



OLAP数据库引擎选型白皮书(2022)

OLAP Selection White Book (2022)

All for data , data for all.

目录

Contents

1.0 背景 01

- 1.1 OLAP数据库引擎概述 ----- 01
- 1.2 OLAP数据库引擎选型过程中存在的问题 ----- 02
- 1.3 如何更好地对数据库引擎进行选型？ ----- 02
- 1.4 报告特色 ----- 03

2.0 测评结果 04

- 2.1 测评结果表 ----- 04
- 2.2 引擎评述 ----- 05
- 2.3 选型指南 ----- 11

附录 14

- 评分标准 ----- 14

1.1 OLAP数据库引擎概述

OLAP（Online Analytical Processing，联机分析处理），主要应用于大规模数据分析及统计计算，为决策提供数据支持。

OLAP数据库引擎贯穿于数据时代发展的 3 个阶段。无论在哪一阶段，当企业需要积累数据并进行数据分析时，就会面临 OLAP数据库引擎的选型问题。



1.2 OLAP数据库引擎选型过程中存在的问题

《中国数据库行业研究报告（2021年）》显示，2020年中国数据库市场总规模达 247.1亿元，同比增长16.2%。2020-2022中国数据库市场预计将呈高增长态势。

数据库行业持续发展的同时，在数据库引擎选型的过程中也存在着以下问题：

数据库引擎采购常见流程



1. 选项繁多，选型流程长

目前市面上可选择的数据库类型越来越多，数据库呈多元选择趋势。在进行数据库引擎选型时，可选项繁多，而目前市面上可参考的选型标准较少，选型流程长，耗费成本高。

2. 开源逐渐成为大趋势，选型时需进一步考虑的维度增加

开源数据库的成熟度开始逐渐超过商业数据库。虽然避免了高昂的服务费及license费用，但开源数据库在易用性、配套能力等方面存在一定缺陷，会产生额外的开发、部署、迁移等成本。因此，在选型过程中，除了考虑性能之外，运维成本、数据安全、生态影响力等维度也需要纳入考虑范围。

1.3 如何更好地对数据库引擎进行选型？

在选型过程中，针对不同的业务场景和环节，分维度对不同数据库引擎进行测试和比对，能有效提升数据库引擎选型的效率。选型时需要综合考虑业务需求、性能、维护成本、数据安全等多种维度的信息。

具体而言，至少需要考虑以下三个层面：

1. 充分考虑业务相关性

测试业务场景要与企业的实际业务场景类似，这样提供的信息才能有效有效辅助决策者做出适合业务的决策；

2. 根据实际业务场景选择合适度量指标

基准测试一般有多个度量指标，不同业务场景关注的指标会有较大区别。在设计指标时候，需要充分考虑业务关注点；

3. 保障测试数据真实有效

构造的数据集过于规则容易测试出较高的指标。但真实的环境，数据是有瑕疵和倾斜的。因此，在数据构造阶段需要充分理解客户的业务场景，并可尽量模拟出贴近真实应用场景的数据。

如何通过测试比对，找到最适合的数据库引擎？



业务相关性



指标多维性



数据真实性

1.4 报告特色

本次测评在数据构造及测试过程中充分考虑到泛零售的业务场景，数据构造方案来源于真实客户业务场景，评测维度充分考虑了引擎的易用性及可维护性。

有以下特色：

1. 来自最佳实践

奇点云深耕泛零售领域，曾服务泛零售行业近 50% 头部客户，在该领域有非常丰富的项目实施经验。测评过程中构造数据时，充分考虑了泛零售行业客户在选型时关注的主要业务场景，如泛零售数据计算、BI报表、不同规模数据即席查询等；

2. 关注更多维度

除基本的性能测评外，本次测评还关注到了数据安全、运维成本、元数据等维度；

3. 客观中立

通过客观的数据建模方式进行计算，整体结论客观中立；

4. 专注于 OLAP

整体指标及选型模型设计时更加贴合 OLAP 场景。



2.0

测评结果

2.1 测评结果表

测评结果如下图，共 11 个分析维度

引擎 \ 维度	全表scan性能	事实表查询性能	宽表查询性能	TP查询性能	AP计算性能	数据批量同步速率	高并发	数据安全	运维成本	生态影响力	元数据
MySQL	D	D	D	C	C	A	D	D	D	C	无
TiDB	A	G	G	A	A	D	G	A	A	D	D
ClickHouse	G	G	C	G	G	G	G	D	A	C	G
StarRocks	G	G	A	A	G	G	C	C	A	C	A
Impala+Kudu	G	G	G	A	A	C	C	D	B	C	G
Presto+Kudu	A	D	D	D	B	C	C	D	B	D	G

