HUAWEI

5G MEC IP Network White Paper

2020

Carrier Network for Edge Computing in the 5G Era: Challenges and Advice

Abstract

Digital transformation of industries is booming around the world, with digitalization being the foundation, network-based connectivity being the support, and intelligence being the goal. The industry intelligence era is here. Intelligence technologies have been applied in various industries, such as manufacturing, electric power, transportation, healthcare, and agriculture.

The Multi-access Edge Computing (MEC) system is a distributed open platform that converges and moves network, computing, storage, and application capabilities to edge nodes that are close to things and data sources. This platform delivers intelligence-enabled services at the network edge to meet the key requirements of industry digitalization in terms of agile connection, real-time service processing, data optimization, application intelligence, and security and privacy protection. It is estimated that by 2022, more than 50% of enterprise-generated data will be created and processed at the edges outside DCs or clouds.

5G provides a good network foundation for the development of the edge computing industry, exemplified by support for the three major scenarios (eMBB, URLLC, and mMTC), flexible deployment of core network user plane functions, and network capability openness.

"5G + MEC + AI" is the key for 5G to better enable various industries at the network edge. It is a new model used by carriers to help vertical industries achieve digital and intelligent transformation. It provides opportunities and key scenarios for carriers to enter vertical industries and serves as an important indicator of whether 5G applications are successful.

5G MEC moves cloud computing and the 5G core network to the network edge, bringing new traffic models and deployment models. If carriers continue to use the design idea of 4G mobile transport networks, they will face edge computing challenges in the 5G era. How can these challenges be overcome and 5G MEC-ready networks be built? This is a problem that must be solved by carriers during network planning.

This document analyzes the four major challenges brought by 5G MEC and six key points in 5G MEC network planning and provides the related advice and reference network model.

Contents

Figures

1 Edge Computing: New Model for Carriers to Help Vertical Industries Achieve Digital and Intelligent Transformation in the 5G Era

1.1 Edge Computing in the Industry Intelligence Era

Digital transformation of industries is booming around the world, with digitalization as the foundation, network-based connectivity as the support, and intelligence as the goal. Digitalization is to generate data from people, things, environments, and processes through digitalization, enable valuable data flow through network-based connectivity, and use data as an essential production factor for creating both economic and social value in various industries through intelligence. Intelligence enables intelligent decision making and operations through intelligent data analysis and achieves continuous, intelligent optimization of business processes through closed-loop control.

Intelligence technologies represented by big data, machine learning, and deep learning have been applied in various fields, such as speech recognition, image recognition, and user profiling, and have made great progress in terms of algorithms, models, and architectures. The industry intelligence era is here. Intelligence technologies have been applied in various industries, such as manufacturing, electric power, transportation, healthcare, and agriculture.

The MEC system is a distributed open platform that converges and moves network, computing, storage, and application capabilities to edge nodes that are close to things and data sources. This platform delivers intelligence-enabled services at the network edge to meet the key requirements of industry digitalization in terms of agile connection, real-time service processing, data optimization, application intelligence, and security and privacy protection. It connects the physical and digital worlds by enabling smart assets, gateways, systems, and services.

Edge computing is regarded as an important link between 5G and systems such as the Industrial Internet and IoT, and is expected to bring more disruptive service models. It is estimated that by 2022, more than 50% of enterprise-generated data will be processed outside DCs or clouds, 20% of new industrial control systems will possess analysis and AI edge inference capabilities, and at least 50% of high-end industrial IoT gateways will provide optional 5G modules. The European Telecommunications Standards Institute (ETSI) established the Industry Specification Group (ISG) for Mobile Edge Computing in 2014 to focus on carrier network edge computing standards and industry enablement. In 2016, the ISG renamed Mobile Edge Computing as Multi-access Edge Computing. In the carrier field, MEC is generally used to refer to edge computing systems. In the 5G era, MEC provides a new way of application for carriers to help vertical industries achieve digital and intelligent transformation.

1.2 Device Models and Values of Edge Computing

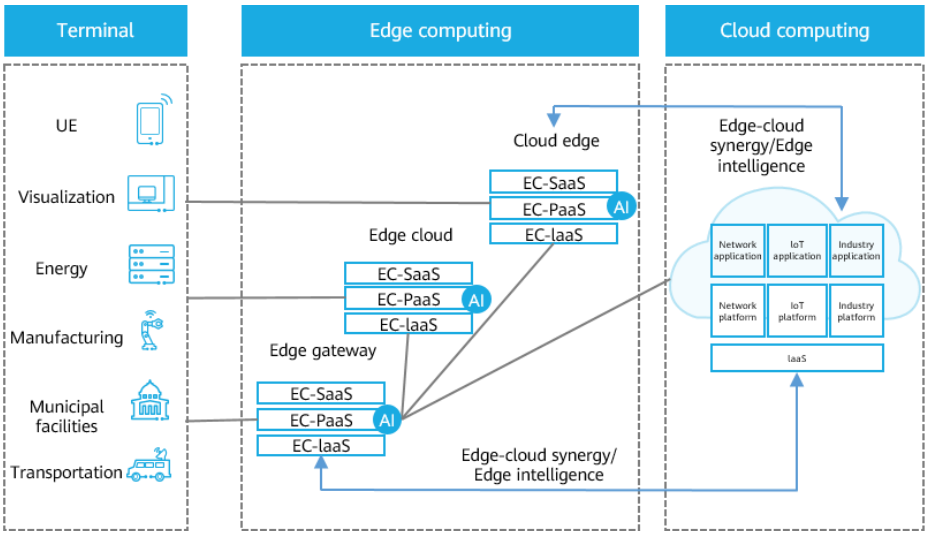
1.2.1 Device Models

Edge computing is essentially the extension and evolution of cloud computing at edge nodes outside DCs, with core capabilities focusing on edge-cloud synergy and edge intelligence. The introduction of cloud computing concepts, architecture, and technologies must be considered for software platforms to provide E2E, real-time, and collaborative intelligence, reliability, and dynamic reconfiguration capabilities, and heterogeneous computing capabilities must be considered for hardware platforms, such as Kunpeng, Ascend, ARM, x86, GPU, NPU, and FPGA.

The Edge Computing Consortium (ECC) classifies the implementation of edge computing systems into three types: cloud edge, edge cloud, and edge gateway, as shown in Figure 1-1.

- Cloud edge: Edge computing is an extension of cloud services at the network edge. Logically, it is still a cloud service. Its capability provisioning depends on cloud services or requires collaboration with cloud services. It is mainly used in public cloud scenarios, such as the IEF solution provided by HUAWEI CLOUD and the IoT Greengrass solution provided by AWS.
- Edge cloud: Edge computing provides small- and medium-scale cloud service capabilities at the network edge. The edge service capabilities are mainly provided by the edge cloud. Edge cloud resources are managed and scheduled by systems deployed in the central cloud. Carrier MECs and CDNs are application examples of edge clouds.
- Edge gateway: The original embedded gateway systems are reconstructed with cloudification technologies and capabilities. Edge gateways provide capabilities such as protocol/interface conversion and edge computing at the network edge. The controller deployed in the cloud provides capabilities such as resource scheduling, application management, and service orchestration for edge nodes. Edge gateways are mainly applied to scenarios such as industrial Internet and Internet of Vehicles (IoV).

Figure 1-1 Implementation modes of edge computing



1.2.2 Edge Computing Values and Features (CROSS)

- Massive number of **C**onnections

  Networks are the cornerstone of system interconnection and data collection and transmission. As the number of connected devices surges, flexible network expansion, low-cost O&M, and reliability assurance are facing great challenges.

- **R**eal-time services

  Industrial system detection, control, and execution and emerging VR/AR applications pose high real-time performance requirements. In some scenarios, the service latency is required to be less than 10 ms or even lower. If data analysis and processing are all conducted on clouds, the real-time requirements of services will fail to be met, severely affecting the service experience of end users.

- Data **O**ptimization

  Currently, industrial sites and IoT endpoints contain a large amount of heterogeneous data. Data optimization must be conducted for centralized data aggregation, presentation, and openness, so that the data can serve intelligent edge applications in a flexible and efficient manner.

- **S**mart applications

  Service process optimization, O&M automation, and service innovation drive applications to become intelligent. Edge intelligence can bring significant efficiency and cost advantages.

- **S**ecurity and privacy protection

  Security, which is critical to cloud and edge computing, requires E2E protection. The network edge is close to IoT devices, making access control and threat prevention extremely difficult. Edge security covers device, network, data, and application security. The integrity and confidentiality of key data, such as mass production data and personal data, are also the focus of security protection.

