12 Introduction to Data Center

www.huawei.com

HUAWEI

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.

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.

Contents

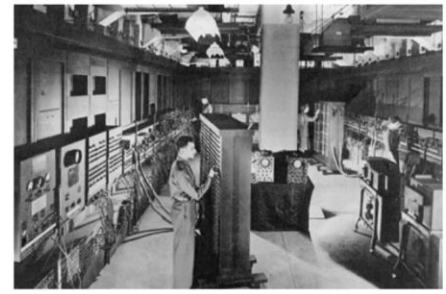
HUAWEI

Evolution of Data Centers

HUAWEI

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.

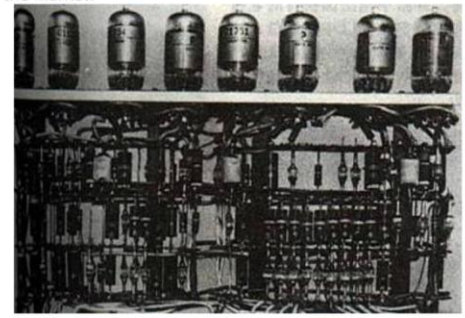


- The idea of developing an electronic computer was generated during the period of World War II. During those times, the war was going on actively and the military equipment of all the countries was still very primitive, and the main strategic weapons used at the time was planes and artillery. Hence, 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.

- It was in the midst of World War II that computers began to be developed by scientists. As mentioned earlier, there was an urgent need for a high-speed computing tool in order to provide an accurate and timely ballistic firing tables for the military tests. Therefore, with the strong support of the U.S. military, ENIAC, the first electronic computer in the world, started its development in 1942. Participating in the development work of ENIAC was the research team lead by John Mauchly and J. Presper Eckert from the Moore School of Electrical Engineering of the University of Pennsylvania.
- But even at the time of ENIAC's creation, it was filled with flaws and disadvantages such as :
    - In addition to its large size, and high energy consumption, and the high heat generated by the machines while running can cause the vacuum tubes to be damaged.
    - Whenever a single vacuum tube is damaged, the whole machine could not function properly unless if you first find the one damaged vacuum tubes out of the 18000 existing vacuum tubes and replace it. This proves to be very inefficient, troublesome and time wasting.

Commercialization of Virtualization Technologies in Data Centers - TRADIC

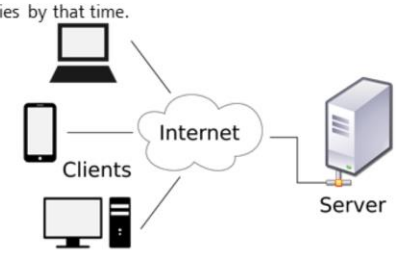
- Unlike vacuum tube systems, the first transistor based computer (TRADIC) was developed in 1954, but more advanced commercial systems appeared in the 1960s and lead to a breakthrough in mainframes technology such as the IBM System Mainframe. Mainframes are primarily developed for government and military 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.



- 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.

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.



- 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.

- B/S Architecture:
  - Browser/Server architecture was introduced with the rise of Internet technologies, it can be considered as an improvement or evolution of the existing C/S architecture. In this architecture, the user interface is fully implemented by the WWW browser. No applications need to be installed and the program under B/S architecture can run on existing browsers on the computers via the internet.
  - Basically, the client does not have a dedicated application and the application is installed on the server side. As the client don't have the application installed on their side, application upgrade and maintenance can be done solely on server side, making the upgrade and maintenance process easier.
  - Due to the fact that users are using the browser as the interface, the user interface can be customized and built with multitude of different designs. However, certain functions such as data printout and other functions are limited on the client side because the data and the application is on server side, while only a browser is used as program interface. In order to overcome this limitation, any functions that are unable to be completed in browser are usually developed separately as a distributable plugin that can be released to the users, and these functions can be completed via a program call on the client side.

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.

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.

- A container Data Center (CDC) is a larger containerized (modular) Data Center system that typically ships in a standard 20ft/40ft containers or structure of similar sizes, and it is suitable for temporary deployments in open areas and harsh environments.
- Modular Data Centers can meet the urgent needs of IT business units towards the future construction of Data Center infrastructure such as the needs for standardized design, factory prefabrication of components, rapid on-line deployment, effective initial investment reduction, energy pooling and management within modules, high dynamic IT infrastructure resources utilization, intelligent operation and maintenance management, guaranteeing of important business continuity, providing shared IT services (such as cross-business infrastructure, information, application sharing, etc.), rapid response to changes in business needs, and building green energy-saving data centers.
- Outstanding advantages of Modular Data Centers:
    - Highly Reliable Standard Modules
        - Modular Data Center uses modular, standardized and highly reliable design, making the stability of the whole system higher. Based on the customer requirements and actual situations faced, we can provide N+1, N+X, 2N solutions on the configurations and implementations of the core power distribution and cooling equipment. It is safe and reliable while able to fulfill standard requirements to meet Tier 3 and Tier 4 levels of requirements in terms of Data Center design.
        - A Tier 3 data center is concurrently maintainable, allowing for any planned maintenance activity of power and cooling systems to take place without disrupting the operation of hardware located in the data center. In terms of redundancy, Tier 3 offers N+1 availability. Any unplanned activity such as operational errors or spontaneous failures of infrastructure components can still cause an outage in Tier 3 Data Centers. In other words, Tier 3 isn't completely fault tolerant.
        - A Tier 4 data center is fault-tolerant in addition of having all the benefits of Tier 3, allowing for the occurrence of any unplanned activity while still maintaining operations. Tier 4 facilities have no single points of failure. Tier 4 offers 2N+1 fully redundant infrastructure.

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.

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.

Contents

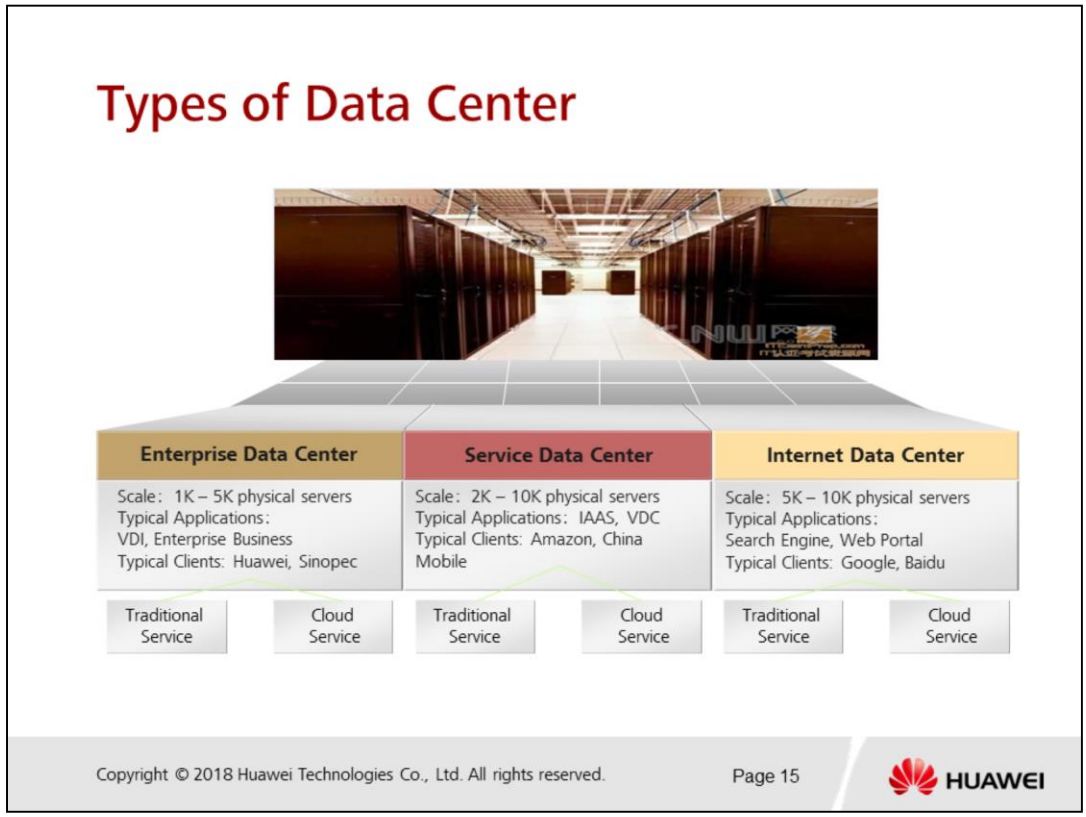
HUAWEI

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.

