

SESSION 1

An Introduction to Big Data and Cloud

Objectives

2

- ❑ Explain the basics of Big Data
- ❑ Describe Big Data and Cloud



Introduction to Big Data (1-2)

3

- Huge volumes of diverse data being produced at high speeds
- Some definitions of Big Data:



A large volume of scientific data for visualization

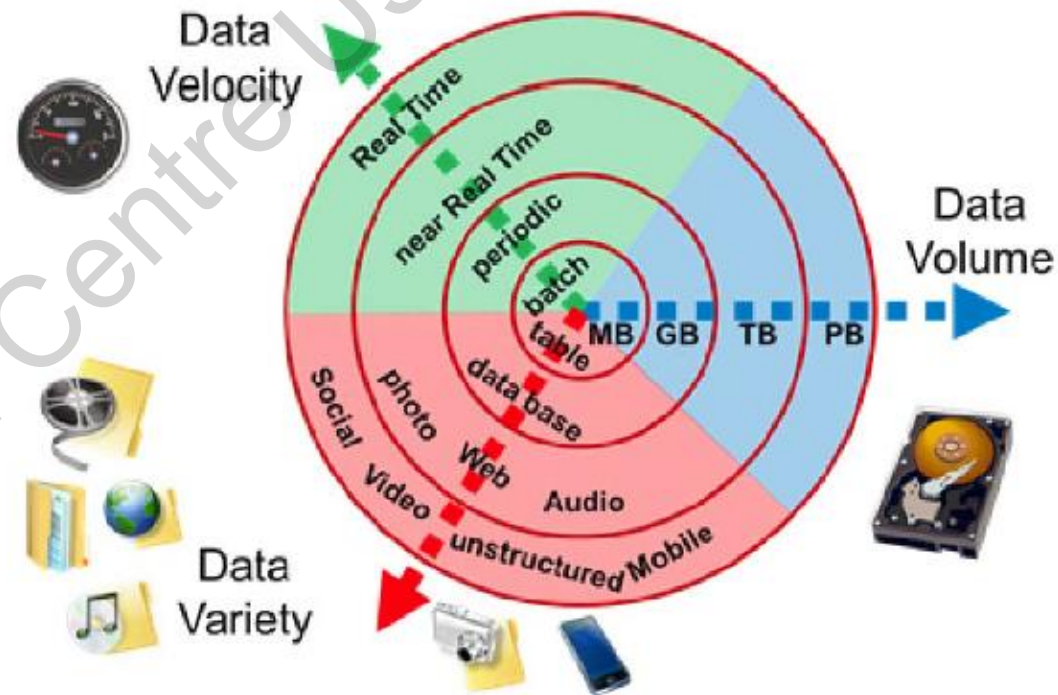
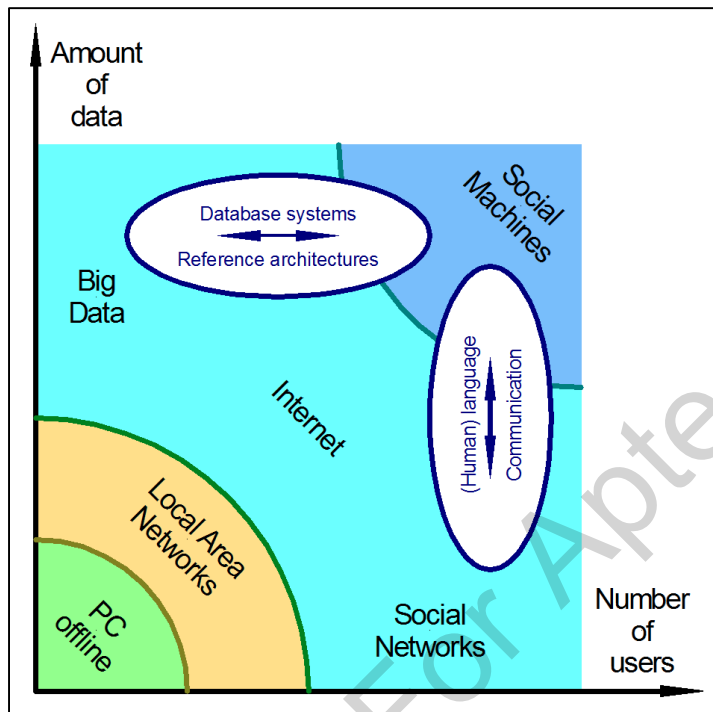
Big Data may also extend to four Vs, namely, volume, variety, velocity, and value

The amount of data, which surpasses existing technology's capability to store, manage, and process efficiently



Introduction to Big Data (2-2)

4



Benefits of Big Data (1-2)

5

Identify new customers

- Data stored by social networks, offers huge insights into users' behavior and interests

Find new marketing opportunities

- Any product or service must be upgraded constantly to survive in the market

Give recommendations to existing customers

- Provide customers with innovative products and services which might be of their interest



Benefits of Big Data (2-2)

6

Make more profitable ads

- Data from social media are used to generate appealing and effective ad campaigns

