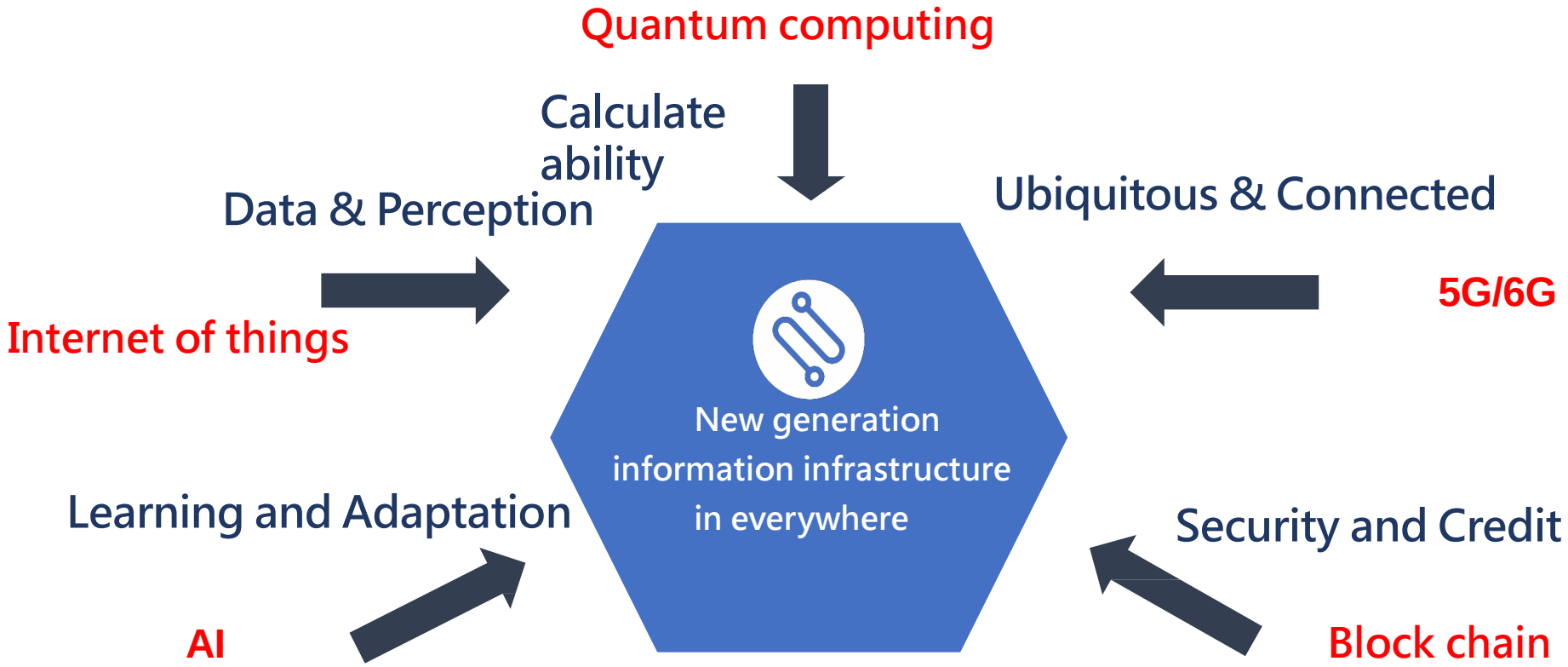
Edge Cloud Infrastructure for the future network

YUAN ZHANG
Chinatelecom
2019-5-21

Technology convergence promotes the next-generation information infrastructure



Network Evolution: Internet - > Internet of Things - > Value Network

Internet

Information sharing

—connection & sharing

Before 2005

Portal for e-commerce and web

Web 1.0

Today

User relationship & Data monopolist

Web 2.0

Internet of Things

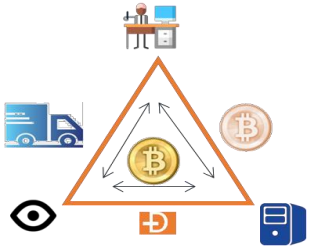
Data exchange —data & perception

IoT connectivity will be more complete, and data collection will be more real-time with the boost of 5G

Value Network

Value transfer —credit & exchange

The distributed and autonomous globalized economic era based on "trust" will gradually emerge



Architecture Evolution: Centralization -> Decentralization

From 1940s to 1990s

Mainframe

1990s-2000

Distributed cluster

2002

CDN

P2P/Grid

Seeking development opportunities in streaming media field from 2006 to 2016

Cloud

Edge computation

Rapid development since 2008

Block chair

It has gradually developed and innovated since the concept formation in 2016.