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.

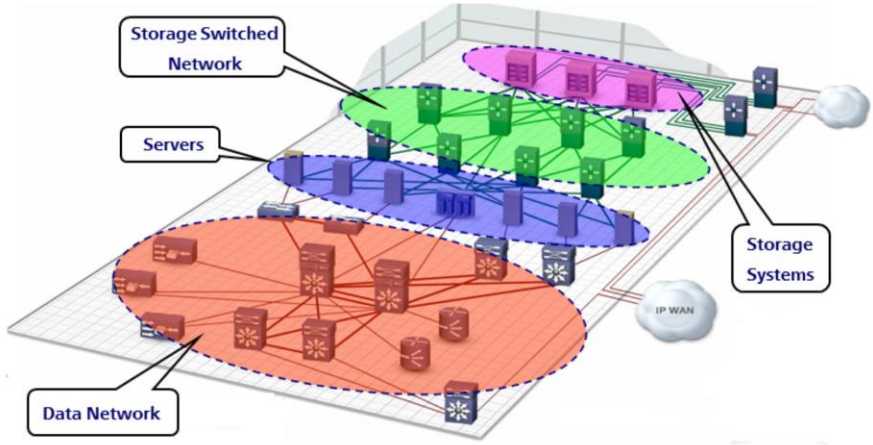


- Data Centers are facilities built to host and access the server's services, and mainly it can be categorized into a few types such as: Enterprise Data Center, Service Data Center and Internet Data Centers.
- Each type of these data centers have different needs that they cater to, and the scale of these 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.

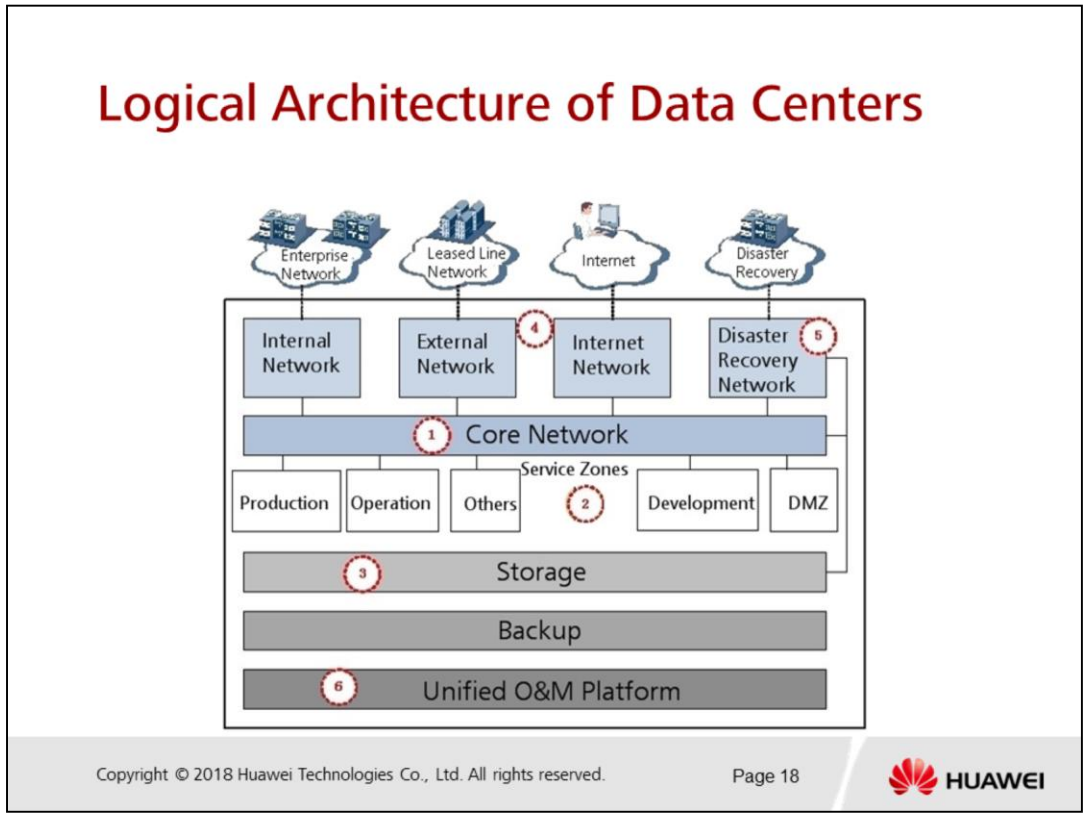
Components of Data Center

- 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.

