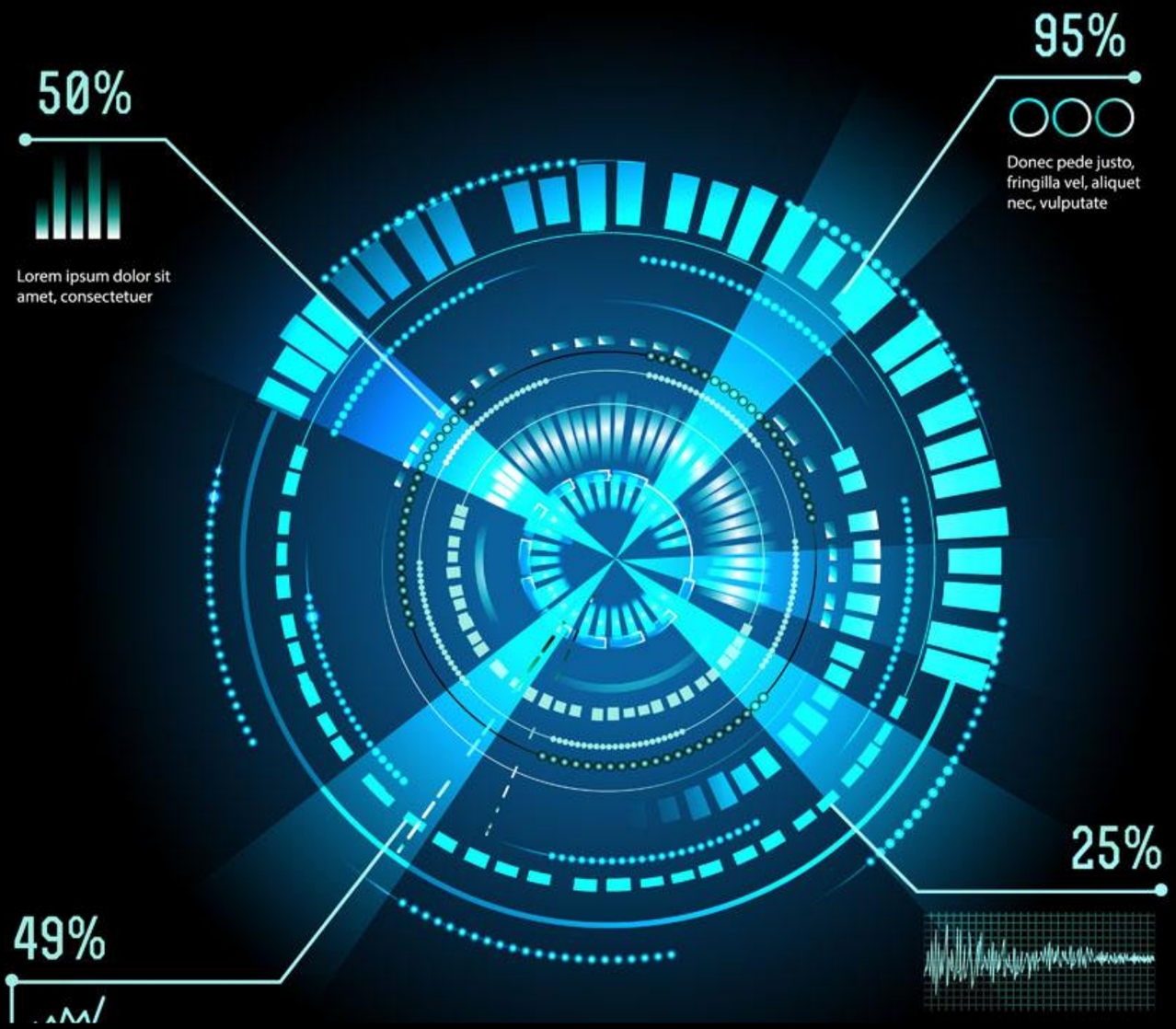


XSQL查询引擎简介

——系统部 统一计算组

What is XSQL

- 基于Spark SQL进行二次开发
- SQL Everything，支持Hive、HDFS、ElasticSearch、Mysql、Druid、MongoDB、HBase、Redis
- 支持查询下推，兼具毫秒级查询与批处理查询的能力
- 动态调整分区，自动解决数据倾斜
- 打通存储边界，跨数据源查询不再困难
- 众多连接方式，支持Cli、Restful、Thrift、JDBC、Spark API

