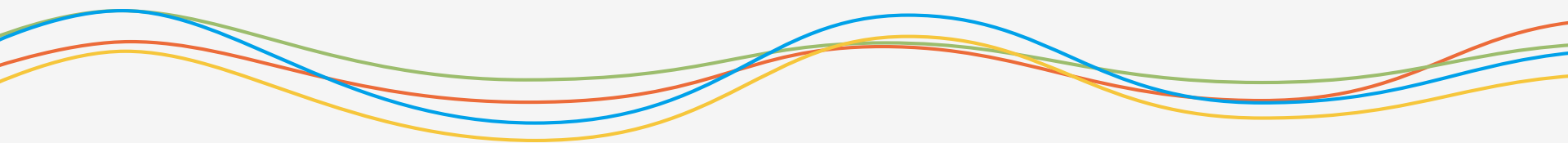


阿里巴巴数据库智能优化系统 探索与实践

阿里巴巴集团-数据库事业部
乔红麟(洪林)



乔红麟，花名洪林。2015年3月加入阿里巴巴数据库技术团队，目前负责阿里巴巴数据库智能优化产品CloudDBA开发。2007年毕业于北京邮电大学，先后从事过服务器虚拟化(Xen/KVM/VMware vSphere)，存储备份(EMC/DataDomain)，分布式存储等相关技术开发工作。

Mail: kenny.ghl@alibaba-inc.com
qiaohonglin@gmail.com



Why

阿里数据库优化服务诉求

How

CloudDBA关键技术

Now

数据库智能优化探索

Q & A

Q & A

Why

阿里数据库优化服务诉求

How

CloudDBA关键技术

Now

数据库智能优化探索

Q & A

Q & A

数据库优化服务诉求 – 业务诉求

- 业务发展速度远超DBA人员发展速度
- 数据库规模大，成本压力
- 数据库性能带来的稳定性风险(如烂SQL)
- 业务负载变化常态化，报警前故障已发生
- 业务场景复杂，数据库性能容量预估难
-

服务产品化

全局规模优化

主动诊断

智能异常发现

容量预估

数据库优化服务诉求 – 用户诉求

- “数据库运行的怎么样？”
- “数据库调用时延为啥变高了？”
- “数据库刚优化了怎么又慢了？”
- “上次的优化效果怎么样？”
- “我们需要DBA啊....”
-

