Fundamentals of Cloud Computing and Big Data

Dr. Preeti Mishra (Ph. D (MNIT Jaipur)), Associate Professor
Dept. of Computer Science & Engineering
Graphic Era Deemed to be University Dehradun

https://scholar.google.com/citations?user=qRKv67 AAAAAJ&hl=en

Cloud Definitions

- ▶ Definition from **NIST** (National Institute of Standards and Technology)
 - Cloud computing is a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction.
 - This cloud model promotes availability and is composed of five essential characteristics, three service models, and four deployment models.



Different Perspectives on Cloud Computing

- From an **engineering perspective** the cloud is a computing architecture characterized by a large number of interconnected identical computing devices that can scale on demand and that communicate via an IP network.
- From a **business perspective** it is computing services that are scalable and billed on a usage basis.
- Allows customers to shift traditional Capital Expenditures (CapEx) into their Operating Expenditure (OpEx) budgets



Different Perspectives on Cloud Computing

Perspective from user :

- Users do not care about how the works are done
 - Instead, they only concern about what they can get
- Users do not care about what the provider actually did
 - Instead, they only concern about their quality of service
- Users do not want to own the physical infrastructure
 - Instead, they only want to pay as many as they used
- What dose user really care ?
 - They only care about their "Service"





Some Advantages

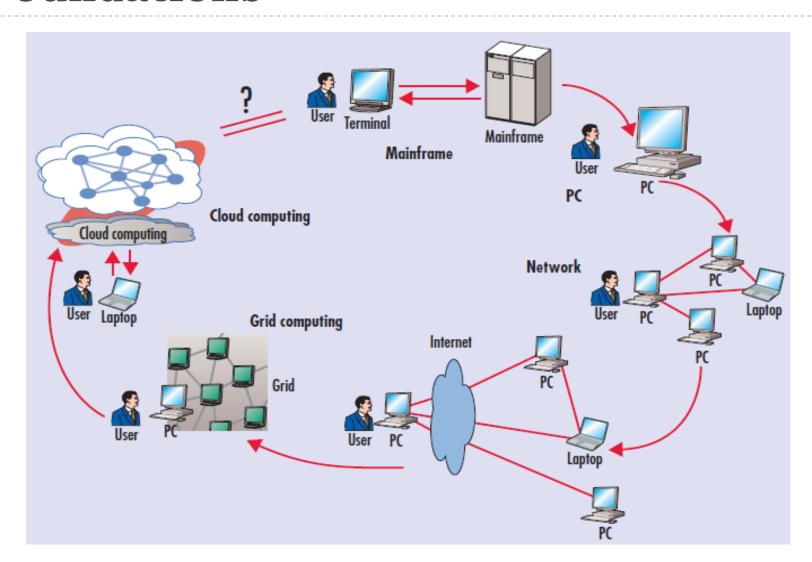
- No up-front investment
- Lowering operating cost
- Highly scalable
- Easy access
- Reducing business risks and maintenance expenses

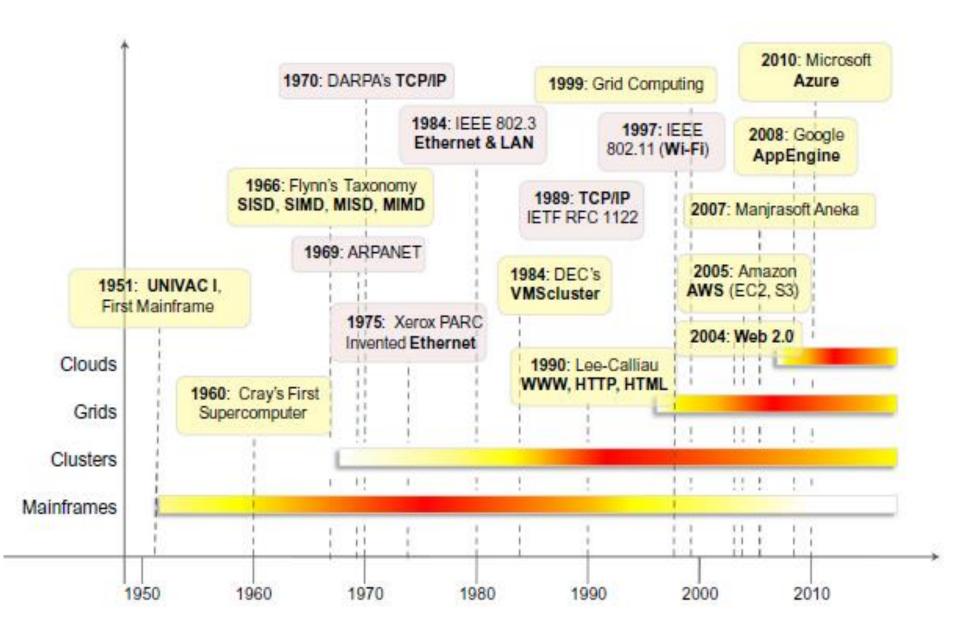
Some disadvantages

- No Physical Control on data
- Security threats
- Technical Issue
- May be not all features are supported
- Support Problem
- Network Connectivity required etc



Foundations





Foundations

Technologies	Description
Grid Computing	Form of distributed computing which applies the resources of numerous computers in a network to work on a single complex task at the same time
Utility Computing	Service provisioning model that offers computing resources as a metered service
Virtualization	 Provides improved utilization of resources Enables optimization of resources by over subscription
Service Oriented Architecture (SOA)	 An architectural approach in which applications make use of services available in the network Each service provides a specific function

A. Distributed Systems

- Cloud computing uses the concept of distributed computing.
- "A distributed system is a collection of independent computers that appears to its users as a single coherent system."
- ▶ The major milestones that led to cloud computing:
- (i) Mainframe (ii) Clusters (iii) Grids

Mainframes

- Mainframes are very powerful, highly reliable computers, specialized for large data movement and massive IO operations.
- Used by organization for bulk data processing.
- Based on batch-processing

Clusters

- Low cost alternative to the use of mainframes and supercomputers
- Cheap commodity machines can be connected by highbandwidth networks.
- Controlled by specific software tools that manage them a single system.
- By 1980s, Standard technology for parallel and high performance computing.

Grids

- Early 90s, as a evaluation of cluster computing.
- An aggregation of geographically dispersed clusters by means of internet connections.
- Clusters belongs to different organizations and arrangements are made between them to share computational power.
- Scale is nation wide or world wide.

Cloud is a successor of it.

