International Telecommunication Union

ITU-T

TELECOMMUNICATION
STANDARDIZATION SECTOR
OF ITU

Y.3508

(08/2019)

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

Cloud Computing

Cloud computing – Overview and high-level requirements of distributed cloud

Recommendation ITU-T Y.3508

ITU-T

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
Network control architectures and protocols	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.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.3508

Cloud computing – Overview and high-level requirements of distributed cloud

Summary

Recommendation ITU-T Y.3508 provides an overview and high-level requirements for distributed cloud. This Recommendation introduces the concept of the distributed cloud and identifies the characteristics of distributed cloud. Based on concept and characteristics, configuration models are illustrated. Deployment considerations of distributed cloud are provided in perspective of infrastructure, network, service, management and security. From use cases, high-level requirements of the distributed cloud are derived.

History

Edition	Recommendation	Approval	Study Group	Unique ID*
1.0	ITU-T Y.3508	2019-08-13	13	11.1002/1000/13988

Keywords

Cloud computing, distributed cloud, high-level requirement.

* 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, 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 TSB patent database at http://www.itu.int/ITU-T/ipr/.

Table of Contents

Recommendation ITU-T Y.3508

Cloud computing – Overview and high-level requirements of distributed cloud

1 Scope

This Recommendation provides an overview and high-level requirements of the distributed cloud. It addresses the following subjects:

– definition of distributed cloud;
– concept of distributed cloud;
– characteristics of distributed cloud;
– configuration models of distributed cloud;
– deployment considerations of distributed cloud; and
– high-level requirements of distributed cloud.

Use cases of distributed cloud are provided in Appendices I and II. Also, a comparison between distributed cloud and other technologies is described in Appendix III.

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

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

[ITU-T Y.3511] Recommendation ITU-T Y.3511 (2014), *Framework of inter-cloud computing*.

3 Definitions

3.1 Terms defined elsewhere

This Recommendation uses the following terms defined elsewhere:

3.1.1 cloud capabilities type [IUT-T Y.3500]: Classification of the functionality provided by a cloud service to the cloud service customer, based on resources used.

NOTE – The cloud capabilities types are application capabilities type, infrastructure capabilities type and platform capabilities type.

3.1.2 cloud computing [IUT-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.

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

3.1.4 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.

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

3.2 Terms defined in this Recommendation

This Recommendation defines the following terms:

3.2.1 cloud service image: An executable code with state information of virtual machine or container.

NOTE 1 – State information includes execution states, disk states and memory states of system, such as program counter, registry, disk volume number, disk volume size, memory stack and information of allocated memory.

NOTE 2 – Cloud service image includes operating system, libraries, data files, applications, etc.

3.2.2 core cloud: 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).

3.2.3 distributed cloud: 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.

3.2.3 edge cloud: 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.

3.2.5 regional cloud: 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).

4 Abbreviations and acronyms

This Recommendation uses the following abbreviations and acronyms:

AR Augmented Reality

CC Core Cloud

CSC Cloud Service Customer

CSP Cloud Service Provider

EC Edge Cloud

IoT Internet of Things

MEC Multi-access Edge Computing

ML Machine Learning

QoS Quality of Service

RC Regional Cloud

SLA Service Level Agreement

VR Virtual Reality

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 distributed cloud

6.1 Concept of distributed cloud

Cloud computing is a 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 [ITU-T Y.3500]. Services in Internet of things (IoT), augmented reality (AR)/virtual reality (VR), artificial intelligence (AI) and 5G application domains requiring cloud infrastructure are gradually increasing and real-time services in the above domains are also demanding especially for safety and surveillance applications.

Cloud computing has challenges to support the real-time services due to (i) the network latency between cloud service provider (CSP) and cloud service customer (CSC), and (ii) the load congestion on the data centre. For these challenges, what is needed is the delivery of cloud services to nearby CSCs in order to meet real-time service delivery. Also, the distribution of cloud capabilities types to the edge of the network accessed by CSCs are needed for load congestion.

This Recommendation introduces the distributed cloud, which is 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 a pool of physical or virtual resources.

Figure 6-1 shows a concept of distributed cloud. The distributed cloud includes core, regional and edge clouds, which meet the cloud capabilities types described in [ITU-T Y.3500]. Cloud services are deployed to the core, regional and edge clouds, interwork with one another, and provide a single system view to the CSCs for location transparency. Thus, the distributed cloud provides low latency and fast response to access cloud services by CSCs to satisfy their need for real-time services in various areas.

Global management manipulates both distributed cloud resources and cloud service distribution appropriately for CSC's demands. The distributed cloud resource is the aggregated infrastructure from core, regional and edge clouds, such as physical or virtual resources of compute, storage and network.

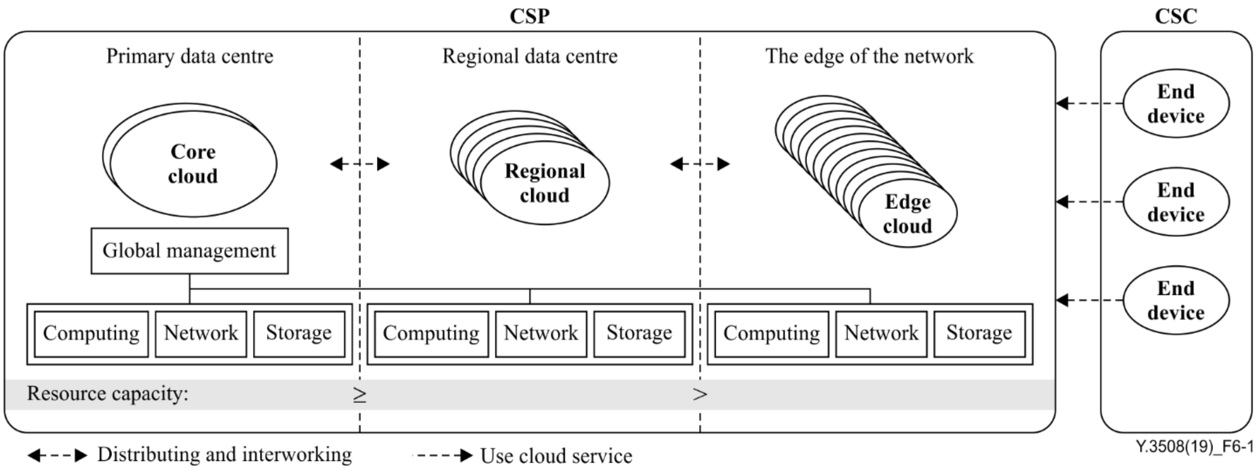


Figure 6-1 – Concept of distributed cloud

– **Core cloud (CC)**: The CC has large resource capacity and global management point to control cloud resources in the distributed cloud. The core cloud supports cloud services with high computing intensive and geographic independence;

– **Regional cloud (RC)**: The RC is optionally deployed in particular regions from the core cloud for load sharing and service quality enhancement. The regional cloud handles cloud service requests from the region controlled by global management of the core cloud;

NOTE 1 – The regional cloud supports lower latency than the core cloud by executing customized cloud services for CSCs in a particular region. It is assumed that the network latency from the CSC to the regional cloud is lower than from the CSC to the core cloud and that the difference of cloud service execution time between the core and regional cloud is negligible.

NOTE 2 – The regional cloud performs buffering load of cloud service and caching data from the core cloud and provides them to CSCs in the region.

– **Edge cloud**: The edge cloud is deployed at the edge of the network accessed by CSCs and has a small resource capacity. The edge cloud requires specialized hardware resources on purpose; i.e., the resources in the edge cloud are constrained due to limitations of space or power. The edge cloud may have different configurations of resources and cloud capabilities types with physical and virtual resources depending on a CSC's requirements of cloud services and conditions in the deployment environment.

