International Telecommunication Union

ITU-T

TELECOMMUNICATION
STANDARDIZATION SECTOR
OF ITU

Y.3526

(11/2021)

SERIES Y: GLOBAL INFORMATION INFRASTRUCTURE, INTERNET PROTOCOL ASPECTS, NEXT-GENERATION NETWORKS, INTERNET OF THINGS AND SMART CITIES

Cloud Computing

Cloud computing – Functional requirements of edge cloud management

Recommendation ITU-T Y.3526

ITU-T Y-SERIES RECOMMENDATIONS

GLOBAL INFORMATION INFRASTRUCTURE, INTERNET PROTOCOL ASPECTS, NEXT-GENERATION NETWORKS, INTERNET OF THINGS AND SMART CITIES

GLOBAL INFORMATION INFRASTRUCTURE	
General	Y.100–Y.199
Services, applications and middleware	Y.200–Y.299
Network aspects	Y.300–Y.399
Interfaces and protocols	Y.400–Y.499
Numbering, addressing and naming	Y.500–Y.599
Operation, administration and maintenance	Y.600–Y.699
Security	Y.700–Y.799
Performances	Y.800–Y.899
INTERNET PROTOCOL ASPECTS	
General	Y.1000–Y.1099
Services and applications	Y.1100–Y.1199
Architecture, access, network capabilities and resource management	Y.1200–Y.1299
Transport	Y.1300–Y.1399
Interworking	Y.1400–Y.1499
Quality of service and network performance	Y.1500–Y.1599
Signalling	Y.1600–Y.1699
Operation, administration and maintenance	Y.1700–Y.1799
Charging	Y.1800–Y.1899
IPTV over NGN	Y.1900–Y.1999
NEXT GENERATION NETWORKS	
Frameworks and functional architecture models	Y.2000–Y.2099
Quality of Service and performance	Y.2100–Y.2199
Service aspects: Service capabilities and service architecture	Y.2200–Y.2249
Service aspects: Interoperability of services and networks in NGN	Y.2250–Y.2299
Enhancements to NGN	Y.2300–Y.2399
Network management	Y.2400–Y.2499
Computing power networks	Y.2500–Y.2599
Packet-based Networks	Y.2600–Y.2699
Security	Y.2700–Y.2799
Generalized mobility	Y.2800–Y.2899
Carrier grade open environment	Y.2900–Y.2999
FUTURE NETWORKS	Y.3000–Y.3499
CLOUD COMPUTING	**Y.3500–Y.3599**
BIG DATA	Y.3600–Y.3799
QUANTUM KEY DISTRIBUTION NETWORKS	Y.3800–Y.3999
INTERNET OF THINGS AND SMART CITIES AND COMMUNITIES	
General	Y.4000–Y.4049
Definitions and terminologies	Y.4050–Y.4099
Requirements and use cases	Y.4100–Y.4249
Infrastructure, connectivity and networks	Y.4250–Y.4399
Frameworks, architectures and protocols	Y.4400–Y.4549
Services, applications, computation and data processing	Y.4550–Y.4699
Management, control and performance	Y.4700–Y.4799
Identification and security	Y.4800–Y.4899
Evaluation and assessment	Y.4900–Y.4999

For further details, please refer to the list of ITU-T Recommendations.

Recommendation ITU-T Y.3526

Cloud computing – Functional requirements of edge cloud management

Summary

Recommendation ITU-T Y.3526 provides requirements for edge cloud management. It presents an overview of edge cloud management including advantages of edge cloud management and the relationship with global management in a distributed cloud. It describes the edge cloud management local functions and mode. Additionally, this Recommendation provides edge cloud management functional requirements derived from use cases.

History

Edition	Recommendation	Approval	Study Group	Unique ID*
1.0	ITU-T Y.3526	2021-11-06	13	11.1002/1000/14759

Keywords

Distributed cloud, edge cloud, functional requirements, management mode.

* To access the Recommendation, type the URL http://handle.itu.int/ in the address field of your web browser, followed by the Recommendation's unique ID. For example, http://handle.itu.int/11.1002/1000/11830-en.

FOREWORD

The International Telecommunication Union (ITU) is the United Nations specialized agency in the field of telecommunications, information and communication technologies (ICTs). The ITU Telecommunication Standardization Sector (ITU-T) is a permanent organ of ITU. ITU-T is responsible for studying technical, operating and tariff questions and issuing Recommendations on them with a view to standardizing telecommunications on a worldwide basis.

The World Telecommunication Standardization Assembly (WTSA), which meets every four years, establishes the topics for study by the ITU-T study groups which, in turn, produce Recommendations on these topics.

The approval of ITU-T Recommendations is covered by the procedure laid down in WTSA Resolution 1.

In some areas of information technology which fall within ITU-T's purview, the necessary standards are prepared on a collaborative basis with ISO and IEC.

NOTE

In this Recommendation, the expression "Administration" is used for conciseness to indicate both a telecommunication administration and a recognized operating agency.

Compliance with this Recommendation is voluntary. However, the Recommendation may contain certain mandatory provisions (to ensure, e.g., interoperability or applicability) and compliance with the Recommendation is achieved when all of these mandatory provisions are met. The words "shall" or some other obligatory language such as "must" and the negative equivalents are used to express requirements. The use of such words does not suggest that compliance with the Recommendation is required of any party.

INTELLECTUAL PROPERTY RIGHTS

ITU draws attention to the possibility that the practice or implementation of this Recommendation may involve the use of a claimed Intellectual Property Right. ITU takes no position concerning the evidence, validity or applicability of claimed Intellectual Property Rights, whether asserted by ITU members or others outside of the Recommendation development process.

As of the date of approval of this Recommendation, ITU had not received notice of intellectual property, protected by patents/software copyrights, which may be required to implement this Recommendation. However, implementers are cautioned that this may not represent the latest information and are therefore strongly urged to consult the appropriate ITU-T databases available via the ITU-T website at http://www.itu.int/ITU-T/ipr/.

Table of Contents

Recommendation ITU-T Y.3526

Cloud computing – Functional requirements of edge cloud management

1 Scope

This Recommendation provides an overview and requirements of edge cloud management. It addresses the following subjects:

– Overview of edge cloud management
  - Challenges for management of a distributed cloud;
  - Advantages of introducing edge cloud management;
  - Relationships with edge cloud and edge node; and
  - Relationships with global management in a distributed cloud.

– Edge cloud management mode

– Edge cloud local management functions

– Functional requirements of edge cloud management

2 References

The following ITU-T Recommendations and other references contain provisions which, through reference in this text, constitute provisions of this Recommendation. At the time of publication, the editions indicated were valid. All Recommendations and other references are subject to revision; users of this Recommendation are therefore encouraged to investigate the possibility of applying the most recent edition of the Recommendations and other references listed below. A list of the currently valid ITU-T Recommendations is regularly published. The reference to a document within this Recommendation does not give it, as a stand-alone document, the status of a Recommendation.

[ITU-T J.1301] Recommendation ITU-T J.1301 (2021), *Specification of cloud-based converged media service to support Internet protocol and broadcast cable television – Requirements.*

[ITU-T X.1361] Recommendation ITU-T X.1361 (2018), *Security framework for the Internet of things based on the gateway model*.

[ITU-T X.1601] Recommendation ITU-T X.1601 (2015), *Security framework for cloud computing*.

[ITU-T X.1642] Recommendation ITU-T X.1642 (2016), *Guidelines for the operational security of cloud computing*.

[ITU-T Y.3500] Recommendation ITU-T Y.3500 (2014), *Cloud computing – Overview and vocabulary*.

[ITU-T Y.3502] Recommendation ITU-T Y.3502 (2014), *Information technology – Cloud computing – Reference architecture.*

[ITU-T Y.3508] Recommendation ITU-T Y.3508 (2019), *Cloud computing – Overview and high-level requirements of distributed cloud*.

3 Definitions

3.1 Terms defined elsewhere

This Recommendation uses the following terms defined elsewhere:

3.1.1 cloud computing [ITU-T Y.3500]: Paradigm for enabling network access to a scalable and elastic pool of shareable physical or virtual resources with self-service provisioning and administration on-demand.

NOTE – Examples of resources include servers, operating systems, networks, software, applications, and storage equipment.

3.1.2 cloud service [ITU-T Y.3500]: One or more capabilities offered via cloud computing invoked using a defined interface.

3.1.3 cloud service customer [ITU-T Y.3500]: Party which is in a business relationship for the purpose of using cloud services.

NOTE – A business relationship does not necessarily imply financial agreements.

3.1.4 cloud service provider [ITU-T Y.3500]: Party which makes cloud services available.

3.1.5 core cloud [ITU-T Y.3508]: A cloud computing, which manages resource pools including resources in the edge of the network and enables cloud service.

NOTE – Enabled cloud service on the core cloud is provided by a cloud service provider (CSP).

3.1.6 distributed cloud [ITU-T Y.3508]: Distribution of cloud capabilities types to the edge of the network for enabling cloud services with low latency and real time processing on limited bandwidth by interworking among pools of physical or virtual resources.

3.1.7 edge cloud [ITU-T Y.3508]: A cloud computing deployed to the edge of the network accessed by cloud service customers (CSCs) with small capacity resources enabling cloud service.

NOTE 1 – Enabled cloud service on the edge cloud is lightweight cloud service provided by a cloud service provider (CSP) depending on cloud service category.

NOTE 2 – Lightweight cloud service refers to a portion of cloud service to reconfigure the functionality of cloud service to fit on edge cloud such as base station and gateway with small capacity resource.

3.1.8 regional cloud [ITU-T Y.3508]: A cloud computing hosted from core cloud to particular geographical regions.

NOTE – Enabled cloud service on the regional cloud is entire or partial cloud service of core cloud provided by a cloud service provider (CSP).

3.1.9 microservice [ITU-T J.1301]: An architectural and organizational approach to software development where software is composed of small independent microservices that communicate over well-defined application programming interfaces. Microservices architecture makes applications easier to scale and faster to develop, enabling innovation and accelerating time-to-market for new features.

3.2 Terms defined in this Recommendation

This Recommendation defines the following terms:

3.2.1 edge cloud management: An efficient and effective way to provide management of resources, data, services, applications, operation and maintenance and risks of edge cloud.

NOTE 1 – The scope of edge cloud management is within an edge cloud. Each edge cloud has one edge cloud management.

NOTE 2 – Edge cloud management coordinates with global management and other edge cloud managements.

3.2.2 edge node: Resource pool with limited computing, storage, network and other software and tools within one edge cloud.

NOTE 1 – Each edge node belongs to only one edge cloud. Each edge cloud has at least one edge node.

NOTE 2 – Edge node uses edge cloud management to expose its abilities and achieve coordination with other edge nodes.

