



## UNIT 1 Introduction of Cloud Computing

### Content :

- **Introduction and Overview of Cloud Computing**
- **Evolution of Cloud Computing**
- **Underlying principle of Parallel and Distributed Computing**
- **History of Cloud Computing**
- **Cloud Architecture**
- **Types of Cloud**
- **Major Player of Cloud computing**
- **Issues in Cloud**
- **Eucalyptus**
- **Nimbus**
- **Open Nebula**
- **Cloud Sim**

## 1.1 Introduction of Cloud Computing

### What is Cloud Computing

The term cloud refers to a network or the internet. It is a technology that uses remote servers on the internet to store, manage, and access data online rather than local drives. The data can be anything such as files, images, documents, audio, video, and more.

There are the following operations that we can do using cloud computing:

- Developing new applications and services
- Storage, back up, and recovery of data
- Hosting blogs and websites
- Delivery of software on demand



- Analysis of data
- Streaming videos and audios

## Why Cloud Computing?

Small as well as large IT companies, follow the traditional methods to provide the IT infrastructure. That means for any IT company, we need a Server Room that is the basic need of IT companies.

In that server room, there should be a database server, mail server, networking, firewalls, routers, modem, switches, QPS (Query Per Second means how much queries or load will be handled by the server), configurable system, high net speed, and the maintenance engineers.

To establish such IT infrastructure, we need to spend lots of money. To overcome all these problems and to reduce the IT infrastructure cost, Cloud Computing comes into existence.

### Characteristics of Cloud Computing

The characteristics of cloud computing are given below:

#### 1) **Agility**

The cloud works in a distributed computing environment. It shares resources among users and works very fast.

#### 2) **High availability and reliability**

The availability of servers is high and more reliable because the chances of infrastructure failure are minimum.

#### ) **High Scalability**

Cloud offers "**on-demand**" provisioning of resources on a large scale, without having engineers for peak loads.

#### 4) **Multi-Sharing**

With the help of cloud computing, **multiple users and applications can work more efficiently** with cost reductions by sharing common infrastructure.

#### 5) **Device and Location Independence**

Cloud computing enables the users to access systems using a web browser regardless of their location or what device they use e.g. PC, mobile phone, etc. **As infrastructure is off-site** (typically provided by a third-party) **and accessed via the Internet, users can connect from anywhere.**



### 6) Maintenance

Maintenance of cloud computing applications is easier, since they **do not need to be installed on each user's computer and can be accessed from different places**. So, it reduces the cost also.

### 7) Low Cost

By using cloud computing, the cost will be reduced because to take the services of cloud computing, IT company need not to set its own infrastructure and pay-as-per usage of resources.

### 8) Services in the pay-per-use mode

Application Programming Interfaces (APIs) are provided to the users so that they can access services on the cloud by using these APIs and pay the charges as per the usage of services.

## 1.3 History of Cloud Computing

Before emerging the cloud computing, there was Client/Server computing which is basically a centralized storage in which all the software applications, all the data and all the controls are resided on the server side.

If a single user wants to access specific data or run a program, he/she need to connect to the server and then gain appropriate access, and then he/she can do his/her business.

Then after, distributed computing came into picture, where all the computers are networked together and share their resources when needed.

On the basis of above computing, there was emerged of cloud computing concepts that later implemented.

At around in 1961, John MacCharty suggested in a speech at MIT that computing can be sold like a utility, just like a water or electricity. It was a brilliant idea, but like all brilliant ideas, it was ahead if its time, as for the next few decades, despite interest in the model, the technology simply was not ready for it.

But of course time has passed and the technology caught that idea and after few years we mentioned that:

In 1999, Salesforce.com started delivering of applications to users using a simple website. The applications were delivered to enterprises over the Internet, and this way the dream of computing sold as utility were true.



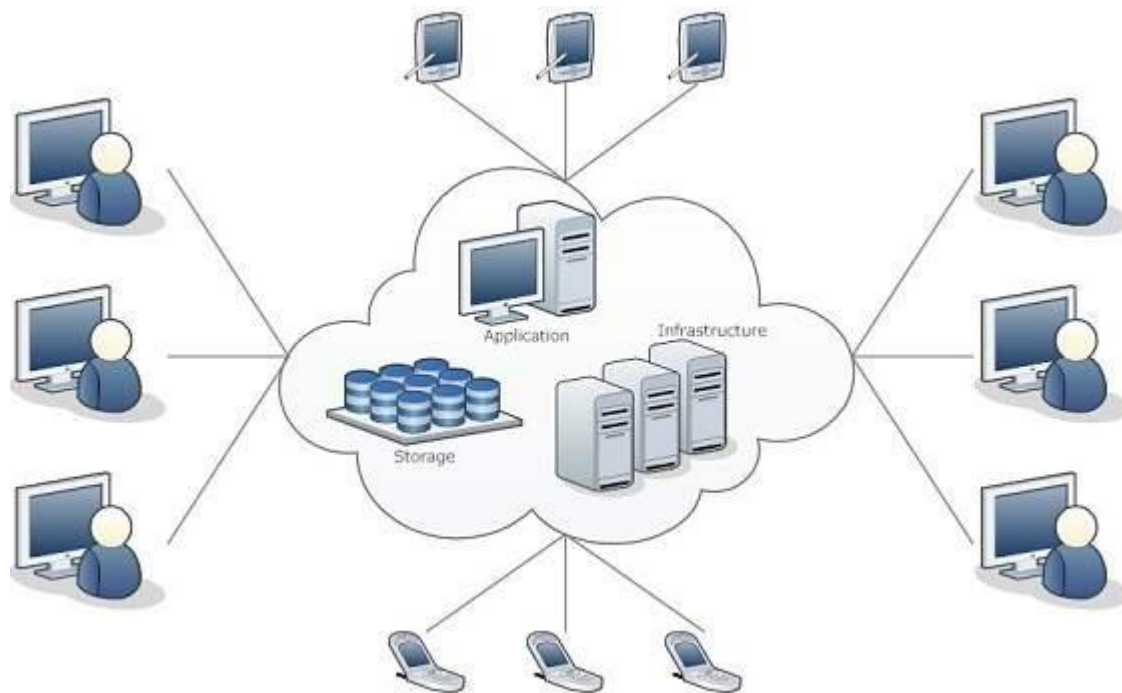
In 2002, Amazon started Amazon Web Services, providing services like storage, computation and even human intelligence. However, only starting with the launch of the Elastic Compute Cloud in 2006 a truly commercial service open to everybody existed.

