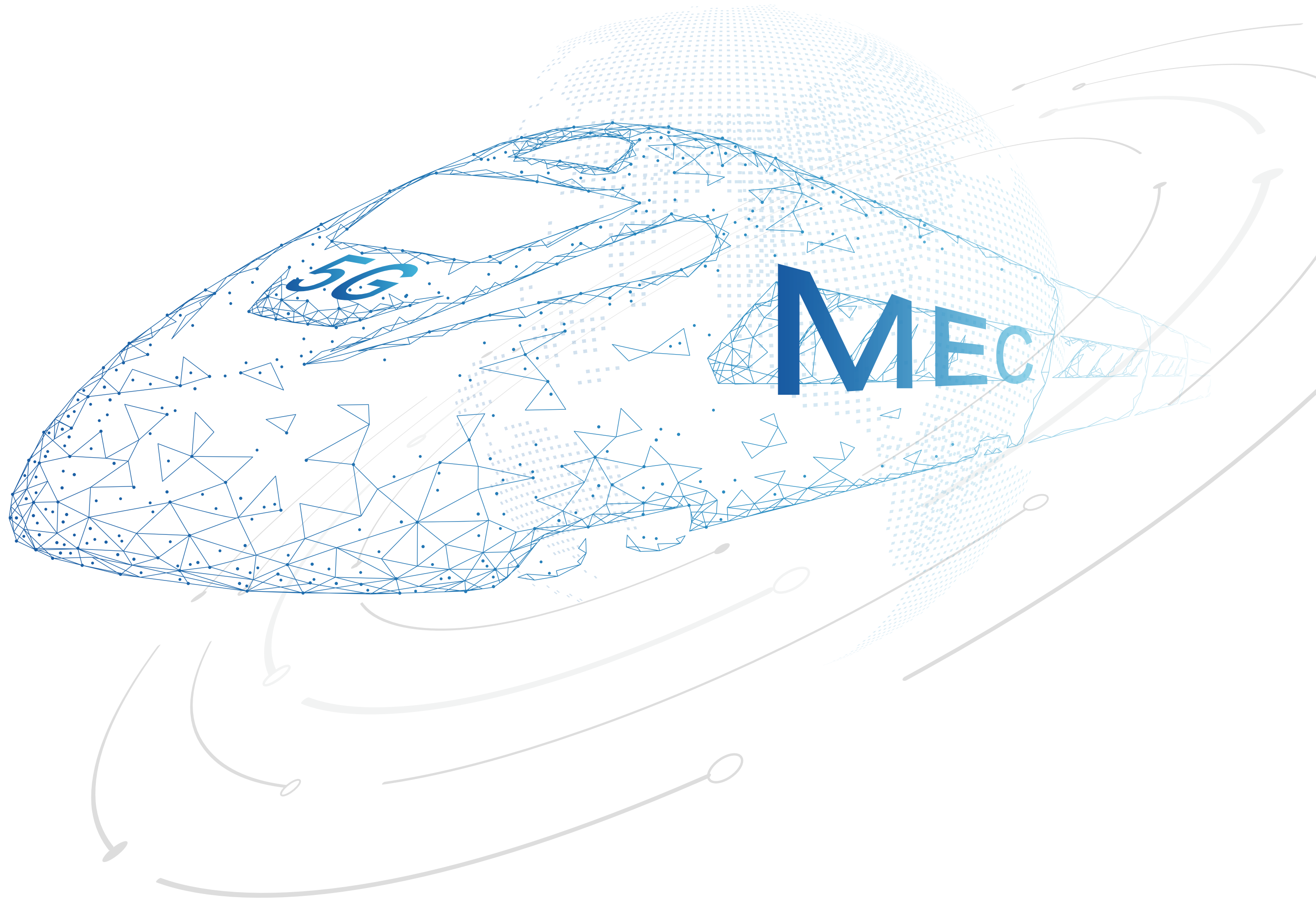
Edge Native Technical Architecture White Paper



February 2021

Contributors

China Academy of Information and Communications Technology (CAICT)

China Mobile Research Institute

China Telecom Research Institute

China Unicom Research Institute

DALU Robotech Technology (Beijing) Co., Ltd.

Haier Institute of Industrial Intelligence (Qingdao)

Huawei Technologies Co., Ltd.

Purple Mountain Laboratories

China Mobile Research Institute

Qingdao Power Plant of Shandong Electric Power Company, State Grid Corporation of China (SGCC)

ROOTCLOUD Internet Technology Co., Ltd.

XENIRO (Shanghai) Co., Ltd.

Xidian University

YunEx (Nanjing) Co., Ltd.

99cloud Co., Ltd.

CONTENTS

Yu Xiaohui

Edge computing is trending as we transition into the new digital era of 5G. A technology that pushes computing, storage, and transmission capabilities to the network edge, it is a source of innovation in computing, signals new business opportunities in the information field, drives the digital economy, and lays the foundation for the digital transformation of industries. Both academia and the private sector have vested interest in edge computing. Authoritative organizations, such as Gartner and IEEE, highlighted edge computing as one of the top 10 strategic technology trends for 2020; while the Chinese Academy of Sciences and the Chinese Academy of Engineering named it as one of the ten most promising research topics in the Research Fronts 2020 and Engineering Fronts 2020, respectively. At the same time, CB Insights estimates that the total value of the global edge computing market will reach US$34 billion by 2023.

Today, we already see the practical application of edge computing. Several national development plans, such as the Industrial Internet Development Action co-initiated by the Ministry of Industry and Information Technology (MIIT) and other ministries in China, call for the application of edge computing. Meanwhile, various industries are pairing edge computing with their services, unlocking diverse applications in different scenarios. For example, Sany Heavy Industry, Haier, and Commercial Aircraft Corporation of China, Ltd. (COMAC) have applied edge computing to upgrade conventional manufacturing. At the same time, innovation in edge computing is driving more and new applications, and brings a new concept called Edge Native. By working with other capabilities, such as computing, Edge Native creates an intelligent and trustworthy network edge, delivering the agile connectivity, rock-solid reliability, optimized data, and intelligent applications that industries require. A critical part of the edge technology system, Edge Native will continue to evolve in collaboration with other key technologies like Cloud Native and Edge Network. Together, they will create a sustainable edge ecosystem and stimulate business innovation.

Within this context, the Alliance of Industrial Internet (AII), Edge Computing Consortium (ECC), 5G Deterministic Networking Alliance (5GDNA), and EdgeGallery Open Source Community have formulated this white paper. The paper introduces the concept, technical architecture, and industry practices of Edge Native, aiming to promote the development of Edge Native technologies and help industries unleash the disruptive potential of edge computing.

Despite remarkable progress, Edge Native has yet to achieve maturity, and its potential remains largely untapped. As such, all industry stakeholders need to work together, continuing to promote the research and practical implementation of Edge Native in different industries. This will help capitalize on its value and ultimately digitally transform industries.

Yu Xiaohui
CAICT Director, AII Director-General, and ECC Deputy Director-General

Introduction

ETSI first unveiled the concept of Multi-access Edge Computing (MEC) in 2015. Since then, the telecom field has widely adopted edge computing. However, the number of 5G applications is growing rapidly, and the current capability of edge computing cannot satisfy their unique requirements; and 5G MEC deployment is also facing a variety of challenges. That said, we are now moving towards Edge Computing 2.0. And to navigate the above-mentioned difficulties, we need to adapt MEC platforms to 5G networks, ensuring that computing and networking are collaboratively implemented. Edge Native is the key in this process.

Edge Native strengthens edge capabilities, including service orchestration, AI, and enhanced security on the network edge; it effectively integrates computing and connectivity. Therefore, Edge Native will be the foundation for industry-oriented networks, those planned by the operators in the medium-term.

The four chapters of this white paper cover various aspects of Edge Native:

- Chapter 1: Introduces the development of edge computing and the relevant 5G industry requirements.
- Chapter 2: Explains the concept and main technical capabilities of Edge Native.
- Chapter 3: Looks at the ways industry stakeholders, including the EdgeGallery Open Source Community,
- China Mobile, and China Unicom, have promoted Edge Native.
- Chapter 4: Illustrates the practical use cases of the Edge Native in different industries, such as smart
- factory and smart robotics, and looks at the prospects for this technology.

01

The Development of Edge Computing and 5G Industry Requirements

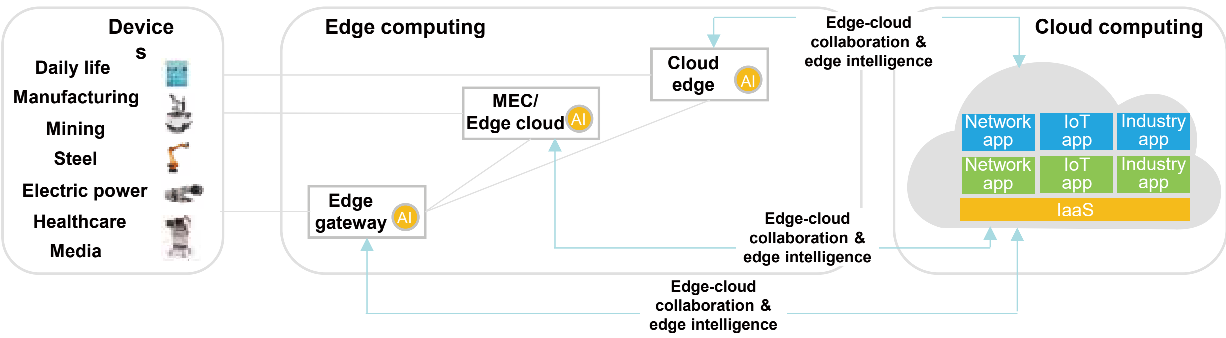
1.1 Edge Computing 2.0 and 5G MEC

The ECC defined Edge Computing 1.0 and its reference architecture in the White Paper of
Edge Computing Consortium released in 2017 as follows:

Industry stakeholders are turning to edge computing as a way to grow business. Instead of technical implementation aspects, they want to know how to apply edge computing to commercial services, the development trends of relevant technologies, and which key capabilities do the hardware and software platforms need to provide. Edge Computing 2.0 can address these key questions.

Edge Computing 2.0 emphasizes edge-cloud collaboration and edge intelligence. It introduces the concept, architecture (see Figure 1-1), and technologies of cloud computing to the edge computing software platform, facilitating real-time, AI-powered data processing as well as reliable data configurations that are easy to reset. Furthermore, it supports the different models of CPUs (x86 and Arm), GPUs, and NPUs to provide heterogeneous computing power.

