INTERNATIONAL STANDARD

ISO/IEC 22123-2

First edition
2023-09

Information technology — Cloud computing —

Part 2: Concepts

Technologies de l'information — Informatique en nuage —

Partie 2: Concepts

Reference number
ISO/IEC 22123-2:2023(E)

ISO copyright office
CP 401 • Ch. de Blandonnet 8
CH-1214 Vernier, Geneva
Phone: +41 22 749 01 11
Email: copyright@iso.org
Website: www.iso.org

Published in Switzerland

Contents

Page

Foreword

ISO (the International Organization for Standardization) and IEC (the International Electrotechnical Commission) form the specialized system for worldwide standardization. National bodies that are members of ISO or IEC participate in the development of International Standards through technical committees established by the respective organization to deal with particular fields of technical activity. ISO and IEC technical committees collaborate in fields of mutual interest. Other international organizations, governmental and non-governmental, in liaison with ISO and IEC, also take part in the work.

The procedures used to develop this document and those intended for its further maintenance are described in the ISO/IEC Directives, Part 1. In particular, the different approval criteria needed for the different types of document should be noted. This document was drafted in accordance with the editorial rules of the ISO/IEC Directives, Part 2 (see www.iso.org/directives or www.iec.ch/members_experts/refdocs).

ISO and IEC draw attention to the possibility that the implementation of this document may involve the use of (a) patent(s). ISO and IEC take no position concerning the evidence, validity or applicability of any claimed patent rights in respect thereof. As of the date of publication of this document, ISO and IEC had not received notice of (a) patent(s) which may be required to implement this document. However, implementers are cautioned that this may not represent the latest information, which may be obtained from the patent database available at www.iso.org/patents and https://patents.iec.ch. ISO and IEC shall not be held responsible for identifying any or all such patent rights.

Any trade name used in this document is information given for the convenience of users and does not constitute an endorsement.

For an explanation of the voluntary nature of standards, the meaning of ISO specific terms and expressions related to conformity assessment, as well as information about ISO's adherence to the World Trade Organization (WTO) principles in the Technical Barriers to Trade (TBT) see www.iso.org/iso/foreword.html. In the IEC, see www.iec.ch/understanding-standards.

This document was prepared by Joint Technical Committee ISO/IEC JTC 1, *Information technology*, Subcommittee SC 38, *Cloud computing and distributed platforms*.

This first edition of ISO/IEC 22123-2, together with ISO/IEC 22123-1 cancels and replaces ISO/IEC 17788:2014, which has been technically revised.

The main changes are as follows:

- cloud computing terminology has been moved to ISO/IEC 22123-1;
- the descriptions of the key characteristics have been expanded;
- the number and descriptions of the cloud service categories have been expanded;
- the cloud deployment model descriptions have been expanded and corrected;
- added differentiation between cloud computing parties and role;
- the descriptions of the cross-cutting aspects have been expanded;
- a new Clause 8 was added to address data and cloud services concepts;
- a new Clause 9 was added to address virtualization concepts;
- a new Clause 10 was added to address considerations when using multiple CSPs;
- a new Clause 11 was added to address logical and physical organization of cloud computing;
- Annex A was expanded to identify additional cloud service categories, not described in this document.

A list of all parts in the ISO/IEC 22123 series can be found on the ISO and IEC websites.

Any feedback or questions on this document should be directed to the user's national standards body. A complete listing of these bodies can be found at www.iso.org/members.html and www.iec.ch/national-committees.

Information technology — Cloud computing —

Part 2: Concepts

1 Scope

This document specifies concepts used in the field of cloud computing. These concepts expand upon the cloud computing vocabulary defined in ISO/IEC 22123-1 and provide a foundation for other documents that are associated with cloud computing.

This document also provides detailed descriptions on the application of these concepts in cloud computing.

2 Normative references

The following documents are referred to in the text in such a way that some or all of their content constitutes requirements of this document. For dated references, only the edition cited applies. For undated references, the latest edition of the referenced document (including any amendments) applies.

ISO/IEC 22123-1, *Information technology — Cloud computing — Part 1: Vocabulary*

3 Terms and definitions

For the purposes of this document, the terms and definitions given in ISO/IEC 22123-1 and the following apply.

ISO and IEC maintain terminology databases for use in standardization at the following addresses:

— ISO Online browsing platform: available at https://www.iso.org/obp

— IEC Electropedia: available at https://www.electropedia.org/

3.1
PII principal
natural person to whom the personally identifiable information (PII) relates

Note 1 to entry: Depending on the jurisdiction and the particular data protection and privacy legislation, the synonym "data subject" can also be used instead of the term "PII principal."

[SOURCE: ISO/IEC 29100:2011, 2.11]

3.2
PII controller
privacy stakeholder (or privacy stakeholders) that determines the purposes and means for processing personally identifiable information (PII) other than natural persons who use data for personal purposes

Note 1 to entry: A *PII controller* sometimes instructs others [e.g. *PII processors* (3.3)] to process PII on its behalf while the responsibility for the processing remains with the *PII controller*.

[SOURCE: ISO/IEC 29100:2011, 2.10]

3.3
PII processor

privacy stakeholder that processes personally identifiable information (PII) on behalf of and in accordance with the instructions of a *PII controller* (3.2)

[SOURCE: ISO/IEC 29100:2011, 2.12]

4 Symbols and abbreviated terms

API	application programming interface
CaaS	communications as a service
CDN	content distribution network
CompaaS	compute as a service
CPU	central processing unit
CSA	cloud service agreement
CSC	cloud service customer
CSN	cloud service partner
CSP	cloud service provider
CSU	cloud service user
DSA	data sharing agreement
DSaaS	data storage as a service
FaaS	function as a service
IaaS	infrastructure as a service
ICT	information and communication technology
NaaS	network as a service
PaaS	platform as a service
PII	personally identifiable information
PIMS	privacy information management system
PSTN	public switched telephone network
RAM	random access memory
SaaS	software as a service
SLA	service level agreement
SLO	service level objective
SQO	service qualitative objective
TCP/IP	transmission control protocol/internet protocol

TDM	time division multiplexing
VM	virtual machine
VPN	virtual private network

5 Cloud computing foundational concepts

5.1 General

ISO/IEC 22123-1 defines cloud computing and notes that examples of resources include servers, operating systems, networks, software, applications, and storage equipment.

The concepts in this document expand upon the cloud computing vocabulary defined in ISO/IEC 22123-1 and provide a foundation for other documents that are associated with cloud computing.

In this document, a name such as cloud service customer (CSC) or cloud service provider (CSP) represents a cloud computing party while CSC role or CSP role indicates a cloud computing role.

5.2 Key characteristics of cloud computing

5.2.1 General

Subclauses 5.2.2 to 5.2.7 identify and describe key characteristics of cloud computing.

The concept of key characteristics refers to the fundamental properties of cloud computing that differentiate it from other Information Technology paradigms. Each key characteristic covers specific properties that are needed by users of cloud computing.

The key characteristics of cloud computing provide a high-level statement of the distinguishing features of cloud computing. The key characteristics are decomposed in order to understand the concepts of cloud computing for typical delivery scenarios.

The analysis of a key characteristic is not always definitive because the requirements for delivering a cloud service can vary depending on the CSC. All the involved parties in the use and provision of cloud services benefit from a verifiable statement describing what the characteristic means.

5.2.2 Broad network access

Broad network access is a characteristic in which the CSP’s physical and virtual resources are available over a network and accessed through standard mechanisms that promote use the CSC. The focus of this key characteristic is that cloud computing offers an increased level of convenience in that users can access resources from wherever they work, as long as it is network accessible, using a wide variety of devices such as mobile phones, tablets, laptops, and workstations.

Cloud services are widely accessible using network services from a variety of network providers. This can include the public internet, an exchange provider’s network or the CSP’s own network. This characteristic can apply to all cloud deployment models. Access is provided to cloud computing resources at all required times and locations from any CSC, within policy and security constraints.

Broad network access includes accessibility and interoperability for many forms of cloud service network including:

- user (client) access to cloud services;
- application access to cloud services;
- peer cloud service interaction (intra- and inter-cloud); and

— cloud management and control interaction including the use of application programming interfaces (APIs).

5.2.3 Measured service

Measured service is a characteristic in which the metered delivery of cloud services is such that usage can be monitored, controlled, reported, and billed. This is an important feature needed to optimise and validate the delivered cloud service. The focus of this key characteristic is that the customer only pays for the resources that they use. From the customers' perspective, cloud computing offers the users value by enabling a switch from a low efficiency and asset utilization business model to a high efficiency one.

Measured service can refer to a wide variety of metering functions that can be required for service operations, administration, maintenance, provisioning, and security. Consumption-based billing requires that cloud service use be measured using an agreed upon measuring algorithm which can be specified in a service level agreement (SLA) (see ISO/IEC 19086). Metered cloud services provide sufficient detail to meet cloud SLA requirements. This can include measurements for the underlying virtual and physical resources (see ISO/IEC TR 23613[15]).

