H-CLOUD

Cloud Computing in Europe

Appendix 1

Technical Definitions

20th April 2020

h-cloud.eu

Title	Cloud Computing in Europe. Appendix 1: Technical Definitions
Lead Editor	Mark Dietrich
Contributors (in alphabetical order)	Carla Arend, Mark Dietrich, Federico Facca
Version	0.4
Date	17 April 2020
Confidentiality Notice	**Confidential** - The information contained in this document and any attachments are confidential. It is governed according to the terms of the project consortium agreement

Important Notice: Working Document

This briefing is an annex to the Green Paper v 0.4. The Green Paper is an outcome of the H-CLOUD project, a Coordination and Support Action that has received funding from the European Union's Horizon 2020 Research and Innovation programme under grant agreement n. 871920.

CONTENTS

TECHNICAL DEFINITIONS USED IN THE H-CLOUD GREEN PAPER

The term "cloud" has many meanings. There is over ubiquitous use of the phrase cloud, with it being added to many other words. This means that, in many cases, the deeper meaning and implication of a description involving some cloud element is often lost or not conveyed. In some cases, people may add cloud or edge to a description to make it sound more interesting.

This section intends to identify and break down these various meanings and uses, so we have clear distinctions between the various words and their uses.

Our analysis: Given this wide range of explicit and implicit meanings of the term “cloud,” in this document we attempt to be very explicit about our meaning whenever using that term, and we recommend that the EC maintain the same level of clarity in its own documentation.

1 DEFINING "CLOUD"

The original meaning of cloud computing related to having characteristics such as elasticity of resource supply and automated provision of services, rather than any specific IT technologies.

NIST's definition[1] of the five essential characteristics of the "cloud" is widely accepted:

- on-demand self-service,
- broad network access,
- resource pooling,
- rapid elasticity or expansion,
- measured service.

In this definition, ICT services can be provided in a "cloudy" way, without specifying the technology or business model to be used.

1.1 Cloud deployment types

Cloud deployment models describe how a cloud IT service is built and delivered to consumers of the service. The factors that determine the type of cloud deployment model are:

- The physical location of the hardware infrastructure systems on which the service is running
- Whether or not the service is dedicated to one organization or shared across multiple independent organizations
- The owner of the hardware infrastructure systems on which the service is running

NIST also defines several forms of cloud deployment from an organizational perspective:

- **Private**: "cloudy" resources reserved for the use of a single organization.
- **Community**: "cloudy" resources available to a community of users (e.g. the research community).
- **Public**: "cloudy" resources available for use by anyone (i.e. anyone who is able to pay for the service).
- **Hybrid**: "cloudy" resources that include two of more of the above deployment types.

Looking at how these services are owned, shared and access is controlled:

- **Public cloud services** are shared among unrelated enterprises and/or consumers, open to a largely unrestricted universe of potential users, and designed for a market, not a single enterprise.
- **Private cloud services** are shared within a single enterprise or an extended enterprise, with restrictions on access and level of resource dedication, and defined/controlled by the enterprise, beyond the control available in public cloud offerings.

Note that the NIST definitions do not specify where the resources are located physically or who owns, controls or operates them.

Figure 1 provides a taxonomy for these various deployment models.

