

SQL计算引擎综述

吴建军

2019.5

目录

- 数据库概述
- 关系数据库及其扩展
- SqlOnHadoop
- NoSQL与NewSQL
- 其他SQL数据库
- 周边组件简介
- 参考文献

wujianjunml@outlook.cn

数据库概述

数据库概述

- SQL简介：
 - SQL是关系模型的第一个实现。
 - 最早在1974被IBM实现。
 - 在1986年成为ANSI标准，在1987年成为ISO标准。
 - 有许多基于标准SQL且类似的语言，如pgSQL，HSQL等。
 - SQL目前非常流行[13]：

May 2019	May 2018	Change	Programming Language	Ratings	Change
1	1		Java	16.005%	-0.38%
2	2		C	14.243%	+0.24%
3	3		C++	8.095%	+0.43%
4	4		Python	7.830%	+2.64%
5	6	▲	Visual Basic .NET	5.193%	+1.07%
6	5	▼	C#	3.984%	-0.42%
7	8	▲	JavaScript	2.690%	-0.23%
8	9	▲	SQL	2.555%	+0.57%
9	7	▼	PHP	2.489%	-0.83%
10	13	▲	Assembly language	1.816%	+0.82%

数据库概述

- 查询分类:

- OLTP: 一般只会访问少量的记录, 且大多时候都会利用索引。
- OLAP: 一般需要Scan大量数据, 常只访问部分列, 聚合 (Sum/Count等)计算多于明细。

- On-line Transaction Processing:

- Short-lived txns.
- Small footprint.
- Repetitive operations.

```
SELECT * FROM useracct  
WHERE userName = ?  
AND userPass = ?
```

```
UPDATE useracct  
SET lastLogin = NOW(),  
hostname = ?  
WHERE userID = ?
```

```
SELECT P.*, R.*  
FROM pages AS P  
INNER JOIN revisions AS R  
ON P.latest = R.revID  
WHERE P.pageID = ?
```

```
INSERT INTO revisions  
VALUES (?, ?, ?)
```

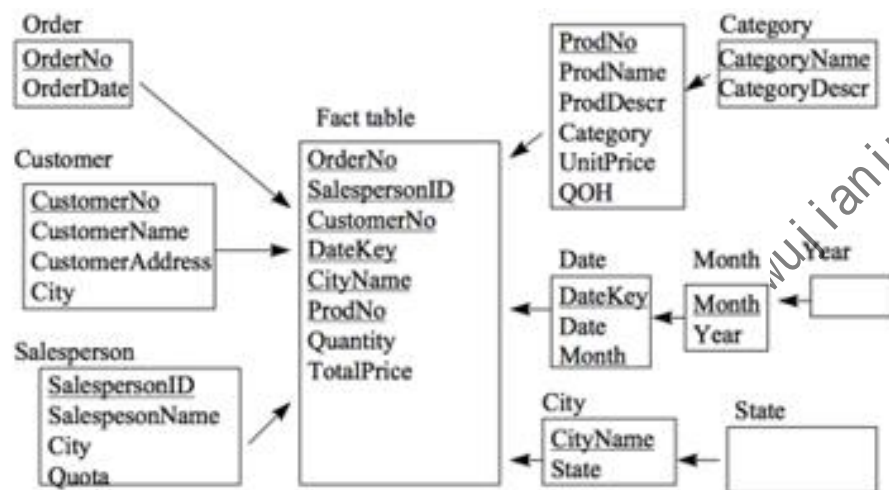
- On-line Analytical Processing:

- Long running queries.
- Complex joins.
- Exploratory queries.

```
SELECT COUNT(U.lastLogin),  
EXTRACT(month FROM U.lastLogin) AS month  
FROM useracct AS U  
WHERE U.hostname LIKE '%.gov'  
GROUP BY EXTRACT(month FROM U.lastLogin)
```

数据库概述

- 数仓模型：
 - 星型模型: 星型模型中的表分为事实表和维表。
 - 雪花模型: 将维表进一步拆为子维表。



dim_product table

product_sk	sku	description	brand	category
30	OK4012	Bananas	Freshmax	Fresh fruit
31	KA9511	Fish food	Aquatech	Pet supplies
32	AB1234	Croissant	Dealicious	Bakery

dim_store table

store_sk	state	city
1	WA	Seattle
2	CA	San Francisco
3	CA	Palo Alto

fact_sales table

date_key	product_sk	store_sk	promotion_sk	customer_sk	quantity	net_price	discount_price
140102	31	3	NULL	NULL	1	2.49	2.49
140102	69	5	19	NULL	3	14.99	9.99
140102	74	3	23	191	1	4.49	3.89
140102	33	8	NULL	235	4	0.99	0.99

dim_date table

date_key	year	month	day	weekday	is_holiday
140101	2014	jan	1	wed	yes
140102	2014	jan	2	thu	no
140103	2014	jan	3	fri	no

dim_customer table

customer_sk	name	date_of_birth
190	Alice	1979-03-29
191	Bob	1961-09-02
192	Cecil	1991-12-13

dim_promotion table

promotion_sk	name	ad_type	coupon_type
18	New Year sale	Poster	NULL
19	Aquarium deal	Direct mail	Leaflet
20	Coffee & cake bundle	In-store sign	NULL