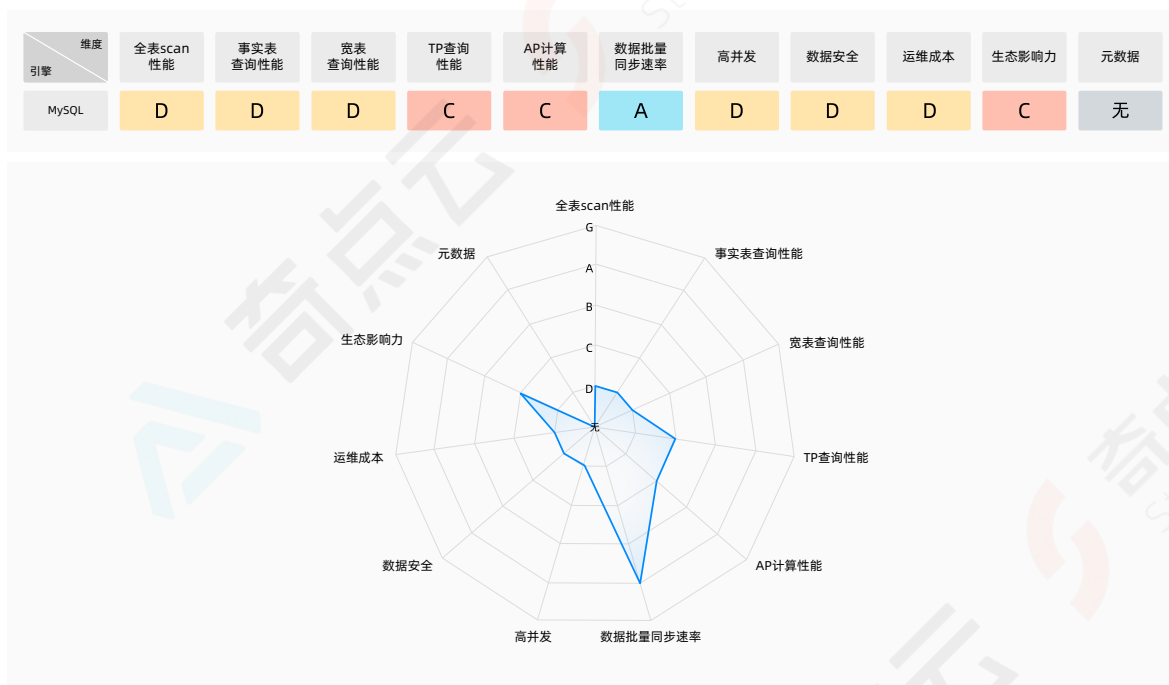
以上评分从高到低依次为：

G A B C D

2.2 引擎评述

2.2.1 MySQL

1. 测评结果



2. 分析引擎评述

基本介绍：

MySQL是目前最流行的开源数据库，其本身属于 OLTP引擎，但是也具备一些 OLAP计算的能力。协议被众多其他引擎采用，被广泛地应用在 Internet上的中小型网站中。

总评：

MySQL在即席查询性能和数据批量同步维度表现较好，其他维度表现一般，不支持统一元数据管理。整体而言，MySQL体积小、速度快、成本低，开源。单机版易操作，单机性能好，适合中小型表，业务量低于亿级别，对计算和查询性能都有一定需求的客户可以选择 MySQL。

优点：

1. 支持事务；
2. 低成本易上手：体积小，部署方便，资源占用少。技术成熟，使用标准 SQL数据语言形式，易上手；
3. 兼容性较好，适配性强：支持 Linux、Mac OS、OS/2 Wrap、Solaris、Windows等多种操作系统，BI软件基本都支持 MySQL数据源。