Virtualization

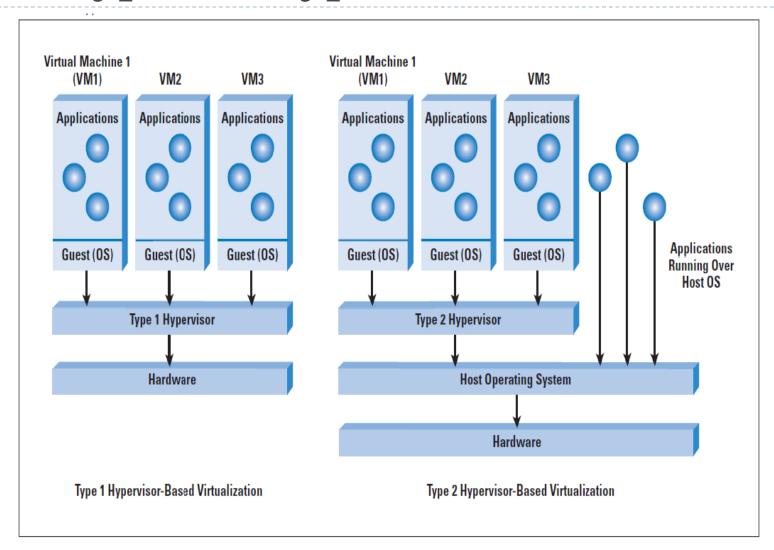
- Transforms from "one server- one application" to multiple virtual machines on each physical machine
- 2. Run Multiple OS on a single machine
- 3. A thin layer is introduced over either the hardware or on top of the OS
- ▶ Leaders in Virtualization VMware, Citrix Xen

Hypervisor

▶ Hypervisor or Virtual Machine Monitor (VMM) is a software that creates and runs VMs. It can access any of the VM spawned by it. It emulates the physical hardware and prevents direct access to physical hardware.

Multiple instances of a variety of OS'es can share a virtualized resource.

Two types of Hypervisors



Hypervisor



Hardware

Standard Computer
Hardware + OS





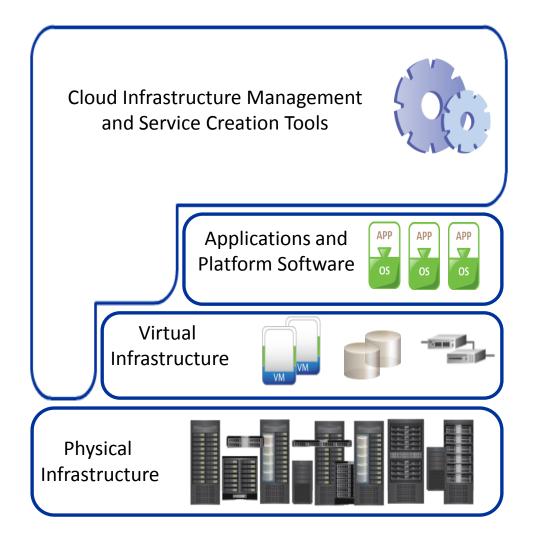
Hardware

Virt. Computer Hardware + Xen

Benefits of Hypervisor

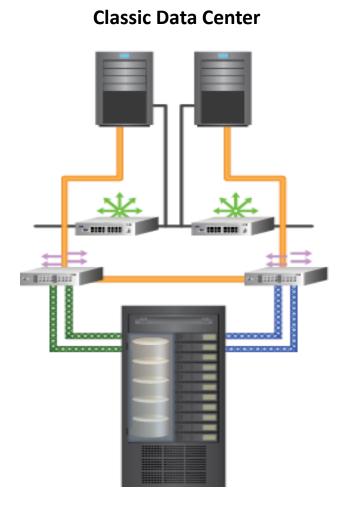
- Increase server utilization
- Consolidate server farms
- Decrease complexity
- Decrease Total Cost of Ownership (TCO)
 - Server consolidation refers to the use of a physical server to accommodate one or more server applications or user instances.
 - Sharing a server's compute resources among multiple applications and services simultaneously.
 - It is mainly used to reduce the number of server required in an organization.

Cloud Infrastructure



Classic Data Center

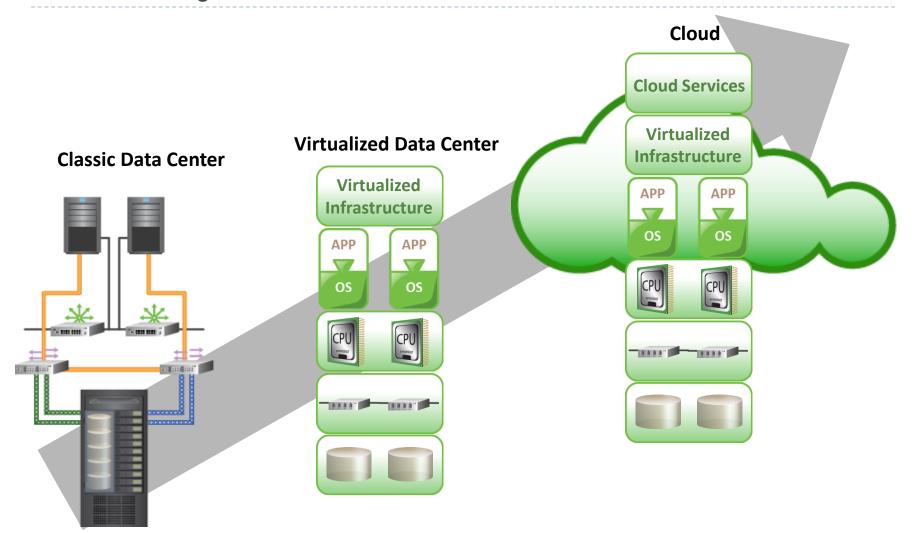
- Core elements of classic data center are:
 - Compute
 - Storage
 - Network
 - Operating System
 - **DBMS**
 - Application