Why use XSQL

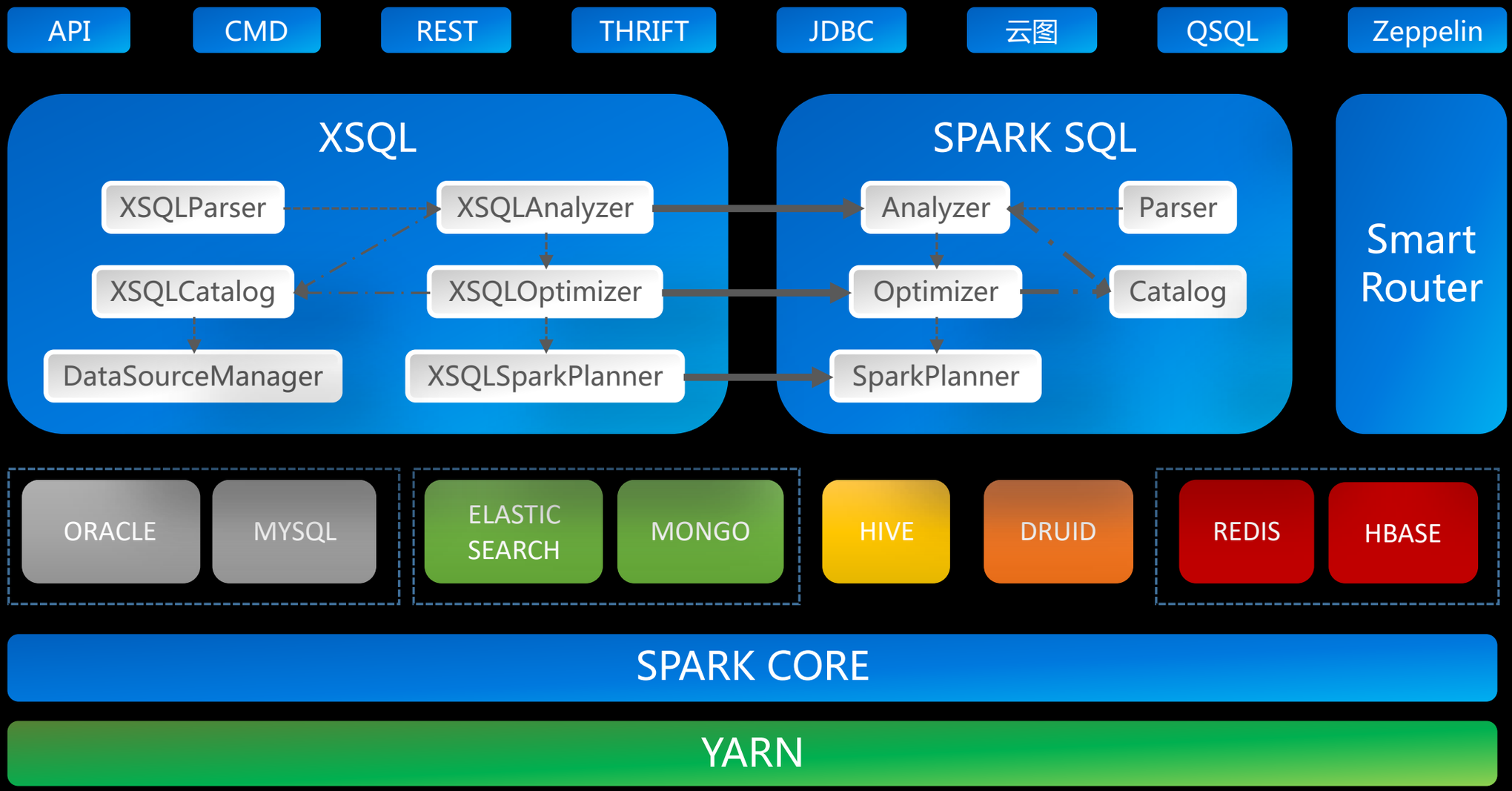
- 降低学习成本，~~pyspark、scala、bash、数据源API~~
- 简化交接流程，降低代码复杂度、提高代码可读性、配置简单
- 零升级成本，随XSQL升级自动获得持续的性能提升
- 避免无意义的数据转储，轻松实现数据联邦

