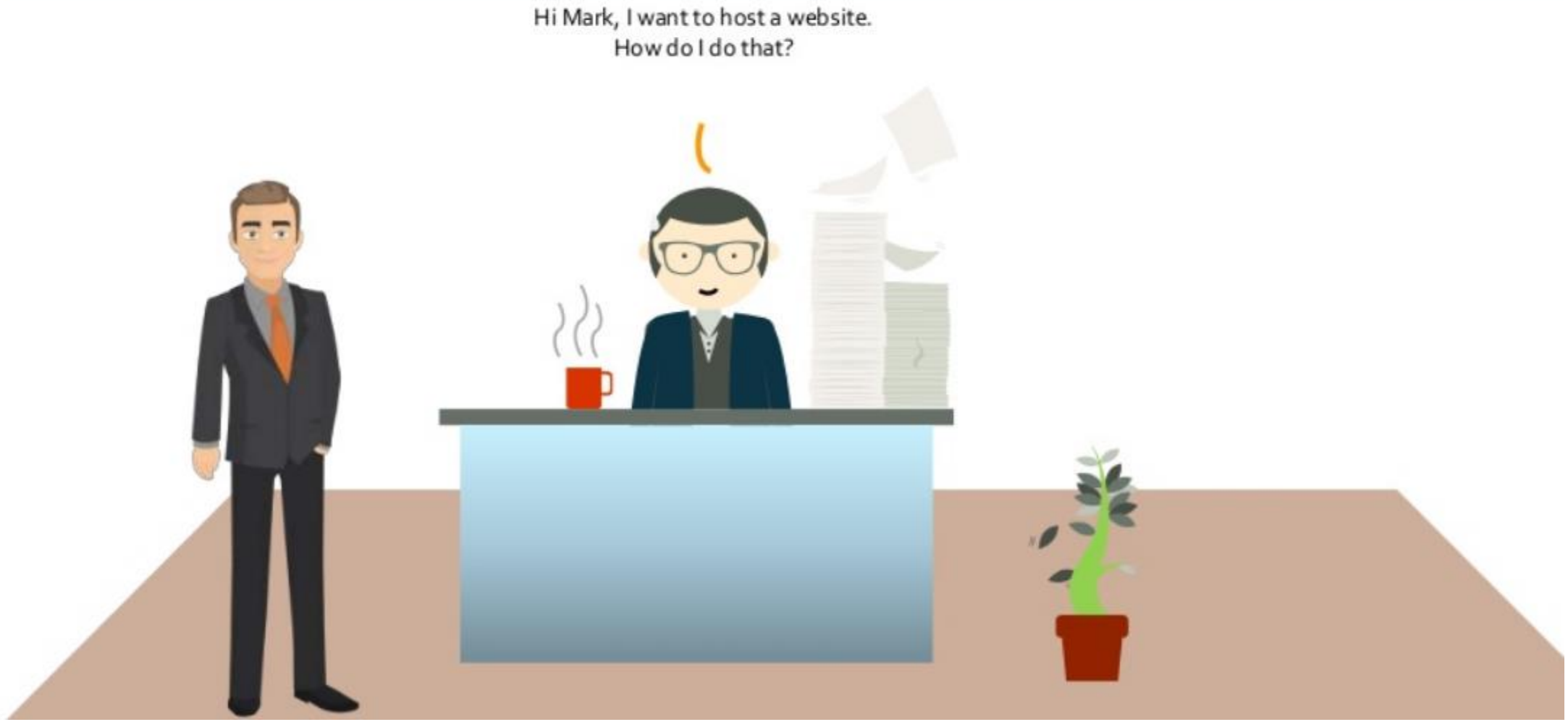
Cloud Computing

Introduction

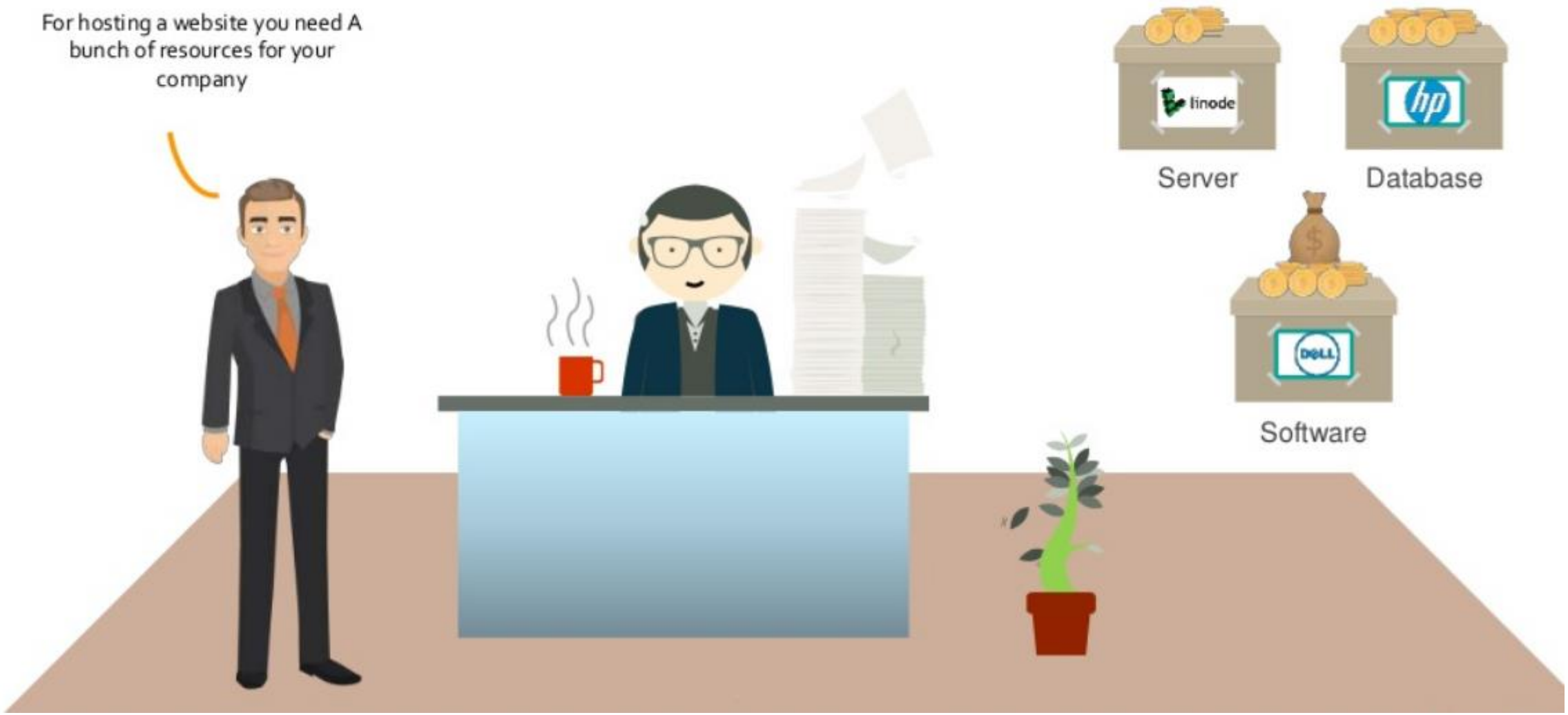
What's in it for you?

1. Before Cloud Computing
2. What is Cloud Computing?
3. Benefits of Cloud Computing
4. Types of Cloud Computing
5. Who uses Cloud Computing?

Before Cloud Computing



Before Cloud Computing



Before Cloud Computing

Before Cloud Computing

Before Cloud Computing

Suppose you want to host a website, these are the following things that you would need to do:

Buy a stack of servers.

Monitoring and Maintain servers.

High traffic? More servers.

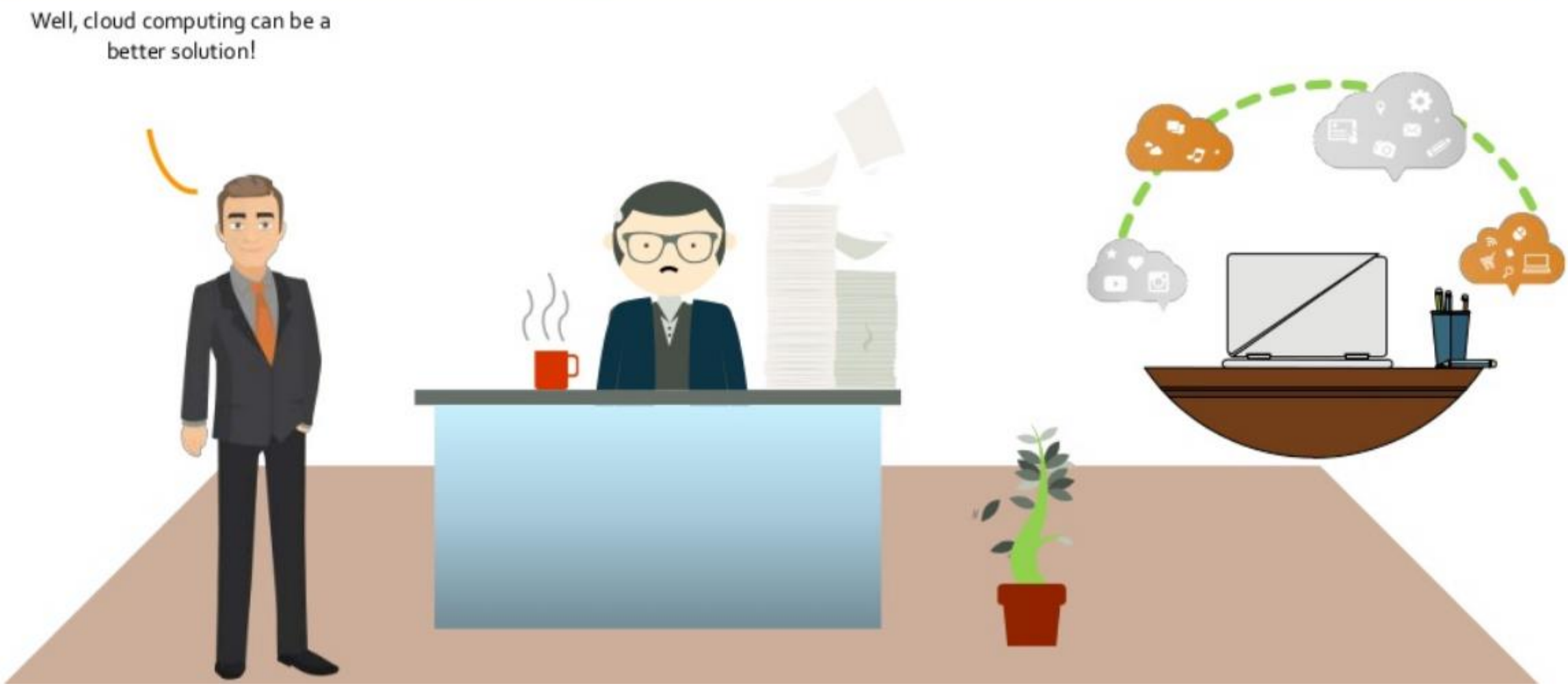
Disadvantages

If you consider costs then this setup is expensive.

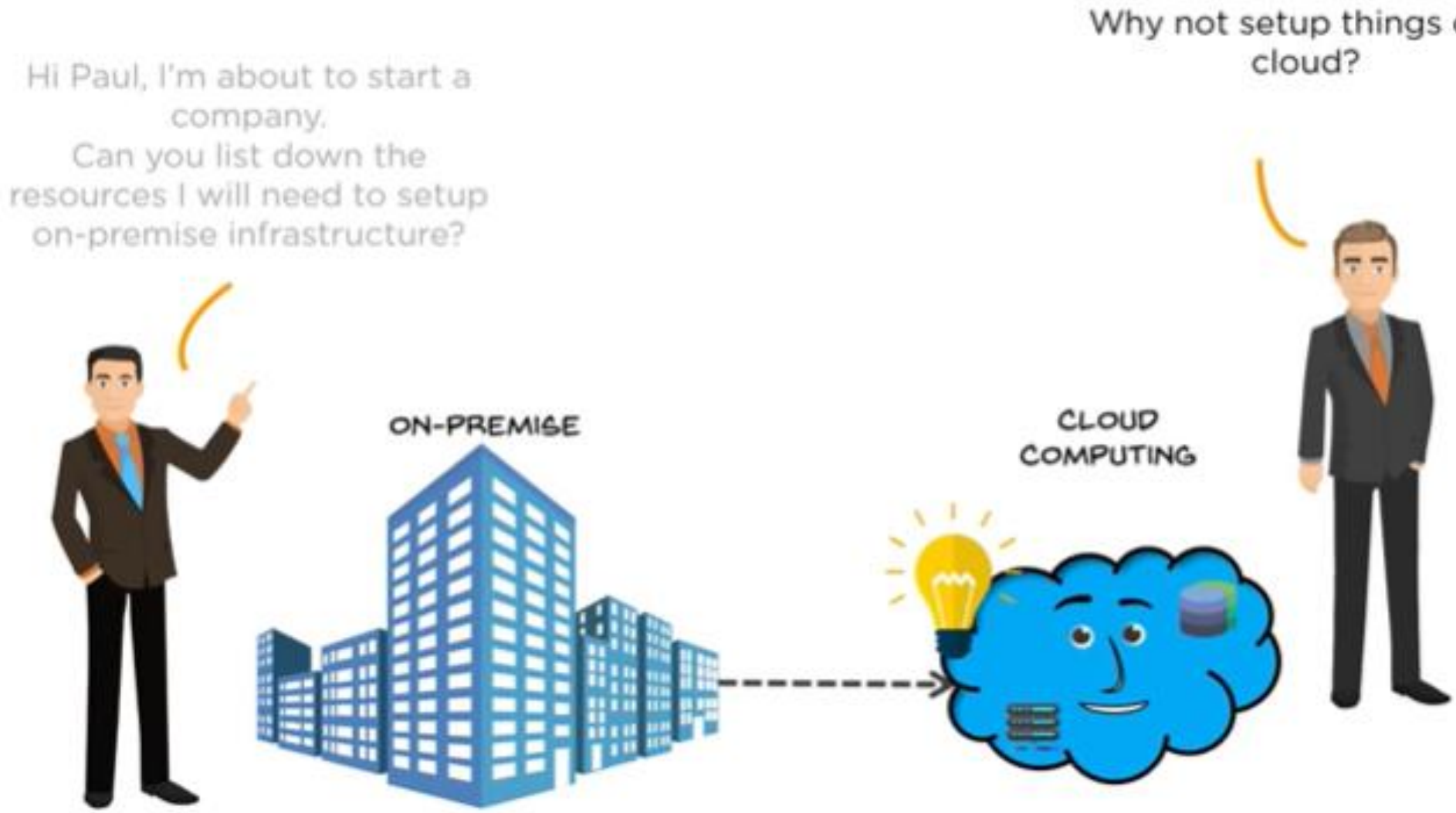
Troubleshooting problems can be tedious and may conflict with your business goals.

Since the traffic is varying, your servers will be idle most of the time.

Before Cloud Computing



Why Cloud Computing?



Why Cloud?

Files

music

E-books

Videos

Applications

Podcasts

What Is Cloud?

Just move your data to Cloud

Local system with limited space

Cloud with unlimited space

What is Cloud Computing?

Rather than managing files on a local storage device, cloud computing makes it possible to save them over internet

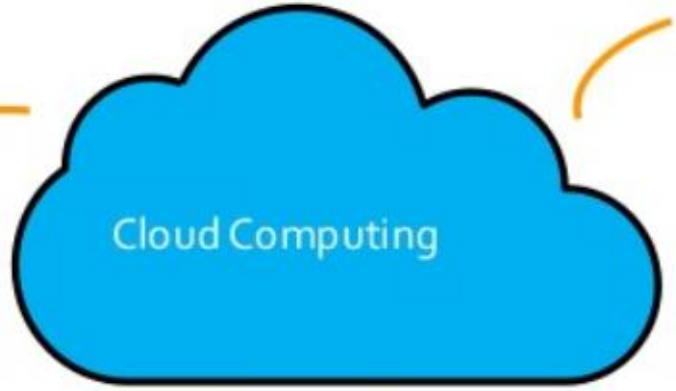
What is Cloud Computing?

