International Telecommunication Union

ITU-T

TELECOMMUNICATION
STANDARDIZATION SECTOR
OF ITU

Y.3511

(03/2014)

SERIES Y: GLOBAL INFORMATION INFRASTRUCTURE, INTERNET PROTOCOL ASPECTS AND NEXT-GENERATION NETWORKS

Cloud Computing

Framework of inter-cloud computing

Recommendation ITU-T Y.3511

ITU-T Y-SERIES RECOMMENDATIONS

GLOBAL INFORMATION INFRASTRUCTURE, INTERNET PROTOCOL ASPECTS AND NEXT-GENERATION NETWORKS

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**

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

Recommendation ITU-T Y.3511

Framework of inter-cloud computing

Summary

Recommendation ITU-T Y.3511 describes the framework for interactions of multiple cloud service providers (CSPs), which is referred to as inter-cloud computing. Based on several use case, and after considering the different types of service offerings, this Recommendation describes the possible relationships (peering, federation or intermediary) among multiple CSPs. By introducing the concept of primary CSP and secondary CSP, the Recommendation further describes CSP interactions for the cases of federation and intermediary patterns. Finally, relevant functional requirements are derived.

History

Edition	Recommendation	Approval	Study Group	Unique ID*
1.0	ITU-T Y.3511	2014-03-09	13	11.1002/1000/12078

Keywords

Cloud computing, infrastructure, inter-cloud computing, network, primary CSP, requirement, secondary CSP, use case.

* 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.3511

Framework of inter-cloud computing

1 Scope

This Recommendation describes the framework for interactions between multiple cloud service providers (CSPs), which is referred to as inter-cloud computing. Based on use cases involving several CSPs and after considering different types of service offerings (given in the appendices), this Recommendation describes the possible relationships among multiple CSPs, interactions between CSPs and relevant functional requirements.

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 X.1601] Recommendation ITU-T X.1601 (2014), *Security framework for cloud computing*.

[ITU-T Y.3501] Recommendation ITU-T Y.3501 (2013), *Cloud computing framework and high-level requirements*.

[ITU-T Y.3520] Recommendation ITU-T Y.3520 (2013), *Cloud computing framework for end to end resource management*.

3 Definitions

3.1 Terms defined elsewhere

This Recommendation uses the following terms defined elsewhere:

3.1.1 cloud service customer [ITU-T Y.3501]: A person or organization that consumes delivered cloud services within a contract with a cloud service provider.

3.1.2 cloud service provider [ITU-T Y.3501]: An organization that provides and maintains delivered cloud services.

3.1.3 resource management [ITU-T Y.3520]: A way to access, control, manage, deploy, schedule and bind resources when they are provided by service providers and requested by customers.

3.1.4 service level agreement [b-ISO/IEC 20000-1:2011]: Documented agreement between the service provider and customer that identifies services and service targets

NOTE 1 – A service level agreement can also be established between the service provider and a supplier, an internal group or a customer acting as a supplier.

NOTE 2 – A service level agreement can be included in a contract or another type of documented agreement.

3.2 Terms defined in this Recommendation

This Recommendation defines the following terms:

3.2.1 inter-cloud computing: The paradigm for enabling the interworking between two or more cloud service providers.

NOTE – Inter-cloud computing is also referred as inter-cloud.

3.2.2 primary cloud service provider: In inter-cloud computing, a cloud service provider which is making use of cloud services of peer cloud service providers (i.e., secondary cloud service providers) as part of its own cloud services.

3.2.3 secondary cloud service provider: In inter-cloud computing, a cloud service provider which provides cloud services to a primary cloud service provider.

NOTE – The primary cloud service provider can use the services of secondary cloud service providers as part of its services offered to cloud service customers.

4 Abbreviations and acronyms

This Recommendation uses the following abbreviations and acronyms:

API	Application Programming Interface
B2B	Business-to-Business
CaaS	Communications as a Service
CDN	Content Distribution Network
CPU	Central Processing Unit
CSC	Cloud Service Customer
CSP	Cloud Service Provider
DRM	Digital Rights Management
IaaS	Infrastructure as a Service
ID	Identifier
IT	Information Technology
LAN	Local Area Network
NaaS	Network as a Service
P-CSP	Primary CSP
PaaS	Platform as a Service
QoS	Quality of Service
S-CSP	Secondary CSP
SaaS	Software as a Service
SDP	Service Delivery Platform
SLA	Service Level Agreement
VM	Virtual Machine
VPN	Virtual Private Network

5 Conventions

In this Recommendation:

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

In the body of this Recommendation and its appendices, the words should, and may sometimes appear, in which case they are to be interpreted, respectively, as is recommended, and can optionally. The appearance of such phrases or keywords in an appendix or in material explicitly marked as informative are to be interpreted as having no normative intent.

6 Introduction

Inter-cloud computing describes the interworking of cloud service providers (CSPs) in order to deliver services for the users. Inter-cloud relationship between CSPs can be realized by using a service from the peer CSP or by providing a service to the peer CSP.

The case where two peer CSPs interact with each other through an inter-cloud relationship is illustrated in Figure 6-1. As shown by a pair of two arrows pointing in opposite directions in Figure 6-1, CSP A is making use of the services provided by CSP B. In this relationship, CSP A is considered as being the primary CSP while CSP B is the secondary CSP. Note that the reverse situation where CSP B is using the services offered by CSP A may also exist and in such case CSP A and CSP B are involved in two inter-cloud relationships, one for providing services to the peer CSP and the other for using the services of the peer CSP.

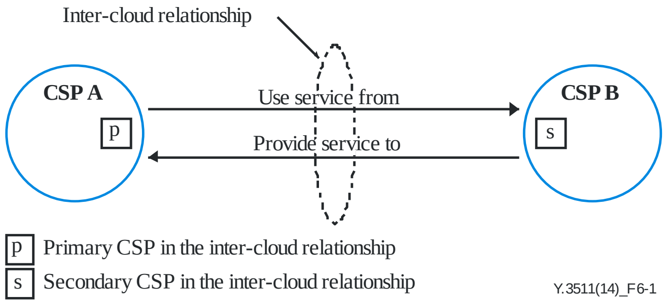

Inter-cloud relationship
CSP A
p
Use service from
Provide service to
CSP B
s
p Primary CSP in the inter-cloud relationship
s Secondary CSP in the inter-cloud relationship
Y.3511(14)_F6-1

Figure 6-1 – Inter-cloud relationship between peer CSPs

Figure 6-2 illustrates in a different way the inter-cloud relationship between CSP A and CSP B, i.e., CSP A uses the service of CSP B through the application programming interface (API) provided by CSP B, shown as API (B). Although the arrow in Figure 6-2 is only one and is shown as unidirectional from CSP A to CSP B, it should be understood as being equivalent to the inter-cloud relationship shown in Figure 6-1, i.e., covering the "use service from" and "provide service to" arrows in Figure 6-1.

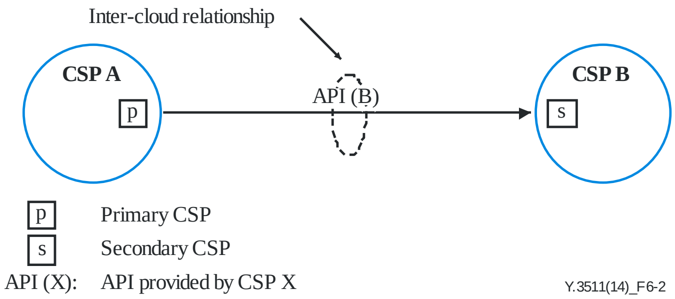


Figure 6-2 – Inter-cloud relationship using API

7 Patterns of inter-cloud

This clause introduces three patterns of inter-cloud for describing relations and interactions involving multiple CSPs, i.e.,:

- The inter-cloud peering;
- The inter-cloud federation;
- The inter-cloud intermediary.

7.1 Inter-cloud peering

In inter-cloud peering, two CSPs interwork directly with each other in order to use the services provided by the peer CSP.

NOTE 1 – Inter-cloud peering does not necessarily imply reciprocal relationships in terms of service use and service providing between the two CSPs.

NOTE 2 – Inter-cloud peering is a fundamental pattern, which may exist on its own or can be used in the two patterns described in clauses 7.2 (inter-cloud federation) and 7.3 (inter-cloud intermediary).

In inter-cloud peering, each CSP exposes its own API for cloud interworking and the CSPs interwork with each other directly by using the other CSP's API. As shown in Figure 7-1, CSP A interworks with CSP B using the API provided by CSP B and vice versa. Since the inter-cloud peering pattern can be used in the other pattern described in clauses 7.2 and 7.3, use of a common API between CSP A and CSP B is not precluded (see Figure 7-2).

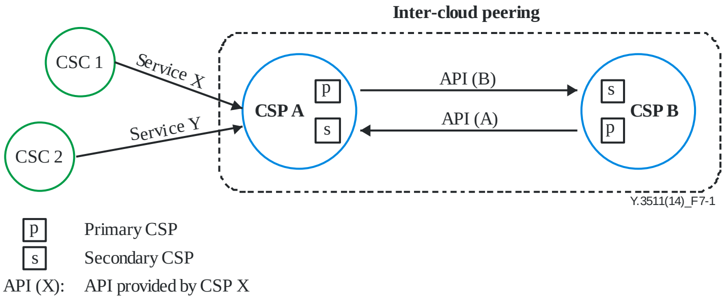


Figure 7-1 – Inter-cloud peering

As shown in Figure 7-1, the inter-cloud peering pattern consists of two inter-cloud relationships, CSP A to CSP B relationship and CSP B to CSP A relationship. CSP A is a primary CSP when using the services of CSP B provided by API (B) for providing services to its own customers (CSC1 and CSC2) and is also a secondary CSP when providing services to CSP B through its own API (A).

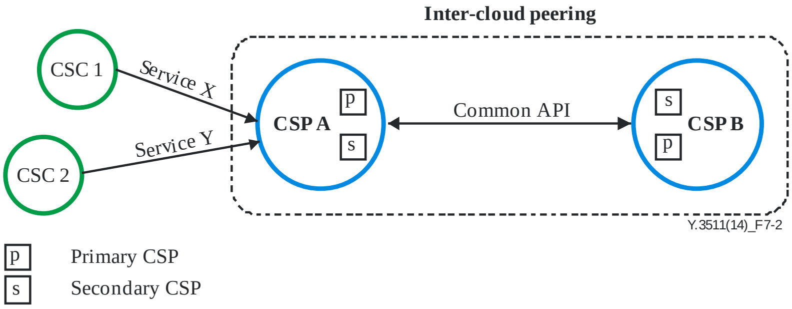


Figure 7-2 – Common API in the peering pattern

Figure 7-2 illustrates the case where a common API is used between CSP A and CSP B, i.e., API (A) and API (B) in Figure 7-1 are the same.

7.2 Inter-cloud federation

Inter-cloud federation involves using the cloud services within a group of peer CSPs who mutually combine their service capabilities in order to provide the set of cloud services required by cloud service customers (CSCs).

The multiple CSPs, which form the inter-cloud federation, establish and share the common agreement which may range from service-related policies, service level agreements (SLAs) and procedures which are relevant to the service offering and resource handling.

Based on the agreement, a CSP in the inter-cloud federation can offer its cloud services with the help of other CSPs.

A common API for cloud interworking is defined in the inter-cloud federation. As shown in Figure 7-3, each CSP interworks with the other CSPs in the inter-cloud federation through this common API.

It should be noted that the inter-cloud federation pattern does not necessarily require the fully meshed configuration of CSPs interacting with each other as shown in Figure 7-3.

