

Cloud Computing Fundamentals

**Prepared By:
Dr. Fatma Omara**

Books:

Hand-Book of Cloud: Chapters 1-3

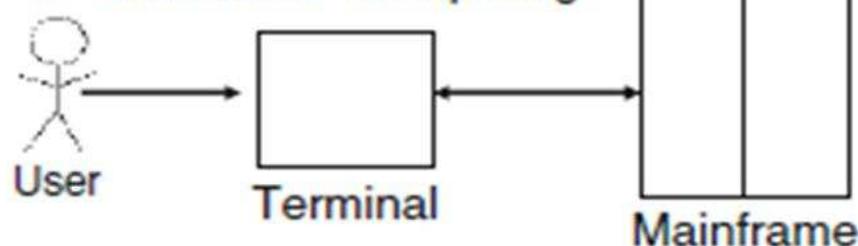
CLOUD COMPUTING Principles and Paradigms: Chapter 1

Other Materials

Computing Generations

Phases

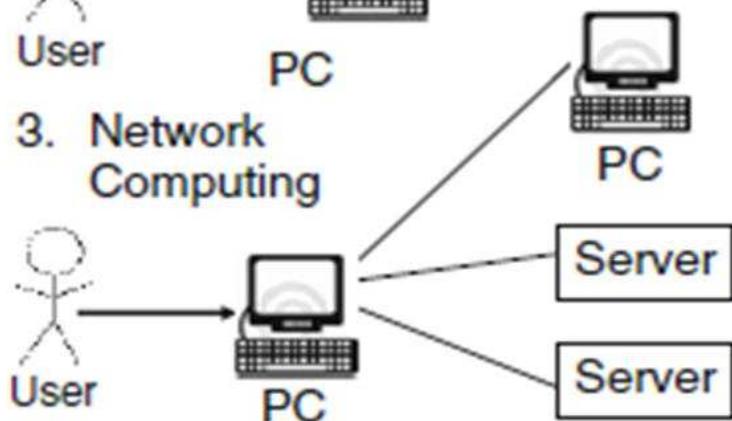
1. Mainframe Computing



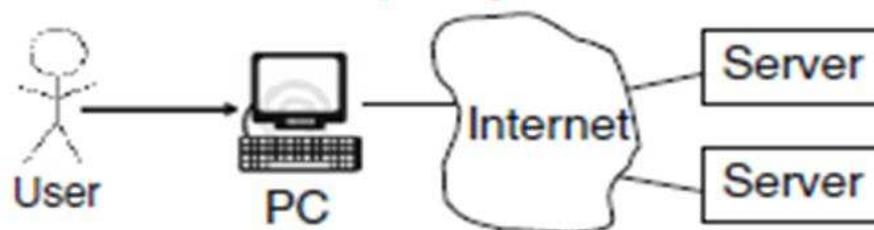
2. PC Computing



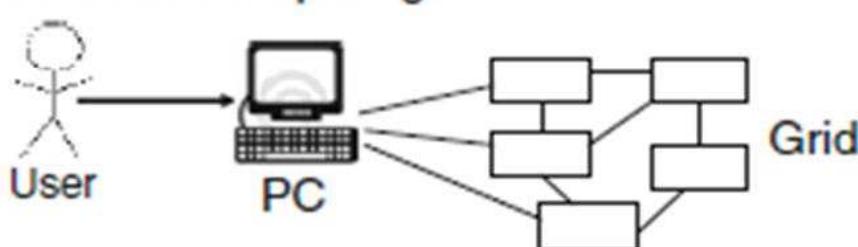
3. Network Computing



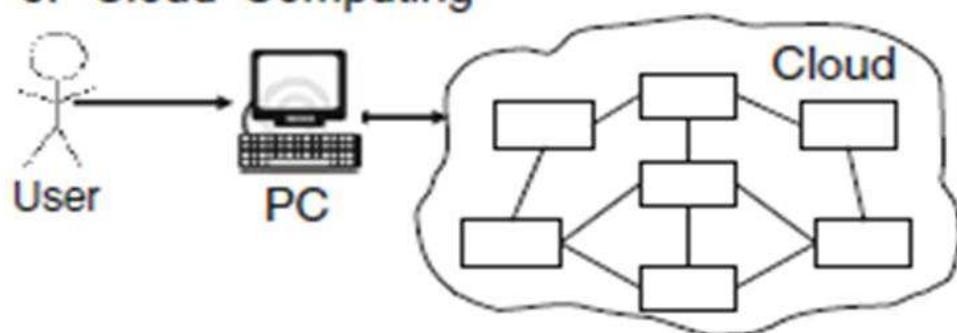
4. Internet Computing



5. Grid Computing

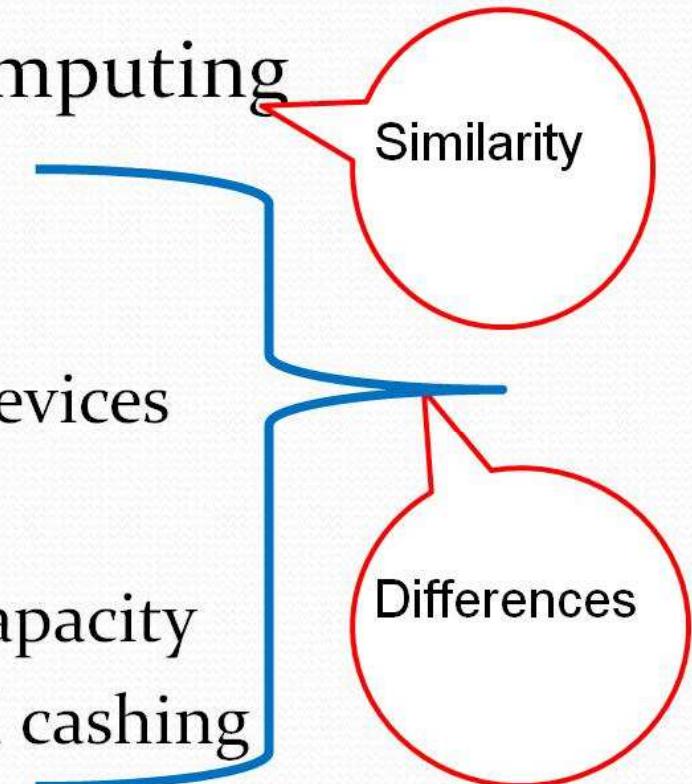


6. Cloud Computing



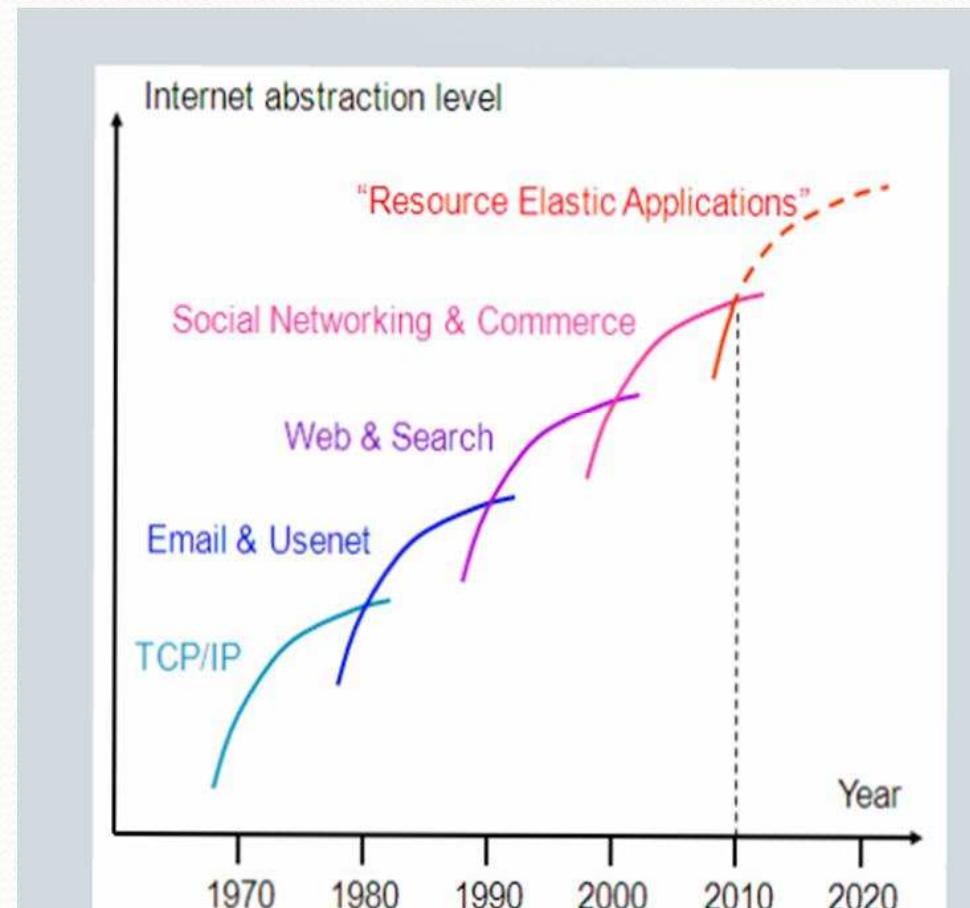
Mainframe Vs. Cloud Computing

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



History of The Internet Revolution

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



Technologies Created Storm of Cloud Computing

- Broadband Internet
- Web services
- Computing systems
- Application software.
- Network
 - Center of the cloud delivery and consumption.
 - The linkage between end users and provider's data centers.
- Data center network
 - Connects thousands of compute and storage nodes.

Cloud Computing Definition

- ❑ 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]



Cloud Computing Definition

The NIST Definition of Cloud Computing

- Cloud computing is 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*.



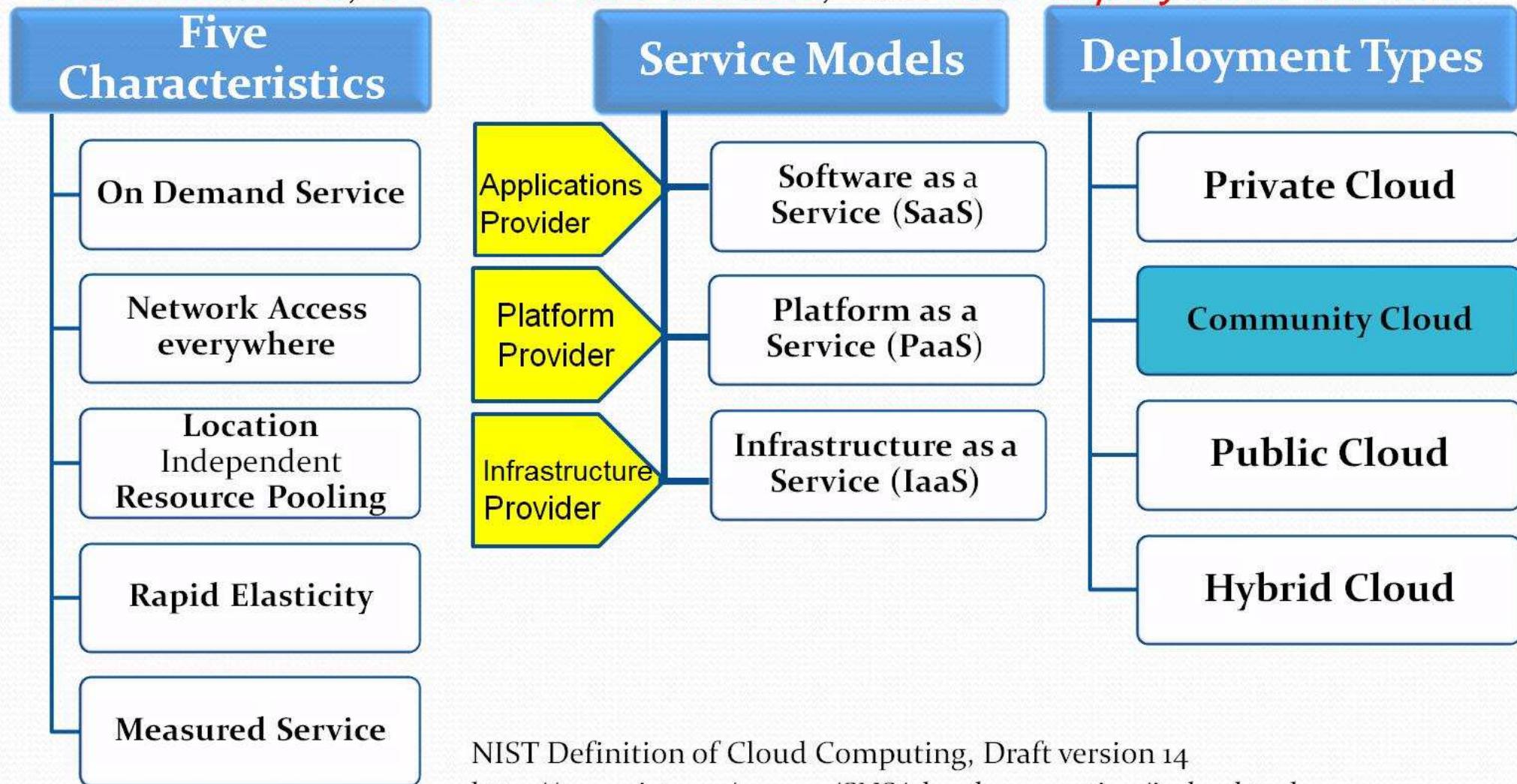
NIST: National Institute of Standards and Technology

<http://nvlpubs.nist.gov/nistpubs/Legacy/SP/nistspecialpublication800-145.pdf>

Cloud Computing Definition

NIST definition:

cloud model promotes availability and is composed of **five essential characteristics**, **three service models**, and **four deployment models**.