1.3 MEC, New Opportunity for Carriers to Enter Vertical Industries in the 5G Era

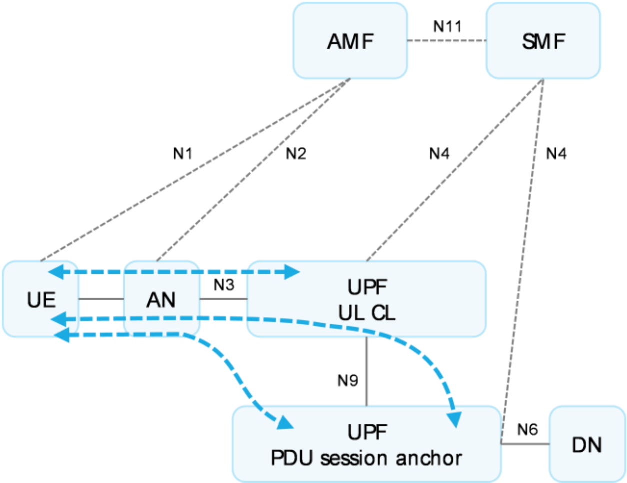
MEC enables carriers to distribute services at the network edge. MEC-empowered E2E solutions enable carriers to provide services with lower latency, higher bandwidth, and lower costs, quickly respond to user requests, and improve service quality. MEC enables carriers to provide high-quality services closer to users and even to enterprise campuses, further promoting the in-depth convergence of carriers' communication networks and enterprise services to improve network values.

5G provides a solid network foundation for the development of the edge computing industry, exemplified by support for the three major 5G scenarios, flexible deployment of user plane functions, and network capability openness.

The three major 5G scenarios are closely related to edge computing. URLLC, eMBB (especially super uplink), and mMTC support edge computing scenarios with diversified requirements, such as industrial control scenarios that have extremely high requirements on latency, AR/VR and live broadcast scenarios that have relatively high requirements on bandwidth, and emerging service (such as IoT) scenarios that require massive numbers of connections. In addition, to meet the continuity requirements of mobile services, the 5G network introduces three service and session continuity modes to ensure user experience in different scenarios, for example, in Internet of Vehicles (IoV) scenarios.

5G user plane functions (UPFs) can be flexibly deployed close to users for local traffic offloading. Edge computing nodes can be flexibly deployed at different network positions to meet the diversified latency and bandwidth requirements of edge computing services. The 5G core (5GC) network adopts the control and user plane separation (CUPS) architecture where the session management functions (SMFs) are decoupled from UPFs. Specifically, the 5G control plane is deployed in a centralized manner. One control plane (SMF) can manage multiple UPFs at the same time without affecting the performance of the 5GC network. The 5G user plane is deployed in a distributed manner. UPFs can be flexibly deployed at the network edge to support edge computing. Unlike the EPC network, the 5GC network can have UPFs deployed hierarchically to provide flow-based hierarchical routing capabilities. Uplink classifiers (UL CLs) can be dynamically started on the user plane as required for service traffic steering. Service traffic can be either locally offloaded or sent to the anchor UPF, and UEs are unaware of service traffic steering. UPFs deployed at the network edge can be viewed as a lightweight specialized UPF. As shown in Figure 1-2, the CUPS architecture and hierarchical UPF deployment bring great flexibility and powerful communication capabilities for 5G to support edge computing. Service traffic from a UE can be steered to local UPFs (for enterprise application traffic and other important service traffic) or directly steered to an anchor UPF (for common Internet access traffic). UL CLs can be dynamically started on the user plane for on-demand traffic splitting. Therefore, base stations in an enterprise campus can support both local enterprise applications and common Internet access applications.

Figure 1-2 5GC architecture: CUPS and hierarchical UPF deployment



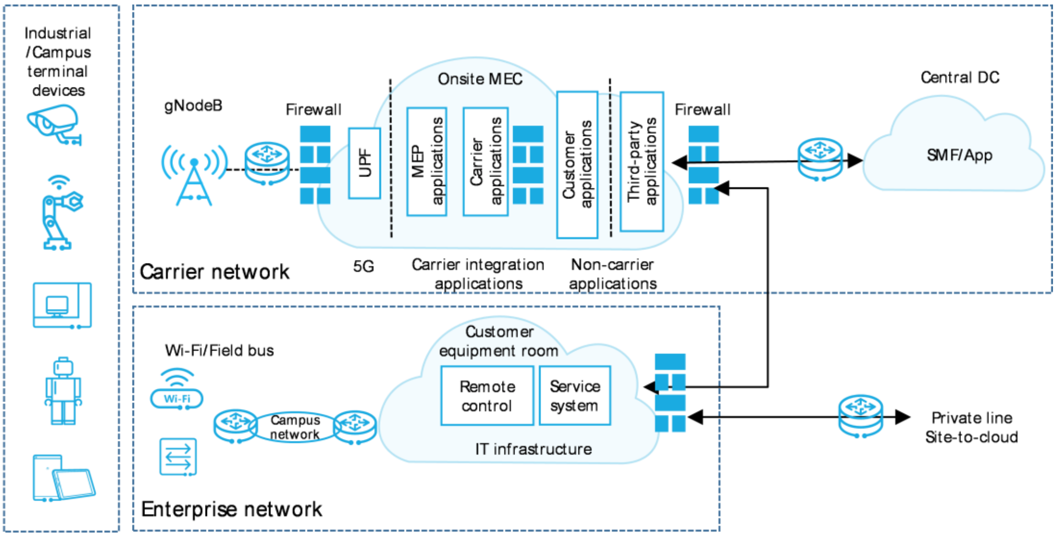
5G allows network capabilities to be opened to edge applications. Network capabilities such as wireless network information, location, and QoS services, can be encapsulated into the APIs of the edge computing PaaS and opened to applications.

The combination of 5G and edge computing offers carriers a unique advantage than before. Edge computing also serves as an important tool for carriers to use 5G to serve vertical industries and give full play to new 5G network features.

Specifically, the combination of 5G functions and features with edge computing brings the following benefits to carriers in their efforts to achieve digital and intelligent transformation:

1. UPFs are deployed on the enterprise campus (in onsite MEC scenarios, generally oriented to large enterprises, as shown in Figure 1-3). This ensures that key service data is not transmitted out of the campus and provides a low-latency transport solution. Carriers can configure an independent UPF for each enterprise user to provide customized 5G services for these users.
2. The programmable capabilities of 5G communications services (such as positioning, wireless communication, and bandwidth management) opened by carriers through APIs can be integrated into enterprises' production service systems, enabling enterprises to customize their own 5G applications.
3. The 5G MECs deployed close to users can directly interconnect with the enterprise network, enabling service systems distributed on the enterprise and carrier networks to interwork in real time. With 5G communication functions oriented to industry applications (such as URLLC, IoT mMTC, wireless super uplink, and service continuity), industries can develop many innovative applications.

Figure 1-3 Onsite MEC, New 5G Application Scenarios



5G MEC brings new service scenarios and business models for carriers to enter vertical industries. Carriers usually construct and maintain 5G MECs in enterprise campuses to provide edge cloud computing services, including IaaS, PaaS (also called MEP), and SaaS (combined with carriers' cloud computing services). This allows carriers to shift their revenue streams from pipes to software and services. Carriers can deeply explore the ICT systems and application fields of vertical industries and better serve enterprises by providing a full set of ICT services and cloud computing applications for enterprises to achieve digital, networked, and intelligent transformation. Compared with traditional enterprise private line services, these services are more comprehensive and customer-centric. This explains why carriers are actively developing 5G MEC enterprise services. 5G MEC services can help carriers attract more enterprise customers.

