HUAWEI TECH

2022 ISSUE 02

Eight Innovations in China Unicom's 2025 Network Architecture

Eastern Data and Western Computing: Building New Computing-first Networks

A Hospital's Journey to Become Intelligent and Low-carbon


HUAWEI
Building a Fully Connected, Intelligent World
Soar into the Cloud with
Huawei Intelligent Cloud-Network
Powered by
AirEngine
Wi-Fi 6
CloudEngine
Switches
NetEngine
Routers
HiSecEngine
Security Gateways

Publisher:
ICT Strategy & Marketing Dept.
Huawei Technologies Co., Ltd.

Editor-in-Chief:
Peng Yuguo

Deputy Editor-in-Chief:
Xing Jingfan

Editor:
Gary Maidment

Art Editor:
Zhou Shumin

Contributors:
Zhang Qin, Tang Xinbing, Yuan Yaru, Luo Jianggan, Zhang Fupeng, Gu Xuejun, Ding Zhibin, Dong Libin, Li Mingxiao, Nie Yi, Feng Guojie, Xing Song, Zhou Bo, Wang Yu, Zhao Shaolong, Du Wei

E-mail: HWtech@huawei.com

Tel: +86 755 89241326
Address: G1, Huawei Industrial Base, Bantian, Longgang, Shenzhen 518129, China

Publication Registration No.:
Yue B No. L015060003

Unlocking the Metaverse with Digital Infrastructure

Suddenly everyone is talking about the metaverse. The word conjures up images of a new digital world in which the physical world is mirrored, a world in which new economic systems, social relationships, and identity systems are created, where you can navigate a space of infinite possibilities with your digital avatar.

The metaverse is a concept that neatly embodies aspirations for the digital lifestyles of the future — a future in which the ideas behind *Ready Player One* become reality.

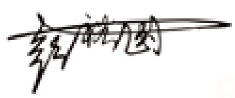
The development of the metaverse needs to be progressive and underpinned by shared digital infrastructure, standards, and protocols, and the increasing convergence and evolution of tools and platforms.

It also makes me recall Second Life, the online virtual world that was popular more than a decade ago. Second Life lacked mass appeal because the technology and digital infrastructure were not ready at the time, limiting user experience. To digitalize the real world, the metaverse needs to provide immersive experiences using extended reality technologies like VR, AR, and mixed reality. It needs to mirror the real world with digital twins powered by big data, AI, smart applications, digital platforms, cloud computing, and connectivity. And its economic systems require blockchain technology.

Enabled by digital infrastructure, the immersion, low latency, and ubiquity expected of the metaverse will require jumps in user experience, network transmission, high-performance cloud-edge computing, and streaming technology. As a more realistic metaverse will generate more data, computing and storage capabilities will determine its scale and level of completeness.

In 2021, digital technology and infrastructure like connectivity, computing, and cloud saw rapid breakthroughs and wide application. AI, 5G, gigabit optical networks, and IPv6 gathered momentum and new infrastructure like supercomputing, autonomous vehicles, and the industrial Internet continued to advance. Digital transformation accelerated in various industries, including energy, manufacturing, healthcare, finance, giving us a glimpse of the digital world ahead.

We have entered a period of rapid digitalization. In 2022 and beyond, we hope to further unlock the value of ICT in the metaverse, as innovation breaks down barriers and digital infrastructure brings us closer to our vision: To bring digital to every person, home and organization for a fully connected, intelligent world.



Peng Yuguo

Contents

Expert Voices

04
ICT in 2021: The Dawn of the Metaverse and New Digital Frontiers

10
Ultra-high-speed Lossless Networks for Exascale Computing Power

Industry Trends

14
Eight Innovations in China Unicom's 2025 Network Architecture

23
Five Trends Shaping Next-gen Data-intensive Supercomputing

38
Native Hard Pipe Networks: An Optical Foundation for Industries to Go Digital

New Solutions

28
Eastern Data and Western Computing: Building New Computing-first Networks

Use Cases

46
A Hospital's Journey to Become Intelligent and Low-carbon

42
How Intelligent All-optical Commercial buildings Enable SME Digitalization

34
All-optical Sensing Brings Intelligent Automation to Oil & Gas Pipeline Inspections

51
AirEngine AP & IdeaHub: The Next Step in Collaborative Digital Offices

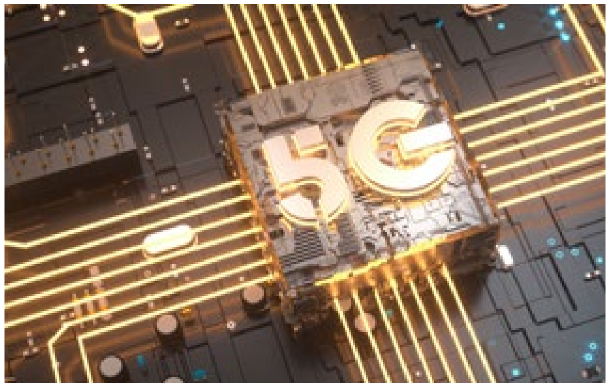
ICT in 2021: The Dawn of the Metaverse and New Digital Frontiers

In 2021, the metaverse took center stage, promising a future where digital lifestyles and the physical world will intertwine. Despite current infrastructural immaturity and a lack of standards, tech giants are scrambling for a foothold in the virtual future. What can ICT developments in 2021 tell us about the future?

By Peng Yugguo, Huawei & Xing Jingfan, Huawei

Regarded by many as next-generation Internet and the successor to today's mobile Internet, a battle of imagination is underway as companies seek opportunities in the metaverse. It may be early days, but what is certain is that the emergence of the metaverse will accelerate the development of ICT infrastructure.

What we can learn from 2021



1 The continuing evolution of commercial 5G

Over the next five years, China is expected to transition from the introductory phase of 5G to one of large-scale expansion. 5G standards will also continue to evolve.

On November 13, 2021, UNISOC announced at the E-surfing Smart Ecosystem Expo that, based on China Telecom's 5G SA network, it had completed testing E2E 5G network slicing services on the cloud phone e-Surfing One 2021, indicating that 5G network slicing was technically ready for commercial use. Network slicing can be viewed in a similar way to dedicated bus lanes on roads – each subscriber to the slicing service receives a dedicated line that ensures smooth, jitter-free access for things live streaming, watching videos, or gaming.

At the same expo, Qualcomm demonstrated 8K video backhaul using the 5G mmWave large uplink frame structure, achieving an uplink rate of 930 Mbit/s. 5G mmWave technology is regarded as the foundation of the ultimate 5G experience and its technical characteristics make it best suited for large open spaces. China Unicom is blazing a trail, providing 5G mmWave connectivity for venues at the Beijing Olympic Winter Games. In Zhangjiakou, where temperatures sometimes dip below minus 30 degrees Celsius, the operator has conducted the industry’s widest ranging tests of mmWave network-device access to date. Test results show that the mmWave solution covers 800 MHz of bandwidth, with a downlink rate of more than 4 Gbit/s and an uplink rate of over 1 Gbit/s.

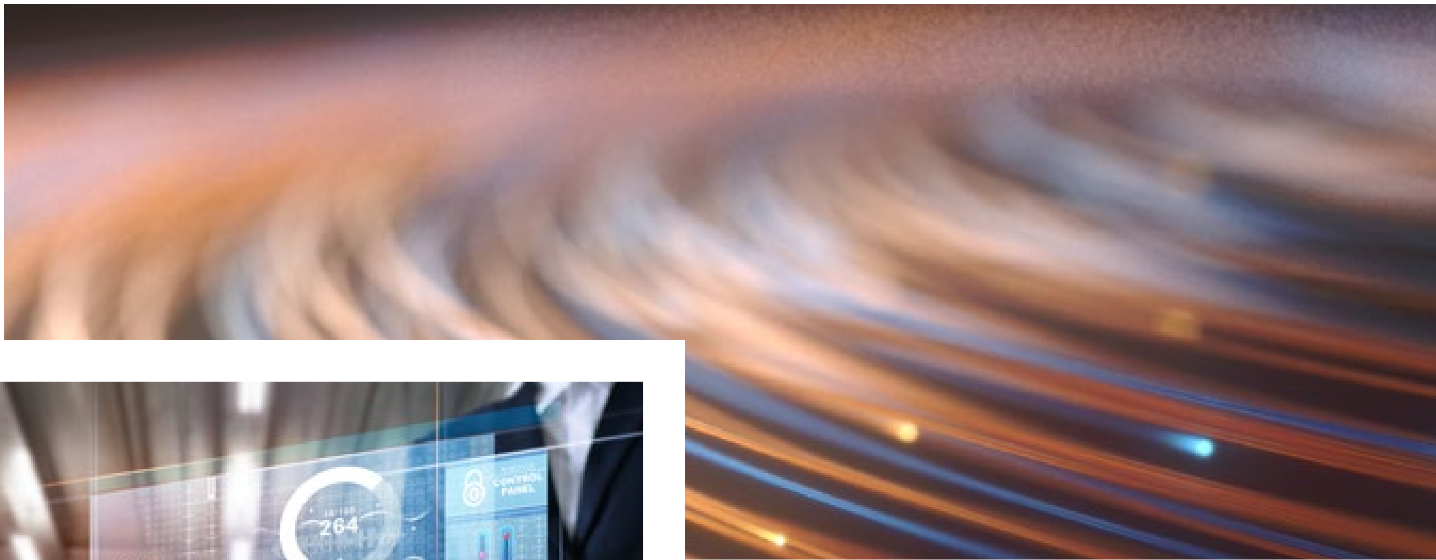
At Xiamen Yuanhai Container Terminal, the combination of a 5G network, the BeiDou satellite navigation system, and driverless technology powers autonomous container trucks and automated rubber-tired gantry cranes. Few workers are seen onsite and the trucks do not require a cockpit.

To support new types of services and the future development of networks, 5G standards have continued to evolve and 5G-Advanced standards are expected soon. On December 22, 2021, China Mobile and Huawei held the Launch Event of Wireless Innovation in 5G-Advanced Dual-chain Convergence. They announced the results of their research into 5G-Advanced (5.5G) dual-chain convergence, forecasting a tenfold increase in the capabilities of existing 5G networks and new capabilities that will meet diverse consumer demand.

2 Accelerated deployment of gigabit optical networks

The ultra-high bandwidth of 5G has brought a smooth mobile network experience to consumers, while wired networks still dominate home scenarios. In 2021, China's Ministry of Industry and Information Technology (MIIT) issued its *Action Plan for Coordinated Development of "Dual-gigabit" Networks (2021-2023)*, which sets out the specifications and direction of broadband development in China. In the same year, the MIIT issued the *Information and Communications Industry Development Plan for the "14th Five-Year Plan"*, which calls for the acceleration of extended deployment of optical transport network (OTN) equipment to integrated access nodes and users.

Aligned with these policies, China Unicom Beijing announced its plan to be the first carrier to offer FTTR gigabit-level connectivity in Beijing using the only networking technology that can support 1,000+ Mbit/s whole-house coverage. Optical fibers can be deployed in every room, closing the gap between the access rates of the network and what users actually experience. This product is especially suited for large apartments, duplexes, and detached houses, representing



the best choice for high-end residential network access.

Following this, China Telecom and China Mobile launched FTTR gigabit optical network solutions in 11 and 10 provinces and cities, respectively. To date, telecom infrastructure companies have deployed 5.5 million 10-Gbit/s passive optical network (10G-PON) ports, and more than 220 million households now have access to gigabit optical networks.

In addition to providing higher-quality experience to homes, gigabit optical networks underpin enterprise digitalization. Currently, there is great demand for premium OTN private lines among enterprise users in a number of different industries. By 2025, 85% of enterprises will have migrated to cloud, and will in turn require high-quality and high-bandwidth networks.

The commercial application of F5G in the primary sector has been accelerating. In the October 2021 release of the Regulations on Intelligent Coal Mine Acceptance (Trial), the National Energy Administration encourages the use of F5G in the wired backbone networks for mines. F5G has also made its way to rural areas, with AirPON deployed in more than 3,000 villages across the country. The solution benefits more than 10 million villagers and has enabled new business models and ecosystems to thrive, including digital husbandry, digital tourism, and digital orchards.

3 Deepening cloud-network convergence

As commercial 5G and enterprises accelerate migration to cloud, the industrial Internet is rapidly developing.

On March 23, China Unicom released the CUBE-Net 3.0 network innovation system in Beijing. CUBE-Net 3.0 features stable architecture, all-optical switching, scalable high bandwidth, and one-hop cloud access for all services. Services can be flexibly migrated to cloud in multiple ways, including PON, IPRAN, and OTN, while automated O&M management and E2E slicing allow a single network to support multiple differentiated services.

At the 2021 International Digital Technology Exhibition & E-surfing Smart Ecosystem Expo, Ke Ruiwen, Chairman of China Telecom, spoke of the need to build digital information infrastructure that can deliver high speeds, ubiquitous access, wired-wireless integration, cloud-network convergence, intelligent agility, energy-efficiency, and controllable security.

This infrastructure will serve as the foundation of a digital China and facilitate digital government. And its most prominent characteristic will be cloud-network convergence.

For cloud computing service providers, cloud-network convergence can boost service capabilities and quality. Before they upgrade their data centers, the networks that connect cloud data centers to each other and to users need to be upgraded. For network operators, cloud-network convergence can lower procurement and network deployment costs and facilitate service innovation.

4 Deeper IPv6 transition

5G, cloud computing, big data, AI, and IoT applications will all be IPv6-based in the next decade.

By late September 2021, IPv6 accounted for 22.87% of all traffic running on mobile communication networks in China, reaching the set target of 20% ahead of schedule and indicating that IPv6 in China has advanced to the extent of increasing traffic. As the *Three-year Special Action Plan for Increased IPv6 Traffic (2021-2023)* issued by the MIIT and Cyberspace Administration of China (CAC) is implemented, the ICT industry will address issues like the limited adoption of IPv6 in commercial Internet applications, IPv6's lack of support for home devices, and the inadequate service performance of IPv6-based application infrastructure.

Huawei's IPv6+ technology, which has been developed on top of IPv6, enhances network intelligence capabilities and supports the intelligent identification, analysis, control, and O&M of network connections. Incorporating technologies like SRv6, network slicing, on-flow detection, innovative multicast, and application awareness, IPv6+ can bring enterprises the benefits of deterministic networks, quality assurance, application awareness, service visibility, and controllable security.

5G and cloud computing have triggered a new wave of new infrastructure construction that will sit on the foundation provided by IPv6 and related standards. IPv6 is a major area for global competition in next-gen commercial Internet application solutions and Internet evolution.

