Edge Computing Reference Architecture 2.0

2017

Jointly issued by the Edge Computing Consortium (ECC)
and Alliance of Industrial Internet (AII)
November, 2017

Contents

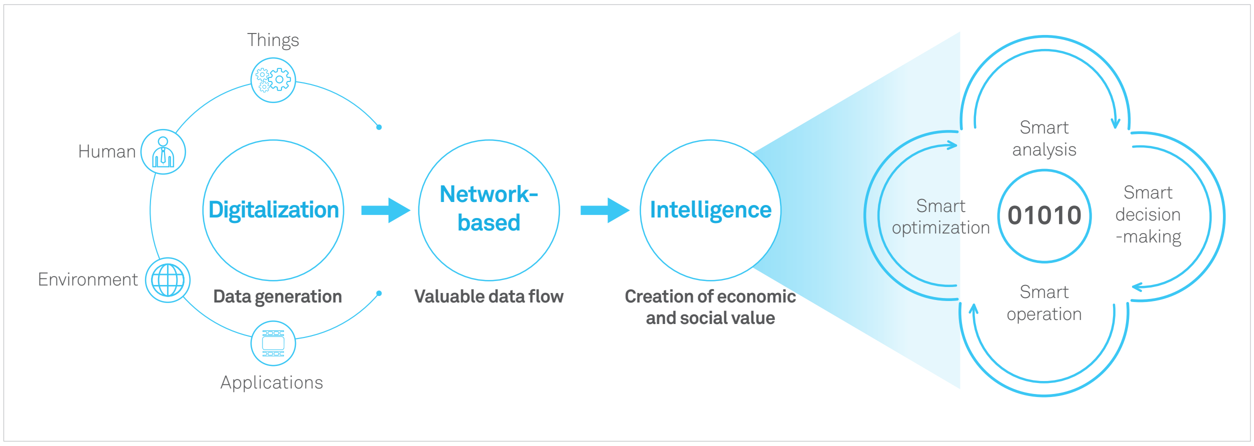
01
Embracing the Industry Intelligence Era

Embracing the Industry Intelligence Era

1.1 The Industry Intelligence Era Has Come

A wave of digital transformation of industries has started around the world. Digitalization is the foundation, network-based connectivity is the support, and intelligence is the goal. Digital transformation is to generate data from people, things, environments, and processes through digitalization, realize 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. With smart analysis of data as the basis, intelligence implements smart decision-making and operation, and achieves continuous, intelligent optimization of the business process through closed-loop control.

Figure 1-1 Digital transformation of industries



Intelligence technologies represented by Big Data, machine learning, and deep learning have been used in applications such as speech recognition, image recognition, and user profiling, and have seen significant progress in terms of algorithms, models, and architectures. Industries such as manufacturing, power, transportation, healthcare, and agriculture have begun to apply such technologies, which has created new requirements and challenges to intelligence technologies. The industry intelligence era has arrived.

Industry intelligence is defined by the following two phases:

1) Industry intelligence 1.0

Industry intelligence 1.0 is oriented to the business process covering market leads, marketing, purchase, logistics, and after-sales. It digitalizes the behavior and status of users, applications, and business process, and builds industry information maps based on multidimensional data analysis and scenario awareness. This allows it to provide personalized resource allocation and services for industry users.

Information and Communications Technology (ICT) supports fast development of industry intelligence 1.0:

- Ubiquitous network connections enable fast flow of data.
- Cloud computing provides low-cost infrastructure services on demand to adapt to service load changes.
- Big Data mining and analytics as well as management of large amounts of data help improve enterprises' business decision-making capabilities.
- Algorithms, data, and computing capabilities unleash the potential value of industry intelligence.

2) Industry intelligence 2.0

Industry intelligence 2.0 is oriented to the production process covering product planning, design, manufacturing, and operation. Products, production equipment, and manufacturing process have already started to become digitalized and network-based. Industry intelligence 2.0 has already met basic conditions. The concept of products and equipment referred to here is broad and includes not only the manufacturing lines and products produced in the manufacturing industry but also the assets on which industries such as energy, transportation, agriculture, and public utilities rely to provide services such as meters, vehicles, agricultural machinery, and environment monitoring equipment.

Industry intelligence 2.0 aims to:

- Improve agility and collaboration of the production and service processes.
- Increase resource sharing and save energy.
- Reduce uncertainties in production and operation.
- Collaborate with industry intelligence 1.0 to build E2E industry intelligence covering production, sales, and services.

The industry intelligence 2.0 era requires four key industry transformations:

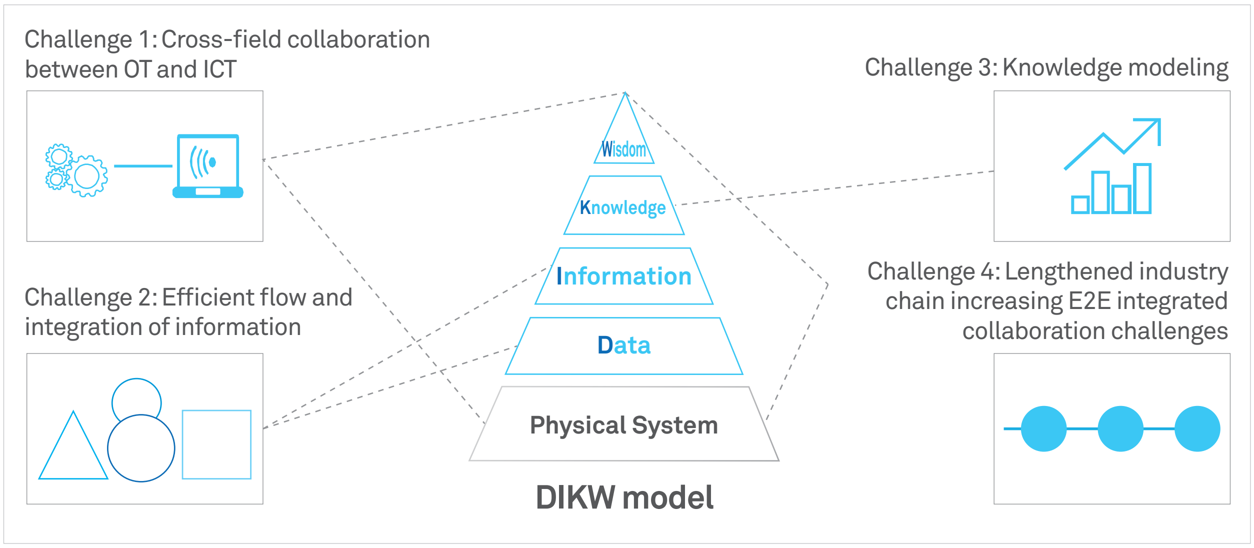
- Transformation from separation of the physical world and the digital world to integrated collaboration between the two worlds

- Transformation from obscure experience-based decision-making to scientific decision-making based on digitalization and modeling
- Transformation from process separation to full-process collaboration
- Transformation from unilateral innovation of enterprises to multilateral open innovation in an industry ecosystem

1.2 Challenges to Industry Intelligence 2.0

Industry intelligence 2.0 faces four challenges from the perspective of the DIKW model.

Figure 1-2 Challenges faced by industry intelligence 2.0



- **Cross-field collaboration between OT and ICT**

  Operational Technology (OT) is different from ICT. OT focuses on physical and business constraints and personal security, whereas ICT focuses on business constraints and information security. OT and ICT have major differences in terms of industry language, knowledge background, and cultural background. Integrated collaboration between the fragmented, specialized OT system and the standardized, open ICT system faces a number of challenges. Additionally, the integrated collaboration between OT and ICT will also bring security challenges.

  Cross-field collaboration between OT and ICT requires connectivity and convergence between the physical world and the digital world.

- **Efficient flow and integration of information**

  Currently, the industry has more than six industrial real-time Ethernet technologies and over 40 types of industrial buses. However, there are no uniform information and service models. Siloed systems lead to data silos, impeding efficient information flow and interaction.

  Efficient information flow and integration are the basis for supporting data innovation and service innovation. Full-lifecycle data management is required.

- **Knowledge modeling**

  The Knowledge Model solves problems associated with the expression, organization and interactive relationship, ordering, and processing modeling of knowledge. It implements formalized and structured abstraction of knowledge. The Knowledge Model is not knowledge; it is the abstraction of knowledge to facilitate interpretation and processing by computers.

  Knowledge Model input faces challenges in terms of incomplete, inaccurate, and insufficient information. Algorithms and modeling of Knowledge Model processing need to be continuously improved and optimized. Applications of Knowledge Model output are limited and need to be continuously accumulated.

  Knowledge modeling is a key element to efficiently realize industry intelligence with low costs.

- **Lengthened industry chain increasing E2E integrated collaboration challenges**

  Collaboration of the industry chain for the physical world and the digital world is required. Integration of full-lifecycle data on products is needed. Industry roles on the value chain need to establish a collaborative ecosystem. Such collaboration and integration of multiple links impose higher requirements on E2E data flow and full-lifecycle data management.

1.3 Edge Computing Enables Industry Intelligence 2.0

Edge computing needs to provide the following key capabilities to address challenges for industry intelligence 2.0:

1) Establish connections and interactions between the physical world and the digital world

With digital twins, a real-time mapping of diverse protocols, numerous devices, and physical assets across systems is established in the digital world to facilitate understanding the state of things or systems, dealing with changes, improving operations, and adding value.

Over the past 10 years, the network, computing, and storage fields, which are the three pillars of the ICT industry, have increased exponentially in terms of technical and economic feasibility.

- **Changes in the network field:** The bandwidth has increased more than 1000-fold while the cost has decreased 40-fold.
- **Changes in the computing field:** The cost of computing chips has reduced 60-fold.
- **Changes in the storage field:** The capacity of a single disk has increased more than 10,000-fold while the cost has reduced 17-fold.

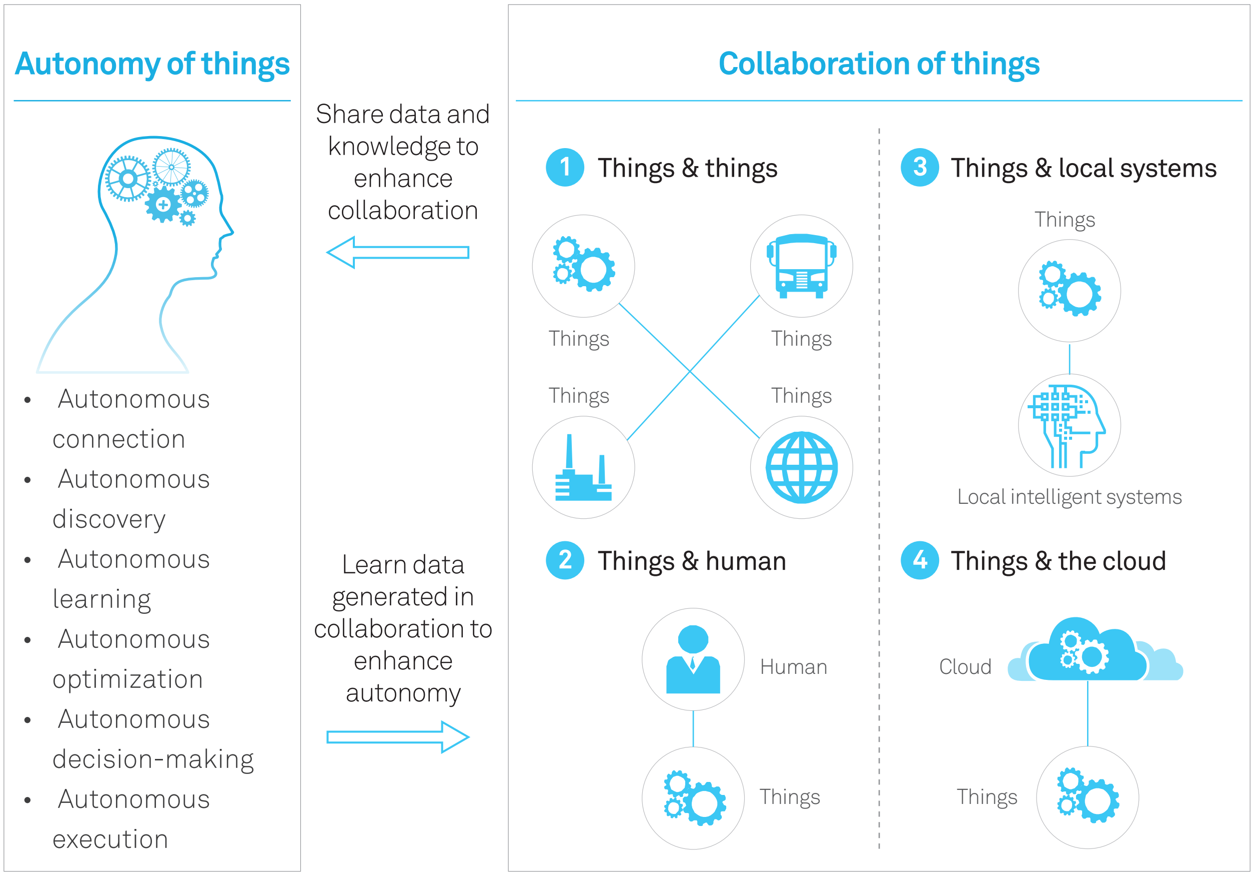
Due to reduction in connection costs, increase in computing capabilities, and large amounts of data, digital twins can play a significant role in the industry intelligence 2.0 era.

2) Provide a model-driven intelligent distributed architecture and platform

On the intelligent distributed architecture and platform on the network edge side, the Knowledge Model drives intelligence capabilities to enable autonomy and collaboration of things.

- **Autonomy of things:** Things can perform autonomous connection, discovery, learning, optimization, decision-making, and execution.
- **Collaboration of things:** including collaboration between things, between things and human, between things and local smart systems, and between things and the cloud.