5.2.4 Multi-tenancy

Multi-tenancy is a characteristic in which physical or virtual resources are allocated in such a way that multiple tenants and their computations and data are isolated from and inaccessible to one another. Typically, and within the context of multi-tenancy, the group of cloud service users (CSUs) that form a tenant all belong to the same CSC. Some cloud computing deployments, particularly public cloud and community cloud, can have a group of CSUs that are from multiple different CSCs. However, a given CSC can have many different tenancies with a single CSP representing different groups within the organization such as by department, division, or subsidiary. In some cases, this is for internal security and confidentiality. In other cases, it can be for regulatory compliance reasons. This can require identity and access management.

5.2.5 On-demand self-service

On-demand self-service is a characteristic in which a CSC can provision cloud services, as needed, automatically or with minimal interaction with the CSP. The focus of this key characteristic is that cloud computing offers users a relative reduction in costs, time, and effort needed to take an action, since it grants the user the ability to do what they need, when they need it, without requiring additional human user interactions or overhead.

The cloud services can be provisioned and configured by the CSC without operator interaction with the CSP. For example, changing the random access memory (RAM) available or disk space available can be done without human intervention.

5.2.6 Rapid elasticity and scalability

Cloud services can be rapidly and elastically adjusted, in some cases automatically, to quickly increase or decrease capacity. For the CSC, the resources of cloud services available for provisioning often appear to be unlimited and can be purchased in any quantity at any time, subject to constraints of service agreements. From the perspective of the CSC, there is no longer a concern about limited resources or possibly capacity planning.

There are two possible directions for scalability with respect to cloud computing. Horizontal scaling is the term used for scalability where more instances of a given resource are allocated [e.g. running more virtual machines (VMs) or containers in parallel, each running an instance of the same application]. Vertical scaling is when an increase is made in the size of a resource allocated to a cloud service, for example when the amount of RAM or the number of central processing units (CPUs) allocated to a single virtual machine is increased, or the storage capacity of a single storage resource is increased. This can sometimes necessitate some delay while new capacity is added to an existing resource, in contrast to horizontal scaling which often has less latency. For a full description of elasticity and scalability, see ISO/IEC TS 23167[11].

The CSP describes the cloud services scalability features including any associated latency and any limitations. The CSC determines that the cloud service's scalability features, associated latency and limitations meet its requirements based on the CSP's description.

To the CSC, the resources available to a cloud service can be increased or decreased by any amount at any time, subject to any limitations imposed by the CSP or according to the pre-arranged policies in a cloud SLA. For the detailed information, refer to ISO/IEC 19086-1[3].

5.2.7 Resource pooling

Resource pooling is a characteristic in which a CSP's physical or virtual resources can be aggregated to serve one or more CSCs. CSPs are able to support multi-tenancy while also using abstraction to mask the complexity of the process from the CSC.

From the CSC's perspective, all they know is that the service works; they generally have no control or knowledge over how the resources are being provided or where the resources are located. This offloads some of the CSC's original workload, such as maintenance requirements, to the CSP.

Specifying a location at a higher level of abstraction is also possible in some environments.

Resources of a similar type (e.g. compute or storage) can be pooled in support of cloud service provision, but resources of different types cannot be pooled. The CSC can stipulate that the cloud resources are not shared by multiple CSCs or by multiple tenants.

Resource pooling includes but is not limited to:

— Two or more share cloud resources from a common resource pool.

— Two or more tenants share cloud resources from a common pool, using a multi-tenant model, regardless of how many CSCs are served.

The cloud service can appear to the CSC to be location independent because the CSC generally has no control or knowledge of the precise geographical location where the cloud service is being run. However, CSCs can generally specify a location for their instances of the cloud service at an abstract level.

5.3 Cloud capabilities types

A cloud capabilities type is a classification of the functionality provided by a cloud service to the CSC, based on the resources used. Cloud capabilities types follow the principle of separation of concerns, i.e. they have minimal functionality overlap between each other.

The cloud capabilities types are:

— application capabilities type: A cloud capabilities type in which the CSC can use the CSP's applications;

— infrastructure capabilities type: A cloud capabilities type in which the CSC can provision and use processing, storage or networking resources;

— platform capabilities type: A cloud capabilities type in which the CSC can deploy, manage and run customer-created or customer-acquired applications using one or more programming languages and one or more execution environments supported by the CSP.

NOTE - In this context "applications" includes scripts, containers, complete programs, partial programs, code and function libraries, microservices, AI training data, and other forms of compliable or executable software. There are only three cloud capabilities types defined in this document. These cloud capabilities types should not be confused with other categorizations of cloud services.

5.4 Cloud service categories

5.4.1 General

A cloud service category is a group of cloud services that possess some common set of qualities. A cloud service category can include capabilities from one or more cloud capabilities types.

The primary determining factors for categorizing a cloud service are:

- the cloud computing capabilities types that are provisioned (application, platform or infrastructure);
- its intended use.

Cloud service categories are typically referred to using *something* “as a service.”

The three best known cloud service categories are:

- software as a service (SaaS), which offers application capabilities types (5.4.2);
- platform as a service (PaaS), which offers platform capabilities types (5.4.3);
- infrastructure as a service (IaaS), which offers infrastructure capabilities types (5.4.4).

However, there are many other examples of cloud service categories. One example often used in the telecom industry is network as a service (NaaS) (5.4.5) which offers networking-related application, platform or infrastructure capabilities types.

Some cloud service categories can offer two or all three of the cloud capabilities types. For example, communications as a service (CaaS) (5.4.6) can offer both platform and application capabilities types (see Annex A for more examples).

5.4.2 Software as a service (SaaS)

SaaS is a cloud service category in which the cloud capabilities type (5.3) provided to the CSC is an application capabilities type.

The cloud service provisioned for the CSC uses the CSP’s software application running on CSP resources. The use and provision of the cloud service category are in accordance with the cloud service agreement and its associated cloud SLAs. The applications are accessible from various CSC devices through either a thin client interface, such as a web browser (e.g. web-based email), or an Application Programming Interface (API). The customer does not manage or control the underlying resources including network, servers, operating systems, storage, or even individual application capabilities, with the possible exception of limited user-specific application configuration settings.

The CSP providing the SaaS product is typically responsible for making all aspects of the software service available including deploying, configuring, maintaining and updating the operation of the software applications on the CSP resources. It is worth noting that the entity responsible for making the service available can be different from the SaaS application developer.

Note that some SaaS services are extensible in that they include limited customer scripting or other code execution within their own functionality, however the execution of such code is not central to the service being offered.

5.4.3 Platform as a service (PaaS)

PaaS is a cloud service category in which the cloud capabilities type (5.3) provided to the CSC is a platform capabilities type.

The capability provided to the CSC is to develop or deploy onto the CSP resources customer-created or acquired applications created using programming languages, libraries, services, and tools supported by the CSP. This capability does not necessarily preclude the use of compatible programming languages, libraries, services, and tools from other sources.

The term "platform" in the PaaS context refers to a development or deployment platform for cloud-enabled applications. The term "platform" is broadly used in the computing industry. It therefore helps to understand the context of the term regarding PaaS. PaaS is distinguished from an extensible SaaS or web application by its primary customers: developers and operations staff versus end users.

The CSC does not manage or control the underlying resources including network, servers, operating systems, or storage, but has control over the deployed applications and possibly configuration settings for the application-hosting environment. The CSC can verify the service meets the requirements in accordance with the cloud service agreement and its associated SLAs. The CSUs of the CSC primarily design, implement, and deploy applications into the cloud computing environment.

5.4.4 Infrastructure as a service (IaaS)

IaaS is a cloud service category in which the cloud capabilities type (5.3) provided to the CSC is an infrastructure capabilities type.

The capability provided to the CSC is to provision physical or virtual processing, storage, networks, and other fundamental computing resources where the customer is able to deploy and run arbitrary software, which can include operating systems and applications. The CSC does not manage or control the underlying resources but has control over operating systems, storage, and deployed applications. They possibly have limited control of select networking components (e.g. host firewalls).

An IaaS service provides, for example, hosting of CSC-defined virtual machine images on a CSP-provided and operated hypervisor (see Clause 9). Because, in this example, each VM runs directly on virtualised hardware, there are fewer limits on the software choices available to the CSC, however this flexibility comes at the cost of requiring management and maintenance of all the software components they select and deploy.

The CSCs can create, install, monitor, and manage applications deployed in an IaaS cloud service. The CSC can verify the service meets their requirements in accordance with the cloud service agreement and its associated SLAs.

The terms "software" and "application" in the IaaS context refers to software and applications sourced and deployed by the CSC and which remain under the control of the CSC. It is typical that the CSP is unaware of what this software is and has no control over it. The term "arbitrary software" in this context means that the CSC can deploy and run any type of software, subject only to any limitations imposed by the nature of the environment made available by the cloud service.

5.4.5 Network as a service (NaaS)

NaaS is a cloud service category in which the capability provided to the CSC is transport connectivity and related network capabilities. NaaS can provide the application, platform and infrastructure cloud capabilities types.

The capability provided to the CSC is to provision and manage physical or virtual network connections. The CSC does not manage or control the underlying physical network infrastructure, but the CSC can control the creation, management and removal of network connections between their own choice of endpoints.

