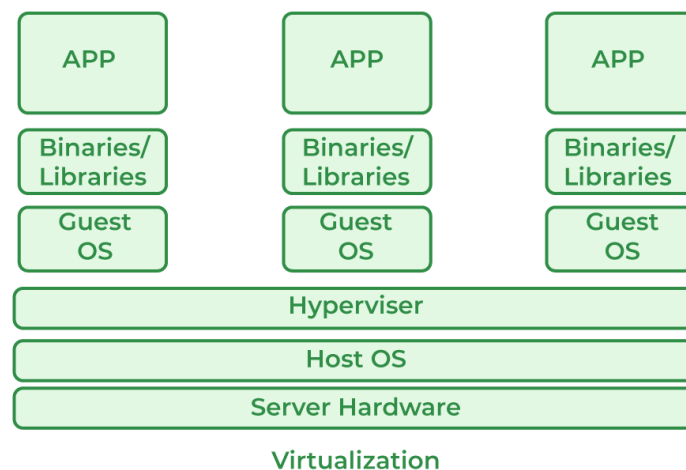


Virtualization

Virtualization is a technique how to separate a service from the underlying physical delivery of that service. It is the process of creating a virtual version of something like computer hardware. It was initially developed during the mainframe era. It involves using specialized software to create a virtual or software-created version of a computing resource rather than the actual version of the same resource. With the help of Virtualization, multiple operating systems and applications can run on the same machine and its same hardware at the same time, increasing the utilization and flexibility of hardware.



Host Machine: The machine on which the virtual machine is going to be built is known as Host Machine.

Guest Machine: The virtual machine is referred to as a Guest Machine.

Virtualization has a prominent impact on Cloud Computing. In the case of cloud computing, users store data in the cloud, but with the help of Virtualization, users have the extra benefit of sharing the infrastructure. Cloud Vendors take care of the required physical resources, but these cloud providers charge a huge amount for these services which impacts every user or organization. Virtualization helps Users or Organisations in maintaining those services which are required by a company through external (third-party) people, which helps in reducing costs to the company. This is the way through which Virtualization works in Cloud Computing.

Benefits of Virtualization

- More flexible and efficient allocation of resources.

- Enhance development productivity.
- It lowers the cost of IT infrastructure.
- Remote access and rapid scalability.
- High availability and disaster recovery.
- Pay per use of the IT infrastructure on demand.
- Enables running multiple operating systems.

Drawback of Virtualization

- **High Initial Investment:** Clouds have a very high initial investment, but it is also true that it will help in reducing the cost of companies.
- **Learning New Infrastructure:** As the companies shifted from Servers to Cloud, it requires highly skilled staff who have skills to work with the cloud easily, and for this, you have to hire new staff or provide training to current staff.
- **Risk of Data:** Hosting data on third-party resources can lead to putting the data at risk, it has the chance of getting attacked by any hacker or cracker very easily.

Characteristics of Virtualization

- **Increased Security:** The ability to control the execution of a guest program in a completely transparent manner opens new possibilities for delivering a secure, controlled execution environment. All the operations of the guest programs are generally performed against the virtual machine, which then translates and applies them to the host programs.
- **Managed Execution:** In particular, sharing, aggregation, emulation, and isolation are the most relevant features.
- **Sharing:** Virtualization allows the creation of a separate computing environment within the same host.
- **Aggregation:** It is possible to share physical resources among several guests, but virtualization also allows aggregation, which is the opposite process.

Types of Virtualization

1. Application Virtualization
2. Network Virtualization
3. Desktop Virtualization

4. Storage Virtualization
5. Server Virtualization
6. Data virtualization

1. Application Virtualization:

Application virtualization helps a user to have remote access to an application from a server. The server stores all personal information and other characteristics of the application but can still run on a local workstation through the internet. An example of this would be a user who needs to run two different versions of the same software. Technologies that use application virtualization are hosted applications and packaged applications.

2. Network Virtualization:

The ability to run multiple virtual networks with each having a separate control and data plan. It co-exists together on top of one physical network. It can be managed by individual parties that are potentially confidential to each other. Network virtualization provides a facility to create and provision virtual networks, logical switches, routers, firewalls, load balancers, Virtual Private Networks (VPN), and workload security within days or even weeks.

3. Desktop Virtualization:

Desktop virtualization allows the users' OS to be remotely stored on a server in the data center. It allows the user to access their desktop virtually, from any location by a different machine. Users who want specific operating systems other than Windows Server will need to have a virtual desktop. The main benefits of desktop virtualization are user mobility, portability, and easy management of software installation, updates, and patches.