NIST Definition of Cloud Computing, Draft version 14

<http://csrc.nist.gov/groups/SNS/cloud-computing/index.html>

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

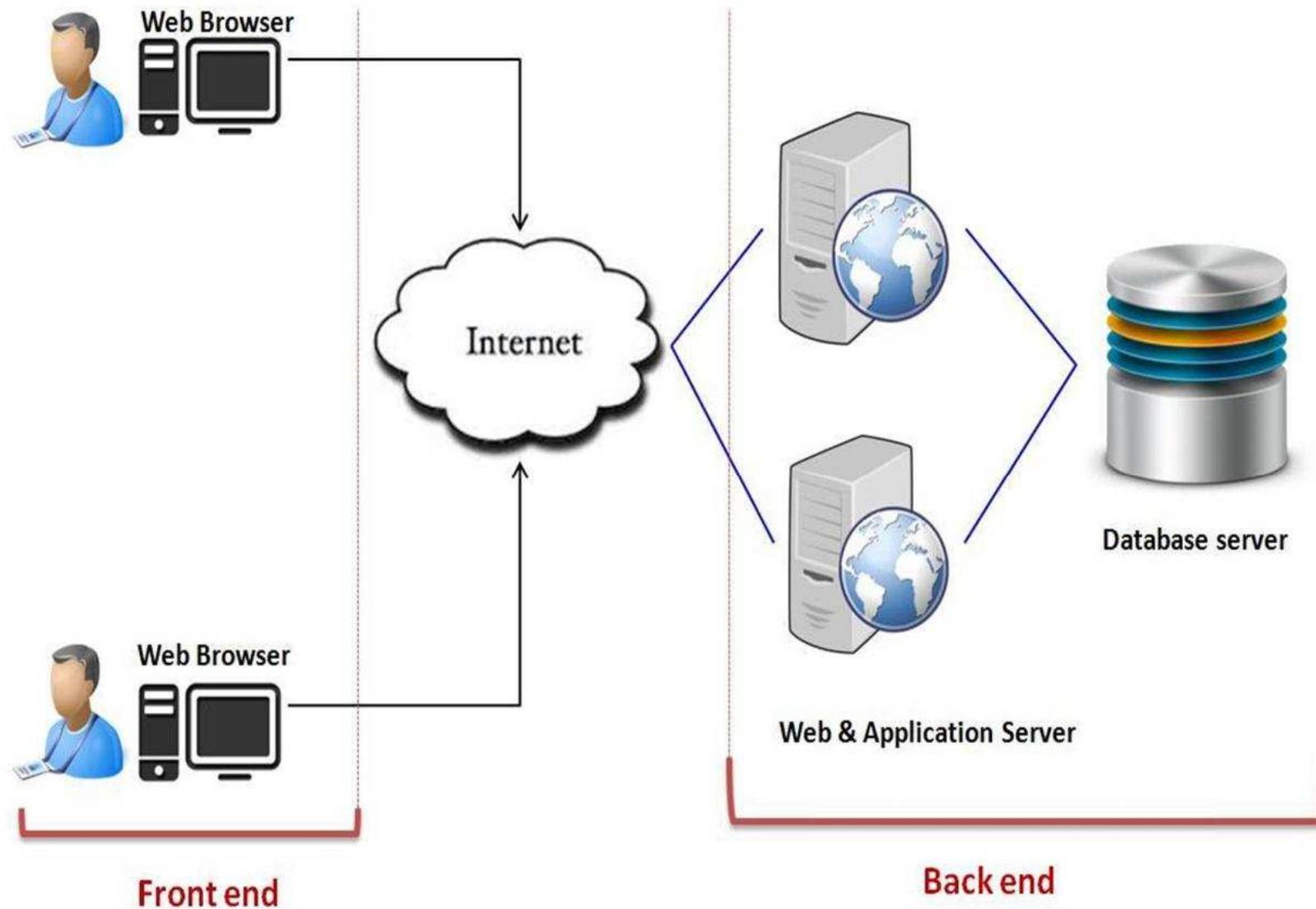
Computing resources are pooled to serve multiple consumers according to their demand.

Rapid Elasticity

Capabilities can *elastically provisioned and released*, in some cases automatically, to scale up and down according to consuming rate.

Measured Service

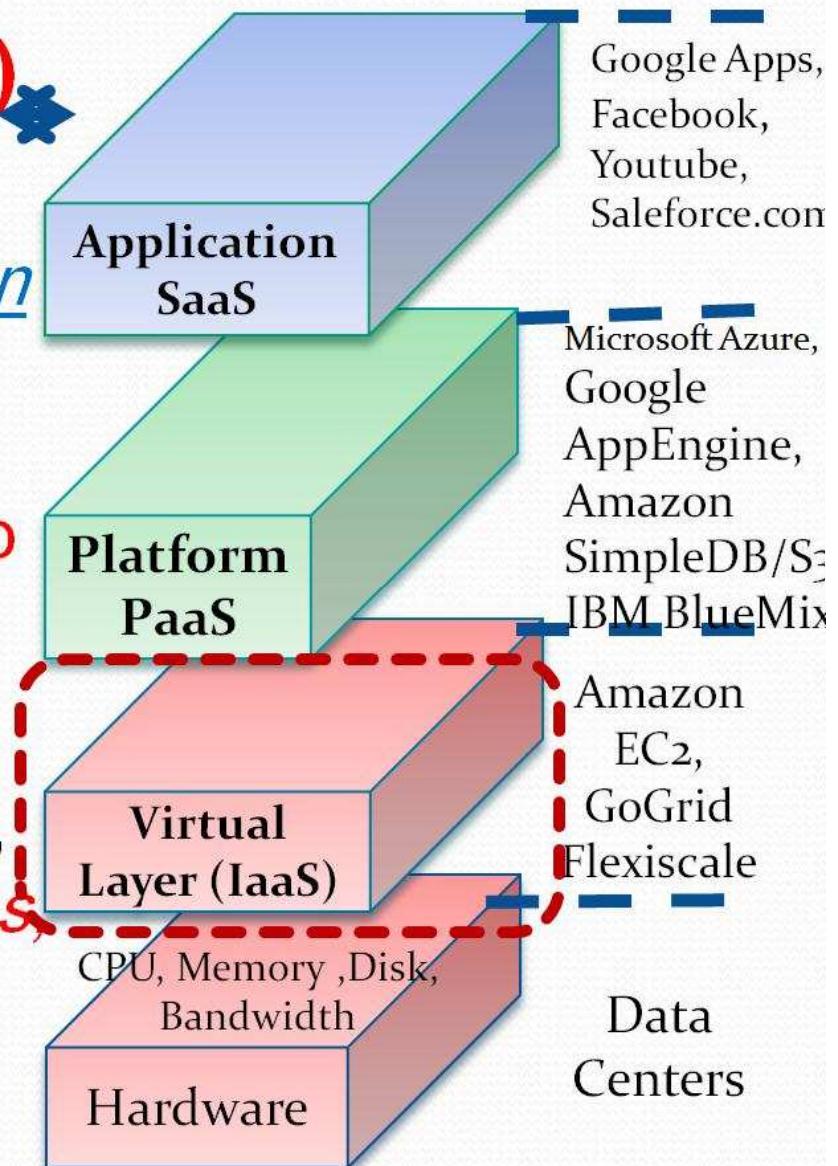
Resource usage can monitor, control, and report, providing transparency for both provider and consumer of the utilized service.



Service Models

Infrastructure as a Service (IaaS) *

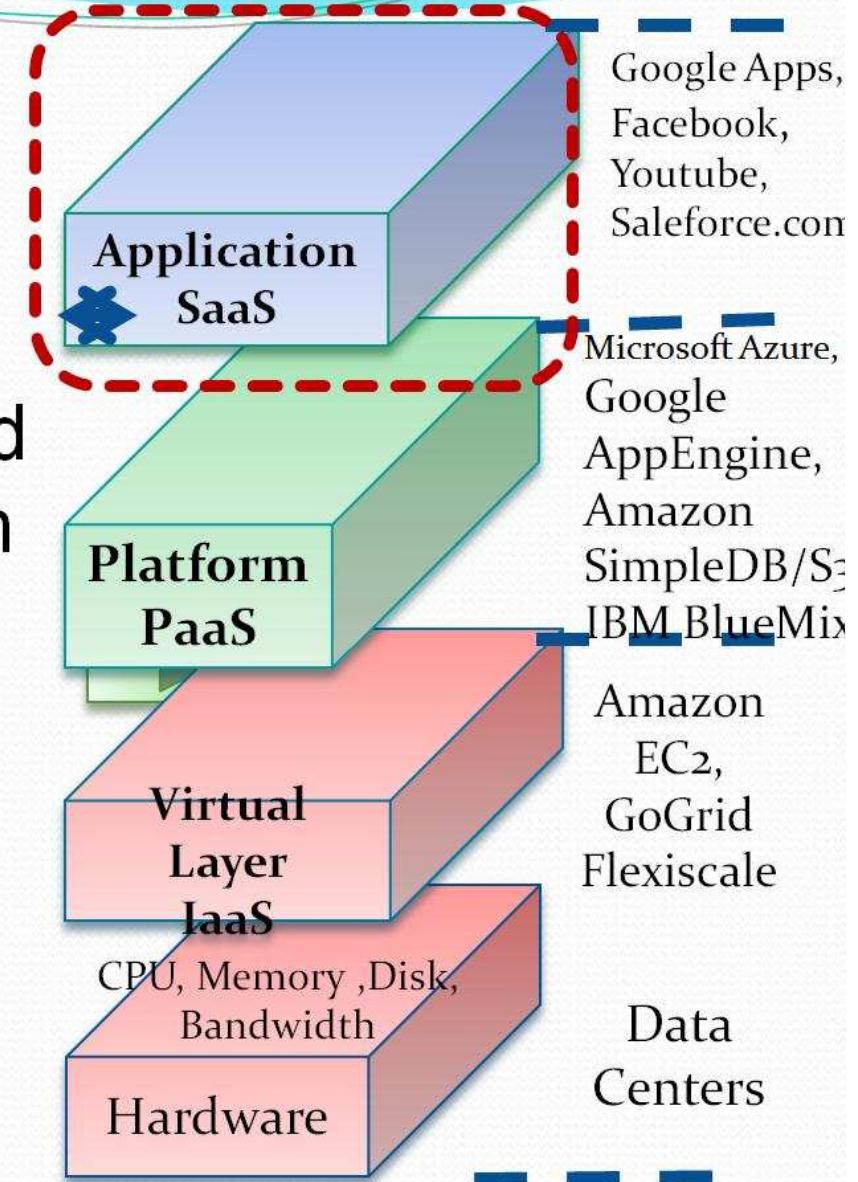
- The capability provided to the consumer (*organizations*) to provision processing, storage, networks, and other fundamental computing resources where the consumer able to deploy and run arbitrary software.
- The *consumer* does not manage or control underlying cloud infrastructure, but has *control over operating systems, storage, and deployed applications; and possibly limited control of select networking components* (e.g., host firewalls).