Standard Data Center Architecture

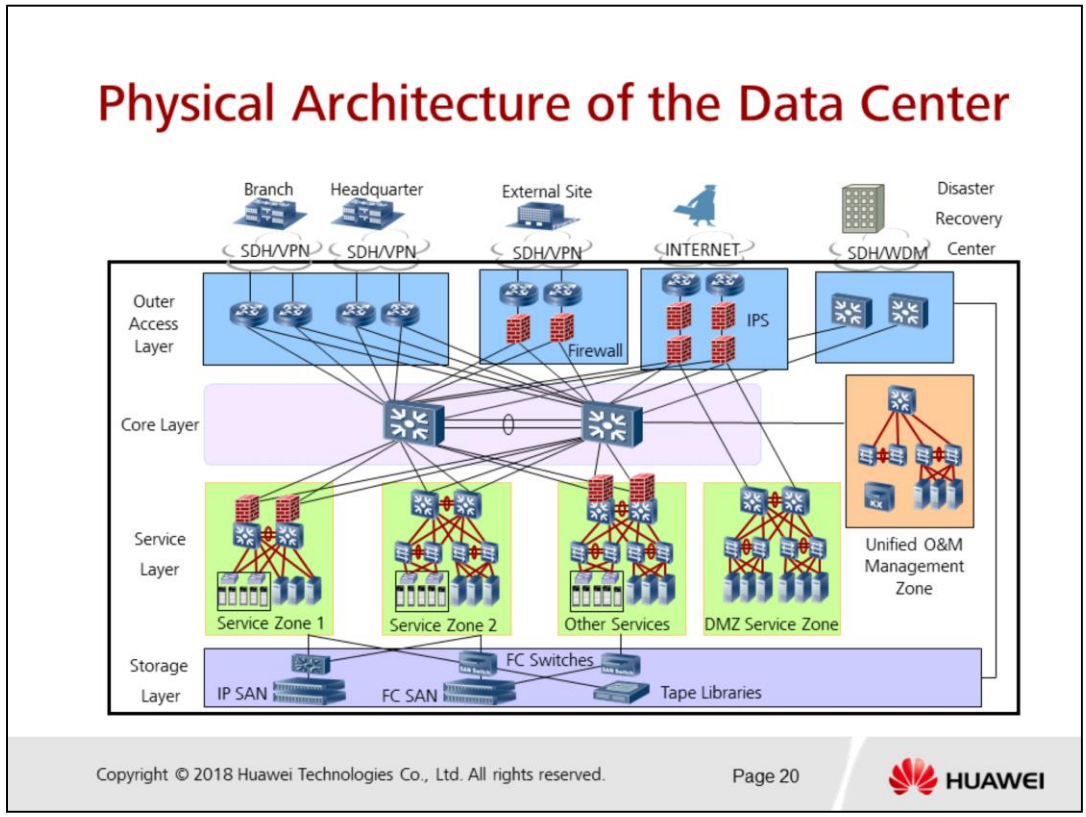


HUAWEI



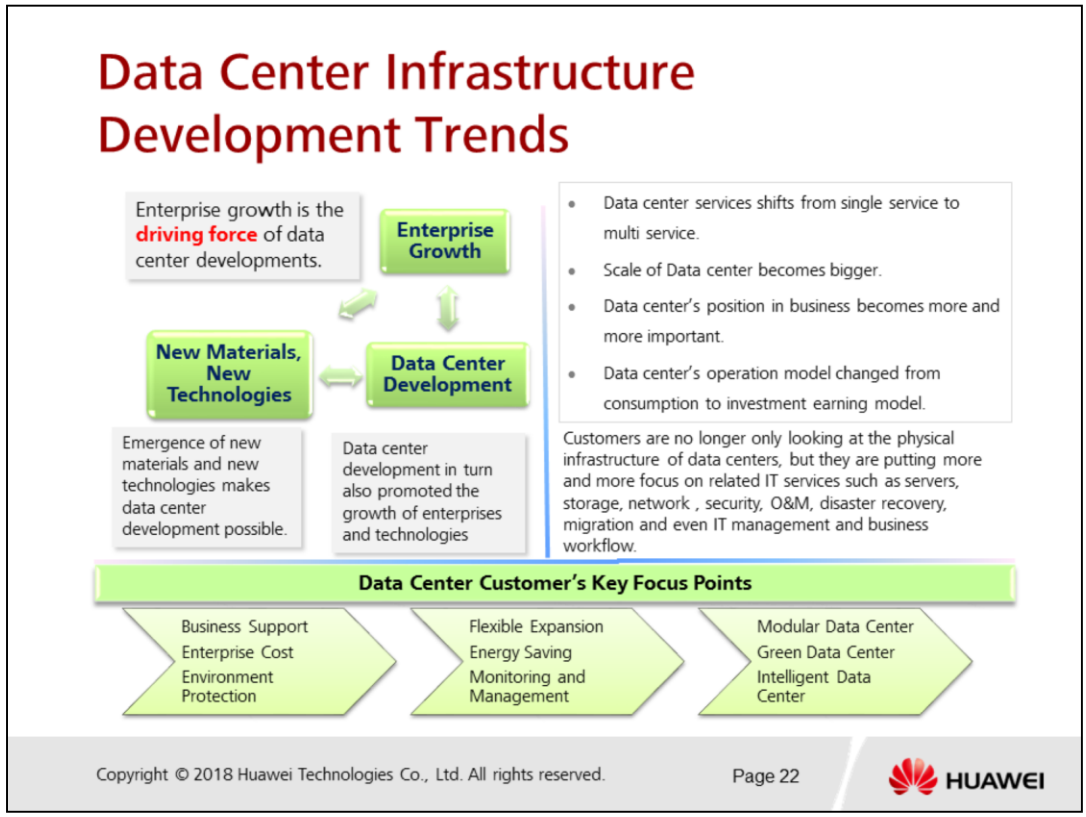
- Core Network Zone: It is the core of the data center network. It connects the internal server areas, enterprise internal network, partner's network, disaster recovery center and the external user access network.
- Server Zone: This the area for the deployment of servers and application systems. Based 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).

- Internet Zone: This is the area where internal and external users accesses the data center. Based on security and scalability considerations, this area can be further divided into different networks according to the types of users connected such as: Internal Network, Partner Network, and Internet Network. Internal Network refers to the network connection that connects the headquarters and branch's network through Campus Network or Wide Area Network(WAN). Partner Network connects to the data center through the use of leased lines, wide area leased lines, and the network existing in the partner's location. Internet network allows public users to access the data center, and allows outstation employees to safely access the data center via the Internet without the need of connecting to the WAN network in the office.
- Disaster Recovery Center Zone: This is the area where the data center is connected to the Disaster Recovery Center (DRC). Transmission equipment is mainly used for the connection to the DRC within the city, and WAN technologies is used for connecting between the data center and DRC located out of the city or remote regions.
- Operation and Maintenance Management Zone: This is the area where management of the network, servers, applications, storages is implemented. The types of management includes fault management, configuration management, performance management and security management.



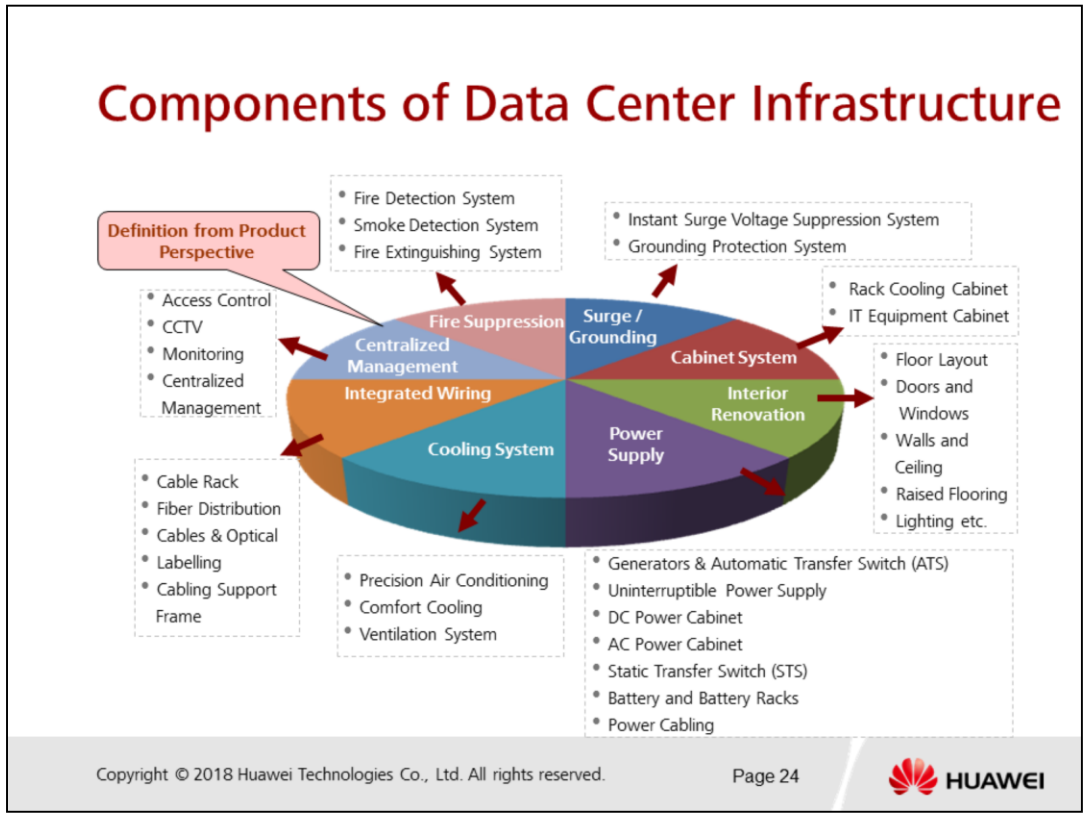
- The diagram above shows the physical architecture of the data center. Generally, the data center is segregated into different layers that that has different functions for better scalability and security.
- The Access layer provides the connectivity for the data center equipment to the devices and networks 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.

    - Wavelength division multiplexing, WDM, has long been the technology of choice for transporting large amounts of data between sites. It increases bandwidth by allowing different data streams to be sent simultaneously over a single optical fiber network. In this way WDM maximizes the usefulness of fiber and helps optimize network investments.
    - Security of the data centers is usually protected by implementation of firewall and Intrusion Prevention System(IPS) on networks that connects externally to the Internet or remote sites.
- The Core layer is the high speed switching backbone that mainly focuses on switching and routing the packets as fast as possible and aggregates the service layer. Usually it consist of high performance equipment to manage large amount of workloads and fully redundant to avoid single point of failure ensuring critical business service continuity.
- The Service layer consist of all the servers and equipment necessary to run the business services and applications. Different sets of equipment can be used to run different services by different data center tenants.
- The Storage layer consist of all the storage devices and network devices for FC SAN and IP SAN. All the data intended for data access or data storage is transmitted between layers via the storage area network through the FC or IP switches. Data usually are stored in storage arrays of hard drives or tape libraries in data centers.



