Distributed Applications

Group 6

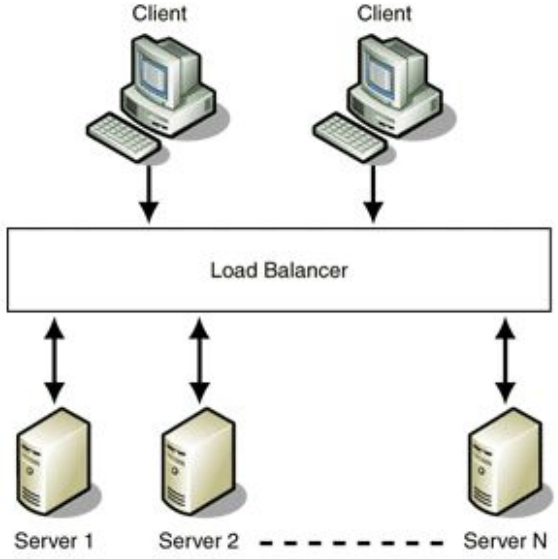
- Eric Han
- Mayuri Wadkar
- Sonali Mishra

Load Balancers : What are they?

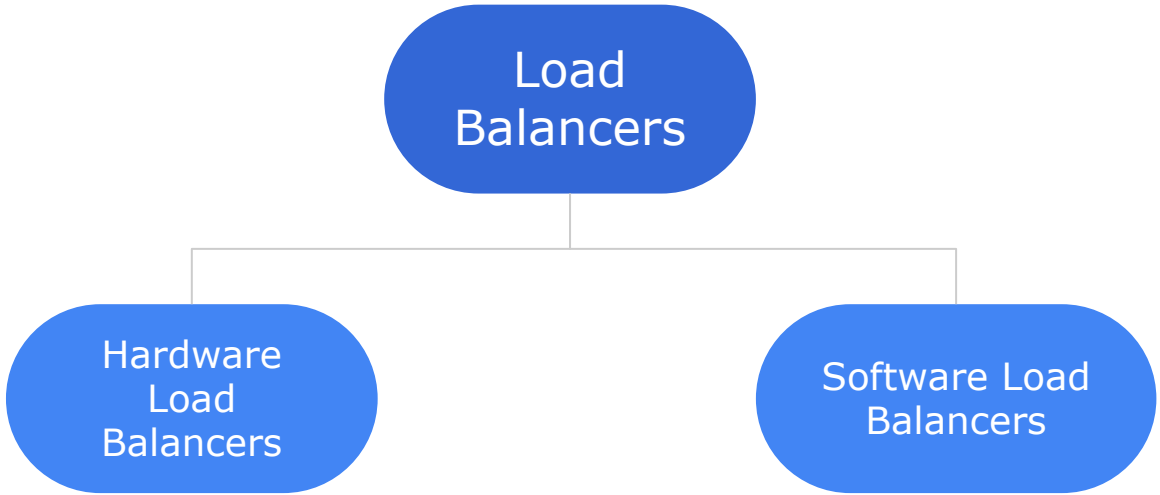
- For applications or websites that get a lot of hits, server might be under heavy load.
- It might be required to distribute the load across multiple servers in this case.
- Load Balancers distribute the workload of a system to multiple systems.
- Amount of workload on an individual system reduced.
- Increased reliability, efficiency, availability of the application or website.

Role of Load Balancers

- To balance load amount among a cluster of nodes so no one node receives too much traffic/workload.
- Load Balancers ensure availability, performance, and stability to end customers/applications.



Types of Load Balancers



Hardware Load Balancers

- Are deployed between the servers and the client.
- Often referred to as specialised routers or switches.
- However, can also be a dedicated system between the client and server, to balance load.
- Hardware Load Balancers are implemented on
    - Layer4 - Transport Layer
    - Layer 7 - Application Layer of OSI model.

Layer 4 - Hardware Load Balancer

DNS Load Balancing:

- The Domain Name Servers are configured to return different ip-address for different systems.
- This approach creates a load balancing effect whenever there is a dns lookup.

Layer 4 - Hardware Load Balancer

Direct routing:

- All the incoming traffic is routed by the load balancer, however all the outgoing traffic directly reaches the client.
- This makes it a super fast load balancing configuration.

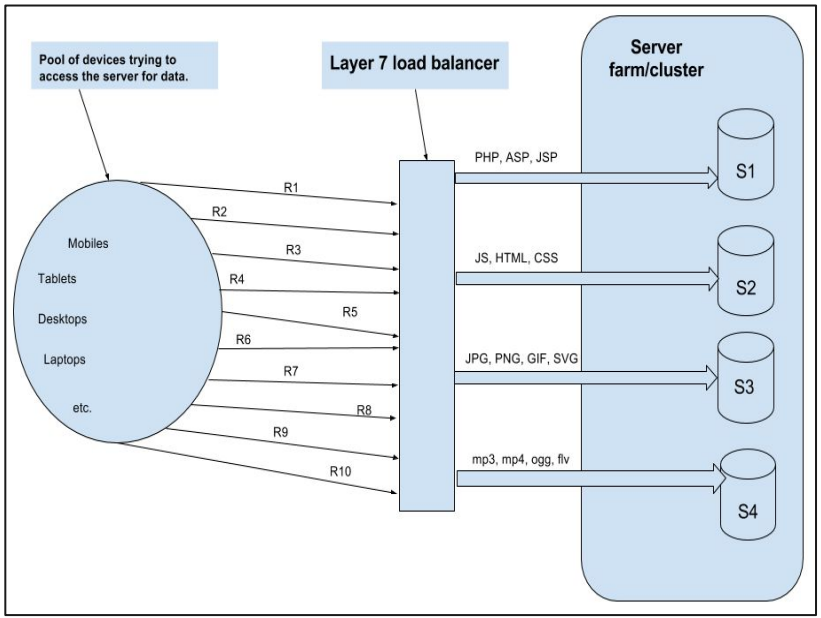
Layer 4 - Hardware Load Balancer

Tunnel or IP Tunnelling:

- Client sends request to the virtual IP of load balancer.
- Load Balancer encapsulates the IP packets, keeps a hash table and distributes it to the different servers.

