Cloud Computing

UNIT CODE : T/615/1644 UNIT LEVEL: 5 CREDIT VALUE: 15

WHAT YOU ARE GOING TO STUDY IN THIS SUBJECT?

LO1. Demonstrate an understanding of the fundamentals of Cloud Computing and its architectures.

LO2. Evaluate the deployment models, service models and technological drivers of Cloud Computing and validate their use.

LO3. Develop Cloud Computing solutions using service provider’s frameworks and open source tools.

LO4. Analyse the technical challenges for cloud applications and assess their risks.

L01 : Demonstrate an understanding of the fundamentals of Cloud Computing and its architectures.

What you are going to study in this chapter?

- Networking Paradigm:
  - Peer-to-peer Computing
  - Client-server Computing
  - Distributed Computing
  - Cluster Computing
  - High-performance Computing
  - Parallel Computing
  - Grid Computing
- Cloud Computing Fundamentals:
  - Definition Of Cloud Computing,
  - Principles Of Cloud Computing,
  - Cloud Ecosystem,
  - Cloud Architecture And Infrastructure,
  - Virtualisation, Network Connectivity,
  - Managing The Cloud, Application Migration To The Cloud

Cloud Computing

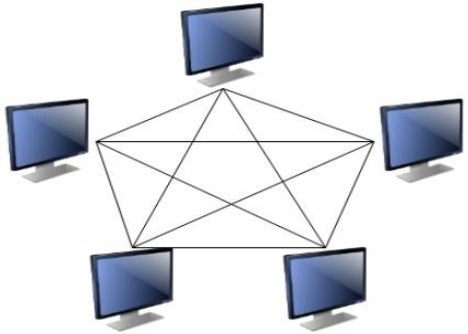
- Cloud computing is a general term for anything that involves delivering hosted services over the internet.
- Cloud infrastructure involves the hardware and software components required for proper implementation of a cloud computing model.
- Cloud computing can also be thought of as utility computing or on-demand computing.

- Cloud computing is the delivery of on-demand IT resources over the internet. The companies that offer these computing services are called **Cloud Service Providers** (CSPs).

Peer-to-peer Computing

The peer to peer computing architecture contains nodes that are equal participants in data sharing. All the tasks are equally divided between all the nodes. The nodes interact with each other as required as share resources.

A diagram to better understand peer to peer computing is as follows −



Characteristics of Peer to Peer Computing

The different characteristics of peer to peer networks are as follows −

- Peer to peer networks are usually formed by groups of a dozen or less computers. These computers all store their data using individual security but also share data with all the other nodes.
- The nodes in peer to peer networks both use resources and provide resources. So, if the nodes increase, then the resource sharing capacity of the peer to peer network increases. This is different than client server networks where the server gets overwhelmed if the nodes increase.
- Since nodes in peer to peer networks act as both clients and servers, it is difficult to provide adequate security for the nodes. This can lead to denial of service attacks.
- Most modern operating systems such as Windows and Mac OS contain software to implement peer to peer networks.

Advantages of Peer to Peer Computing

Some advantages of peer to peer computing are as follows −

- Each computer in the peer to peer network manages itself. So, the network is quite easy to set up and maintain.
- In the client server network, the server handles all the requests of the clients. This provision is not required in peer to peer computing and the cost of the server is saved.
- It is easy to scale the peer to peer network and add more nodes. This only increases the data sharing capacity of the system.
- None of the nodes in the peer to peer network are dependent on the others for their functioning.

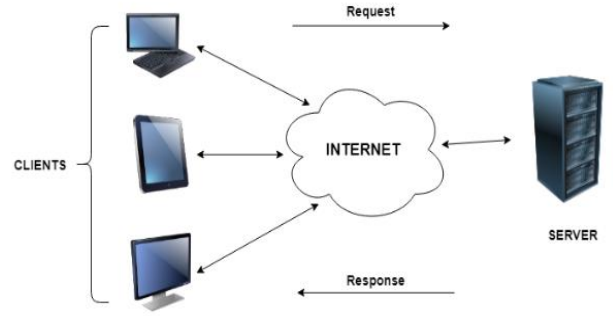
Disadvantages of Peer to Peer Computing

Some disadvantages of peer to peer computing are as follows −

- It is difficult to back up the data as it is stored in different computer systems and there is no central server.
- It is difficult to provide overall security in the peer to peer network as each system is independent and contains its own data.