4 Abbreviations and acronyms

This Recommendation uses the following abbreviations and acronyms:

AI	Artificial Intelligence
API	Application Programming Interface
AR	Augmented Reality
CC	Core Cloud
CDN	Content Delivery Network
CSC	Cloud Service Customer
CSN	Cloud Service Partner
CSP	Cloud Service Provider
EC	Edge Cloud
FCAPS	Fault, Configuration, Accounting, Performance, Security
ICT	Information and Communication Technology
IoT	Internet of Things
ML	Machine Learning
NFV	Network Function Virtualization
NFVO	Network Function Virtualization Orchestrator
RC	Regional Cloud
SLA	Service Level Agreement
VR	Virtual Reality
V2X	Vehicle to Everything

5 Conventions

In this Recommendation:

The keywords "**is required to**" indicate a requirement which must be strictly followed and from which no deviation is permitted if conformance to this document is to be claimed.

The keywords "**is recommended**" indicate a requirement which is recommended but which is not absolutely required. Thus this requirement need not be present to claim conformance.

6 Overview of edge cloud management

6.1 Challenges for management of distributed cloud

Services in Internet of things (IoT), augmented reality/virtual reality (AR/VR), artificial intelligence (AI) and 5G application domains, which require low network latency between a cloud service provider (CSP) and cloud service customer (CSC) and wide bandwidth, need a new cloud computing concept of distributed cloud, which is the distribution of cloud capabilities types to the edge of the network for enabling cloud service with low latency and real-time processing on limited bandwidth by interworking among pool of physical or virtualized resources [ITU-T Y.3508].

According to the concept and structure of distributed cloud, refer to Figure 6-1 of [ITU-T Y.3508], global management in core cloud (CC) provides cloud resource management for distributed cloud, which mainly refers to CC, regional cloud (RC) and edge cloud (EC). However, as 5G and edge computing technologies are widely used on the edge side, the number and type of ECs will increase remarkably. Using global management alone to manage distributed cloud with a large amount of ECs will face the following challenges:

– **Heavy operation and management burden**: With the rapid growth of the number of ECs, global management will handle a huge amount of fault, configuration, accounting, performance and security (FCAPS) data coming from the distributed cloud system. Processing all those data will place a heavy burden on global management. The transmission of FCAPS data from ECs to a centralized global management will occupy a wide bandwidth of the transmission network.

– **Edge autonomy management**: Global management has difficulty in maintaining autonomy management on the edge side. Firstly, ECs are far away from global management, which might lead to insecure and inefficient operation and management if ECs lose connection with centralized global management. Secondly, EC services have regional characteristics. Different areas might have diversified edge services, not all services need to be controlled by global management, e.g., some IoT or some local vertical industry services. Global management cannot achieve the flexibility and efficiency of edge autonomy management in location-based edge services.

– **Heterogeneous edge resources**: EC resources have two major types: One is a general type which carries edge services that have general resource requirements such as vehicle to everything (V2X), content delivery network (CDN), cloud games, etc. The other is a customized type which carries typical vertical industry services that have special resource requirements and are only used by certain industries, such as IoT. Different use cases have different resource requirements such as virtual machines, containers, physical servers, and all-in-one solutions. It is difficult for centralized global management to cover all of the different features and requirements with resource diversification at edge side. In addition cloud resources and services on edge nodes are provided by multiple vendors and in different versions. It is hard for global management to hide the heterogeneity of the underlying resources and offer unified services to CSCs.

– **Wide fault domain**: Upgrading, updating, breakdown or any other transient offline of global management would leave the whole distributed cloud unmanaged. So in order to ensure manageability of distributed cloud and make sure that faults in different edge regions do not affect each other, distributed cloud needs isolation from faults in the domain of local edge management.

– **Information and communication technology (ICT) united management**: To ensure low latency and wide bandwidth on edge side, the telecommunication user plane functions provided by telco-operators and cloud services provided by CSP should be deployed in the same EC. Telecommunication user plane functions are managed and orchestrated by network function virtualization orchestrator (NFVO) in a network function virtualization

(NFV) structure, which is defined in [b-ETSI NFV]. While general cloud services are managed by global management, the merging of NFVO functions and global management is complicated and should receive careful consideration in edge cloud (EC).

– **Edge coordination**: EC is limited in abilities and resources. Consequently the applications use abilities and resources from multiple ECs, and may also be deployed or migrated between multiple ECs or edge nodes. These cases require tight coordination and flexible dispatching between ECs or edge nodes. It is a challenge for global management to achieve in-time operations.

6.2 Advantages of introducing edge cloud management

To solve the challenges faced by global management to manage distributed cloud, edge cloud management is introduced into distributed cloud to manage edge cloud and support global management.

Advantages of introducing edge cloud management are as follows:

– **Relieve operation and maintenance pressure**: Edge cloud management can help transfer information between ECs and global management. It collects FCAPS data, filters important information (such as resource capacity, resource usage, faults, etc.) and reports to global management. This can help global management obtain a global view of distributed cloud, and relieve the pressure on the transmission network from global management.

– **Provide edge autonomy management**: Edge cloud management provides local management for ECs to reduce the unreliability and inflexibility of remote management.

– **Shield differences of edge cloud services**: Edge cloud management ensures unified application programming interfaces (APIs), unified services and cloud resources dispatch for CSCs by shielding the differences of edge nodes, such as different APIs, software versions, data formats, etc.

– **Support flexible ICT management**: Edge cloud management provides APIs and abilities to CSCs and other management systems such as NFVO, global management, etc. It integrates abilities required by different ICT management systems, and maintains coordination between them. This helps to merge different requirements of ICT management.

– **Promote edge coordination**: Edge cloud management provides coordination for a limited number of edge nodes to balance the flexibility and efficiency by using different ECs.

– **Cooperate with global management**: Edge cloud management cooperates with global management to achieve coordination between ECs and core cloud (CC), ECs and regional cloud (RC).

6.3 Relationship of edge cloud management with edge node

Edge cloud is a kind of cloud computing deployed to the edge of the network accessed by CSCs with small capacity resources enabling cloud service as defined in [ITU-T Y.3508]. It can be managed by edge cloud management and includes several edge nodes. The relationships between edge cloud management and edge nodes, shown in Figure 6-1, are described as follows:

– Edge cloud management is focused on offloading the operation and maintenance pressure from global management, executing edge autonomy management, shielding differences of EC services, supporting flexible ICT management and promoting edge coordination between EC to RC and CC.

– Edge nodes include different resources, such as computing, networking, storage and other special industry resources.

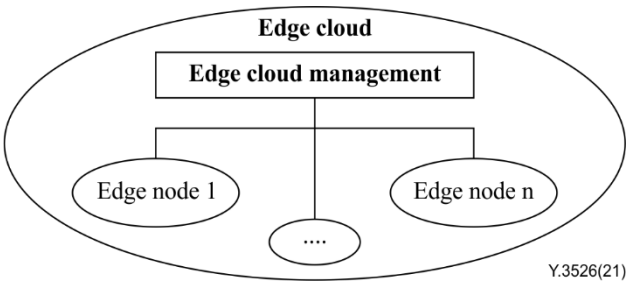


Figure 6-1 – Conceptual relationships between edge cloud management and edge nodes

Physically, edge cloud management can be deployed into a single node or several nodes as a set of functions by an administrator. Physical relationships of edge cloud management with edge nodes are shown in Figure 6-2 and Figure 6-3. Edge cloud management is deployed in single edge node with a set of functions as shown in Figure 6-2. Edge cloud management can be divided into several functions and deployed to multiple edge nodes as shown in Figure 6-3.

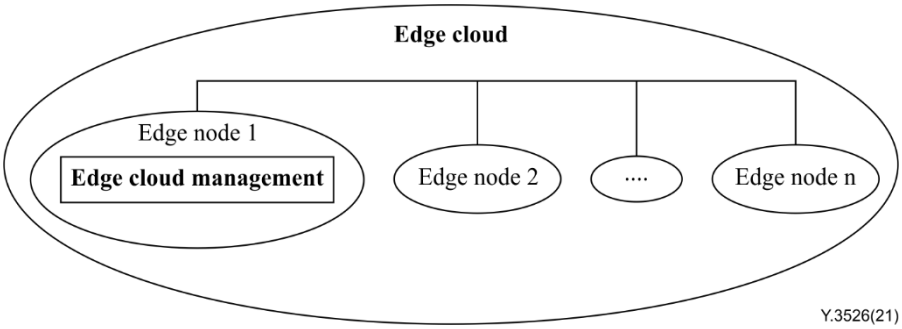


Figure 6-2 – Edge cloud management deployed in a single edge node

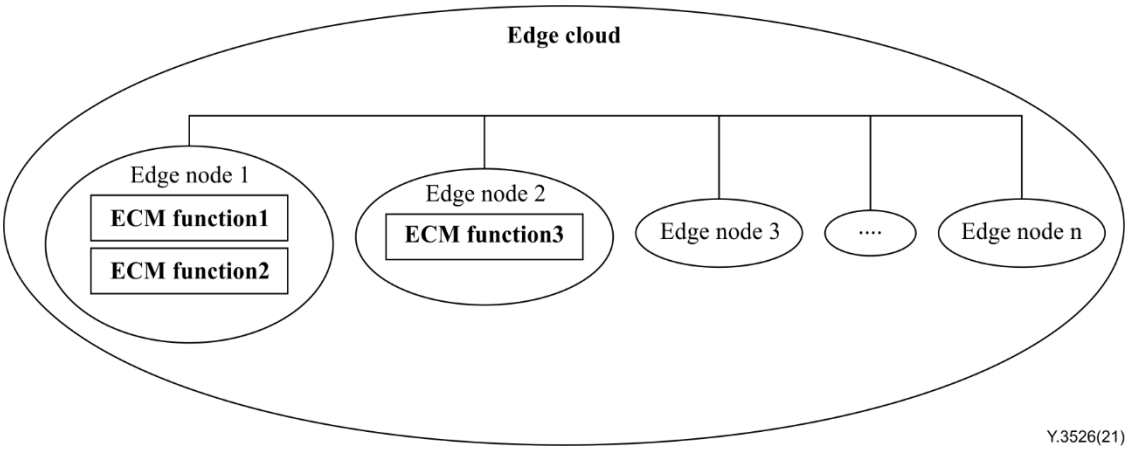


Figure 6-3 – Edge cloud management deployed in multiple nodes with a set of functions

6.4 Relationship of edge cloud management with global management in distributed cloud

In [ITU-T Y.3508], distributed cloud has a global view and local view of operations and administration management for distributed cloud resources and services. Global management works in coordination with CC, RC as well as EC from a global view. However from the local view edge cloud management is necessary to execute autonomy management, different edge nodes management, and also to interact with global management or other systems.

This Recommendation focuses on edge cloud management which provides local management for cloud resources and service in ECs, and will cooperate with global management to achieve hierarchical and automatic management of a distributed cloud.

