DESIGN A NEWS FEED SYSTEM

News feed is the constantly updating list of stories in the middle of your home-page. News feed includes status updates, photos, videos, links, app activity or likes from people, pages and groups that you follow on FB. Similar are → FB news feed, Instagram feed, twitter timeline.

Clarification questions

1) Mobile App (or) Web App?
   Both

2) What are the important features?
   A user can publish a post and see his friend's post on the news feed page.

3) Is the news feed sorted by reverse chronological order or any particular order such as topic scores? For instance, posts from close friends have higher scores.
   To keep things simple, let us assume it is sorted by reverse chronological order.

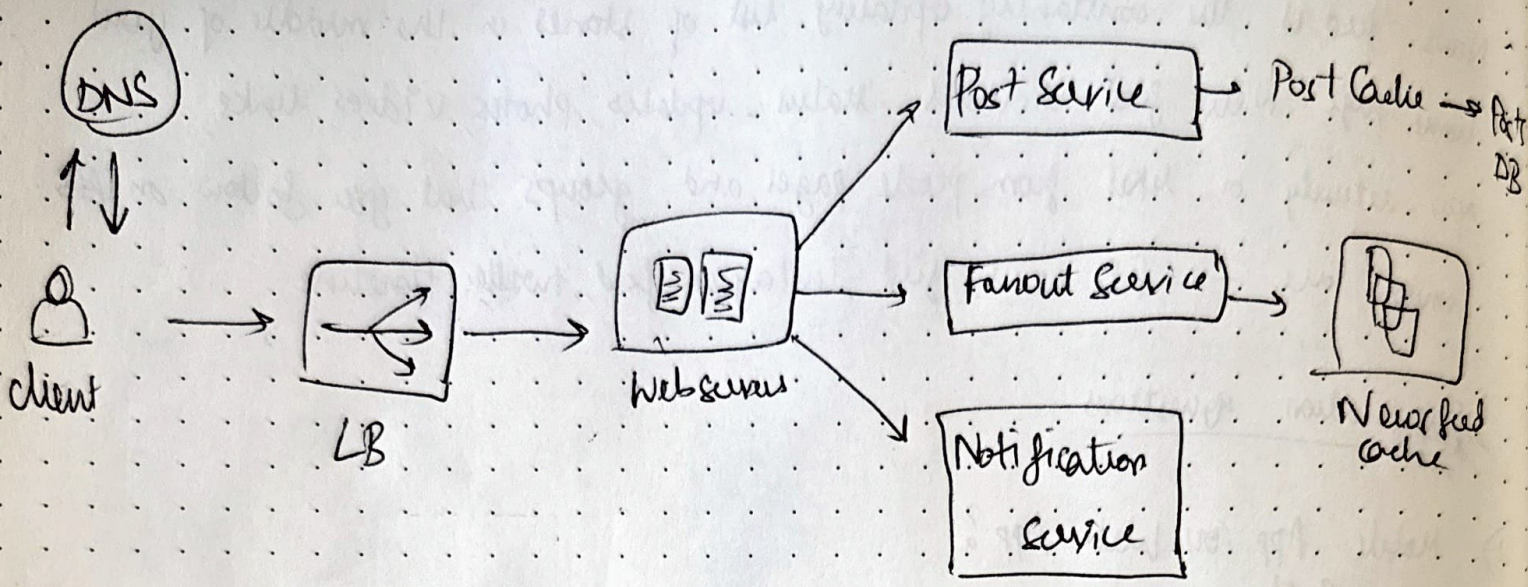
4) How many friends can a user have?
   5000

5) What is the traffic volume?
   10×10^6 DAU

6) Can the feed contain images, videos or just text?
   Yes, both images and videos, text

Dsi

High Level design



The Design is divided into 2 flows: -

(i) feed publishing

(ii) News feed building

→ When a user publishes a post, corresponding data is written into cache and database. A post is of all friend's news feed.

→ Built by aggregating of all friends post in reverse chronological order

Fanout service → takes an update and distribute to all friends.