Cloud Computing is the use of a network of remote servers hosted on the Internet to store, manage and process data rather than a local server

Pay for what you use

In simple terms, it means storing or accessing your data over the internet

It is the delivery of computing services (servers, databases, networking, software etc.) over the Internet

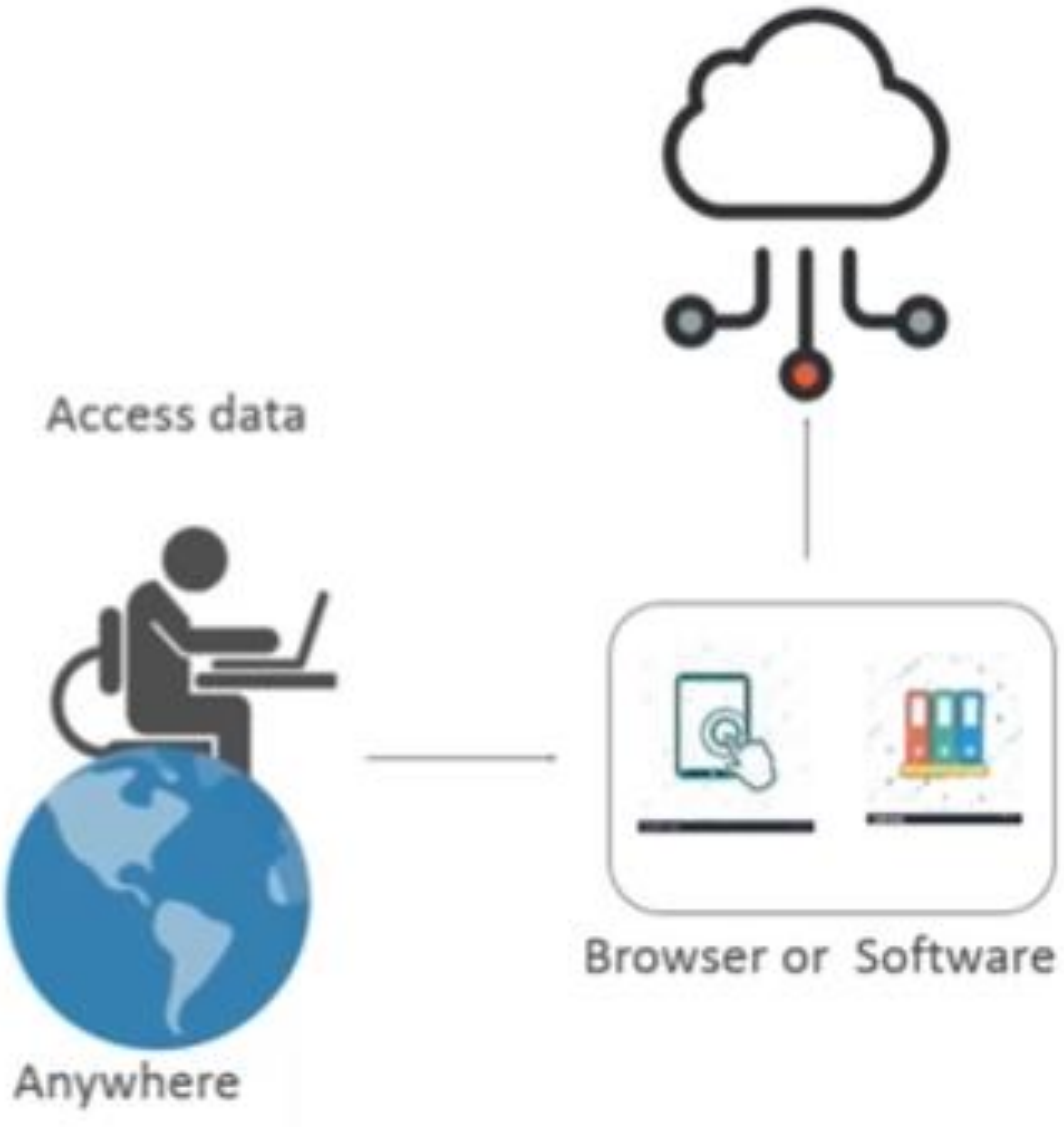


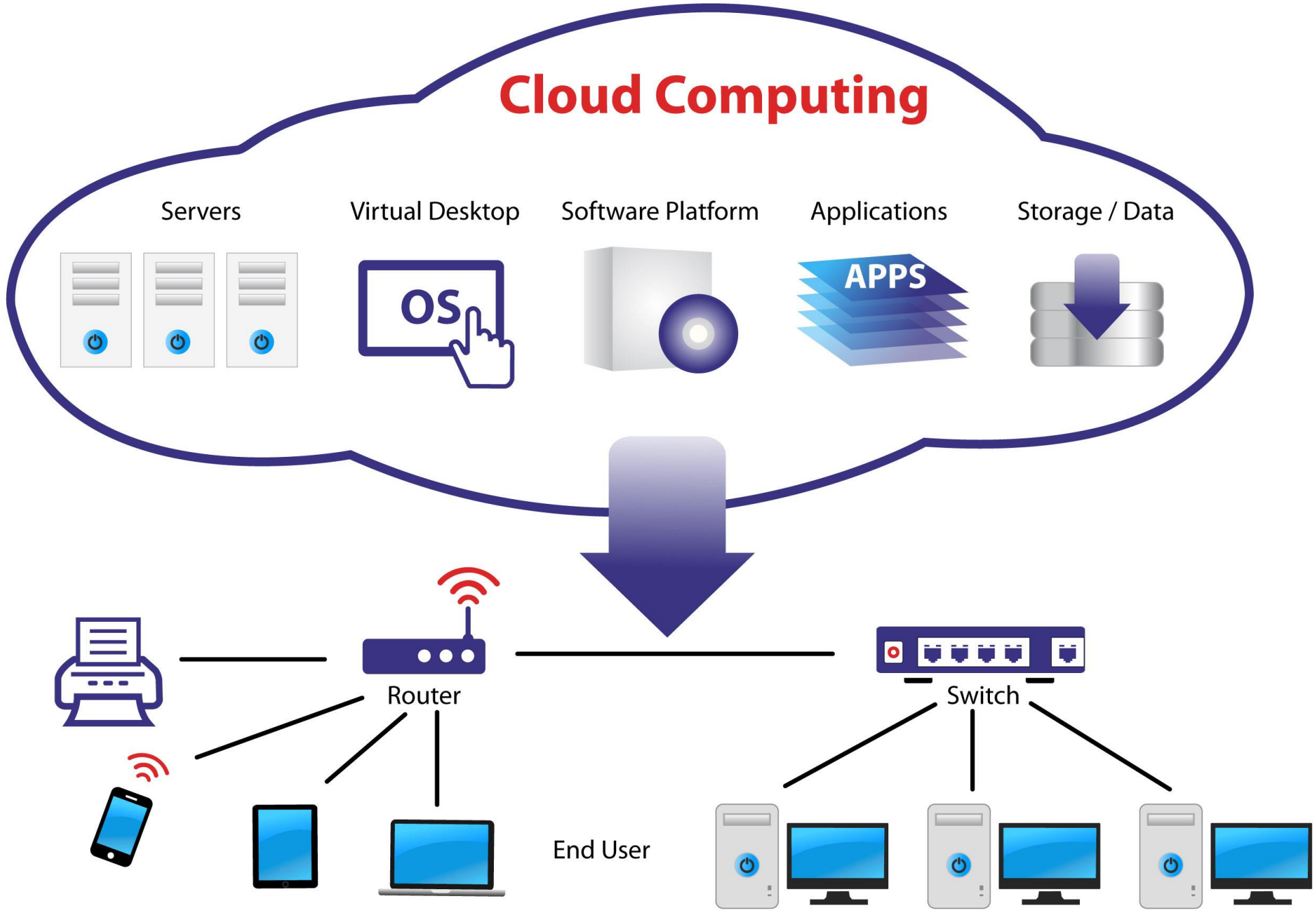
Note: Companies offering these computing services are called cloud providers

What Is Cloud Computing?

Cloud computing is:

- Storing data/applications on remote servers
- Processing data/applications from servers
- Accessing data/applications via Internet




Cloud Computing
Servers
Virtual Desktop
Software Platform
Applications
Storage / Data
OS
APPS
Router
Switch
End User

Definitions

- Cloud computing is using the internet to access someone else's software running on someone else's hardware in someone else's data center.

Lewis Cunningham

Definitions

- A large-scale distributed computing paradigm that is driven by economies of scale, in which a pool of abstracted, virtualized, dynamically scalable, managed computing power, storage, platforms, and services are delivered on demand to external customers over the Internet.

Ian Foster

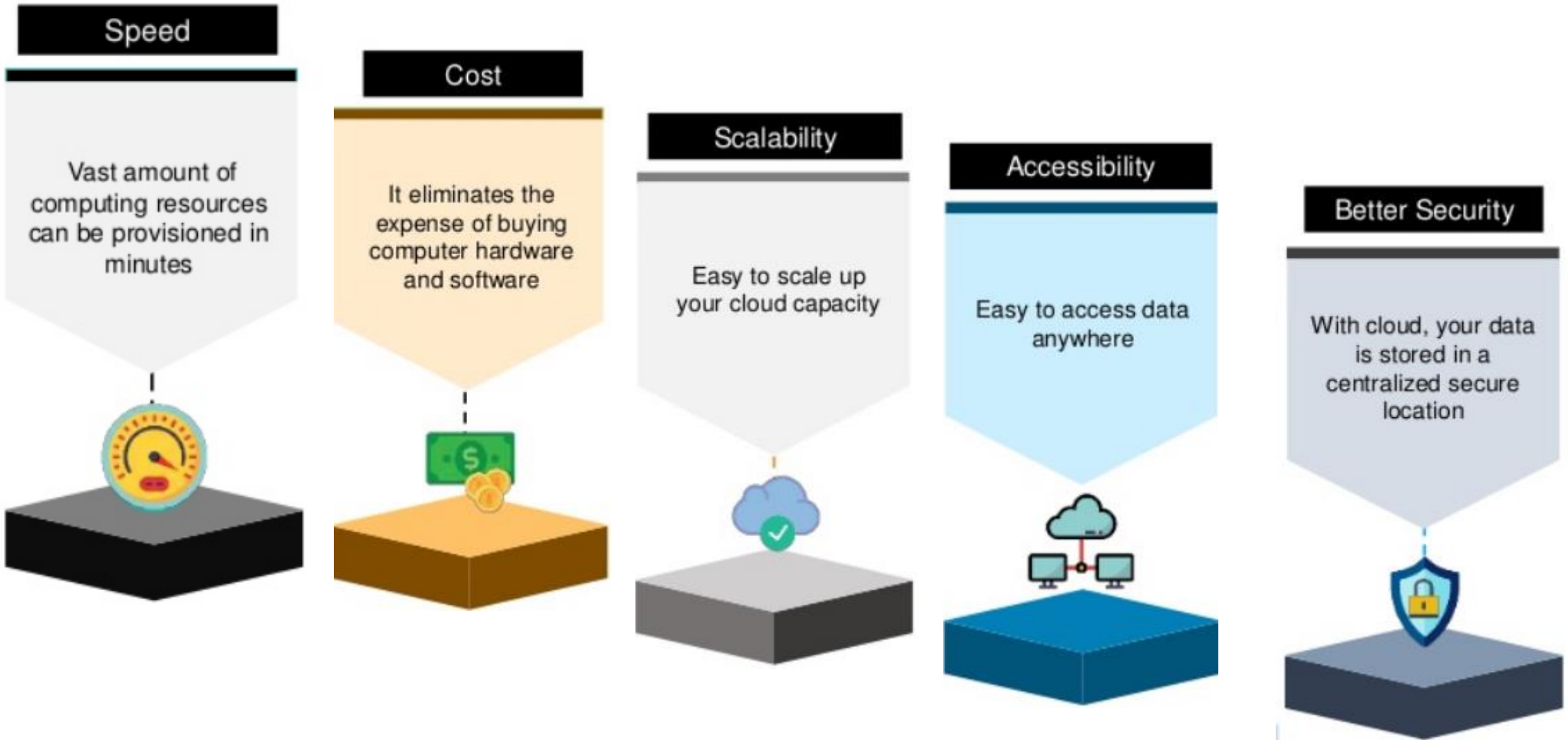
Definition:

*Cloud computing is **a model** for enabling 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.*

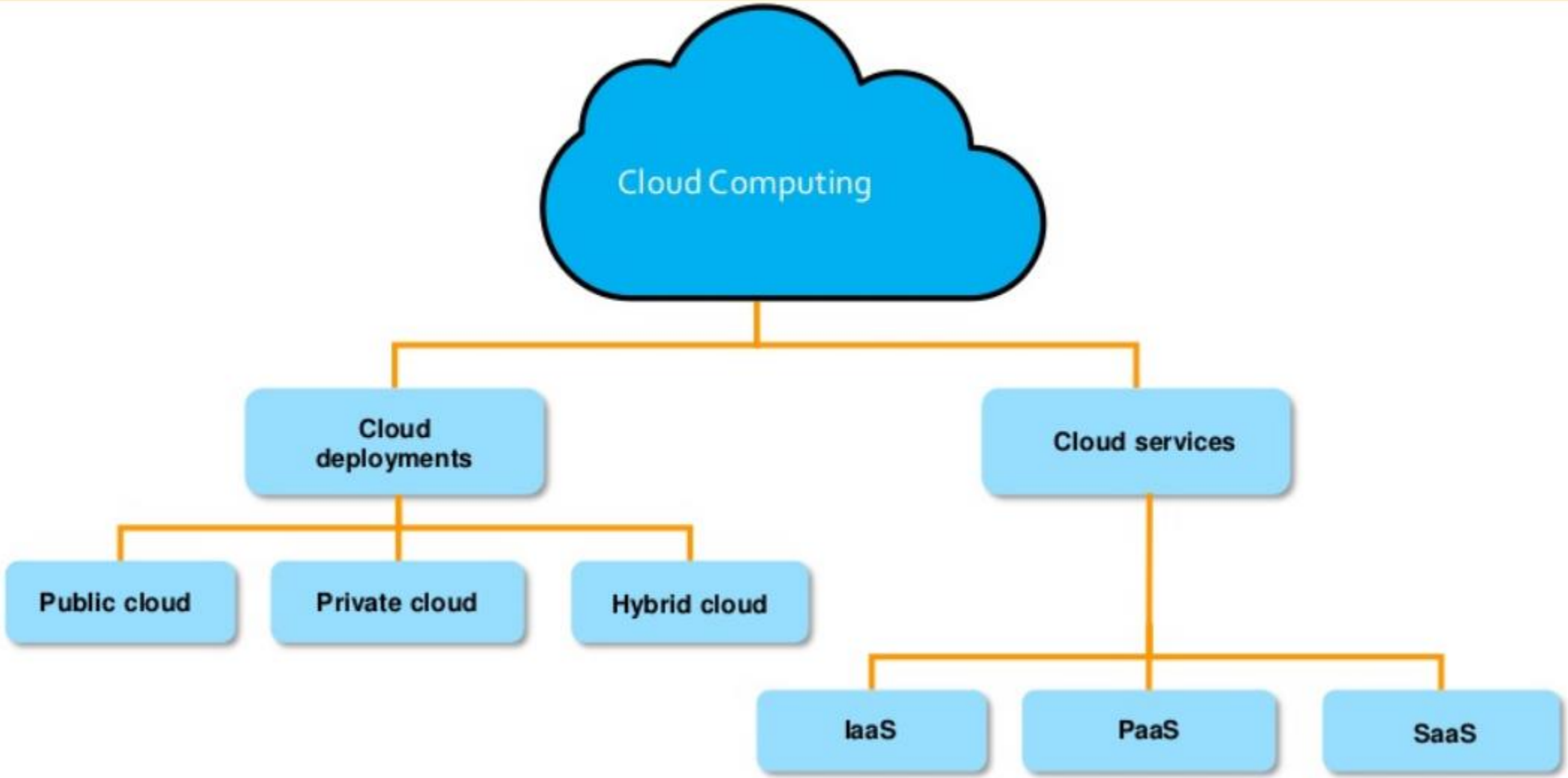
Definition by the U.S. Government's National Institute of Standards and Technology

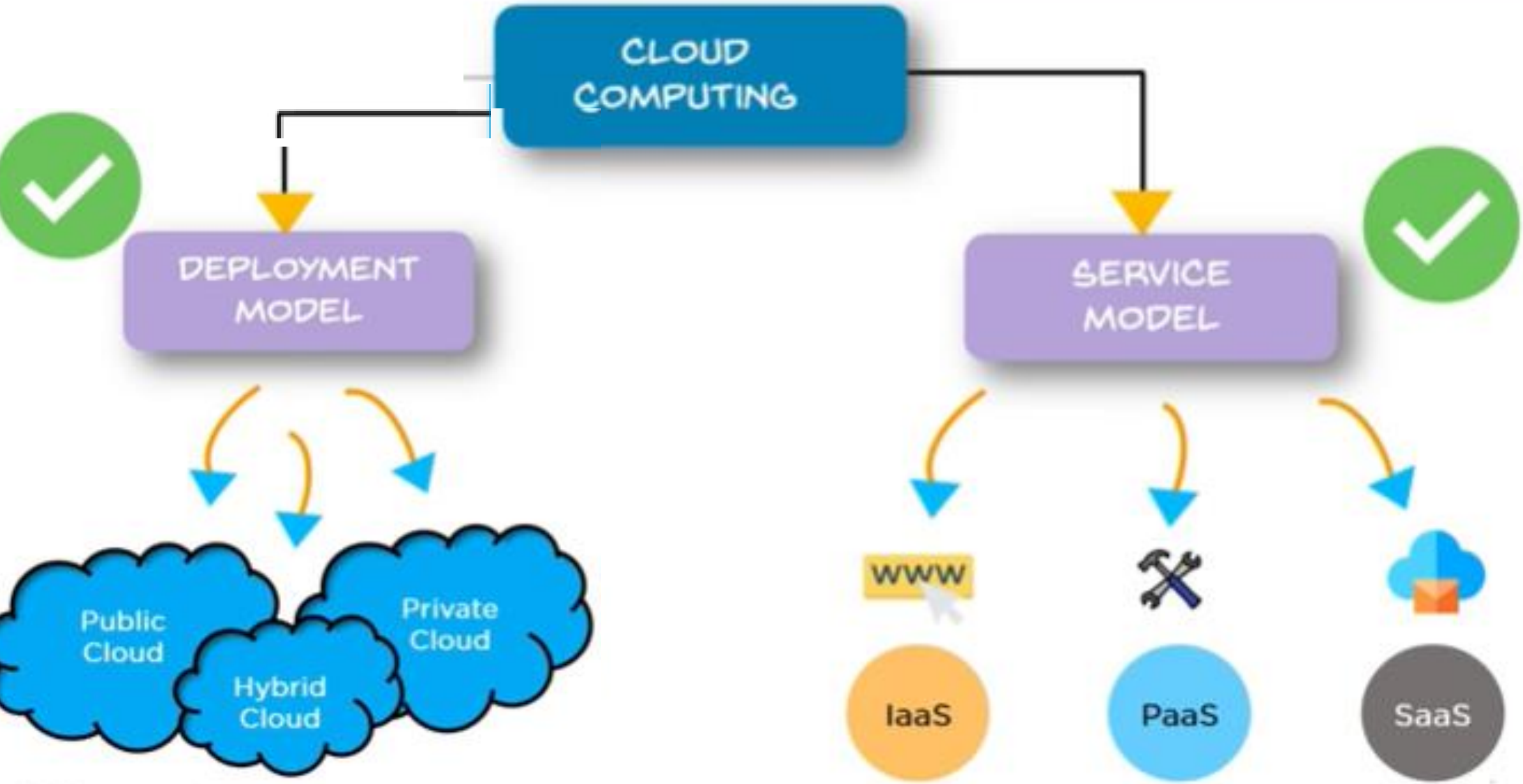
Source	Definition
Gartner	"a style of computing in which massively scalable IT-related capabilities are provided "as a service" using Internet technologies to multiple external customers" (Gartner 2008b)
IDC	"an emerging IT development, deployment and delivery model, enabling real-time delivery of products, services and solutions over the Internet (i.e., enablin cloud services)" (Gens 2008)
The 451 Group	"a service model that combines a general organizing principle for IT delivery, infrastructure components, an architectural approach and an economic model – basically, a confluence of grid computing, virtualization, utility computing, hosting and software as a service (SaaS)" (Fellows 2008)
Merrill Lynch	"the idea of delivering personal (e.g., email, word processing, presentations.) and business productivity applications (e.g., sales force automation, customer service, accounting) from centralized servers" (Merrill Lynch 2008)

