RAID levels based on actual needs because RAID levels vary in performance, cost, and reliability.
- Planning file systems: File systems must be properly planned to optimize storage system space utilization and effectively accelerate storage system responses.
    - File system block size: The sizes of file system blocks are determined by data blocks written, so that the storage space can be fully utilized.
    - Checksum: Verifies data integrity. After the Checksum function is enabled, Checksum of data is automatically calculated when the data is being written. When the data is being read, Checksum ensures the integrity of the data. You can enable this function based on site requirements.
    - Atime: Updates file system time. Because this function affects system performance, determine whether to enable this function based on site requirements.

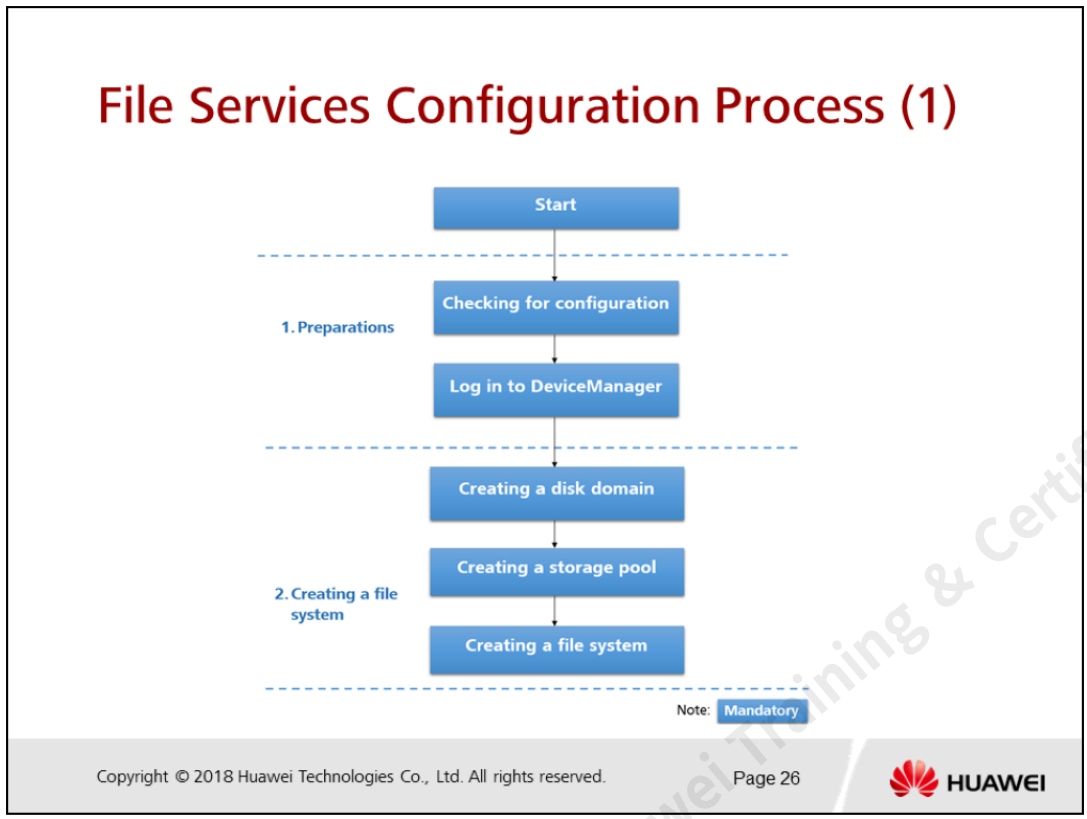
Huawei Training & Certification



- Quota configuration: Different space sizes and quantities can be configured for different directories to achieve efficient storage resource utilization.
- Permission control: Different permissions of the same directory can be configured for different users, so that the users can only access the directory within their specified permissions.

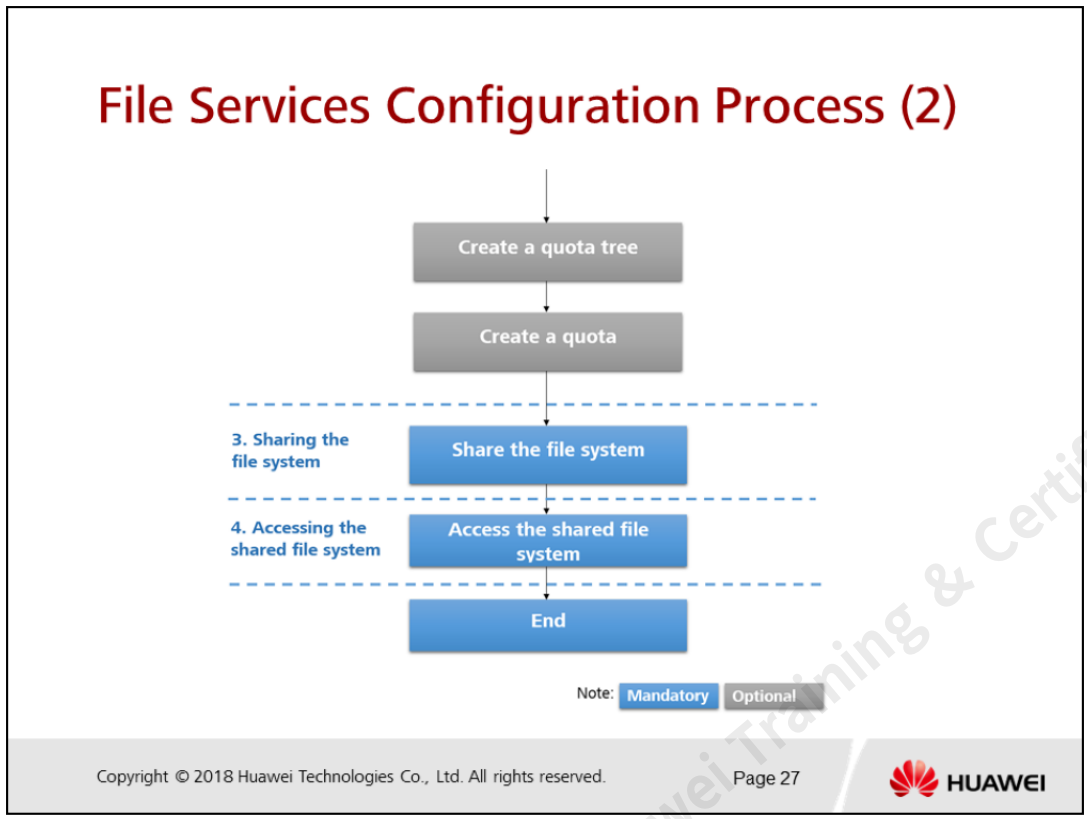
- (Optional) Planning authenticated users: When CIFS shares are being matched, different authenticated users or user groups need to be selected for easy user permission control and management.
  - User: Different users need to be specified for shared directories.
  - User group: Each user can only belong to a user group to facilitate user group permission control.

Huawei Training & Certification



- Preparing for configuration:
  - Logging in to the DeviceManager: Users can manage and maintain the storage systems through the DeviceManager software from a maintenance terminal that is connected to the storage management network ports or SVP(Service Processor in 18000 series or 18000F series). DeviceManager is the integrated management software developed by Huawei.
- Creating the file system:
  - Creating a disk domain: A disk domain provides storage space for storage pools, whose storage tiers and available capacities depend on the disk types, capacity, and hot spare policy of the disk domain. The storage system automatically allocates hot spare space with different capacities based on hot spare policies to take over data from failed member disks
  - Creating a storage pool: The storage space used by application servers is provided by the storage pools on the storage system.
  - Creating a file system: Application servers require a file system to access storage system space.

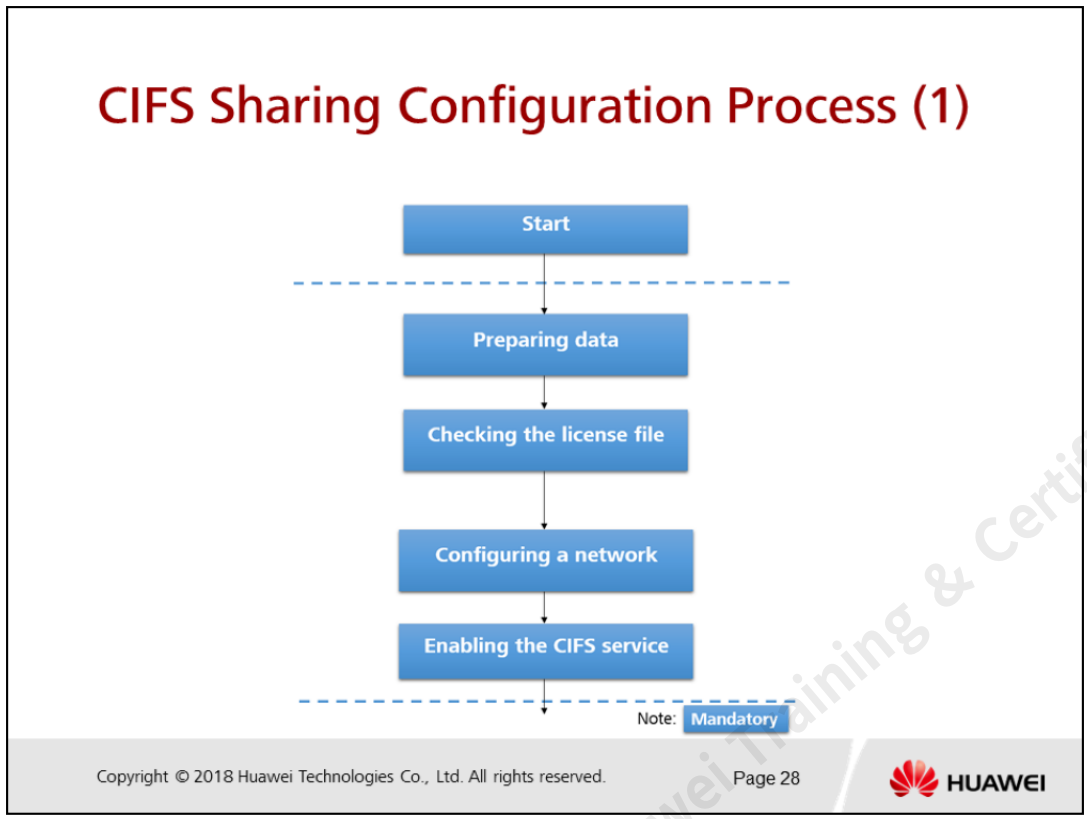
Huawei Training & Certification



- (Optional) Creating a quota tree: Quota tree is the first level directory of the file system that can be managed by quota. You can set the directory quota, user quota, and user group quota to manage the storage space used by all the files within those directories.
- (Optional) Creating a quota: By creating a quota, you can restrict the amount of storage space used or number of files that can be stored by users or group of users within the first level directory of the file system.