FanOut on write (PUSH)

News feed cache is pre-populated during write

+ very fast
- If use has many friends, then fetching and generating news feed is slow.
- for inactive users, wastes a lot of resources

FanOut on READ (PULL)

News feed cache is populated during read

Time

+ memory efficient
+ for inactive users, read works better

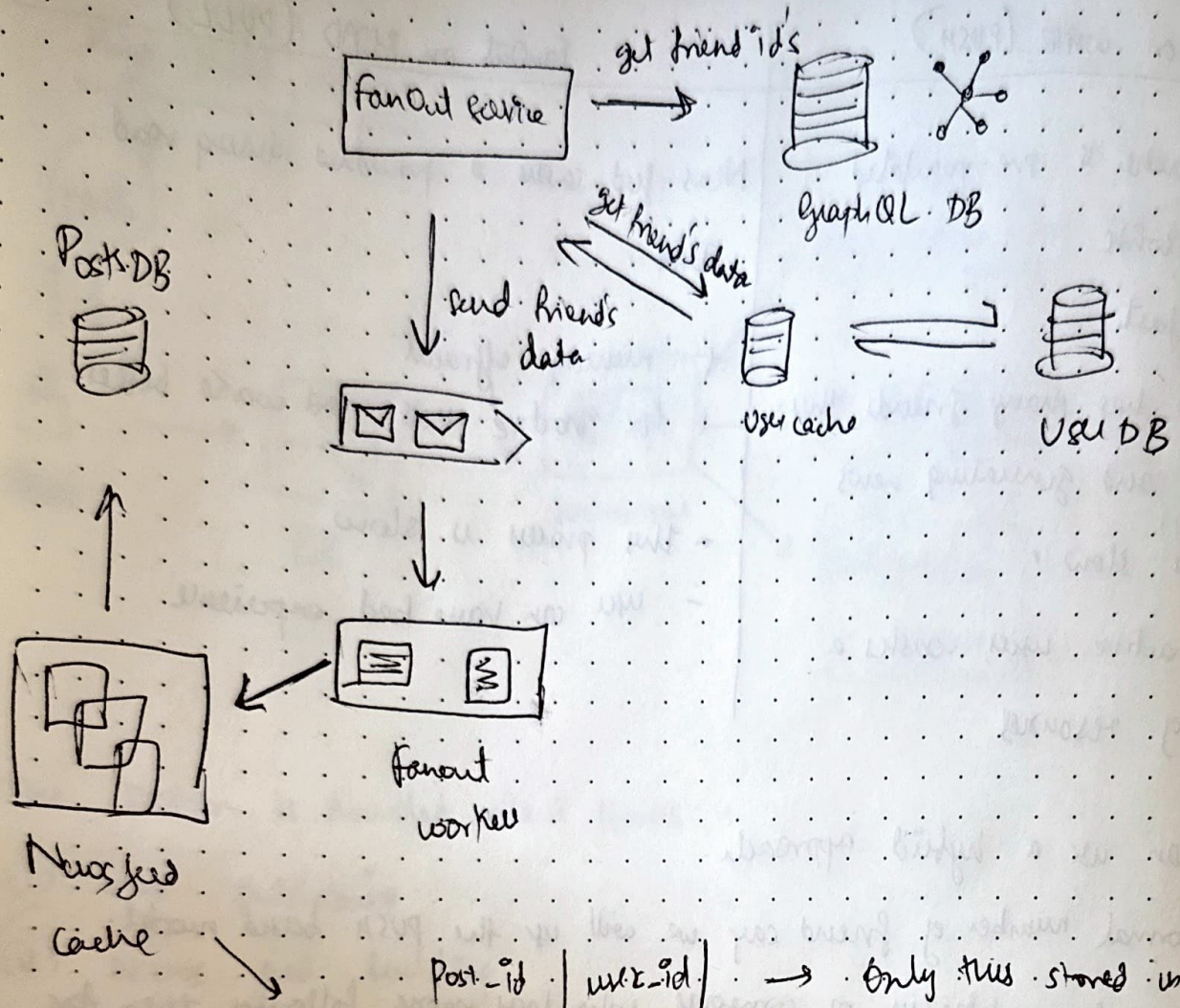
- this process is slow
- use can have bad experience

So, we can use a hybrid approach.

→ for normal number of friend say we will use the PUSH based model.

→ and say for celebraties or someone who has more followers then for them we use a pull based approach.

So we can modify based on our use case



post_id	user_id

→ Only this stored in the cache, coz if we store entire user info and post info and all then memory consumption can become very large.