PS：建议优先将新业务切换到XSQL，早使用早受益

Who use XSQL

- 360手机助手
- 360金融
- 360OS
- 360搜索
- 大数据中心QSQL
- More Coming…

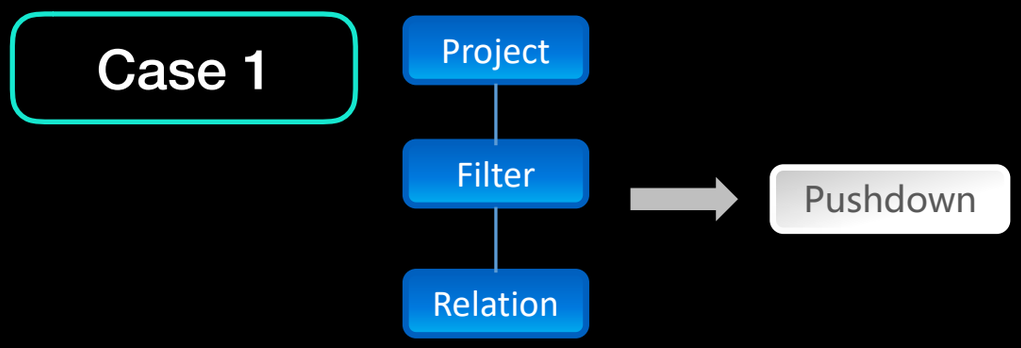
Architecture of XSQL

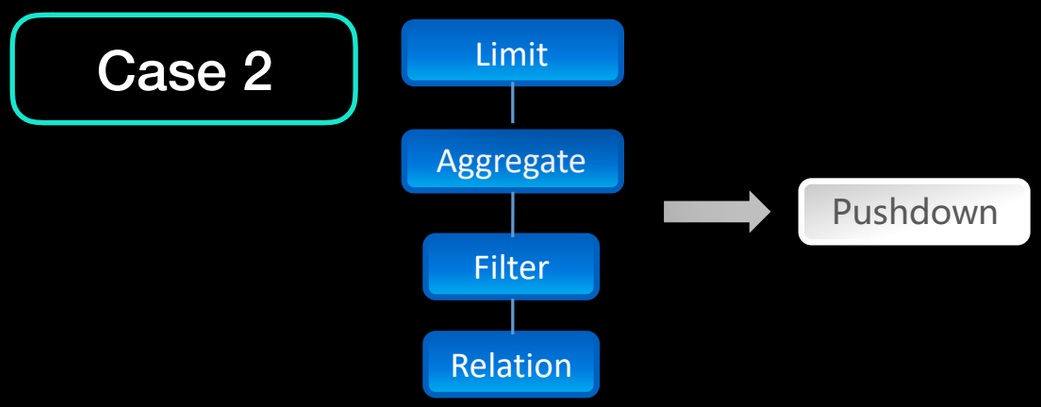


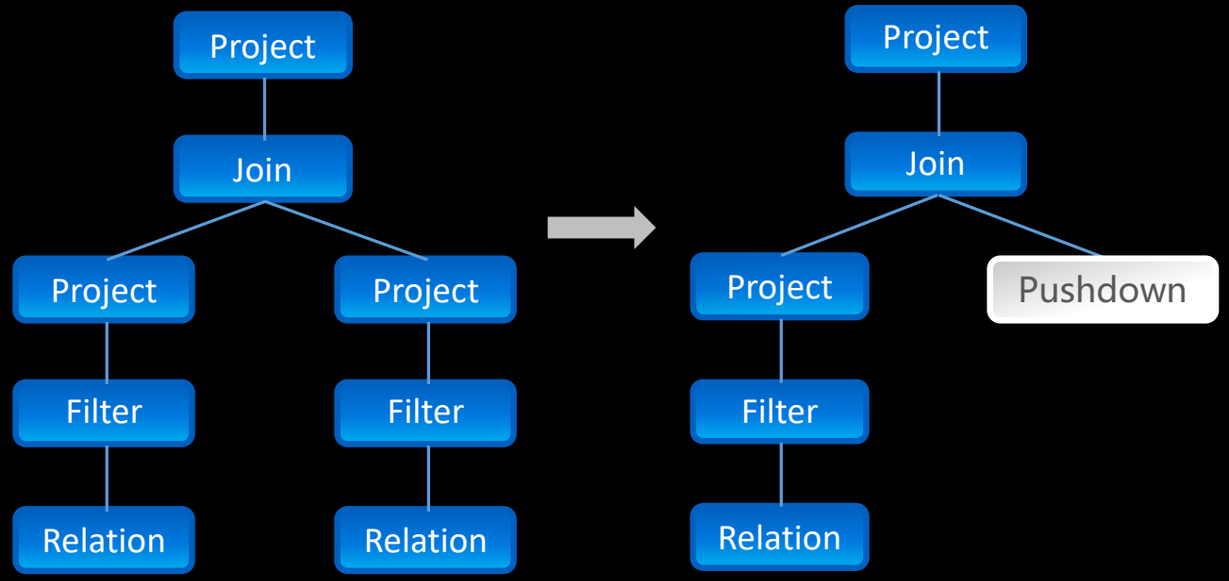
Architecture of XSQL

1. 支持查询下推

下推：将原始查询或原始查询的一部分交给数据源执行，从而最大程度地减少网络传输，同时充分利用不同数据源的性能优势，具体分为：列裁剪、谓词下推、聚合函数下推（Druid、ElasticSearch）、Limit下推。