数据库概述

- 数据库的演进:
 - SQL数据库：关系模型是在1970年E.F.Codd博士首先提出的，通过SQL语言进行CURD操作，并保证事物的ACID性质。典型软件有MySQL，PostgreSQL。
 - NoSQL数据库：泛指非关系型数据库，一般满足CAP和BASE两条原则，2000年开始流行。NoSQL又分为：KV存储(Memcache/Redis)，列存储(Hbase/Cassandra)，文档数据库(MongoDB)，图数据库(Neo4j)，搜索引擎(Elasticsearch/Solr)等。
 - NewSQL数据库：2011年被首次提出。NoSQL对海量数据的存储管理能力强大，但是对ACID和SQL支持不佳。而RDBMS虽然有着ACID和SQL，但是对海量数据比较乏力。NewSQL希望结合两者的优点。NewSQL系统包括：Spanner，TiDB，OceanBase。

数据库概述

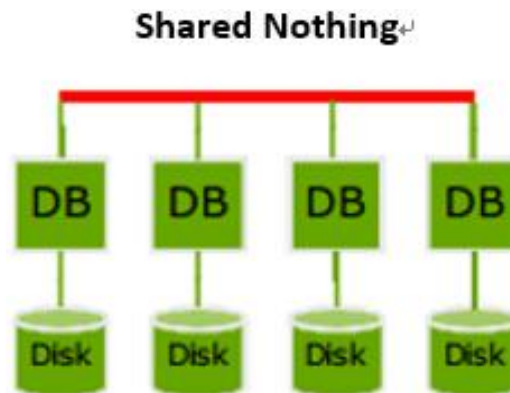
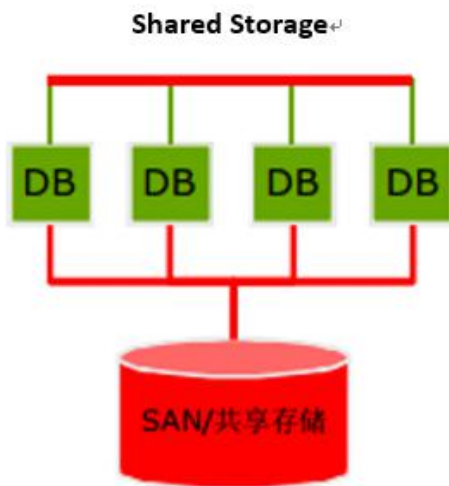
- 明星公司：
 - Cloudera: 2008 年成立, 2009年hadoop的创始人Doug Cutting加入。主要产品为CDH(Cloudera的hadoop发行版) 等。
 - Hortonworks: 2011年成立, 雅虎工程副总裁、Hadoop开发团队负责人出任Hor首席执行官。主打产品是Hortonworks Data Platform (HDP) 。
 - MapR: 2009年成立。MapR认为Hadoop存在很多问题, 他们用新架构重写HDFS。
 - Databricks: 由Spark创建者们于2013年成立, 主要从事Spark的产品化, 已与微软联合, 在Azure上推出了基于Spark的数据分析产品。
 - Pivotal: 成立于2013年, 主要产品是Greenplum, HAWQ, 也是许多java开发组件的主要开发方, 如Spring, Tomcat, Redis, RabbitMQ。

数据库概述

- 学术会议：
 - SIGMOD：数据库的最好会议。已经有30年的历史。SIGMOD在数据库领域具有最高学术地位的国际性学术会议。
 - VLDB：也已经有30年的历史，它是唯一能接近SIGMOD的会议。
 - ICDE：明显比SIGMOD/VLDB差一个档次，但又明显比其他的数据库会议高一个档次。
 - CIDR：数据库领域的新兴会议，强调创新性的，已成为很受重视的一个数据库会议。
- DB-Engines：
 - DB-Engines搜集了市面上的数据库管理系统(DBMS)。
 - DB-Engines排行榜按目前的流程度即人气指数来对数据库进行排名。
 - DB-Engines这个排名在业界引用得非常多，权威性也很高，总体来说比较客观。

数据库概述

- DB集群架构：
 - Shared Storage: 多个Node共享一个Storage, 但各自拥有自己的Memory/CPU。缺点是多个节点在Storage的竞争降低整个并行数据库系统的性能。
 - Share Nothing: 每个node拥有自己的memory/disk/CPU, 彼此之间通过网络进行通讯和数据传输。缺点是node之间的通讯开销就会变得很大。

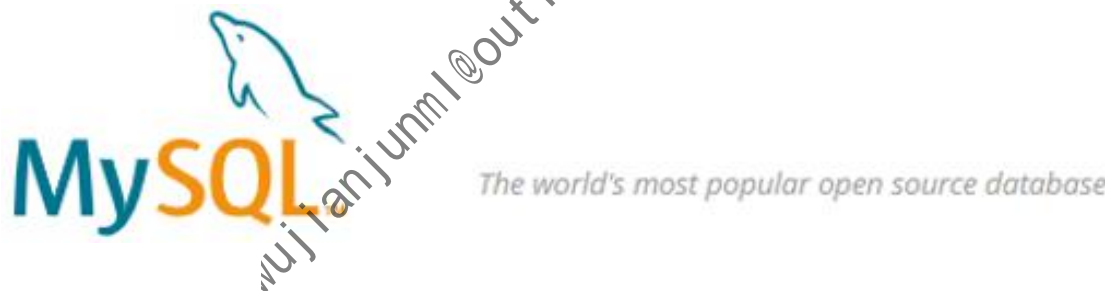


关系数据库及其扩展

关系数据库及其扩展

- **MySQL**

- 原本是一个开源的关系数据库，1995年5月发布初始版本。原开发者为瑞典的MySQL AB公司，MySQL AB公司于2008年被Sun收购。2009年，Oracle收购Sun。被Oracle收购后，MySQL的创始人以MySQL为基础，成立MariaDB。