Hadoop

It was proposed in 2004 and achieved a scale of production in 2011.

Centralized architecture

Decentralized architecture

New business promotes the evolution and development of DC technology

- **Data center transforms from "resource carrier" to "service carrier"**

Data center for rental of basic resources

- Rental of infrastructure resources

Hosting data center

- IT infrastructure services

Cloud data center

SaaS

PaaS

IaaS

- Large, virtualized, integrated, and on-demand cloud services

Converged data center

- Clouding platform , resource heterogeneity, cloud edge collaboration
- Ecologicalization and fragmentation of business
- Data-oriented operation and intelligence
- Trade-off between Reliability and availability
- Green technology

Decentralization mode will push edge DC into a new form of DC

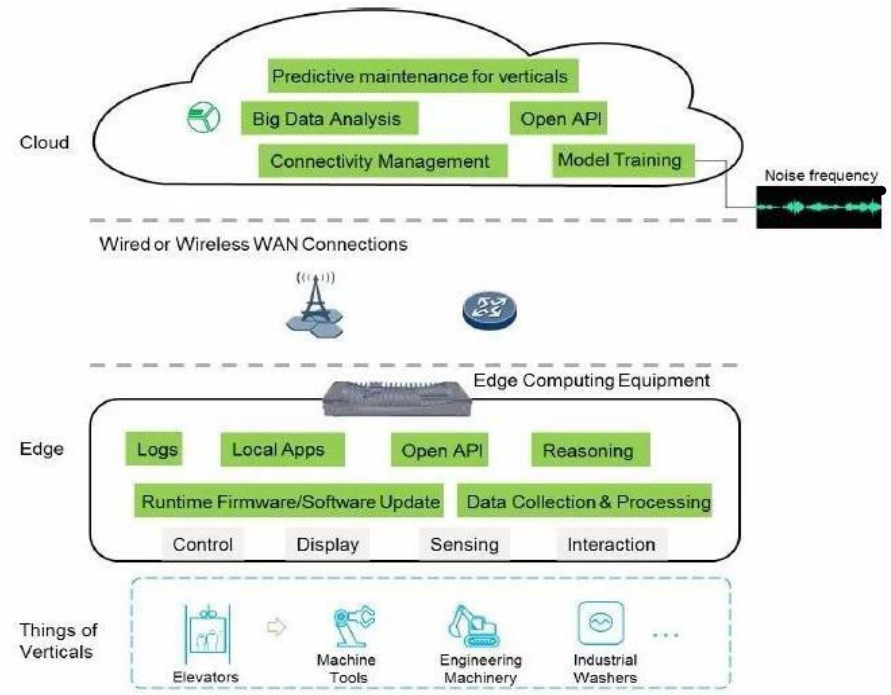
The edge of the infrastructure from the carrier's perspective

➢The IT resources of network operators or service providers located in the "last mile" are mainly constructed by edge DC.

