



超人学院 hadoop 面试葵花宝典(V1.0)

www.crxy.cn

QQ 群：204069925

为了让大家更有针对性的学习和复习,对 hadoop 面试题有个大概的了解,超人学院特将学员面试过程中遇到的面试题汇集成篇。在此,非常感谢大家对超人学院工作的鼎力支持与配合,尤其要感谢(包括但不限于)飞哥、然月枕流君、北京-南桑、彩虹伴相思雨、Clouds、小萝卜、北京-大数、象夫、随心、mo•mo•ring、aboutyun、happy、闪客、找自己、炎帝初始化.....截至 2014 年 7 月 25 日为止,本次共收集了 336 道 hadoop 相关的面试题。随着面试学员数的增加,我们将不定期更新面试题库,欢迎大家持续关注超人学院的官网 www.crxy.cn 和 QQ 交流群 204069925,确保第一时间获取免费的公开课信息和其他学习资料。

更多学员面试经验请关注超人学院官网：<http://www.crxy.cn/>。

- 1、10 年工作经验罗同学由月薪 18k 飙升至 45k: <http://www.crxy.cn/detail/news/12> ;
- 2、14 届应届本科生入职中国航天集团,基本年薪 20w,年终奖 10w :
<http://www.crxy.cn/detail/jobinfo/31>;
- 3、入职腾讯学员告诉你 hadoop 学习方法 : <http://www.crxy.cn/detail/jobinfo/10> ;
- 4、揭秘大专生月薪 6.5k 翻番至 13k : <http://www.crxy.cn/detail/jobinfo/8> ;
- 5、美女研究生 hadoop 工作经验分享 : <http://www.crxy.cn/detail/jobinfo/6>。

一、来自****提供的面试题 14 道:

HADOOP 工程师面试题

1. 简要描述如何安装配置一个 apache 开源版 hadoop，只描述即可，无需列出完整步骤，能列出步骤更好。

2. 请列出正常工作的 Hadoop 集群中 Hadoop 都分别需要启动哪些进程，他们的作用分别是什么，尽可能写的全面些。

3. 启动 Hadoop 时报如下错误，如何解决

```
ERROR                                org.apache.hadoop.hdfs.server.namenode.NameNode:
org.apache.hadoop.hdfs.server.common.InconsistentFSStateException:      Directory
/tmp/hadoop-root/dfs/name is in an inconsistent state: storage directory does not exist or is not
accessible.
    at
    org.apache.hadoop.hdfs.server.namenode.FSImage.recoverTransitionRead(FSImage.java:303)
    at
    org.apache.hadoop.hdfs.server.namenode.FSDirectory.loadFSImage(FSDirectory.java:100)
    at
    org.apache.hadoop.hdfs.server.namenode.FSNamesystem.initialize(FSNamesystem.java:388)
    at
    org.apache.hadoop.hdfs.server.namenode.FSNamesystem.<init>(FSNamesystem.java:362)
    at
    org.apache.hadoop.hdfs.server.namenode.NameNode.initialize(NameNode.java:276)
    at org.apache.hadoop.hdfs.server.namenode.NameNode.<init>(NameNode.java:496)
    at
    org.apache.hadoop.hdfs.server.namenode.NameNode.createNameNode(NameNode.java:1279)
    at org.apache.hadoop.hdfs.server.namenode.NameNode.main(NameNode.java:1288)
```

4. 请写出以下执行命令

- 1) 杀死一个 job
- 2) 删除 hdfs 上的/tmp/aaa 目录
- 3) 加入一个新的存储节点和删除一个计算节点需要刷新集群状态命令

5. 请列出你所知道的 hadoop 调度器，并简要说明其工作方法。

6. 请列出在你以前的工作中所使用过的开发 map/reduce 的语言
7. 当前日志采样格式为
a,b,c,d
b,b,f,e
a,a,c,f
请用你最熟悉的语言编写一个 map/reduce 程序，计算第四列每个元素出现的个数。
8. 你认为用 Java, Streaming, pipe 方式开发 map/reduce，各有那些优缺点。
9. Hive 有哪些方式保存元数据的，各有那些特点。
10. 请简述 hadoop 怎样实现二级排序。
11. 简述 hadoop 实现 join 的几种方法。
12. 请用 java 实现非递归二分查找。
13. 请简述 MapReduce 中 combiner, partition 作用
14. 某个目录下有两个文件 a.txt 和 b.txt，文件格式为(ip username)，例如：
- a.txt
- 127.0.0.1 zhangsan
127.0.0.1 wangxiaoer
127.0.0.2 lisi
127.0.0.3 wangwu
- b.txt
- 127.0.0.4 lixiaolu
127.0.0.1 lisi
- 每个文件至少有 100 万行，请使用 linux 命令行完成如下工作：
- 1) 两个文件各自的 ip 数，以及总 ip 数
- 2) 出现在 b.txt 而没有出现在 a.txt 的 ip
- 3) 每个 username 出现的次数 以及 每个 username 对应的 ip 数



第一题：1.创建 hadoop 帐户。

2.setup.改 IP。

3.安装 java，并修改/etc/profile 文件，配置 java 的环境变量。

4.修改 Host 文件域名。

5.安装 SSH，配置无密钥通信。

6.解压 hadoop。

7.配置 conf 文件下 hadoop-env.sh、core-site.sh、mapre-site.sh、hdfs-site.sh。

8.配置 hadoop 的环境变量。

9.Hadoop namenode -format

10.Start-all

第二题：namenode：管理集群，并记录 datanode 文件信息。

Secondname:可以做冷备，对一定范围内数据做快照性备份。

Datanode:存储数据

Jobtracker :管理任务，并将任务分配给 tasktracker。

Tasktracker:任务执行方。

第三题：可能的原因：1.hdfs 没有启动成功，通过查看 jps 确认下。

2.确认文件是否存在。

第四题：hadoop job -list 拿到 job-id ,hadoop job -kill job-id

Hadoop fs -rmr /tmp/aaa

加新节点时：

Hadoop-daemon.sh start datanode



Hadoop-daemon.sh start tasktracker

删除时：

Hadoop mradmin -refreshnodes

Hadoop dfsadmin -refreshnodes

第五题：

Fifo scheduler :默认，先进先出的原则

Capacity scheduler :计算能力调度器，选择占用最小、优先级高的先执行，依此类推。

Fair scheduler:公平调度，所有的job 具有相同的资源。

第六题：java、python、hive

第七题：wordcount。。

第八题：就用过java 和 hiveQL。

Java 写 mapreduce 可以实现复杂的逻辑，如果需求简单，则显得繁琐。

HiveQL 基本都是针对 hive 中的表数据进行编写，但对复杂的逻辑很难进行实现。写起来简单。

第九题：三种：内存数据库 derby，挺小，不常用。

本地 mysql。。常用

远程端 mysql。。不常用

上网上找了下专业名称：single user mode..multi user mode...remote user mode

第十题：在源码中有个例子。。不过我没看。。

第十一题：貌似好几种来着，像 mapjoin ..reducejon..还有其它的来着吧。。可以去网上查一下，我常用的就是 mapjoin，可以将小表的数据加载到内存中使用，然后匹配的大表的



数据，加快效率。

第十二题：用 java。。我的第一思路就是排序后从中间查询呗，for 循环的事。。

第十三题：

combiner :实现的功能跟 reduce 差不多，接收 map 的值，经过计算后给 reduce，它的 key,value 类型要跟 reduce 完全一样，当 reduce 业务复杂时可以用，不过它貌似只是操作本机的数据。。

Partition：将输出的结果分别保存在不同的文件中。。

第十四题：不会 shell。。。

二、来自彩虹伴相思雨提供的面试题 31 道：

笔试题：



Java 笔试题 (30 分钟)

题目 1: 字符串统计

现有字符串需要处理，请按程序压缩字符的个数，例如

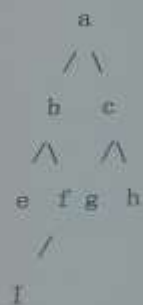
字符串: aaaaabbbbccccdddddlaa

输出结果: 5a4b1a5c4d3a

请编写 Java 程序实现字符串的压缩功能，程序要求可以处理类似字符串。

题目 2: 二叉树遍历 (后序)

根据如下二叉树，编写 Java 程序，采用二叉树后序遍历算法实现，并输出结果。



题目 3: 时间段合并

现有多个用户，任意的时间段有上网行为，需要将相同用户的上网时间段合并，例如：

用户 1, 上网时间段一: 9 点 30 到 9:50

用户 2, 上网时间段一: 9 点 10 到 9 点 20

用户 1, 上网时间段二: 9 点 40 到 10:00

用户 2, 上网时间段二: 9 点 00 到 9:15

上网 1, 上网时间段三: 10:10 到 10:20

...

输出结果:

