Scalability

Vertical Scalability

* Increase the resource (CPU,RAM) of a single machine to handle more load

Horizontal scalability

* Add more machines to the system to handle increasing load

Load Balancing

* Distribute the incoming load across multiple machines to ensure optimal resource utilization

Database sharding

* Partition a large database into smaller manageable shards to improve performance and scalability.

Caching

* Store frequently accessed data in a cache to reduce the load on the backend system

Asynchronous Processing

* Handle long-running tasks asynchronously to free up resources and improve system responsiveness.