Figure 6-2 shows an example of providing real-time service on distributed cloud for the case of a machine learning (ML) service. The example has four phases.

– **Phase 1 – collecting**: End device 1 transfers sensor data to an endpoint of ML training service in the core cloud to train the data;

– **Phase 2 – training**: ML training service trains the data and gets a trained rule at the core cloud;

– **Phase 3 – caching**: The trained rule is cached from core cloud to regional cloud; the trained rule is then cached to edge clouds 1 and 2;

– **Phase 4 – forecasting**: End devices 1 and 2 transfer sensor data to edge clouds 1 and 2, and then edge clouds 1 and 2 perform ML forecasting in real-time processing. The forecasting result is quickly delivered to end devices 1 and 2.

If the ML forecasting service is also processed at the core cloud without the edge cloud, then the end device of the CSC will receive the forecasting results from the core cloud with high latency. The example in Figure 6-2 highlights how the distributed cloud provides benefit to the cloud service in the edge cloud by providing low latency and real-time processing.

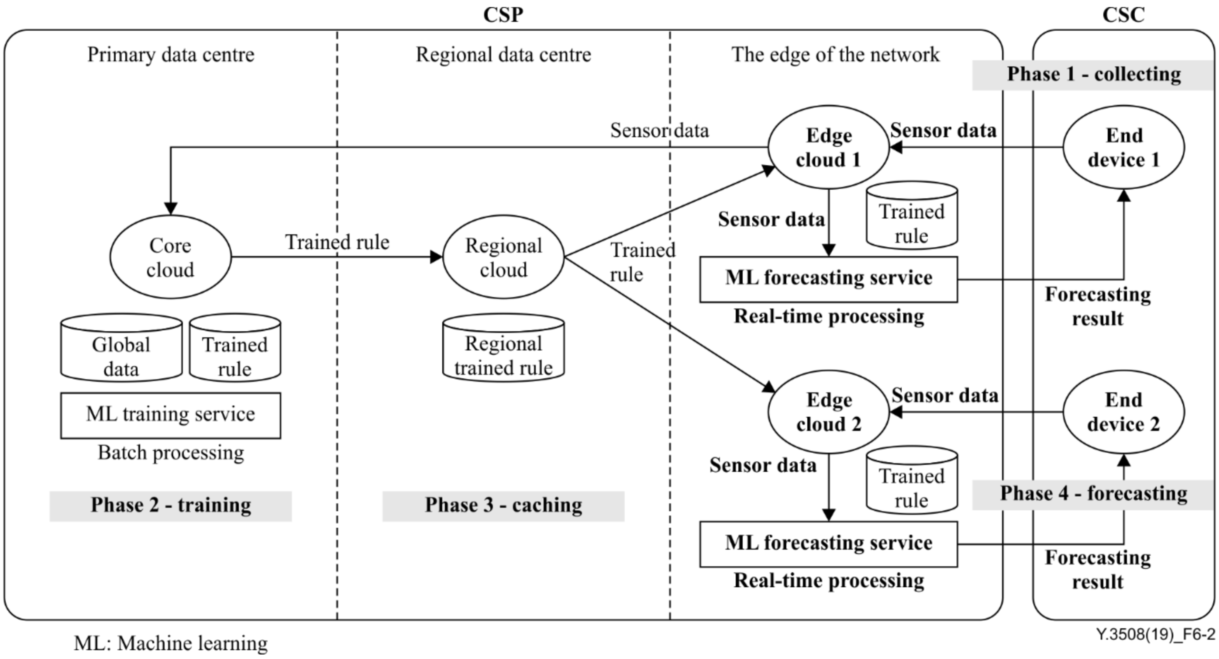


Figure 6-2 – Example of providing real-time service on distributed cloud for the case of a machine learning service

NOTE 3 – In this example, it is assumed that the network latency from the end devices to the edge cloud is much lower than from the end devices to the core cloud, and that the execution time of the ML forecasting service is faster than the ML training service.

6.2 Characteristics of distributed cloud

The characteristics of cloud computing [ITU-T Y.3500] are inherited to distributed cloud as:

– broad network access;
– measured service;
– multi-tenancy;
– on-demand self-service;
– rapid elasticity and scalability;
– resource pooling.

The distributed cloud also has the following additional characteristics:

– **distributed cloud resource**: The resources of distributed cloud are geographically distributed at various types of data centres, base stations and IoT gateways. The distributed cloud resources are pooled on demand to meet CSC requirements by using different physical and virtual resources dynamically;
– **heterogeneous infrastructure**: The distributed cloud has different scale infrastructures with large resource capacity in core/regional cloud and small resource capacity in edge cloud. The distributed cloud needs to utilize heterogeneous infrastructure as a single system to provide various services to CSCs;
– **context awareness network**: The distributed cloud has on-demand network control according to context information for CSCs. Through context awareness network, mirrored cloud service to edge cloud is provided directly to CSCs;

  NOTE – Context information includes cloud service context information and network service context information. Examples of cloud service context information are capacity of cloud, resource usage, service ID/name, service location, status of neighbour. Examples of network service context information are latency, bandwidth, quality of service (QoS), routing path.

– **agile service**: Cloud services can be deployed to the distributed cloud resource quickly and can be easily developed. The developed cloud services can be rapidly and dynamically provided on-demand across various types of distributed cloud resources;

– **autonomous management**: The distributed cloud has global management in core cloud for the management of distributed cloud resources and cloud service with autonomous resource management for edge cloud. Edge cloud can manage its own resources and services when the global management system of distributed cloud is not applicable.

6.3 Configuration models of distributed cloud

The cloud computing needs to be distributed to nearby CSCs to support characteristics of distributed cloud. Depending on a CSC's requirements and characteristics of distributed cloud, cloud services are deployed to one or more cloud among edge, regional and core clouds with network functions. For each cloud service with the CSC's requirements, network resources are allocated by network slicing which allows multiple logical networks to run on top of a shared physical network infrastructure.

