

rh phone + vo

applying for sd2 - toronto - 4yoe

phone screen - referral count

VO

architecture - job scheduler

project deep dive

coding - leetcode.com

Requirements for job scheduler listed below. Interviewer copy pasted this into the miro board.

1. Should be able to create jobs
2. Should be able to schedule and run jobs
3. Should be able to report failures and successes
4. Should be reliable and have strong guarantees about its job runs
5. Should be able to view logs and status of running jobs, as well as previously finished jobs
6. Should be able to handle when a job takes longer to run than expected (SLA)

focus on fault tolerance and schema. i got hung up on the last requirement where my schema didn't really support the last req for example if a job is to be run once every 5 minutes and the first one is taking too long, the first job should not impact the next job running at the next scheduled time.

coding

fractional inventory

leetcode.com

<https://leetcode.com/discuss/interview-question/5461047/Robinhood-Phonescreen-or-L4-or-July-2024>

pretty easy. finished it in 20 minutes and chatted the rest of the time.

pdd

mix of behavioural and technical questions about [system design](#). i think conversation was ok but im not sure how well it went.

other potential questions

architecture

design an API that can view real-time stock prices and historical prices

- get live price of a ticker

- get historical price with a date range of a tickergiven 2 hoses of live data streamed directly from NASDAQ

limit order sell

coding

offset ordering - leetcode.com (required to do in $O(1)$ space)

margin call - leetcode.com

match orders/house and street - leetcode.com

candlestick and service dag - leetcode.com

portfolio optimization - leetcode.com

referral count- leetcode.com

please give rice! im very poor :(

补充内容 (2024-08-06 04:04 +08:00):

additional SD: limit order sd

System design:

- job scheduler
 - 1. Should be able to create jobs
 - 2. Should be able to schedule and run jobs
 - 3. Should be able to report failures and successes
 - 4. Should be reliable and have strong guarantees about its job runs
 - 5. Should be able to view logs and status of running jobs, as well as previously finished jobs
 - 6. Should be able to handle when a job takes longer to run than expected (SLA)

-
- focus on fault tolerance and schema. i got hung up on the last requirement where my schema didn't really support the last req for example if a job is to be run once every 5 minutes and the first one is taking too long, the first job should not impact the next job running at the next scheduled time.
-
-
- design an API that can view real-time stock prices and historical prices
 - - get live price of a ticker
 - - get historical price with a date range of a tickergiven 2 hoses of live data streamed directly from NASDAQ
-
-
- limit order
 - client通过robinhood来place limited order, robinhood 通过call第三方market的API, 得到stock的具体信息, 最后完成交易。面试官没直接给出具体的需求, 我大概通过交流概括了一下:
 -
 - 1. support placing limited orders: during market hours, expire at each trading day
 - 2. support accounting logic: to ensure the order is valid and consistent
 - 3. support multiple/concurrent orders
 - 关于accounting logic需求, 举个例子, 一个用户账户里只有100块, 但是他打算买120块的股票, 那这种order就是无效的。我当时说的是因为系统只要支持limited order而不需要实时的order, 所以可以存下来所有的limited order, 然后after hour时候跑MR去校验order (稍微往lambda的batch layer套了套)
-
-
- Compute ads campaign ROI
 - 系统设计的面试
 - 问题是假设RH在很多外部的广告平台投放广告, 如何设计一个数据pipeline 日常的统计每个平台的CPA(cost per action).
 - 假设有一下几个数据源
 - - 每个平台会定时给RH发送看到他们广告的用户设备ID, 时间戳。
 - - RH自己统计每天注册自己网站的用户ID, 设备ID, 和时间戳
 - - 每天结束的时候RH可以向不同的广告平台发送一个请求, 拿到每天广告花费数额
 -
 - 输出
 - - 统计处每个平台每天的CPA(产生一个注册用户所花费的钱)
 -
 - 需要设计整个数据系统
 - - 怎么拿到数据, 怎么存储实时数据(卡夫卡)
 - - 拿到后怎么整理, 怎么存储, 怎么备份
 - - 从数据上有什么问题(回答:设备ID可能有多个, 或者一个设备ID匹配到多个平台)
 - - 怎么防止failure
 - - 怎么检测 在哪里加检测
 - - 如果把每天的job变成realtime job 怎么弄?