Distributed Computing

Computer Engg

Semester – 6th

Unit 1

Cloud Computing

Topics -

- Overview of Cloud Computing
- Characteristics of Cloud Computing
- Advantages of Cloud Computing
- Challenges of Cloud Computing
- Applications of Cloud Computing

Distributed Computing

- Field of **computer** science that studies **distributed systems**
- A **distributed system** is a model in which **components** located on networked computers communicate and coordinate their actions by passing messages.
- The **components** interact with each other in order to achieve a common goal.


Cloud
Computing

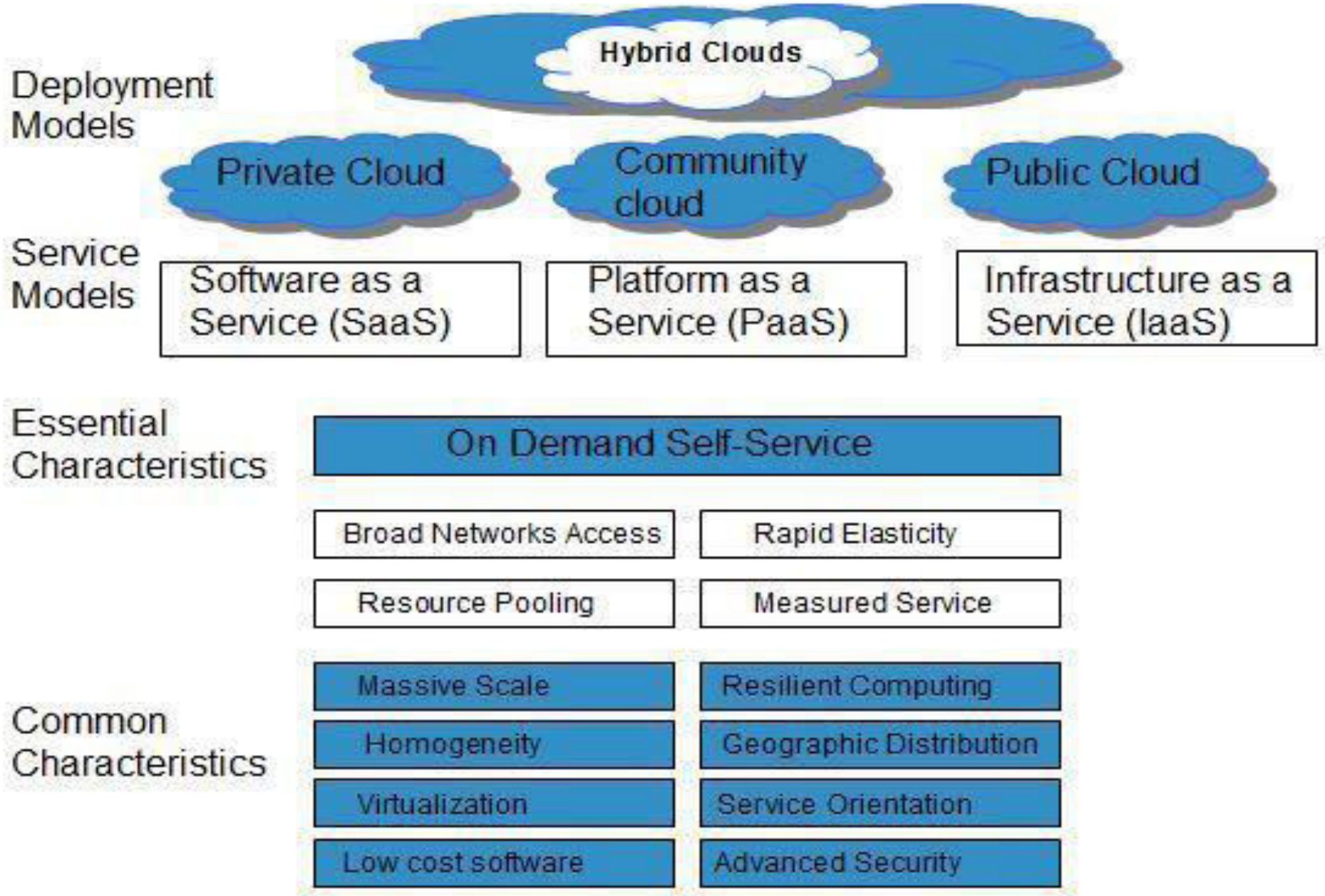
CLOUD COMPUTING

- Cloud Computing can be defined as delivering computing power (CPU, RAM, Network Speeds , Storage OS software) as a service over a network (usually on the internet) rather than physically having the computing resources at the customer location.
- Cloud computing means ***on demand delivery of IT resources via*** **the internet with pay-as-you-go pricing.**
- It provides a solution of IT infrastructure in low cost.