用户 1: 9 点 30 到 10:00 和 10:10 到 10:20

用户2: 9点00到9:20

请编写 Java 程序实现相同用户的上网时间段合并，程序要求可以处理多用户任意情况时间段合并，不限于示例数据，并且需要考虑亿级时间段的处理效率及内存开销。

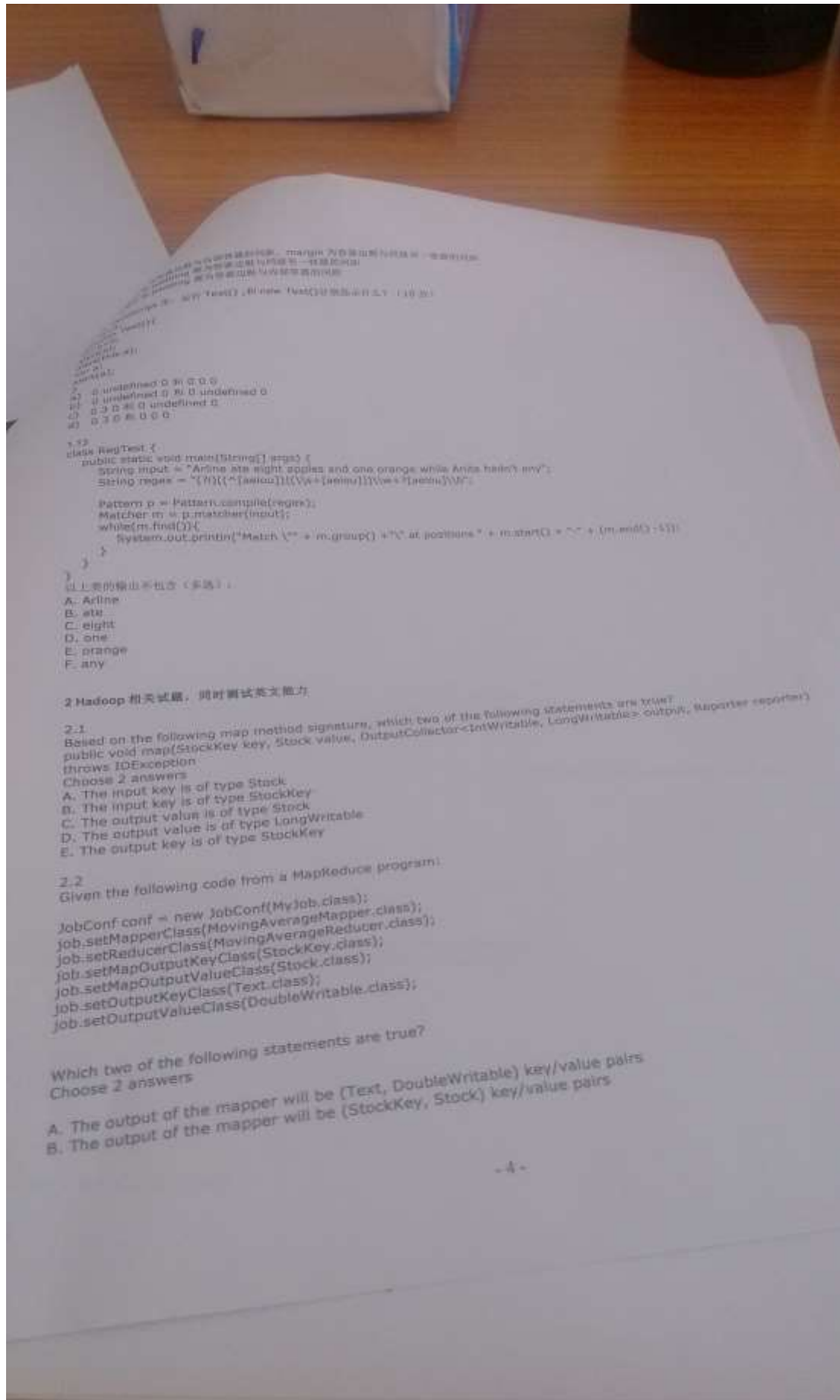
题目 4: SQL 数据去重

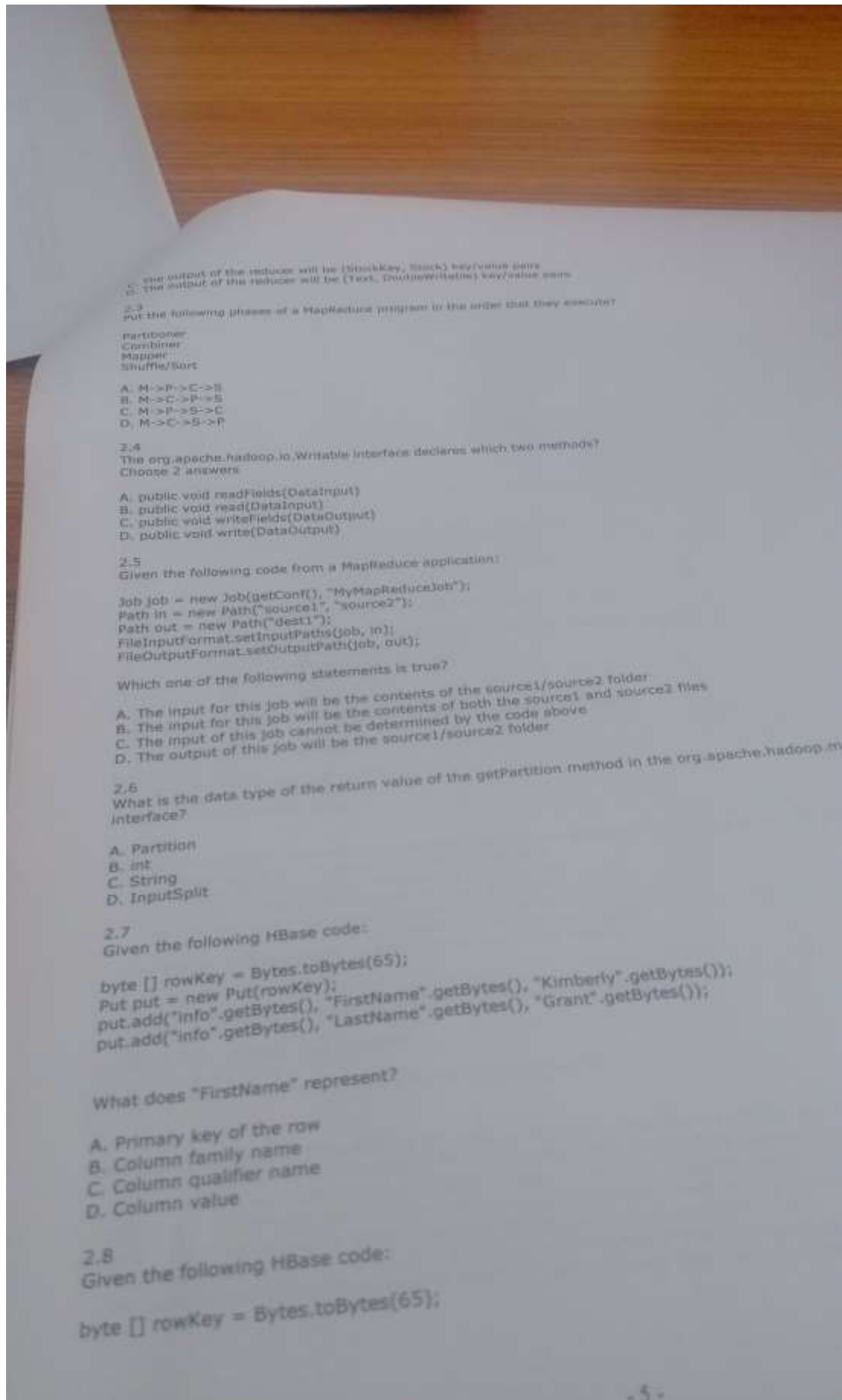
现有 test 表，字段包括 id (主键，自增长)，name(商品名)，price (价格)，type (种类)。

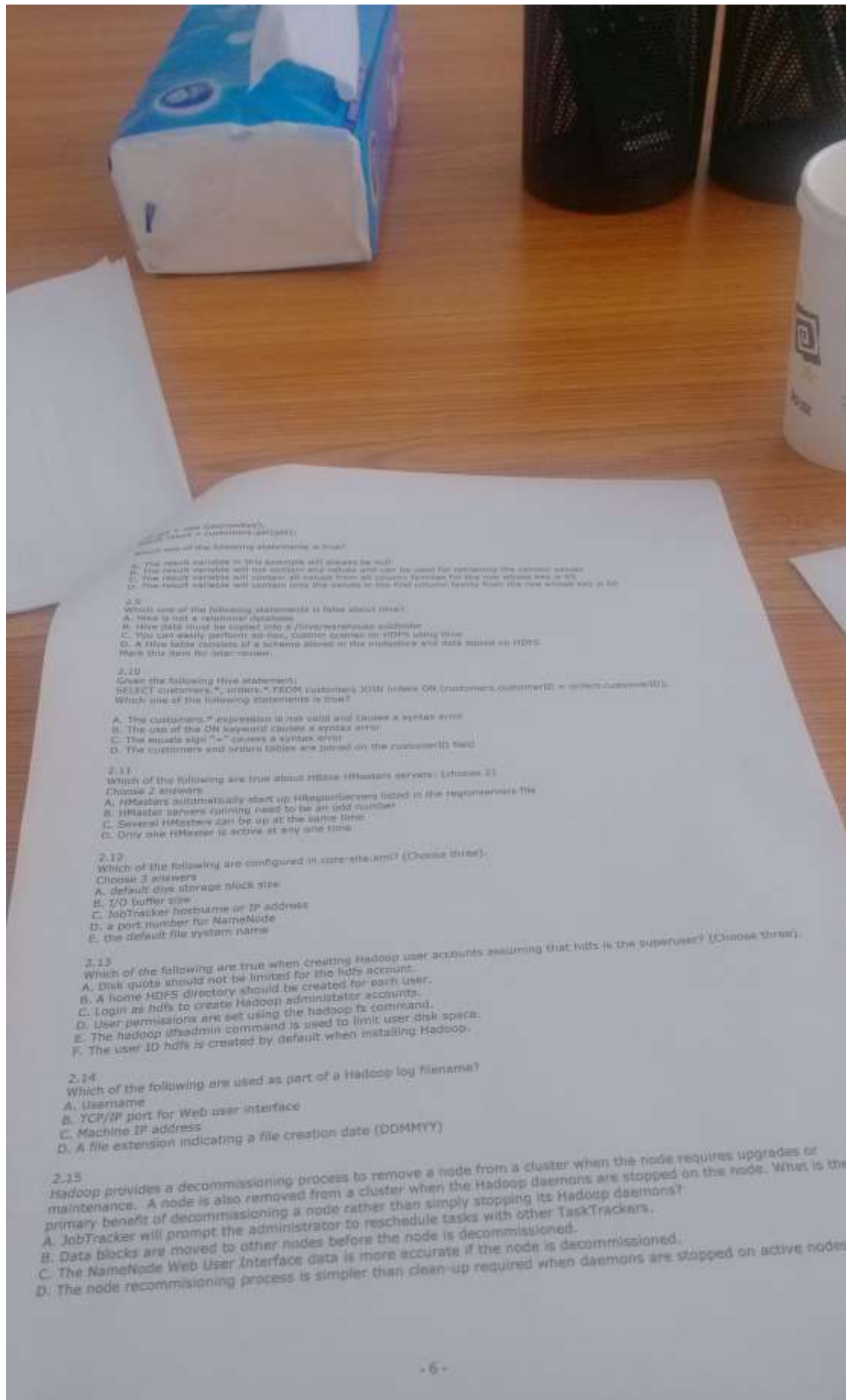
例如：

| id | name | type | price |
|----|------|------|-------|
| 1 | 商品 A | 1 | 10 |
| 2 | 商品 B | 2 | 20 |
| 3 | 商品 A | 1 | 10 |
| n | ... | ... | ... |

请编写 1 条 SQL 语句删除同种类下商品名、价格相同的重复数据。









(重点面试题)

- 1、Hive 内部表和外部表的区别？
- 2、Hbase 的 rowkey 怎么创建比较好？列族怎么创建比较好？
- 3、用 mapreduce 怎么处理数据倾斜问题？
- 4、hadoop 框架中怎么来优化？
- 5、Hbase 内部是什么机制？
- 6、我们在开发分布式计算 job 的,是否可以去掉 reduce()阶段？
- 7、hdfs 的数据压缩算法
- 8、mapreduce 的调度模式
- 9、hive 底层与数据库交互原理
- 10、hbase 过滤器实现原则
- 11、reduce 后输出的数据量有多大？
- 12、现场出问题测试 mapreduce 掌握情况和 HIVE 的 Hql 语句掌握情况？

三、来自 happy 提供的面试题 9 道：

- 1、datanode 在什么情况下不会备份？
- 2、combine 出现在那个过程？
- 3、hdfs 得体系结构？
- 4、flush 的过程？
- 5、什么是队列？
- 6、List 与 Set 的区别？
- 7、数据库的三大范式？