- **PostgreSQL**

- 1994年，两名UC Berkeley学生在Ingres的基础上创建了Postgres95，1996年被重命名为PostgreSQL。PostgreSQL每年一个主要版本发布。PostgreSQL从编程角度是最高级的数据库系统，具有非常多的高级特性。

PostgreSQL: The World's Most Advanced Open Source
Relational Database

关系数据库及其扩展

• PostgreSQL vs MySQL

- PG支持更多数据类型，支持json, array, jsonb 等。
- PG编程能力更强悍：支持OVER 子句，有非常丰富的统计函数(如grouping sets)和统计语法支持，支持多种语言编写存储过程(Python, psql等)。
- PG大表join能力大大强于MySQL。
- PG的GIS功能碾压MySQL。
- PG有很强大的fdw功能，支持查询70多种外部数据源。
- PG的性能在不少场景下是MySQL的3-4倍，其他场景下两者性能相当。
- GP现在未受商业公司控制。
- GP编程难度稍高一点，学习成本比MySQL高，语法比MySQL严格。

关系数据库及其扩展

• Greenplum

- Greenplum是基于PostgreSQL的MPP扩展。
- 国内不少使用Teradata的企业已经迁移到GP上。
- 查询性能很高(跟Impala一个级别), 且SQL能力强大。
- 与hadoop兼容性较差, 很难平滑扩容。
- 非apache顶级项目。



• HAWQ

- HAWQ也是基于PostgreSQL的MPP扩展。
- 存储使用HDFS。
- 性能号称很高(但是缺乏严谨的第三方证据)。
- 用户不多(几乎没有国外用户), 资料较少(维基百科都没有), 18年底才从apache毕业, 开发团队已离开Pivotal单干。



关系数据库及其扩展

- 还存在非常多的基于PostgreSQL的数据库[1]:

Name	Vendor	License	Availability	Notes
AgensGraph	Bitnine	Apache2	2016	PostgreSQL + Graph Model features
HadoopDB	Yale University	Apache License V2.0	2009-....	PostgreSQL + shared-nothing cluster + Hadoop
PipelineDB	PipelineDB	GPL v3	2015-....	Streaming SQL
Redshift	Amazon	Private/Cloud-based	2013-....	Data Warehouse on AWS
Vertica	HP	proprietary	2005-....	Column-oriented DataWarehouse

SQL-On-Hadoop

SQL-On-Hadoop

- SQL-On-Hadoop简介

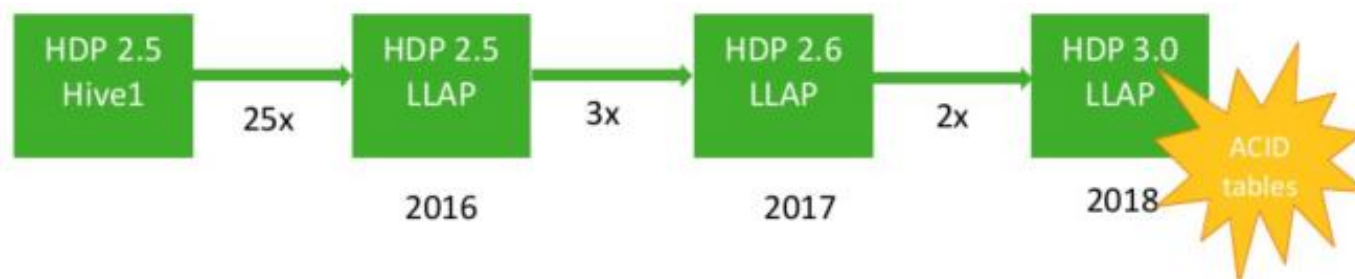
- 定义：指那些支持以类SQL语法查询存储于HDFS中数据的一类引擎。
- Hive就是最早的SQL-on-Hadoop系统。
- 一般不能支持全部SQL语法，但是要求至少支持group by, join, UDF等。
- 有很多被广泛使用的开源方案：Hive, Impala, Presto, SparkSQL。
- 还有一些其他不是很出名的方案：Drill(mapR主导, 2015年5月推出首个正式版本)。Tajo(韩国公司Gruter主导, 2015年10月以后再无新版本)。
- 右图显示了几个SQL数据库的排名[2].

Rank			DBMS	Database Model	Score		
May 2019	Apr 2019	May 2018			May 2019	Apr 2019	May 2018
9.	↑ 10.	↑ 11.	Hive 🏠	Relational	77.90	+3.19	+20.93
22.	↓ 21.	22.	Spark SQL	Relational	16.28	-0.90	+4.88
23.	↓ 22.	↓ 20.	Impala	Relational, Multi-model ⓘ	14.93	+0.41	+2.06
24.	↓ 23.	24.	Greenplum	Relational, Multi-model ⓘ	12.58	+0.13	+2.22
36.	↑ 37.	↑ 65.	Presto	Relational	5.05	+0.76	+4.11
44.	↑ 46.	↓ 38.	Apache Drill	Multi-model ⓘ	2.97	+0.27	+0.18
78.	↓ 77.	↓ 77.	HAWQ	Relational	0.93	-0.03	+0.23
111.	↓ 108.	↑ 121.	Tajo	Relational	0.30	-0.03	+0.23

