**LOAD BALANCING**

***WHAT IS LOAD BALANCING:***

Load balancing is the process of distributing incoming network traffic or computing workload across multiple servers or resources to ensure optimal performance, reliability, and availability. It prevents any single server from becoming overwhelmed, reducing the risk of slow response time or system failures.

***TYPES OF LOAD BALANCING:***

\*Hardware Load Balancing – Uses physical devices to distribute traffic

\*Software Load Balancing – Uses software-based solutions running on standard servers

\*DNS Load Balancing – Uses the **Domain Name System** (DNS) to distribute traffic across multiple IP addresses.

***BENEFITS OF LOAD BALANCING:***

\*Improves scalability and efficiency

\*Enhances fault tolerance and redundancy

\*Optimizes resource utilization

\*Ensures high availability and reliability

Load balancers are commonly used in web applications, cloud computing, and enterprise IT infrastructures.

***HOW AWS LOAD BALANCING WORK:***

\*A user visit your site

\*The load balancing selects an available server

\*It forwards the request to that server

\*If a server fails, traffic is redirected