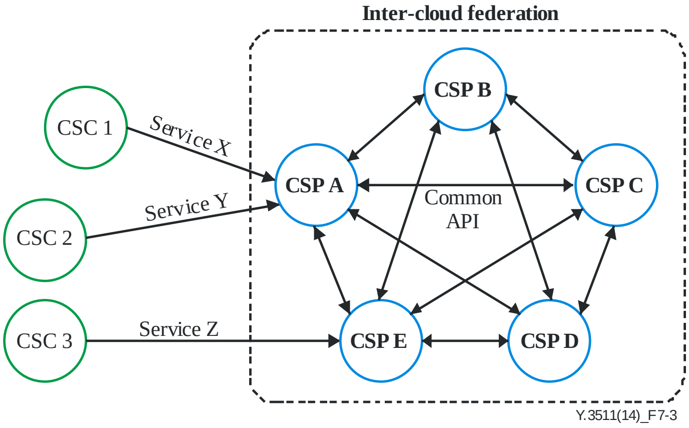


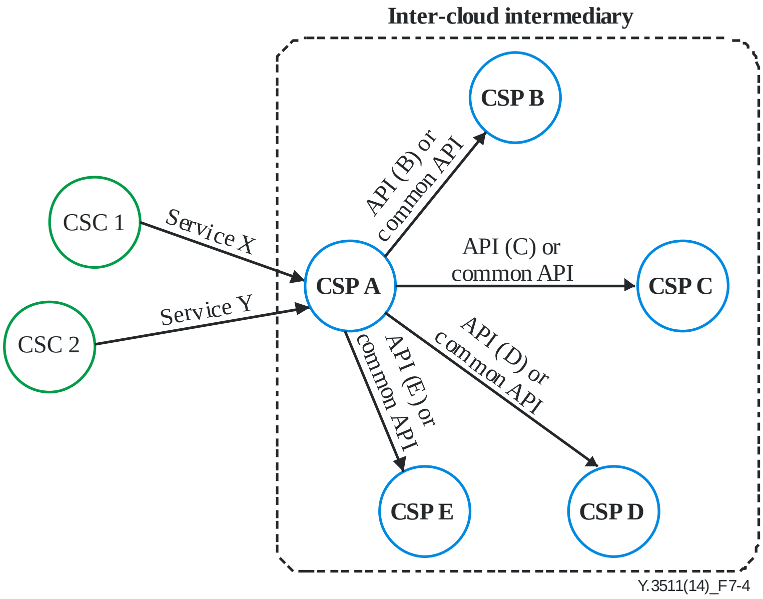
Figure 7-3 – Inter-cloud federation

7.3 Inter-cloud intermediary

In the inter-cloud intermediary pattern, the CSP interworks with one or more peer CSPs and provides intermediation, aggregation and arbitrage of services provided by these CSPs.

Service intermediation consists in conditioning or enhancing the cloud service of a peer CSP. Service aggregation relates to the composition of a set of services provided by the peer CSPs. Service arbitrage is about selecting one service offering from a group of services offered by the peer CSPs.

Interworking between a CSP providing service intermediation, aggregation and arbitrage and the peer CSPs can rely on either the inter-cloud peering pattern or the inter-cloud federation pattern. Figure 7-4 illustrates the inter-cloud intermediary pattern where CSP A provides intermediation, aggregation and arbitrage of cloud services provided by CSPs B, C, D and E.



NOTE – API (X): API provided by cloud service provider X

Figure 7-4 – Inter-cloud intermediary

8 Overview of inter-cloud computing

8.1 Relationship between intra-cloud and inter-cloud handling of resources

For collaboration among CSPs, two types of resources can be distinguished. One includes underlying physical resources of a cloud infrastructure which are controlled and managed by the CSP who owns these resources. The other type includes resources which are abstracted from the underlying physical resources and are offered as services to CSPs. During the collaboration of CSPs, such abstracted resources are also utilized during interactions between these CSPs.

Based on the abstraction, underlying physical resources will become abstracted resources. Detailed information of underlying physical resources, such as total central processing unit (CPU) cores and memories available in the infrastructure, will be hidden. During the collaboration among CSPs, only the information of abstracted resources, such as the CPU cores and memories dedicated to the given service, will be subject for interactions.

Figure 8-1 shows the relationship between intra-cloud and inter-cloud and their handling of resources.

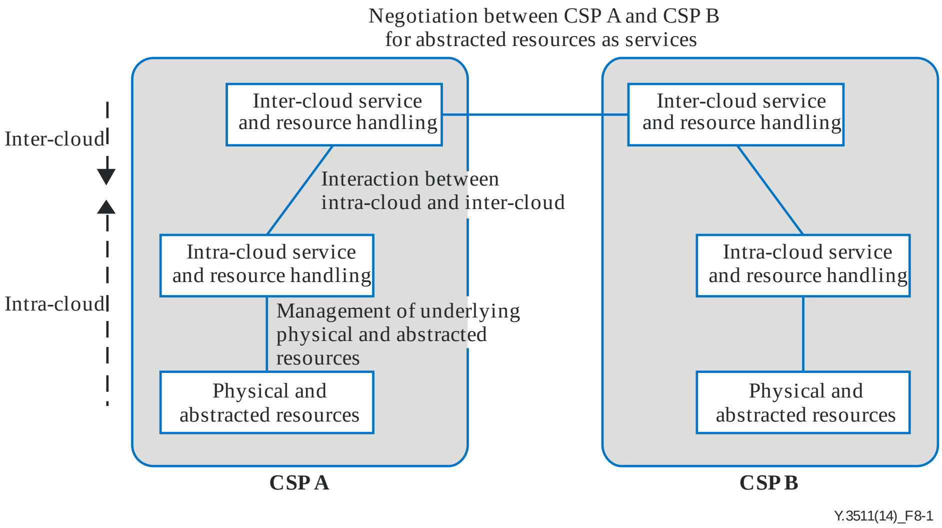

Negotiation between CSP A and CSP B for abstracted resources as services
Inter-cloud
Intra-cloud
Inter-cloud service and resource handling
Interaction between intra-cloud and inter-cloud
Intra-cloud service and resource handling
Management of underlying physical and abstracted resources
Physical and abstracted resources
Inter-cloud service and resource handling
Intra-cloud service and resource handling
Physical and abstracted resources
CSP A
CSP B
Y.3511(14)_F8-1

Figure 8-1 – Intra-cloud and inter-cloud relationship and handling of resources

In Figure 8-1, intra-cloud service and resource handling allows a CSP (A or B) to manage its own resources including underlying physical resources and abstracted resources. Inter-cloud service and resource handling allows a CSP to negotiate for the use of peer CSPs' abstracted resources provided as cloud services.

For example, when CSP A decides to use resources from CSP B, its intra-cloud service and resource handling will interact with its inter-cloud service and resource handling, which will then interact with CSP B. When the inter-cloud service and resource handling of CSP B receives the request from CSP A, the request will be relayed towards its own intra-cloud service and resource handling in order for CSP B to decide whether to provide the service and associated abstracted resources to the CSP A.

8.2 Overview of inter-cloud federation

8.2.1 Introduction

In the inter-cloud federation pattern, a number of CSPs provide services to CSCs. When needed (e.g., in the event of a serious shortage in resource pool), CSPs within the inter-cloud federation utilize other CSPs' resources to provide services to their customers.

Figure 8-2 illustrates the inter-cloud federation pattern.

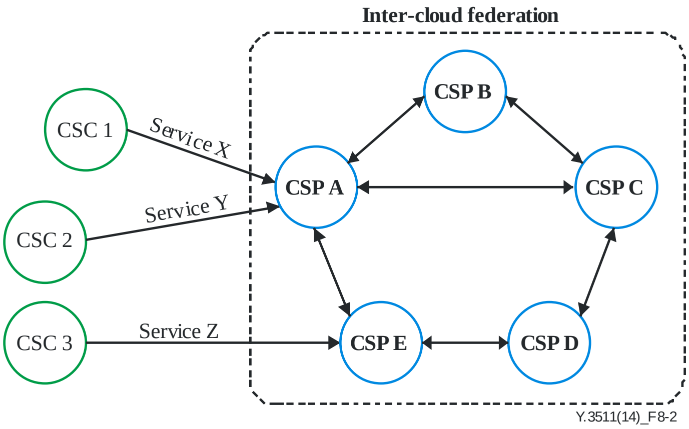


Figure 8-2 – Inter-cloud federation

As shown in Figure 8-2, CSC 1 and CSC 2 use services X and Y provided by CSP A, but the resources used for services X and Y may actually be provided by CSPs B, C or E.

8.2.2 Primary CSP and secondary CSP

In an inter-cloud federation, two or more CSPs interact to provide cloud services to CSCs. The CSP that is responsible for providing the services to a given CSC is called the primary CSP while the peer CSPs in the inter-cloud federation offering their own resources as services to the primary CSP are called secondary CSPs.

When needed, the primary CSP will make request in order to utilize the resources of secondary CSPs. The primary CSP determines which secondary CSPs will actually provide such resources (e.g., in terms of processing power, storage and networks) to the CSC. In some cases, the primary CSP may provide none of its own resources and will have to obtain all resources required for the support of services from secondary CSPs.

The roles of primary CSP and secondary CSP depend on the individual service. For example, in Figures 8-3 and 8-4, CSP A is the primary CSP and CSPs B, C and E are secondary CSPs for services X and Y. For service Z, CSP E is the primary CSP and CSPs A and D are secondary CSPs.

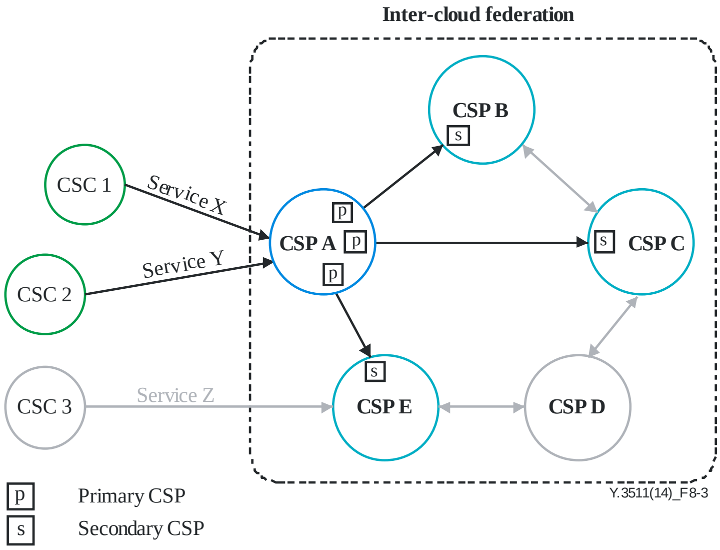


Figure 8-3 – CSP A offering services as the primary CSP with the help of secondary CSPs

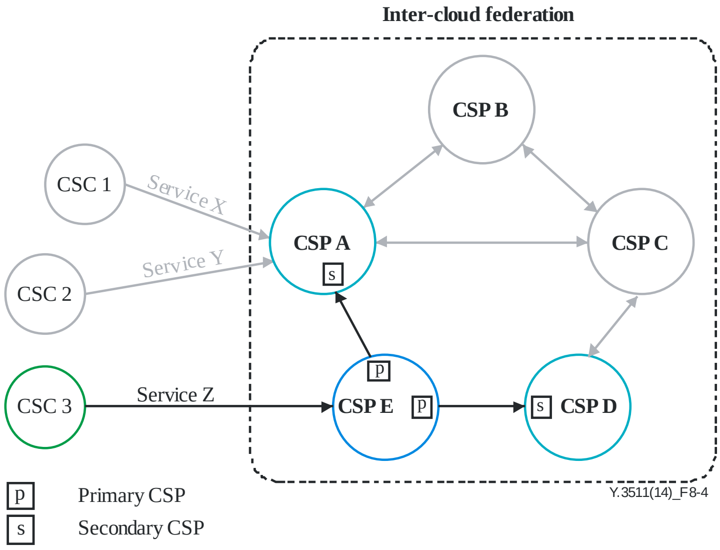


Figure 8-4 – CSP E offering services as the primary CSP with the help of secondary CSPs

A given CSP can act as a primary CSP and secondary CSP, i.e., using the services of secondary CSPs (e.g., CSP A in Figure 8-3) and providing services to a primary CSP (e.g., CSP A in Figure 8-4).

8.2.3 Network connectivity

In order to deliver cloud services based on the use of inter-cloud computing, network connectivity is required among the involved CSCs, primary CSPs and secondary CSPs. Specifically, this includes:

– Connectivity between peer CSPs. By means of this connectivity, a primary CSP can interwork with the secondary CSPs to request a service (e.g., backup of data or transfer of virtual machines between connected CSPs). In some cases, this connectivity is provided on-demand; the connectivity is established instantaneously as the need arises and removed after the need disappears;

– Connectivity between CSCs and CSP. By means of this connectivity, CSCs can use, operate, and manage their cloud services provided by CSPs. CSCs do not know on which CSPs their services actually run, but network connectivity facilitates access of CSCs to the appropriate CSP.

Figure 8-5 illustrates an example of configuration involving a Software as a Service (SaaS) CSP and multiple Infrastructure as a Service (IaaS) CSPs forming an inter-cloud federation. The SaaS CSP uses virtual machines (VMs) provided by IaaS CSPs which are members of the inter-cloud federation in order to provide SaaS services (e.g., applications such as e-commerce) to its SaaS CSCs.