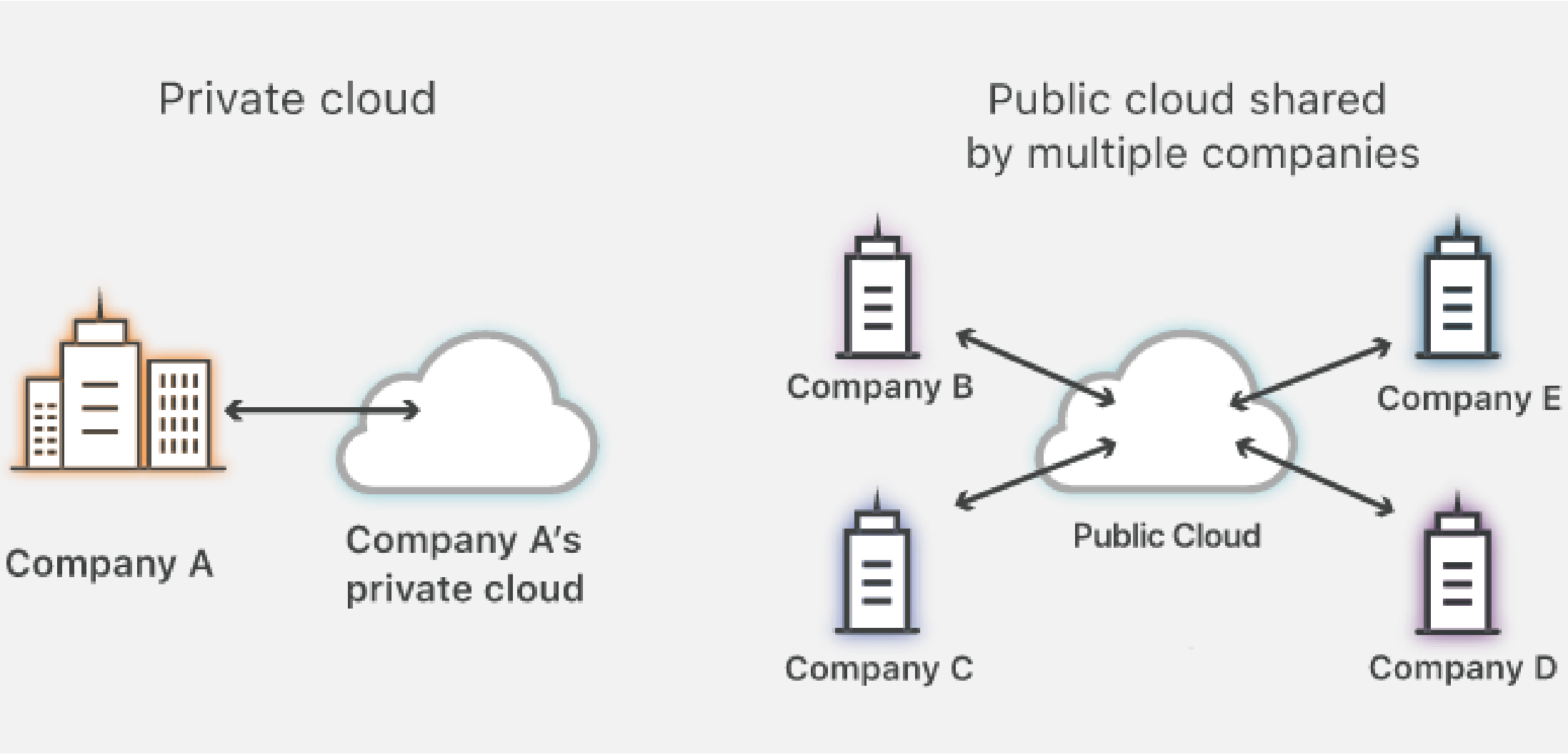
Benefits of Cloud Computing



Types of Cloud Computing






Private cloud
Company A
Company A's
private cloud
Public cloud shared
by multiple companies
Company B
Company E
Public Cloud
Company C
Company D

- **Private Cloud:** Cloud services are used by a single organization, so only the organization has access to its data and can manage it.
- **Public Cloud:**Cloud services are accessible via a network like internet and can be used by multiple clients. MS Azure, AWS are some of the examples of public cloud hosting.
- **Hybrid Cloud:**Cloud services are distributed among public and private clouds. Its users can use them as per their requirement i.e.private as well as public cloud services in day to day operations.

Tips to Choose the Best Cloud Model for Your Business

Deployment Models - Cloud Computing

01 Public Cloud

02 Private Cloud

03 Hybrid Cloud

- Here, the services are stored off-site and accessed over the internet
- It can be used by general public
- All hardware, software and other supporting infrastructure is owned and managed by the cloud provider
- Example: Amazon Web Service and Microsoft Azure

Deployment Models

- A service provider makes resources, such as applications and storage, available to the general public over the Internet.
- Easy and inexpensive set-up because hardware, application and bandwidth costs are covered by the provider.
- No wasted resources because you pay for what you use

Deployment Models - Cloud Computing

01 Public Cloud

02 Private Cloud

03 Hybrid Cloud

- The cloud infrastructure is used exclusively by a single organization
- The organization may run its private cloud or outsource it to a hosting company
- The services and infrastructure are maintained on a private network
- Example: AWS, VMware

Deployment Models

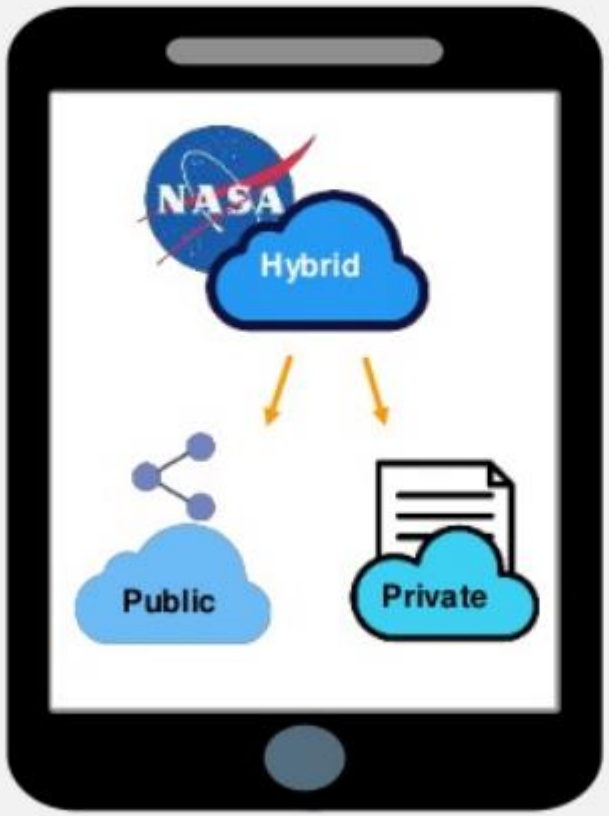
- Offers hosted services to a limited number of people behind firewall, so it minimizes the security concerns.
- Private cloud gives companies direct control over their data.

Deployment Models - Cloud Computing

01 Public Cloud

02 Private Cloud

03 Hybrid Cloud



- It consists the functionalities of both public and private cloud
- Example:

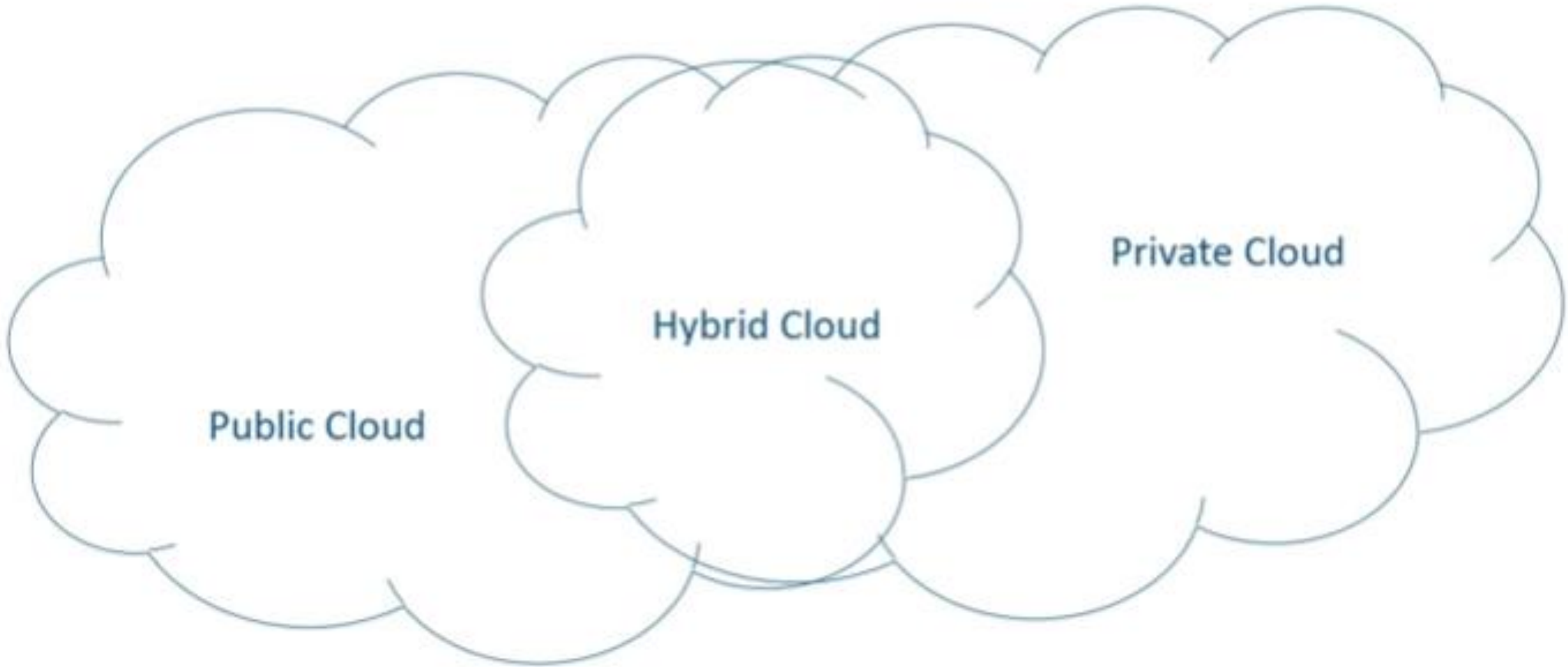
  For a project, NASA is using the Hybrid cloud computing deployment

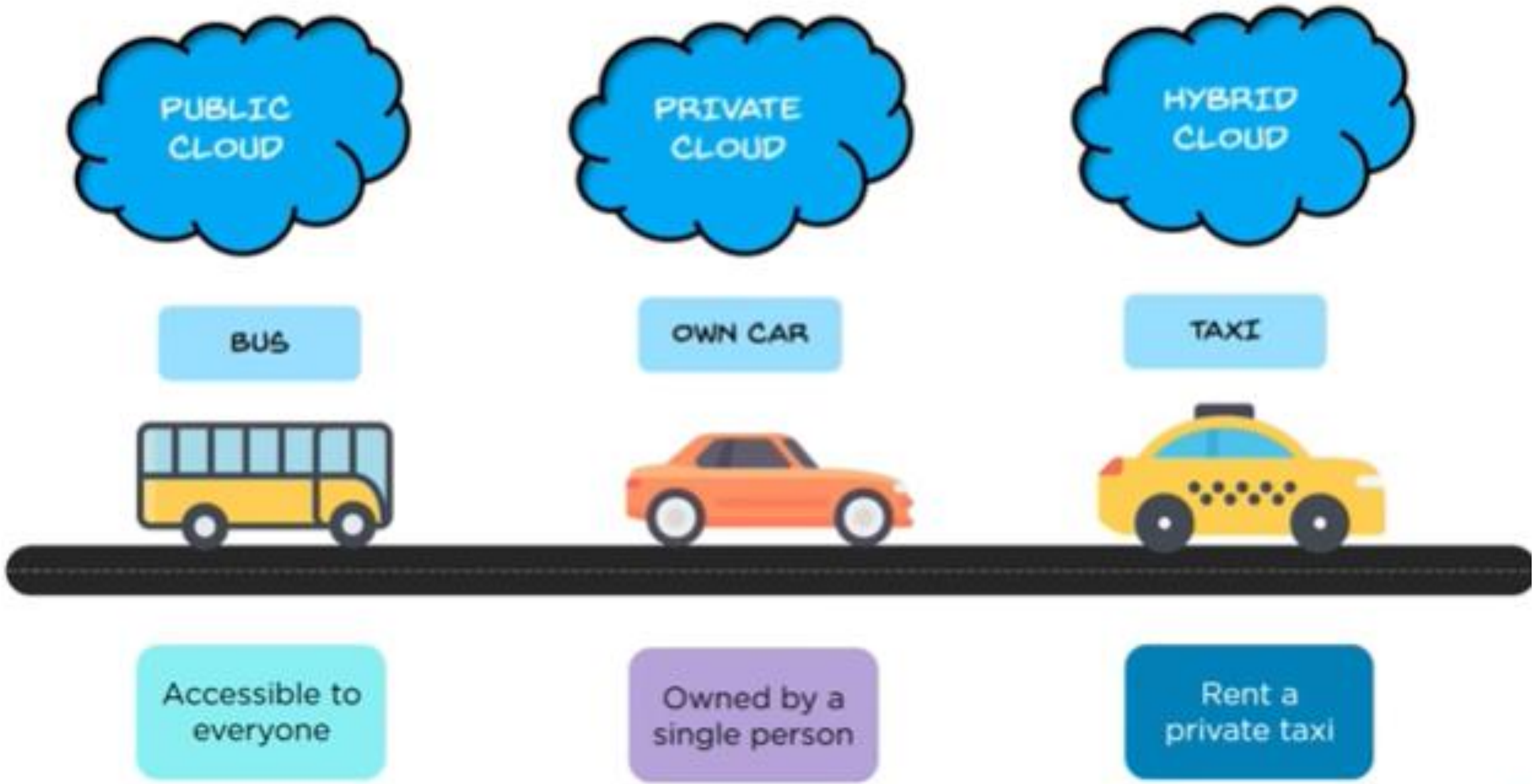
  Its open-source cloud registered project NEBULA uses a private cloud for research and development and in addition a public cloud to share datasets with external partners and the public

Deployment Models

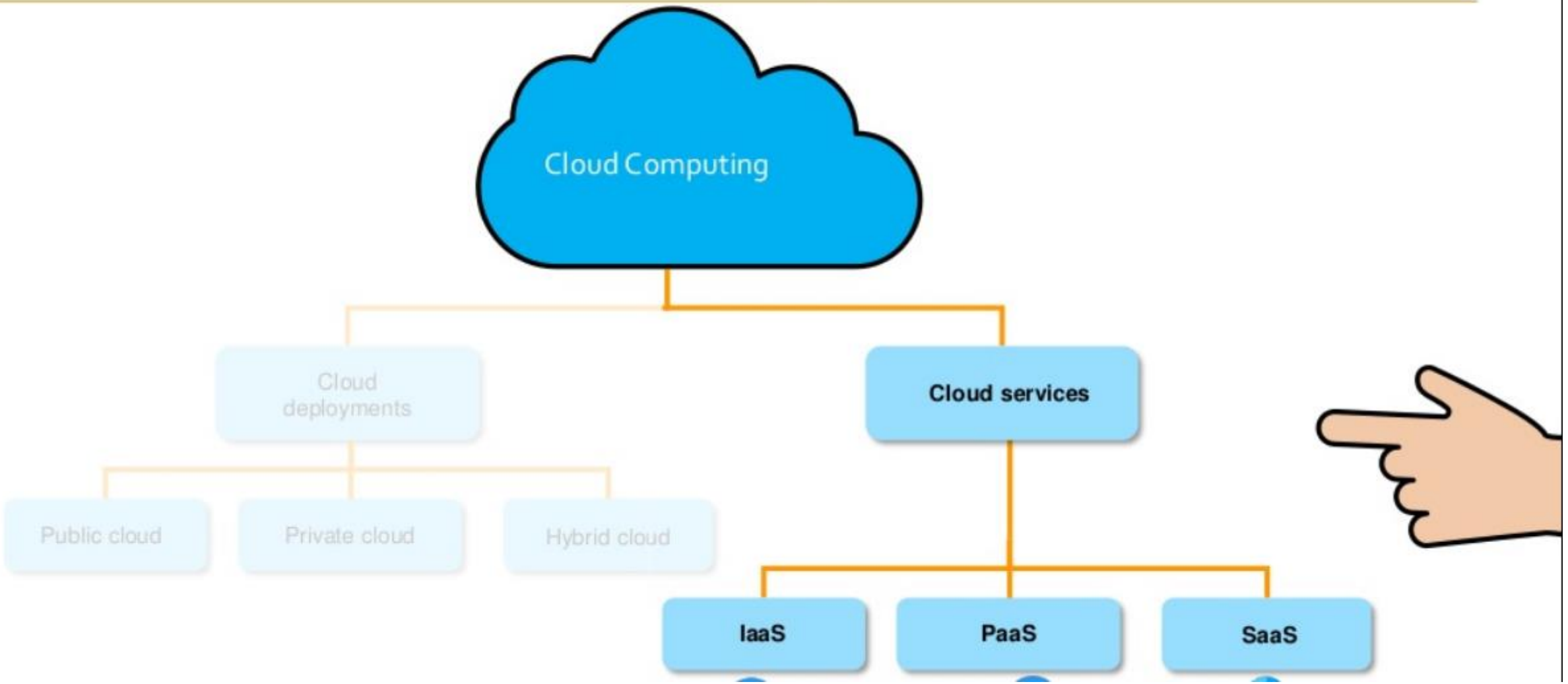
- A cloud computing environment which uses a mix of on-premises, private cloud and third-party, public cloud services
- It helps you leverage the best of both worlds