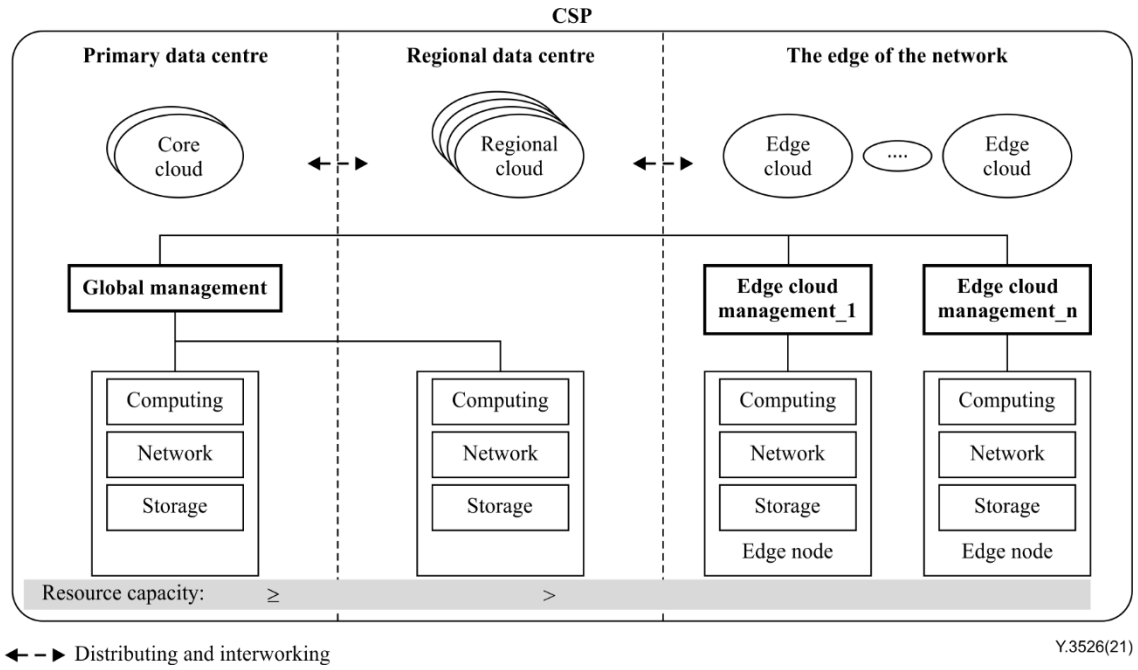


Figure 6-4 – Concept of edge cloud management in distributed cloud

Figure 6-4 shows the concept of edge cloud management in a distributed cloud. The ECs can be grouped according to different characteristics such as industry, features, locations, etc. Each group contains one or several edge nodes, and is managed by edge cloud management. Edge cloud management can manage EC independently, and can also cooperate with global management.

Relationships between global management and edge cloud management can be described as follows:

– Global management is responsible for managing the overall distributed cloud, including CC, RC and EC, which manages EC indirectly through edge cloud management.

– Edge cloud management is responsible for managing EC, and mainly focuses on management of coordination between edge nodes within one EC, and between different ECs. Edge cloud management manages EC directly.

NOTE 1 – Global management subscribes to edge cloud management to collect data of EC.

NOTE 2 – Instructions of global management given to EC will be illustrated by edge cloud management firstly and then send to target edge node for execution.

7 Edge cloud management mode in configuration models

The term "model" used here specifies structures of management. Distributed cloud configuration models are used to indicate different combinations of CC, RC, and EC, refer to Figure 6-3 of [ITU-T Y.3508]. It can be used to instruct deployment and construction. The rem "mode" used here specifies the method of management. Management mode of edge cloud management is used to indicate different kinds of management method in EC. Different management modes will be abstracted from the configuration models, and different management and interactions will be defined. This clause mainly describes the edge cloud management mode.

7.1 Edge cloud management in configuration models of distributed cloud

There are four configuration models of distributed cloud defined in Figure 6-3 and Table 6-1 of [ITU-T Y.3508]. Specifically, only model 2 (Edge-Core) and model 3 (Edge-Regional-Core) which includes EC have interactions with edge cloud management as shown in Figure 7-1.

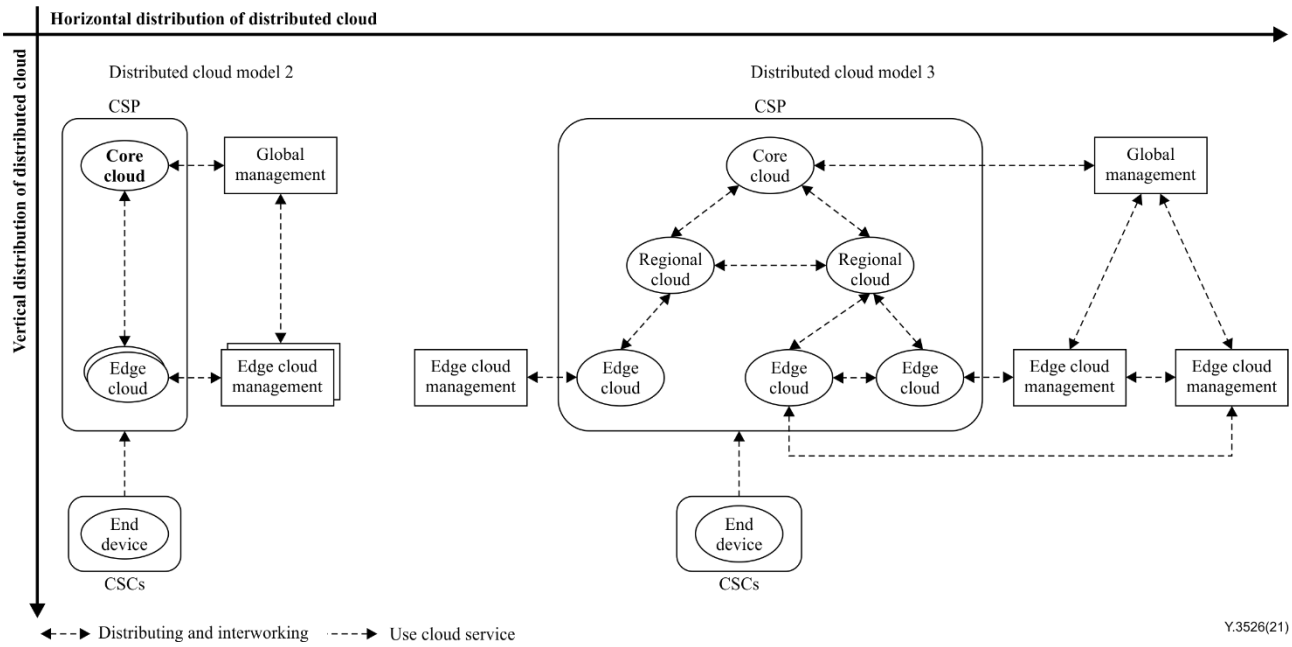


Figure 7-1 – Edge cloud management in configuration models of distributed cloud

Distributed cloud model 2 is a type of distributed cloud in which the EC and the CC are configured together. In this model, edge cloud management manages the EC, while global management manages the CC. Cloud service in this model is executed on both edge and core clouds by interworking between edge cloud management and global management.

Distributed cloud model 3 is a type of distributed cloud in which edge, regional and core clouds are configured together. In this model, edge cloud management manages the EC while global management manages the CC and RC. As cloud service in this model is deployed on edge, regional and core clouds, edge cloud management interacts with global management to support coordination among edge, regional and core clouds, and edge cloud management supports coordination between RC and CC.

7.2 Edge cloud management modes

Interactions between edge nodes, EC, and CC or RC are frequent. Therefore, edge cloud management should interact with global management in a distributed cloud. According to the illustration of edge cloud management in configuration model 2 and model 3 of a distributed cloud, there are three management modes of edge cloud management. As shown in Figure 7-2, the three management modes of edge cloud management are:

– EC local management mode,

– EC vertical management mode,

– EC horizontal management mode.

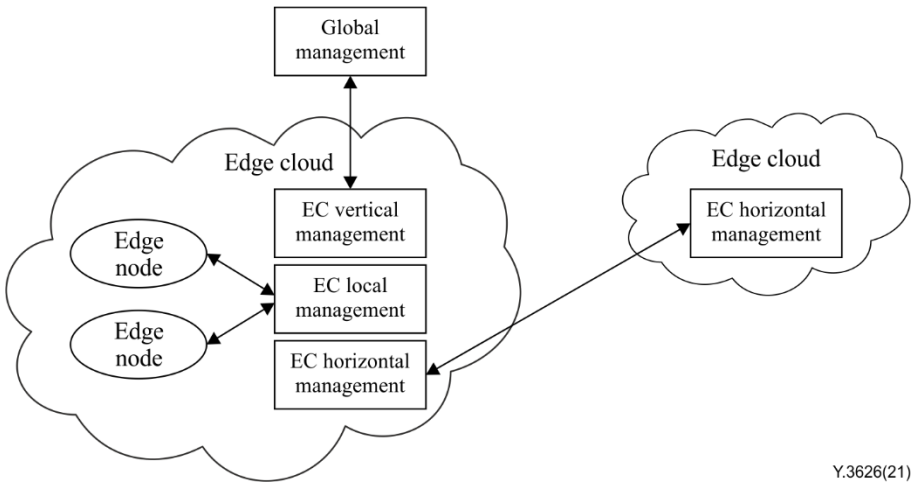


Figure 7-2 – Edge cloud management modes

– **EC local management**: EC local management manages edge nodes within a given EC under local management mode, which includes the following objects:

  • **Resource**: Infrastructures of the EC within the jurisdiction of edge cloud management, including physical and virtualized resources.

  • **Service**: Cloud services and functionalities to deliver cloud services of an EC within the jurisdiction of edge cloud management. Local management on services supports life cycle management of services deployed on an EC.

  • **Product**: Cloud service products defined in [ITU-T Y.3502] and functionalities to support product delivery of an EC within the jurisdiction of edge cloud management. EC local management of edge cloud management on products maintains the existing product catalogue defined in [ITU-T Y.3502] of a managed EC.

  • **Customer**: The customer purchases resource/service/products both in an EC and a distributed cloud. For both situations, global management has a global view of customer information while edge cloud management only focuses on customers' information and requirements on EC.

– **EC vertical management**: EC vertical management supports interactions between edge cloud management and global management. Objects managed under vertical management mode include interactions of services and customers existing across CC, RC and EC. Distributed cloud model 2 and model 3 need the vertical management mode because edge cloud management should interact with global management to support coordination among EC, CC and RC.

– **EC horizontal management**: Edge cloud management supports cooperation with between ECs. Objects managed under horizontal management mode include interactions of services, products and customers existing cross ECs, such as data exchanging, service continuity guarantees, etc. Distributed cloud model 3 needs horizontal management mode because edge cloud management should interact with other edge cloud management objects to support coordination among ECs.