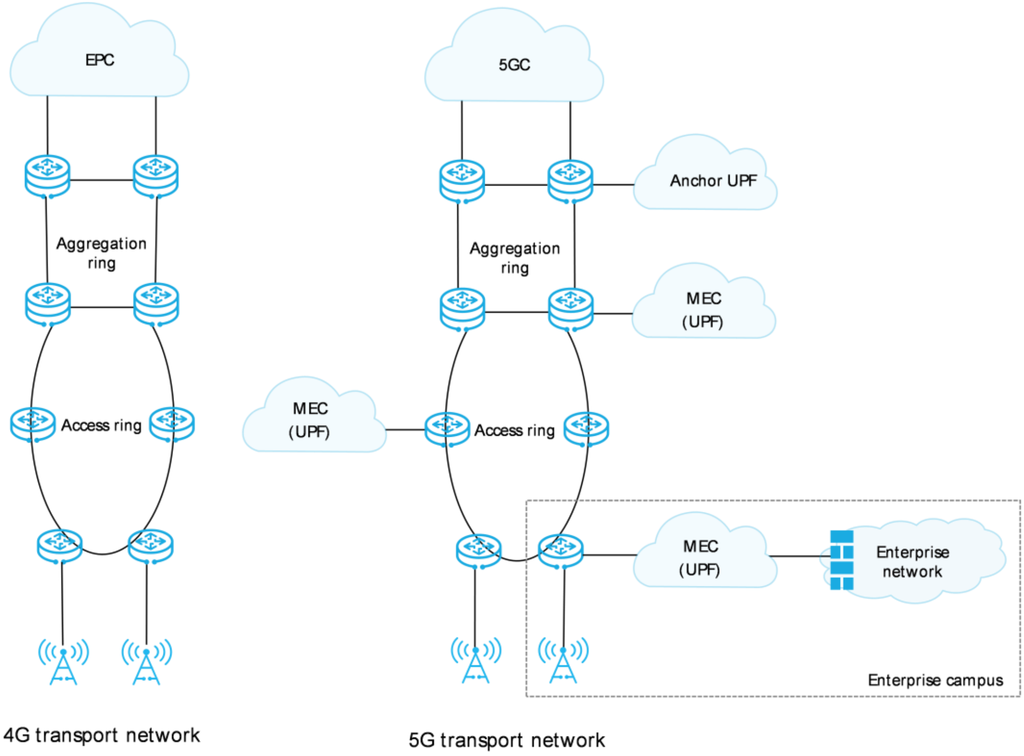
2 Edge Computing Challenges Facing Carrier Networks in the 5G Era

2.1 5G MEC Mobile Transport Network, Not a Simple Upgrade of the 4G Mobile Transport Network

The EPC networks are deployed in a centralized manner. Generally, one EPC network is deployed in one province or region. As such, the traffic model of the 4G transport network features north-south traffic. Carriers tend to use a simple access network design. For example, many carriers use the L2VPN + L3VPN networking mode, with the access network being a relatively simple L2VPN.

The 5GC network uses the CUPS architecture. The control plane is deployed in a centralized manner. Generally, one control plane is deployed in one province or region. UPFs are deployed in a distributed manner. Generally, one anchor UPF and multiple MEC UPFs are deployed in a city. 5G MEC devices can be deployed in either carriers' edge equipment rooms or enterprise equipment rooms in enterprise campuses, as shown in Figure 2-1. The distributed deployment of 5G UPFs on a mobile transport network changes the data and transport models. In the 4G era, wireless core network traffic is carried over the IP backbone network instead of the mobile transport network. In addition, 5G MECs often connect to the access network (for example, in onsite MEC scenarios), which poses additional access network requirements for the 5G mobile transport network. For detailed analysis of the UPF service flow requirements, see section 2.2.2 5GC Network Deployed Closer to Users"2.2.2 5GC Network Deployed Closer to Users." 5G MECs require a more powerful network architecture that supports enterprise services. This architecture cannot be a simple bandwidth upgrade of the existing 4G mobile transport network architecture.

Figure 2-1 Centralized deployment of EPC NEs and distributed deployment of 5G MEC UPFs



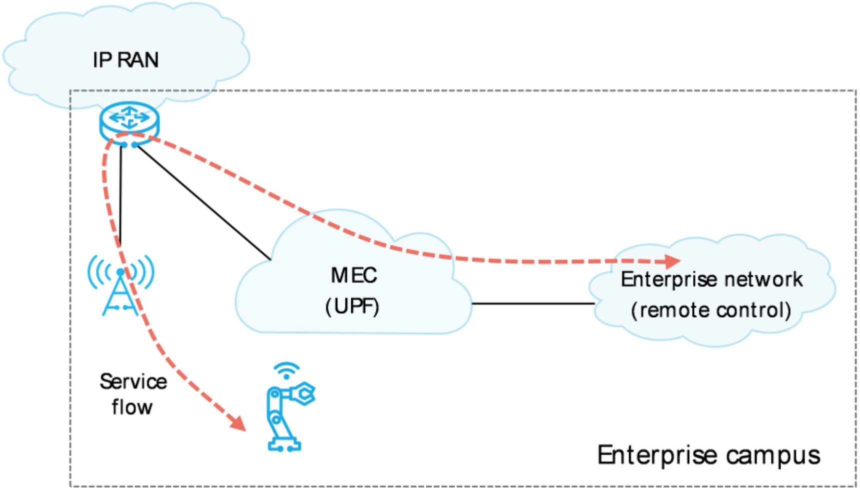
2.2 Four Major Challenges Facing 5G MECs

5G MEC brings new application scenarios and communication requirements. In the 5G era, the edge computing challenges facing carrier networks mainly come from the following new aspects.

2.2.1 Onsite MEC Scenario

Onsite MEC (deployed in an enterprise campus) is a new application scenario brought by 5G MEC. As shown in Figure 2-2, 5G MEC devices are located in the equipment room of an enterprise campus. These devices are generally installed and maintained by the carrier. Enterprises use 5G MECs for production control, remote monitoring, logistics management, and smart security protection. Many production services have strict requirements on latency. For example, the E2E control information flow latency of a remote tower crane must be less than 18 ms. In other words, traffic from production devices (such as tower cranes) must be transmitted through wireless base stations, the IP RAN, and 5G MEC to the enterprise application system (for remote control) at a low latency. The requirements for the carrier network are that the network between the 5G base stations and 5G MEC in the enterprise campus and the connection between the 5G MEC and enterprise network must have low latency.

Figure 2-2 Onsite MEC scenario of a large enterprise



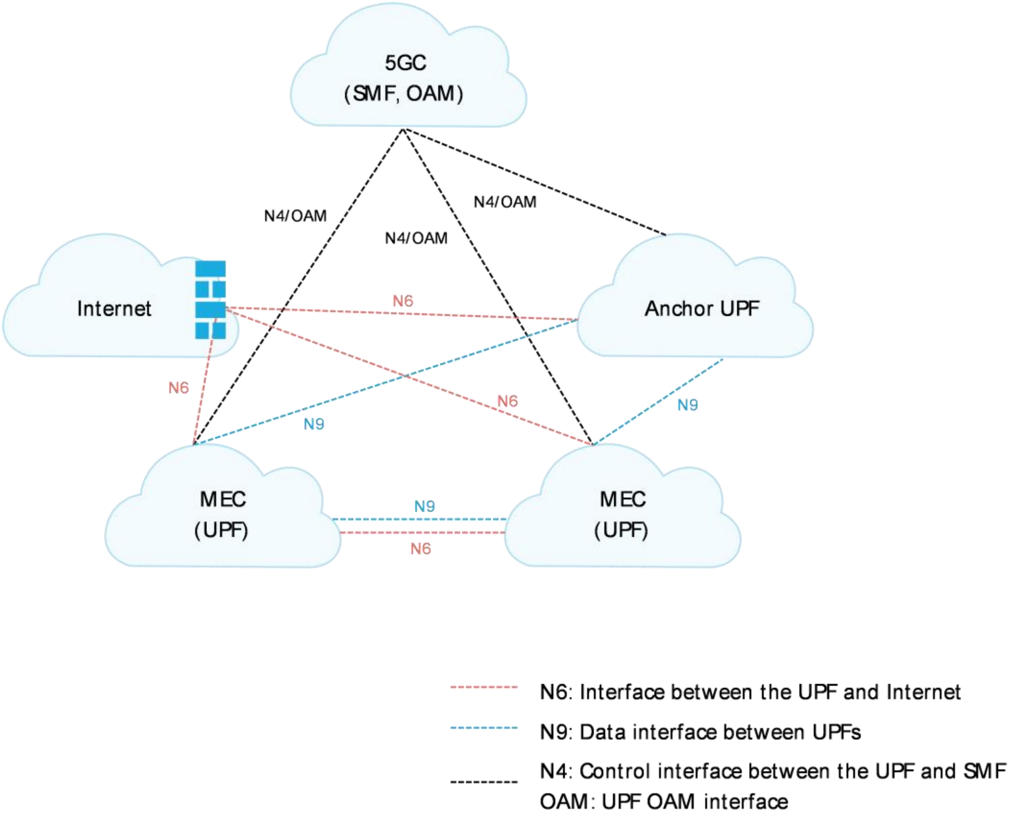
In addition, due to data security reasons, important service data of enterprises cannot be transmitted outside of the campus. Most enterprises raise this requirement during MEC pilots.

The onsite MEC scenario poses new challenges to the access network of carriers. The access network needs to provide SLA assurance in terms of latency and ensure that data is not transmitted outside of the campus.

