Distributed Web Infrastructure - Key Concepts

1. Distributed Web Infrastructure:

A distributed web infrastructure involves multiple interconnected servers collaborating to host a website, enhancing reliability and performance.

Design:
- 3 Servers
- 1 Web Server (Nginx)
- 1 Application Server
- 1 Load Balancer (HAproxy)
- 1 Set of Application Files (Code Base)
- 1 Database (MySQL)

2.Load Balancer:

Purpose:

- Manages traffic distribution to servers, preventing overload and enhancing user experience.

- Distribution Algorithm: Commonly Round Robin, ensuring fair sharing of requests among servers.

Active-Active vs. Active-Passive:

- Active-Active: Both load balancers actively handle traffic, providing higher availability.

- Active-Passive: One load balancer is active, while the other is on standby, used if the active one fails.

3.Primary-Replica (Master-Slave) Cluster:

Purpose:

- Enhances database reliability, provides read scalability.
- Primary Node**: Accepts write operations, synchronizes data to replicas.

- Replica Node: Handles read operations, asynchronously copies data from primary.

Application Interaction:

- Primary Node: Handles write-intensive tasks, updates data.
- Replica Node: Offers read operations, alleviating primary's load.

Issues with this Infrastructure:

1.Single Point of Failure (SPOF):

- A single failure can disrupt the entire infrastructure.
- Addressed with redundancy and failover mechanisms.

2.Security Issues:

- Lack of firewall exposes servers to potential attacks.
- Lack of HTTPS jeopardizes data integrity and user trust.

3.No Monitoring:

- Absence of monitoring tools means limited visibility into system health.

- Monitoring is crucial for proactive issue identification and resolution.