OVERVIEW OF CLOUD COMPUTING

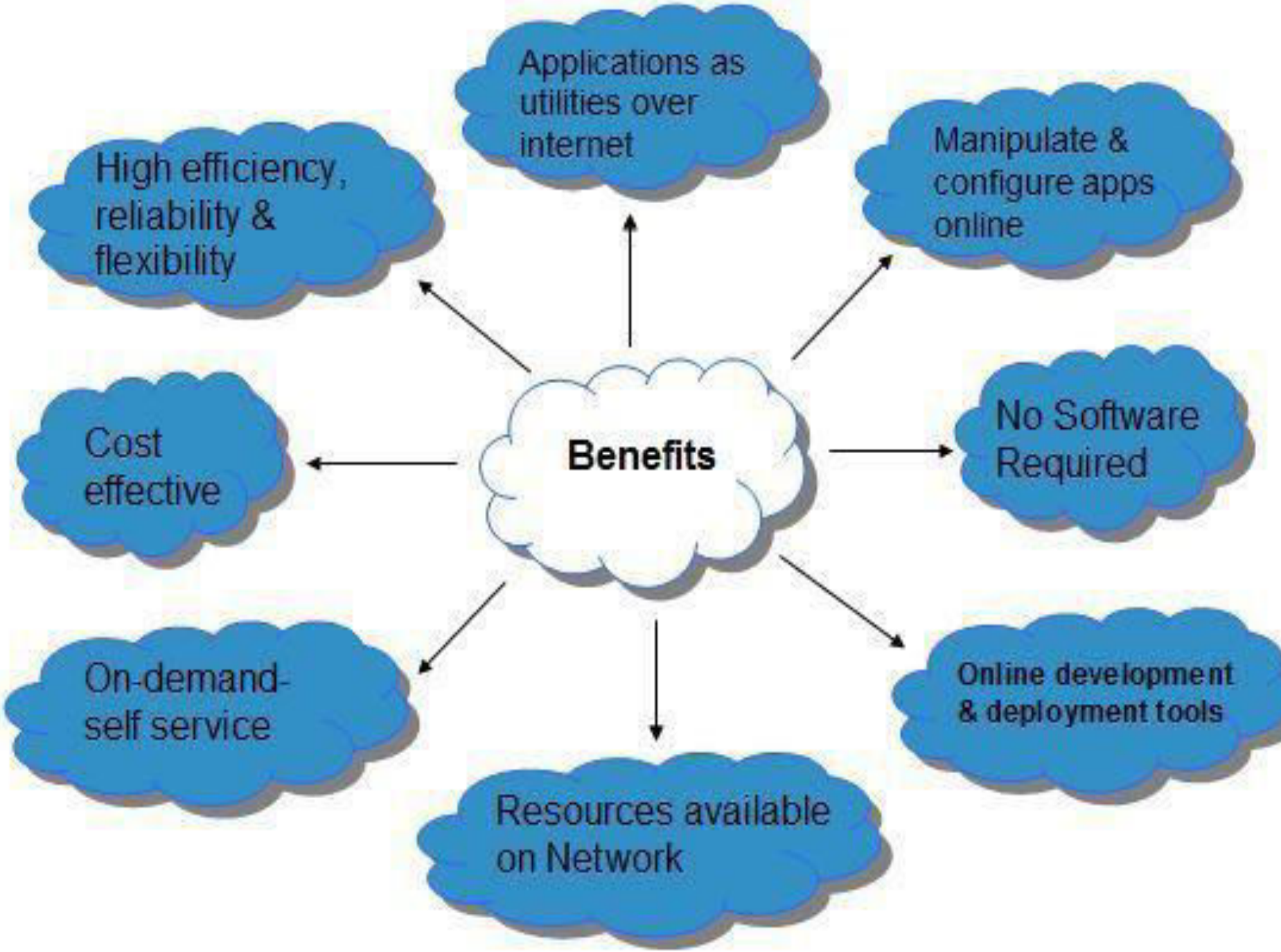
- **Cloud computing** refers to applications and services that run on a distributed network using virtualized resources
- Can be accessed by common Internet protocols and networking standards.
- It brings the user access to data, applications and storage that are not stored on their **computer**.
- It can be understood as a delivery system that delivers computing the same way a power grid delivers electricity.

Charcteristics of Cloud Ccomputing



CHARACTERISTICS

- On-demand Self Services
- Resource Pooling
- Rapid Elasticity
- Measured Service(pay-per-use mode)
- Multi-tenacity
- High Scalability
- High Availability and Reliability
- Low cost
- Location Independence


Applications as utilities over internet
High efficiency, reliability & flexibility
Manipulate & configure apps online
Cost effective
Benefits
No Software Required
On-demand-self service
Online development & deployment tools
Resources available on Network

ADVANTAGES

- Cost Efficient
- Increased Storage
- Flexibility
- Backup and Recovery
- Automatic Software Integration
- Easy access to Information
- Quick Deployment
- Increased Data Safety
- Improved Performance
- Scalability

CHALLENGES

- Data Protection(Security)
- Data Recovery and Availability
- Management Capabilities
- Regulatory and Compliance Restrictions
- Technical Issues
- Security Issues
- Prone to Attack
- High speed Internet Required
- Requires Constant Internet Connection

APPLICATIONS

- Infrastructure as a Service
- Platform as a Service
- Private cloud
- Hybrid cloud
- Test and Development
- Bid Data Analytics
- File Storage
- Disaster Recovery
- Backup

Unit 2

Cloud Computing
Service Models
and
Deployment Models

Topics -

- Service Model- Saas, Paas, Iaas
- Deployment Models : Private Cloud, Public Cloud, Hybrid Cloud, Community Cloud

CLOUD COMPUTING MODELS

- Cloud computing has two types of models :-
  - **Service Models**
  - **Deployment Models**

Major Types of cloud service models

SaaS : Software as a Service

PaaS: Platform as a Service

IaaS: Infrastructure as a Service

Software as a Service

- SaaS is a software distribution model in which applications are hosted by a cloud service provider and made available to customers over internet.
- SaaS is also known as "On-Demand Software".
- In SaaS, software and associated data are centrally hosted on the cloud server.
- SaaS is accessed by users using a thin client via a web browser.

