

知 MPP1数据加载方式有哪些?

厉梦如 2019-05-19 发表

问题描述

Q: MPP1数据加载方式有哪些?

解决方法

A:

1、insert

这种数据加载方式效率最差，只适合加载极少量的数据。

2、copy

copy这种加载方式源于PostgreSQL，较SQL的insert方式效率大大提升，但数据仍然需要通过Master节点，无法实现并行高效数据加载。

3、gpfdist

支持通过外表，在Segment上并发地从gpfdist导入数据，由于是直接从Segment上导入数据，所以效率非常高。一个gpfdisk实例加载速度可以达到200MB/s

4、gpload

gpload是GP使用可读外表和gpdist装载数据的一个命令包装，本质与gpfdist相同。