信息透明

自助化诊断优化

持续优化

量化跟踪，流程闭环

产品化输出

CloudDBA是一个数据库智能优化产品，面向开发人员提供自助化诊断优化服务，致力于成为用户身边的数据库专家。

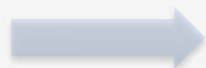


人

工具化

自动化

智能化



CloudDBA

用户身边的数据库专家

Why

阿里数据库优化服务诉求

How

CloudDBA关键技术

Now

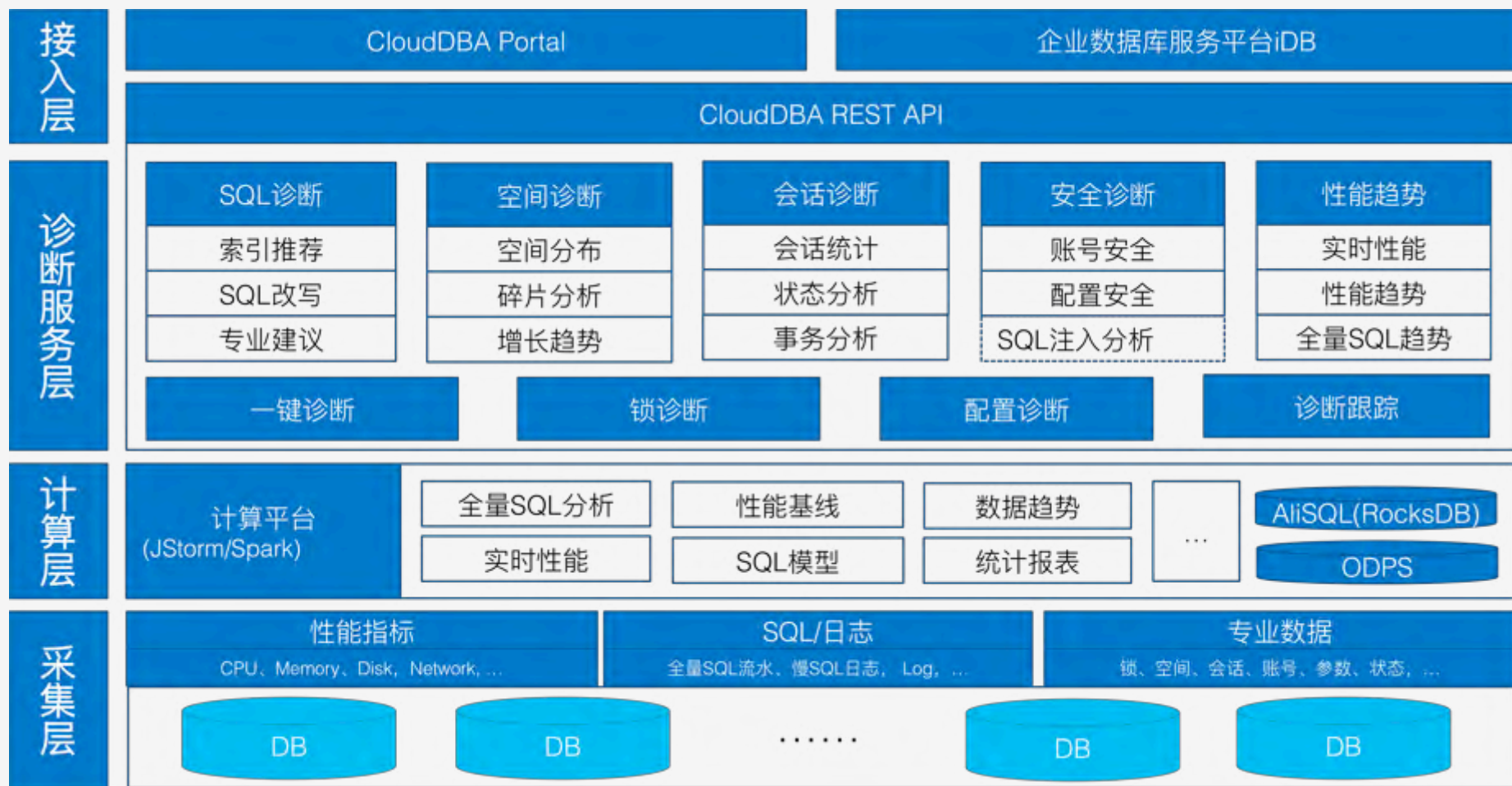
数据库智能优化探索

