**2020年安徽省大数据与人工智能应用初赛**

**学校名称： 队伍名称： 队伍编号：**

**说明：本次比赛为操作题。请在相应的虚拟机上作答，请遵守考试规则。**

# 第一部分：大数据平台部署（10分）

**场景说明**

数据中台的概念是最早由阿里巴巴首次提出，是为了应对像双十一这样的业务高峰、应对大规模数据的线性可扩展问题、应对复杂业务系统的解耦问题，而在技术、组织架构等方面采取的一些变革，其本质上还是一个大数据平台，开源框架中由Hadoop和Hive为主要核心框架组成，下面请你完成hadoop和hive的安装：

## 任务一、Hadoop安装（按下面要求提供核心步骤）（5分）

### 配置SSH免密登录（截图成功ssh登录到子节点截图）（0.5分）





### 解压Hadoop安装包并修改配置文件（截图Slaves文件中内容）（0.5分）



### 3、配置hadoop环境变量（截图profile文件配置的关键位置）（1分）



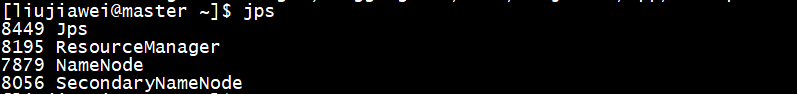
### 4、拷贝hadoop到其他的机器上（截图拷贝命令）（0.5分）

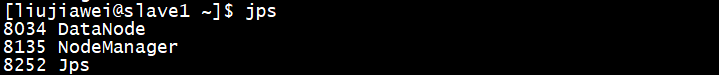


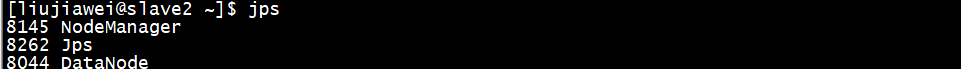
### 5、初始化hadoop集群（截图执行命令）（1分）



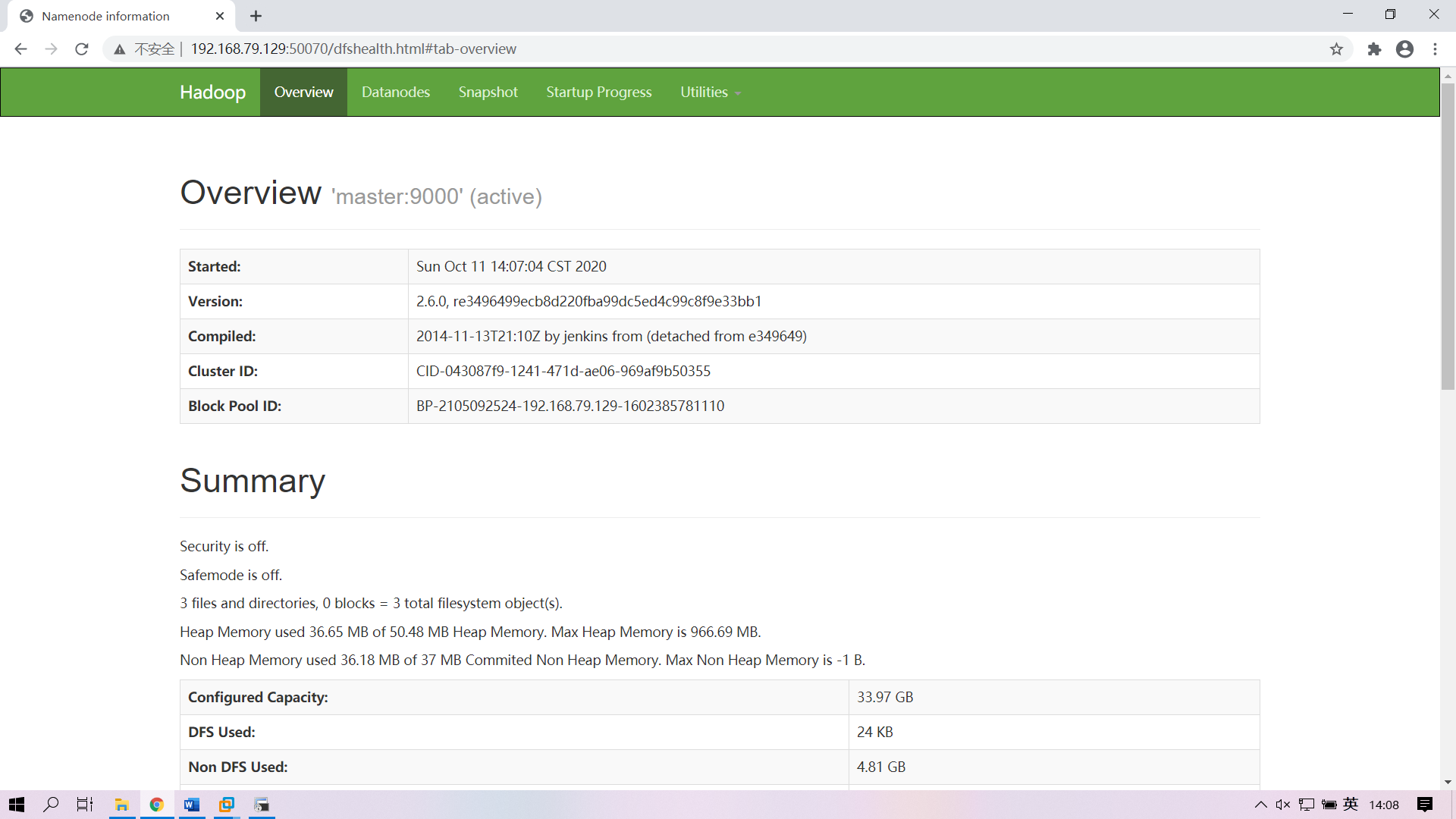
### 6、启动Hadoop集群（启动成功后执行jps命令，截图主节点和子节点的进程）（0.5分）





