Chapter 4

Overview of Cloud Computing

4.1 Background

National Institute of Standards and Technology, U. S. Department of Commerce initiated the standardization processes for cloud computing and issued 'NIST Cloud Computing Standards Roadmap' in July 2011 which was followed by version 2.0 in July 2013. The NIST Cloud Computing Program has developed a United States Government (USG) Cloud Computing Technology Roadmap, as one of many mechanisms in support of USG's secure and effective adoption of the Cloud Computing model to reduce costs and improve services.

International Telecommunication Union (ITU) established Focus Group on Cloud Computing (FG Cloud) further to ITU-T Telecommunications Standardization Advisory Group (TSAG) agreement at its meeting in Geneva, 8-11 February 2010 followed by ITU-T study groups and membership consultation with the objective to scout standardization landscape, with a closure in December 2011. The Focus Group was established to collaborate with worldwide cloud computing communities (e.g., research institutes, forums and academia) including other Standards Developments Organizations (SDOs) and consortia. The structure of working groups (WG) and work areas (WA) are given below:

- **WG1: Cloud computing benefits and requirements**
  - WA 1-1 Cloud Definition, Ecosystem & Taxonomy
  - WA 1-2 Uses cases Requirements & Architecture
  - WA 1-3 Cloud security
  - WA 1-4 Infrastructure & Network enabled Cloud
  - WA 1-5 Cloud Services and Resource Management, Platforms and Middleware
  - WA 1-6 Cloud computing benefits and first Requirements from ICT perspectives
- **WG2: Gap Analysis and Roadmap on Cloud Computing Standards development in ITU-T**

- WA 2-1 Overview of cloud computing SDOs activities
- WA 2-2 Gap analysis and Action plan for development of relevant ITU-T Cloud Standard

➢ **Focus Group was mandated to have seven output documents:**

- Overview of Standards Development Organizations involved in Cloud Computing
- Introduction to the Cloud Ecosystem
- Benefits of Cloud Computing from Telecom/ICT Perspectives
- Cloud Security, Threat & Requirements
- Functional Requirements and Reference Architecture
- Infrastructure and Network Enabled Cloud
- Cloud Resources Management Gap Analysis

ITU has since issued recommendations on various issues concerning cloud computing. Cloud computing has been described as a paradigm for enabling network access to a scalable and elastic pool of shareable physical or virtual resources with self-service provisioning and on-demand administration. The **cloud computing** paradigm is composed of **key characteristics**, cloud **deployment models,** cloud **capabilities types** and cloud **service categories,** cloud computing **roles and activities**, and cloud computing **cross cutting aspects.**

4.2 Key characteristics

Cloud computing is an evolving paradigm and the key characteristics of cloud computing are:

- *Broad network access*: A feature where the physical and virtual resources are available over a network and accessed through standard mechanisms that promote use by heterogeneous client platforms. The focus of this key characteristic is that cloud computing offers an increased level of convenience in that users can access physical and virtual resources from wherever they need to work, as long as it is network accessible, using a wide variety of clients including devices such as mobile phones, tablets, laptops, and workstations;

- *Measured service*: A feature where the metered delivery of cloud services is such that usage can be monitored, controlled, reported, and billed. This is an important feature needed to optimize and validate the delivered cloud service. The focus of this key characteristic is that the customer may only pay 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;
- *Multi-tenancy*: A feature where 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 that form a tenant will all belong to the same cloud service customer organization. There might be cases where the group of cloud service users involves users from multiple different cloud service customers, particularly in the case of public cloud and community cloud deployments. However, a given cloud service customer organization might have many different tenancies with a single cloud service provider representing different groups within the organization;
- *On-demand self-service*: A feature where a cloud service customer can provision computing capabilities, as needed, automatically or with minimal interaction with the cloud service provider. 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;
- *Rapid elasticity and scalability*: A feature where physical or virtual resources can be rapidly and elastically adjusted, in some cases automatically, to quickly increase or decrease resources. For the cloud service customer, the physical or virtual resources available for provisioning often appear to be unlimited and can be purchased in any quantity at any time automatically, subject to constraints of service agreements. Therefore, the focus of this key characteristic is that cloud

computing means that the customers no longer need to worry about limited resources and might not need to worry about capacity planning;

