MODULE 4

CLOUD COMPUTING CHARACTERISTICS AND BENEFITS

4.1 Drivers for cloud computing

- Business requirements
  - Transformation of IT processes to achieve more with less
  - better agility and higher availability at reduced expenditures
  - Reduced time-to-market
  - Accelerated pace of innovation
- IT challenges to meet business requirements are:
  - Serving customers worldwide round the clock, refreshing technology quickly, faster provisioning of IT resources-all at reduced cost.
- These challenges are addressed with the emergence of cloud computing.

4.2 Definition

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.

4.3 Characteristics of Cloud Computing

Cloud infrastructure should have five essential characteristics:

- **On-demand self-service:** A consumer can unilaterally provision computing capabilities, such as server time and network storage, as needed, automatically without requiring human interaction with each service provider.

  A cloud service provider publishes a service catalogue, which contains information about all cloud services available to consumers. The service catalogue includes information about service attributes, prices, and request processes. Consumers view the service catalogue via a web-based user interface and use it to request for a service. Consumers can either leverage the “ready-to-use” services or change a few service parameters to customize the services.

- **Broad network access:** Capabilities are available over the network and accessed through standard mechanisms that promote use by heterogeneous thin or thick client platforms (for example, mobile phones, tablets, laptops, and workstations).

- **Resource pooling:** The provider's computing resources are pooled to serve multiple consumers using a multitenant model, with different physical and virtual resources dynamically assigned and reassigned according to consumer demand. There is a sense of location independence in that the customer generally has no control or knowledge over the exact location of the provided resources but may be able to specify location at a higher level of abstraction (for example, country, state, or data center). Examples of resources include storage, processing, memory, and network bandwidth.

- **Rapid elasticity:** Capabilities can be elastically provisioned and released, in some cases automatically, to scale rapidly outward and inward commensurate with demand. To the consumer, the capabilities available for provisioning

.

- **Measured service:** Cloud systems automatically control and optimize resource use by leveraging a metering capability at some level of abstraction appropriate to the type of service (for example, storage, processing, bandwidth, and active user accounts). Resource usage can be monitored, controlled, and reported, providing transparency for both the provider and consumer of the utilized service.

4.4 Benefits of cloud computing

Benefits	Description
Reduced IT cost	• Reduces the up-front capital expenditure (CAPEX)
Business agility	• Provides the ability to deploy new resources quickly • Enables businesses to reduce time-to-market
Flexible scaling	• Enables consumers to scale up, scale down, scale out, or scale in the demand for computing resources easily • Consumers can unilaterally and automatically scale computing resources
High availability	• Ensures resource availability at varying levels, depending on consumer's policy and priority

Fig 4.1: Cloud enabling technologies

4.5 Cloud Enabling Technologies

Technologies	Description
Grid computing	• Form of distributed computing • Enables resources of numerous computers in a network to work on a single task at the same time
Utility computing	• Service provisioning model that offers computing resources as a metered service
Virtualization	• Abstracts physical characteristics of IT resources from resource users • Enables resource pooling and creating virtual resources from pooled resources
Service-oriented architecture (SOA)	• Provides a set of services that can communicate with each other

4.6.1 Infrastructure-as-a-Service

- The capability provided to the consumer is to provision processing, storage, networks, and other fundamental computing resources where the consumer is able to deploy and run arbitrary software, which can include operating systems and applications.
- The consumer does not manage or control the underlying cloud infrastructure but has control over operating systems and deployed applications; and possibly limited control of select networking components (for example, host firewalls).
- IaaS is the base layer of the cloud services stack (see Fig 4.3[a]). It serves as the foundation for both the SaaS and PaaS layers.
- Amazon Elastic Compute Cloud (Amazon EC2) is an example of IaaS that provides

scalable compute capacity, on-demand, in the cloud. It enables consumers to leverage Amazon's massive computing infrastructure with no up-front capital investment.