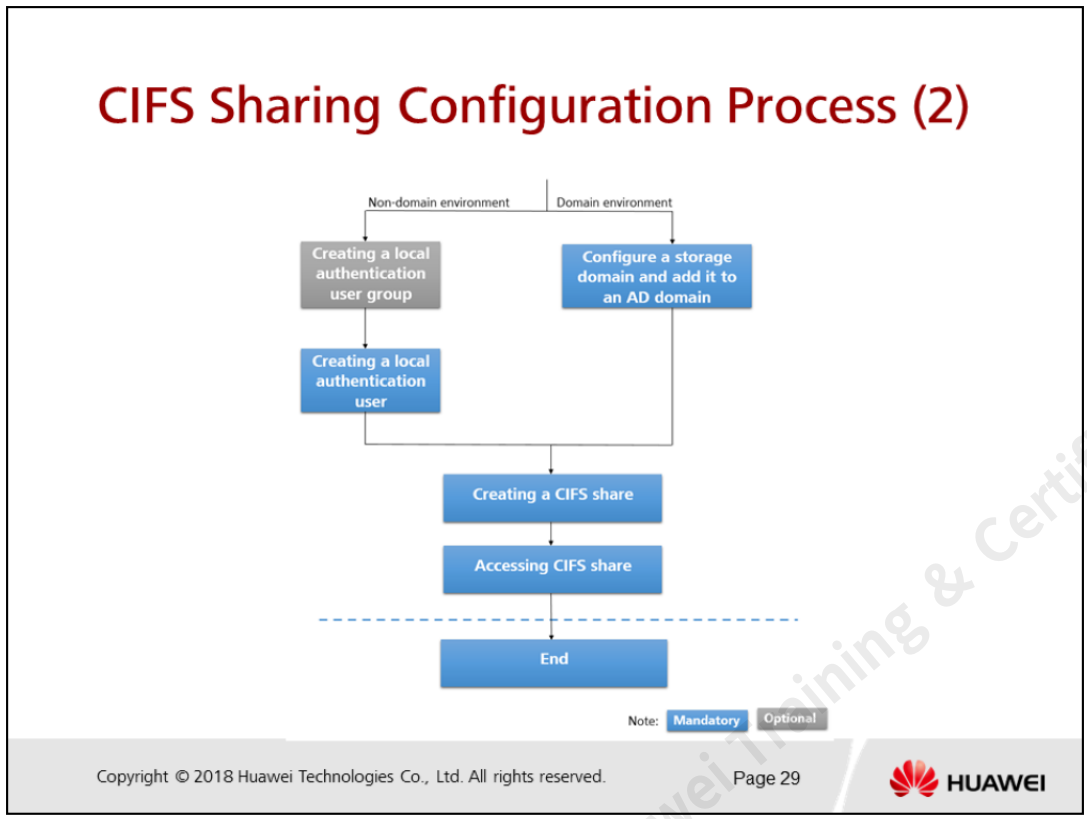
- Sharing storage space and accessing a shared file system:
  - Sharing a file system and using an application server to access a shared file system: A file system must be shared before application servers access storage system space.

Huawei Training & Certification



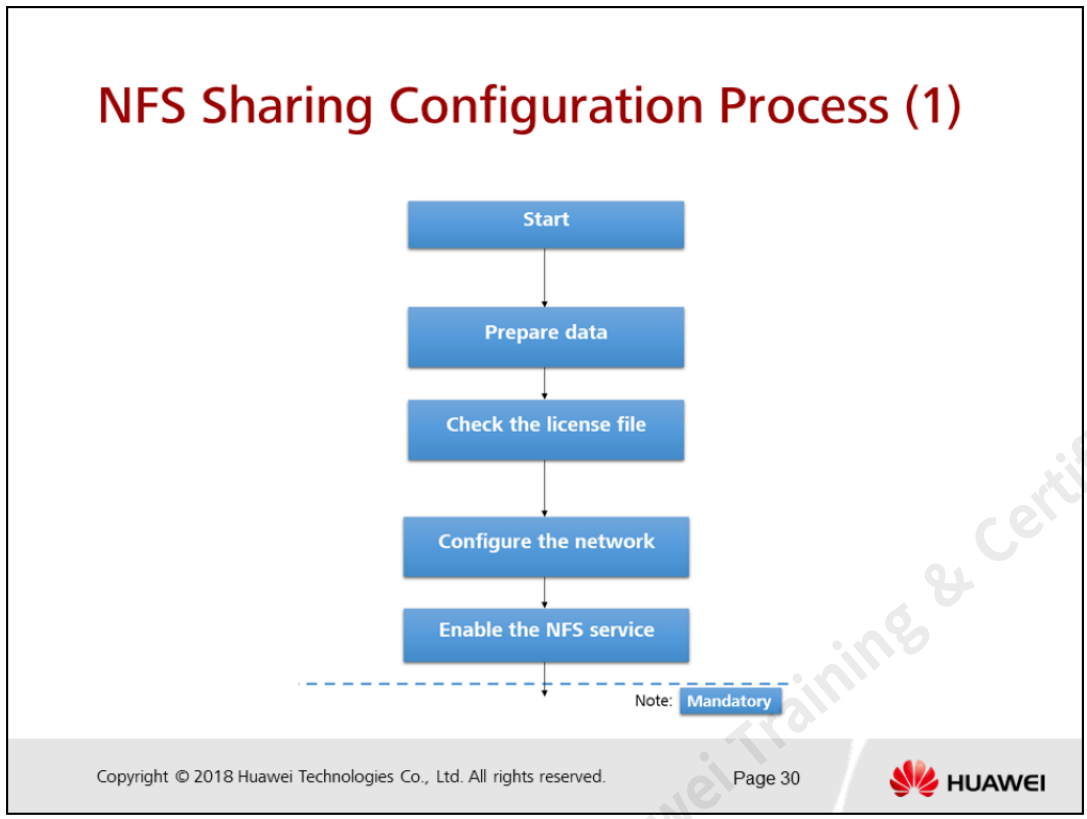
- Before configuring the CIFS share, please prepare the logical IP address for the storage system, file system for the CIFS share directories, CIFS sharename, user rights, local authentication users for non domain environment and Active Directory information for domain environment, for a quicker and smooth configuration process.
- After the file system is shared using the CIFS mode, users can access the storage space and shared directories.
- Preparing Data:
  - Before configuring a CIFS share, obtain information about storage system IP address, local users, quotas, permissions, and AD domain to assist in the follow-up configuration.
- Checking the License File:
  - Each value-added feature requires a license file for activation. Before configuring a value-added feature, ensure that its license file is valid for the feature.
- Configuring a Network:
  - Use the DeviceManager to configure the IP address of the storage ports, gateway, subnet mask, and floating IP address of the storage system.
- Enabling the CIFS Service:
  - Before creating a CIFS share, check whether the CIFS service has been enabled and whether parameters are correct.

Huawei Training & Certification



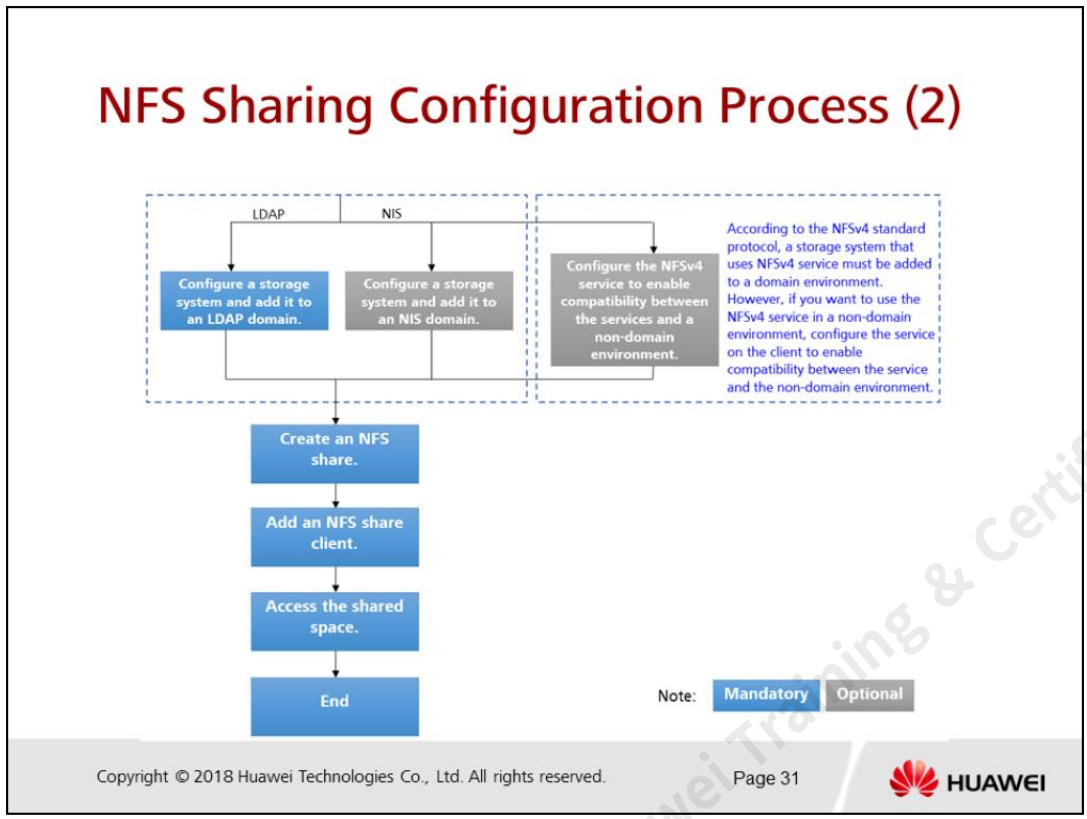
- Configuring a Local Authentication User Group:
  - In a non-domain environment, you must configure a local authentication user (or group). The storage system enables you to allocate different CIFS share access permissions to different user (or group).
- Configuring a Storage System to Add It to an AD Domain:
  - After the storage system is added to an AD domain, domain users can access CIFS shares that are allocated to the domain.
- Creating a CIFS share:
  - You may share the file system through CIFS, and afterwards users can access the shared storage space.
- Accessing CIFS Shares:
  - By accessing a CIFS share, different users can access the shared directory based on their user rights and permissions.

Huawei Training & Certification



- Preparing Data:
  - Before configuring an NFS share in a storage system, plan and collect required data to assist in the follow-up service configuration.
- Checking the License File:
  - Each value-added feature requires a license file for activation. Before configuring a value-added feature, ensure that its license file is valid for the feature.
- Configuring a Network:
  - Use the DeviceManager to configure the IP address of the storage ports, gateway, subnet mask, and floating IP address of the storage system.
- Enabling the NFS Service:
  - Before configuring an NFS share, enable the NFS service for clients to access the NFS share. The storage system supports NFSv3 and NFSv4.

Huawei Training & Certification



- (Optional) Configuring a Storage System to Add It to a Domain:
  - Add the storage system to a domain such as an LDAP or NIS domain. Afterwards, the domain users can access the NFS shares allocated to the domain. The LDAP (Lightweight Directory Access Protocol) is an open, vendor-neutral, industry standard application protocol for accessing and maintaining distributed directory information services over an Internet Protocol network. The NIS (Network Information Service), is a client–server directory service protocol for distributing system configuration data such as user and host names between computers on a computer network.
- (Optional) Configuring the NFSv4 Service to Enable It to Be Used in a Non-Domain Environment:
  - According to the NFSv4 standard protocol, the NFSv4 service must be used in a domain environment to ensure that the NFSv4 service functions properly. However, if you want to use the NFSv4 service in a non-domain environment, configure the user_name@domain_name mapping mechanism used by the NFSv4 service on your client. After the configuration is complete, the NFSv4 service will use UIDs and GIDs to transfer information about files during service transactions between your storage system and client.