- *Resource pooling*: A feature where a cloud service provider's physical or virtual resources can be aggregated in order to serve one or more cloud service customers. The focus of this key characteristic is that cloud service providers can support multi-tenancy while at the same time using abstraction to mask the complexity of the process from the customer. From the customer's perspective, all they know is that the service works, while 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 customer's original workload, such as maintenance requirements, to the provider. Even with this level of abstraction, it should be pointed out that users might still be able to specify location at a higher level of abstraction (e.g., country, state, or data centre).

4.3 Cloud deployment models

Cloud deployment models represent how cloud computing can be organized based on the control and sharing of physical or virtual resources. The cloud deployment models include:

- *Public cloud*: Cloud deployment model where cloud services are potentially available to any cloud service customer and resources are controlled by the cloud service provider. A public cloud may be owned, managed, and operated by a business, academic, or government organization, or some combination of them. It exists on the premises of the cloud service provider. Actual availability for specific cloud service customers may be subject to jurisdictional regulations. Public clouds have very broad boundaries, where cloud service customer access to public cloud services has few, if any, restrictions;
- *Private cloud*: Cloud deployment model where cloud services are used exclusively by a single cloud service customer and resources are controlled by that cloud service customer. A private cloud may be owned, managed,

and operated by the organization itself or a third party and may exist on premises or off premises. The cloud service customer may also authorize access to other parties for its benefit. Private clouds seek to set a narrowly controlled boundary around the private cloud based on limiting the customers to a single organization;

- *Community cloud*: Cloud deployment model where cloud services exclusively support and are shared by a specific collection of cloud service customers who have shared requirements and a relationship with one another, and where resources are controlled by at least one member of this collection. A community cloud may 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 may exist on or off premises. Community clouds limit participation to a group of cloud service customers 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 compliance considerations;
- *Hybrid cloud*: Cloud deployment model using at least two different cloud deployment models. The deployments involved remain unique entities but are bound together by appropriate technology that enables interoperability, data portability and application portability. A hybrid cloud may be owned, managed, and operated by the organization itself or a third party and may exist on premises or off premises. Hybrid clouds represent situations where interactions between two different deployments may be needed but remained linked via appropriate technologies. As such the boundaries set by a hybrid cloud reflect its two base deployments.

4.4 Cloud capabilities types and cloud service categories

A cloud capabilities type is a classification of the functionality provided by a cloud service to the cloud service customer, based on the resources used. There are three different cloud capabilities types: application capabilities type, infrastructure capabilities type, and platform capabilities type, which are different

because they 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 cloud service customer can use the cloud service provider's applications;
- *Infrastructure capabilities type*: A cloud capabilities type in which the cloud service customer can provision and use processing, storage or networking resources;
- *Platform capabilities type*: A cloud capabilities type in which the cloud service customer 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 cloud service provider.

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.

Representative cloud service categories are:

