📸 **Instagram Feed Caching Strategy**

Instagram feed =

❌ simple list না

❌ fully dynamic

✅ **hybrid caching + ranking system**

1. Feed Types (First understand this)

Instagram-এ basically 3 ধরনের feed থাকে:

1. **Home Feed** (followers posts)
2. **Explore Feed** (recommended)
3. **Profile Feed** (user's own posts)

→ সবগুলোর caching strategy আলাদা

2. Core Challenges

Interviewers এগুলো expect করে 👇

- Feed is **highly personalized**

- Real-time updates (likes, comments)

- Ranking changes often

- Huge traffic (millions QPS)

3. High-Level Architecture

User

↓

API Gateway

↓

Feed Service

↓

Redis Cache ↔ Feed DB

↓

Response

4. Feed Generation Strategy (IMPORTANT)

🔥 Fan-out on Write (Primary)

When a user posts:

User A posts → push postId into followers' feed cache

Stored as:

feed:userId → Sorted Set(postId, timestamp)

📌 Redis **Sorted Set** is key here

❓ Why Fan-out on Write?

✅ Read is super fast
❌ Write expensive (celebrity problem)

📌 Interview line:

"Instagram primarily uses fan-out on write for fast feed reads."

5. Redis Data Structures Used

🔷 **Feed Cache**

Key: feed:user123

Value: Sorted Set (postId, score)

Score =

timestamp + ranking weight

◆ **Post Cache**

post:postId → {imageURL, caption, owner}

◆ **Like Count Cache**

likes:postId → counter

TTL based or write-through

6. Read Flow (Cache Hit Path)

User opens app

↓

Feed Service checks Redis

↓

Fetch top N postIds

↓

Fetch post data (cache)

↓

Return feed

✅ Multiple cache hits

❌ DB rarely touched

7. Cache Miss Handling

If feed cache missing:

Rebuild feed from DB

Store in Redis

Serve user

📌 Happens on:

- New user

- Cache eviction

- Cold start

8. Cache Invalidation Strategy (VERY IMPORTANT)

🔷 **Feed Cache**

- Time-based TTL (minutes)

- Partial update (push new posts)

- Background refresh

🔷 **Post Updates**

- Likes/comments → update counters only

- No full feed invalidation

📌 Interview line:

"We avoid full feed invalidation; we update deltas instead."

9. Celebrity Problem (🔥 Interview Favorite)

Problem:

- User with millions of followers

Solution:

- **Hybrid approach**

    o Normal users → Fan-out on write

    o Celebrities → Fan-out on read

📌 Line:

"For high-fanout users, we generate feed at read time."

10. Explore Feed Caching

- Not personalized per user

- Cached per region / cohort

- Longer TTL

11. CDN + Media Caching

- Images/videos → CDN

- Thumbnails → Edge cache

- Adaptive resolution

12. Cache Hit Ratio Goal

- Feed list → 95%+

- Post metadata → 90%+

- Media → 99%

🏆 FAANG Interview Killer Summary 💎

Use this 👇

"Instagram uses Redis sorted sets to cache user feeds, primarily fan-out on write for fast reads, with a hybrid strategy to handle celebrity users. Media is served via CDN while dynamic metadata is cached with controlled TTL."

🔚 One-Line Mental Model

Feed = cached list of postIds + dynamic counters