**Stack overflow is a Question and Answer site**

**Functional Requirement.**

* Post question
  + User should be logged in order to post question( can be text or image)
  + Associate question with tags (so that will be help full for search)
  + Upvote or downvote the question if it the interested for you
  + Add comment to the questions or answer (need to logged in and cannot add comment on comment just on answer or question)
* Post Answer
  + User should be logged in order to post answer( can be text or image)
  + Multiple answer are allowed to a single question
  + Upvote or downvote the Answer if it the interested for you
  + Single user can write single answer if exiting user want to add the another answer he should edit the earlier answer .
  + Add comment to the answer (need to logged in and cannot add comment on comment and it will be singular)
* Question news feed (on homepage there are list of new question)
* User should be notified for the questions if it anyone has posted answer or comment for his question.
* Search for questions
* User profile page like anality page where user can go check what all the question I have asked on this site and its questions.
  + Rank answers basis of upvote or downvote
  + Views for the question and answers

**Non-functional requirement**

* Availablity
* Latency
* Scalability can be less than 2 seconds
* Eventual consistency is ok
* Durability - QandA should be persisted in the system unless deleted by owner or moderators of the question and answer.

**Capacity planning :**

Traffic on site : around 270 Million per month

Read per second : 270,000,000 / (86400 \* 30) = 104.1667 per second (approximate 110)

Write per second : 20% of read per second = (110 / 100) \* 20 = 22 (approximate 20 write per second for asking the question)

: 20% of read per second = (110 / 100) \* 20 = 22 (approximate 20 write per second for asking the answer)

Storage :

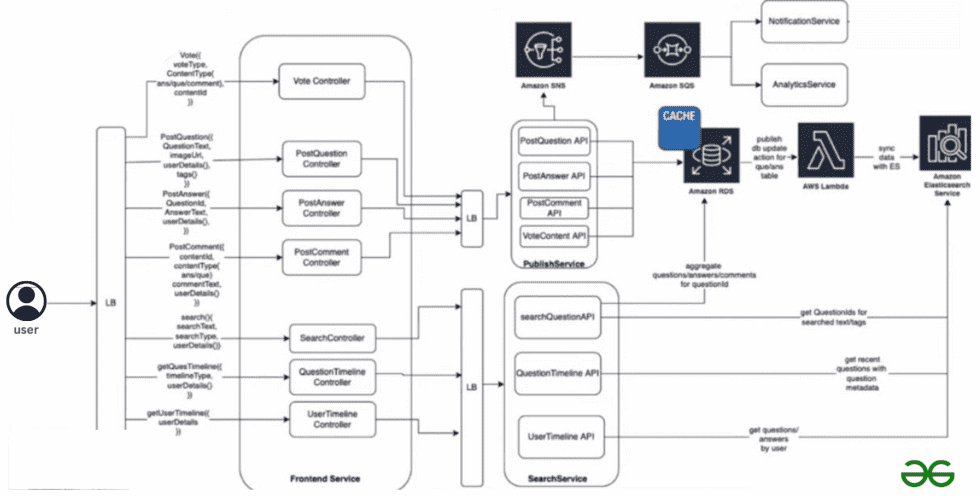
So let say if have 100 KB for each Q or A then 40 \* 100 = 4,000 KB per second

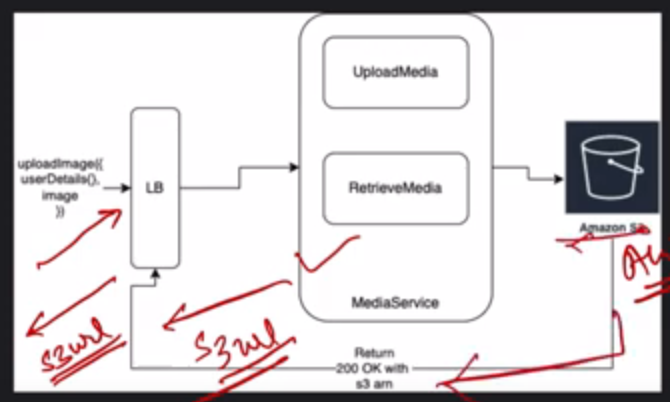
Which makes 4000 / 1000 = 4 MB per second

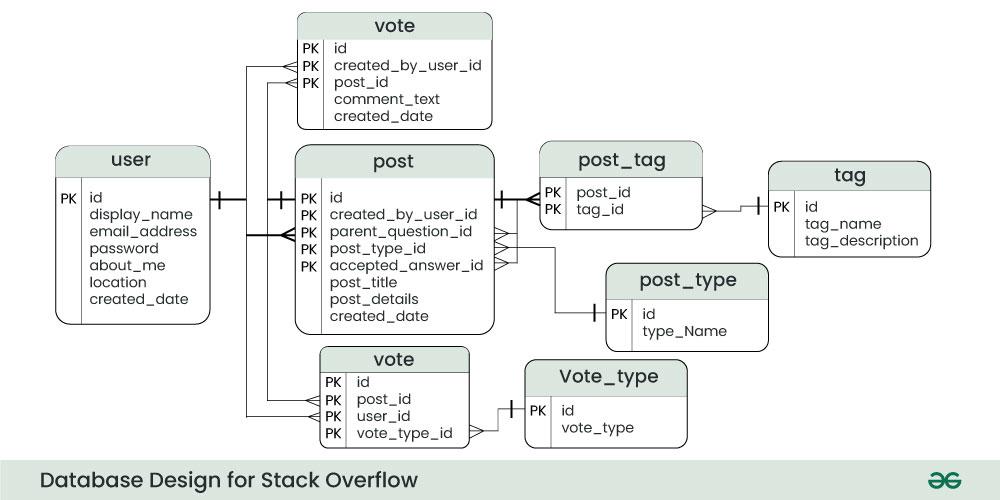
Which makes 4 \* (86400 \* 365) = 126144000 MB per year

(126144000 / 1000)/1000 = 126.144 TB per year (approximate 125 TB per year)

**High level design :**







**Points to be considered for enhancement or improvement**

* Elastic search schema
* Why RDS and not NoSQL
* RDS Table relationships
* Sharding in database
* Can any one of the APIs be asynchronous
* Requirement of admin user- as moderator
* Caching requirement and strategies
* Background scheduler job to update user badges