**浙 江 理 工 大 学**

**实**

**践**

**工**

**作**

**日**

**志**

计算机科学与技术（人工智能）学院

学 院：\_\_\_\_\_\_\_\_\_\_\_\_\_ \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

年 级：

计算机科学与技术

2021

专 业：

姓 名：

陈昊天

学 号：

2021329600006

实践时间： 2024 年 6 月 21 日至 2024 年 6 月 23 日

**实 践 工 作 日 志**

|  |  |
| --- | --- |
| |  | | --- | | 2024年06月21日 |   **安装 MySQL**  1. 安装MySQL 服务器  在node1上执行以下命令：  sudo apt update  sudo apt-get install mysql-server    2. 配置MySQL  启动MySQL服务并进行安全设置：  sudo systemctl start mysql  sudo mysql\_secure\_installation    按照提示完成安全设置，创建root用户密码等。  3. 创建Hive数据库  使用MySQL登录并创建Hive所需的数据库和用户：  sudo mysql -u root -p    在MySQL shell中执行以下命令：  CREATE DATABASE metastore;  CREATE USER 'hiveuser'@'localhost' IDENTIFIED BY 'hivepassword';  GRANT ALL PRIVILEGES ON metastore.\* TO 'hiveuser'@'localhost';  FLUSH PRIVILEGES;  EXIT; |
| 指导老师评语： |

**实 践 工 作 日 志**

|  |  |
| --- | --- |
| |  | | --- | | 2024年06月22日 |   **安装 Hive**  1. 下载并解压Hive  在node1上执行以下命令来下载并解压Hive：  cd /usr/local  wget https://downloads.apache.org/hive/hive-3.1.3/apache-hive-3.1.3-bin.tar.gz  sudo tar -zxvf apache-hive-3.1.3-bin.tar.gz  sudo mv apache-hive-3.1.3-bin hive  rm apache-hive-3.1.3-bin.tar.gz    2. 配置环境变量  在~/.bashrc文件中添加以下内容：  export HIVE\_HOME=/usr/local/hive  export PATH=$PATH:$HIVE\_HOME/bin    使更改生效：  source ~/.bashrc    3. 配置Hive    创建Hive配置目录并编辑配置文件：  sudo mkdir -p $HIVE\_HOME/conf  cd $HIVE\_HOME/conf  # sudo cp $HIVE\_HOME/conf/hive-default.xml.template $HIVE\_HOME/conf/hive-site.xml  sudo vim $HIVE\_HOME/conf/hive-site.xml    `hive-site.xml`：  <configuration>  <property>  <name>javax.jdo.option.ConnectionURL</name>  <value>jdbc:my://localhost/metastore</value>  <description>JDBC connect string for a JDBC metastore</description>  </property>  <property>  <name>javax.jdo.option.ConnectionDriverName</name>  <value>com.my.cj.jdbc.Driver</value>  <description>Driver class name for a JDBC metastore</description>  </property>  <property>  <name>javax.jdo.option.ConnectionUserName</name>  <value>hiveuser</value>  <description>Username to use against metastore database</description>  </property>  <property>  <name>javax.jdo.option.ConnectionPassword</name>  <value>hivepassword</value>  <description>Password to use against metastore database</description>  </property>  <property>  <name>hive.metastore.warehouse.dir</name>  <value>/user/hive/warehouse</value>  <description>location of default database for the warehouse</description>  </property>  <property>  <name>hive.cli.print.header</name>  <value>true</value>  </property>  <property>  <name>hive.cli.print.current.db</name>  <value>true</value>  </property>  </configuration>      4. 下载MySQL JDBC驱动  下载MySQL JDBC驱动并放到Hive的lib目录：  cd /usr/local  wget  https://dev.mysql.com/get/Downloads/Connector-J/mysql-connector-java-8.0.26.tar.gz  tar -zxvf mysql-connector-java-8.0.26.tar.gz  sudo cp mysql-connector-java-8.0.26/mysql-connector-java-8.0.26.jar $HIVE\_HOME/lib/  rm mysql-connector-java-8.0.26.tar.gz  # rm -r mysql-connector-java-8.0.26    5. 初始化Hive Metastore  运行以下命令初始化Hive Metastore：  schematool -initSchema -dbType mysql    6. 验证安装  启动Hive CLI并验证安装：  hive    7. 修改HDFS权限  hdfs dfs -mkdir -p /user/hive/warehouse  hdfs dfs -chmod -R 777 /user/hive/warehouse    **验证 Hive**  1. 创建一个目录来存放数据文件  hdfs dfs -mkdir -p /user/hive/warehouse/user\_data    2. 上传数据文件到新目录  创建一个数据文件：  echo -e "1,John\n2,Jane\n3,Bob" > data.csv  然后将数据文件上传到 HDFS 的 `/user/hive/warehouse/user\_data` 目录：  hdfs dfs -put data.csv /user/hive/warehouse/user\_data/    3. 在Hive CLI中，运行以下命令创建一个示例表并查询：  CREATE EXTERNAL TABLE IF NOT EXISTS user\_data (  id INT,  name STRING  )  ROW FORMAT DELIMITED  FIELDS TERMINATED BY ','  STORED AS TEXTFILE  LOCATION 'hdfs://node1:8020/user/hive/warehouse/user\_data';    SELECT \* FROM user\_data;  IMG_256 |
| 指导老师评语： |

