



十年架构 成长之路

SACC 第十届中国系统架构师大会

SYSTEM ARCHITECT CONFERENCE CHINA 2018

2018年10月17-10月21日 北京海淀永泰福朋喜来登酒店



百度云Spark优化及大规模应用

张建伟

大数据平台架构



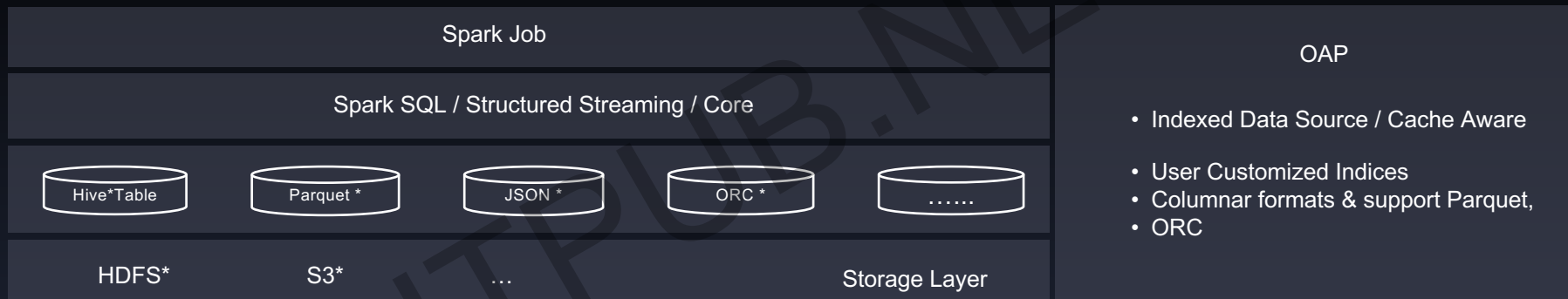
发展历程



大数据计算核心引擎 Spark



Spark 性能优化 – OAP (Optimized Analytics Package)