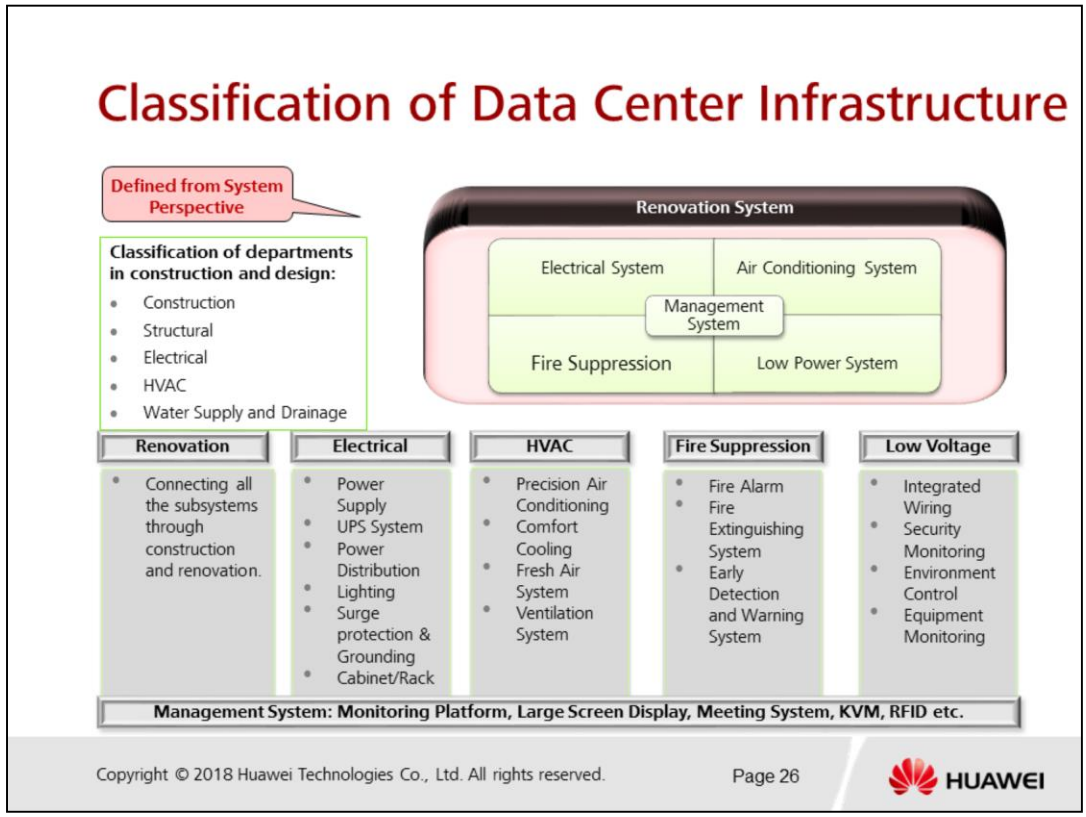
- IT systems and information construction have become the driving force of enterprise growth and development, which means that enterprise growth or more specifically its business growth has generated continuous demand on IT systems and in turn became the primary driving force on data center development.
- 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.

- To put these concepts in example, imagine an enterprise that is expanding its business, which makes its scale of the data center much bigger, its daily operation and maintenance of the data center getting more complex and tedious with higher costs and risks. This has created a demand for an intelligent management of the data center. The demand of intelligent management in turn created a demand for a corresponding material and technology such as a database with auto data collection feature, application that works with this database and the front end platform technology required for managing these technologies. With the maturity of these technologies and materials, intelligent management of the data center is realized, and while it increases the performance of data centers, it also eliminated the bottleneck faced by the enterprise which promoted the next level growth of the enterprise's business.
- From the example shown above, we can see that enterprise growth, new materials and technologies, and data center development are interrelated and has a relationship that both benefit, restrict and promote each other's growth. At the same time, enterprise growth is the primary driving factor for data center development as demand decidedly creates growth and development.
- In the recent years, what are the changes or development that data center has gone through?
- At first, lets discuss on the changes on data center itself, the following are the key changes of the data center:
    - Data center services has changed and grows from single service to multi service development trend. In the past, data center services was quite specialized for single service, for example, it was a computing center that mainly provides support for enterprise business production. But nowadays, data centers has a wider range of services and features that not only able to support enterprise production but also support in enterprise management, enterprise training and enterprise daily operational activities, and even can provide external services to others.



- 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.

- We noticed that when all 8 of the different systems is usually defined, much focus is placed on the products contained within the system rather than the system itself. For example when the power supply system is defined, focus in placed on the diesel generators, ATS, UPS and other products. Although in a sense we could define those systems by the component products. However, data center design and building is a project that satisfy a certain function or customer requirement, which means that we must consider all the factors from a system's perspective instead of considering from a product perspective that may not be able to meet the requirement of the customer. We must consider the suitability, connection and the interrelated relationship of those devices with the matching system design and installation so that we could satisfy the final requirements.
- Air conditioning covers many things but is mainly, heating, cooling, ventilation, fresh air, filtration and control of humidity. Comfort cooling mainly relates to provision of just cooling to improve comfort. Precision Air Conditioning is designed for a wide range of applications where close control, high precision air conditioning is essential, including data center cooling, medium and low density server environments, telecom switching stations, and clean room environments.
- An automatic transfer switch (ATS) is a device that automatically transfers a power supply from its primary source to a backup source when it senses a failure or outage in the primary source. When a failure occurs in a primary power system, the ATS invokes a standby power source, such as an uninterruptable power supply (UPS). An ATS can also start up more long-term backup power systems, such as local diesel generators, to run electric equipment until utility power is restored.
- STS is based on static electronic components (SCR) therefore allowing for a fast and precise control of the switching between one line and the other. This solution permits to obtain a perfect Break Before Make (BBM) behavior by never permitting a source overlap. Moreover, it is also capable of very fast switching between the two sources with a max delay of less than 5 msec (typically 4 msec). ATS is based on electromechanical components where the BBM switching is actually made by controlling the relays on each source line. This kind of technology can still make a perfect Break Before Make change of supply sources both in synchronous and asynchronous conditions but it is certainly slower than the static solution.

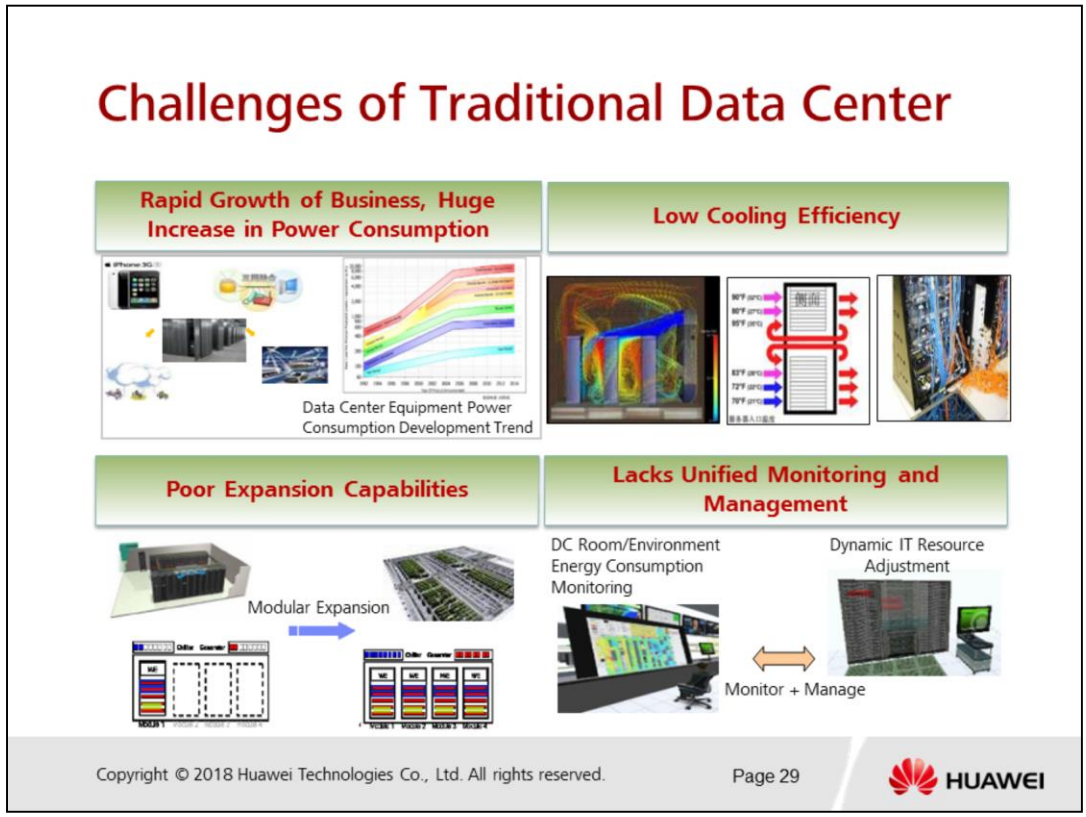


- For the following segment, lets look at the content of the Data Center L1 Physical Infrastructure Design from a system perspective without focusing on products or equipment.
- Building a data center involves the expertise of multiple different fields in order to build and interconnect all the different 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).