5 AI as infrastructure

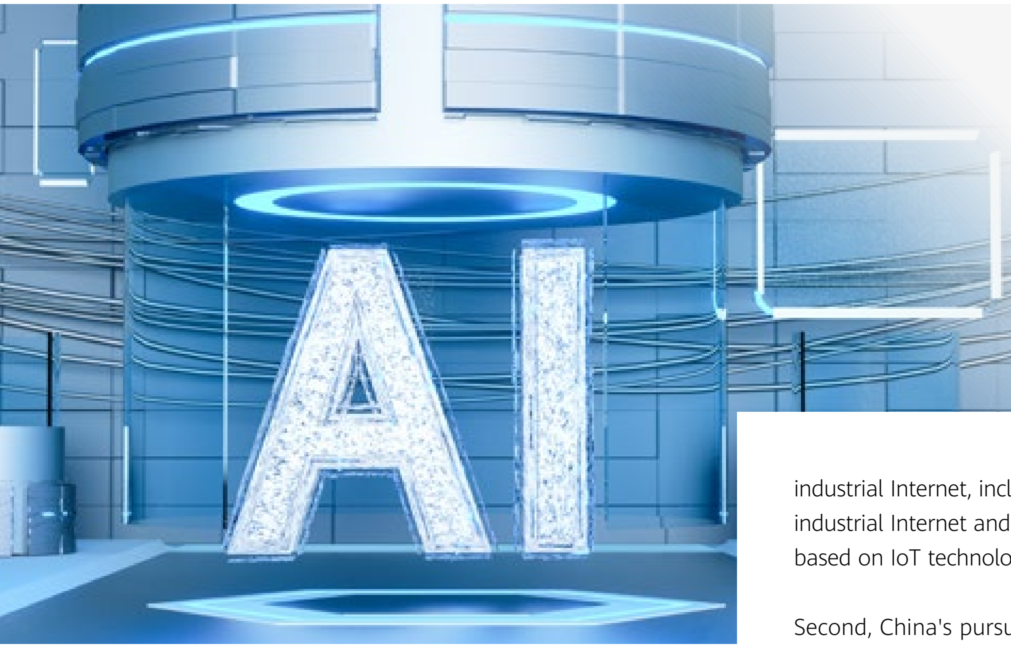
5G, cloud, and AI are regarded as the three engines of the modern ICT industry. Integrating AI into digital infrastructure and implementing national AI strategies are both clear global trends.

In April 2021, Huawei released the Pangu Model. The Pangu NLP Model represents the industry's first model of Chinese natural language processing (NLP). To train it, the development team used 40 TB of text data, including general and domain-specific text. Featuring hundreds of billions of parameters, it ranks top of the total Chinese Language Understanding Evaluation (CLUE) index. The Pangu CV model is the first solution to extract models on demand, where different-sized models can be extracted depending on the scenario. The scope can be dynamically adjusted on demand, ranging from specific and small to complex and large. The solution also supports comparative learning based on sample similarity, making it the industry leader in terms of few-shot learning on the ImageNet database.

The Pangu Model can help solve the problem of standardization and makes AI capabilities available on-demand to SMEs, so they can focus on innovation and their core businesses.

6 Open-source operating systems for digital infrastructure

In addition to basic AI models, the ICT industry has also seen major breakthroughs in basic software, at the core of which sits operating systems (OS). The OS also serves as the core of computing systems and determines the level at which digital infrastructure can advance. At the Operating System Industry



Summit 2021, Huawei announced that it would donate OpenEuler to the OpenAtom Foundation. OpenEuler is an open source operating system that supports the full range of digital infrastructure scenarios, including IT, CT, and OT. It can be deployed on servers, cloud computing, edge computing, embedded systems, and various types of devices.

The donation means that OpenEuler has evolved from a Huawei-led open source project to a more decentralized project, which many developers and coders can contribute to and which is independently managed by the community.

ICT trends in 2022

In 2022, the growing demand for digital innovation will drive the upgrade of digital infrastructure in three main ways:

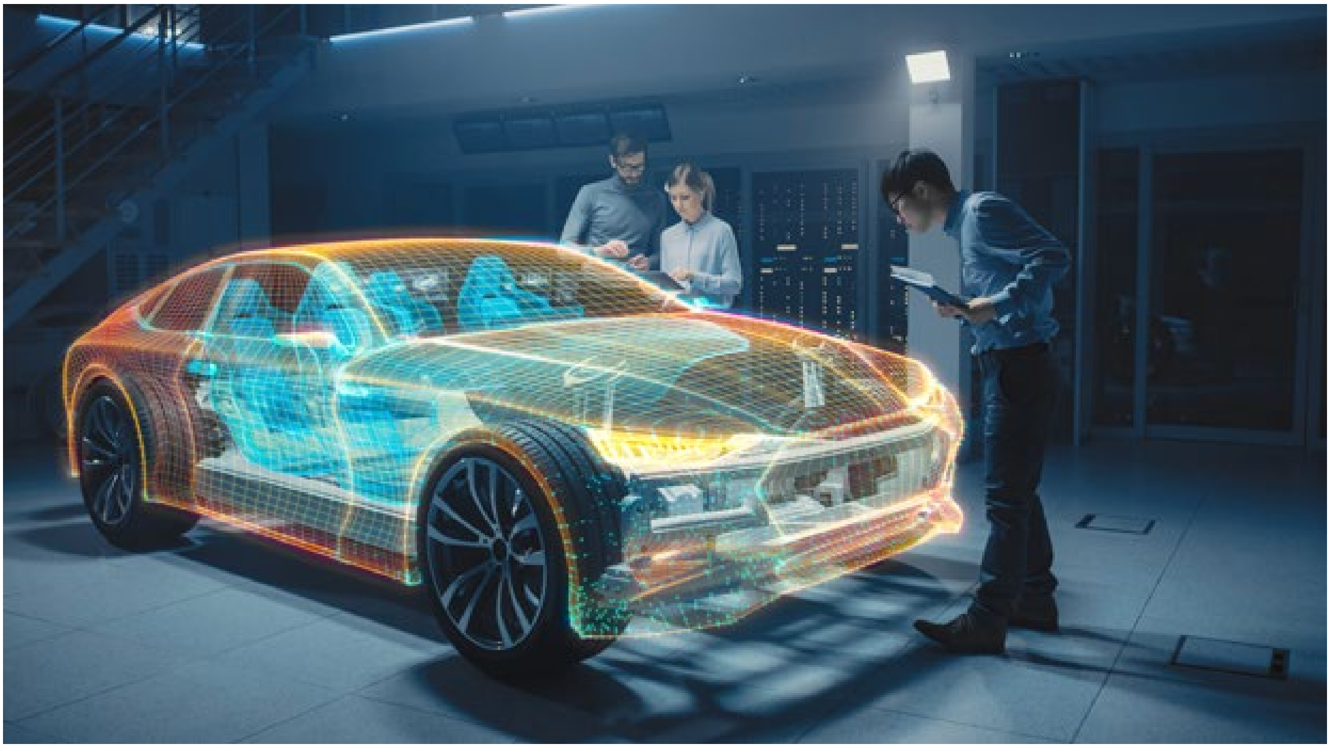
First, China's industrial Internet will develop at a faster pace. To date, four national-level industrial Internet industry demonstration bases in China are being constructed. A total of 258 pilot demonstration projects have been selected, and 31 provinces, autonomous regions, and municipalities have issued industrial Internet development policies. This requires the ongoing optimization of digital technologies to meet the needs of the industrial Internet, including communication protocols for the industrial Internet and data ingestion and monitoring systems based on IoT technologies.

Second, China's pursuit of carbon peak and carbon neutrality goals will accelerate national energy transformation. Since China announced its targets of reaching peak carbon by 2030 and achieving carbon neutrality by 2060, major enterprises in the energy industry have drawn up energy transformation roadmaps and action plans. The ICT industry needs to support this by developing technologies to help the energy industry improve efficiency and cut emissions, as well as making efforts to minimize the emissions of the ICT industry itself.

Third, the wide adoption of ICT has made the security of critical information infrastructure a key issue. In 2021, a number of laws and regulations were adopted to ensure the security of information infrastructure. These included the Regulations on Critical Information Infrastructure Security Protection, Data Security Law, and Personal Information Protection Law. As the ICT industry creates the underlying infrastructure for the digital development of other industries, zero-trust mechanisms must be adopted to ensure security. Therefore, the implementation of data security regulations in various industries will become a major area of focus.

In conclusion

Huawei is committed to accelerating the inception of new information infrastructure and application ecosystems, supporting the digital transformation of the economy and society, and driving the ICT industry to support global development.



Ultra-high-speed Lossless Networks for Exascale Computing Power

The increasingly digital world demands tremendous computing power. While today's exascale supercomputing capabilities are 1,000 times as powerful as the computer that powered AlphaGo, computing on this scale has long been held back by packet loss issues with interconnect technologies such as Ethernet.

Huawei's R&D efforts have resulted in a breakthrough that has greatly improved the performance computing and storage networks: fully lossless Ethernet.

By Li Xinyuan, Huawei

At Huawei Connect 2021, Wang Lei, President of the Data Center Network Domain of Huawei Data Communication Product Line, invited customer experts to discuss how Huawei's hyper-converged data center network can help improve supercomputing power and artificial intelligence. The experts also shared their insights into and visions for supercomputing applications.

In Liu Cixin's science-fiction novel *The Three-Body Problem*, "the world's most powerful computers," used to simulate nuclear explosions, "can perform 500 trillion floating-point operations per second." That seems like an incredibly huge number. However, on July 1, 2021, China's Pengcheng Cloud Brain II topped the IO500 ranking at the World Supercomputing Conference, exceeding one quintillion floating-point operations per second (FLOPS), 2,000 times faster than the most powerful computers in *The Three-Body Problem*.

Real-world supercomputing has gone beyond the imaginings of science fiction.

Q

Huawei Tech: We know that supercomputers have tremendous computing power. In layman's terms, what does that mean?

Wang Lei: Exascale computing is truly powerful. Technically speaking, it is 1018 FLOPS, equivalent to what half a million of our latest laptops can do combined. AlphaGo, the AI that beat Lee Sedol at Go a few years ago, was powered by a petaflop-level [more than one quadrillion] FLOPS supercomputer. Today's exascale computers are 1,000 times as powerful as the computer that powered AlphaGo.

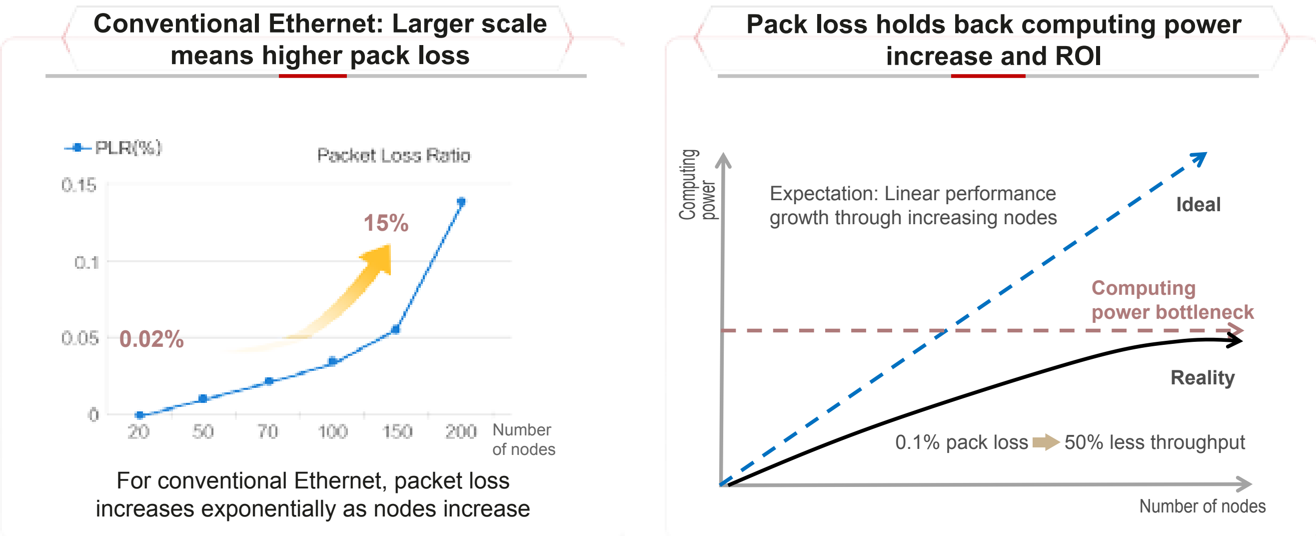
Q

Huawei Tech: Is building a supercomputer simply a matter of stacking many computers together?

Customer 1: The consensus in the supercomputing world is that "1 + 1 < 2". For example, when we bind 100 computers together, do they deliver 100 times the computing power of a single computer? The answer is definitely no. For two connected computing units to work together, complex coordination is required, and the underlying condition is ensuring unimpeded communication between them. Lossless communication channels mean low latency and high efficiency for communications. Huawei's lossless network was developed to solve these kinds of communication problems.

Wang Lei: When we say "supercomputer", we are referring to a super computing cluster. It appears to be just a bunch of computing units clustered together, but the underlying super-fast network connecting them is what enables them to run at high speeds. Therefore, a supercomputer is not just a stack of regular computers, but an integrated system supported by an ultra-high-speed lossless network.

Customer 2: Both supercomputing centers and AI computing centers need high-speed



Until now, supercomputing clusters that relied on Ethernet interconnect have always suffered from packet loss

networks and software to centrally manage physically scattered computers to form a logically unified computing cluster which serves as a computing resource pool that can be used on-demand.

Q

Huawei Tech: Ethernet is inherently prone to packet loss. What led Huawei to believe that it could solve a problem that has been around for 40 years? Has Huawei overcome this challenge?

Wang Lei: Huawei overcame this technical challenge two years ago. Back when we just started investment and research in computing, our researchers found that simply binding servers together could not create a linear increase in computing power. For example, they found that doubling the number of GPU servers increased computing power by just 4%. Through analyzing the computing process, we found that the problem was caused by packet loss, an inherent problem with conventional Ethernet. Packet loss of just 0.1% can result in computing power loss of 50%, meaning half the server computing power is wasted. To address this, Huawei began examining the question of how we might create lossless Ethernet networks. We finally solved this problem two years ago, thereby realizing 100% utilization of servers' computing power.

Q

Huawei Tech: The digital economy is said to have entered the computing era. But will supercomputing have any impact on the daily lives of ordinary people?

Wang Lei: For most people, the whole topic of supercomputing seems like something very remote, because this technology is mainly used in relatively high-level applications such as weather forecasting, earthquake monitoring, and human genetic testing. However, supercomputing is much closer to our daily lives than most of us are aware of.

For example, in recent years, it's played a role in increasing the variety of new, affordable cars. Vehicle crash testing is one of the most time- and investment-intensive processes in automobile manufacturing. Using physical vehicles for testing means each crash results in a scrapped testing vehicle, and the cost can add up to millions of RMB. However, using supercomputers to simulate crash tests can shorten the development cycle of new cars from 36 months to 12 months. Now, with Huawei's hyper-converged data center network, that process can be expedited even further.