SQL-On-Hadoop

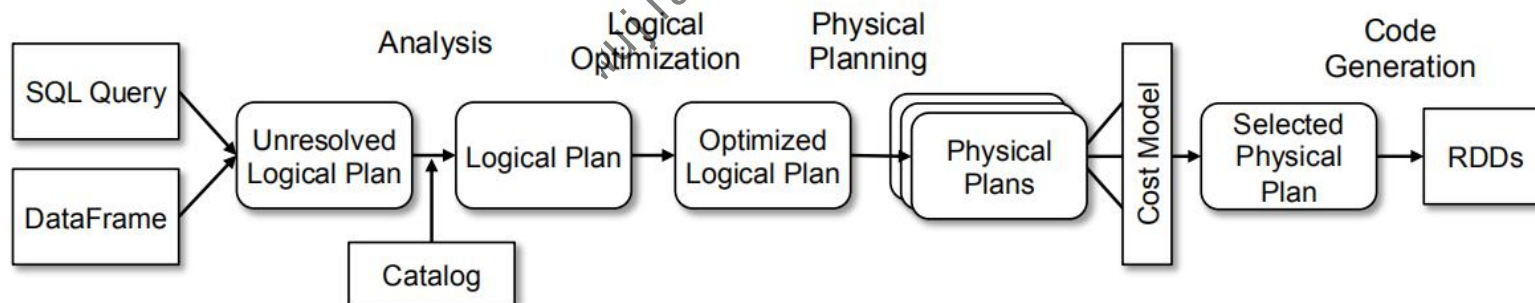
- Hive

- Hive最初由Facebook发起，2010年10月发布首个版本(0.3)。2016年2月发布了2.0版本。2018年5月发布了3.0版本。
- 支持的存储系统包括HDFS，HBase。
- 支持MapReduce，Tez和Spark三种计算框架。
- 支持很多种文件格式：文本, RCFile, ORC, Parquet等。
- 具有极佳的扩展性和容错性。
- 在2.0种引进了 LLAP[3]，可以大幅提升一些短小型查询的性能。
- Hive支持许多加速技术，如向量化，CBO等，其性能一直在提高。



SQL-On-Hadoop

- SparkSQL
 - Spark的一个组件。
 - 在Spark1.0中被开发，Spark1.3后可以被生产使用。
 - 兼容Hive。
 - 最初采用Shark作为执行引擎。
 - 2014年单独开发了一套执行引擎。
 - Catalyst[4]是其核心，可以利用它实现许多优化策略。

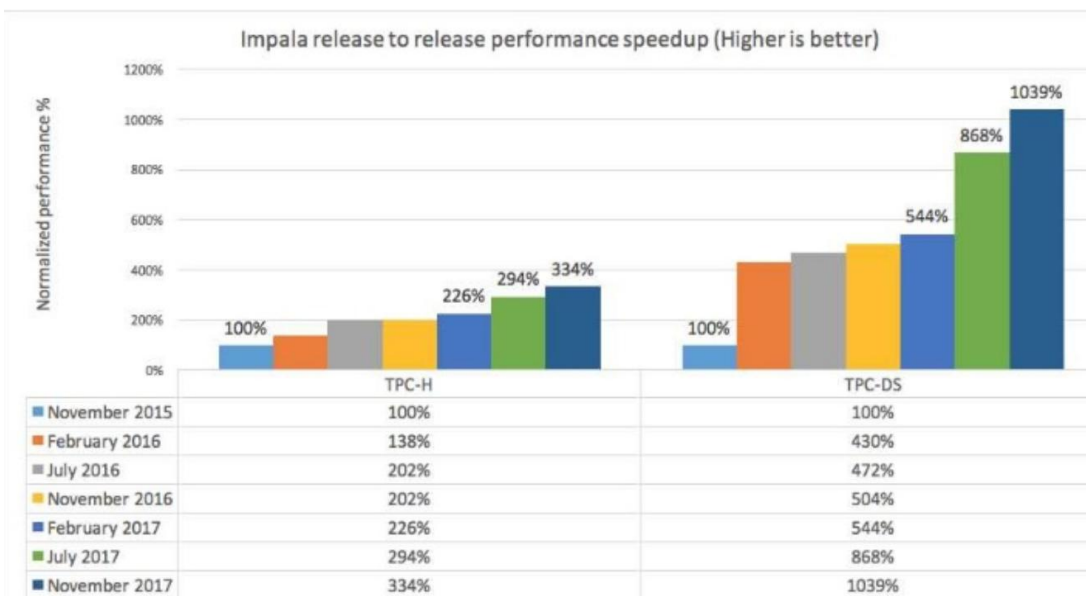


- 在Spark2.2中增加CBO。

SQL-On-Hadoop

- Impala

- 最初由cloudera主导，2012年开始开发，2013年5月出现首个可用版本。
- 应该是最早的生产级别的hadoop-MPP方案。
- 支持的存储系统有：HDFS，HBase等。
- 被广泛使用，得到AWS，MapR等公司支持。
- 支持各类Join算法，窗口函数，子查询。
- C++实现，充分利用硬件特性。
- 利用LLVM实现了动态代码生成。
- Impala性能一直稳步增长[5]:



SQL-On-Hadoop

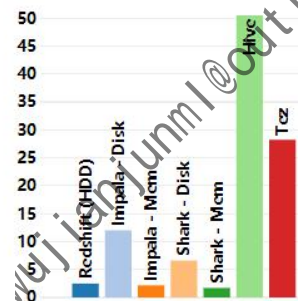
- Presto
 - 2013年底由Facebook开源。
 - 支持丰富的数据源，HDFS，MySQL，Cassandra，MongoDB，Elasticsearch等。
 - 非Apache项目，目前主要由Presto Software Foundation和StarBurst维护。
 - 在国内外被广泛使用。
 - 采取很多加速手段，列式存储，向量化，CBO等[6]。
 - MPP-style **pipelined** in-memory execution
 - **Columnar** and **vectorized** data processing
 - Runtime query **bytecode compilation**
 - Memory **efficient** data structures
 - Multi-threaded multi-core execution
 - Optimized readers for **columnar formats** (ORC and Parquet)
 - Now also **Cost-Based Optimizer**

SQL-On-Hadoop

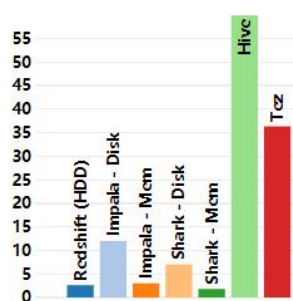
- 测试对比
 - 我们先看看UC Berkeley AMPLab在2014年做的测试[7]:
 - Scan Query

- Aggregation Query

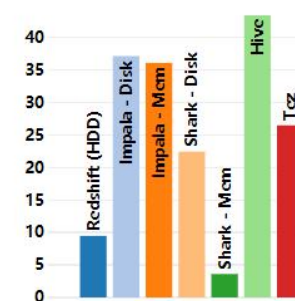
Query 1A
32,888 results



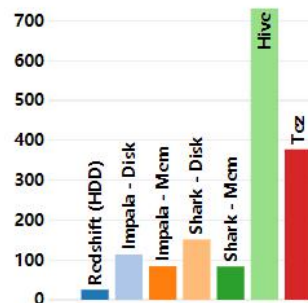
Query 1B
3,331,851 results



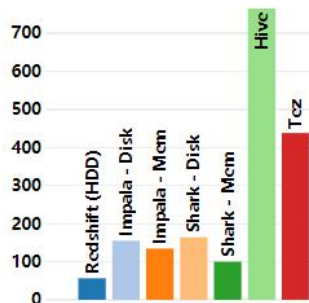
Query 1C
89,974,976 results



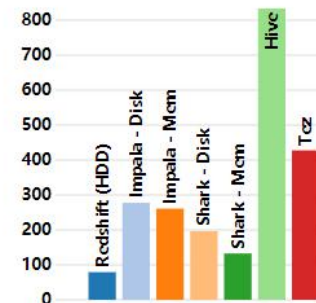
Query 2A
2,067,313 groups



Query 2B
31,348,913 groups



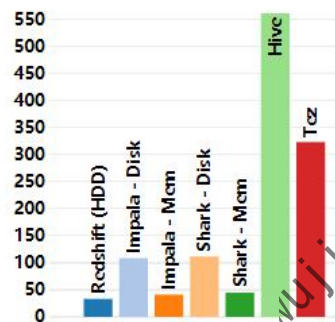
Query 2C
253,890,330 groups



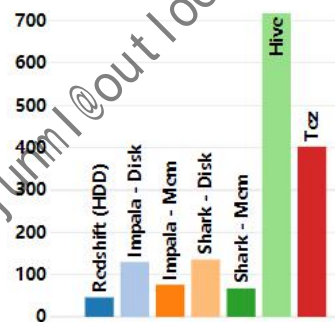
SQL-On-Hadoop

- 测试对比:
 - Join Query

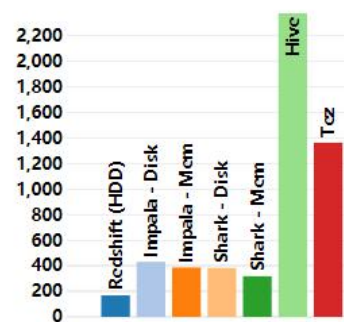
Query 3A
485,312 rows



Query 3B
53,332,015 rows



Query 3C
533,287,121 rows



- 这是Shark团队做的，所以不能证明shark很强，但是可以证明Impala远超hive/hive on tez，也可以看出很多场景跟Redshift不遑多让。

SQL-On-Hadoop

- 测试对比：
 - 这里我们再看看一份最新的(2018年)来自 Vanderbilt University 的对比测试[8]。
 - 对比测了 Impala, SparkSQL, Drill。
 - 可以看到 Impala+Parquet 在大多数用例下都是性能最高的。
- 哪怕在最不利的测试中, Impala 都未曾处于下风。

