超人学院 hadoop 面试葵花宝典(V1.1)

www.crxy.cn

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（包含但不限于）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 2015 年 1 月 21 日为止，本次共收集了将近 500 道 hadoop 相关的面试题。随着面试人数的增加，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，确保第一时间获取免费的公开课信息和其他学习资料。

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，各有哪些优缺点。

9. Hive 有哪些方式保存元数据的，各有哪些特点。

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.sh

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 schedular :默认，先进先出的原则

Capacity schedular :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 schedular: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：

二、来自彩虹伴相思雨提供的面试题（31 道）：

15、Hive 内部表和外部表的区别？

16、Hbase 的 rowkey 怎么创建比较好？列族怎么创建比较好？

17、用 mapreduce 怎么处理数据倾斜问题？

18、hadoop 框架中怎么来优化？

19、Hbase 内部是什么机制？

20、我们在开发分布式计算 job 的,是否可以去掉 reduce()阶段？

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、reduce 后输出的数据量有多大？

26、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）：

27、datanode 在什么情况下不会备份？

28、combine 出现在那个过程？

29、hdfs 得体系结构？

30、flush 的过程？

31、什么是队列？

32、List 与 Set 的区别？

33、数据库的三大范式？

34、三个 datanode，当有一个 datanode 出现错误会怎样？

35、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）：

36、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

Linux 脚本能力考察：

37、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

38、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:

...

Hadoop 基础知识与问题分析的能力：

39、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

40、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？

41、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）

42、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

43、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？

44、Java 开发能力

45、 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

46、 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）：

47、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

48. 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。

49、 MapReduce 大致流程？

50、 combiner, partition 作用？

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

52、用 mapreduce 如何实现两张表连接，有哪些方法？

53、知道 MapReduce 大致流程，map, shuffle, reduce

54、知道 combiner, partition 作用，设置 compression

55、搭建 hadoop 集群，master/slave 都运行那些服务

56、HDFS，replica 如何定位

57、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？

58、Hadoop 参数调优，cluster level: JVM, map/reduce slots, job level: reducer #,memory, use combiner? use compression?

59、pig latin, Hive 语法有什么不同

60、描述 HBase, zookeeper 搭建过程

61、hadoop 运行的原理?

62、mapreduce 的原理?

63、HDFS 存储的机制?

64、举一个简单的例子说明 mapreduce 是怎么来运行的 ？

65、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。

66、hadoop 中 Combiner 的作用?