8、三个 datanode 当有一个 datanode 出现错误会怎样？

9、sqoop 在导入数据到 mysql 中，如何让数据不重复导入？如果存在数据问题 sqoop 如何处理？

四、来自*****提供的面试题 7 道：

1、使用 Hive 或者自定义 MR 实现如下逻辑

| product_no | lac_id | moment | start_time | user_id | county_id | staytime | city_id |
|-------------|--------|--------|-------------------------------|---------|-----------|----------|---------|
| 13429100031 | 22554 | 8 | 2013-03-11 08:55:19.151754088 | 571 | 571 | 282 | 571 |
| 13429100082 | 22540 | 8 | 2013-03-11 08:58:20.152622488 | 571 | 571 | 270 | 571 |
| 13429100082 | 22691 | 8 | 2013-03-11 08:56:37.149593624 | 571 | 571 | 103 | 571 |
| 13429100087 | 22705 | 8 | 2013-03-11 08:56:51.139539816 | 571 | 571 | 220 | 571 |
| 13429100087 | 22540 | 8 | 2013-03-11 08:55:45.150276800 | 571 | 571 | 66 | 571 |
| 13429100082 | 22540 | 8 | 2013-03-11 08:55:38.140225200 | 571 | 571 | 133 | 571 |
| 13429100140 | 26642 | 9 | 2013-03-11 09:02:19.151754088 | 571 | 571 | 18 | 571 |
| 13429100082 | 22691 | 8 | 2013-03-11 08:57:32.151754088 | 571 | 571 | 287 | 571 |
| 13429100189 | 22558 | 8 | 2013-03-11 08:56:24.139539816 | 571 | 571 | 48 | 571 |
| 13429100349 | 22503 | 8 | 2013-03-11 08:54:30.152622440 | 571 | 571 | 211 | 571 |

字段解释：

product_no：用户手机号；

lac_id：用户所在基站；

start_time：用户在此基站的开始时间；

staytime：用户在此基站的逗留时间。

需求描述：

根据 lac_id 和 start_time 知道用户当时的位置，根据 staytime 知道用户各个基站的逗留时长。根据轨迹合并连续基站的 staytime。

最终得到每一个用户按时间排序在每一个基站驻留时长

期望输出举例：

| | | | | | | | |
|-------------|-------|---|-------------------------------|-----|-----|-----|-----|
| 13429100082 | 22540 | 8 | 2013-03-11 08:58:20.152622488 | 571 | 571 | 270 | 571 |
| 13429100082 | 22691 | 8 | 2013-03-11 08:56:37.149593624 | 571 | 571 | 390 | 571 |



| | | | | | | | |
|-------------|-------|---|-------------------------------|-----|-----|-----|-----|
| 13429100082 | 22540 | 8 | 2013-03-11 08:55:38.140225200 | 571 | 571 | 133 | 571 |
| 13429100087 | 22705 | 8 | 2013-03-11 08:56:51.139539816 | 571 | 571 | 220 | 571 |
| 13429100087 | 22540 | 8 | 2013-03-11 08:55:45.150276800 | 571 | 571 | 66 | 571 |

2、Linux 脚本能力考察

2.1 请随意使用各种类型的脚本语言实现：批量将指定目录下的所有文件中的

\$HADOOP_HOME\$替换成/home/ocetl/app/hadoop

2.2 假设有 10 台主机，H1 到 H10，在开启 SSH 互信的情况下，编写一个或多个脚本实现

在所有的远程主机上执行脚本的功能

例如：runRemoteCmd.sh "ls -l"

期望结果：

H1:

XXXXXXXX

XXXXXXXX

XXXXXXXX

H2:

XXXXXXXX

XXXXXXXX

XXXXXXXX

H3:

...

3 Hadoop 基础知识与问题分析的能力

3.1 描述一下 hadoop 中，有哪些地方使用了缓存机制，作用分别是什么

3.2 请描述 <https://issues.apache.org/jira/browse/HDFS-2379> 说的是什么问题，最终解决的思路是什么？



4、MapReduce 开发能力

请参照 wordcount 实现一个自己的 map reduce，需求为：

a 输入文件格式：

xxx,xxx,xxx,xxx,xxx,xxx,xxx

b 输出文件格式：

xxx,20

xxx,30

xxx,40

c 功能：根据命令行参数统计输入文件中指定关键字出现的次数，并展示出来

例如：hadoop jar xxxxx.jar keywordcount xxx,xxx,xxx,xxx(四个关键字)

5、MapReduce 优化

请根据第五题中的程序，提出如何优化 MR 程序运行速度的思路

6、Linux 操作系统知识考察

请列举曾经修改过的/etc 下的配置文件，并说明修改要解决的问题？

7、Java 开发能力

7.1 写代码实现 1G 大小的文本文件，行分隔符为\x01\x02,统计一下该文件中的总行数，

要求注意边界情况的处理

7.2 请描述一下在开发中如何对上面的程序进行性能分析，对性能进行优化的过程。

五、来自*****提供的 hadoop 面试题 21 道：

1、设计一套系统，使之能够从不断增加的不同的数据源中，提取指定格式的数据。

要求：



1)、运行结果要能大致得知提取效果，并可据此持续改进提取方法；

2)、由于数据来源的差异性，请给出可弹性配置的程序框架；

3)、数据来源可能有 Mysql,sqlserver 等；

4)、该系统具备持续挖掘的能力，即，可重复提取更多信息

2. 经典的一道题：

现有 1 亿个整数均匀分布，如果要得到前 1K 个最大的数，求最优的算法。

(先不考虑内存的限制，也不考虑读写外存，时间复杂度最少的算法即为最优算法)

我先说下我的想法:分块，比如分 1W 块，每块 1W 个，然后分别找出每块最大值，从这最大的 1W 个值中找最大 1K 个，

那么其他的 9K 个最大值所在的块即可扔掉，从剩下的最大的 1K 个值所在的块中找前 1K 个即可。那么原问题的规模就缩小到了 1/10。

问题：

(1) 这种分块方法的最优时间复杂度。

(2) 如何分块达到最优。比如也可分 10W 块，每块 1000 个数。则问题规模可降到原来 1/100。但事实上复杂度并没降低。

(3) 还有没更好更优的方法解决这个问题。

3. MapReduce 大致流程？

4. combiner, partition 作用？

5. 用 mapreduce 实现 sql 语句 select count(x) from a group by b？

6. 用 mapreduce 如何实现两张表连接，有哪些方法？

7. 知道 MapReduce 大致流程，map, shuffle, reduce



8.知道 combiner, partition 作用，设置 compression

9.搭建 hadoop 集群，master/slave 都运行那些服务

10.HDFS，replica 如何定位

11.版本 0.20.2->0.20.203->0.20.205, 0.21, 0.23, 1.0. 1

新旧 API 有什么不同

12.Hadoop 参数调优，cluster level: JVM, map/reduce slots, job level: reducer

#,memory, use combiner? use compression?

13.pig latin, Hive 语法有什么不同

14.描述 HBase, zookeeper 搭建过程

15.hadoop 运行的原理?

16.mapreduce 的原理?

17.HDFS 存储的机制?

18.举一个简单的例子说明 mapreduce 是怎么来运行的？

19.使用 mapreduce 来实现下面实例

实例:现在有 10 个文件夹,每个文件夹都有 1000000 个 url.现在让你找出 top1000000url。

20.hadoop 中 Combiner 的作用?

21.如何确认 Hadoop 集群的健康状况。

六、来自****提供的 hadoop 面试题 9 道:

1.使用的 hadoop 版本都是什么？

编者 QQ: 1040195253



- 2.mapreduce 原理是什么？
- 3.mapreduce 作业，不使用 reduce 来输出，用什么能代替 reduce 的功能
- 4.hive 如何调优？
- 5.hive 如何权限控制？
- 6.hbase 写数据的原理是什么？
- 7.hive 能像关系数据库那样，建多个库吗？
- 8.hbase 宕机如何处理？
- 9.假设公司要建一个数据中心，你会如何规划？

七、hadoop 选择判断题 33 道：

单项选择题

1. 下面哪个程序负责 HDFS 数据存储。
a)NameNode b)Jobtracker c)Datanode d)secondaryNameNode e)tasktracker
2. HDFS 中的 block 默认保存几份？
a)3 份 b)2 份 c)1 份 d)不确定
3. 下列哪个程序通常与 NameNode 在一个节点启动？
a)SecondaryNameNode b)DataNode c)TaskTracker d)Jobtracker
4. Hadoop 作者
a)Martin Fowler b)Kent Beck c)Doug cutting
5. HDFS 默认 Block Size
a)32MB b)64MB c)128MB
6. 下列哪项通常是集群的最主要瓶颈



a)CPU b)网络 c)磁盘 d)内存

7. 关于 SecondaryNameNode 哪项是正确的？

a)它是 NameNode 的热备 b)它对内存没有要求

c)它的目的是帮助 NameNode 合并编辑日志，减少 NameNode 启动时间

d)SecondaryNameNode 应与 NameNode 部署到一个节点

多选题：

8. 下列哪项可以作为集群的管理工具

a)Puppet b)Pdsh c)Cloudera Manager d)d)Zookeeper

9. 配置机架感知的下面哪项正确

a)如果一个机架出问题，不会影响数据读写

b)写入数据的时候会写到不同机架的 DataNode 中

c)MapReduce 会根据机架获取离自己比较近的网络数据

10. Client 端上传文件的时候下列哪项正确

a)数据经过 NameNode 传递给 DataNode

b)Client 端将文件切分为 Block，依次上传

c)Client 只上传数据到一台 DataNode，然后由 NameNode 负责 Block 复制工作

11. 下列哪个是 Hadoop 运行的模式

a)单机版 b)伪分布式 c)分布式

12. Cloudera 提供哪几种安装 CDH 的方法

a)Cloudera manager b)Tar ball c)Yum d)Rpm



判断题：

- 13. Ganglia 不仅可以进行监控，也可以进行告警。()
- 14. Block Size 是不可以修改的。()
- 15. Nagios 不可以监控 Hadoop 集群，因为它不提供 Hadoop 支持。()
- 16. 如果 NameNode 意外终止，SecondaryNameNode 会接替它使集群继续工作。()
- 17. Cloudera CDH 是需要付费使用的。()
- 18. Hadoop 是 Java 开发的，所以 MapReduce 只支持 Java 语言编写。()
- 19. Hadoop 支持数据的随机读写。()
- 20. NameNode 负责管理 metadata，client 端每次读写请求，它都会从磁盘中读取或则会写入 metadata 信息并反馈 client 端。()
- 21. NameNode 本地磁盘保存了 Block 的位置信息。()
- 22. DataNode 通过长连接与 NameNode 保持通信。()
- 23. Hadoop 自身具有严格的权限管理和安全措施保障集群正常运行。()
- 24. Slave 节点要存储数据，所以它的磁盘越大越好。()
- 25. `hadoop dfsadmin -report` 命令用于检测 HDFS 损坏块。()
- 26. Hadoop 默认调度器策略为 FIFO ()
- 27. 集群内每个节点都应该配 RAID，这样避免单磁盘损坏，影响整个节点运行。()
- 28. 因为 HDFS 有多个副本，所以 NameNode 是不存在单点问题的。()
- 29. 每个 map 槽就是一个线程。()
- 30. Mapreduce 的 input split 就是一个 block。()
- 31. NameNode 的 Web UI 端口是 50030，它通过 jetty 启动的 Web 服务。()



32. Hadoop 环境变量中的 HADOOP_HEAPSIZE 用于设置所有 Hadoop 守护线程的内存。它默

认是 200 GB。()

33. DataNode 首次加入 cluster 的时候，如果 log 中报告不兼容文件版本，那需要 NameNode

执行 “Hadoop namenode -format” 操作格式化磁盘。()