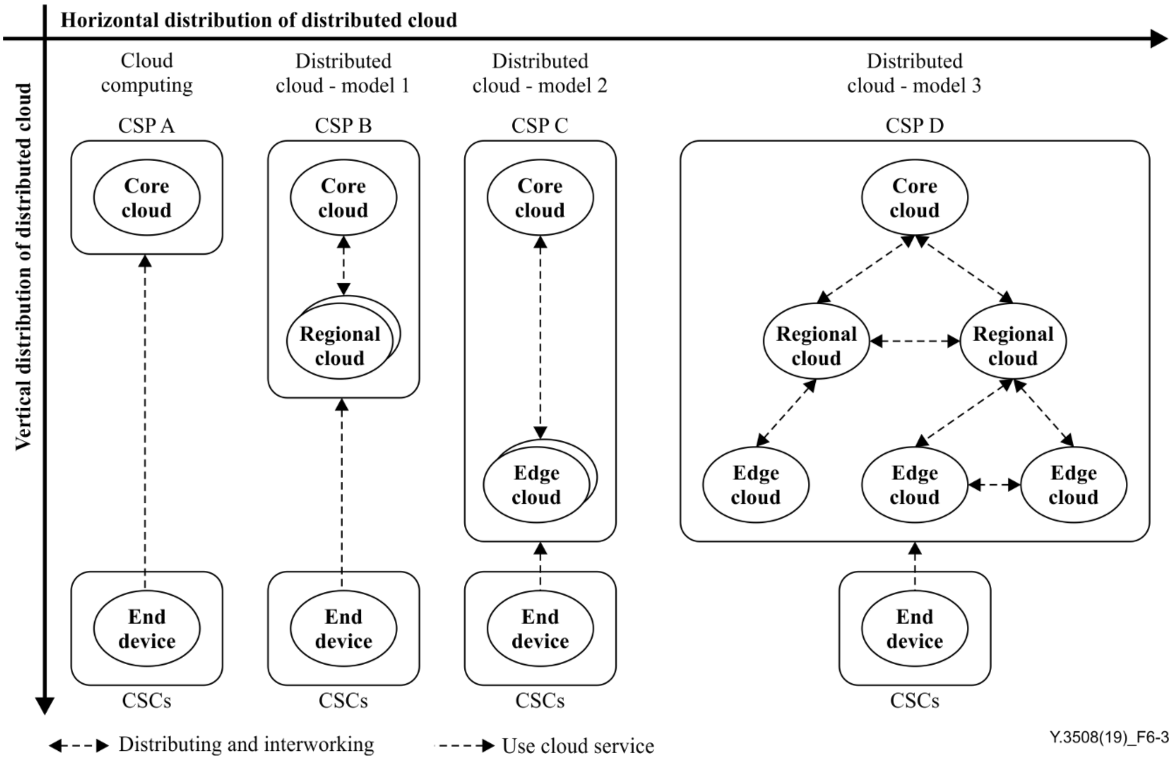


Figure 6-3 – Edge, regional, core cloud configuration in distributed cloud

Table 6-1 describes the generic model of distributed cloud by composing two or more clouds among core, regional and edge cloud. The end device in the figure represents CSCs' devices such as mobile phones, personal computers (PCs), laptops, etc.

Table 6-1 – Distributed cloud configuration model

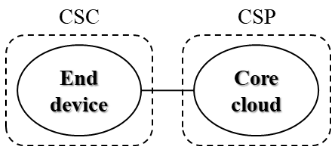
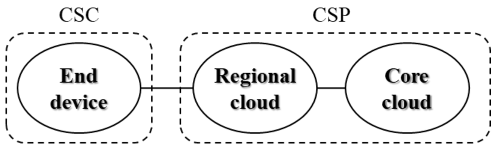
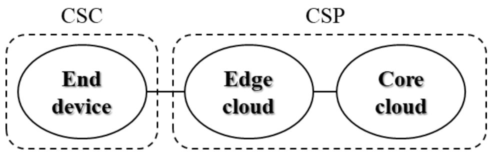
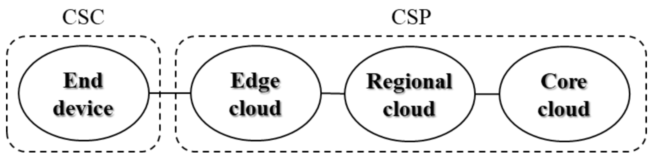
Model	Description
Model 0	 Model 0 is a shape of cloud computing.

Table 6-1 – Distributed cloud configuration model

Model	Description
Model 1 (Regional – Core)	 Model 1 is a shape of a distributed cloud in which the core cloud and the regional cloud are configured together. Cloud services are provided from one or more clouds among the regional cloud and the core cloud. The regional cloud is hosted for a cloud service from the core cloud to a particular region to support localization of cloud services. The regional cloud assists the core cloud by mitigating traffic of cloud service.
Model 2 (Edge – Core)	 Model 2 is a shape of a distributed cloud in which the edge cloud and the core cloud are configured together. A cloud service is functionally reconfigured and is executed on both edge and core clouds by interworking. The edge cloud is configured for real-time service delivery. The edge cloud assists end devices by offloading cloud services or caching the data of cloud services.
Model 3 (Edge – Regional – Core)	 Model 3 is a shape of a distributed cloud in which edge, regional and core clouds are configured together. A cloud service is reconfigured and deployed to edge, regional and core clouds and is executed by interworking. This model is a combination of model 1 and model 2. This model provides cloud services efficiently by mitigating the traffic of the cloud service, caching data of the cloud service and by offloading cloud services for the CSC's end devices.

6.4 Deployment considerations of distributed cloud

This clause provides consideration aspects for the deployment of cloud services on a distributed cloud. These consideration aspects include:

– infrastructure considerations of distributed cloud;

– network considerations of distributed cloud;

– service considerations of distributed cloud;

– management considerations of distributed cloud; and

– security considerations of distributed cloud.

6.4.1 Infrastructure considerations of distributed cloud

Infrastructure in a distributed cloud is heterogeneous and geographically distributed. Thus, consideration aspects are as follows:

– variety of physical resources, which have different hardware structure and instruction, for pooling the resources and deploying cloud services execution efficiently;
– energy efficient resources and power consumption control for edge cloud with small resource capacity; and
– infrastructure automatic provision, configuration and updating, as well as monitoring functionalities for providing a single system view to CSCs.

6.4.2 Network considerations of distributed cloud

The network in a distributed cloud supports connecting distributed cloud resources in broad areas among core, regional and edge clouds and is aware of cloud service. Thus, network considerations are as follows:

– broad network connectivity of distributed cloud resources among edge, regional and core clouds for managing resources and providing cloud service efficiently; and
– deployment of virtual network functions to edge cloud for controlling network on-demand and being aware context of cloud service.

6.4.3 Service considerations of distributed cloud

Services in a distributed cloud have various types based on the resource capacity of core, regional and edge clouds and on the CSC's demands. Thus, service considerations are as follows:

– the scale of a cloud service for deploying cloud services in different-scale edge, regional and core clouds; and
– automatic cloud service delivery for supporting the mobility of CSC’s rapidly.

6.4.4 Management considerations of distributed cloud

Management in a distributed cloud has a global view and a local view of operations, and administration for distributed cloud resources and services. Thus, management considerations are as follows:

– global management of both distributed cloud resources and cloud services for interworking among core, regional and edge clouds; and
– provisioning, operation and administration of distributed cloud resources and cloud services.

6.4.5 Security considerations of distributed cloud

Security in a distributed cloud has adaptive security control for core, regional and edge clouds. Thus, security considerations are as follows:

– identification and authentication of CSCs using multiple distributed cloud infrastructures at any distributed cloud;
– different scales of security functions, such as availability and data integrity, based on the resource capacity of core, regional and edge clouds; and
– different security suites for confidentiality, policy and key management appropriate to heterogeneous distributed cloud resources.

7 High-level requirements for distributed cloud

7.1 Infrastructure requirements for distributed cloud

– **high scalability**: It is required that a distributed cloud provide distributed cloud resources with scale-out to edge, regional and core clouds to satisfy a large number of demands for cloud services;

– **high reliability**: It is required that a distributed cloud provide redundant deployment of cloud services among edge, regional and core clouds against catastrophic events;

– **low power consumption**: It is required that a distributed cloud provide energy-efficient distributed cloud resources for low power consumption at an edge cloud in order to save power of the edge cloud;

– **adaptive resource composition**: It is recommended that a distributed cloud provide adaptive resource composition of the distributed cloud among edge, regional and core clouds;

NOTE 1 – Adaptive resource composition refers to preparation of distributed cloud resources appropriately to support changes of a CSC's condition, such as environmental changes by the CSC's mobility.

– **automatic infrastructure provision**: It is recommended that a distributed cloud provide automatic setup of infrastructure, including hardware initiation, network configuration and software installation;

– **heterogeneous devices access**: It is recommended that a distributed cloud provide heterogeneous access mechanisms for end devices to enrich cloud services.

NOTE 2 – Examples of heterogeneous access mechanisms are Wi-Fi, Bluetooth, wireless sensor network (WSN) and radio access network (RAN).

