Welcome to Cloud Computing

Krishna Chakravarty

School of Computer Engineering

Module Outline

- Fundamental concepts of Distributed Systems
- Cluster Computing
- Grid Computing
- Cloud Computing
- Mobile Computing

What is computing?

- **Computing is the process of using computer technology to complete a given goal-oriented task.**
- Computing may encompass the **design and development of software and hardware systems** for a broad range of purposes - often structuring, processing and managing any kind of information
- to aid in the pursuit of scientific studies, making intelligent systems, and creating and using different media for entertainment and communication.
- the use of a computer to process data or perform calculations.
- **It has scientific, engineering, mathematical, technological and social aspects.**
- The list is virtually endless, and the possibilities are vast

What is Distributed sysem?

- **A distributed system, also known as distributed computing, is a system with multiple components located on different machines that communicate and coordinate actions in order to appear as a single coherent system to the end-user.**
- Example -
- Telecommunication networks: Telephone networks and Cellular networks.
- A massively multiplayer online game (MMOG, or more commonly, MMO) is an online game with large numbers of players, often hundreds or thousands
- World wide web

Distributed Systems : A brief introduction

- A collection of independent computers that appears to its users as a single coherent system.
- Features:
  - No shared memory
  - Message-based communication
  - Each runs its own local OS
  - Heterogeneity
- Ideal: to present a single-system image:
  - The distributed system "looks like" a single computer rather than a collection of separate computers.

Distributed Systems : A brief introduction

Distributed system: *the hardware and software of a collection of independent computers that cooperate to realize some functionality*

- –usually, individual parts cannot realize that functionality just by themselves
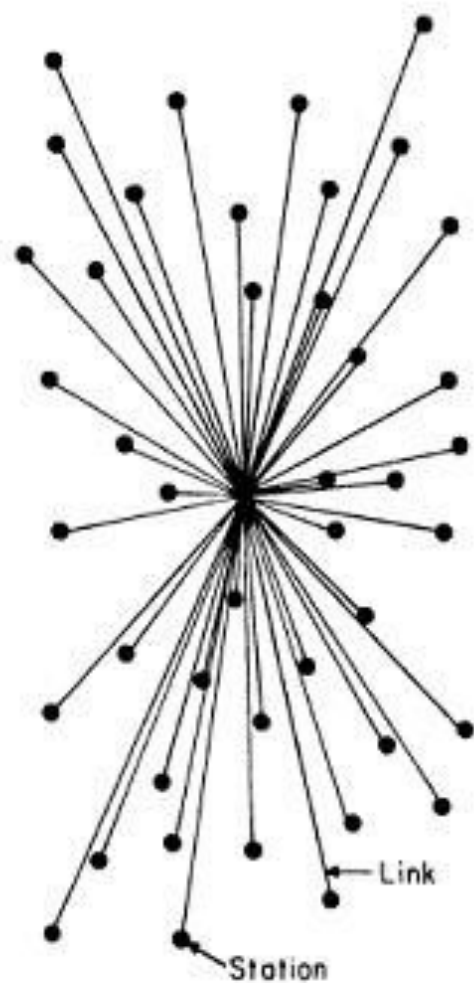- –may be serving a user (…ultimately)

Centralized System Characteristics
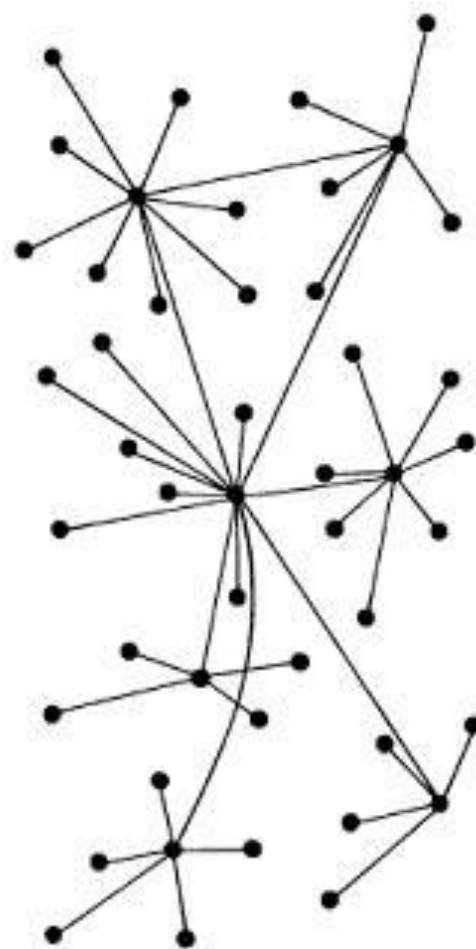## On-Cs-Ra-S-P-P

- One component with Non-autonomous parts

- Components shared by users all the time

- All Resources accessible

- Software runs in a single process

- Single point of control

- Single point of failure

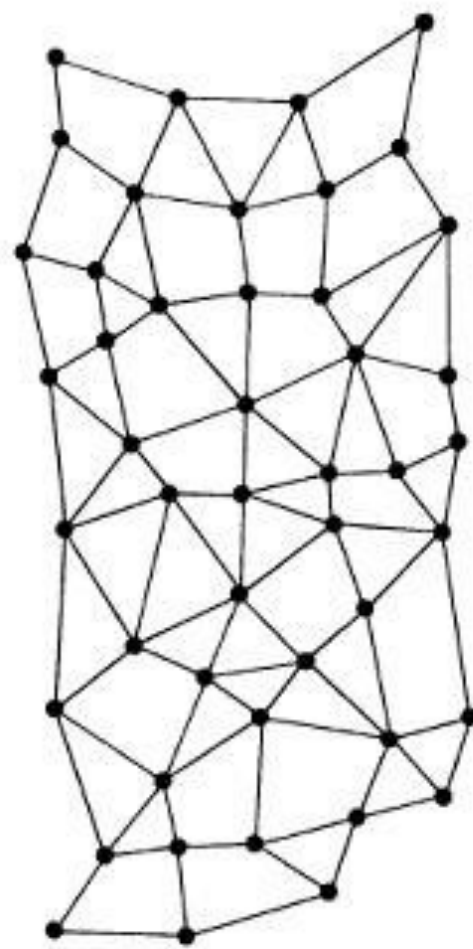# Distributed System Characteristics opposite of Centralised..

- Multiple autonomous components

- Components are not shared by all users

- Resources may not be accessible

- Software runs in concurrent processes on different processors

- Multiple points of control

- Multiple points of failure

CENTRALIZED
(A)

DECENTRALIZED
(B)

DISTRIBUTED
(C)

Distributed System Characteristics (SEAM)

- To present a Single-system image:
  - Hide internal organization, communication details
  - Provide uniform interface
- Easily Expandable
  - Adding new computers is hidden from users
- Continuous Availability
  - Failures in one component can be covered by other components
- Supported by Middleware

Advantages of Distributed System over Centralized System(SIRI)

- **Speed** - When used to implement parallel processing where only goal is to achieve maximum speed on a single problem, distributed systems can achieve very high speed as compared to the centralized ones.

- **Inherent Distribution-** Another reason for building a distributed system is that some applications are inherently distributed. Banking, Airline reservation etc. are examples of the applications that are inherently distributed . When all the branches of a bank are connected, we have a commercial distributed system.

Contd ..

- **Reliability**- By distributing the workload over many machines, a single chip failure will bring down at most one machine, leaving the rest intact.