八、mr 和 hive 实现手机流量统计面试题 6 道：

- 1.hive 实现统计的查询语句是什么？
- 2.生产环境中为什么建议使用外部表？
- 3.hadoop mapreduce 创建类 DataWritable 的作用是什么？
- 4.为什么创建类 DataWritable？
- 5.如何实现统计手机流量？
- 6.对比 hive 与 mapreduce 统计手机流量的区别？

九、来自 aboutyun 的面试题 1 道：

最近去面试，出了个这样的题目，大家有兴趣也试试。

用 Hadoop 分析海量日志文件，每行日志记录了如下数据：

TableName(表名)，Time(时间)，User(用户)，TimeSpan(时间开销)。

要求：

编写 MapReduce 程序算出高峰时间段（如上午 10 点）哪张表被访问的最频繁，以及这段时间访问这张表最多的用户，以及这个用户的总时间开销。



十、来自 aboutyun 的面试题 6 道：

前段时间接到阿里巴巴面试云计算,拿出来给我们共享下

- 1、hadoop 运转的原理？
- 2、mapreduce 的原理？
- 3、HDFS 存储的机制？
- 4、举一个简略的比方阐明 mapreduce 是怎么来运转的？
- 5、面试的人给你出一些疑问,让你用 mapreduce 来完成？

比方:如今有 10 个文件夹,每个文件夹都有 1000000 个 url.如今让你找出 top1000000url。

- 6、hadoop 中 Combiner 的效果？

论坛某网友的回复：

- 1.hadoop 即是 mapreduce 的进程，服务器上的一个目录节点加上多个数据节点，将程序传递到各个节点，再节点上进行计算。
- 2.mapreduce 即是把数据存储到不一样的节点上，用 map 方法对应办理，在各个节点上进行计算，最后由 reduce 进行合并。
- 3.java 程序和 namenode 合作，把数据存放在不一样的数据节点上
- 4.怎么运转用图来表明最好了。图无法画。谷歌下
- 5.不思考歪斜，功能，运用 2 个 job，第一个 job 直接用 filesystem 读取 10 个文件夹作为 map 输入，url 做 key，reduce 计算个 url 的 sum，
下一个 job map 顶用 url 作 key，运用-sum 作二次排序，reduce 中取 top10000000



第二种方法，建 hive 表 A，挂分区 channel，每个文件夹是一个分区。

```
select x.url,x.c from(select url,count(1) as c from A where channel =" group by url)x order by x.c desc limit 1000000;
```

6 combiner 也是一个 reduce，它可以削减 map 到 reduce 的数据传输，进步 shuffle 速度。

牢记平均值不要用。需求输入=map 的输出，输出=reduce 的输入。

十一、小萝卜(hadoop 月薪 13k)的笔试和面试题 11 道：

一、笔试

1、java 基础类：

1) 继承：写的一段代码，让写出结果；

2) 引用对象和值对象；

Java 基础类记不太清了，有很多都是基础。

2、linux 基础：

1) find 用法

2) 给出一个文本：比如 http://aaa.com

http://bbb.com

http://bbb.com

http://bbb.com

http://ccc.com

http://ccc.com

让写 shell 统计，最后输出结果：aaa 1

Ccc 2



Bbb 3

要求结果还要排序

还有别的，也是比较基础的

3、数据库类：oracle 查询语句

二、面试

讲项目经验：问的很细，给纸，笔，让画公司 hadoop 的项目架构,最后还让自己说几条业务数据，然后经过平台后，出来成什么样子。

java 方面：io 输入输出流里有哪些常用的类，还有 webService,线程相关的知识

linux：问到 jps 命令，kill 命令，问 awk,sed 是干什么用的、还有 hadoop 的一些常用命令

hadoop：讲 hadoop1 中 map,shuffle,reduce 的过程，其中问到了 map 端和 reduce 端溢写的细节（幸好我之前有研究过）

项目部署：问了项目是怎么部署，代码怎么管理

Hive 也问了一些，外部表，还有就是 hive 的物理模型跟传统数据库的不同

三、某互联网公司的面试：

问到分析人行为的算法：我当时想到我们做的反洗钱项目中，有用到。我就给举例：我们是怎么筛选出可疑的洗钱行为的。

十二、闪客、找自己、大数等提供的面试题 26 道：

****信 Hadoop 面试笔试题（共 14 题，还有一题记不住了）

1、hadoop 集群搭建过程，写出步骤。

2、hadoop 集群运行过程中启动那些线程，各自的作用是什么？



3、/tmp/hadoop-root/dfs/name the path is not exists or is not accessible.

NameNode main 中报错，该怎么解决。（大意这样 一个什么异常）

4、工作中编写 mapreduce 用到的语言，编写一个 mapreduce 程序。

5、hadoop 命令

1) 杀死一个 job 任务 （杀死 50030 端口的进程即可）

2) 删除/tmp/aaa 文件目录

3) hadoop 集群添加或删除节点时，刷新集群状态的命令

6、日志的固定格式：

a,b,c,d

a,a,f,e

b,b,d,f

使用一种语言编写 mapreduce 任务，统计每一列最后字母的个数。

7、hadoop 的调度器有哪些，工作原理。

8、mapreduce 的 join 方法有哪些？

9、Hive 元数据保存的方法有哪些，各有什么特点？

10、java 实现非递归二分法算法。

11、mapreduce 中 Combiner 和 Partition 的作用。

12、用 linux 实现下列要求：

ip username

a.txt

210.121.123.12 zhangsan

34.23.56.78 lisi



11.56.56.72 wanger

.....

b.txt

58.23.53.132 liuqi

34.23.56.78 liba

.....

a.txt,b.txt 中至少 100 万行。

1) a.txt,b.txt 中各自的 ip 个数，ip 的总个数。

2) a.txt 中存在的 ip 而 b.txt 中不存在的 ip。

3) 每个 username 出现的总个数，每个 username 对应的 ip 个数。

13、大意是 hadoop 中 java、streaming、pipe 处理数据各有特点。

14、如何实现 mapreduce 的二次排序。

大数遇到的面试题：

15、面试官上来就问 hadoop 的调度机制；

16、机架感知；

17、MR 数据倾斜原因和解决方案；

18、集群 HA。

@找自己 提供的面试题：

19、如果让你设计，你觉得一个分布式文件系统应该如何设计，考虑哪方面内容；

每天百亿数据入 hbase，如何保证数据的存储正确和在规定的时间内全部录入完毕，不残留数据。

20、对于 hive，你写过哪些 UDF 函数，作用是什么



- 21、hdfs 的数据压缩算法
- 22、mapreduce 的调度模式
- 23、hive 底层与数据库交互原理
- 24、hbase 过滤器实现原则
- 25、对于 mahout，如何进行推荐、分类、聚类的代码二次开发分别实现那些借口
- 26、请问下，直接将时间戳作为行键，在写入单个 region 时候会发生热点问题，为什么呢？

十三、飞哥(hadoop 月薪 13k)提供的面试题 17 道：

- 1、hdfs 原理，以及各个模块的职责
- 2、mr 的工作原理
- 3、map 方法是如何调用 reduce 方法的
- 4、shell 如何判断文件是否存在，如果不存在该如何处理？
- 5、fsimage 和 edit 的区别？
- 6、hadoop1 和 hadoop2 的区别？

笔试：

- 1、hdfs 中的 block 默认报错几份？
- 2、哪个程序通常与 nn 在一个节点启动？并做分析
- 3、列举几个配置文件优化？
- 4、写出你对 zookeeper 的理解
- 5、datanode 首次加入 cluster 的时候，如果 log 报告不兼容文件版本，那需要 namenode 执行格式化操作，这样处理的原因是？



- 6、谈谈数据倾斜，如何发生的，并给出优化方案
- 7、介绍一下 hbase 过滤器
- 8、mapreduce 基本执行过程
- 9、谈谈 hadoop1 和 hadoop2 的区别
- 10、hbase 集群安装注意事项
- 11、记录包含值域 F 和值域 G，要分别统计相同 G 值的记录中不同的 F 值的数目，简单编写过程。

十四、飞哥(hadoop 月薪 13k)提供的面试题 3 道：

- 1、算法题：有 2 个桶，容量分别为 3 升和 5 升，如何得到 4 升的水，假设水无限使用，写出步骤。
- 2、java 笔试题：忘记拍照了，很多很基础的 se 知识。后面还有很多 sql 相关的题,常用的查询 sql 编写，答题时间一个小时。
- 3、Oracle 数据库中有一个表字段 name，name varchar2(10),如何在不改变表数据的情况下将此字段长度改为 varchar2 (2) ？

十五、海量数据处理算法面试题 10 道：

第一部分：十道海量数据处理面试题

- 1、海量日志数据，提取出某日访问百度次数最多的那个 IP。

首先是这一天，并且是访问百度的日志中的 IP 取出来，逐个写入到一个大文件中。注意到 IP 是 32 位的，最多有个 2^{32} 个 IP。同样可以采用映射的方法，比如模 1000，把整个大文件映射为 1000 个小文件，再找出每个小文中出现频率最大的 IP(可以采用 hash_map



进行频率统计，然后再找出频率最大 的几个) 及相应的频率。然后再在这 1000 个最大的 IP 中，找出那个频率最大的 IP，即为所求。

或者如下阐述 (雪域之鹰)：

算法思想：分而治之+Hash

- (1) .IP 地址最多有 $2^{32}=4G$ 种取值情况，所以不能完全加载到内存中处理；
- (2) .可以考虑采用“分而治之”的思想，按照 IP 地址的 Hash(IP)%1024 值，把海量 IP 日志分别存储到 1024 个小文件中。这样，每个小文件最多包含 4MB 个 IP 地址；
- (3) .对于每一个小文件，可以构建一个 IP 为 key，出现次数为 value 的 Hash map，同时记录当前出现次数最多的那个 IP 地址；
- (4) .可以得到 1024 个小文件中的出现次数最多的 IP，再依据常规的排序算法得到总体上出现次数最多的 IP；

2、搜索引擎会通过日志文件把用户每次检索使用的所有检索串都记录下来，每个查询串的长度为 1-255 字节。

假设目前有一千万个记录 (这些查询串的重复度比较高，虽然总数是 1 千万，但如果除去重复后，不超过 3 百万个。一个查询串的重复度越高，说明查询它的用户越多，也就是越热门。)，请你统计最热门的 10 个查询串，要求使用的内存不能超过 1G。

典型的 Top K 算法，还是在这篇文章里头有所阐述，详情请参见：十一、从头到尾彻底解析 Hash 表算法。

文中，给出的最终算法是：

第一步、先对这批海量数据预处理，在 $O(N)$ 的时间内用 Hash 表完成统计 (之前写成了排序，特此订正。July、2011.04.27)；



第二步、借助堆这个数据结构，找出 Top K，时间复杂度为 $N \cdot \log K$ 。

即，借助堆结构，我们可以在 \log 量级的时间内查找和调整/移动。因此，维护一个 K(该题目中是 10)大小的小根堆，然后遍历 300 万的 Query，分别和根元素进行对比所以，我们最终的时间复杂度是： $O(N) + N' \cdot O(\log K)$ ，(N 为 1000 万，N' 为 300 万)。ok，更多，详情，请参考原文。