&

The edge of the device from the end user's perspective

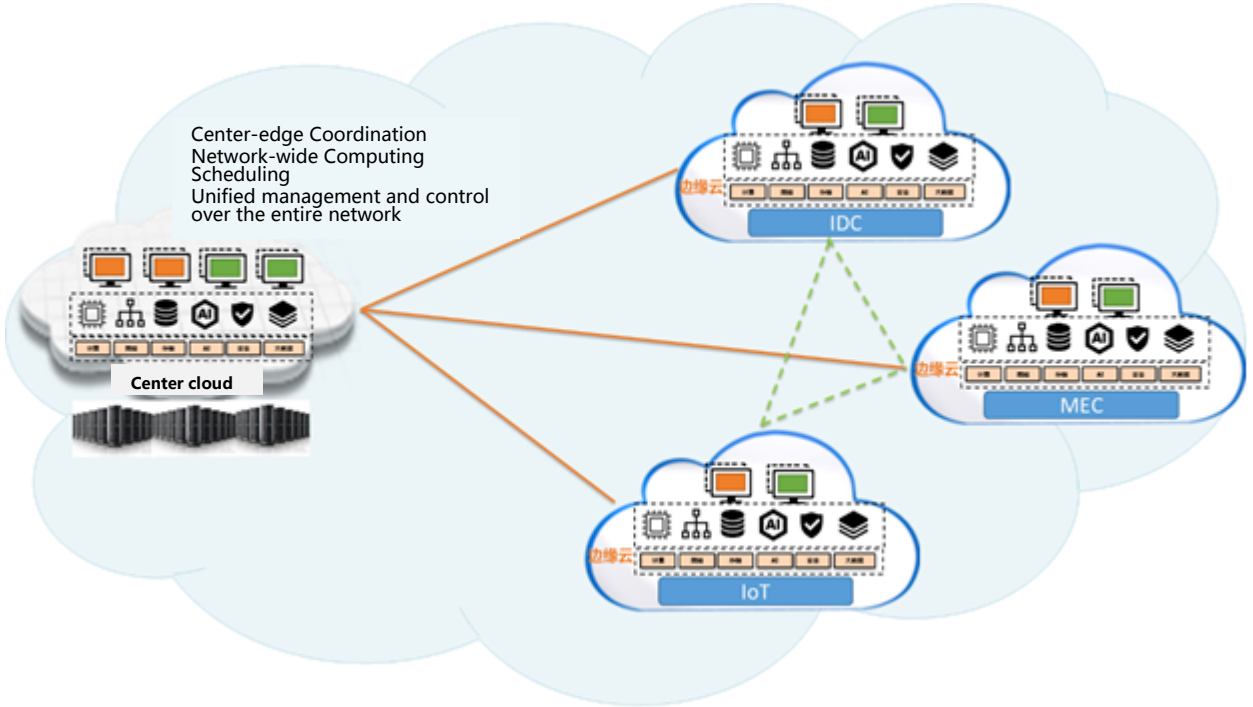
➢Edge computing resources on the network terminal or device side, including traditional Internet devices (such as mobile phones) and new smart devices (such as smart cars)



Edge DC emphasizes the distribution of nodes rather than centralization, which is highly complementary to cloud computing data center.

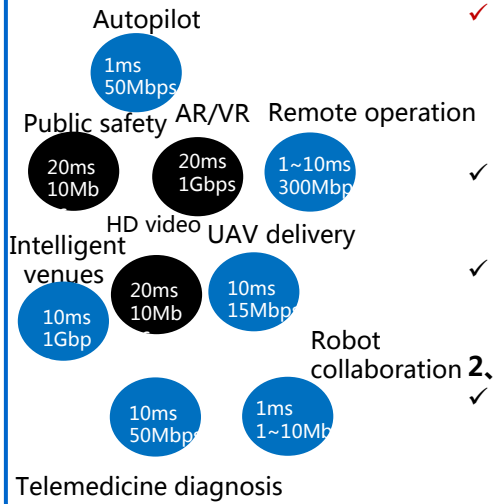
Edge cloud is the extension from center cloud DC to edge DC

- Edge cloud extends some services or capabilities of the cloud (including but not limited to storage, computing, network, AI, large data, security, etc.) to edge DC.
- Center DC and edge DC cooperate with each other to realize the "ubiquitous" cloud.
- Solve the problem of long latency and bandwidth occupancy caused by centralized deployment of cloud computing



Edge cloud from the perspective of service

New service in the future



Characteristics of the service

1、Massive Data

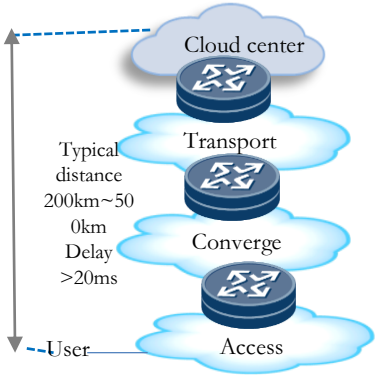
- ✓ 50 billion connected wireless devices, generating 600 terabytes of data, but only a small amount of key data in 2020, according to Cisco's forecast.
- ✓ Sensors and cameras on driverless cars produce 1GB of data per second.
- ✓ "Skynet" has installed more than 20 million high-definition cameras nationwide.。

2、Real-time processing

- ✓ Services such as autopilot and telemedicine whose data requires real-time processing and timely response.

