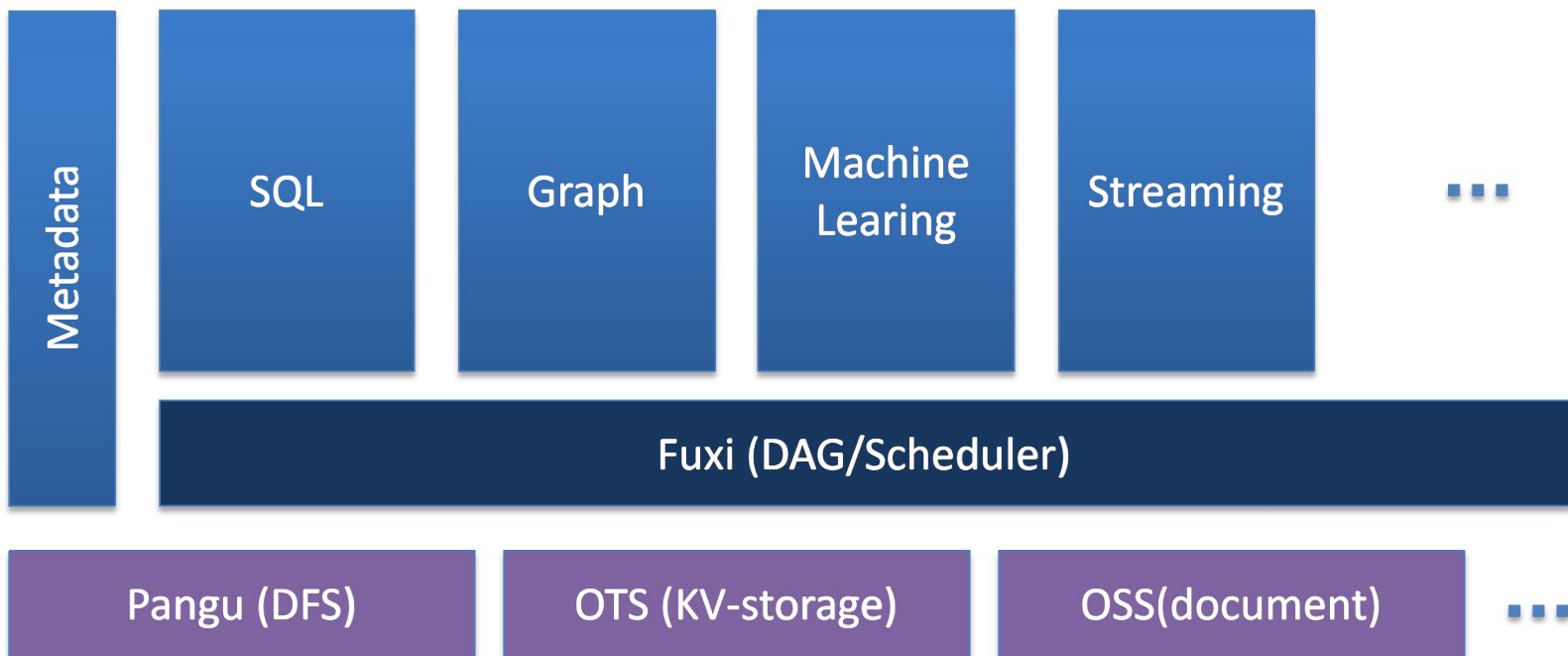


MaxCompute 复杂数据分布的查询优化实践

Agenda

- MaxCompute简介
- 数据分布问题
- MaxCompute实践

MaxCompute架构



MaxCompute SQL

Compiler

- 灵活语法兼容的解析器
- 无缝整合效率工具

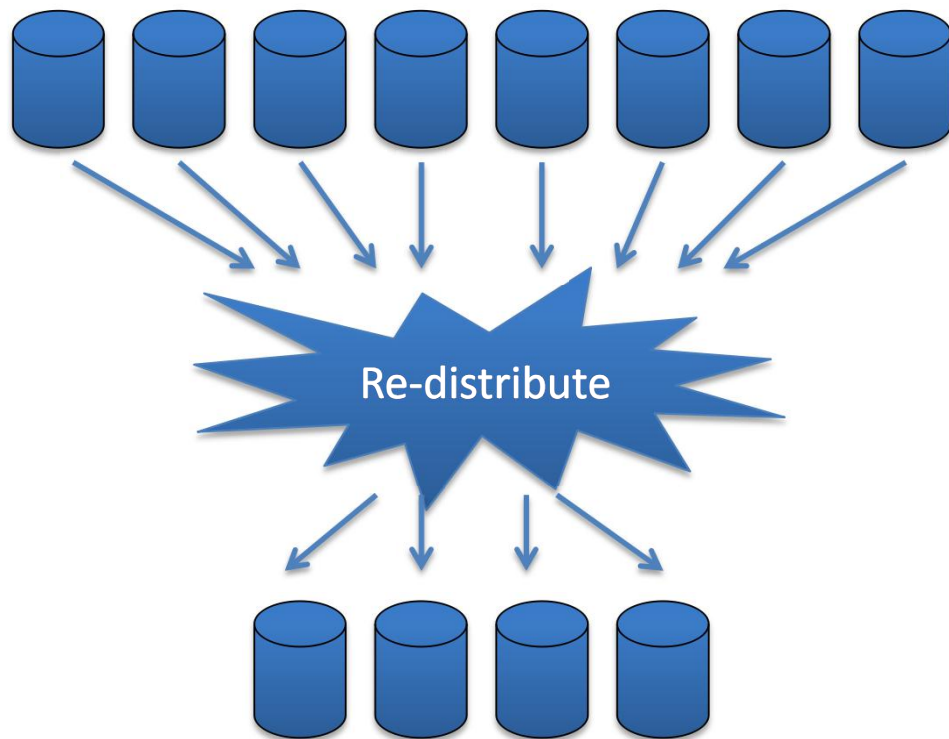
Optimizer

- 基于代价的优化器

Runtime

- 编译执行的高效运行时
- 向量执行

数据分布

