# DISTRIBUTED COMPUTING AND TYPES

Prepared by
AVITA KATAL
Assistant Professor
Systemics Cluster
School of Computer Science, UPES

# COURSE OUTLINE

- **Unit 1 Introduction to Virtualization**

Traditional IT Infrastructure, Benefits of Virtualization, Types of Virtualization, History of Virtualization.

- **Unit 2 Server, Storage, Network and Application Virtualization**

Types of Server Virtualization, Hypervisors, Anatomy of Server Virtualization, Benefits of Storage Virtualization, Types of Storage Virtualization, VPN, VLAN, Benefits of Application Virtualization.

- **Unit 3 Introduction to Cloud Computing**

History, Importance of Virtualization in Cloud, Anatomy of Cloud, Cloud deployment models, Cloud delivery models, stepping stones for the development of cloud, Grid Computing, Cloud Computing.

# COURSE OUTLINE

- **Unit 4 Cloud Implementations / Cloud Deployment Models, Cloud Delivery Models**

Decision Factors for Cloud Implementations, Public, Private and Hybrid Cloud, Overview, Infrastructure as a Service (IaaS) Cloud Delivery Model, Platform as a Service (PaaS) Cloud Delivery Model, Software as a Service (SaaS) Cloud Delivery Model.

- **Unit 5 Case Study on Virtualization, Cloud Workloads**

Customer IT Landscape, Triggers of Virtualization, Preparation for Virtualization, Transition Tools for Virtualization, Cost savings , Cloud workload Overview, Workloads most suitable for Cloud, Workloads not suitable for Cloud.

# CONTENT...

- **TEXT BOOKS:**

1. Virtualization and Cloud Computing (IBM ICE Publication)
2. Mastering in Cloud Computing, Rajkumar Buyya et al.

- **REFERENCE BOOKS**

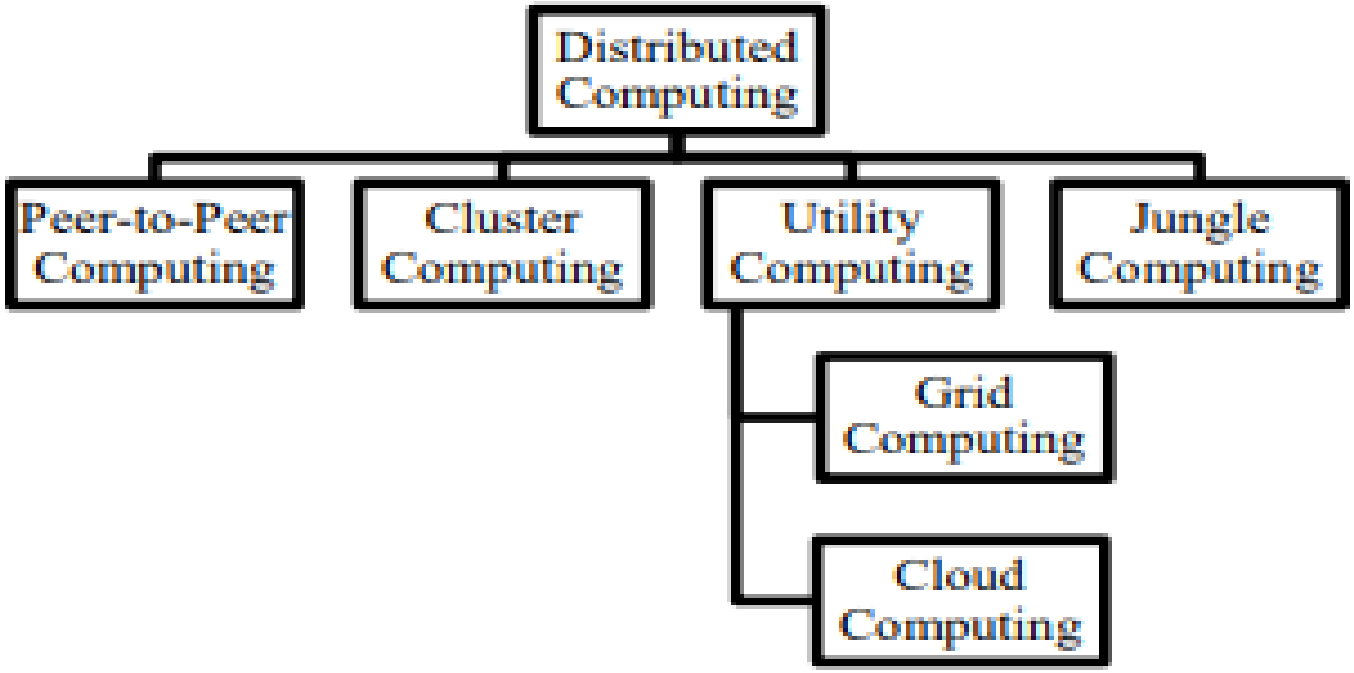
1. Cloud Computing: Fundamentals, Industry Approach and Trends, by Rishab Sharma, Wiley Publication