In 2009, Google Apps also started to provide cloud computing enterprise applications.

Of course, all the big players are present in the cloud computing evolution, some were earlier, some were later. In 2009, Microsoft launched Windows Azure, and companies like Oracle and HP have all joined the game. This proves that today, cloud computing has become mainstream.

## 1.4 What is Cloud Computing?

Cloud Computing refers to manipulating, configuring, and accessing the hardware and software resources remotely. It offers online data storage, infrastructure, and application.



Cloud computing offers platform independency, as the software is not required to be installed locally on the PC. Hence, the Cloud Computing is making our business applications mobile and collaborative.

### Basic Concepts

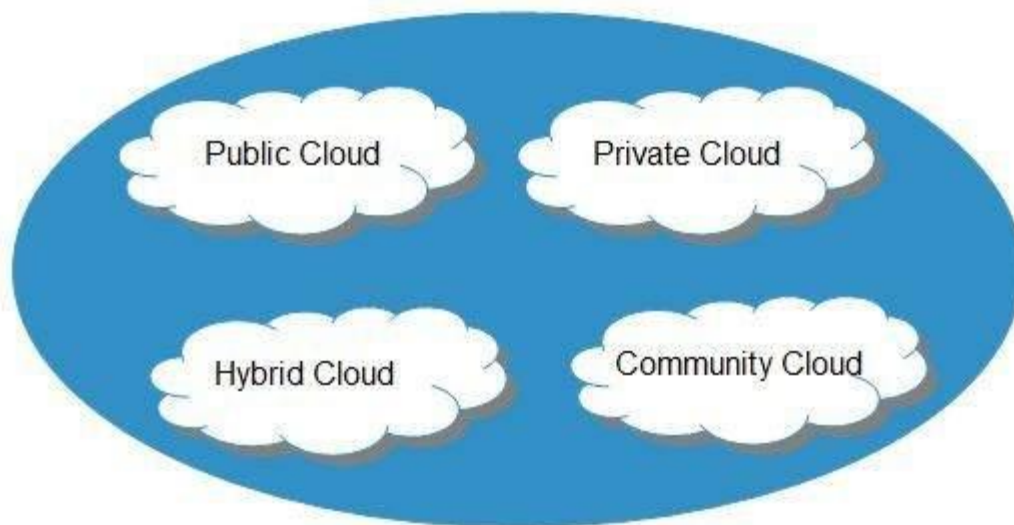


There are certain services and models working behind the scene making the cloud computing feasible and accessible to end users. Following are the working models for cloud computing:

- Deployment Models
- Service Models

#### Deployment Models

Deployment models define the type of access to the cloud, i.e., how the cloud is located? Cloud can have any of the four types of access: Public, Private, Hybrid, and Community.



#### **Public Cloud**

The public cloud allows systems and services to be easily accessible to the general public. Public cloud may be less secure because of its openness.

#### **Private Cloud**

The private cloud allows systems and services to be accessible within an organization. It is more secured because of its private nature.

#### **Community Cloud**

The community cloud allows systems and services to be accessible by a group of organizations.

#### **Hybrid Cloud**

The hybrid cloud is a mixture of public and private cloud, in which the critical activities are performed using private cloud while the non-critical activities are performed using public cloud.



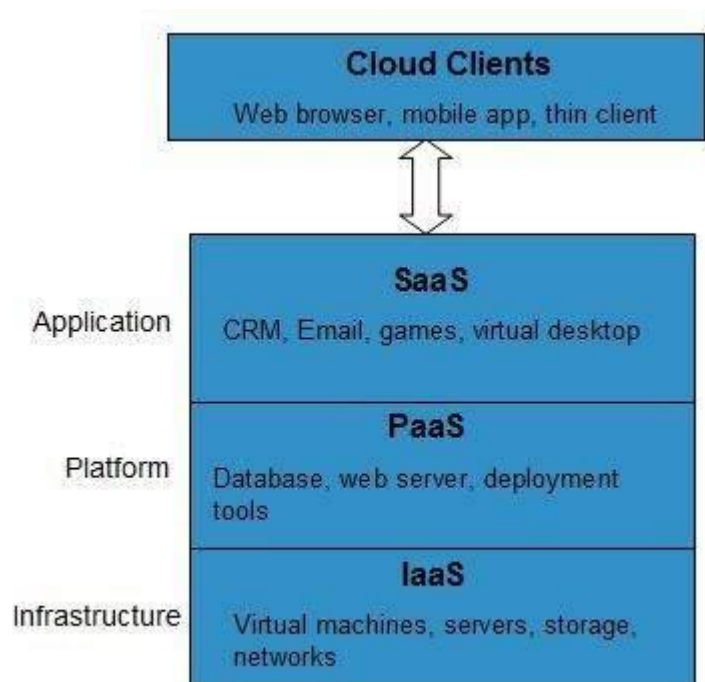
## 1.5 Service Models

Cloud computing is based on service models. These are categorized into three basic service models which are -

- Infrastructure-as-a-Service (IaaS)
- Platform-as-a-Service (PaaS)
- Software-as-a-Service (SaaS)

Anything-as-a-Service (XaaS) is yet another service model, which includes Network-as-a-Service, Business-as-a-Service, Identity-as-a-Service, Database-as-a-Service or Strategy-as-a-Service.

The Infrastructure-as-a-Service (IaaS) is the most basic level of service. Each of the service models inherit the security and management mechanism from the underlying model, as shown in the following diagram:



### Infrastructure-as-a-Service (IaaS)

IaaS provides access to fundamental resources such as physical machines, virtual machines, virtual storage, etc.