Measure campaign results more accurately

- Calculates campaign's performance and individual media channel's effectiveness

Required tools and skills

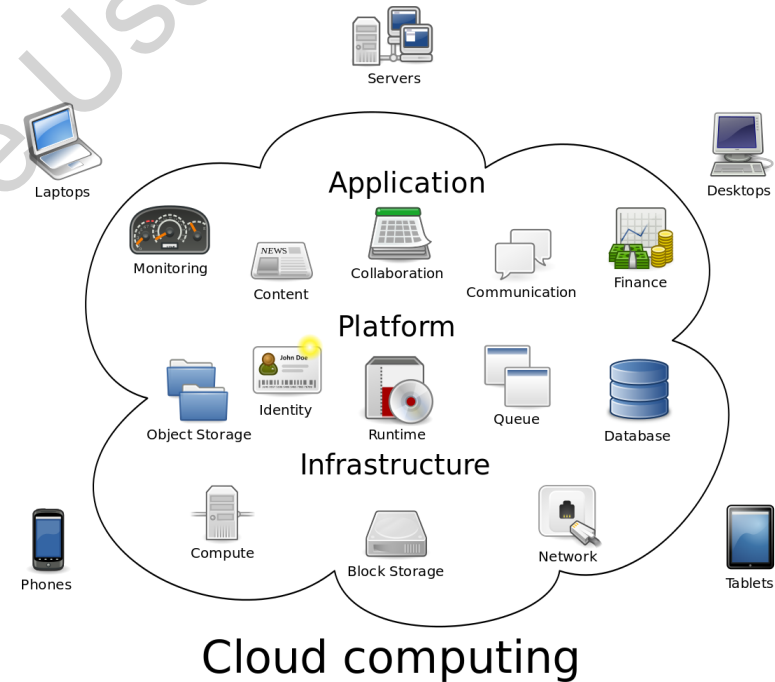
- Identifying the correct tools that will make Big Data more liable



Cloud Computing

7

- Refers to the process of manipulating, configuring, and accessing the applications online.
- Provides users with online data storage facilities, infrastructure, and applications.



Cloud Computing Technologies

8

Technologies of Cloud Computing are:

Virtualization

Service-Oriented
Architecture (SOA)