TPC-H Query No.	TPC-H Scale Factor - 125					
	Impala		Spark SQL		Drill	
	Text	Parquet	Text	Parquet	Text	Parquet
Q1	37.68	37.79	24.8	4.86	33.33	18.06
Q2	3.59	2.58	43.4	16.03	Failed	Failed
Q3	9.08	7.34	28.7	7.3	52.16	16.5
Q4	7.94	7.31	44.2	25.56	54.76	18.36
Q5	13.3	10.44	46.1	25.16	50.23	13.8
Q6	3	1.81	21.63	2.26	29.93	5.26
Q7	13.47	13.29	75.81	22.4	52.1	17.96
Q8	5.68	3.49	47.33	18.86	42.23	5.53
Q9	14.79	12.29	54.26	30.36	46.7	12.4
Q10	7.27	5.68	32.83	8.56	40.73	12.86
Q11	Failed	Failed	33.23	21.33	11.93	1.66
Q12	6.58	4.97	26.76	4.76	37.43	18.1
Q13	9.01	10.75	16.43	9.26	Failed	9.36
Q14	4.44	3.88	23.26	3.73	30.2	4.5
Q15	11.26	8.4	49.5	4.96	30.6	6
Q16	6.37	6.32	68.93	81	Failed	8.46
Q17	29.65	31.19	53.8	17.06	95.26	14.96
Q18	17.8	17.82	63.6	15.13	98.76	25.93
Q19	49.69	52.22	22.96	4.46	Failed	Failed
Q20	8.54	4.53	38.26	21.43	45.13	9.86
Q21	24.19	23.15	131.7	83.63	Failed	Failed
Q22	3.36	2.86	22.86	13.3	Failed	Failed
AM	-	-	44.11	20.06	-	-
AM-Q{2,11,13,16,19,21,22}	12.86	11.73	41.37	14.15	50.15	12.76
Normalized AM-Q{2,11,13,16,19,21,22}	1.26	1	5.83	1.87	6.75	1.68
Text (AM) over Parquet (AM)	1.26		3.11		4.01	

NoSQL与NewSQL

wujianjunm1@outlook.com

NoSQL与NewSQL

- NoSQL概览：
 - NoSQL一词自从1998年发明后，经历了长期的技术繁荣(或者说混乱，几乎任何能载入数据的软件都可以说成NoSQL)，时至今日，有非常多的NoSQL数据库。
 - 主要特性是：非关系型，无ACID保证，几乎无SQL支持，分布式，可扩展。
 - 又可以再分为几个类别：
 - 文档数据库：MongoDB, Couchbase等。
 - 图数据库：Neo4j, AllegroGraph等。
 - KV存储：memcached, Redis等。
 - 时序数据库：InfluxDB, OpenTSDB等。
 - 搜索引擎：ElasticSearch, Solr等。
 - 列式存储：HBase, Cassandra, Kudu等。
 - 还有一些类别难说的数据库：Druid(时序，列存储)，Kylin(预计算)等。
 - 这些NoSQL往往都提供SQL接口(他们SQL支持得往往不强)。

NoSQL与NewSQL

- Calcite:
 - 许多NoSQL没有自身的解析器，他们往往借助Calcite实现对SQL的解析。
 - Calcite是一款SQL解析器，他首先将SQL解析成AST，然后校验AST是否合法(如字段是否存在)并生成RelNode树，接着基于规则或基于代价优化RelNode树并将其转化成物理执行计划，最后将物理执行计划转化成可在**特定平台**执行的程序。
 - Calcite被许多项目使用[9]:

Used by



Connects to



NoSQL与NewSQL

- Druid SQL:

- 原生基于json的查询。
- 2017年支持SQL。
- 无法join。

```
[ EXPLAIN PLAN FOR ]  
[ WITH tableName [ ( column1, column2, ... ) ] AS ( query ) ]  
SELECT [ ALL | DISTINCT ] { * | exprs }  
FROM table  
[ WHERE expr ]  
[ GROUP BY exprs ]  
[ HAVING expr ]  
[ ORDER BY expr [ ASC | DESC ], expr [ ASC | DESC ], ... ]  
[ LIMIT limit ]  
[ UNION ALL <another query> ]
```

- Elasticsearch SQL

- 原生也是基于json的查询。
- 2018年9月支持SQL。
- 无法join。

```
SELECT select_expr [, ...]  
[ FROM table_name ]  
[ WHERE condition ]  
[ GROUP BY grouping_element [, ...] ]  
[ HAVING condition]  
[ ORDER BY expression [ ASC | DESC ] [, ...] ]  
[ LIMIT [ count ] ]
```

NoSQL与NewSQL

- Kylin SQL:
 - 一直使用SQL查询。
 - 仅仅支持INNER JOIN和LEFT JOIN。
 - 主要支持Dimension-Fact之间的join。
- Phoenix
 - 一个运行在HBase上的SQL框架。
 - 对SQL的支持比较全面，支持OLTP。
 - 对join的支持比较晚，支持Hash Join和Sort-Merge Join两种join算法[10]。
 - 然而，其性能不高(特别是超大表join)[8]，且需要将数据载入Hbase。

QUERY SYNTAX
SELECT
STATEMENT
EXPRESSION
SUBQUERY
JOIN
INNER JOIN
LEFT JOIN
UNION
UNION ALL

WDA Query No.		# of worker nodes - 2				# of worker nodes - 4				# of worker nodes - 8			
		Impala	Spark	Drill	P-HBase	Impala	Spark	Drill	P-HBase	Impala	Spark	Drill	P-HBase
Q1	Execution	0.28	5.5	9	0.3	0.28	6	16.5	0.34	0.28	7	30	0.62
	Write Output	3.13	0.5	2	8.7	6.03	0.9	3.5	14.66	11.52	2.1	5	26.38
	Total	3.41	6	11	9	6.31	6.9	20	15	11.8	9.1	35	27
Q2	Total	12.6	68.6	33	240	12.4	68	31.8	228	13.15	69.1	68	231
Q3	Total	14.8	67.1	34	14	15	70.1	35.1	17	15.12	84.6	72	40
AM		10.27	47.23	26	87.67	11.24	48.33	28.97	86.67	13.36	54.27	58.33	99.33
Normalized AM		1	3.89	2.7	7.52	1.28	4.04	3.58	7.84	1.84	4.61	6.79	9.63