4. Storage Virtualization:

Storage virtualization is an array of servers that are managed by a virtual storage system. The servers aren't aware of exactly where their data is stored and instead function more like worker bees in a hive. It makes managing storage from multiple sources be managed and utilized as a single repository. storage virtualization software maintains smooth operations, consistent performance, and a continuous suite of advanced functions despite changes, breaks down, and differences in the underlying equipment.

5. Server Virtualization:

This is a kind of virtualization in which the masking of server resources takes place. Here, the central server (physical server) is divided into multiple different virtual servers by changing the identity number, and processors. So, each system can operate its operating systems in an isolated manner. Where each sub-server knows the identity of the central server. It causes an increase in performance and reduces the operating cost by the deployment of main server resources into a sub-server resource. It's beneficial in virtual migration, reducing energy consumption, reducing infrastructural costs, etc.

6. Data Virtualization:

This is the kind of virtualization in which the data is collected from various sources and managed at a single place without knowing more about the technical information like how data is collected, stored & formatted then arranged that data logically so that its virtual view can be accessed by its interested people and stakeholders, and users through the various cloud services remotely. Many big giant companies are providing their services like Oracle, IBM, At scale, Cdata, etc.

Load Balancing

Load balancing is the method that allows you to have a proper balance of the amount of work being done on different pieces of device or hardware equipment. Typically, what happens is that the load of the devices is balanced between different servers or between the CPU and hard drives in a single cloud server.

Load balancing was introduced for various reasons. One of them is to improve the speed and performance of each single device, and the other is to protect individual devices from hitting their limits by reducing their performance.

Cloud load balancing is defined as dividing workload and computing properties in cloud computing. It enables enterprises to manage workload demands or application demands by distributing resources among multiple computers, networks or servers. Cloud load balancing involves managing the movement of workload traffic and demands over the Internet.

Traffic on the Internet is growing rapidly, accounting for almost 100% of the current traffic annually. Therefore, the workload on the servers is increasing so rapidly, leading to

overloading of the servers, mainly for the popular web servers. There are two primary solutions to overcome the problem of overloading on the server-

First is a single-server solution in which the server is upgraded to a higher-performance server. However, the new server may also be overloaded soon, demanding another upgrade. Moreover, the upgrading process is arduous and expensive.

The second is a multiple-server solution in which a scalable service system on a cluster of servers is built. That's why it is more cost-effective and more scalable to build a server cluster system for network services.

Cloud-based servers can achieve more precise scalability and availability by using farm server load balancing. Load balancing is beneficial with almost any type of service, such as HTTP, SMTP, DNS, FTP, and POP/IMAP.

It also increases reliability through redundancy. A dedicated hardware device or program provides the balancing service.

Different Types of Load Balancing Algorithms in Cloud Computing:

1. Static Algorithm

Static algorithms are built for systems with very little variation in load. The entire traffic is divided equally between the servers in the static algorithm. This algorithm requires in-depth knowledge of server resources for better performance of the processor, which is determined at the beginning of the implementation.

However, the decision of load shifting does not depend on the current state of the system. One of the major drawbacks of static load balancing algorithm is that load balancing tasks work only after they have been created. It could not be implemented on other devices for load balancing.

2. Dynamic Algorithm

The dynamic algorithm first finds the lightest server in the entire network and gives it priority for load balancing. This requires real-time communication with the network which can help increase the system's traffic. Here, the current state of the system is used to control the load.

The characteristic of dynamic algorithms is to make load transfer decisions in the current system state. In this system, processes can move from a highly used machine to an underutilized machine in real time.

3. Round Robin Algorithm

Round robin load balancing algorithm uses round-robin method to assign jobs. First, it randomly selects the first node and assigns tasks to other nodes in a round-robin manner. This is one of the easiest methods of load balancing.

Processors assign each process circularly without defining any priority. It gives fast response in case of uniform workload distribution among the processes. All processes have different loading times. Therefore, some nodes may be heavily loaded, while others may remain under-utilised.

4. Weighted Round Robin Load Balancing Algorithm

Weighted Round Robin Load Balancing Algorithms have been developed to enhance the most challenging issues of Round Robin Algorithms. In this algorithm, there are a specified set of weights and functions, which are distributed according to the weight values.

Processors that have a higher capacity are given a higher value. Therefore, the highest loaded servers will get more tasks. When the full load level is reached, the servers will receive stable traffic.

5. Opportunistic Load Balancing Algorithm

The opportunistic load balancing algorithm allows each node to be busy. It never considers the current workload of each system. Regardless of the current workload on each node, OLB distributes all unfinished tasks to these nodes.