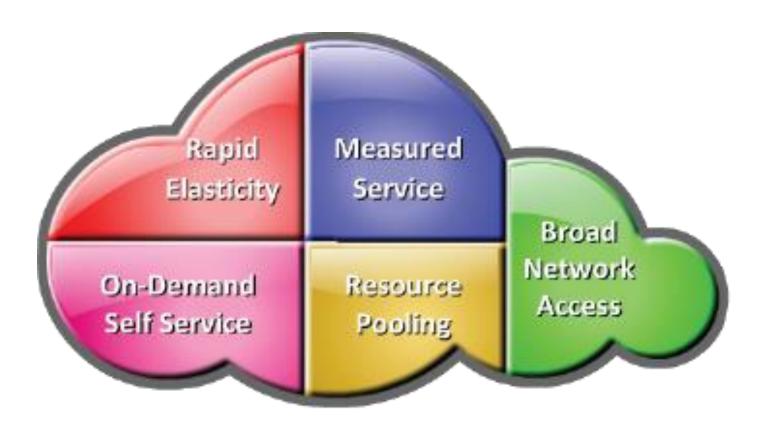
Virtualized Data Center

- Virtualization may be implemented at compute, storage, network, and/or application layers
- Virtualization Benefits:
 - Optimizes utilization of IT infrastructure
 - Reduces cost and management complexity
 - Reduces deployment time
 - Increases flexibility

Journey to the Cloud



Essential Characteristics as per NIST



On-Demand Self-Service

- Enables consumers to get computing resources as and when required, without any human intervention
- Facilitates consumer to leverage "ready to use" services or, enables to choose required services from the service catalog
- Allows provisioning of resources using self-service interface

Self-service interface should be user-friendly

On-Demand
Self Service

Broad Network Access

- Cloud services are accessed via the network, usually the internet, from a broad range of client platforms such as:
 - Desktop computer
 - Laptop
 - Mobile phone
 - Thin Client
- Eliminates the need for accessing a particular client platform to access the services
- Enables accessing the services from anywhere across the globe



Resource Pooling

- IT resources (compute, storage, network) are pooled to serve multiple consumers
 - Based on multi-tenant model
- Consumer has no knowledge about the exact location of the resources provided
- Resources are dynamically assigned and reassigned based on the consumer demand



Pooling

Rapid Elasticity

- Ability to scale IT resources rapidly, as required, to fulfill the changing needs without interruption of service
 - Resources can be both scaled up and scaled down dynamically
- ▶ To the consumer, the Cloud appears to be infinite
 - Consumers can start with minimal computing power and can expand their environment to any size

Rapid

Elasticity

Metered Service

- Consumers are billed based on the metered usage of Cloud resources
 - Cost incurred on a pay-per-use basis
 - Pricing/billing model is tied up with the required service levels
- Resource usage is monitored and reported
- Provides transparency for chargeback to both
 - Cloud service provider
 - Consumer about the utilized service



Why the Cloud

- Cloud Architectures address key difficulties surrounding large-scale data processing.
- In traditional data processing it is difficult to get as many machines as an application needs.
- Second, it is difficult to get the machines when one needs them.
- Third, it is difficult to distribute and coordinate a large-scale job on different machines, run processes on them, and provision another machine to recover if one machine fails.

Why the Cloud

- Fourth, it is difficult to auto scale up and down based on dynamic workloads.
- Fifth, it is difficult to get rid of all those machines when the job is done.

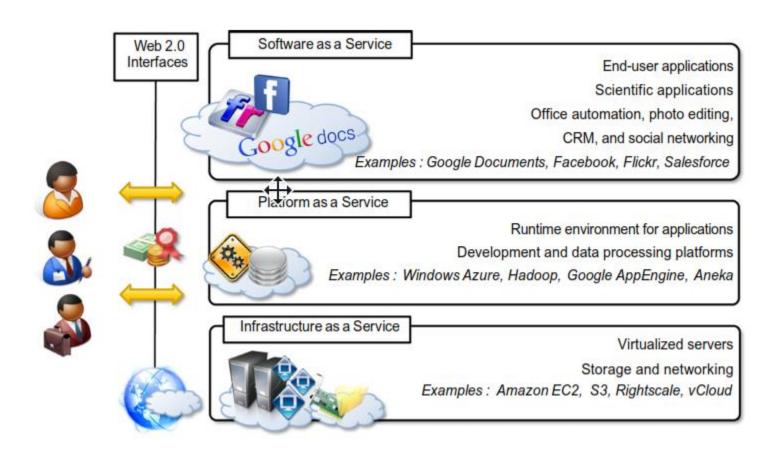
etc....

Cloud Architectures solve such difficulties!!!

Cloud Reference Architecture

SaaS, PaaS, IaaS

Cloud Reference Architecture*





Cloud Service Delivery Models

SaaS, PaaS, IaaS

Cloud Service Models

Cloud Service can be classified into three categories:

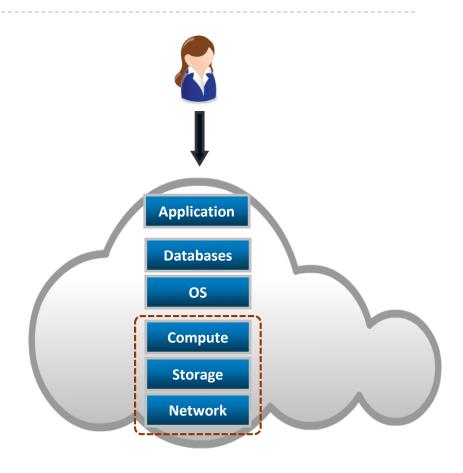
- Infrastructure-as-a-Service (laaS)
- Platform-as-a-Service (PaaS)
- Software-as-a-Service (SaaS)

Infrastructure as a service

Dr. Preeti Mishra

Infrastructure-as-a-Service

- Provides capability to the consumer to hire infrastructure components such as servers, storage, and network
- Enables consumers to deploy and run software, including OS and applications
- Pays for infrastructure components usage, for example, Storage capacity, CPU usage, etc.



IaaS Examples

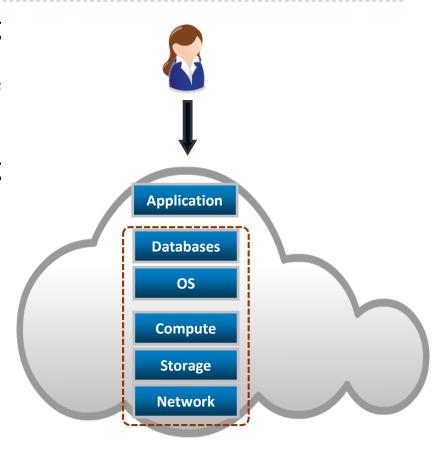
- Amazon Elastic Compute Cloud (EC2) is an laaS model that provides resizable compute capacity on a pay-per-use basis
 - Allows consumers to hire virtual compute on which they run their own applications
- ▶ EMC Atmos Online provides Storage as a service
 - Internet accessible, on demand storage
- ▶ Common Examplë: DigitalOcean, Linode, Rackspace, Amazon Web Services (AWS), Cisco Metapod, Microsoft Azure, Google Compute Engine (GCE)

Platform as a service

Dr. Preeti Mishra

Platform-as-a-Service

- It provides a <u>platform</u> allowing customers to develop, run, and manage applications without the complexity of building and maintaining the infrastructure typically associated with developing and launching an app
- Consumer has control over
 - Deployed applications
 - Possible application hosting environment configurations
- Consumer is billed for platform software components
 - OS, Database, Middleware ((e.g. <u>Java runtime</u>, <u>.NET</u> runtime, integration, etc.))