Deployment Models – Cloud Computing

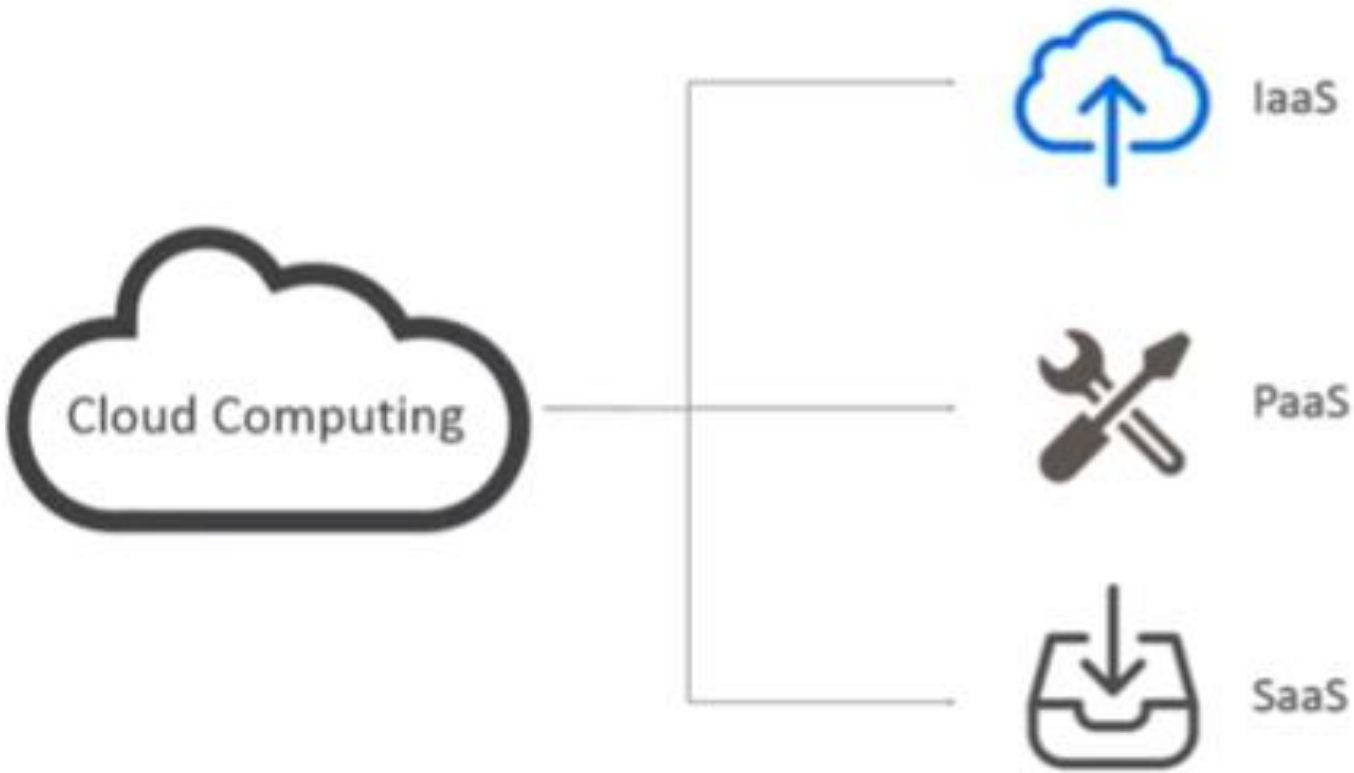




Service Models - Cloud Computing



Service Models : IaaS?



Infrastructure as a Service

- Provides virtualized computing resources over the Internet
- No worries about the underlying physical machine.
- Abstract the user from the physical machine

Service Models - Cloud Computing

01 Infrastructure as a service (IaaS)

02 Platform as a service (PaaS)

03 Software as a service (SaaS)

- You rent IT infrastructure (servers and networks etc.) from a cloud provider on a pay-as-you-go basis
- Users of IaaS can outsource and build a "virtual data center" in the cloud and have access to the resources as well
- Example: AWS Elastic Compute Cloud (EC2)

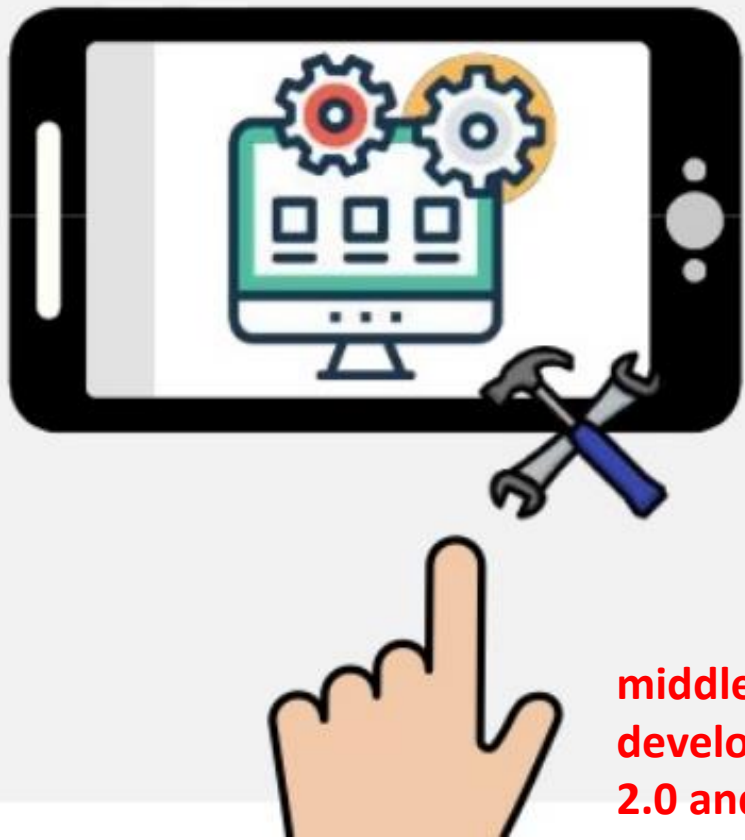
servers, storage, networks, and the data center fabric

Service Models - Cloud Computing

01 Infrastructure as a service (IaaS)

02 Platform as a service (PaaS)

03 Software as a service (SaaS)



- Provides a platform on which software can be developed and deployed
- The cloud provider allows the customer to deploy their own application using programming languages, tools etc.
- Example: AWS Elastic Beanstalk

middleware, databases, development tools, and some runtime support such as Web 2.0 and Java

Service Models - Cloud Computing

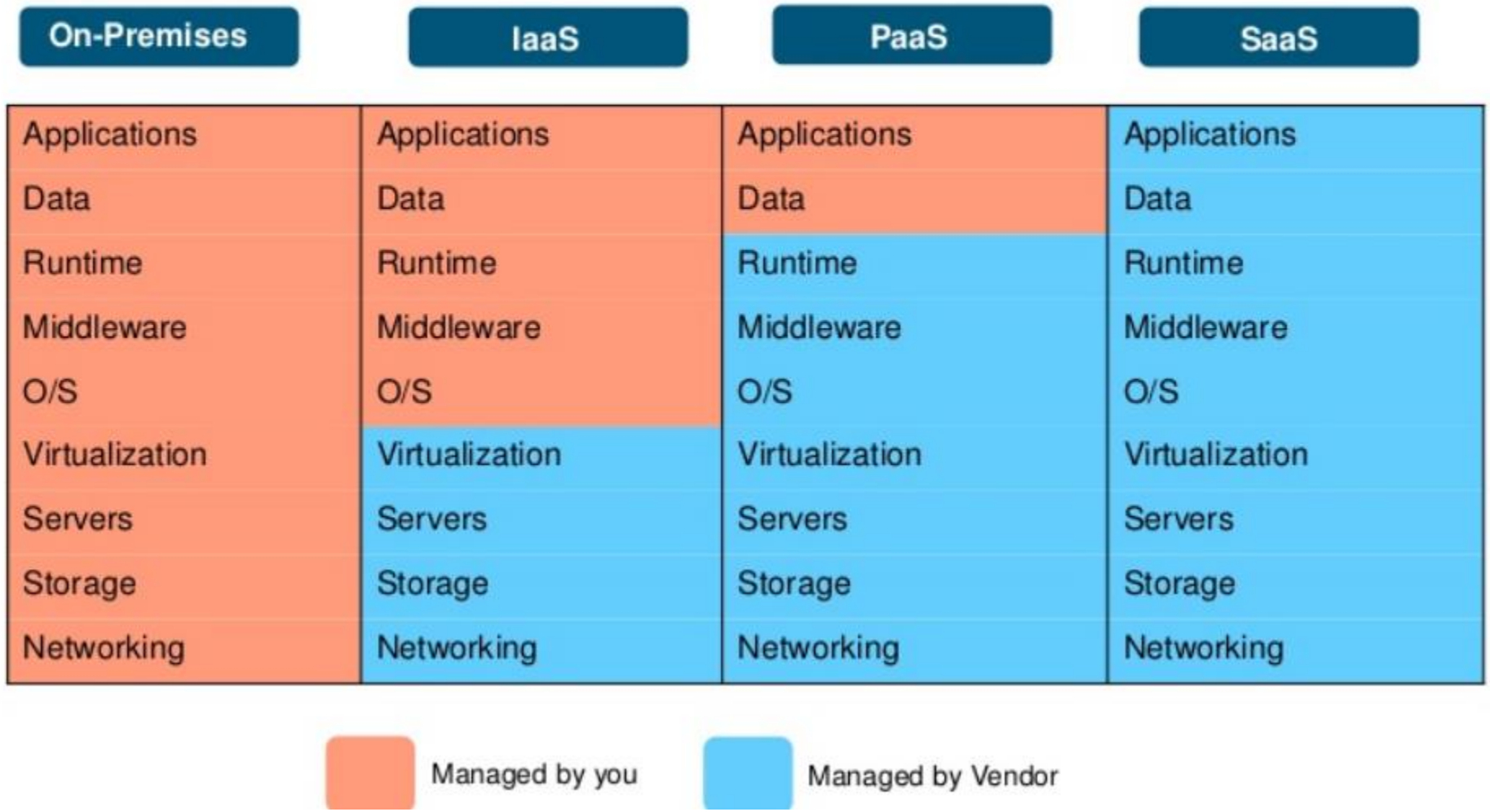
01 Infrastructure as a service (IaaS)

02 Platform as a service (PaaS)

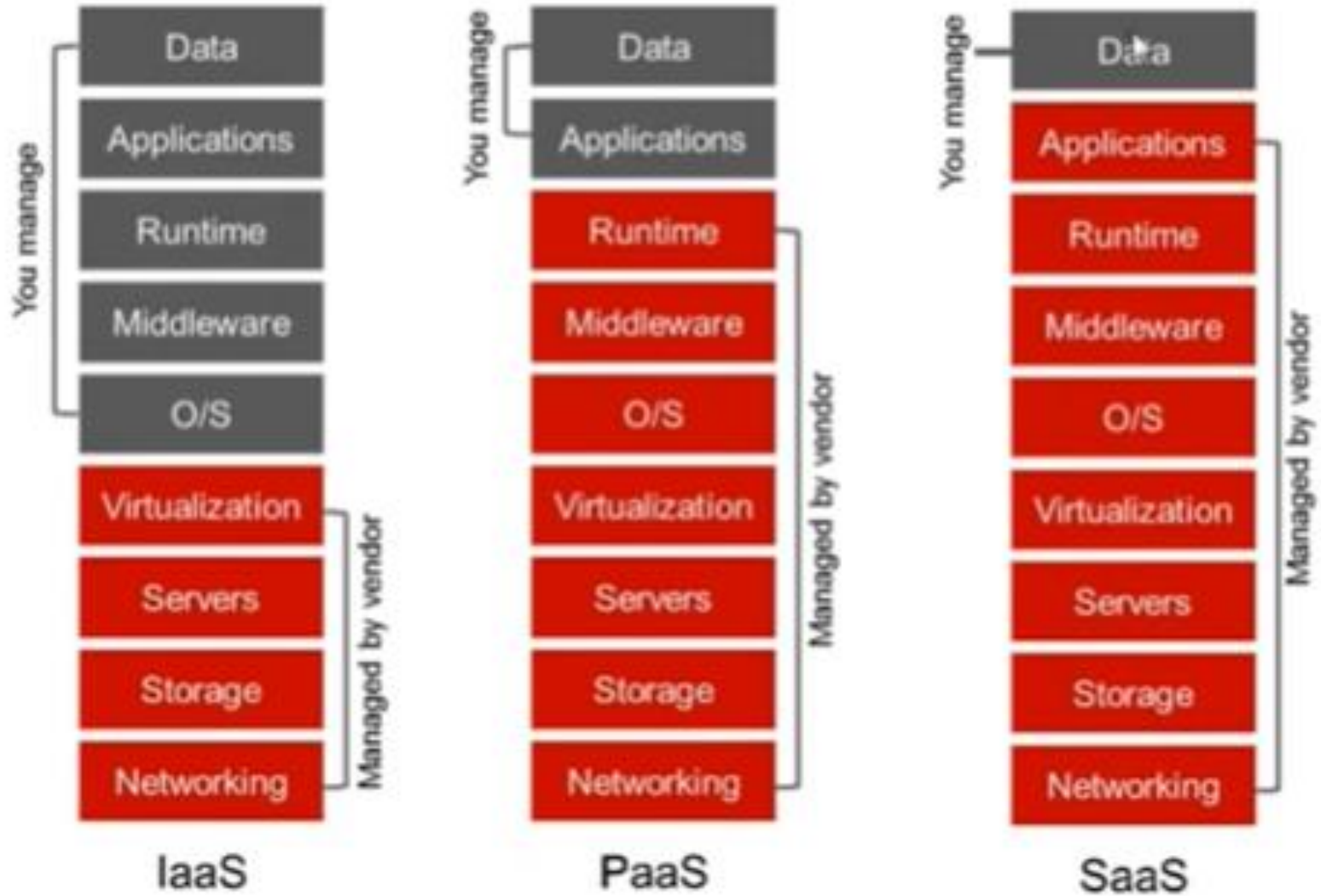
03 Software as a service (SaaS)

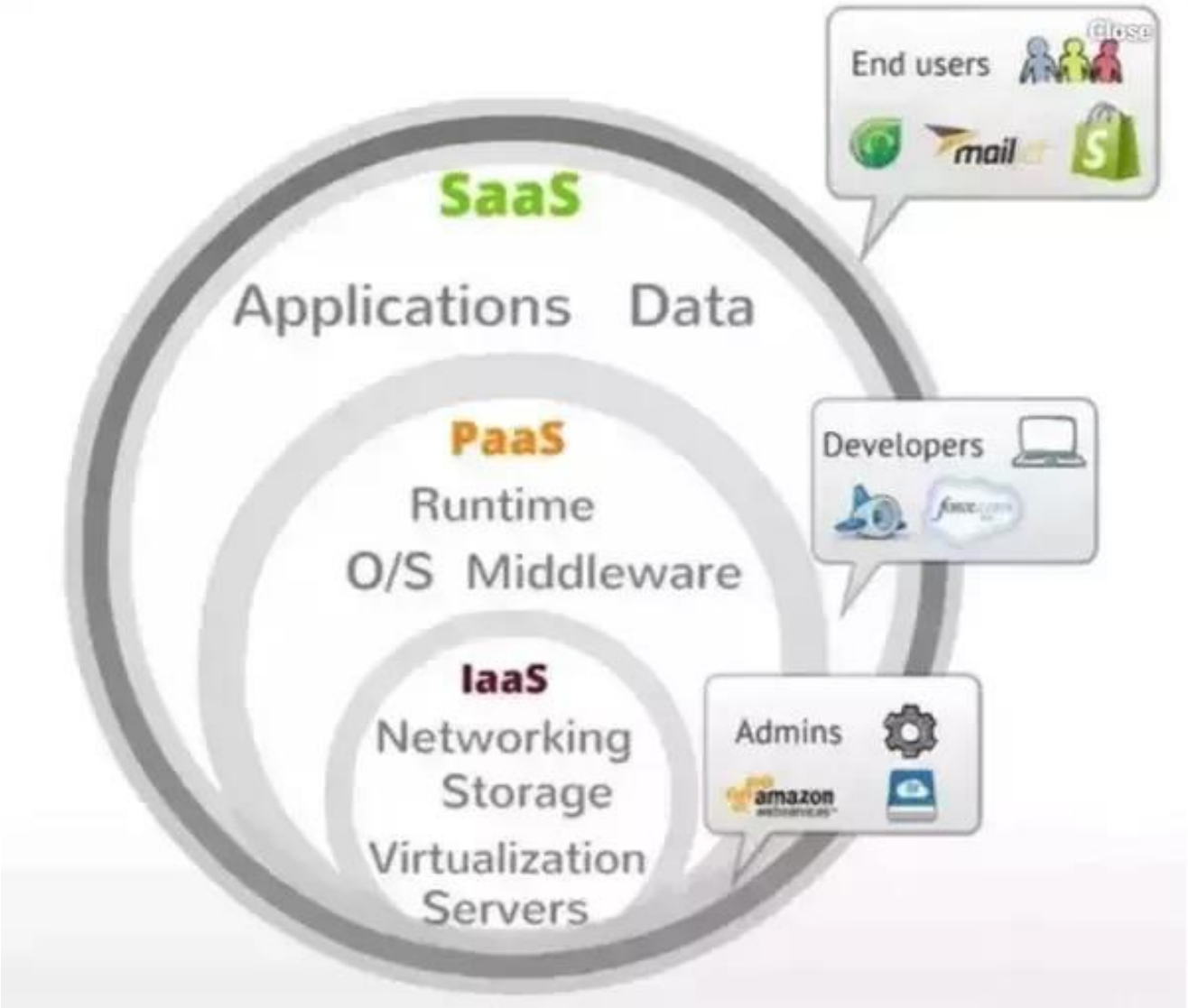
- Cloud providers host and manage the software application on a subscription basis
- Client maintains the control of a software environment but does not maintain any equipment
- Example: Amazon Web Services

Difference between IaaS, PaaS and SaaS



Service Models





Cloud Clients

Apps, Browsers, Mobiles

SAAS

Applications like SharePoint Online, O365

PAAS

Operating system like Windows Azure, Database like SQL Azure, Development tools like NAPA.