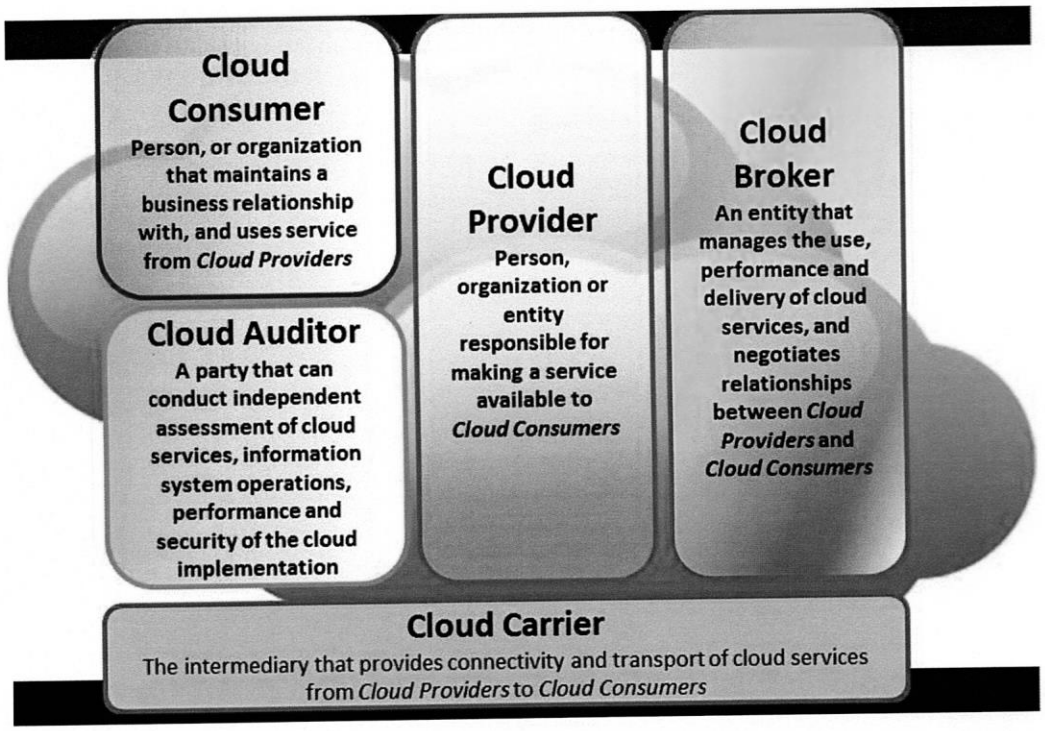
- *Communications as a Service (CaaS)*: A cloud service category in which the capability provided to the cloud service customer is real time interaction and collaboration;
- *Compute as a Service (CompaaS)*: A cloud service category in which the capabilities provided to the cloud service customer are the provision and use of processing resources needed to deploy and run software;
- *Data Storage as a Service (DSaaS)*: A cloud service category in which the capability provided to the cloud service customer is the provision and use of data storage and related capabilities;
- *Infrastructure as a Service (IaaS)*: A cloud service category in which the cloud capabilities type provided to the cloud service customer is an infrastructure capabilities type;

- *Network as a Service (NaaS)*: A cloud service category in which the capability provided to the cloud service customer is transport connectivity and related network capabilities;
- *Platform as a Service (PaaS)*: A cloud service category in which the cloud capabilities type provided to the cloud service customer is a platform capabilities type;
- *Software as a Service (SaaS)*: A cloud service category in which the cloud capabilities type provided to the cloud service customer is an application capabilities type.

It is expected that there will be additional cloud service categories.

4.5 Cloud computing roles and activities

The NIST cloud computing reference architecture defines five major actors: cloud consumer, cloud provider, cloud auditor, cloud broker, and cloud carrier.



(Source: NIST, 2011)