### Platform-as-a-Service (PaaS)



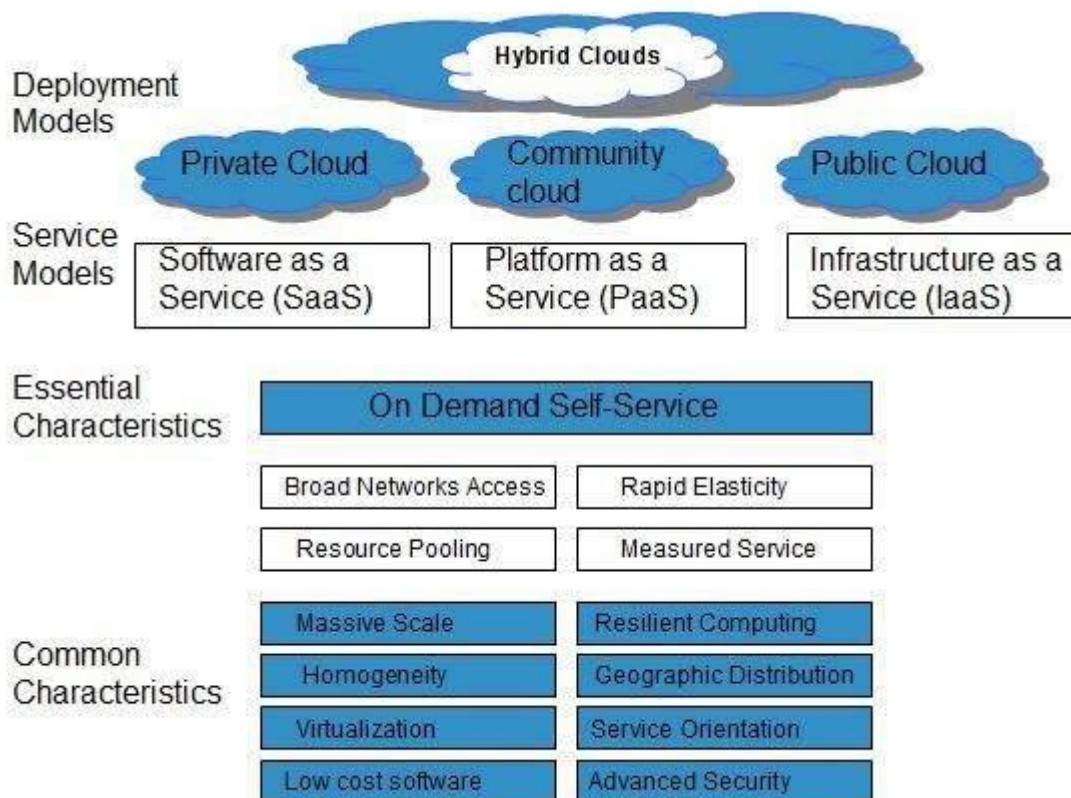
PaaS provides the runtime environment for applications, development and deployment tools, etc.

## Software-as-a-Service (SaaS)

SaaS model allows to use software applications as a service to end-users.

## 1.6 Characteristics of Cloud Computing

There are four key characteristics of cloud computing. They are shown in the following diagram:



### On Demand Self Service

Cloud Computing allows the users to use web services and resources on demand. One can logon to a website at any time and use them.

### Broad Network Access

Since cloud computing is completely web based, it can be accessed from anywhere and at any time.



## Resource Pooling

Cloud computing allows multiple tenants to share a pool of resources. One can share single physical instance of hardware, database and basic infrastructure.

There are basically 5 essential characteristics

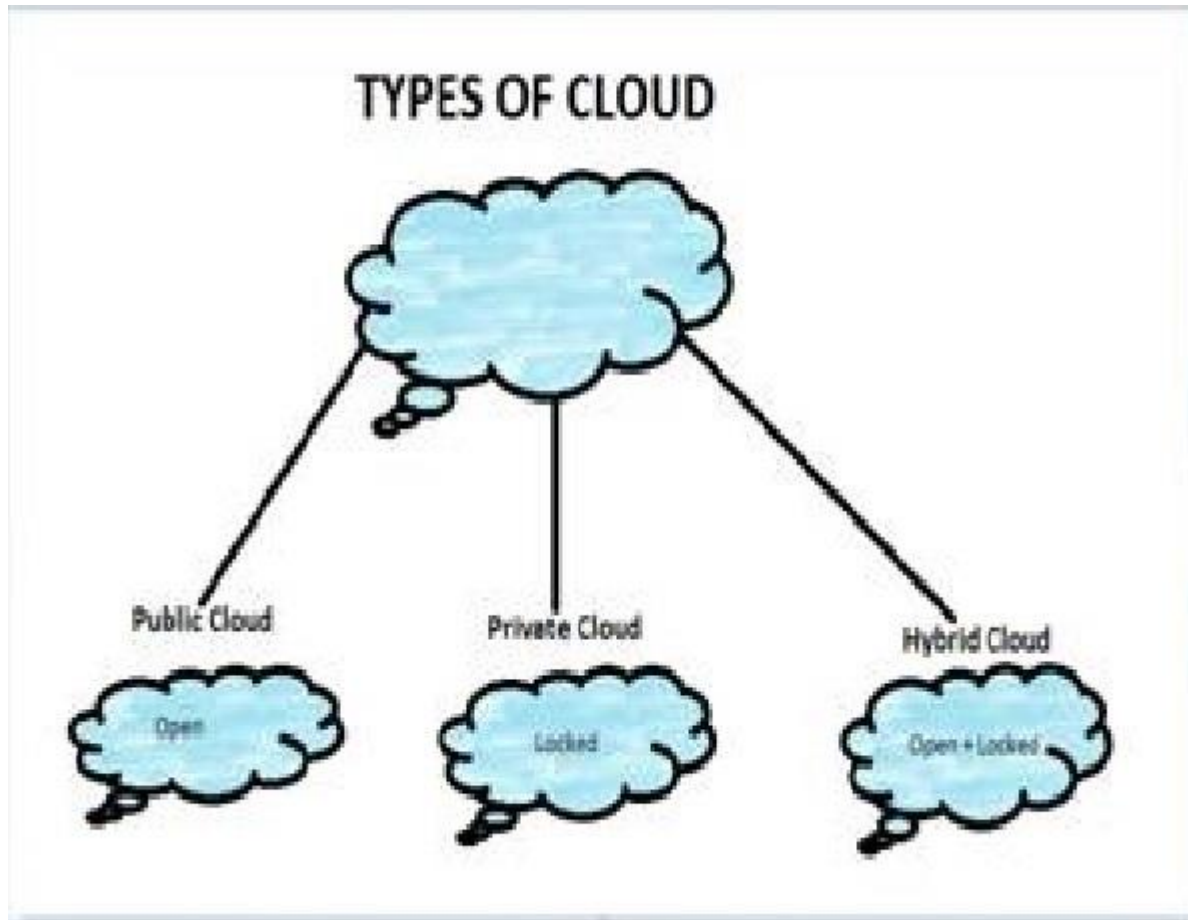
1. **On-demand self-services:** The Cloud computing services does not require any human administrators, user themselves are able to provision, monitor and manage computing resources as needed.
2. **Broad network access:** The Computing services are generally provided over standard networks and heterogeneous devices.
3. **Rapid elasticity:** The Computing services should have IT resources that are able to scale out and in quickly and on as needed basis. Whenever the user require services it is provided to him and it is scale out as soon as its requirement gets over.
4. **Resource pooling:** The IT resource (e.g., networks, servers, storage, applications, and services) present are shared across multiple applications and occupant in an uncommitted manner. Multiple clients are provided service from a same physical resource.
5. **Measured service:** The resource utilization is tracked for each application and occupant, it will provide both the user and the resource provider with an account of what has been used. This is done for various reasons like monitoring billing and effective use of resource.
6. **Multi-tenancy:** Cloud computing providers can support multiple tenants (users or organizations) on a single set of shared resources.
7. **Virtualization:** Cloud computing providers use virtualization technology to abstract underlying hardware resources and present them as logical resources to users.