Customer 1: We can look at supercomputing and its implications from the public health perspective. In the early days of the pandemic, we didn't understand COVID-19 so well. Through extensive analysis, it was later found that the cytokine storm was an important factor increasing the morbidity rate. Supercomputing played a major role in the process of deepening our knowledge. Scientists and doctors working together found that the overreaction of the human immune system to the invading virus affected certain normal bodily functions and led to the failure of those functions. With the support of supercomputing, a way to cut off the cytokine storm signal pathways was discovered. This knowledge was put to good use in Wuhan, where it saved lives.

Customer 2: If we compare AI computing power to electric power, the AI computing center is like a large-scale power station. AI applications, like electricity, will be widely used in numerous industries and households. The use of AI will make urban management more precise – self-driving vehicles and license plate recognition are examples of AI in our daily lives and urban management.

A more powerful future

Albert Einstein once said, "We cannot solve our problems with the same level of thinking that created them." The history of humanity has been a long process of creating and solving problems. Our understanding of supercomputers and AI will become clearer as technology advances, and technological innovation in supercomputing will play a key role in this process. We believe that Huawei will provide high-quality computing infrastructure for scientific research in all kinds of industries and key fields, powering economic growth and social development, and enabling everyone to step into the computing era.

Using supercomputers to simulate crash tests can shorten the development cycle of new cars from 36 months to 12 months. Now, with Huawei's hyper-converged data center network, that process can be expedited even further.

Eight Innovations in China Unicom's 2025 Network Architecture

China Unicom has proposed its CUBE-Net 3.0 network transformation strategy for network architecture innovation and development for 2025, reflecting its commitment to building next-generation digital infrastructure to underpin the digital economy.

By Tang Xiongyan, Chief Scientist of China Unicom, Vice President of China Unicom Research Institute, and Chief Leader of CUBE-Net 3.0 Technical Architecture

In March, 2021, China Unicom released a series of white papers on next-generation digital infrastructure, or CUBE-Net 3.0, and set up a demonstration base in the Guangdong–Hong Kong–Macao Greater Bay Area to promote innovation.

Key innovations of CUBE-Net 3.0

The CUBE-Net 3.0 technical architecture defined by China Unicom (Figure 1) consists of eight parts:

1. Ubiquitous and elastic mobile broadband

With the steady development of 5G networks and applications, 5G services are becoming increasingly differentiated, convergent, and diversified. However, uplink bandwidth remains a bottleneck.

Mobile broadband communications need to evolve towards higher spectrum bands and bandwidth. Due to the massive amounts of connections, a large number of low- and medium-rate services will coexist with services that require large bandwidth for a long time. As mobile broadband requires more intelligent, flexible, and elastic bandwidth management, it will become ubiquitous, elastic, intelligent, and green over the next decade.

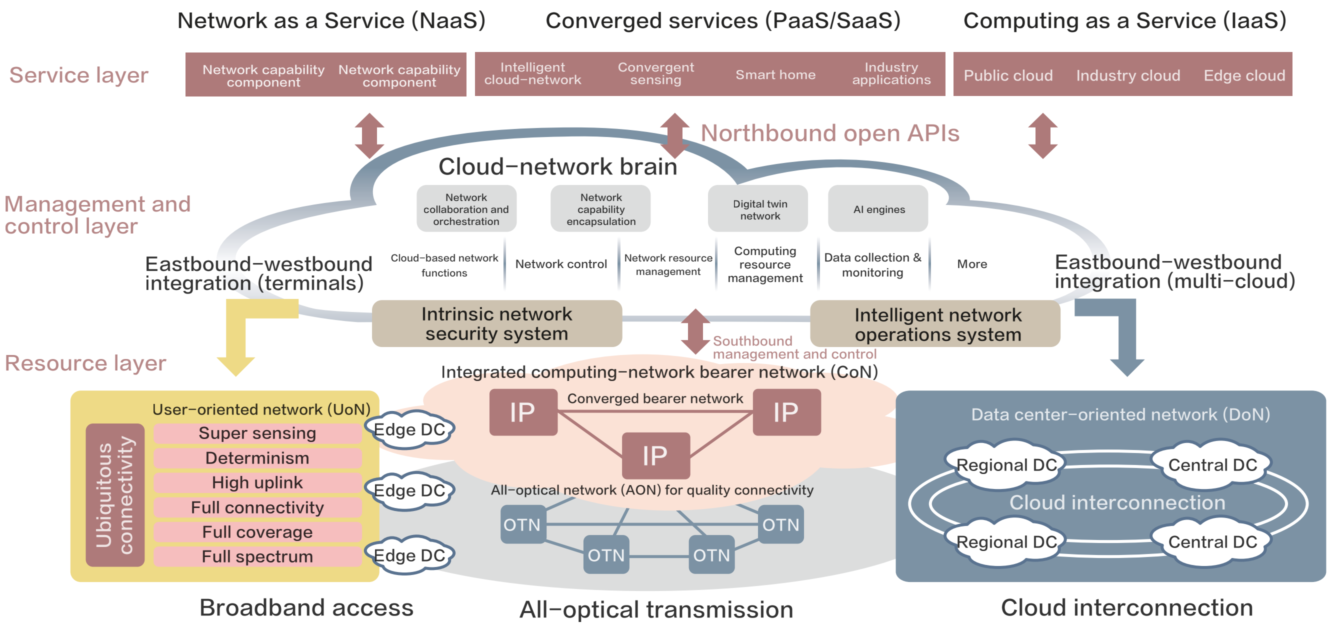
Flexible bandwidth: Full-spectrum 5G reconstruction will lay the foundation for the continuous evolution and development of 5G networks. To achieve that, three major requirements must be fulfilled: First, driving evolution towards operating bandwidth above 1 GHz with large single-carrier bandwidth above 400 MHz in millimeter wave and terahertz bands. Second, creating spectrum resource pools where cross-band resources are managed in pools by coordinated on-demand access to high-, medium-, and low-frequency bands. Third, matching bandwidth resources to users' service demands through intelligent bandwidth resource allocation.

Flexible timeslots: High uplink and downlink demands for B2B and B2C services are completely different. Alongside interference avoidance, flexible timeslot configuration is required for large uplink bandwidth to achieve optimal resource utilization and user experience. AI can accurately predict service demand trends and thus determine the necessary timeslot configuration. Real-time adjustments to the timeslot configuration based on service demand are required to achieve symbol-level flexible configuration. We can use AI to identify interference features and prevent interference to match services with functions.

Dedicated smart uplink: Built on smart timeslots, carrier aggregation, and dedicated devices, dedicated smart uplink works with AI to boost time, frequency, power, and antenna concurrency. It also opens up the possibility for future networks to offer gigabit uplink.

Wireless AI: 5G and AI can accurately match radio resources to users with converged scheduling algorithms based on service demand, user grouping, scheduling

With the steady development of 5G networks and applications, 5G services are becoming increasingly differentiated, convergent, and diversified. However, uplink bandwidth remains a bottleneck.



CUBE-Net 3.0 technical architecture

prediction, and network MCS/RANK quality. Technologies for network, user, and service sensing can understand service intentions and intelligently select and coordinate resources between networks. This contributes to optimized mobility management, interference management, and load balancing to achieve intelligent service navigation.

2. Ultra-broadband, high-quality optical fiber access

Fixed access networks are evolving from 1G PON and Wi-Fi 5 to 10G PON and Wi-Fi 6, providing gigabit access to homes. In the next 5 to 10 years, optical access networks will develop towards the following three directions.

Direction 1: extending all-optical access to the end. OLTs will need to support multiple forms of flexible networking for the diverse services of the future. Therefore, future optical fibers will continue to be extended to the end. Gigabit FTTR will be deployed in rooms to achieve stable access with high bandwidth and low latency.

Direction 2: providing differentiated slice bearing and open capabilities. OLTs will support end-to-end slicing to meet the differentiated bearing requirements of services such as home broadband, government and enterprise services, and industry applications. The uplink of OLTs supports VXLAN, VLAN, ODU/OSU, SRv6, and other forwarding and routing selection modes. Adopting technologies such as HQoS, hard isolation in chips and PON ports, and Wi-Fi air interface slicing, OLTs work together with bearer network slices to enable end-to-end slicing. Traffic is directed to different slices based on services and applications to create greater synergies with service capabilities on the cloud.

Direction 3: opening up embedded computing power. Computing resources deployed on devices and the edge can help achieve application sensing and real-time pipe optimization. AI-enabled cloud-edge-device collaboration supports application scenarios such as experience management, identifying potential customers, and intelligent O&M. Video optimization, video surveillance backhaul, and industrial IoT can be migrated to cloud by opening up the embedded computing capabilities of OLTs and ONTs and collaborating with the cloud ecosystem.

3. Smart and open all-optical base

The CUBE-Net 3.0 all-optical base has four key features: (1) Stable architecture, all-optical switching, large bandwidth, and a high level of scalability. (2) All-optical anchor and all-service access support one-hop access to the cloud. (3) Intelligent management and control for automatic O&M. (4) End-to-end slicing that provides differentiated SLA services for multiple services on the same network.

The CUBE-Net 3.0 all-optical base also adopts four key technologies:

1.All-optical cross-connection (OXC). OXC equipment that adopts wavelength selective switches and optical backplanes can achieve zero fiber connections, plug-and-play, flexible scheduling, smooth capacity expansion, and ultra-large-capacity wavelength grooming, greatly saving equipment room space and reducing power consumption.

2.Optical service units (OSUs). Service-oriented OSUs can efficiently transport low-rate services and provide a slicing solution for the all-optical base, representing a potential evolution direction for OTN technologies.

3.Centralized and distributed control plane protocol architecture. Centralized SDNs use global resources to compute optimal service paths. Distributed control protocols enable agile and efficient execution, millisecond-level fault detection, and quick response. The network control protocol for all-optical services based on centralized and distributed architecture provides massive connections, ultra-high reliability, and intelligent O&M.

4.Layered intelligent management and control as well as end-to-end orchestration. Using standard ACTN interfaces, the layered architecture helps achieve cross-domain and cross-vendor E2E automatic orchestration and collaboration, providing open, fast, and layered optical network service provisioning and O&M, including fast service provisioning, latency management, and service availability management.

4. Integrated computing-network bearer networks

As IP networks evolve towards IPv6+/SRv6 alongside integrated computing-network architecture, computing-aware networking can enable collaboration between cloud, networks, edge, devices, and services. Enabling technologies include Application-aware Networking (APN), Service Function Chaining (SFC), and in-situ Flow Information Telemetry (iFIT).

The IPv6 extension header transmits application information to networks using

The network control protocol for all-optical services based on centralized and distributed architecture provides massive connections, ultra-high reliability, and intelligent O&M.

APN technology. Using this information, the network can deploy services and adjust resources to satisfy SLA requirements of applications. When a site is deployed at the edge of a network (i.e., edge computing), APN technology can connect the network and applications, adapt to the requirements of edge services, and direct traffic to the right network path, bringing the advantages of edge services into full play.

SFC – an ordered collection of service functions – enables service chains to be created with different types of computing power, so that new types of services can be quickly provisioned. Service flows pass through specified value-added service devices in a specified sequence to acquire one or more value-added services. In a computing network, SFC underpins intent-driven computing services by connecting with different computing services based on customer intention. Alongside SRv6 SIDs and related services, SFC helps build a computing transaction platform. Computing power provided by ecosystem partners is registered with the network as an SRv6 SID. Buyers can purchase the computing power they need, while the network connects computing power services through SFC and provides services to buyers hassle-free.

A visualized and user-friendly interface allows customers to monitor services against SLAs in real time and accurately sense service indicators.

Visualized computing paths and measurable performance have become key capabilities of integrated computing-network architecture. iFIT supports accurate on-demand packet-by-packet detection of performance indicators such as latency, packet loss, and jitter for each service. Second-level telemetry data sampling displays the SLA of service flows in real time. The hop-by-hop deployment mode enables ms-level fault recovery, ensuring the lossless transfer of computing power.

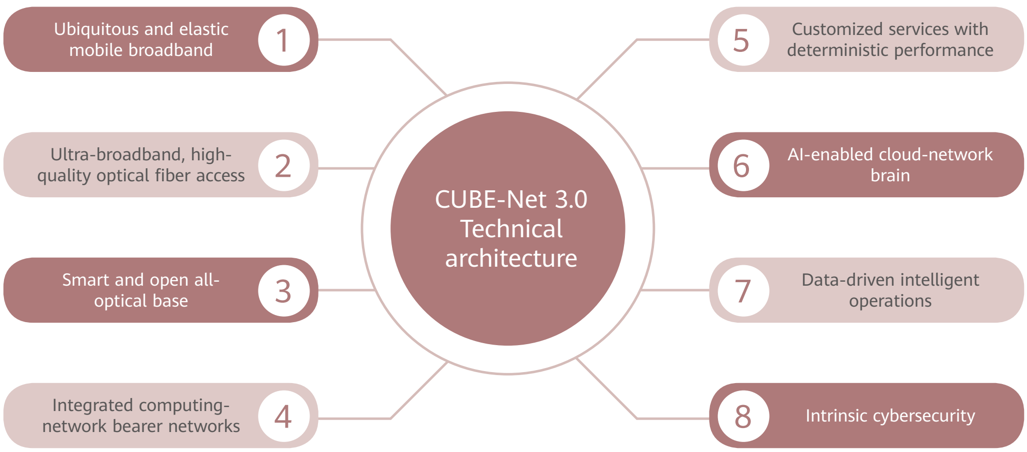
5. Customized services with deterministic performance

Many vertical industry applications, such as industrial control in power grids, manufacturing, IoV, and telemedicine, have precise requirements on network latency, reliability, timing accuracy, and data isolation. Increasing the adoption of 5G industry applications requires differentiated, deterministic, and stable QoS for vertical industry customers.

Using the CUBE-Net 3.0 architecture, China Unicom proposed providing customized networks with deterministic services, which can be approached in the following two ways:

First, unified SLA parameters applied industry-wide on deterministic services. Key requirements of vertical industry applications are classified into different levels to unify and standardize application scenario requirements, network indicators, and deterministic service indicators such as latency, jitter, bandwidth, reliability, positioning, and clock precision. A visualized and user-friendly interface allows customers to monitor services against SLAs in real time and accurately sense service indicators.

Second, removing technical bottlenecks, especially wireless air interface jitter, wide-area determinism, and the ability to integrate industry technologies. 3GPP has proposed URLLC technology to provide low latency and high reliability over the air interface. A multi-service co-existence mechanism has also been introduced. The feasibility and performance of



both technologies are yet to be verified.

The capabilities of air interfaces need to be enhanced. AI-enabled real-time control can offset the adverse effects of having a large number of coexistent, complex services and mitigate the impact of traffic burst on air interface performance. Currently, the technologies for delivering deterministic services have many limitations. For example, applications like connected vehicles and remote control still struggle to deliver deterministic wide-area services across regions. Delivering these services places high demands on network equipment and requires large-scale network upgrades, meaning that cost will become a limiting factor. Currently, 5G systems support latency-sensitive services mainly by serving as a black box and integrating time-sensitive networking (TSN). Realizing native TSN functions and deeply integrating them with industrial networks is a key technical challenge for customized 5G networks.