Huawei Training & Certification

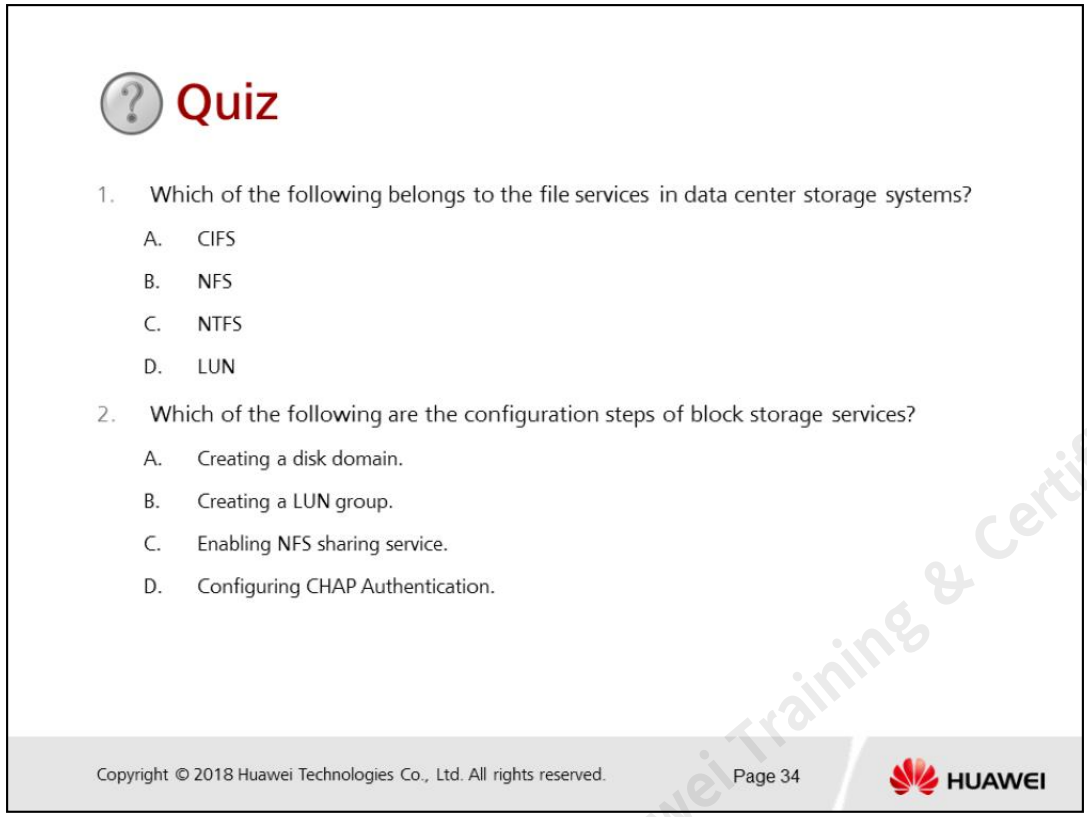
Summary

- This module introduced:
  - The Initialization Process and Configuration Methods Of Storage Systems In Data Centers.
  - The Configuration Process, Configuration Methods And Application Scenarios Of Block Storage Services In Data Centers.
  - The Configuration Process, Configuration Methods And Application Scenarios Of File Storage Services In Data Centers.

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

Huawei Training & Certification
Huawei Training & Certification



- Answers:
  - AB.
  - ABD.

Huawei Training & Certification
Huawei Training & Certification

Thank You

www.huawei.com

Huawei Training & Certification
Huawei Training & Certification

14 Operation & Maintenance of Datacenter Storage Systems

www.huawei.com

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

Huawei Training & Certification
Huawei Training & Certification

Foreword

- This module mainly introduces:
  - Daily operations of Datacenter Storage Administrators.
  - The routine maintenance items for Datacenter Storage Systems.
  - The daily management items for Datacenter Storage Systems.
  - The common Operation & Maintenance (O&M) tools used on Datacenter Storage Systems.

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

Huawei Training & Certification
Huawei Training & Certification

Objectives

- Upon completion of this module, you will be able to:
  - Describe the work scope of Datacenter Storage Administrators.
  - Describe the routine maintenance items and the maintenance methods for Datacenter Storage Systems.
  - Describe the content of daily management items for Datacenter Storage Systems.
  - Describe the functions and application scenario of the common O&M tools used on Datacenter Storage Systems.

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

Huawei Training & Certification
Huawei Training & Certification

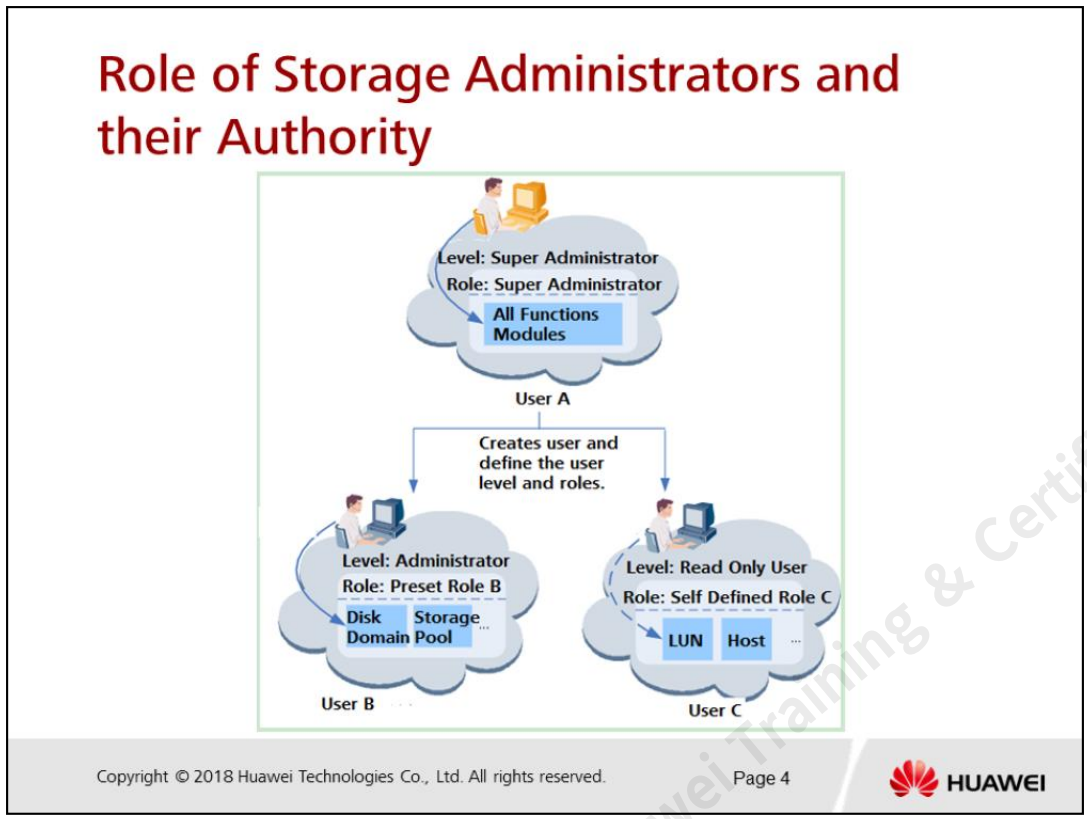
Contents

1. **Administrator Guide.**
2. Routine Maintenance.
3. Daily Management.
4. Common O&M Tools.

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

Huawei Training & Certification
Huawei Training & Certification



- In order to prevent wrong operations from affecting the stability of the business system and the security of business data, the system controls the user's operating authority and scope through levels and roles.
- Super administrator: has full control over the storage device and can create users at all levels.
- Administrator: Has a certain level of control over the storage device, but cannot perform operations such as user management, system upgrade, system modification, device power off, and device restart.
- Read-only user: Only has access to the storage device. After logging in to the storage device, read-only users can only perform query operations.

Huawei Training & Certification

Common Management Software

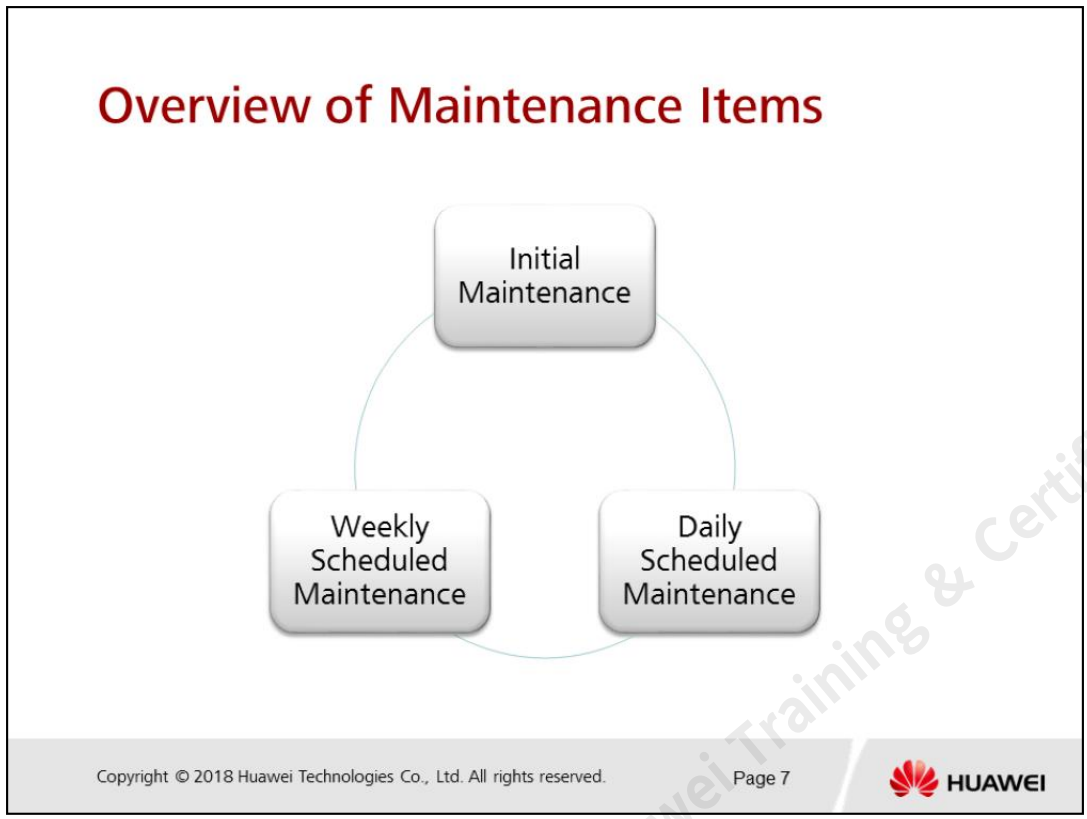
Management Software	Software Functionality:
DeviceManager	Storage resource allocation, user management, data protection feature management, device performance monitoring, and alarm management etc.
CLI Management Interface	Storage resource allocation, user management, data protection feature management, device performance monitoring, and alarm management etc.
Toolkit	Routine maintenance, upgrades/patches, troubleshooting, capacity expansion, and component replacement etc.
eService	Alarm reporting, file upload, and remote access etc.

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

- Users can query, set, manage, and maintain the storage system through the DeviceManager and CLI management interfaces. Serviceability management tools such as Toolkit and eService can effectively improve the operation and maintenance efficiency. The operation and maintenance management software can help administrators manage and monitor the storage system. The commonly used software for managing storage systems is shown in the table.
- DeviceManager is a Huawei storage device management software that can easily configure, manage, and maintain storage devices.
- The CLI (Command Line Interface) provides command line management and maintenance of the storage system. Configuration commands are input through the keyboard and are compiled and executed by the program. The execution result is displayed in the CLI operation interface as text or graphics.

Huawei Training & Certification



- The diagram above shows the three main elements in the maintenance process of storage systems in a datacenter. It consists of initial maintenance when all the systems are installed for the first time, then followed by daily and weekly scheduled maintenance to ensure optimal operation and to locate any faults in a timely manner.

Huawei Training & Certification

Initial Maintenance

Maintenance Items	Maintenance Operations
Checking the installation of Toolkit	Check whether the Toolkit and its sub-tools have been installed on the maintenance terminal. The sub-tool features include: • Device File Collection • Information Collection • Hard Disk Health Analysis • Inspection • Patch tool
Checking the installation and configuration of eService	On the maintenance terminal, check whether the eService tool is installed and complete the alarm policies configuration.
Checking the Alarm policy configuration	On the DeviceManager, check whether the related alarm policy has been configured. The alarm policy is configured for it to be used to report alarms to client servers or mobile phones in a timely manner so that they can be viewed and processed in a timely manner. It Includes: • Email notification • SMS notification • System notification • Alarm dump • Trap IP address management • USM User Management • Alarm masking • Syslog notification

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

- Through diligent maintenance and weekly reporting, the system administrator can ensure the optimal working environment and status of the equipment. If any abnormalities occur, then it can be processed and maintained in a timely manner, which ensures the continuous healthy operation of the storage system.
- Note: If you haven't installed the Toolkit or the eService tool, kindly login to Huawei Enterprise website (http://enterprise.huawei.com), to search and download the "OceanStor Toolkit" or "eService" software installers and the operational guide documents. You may install the tools based on the guide document provided.