7.2 Network requirements for distributed cloud

– **low latency**: It is required that a distributed cloud provide network resources to support low latency for cloud services in edge clouds;

NOTE 1 – Examples of network resources are access point, switch, router and gateway.

– **on-demand network traffic control**: It is required that a distributed cloud provide on-demand network traffic control to mitigate congested network links among core, regional and edge clouds by reallocating routing paths and allocating bandwidth;

NOTE 2 – The network bandwidth utilization inside a distributed cloud is substantially reduced with on-demand network traffic control.

– **service routing**: It is required that a distributed cloud support service routing for provisioning cloud service flowing to CSCs;

NOTE 3 – Service routing is a unified service supporting platforms built on the distributed service networking (DSN). It supplies the service registration, publication, discovery, triggering and access mechanisms, and enhanced capabilities to optimize the service. [b-ITU-T Y.2085]

– **context awareness**: It is required that a distributed cloud provide discovery of context information for deploying and delivering cloud service requests to the best location among core, regional and edge clouds;

NOTE 4 – Discovery of context information is enabled by network functions.

– **high-speed network connectivity**: It is recommended that a distributed cloud provide high-speed network connectivity to satisfy significantly higher data rates and data transportation among core, regional and edge clouds;

– **reliable network connectivity**: It is recommended that a distributed cloud provide reliable network connectivity for seamless network links with low error rates and fast connection recovery from failures among core, regional and edge clouds;

– **network virtualization**: It is recommended that a distributed cloud support network virtualization for dynamic network deployment;

NOTE 5 – Network virtualization is a technology that enables the creation of logically isolated network partitions over shared physical networks so that heterogeneous collections of multiple virtual networks can simultaneously coexist over the shared networks. This includes the aggregation of multiple resources in a provider and appearing as a single resource. [b-ITU-T Y.3011]

– **network services delivery**: It is recommended that a distributed cloud provide network services delivery wherever CSC's use cloud services.

NOTE 6 – Network service delivery includes virtualized network functions of switching, tunnelling, traffic analysis, and security such as load balancer, firewall and intrusion detection.

7.3 Service requirements for distributed cloud

– **service mobility**: It is required that a distributed cloud provide service mobility for supporting service level agreements (SLAs);

NOTE 1 – For service mobility, the CSCs' location detection and delivery of corresponding cloud services to the suitable distributed cloud resource automatically are needed for fast response time.

– **lightweight service**: It is required that a distributed cloud provide lightweight cloud services for low power consumption resources at edge clouds;

– **context information update**: It is required a distributed cloud support the updating of the associated context information periodically or aperiodically;

NOTE 2 – See clause 6.2 for context information.

– **caching of cloud service image**: It is required that a distributed cloud provide caching of cloud service images;

NOTE 3 – Frequently used cloud service images need to be cached in a regional cloud to deploy cloud service quickly to an edge cloud.

– **service migration**: It is required that a distributed cloud provide service migration from one distributed cloud resource to another one;

NOTE 4 – Service migration refers to running cloud services reallocated to another distributed cloud resource by taking a snapshot.

– **agile service**: It is recommended that a distributed cloud provide rapid development and delivery of cloud services to distributed cloud resources to support various CSC's requests;

– **automatic service displacement**: It is recommended that a distributed cloud provide cloud service displacement to deploy, maintain and upgrade automatically based on SLAs;

– **autonomous cloud service provisioning**: It is required a distributed cloud provide autonomous cloud service provisioning to initially launch the cloud service by using context information for fast provisioning of distributed cloud resources;

NOTE 5 – Autonomous cloud service provisioning refers to deploying and providing cloud services on edge clouds without intervention of global management in the core cloud by using context information in edge or regional clouds.

– **massive data handling on edge cloud**: It is recommended that a distributed cloud provide massive data handling on edge clouds to reduce the interaction between core/regional clouds and edge clouds.

NOTE 6 – Data handling on edge clouds includes filtering and delivery process to prevent noise from multiple end devices, and batched data delivery to core or regional clouds, respectively.

7.4 Management requirements for distributed cloud

– **global management**: It is required that a distributed cloud provide global management of distributed cloud resources collected from core, regional and edge clouds;

NOTE 1 – Core cloud aggregates information of distributed cloud resources, has global knowledge and provides cloud services appropriately to meet CSC's requirements.

– **real-time service management**: It is required that a distributed cloud provide management for real-time cloud services by finding feasible distributed cloud resources, predicting cloud service latency and scheduling cloud services;

– **heterogeneous resource management**: It is required that a distributed cloud provide management of heterogeneous distributed cloud resources to accommodate and utilize distributed cloud resources easily and efficiently;

NOTE 2 – Examples of heterogeneous distributed cloud resources includes different capacity servers, different types (block, object, file) of storages and different speeds of network devices.

– **end-device connectivity management**: It is required that a distributed cloud provide management of network connectivity to end devices supporting various access protocols to control network latency and bandwidth;

– **context information management**: It is required that a distributed cloud provide management of context information to support service routing and context awareness of the distributed cloud;

– **synchronization management**: It is required that a distributed cloud provide synchronization management for the status of cloud services and distributed cloud resources to perform global management accurately.

7.5 Security requirements for distributed cloud

– **resource isolation and protection**: It is recommended that a distributed cloud provide isolation and protection of distributed cloud resources and cloud services for tenants to guarantee security;

– **adaptive security function control**: It is recommended that a distributed cloud provide adaptive security functions to meet a CSC's requirements with different resource capacity among core, regional and edge clouds;

NOTE 1 – Examples of security function include encryption, decryption, key management, authentication and policy management.

– **adaptive security suites control**: It is recommended that a distributed cloud provide adaptive security suites to apply the suites on different resource capacities among core, regional and edge clouds.

NOTE 2 – Examples of security suites in case of encryption include data encryption standard (DES), advanced encryption standard (AES), international data encryption algorithm (IDEA) and Rivest, Shamir, Adleman (RSA).

8 Security considerations related to ITU-T Recommendations

It is recommended that the security framework for cloud computing described in [b-ITU-T X.1601] be considered for the distributed cloud. [b-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 [b-ITU-T X.1642] be considered for the distributed cloud. [b-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 [b-ITU-T X.1361] be considered for edge clouds in the distributed cloud. In particular, [b-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

General use cases for distributed cloud

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

This appendix identifies use cases of distributed cloud. The table below shows the template used for the description of the use cases.

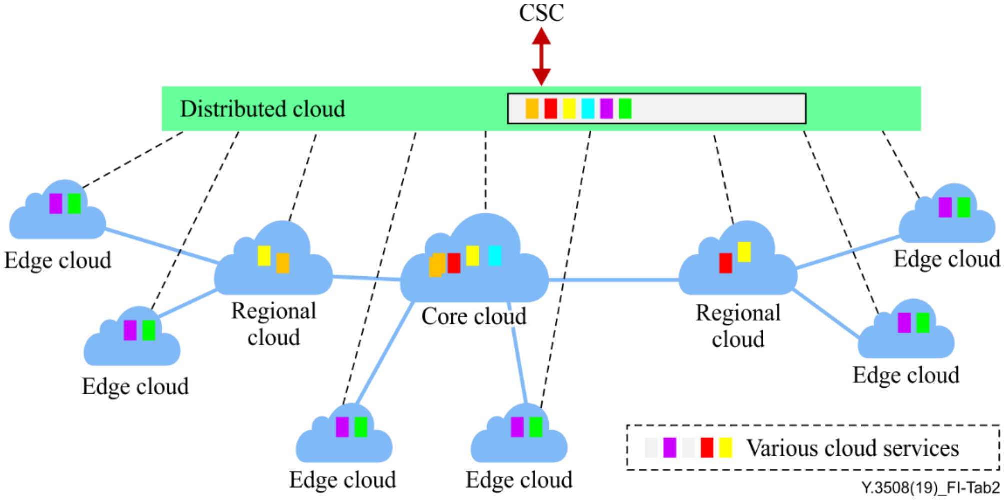
I.1 Use case template

Table I.1 – Template for the description of a use case

Use case	
Name	Title of use case.
Abstract	Overview and features of use case.
Figure	Figure to present the use case. (A unified modeling language (UML)-like diagram is suggested for clarifying relations between roles).
Pre-conditions (optional)	Pre-conditions represent the necessary conditions or use cases that should be achieved before starting the described use case. (Note)
Post-conditions (optional)	As the same for pre-condition, the post-condition describes conditions or use cases that will be carried out after the termination of a currently described use case.
Requirements	The title of requirements derived from the use case. For example: – Large-scale migration.
NOTE – As dependency may exist among different use cases, pre-conditions and post-conditions are introduced to help understand the relationships among use cases.	

