

精选 | Apache Impala引领传统分析数据库技术的发展

2017-05-12 Cloudera中国



点击上方“公众号”可以订阅哦！

作者：Greg Rahn和Mostafa Mokhtar
Greg Rahn是Cloudera公司Impala团队的产品管理总监。
Mostafa Mokhtar是Cloudera公司的一名软件工程师，负责Impala团队的工作。
网址：<http://blog.cloudera.com/blog/2017/04/apache-impala-leads-traditional-analytic-database/>

与传统的分析数据库（Greenplum）相比，未经修改的基于TPC-DS的性能基准测试表现出了Impala的领导地位，特别是对于多用户并发工作负载而言。此外，基准测试还进一步证明了分析数据库与Hive LLAP、Spark SQL和Presto等SQL-on-Hadoop引擎之间存在的显著性能差距。

过去一年是Apache Impala（正在孵化中）发展变化最大的一年。Impala团队不仅继续努力不断扩大其规模和稳定性，而且还推出了一系列的关键功能，进一步巩固了Impala作为高性能商务智能（BI）和SQL分析的开放标准地位。对于云计算和混合部署而言，Impala现在可以提供云端-本地部署弹性、灵活性，以及直接从Amazon S3对象存储中（以及为未来一年制定的其他对象存储）读取/写入的能力。随着Apache Kudu的GA，用户现在可以使用Impala对接收到或更新的数据立即进行高性能分析。另外，也很容易将现有的商务智能（BI）工作负载从传统分析数据库或数据仓库迁移至由Impala构建的Cloudera分析数据库中，同时可以使用Navigator Optimizer优化其性能。而且如同以往一样，对于更大的并发性工作负载的性能改进仍然是全年工作的重中之重。

除了这些性能改进之外，随着越来越多的企业机构（例如纽约证券交易所（NYSE）和奎斯特诊断公司（Quest Diagnostics））已经注意到Cloudera现代分析数据库（而不是传统分析数据库）的灵活性、可扩展性和支持SQL及非SQL工作负载（例如数据科学、机器学习和操作性工作负载）的开放式架构，Impala的采用率也在不断增长。

对于该基准测试而言，我们使用未经修改的多用户TPC-DS查询对具有Impala的Cloudera现代分析数据库与传统分析数据库（Greenplum）进行了性能比较。我们还研究了分析数据库与SQL-on-Hadoop引擎，例如：Hive LLAP、Spark SQL和Presto的对比。总的来说，我们发现：

- Impala相对于传统分析数据库而言性能更为先进，包括超过8倍的高并发工作负载性能。
- 分析数据库和其他SQL-on-Hadoop引擎之间存在显著的性能差异，使用Impala可以使多用户工作负载的性能提高近22倍。
- 其他SQL-on-Hadoop引擎也无法完成大规模基准测试来与分析数据库进行比较，因此需要一个简化的、规模较小的基准测试（Hive甚至还需要修改，Presto无法完成多用户测试）。

比较集

分析数据库（采用10TB和1TB级别的数据进行测试，未经修改的查询）。

- Impala 2.8 from CDH 5.10；
- Greenplum Database 4.3.9.1。

附加的SQL-on-Hadoop引擎（采用1TB级别的数据进行测试，并对Hive进行了一些查询修改）。

- Spark SQL 2.1；
- Presto 0.160；
- Hive 2.1 with LLAP from HDP 2.5。

配置

每一个集群由七个工作节点组成，每个节点采用以下配置：

- CPU：2 块E5-2698 v4 @ 2.20GHz；
- 存储器：8 块2TB硬盘；
- 内存：256GB内存。

我们配置了三个由相同硬件组成的集群，其中一个用于Impala、Spark和Presto（负责运行CDH），另一个用于Greenplum，还有一个用于具有LLAP（负责运行HDP）的Hive。每个集群都装载了相同的TPC-DS数据：针对Impala和Spark的Parquet/Snappy，以及针对Hive和Presto的ORCFile/Zlib，而Greenplum使用内部的柱状格式与QuickLZ压缩文件。

查询

工作负载：

- 数据：TPC-DS 10TB和1TB（比例系数）；
- 查询：TPC-DS v2.4查询模板（未经修改的TPC-DS）。

我们运行了77个查询，所有引擎的运行都具有语言支持，无需修改TPC-DS规范（Hive除外）。1其中22个已排除的查询都使用以下几个不常见的SQL功能：

- 使用ROLLUP进行的11个查询（TPC-DS允许的变体在本测试中未使用）；
- 3个INTERSECT或EXCEPT查询；
- 8个具有高级子查询位置的查询（例如HAVING子句中的子查询等）。