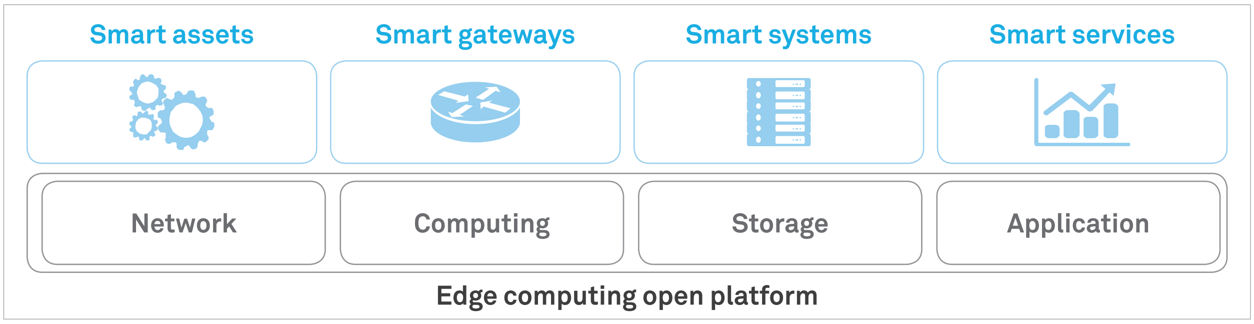
Figure 1-3 Intelligent distributed architecture



The intelligent distributed architecture enables:

- **Smart assets:** provide autonomy and collaboration capabilities through integrating ICT capabilities such as network, computing, and storage.
- **Smart gateways:** connect the physical world and the digital world through functions such as network connection and protocol conversion, and provide lightweight connection management, real-time data analysis, and application management.
- **Smart systems:** Smart systems are built based on collaboration of multiple distributed smart gateways or servers, and provide elastic network, computing, and storage capabilities.
- **Smart services:** Provide the development service framework and the deployment and operation service framework for various roles, including system operation and maintenance personnel, business decision-makers, system integrators, and application developers, based on the model-driven unified service framework.

Figure 1-4 Edge computing open platform enables industry intelligence 2.0



3) Provide the development service framework and the deployment and operation service framework

The development service framework mainly includes solution development, integration, verification, and release. The deployment and operation service framework mainly includes service orchestration, application deployment, and application market of solutions. The development service framework must closely collaborate with the deployment and operation service framework to support efficient development, automatic deployment, and centralized operation of solutions.

4) Collaborate with cloud computing

The network edge side needs to support multiple types of network interfaces, protocols, and topologies, real-time service processing, fixed latency, data processing and analysis, distributed intelligence, and security and privacy protection. The cloud side cannot meet such requirements. Edge computing needs to collaborate with cloud computing in networks, services, applications, and intelligence.

1.4 Current Progress of Industrialization of Edge Computing

Edge computing was added to Gartner's Hype Cycle in 2015.

A wave of industrialization of edge computing has been set off. Various industrial organizations and commercial organizations are actively initiating and promoting research, standardization, and industrialization of edge computing. Representative activities include:

- **Academic research**

  In October 2016, the first IEEE/ACM Symposium on Edge Computing was hosted by IEEE and ACM. It gained academic, industrial, and government (National Science Foundation of the United States) recognition. This academic symposium focuses on the application values and research orientations of edge computing.

- **Standardization**

  IEC released the Vertical Edge Intelligence (VEI) White Paper in 2017. The white paper describes the values of edge computing to vertical industries such as manufacturing.

  ISO/IEC JTC 1/SC 41 established the edge computing research group to promote standardization of edge computing.

- **Industry consortium**

  In November 2016, the Edge Computing Consortium (ECC) was set up by organizations including Huawei, Shenyang Institute of Automation (SIA) of Chinese Academy of Sciences, China Academy of Information and Communications Technology (CAICT), Intel, ARM, and iSoftStone.

  In 2017, Industrial Internet Consortium (IIC), a global industrial organization, established the Edge Computing TG to also define the reference architecture of edge computing.

02
Edge Computing

Edge Computing

2.1 Concept

Edge computing is a distributed open platform at the network edge, close to the things or data sources, integrating the capabilities of networks, storage, and applications. By delivering edge intelligence services, edge computing meets the key requirements of industry digitalization for agile connectivity, real-time services, data optimization, application intelligence, security and privacy protection.

Serving as a bridge between the physical and digital worlds, edge computing enables smart assets, smart gateways, smart systems, and smart services.

2.2 Basic Characteristics and Attributes

- **Connectivity**

  Connectivity is the basis of edge computing. The diversity of connected physical objects and application scenarios requires that edge computing provide abundant connection functions such as various network interfaces, protocols, topologies, network deployment and configuration, and network management and maintenance. Connectivity needs to fully draw on the advanced research achievements in the network field, such as TSN, SDN, NFV, NaaS, WLAN, NB-IoT, and 5G. Additionally, connectivity needs to consider interoperability with a variety of existing industrial buses.

- **First Entry of Data**

  As a bridge between the physical and digital worlds, edge computing is the first entry of data. With mass, real-time, and complete data, edge computing implements data management and creates values based on the data E2E lifecycle, supporting innovative applications such as predictive maintenance, asset efficiency and management. In addition, as the first entry of data, edge computing also faces the challenges caused by real-time, determinacy, and diversity.

- **Constraint**

  Edge computing products need to adapt to harsh working conditions and operating environments at industrial sites, such as anti-electromagnetic interference, anti-dust, anti-explosion, anti-vibration, and anti-current/voltage fluctuations. Moreover, in industrial interconnection scenarios, high requirements are imposed on the power consumption, cost, and space of edge computing devices.

Edge computing products need to be integrated and optimized through hardware and software to adapt to various conditions and constraints and support diverse scenarios of industry digitalization.

- **Distribution**

  In actual deployment, edge computing needs to support distributed computing and storage, achieve dynamic scheduling and unified management of distributed resources, support distributed intelligence, and deliver distributed security capabilities.

- **Convergence**

  Convergence of the Operational Technology (OT) and Information and Communications Technology (ICT) is an important foundation for the digital transformation of industries. As the key carrier of "OICT" convergence and collaboration, edge computing must support collaboration in connection, data, management, control, application, and security.

2.3 "CROSS" Value of Edge Computing

- **Mass and Heterogeneous Connection**

  Networks are the cornerstone of system interconnection and data aggregation transmission. With the surge in the number of connected devices, networks face enormous challenges in terms of Operations and Maintenance (O&M), management, flexible expansion, and reliability. In addition, a large number of heterogeneous bus connections have long existed at industrial sites, and multi-standard industrial Ethernet coexists. It is a tough issue that must be solved to achieve compatibility among multiple connections and ensure real-time reliability of connections.

- **Real-Time Services**

  Industrial system testing, control, and implementation have high real-time requirements, even within 10 milliseconds in some scenarios. If data analysis and control logic is implemented only on the cloud, it is difficult to meet the real-time requirements of services.

- **Data Optimization**

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

- **Smart Applications**

  Business process optimization, O&M automation, and service innovation drive applications to be smart. Edge intelligence delivers significant efficiency and cost advantages. Intelligent applications represented by predictive maintenance are driving industries to transition to new service models and business models.

- **Security and Privacy Protection**

  Security is critical to cloud and edge computing, requiring end-to-end protection. The network edge is close to Internet of Things (IoT) devices, making access control and threat protection difficult. Edge security includes device security, network security, data security, and application security. The integrity and confidentiality of key data, as well as protection of mass production or personal data are also key areas of focus for security.

2.4 Collaboration of Edge Computing and Cloud Computing

Cloud computing is suitable for non-real-time, long-period data and business decision-making scenarios, while edge computing plays an irreplaceable role in scenarios such as real-time, short-period data and local decision-making.

Edge and cloud computing are two important foundations for the digital transformation of industries. The collaboration between them in respect of network, service, application, and intelligence will help support more scenarios and unleash greater value in industry digitalization.

Figure 2-1 Points of collaboration between edge computing and cloud computing

Point of Collaboration	Edge Computing	Cloud Computing
Network	Data aggregation (TSN + OPCUA)	Data analysis
Service	Agent	Service orchestration
Application	Micro applications	Lifecycle management of applications
Intelligence	Distributed reasoning	Centralized training

03

Edge Computing Reference Architecture

Edge Computing Reference Architecture

3.1 Model-Driven Reference Architecture

The reference architecture is designed based on Model-Driven Engineering (MDE) principles. The model used for the reference architecture enables knowledge in the physical and digital worlds to be modeled to achieve:

- **Coordination Between the Physical and Digital Worlds**

  A real-time, systematic cognitive model of the physical world is established. The status of the physical world is predicted, and the running of the physical world emulated in the digital world. This approach simplifies reconstruction of the physical world and drives the physical world to optimize the operation of physical systems. Full lifecycle data of the physical world can coordinate with business process data to enable collaboration between business and production processes.

- **Cross-Industry Collaboration**

  Based on the modeling approach, the Information and Communications Technology (ICT) industry and vertical industries can build and reuse knowledge modeling systems in their own realms. The ICT industry shields the complexity of ICT technologies using the horizontal edge computing model and reference architecture. Each vertical industry performs modeling encapsulation of the industry know-how, achieving effective collaboration between ICT vendors and vertical industries.

- **Reduced System Heterogeneity and Simplified Cross-Platform Migration**

  Model-based interfaces between systems, subsystems, services, and new and legacy systems enable interaction, simplifying integration of these systems. Using the model, software interfaces can be decoupled from development languages, platforms, tools, and protocols, which reduces the complexity of cross-platform migration.

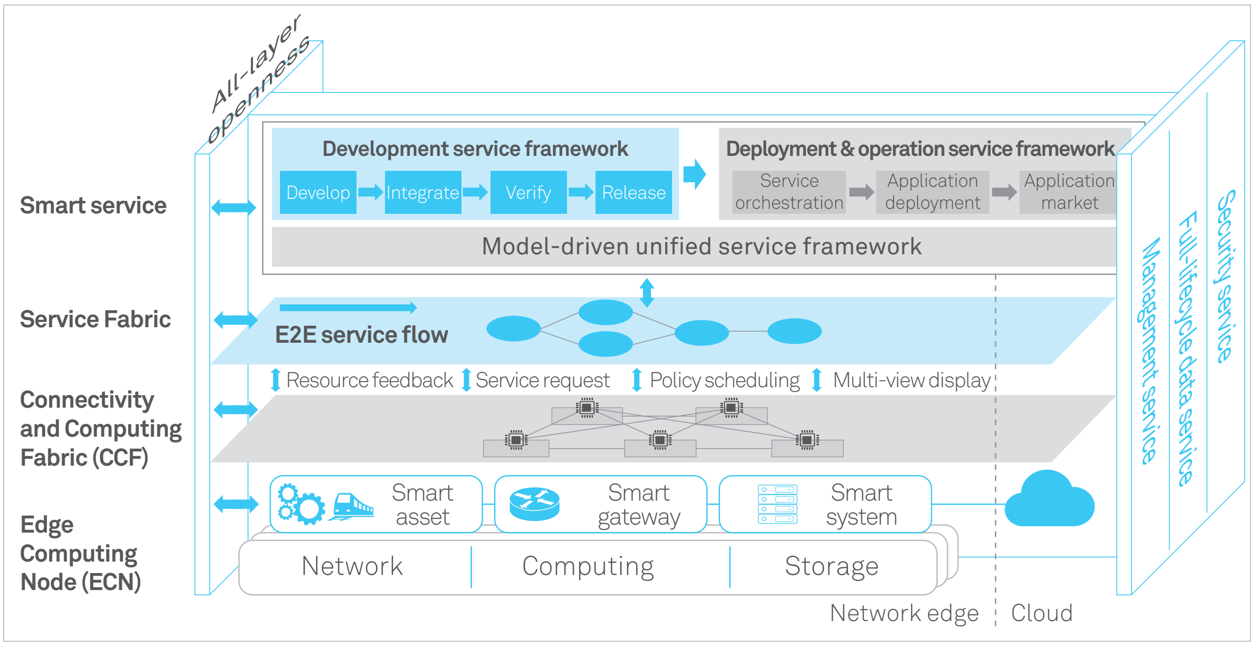
- **Effective Support for System Lifecycle Activities**

  System lifecycle activities include full lifecycle activities of application development services, deployment and operation services, data processing services, and security services.

The ICT industry faces challenges such as the need to simplify architectures, establish service intelligence, and reduce Capital Expenditure (CAPEX) and Operating Expense (OPEX) in the areas of networking, computing, and storage. To address these challenges, the ICT industry is adopting technological innovations such as virtualization, Software-Defined Networking (SDN), model-driven Service Orchestrator (SO), and microservices. Edge computing is an Operation Technology (OT) and ICT convergence industry, and the edge computing reference architecture design needs to draw on the new technologies and concepts. In addition, edge computing and cloud computing coordinate with each other yet differ in many ways. Therefore, edge computing faces unique challenges and requires unique and innovative technologies.

Based on the above concepts, the Edge Computing Consortium (ECC) proposes Edge Computing Reference Architecture 2.0.

Figure 3-1 Edge computing reference architecture 2.0



From the horizontal perspective, the architecture has the following characteristics:

- Smart services are based on the model-driven unified service framework. Intelligent coordination between service development and deployment is achieved through the development service framework and deployment & operation service framework. These frameworks enable consistent software development interfaces and automatic deployment and operations.
- Smart service orchestration defines the E2E service flow through the service fabric (SF) to realize service agility.
- Use of a Connectivity and Computing Fabric (CCF) enables a simplified architecture and simplifies the

distributed edge intelligence architecture for services. The CCF also enables automatic and visualized deployment and operations of the OT & ICT (OICT) infrastructure, supporting coordination between edge computing resource services and service needs of industries.

- Intelligent Edge Computing Nodes (ECNs) are compatible with a variety of heterogeneous connections, support real-time processing and response, and deliver integrated hardware and software security.

Edge Computing Reference Architecture2.0 provides model-based open interfaces at each layer, enabling full-layer openness. Vertically, the architecture uses management services, lifecycle data services, and security services to deliver smart services in the entire service process and full lifecycle.

3.2 Multi-View Display

Guided by international standards defined by ISO/IEC/IEEE 42010:2011, the architecture systematically addresses the industry's concerns about edge computing and presents solutions and frameworks. Edge Computing Reference Architecture2.0 is demonstrated using the following views:

- **Concept View**

  Describes the domain models and key concepts of edge computing.

- **Function View**

  Describes the functions and design concepts of the development service framework, deployment and operation framework SF, CCF, and ECN in the horizontal direction, as well as of cross-layer open services, management services, lifecycle data services, and security services in the vertical direction.