Q & A

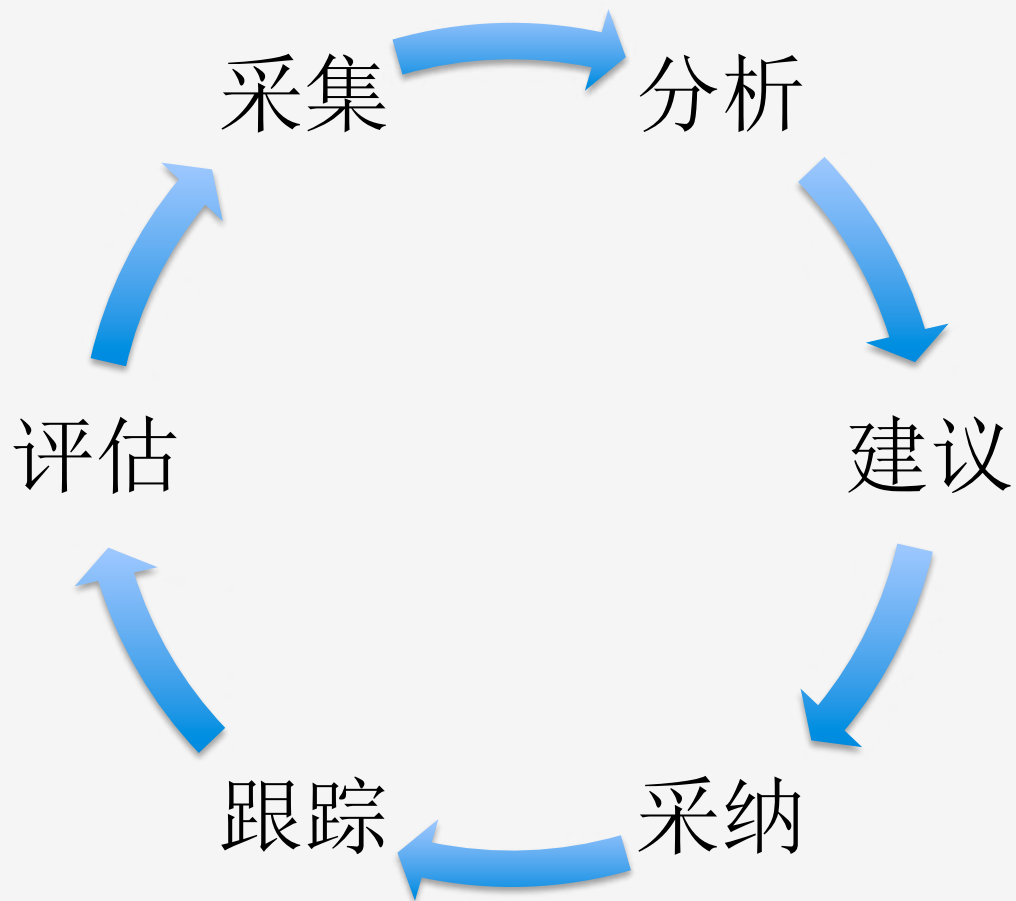
Q & A

CloudDBA整体架构

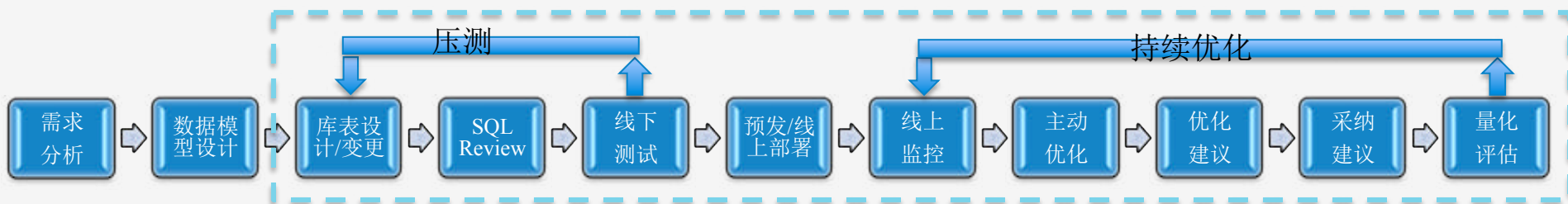
APMCon



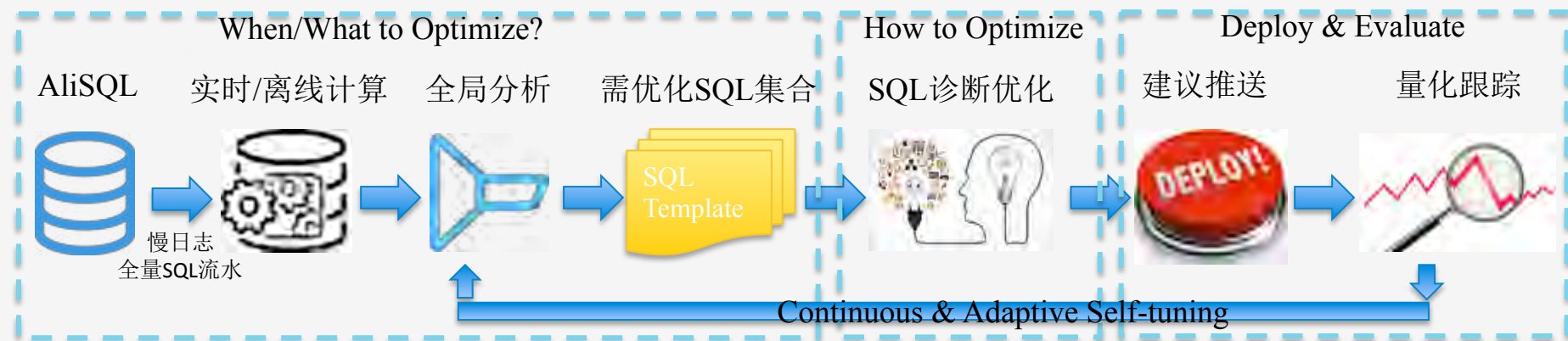
- 自助化
- 流程闭环
- 主动诊断
- 规模优化
- 数据驱动
- 智能化



开发流程闭环



优化流程闭环



SQL优化 – What-if Optimizer

➤ 全局视角

➤ 代价评估(CBO)