Client-server Computing

In client server computing, the clients request a resource and the server provides that resource. A server may serve multiple clients at the same time while a client is in contact with only one server. Both the client and server usually communicate via a computer network but sometimes they may reside in the same system.



An illustration of the client server system is given as follows −

Characteristics of Client Server Computing

The salient points for client server computing are as follows:

- The client server computing works with a system of request and response. The client sends a request to the server and the server responds with the desired information.
- The client and server should follow a common communication protocol so they can easily interact with each other. All the communication protocols are available at the application layer.
- A server can only accommodate a limited number of client requests at a time. So it uses a system based to priority to respond to the requests.
- Denial of Service attacks hindera server’s ability to respond to authentic client requests by inundating it with false requests.
- An example of a client server computing system is a web server. It returns the web pages to the clients that requested them.

Difference between Client Server Computing and Peer to Peer Computing

The major differences between client server computing and peer to peer computing are as follows:

- In client server computing, a server is a central node that services many client nodes. On the other hand, in a peer to peer system, the nodes collectively use their resources and communicate with each other.
- In client server computing the server is the one that communicates with the other nodes. In peer to peer to computing, all the nodes are equal and share data with each other directly.
- Client Server computing is believed to be a subcategory of the peer to peer computing.

Advantages of Client Server Computing

The different advantages of client server computing are −

- All the required data is concentrated in a single place i.e. the server. So it is easy to protect the data and provide authorisation and authentication.
- The server need not be located physically close to the clients. Yet the data can be accessed efficiently.
- It is easy to replace, upgrade or relocate the nodes in the client server model because all the nodes are independent and request data only from the server.
- All the nodes i.e clients and server may not be build on similar platforms yet they can easily facilitate the transfer of data.

Disadvantages of Client Server Computing

The different disadvantages of client server computing are −

- If all the clients simultaneously request data from the server, it may get overloaded. This may lead to congestion in the network.
- If the server fails for any reason, then none of the requests of the clients can be fulfilled. This leads of failure of the client server network.
- The cost of setting and maintaining a client server model are quite high.

Distributed Computing

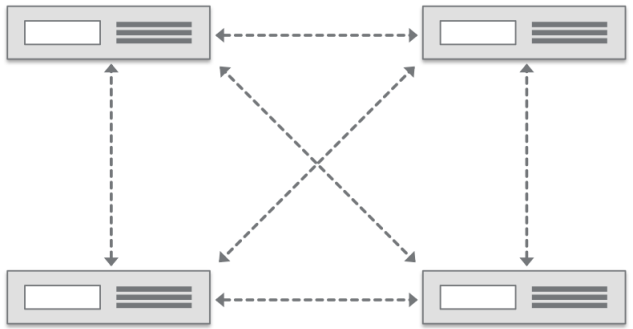
- **Distributed computing** (or distributed processing) is the technique of linking together multiple computer servers over a network into a cluster, to share data and to coordinate processing power.

- Such a cluster is referred to as a “distributed system.” Distributed computing offers advantages in scalability (through a “scale-out architecture”), performance (via parallelism), resilience (via redundancy), and cost-effectiveness (through the use of low-cost, commodity hardware).

- As data volumes have exploded and application performance demands have increased, distributed computing has become extremely common in database and application design.

- This is why it is especially valuable for scaling so that as data volumes grow, that extra load can be handled by simply adding more hardware to the system.

- Contrast this to traditional “big iron” environments consisting of powerful computer servers, in which load growth must be handled by upgrading and replacing the hardware.



In distributed computing, multiple computer servers are tied together across a network to enable large workloads that take advantage of all available resources.

Distributed Computing in Cloud Computing

The growth of cloud computing options and vendors has made distributed computing even more accessible. Although cloud computing instances themselves do not automatically enable distributed computing, there are many different types of distributed computing software that run in the cloud to take advantage of the quickly available computing resources.

Previously, organizations relied on database administrators (DBAs) or technology vendors to link computing resources across networks within and across data centers to be able to share resources. Now, the leading cloud vendors make it easier to add servers to a cluster for additional storage capacity or computing performance.

With the ease and speed in which new computing resources can be provisioned, distributed computing enables greater levels of agility when handling growing workloads. This enables “elasticity,” in which a cluster of computers can be expanded or contracted easily depending on the immediate workload requirements.

