



九章算法 帮助更多中国人找到好工作

扫描二维码，获取“简历”“冷冻期”“薪资”等求职必备干货

九章算法，专业的IT 求职面试培训。团队成员均为硅谷和国内顶尖IT企业工程师。目前开设课程有《九章算法班》《系统设计班》《Java入门与基础算法班》《算法强化班》《Android 项目实战班》《Big Data 项目实战班》。

Amazon 系统设计题目汇总

Amazon 系统设计题目汇总

- 1.在SQL里面一个产品有多个图片多个价格的话怎么设计数据库。后台提取数值render到页面上得时候，class怎么设计，服务器怎么安排,怎样给suggest product。
- 2.对于key,value pairs，在给定的文件系统中实现 put, get, delete 的方法。其中key比较小，全部key可以放在内存中，value有的会比较大。已知一个文件系统，可以create files, delete files, sequentially scan file content, read file content randomly, append file content..
- 3.已知一个函数，输入用户ID，可以返回该用户的所有友好（degree 1 friends），按友好ID从小到大排序。要求实现函数来输出返回一个用户的所有好友的好友(degree 2 friends), 以及 degree 3 friends。
- 4.社交网站上的文章转发，如何设计系统可以得到实时的转发量榜单和weekly digest，要求数据库的设计，有人转发一个文章时request是什么样的，如何快速得到实时的转发量榜单，如何得到weekly digest等。
- 5.设计一个 Netflix/Spotify，如何限制同一个用户多处登录不能同时看一个资源，如何实现根据用户的网速调整清晰度。
- 6.如何设计一个简单的协同编辑。协同编辑就是能够让多人同时编辑同一份文档，用版本控制的一个好处就是，可以方便的多人共同去做一个项目，但有时大家会对同一个文件的同一行进行了修改，这时merge代码的时候，就需要你去手工去解决冲突，而协同编辑同样也面临冲突的问题，这时候就是程序实时的自动去解决冲突了

-
- 7.设计一个service, 任何人call 这个service , return 一个unique id, 不能和其他人重复
- 8.求一个论坛的在线人数, 假设有一个论坛, 其注册ID有两亿个, 每个ID从登陆到退出会向一个日志文件中记下登陆时间和退出时间, 要求写一个算法统计一天中论坛的用户在线分布, 取样粒度为秒。
- 9.设计top 10 most page view product, 有dvd, 显示page view最多的10个dvd, 有book, 显示page view最多的10个book。traffic访问量很高怎么办。如果要实现上个月page view最多的10个dvd怎么办
- 10.在作 consistent hash 的时候通常一台机器是对应好几个 virtual nodes. 然后把这些 nodes 随机撒在环上. 在作 replicas 时, 通常也是拿环上顺时针方向的下三个 nodes (given RF=3) 来当 replicas., 怎么避免 virtual nodes 被撒在环上的连续空间.譬如一台机器有 5 个 virtual nodes. $A \Rightarrow A1, A2, \dots, A5$, 如果他们刚好被放在环上的连续空间, 作 replicas 的时候, 其实都作在同一台机器上, 一台挂了就所有 replicas 一起挂。
- 11.We've 1 book left in the inventory. and two people are trying to get the same book (say person x and person y). Person x has added book to the cart and about to make payment and person y has also added book to the cart. How would you solve this concurrency problem ?
- 12.Let's say you have a simple function (fibonacci/factorial) that you need to run constantly. The largest number that you will receive as input will be 1,000. How can you improve the performance of this function call?
- 13.How would you design a social network and find or keep track of someone's oldest friend in a social network? Oldest friend means the friend that you have added for the longest time period.
- 14.Design a conference room booking system for a company which can have offices in multiple cities, each city can have multiple buildings, each building can have multiple floors, each floor can have multiple rooms. Each room can have features like capacity, video conferencing available, etc.
- 15.You have a website which has 5 reads on a database when it is loaded. Each read takes 7 seconds for a total of 35 seconds, discuss how you could improve the performance?
- 16.Given that an external service gives a list of credit cards that have become fraud, design a fraud management system for a shopping website for bookings with fraud credit cards
- 17.You have a cluster with 100 machines that need time to be synced. The central time server can only handle 10 requests at a time. How will you set this up?
- 18.Design a system that processes an infinite stream of data: each record comes as a tuple(url, html content) ,extract and store the occurrences of: urls, hosts, top-level-

domains, in/out links of the page .We have one machine that has enough disk space but limited memory. What data structures as building blocks would you use?

19.Imagine that there are 7 servers running in parallel. What happens when you need to expand to 20 live? What are issues? What could you do to fix this issue in the future?

20.Design a logging system. The system contains multiple application servers which are logging the information to file system. In this scenario, we want to check and alarm in case an exception is thrown in any of the servers. We want a system that checks for appearance of specific words, "Error", "Exception", "Disk Full" etc. in the logs of any of the servers. How would you design this system? What if we want to scale the system in future?

九章算法总结了一些系统面试的建议:

1.系统设计面试内容

并发性(concurrency)。你知道线程(threads)、死锁 (deadlock) 和starvation吗? 你知道如何并行化算法吗? 你了解一致性(consistency)和连贯性(coherence)吗? 你大概了解IPC和TCP/IP吗? 你知道吞吐量 (throughput) 和延迟 (latency) 的区别吗?