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.



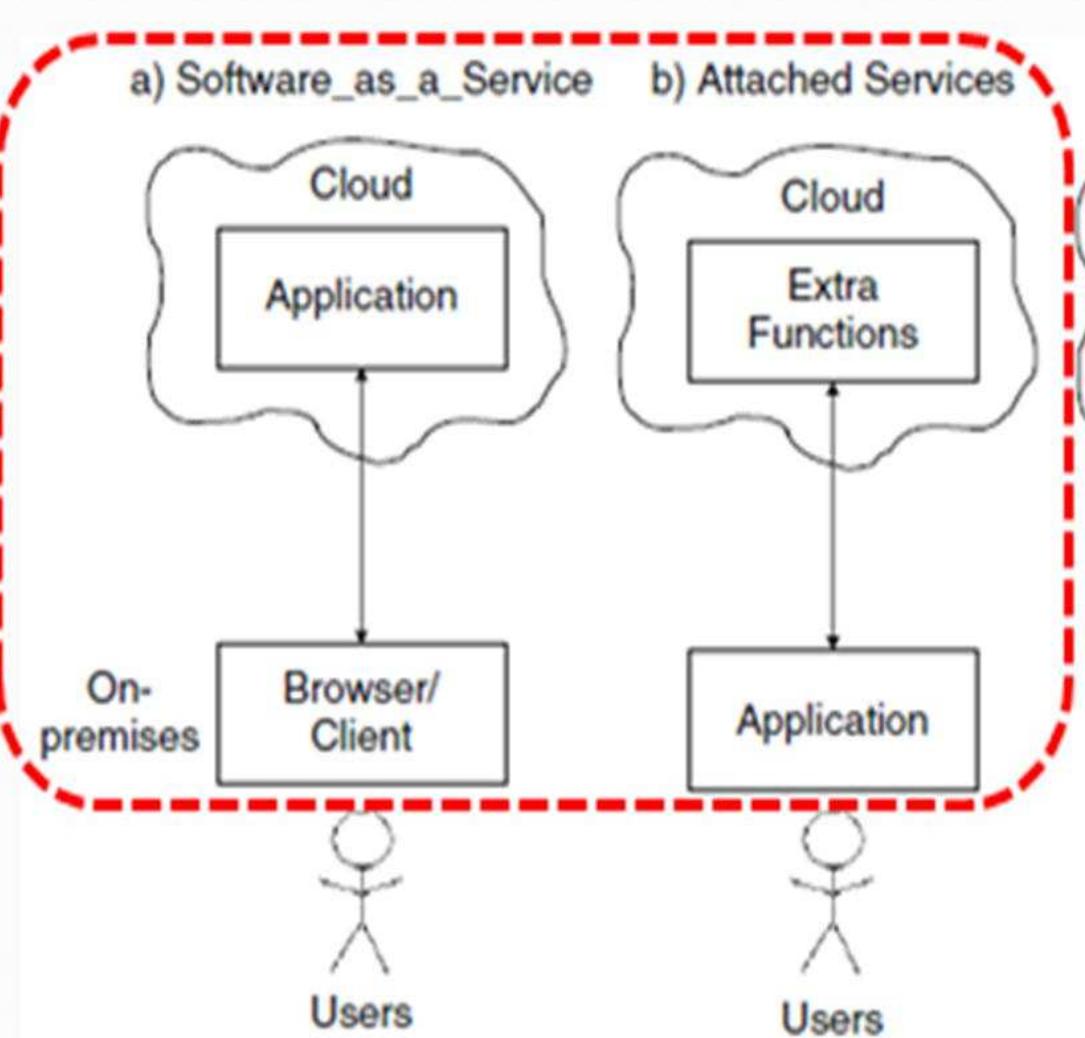
Software as a Service (SaaS)-Cont..

SaaS

- Entire application is running in the cloud
- Client contains a browser to access application
- Example: salesforce.com

Attached services

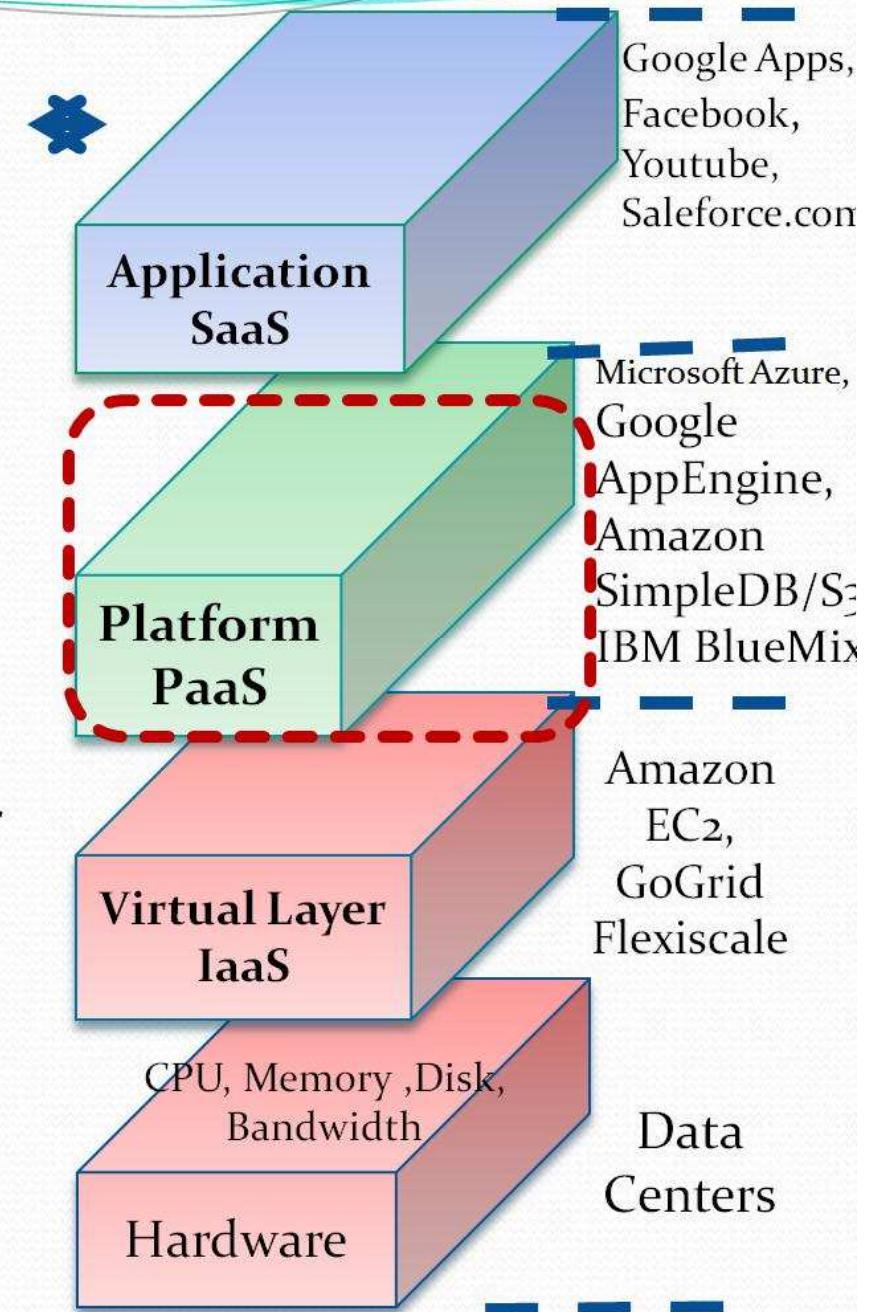
- Application runs on the client premises
- Accesses functions and services provided in cloud
- Example: Apple's iTunes



Service Models

Platform as a Service (PaaS)

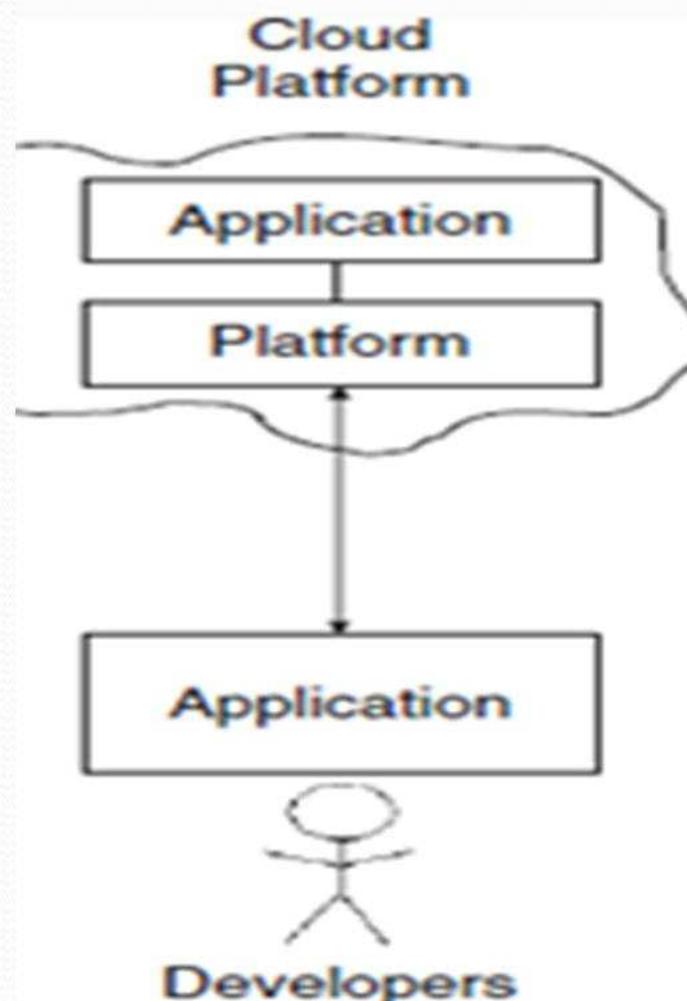
- The capability provided to the consumer (*developer*) by needed platform environment to develop his application using *programming languages, libraries, services, and tools* supported by the provider.
- The consumer **does not manage or control** underlying cloud infrastructure, but has control over deployed applications and possibly configuration settings for the application-hosting environment.



Platform as a Service (PaaS)

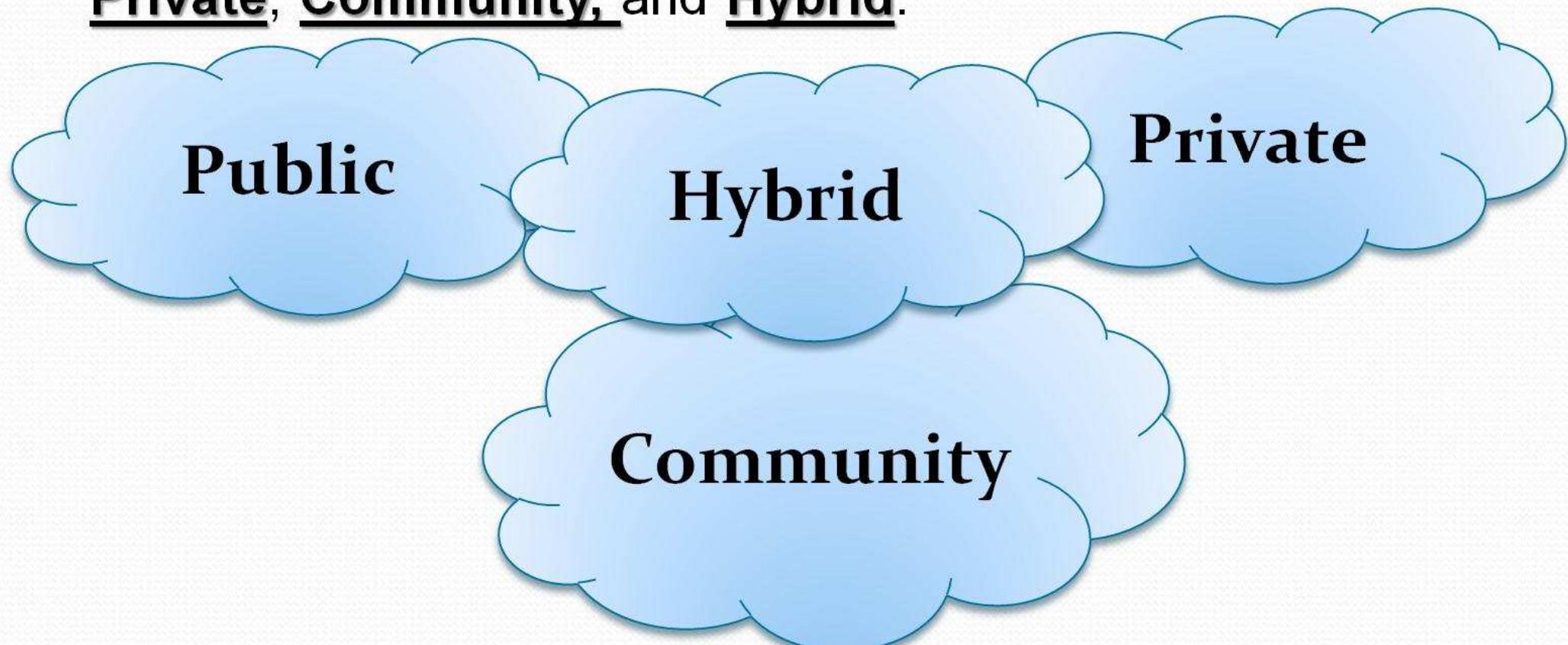
Cloud Platform

- Used by developers for creating a SaaS applications



Deployment Types

- 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, Community, and Hybrid.