Infrastructure as a Service

- IaaS is one of the layers of cloud computing platform wherein the customer organization outsources its IT infrastructure such as servers, networking, processing, storage, virtual machines and other resources.
- Customers access these resources over internet i.e. cloud computing platform, on a pay-per-use model.
- IaaS cloud computing platform layer eliminates the need for every organization to maintain the IT infrastructure.
- IaaS is offered in three models: public, private, and hybrid cloud.

Platform as a Service

- PaaS cloud computing platform is a developer programming platform which is created for the programmer to develop, test, run and manage the applications.
- A developer is able to write the application as well as deploy it directly into this layer easily.
- PaaS extend and abstract the IaaS layer by removing the hassle of managing the individual virtual machine.
- In PaaS cloud computing platform, back end scalability is handled by the cloud service provider and the end user does not have to worry about to manage the infrastructure.

Cloud Deployment Models

- Private cloud

  -Enterprise owned or leased

- Public cloud

  -Sold to the public, mega-scale infrastructure

- Community cloud

  -Shared infrastructure for specific community

- Hybrid cloud

  -composition of two or more clouds

PUBLIC CLOUD

- Public cloud allows the accessibility of systems and services easily to general public.
- It is constructed with a view to offer unlimited storage space and increased bandwidth via Internet to all Businesses.
- They are owned, hosted and operated by third-party service providers.
- It is based on “pay-as-you-go” model.
- Eg: Amazon, IBM, Microsoft, Google, Rackspace etc.

PRIVATE CLOUD

- The Private cloud allows the accessibility of systems and services within the organization.
- Private cloud is operated only within a particular organization.
- It can be managed and hosted internally or by third party.
- Users are charged on basis of per Gigabytes usage along with bandwidth transfer fees.
- E.g. Amazon VPC, IBM SmartCloud Foundation and Microsoft Private Cloud.

HYBRID CLOUD

- The Hybrid cloud is the mixture of public and private cloud.
- Non-critical activities are performed by public cloud while critical activities are performed by private cloud.
- It has advantages of both private and public cloud, offer flexibility, control and security of multiple deployment models.
- IT organizations use hybrid clouds to employ cloud bursting for scaling across clouds.

COMMUNITY CLOUD

- The cloud infrastructure is shared between the organizations having similar interests and requirements.
- It can be managed and hosted internally or by a third party.
- This model is built when the organizations are ready to share the benefits of cloud service.

Cloud Efficiencies and Improvements

- Cost efficiencies
- Time efficiencies
- Power efficiencies
- Improved process control
- Improved security
- Unlimited capacity

Unit 3

Grid Computing

Topics –

Overview

Advantages

Virtual Organizations

Applications

Introduction

- The term Grid comes from an analogy to the Electric Grid.
  - Pervasive access to power.
  - Similarly, Grid will provide pervasive, consistent, and inexpensive access to advanced computational resources.
- Grid computing is all about achieving greater performance and throughput by pooling resources on a local, national, or international level.

GRID Computing

- Grids are about large-scale resource sharing.
  - Spanning administrative boundaries.
    - Central processors, storage, network bandwidth, databases, applications, sensors and so on
- Problem solving in dynamic, multi-institutional environment.
- Organizing geographically distributed computing resources
  - So that they can be flexibly and dynamically allocated and accessed

- Providing such capabilities, where Sharing is highly controlled, clear definitions of exactly what is shared, who is allowed to share, and the conditions under which sharing occurs.

Elements of Grid Computing

- Resource sharing
  - Computers, data, storage, sensors, networks, …
  - Sharing always conditional: issues of trust, policy, negotiation, payment, …
- Coordinated problem solving
  - Beyond client-server: distributed data analysis, computation, collaboration, …
- Dynamic, multi-institutional *virtual organizations*
  - Community overlays on classic org structures
  - Large or small, static or dynamic

Virtual Organizations

- A set of individuals and/or institutions defined by a set of sharing rules
- The sharing is highly controlled, with resource providers and consumers defining clearly and carefully just what is shared

An example: the set of application service providers, storage service providers, cycle providers and consultants engaged by a car manufacturer to plan for a new factory

Another example: industrial consortium building a new aircraft

More Formal Definition of Grids

- A grid is a system that:
  - Coordinates resource sharing in a de-centralized manner (i.e., different VOs).
  - Uses standard, open, general purpose protocols and interfaces.
  - Delivers non-trivial qualities of service.
    - Guaranteed bandwidth for application.
    - Guaranteed CPU cycles.
    - Guaranteed latency.

