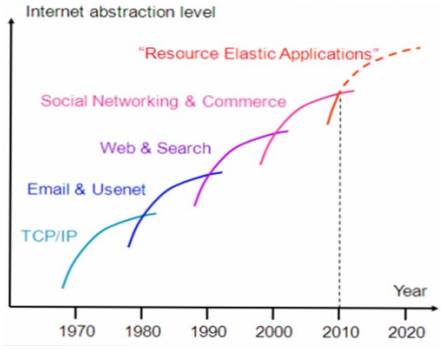
# Chapter 2. Cloud Computing Principles and Paradigms

***Mainframe Verses Cloud Computing***

| | Mainframe | Cloud Computing |
|---|---|---|
| **Differences** | Offers finite computing power | Provides almost infinite power and capacity |
| | Dummy terminals as user interface devices | PCs can provide local computing and cashing |
| **Similarity** | Cloud is a return to mainframe computing | |

**History of The Internet Revolution**

- The Internet has gone through four revolutions since its inception.
- Every revolution takes about ten years to be internalized.
- Cloud Computing is a major part of the *fifth revolution*.




## Cloud Computing

***Definitions:***

- *It is a computational environment that provides transparent access to a shared pool of computing resources matching the user's needs,*
  - e.g., processing power, Storage, Applications, and Components.

  - *[Egypt's Cloud Computing Strategic Plan-2011]*

***National Institute of Standards and Technology [NIST] Definitions***

- *It's a Computing model for enabling everywhere, 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

  - *NIST Definition of Cloud Computing.*
- Cloud Model promotes *availability* and is composed of *five essential characteristics, three service models*, and *four deployment models*.

| Five Essential Characteristics | Three Service | Four Deployment models |
|---|---|---|
| On Demand Service | Software as a Service (SaaS) | Private Cloud |
| Network Access Everywhere | Platform as a Service (PaaS) | Community Cloud |
| Resource Pooling | Infrastructure as a Service (IaaS) | Public Cloud |
| Rapid Elasticity | | Hybrid Cloud |
| Measured Service | | |

- **Essential Characteristics:**
  - **On-demand Self-Service:**
    A consumer can provision computing capabilities (e.g., server time, network, and storage) as needed automatically without requiring human interaction with each service provider.
  - **Broad Network Access**
    Capabilities available over network and accessed through standard mechanisms using heterogeneous client platforms (e.g., mobile phones, tablets, laptops, and workstations).
  - **Resource Pooling**
    Computing resources are pooled to serve multiple consumers according to their demand.
  - **Rapid Elasticity**
    Capabilities can elastically provision and released, in some cases automatically, to scale up and down according to consuming rate.
  - **Measured Service**
    Resource usage can monitor, control, and report, providing transparency for both provider and consumer of the utilized service.

- **Service Models:**
  - **Software as a Service (SaaS)**
    The computing capability Provided to consumer (External user) with provider's applications running on a cloud infrastructure. The applications are accessible from various client devices through web browser
    - (e.g., web-based email), or a program interface.
  - **Platform as a Service (PaaS):**
    The capability provided to the consumer (developer) by needed platform environment to develop his application using programming languages, libraries, services, and tools supported by the provider.
    - The consumer *doesn't manage or control* underlying cloud infrastructure but has control over deployed applications and possibly configuration settings for the application-hosting environment.
  - **Infrastructure as a Service (IaaS)**
    The capability provided to the consumer (organizations) to provision processing, storage, networks, and other fundamental computing resources where the consumer able to deploy and run arbitrary software.
    - The consumer does not manage or control underlying cloud infrastructure, but has control over operating systems, storage, and deployed applications; and possibly limited control of select networking components (e.g., host firewalls).

- **Deployment Types:**
  - **Private Cloud**
    Cloud infrastructure provisioned for exclusive use by a single organization comprising multiple consumers (e.g., 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.