Note that these network connections can be quite sophisticated and can include the use of complex network resources such as physical or virtual switches, routers, transmission links, satellite uplinks and transponders, content distribution networks (CDNs), caches, proxies, firewalls, redundant links, relays, repeaters, multiplexors, or other network resources.

Note that, while the NaaS itself operates as a cloud service, the networks that it manages can include non-IP networking technologies such as time division multiplexing (TDM) connections, optical connections, or satellite transmission links. For example, a TV broadcaster can employ a NaaS to establish a high-definition television contribution[1] link over terrestrial and satellite resources from a remote location to their studio for coverage of a sporting event.

A NaaS service is primarily concerned with the provision of connectivity services (including related network capabilities) between endpoints. A NaaS can offer infrastructure capabilities type functions, such as raw connectivity and can offer platform capabilities type functions, such as the ability to execute customer code for dynamic routing, call queuing or redirection, or specialist CDN, caching or firewall behaviours. A NaaS can offer application capabilities type functions, by offering CSP-built network-oriented applications that simplify or coordinate management of the connections.

The network connections it controls need not be Internet or cloud-related. In the context of the definition of NaaS, "transport" is not to be assumed to mean the TCP/IP transport layer. It can refer to any kind of network connection, using any network protocol or none. For example, a NaaS can be used to allocate capacity on a satellite transponder for a regional television broadcast, or to place or route telephone calls through the public switched telephone network (PSTN).

5.4.6 Communications as a service (CaaS)

CaaS is a cloud service category in which the capability provided to the CSC is real-time interaction and collaboration.

For the purposes of CaaS, "real-time" refers to sufficiently low latency in communication that it occurs without perceivable delay for the average human user.

The capability provided to the CSC is to enable real-time communications between humans or with software entities that act like humans (e.g. bots). The CSC does not manage or control the underlying network connections or communications features.

CaaS services can include, but are not limited to the following:

- text chat;
- screen sharing (whole or partial);
- spoken voice;
- video;
- file transfer;
- presence[2];
- shared editing; or
- other communication modes.

CaaS services typically support calls between more than two endpoints, typically referred to as conferencing.

CaaS services can permit storage of shared files.

CaaS services can permit recording of communication for later reference.

CaaS services can interface with a non-cloud communications system such as the PSTN, and can use a NaaS to control the interaction between the two.

[1] In television, "contribution" is the format in which companies transfer broadcasts and materials between companies and locations, as opposed to "distribution" which is the format used in transmission to local broadcasters and to the public. Contribution links are typically much higher in quality, bandwidth, and fault tolerance.

[2] Presence means the indication to others of a user being online, available, away, busy, their location. etc.

CaaS application can include complex functionality such as private branch exchange/Centrex features, call routing, queuing, and connection security functions.

5.4.7 Compute as a service (CompaaS)

CompaaS is a cloud service category in which the capability provided to the CSC is the provision and use of processing resources needed to deploy and run arbitrary software.

Some software needs capabilities other than processing resources to run, however these capabilities are often obtained by employing other cloud services such as data storage as a service outside of the CompaaS service itself, often but not always from the same CSP.

A VM-based cloud service where all storage is external to the cloud service can qualifies as CompaaS.

The primary stakeholders are information and communication technology (ICT) operators or developers creating, installing, monitoring, and managing software executing in a CompaaS cloud service.

CompaaS is distinctly different from the underlying processing resources. The compute service can optionally include a pre-installed operating system and other support for software and applications. The term “arbitrary software” in this context means that the customer can deploy and run any type of software, subject only to any limitations imposed by the nature of the environment made available by the cloud service. If the service includes data storage functions or anything more than basic connectivity in addition to processing, it is usually better categorised as an IaaS.

5.4.8 Data storage as a service (DSaaS)

DSaaS is a cloud service category in which the capability provided to the CSC is the provision and use of data storage and related capabilities.

These capabilities can range from the most basic storage of files or blocks of raw binary data, up through relational databases, up to advanced big data platforms. The storage can be relatively simple in a single location, or can offer automatic backup, disaster recovery, geographical redundancy, advanced failover, and other more sophisticated features. See ISO/IEC TS 23167 and ISO/IEC 27040[23].

For a basic file storage service, the DSaaS exhibits only infrastructure capabilities type functions. For a more sophisticated storage service such as a relational database, the DSaaS can support execution of scripts and software with exposed APIs, thus also exhibiting platform capabilities type functions.

The DSaaS often also provides a user interface for database administrators and storage users, thus exhibiting some application capabilities type functions. A DSaaS is often employed alongside other cloud service categories such as CompaaS, IaaS, or PaaS, providing secure and reliable storage for CSC code that is running on those services.

The primary stakeholders are end users, developers, and ICT Operations roles creating and maintaining cloud data storage structures, uploading, downloading and manipulating data, and managing the data stored in the DSaaS. The CSC can verify the cloud service category provided meets the requirements as stipulated in accordance with the cloud service agreement and its associated SLAs.

5.5 Cloud deployment models

5.5.1 General

A cloud deployment model represents the way in which cloud computing can be organized based on the control and sharing of physical or virtual resources.

Cloud deployment models illustrate the basic patterns for interaction between CSCs and CSPs. Cloud deployment models do not specify implementation details or how resources are configured.

5.5.2 Private cloud deployment model

A private cloud is a cloud deployment model in which the cloud services are used exclusively by a single CSC and the underlying resources are controlled by that CSC.

The CSC can authorize access to other parties for its benefit.

A private cloud can be owned, managed, or operated by the CSC itself or a third party and can exist on premises or off premises.. When a private cloud is hosted off premises, the cloud resources can be owned, managed, and operated by the CSP.

Private clouds seek to set a narrowly controlled boundary around the private cloud based on limiting the CSUs to a single CSC. The CSC can provide stricter control over security and data protection, assurance over data location, and potential simplification of legal requirements across jurisdictions. A private cloud gives a single CSC the exclusive access to and usage of the cloud services and related infrastructure.

Another term for "off premises private cloud" is “dedicated cloud."

A private cloud can operate as a multi-tenant deployment and support multiple groups of CSUs, with logical partitioning between the CSUs. Many CSCs can choose to partition their users and associated data and functions into isolated tenant groups for business, legal, or compliance reasons.

A private cloud gives a single CSC the exclusive access to and usage of the cloud services and related infrastructure, and thereby it allows the CSC much greater control over data, underlying systems and applications. The CSC can verify the service meets their requirements with the private cloud in their control. The CSC can determine the size (and hence the cost) of the reserve capacity needed to guarantee that the desired level of rapid elasticity and scaling.

Clouds require substantial surplus idle capacity, or significant pre-emptible and reserve workloads, to provide elasticity and scalability. The cost of that surplus capacity has to be borne somehow. Usually, in the case of public cloud deployments, this is directly handled by the CSP and then amortized across their CSCs. If the surplus capacity underpinning a private cloud is potentially used elsewhere, the mechanisms protecting those reserve private cloud resources from the consequences of their prior or subsequent use outside that private cloud can be communicated to the CSC.

When a CSP hosts a dedicated private cloud off premises for a CSC, the cloud resources are owned, managed, and operated by that CSP. From a CSC perspective, many aspects of this off premises private cloud are very similar to those of a public cloud. The implication of this is that combining an on premises private cloud with an off premises private cloud has similar challenges to that of creating a hybrid cloud (see 5.5.5) that combines a private cloud with a public cloud.

Figure 1 and Figure 2 illustrate an example of an on premises private cloud and an off premises private cloud hosted by a CSP, respectively.

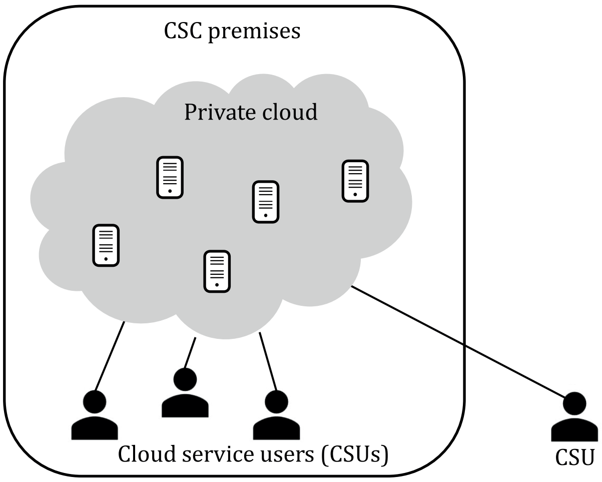

Figure 1 — Example of an on premises private cloud

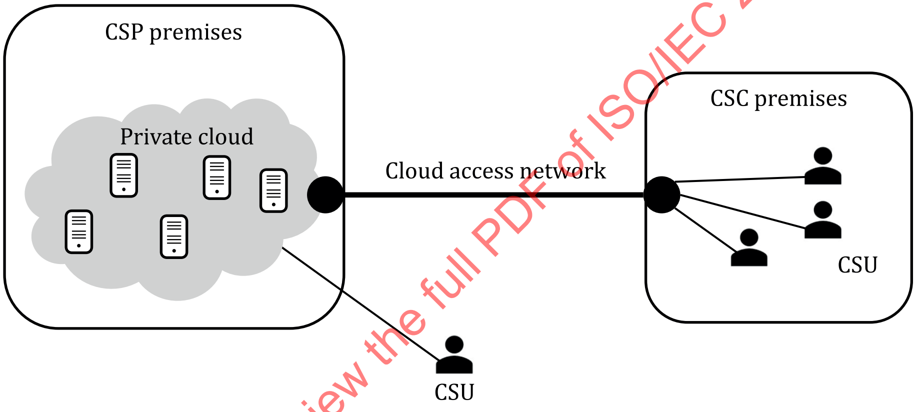

Figure 2 — Example of an off premises private cloud

5.5.3 Public cloud deployment model

A public cloud is a cloud deployment model in which cloud services (and resources) are potentially available to (almost) any CSC and in which the resources are potentially made available to any CSC, typically over a public network. Figure 3 presents a simple view of a public cloud and its customers.