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.



PRIVATE CLOUD

A private cloud is typically defined as everything behind a company's walls. These kinds of systems operate in a company's local data centers, although some companies prefer to use colocated data center facilities.

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).



PUBLIC CLOUD

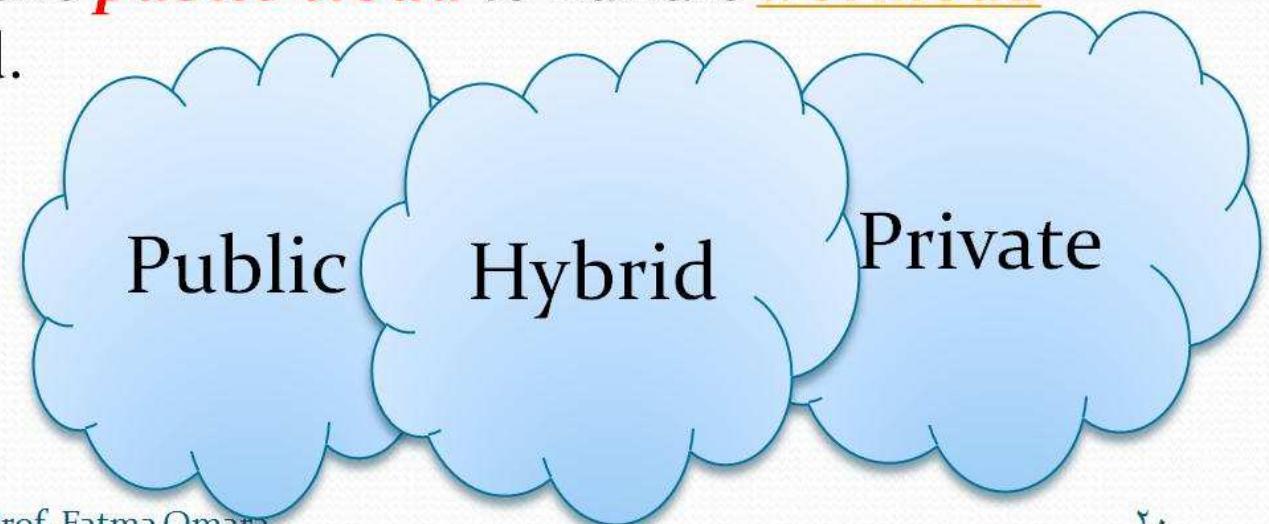
The public cloud includes a whole host of services and companies. The most common names are AWS and Microsoft Azure, among others. However, you may also hear software as a service, such as Microsoft Office 365, included in the definition.

Community Cloud

- Cloud infrastructure is provisioned for exclusive use by **a specific community of consumers from organizations** that have **shared concerns**
 - (e.g., mission, security requirements, policy, and other 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.

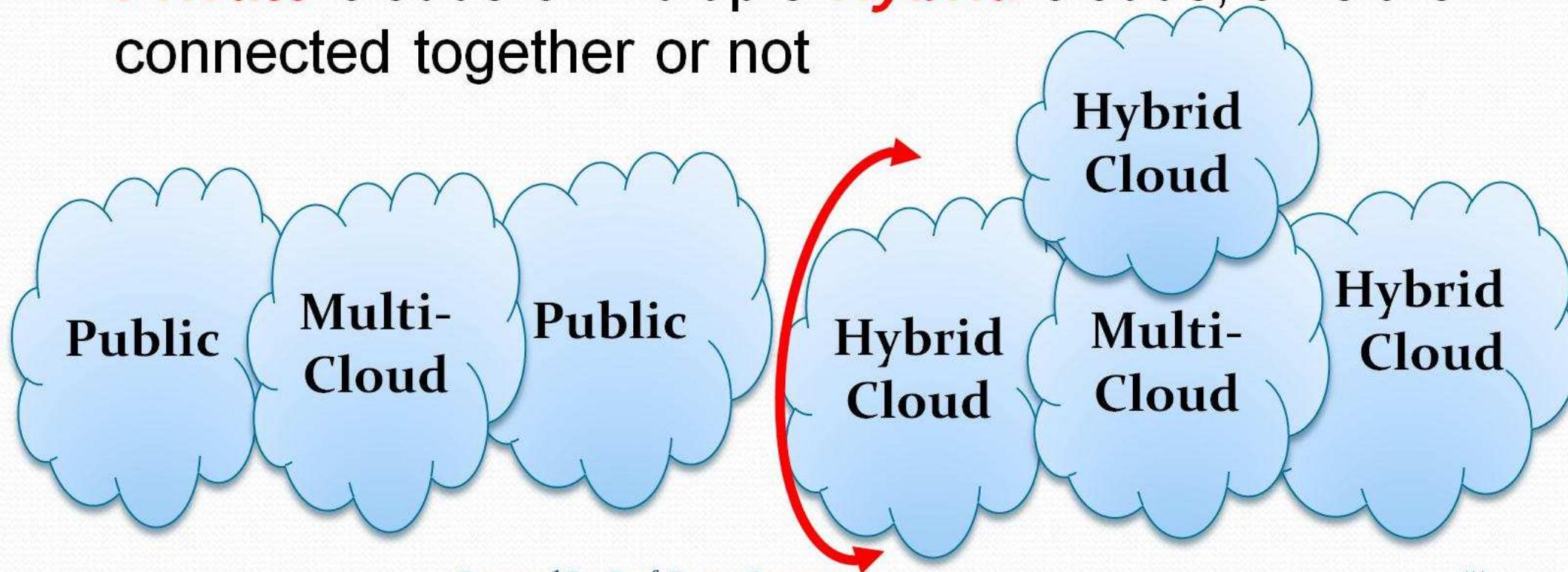
Hybrid Cloud

- Cloud infrastructure is a composition ***of two or more distinct cloud infrastructures*** (private, community, or public) ***bound together by standardized or proprietary {p(r)e 'prīə, terē} technology*** that enables ***data and application portability*** (e.g., cloud bursting for load balancing between clouds).
- 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 **workload bursts** or spikes in demand.



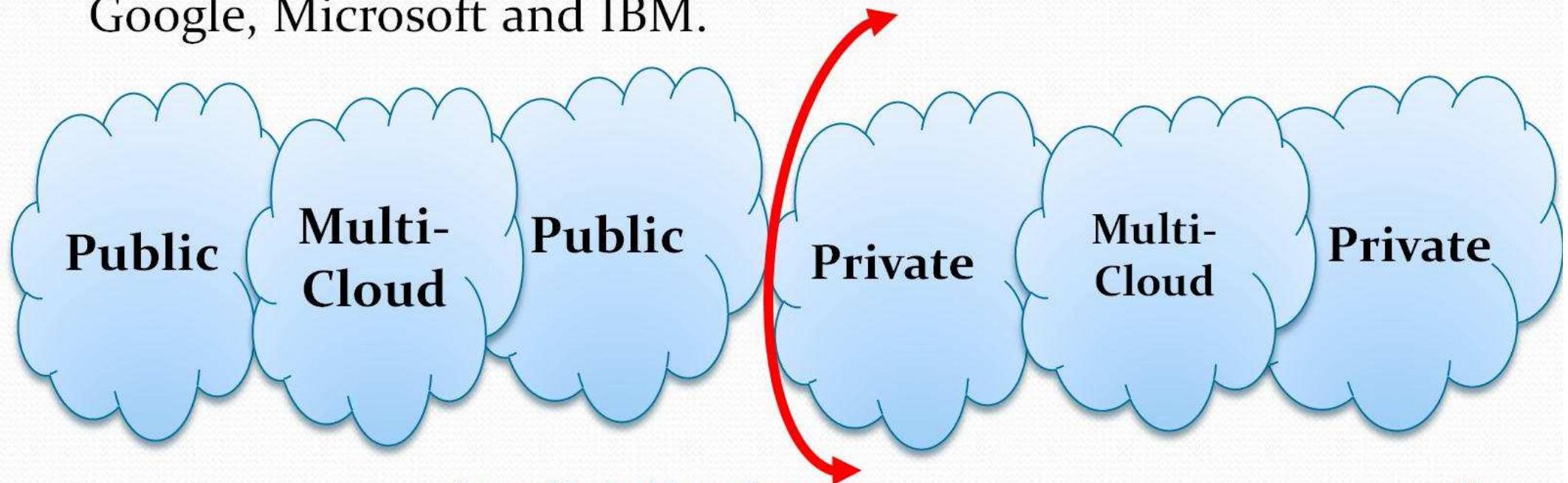
Multi-Cloud

- 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-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.
- There is no single multi-cloud infrastructure vendor. Instead, a multi-cloud strategy involves a mix of the major public cloud providers, such as Amazon Web Services (AWS), Google, Microsoft and IBM.



Multi-Cloud Challenges

- ❖ Clouds must have **compatible interfaces** to **migrate workloads** between them
- ❖ Clouds with **different APIs** can **not able to talk** to each other.
- ❖ Using cloud **storage services** in one cloud might not be available in another, which can make it impossible to move workloads easily between clouds.
- ❖ You might find yourself working from several **different management windows**, which increases chances of **getting lost** while configuring the environment.
- ❖ Every cloud comes with its own **monitoring tools**. It is difficult to monitor the output of different monitoring tools.
- ❖ Managing every cloud as a separate entity, it's difficult to avoid application ***stretch using services from different cloud***.
- ❖ When building a multi-cloud environment, it should make sure the clouds are **highly compatible** to simplify **interoperability**.

Approaches helps multi-cloud environment to run smoothly

1. Standardize as much as possible

- A standardized approach makes multi-cloud environments run smoother. At least, it ensures workloads can be easily migrated among different clouds.

2. Consolidate Management

- A single multi-cloud management tool can solve challenges ranging from multiple management tools and multiple management portals to dealing with application and VM stretching over different clouds.
- Providing a single interface to manage the multi-cloud environment and make it look as working with one cloud.

3. Monitor the results

- An independent monitoring tool can be helpful.

Multi-Cloud vs. Hybrid Cloud

Hybrid Cloud

- An amalgamation of a private cloud with one or more public cloud.
- It can be any **combination** of SaaS, IaaS, PaaS and any other as-a-service
- It is a **singular entity** (i.e., the cloud components are integrated to form singular entity)
- a hybrid cloud could be part of a multi-cloud deployment.

Multi-Cloud

- An amalgamation of two or more public/hybrid cloud under centralized management
- A multi-cloud strategy offers the **ability to select different cloud services from different providers**
- It **isn't single entity** “individual clouds may not be integrated together”
- Enables organizations to locate IT resources closely to end users to achieve **optimal performance and minimal latency**.

Multi-cloud and hybrid cloud computing are similar, but different IT infrastructure models

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.

Cloud Benefits (Cont.)

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.

Cloud Benefits (Cont.)