[1] https:/nvlpubs.nist.gov/nistpubs/Legacy/SP/nistspecialpublication800-145.pdf

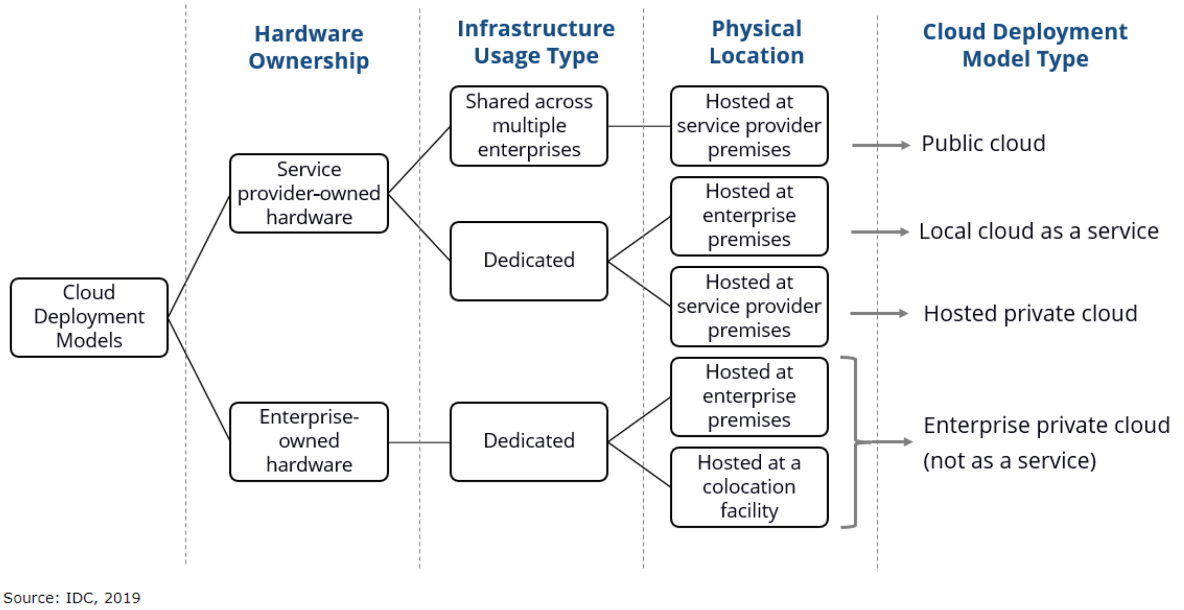

Figure 1. Taxonomy of Cloud deployment models.

The phrase "Cloud" is also loosely used to refer to "infrastructure hosted outside of the organization." Most commercial cloud suppliers are providing “cloudy” service using their own off-site infrastructure, so the term “cloud” sometimes signals the move away from on-premise infrastructure.

1.2 Cloud business models

The business model of public (commercial) cloud suppliers is often a "pay as you go" revenue model. Therefore, the term “cloud” can also refer to a shift from CapEx and OpEx to a more pure OpEx cost model. (This shift is not unique to IT service providers, e.g. General Electric can supply jet engines in a similar “propulsion as a service” business model.)

Public (commercial) cloud suppliers achieve economies of scale in part by “reselling” the same physical infrastructure to multiple clients over time. For example, they assign free capacity to a client when they request it, and releasing that capacity when the client releases it. This applies to all three cloud service models (defined below): from Infrastructure as a Service (IaaS) to Software as a Service (SaaS).

Public (commercial) cloud suppliers also achieve economies of scale in part by consolidating their investments in physical infrastructure, allowing them to serve more and more clients from one location, and amortizing certain fixed costs (e.g. administration, physical security, facilities management, some power and cooling components) over a larger user base. The benefits of such consolidation are offset by two main factors:

- increased and/or more variable network latency. This limits application response times for consumer-facing services, particularly network-intensive applications such as content delivery. This is also a concern for IoT and edge-related services, where edge-based components may not tolerate longer communication delays.
- legal, regulatory and business control concerns around where data is stored.

These concerns apply to all three cloud service models (defined below): from Infrastructure as a Service (IaaS) to Software as a Service (SaaS). Most major IaaS suppliers require clients to specify, broadly, where their infrastructure should be set up, but it is up to the client to architect their solution to work across multiple IaaS locations to meet their own latency and/or data storage requirements.

In SaaS service models, the service itself has been designed in a “multi-tenant” model, storing all client data in the same database and giving all clients access to the same running software, but separating each client's experiences and possible actions by restricting that client to their own data and related processing steps. (Again, this business model is not unique to the cloud. IaaS resources are analogous to airplane seats or hotel rooms. Financial institutions "resell" currency deposits through multiple loans.)

1.1 Cloud as software technologies

"Cloud" is also used to refer to the various software technologies used to deploy IT workloads in a "cloudy" way (typically using virtual machines or containers).