Problems

The existing cloud is difficult to meet the low-latency and real-time processing requirements of massive data services.



Trends

Application localization

"low cost"

Content distribution

"Large bandwidth"

Calculation marginalization

"Ultra low latency"

With the vigorous development of IoT, big data and big video, the existing network and cloud centralized deployment are difficult to meet the new business needs. It is an industry trend to make the network and cloud migrate to the edge with "business flow" through edge computing.

Edge computing (MEC) scenarios defined by ETSI

For service scenarios such as enterprise parks, venues, manufacturing and households, edge computing brings potential business innovation while satisfying and optimizing experience needs.

Video optimization

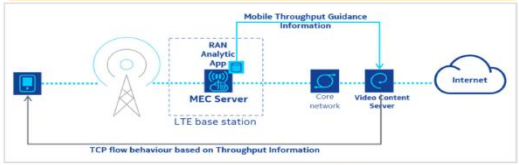


Figure 1: Intelligent Video Acceleration

Video stream analysis

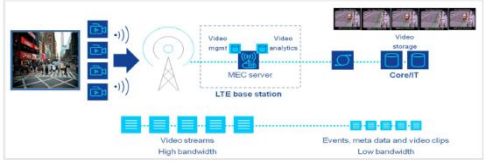


Figure 2: Video stream analysis

Internet of vehicles

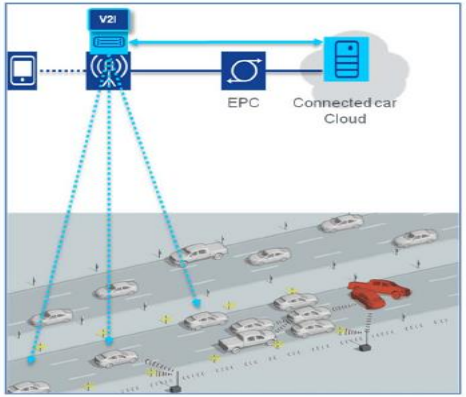


Figure 5: Connected vehicles

Augmented reality

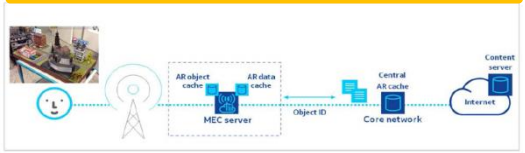


Figure 3: Augmented reality

Assisted Sensitive Computing

Edge applications provide high-performance computing power, perform time-sensitive data processing, and feedback the results to end devices, such as intelligent robots.。