IAAS

Windows Azure Virtual Machines and Networks , Storage


aws
IBM Cloud
Google Cloud Platform
CLOUD
COMPUTING
Microsoft Azure
vmware
DigitalOcean

Which cloud service is suitable for you?

IAAS

If your business needs a virtual machine, opt for Infrastructure as a Service

PAAS

If your company requires a platform for building software products, pick Platform as a Service

SAAS

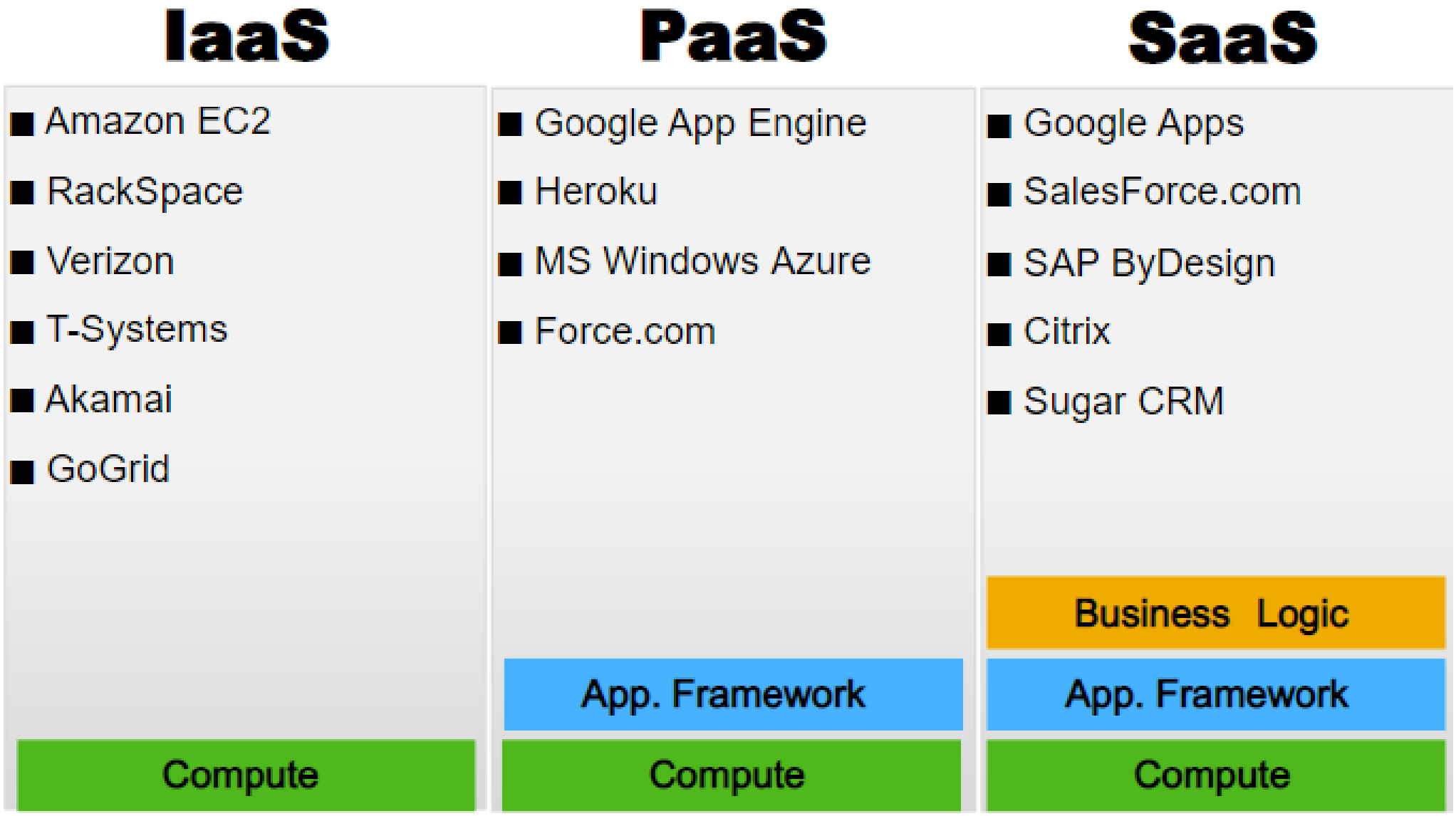
If your business doesn't want to maintain any IT equipment, then choose Software as a Service

On-Premises	IaaS	PaaS	SaaS
Made at Home	Buy & bake	Cake delivery	Dine out
Dinning table	Dinning table	Dinning table	Dinning table
Water	Water	Water	Water
Electricity	Electricity	Electricity	Electricity
Oven	Oven	Oven	Oven
Cake Pan	Cake Pan	Cake Pan	Cake Pan
Flour	Flour	Flour	Flour
Sugar	Sugar	Sugar	Sugar
Butter	Butter	Butter	Butter
Eggs	Eggs	Eggs	Eggs

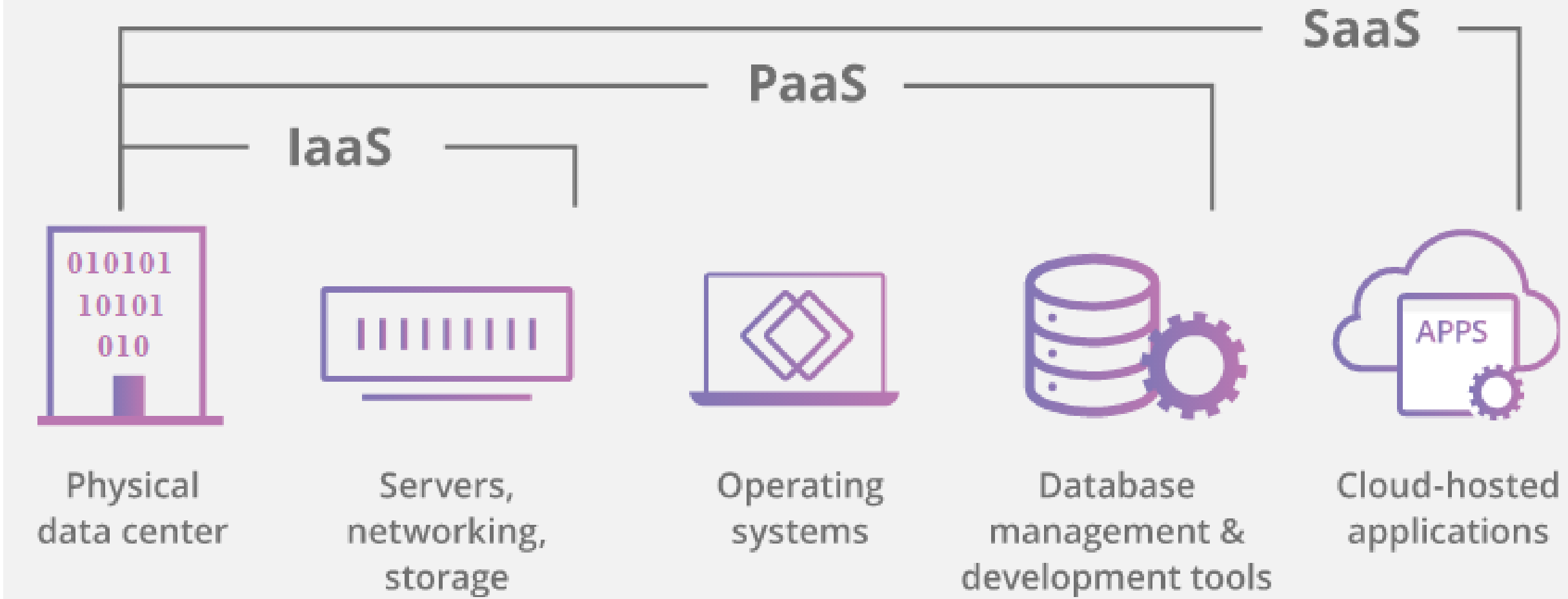
Managed by you

Managed by Vendor

Basic Models > Service Models



	Software-as-a-Service (SaaS)	Platform-as-a-Service (PaaS)	Infrastructure-as-a-Service (IaaS)
Definitions	Cloud provider installs and operates the application software in the cloud where users can access the software from cloud clients.	Cloud provider offers a computing platform including OS, programming language execution environment, database, and web server where developers develop and run software.	Cloud provider offers the physical infrastructure including virtual machines, servers, storage, load balancers, network, and datacenter space with power and cooling.
Control	• Provides controls underlying infrastructure • Provider controls over the configuration of their application	• Provider controls underlying infrastructure • Customer controls over the configuration of their application	• Provider controls the physical assets • Customer controls all the other aspects of infrastructure
Examples	• Salesforce • Google Apps • Microsoft Office 365 • Workday	• Microsoft Azure • Google App Engine • Heroku • Force.com	• Amazon Web Services (AWS) • Rackspace • Go-Grid • OpenStack


SaaS
PaaS
IaaS
010101
10101
010
APPS
Physical data center
Servers, networking, storage
Operating systems
Database management & development tools
Cloud-hosted applications

Public/Internet Clouds

3rd party, multi-tenant Cloud infrastructure & services:

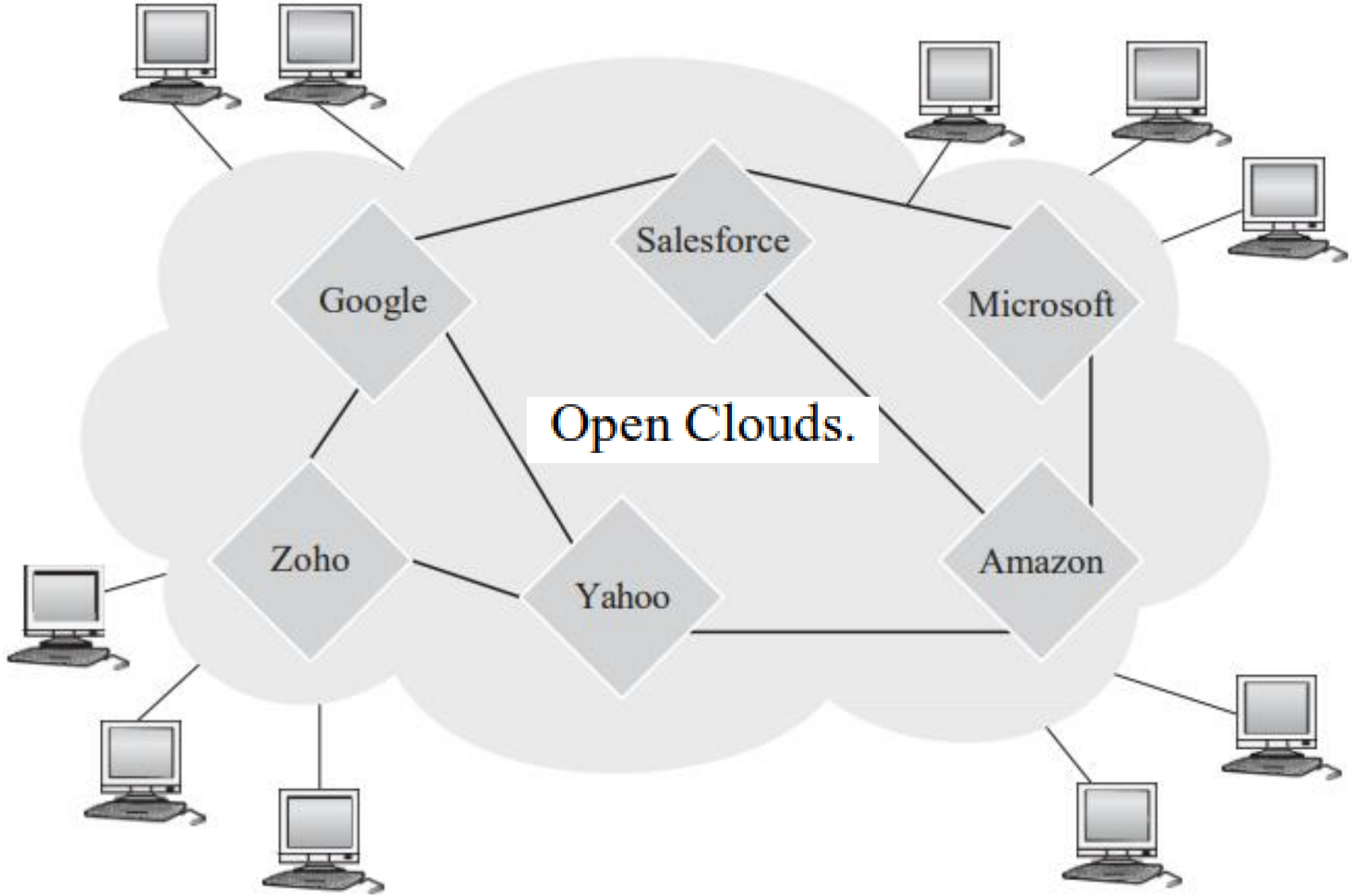
* available on subscription basis (pay as you go)

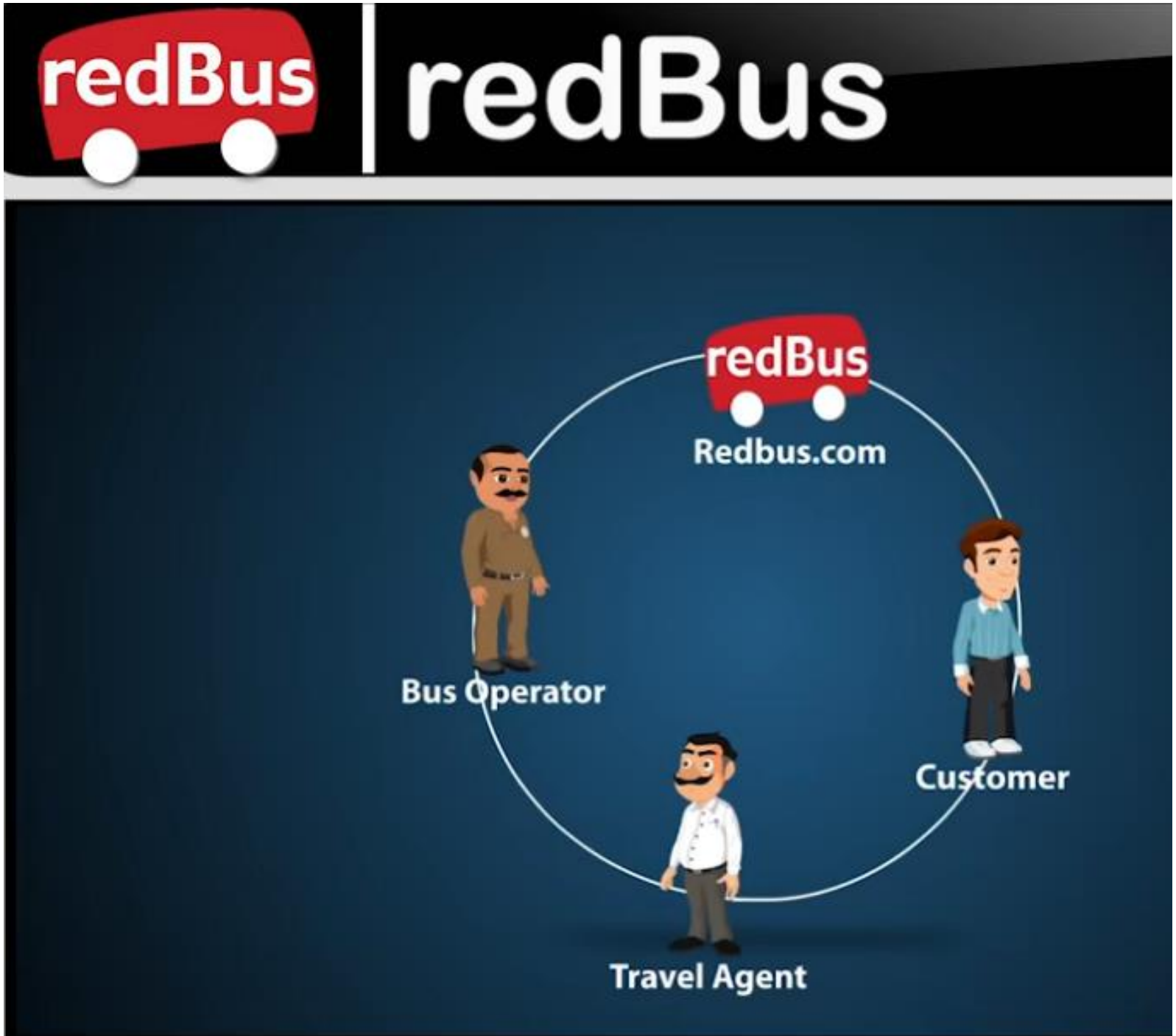
Private/Enterprise Clouds

Cloud computing model run within a company's own Data Center/ infrastructure for internal and/or partners use.