由于Hive对子查询位置的更大限制，我们被迫进行了一些修改以创建语义上等同的查询。我们针对Hive运行了这些经过修改的查询。

虽然Greenplum、Presto和Spark SQL也声称支持所有99个未经修改的查询，但是即使没有并发执行，Spark SQL和Presto也无法成功完成10TB级别的99个查询。Greenplum随着多用户并发性的增加而出现越来越多的查询失败（详见下文）。

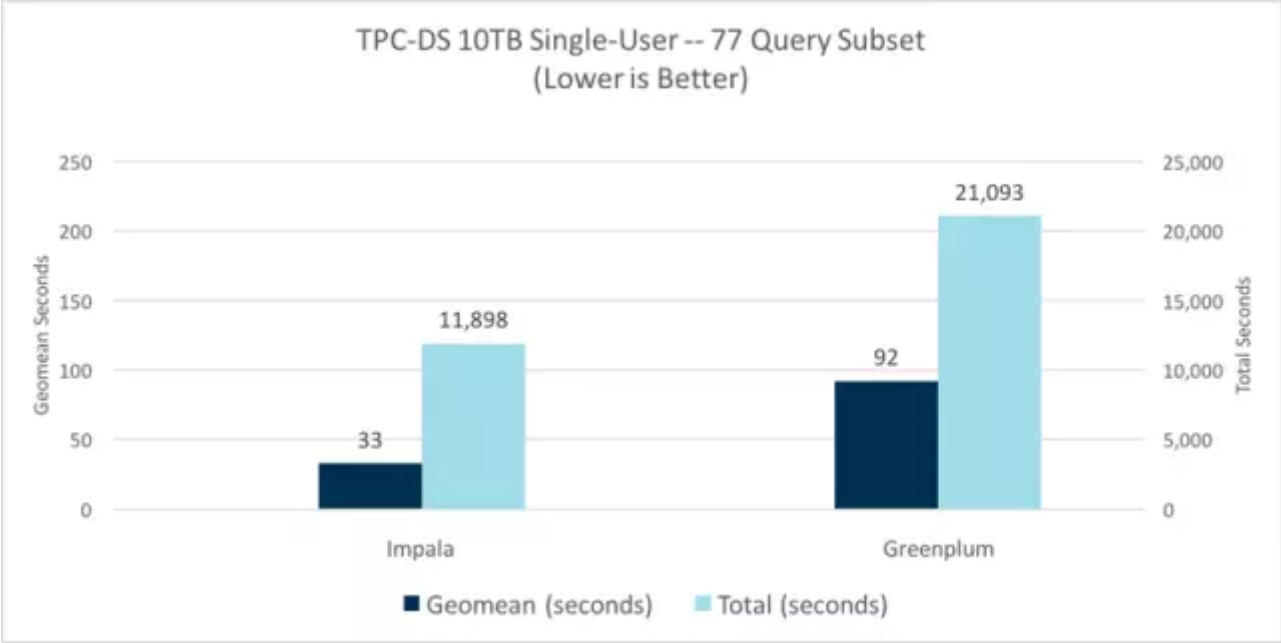
分析数据库基准测试结果

10 TB级别上Impala与Greenplum的比较

我们使用常见的77个未修改的TPC-DS查询在10TB级别数据下对Impala和Greenplum进行了测试。在单用户测试和更实际的多用户测试集上比较了两个、四个和八个并发流。总结如下：