A public cloud can be owned, managed, and operated by a business, academic, or government organization, or some combination of them. Public clouds can also be owned by the CSP providing the cloud services and located on the premises of the CSP. Actual availability for specific CSCs can be subject to jurisdictional regulations.

Public clouds have very broad boundaries, where CSC access to public cloud services has few, if any, restrictions. While the CSP can limit access to a cloud service, the CSC has no control and no visibility over the use of the public cloud service and no control over which other CSCs and their CSUs are also users of the same cloud service.

One characteristic of cloud services is that two or more CSCs can share the resources associated with the cloud service.

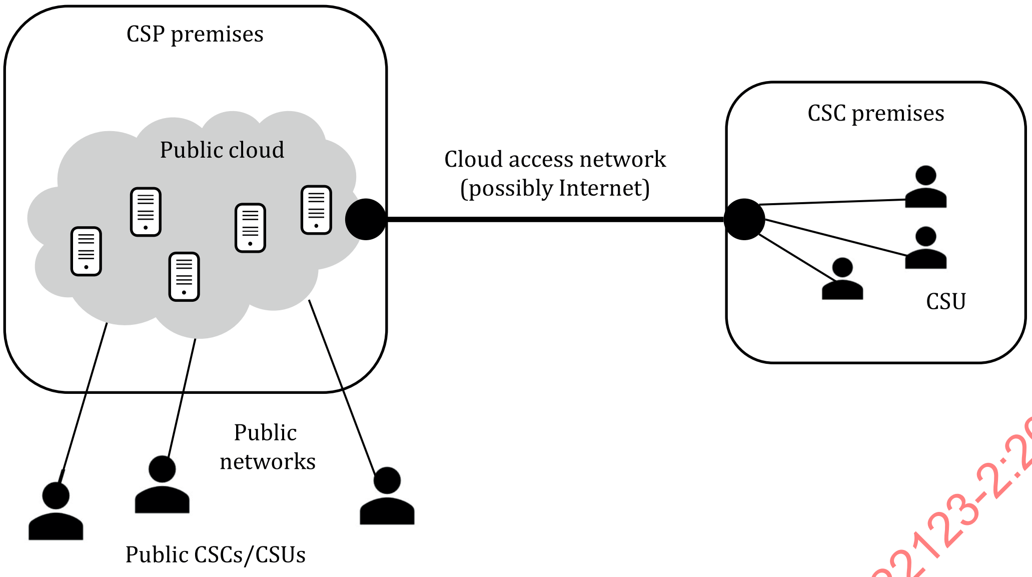

Figure 3 — Example of a public cloud

5.5.4 Community cloud deployment model

A community cloud is a cloud deployment model in which the cloud services exclusively support and are shared by a specific collection of CSCs who have shared requirements and a relationship with one another, and where resources are controlled by at least one member of this collection. Figure 4 depicts an on premises community cloud comprised of a number of participant organizations.

A community cloud can be owned, managed, and operated by one or more of the organizations in the community, a third party, or some combination of them, and it can exist on or off premises.

Community clouds limit participation to a group of CSCs who have a shared set of concerns, in contrast to the openness of public clouds, while community clouds have broader participation than private clouds. These shared concerns include, but are not limited to, mission, information security requirements, policy, and regulatory compliance considerations.

As Figure 4 illustrates, a CSC can access the local cloud resources, and also the resources of other participating organizations, through the connections between the associated organizations.

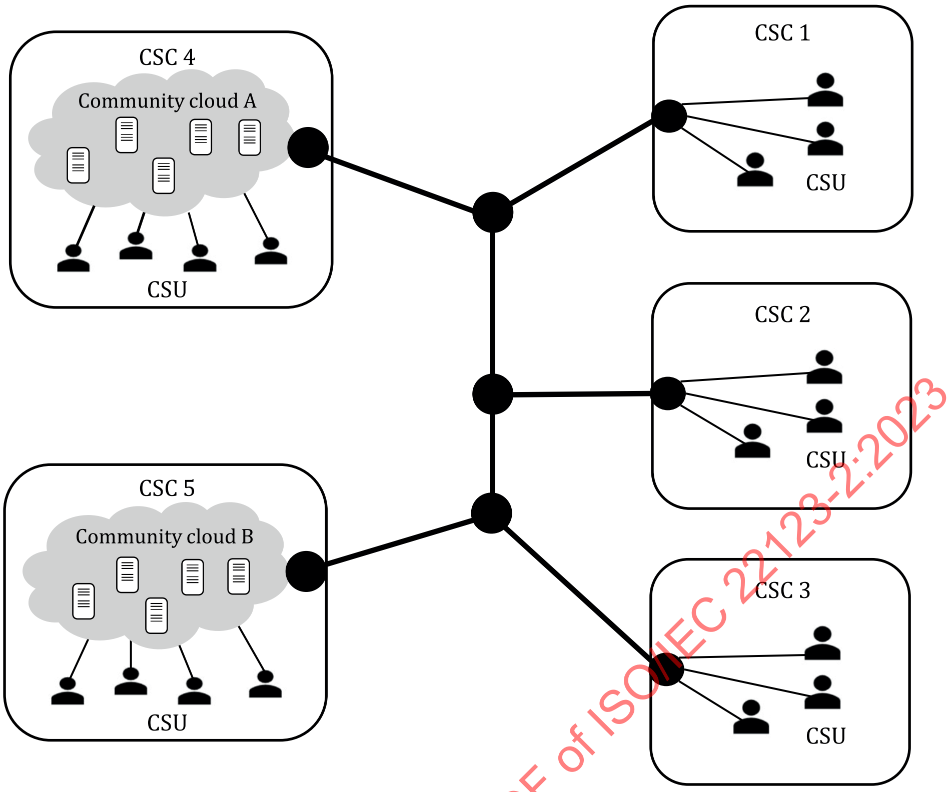

Figure 4 — Example of an on premises community cloud

Figure 5 shows an off premises community cloud, where the server side is off premises. In this case, an off premises community cloud builds its infrastructure off premise and serves a set of organizations that request and consume cloud services. In addition, various combinations of Figure 4 and 5 are also suitable for use as a community cloud.

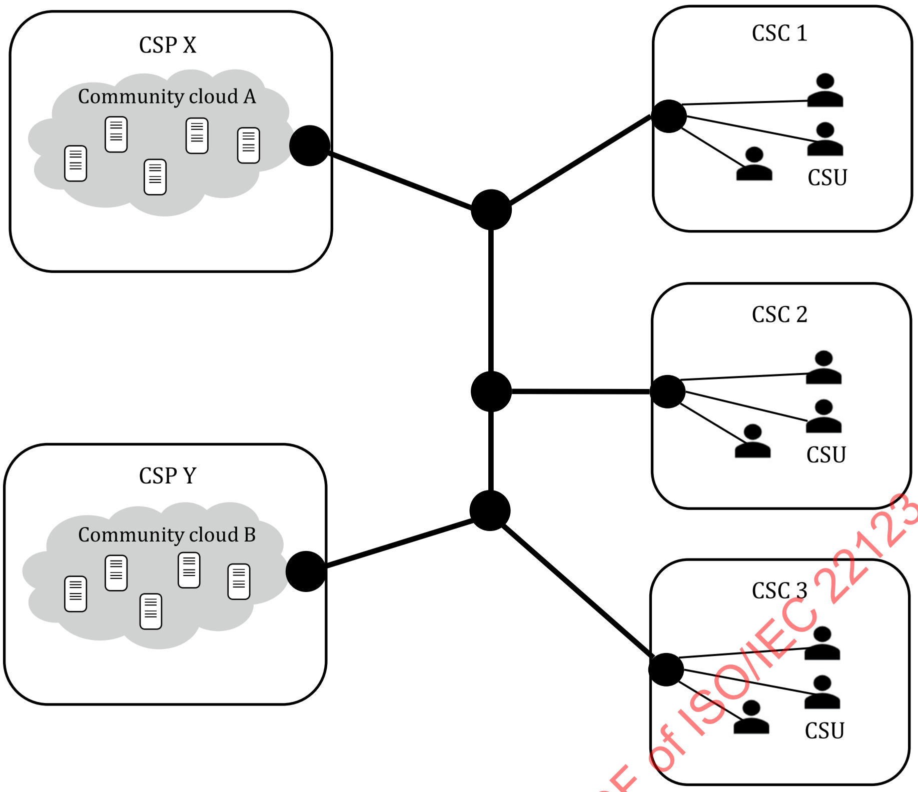

Figure 5 — Example of an off premises community cloud

The community of CSCs verifies its membership. The CSP verifies that only authorized CSUs have access to the cloud services.

