

苏宁 OLAP 引擎发展之路

赵云



全球技术领导力峰会

Geekbang> | TGO 鲲鹏会
极客邦科技

500+ 高端科技领导者与你一起探讨 技术、管理与商业那些事儿



🕒 2019年6月14-15日 | 📍 上海圣诺亚皇冠假日酒店



扫码了解更多信息

苏宁OLAP引擎负责人
苏宁大数据平台负责人



统一报表门户



统一指标口径



统一模型视图



统一SQL引擎

01

演进和优化

02

离线&实时任务

03

整体架构



Druid

时序汇总



Druid优化



实时peon OOM优化



动态查询优先级



大查询限制



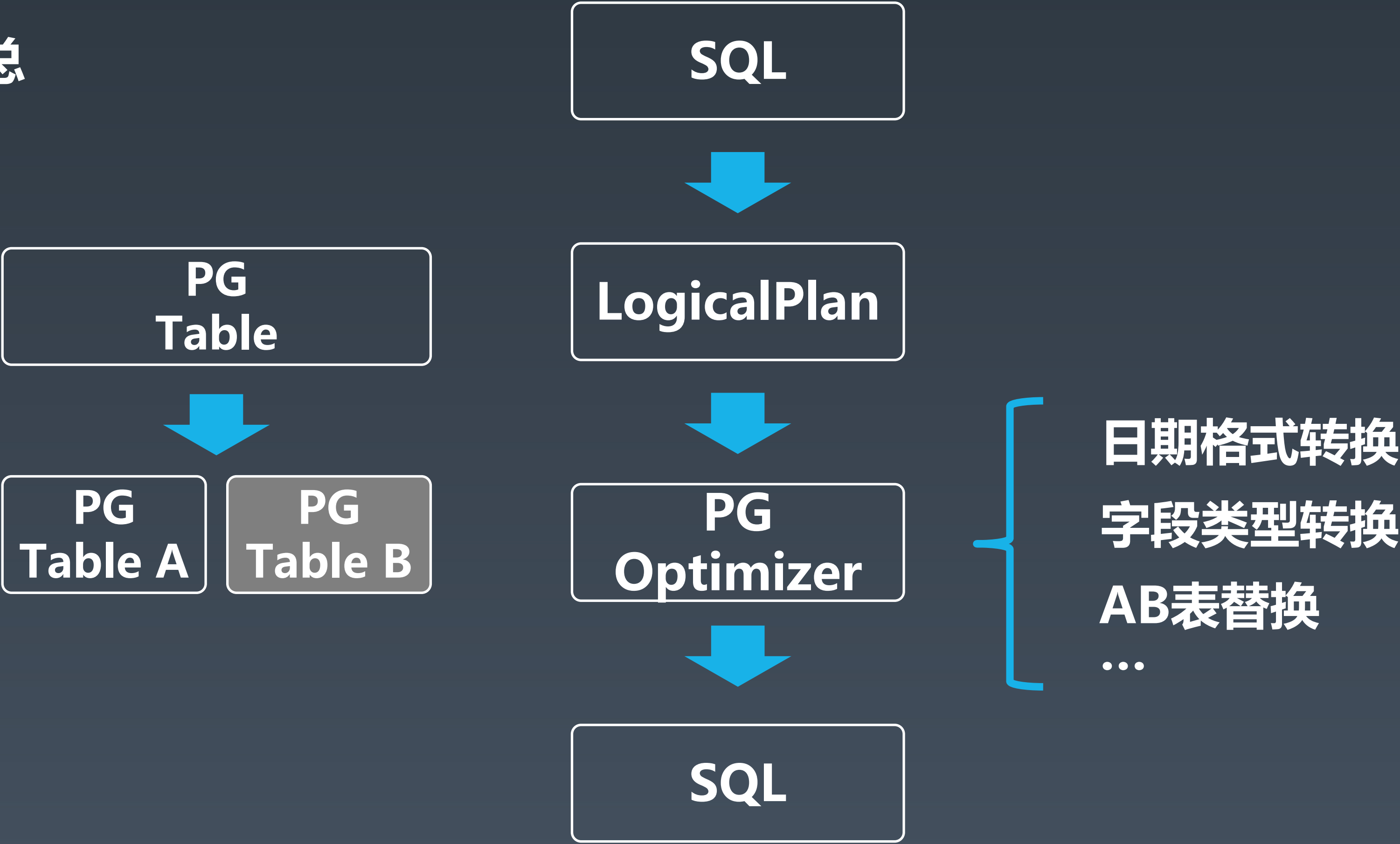
ES

明细



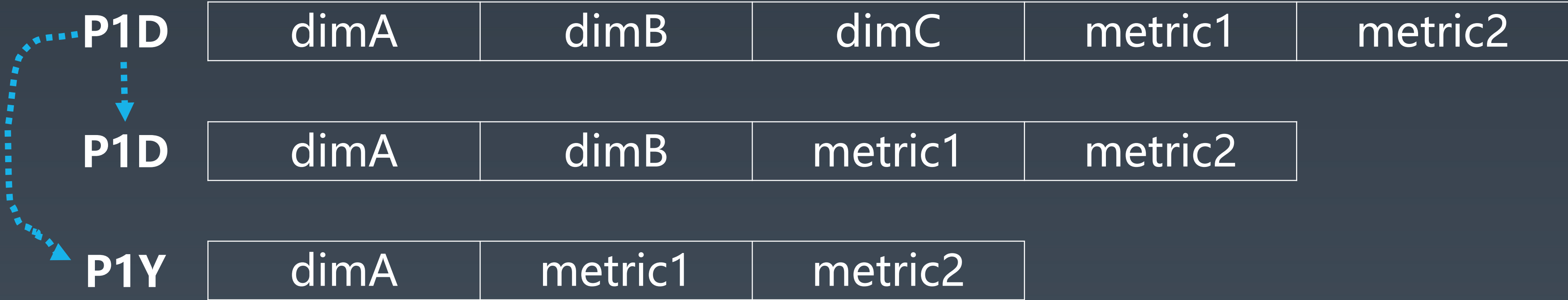
PG

非时序汇总



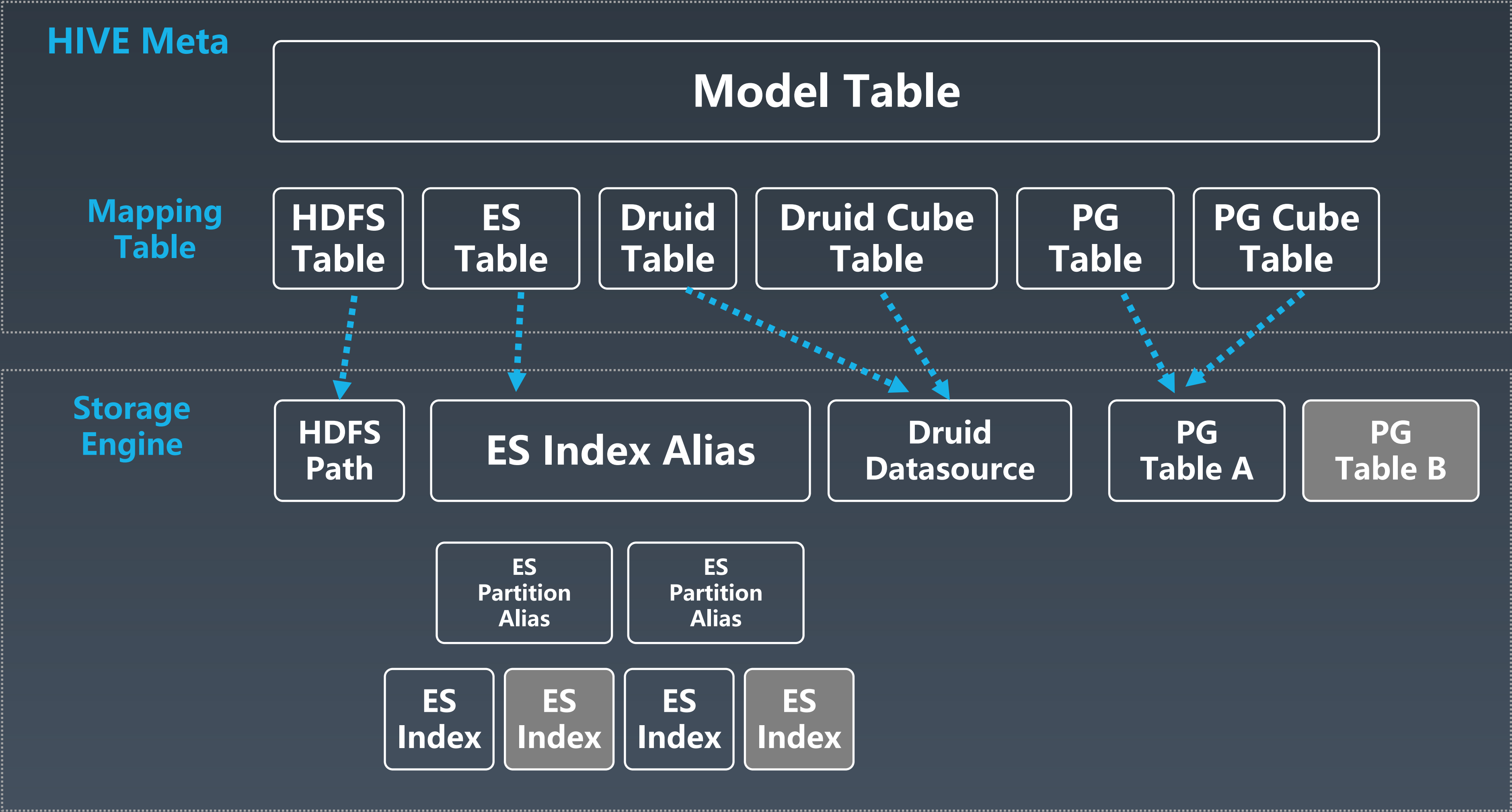