67、如何确认 Hadoop 集群的健康状况。

六、来自****提供的 hadoop 面试题（9 道）：

68、使用的 hadoop 版本都是什么？

69、mapreduce 原理是什么？

70、mapreduce 作业，不使用 reduce 来输出，用什么能代替 reduce 的功能

71、hive 如何调优？

72、hive 如何权限控制？

74、 hbase 写数据的原理是什么？

74、hive 能像关系数据库那样，建多个库吗？

75、hbase 宕机如何处理？

76、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

77、 下面哪个程序负责 HDFS 数据存储。

a)NameNode b)Jobtracker c)Datanode d)secondaryNameNode e)tasktracker

78、HDfS 中的 block 默认保存几份？

a)3 份 b)2 份 c)1 份 d)不确定

79、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

80、Hadoop 作者

a)Martin Fowler b)Kent Beck c)Doug cutting

81、HDFS 默认 Block Size

a)32MB b)64MB c)128MB

82、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

83、 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：

84、下列哪项可以作为集群的管理工具

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

85、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

86、 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

87、下列哪个是 Hadoop 运行的模式

a)单机版 b)伪分布式 c)分布式

88、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：

89、Ganglia 不仅可以进行监控，也可以进行告警。（ ）

90、Block Size 是不可以修改的。（ ）

91、Nagios 不可以监控 Hadoop 集群，因为它不提供 Hadoop 支持。（ ）

92、如果 NameNode 意外终止，SecondaryNameNode 会接替它使集群继续工作。（ ）

93、Cloudera CDH 是需要付费使用的。（ ）

94、Hadoop 是 Java 开发的，所以 MapReduce 只支持 Java 语言编写。（ ）

95、Hadoop 支持数据的随机读写。（ ）

96、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。（ ）

97、NameNode 本地磁盘保存了 Block 的位置信息。（ ）

98、DataNode 通过长连接与 NameNode 保持通信。（ ）

99、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。（ ）

100、 Slave 节点要存储数据，所以它的磁盘越大越好。（ ）

101、hadoop dfsadmin –report 命令用于检测 HDFS 损坏块。（ ）

102、Hadoop 默认调度器策略为 FIFO（ ）

103、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。（ ）

104、因为 HDFS 有多个副本，所以 NameNode 是不存在单点问题的。（ ）

105、每个 map 槽就是一个线程。（ ）

106、Mapreduce 的 input split 就是一个 block。（ ）

107、NameNode 的 Web UI 端口是 50030，它通过 jetty 启动的 Web 服务。（ ）

108、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。（ ）

109、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）：

110、hive 实现统计的查询语句是什么？

111、生产环境中为什么建议使用外部表？

112、hadoop mapreduce 创建类 DataWritable 的作用是什么？

113、为什么创建类 DataWritable？

114、如何实现统计手机流量？

115、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：

116、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

117、hadoop 运转的原理?

118、mapreduce 的原理?

119、HDFS 存储的机制?

120、举一个简略的比方阐明 mapreduce 是怎么来运转的 ?

121、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。

122、hadoop 中 Combiner 的效果?

论坛某网友的回复:

117.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。

118.mapreduce 即是将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。

119.java 程序和 namenode 合作,把数据存放在不一样的数据节点上

120.怎么运转用图来表明最好了。图无法画。谷歌下

121.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 limie 1000000;

122. combiner 也是一个 reduce,它可以削减 map 到 reudce 的数据传输,进步 shuff 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、来自小萝卜的笔试和面试题：（11 道）

一、笔试

123、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。

124、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

125、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、来自闪客、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）

126、hadoop 集群搭建过程，写出步骤。

127、hadoop 集群运行过程中启动那些线程，各自的作用是什么？

128、/tmp/hadoop-root/dfs/name　　the path is not exists or is not accessable.

NameNode main 中报错，该怎么解决。（大意这样 一个什么异常）

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

130、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

131、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。

132、hadoop 的调度器有哪些，工作原理。

133、mapreduce 的 join 方法有哪些？

134、Hive 元数据保存的方法有哪些，各有什么特点？

135、java 实现非递归二分法算法。

136、mapreduce 中 Combiner 和 Partition 的作用。

137、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。

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

139、如何实现 mapreduce 的二次排序。

大数遇到的面试题：

140、面试官上来就问 hadoop 的调度机制；

141、机架感知；

142、MR 数据倾斜原因和解决方案；

143、集群 HA。

@找自己 提供的面试题：

144、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里全部录入完毕，

不残留数据。

145、对于 hive，你写过哪些 UDF 函数，作用是什么

146、hdfs 的数据压缩算法

147、mapreduce 的调度模式

148、hive 底层与数据库交互原理

149、hbase 过滤器实现原则

150、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

151、请问下，直接将时间戳作为行健，在写入单个 region 时候会发生热点问题，为什么

呢？

十三、来自飞哥提供的面试题（17 道）：

152、hdfs 原理，以及各个模块的职责？

153、mapreduce 的工作原理？

154、map 方法是如何调用 reduce 方法的？

155、shell 如何判断文件是否存在，如果不存在该如何处理？

156、fsimage 和 edits 的区别？

157、hadoop1 和 hadoop2 的区别？

笔试：

158、hdfs 中的 block 默认报错几份？

159、哪个程序通常与 namenode 在一个节点启动？并做分析

160、列举几个配置文件优化？

161、写出你对 zookeeper 的理解

162、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？

163、谈谈数据倾斜，如何发生的，并给出优化方案

164、介绍一下 hbase 过滤器

165、mapreduce 基本执行过程

166、谈谈 hadoop1 和 hadoop2 的区别

167、hbase 集群安装注意事项

168、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、来自飞哥提供的面试题（3 道）：

169、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。

170、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。

171、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

172、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；

173、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‘logK。

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’ *O（logK）,（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 个元素的最小推来对出现频率进行排序。