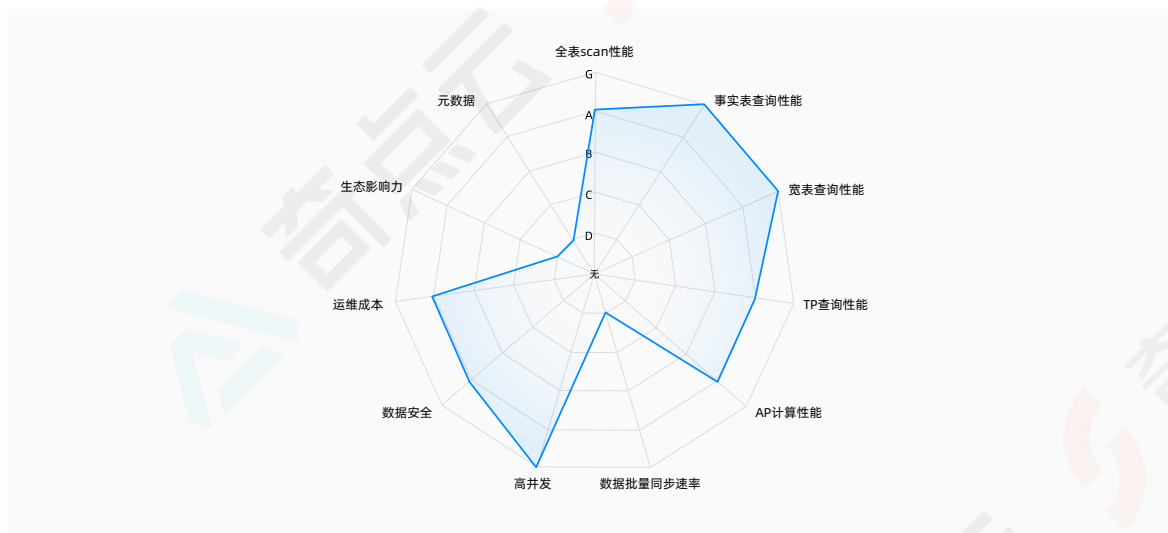
缺点：

1. 开发运维成本高：对存储过程和触发器支持不够良好，开发和维护存储过程比较难。且不支持热备份；
2. 海量数据处理的时候效率不高，单机连接并发数上线较低，单表数据量千万查询性能下降，复杂 SQL执行效率低；
3. 版本更新慢。

2.2.2 TiDB

1. 测评结果

引擎 \ 维度	全表scan性能	事实表查询性能	宽表查询性能	TP查询性能	AP计算性能	数据批量同步速率	高并发	数据安全	运维成本	生态影响力	元数据
TiDB	A	G	G	A	A	D	G	A	A	D	D



2. 分析引擎评述

基本介绍:

TiDB由 PingCAP公司开发和支持,是一个开源的 NewSQL数据库,支持混合事务和分析处理 (HTAP) 工作负载,与 MySQL兼容,并且可以提供水平可扩展性、强一致性和高可用性。TiDB是一个综合性的引擎,兼顾 OLAP和 OLTP,本次测评时主要考虑 TiDB的 OLAP场景。

总评:

TiDB事实表查询、宽表查询、即席查询和 AP计算性能方面表现亮眼,运维成本较低,无统一元数据管理。TiDB解决了 MySQL单机和中小型数据库表的瓶颈,高并发、AP、TP、即席查询性能均衡。

优点:

1. 支持一键水平扩容或者缩容,支持高可用;
2. 云原生的分布式数据库;
3. 兼容 MySQL 5.7 协议和 MySQL 生态;
4. 自带运维监控系统。

缺点:

1. 布署成本高,集群模式决定了需要更多的硬件成本;
2. 不支持分区,存储过程。

2.2.3 ClickHouse

1. 测评结果



2. 分析引擎评述

基本介绍：

ClickHouse是用于联机分析处理的开源列式数据库，允许分析实时更新的数据，以高性能为目标。该引擎为 OLAP查询而设计，常见应用场景有服务器日志分析等。该项目于 2016年 6月开源。

总评：

ClickHouse在全表 scan、事实表查询、TP查询、AP计算、数据批量同步速率性能方面表现亮眼，支持统一元数据并且有落地场景，运维成本相对低。该引擎在并发和 TP查询维度一般，DataX集成速度快，适合大规模批处理计算场景。