  For critical applications, such as control of nuclear reactors or aircraft, using a distributed system to achieve high reliability may be a dominant consideration

- **Incremental Growth**- Often a company will buy a mainframe with the intention of doing all its work on it.

  If the company prospers and the workload grows, at a certain point the mainframe will no longer be adequate.
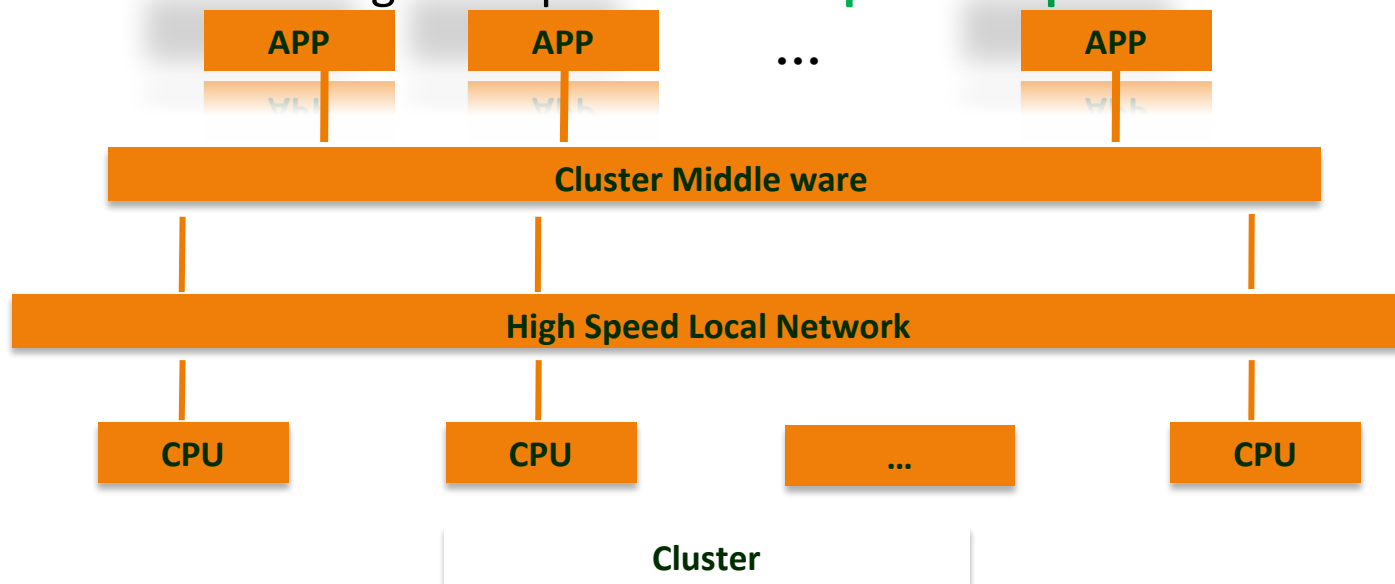
Contd ..

- The only solutions are to either replace the mainframe with a larger one (if it exists), or add a second mainframe. Both of these can cause management difficulties with the company's operations.

- In contrast, with a distributed system, it may be possible to simply add more processors to the system, thus allowing it to expand gradually as the need arises.

Disadvantages of Distributed System

- **Software**: difficult to develop software for distributed systems

- **Control :**Distribution of control

- **Network:** saturation, lossy transmissions

- **Security:** easy access also applies to secret data

- Hard to detect faults

- Administration issues

What is Cluster Computing?

- A **computer cluster** is a **group of linked computers**, **working together** closely so that in many respects they form a single computer.

- The components of a cluster are commonly, but not always, connected to each other through fast local area networks.

- Clusters are usually deployed to improve **performance and/or availability** over that provided by a single computer, while typically being **much more cost-effective** than single computers of **comparable speed.**

APP	APP	...	APP

Cluster Middle ware

High Speed Local Network

CPU	CPU	...	CPU

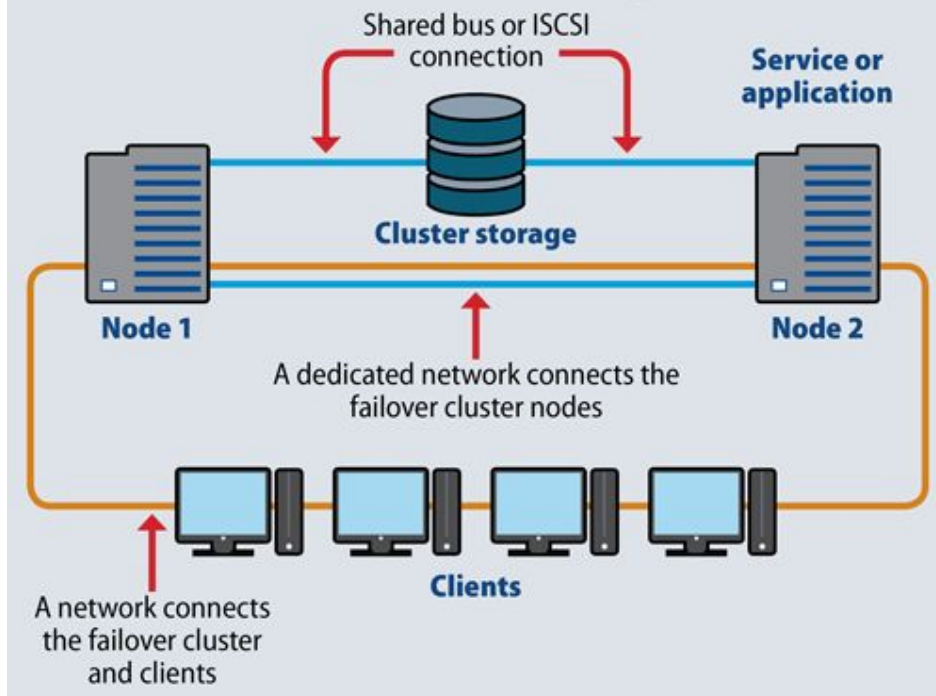**Cluster**

What is Cluster Computing?

Cluster consists of :

- Nodes(master + computing)
- Network
- OS
- Cluster middleware: Middleware such as MPI (Message Passing Interface) which permits compute clustering programs to be portable to a wide variety of clusters