- **Public Cloud**

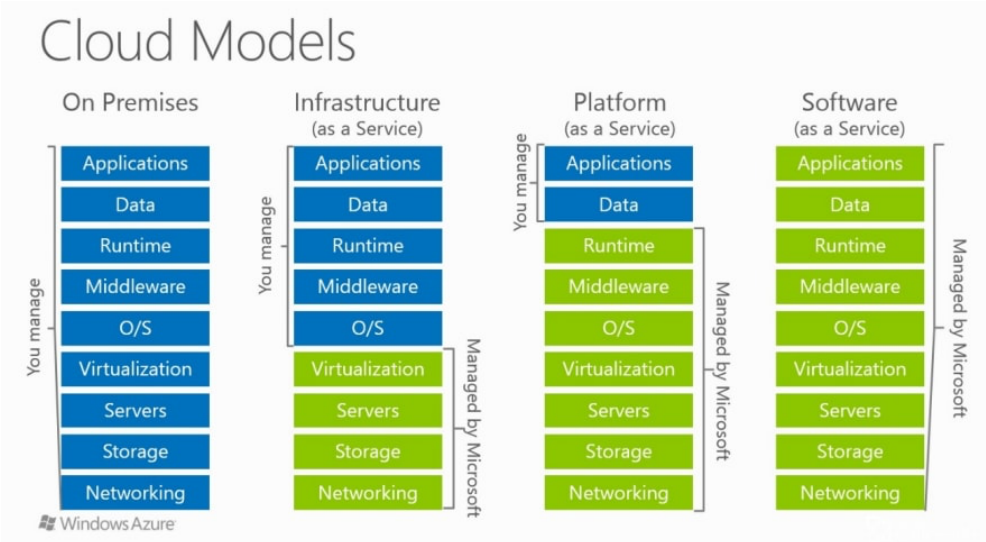
  Cloud infrastructure is provisioned for open use by the general public.
  - It may be owned, managed, and operated by a business, academic, or government organization, or some combination of them.
  - It exists on the premises of the cloud provider (i.e., Data Centers).

- **Community Cloud**

  It's provisioned for exclusive use by a specific community of consumers from organizations that have shared concerns (e.g., mission, security requirements, policy, and compliance considerations).
  - 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.

- **Hybrid Cloud**

  it's a composition of two or more distinct cloud infrastructures (private, community, or public) bound together by standardized or proprietary technology that enables data and application portability.
  - A cloud computing environment uses a mix of on-premises, private cloud and third-party, public cloud services with orchestration between their platforms.
  - Companies can run critical workloads or sensitive applications on the private cloud and use the public cloud to handle bursts or spikes in demand.




---

## *Multi-Cloud*

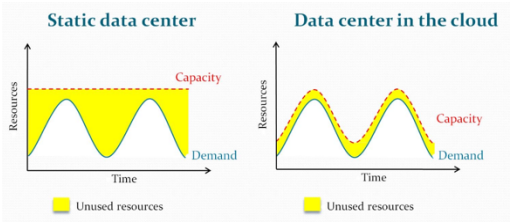
- It refers to the presence of more than one cloud deployment of the *same type* (Public or Private or Hybrid), sourced from different vendors.
- It refers to the presence of more than one cloud deployment of the same type (Public or Private or Hybrid), sourced from different vendors.
- A multi-Cloud strategy is an approach that operates any combination of Private, Public, and Hybrid clouds sourced from different vendors.
- An organization may have multiple Public and Private clouds or multiple Hybrid clouds, all either connected together or not.

## *Multi-Clouds Verses Hybrid Cloud*

| Compression | Hybrid Cloud | Multi-Cloud |
|---|---|---|
| Differences | An amalgamation of a private cloud with one or more public cloud. | An amalgamation of two or more public/hybrid cloud under centralized management |
| | It can be any combination of SaaS, IaaS, PaaS and any other as-a-service | A multi-cloud strategy offers the ability to select different cloud services from different providers |
| | It is a singular entity (ie., the cloud components are integrated to form singular entity) | It isn't single entity "individual clouds may not be integrated together" |
| | a hybrid cloud could be part of a multi-cloud deployment. | Enables organizations to locate IT resources closely to end users to achieve optimal performance and minimal latency |
| Similarity | Multi-cloud and hybrid cloud computing are similar, but different IT infrastructure model. | |