Huawei Training & Certification

Daily and Weekly Scheduled Maintenance

Daily Maintenance Items	Maintenance Operations
Check and Process Alarms	Log in to the DeviceManager or view the alarms through the configured alarm reporting mode and handle the alarms in timely manner according to the repair suggestions. NOTE: If the alarm still exists, use the Toolkit to collect related information and contact Huawei technical support.

Weekly Maintenance Items	Maintenance Operations
Inspect Storage Devices	Use the Toolkit's "inspection" tool to perform inspections on the maintenance terminal. The inspection items include: • Hardware status • Software status • Value-added services • Alarm check
Check Server Room Environments	Check the room environment. If is doesn't meet requirements, adjust the room environment based on the relevant specification in a timely manner.
Check The Internal Environments of the Racks/Cabinets	Check the rack internal environment to see if it fits the relevant requirements. If it doesn't meet the requirements, adjust the internal rack environment accordingly based on relevant specifications in a timely manner.

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

- The table above shows the list of items in the maintenance plans for daily and weekly maintenance process along with the maintenance operations conducted for each maintenance item.
- Daily maintenance involves checking and processing alarms found via the DeviceManager or other notification channels.
- Weekly maintenance involves maintenance checks for storage devices and its environments to ensure that each devices are operating normally in the working environment specified. If the environment of the equipment doesn't meet the requirements, action needs to be taken immediately to rectify the issue before any damage occurs to the equipment.

Huawei Training & Certification

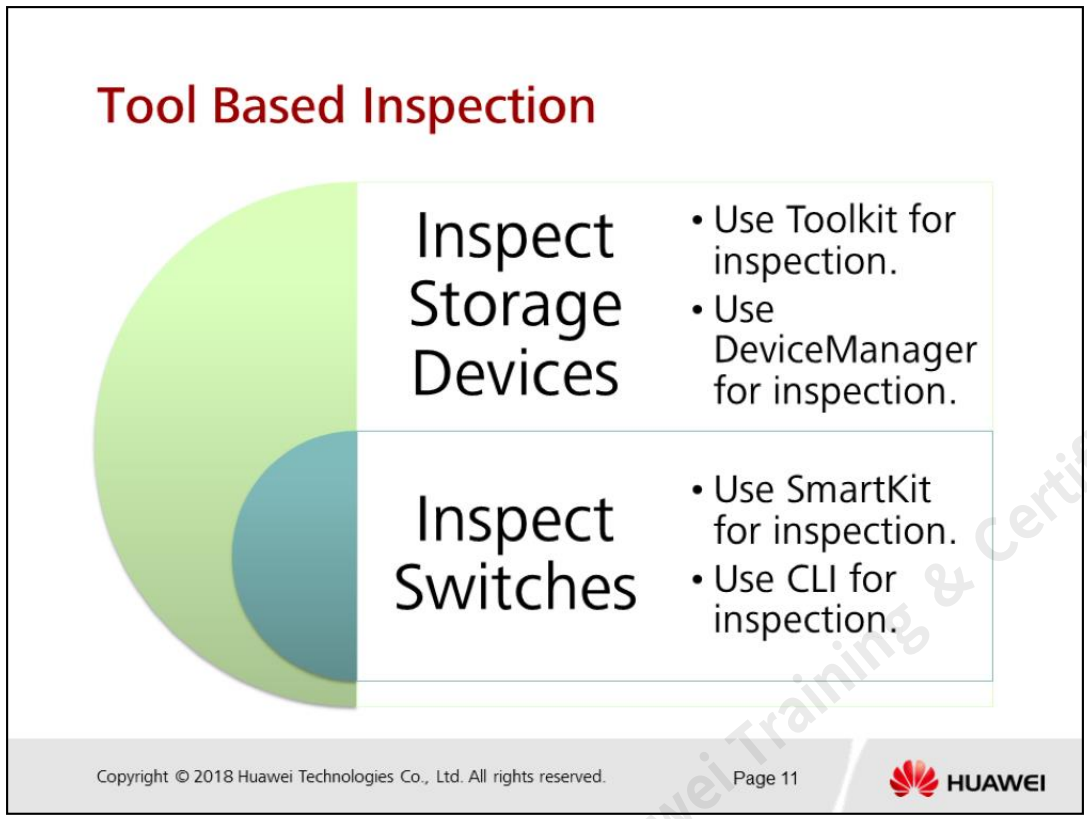
Contents

1. Administrator Guide.
2. **Routine Maintenance.**
3. Daily Management.
4. Common O&M Tools.

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

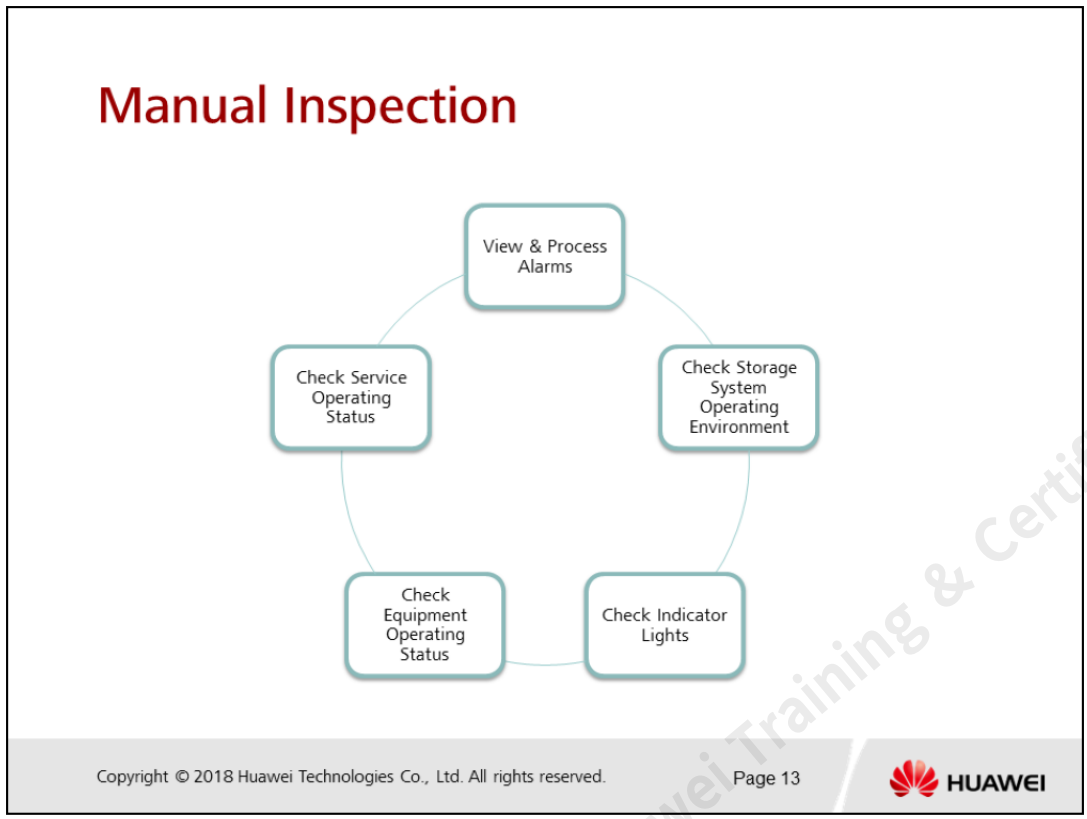
HUAWEI

Huawei Training & Certification
Huawei Training & Certification



- The O&M tool can inspect the current state of the equipment to know the running status of the equipment and the network in real time, understand the operational trends of the equipment and network, and improve the handling efficiency of maintenance personnel in the event of emergencies.
- Inspection of storage device: Inspection of storage device can be done via the Toolkit or the DeviceManager.
- Inspection of switches: You can use the SmartKit tool to inspect the switch or log in to the CLI through the serial port to check the health and running status of the switch.
- Using the Toolkit to perform inspection: When the OceanStor Toolkit inspection tool is installed on the storage system, you can use the inspection tool to formulate a corresponding inspection strategy and perform inspections on the equipment so that you can know the running status of the equipment in real time.

Huawei Training & Certification



Huawei Training & Certification

- View and process alarms: By checking detailed alarm information and repair suggestions in the alarm list, it can help you to repair and act upon the storage system alarms, and ensure normal operation of devices and services.
- Checking the operating environment of the storage system: The operating environment of the storage system that meets the business requirements is a necessary condition for the storage system to be able to operate in stable conditions for a long period time.
- Check indicator lights: The indicator lights can reflect the working status of the corresponding hardware device in real time. By checking the indicator lights, you can quickly and accurately determine whether each hardware device is in the normal working condition.
- Check the operating status of the device: By checking the operating status of the storage system on the storage system and DeviceManager, you can find faults in time and ensure the normal operation of the storage system.
- Check the service operating status: Check the service operating status on the DeviceManager management interface to detect faults in time and avoid data loss due to service interruption.

Collecting Storage System Information (1)

Types of Information	Names	Description:
Basic Information	Device Serial Number and Version	Provides the serial number and version info of the device.
	Customer Info	Provides actual contact and contact information.
Fault Information	Time of Fault	Records the accurate time when the faults occur.
	Fault Symptoms	Records in detail on the symptoms of the faults such as the content of the error dialog box, and the content of the notification received during fault.
	Operation Prior to Fault Occurrence	Recall and records the operation done on the system before the fault occurs.
	Operation After the Fault Occurred	Recall and records the operation done after the fault has occurred.
Storage Device Information	Configuration of Hardware Module	Records in detail the configuration information of the hardware module.
	Indicator Light Status	Record the indicator status of the storage device, especially the orange and red indicator light status information. For the indicator status of each component of the storage device, see "Product Description" in the documentation of the corresponding product model.
	Storage System Data	Manually export the operating data and system logs of the storage device.
	Alarms and Logs	Manually export the alarm and logs information of the storage device.

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

- The maintenance information that needs to be collected includes basic information, fault information, storage device information, networking, and application server information.
- DeviceManager can be used to collect system data, alarms, and event information.
- Use the CLI to back up or import device configuration data.
- Use Toolkit to collect device system data, file information, host information, and more.

Huawei Training & Certification

Collecting Storage System Information (2)

Types of Information	Names	Description:
Network Information	Connection Method	Describe the network connection between application servers and storage devices, such as FC networking and iSCSI networking.
	Switch Model	Records the model number of the existing switches in the network.
	Switch Diagnostics Information	Manually export diagnostic information of the currently running switch, including startup configuration, current configuration, interface information, time, and system version.
	Network Topology	Describes the network topology between application servers and storage devices or provide networking diagrams of the current environment.
	IP Address Information	If iSCSI networking is used between the application server and the storage device, it describes the principle of IP address allocation or provides an IP address allocation list of the current network.
Application Server Information	OS Version	Records the operating system type and its version number installed in the application server.
	Port Rate	Record the application server port rate connected to the storage device. For how to check the port rate, see the documentation for more information.
	OS Logs	View and exports the operating system logs information.

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

- The table above shows the continuation of the information collected from the storage system as part of the maintenance process.
- Collection of these information allows O&M personnel to have better understanding of the current running status of the storage devices and their configurations, which helps in locating faults or bottlenecks in the performance of the storage systems.

