

知 什么是Hive?

厉梦如 2019-05-19 发表

问题描述

Q: 什么是Hive?

解决方法

A: Hive 可以被认为是一种数据仓库，它包括了数据存储和数据查询。Hive可以用来进行数据提取转化加载（ETL）。HQL是Hive下的查询语言，是Hibernate Query Language的缩写，提供更加丰富灵活、更为强大的查询能力；HQL更接近SQL语句查询语法。完整的HQL语句形式如下：Select/update/delete..... from where group by having order by asc/desc 其中的update/delete为Hibernate3中所新添加的功能。

Hive内部表都属于缺省库default，在HDFS的目录为/user/hive/warehouse/

I Hive是建立在Hadoop HDFS上的数据仓库工具

I Hive定义HQL允许熟悉SQL的用户查询数据

I Hive的表其实就是HDFS上的目录/文件

I Hive是SQL解析引擎，他将SQL语句转移成M/R Job然后在Hadoop执行