## 1.8 Types of Cloud

The storage options on cloud is in 3 forms –

- Public
- Private
- Hybrid





**Public Cloud** – A service provider makes the clouds available to the general public which is termed as a public cloud. These clouds are accessed through internet by users. These are open to public and their infrastructure is owned and operated by service providers as in case of Google and Microsoft.

**Private Cloud** – These clouds are dedicated to a particular organization. That particular organization can use the cloud for storing the company's data, hosting business application, etc. The data stored on private cloud can't be shared with other organizations. The cloud is managed either by the organization itself or by the third party.

**Hybrid Cloud** – When two or more clouds are bound together to offer the advantage of both public and private clouds, they are termed as Hybrid Cloud. Organizations can use private clouds for sensitive application, while public clouds for non-sensitive applications. The hybrid clouds provide flexible, scalable and cost-effective solutions to the organizations.

## 1.9 Benefits of Cloud

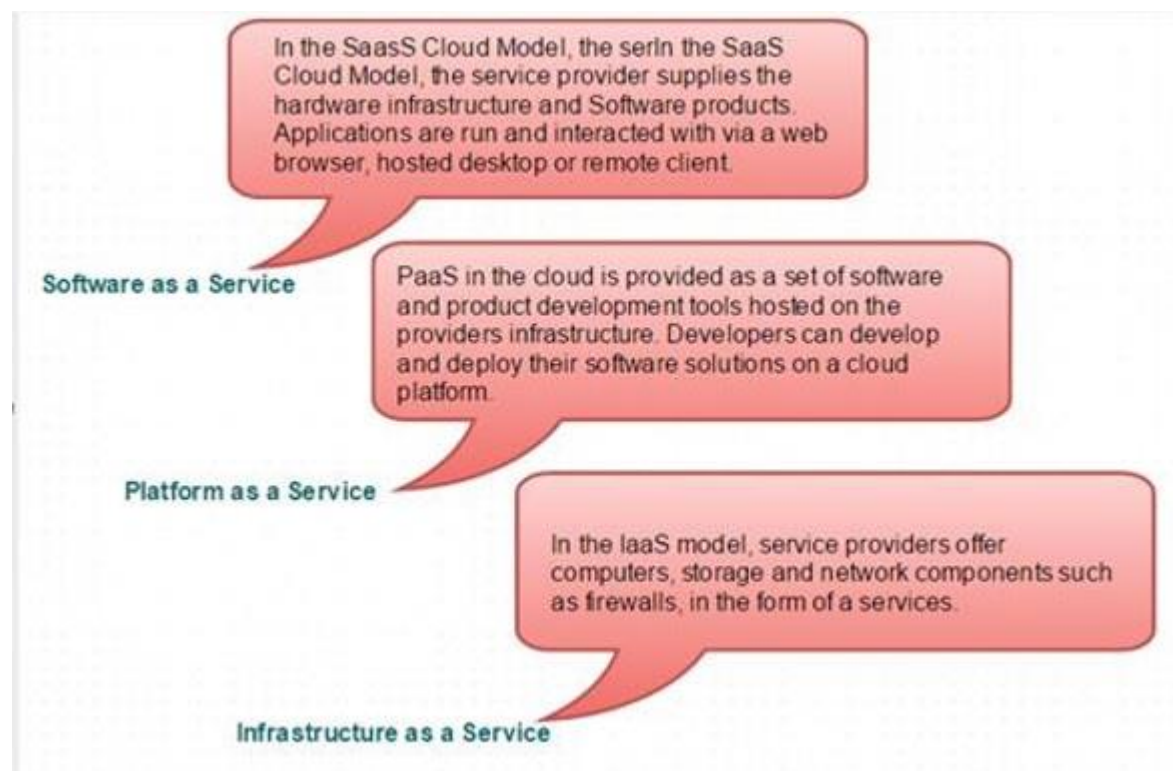


There are many benefits of clouds. Some of them are listed below.