PaaS Examples

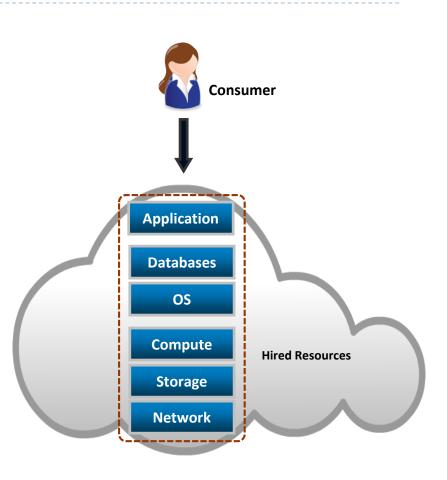
- Google App Engine provides platform for consumers to deploy or create their own applications
 - Allows dynamic allocation of system resources for an application based on the actual demand
 - Provides Java and Python environment to create and deploy application
- Microsoft Azure Platform provides diverse functionalities to build applications
 - Uses existing skills with Visual Studio and .Net to build applications
 - Builds applications also in Java and PHP using Eclipse and other tools
- Common Examples: AWS Elastic Beanstalk, Windows Azure,
 Heroku, Force.com, Google App Engine, Apache Stratos

Cloud Service Delivery Models

Software as a Service (SaaS)

Software-as-a-Service

- Capability provided to the consumer to use provider's applications running in a Cloud infrastructure
- Complete stack including application is provided as a service
- Application is accessible from various client devices, for example, via a thin client interface such as a Web browser
- Billing is based on the application usage



SaaS Examples

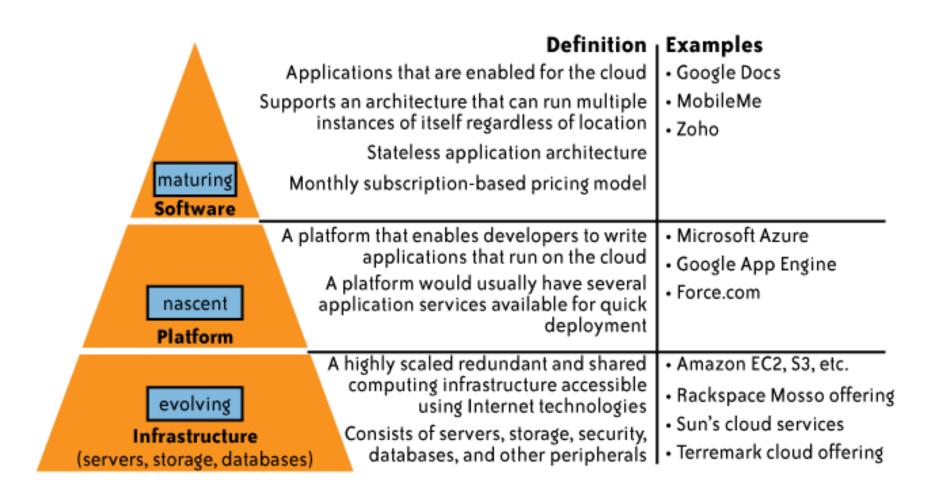
- EMC Mozy is a Software-as-a-Service solution for on-line backup
 - Consumers can leverage the Mozy console to perform automatic, secured, online backup and recovery of their data with ease
- Salesforce.com is a Software-as-a-Service solution for CRM application
 - Consumers can access CRM applications from anywhere, any time
- Common Examples: Google Apps, Microsoft office365, Google docs, Gmail

Everthing-As-A-Service (XaaS)

- Storage-as-a-service
- Database-as-a-service
- Information-as-a-service
- Process-as-a-service
- Application-as-a-service
- Platform-as-a-service
- Integration-as-a-service
- Security-as-a-service
- Management/Governance-as-a-service
- Testing-as-a-service
- Infrastructure-as-a-service



Taxonomy of Cloud Service Models



Amazon Web Services (AWS)

AWS stands for Amazon Web Service which is a collection of remote computing services also known as cloud computing. This technology of cloud computing is also known as laaS or Infrastructure as a Service.

Key Components of AWS

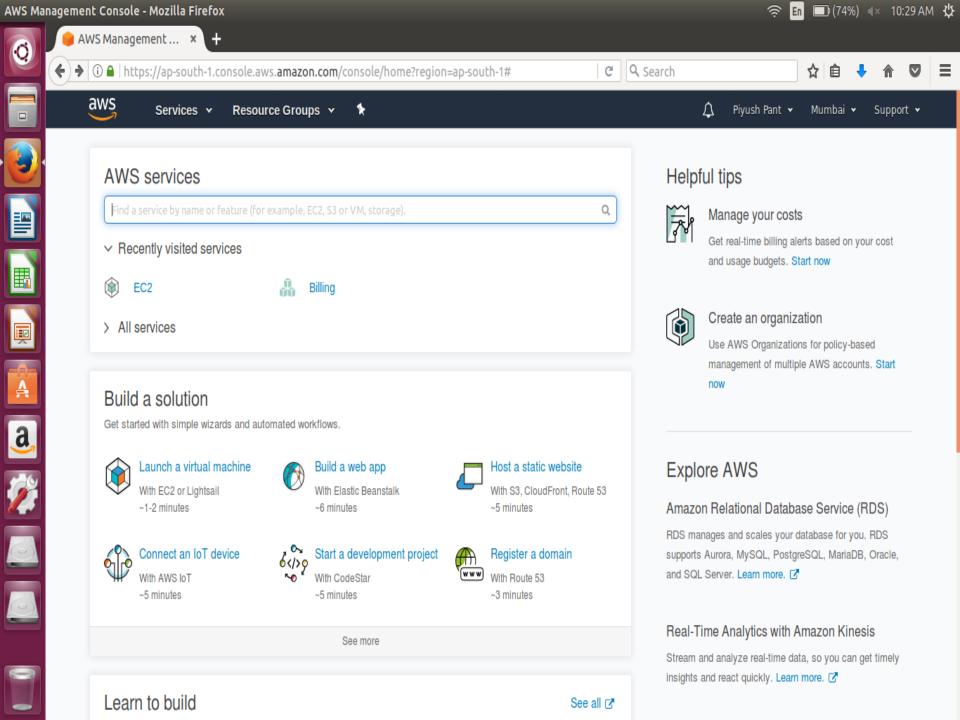
- ▶ The key components of AWS are as follows:
- ▶ Route 53: A DNS (Domain Name SERVER) web based service platform.
- Simple E-mail Service: Sending of E-mail is done by using RESTFUL API call or via regular SMTP (Simple Mail Transfer Protocol).
- Identity and Access Management: Improvised security and Identity management is provided for AWS account.
- ▶ Simple Storage Device or (S3): It is a huge storage medium, widely used for AWS services.