优点：

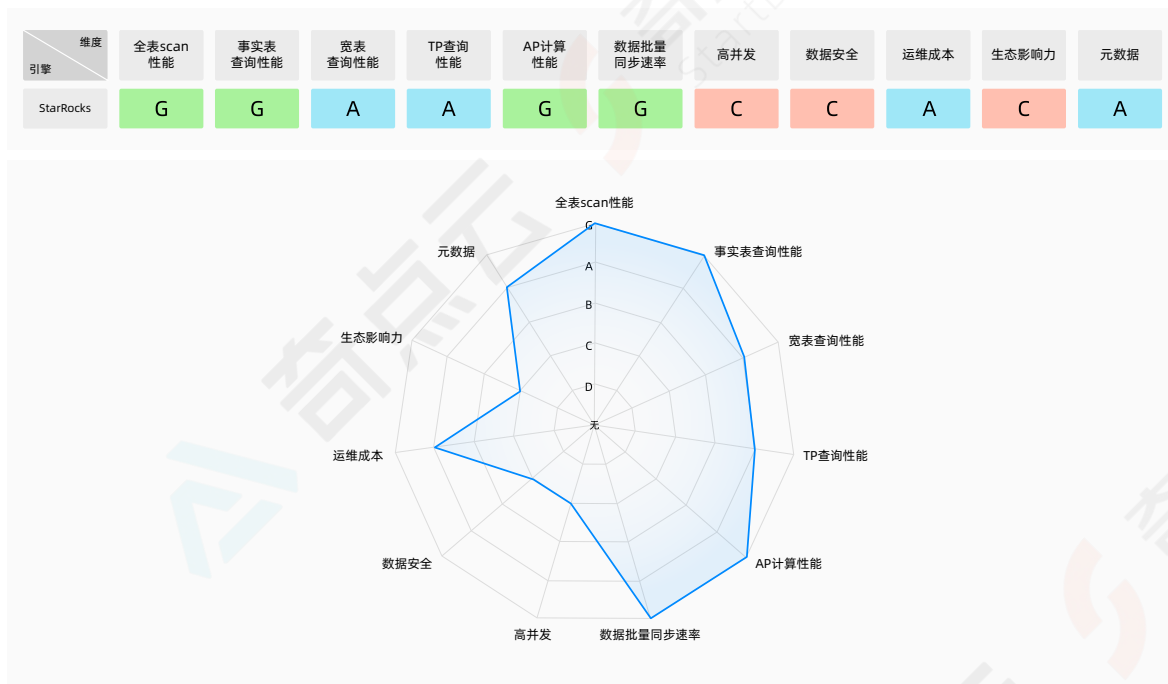
1. 列式数据库管理系统：写入速度快，在一些其他系统中也可以将不同的列分别进行存储；
2. 数据压缩：数据压缩空间大，减少 IO，处理查询高吞吐量，每台服务器秒级数十亿行；
3. 资源利用率高，适合在线查询；
4. 实时的数据更新，支持近似计算；
5. 不依赖 Hadoop复杂生态。

缺点：

1. 不支持事务；
2. 不擅长根据主键按行粒度查询；
3. 不擅长按行删除数据。

2.2.4 StarRocks

1. 测评结果



2. 分析引擎评述

基本介绍:

StarRocks主要应用于实时分析场景，于 2018年在 Apache社区开源。是一款面向多种数据分析场景、兼容 MySQL协议的分布式关系型列式数据库。

总评:

StarRocks在数据查询处理性能方面整体表现亮眼，支持统一元数据管理，运维成本相对低。该引擎 load接口不稳定，常出现删除数据库空间无法释放的情况，并发性能中等，其他表现不弱于TiDB。

优点:

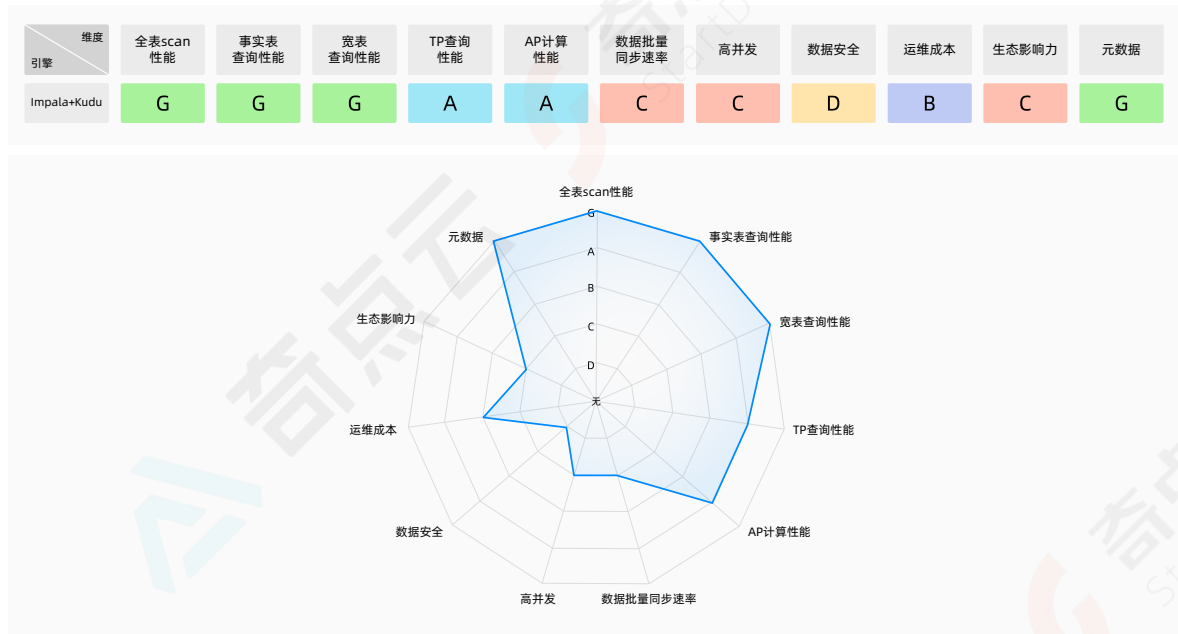
1. StarRocks 并不依赖于大数据生态，但其外表联邦查询可兼容大数据生态；
2. 提供了多种不同模型，能够支持不同维度的数据建模；
3. 支持在线弹性扩缩容，可以自动负载均衡；
4. 支持高并发分析查询，在单表查询及多表 join 方面综合性能均较优；
5. 实时性好，支持数据秒级写入；
6. 兼容 MySQL 5.7 协议和 MySQL 生态。