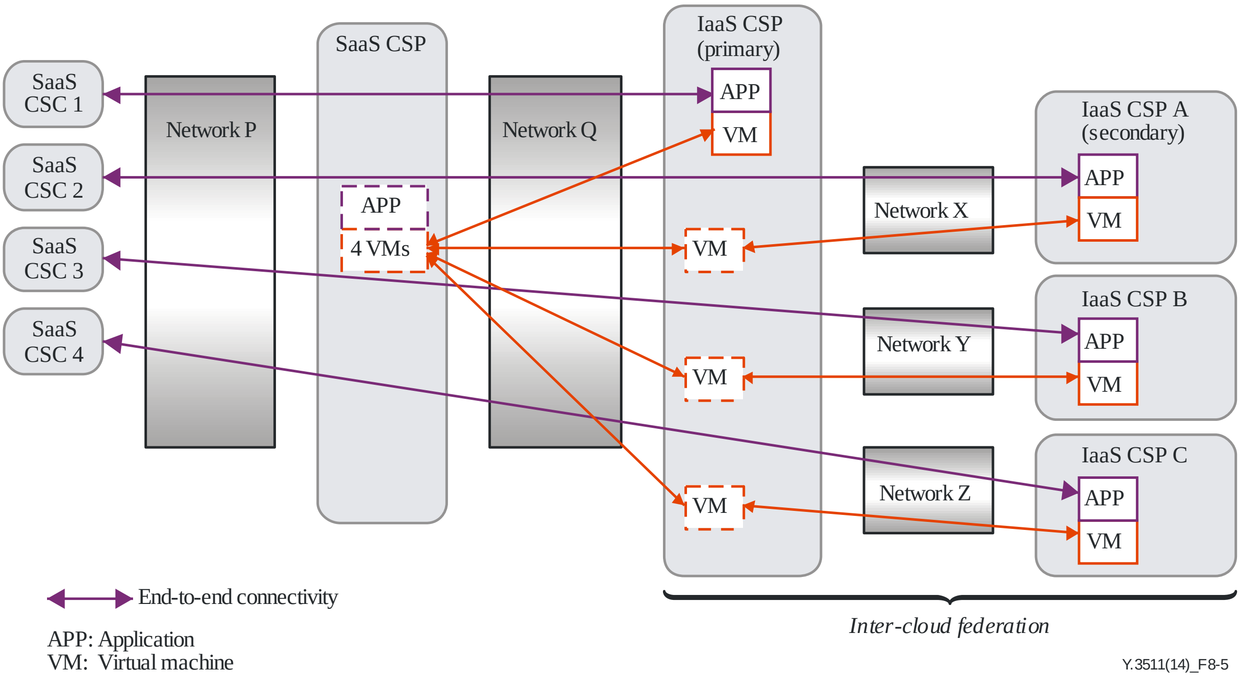


Figure 8-5 – View highlighting actual running VM and application locations

For the sake of simplicity, Figure 8-5 also shows network connectivity between different CSCs and CSPs (illustrated with "Network" boxes). These "Network" boxes may be under the responsibility of third-party providers different from the IaaS CSPs and SaaS CSPs, or may be provided by the IaaS and SaaS CSPs themselves. These "Networks" are involved in the support of end-to-end network connectivity (between the SaaS CSC and the "Application" running in the IaaS CSPs). Network connectivity supported by these "Networks" may be provided as a cloud service of the NaaS service category. Network capabilities for the support of NaaS service category are for further study.

By means of the connectivity provided by these networks, the SaaS CSCs can access the IaaS CSP on which the VMs providing the service run. The SaaS CSCs do not know on which IaaS CSPs the VMs exist, but the networks facilitate each access from SaaS CSCs to the appropriate IaaS CSP.

For the IaaS primary CSP, in order to achieve optimal resource use, it is desirable to handle information about VM availability as well as their connectivity (including bandwidth, quality of service (QoS) and cost). The IaaS CSP may choose to provide both computing and network resources jointly.

Multiple service offering scenarios including network contributions to the cloud services are described in Appendix III.

8.2.4 Interactions in the case of inter-cloud federation

Figure 8-6 shows the interactions involving multiple CSPs in the case of the inter-cloud federation pattern. In the inter-cloud federation, secondary CSPs provide their resources as one kind of service to the primary CSP.

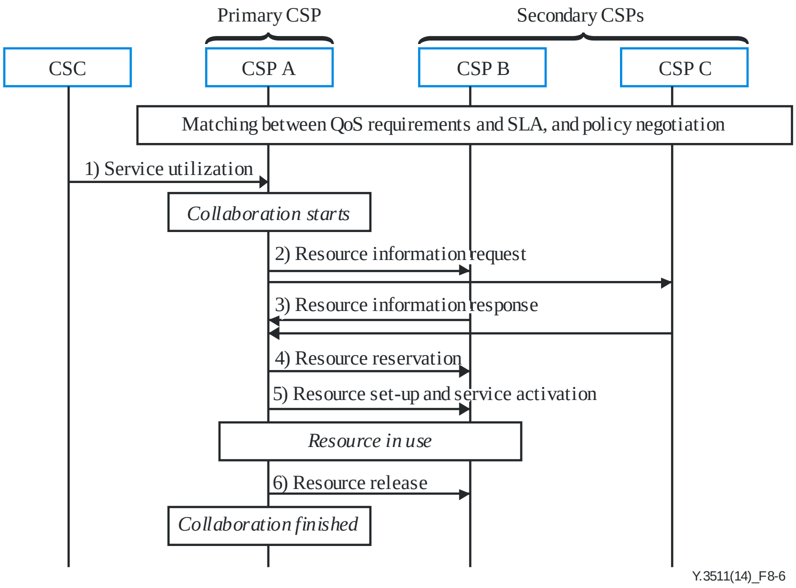


Figure 8-6 – Interaction among CSPs in inter-cloud federation

When the federation is established, the primary and the secondary CSPs perform matching between QoS requirements and SLA, and policy negotiation. The next steps as shown in Figure 8-6 are as follows:

1) The CSC starts to use the service of CSP A. CSP A is the primary CSP for this CSC;

2) The primary CSP (i.e., CSP A) decides to initiate an interaction due to a shortage of resources that results in service quality degradation. CSP A requests information about the resources (e.g., VMs and storage) from CSP B and CSP C. CSP B and CSP C are the secondary CSPs for this interaction;

3) CSP B and CSP C provide responses to CSP A regarding available resources;

4) On the basis of the received responses, CSP A reserves the resources of CSP B. In this step, CSP A estimates the performance of the resources available in CSP B and confirms that the estimated performance is acceptable;

5) CSP A sets up the resources of CSP B and activates the service (this action can be considered as VM migration or application rebuilding). As a result, service quality is maintained.

6) CSP A decides to end the collaboration with CSP B (e.g., CSP A has sufficient resources to provide services by itself or demand for services has decreased). CSP A releases the resources provided by CSP B.

These steps can be applied to every CSP involved in the inter-cloud federation. Each CSP in the federation is the primary CSP for its own CSCs and the CSPs providing resources to a primary CSP can be considered as the secondary CSPs.

The way a primary CSP requests resources from the secondary CSPs may vary from:

– "More-specific" resources request which includes detailed descriptions and may limit the number of candidate resources and resulting responses from the secondary CSPs but may obtain optimal resources if responded. However, complicated resource calculation and performance estimation may be necessary. A "more-specific" resources request is adequate when the cost associated with the offered resources is high and sensitive;

- "Less-specific" resources request soliciting more offerings and resulting in a simple and quick decision, although the offered resources may not be optimal.

A primary CSP may only receive and use a certain number of responses to reduce the processing burden caused by a large number of responses.

Details related to "more-specific" and "less-specific" resource requests are for further study.

8.3 Overview of inter-cloud intermediary

8.3.1 Introduction

The inter-cloud intermediary pattern provides the capability for CSPs to offer additional services to the CSCs and to other CSPs.

As shown in Figure 8-7, one of the central components of the inter-cloud intermediary pattern is the catalogue of service offerings. This CSP's catalogue is a registry that includes the services that the CSP offers to CSCs and to other CSPs. The catalogue provides the capability for CSCs and CSPs to obtain services from the CSP offering the services. This catalogue may be accessible through a portal and/or via a well-defined interface or API.

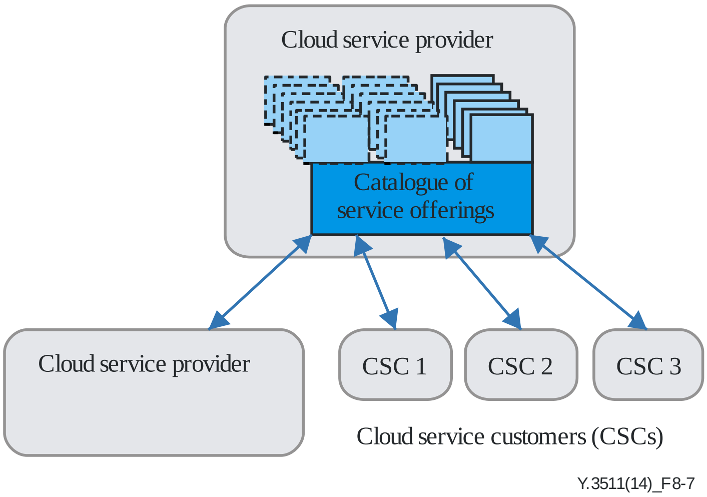


Figure 8-7 – CSCs and CSPs accessing cloud services offerings through a catalogue

In addition to the catalogue of service offerings, the CSP can include functions for the support of service intermediation, service aggregation and service arbitrage as described in clause 7.3.

8.3.2 Primary CSP and secondary CSP

The CSP that is responsible for offering the service to the CSC is the primary CSP. The CSPs that support the primary CSP by offering their services are the secondary CSPs.

In an inter-cloud intermediary pattern (see Figure 8-8), services listed in the primary CSP catalogue of service offerings may include the services hosted by the primary CSP itself or services that are provided by secondary CSPs. In most cases, the primary CSP catalogue of service offerings will be a combination of services hosted by the primary CSP and services offered by the secondary CSPs.

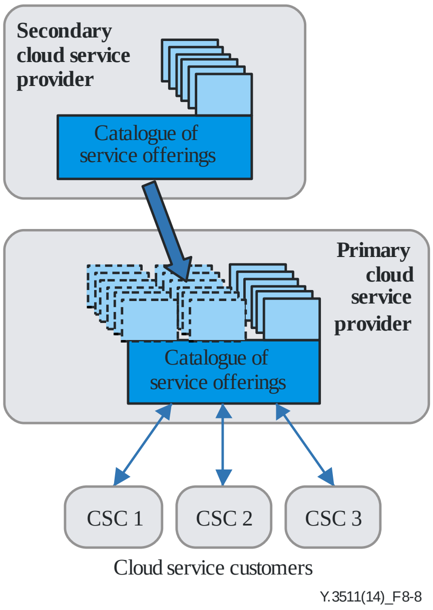

Secondary cloud service provider
Catalogue of service offerings
Primary cloud service provider
Catalogue of service offerings
CSC 1
CSC 2
CSC 3
Cloud service customers
Y.3511(14)_F8-8

Figure 8-8 – Services provided from a secondary CSP to a primary CSP

The primary CSP may offer services from multiple secondary CSPs. Some of the services offered by the secondary CSPs may themselves be their own services or services from other CSPs. Note that the roles of primary and secondary service providers may change depending on the service under consideration.

The primary CSP serving the CSC provides the service level agreement (SLA) to the CSC and is responsible for ensuring that both services hosted by that primary CSP and services offered from a secondary CSP will meet the SLA between the CSC and the primary CSP.

8.3.3 Network connectivity

In considering network connectivity for the inter-cloud intermediary pattern, multiple levels of connectivity may be required.

In the simplest case, the CSP providing services to the CSC (the primary CSP) is hosting all of the services and is providing the network connectivity to the CSC. In this case, the CSP can offer an SLA covering both the cloud and network services.

In a more common case, the CSP providing services to the CSC (the primary CSP) will offer some of the hosted services but will also be offering services from one or more secondary CSPs. The network connectivity between the primary and the secondary CSPs may be offered as one of the primary CSP's services or it may be provided by a distinct third party network provider.

As illustrated by Figure 8-9, the primary CSP is responsible for ensuring that the SLA between the primary CSP and the CSC is met taking into account:

1) the network connectivity between the primary CSP and the CSC;
2) the services from the primary CSP;
3) the network connectivity between the primary CSP and the involved secondary CSPs;
4) the services from the secondary CSPs.

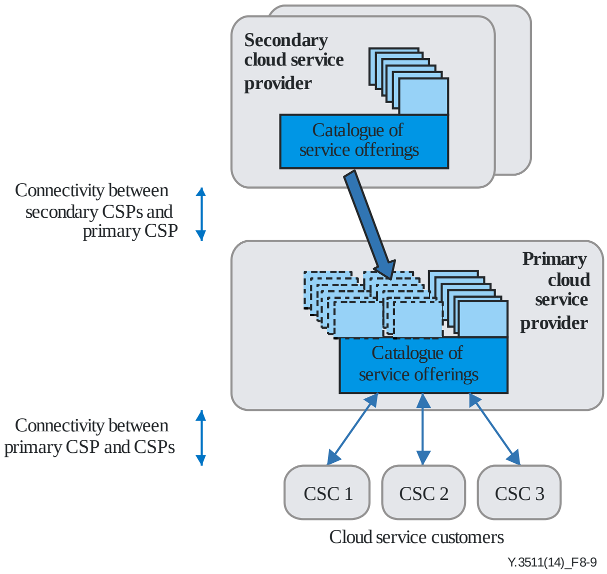

Secondary cloud service provider
Catalogue of service offerings
Connectivity between secondary CSPs and primary CSP
Primary cloud service provider
Catalogue of service offerings
Connectivity between primary CSP and CSPs
CSC 1
CSC 2
CSC 3
Cloud service customers
Y.3511(14)_F8-9

Figure 8-9 – Primary and secondary CSPs and relevant network connectivity

As illustrated in Figure 8-9, the CSCs use services provided by the primary CSP. The catalogue of service offerings in the primary CSP includes services from the secondary CSP. Since the CSCs use services through the primary CSP, the primary CSP is responsible for ensuring that the service levels provided to the CSC meet the SLA.

