

1+X 大数据平台运维初级模拟题

一、单选题

1. Centos 中修改文件或目录的访问权限（ C ）命令
A. chown
B. passwd
C. chmod
D. clear
2. 在文件属性中，文件的权限用第(B)列字符表示。
A. 1-5
B. 2-10
C. 3-9
D. 1-8
3. Hadoop 安装在一台计算机上，需修改相应的配置文件，用一台计算机模拟多台主机的集群是（B ）模式。
A. 全分布模式
B. 伪分布模式
C. 单机模式
D. 全分布 HA 模式
4. 配置 Hadoop 环境变量修改（A ）文件
A. vi /etc/profile
B. vi /etc/profiles

- C. vi /etc/hosts
 - D. vi ~/input/data
5. 配置 SSH 服务需配置 (B) 文件
- A. vi /etc/ssh/sshd
 - B. vi /etc/ssh/sshd_config
 - C. C.vi /etc/sysconfig/network-scripts/ifcfg-eth0
 - D. D.vi ~/.bash_profile
6. hadoop2.0 与 hadoop1.0 区别 (B)
- A. 增加 MapReduce2
 - B. 增加 YARN
 - C. 增加 HDFS2
 - D. 增加容错机制
7. 在 hadoop 配置中 core-site.xml 的配置是 (A) 参数。
- A. 集群全局参数
 - B. HDFS 参数
 - C. Mapreduce 参数
 - D. 集群资源管理系统参数
8. 在 hadoop 配置中 yarn-site.xml 作用是 (D)
- A. 用于定义系统级别的参数
 - B. 用于名称节点和数据节点的存放位置
 - C. 用于配置 JobHistory Server 和应用程序参数
 - D. 配置 ResourceManager, NodeManager 的通信端口

9. 如没有配置 `hadoop.tmp.dir` 参数, 此时系统默认的临时目录为(A)

- A. `/tmp/hadoop-hadoop`
- B. `/usr/hadoop-hadoop`
- C. `/usr/local/hadoop-hadoop`
- D. `/usr/local/src/hadoop-hadoop`

10. 以下哪个文件是能够修改 HDFS 的副本保存数量参数 (D)

- A. 修改 `mapred-site.xml` 配置文件
- B. 修改 `core-site.xml` 配置文件
- C. 修改 `yarn-site.xml` 配置文件
- D. 修改 `hdfs-site.xml` 配置文件

11. 在 MapReduce 中 Shuffle 的主要作用是 (B)

- A. 将数据进行拆分
- B. 对映射后的数据进行排序, 然后输入到 Reducer
- C. 经过映射后的输出数据会被排序, 然后每个映射器会进行分区
- D. 通过实现自定义的 Partitioner 来指定哪些数据进入哪个 Reducer

12. HDFS 启动成功的标识是 (D)

- A. `jps` 显示 NameNode
- B. `jps` 显示 DataNode

C. jps 显示 SecondaryNameNode

D. jps 显示

NameNode\DataNode\SecondaryNameNode

13. Hive 适合 (C) 环境

A. Hive 适合用于联机(online)事务处理

B. 提供实时查询功能

C. 适合应用在大量不可变数据的批处理作业

D. Hive 适合关系型数据环境

14. Hive 默认的元存储是保存在内嵌的 (B) 数据库中

A. MySQL

B. Derby

C. Oracle

D. SqlServer

15. 典型的 NoSQL 数据库是 (C)

A. Hive

B. MySQL

C. Hbase

D. Oracle

16. HMaster 的主要作用是 (A)

A. 用于启动任务管理多个 HRegionServer

B. 负责响应用户 I/O 请求, 向 HDFS 文件系统读写数据

C. 负责协调集群中的分布式组件

D. 最终保存 HBase 数据行的文件

17. ZooKeeper 的特点不包括 (D)

A. 顺序一致性

B. 原子性

C. 可靠性

D. 复合系统映像

18. Sqoop 的作用是 (A)

A. 用于传统关系型数据库和 Hadoop 之间传输数据

B. 提供对日志数据进行简单处理的能力

C. 是 Hadoop 体系中数据存储管理的基础

D. 是一个建立在 HDFS 之上, 面向列的针对结构化和半结构化数据的动态数据库

19. 测试 Sqoop 是否能够正常连接 MySQL 数据库命令是(D)

A. sqoop list-database --connect jdbc:mysql://127.0.0.1:3306/ --username root -P

B. sqoop list-databases --connection jdbc:mysql://127.0.0.1:3306/ --username root -P

C. sqoop list-databases --connect jdbc:mysql://127.0.0.1:3306/ --username root -L

D. sqoop list-databases --connect jdbc:mysql://127.0.0.1:3306/ --username root -P

20. 大数据解决方案确认需求分析准备工作不包括 (C)