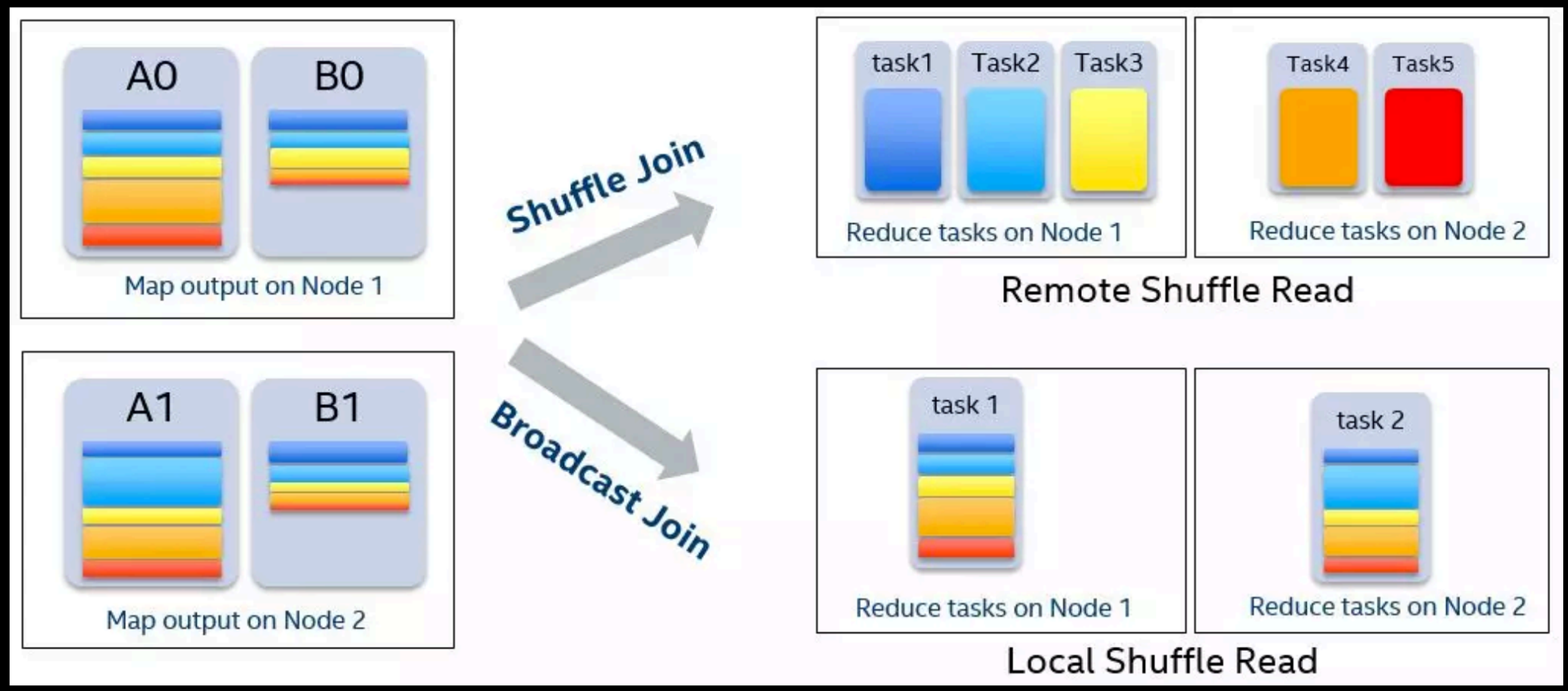




Architecture of XSQL

2. Spark SQL自适应执行

- 执行阶段根据中间结果大小将Shuffle Join转化为Broadcast Join

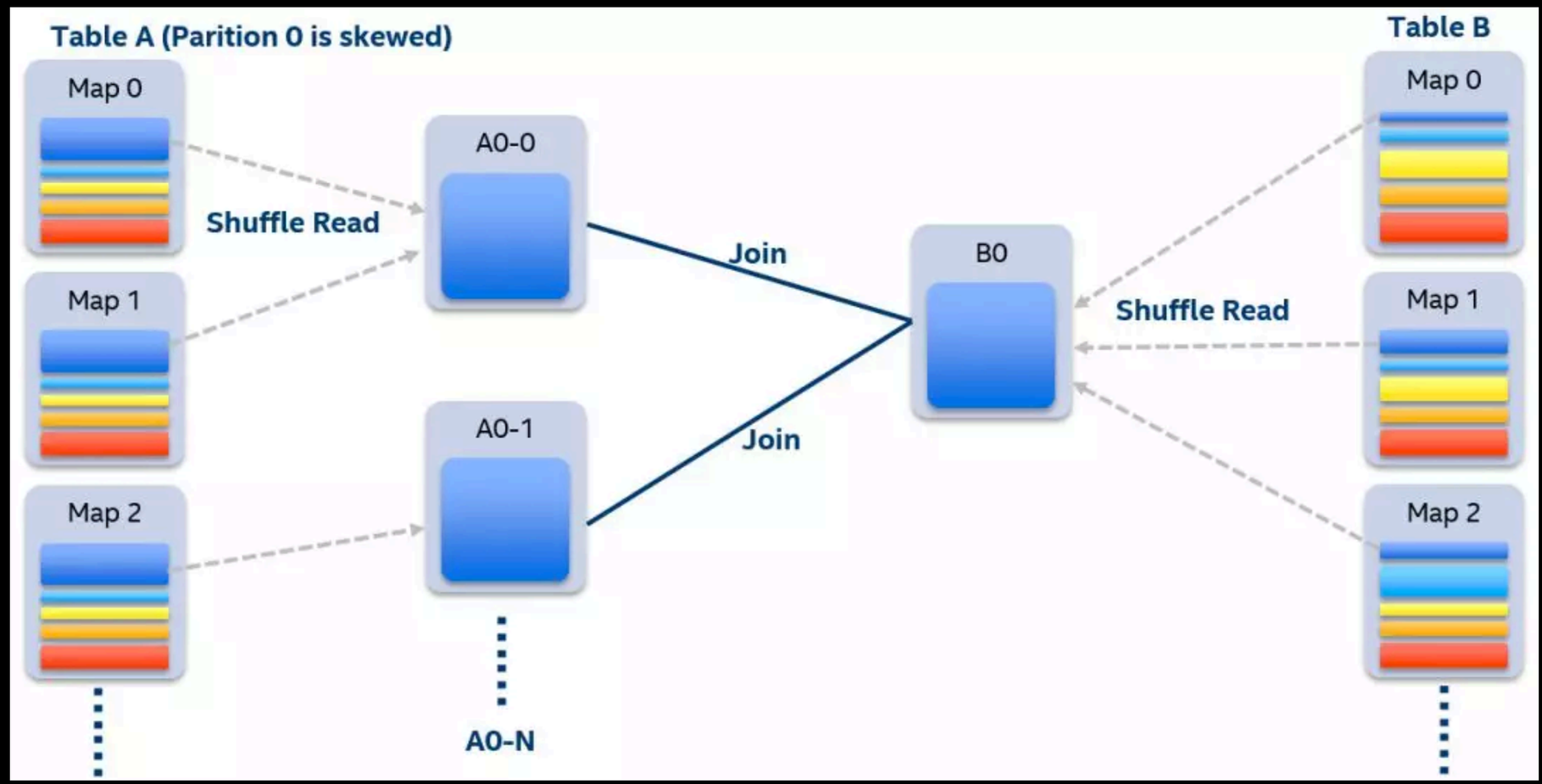


- 动态处理数据倾斜，自动拆分数据倾斜的分区

Architecture of XSQL

2. Spark SQL自适应执行

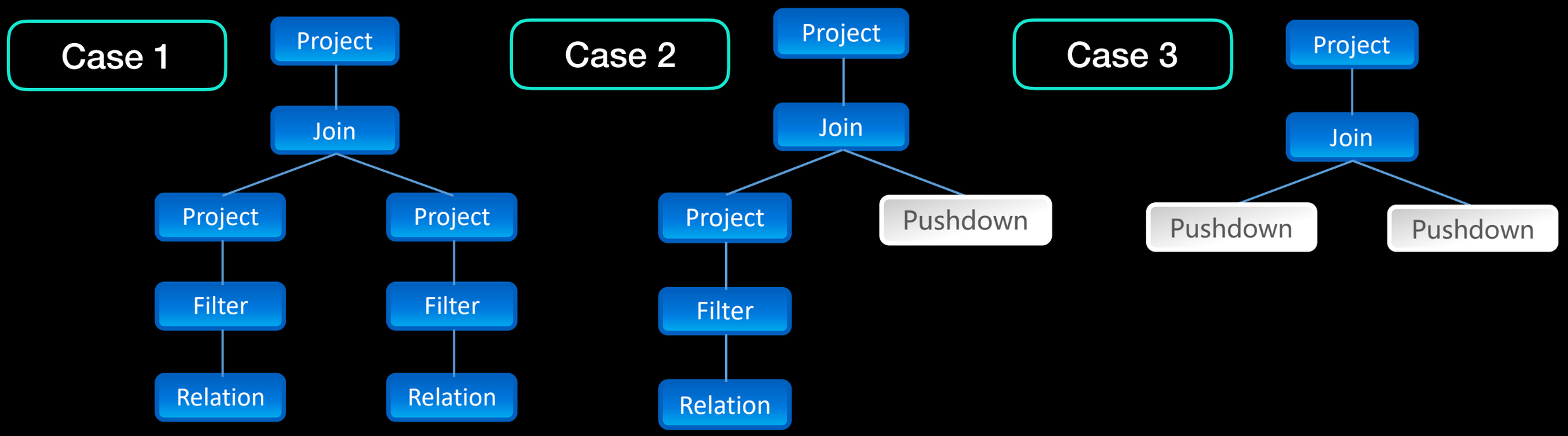
- 执行阶段根据中间结果大小将Shuffle Join转化为Broadcast Join
- 动态处理数据倾斜，自动拆分数据倾斜的分区



Architecture of XSQL

3. 跨数据源连接

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



Usage of XSQL

XSQL.conf 配置示例：

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

spark.xsql.datasource.default.type = hive
spark.xsql.datasource.default.metastore.uri = thrift://42.236.102.104:9083
spark.xsql.datasource.default.user = test
spark.xsql.datasource.default.password = test
spark.xsql.datasource.default.version = 1.2.1

spark.xsql.datasource.customer.type = mysql
spark.xsql.datasource.customer.url = jdbc://127.0.0.1:3306
spark.xsql.datasource.customer.user = root
spark.xsql.datasource.customer.password = secret
spark.xsql.datasource.customer.version = 5.6.19

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

Usage of XSQL

DDL 示例：

- 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;
- alter table customer.web.clicks set tblproperties (‘note' = ‘备注');
- use customer.web