Hybrid/Mixed Clouds

Mixed usage of private and public Clouds: Leasing public cloud services when private cloud capacity is insufficient

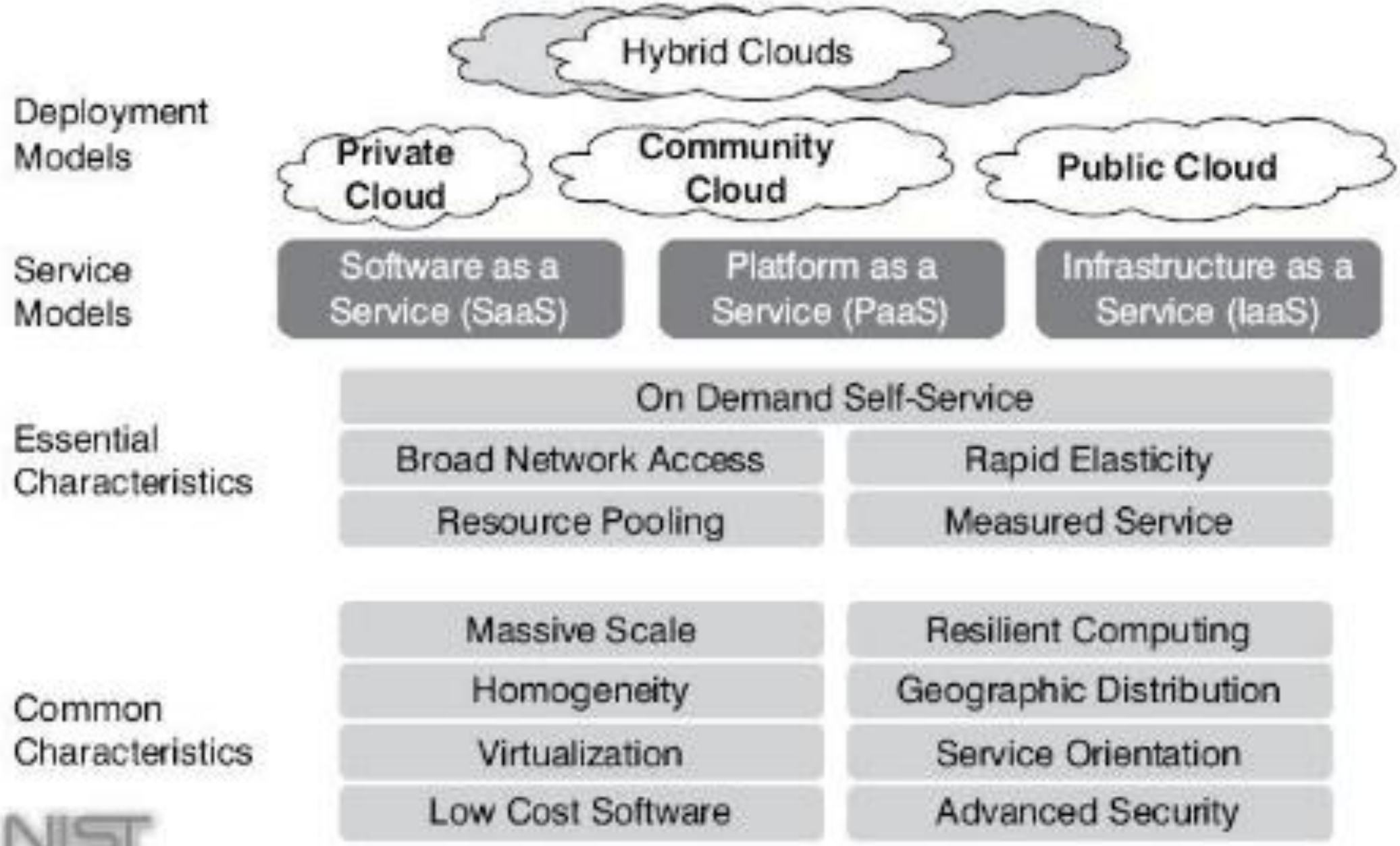

Salesforce
Google
Microsoft
Open Clouds.
Zoho
Yahoo
Amazon


redBus
redBus
redBus
Redbus.com
Bus Operator
Customer
Travel Agent

Cloud Computing – NIST

Definition

- *"Cloud computing is a model for enabling 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."**


Deployment Models
Hybrid Clouds
Private Cloud
Community Cloud
Public Cloud
Service Models
Software as a Service (SaaS)
Platform as a Service (PaaS)
Infrastructure as a Service (IaaS)
Essential Characteristics
On Demand Self-Service
Broad Network Access
Rapid Elasticity
Resource Pooling
Measured Service
Common Characteristics
Massive Scale
Resilient Computing
Homogeneity
Geographic Distribution
Virtualization
Service Orientation
Low Cost Software
Advanced Security
NIST

ADVANTAGES OF CLOUD COMPUTING

- On-demand self-service.
- Broad network access & High Availability.
- Resource pooling
- Rapid elasticity
- Measured Services

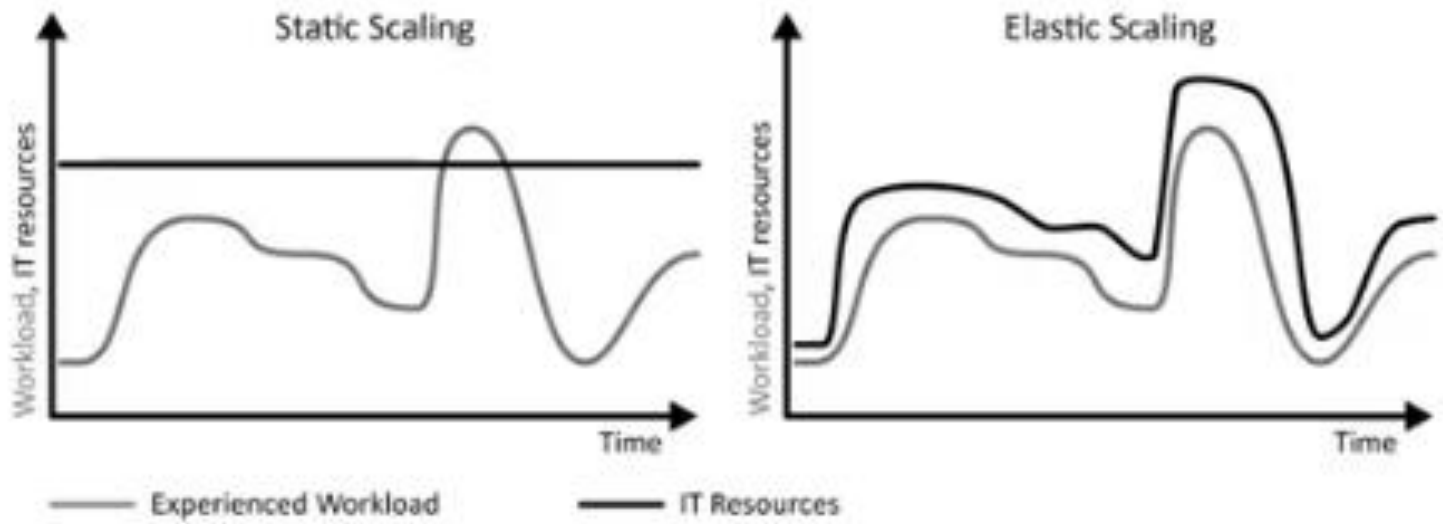
NIST Essential Characteristics

On-demand self-service

- a consumer can *unilaterally* provision computing capabilities *without human interaction* with the service provider
- computing capabilities
  - server time, network storage, number of servers etc.

Elasticity

- Major pattern which benefits from cloud computing
- As your workload changes, resources can be changed to compensate (up or down)
- Example: Seasonal demand for retail web site

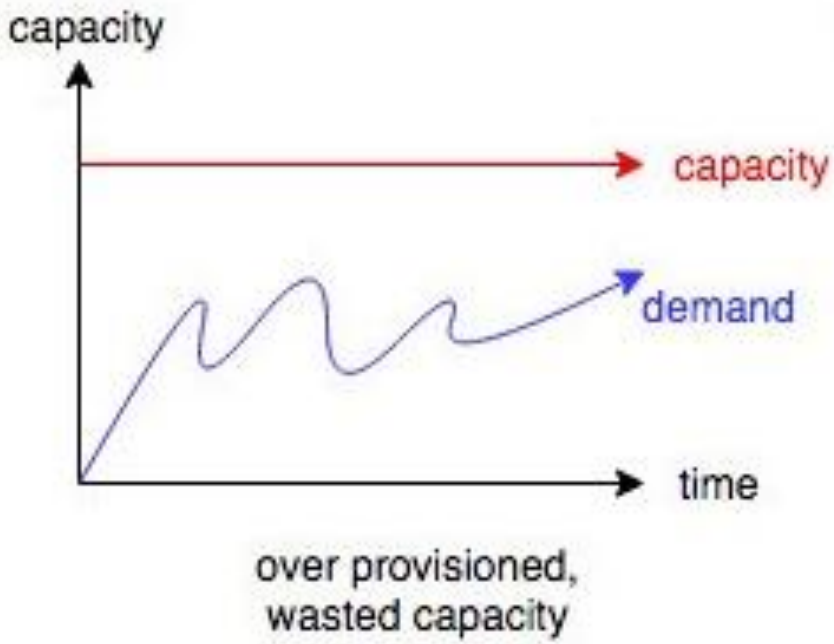


Resource Limit

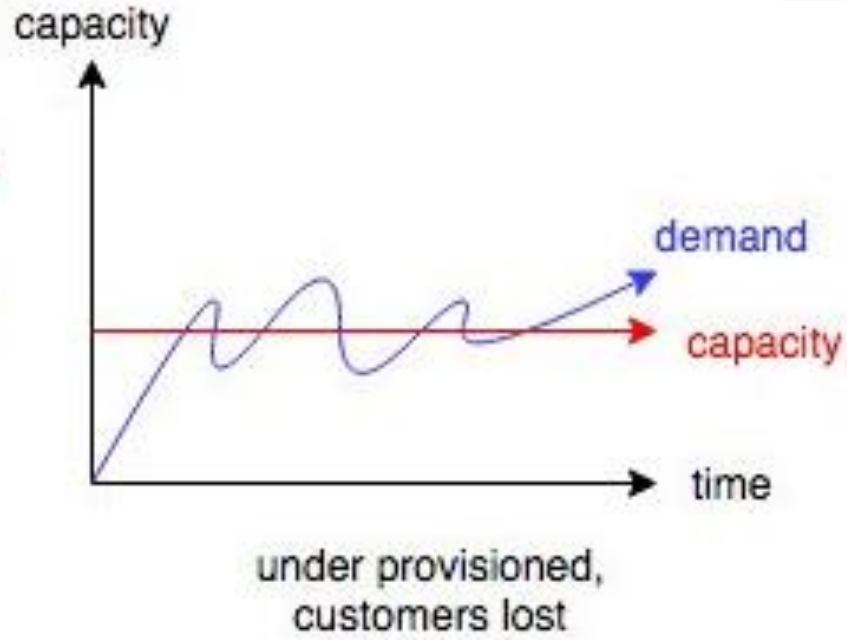
Applications

Unavailable Resources

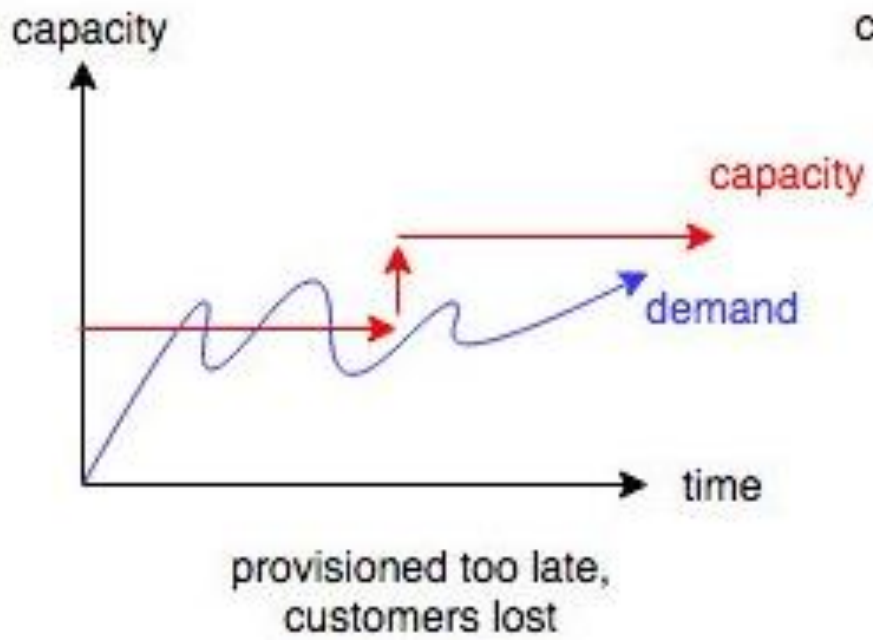
The key factor on Cloud economics



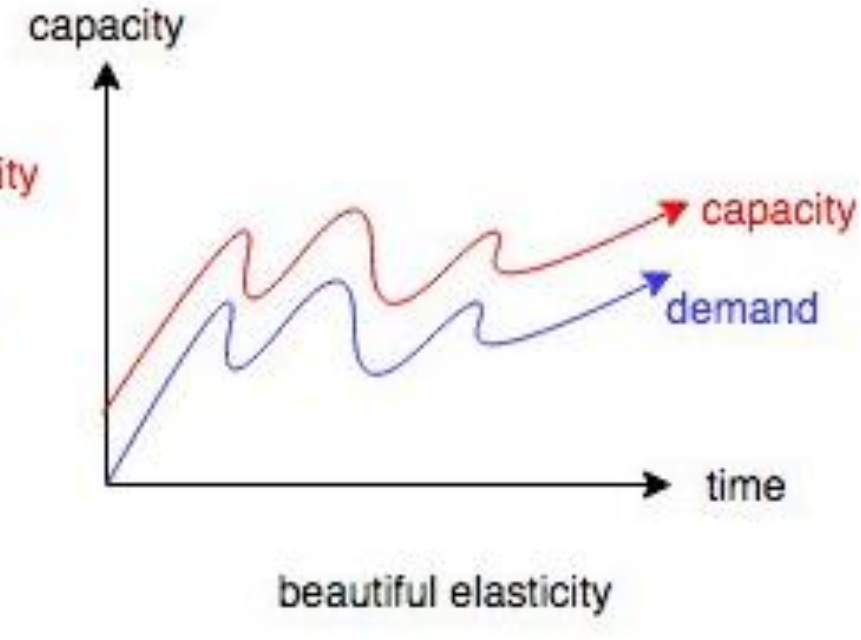
over provisioned,
wasted capacity



under provisioned,
customers lost



provisioned too late,
customers lost



beautiful elasticity

NIST Essential Characteristics

Broad network access

- capabilities are
  - available over the network
  - accessed through standard mechanisms
- promote use by
  - heterogeneous thin or thick client platforms

NIST Essential Characteristics

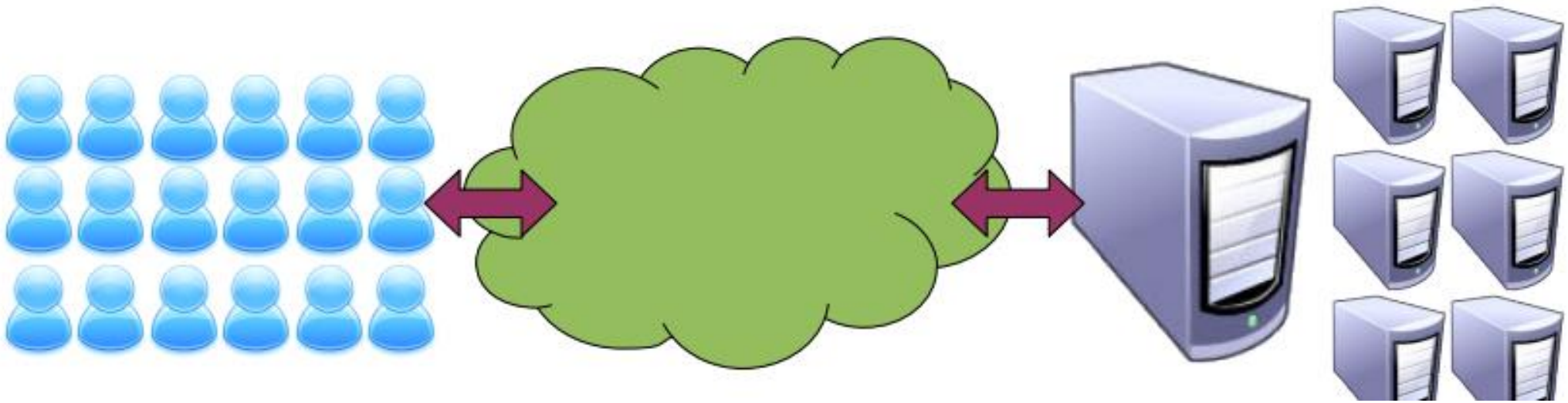
Multi-tenancy / Resource pooling

- provider's computing resources are pooled to serve multiple consumers
- computing resources
  - storage, processing, memory, network bandwidth and virtual machines
- location independence
  - no control over the exact location of the resources
- has major implications

NIST Essential Characteristics

Rapid elasticity

- capabilities can be rapidly and elastically provisioned
- unlimited virtual resources
- predicting a ceiling is difficult

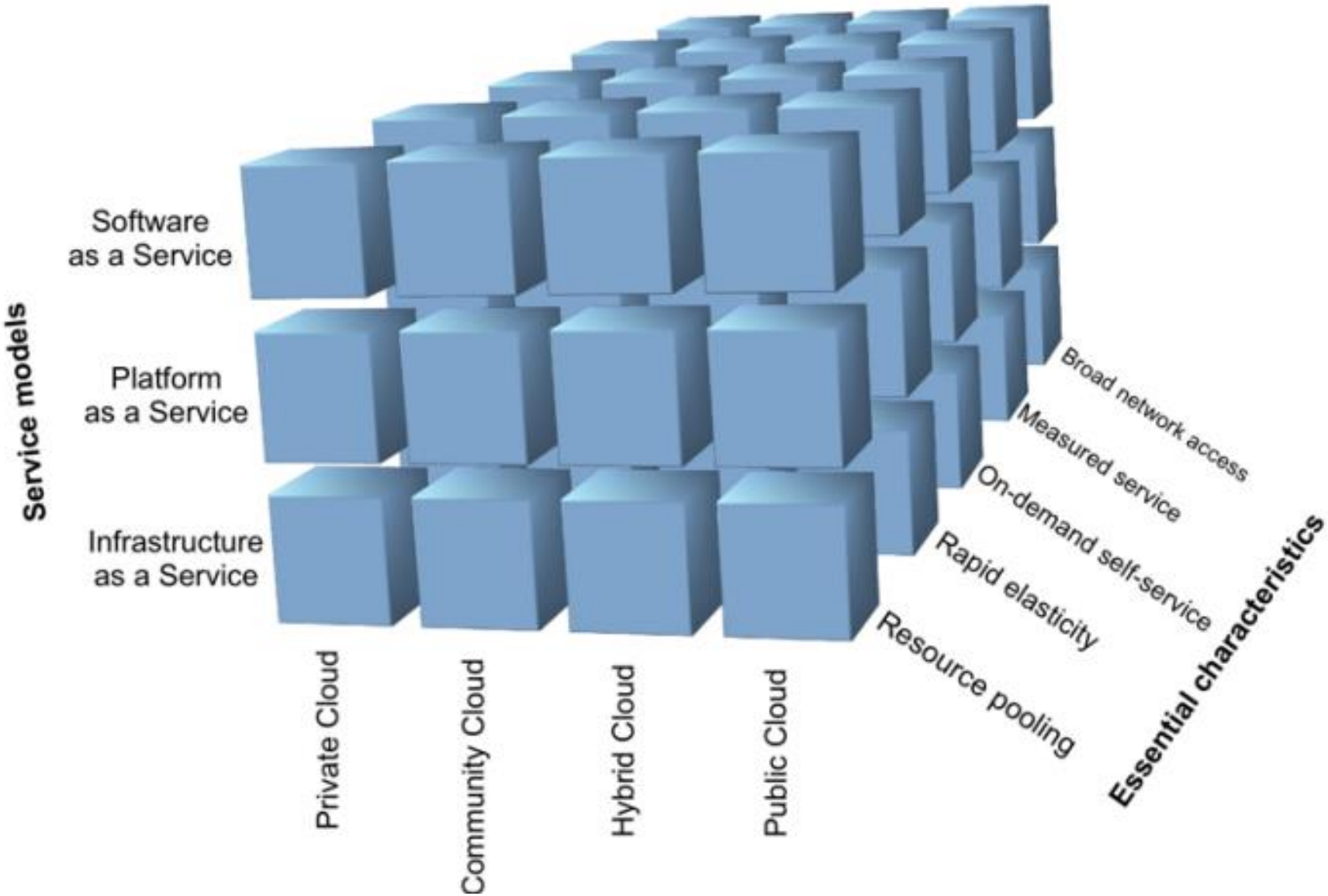


NIST Essential Characteristics

Measured service

- metering capability of service/resource abstractions
  - storage
  - processing
  - bandwidth
  - active user accounts
- **OK so what happened to utility computing – pay as you go model??**
  - more on this later when we discuss deployment models