To provide complete management, the three management modes should cooperate under different scenarios. Table 7-1 presents four scenarios of management modes.

Table 7-1 – Scenarios of management modes

Scenario	Description	Calling sequence
Scenario 1: Local management	If CSC only orders services in a specific EC, edge cloud management should possess local management mode to provide local EC service.	—
Scenario 2: Vertical management + Local management	If CSC orders services in core/regional cloud and a specific EC, only vertical interaction would exist between the core/regional cloud and EC. Edge cloud management possess vertical management mode to support vertical interaction, and local management mode to support any modification on EC in the vertical interaction.	Vertical -> Local
Scenario 3: Vertical management + Horizontal management + Local management	If CSC orders services in core/regional cloud and multiple ECs, vertical interaction would exist while horizontal interaction would exist among multiple ECs. Edge cloud management possess vertical management mode to support vertical interaction, horizontal management mode to support horizontal interaction, and local management to support any resource or service change on EC in the vertical interaction and horizontal interaction.	Vertical -> Horizontal/Local -> Local/Horizontal
Scenario 4: Horizontal management + Local management	If CSC orders services among multiple ECs, only horizontal interaction would exist among ECs. Edge cloud management possess horizontal management mode to support horizontal management, and local management mode to support any resource or service change on ECs in horizontal interaction.	Local -> Horizontal -> Local

8 Edge cloud local management functions

Edge cloud local management consists of the following functions:

– Service awareness: The function is derived from the "context awareness" requirement in clause 7.2 of [ITU-T Y.3508] and aims to inspect the requested cloud service by CSC based on context information described in clause 6.2 of [ITU-T Y.3508]. If the requested cloud service is inspected and exists in EC, the requested service is deployed and provided directly from an edge node to support low-latency. If not, the request is passed and executed to RC or CC.

– Service discovery: The function aims to search, trace, and maintain the service information of cloud services to deliver the CSC's request to the available endpoint. The service discovery function assigns the endpoint of the cloud service.

– Service deployment: The function aims to configure and deploy cloud services to one or multiple edge nodes to meet service requirements (e.g., low-latency, high speed network, etc.).

– Offloading: The function aims to divide the task of the cloud service into multiple sub-tasks and offload them to multiple edge nodes. Since edge nodes in EC have limited resources, the requested cloud services may be too heavy to handle and provide with low-latency to CSC. Edge cloud management transfers sub-tasks of the cloud services to one or multiple edge nodes dynamically or statically in EC. The tasks are partitioned unequally or equally according to the resource status of edge nodes dynamically in running time or statically based on the predefined policy.

– Edge node monitoring: The function aims to observe the status of edge nodes in EC such as temperature, computing resources utilization, network resources bandwidth, and services status, etc. Edge cloud management uses the monitored information to adjust unhealthy cloud services and edge nodes efficiently.
– Service integration: The function aims to integrate the completed tasks of the offloaded cloud service to deliver to the CSC. Edge cloud management maintains all sub-tasks carried on multiple edge nodes and merges them into the target of the cloud service.
– Remote backup and recovery: The function aims to backup and recover data of the edge node remotely. As EC has limited resources, edge cloud management needs to provide a remote backup and recovery mechanism by using regional or core cloud with enough resources. Edge cloud management supports set policies and triggers them to backup or recover data based on the CSC's demand.
– Software version management: The function includes delivery, deletion, and querying of the software package and aims to manage the software version of edge nodes. Since the total number of edge nodes is huge in a distributed cloud, direct software management from core to edge node is difficult. A software version update of edge nodes is controlled remotely by edge cloud management and pushed from CC by global management.
– Scaling: The function aims to scale in or out the edge nodes in EC by reducing or adding resources to ensure better utilization depending on traffic balance. If resources in EC are limited to meet the service level agreement (SLA) requirements of CSCs, edge cloud management interacts with global cloud management for scaling out.
– Service migration: The function aims to migrate cloud services running on one edge node to another edge node by copying the status of the process of the cloud service. Edge cloud management resumes the copied image of the cloud service online or offline.
– Edge node provisioning: The function aims to authenticate and authorize resources of edge nodes, and configure environment variables and system resources to be accessed by cloud services at any time. Edge cloud management provides scalable resource pools in EC. This function needs to support unified, flexible, and secure interfaces for unregistered edge nodes to involve the resource pool in EC.
– Service mash-up: The function aims to compose a cloud service from multiple microservices provided by CSPs. The cloud service has limitations to be deployed as one cloud service due to the lesser capacity of the edge node. Edge cloud management supports cloud services composed by microservices running on multiple edge nodes.
– Service customization: The function aims to configure, register, and publish the cloud services in EC according to specific CSC's requests. Cloud services on EC are not diverse enough to fulfil different requirements subscribed by various CSCs. Thus, edge cloud management supports CSCs to load customized services such as customized monitoring service, customized load balancing, etc. and instantiates them.

9 Functional requirements of edge cloud management

9.1 Resource requirements

9.1.1 Image transmission

It is recommended that CSP supports remote service image repository by edge cloud management.

NOTE – Service image includes AI reasoning model images, AI training model images, etc.

It is recommended that CSP monitors service image transmission status and progress by edge cloud management when pushing service images to an edge node.

It is recommended that CSP resumes service image transmission at break-point by edge cloud management if the transmission is interrupted due to a bad network connection on an edge node.

9.1.2 Remote software version management

It is recommended that CSP manages software versions and packages on edge node remotely by edge cloud management.

It is required that CSP supports software upgrading and rollback by edge cloud management.

NOTE – Software on edge node includes edge management software and service software, etc.

9.1.3 Remote scaling

It is recommended that CSP supports remote scaling based on edge node resources by edge cloud management to adapt to traffic workflow and improve resource usage rate.

9.1.4 Resource alarm format

It is recommended that CSP provides alarm format templates by edge cloud management to unify the alarm format.

It is recommended that CSP supports CSC or CSN to subscribe alarm reports in customized format by edge cloud management.

9.1.5 Remote health check and resume

It is required that CSP supports checking health status of edge nodes remotely and resuming service in healthy edge nodes by edge cloud management once the original nodes are unreachable.

9.1.6 Remote backup and recovery

It is recommended that CSP provides a remote backup and recovery mechanism to support EC resources and data backup and to recover services in resource-sufficient edge nodes or other clouds by edge cloud management.

9.1.7 Heterogeneous resource management

It is recommended that CSP hides heterogeneous hardware differences and provides unified services by edge cloud management.

9.2 Service requirements

9.2.1 Service inspection

It is required that CSP inspects the CSC's service request context for deploying cloud services to edge nodes by edge cloud management.

9.2.2 Service location information maintenance

It is required that CSP maintains location and access information of cloud services by edge cloud management to assign endpoints for the CSC's requests.

9.2.3 Deployment of latency sensitive service

It is recommended that CSP deploys latency sensitive cloud services by edge cloud management based on EC availability to meet SLAs.

NOTE – Examples of latency sensitive cloud service include an augmented reality (AR) service, a virtual reality (VR) service, a massive disaster alert service, and a self-driving public transportation service.

9.2.4 Context information based service offloading

It is recommended that CSP offloads the CSC's service by edge cloud management based on context information from a CC, RC, or neighbour EC to meet SLAs.

9.2.5 Delegation of service authority

It is required that CSP obtains delegation of a CSC's service authority by edge cloud management to access cloud services used by the CSC in a neighbour EC.

9.2.6 Task based service integration

It is required that CSP integrates tasks of offloaded services based on context information by edge cloud management to provide the CSC's requested cloud service.

NOTE – An example of task of offloaded services is machine learning model generation.

9.2.7 Service mash-up

It is recommended that CSP composes a cloud service by edge cloud management using multiple microservices.

9.2.8 Simplified and customizable management process

It is recommended that CSP provides a simplified and customizable management process by edge cloud management to satisfy a CSC's customized requests.

9.2.9 Customized service

It is recommended that CSP supports a CSC to request, to load and instantiate customized services by edge cloud management.

It is recommended that CSP supports a CSC to publish customized services and makes these customized services available for other CSC by edge cloud management.

10 Security considerations

It is recommended that the security framework for cloud computing described in [ITU-T X.1601] be considered for the EC. [ITU-T X.1601] analyses security threats and challenges in the cloud computing environment and describes security capabilities that could mitigate these threats and meet security challenges.

It is recommended that the guidelines for the operational security of cloud computing described in [ITU-T X.1642] be considered for the EC. [ITU-T X.1642] clarifies the security responsibilities between CSPs and CSCs, and analyses the requirements and categories of security metrics of operational security for cloud computing.

It is recommended that the security framework for the IoT based on the gateway model described in [ITU-T X.1361] be considered for ECs. In particular, [ITU-T X.1361] analyses security threats to the IoT gateway and to the network and describes security capabilities for gateways and the network that address and mitigate these security threats and challenges.

Appendix I

Use case of edge cloud management in distributed cloud

(This appendix does not form an integral part of this Recommendation.)

I.1 Use case template

The use cases developed in Appendix I should adopt the unified format presented in Table I.1 for better readability and convenient material organization.

Table I.1 – Use case template

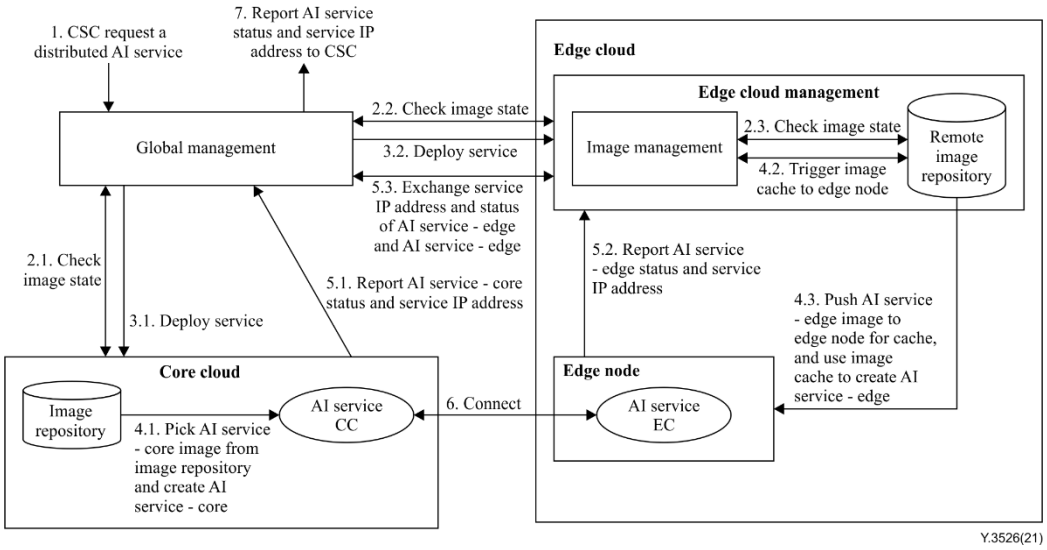
Title	Note: The title of the use case
Description	Note: Scenario description of the use case
Roles	Note: Roles involved in the use case
Figure (optional)	Note: Figure to explain the use case, but not mandatory
Pre-conditions (optional)	Note: The necessary pre-conditions that should be achieved before starting the use case.
Post-conditions (optional)	Note: The post-condition that will be carried out after the termination of current use case.
Derived requirements	Note: Requirements derived from the use cases, whose detailed description is presented in the dedicated chapter