8.3.4 Interactions in the case of inter-cloud intermediary

Figure 8-10 shows the interactions involving a CSC and multiple CSPs in the case of the inter-cloud intermediary pattern.

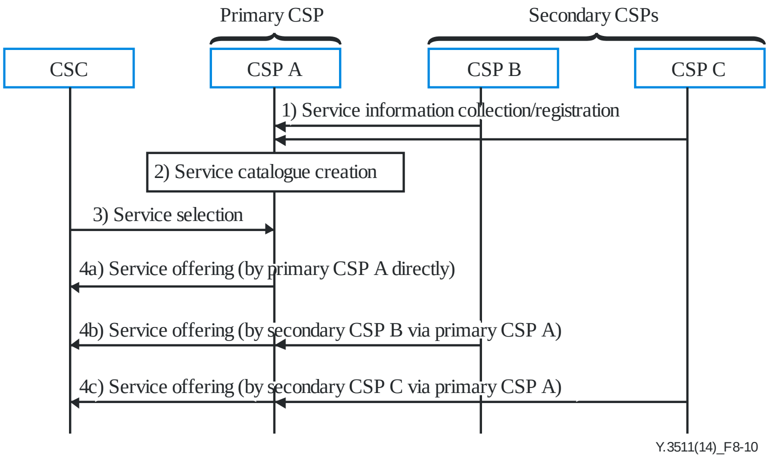

Primary CSP
Secondary CSPs
CSC
CSP A
CSP B
CSP C
1) Service information collection/registration
2) Service catalogue creation
3) Service selection
4a) Service offering (by primary CSP A directly)
4b) Service offering (by secondary CSP B via primary CSP A)
4c) Service offering (by secondary CSP C via primary CSP A)
Y.3511(14)_F8-10

Figure 8-10 – Interactions in inter-cloud intermediary pattern

Along with the service offering, the description hereafter shows the interactions involving multiple CSPs. The steps as shown in Figure 8-10 are as follows:

1) The primary CSP collects service information from the secondary CSPs, or the secondary CSPs register their services to the primary CSP. Difference in terms of SLAs is negotiated as well;

2) The primary CSP creates the catalogue of service offerings by combining the list of hosted services and the services provided by the secondary CSPs;

3) The CSC accesses the primary CSP's catalogue of service offerings and selects one or some of the services in the catalogue;

4) The primary CSP arranges the service selected by the CSC. The service may be provided by the primary CSP, by (some of) the secondary CSP(s) or by a combination of these CSPs. The primary CSP intermediates, aggregates and/or arbitrages the services (see clause 7.3). The CSC starts using the service.

In practice, the actual process will be more complex. There are different ways for the CSC to use the services. If the requested service is hosted in the primary CSP, the CSC may access to this service directly (in the case of 4a)). If the requested service is offered by a secondary CSP, the CSC may access to the service in the secondary CSP via the primary CSP (in the cases of 4b) and 4c)). In the latter case, the network conditions, service characteristics and service agreement among the CSC, the primary CSP and the involved secondary CSP should be considered in order that the CSC accesses to the service in an appropriate manner.

9 Functional requirements for inter-cloud

This clause describes the CSPs' capabilities necessary to support different inter-cloud computing patterns described in clause 8.

The capabilities and the relevant requirements identified in this clause are complementary to the general requirements applicable to a CSP involved in inter-cloud as specified in [ITU-T Y.3501].

9.1 SLA and policy negotiation

The SLA and policy negotiation capability deals with the matching made by a primary CSP between SLA requirements of a CSC and the SLAs of secondary CSPs involved in the considered inter-cloud pattern. The subject includes the QoS related aspects. This capability also deals with the negotiation of service provisioning policies associated with the different CSPs involved in the inter-cloud pattern.

The SLA requirements (including QoS) of a CSC for a given cloud service are expected to be met by appropriate interworking with selected CSPs, even in the event of service performance degradation or a disaster.

The SLA and policy negotiation capability is required to:

– be aware of the SLA information related to the QoS and performance aspects of the CSPs involved in the inter-cloud using standard formats.

The SLA and policy negotiation capability is recommended to:

– allow comparing, negotiating and settling down service provisioning policies between multiple CSPs (for example, based on the settlement, these CSPs may be considered as a trusted group for inter-cloud support).

NOTE – In this clause, policy refers to a way for a CSP to provide services in terms of presumed reliability, including its backup scheme and target service levels. The policy affects SLAs. Policies may be different among CSPs. Policies may be negotiated beforehand and settled. This process is referred to as the policy negotiation.

9.2 Resource monitoring

The resource monitoring capability deals with the monitoring by the primary CSP of the resources of the secondary CSPs and the status attributes of these resources (e.g., usage amount, performance and quality aspects). The primary CSP collects and monitors the information from the secondary CSPs in a secure way. By monitoring the secondary CSPs' resources status (e.g., availability and dead/alive status of machines) and detecting service level performance degradation (in terms of delay and response time), the primary CSP can initiate actions to maintain service availability with the help of other secondary CSPs.

The resource monitoring capability is recommended to:

- allow describing and expressing the resource information (e.g., resource type, configuration and status) in a standard manner in order to be able to monitor these resources across multiple CSPs;
- allow updating the resource information across multiple CSPs in synchronization with the events (e.g., reserve or release of resources) involving the CSPs;
- allow periodically, or on a request basis, collecting information about the usage and performance status of the resources of multiple CSPs;
- allow periodically, or on a request basis, collecting information about the resource availability (e.g., dead or alive status of machines) of multiple CSPs;
- allow exchange of monitoring information in commonly defined ways across multiple CSPs.

9.3 Resource performance estimation and selection

The resource performance estimation and selection capability deals with the selection of resources from the candidate resources that have already been reserved in peer CSPs. This capability estimates the achievable performance of available reserved resources and assists the CSP in the selection of resources to be effectively used.

The resource performance estimation and selection capability is recommended to:

- allow estimating the achievable performance of available reserved resources (e.g., computing resources, storage resources, input/output capacity between storage resources, network bandwidth) in the secondary CSPs.

9.4 Resource discovery and reservation

The resource discovery and reservation capability deals with search, discovery and reservation of available resources in the peer CSPs. This capability also deals with reservation acknowledgement for the candidate resources that have been tentatively reserved in the peer CSPs.

The resource discovery and reservation capability is recommended to:

- enable discovery of resources available in the peer CSPs;
- allow the reservation of discovered resources in the peer CSPs;
- allow provisional reservation of discovered resources, i.e., to keep the resources to be used (as candidates), for later acknowledgement (for some of them) or release (for others);
- allow finding available resources in the peer CSPs based on different priorities (e.g., in a different order of searching);

  NOTE 1 – Quality requirements may vary from service to service and each resource contribution to the service quality may vary as well. For example, if latency is critical, it should be possible to first reserve resources in the servers that are near to the user and then reserve the network resources. In contrast, if bandwidth is critical, it should be possible to first reserve resources of the networks that can provide sufficient bandwidth and then search for available resource in the servers that are connected to those networks.
- allow reservation of available resources in the peer CSPs on the basis of different priorities (e.g., early recovery, required quality guarantee, service type, etc.).

  NOTE 2 – For example, a vast quantity of resources is required for recovery from a large-scale disaster. However, all required resources may not necessarily be available. In that case, it should be possible to forcefully reserve resources for lifeline services rather than for other services.

9.5 Resource set-up and activation

The resource set-up and activation capability deals with the set up and activation of reserved resources in the peer CSPs. This includes connecting to the peer CSPs via networks, remotely activating (i.e., invoking) software and transferring or copying data to enable the use of resources in the peer CSPs.

The resource set-up and activation capability is recommended to:

– allow the establishment of reserved resources in a peer CSP;

– allow accessing to the configuration and policy settings of reserved resources in the peer CSPs.

9.6 Cloud services switchover and switchback

The cloud services switchover and switchback capability deals with the switchover of the CSC's end-user access to cloud services from the primary CSP to a peer CSP to which the services may be provided in order to cope with degradation in service performance or a serious problem. It also deals with the switchback to the primary CSP when it is able to provide the services again. It should be noted that the reasons for switch-over ranges from a load distribution between CSPs, in which the primary CSP role is maintained as it is, to a serious problem, in which the primary CSP role is delegated to the peer CSP. The capability differences to reflect those reasons need further study.

The cloud service switchover and switchback capability is recommended to:

– allow switching over CSC's end-user access to a peer CSP (acting as primary CSP) without manual operation from the CSC, in order to allow the CSC's end user to use services in a similar manner to the way they did before the access was switchover;

– allow switching back the CSC's end-user access to the primary CSP when this CSP has recovered from the reasons that led to the switchover (e.g., a disaster or load distribution between peer CSPs is no longer needed).

9.7 Resource release

The resource release capability deals with the release by a CSP of the peer CSPs' (reserved and/or used) resources after determining that these resources are no longer needed, e.g., based on monitoring the results such as disaster recovery has been completed or load has been reduced.

The resource release capability is recommended to:

– allow releasing by the CSP of resources reserved, activated and/or set up in the peer CSPs;

– allow updating the peer CSP's resource configuration information;

– allow erasing and/or transferring back cloud application data received during the resource reservation.

9.8 CSC information exchange

The CSC information exchange capability deals with exchanging CSC profiles and associated information between a primary CSP and the secondary CSPs. The information associated with a CSC is initially maintained by the primary CSP. When the primary CSP requests that the secondary CSPs should provide additional resources and run applications over the secondary CSP resources, the secondary CSPs may need to perform customer management by inheriting the CSC profiles and associated information given by the primary CSP. Activation of CSC information exchange needs prior agreement of the CSC.

The CSC information exchange capability is required to:

– be activated only with the prior agreement of the CSC;

– be able to manage CSC profiles and associated information.

The CSC information exchange capability is recommended to:

– be able to exchange CSC profiles and associated information among multiple CSPs according to a pre-determined protocol and format, with the condition that the CSC is informed of and agrees to the exchange.

9.9 Primary CSP role delegation

The primary CSP role delegation capability deals with transferring the primary CSP role to one of the secondary CSPs, e.g., in the event of a serious problem caused by natural disasters or permanent service termination occurring at the current primary CSP. In preparation for the serious problem or permanent service termination at the primary CSP, all management information associated with the primary CSP is shared with the secondary CSPs, while the absolute controllability of the information, i.e., permission to update the information, is still held by the primary CSP. When the serious problem or service termination occurs, the absolute controllability for a given primary CSP role, i.e., permission, is transferred to one of the designated secondary CSPs. By transferring the responsibility of the primary CSP role with associated management information, the service can continue even if the primary CSP's systems are seriously damaged, e.g., due to a natural disaster or the CSP stops a service due to economic decisions (refer to use cases in clauses I.4 and I.5). Activation of the primary CSP role delegation needs prior agreement of the CSC.

The primary CSP role delegation capability is required to:

– be activated only with the prior agreement of the CSC.

The primary CSP role delegation capability is recommended to:

– allow a CSP to discover peer CSPs that are capable of inheriting the primary CSP role, and enable the CSP to negotiate with these peer CSPs as to whether they can accept the inheritance;

– allow a CSP to transfer its management information associated with the primary CSP role in a reliable manner (e.g., periodically) to the peer CSPs that have accepted the permission transfer with that CSP;

– allow the controllability of the information associated with the primary CSP role to be transferred to the secondary CSPs with minimum interruptions;

– allow a CSP to cancel the permission transfer arrangements.

9.10 Inter-cloud service handling

The inter-cloud service handling capability deals with the primary CSP offering cloud services to its CSCs based on the handling of services provided by the secondary CSPs. This capability can be used for inter-cloud intermediary pattern.

The inter-cloud service handling capability is required to:

– support service intermediation, i.e., conditioning or enhancing the cloud service of a peer CSP;

– support service aggregation, i.e., providing the composition of a set of services provided by the CSPs;

– support service arbitrage, i.e., selecting one service offering from a group offered by the peer CSPs.

10 Security considerations

The security framework for cloud computing is described in [ITU-T X.1601] covering security challenges for CSPs. In particular, [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.