2.2.2 5GC Network Deployed Closer to Users

As UPFs are moved downwards along with MEC, UPF-related interfaces (such as N4, N6, N9, and 5GC OAM interfaces) are moved downwards to the 5G mobile transport network. As shown in Figure 2-3, MEC UPFs need to receive control information from the SMF over N4 interfaces and receive management information over OAM interfaces. N9 interfaces are data interfaces between UPFs. They can be used between the MEC UPF and anchor UPF, and between MEC UPFs. N6 interfaces serve as the Internet data egress for UPFs. Data to the enterprise network or MEC applications is transmitted over N6 interfaces. Data from the wireless core network to the Internet is generally aggregated to a unified egress and transmitted to the Internet through a firewall. Service data flows between MECs can be transmitted over either N6 or N9 interfaces. N6 interfaces are used for communication between local MEC UPFs and the peer MEC application layer, and N9 interfaces are used for communication between local and peer MEC UPFs.

Figure 2-3 Interfaces between core network NEs in 5G MEC



The EPC network is deployed on the provincial or national IP backbone network. EPC NEs communicate through a VPN provided by the IP backbone network and do not rely on the 4G mobile transport network (IP RAN). Reliable communication for 5G UPF interfaces is a new requirement of 5G MEC for the mobile transport network (IP RAN). Some carriers use one 5GC control plane for a large area. As a result, some service interfaces (such as N4 and 5GC OAM interfaces) need to communicate across the mobile transport network and IP backbone network.

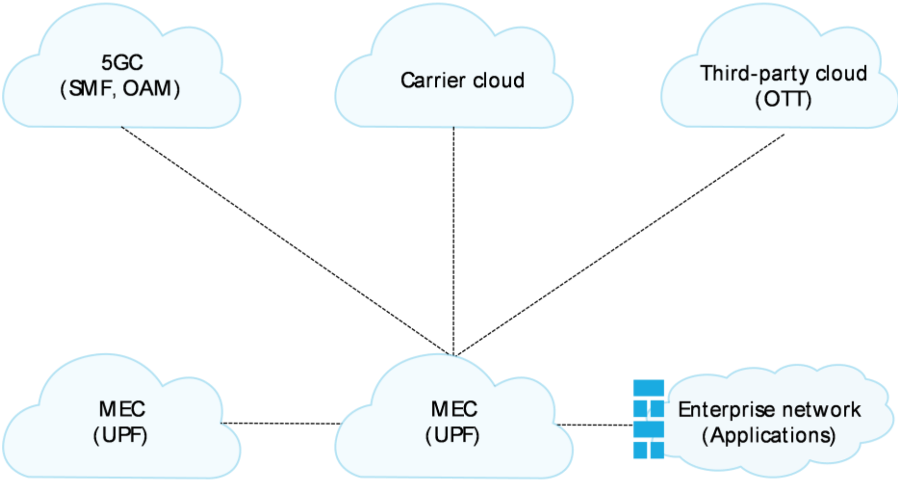
The large number of distributed UPFs and the interconnection complexity of UPF interfaces add to the complexity of the 5G mobile transport network traffic model and increase the multi-point communication coverage (generally the entire network). In the 4G era, the L2+L3 network design provides multi-point communication capabilities only above the aggregation layer. In addition, some service interfaces, such as N6 and N9 interfaces, have low latency requirements and require the transport network to provide SLA assurance.

Moving the 5G core network closer to users extends the transport scope of the wireless core network from the backbone network to the mobile transport network, posing new challenges, such as complex multi-point communication and SLA assurance, to the 5G MEC mobile transport network.

2.2.3 Cloud-Edge Synergy

5G MEC includes 5GC UPFs deployed close to users and (cloud) computing applications. 5G MEC UPFs need to communicate with the control plane and management application systems of the 5GC network in the central cloud. Some applications deployed in the 5G MEC may be part of the central cloud (such as the carrier central cloud and OTT central cloud), and some may need to collaborate with enterprise application systems or other MEC application systems to implement a complete service application. These communication connections may be established in real time on demand, and some have SLA assurance requirements. For details, see Figure 2-4.

Figure 2-4 5G MEC edge-cloud synergy and edge-edge synergy

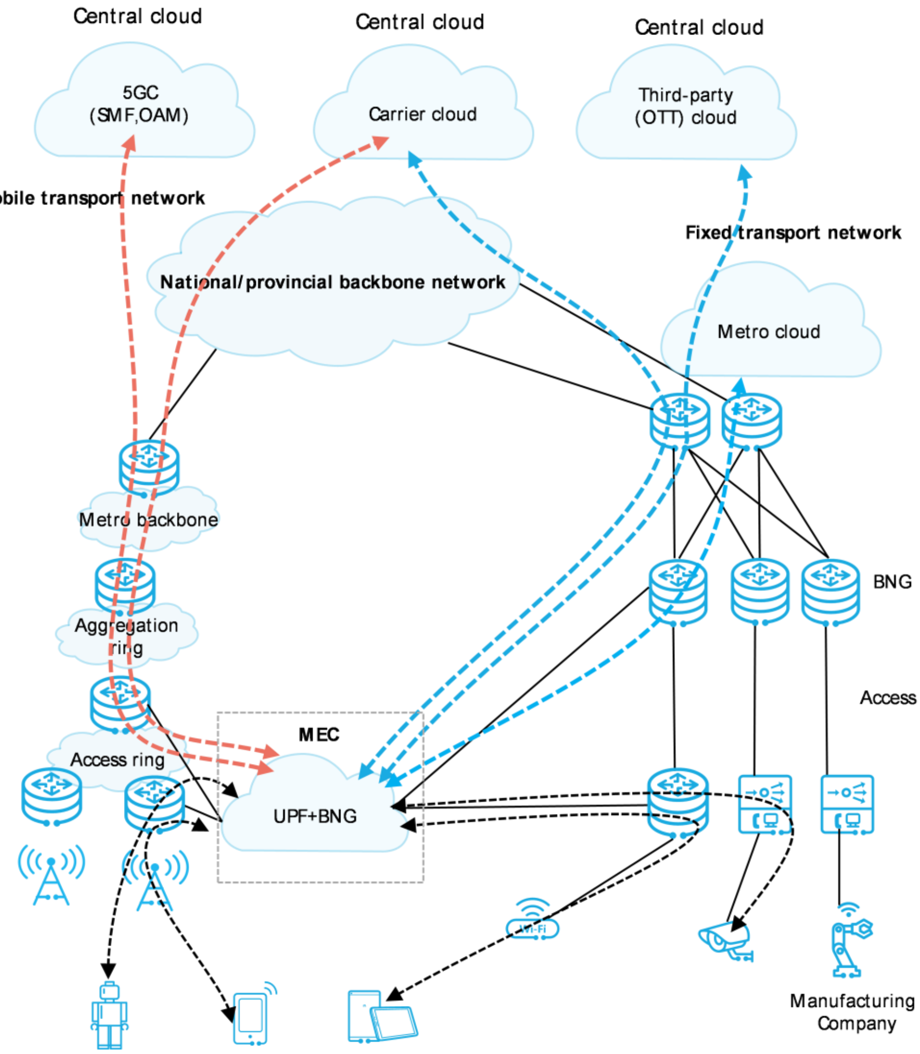


These communication requirements are new requirements brought by 5G MEC and pose new challenges to carrier networks by requiring edge-cloud synergy and edge-edge synergy.

2.2.4 FMC MEC

Devices and application systems may access the MEC over either 5G or fixed network connections, such as xPON, private line, and Wi-Fi connections. FMC MEC aims to connect all related application components to provide a complete MEC application, that is, to provide seamless FMC service, as shown in Figure 2-5.

Figure 2-5 FMC MEC



The MEC access network may be comprised of both mobile broadband (MBB) and fixed broadband (FBB) networks and require interconnection with the two network planes. The MEC communicates with central clouds (including the 5GC, carrier clouds, and third-party clouds) and service clouds (which may be deployed on a fixed metro network) over either the FBB or MBB network. Rather than being limited to the mobile transport network, the MEC connections may involve the mobile transport network, fixed network transport network, and IP backbone network.

MEC poses new challenges for FMC in terms of network architecture and interconnection. This is especially true for carriers who have both FBB and MBB networks, which is the case for the three major carriers in China.