I.2 Service image management in distributed AI service deployment

Table I.2 – Service image management in distributed AI service deployment

Title	Service image management in distributed AI service deployment
Description	This use case describes the edge cloud management to provide image management when deploying distributed AI service. Distributed AI service for distributed cloud is a kind of separated service combined with CC and EC. In Figure I.1 AI service – CC is in charge of CPU-costing model training, and AI service – EC is in charge of data collection and data cleaning and reasoning, which requires less compute resources. The process of CSC requesting distributed AI service to CC and edge is as follows: – Global management checks service image status in CC, edge cloud management checks service image status in service image repository on edge cloud management (step 2.1, 2.2, 2.3 in Figure I.1) – Trigger service deployment (step 3.1, 3.2 in Figure I.1) – CC picks AI service –CC created by image from service image repository on CC (step 4.1 in Figure I.1). Edge cloud management triggers service image cache to edge node (step 4.2 in Figure I.2). AI service –EC image is pushed to edge node form remote service image repository on EC and create AI service in EC (step 4.3 in Figure I.1). – After AI service has been created in core and edge cloud, service status and IP address would be reported to global management and edge cloud management for service connection. (step 5.1, 5.2, 5.3, 6 in Figure I.1) Report the whole AI service status and IP address to CSC (step 7 in Figure I.1) In this use case, edge cloud management is required to support remote service image repository as edge node is limited in resources. When pushing service image to edge node for cache, edge cloud management is required to monitor service image

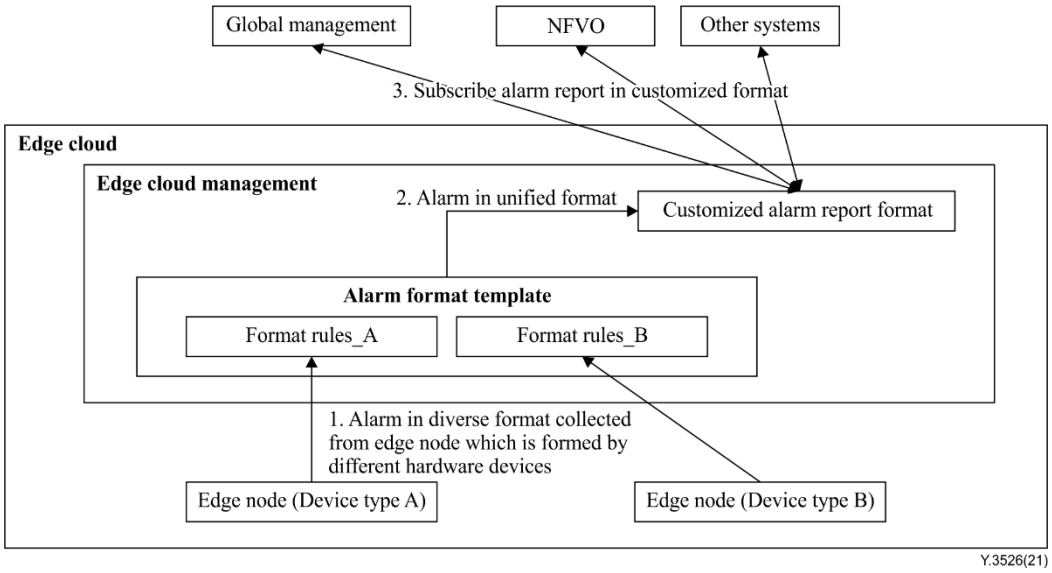
Table I.2 – Service image management in distributed AI service deployment

	transmission status and support transmission resuming at break-point when bad network connection causes transmission interruption. After AI service – CC/EC has been created, edge cloud management is required to maintain service access information (e.g., IP address) so that AI service – EC can push structured data to AI service – core for model training.
Roles	CSC, CSP
Figure (optional)	 **Figure I.1 – Service image management in distributed AI service deployment**
Pre-conditions (optional)	AI service is packaged into related service images and stored separately in CC image repository and EC image repository.
Post-conditions (optional)	
Derived requirements	Image transmission (refer to clause 9.1.1) Service location information maintenance (refer to clause 9.2.2)

I.3 Alarm format management

Table I.3 – Alarm format management

Title	Alarm format management
Description	This use case, shown in Figure I.2, describes the alarm format management requirements of EC. Edge cloud management manages edge nodes from different hardware devices, of which alarm format are different. To implement unified alarm display and storage, it is recommended edge cloud management to provide alarm format template, which contains device-specific format rules and can reformat alarms into unified format. As edge cloud management connects with other management systems including global management, NFVO, etc., it is recommended edge cloud management support customized alarm report format, through which other management systems can subscribe alarms based on their formats, and edge cloud management will package and report alarms in customized format.
Roles	CSC, CSP

Figure (optional)	 **Figure I.2 – Alarm format management in EC**
Pre-conditions (optional)	
Post-conditions (optional)	
Derived requirements	– Alarm format (refer to clause 9.1.4)

I.4 Edge cloud service to a stationary CSC

Table I.4 – Edge cloud service to a stationary CSC

Title	Edge cloud service to a stationary CSC
Description	This use case describes the edge cloud management to provide a cloud service to a stationary CSC. When a stationary CSC requests a cloud service to EC, the edge cloud management in EC handles the CSC's request based on service awareness, service discovery, and service deployment functions. Service awareness function inspects the requested cloud service by CSC. Service discovery function tracks and maintains the service information of cloud services (e.g., the service endpoint) to deliver the CSC's request to the available endpoint. Service discovery function assigns the endpoint of the cloud service. Service deployment function configures and deploys a cloud service to EC which supports low-latency. There are several options as follows. <Option 1: The cloud service provided directly from CC> (See Figure I.3) If the requested cloud service by CSC does not exist in EC, the service discovery function delivers the service request to CC and provides the cloud service. <Option 2: The cloud service provided directly in EC> (See Figure I.4) If the requested cloud service by CSC exists in EC, then the service discovery function delivers the cloud service request to edge node in EC. The edge node provides the cloud service. <Option 3: The cloud service deployed in CC, displaced to EC and provided from EC > (See Figure I.5) If the requested cloud service by CSC does not exist in EC, the service deployment function deploys the cloud service to edge node in EC. The edge node provides the cloud service with low-latency.

Table I.4 – Edge cloud service to a stationary CSC

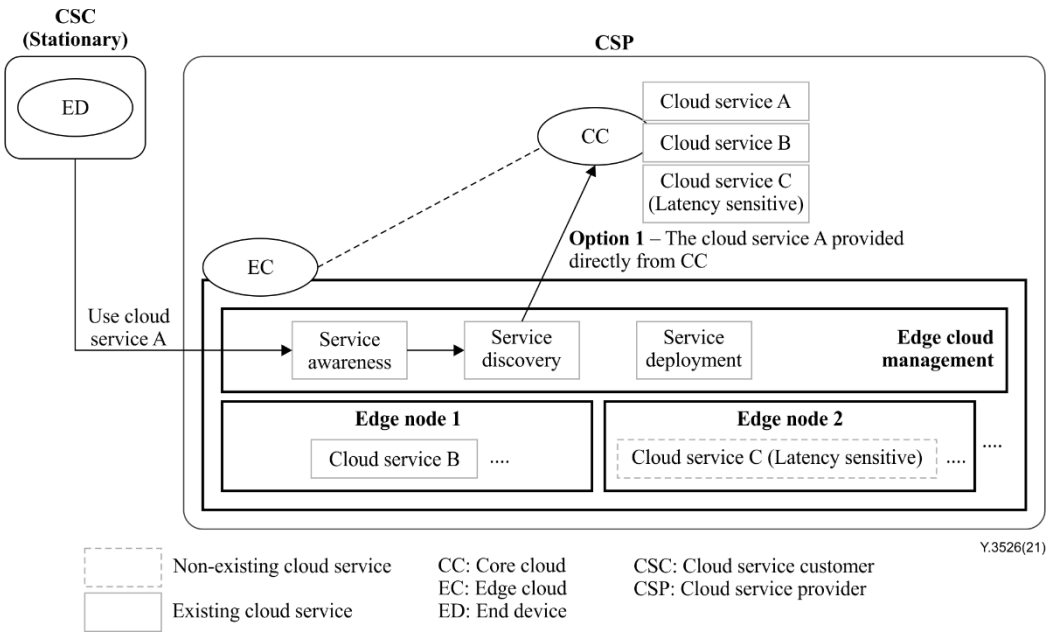
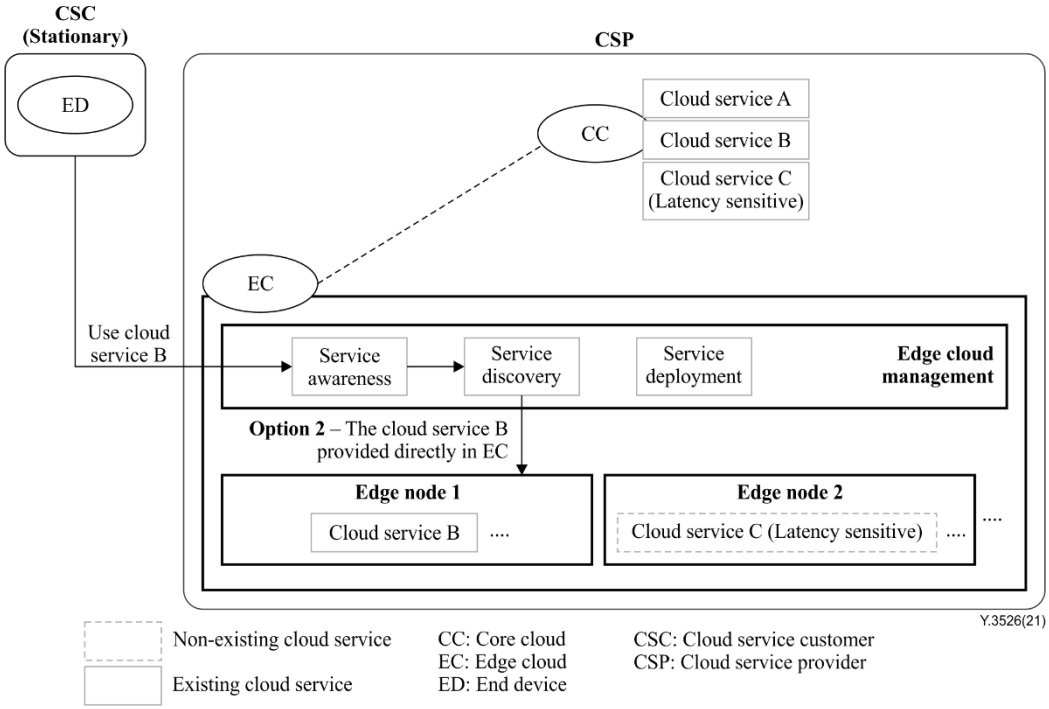
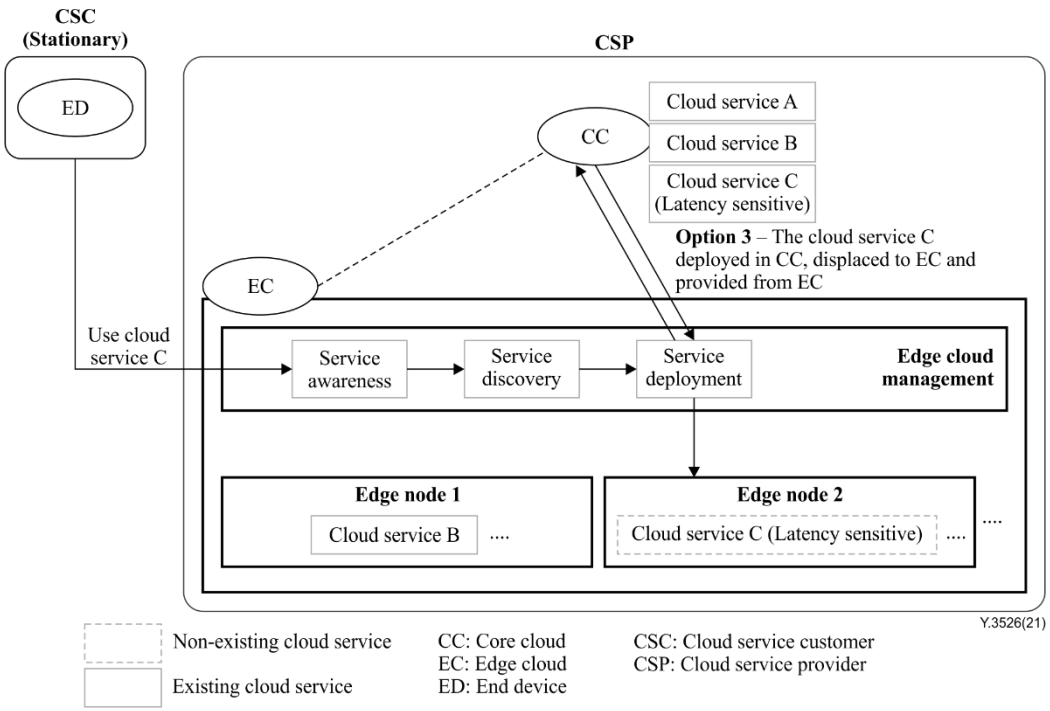
	NOTE 1 – Service awareness is derived from "context awareness" requirement in [ITU-T Y.3508] clause 7.2. NOTE 2 – This use case is derived from use case II.1 in [ITU-T Y.3508].
Roles	CSP, CSC
Figure	 **Figure I.3 – An example of providing service directly from CC**  **Figure I.4 – An example of providing service directly in EC**