Computational Grid Applications

- Biomedical research
- Industrial research
- Engineering research
- Studies in Physics and Chemistry

Applications

- National Virtual Observatory
  - Astronomical surveys produce terabytes of data.
  - Data sets will cover sky in different wave bands (x-rays, optical, infrared, radio).
  - Challenge is to make this accessible to general research community.
    - Heterogeneous data producers and consumers.
  - Resources in this Grid are data sets rather than compute engines.

High-Energy Physics

- Large-scale collaborations for CERN's Large Hadron Collider.
- Involves 4000 physicists, 150 institutions, in more than 30 countries.
- Data sets now at petabyte level. Predicted to generate data at the exabyte level in this decade.
- Challenges:
  - Providing rapid access to subsets of data.
  - Secure access to distributed computing and data handling resources.

Open Grid Services Architecture

- Developed by the Global Grid Forum to define a common, standard, and open architectures for Grid-based applications.
  - Provides a standard approach to all services on the Grid.
    - VO Management Service.
    - Resource discovery and management service:
    - Job management service.
    - Security services.
    - Data management services.
- Built on top of and extends the Web Services architecture, protocols, and interfaces.

A stateless Web Service invocation

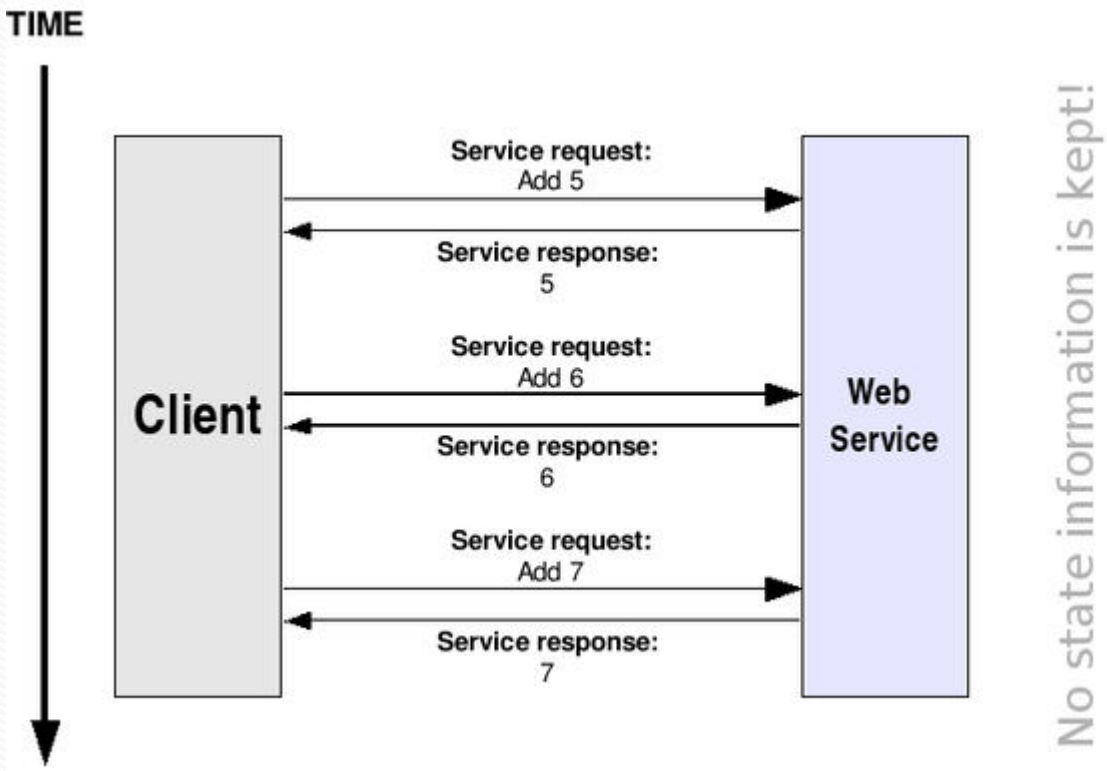
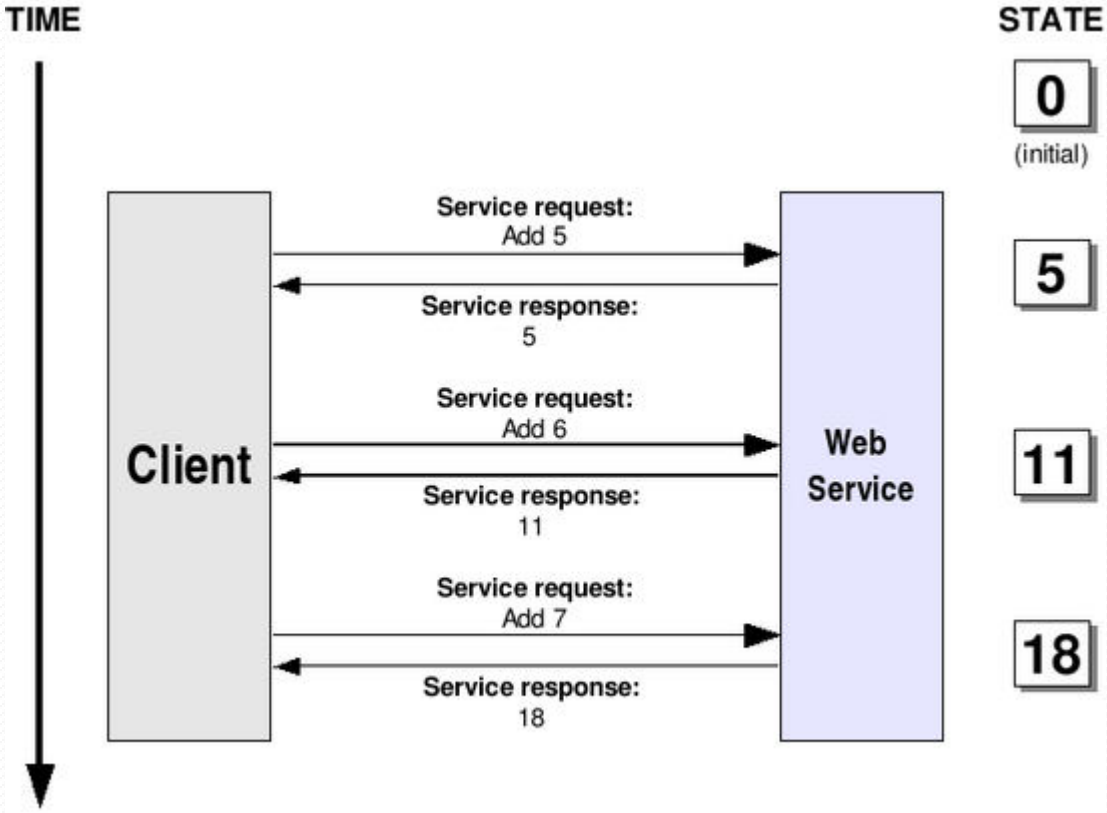