- A. 充分了解客户的要求，客户的业务，客户的关注重点
- B. 和客户负责人沟通，确定客户需求小组成员
- C. 编制确定客户实施计划
- D. 根据事先了解的初步用户需求，列出这些难点

21. 大数据平台项目实施计划不包括（D ）

- A. 基础平台部署
- B. 用户环境准备
- C. 完成大数据平台的安装调试
- D. 完成大数据平台操作开发

22. 大数据平台操作演示不包括（ C ）

- A. 在大数据平台操作演示前需要做好准备工作
- B. 分析本次演示的目的和听众群体
- C. 实现呈现技巧提升
- D. 准备好演示环境

23. Hbase 的主要作用是（ A ）

- A. 是一个分布式的、面向列的开源数据库
- B. 是一种编程模型，用于大规模数据集（大于 1TB）的并行运算
- C. 是 Hadoop 集群当中的资源管理系统模块
- D. 将要储存的文件分散在不同的硬盘上，并记录他们的位置

24. 在配置好 Hadoop 集群之后, 通过浏览器登录查看

Hadoop 状态, 以下是查看的 URL (B)

- A. <http://master:8088>
- B. <http://192.168.1.6:50070>
- C. <http://master:16010>
- D. <http://master:9999>

25. 通过浏览器界面监控 Hive 的状态, 以下是查看的端口

(D)

- A. <http://master:8088>
- B. <http://192.168.1.6:50070>
- C. <http://master:16010>
- D. <http://master:9999>

26. 日志消息的级别中哪个是比较严重的情况(B)

- A. ALEKKT
- B. CRIT
- C. ERR
- D. DEBUG

27. 大数据处理中数据处理层不包括 (D)

- A. 数据清洗
- B. 数据规约
- C. 数据提取
- D. 数据分析

28. 大数据处理中数据应用层不包括（ C ）

- A. 数据挖掘
- B. 数据建模
- C. 数据可视化
- D. 数据分析

29. 日志消息的级别中那个是不会影响系统但值得注意(C)

- A. ALEKKT
- B. CRIT
- C. NOTICE
- D. DEBUG

30. 在进行 Hive 组件配置时，JAR 包冲突导致异常，出现以下错误信息(A)

- A. Failed to set setXIncludeAware(true) for parser
- B. mkdir: `input' : No such file or directory
- C. E45: 'readonly' option is set (add ! to override)
- D. ERROR tool.ImportTool:Imported Failed:Character 8216 is an out-of-range delimiter

31. Hadoop 官方真正支持的操作系统是 (D)

- A. UNIX
- B. Windows
- C. DOS
- D. Linux

32. 显示当前所在目录的命令是 (A)

A.pwd

B.ls

C.cd

D.mkdir

33. 以下选项哪个可以修改文件或目录的访问权限 (C)

A.passwd

B.su

C.chmod

D.useradd

34. 以下选项哪个是 HDFS 的名称节点 (D)

A. DataNode

B. SecondaryNode

C. Node

D. NameNode

35. 以下选项哪个是 MapReduce 正确的运行模型 (C)

A.Reduce-Map-Shuffle

B.Shuffle-Map-Reduce

C..Map-Shuffle-Reduce

D.Map-Reduce-Shuffle

36. 以下选项哪个是 YARN 中动态创建的资源容器 (D)

A.ResourceManager

B.NodeManager

C.ApplicationMaster

D.Container

37. 以下选项哪个命令可以生成公钥 (A)

A.ssh-keygen

B.ifup team0

C.nmcli dev

D.nmcli con sh

38. 以下选项哪个命令是用 hadoop 用户登录, 通过 scp 命令实现密钥拷贝到 slave1 节点上 (C)

A.scp ~/.ssh/id_rsa.pub hadoop@slave2:~/

B.scp ~/.ssh/id_rsa.pub root@slave1:~/

C.scp ~/.ssh/id_rsa.pub hadoop@slave1:~/

D.scp ~/.ssh/id_rsa.pub slave1@root:~/

39. 以下选项哪个命令可以查看主机信息(A)

A.nmcli dev

B.nmcli con sh

C.ifup team0

D.teamdctl team0 st

40. 在 Hadoop 配置文件中, 以下属于 core-site.xml 文件主要内容的是 (A)