Enterprise diversion

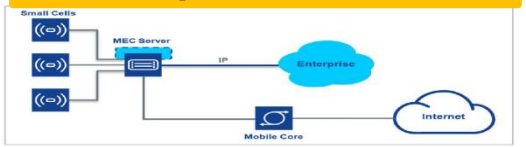


Figure 4: MEC platform with breakout to Enterprise network

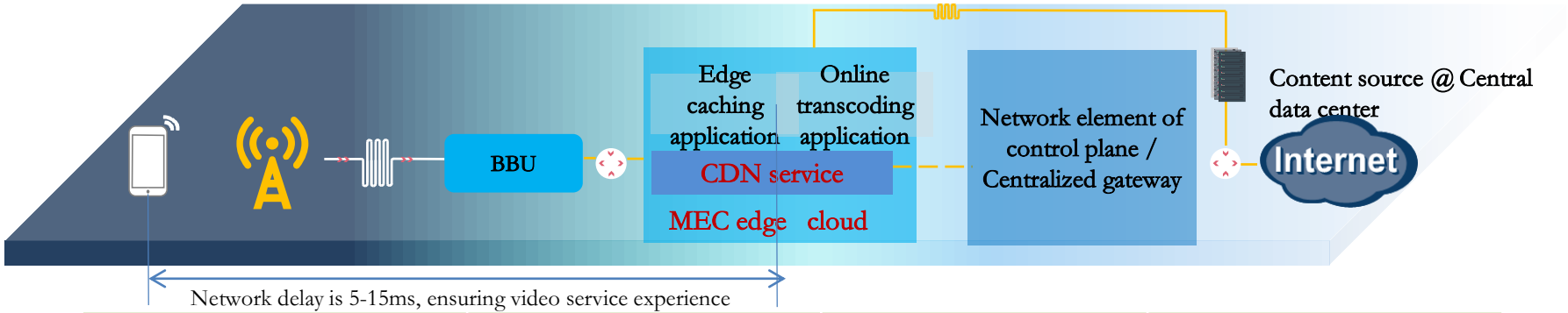
IoT/Industrial interconnection

Figure 6: IoT gateway

Scenario of the edge cloud—— CDN sinking

The CDN cache node sinks to the network edge DC, reducing network backhaul cost and central node pressure, reducing latency and improving user experience: :

- Fixed-line CDNs generally sink to the core area of each metropolitan now. With the development of large video services, there is a need for further sinking; ;
- mobile network traffic generally needs to be exported to the IDC through the provincial center of the mobile core network, and the relative delay is large, requiring the sinking demand; ;
- **sinking increases the cost of edge nodes, requiring comprehensive analysis of experience, cost, and benefits ;**



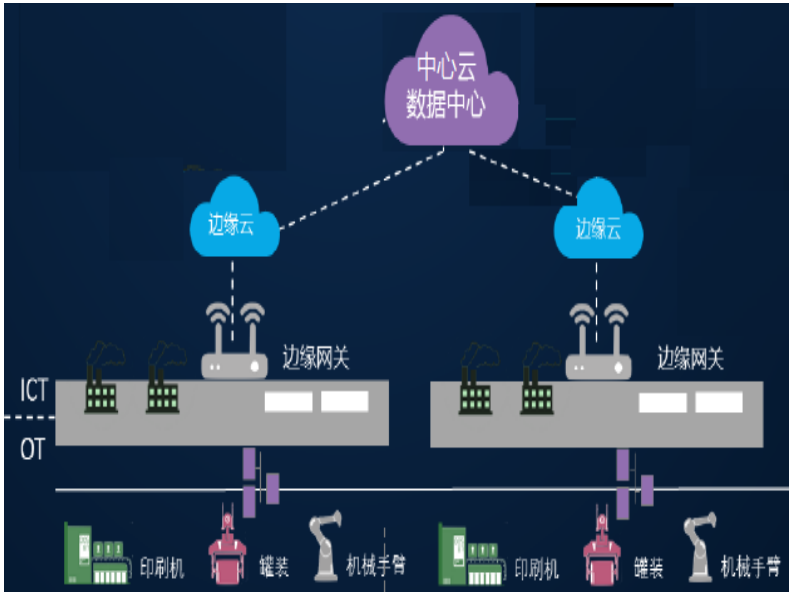
KPI	HD video(720p)	HD video(1080p)	VR video
Increase of video download rate	14.71% ↑	47.41% ↑	7.76% ↑
Decrease of round trip delay	19.04% ↓	32.58% ↓	27.10% ↓
Decrease of dragging wait delay	61.18% ↓	14.65% ↓	12.12% ↓

Note: this is the preliminary test result and further testing and verification is ongoing;

Scenario of edge cloud-industrial internet

In intelligence industrial manufacturing industry, The factory intelligence is implemented on the edge.

- An example: Taking the quality inspection of intelligent machine vision in industrial manufacturing , making real-time analysis based on product images captured by the camera on the production line, detecting product defects and adjusting the parameters of control manipulator in real time ;
- Edge computing cooperates with cloud computing : the central cloud is responsible for AI model training, and the factory edge performs local reasoning on the trained AI model