6. AI-enabled cloud-network brain

Advanced technologies like IT, data technology, and AI can help build secure, reliable capabilities like fast service provisioning across clouds or networks, domains, and vendors, integrated fault diagnosis, resource visualization, automatic load balancing, and coordinated SLA monitoring and prediction. These will combine to form a CUBE-Net 3.0-based cloud-network brain.

End-to-end cloud-network scheduling and management require resource abstraction and a scheduling engine.

Resource abstraction means abstracting clouds or networks in different scenarios and technologies into standard basic resource models based on common data models and providing unified capabilities for acquiring, converting, storing, and managing resources.

Developing a service-independent automatic orchestration engine where design and implementation are decoupled can: (1) Dynamically inject resources and service models, (2) Automatically analyze service intent, receive management policies and configuration items automatically converted by B/O through scenario-based interfaces, intelligently control networks enabled by

CUBE-Net 3.0 will provide network digital twin management capabilities with traceable status, predictable goals, and easy adaptation to changes.

> *CUBE-Net 3.0 will include a blockchain-based trusted computing service system, and a more efficient and reliable intelligent twin system enabled by AI.*

AI inference, and sense and respond to network status changes in real time, and (3) Support highly reliable distributed transaction mechanisms, retries at failure breakpoints, and automatic or manual rollback based on time snapshots. The orchestrator allows for visualization of resources at different layers, service analysis and association, and lightweight OSSs.

CUBE-Net 3.0 will provide network digital twin management capabilities with traceable status, predictable goals, and easy adaptation to changes. It will also offer virtual support for network diagnosis, prediction, decision-making, and sharing, and enable real-time interaction between logical and physical objects, data privacy, and security.

The real-time interaction between logical and physical objects requires real-time sensing and dynamic modeling technologies that can help precisely map logical and physical networks. Software and AI technologies can enable interaction between physical and logical networks, assist with decision-making, and support continuous improvement. This will help dynamically monitor and simulate the real status, behavior, and rules of physical networks, which will in turn support dynamic design and programming, fault simulation, and cutover simulation.

Digital twins create a new digital space outside of closed-off physical resources. This space will become increasingly open and accessible, which will cause more security and privacy issues. Blockchains are decentralized, difficult to tamper with, and provide non-repudiation, making them ideal for data privacy and security. CUBE-Net 3.0 will include a blockchain-based trusted computing service system, and a more efficient and reliable intelligent twin system enabled by AI.

AI is the technical foundation for running and decision-making in autonomous networks. Going forward, AI capabilities will be everywhere at the network element layer. With efficient collaboration between on-premises and the cloud, the CUBE-NET 3.0 network architecture enables AI models to take effect with just one click and be replicated across the entire network, addressing the challenge of large-scale replication of AI applications that has long plagued the industry.

The cloud has strong computing power and can fit all AI application scenarios across networks. As networks are deployed far from end users, AI inference focuses on meeting time-insensitive services (seconds or longer) and cross-domain/cross-vendor scenarios that require global collaboration, like service quality forecasting over a long cycle based on performance indicators, IP and optical collaboration, and wireless and bearer network collaboration. Expert experience and product knowledge can be aggregated on the cloud to help build a powerful AI training platform. Based on this platform, a network knowledge repository can be developed using knowledge graphs and fed into on-premises, which includes the network equipment layer and the management layer, for intelligent network O&M.

On-premises are distributed across different geographical locations and are highly time-sensitive. Due to limited computing power, they focus on providing real-time AI inference capabilities. When a lack of samples of on-premises can meet the precision requirements of training models, the models can be retrained on cloud and the results can be fed into the on-premises inference framework. To address the aging of on-premises models, a regular update and assessment mechanism is used to select optimal models.

7. Data-driven intelligent operations

Data-driven intelligent network operations comprise three parts: data convergence and the opening up of capabilities, data intelligence and building the driver layer, and innovative data applications.

The first part is data convergence and the opening up of capabilities. All-domain convergence enables data source identification and management based on standardized and ID-based data and supports the establishment of data analysis chains and data lineage tracing. Scenario-based data aggregation and public data models can help build data interfaces that are scenario-based, standardized, automated, and that translate data into information. The layered loop provides intent-based interfaces externally, so the upper layer will have lower network expertise requirements. Unified APIs and data make it easier for the upper layer to acquire network data, helping build a layered data governance architecture that supports data collection, transmission, and the analysis of different levels of time sensitivity at different scales and different levels of precision.

With a data security management framework that covers the entire data lifecycle, data management focuses on the highest-risk data transfers. A security management policy created based on a review of existing data assets and data risks can quickly satisfy the customized data security needs of businesses and households.

The second part is data intelligence and building the driver layer. General algorithm models, data simulation, and a network operation knowledge repository are used to build a data-driven middle layer for the research and development of upper-layer applications.

The third part is innovative data applications. Using converged data and open capabilities, network AI algorithms, and network operation knowledge repositories, we can build a series of intelligent data applications such as dynamic network monitoring and prediction and network configuration, maintenance, and optimization. These are used to track and sense service objects in real time while dynamically and intelligently adjusting resources to meet diverse customer needs.

8. Intrinsic cybersecurity

The CUBE-Net 3.0 intrinsic network security module uses the following five technologies:

Channel-associated trusted identity technology optimizes communication protocols and network equipment, and embeds trusted identifiers and password credentials into packet headers. Network equipment can verify the authenticity and legitimacy of requests through identity verification, preventing fraud and building fine-grained access authentication and sourcing capabilities.

Intrinsic NE security technology based on trusted boot and abnormal behavior

Software-defined and integrated security capability orchestration provides security functions in the form of pools and microservices, making integrated orchestration and the opening up of capabilities possible.

detection introduces chip-level trusted computing technology into NEs, which helps build a trusted, secure software and hardware running environment based on the bottom layer of NEs. This allows for verification at all layers, from hardware platforms and operating systems to applications, ensuring the confidentiality and integrity of systems as a whole.

AI-enabled dynamic planning of security policies strengthens capabilities like self-learning and the modeling of traffic and service characteristics, risk prediction, and security policy orchestration based on feature models. This function also enables security policy conflict detection and automatic optimization.

Blockchain-enabled security management of digital network resources helps build a trusted system for digital network resources such as IP addresses, domain names, and AS numbers. A distributed accounting and consensus mechanism ensures the authenticity of resource ownership and mappings and prevents security issues such as IP tampering, route hijacking, and domain spoofing.

Software-defined and integrated security capability orchestration provides security functions in the form of pools and microservices, making integrated orchestration and the opening up of capabilities possible. This will allow users to flexibly define security policies, invoke security resources on demand, and agilely deploy and roll out security capabilities. Intelligent policy planning enables security on any cloud and any network.

Enabling industry digitalization with an open ecoystem

China Unicom will use CUBE-Net 3.0 to create an ecosystem with capabilities opened up on demand and easily integrated, build a unified platform for opening up network capabilities, implement a strategy of network service as a platform, provide standardized network capability APIs, open up more network service elements in the form of applications, and make network as a service (NaaS) a reality.

It will also work closely with vertical industries, industry application developers, and industry device providers to integrate applications, computing, networks, and devices, and provide intelligent, converged services for vertical industries, thus building up its strengths in connectivity, computing, and intelligence, enabling industry digitalization.

With CUBE-Net 3.0, China Unicom will also work to build a controllable open source technology ecosystem. We will continue to pursue open innovation, make full use of integrated applications and consolidated resources, expand the open source ecosystem, and work with industry partners to build a technical alliance for the open source industry. To become self-sufficient in basic software and hardware and core technologies, China Unicom will intensify efforts to pursue technical partnerships with domestic leaders.

Together, we will build a full-stack open-source technology ecosystem that covers chips, devices, networks, operating systems, and security, over which we will have full autonomy.

Five Trends Shaping Next-gen Data-intensive Supercomputing

Underpinned by data intensification, five major trends are shaping today's supercomputing industry and driving the balanced development of data ingestion, storage, computing, transmission, and application.

By Chen Mo, Chief Solution Architect, Data Storage and Intelligent Vision Product Line, Huawei

Supercomputing is a key technology that can transform society and open the door to humanity's next stage of technological evolution. Many nations are prioritizing supercomputing R&D, including the US, Japan, the EU, Russia, and China. China, for example, has 188 entries on the June 2021 release of the TOP500 most powerful supercomputers, with Tianhe-2A ranking 7th. Before that, Tianhe 2, Tianhe 3, and then Sunway Taihu Light topped the list ten consecutive times.

Nevertheless, new technologies like cloud computing, big data, artificial intelligence, and blockchain have drawn attention away from supercomputing. And alongside a limited application ecosystem coupled with an insufficiently diverse machine-hour supercomputing service model, China's supercomputing sector has much room to develop.

Seeing its broader value in terms of socioeconomic progress, an increasing number of provinces and cities in China are establishing supercomputing centers and deploying next-gen supercomputing systems. China currently runs 10 national-level supercomputing centers in major cities, including Tianjin, Shenzhen, Guangzhou, and Xi'an, with many more planned.

Five trends

To transform supercomputing centers from computing service providers to integrated data value providers, China is prioritizing five supercomputing trends: diversified computing, all-optical networks, intensified data, containerized applications, and converged architecture.

Remote Direct Memory Access over Converged Ethernet (RoCE) and lossless data technologies make it possible for the computing, storage, and management networks in a supercomputing center to be integrated into a single box.

Diversified computing is becoming mainstream. Traditional high-performance computing (HPC) systems use CPUs for double-precision floating-point computing, while emerging supercomputing systems use CPUs, GPUs, and FPGAs for more powerful parallel computing. Today's industry in China is stepping up R&D and the deployment of homegrown microprocessors and accelerators, improving the efficiency of diversified heterogeneous computing and improving the efficiency of diversified hybrid applications.

Optical switching technology is maturing and networks are becoming all-optical. Remote Direct Memory Access over Converged Ethernet (RoCE) and lossless data technologies make it possible for the computing, storage, and management networks in a supercomputing center to be integrated into a single box. An all-optical supercomputing Internet between supercomputing centers has been proposed to facilitate resource sharing.

Data is becoming intensified. Traditional supercomputing applications, such as weather forecasting, energy exploration, and satellite remote sensing, will generate increasing amounts of data as precision improves. Moreover, more than 80% of emerging supercomputing applications, including autonomous vehicles, genetic testing, and brain science, generate data at the petabyte scale. Larger data volumes, more data types and concurrent tasks, and higher reliability requirements demand more from supercomputing storage to deliver higher bandwidth, IOPS, reliability, and support for massive concurrent access.

Containerized applications. Containerization technology can encapsulate the supercomputing operating environment to decouple supercomputing applications from the underlying hardware, making supercomputing easier to use for the majority of non-expert users. Containerization technology is currently open-source, making ecosystem development more viable.

Supercomputing architecture is converging. Aligned with the first four trends, supercomputing will adopt a heterogeneous, multi-state composite architecture to converge siloed resources, data, and applications. This means a unified, converged heterogeneous system in which CPUs, GPUs, and other dedicated computing power systems are scheduled on a unified service scheduling platform. Various supercomputing applications will be managed on a unified application platform and data assets carried by a unified data foundation. Data silos will be broken down, ensuring that no data needs to be migrated if the data foundation remains unchanged, optimizing TCO and boosting ROI.

Data-intensive

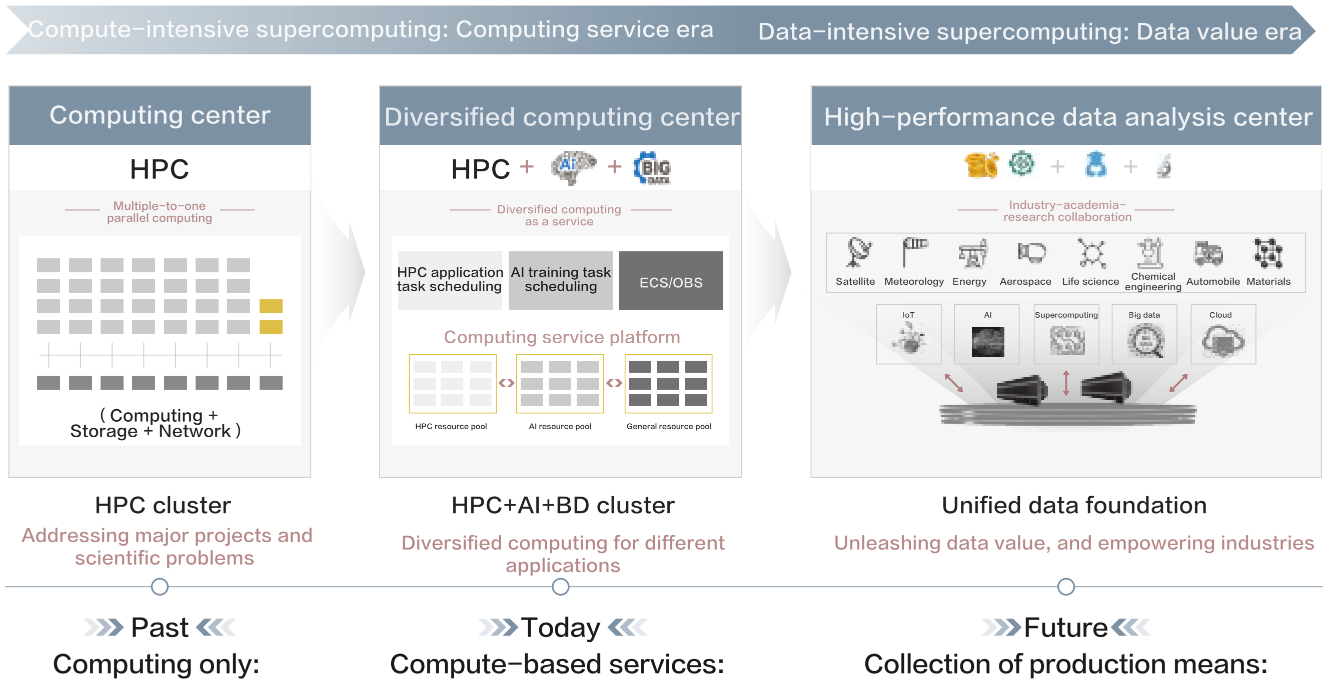


Figure 1: Supercomputing entering the data-value era

supercomputing at the core

Of the five trends, data intensification is the most significant. Traditionally, supercomputing is mainly used to solve computing problems. Customers bring data on hard drives to a supercomputing center and copy the results back onto hard drives, leaving no data for long-term storage in the center. However, the evolution of supercomputing has led to both changes and new challenges.

First, the amount of data involved in computing has increased dramatically. For example, the improvement in precision by applications like weather forecasting and satellite remote sensing has doubled data volumes. More types of data, both structured and unstructured, are involved in computing. For example, image data is directly used for computing in applications like brain science and electron cryomicroscopy (CryoEM).