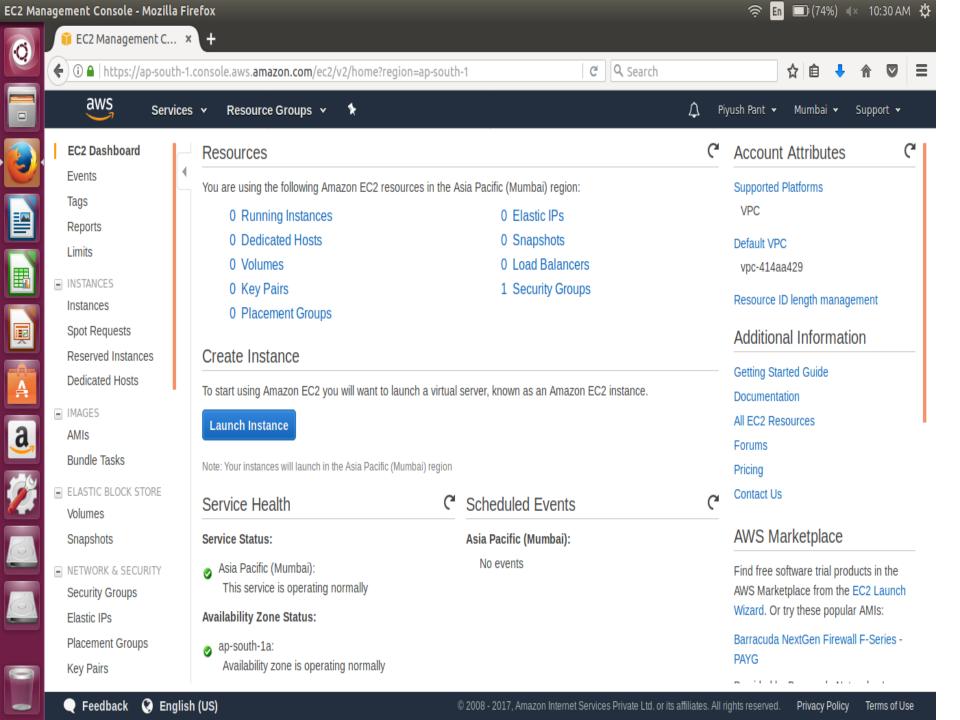
Key Components of AWS

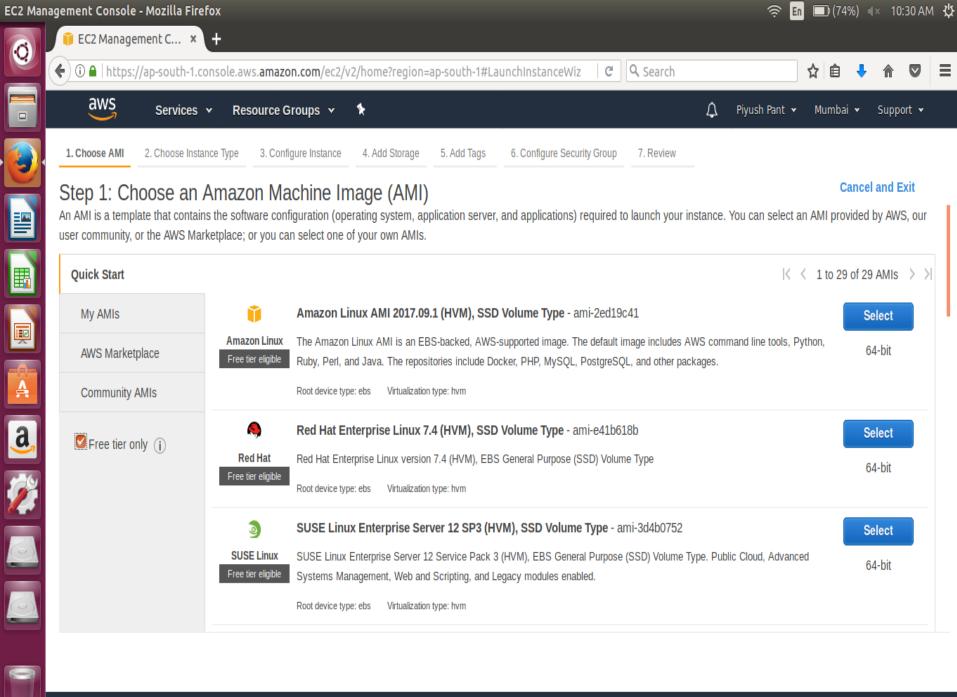
- Elastic Compute Cloud (EC2): Allows on-demand computing resources for hosting applications and essentially useful for unpredictable workloads
- ▶ Elastic Block Store (EBS): Storage volumes which is being attached to EC2 and allows the data lifespan of a single EC2
- Cloud Watch: It is used to monitor AWS resources and it allows administrators to view and collect keys required. Access is provided so that one can set a notification alarm in case of trouble.