Huawei Training & Certification

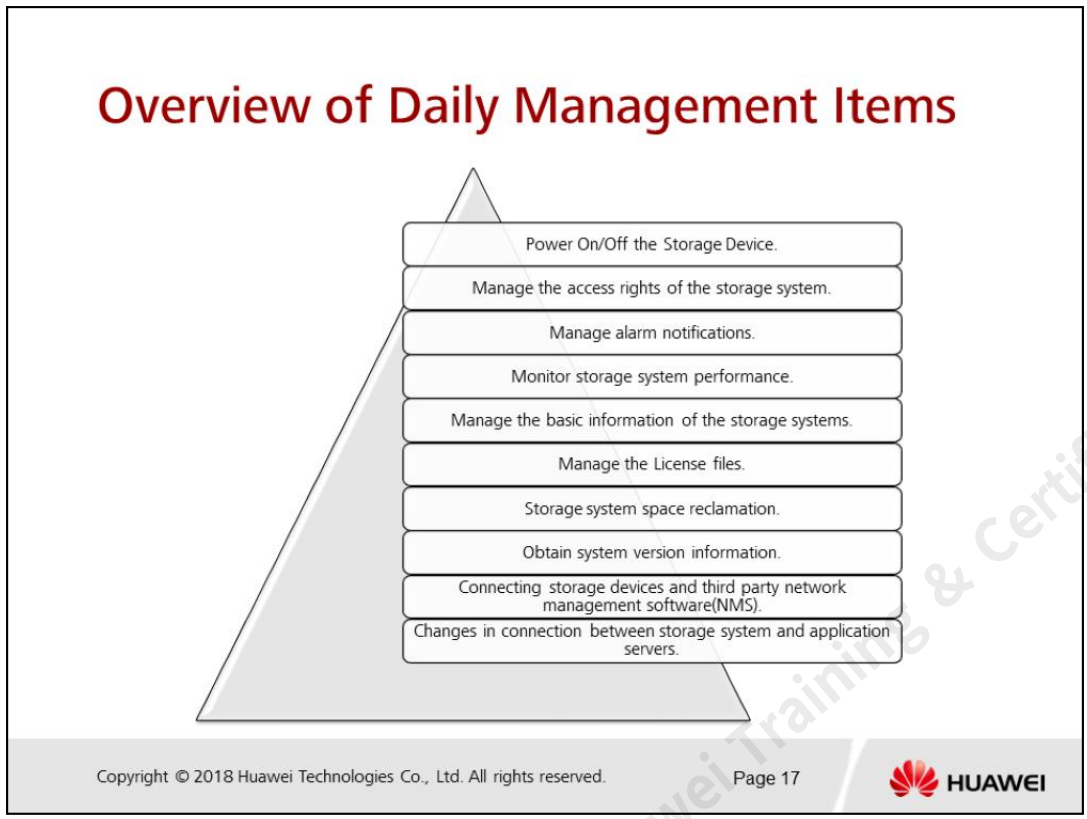
Contents

1. Administrator Guide.
2. Routine Maintenance.
3. **Daily Management.**
4. Common O&M Tools.

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

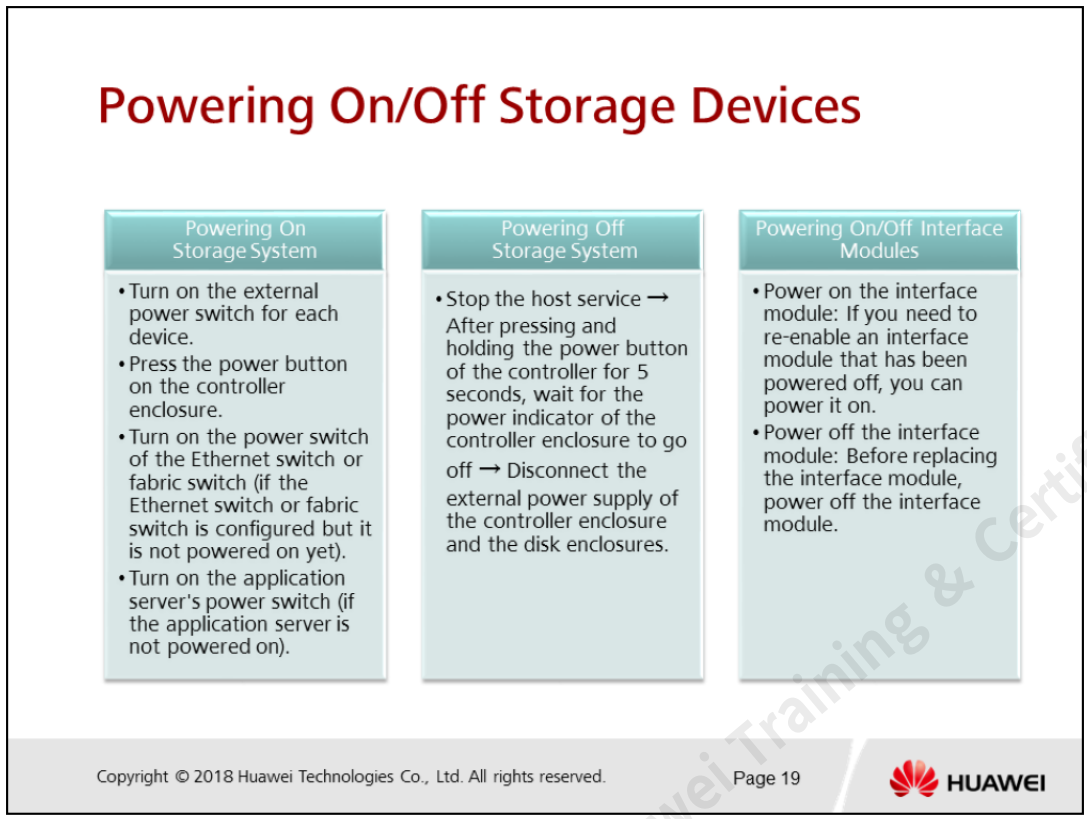
HUAWEI

Huawei Training & Certification
Huawei Training & Certification



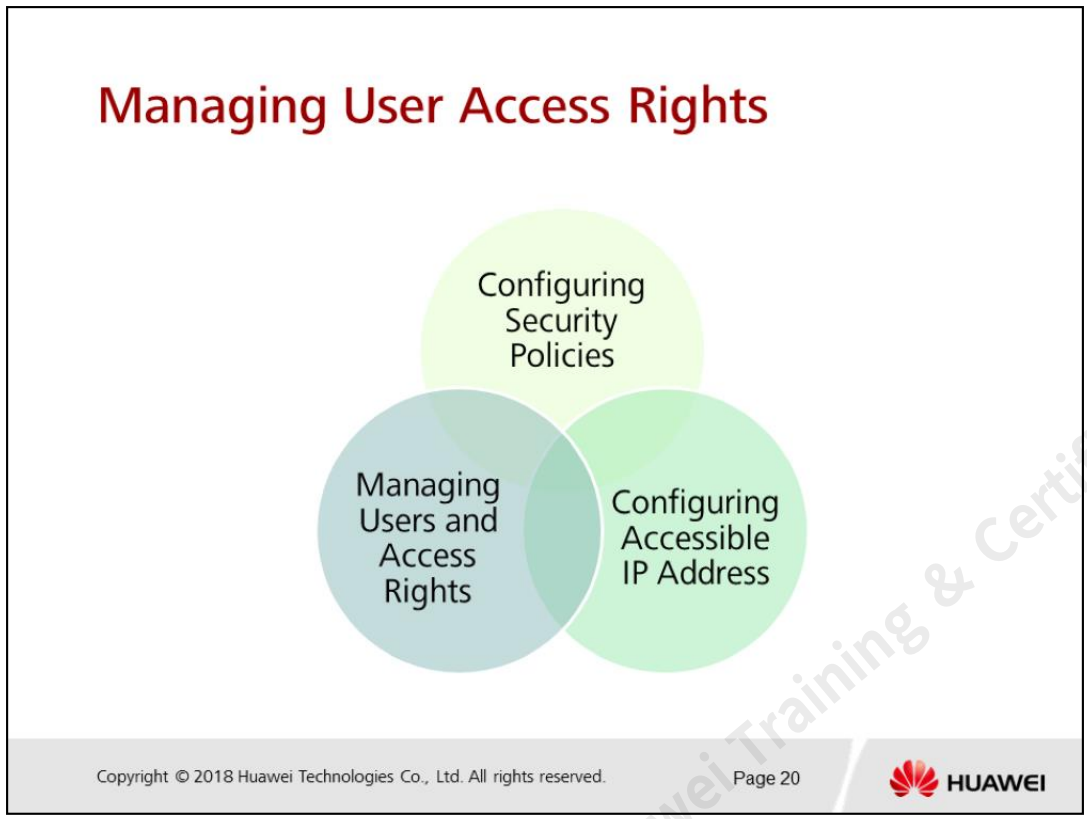
- Powering on and off the storage device: It includes powering on and off the storage system and powering on and off the interface module. Proper power-on and power-off prevents equipment damage due to improper operation.
- Managing access rights of the storage system: The storage system provides security policy adjustment such as access IP address management, and user management to ensure the security of storage devices and service data.
- Management of alarm and notifications: The storage system provides remote alarm notification. This chapter describes how to manage remote alarm notifications to ensure that administrators receive and process alarms in a timely manner.
- Monitor storage system performance: Monitor and view performance monitoring data to understand the system performance status and use related information for tuning the configurations to improve system performance.
- Manage basic information about the storage system: Basic information such as device name and system time can be modified according to service requirements.
- Management of License Files: A license file is a credential that allows the usage of various value-added features. During the daily management of the device, you need to pay attention to whether the license files is available and valid, to avoid the situation of unable to use certain value-added features because the license file is unavailable or invalid.

Huawei Training & Certification



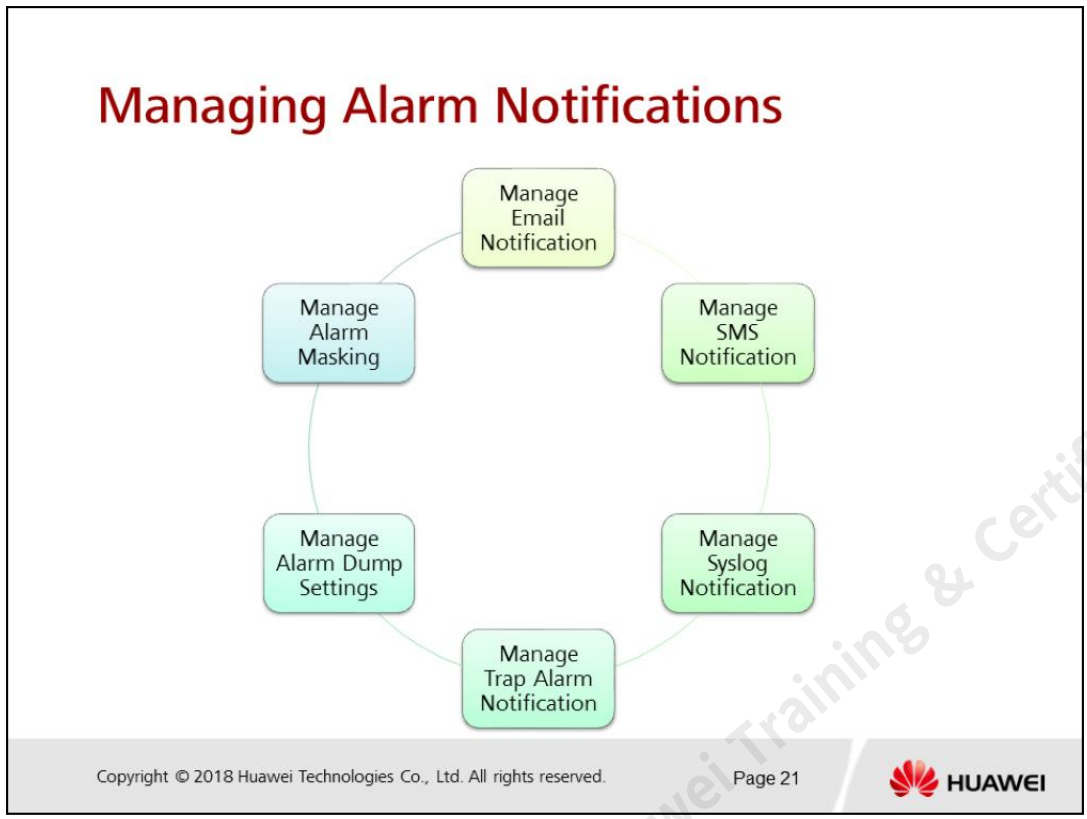
- Restarting the storage system causes service interruption on the storage system. Perform this operation with caution.
- When a fire, smoke, or flooding occurs in the equipment room, you need to do an emergency power-off to the storage system in order to ensure your personal safety or ensuring that the equipment is not damaged. For details on how to emergency power-off the storage system, please follow the user's server room safety guidelines for the power off process.
- The precondition for powering off the interface module is that all services involved in the interface module have already been stopped.