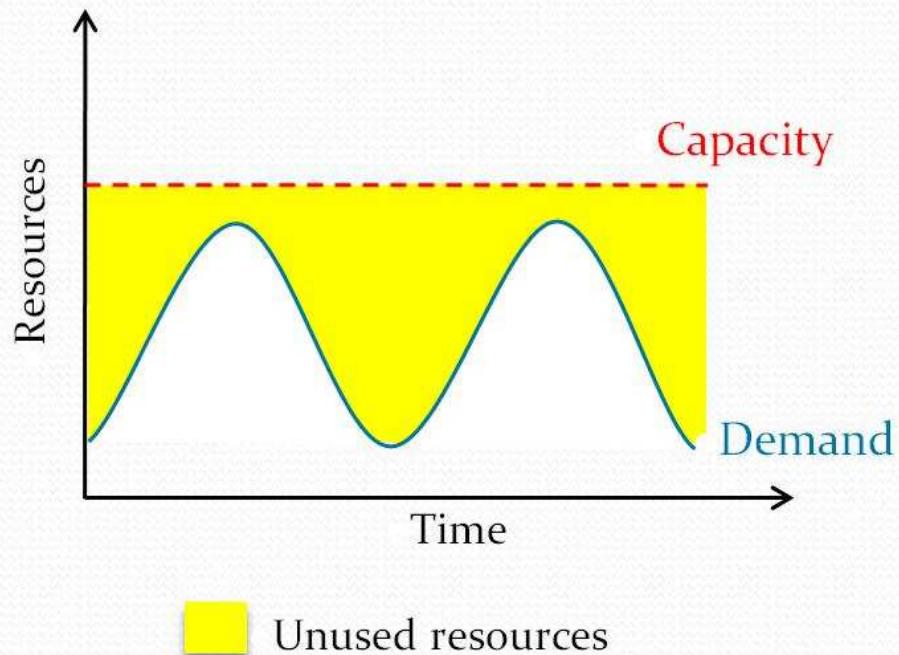
5) Reducing business risks and maintenance expenses (**OpEx**)

- Outsource the service infrastructure to the cloud
- Service providers shift business risks (such as hardware failures) to infrastructure providers
- A service provider can cut down the hardware maintenance and the staff training costs.
- Infrastructure providers often have better expertise and are better equipped for managing these risks.

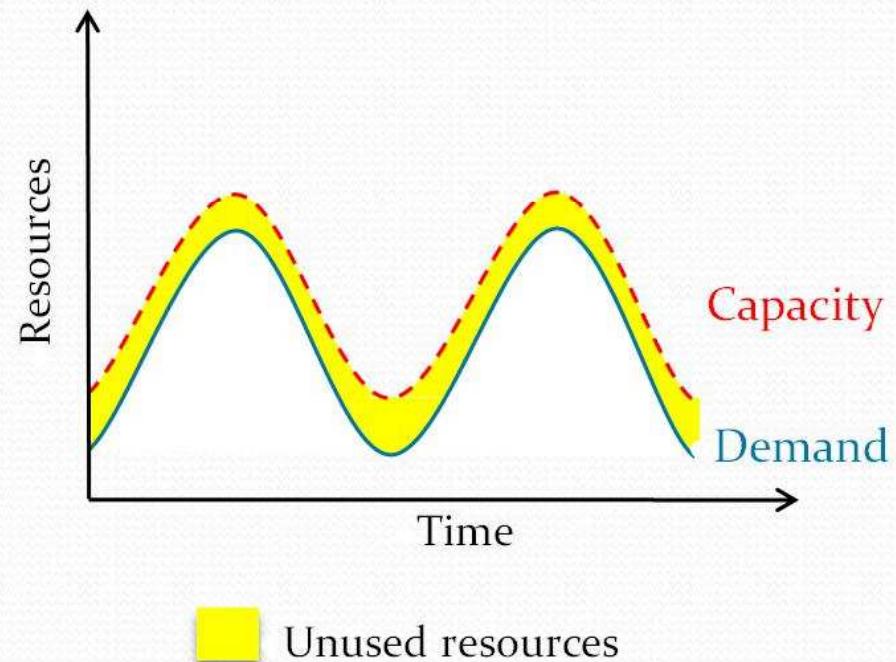
Cloud Benefits (Cont.)

Pay by use instead of provisioning for peak
(Elasticity)

Static data center



Data center in the cloud



Slide Credits: Berkeley RAD Lab

Cloud Benefits (Cont..)

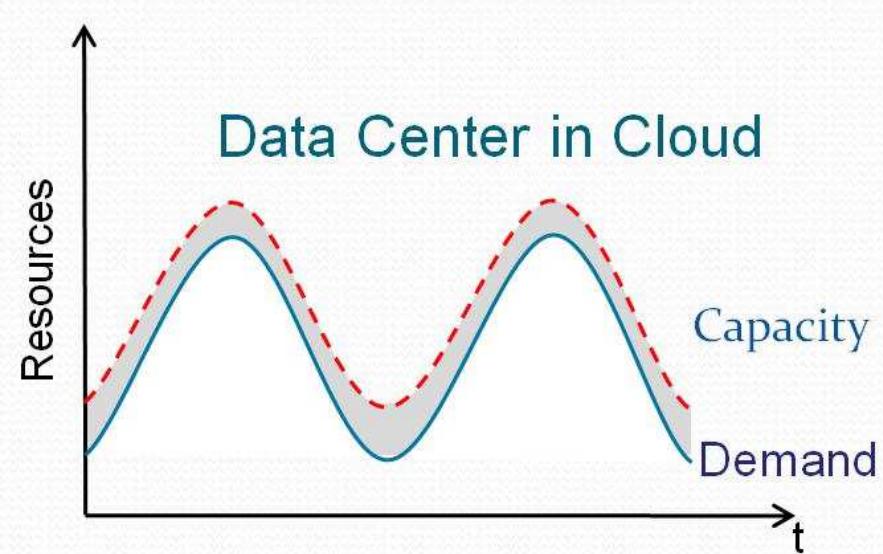
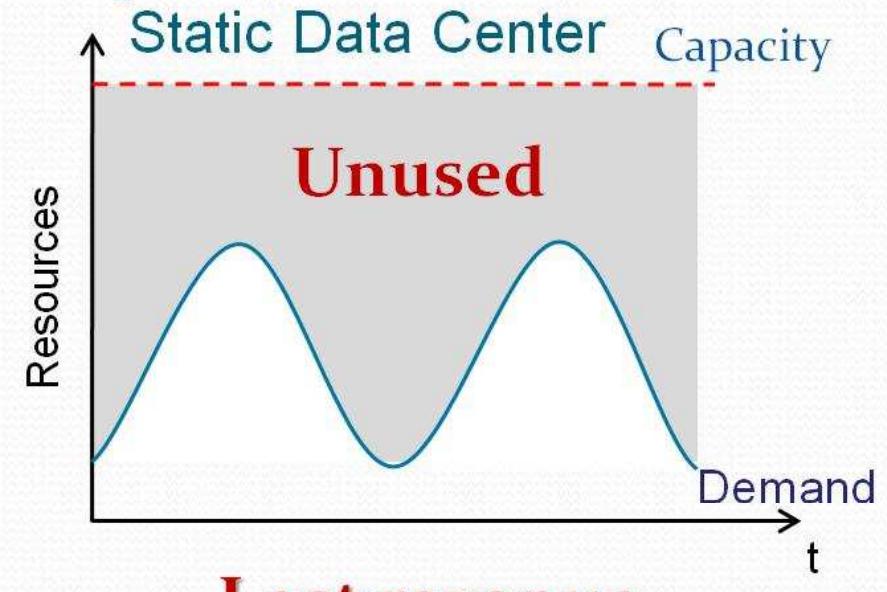
6) Elasticity

- Degree to which a system is able to adapt to workload changes by provisioning and de-provisioning 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- or Under-Provisioning.*

Cloud Benefits (Cont.)

Over-Provisioning

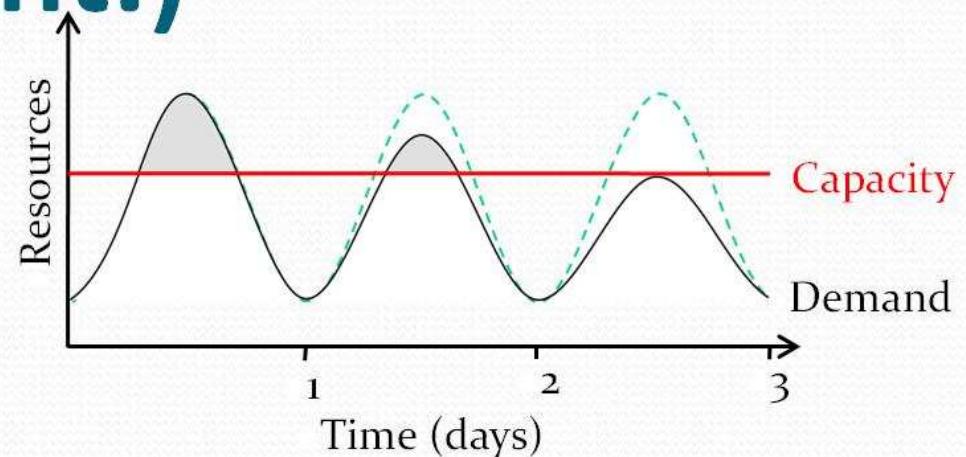
- Allocating more resources than required
- It should be avoided as the service provider often has to pay for the resources that are allocated to the service (Lost revenue)



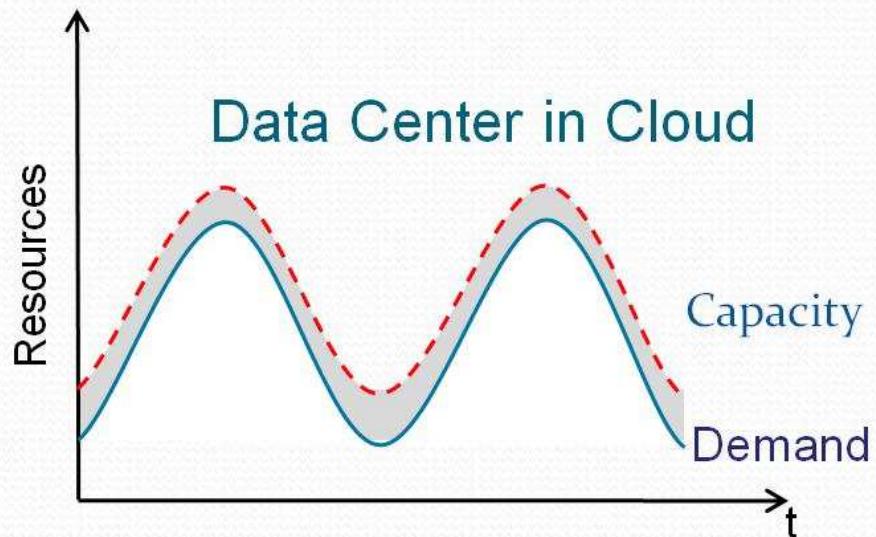
Cloud Benefits (Cont.)

Under-Provisioning

- Allocating fewer resources than required due to:
 - People don't understand how virtualization works
 - They don't realize how much of a computing resource they're really buying
- This must be avoided; otherwise the service cannot serve its users with a good service (**Lost Users**).



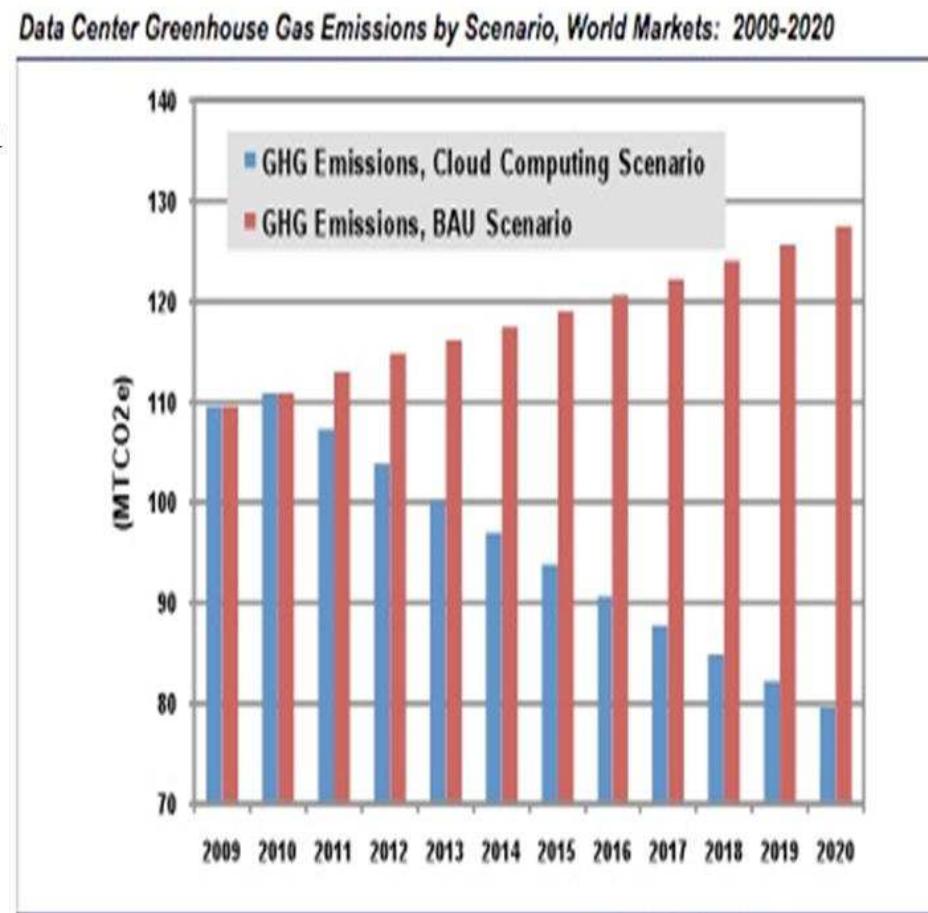
Lost users



Cloud Benefits (Cont.)