3 Key Points of the Solution to Edge Computing Challenges in the 5G Era

3.1 ECNI's Edge Computing Network Model

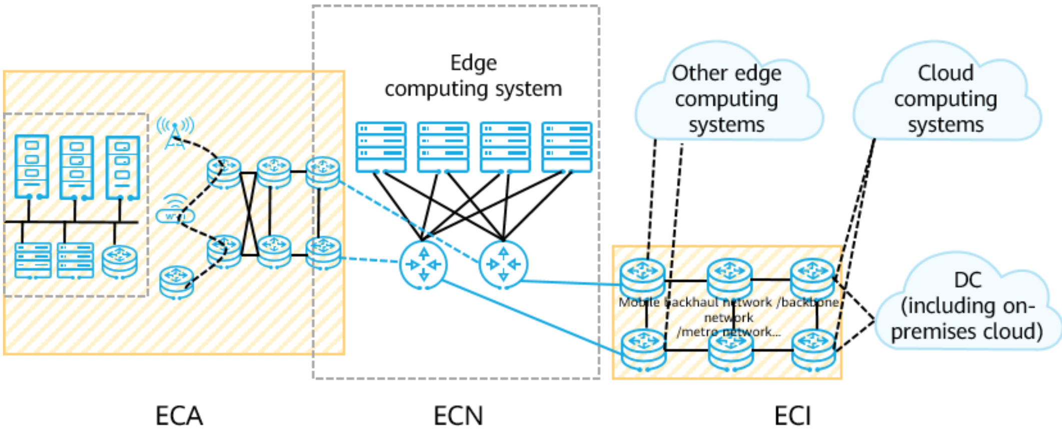
With the aim of promoting the development of edge computing networks and the edge computing industry, the Edge Computing Consortium (ECC) and Network 5.0 Industry and Technology Innovation Alliance jointly set up the Edge Computing Network Infrastructure (ECNI) SIG in 2019, which is the first industry organization to focus on edge computing networks. In November 2019, the ECNI released the *Technical White Paper on Carrier Edge Computing Networks*, the first white paper on carrier edge computing networks in the industry.

This white paper proposes a new edge computing network abstraction model that divides the edge computing network into three logical network segments, as shown in Figure 3-1.

1. Edge computing access (ECA): refers to the network infrastructure between user systems and the MEC.
2. Edge computing network (ECN): refers to the MEC internal network infrastructure.
3. Edge computing interconnect (ECI): refers to the network infrastructure between an MEC and cloud computing systems (such as public clouds, private clouds, communication clouds, and on-premises clouds), other MECs, and DCs.

The mapping between ECA and ECI varies according to where the MEC is located on the physical network, and the ECI may span multiple network segments of a carrier. As such, the ECA/ECN/ECI network model can better describe relationships between complex MEC networks and diversified physical networks. It is based on this model that the white paper describes MEC issues in subsequent sections.

Figure 3-1 Abstract edge computing network model proposed by the ECNI



3.2 Key Points of the Solution to Edge Computing Challenges

5G MEC raises new requirements and challenges that lie beyond the capabilities of the data transmission model and design solution of existing 4G mobile transport networks. In the 5G era, to deal with edge computing challenges and build a 5G MEC-ready network, carriers need to focus on the following six key points during network planning and design.

3.2.1 ECA: Shortest Paths

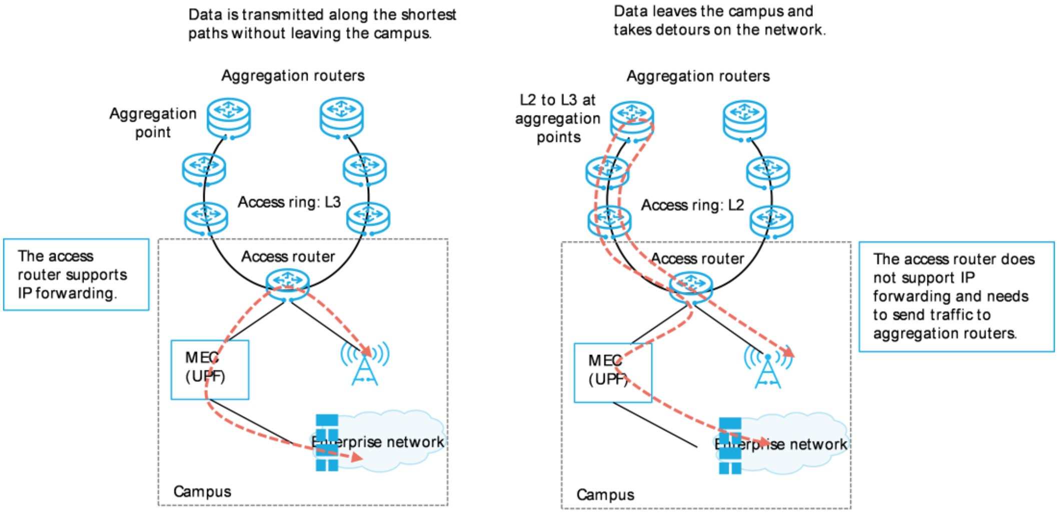
Carriers need to provide the shortest transmission paths for N3 data flows from base stations to MEC UPFs. For example, in onsite MEC scenarios, N3 data flows should be directly forwarded to MEC UPFs through mobile transport routers in the campus instead of taking detours on the carrier network.

In addition to ensuring low latency and saving bandwidth resources on the carrier network, this approach also ensures that enterprises' key service data does not leave the campus. As shown in Figure 3-2, the access router on the left can directly forward traffic from base stations at L3.

N3 data flows can then be forwarded to the nearest MEC while passing through the fewest possible NEs without leaving the campus. The access router on the right needs to forward N3 data flows from base stations to the aggregation router for IP forwarding. These data flows take a detour to the carrier network and leave the enterprise campus.

Network function architecture design is essential to the implementation of ECA shortest paths. In this design, the IP forwarding capability is moved closer to the edge (L3 to edge), following the MEC, and the access point router can forward data packets along the shortest paths at L3. Such changes to the architecture make the L2 access ring design that was used for the 4G mobile transport network unsuitable for the 5G MEC transport network.

Figure 3-2 MEC requiring a low-latency access network without traffic detours



3.2.2 ECA and ECI: Low-Latency Slices

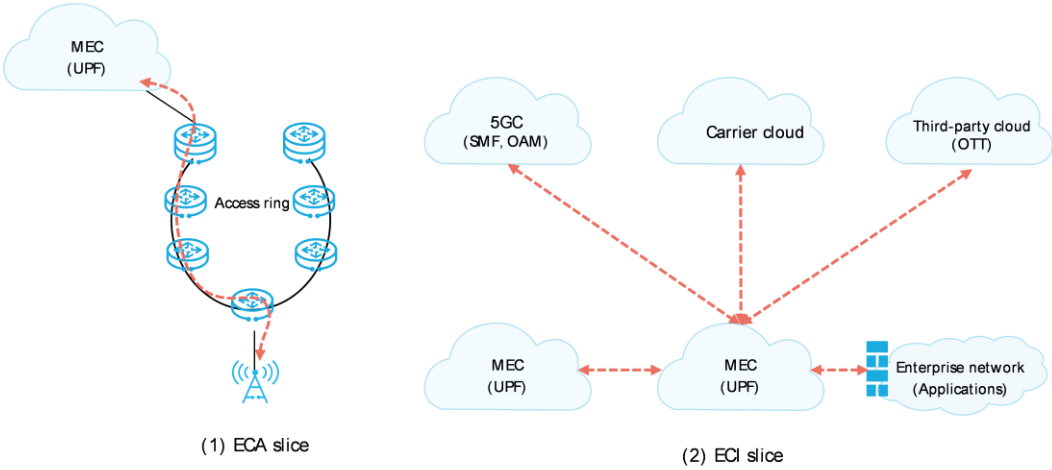
To meet the low-latency and security reliability requirements of MEC applications, some enterprise services require the carrier network to provide low-latency slice services, as shown in Figure 3-3.

ECA slice: The whole slice system includes the wireless base stations, NEs of the mobile transport network (between base stations and the MEC), and UPFs — all NEs traversed by enterprise service flows before reaching the MEC. An ECA slice involves IP network NEs, from base station to MEC.

ECI slice: The MEC needs to interconnect with the enterprise network, clouds, and other MECs across the ECI slice for SLA assurance, security, and reliability. The ECI slice may span multiple network segments.

When designing low-latency network slices, it is critical that the network slices are as small as possible, especially for an ECA network that has strict low-latency requirements. In particular, fewer NEs mean lower slice complexity and better low latency assurance. The network design requirements here are consistent with the ECA shortest path requirements in the preceding section. Complex ECI network connections and cross-network-segment transmission add more complexity to network slicing. The ECA and ECI networks can use different slicing solutions.

