Cloud computing & Scalability

↳ AWS → Auto Scaling, Load Balancer, RDS Replicas, Elasticache, NoSQL DB Auto Scaling

↳ Azure → Azure Auto Scale, SQL Elastic Pools, Cache, Redis, Load Balancing

↳ GCP (Google Cloud Platform) → Autoscaling, Cache, Load Balancing

↳ IBM cloud → Load Balancer, Auto Scaling, DB2 warehouse, cloudant

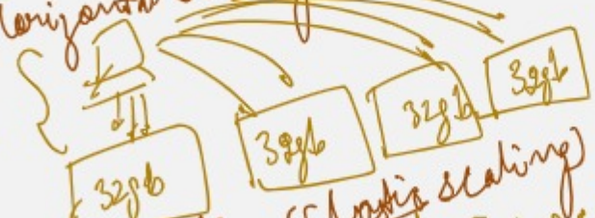
Scalable System
↳ Handling user traffic
↳ DB storage
↳ Enable businesses growth
↳ Fault tolerance
↳ user experience good
↳ HTML, CSS → Reactjs
↳ High volume transaction
↳ Infrastructure

req/mins
10,000 → Web
15,000 →
↳ 1 table → user data in DB
↳ 3 tables → DB ?
100gb - 200gb

Types of Scalability

① Vertical Scaling (single server)

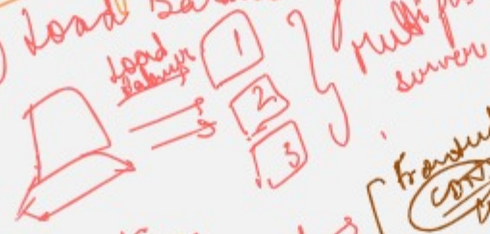
② Horizontal Scaling (multiple server/machine)



③ Autoscaling (Elastic Scaling)
↳ dynamic approach
↳ Horizontal & vertical scaling.
scaling when needed.

Scalability Techniques

① Load Balancing

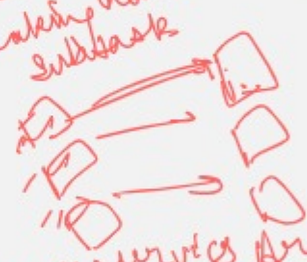


② Caching
↳ CDN → [Frontend caching, CDN, Static, CSS]
↳ Distributed caching
↳ DB caching
↳ data usage

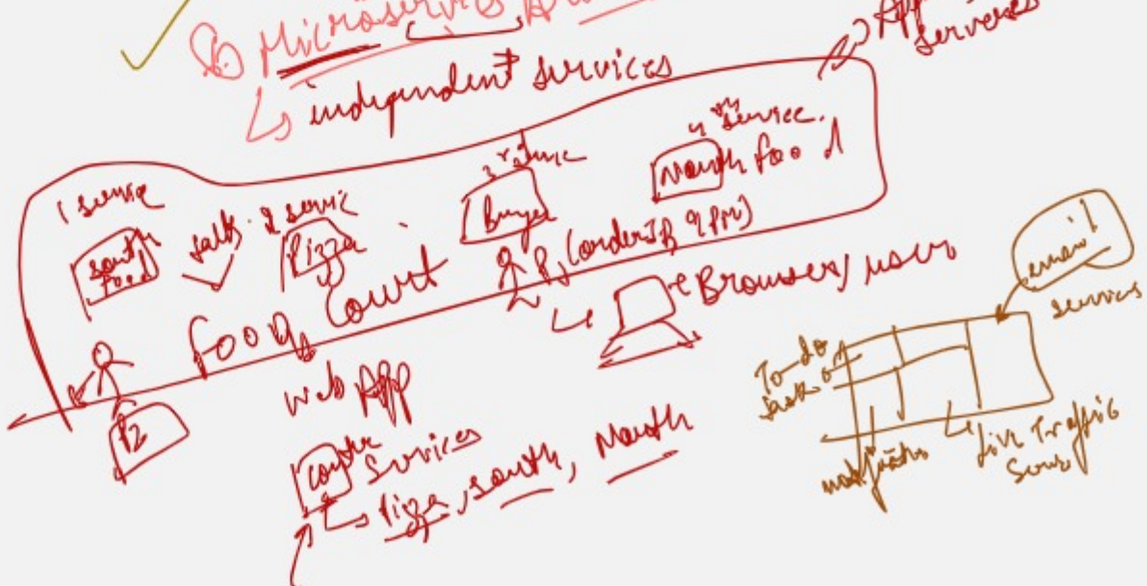
③ Database sharding

④ Asynchronous Processing
↳ Task which is not dependent

⑤ Distributed System
↳ breaking down complex task into subtasks



⑥ Microservices Architecture
↳ independent services → Application services

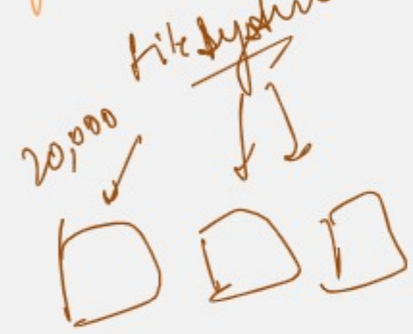


Scalability in Web Development
(Distributed network of servers)

① CDN
② Caching

Scalability for frontend & Backend

Frontend
↳ Performance optimization
↳ Responsive
↳ CDN
↳ caching
↳ Asynchronous operations

Backend
① Load Balancer
② Horizontal Scaling
③ Data optimization
↳ Indexing, db queries, caching, sharding.
④ Distributed file systems

File System