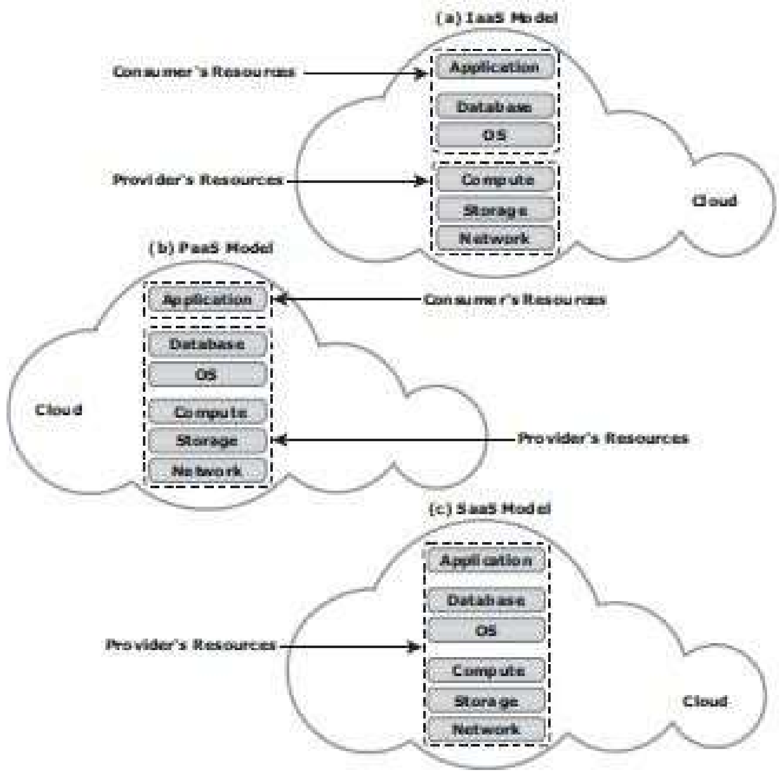

Fig 4.3 IaaS, PaaS, and SaaS models

4.6.2 Platform-as-a-Service

- The capability provided to the consumer is to deploy onto the cloud infrastructure consumer-created or acquired applications created using programming languages, libraries, services, and tools supported by the provider.
- The consumer does not manage or control the underlying cloud infrastructure including network, servers, operating systems, or storage, but has control over the deployed applications and possibly confi guration settings for the application-hosting environment. (See Fi 4.3 [b]).
- PaaS is also used as an application development environment, offered as a service by the cloud service provider.
- The consumer may use these platforms to code their applications and then deploy the applications on the cloud.
- Because the workload to the deployed applications varies, the scalability of

computing resources is usually guaranteed by the computing platform, transparently.

4.6.3 Software-as-a-Service

- The capability provided to the consumer is to use the provider's applications running on a cloud infrastructure.
- The applications are accessible from various client devices through either a thin client interface, such as a web browser (for example, web-based e-mail), or a program interface.
- The consumer does not manage or control the underlying cloud infrastructure including network, servers, operating systems, storage, or even individual application capabilities, with the possible exception of limited user-specifi c application configuration settings. (Fig 4.3[c]).
- In a SaaS model, applications, such as customer relationship management (CRM), e-mail, and instant messaging (IM), are offered as a service by the cloud service providers.
- The cloud service providers exclusively manage the required computing infrastructure and software to support these services
- . The consumers may be allowed to change a few application configuration settings to customize the applications.
- EMC Mozy is an example of SaaS. Consumers can leverage the Mozy console to perform automatic, secured, online backup and recovery of their data with ease. Salesforce.com is a provider of SaaS-based CRM applications, such as Sales Cloud and Service Cloud.Google App Engine and Microsoft Windows Azure Platform are examples of PaaS.

<u>4.7 Cloud Deployment Models</u>

Cloud computing is classified into four deployment models : public, private, community, and hybrid — which provide the basis for how cloud infrastructures are constructed and consumed.

4.7.1 Public Cloud

- In a public cloud model, the cloud infrastructure is provisioned for open use by the general public. It may be owned, managed, and operated by a business, academic, or government organization, or some combination of them.