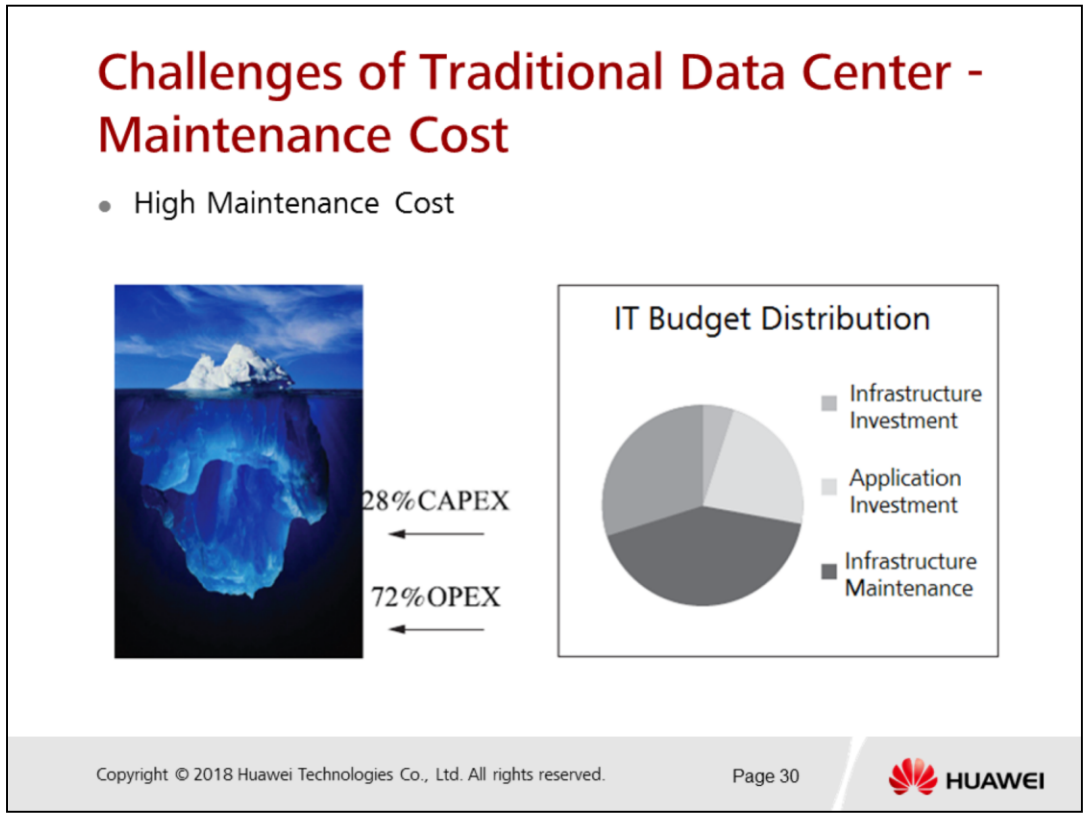
- Each of the system has its own functionality and plays an important role in the data center.
- The electrical system satisfies the needs for power supply by the data center equipment (including IT equipment, power equipment, support equipment). Air conditioning system satisfies the data center needs for different kinds of environmental, temperature and ventilation control. Fire suppression system satisfies the needs for fire alarm and fire extinguishing in data centers. Low power system refers to low voltage distribution system (lower than 36 volts in AC or lower than 24 volts in DC) such as the network cables, CCTV cables or other cabling or equipment that requires or runs on low voltage current in the data center. Low power system completes the needs of data center in the aspects of data communication, security, environment and equipment monitoring. All these different subsystems requires the need of renovation (including partitioning and deployment) to interconnect with each other and operate optimally through the management system.
- Depending on the customer's requirement, we need to reasonably plan, deploy and coordinate with each other in order to build the physical infrastructure of a L1 Data Center. Without careful planning and design, all the subsystem involved within the data center might not be deployed and operate at the optimum intended performance, which in turn might not be able to meet the requirements of the customer.
- In the following, lets introduce the content of all these different systems.
- Renovation System: includes all the renovation and decoration works for the ceiling, floor and walls for the rooms in the data center facility.
- Electrical System: includes the power supply, UPS, power distribution, lighting system, lightning surge protection and grounding of systems and cabinets. The cabinets mentioned here are quite special as some parties would consider cabinets or racks as part of the renovation or low power system.
- Air Conditioning System: includes precision air conditioning, comfort cooling, fresh air system and ventilation system.
- Fire Suppression System: includes fire alarm system, fire extinguishing and early detection and warning system.
- Low Power Systems: includes the integrated wiring, security and monitoring system, environmental and equipment monitoring systems.
- Management System: includes the monitoring platform, large screen display to show the status of the data center in the control room , meeting system, KVM(Keyboard, Video, Mouse), RFID (Radio Frequency Identification).