Second, computing power has increased dramatically. Currently, few single tasks can consume all the computing power in a cluster, so multiple tasks run concurrently in most cases. The 100 PFLOPS HPC Center at Shanghai Jiao Tong University can run nearly 50 concurrent tasks, some of which require high bandwidth while others require high I/O performance. Therefore, more balanced storage capabilities are needed.

Third, higher reliability is required. When traditional supercomputing was applied to research projects, users could tolerate iterative operations before a reliable result was obtained. However, today's supercomputing is mostly applied to production systems that have higher requirements on the reliability of both results and processes. Storage systems

> *Data-intensive supercomputing serves as a data-centric, high-performance data analytics platform with the analytics capabilities of traditional supercomputing, big data analytics, and AI.*

therefore need to be extraordinarily reliable.

Fourth, supercomputing centers and data centers need to converge. In recent years, supercomputing centers are exploring more diversified services such as AI computing, big data analytics, virtualization, and disaster recovery. During this process, centers have found data mobility to be the biggest problem, as storage is split between supercomputing files, virtualized blocks, machine learning objects, and big data HDFS. Mobilizing stored data is the biggest challenge facing supercomputing centers.

These developments are both challenges and opportunities for the data storage industry that is key to transforming supercomputing from computing-intensive to data-intensive.

Data-intensive supercomputing serves as a data-centric, high-performance data analytics platform with the analytics capabilities of traditional supercomputing, big data analytics, and AI. It supports end-to-end scientific computing services through application-driven, unified data sources. It also provides diversified computing power for both research and business, and provides high-level data value services leveraging accumulated knowledge about data.

Data-intensive supercomputing transitions computing centers to centers delivering diversified computing services. Ultimately, diversified computing convergence and a unified storage foundation for massive data will enable high-performance data analytics, driving supercomputing from the computing-service era to the data-value era (Figure 2).

Data-intensive supercomputing delivers the following value:

Research: The architecture that converges HPC, AI, and big data technologies is an interdisciplinary innovation fueled by data-intensive research. It facilitates the evolution of research from the third paradigm (computational science) to the fourth paradigm (data science).

Business: The unified data foundation is converged, efficient, secure, and low-carbon, reducing the lifecycle management costs of massive structured and unstructured data and improving the data utilization efficiency of applications that converge scientific computing, big data, and AI.

Industry: Chinese-made high performance data analytics (HPDA) software, parallel file systems, and data storage and data management systems have boosted the development of China's supercomputing storage industry and application technology ecosystem.

Wide adoption of data-intensive supercomputing

Data-intensive supercomputing is widely used in scientific research, manufacturing, and business.

For example, in gene sequencing applications, MGI's DNBSEQ-T7 sequencer generates 4.5 TB of data every 24 hours. At full load, that totals 1.7 PB of data a year. Analyzing biological information generally creates intermediate files and results that are about five times the volume of the original data. West China Hospital adopted data-intensive supercomputing to improve gene sequencing efficiency, shortening the time required for a single gene sequencing

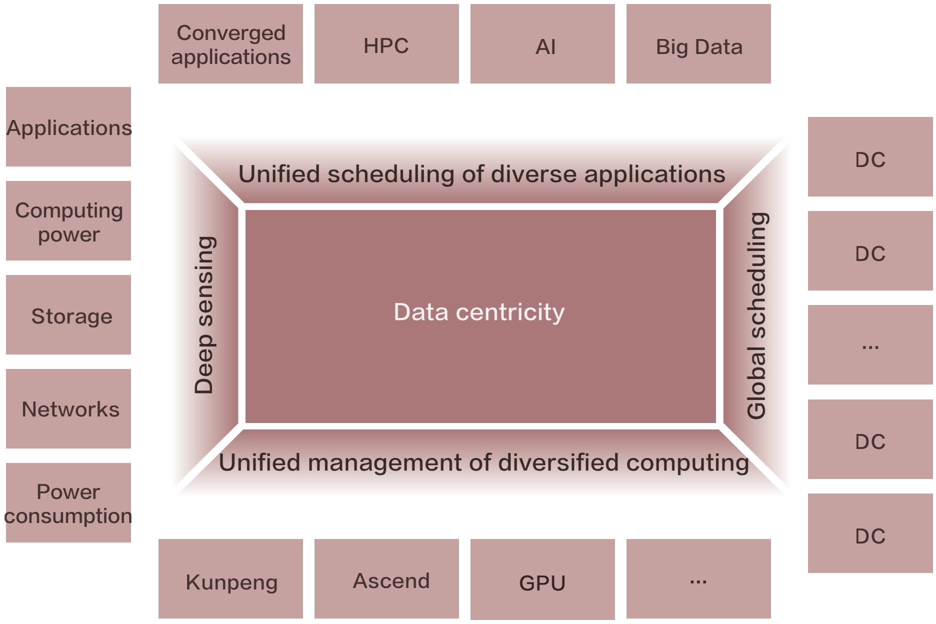


Figure 2: Data-centric, data-intensive supercomputing

task from 3 hours to just minutes.

Autonomous applications are highly complex and involve more than 10 steps, including data import, preprocessing, training, simulation, and results analysis. Each requires different protocols such as object, NAS, and HDFS. Data silos are also a major issue, because it takes twice as long to copy data than it does to process and analyze it. The auto maker Geely-Volvo has adopted data-intensive supercomputing and uses a single data foundation to support multi-protocol interoperability and adapt to the entire service process, reducing data storage costs while improving data analysis efficiency.

In university supercomputing applications, the π supercomputer series at Shanghai Jiao Tong University and the Hanhai supercomputer series at the University of Science and Technology of China provide more balanced data access capabilities by adopting data-intensive supercomputing. The series can support more than 50 concurrent supercomputing tasks with different load demands. When computing facilities are upgrading, legacy data can be stored for a long time without migration to provide greater support for research tasks.

At the Ninth Supercomputing Innovation Alliance Conference held in China in September 2021, a data-intensive supercomputing working group was established, recognizing data as equally as important as computing power. Data intensification was also a key topic at the Seventh China Scientific Data Conference held in Hohhot, Inner Mongolia a month later. And at CCF HPC China 2021, Huawei and the CCF HPC Profession Committee jointly released the *Data-intensive Supercomputing Technology White Paper.*

The industry consensus is that data-intensive computing can unlock a thriving supercomputing industry covering data collection, storage, computing, transmission, and utilization capabilities.

West China Hospital adopted data-intensive supercomputing to improve gene sequencing efficiency, shortening the time required for a single gene sequencing task from 3 hours to just minutes.

Eastern Data and Western Computing: Building New Computing-first Networks

The Eastern Data and Western Computing initiative aims to address regional computing supply imbalances. As part of the initiative, the national-level integrated big data center project poses a huge challenge for networks – one that requires intelligent IP network solutions to overcome.

By Li Jun, Chief Architect, Data Center Networks, Huawei

In May 2021, China's National Development and Reform Commission (NDRC), Cyberspace Administration of China (CAC), Ministry of Industry and Information Technology (MIIT), and National Energy Administration (NEA) jointly issued the *Implementation Plan for Computing Hubs of the National Integrated Big Data Center Collaborative Innovation System*. The plan proposes a new computing network system that integrates data centers, cloud computing, and big data, as well as Eastern Data and Western Computing demonstration projects that will enable high-quality, green data centers.

Overall, the project will create an integrated national system of big data centers and strengthen the planning and intelligent scheduling of computing power in line with China's National Economic and Social Development and Vision 2035.

Integrated arrangement: Networks are key

As China's digital economy has accelerated, the imbalanced distribution of computing facilities between the eastern and western regions of the country has become an increasing problem. The east has high demand for computing applications and strong innovation capacity, but a shortage of land, hydropower, and other auxiliary resources. The west has a favorable climate and abundant energy sources, but its digital industry is lagging. The project can optimize the distribution of computing power and applications and help form a nationwide market where data is shared, flows freely, and is allocated on-demand.

China will deploy and build hub nodes in the following areas: Beijing-Tianjin-Hebei Metropolitan Region, the Yangtze River Delta, and the Guangdong–Hong Kong–Macau Greater Bay Area in eastern China, and Chengdu, Chongqing, Guizhou, Inner Mongolia, Gansu, and Ningxia in central and western China.

The hub nodes will consist of data center computing clusters. Data center networks will connect ultra-large-scale computing and storage resources to take advantage of the intensiveness and scale of networks, and ensure the abundant supply of computing power. The hub nodes can exchange data and computing power over a wide area network (WAN), satisfying the needs of high-frequency, real-time interactive services locally. It will also remotely support non-real-time computing tasks like offline analysis, storage, and backup.

Networks will play a key role throughout computing power generation, transmission, and consumption, and in building the integrated big data center system.

The project will create an integrated national system of big data centers and strengthen the planning and intelligent scheduling of computing power.

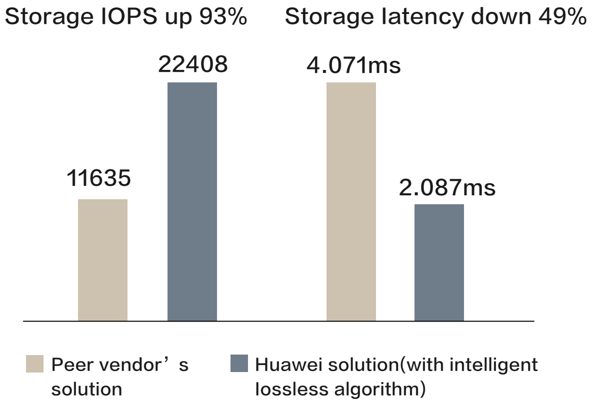


Figure 1: Distributed adaptive routing

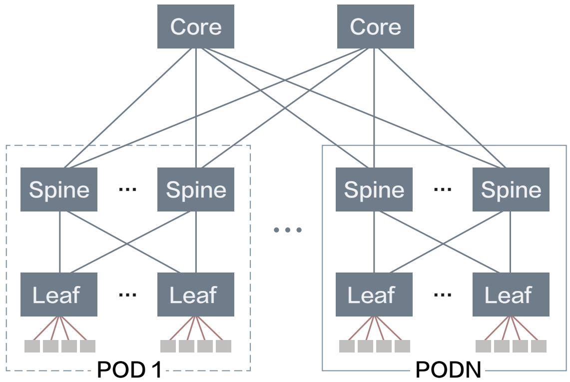


Figure 2: Conventional data center Clos networking

Computing power generation and scheduling: Networks as a challenge

The storage and processing of data in the west of China that comes from the east poses a daunting challenge for network performance. The dependency of large-scale server clusters on networks for interconnection results in packet loss when multiple servers send large numbers of packets to one server simultaneously, exceeding the caching capacity of the switch. Data retransmission will in turn greatly compromise computing and storage efficiency. On RDMA over Converged Ethernet (RoCE) networks, a packet loss rate of 0.1% would result in a 50% decrease in computing power, a massive waste of server CPUs that presents a major barrier to improving computing power.

WANs act as channels connecting data in eastern China with computing power in western China, bearing hundreds of services for massive enterprises. Each service has unique requirements on key capabilities such as bandwidth, latency, and computing power. WANs must be able to efficiently schedule carried services to deliver the expected value of the Eastern Data and Western Computing project. The best-effort service model of traditional IP networks can barely differentiate services for scheduling and thus fails to meet different service requirements. Network and cloud pool resources cannot be fully utilized and enterprises cannot be allocated the optimal cloud pool, increasing the cost of enterprise migration to the cloud and degrading service experience.

Three innovations: Building IP networks for cross-regional computing power scheduling

Using its expertise in network technology

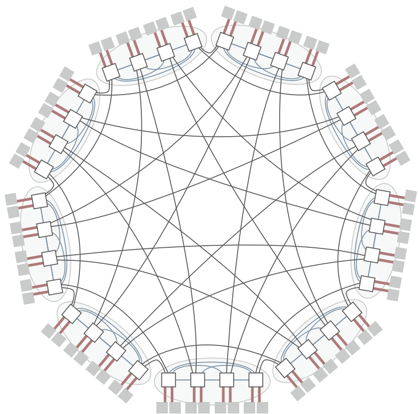


Figure 3: New directly connected topologies in data centers

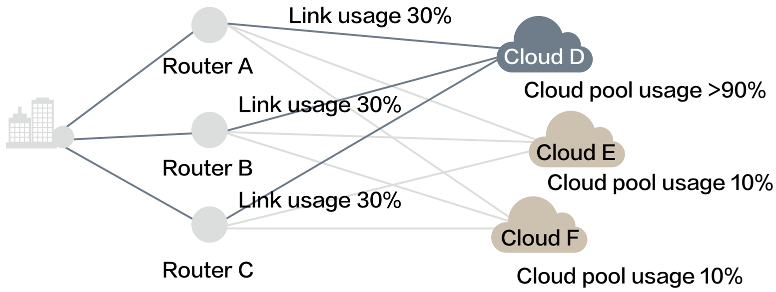


Figure 4: Conventional WAN scheduling

R&D and commercial projects over the past 30 years, Huawei has launched an industry-leading IP network solution that supports cross-regional computing power scheduling. The solution comprises CloudEngine 16800 series data center switches and NetEngine 8000 F8 series WAN routers. The solution uses the innovative intelligent lossless algorithm and intelligent cloud-map algorithm to optimize the scheduling of computing power on the generation side and transportation side, thus achieving minimum loss, best performance, and the optimum transportation of computing power.

Innovation 1: Intelligent lossless algorithm

The key to addressing packet loss in data center networks involves setting an appropriate congestion indication threshold for switch buffer queues. If the threshold is set too high, the transmission rate of the sender server and network congestion cannot be reduced in time, greatly increasing packet loss or latency. If the threshold is set too low, the transmission rate of the sender server greatly declines and the network cannot achieve 100% throughput, resulting in wasted resources.

Service traffic models now vary greatly and even with extensive testing and simulation, it is difficult to determine the optimal threshold based on human experience. To address this, Huawei has developed intelligent algorithms for data center network switches. These switches can dynamically set the optimal queue threshold using intelligent lossless algorithms based on real-time network status information such as queue depth, bandwidth throughput, and traffic model. To ensure that the algorithm can adapt to any scenario and traffic model, Huawei has used millions of real service samples and tens of millions of random samples to train

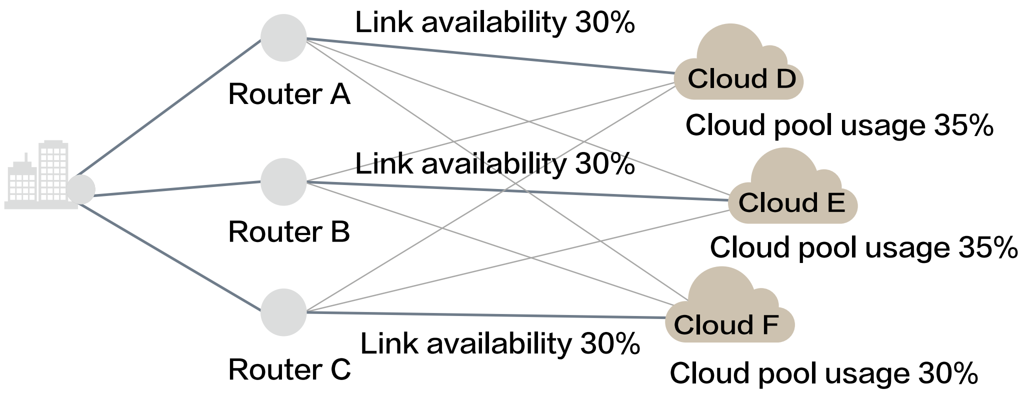


Figure 5: WAN-based cloud-network integrated scheduling

the algorithm model, achieving the optimal balance between zero packet loss, high performance, and low latency, as shown in Figure 1.