- **Deployment View**

  Describes the system deployment process and typical deployment scenarios.

In addition, the architecture needs to meet typical cross-industry, non-functional requirements, including real-time performance, certainty, and reliability. To this end, related technical solution recommendations are provided in the function view and deployment view.

3.3 Concept View

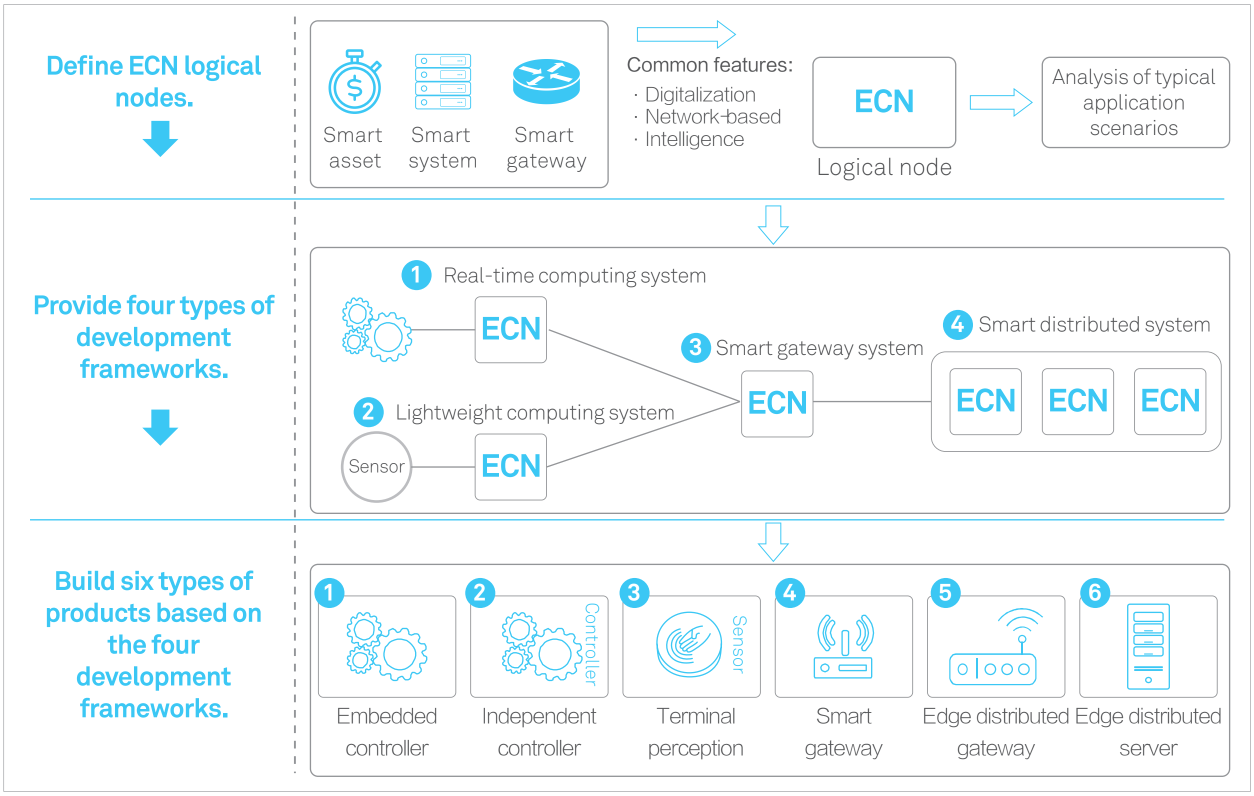
3.3.1 ECNs, Development Frameworks, and Product Implementation

Smart edge computing assets, systems, and gateways are digitalized, network-based, and intelligent. They provide ICT resources such as networks, computing, and storage, and can be logically abstracted as ECNs.

To suit the typical application scenarios of ECNs, the architecture defines four types of ECN development frameworks. Each framework includes an operating system, functional modules, and integrated development environment to meet the needs of various scenarios.

Based on the ECN development frameworks combined with the specific hardware platform required by ECNs, six types of products can be built, as outlined in Figure 1-2.

Figure 3-2 Concept view: ECNs, development frameworks, and product implementation



Typical functions of ECNs include:

- **Bus protocol adaptation**
- **Real-time connection**
- **Streaming real-time data analysis**

- **Sequential data access**
- **Policy execution**
- **Device plug-and-play**
- **Resource management**

Development frameworks of ECNs include:

- **Real-Time Computing System**

  For digital physical assets; meets the needs of real-time applications.

- **Lightweight Computing System**

  For sensing terminals with limited resources; meets the needs of low power consumption.

- **Smart Gateway System**

  Supports multiple network interfaces, bus protocols, and network topologies; enables interconnection of local edge systems; delivers local computing and storage capabilities; and enables interworking with cloud systems.

- **Smart Distributed System**

  Based on a distributed architecture, this framework flexibly expands network, computing, and storage capabilities at the network edge; supports service-oriented dynamic resource management and scheduling; and enables interworking with cloud systems.

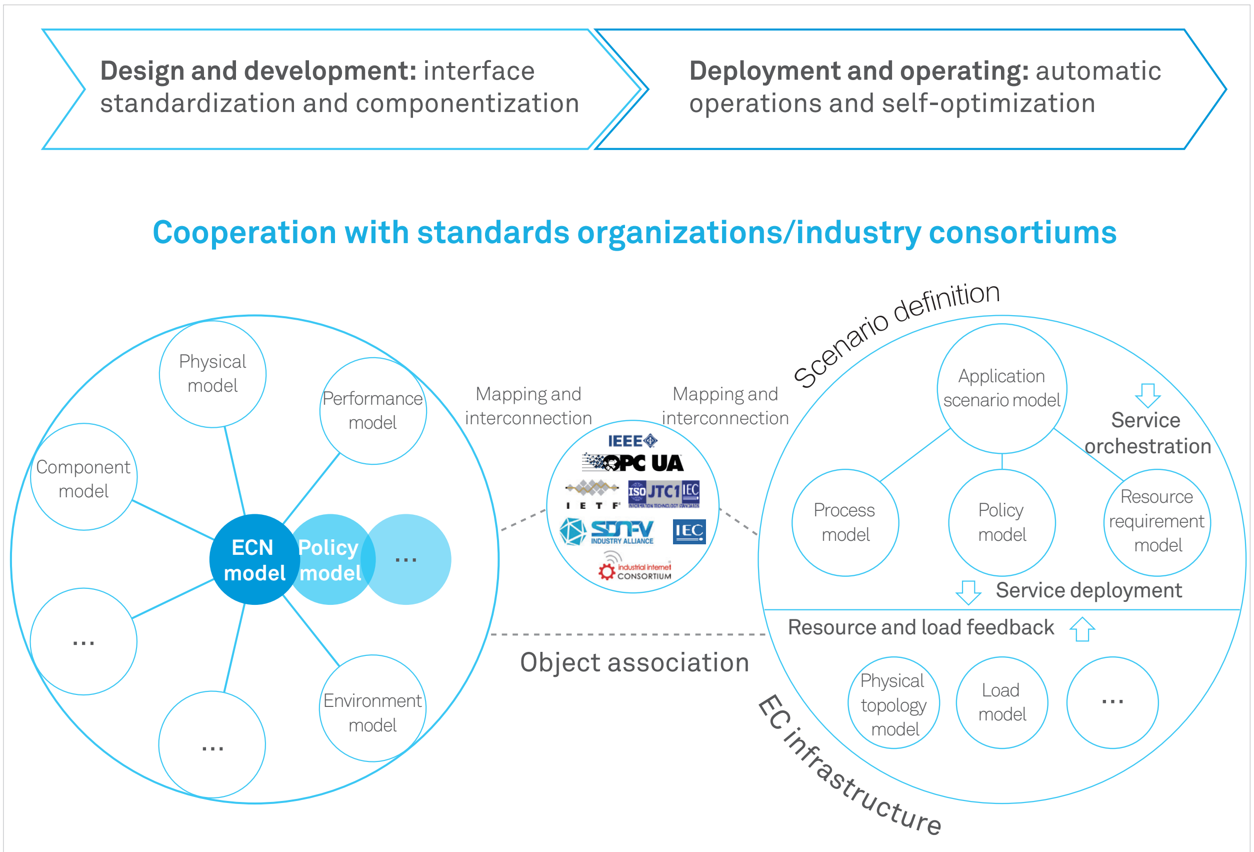
ECN product implementation:

Product	Typical Scenario
ICT-converged gateway	Connection of elevators, smart street lamp
Independent controller	Industrial Programmable Logic Controller (PLC)
Embedded controller	Virtual Programmable Logic Controller (vPLC), robot
Sensing terminal	Computer Numerical Control (CNC), instrument
Distributed service gateway	Smart power distribution
Edge cluster (edge cloud)	Digital workshop

3.3.2 Edge Computing Domain Models

Edge computing domain models are defined from the ICT perspective of edge computing.

Figure 3-3 Concept view: full lifecycle-oriented models



- **Models in the Design Phase**

  Define the IDs, attributes, functions, performance, and derivation and inheritance relationships of ECNs, providing valuable information for the deployment and operation phases.

- **Models in the Deployment Phase**

  Include service policy and physical topology models. The service policy model describes service rules and constraints using a service language rather than machine language, allowing services to drive the edge computing infrastructure. The service policy model is reusable and changeable, enabling service agility.

- **Models in the Operation Phase**

  Include the connection computing fabric model and operation load model. Based on these models, the system operating status can be monitored and optimized, and the deployment of loads optimized on the distributed edge architecture.

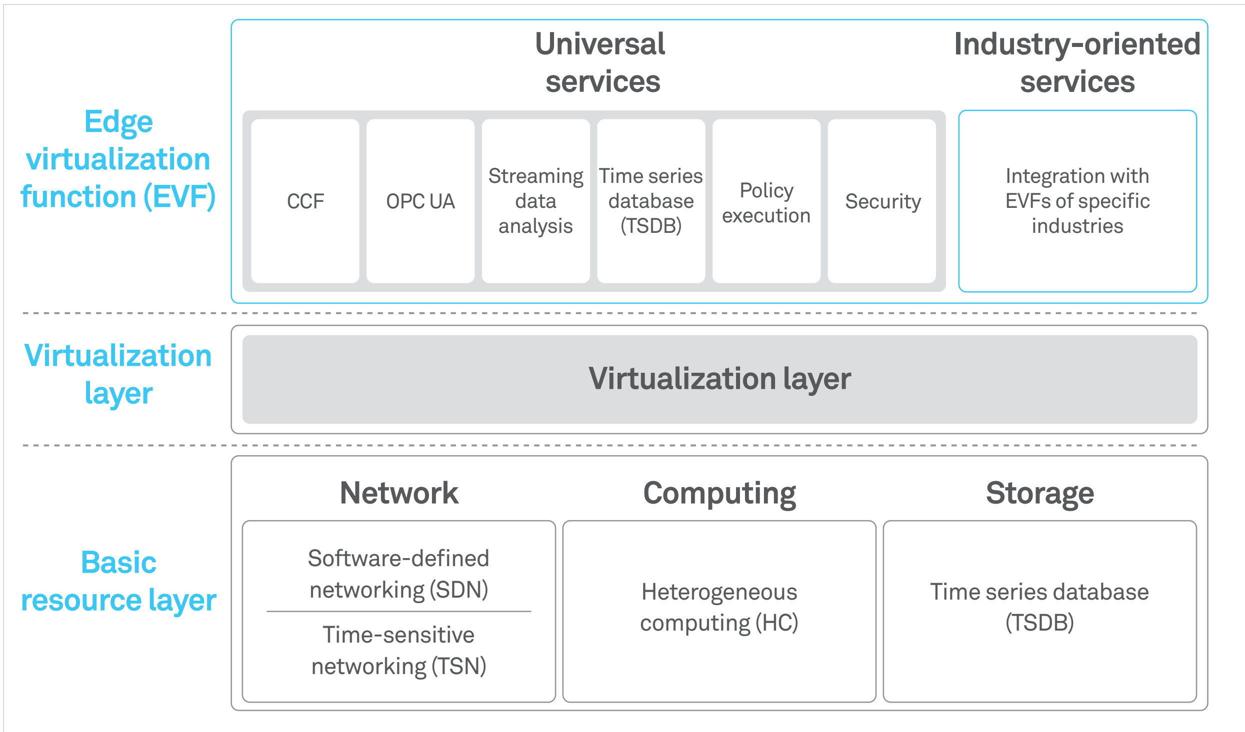
Through the model-driven unified service framework, edge computing domain models and vertical industry models can be mapped to each other and centrally managed. In this way, vertical industry models such as OPC Unified Architecture (OPC UA) and its ecosystem can be reused, allowing for easy integration of the edge computing reference architecture with industry platforms and applications.

3.4 Function View

3.4.1 ECN

The ECN architecture comprises three main layers, as outlined in Figure 1-4.

Figure 3-4 Function view: ECN functional layers



1) Basic Resource Layer

This layer includes network, computing, and storage modules.

- **Network**

  SDN architectures separate a network's control plane from the forwarding plane to make the network programmable. When SDN is applied to edge computing, the network can support access to millions of network devices as well as flexible expansion, enabling efficient and low-cost automatic O&M. Additionally, this approach helps achieve network and security policy association and integration.

Network connections need to accommodate requirements for transmission time certainty and data integrity. Time-Sensitive Networking (TSN) series standards define unified technical specifications for key services such as real-time priorities and clocks. TSN is the future development direction of industrial Ethernet connectivity.

- **Computing**

Heterogeneous Computing (HC) is a crucial aspect of the computing hardware architecture at the network edge. Even as Moore's Law continues to hold true for breakthroughs in chip technologies, the popularity of IoT applications has brought explosive growth in information volume, and the application of Artificial Intelligence (AI) has increased computing complexity. These developments place higher requirements on computing capabilities. The types of data to be processed are also becoming more diversified. As a result, edge devices need to process both structured and unstructured data. In addition, as ECNs contain more compute units and different types of compute units, costs become a concern.

Therefore, a new computing architecture is proposed that combines compute units that handle different types of instruction sets and have different architectures, that is, heterogeneous computing. Such an architecture gives full play to the advantages of various compute units, achieving a balance between performance, cost, power consumption, and portability.

