

Cache - Discussion Topics

Architecture & Design

1. When to use cache-aside vs read-through vs write-through?

- Trade-offs between consistency and performance
- Use cases for each pattern

2. How do you handle cache invalidation in a distributed system?

- Event-driven invalidation
- TTL strategies
- Cache coherence protocols

3. What are the risks of over-caching?

- Memory pressure
- Stale data issues
- Complexity in debugging

Real-World Scenarios

1. Design a caching strategy for a social media feed

- User-specific vs global cache
- Freshness requirements
- Cache warming strategies

2. How would you prevent cache stampede during high traffic?

- Request coalescing
- Probabilistic early expiration
- Lock-based approaches

Performance & Optimization

1. How do you measure cache effectiveness?

- Hit rate vs miss rate
- Latency improvements
- Cost-benefit analysis

2. What's your approach to cache sizing?

- Working set estimation
- Memory vs performance trade-offs
- Monitoring and adjustment strategies

Advanced Topics

1. Multi-level caching: When and how?

- L1 (local) vs L2 (distributed) cache
- Consistency challenges
- Eviction coordination

2. How do you handle cache in a multi-region deployment?

- Regional vs global cache
- Replication strategies
- Consistency models