- A cluster has two types of computers - **a master computer, and node computers.**
- When a **large problem** or set of data is **given to a cluster**, the **master computer** first runs a program that **breaks** the problem into small discrete pieces; it then sends a piece to **each node to compute**. As nodes finish their tasks, the master computer **continually sends more pieces to them** until the entire problem has been computed.
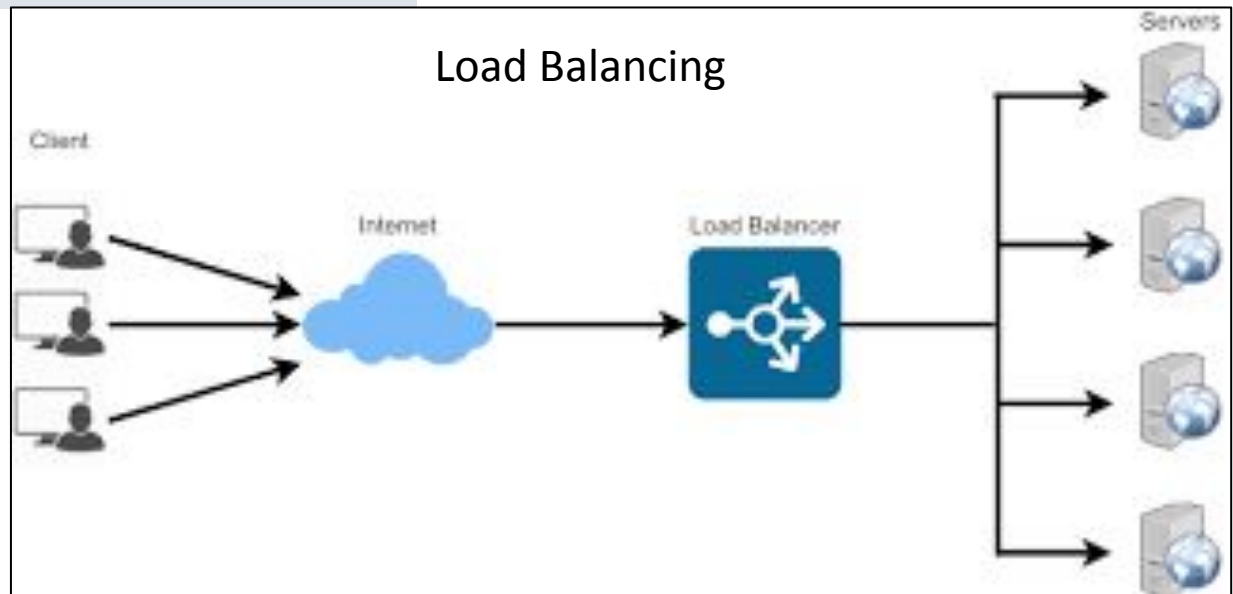
Types of cluster computing

- <u>High Availability (HA) and failover clusters</u> –
1. These models are built to provide an availability of **services and resources in an uninterrupted manner** through the **use of implicit redundancy to the system**.
2. If a cluster node fails (failover), applications or services may be **available in another node**.
3. These types are used to cluster **data base** of **critical missions, mail, file and application servers.**
- <u>Load balancing (LB)</u> –
1. This model **distributes incoming traffic** or **requests for resources from nodes** that **run the same programs** between machines that make up the cluster.
2. If a **node fails**, the requests **are redistributed** among the nodes available.
3. This type of solution is usually used for **Web servers** .

Failover Clustering

Shared bus or ISCSI connection

Service or application

Cluster storage

Node 1

Node 2

A dedicated network connects the failover cluster nodes

A network connects the failover cluster and clients

Clients

Load Balancing

Client

Internet

Load Balancer

Servers

Types of cluster computing

- <u>HA &LB combination</u> –

1. As its name says, it combines the features of both types of cluster, thereby **increasing the availability and scalability of services and resources**.

2. This type of cluster configuration is widely used in **web, email, news, or ftp servers.**

Cluster Computing

Advantages
- Increasing speed – better performance
- optimized resources utilization
- Can execute Large applications

Disadvantages
- Complex programming models
- Difficult for debug and development

Grid Computing

- **Grid Computing** was first coined in 1990s by Cart Kesselman & Ian Foster and defined as –

  "*Grid computing is* **coordinated resource sharing and problem solving** *in* **dynamic, multi-institutional virtual organizations.**"
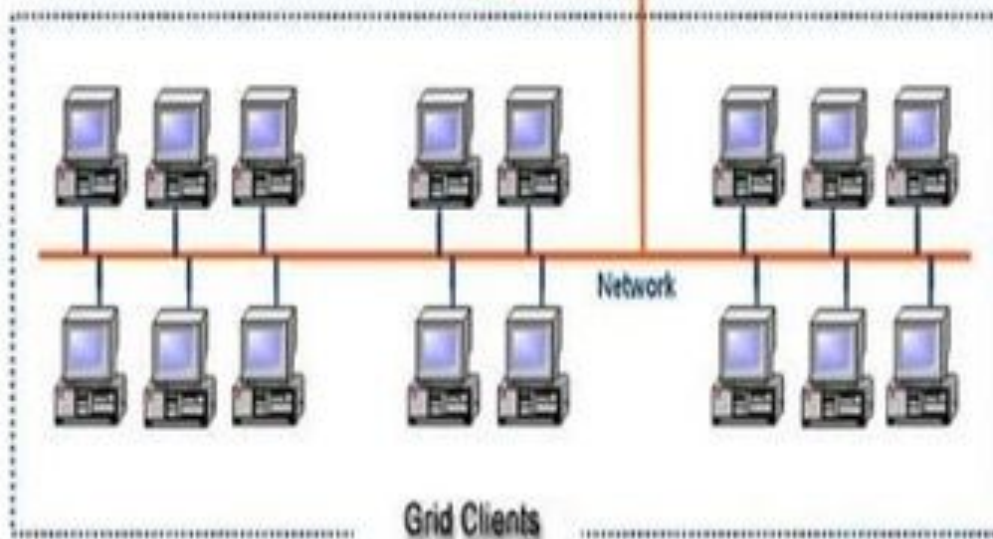
-

Grid Computing

- Computing grids are conceptually and logically like electrical grids. In an electrical grid, wall outlet allow us to connect to an infrastructure of resources which generate and distribute the electricity. Likewise, in the IT industry, grid computing uses middleware to coordinate distinct IT resources over the network, allowing them to   function and work as a virtual whole.

- The goal of a computing grid , like that of a electrical grid, is to provide users with access to the resources they need, when they need them and to provide remote access to IT assets and aggregating processing power.

- Grid provides a series of distributed computing resources via LAN or WAN to the terminal user's application, as if he is using a super virtual computer.

- Grid computing is the integration of computer resources for achieving similar objective.
- Grids are frequently created with middleware software libraries of a common grid.
- Grids are a type of dispersed computing system, whereas a **virtualized super computer** is made from various networked.
- Grids focus on two different but associated objectives—**supplying isolated access to IT resources and building up processing control.**
- The grid is a technology which controls two factors—**allocation and trust.**

How Grid computing works ?

Users

Grid Server

Network

Grid Clients

In general, a grid computing system requires:

- **At least one computer, usually a server, which handles all the administrative duties for the System**
- **A network of computers running special grid computing network software.**
- **A collection of computer software called middleware**

INTRODUCTION:

* These technology opportunities have led to the possibility of using distributed computers as a single, unified computing resource, leading to what is popularly known as Grid computing.

* The term *Grid* is chosen as an analogy to a Power Grid that provides *consistent, pervasive, dependable, transparent* access to electricity irrespective of its source.

* This new approach to network computing is known by several names, such as *metacomputing, scalable computing, global computing, Internet computing*, and more recently *peer to peer (P2P) computing*.

THE GRID PROBLEM:

- The real and specific problem that underlies the Grid concept is *coordinated resource sharing and problem solving in dynamic, multi-institutional virtual organization*s.

- The sharing that we are concerned with is not primarily file exchange but rather direct access to computers, software, data, and other resources.

- This sharing is, necessarily, highly controlled, with resource providers and consumers defining clearly and carefully just what is shared, who is allowed to share, and the conditions under which sharing occurs.

- Grid Computing is a subset of distributed computing, where a virtual super computer comprises of machines on a network connected by some bus, mostly Ethernet or sometimes the Internet.

Components in Grid

A Grid computing network mainly consists of these three types of machines

Control Node:

- A computer, usually a server or a group of servers which administrates the whole network and keeps the account of the resources in the network pool.

Provider:

- The computer which contributes it's resources in the network resource pool.

User:

- The computer that uses the resources on the network.