Appendix IV identifies important aspects that should be considered when developing Recommendations addressing inter-cloud security aspects.

Appendix I

Use cases from the inter-cloud perspective

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

This appendix describes use cases in which multiple cloud computing systems interact with each other to satisfy the specified requirements and how cloud systems work in each use case.

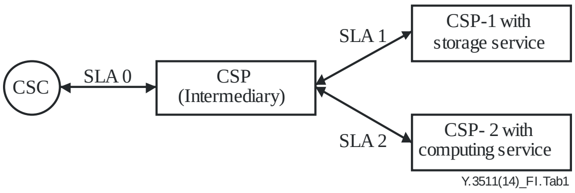
I.1 SLA mapping in intermediary pattern

This use case illustrates the SLA mapping between the primary CSP in inter-cloud intermediary pattern (called CSP-Intermediary) and other secondary CSPs.

Multiple CSPs contribute to, or impact concurrently, the SLA between the CSP-Intermediary and the CSC when an orchestrated service is provided.

Table I.1 shows SLA mapping in an inter-cloud intermediary pattern.

Table I.1 – SLA mapping in an inter-cloud intermediary pattern

Use case	
Use case title	SLA mapping in an inter-cloud intermediary pattern
Relevant roles	CSC and CSP
Use case description	– The primary CSP in an inter-cloud intermediary pattern (CSP-Intermediary) is the contact point for CSC and there is SLA (SLA0) between them. – The CSP-Intermediary integrates services from multiple CSPs, for instance, storage service from CSP-1 and computing service from CSP-2. There are business-to-business (B2B) level SLAs between CSP-Intermediary and CSP-1, CSP-2 respectively (SLA1, SLA2). – For the CSP-Intermediary, in order to guarantee SLA0 for CSC, it is necessary to map SLA0 to SLA1 and SLA2, because SLA0 is actually implemented by SLA1 and SLA2.
Information flow	SLA mapping may be performed via explicit information exchange or off-line negotiation.
High-level figure describing the use case	 SLA 1 CSP-1 with storage service CSC SLA 0 CSP (Intermediary) SLA 2 CSP- 2 with computing service Y.3511(14)_FI.Tab1
Derived requirements for cloud capability	– The capability to support SLA negotiation between CSP-Intermediary and other CSPs is recommended. – The capability to support coordination of the SLAs from multiple CSPs (which is related to a business decision) is recommended.

I.2 Performance guarantee against an abrupt increase in load (offloading)

Table I.2 shows an inter-cloud use case where performance is guaranteed in case of an abrupt increase in load.

NOTE – The following legend applies to the Figures in Tables I.2 to I.5:

🟩 Virtual resources (i.e., virtual machine, virtual storage, and virtual network)

o Ongoing applications (e.g., snapshot image of the main memory)

n Newly invoked applications

Y.3511(14)_FI.Lgnd

Table I.2 – Inter-cloud use case: Performance guarantee against an abrupt increase in load

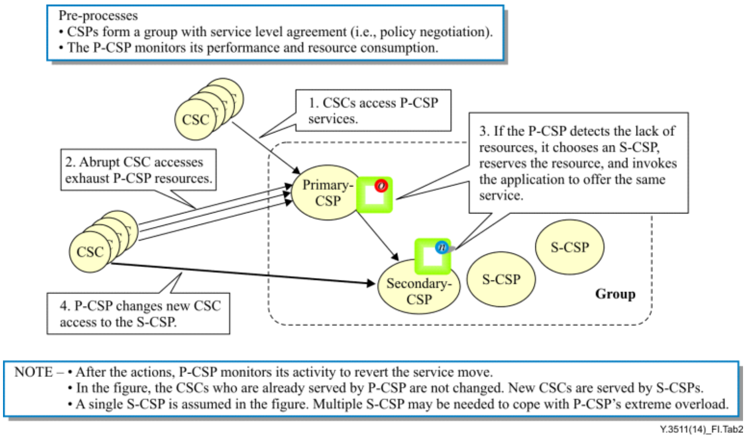
Use case	
Use case title	Inter-cloud use case: Performance guarantee against an abrupt increase in load
Relevant roles	CSP and CSC
Use case description	– A CSP guarantees its service performance, even when an unexpected surge in access to the service arises, by using cloud resources provided by other CSPs on a temporary basis. – When overload is detected at a CSP, available resources in other CSPs are autonomously discovered and reserved through the inter-cloud federation. – Network connections among interworking CSPs are instantaneously established or reconfigured. Then service-related data including user identifier (ID), user data and application data are transferred from the original CSP to the CSP that is leasing the resources. – Access from CSCs is appropriately changed to the interworking CSPs so as to distribute the load and thus mitigate the overload of the original CSP.
Information flow	– Relevant CSPs are supposed to join a common trusted alliance (i.e., federation) in advance and set up the service level agreements (SLAs). – A CSP inquires about the resource availability of other CSPs in the federation and requests reservation of the available resources that meet the quality requirements of the CSC. The requested CSPs reply whether or not they are able to lease the requested resources. – The cloud resource management (e.g., CRUD: create, read, update and delete) are operated across multiple CSPs. The management is to enable cloud resources to be leased from different CSPs in the federation. – The relevant CSPs exchange monitoring and auditing information of the leased resources.
High-level figure describing the use case	 Pre-processes • CSPs form a group with service level agreement (i.e., policy negotiation). • The P-CSP monitors its performance and resource consumption. 1. CSCs access P-CSP services. 2. Abrupt CSC accesses exhaust P-CSP resources. 3. If the P-CSP detects the lack of resources, it chooses an S-CSP, reserves the resource, and invokes the application to offer the same service. 4. P-CSP changes new CSC access to the S-CSP. CSC CSC Primary-CSP Secondary-CSP S-CSP S-CSP Group NOTE – • After the actions, P-CSP monitors its activity to revert the service move. • In the figure, the CSCs who are already served by P-CSP are not changed. New CSCs are served by S-CSPs. • A single S-CSP is assumed in the figure. Multiple S-CSP may be needed to cope with P-CSP's extreme overload. Y.3511(14)_FI.Tab2

Table I.2 – Inter-cloud use case: Performance guarantee against an abrupt increase in load

Use case	
Derived requirements for the cloud capability	The capability is required to support: – Policy negotiation including SLA management among the multiple CSPs within a pre-established group (i.e., federation); NOTE – Policy refers to a way for a CSP to provide services in terms of presumed reliability of a machine, including its backup scheme and target service levels. The policy may be different with each CSP. To maintain the same quality of service to the CSC even when the CSP changes, the difference should be negotiated beforehand and settled. This process is referred to as policy negotiation. The same note applies to the other use cases. – Self-performance monitoring at a CSP. If the performance degrades, the CSP should initiate the next configured actions; – Discovery, reservation, use and release of cloud resources in a dynamic manner (i.e., not relying on the pre-configuration) on other CSPs within a federation; – Application invocation over the reserved resources on other CSPs within a federation; – Alteration and reversion (i.e., switchover and switchback) of CSC access from one CSP to another CSP in a dynamic manner (i.e., not relying on the pre-configuration) within a federation; – Exchange of monitoring and auditing information among the multiple CSPs within a federation; – Exchange of authentication information about CSC (user/enterprise) authentication status among the multiple CSPs within a federation.

I.3 Performance guarantee regarding delay (optimization for user location)

Table I.3 shows inter-cloud use case for performance guarantee regarding delay.

Table I.3 – Inter-cloud use case: Performance guarantee regarding delay

Use case	
Use case title	Inter-cloud use case: Performance guarantee regarding delay
Relevant roles	CSP and CSC
Use case description	– CSPs guarantee their service performance (in particular, network delay and response time), even when a CSC moves to a remote location (e.g., on a business trip), by using cloud resources provided by another CSP located close to the CSC on a temporary basis. – When degradation in the response time is detected for a CSC at a CSP, available resources are autonomously discovered and reserved in another CSP that is near the CSC based on the user's location information.

Table I.3 – Inter-cloud use case: Performance guarantee regarding delay

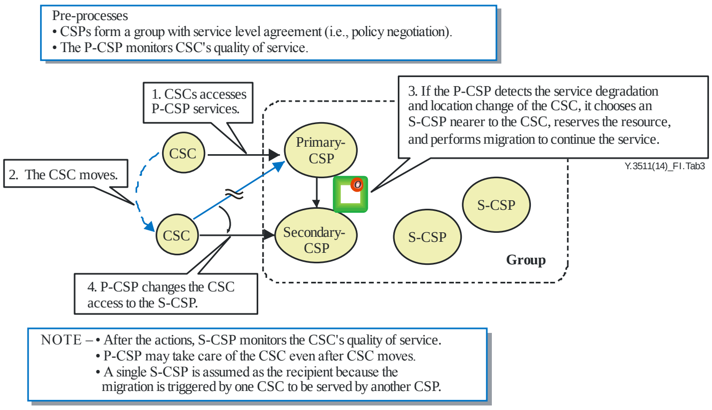
Use case	
	– Network connections among interworking CSPs are instantaneously established or reconfigured. Then service-related data including user identifier (ID), user data and application data are transferred from the original CSP to the CSP that is leasing the resources. – Access from CSCs is appropriately changed to the interworking CSP so as to achieve route optimization and thus mitigate the performance degradation caused by the distance from the original CSP. – As a result, the CSC, who keeps the same user ID, can continuously access the service at the same level of response time as before.
Information flow	– Relevant CSPs are supposed to join a common trusted alliance (federation) in advance and set up the service level agreements (SLAs). – A CSP inquires about the resource availability of other CSPs in the federation, and requests a reservation of available resources that meet the quality requirements of the CSC. The requested CSPs reply whether or not they are able to lease the resources. – The cloud resource management (e.g., CRUD: create, read, update and delete) are operated across multiple CSPs. The management is to enable the leasing of cloud resources from different CSPs in the federation. – The relevant CSPs exchange monitoring and auditing information of the leased resources.
High-level figure describing the use case	 Pre-processes • CSPs form a group with service level agreement (i.e., policy negotiation). • The P-CSP monitors CSC's quality of service. 1. CSCs accesses P-CSP services. 2. The CSC moves. 3. If the P-CSP detects the service degradation and location change of the CSC, it chooses an S-CSP nearer to the CSC, reserves the resource, and performs migration to continue the service. Y.3511(14)_FI.Tab3 CSC Primary-CSP CSC Secondary-CSP S-CSP S-CSP Group 4. P-CSP changes the CSC access to the S-CSP. NOTE – • After the actions, S-CSP monitors the CSC's quality of service. • P-CSP may take care of the CSC even after CSC moves. • A single S-CSP is assumed as the recipient because the migration is triggered by one CSC to be served by another CSP.

Table I.3 – Inter-cloud use case: Performance guarantee regarding delay

Use case	
Derived requirements for the cloud capability	The capability is required to support: – Policy negotiation including SLA management among the multiple CSPs within a pre-established group (i.e., federation); – CSC service level monitoring at CSP. If the service level degrades, the CSP should initiate the next configured actions; – Discovery, reservation, use and release of cloud resources, based on CSC location, in a dynamic manner (i.e., not relying on the pre-configuration) on other CSPs within the federation; – Capability migration (e.g., virtual machine (VM) and applications) over the reserved resources on other CSPs within the federation; – Alteration and reversion (i.e., switchover and switchback) of CSC access to one CSP to another CSP in a dynamic manner (i.e., not relying on the pre-configuration) within the federation; – Exchange of monitoring and auditing information among the multiple CSPs within the federation; – Exchange of authentication information about CSC (user/enterprise) authentication status among the multiple CSPs within the federation.

I.4 Guaranteed availability in the event of a disaster or large-scale failure

Table I.4 shows inter-cloud use case for guaranteed availability in the event of a disaster or large-scale failure.

Table I.4 – Inter-cloud use case: Guaranteed availability in the event of a disaster or large-scale failure

Use case	
Use case title	Inter-cloud use case: Guaranteed availability in the event of a disaster or large-scale failure
Relevant roles	CSP and CSC
Use case description	– CSPs continue offering their service using the resources leased from each other, even when systems in one CSP are damaged due to natural disasters or large-scale failures. – Available resources in other CSPs are autonomously discovered and reserved through the inter-cloud federation. – The services with a high priority are only recovered if available resources are not sufficient to recover all services. In examining the availability of resources provided by other CSPs, the guaranteed level of quality of the resources is taken into account.