Contents

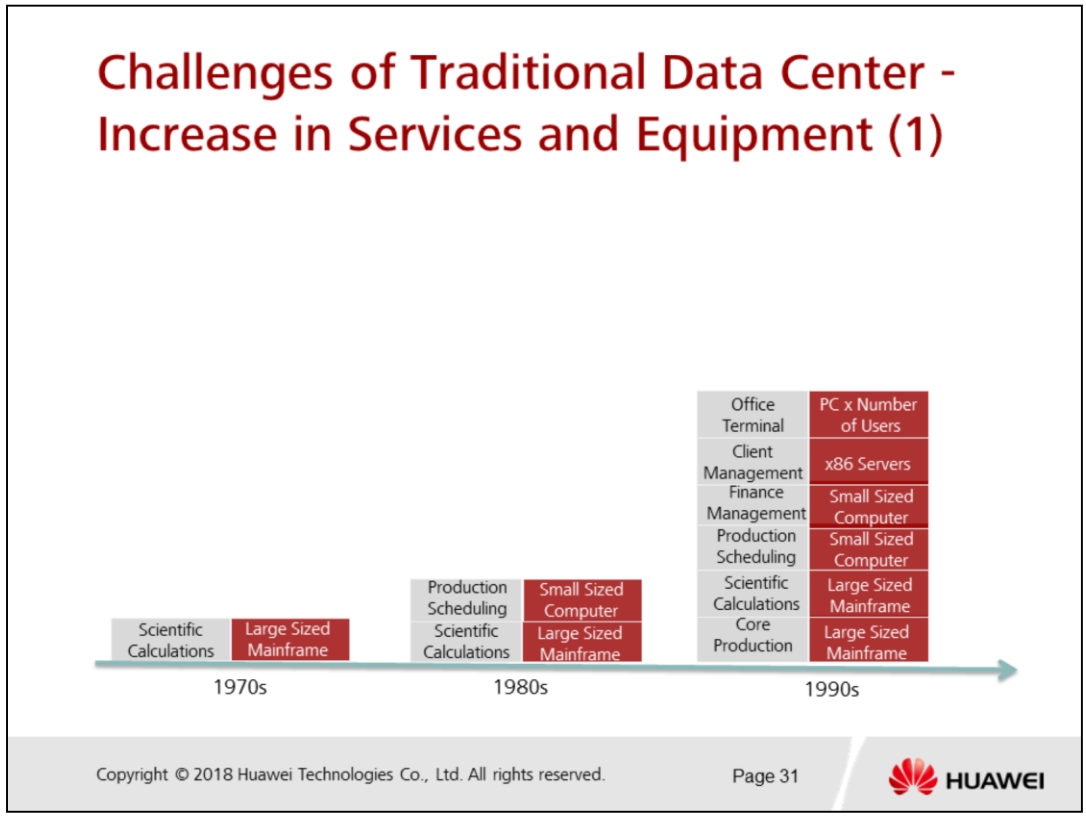
HUAWEI



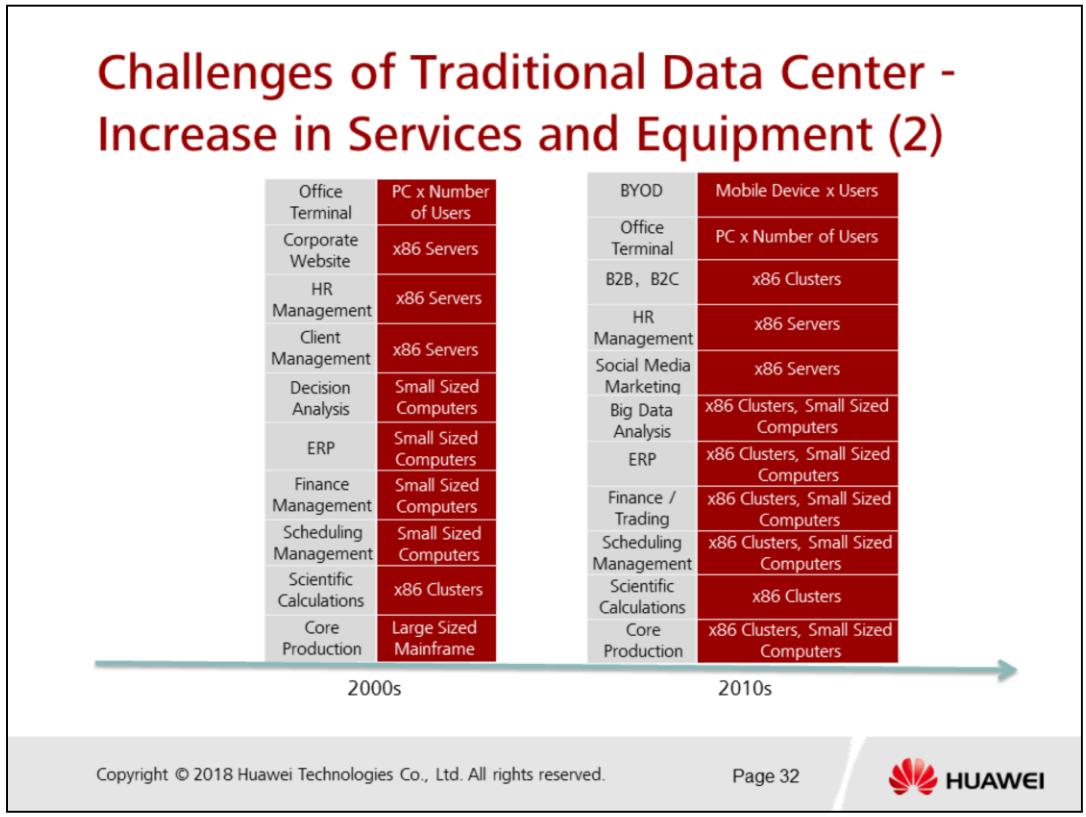
- With the growth of 3G, Cloud Computing, and Internet of Things, the power consumption of data center dramatically increased. Traditional data center power density increased from 3-5kW/cabinet to 10kw+ and overheating of certain areas in the data center became the biggest risk of safe operation of data 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.



- 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.



- 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.



- 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!

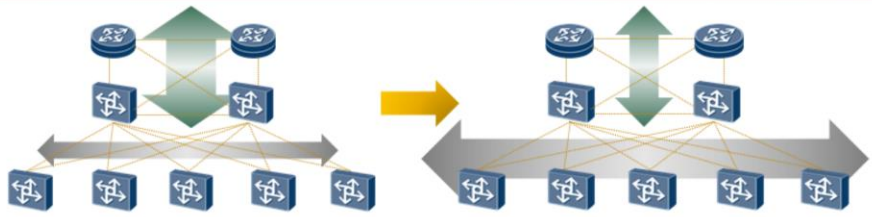
Challenges of Traditional Data Center - Increase in Data Traffic (1)

- 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.

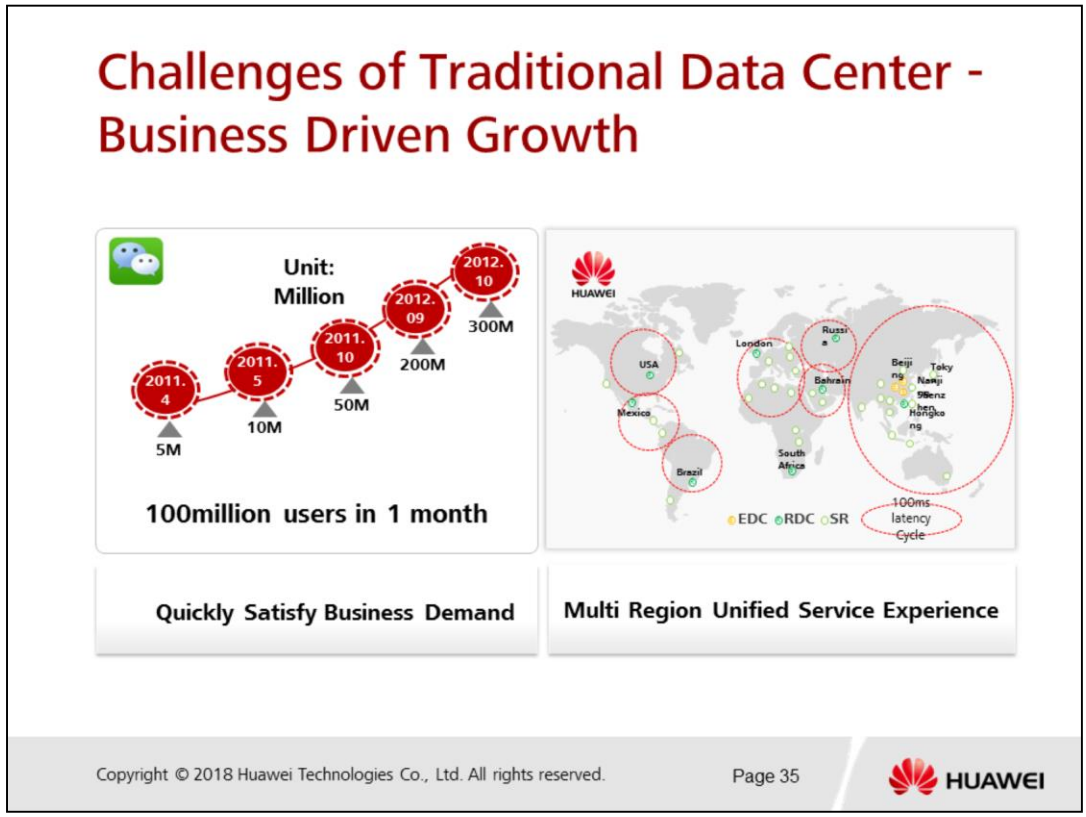
HUAWEI

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

Changes in Data Traffic: Increased Horizontal Flow

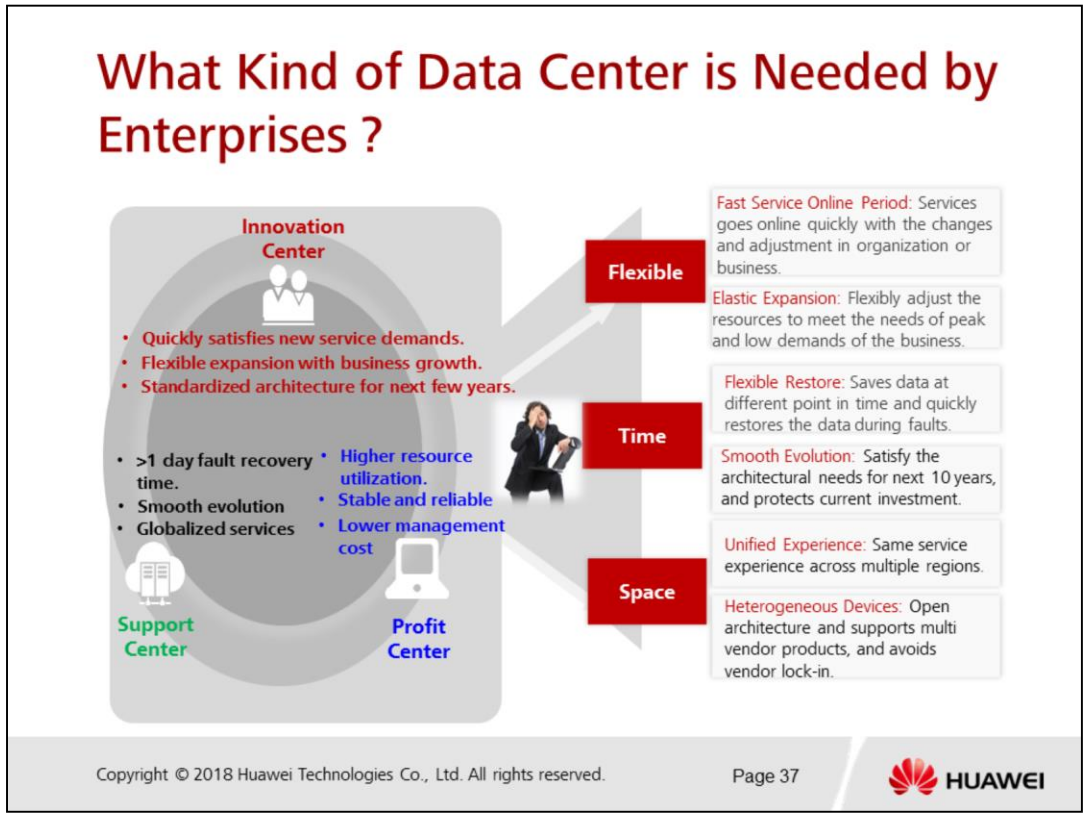


- Big Data, Parallel Processing, 3D Rendering, and Search Engine services requires the combined processing of clustered servers and generates high amount of horizontal flow of data within the network.
- Elastic scaling and online migration of virtual machines, and high amount of data transferred between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