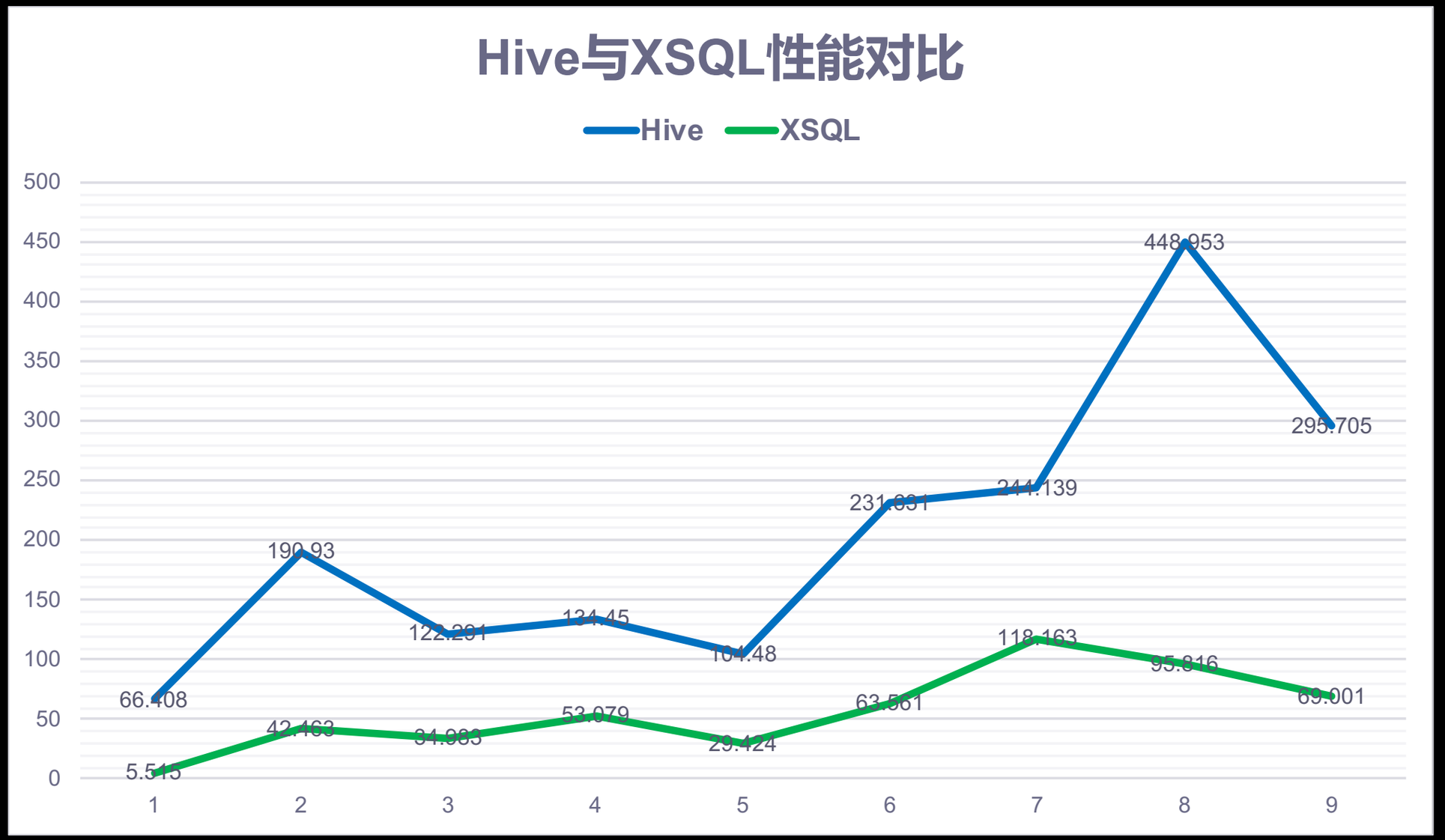
DML 示例：

- select s.*,c.* from default.xitong.student s, myes.xsql_test.class_room c where s.class_id = c.id
- 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’);

更多细节请参看：http://xsql.qihoo.net/

Performance of XSQL

XSQL vs Hive



Performance of XSQL

XSQL vs Nature Datasource API