Cache hit / miss

🧠 Cache Hit কী?

👉 যখন requested data cache-এই পাওয়া যায়
→ DB-তে যেতে হয় না

🔁 Flow (Hit)

Client → Server → Cache ✅ → Response

✅ Benefits

- Very low latency
- DB load zero
- Fast user experience

📌 Interview line:

“A cache hit means the data is served directly from cache with minimal latency.”

🧠 Cache Miss কী?

👉 যখন cache-এ data নাই বা expire হয়ে গেছে
→ DB-তে যেতে হয়

🔁 Flow (Miss)

Client → Server → Cache ❌

↓

Database

↓

Cache store

↓

Response

❌ Problems

- Slower response
- DB load increases

📌 Interview line:

“On a cache miss, the system falls back to the database and populates the cache.”

⚖️ Hit vs Miss (Quick Compare)

Topic	Cache Hit	Cache Miss
Speed	⚡ Fast	🐢 Slow
DB Access	❌	✅
Latency	Low	High
Cost	Cheap	Expensive

📊 Cache Hit Ratio (VERY IMPORTANT)

📌 Formula

Hit Ratio = Cache Hits / Total Requests

Example:

900 hits / 1000 requests = 90% hit ratio

📌 FAANG rule of thumb:

Higher hit ratio = better performance

🎯 How to Improve Cache Hit?

- Cache frequently used data
- Proper TTL
- Right cache key design
- Avoid over-invalidation

⚠️ Cache Miss Types (Interview Bonus)

1. Cold Start Miss

- Cache just started
- No data yet

2. Capacity Miss

- Cache memory full
- Data evicted

3. Invalidation Miss

- Data expired or removed

🏆 FAANG Cheat Sentences 💎

- “We aim for a high cache hit ratio to reduce database load.”
- “Cache miss triggers a fallback to the database.”
- “Cache hit improves latency significantly.”

🔚 One-Line Memory Hook

Hit = Happy path
Miss = DB path