Turning again to NIST:

> *"The cloud infrastructure can be viewed as containing both a physical layer and an abstraction layer.*
>
> - **The physical layer** consists of the hardware resources that are necessary to support the cloud services being provided, and typically includes server, storage and network components.
> - **The abstraction layer** consists of the software deployed across the physical layer, which manifests the essential cloud characteristics. Conceptually the abstraction layer sits above the physical layer."

For the purposes of this paper, we refer to the technologies involved in this abstraction layer as "cloud-style". This means we can refer to "cloud style deployment" and "cloud style technologies" etc..

2 CLOUD SERVICE MODELS

The many cloud suppliers -- public, community and private -- can be analyzed through the common lens of the services they provide to clients, organized in the three models defined by NIST in 2011:

2.1 Infrastructure as a Service (IaaS):

IaaS refers to services closely tied to the physical technologies themselves, notably various types of compute, storage and file systems, and networking, typically exposed as virtual machines ("instances") and managed with virtualization tools such as OpenStack.

2.2 Platform as a Service (PaaS):

In PaaS, cloud-based business functions are developed in the selected PaaS environment and can be configured to scale up or down automatically -- the client will be charged for this but the client doesn't have to track utilization and adjust configurations directly.

2.3 Software as a Service (SaaS):

SaaS involves fully integrated software suites centred around specific activities. Major examples include customer relationship management (CRM), enterprise resource planning (ERP) and financial accounting, as well as tools like Google Office and DropBox. Microsoft's Azure offerings are largely SaaS, supported by flexible IaaS capabilities that are also sold to clients.

Many of the services provided in research clouds could be classed as SaaS, including science gateways and research platforms, and data/artefact management tools (archiving and replication, search and discovery, policy management).

2.4 Hybrid Cloud, Multi-Cloud, Industry Cloud, Federated Cloud

These terms have varying meanings in practice, but in the context of the H-CLOUD project, we adopt definitions from IDC.

Hybrid Cloud: IDC defines hybrid cloud as the usage of IT services (including IaaS, PaaS, SaaS apps, and SaaS-SIS cloud services) across one or more deployment models using a unified framework. The cloud services used leverage more than one cloud deployment model. Hybrid cloud services include "public-public," "public-private," and "private-private" combinations. Cloud and noncloud combinations (sometimes referred to hybrid IT) where the noncloud applications are front ended with cloud services interfaces (e.g., RESTful APIs) are also included.

Hybrid cloud arrangements are often used to integrate technically similar infrastructure capabilities with differing economic or regulatory characteristics. For example, combining fixed cost dedicated private cloud facilities with variable cost public cloud resources allows clients to balance the lower TCO of dedicated equipment against the elastic provisioning, yet higher variable costs, of public cloud. Similarly, dedicated private facilities can be earmarked for storage and processing of sensitive data, while data and workloads with lower sensitivity can be deployed on public cloud.

Multicloud: The term multicloud is sometimes used in the industry as an alternative to hybrid cloud. However, for IDC, multicloud is a description of an organizational strategy or the architectural approach to the design of a complex digital service that involves consumption of cloud services from more than one cloud service provider. These may be directly competing cloud services such as hosted private cloud versus public cloud compute services, public object storage from more than one public cloud service provider, or IaaS and SaaS from one or more cloud service providers. In both of these contexts, multicloud encompasses a much larger universe than hybrid cloud and is only gated by the cost/complexity associated with enabling consistent management/ governance of many different cloud options.

Multicloud arrangements allow clients to integrate “best fit” functional capabilities from multiple SaaS providers (for example SAP ERP and Salesforce.com CRM), with customized applications developed by clients to deliver differentiated, business-specific solutions, such as product configuration, integrated manufacturing management, etc. In addition to contracting for the various functional services and developing and maintaining its business-specific processes, the client must also integrate those functions to enable seamless operations. The client orchestrates the different components of its complete capability through a range of contractual agreements governing purchased cloud services and potentially software development services.