Additionally, a new generation of AI technology, represented by deep learning, needs new technical optimizations at the network edge. Currently, the processing of a single picture at the inference stage often requires more than 1 billion computations, so standard deep learning algorithms are obviously not suitable for embedded computing environments at the network edge. Ongoing optimization for deep learning in the industry includes top-down optimization, which compresses learned deep-learning models to reduce computational load at the inference stage. Bottom-up optimization that redefines an algorithm architecture oriented for edge-side embedded system environment is also being attempted.

- **Storage**

The digital world needs to keep track of the dynamics of the physical world in real time and store complete historical data in chronological order. A new generation of Time Series Database (TSDB) offers efficient storage for time series data (including information such as timestamps of the data). TSDBs need to support basic functions of time series data, such as fast write, persistence, and multi-dimensional aggregated query. To ensure data accuracy and completeness, TSDBs need to continuously add new time series data instead of updating the original data. These requirements raise the following challenges:

- The time series data write function must support writing tens or hundreds of millions of data points per second.
- The time series data read function must support packet aggregation operations on hundreds of millions of data values within seconds.
- The cost-sensitive nature of most scenarios means that the top priority of TSDBs is to reduce the cost of mass data storage

2) Virtualization Layer

Virtualization technology reduces system development and deployment costs, and has been adopted into embedded system applications from server applications. Typical virtualization technologies include the bare metal architecture and host architecture. In the bare metal architecture, virtualization-layer functions such as the hypervisor run directly on the system hardware platform, and then operating system and virtualization functions run under the hypervisor. In the host architecture, virtualization-layer functions run under the host operating system. The bare metal architecture has better real-time performance and is generally used by smart assets and smart gateways.

3) EVF Layer

Edge Virtualization Functions (EVFs) are software-based and service-based functions that are decoupled from a proprietary hardware platform. Based on virtualization technology, the hardware, system, and specific EVFs can be vertically combined based on services on the same hardware platform. In this manner, multiple independent service zones can be virtualized and isolated from each other. ECNs' service scalability reduces CAPEX and extends a system's lifecycle.

EVFs can be flexibly combined and orchestrated, and migrated and expanded on different hardware platforms and devices, enabling dynamic resource scheduling and service agility.

The EVF layer delivers the following basic services that can be tailored:

- Distributed CCF service
- OPC UA service
- Streaming real-time data analysis service
- TSDB service
- Policy execution service
- Security service

Key ECN Technologies:

1) SDN

SDN uses a completely different control architecture from traditional networks. It separates the network control plane from the forwarding plane, replaces the original distributed control with centralized control, and implements "software-defined" through open, programmable interfaces. SDN, as a new technology, changes the way a network is built and operated: building networks from an application perspective and operating networks using Information Technology (IT).

The SDN architecture includes controllers, southbound/northbound interfaces, and various application-layer applications and infrastructure-layer Network Elements (NEs). The most important part of the architecture is the SDN controller. It implements configuration and management of forwarding policies at the infrastructure layer and supports forwarding control based on multiple flow tables.

SDN's unique benefits for edge computing include:

- **Mass Connections**

  SDN supports access to millions of network devices and flexible expansion. SDN also integrates and adapts to the management of multi-vendor network devices.

- **Model-Driven Policy Automation**

  SDN provides flexible network automation and management frameworks; enables service-based infrastructure and service delivery functions; and implements plug-and-play smart assets, gateways, and systems. These capabilities greatly reduce the technical requirements for network administrators.

- **E2E Service Protection**

  SDN delivers E2E tunnel services, such as Generic Routing Encapsulation (GRE), Layer 2 Tunneling Protocol (L2TP), Internet Protocol Security (IPSec), and Virtual Extensible LAN (VXLAN). SDN also optimizes Quality of Service (QoS) scheduling, helps meet key requirements such as E2E bandwidth and delay specifications, and implements edge-to-cloud service coordination.

- **Lifecycle Management of Applications**

  SDN supports lifecycle management tasks such as application deployment, loading, update, uninstallation, and deletion. SDN also supports multi-application resource scheduling and management, including priority enforcement, security, and QoS.

- **Architecture Openness**

  SDN opens centralized network control and network status information to smart applications so that they can flexibly and quickly drive network resource scheduling.

  Edge computing SDN technology has been successfully applied to smart buildings, smart elevators, and many other industry scenarios.

2) TSN

Standard Ethernet technologies have been widely implemented, with advantages such as high transmission speed, flexible topology, long transmission distance, and cost-effectiveness. Meanwhile, due to constraints from the traditional Quality of Service (QoS) mechanism and the Carrier Sense Multiple Access with Collision Detection (CSMA/CD) mechanism, Ethernet technologies cannot meet key industry requirements for timeliness and determinism. The industry optimizes standard Ethernet technologies and offers commercial implementations of multiple industrial real-time Ethernet technologies. Consequently, a variety of industrial real-time Ethernet networks coexist, creating obstacles and challenges for interoperation.

In recent years, IEEE802.1 defined the TSN technical standards, aiming to promote standardization and interoperability of real-time Ethernet networks, and ultimately merging Operational Technologies (OT) and ICT using 'one network.' This also brings the following advantages:

- Ensures determinism with μs-level latency and jitter of less than 500 ns.
- Meets large bandwidth requirements for scenarios such as industrial machine vision, with interface bandwidth of larger than 1 Gbit/s.
- Achieves reliable data transmission through multiple paths or redundant paths.
- Coordinates with SDN technologies to achieve unified scheduling management of TSN and non-TSN networks.

TSN is designed to provide a unified low-latency queue scheduling mechanism, resource reservation mechanism, clock synchronization mechanism, path control mechanism, and configuration management model at the Media Access Control (MAC) layer of Ethernet networks, to achieve interoperation between TSN networks and standard Ethernet networks.

Currently, a good industrial collaboration ecosystem for TSN has been established. For example, IEEE is responsible for standards establishment, Avnu Alliance is responsible for interoperability certification, and industrial organizations represented by the ECC and the Industrial Internet Consortium (IIC) are performing industry demonstrations and promotion through test beds and other activities.

3) HC

Heterogeneous computing architecture is designed to coordinate and bring into play unique advantages of various computing units: The Central Processing Unit (CPU) manages system control, task decomposition, and scheduling; the Graphics Processing Unit (GPU) has strong floating-point and vector computing capabilities, and performs well with parallel computing such as matrix computing and vector computing; the Field Programmable Gate Array (FPGA) unit provides advantages such as hardware programmability and low latency; the Application-Specific Integrated Circuit (ASIC) unit offers advantages such as low power consumption, high performance, and cost-effectiveness.

The objective of heterogeneous computing is to integrate separate processing units of the same platform to collaboratively execute different types of computing loads. Moreover, heterogeneous computing achieves cross-platform deployment of software through open, unified programmable interfaces.

The heterogeneous computing architecture uses the following key technologies:

- **Memory processing optimization**

  With a traditional architecture, data transfer between different computing units requires data replication, which not only occupies processor resources, but also occupies a large amount of system bus bandwidth. Heterogeneous computing enables uniform memory access of multiple computing units. Data of any processing unit can be easily accessed by other processing units, without the need of copying the data to each other's memory area, which greatly improves system performance.

- **Task scheduling optimization**

  The relationship of various computing units changes from master-slave relationship to equal partnership. The most appropriate computing unit is dynamically determined to execute the workload based on tasks. Heterogeneous computing involves a series of optimizations, including those of scheduling algorithms, instruction sets, and compilers.

- **Tool chain for development**

  The tool chain provides application programmers with hardware and software interfaces as well as a basic runtime environment; encapsulates and hides complicated bottom-layer details such as memory consistency and task scheduling management; supports optimization of architecture parameters and task scheduling; and minimizes the application porting workload. For Artificial Intelligence (AI) applications, this tool chain integrates multiple open AI training and reasoning platforms and is compatible with multi-vendor computing units.

Heterogeneous computing is currently used in both chip design and edge computing platform design. In terms of chips, heterogeneous computing integrates CPU and GPU resources to accelerate video encoding and decoding. In terms of computing platforms, heterogeneous computing uses CPU and FPGA (or GPU) resources to achieve implementation of AI functions in areas such as smart transportation and smart robots.

4) Time Series Database (TSDB)

Efficient writing, query, and distributed storage of large amounts of data are key challenges to TSDBs. TSDBs use the following key technologies:

- **Distributed storage**

  The core for distributed storage is how to distribute data to multiple machines, that is, data fragmentation. Data fragmentation can be implemented based on timestamps, tags, and priorities. Data fragments that have the same tag (one or more identical fields), are generated within the same time range, and match priority conditions are stored on the same machine. Data can be compressed before it is stored, which improves data writing efficiency and saves storage space.

- **Priority-based storage**

  Using timestamps of time series data as the priority division basis is very efficient. Data that was processed recently is queried for more times and is considered hot data. Data that was processed a long time ago is queried less often and is considered cold data. In addition, factors such as storage costs are often considered in priority-based storage. Data with different priorities is stored on storage media with different costs (including memories, HDDs, and SSDs).

- **Fragment-based query optimization**

  During data queries, all data segments are queried based on query conditions. All of the fragments are merged based upon the timestamp conditions to generate the original data results. If the query conditions include aggregation operations on data, these operations are performed based on the time sampling window for returning the results.

Besides commercial versions, the industry has a large number of open-source TSDBs, such as OpenTSDB, KairosDB, and InfluxDB. In addition to meeting performance challenges, TSDBs need to provide industry data modeling and visualization tools and to support rapid integration with industry application systems.

3.4.2 Service Fabric

A service fabric is a model-based workflow that digitally represents service requirements. It consists of multiple types of logically related functional services that collaborate with each other to implement specific service requirements.

The service model includes the following information:

- Service name
- Function to be executed or provided
- Nesting, dependency, and inheritance relationships between services
- Input and output of each service
- Service constraints such as QoS, security, and reliability

The service types include universal services provided by edge computing and specific industry services defined by vertical industries.

A service fabric has the following features:

- Focus on service processes and shield technical details, helping service departments, development departments, and deployment operation departments establish effective cooperation.
- Decouple from the OICT infrastructure and hardware platform to implement cross-technology platforms and support service agility.
- As a service description model, it can be inherited and reused to implement fast modeling.

A service fabric provides the following functions:

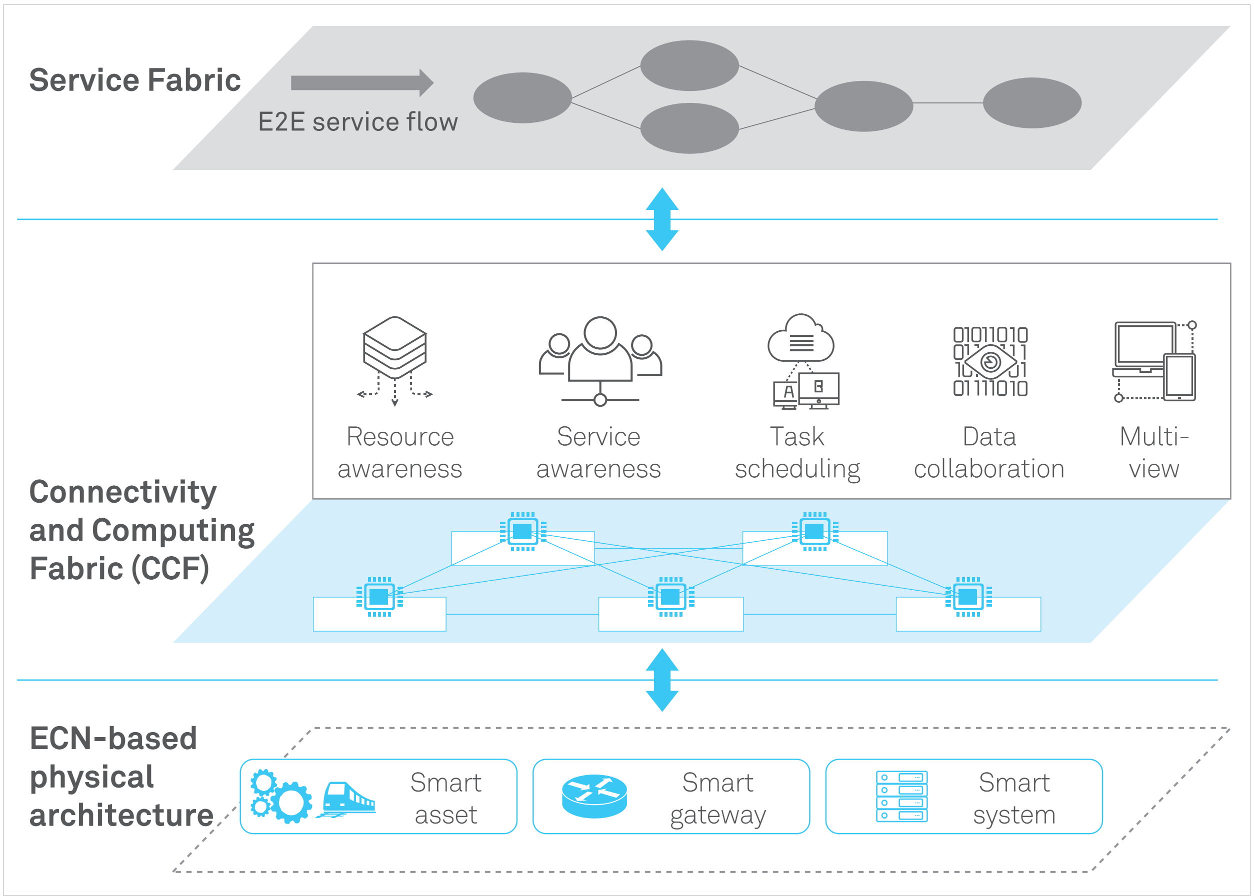
- Workflow and workload definition
- Visualized display
- Semantic check and policy conflict detection
- Version management of fabric and service models

3.4.3 CCF

The CCF is a virtualized connectivity and computing service layer. It has the following features:

- Shield the heterogeneity between ECNs.
- Reduce the complexity of the smart distributed architecture in terms of data consistency and error tolerance.
- Implement the discovery, unified management, and orchestration of resource services.
- Support the sharing of data and knowledge models between ECNs.
- Support dynamic scheduling and optimization of the service load.
- Support distributed decision-making and policy execution.