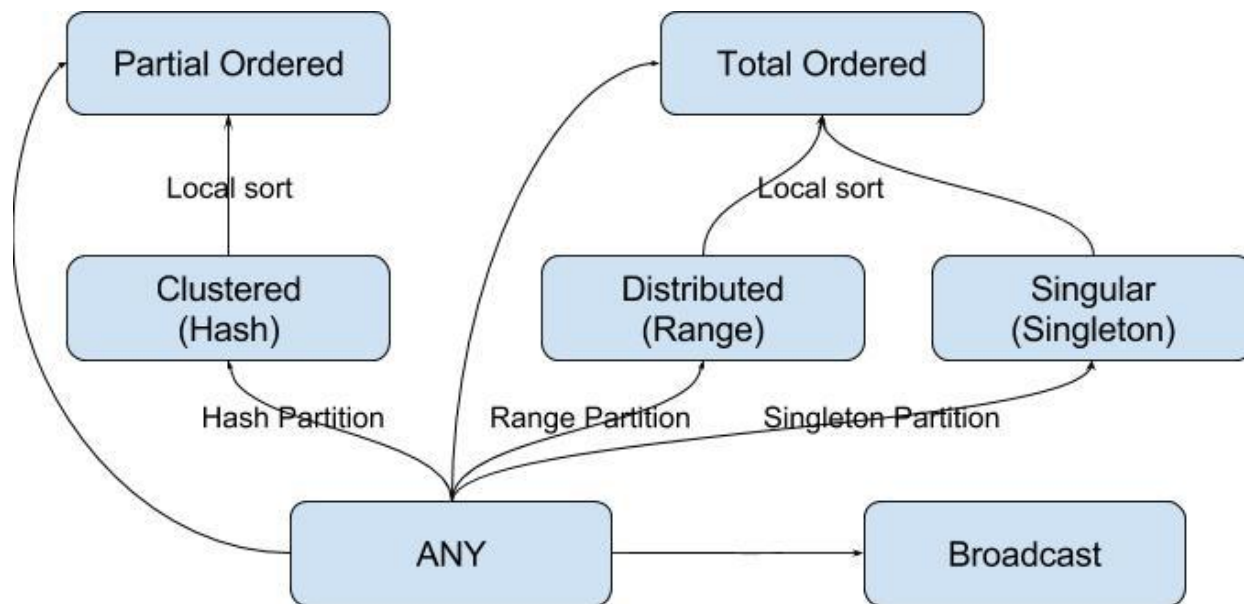


- 数据分布是大数据处理的
原生需求
 - 数据量
 - 并行化
- 重新审视“简单一致性
模型”

数据分布分类

类型	键值	参数	范例
ANY	-	-	ANY
HASH	keys	numBuckets	HASH(c1)[100]
RANGE	keys	boundaries	RNG(c1){(100, 200], (200, 300]}
BROADCAST	-	-	BROADCAST
SINGLETON	-	-	SINGLETON

数据分布转换关系

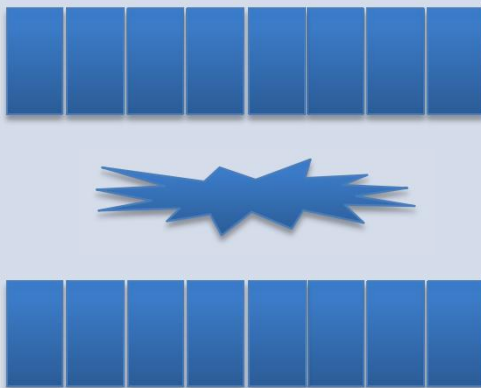