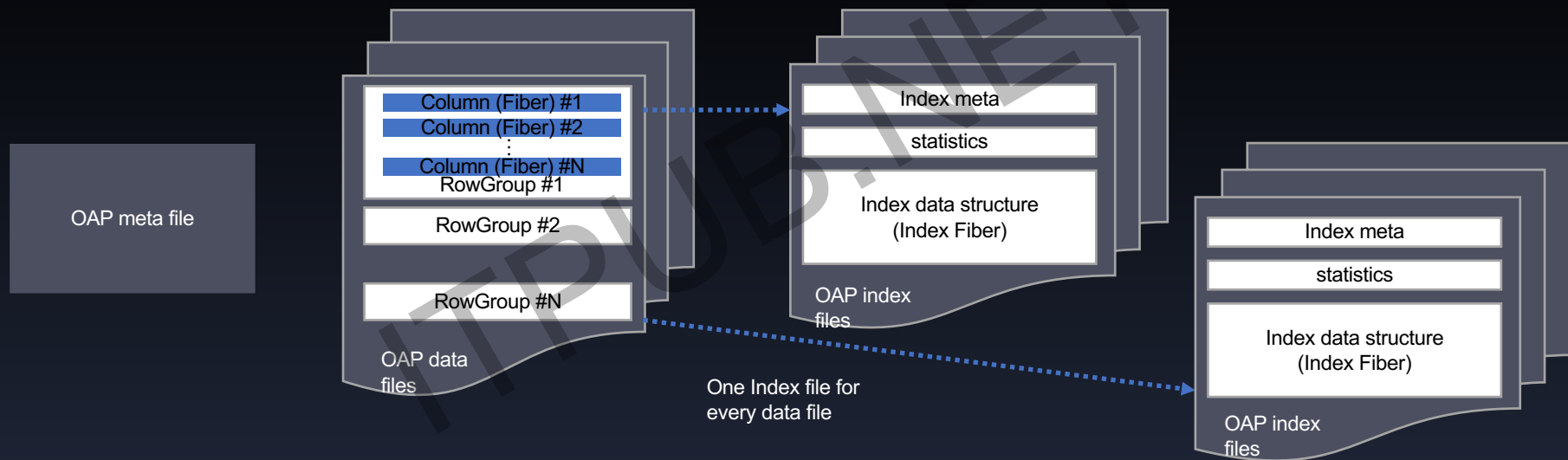


• 开源：与 Intel 联合开源

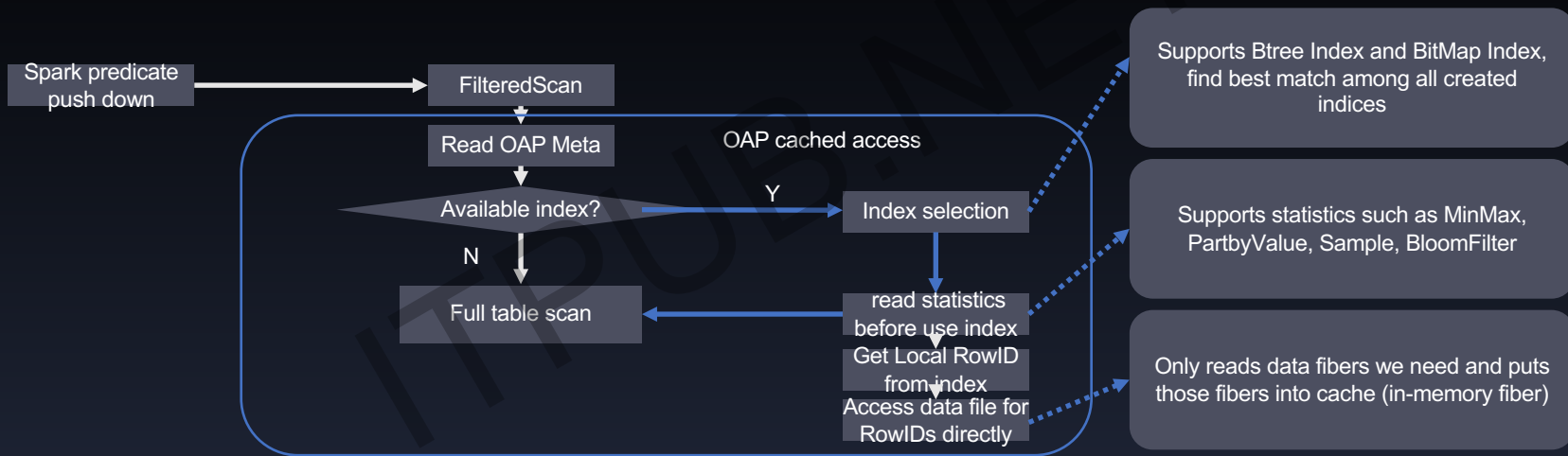
• 高性能：5 倍+性能提速

• 易用：插件式易部署；类似传统数据库的索引和 cache，易于使用

OAP Files and Fibers



OAP Internals - index



Select xxx from xxx where age > 29 and department in (Baidu, AI-Lab)

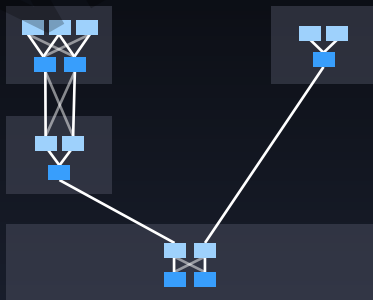
Spark 性能优化 – AE (Adaptive Execution)