Cloud edge collaboration
in intelligent industrial manufacturing

The requirements of edge cloud applications for future network

① **Intelligent connection capabilities**

② **Intrinsic security capabilities**

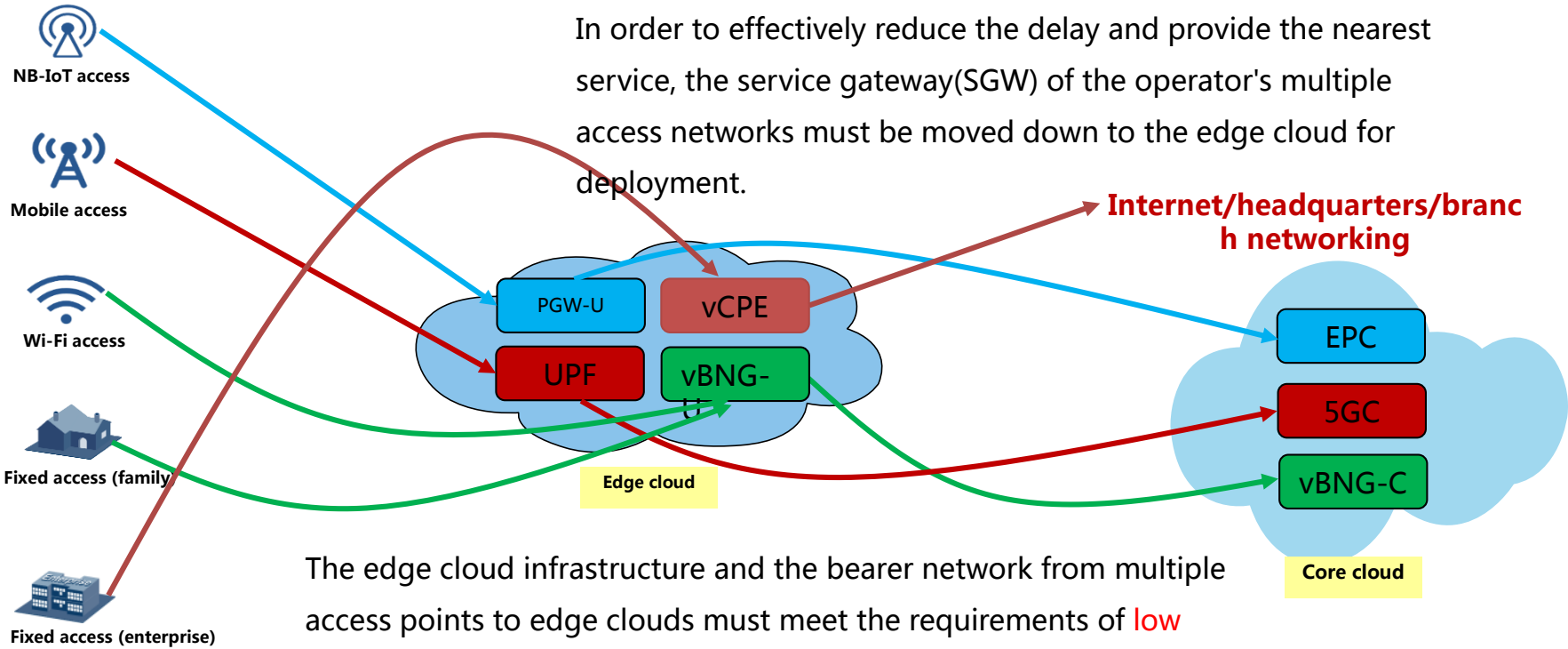
③ **Cloud & edge collaboration capabilities**

④ **Smart operation & maintenance capabilities :**

⑤ **Low-cost networking capabilities**

The impact on future networks 1: SGW of edge cloud will move close to the user

In order to effectively reduce the delay and provide the nearest service, the service gateway(SGW) of the operator's multiple access networks must be moved down to the edge cloud for deployment.



The edge cloud infrastructure and the bearer network from multiple access points to edge clouds must meet the requirements of low latency, high bandwidth, security and stability for various edge computing applications.

The impact on future networks 2: Edge server towards customization

- Edge nodes have strong physical constraints in space and power supply, load bearing, etc,the deployment of edge cloud infrastructure has to meet the performance requirements of the business on the premise of limited space, power resources, and load-bearing capacity.