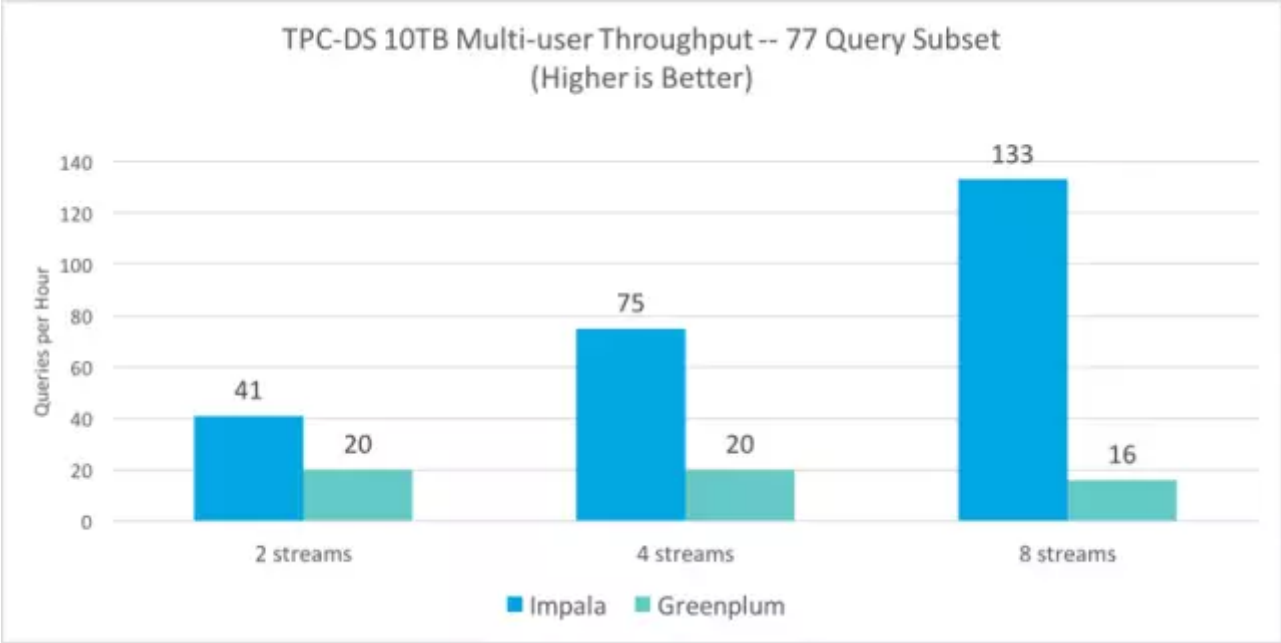
- 总体来说，Impala在单用户和多用户并发测试方面优于Greenplum。
- 相比Greenplum而言，Impala 线性扩展表现更优异，随着并发度增加，Impala与Greenplum的性能比率从2倍上升到了8.3倍，，同时保持了更高的成功率。

在单用户测试中，当比较查询中的几何平均值时，Impala的性能是Greenplum的2.8倍；完成查询流的总时间是Greenplum的1.8倍：



对于多用户吞吐量比较，我们使用TPC-DS dsqgen工具运行同一组77个未修改的查询来生成并发查询流。每个查询流由随机排序的77个通用查询组成，并且每个查询流使用不同的查询替换值。我们运行了多个测试，增加了超过系统饱和点的查询流数量，并且测量了各个并发级别的所有后续查询的吞吐量。

如下图所示，与Greenplum相比，Impala的性能指标随着并发速度从2个查询流2倍的速度提升加速到8个查询流8.3倍的速度提升。



鉴于集群的规模与数据集大小和并发性相比较而言较小，对于Impala和Greenplum这两个系统而言，预期在并发性增加时会发现一些查询失败。Impala和Greenplum这两个系统在两个查询流测试中达到了100%的成功率。对于四个和八个查询流测试，Impala系统的平均成功率为97%，而Greenplum系统的成功率下降到50%。如果这些测试在大于7节点集群的集群上运行，则可以预期这两个系统的成功率都会相应提高。

