Contents

Scenario: Replicate Data Across Regions

Describe your approach to cross-region data replication, consistency, and failover.

Purpose

- Tests your understanding of distributed systems and data consistency
- Evaluates your ability to balance latency, availability, and durability
- Assesses your approach to conflict resolution and disaster recovery
- Checks your awareness of trade-offs in global architectures

---

Summary

Choose async replication for low-latency writes. Use quorum-based reads for stronger consistency. Resolve conflicts using timestamps or version vectors.

Replication Strategies

- **Asynchronous Replication:**
  - Primary region handles writes, replicates changes to secondary regions in the background
  - Pros: Low write latency, high availability
  - Cons: Risk of data loss on failover, eventual consistency
- **Synchronous Replication:**
  - Writes are committed only after all (or a quorum of) regions acknowledge
  - Pros: Stronger consistency, no data loss on failover
  - Cons: Higher write latency, lower availability during network partitions
- **Hybrid/Quorum-Based:**
  - Use a quorum of regions for reads/writes (e.g., majority must agree)
  - Balances consistency and availability

Consistency Models

- **Eventual Consistency:** All regions will converge to the same state, but reads may be stale
- **Strong Consistency:** Reads always return the latest committed write (requires synchronous/quorum replication)
- **Tunable Consistency:** Allow clients to choose consistency level per operation (e.g., Cassandra, DynamoDB)

CAP Theorem

The CAP theorem states that in a distributed data system, you can only guarantee two out of the following three properties in the presence of a network partition: - **Consistency (C):** Every read receives the most recent write or an error. - **Availability (A):** Every request receives a (non-error) response, regardless of the state of other nodes. - **Partition Tolerance (P):** The system continues to operate despite network partitions (communication breakdowns between nodes/regions).

Key clarification: - When there is no network partition, a system can provide both Consistency and Availability. - When a partition occurs (which is inevitable in distributed, cross-region systems), the system must choose to either: - Remain **consistent** (some requests may be denied to prevent stale reads), or - Remain **available** (all requests are served, but some may return stale data).

In cross-region replication: - Partition tolerance is required due to the nature of wide-area networks. - The trade-off between Consistency and Availability only arises during a partition event.

PACELC Theorem

PACELC extends CAP by considering system behavior not only during partitions, but also under normal operation: - **If there is a Partition (P):** choose between Availability (A) and Consistency (C) — just like CAP. - **Else (E), when the system is running normally:** choose between Latency (L) and Consistency (C).

In practice: - Systems like DynamoDB, Cassandra, and Cosmos DB allow tuning this trade-off. - You may choose lower latency (faster responses) at the cost of weaker consistency, or stronger consistency with higher latency, even when there is no partition.

Summary Table: | Theorem | Partition? | Trade-off | |———|————|———| | CAP | Yes | C vs. A | | PACELC | Yes | C vs. A | | PACELC | No | C vs. L |

Conflict Resolution

- **Last Write Wins (timestamp-based):** Simple, but may lose updates
- **Version Vectors:** Track causality and resolve conflicts more accurately
- **Custom Merge Logic:** For complex data types (e.g., CRDTs)

Failover & Disaster Recovery

- **Automated Failover:** Detect region failure and promote a new primary
- **Data Reconciliation:** Sync missed updates after recovery
- **Backup & Restore:** Regular cross-region backups for disaster recovery

---

Network & Latency Considerations

- Use geo-DNS or global load balancers to route users to the nearest healthy region
- Minimize cross-region traffic for latency-sensitive operations
- Compress and batch replication traffic

Security & Compliance

- Encrypt data in transit and at rest across all regions
- Ensure compliance with data residency and privacy laws (e.g., GDPR)

Metrics for Success

- Replication lag (seconds behind primary)
- Data consistency error rate
- Recovery time objective (RTO) and recovery point objective (RPO)
- User-perceived latency and availability