Table I.4 – Edge cloud service to a stationary CSC

	 **Figure I.5 – An example of providing service by deployment from CC to EC**
Pre-conditions (optional)	
Post-conditions (optional)	
Derived requirements	– Service inspection (refer to clause 9.2.1) – Service location information maintenance (refer to clause 9.2.2) – Deployment of latency sensitive service (refer to clause 9.2.3)

I.5 Edge cloud service seamlessly to mobile CSC

Table I.5 – Edge cloud service seamlessly to mobile CSC

Title	Edge cloud service seamlessly to mobile CSC
Description	This use case describes the edge cloud management to provide a cloud service seamlessly to mobile CSC. The mobile CSC connected to EC 1 moves and connects to EC 2. The edge cloud management in EC 2 authorizes the CSC to delegate authority of the cloud service, is aware service context of the used cloud service, and then provides the cloud service. The several workflow options are as follows. <Option 1: The cloud service provided directly from CC> (See Figure I.6) If the requested cloud service by CSC does not exist in EC 2, the edge cloud management delivers the service request to CC and CC provides the cloud service. <Option 2: The cloud service provided directly in EC 2> (See Figure I.7) If the requested cloud service by CSC exists in EC 2, then the edge cloud management delivers the service request to edge node 1 in EC 2. The edge node 1 provides the cloud service. <Option 3: The cloud service deployed in CC, displaced to EC 2 and provided from EC 2> (See Figure I.8)

Table I.5 – Edge cloud service seamlessly to mobile CSC

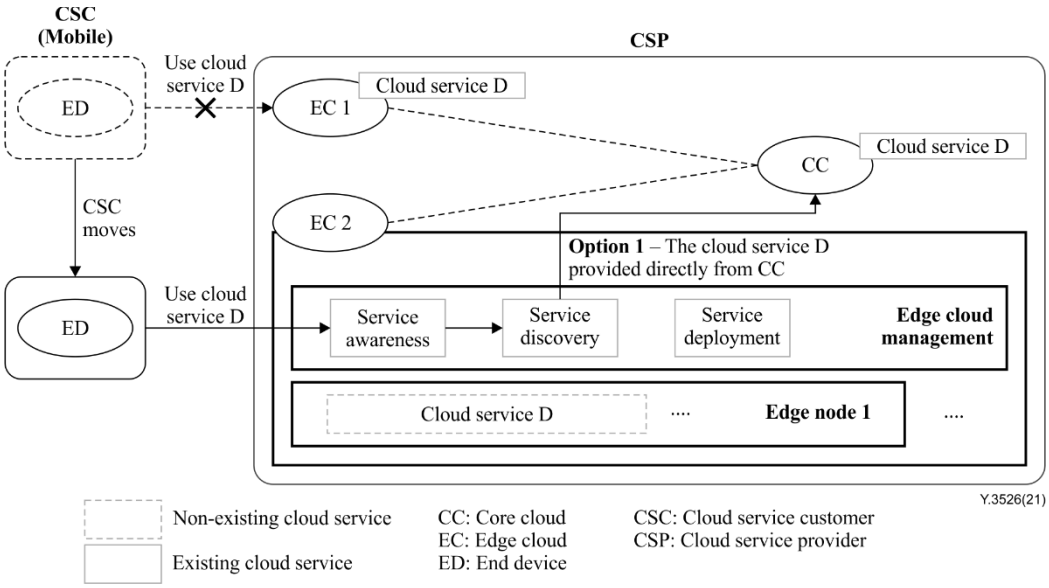
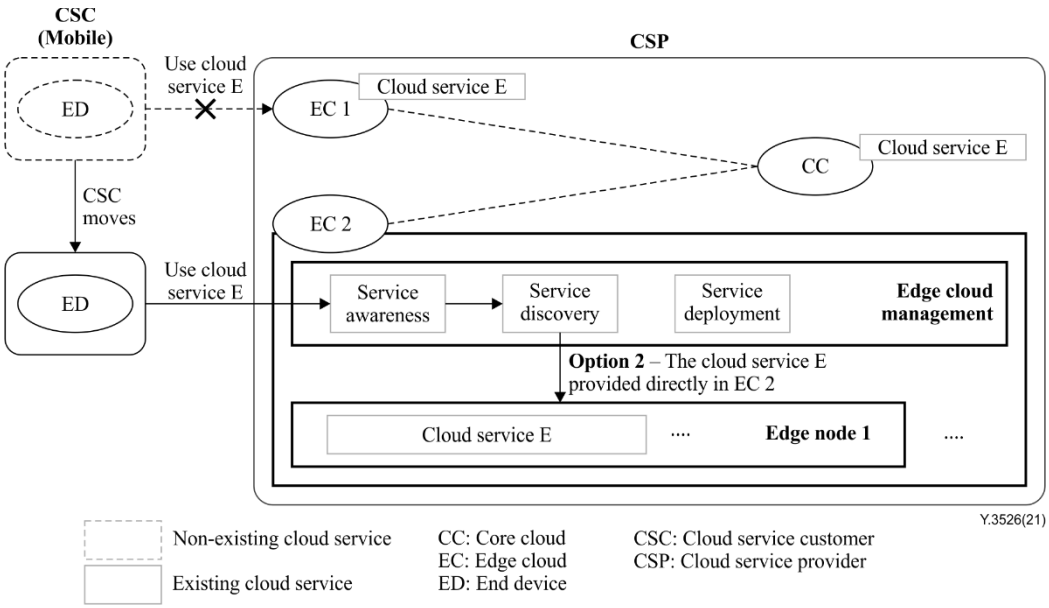
	If the requested cloud service by CSC does not exist in EC 2, the edge cloud management deploys the cloud service to edge node 1 in EC 2. The edge node 1 provides the cloud service with low-latency. <Option 4: The cloud service provided from EC 1> (See Figure I.9) If the requested cloud service by CSC does not exist in EC 2, the edge cloud management in EC 2 obtains delegation of CSC's authority, requests the cloud service to EC 1 based on cloud service context information (described in [ITU-T Y.3508]), and then EC 1 provides the cloud service. NOTE – This use case is derived from use case II.6 in [ITU-T Y.3508].
Roles	CSP, CSC
Figure	 **Figure I.6 – An example of providing service directly from CC**  **Figure I.7 – An example of providing service directly in EC 2**