缺点:

1. 周边生态比较不完善；
2. 部分 SQL语法不支持。

2.2.5 Impala+Kudu

1. 测评结果



2. 分析引擎评述

基本介绍：

Impala是 Cloudera主导开发的查询系统，Impala并没有自己的存储引擎，不提供数据存储服务，但底层可集成多个数据源。Kudu 是针对 Apache Hadoop 研发的列式存储管理器，支持横向扩展和高可用。Impala和 Kudu深度集成，数据可实时写入 Kudu，Impala提供 BI分析 SQL查询。

总评：

在即席查询场景下，Impala的稳定性和速度已在工业界经过广泛验证。Impala+Kudu架构支持随机读写，有良好的 Scan性能，对 Spark等流式计算框架有官方客户端支持，为实时数据仓库存储提供了良好的解决方案。适用于数据量不大、有一定的并发量、需要快速响应的场景。

优点：

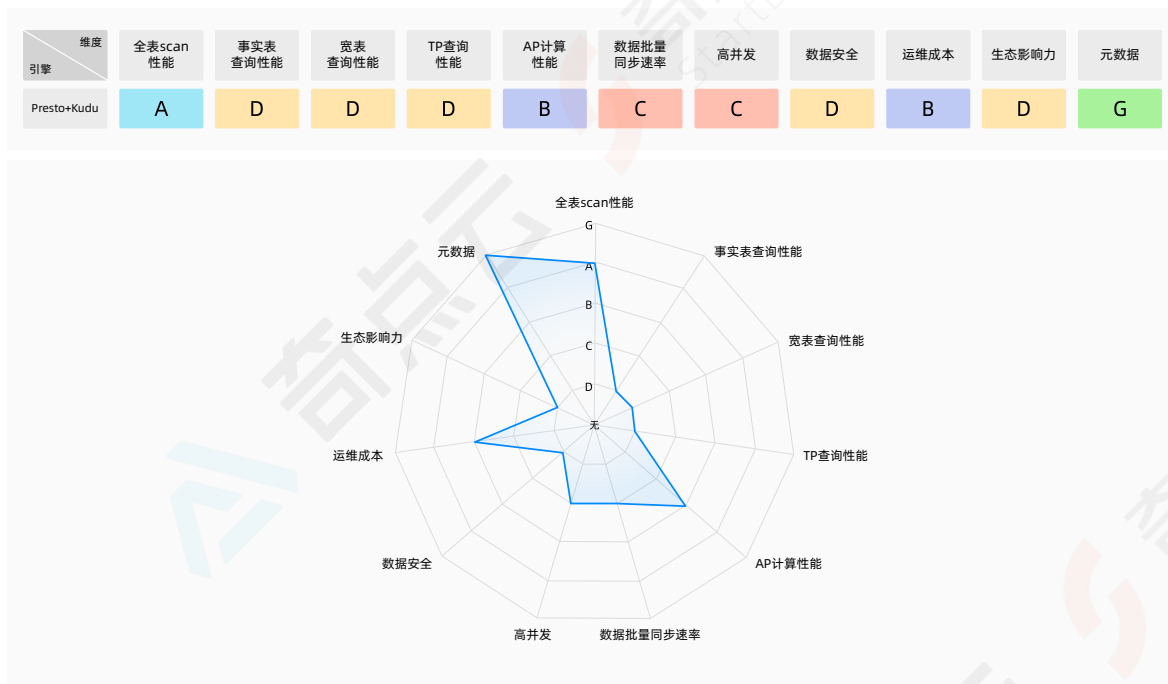
1. Impala SQL语法与和 Hive SQL高度相似，学习成本低；
2. Impala具备超大数据规模 SQL解析能力，能够高效利用 CPU与内存，快速返回结果；
3. Impala与 Hue深度集成，提供可视化的 SQL操作以及 work flow；
4. Kudu能完成与关系型数据库类似的操作，数据可以存储在 Kudu里并随时更新。

缺点：

1. Impala基于内存计算，整体而言对内存依赖性比较大；
2. Impala基于 C++编写，引擎维护难度较大；
3. Impala不支持 ANSI SQL，在 SQL聚合等方面功能相对较弱；
4. Kudu提高读性能的同时牺牲了写性能。其表必须设置主键，对于非主键列的过滤条件，只能进行全表扫描，性能较差。