Figure 3-5 Function view: CCF



A CCF provides the following functions:

a. Resource awareness

The CCF can detect the ICT resource status (such as network connection quality and CPU usage), performance specifications (such as real-time performance), and physical information (such as location) of each ECN, providing key input for resource allocation and scheduling at the edge.

b. EVF awareness

The CCF can detect the EVFs provided by the system, ECNs where the EVFs are distributed, computing tasks that each EVF serves, and task execution status.

c. Workload scheduling

The CCF supports proactive task scheduling. It can automatically divide the workload into multiple subtasks based on resource awareness, service awareness, and service constraints (such as the connection bandwidth and delay requirement between ECNs) and allocate them to multiple ECNs for collaborative computing. In addition, the CCF can open the resource and service information for the service fabrics through open interfaces so that the service fabrics can automatically control the scheduling process of the workload.

d. Data collaboration

ECNs adapt to the southbound multi-bus protocol. The east-west connections between ECNs use a unified data connection protocol. Through data collaboration, ECNs can exchange data and knowledge models with each other. The exchange modes include simple broadcast and Pub/Sub.

e. Multi-view display

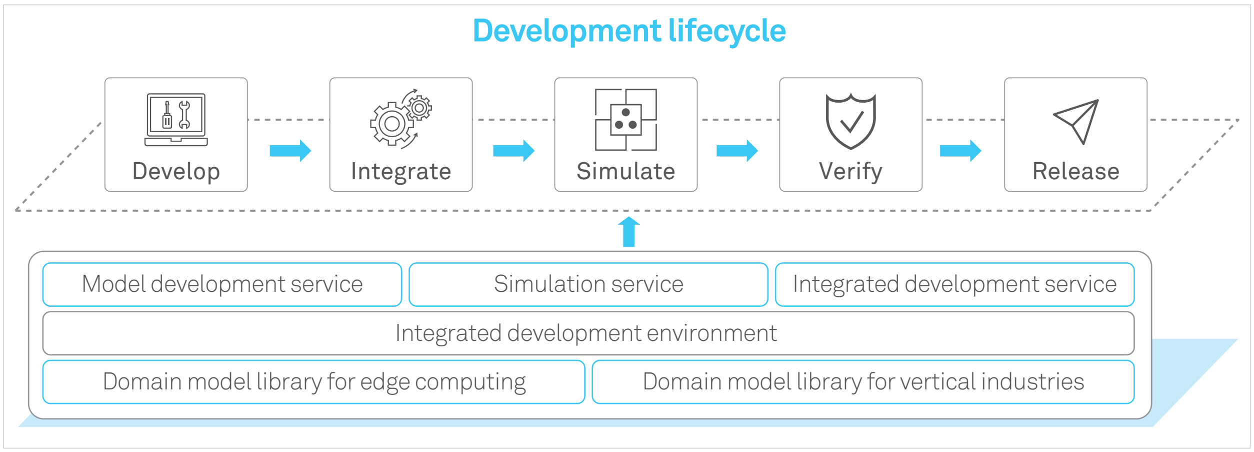
Services can be displayed by tenant or service logic, shielding physical connection complexity. For example, each tenant only needs to view its own workload, distribution of the workload on the CCF, and resource usage.

f. Open service interfaces

The CCF provides workload requests, resource status feedback, and scheduling and execution status feedback through open interfaces to shield physical differences between smart assets, smart gateways, and smart systems.

3.4.4 Development Service Framework (Smart Service)

Figure 3-6 Function view: development service framework



The edge computing model library and vertical industry model library are integrated through the integrated development platform and tool chain to provide full-lifecycle services for the development, integration, simulation, verification, and release of models and apps.

The development service framework supports the following key services:

a.Model-based development service

- Support model definitions including the architecture, function requirements, and interface requirements.
- Support visualized display of the model and service processes.
- Support generation of multi-language codes based on models.
- Support integration and mapping of the edge computing domain models and vertical industry domain models.
- Support model library version management.

b.Emulation service

- Support software and hardware simulation of ECNs, allowing the specifications such as memory and storage space of ECNs to be simulated in the target application scenarios. The system needs to support fine-grained componentization and component tailoring and re-packaging (system reset) to match ECN specifications.
- Based on simulation nodes, build networks and systems based on application scenarios, and perform low-cost and automated function verification on the developed models and apps in a simulation environment.

c. Integrated release service

- Obtain the release version from the baseline library, invoke the deployment operation service, and deploy the models and apps to real ECNs.

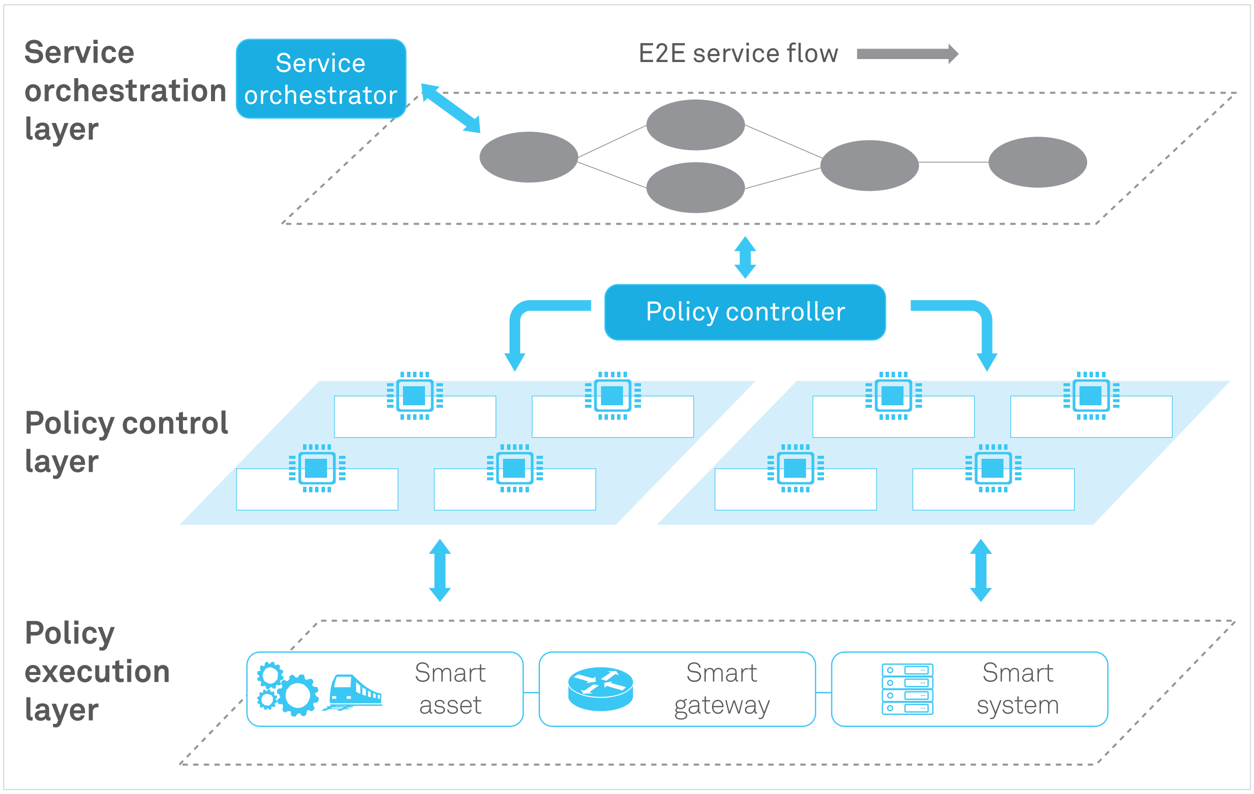
3.4.5 Deployment Operation Service Framework (Smart Service)

The framework includes the following key services: service orchestration, app deployment (not described in this document), and app market.

1) Service orchestration service

Generally, the service orchestration service is based on a three-layer architecture.

Figure 3-7 Function view: service orchestration layer



- **Service orchestrator**

  The orchestrator defines service fabrics, which are generally deployed on the cloud (public or private cloud) or local systems (smart systems). The orchestrator provides a visualized workflow definition tool and supports Create/Retrieve/Update/Delete (CRUD) operations. The orchestrator can orchestrate services based on the service templates and policy templates defined by the development service framework and reuse the templates. Before delivering a service fabric to the policy controller, the orchestrator can perform semantic check and policy conflict detection on the workflow.

- **Policy controller**

  To ensure real-time service scheduling and control, the policy controller is deployed at the network edge to implement local control.

  Based on certain policies and the services and capabilities supported by the local CCFs, the policy controller allocates the service flows defined by a service fabric to a local CCF for scheduling and execution.

  Because the edge computing domain and vertical industry domain require different domain knowledge and system implementation, the controller design and deployment are generally implemented by domain. The edge computing domain controller is responsible for deploying edge computing services such as security and data analysis. If the vertical industry service logic is involved, the vertical industry domain controller is responsible for distribution and scheduling.

- **Policy executor**

  Each ECN has a built-in policy executor, which is responsible for translating policies into local device commands and performing local scheduling and execution. ECNs can either passively receive policies delivered by the controller or proactively request policies from the controller.

  The policies need to focus only on high-level service requirements without implementing fine-grained control on ECNs. This ensures the autonomy of ECNs and real-time responses to local events.

2) App market service

The app market service connects consumers and suppliers, and can transform the unilateral innovation of an enterprise into the multilateral open innovation in an industry ecosystem. Suppliers can encapsulate the industry know-how into apps and register them with the app market for quick publishing. Consumers can easily find an app matching their requirements from the app catalog and subscribe to the app.

The app market service supports a wide range of apps, including the mechanism models constructed based on industrial knowledge, algorithm models constructed based on data analysis methods, service fabric models that can be inherited and reused, and apps that support specific functions (such as fault diagnosis). These apps can be directly used by end users or used for secondary development through model-based open interfaces.

3.4.6 Management Service

- Support the unified management service oriented to terminals, network devices, servers, storage, data, isolation between services and apps, security, and distributed architecture.
- Support the full-lifecycle management of engineering design, integration design, system deployment, service and data migration, integration testing, and integration verification and acceptance.

3.4.7 Full-Lifecycle Data Service

1) Edge data characteristics

Edge data is generated at the network edge and includes machine running data, environment data, and information system data. It features high throughput (large transient traffic), fast flow movement, diversity, strong correlation, and high requirements on real-time analysis and processing.

Compared with business big data scenarios such as Internet, smart analysis of edge data has the following characteristics and differences:

- **Causal relationship vs. association relationship**

  Edge data is mainly targeted at smart assets that generally run with explicit input and output causal relationships, whereas business big data focuses on data association relationships.

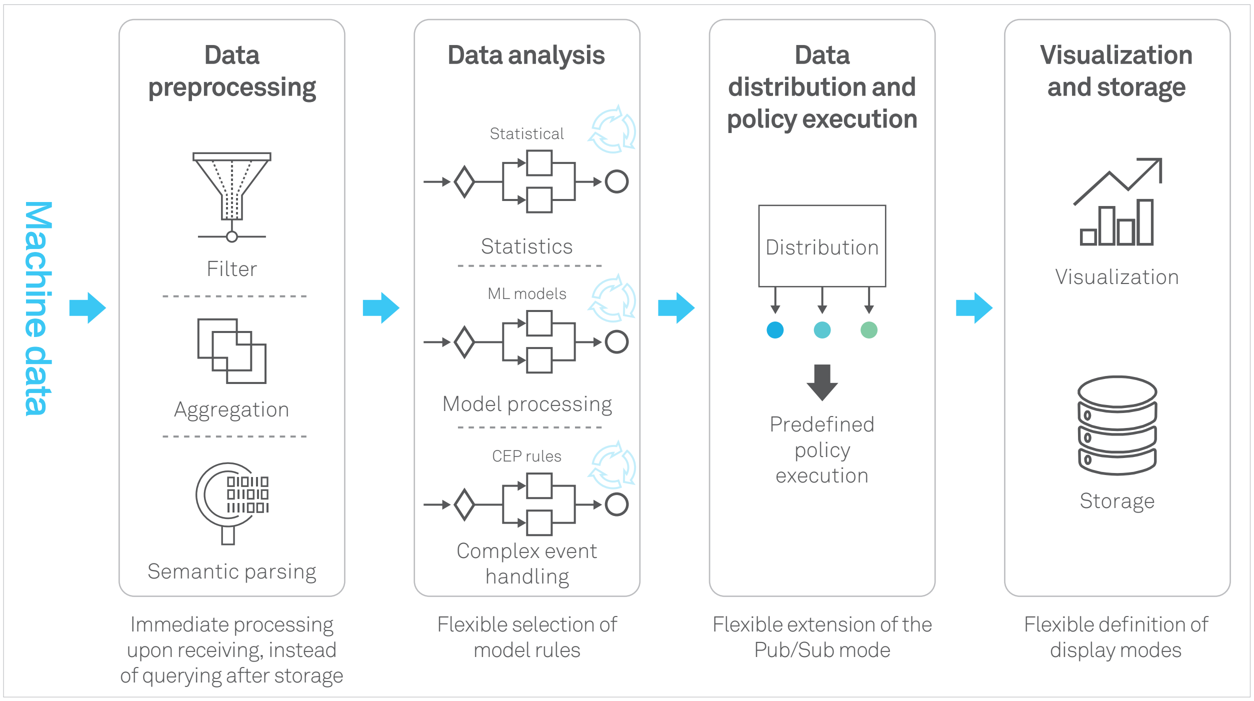
- **High reliability vs. low reliability**

  Industries such as manufacturing and transportation have high requirements on the accuracy and reliability of models. If the accuracy or reliability is low, property loss or even personal injury may occur. In contrast, business big data analysis generally has low requirements on reliability. It is required that the edge data analysis result be explainable. Therefore, black-box deep learning is restricted in some application scenarios. The combination of the traditional mechanism models and data analysis methods is the innovation and application direction of smart analysis.

- **Small data vs. big data**

  Assets such as machine tools and vehicles are designed and manufactured by people. Most data in their running process is predictable, and only data generated in abnormal or critical conditions is truly valuable. On the other hand, business big data analysis requires mass data.

Figure 3-8 Function view: full-lifecycle data service



2) Full-lifecycle data service

Service fabrics are used to define the full-lifecycle data service logic, including specifying the data analysis algorithms. CCFs are used to optimize the data service deployment and running to meet real-time service requirements.