- Cloud service offers scalability. Allocation and de-allocation of resources is dynamically as per demand.
- It saves on cost by reducing capital infrastructure.
- It allows the user to access the application independent of their location and hardware configuration.
- It simplifies the network and lets the client access the application without buying license for individual machine.
- Storing data on clouds is more reliable as it is not lost easily.

## SPI

Next comes how cloud services are categorized. S stand for Software, P stands for Platform and I for Infrastructure in SPI. SaaS is Software as a service; PaaS is Platform as a service and IaaS is Infrastructure as a Service.





Following are the live examples of these models.

- **SAAS Model** – E-mail (Gmail, Yahoo, etc.)
- **PAAS Model** – Microsoft Azure
- **IAAS Model** – Amazon S3

## 1.10 Major Cloud Service provider

There are the following Cloud Service Providers Companies -

### Amazon Web Services (AWS)

**AWS** (Amazon Web Services) is a **secure cloud service platform** provided by **Amazon**. It offers various services such as database storage, computing power, content delivery, Relational Database, Simple Email, Simple Queue, and other functionality to increase the organization's growth.



- AWS is scalable because it has an ability to scale the computing resources up or down according to the organization's demand.
- AWS is cost-effective as it works on a pay-as-you-go pricing model.
- It provides various flexible storage options.
- It offers various security services such as infrastructure security, data encryption, monitoring & logging, identity & access control, penetration testing, and DDoS attacks.
- It can efficiently manage and secure Windows workloads.



○

## ○ 2. Microsoft Azure

- Microsoft Azure is also known as Windows Azure. It supports various operating systems, databases, programming languages, frameworks that allow IT professionals to easily build, deploy, and manage applications through a worldwide network. It also allows users to create different groups for related utilities.



**Microsoft Azure**

### Features of Microsoft Azure

- Microsoft Azure provides scalable, flexible, and cost-effective
- It allows developers to quickly manage applications and websites.
- It managed each resource individually.
- Its IaaS infrastructure allows us to launch a general-purpose virtual machine in different platforms such as Windows and Linux.
- It offers a Content Delivery System (CDS) for delivering the Images, videos, audios, and applications.

## ○ 3. Google Cloud Platform

- Google cloud platform is a product of **Google**. It consists of a set of physical devices, such as computers, hard disk drives, and virtual machines. It also helps organizations to simplify the migration process.



## Google Cloud Platform

### Features of Google Cloud

- Google cloud includes various big data services such as Google BigQuery, Google CloudDataproc, Google CloudDatalab, and Google Cloud Pub/Sub.
- It provides various services related to networking, including Google Virtual Private Cloud (VPC), Content Delivery Network, Google Cloud Load Balancing, Google Cloud Interconnect, and Google Cloud DNS.
- It offers various scalable and high-performance
- GCP provides various serverless services such as Messaging, Data Warehouse, Database, Compute, Storage, Data Processing, and Machine learning (ML)
- It provides a free cloud shell environment with Boost Mode.
- 

### 4. IBM Cloud Services

- 

IBM Cloud is an open-source, faster, and more reliable platform. It is built with a suite of advanced data and AI tools. It offers various services such as Infrastructure as a service, Software as a service, and platform as a service. You can access its services like compute power, cloud data & Analytics, cloud use cases, and storage networking using internet connection



**IBM Cloud Services**

### Feature of IBM Cloud

- IBM cloud improves operational efficiency.
- Its speed and agility improve the customer's satisfaction.
- It offers Infrastructure as a Service (IaaS), Platform as a Service (PaaS), as well as Software as a Service (SaaS)
- It offers various cloud communications services to our IT environment.
- 
- 5. VMware Cloud
- 
- VMware cloud is a Software-Defined Data Center (SSDC) unified platform for the Hybrid Cloud. It allows cloud providers to build agile, flexible, efficient, and robust cloud services.



**VMware**

### Features of VMware



- VMware cloud works on the pay-as-per-use model and monthly subscription
- It provides better customer satisfaction by protecting the user's data.
- It can easily create a new VMware Software-Defined Data Center (SDDC) cluster on AWS cloud by utilizing a RESTful API.
- It provides flexible storage options. We can manage our application storage on a per-application basis.
- It provides a dedicated high-performance network for managing the application traffic and also supports multicast networking.
- It eliminates the time and cost complexity.

## 6. Oracle cloud

Oracle cloud platform is offered by the Oracle Corporation. It combines Platform as a Service, Infrastructure as a Service, Software as a Service, and Data as a Service with cloud infrastructure. It is used to perform tasks such as moving applications to the cloud, managing development environment in the cloud, and optimize connection performance.



### Features of Oracle cloud

- Oracle cloud provides various tools for build, integrate, monitor, and secure the applications.
- Its infrastructure uses various languages including, Java, Ruby, PHP, Node.js.
- It integrates with Docker, VMware, and other DevOps tools.
- Oracle database not only provides unparalleled integration between IaaS, PaaS, and SaaS, but also integrates with the on-premises platform to improve operational efficiency.



- It maximizes the value of IT investments.
- It offers customizable Virtual Cloud Networks, firewalls, and IP addresses to securely support private networks.

## 7. Red Hat

Red Hat virtualization is an open standard and desktop virtualization platform produced by Red Hat. It is very popular for the Linux environment to provide various infrastructure solutions for virtualized servers as well as technical workstations. Most of the small and medium-sized organizations use Red Hat to run their organizations smoothly. It offers higher density, better performance, agility, and security to the resources. It also improves the organization's economy by providing cheaper and easier management capabilities.



**Red Hat**

### Features of Red Hat

- Red Hat provides secure, certified, and updated container images via the Red Hat Container catalog.
- Red Hat cloud includes OpenShift, which is an app development platform that allows developers to access, modernize, and deploy apps
- It supports up to 16 virtual machines, each having up to 256GB of RAM.
- It offers better reliability, availability, and serviceability.
- It provides flexible storage capabilities, including very large SAN-based storage, better management of memory allocations, high availability of LVMs, and support for particularly roll-back.