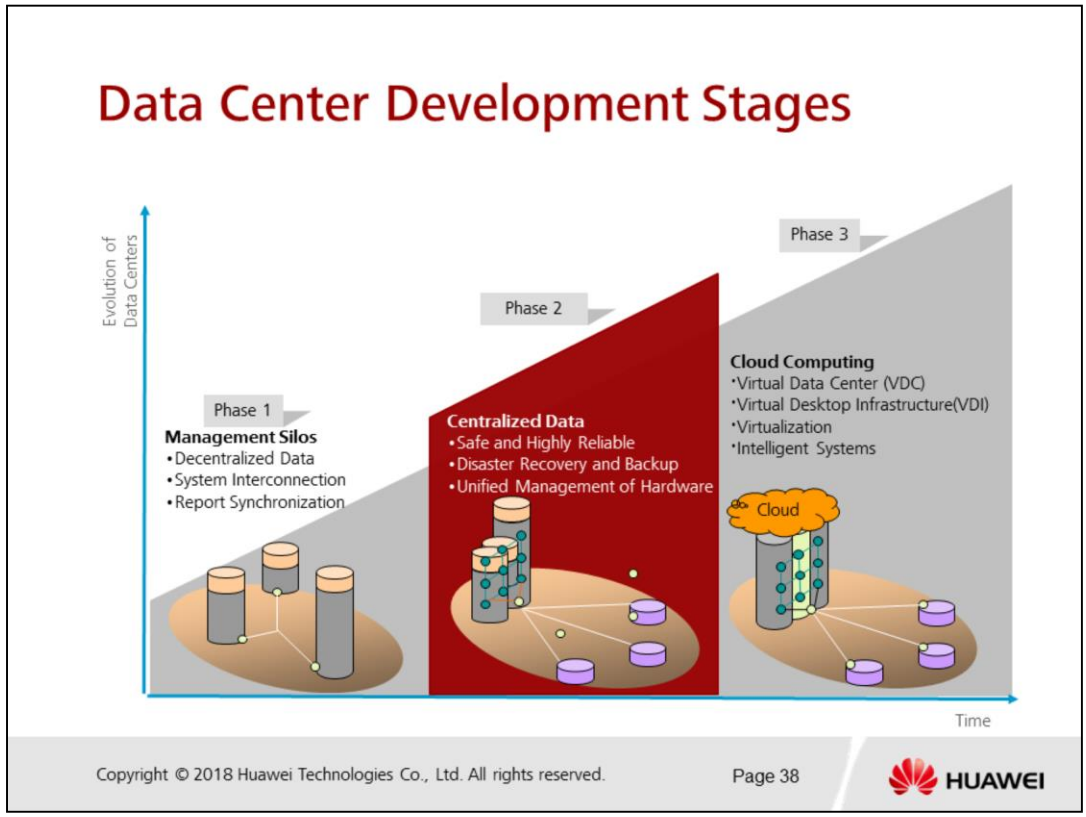


- Wechat launched its first version in 2011, and its user base increased from 5 million to 50 million in 6 months, however it only used 1 month to expand from 200 million users to 300 million users. This is an example that shows businesses are growing in a much faster pace, and how to quickly respond 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.

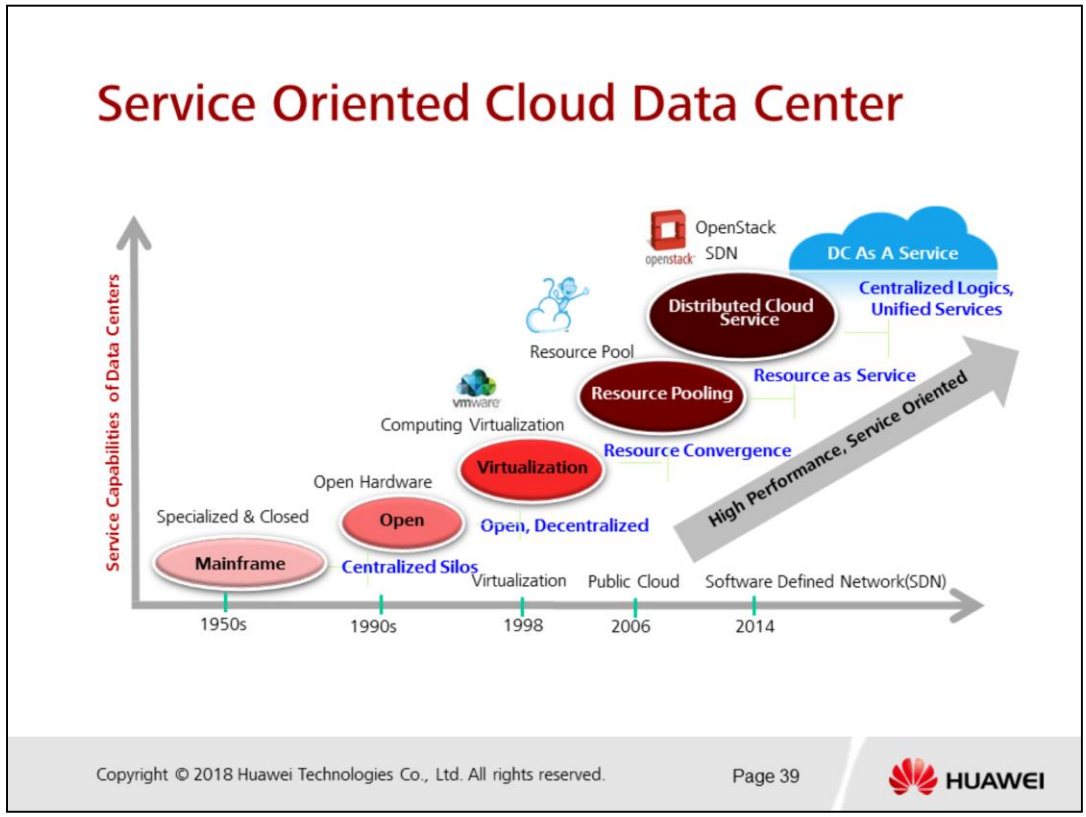
- Lets take Huawei data center as an example:
- Enterprise Data Center (EDC):
  - Dongguan Main Data Center and Same City Disaster Recovery Center
  - Nanjing Remote Disaster Recovery Center
- Regional Data Center (RDC):
  - Deployed 8 regional data center with 100ms latency cycle to support the services within each region. The RDC are setup at: England, Russia, Berlin, South Africa, Hong Kong, United States, Mexico, Brazil.
- Server Room (SR) in Huawei is categorized into 3 types:
  - Base (SR1) : Shanghai, Beijing, Chengdu, Hangzhou, India, Xian, Wuhan, Suzhou.
  - Law Compliance (SR2): Toronto, Australia.
  - Others (SR3)



- 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.



- 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.