Reduce 数量该如何设置？

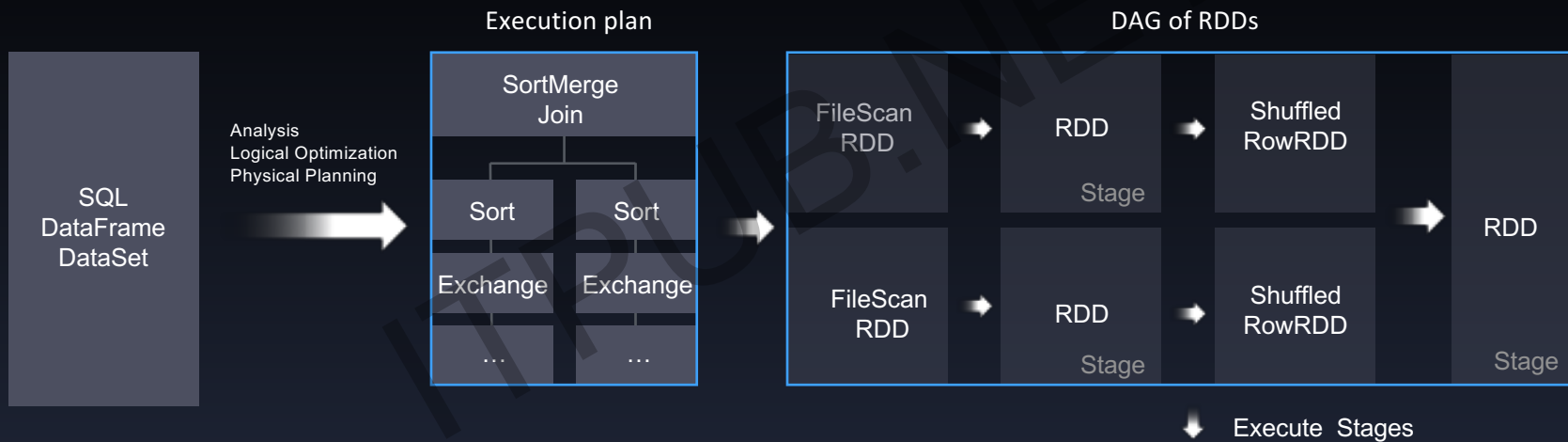
太小：Spill, OOM

太大：调度开销，更多 IO，输出更多小文件

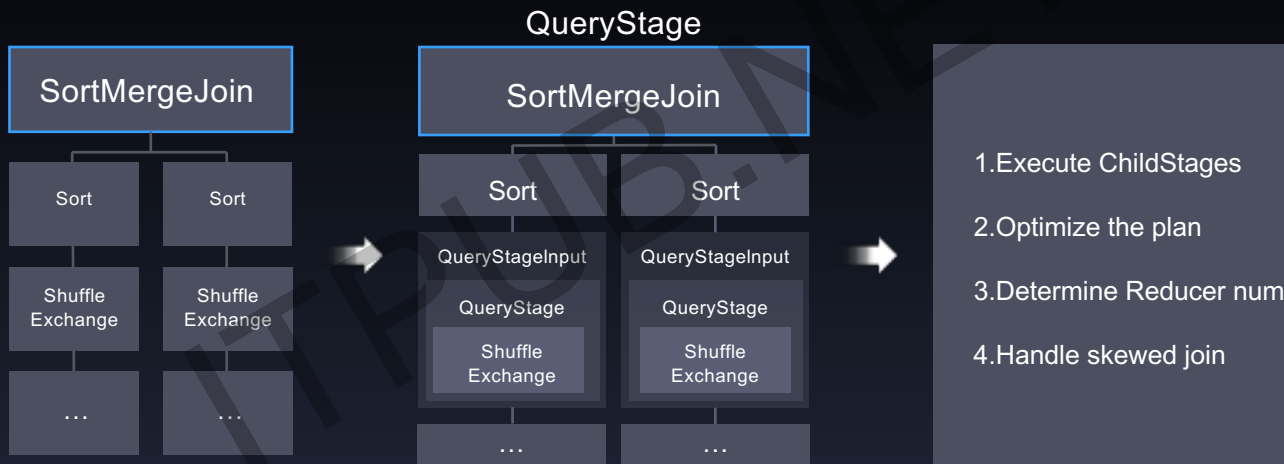
同样的 Reduce 个数不适合每个 Stage



Spark 性能优化 – AE – Normal Execution Mode

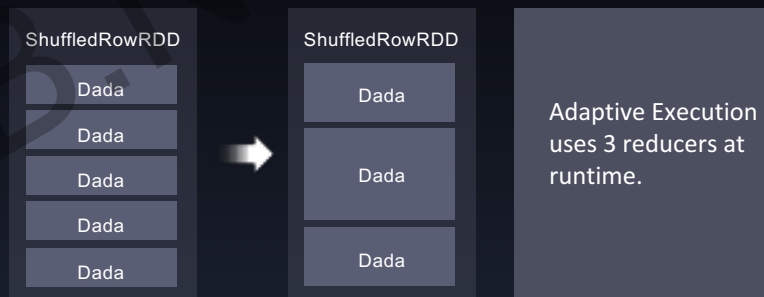


Spark 性能优化 – AE - Adaptive Execution Mode

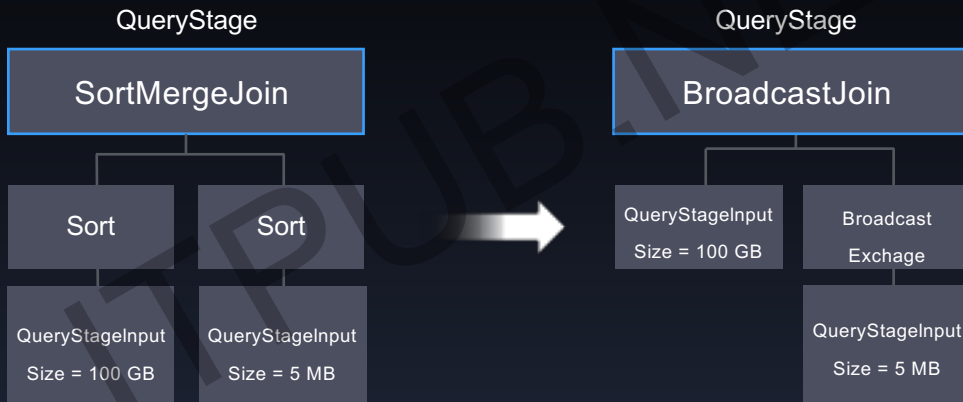


Spark 性能优化 – AE – auto setting the reduce number

- Target size per reducer = 64 MB.
