

CAP470: Fundamentals of Cloud Computing

Lecture 1

Cloud Computing Fundamentals

- Definition of Cloud Computing
- Components of Cloud
- The NIST Model
- The Cloud Cube Model

What is Cloud Computing?

Cloud Computing is a construct (infrastructure) that allow you to access application that actually resides at a remote location of other internet connected device, most often, this will be a distant datacenter.

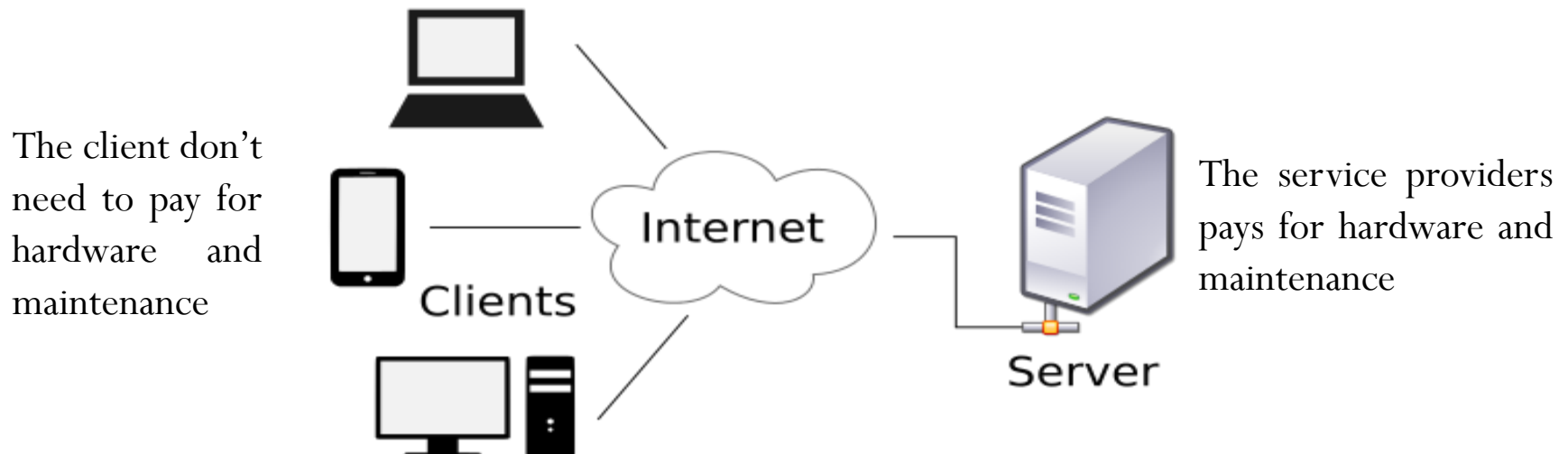
e.g. suppose we want to install MS-Word in our organization's computer. We have to bought the CD/DVD of it an install it or can setup a S/W distribution server to automatically install this application on your machine. Every time Microsoft issued a new version we have to perform same task.

What about the cost?

If someone is using this software only once or twice in a month but he/she has to pay the license cost of the software.

- If some other company hosts your application i.e they handles the cost of servers, manage the software update.
- They charge the customer as per their utilization i.e as per the usage you will pay them
- This will reduce the cost of using that software.
- This will reduce the cost of installation of heavy servers.
- This will reduce the cost of electricity bills.

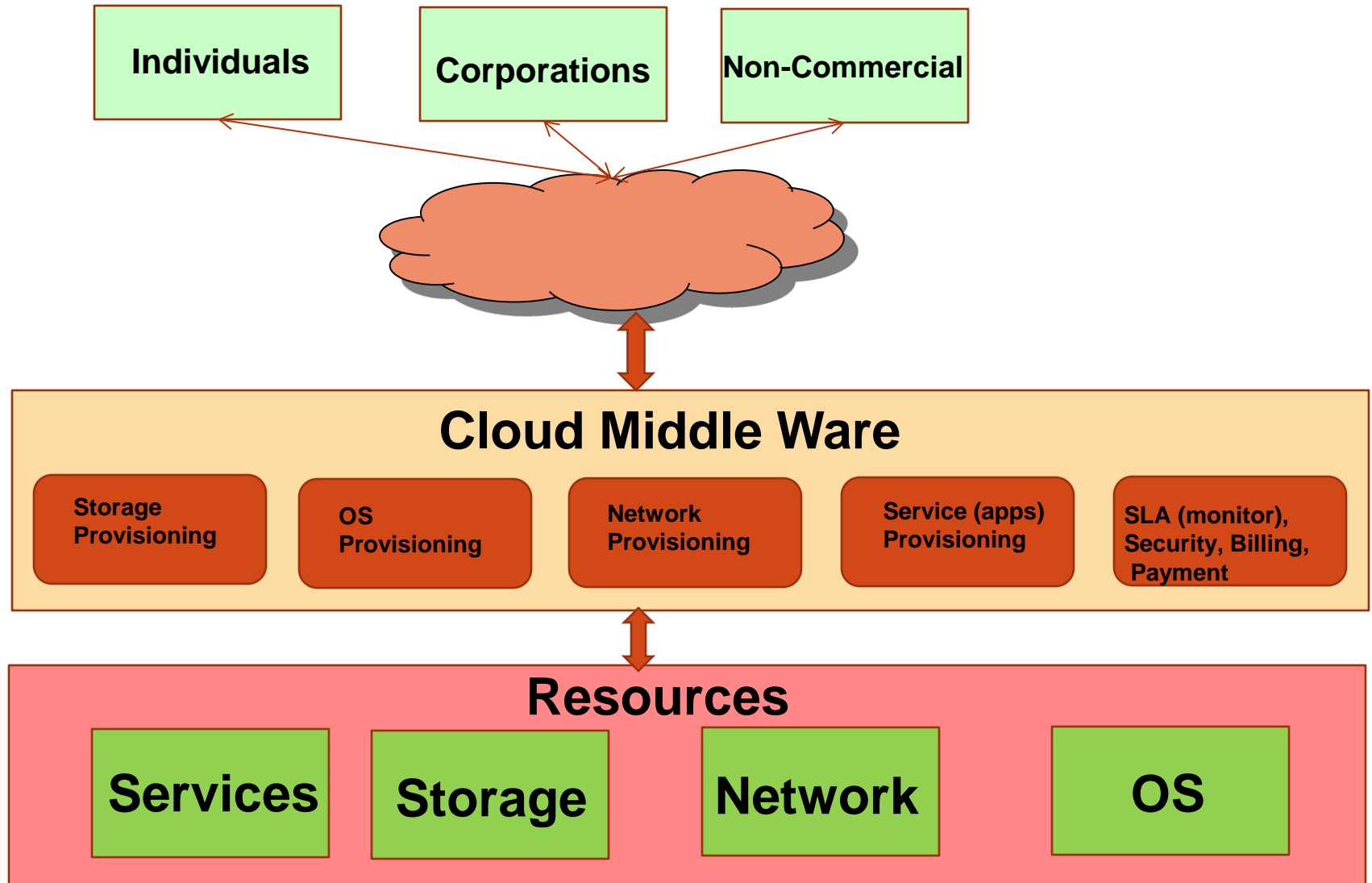
This is the use of cloud computing.



This approach (cloud Computing) have some problems also:

1. In case of internet problem, you cannot access your application and perform your task.
2. There is a issue of data security, because our data will be someone else's control.
3. It is more difficult to integrate your applications if, they are geographically dispersed.

What is a Cloud?



Evolution of Cloud Computing

- 1960 - Benjamin Curley developed the 1st minicomputer.
- 1961 - John McCarthy introduces the cloud computing concept
- 1964 - Douglas Engelbart creates first Windows UI
- 1968 - Foundation of Intel
- 1969 - Development of ARPANET and UNIX
- 1970 - ARPANET developed into the Internet
- 1971 - 1st e-mail sent by Ray Tomlinson
- 1974 - Foundation of Microsoft
- 1976 - Foundation of Apple Computers
- 1977 - Foundation of Oracle
- 1980 - World wide boom in Computers market
- 1981 - Personal Computer launched by IBM
- 1982 - MS launched MS-DOS

- 1984
 - Apple launch Macintosh Computer
 - Foundation of Dell
 - William Gibson coined the term “Cyberspace”
- 1985
 - Windows 1.0 launched
- 1989
 - Compaq release first notebook
- 1990
 - Internet age begins
 - First web browser invented by Sir Tim Berners-Lee.
 - His browser was called **WorldWideWeb** and later renamed Nexus.
- 1991
 - CERN release the internet for general use
- 1993
 - Mosaic browser launched
- 1994
 - Netscape founded
- 1995
 - Foundation of Amazon and eBay
- 1996
 - Palm Pilot PDA launched
- 1999
 - Salesforce.com delivers business app Napster launched
- 2000
 - Cloud era started

- 2002 - Amazon launch Mechanical turk
RIM launch Blackberry
- 2004 - Foundation of facebook
- 2006 - Amazon launch EC2/S3 (pay-as-you-go)
- 2007 - Salesforce launch fore.com
Apple launch iPhone
- 2008 - HTC launch first android phone
MS announced Azure
- 2009 - Google apps launch
- 2010 - Salesforce launch databox.com and chatter
MS released Windows Azure
Samsung launch 1st android tablet
- 2011 - IBM announced the IBM SmartCloud framework to support Smarter Planet
- 2012 - Oracle announced the Oracle Cloud

Cloud Computing

Cloud computing takes the technology, services, and applications that are similar to those on the Internet and turns them into a self-service utility. The use of the word “cloud” makes reference to the two essential concepts:

- **Abstraction:** Cloud computing abstracts the details of system implementation from users and developers. Applications run on physical systems that aren't specified, data is stored in locations that are unknown, administration of systems is outsourced to others, and access by users is ubiquitous.
- **Virtualization:** Cloud computing virtualizes systems by pooling and sharing resources. Systems and storage can be provisioned as needed from a centralized infrastructure, costs are assessed on a metered basis, multi-tenancy is enabled, and resources are scalable with agility.

- Cloud computing is an abstraction based on the notion of pooling physical resources and presenting them as a virtual resource.
- It is a new model for provisioning resources, for staging applications, and for platform-independent user access to services.
- Clouds can come in many different types, and the services and applications that run on clouds may or may not be delivered by a cloud service provider.
- These different types and levels of cloud services mean that it is important to define what type of cloud computing system you are working with.