Grid Computing

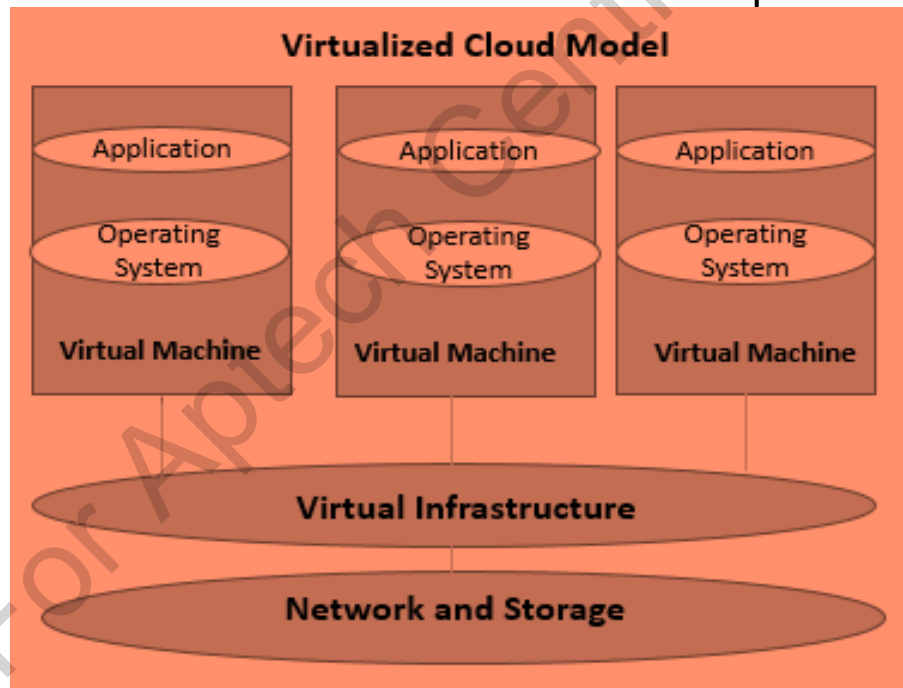
Utility Computing



Virtualization

9

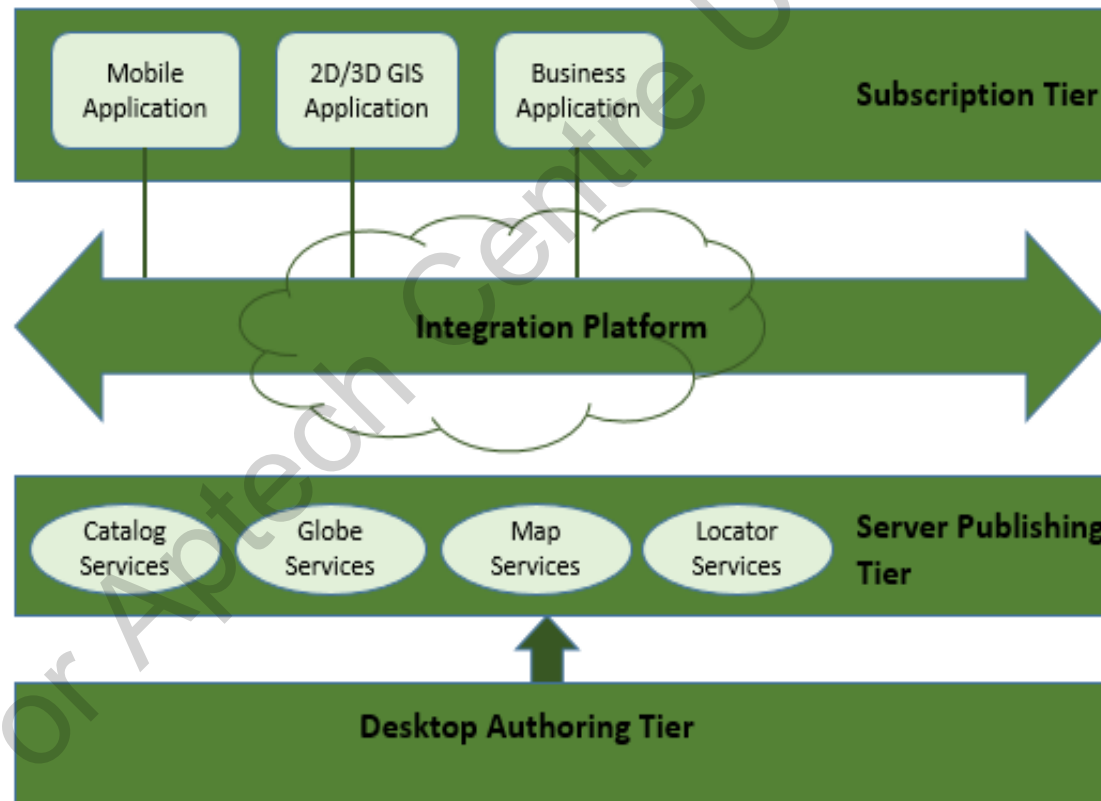
- ❑ A technique that allows users to share a physical instance or resource with many organizations
- ❑ Achieved by allocating a logical name to a physical resource
- ❑ It provides virtual isolation between multiple tenants



Service-Oriented Architecture

10

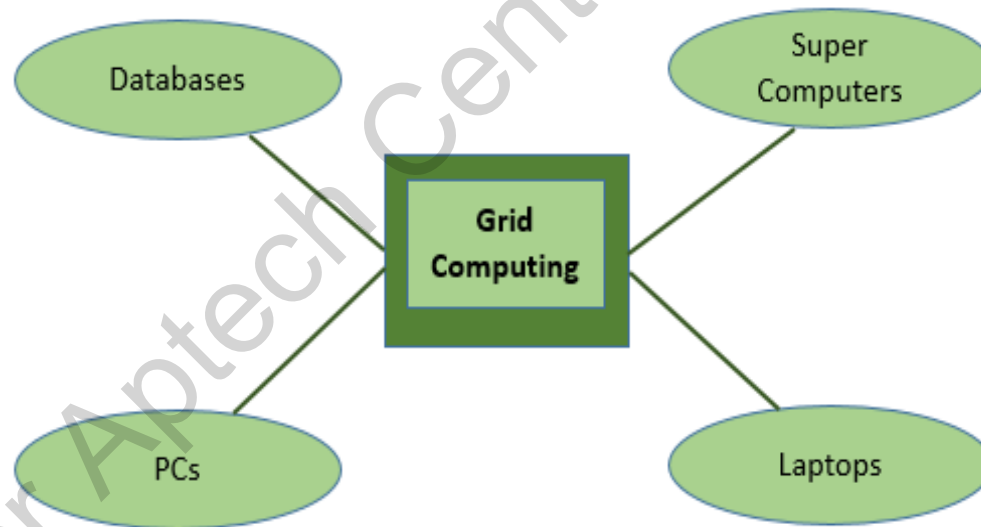
- It allows users to utilize applications and exchange data among applications of different vendors



Grid Computing

11

- ❑ A form of distributed computing involving a group of computers
- ❑ Heterogeneous and distributed geographically computers
- ❑ Fragments a compound task into smaller tasks



Utility Computing

12

- ❑ Is built on Pay per Use model
- ❑ Provides computational resources in metered service form