NoSQL与NewSQL

- NewSQL概览：
 - 目标是支持NoSQL的高性能和扩展性，同时支持传统DB中的事务ACID。
 - 主要针对OLTP，不适合OLAP。
 - 对SQL的支持很弱。
 - 分区，复制，MVCC等技术是其核心。
 - google的Spanner，F1是NewSQL的经典之作。
 - 开源的数据库有：oceanbase、TiDB，VoltDB等。

其他SQL数据库

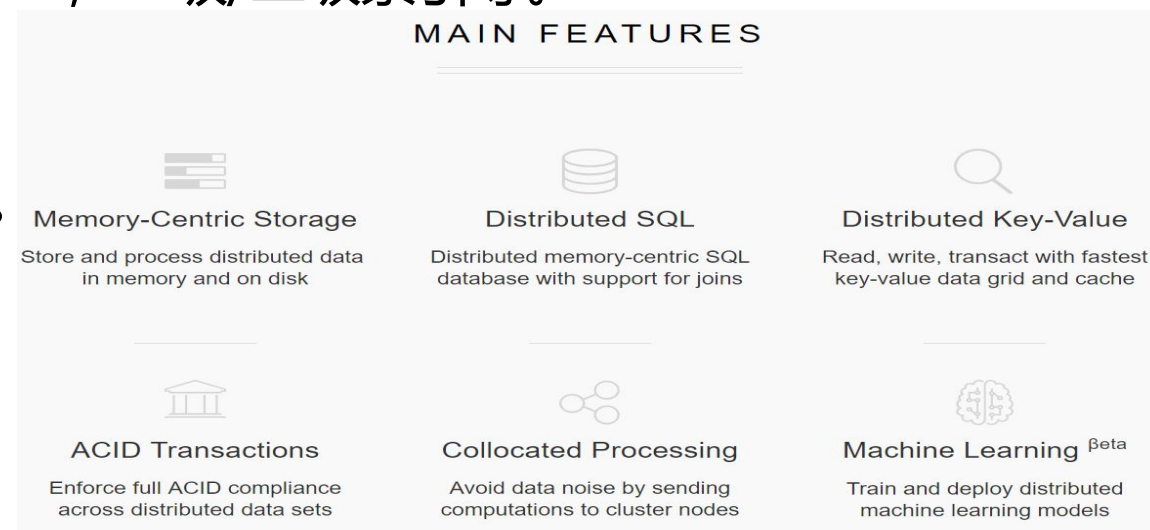
wujianjunm1@outlook.com

其他SQL数据库

- ClickHouse:
 - Yandex在2016年6月开源。
 - 主要针对时序数据分析。
 - 性能据说很高。
 - 文档资料较少，对维护者要求较高(用起来需要较强的开发能力)。
 - 不支持大表之间的join，且一次查询中仅支持一个join[11]。
- SQLite & Derby
 - 二者均是单机，嵌入式，基于内存的数据库。
 - 多用于在线程序数据读写，几乎不用于离线分析。

其他SQL数据库

- Ignite:
 - 一款分布式的内存SQL数据库。
 - 源于GridGain公司，2015年1月进入 Apache 的孵化器，当年8月毕业。
 - 完整的SQL支持，可以水平扩展，容错能力好。
 - 特别是对分布式Join得支持是其核心优势。
 - 完善的数据库特性：更新持久化，事务ACID，一级/二级索引等。
 - 号称真正的分布式关系数据库系统。
 - 可以支持高并发读写。
 - 还可以支持机器学习应用，甚至深度学习。
 - 比较复杂，需要更多的研究。
 - 目前使用不算广泛。



其他SQL数据库

- MapD:
 - 一款基于GPU的开源数据库。
 - 号称OLAP性能超过传统数据库一个数量级。
 - 目前仅仅运行于单机。
 - 被分析数据必须装的进显存。
 - 未见join测试结果[12]。