Layer 7 - Hardware Load Balancer

- These type of load balancers make decisions according to the actual content on the message.
- For example an image will go to an image server, request for PHP scripts may go to another server, media content may go to another server, and so on.

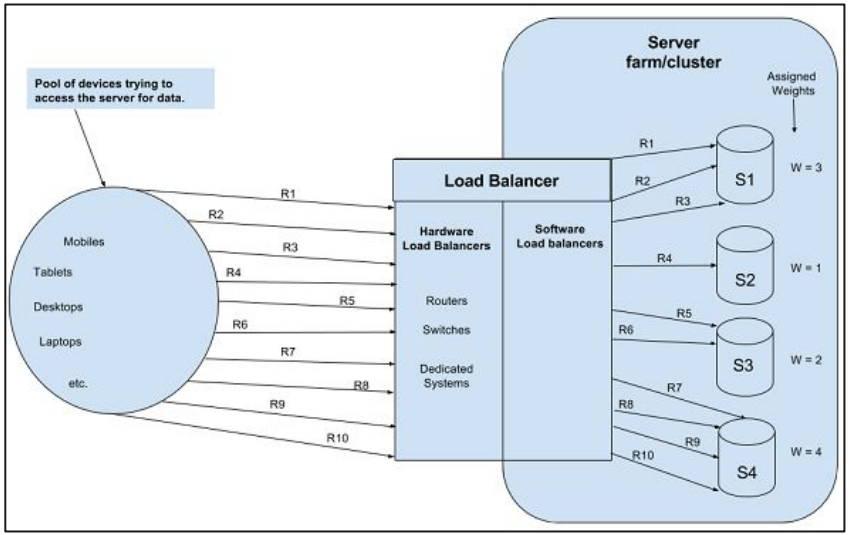


Software Load Balancers

- Software load balancers are usually implemented as a combination of one or more scheduling algorithms.
- Three basic scheduling algorithms used by load balancers are:
  - Weighted Scheduling Algorithm
  - Round Robin Scheduling
  - Least Connection First Scheduling

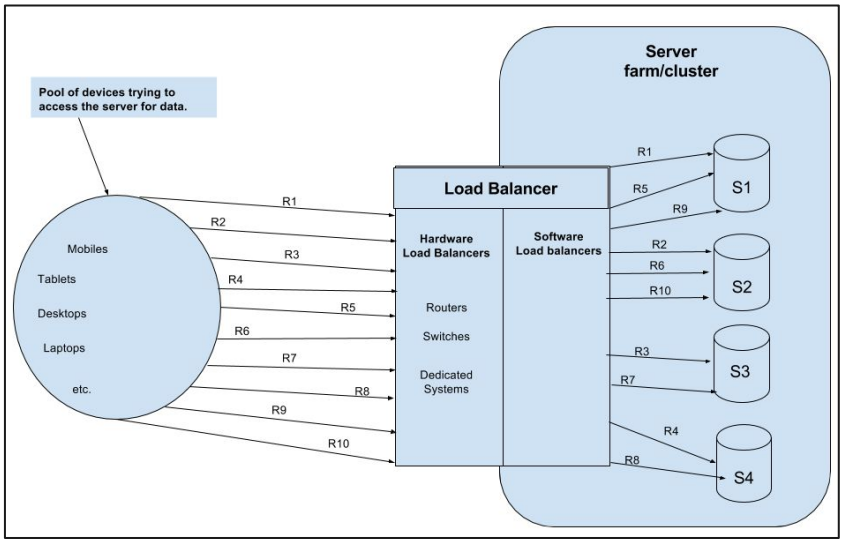
Software Load Balancer - Weighted Scheduling Algorithm

- Scheduling is done according to the weight assigned to the server.
- Different servers are assigned different weights.
- Weights are assigned to the servers by system admins after considering the hardware of the system server.



Software Load Balancer - Round Robin Scheduling

- Requests are assigned to server sequentially one after another.
- After sending request to the last server, the scheduling starts from the first server again.
- This algorithm is used when servers are of equal specifications.



Software Load Balancer - Least Connection First Scheduling

- Requests are assigned to the server which is currently serving the least number of requests.
- This type of scheduling is used when there are large number of persistent connections in an unevenly distributed traffic.

Examples of Hardware/Software Load Balancers

Software Load Balancers

1. HA Proxy - TCP Load Balancer
2. NGINX - HTTP Load Balancer
3. Mod_athena - HTTP Load Balancer(Apache)
4. Varnish - Reverse-proxy based
5. Balance - open-source TCP Load Balancer
6. LVS (Linux Virtual Server) - Layer 4

Hardware Load Balancers

1. Citrix Netscaler
2. F5-BIGIP
3. CISCO System Catalyst
4. Barracuda Load Balancer
5. CoyotePoint Load Balancer

References :

1. http://www.thegeekstuff.com/2016/01/load-balancer-intro/

2. https://www.nginx.com/resources/glossary/load-balancing/