The CSP organization's cloud resources can be operated by one or more of the organizations in the community or a third-party. Community clouds generally get the cost benefits of a public cloud while providing heightened privacy, security and regulatory compliance. A cloud auditor can conduct independent assessment of cloud services to verify the scope of the group and verify that the service and underlying infrastructure are exclusive to the group and its users.

5.5.5 Hybrid cloud deployment model

A hybrid cloud is a cloud deployment model which uses a private cloud and a public cloud. Figure 6 presents a simple example of a hybrid cloud.

The individual clouds involved remain unique entities but are bound together by appropriate technology that enables interoperability, data portability and application portability. A hybrid cloud can be owned, managed, and operated by the organization itself or a third party and can exist on premises or off premises. Hybrid clouds represent situations where interactions between two different deployments are needed but remain linked via appropriate technologies. As such the boundaries set by a hybrid cloud reflect its two base deployments.

In practice, a hybrid cloud can be a composition of a private cloud and a public cloud that remain as distinct entities but are bound together by standardized or proprietary technology that enables data and application portability.

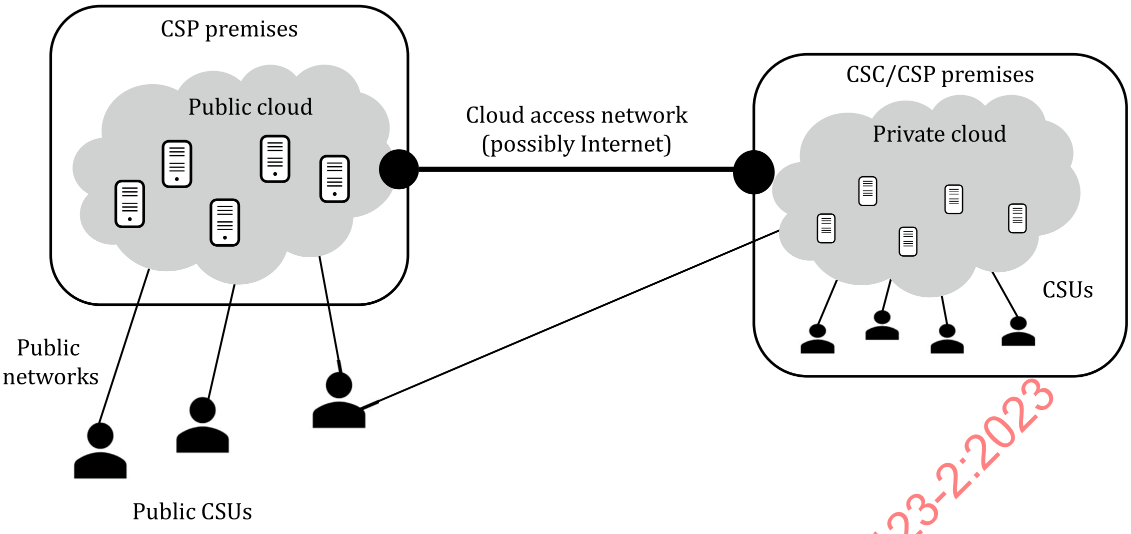

Figure 6 — Example of hybrid cloud

A CSC can combine cloud services from an on premises private cloud and a public cloud and can be using cloud services from two or more CSPs. It is possible that the CSPs are not aware of the CSC's hybrid cloud or the other CSPs.

6 Cloud computing parties and roles

6.1 Cloud computing parties

Parties are entities that play roles (see 6.2). A party can play more than one role at any given point in time and can only engage in a specific subset of activities of that role.

Within the context of cloud computing, it is often necessary to differentiate requirements and issues for certain parties. The major parties of cloud computing are:

- Cloud service customer (CSC): A party which is in a business relationship with cloud service provider(s) for the purpose of using cloud services. This party can also have a business relationship with one or more cloud service partner for a variety of purposes.
- Cloud service partner (CSN): A party which is engaged in support of, or auxiliary to, activities of either the cloud service provider or the cloud service customer, or both.
- Cloud service provider (CSP): A party which is in a business relationship for the purpose of providing cloud services. The business relationship is with a cloud service customer or a cloud service partner.

6.2 Cloud computing roles

6.2.1 General

The major roles of cloud computing are cloud service customer role (CSC role), cloud service partner role (CSN role), and cloud service provider role (CSP role). These roles can be further organized into sub-roles (see ISO/IEC 22123-3[7]).

Roles and sub-roles, in turn, are sets of activities (see ISO/IEC 22123-3). All cloud computing activities can be categorized into three main groups: activities that use cloud services, activities that provide cloud services and activities that support cloud services.

As an example, organizations operate an internal private cloud and also use cloud services; they are both a provider of cloud services as well as a customer of the same or different services. The organization therefore serves both as a CSP and as one or more CSCs.

Another example can involve an organization that has multiple independent divisions, with each serving as a CSC. This leads to a single organization acting as multiple CSCs at the same time.

Scenarios can arise in which a CSC chooses to separate the activities of the CSC role such that administration and management is centralized while user control and users are distributed. In this configuration, not all CSC sub-roles are implemented in each CSC role within a CSC.

6.2.2 Cloud service customer role

The CSC role includes activities for the purpose of using cloud services. The CSC role can include ensuring the smooth operation of the cloud services, acquisition and use of cloud services, and integration of cloud services with a CSC's existing ICT systems. Specific descriptions of sub-roles for the CSC role and associated activities are described in ISO/IEC 22123-3.

6.2.3 Cloud service provider role

A CSP is a party that is acting in the cloud service provider role (CSP role). The CSP is in a business relationship with a CSC or a CSN for the purpose of providing cloud services.

The CSP role is a set of activities that make cloud services available. The CSP role focuses on activities necessary to provide a cloud service and activities necessary to ensure its delivery to the CSC as well as cloud service maintenance. The CSP role includes an extensive set of activities including providing services, deploying and monitoring services, managing business plans, providing audit data, etc. as well as numerous sub-roles, e.g. business manager, service manager, network provider, security and risk manager, etc. Specific descriptions of sub-roles for the CSP role and associated activities are described in ISO/IEC 22123-3.

6.2.4 Cloud service partner role

The CSN is a party that is acting in the cloud service partner role (CSN role).

The CSN role is a set of activities that support, or are auxiliary to, either the CSP role or the CSC role, or both. Activities of a CSN role vary depending on the type of partner and their relationship with the CSP role and the CSC role (see ISO/IEC TR 23187[13]). Specific descriptions of sub-roles for the CSN role and associated activities are described in ISO/IEC 22123-3.

7 Cloud computing cross-cutting aspects

7.1 General

In the context of cloud computing, cross-cutting aspects include both architectural and operational considerations. Cross-cutting aspects are behaviours or capabilities which are coordinated across roles and implemented consistently in a cloud computing system. Such aspects can impact multiple roles, activities, and functional components (a functional building block needed to engage in an activity, backed by an implementation), in such a way that it is not possible to clearly assign them to individual roles or functional components, and thus become shared issues across the roles, activities and functional components.

Cross-cutting aspects often affect the cloud computing activities performed by roles. Roles can coordinate supporting a cross-cutting aspect amongst themselves and their cloud computing activities. Supporting cross-cutting aspects also needs functional components to provide support for the cloud computing activities, technical capabilities and implementations.

Many of these cross-cutting aspects, when combined with the key characteristics of cloud computing, represent good reasons for using cloud computing. However, cross cutting aspects like security, protection of PII, and governance have been identified as major concerns and in some cases an impediment to the adoption of cloud computing.

For each cross-cutting aspect a set of cloud computing activities and functional components are defined to support them. Different roles and solutions can use different subsets of these.

The following clauses describe key cross-cutting aspects of cloud computing. Many of these cross-cutting aspects are also the subject of one or more other standards, which are referenced in the relevant clauses.

7.2 Auditability

Auditability is the capability of collecting and making available necessary evidential information related to the operation and use of a cloud service, for the purpose of conducting an audit.

Auditability is concerned with the capability of being able to gather necessary evidential information related to the operation and use of a cloud service. Any issues with achieving this capability jeopardize the completion of audits (see ISO/IEC TR 3445[1]). An audit evaluates the evidence objectively to determine the extent to which the audit criteria are fulfilled. Audits are generally performed to check that processes and systems meet requirements (e.g. requirements of a standard) and thus provide assurance such processes and systems are operating to meet expectations (typically expectations of a governing body or of senior management).

Audit evidence is data gathered relating to the processes and systems being audited – the precise set of data required varies depending on the type of audit concerned. In general, the data relates to the processes involved and to the usage, the environment, the performance of cloud services and their associated resources. The data can take various forms but generally includes records and logs related to the use and provision of cloud services, potentially from both the CSC and from the CSP. These records and logs need to be collected and maintained in a secure manner.