# CONTENTS



# 1. PATH TOWARDS CLOUD

**Concepts**



# 1.1 DISTRIBUTED COMPUTING

- A distributed system is a network that consists of **autonomous computers** that are connected using a **distribution middleware.** They help in **sharing different resources** and **capabilities** to provide users with a **single and integrated coherent network.**
- Some of the goals of Distributed systems are:

Transparency, Openness, Reliability, Performance & Scalability

# 1.1 DISTRIBUTED COMPUTING CONTD...

The four important goals that should be met for an efficient distributed system are as follows:

**1. Connecting Users and Resources:**

- The main goal of a distributed system is to make it easy for users to access remote resources and to share them with others in a controlled way.
- It is cheaper to share a printer by several users than buying and maintaining printers for each user.
- Collaborating and exchanging information can be made easier by connecting users and resource.

**2. Transparency:**

- It is important for a distributed system to hide the location of its process and resource. A distributed system that can portray itself as a single system is said to be transparent.
- The various transparencies need to be considered are access, location, migration, relocation, replication, concurrency, failure and persistence.
- Aiming for distributed transparency should be considered along with performance issues.

# 1.1 DISTRIBUTED COMPUTING CONTD...

**3. Openness:**

- Openness is an important goal of distributed system in which it offers services according to standard rules that describe the syntax and semantics of those services.
- Open distributed system must be flexible making it easy to configure and add new components without affecting existing components.
- An open distributed system must also be extensible.

**4. Scalable:**

- Scalability is one of the most important goals which are measured along three different dimensions.
- First, a system can be scalable with respect to its size which can add more user and resources to a system.
- Second, users and resources can be geographically apart.
- Third, it is possible to manage even if many administrative organizations are spanned.

# 1.1. CLIENT SERVER NETWORK

- A Computer networking model where one or more powerful computers (servers) provide the different computer network services and all other user of computer network (clients) access those services to perform user's tasks is known as client/server computer networking model.
- In such networks, there exists a central controller called server. A server is a specialized computer that controls the network resources and provides services to other computers in the network.

All other computers in the network are called clients. A client computer receives the requested services from a server.

•A server performs all the major operations like security and network management.