CUBE

在汇总模型上建立cube

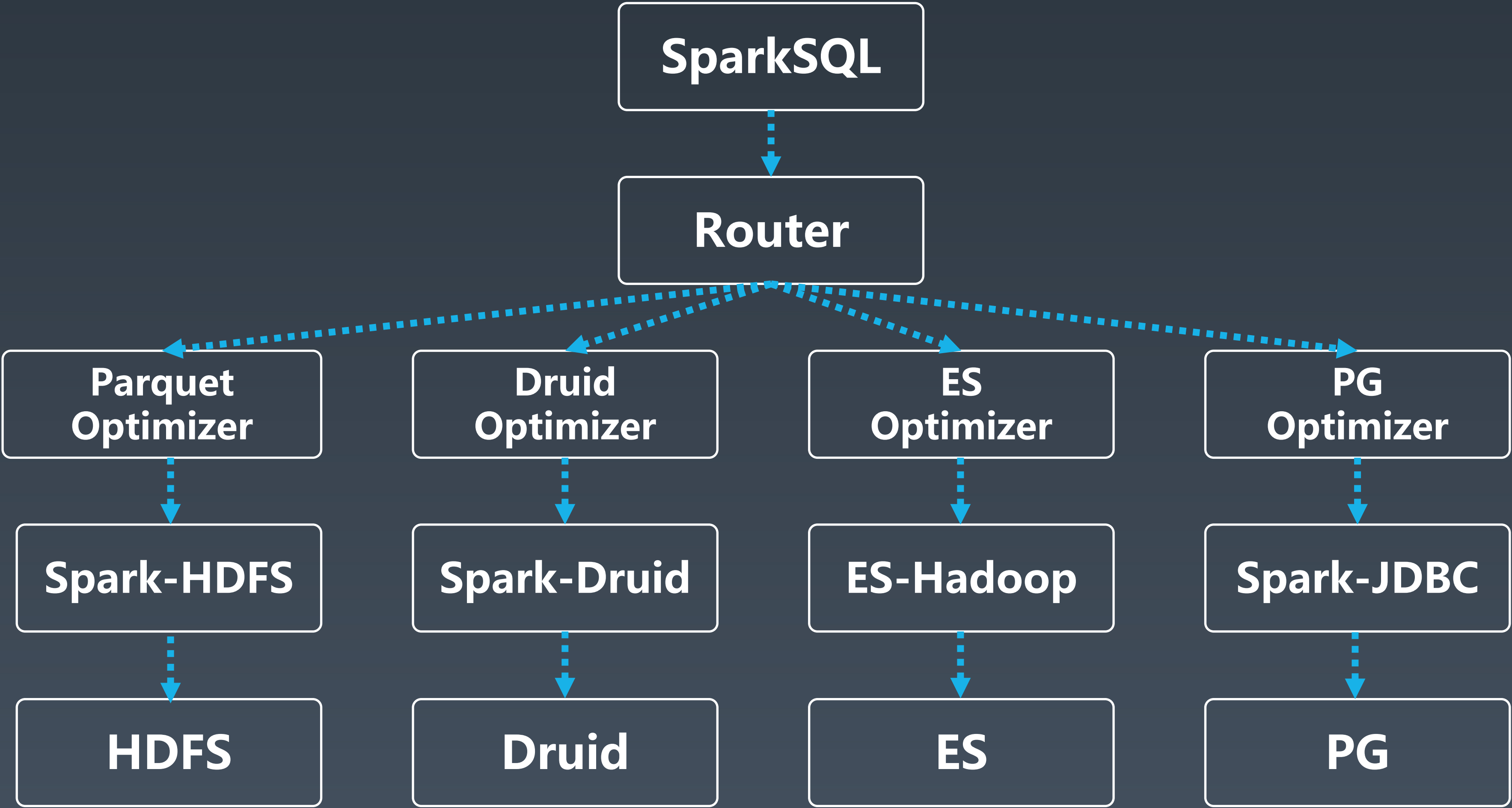


根据粒度、维度等路由

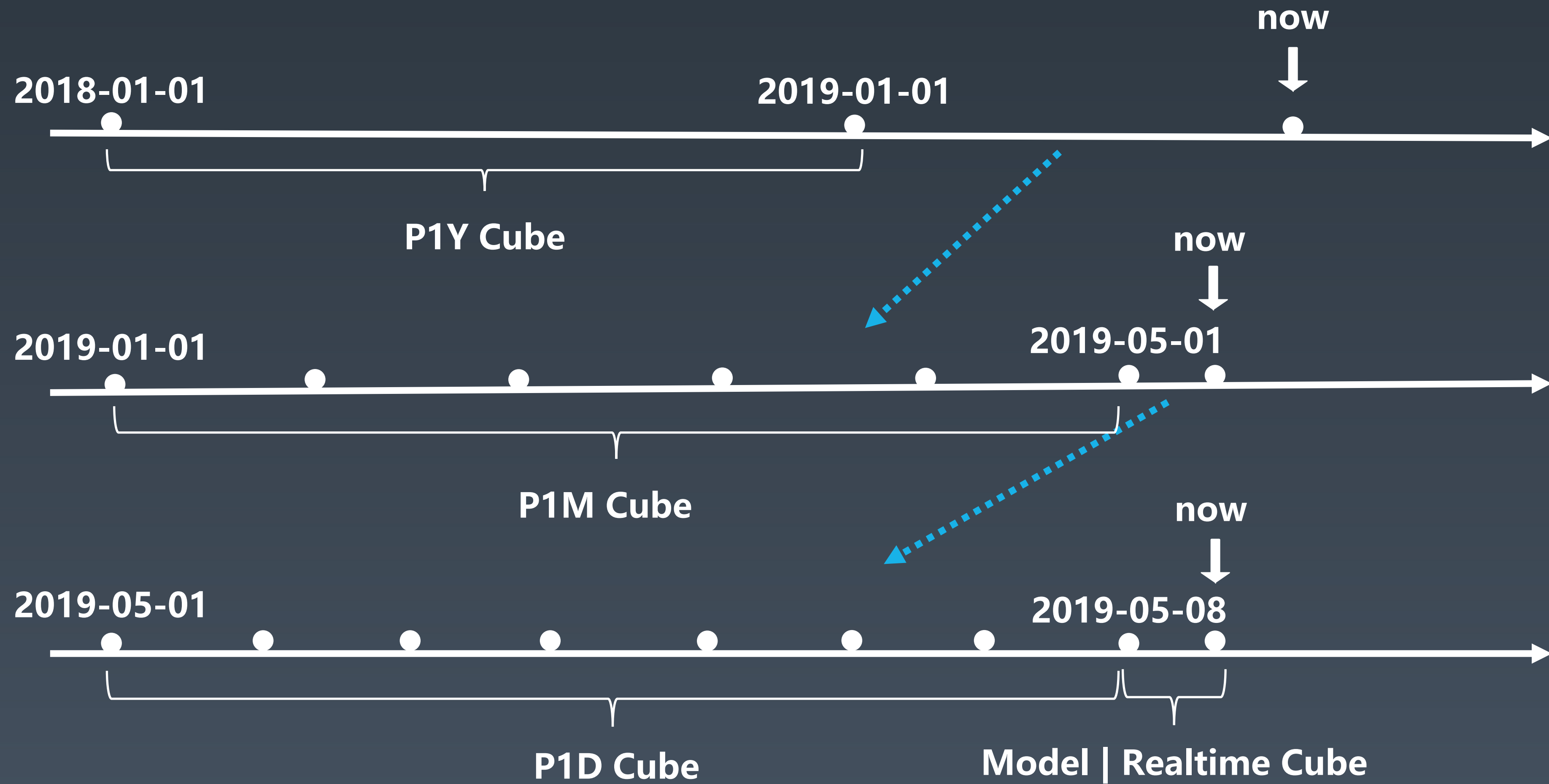
元数据



查询引擎



分段汇总



预计算

复杂查询

dimA, dimB, dimC	111
dimC	001
dimA, dimC	101

预定义SQL

```
select ...  
from ...  
...  
group by ??  
...
```

cuboid	dimA	dimB	dimC	metric1	metric2
111	xx	xx	xx	xx	xx
111	xx	xx	xx	xx	xx
001			xx	xx	xx
001			xx	xx	xx
101	xx		xx	xx	xx

COST

采集



Count



Count Distinct(HLL)