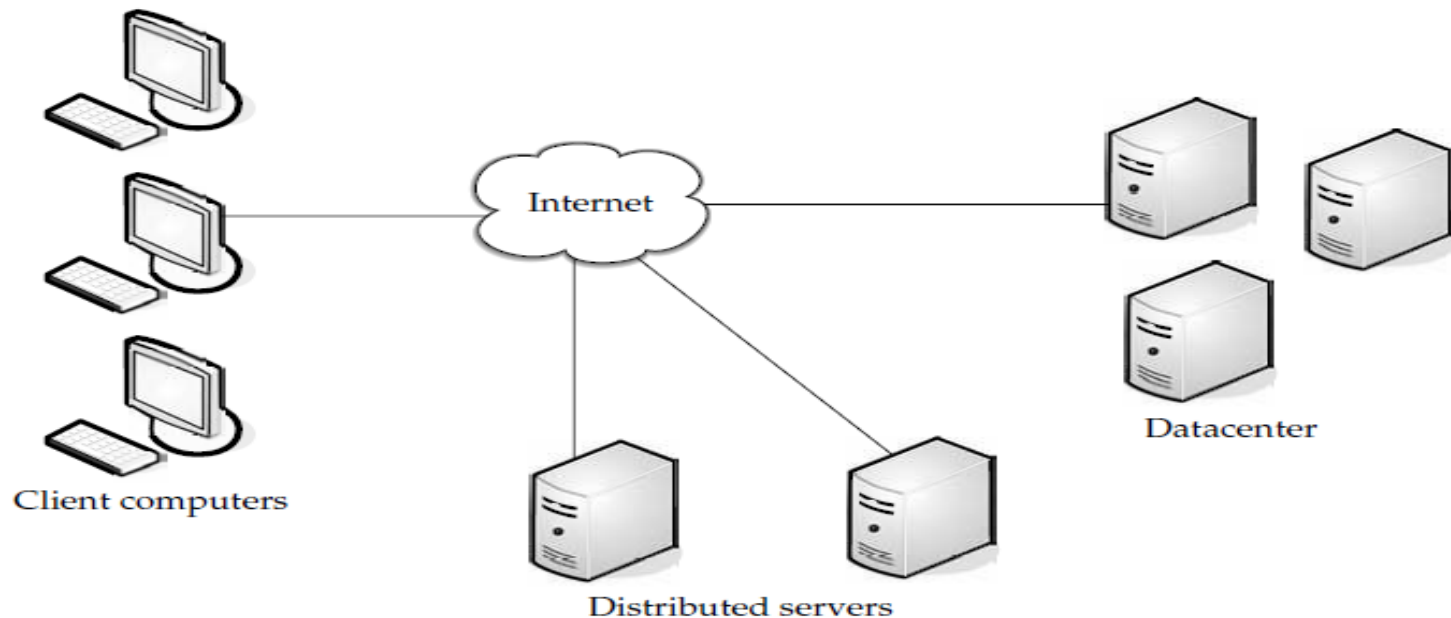
Definition as per NIST

“Cloud computing is a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computer resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction.”

Components of Cloud

In a simple, topological sense, a cloud computing solution is made up of several elements and these elements make up the three components of a cloud computing solution.:

- clients
- the data center, and
- distributed servers



A. Clients

Clients are the devices that the end users interact with to manage their information on the cloud.

Clients generally fall into three categories:

- **Mobile Clients:** Mobile devices include PDAs or smartphones, like a Blackberry, Windows Mobile Smartphone, or an iPhone.
- **Thin Clients:** Clients are computers that do not have internal hard drives, but rather let the server do all the work, but then display the information.
- **Thick Clients:** This type of client is a regular computer, using a web browser like Firefox or Internet Explorer to connect to the cloud.

Thin clients are becoming an increasingly popular solution, because of their price and effect on the environment. Some benefits to using thin clients include

- **Lower hardware costs** Thin clients are cheaper than thick clients because they do not contain as much hardware. They also last longer before they need to be upgraded or become obsolete.
- **Lower IT costs** Thin clients are managed at the server and there are fewer points of failure.
- **Security** Since the processing takes place on the server and there is no hard drive, there's less chance of malware invading the device. Also, since thin clients don't work without a server, there's less chance of them being physically stolen.

- **Data security** Since data is stored on the server, there's less chance for data to be lost if the client computer crashes or is stolen.
- **Less power consumption** Thin clients consume less power than thick clients. This means you'll pay less to power them, and you'll also pay less to air-condition the office.
- **Ease of repair or replacement** If a thin client dies, it's easy to replace. The box is simply swapped out and the user's desktop returns exactly as it was before the failure.
- **Less noise** Without a spinning hard drive, less heat is generated and quieter fans can be used on the thin client.

B. Datacenter

- The *datacenter* is the collection of servers where the application to which you subscribe is housed.
 - It could be a large room in the basement of your building or a room full of servers on the other side of the world that you access via the Internet.
- A growing trend in the IT world is virtualizing servers. That is, software can be installed allowing multiple instances of virtual servers to be used.
 - In this way, you can have half a dozen virtual servers running on one physical server.

C. Distributed Servers

- But the servers don't all have to be housed in the same location.
 - Often, servers are in geographically disparate locations.
 - But for the cloud subscriber, these servers act as if they're humming away right next to each other.
- This gives the service provider more flexibility in options and security.
 - For instance, Amazon has their cloud solution in servers all over the world.
 - If something were to happen at one site, causing a failure, the service would still be accessed through another site.
 - Also, if the cloud needs more hardware, they need not throw more servers in the safe room—they can add them at another site and simply make it part of the cloud.

Characteristics of Cloud Computing

(As per National Institute of Standards and Technology)

NIST identifies five essential characteristics of the cloud, summarized here:

- 1. On-demand self-service**
- 2. Broad network access**
- 3. Resource pooling**
- 4. Rapid elasticity**
- 5. Measured service**

1. **On-demand self-service** – A user can provision computing capabilities, such as server time and storage, as needed without requiring human interaction.
2. **Broad network access** – Capabilities are available over a network and typically accessed by the users' mobile phones, tablets, laptops, and workstations.
3. **Resource pooling** – The provider's computing resources are pooled to serve multiple users using a multi-tenant model, with different physical and virtual resources dynamically assigned and reassigned according to consumer demand.
 - Examples of resources include storage, processing, memory, and network bandwidth.

4. **Rapid elasticity** – Capabilities can be elastically provisioned and released, in some cases automatically, to scale rapidly outward and inward as needed.

- For the user, the capabilities available for provisioning often appear to be unlimited and can be appropriated in any quantity at any time.

5. **Measured service** – Cloud systems automatically control and optimize resource use by leveraging a metering capability appropriate to the type of service (e.g., storage, processing, bandwidth, and active user accounts).

- Resource usage can be monitored, controlled, and reported, providing transparency for both the provider and user of the service.
- This cloud characteristic enables a cloud user to consume the service in a “pay as you grow” model or for internal IT departments to provide IT chargeback capabilities.

Cloud Model Types

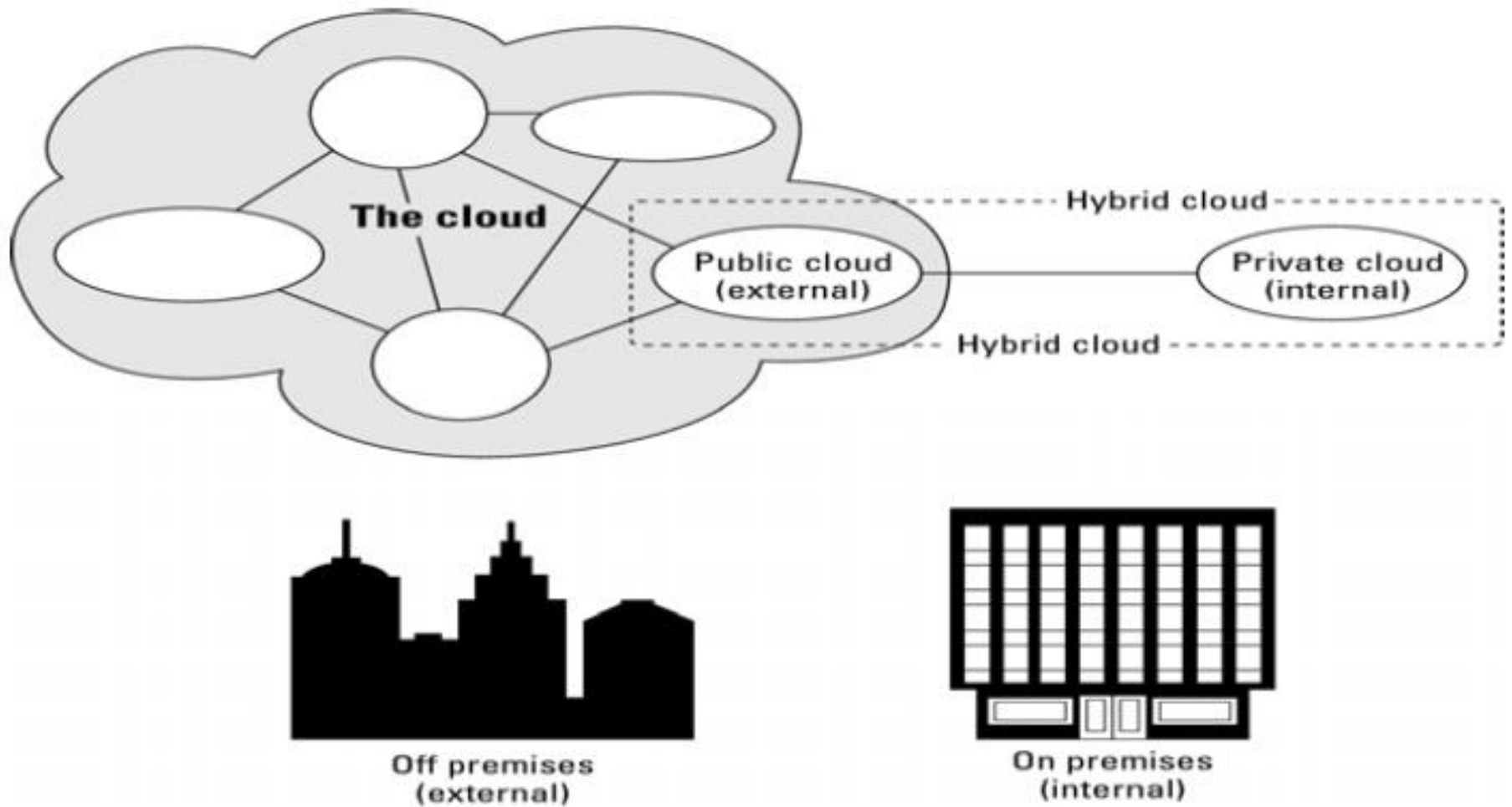
Cloud computing is divided into two distinct sets of models:

- **Deployment models:** This refers to the location and management of the cloud's infrastructure.
 - Public cloud
 - Private cloud
 - Hybrid cloud
 - Community cloud
- **Service models:** This consists of the particular types of services that you can access on a cloud computing platform.
 - Infrastructure as a Service
 - Platform as a Service
 - Software as a Service

Deployment Models

The NIST definition for the four deployment models is as follows:

- **Public cloud:** The public cloud infrastructure is available for public use alternatively for a large industry group and is owned by an organization selling cloud services.
- **Private cloud:** The private cloud infrastructure is operated for the exclusive use of an organization. The cloud may be managed by that organization or a third party. Private clouds may be either on or off premises.
- **Hybrid cloud:** A hybrid cloud combines multiple clouds (private, community or public) where those clouds retain their unique identities, but are bound together as a unit. A hybrid cloud may offer standardized or proprietary access to data and applications, as well as application portability.
- **Community cloud:** A community cloud is one where the cloud has been organized to serve a common function or purpose.



Deployment Locations of Different Cloud types

Service Models

Three service models have been universally accepted:

- **Infrastructure as a Service:** IaaS provides virtual machines, virtual storage, virtual infrastructure, and other hardware assets as resources that clients can provision.
 - The IaaS service provider manages all the infrastructure, while the client is responsible for all other aspects of the deployment. This can include the operating system, applications, and user interactions with the system.
- **Platform as a Service:** PaaS provides virtual machines, operating systems, applications, services, development frameworks, transactions, and control structures. The client can deploy its applications on the cloud infrastructure or use applications that were programmed using languages and tools that are supported by the PaaS service provider.
 - The service provider manages the cloud infrastructure, the operating systems, and the enabling software. The client is responsible for installing and managing the application that it is deploying.
- **Software as a Service:** SaaS is a complete operating environment with applications, management, and the user interface.

The three different service models taken together have come to be known as the SPI model of cloud computing.

Many other service models have also been mentioned:

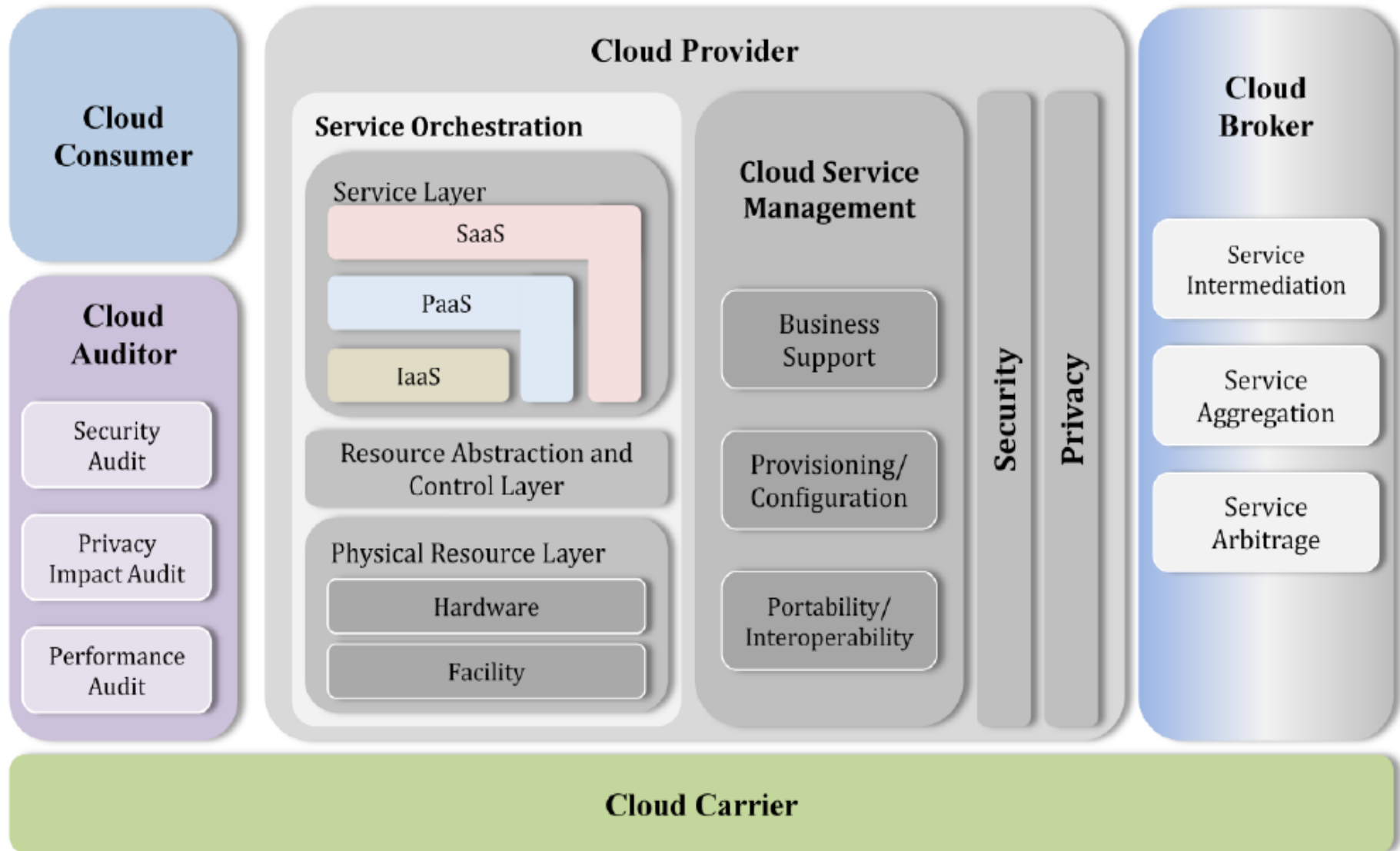
- StaaS - Storage as a Service;
- IdaaS - Identity as a Service;
- CmaaS - Compliance as a Service; and so forth.

However, the SPI services encompass all the other possibilities.

Business Models

- NIST Cloud Computing Reference Model
- Cloud Cube Model

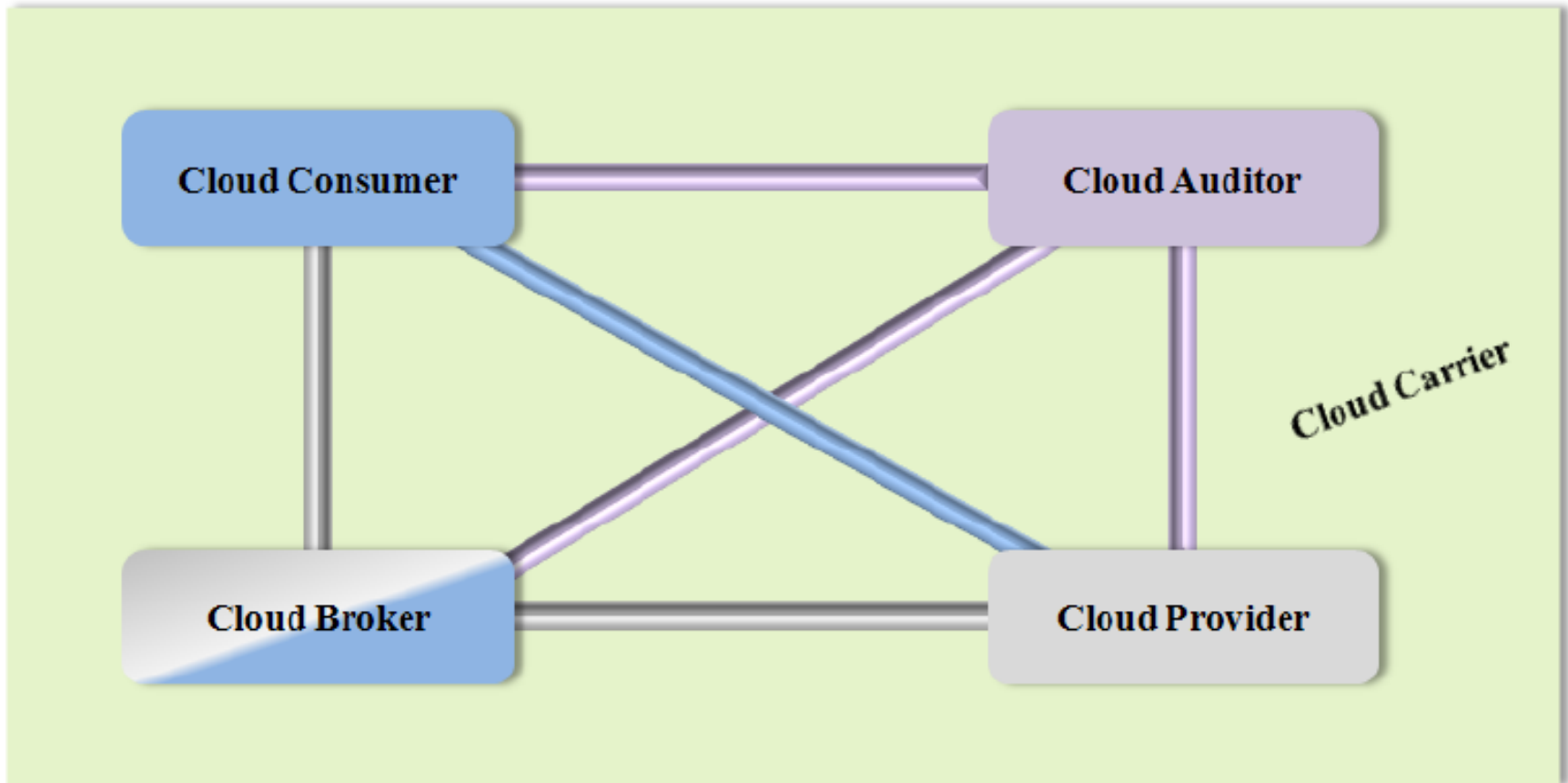
NIST Cloud Computing Reference Model






Actors in Cloud Computing

| Actor | Definition |
|-----------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Cloud Consumer | A person or organization that maintains a business relationship with, and uses service from, <i>Cloud Providers</i> . |
| Cloud Provider | A person, organization, or entity responsible for making a service available to interested parties. |
| Cloud Auditor | A party that can conduct independent assessment of cloud services, information system operations, performance and security of the cloud implementation. |
| Cloud Broker | An entity that manages the use, performance and delivery of cloud services, and negotiates relationships between <i>Cloud Providers</i> and <i>Cloud Consumers</i> . |
| Cloud Carrier | An intermediary that provides connectivity and transport of cloud services from <i>Cloud Providers</i> to <i>Cloud Consumers</i> . |