174、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 hash(x)%5000，然后按照该值存到 5000 个小文件（记为 x0,x1,...x4999）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 个结点的最小堆），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。

175、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 hash(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 hash_map(query, 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_cout 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。

176、 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×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 hash(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 a0,a1,...,a999）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 b0,b1,...,b999）。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。

177、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 bit=1 GB 内存，还可以接受。然后扫描这 2.5 亿个整数，查看 Bitmap 中相对应位，如果是 00 变 01，01 变 10，10 保持不变。所描完事后，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。

178、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(logn)，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。

179、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）。

180、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。

181、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*lg10)。所以总的时间复杂度，是 O(n*le)与 O(n*lg10)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*lg100)。

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/yanxionglu/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=(ln2)*(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>=nlg(1/E)*lge 大概就是 nlg(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，h1 和 h2。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这个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，在将该区域分成 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）：

182.说说值对象与引用对象的区别？

183.谈谈你对反射机制的理解及其用途？

184.ArrayList、Vector、LinkedList 的区别及其优缺点？HashMap、HashTable 的区别及其优缺点？

185.列出线程的实现方式？如何实现同步？

186.sql 题,是一个图表，具体忘了

187、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？

188、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）：

189、

用户手机号　出现的地点　出现的时间　　逗留的时间

111111111 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

111111111 ~~~~~~~~

333333333 ~~~~~~~

十八、来自象夫提供的面试题（7 道）：

190、文件大小默认为 64M，改为 128M 有啥影响？

191、RPC 原理？

192、NameNode 与 SecondaryNameNode 的区别与联系？

193、介绍 MadpReduce 整个过程，比如把 WordCount 的例子的细节将清楚（重点讲解 Shuffle）？

194、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）

195、MapReduce 出现单点负载多大，怎么负载平衡？（可以用 Partitioner）

196、MapReduce 怎么实现 Top10？

十九、来自 mo•mo•ring 提供的面试题（13 道）：

197、你胜任该职位有什么优势

198、java 优势及原因（至少 3 个）

199、jvm 优化

200、写一个冒泡程序

201、hadoop 底层存储设计

202、说说你最近 2~3 年的职业规划

一.数据库

203、第一范式，第二范式和第三范式

204、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）

205、以你的实际经验，说下怎样预防全表扫描

二、网络七层协议

206、多线程

207、集合 HashTable 和 HashMap 区别

208、操作系统碎片

209、zookeeper 优点，用在什么场合

310、Hbase 中的 metastore 用来做什么的？

二十、来自 Clouds 提供的面试题（18 道）：

311、在线安装 ssh 的命令以及文件解压的命令？

312、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？

313、HadoopHA 集群中，各个服务的启动和关闭的顺序？

314、HDFS 中的 block 块默认保存几份？默认大小多少？

315、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？

316、下列那个程序通常与 NameNode 在一个节点启动？

317、下面那个程序负责 HDFS 数据存储？

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

319、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？

320、如何提高 HBase 客户端的读写性能？请举例说明。

321、基于 HadoopHA 集群进行 MapReduce 开发时，Configuration 如何设置 hbase.zookeeper,quorum 属性的值？

322、 在 hadoop 开发过程中使用过哪些算法？其应用场景是什么？

323、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？

324、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。

325、hadoop 中 combiner 的作用?

326、IO 的原理，IO 模型有几种?

327、Windows 用什么样的模型，Linux 用什么样的模型？

328、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）：

329、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

块信息是先写 dataNode 还是先写 nameNode?