2.2.6 Presto+Kudu

1. 测评结果



2. 分析引擎评述

基本介绍：

Presto是用于大数据场景的高性能分布式 SQL 查询引擎，用户能基于该架构查询各种数据源，并且可以在单个查询中查询来自多个数据源的数据。Presto本身不存储数据，通常配合存储管理器使用（本次测试与 Kudu 配合使用）。

总评：

Presto支持 GB到 PB级数据的秒级查询场景，通过使用分布式查询，可以高效完成海量数据的查询，擅长对海量数据进行复杂分析，在遍历全表、元数据管理维度表现较好。Presto可以接入多种数据源，支持跨数据源的级联查询。

优点：

1. 数据源支持丰富：支持的数据源包括传统关系型数据库、图数据库、Hive、Redis等；
2. 支持连接多个数据源，并且能够进行跨数据源连表查询，可以在一条查询中对来自多个数据源的数据进行合并分析；
3. Presto根据场景进行聚合运算，基于内存运算，根据场景聚合运算（如 Count，Avg等），边读数边计算，再清内存，再读数据计算。和传统的 MapReduce相比，消除了延迟和磁盘 IO 开销。

缺点：

1. Presto支持 ANSI SQL，但在对 DML的支持方面相对较弱；
2. 跨数据源连表查询时有可能产生大量临时数据，查询时占用存储多，查询速度会变慢。

2.3 选型指南

在项目交付实施过程中，我们总结出了进行 OLAP数据库引擎选型的通用流程。参考下方流程，基本可选出符合业务需求的 OLAP数据库引擎。



1. 引擎与 BI工具是否适配

企业通常已有正在使用的 BI工具，在进行 OLAP 选型的时候需要首先考虑对 BI的支持度。BI使用 OLAP主要用于探索性报表、报表深度分析等。

1. 首先考虑 BI工具是否能够顺利连接到该 OLAP 引擎；
2. 在能够顺利连接的基础之上，看哪一个引擎对该 BI工具支持得最好，查询数据响应最快。建议重点关注的指标有：「全表 scan性能」、「事实表查询性能」、「宽表查询性能」、「TP 查询性能」、「AP 计算性能」。

引擎 \ 维度	全表scan性能	事实表查询性能	宽表查询性能	TP查询性能	AP计算性能
MySQL	D	D	D	C	C
TiDB	A	G	G	A	A
ClickHouse	G	G	C	G	G
StarRocks	G	G	A	A	G
Impala+Kudu	G	G	G	A	A
Presto+Kudu	A	D	D	D	B

BI工具适配场景建议关注指标

2. 成本：包含硬件成本、运维成本等

1. 硬件成本：泛零售领域对硬件成本控制比较严格，因此需要重点考虑部署引擎的硬件成本（本次测评暂不涉及硬件成本数据，如对相关信息感兴趣，可联系奇点云相关人员）；
2. 运维成本：考虑运维便利程度，如交付后续客户无法自主运维，则还需要后续的运维支出。建议关注指标：「运维成本」，运维成本指标等级越高，运维便利性越好。

引擎 \ 维度	运维成本
MySQL	D
TiDB	A
ClickHouse	A
StarRocks	A
Impala+Kudu	B
Presto+Kudu	B

成本场景建议关注指标

3. 批处理场景（离线计算）

批处理场景对时效性要求并不高，但需处理的数据量极大，因此对查询性能及数据同步速率等有要求。典型的应用场景有人群计算、产出标签和产出报表等。

1. 优先考虑即席查询相关指标，包括「事实表查询性能」、「宽表查询性能」、「高并发」。OLAP本身就面向即席查询，因此即席查询相关的性能越高越好。

引擎 \ 维度	事实表查询性能	宽表查询性能	高并发
MySQL	D	D	D
TiDB	G	G	G
ClickHouse	G	C	G
StarRocks	G	A	C
Impala+Kudu	G	G	C
Presto+Kudu	D	D	C

批处理场景建议关注指标（1）

2. 在即席查询相关指标满足业务需求的基础上，进一步考虑「TP查询性能」、「数据批量同步速率」；

引擎 \ 维度	TP查询性能	数据批量同步速率
MySQL	C	A
TiDB	A	D
ClickHouse	G	G
StarRocks	A	G
Impala+Kudu	A	C
Presto+Kudu	D	C

批处理场景建议关注指标（2）

3. 进一步考虑「全表scan性能」。

引擎 \ 维度	全表scan性能
MySQL	D
TiDB	A
ClickHouse	G
StarRocks	G
Impala+Kudu	G
Presto+Kudu	A