Audits can be performed on the CSC's use of cloud services, or they can be performed on the CSPs provision of cloud services. An audit is often associated with the CSP provision of cloud services, as a means for the CSP to advertise to CSCs that their cloud services meet significant requirements, e.g. for information security, for PII protection, etc.

Another use of audits is to verify that the CSP is providing the cloud services in line with the cloud service agreement (CSA) and its associated cloud SLA, as described in ISO/IEC 19086.

7.3 Availability

Availability, which is the ability to be in a state to perform as required, depends upon the combined characteristics of the reliability, recoverability, and maintainability.

Availability is usually a significant cloud service level objective (SLO) in the cloud SLA for a cloud service, as described in ISO/IEC 19086-1. The typical availability SLO is “the amount or percentage of time in a given period that the cloud service is accessible and usable.” Availability metrics can include some element of planned (or *allowable*) downtime (e.g. for maintenance purposes).

Availability is also a key aspect of information security relating to cloud services, as described in ISO/IEC 27017[19].

7.4 Governance

Governance is the system by which the provision and use of cloud services are directed and controlled by the decision-makers of the organisation (CSP or CSC). Cloud governance is cited as a cross-cutting aspect because of the requirement for transparency and the need to rationalize governance practices with SLAs and other contractual elements of the CSC to CSP relationship. The term internal cloud governance is used for the application of design-time and run-time policies to ensure that cloud computing based solutions are designed and implemented, and cloud computing based services are delivered, according to specified expectations. The term external cloud governance is used for cloud service agreement between the CSC and the CSP concerning the provision of cloud services by the CSP and the use of cloud services by the CSC.

Governance of cloud services has no specific standard, but ISO/IEC 38500[30] applies to ICT systems in general and so covers both the use and the provision of cloud services. Central to governance are the creation and maintenance of policies which cover the use and provision of cloud services and the effective monitoring and reporting in relation to cloud services.

7.5 Interoperability

Cloud interoperability is the ability of a CSC's system to interact with a cloud service, or the ability for one cloud service to interact with other cloud services, by exchanging information according to a prescribed method to obtain predictable results. This is based on the general definition of interoperability as ability of two or more systems or applications to exchange information and to mutually use the information that has been exchanged.

Cloud interoperability covers both the interaction of the CSCs with a cloud service and also the interaction of one cloud service with another cloud service. In each of the cases, there are typically multiple interfaces involved in the interaction (e.g. service interface, administration interface, business interface).

7.6 Maintenance and versioning

Maintenance is the process of modifying a system or component after delivery to correct flaws, improve performance or other attributes, or adapt to a changed environment (ISO/IEC 21827[8]).

Versioning is the assignment of either unique version names or unique version numbers to unique states of software configuration items, usually for a specific purpose, such as a release of the software product to an external group or the identification of a specific baseline (ISO/IEC/IEEE 24765[17]).

Maintenance can take place for a variety of reasons, including fixing faults and also upgrading or extending facilities for business reasons. Maintenance actions can have the effect of changing the behaviour of cloud services – in particular, changes can affect how a service operates when used by a customer.

Maintenance of cloud services and of the resources that cloud services use is advantageous. This is a security-enhancing aspect of cloud services, with the CSP ensuring that all resources use by the cloud service are up to date with the latest security fixes.

It is important to distinguish between maintenance performed by the CSP and maintenance performed by the CSC. This depends on the cloud service capabilities type and on the responsibility for ownership and control of relevant components.

Through the use of techniques such as DevOps and continuous delivery, it has become common for cloud services to be updated frequently with rapid cycles, including both fixes and updates to functionality. This can be good for maintenance, but it puts more emphasis on the need for clear versioning.

An important area for careful versioning is any APIs provided by the cloud service. User interfaces can be changed without too great an impact on the use of the cloud service. API changes can have significant impact on client software. It is important to aim to achieve forward compatibility when updating APIs to avoid breaking client software. Any breaking API changes need to be introduced cautiously and with warnings given to CSCs, and only introduced when absolutely necessary, for example when caused by fixing bugs or solving security issues.

Maintenance and versioning procedures and targets can be described in the CSA and cloud SLA, as described in ISO/IEC 19086.

7.7 Performance

Performance is a general term covering a number of non-functional aspects of a cloud service. The most common aspect is the response time to complete service requests, but it can include other aspects such as

- transaction rate at which service requests are executed;
- latency for service requests;
- data throughput rate;
- number of concurrent service requests;
- data storage capacity;
- number of concurrent execution threads;
- amount of runtime memory (RAM) available;
- limits on the rapidity and extent of elasticity and scaling.

Cloud service performance is covered by a number of components in the cloud SLA, as described in ISO/IEC 19086-1. The definition of performance can vary depending on the cloud service, the CSP and (potentially) the CSC. Performance considerations are vital for the overall use of cloud services in consideration of reliability, scalability and optimum cost of cloud services.

7.8 Portability

Portability for cloud services involves cloud application portability and cloud data portability, either individually or together, depending on the nature of the cloud service. Cloud application portability is the ability to migrate an application from one cloud service to another cloud service or between a CSC's system and a cloud service. Cloud data portability is data portability from one cloud service to another cloud service or between a CSC's system and a cloud service – where data portability is the ability to easily transfer data from one system to another without being required to re-enter data.

Portability in cloud computing is rarely confined to a binary decision of possible or impossible. More often, portability is possible subject to switching costs (see ISO/IEC 19941[5]). A cost/benefit analysis is required to determine whether porting applications or data is worthwhile. The similarity of the CSC and CSP's systems is therefore more of a matter of lowering the switching cost than of "enabling" portability to take place, since almost any portability is possible if the customer is willing and able to pay for it. Switching concerns are not limited to costs; it also usually involves some risks and usually entails the CSC spending effort and time and perhaps a period of service interruption.

There is the potential to port data or applications to and from the CSC's system and a cloud service.

7.9 Protection of PII

Protection of personally identifiable information (PII) involves the appropriate collection, processing and disposal of PII. A privacy information management system (PIMS) can be used in the context of an organisation's approach to protection and management of PII (ISO/IEC 27701[25]).

PII is any information that can be used to establish a link between the information and the natural person to whom such information relates, or is or can be directly or indirectly linked to a natural person. The person to whom the PII relates is termed the "PII principal" or "data subject".

The PII controller is an entity (typically an organization) that determines the purposes and means for processing PII. A PII processor is an entity (typically an organization) that processes PII on behalf of and in accordance with the instructions of a PII controller.

In cloud computing, the most common situation is that the CSC is a PII controller and the CSP is a PII processor, with the PII being contained within CSC data within the cloud service. In this case, it is the CSC organization that has the relationship with the PII principals and which is responsible for collecting and processing the PII. There are some cases where the CSP acts as a PII controller – for example, where the cloud service is offered directly to customers and is used to hold the customers' PII or where the CSP generates, holds and processes per-customer data. ISO/IEC 27018[20] establishes commonly accepted control objectives, controls and guidelines for implementing measures to protect Personally Identifiable Information (PII) in line with the privacy principles in ISO/IEC 29100[26] for the public cloud computing environment.

The appropriate collection, processing and disposal of PII is covered by the privacy principles of ISO/IEC 29100, but it is also subject to many laws and regulations in different jurisdictions around the world. Some types of PII are subject to more stringent control. For example, personal health data is typically subject to strict controls. It is necessary for both the CSC and the CSP to be aware of the privacy principles and the applicable laws and regulations.

A PIMS is applicable to both PII controllers and to PII processors. ISO/IEC 27701 defines a set of requirements and sets of controls for both PII controllers and PII processors.

Where a CSC uses cloud services to process PII, the interface between the CSC and the CSP is of vital significance. It is necessary to clearly understand the roles of the CSC and the CSP, typically acting as PII controller and PII processor respectively. ISO/IEC 27018 addresses the interface between the CSC and the CSP for public cloud services and describes a set of controls that can be applied in the context of that interface.

One of the major items that forms part of the interface between a CSC and a CSP that relates to privacy and protection of PII is the cloud service agreement and its associated SLAs. The CSA and SLAs elements specifically relating to privacy and security are described in ISO/IEC 19086-4[4]. It is advisable for clarity and transparency purposes to use the SLOs, Service qualitative objectives (SQOs) and associated elements described in 19086-4 where a cloud service is being used to process PII.

Another factor to consider when processing PII is the need to provide transparency about the processing of PII to the PII principals. It is advisable and can be legally necessary to provide the PII principals with data use statements which cover the various elements of PII processing which take place within cloud services. Where the PII controller is the CSC, the CSP provides such data use statements to the CSC. The CSC in turn provides data use statements to the PII principals whose PII the CSC is processing. ISO/IEC 19944-1[6] provides a comprehensive set of elements for creating appropriate data use statements. Similarly, ISO/IEC 29184[28] provides a structure for online privacy notices to PII principals and separately for gathering consent from PII principals, if consent is the basis used for lawful processing of PII.

7.10 Regulatory

The use and provision of cloud services are influenced by a range of legal and regulatory provisions, which naturally tend to vary from one jurisdiction to another. Both CSCs and CSPs need to be fully aware of the legal and regulatory provisions which apply. This can be made more complex by the cross-jurisdictional nature of cloud services, many of which are provided on a global or on a regional basis.