330、快排现场写程序实现？

331、jvm 的内存是怎么分配原理？

332、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？

333、用栈实现队列？

334、链表倒序实现？

335、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？

336、synchonized 是同步悲观锁吗？互斥？怎么写同步提高效率？

337、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。

338、java 是传值还是传址？

339、 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）：

340、一个网络商城 1 天大概产生多少 G 的日志？

341、大概有多少条日志记录（在不清洗的情况下）？

342、日访问量大概有多少个？

343、注册数大概多少？

344、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？

345、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？

346、问：你们的服务器有多少台？

347、问：你们服务器的内存多大？

348、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）

349、问：你平常在公司都干些什么（一些建议）

350、hbase 怎么预分区？

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

352、.htable API 有没有线程安全问题，在程序中是单例还是多例？

353、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？

354、hbase 的并发问题？

Storm 问题：

355、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？

356、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）

357、storm 其他的一些面试经常问的问题？

二十三、来自飞哥提供的面试题（18 道）：

358、你们的集群规模？

开发集群：10 台（8 台可用）8 核 cpu

359、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。

360、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）

361、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

362、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

363、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

364、hive 跟 hbase 的区别是？

365、你在项目中主要的工作任务是？

利用 hive 分析数据

366、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

367、你自己写过 udf 函数么？写了哪些？

这个我没有写过

368、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

369、reduce 后输出的数据量有多大？

370、一个网络商城 1 天大概产生多少 G 的日志？ 4tb

371、大概有多少条日志记录（在不清洗的情况下）？ 7-8 百万条

372、日访问量大概有多少个？百万

373、注册数大概多少？不清楚　几十万吧

374、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

375、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)：

376、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 nlgn。

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、来自北京-南桑提供的面试题（5 道）：

377、job 的运行流程(提交一个 job 的流程)？

378、Hadoop 生态圈中各种框架的运用场景？

379、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.还有 Hadoop 中的一个 HA 压缩。

380、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。

381：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

382：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_firstime=1374654651270;
CNZZDATA30017898=cnzz_eid%3D2077433986-1374654656-http%253A%252F%252Fsh.58.com%26ntime%3D1400928250%26cnzz_a%3D0%26ltime%3D1400928244483%26rtime%3D63;
Hm_lvt_f5127c6793d40d199f68042b8a63e725=1395547468,1395547513,1395758399,1395759468; 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）：

383、 解释“hadoop”和“hadoop 生态系统”两个概念。

384、 说明 Hadoop 2.0 的基本构成。

385、 相比于 HDFS1.0, HDFS 2.0 最主要的改进在哪几方面？

386、 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。

387、 “MapReduce 2.0”与“YARN”是否等同，尝试解释说明。

388、MapReduce 2.0 中，MRAppMaster 主要作用是什么，MRAppMaster 如何实现任务容错的？

389、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）：

390、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？

391、自动化运维了解过吗,你们是否是自动化运维管理？

392、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？

393、怎么提升多个 JOB 同时执行带来的压力,如何优化,说说思路？

394、你们用 HBASE 存储什么数据？

395、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：

396、请说说 hadoop1 的 HA 如何实现？

三十、来自李同学的面试总结

397、Hadoop 体系结构（HDFS 与 MapReduce 的体系结构）、Hadoop 相比传统数据存储方式（比如 mysql）的优势？

398、Hadoop 集群的搭建步骤、Hadoop 集群搭建过程中碰到了哪些常见问题（比如 datanode 没有起来）、Hadoop 集群管理（如何动态增加和卸载节点、safe mode 是什么、常用的命令 kill 等）？

399、HDFS 的 namenode 与 secondarynamenode 的工作原理（重点是日志拉取和合并过程）、hadoop 1.x 的 HDFS 的 HA 方案（namenode 挂掉的情况如何处理、datanode 挂掉的情况如何处理）？

400、HDFS 的常用 shell 命令有哪些？分别对应哪些 Client Java API？：显示文件列表、创建目录、文件上传与下载、文件内容查看、删除文件

401、HDFS 的文件上传与下载底层工作原理（或 HDFS 部分源码分析）：FileSystem 的 create()和 open()方法源码分析？

402、MapReduce 计算模型、MapReduce 基础知识点（MapReduce 新旧 API 的使用、在 linux 命令行运行 MapReduce 程序、自定义 Hadoop 数据类型）？

403、MapReduce 执行流程："天龙八步"，计数器、自定义分区、自定义排序、自定义分组、如何对 value 进行排序：次排序+自定义分组、归约？

404、MapReduce 的 shuffle 工作原理、MapReduce 工作原理（MapReduce 源码、InputStream 源码、waitForCompletion()源码）、jobtracker 如何创建 map 任务和 reduce 任务是面试的重点。

405、MapReduce 进阶知识：Hadoop 的几种文件格式、常见输入输出格式化类、多输入多输出机制、MapReduce 的常见算法（各种 join 原理和优缺点、次排序和总排序）？