Table I.5 – Edge cloud service seamlessly to mobile CSC

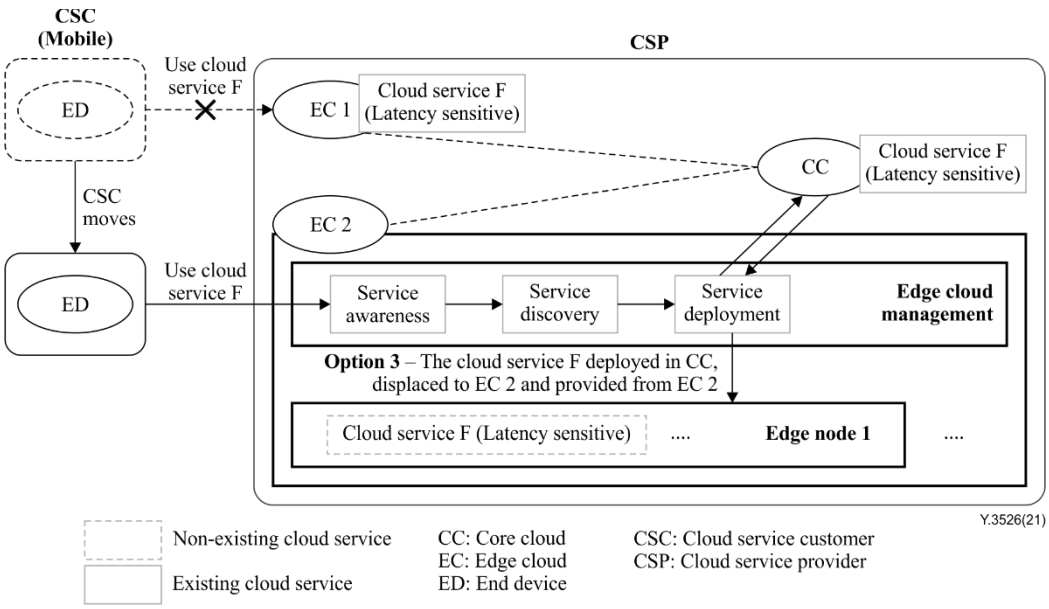


Figure I.8 – An example of providing service by deployment from CC to EC 2

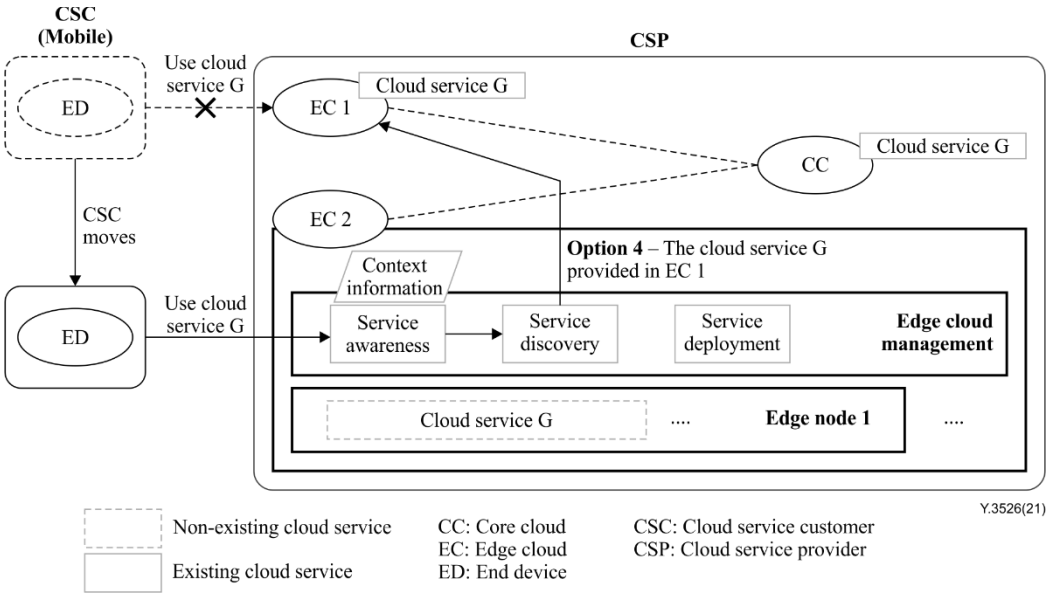


Figure I.9 – An example of providing service in EC 1

Pre-conditions (optional)	A CSP maintains information of CSC's authority for delegation. NOTE – Information of CSC's authority refers to information of access rights and privileges.
Post-conditions (optional)	
Derived requirements	– Service inspection (refer to clause 9.2.1) – Service location information maintenance (refer to clause 9.2.2) – Deployment of latency sensitive service (refer to clause 9.2.3) – Context information based service offloading (refer to clause 9.2.4) – Delegation of service authority (refer to clause 9.2.5)

I.6 Edge cloud service offloading

Table I.6 – Edge cloud service offloading

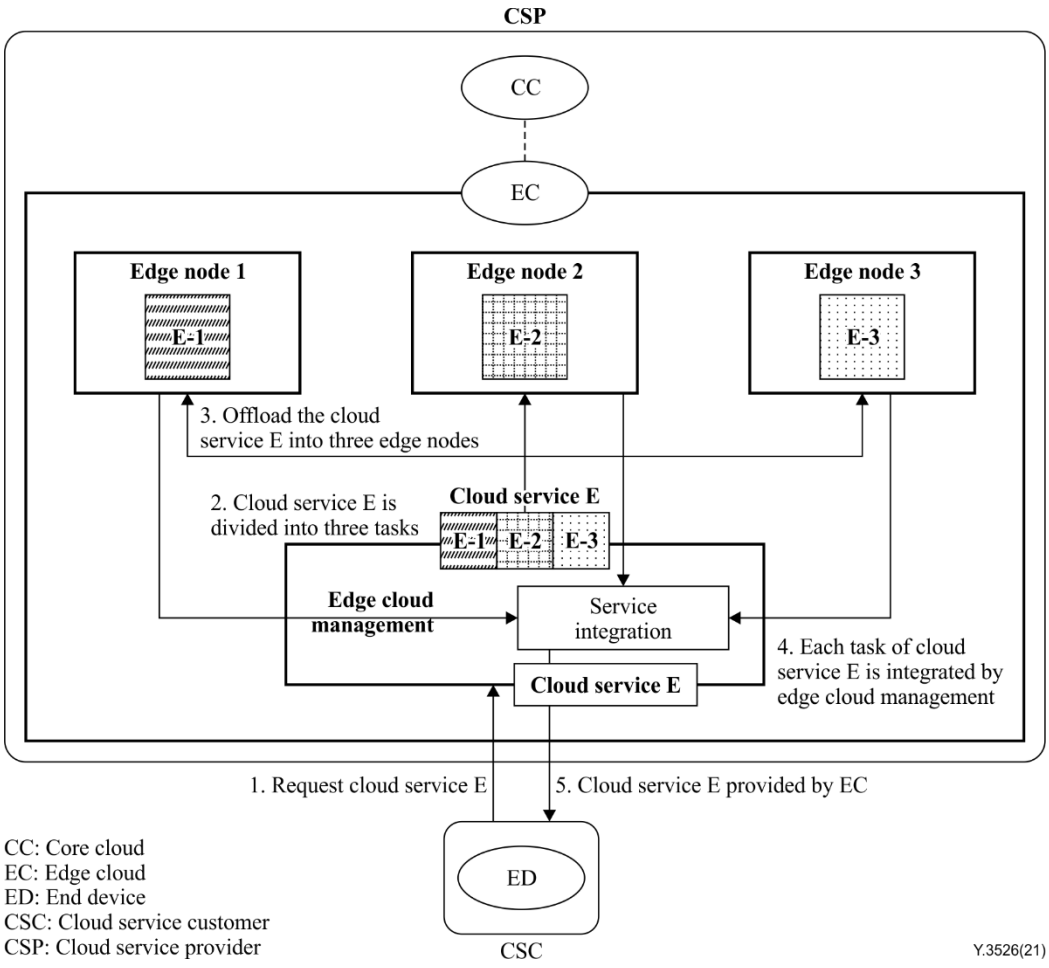
Title	Edge cloud service offloading
Description	This use case, shown in Figure I.10, describes the edge cloud management to provide a cloud service with service offloading. When CSC requests a cloud service to EC, the edge cloud management handles the CSC's request. Service awareness function in the edge cloud management inspects the CSC's service based on context information. The tasks of the offloaded cloud service are executed in each edge node. When the executions are done, service integration function in edge cloud management integrates the tasks of the offloaded cloud service and provides the cloud service to CSC. An example of the cloud service provided with offloading is machine learning (ML) model training. ML model training service is offloaded to edge nodes with divided data. When each offloaded service is done, then the edge cloud management integrates each trained model and provides a global trained model to CSC.
Roles	CSP, CSC
Figure	 **Figure I.10 – Process of edge cloud management service offloading**
Pre-conditions (optional)	
Post-conditions (optional)	

Table I.6 – Edge cloud service offloading

Derived requirements	– Service inspection (refer to clause 9.2.1) – Service location information maintenance (refer to clause 9.2.2) – Context information based service offloading (refer to clause 9.2.4) – Task based service integration (refer to clause 9.2.6)

I.7 Edge cloud management as a service broker with microservices

Table I.7 – Edge cloud management as a service broker with microservices

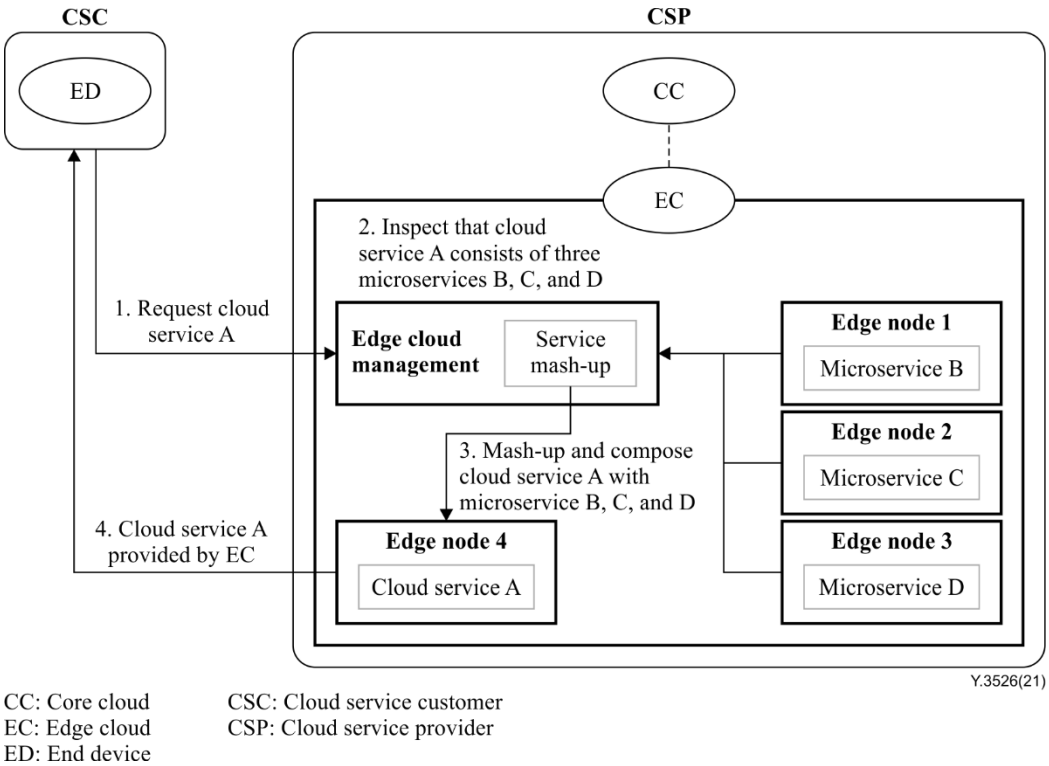
Title	Edge cloud management as a service broker with microservices
Description	This use case, shown in Figure I.11 describes the edge cloud management as a service broker. When CSC requests the cloud service consisting of microservices, the edge cloud management handles the CSC's request with service mash-up function. If the requested cloud service is composed of several microservices, then the CSC's request delivered to the edge nodes that the microservices are deployed. Service mash-up function composes the microservices to produce the cloud service. Figure I.11 shows the splitting mechanism of cloud service by service mash-up as service broker in edge cloud management. For example, the video editing cloud service consists of three microservices such as format conversion, caption design, and speed control.
Roles	CSP, CSC
Figure	