Key Advantages

Distributed computing makes all computers in the cluster work together as if they were one computer. While there is some complexity in this multi-computer model, there are greater benefits around:

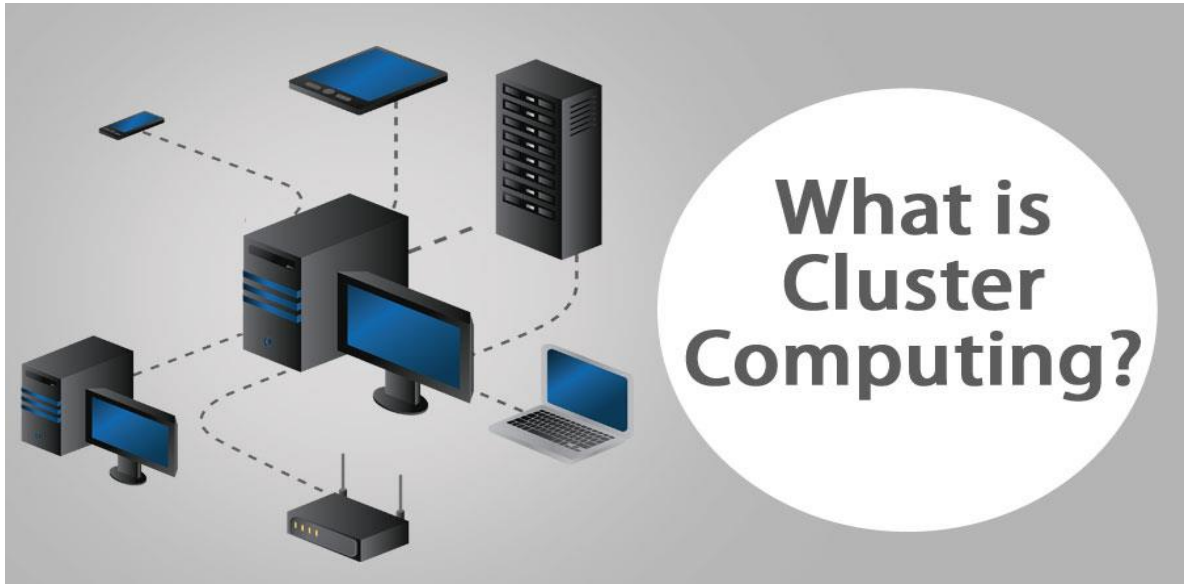
- **Scalability.** Distributed computing clusters are easy to scale through a “scale-out architecture” in which higher loads can be handled by simply adding new hardware (versus replacing existing hardware).
- **Performance.** Through parallelism in which each computer in the cluster simultaneously handles a subset of an overall task, the cluster can achieve high levels of performance through a divide-and-conquer approach.
- **Resilience.** Distributed computing clusters typically copy or “replicate” data across all computer servers to ensure there is no single point of failure. Should a computer fail, copies of the data on that computer are stored elsewhere so that no data is lost.
- **Cost-effectiveness.** Distributed computing typically leverages low-cost, commodity hardware, making initial deployments as well as cluster expansions very economical.

Cluster Computing

Cluster computing is the process of sharing the computation tasks among multiple computers and those computers or machines form the cluster.

It works on the distributed system with the networks. Several types of cluster computing are used based upon the business implementations, performance optimization, and architectural preference such as load balancing clusters, high availability (HA) clusters, high performance (HP) clusters.

Some of the advantages are processing speed, cost efficiency, scalability, high availability of resources. Some of the popular implementations of cluster computing are Google search engine, Earthquake Simulation, Petroleum Reservoir Simulation, and Weather Forecasting system.



Types of Cluster computing

The types of cluster computing are described below.

1. Load-balancing clusters: Here workload is equally distributed across multiple installeservers in the cluster network.

2. High availability (HA) clusters: A group of clusters which ensure to maintain very high availability. computers pulled from these systems are considered to be very much reliable and may not face a downtime even possibly on any instance.

3. High performance (HP) clusters: This computer networking tactic use supercomputers and Cluster computing to resolve complex and highly advanced computation problems.

Advantages

The advantages are mentioned below.

1. Cost efficiency: Compared to highly stable and more storage mainframe computers these form of cluster computing systems are considered to be largely cost-efficient and cheaper. Moreover, most of these systems offer higher performance than mainframe computer systems.