Figure 3-3 Low-latency ECA and ECI slices

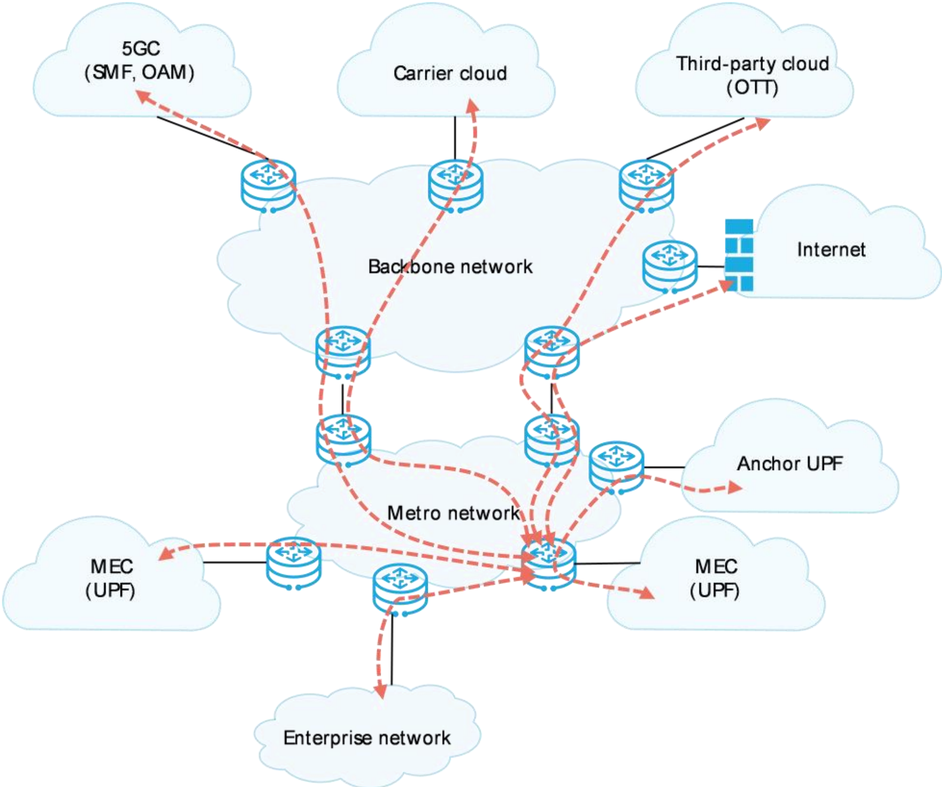


3.2.3 ECI: Flexible Multipoint Communication

The ECI network needs to support service communication between 5G MECs and the 5G core network (N4, OAM), other MECs (N6, N9), Internet egresses (N6), anchor UPFs (N9), carrier clouds, enterprise networks, third-party clouds (OTT), and the like, as shown in Figure 3-4. The ECI service flow model is complex and presents a multipoint-to-multipoint communication characteristic. The ECI network is a logical network concept from the perspective of MECs. Once mapped to a physical network, it may span multiple network segments, such as the metro network and the backbone network. The network connections of the ECI network for each MEC vary according to MEC locations and applications. To support dynamic deployment of applications in the MEC, it is required to deploy ECI network connections as soon as possible. For example, establish a VPN channel from MEC to a third-party cloud in real time as required.

It is recommended that the ECI network design be based primarily on L3VPN, with L2VPN as a supplement. Due to distributed and flexible MEC deployment, the 5G MEC requires the planning of L3VPN capabilities on the entire network, including the access network (L3VPN capabilities are extended to the edge). In addition, the VPN needs to span multiple network segments, such as the metro and backbone networks. To quickly support MEC service deployment, planning a unified logical ECI network (overlay network) is recommended. This also facilitates network slice deployment on the ECI network.

Figure 3-4 ECI multipoint communication network



Ethernet virtual private network (EVPN) integrates L2VPN and L3VPN services. Its control plane uses MP-BGP to advertise EVPN routing information, and its data plane supports multiple encapsulation modes, such as VXLAN, MPLS, and SRv6, in packet forwarding. The introduction of EVPN to the ECI network can meet the requirements of the edge computing system for scalability, reliability, and simplified O&M, reduce the management costs of MECs, and flexibly expand the services of the edge computing system.

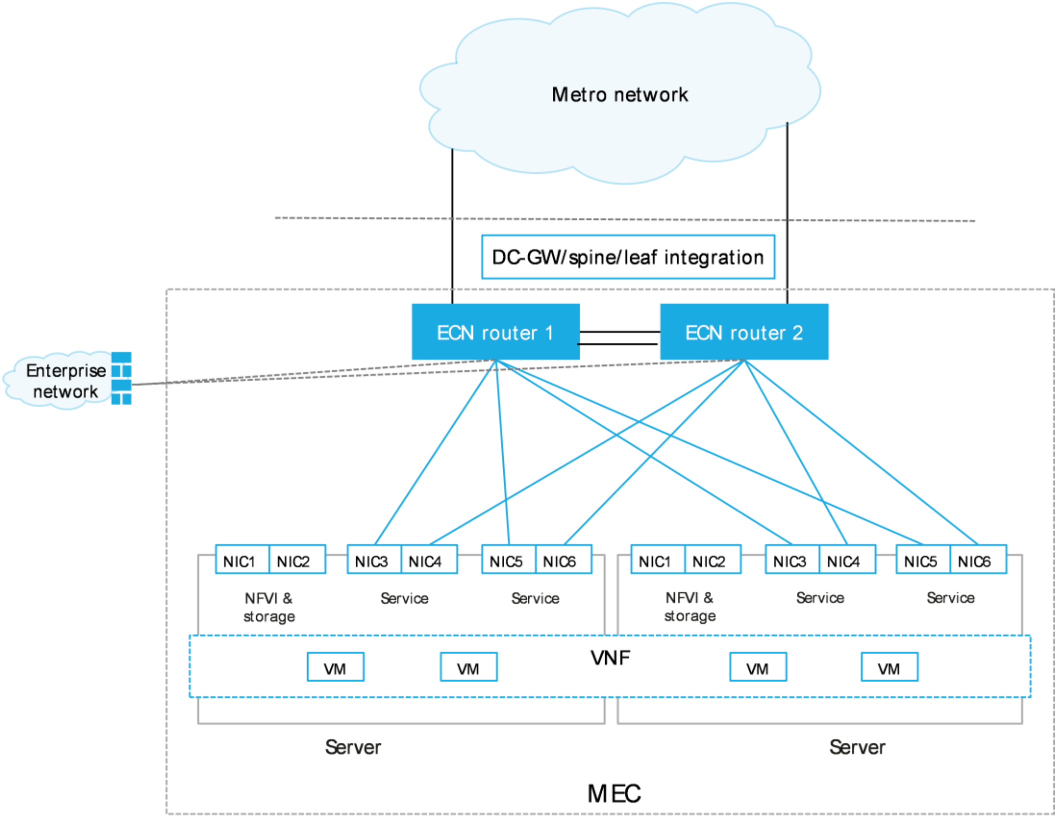
SRv6 is a network forwarding technology that combines the advantages of Segment Routing and IPv6. Oriented to the SDN architecture, SRv6 has powerful programmability and can work with the SDN controller to directly schedule network forwarding resources based on service requirements, meeting the SLA requirements of different services and reducing the number of routing protocols. On the ECI network, SRv6 enables fast service provisioning, fast VPN connection establishment across network segments, simplified protocol stacks, and simplified system integration.

Due to the complexity of ECI, ECI network design is a key part of 5G MEC planning for carriers. The ECI network design of each carrier varies according to its network architecture.

3.2.4 ECN: Integrated Network Architecture

The mainstream 5G MECs are small and mini MECs, which cost little to construct and have few servers. What's more, these MECs do not use the complex multi-layer network architecture (CE/DC-GW/spine/leaf) adopted by large DCs. Instead, ECNs generally use a one-layer integrated network model, as shown in Figure 3-5. ECNs, however, still need to provide most of the communication functions offered by multi-layer DC networks.

Figure 3-5 ECN reference model



ECN functions can be divided into two parts: LAN-side functions (for internal interconnection) and WAN-side functions (for external interconnection).

The LAN side of the ECN interconnects internal servers (forwarding, compute, and storage servers) with related devices and provides reliable L2 and L3 connections between VMs on internal servers. A UPF VNF instance can run on multiple VMs at the same time to improve performance and reliability. The ECN routers need to provide the equal-cost multi-path (ECMP) communication capability for UPF VMs, so that UPF service traffic can be evenly distributed among UPF VMs and UPF service traffic can be evenly redistributed when a UPF VM fails.

UPFs are routed VNFs deployed in the telco cloud. Routed VNFs need to exchange user route information with the ECN and transport network. Therefore, intra-ECN routes include UPF mobile user routes and VM routes. The number of ECN internal routes is determined by the number of end users and VMs. If there are a large number of routes, ECN devices must provide high routing capabilities.

The WAN side of the ECN implements edge-cloud synergy interworking and reliable communication between the MEC and external IP network (IP RAN). The WAN-side functions vary according to the ECI networking mode.

During ECN design, it is vital that a cost-effective integrated network solution is provided. Generally, one layer of routers are used as MEC gateways to provide CE/DC-GW/spine/leaf functions and meet the requirements of the LAN side and WAN side of the ECN. The ECI design varies according to carriers. Some carriers use ECN routers as service PEs. This reduces the impact of MEC services on the external IP network (IP RAN) and improves the flexibility of ECI connection setup (for example, SRv6 is deployed across multiple network segments between ECN routers and the central cloud), but raises higher function requirements for ECN gateways.