- In the Desktop environment, it includes features like New on-screen keyboard, GNOME software, which allows us to install applications, update application, as well as extended device support.

## 8. DigitalOcean

DigitalOcean is the unique cloud provider that offers computing services to the organization. It was founded in 2011 by Moisey Uretsky and Ben. It is one of the best cloud provider that allows us to manage and deploy web applications.



### Features of DigitalOcean

- It uses the KVM hypervisor to allocate physical resources to the virtual servers.
- It provides high-quality performance.
- It offers a digital community platform that helps to answer queries and holding feedbacks.
- It allows developers to use cloud servers to quickly create new virtual machines for their projects.
- It offers one-click apps for droplets. These apps include MySQL, Docker, MongoDB, Wordpress, PhpMyAdmin, LAMP stack, Ghost, and Machine Learning.
-



- **Rackspace**
- Rackspace offers cloud computing services such as hosting web applications, Cloud Backup, Cloud Block Storage, Databases, and Cloud Servers. The main aim to designing Rackspace is to easily manage private and public cloud deployments. Its data centers operating in the USA, UK, Hong Kong, and Australia.



### Features of Rackspace

- Rackspace provides various tools that help organizations to collaborate and communicate more efficiently.
- We can access files that are stored on the Rackspace cloud drive, anywhere, anytime using any device.
- It offers 6 globally data centers.
- It can manage both virtual servers and dedicated physical servers on the same network.
- It provides better performance at a lower cost.

### 10. Alibaba Cloud

Alibaba Cloud is used to develop data management and highly scalable cloud computing services. It offers various services, including Elastic Computing, Storage, Networking, Security, Database Services, Application Services, Media Services, Cloud Communication, and Internet of Things.



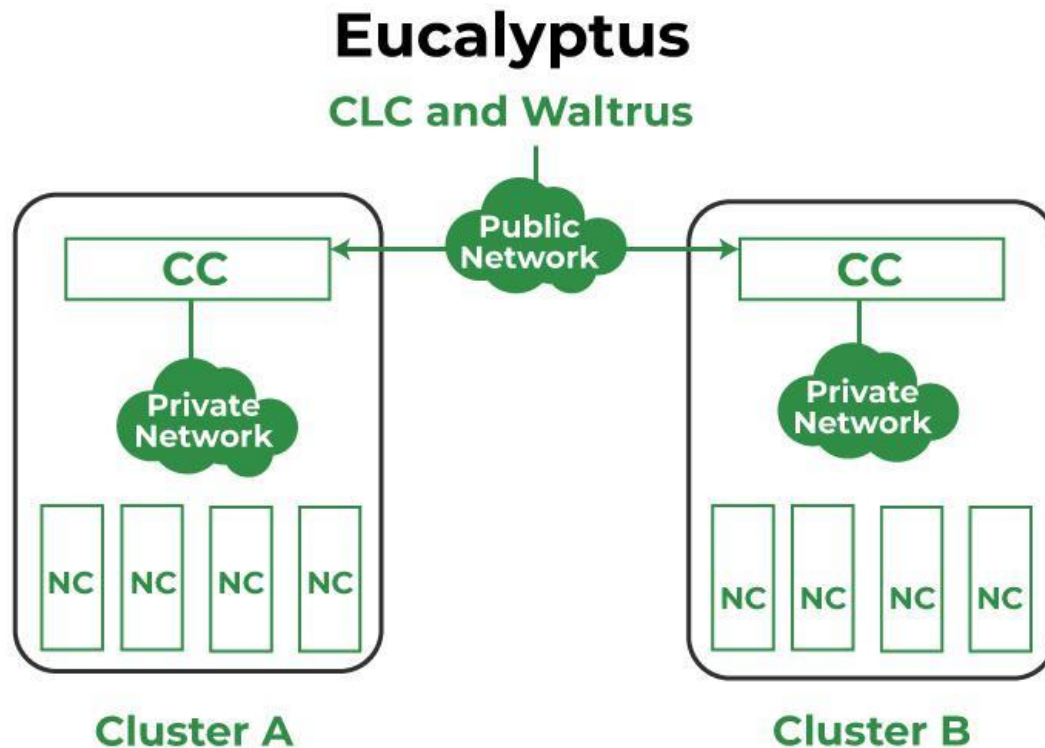
### Features of Alibaba Cloud

- Alibaba cloud offers a suite of global cloud computing services for both international customers and Alibaba Group's e-commerce ecosystem.
- Its services are available on a pay-as-per-use basis.
- It globally deals with its 14 data centers.
- It offers scalable and reliable data storage.

## 1.11 Eucalyptus

The open-source cloud refers to software or applications publicly available for the users in the cloud to set up for their own purpose or for their own purpose or for their organization.

Eucalyptus is a Linux-based open-source software architecture for cloud computing and also a storage platform that implements Infrastructure as a Service (IaaS). It provides quick and efficient computing services. Eucalyptus was designed to provide services compatible with Amazon's EC2 cloud and Simple Storage Service(S3).



### Eucalyptus Architecture

Eucalyptus CLIs can handle Amazon Web Services and their own private instances. Clients have the independence to transfer cases from Eucalyptus to Amazon Elastic Cloud. The virtualization layer oversees the Network, storage, and Computing. Occurrences are isolated by hardware virtualization.

#### Important Features are:-

1. Images: A good example is the Eucalyptus Machine Image which is a module software bundled and uploaded to the Cloud.
2. Instances: When we run the picture and utilize it, it turns into an instance.
3. Networking: It can be further subdivided into three modes: Static mode(allocates IP address to instances), System mode (assigns a MAC address and imputes the instance's network interface to the physical network via NC), and Managed mode (achieves local network of instances).
4. Access Control: It is utilized to give limitations to clients.
5. Elastic Block Storage: It gives block-level storage volumes to connect to an instance.
6. Auto-scaling and Load Adjusting: It is utilized to make or obliterate cases or administrations dependent on necessities.