或者：采用 trie 树，关键字域存该查询串出现的次数，没有出现为 0。最后用 10 个元素的最小堆来对出现频率进行排序。

3、有一个 1G 大小的一个文件，里面每一行是一个词，词的大小不超过 16 字节，内存限制大小是 1M。返回频数最高的 100 个词。

方案：顺序读文件中，对于每个词 x，取 $\text{hash}(x) \% 5000$ ，然后按照该值存到 5000 个小文件（记为 $x_0, x_1, \dots, x_{4999}$ ）中。这样每个文件大概是 200k 左右。

如果其中的有的文件超过了 1M 大小，还可以按照类似的方法继续往下分，直到分解得到的小文件的大小都不超过 1M。

对每个小文件 统计每个文件中出现的词以及相应的频率（可以采用 trie 树/hash_map 等），并取出出现频率最大的 100 个词（可以用含 100 个结点的堆），并把 100 个词及相应的频率存入文件，这样又得到了 5000 个文件。下一步就是把这 5000 个文件进行归并（类似与归并排序）的过程了。

4、有 10 个文件，每个文件 1G，每个文件的每一行存放的都是用户的 query，每个文件的 query 都可能重复。要求你按照 query 的频度排序。

还是典型的 TOP K 算法，解决方案如下：



方案 1：

顺序读取 10 个文件，按照 $\text{hash}(\text{query})\%10$ 的结果将 query 写入到另外 10 个文件（记为）

中。这样新生成的文件每个的大小大约也 1G（假设 hash 函数是随机的）。

找一台内存在 2G 左右的机器，依次对用 $\text{hash_map}(\text{query}, \text{query_count})$ 来统计每个 query 出现的次数。利用快速/堆/归并排序按照出现次数进行排序。将排序好的 query 和

对应的 query_count 输出到文件中。这样得到了 10 个排好序的文件（记为）。

对这 10 个文件进行归并排序（内排序与外排序相结合）。

方案 2：

一般 query 的总量是有限的，只是重复的次数比较多而已，可能对于所有的 query，一次性就可以加入到内存了。这样，我们就可以采用 trie 树/hash_map 等直接来统计每个 query 出现的次数，然后按出现次数做快速/堆/归并排序就可以了。

方案 3：

与方案 1 类似，但在做完 hash，分成多个文件后，可以交给多个文件来处理，采用分布式的架构来处理（比如 MapReduce），最后再进行合并。

5、给定 a、b 两个文件，各存放 50 亿个 url，每个 url 各占 64 字节，内存限制是 4G，让你找出 a、b 文件共同的 url？

方案 1：可以估计每个文件安的大小为 $5G \times 64 = 320G$ ，远远大于内存限制的 4G。所以不可能将其完全加载到内存中处理。考虑采取分而治之的方法。

遍历文件 a，对每个 url 求取 $\text{hash}(\text{url})\%1000$ ，然后根据所取得的值将 url 分别存储到 1000 个小文件（记为 a_0, a_1, \dots, a_{999} ）中。这样每个小文件的大约为 300M。

遍历文件 b，采取和 a 相同的方式将 url 分别存储到 1000 小文件（记为 b_0, b_1, \dots, b_{999} ）。



这样处理后，所有可能相同的 url 都在对应的小文件（a0vsb0,a1vsb1,...,a999vsb999）中，不对应的小文件不可能有相同的 url。然后我们只要求出 1000 对小文件中相同的 url 即可。

求每对小文件中相同的 url 时，可以把其中一个小文件的 url 存储到 hash_set 中。然后遍历另一个小文件的每个 url，看其是否在刚才构建的 hash_set 中，如果是，那么就是共同的 url，存到文件里面就可以了。

方案 2：如果允许有一定的错误率，可以使用 Bloom filter，4G 内存大概可以表示 340 亿 bit。将其中一个文件中的 url 使用 Bloom filter 映射为这 340 亿 bit，然后挨个读取另外一个文件的 url，检查是否与 Bloom filter，如果是，那么该 url 应该是共同的 url（注意会有一定的错误率）。

Bloom filter 日后会在本 BLOG 内详细阐述。

6、在 2.5 亿个整数中找出不重复的整数，注，内存不足以容纳这 2.5 亿个整数。

方案 1：采用 2-Bitmap（每个数分配 2bit，00 表示不存在，01 表示出现一次，10 表示多次，11 无意义）进行，共需内存 $2^{32} * 2 \text{ bit} = 1 \text{ GB}$ 内存，还可以接受。然后扫描这 2.5 亿个整数，查看 Bitmap 中相对位，如果是 00 变 01，01 变 10，10 保持不变。扫描完后，查看 bitmap，把对应位是 01 的整数输出即可。

方案 2：也可采用与第 1 题类似的方法，进行划分小文件的方法。然后在小文件中找出不重复的整数，并排序。然后再进行归并，注意去除重复的元素。

7、腾讯面试题：给 40 亿个不重复的 unsigned int 的整数，没排过序的，然后再给一个数，如何快速判断这个数是否在那 40 亿个数当中？



与上第 6 题类似，我的第一反应时快速排序+二分查找。以下是其它更好的方法：

方案 1：oo，申请 512M 的内存，一个 bit 位代表一个 unsigned int 值。读入 40 亿个数，设置相应的 bit 位，读入要查询的数，查看相应 bit 位是否为 1，为 1 表示存在，为 0 表示不存在。

dizengrong：

方案 2：这个问题在《编程珠玑》里有很好的描述，大家可以参考下面的思路，探讨一下：

又因为 2^{32} 为 40 亿多，所以给定一个数可能在，也可能不在其中；

这里我们把 40 亿个数中的每一个用 32 位的二进制来表示

假设这 40 亿个数开始放在一个文件中。

然后将这 40 亿个数分成两类：

1.最高位为 0

2.最高位为 1

并将这两类分别写入到两个文件中，其中一个文件中数的个数 ≤ 20 亿，而另一个 ≥ 20 亿

(这相当于折半了)；

与要查找的数的最高位比较并接着进入相应的文件再查找

再然后把这个文件为又分成两类：

1.次最高位为 0

2.次最高位为 1

并将这两类分别写入到两个文件中，其中一个文件中数的个数 ≤ 10 亿，而另一个 ≥ 10 亿

(这相当于折半了)；

与要查找的数的次最高位比较并接着进入相应的文件再查找。

.....



以此类推，就可以找到了,而且时间复杂度为 $O(\log n)$ ，方案 2 完。

附：这里，再简单介绍下，位图方法：

使用位图法判断整形数组是否存在重复

判断集合中存在重复是常见编程任务之一,当集合中数据量比较大时我们通常希望少进行几次扫描，这时双重循环法就不可取了。

位图法比较适合于这种情况,它的做法是按照集合中最大元素 \max 创建一个长度为 $\max+1$ 的新数组，然后再次扫描原数组，遇到几就给新数组的第几位置上 1，如遇到 5 就给新数组的第六个元素置 1，这样下次再遇到 5 想置位时发现新数组的第六个元素已经是 1 了，这说明这次的数据肯定和以前的数据存在着重复。这种给新数组初始化时置零其后置一的做法类似于位图的处理方法故称位图法。它的运算次数最坏的情况为 $2N$ 。如果已知数组的最大值即能事先给新数组定长的话效率还能提高一倍。

欢迎，有更好的思路，或方法，共同交流。

8、怎么在海量数据中找出重复次数最多的一个？

方案 1：先做 hash，然后求模映射为小文件，求出每个小文件中重复次数最多的一个，并记录重复次数。然后找出上一步求出的数据中重复次数最多的一个就是所求（具体参考前面的题）。

9、上千万或上亿数据（有重复），统计其中出现次数最多的钱 N 个数据。

方案 1：上千万或上亿的数据，现在的机器的内存应该能存下。所以考虑采用 `hash_map/` 搜索二叉树/红黑树等来进行统计次数。然后就是取出前 N 个出现次数最多的数据了，可以用第 2 题提到的堆机制完成。



10、一个文本文件，大约有一万行，每行一个词，要求统计出其中最频繁出现的前 10 个词，请给出思想，给出时间复杂度分析。

方案 1：这题是考虑时间效率。用 trie 树统计每个词出现的次数，时间复杂度是 $O(n * le)$ (le 表示单词的平准长度)。然后是找出出现最频繁的前 10 个词，可以用堆来实现，前面的题中已经讲到了，时间复杂度是 $O(n * \lg 10)$ 。所以总的时间复杂度，是 $O(n * le)$ 与 $O(n * \lg 10)$ 中较大的哪一个。

附、100w 个数中找出最大的 100 个数。

方案 1：在前面的题中，我们已经提到了，用一个含 100 个元素的最小堆完成。复杂度为 $O(100w * \lg 100)$ 。

方案 2：采用快速排序的思想，每次分割之后只考虑比轴大的一部分，知道比轴大的一部分在比 100 多的时候，采用传统排序算法排序，取前 100 个。复杂度为 $O(100w * 100)$ 。

方案 3：采用局部淘汰法。选取前 100 个元素，并排序，记为序列 L。然后一次扫描剩余的元素 x，与排好序的 100 个元素中最小的元素比，如果比这个最小的要大，那么把这个最小的元素删除，并把 x 利用插入排序的思想，插入到序列 L 中。依次循环，知道扫描了所有的元素。复杂度为 $O(100w * 100)$ 。

致谢：<http://www.cnblogs.com/youwang/>。

第二部分：十个海量数据处理方法大总结

ok，看了上面这么多的面试题，是否有点头晕。是的，需要一个总结。接下来，本文将简单总结下一些处理海量数据问题的常见方法，而日后，本 BLOG 内会具体阐述这些方法。



下面的方法全部来自 <http://hi.baidu.com/yanxionggu/blog/> 博客，对海量数据的处理方法进行了一个一般性的总结，当然这些方法可能并不能完全覆盖所有的问题，但是这样的一些方法也基本可以处理绝大多数遇到的问题。下面的一些问题基本直接来源于公司的面试笔试题目，方法不一定最优，如果你有更好的处理方法，欢迎讨论。

一、Bloom filter

适用范围：可以用来实现数据字典，进行数据的判重，或者集合求交集

基本原理及要点：

对于原理来说很简单，位数组+k 个独立 hash 函数。将 hash 函数对应的值的位数组置 1，查找时如果发现所有 hash 函数对应位都是 1 说明存在，很明显这个过程并不保证查找的结果是 100%正确的。同时也不支持删除一个已经插入的关键字，因为该关键字对应的位会牵动到其他的关键字。所以一个简单的改进就是 counting Bloom filter，用一个 counter 数组代替位数组，就可以支持删除了。

还有一个比较重要的问题 如何根据输入元素个数 n 确定位数组 m 的大小及 hash 函数个数。当 hash 函数个数 $k = (\ln 2) * (m/n)$ 时错误率最小。在错误率不大于 E 的情况下，m 至少要等于 $n * \lg(1/E)$ 才能表示任意 n 个元素的集合。但 m 还应该更大些，因为还要保证 bit 数组里至少一半为 0，则 m 应该 $\geq n \lg(1/E) * \lg e$ 大概就是 $n \lg(1/E) 1.44$ 倍 (\lg 表示以 2 为底的对数)。