"Edge computing is performed on an open platform at the network edge near things or data sources, integrating network, computing, storage, and application core capabilities and providing edge intelligent services. Edge computing meets the requirements of industrial digitalization for agile connections, real-time services, data optimization, smart application, security and privacy protection".



There are three ways to implement Edge computing 2.0:

Cloud edge	The cloud edge is a logical part and natural extension of the central cloud. It relies on the central cloud to provide capabilities and services. Representative solutions include IEF (Huawei), Link IoT Edge (Alibaba Cloud), and Outposts (AWS).
MEC	This approach is also called edge cloud. In this case, small- or medium-scale cloud services or equivalent network capabilities are deployed at the network edge, and services are processed locally, eliminating the need to go through the central cloud. A content distribution network (CDN) is a typical MEC application.
Edge gateway	This approach reconstructs conventional embedded gateways by utilizing cloud technologies and capabilities. Edge gateways provide comprehensive APIs to adapt to various protocols used for data transmission to the network edge, and also offer computing capabilities to process services. In addition, cloud-based controllers assist edge gateways by flexibly scheduling resources, managing applications, and orchestrating services for edge nodes.

It is important to highlight that MEC tends to be the preferred approach. This is because it ideally matches IT computing power with CT networks; plus, ETSI and 3GPP have clearly defined its value and functionality. As such, 5G network operators tend to prefer MEC over other approaches.

A standard 5G MEC architecture consists of the user plane function (UPF), Multi-access Edge Platform (MEP), industry applications, and virtualized infrastructure. It is usually deployed with an all-in-one hardware set and maintains close proximity with end users. It provides large bandwidth, ultra-low latency, strong computing capabilities (AI and image rendering), and rock-solid security. As such, MEC significantly reduces the costs and management complexity of 5G devices, benefiting both individuals and industries.

1.2 Value of 5G MEC

Statistics for 2020 indicate that 60% of the over 5000 5G-based industry pilot projects around the world used MEC. For example, China Mobile, China Telecom, and China Unicom have piloted over 100 5G MEC projects in more than 40 cities, covering a variety of services in different industries, including manufacturing, mining, and energy. During the Zhan Fang Cup 5G application contest held by the MIIT in 2020 (see Figure 1-2), 43% of submitted projects used MEC — 10% more than in 2019; and over 80% of the projects that made it to the finals used MEC. All this to say that MEC is evidently key for promoting 5G applications.

It is obvious that 5G MEC will serve as a powerful driving force behind the digital transformation of various industries.

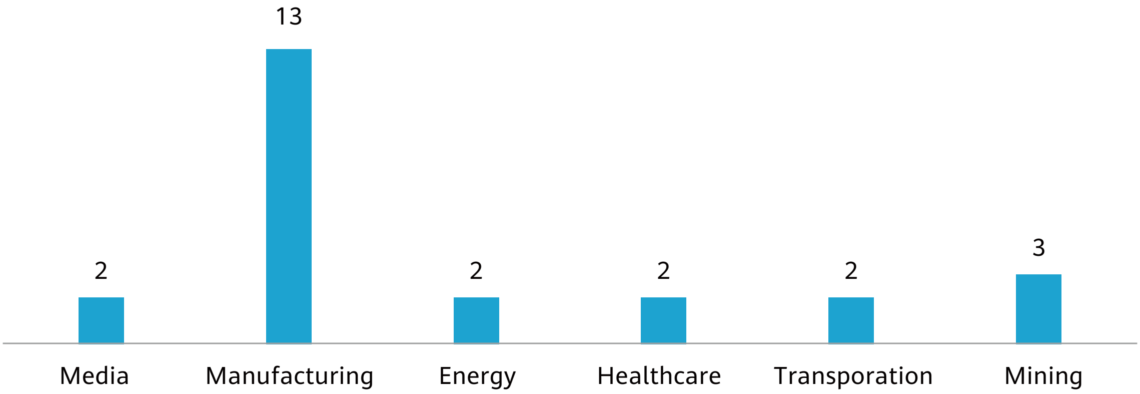


Figure 1-2 MEC-based projects in the finals of the Zhan Fang Cup contest

Smart Manufacturing

5G-based smart manufacturing is essential in modern factories, which often use a wide range of devices that generate various types of data. They require highly secure and efficient data transmission — requirements that cannot be satisfied without MEC.

In 2020, most 5G-based industrial manufacturing pilot projects applied MEC. These projects focus on auxiliary production systems, especially the application of automated guided vehicles (AGVs). For example, China Telecom helped Sany Heavy Industry deploy MEC in its factory, replacing conventional tracked vehicles with mobile workbenches and AGVs. They also integrated the 5G MEP with software to collect and analyze the massive amount of video data captured for AGVs, using it to adjust AGV trajectories and schedule related resources. This has improved the work efficiency of AGVs while reducing their costs, and has also ensured almost real-time data transmission (with a latency of approximately 20 ms), high security, and strong computing capability.

Smart Mining

Safe production is essential to the mining industry. A 5G network (see Figure 1-3) can stretch above and below ground, covering every corner of a mine. It also delivers large bandwidth, massive connections, and low latency to facilitate the analysis of video content and remote control of various devices such as sensors. With MEC deployed in the mine, local servers can directly process device data, avoiding the need to first send data to the public network. In this way, comprehensive, accurate production data and HD video footage can be efficiently obtained across the mine; and a series of novel applications, such as HD audio and video calls, instant transmission of various data, and intelligent remote control of devices, can be implemented, facilitating communications between overground and underground areas. This lays a solid foundation for the mining industry to reduce their workforce, improve efficiency, guarantee safety, and escalate digitalization.

Moreover, an emergency 5G Core can be deployed with the MEC system to meet the stringent requirements on security. If the path between the network edge and public network goes down, the emergency 5G Core will serve as a backup of the public 5G Core. It will continue to provide services, stabilizing wireless communications and data transmission as well as ensuring secure isolation above and below ground.

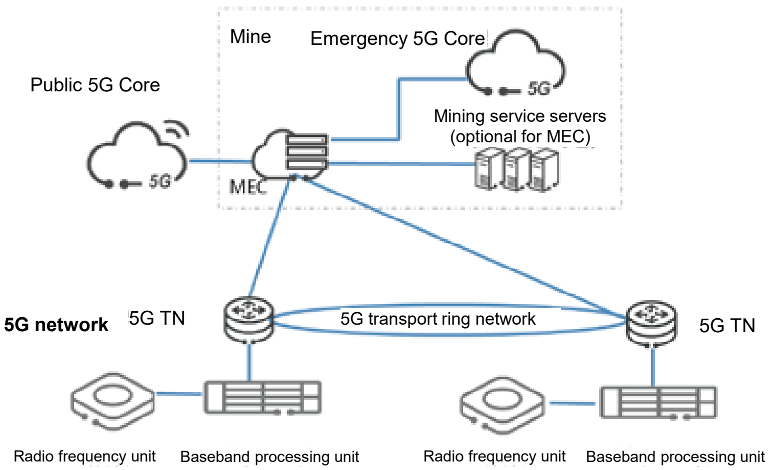
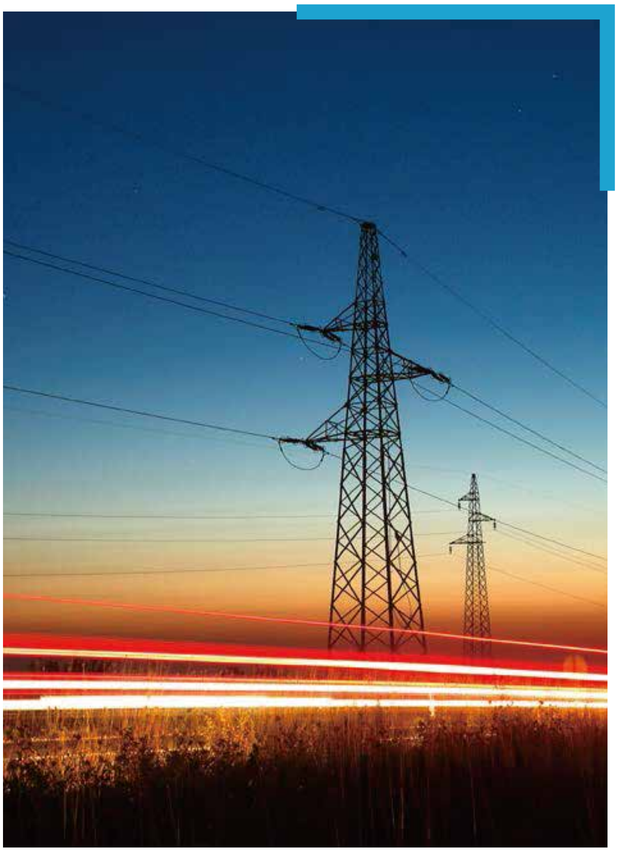


Figure 1-3 Smart mine networking

Smart Energy



The electric power industry is transitioning to smart energy powered by 5G. Previously, optical fibers covered power transmission and transformation to offer low latency, large bandwidth, and high security. However, they were not comprehensively used for power distribution and consumption due to terrain and heavy costs. The power distribution and consumption phases require wide network coverage and involve a large number of devices, and therefore they used to rely on 2G, 3G, and 4G public telecom networks. Yet these public networks cannot provide the ultra-high security that electric power services require, particularly those with the control functionality. Neither can they cope with the massive connections generated by the increasing number of devices.

5G and MEC can address these paint points. MEC and UPFs can ensure data transmission only takes place within a power grid, thereby securing power service data. In addition, substations utilize a panoramic visualization system that takes the massive raw data they generate locally turning them into various measured values and audio/visual materials. MEC also enables strong connectivity and computing capabilities, allowing this data to first be securely calculated within the grid campus; and then be efficiently transferred to the primary station without occupying the bandwidth of the metro network used for electric power services.

Smart Transportation

The transportation industry used to manually monitor and maintain roads using a backhaul of images and videos. Now, the industry is looking for ways to digitalize and automate operations and management. MEC is an ideal solution for this.