NIST Cloud Computing



Comparison between Three

	Cluster	Grid	Cloud
On demand Self Service	NO	NO	YES
Broad Network Access	YES	YES	YES
Resource Pooling	Yes	Yes	Yes
Rapid Elasticity	NO	NO	YES
Measured Service	NO	YES	YES

Traditional IT	Cloud Computing
• Hardware is hosted on the premises of the organization and/or manage hosted. • Hardware and software is provisioned for peak demand. • Service management monitoring is used to generate forecasts of demand usage and current SLA performance. • Chargebacks and compensations are used to adjust usage and payments. • Under-provisioning and over-provisioning of capacity can result from unforeseen demand changes. • Business invests in ownership of assets that can be enhanced and extended through IT programs and development. • Changes to IT involve migration and divestment/investment issues and programs.	• Hardware and/or software is hosted off-premise (public or hybrid) or on-premise as a private Cloud service. • Services are provisioned and used based on actual demand, providing this elasticity as a managed service. • Services are typically focused on short-term "burst" demand to gain cost savings over provisioning and owning the assets. • Statistical automated scaling is used to optimize the shared virtual assets. • Risk is transferred from the buyer to the seller/provider of the Cloud service. • Cloud sellers and providers seek to grow amortized economies of scale through increasing the numbers of users of the shared resources. • The IT infrastructure and operation is masked from the service user. Cloud is more than just SaaS.

Client/Server Computing (cont.)

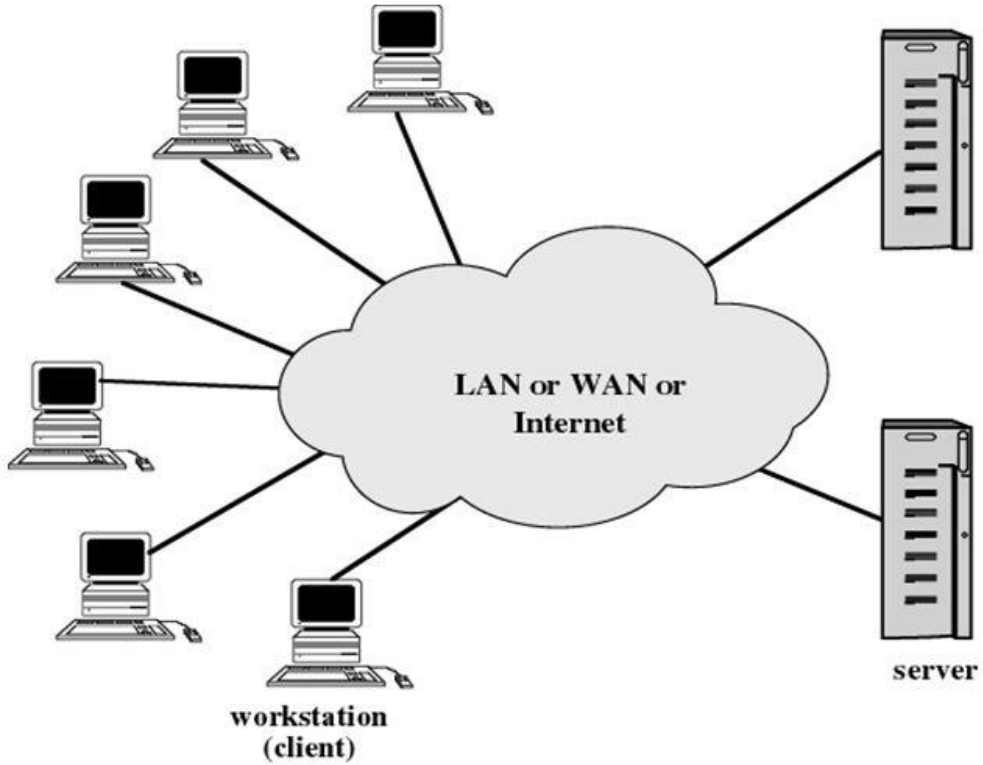
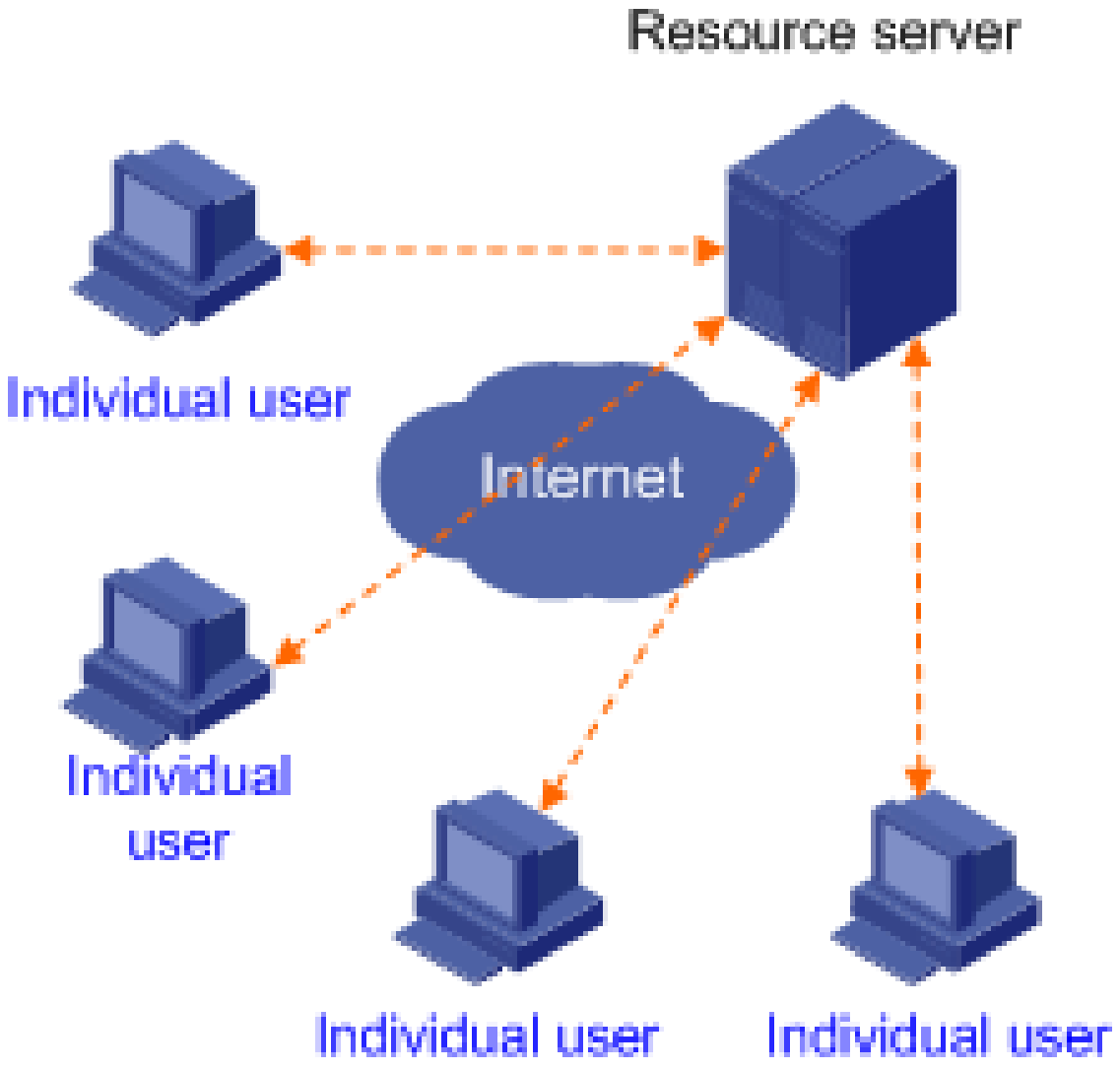
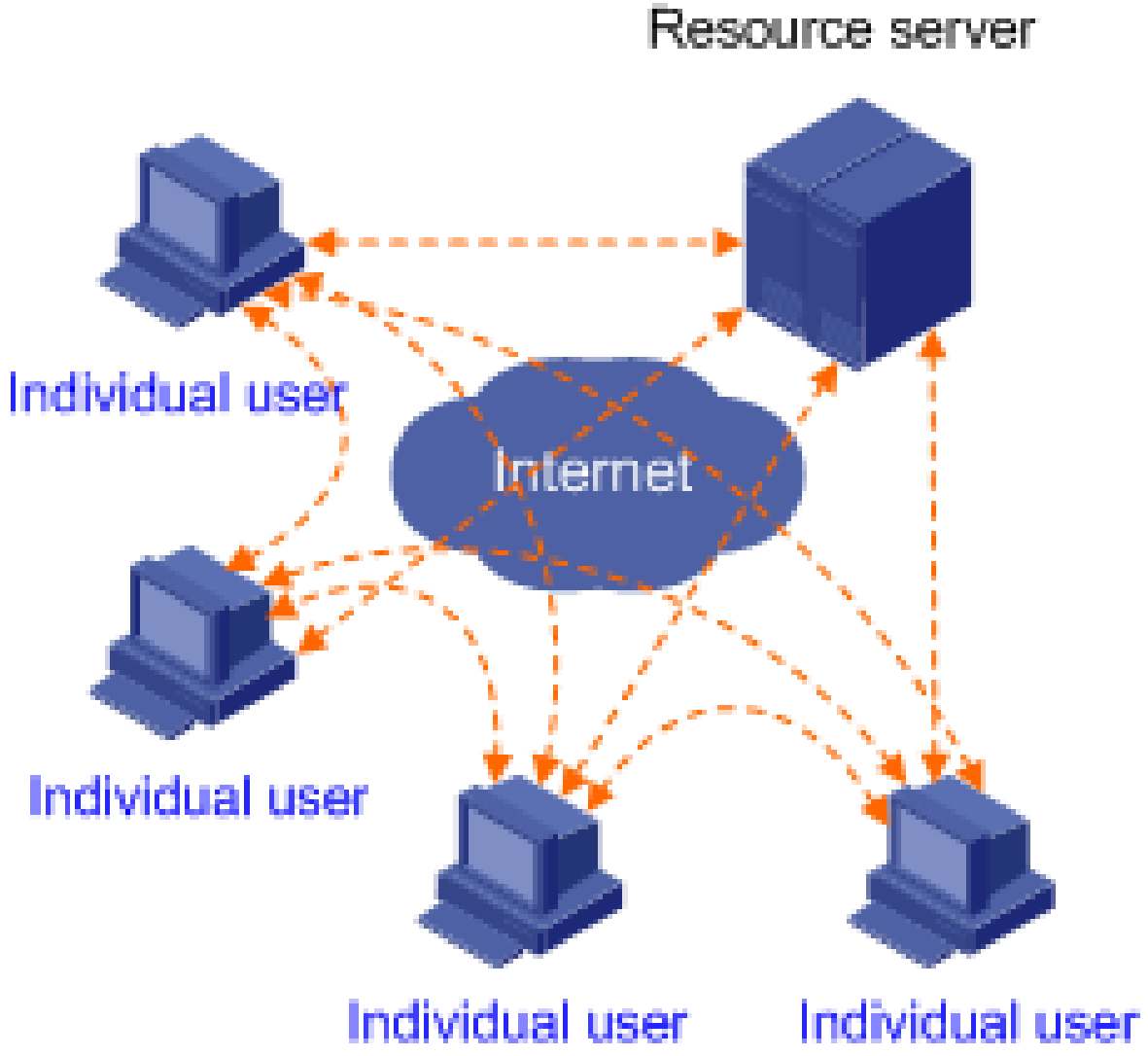


Figure 13.1 Generic Client/Server Environment

Centralized model



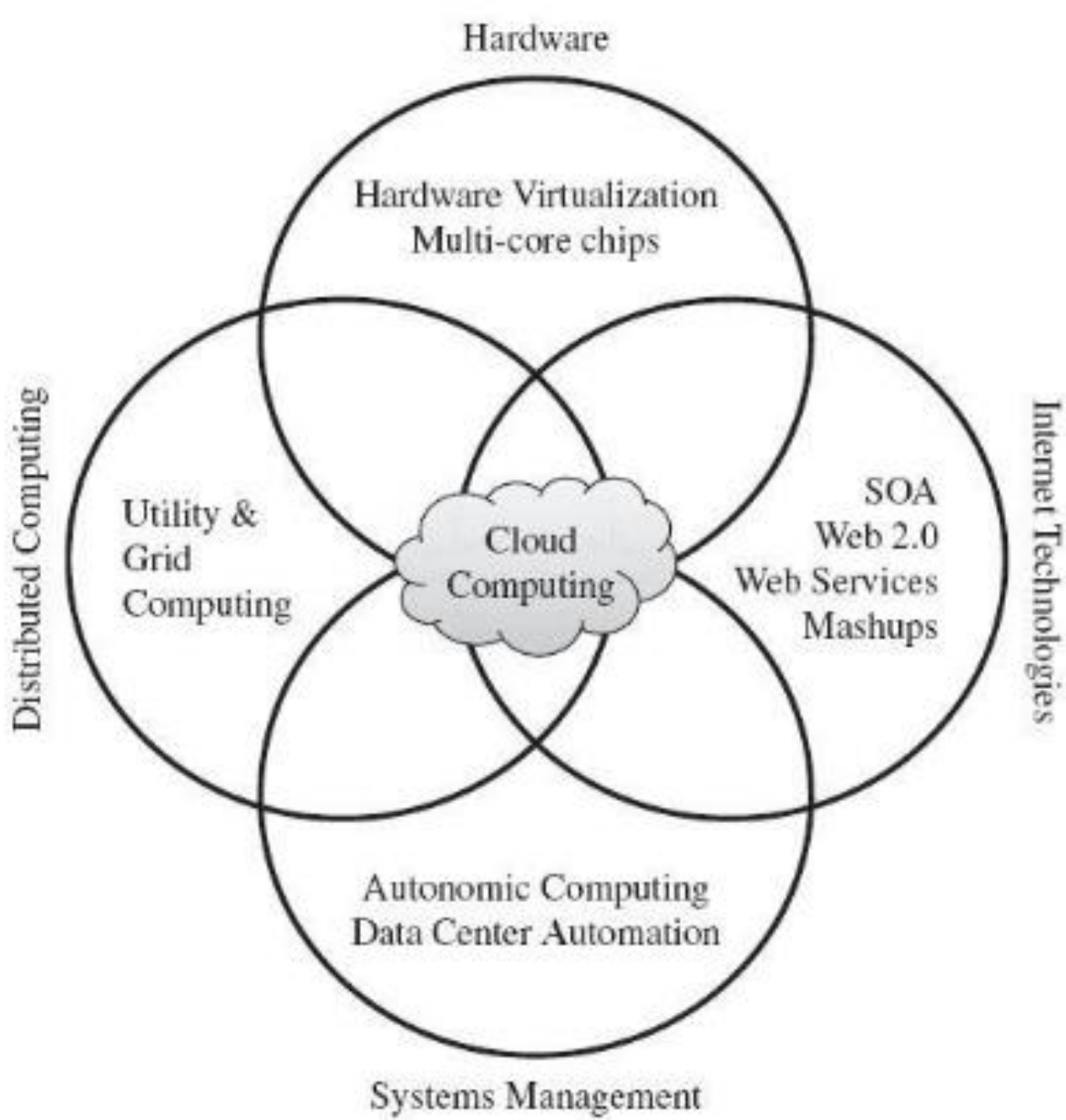
Traditional traffic model



P2P traffic model

Computing Paradigm

- **Centralized computing** This is a computing paradigm by which all computer resources are centralized in one physical system. All resources (processors, memory, and storage) are fully shared and tightly coupled within one integrated OS. Many data centers and supercomputers are *centralized systems*, but they are used in parallel, distributed, and cloud computing applications [18,26].
- **Parallel computing** In parallel computing, all processors are either tightly coupled with centralized shared memory or loosely coupled with distributed memory. Some authors refer to this discipline as *parallel processing* [15,27]. Interprocessor communication is accomplished through shared memory or via message passing. A computer system capable of parallel computing is commonly known as a *parallel computer* [28]. Programs running in a parallel computer are called *parallel programs*. The process of writing *parallel programs* is often referred to as *parallel programming* [32].
- **Distributed computing** This is a field of computer science/engineering that studies distributed systems. A *distributed system* [8,13,37,46] consists of multiple autonomous computers, each having its own private memory, communicating through a computer network. Information exchange in a distributed system is accomplished through *message passing*. A computer program that runs in a distributed system is known as a *distributed program*. The process of writing distributed programs is referred to as *distributed programming*.
- **Cloud computing** An *Internet cloud* of resources can be either a centralized or a distributed computing system. The cloud applies parallel or distributed computing, or both. Clouds can be built with physical or virtualized resources over large data centers that are centralized or distributed. Some authors consider cloud computing to be a form of *utility computing* or *service computing* [11,19].


