Overview:
Design a autocomplete system; or “design top k” or design top k most searched query

Requirement:
- Fast response time
- Relevant (should be relevant to search item)
- Sorted: results returned by the system must be sorted by popularity other than ranking model
- Scalable (handles high traffic volume)
- Highly available: the system should remain available and accessible when part of the system is offline, slows down or experiences unexpected network errors.

Back of Envelope Estimation
- Assume 10 million daily active users (DAU)
- An average person performs 10 searches per day
- 20 bytes of data per query string
  - Assume we use ASCII character encoding, 1 character = 1 byte
  - Assume a query contains 4 words, and each word contains 5 characters on average
  - That’s 4 x 5 = 20 bytes per query
- 在autocomplete的时候，每个新输入进去的数字都要有个单独的query，如果每个query是20byte的长度，那么就有20个query
- 20query * 10,000,000 DAU * 10 queries/day per user / (24 * 3600) ~= 48000 queries per second
- Assuming 20% of the daily queries are new, then 20% 10 million * 10 queries day = 0.4 GB of new data everyday

Data gathering Service
1. First step is to create the frequency table:
   Something like:

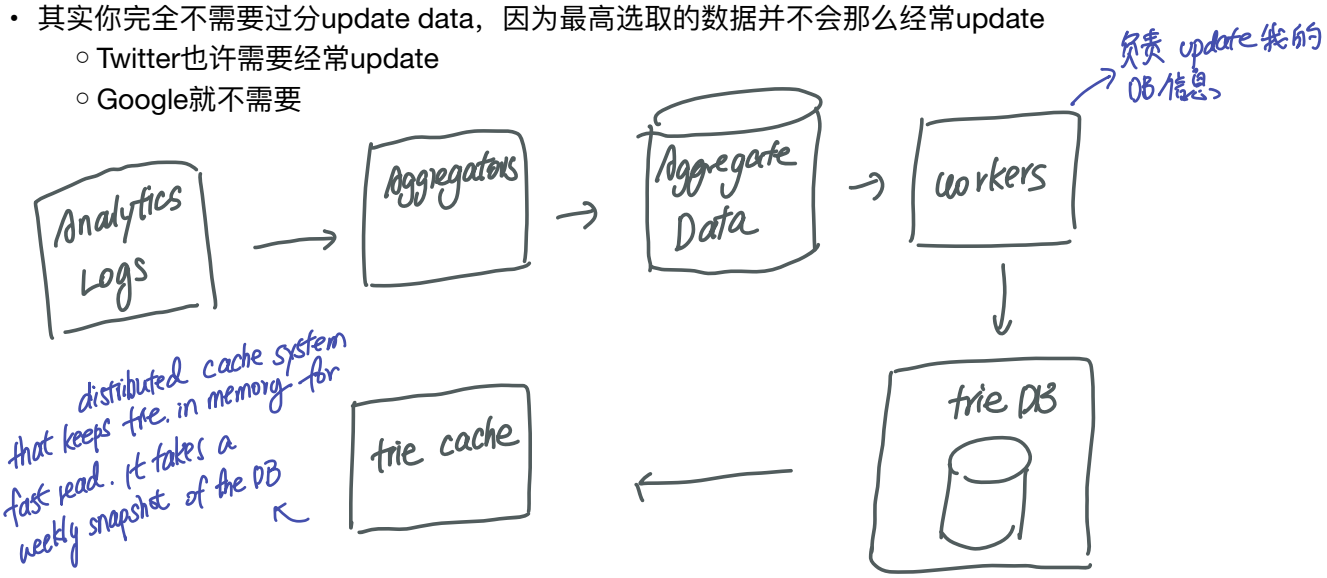
word	Frequency
twitch	5
twitter	34
twist	10
….	

2. Then, you can enter a SQL query to something like
SELECT * FROM frequency_table WHERE query LIKE `prefix%` ORDER BY frequency DESC LIMIT 5
This is an acceptable solution when the dataset is small

Deep dive
明显之前那个data gathering service不太行，dataset大了以后这个query能死人，所以deep dive来refine一下。
Tri-data structure:
- 之前的relational database不是最好的存储方法。这里用上一个trie会舒服很多（reTRIeval）这里除了一个基本的trie，还需要保存一个frequency和顺序上的问题
- 这里找一个prefix的children需要 O(p), 找subtree children需要O(c)，(c being ALL children).sort children找top k需要O(clogc)，所以是O(p) + O(c) + O(clogc)
  - 相比之前relational的找寻所有的elements O(n*p)，然后找出满足prefix的进行sorting，要快速一些
- 一些比较常见的optimize方法
  - 把p的长度limit到50，因为user不太经常输入过长的query
  - **Cache top searched queries at each node:** store top 5 queries (since they are most commonly used). This is trading space with memory efficiency
  - This way we reduce the search time to O(1) [limited length + cached top results]

Data Gathering Service

- 其实你完全不需要过分update data，因为最高选取的数据并不会那么经常update
  - Twitter也许需要经常update
  - Google就不需要



Analytics Log

- You can choose to have a analytics log that just act like a blob. It stores all the raw data about search queries. Logs are append-only and are not index. You then aggregate these analytics log **weekly, or biweekly.** Depending on the needs

Trie Store 的databae咋整?

有两个选择:

- **Document Store**: since a new trie is built weekly, we can periodically take a snapshot of it, serialized it, and store the serialized data in the database. Document stores like MongoDB are good fits for serialized data.
- **Key-value store:** a trie can be represented in a hash table form
  - Each prefix is mapped to a key in a hash table
  - Data on each trie node is mapped to a value in the hash table

Query

1. A search query is sent to load balancer
2. The LB routes the request to API servers
3. API servers get trie data from Trie Cache and construct autocomplete suggestions for the client
4. When data not in cache, replenish the data back to the cache

Query Service:

1. **AJAX request:** browser usually sends AJAX request to fetch autocomplete results.The main benefit of AJAX is that sending/receiving a request/response does not refresh the whole web page.
2. **Browser cache:** the browser can cache part of the service too. (Google caches the results in the browser for 1 hour)
   A. Private in cache-control means results ar intended for a single user and must not be cached a shared cache. Max-age=3600 means the cache is valid for 3600 seconds, aka, an hour.

Trie operation

1. **Create:** done by the workers from aggregating data
2. **Update:**
   A. Update it weekly
   B. Update the individual trie node directly (OK with small trie)

3. **Delete**

Scaling:

1. One way to scale is to parse the input based on the starting character of the word
    A. Say ‘a’ - ‘m’, ‘m’ - ‘z’
    B. We can then split them up to 26 servers
    C. We can then even shard more later on