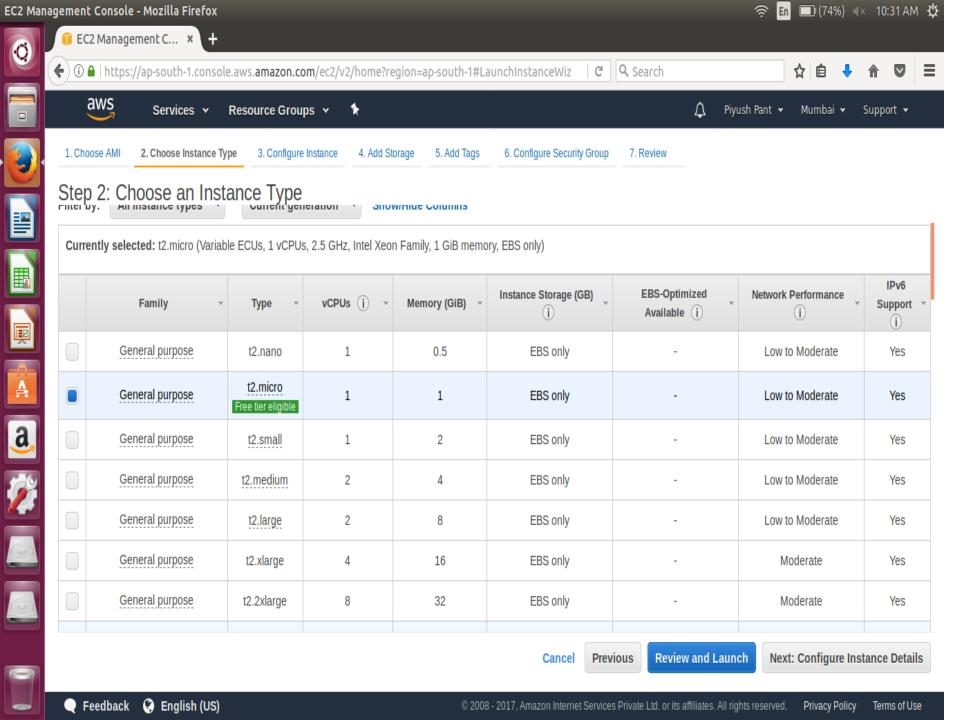
Hands-on: Demo of AWS

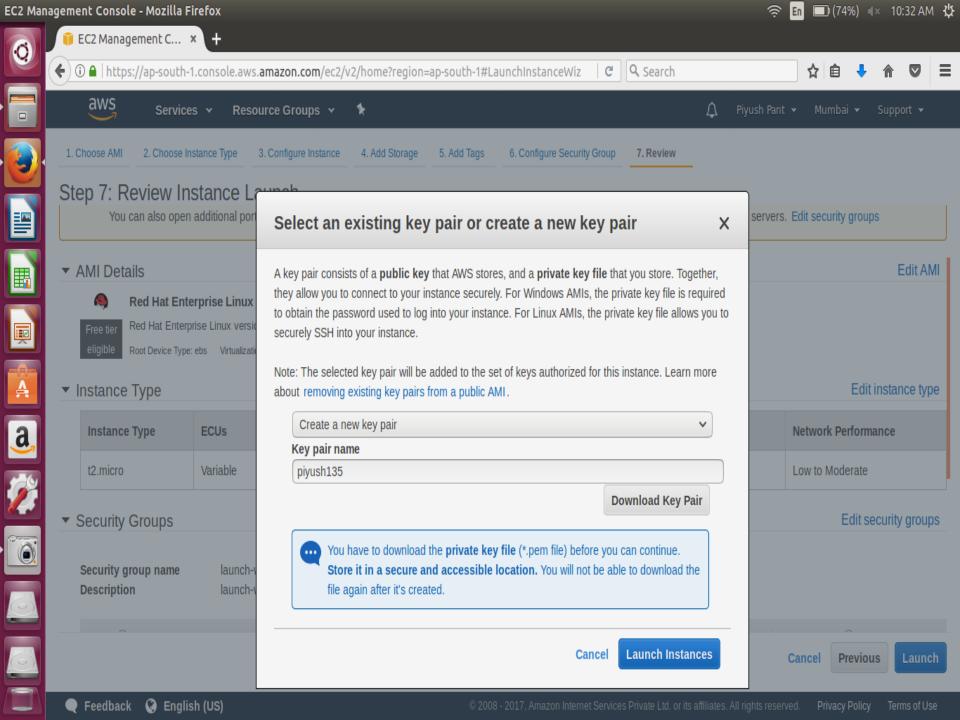
► AWS EC2 service to launch and access VM from client machine using ssh.

