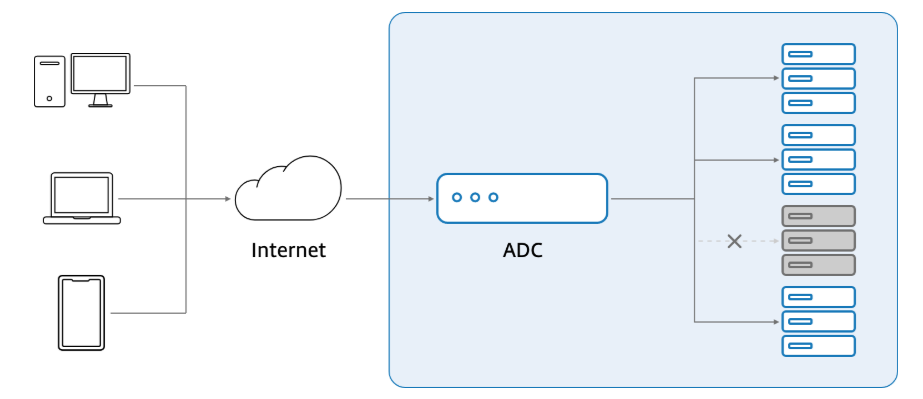
Mechanism /methodology in which we effectively distribute the network traffic across multiple servers in a server pool/farm.

Load balancing site between the client and backend server receive the request from multiple clients and distribute the traffic among multiple server.

What are load balancers and how do they work?

1. Physical device or virtualized instance running on the specific hardware or software process.
2. Incorporated into ADCs application deliver controller designed for improving a the application performance and security in 3 tire application and microservices based application
3. Able to leverage many load balancing mechanism round robin methods, server response time and least connection method.



Load balancer detects health of the backend components and resource and send the traffic to the server which can fullfill the request.

It minimizes the server response time and maximizes the throughput of the application and increases the availability of the application.

It acts as application traffic cop that systematically rout the request to the right location at any given point of time there by preventing costly bottleneck and unforeseen incidents.

### **Hardware- vs software-based load balancers**

1.High performance device capable of securely handling multiple traffic from the client.

2. Also contained built in visualization capability which virtualize several instance of load balancers.

3. Allow multi-tenant architecture with complete isolation between the systems.

1. Completely replaces hardware based load balancing with superior functionality.
2. Reduce space and hardware cost
3. Run on common hypervisor, in container with minimal over head

L4, L7 and GSLB load balancers, explained

Digital work places are highly application driven so there is a high demand for the availability of the application. So that customer is not facing any down time.

L4: Layer 4 where in routing decision are made in the transport layer using tcp/upd ports that takes packets from source IP to target IP

L7: Routing decision are made at the top level in layer 7 .Evaluate wider range of the data

GSLB: Global server load balancing manages the load balancing across multiple servers Across multiple data centres so that huge traffic can be addressed.

GSLB used for the applications hosted on the cloud servers.

## What are some of the common load balancing algorithms?

1. Round robin

Easy to implement

Virtual server assign tasks to other server in an ordered list

But won’t consider if the server is already overloaded with tasks.

1. Least connection method

Sophisticated algorithm where we check for the current health of the server and decided if this server can service the request

Server with least request will be assigned with new task

1. Least response time

Based on the server response time to the health monitor request we decide which server can basically serve for this request.

1. Least bandwidth

Based on the amount pf traffic on the server its measured in MBps who can serve the request

1. Hashing method

Varies data sent over the network will be hashed like Ip, header etc

1. Custom load

Query load on individual severe and then one with least load will be assigned with the task.

And ADc with incorporated load balance will help The it industry to leverage the availability and scalability of the application.

ADC offers many other functionalities like securing, managing, monitoring application environments and ensuring the best end user experience

Makes sure the load is evenly distributed so the user never experiences down time and application always available for the user

Continuously checks for the server health to verify if the server is ready to serve the request.

Some load balances will also trigger the new virtualized server to meet the increase in demand of the resource.

Distributing traffic across multiple severs makes sure the no single sever is overloaded with too many tasks or loads.

I improve the responsiveness of the application and makes your application available to user all the time

Modern software load balances have leveraged the capability security of the application.

Load balancer manages the flow of information between the client system and the server.

Makes sure server moves the data efficiently, Optimize the use of application delivery resource and improve the responsiveness of the application

Continuously monitors application performance and ensures that they handle the request. Even thy can remove server from the pool if it is not performing well.