406、MapReduce 性能优化（shuffle 调优、压缩算法、更换调度器、设置 InputSplit 大小减少 map 任务数量、map 和 reduce 的 slot 如何设置、数据倾斜原理和如何解决）？

407、HBase 的体系结构和搭建步骤、shell 命令与 Java API、HBase 作为 MapReduce

的输入输出源、高级 Java API、工作原理（重点是 combine 和 split 原理）、行键设计原则、性能优化？

408、Hive 的工作原理、两种元数据存放方式、几种表之间的区别、数据导入的几种方式、几种文件格式、UDF 函数、性能调优（重点是 join 的时候如何放置大小表）？

409、Zookeeper、Flume、Pig、Sqoop 的基本概念和使用方式，ZooKeeper 被问到过其如何维护高可用（如果某个节点挂掉了它的处理机制）？

410、Hadoop2：体系结构、HDFS HA、YARN？

三十一、来自刘同学的面试题

411、关系型数据库和非关系型数据库的区别？

提示：

关系型数据库通过外键关联来建立表与表之间的关系，非关系型数据库通常指数据以对象的形式存储在数据库中，而对象之间的关系通过每个对象自身的属性来决定。

对数据库高并发读写、高可扩展性和高可用性的需求，对海量数据的高效率存储和访问的需求，存储的结构不一样，非关系数据库是列式存储，在存储结构上更加自由。

412、hive 的两张表关联，使用 mapreduce 是怎么写的？

提示：打标记笛卡尔乘积

413、hive 相对于 Oracle 来说有那些优点？

提示：

hive 是数据仓库，oracle 是数据库，hive 能够存储海量数据，hive 还有更重要的作用就是数据分析，最主要的是免费。

414、现在我们要对 Oracle 和 HBase 中的某些表进行更新，你是怎么操作？

提示：

disable '表名'

alter '表明', NAME => '列名', VERSIONS =>3

enable '表名'

415、HBase 接收数据，如果短时间导入数量过多的话就会被锁，该怎么办？ 集群数 16 台 ，高可用性的环境。

参考：

通过调用 HTable.setAutoFlush(false)方法可以将 HTable 写客户端的自动 flush 关闭，这样可以批量写入数据到 HBase，而不是有一条 put 就执行一次更新，只有当 put 填满客户端写缓存时，才实际向 HBase 服务端发起写请求。默认情况下 auto flush 是开启的。

416、 说说你们做的 hadoop 项目流程？

417、你们公司的服务器架构是怎么样的（分别说下 web 跟 hadoop）？

418、假如有 1000W 用户同时访问同一个页面，怎么处理？

提示：优化代码、静态化页面、增加缓存机制、数据库集群、库表散列。。。

419、怎样将 mysql 的数据导入到 hbase 中？ 不能使用 sqoop，速度太慢了

提示：

A、一种可以加快批量写入速度的方法是通过预先创建一些空的 regions，这样当数据写入 HBase 时，会按照 region 分区情况，在集群内做数据的负载均衡。

B、hbase 里面有这样一个 hfileoutputformat 类，他的实现可以将数据转换成 hfile 格式，通过 new 一个这个类，进行相关配置,这样会在 hdfs 下面产生一个文件，这个时候利用 hbase 提供的 jruby 的 loadtable.rb 脚本就可以进行批量导入。

420、在 hadoop 组中你主要负责那部分？

提示：负责编写 mapreduce 程序，各个部分都要参加

421、怎么知道 hbase 表里哪些做索引？哪些没做索引？

提示：

有且仅有一个：rowkey，所以 hbase 的快速查找建立在 rowkey 的基础的，而不能像一般的关系型数据库那样建立多个索引来达到多条件查找的效果。

422、 hdfs 的原理以及各个模块的职责

423、 mapreduce 的工作原理

424、 map 方法是如何调用 reduce 方法的

425、fsimage 和 edit 的区别？

提示：fsimage：是存储元数据的镜像文件，而 edit 只是保存的操作日志。

426、hadoop1 和 hadoop2 的区别？

提示：

（1） hdfs 的 namenode 和 mapreduce 的 jobtracker 都是单点。

（2） namenode 所在的服务器的内存不够用时，那么集群就不能工作了。

（3）mapreduce 集群的资源利用率比较低。

单 NN 的架构使得 HDFS 在集群扩展性和性能上都有潜在的问题，在集群规模变大后，NN 成为了性能的瓶颈。Hadoop 2.0 里的 HDFS Federation 就是为了解决这两个问题而开发的。扩大 NN 容量，共享 DN 数据，且方便客户端访问。