For controlling the network and it's resources a software/networking protocol is used generaly known as Middleware. This is responsible for administrating the network and the control nodes are merely it's executors. As a grid computing system should use only unused resources of a computer, it is the job of the control node that any provider is not overloaded with tasks.

- Another job of the middleware is to authorize any process that is being executed on the network. In a grid computing system, a provider gives permission to the user to run anything on it's computer, hence it is a huge security threat for the network. Hence a middleware should ensure that there is no unwanted task being executed on the network.
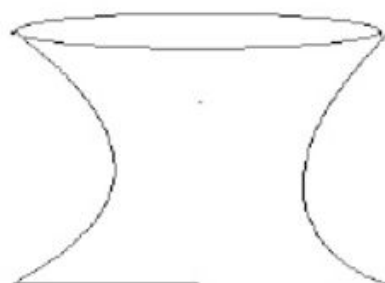
THE GRID ARCHITECTURE DESCRIPTION

- Our goal in describing our Grid architecture is not to provide a complete enumeration of all required protocols but rather to identify requirements for general classes of component.
- Our architecture and the subsequent discussion organize components into layers, as shown in Figure.
- Components within each layer share common characteristics but can build on capabilities and behaviors provided by any lower layer.
- In specifying the various layers of the Grid architecture, we follow the principles of the "hourglass model".

Applications
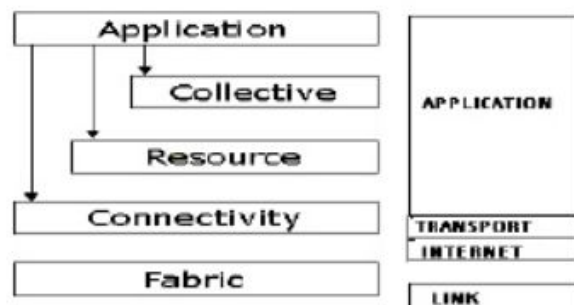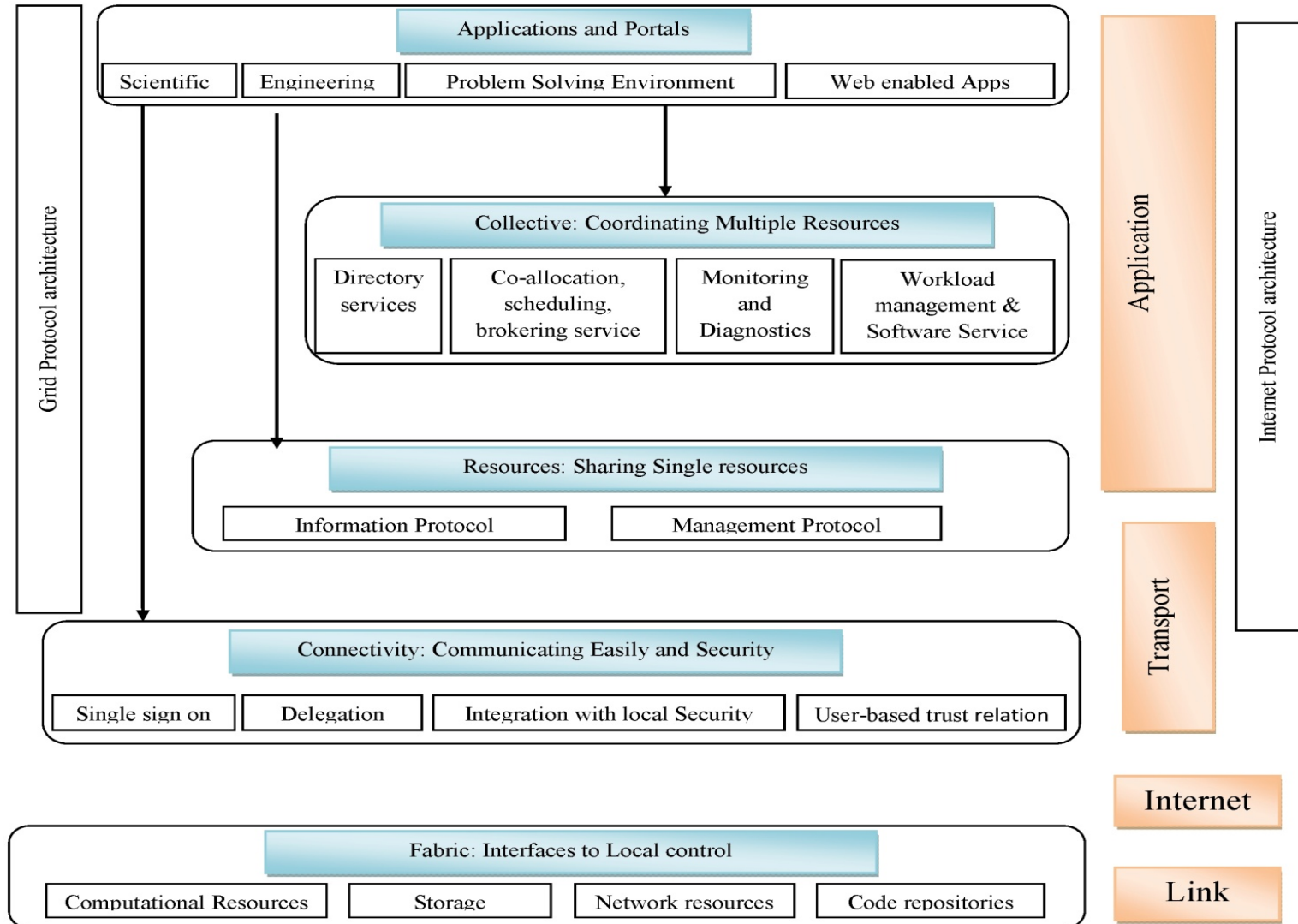
Diverse global

services

Local OS

Core services

Application
Collective
Resource
Connectivity
Fabric

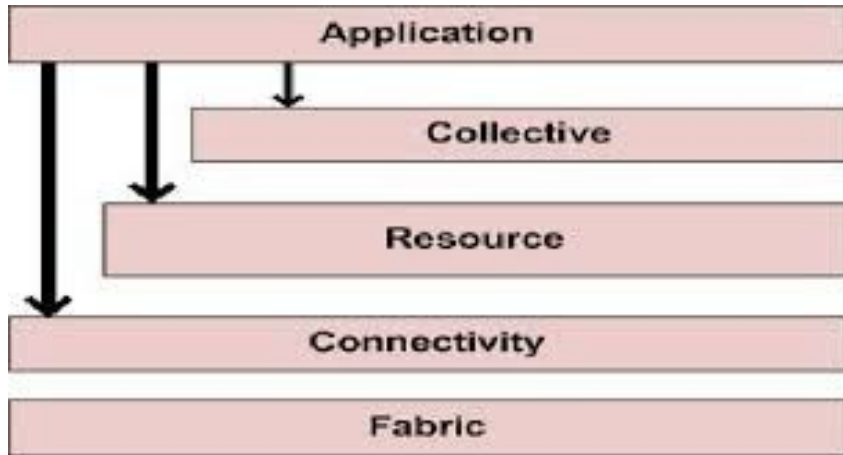| APPLICATION |
|-------------|
| TRANSPORT |
| INTERNET |
| LINK |

