

X S Q L

360分布式查询引擎

——技术揭秘与使用介绍

序

言

目前，数据库技术风起云涌。使用这些技术的方式差别很大，一些技术有非常陡峭的学习曲线，XSQL将致力于降低数据使用的门槛；另一方面，XSQL将致力于提供更加稳定的分布式查询服务。



目录页 | Contents

01



数据开发之痛

XSQL主要特点



02

03



XSQL技术架构

XSQL使用介绍



04



01

数据开发之痛

- 高门槛
- 不稳定



曲线

学习曲线陡峭

标准SQL
SQL-like
No-SQL

维护

可维护性差

依赖繁杂
可读性差

部署

部署不可控

部署频繁
向前兼容性

不稳定

环境问题

比如：内存限制、网络I/O限制、带宽打满、跨机房传输、CPU跑满、资源隔离等。



比如：数据倾斜、内存溢出、Full GC、空值等。

数据问题



02

XSQL主要特点

- 低门槛
- 更稳定
- 增效节能
- 数据联邦

低门槛

XSQL基于ANSI SQL-2003标准，现已支持绝大多数Hive SQL，并增加了很多MySQL方言。

XSQL目前已经提供了Rest和Thrift两种服务。



服务化



平台接入

XSQL跟云图、搜索的智能分析系统合作，提供了更加友好的大数据管理、查询功能。

XSQL 提供了类似于Hive、MySQL的命令行工具。



统一Cmd



统一API

XSQL继承了Spark的编程API，包括Java和Scala两种语言。



可视化

XSQL提供的可视化服务包括：云图、Zeppelin。搜索部门的THOR

更稳定

减少环境问题

未来将提供任务监控服务。



比如：自适应broadcast、分区合并、下推优化、CBO等。未来还将提供内存建议服务。

自适应优化



效率 (Efficiency)

提升开发和数据处理效率。

节能 (Economize)

节省计算和存储资源。

MongoDB数据源

```
SELECT s.*, c.*  
FROM mymongo.xsql_mongodb.student s,  
myes.xsql_index.class_room c  
WHERE s.class_id = c.id
```