427、hdfs 中的 block 默认报错几份？

提示：3 份

428、哪个程序通常与 nn 在一个节点启动？并做分析

提示：jobtrack，将两者放在一起，减少网络访问，IO 访问的时间，提高了效率。

429、列举几个配置文件优化？

提示：

http://blog.csdn.net/wangqiaoshi/article/details/18142841

http://www.3lian.com/edu/2013/01-15/53867.html

430、写出你对 zookeeper 的理解

提示：大部分分布式应用需要一个主控、协调器或控制器来管理物理分布的子进程（如资源、任务分配等）。目前，大部分应用需要开发私有的协调程序，缺乏一个通用的机制协调程序的反复编写浪费，且难以形成通用、伸缩性好的协调器。

ZooKeeper：提供通用的分布式锁服务，用以协调分布式应用。

431、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？

提示：

这样处理是不合理的，因为那么 namenode 格式化操作，是对文件系统进行格式化，namenode 格式化时清空 dfs/name 下空两个目录下的所有文件，之后，会在目录 dfs.name.dir 下创建文件。

文本不兼容，有可能时 namenode 与 datanode 的 数据里的 namespaceID、clusterID 不一致，找到两个 ID 位置，修改为一样即可解决。

432、谈谈数据倾斜，如何发生的，并给出优化方案。

原因：

（1）key 分布不均匀

（2）业务数据本身的特性

（3）建表时考虑不周

（4）某些 SQL 语句本身就有数据倾斜

map 处理数据量的差异取决于上一个 stage 的 reduce 输出，所以如何将数据均匀的分配到各个 reduce 中，就是解决数据倾斜的根本所在。

优化：参数调节；

SQL 语句调节：http://blog.sina.com.cn/s/blog_9402246001013kxf.html

433、介绍一下 HBase 过滤器

参考：http://blog.sina.com.cn/s/blog_ae33b83901017km4.html

434、 mapreduce 基本执行过程

435、 谈谈 hadoop1 和 hadoop2 的区别

436、谈谈 HBase 集群安装注意事项？

提示：

需要注意的地方是 ZooKeeper 的配置。这与 hbase-env.sh 文件相关，文件中 HBASE_MANAGES_ZK 环境变量用来设置是使用 hbase 默认自带的 Zookeeper 还是使用独立的 ZooKeeper。HBASE_MANAGES_ZK=false 时使用独立的，为 true 时使用默认自带的。

某个节点的 HRegionServer 启动失败 ,这是由于这 3 个节点的系统时间不一致相差超过集群的检查时间 30s。

判断题：

437、Ganglia 不仅可以进行监控，也可以进行告警。（ ）

438、Nagios 不可以监控 Hadoop 集群，因为它不提供 Hadoop 支持。（ ）

439、 如果 NameNode 意外终止，SecondaryNameNode 会接替它使集群继续工作。（ ）

440、 Cloudera CDH 是需要付费使用的。( )

441、 Hadoop 是 Java 开发的，所以 MapReduce 只支持 Java 语言编写。( )

442、 Hadoop 支持数据的随机写。( )

443、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。( )

444、NameNode 本地磁盘保存了 Block 的位置信息。( )

445、Slave 节点要存储数据，所以它的磁盘越大越好。( )

446、Hadoop 默认调度器策略为 FIFO，并支持多个 Pool 提交 Job。( )

447、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。( )

448、因为 HDFS 有多个副本，所以 NameNode 是不存在单点问题的。( )

449、每个 map 槽就是一个线程。( )

450、Mapreduce 的 input split 就是一个 block。( )

451、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 200MB。( )

452、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。( )

453、Hadoop1.0 和 2.0 都具备完善的 HDFS HA 策略。( )

454、GZIP 压缩算法比 LZO 更快。( )