Histogram

计算



where



group by



having



order by



select

应用

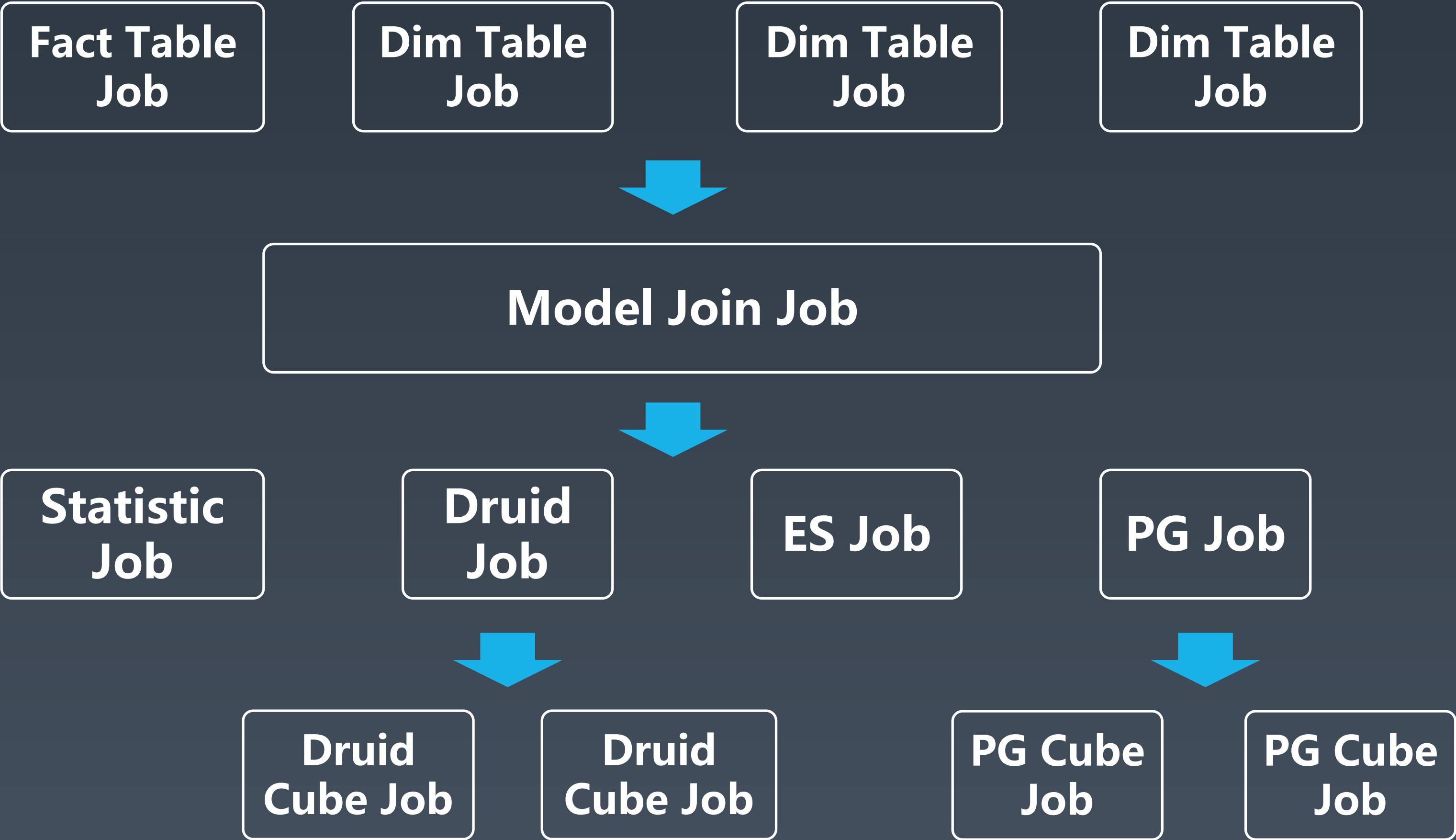
报表开发

路由

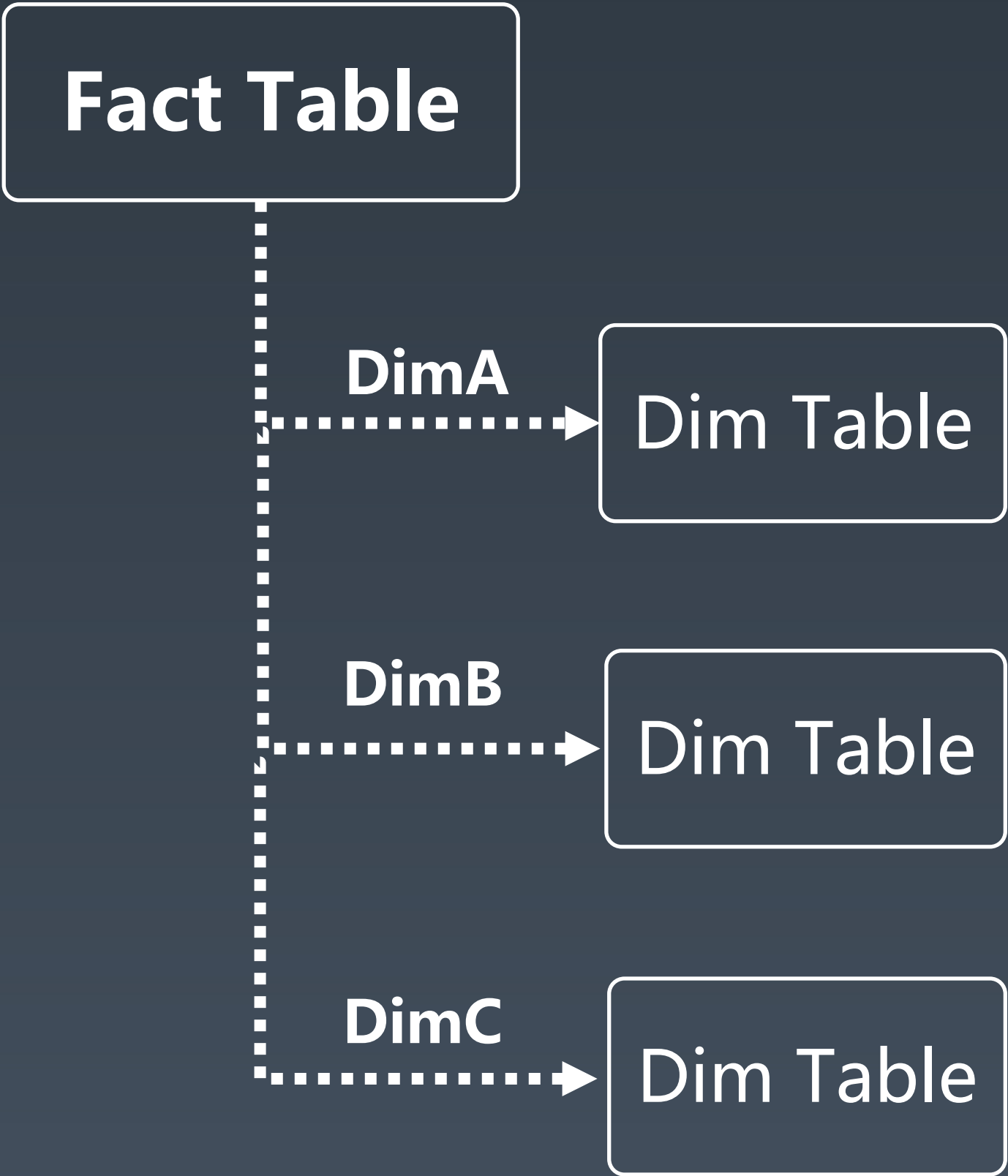
隔离

熔断

离线任务



数据倾斜优化

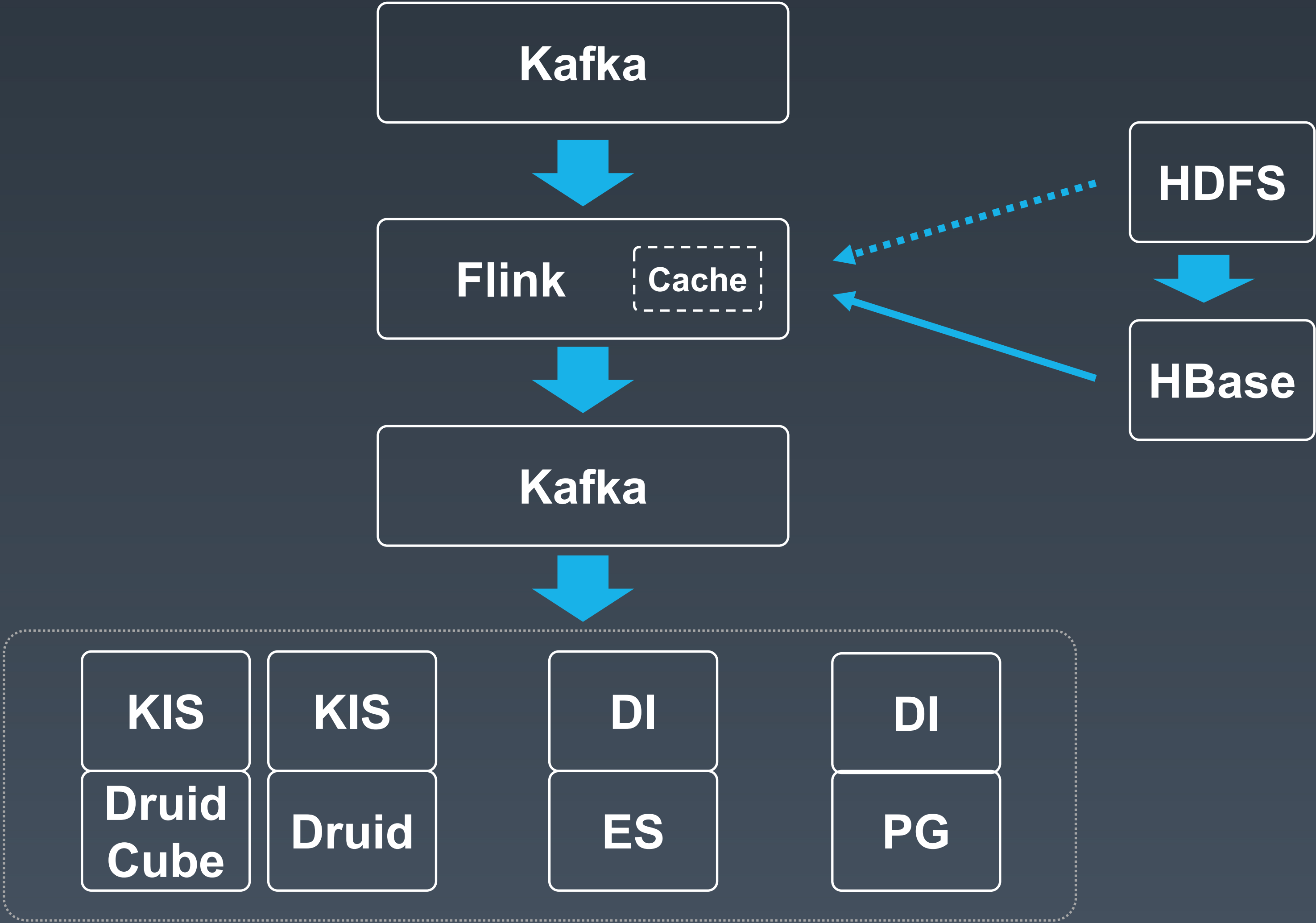