举个例子我们假设错误率为 0.01，则此时 m 应大概是 n 的 13 倍。这样 k 大概是 8 个。

注意这里 m 与 n 的单位不同，m 是 bit 为单位，而 n 则是以元素个数为单位(准确的说是不同元素的个数)。通常单个元素的长度都是有很多 bit 的。所以使用 bloom filter 内存上通常都是节省的。



扩展：

Bloom filter 将集合中的元素映射到位数组中，用 k (k 为哈希函数个数) 个映射位是否全 1 表示元素在不在这个集合中。Counting bloom filter (CBF) 将位数组中的每一位扩展为一个 counter，从而支持了元素的删除操作。Spectral Bloom Filter (SBF) 将其与集合元素的出现次数关联。SBF 采用 counter 中的最小值来近似表示元素的出现频率。

问题实例：给你 A,B 两个文件，各存放 50 亿条 URL，每条 URL 占用 64 字节，内存限制是 4G，让你找出 A,B 文件共同的 URL。如果是三个乃至 n 个文件呢？

根据这个问题我们来计算下内存的占用， $4G = 2^{32}$ 大概是 40 亿 * 8 大概是 340 亿， $n=50$ 亿，如果按出错率 0.01 算需要的大概是 650 亿个 bit。现在可用的是 340 亿，相差并不多，这样可能会使出错率上升些。另外如果这些 urlip 是一一对应的，就可以转换成 ip，则大大简单了。

二、Hashing

适用范围：快速查找，删除的基本数据结构，通常需要总数据量可以放入内存

基本原理及要点：

hash 函数选择，针对字符串，整数，排列，具体相应的 hash 方法。

碰撞处理，一种是 open hashing，也称为拉链法；另一种就是 closed hashing，也称开地址法，opened addressing。

扩展：

d-left hashing 中的 d 是多个的意思，我们先简化这个问题，看一看 2-left hashing。2-left hashing 指的是将一个哈希表分成长度相等的两半，分别叫做 T1 和 T2，给 T1 和 T2 分别配备一个哈希函数， h_1 和 h_2 。在存储一个新的 key 时，同时用两个哈希函数进行计算，



得出两个地址 $h1[key]$ 和 $h2[key]$ 。这时需要检查 T1 中的 $h1[key]$ 位置和 T2 中的 $h2[key]$ 位置，哪一个位置已经存储的（有碰撞的）key 比较多，然后将新 key 存储在负载少的位置。如果两边一样多，比如两个位置都为空或者都存储了一个 key，就把新 key 存储在左边的 T1 子表中，2-left 也由此而来。在查找一个 key 时，必须进行两次 hash，同时查找两个位置。

问题实例：

1). 海量日志数据，提取出某日访问百度次数最多的那个 IP。

IP 的数目还是有限的，最多 2^{32} 个，所以可以考虑使用 hash 将 ip 直接存入内存，然后进行统计。

三、bit-map

适用范围：可进行数据的快速查找，判重，删除，一般来说数据范围是 int 的 10 倍以下

基本原理及要点：使用 bit 数组来表示某些元素是否存在，比如 8 位电话号码

扩展：bloom filter 可以看做是对 bit-map 的扩展

问题实例：

1) 已知某个文件内包含一些电话号码，每个号码为 8 位数字，统计不同号码的个数。

8 位最多 99 999 999，大概需要 99m 个 bit，大概 10 几 m 字节的内存即可。

2) 2.5 亿个整数中找出不重复的整数的个数，内存空间不足以容纳这 2.5 亿个整数。

将 bit-map 扩展一下，用 2bit 表示一个数即可，0 表示未出现，1 表示出现一次，2 表示出现 2 次及以上。或者我们不用 2bit 来进行表示，我们用两个 bit-map 即可模拟实现这个 2bit-map。



四、堆

适用范围：海量数据前 n 大，并且 n 比较小，堆可以放入内存

基本原理及要点：最大堆求前 n 小，最小堆求前 n 大。方法，比如求前 n 小，我们比较当前元素与最大堆里的最大元素，如果它小于最大元素，则应该替换那个最大元素。这样最后得到的 n 个元素就是最小的 n 个。适合大数据量，求前 n 小， n 的大小比较小的情况，这样可以扫描一遍即可得到所有的前 n 元素，效率很高。

扩展：双堆，一个最大堆与一个最小堆结合，可以用来维护中位数。

问题实例：

1) 100w 个数中找最大的前 100 个数。

用一个 100 个元素大小的最小堆即可。

五、双层桶划分——其实本质上就是【分而治之】的思想，重在“分”的技巧上！

适用范围：第 k 大，中位数，不重复或重复的数字

基本原理及要点：因为元素范围很大，不能利用直接寻址表，所以通过多次划分，逐步确定范围，然后最后在一个可以接受的范围内进行。可以通过多次缩小，双层只是一个例子。

扩展：

问题实例：

1) 2.5 亿个整数中找出不重复的整数的个数，内存空间不足以容纳这 2.5 亿个整数。

有点像鸽巢原理，整数个数为 2^{32} ，也就是，我们可以将这 2^{32} 个数，划分为 2^8 个区域(比如用单个文件代表一个区域)，然后将数据分离到不同的区域，然后不同的区域在利用 bitmap 就可以直接解决了。也就是说只要有足够的磁盘空间，就可以很方便的解决。

2) 5 亿个 int 找它们的中位数。



这个例子比上面那个更明显。首先我们将 int 划分为 2^{16} 个区域，然后读取数据统计落到各个区域里的数的个数，之后我们根据统计结果就可以判断中位数落到那个区域，同时知道这个区域中的第几大数刚好是中位数。然后第二次扫描我们只统计落在该区域中的那些数就可以了。

实际上，如果不是 int 是 int64，我们可以经过 3 次这样的划分即可降低到可以接受的程度。即可以先将 int64 分成 2^{24} 个区域，然后确定区域的第几大数，再将该区域分成 2^{20} 个子区域，然后确定是子区域的第几大数，然后子区域里的数的个数只有 2^{20} ，就可以直接利用 direct addr table 进行统计了。

六、数据库索引

适用范围：大数据量的增删改查

基本原理及要点：利用数据的设计实现方法，对海量数据的增删改查进行处理。

七、倒排索引(Inverted index)

适用范围：搜索引擎，关键字查询

基本原理及要点：为何叫倒排索引？一种索引方法，被用来存储在全文搜索下某个单词在一个文档或者一组文档中的存储位置的映射。

以英文为例，下面是要被索引的文本：

T0 = "it is what it is"

T1 = "what is it"

T2 = "it is a banana"

我们就能得到下面的反向文件索引：



"a" : {2}

"banana" : {2}

"is" : {0, 1, 2}

"it" : {0, 1, 2}

"what" : {0, 1}

检索的条件" what" ," is" 和" it" 将对应集合的交集。

正向索引开发出来用来存储每个文档的单词的列表。正向索引的查询往往满足每个文档有序频繁的全文查询和每个单词在校验文档中的验证这样的查询。在正向索引中，文档占据了中心的位置，每个文档指向了一个它所包含的索引项的序列。也就是说文档 指向了它包含的那些单词，而反向索引则是单词指向了包含它的文档，很容易看到这个反向的关系。

扩展：

问题实例：文档检索系统，查询那些文件包含了某单词，比如常见的学术论文的关键字搜索。

八、外排序

适用范围：大数据的排序，去重

基本原理及要点：外排序的归并方法，置换选择败者树原理，最优归并树

扩展：

问题实例：

1).有一个 1G 大小的一个文件，里面每一行是一个词，词的大小不超过 16 个字节，内存限制大小是 1M。返回频数最高的 100 个词。

这个数据具有很明显的特点，词的大小为 16 个字节，但是内存只有 1m 做 hash 有些不够，

所以可以用来排序。内存可以当输入缓冲区使用。



九、trie 树

适用范围：数据量大，重复多，但是数据种类小可以放入内存

基本原理及要点：实现方式，节点孩子的表示方式

扩展：压缩实现。

问题实例：

- 1).有 10 个文件，每个文件 1G，每个文件的每一行都存放的是用户的 query，每个文件的 query 都可能重复。要你按照 query 的频度排序。
- 2).1000 万字符串，其中有些是相同的(重复),需要把重复的全部去掉，保留没有重复的字符串。请问怎么设计和实现？
- 3).寻找热门查询：查询串的重复度比较高，虽然总数是 1 千万，但如果除去重复后，不超过 3 百万个，每个不超过 255 字节。

十、分布式处理 mapreduce

适用范围：数据量大，但是数据种类小可以放入内存

基本原理及要点：将数据交给不同的机器去处理，数据划分，结果归约。

扩展：

问题实例：

- 1).The canonical example application of MapReduce is a process to count the appearances of each different word in a set of documents:
- 2).海量数据分布在 100 台电脑中，想个办法高效统计出这批数据的 TOP10。



3).一共有 N 个机器，每个机器上有 N 个数。每个机器最多存 $O(N)$ 个数并对它们操作。如何找到 N^2 个数的中数(median)？

经典问题分析

上千万 or 亿数据（有重复），统计其中出现次数最多的前 N 个数据,分两种情况：可一次读入内存，不可一次读入。

可用思路：trie 树+堆，数据库索引，划分子集分别统计，hash，分布式计算，近似统计，外排序

所谓的是否能一次读入内存，实际上应该指去除重复后的数据量。如果去重后数据可以放入内存，我们可以为数据建立字典，比如通过 map，hashmap，trie，然后直接进行统计即可。当然在更新每条数据的出现次数的时候，我们可以利用一个堆来维护出现次数最多的前 N 个数据，当然这样导致维护次数增加，不如完全统计后在求前 N 大效率高。

如果数据无法放入内存。一方面我们可以考虑上面的字典方法能否被改进以适应这种情形，可以做的改变就是将字典存放到硬盘上，而不是内存，这可以参考数据库的存储方法。

当然还有更好的方法，就是可以采用分布式计算，基本上就是 map-reduce 过程，首先可以根据数据值或者把数据 hash(md5)后的值，将数据按照范围划分到不同的机子，最好可以让数据划分后可以一次读入内存，这样不同的机子负责处理各种的数值范围，实际上就是 map。得到结果后，各个机子只需拿出各自的出现次数最多的前 N 个数据，然后汇总，选出所有的数据中出现次数最多的前 N 个数据，这实际上就是 reduce 过程。



实际上可能想直接将数据均分到不同的机子上进行处理，这样是无法得到正确的解的。

因为一个数据可能被均分到不同的机子上，而另一个则可能完全聚集到一个机子上，同时还可能存在具有相同数目的数据。比如我们要找出现次数最多的前 100 个，我们将 1000 万的数据分布到 10 台机器上，找到每台出现次数最多的前 100 个，归并之后这样不能保证找到真正的第 100 个，因为比如出现次数最多的第 100 个可能有 1 万个，但是它被分到了 10 台机子，这样在每台上只有 1 千 个，假设这些机子排名在 1000 个之前的那些都是单独分布在一台机子上的，比如有 1001 个，这样本来具有 1 万个的这个就会被淘汰，即使我们让每台机子选 出出现次数最多的 1000 个再归并，仍然会出错，因为可能存在大量个数为 1001 个的发生聚集。因此不能将数据随便均分到不同机子上，而是要根据 hash 后的值将它们映射到不同的机子上处理，让不同的机器处理一个数值范围。

