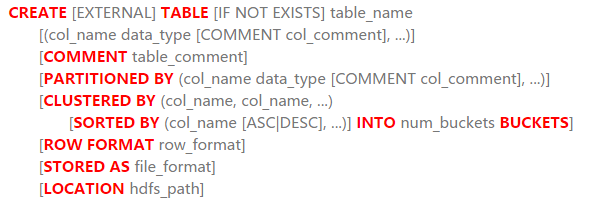
1. 创建表



Hive 主要处理的文件：hdfs上的文本文件。Hive要将文本文件的内容要以表的形式进行处理。那也就是hive要将文件内的数据转换成表的形式。文件的内容如何和虚拟出来的表如何映射。

### 内部表（1.删除表的时候，同时删除文件夹（数据也没了））

Hive

1.数据（hdfs）

2.hive是需要像mysql一样可以执行类sql语句，就需要有表名，列名。。。

3.hive 元数据 🡺 mysql

内部表：删除表的时候，会删除原始数据（该表加载到该表所以对应的文件夹中）

1. 加载数据 🡺 将数据挪移到指定目录。

问题： 加载数据的时候，数据位置在本地或者hdfs有什么区别？

1. 本地 🡺 put
2. Hdfs 🡺 mv

大数据分析当中，数据可能进行多次分析。如果将原始数据储存在内部表当中，一旦删除

内部表会连带原始数据一同删除。

适合：用完就不要的数据。

数据存储在Hive的数据仓库目录下，删除表时，除了删除元数据，还会删除实际表文件，内部表数据存储的位置是hive.metastore.warehouse.dir

创建内部表

hive> create table if not exists abc (id bigint, name string, age int) row format delimited fields terminated by ',';

hive> show tables;

本地导入数据

[hadoop@h201 ~]$ cat h1.txt

10001,zs,20

20002,ls,30

30003,ww,40

40004,zl,50

50005,tq,60

60006,qq,70

hive> load data local inpath '/home/hadoop/h1.txt' into table abc;

hive> select \* from abc;

复制表

hive> create table t1 as select id,name from abc; //复制数据。文件的分割和合并。

hive> dfs -ls /hive/warehouse/t1;

创建表同时载入原有表结构

hive> create table t2 like abc;

hive> desc t2;

OK

id bigint

name string

age int

内部表存储在warehouse，外部表存储在warehouse 外，这不是语法规定，按规范来说的。

### 临时表（属于客户端，断开连接消失）

hive> create temporary table t3 as select \* from abc;

### 1.3外部表

1.原始数据可以存储在外部表当中。create table t1 as select \* from abc;将原始数据拆分成多个内部表进行操作。

带external参数为外部表，数据并不存储在Hive的数据仓库目录下，删除表时，只是删除元数据，并不删除实际表文件。

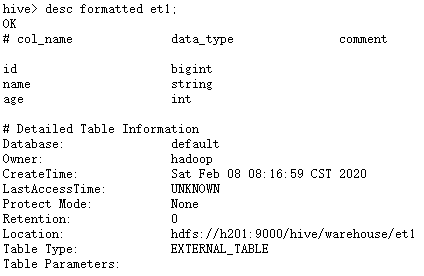
hive> create external table et1(id bigint,name string,age int) row format delimited fields terminated by ',';

导入数据

hive> load data local inpath '/home/hadoop/h1.txt' overwrite into table et1;

hive> select \* from et1;

hive> desc formatted et1;



验证外部表

hive> drop table et1;

hive> desc det1;

FAILED: SemanticException [Error 10001]: Table not found det1

(表已经不存在了，证明元数据已经被删除)

hive> dfs -ls /hive/warehouse/et1;

Found 1 items

-rwxr-xr-x 2 hadoop supergroup 72 2020-02-08 08:19 /hive/warehouse/et1/h1.txt

(hdfs上的文件依然存在)

### 表管理（hive的元数据）

hive> select \* from t1;

OK

10001 zs 20

20002 ls 30

30003 ww 40

40004 zl 50

50005 tq 60

60006 qq 70

* 1. 添加字段

hive> alter table t1 add columns(sal bigint);

hive> desc t1;

OK

id bigint

name string

age int

sal bigint

* 1. 删除字段

使用替换方式删除字段

hive> alter table t1 replace columns(id bigint,name string);

d