# 下载指定位置的数据

下载附件中的数据文件【任务3清洗后的数据.zip】，解压，使用SFTP相关工具将数据传至slave2服务器的/home目录下

# 使用Hadoopshell命令创建目录/data/

Hadoop fs –mkdir /data

# 上传数据至HDFS的/data目录下

使用命令：hadoop fs –put /home/任务3清洗后的数据 /data/

# Hive创建数据库

Create database db\_phone\_raw\_3;

# Hive创建表

## 原始数据表

Create table tbl\_phone\_data\_3(fld\_phone\_os string, fld\_Phone\_cpu string, fld\_phone\_ram string, fld\_phone\_rom string) row format delimited fields terminated by '|';

## 手机RAM销售量统计表

Create table tbl\_sales\_ram\_count (fld\_ram\_name string, fld\_sale\_count int);

## 手机操作系统统计表

Create table tbl\_sales\_os\_count(fld\_os\_name string,fld\_sale\_count int);

## 手机CPU销售量统计表

Create table tbl\_sales\_cpu\_count(fld\_cpu\_name string, fld\_sale\_count int);

## 手机ROM销售量统计表

Create table tbl\_sales\_rom\_count(fld\_rom\_name string,fld\_sale\_count int);

# 数据加载命令

Load data inpath '/Clean\_Data/\*' into table tbl\_phone\_data\_3;

# 查询分析

## 手机RAM销量查询

insert overwrite table tbl\_sales\_ram\_count select fld\_phone\_ram,count(1) as fld\_sale\_count from tbl\_phone\_data\_3 group by fld\_phone\_ram;

## 手机cpu销量查询

insert overwrite table tbl\_sales\_cpu\_count select fld\_phone\_cpu,count(1) as fld\_sale\_count from tbl\_phone\_data\_3 group by fld\_phone\_cpu;

## 手机操作系统销量查询

insert overwrite table tbl\_sales\_os\_count select fld\_phone\_os,count(1) as fld\_sale\_count from tbl\_phone\_data\_3 group by fld\_phone\_os;

## 手机rom销量查新

insert overwrite table tbl\_sales\_rom\_count select fld\_phone\_rom,count(1) as fld\_sale\_count from tbl\_phone\_data\_3 group by fld\_phone\_rom;