Architecture of Grid
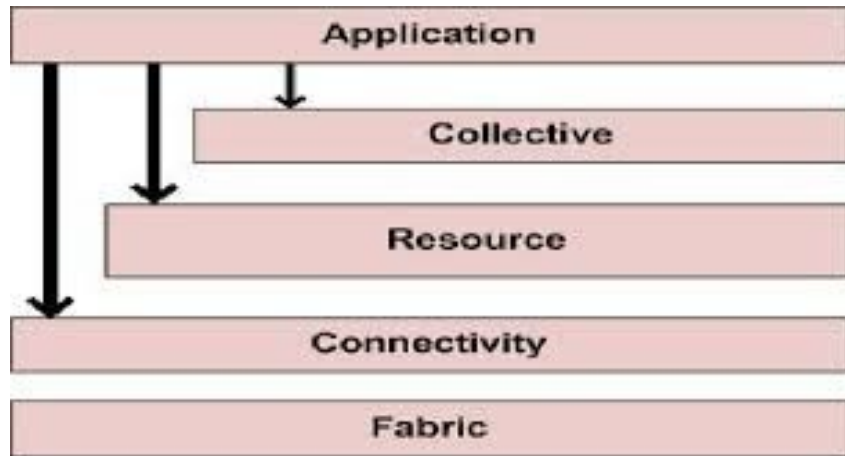
Grid Computing

Layers of Grid System architecture are as follows:

(i)Fabric Layer: Lowest layer

i.   It comprises geographically **distributed and sharable resources** such as **storage devices, computing devices, high bandwidth networks and scientific instruments, etc.**

ii.   These resources are **accessible from anywhere** in the internet.

(ii)Connectivity Layer:

i.   It provides an **easy and secure access** to the resources through **some standardized protocols.**

ii.   The **authentication protocols** are used to provide proper **verification of users and resources.**

iii.   The **communication protocols** provide a mechanism for **transmission of data between the fabric and the resource layer.**

Cont …

(iii)Resource Layer:
i.   It specifies the **necessary protocols** to operate **among shared resources.**
ii.   It is primarily **built** upon the **communication and authentication protocols of connectivity layer.**
iii.   It defines the **APIs** (Application Program Interfaces) and **SDK** (Software Development Kit) to make **secure the initiation, accounting, negotiation, controlling and monitoring of resources.**

(iv)Collective Layer:
i.   It is used for common functional utilities.
ii.   The operations such as **collaboration among the shared resources** are performed in this layer.
iii.   It provides services like brokering, monitoring, scheduling, discovery, replication and co-allocation etc.

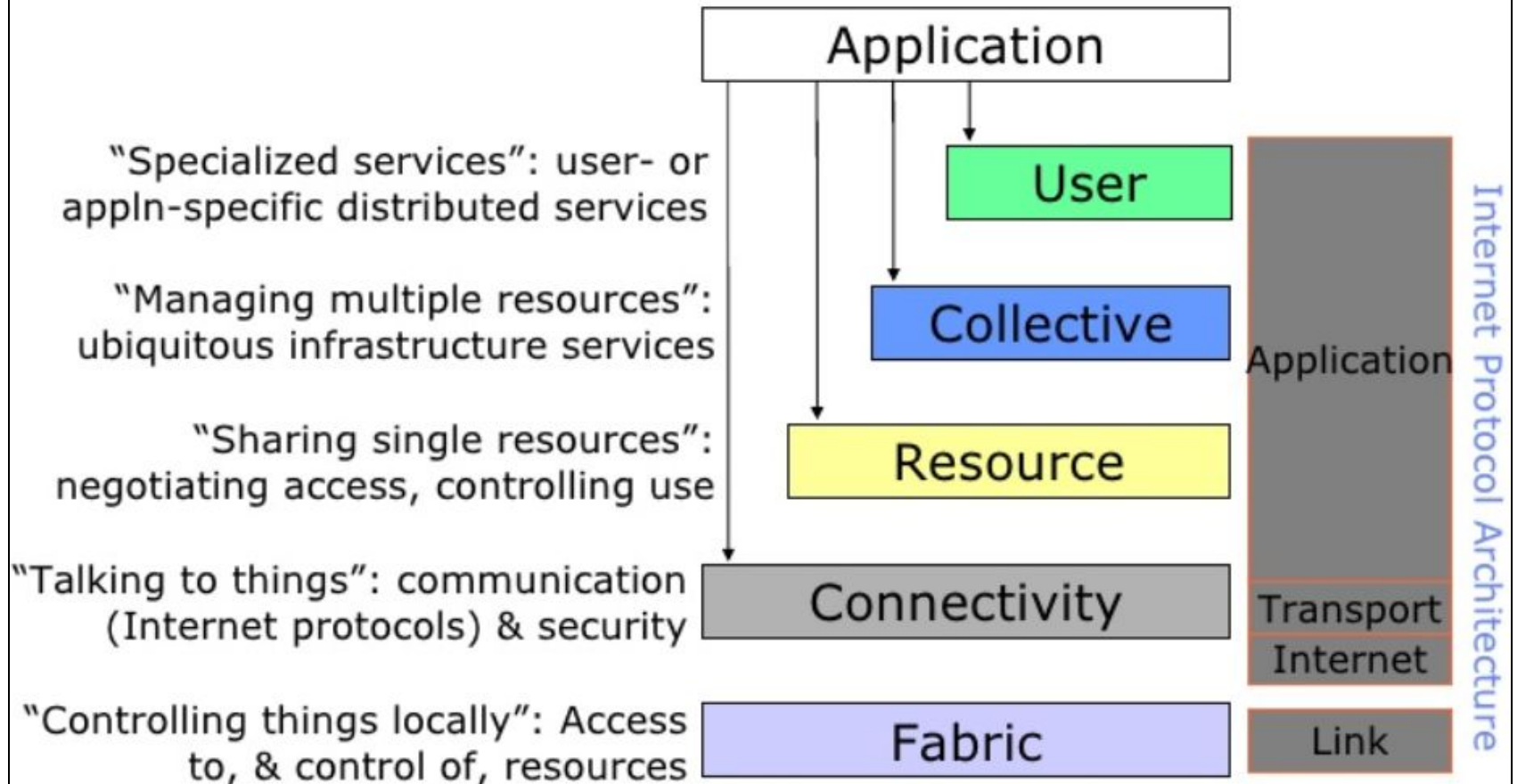# Grid Computing

(v)Application Layer:

This layer offers **communication interfaces** to the users and the administrators for **interacting with the Grid.**
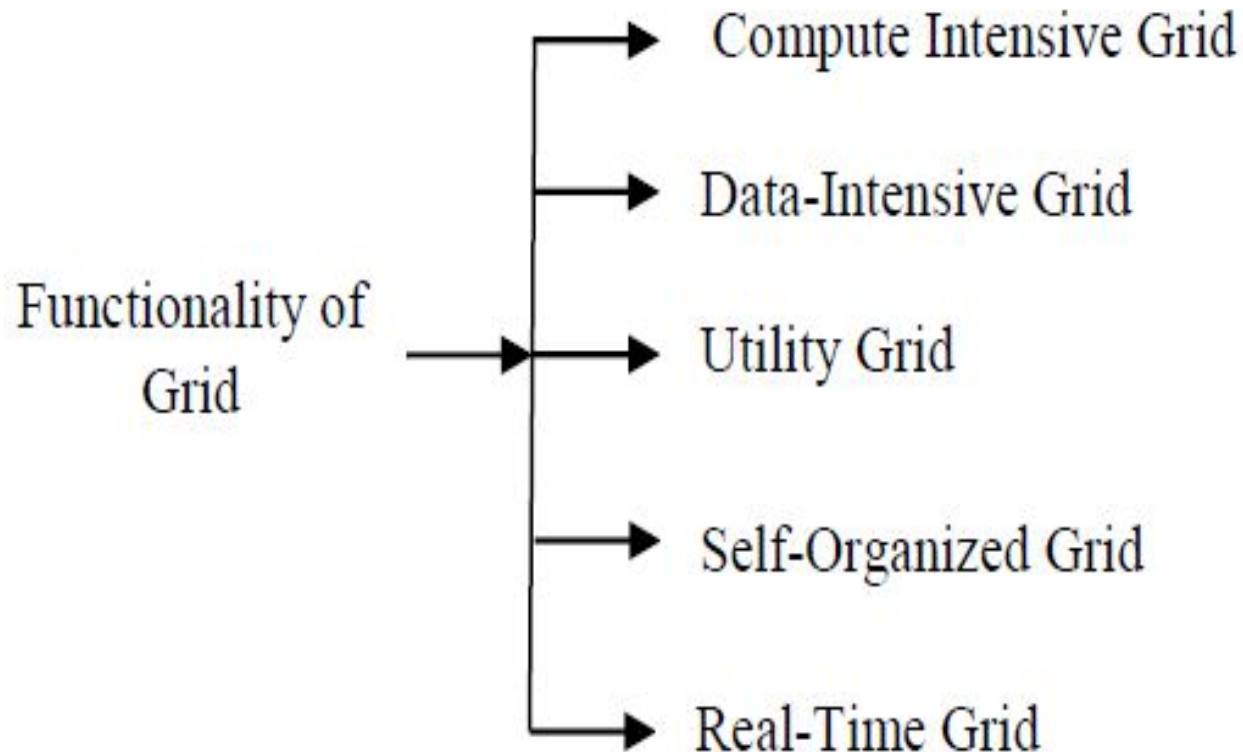
Working of layers

- **Fabric.** The lowest layer job is used to make a common interface on all possible kinds of resources available. Access by higher layers is granted via standardized processes.
- **Resource and connectivity protocols:** The connectivity layer defines the basic communication- and authentication protocols which are needed by the grid. While the communication protocols allow the exchange of files between different resources connected by the first layer, the authentication protocols allow to communicate confidentially and to ensure the identity of the two partners.
- **Collective services:** The purpose of this layer is the coordination of multiple resources. Access to these resources doesn't happen directly but merely via the underlying protocols and interfaces.
- **User applications:** To this layer belong all those applications which are operating in the environment of a virtual organization. Jobs of the lower layers get called by applications and can use resources transparently.

Layered grid architecture

Application

"Specialized services": user- or appln-specific distributed services

User

"Managing multiple resources": ubiquitous infrastructure services

Collective

Application

"Sharing single resources": negotiating access, controlling use

Resource

"Talking to things": communication (Internet protocols) & security

Connectivity

Transport

Internet

"Controlling things locally": Access to, & control of, resources

Fabric

Link

Internet Protocol Architecture

Types of Grid Systems

.0. Taxonomy of Grid Systems based on the type of Functionality

Types of Grid Systems

- **Compute Intensive:**
  i.   Major functionality of this Grid is to **offer powerful CPU resources**, whereas other characteristics are trivial.
  ii. It intends to **provide aggregated CPU cycles** of several resources for **fulfilling computational need of a particular application.**

- **Data-Intensive Grid:**
  It offers **large storage space** for **data discovery, data management and data processing.**

- **Utility Grid:**
  i.   It **pools dynamically available resources** to **match the requirements of the applications**.
  ii. This enables the Utility Grid to provide services that **cannot be provided by any single machine.**

Cont …

•Self-Organized Grid:

i.   Grid management is extremely challenging due to its **complex dynamic environment.**

ii.   However, Self-Organized Grid includes **some intelligence embedded into its infrastructure** to **automate its control, organization, and monitoring procedure**.

iii.   It can **heal itself** in **changing and uncertain environments for maximizing reliability and resource utilization.**

•Real Time Grid:

i.   **Traditional Grid** could **not support certain requirements** of many real time applications.

ii.   **Technological advancement** with increasing number of real time applications initiated the _emergence infrastructure_ that can deal with the rich variety of real time _interactions_ between the users and the applications through virtual workspace.

iii.   Therefore, Grids that support Real time applications such as **disaster management, e-healthcare, e-learning** etc. can be termed as Real-Time Grid.

Examples

- **a) Health**
- World Community Grid
- World Community Grid's mission is to create the largest public computing grid benefiting humanity, which is funded and operated by IBM. Using the idle time of computers around the world, World Community Grid's research projects have analyzed aspects of the human genome, HIV, dengue, muscular dystrophy, and cancer.

- **B) Engineering Design**
- Analysis of real-time data to find a particular pattern.
- • Experiment modeling to create new designs.
- • Verifying existing models for accuracy using simulation activities.

- **C) Grid Computing in Government**
- C-DAC launched the Indian National Grid Computing Initiative - GARUDA, with a vision to facilitate solutions to scientific, engineering and socio-economic development. GARUDA is a collaboration of scientific and technological researchers on a nation wide grid comprising of computational nodes, mass storage and scientific instruments. The aim is to provide the technological advances required to enable data and compute intensive science and engineering applications for the 21st Century.

- **D) Education, Science and beyond**
- SETI (Search for Extraterrestrial Intelligence) @Home project
- SETI@home is a scientific experiment that uses Internet-connected computers in the Search for Extraterrestrial Intelligence (SETI). You can participate by running a free program that downloads and analyzes radio telescope data.
- Education - e-learning

- **E) Grid Computing in the Movie Industry**
- Many films can't be made without the use of grid computing not only because of special effects but also because of the fact that grid computing enables faster production of a film.
- Grid computing is an enhancement tool for the studios.

Mobile Computing

- Mobile Computing is a technology that allows transmission of data, voice and video via a computer or any other wireless enabled device without having to be connected to a fixed physical link. The main concept involves –

- Mobile communication
- Mobile hardware
- Mobile software

Mobile Computing

- The mobile communication in this case, refers to the infrastructure put in place to ensure that seamless and reliable communication goes on. These would include devices such as protocols, services, bandwidth, and portals necessary to facilitate and support the stated services. The data format is also defined at this stage.

Mobile Computing

Mobile Hardware
Mobile hardware includes mobile devices or device components that receive or access the service of mobility. They would range from portable laptops, smartphones, tablet Pc's, Personal Digital Assistants.

Mobile software
Mobile software is the actual program that runs on the mobile hardware. It deals with the characteristics and requirements of mobile applications. This is the engine of the mobile device. In other terms, it is the operating system of the appliance. It's the essential component that operates the mobile device.

What is Cloud Computing?

- **Cloud Computing** is a general term used to describe a new class of network based computing that takes place over the Internet,

  ☐ basically a step on from Utility Computing

  ☐ a collection/group of integrated and networked hardware, software and Internet infrastructure (called a platform).

  ☐ Using the Internet for communication and transport provides hardware, software and networking services to clients

Contd..

- These platforms hide the complexity and details of the underlying infrastructure from users and applications by providing very simple graphical interface or API

- The platform provides on demand services, that are always on, anywhere, anytime and any place.

- Pay for use

- Scalable as needed

    -scale up and down in capacity and functionalities

- The hardware and software services are available to general public, enterprises, corporations and businesses markets

Difference between cluster and cloud computing (SCHOL)

- **Spanning area** -Cluster is a group of computers connected by a **local area network (LAN),** whereas cloud and grid are more **wide scale and can be geographically distributed.**

- **Coupling** - cluster is **tightly** coupled, whereas a grid or a cloud is **loosely** coupled.