数据分布优化问题

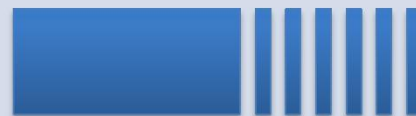
```
SELECT * FROM src  
WHERE uid = 'foobar'
```



```
SELECT src.uid, count(*)  
FROM src JOIN dim  
ON src.uid = dim.uid  
GROUP BY src.uid
```



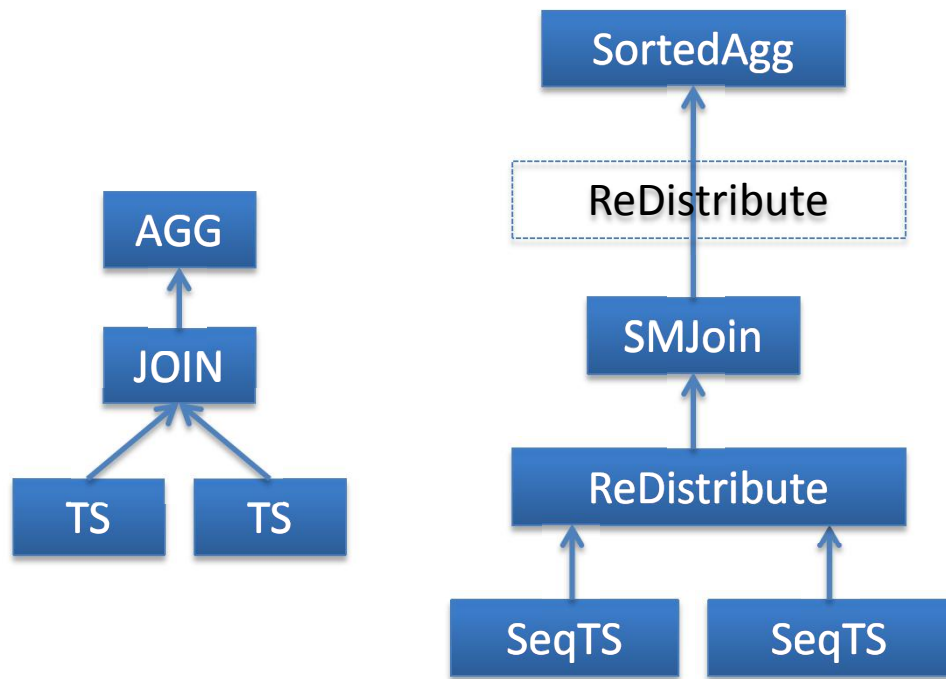
```
SELECT *  
FROM src JOIN dim  
ON src.uid = dim.uid
```