Figure 1.11. A stateful Web Service invocation



Standards Bodies

The primary standards-setting body is[1]:

- Global Grid Forum (GGF)
  - Started in 1998
  - More than 40 organizations involved and growing …

Others:

- W3C consortium (Worlds Wide Web Consortium)
  - Working on standardization of web-related technologies such as XML
  - See http://www.w3.org
- OASIS (Organization for the Advancement of Structured Information Standards)
- IETF, DMTF

Globus Project

- Open source software toolkit developed for grid computing.
- Roots in I-way experiment.
- Work started in 1996.
- Four versions developed to present time.
- Reference implementations of grid computing standards.
- Defacto standard for grid computing.

Globus Version 4

- A “toolkit” of services and packages for creating the basic grid computing infrastructure
- Higher level tools added to this infrastructure
- Version 4 is web-services based
- Some non-web services code exists from earlier versions (legacy) or where not appropriate (for efficiency, etc.).

Unit 4

Other Technologies

Topics -

Cluster Computing

Peer to Peer Networks

Utility Computing

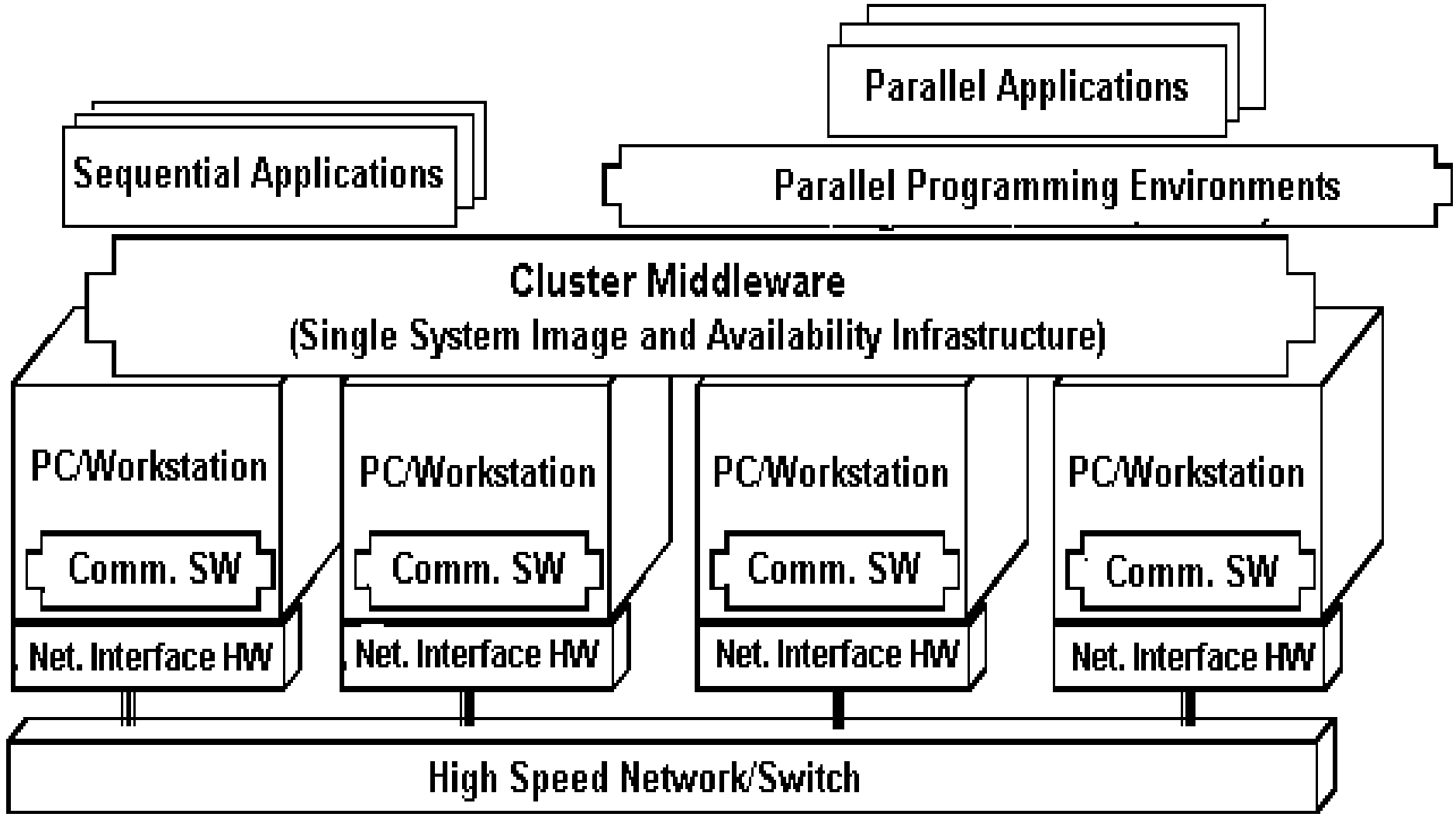
Ubiquitous Computing

Comparison of Grid, Cluster and Cloud Computing

Cluster Computing

- Clustering gained momentum when following technologies converged:
  - Very HP Microprocessors
    - workstation performance = yesterday supercomputers
  - High speed communication