- It exists on the premises of the cloud provider.
- Consumers use the cloud services offered by the providers via the Internet and pay metered usage charges or subscription fees.
- An advantage of the public cloud is its low capital cost with enormous scalability. However, for consumers, these benefits come with certain risks: no control over the resources in the cloud, the security of confidential data, network performance, and interoperability issues.
- Popular public cloud service providers are Amazon, Google, and Salesforce.com.
- Fig 4.2 shows a public cloud that provides cloud services to organizations and individuals.

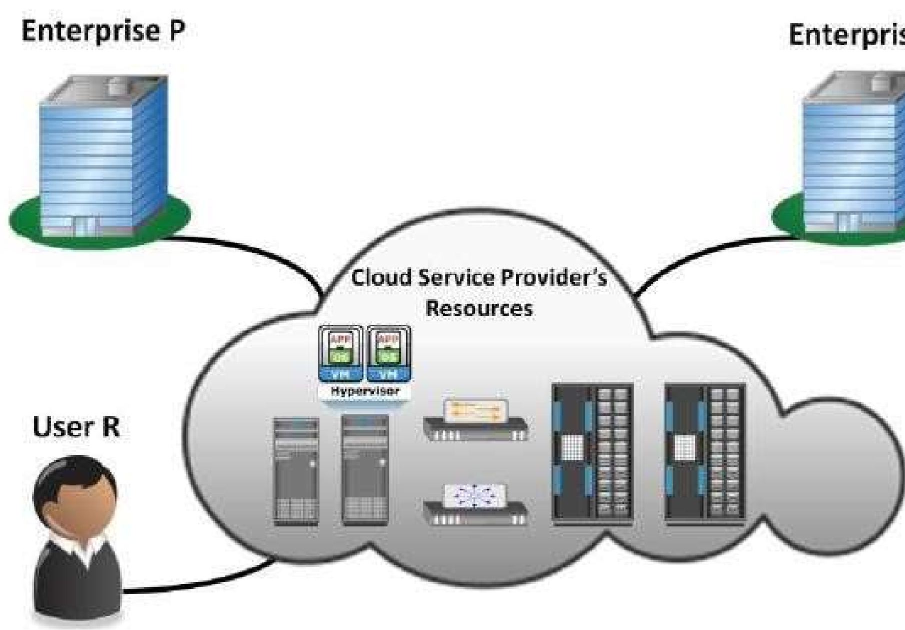

Fig 4.4 Public cloud

4.7.2 Private Cloud

- In a private cloud model, the cloud infrastructure is provisioned for exclusive use by a single organization comprising multiple consumers (for example, business units).
- It may be owned, managed, and operated by the organization, a third party, or some combination of them, and it may exist on or off premises.

- Following are two variations to the private cloud model:
    1. **On-premise private cloud:** The on-premise private cloud, also known as internal cloud, is hosted by an organization within its own data centers (see Fig 4.5[a]). This model enables organizations to standardize their cloud service management processes and security, although this model has limitations in terms of size and resource scalability.
    2. **Externally hosted private cloud:** This type of private cloud is hosted external to an organization (see Fig 4.4[b]) and is managed by a third party organization. The third-party organization facilitates an exclusive cloud environment for a specifi c organization with full guarantee of privacy and confi dentiality.