• All the clients communicate with each other via centralized server

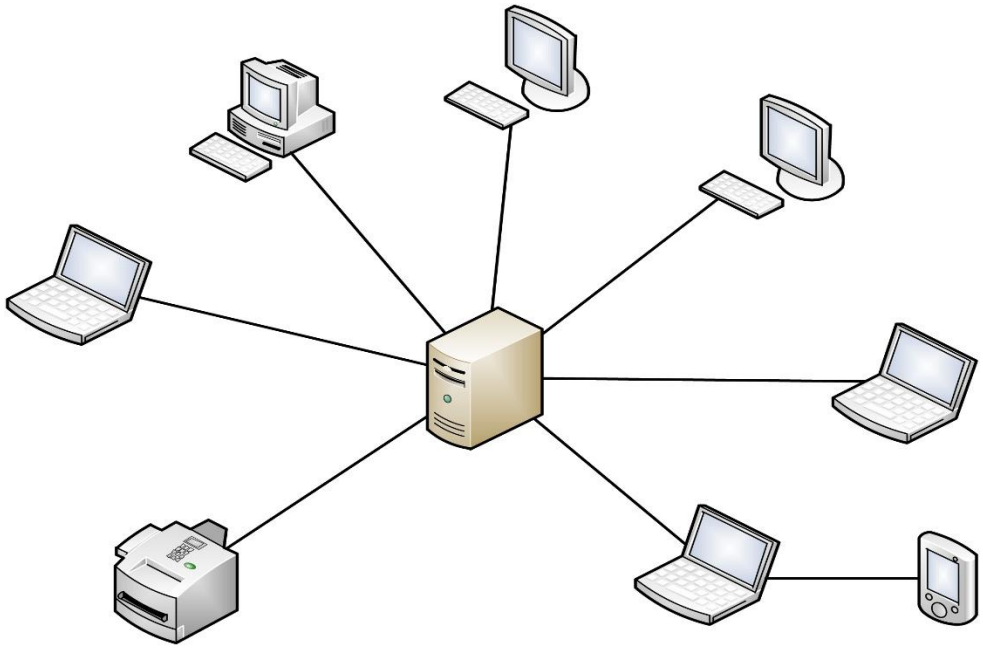
• If client 1 wants to send data to client 2, it first sends request to server to seek permission for it. The server then sends a signal to client 1 allowing it to initiate the communication.

• A server is also responsible for managing all the network resources such as files, directories, applications & shared devices like printer etc.

• If any of the clients wants to access these services, it first seeks permission from the server by sending a request.

- Most Local Area Networks are based on client server relationship.

# 1.1. CLIENT SERVER NETWORK

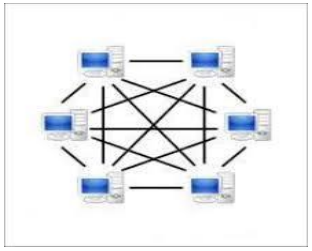


# 1.2 PEER TO PEER NETWORKS

- Peer-to-peer (P2P) networking has been working primarily on the scalability issues inherent in distributing resources over a large number of networked processes.
- In a P2P system, every node acts as **both a client and a server**, providing part of the system resources.
- Peer machines are simply client computers connected to the Internet. All client machines act autonomously to join or leave the system freely.

# 1.2. EXAMPLES OF P2P

- This implies that **no master-slave relationship** exists among the peers.
- **No central coordination** or no central database is needed.
- The concept was popularized by file sharing systems such as the music-sharing application **Napster (originally released in 1999).**



napster

BitTorrent

# 1.3. CLUSTER COMPUTING

A cluster computing comprises **a set of independent or stand-alone computers** and **a network** interconnecting them.

- It works cooperatively together as a **single integrated computing resource.** A cluster is local in that all of its component subsystems are supervised within **a single administrative domain**, usually residing in a single room and managed as a single computer system.
- The components of a cluster are connected to each other through **fast local area networks.** To handle heavy workload with large datasets, clustered computer systems have demonstrated impressive results in the past.
- From the users view point they are multiple machines, but they function as **a single virtual machine**.
- The user's request are received and distributed among all the standalone computers to form a cluster. This results in balanced computational work among different machines, improving the performance of the cluster systems

# POPULAR CLUSTERS

The Borg, a 52-node Beowulf cluster used by the McGill University pulsar group to search for pulsations from binary pulsars

A VAX 11/780, c. 1977

# 1.4. UTILITY COMPUTING

- Utility computing is a **service <u>provisioning</u> model** in which a service provider makes **computing resources** and **infrastructure management** available to the customer as needed, and charges them for specific usage rather than a flat rate. Like other types of **<u>on-demand computing</u>** , the utility model seeks to ***maximize the efficient use of resources and/or minimize associated costs.***