**实 践 工 作 日 志**

|  |  |
| --- | --- |
| |  | | --- | | 2024年06月23日 |   **外部表 案例实操**  分别创建部门和员工外部表，并向表中导入数据。  1 原始数据  注意数据以制表符分隔  dept.txt  10 ACCOUNTING 1700  20 RESEARCH 1800  30 SALES 1900  40 OPERATIONS 1700  emp.txt  7369 SMITH CLERK 7902 1980-12-17 800.00 0.00 20  7499 ALLEN SALESMAN 7698 1981-2-20 1600.00 300.00 30  7521 WARD SALESMAN 7698 1981-2-22 1250.00 500.00 30  7566 JONES MANAGER 7839 1981-4-2 2975.00 0.00 20  7654 MARTIN SALESMAN 7698 1981-9-28 1250.00 1400.00 30  7698 BLAKE MANAGER 7839 1981-5-1 2850.00 0.00 30  7782 CLARK MANAGER 7839 1981-6-9 2450.00 0.00 10  7788 SCOTT ANALYST 7566 1987-4-19 3000.00 0.00 20  7839 KING PRESIDENT NULL 1981-11-17 5000.00 0.00 10  7844 TURNER SALESMAN 7698 1981-9-8 1500.00 0.00 30  7876 ADAMS CLERK 7788 1987-5-23 1100.00 0.00 20  7900 JAMES CLERK 7698 1981-12-3 950.00 0.00 30  7902 FORD ANALYST 7566 1981-12-3 3000.00 0.00 20  7934 MILLER CLERK 7782 1982-1-23 1300.00 0.00 10  2 建表语句  创建部门表  create external table if not exists dept(  deptno int,  dname string,  loc int  )  row format delimited fields terminated by '\t';  创建员工表  create external table if not exists emp(  empno int,  ename string,  job string,  mgr int,  hiredate string,  sal double,  comm double,  deptno int)  row format delimited fields terminated by '\t';  3 查看创建的表  show tables;  IMG_256  4 向外部表中导入数据  导入数据，数据文件位于虚拟机上  load data local inpath '/root/dept.txt' into table default.dept;  load data local inpath '/root/emp.txt' into table default.emp;  IMG_256  查询结果  select \* from dept;  select \* from emp;  IMG_256  5 查看表格式化数据  desc formatted dept;  IMG_256 |
| 指导老师评语： |