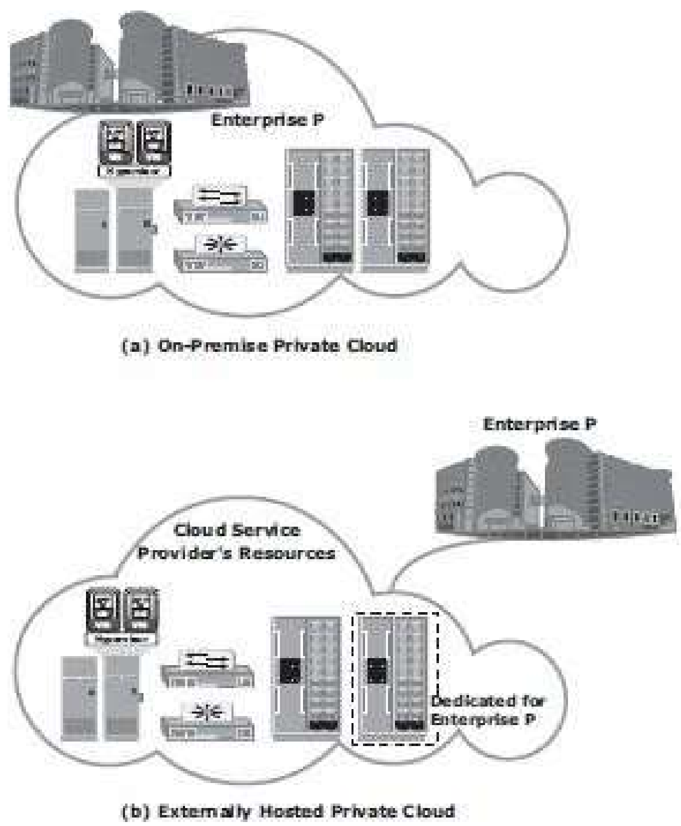

Fig 4.4 On-premise and externally hosted private clouds

4.7.3 Community Cloud

- In a community cloud model, the cloud infrastructure is provisioned for exclusive use by a specific community of consumers from organizations that have shared concerns.
- It may be owned, managed, and operated by one or more of the organizations in the

community, a third party, or some combination of them, and it may exist on or off premises. (Fig 4.5).

- In a community cloud, the costs spread over to fewer consumers than a public cloud. Hence, this option is more expensive but might offer a higher level of privacy, security, and compliance.
- The community cloud also offers organizations access to a vast pool of resources compared to the private cloud.
- An example in which a community cloud could be useful is government agencies. If various agencies within the government operate under similar guidelines, they could all share the same infrastructure and lower their individual agency's investment.

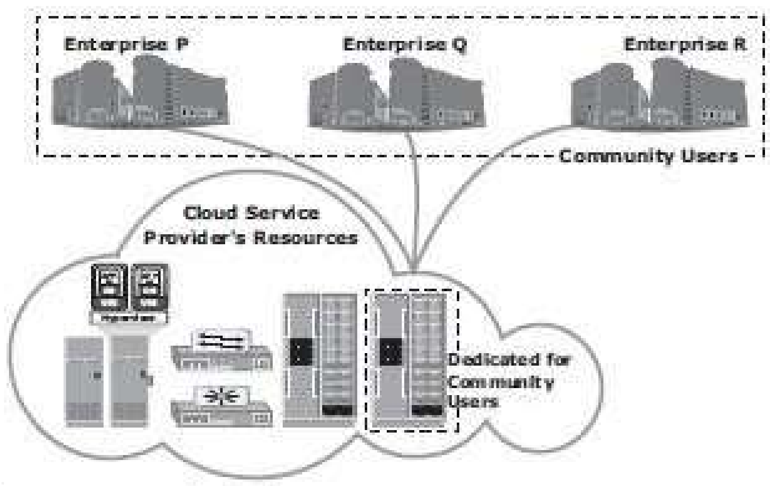

Fig 4.5 Community cloud

4.7.4 Hybrid Cloud

- In a hybrid cloud model, the cloud infrastructure is a composition of two or more distinct cloud infrastructures (private, community, or public) that remain unique entities, but are bound together by standardized or proprietary technology that enables data and application portability (for example, cloud bursting for load balancing between clouds).
- The hybrid model allows an organization to deploy less critical applications and data to the public cloud, leveraging the scalability and cost-effectiveness of the public cloud.
- The organization's mission-critical applications and data remain on the private cloud that provides greater security. Fig 4.6 shows an example of a hybrid cloud.