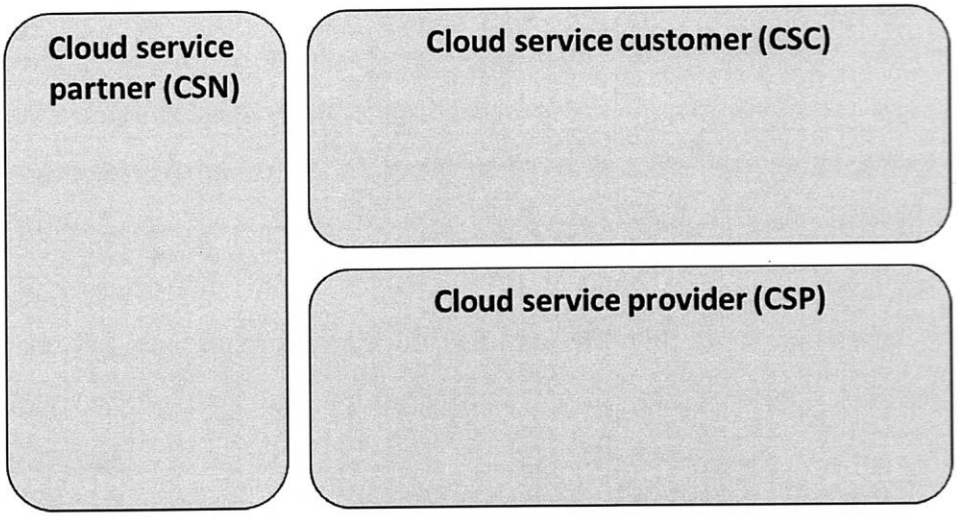
Figure 4.1: Cloud Actors

Given that distributed services and their delivery are at the core of cloud computing, all cloud computing related activities can be categorized into three

main groups: activities that use services, activities that provide services and activities that support services. A cloud computing activity is defined as a specified pursuit or set of tasks. Cloud computing activities need to have a purpose and deliver one or more outcomes.

Within the context of cloud computing, it is often necessary to differentiate requirements and issues for certain parties. A party is a natural person or legal person, whether or not incorporated, or a group of either. Parties in a cloud computing system are its stakeholders.

These parties are entities that play roles (and sub-roles). A role is a set of cloud computing activities that serve a common purpose. Roles, in turn, are sets of activities and activities themselves are implemented by components. All cloud computing related activities can be categorized into three main groups:



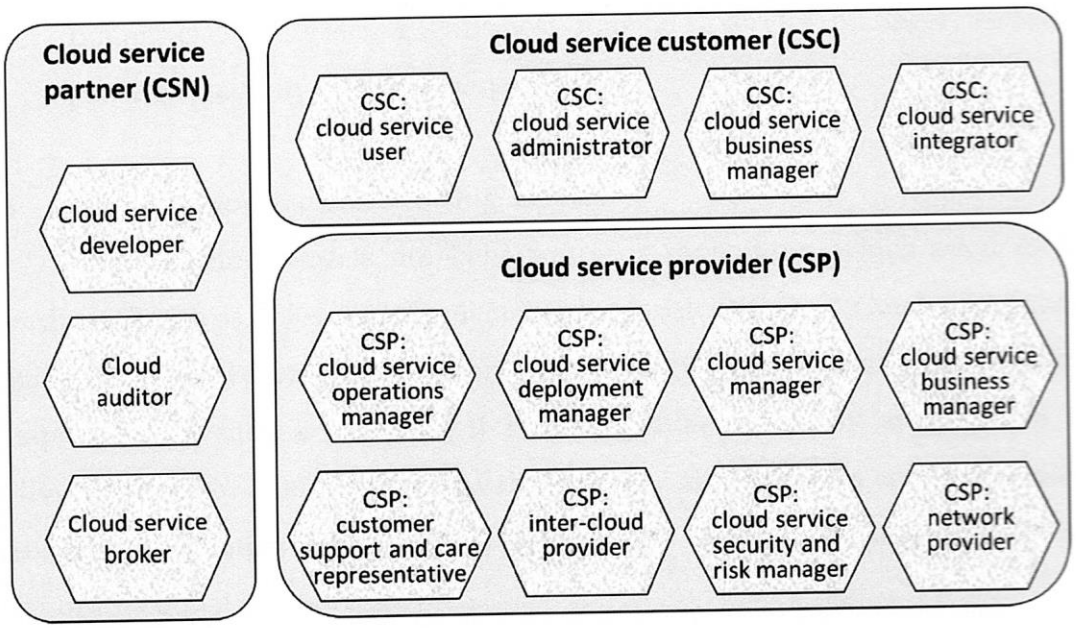
(Source: ITU Regulations)

Figure 4.2: Cloud Computing Roles

A sub-role is a sub-set of the cloud computing activities for a given role. Different sub-roles can share the cloud computing activities associated with a given role.