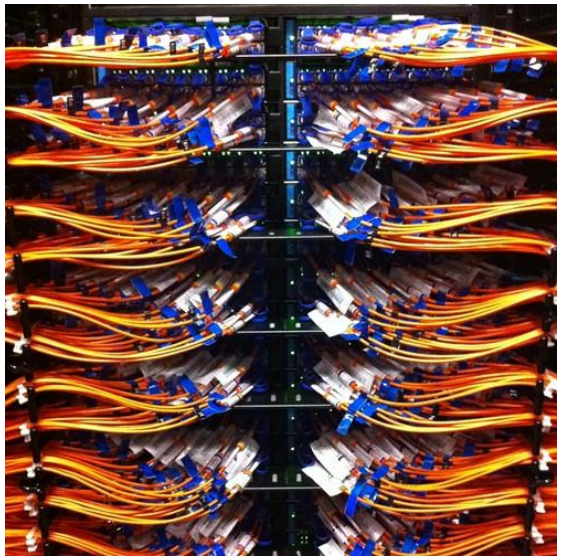
2. Processing speed: The speed of processing also equitable to the mainframe systems and other forms of supercomputers in the market.

3. Expandability: Scalability and expandability is the next key advantage of these clustered systems. because they instantiate the opportunity to add any number of additional resources or systems to the existing computer network.

4. High availability of resources: computers face failure very often. High Availability is concurrent in a straight line to our increasing dependence on computers because at the present they include a vital role mainly in companies whose most important functionality is accurately the offer of some stable computing service, such as e-business, databases, among others. Availability plays the next key role in these systems. failure of one of the currently active nodes may be passed on to the other live nodes and on receiving this message the other set of the node will operate as a proxy for the dead node. so this ensures enhanced availability of these systems.

High-performance Computing

High Performance Computing most generally refers to the practice of aggregating computing power in a way that delivers much higher performance than one could get out of a typical desktop computer or workstation in order to solve large problems in science, engineering, or business.



Scientist need HPC because they hit a tipping point.

At some point in research, there is a need to:

- Expand the current study area (regional → national → global)
- Integrate new data
- Increase model resolution

Some typical computational barriers:

- Time – processing on local systems is too slow or not feasible
- CPU Capacity -- Can only run one model at a time
- Develop, implement, and disseminate state-of-the-art techniques and tools so that models are more effectively applied to today's decision-making
- Management of Computer Systems – Science Groups don't want to purchase and manage local computer systems – they want to focus on science

Parallel Computing

Parallel computing is a type of computation in which many calculations or processes are carried out simultaneously. Large problems can often be divided into smaller ones, which can then be solved at the same time. There are several different forms of parallel computing: bit-level, instruction-level, data, and task parallelism.

It is the use of multiple processing elements simultaneously for solving any problem. Problems are broken down into instructions and are solved concurrently as each resource that has been applied to work is working at the same time.

Advantages of Parallel Computing over Serial Computing are as follows:

1. It saves time and money as many resources working together will reduce the time and cut potential costs.
2. It can be impractical to solve larger problems on Serial Computing.
3. It can take advantage of non-local resources when the local resources are finite.
4. Serial Computing ‘wastes’ the potential computing power, thus Parallel Computing makes better work of the hardware.

Why parallel computing?

- The whole real-world runs in dynamic nature i.e. many things happen at a certain time but at different places concurrently. This data is extensively huge to manage.
- Real-world data needs more dynamic simulation and modeling, and for achieving the same, parallel computing is the key.
- Parallel computing provides concurrency and saves time and money.
- Complex, large datasets, and their management can be organized only and only using parallel computing's approach.
- Ensures the effective utilization of the resources. The hardware is guaranteed to be used effectively whereas in serial computation only some part of the hardware was used and the rest rendered idle.
- Also, it is impractical to implement real-time systems using serial computing.

Applications of Parallel Computing:

- Databases and Data mining.
- Real-time simulation of systems.
- Science and Engineering.
- Advanced graphics, augmented reality, and virtual reality.

Limitations of Parallel Computing:

- It addresses such as communication and synchronization between multiple sub-tasks and processes which is difficult to achieve.
- The algorithms must be managed in such a way that they can be handled in a parallel mechanism.
- The algorithms or programs must have low coupling and high cohesion. But it's difficult to create such programs.
- More technically skilled and expert programmers can code a parallelism-based program well.