数据分布优化问题：范例

- 逻辑层面

- 物理层面

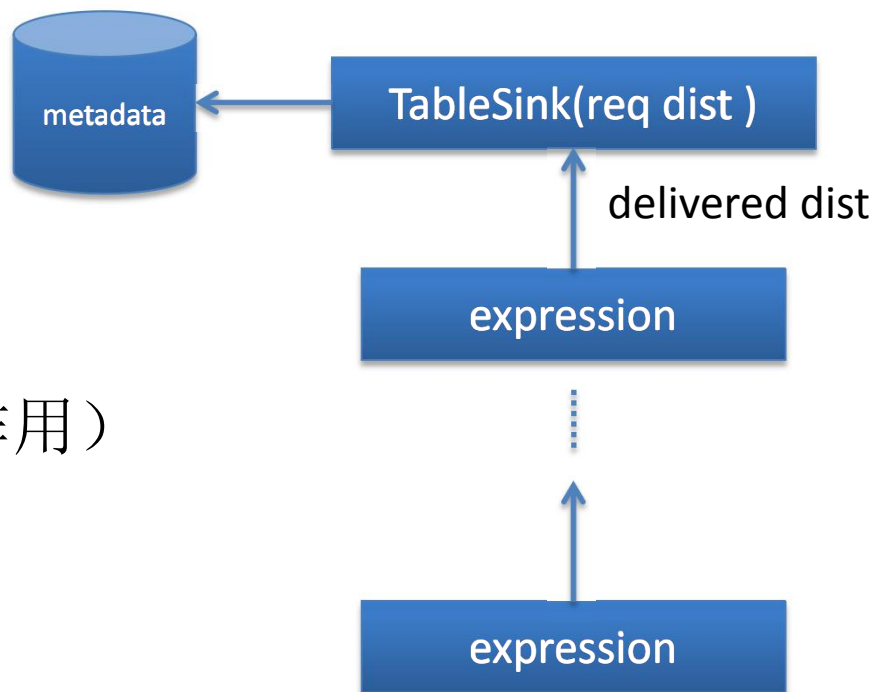


优化数据分布

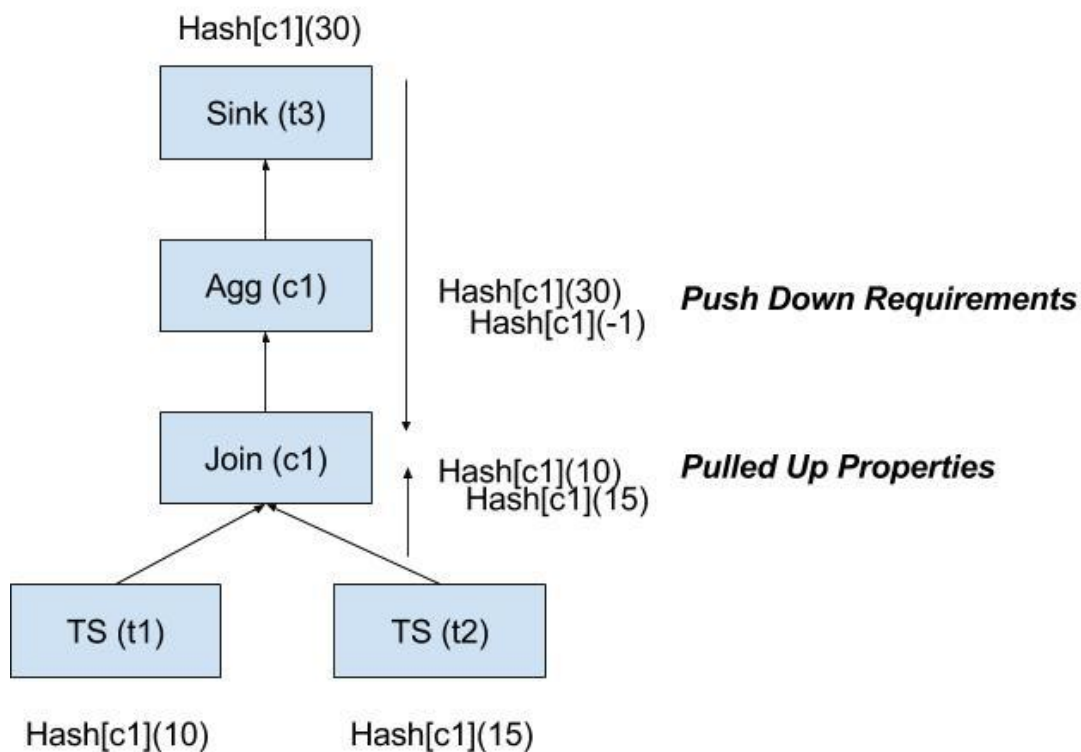
- 收集分布特性
- （局部关系代数编译）选择合适的分布特性
- （全局代价计算）规避不合适的分布特性

收集分布特性

- 产生分布
 - 用户指定
 - 逻辑计划
 - 物理计划（算法的副作用）
- 消费分布
 - Aggregate/Window
 - Join
 - Sort