Figure I.11 – Service mash-up as a service broker

Table I.7 – Edge cloud management as a service broker with microservices

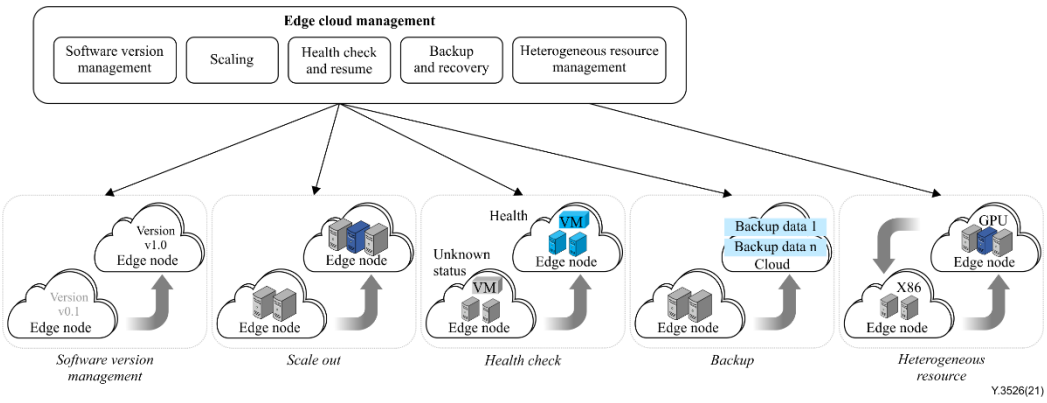
Pre-conditions (optional)	
Post-conditions (optional)	
Derived requirements	– Service inspection (refer to clause 9.2.1) – Service location information maintenance (refer to clause 9.2.2) – Service mash-up (refer to clause 9.2.7)

I.8 Remote management for edge nodes

Table I.8 – Remote management for edge nodes

Title	Remote management for edge nodes
Description	This use case, shown in Figure I.12, describes the edge cloud management to manage edge nodes remotely. Six examples have been provided to derive requirements. **Example 1**: Edge nodes of EC are usually large in number and widely distributed. This makes direct software management on edge node impossible, which recommends edge cloud management to manage edge nodes software remotely within stable and remote management network. All the software packages can be delivered, deleted, queried remotely. And the software upgrading and rollback in edge nodes can be controlled remotely by edge cloud management. **Example 2**: The business traffic on EC of single edge node may change suddenly or regularly on different locations, such as during World Cup period, workday, weekend. To deal with traffic growth and ensure service quality, edge cloud management should scale out edge node remotely by adding resources. To ensure better utilization of resources during period with less traffic, edge cloud management should scale in edge node remotely by releasing idle resources. These resources can come from other edge nodes in the same edge, etc. **Example 3**: Edge cloud is usually located in harsh environment instead of in data centre, for example high-temperature and dusty environment. It is recommended that edge cloud management to support remote health check on edge nodes. Once the edge node is in unknown status and application is unreachable, edge cloud management should pick another edge node to resume the application. **Example 4**: Edge cloud has limited resources, it is recommended to reserve more resources for CSC's service and local backup is not suggested. Edge cloud management is recommended to provide remote backup and recovery mechanism. Edge cloud management supports to set backup policies for edge nodes and trigger them to backup data on demand. Edge nodes resource or data could be backup to RC or CC with enough resources. **Example 5**: CC and RC have huge amount of resources to serve multiple CSCs, for which complementary and unified process is easier to maintain. However, EC has limited resources, sometimes can only serve one CSC in local industry environment like private cloud. It is recommended that edge cloud management could support simplified and customizable management process to fulfil different CSC and various CSC's service requirements on demand.

Table I.8 – Remote management for edge nodes

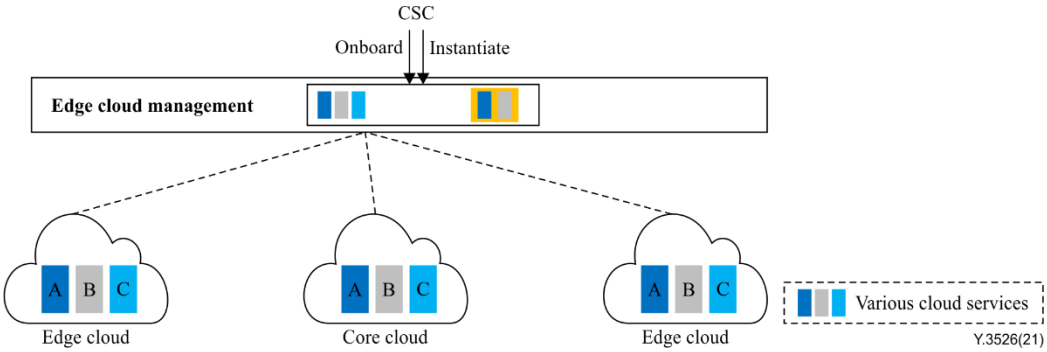
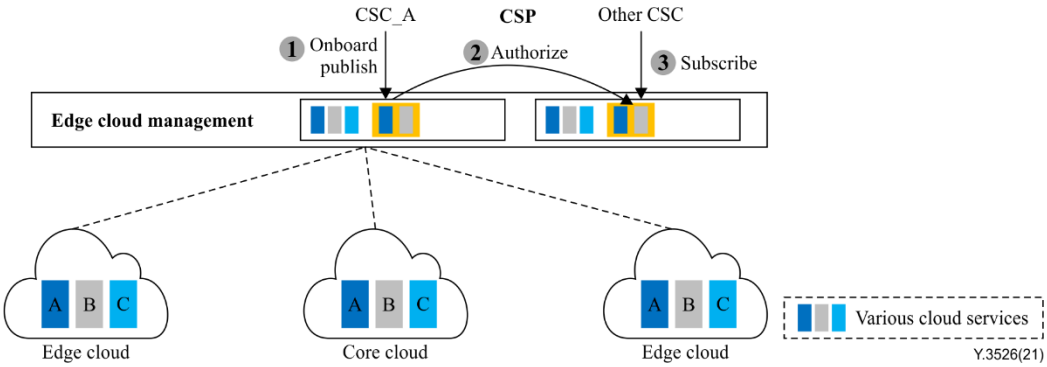
	Example 6: The CSC's service could use different type of cloud resource in EC. Taking one AI service for example, web service module needs VM or container, but video codec module needs GPU. So different edge nodes might have diversified device type. It is recommended that edge cloud management hides heterogeneous hardware differences and provide unified services.
Roles	CSP, CSC
Figure	 **Figure I.12 – Remote management for edge nodes**
Pre-conditions (optional)	
Post-conditions (optional)	
Derived requirements	– Remote software version management (refer to clause 9.1.2) – Remote scaling (refer to clause 9.1.3) – Remote health check and resume (refer to clause 9.1.5). – Remote backup and recovery (refer to clause 9.1.6) – Heterogeneous resource management (refer to clause 9.1.7) – Simplified and customizable management process (refer to clause 9.2.8)

I.9 Customized service management

Table I.9 – Customized service management

Title	Customized service management
Description	This use case describes the edge cloud management to provide customized service management. **Example 1**: As EC has limited resources, and CSC requirements on EC are not universal, services on EC may not be diverse enough to fulfil different CSCs requirements. To balance between limited service types in EC and non-universal edge service requirements of CSC, it is recommended that edge cloud management supports CSC to load customized services and instantiates them. These customized services could be provided by CSC, for example customized monitoring service, customized load balancer, etc., used to support multiple CSC's service designing, running, monitoring, management, etc. (see Figure I.13)

Table I.9 – Customized service management

	Example 2: Following Example 1, if the customized service loaded by one CSC ("CSC_A" in Figure I.14) can represent requirements of a group of CSCs ("other CSC" in Figure I.14), then CSC_A can choose to make its customized service publicly available (step ① in Figure I.14), and let other CSC to subscribe and use this service (step ③ in Figure I.14). In step ②, CSP should authenticate CSC_A as a legal customized service provider, so that CSC_A could load and publish customized service. Edge cloud management authenticates the published service as a legal and public available service on EC, and authorize other CSC to subscribe this service. This example recommends edge cloud management to support customized service publishing.
Roles	CSP, CSC
Figure	 **Figure I.13 – Edge cloud management supports CSC to load customized service**  **Figure I.14 – Edge cloud management supports CSC to publish customized service**
Pre-conditions (optional)	
Post-conditions (optional)	
Derived requirements	– Customized service (refer to clause 9.2.9)

Bibliography

[b-ETSI NFV] ETSI NFV White Paper 2: Network Functions Virtualisation – Introductory White Paper
https://portal.etsi.org/NFV/NFV_White_Paper2.pdf

SERIES OF ITU-T RECOMMENDATIONS

Series A	Organization of the work of ITU-T
Series D	Tariff and accounting principles and international telecommunication/ICT economic and policy issues
Series E	Overall network operation, telephone service, service operation and human factors
Series F	Non-telephone telecommunication services
Series G	Transmission systems and media, digital systems and networks
Series H	Audiovisual and multimedia systems
Series I	Integrated services digital network
Series J	Cable networks and transmission of television, sound programme and other multimedia signals
Series K	Protection against interference
Series L	Environment and ICTs, climate change, e-waste, energy efficiency; construction, installation and protection of cables and other elements of outside plant
Series M	Telecommunication management, including TMN and network maintenance
Series N	Maintenance: international sound programme and television transmission circuits
Series O	Specifications of measuring equipment
Series P	Telephone transmission quality, telephone installations, local line networks
Series Q	Switching and signalling, and associated measurements and tests
Series R	Telegraph transmission
Series S	Telegraph services terminal equipment
Series T	Terminals for telematic services
Series U	Telegraph switching
Series V	Data communication over the telephone network
Series X	Data networks, open system communications and security
Series Y	**Global information infrastructure, Internet protocol aspects, next-generation networks, Internet of Things and smart cities**
Series Z	Languages and general software aspects for telecommunication systems

Printed in Switzerland
Geneva, 2021