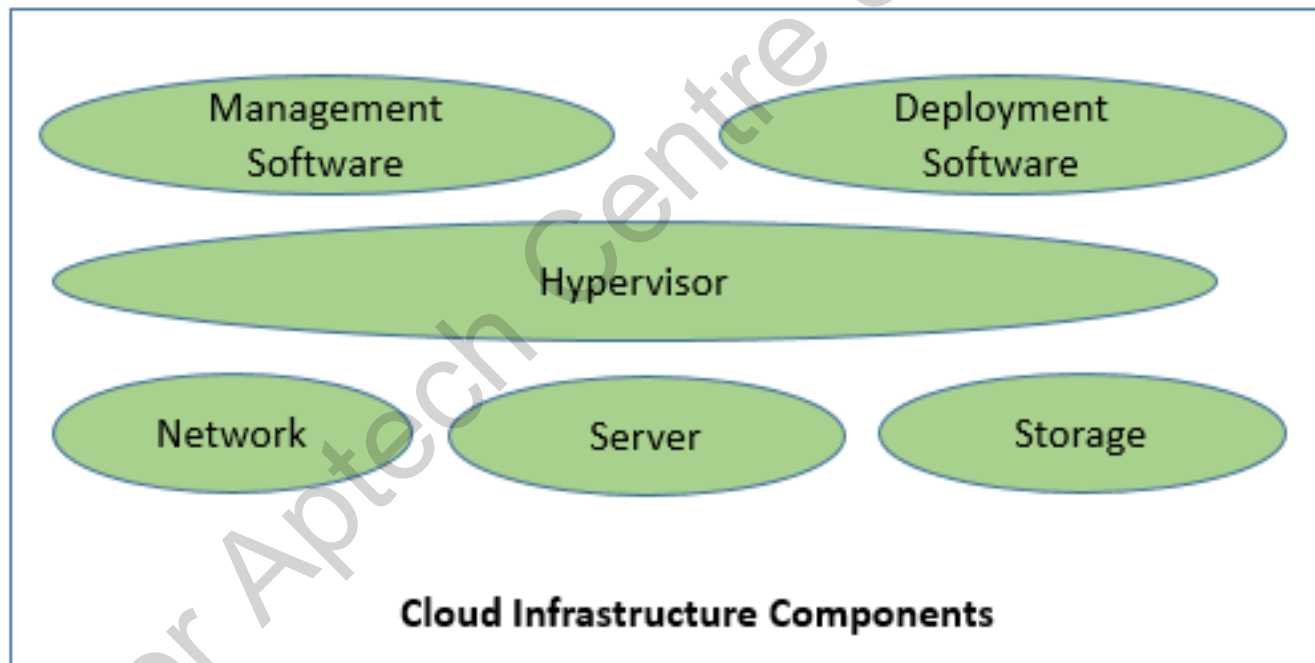
Managed IT services, cloud computing, and grid computing are all based on the utility computing concept.



Cloud Infrastructure

13

- Includes servers, storage, network, deployment software, and platform virtualization.



Cloud Deployment Models

14

- Following are cloud deployment models:

Public Cloud Model

Private Cloud Model

Hybrid Cloud Model

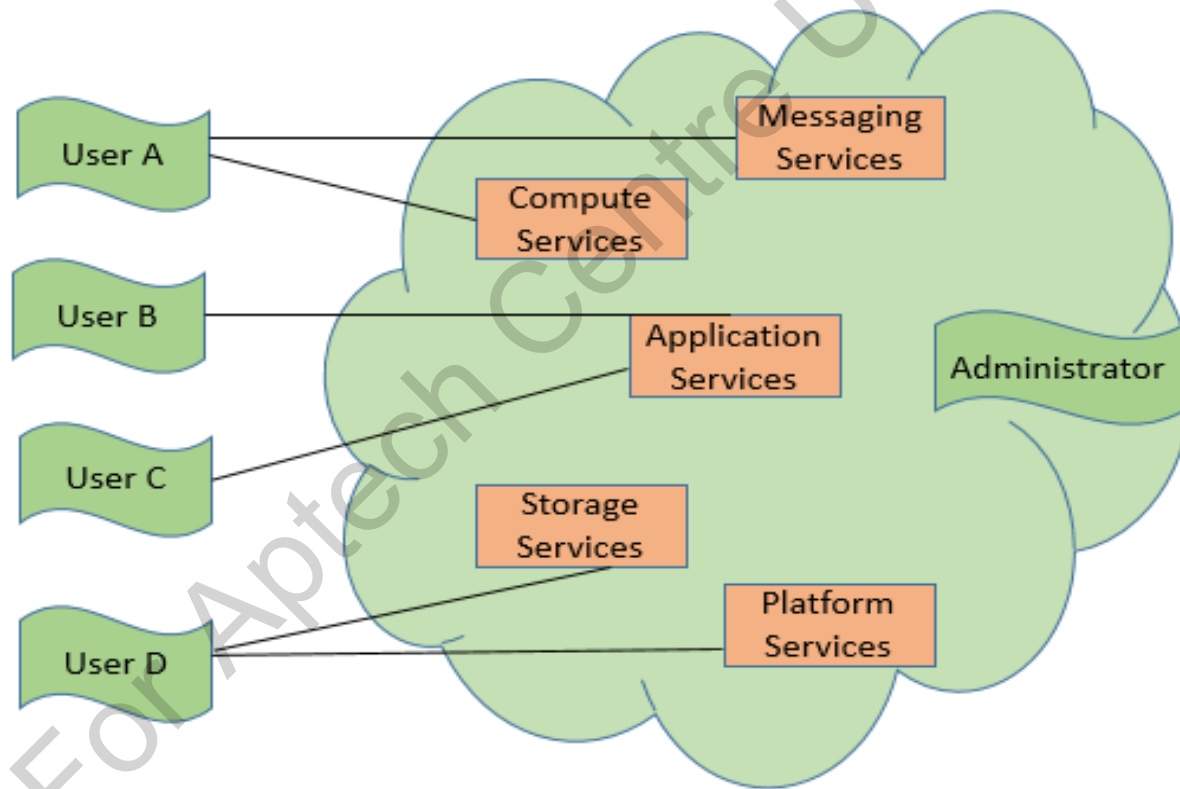
Community Cloud Model



Public Cloud Model

15

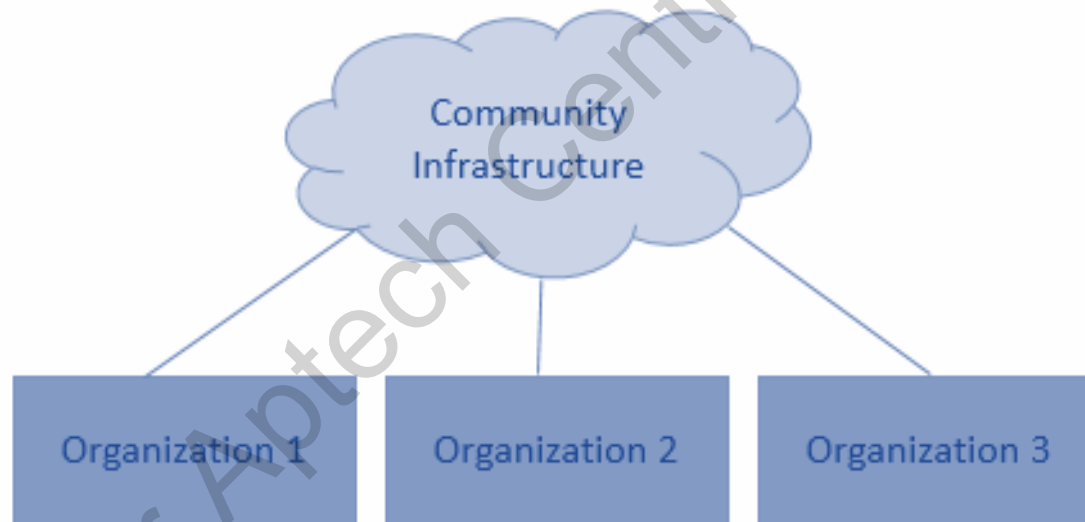
- Ensures systems and services are available easily to the public



Community Cloud Model

16

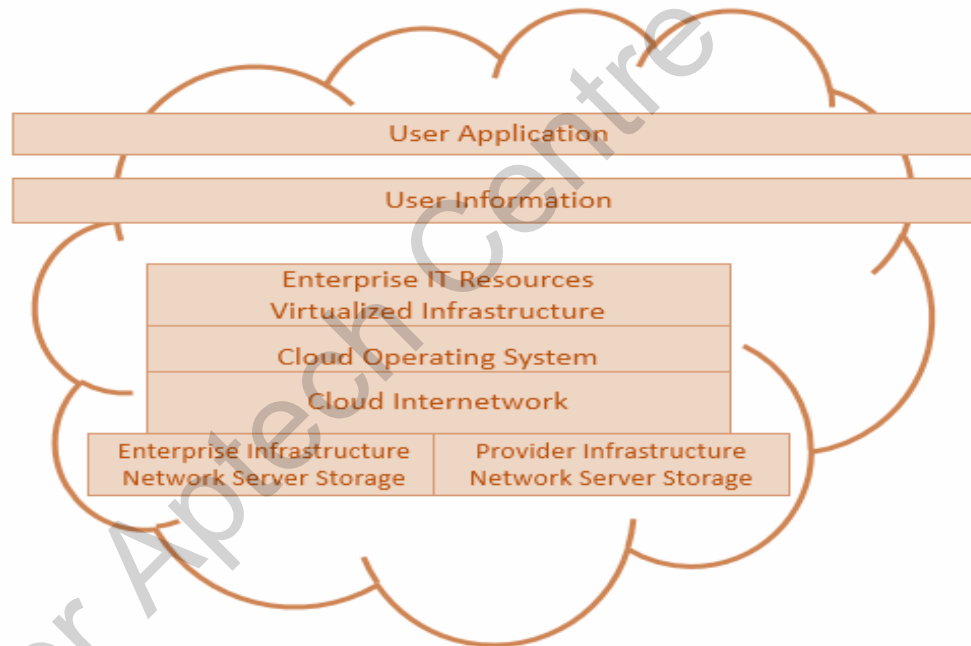
- Ensures system and services are accessible by a cluster of organizations
- Shares the infrastructure among multiple organizations



Private and Hybrid Cloud Models

17

- ❑ **Private cloud model** ensures systems and services are available within an organization



- ❑ **Hybrid Cloud Model** - Combination of public cloud and private cloud.



Cloud Service Models

18

Following are Cloud Service Models:

Infrastructure as
a Service (IaaS)
Model

Platform as a
Service (PaaS)
Model

Software as a
Service (SaaS)
Model

Identity as a
Service (IDaaS)
Model

Network as a
Service (NaaS)
Model



IaaS (1-2)

19

Provides users access to basic resources:

Virtual
machine disk
storage

Virtual Local
Area Network
(VLANs)

Load
balancers

IP addresses

Software
bundles

Its characteristics are:

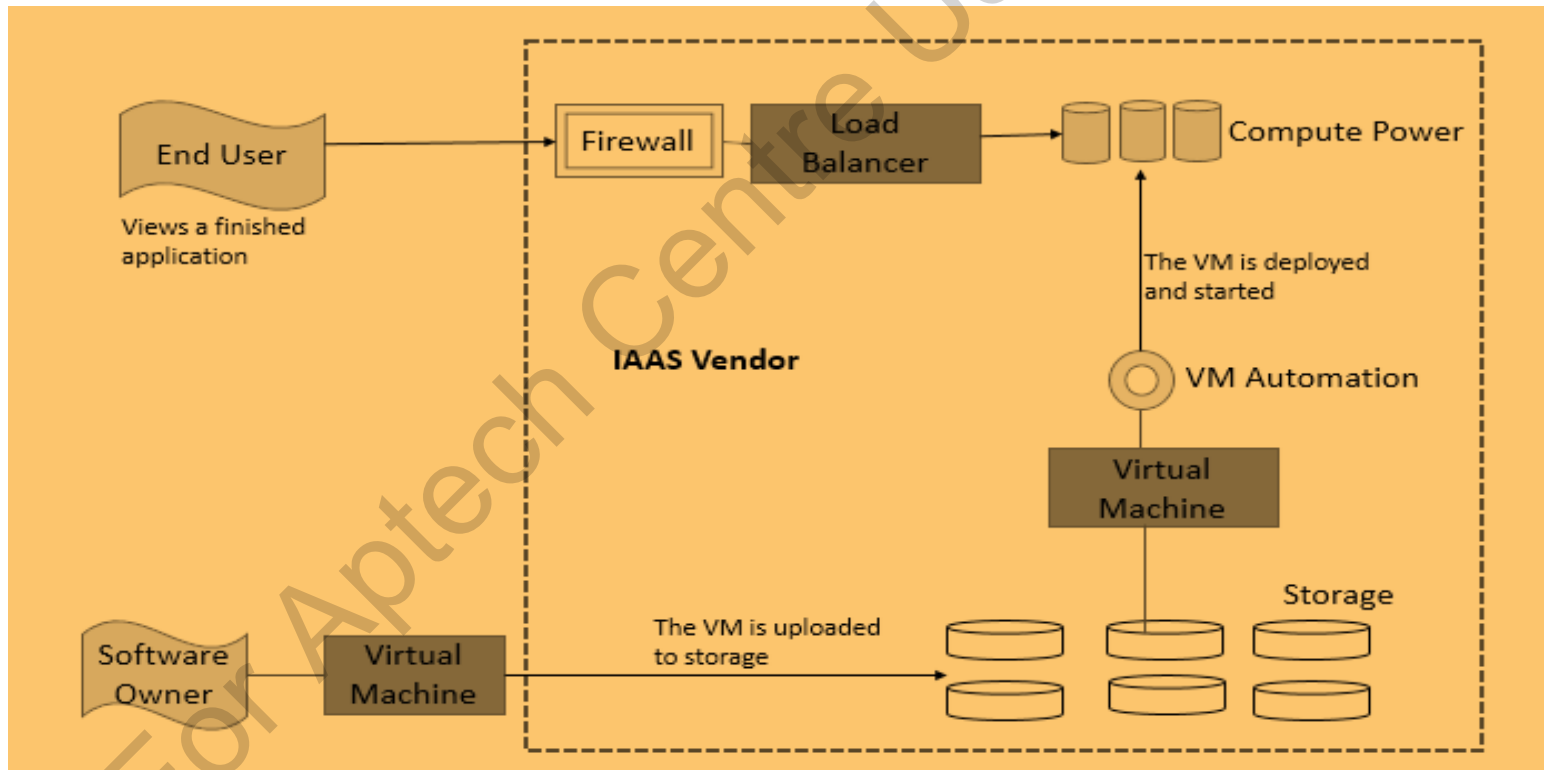
- Supports virtual machines having pre-installed software operating systems
- Offers on-demand availability of resources
- Provision to save specific data copies in different locations
- Allows scaling up and down of computing resources



IaaS (2-2)

20

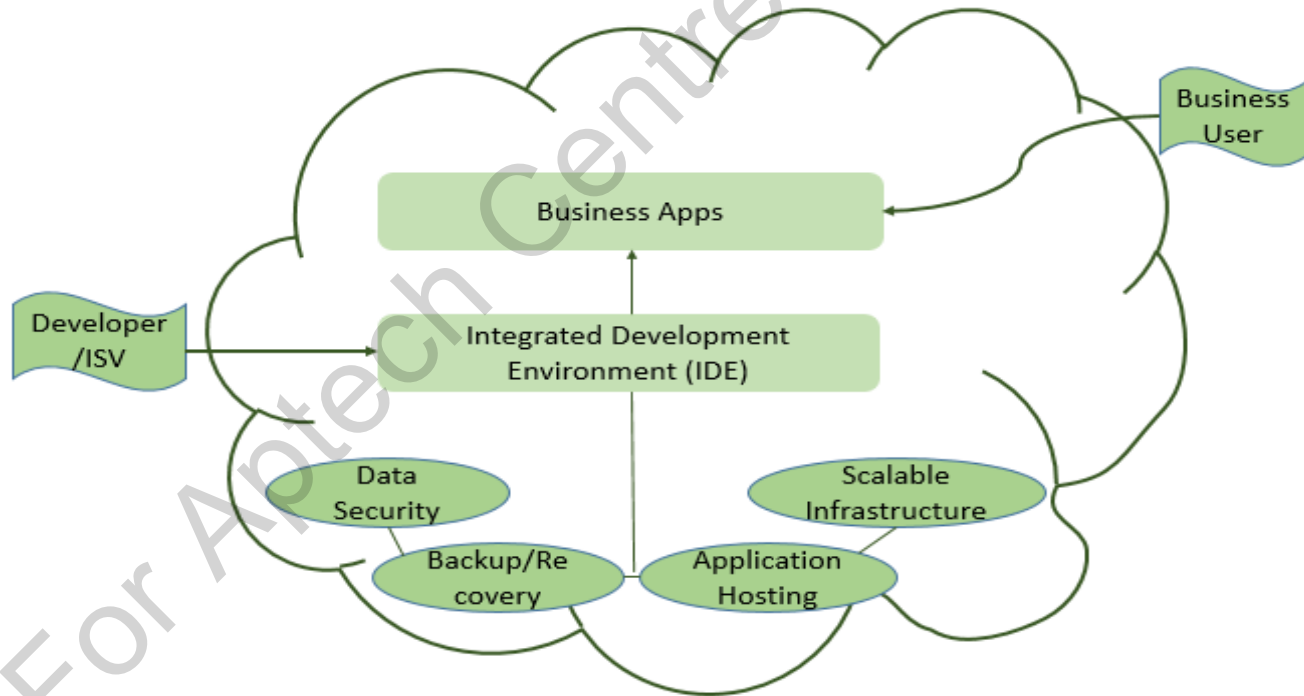
- The resources are offered to end users through server virtualization