High-performance computing and parallel computing require extensive computing power. For high-speed interconnections between massive computing nodes, networks must support ultra-large-scale networking and low latency, as well as zero packet loss. Today's mainstream data center network architecture is Clos (as shown in Figure 2), where networking scale is subject to the port density of the switch. For a 64-port switch, a 3-level Clos architecture supports a maximum of 65,536 server ports. More than 200,000 compute nodes need to be connected to build 10-EB-level computing capabilities. Forwarding between computing nodes on a level-3 Clos network takes five hops. On a level-4 Clos network, networking cost increases greatly and forwarding takes seven hops. As a result, computing efficiency is greatly reduced by the latency increase.

To boost networking scale while minimizing costs and network latency, Huawei has introduced directly connected topology (as shown in Figure 3) for Ethernet networking, breaking the limitations of Clos architecture and realizing ultra-large-scale networking with low network diameter (small number of hops). Distributed adaptive routing technology utilizes unequal cost paths to realize dynamic routing, ensuring low latency while improving bandwidth utilization. Once upgraded, the Huawei

CloudEngine series switches support directly connected topologies and adaptive routing. For example, 64-port switches support zero packet loss networking with up to 270,000 servers, a scale that is four times the industry average with level-3 Clos. The number of network hops and latency can be reduced by 25%. At the same scale of server networking, NE nodes can be reduced by 20% to 40% compared with level-3 or level-4 Clos networking.

Innovation 3: Intelligent cloud-map algorithm

Conventional WANs use shortest-path scheduling, which results in unbalanced link utilization when services are transmitted on the same path. Multi-path load balancing improves network utilization, but does not satisfy services' different network requirements such as latency, jitter, and reliability. Moreover, this method only considers network factors and not cloud pool factors such as computing usage, cost, and storage, resulting in the unbalanced use of cloud computing resources and the inefficient scheduling of computing resources (as shown in Figure 4).

To address this issue, Huawei has introduced the Edge-Disjoint KSP algorithm. This algorithm integrates network factors, such as latency, bandwidth, reliability, and availability, and cloud factors, like computing usage, storage resources, and cost, to implement cloud-network map modeling. The algorithm gives the recommended optimal path for different services through dynamic parallel computation based on multidimensional constraints. The algorithm then defines and orchestrates the path labels of SRv6 packets according to these recommendations and includes service data in the SRv6 packets. WAN routers across the network can forward services in the optimal paths based on service types, which optimizes cloud pools, networks, and services simultaneously. Based on the integrated scheduling of cloud factors and network factors, the algorithm can select the optimal cloud pool according to enterprise requirements and balance cloud-network resources from multiple sources to multiple destinations, improving computing power transmission efficiency by more than 30%, as shown in Figure 5.

Since the Chinese government proposed the guiding principle for building integrated big data centers back in 2016, the concept of new computing-first networks has gained popularity in the industry. Huawei is actively involved in designing and building nationwide hub nodes and is thus committed to achieving innovations in IP network solutions.

Based on the integrated scheduling of cloud factors and network factors, the algorithm can select the optimal cloud pool according to enterprise requirements.

All-optical Sensing Brings Intelligent Automation to Oil & Gas Pipeline Inspections

While sensing technologies deployed on oil and gas pipelines aren't new, they tend to be plagued by issues such as false positives, false negatives, and misidentification. Huawei has taken the optical technologies it uses in fiber communications and applied fiber sensing, resulting in the Huawei OptiXsense EF3000 – Huawei's first intelligent optical sensing product.

By Zheng Kai, tech journalist

Petroleum is one of the most important economic resources in the world today. However, uneven geographical distribution and the physical properties of petroleum mean that transporting the resource to where it is needed presents a tough logistical challenge.

Oil and gas are typically transported by rail, road, water, air, or pipelines. Of these, pipes are the least restrictive method. However, inspections are notoriously difficult on pipe networks that can traverse tens of thousands of kilometers. A 200-km to 300-km stretch of pipeline typically requires a dedicated team of three inspectors in one vehicle, assuming navigable terrain free of obstacles like reservoirs, lakes, forests, and steep slopes.

At HUAWEI CONNECT 2021, David Wang, Executive Director of the Board and President of ICT Products & Solutions, said that the digital infrastructure of the future would need to be hyper secure, reliable, and deterministic, as well as equipped with efficient data circulation and computing power capabilities. Huawei has taken the optical technologies that it usually uses in fiber communications, applied them to fiber sensing, and built the Huawei OptiXsense EF3000, Huawei's first intelligent optical sensing product.

The Challenge of inspecting 90,000 km of pipeline

China's oil and gas resources are mainly concentrated in the northeast and northwest of the country, but the major fuel consumers are located in the cities of the southeast coast and south-central area. Oil and gas must be transported over vast distances, with pipeline transportation providing the best solution, but

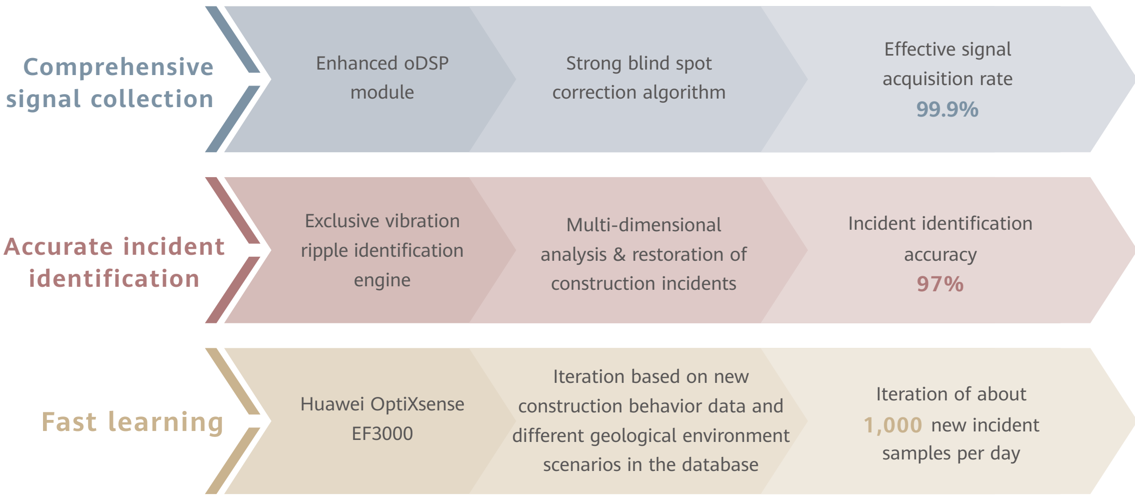
one that presents the most challenges when it comes to inspections.

China Oil & Gas Pipeline Network Corporation (PipeChina) builds and operates various infrastructure, including oil and gas trunk pipeline networks as well as gas storage and peak shaving. It also works on interconnecting trunk pipeline networks and docking with local pipelines, operating and scheduling nationwide networks, and ensuring fair and open user access to the infrastructure. As of 2021, the total length of PipeChina's oil and gas pipelines exceeded 90,000 km, and by 2025, the total length is expected to exceed 120,000 km.

PipeChina employs a pipeline inspection team of about 15,000 people. However, as the pipelines become longer, inspection management and operations are becoming increasingly difficult.

To address this problem, PipeChina plans to combine both manual and automatic pipeline inspections and use digital technologies to automate the entire inspection process, from detection and analysis to decision-making and warnings. Zero false negatives and minimal false alarms are expected, and reliance on manual inspections will be gradually reduced. Other problems will also be solved such as losses resulting from issues like inadequate repairs to damage caused by other construction projects, inefficient pipeline management methods, and poor accuracy of early warnings.

Yu Hao, a senior expert at PipeChina's Digitalization Department, said, "Huawei has been actively involved in the pipeline industry and has innovated technologies



Three core technologies of Huawei OptiXsense

such as optical sensing and big data. These technologies have shown us the possibility of automated unmanned inspection of pipelines. We hope that Huawei continues to invest in this area to contribute to ensuring safe, reliable, and efficient pipeline transportation in China."

Three core technological capabilities

Some digital technologies have already been applied to pipeline inspections such as sensor technologies. However, problems still remain such as false negatives where damage goes undetected, damage is detected where there is none, or misidentification occurs where construction is mistakenly categorized.

Some technologies, such as aerial drones used for inspecting power grids, are ill-suited to oil and gas pipeline scenarios, because most pipelines are built underground.

Huawei's OptiXsense is an optical-fiber vibration sensor with best-in-class identification accuracy. Three major capabilities set the OptiXsense EF3000 apart from rival products:

Signal collection: Huawei's enhanced oDSP module has a strong, built in correction algorithm for blind spots, which can correct and shape the phase of weak signals and raise the effective signal acquisition rate to 99.9%.

Accurate identification of incidents: Huawei's exclusive vibration ripple identification engine can analyze construction incidents across multiple dimensions. For each construction vibration point, a minimum of 32 pieces of phase information are obtained. Multiple features can be extracted, including voiceprint, frequency, space, time sequence, and duration. Samples are identified and compared through multi-dimensional deep convolution, raising incident identification accuracy to 97%.

Fast learning: Huawei OptiXsense EF3000 can perform iteration based on new construction behavior data and different geological environment scenarios in the database. Huawei works with universities that provide large amounts of geological data – up to 1,000 new incident samples each day. Huawei uses this data to improve the sensing and warning accuracy of its OptiXsense products.

With these three technical characteristics, we believe that OptiXsense products can help digitalize and automate the E2E inspection process with zero false negatives and minimal false alarms.

All-optical innovation: Technology and use cases

Huawei has drawn on its 30 years of cumulative expertise in optical technology to apply all-optical sensing technology to the pain points that have existed in oil and gas pipeline transport for years.

OptiXsense EF3000 marks Huawei's first foray into fiber sensing beyond communications networks. We plan to develop additional optical technologies, such as the distributed sensing of optical fiber vibration, temperature, stress, and water quality, for many more industries, including electric power, transportation, government, and sanitation. In conjunction with big data and GIS mapping, these technologies will support differentiated, multi-dimensional, and intelligent detection and warning solutions.

Shandong Jihua Gas was one of the first companies to adopt the OptiXsense EF3000. The company operates gas pipeline networks at various levels, comprising more than 3,000 km of piping that supplies gas to 920,000 households and more than 3,000 industrial, boiler, and public welfare users in Jinan city. To ensure that the fiber sensing and adaptive AI identification solutions can be adopted in pipeline construction, Huawei has worked with Shandong Jihua Gas to deploy a 20-km underground pipeline that traverses complex environments such as suburbs, national highways, and rural areas. Construction samples from excavators, rammers, ditchers, and manual excavation were collected onsite for training and model creation. Based on these, 56 tests were conducted on different road segments. The initial phase of technical verification has been completed, proving the efficacy of Huawei's first intelligent optical sensing product.

As major energy arteries of the world economy, oil and gas pipelines will inevitably become digitalized. And in the future, fiber sensing technologies will continue to evolve, benefiting more industries by addressing particular requirements.

Huawei has drawn on its 30 years of cumulative expertise in optical technology to apply all-optical sensing technology to the pain points that have existed in oil and gas pipeline transport for years.

Native Hard Pipe Networks: An Optical Foundation for Industries to Go Digital

As industries digitalize, the demand for integrating multiple coexisting networks is increasing. In response, Huawei has launched Hybrid OTN (H-OTN), the industry's first optical communications product for native hard pipe (NHP) networks. These simplified, secure networks provide E2E physical isolation of services, making them perfect for a variety of industries, including electric power and transportation.

By Zhang Chi, Optical Marketing Execution Dept, Huawei

Electric power and transportation provide essential services for national economies and people's livelihoods and as such form a vital part of the digital transformation process.

In the electric power industry, the energy Internet emerged on third-generation power grids at the end of the 20th century, supporting new power systems that use clean energy and run on smart, ultra-high voltage power grids and architecture. In the transportation industry, China's 2019 Program of Building National Strength in Transportation set similar requirements: first, deeply integrate advanced IT into transportation to digitalize the system. And second, establish an integrated transportation system that connects sea, air, road, rail, and freight transportation.

Industries need to digitalize both infrastructure and services. Digital services need to rely on digital infrastructure equipped with comprehensive sensing and connectivity capabilities. Native hard pipes are an underlying technology for digital infrastructure that supports secure, real-time services for multiple industries.

Multiple coexisting networks result from new digital services

Industrial digitalization is a long-term, continuous process that requires regular adjustments to match social development and business changes. As new services emerge, we have seen increasing numbers of communication networks being built to support them.

For example, rapidly developed new services in the electric power industry, including distribution automation, intelligent inspection, and distributed clean energy, have resulted in the construction of multiple networks. Distribution automation in particular, has become indispensable to modern power grid management and most customers now prefer to build independent, fiber-optic communication networks for power distribution.

As quickly expanding power facilities increase installed generation capacity and transmission line length, manual inspection is no longer enough. This has driven the adoption of intelligent inspection that requires networks with high bandwidth, low latency, and high security, so many customers choose to build new networks for this service.

Under carbon peak and carbon neutrality goals, more distributed clean energy systems need to be connected and the underlying power IoT must support power grid information collection and industrial control services. Independent IoT networks that provide end-to-end SLAs and support service isolation are necessary to meet the high security requirements of these services.

Similar problems exist in the transportation industry. For example, metro station operations systems run multiple services, including passenger information, automated ticketing, and video surveillance. These service systems are usually built separately, meaning each has its own communication network.

The passenger information system, for example, provides passenger notices, service times, emergency warnings, and other dynamic information, making it a critical production service requiring an independent secure network with real-time capabilities. The automated ticketing system, which has evolved over the past 20 years from reading magnetic

Under carbon peak and carbon neutrality goals, more distributed clean energy systems need to be connected and the underlying power IoT must support power grid information collection and industrial control services.

cards to processing mobile payments, also needs a separate network due to strict security requirements. With maturing video surveillance technology, yet another network is needed for real-time surveillance, dispatch and command, video conferencing, and other modern video services. Each needs an underlying communication system that is secure, stable, and reliable, and that enables simple O&M.

Integrated, simplified optical networks

The emerging global consensus on low-carbon, environmentally friendly development is heavily influencing evolution plans for communication networks.