### Components of Architecture

- Node Controller is the lifecycle of instances running on each node. Interacts with the operating system, hypervisor, and Cluster Controller. It controls the working of VM instances on the host machine.
- Cluster Controller manages one or more Node Controller and Cloud Controller simultaneously. It gathers information and schedules VM execution.
- Storage Controller (Walrus) Allows the creation of snapshots of volumes. Persistent block storage over VM instances. Walrus Storage Controller is a simple file storage system. It stores images and snapshots. Stores and serves files using S3(Simple Storage Service) APIs.
- Cloud Controller Front-end for the entire architecture. It acts as a Complaint Web Services to client tools on one side and interacts with the rest of the components on the other side.

### Operation Modes Of Eucalyptus

- Managed Mode: Numerous security groups to users as the network is large. Each security group is assigned a set or a subset of IP addresses. Ingress rules are applied through the security groups specified by the user. The network is isolated by VLAN between Cluster Controller and Node Controller. Assigns two IP addresses on each virtual machine.
- Managed (No VLAN) Node: The root user on the virtual machine can snoop into other virtual machines running on the same network layer. It does not provide VM network isolation.
- System Mode: Simplest of all modes, least number of features. A MAC address is assigned to a virtual machine instance and attached to Node Controller's bridge Ethernet device.
- Static Mode: Similar to system mode but has more control over the assignment of IP address. MAC address/IP address pair is mapped to static entry within the DHCP server. The next set of MAC/IP addresses is mapped.

### Advantages Of The Eucalyptus Cloud

1. Eucalyptus can be utilized to benefit both the eucalyptus private cloud and the eucalyptus public cloud.
2. Examples of Amazon or Eucalyptus machine pictures can be run on both clouds.
3. Its API is completely similar to all the Amazon Web Services.
4. Eucalyptus can be utilized with DevOps apparatuses like Chef and Puppet.
5. Although it isn't as popular yet but has the potential to be an alternative to OpenStack and CloudStack.
6. It is used to gather hybrid, public and private clouds.
7. It allows users to deliver their own data centers into a private cloud and hence, extend the services to other organizations.



## 1.12 Parallel computing

In parallel computing multiple processors performs multiple tasks assigned to them simultaneously. Memory in parallel systems can either be shared or distributed. Parallel computing provides concurrency and saves time and money.

## Distributed Computing

In distributed computing we have multiple autonomous computers which seems to the user as single system. In distributed systems there is no shared memory and computers communicate with each other through message passing. In distributed computing a single task is divided among different computers.

### Difference between Parallel Computing and Distributed Computing:

Parallel Computing	Distributed Computing
Many operations are performed simultaneously	System components are located at different locations
Single computer is required	Uses multiple computers
Multiple processors perform multiple operations	Multiple computers perform multiple operations
It may have shared or distributed memory	It have only distributed memory
Processors communicate with each other through bus	Computer communicate with each other through message passing.
Improves the system performance	Improves system scalability, fault tolerance and resource sharing capabilities

### Difference between Parallel Computing and Distributed Computing

There are mainly two computation types, including parallel computing and distributed computing. A computer system may perform tasks according to human instructions. A single processor executes only one task in the computer system, which is not an effective way. Parallel computing solves this problem by allowing numerous processors to accomplish tasks simultaneously. Modern computers support parallel processing to improve system



performance. In contrast, distributed computing enables several computers to communicate with one another and achieve a goal. All of these computers communicate and collaborate over the network. Distributed computing is commonly used by organizations such as Facebook and Google that allow people to share resources.

### **What is Parallel Computing?**

It is also known as parallel processing. It utilizes several processors. Each of the processors completes the tasks that have been allocated to them. In other words, parallel computing involves performing numerous tasks simultaneously. A shared memory or distributed memory system can be used to assist in parallel computing. All CPUs in shared memory systems share the memory. Memory is shared between the processors in distributed memory systems.

Parallel computing provides numerous advantages. Parallel computing helps to increase the CPU utilization and improve the performance because several processors work simultaneously. Moreover, the failure of one CPU has no impact on the other CPUs' functionality. Furthermore, if one processor needs instructions from another, the CPU might cause latency.

### **Advantages and Disadvantages of Parallel Computing**

There are various advantages and disadvantages of parallel computing. Some of the advantages and disadvantages are as follows:

#### **Advantages**

1. It saves time and money because many resources working together cut down on time and costs.
2. It may be difficult to resolve larger problems on Serial Computing.
3. You can do many things at once using many computing resources.
4. Parallel computing is much better than serial computing for modeling, simulating, and comprehending complicated real-world events.

#### **Disadvantages**

1. The multi-core architectures consume a lot of power.
2. Parallel solutions are more difficult to implement, debug, and prove right due to the complexity of communication and coordination, and they frequently perform worse than their serial equivalents.

### **What is Distributing Computing?**



It comprises several software components that reside on different systems but operate as a single system. A distributed system's computers can be physically close together and linked by a local network or geographically distant and linked by a wide area network (WAN). A distributed system can be made up of any number of different configurations, such as mainframes, PCs, workstations, and minicomputers. The main aim of distributed computing is to make a network work as a single computer.

There are various benefits of using distributed computing. It enables scalability and makes it simpler to share resources. It also aids in the efficiency of computation processes.

#### Advantages and Disadvantages of Distributed Computing

There are various advantages and disadvantages of distributed computing. Some of the advantages and disadvantages are as follows:

##### Advantages

1. It is flexible, making it simple to install, use, and debug new services.
2. In distributed computing, you may add multiple machines as required.
3. If the system crashes on one server, that doesn't affect other servers.
4. A distributed computer system may combine the computational capacity of several computers, making it faster than traditional systems.

##### Disadvantages

1. Data security and sharing are the main issues in distributed systems due to the features of open systems
2. Because of the distribution across multiple servers, troubleshooting and diagnostics are more challenging.
3. The main disadvantage of distributed computer systems is the lack of software support.