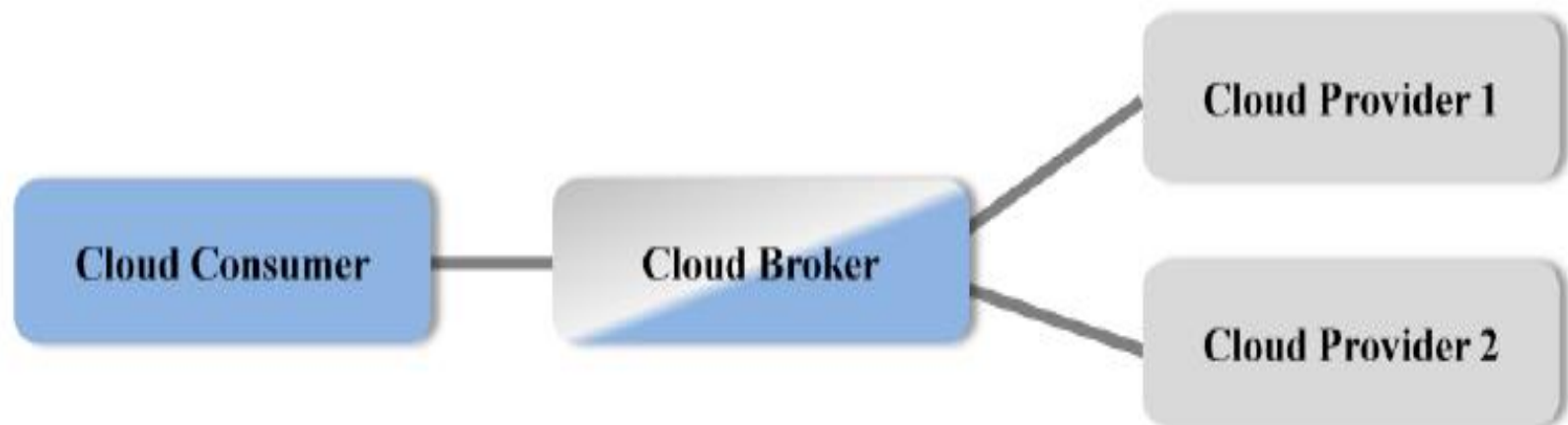
Interactions between the Actors in Cloud Computing



-  The communication path between a cloud provider and a cloud consumer
-  The communication paths for a cloud auditor to collect auditing information
-  The communication paths for a cloud broker to provide service to a cloud consumer

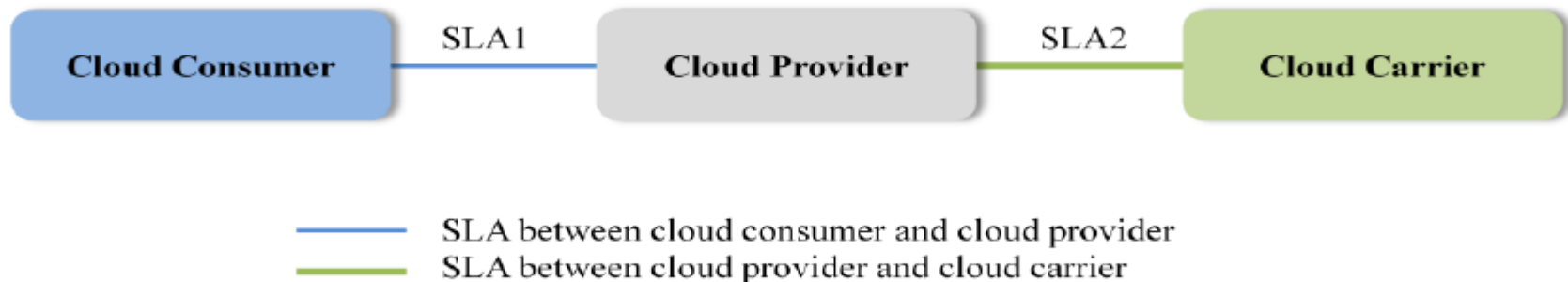
Usage Scenario for Cloud Brokers

A cloud consumer may request service from a cloud broker instead of contacting a cloud provider directly. The cloud broker may create a new service by combining multiple services or by enhancing an existing service. In this example, the actual cloud providers are invisible to the cloud consumer and the cloud consumer interacts directly with the cloud broker.



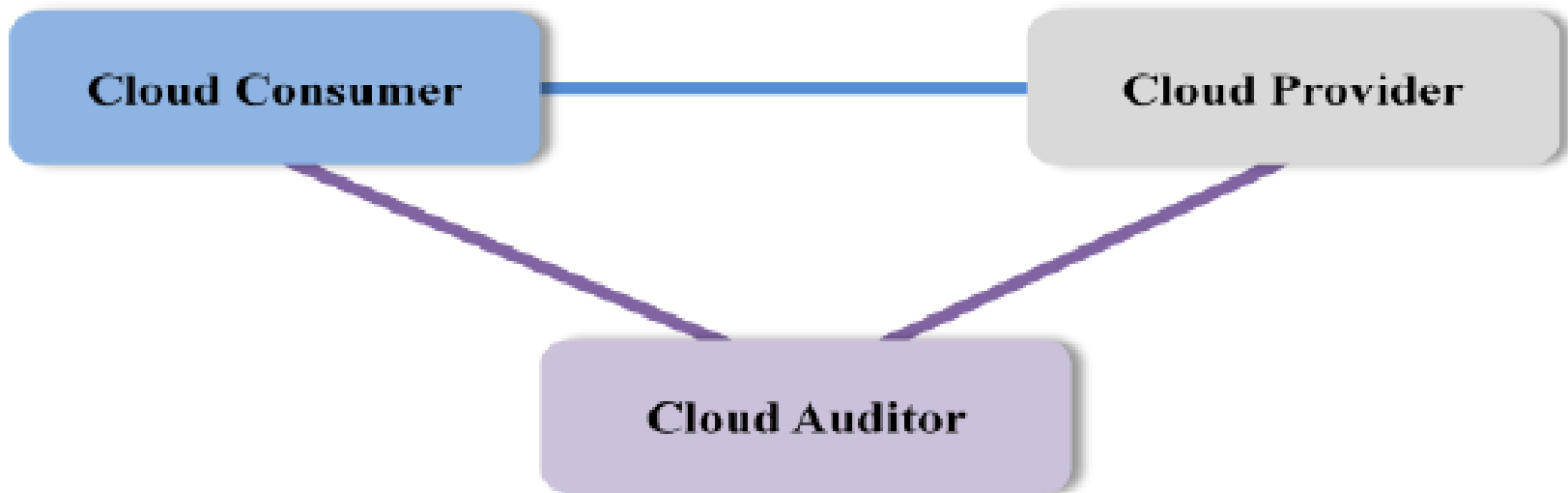
Usage Scenario for Cloud Carriers

Cloud carriers provide the connectivity and transport of cloud services from cloud providers to cloud consumers. As illustrated in Figure, a cloud provider participates in and arranges for two unique service level agreements (SLAs), one with a cloud carrier (e.g. SLA2) and one with a cloud consumer (e.g. SLA1). A cloud provider arranges service level agreements (SLAs) with a cloud carrier and may request dedicated and encrypted connections to ensure the cloud services are consumed at a consistent level according to the contractual obligations with the cloud consumers. In this case, the provider may specify its requirements on capability, flexibility and functionality in SLA2 in order to provide essential requirements in SLA1.