These are just a few examples of how the new services brought by industrial transformation often lead to multiple coexisting networks, which results in repeated network investment, complex O&M, low utilization, and insufficient security.

They need to be integrated for industrial digitalization to proceed for multiple reasons.

First, the emerging global consensus on low-carbon, environmentally friendly development is heavily influencing evolution plans for communication networks. Multiple coexisting networks simply waste too many resources. Second, simplifying the architecture of communication networks to avoid complex operations across multiple layers will speed up the rollout of new services as they emerge. Finally, the complex O&M and low efficiency generated by these redundant networks are too costly for many enterprises seeking to improve management efficiency by going digital.

Therefore, these networks need to be integrated, which requires the ability to create a single simplified network that supports the end-to-end physical isolation of services.

From NHP to H-OTN

Hard pipe optical transmission technologies, like PDH and SDH, have been used in the electric power and transportation industries over the past 40 years to ensure the stable and secure operations of production transport networks.

However, as more industries go digital, new problems, like those mentioned above, have occurred. Given this, Huawei has launched the native hard pipe (NHP) network solution supported by optical service unit (OSU) technology, a hard pipe technology similar to PDH and SDH. The NHP network is a simplified optical network that supports end-to-end hard pipe isolation by extending the use of OSU technology from WAN to access networks. As a hard pipe technology, OSU can physically isolate different services.

With OSU-A technology, NHP networks replace multiple traditional access networks with one P2MP fiber-optic network. Multiple coexisting access networks – for example, in electric power scenarios for distribution automation and intelligent inspections and in transportation scenarios for passenger information and automated ticketing – are integrated and multiple services are transmitted through different OSUs, ensuring each service is physically isolated. This saves resources and helps enterprises become green and low-carbon.

NHP networks have three distinct advantages:

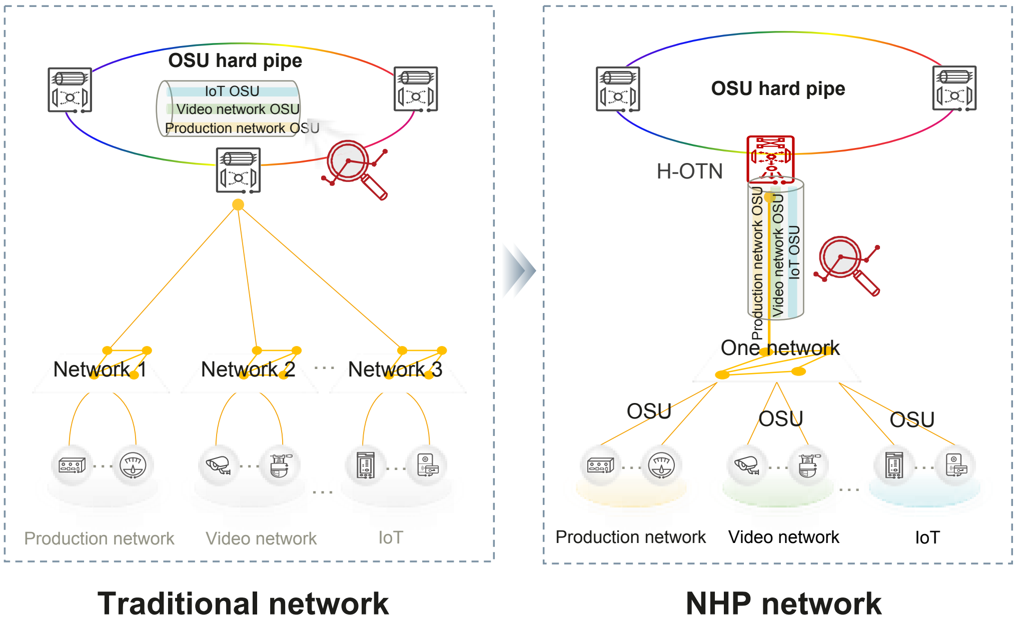


Figure: Comparison between traditional network architecture and NHP network architecture

First, they are more secure. Digital services are carried through hard pipe OSUs from terminals to the access network and then to the WAN, which enables end-to-end physical isolation, effectively ensuring service security.

Second, they are faster. As a native hard pipe technology, OSU provides deterministic ultra-low latency and zero jitter, enabling real-time services.

Third, they are more efficient. The simplified architecture of NHP networks can reduce enterprise management costs by improving O&M efficiency and enabling rapid service rollout.

At HUAWEI CONNECT 2021, Huawei launched a key component of NHP networks: H-OTN, an optical communication product that integrates hard pipes. H-OTN is based on OSU hard pipe technology, and has an innovative point-to-multipoint (P2MP) OTN architecture.

We have been working with State Grid Shanxi Information & Telecommunications to explore H-OTN and the SDH + OSU dual-plane solution since 2016. This new solution further deploys OSU on terminal access points of the power distribution network to build an end-to-end native hard pipe network.

Native hard pipe technology is the ideal tool for deepening industrial digitalization. It can provide high-security, high-reliability, low-latency, and high-bandwidth transport networks for various industries such as electric power and transportation. The launch of the H-OTN will also open up infinite possibilities for digital network infrastructure, lighting the way forward for industrial digitalization.

How Intelligent All-optical Commercial Buildings Enable SME Digitalization

Commercial buildings house a large number of SMEs and as they become increasingly digitalized and cloud-based, private lines for basic access are no longer sufficient. Carriers need to provide SMEs with high-quality private lines that bring convenience and a premium experience.

By Zhang Han, Solution Architect, All-optical Network Architecture Innovation Lab, Optical Architecture & Technology Planning Dept, Huawei

New infrastructure underpinned by technologies like 5G, F5G, IoT, industrial Internet, big data, cloud computing, and AI is becoming deeply integrated with the real economy in industries such as urban governance, transportation, campuses, enterprise operations, and manufacturing. In turn, this is boosting the economy and promoting the digital transformation of industries.

According to 2021 data from the State Administration for Market Regulation (SAMR), 95% of China's 42 million enterprises are SMEs. They play a key role in socioeconomic development and serve as the foundation for modern economies. To help SMEs reduce OPEX and become digital, carriers need to provide them with high-quality private-line services.

The White Paper on High-quality Broadband Network Service Experience for SMEs released by the Broadband Development Alliance in 2020 reports that SMEs have specific requirements for private-line quality. These include guaranteed bandwidth, availability protection, hard service isolation, provisioning and recovery time, and SLA visibility (see Table 1). SMEs are also price-sensitive due to OPEX pressures, so they need inclusive OTN private-line services.

Major issues with conventional private line services

SMEs are mostly found in commercial buildings. Their tremendous need for private lines makes commercial buildings a key area of competition for carriers. However, conventional private-line services for commercial buildings face the following issues: 1) A shortage of optical cables in most buildings: Deploying backbone optical fibers in buildings is costly and approval is difficult. 2) Slow delivery: To minimize costs, enterprises need private lines as fast as possible, ideally on the same day as the application or the next day. However, conventional private lines are typically delivered in about 30 days. This creates great pressure on carriers. 3) Homogeneous competition: PON-based delivery is widely used for private lines in buildings, and the backhaul network is

Type	SLA Index	3A	2A	1A	0A (Internet)
Network indexes	Guaranteed bandwidth	Committed bandwidth = 100% package bandwidth	Committed bandwidth = 100% package bandwidth	Committed bandwidth = 100% package bandwidth	Not committed
	Reliability	Availability: 99.95%	Availability: 99.9%	Availability: 99%	Depends on Internet private line
		Gateway-side uplink dual-protection	N/A	N/A	N/A
		Hard service isolation	Hard service isolation	Soft service isolation	No isolation
Service indexes	Provisioning time	24 hours	48 hours	72 hours	Depends on Internet private line
	Restoration time	< 4 hours	< 8 hours	< 24 hours	Depends on Internet private line
	SLA operation	Network-wide quality indicator visibility	Network-wide quality indicator visibility	N/A	N/A
		Network operations quality report	Network operations quality report	N/A	N/A

Table 1: Private line network and service indicators from the *White Paper on High-quality Broadband Network Service Experience for SMEs*

based on IP MAN. During peak daytime hours, network congestion often occurs due to heavy traffic, resulting in a poor user experience. Carriers need to focus on user experience to gain differentiated competitive advantages.

4) Low O&M efficiency and a lack of value-added services: The PON+IP transport solution separates service provisioning and O&M into two segments, making fast and integrated service provisioning impossible. As a result, fault location and demarcation are complex and inefficient, value-added intelligent services such as service visibility and adjustable bandwidth cannot be provided, and services are not intelligently scalable.

OTN P2MP private lines enable all-optical digital commercial buildings

In all-optical digital commercial buildings that have deployed the OTN point-to-multipoint (P2MP) solution, SMEs can access premium, premium cloud-access, and Internet private-line services. In the OTN P2MP private-line solution, the OLT and OTN share a site, directly connecting the gigabit access network to the all-optical network. Services are no longer transmitted through the MAN like in conventional PON private lines. This solves problems such as unguaranteed service bandwidth and high delay variation caused by congestion in the MAN during peak hours. The OTN P2MP solution uses point-to-multipoint architecture based on a passive optical splitter at the end of the ODN. This realizes 5A-quality OTN enterprise private lines, while retaining the wide coverage of the ODN, meeting SLA demands of SMEs for bandwidth, stable latency, and quick provisioning, while also providing service quality differentiation.

The OTN P2MP solution delivers the following advantages:

1. Fewer optical fibers required

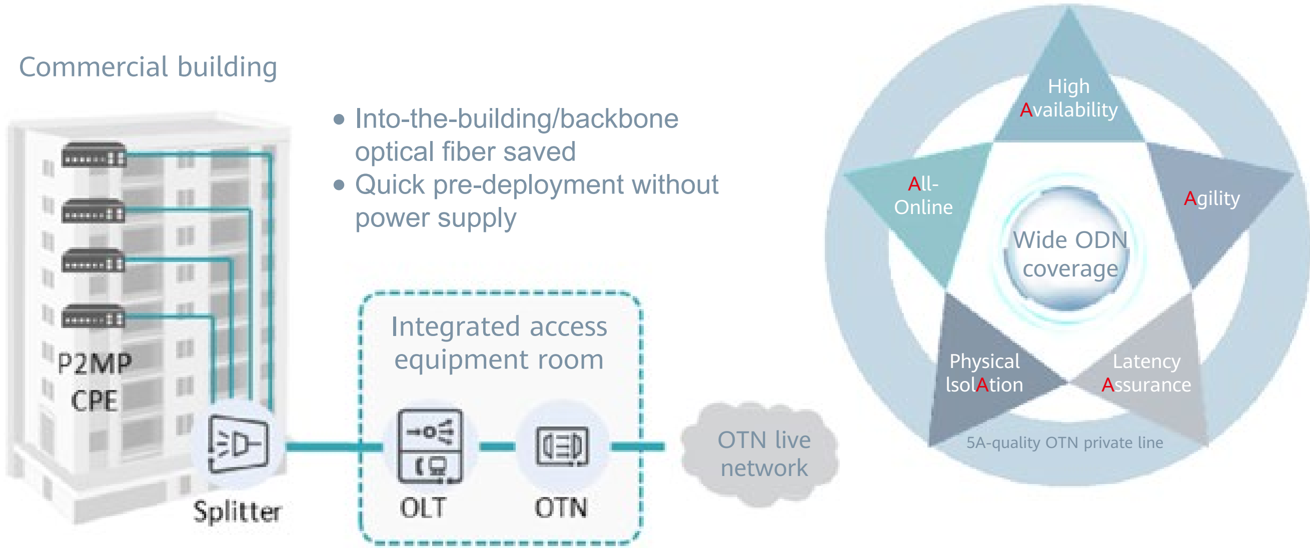


Figure 1: OTN P2MP solution and its innovations

To ensure service quality for enterprise customers, typical private line solutions use direct optical fiber connections, resulting in great demand for backbone optical cables. The OTN P2MP private line solution requires only one optical fiber to connect to a commercial building and uses 1:16 optical splitting at the end, greatly reducing the use of fibers in the backbone cable. As the OTN P2MP solution reuses existing ODN networks, far fewer optical fibers are required, saving carriers the time and cost of deploying additional cables.

2. Hard pipes

The OTN P2MP solution is the first in the industry to enable an OTN board on the OLT to allow services to be transmitted through OTN-exclusive channels. The OLT isolates regular home broadband services from private line services through network slicing to prevent interference between the two types of services.

3. Low latency

OTN P2MP uses multiple optimization technologies to ensure stable low latency for enterprise private lines.

(i) Dual-wavelength isolation for reducing delay variation

The dual-wavelength isolation mechanism separates the wave channels for services and administration. The service wave channel transmits data and the administration wave channel manages signals. This enables the independent detection of CPEs that go live, preventing delay variations caused by the ONT that detects terminal access and reducing the uplink delay variation in the access segment to less than 150 μs.

(ii) Fixed timeslot allocation for lowering latency

The fixed timeslot allocation mechanism fixes the timeslots used by each enterprise CPE for transmission and these are automatically allocated by the OLT in advance. The OLT and the CPE do not need repeated handshakes to dynamically apply for timeslots, so the highest transmission latency in between can be limited to about 200 µs.

4. Intelligent management and control

The OTN P2MP solution fully upgrades the way networks are managed and controlled and provides a northbound interface for connecting with carriers' BSSs/OSSs to deliver intelligent service capabilities.

In conventional private line solutions, PONs and OTNs are separately managed, so end-to-end management is not possible and the performance metric data of private lines cannot be centrally viewed.

With the OTN P2MP solution, users can directly access the iMaster NCE-FAN by installing cross-domain components in iMaster NCE-T, realizing the centralized management and control of PONs and OTNs through smooth upgrades. The iMaster NCE intelligent management and control system supports the cross-domain management and control of transport OTN and access PON devices. It also delivers functions such as quick provisioning, SLA visibility, and the self-service bandwidth adjustment of OTN P2MP private line services. The resulting self-service experience is similar to online shopping, and includes temporary network acceleration for livestream events, checking private line quality, and adjusting speeds.

5. Diversified protection

The OTN P2MP solution provides a range of protection mechanisms such as Type B, SNCP, and OTN ASON, realizing 99.99% reliability.

OTN P2MP: The best inclusive premium private line solution for carriers

On July 7, 2021, China Unicom Guangdong and Huawei jointly launched the world's first OTN P2MP premium private-line service in the Shenzhen-Dongguan metropolitan area. The service will bring high-quality connectivity to industries and help boost the digital economy in the Guangdong–Hong Kong–Macau Greater Bay Area. China Unicom Jiangsu also launched the OTN P2MP premium private-line service in Suzhou in the same month.

On September 10, 2021, China Mobile Guangdong and Huawei jointly launched the first OTN P2MP cloud-access private line in China. This service facilitates the digital and cloud transformation of SMEs in China, and enriches China Mobile's enterprise private line portfolio.