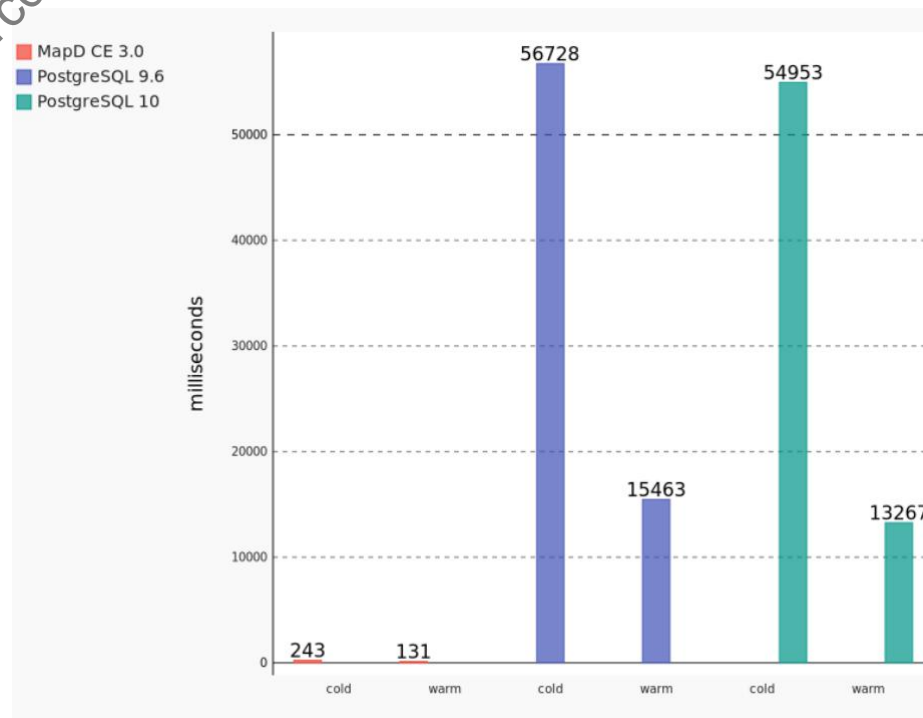


Figure 4.10.: Query 9

周边组件简介

数据库前端

- Hue:
- Hue是一个开源的Hadoop UI系统。
- 由Cloudera Desktop演化而来。
- 可以在浏览器上与Hadoop进行交互分析。
- 可以编写和运行Hive脚本以及查看运行结果
- 提供Impala的应用进行数据交互查询。

The screenshot displays the Hue web interface. On the left, a sidebar lists various tables and databases, including 'u_omain', 'free_logs', 'hue_commits', and 'jira_c'. The main area shows a Hive query being executed. The query is a SELECT statement that filters for tickets with multiple SFD cases and a specific status. Below the query, there's a section for 'Query History' and 'Saved Queries'. The 'Results' section shows a table with 19 rows, including columns for 'name', 'jira_summary_c', and 'tickets'. The table lists various tickets and their associated summaries.

Some high risks were detected.

```
-- the objective is to find the JIRAs in Hue where there are multiple SFD cases linked
-- it reveals the soft spots in the product

SELECT sfdc.jira_c.name,
       sfdc.jira_c.jira_summary_c,
       count(jira_c.name) AS tickets
FROM sfdc.cases, sfdc.jira_c, jira.ticket
WHERE sfdc.cases.component_c IN ('Hue')
AND sfdc.jira_c.case_c = sfdc.cases.id
AND jira.ticket.issuekey = sfdc.jira_c.name
AND jira.ticket.statusname NOT IN ('Resolved', 'Closed')
AND sfdc.jira_c.name NOT LIKE 'CLIK'
GROUP BY jira_c.name, jira_c.jira_summary_c
HAVING count(jira_c.name) > 1
ORDER BY count(jira_c.name) DESC
```

name	jira_summary_c	tickets
1	CDH-45011 Improve interaction between Hue and Impala	66
2	CDH-51313 Tracking Jira document2 upgrade 5.7 and below to 5.8 and above	47
3	OPSAPS-25666 Offer option in add service wizard to automatically set as a dependency for another service	16
4	CDH-46194 Security analysis for Hue security Jiras	15
5	CDH-46197 Improve integration between Hue and HiveServer2	14
6	OPSAPS-27028 Ease of Embedded DB Causing Frustration and Database Migration Asks	11
7	OPSAPS-39656 Hue needs Load Balancer parameter for SPNEGO auth	10
8	OPSAPS-28330 We should automatically add the value of ldap_username in Hue to hive and impala proxy users	8
9	OPSAPS-26574 Add 1 DAP Priorities for HRS and Impala to HUE	8

Suggestions
Query on partitioned table is missing filters on partitioning columns.
Rewrite query to add filtering conditions.

Improve Analysis

数据库前端

- Zeppelin:
 - 一款基本web的notebook工具。
 - 可以接入各种解释器来支持许多后端。
 - 支持不少后端：Hive，Spark等。
- Superset:
 - 由Airbnb开源的数据分析与可视化平台。
 - 数据可视化能力强大。
 - 支持许多后端：MySQL，Greenplum，Druid等。

周边组件简介

- Hadoop数据采集：
 - Sqoop：用于在关系型数据库和HDFS之间互相传输数据。
 - Flume：分布式、高可用的日志数据采集系统，可以写入HDFS。
 - Kafka：高性能消息中间件，支持HDFS写入。
- 流式处理引擎：
 - Storm：最早的流式引擎(2011年)。
 - Spark Streaming：Spark的一部分，用批处理模拟流式处理。
 - Flink：主攻流式计算，性能很高。
- 任务调度：
 - Oozie：xml的形式写调度流程。
 - azkaban：KV文件格式来描述依赖关系。
 - Airflow：通过 DAG 定义整个工作流。

参考文献

- [1]. https://wiki.postgresql.org/wiki/PostgreSQL_derived_databases.
- [2]. <https://db-engines.com/en/ranking/relational+dbms>
- [3]. <https://cwiki.apache.org/confluence/display/Hive/LLAP>
- [4]. Spark SQL: Relational Data Processing in Spark, [SIGMOD15](#)
- [5]. Performance Optimizations in Apache Impala, slideshare, 2018-01
- [6]. Presto query optimizer: pursuit of performance, slideshare, 2018-06
- [7]. <https://amplab.cs.berkeley.edu/benchmarks>
- [8]. A comparative analysis of state-of-the-art SQL-on-Hadoop systems for interactive analytics, arXiv18.
- [9]. https://calcite.apache.org/docs/powered_by.html
- [10]. <https://phoenix.apache.org/joins.html>
- [11]. https://clickhouse.yandex/docs/en/query_language/select/
- [12]. https://wiki.hsr.ch/Datenbanken/files/MSE_DB-Seminar_HSR_HS1718_GPU_Databases_S_Kurath.pdf
- [13]. <https://www.tiobe.com/tiobe-index/>

Q&A

wujianjunm1@outlook.com