Table I.4 – Inter-cloud use case: Guaranteed availability in the event of a disaster or large-scale failure

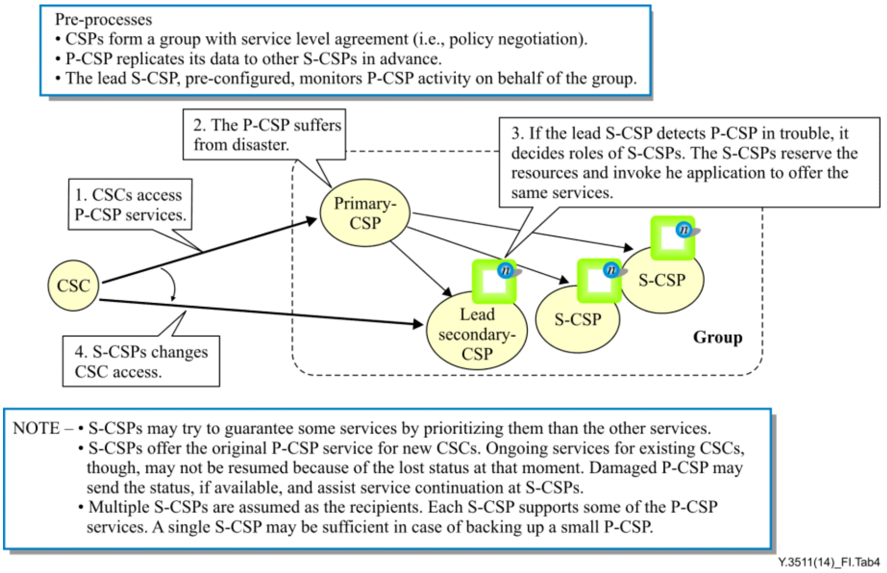
Use case	
	– The services requiring early recovery are recovered using available resources on a best-effort basis even if their quality requirements are partly satisfied. – Network connections among interworking CSPs are instantaneously established or reconfigured. The lead CSP, which is preconfigured and governs the recovery procedure, manages the roles of available CSPs and instructs service continuation based on the original CSP data. – Access from CSCs is appropriately distributed to the interworking CSPs so as to achieve the disaster recovery and thus mitigate the service discontinuity.
Information flow	– Relevant CSPs are supposed to join a common trusted alliance (federation) in advance and set up the service level agreements (SLAs). – The lead CSP, which is preconfigured and governs the recovery procedures, inquires about the resource availability of other CSPs in the alliance to recover its cloud services to meet quality requirements of the CSCs. The requested CSPs reply whether or not they are able to lease the resources. – The cloud resource management (e.g., CRUD: create, read, update and delete) is operated across multiple CSPs. The management is to enable leasing of cloud resources from different CSPs in the alliance. – The relevant CSPs exchange monitoring and auditing information of the leased resources.
High-level figure describing the use case	 Pre-processes • CSPs form a group with service level agreement (i.e., policy negotiation). • P-CSP replicates its data to other S-CSPs in advance. • The lead S-CSP, pre-configured, monitors P-CSP activity on behalf of the group. 2. The P-CSP suffers from disaster. 3. If the lead S-CSP detects P-CSP in trouble, it decides roles of S-CSPs. The S-CSPs reserve the resources and invoke he application to offer the same services. 1. CSCs access P-CSP services. Primary-CSP CSC Lead secondary-CSP S-CSP S-CSP Group 4. S-CSPs changes CSC access. NOTE – • S-CSPs may try to guarantee some services by prioritizing them than the other services. • S-CSPs offer the original P-CSP service for new CSCs. Ongoing services for existing CSCs, though, may not be resumed because of the lost status at that moment. Damaged P-CSP may send the status, if available, and assist service continuation at S-CSPs. • Multiple S-CSPs are assumed as the recipients. Each S-CSP supports some of the P-CSP services. A single S-CSP may be sufficient in case of backing up a small P-CSP. Y.3511(14)_FI.Tab4
Derived requirements for cloud capability	The system is required to support: – Policy negotiation including SLA management among the multiple CSPs within a pre-established group; – Self-activity monitoring at a CSP or mutual activity monitoring among the CSPs in a pre-established group. If the activity disappears, the detecting CSP should initiate the pre-configured actions; – Discovery, reservation, use and release of cloud resources in a dynamic manner (i.e., not relying on the pre-configuration) on other CSPs within the federation;

Table I.4 – Inter-cloud use case: Guaranteed availability in the event of a disaster or large-scale failure

Use case	
	– Application invocation over the reserved resources on other CSPs within the federation; – Alteration and reversion (i.e., switchover and switchback), in a dynamic manner (i.e., not relying on the pre-configuration), of CSC access to any CSP within the federation; – Exchange of monitoring and auditing information among the multiple CSPs within the federation; – Exchange of authentication information about CSC (user/enterprise) authentication status among the multiple CSPs within the federation.

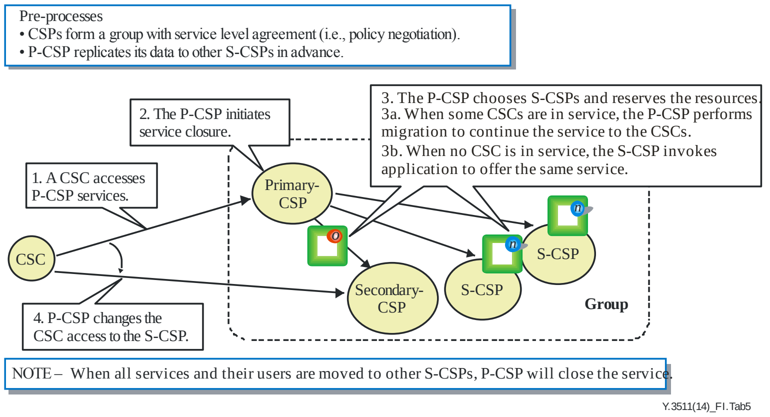
I.5 Service continuity (in the case of service termination of the original CSP)

Table I.5 shows an inter-cloud use case for service continuity in the case of service termination of the original CSP.

Table I.5 – Inter-cloud use case: Service continuity

Use case	
Use case title	Inter-cloud use case: Service continuity
Relevant roles	CSP and CSC
Use case description	– The cloud service offering continues through the collaboration with other CSPs, even when the original CSP terminates its business. – Available resources in the CSPs other than the service-terminating CSP are discovered and reserved in advance. – Network connections among the interworking CSPs are established or reconfigured. Then service-related data including user identifier (ID), user data and application data are transferred from the original CSP to the new CSPs. – Access from CSCs is appropriately changed to the interworking CSPs so that the same service is continuously offered. – If the capabilities (VM and applications) at the original CSP migrate to other CSPs, the CSC, who keeps the same user ID, can continuously access the service at the same level of performance as before.
Information flow	– The relevant CSPs are supposed to join a common trusted alliance in advance and set up the service level agreements (SLAs). – The terminating CSP inquires about the resource availability of other CSPs in the alliance and requests a reservation of the available resources to continue the services. – The cloud resource management (e.g., CRUD: create, read, update and delete) are operated across multiple CSPs. The management is to enable leasing of the cloud resources from different CSPs in the federation.

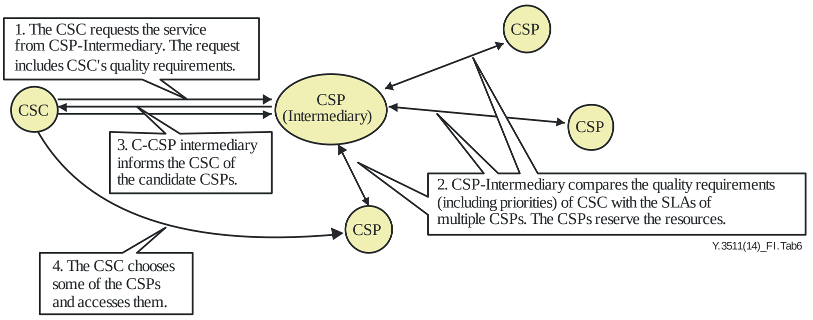
Table I.5 – Inter-cloud use case: Service continuity

Use case	
High-level figure describing the use case	 Pre-processes • CSPs form a group with service level agreement (i.e., policy negotiation). • P-CSP replicates its data to other S-CSPs in advance. 2. The P-CSP initiates service closure. 3. The P-CSP chooses S-CSPs and reserves the resources. 3a. When some CSCs are in service, the P-CSP performs migration to continue the service to the CSCs. 3b. When no CSC is in service, the S-CSP invokes application to offer the same service. 1. A CSC accesses P-CSP services. Primary-CSP CSC S-CSP S-CSP Secondary-CSP Group 4. P-CSP changes the CSC access to the S-CSP. NOTE – When all services and their users are moved to other S-CSPs, P-CSP will close the service. Y.3511(14)_FI.Tab5
Derived requirements for cloud capability	The system is required to support: – Policy negotiation including SLA management among the multiple CSPs within a pre-established group (federation); – Discovery, reservation, use and release of the cloud resources in a dynamic manner (i.e., not relying on the pre-configuration) across the multiple CSPs within the federation; – Capability migration (e.g., VM and applications) among multiple CSPs within the federation; – Alteration (i.e., switchover) of the CSC access, in a dynamic manner (i.e., not relying on the pre-configuration), from one CSP to another CSP within the federation; – Exchange of authentication information about CSC (user/enterprise) authentication status among the multiple CSPs within the federation.

I.6 Market transactions in inter-cloud intermediary pattern

Table I.6 shows inter-cloud use case for market transactions in inter-cloud intermediary pattern.

Table I.6 – Inter-cloud use case: Market transactions in inter-cloud intermediary pattern

Use case	
Use case title	Inter-cloud use case: Market transactions in inter-cloud intermediary pattern
Relevant roles	CSP and CSC
Use case description	– The primary CSP in an inter-cloud intermediary pattern (CSP-Intermediary) mediates between CSPs meeting the CSC's quality requirements and provides the list of selected CSPs to the CSC. – The CSP-Intermediary coordinates multiple services offered by other CSPs.
Information flow	– The SLAs of the CSPs are submitted to the CSP-Intermediary in advance. – A CSC requests the CSP-Intermediary to select CSPs that provide a service which satisfies the CSC's quality requirements. – The CSP-Intermediary compares the CSC quality requirements with the SLAs of other CSPs. Then the CSP-Intermediary discovers and reserves the CSP resources that meet the CSC's quality requirements. – The CSP-Intermediary returns the CSP candidate list to the CSC. – The CSC selects a CSP or CSPs on the list. – The CSP-Intermediary sends a cloud service adaptation request to the selected CSP to invoke the service and adapt it to concrete cloud services and resources. – The CSP returns an adaptation response to the CSP-Intermediary.
High-level figure describing the use case	 1. The CSC requests the service from CSP-Intermediary. The request includes CSC's quality requirements. 2. CSP-Intermediary compares the quality requirements (including priorities) of CSC with the SLAs of multiple CSPs. The CSPs reserve the resources. 3. C-CSP intermediary informs the CSC of the candidate CSPs. 4. The CSC chooses some of the CSPs and accesses them. Y.3511(14)_FI.Tab6
Derived requirements for cloud capability	The system is required to support: – Policy negotiation including SLA management among the multiple CSPs including CSP-Intermediary in a pre-established group; – Discovery, reservation, use and release of cloud resources in a dynamic manner (i.e., not relying on the pre-configuration) on other CSPs within the federation; – Creation of the network connections in a dynamic manner (i.e., not relying on the pre-configuration) from the CSC to the selected CSP that provides the resources; – Flexible reallocation of applications, to meet requirements at different stages in its lifecycle, across multiple CSPs.

Appendix II

Use cases from cloud service providers' views

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

This appendix describes nine inter-cloud related use cases from the perspective of the cloud service provider.

Introduction to participants

For the purpose of this analysis, the following participants are considered. Each of the boxes in Figure II.1 represents a cloud service provider (CSP).

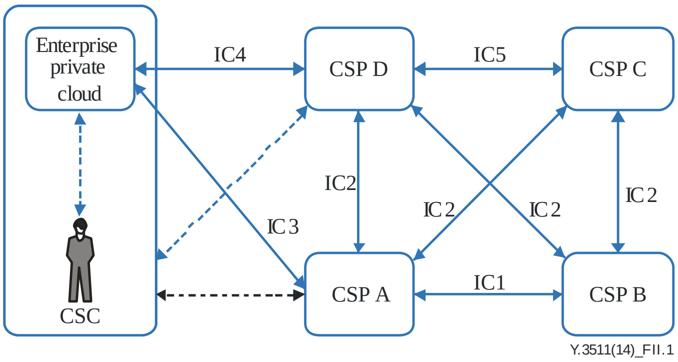


Figure II.1 – Inter-cloud participants and relationships

Participant	Description
User	The human or machine end-user of the overall cloud computing service.
CSP	Cloud service provider: Party (e.g., information technology (IT) or telecom organization) which makes cloud services available. This may include any of the cloud services (SaaS, CaaS, PaaS, IaaS, and NaaS).
Enterprise private cloud	IT resources within an enterprise that are constructed using cloud computing technologies, but which are owned and operated by the enterprise for their own internal use.
Not included for inter-cloud	– Hosting services using non-cloud technologies. – Connectivity services not employing cloud technologies.

