ONE PAGE SUMMARY

IT & Software › System Design › Caching Strategies

Introduction to Caching Strategies

Boost Performance with Effective Caching Strategies

DEFINITION

> Caching strategies store frequently accessed data temporarily to improve system performance and reduce latency.

KEY CONCEPTS

Cache Types

Different types of caches include in-memory, disk, and distributed caches, each serving unique purposes.

Cache Hit vs. Miss

A cache hit occurs when data is found in the cache, while a miss indicates it must be fetched from the source.

Eviction Policies

Eviction policies like LRU (Least Recently Used) determine which data to remove when the cache is full.

TTL (Time to Live)

TTL defines how long data remains in the cache before it is considered stale and removed.

Cache Coherency

Cache coherency ensures that all caches in a system have the most recent data to avoid inconsistencies.

EXAMPLES

→ Web browsers cache images to load pages faster on subsequent visits.

→ A CDN caches website content closer to users for quicker access.

→ Database query results are cached to reduce load times for repeated queries.

MEMORY TIPS

★ Think 'Cache is Cash' for quick access to valuable data.

★ Use 'Hit or Miss' to remember the importance of cache performance.

★ Remember 'LRU: Last Recently Used' to recall how data is evicted.

COMMON MISTAKES

✗ Overlooking cache invalidation, leading to stale data.

✗ Not considering the size of the cache, causing frequent evictions.

✗ Ignoring the choice of eviction policy based on data access patterns.

QUICK RECAP

Caching strategies enhance system efficiency by temporarily storing data for quick access. Understanding cache types, hit/miss concepts, and eviction policies is crucial for effective implementation.