而外排序的方法会消耗大量的 IO，效率不会很高。而上面的分布式方法，也可以用于单机版本，也就是将总的数据根据值的范围，划分成多个不同的子文件，然后逐个处理。处理完毕之后再对这些单词的及其出现频率进行一个归并。实际上就可以利用一个外排序的归并过程。

另外，还可以考虑近似计算，也就是我们可以通过结合自然语言属性，只将那些真正实际中出现最多的那些词作为一个字典，使得这个规模可以放入内存。

十六、来自 aboutyun 的面试题 6 道：

- 1.说说值对象与引用对象的区别？
- 2.谈谈你对反射机制的理解及其用途？
- 3.ArrayList、Vector、LinkedList 的区别及其优缺点？HashMap、HashTable 的区别及其



优缺点？

3.列出线程的实现方式？如何实现同步？

4.sql 题,是一个图表，具体忘了

5.列出至少五种设计模式？用代码或 UML 类图描述其中两种设计模式的原理？

6.谈谈你最近正在研究的技术，谈谈你最近项目中用到的技术难点及其解决思路。

十七、来自巴图提供的算法面试题 1 道：

| 用户手机号 | 出现的地点 | 出现的时间 | 逗留的时间 |
|------------|-------|----------------------------|-------|
| 1111111111 | 2 | 2014-02-18 19:03:56.123445 | 133 |
| 222222222 | 1 | 2013-03-14 03:18:45.263536 | 241 |
| 333333333 | 3 | 2014-10-23 17:14:23.176345 | 68 |
| 222222222 | 1 | 2013-03-14 03:20:47.123445 | 145 |
| 333333333 | 3 | 2014-09-15 15:24:56.222222 | 345 |
| 222222222 | 2 | 2011-08-30 18:13:58.111111 | 145 |
| 222222222 | 2 | 2011-08-30 18:18:24.222222 | 130 |

按时间排序

期望结果是：

| | | | |
|------------|-------|----------------------------|-----|
| 222222222 | 2 | 2011-08-30 18:13:58.111111 | 145 |
| 222222222 | 2 | 2011-08-30 18:18:24.222222 | 130 |
| 222222222 | 1 | 2013-03-14 03:18:45.263536 | 24 |
| 1111111111 | ~~~~~ | | |
| 333333333 | ~~~~~ | | |



十八、来自象夫提供的面试题 7 道：

Hdfs:

- 1.文件大小默认为 64M，改为 128M 有啥影响？
- 2.RPC 原理？
- 3.NameNode 与 SecondaryNameNode 的区别与联系？

MapReduce:

- 4.介绍 MapReduce 整个过程，比如把 WordCount 的例子细节讲清楚（重点讲解 Shuffle）？
- 5.对 Hadoop 有没有调优经验，没有什么使用心得？（调优从参数调优讲起）
- 6.MapReduce 出现单点负载多大，怎么负载平衡？（可以用 Partitioner）
- 7.MapReduce 怎么实现 Top10？

十九、来自 mo•mo•ring 提供的面试题 13 道：

xxxx 软件公司

- 1.你胜任该职位有什么优势
- 2.java 优势及原因（至少 3 个）
- 3.jvm 优化
- 4.写一个冒泡程序
- 5.hadoop 底层存储设计
- 6.职业规划

xxx 网络公司

编者 QQ: 1040195253



1.数据库

1.1 第一范式，第二范式和第三范式

1.2 给出两张数据表，优化表（具体字段不记得了，是关于商品定单和供应商方面的）

1.3 以你的实际经验，说下怎样预防全表扫描

2.网络七层协议

3.多线程

4.集合 HashTable 和 HashMap 区别

5.操作系统碎片

6.zookeeper 优点，用在什么场合

7.Hbase 中的 metastore 用来做什么的？

二十、来自 Clouds 提供的面试题 18 道：

1，在线安装 ssh 的命令以及文件解压的命令？

2，把公钥都追加到授权文件的命令？该命令是否在 root 用户下执行？

3，HadoopHA 集群中哥哥服务的启动和关闭的顺序？

4，HDFS 中的 block 块默认保存几份？默认大小多少？

5，NameNode 中的 meta 数据是存放在 NameNode 自身，还是 DataNode 等其他节点？

DatNode 节点自身是否有 Meta 数据存在？

6，下列那个程序通常与 NameNode 在一个节点启动？

7，下面那个程序负责 HDFS 数据存储？

8，在 HadoopHA 集群中 Zookeeper 的主要作用，以及启动和查看状态的命令？

9，HBase 在进行模型设计时重点在什么地方？一张表中国定义多少个 Column Family



最合适？为什么？

10，如何提高 HBase 客户端的读写性能？请举例说明。

11，基于 HadoopHA 集群记性 MapReduce 开发时，Configuration 如何设置 hbase.zookeeper.quorum 属性的值？

12，在 hadoop 开发过程中使用过哪些算法？其应用场景是什么？

13，MapReduce 程序如何发布？如果 MapReduce 中涉及到了第三方的 jar 包，该如何处理？

14，在实际工作中使用过哪些集群的运维工具，请分别阐述其作用。

15，hadoop 中 combiner 的作用？

16，IO 的原理，IO 模型有几种？

17，Windows 用什么样的模型，Linux 用什么样的模型？

18，一台机器如何应对那么多的请求访问，高并发到底怎么实现，一个请求怎么产生的，在服务端怎么处理的，最后怎么返回给用户的，整个的环节操作系统是怎么控制的？

二十一、来自****提供的面试题 11 道：

1.hdfs 的 client 端，复制到第三个副本时宕机，hdfs 怎么恢复保证下次写第三副本？block 信息是先写 dataNode 还是先写 nameNode？

2.快排现场写程序实现？

3.jvm 的内存是怎么分配原理？

4.毒酒问题---1000 桶酒，其中 1 桶有毒。而一旦吃了，毒性会在 1 周后发作。问最少需要多少只老鼠可在一周内找出毒酒？

5.用栈实现队列？



6.链表倒序实现？

7.多线程模型怎样（生产，消费者）？平时并发多线程都用哪些实现方式？

8.synchronized 是同步悲观锁吗？互斥？怎么写同步提高效率？

9.4 亿个数字，找出哪些重复的，要用最小的比较次数，写程序实现。

10.java 是传值还是传址？

11.java 处理多线程，另一线程一直等待？

二十二、来自****提供的面试题 18 道：

1.一个网络商城 1 天大概产生多少 G 的日志？

2.大概有多少条日志记录（在不清洗的情况下）？

3.日访问量大概有多少个？

4.注册数大概多少？

5.我们的日志是不是除了 apache 的访问日志是不是还有其他的日志？

6.假设我们有其他的日志是不是可以对这个日志有其他的业务分析？这些业务分析都有什么？

7、问：你们的服务器有多少台？

8、问：你们服务器的内存多大？

9、问：你们的服务器怎么分布的？（这里说地理位置分布，最好也从机架方面也谈谈）

10、问：你平常在公司都干些什么（一些建议）

下面是 HBASE 我非常不懂的地方：

11、hbase 怎么预分区？



12、hbase 怎么给 web 前台提供接口来访问（HBASE 可以提供对 HBASE 的访问，但是怎么查询同一条记录的多个版本数据）？

13、.htable API 有没有线程安全问题，在程序中是单例还是多例？

14、我们的 hbase 大概在公司业务中（主要是网上商城）大概都几个表，几个表簇，大概都存什么样的数据？

15、hbase 的并发问题？

下面的 Storm 的问题：

16、metaq 消息队列 zookeeper 集群 storm 集群（包括 zeromq,jzmq和 storm 本身）就可以完成对商城推荐系统功能吗？还有没有其他的中间件？

17、storm 怎么完成对单词的计数？（个人看完 storm 一直都认为他是流处理，好像没有积攒数据的能力，都是处理完之后直接分发给下一个组件）

18、storm 其他的一些面试经常问的问题？

二十三、飞哥(hadoop 月薪 13k)提供的面试题 18 道：

1、你们的集群规模？

开发集群：10 台（8 台可用）8 核 cpu

2、你们的数据是用什么导入到数据库的？导入到什么数据库？

处理之前的导入：通过 hadoop 命令导入到 hdfs 文件系统

处理完成之后的导出：利用 hive 处理完成之后的数据，通过 sqoop 导出到 mysql 数据库中，以供报表层使用。

3、你们业务数据量多大？有多少行数据？(面试了三家，都问这个问题)

开发时使用的是部分数据，不是全量数据，有将近一亿行（8、9 千万，具体不详，一般开



发中也没人会特别关心这个问题)

4、你们处理数据是直接读数据库的数据还是读文本数据？

将日志数据导入到 hdfs 之后进行处理

5、你们写 hive 的 hql 语句，大概有多少条？

不清楚，我自己写的时候也没有做过统计

6、你们提交的 job 任务大概有多少个？这些 job 执行完大概用多少时间？(面试了三家，都问这个问题)

没统计过，加上测试的，会与很多

7、hive 跟 hbase 的区别是？

8、你在项目中主要的工作任务是？

利用 hive 分析数据

9、你在项目中遇到了哪些难题，是怎么解决的？

某些任务执行时间过长，且失败率过高，检查日志后发现没有执行完就失败，原因出在 hadoop 的 job 的 timeout 过短（相对于集群的能力来说），设置长一点即可

10、你自己写过 udf 函数么？写了哪些？

这个我没有写过

11、你的项目提交到 job 的时候数据量有多大？(面试了三家，都问这个问题)

不清楚是要问什么

12、reduce 后输出的数据量有多大？

13、一个网络商城 1 天大概产生多少 G 的日志？ 4tb

14、大概有多少条日志记录（在不清洗的情况下）？ 7-8 百万条

15、日访问量大概有多少个？百万



16、注册数大概多少？不清楚 几十万吧

17、我们的日志是不是除了 apache 的访问日志是不是还有其他的日志？关注信息

18、假设我们有其他的日志是不是可以对这个日志有其他的业务分析？这些业务分析都有什么？

二十四、来自 aboutyun 提供的面试题 1 道：

有一千万条短信，有重复，以文本文件的形式保存，一行一条，有重复。

请用 5 分钟时间，找出重复出现最多的前 10 条。

分析：

常规方法是先排序，在遍历一次，找出重复最多的前 10 条。但是排序的算法复杂度最低为 $n\lg n$ 。

可以设计一个 `hash_table`, `hash_map<string, int>`，依次读取一千万条短信，加载到 `hash_table` 表中，并且统计重复的次数，与此同时维护一张最多 10 条的短信表。

这样遍历一次就能找出最多的前 10 条，算法复杂度为 $O(n)$ 。

二十五、北京-南桑（hadoop 月薪 12k）提供的面试题 5 道：

1、job 的运行流程(提交一个 job 的流程)？

2、Hadoop 生态圈中各种框架的运用场景？

3、还有很多的选择題

4、面试问到的

hive 中的压缩格式 `RCFile`、`TextFile`、`SequenceFile` 各有什么区别？

以上 3 种格式一样大的文件哪个占用空间大小..等等



还有 Hadoop 中的一个 HA 压缩。

5、假如：Flume 收集到的数据很多个小文件,我需要写 MR 处理时将这些文件合并

(是在 MR 中进行优化,不让一个小文件一个 MapReduce)