Industry Cloud: IDC defines industry clouds (ICs) as cloud platforms that enable multiple organizations in an industry to transact or collaborate toward common goals, such as improving industry insights, efficiencies, and/or capabilities. ICs can be controlled by multiple organizations/participants or by one primary owner. An industry cloud often resides on a public cloud but may reside in a private cloud environment when there is a smaller number of participants. Some ICs are open platforms or services that organizations may join freely, while others are commissioned and controlled by a group of named enterprises (occasionally called community clouds in some industries).

An industry cloud is defined relative to the set of organizations for which the IC is designed and the common goal being pursued. Mutual benefit is a key organizing principle. An IC could use any underlying cloud implementation model, from public cloud, to hybrid cloud, to multicloud. However, in an IC, the IC participants often provide services to other participants, so participants can be both service providers and service consumers. The creation of the industry cloud recognizes that each of the IC's services can benefit multiple parties, creating greater value for the group than if each service provider performed the service only for itself. Among the participants, the industry cloud operates like a multicloud, while also integrating various IC services for the benefit of all participants, and amortizing the cost of that integration across a greater number of clients.

Federated Cloud: Federated cloud is not defined by IDC, but a definition is presented here to highlight the differences from and similarities with the other three flavours of cloud.

- Like an industry cloud and a multi-cloud, federated cloud brings together services from multiple providers.
- Like an industry cloud, but unlike multi-cloud, federated cloud also integrates those services so that they are interoperable. (In multi-cloud, this integration is the client's responsibility.)
- Unlike many industry clouds, federated clouds employ a specific organizational structure, with a central coordinating body and formal governance processes.
- Like industry clouds, federated cloud services are primarily provided by federation members to other federation members, rather than contracted providers.
- Unlike most industry clouds but similar to some hybrid cloud structures, federated clouds often provide resource pooling across multiple providers (other federation members).

Appendix 14 explores federated clouds in more detail, and returns to this comparison of cloud flavours to examine use cases that are appropriate for each type of cloud.

3 EDGE AND FOG DEFINITIONS, DISTINCTIONS WITH CLOUD COMPUTING

The following definitions are drawn from the NIST Fog Computing Conceptual Model https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.500-325.pdf:

- Sensors/Actuators can include all IoT devices, wearables, smart devices (phones), mobile devices (on cars) etc.
- Edge computing is defined as being connected directly to sensors/actuators, and provided with enough compute/storage/network resources to manage connected sensors/actuators. Edge computing resources have well defined geographic ("I am here") and logical ("I am connected to these things in the real world") attributes.
- By contrast, cloud-style computing is not tightly defined from either a geographic or logical standpoint, and typically there are latencies/limitations when communicating with cloudy resources. Physically, cloudy resources are located in specific data centres, but since these resources are presented to clients in a virtualized or abstracted way, clients cannot rely on any physical attributes when developing solutions with cloud-style resources.
- Fog computing resources may be connected between edge/sensors/actuators and cloud-style resources.
    - Fog computing is intended to be more geographically and logically aware than cloudy-computing. Fog computing allows solutions to be architected that meet constraints on processing time, response time, etc.
    - Fog computing shares many characteristics with cloud-style computing (as defined by NIST). For example it can be deployed with public, private, community, and hybrid models. It can also operate in similar service models (IaaS, PaaS, SaaS).
    - Like cloud-style computing, fog computing nodes can be provisioned on demand, but provisioning requests also need to specify the required geographic/logical characteristics. For example, if new sensors/actuators are added to a solution (e.g. a client is increasing its service territory), but their location requires additional processing/storage/network buffering resources nearby in order to operate correctly, new fog computing nodes, with the required geographic/logical characteristics, can be provisioned from fog computing providers (such as a mobile network operator) and linked to the new devices.

Cloud computing assumes that the processing and storage of data take place within the boundaries of a cloud and its underlying infrastructure. While conventional cloud computing infrastructures and applications are designed to scale, and - eventually - to replicate geographically, they are not meant to be geographically dispersed across a high number of points of presence and leveraging heterogeneous end points.

A geographic dispersion of coordinated data processing and storage services enables low latency and data localization, typical requirements of some of the most recent use case scenarios in Internet of Things and Media application areas.