- Min-Max shuffle partition number = 1 to 5

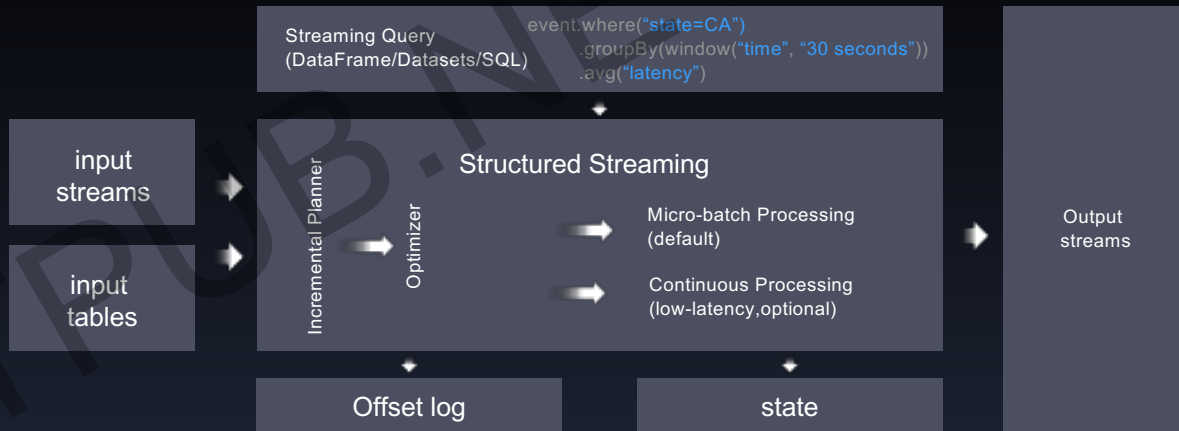


Spark 性能优化 – AE – Optimize Join Strategy at Runtime

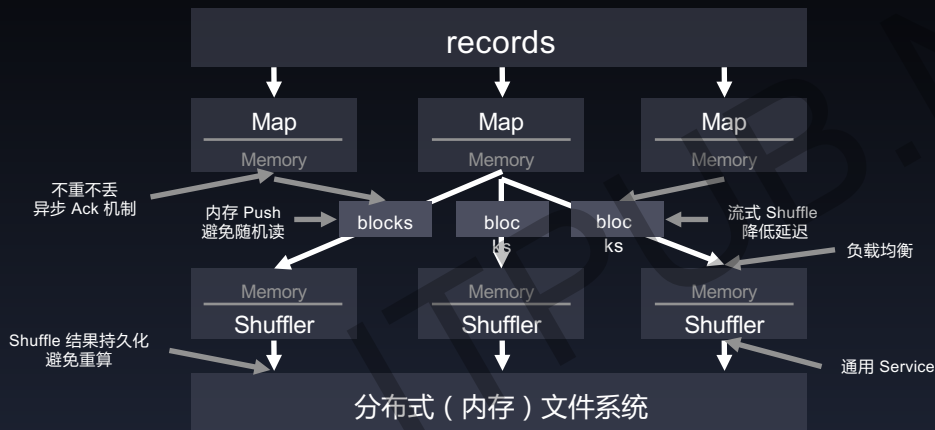


Spark功能优化 – CP (continuous processing) – 与社区联合开发

Structured Streaming processing modes:
Micro-batch and Continuous Processing



Spark 稳定性优化 – 独立 Shuffle 服务



收益

1. 流式shuffle：减少旧 shuffle map merge、reduce pull 时间消耗