Usage Scenario for Cloud Auditors

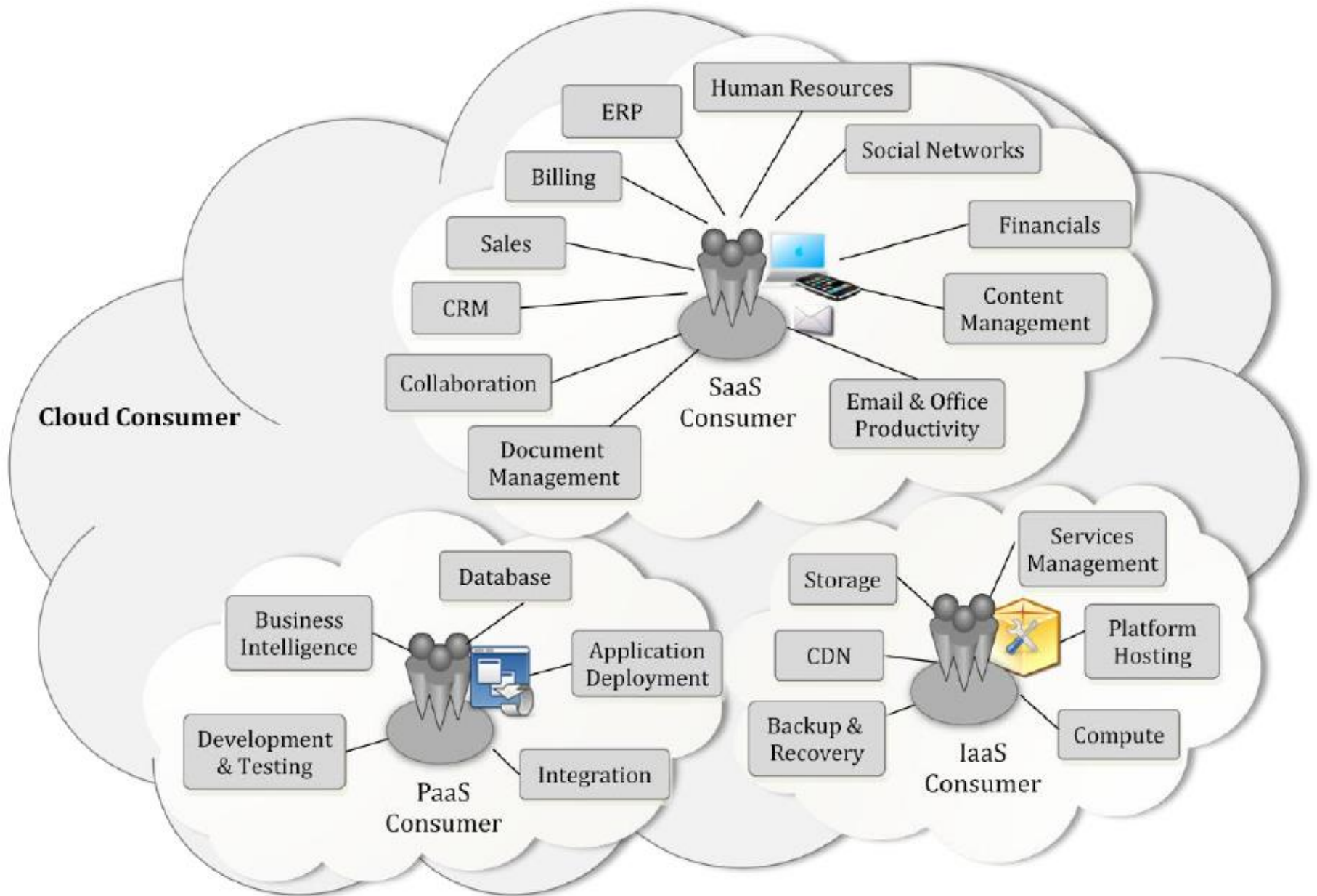
For a cloud service, a cloud auditor conducts independent assessments of the operation and security of the cloud service implementation. The audit may involve interactions with both the Cloud Consumer and the Cloud Provider.



1. Cloud Consumer

The cloud consumer is the principal stakeholder for the cloud computing service.

- A cloud consumer represents a person or organization that maintains a business relationship with, and uses the service from a cloud provider.
- A cloud consumer browses the service catalogue from a cloud provider, requests the appropriate service, sets up service contracts with the cloud provider, and uses the service.
- The cloud consumer may be billed for the service provisioned, and needs to arrange payments accordingly.



Example Services Available to a Cloud Consumer

2. Cloud Provider

A cloud provider is a person, an organization; it is the entity responsible for making a service available to interested parties. A Cloud Provider:

- acquires and manages the computing infrastructure required for providing the services,
- runs the cloud software that provides the services, and
- makes arrangement to deliver the cloud services to the Cloud Consumers through network access.

A Cloud Provider's activities can be described in five major areas:

- *service deployment,*
- *service orchestration,*
- *cloud service management,*
- *security, and*
- *privacy.*

3. Cloud Auditor

A cloud auditor is a party that can perform an independent examination of cloud service controls with the intent to express an opinion thereon.

- Audits are performed to verify conformance to standards through review of objective evidence.
- A cloud auditor can evaluate the services provided by a cloud provider in terms of security controls, privacy impact, performance, etc.
- The auditor may ensure that fixed content has not been modified and that the legal and business data archival requirements have been satisfied.

4. Cloud Broker

As cloud computing evolves, the integration of cloud services can be too complex for cloud consumers to manage.

- A cloud consumer may request cloud services from a cloud broker, instead of contacting a cloud provider directly.
- A cloud broker is an entity that manages the use, performance and delivery of cloud services and negotiates relationships between cloud providers and cloud consumers.

In general, a cloud broker can provide services in three categories:

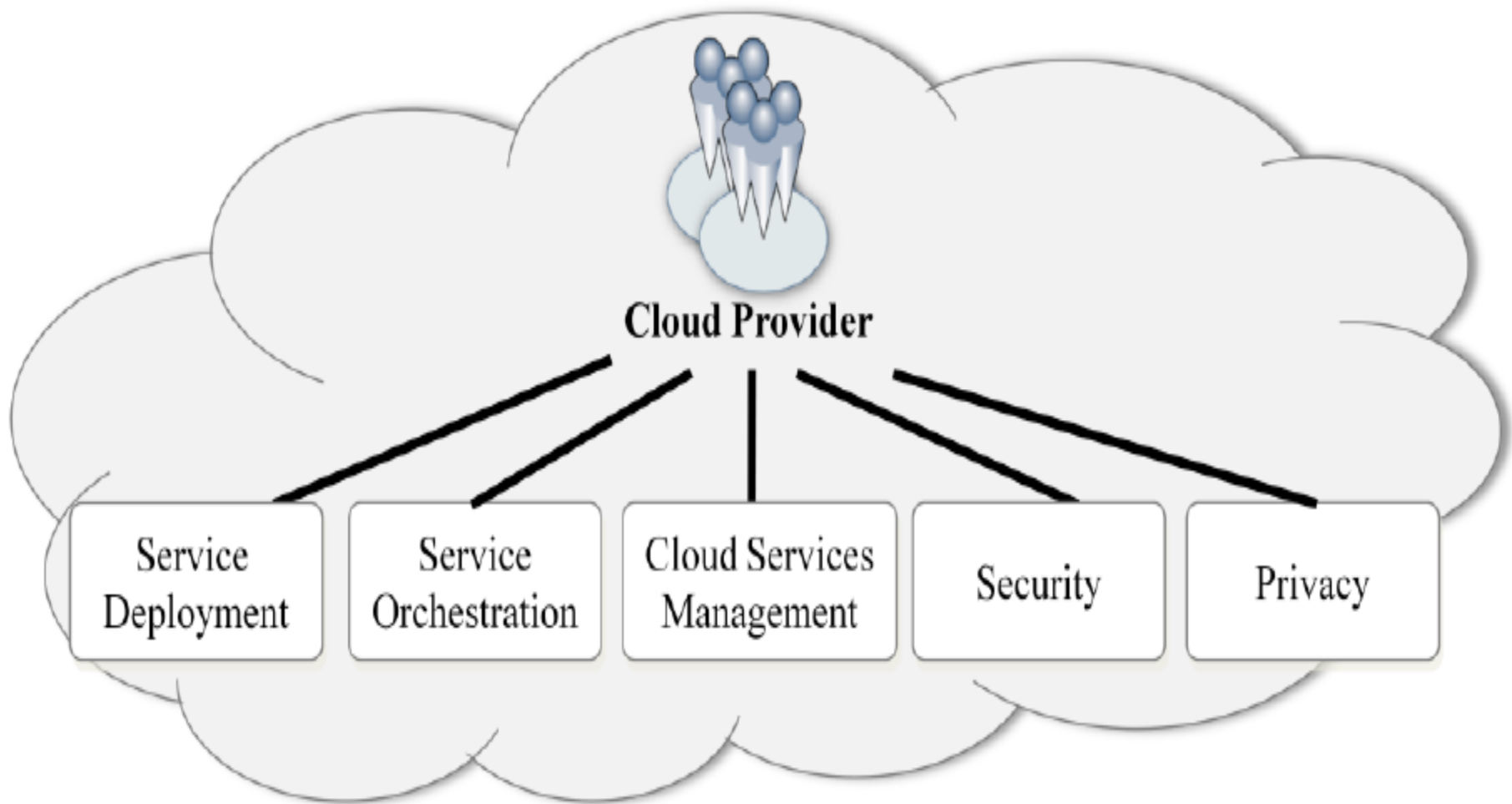
- **Service Intermediation:** A cloud broker enhances a given service by improving some specific capability and providing value-added services to cloud consumers.
 - The improvement can be managing access to cloud services, identity management, performance reporting, enhanced security, etc.
- **Service Aggregation:** A cloud broker combines and integrates multiple services into one or more new services.
 - The broker provides data integration and ensures the secure data movement between the cloud consumer and multiple cloud providers.
- **Service Arbitrage:** Service arbitrage is similar to service aggregation except that the services being aggregated are not fixed. Service arbitrage means a broker has the flexibility to choose services from multiple agencies.
 - The cloud broker, for example, can use a credit-scoring service to measure and select an agency with the best score.

5. Cloud Carrier

A cloud carrier acts as an intermediary that provides connectivity and transport of cloud services between cloud consumers and cloud providers.

- Cloud carriers provide access to consumers through network, telecommunication and other access devices.
 - For example, cloud consumers can obtain cloud services through network access devices, such as computers, laptops, mobile phones, mobile Internet devices (MIDs), etc.
- The distribution of cloud services is normally provided by network and telecommunication carriers or a transport agent, where a transport agent refers to a business organization that provides physical transport of storage media such as high-capacity hard drives.
 - Note that a cloud provider will set up SLAs with a cloud carrier to provide services consistent with the level of SLAs offered to cloud consumers, and may require the cloud carrier to provide dedicated and secure connections between cloud consumers and cloud providers.