The processing task will be executed slowly as an OLB, and it does not count the implementation time of the node, which causes some bottlenecks even when some nodes are free.

6. Minimum to Minimum Load Balancing Algorithm

Under minimum to minimum load balancing algorithms, first of all, those tasks take minimum time to complete. Among them, the minimum value is selected among all the functions. According to that minimum time, the work on the machine is scheduled.

Other tasks are updated on the machine, and the task is removed from that list. This process will continue till the final assignment is given. This algorithm works best where many small tasks outweigh large tasks.

Load balancing solutions can be categorized into two types -

Software-based load balancers: Software-based load balancers run on standard hardware (desktop, PC) and standard operating systems.

Hardware-based load balancers: Hardware-based load balancers are dedicated boxes that contain application-specific integrated circuits (ASICs) optimized for a particular use. ASICs allow network traffic to be promoted at high speeds and are often used for transport-level load balancing because hardware-based load balancing is faster than a software solution.

Major Examples of Load Balancers

Direct Routing Request Dispatch Technique: This method of request dispatch is similar to that implemented in IBM's NetDispatcher. A real server and load balancer share a virtual IP address. The load balancer takes an interface built with a virtual IP address that accepts request packets and routes the packets directly to the selected server.

Dispatcher-Based Load Balancing Cluster: A dispatcher performs smart load balancing using server availability, workload, capacity and other user-defined parameters to regulate where TCP/IP requests are sent. The dispatcher module of a load balancer can split HTTP requests among different nodes in a cluster. The dispatcher divides the load among multiple servers in a cluster, so services from different nodes act like a virtual service on only one IP address; Consumers interconnect as if it were a single server, without knowledge of the back-end infrastructure.

Linux Virtual Load Balancer: This is an open-source enhanced load balancing solution used to build highly scalable and highly available network services such as HTTP, POP3, FTP,

SMTP, media and caching, and Voice over Internet Protocol (VoIP) is done. It is a simple and powerful product designed for load balancing and fail-over. The load balancer itself is the primary entry point to the server cluster system. It can execute Internet Protocol Virtual Server (IPVS), which implements transport-layer load balancing in the Linux kernel, also known as layer-4 switching.

Types of Load Balancing

Network Load Balancing

Cloud load balancing takes advantage of network layer information and leaves it to decide where network traffic should be sent. This is accomplished through Layer 4 load balancing, which handles TCP/UDP traffic. It is the fastest local balancing solution, but it cannot balance the traffic distribution across servers.

HTTP(S) load balancing

HTTP(s) load balancing is the oldest type of load balancing, and it relies on Layer 7. This means that load balancing operates in the layer of operations. It is the most flexible type of load balancing because it lets you make delivery decisions based on information retrieved from HTTP addresses.

Internal Load Balancing

It is very similar to network load balancing, but is leveraged to balance the infrastructure internally.

Load balancers can be further divided into hardware, software and virtual load balancers.

Hardware Load Balancer

It depends on the base and the physical hardware that distributes the network and application traffic. The device can handle a large traffic volume, but these come with a hefty price tag and have limited flexibility.

Software Load Balancer

It can be an open source or commercial form and must be installed before it can be used. These are more economical than hardware solutions.

Virtual Load Balancer

It differs from a software load balancer in that it deploys the software to the hardware load-balancing device on the virtual machine.

WHY CLOUD LOAD BALANCING IS IMPORTANT IN CLOUD COMPUTING?

Here are some of the importance of load balancing in cloud computing.

Offers better performance

The technology of load balancing is less expensive and also easy to implement. This allows companies to work on client applications much faster and deliver better results at a lower cost.

Helps Maintain Website Traffic

Cloud load balancing can provide scalability to control website traffic. By using effective load balancers, it is possible to manage high-end traffic, which is achieved using network equipment and servers. E-commerce companies that need to deal with multiple visitors every second use cloud load balancing to manage and distribute workloads.

Can Handle Sudden Bursts in Traffic

Load balancers can handle any sudden traffic bursts they receive at once. For example, in case of university results, the website may be closed due to too many requests. When one uses a load balancer, he does not need to worry about the traffic flow. Whatever the size of the traffic, load balancers will divide the entire load of the website equally across different servers and provide maximum results in minimum response time.

Greater Flexibility

The main reason for using a load balancer is to protect the website from sudden crashes. When the workload is distributed among different network servers or units, if a single node

fails, the load is transferred to another node. It offers flexibility, scalability and the ability to handle traffic better. Because of these characteristics, load balancers are beneficial in cloud environments. This is to avoid heavy workload on a single server.