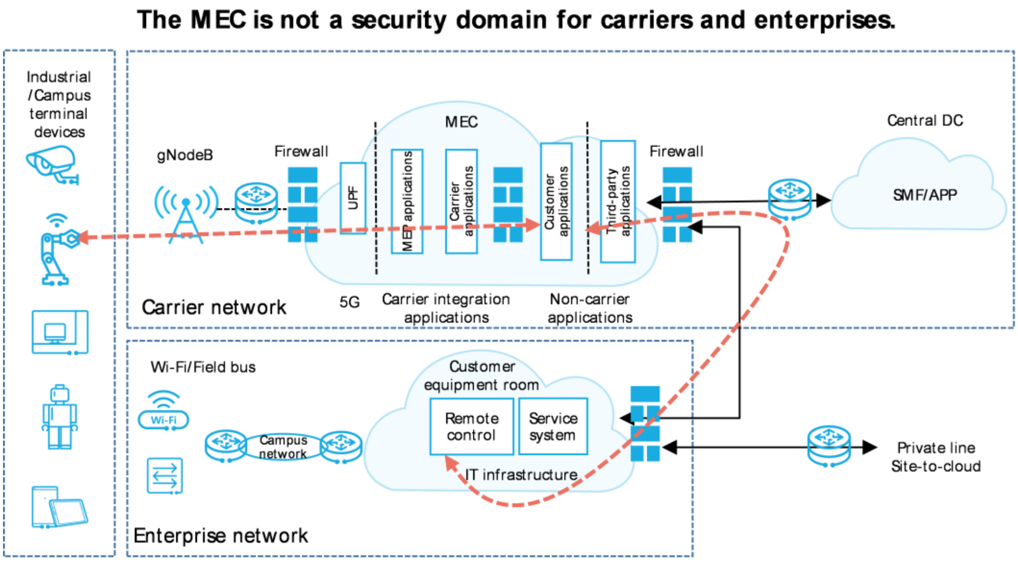
3.2.5 Carrier Network and Enterprise Network: Secure Interconnection Through the MEC

The 5G MEC is not a security domain for carriers and enterprises. It brings new network security risks, as shown in Figure 3-6.

From the perspective of carriers, the 5G MEC has non-carrier applications and network connections. For example, the MEC directly interconnects with enterprise networks and does not belong to the telecom security domain. In addition, the MEC changes the closed service transport nature of the IP RAN on the original mobile transport network.

From the perspective of enterprises, service data and applications pass through the external network and IT systems. In other words, service data and applications pass through non-security zones. In addition, additional points are added by MEC to the enterprise network for interconnection with the carrier network.

Figure 3-6 Secure interconnection between the carrier network and enterprise network through the MEC



Currently, MEC projects use firewall-based network security solutions. Carriers and enterprises both deploy firewalls on the interconnection channel to ensure network security.

A 5G MEC may have non-carrier applications. The ECN generally uses firewalls to divide the internal network into different security domains for isolation. For example, UPFs, carrier MEC applications, and third-party applications are distributed in different security domains with different security levels. The telecom systems and NEs of carriers are the key targets of security protection. Because there are large numbers of geographically dispersed 5G MECs, carriers need to plan an overall network security solution based on firewalls for network security deployment and security policy management.

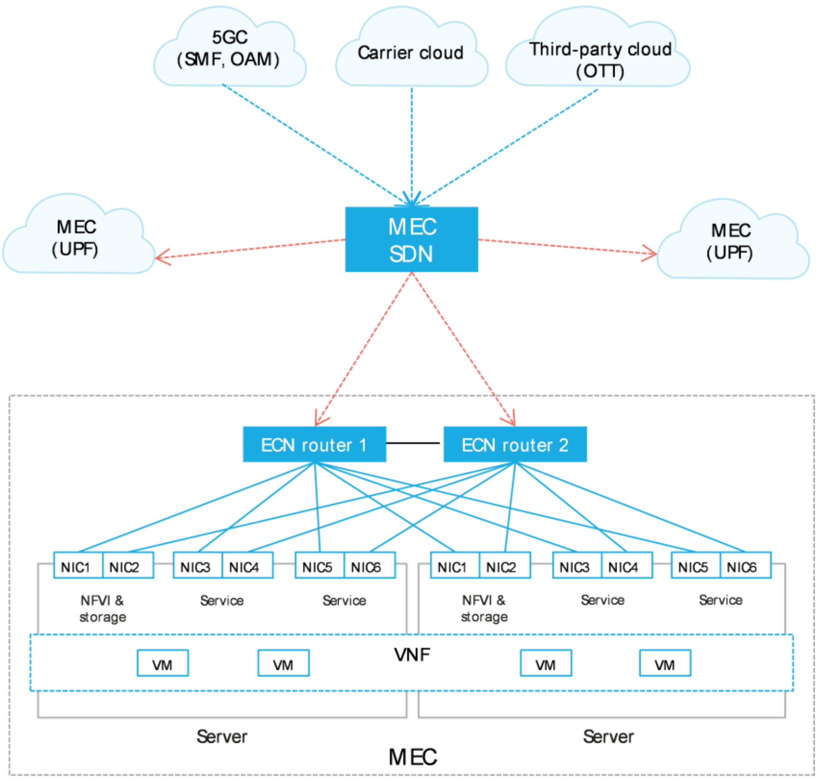
In addition to firewalls, enterprises can also use information security encryption solutions to protect key service flows. That is, end-to-end information encryption is used to encrypt traffic transmitted through carrier networks and MEC. However, the latency caused by encryption and firewall traversal is a problem to be considered for latency-sensitive applications.

Enterprises and carriers are both highly concerned about secure interworking between the two networks.

3.2.6 Network Support for Cloud-Edge Synergy

The 5G MEC needs to support automated deployment and online scaling of 5G UPFs. For example, the performance of a UPF VNF can be improved by adding more VMs. The MEC can dynamically take service requirements and automatically deliver network configurations to ensure fast online scaling of UPFs, thereby supporting cloud-edge synergy. If edge computing applications deployed on the MEP are related to the central cloud, edge-cloud synergy is also required. One of the basic communication requirements of cloud-edge synergy is to quickly establish VPN channels between the cloud and edge based on service requirements.

Figure 3-7 Reference network model for automated service deployment and cloud-edge synergy



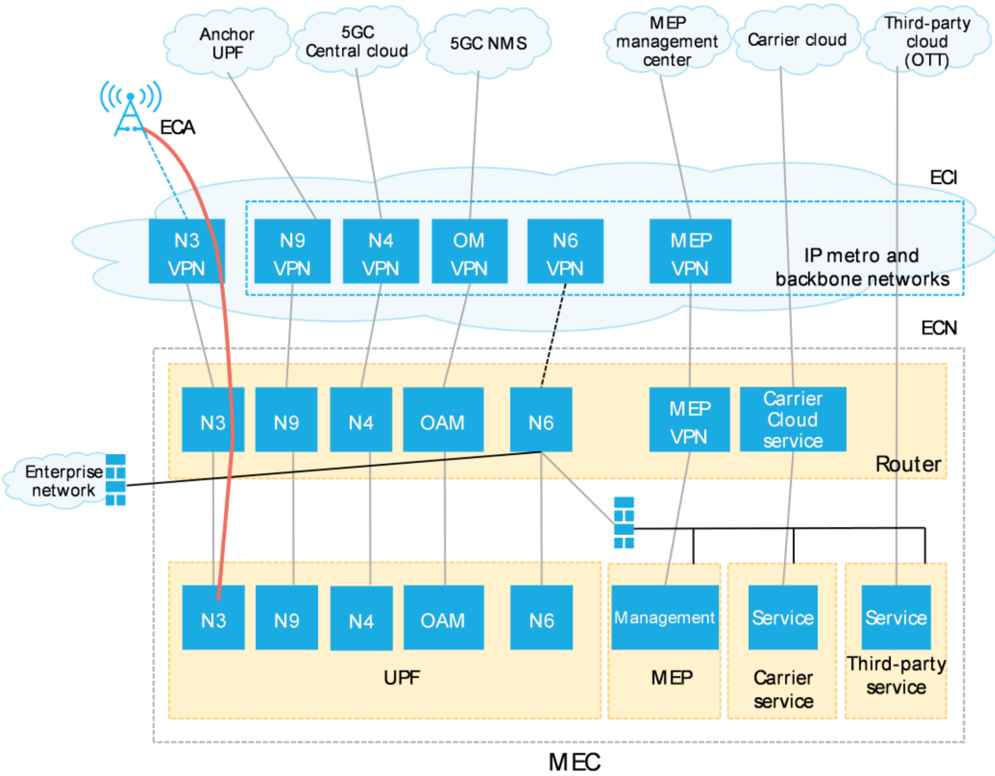
As aforementioned, the 5G MEC supports cloud-edge synergy. To support automated deployment of edge-cloud synergy for the telco cloud and central cloud, a unified SDN controller can be included in the design. In the reference model shown in Figure 3-7, an MEC SDN controller centrally controls MEC transport network devices (red line in Figure 3-7), and works with the management and orchestration (MANO) systems (blue line in Figure 3-7) of the 5G core network, carrier cloud, and third-party applications to dynamically accept service requirements and centrally configure related network devices for automated deployment and O&M of cloud-edge services, and dynamic establishment of VPN channels between clouds and MECs.