分析数据库与SQL-on-Hadoop引擎

1TB基准测试

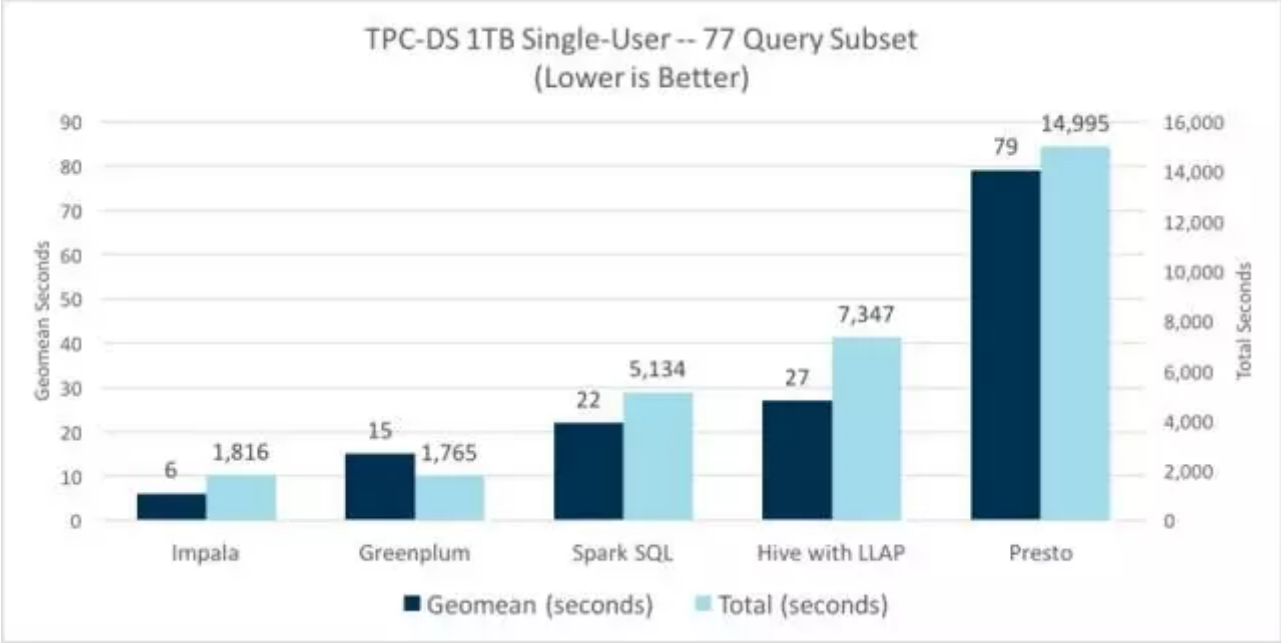
我们已经尝试针对SQL-on-Hadoop引擎使用相同的77个查询和10TB级别基准测试，但是，Hive、Presto和Spark SQL都无法成功完成77个未修改查询中的大多数查询，甚至仅仅是单用户结果也未能成功，因此无法在10TB级上进行比较。

因此，我们在1TB规模下运行了单独的比较，将分析数据库引擎与其余的SQL-on-Hadoop引擎进行比较。除了Hive之外，所有的引擎都使用相同的77个TPC-DS查询，但是需要进行一些修改，以寻找方法去绕过这些限制条件，从而解决无法解析的子查询。

通过这些简化的标准（对于其他SQL-on-Hadoop引擎来说是非常必要的），我们再次对所有五个引擎进行了单用户测试和更为真实环境的多用户测试。测试结果汇总如下：

- 分析数据库 - Impala和Greenplum系统在各个并发级别展现出的性能都优于所有的SQL-on-Hadoop引擎。
- 随着并发性的提高，再次看到Impala在性能方面拔得头筹，是其他引擎的8.5倍 - 21.6倍。
- 在所有引擎中，Presto在单用户测试中表现出最慢的性能，甚至无法完成多用户测试。

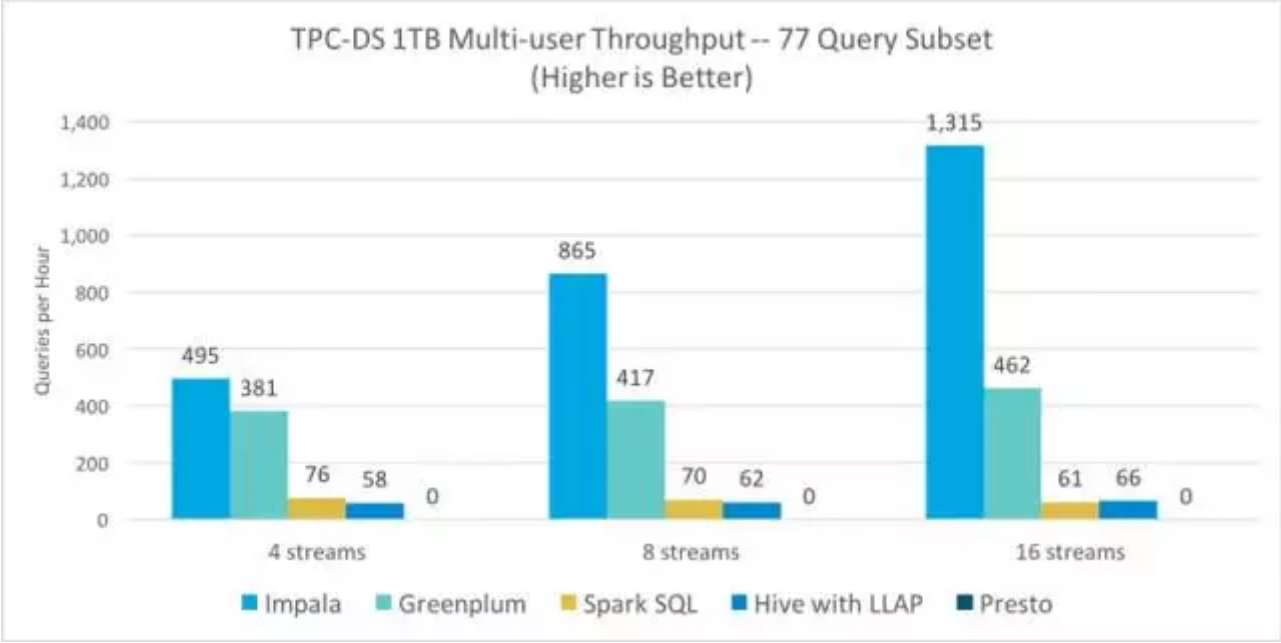
在该单用户测试中，我们再次看到，在几何平均值方面相较而言Impala仍然保持了其性能优势，但是，Greenplum在总时间上略有下降。这两个分析数据库的性能显着优于其他引擎，与其他SQL-on-Hadoop引擎相比，Impala在几何平均值方面性能优势在3.6倍至13倍之间，在总时间方面性能优势在2.8倍-8.3倍之间。