PaaS

21

- ❑ Stands for Platform as a Service
- ❑ Provides platform and tools to create applications and runtime environment



SaaS

22

- ☐ Stands for Software as a Service
- ☐ Offers software applications in the form of a service to end users
- ☐ A software that is installed on a hosted service
- ☐ Applications are:

Billing and Invoicing System

Customer Relationship Management (CRM) Applications

Help Desk Applications

Human Resource (HR) Solutions



IDaaS

23

- Supports management of digital entity identity (information)
- It consists of services:



NaaS

24

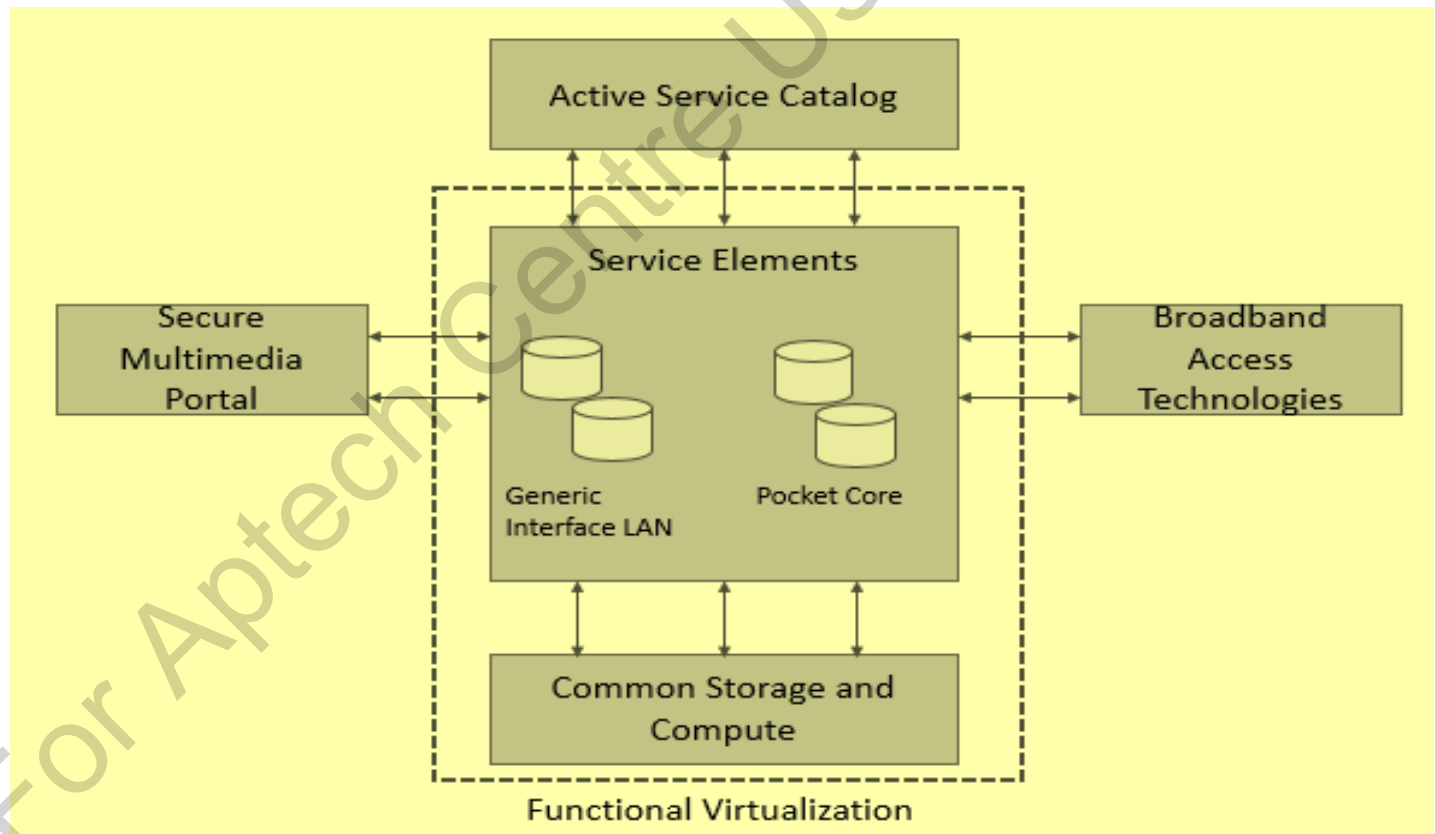
- ❑ Stands for Network as a Service
- ❑ Provides a direct and secure access to network infrastructure
- ❑ Employs virtualized network infrastructure
- ❑ Maintains and supervises the network resources
- ❑ Provides network as a utility and works as per the pay-per-use model