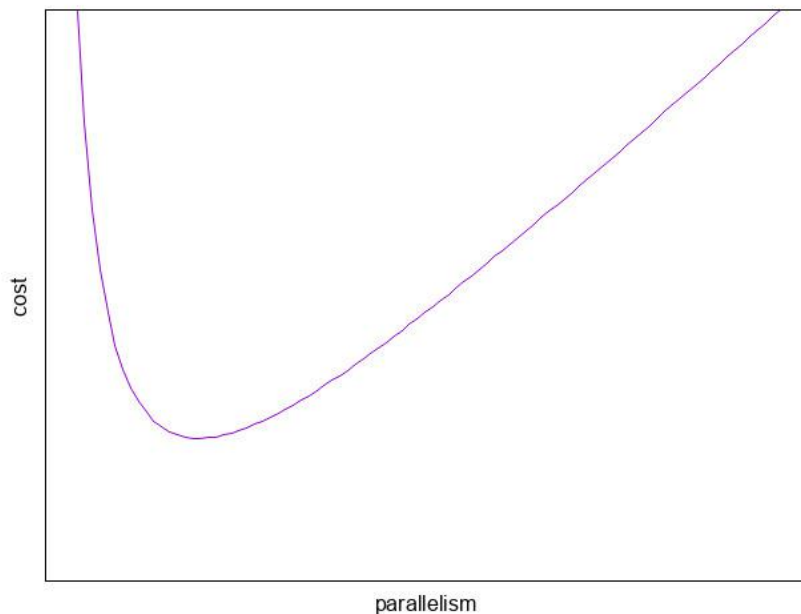
选择分布特性



- 问题的规模
 - 分布类型的选择
 - 分布参数的选择
- 分布对齐
 - Enforce
 - Align
- 启发式的选择算法

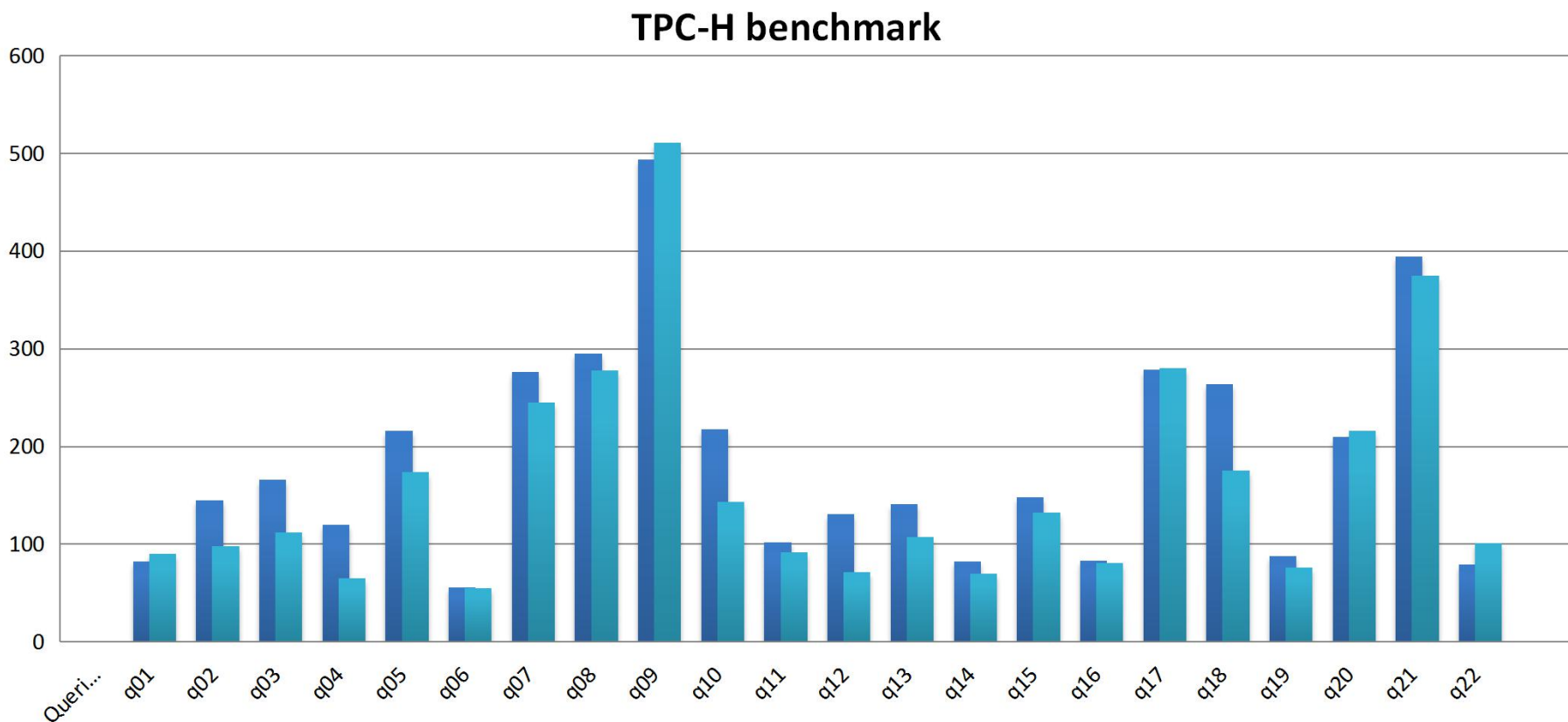
避免最差分布

- 最差分布
 - 数据倾斜 (Skew)
 - 低并发 (Under Partition)
- 代价惩罚



效果

- 预分区：12.8%性能提升，17.3%资源减少



- 利用动态分区特性：线上13.4%的查询受益

总结

- 复杂数据分布
 - 理解分布
 - 分布优化问题
- MaxCompute支持
 - 最新发布已包含Hash和部分Range相关的优化
- Roadmap
 - 利用delivered distribution
 - 复合的re-distribution
 - ...

Q & A



We're hiring 技术讨论（钉钉群）



产品官网

BDTC 2017 中国大数据技术大会
Big Data Technology Conference 2017