I.2 General use case

Table I.2 – General use case of distributed cloud

Use case	
Name	General use case of distributed cloud.
Description	The use case describes the scenarios for the CSC to use a published cloud service. The cloud service has already been published by a CSP and can be browsed through the product catalogue provided by the CSP. The CSC subscribes the cloud service with various requirements, for example, different requirements for service capacity and network latency. The distributed cloud realizes service provisioning to provide cloud service on appropriate core, regional, or edge cloud, according to the CSC's subscription.
Figure	
Pre-conditions (optional)	
Post-conditions (optional)	– The used service should be kept available during the whole invocation. – The SLA should be met for CSC.
Requirements	– High scalability (See clause 7.1) – High reliability (See clause 7.1) – Low power consumption (See clause 7.1) – Low latency (See clause 7.2) – On-demand network traffic control (See clause 7.2)

Appendix II

Use cases for configuration of distributed cloud

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

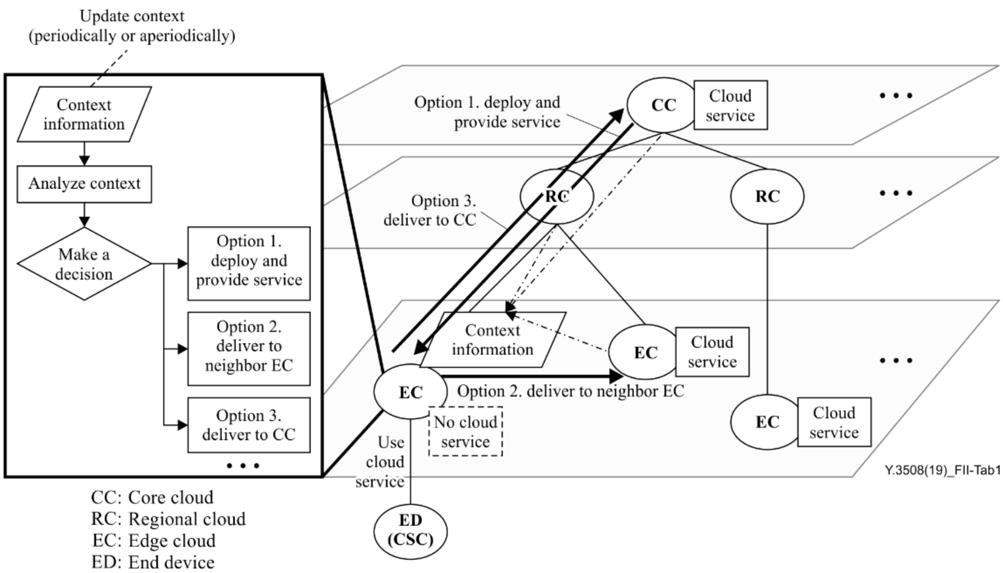
This appendix identifies use cases for configuration of distributed cloud.

II.1 Autonomous cloud service provisioning on distributed cloud

Table II.1 – Autonomous cloud service provisioning on distributed cloud

Use case	
Name	Autonomous cloud service provisioning on distributed cloud.
Abstract	This use case describes the autonomous feature for provisioning a cloud service on distributed cloud. The distributed cloud supports dynamic service provisioning by utilizing core cloud (CC), regional cloud (RC) and edge cloud (EC). EC mainly supports low latency to provide a cloud service. For EC, it is sometimes hard to provide the cloud service due to various reasons (e.g., limited capacity of EC, initially requested service, energy saving). In the case where EC is not ready to provide the corresponding cloud service, EC is required to make a decision to provide the service to the CSC as rapidly as possible. For fast decisions required by time sensitive services in EC, a request to a centralized point (e.g., core cloud and regional cloud) is not appropriate due to network traversal time; the request should be handled in the EC itself. To support autonomous cloud service provisioning on EC, associated context information is required. Associated context information is necessary to be updated on EC periodically and aperiodically. To avoid synchronization burden, associated context information is summarized and updated (e.g., calculating average, grouping value, choosing representative value). To avoid frequent transfer data to the core/regional cloud by edge cloud, the edge cloud filters and deliver to prevent noise from multiple end devices and deliver batched data to core or regional cloud. In conclusion, autonomous provisioning is a main feature that differentiates conventional cloud and distributed cloud. Through the feature, CC, RC and EC on distributed cloud are loosely coupled. The feature also supports fast cloud service provisioning on the distributed cloud. Autonomous provisioning is done by associated context information updated periodically and aperiodically. NOTE – Examples of decision: 1) Deploying the cloud service and provide the cloud service on the EC, 2) Delivering the request to neighbour EC, 3) Delivering the request to the CC, etc.

Table II.1 – Autonomous cloud service provisioning on distributed cloud

<table>
<tr><th colspan="2">Use case</th></tr>
<tr><td>Figure</td><td>


</td></tr>
<tr><td>Pre-conditions (optional)</td><td>Associated context information is updated periodically or aperiodically.</td></tr>
<tr><td>Post-conditions (optional)</td><td></td></tr>
<tr><td>Requirements</td><td>– Adaptive resource composition (See clause 7.1)
– Context information update (See clause 7.3)
– Autonomous cloud service provisioning (See clause 7.3)
– Massive data handling on edge cloud (See clause 7.3)
– Context information management (See clause 7.4)</td></tr>
</table>

II.2 Customer-oriented cloud service provisioning on distributed cloud

Table II.2 – Customer-oriented cloud service provisioning on distributed cloud

<table>
<tr><th colspan="2">Use case</th></tr>
<tr><td>Name</td><td>Customer-oriented cloud service provisioning on distributed cloud.</td></tr>
<tr><td>Abstract</td><td>This use case describes customer-oriented service deployment on distributed cloud.
A main benefit of distributed cloud is provisioning of a cloud service in close proximity to CSC. For the provisioning, methods to support customer-oriented cloud service provisioning on distributed cloud are needed and is called service routing. The service routing includes packet routing, service awareness, etc. By service awareness on distributed cloud, a cloud service is deployed on a cloud in close proximity to the CSC or the request of a cloud service is delivered to another cloud. Accordingly, the service routing on distributed cloud enables:
– scaling or migrating a cloud service for supporting mobility as shown in figure <An example of scaling or migrating a cloud service for supporting mobility>;</td></tr>
</table>

Table II.2 – Customer-oriented cloud service provisioning on distributed cloud

Use case	
	– deploying and providing a cloud service for low response time as shown in figure < An example of deploying and providing a cloud service for low response time>; and – delivering request of a cloud service to neighbour edge cloud as shown in figure <An example of delivering a cloud service to neighbour edge cloud>. Finally, customer-oriented service provisioning (i.e., service routing) method is highly required on distributed cloud in order to improve cloud service QoS (e.g., response time, network bandwidth, mobility). NOTE – Provisioning of a cloud service in close proximity to cloud service customer depends on condition of CSC (e.g., location, QoS level) and condition of surrounding CSC (e.g., availability of edge cloud, capacity of edge cloud, network congestion).