- OTII (Open Telecom IT Infrastructure) : forming server technology solutions and prototype products of deep customization, open standards and unified norms for telecom applications in operator industry

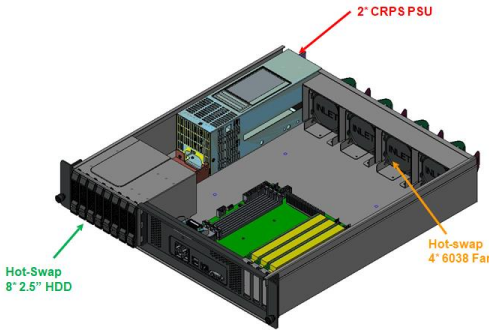
Customized server scheme for edge room

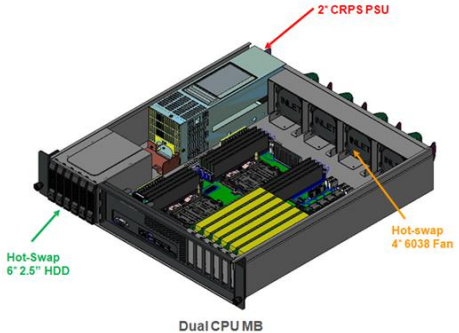
■ **Edge room server of single OTII**

Applicable to scenarios where network traffic is mainly forwarded, and power consumption and space constraints are strict.

■ **Edge room server of dual OTII**

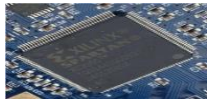
Applicable to scenarios with high CPU and memory requirements, and good room conditions.





Dimensions= 465mm(D) x 438mm(W) x 87mm(H) = 19.7 " x 17.1" x 3.4"

The impact on future networks 3: Edge acceleration to support heterogeneity

Forwarding acceleration : ASIC/NP			Suitable for data packet forwarding with simple business and high throughput	Accelerated resource pool of forwarding service
Service computing acceleration : FPGA			Programmable and suitable for infrequently changing services	Accelerated resource pool of computing service
Graphic computing acceleration: GPU				Good at parallel computing, whole data stream processing	Accelerated resource pool of graphic service
AI computing acceleration: AI chips			Typical AI algorithms	Accelerated resource pool of AI service

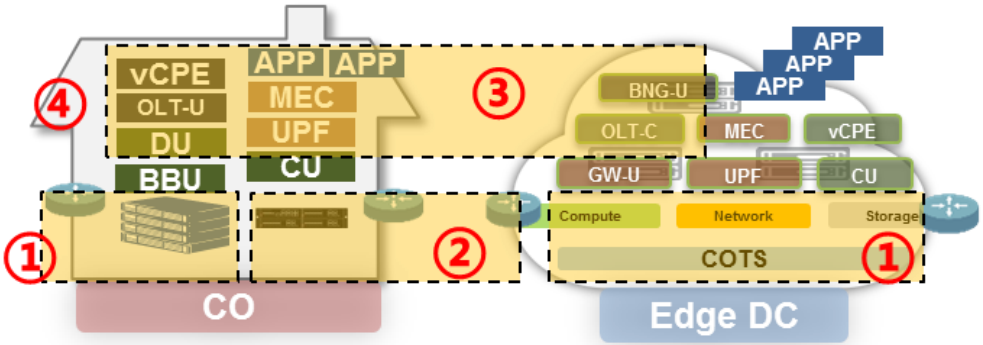
- **Edge cloud drives the evolution of network architecture**
  - It is necessary to consider the cooperation between the edge intra-cloud network (physical/virtual) and the extra-cloud network (access network/metropolitan area network), and the network should be optimized and evolved from the architecture.
  - Edge cloud management requires to build a unified management and control system to achieve the optimal forwarding path of network

① Infrastructure resource management of Edge DC and access site

② Network orchestration of edge cloud

③ Management of edge cloud elements (VM/container/specific)

④ Coordination between edge cloud, public cloud, network and terminal



Summary

- Decentralized cloud infrastructure will be the main development direction in the future
- The edge cloud infrastructure and the basic bearer network from multiple access points to edge clouds must meet the basic requirements of low latency, high bandwidth, security and stability for various edge computing applications.
- The diversity of edge computing services leads to the diversity of edge acceleration devices. More acceleration hardware is needed to form a heterogeneous hardware acceleration platform.
- A unified and coordinated management and control system for “edge could, center cloud , cloud network & teminal “ is needed.

中国电信 CHINA TELECOM | 5G hello 赋能未来

THANKS