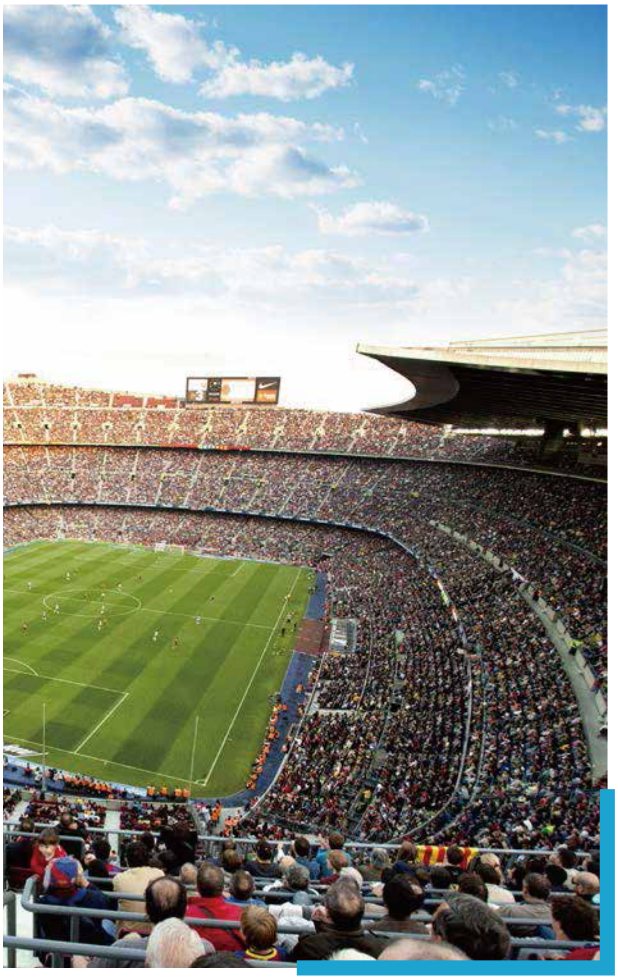
The MEP accelerates the backhaul of HD images and videos in multiple channels to almost real-time. Moreover, it draws on AI-driven technologies, such as video content analysis and machine vision, to dynamically analyze the collected data and quickly make decisions to handle faults or initiate alarms. This ensures that the system will not miss any key status, performance indicator, or loss, quickly diagnosing any issues.

Smart Healthcare

The healthcare industry needs to respond faster and be more secure. Time is literally life. This means that healthcare personnel need to have access to the hospital network anytime, anywhere. This idea has accelerated the development of telemedicine at an astonishing rate. Meanwhile, massive healthcare data (60% of which is made up of images) must stay secure during transmission; for example, it is key to prevent data tampering or leakage out of hospitals.

5G healthcare private networks that combine MEC and network slicing technologies provide the ultra-low latency, high reliability, and all the other specific SLA items required by the healthcare industry. The dedicated MEP used on such networks prevents data from leaving the hospital, ensuring secure isolation. The platform also runs remote consultation services, making sure that diagnosis and treatment data can be transferred with minimal latency.

Smart Media



Broadcasting stadium events live is a major trend in smart media, whether it be for sports, entertainment, or conventions. To provide viewers with a more exciting immersive experience, cutting-edge technologies (such as UHD, VR, and surround sound) need to be used for content development. And this content must be pushed in real time to an extensive range of terminals. This is impossible using previous technologies, such as outside broadcast (OB) vans. Instead, content production is becoming more centralized, remote, and lightweight, where MEC plays a key role.

For small-scale events that do not have strict content production requirements, MEC assists in the agile deployment of the live broadcast system and onsite content production. Large production devices, such as OB vans, are expensive for such events. MEC helps efficiently transfer signals from multiple cameras over 5G networks to the remote production center, which then distributes the content. This improves production quality and reduces costs. Therefore, broadcast personnel can better accommodate different types of events and deliver enhanced services.

MEC also helps move content production capabilities to the network edge, so that content can be processed locally and transferred faster than ever. For example, real-time analysis of sports event data, video playback editing, and transcoding between different video formats can all take place at the edge. This means that all types of rich media content can be presented to viewers at lower cost. Furthermore, it is also possible to move the centralized production process to the edge cloud, making more resources available for lightweight production.

1.3 Challenges for 5G MEC

Though MEC has been highly valued by diverse industries, there is still an array of challenges ahead. These are mainly related to network deployment, device capabilities, and software development.

» MEC performs well in each individual service. However, different services require that MEC be deployed in different locations or in varied formats. Therefore, MEC needs to be improved for complex service scenarios, adapting to heterogeneous devices and various network structures.

» The MEP and the applications on it need to provide stronger data analysis capabilities. AI needs to be infused into both related hardware and software.

» MEC needs to be systematically coordinated with industry-specific applications and operator networks, making these applications more stable and efficient.

» MEC also needs to be strengthened to provide more reliable and secure isolation of services and data.

As for software development of edge applications, developers need to consider how to use different types of MEPs to unleash the full potential of open 5G MEC capabilities in various industries.

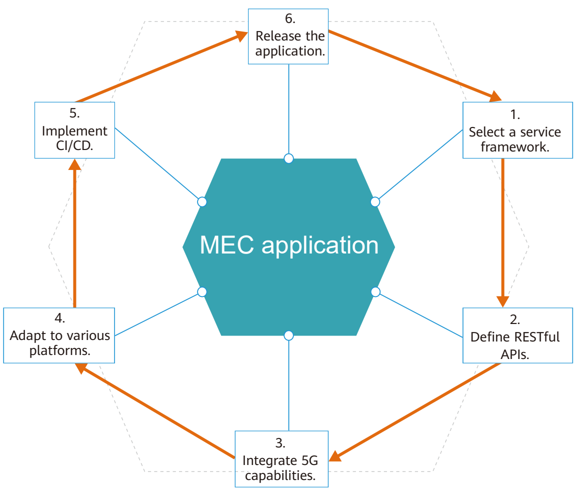


Figure 1-4 MEC application development process

1.Select a service framework.

To begin with, developers need to select an application framework, third-party software, and database based on their skills in development languages and the target environment. They need to ensure that the same MEP can manage the microservices under different frameworks; for example, these microservices can register, discover, and communicate with each other.

2.Define RESTful APIs.

Developers need to define the APIs that the application will open, and determine the capabilities exposed over these APIs.

3.Integrate 5G capabilities.

5G network capabilities, such as high bandwidth and low latency, need to be abstracted into programmable sets, such as software development kits (SDKs) or RESTful APIs, or even into application templates. In this way, developers do not need to focus on the sophisticated details in technical implementation, but directly use these sets and templates during coding as well as using the orchestrator to flexibly manage the coded applications.

4.Adapt to various platforms.

Developers need to adapt the application to different IaaS and PaaS platforms, including x86/Arm hardware platforms and virtualization platforms represented by OpenStack and Kubernetes. In addition, the application will be powered by the MEP, which provides well-planned edge capabilities like basic databases, strengthened security, AI, and blockchain.

5.Implement Continuous Integration (CI) and Continuous Delivery (CD).

Developers need assistance with a tool chain for security testing for DevOps (DevSecOps). Additionally, the capabilities for the application need to be orchestrated from end to end, covering the other applications and third-party capabilities invoked during application deployment.

6.Release the application.

After being proven viable through extensive testing, the application can be placed into the application repository or store.

02 Edge Native Concept and Capabilities

2.1 Concept and Industry Value

To address the challenges of 5G MEC and promote its application in different industries, the Edge Native concept was proposed by 5GDNA, EdgeGallery Open Source Community, ECC, and AII.

Cloud Native was first put forward to develop 12-factor applications. In 2018, it was defined by the Cloud Native Computing Foundation (CNCF) as a paradigm that eliminates the underlying differences in technical implementation between various applications and allows these applications to leverage cloud features, such as elasticity and distributed layout. Under the Cloud Native architecture, applications can be quickly deployed, flexibly scaled, and guaranteed with zero downtime. This approach laid the foundations for Edge Native, which is expected to have greater reach thanks to its ability to adjust to varying industry needs.

As the ICT industry embraces edge services, Cloud Native encounters several challenges at the edge:

» Edge device data must be transferred to the central network first, and the resulting long round-trip time (RTT) degrades user experience.

» The consumption and management of computing, storage, and network resources must be terminated on the central network.

» Costly data transmission and computing.

» Limited privacy protection and data security.

Table 2-1 Cloud Native V.S. Edge Native

Cloud Native		Edge Native
» Computing only		» Computing and connectivity
» Medium- and large-size data centers » General-purpose hardware » Flexible scaling		» Lightweight edge node » Heterogeneous hardware » Limited or unavailable scaling capabilities
» Independent	VS	» Mutually dependent
» Centralized, automatic management » Horizontal orchestration		» Remote management and edge autonomy » Cross-node orchestration
» Internet-oriented » Backbone network as the focus		» Access-oriented » 5G mobile access as the focus

Edge Native encompasses a new series of concepts for IT developers that are related to telecom network connectivity. These concepts, regardless of whether they are conveyed by 3GPP-defined attributes or reflected by the edge network architecture, parameters, and features, must be clearly paraphrased and presented to developers.