I. Grid Computing

II. Cloud Computing

# 1.4.1. GRID COMPUTING

- Grid computing combines computers from **multiple administrative domains** to reach a common goal, to solve a single task, and may then disappear just as quickly.
- It is analogous to the power grid .
- One of the main strategies of grid computing is to use ***middleware to divide and apportion pieces of a program among several computers***.
- Grid computing involves computation in a distributed fashion, which may also involve the aggregation of large-scale cluster computing based systems.
- The size of a grid may vary from small a network of computer workstations within a corporation to large collaborations across many companies and networks.

# 1.4.1. POPULAR GRID PROJECTS


Search for **extraterrestrial life** by analyzing specific **radio frequencies** emanating from space


Scans/analyzes the collection grid from the **NASA Stardust mission** to capture particles from the coma of comet Wild 2

# DIFFERENCE BETWEEN CLUSTER AND GRID COMPUTING

| Cluster Computing | Grid Computing |
| --- | --- |
| Nodes must be homogeneous i.e. they should have same type of hardware and operating system. | Nodes may have different Operating systems and hardwares. Machines can be homogeneous or heterogeneous. |
| Computers in a cluster are dedicated to the same work and perform no other task. | Computers in a grid contribute their unused processing resources to the grid computing network. |
| Computers are located close to each other. | Computers may be located at a huge distance from one another. |
| Computers are connected by a high speed local area network bus. | Computers are connected using a low speed bus or the internet. |

# DIFFERENCE BETWEEN CLUSTER AND GRID COMPUTING CONTD...

| Cluster Computing | Grid Computing |
|---|---|
| Computers are connected in a centralized network topology. | Computers are connected in a distributed or de-centralized network topology. |
| Scheduling is controlled by a central server. | It may have servers, but mostly each node behaves independently. |
| Whole system has a centralized resource manager. | Every node manages it's resources independently. |
| Whole system functions as a single system. | Every node is autonomous, and anyone can opt out anytime. |
| Cluster computing is used in areas such as WebLogic Application Servers, Databases, etc. | Grid computing is used in areas such as predictive modeling, Automation, simulations, etc. |
| It has Centralized Resource management. | It has Distributed Resource Management. |

# 1.4.2. CLOUD COMPUTING

- A cloud is a type of **parallel and distributed system** consisting of a collection of **interconnected and virtualized computers** that are **dynamically provisioned** and presented as one or more unified computing resources based on **service-level agreement.**

- Cloud computing is a model for enabling **ubiquitous**, **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

# 1.4.2. SERVICE PROVIDERS

- **Amazon web Services**
- **Microsoft Azure**
- **Google Cloud Platform**
- **Adobe**
- **Vmware**
- **IBM Cloud**
- **Rackspace**
- **Red Hat**
- **Salesforce**
- **Oracle Cloud**

# DIFFERENCE BETWEEN GRID AND CLOUD??

| BASIS FOR COMPARISON | CLOUD COMPUTING | GRID COMPUTING |
|---|---|---|
| Application focus | business and web-based applications. | Collaborative purposes. |
| Architecture used | Client-server | Distributed computing |
| Management | Centralized | Decentralized |
| Business model | Pay per use | No defined business model |
| Accessibility of services | High because it is real-time | Low because of scheduled services. |
| Programming models | Eucalyptus, Open Nebula, Open stack etc, for Iaas but no middleware exists. | Different middlewares are available such as Globus gLite, Unicore, etc. |

# DIFFERENCE BETWEEN GRID AND CLOUD??

| BASIS FOR COMPARISON | CLOUD COMPUTING | GRID COMPUTING |
|---|---|---|
| Resource usage patterns | Centralized manner | Collaborative manner |
| Flexibility | High | Low |
| Interoperability | Vendor lock-in and integration are some issues | Easily deals with interoperability between providers. |