

知 MPP2 (Vertica) 对于结构化和非结构化数据混合分析如何?

厉梦如 2019-05-19 发表

问题描述

Q: MPP2 (Vertica) 对于结构化和非结构化数据混合分析如何?

解决方法

A: Vertica提供Hadoop Connector, 支持Map任务并发读取Vertica的数据, 以及Reduce 任务把数据保存到Vertica中。

Vertica的Hadoop Connector还允许Pig存取Vertica中的数据, 并支持直接使用 Hadoop Streaming 接口在Vertica与HDFS之间同步数据。

Vertica还可以与HCatalog深度集成。Hadoop上的Hive等定义表信息, 可以直接在Vertica中访问, 从而实现Hadoop上的数据与Vertica内部列式存储数据以标准SQL的方式对外提供关联分析服务。