Major activities of Cloud Provider



A. Service Deployment

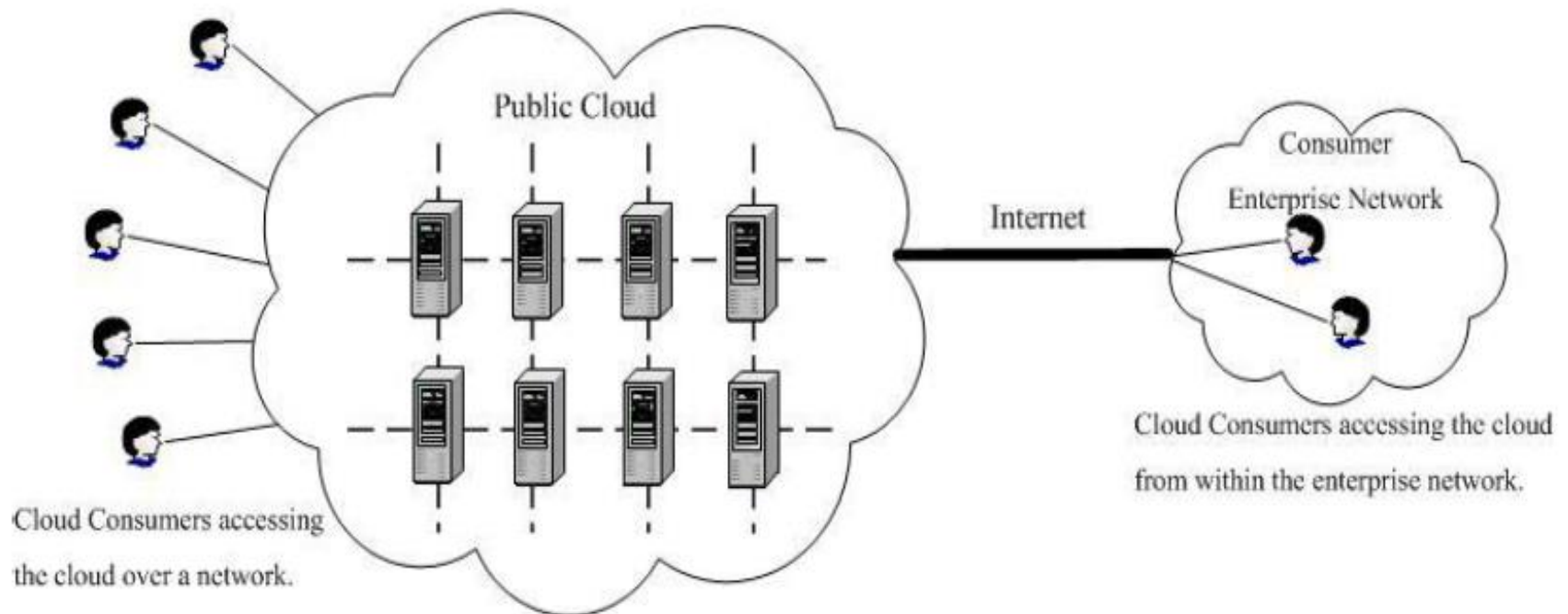
A cloud infrastructure may be operated in one of the following deployment models:

- public cloud,
- private cloud,
- community cloud, or
- hybrid cloud.

The differences are based on how exclusive the computing resources are made to a Cloud Consumer.

Public Cloud

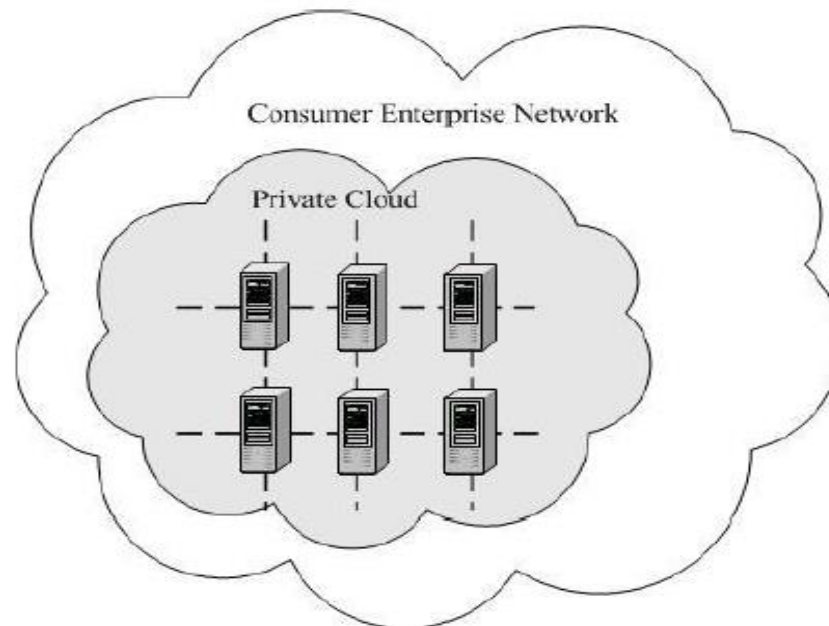
A public cloud is one in which the cloud infrastructure and computing resources are made available to the general public over a public network. A public cloud is owned by an organization selling cloud services, and serves a diverse pool of clients.



Private Cloud

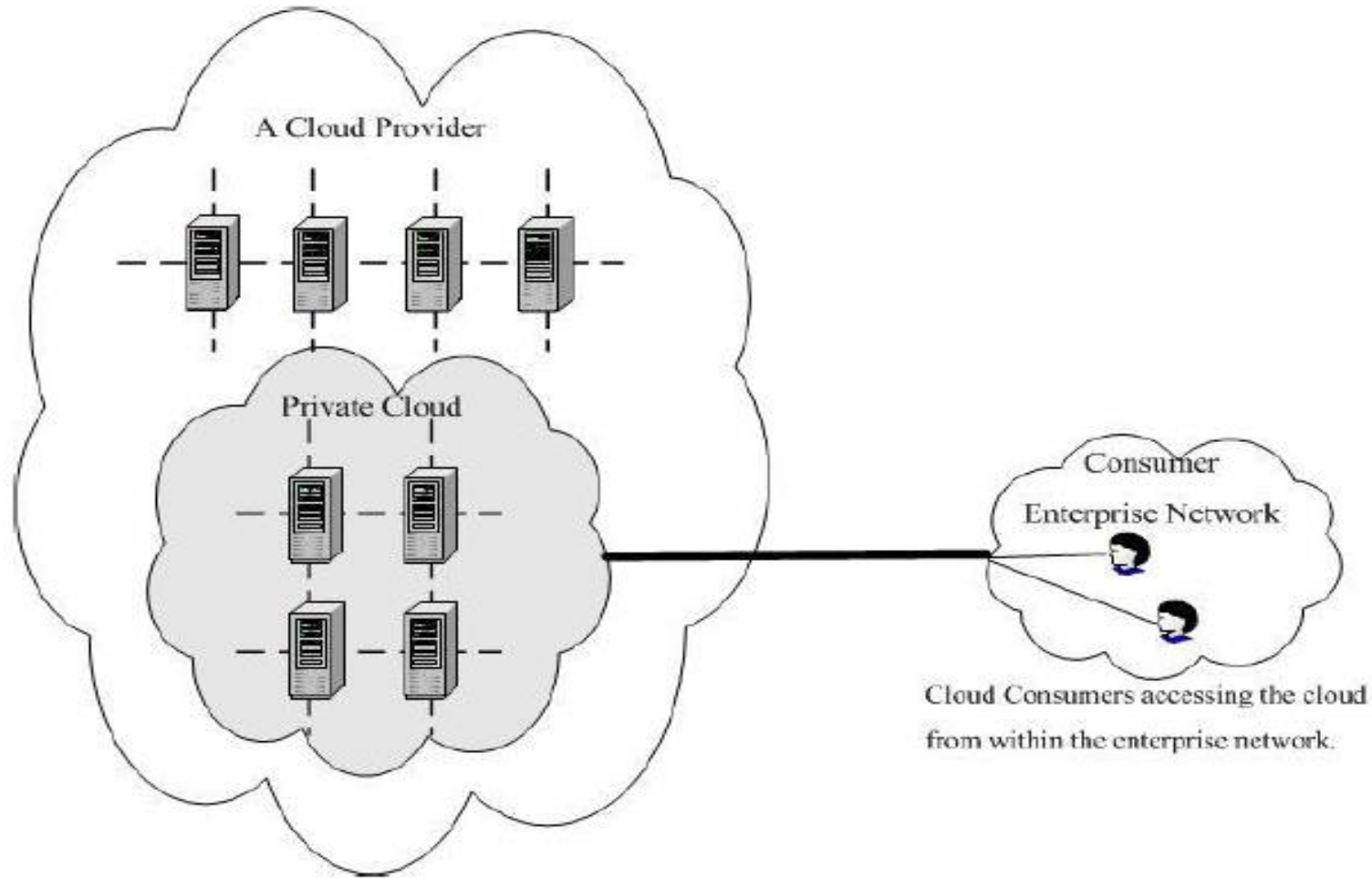
A private cloud gives a single Cloud Consumer's organization the exclusive access to and usage of the infrastructure and computational resources. It may be managed either by:

- The Cloud Consumer organization and may be hosted on the organization's premises (i.e. *on-site private clouds*), or



On-site Private Cloud

- A third party, outsourced to a hosting company (i.e. *outsourced private clouds*).