Huawei Training & Certification



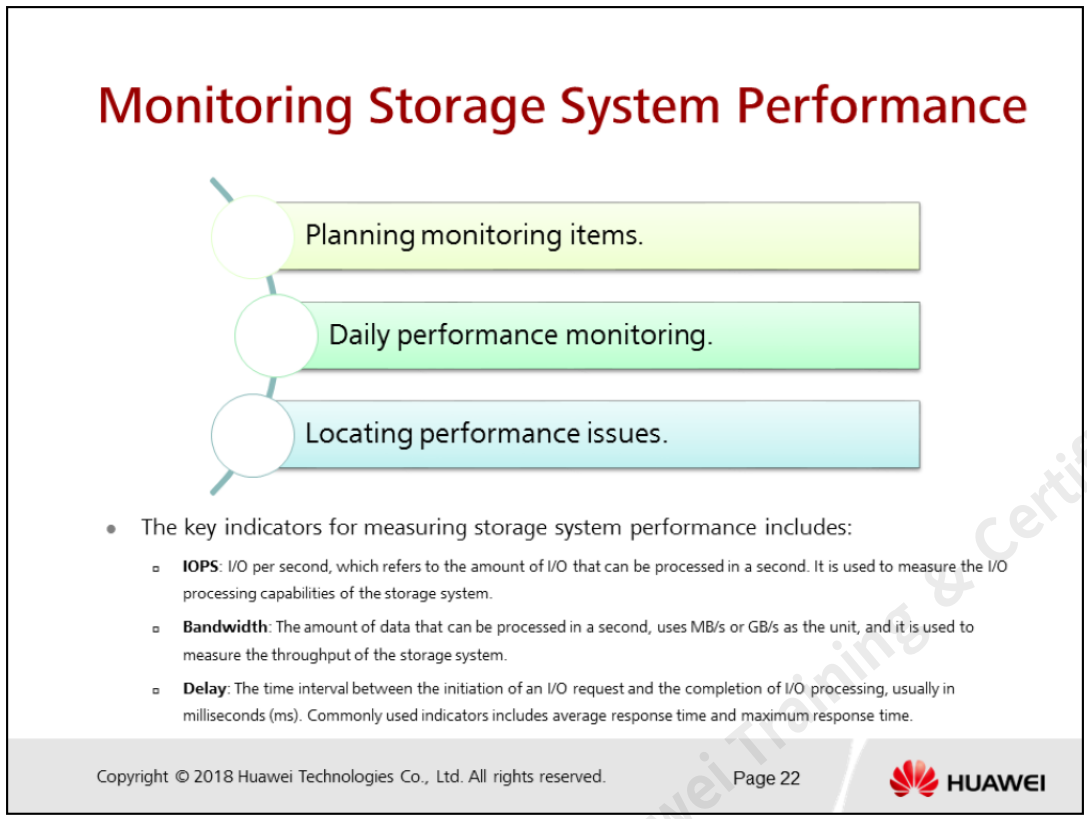
- The storage system provides security policies configuration such as access IP address management, and user management to ensure the security of storage devices and service data.
- Configure the security policy: You can control the complexity of the user name and password of the newly created account by setting the user name policy and password policy. Setting the login policy can lock the account that the system identifies as abnormal.
- Configure Accessible IP Addresses: After the IP address security limit is enabled, only the IP address or address segment in the list can access DeviceManager. Only super administrators can perform the operation of configuring accessible IP addresses.
- Managing users and access rights: In order to prevent incorrect operations from affecting the stability of business systems and the security of business data, the storage system defines three levels of users and restricts the operation rights of users at different levels.

Huawei Training & Certification



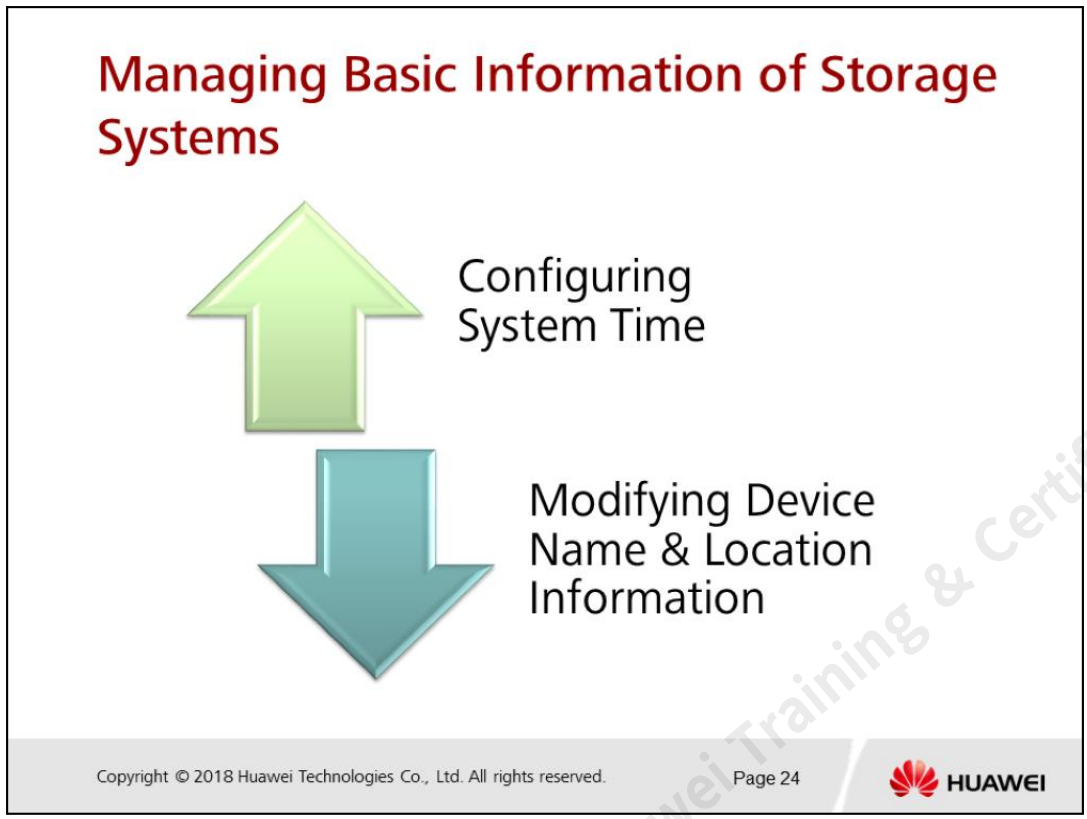
- Management of email notification: In the daily operation and maintenance process, changes in the business may require you to change the backup SMTP server and receive the alarm notification email address in time.
- Management of SMS notification: On the DeviceManager management interface, alarms can be set to be sent to the specified mobile phone as required by SMS to enable users to know and handle the storage system faults in a timely manner.
- Management of Syslog Notification: Users can adjust the server address for receiving syslog notifications and the level and type of notifications received according to the actual situation.
- Management of Trap alarm notification: Users can adjust the address of the Trap alarm notification according to the actual situation. The storage device alarm information will be sent to the network management system or storage device specified by the server.
- Management of alarm dump settings: After you set the alarm dump function, when the alarm information capacity reaches the system-defined threshold, it is automatically dumped to the specified FTP server or SFTP server.
- Management of alarm masking: After the alarm mask is set, you can enable/disable the alarm mask to manage the alarm information of the device without reporting or with reporting to the NMS.

Huawei Training & Certification



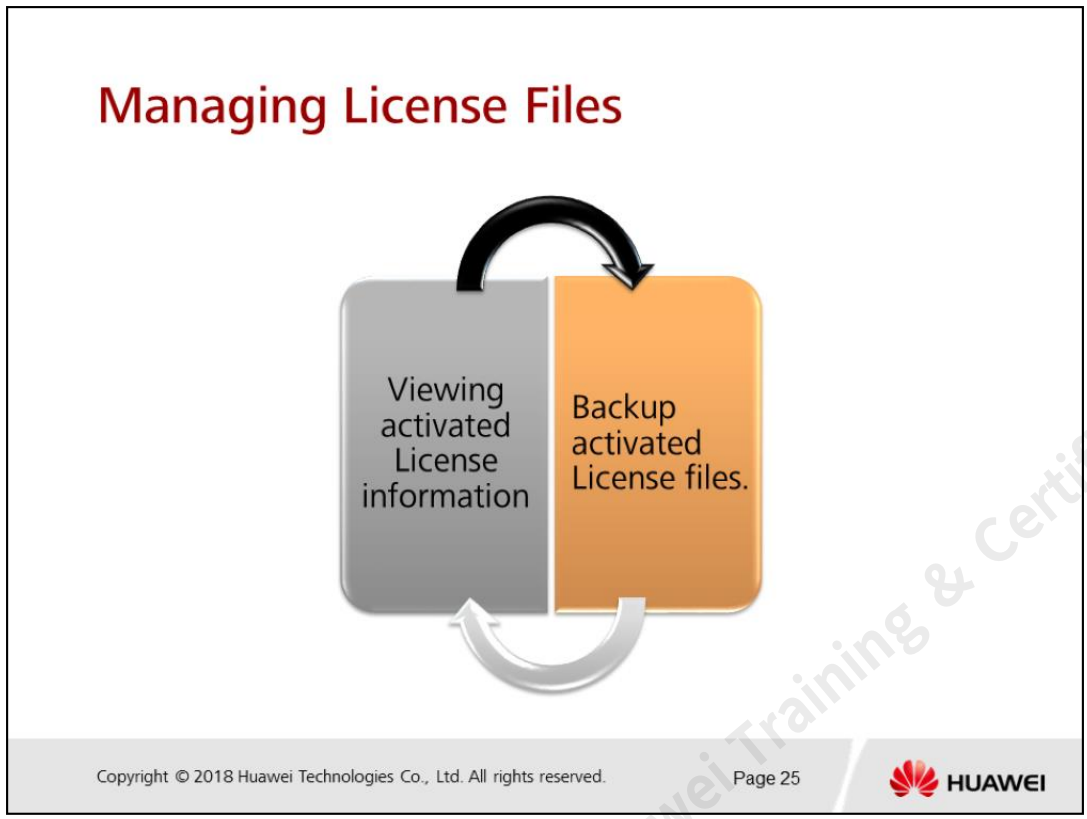
- Performance Monitoring Overview: Performance is a manifestation of the comprehensive capabilities of a storage system. In the service running phase, users can fully understand the performance of the storage system by performing real-time performance monitoring and performance trend analysis on the storage system. When performance problems occur, the analysis and locating of problems can be performed based on the performance monitoring data.
- Planning monitoring items: Under different service scenarios, different types of performance indicators need to be monitored by the users.
- Daily performance monitoring: On Huawei OceanStor V3 storage, daily performance monitoring of the storage system can be performed through the storage system management software DeviceManager and performance monitoring tool SystemReporter, enabling users to view and fully understand the operating status of the storage system.
- Locating performance issues: System performance issues is determined by the presence of a bottleneck in the system. When performing performance optimization and locating of performance issues, it is first necessary to clarify the service scenario currently carried by the system and the user's demand for performance. Followed by the usage of the system I/O path as a clue to determine the module where the performance problem lies. Finally, locate and optimize the performance problem of the storage system.