Mobile NaaS

25

- Provides enhanced efficiency and flexible control over mobile devices



Cloud Computing Advantages

26

Cost Saving

- Helps users to save significant capital costs with no in-house server storage and application constraints

Reliability

- Is more dependable and consistent and a quick failover mechanism

Manageability

- Offers improved and streamlined IT management and maintenance facilities

Strategic Edge

- Ensures users focus more on their main business activities and goals



Cloud Computing Disadvantages

27

Downtime

- Handling numerous clients every day may cause service provider to be flooded, which can result in technical outages

Security

- Easy accessibility gives users chance to misuse ambiguities and vulnerabilities

Vendor Lock-In

- Companies face problems while migrating their services between vendors

Limited Control

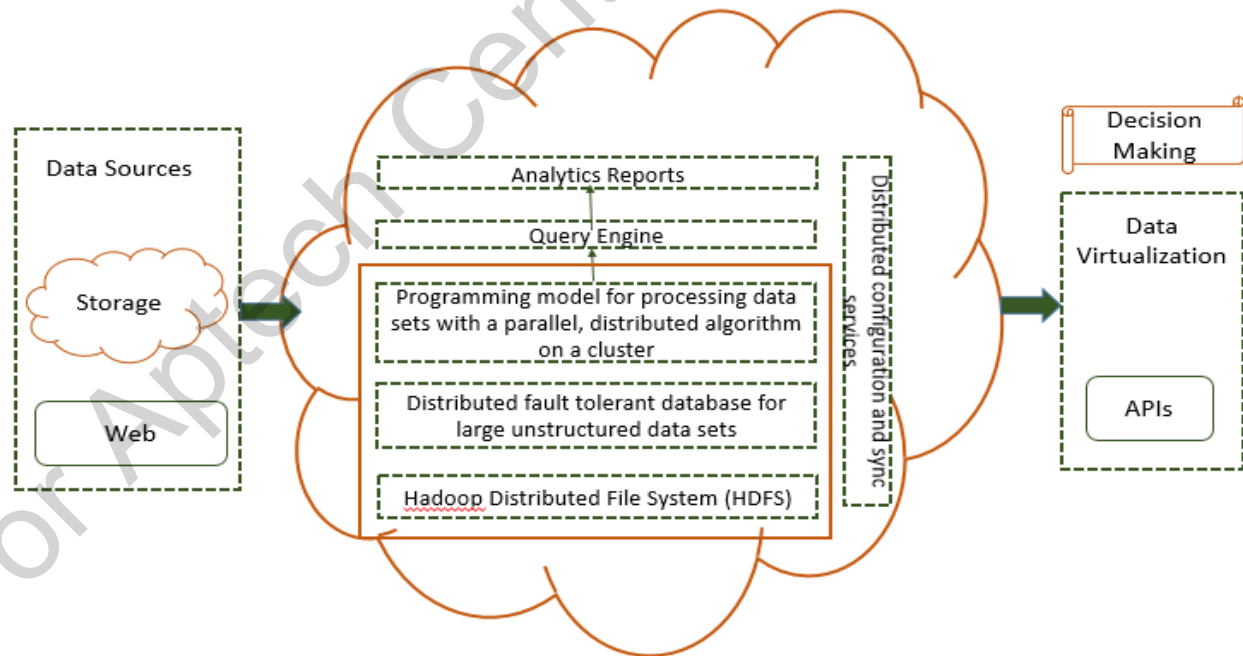
- Service provider holds complete ownership of the cloud infrastructure, very little control is given to users



Big Data and Cloud Computing

28

- ❑ Go hand in hand
- ❑ Big Data gives users the access to utilize commodity computing to process distributed queries
- ❑ Cloud computing offers the triggering engine



Summary (1-2)

29

- ❑ Big Data includes huge volumes of diverse data being produced mostly at high speeds.
- ❑ Some benefits of Big Data include identifying new customers, finding new marketing opportunities, giving great offers to existing customers, and making more profitable ads.
- ❑ Cloud computing can be defined as the process of manipulating, configuring, and accessing the applications online. It provides users with online data storage facilities, infrastructure, and application.
- ❑ Advantages of cloud computing are Cost Savings, Reliability, Manageability, and Strategic Edge.



Summary (2-2)

30

- ❑ Disadvantages of cloud computing include Downtime, Security, Vendor Lock-In, and Limited Control.
- ❑ Cloud computing and Big Data go hand in hand.
- ❑ Big Data gives users the access to utilize commodity computing to process distributed queries across multiple datasets and return resultant sets within appropriate time.
- ❑ Cloud computing offers the triggering engine via Hadoop, which is a class of distributed data-processing platforms.