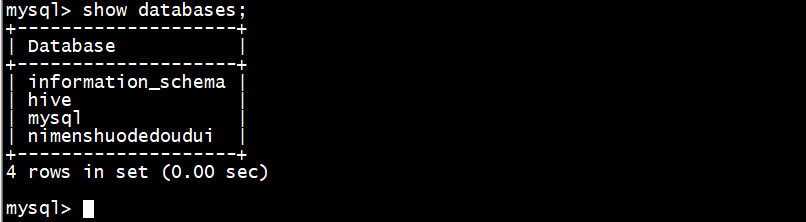


7、在浏览器中访问 主节点:50070端口的webUI页面。并截图（建议Chrome浏览器）（1分）

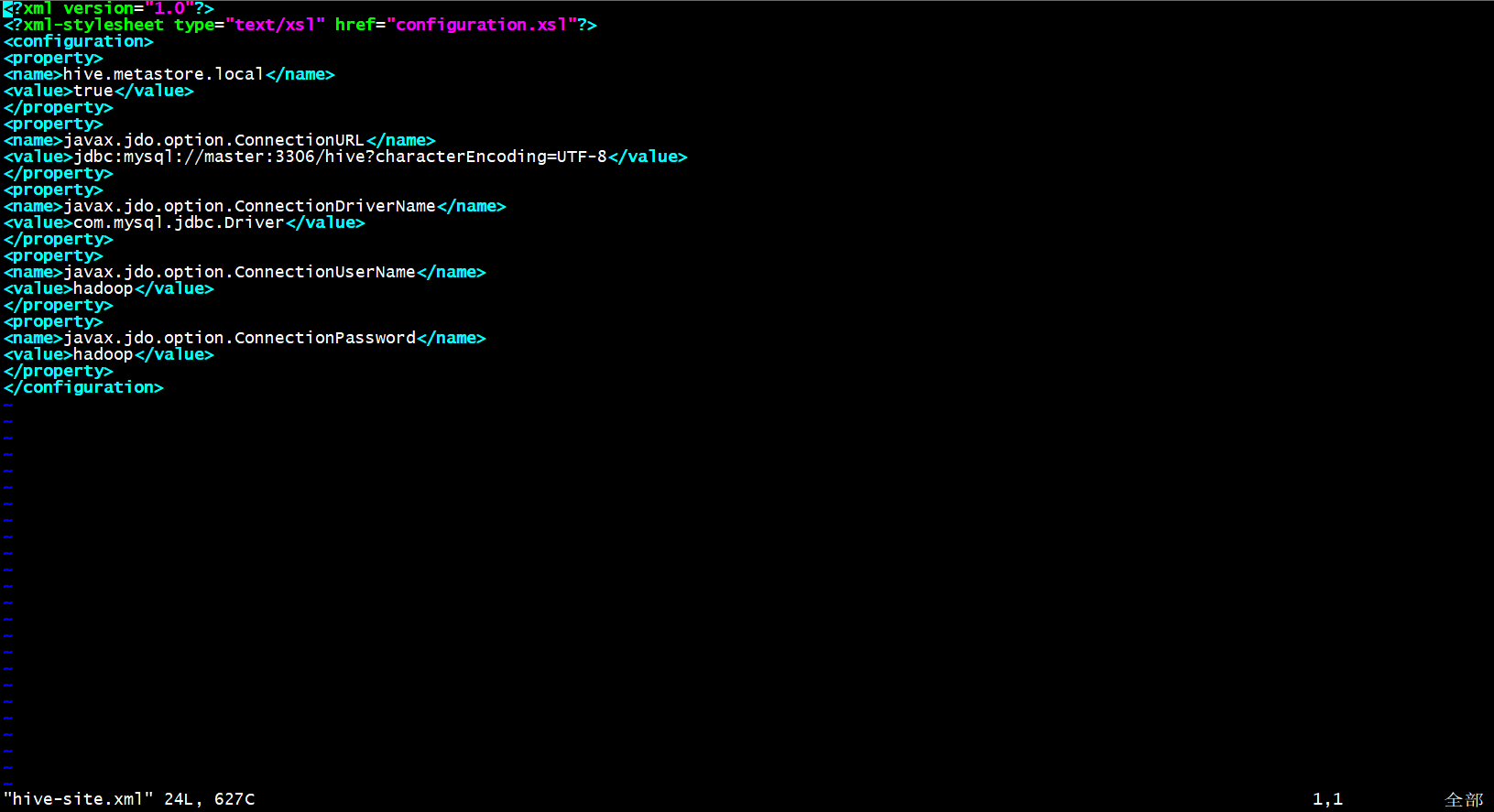


## 任务二、Hive安装（按下面要求提供核心步骤）（5分）

### 基于Linux系统Mysql安装、在Mysql中创建一个以你队伍组名英文全拼命名的数据库（截图结果界面2分）



### Hive搭建，修改hive-site.xml文件中默认元数据库驱动为mysql的驱动（截图2分）



3、启动hive，启动成功后，执行show databases；（截图结果1分）