A party can assume more than one role at any given point in time and can engage in a specific sub-set of activities of that role. Examples of parties include

but are not limited to: large corporations, small and medium sized enterprises, government departments, academic institutions, and private citizens.



(Source: ITU Regulations)

Figure 4.3: Roles and Sub-roles

The major roles of cloud computing are:

- *Cloud service customer*: A party which is in a business relationship for the purpose of using cloud services. The business relationship is with a cloud service provider or a cloud service partner. Key activities for a cloud service customer include, but are not limited to, using cloud services, performing business administration, and administering use of cloud services;
- *Cloud service partner*: 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. A cloud service partner's activities vary depending on the type of partner and their relationship with the cloud service provider and the cloud service customer. Examples of cloud service partners include cloud auditor and cloud service broker;
- *Cloud service provider*: A party which makes cloud services available. The cloud service provider focuses on activities necessary to provide a cloud service and activities necessary to ensure its delivery to the cloud service

customer as well as cloud service maintenance. The cloud service provider includes an extensive set of activities (e.g., provide service, deploy and monitor service, manage business plan, provide audit data, etc.) as well as numerous sub-roles (e.g., business manager, service manager, network provider, security and risk manager, etc.).

4.6 Cloud computing cross cutting aspects

Cross cutting aspects are behaviours or capabilities which need to be coordinated across roles and implemented consistently in a cloud computing system. Such aspects can be shared and may impact multiple roles, activities, and components, in such a way that it is not possible to clearly assign them to individual roles or components, and thus become shared issues across the roles, activities and components. Cross-cutting aspects apply to multiple individual roles or functional components.

Key cross cutting aspects include:

- *Auditability*: 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;
- *Availability*: The property of being accessible and usable upon demand by an authorized entity. The "authorized entity" is typically a cloud service customer;
- *Governance*: The system by which the provision and use of cloud services are directed and controlled. 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 cloud service customer to cloud service provider 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 some

form of agreement between the cloud service customer and the cloud service provider concerning the use of cloud services by the cloud service customer;

- *Interoperability*: Ability of a cloud service customer to interact with a cloud service and exchange information according to a prescribed method and obtain predictable results;
- *Maintenance and versioning*: Maintenance refers to changes to a cloud service or the resources it uses in order to fix faults or in order to upgrade or extend capabilities for business reasons. Versioning implies the appropriate labelling of a service so that it is clear to the cloud service customer that a particular version is in use;
- *Performance*: A set of behaviours relating to the operation of a cloud service, and having metrics defined in a SLA;
- *Portability*: Ability of cloud service customers to move their data or their applications between multiple cloud service providers at low cost and with minimal disruption. The amount of cost and disruption that is acceptable may vary based upon the type of cloud service that is being used;
- *Protection of PII*: Protect the assured, proper, and consistent collection, processing, communication, use and disposal of Personally Identifiable Information (PII) in relation to cloud services;
- *Regulatory*: There are a number of different regulations that may influence the use and delivery of cloud services. Statutory, regulatory, and legal requirements vary by market sector and jurisdiction, and they can change the responsibilities of both cloud service customers and cloud service providers. Compliance with such requirements is often related to governance and risk management activities;
- *Resiliency*: 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;
- *Reversibility*: A process for the cloud service customer to retrieve their cloud service customer data and application artefacts and for the cloud

service provider to delete all cloud service customer data as well as contractually specified cloud service derived data after an agreed period;

- *Security*: Ranges from physical security to application security, and includes requirements such as authentication, authorization, availability, confidentiality, identity management, integrity, non-repudiation, audit, security monitoring, incident response, and security policy management;
- *Service levels and service level agreement*: The cloud computing service level agreement (cloud SLA) is a service level agreement between a cloud service provider and a cloud service customer based on taxonomy of cloud computing specific terms to set the quality of the cloud services delivered. It characterizes quality of the cloud services delivered in terms of: 1) a set of measurable properties specific to cloud computing (business and technical) and 2) a given set of cloud computing roles (cloud service customer and cloud service provider and related sub-roles).

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.