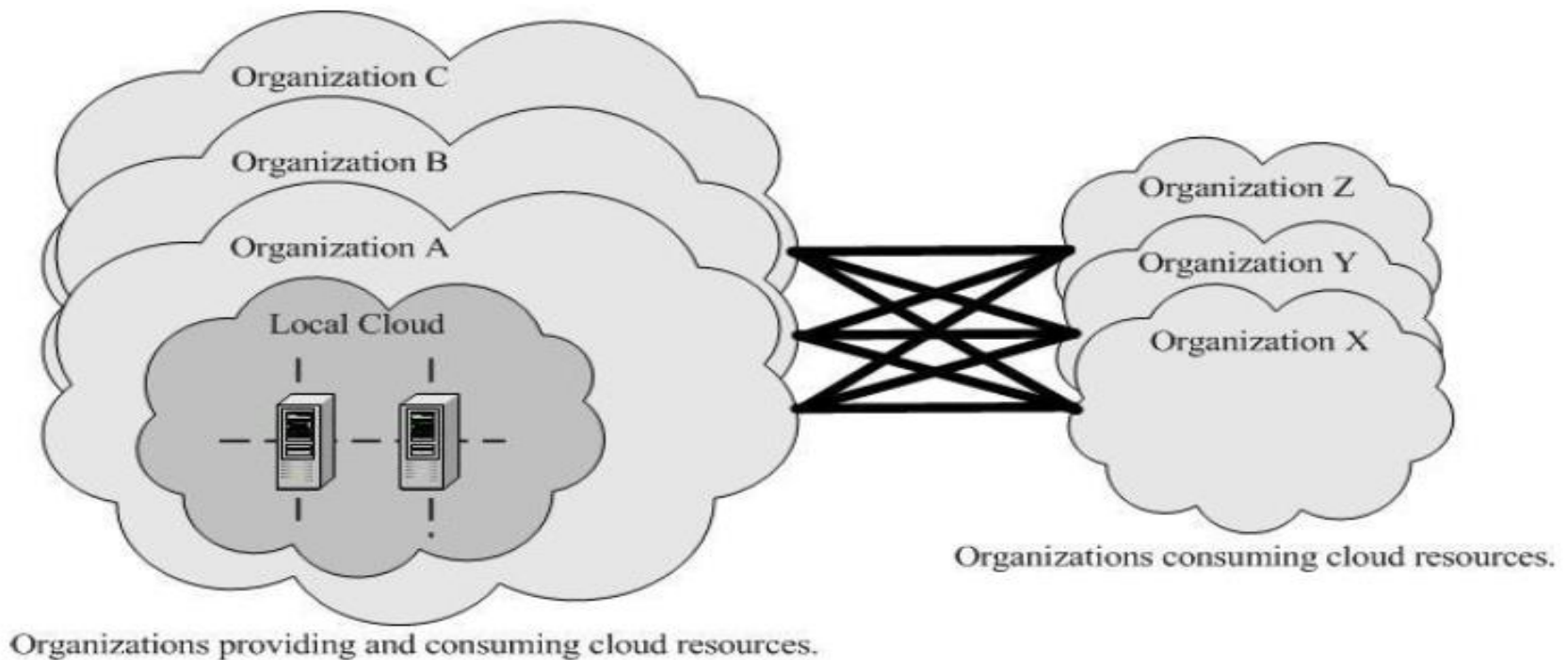
Outsourced Private Cloud

Community Cloud

A community cloud serves a group of Cloud Consumers which have shared concerns such as mission objectives, security, privacy and compliance policy, rather than serving a single organization as does a private cloud. Similar to private clouds, a community cloud may be managed by

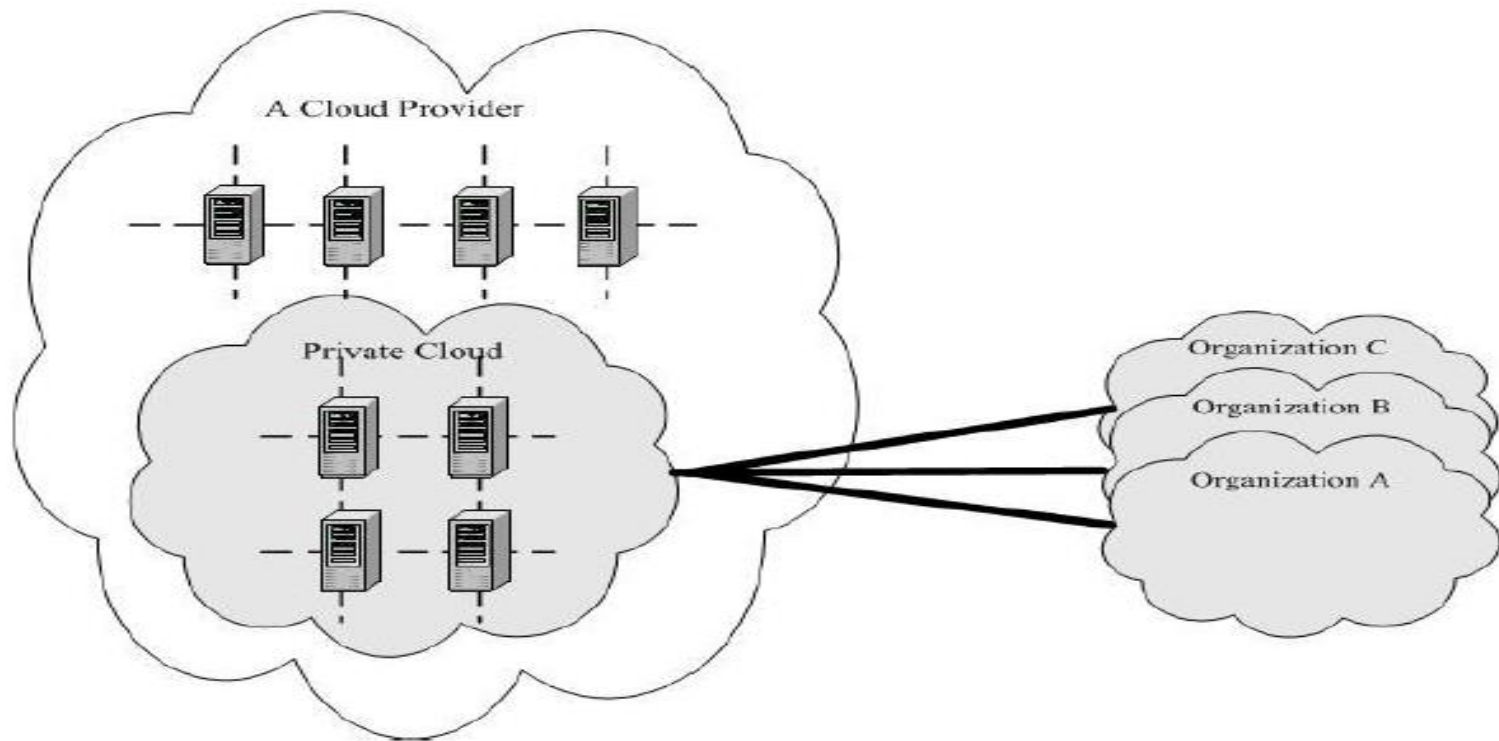
- The organizations and may be implemented on customer premise (i.e. *on-site community cloud*), or
- A third party, outsourced to a hosting company (i.e. *outsourced community cloud*).

Following figure depicts an on-site community cloud comprised of a number of participant organizations. A cloud consumer can access the local cloud resources, and also the resources of other participating organizations through the connections between the associated organizations.



On-site Community Cloud

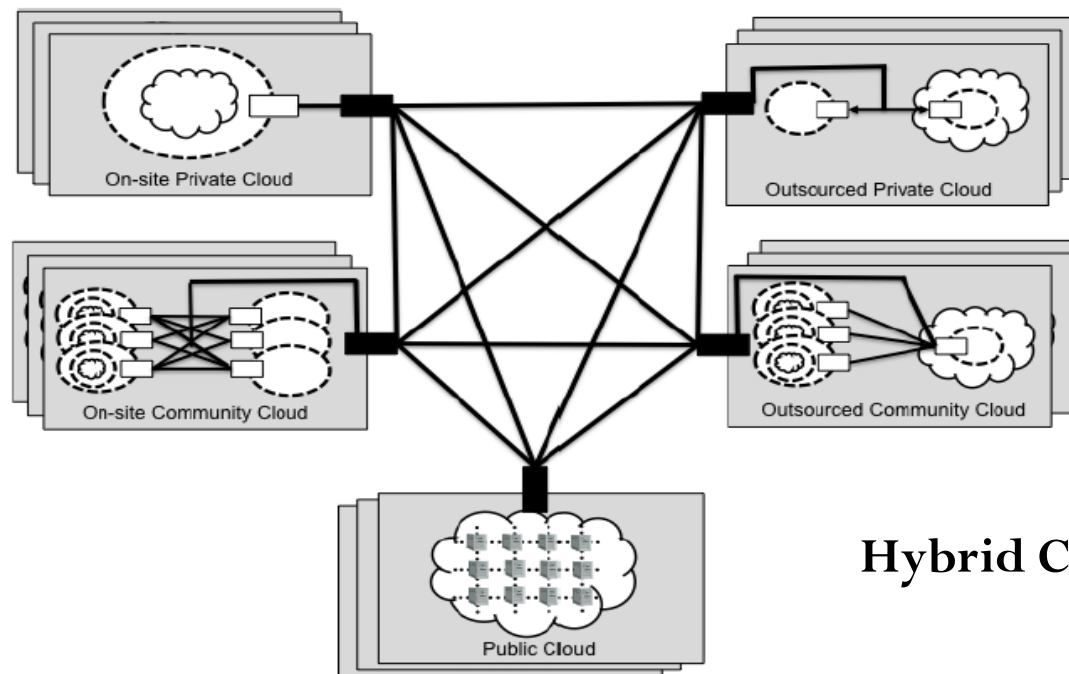
Following figure shows an outsourced community cloud, where the server side is outsourced to a hosting company. In this case, an outsourced community cloud builds its infrastructure off premise, and serves a set of organizations that request and consume cloud services.



Outsourced Community Cloud

Hybrid Cloud

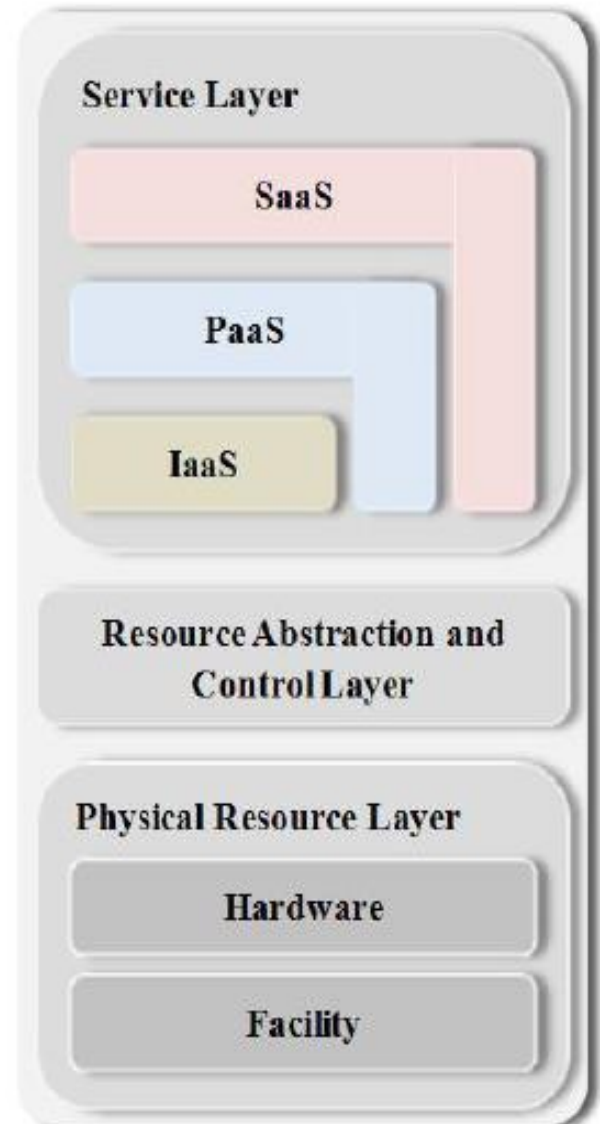
A hybrid cloud is a composition of two or more clouds (on-site private, on-site community, off-site private, off-site community or public) that remain as distinct entities but are bound together by standardized or proprietary technology that enables data and application portability.