Some legal and regulatory provisions are specific to cloud services. Other provisions are not specific to cloud services but do apply generally, such as those relating to security and to privacy. There are also provisions that depend on the particular use that is being made of cloud services, or the area of application covered by the cloud services. This latter case is exemplified by cloud services that have functionality relating to health, which are typically closely regulated in most jurisdictions. The provision of financial services is similarly closely controlled.

There are regulations that target the operations processes of cloud services, which are rising in importance. These regulations relate to logging of information about the use of cloud services, lawful interception, digital evidence, child protection, the use of AI including facial recognition. Standards relating to such regulations include ISO/IEC 27050[24] on electronic discovery and ISO/IEC 27037[22] for identification, collection, acquisition, and preservation of digital evidence.

Also notable in the area of laws and regulations are moves to change tax laws relating to organizations operating across multiple jurisdictions, typical of many CSPs, with the result that careful accounting of cloud service use is necessary.

It is likely that CSCs are fully aware of the use that is being made of cloud services, but this does not necessarily apply to CSPs.

The core mechanism for establishing that a cloud service meets the necessary legal and regulatory provisions is the cloud service agreement (CSA) and its related SLA. It is important that the CSP provides appropriate transparency about the cloud services they offer. In this case, the content of ISO/IEC 19086 helps establish a common vocabulary for the CSA and SLA. Where security and privacy are concerned, additional sources of both requirements and guidance are provided by ISO/IEC 27017, ISO/IEC 27018 and ISO/IEC 27701.

For further transparency about the processing taking place within a cloud service and the use being made of data, ISO/IEC 19944-1 applies.

For those developing or updating governmental laws, regulations, or policies affecting the purchase, deployment, or use of cloud services it is useful to consider ISO/IEC TR 22678[10]. This document provides advice on various aspects of crafting regulatory, advisory, or procurement policies for cloud computing, for all levels of government and where appropriate also for organizations.

7.11 Resiliency

Resiliency is the ability of a system to provide and maintain an acceptable level of service in the face of faults (unintentional, intentional, or naturally caused) affecting normal operation.

Resiliency describes the set of monitoring, preventive and responsive processes that enable a cloud service to provide continuous operations, or predictable and verifiable outages, through failure and recovery actions. These can include hardware, communications and/or software failures, and can occur as isolated incidents or in combination, including serial failure. These processes can include both automated and manual actions, usually spanning multiple systems, and thus their description and realization are part of the overall cloud infrastructure, not an independent function.

Inherent in resiliency is the realization of risk management - since resiliency is determined by the least resilient component in the system, and cost/performance or other factors can limit the extent to which resiliency is possible or practical. The association of risk to value is realized in the implementation choices to provide resiliency.

7.12 Reversibility

Reversibility is the process for the CSC to retrieve their CSC data and application artefacts and for the CSP to delete all CSC data as well as contractually specified cloud service derived data after an agreed upon period.

The activity related to reversibility can in most cases involve a series of steps, typically requiring the cloud service customer to retrieve their data and inform the cloud service provider that the cloud service provider can delete their copies of the cloud service customer data - safeguarding backup copies until that point in case of failures in the exit process. These steps can also apply to any peer services that are used by the cloud service provider to support the cloud service provider's services.

It is generally expected that reversibility is covered by the CSA and related cloud SLA. However, in some jurisdictions and for some types of information such as PII, regulations exist that place specific requirements on cloud services regarding reversibility. CSPs and CSCs need to be aware of such regulations.

ISO/IEC 19086 has SQOs and SLOs which relate to reversibility elements of the CSA and cloud SLA. ISO/IEC 19941 deals with the application portability and cloud data portability issues involved in CSCs retrieving their assets from a cloud service.

7.13 Security

Information security is the preservation of confidentiality, integrity and availability of information. In addition, other properties, such as authenticity, accountability, non-repudiation, and reliability can be involved.

Information security applies both to the provision of cloud services by the CSP and to the use of cloud services by the CSC. Cloud services are significant in that they involve shared responsibilities for information security between the CSC and the CSP, where the split of responsibilities can vary depending on the capabilities provided by the cloud service. For example, it is likely that the majority of responsibilities lie with the CSP for cloud services of application capabilities type. By contrast, for cloud services of infrastructure capabilities type, much more responsibility is likely to lie with the CSC.

As described in ISO/IEC 27017, which deals with information security for cloud services, it is necessary for the CSA to describe the responsibility split between the CSP and CSC for a particular cloud service. ISO/IEC 19086-4 contains a series of SLOs and SQOs which provide the detail necessary to describe the information security components of the CSA and the cloud SLA. It is common for cloud services to have certifications to one or more information security schemes, which can include certification against ISO/IEC 27001[16] using the extended control set defined in ISO/IEC 27017.

ISO/IEC 27036-4[21] provides detail on the supplier relationship aspects between the CSC and the CSP in the area of information security.

Important elements of the relationship between the CSC and the CSP involve monitoring and reporting in relation to information security. Security incidents can be detected and reported by either party. Appropriate processes need to be in place to deal with security incidents and these processes necessarily need appropriate communication to take place between the parties. Security vulnerabilities need to be addressed without delay and affected hardware and software updated appropriately. ISO/IEC 29147[27] deals with the topic of vulnerability disclosure and ISO/IEC 30111[29] covers vulnerability handling processes.

7.14 Service levels and service level agreement

Service level agreements are important components of cloud computing governance and represent measurable elements needed to assure an agreed upon quality of service between a CSC and a CSP.

The cloud service level agreement (cloud SLA) as described in ISO/IEC 19086, all parts, is a part of the cloud service agreement (CSA) that includes cloud service level objectives (SLOs) and cloud service qualitative objectives (SQOs) for the covered cloud service. Details of cloud SLAs, SLOs and SQOs can vary for different cloud service categories (5.4), cloud capabilities types (5.3), key characteristics (5.2), cross-cutting aspects (7) and different cloud deployment models (5.5). CSAs and their associated cloud SLAs vary between CSPs, and in some cases different CSCs can negotiate different contract terms with the same CSP for the same cloud service.

The CSA and cloud SLA define the properties of the cloud service in terms of a set of observable and measurable components. Implementing a cloud SLA involves setting up processes for monitoring and managing cloud service characteristics, reporting any failures to meet SLOs and SQOs and identifying any remedies.

8 Data and cloud services

8.1 General

Data is an important element of any ICT system, and this applies to cloud services. There are a number of topics relating to data and cloud services which need to be addressed and these are covered by a series of standards, which are described in the following clauses.

8.2 Data processing within cloud services

Data is processed within many cloud services. The simplest form of data processing is data storage, which is the basic capability of DSaaS. However, many cloud services process data in various other ways, such as performing analytics on the data.

There is a need to describe the data processing that goes on within a cloud service. The description can assist with providing transparency to CSCs or to CSUs concerning the functionality of the cloud service. The transparency can also be required for some types of data, such as PII, where the interests of the PII principal need to be addressed.

ISO/IEC 19944-1 includes foundational concepts, including a data taxonomy for use in the description of data processing within cloud services. The standard also provides a series of data processing and data use categories. The data taxonomy and data processing categories enable the construction of data use statements that enable the CSP to provide the necessary information about data processing in a standard form.

ISO/IEC 22624[9] extends the data taxonomy and data use categories of ISO/IEC 19944-1 and creates a framework for the creation of data-related policies and practices for cloud services. ISO/IEC 22624 also describes codes of conduct for practices regarding data at rest and in transit, including cross border data transfer, as well as remote access to data. Data-related policies and codes of conduct can apply to both CSPs and CSCs.

8.3 Data flow

Data flow takes place into and out of cloud services. Data flow can involve other cloud services; it can involve devices connected to the cloud services; it can involve nodes within edge computing tiers.

Data flow between cloud services is described in terms of interoperability and data portability by ISO/IEC 19941. Data flow to and from edge computing nodes, described in ISO/IEC 23188[14], involves the same aspects of interoperability and data portability. Data flow between devices and cloud services is addressed by ISO/IEC 19944-1 and also by ISO/IEC 23188 in the context of edge computing.

8.4 Processing of data from multiple sources

The scalability of resources available in association with cloud services, both in terms of large data storage capacity and in terms of high processing capacity, can simplify the bringing together of multiple data sets from a range of different sources and the processing and analysis that spans across those multiple data sets. Such processing can provide insights and outcomes that are not possible without using all of the data sets.

Processing of data from multiple sources involves a series of challenges for the organization concerned. As the data flows through the processing steps to produce some outputs, a series of trust elements are relevant, including:

- data use obligations and controls;
- data provenance records, quality and integrity;
- chain of custody;
- security and privacy;
- immutable proof of regulatory compliance.

ISO/IEC 23186[12] provides organizations with a framework which aims to enable trustworthy processing of data from multiple sources, particularly in the context of cloud computing.

8.5 Data sharing

It is increasingly evident to organisations of the value of being able to share and exchange data. This approach can drive collaboration across the ecosystem of stakeholders, and this can improve and drive enhanced business mission outcomes. Such sharing can be on a no-cost basis, but sharing can also involve significant payments by receiving organizations. Data sharing tends to be a significant part of the processing of multi-sourced data, since it is often the case that a single organization is unable to source all relevant data itself.