Elasticsearch数据源



03

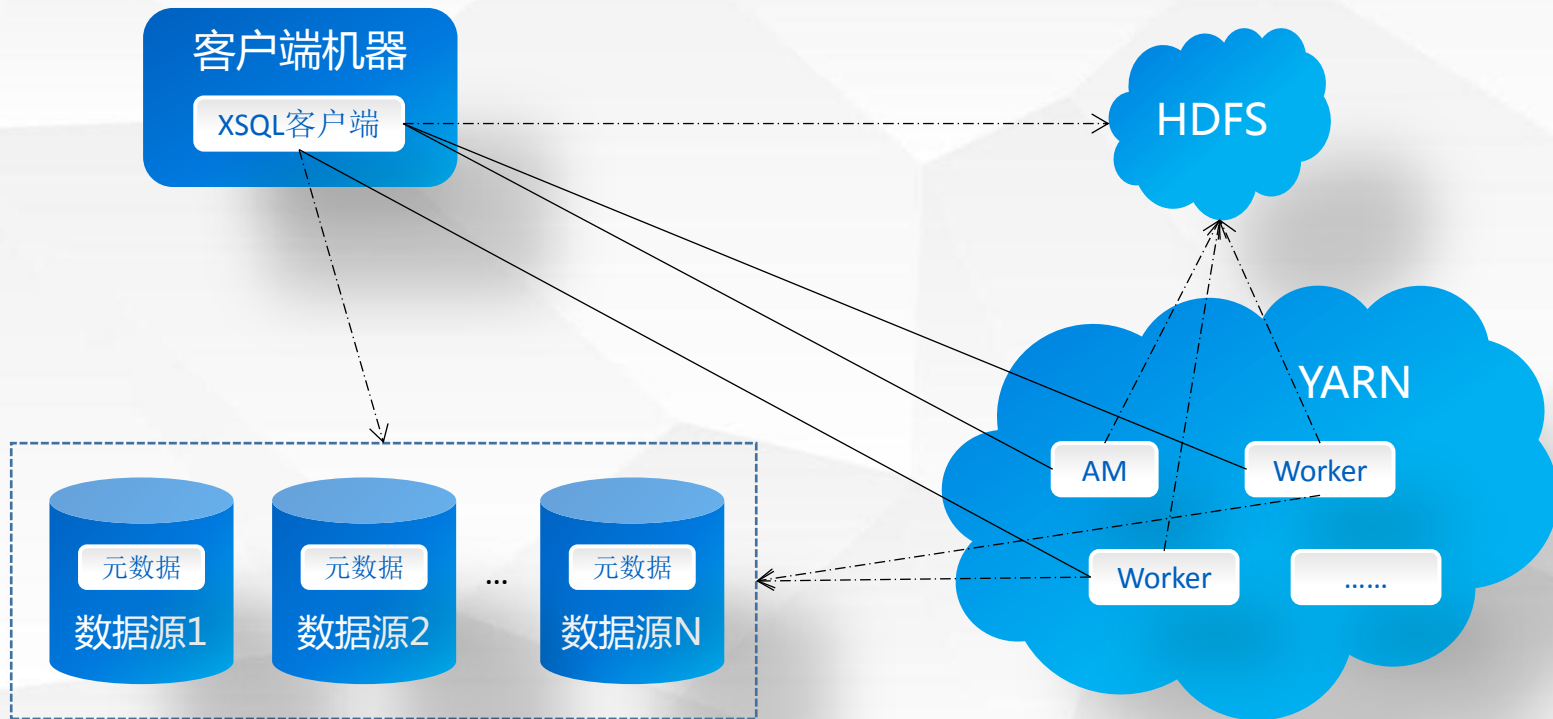
XSQL技术架构

- XSQL整体架构
- XSQL数据源管理

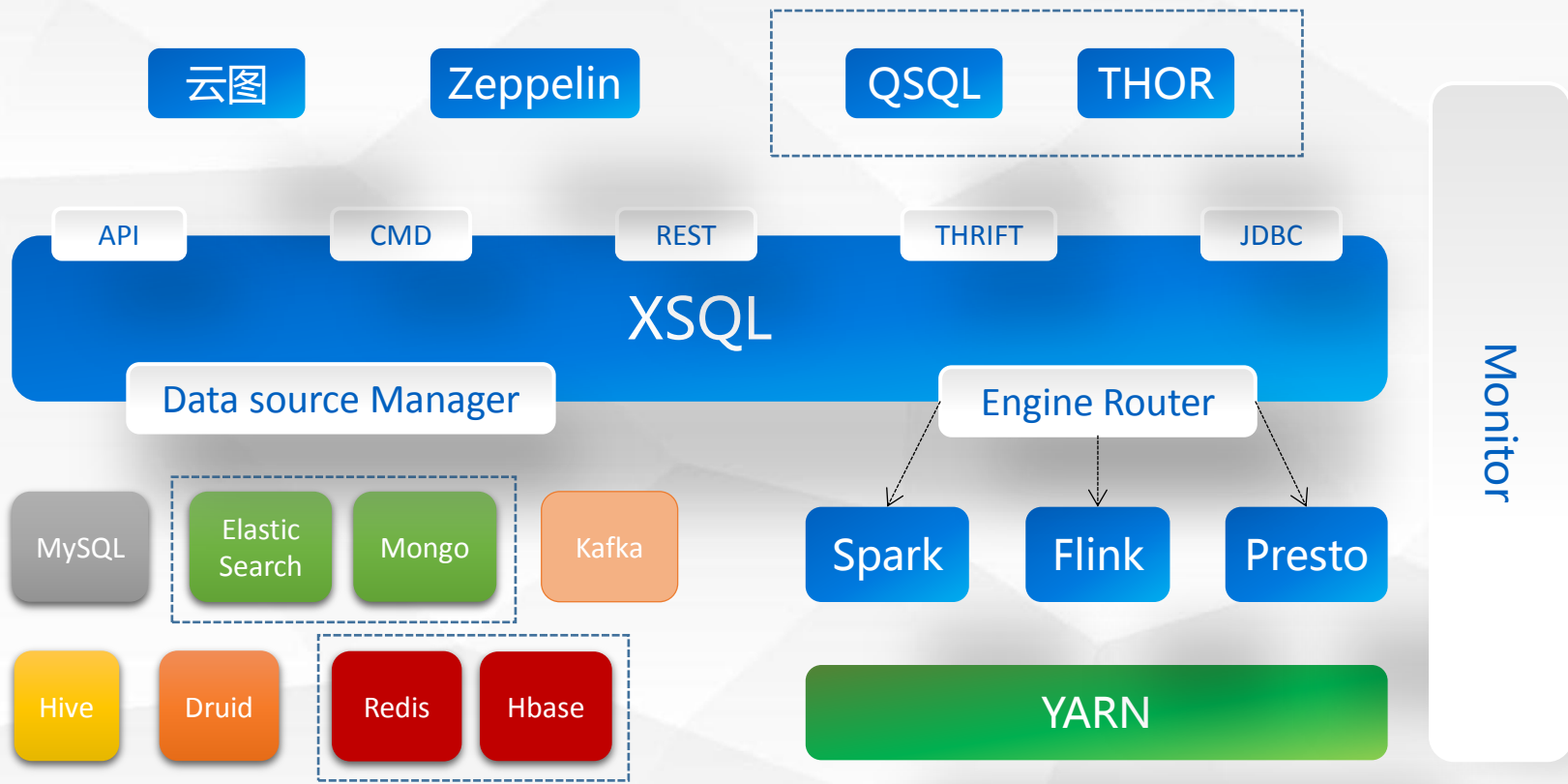
XSQL整体架构



