

What is cloud computing?

[Technology](#)



Public cloud :-Public cloud is same as internet based on the cloud computing standard model. Service provider uses the web to make all the services available to the client. Many popular cloud services are Google app engine, salesforce. com.

Private cloud :-Private clouds are owned by individual organization and give security to its resources.

Hybrid cloud:-The combination of both public and private is hybrid cloud.

These two are mixed to take advantage of both and create more value [8].

Three different services model in cloud computing are :

1. Infrastructure as a service (IAAS): . IaaS include the services in which infrastructure like hardware, network, connectivity and storage is provided to the consumer
2. Platform as a service(PAAS): PaaS readymade environment is provided as a service for managing, developing and testing the software application
3. Software as a service(SAAS): SaaS refers to the category of service in which software are provided over the internet so neither need to install software nor have to pay cost to purchase the software and licence [7].

Various benefits of cloud computing: Scalability, Simplicity, Security and Storage

Load Balancing:-Load balancing is an even division of processing work in the cloud computing which have a great effect on the performance of a framework, dependent on the measure of work allotted to a framework for a given period of time. It is the way toward adjusting the amount of workload

among framework resources for enhancing system performance and resource use [9].

Load balancing can minimize the reaction time, amplify the client's satisfaction, advances the system potency and can reduce the assignment dismissal [10]. So overall load balancing try that there should not be a situation when one of the server or data center is under utilization and other get overloaded [11].

The objectives of the study of the load balancing techniques are as follow:

- Improve both resource deployment and job response time while also avoiding a situation where some of the nodes are having a huge amount of load while other nodes are doing nothing or idle.
- To reduce operational cost, better performance in terms of response time and data processing time, maintain the system consistency.

Example of Load Balancing

Load balancer has to follow some steps to perform the load balancing. These steps are:

- Receive incoming service requests from various clients.
- Calculate requested load size of the incoming load request from clients and build a request queue.
- Check the current load status of the serves in the server pool repeatedly using a server monitor daemon.
- Use a load balancing method/algorithm to select appropriate server.

All the provider use different types of load balancing strategies to achieve the different goals of load balancing. There are different metrics to measure that up to which extent load balancing goals achieved. Following are the metrics in existing load balancing techniques[14], [15]:

- **Throughput:** It is used to calculate the total number of tasks get completed in specific period. High value of throughput shows that load balancing is implemented properly.
- **Response time:** It is the measure of time taken by any load balancing strategy to react or respond in a framework. Low value of this time shows that load balancing is implemented in a good way.
- **Fault tolerance:** It is the ability of a load balancing strategy to continue nonstop when a hardware failure occurs. A load balancing strategy should have good fault tolerance techniques.
- **Scalability:** A system with finite number of resources should have ability to grow and manage increased load. An efficient algorithm should have optimized scalability.

Performance: Performance refers to the overall proficiency or productivity of a framework. This must be enhanced at a sensible cost.

- **Resource utilization:** All the resources should be used in a good manner. It ought to be enhanced for a beneficial load balancing method.