Edge Native is the ideal target network architecture for industries to carry out their digitalization strategies. Leveraging cutting-edge technologies will enable all industry stakeholders to efficiently build, run, maintain, and manage latency-sensitive applications at the network edge. These technologies include:

- 5G network capability openness at the network edge
- Cloud-edge collaboration and edge-edge collaboration
- Service orchestration across different network edges
- Multi-language and serverless design for edge autonomy
- Near-real-time preprocessing and analysis of massive raw data
- Authentication, encryption, and blockchain for a trustworthy edge
- Edge AI for heterogeneous hardware acceleration

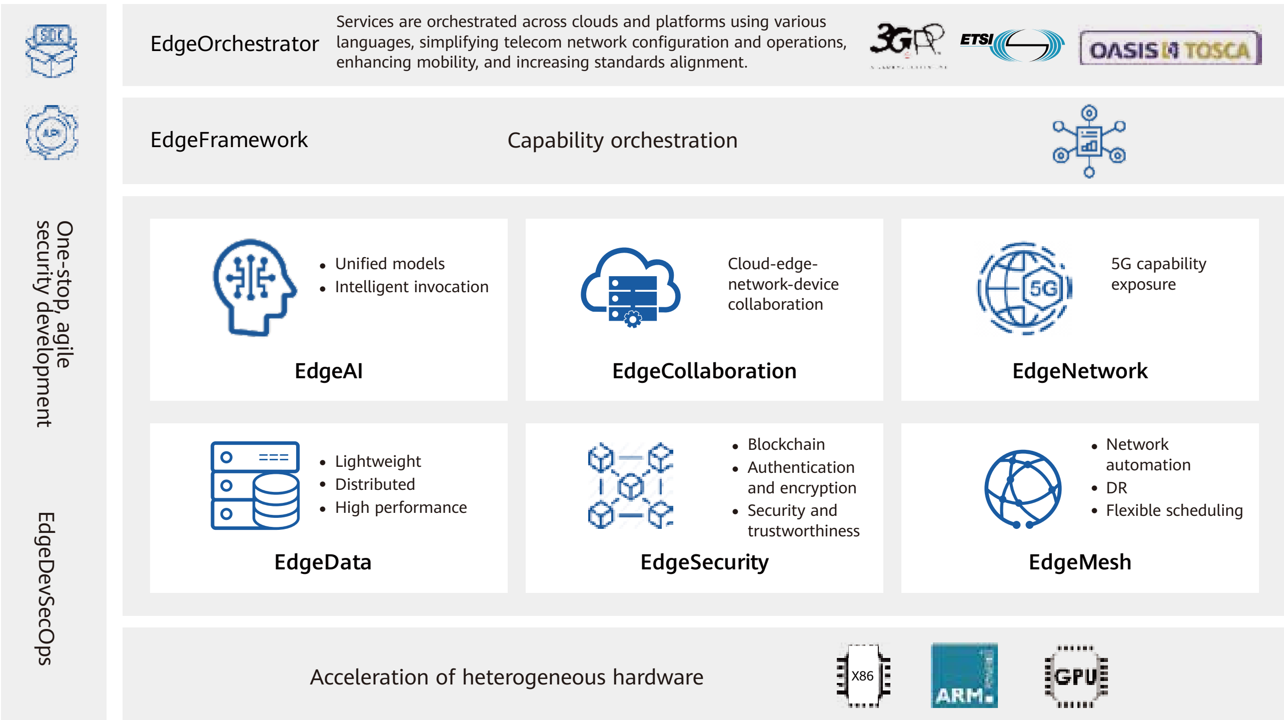


Figure 2-1 Edge Native technical architecture

2.2 Capabilities

2.2.1 EdgeInfra

EdgeInfra provides a solid infrastructure for the network edge.

In terms of hardware, EdgeInfra not only enhances computing power, but also allows for all computing capabilities, such as CPUs (x86 or Arm), GPUs, NPUs, or whatever follows, that facilitate the forwarding of heavy 5G traffic.

As applications are increasingly virtualized and containerized, Kubernetes has been popularized as an open-source software platform for application management. However, this platform alone cannot meet the stringent secure isolation requirements imposed by MEC applications. With the help of EdgeInfra, it is the ideal choice for the needs of telecom networks, and provides:

» **Upgraded support for telecom networks**
In MEC-based telecom service scenarios, different network planes and tenants need to be separated, and telecom technologies, such as time sensitive networking (TSN), need to be applied to reduce latency. EdgeInfra enhances the existing container network interfaces (CNIs) for Kubernetes, such as introducing a series of CNI plug-ins like Multus. In doing so, EdgeInfra enables the attributes related to telecom service isolation and technology implementation to be abstracted and configured through standard Kubernetes APIs or templates. The final attributes then can be flexibly invoked by applications.

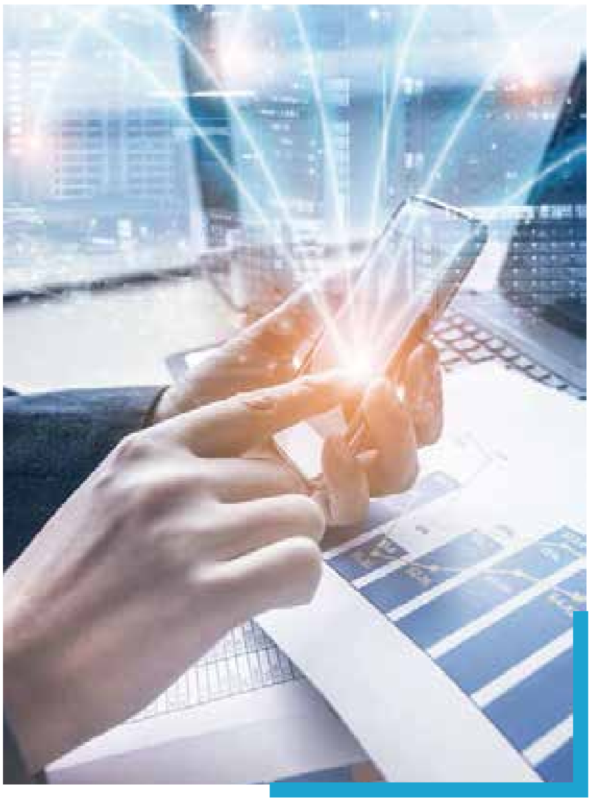
» **Hybrid orchestration of containers and VMs**
Containerizing all telecom network functions and MEC applications will be a slow process, and for this reason, containers and VMs will continue to coexist. EdgeInfra packages VMs into container images, facilitating hybrid orchestration of VMs and containers. The orchestration can be made more efficient with assistance from edge management solutions, such as StarlingX provided by OpenStack.

2.2.2 EdgeNetwork

The MEP functions akin to a bridge connecting 5G networks to applications and end users. In addition to providing strong connectivity, the MEP needs to systematically collaborate with applications and open 5G network capabilities over well-defined APIs. EdgeNetwork provides the necessary capabilities to achieve this.

» **5G network capability exposure**
The ETSI MEC standards stipulate that network capabilities and related information, such as device locations, need to be opened. At the same time, the 3GPP and telecom operators also defined a network exposure function (NEF). EdgeNetwork enables the MEP to quickly launch 5G network capabilities under the architecture specified by the ETSI or 3GPP, and presented them as attributes that can be invoked. Furthermore, EdgeNetwork provides SDKs and sample code to developers for exposing 5G network capabilities, and fully unleash 5G functionality.

» **MEP with converged computing power and enhanced connectivity**
The MEP needs to standardize its underlying 5G network capabilities as well as provide computing and network resources to the upper-layer applications. To efficiently schedule and orchestrate the connectivity, computing, and other deterministic 5G capabilities, the requirements of different industry applications must be clearly defined, ideally as templates. Each template needs to include two parts, with one describing computing resources, such as Helm for Kubernetes and Heat for OpenStack, and the other utilizing EdgeNetwork to abstract and detail telecom network resources.



2.2.3 EdgeOrchestrator

EdgeOrchestrator enables the network edge to provide adequate orchestration capabilities, so that lightweight edge nodes can be flexibly deployed in complex environments to fulfill the different SLA requirements of industries. More specifically, EdgeOrchestrator allows for:

» **Edge autonomy**

If the edge network fails to communicate with the central network, data synchronization between the two sides may fail, and will likely lead to application failures. EdgeOrchestrator enables certain autonomy capabilities on edge nodes, so that they can keep running, even during network faults. Once communication recovers, data synchronization will be re-initiated.

» **Mobility**

With EdgeOrchestrator, service information can be shared among operators. This ensures that subscribers can enjoy a consistent experience when switching between different operator networks.

» **Declarative programming for resource abstraction**

IT languages are used to predefine the resources required for deploying and configuring each application. For example, the platform capabilities, third-party components, and network information can be described as deployment dependencies in an application descriptor (APPD). This declarative approach tackles the low efficiency issue that comes with inconsistent languages adopted in IT and CT domains, and relieves the deployment and maintenance personnel from the labor-intensive configurations based on imperative commands.

» **Adaptive plug-in**

EdgeOrchestrator can be plugged into different infrastructure types, for example, it can pair with Kubernetes and OpenStack in cluster solutions, with lightweight Kubernetes (K3s or MicroK8s) in single-server solutions, and with runC on bare metal.

2.2.4 EdgeCollaboration

EdgeCollaboration allows MEC nodes to collaborate with everything, from other edge nodes to central public networks and end user devices, to effectively deliver computing capabilities and stretch network connections.

» **Edge-edge collaboration**

EdgeOrchestrator centrally orchestrates the MEC nodes on the network. If the current MEC node cannot provide sufficient computing for a task, EdgeOrchestrator will split this task, distribute subtasks to different MEC nodes, and combine them at one MEC node. In this process, resources are flexibly scheduled for optimal utilization. If an end user roams across MEC nodes, EdgeOrchestrator can quickly replicate the images of applications used by this user to the node they have roamed to and then deploy them, ensuring seamless and uninterrupted service.

» **Edge-cloud collaboration**

EdgeOrchestrator contributes to systematic coordination between the network edge and the public, private, and hybrid clouds. MEC applications can be driven by the massive computing power on the clouds. In addition, the public cloud capabilities can be directly opened to the MEC applications over APIs, so that the edge network can offer a similar service experience to the public cloud.

» **Edge-device collaboration**

With the rise of VR, AR, and video services, devices impose strict requirements on latency and image rendering. EdgeOrchestrator enables MEC nodes to efficiently communicate with devices in compliance with appropriate datacom protocols and provide strong image processing capabilities.

2.2.5 EdgeAI

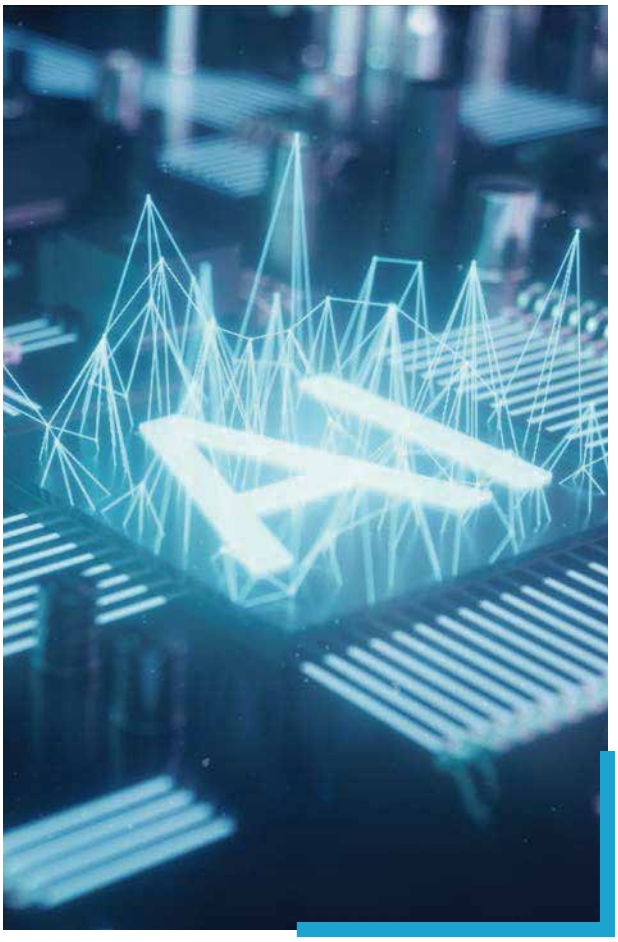
The fusion of MEC and AI has been deeply analyzed and studied in both academic and industrial circles. AI models can be leveraged in diverse MEC service scenarios, both those trained at the edge or in the center. EdgeAI converges the research findings in different scenarios and provides capabilities for creating an intelligent edge.