    - Comm. between cluster nodes >= between processors in an SMP.
  - Standard tools for parallel/ distributed computing & their growing popularity.

Cluster Computer Architecture



Cluster Components Applications

- Sequential
- Parallel / Distributed (Cluster-aware app.)
  - Grand Challenging applications
    - Weather Forecasting
    - Quantum Chemistry
    - Molecular Biology Modeling
    - Engineering Analysis (CAD/CAM)
  - Web servers, data-mining

Key Benefits of Clustering

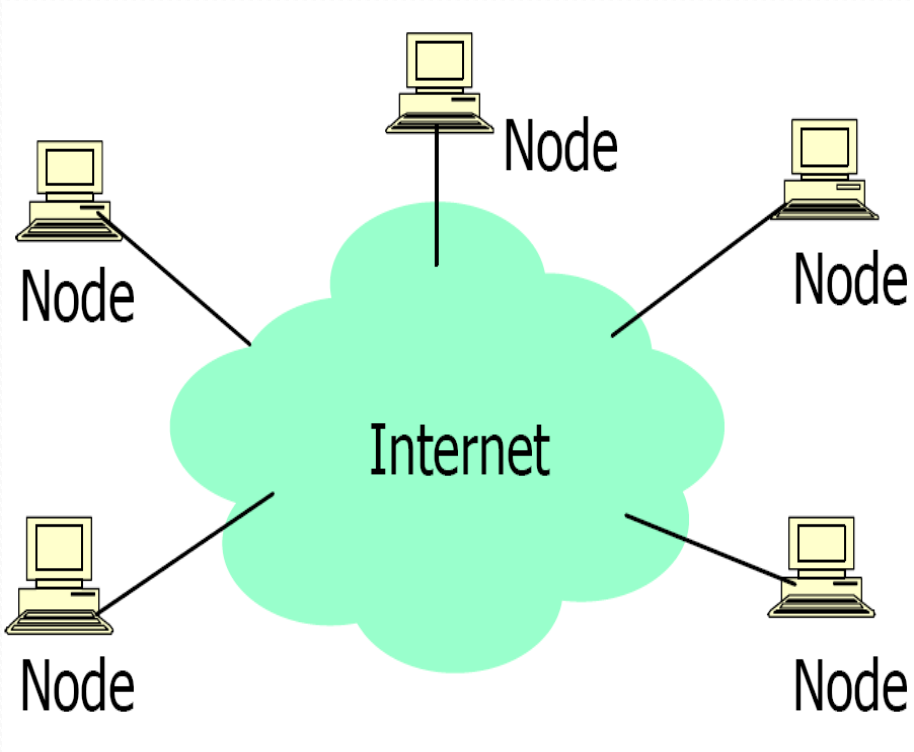
- System availability (HA). offer inherent high system availability due to the redundancy of hardware, operating systems, and applications.
- Hardware Fault Tolerance. redundancy for most system components (eg. disk-RAID), including both hardware and software.
- OS and application reliability. run multiple copies of the OS and applications, and through this redundancy
- Scalability. adding servers to the cluster or by adding more clusters to the network as the need arises or CPU to SMP.
- High Performance. (running cluster enabled programs)

Peer to Peer Networks

- P2P computing is the sharing of computer resources and services by direct exchange between systems.
- These resources and services include the exchange of information, processing cycles, cache storage, and disk storage for files.
- P2P computing takes advantage of existing computing power, computer storage and networking connectivity, allowing users to leverage their collective power to the "benefit" of all.