The full-lifecycle data service includes the following:

- **Data preprocessing**

  Filter, clean, aggregate, and optimize (including dirty data elimination) raw data and perform semantic parsing.

- **Data analysis**

  - Based on streaming data analysis, process data in real time so that events and ever-changing service conditions and requirements can be responded to quickly, accelerating the continuous analysis of data.
  - Provide the common statistical model library, and support the algorithm integration of models such as statistical models and mechanism models.
  - Support model training methods such as lightweight deep learning.

- **Data distribution and policy execution**

  Execute policies locally based on predefined rules and data analysis results, or forward data to the cloud or other ECNs for processing.

- **Data visualization and storage**

  Using technologies such as TSDB can significantly conserve storage space and meet the requirements on high-speed read and write operations. Next-generation interaction technologies such as AR and VR are used to provide a vivid display effect.

3.4.8 Security Service

The security design and implementation of the edge computing architecture are first expected to provide the following features:

- Security functions adapt to the specific architecture of edge computing.
- Security functions can be flexibly deployed and expanded.
- The system can continuously mitigate attacks within a certain period of time.
- The system can tolerate function failures to a certain extent and within a specified range, while basic functions run properly.
- The entire system can quickly recover from failure.

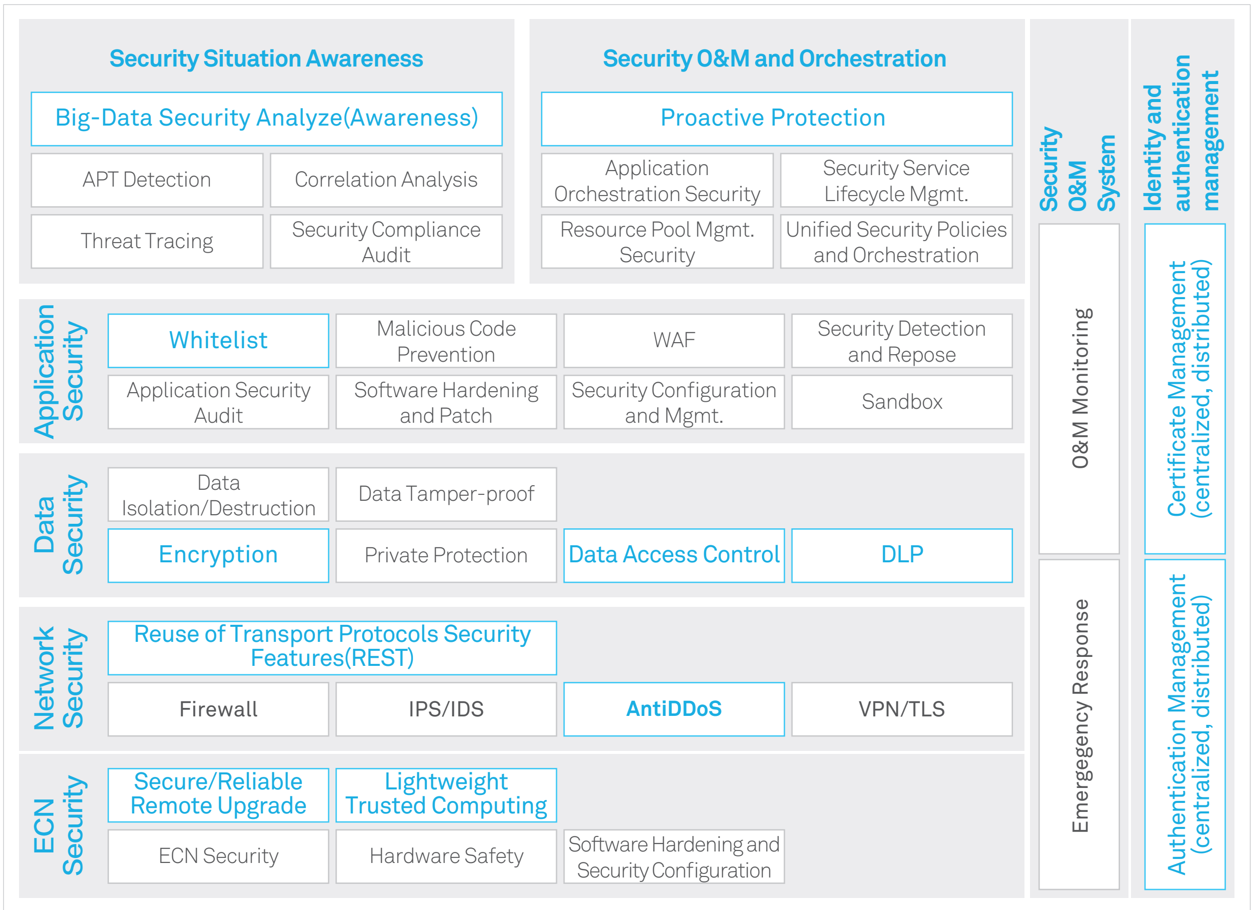
In addition, the security design and implementation need to take the following unique features of edge computing scenarios into consideration:

- Lightweight security functions can be deployed on IoT devices with limited hardware resources.
- The traditional trust-based security model is no longer applicable to the access of a large number of heterogeneous devices. Therefore, the security model (such as the whitelist function) needs to be re-designed based on the minimum authorization principle.
- Isolation between networks and domains is implemented on key nodes (such as smart gateways) to control the scope of security attacks and risks, preventing attacks on a node from spreading to the entire network.

- The security and real-time situation awareness functions are seamlessly embedded into the entire edge computing architecture to achieve continuous detection and response. Automation should be implemented as much as possible, but manual intervention is also required at times.

Security design must cover each layer of the edge computing architecture, and different layers require different security features. In addition, unified situation awareness, security management and orchestration, identity authentication and management, and security O&M are required to ensure maximum security and reliability of the entire architecture.

Figure 3-9 Function view: security service



ECN security: Includes basic ECN security, endpoint security, software hardening and security configuration, secure and reliable remote upgrade, lightweight trusted computing, and hardware safety switch. Secure and reliable remote upgrade can fix vulnerabilities and install patches in time, and prevent system failures after the upgrade. Lightweight trusted computing is applicable to simple IoT devices with limited computing (CPU) and storage resources to solve basic trust problems.

Network (fabric) security: Includes firewalls, IPS/IDS, anti-DDoS, VPN/TLS, and the reuse of security functions of some transport protocols, such as the REST protocol. Anti-DDoS is particularly important in IoT and edge computing. In recent years, a growing number of attacks on IoT devices are DDoS attacks. Attackers control IoT devices with poor security (such as cameras with fixed passwords) to attack specific targets.

Data security: Includes data encryption, data isolation and destruction, data anti-tampering, privacy protection (data anonymization), data access control, and data leakage prevention. Data encryption includes encryption during data transmission and storage. Data leakage prevention for edge computing is different from that of traditional systems because edge computing devices are usually deployed in distributed mode. Special considerations are required to ensure that data will not be leaked even if these devices are stolen.

Application security: Includes security functions such as whitelist, application security audit, malicious code prevention, web application firewall (WAF), and sandbox. The whitelist function is important in the edge computing architecture. The traditional trust-based security model is no longer applicable to the access of a large number of heterogeneous terminals and various services. Therefore, security models (such as the whitelist function) with minimal authorization are used to manage applications and access rights.

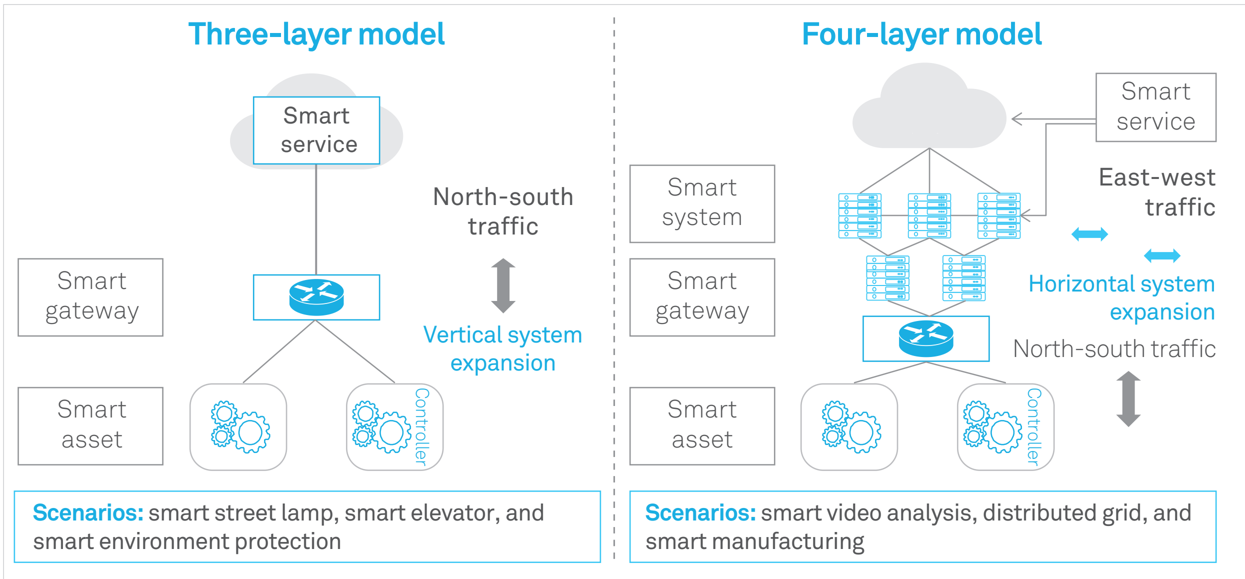
Security situation awareness and security management and orchestration: Since a large number of diversified terminals are connected at the network edge and services carried on the network are complex, passive security defense is ineffective. Therefore, more proactive security defense methods are required, including Big Data–based situation awareness, advanced threat detection, unified network-wide security policy execution, and proactive protection. These can facilitate quick responses. In combination with comprehensive O&M monitoring and emergency response mechanisms, maximum security, availability, and reliability of the edge computing architecture can be ensured.

Identity and authentication management: Covers all function layers. However, accessing a large number of devices at the network edge places much pressure on the performance of the traditional centralized security authentication system, especially when many devices go online within a short period of time. Therefore, the decentralized and distributed authentication and certificate management can be used if needed.

3.5 Deployment View

The edge computing architecture provides the following typical deployment models: three-layer model and four-layer model.

Figure 3-10 Deployment view



1) Three-layer deployment model

This model is applicable to scenarios where services are deployed in one or more scattered areas, each with a low traffic volume.

Typical scenarios include smart street lamps, smart elevators, and smart environmental protection.

After local processing of smart assets, multiple types of or multiple service data flows are aggregated on the smart gateways along the north-south direction. In addition to network functions such as supporting the access and local management of smart assets and bus protocol conversion, the smart gateways provide real-time streaming data analysis, security protection, and small-scale data storage. The gateways process real-time service requirements locally, and aggregate and send non-real-time data to the cloud for processing.

2) Four-layer deployment model

This model is applicable to scenarios where services are deployed centrally and the traffic volume is high.

Typical scenarios include smart video analysis, distributed grid, and smart manufacturing.

The typical differences between four-layer deployment and three-layer deployment are as follows: At the edge, there is a large amount of data and many local application systems are deployed. Therefore, a large amount of computing and storage resources are required. After the most real-time data processing is completed on smart assets and smart gateways, data is aggregated on local distributed smart systems for secondary processing. The distributed ECNs exchange data and knowledge through east-west connections, support horizontal elastic expansion of computing and storage resources, and implement real-time decision-making and optimization operations locally.

04

ECC Industry Development and Business Practice

ECC Industry Development and Business Practice

4.1 ECC Industry Development Overview

The industrial value, reference architecture, and business practices of edge computing are highly recognized by industry partners. The Edge Computing Consortium (ECC) is the largest consortium in the industry to focus on the edge computing field. The ECC was established in 2016 and now has more than 150 members, including Huawei, Intel, ARM, Bosch, Honeywell, ABB, Schneider, Schindler, Infosys, Mitsubishi, HollySys, McAfee, 360, NI, and OSIsoft.

In order to promote vigorous development and quick adoption of edge computing in industries, the ECC has established contact and partnerships with many industry organizations, including the Industrial Internet Consortium (IIC), Alliance of Industrial Internet (AII), Chinese Association of Automation (CAA), SDNFV Industry Alliance, Avnu Alliance, International SSL Alliance (ISA), and Telematics Industry Application Alliance (TIAA).

4.1.1 Cooperation Between the ECC and Industry Organizations

- In April 2017, the ECC and SDNFV Industry Alliance signed an agreement that said that the two organizations will strengthen cooperation in standards and test beds to promote industry development.
- In June 2017, the ECC and IIC established an intergroup collaboration by signing a memorandum of understanding (MoU) and reached agreements on edge computing technologies, test beds, and marketing promotion. Huawei, an ECC member, promoted the establishment of the Networking Task Group (TG) of the IIC, and is the co-chair of the Edge Computing TG.
- In August 2017, the ECC and CAA signed a cooperation agreement and set up the Technical Committee on Edge Computing of the CAA to accelerate open collaboration of OT and ICT industries and incubate industry application best practices.

- In November 2017, the ECC and Avnu Alliance signed a MoU and agreed to identify and share best practices of Industrial Internet of Things (IIoT), strengthen their cooperation in standards, test beds, and R&D projects, and realize interoperability by harmonizing the architecture and other elements.
- In November 2017, the ECC will sign an agreement with the AII to strengthen cooperation in architecture, standards, test beds, and marketing promotion.

4.1.2 Cooperation Between the ECC and Standardization Organizations

- In March 2017, owing to the ECC's efforts, edge computing became an important part of IEEE P2413 (Standard for an Architectural Framework for the Internet of Things).
- In May 2017, the ECC promoted and set up an edge computing research group, with Huawei and China Electronics Standardization Institute (CESI) as the group chair, at IEC/ISO JTC1 SC41 to include edge computing in standards.
- In July 2017, the ECC collaborates with the CESI to promote the application of edge computing architecture and technology in the Made in China 2025 strategy.
- In September 2017, Huawei and Fraunhofer FOKUS led the development and release of the Vertical Edge Intelligence White Paper for the IEC.
- In September 2017, the ECC and ISA set up the ISA-ECC Smart Street Lighting System Committee to jointly formulate smart lighting technology standards and use edge computing to develop innovative solutions.
- In November 2017, the ECC and TIAA set up a joint working group to promote the wide-ranging application of edge computing technologies and solutions on commercial vehicles and special-purpose motor vehicles.