统一关联字段采样

小维表 Broadcast

大维表 分2段处理

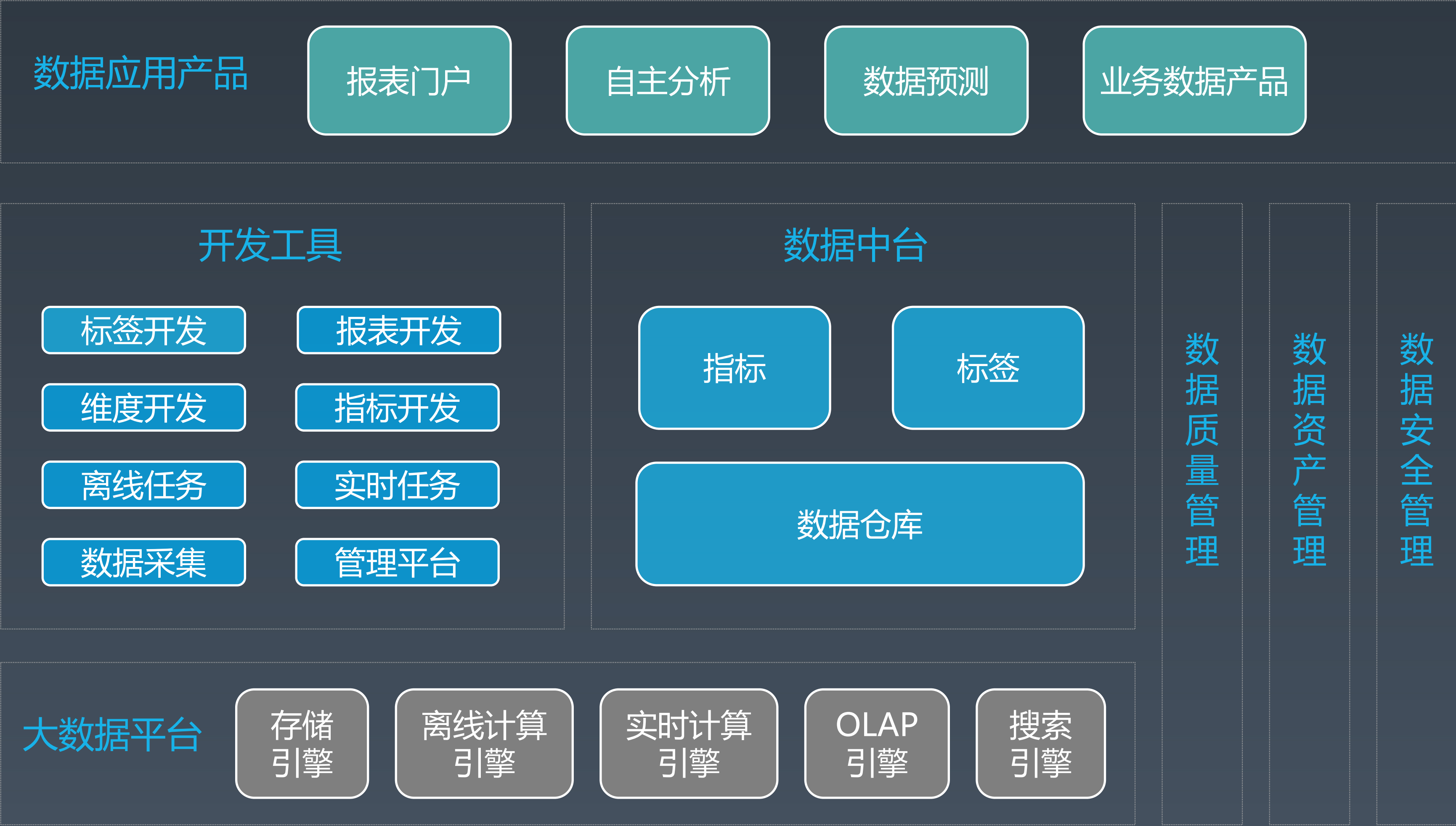
实时任务演进



OLAP整体架构



苏宁大数据



极客邦科技 会议推荐2019

5月

QCon 北京

全球软件开发大会

大会: 5月6-8日
培训: 5月9-10日

QCon 广州

全球软件开发大会

培训: 5月25-26日
大会: 5月27-28日

6月

GTLC
GLOBAL
TECH LEADERSHIP
CONFERENCE

上海

技术领导力峰会

时间: 6月14-15日

GMTC 北京

全球大前端技术大会

大会: 6月20-21日
培训: 6月22-23日

ArchSummit 深圳

全球架构师峰会

大会: 7月12-13日
培训: 7月14-15日

7月

QCon 上海

全球软件开发大会

大会: 10月17-19日
培训: 10月20-21日

10月

ArchSummit 北京

全球架构师峰会

大会: 12月6-7日
培训: 12月8-9日

12月

GMTC 深圳

全球大前端技术大会

大会: 11月8-9日
培训: 11月10-11日

AiCon 北京

全球人工智能与机器学习大会

大会: 11月21-22日
培训: 11月23-24日

11月

THANKS! | QCon th