» **Unified AI framework and models**

EdgeAI provides engines for segmentation of AI models and data as well as machine learning with different methods, such as distributed learning, incremental learning, and transfer learning. It also defines corresponding algorithms and protocol interfaces, facilitating edge-cloud-device synergy. The AI models conveyed by EdgeAI are mostly open-sourced, for example, Open Neural Network Exchange (ONNX), which abstracts resources in a generic approach for developers, and Kubeflow, which fits into various AI frameworks and allows for centralized orchestration.

» **Unified scheduling of heterogeneous AI computing**

EdgeAI offers a hardware abstraction layer (HAL), through which unified hardware resources are scheduled to run different services, allowing applications to operate as efficiently as possible. Moreover, in edge-edge and edge-device collaboration, EdgeAI assists in offloading computing power locally based on the KPIs for services and resources.



2.2.6 EdgeSecurity

EdgeSecurity hardens the edge security for different industries by providing:

» **End-to-end security assurance**

EdgeSecurity offers well-rounded mechanisms, covering every aspect, from data authentication to protocol compliance, anti-attack, access control, service isolation, privacy protection, and virtualization security, to ensure rock-solid security and trustworthiness for infrastructure, networks, data, and applications. For instance, approaches like service isolation, status monitoring, and issue prewarning are used to strengthen network security. Moreover, to protect the data transferred between the MEP and applications from unauthorized access, measures must be taken to protect data confidentiality and integrity against attacks in different formats, such as replaying. To facilitate these measures, EdgeSecurity authenticates and authorizes the interoperations over the interfaces between the MEP and applications, secures the transmission path, and records logs for all access attempts from end users for post-event auditing.

» **Transparent verification platform**

This platform utilizes blockchain technology and cryptography algorithms to provide a series of privacy protection schemes, including cryptographic commitments, zero-knowledge proofs, verifiable encryption, and consistency verification between actual communication content and encrypted content.

» **Tamper-proof consortium blockchains**

EdgeSecurity encrypts the metadata, including the hash, abstracted from the digest of each application image. It also uses blockchain technology to store the encrypted data to a block, and manages each MEC node as private blockchain nodes. Together, these nodes are chained to a central public platform for efficient coordination and end-to-end data security assurance.

2.2.7 EdgeSecurity

EdgeMesh inherits Service Mesh technology to schedule network-wide MEC resources and provide refined management of multi-vendor applications on each MEC node.

» **Flexible computing power scheduling**

EdgeMesh uses the sidecar pattern to deploy components of an application into a separate process or container to facilitate isolation and encapsulation. In this pattern, EdgeMesh can obtain the computing status and resource usage of the current MEC node, and report them to the upper-layer controller or orchestrator as references for adjusting resource scheduling policies, ultimately yielding optimal resource utilization.

» **Fine-grained application optimization**

EdgeMesh works together with the UPFs to present revocable resources through the sidecars. Each application can be arranged with the most suitable resources, and allocated to an appropriate network plane for data transmission based on its users and services.

» **Efficient protocols**

EdgeMesh sidecars leverage the most suitable protocols to exchange data with devices.

2.2.8 EdgeData

Research organizations such as the IEEE have extensively researched data storage at the network edge. Their research has shown that a MEC storage system must be distributed, heterogeneous, lightweight, and perceivable. EdgeData is well positioned to offer this.

» **Distributed**

EdgeData guarantees data consistency among decentralized storage systems. These systems can effectively communicate with each other, and ensure redundancy, DR, and data collaboration between edge sites.

» **Heterogeneous**

Different storage software or medium can interoperate, and various types of data can be abstracted.

» **Lightweight**

Compact storage solutions can be provided for various scenarios. To illustrate this, a lightweight time series database (TSDB), such as TDEngine, can be used to store the monitoring data and service messages in IoT scenarios where a huge number of devices are interconnected.

» **Perceivable**

EdgeData can sense the privacy of core algorithms and data of certain applications, and encrypt them before storage. It then works with EdgeAI to determine whether data is steered locally or saved into other edge nodes.

2.2.9 EdgeFramework

EdgeFramework provides a solid framework for applications to run, and enables edge node owners to define policies, including charging and auditing, to ensure that the application behaves securely and as expected. EdgeFramework has the following features:

» **Generic**

EdgeFramework conforms to the Common API Framework (CAPIF) defined by 3GPP (see Figure 2-2), and supports interoperability over the Mp1 reference point as specified by ETSI. It also develops more functions beyond the 3GPP and ETSI standards, including API invoking across domains and even CAPIFs, distributed deployment, and end-to-end API authentication and authorization policies. EdgeFramework uses standard auditing points, such as invoking times and invoking parties, to provide important charging and auditing references for MEC node owners.

» **Adaptive**

A diverse variety of hardware models are sprawling along the network edge, and they use different protocols to interact with devices. To illustrate this, Industrial IoT (IIoT) services use protocols including Message Queuing Telemetry Transport (MQTT) and OPC Unified Architecture (OPC UA), whereas during cloud games, players use different terminals, such as consoles and keyboards, from which traffic is initiated towards the network edge. To adapt to such ever-varying service scenarios, EdgeFramework can be integrated with different protocol stacks as a plug-in. It can also abstract hardware resources, and hide differences in heterogeneous computing capabilities (x86, Arm, and AI) from applications. On top of this, it can collaborate with EdgeAI to provide a flexible application development framework.

» **Flexible**

EdgeFramework abstracts the MEP into functions and components which can be flexibly selected and combined. More specifically, the MEP can be arranged as core functions, including service registration, discovery, and authentication, as well as the AI, storage, and network components. These functions and components can be used on demand thanks to the design-time framework and Infrastructure as Code (IaC). For instance, if an application only relies on 5G network capabilities, only the core functions and 5G components are required.

2.2.10 EdgeDevSecOps

Developers tend to use simulated environments to verify MEC functions, especially when using actual 5G networks and devices is not feasible. DevSecOps provides everything required by MEC developers, including:

» **Simulated end-to-end environments**

In the end-to-end environments, not only edge networks are simulated, but also network capability exposure and 5G device behavior are imitated. The environments remain operational through testing on all MEC functions, such as local breakout (LBO) by the UPFs and location information exposure by the MEP.

» **Compressive security assurance**

Static security analysis tools ensure that vulnerable components will not be used by any application. The code and target deployment environment are comprehensively scanned to ensure that no known vulnerabilities exist. Similarly, test cases covering all service processes are provided to verify deployed applications, of which identity authentication and authorization are fully automated processes.

» **Customized tool chain or SDK**

The tool chain and SDK helps developers efficiently organize and invoke the lower-layer capabilities of telecom networks and converse code between different platforms (x86 and Arm).

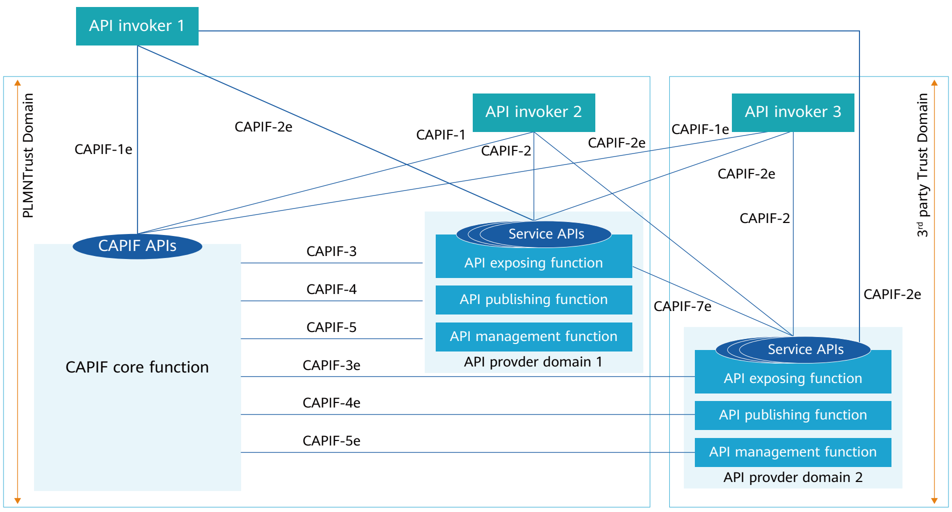


Figure 2-2 CAPIF architecture

03

Open Source and Practical Implementation

3.1 EdgeGallery

EdgeGallery is an open-source 5G MEC project co-initiated by the CAICT, China Mobile, China Unicom, Huawei, Tencent, 99cloud, Purple Mountain Laboratories, and DBAPPSecurity. Since it was first launched, the project has garnered significant attention from industry stakeholders, including equipment vendors, operators, and vertical industry players. Today, the EdgeGallery community includes 8 senior members and 17 regular members.

EdgeGallery aims to utilize Edge Native technologies to build a public MEP that combines connectivity and computing, standardizes the exposure of network capabilities (especially in terms of 5G network capabilities), and streamlines life cycle management processes for MEC applications. These life cycle management processes include development, testing, migration, and operations. In August 2020, the community announced that EdgeGallery was available on Gitee, officially designating it as open source. Since then, the community has been devoted to providing end-to-end services to application developers, edge operations personnel, and O&M engineers.