Data sharing between two or more organizations raises a question regarding the basis on which the sharing takes place. On the side of the data holder (the provider), there is a need to describe what can be done with the data by the data user (the customer), what obligations are placed on the data user, what restrictions exist. For the data user, there is a need for assurance about the origins of the data, for clarity about any processing done on the data.

As a result, it is expected that data sharing is performed under the terms of a data sharing agreement (DSA). It is the DSA that describes the terms on which the data is shared, the obligations for the data holder and for the data user, data provenance, data quality, data integrity and any processing restrictions placed on the data user.

Data sharing agreements, their structure and contents, are described in ISO/IEC 23751[16].

9 Virtualization concepts

9.1 General

One of the key characteristics of cloud computing is its ability to share resources. Sharing of compute resources involves some level of virtualization.

For an additional description of virtualization concepts, see ISO/IEC TS 23167.

9.2 System hardware virtualization

9.2.1 General

One approach to the virtualization of compute resources is the use of virtual machines (VMs), which often involve using a hypervisor to virtualize the physical resources of the system hardware, which enables multiple virtual machines to run on a particular physical system. This permits the system to be shared by the software running inside each VM.

9.2.2 Virtual machines

A virtual machine is an isolated execution environment for running software that uses virtualized physical resources. In other words, this involves the virtualization of the system – and the software within each VM is given carefully controlled access to the physical resources to enable sharing of those resources without interference. Sometimes termed system virtual machines, VMs provide the functionality needed to execute complete software stacks including entire operating systems and the application code that uses the operating system.

The purpose of VMs is to enable multiple applications, or instances of an application, to run at the same time on one hardware system, while remaining isolated from each other. The software running within each VM appears to have its own system hardware, such as processor, runtime memory, storage device(s) and networking hardware.

Each VM running on a system can contain its own operating system – and this permits multiple different operating systems to run on a single system, each in its own VM.

9.2.3 Hypervisors

A hypervisor, sometimes termed a virtual machine monitor, is software that virtualizes physical resources and allows for running virtual machines. Virtualization means control of the abstraction of the underlying physical resources of the system. A hypervisor also manages the operation of the VMs. A hypervisor allocates resources to each running VM including central processing unit (CPU), memory, disk storage and networking capabilities and bandwidth.

9.3 Containers

A container is an isolated execution environment for running software that uses a virtualized operating system kernel (termed the host operating system). Virtualized operating system kernel means access to the resources of the operating system is mediated and the software within the container only gets to see and use a carefully controlled and limited version of the operating system resources.

The key point of the operating system kernel virtualization is that it limits what can be accessed by the software in the container – the software in the container is isolated (it cannot access other software on the same system, or be accessed by other software), but also the software is presented with a strictly limited version of the host operating system itself, in an analogous way to the virtualization of the physical resources for a VM.

Each container contains its own application software, and runs that software in a set of processes using virtualized resources such as memory, CPU, storage and networking, where the resources available to the container are usually a subset of the resources of the host operating system. The software in the container is isolated from software running in other containers running on the same system, but all share the kernel of the host operating system. This means that the operating system used by the software in the containers is compatible with the host operating system kernel.

9.4 Serverless computing

Serverless computing is a cloud service category in which the CSC can use different cloud capabilities types without the CSC having to provision, deploy and manage either hardware or software resources, other than providing CSC application code or providing CSC data. Serverless computing represents a form of virtualized computing.

Functions as a service is a form of serverless computing in which the capability used by the CSC is the execution of CSC application code as one or more functions that are each triggered by a CSC specified event.

Serverless database is a form of serverless computing in which the capability used by the CSC is a database, where the database is provisioned, managed and operated by the CSP and its functions are made available via an API.

9.5 Virtualized networking

Network capabilities are often virtualized and are oriented towards the applications or systems that use the capabilities rather than the underlying networking hardware. The capabilities involved include virtual private networks (VPN), which provide secure networking between specific machines, even where those machines are attached to different networks.

9.6 Virtualized DSaaS

Data storage as a service is discussed in 5.4.8. It is common for DSaaS to be virtualized, in that the storage capabilities offered to CSCs are abstracted from the underlying hardware by software and can offer advanced capabilities without intervention by the CSC, such as distribution, replication, sharding, affinity.

10 Concepts of cloud computing involving multiple CSPs

10.1 General

Instead of accessing the cloud services provided by a single CSP, a CSC can access cloud services directly from more than one CSP. The concept of a CSC using the cloud services of multiple CSPs comes from the desire of the CSC to make use of resources or to utilize capabilities beyond the capabilities of any single CSP. The potential benefits to the CSC include improved reliability, performance optimization, choice of best cloud services and cost savings. As a result, CSCs can adopt solutions that take advantage of multiple cloud services from different CSPs.

Clause 10 introduces basic concepts of cloud computing involving multiple CSPs.

10.2 Types of cloud computing involving multiple CSPs

10.2.1 General

One way to look at the different ways of using and combining cloud services is to examine the ways they can interact with each other. Interactions between CSPs often function differently depending on who is involved, how they are organized, the division of responsibilities and what cloud services are involved.

ISO/IEC 5140[2] provides more detail on multi-cloud and inter-cloud computing.

10.2.2 Multi-cloud computing

A CSC often requires public cloud services from two or more CSPs and can provide for the integration and management of these disparate cloud services.

When cloud services from one or more CSPs are combined or integrated by the CSC for their own use, the result is often referred to as a cloud solution.

Multi-cloud is a cloud deployment model that can be used to deliver application, platform or infrastructure capabilities types.

10.2.3 Inter-cloud computing

A CSP that uses one or more cloud services provided by other CSPs is said to be in an inter-cloud relationship. The CSP using the cloud services is referred to as the primary CSP while a CSP whose cloud services are being used is referred to as a secondary CSP. The purpose of the relationship is to jointly provide cloud services to a CSC.

The use of inter-cloud computing can be invisible to the CSC who thinks they are only using the cloud services from the primary CSP even though some of those cloud services are actually provided by a secondary CSP.

10.2.4 Other types of cloud computing involving multiple CSPs

Other types of cloud computing can involve configurations in which two or more CSPs separately deliver one or more cloud services to a single CSC.

For example, the other CSPs can be used to address elasticity and scalability requirements that exceed the capabilities of an on premises private cloud (see 5.5.2).

A common use case occurs when the CSC uses a cloud service on a private cloud supported by one CSP which is used in combination with one or more cloud services provided by one or more other CSPs.

10.3 Considerations when using multiple CSPs

10.3.1 Identity and access management

In operating with multiple CSPs, it is important to understand that each CSP needs the ability to both identify the users making requests as well as determine if they have the authorization to make those requests.

10.3.2 Policy considerations

In dealing with multiple CSPs, one important consideration is that the CSPs can exist in different policy environments (such as jurisdictional, governance, regulation, etc.). Resolving these different policy concerns can impact the interaction between those CSPs and alter the definition of their cloud services.

10.3.3 Management

Management across multiple CSPs typically focuses on the level of how the CSC and CSPs and the associated cloud services act together. Important issues that need to be addressed by the CSPs, in order to successfully work together, range from dealing with issues of how access is managed, to how billing is handled and to how resources are managed.

10.3.4 Operations

Operational interactions across multiple CSPs focus on the actual provision of cloud services and resources to the CSC. Issues range from how to allocate resources, to broadcasting available resources and capabilities, to monitoring and metering current activities.

11 Organization of cloud computing

11.1 Logical organization of cloud computing

11.1.1 Cloud service instance

When a CSC uses a cloud service, the CSC can allocate one or more instances of that cloud service, as shown in Figure 7.

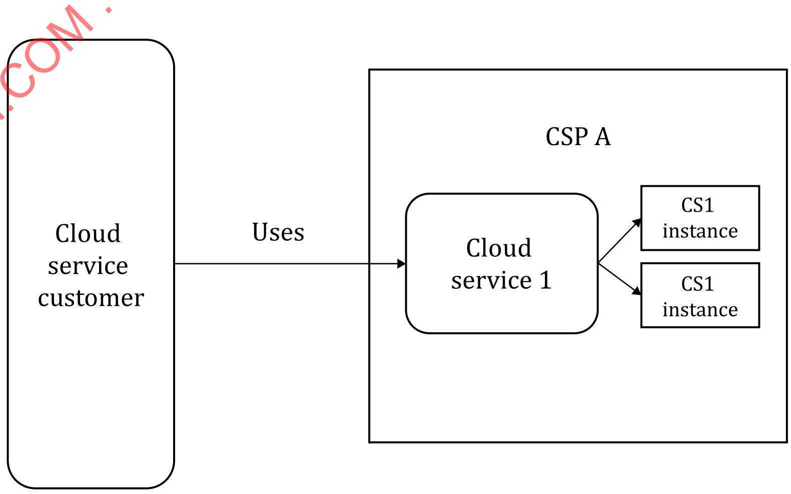

Figure 7 — CSC allocating and using multiple cloud service instances