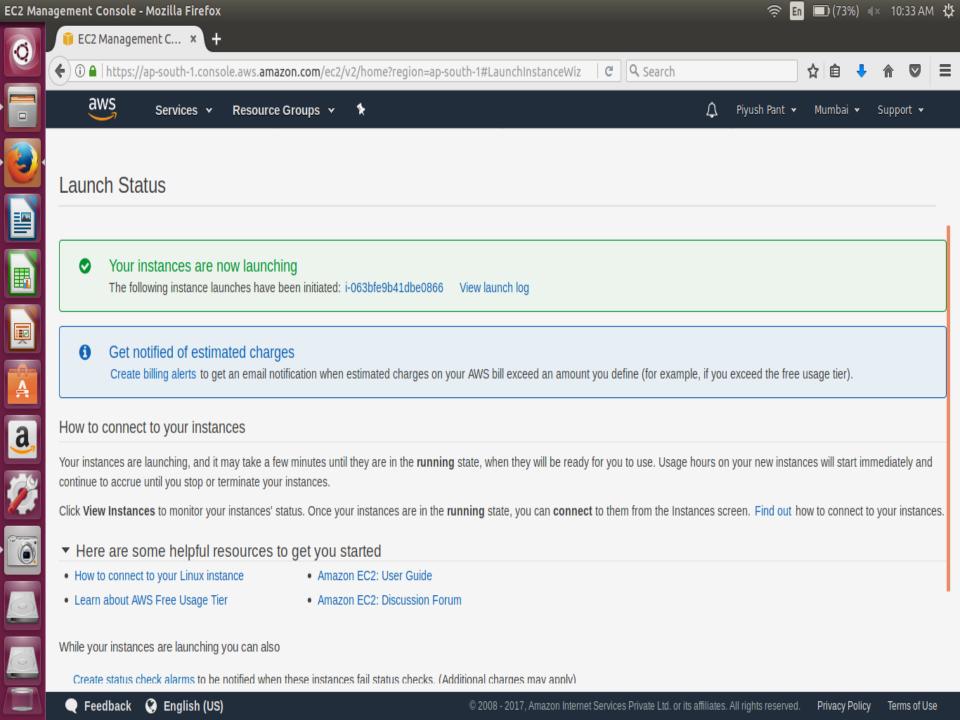


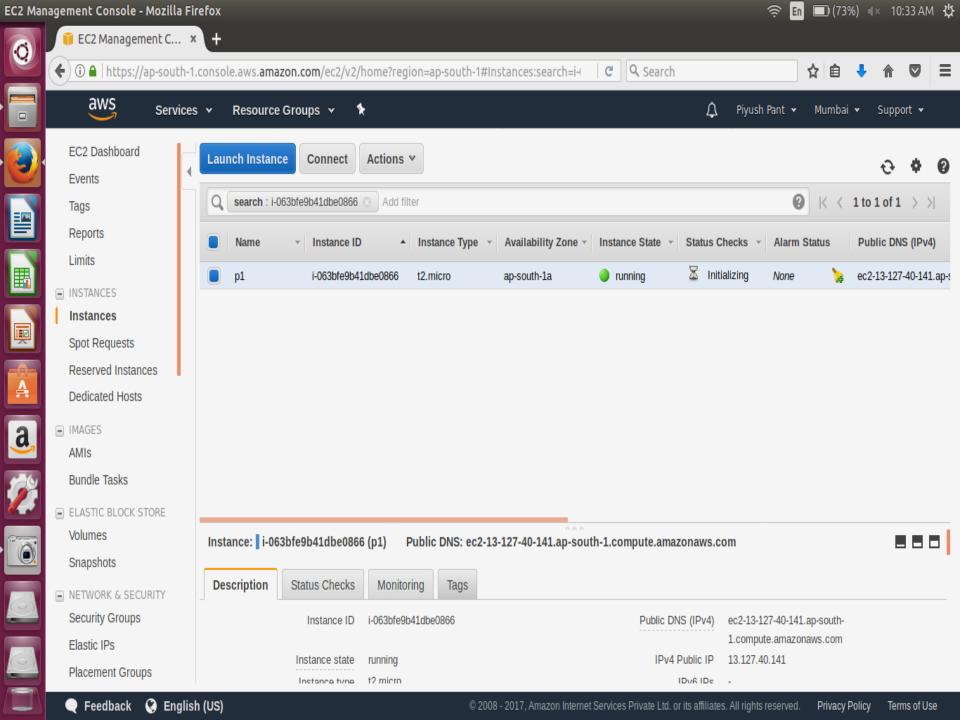


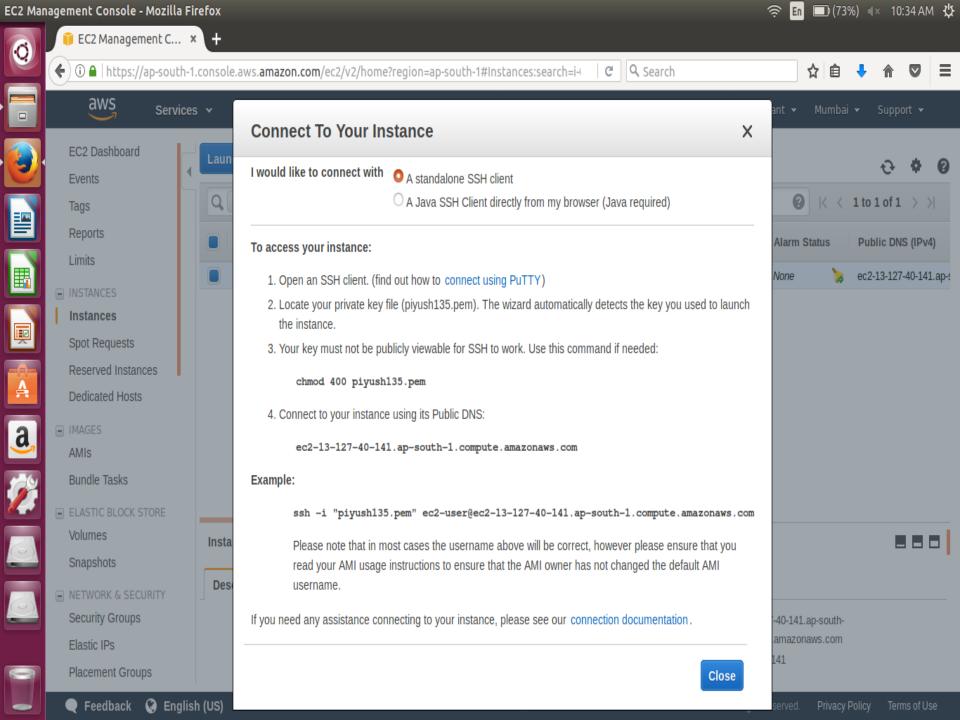


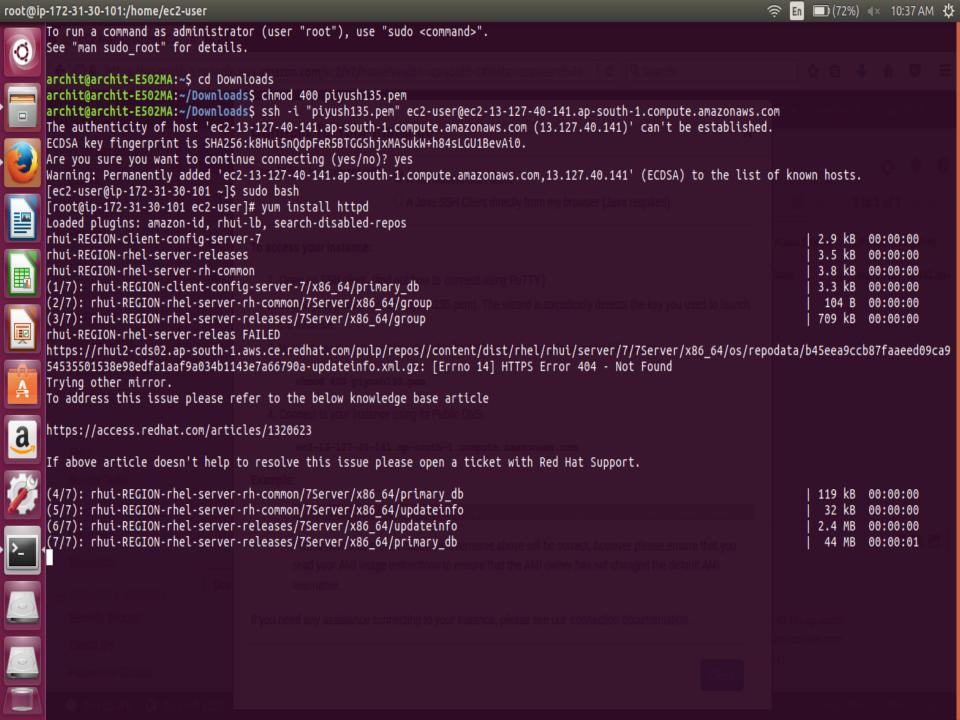








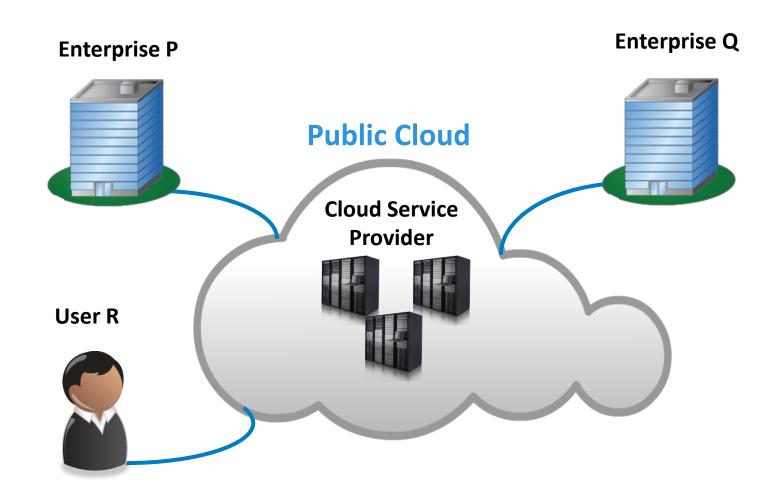




Assignment

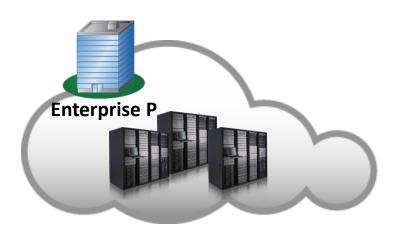
- Install VMWare Workstation in your machine.
 - Create a VM (any Unix flavor)
 - Install a Apache WS on the VM
 - ☐ Create a webpage
 - □ host the webpage on the Apache WS

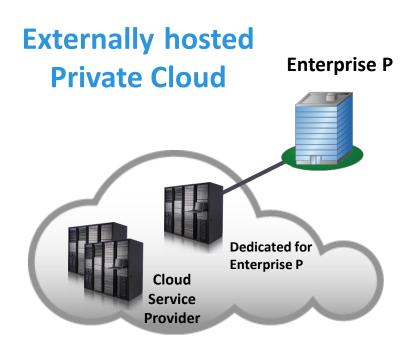
Taxonomy of Deployment Model – Public Cloud



Deployment Model - Private Cloud

On-premise Private Cloud

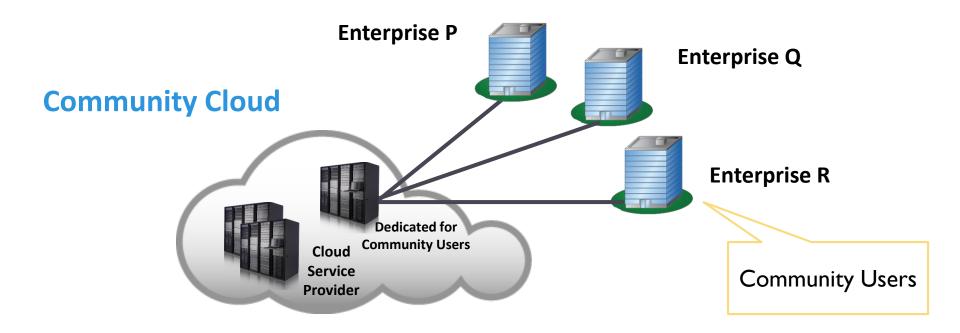




Deployment Model – Hybrid Cloud

Private Cloud Enterprise P Enterprise P Enterprise Q Cloud Service Provider

Deployment Model – Community Cloud



- Cloud infrastructure is shared by several organizations and supports a specific community that has shared concerns
- Managed by the organizations or by a third party

Vendors



Cloud Service Provider

- ▶ Cloud service providers: A cloud service provider, or CSP, is a company/Organization that offers some component of cloud computing -- typically infrastructure as a service (laaS), software as a service (SaaS) or platform as a service (PaaS) -- to other businesses or individuals.
- Google
- Microsoft
- Citrix
- Joyent. San Francisco infrastructure-as-a-service provider,
- CenturyLink.

Cloud Service Provider

- Amazon.
- ► IBM.
- Salesforce
- Rackspace
- Verizon Terremark

Cloud Software

- Give the best example for open source Cloud Computing Management Software: OpenStack
- OpenStack software controls large pools of compute, storage, and networking resources throughout a datacenter, managed through a dashboard or via the OpenStack API.
- Explain what is the full form and usage of "EUCALYPTUS" in cloud computing.