他们公司主要做的是中国电信的流量计费为主,专门写 MR。

二十六、来自炎帝初始化提供的面试题 2 道：

以下题目不必都做完，挑最擅长的即可。

题一：RTB 广告 DSP 算法大赛

请按照大赛的要求进行相应的建模和分析，并详细记录整个分析处理过程及各步骤成果物。

算法大赛主页：<http://contest.ipinyou.com/cn/index.shtml>

算法大赛数据下载地址：

<http://pan.baidu.com/share/link?shareid=1069189720&uk=3090262723#dir>

题二：cookieID 识别

我们有 M 个用户 N 天的的上网日志：详见 58.sample

字段结构如下：

| | | |
|------------|--------|------------|
| ip | string | 客户端 IP |
| ad_id | string | 宽带 ADSL 账号 |
| time_stamp | string | 上网开始时间 |
| url | string | URL |
| ref | string | referer |
| ua | string | User Agent |
| dest_ip | string | 目标 IP |
| cookie | string | cookie |
| day_id | string | 日期 |

58.com 的 cookie 值如：

```
bangbigtip2=1; bdshare_firsttime=1374654651270;
CNZZDATA30017898=cnzz_eid%3D2077433986-1374654656-http%253A%252F%252Fsh.58.com
%26ntime%3D1400928250%26cnzz_a%3D0%26ltime%3D1400928244483%26rtime%3D63;
Hm_lvt_f5127c6793d40d199f68042b8a63e725=1395547468,1395547513,1395758399,13957594
68; id58=05dvZ1HvkL0TNy7GBv7gAg==;
Hm_lvt_3bb04d7a4ca3846dcc66a99c3e861511=1385294705;
__utma=253535702.2042339925.1400424865.1400424865.1400928244.2;
__utmz=253535702.1400424865.1.1.utmcsr=(direct)|utmccn=(direct)|utmcmd=(none); city=sh;
pup_bubble=1; __ag_cm_=1400424864286; myfeet_tooltip=end; ipcity=sh%7C%u4E0A%u6D77
```



其中有一个属性能标识一个用户，我们称之为 cookieID。

请根据样例数据分析出 58.com 的 cookieID。

要求详细描述分析过程。



二十七、来自 aboutyun 提供的面试题 7 道：

- 1、解释 “hadoop” 和 “hadoop 生态系统” 两个概念。
- 2、说明 Hadoop 2.0 的基本构成。
- 3、相比于 HDFS1.0, HDFS 2.0 最主要的改进在哪几方面？
- 4、试使用 “步骤 1，步骤 2，步骤 3.....” 说明 YARN 中运行应用程序的基本流程。
- 5、“MapReduce 2.0” 与 “YARN” 是否等同，尝试解释说明。
- 6、MapReduce 2.0 中，MRAppMaster 主要作用是什么，MRAppMaster 如何实现任务容错的？
- 7、为什么会产生 yarn,它解决了什么问题，有什么优势？

二十八、来自 然月枕流君 提供的面试题 6 道：

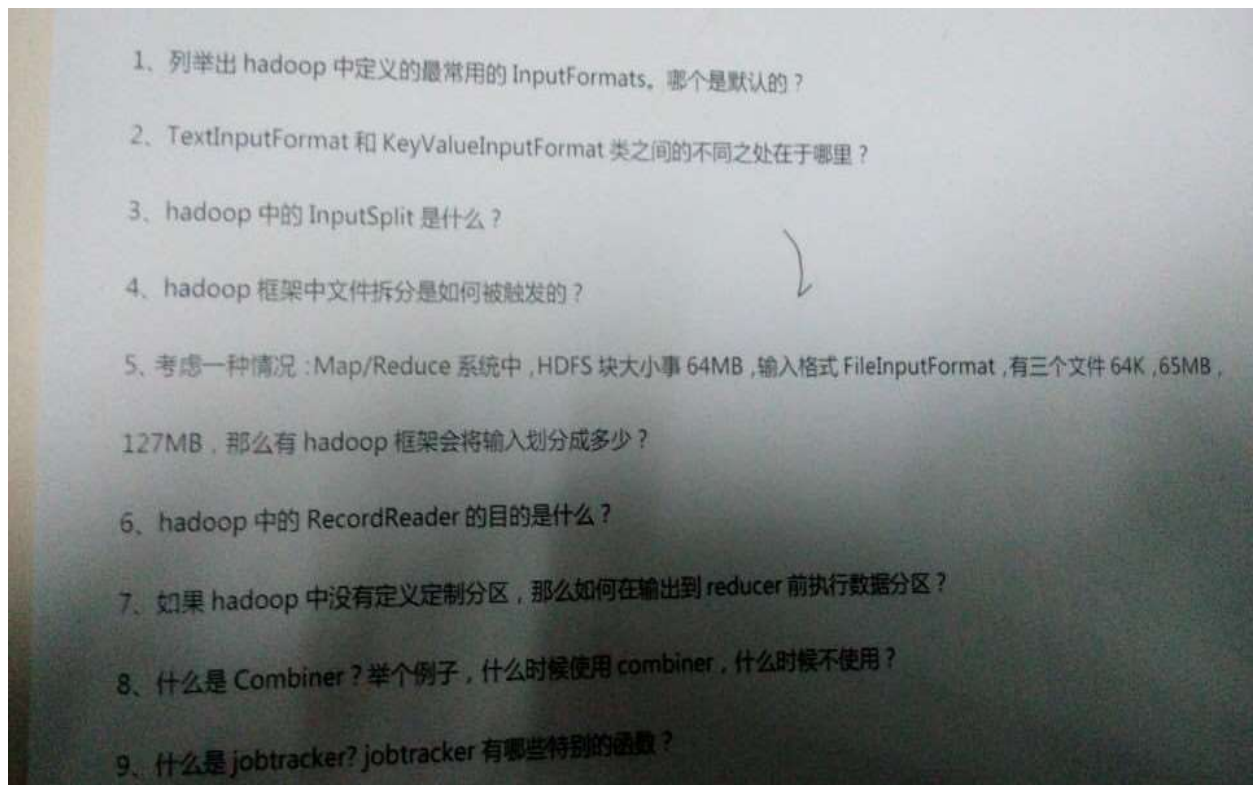
- 1、集群多少台,数据量多大,吞吐量是多大,每天处理多少 G 的数据？
- 2、自动化运维了解过吗,你们是否是自动化运维管理？
- 3、数据备份,你们是多少份,如果数据超过存储容量,你们怎么处理？
- 4、怎么提升多个 JOB 同时执行带来的压力,如何优化,说说思路？
- 5、你们用 HBASE 存储什么数据？
- 6、你们的 hive 处理数据能达到的指标是多少？

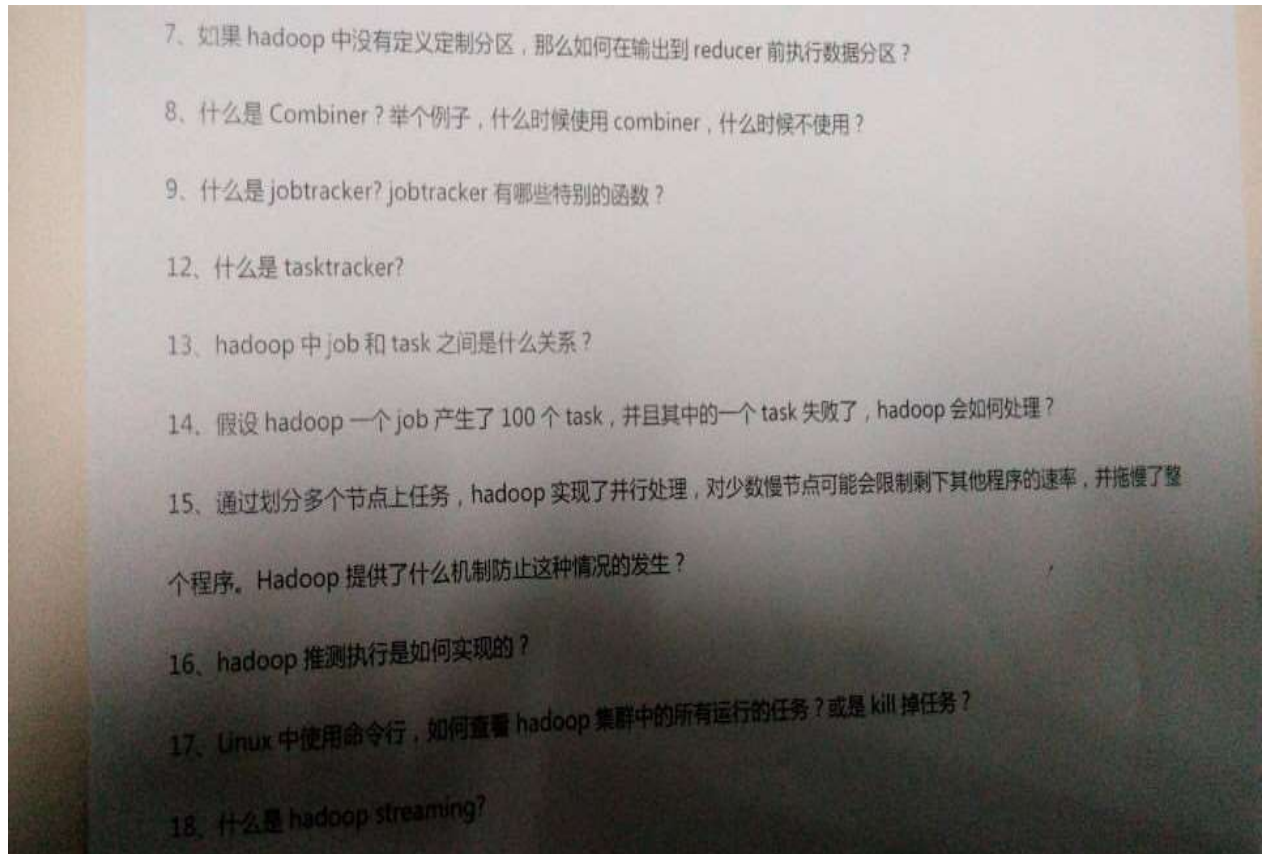


二十九、来自夏天提供的面试题 1 道：

1、请说说 hadoop1 的 HA 如何实现？

三十、来自枫林木雨提供的面试题 18 道：





备注：想知道更多学员的面试经验，欢迎关注官网：www.crxy.cn。

- 1、10 年工作经验罗同学由月薪 18k 飙升至 45k: <http://www.crxy.cn/detail/news/12>；
- 2、14 届应届本科生入职中国航天集团，基本年薪 20w，年终奖 10w：
<http://www.crxy.cn/detail/news/12>；
- 3、入职腾讯学员告诉你 hadoop 学习方法：<http://www.crxy.cn/detail/jobinfo/10>；
- 4、揭秘大专生月薪 6.5k 翻番至 13k：<http://www.crxy.cn/detail/jobinfo/8>；
- 5、美女研究生 hadoop 工作经验分享：<http://www.crxy.cn/detail/jobinfo/6>。