Table II.2 – Customer-oriented cloud service provisioning on distributed cloud

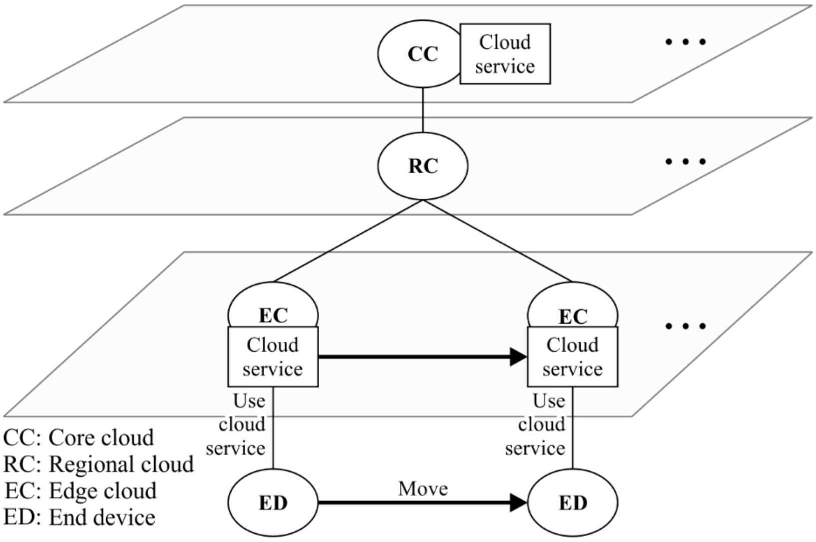
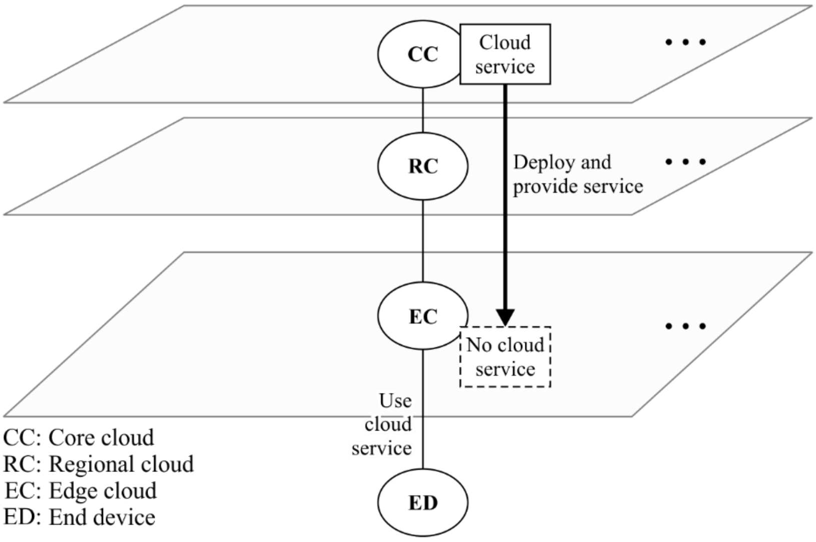
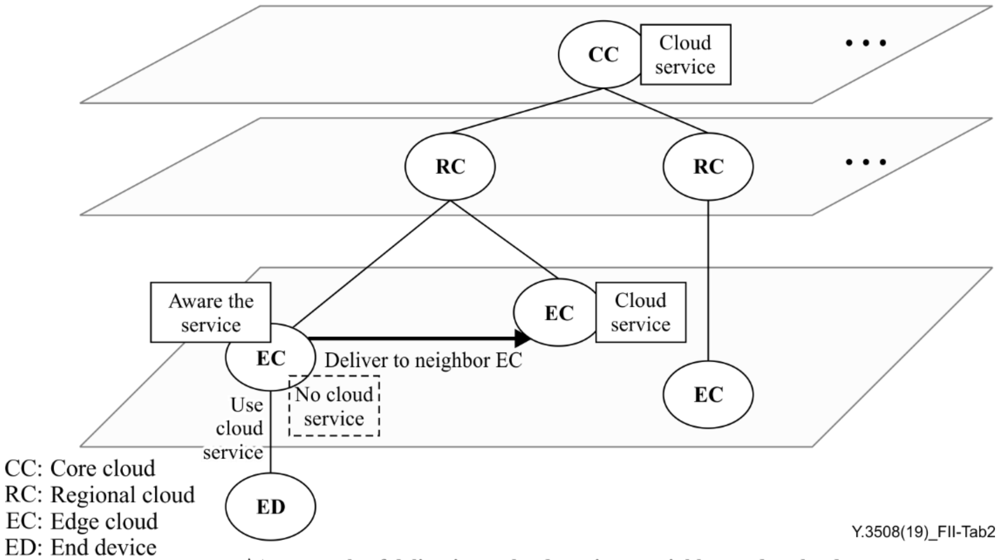
Use case	
Figure	 **An example of scaling or migrating a cloud service for supporting mobility**  **An example of deploying or providing a cloud service for low response time**  **✱An example of delivering a cloud service to neighbour edge cloud**

Table II.2 – Customer-oriented cloud service provisioning on distributed cloud

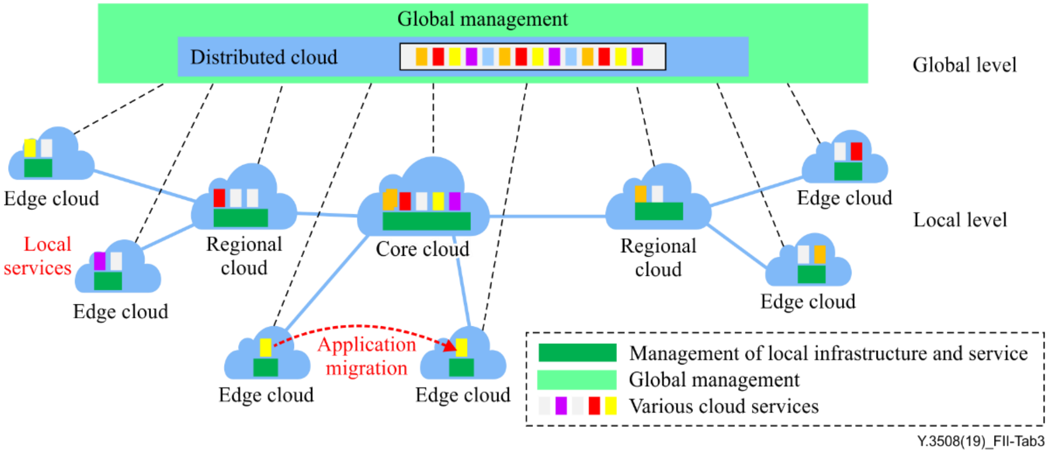
Use case	
Pre-conditions (optional)	
Post-conditions (optional)	
Requirements	– Adaptive resource composition (See clause 7.1) – Service routing (See clause 7.2) – Network services delivery (See clause 7.2) – Context awareness (See clause 7.2) – Context information management (See clause 7.4)

II.3 Distributed cloud infrastructure and service management

Table II.3 – Distributed cloud infrastructure and service management

Use case	
Name	Distributed cloud infrastructure and service management
Description	This use case describes the distributed cloud infrastructure and service management. Distributed cloud infrastructure includes processing, storage, networking and other hardware resources as well as software assets. Distributed cloud infrastructure management is the management of hardware resources and software assets of distributed cloud. Distributed cloud includes core, regional and edge clouds, which are different in size, hardware types and software. For example, a distributed cloud resource may use a graphics processing unit (GPU) or field-programmable gate array (FPGA) hardware for AR/VR scenarios and microprocessor-based server instead of x86 server for some specified application. These resources are distributed across various environments. Thus, resources need to be isolated and protected in secure manner. Distributed cloud provides flexible and agility network connectivity for core, regional and edge clouds. The network connection can be changed on demand or based on policy, such as active-standby policy. The ultimate objective of distributed cloud is to provide distributed cloud services for the customers. The distributed cloud encapsulates the capabilities of distributed cloud infrastructure into various cloud services for the end customers. Distributed cloud is a hierarchical system, in which each core cloud, regional cloud and edge cloud could manage its local infrastructure and service, and the infrastructure and service of the whole distributed cloud could be unified managed and coordinated by the distributed cloud management.