4.2 Business Practices of Edge Computing

4.2.1 Theory and Practice of Edge Computing

From its proposal to its implementation, edge computing has developed rapidly. The edge computing architecture and the following core concepts have already been implemented in projects:

- Heterogeneous computing hardware platforms: x86, ARM, and FPGA
- Distributed open software platforms: SDN, streaming data analysis, container/VM, application lifecycle management and security policy and mechanism
- Experiential development and test cloud: development and test environment, tool chain, simulation environment, developer community, and code library

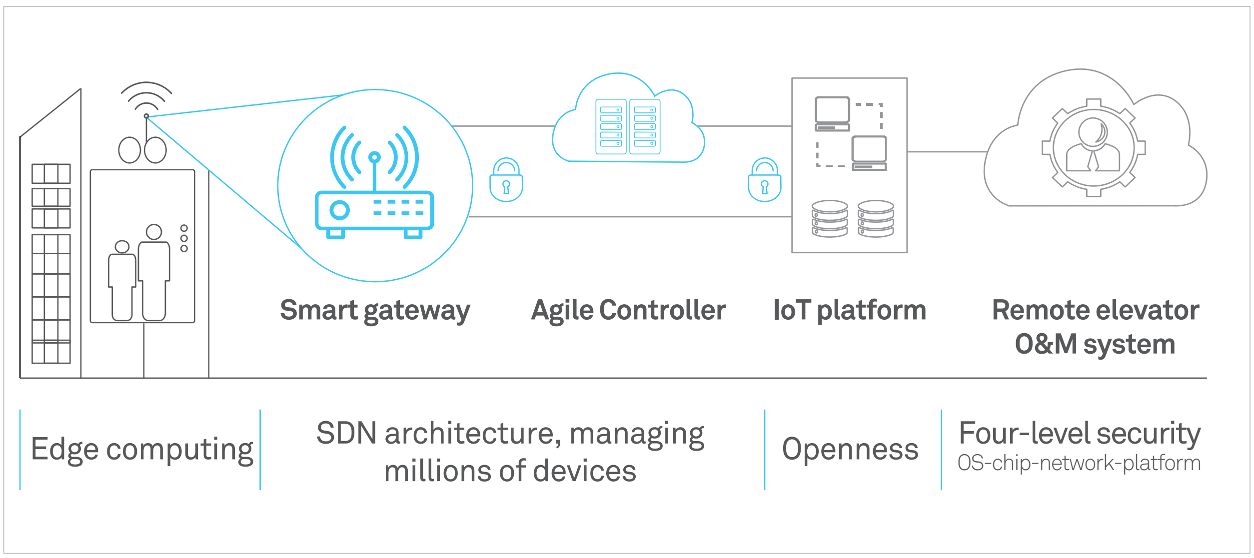
4.2.2 Implementation of Horizontal Solutions in Vertical Industries

By integrating industry usage scenarios and corresponding applications, edge computing implementation has shifted from horizontal solution platforms to vertical industries based on the characteristics and requirements of different industries, building many innovative vertical industry solutions in different industries.

1) Elevators Connection

Currently, more than 15 million elevators are in use around the world and more than a billion people use elevators every day. Based on edge computing, gateways deployed at the elevators IoT edge collect elevators' operational data in real time and preprocess the data. The elevators connection solution coordinates with the cloud-based Big Data analytics platform to monitor the health status of elevator components. In addition, predictive maintenance of elevators helps to make maintenance more efficient and reduces OPEX. Value-added services (VASs) enable elevator manufacturers to transform into service providers and explore new markets for growth.

Figure 4-1 Elevators connection solution architecture



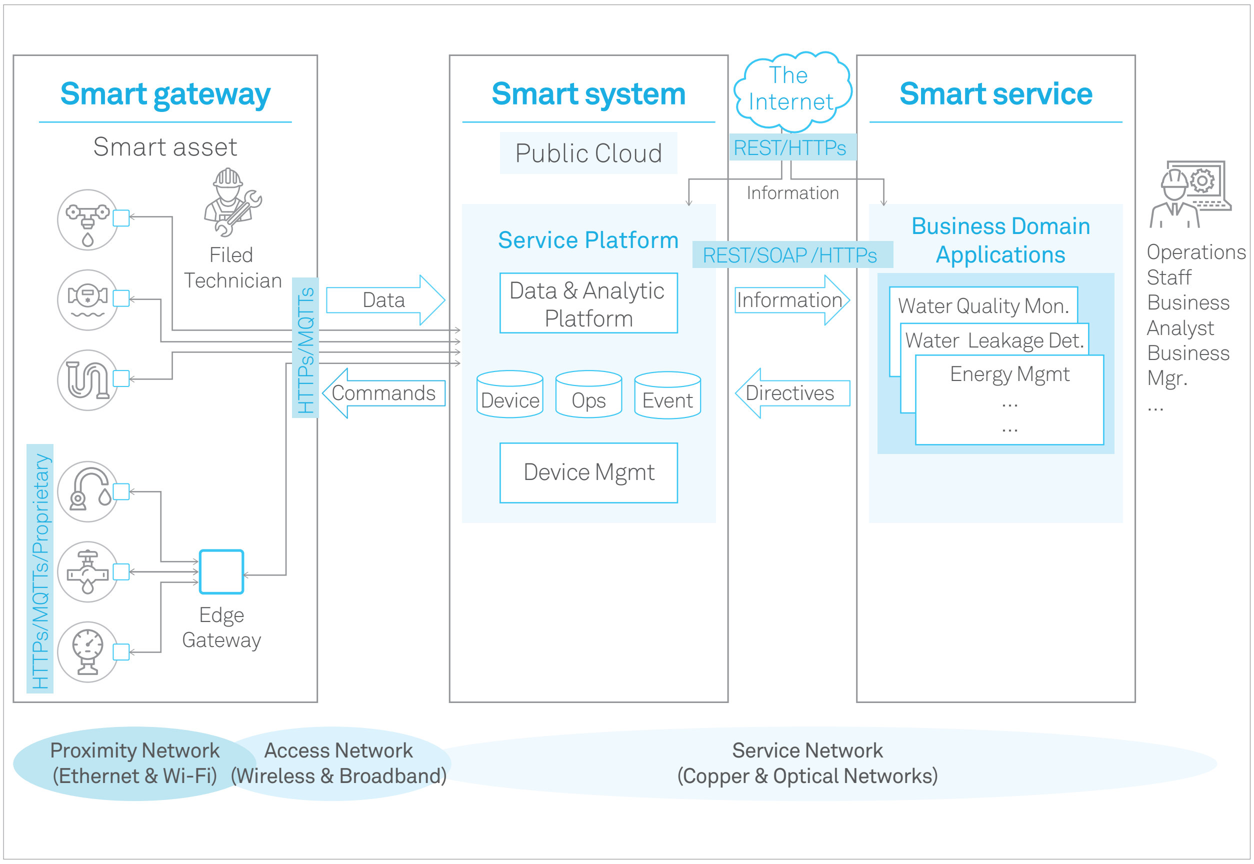
Core values of edge computing in the elevators connection solution:

- Based on software-defined networking (SDN), the solution can manage millions of elevators and supports network visualization and automatic O&M.
- Smart gateways that have an energy-saving and industry-grade design provide network access for elevators in various complicated scenarios.
- A container can have multiple elevator monitoring apps installed and supports lifecycle management of the apps.
- Data analysis modules can locally analyze elevators' operational data and respond to requests in real time.
- The following four-level security is provided: Dedicated TPM chips ensure secure encryption, the operating system has a built-in security module, IPSec encrypted channels are used, and the SDN controller provides security solutions.

2) Smart Water

A water supply system is a key urban infrastructure, and is an important part of the construction of safe cities and smart cities. Based on edge computing platforms, the smart water solution uses advanced sensing, networking, computing, control, and intelligence technologies to monitor equipment such as secondary water supply facilities. The urban water supply equipment, information system, and service process are integrated to support the exchange of a large amount of data between multiple systems on a large scale. This enables full-process control, automatic fault diagnosis, and predictive maintenance, reducing energy consumption and ensuring water security.

Figure 4-2 Smart water solution architecture



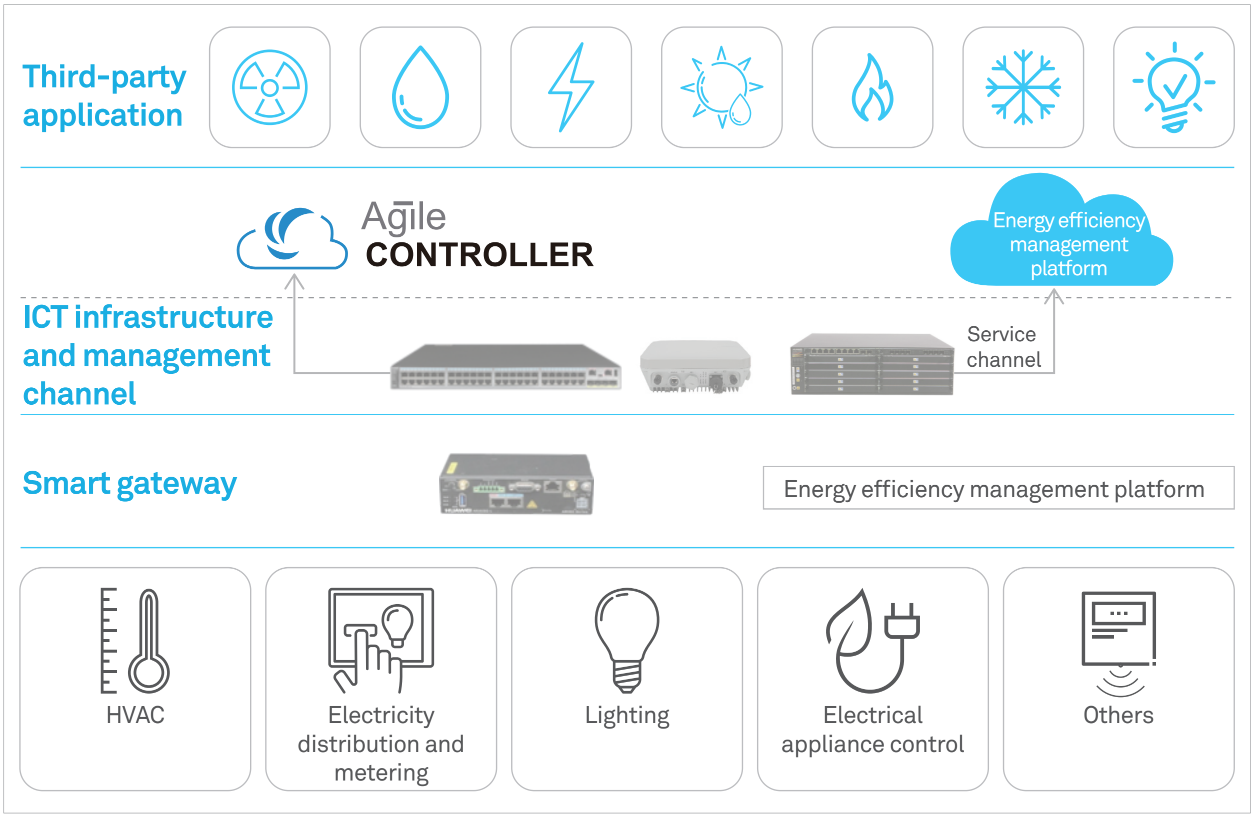
Core values of edge computing in the smart water solution:

- Smart gateways collaborate with cloud-based platforms to receive data analysis models, so that they can analyze data and respond to requests based on policies in real time.
- Edge devices are deployed in distributed mode and can be expanded on demand.
- The solution supports lifecycle management of multiple apps for flexible addition of new services.
- Hardware platforms have an industry-grade design and multiple interfaces that comply with various protocols.

3) Smart Building

The smart building solution uses advanced sensing, networking, computing, control, intelligence, and security technologies to monitor device status, detect faults, and implement predictive maintenance by collecting, monitoring, and analyzing data from sensors in buildings. Multiple systems are orchestrated to provide automatic control and visualized O&M, in order to create intelligent, eco-friendly, and efficient office and living environments for people.

Figure 4-3 Smart building solution architecture



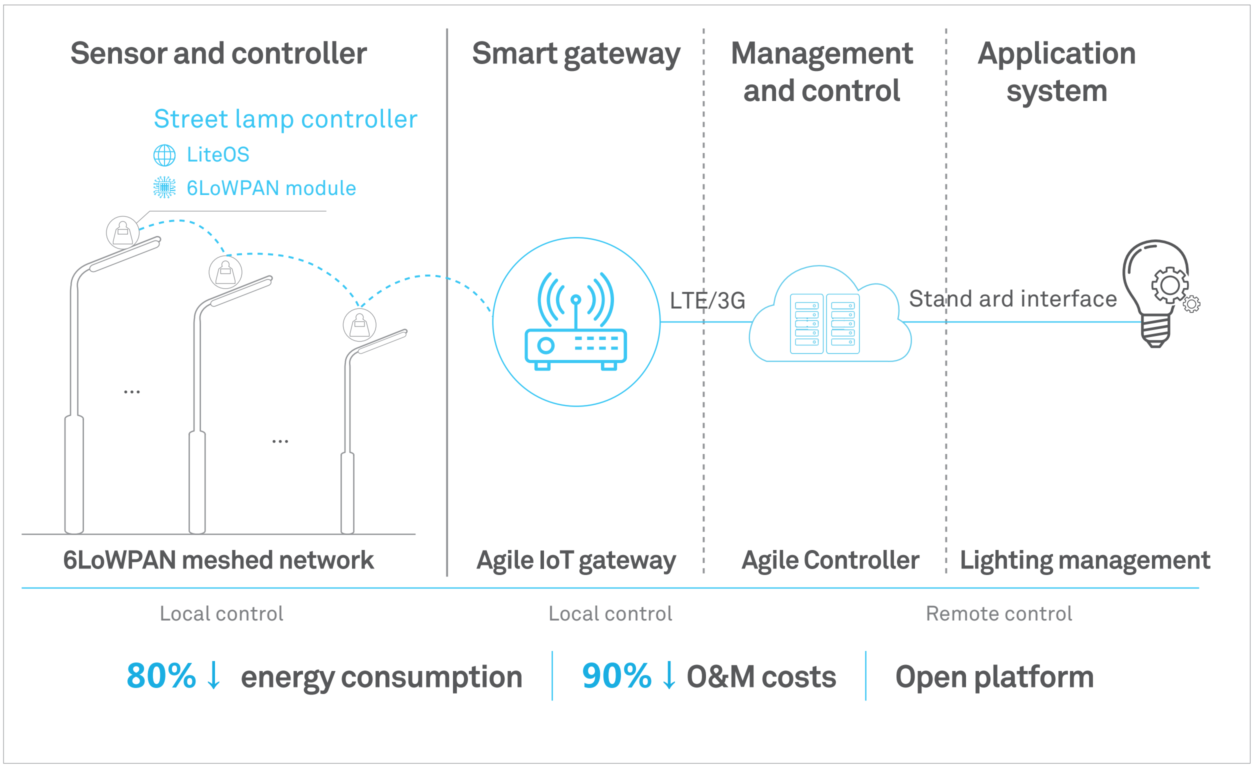
Core values of edge computing in the smart building solution:

- Smart gateways provide various interfaces and features to simplify network connections, and are applicable to multiple building access scenarios.
- Platforms support real-time responses and more than 1000 industry protocols for seamless connections with a variety of devices.
- Gateways, the Agile Controller, and platforms are open, simplifying application development and improving the efficiency.
- Security at multiple levels is provided: Security components ensure container security, VPN channels ensure data transmission security, and platform authorization ensures system security.