批处理场景建议关注指标（3）

4. 流计算场景（实时计算）

流计算场景对数据返回实时性要求比较高，关注瞬时流量带宽，及最长延时的底线。典型场景有注册发券、订单同时等。以注册发券场景为例，用户进行注册之后根据用户所选偏好、基础信息、当前行为等信息立刻推送优惠券，该场景下需要对用户的行为数据进行快速查询计算，数据返回延时低。选型考虑以下方面：

1. 引擎与消息队列的兼容性：首先考虑引擎是否能对接 Kafka之类的消息队列；
2. 进一步关注「事实表查询性能」、「宽表查询性能」；

引擎 \ 维度	事实表查询性能	宽表查询性能
MySQL	D	D
TiDB	G	G
ClickHouse	G	C
StarRocks	G	A
Impala+Kudu	G	G
Presto+Kudu	D	D

流计算场景建议关注指标（1）

3. 此外流计算场景（实时计算）还需关注「AP 计算性能」指标。

引擎	维度 AP计算性能
MySQL	C
TiDB	A
ClickHouse	G
StarRocks	G
Impala+Kudu	A
Presto+Kudu	B

流计算场景建议关注指标（2）

5. CDP产品兼容性

如客户需要对接 CDP 产品（如 DataNuza），则需要考虑 CDP 产品与 OLAP 引擎的兼容性及对接成本。

6. 学习成本

学习成本指的是，选型时需要考虑引擎的生态影响力及学习难度：

1. 生态影响力：生态影响力包含了商业化程度、热度等，生态影响力大的引擎，引擎的周边配套软件多。在行业内应用会较多，学习渠道及素材多，使用过程中出现的问题会有相应的解决方案；

引擎	维度 生态影响力
MySQL	C
TiDB	D
ClickHouse	C
StarRocks	C
Impala+Kudu	C
Presto+Kudu	D

生态影响力指标

2. 学习成本：考虑是否支持 SQL、是否需要学习额外的语言等（如 ClickHouse 需要学习额外语法）。

在各项指标中，数据安全、运维成本、生态影响力、元数据等指标为引擎的基础能力项，其等级越高，引擎本身的安全及易用性越高。在满足业务场景所需达到的性能且成本允许的前提下，建议尽可能选择基础能力等级较高的引擎。

引擎	维度 数据安全	运维成本	生态影响力	元数据
MySQL	D	D	C	无
TiDB	A	A	D	D
ClickHouse	D	A	C	G
StarRocks	C	A	C	A
Impala+Kudu	D	B	C	G
Presto+Kudu	D	B	D	G

基础能力项指标

附录

评分标准

1. 测试后根据以下评分表对每个细项进行评分，得出原始指标

角度/指标		分数	50 (G)	40 (A)	30 (B)	20 (C)	10 (D)	0
遍历全表	遍历全表		平均小于1s	平均小于3s	平均小于10s	平均小于30s	平均大于30s	NA
查询能力	亿级别事实表		小于200ms	小于300ms	小于500ms	小于1000ms	大于1000ms	NA
	十亿级别事实表		小于1s	小于3s	小于6s	小于10s	大于10s	NA
	二十亿级别事实表		小于2s	小于5s	小于10s	小于20s	大于20s	NA
	宽表 (3E)		小于1s	小于3s	小于5s	小于10s	大于10s	NA
OLTP	update		case平均小于5s	case平均小于20s	case平均小于50s	case平均小于100s	case平均大于100s	NA
	group by		case平均小于30s	case平均小于100s	case平均小于300s	case平均小于500s	case平均大于500s	NA
	join查询		case平均小于10s	case平均小于30s	case平均小于100s	case平均小于200s	case平均大于200s	NA
	select		case平均小于30s	case平均小于80s	case平均小于200s	case平均小于500s	case平均大于500s	NA
	with		case平均小于30s	case平均小于80s	case平均小于200s	case平均小于500s	case平均大于500s	NA
	实时计算框架		NA	支持Flink、Spark	NA	支持Flink或Spark	NA	不支持
	DataX导入 (5并发)		大于20M/s	大于15M/s	大于10M/s	大于5M/s	小于5M/s	NA
	存储计算		存储空间：小于1T	存储空间：小于1.5T	存储空间：小于2T	存储空间：小于2.5T	存储空间：小于3T	存储空间：大于3T
高并发	select		qps：平均1万+	qps：平均5000+	qps：平均1000+	qps：平均大于500	qps：平均小于500	NA
	select join		qps：平均100+	qps：平均30+	qps：平均10+	qps：平均3+	qps：平均小于3	NA
事务	事务		NA	支持	NA	NA	NA	不支持
云平台兼容性	云平台兼容性		NA	主流云平台全部兼容	NA	NA	主流云平台存在不兼容	不支持云平台
安全	数据权限分级		支持库、表、行、列、字段	NA	支持库、表、行、列	NA	支持库、表	NA
	用户权限分级		支持库、表、行、列、字段	NA	支持库、表、行、列	NA	支持库、表	NA
	数据审计		NA	系统支持	NA	外部支持	NA	不支持
	自主可控		国产&商业&开源&底层自研	国产&商业&开源	国产开源	国产商业化	NA	国外
	信创认证		NA	已认证	NA	NA	NA	未认证
	存取控制		资源、SQL、数量、字段	资源、SQL、数量	资源&数量	资源	NA	不支持
运维	部署&升级&修改		部署简单、升级简单&修改可以配置无需重启	部署难度一般、升级难度一般&修改可以配置无需重启	部署难度一般、升级难度一般&修改可以配置无需重启	部署难度复杂、升级复杂&修改可以配置无需重启	部署难度复杂、升级复杂&修改需要重启服务	NA
	迁移		命令和工具一键迁移	NA	数据同步迁移	NA	需要导出重新导入	NA
	SQL协议		完全支持	主体支持	NA	只支持部分协议	NA	不支持
稳定性	高可用		多节点部署，自动分发	NA	多节点部署，不支持自动分发	NA	多节点部署，节点小于10个	单节点
	伸缩性		弹性伸缩，热配置	NA	配置伸缩，需要重启	NA	部署伸缩，需要停服务	不支持
	容灾恢复		故障自动恢复，自动保存灾前数据	NA	自动保存灾前数据	NA	手动恢复	不支持
	监控		引擎自带，监控齐全无需配置	引擎自带，监控需要配置	NA	安装第三方	NA	无合适的监控
	自动备份		定时自动备份、故障自动备份	NA	故障自动备份	定时自动备份	NA	无
商业化	商业化		周边配套软件多，专利数多	周边配套软件多	有周边配套软件	有周边配套软件	NA	无
元数据	元数据		支持，并有落地场景	支持，并有落地方案	支持	NA	NA	不支持
热度	百度热度		结果数大于1E，日指数大于1万	结果数大于1E，日指数大于5000	结果数大于3千万，日指数大于1000	结果数大于1千万，日指数大于300	结果数大于1千万，日指数小于300	NA
	github热度		github项目数大于50万	github项目数大于10万	github项目数大于1万	github项目数大于1000	github项目数小于1000	NA