The above categorization of participants is for the purpose of use case analysis only and does not imply specific business or regulatory situations. Not all participants identified will be present in all situations. Some organizations may fulfil multiple participant roles.

Relationship	Description
ICn	Inter-cloud relationship (focus of study)

The relationships labelled on the diagram are used to clarify the use cases and do not necessarily indicate information flows or interfaces that require standardization by ITU. The relationships shown as dashed lines (- - - -) in Figure II.1 are included for completeness and are outside the scope of inter-cloud.

II.1 Use case 1 – Cloud service rebranding

CSP-A wishes to offer browser-based office productivity suite services to their users but does not want to run a data centre or build the applications. CSP-A resells the office suite services built and operated by CSP-D, using CSP-A branding, IP network connectivity (IC2), and customer management, while CSP-D develops and maintains the applications and runs the service.

II.2 Use case 2 – Discovery

CSP-A offers a directory service for cloud services to their users. CSP-D and CSP-C both advertise their cloud service offerings into CSP-A's directory (IC2). Enterprise CSC wishes to find a disaster recovery backup provider, uses CSP-A's directory (IC3) to determine that CSP-D offers this service at a good price, and connects with CSP-D via CSP-A's network (IC4) to use the service.

II.3 Use case 3 – Intermediary

CSP-A offers an intermediary service. Enterprise CSC requests the CSP-A to provide hosting of a virtual machine (IC3), CSP-A determines that CSP-D offers the best match of requirements, reserves the resources at CSP-D and creates the necessary connectivity (IC2). Enterprise CSC might or might not know the identity of CSP-D, depending on the requirements of the SLA.

II.4 Use case 4 – Platforming

CSP-D develops a cloud computing application to host consumer music collections under their own brand. CSP-D subscribes to CSP-C's PaaS offering (IC5) and deploys their SaaS application onto CSP-C's PaaS. Consumers connect their devices to CSP-A's application, which is actually running at CSP-C's datacentre (IC2) via virtual private network (VPN).

II.5 Use case 5 – Offloading

Enterprise CSC runs an engineering simulation package which requires significant amounts of computing power at infrequent intervals. CSC's private cloud does not have sufficient peak capacity to handle this effectively, so they have contracted with CSP-A to provide additional compute power (IC3). Due to the success of CSC's business, they now need more peak computing power than CSP-A can provide from CSP-A's own cloud data centre, so CSP-A reserves additional computing resources from CSP-D, handles the load and bills CSC accordingly.

II.6 Use case 6 – Virtual data centre expansion

CSP-A has encountered resistance to expansion of their cloud data centres due to environmental considerations. CSP-A therefore orders 1000 new virtual machine (VM) instances from CSP-D and establishes a VPN bridge such that the new VMs appear to be on the same virtual local area network (LAN) as used in their own data centre.

II.7 Use case 7 – Distributed media

A broadcaster (CSC) will be hosting a major television sporting event series with a global audience and wishes to offer both live and on-demand streaming of the event to many types of devices. CSC requests CSP-A to provide global distribution. CSP-A establishes connection of the live source feeds to CSP-D, which provides secure media reformatting as part of their PaaS offering (IC2), returning digital rights management (DRM)-protected streams/files suitable for playing on many types of devices. CSP-A also develops a global authentication tool and deploys this on PaaS offerings from other CSPs worldwide (IC1, IC2). CSP-A also books capacity in content distribution network (CDN) services worldwide. When the event begins, millions of consumer devices are able to authenticate themselves on their local network provider and stream the content from an efficient local source.

II.8 Use case 8 – Cloud storage expansion

A scientific organization (CSC) collects very large volumes of data in a short period of time that will take years to study. They have sufficient CPU power to analyse this over time but the volume exceeds the storage capacity of their own cloud. CSC contracts with CSP-D to provide additional storage capacity. CSC requests CSP-A to provide very high bandwidth connectivity between CSC and CSP-D. CSC writes the incoming data directly to CSP-D's cloud data storage, and then reduces the network bandwidth to normal levels. CSC is now able to run queries on their data directly at CSP-D or to download interesting parts of the data to their private cloud for intensive processing.

II.9 Use case 9 – Service delivery platform components

A business conference organization (CSC) wishes to rapidly develop and deploy an interactive media conferencing application to multiple venues for an upcoming event. CSP-A offers a service delivery platform (SDP) that includes pre-built components for such services. The CSC developers write their application using several off-the-shelf PaaS, NaaS and CaaS components provided by the SDP platform and are thus able to quickly and reliably create a complex multimedia application and deploy this to CSP-A's SDP.

Appendix III

Abstract service offering models for inter-cloud computing

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

This appendix describes several abstract service offering models relevant to inter-cloud computing and provides supplementary information to the description provided in the main body of the Recommendation.

Inter-cloud computing performed over multiple CSPs allows a CSP (i.e., the primary CSP) to offer new services with respect to expanded service items (cf. clause III.1) and enhanced service operations (cf. clause III.2).

III.1 Service item expansion

In terms of service variety, there are roughly two types of cloud expansions; one refers to adding more of the same resources that the primary CSP already has, and the other refers to adding features based on the resources that differ from those the primary CSP has.

Figure III.1 shows these two expansions.

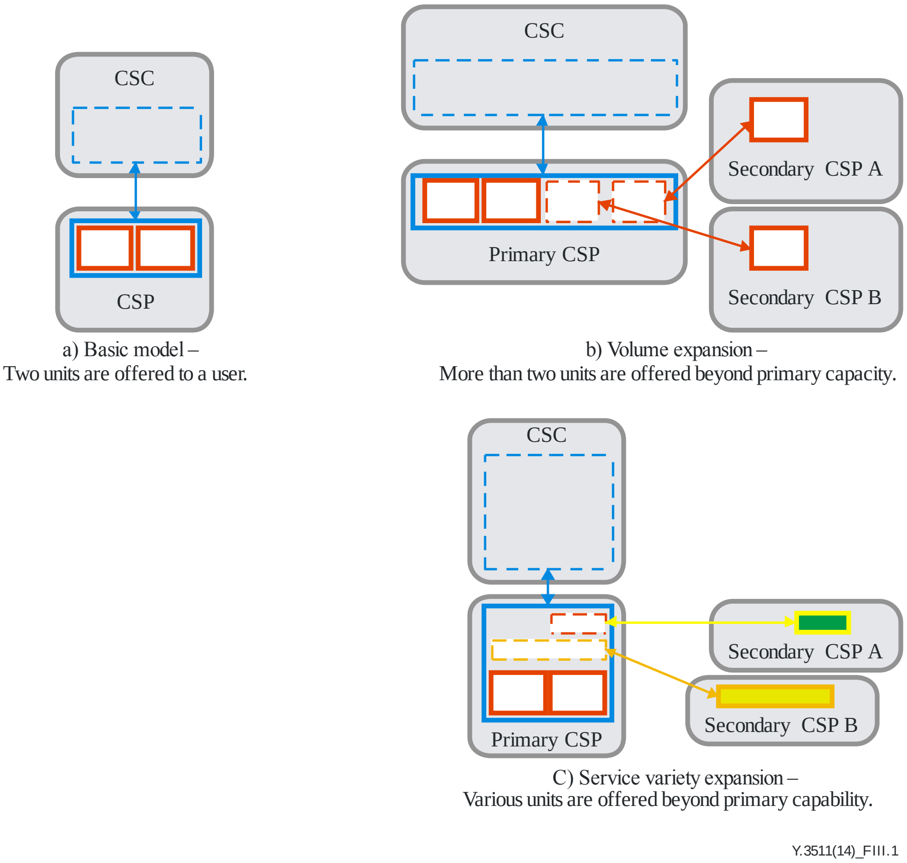

CSC
CSP
a) Basic model –
Two units are offered to a user.
CSC
Primary CSP
Secondary CSP A
Secondary CSP B
b) Volume expansion –
More than two units are offered beyond primary capacity.
CSC
Primary CSP
Secondary CSP A
Secondary CSP B
C) Service variety expansion –
Various units are offered beyond primary capability.
Y.3511(14)_FIII.1

**Figure III.1 – Service item expansion –
Volume and variety expansions in inter-cloud computing**

In the basic model shown in Figure III.1-a, a single CSP offers a service consisting of two resource units. A typical example of such a resource is a virtual machine (VM). The CSP by itself offers two VMs to a given CSC.

In the volume expansion shown in Figure III.1-b, two additional VMs are provided by the secondary CSPs A and B. With their support, the primary CSP can offer a volume-expanded service, which now consists of four VMs.

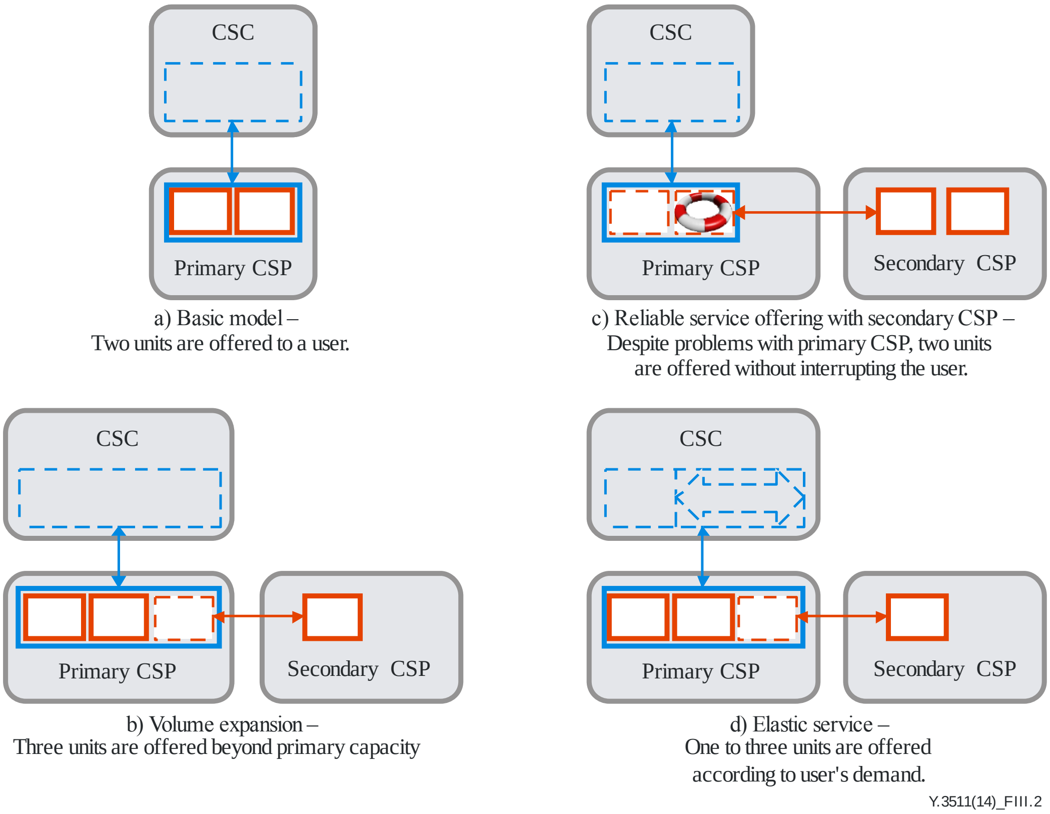
In the service expansion shown in Figure III.1-c, two new resources that are different from the primary CSP's resource are added by secondary CSPs A and B. These might be software packages or platform-type applications. With the support of the secondary CSPs, the primary CSP can offer a wide-variety service, which now consists of various service components.

From the viewpoint of inter-cloud patterns, which are described in clauses 7 and 8, inter-cloud federation is suitable for volume-based service expansion. The description of the inter-cloud federation in clause 8.2 focuses on resource reservation, use and release. Inter-cloud intermediary is suitable for variety-based service expansion. The description of the inter-cloud intermediary in clause 8.3 underlines the significance of the catalogue of service offerings.

III.2 Service operation enhancement

Inter-cloud interaction enables not only service expansion as described in clause III.1, which matters more at the beginning of a service offering, but can also enhance the ways that services are offered. This relates more to the entire process of offering a service.

Figure III.2 shows two such enhancements.



a) Basic model – Two units are offered to a user.

c) Reliable service offering with secondary CSP – Despite problems with primary CSP, two units are offered without interrupting the user.

b) Volume expansion – Three units are offered beyond primary capacity

d) Elastic service – One to three units are offered according to user's demand.

Figure III.2 – Operational enhancement in inter-cloud computing

In Figure III.2, volume expansions are assumed. For the sake of comparison, Figures III.2-a and III.2-b depict item expansion as already discussed in clause III.1, whereas cases c) and d) depict operation enhancement.

The basic scenario is shown in Figure III.2-a again, which offers two resource units (e.g., two VMs) to a user. A simple service item expansion is shown in Figure III.2-b.