4) Smart Lighting

The smart lighting solution collects all useful data from street lamps to provide remote, real-time, and self-adaptive street lamp control, automatic fault diagnosis, and predictive maintenance to connect terminals such as trash can monitors and parking sensors deployed near street lamp poles — which can be considered as wireless network access nodes — providing the basis for building an extendable city IoT framework and promoting evolution of the lighting IoT into city IoT.

Figure 4-4 Smart lighting solution architecture



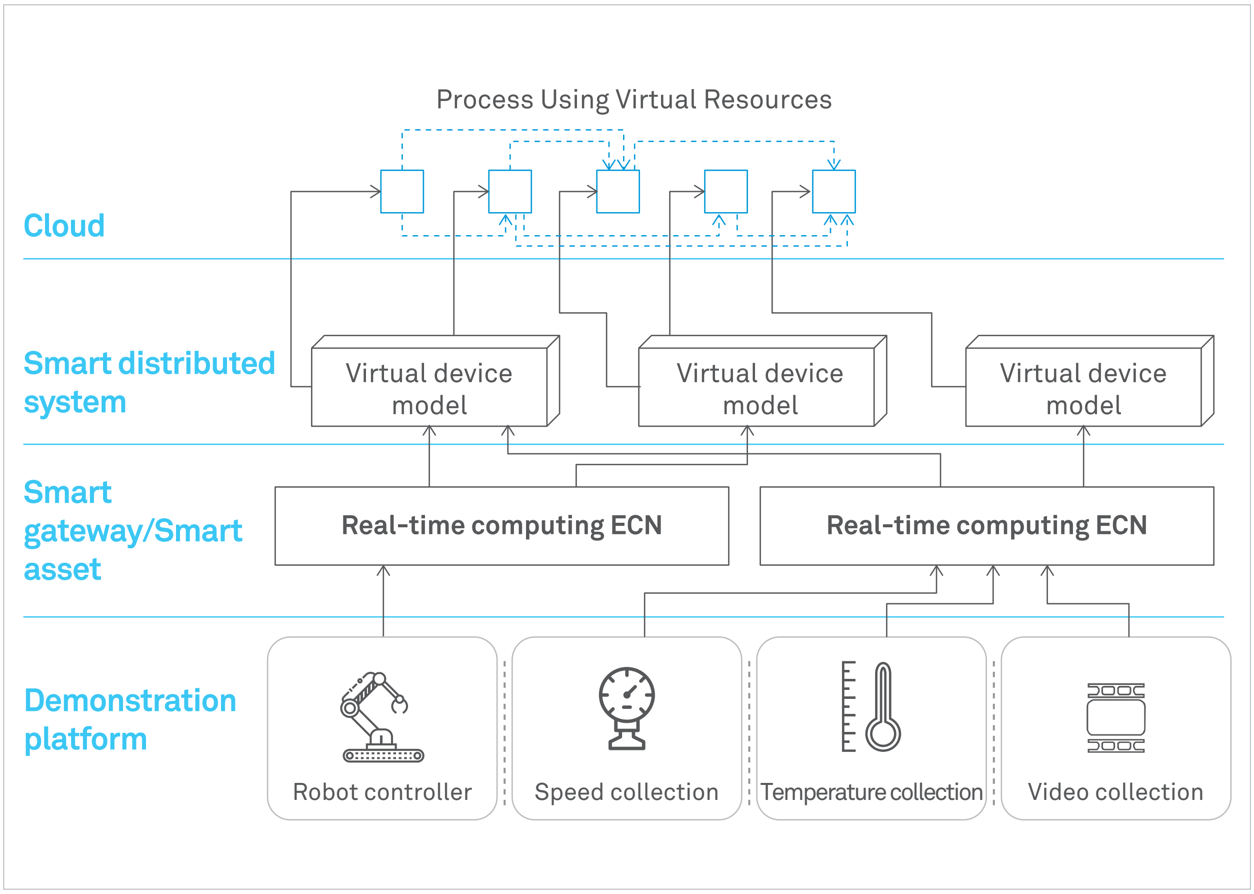
Core values of edge computing in the smart lighting solution:

- Smart gateways and street lamp controllers support multi-level offline control, ensuring system reliability.
- Edge computing gateways support intelligent multi-dimensional analysis for real-time street lamp control, saving energy and improving user experience.
- Data is reliably transmitted over 6LoWPAN-based meshed wireless communication networks in real time.
- Visualized management based on the Geographic Information System (GIS) and active maintenance help to reduce street lamp maintenance and management costs.

5) Virtualization and Service Orchestration Application in Smart Manufacturing Systems

Virtualization and service orchestration in smart manufacturing systems enable resources such as production equipment to be connected to smart distributed systems through smart gateways or smart assets. This achieves virtualization and modeling of production equipment in the digital world. Network resources, production equipment, and production process can be intelligently orchestrated through collaboration of edge and cloud computing, enabling self-perception, self-decision-making, self-implementation, and predictive maintenance in the manufacturing process.

Figure 4-5 Virtualization and service orchestration framework in smart manufacturing systems



Core values of edge computing for virtualization and service orchestration in smart manufacturing systems:

- Sets up an independent, restructurable digital device model (Digital Twins) through real-time connection and perception of edge devices, and enables virtualization, modeling, correlation, and retrieval of production resources.

- Achieves adaptive allocation of network resources using the SDN technology, and provides effective means of information transmission for restructurable devices.

- Uses service fabrics to define the tasks, technological processes, path plans, and control parameters of the processing and assembly phases, implementing fast deployment of service policies and fast processing of multiple types of products.

4.2.3 Requirements and Practices of Edge Computing

During the development of edge computing, key requirements and technologies, reference architecture and platforms, and test beds and business cases interact with, promote, and iterate each other. This continuously improves the competitiveness and adaptability of edge computing.

- Refining scenario-based requirements and identifying key technologies facilitate the construction of the edge computing reference architecture. Industry cooperation is more effective based on unified architecture and language.

- The edge computing reference architecture supports open edge computing software and hardware platforms that allow deployment of test beds and business cases. Best practices accelerate the industry development.

- Test beds and business cases can be used to test key technologies, reference architecture, and software and hardware platforms. The test results are useful in problem identification and subsequent requirement optimization for business application improvement.

05
Appendixes

Glossary

No.	Term	Definition
1	Cloud Computing	Cloud computing is generally referred to as the cloud. It provides on-demand computing resources (from applications to data centers) in pay-per-use mode based on the Internet. Deployment modes include public cloud, private cloud, and hybrid cloud.
2	Edge Computing	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 smart services. It meets the key requirements of industry digitalization in terms of agile connection, real-time service, data optimization, application intelligence, and security and privacy protection.
3	Smart Asset	Smart assets are assets or things that can implement autonomy and collaboration through integration with ICT resources such as network, computing, and storage.
4	Smart Gateway	Smart gateways connect the physical and digital worlds through functions such as network connection and protocol conversion, and provide lightweight connection management, real-time data analysis, and application management functions.
5	Smart System	Smart systems are constructed based on the collaboration of multiple distributed smart gateways and servers to support elastic expansion of network, computing, and storage resources.
6	Smart Service	Smart services are based on a model-driven unified service framework and provide the development service framework and deployment operation service framework for multiple roles such as system O&M personnel, service decision-makers, system integrators, and application development personnel.

7	Software-Defined Networking	Software-defined networking (SDN) is a newer network architecture. It separates the network control plane from the forwarding plane, uses centralized control to replace the original distributed control, and implements software definition through open and programmable interfaces.
8	Time-Sensitive Networking	The IEEE formulates a set of time-sensitive networking (TSN) standards to define unified technical standards for key services such as real-time priorities and clocks. TSN is the future development trend of industrial Ethernet connections.
9	Heterogeneous Computing	Heterogeneous computing is a newer computing architecture that combines different types of instruction sets and computing units of different architectures to fully utilize the advantages of various computing units and balance the performance, cost, power consumption, and portability.
10	Time Series Database	A time series database (TSDB) is a database for storing time series data (including the time stamp). It also supports basic functions such as fast writing, persistency, and multi-dimensional aggregation queries of time series data.
11	Interoperability	Interoperability means that two or more systems exchange information. They can understand the meaning of each other's information and collaborate with each other.
12	Information model	An information model is a representation of concepts, relationships, constraints, and operations to specify data semantics, so that information can be shared and organized.

Acronyms and Abbreviations

No.	Acronym and Abbreviation	Full Name
1	AI	Artificial Intelligence
2	AR	Augmented Reality
3	ASIC	Application-Specific Integrated Circuit
4	CCF	Connectivity and Computing Fabric
5	CSMA/CD	Carrier Sense Multiple Access with Collision Detection
6	CNC	Computer Numerical Control
7	CPS	Cyber-Physical System
8	CRUD	Create/Retrieve/Update/Delete
9	CT	Communication Technology
10	DCS	Distributed Control System
11	DDoS	Distributed Denial of Service
12	DDS	Data Distribution Service
13	DIKW	Data Information Knowledge Wise
14	ECC	Edge Computing Consortium
15	ECN	Edge Computing Node
16	ERP	Enterprise Resource Planning
17	EVF	Edge Virtualization Function
18	FPGA	Field Programmable Gate Array
19	HC	Heterogeneous Computing
20	HDD	Hard Disk Drive
21	HMI	Human-Machine Interface
22	ICT	Information and Communications Technology
23	II	Industry Intelligence
24	IIC	Industrial Internet Consortium

25	IPS/IDS	Intrusion Prevention System/Intrusion Detection System
26	IT	Information Technology
27	MAC	Media Access Control
28	MDE	Model-Driven Engineering
29	MES	Manufacturing Execution System
30	OICT	Operation Technology & Information Communications Technology
31	OPC-UA	OPC Unified Architecture
32	OS	Operating System
33	OT	Operation Technology
34	PC	Policy Controller
35	PE	Policy Executor
36	PLC	Programmable Logic Controller
37	QoS	Quality of Service
38	REST	Representational State Transfer
39	SDN	Software-Defined Networking
40	SF	Service Fabric
41	SLA	Service-Level Agreement
42	SO	Service Orchestrator
43	SSD	Solid-State Drive
44	TLS	Transport Layer Security Protocol
45	TSDB	Time Series Database
46	TSN	Time-Sensitive Networking
47	vPLC	Virtual Programmable Logic Controller
48	VPN	Virtual Private Network
49	VR	Virtual Reality

References

01. Margaret Chiosi, AT&T, Don Clarke, Peter Willis, Andy Reid, BT, Dr. Chunfeng Cui, Dr. Hui Deng, China Mobile, etc. "Network Functions Virtualization — Introductory White Paper"

02. [Weisong Shi, Fellow, IEEE, Jie Cao, Student Member, IEEE, Quan Zhang, Student Member, IEEE, Youhuizi Li, and Lanyu Xu, "Edge Computing: Vision and Challenges"

03. IEC Vertical Edge Intelligence Whitepaper
http://www.iec.ch/whitepaper/edgeintelligence

04. Article series about the Time Series Database (TSDB)
http://www.infoq.com/cn/articles/hierarchical-storage-of-sequential-databases?utm_campaign=rightbar_v2&utm_source=infoq&utm_medium=articles_link&utm_content=link_text

05. IEC 62443-3-3:2013, Industrial communication networks — Network and system security — Part 3-3: System security requirements and security levels

06. OPC Unified Architecture Specification — Part 1: Overview and Concepts. OPC Foundation

07. IEEE Time-Sensitive Networking Task Group
http://www.ieee802.org/1/pages/tsn.html

08. Software-Defined Networking (SDN) Definition
https://www.opennetworking.org/sdn-definition/

09. Open Container Initiative
https://www.opencontainers.org/about.html

10. Vertical Edge Intelligence white paper
http://www.iec.ch/whitepaper/pdf/IEC_WP_Edge_Intelligence.pdf

11. Model-Driven Engineering
https://en.wikipedia.org/wiki/Model-driven_engineering

12. Reference Model for Service Oriented Architecture 1.0
http://docs.oasis-open.org/soa-rm/v1.0/soa-rm.pdf

Scan the QR code to follow the ECC

Edge Computing Consortium (ECC)

Address	Room 1416, No. 5 Building, Splendid International Plaza, Shangdi 10th street, Haidian District, Beijing 100085, China
Website	http://www.ecconsortium.org
Email	info@eccconsortium.net
Phone number	010-57116299

Alliance of Industrial Internet (AII)

Address	No. 52, Hua Yuan North Road, Haidian District, Beijing 100191, China
Website	http://www.aii-alliance.org/
Email	aii@caict.ac.cn
Phone number	010-62305887