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**”.
- The adoption of cloud computing leads to a **38% reduction** in worldwide data center energy expenditures by **2020**.
- Total data center energy expenditures down from **\$23.3 billion** in **2010** to **\$16.0 billion** in **2020**.



(Source: Pike Research)

BAU: Business as Usual

[<http://cloudnumbers.com/energy-savings-by-cloud-computing>, September 8th, 2011]

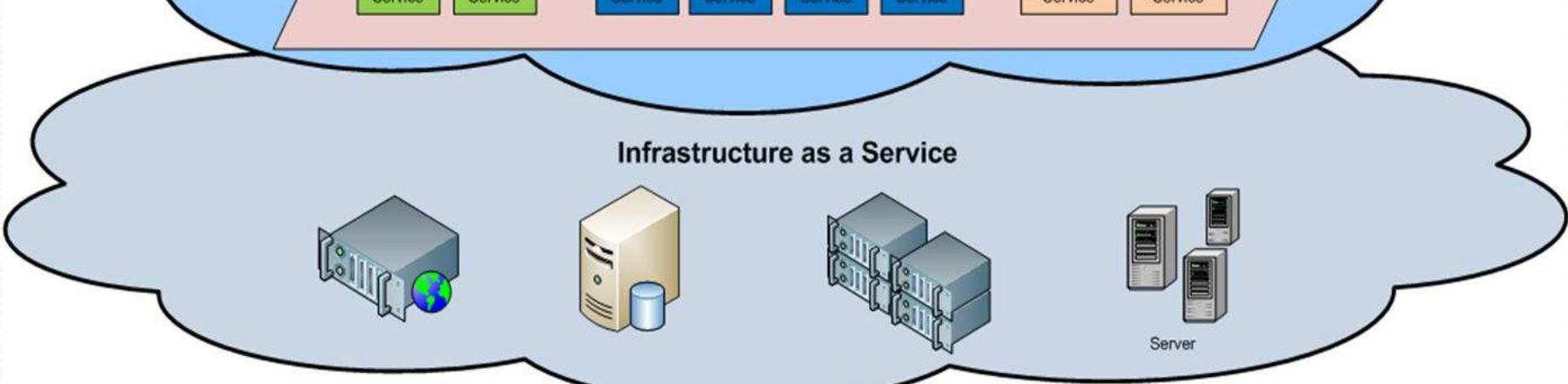
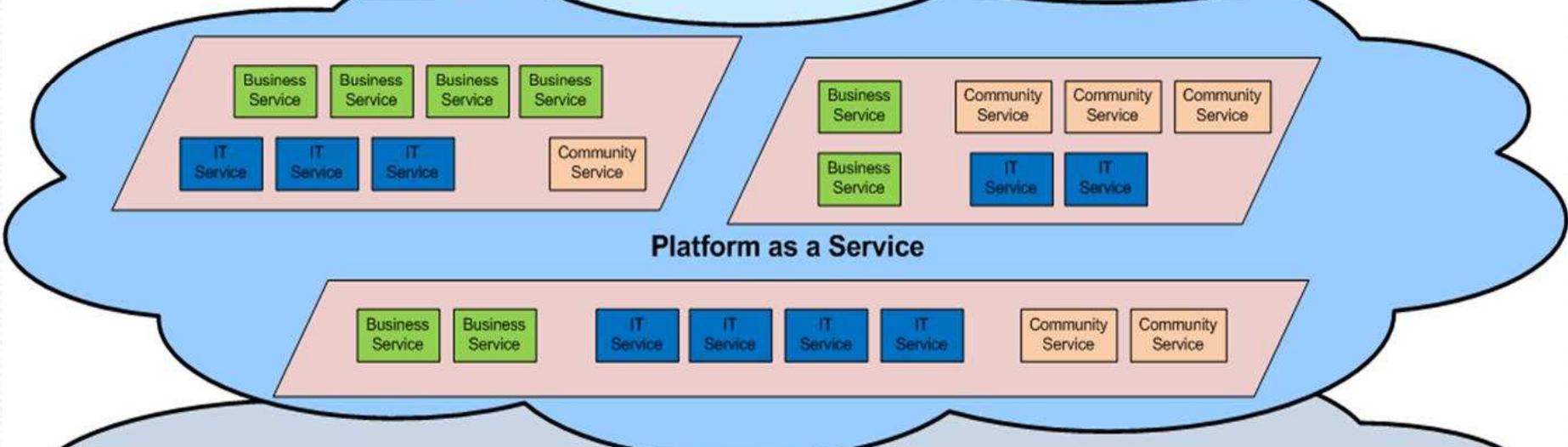
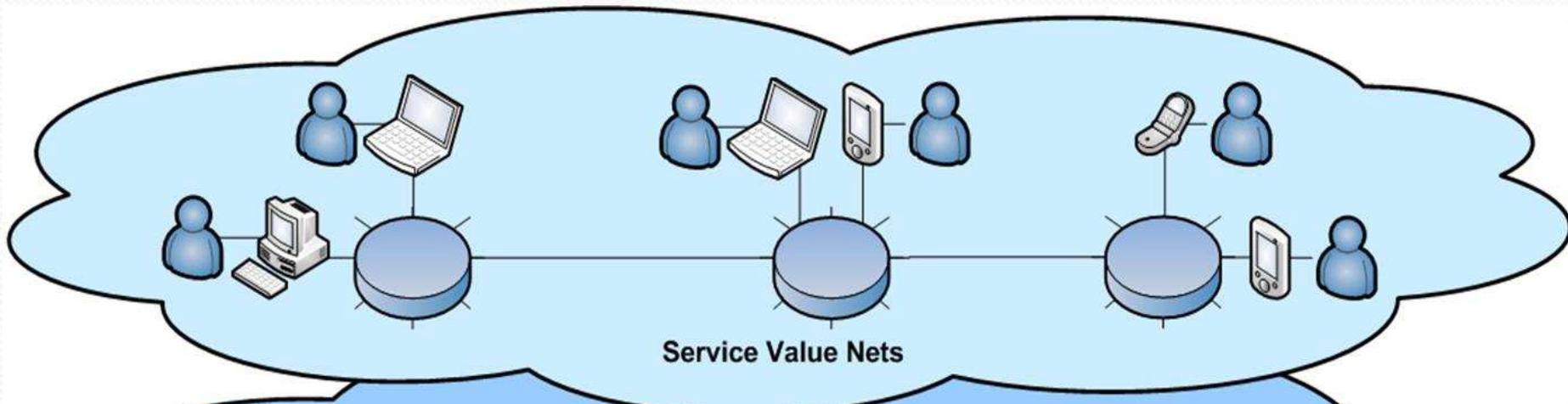
Advantages of Cloud

- 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

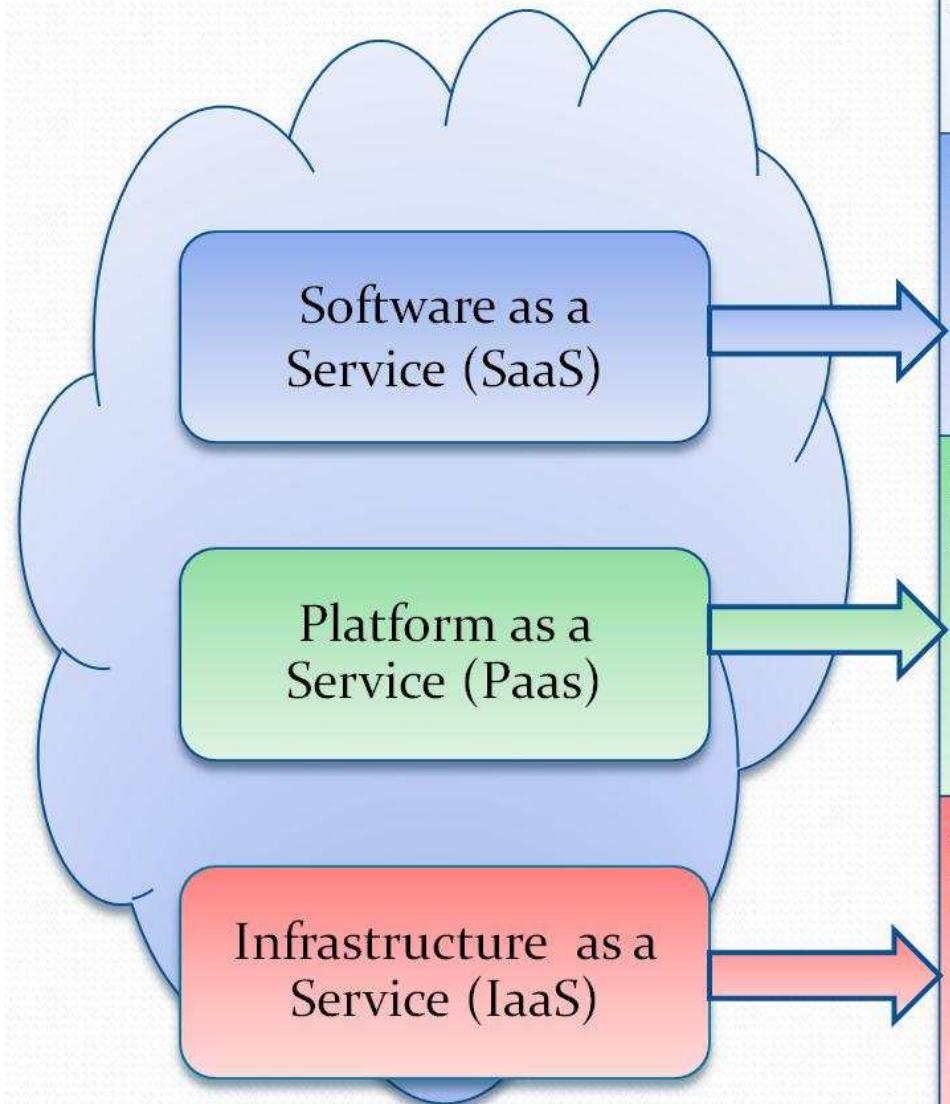
- Requires *a constant Internet connection*
 - Does not work well with low-speed connections
- Features might be limited
 - Can be slow
 - Stored data can be lost
 - Stored data might not be secure
- Large Latency Time