Huawei Training & Certification



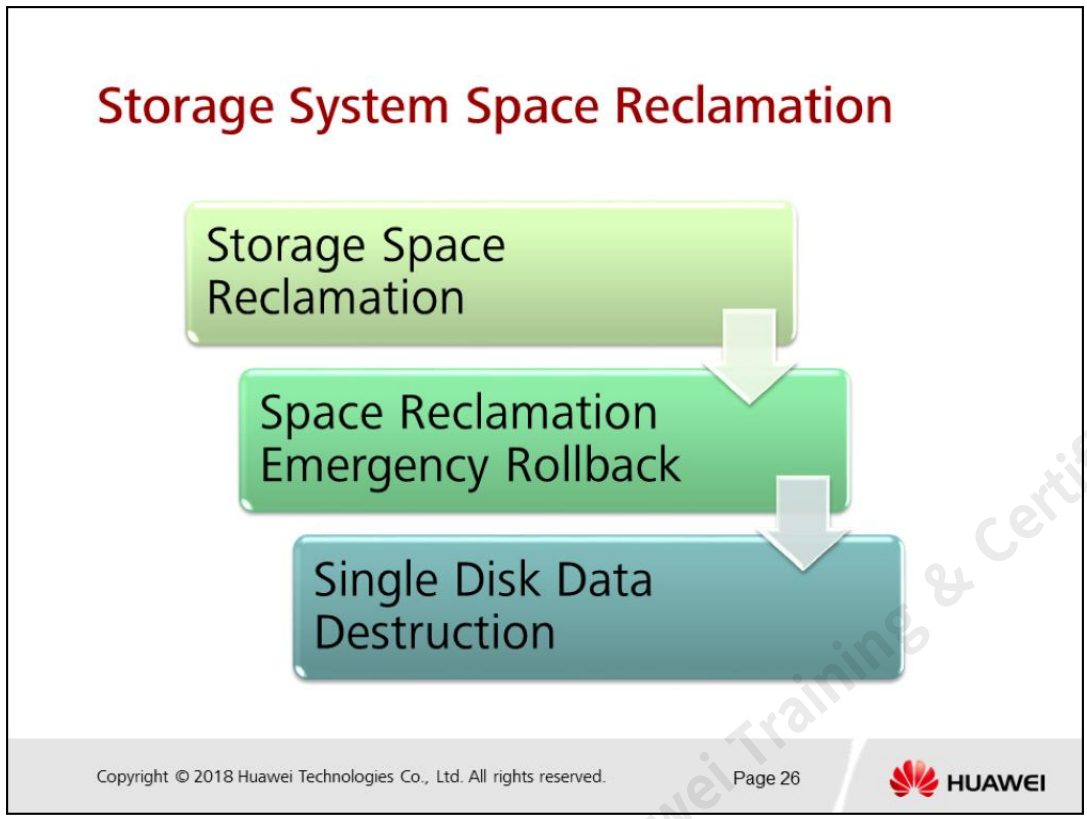
- Configuring the system time: When the time of the storage system deviates, you need to modify the time of the storage system so that the alarm logs can accurately locate the time when the alarm is generated. Through this operation, you can synchronize the client time to the device, set NTP to automatically synchronize, or manually modify the device time.
  - NTP (Network Time Protocol) is a protocol that synchronizes computer system time. It can synchronize the computer's system time to UTC (Universal Time Coordinated). Servers that support the NTP protocol are called NTP servers.
  - By synchronizing the client time, you can adjust the storage system time to match the client time.
  - By setting up NTP automatic synchronization, you can use the NTP server as an external time source to periodically and automatically synchronize the time to the device.
- Modifying the device name and location information: This operation allows you to set the name of the device and the geographical location of the device.

Huawei Training & Certification



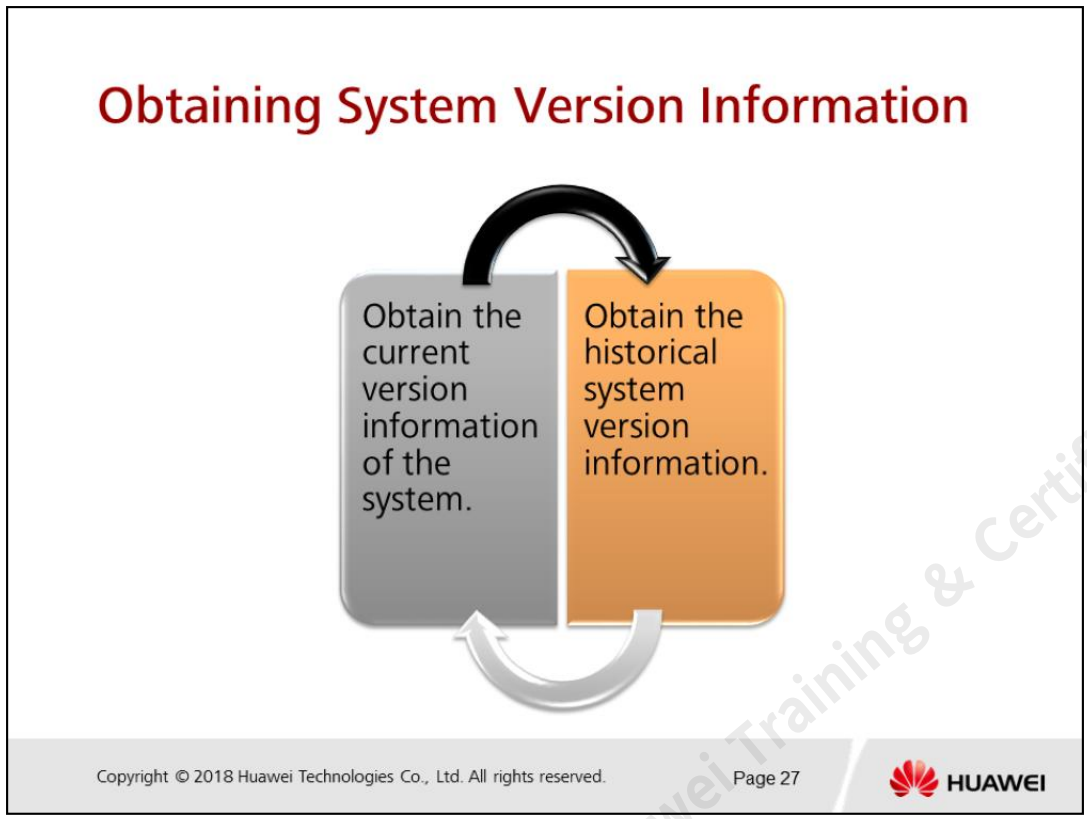
- Viewing activated license information: Before using the value-added service, you need to check the activated license file to ensure that the license file for the feature is available.
- Back up the activated license files: After the license file on the device becomes faulty, you can import the backup license file again for usage.

Huawei Training & Certification



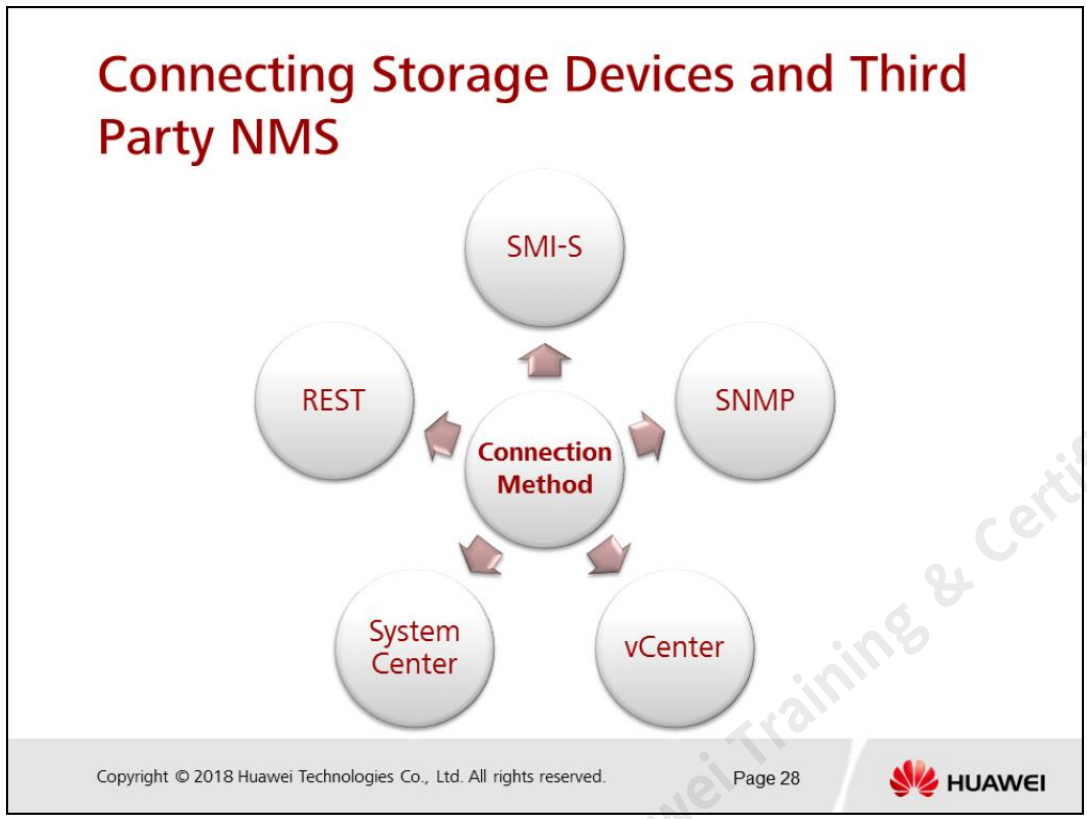
- Storage system space reclamation: When all or part of the storage system services do not need to be in operation, or the allocated capacity of the capacity expansion is not actually used, this part of space can be reclaimed, and these spaces can be used to run new services to improve the utilization of storage space.
    - Space recovery is divided into two types: full recovery and partial recovery.
    - Before implementing space recovery, it is necessary to complete preparations such as running status checks of the host and storage devices and backup of disk information to ensure that the space recovery process can be implemented smoothly.
- Emergency recovery of space reclamation: In the event of an exception, anomaly or failure during the space reclamation process, emergency rollback can be performed. The emergency recovery of space reclamation involves two scenarios which are, full recovery and partial recovery.
- Single disk data destruction: When data on a hard disk is no longer needed, single-disk data destruction can be performed on the hard disk. After this operation is performed, the deleted data cannot be restored to ensure data security. Data destruction can only be performed on encrypted disks that have not been added to the disk domain.

Huawei Training & Certification



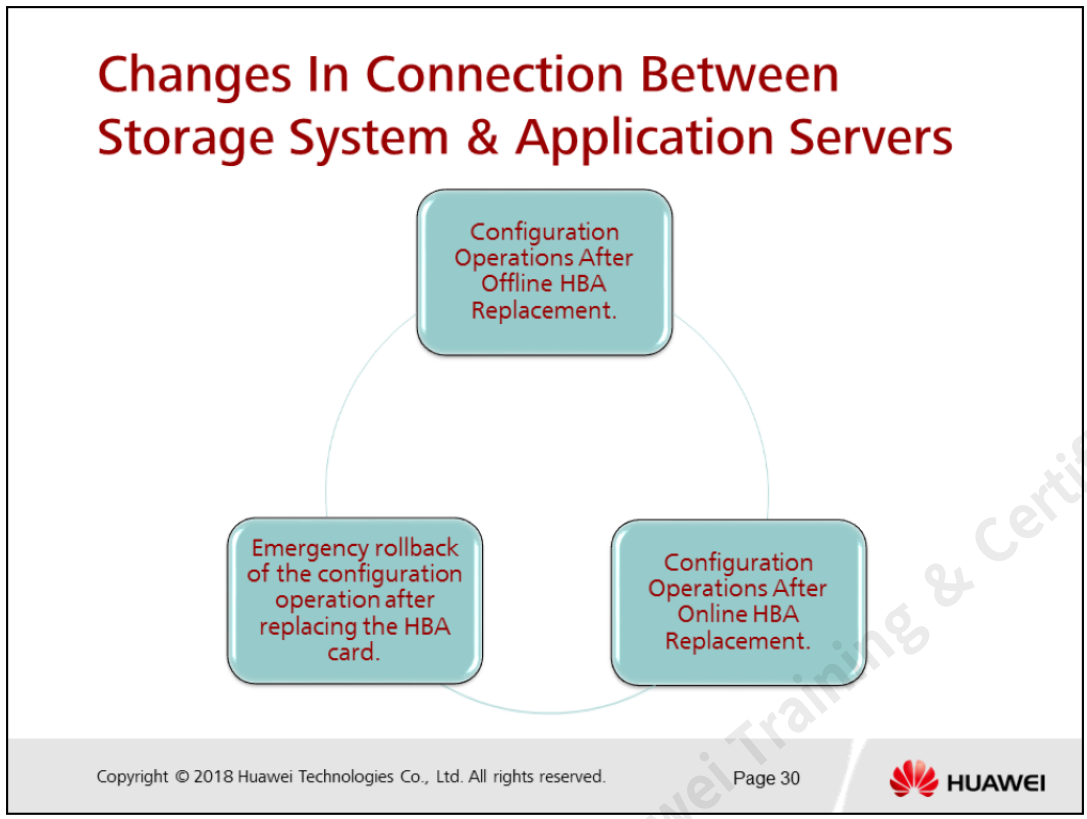
- Obtain the current version information of the system: Users can query and learn the current version information of the storage system through the DeviceManager.
- Obtaining historical system version information: Users can query and learn about the historical version information of the storage system through the CLI.