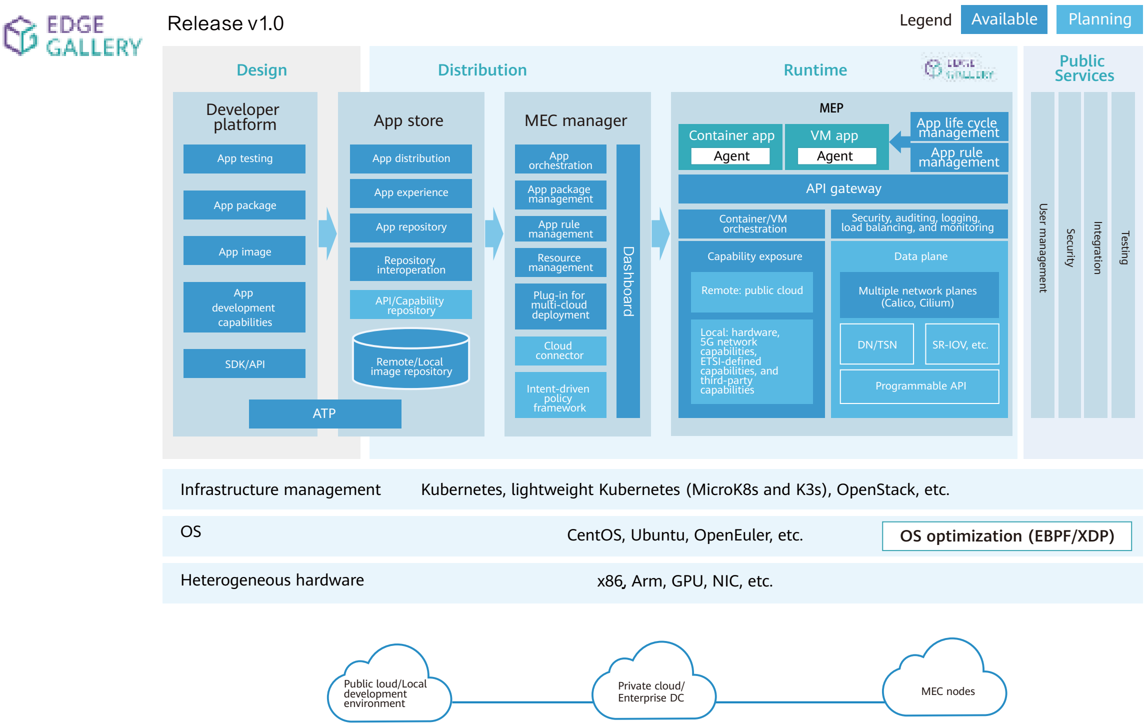


Figure 3-1 EdgeGallery architecture

EdgeGallery is comprised of four components:

» Developer platform

The platform uses EdgeFramework and EdgeDevSecOps to provide a unified entry for developers. It is adaptive to MEP APIs and conveys useful tools and processes for code integration as well as the packaging and testing of applications.

MEC application store

This component relies on EdgeSecurity. MEC applications are placed in a central repository, which is compatible with the application stores of community partners. Once feasibility is verified, a single application can be obtained from the central repository and launched in every store connected to the repository.

» MEC manager

The MEC manager, which uses both EdgeOrchestrator and EdgeCollaboration and connects to EdgeInfra, uniformly manages the life cycles of applications. It also provides intuitive UIs for engineers to monitor resources and applications, facilitating routine O&M.

» MEP

The MEP is a standard element in the MEC architecture defined by the ETSI, and pairs with EdgeNetwork, EdgeAI, EdgeData, and EdgeMesh. Open to a diverse array of hardware models and virtualized computing technologies, the MEP centrally provides APIs for managing applications and services as well as opening network capabilities.

To facilitate EdgeGallery application development across different platforms, and to contribute to a beneficial edge ecosystem, all industry stakeholders should promote the specifications for developing MEC applications. As shown in Figure 3-3, these specifications cover application packaging, registration, deployment, certification, security, and APIs, and lay a solid foundation for the formulation of de facto 5G MEC standards.

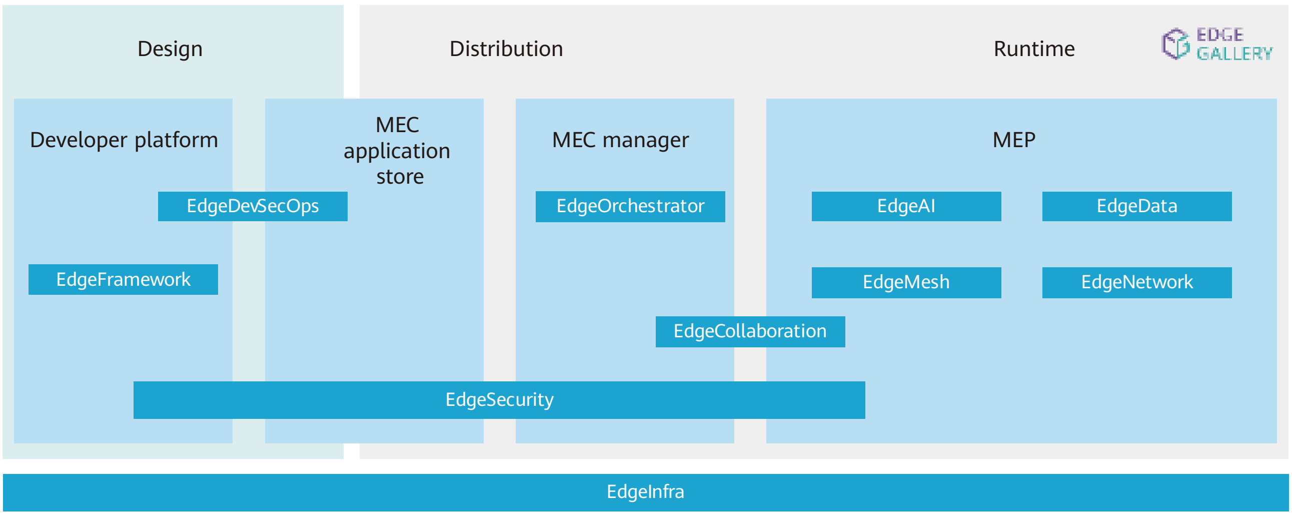


Figure 3-2 Mappings between EdgeGallery and Edge Native

Furthermore, EdgeGallery converges extensive lab resources (see Figure 3-4), including China's national future network experiment infrastructure, the 5G lab (sponsored by Purple Mountain Laboratories) in Nanjing, and the 5G experiment network in Shenzhen. These resources are available to labs run by community members where EdgeGallery is deployed, but where testing environments are independent from the community.

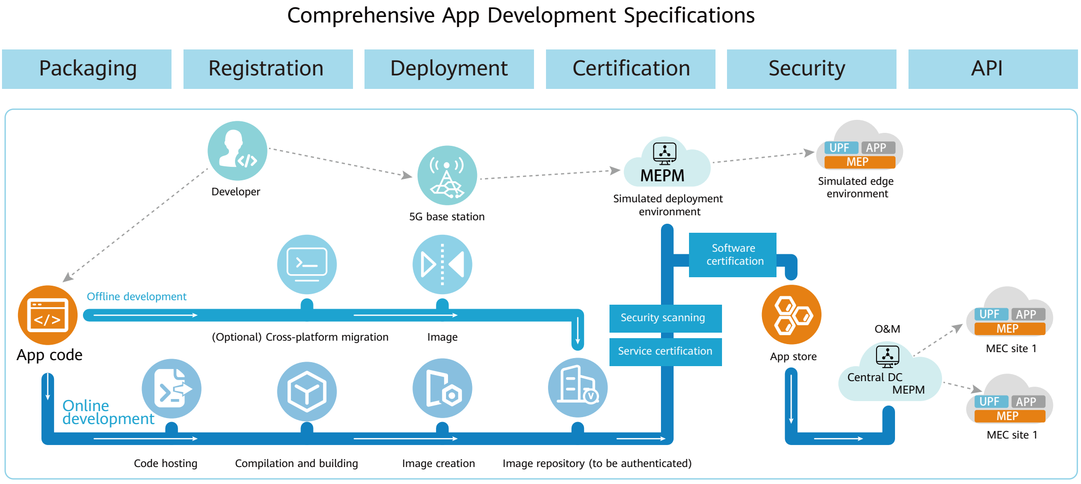


Figure 3-3 EdgeGallery application development specifications

The community has built a central platform, on which developers can check, apply for, and release lab resources for verification of the MEP functionality and applications as well as certification of services and software. An upgraded lab management platform will be launched in the upcoming EdgeGallery Release 1.1, capable of dynamically allocating testing resources to developers anytime, anywhere, and supporting the development of services in AR/VR, IoV, IIoT, and Smart City scenarios. Available resources include sandboxes and real 5G MEC environments, test SIM cards, computing resources (CPUs and GPUs), servers (x86 and Arm), as well as network and security test tools.

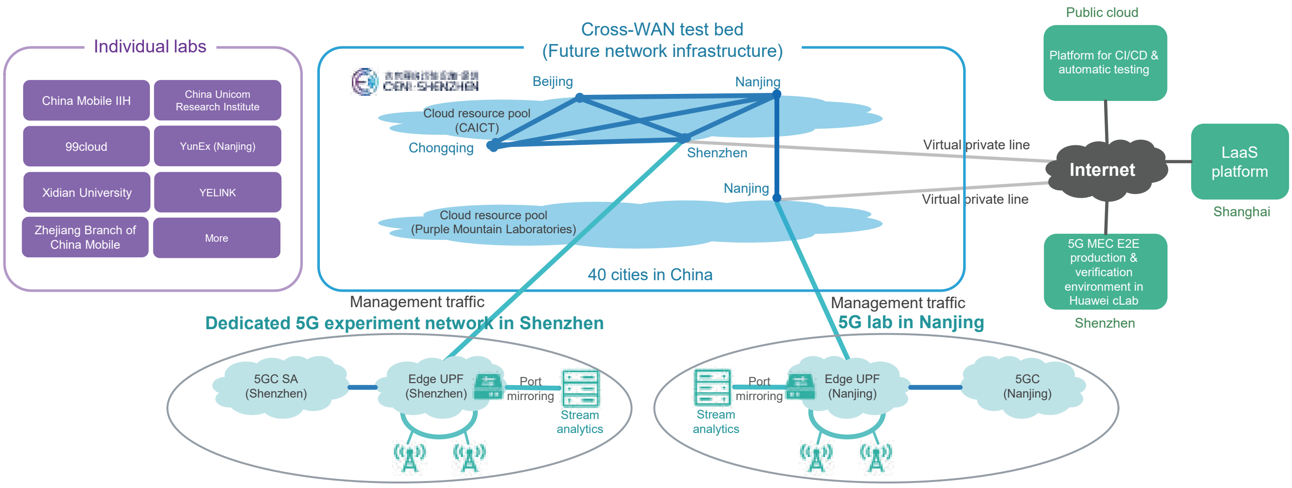


Figure 3-4 EdgeGallery lab resources

3.2 OpenSigma

To cope with 5G development trends and maximize each operator's unique advantages, China Mobile launched the 5G+E project, aiming to promote full-stack technologies, products, and capabilities for network-edge-cloud synergy. During this project, China Mobile recognized the increasingly prominent role played by MEC, and has since invested heavily in the R&D of MEC products. OpenSigma, unveiled at the 2020 China Mobile Global Partner Conference, represents one such product.