Table II.3 – Distributed cloud infrastructure and service management

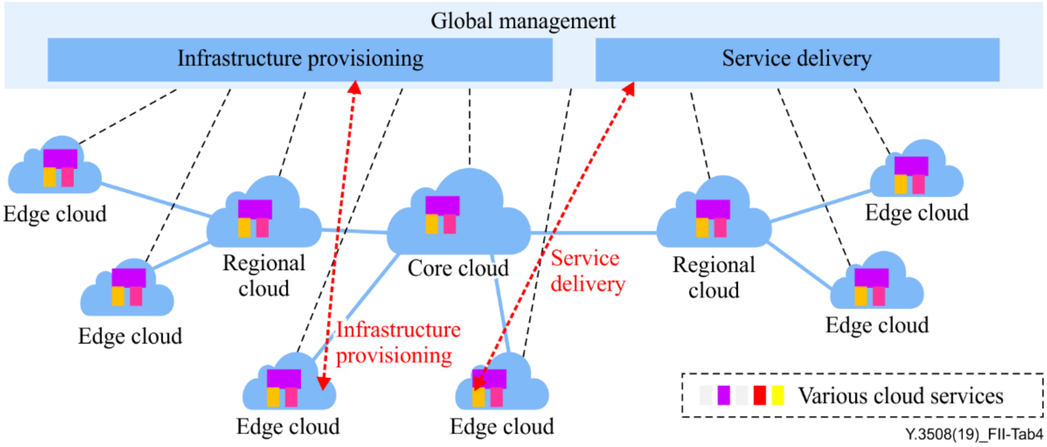
<table>
<tr><th colspan="2">Use case</th></tr>
<tr><td>Figure</td><td>
</td></tr>
<tr><td>Pre-conditions (optional)</td><td></td></tr>
<tr><td>Post-conditions (optional)</td><td></td></tr>
<tr><td>Requirements</td><td>– Heterogeneous device access (See clause 7.1)
– Global management (See clause 7.4)
– Heterogeneous resource management (See clause 7.4)
– End device connectivity management (See clause 7.4)
– Resource isolation and protection (See clause 7.5)</td></tr>
</table>

II.4 Distributed cloud infrastructure and service provisioning

Table II.4 – Distributed cloud infrastructure and service provisioning

<table>
<tr><th colspan="2">Use case</th></tr>
<tr><td>Name</td><td>Distributed cloud infrastructure and service provisioning.</td></tr>
<tr><td>Description</td><td>This use case describes the distributed cloud infrastructure and service provisioning.
Distributed cloud can provision infrastructure and deliver service among core, regional and edge cloud.
The distributed cloud infrastructure contains various hardware and software, including servers, storage, network devices and software. The edge cloud especially is resource-constrained due to space and power. The infrastructure provision process includes hardware initiation, network configuration and software installation. After provisioning, the services including security functions and suites can be provided based on the infrastructure. Also, if a cloud service needs to use an edge cloud resource, it is reconfigured or developed to run on the edge cloud with small capacity and heterogeneous resources.
The distributed cloud infrastructure can be maintained and upgraded by distributed cloud management. The provisioning is a policy-based process so it can be done automatically on distributed cloud resources.
Once the distributed cloud infrastructure has been provisioned, the cloud service should be deployed and delivered rapidly to the suitable distributed cloud resource by high-speed and reliable network connectivity based on service's SLA, such as required resources, maximum latency, required or useful services, etc. The deployed service can be maintained and upgraded based on global management. For accurate global management, status information of cloud service and distributed cloud resource are synchronized from edge/regional cloud to core cloud.</td></tr>
</table>

Table II.4 – Distributed cloud infrastructure and service provisioning

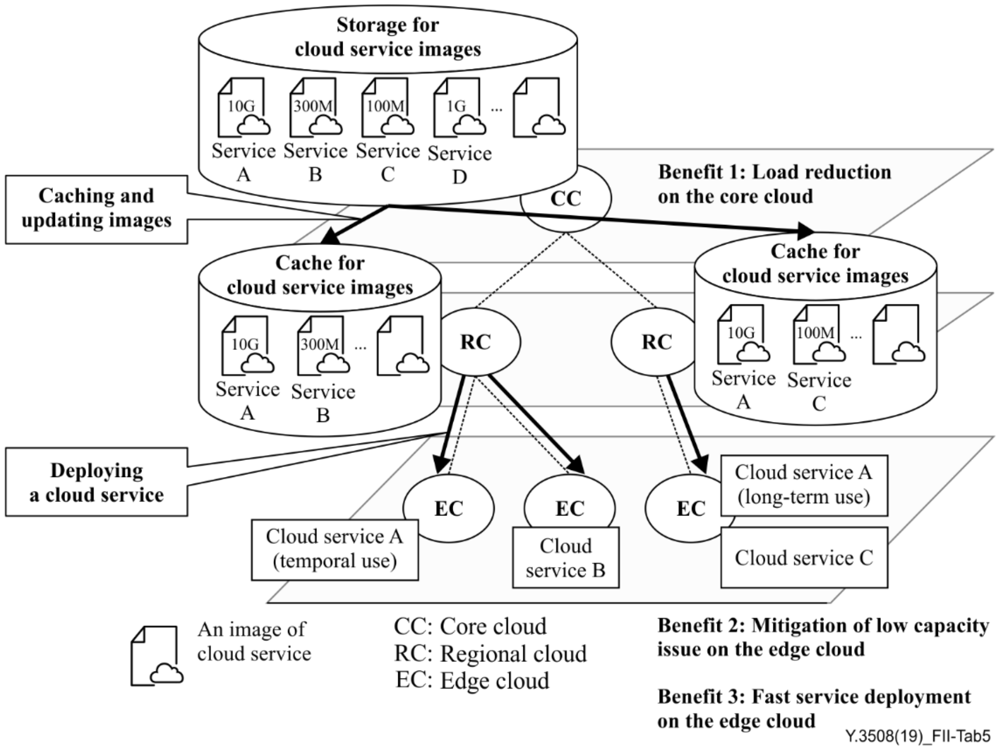
Use case	
Figure	
Pre-conditions (optional)	
Post-conditions (optional)	
Requirements	– Automatic infrastructure provision (See clause 7.1) – High speed network connectivity (See clause 7.2) – Reliable network connectivity (See clause 7.2) – Network virtualization (See clause 7.2) – Network services delivery (See clause 7.2) – Agile service (See clause 7.3) – Lightweight service (See clause 7.3) – Automatic service displacement (See clause 7.3) – Synchronization management (See clause 7.4) – Adaptive security function control (See clause 7.5) – Adaptive security suites control (See clause 7.5)