Cloud Architecture



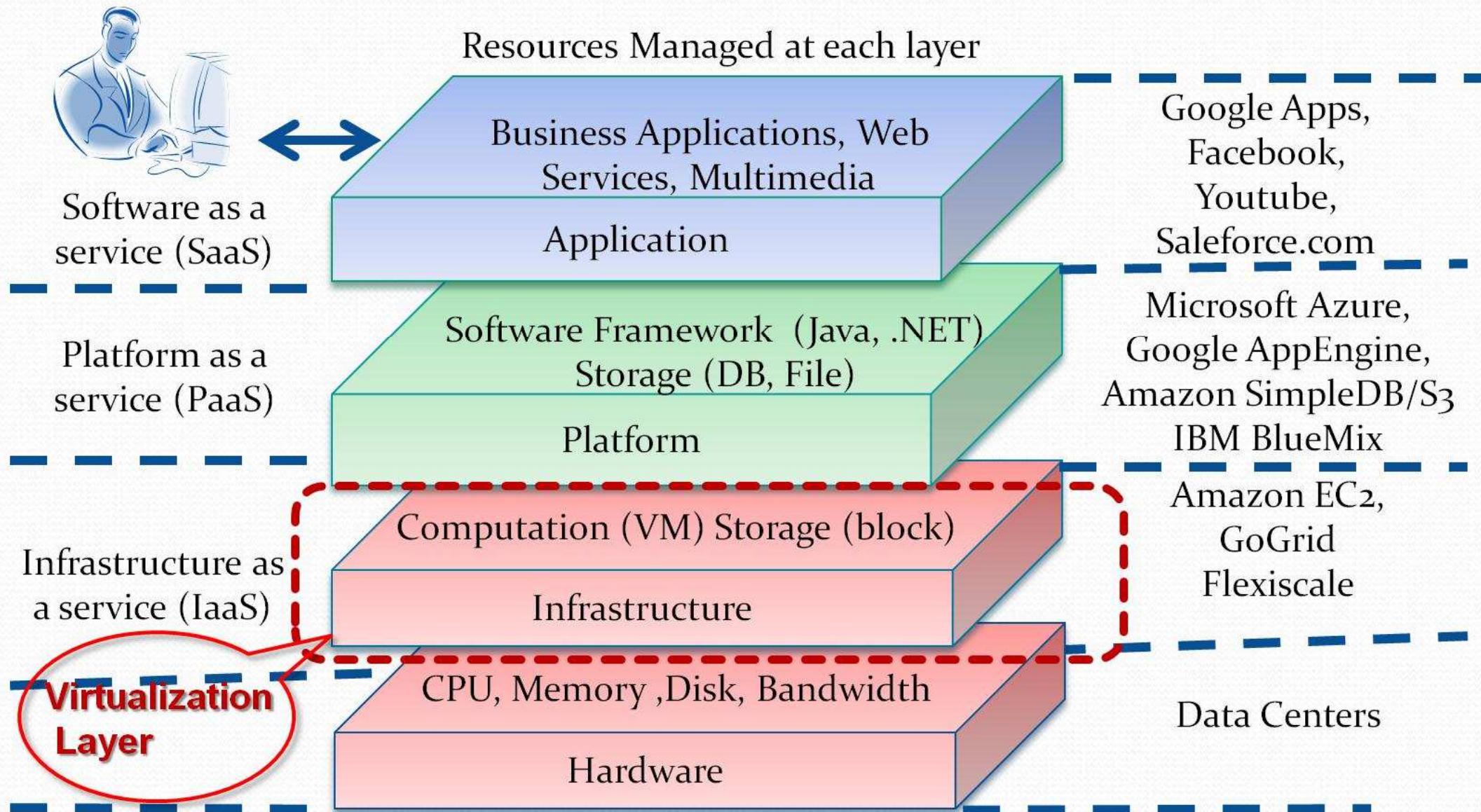
Cloud Computing Services

(CRM): Customer Relationship Management

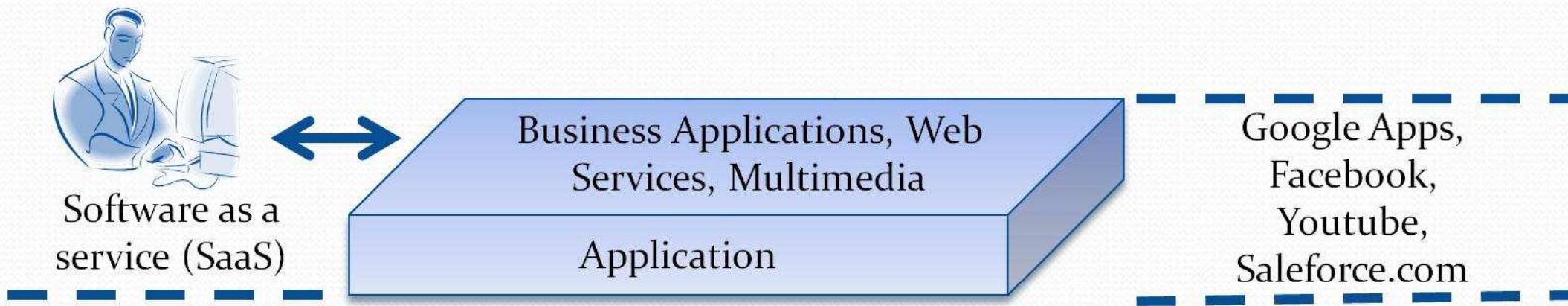


Who Uses It	What Services are available	Why use it?
Business Users	E-Mail, Office Automation, CRM, Website Testing, Wiki, Blog, Virtual Desktop ...	To complete business tasks
Developers and Deployers	Service and application test, development, integration and deployment environment	Create or deploy applications and services for users
System Managers	Virtual machines, operating systems, message queues, networks, storage, CPU, memory, backup services	Create platforms for service and application test, development, integration and deployment

Cloud Computing Hierarchy



Application Layer



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

SaaS Examples



Microsoft Online Services: Business Productivity Online Suite

Microsoft SharePoint Online

Office Communications Online

Microsoft Exchange Online

Office Live Meeting



facebook.

Platform Layer

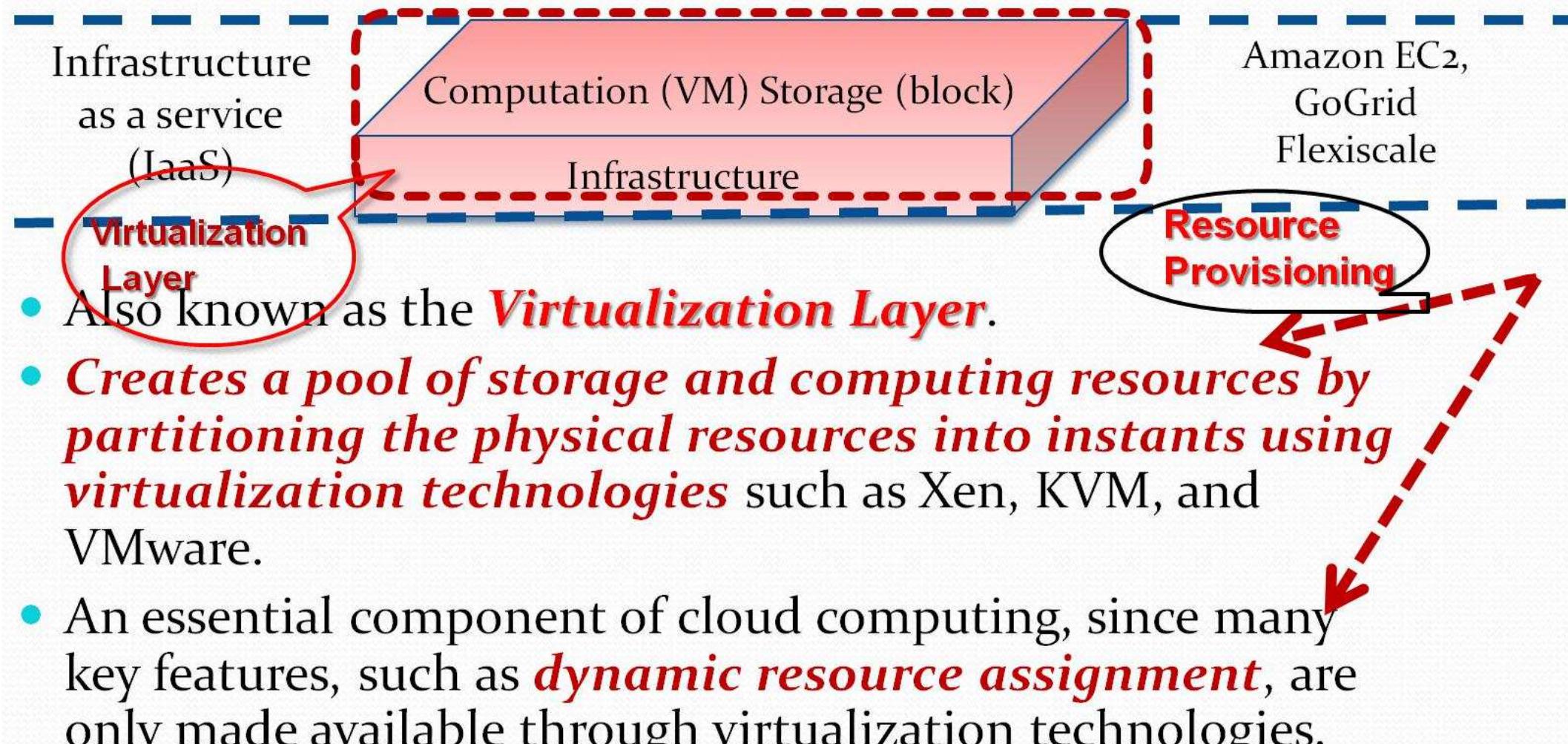


- Built on top of the infrastructure layer.
- Consists of OSs and application frameworks by ***providing Virtual Machines with specific configuration.***
- Minimize the burden of deploying applications directly into VM containers.
- E.g. Google App Engine provides API support for implementing storage,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

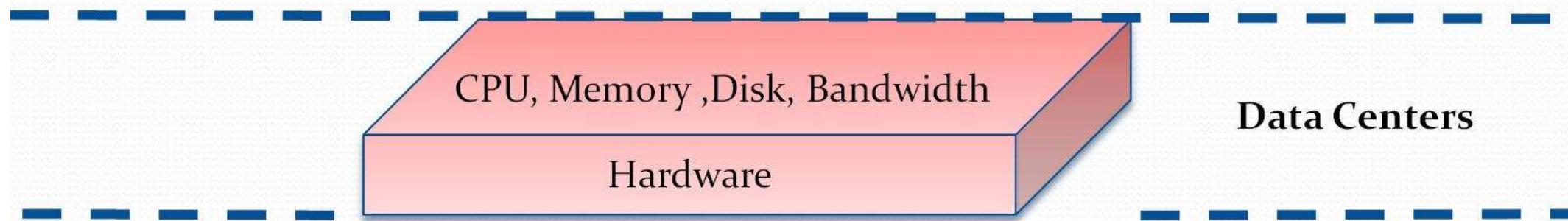
PaaS Examples



Infrastructure Layer



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 management,
 - Power and cooling, resource management.

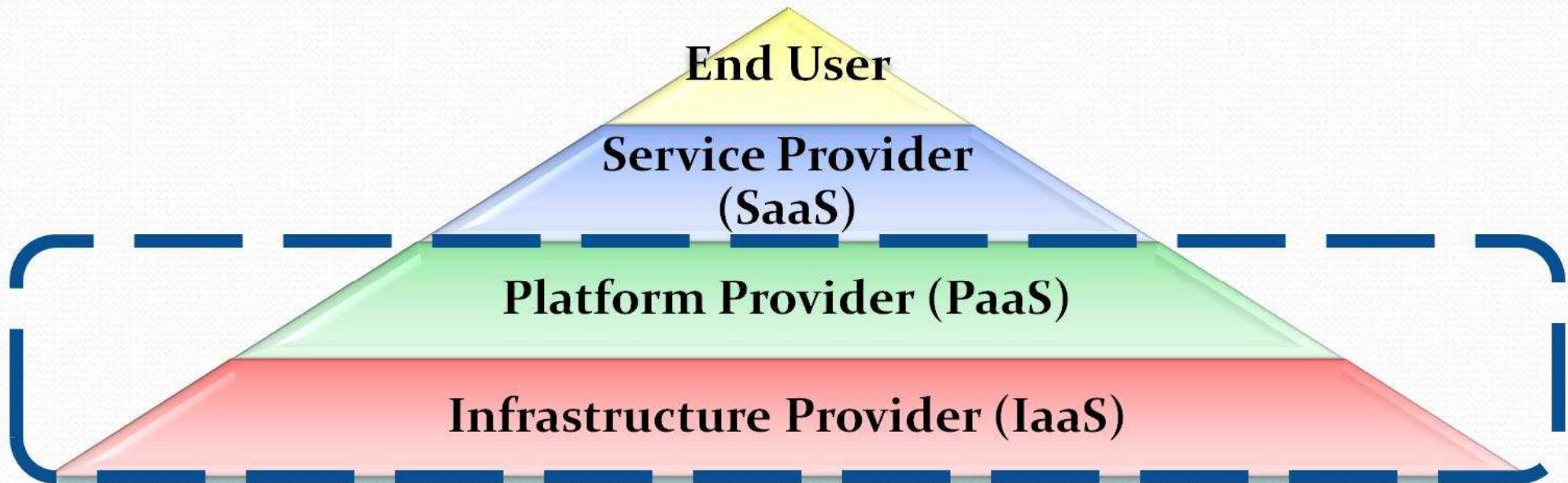
IaaS Examples



Do you Use the Cloud?