In September 2021, China Telecom Shandong and Huawei jointly launched China Telecom's first premium private line, providing cloud-based video backhaul private-line services for Shandong Land Development Group. In October 2021, China Telecom Sichuan also launched the OTN P2MP private-line service. Verified performance metrics, such as latency and jitter, demonstrate the solution's ability to meet the premium experience needs of enterprise customers.

As the digital economy thrives, equipment providers like Huawei are committed to strengthening their technical solutions to support information-based transformation and bring convenience, a premium experience, and high-quality services to SMEs.

The OTN P2MP private line solution requires only one optical fiber to connect to a commercial building and uses 1:16 optical splitting at the end, greatly reducing the use of fibers in the backbone cable.

A Hospital's Journey to Become Intelligent and Low-carbon

In the digital era, hospital infrastructure must adapt to two trends: emerging medical technologies and new business models. Beijing Tongren Hospital has deployed all-flash data centers to transform from digital to intelligent, save energy, and reduce carbon emissions.

By Liu Yanting, CPC Committee Member and Director of the Information Center, Beijing Tongren Hospital Affiliated to Capital Medical University
By Zhang Yang, Marketing Expert, Data Storage and Intelligent Vision Product Line, Huawei

Data centers: The key to making hospitals intelligent

In June 2021, China's State Council issued guidelines on the high-quality development of public hospitals, highlighting the role of IT in this process and calling for efforts to further integrate next-gen IT, such as cloud computing, big data, IoT, blockchain, and 5G, into medical services.

Driven by data-centricity, new hospital services are constantly increasing requirements on IT, as intelligent healthcare, management, and services continue to develop.

Intelligent healthcare: As well as meeting the basic service requirements of healthcare professionals, hospital IT systems can combine medical knowledge with IT, like big data and AI, to provide hospital staff with functions such as data queries, medical knowledge reminders, and guidance on clinical decisions.

Hospital operations and management: With intelligent IT systems, hospital management can view the statistical reports of different departments and acquire a full view in real time at multiple levels and from multiple perspectives through indicators such as inputs, outputs, and cost. Big data analytics and AI provide deeper insights into daily activities and enable an agile response during hospital operations and management by supporting data-driven decision-making.

Intelligent services: An intelligent hospital covers end-to-end services ranging from booking appointments, telemedicine, mobile payments, and medicine delivery to AI-powered health management and intelligent consultations.

Internet-based medical services have tremendous development potential and data centers are therefore vital to new hospital infrastructure. Building new hospital data centers to unleash the potential of new infrastructure is crucial to realizing intelligent, green, and low-carbon hospitals in the new era.

So how has Beijing Tongren Hospital managed to transform itself from digital to green and intelligent?

Two major challenges

Founded in 1886, Beijing Tongren Hospital Affiliated to Capital Medical University is a large tertiary hospital that focuses on ophthalmology (eye disorders and treatment) and otorhinolaryngology (ear, nose, and throat). It has 68 clinical and technical departments and handles 2.9 million outpatient and emergency visits every year.

The hospital understands the value of next-generation IT in development and management. It is working to build an intelligent hospital system that integrates healthcare, services, and management, including maintaining electronic medical records (EMRs), and enable intelligent management. After transformation, it hopes to provide patients with more efficient, secure, convenient, and high-quality services.

Beijing Tongren Hospital's infrastructure needed to adapt to two trends on the

Building new hospital data centers to unleash the potential of new infrastructure is crucial to realizing intelligent, green, and low-carbon hospitals in the new era.

The hospital's legacy information system had outdated architecture, poor reliability, and was vulnerable to single point of failure risks.

path to becoming intelligent: new medical technologies and new business models.

The hospital's legacy information system had outdated architecture, poor reliability, and was vulnerable to single point of failure risks. The storage arrays still used hard disk drives (HDDs), seriously compromising the user experience of core service systems and hindering the launch of new services such as EMRs and mobile apps. The hospital's gradual launch of more than 30 systems, including the Hospital Information System (HIS) and the Picture Archiving and Communication System (PACS), greatly increased the number of physical devices, which created numerous issues. For example, resources were not centrally managed, resource utilization was low, management and maintenance were complex, and O&M costs were increasing every year.

Second, new business models, such as hierarchical medical care, medical big data, and Internet-based healthcare services, were driving change across the entire healthcare industry, with the wide adoption of cloud computing, big data, AI, IoT, and Internet technologies increasingly seen in hospitals.

Building intelligence with all-flash data centers

To address these challenges and transform into an intelligent hospital, Beijing Tongren Hospital came up with a top-level design for its information system. The system architecture mainly covers six layers: portal, platform application, resources, hospital information platform, application system, and intelligent infrastructure. Based on this system design, Beijing Tongren Hospital made four transformations:

First, service data: The scope of service data expanded from inpatients and research to include outpatients.

Second, app access: The scope of apps expanded from medical staff to patients.

Third, network architecture: Traditional intranet architecture was expanded to an open Internet interface.

Fourth, the diagnosis and treatment model: The model switched from outpatient diagnosis and treatment to hierarchical medical care or AI-assisted diagnosis and treatment.

Data storage is the foundation of any intelligent hospital and data centers are the key to intelligent transformation. All-flash data centers support data acceleration, enable intelligent O&M, and reduce carbon emissions, allowing them to play a vital role in helping Beijing Tongren

Hospital become intelligent.

During its transformation, the hospital focused on three areas:

Performance improvement

Beijing Tongren Hospital migrated all its legacy and outdated core services and data, such as HIS and PACS, to high-end flash storage systems. Doing so improved the performance of the intelligent service system and addressed the performance bottlenecks of offline services. The hospital also loaded the cloud resource pools and databases of new services, such as Internet access and registration, onto new high-end flash storage resource pools.

The solution enables new services to be launched and expanded and facilitates data splitting for online services. By improving the storage performance of the overall intelligent service system, Beijing Tongren Hospital has also greatly shortened outpatient registration time and improved the work efficiency of hospital staff.

The hospital’s storage resource pool supports elastic capacity expansion, meaning more hospital services can be launched in the future as intelligent services continue to experience rapid development in hospitals.

Data protection

In today's big data era, service data should be stored through flash storage acceleration, while hospital data must be protected in all scenarios to enhance business continuity and security.

Beijing Tongren Hospital uses Huawei's OceanStor all-flash storage systems to support active-active disaster recovery for databases and the virtualization layers of core services such as HIS and PACS. This eliminates single point of failure risks while the hospital's core services run 24/7 without interruption. Continuous Data Protection (CDP) and the remote replication functions of the all-flash storage systems allow important data to be stored on backup devices, ensuring zero data loss. The disaster recovery management software OceanStor BCManager implements end-to-end disaster recovery management and one-click drills in all scenarios, greatly reducing the workload of O&M staff in the hospital.

Security and reliability

Beijing Tongren Hospital's data protection policy covers the full disaster recovery of core data and hot backup of important data. It has achieved this based on Huawei's all-flash data center target architecture, with active-active configuration enabled for key service storage: HIS uses Huawei's

All-flash data centers support data acceleration, enable intelligent O&M, and reduce carbon emissions, allowing them to play a vital role in helping Beijing Tongren Hospital become intelligent.

For important hospital data, the data center uses Huawei's storage systems to implement more secure data backup and protection and reduce the backup period from one day to three seconds.

OceanStor Dorado 18000 all-flash storage active-active system, while PACS uses Huawei's OceanStor 18510F all-flash storage acceleration active-active system, which supports block/file active-active configuration. The databases and virtualization layer also use Huawei's OceanStor Dorado 18000 all-flash active-active system, while Huawei's OceanStor 5310F active-active system and other products are used in the intelligent medical education system.

Active-active storage configuration means that two all-flash arrays provide redundancy for each other. Even if one array becomes faulty, core services can continue to run 24/7 without interruption or data loss. For important hospital data, the data center uses Huawei's storage systems to implement more secure data backup and protection and reduce the backup period from one day to three seconds, providing assurance against any future risks that the hospital may face.

Intelligent and low-carbon

Hospitals are an important part of China’s stated peak carbon emissions and carbon neutrality goals. While becoming intelligent, hospitals must save energy and reduce carbon emissions, a key starting point for which is to reduce carbon emissions from data centers, which are highly energy-intensive.

According to forecasts by think tanks in the digital power industry, the power consumption of data centers worldwide will increase from 670 billion kWh in 2020 to 950 billion kWh in 2025, accounting for 3% of all power consumed globally.

So how can hospitals build efficient, low-carbon data centers?

All-flash data centers are the first step. An all-flash data center uses solid state drives (SSDs) for at least 90% of its storage capacity, while delivering high-density, high-reliability, low-latency, and high-energy efficiency. At the same capacity, flash drives use 70% less power and require 50% less space than HDDs. It is estimated that the annual CO_2 emissions reduced by replacing a single HDD with a flash drive are equivalent to the annual emissions absorbed by 150 trees, and that new all-flash data centers can reduce power consumption by 21%.

Since becoming intelligent, Beijing Tongren Hospital has been able to provide patients with more efficient, secure, convenient, and high-quality services. Its adoption of all-flash data centers to build a green foundation has proven to be a valuable exploration into how to achieve peak carbon emissions and carbon neutrality in the healthcare industry.

AirEngine AP & IdeaHub: The Next Step in Collaborative Digital Offices

Powered by gigabit wireless networks, Huawei's AirEngine AP + IdeaHub digital office solution provides a superior experience for working anytime, anywhere and enabling seamless collaboration across geographic locations and devices.

Xu Dimiao, IdeaHub
Marketing Expert,
Intelligent Collaboration
SPDT, Huawei

Yuan Yunlian,
Campus Network Senior Marketing
Manager, Data Communication
Product Line, Huawei

Figure 1: Huawei AirEngine AP and IdeaHub

Offices are the engines that drive enterprise innovation.

As recorded in the biography *Steve Jobs*, it is hard not to be impressed by how Jobs founded Apple in a garage. This is not simply admiration for the creator of a legendary enterprise, but an aspiration for a working environment where we can closely collaborate, work in the ways we want, and create infinite possibilities.

How can we make office work as efficient as a garage-based partner start-up? To find the answer, we need to know what the pain points of today's offices are:

Network: Online offices, remote collaboration, and device intelligence are becoming increasingly common. But without effective networks, we cannot work effectively. Networks that can cover all smart devices and enable office work anytime, anywhere are therefore crucial for modern enterprises.

Devices: Conventional office devices and tools reduce work efficiency, requiring the shift to a digital-led approach.

Application: Seamless communication is central to successfully conducting business, with today's world requiring far more from remote collaboration powered by high-quality connectivity.

Huawei believes that developing ICT infrastructure for intelligent collaboration and device-network convergence can enable digital offices and improve work efficiency and creativity. That is the thinking behind the two digital tools AirEngine Wi-Fi 6 AP and IdeaHub.

AirEngine Wi-Fi 6 AP: Gigabit connections anytime, anywhere

Hot-desking depends as much on the company's network as on corporate or personal preference. As well as freeing offices from cables, AirEngine Wi-Fi 6 AP helps enterprises build high-quality office networks with full coverage, zero dead spots, zero packet loss, and zero delays.

Strong signals everywhere

How does the AirEngine AP extend its reach to every corner? Thanks to intelligent antenna algorithms and beamforming technology, the AirEngine AP increases signal strength by 100% and extends coverage by more than 20%, ensuring signal coverage everywhere.

The solution’s dynamic-zoom smart antenna can sense dynamic changes in the density of mobile office terminals accessing the network. When low, the AP works in wide-angle mode by default to improve signal coverage. When high, the AP reduces the signaling angle of its antenna to reduce signal boundary interference, increasing the access rate for users by 20% and solving the problem of poor experience for high-density indoor access.

Intelligent roaming ensures always-online mobile office

The AirEngine AP's intelligent roaming technology ensures uninterrupted roaming by learning the behavior of each device, enabling office workers to work seamlessly on the go. The dual-fed and selective-receiving technology greatly improves the stability and reliability of Wi-Fi connections and prevents disconnections regardless of location or usual dead spots like corridors, elevators, and stairs. The AP can also easily cope with quickly moving devices, giving a worry-free office Wi-Fi experience.

Intelligent multimedia scheduling and Wi-Fi & IoT convergence

The AirEngine AP's intelligent multimedia scheduling algorithm automatically identifies multimedia applications, dynamically senses latency and adjusts bandwidth, and provides private line-level access assurance while delivering unlimited user rates. With built-in containers supporting software-defined IoT (SD-IoT), the AirEngine AP also provides IoT device connection and sensing capabilities. This enables applications such as door access through facial recognition, automatic lighting and temperature adjustments, scheduling meetings with QR code scans, and one-touch projection.

IdeaHub creates a one-stop collaborative digital office

More often than not, work projects are collaborative, which in turn greatly benefits from digital tools. The Huawei IdeaHub is an intelligent office terminal designed to enable teamwork. Integrating HD projection,

The AirEngine AP's intelligent roaming technology ensures uninterrupted roaming by learning the behavior of each device, enabling office workers to work seamlessly on the go.

intelligent writing, cloud conferencing, and intelligent applications, it enables the free flow of information, efficient team collaboration, and quick access to applications.

Integrated super office terminal

IdeaHub integrates a range of devices including a projector, whiteboard, microphone, speakers, and conference terminals. Light in weight and with a unibody design, it supports wireless access and serves as a mobile conference room whenever needed.

Smart devices such as mobile phones, computers, and tablets can share the screen with IdeaHub with one click. Users can make notes with ease, write smoothly, and share information through QR code scans, allowing them to enjoy the convenience of a paperless office and the free flow of information in the most natural way.

IdeaHub integrates professional audio and video capabilities with scenario-based intelligent technologies to support intelligent noise reduction and intelligent view, enabling teams that are thousands of miles apart to hear and see each other clearly.

An intelligent assistant for office collaboration

IdeaHub integrates professional audio and video capabilities with scenario-based intelligent technologies to support intelligent noise reduction and intelligent view, enabling teams that are thousands of miles apart to hear and see each other clearly. IdeaHub also provides functions such as intelligent sign-in, real-time captioning, translation, and meeting recordings to improve meeting efficiency.

IdeaHub has a built-in office application market and can be quickly integrated with enterprise office automation (OA) and more service applications to serve different functions, including daily operations, command capabilities, and technical demonstrations.

Powerful tools combined: Intelligent collaboration

The potential of the AirEngine AP and IdeaHub can be fully unlocked through intelligent collaboration and device-network integration to create a digital office space with a superior experience.

When an IdeaHub is connected to an AirEngine AP, it is identified as a smart screen device that enables intelligent collaboration.

The AirEngine AP preferentially schedules, accesses, and reserves bandwidth for IdeaHub, and automatically optimizes the roaming of devices using screen casting. Moreover, continuous network packet loss measurements and loss rates can be viewed.

The integration of the AirEngine AP and IdeaHub and the collaborative innovation in both products have created digital offices with ubiquitous gigabit access for seamless remote collaboration. The solutions can help accelerate global transformation towards efficient and collaborative digital offices.

Building a Fully Connected, Intelligent World

All-Optical Smart City

Gigabit optical networks build an all-optical base for smart cities