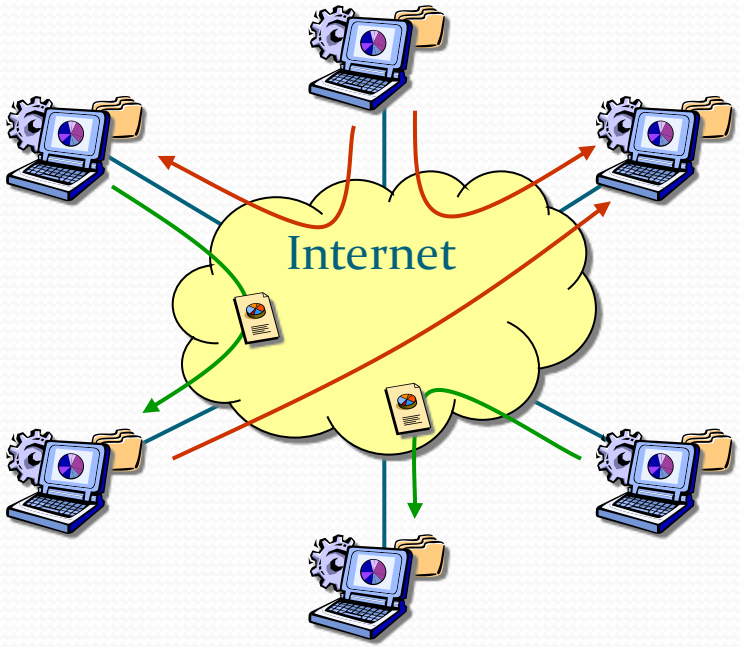
P2P Architecture

- All nodes are both clients and servers
  - Provide and consume data
  - Any node can initiate a connection
- No centralized data source
  - "The ultimate form of democracy on the Internet"
  - "The ultimate threat to copy-right protection on the Internet



What is P2P?

- A distributed system architecture
  - No centralized control
  - Typically many nodes, but unreliable and heterogeneous
  - Nodes are symmetric in function
  - Take advantage of distributed, shared resources (bandwidth, CPU, storage) on peer-nodes
  - Fault-tolerant, self-organizing
  - Operate in dynamic environment, frequent join and leave is the norm



P2P Network Characteristics

- Clients are also servers and routers
  - Nodes contribute content, storage, memory, CPU
- Nodes are autonomous (no administrative
- authority)
- Network is dynamic: nodes enter and leave the network "frequently"
- Nodes collaborate directly with each other (not through well-known servers)
- Nodes have widely varying capabilities

P2P vs. Client/Server

- Pure P2P:
  - No central server
  - For certain requests any peer can function as a client, as a router, or as a server
  - The information is not located in a central location but is distributed among all peers
  - A peer may need to communicate with multiple peers to locate a piece of information

Both demand and capacity of the network increases as more peers are added.

P2P Benefits

- Efficient use of resources
  - Unused bandwidth, storage, processing power at the edge of the network
- Scalability
  - Consumers of resources also donate resources
  - Aggregate resources grow naturally with utilization
- Reliability
  - Replicas
  - Geographic distribution
  - No single point of failure
- Ease of administration
  - Nodes self organize
  - No need to deploy servers to satisfy demand (c.f. scalability)
  - Built-in fault tolerance, replication, and load balancing