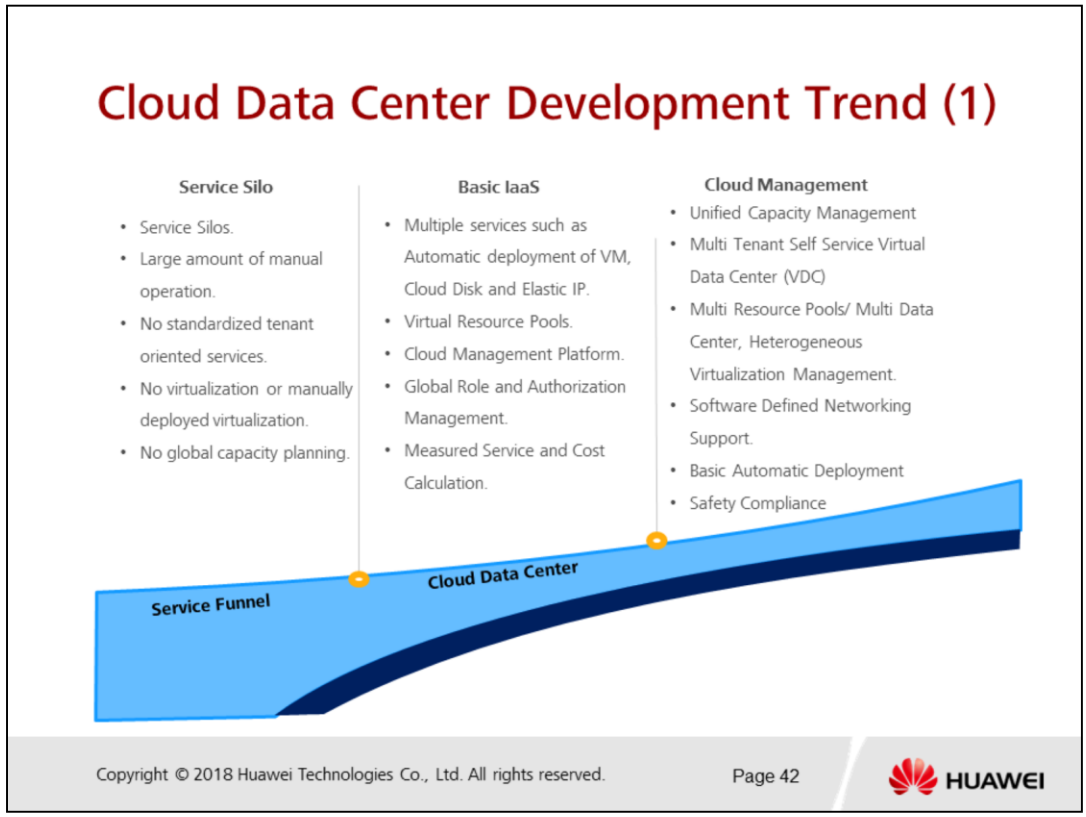
- The trend in data center technologies moves toward open architecture, and flexible service oriented data centers.
- Centralized computing is computing done at a central location, using terminals that are attached to a central computer. The computer itself may control all the peripherals directly (if 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

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.

- 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.

- Fast Deployment of Services: It allows business services to be deployed quickly and allows much agile support of business demands. With the changes in the market happening at a fast pace, Enterprise IT need to be able to adapt and change in order to support those ever changing business needs. Since, the IT equipment in a cloud data center are formed into resource pools, the IT resource needed by new services can be applied online via templates and deployment can be completed upon approval by the management. This totally changes the complex process in traditional data where you would need to apply, approve, purchase, and commission the equipment in order to support new services. This change in resource application process in cloud data center has shorten the period of deploying new services from 3 months using traditional process to few days using the new simplified process.
- Energy Saving: As the increase in utilization rate of servers, the numbers of servers required in the data center decreases, which lowers the power consumption of servers, lowers thermal load and lowers the power consumptions of other plugged in hardware. As newer types of cooling technologies appears in the market such as evaporative cooling, natural cooling, hot aisle cold aisle cooling can effectively lower the energy consumption for the data center cooling and ventilation. All these different energy saving procedures greatly lowers the energy consumption of the data center, traditional data center Power Usage Effectiveness(PUE) which defines the data center energy efficiency indicator is commonly between 2.5-3.0, however cloud data center can have a PUE of 1.5 or lower.



- 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).

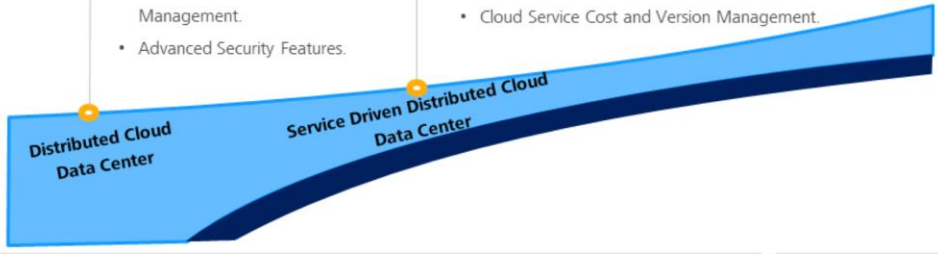
Cloud Data Center Development Trend (2)

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.

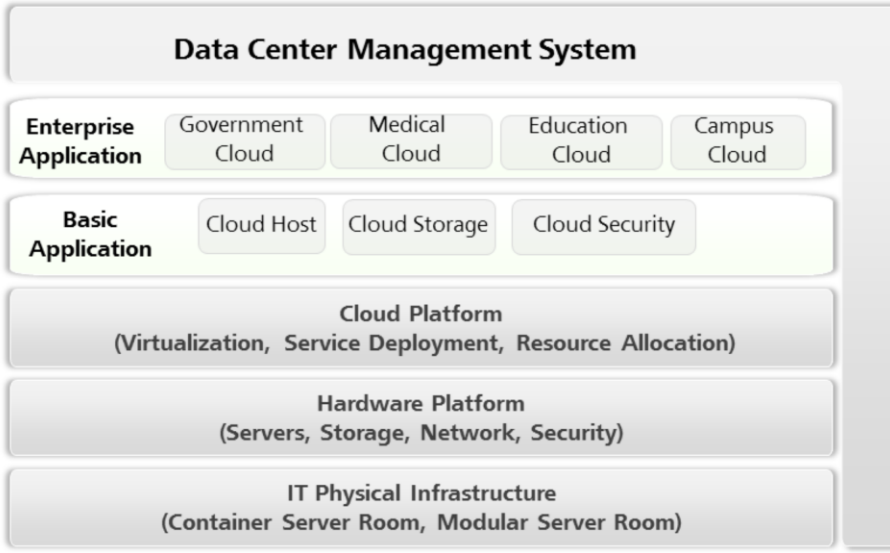
Service Driven Cloud Service

- 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.

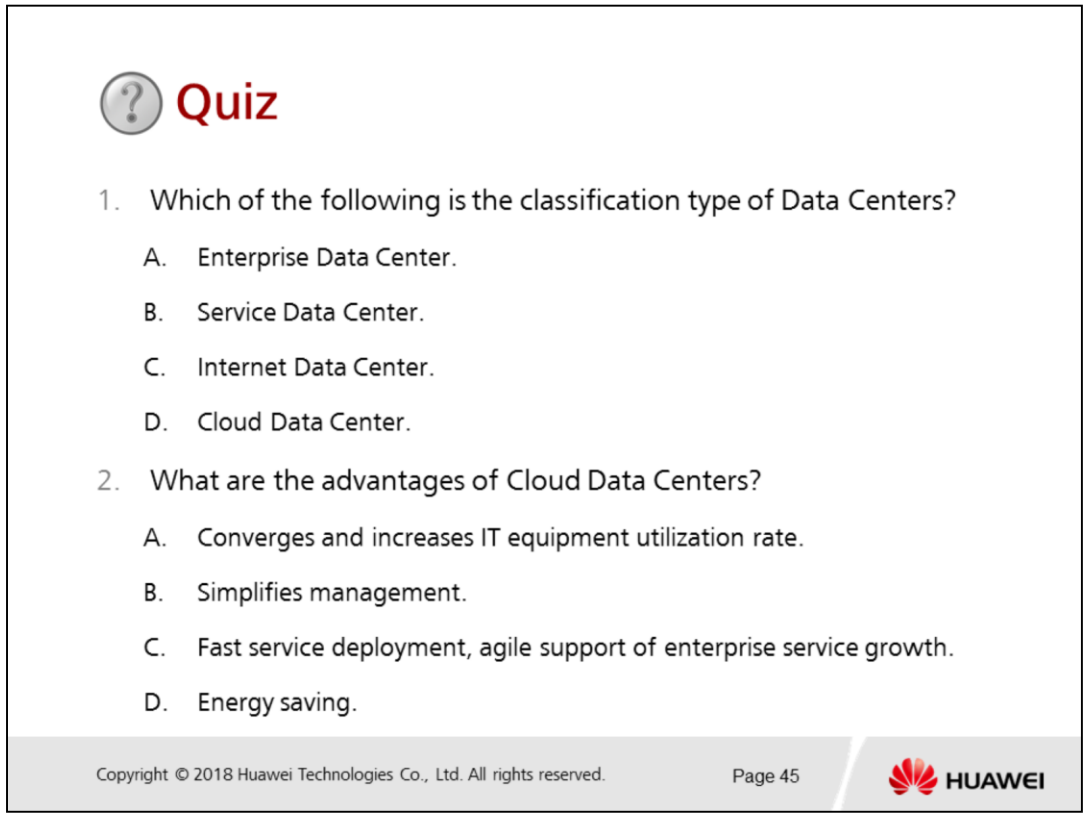


HUAWEI

Next Generation Data Center Architecture



HUAWEI



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

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.

Thank You

www.huawei.com