Cloud Computing Providers



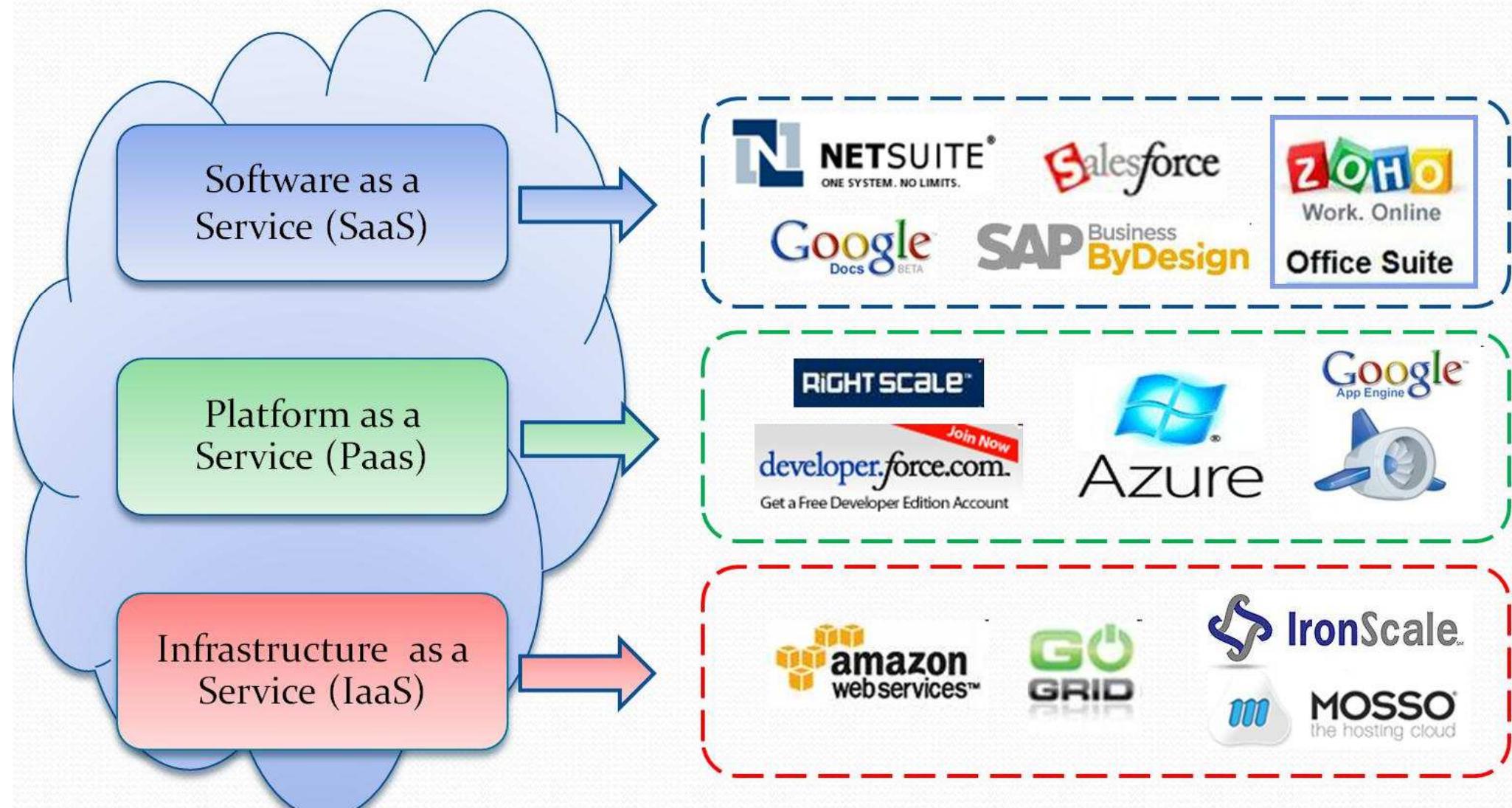
- IaaS and PaaS providers are often parts of the same organization (e.g. Google, Salesforce).
- PaaS and IaaS providers are often called the **infrastructure providers** or **cloud providers**

THE BOTTOM LINE

Cloud computing is
a simple idea,
but it can have
a huge impact on
the business.



Market Overview



Some Commercial Clouds

Report

❖ Windows Azure (PaaS)

- Hosting service for .NET applications and SQL databases

❖ Google AppEngine (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.

Service Level Agreement (SLA)

- Cloud places a strong emphasis on service management.
- *IaaS, PaaS and SaaS providers offer service according to Service Level Agreement (SLA) with customers.*
- SLA assurance is a critical objective of every provider.
- SLAs suppose to keep customers and providers happy, but writing one for the cloud means setting boundaries to account for more players and possibilities.



Service Level Agreement (SLA)-Cont..

- User needs:
 - **Guarantee** the resources **availability, reliability**, and **QoS**.
- Provider needs:
 - **Guarantee** the resource **utilization** and **revenue**
- This guarantee provided through a **negotiation** with some criteria called ***Service Level Agreements (SLAs)***.

The 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 users' needs and creates a relationship between the user and the service provider.***



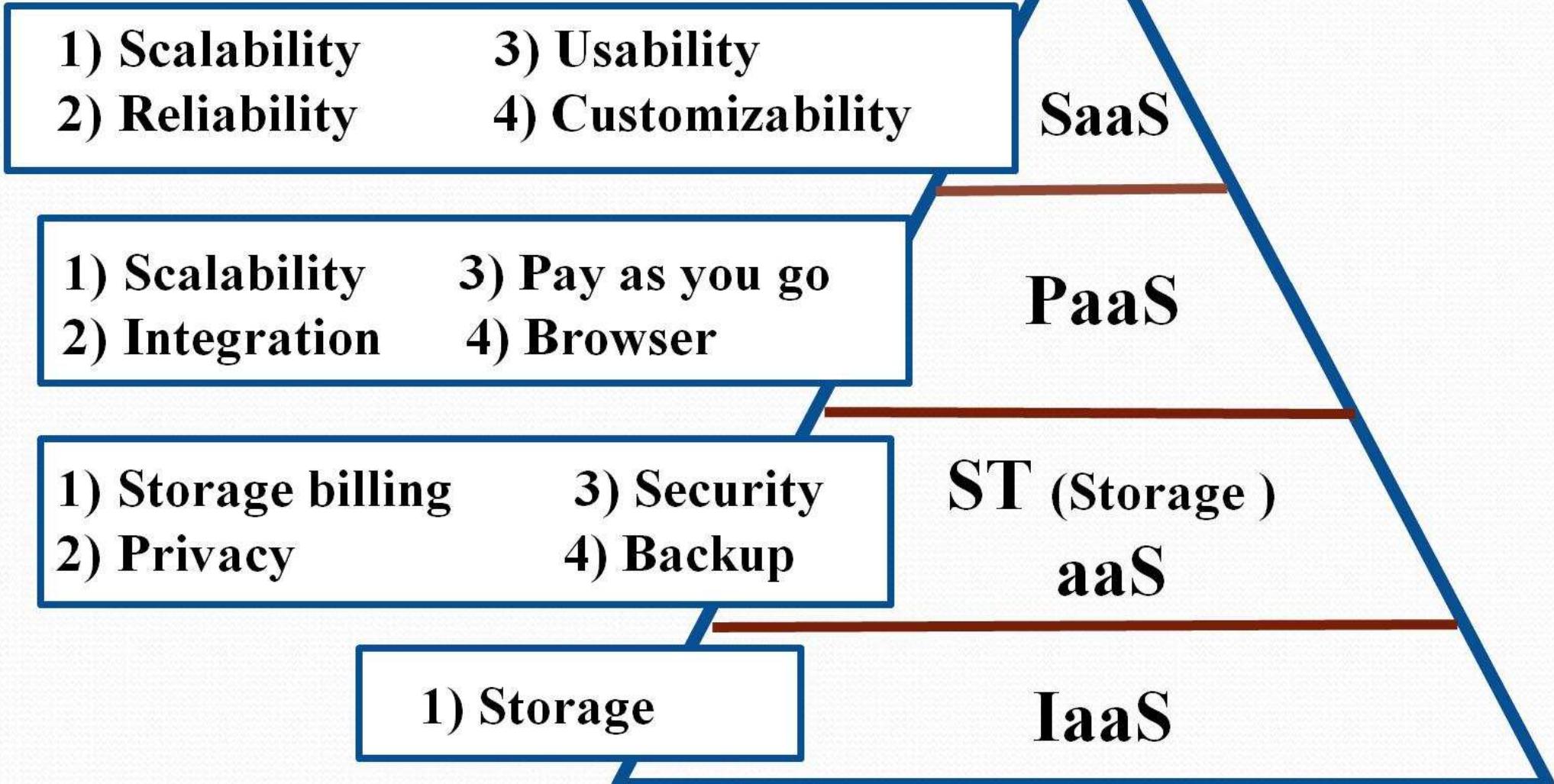
Service Level Agreement (SLA)-Cont..

- SLA describes a set of ***non-functional requirements*** of the service the customer is rendering.
- SLA usually ***contains penalties when the requirements are not met.***
- **If one or more conditions are met, then appropriate actions are triggered.**



Cloud Layers

Example: SLA Metrics for E-Education Service



Service Level Agreement (SLA)- Cont..

A number of requirements that service users should consider for SLAs are:

Component	Description
Responsibilities	Cloud service users should be responsible for limits on system usage and restrictions on the type of data that can be stored
Business continuity and disaster recovery	Cloud service users should ensure their cloud providers have adequate protection in case of a disaster.
System redundancy	Cloud service users moving data and applications that must be constantly available should consider the redundancy of their provider's systems.
Maintenance	Cloud service users should understand how and when their providers will do maintenance tasks
Location of Data	Cloud service users must be able to audit the provider to prove that regulations are being followed if a cloud service provider promises to enforce data location regulations,
Security	Cloud service users must understand their security requirements and what controls and federation patterns are necessary to meet those requirements.
Transparency	Cloud service users bear the burden of proving that the provider failed to live up to the terms of the SLA under the SLAs of some cloud providers.
Certification	Cloud service users might have the certification requirement that their cloud provider be ISO 27001 certified.

Service Level Agreement (SLA)- Cont..

a number of requirements providers should consider for SLAs are:

Component	Description
Security	Provider must understand what they must deliver to the service users to enable the appropriate controls and federation patterns.
Data Encryption	The details of the encryption algorithms and access control policies should be specified in the SLA.
Privacy	An SLA should make it clear how the cloud provider isolates data and applications in a multi-tenant environment.
Data Retention and Deletion	Cloud providers must be able to keep data for a certain period of time and delete data after a certain period of time.
Hardware Erasure and Destruction	Cloud providers should offer the added protection of zeroing out memory space after a consumer powers off a VM.
Regulatory Compliance	Cloud providers must be able to prove their compliance if regulations must be enforced.
Transparency	Cloud providers must be proactive in notifying consumers when the terms of the SLA are breached for critical data and applications.
Certification	Cloud provider would be responsible for proving their certification and keeping it up-to-date.

Cloud Computing Transparency

- The cloud transparency shields complexities related to utility computing and Software as a Service (SaaS) from both ***developers*** and ***end-users***.
- Clouds transparency addressed eight aspects:
 - ***Access, location, concurrency, replication, failure, mobility, performance, and scaling.***

[Egypt's Cloud Computing Strategic Plan-2011]

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

Cloud Computing Challenges

- **Bandwidth Costs**
 - Low for smaller Internet-based applications
 - Significantly grow for data-intensive applications
- **Reliability**
 - Not always offer round-the-clock reliability
 - **HP Labs, Intel, and Yahoo launched the distributed Cloud Research Test Bed to develop innovations including cloud computing specific chips**
 - **IBM has launched the *Research Computing Cloud*, which is an *on-demand, globally accessible set of computing resources* that support business processes**