A generic MEP, OpenSigma provides all the basic MEC functionality, including LBO, application hosting, and domain name service (DNS) for enterprises and other industry players. In addition, it offers special edge services, manages edge cloud resources, and allows registration, charging, and launching of capabilities and applications. On the intuitive OpenSigma portal, industry partners can easily obtain required services and quickly monetize them.

China Mobile is actively contributing to the EdgeGallery community, which in turn provides valuable references for the construction of OpenSigma technical architecture (see Figure 3-5).

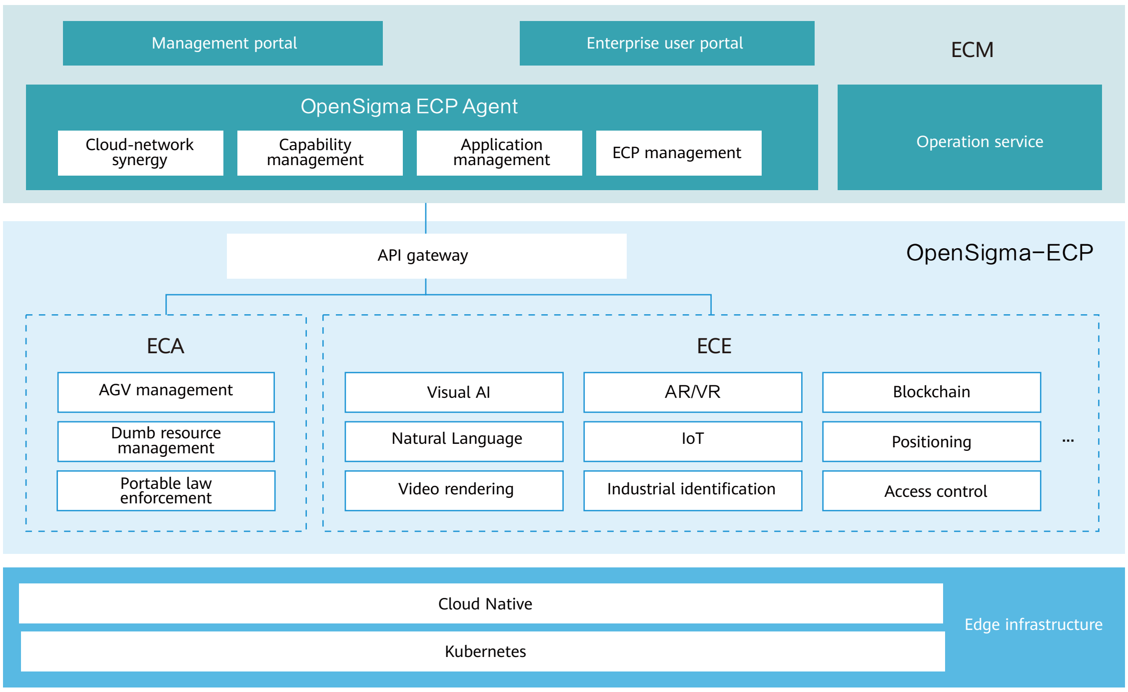


Figure 3-5 OpenSigma technical architecture

OpenSigma consists of two major elements:

» Edge Computing Management (ECM)

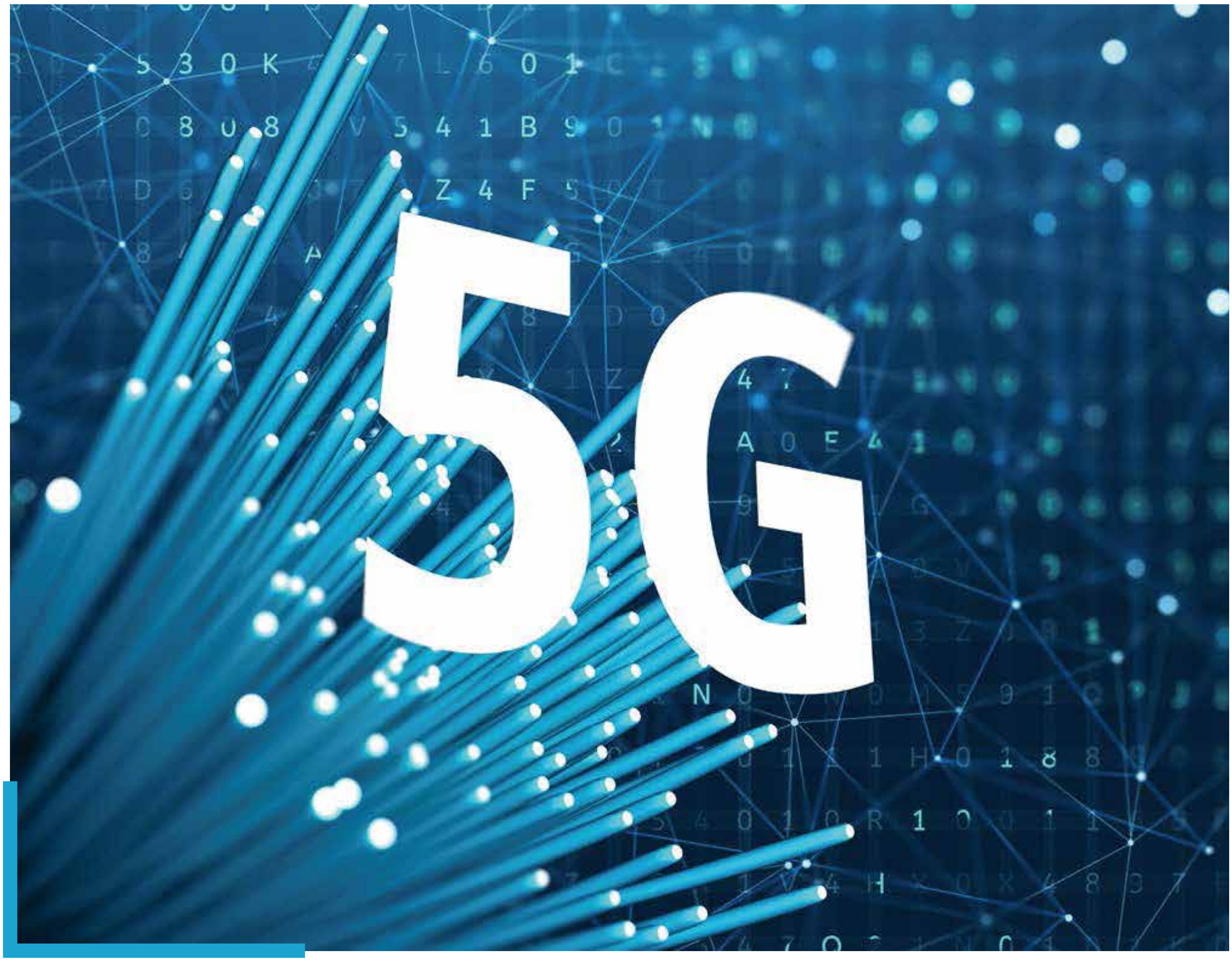
The ECM provides both management and service portals, and uses EdgeOrchestrator and EdgeCollaboration to manage the PaaS and support operations for network capabilities and services

» Edge Computing Platform (ECP)

The ECP relies on EdgeInfra to provide basic environments for IT applications, utilizes EdgeFramework to collaborate with the network capability exposure platform, and leverages EdgeNetwork and EdgeAI to open capabilities oriented to vertical industries through standard APIs. In addition, EdgeDevSecOps enables the ECP to iterate and update synchronously with the live network nodes through incubators, and to provide the environments and tools for connecting to third-party network capabilities and applications.

China Mobile has now built 156 MEC nodes in 22 provinces and autonomous regions across China, providing partners with various forms of super computing power over high-quality 5G networks. In addition, the company has deployed MEC service incubators in Beijing, Jiangsu, Zhejiang, Fujian, and Guangdong. China Mobile is currently recruiting partners from all over the world for their MEC application incubation project, with a goal to provide one-stop MEC development services, covering requirement clarification, environment adaptation, distribution, deployment, and commercial launch.

China Mobile and EdgeGallery will continue to improve the developer platform, application store, and application integration and testing; while also engaging in the research and application of core MEP capabilities. Together, they will standardize the MEC architecture, making it more adaptive to operator networks and more friendly to industry-specific applications; foster a sustainable edge ecosystem; and capitalize on the business opportunities introduced by 5G MEC.



3.3 CUC-MEC

5G is now fully operational around the world, and MEC represents a tremendous driving force. MEC-based edge clouds are unlocking new applications and services, which in turn fuel the digital transformation of all industries. Having pioneered in MEC development, China Unicom remains devoted to creating an agile, elastic, efficient, and open edge network, and has launched an in-house MEP — CUC-MEC. The CUC-MEC platform converges CT, IT, and operational technology (OT) to facilitate development, integration, operations, and innovation of applications from end to end, and to ensure that these applications are commercially feasible and easy to operate and maintain across various service scenarios, such as smart manufacturing, smart port, IoV, and smart healthcare.

China Unicom has built the world's largest MEC test bed and established an open lab, laying sound groundwork for the verification and promotion of the CUC-MEC platform and MEC applications. Compliant with the Edge Native architecture, the CUC-MEC platform can manage edge cloud resources and services for each MEC node, while also enabling customers or third-party developers to contribute applications to the application store through open APIs. Under the CUC-MEC-driven edge architecture, all MEC applications processes — including planning, building, maintenance, development, and operations — are standardized and streamlined.