Presto对除了过滤、分组和聚合的简单单表扫描之外的其他常见的 SQL 查询表现的很挣扎。对于非常简单的查询类型，它更符合Spark SQL的性能，但是如上所述，对于使用更多标准SQL（包括连接）的更典型的商务智能（BI）查询，是执行效果最差的SQL-on-Hadoop集群。

使用TPC-DS通过四个、八个和十六个并发流运行更具代表性的多用户比较测试，以生成与上述的10TB分析数据库比较一样的随机查询流。除Presto之外，所有引擎都能够在三个并发级别的1TB级别下完成流，而不会出现任何查询失败。即使只运行四个并发查询，Presto也可能由于内存不足错误而使大多数查询失败。

对于能够成功完成多用户并发测试的引擎，分析数据库组群和SQL-on-Hadoop组群之间的性能差异变得更加明显。Impala在每一个并发级别上都展示出了优异的吞吐量 - 不仅比Greenplum快1.3-2.8倍，与Spark SQL相比，其速度快达6.5-21.6倍，并且比Hive快8.5-19.9倍。



结论

各企业机构越来越期待现代化改造其系统架构，但是不愿意牺牲重要商务智能（BI）和SQL分析所需的交互式、多用户性能。Impala作为Cloudera公司平台的一部分，能够独特地提供一个现代化分析数据库。通过设计，Impala可以灵活地支持更多种类的数据和使用案例，而无需任何前期建模工作；Impala可以有弹性地和成本高效地在公司内部部署和云端部署方式下按需进行扩展；并且，作为共享平台的一部分，这些相同的数据可用于其他团队和工作负载，而不仅仅只是SQL分析，因此可以进一步拓展其价值。此外，从上述基准测试结果可以看出，与传统分析数据库相比，Impala还提供了领先的性能。

无论整体性能还是大规模运算以及不断激增的并发性工作负载能力方面，分析数据库群（Impala、Greenplum）和SQL-on-Hadoop组群（Hive，Presto，Spark）之间的差异也变得非常明显。虽然其他SQL-on-Hadoop引擎不能满足分析数据库工作负载的要求，但这并不意味着对其他工作负载没有价值。事实上，绝大多数

Cloudera客户充分利用平台的开放架构，通过Hive准备数据，通过Spark建立和测试模型，通过Impala运行商务智能（BI）并提供报告，而无需在不同的孤岛中复制数据。

在接下来的一年中，我们将以Impala为核心继续推动现代化分析数据库的重大性能改进，包括增加商务智能（BI）体验的智能化和自动化，并且不断扩大云计算支持，进一步提高多租户能力和可扩展性。请点击此博客了解更多详情。

像往常一样，我们鼓励您通过基于开放式基准测试工具包运行您自己的基准测试以独立验证这些结果。



请点击 **“阅读全文”** 进入微站
(更多技术干货、行业动态，请关注【微站】，不定时更新)



[阅读原文](#)