- 2.将原始指标在每个维度下进行加总平均，平均分即为每个维度下的分数。

得分 = 维度下的原始指标分数总和 / 维度下的原始指标数量，例如，事实表查询性能分数 = (亿级别得分 + 十亿级别得分 + 二十亿级别得分) / 3

角度	原始指标	分析维度（因素）
遍历全表	遍历全表	全表scan性能
查询能力	亿级别事实表	事实表查询性能
	十亿级别事实表	事实表查询性能
	二十亿级别事实表	事实表查询性能
	宽表（3E）	宽表查询性能
OLTP	update	TP查询性能
	group by	TP查询性能
	join语句	TP查询性能
OLAP	select	AP计算性能
	with	AP计算性能
	实时计算框架	AP计算性能
	DataX导入（5并发）	数据批量同步速率
高并发	select	高并发
	select join	/
事务	事务	/
安全	数据权限分级	数据安全
	用户权限分级	数据安全
	数据审计	数据安全
	自主可控	数据安全
	信创认证	数据安全
	存取控制	数据安全
运维	部署&升级&修改	运维成本
	迁移	运维成本
	存储计算	运维成本
	SQL协议	运维成本
云平台兼容性	云平台兼容性	运维成本
稳定性	高可用	运维成本
	伸缩性	运维成本
	容灾恢复	运维成本
	监控	运维成本
	自动备份	运维成本
商业化	商业化	生态影响力
热度	百度热度	生态影响力
	github热度	生态影响力
元数据	元数据	元数据

► 3. 将上表结果中每个维度的得分换算为等级，得到最终评级

等级	分数段
G	[40,50]
A	[30,40)
B	[25,30)
C	[20,25)
D	(0,20)
无	0

维度 引擎	全表scan 性能	事实表 查询性能	宽表 查询性能	TP查询 性能	AP计算 性能	数据批量 同步速率	高并发	数据安全	运维成本	生态影响力	元数据
MySQL	D	D	D	C	C	A	D	D	D	C	无
TiDB	A	G	G	A	A	D	G	A	A	D	D
ClickHouse	G	G	C	G	G	G	G	D	A	C	G
StarRocks	G	G	A	A	G	G	C	C	A	C	A
Impala+Kudu	G	G	G	A	A	C	C	D	B	C	G
Presto+Kudu	A	D	D	D	B	C	C	D	B	D	G



400-080-0326

www.startdt.com

business@startdt.com

浙江省杭州市余杭区良睦路 1399 号梦想小镇互联网村 15、17 幢