# 前置环境

## Hadoop环境

注：使用Hive和查看数据文件2、3都需要登录Hadoop用户

### 1、登录机器

IP: 128.32.61.66 武数

huangjiande.zh

### 2、使用applaud 用户

武汉数据中心

N-DI数据集成

WN1QETLWLAP0001

选择applaud@11.38.83.1:22

### 3、登录Hadoop用户

source /home/ap/dip/hadoop/hadoop\_client/bigdata\_env

kinit -k TS150 -t /home/ap/dip/hadoop/keytab/TS150.keytab

### 4、查看验证当前用户

使用用户名为 TS150

查看命令 klist

# 查看HIVE数据库

## 1、进入Hive 命令控制台

beeline

## 2、切换数据库到 ts150

use ts150;

## 3、查看所有表

根据文件的名字可知 其对应的表名

show tables;

## 4、查看表的结构

以ETL\_WEBSENSE\_DLP 表为例子

desc etl\_websense\_dlp;

## 5、核对表的数据与结构

列格式查看表的单条全字段数据：

!? outputformat vertical

Select \* from ct\_toddc\_crcrdcrd\_h limit 1;

返回默认查看格式：

!? outputformat table

**用于和GP对比字段值是否相同**

## 6、查询表记录数

使用SQL语句查询，语句以英文分号结束

如 p9\_start\_date 为表的日期字段 ，按日期查看 所有按日期分组的记录数

select p9\_start\_date,count(\*) from ct\_toddc\_crcrdcrd\_h group by p9\_start\_date order by p9\_start\_date;

查看 某一天的记录数

select count(\*) from ct\_toddc\_crcrdcrd\_h where p9\_start\_date ='20150501';