II.5 Hierarchical caching of cloud service images

Table II.5 – Hierarchical caching of cloud service images

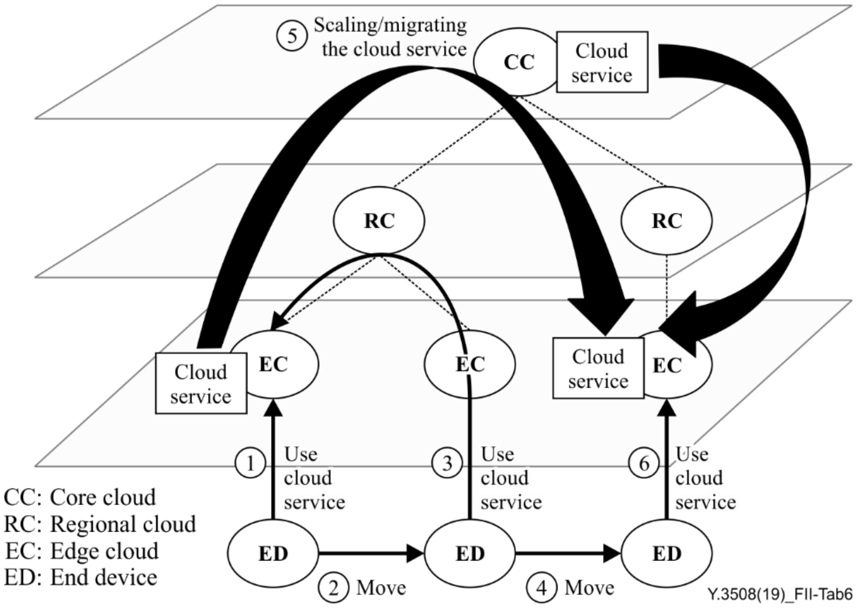
Use case	
Name	Hierarchical caching of cloud service images.
Abstract	This use case shows an example of hierarchical caching of cloud service images. The core cloud mainly manages cloud service images; the size of a cloud service image varies from tens of megabytes to hundreds of gigabytes. A cloud service image is needed to run a cloud service. For example, when a cloud service is run on the edge cloud, the image is copied from the core cloud by request. Thus, depending on the number of requests, the core cloud can be overloaded. In addition, copying a cloud service image from the core cloud to the edge cloud is inefficient. Also, the edge cloud has relatively lower storage capacity than the core cloud or the regional cloud. It is inefficient for the edge cloud to maintain all cloud service images.

Table II.5 – Hierarchical caching of cloud service images

Use case	
	Therefore, hierarchical caching is needed. Cloud service images are cached on the regional cloud according to the initial request or frequency. It is helpful to reduce the load on the core cloud because multiple regional clouds handle the requests respectively. It also mitigates the capacity issue on the edge cloud, and it provides fast service deployment to support real-time performance by reducing latency than from the core cloud.
Figure	
Pre-conditions (optional)	
Post-conditions (optional)	
Requirements	– High speed network connectivity (See clause 7.2) – Caching of cloud service image (See clause 7.3) – Real-time service management (See clause 7.4)

II.6 High mobility support on distributed cloud

Table II.6 – High mobility support on distributed cloud

Use case	
Name	High mobility support on distributed cloud.
Abstract	This use case illustrates an example of high mobility support on distributed cloud. A typical mobility support environment follows a method of modifying the packet routing path while the location of the service is fixed. On the distributed cloud, the location of the service is different from the conventional cloud and is located near the user. Accordingly, the distributed cloud needs high mobility. As an example, the below figure shows two ways of mobility while an end device moves with different type of access. One way is that the regional cloud routes the packet to the previous edge cloud as shown in step 3 in the figure below. (It is also possible to route the cloud service to the current edge cloud.) Another way is that the core cloud routes the cloud service to the edge cloud near the end device by scaling or migrating the cloud service with high-speed and reliable network connectivity as shown in step 5 and 6. The cloud service is from the previous edge cloud or the core cloud. Therefore, to efficiently support high mobility on the distributed cloud, both ways are selected based on service level agreement (SLA) for the cloud service. SLA includes latency, quality of service, etc.
Figure	
Pre-conditions (optional)	
Post-conditions (optional)	
Requirements	– Heterogeneous device access (See clause 7.1) – High speed network connectivity (See clause 7.2) – Reliable network connectivity (See clause 7.2) – Context awareness (See clause 7.2) – Service mobility (See clause 7.3) – Service migration (See clause 7.3) – Context information management (See clause 7.4)

Appendix III

A comparison of distributed cloud with related technology

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

This Appendix identifies a comparison of the distributed cloud, edge computing, fog computing, multi-access edge computing and cloudlet. There are various technologies related to the edge of the network, such as edge computing, fog computing, cloudlet and multi-access edge computing. The edge-related technologies fully or partially utilize cloud computing technology. In the perspective of multi-cloud, distributed cloud extends a cloud to the edge of the network. Accordingly, the distributed cloud and the edge-related technologies definitely have a common point. Thus, it is necessary to compare key features and show the coverage of them among the distributed cloud and the edge related technologies. In this Appendix, the coverage of key features among the distributed cloud and the edge-related technologies are analysed and illustrated as follow:

– **Cloud computing** focuses on non-real-time services and supports periodic maintenance and service decision – making while distributed cloud extends the cloud computing capabilities to the edge and supports real-time smart processing and execution of local services. Distributed cloud takes core, regional and edge as a whole and closely interacts with one another. Resource-constrained, unreliable or bandwidth-limited network connections, or applications that demand very high bandwidth, low latency, or widespread compute capacity across many sites are considered by distributed cloud but not involved with cloud computing.

– **Edge computing** is performed on an open platform at the edge of the network near things or data sources, integrating network, computing, storage and application core capabilities and providing edge intelligent services. Edge computing meets the requirements of industrial digitalization for agile connections, real-time services, data optimization, smart application, security and privacy protection. Local devices and their capabilities are a part of edge computing while distributed cloud is not involved with local devices. The major scenario of edge computing focuses on IoT, but distributed cloud has many scenarios beside IoT. Big data analytics is considered as one of the key capabilities for edge computing while distributed cloud treats it as a service which it’s carried on.

• **Fog computing** is a horizontal, physical or virtual resource paradigm that resides between smart end-devices and traditional cloud or data centres. Fog computing supports vertically isolated, latency-sensitive applications by providing ubiquitous, scalable, layered federated and distributed computing, storage and network connectivity. Rather than a substitute, fog computing often serves as a complement to cloud computing. Fog computing is mainly for IoT scenarios while distributed cloud has many scenarios beside IoT. It is similar to edge computing where the local devices are involved for fog computing but the infrastructure of distributed cloud only includes data centres.

– Multi-access edge computing **(MEC)** offers cloud-computing capabilities and an IT service environment at the edge of the network and provides ultra-low latency and high bandwidth as well as real-time access to radio network information that can be leveraged by applications. MEC is mainly for mobile scenarios and is extending to fixed mobile convergence scenarios. The scope of MEC is involved in distributed cloud.

– **Cloudlet** can be viewed as a "data centre in a box", which represents the middle tier of a three-tier hierarchy "mobile or IoT device – cloudlet – cloud", whose goal is to bring the cloud closer. It is mainly for scenarios of convergence of mobile and cloud computing. The scope of cloudlet is involved in distributed cloud.

Bibliography

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

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

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

[b-ITU-T Y.2085] Recommendation ITU-T Y.2085 (2016), *Distributed service networking service routing*.

[b-ITU-T Y.3011] Recommendation ITU-T Y.3011 (2012), *Framework of network virtualization for future networks*.

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, 2019