XSQL整体架构



XSQL整体架构

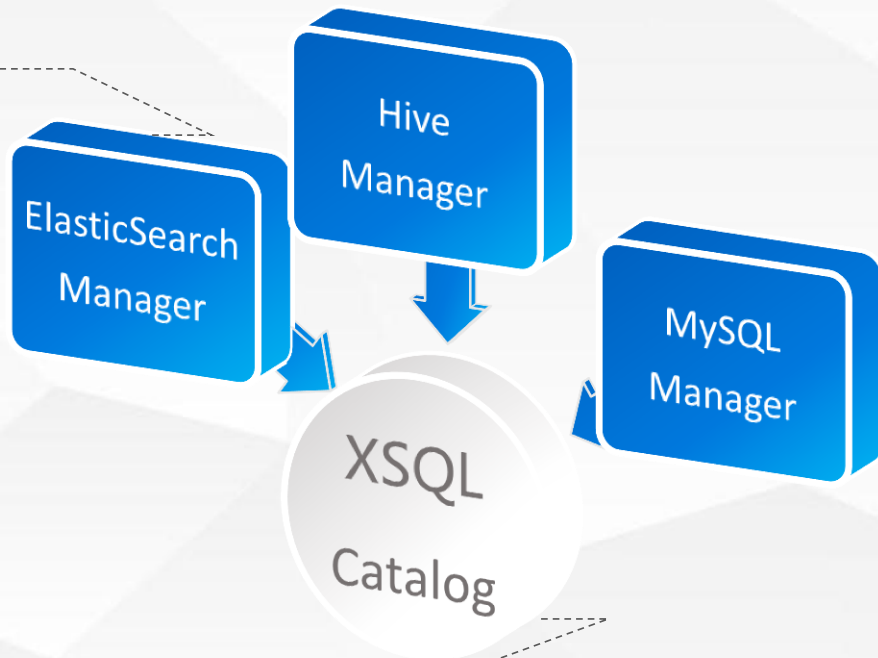


数据源管理器

通过实现DataSourceManager接口，对具体数据源的元数据进行管理。XSQL目前内置了七种不同的数据源管理器实现。数据源管理器可以以插件的方式整合到XSQL中。