A、用于定义系统级别的参数, 如 HDFS URL、Hadoop

的临时目录等。

B、名称节点和数据节点的存放位置、文件副本的个数、文件读取权限等。

C、配置 ResourceManager, NodeManager 的通信端口, web 监控端口等。

D、包括 JobHistory Server 和应用程序参数两部分, 如 reduce 任务的默认个数、任务所能够使用内存的默认上下限等。

41. 以下 (D) 文件中主要用来配置 ResourceManager, NodeManager 的通信端口, web 监控端口等。

A、core-site.xml

B、mapred-site.xml

C、hdfs-site.xml

D、yarn-site.xml

42. 在 HDFS 分布式文件系统中, 一般采用冗余存储, 冗余因子通常设置为 (C)

A、1

B、2

C、3

D、4

43. 用于检查节点自身的健康状态的进程是 (B)

A、ResourceManager

B、 NodeManager

C、 NameNode

D、 DataNode

44. ResourceManager 对 web 服务提供地址，用户可通过

ip 地址：(A) 在浏览器中查看集群各类信息

A、 8088

B、 8033

C、 8080

D、 8032

45. 启动 HDFS 至少需要启动 (ABD) 进程

A、 NameNode

B、 ResourceManager

C、 SecondaryNameNode

D、 DataNode

46. 在 HDFS 文件操作中，上传文件的命令是 (A)

A、 put

B、 input

C、 get

D、 up

47. MySQL 数据库驱动文件放置于 hive 哪个目录下 (B)

A、 jar

B、 lib

C、 bin

D、 sbin

48. 下列哪些不是 HBase 的特点 (D)

A. 高可靠性

B. 高性能

C. 面向列

D. 紧密性

49. HBase 与下列哪个选项属于同一种类型的数据库 (A)

A. MongoDB

B. MariaDB

C. MySQL

D. Oracle

50. 在 HBase 的组件中, 哪一个负责日志记录 (D)

A. HRegion

B. HFile

C. MemStore

D. WAL

51. 下列哪一种命令格式是不正确的 (B)

A. get 表 行键 列族

B. scan 表 时间戳 起始行键 结束行键

C. alter 表 列族

D. put 表 行键 列族:列 值

52. 下列哪些不是 ZooKeeper 的特点 (C)

- A. 顺序一致性
- B. 原子性
- C. 多样系统映像
- D. 可靠性

53. 在 ZooKeeper 中有三种角色, 下列选项中不属于 ZooKeeper 中的角色 (C)

- A. Observer
- B. Leader
- C. Obeyer
- D. Follower

54. 下列命令中(在 zookeeper 安装文件夹的 bin 目录下执行), 哪项是停止 ZooKeeper 的正确命令 (C)

- A. start-all.sh
- B. start-stop.sh
- C. zkServer.sh stop
- D. zkServer.sh start

55. ZooKeeper 是一个典型的分布式数据一致性解决方案, 下列哪项是不属于它实现的功能 (C)

- A. 分布式锁
- B. 数据发布/订阅
- C. 响应用户 I/O 请求

D. 负载均衡

56. Sqoop 是 Hadoop 和关系数据库服务器之间传送数据的工具，下列哪种数据库是不支持使用 Sqoop 直接进行数据传送 (D)

A. Postgres

B. Oracle

C. MySQL

D. Redis

57. 下列描述中，哪项不属于 Sqoop 的缺点 (A)

A. 无法控制任务的并发度

B. 格式紧耦合

C. 安全机制不够完善

D. connector 必须符合 JDBC 模型

58. Sqoop 将导入或导出命令翻译成什么程序来实现导入导出 (C)

A. Spark

B. Storm

C. MapReduce

D. Tez

59. 在导入开始之前,Sqoop 检索出表中所有的列以及列的 SQL 数据类型,这些 SQL 类型被映射什么类型 (C)

A. C

- B. C ++
- C. Java
- D. Python

60. 关于 Flume 的三大组件以下说法正确的是(A)。

- A. channel 可以和任意数量的 source 和 sink 链接。
- B. channel 只能链接单一的 source 和 sink。
- C. sink 在三大组件之间起着桥梁的作用。
- D. sink 从 channel 消费数据并将其传递给目标地，目标地只能是 hdfs。

61. 使用 root 用户解压 Flume 安装包到 “/usr/local/src” 路径,以下哪条是正确的命令 (C)。

- A. [root@master ~]#tar zxvf /opt/software/apache-flume-1.6.0-bin.tar.gz /usr/local/src
- B. [hadoop@master ~]#tar zxvf /opt/software/apache-flume-1.6.0-bin.tar.gz -C /usr/local/src
- C. [root@master ~]#tar zxvf /opt/software/apache-flume-1.6.0-bin.tar.gz -C /usr/local/src
- D. [root@master ~]#unzip /opt/software/apache-flume-1.6.0-bin.tar.gz -C /usr/local/src