- **Homogeneity** -clusters are made up of **machines with similar hardware**, whereas clouds and grids are made up of machines with possibly very **different hardware configurations.**

- **Ownership**: A grid is a collection of computers which is owned by **multiple parties in multiple locations.** Whereas a cloud is a collection of computers usually owned by a **single party.**

- **Load distribution**: In cluster the load is distributed among **all nodes** but in case of Cloud, the **resources are allocated as per the request**. The whole data center is not occupied by one job.

What is Cloud Computing?

Cloud Computing Architecture

Cloud Computing - Overview

Cloud Computing provides us a means by which we can access the applications as utilities, over the Internet. It allows us to create, configure, and customize applications online.

What is Cloud?

The term **Cloud** refers to a **Network** or **Internet**. In other words, we can say that Cloud is something, which is present at remote location. Cloud can provide services over network, i.e., on public networks or on private networks, i.e., WAN, LAN or VPN.

Applications such as **e-mail, web conferencing, customer relationship management (CRM)**, all run in cloud.

What is Cloud Computing?

Cloud Computing refers to **manipulating, configuring,** and **accessing** the applications online. It offers online data storage, infrastructure and application.

- We need not to install a piece of software on our local PC and this is how the cloud computing overcomes platform dependency issues. Hence, the Cloud Computing is making our business application mobile and collaborative.
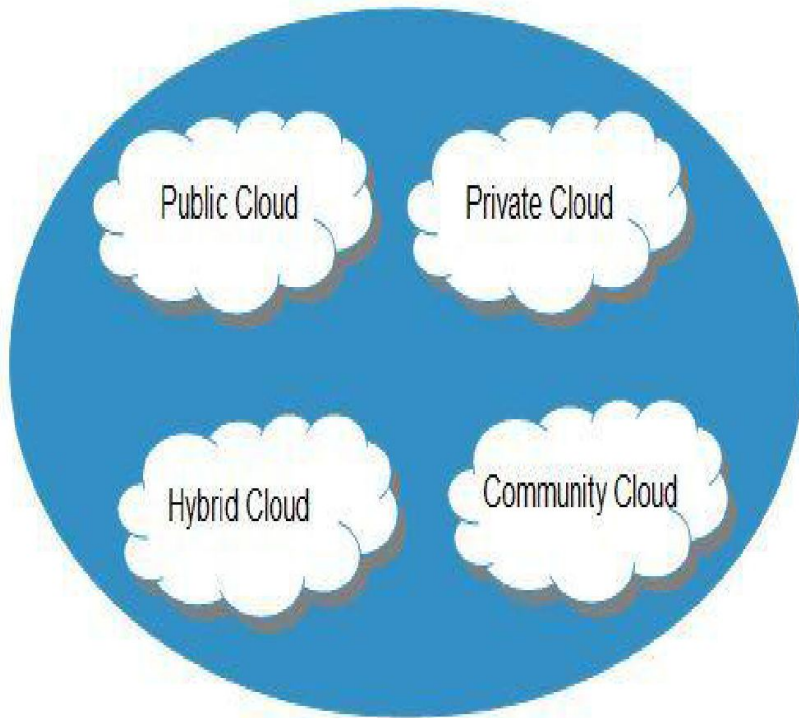
Basic Concepts

There are certain services and models working behind the scene making the cloud computing feasible and accessible to end users. Following are the working models for cloud computing:

- # Deployment Models

- # Service Models

- DEPLOYMENT MODELS
- 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, Hybrid and Community

Deployment Models

- The **Public Cloud** allows systems and services to be easily accessible to the general public. Public cloud may be less secure because of its openness, e.g., e-mail.
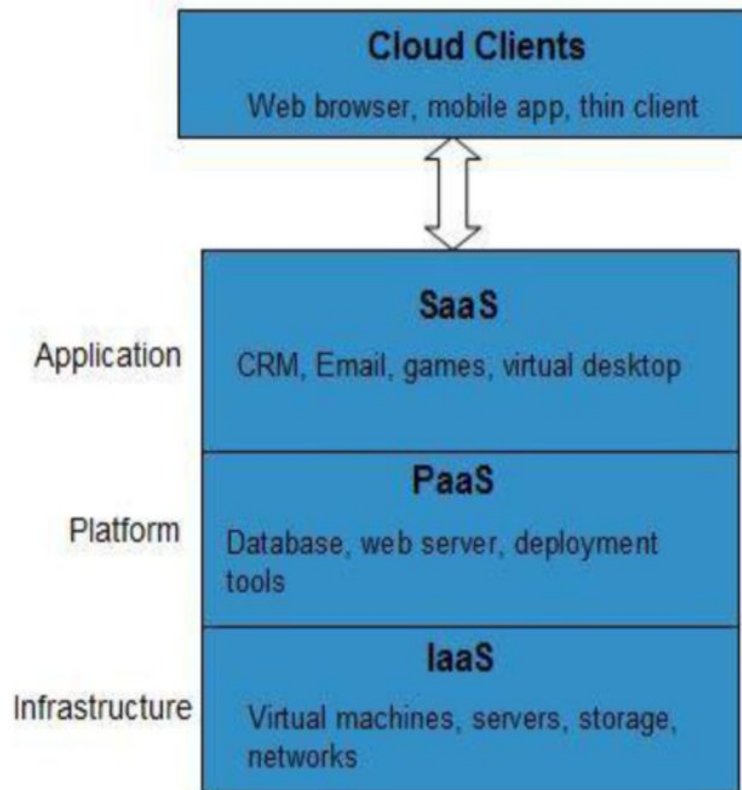
The **Private Cloud** allows systems and services to be accessible within an organization. It offers increased security because of its private nature.

Cummunity Cloud -The Community Cloud allows systems and services to be accessible by group of organizations.

The **Hybrid Cloud** -is mixture of public and private cloud. However, the critical activities are performed using private cloud while the non-critical activities are performed using public cloud.

Service Models

INFRASTRUCTURE AS A SERVICE (IAAS)