现实表现(real-word performance)。你应当熟悉你电脑的速度和性能, 包括RAM、硬盘、SSD以及你的网络状况。

估计(estimation)。估计在帮助你缩小可能性解决方案的范围时起到了重要的作用。这样, 你就只需写少数几个原型或微基准。

可用性和可靠性(availability and reliability)。你是否考虑过系统什么时候会出现bug无法运行吗 (特别是在分散式的环境中)? 你知道如何设计一个系统以应当网络故障吗? 你了解持久性吗? 切记, 我们并不是要寻找一个熟悉以上所有的问题的“全才”。我们想衡量的是你的熟练程度。我们只需要你对系统设计方面有一定的基础, 并且知道什么时候应该寻求专家的帮助。

2.系统面试评分标准

优秀的面试者

类别: 极少见

是否聘用建议: 一旦发现, 马上聘用

一个优秀的面试者, 在面试系统设计题时, 只需通过自主地向面试官询问并厘清系统设计需求, 就可以相对独立地完成面试题, 并不需要面试官的太多帮助。通常, 这类优秀的面试者, 他们这么答系统设计题: 先描述一个大的设计框架, 然后给出更多的

details。这类优秀的面试者，给出的系统设计常常是非常make sense的，以致于面试官都很难“找茬”。当面试官要求该面试者解释部分的设计时，该面试者可以非常清楚、准确地描述这个设计部分是如何works的。如果面试官要求面试者拓展整个设计(expand the design)，面试者也能非常清楚地修正/或重新设计他先前的设计。

合格的面试者

类型：常见

是否聘用建议：如果其他的面试轮表现得也不错，就聘用

一个合格的面试者会尝试跟面试官沟通并厘清系统设计需求，然后从一个high level的层面开始系统设计。这类求职者在面试过程中或多或少会犯一些错误。但是，一旦面试官指出他系统设计中的错误时，这类面试者可以非常快地修正他的设计。当面试官要求他解释其中的某部分设计时，这类求职者也通常可以不费力地描述和解释自己的设计。这类求职者，对于系统设计有一个high level的想法，同时也可以针对新需求进行改进。这类求职者提供的系统设计或许看起来有些蹩脚，但总体来说，应该是一个看起来在现实中可work的产品。

比较差的面试者

类型：常见

是否聘用建议：如果是一个完全没接触过系统设计的面试者，比如New Grads，可以考虑hire

一个比较差的面试者，在拿到系统设计题目时，会直接开始设计，而忽略了与面试官沟通并澄清设计需求的过程。所以他们在设计时，通常会有很多错误。当面试官指出其中的设计错误时，比较差的求职者，他们往往没有能力进行修正。当面试官要求面试者修正或拓展整个设计时，比较差的求职者也没办法去改进原先的设计，或者缺乏相关的知识。比较差的面试者，他们不太能给出一个完善的设计，而常常是非常蹩脚的。

糟糕的面试者

类型：极少见

是否聘用建议：不聘用

一个糟糕的面试者在拿到系统设计题时，通常不知道如何入手。他们不能自己把系统设计问题break down，而需要面试官来帮忙。当面试官让他们只设计某一小部分时，他们可以做出很好的设计，但是无法全面、完整地考虑整个project的设计(think the project as a whole)。这类面试者需要面试官给很多很多的hints才能给出一个设计，或者他们给出的设计，在现实中完全就是不可行的。

3.系统设计面试准备

模拟系统设计面试。邀请一个工程师帮你模拟面试。让他提出一个系统设计问题，如果正好是他正在做的项目那就再好不过了。不要把它当成是一个面试，而是放轻松地去思考问题，并提出你能想到的最佳解决方案。

在实际的系统中去实践。你可以在既有的OSS中去练习，也可以与朋友合作搭建一个系统。对于课堂中的系统设计作业，不再把它仅仅当成一个学术训练，而是把它当成实际问题，思考系统设计过程中的架构和博弈。正如我们生活中遇到的大多事情一样，只有做了才知道其中会遇到什么问题，从而真正学到东西。

深挖开源系统的运行特点。例如，你可以看看levelDB。这是一个干净、小、且编写良好的系统。你可以读读执行命令，了解它是如何在硬盘中存储数据的，如何将数据压缩成不同的层？你也可以多多反思一下的博弈问题：哪种数据和大小是最优的？什么情况下会降低读写速度？（提示：比较一下随机写和顺序写）

多了解一下系统中数据库和操作系统是如何运行的。这些技术并不只是你口袋中的工具，它们往往会在你设计系统的时候给你带来启发。如果你经常像DB或OS一样思考它们如何处理各自的问题，你也会把这些思考方式应用到其它的系统设计中。

课程推荐

九章算法《系统设计班》主要针对北美互联网公司的系统设计面试内容和要求而设计，包括如何解决系统的 Concurrency, Consistency, Reliability等问题。为缺乏工业界经验的面试者而准备，从0到1，无需任何先修基础。

通过实战面试真题，讲解实际面试中会遇到的系统设计知识和面试技巧，包括：**数据系统设计、爬虫系统与搜索建议系统、Tiny URL 设计、Google 三驾马车(Big Table / Google File System / MapReduce)、分布式系统设计、WhatsApp聊天系统设计、实时位置信息系统等。**

本课程不是项目实战课程，主要是系统实践基础知识的讲解和巩固，面试中的系统设计面试，主要是通过面试官，理解面试官的设计需求，进行口头上的设计和优化为主。