Huawei Training & Certification



- Huawei storage devices support interface protocols such as SNMP and SMI-S. Users can manage storage devices through a third-party network management system based on these protocols.
- SMI-S: By installing the SMI-S provider on a third-party Windows/Linux server, users can manage the Huawei storage system through the SMI-S provider.
- SNMP: The third-party NMS can view information about storage devices such as LUNs, ports, and storage pools through the SNMP protocol.
- vCenter: OceanStor VMware vCenter Plug-in (referred to as vCenter plug-in) is a storage management plug-in developed based on the vSphere Web Services SDK (Software Development Kit) to implement management of Huawei storage devices through the vSphere client.

Huawei Training & Certification



- After the connection between the storage system and the application server is changed, you need to change the configuration on the storage system and application server so that the application server uses the storage space of the storage system through the new connection channel.
- After the HBA of the host is replaced, configure the storage system and the host to make the replaced HBA work normally. Before configuring the replacement HBA card, you need to complete preparations such as host multipath and disk information backup and running status check on storage device to ensure that the space configuration process can be implemented smoothly.
- After an HBA is replaced, an emergency rollback can be performed if an exception or fault occurs during the configuration operation on the host or storage side.

Huawei Training & Certification

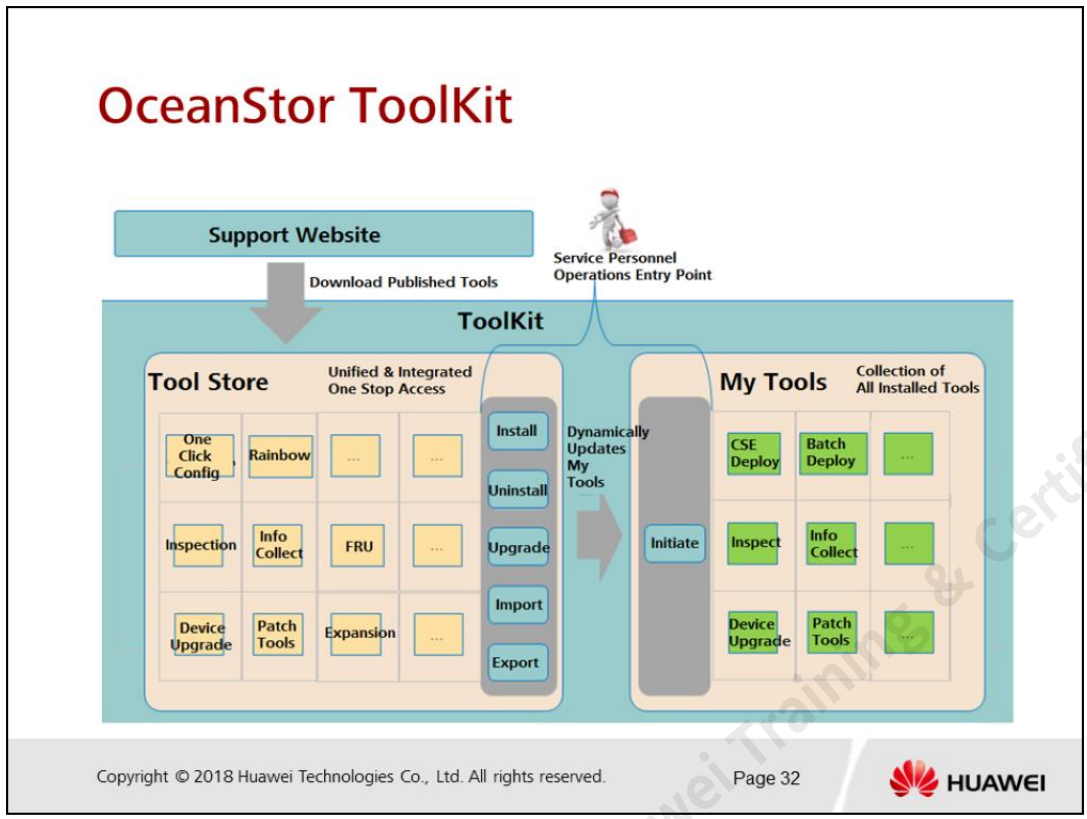
Contents

1. Administrator Guide.
2. Routine Maintenance.
3. Daily Management.
4. **Common O&M Tools.**

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

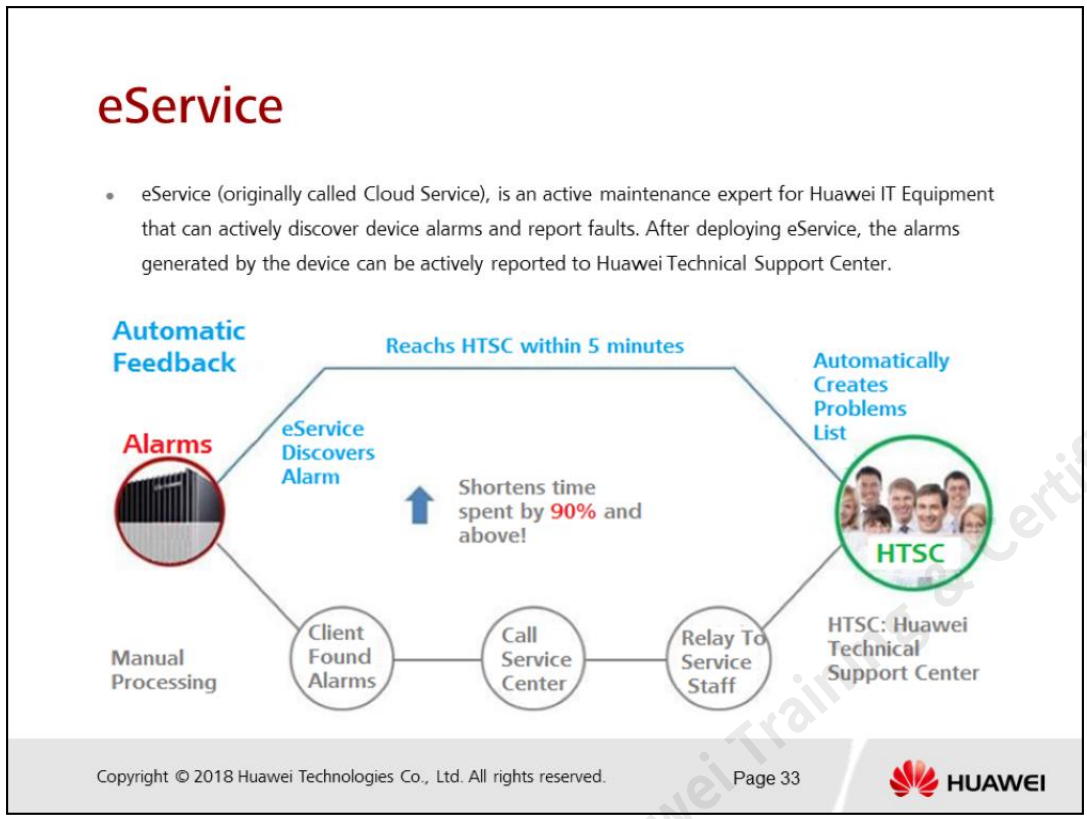
HUAWEI

Huawei Training & Certification
Huawei Training & Certification



- OceanStor Toolkit is a serviceability tool developed by Huawei. It is a unified desktop management platform for all IT tools. Its built-in tool store can quickly download, install, or upgrade tools, and can automatically check installed systems to see whether the installed tools needs to be upgraded.
- The Toolkit contains various tools required for the deployment, maintenance, and upgrade of IT equipment. It can help service engineers and maintenance engineers perform accurate operations on the equipment in the O&M process, reducing the difficulty of operation and maintenance, thus improving work efficiency.
- The maintenance tool Toolkit of the storage product is a one-stop management tool used in various operation and maintenance scenarios to improve the efficiency of the various service operations such as product deployment, daily maintenance, and upgrade, and reduce the operation difficulty.
- Due to the fact that maintenance operation has the operating rights and professional skill requirements, the tool settings have permission/access right control. For more details on the specific tool permission/access rights, please contact technical support engineers.
- You can search for product documents related to the OceanStor ToolKit at http://e.huawei.com.

Huawei Training & Certification



- The traditional service support method is fully manual local service which means that the problem need to reported to the local support center before support personnel can act on those problems. In the fault discovery process, technical service personnel are faced with the challenge of not finding the problem in time and failing to deliver the information. When a device fails, the alarm reporting function provided by the eService can transmit all the fault information to the Huawei technical support center in a timely manner, shortening the time for troubleshooting and processing, and meeting the current urgent needs of the enterprise.
- The eService is deployed on a customer-supplied server to receive device alarm information and send it to Huawei technical support center. It also supports uploading device logs and inspection reports to Huawei technical support centers.
- You can search for http://e.huawei.com to obtain product documentation related to the eService tool.

Huawei Training & Certification

Summary

- This module mainly introduced:
  - The daily operation guide for Storage Administrators in the datacenter.
  - The daily maintenance items for datacenter storage systems.
  - The daily management tasks for datacenter storage systems.
  - The common operation and maintenance (O&M) tools for datacenter storage systems.

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

Huawei Training & Certification
Huawei Training & Certification

Quiz

1. (True or False) OceanStor eService is a Huawei developed serviceability tool, an unified desktop management platform for all IT tools, with the inbuilt Tool Store that can quickly download, install and upgrade tools and also automatically checks whether the installed tools needs an upgrade. ( )
2. Which of the followings is not an item in the weekly maintenance of datacenter storage systems ? ( )
   A. View and Process Alarms.
   B. Inspect Storage Devices.
   C. Inspect Server Room Environments.
   D. Check Internal Environments of Racks/Cabinets

Copyright © 2018 Huawei Technologies Co., Ltd. All rights reserved.

HUAWEI

- Answers:
  - False.
  - A.

Huawei Training & Certification
Huawei Training & Certification

Thank You

www.huawei.com

Huawei Training & Certification
Huawei Training & Certification

Recommendations

- Huawei Learning Website
  - http://learning.huawei.com/en
- Huawei e-Learning
  - http://support.huawei.com/learning/NavigationAction!createNavi?navId=MW000001_term1000025144&lang=en
- Huawei Certification
  - http://support.huawei.com/learning/NavigationAction!createNavi?navId=_31&lang=en
- Find Training
  - http://support.huawei.com/learning/NavigationAction!createNavi?navId=_trainingsearch&lang=en

More Information

- Huawei learning APP

Huawei Training & Certification

版权所有© 2018 华为技术有限公司