In recent years, China Unicom has proactively participated in open-source projects carried out by various organizations, including the Open Networking Foundation (ONF), the Global System for Mobile Communications Association (GSMA), Akraino, EdgeGallery, ETSI, and 3GPP. Embracing Edge Computing 2.0, China Unicom aims to strengthen their focus on the building of open MEC networks, platforms, and ecosystems through deeper cooperation. In this regard, the company will collaborate with industry peers to improve the basic MEC architectures of resources, applications, security, and management, standardize network capability exposure, and synergize the network edge with public, industry, and private clouds. With help from EdgeGallery, China Unicom will continue to foster a sustainable edge ecosystem based on a diverse range of hardware resources. Applications will increasingly be born out of this ecosystem and will become easier to integrate, fast tracking the commercialization of 5G MEC.

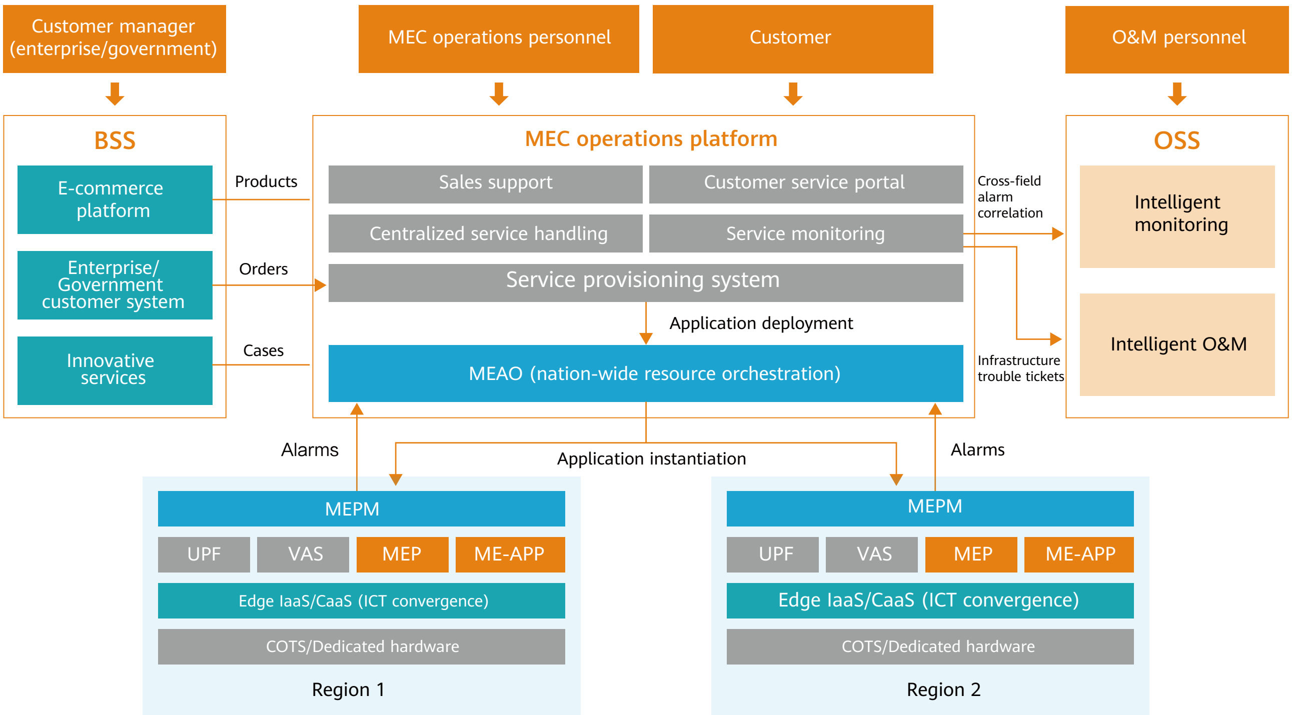


Figure 3-6 CUC-MEC architecture

04 Application Benchmarks and Outlook

Though still in its infancy, Edge Native has already been applied by a handful of leading enterprises, with both Haier and DALU Robotech setting industry benchmarks.

Haier: 5G+MEC Service Platform

Haier is one of world's largest white goods manufacturers and a Fortune 500 company. Similar to other such manufacturers, Haier previously faced a number of pain points relating to conventional manufacturing approaches — long product supply periods, delayed responses, lack of collaborative management methods, and deteriorating customer satisfaction. To address these challenges, Haier leveraged Edge Native and built a 5G+MEC service platform oriented to the home appliance industry. This platform now serves in a variety of industrial scenarios, including production optimization, collaborative production, quality management, flexible manufacturing, intelligent security protection, and energy management. It allows customers to influence products at the very beginning of the manufacturing process by transforming customer requirements directly into production orders, thereby enabling products to be personalized and mass-manufactured while striking a balance between costs, quality, production flexibility, and delivery period.

At the core of this platform is the machine vision-based quality inspection solution, powered by Edge Native technologies and delivered by China Mobile. This solution automatically identifies defects in various products, such as scratches and dents on stainless steel surfaces, imperfections in product doors, and damages to the exterior of products. Machine vision delivers a more thorough quality assurance process, with less risk of missed defects when compared with manual inspection. However, this approach demands much larger bandwidth and ultra-low latency, and the requirements vary significantly for different products. Such high requirements can only be satisfied by the combination of Edge Native and MEC.

Furthermore, the large network pipe and real-time transmission paths yielded by Edge Native facilitate effective collaboration of industrial cameras, visual processors, programmable logic controllers (PLCs), and robotic arms over wireless connections — enabling automation of the entire manufacturing process, reducing production costs, and enhancing product quality.

DALU Robotech: Intelligent Robots

DALU Robotech is a rising start-up specializing in the R&D of special robots and AI. Having recognized the key role played by MEC in cloud computing, intelligent work effect evaluation, and cloud-based storage of key data for robots, DALU Robotech saw the disruptive potential of emerging open-source technologies. Consequently, they placed their robot operating system and database onto the EdgeGallery platform. Thanks to the abundant network capabilities converged by EdgeGallery, robots can now operate in a more secure and stable manner, efficiently fulfilling their purposes with real time exchange of big data analytics, AI-powered analysis of video data, and identification of event behavior. In addition, thanks to capability exposure, customers can quickly deploy cloud and AI applications on the robot platform. In this way, robots can be more easily managed while efficiently responding to customer needs.

Since DALU Robotech began cooperating with EdgeGallery in April 2020, the two parties have completed a series of tests on MEC-based robot operations and carried out a number of practices. For example, robots can execute such tasks as automatic patrolling, intelligent charging, key information identification, and alarm reporting for security protection; they can monitor body temperatures in real time and disinfect specific areas to help prevent the spread of COVID-19 or other coronaviruses; and they can effectively assist in fire-fighting and riot control. Figure 4-1 shows the technical architecture adopted by DALU Robotech for their robots.

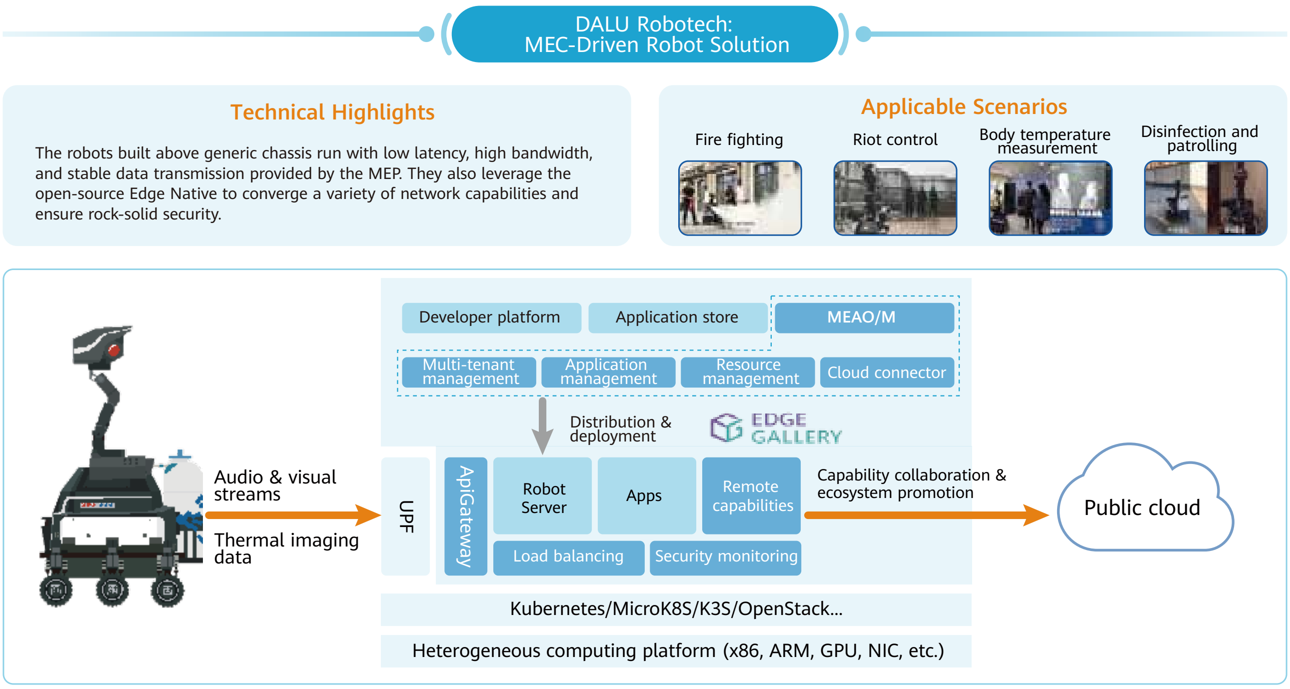


Figure 4-1 Technical architecture for MEC-based robot operations

The immense value offered by MEC and Edge Native is obvious, and the pilots conducted during 2020 represent an exciting first step towards a prosperous Edge Native future. Looking ahead, we invite all industry chain participants, including enterprises, communications equipment vendors, developers, and especially operators, to unite under a common goal — further advancing the concept, architecture, and technologies of Edge Native, and using these ever-advancing technologies to foster a sustainable ecosystem and stimulate business growth. Edge Native will dictate future network development and create unstoppable momentum for the digital upgrade of all industries.