3.3 5G MEC Planning Advice and Network Architecture Reference Model

Based on in-depth analysis and research of the four major challenges and six key points of 5G MEC planning, this document provides the following 5G MEC planning advice for carrier network design.

1. Design the network solution segment by segment based on ECA, ECN, and ECI models. The ECA and ECI models can be deployed on different physical networks (for example, in scenarios where two carriers share the same 5G transport network).
2. Isolate the MEC from the external network as much as possible, so that service changes (for example, new UPF deployment) and network connection changes inside the MEC have little impact on the external network.
3. Align network solution planning with the responsibility matrix of the carrier's internal network O&M team to minimize work overlapping between different O&M teams. For example, if the ECN and IP RAN O&M teams are two independent teams, the O&M responsibilities of the two teams must be clear during network function design.
4. Consider incremental and on-demand deployment of the 5G MEC during network design. Try to minimize the impact of 5G MEC deployment on the network.
5. Construct the ECI network as a logical network with unified control and management. This ensures that network connections can be quickly established and the SLA assurance can be provided and 5G MEC services can be quickly deployed even when multiple networks are involved.

Figure 3-8 Carrier network architecture reference model from the MEC perspective



The network architectures of different carriers differ greatly. This document provides a network architecture reference model from the MEC perspective, as shown in Figure 3-8.

ECA: refers to the network between 5G base stations and the MEC, which mainly connects to the N3 VPN of the 5G mobile transport network. The ECA network of MEC B (carrier B) can use the network of another carrier, for example, the 5G mobile transport network of carrier A (N3 VPN of carrier A). In this case, carrier A provides SLA assurance for the ECA network.

ECN: The ECN interworks with the ECA network to direct 5G traffic to UPFs. The ECN interworks with the ECI network to guide service flows from service interfaces (N4, N6, N9, and OAM), MEPs, carrier clouds, and third-party clouds to the corresponding VPN of the ECI network. The roles of ECN routers in the ECI have great impact on the ECI networking. The network design varies according to carriers. To reduce the impact of device and network connection changes inside an MEC on external networks, the MEC may use an independent AS.

ECI: The ECI network is a complex logical network. It uses a multi-domain VPN to connect MECs to each other and to clouds. For service flows from 5GC network interfaces and the MEP on the ECI network, the corresponding connections are stable. For service flows from cloud applications on the ECI network, VPN connections may be needed for interconnection with the carrier clouds, OTT networks, and enterprise networks. These connections vary according to specific applications and need to be dynamically changed. It is recommended that carriers use their own networks to construct ECI. That is, the ECN and ECI should belong to the same carrier for rapid service deployment and SLA assurance.

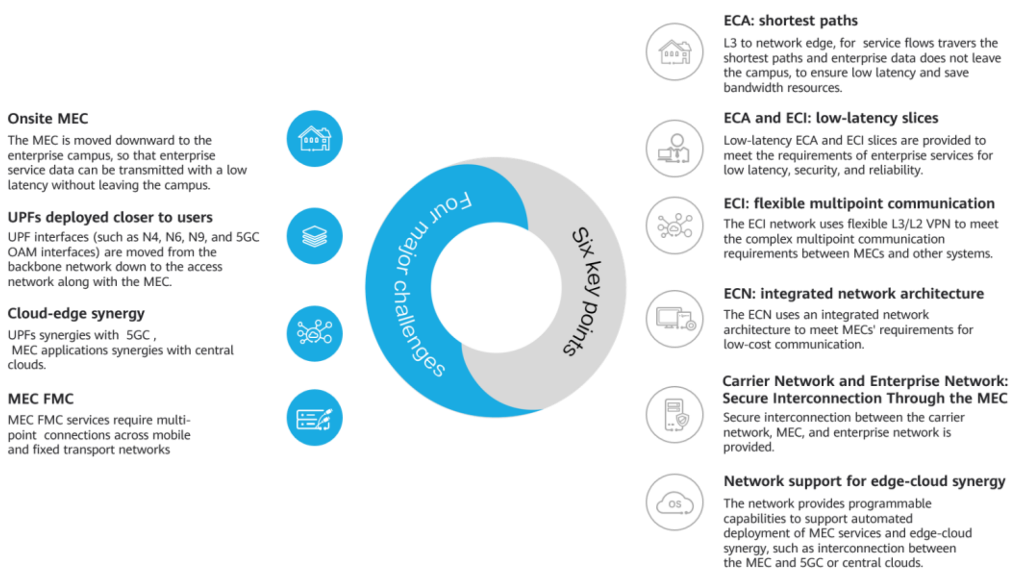
4 Conclusion

The 5G mobile communications system has made breakthroughs in the support of three major 5G scenarios, flexible deployment of user plane functions, and network capability exposure. "5G+MEC+AI" has become a new model that carriers are using to help vertical industries achieve digital and intelligent transformation, as well as an important indicator of the success of 5G applications.

In the 5G era, edge computing challenges facing carrier networks are mainly caused by cloud computing applications and the deployment of 5G gateways closer to users, which raise new requirements for service models and network functions. Figure 4-1 shows the four major network challenges. The design principles and network architecture of 4G mobile transport networks fail to meet the requirements of 5G MEC networks. In short, 5G MEC transport network construction is not a simple bandwidth upgrade of 4G networks.

How to build a 5G MEC-ready network is a question that must be answered by carriers in their efforts to tackle edge computing challenges. This document describes the six key points of 5G MEC planning (as shown in Figure 4-1) and provides network planning advice and the reference network model.

Figure 4-1 Four challenges and six key points of 5G MEC for carrier networks



5G MEC transport involves a wide range of networks, not only traditional mobile transport networks. Service flows may cross multiple network segments, including the metro and backbone networks, and the MECs need to be deployed on demand and incrementally. Carriers need to research and formulate the overall planning and deployment policy of the 5G MEC transport network as soon as possible.

5 Acronyms and Abbreviations

Acronym or Abbreviation	Full Name
4G	4th Generation
5G	5th Generation
5GC	5G Core
AI	Artificial Intelligence
AN	Access Network
AMF	Access and Mobility Management Function
API	Application Programming Interface
AS	Autonomous System
BGP	Border Gateway Protocol
BNG	Broadband Network Gateway
CE	Customer Equipment
CDN	Content Delivery Network
CUPS	Control and User Plane Separation
DC	Data Center
DC-GW	Data Center Gateway
DN	Data Network
EC	Edge Computing
ECA	Edge Computing Access
ECC	Edge Computing Consortium
ECI	Edge Computing Interconnect
ECMP	Equal-cost multi-path
ECN	Edge Computing Network

Acronym or Abbreviation	Full Name
ECNI	Edge Computing Network Infrastructure
EPC	Evolved Packet Core
EVPN	Ethernet VPN
FBB	Fixed Broadband
FW	Firewall
FMC	Fixed Mobile Convergence
IaaS	Infrastructure as a Service
IEF	Intelligent Edge Fabric
L2	Layer 2
L3	Layer 3
LAN	Local Area Network
MAN	Metro Area Network
MBB	Mobile Broadband
MEC	Multi-access Edge Computing
MEP	MEC Platform
MP-BGP	Multi-Protocol Extensions for BGP
MPLS	Multi-Protocol Label Switching
mMTC	massive Machine Type Communications
NFV	Network Functions Virtualization
OAM	Operation, Administration and Maintenance
OTT	Over The Top
PaaS	Platform as a Service
PE	Provider Equipment
PLC	Programmable Logic Controller
PDU	Packet Data Unit
SaaS	Software as a Service
SDN	Software Defined Network
SLA	Service-Level Agreement
SMF	Session Management Function
SRv6	Segment Routing over IPv6
UL CL	Uplink Classifier

Acronym or Abbreviation	Full Name
UPF	User Plane Function
URLLC	Ultra-Reliable and Low Latency Communications
VM	Virtual Machine
VNF	Virtual Network Function
VPN	Virtual Private Network
VXLAN	Virtual Extensible LAN
WAN	Wide Area Network
xPON	x Passive Optical Network

Contacts：

E-mail： song.jun@huawei.com