## 1.13 Open Nebula

OpenNebula is a simple, feature-rich and flexible solution for the management of virtualised data centres. It enables private, public and hybrid clouds. Here are a few facts about this solution.

OpenNebula is an open source cloud middleware solution that manages heterogeneous distributed data centre infrastructures. It is designed to be a simple but feature-rich, production-ready, customisable solution to build and manage enterprise clouds—simple to





install, update and operate by the administrators; and simple to use by end users. OpenNebula combines existing virtualisation technologies with advanced features for multi-tenancy, automated provisioning and elasticity. A built-in virtual network manager maps virtual networks to physical networks. Distributions such as Ubuntu and Red Hat Enterprise Linux have already integrated OpenNebula. As you'll learn in this article, you can set up OpenNebula by installing a few packages and performing some cursory configurations. OpenNebula supports Xen, KVM and VMware hypervisors.

### The OpenNebula deployment model

An OpenNebula deployment is modelled after the classic cluster architecture. Figure 1 shows the layout of the OpenNebula deployment model.

**Master node:** A single gateway or front-end machine, sometimes also called the master node, is responsible for queuing, scheduling and submitting jobs to the machines in the cluster. It runs several other OpenNebula services mentioned below:

- Provides an interface to the user to submit virtual machines and monitor their status.
- Manages and monitors all virtual machines running on different nodes in the cluster.
- It hosts the virtual machine repository and also runs a transfer service to manage the transfer of virtual machine images to the concerned worker nodes.
- Provides an easy-to-use mechanism to set up virtual networks in the cloud.
- Finally, the front-end allows you to add new machines to your cluster.

**Worker node:** The other machines in the cluster, known as 'worker nodes', provide raw computing power for processing the jobs submitted to the cluster. The worker nodes in an OpenNebula cluster are machines that deploy a virtualisation hypervisor, such as VMware, Xen or KVM.

### OpenNebula installation

OpenNebula is one of the easiest cloud systems to install and configure. Binary packages exist for several distributions, including Red Hat Enterprise Linux, Ubuntu, Fedora and openSUSE. Distributions like Ubuntu also include packages in their standard repositories, which makes it simple to install OpenNebula.

We will build a private cloud with three Ubuntu 10.04 machines. One machine will serve as the front-end, called the nebula-cloud-server. Two machines will be worker nodes deploying KVM. The worker nodes are open-nebula-wn and open-nebula-wn2.

To install the front-end, you need to install a single package by typing the following command:

```
sudo apt-get install opennebula
```

This package installs the daemon `oned` that manages all OpenNebula services. The installation also creates a new user called `oneadmin`, which OpenNebula retains for its own use. The installation of this package also produces a new pair of SSH RSA keys. OpenNebula uses SSH to communicate with other machines in the cluster, so a public key is used to



identify the front-end. All worker nodes need to authorise the front-end SSH key. After the opennebula package has been installed, the next step is to add the worker nodes, but before doing that, you need to install the worker node OpenNebula package on the worker node machines. At both worker nodes, you need to install the opennebula-node package by typing the following command:

```
sudo apt-get install opennebula-node
```

This package installs the necessary worker node packages and creates the oneadmin user. The following command is used to perform any operation related to the hosts:

```
onehost [options] command [parameters]
```

The command argument specifies the operation you want to carry out. You can add new machines to your cluster, retrieve monitoring information, enable or disable scheduling of virtual machines to the host, and list the currently active machines.

The onehost create command assists with adding new worker nodes to a cluster, as shown below:

```
onehost create hostname im_mad vmm_mad tm_mad
```

The hostname argument is the hostname or the IP of the machine; the next three parameters relate to the hypervisor and the transfer manager service. OpenNebula uses different drivers to access different hypervisors. These three parameters essentially tell the front-end that the machine you are adding deploys this specific hypervisor, and this mechanism should be used to transfer data to this machine. In this example, I add a host deploying KVM:

```
onehost add open-nebula-wn im_kvm vmm_kvm tm_ssh
```

The final arguments relate to the way virtual machines will be transferred to the node. In this example, SSH is used to transfer virtual machines. Note that the oneadmin user on this host needs to be able to SSH to oneadmin@open-nebula-wn; so, on this host, run the following command:

## 1.14 Nimbus – cloud computing for science

Nimbus is an open-source toolkit that, once installed on a cluster, provides an infrastructure as a service cloud to its client via WSRF-based or Amazon EC2 WSDL web service APIs. The mission of the software is to evolve the infrastructure with emphasis on the needs of science, but many non-scientific use cases are supported as well.

Nimbus supports the Xen hypervisor or KVM and virtual machine schedulers PBS and SGE. It allows deployment of self-configured virtual clusters via contextualization. It is configurable with respect to scheduling, networking leases, and usage accounting.

Nimbus clients can deploy, pause, restart and shutdown VMs.



**Features include:**

- Storage cloud implementation called Cumulus that has been tightly integrated with the other central services.
- Cumulus is compatible with the Amazon Web Services S3 REST API.
- Intuitive Cloud client – the workspace cloud client allows authorized clients to access many Workspace Service features in a user friendly way. cloud-client is an easy way to use both a storage cloud and IaaS.
- Provides an implementation of Amazon’s Elastic Compute Cloud (EC2) that allows you to use clients developed for the real EC2 system against Nimbus based clouds.
- WSRF based: protocol implementation.
- S3 REST API clients can also be used for managing VM storage with the Nimbus system.
- Allows clients to create auto-configuring clusters.
- Resource management options including schedulers such as PBS.
- LANTorrent – a file distribution protocol which act as a means to multi-cast virtual machine images to many backend nodes.
- Flexible group management.
- Per-client usage tracking.
- Per-user Storage Quota.
- Flexible request authentication and authorization.
- Simple user management.
- Configuration management.
- One-click clusters.
- Workspace client.
- VM network configuration.
- Local resource management plugin.
- Xen and KVM plugins.
- Site scheduler integration.
- VM resource usage enforcement (resource allocation).