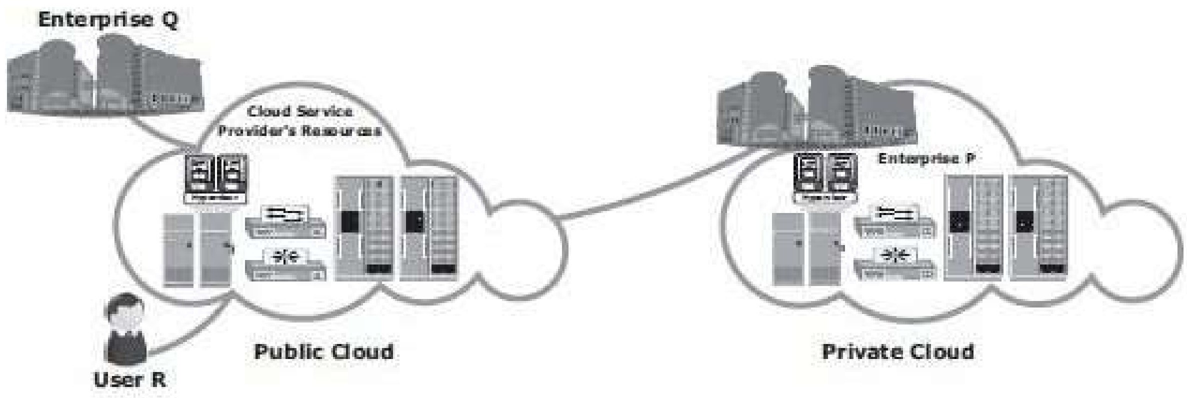

Fig 4.6 Hybrid cloud

4.8 Cloud Computing Infrastructure

A cloud computing infrastructure is the collection of hardware and software that enables the five essential characteristics of cloud computing. Cloud computing infrastructure usually consists of the following layers:

- ✓ Physical infrastructure
- ✓ Virtual infrastructure
- ✓ Applications and platform software
- ✓ Cloud management and service creation tools

The resources of these layers are aggregated and coordinated to provide cloud services to the consumers (see Fig 4.6).

4.8.1 Physical infrastructure

- ➢ The physical infrastructure consists of physical computing resources, which include physical servers, storage systems, and networks.
- ➢ Physical servers are connected to each other, to the storage systems, and to the clients via networks, such as IP, FC SAN, IP SAN, or FCoE networks.
- ➢ Cloud service providers may use physical computing resources from one or more data centers to provide services.
- ➢ If the computing resources are distributed across multiple data centers, connectivity must be established among them.

- The connectivity enables the data centers in different locations to work as a single large data center.
- This enables migration of business applications and data across data centers and provisioning cloud services using the resources from multiple data centers.