62. 在本次项目实施中，需求调研前的准备不包括 (B)。

- A. 充分了解客户的要求，客户的业务，客户的关注重

点。

B. 接到调研任务，将计划写好，立即就开始和用户沟通。

C. 编制确定客户需求调研计划和调研提纲。

D. 和客户负责人沟通，确定客户需求小组成员。

63. 大数据平台实施方案流程中,建议整个项目过程顺序是 (A)。

A. 与客户确认实施方案

B. 确认客户需求

C. 系统测试

D. 编写实施方案

E. 用户培训

F. 执行实施方案

G. 项目验收

A. BDAFCEG B. ABDFCEG

C. DEABCFG D. BDAFEGC

64. 在确认客户需求，进行确认需求调研的时候，以下说法正确的是 (A)。

A. 按计划有步骤的确定客户需求调研。

B. 先了解细节需求，再了解宏观需求。

C. 将客户的想法记录下来，就能解决客户的实际问题。

D. 确定客户需求调研的过程中，不可避免的会出现客户

提出一些现有条件下根本无法实现或者即使实现也非常困难的要求，我们要无条件去满足客户的需求。

65. 在项目实施人员中，适时与客户沟通，确立项目目标，建立组织机构，编制计划，负责项目全面管理的角色是 (B)。

- A. 实施工程师
- B. 项目经理
- C. 产品经理
- D. 财务人员

66. 大数据分析平台的实施流程顺序是 (A)。

- A. 与甲方确定总体计划
- B. 组建项目团队
- C. 部署环境准备
- D. 应用集成及数据集成
- E. 系统安装和调试

A. BACED B. CABED C.ADCBE D. DABCE

67. 关于培训的总体目标、定制培训课程和内容 and 培训形式，以下说法正确的是 (D)。

- A. 系统使用人员的培训目的是使其可以独立完成大数据平台的日常维护，解决一般问题。
- B. 培训过程中将针对不同的用户和不同的培训内容采用相同的培训方案，以达到最佳的培训效果。

C. 实践培训主要针对于技术开发人员及系统维护和客户。

D. 培训形式包括集中培训、现场培训、针对性培训。

68. 大数据平台 Hadoop 的组件中，提供分布式协作服务的组件是 (B)

A. HBase

B. ZooKeeper

C. Sqoop

D. Hive

69. 以下哪个不是 HDFS 的进程 (C)

A. SecondaryNode

B. NameNode

C. ResourceManager

D. DataNode

70. 查看 HDFS 的报告命令是 (A)

A. bin/hdfs dfsadmin -report

B. bin/hdfs -report

C. bin/hdfs deadmin -report

D. bin/hdfs dfs -report

二、多选题

1. Linux 版本分为以下哪几类 (AC)

- A. 发行版
- B. 测试版
- C. 内核版
- D. 开发版

2. Linux 的特点是以下 (ACD) 方面

- A. 开放源代码的程序软件，可自由修改。
- B. 与 Unix 系统不兼容，具备几乎所有 Unix 的优秀特性。
- C. 可自由传播，免费使用，无任何商业化版权制约。
- D. 适合 Intel 等 x86 CPU 系列架构的计算机。

3. Hadoop 的运行模式分为以下 (ABC) 种

- A. 单机模式
- B. 伪分布模式
- C. 全分布模式
- D. 主备模式

4. Hadoop 优点是以下 (ABCD) 种。

- A. 扩容能力强
- B. 成本低
- C. 高效率
- D. 高可靠性

5. 验证 SSH 无密码登录需要 (ABCD) 步骤

- A. 修改 Master 节点 authorized_keys 文件
 - B. 修改 Slave 节点 authorized_keys 文件
 - C. 验证 Master 到每个 Slave 节点无密码登录
 - D. 验证两个 Slave 节点到 Master 节点无密码登录
6. 查询获取 Hadoop 全部的配置信息官方网址是 (ABCD)
- A. <http://hadoop.apache.org/docs/current/hadoop-project-dist/hadoop-common/core-default.xml>
 - B. <http://hadoop.apache.org/docs/current/hadoop-project-dist/hadoop-hdfs/hdfs-default.xml>
 - C. <http://hadoop.apache.org/docs/current/hadoop-mapreduce-client/hadoop-mapreduce-client-core/mapred-default.xml>
 - D. <http://hadoop.apache.org/docs/current/hadoop-yarn/hadoop-yarn-common/yarn-default.xml>
7. 以下哪些是 yarn-site.xml 配置文件主要参数 (ABC)
- A. yarn.resourcemanager.address
 - B. yarn.resourcemanager.scheduler.address
 - C. yarn.resourcemanager.admin.address
 - D. yarn.nodemanager.oux-services
8. Hadoop 平台的主要配置文件是 (ABC)
- A. core-site.xml
 - B. hdfs-site.xml