Utility Computing

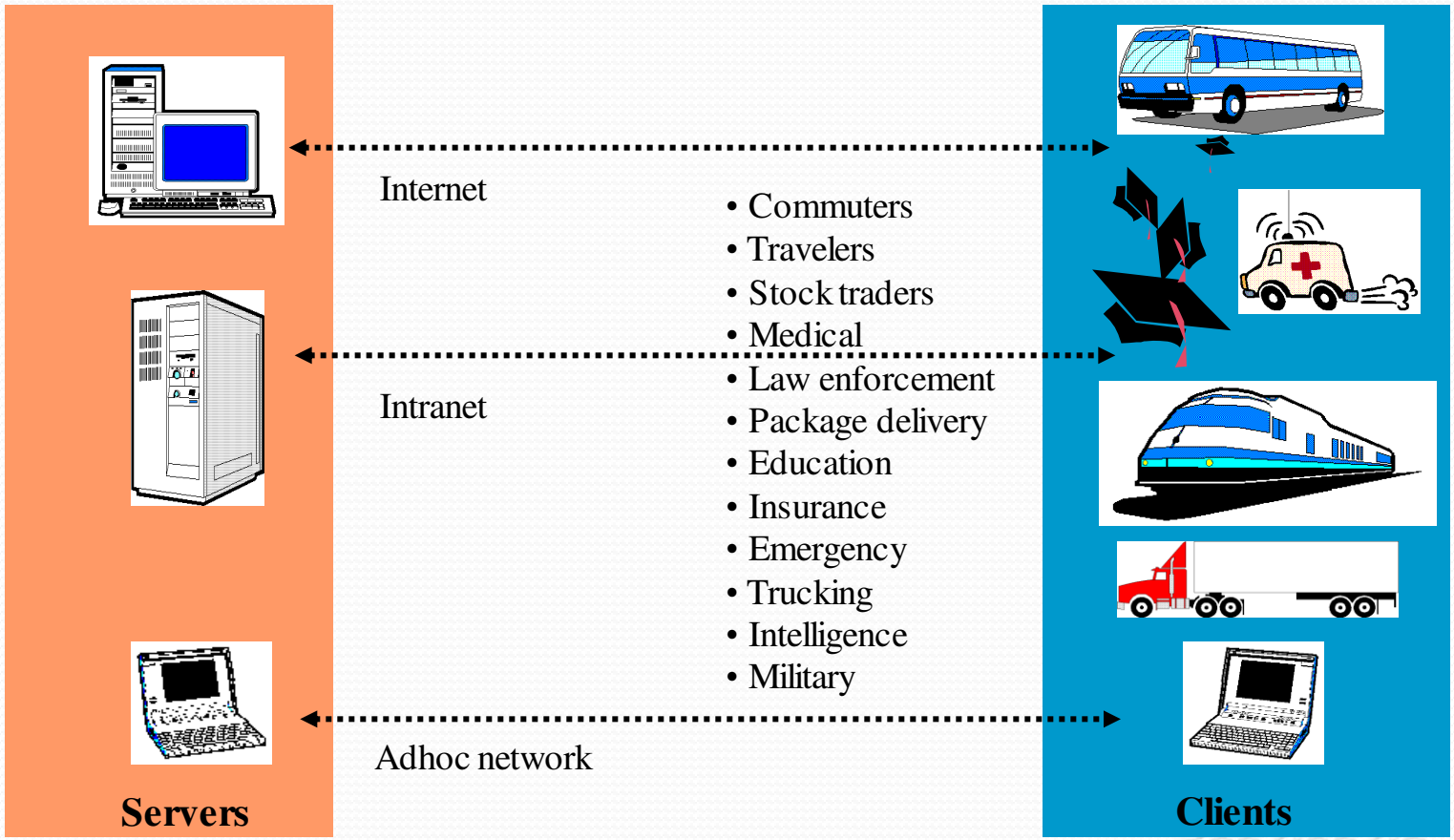
Packaging of computing resources such as:

- computation and storage
- as a metered service
- similar to a traditional public utility

Ubiquitous Computing

- integrates computation into the environment, rather than having computers which are distinct objects.
- Another term for ubiquitous computing is pervasive computing.
- Promoters of this idea hope that embedding computation into the environment would enable people to move around and interact with computers more naturally than they currently do.

Beneficiaries of Ubiquitous Computing



Cluster vs. Grid vs. Cloud

Cluster Computing	Grid Computing	Cloud Computing
Characteristics 1:Tightly coupled systems 2: Single system image 3: Centralized Job management & scheduling system	Characteristics 1: Loosely coupled (Decentralization) 2 : Diversity and Dynamism 3: Distributed Job Management & scheduling	Characteristic 1: Dynamic computing infrastructure 2: IT service-centric approach 3: Self-service based usage model 4: Minimally or self-managed platform 5: Consumption-based billing
A bunch of similar computers are hooked up locally to operate as a single computer	The computers do not have to be in the same physical location and can be operated independently	The computers need not to be in the same physical location.

Cluster vs Grid vs Cloud

Cluster Computing	Grid Computing	Cloud Computing
The cluster computers all have the same hardware and OS.	The computers that are part of a grid can run different operating systems and have different hardware	The memory, storage device and network communication are managed by the operating system of the basic physical cloud units.
The whole system (all nodes) behaves like a single system view and resources are managed by centralized resource manager	Every node is autonomous i.e. it has its own resource manager and behaves like an independent entity	Every node acts as an independent entity.
The computers in the cluster are normally contained in a single location or complex	Grid are inherently distributed by its nature over a LAN, metropolitan or WAN	Clouds are mainly distributed over MAN

Cluster vs Grid vs Cloud

Cluster Computing	Grid Computing	Cloud Computing
More than 2 computers are connected to solve a problem	A large project is divided among multiple computers to make use of their resources.	It does just the opposite. It allows multiple smaller applications to run at the same time.
Areas of cluster computing 1. Educational resources 2. Commercial sectors for industrial promotion 3. Medical research	Areas of Grid Computing 1.Predictive Modeling and Simulations 2.Engineering Design and Automation 3.Energy Resources Exploration 4.Medical, Military and Basic Research 5.Visualization	Areas of cloud Computing 1.Banking 2.Insurance 3.Weather Forecasting 4.Space Exploration 5.Software as a service 6.PaaS 7.Infrastructure- as -a-Service

Cluster vs Grid vs Cloud

Cluster Computing	Grid Computing	Cloud Computing
Single Ownership	Multiple Ownership	Single Ownership
Resource management is centralized	Resource management is distributed	Resource management is centralized/distributed
User management is centralized	User management is decentralized and also virtual organization (VO)-based	User management is centralized or can be delegated to third party

End