Hardware
Hardware Virtualization
Multi-core chips
Distributed Computing
Utility &
Grid
Computing
Cloud
Computing
SOA
Web 2.0
Web Services
Mashups
Internet Technologies
Autonomic Computing
Data Center Automation
Systems Management

Grid Computing Vs Cloud Computing

1. **Main objective: –** The main objective of cloud computing is to offer the Service at a lower rate. It also offers scalability and flexibility so that the customer efficiently uses cloud computing with increased security and availability.

However, the grid computing objective is to focus on the network to solve complicated problems; it also provides a computer as a utility.

GRID COMPUTING	CLOUD COMPUTING
Grid computing is for Application-oriented.	Cloud computing is for Service-oriented.
In Grid computing, resources are shared among multiple computing units for processing a single task.	In cloud computing, all the resources are managed centrally and are place over different servers in clusters.
Grid computing is a collection of Interconnected computers and networks that can be called for large scale processing tools.	In cloud computing, more than one computer coordinates to resolve the problem together.
Grid computing is operated within a corporate network	Cloud computing can be accessed via the Internet.
In this, Grids are mainly owned and managed by an organization within its premises.	The cloud servers are owned by infrastructure providers and are placed in physically various locations.
It offers a shared pool of computing resources based on needs.	Cloud computing includes dealing with a common problem using a varying number of computing resources

Region & Number of Availability Zones

US East
N. Virginia (6), Ohio (3)

US West
N. California (3), Oregon (3)

Asia Pacific
Mumbai (2), Seoul (2), Singapore (2), Sydney (3), Tokyo (3)

Canada
Central (2)

China
Beijing (2)

Europe
Frankfurt (3), Ireland (3), London (2)

South America
São Paulo (3)

AWS GovCloud (US-West) (2)

New Region (coming soon)

Bahrain

China

France

Hong Kong

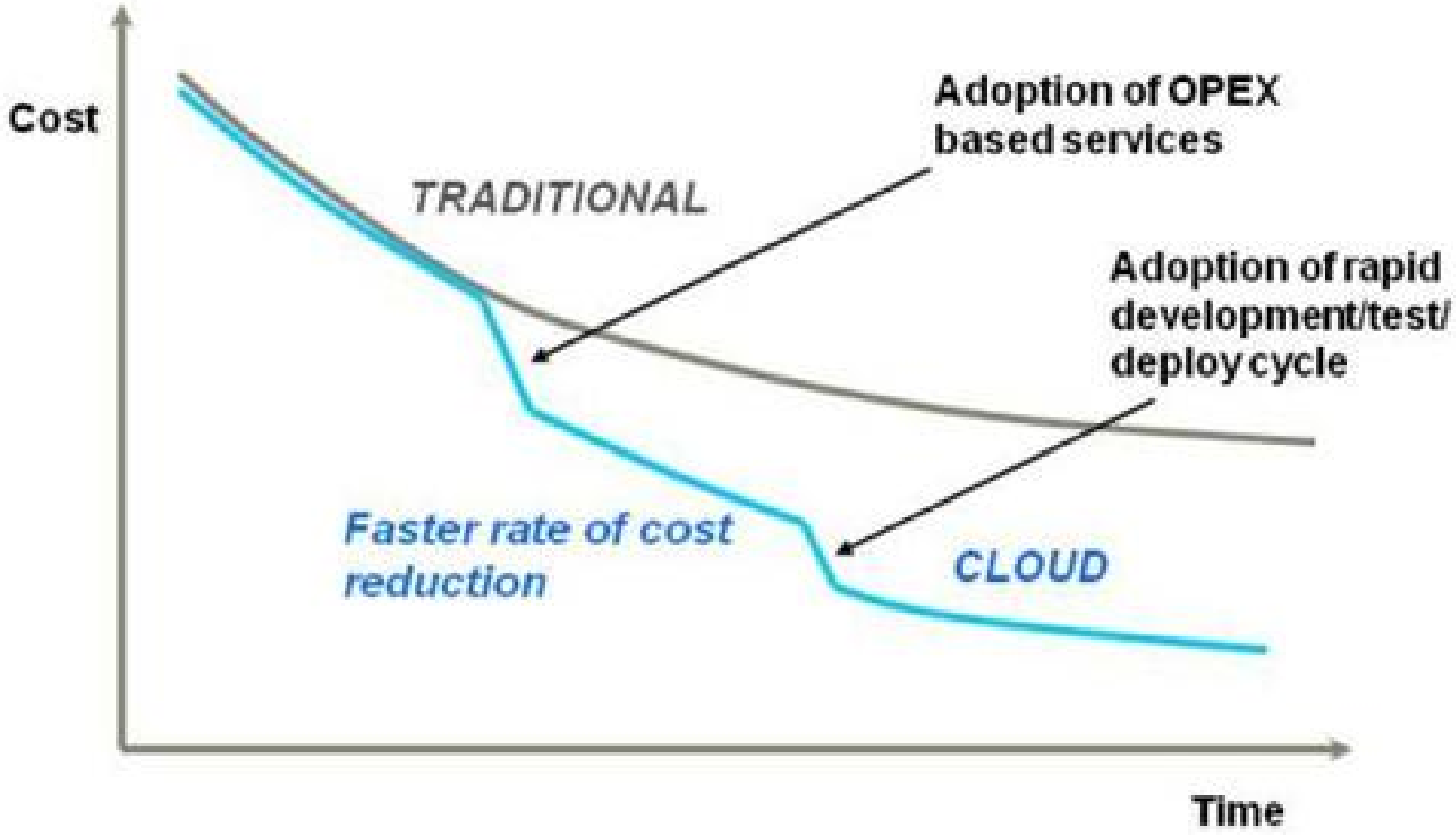
Sweden

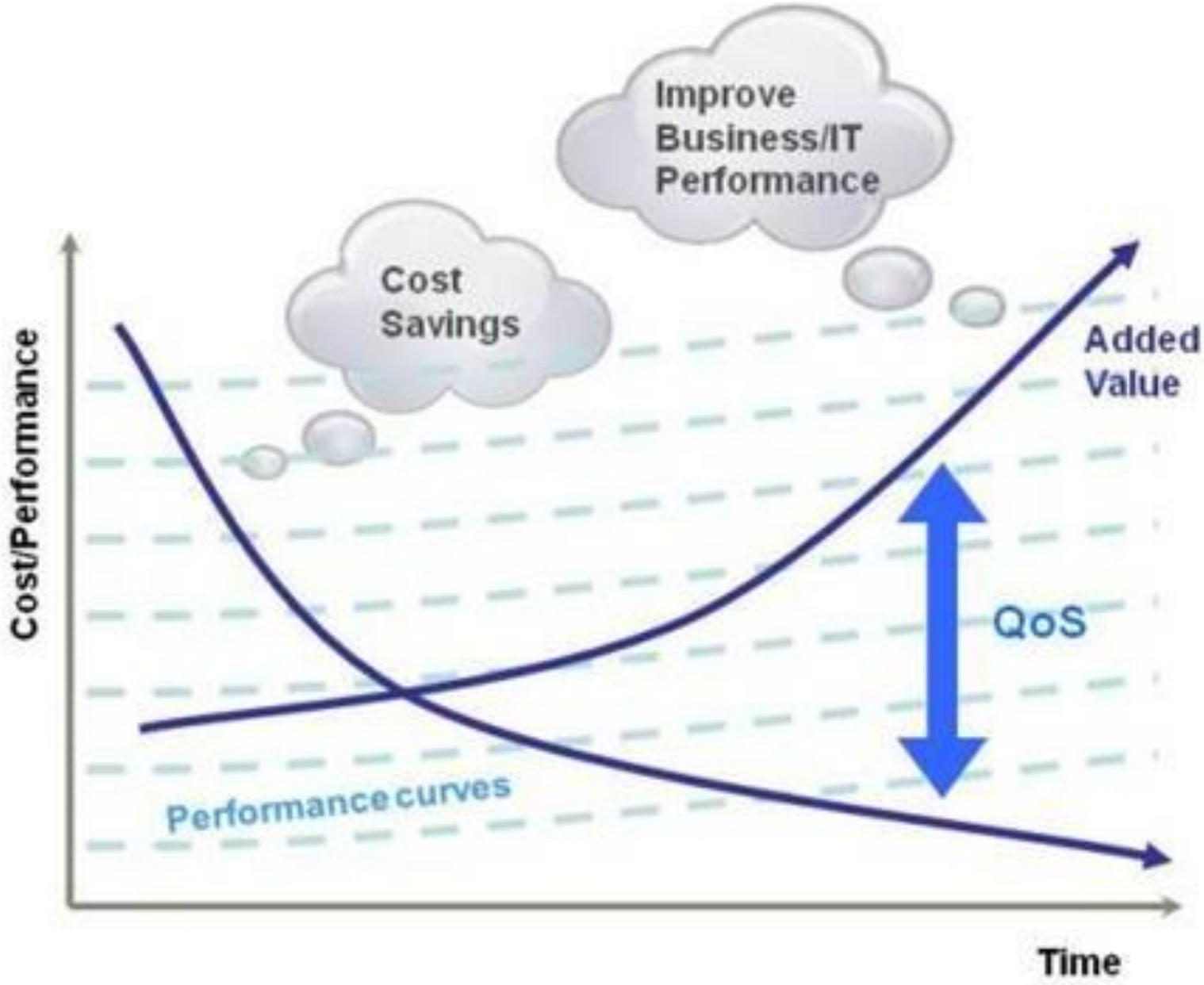
AWS GovCloud (US-East)

16 Regions & 44 Availability Zones till date,

Lifetime Cost Models

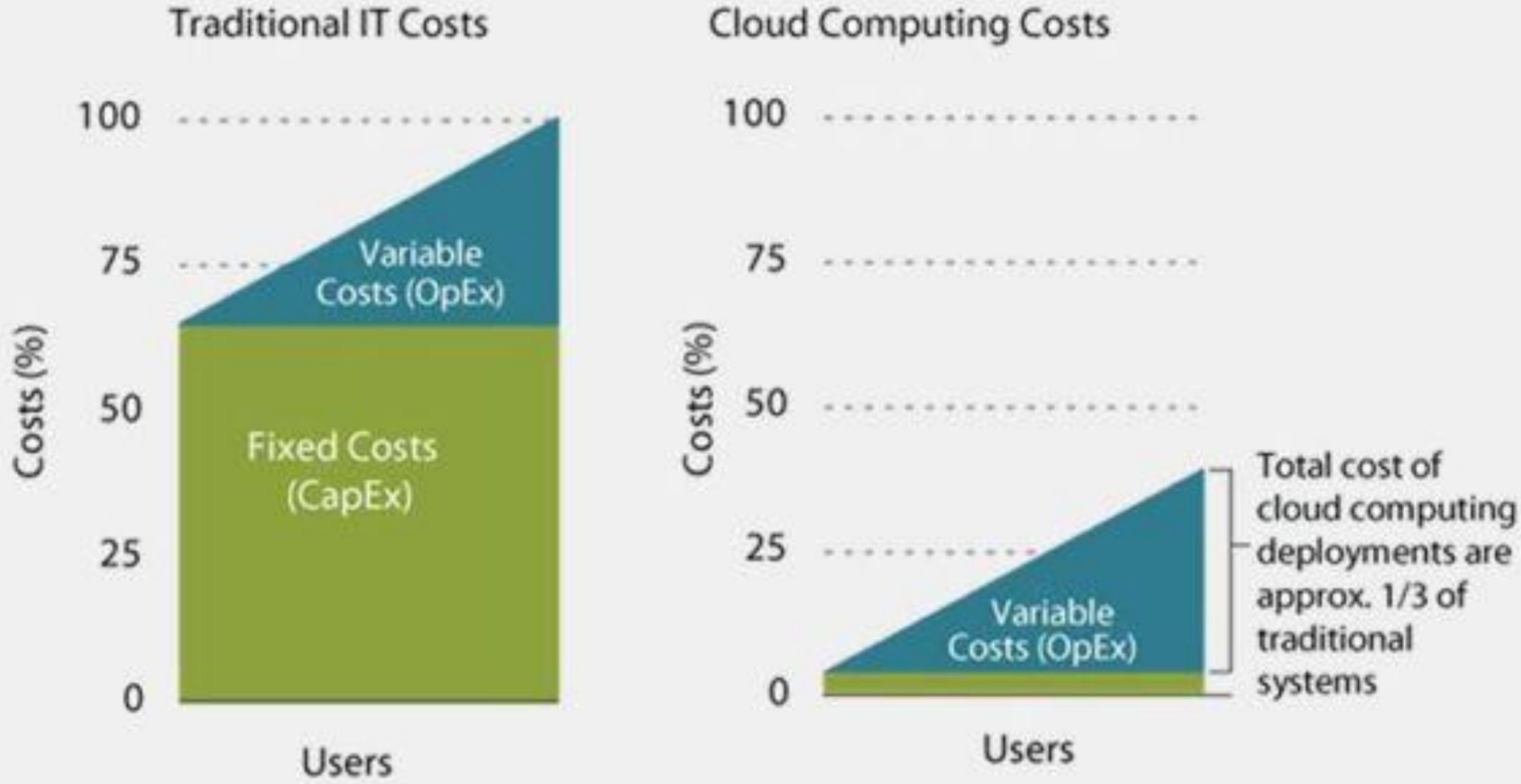
Increased speed of execution has a positive impact on lifetime cost models. Typically, cost is reduced over the lifetime of a product or service as the depreciation cost of purchased assets decreases and as efficiencies are introduced. The speed of cost reduction can be much higher using cloud computing than traditional investment and divestment of IT assets, as illustrated in the figure below.





Competitive Pressure

Traditional IT Costs vs. Cloud Computing Costs



Calculating system availability

System availability is calculated by dividing uptime by the total sum of uptime and downtime.

Availability = Uptime ÷ (Uptime + downtime)

For example, let's say you're trying to calculate the availability of a critical production asset. That asset ran for 200 hours in a single month. That asset also had two hours of unplanned downtime because of a breakdown, and eight hours of downtime for weekly PMs. That equals 10 hours of total downtime.

Here is how to calculate the availability of that asset:

Availability = 200 ÷ (200 + 10)

Availability = 200 ÷ 210

Availability = 0.952

Availability = 95.2%

World-class availability is considered to be 90% or higher.

Calculating availability

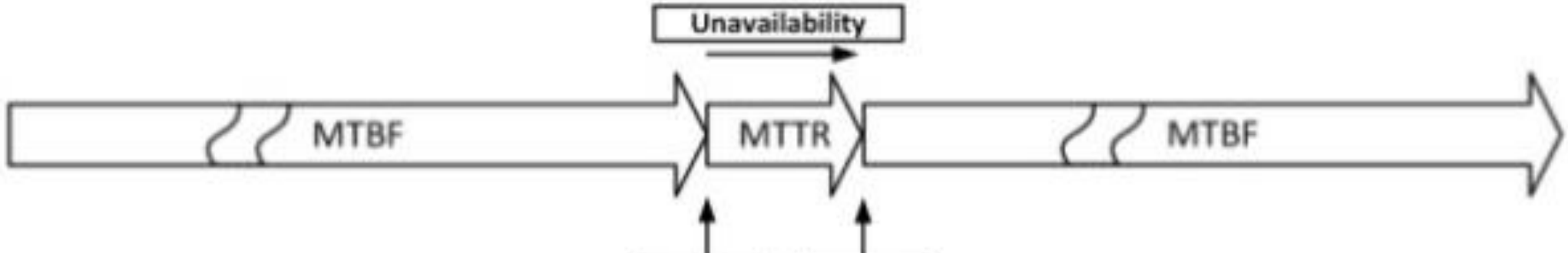
- The availability of a system is usually expressed as a percentage of uptime in a given time period
  - Usually one year or one month

- Example for downtime expressed as a percentage per year:

Availability %	Downtime per year	Downtime per month	Downtime per week
99.8%	17.5 hours	86.2 minutes	20.2 minutes
99.9% ("three nines")	8.8 hours	43.2 minutes	10.1 minutes
99.99% ("four nines")	52.6 minutes	4.3 minutes	1.0 minutes
99.999% ("five nines")	5.3 minutes	25.9 seconds	6.1 seconds

MTBF and MTTR

- Mean Time Between Failures (MTBF)
  - The average time that passes between failures
- Mean Time To Repair (MTTR)
  - The time it takes to recover from a failure



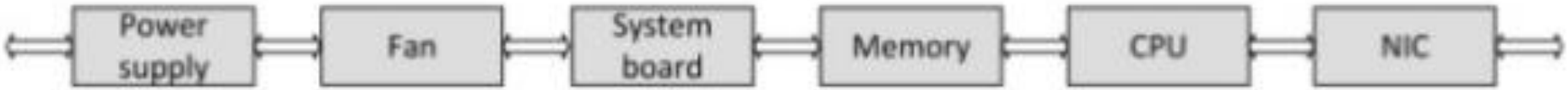
Calculation examples

$$\text{Availability} = \frac{\text{MTBF}}{(\text{MTBF} + \text{MTTR})} \times 100\%$$

Component	MTBF (h)	MTTR (h)	Availability	in %
Power supply	100,000	8	0.9999200	99.99200
Fan	100,000	8	0.9999200	99.99200
System board	300,000	8	0.9999733	99.99733
Memory	1,000,000	8	0,9999920	99.99920
CPU	500,000	8	0.9999840	99.99840
Network Interface Controller (NIC)	250,000	8	0.9999680	99.99680

Calculation examples

- Serial components: One defect leads to downtime



- Example: the above system's availability is:

$$0.9999200 \times 0.9999200 \times 0.9999733 \times 0.9999920 \times 0.9999840 \times 0.9999680 = 0.99977 = \mathbf{99.977}\%$$

(each components' availability is at least 99.99%)

Capacity Utilization Rate Formula = $\frac{\text{Actual output}}{\text{Maximum possible output}} \times 100$

$$Resource\ Utilization = \left(\frac{Busy\ Time}{Available\ Time}\right) \times 100\%$$

Utilization Rate = $\frac{\text{Hours Worked}}{\text{Total Available Hours}}$

$$\frac{\text{Billable hours}}{\text{Available hours}} \text{X100}$$