2. Shuffler 内存聚合：聚合度高，减少 IO
3. 中间数据持久化：避免重算（对大 dag 作业尤为重要）

基于 Spark 的可信数据计算平台

可信数据计算平台 (SaaS 服务)

在安全可信的环境下，探索数据价值

可信计算环境
(AC)

Sgx 加固的 spark
计算框架

数仓之间的
多方计算($A \cap B$)

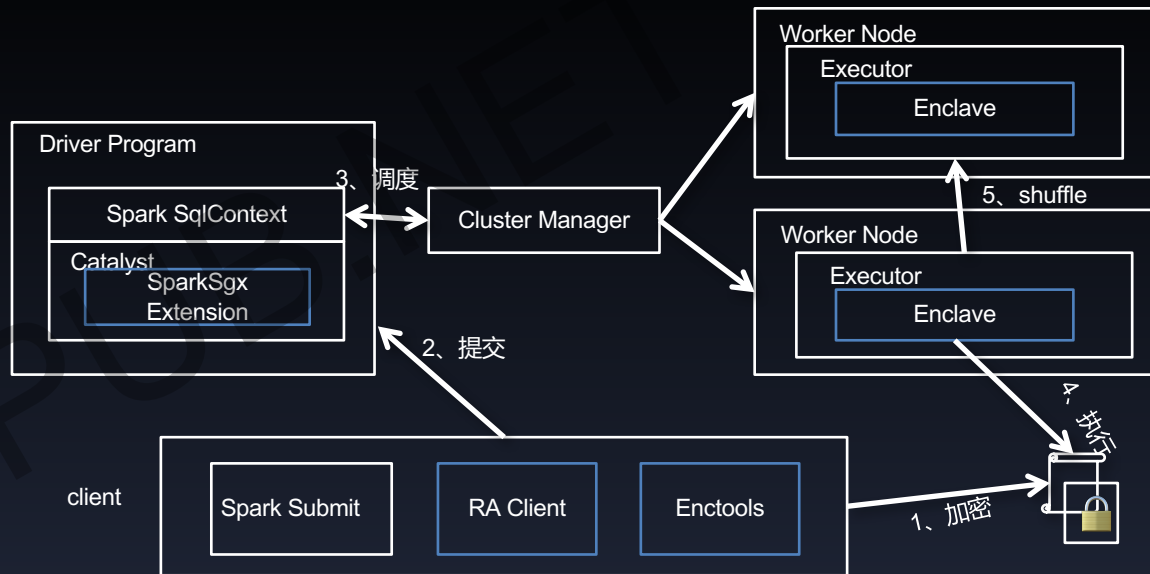
Sql 版本的联合
多方计算服务

隐私数据查询
分析($A \rightarrow B$)

差分隐私解决数据
查询隐私泄漏

SparkSgx整体架构

- 集群
 - SparkSgx Extension
Catalyst的扩展转换执行计划 cluster
 - Enclave
Spark算子实现
- 客户端
 - RA Client
远程认证客户端
 - Enctools
加密工具



THANK YOU