455、PIG 是脚本语言，它与 mapreduce 无关。( )

456、例举 hadoop 中定义的最常用的 InputFormats，哪个是默认的？

提 示 ： DBInputFormat 、 FileInputFormat （ KeyValueTextInputFormat 、 NlineInputFormat、TextInputFormat)

默认：TextInputFormat

457、TextInputFormat 和 KeyValueInputFormat 类之间的不同之处在于哪里？

提示：TextInputFormat 中的 key 表示行的偏移量，value 是行文本内容

KeyValueInputFormat 的 key value 是通过第一个制表符进行划分的

458、hadoop 中的 InputSplit 是什么？

提示：每一个 map 任务单独处理的数据单位，可以决定单个 mapper 任务处理的大小。

默认大小与 block 一样大

459、hadoop 框架中文件拆分时如何被触发的？（block 是怎么触发）

提示：客户端上传文件时为从 namenode 申请的 ID 和位置

460、hadoop 中 RecordReader 的目的是什么？

提示：将 InputSplit 的数据解析成键值对

461、如果 hadoop 中没有定义定制分区，那么如何在输出到 reducer 前执行数据分区？

462、什么是 Combiner？举个例子，什么时候使用 combiner，什么时候不使用？

463、什么是 jobtracker？jobtracker 有哪些特别的函数

提示：jobtracker 负责接收用户提交的作业，负责启动、跟踪任务执行。是一个 rpc 服务端

jobtracker 有哪些特别的函数：JobSubmitProtocol

464、什么是 tasktracker？

提示：mr 的客户端，接收 jobtracker 的发出的指令，用来执行任务的

465、hadoop 中 job 和 task 之间是什么关系？

提示：执行一次 mr 程序就是一个 job，job 再执行时会划分 maptask，reducetask。task 是 job 运行作业的一个重要组成部分。

466、假设 hadoop 一个 job 产生 100 个 task，并且其中的一个 task 失败了，hadoop 会怎样处理？

提示：hadoop 容错机制，当一个任务执行失败，jobTracker 发送命令重新执行，如果重新执行四次也不行，任务执行失败

（mapred-site.xml 配置文件里<name>mapred.max.attemp<name>

<value>4</value>）

467、hadoop 推测执行时如何实现的？

468、Linux 中使用命令行 如和查看 hadoop 集群中所有运行的任务？或是 kill 掉任务？

提示：hadoop job -list

hadoop job -kiil 进程名

469、什么 hadoop streaming？

提示：指的是用其它语言处理

三十二、来自李同学提供的面试题：

470、谈谈整个 hadoop 的生态圈

471、Hadoop 集群的优化

472、 Hive 优化

473、 Hadoop1 与 hadoop2 有何异同点

474、有关项目细节，产生的日志量多大，集群规模，处理数据要用多久

三十三、来自何同学提供的面试题：

475、Hive 的 sort by 和 order by 的区别

http://www.cnblogs.com/ggjucheng/archive/2013/01/03/2843243.html

476、Hive 里面用什么代替 in 查询

提示：Hive 中的 left semi join 替换 sql 中的 in 操作

477、简述 HBase 的瓶颈

提示：HBase 的瓶颈就是硬盘传输速度。HBase 的操作，它可以往数据里面 insert,也可以 update 一些数据，但 update 的实际上也是 insert，只是插入一个新的时间戳的一行。Delete 数据，也是 insert，只是 insert 一行带有 delete 标记的一行。Hbase 的所有操作都是追加插入操作。Hbase 是一种日志集数据库。它的存储方式，像是日志文件一样。它是批量大量的往硬盘中写，通常都是以文件形式的读写。这个读写速度，就取决于硬盘与机器之间的传输有多快。而 Oracle 的瓶颈是硬盘寻道时间。它经常的操作时随机读写。要 update 一个数据，先要在硬盘中找到这个 block，然后把它读入内存，在内存中的缓存中修改，过段时间再回写回去。由于你寻找的 block 不同，这就存在一个随机的读。硬盘的寻道时间主要由转速来决定的。而寻道时间，技术基本没有改变，这就形成了寻道时间瓶颈。