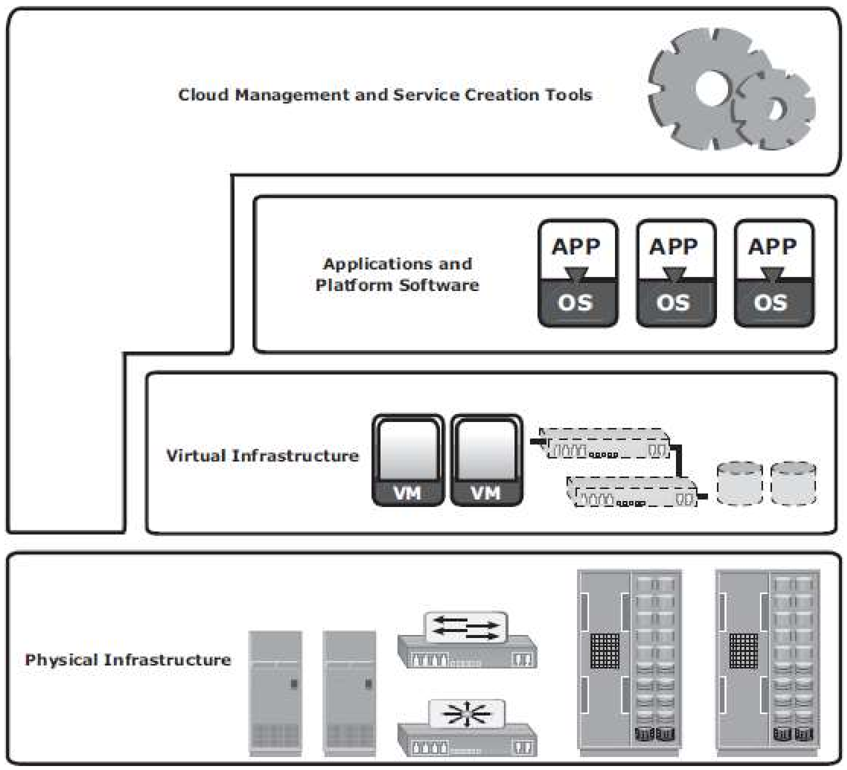

Fig 4.7 Cloud infrastructure layers

4.8.2 Virtual Infrastructure

- Cloud service providers employ virtualization technologies to build a virtual infrastructure layer on the top of the physical infrastructure.
- Virtualization enables fulfilling some of the cloud characteristics, such as resource pooling and rapid elasticity.
- It also helps reduce the cost of providing the cloud services.
- Some cloud service providers may not have completely virtualized their physical infrastructure yet, but they are adopting virtualization for better efficiency and optimization.
- Virtualization abstracts physical computing resources and provides a consolidated

view of the resource capacity.

- The consolidated resources are managed as a single entity called a resource pool.
- For example, a resource pool might group CPUs of physical servers within a cluster.
- The capacity of the resource pool is the sum of the power of all CPUs (for example, 10,000 megahertz) available in the cluster.
- In addition to the CPU pool, the virtual infrastructure includes other types of resource pools, such as memory pool, network pool, and storage pool.
- Apart from resource pools, the virtual infrastructure also includes identity pools, such as VLAN ID pools and VSAN ID pools.
- The number of each type of pool and the pool capacity depend on the cloud service provider's requirement to create different cloud services.
- Virtual infrastructure also includes virtual computing resources, such as virtual machines, virtual storage volumes, and virtual networks.
- These resources obtain capacities, such as CPU power, memory, network bandwidth, and storage space from the resource pools.
- Virtual networks are created using network identifiers, such as VLAN IDs and VSAN IDs from the respective identity pools.
- Virtual computing resources are used for creating cloud infrastructure services.

4.8.3 Applications and Platform Software

- This layer includes a suite of business applications and platform software, such as the OS and database.
- Platform software provides the environment on which business applications run.
- Applications and platform software are hosted on virtual machines to create SaaS and PaaS. For SaaS, both the application and platform software are provided by cloud service providers.
- In the case of PaaS, only the platform software is provided by cloud service providers; consumers export their applications to the cloud.

4.9 Cloud Adoption Considerations

Following are some key considerations for cloud adoption:

- **Selection of a deployment model:** Risk versus convenience is a key consideration for deciding on a cloud adoption strategy. This consideration also forms the basis for choosing the right cloud deployment model. A public cloud is usually preferred by individuals and start-up businesses. The tier 1applications should run on the private cloud, whereas less critical applications such as backup, archive, and testing can be deployed in the public cloud.

- **Application suitability:** Not all applications are good candidates for a public cloud. If an application workload is network traffi c-intensive, its performance might not be optimal if deployed in the public cloud. Also if the application communicates with other data center resources or applications, it might experience performance issues.

- **Financial advantage:** A careful analysis of financial benefits provides a clear picture about the cost-savings in adopting the cloud. The analysis should compare both the Total Cost of Ownership (TCO) and the Return on Investment (ROI) in the cloud and noncloud environment and identify the potential cost benefit. While calculating TCO and ROI, organizations and individuals should consider the expenditure to deploy and maintain their own infrastructure versus cloud-adoption costs.

- **Selection of a cloud service provider:** The selection of the provider is important for a public cloud. Consumers need to fi nd out how long and how well the provider has been delivering the services. They also need to determine how easy it is to add or terminate cloud services with the service provider. The consumer should know how easy it is to move to another provider, when required. They must assess how the provider fulfills the security, legal, and privacy requirements. They should also check whether the provider offers good customer service support.

- **Service-level agreement (SLA):** Cloud service providers typically mention quality of service (QoS) attributes such as throughput and uptime, along with cloud services. The QoS attributes are generally part of an SLA, which is the service contract between the provider and the consumers. The SLA serves as the foundation for the expected level of service between the consumer and the provider. Before adopting the cloud services, consumers should check whether the QoS attributes meet their requirements.