XSQLCatalog通过调用数据源管理器提供的接口，对各种数据源进行统一管理。

XSQL数据字典





05 XSQL使用介绍

- 多数据源配置
- XSQL帮助文档
- XSQL的上帝
- 应用举例
- XSQL里程碑

多数据源配置

数据源列表

`spark.xsql.datasources=default,customer,order`

Hive数据源示例

`spark.xsql.datasource.default.type=hive`

`spark.xsql.datasource.default.metastore.url=thrift://127.0.0.1:9083`

`spark.xsql.datasource.default.user=test`

`spark.xsql.datasource.default.password=test`

`spark.xsql.datasource.default.version=1.2.1`

MySQL数据源示例

`spark.xsql.datasource.customer.type=mysql`

`spark.xsql.datasource.customer.uri=jdbc://127.0.0.1:3306`

`spark.xsql.datasource.customer.user=user_test`

`spark.xsql.datasource.customer.password=secret`

`spark.xsql.datasource.customer.version=5.6.19`

ES数据源示例

`spark.xsql.datasource.order.type=elasticsearch`

`spark.xsql.datasource.order.uri=http://127.0.0.1:9425`

`spark.xsql.datasource.order.user=root`

`spark.xsql.datasource.order.password=secret`

`spark.xsql.datasource.order.version=5.2`

1 SQL-2003

2 Data Source SQL

3 Hive SQL

4 MySQL dialect

```
CREATE DATASOURCE customer (  
    type='mysql',  
    url='jdbc:mysql://127.0.0.1:2336',  
    user='test',  
    password='test',  
    version='5.6.19');  
SHOW SCHEMAS FROM customer;  
SHOW TABLES FROM customer.web;  
DESCRIBE customer.web.clicks;  
SHOW COLUMNS FROM customer.web.clicks;  
SELECT * FROM customer.web.clicks;  
ALTER TABLE customer.web SET TBLPROPERTIES ('note' = '备注');  
INSERT OVERWRITE TABLE customer.web.clicks select * from other;  
INSERT OVERWRITE DIRECTORY '/home/customer/web/clicks' SELECT *  
from customer.web.clicks;  
INSERT INTO TABLE clicks values('a','A','ao'),('b','B','bo');  
ALTER TABLE clicks RENAME TO clicks_new;  
RENAME TABLE clicks TO clicks_new;
```

Getting Started

XSQL帮助文档提供了快速入门的内容。从应用场景、基本概念到安装、配置、运行。



A

Data Source Configuration

XSQL帮助文档提供了详细的配置介绍。除了通用的配置外，还对各个数据源的特有配置进行介绍。



B

<http://xsql.qihoo.net>

Command & API

XSQL帮助文档展示了一些基于命令行的使用示例。此外还有各类API使用方面的详尽内容。



C

Performance & Utils

XSQL帮助文档展示了一批SQL查询的性能测试报告。此外还提供了一些常用的工具。



D

XSQL里程碑

2018.4
XSQL立项

2018.7
XSQL内核
Hive支持

2018.10.19
HBASE支持

2018.11.13
MongoDB支持

2019.1.18
第一个用户
手机助手

2019.1.31
0.1.0发布
Hive 90%覆盖率

2019.4.4
0.3.0发布
插件化等特性

2018.5
XSQL架构设计

2018.8
ES支持

2018.10.29
DRUID支持

2018.12.7
REST支持
THRIFT支持
JDBC支持

2019.1.24
第一个ES用户
云OS

2019.3.6
0.2.0发布
缓存级别

2019.4.30
金融业务线
全面迁移到
XSQL



360 手机助手

01. 360手机助手

手机助手是XSQL的第一个使用用户。目前手机助手使用XSQL部分替代Spark任务，并且利用XSQL数据导入导出功能为Web界面提供实时查询。



360 手机OS

02. 360手机OS

手机OS是XSQL的第一个使用Elasticsearch的用户。目前手机OS使用XSQL来对Elasticsearch数据进行日常的验证工作。



360 金融

03. 360金融

360金融是XSQL的特殊用户。XSQL为金融特别增加了数据权限相关的控制，以提高其数据安全。XSQL通过对数据倾斜的优化，也让金融的数据产出更加稳定。



360 搜索

04. 360搜索

360搜索的智能分析系统THOR目前已经接入XSQL。360搜索正在全面迁移到XSQL中。

其他

05. 其他

在公司的其他部门，我们将继续推广、应用。

X S Q L

演示完毕，谢谢观看

360系统部——统一计算组