

Hadoop 测试题

姓名: _____

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) Rsync + ssh + scp
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, 并支持多个 Pool 提交 Job。()
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。()