---

## *Cloud Benefits:*




1. **No up-front investment [Capital Expense (CapEx)]**
   - Pay-as you-go pricing model.
   - No need to invest in the infrastructure.
   - Resources are rented according to needs.
2. **Lowering operating cost [Operational Expense (OpEx)]**
   - Resources are allocated and de-allocated on demand.
   - No need to provide capacities according to peak load.
   - Resources can be released to save on operating costs when service demand is low.
3. **Highly scalable**
   - **Infrastructure providers:** provide pool large amount of resources from data centers and make them easily accessible.
   - **Service providers:** can easily expand its service to large scales to handle rapid increase in service demands.
4. **Easy access**
   - Services hosted in the cloud are generally web-based.
   - Accessible through devices with Internet connections.
   - Devices: desktop, laptop, cell phones and PDAs.
5. **Reducing business risks and maintenance expenses (OpEx):**
   - Outsource the service infrastructure to the cloud.
   - ***Service providers*** shift business risks (such as hardware failures) to ***infrastructure providers.***
   - ***Infrastructure providers*** often have better expertise and are better equipped for managing these risks.
   - A ***service provider*** can cut down the hardware maintenance and the staff training cost.
6. **Elasticity**
   Degree to which a system can adapt to workload changes by provisioning and deprovisioning resources in an autonomic manner.
   - such that at each point in time the available resources match the current demand as closely as possible to avoid *Over- Provisioning* or *Under-Provisioning.*
   - *Over-Provisioning:* Allocating more resources than required.
   - *Under-Provisioning:* Allocating fewer resources than required.

7. **Energy Savings**
   Energy efficiency benefits of cloud computing are substantial, and growth in the market will have important implications for both energy consumption and Green House Gas [GHG] emissions.

## *Advantages of Cloud:*

- Lower computer costs
- Improved performance
- Reduced software costs
- Instant software updates
- Unlimited storage capacity
- Increased data reliability
- Universal document access
- Latest version availability
- Easier group collaboration
- Device independence

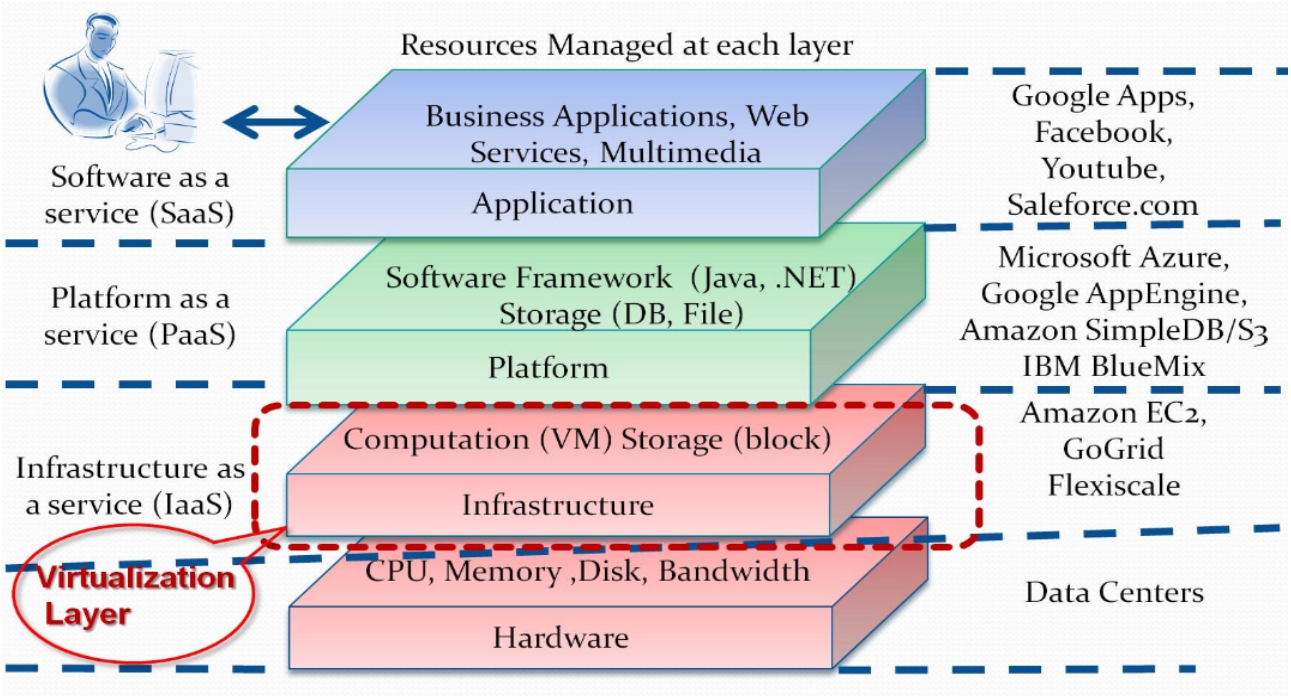
## *Disadvantages of Cloud:*

- *Requires a constant Internet connection:* Doesn't work well with low-speed connections.
- *Features might be limited:* Can be slow, Stored data can be lost, Stored data might not be secure

