Friday, 19 Jan 2024 5:31 PM

# Last class learnings

1. Need of storage layer
2. What is Cache
3. Cache Eviction
4. Layers of caching

We will start at 4:05 PM

# Agenda

1. Case study to understand different form of cache
2. Real world case study to understand & solve thinking back

# Case Study!

- User
  | User id | name | email | ...
- Post
  | post id | user id | content | media_url | timestamp
- Friends
  | User1 id | Friend2 id | timestamp
- Likes
  | post id | user id
- Comments
  | Comment id | post id | user id | content | timestamp

Feeds: Collection of all posts made by friends of a user

world → 5B

select * from post where user_id in (select friendid from friends where user id=123)

## Feed Generation Query

# Data Storage Estimation

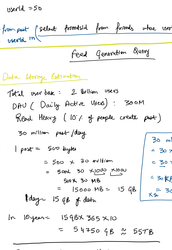
1. Total user base : 2 Billion users
2. DAU (Daily Active Users) : 300M
3. Read Heavy (10:1 of people create post)
4. 30 million post/day
5. 1 post = 500 bytes
   = 500 x 30 million
   = 500 x 10000 x 3000
   500 x 30 MB
   = 15000 MB = 15 GB

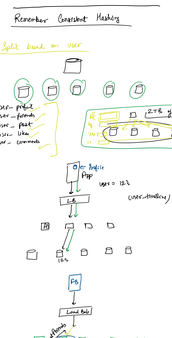
   1day = 15 GB of data.
6. In 10years = 15GB x 365 x 10
   = 54750 GB ≈ 55TB

30 million
= 30 x10^6
= 30 x 10000 x 3000
= 30 MB x 500
KB = 30MB

## Remember Consistent Hashing

# Split based on user





Db
Page
QPS = 1B
(user-timeline)
LB
Db
Db
Load Bal
Db
Db
Db