With the help of secondary CSPs, the primary CSP can keep offering the service even if something unexpected happens to the primary CSP, which is shown in Figure III.2-c. Due to the availability of cloud technology across multiple CSPs, the secondary CSP can compensate for the unavailable resources by offering alternative resources on behalf of the primary CSP. The primary CSP can offer the same service continuously with minimal or no interruption to the user.

Another scenario, shown in Figure III.2-d is to offer an elastic service, in which the service capacity is adjusted in accordance with the user's demand. In this scenario, the secondary CSP should start and stop its resource offering according to the primary CSP's control. Interaction with the user involves changing the resource offering.

III.2.1 CSC-initiated operation and CSP-initiated operation

Looking at the scenarios in Figure III.2 from the viewpoint of who initiates operations, the reliable service in Figure III.2-c and the elastic service in Figure III.2-d are different.

With the reliable service, the service continuity should be achieved without disturbing the user. The problem should be solved on the CSP side, without necessarily revealing the problem to the user. This may impose a specific requirement. In order not to disturb the user, the primary and secondary CSPs should move the user application, if necessary, and continue the service by themselves. This may include installation and activation of the user application.

The requirement for resource set-up and activation, which is described in clause 9.5, corresponds to this case.

With the elastic service, the CSC may change the use of the CSP's resources explicitly or the primary CSP may change the resource offerings by somehow sensing the CSC's demand.

III.3 Consideration on network connectivity

The description here is meant to supplement clauses 8.2.3 and 8.3.3 on network connectivity.

Networks should, at least, support connectivity between the CSC and the CSP, between the CSPs, and within the CSPs. Based on the primary-secondary model of inter-cloud computing over multiple CSPs, these network parts correspond to:

1) a network between the CSC and the primary CSP;

2) networks between the primary and the secondary CSPs;

3) a network in the primary CSP, and;

4) networks in the secondary CSPs.

From the CSP's perspective, networks 1) and 2) are external, whereas networks 3) and 4) are internal.

Figure III.3 explicitly shows the external networks of 1) as "Network Q" and 2) as "Network X", "Network Y", and "Network Z".

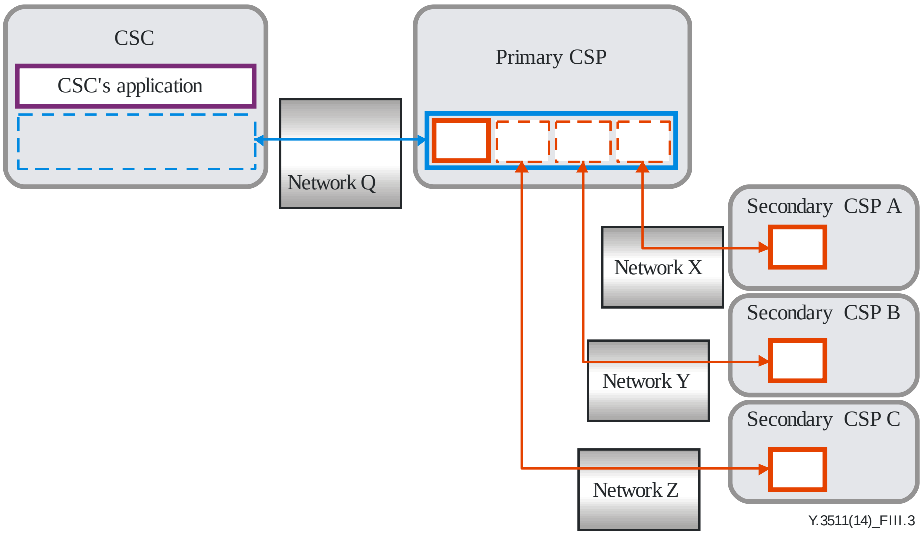


Figure III.3 – Networking in inter-cloud

More networks can be involved as shown in the example provided in Figure III.4.

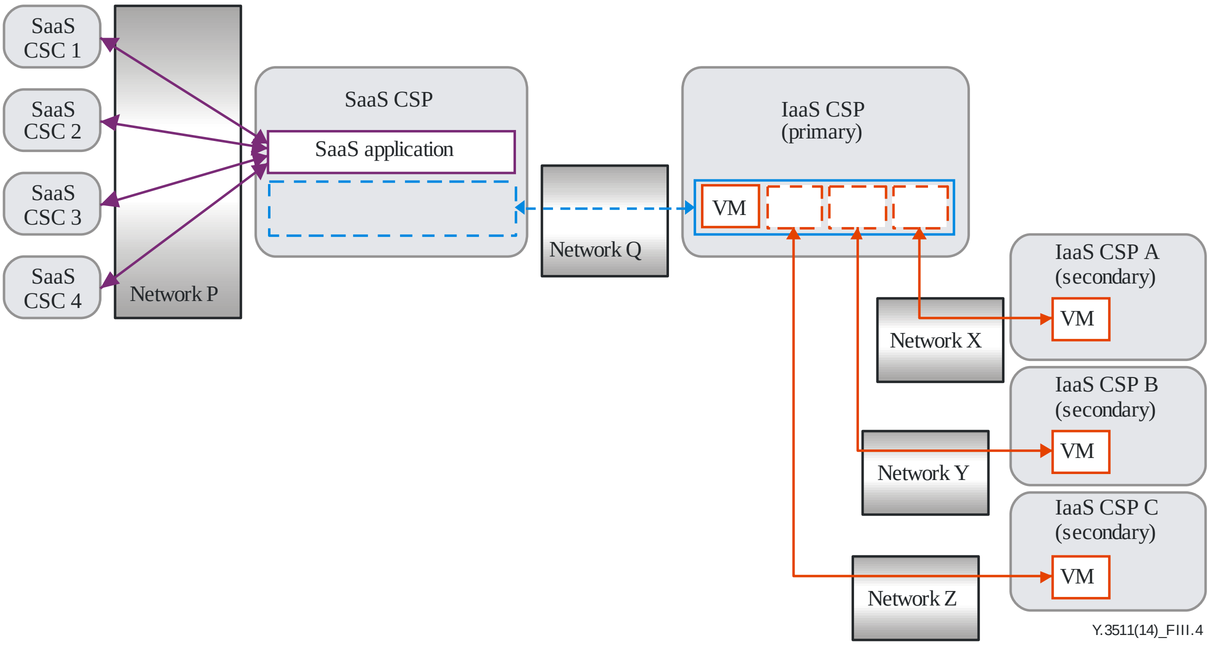


Figure III.4 – Inter-cloud view including SaaS CSCs

In Figure III.4, the IaaS CSP (right side of the figure) provides VMs for the SaaS CSP in order to provide its SaaS application for SaaS CSCs (left side of the figure). In this case, the SaaS CSCs use the SaaS application provided by the SaaS CSP. The SaaS CSP uses VMs, which are provided by IaaS CSPs, on which the SaaS application is executed. Some of the actual VMs are provided by the primary CSP, and some are provided by the secondary CSPs.

Figure III.5 shows the same service offering in a different representation.

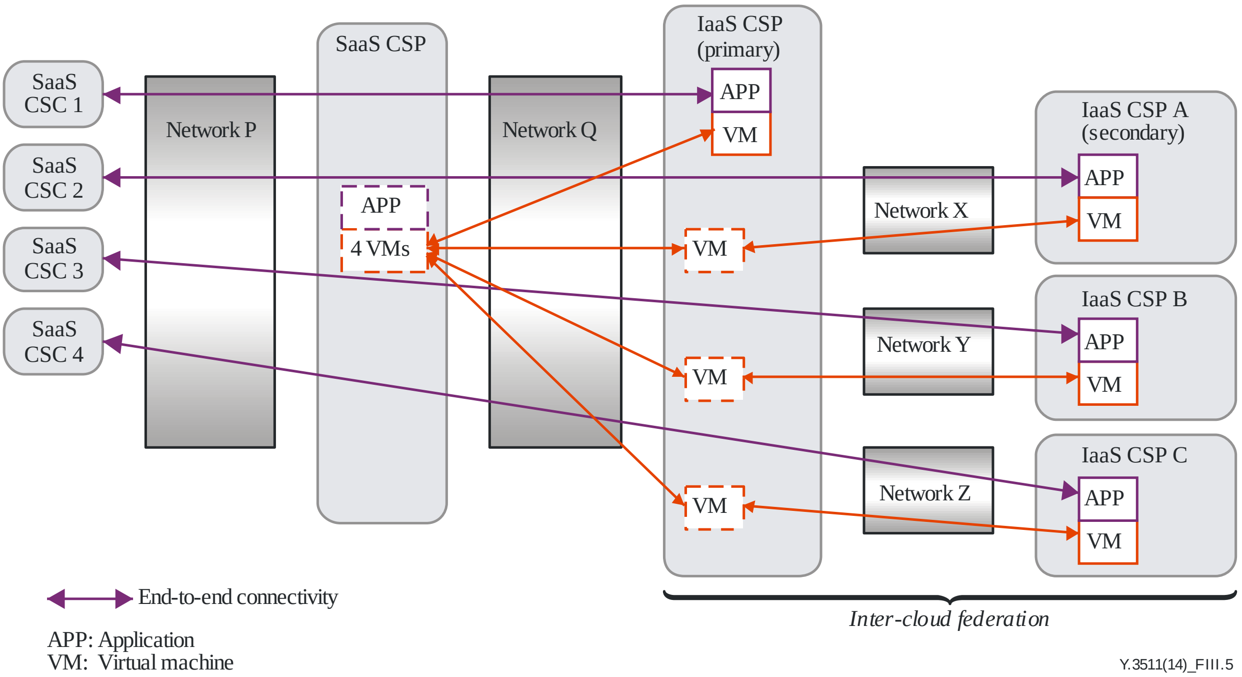


Figure III.5 – View highlighting actual running VMs and application locations (same Figure 8-5)

The SaaS CSCs expect an application service to be provided by the SaaS CSP. In reality, the SaaS CSP relies on IaaS CSP resources and runs its application over the IaaS-CSP's resources. From the IaaS CSP's perspective, the IaaS CSP should, at least:

1) provide VMs, which are now distributed over multiple secondary IaaS CSPs,

2) keep running the application over the distributed VMs, and

3) allow continuous access to the distributed applications from the user (i.e., SaaS CSC users).

In response to the expectations of 1), 2), and 3) above, specific requirements are derived.

The requirements on general resource handling in clauses 9.1, 9.2, 9.3, 9.4, and 9.7 relate to 1).

The requirements on resource set-up and activation in clause 9.5 relate to 2).

The requirements on switchover and switchback of the cloud service user access in clause 9.6 relate to 3).

In a simple implementation, these different networks are designed and operated independently. Such a network configuration is straightforward and easy to operate. However, it may cause inefficient operation, where the user traffic path traverses through an unnecessary route with longer delay. In more sophisticated implementations, the locations of these users, providers and VMs are taken into account and more efficient operation will be achieved.

The capabilities of the networks shown in Figure III.5 may be further offered as NaaS. The detailed requirements and functions for NaaS are under study.

Appendix IV

Inter-cloud security aspects

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

This appendix provides important aspects to be considered regarding inter-cloud security matters.

One important aspect is the multiple and sometimes complicated CSP and CSC inter-cloud relationships such as those described in clause 7 of this Recommendation. For these multiple inter-cloud relationships, appropriate secured mechanisms should be supported during the peer CSPs interactions such as the services request phase (e.g., access control), service usage phase as well as the security of network connectivity between the CSPs.

Other aspects to be considered include:

- Establishment of a trust relationship between CSPs is important given that the multiple CSPs involved in inter-cloud may be administrated by different parties. In case of an inter-cloud federation, the involved CSPs may establish trust relationships among them prior to any interactions between them or during inter-cloud interactions (e.g., service requests between CSPs);
- CSC profiles may be shared among the CSPs involved in the federation. In this case the CSC profile should be handled in a secure manner and in respect of privacy rules and regulations.

Bibliography

[b-ITU-T Y.3510] Recommendation ITU-T Y.3510 (2013), *Cloud computing infrastructure requirements*.

[b-ISO/IEC 20000-1:2011] ISO/IEC 20000-1:2011, *Information technology – Service management – Part 1: Service management system requirements*.

[b-FG Cloud TR-Part 1] FG Cloud TR-Part 1 (2012), *Technical Report: Part 1: Introduction to the cloud ecosystem: definitions, taxonomies, use cases and high-level requirements*, http://www.itu.int/pub/T-FG-CLOUD-2012-P1.

SERIES OF ITU-T RECOMMENDATIONS

Series A	Organization of the work of ITU-T
Series D	General tariff principles
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	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	Terminals and subjective and objective assessment methods
Series Q	Switching and signalling
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 and next-generation networks**
Series Z	Languages and general software aspects for telecommunication systems

Printed in Switzerland
Geneva, 2014