Hybrid Cloud

B. Cloud Service Orchestration

Service Orchestration refers to the composition of system components to support the Cloud Providers activities in arrangement, coordination and management of computing resources in order to provide cloud services to Cloud Consumers. Figure shows a generic stack diagram of this composition that underlies the provisioning of cloud services. A three-layered model is used in this representation, representing the grouping of three types of system components Cloud Providers need to compose to deliver their services.



Service layer

This is where Cloud Providers define interfaces for Cloud Consumers to access the computing services.

- Access interfaces of each of the three service models are provided in this layer. It is possible, though not necessary, that SaaS applications can be built on top of PaaS components and PaaS components can be built on top of IaaS components.
- The optional dependency relationships among SaaS, PaaS, and IaaS components are represented graphically as components stacking on each other; while the angling of the components represents that each of the service component can stand by itself.

Resource Abstraction and Control Layer

This layer contains the system components that Cloud Providers use to provide and manage access to the physical computing resources through software abstraction.

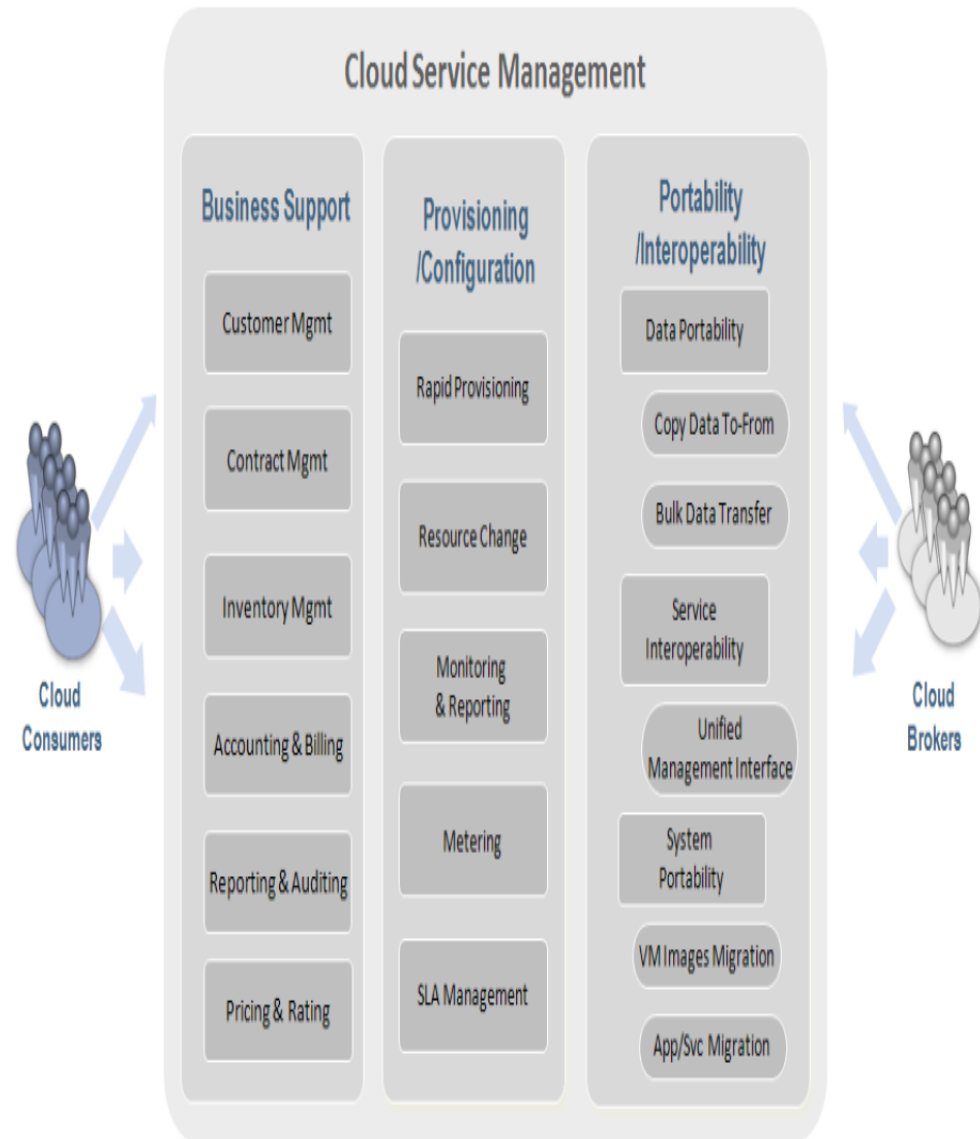
- The **resource abstraction** needs to ensure efficient, secure, and reliable usage of the underlying physical resources. While virtual machine technology is commonly used at this layer, other means of providing the necessary software abstractions are also possible.
 - Examples of resource abstraction components include software elements such as hypervisors, virtual machines, virtual data storage, and other computing resource abstractions.
- The **control** aspect of this layer refers to the software components that are responsible for resource allocation, access control, and usage monitoring.
 - This is the software fabric that ties together the numerous underlying physical resources and their software abstractions to enable resource pooling, dynamic allocation, and measured service.
 - Various open source and proprietary cloud software are examples of this type of middleware.

Physical Resource Layer

- This layer includes **hardware resources**, such as computers (CPU and memory), networks (routers, firewalls, switches, network links and interfaces), storage components (hard disks) and other physical computing infrastructure elements.
- It also includes **facility resources**, such as heating, ventilation and air conditioning (HVAC), power, communications, and other aspects of the physical plant.

C. Cloud Service Management

Cloud Service Management includes all of the service-related functions that are necessary for the management and operation of those services required by or proposed to cloud consumers. As illustrated in Figure, cloud service management can be described from the perspective of *business support*, *provisioning and configuration*, and from the perspective of *portability and interoperability* requirements.



Business Support

- *Business Support* entails the set of business-related services dealing with clients and supporting processes. It includes the components used to run business operations that are client-facing.
- *Customer management*: Manage customer accounts, open/close/terminate accounts, manage user profiles, manage customer relationships by providing points-of-contact and resolving customer issues and problems, etc.
- *Contract management*: Manage service contracts, setup/negotiate/close/terminate contract, etc.
- *Inventory Management*: Set up and manage service catalogues, etc.
- *Accounting and Billing*: Manage customer billing information, send billing statements, process received payments, track invoices, etc.
- *Reporting and Auditing*: Monitor user operations, generate reports, etc.
- *Pricing and Rating*: Evaluate cloud services and determine prices, handle promotions and pricing rules based on a user's profile, etc.

Provisioning and Configuration

- *Rapid provisioning*: Automatically deploying cloud systems based on the requested service/resources/capabilities.
- *Resource changing*: Adjusting configuration/resource assignment for repairs, upgrades and joining new nodes into the cloud.
- *Monitoring and Reporting*: Discovering and monitoring virtual resources, monitoring cloud operations and events and generating performance reports.
- *Metering*: Providing a metering capability at some level of abstraction appropriate to the type of service (e.g., storage, processing, bandwidth, and active user accounts).
- *SLA management*: Encompassing the SLA contract definition (basic schema with the QoS parameters), SLA monitoring and SLA enforcement according to defined policies.

Portability and Interoperability

The proliferation of cloud computing promises cost savings in technology infrastructure and faster software upgrades. The US government, along with other potential cloud computing customers, has a strong interest in moving to the cloud. However, the adoption of cloud computing depends greatly on how the cloud can address users concerns on security, portability and interoperability.

- For portability, prospective customers are interested to know whether they can move their data or applications across multiple cloud environments at low cost and minimal disruption.
- From an interoperability perspective, users are concerned about the capability to communicate between or among multiple clouds.

Cloud providers should provide mechanisms to support *data portability*, *service interoperability*, and *system portability*.

- **Data portability** is the ability of cloud consumers to copy data objects into or out of a cloud or to use a disk for bulk data transfer.
- **Service interoperability** is the ability of cloud consumers to use their data and services across multiple cloud providers with a unified management interface.
- **System portability** allows the migration of a fully-stopped virtual machine instance or a machine image from one provider to another provider, or migrate applications and services and their contents from one service provider to another.

It should be noted that various cloud service models may have different requirements in related with portability and interoperability.

For example,

- IaaS requires the ability to migrate the data and run the applications on a new cloud. Thus, it is necessary to capture virtual machine images and migrate to new cloud providers which may use different virtualization technologies. Any provider-specific extensions to the VM images need to be removed or recorded upon being ported.
- While for SaaS, the focus is on data portability, and thus it is essential to perform data extractions and backups in a standard format.

D. Security

- It is critical to recognize that security is a cross-cutting aspect of the architecture that spans across all layers of the reference model, ranging from physical security to application security.
- Therefore, security in cloud computing architecture concerns is not solely under the purview of the Cloud Providers, but also Cloud Consumers and other relevant actors.
- Cloud-based systems still need to address security requirements such as authentication, authorization, availability, confidentiality, identity management, integrity, audit, security monitoring, incident response, and security policy management.

While these security requirements are not new, we discuss cloud specific perspectives to help discuss, analyse and implement security in a cloud system.

- **Cloud Service Model Perspectives**
- **Implications of Cloud Deployment Models**
- **Shared Security Responsibilities**

E. Privacy

Cloud providers should protect the assured, proper, and consistent collection, processing, communication, use and disposition of personal information (PI) and personally identifiable information (PII) in the cloud.

- PII is the information that can be used to distinguish or trace an individual's identity, such as their name, social security number, biometric records, etc. alone, or when combined with other personal or identifying information that is linked or linkable to a specific individual, such as date and place of birth, mother's maiden name, etc.
- Though cloud computing provides a flexible solution for shared resources, software and information, it also poses additional privacy challenges to consumers using the clouds.