- IaaS provides access to fundamental resources such as physical machines, virtual machines, virtual storage, etc.
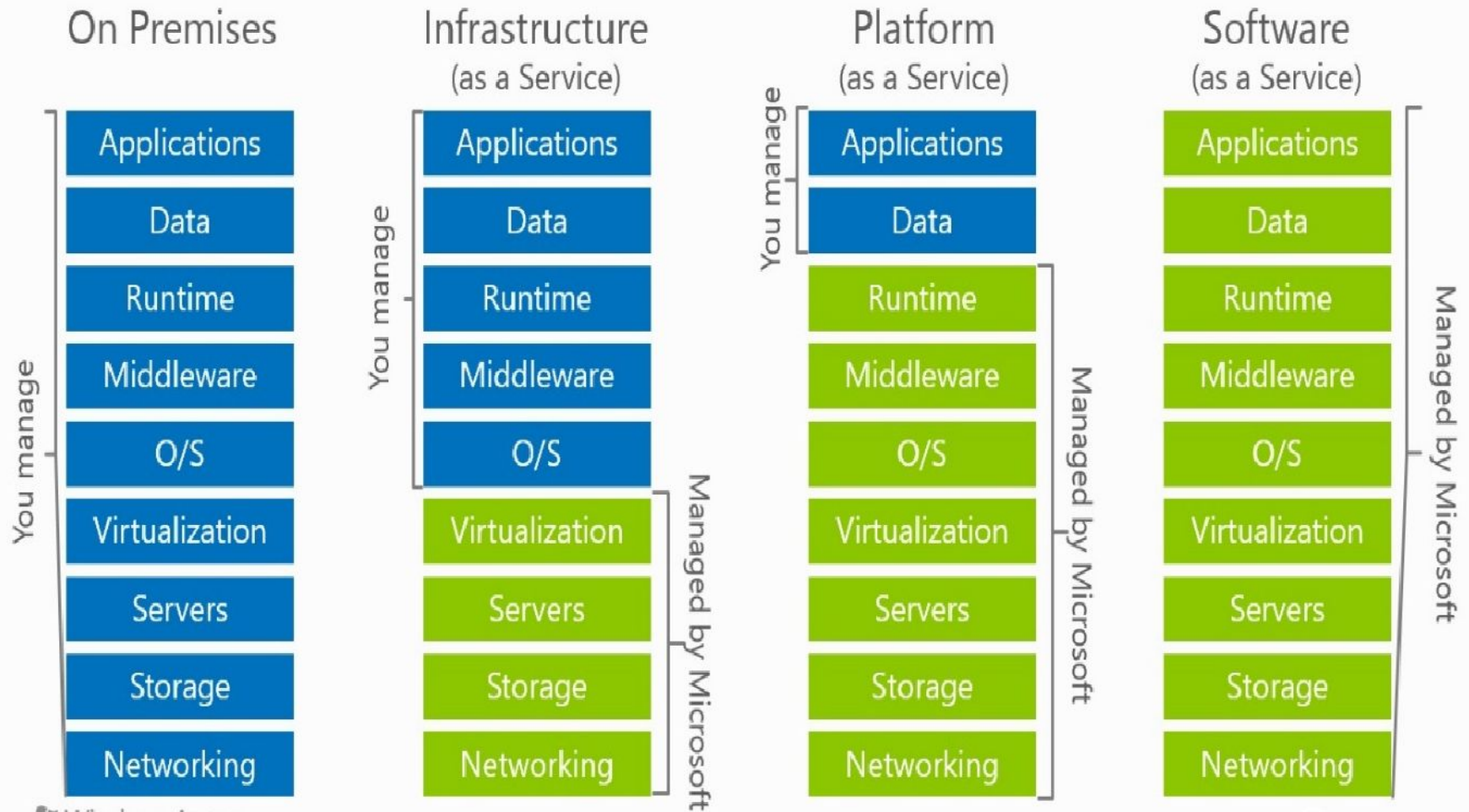
PLATFORM AS A SERVICE (PAAS)

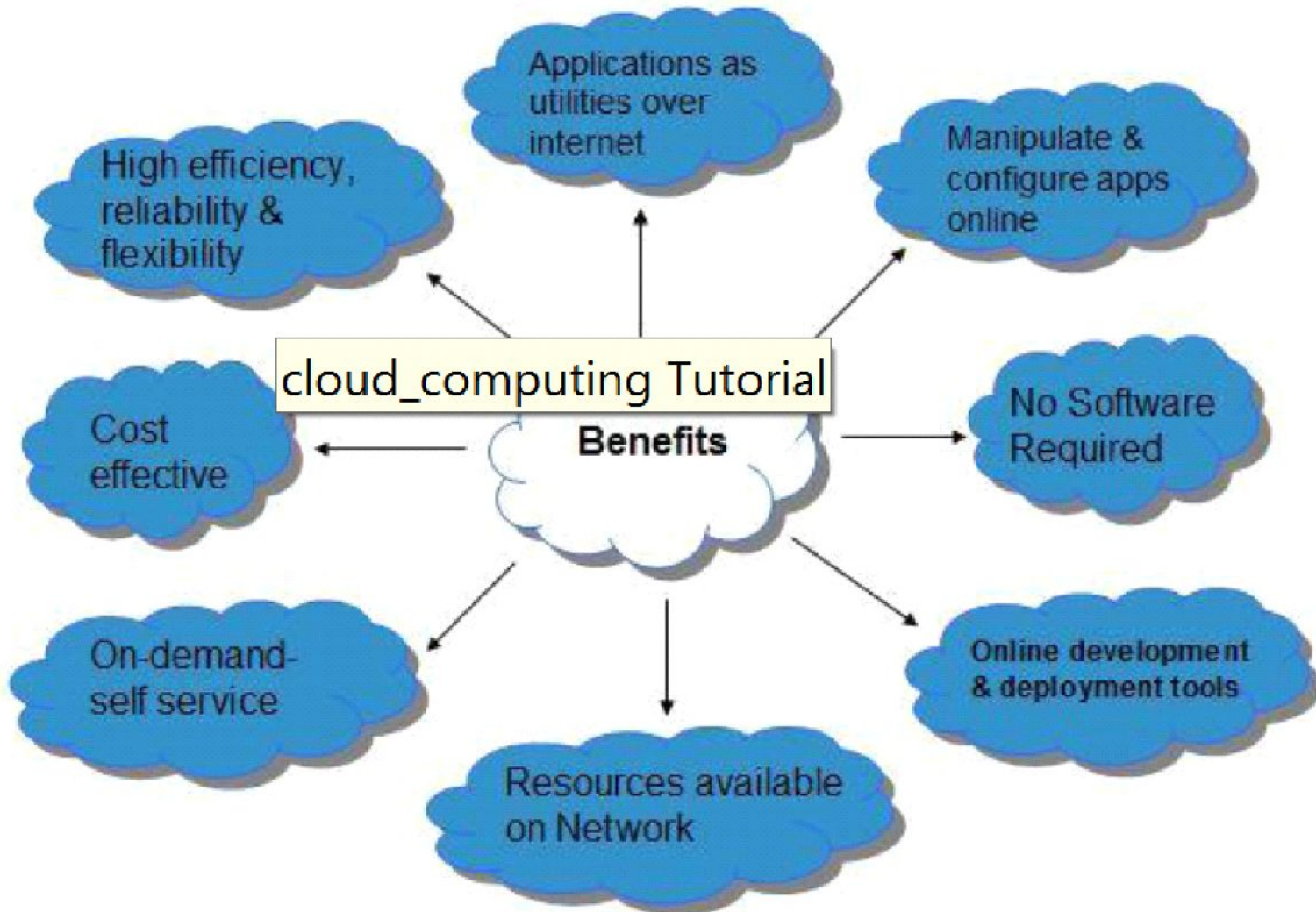- PaaS provides the runtime environment for applications, development & deployment tools, etc.

SOFTWARE AS A SERVICE (SAAS)

- SaaS model allows to use software applications as a service to end users

Cloud Models

On Premises	Infrastructure (as a Service)	Platform (as a Service)	Software (as a Service)
Applications	Applications	Applications	Applications
Data	Data	Data	Data
Runtime	Runtime	Runtime	Runtime
Middleware	Middleware	Middleware	Middleware
O/S	O/S	O/S	O/S
Virtualization	Virtualization	Virtualization	Virtualization
Servers	Servers	Servers	Servers
Storage	Storage	Storage	Storage
Networking	Networking	Networking	Networking

You manage (On Premises)

You manage (Infrastructure) — Managed by Microsoft

You manage (Platform) — Managed by Microsoft

Managed by Microsoft (Software)

Windows Azure

Benefits

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 Networks Access | Rapid Elasticity |

| Resource Pooling | Measured Service |

Common Characteristics

| Massive Scale | Resilient Computing |

| Homogeneity | Geographic Distribution |

| Virtualization | Service Orientation |

| Low cost software | Advanced Security |