478、Jvm 的 4 个引用

提示：强引用就是不会被 GC 回收 软引用在 jvm 报告内存不足时候才会被回收 弱引用于软引用相似 虚引用是跟踪对象被 GC 回收的状态。

479、 mapreduce 的一些配置参数优化

480、ArrayList、Vector、LinkedList 的区别及其优缺点？HashMap、HashTable 的区别及优缺点？

提示：

ArrayList 和 Vector 是采用数组方式存储数据的,是根据索引来访问元素的 ,都可以根据需要自动扩展内部数据长度，以便增加和插入元素，都允许直接序号索引元素，但是插入数据要涉及到数组元素移动等内存操作，所以索引数据快插入数据慢，他们最大的区别就是 synchronized 同步的使用。

LinkedList 使用双向链表实现存储，按序号索引数据需要进行向前或向后遍历，但是插入数据时只需要记录本项的前后项即可，所以插入数度较快！

如果只是查找特定位置的元素或只在集合的末端增加、移除元素，那么使用 Vector 或 ArrayList 都可以。如果是对其它指定位置的插入、删除操作，最好选择 LinkedList

HashMap、HashTable 的区别及其优缺点：

HashTable 中的方法是同步的 HashMap 的方法在缺省情况下是非同步的 因此在多线程环境下需要做额外的同步机制。

HashTable 不允许有 null 值 key 和 value 都不允许，而 HashMap 允许有 null 值 key 和 value 都允许 因此 HashMap 使用 containKey（）来判断是否存在某个键。

HashTable 使用 Enumeration ，而 HashMap 使用 iterator。

Hashtable 是 Dictionary 的子类，HashMap 是 Map 接口的一个实现类。

481、MapReduce 中排序发生在哪几个阶段？这些排序是否可以避免？为什么？

提示：

一个 MapReduce 作业由 Map 阶段和 Reduce 阶段两部分组成，这两阶段会对数据排序，从这个意义上说，MapReduce 框架本质就是一个 Distributed Sort。在 Map 阶段，在 Map 阶段，Map Task 会在本地磁盘输出一个按照 key 排序（采用的是快速排序）的文件（中间可能产生多个文件，但最终会合并成一个），在 Reduce 阶段，每个 Reduce Task 会对收到的数据排序，这样，数据便按照 Key 分成了若干组，之后以组为单位交给 reduce（）处理。很多人的误解在 Map 阶段，如果不使用 Combiner 便不会排序，这是错误的，不管你用不用 Combiner，Map Task 均会对产生的数据排序(如果没有 Reduce Task 则不会排序，实际上 Map 阶段的排序就是为了减轻 Reduce 端排序负载)。由于这些排序是 MapReduce 自动完成的，用户无法控制，因此，在

hadoop 1.x 中无法避免，也不可以关闭，但 hadoop2.x 是可以关闭的。

482、编写 MapReduce 作业时，如何做到在 Reduce 阶段，先对 Key 排序，再对 Value 排序？

提示：

该问题通常称为”二次排序“，最常用的方法是将 Value 放到 Key 中，实现一个组合 Key，然后自定义 Key 排序规则（为 Key 实现一个 WritableComparable）。

483、如何使用 MapReduce 实现两个表 join，可以考虑一下几种情况：（1）一个表大，一个表小（可放到内存中）；（2）两个表都是大表？

提示：

第一种情况比较简单，只需将小表放到 DistributedCache 中即可；第二种情况常用的方法有：map-side join（要求输入数据有序，通常用户 Hbase 中的数据表连接），reduce-side join，semi join（半连接），具体资料可网上查询。

484、讲一下垃圾回收算法

485、大数据建模和清洗

486、聚类算法

487、ssh 锁涉及到权限

488、hive 如何优化

489、linux 如何合并文件

490、两个文件，每一个都是几百个亿条数据，都有订单字段，这两个表，如何关联，效率最高？

491、HBase 如果只向一个 RegionServer 写入数据，有什么优点？

492、java 如何实现高并发？

493、HashMap、TreeMap 区别，以及 TreeMap 原理

494、HBase 一行数据如何存储？

495、Spring 用过哪些组件？

496、如果有几百亿条数据，如果在表中存放？

497、mapreduce 全排序原理。