- EUCALYPTUS" full form stands for Elastic Utility Computing Architecture for Linking Your Programs to Useful Systems".

Cloud Software

- Both are open source software infrastructure in cloud computing, which enable us to implement clusters in cloud computing platform. It is mainly used to build public, hybrid and private clouds.
- List the platforms which are used for large scale cloud computing: The platforms that are used for large scale cloud computing are:
 - Apache Hadoop
 - MapReduce

Large Cloud Databases

- Mention the name of some large cloud providers and databases.
- □ Google Big table
- Amazon Simple Database
- Cloud based SQL(Sequential Query Language)

Cloud Vs Traditional

- The expenditure of the traditional data center is expensive due to heating and hardware/software issues.
- Cloud being scaled when there is an increase in demand. Mostly the expenditure is on the maintenance of the data centers, while this issues are not faced in cloud computing.

▶ Etc....

API's

- What are the uses of API's in cloud services?
- □ API's (Application Programming Interface are used to eliminate the necessity to write the complete programs.
- The instructions are provided to make communication between one or more applications.
- □ Creation of applications is made easy and access for the link of cloud services with other systems.

Benefits of Cloud

- Cost & management
- ▶ Economies of scale, "out-sourced" resource management
- Reduced Time to deployment
- Ease of assembly, works "out of the box"
- Scaling
- On demand provisioning, co-locate data and compute
- Reliability
- Massive, redundant, shared resources
- Sustainability

Benefits of Cloud

- Hardware not owned
- What are the benefits of cloud computing?
- Data backup and storage of data.
- Powerful server capabilities.
- Incremented productivity.
- Cost effective and time saving.