Grid Computing

- **Grid Computing** can be defined as a network of computers working together to perform a task that would rather be difficult for a single machine.
- All machines on that network work under the same protocol to act as a virtual supercomputer.
- The task that they work on may include analyzing huge datasets or simulating situations that require high computing power.
- Computers on the network contribute resources like processing power and storage capacity to the network.

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

- It can also be seen as a form of Parallel Computing where instead of many CPU cores on a single machine, it contains multiple cores spread across various locations.

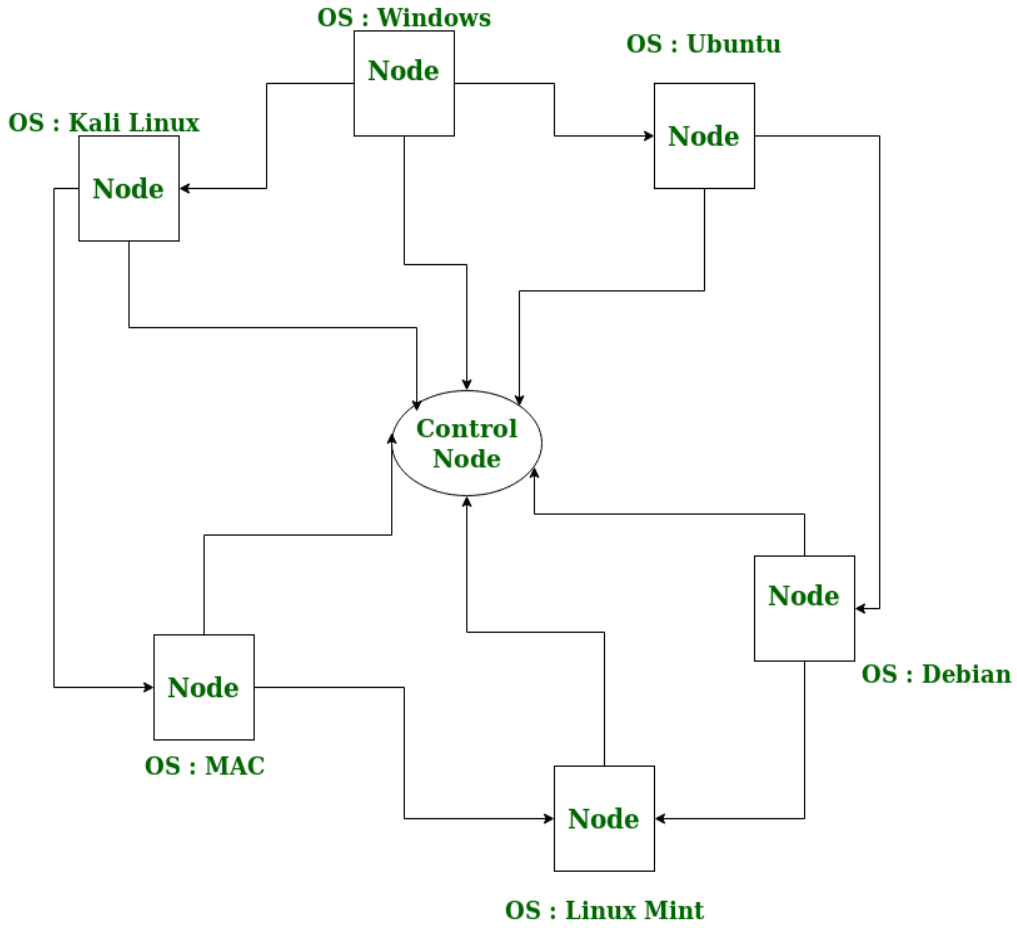
- The concept of grid computing isn't new, but it is not yet perfected as there are no standard rules and protocols established and accepted by people.

Advantages of Grid Computing:

- It is not centralized, as there are no servers required, except the control node which is just used for controlling and not for processing.
- Multiple heterogeneous machines i.e. machines with different Operating Systems can use a single grid computing network.
- Tasks can be performed parallels across various physical locations and the users don’t have to pay for them (with money).

Disadvantages of Grid Computing:

- The software of the grid is still in the involution stage.
- A super-fast interconnect between computer resources is the need of hour.
- Licensing across many servers may make it prohibitive for some applications.
- Many groups are reluctant with sharing resources.



Topology in Grid Computing