## *Cloud Computing Services*

| | Who Uses it | What Services are available | Why use it? |
|---|---|---|---|
| **SaaS** | Business Users | Email, Office, CRM, Blogs. | To complete business tasks. |
| **PaaS** | Developers | Services and Applications test, development, integrations and deployment environment. | Create and deploy service for users |
| **IaaS** | System Manager | Virtual Machines, Operations Systems, Massaging Queue, Networks, Storage, CPU, Memory backup. Services. | Create platforms for services and applications test, development, integration and deployment |

## *Cloud Computing Hierarchy*



1. **Application Layer:**
   - Highest level of the hierarchy
   - Consists of the actual cloud applications
   - Cloud applications can better affect the automatic-scaling feature to achieve better performance, availability, and lower operating cost.

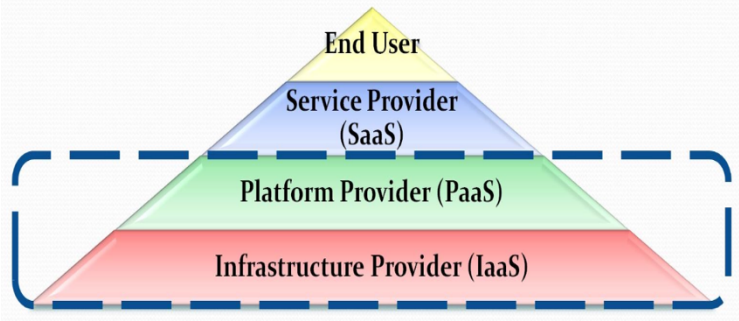
2. **Platform Layer**
   - Built on top of the infrastructure layer.
   - Consists of OSs and application frameworks by providing Virtual Machines with specific configuration.
   - Minimize the burden of deploying applications directly into VM containers.
   - E.g., Google App Engine provides API support for implementing storage, database, and business logic.

3. **Infrastructure Layer *(aka: Virtualization Layer)***
   - Creates a pool of storage and computing resources by partitioning the physical resources into instants using virtualization technologies such as Xen, KVM, and VMware.
   - An essential component of cloud computing, since many key features, such as dynamic resource assignment, are only made available through virtualization technologies.

4. **Hardware Layer:**
   - Typically implemented in data centers
   - Responsible for ***managing physical resources of the cloud:***
     - Physical servers, routers, switches, power cooling systems.
   - Typical issues at hardware layer include:
     - Hardware configuration, Fault tolerance, Traffic and resource management.

## *Cloud Computing Providers*

- IaaS and PaaS providers are often parts of the same organization.
- PaaS and IaaS providers are often called the infrastructure providers or cloud providers.




## *Some Commercial Cloud*

- **Windows Azure (PaaS):** Hosting service for .NET applications and database.

- **Google App Engine [GAE] (PaaS):** Automatic scaling and reliability at the price of a highly constrained application structure (3-tier Web application).

- **BlueMix (PaaS):** Provides a *dashboard to create, view, and manage applications* and services, as well as monitor application's resource usage.

- **Amazon Web Services (IaaS):**
  - includes Scalable Storage Service (S3), Elastic Computing Cloud (EC2), Elastic Block Store (EBS)
  - Comes with many secondary tools: e-commerce, Content- Distribution Network (CDN), etc.

***Cloud Computing Challenges:***

- **Performance**
  - Data-intensive applications.
  - Latency and delays for users far from cloud providers.
- **Security and Privacy**
  - Companies are still concerned about security.
  - Customers are worried about vulnerability to attacks.
- **Control**
  - Providers have control of platforms.
  - No platforms for specific users &their business practices.
- **Bandwidth Costs**
  - Low for smaller Internet-based applications.
  - Significantly grow for data-intensive applications.
- **Reliability**
  - Not always offer round-the-clock reliability

***Service Level Agreement (SLA):*** A contract between the user and the Provider.

- **SLA** contains terms and conditions to ensure the rights of the users, as well as the providers.
- Its role is identifying the user's needs and creates a relationship between the user and the service provider.
  - *User Needs:* Guarantee the resources availability, reliability, and QoS.
  - *Provider needs:* Guarantee the resource utilization and revenue.
- SLA describes a set of non-functional requirements of the service the customer is rending.

***Service Level Agreement Metrics***