### Load Balancing and SSL

Secured socket layer is standard technology to establish secure link between the client computer and server

Load balancing scenarios

SSL decryption happens at the load balancers end before the traffic is distributed to the client browser. Its called SSL termination. With this it eliminates the overhead on the server to perform decryption.

This will lead to risk of request being intercepted between the load balance and servers because traffic between load balancer and server is not secured.

This will be lessened if both server and load balancer hosted on same data centre.

Or else always transmit the request to the server to put over head on him this might need some extra process cycle on the server.

### Load Balancing and Security

With the cloud architecture we can move the over head of the security to the cloud. Because the offloading job of the load balancer will be performed at public cloud which will take the over head of the security.

## Load Balancing Benefits

## More than a network traffic cop it does lot of other stuffs.

## Software based load balancer proved the predictive analytics that determines the bottleneck for server load before it happens. SO it gives and actionable insight for the organization.

## In seven layers of OSI model

## OSI- Operating system interconnection.

## Firewall present between

## Physical

## Data

## Network layer

## Between layer 4 to 7 load balance is present

## Transport layer

## Session layer

## Presentation layer

## Application layer.

## Load balancer can be present in any of these 4 layer.

## Layer 4: transmit the traffic based on the data and TCP protocol.

## L7: Content switching to lad balancer. Routing decision based on HTTP header , Uri, SSL session id HTML data.

## GSLB: Global server load balancing extends the capabilities of L$ and L7 in different geographical location. Cloud

## Software Load Balancers vs. Hardware Load Balancers

## Hardware based always runs on top of appliances proprietary software optimized to run on customer processor. As traffic increases vendor simply adds more loads.

## Software defined always run with the less coast.

* **SDN**— Load balancing using [SDN (software-defined networking)](https://avinetworks.com/glossary/sdn-load-balancing/) separates the control plane from the data plane for application delivery. This allows the control of multiple load balancing. It also helps the network to function like the virtualized versions of compute and storage. With the centralized control, networking policies and parameters can be programmed directly for more responsive and efficient application services. This is how networks can become more agile.
* **UDP**— A UDP load balancer utilizes User Datagram Protocol (UDP). UDP load balancing is often used for live broadcasts and online games when speed is important and there is little need for error correction. UDP has low latency because it does not provide time-consuming health checks.
* **TCP**— A TCP load balancer uses transmission control protocol (TCP). [TCP load balancing](https://avinetworks.com/glossary/tcp-load-balancing/) provides a reliable and error-checked stream of packets to IP addresses, which can otherwise easily be lost or corrupted.
* **SLB**— Server Load Balancing (SLB) provides network services and content delivery using a series of load balancing algorithms. It prioritizes responses to the specific requests from clients over the network. Server load balancing distributes client traffic to servers to ensure consistent, high-performance application delivery.
* **Virtual** — Virtual load balancing aims to mimic software-driven infrastructure through virtualization. It runs the software of a physical load balancing appliance on a virtual machine. [Virtual load balancers](https://avinetworks.com/glossary/virtual-load-balancer/), however, do not avoid the architectural challenges of traditional hardware appliances which include limited scalability and automation, and lack of central management.
* **Elastic** — [Elastic Load Balancing](https://avinetworks.com/glossary/elastic-load-balancer/) scales traffic to an application as demand changes over time. It uses system health checks to learn the status of application pool members (application servers) and routes traffic appropriately to available servers, manages fail-over to high availability targets, or automatically spins-up additional capacity.
* **Geographic** — Geographic load balancing redistributes application traffic across data centers in different locations for maximum efficiency and security. While local load balancing happens within a single data center, [geographic load balancing](https://avinetworks.com/glossary/geographic-load-balancing/) uses multiple data centers in many locations.
* **Multi-site** — Multi-site load balancing, also known as global server load balancing (GSLB), distributes traffic across servers located in multiple sites or locations around the world. The servers can be on-premises or hosted in a public or private cloud. [Multi-site load balancing](https://avinetworks.com/glossary/multi-site-load-balancing/) is important for quick disaster recovery and business continuity after a disaster in one location renders a server inoperable.
* **Load Balancer as a Service (LBaaS)** — Load Balancer as a Service (LBaaS) uses advances in load balancing technology to meet the agility and application traffic demands of organizations implementing private cloud infrastructure. Using an as-a-service model, [LBaaS](https://avinetworks.com/glossary/load-balancing-as-a-service/) creates a simple model for application teams to spin up load balancers.