C. mapred-site.xml

D. slaves.xml

9. core-site.xml 文件核心参数是 (ABD)

A. mapreduce.framework.name

B. mapreduce.jobhistory.address

C. yarn.resourcemanager.address

D. mapreduce.jobhistory.webapp.address

10. MapReduce 是 Hadoop 中集群数据处理的核心, 分为
(AB)

A. 分为映射, 映射为 Map 阶段

B. 分为减速, 减速为 Reduce 阶段

C. 分为映射, 映射为 Reduce 阶段

D. 分为减速, 减速为 Map 阶段

11. Hive 优点是 (ABCD)

A. 提供了类 SQL 的查询语言 HQL。

B. 可扩展强

C. 提供统一的元数据管理

D. 容错行强

12. Hbase 组件部分包括 (ABCD)

A. hbase-client

B. Zookeeper

C. HMaster

D. HRegionServer

13. Flume 将采集到的数据输出到(abcd)

A. HDFS

B. Hbase

C. Hive

D. Kafka

14. 大数据平台安装部署步骤包括 (ABCD)

A. 服务器上架

B. 服务器操作系统安装

C. 网络环境分配

D. Hadoop 平台安装

15. 大数据培训形式分 (ABC)

A. 集中培训

B. 现场培训

C. 针对性培训

D. 专业性培训

16. 大数据运维中的日志分类为 (ABC)

A. 内核及系统日志：由系统服务 syslog 统一进行管理,日志格式基本相似。

B. 用户日志：记录系统用户登录及退出系统的相关信息。

C. 程序日志：由各种应用程序独立管理的日志文件,记录格式不统一

D. 错误日志：由各种错误信息组成的日志信息

17. Hadoop 集群验证步骤包括 (ABCD)

A. 初始化集群，使用 Hadoop 命令启动集群。

B. 使用 Hadoop 命令，创建 HDFS 文件夹。

C. 使用 HDFS 命令查看文件系统 "/" 路径下是否存在文件。

D. 用 Hadoop 自带的 WordCount 程序去测试
MapReduce

18. 大数据业务处理系统包含 (ABCD)

A. 数据源层

B. 数据处理层

C. 数据应用层

D. 数据访问层

19. ZooKeeper 角色选举角色包括(ABCD)

A. 领导者 (leader)

B. 观察者 (observer)

C. 客户端 (client)

D. 跟随者 (follower)

20. 主流虚拟化软件包括 (ABCD)

A. VMware 系列虚拟化软件

B. KVM 二次开发虚拟化软件

C. 微软 Hyper-V 虚拟化软件

D. Citrix 虚拟化软件

21. 以下哪些是 Linux 的特点 (ABD)

- A. 开放源代码的程序软件，可自由修改。
- B. 与 Unix 系统兼容，具备几乎所有 Unix 的优秀特性。
- C. 可自由传播，收费使用，无任何商业化版权制约。
- D. 适合 Intel 等 x86 CPU 系列架构的计算机。

22. 以下哪些是 Linux 操作系统组成部分 (ABC)

- A. Linux 内核
- B. GNU 软件及系统软件
- C. 必要的应用程序
- D. CPU

23. 以下哪些组件属于 Hadoop 生态圈 (BCD)

- A. MySQL
- B. Hive
- C. HDFS
- D. Sqoop

24. Hadoop 的优点有哪些 (ABCD)

- A. 扩容能力强
- B. 成本低
- C. 高效率
- D. 高可靠性

25. 以下选项中哪些是 SSH 涉及到的两个验证方式 (AB)

- A. 主机验证

B.用户身份验证

C.文件系统验证

D.操作系统验证

26. 以下选项中哪些是常用的 ss 命令 (ABC)

A.scp

B.sftp

C.ssh-copy-id

D.ifconfig

27. Hadoop 核心主要由 (AC) 部分组成

A、 HDFS

B、 DFS

C、 MapReduce

D、 RDD

28. Hive 作为大数据分析的重要组件, 常用在对 (ABC) 数据进行分析

A、 HDFS

B、 Mysql

C、 Hbase

D、 Hadoop

29. master 启动 Hadoop 相关进程包含 (BCD)

A、 jps

B、 NameNode

C、 SecondaryNameNode

D、 ResourceManager

30. 以下哪条命令能退出 mysql 状态 (AB)

A、 quit

B、 exit

C、 Ctrl+c

D、 Ctrl+z