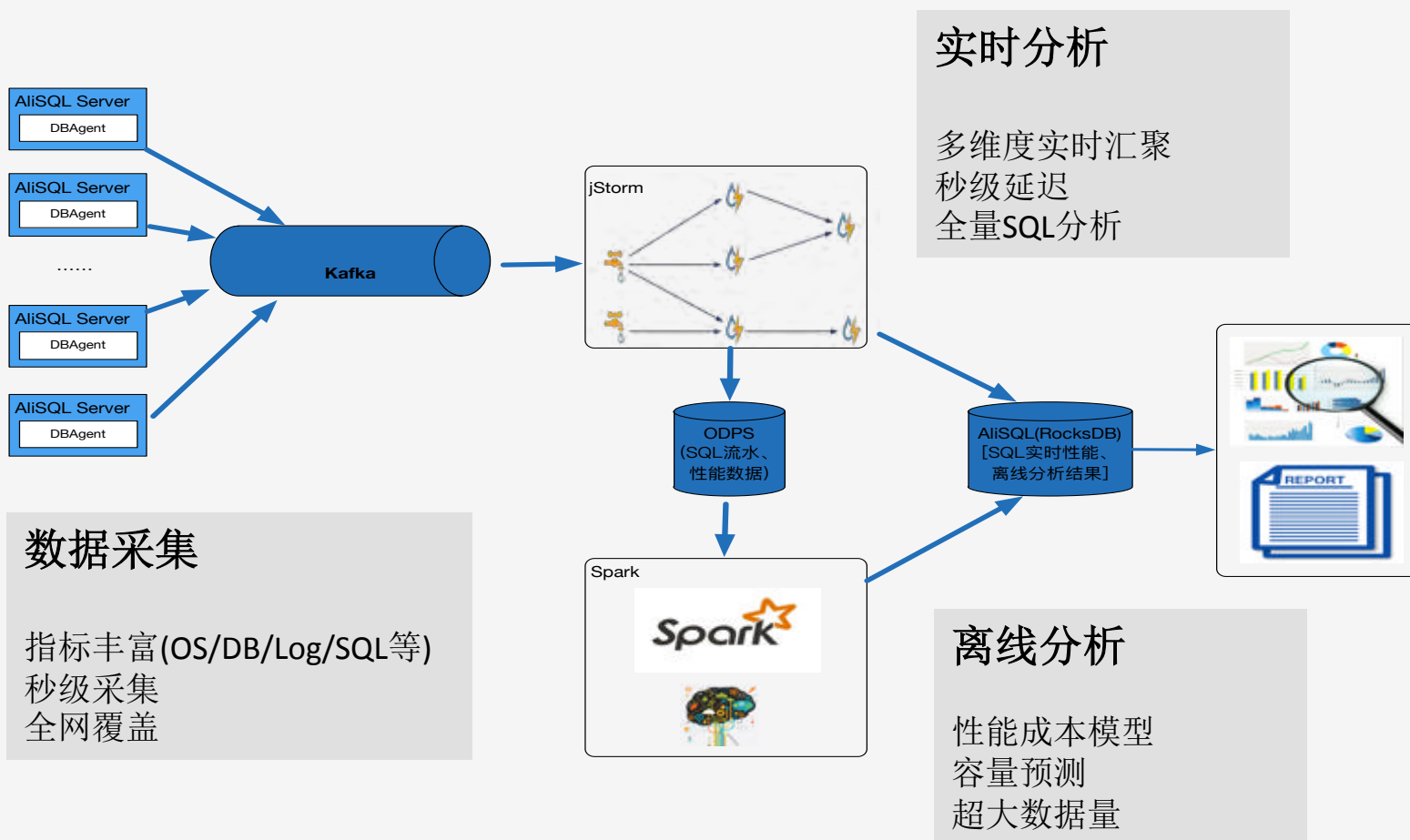
➤ 动态采样



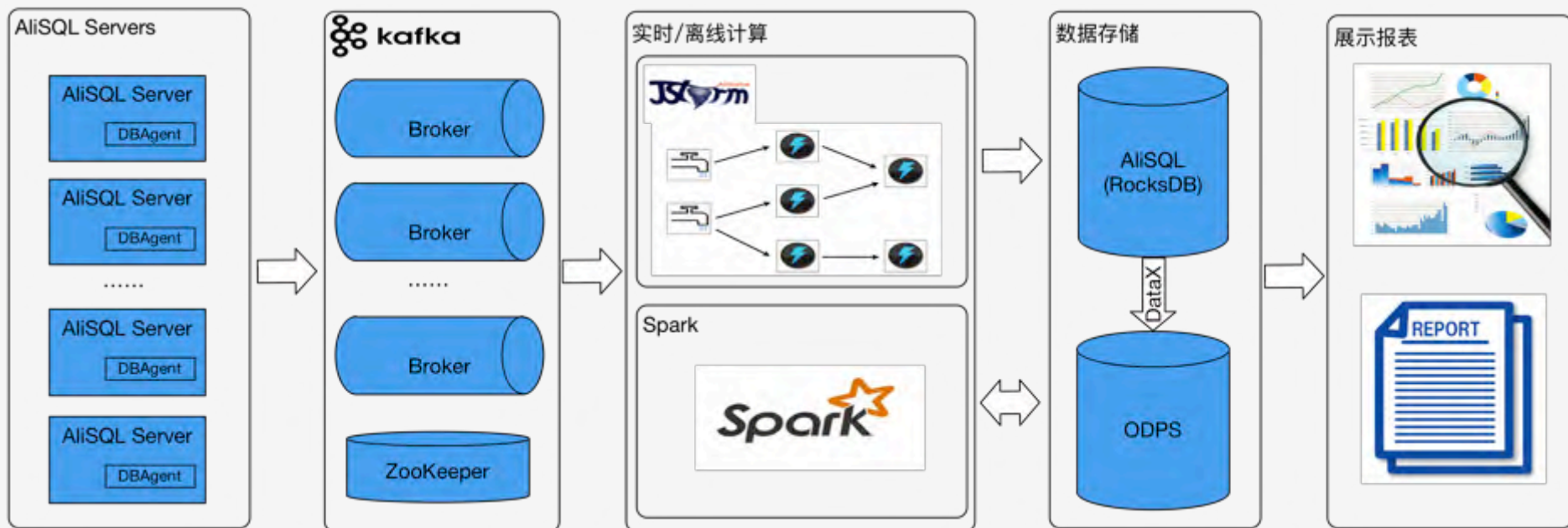
- 实例迁移？
- 紧急扩容？
- 删除数据？



数据驱动 - 数据通道



数据驱动 – 全量SQL分析



SQL采集

- 内核轻量级SQL采集
- 全部SQL流水，全网覆盖

SQL计算/分析/存储

- 实时计算千万级SQL/秒
- SQL流水离线计算百TB级/天

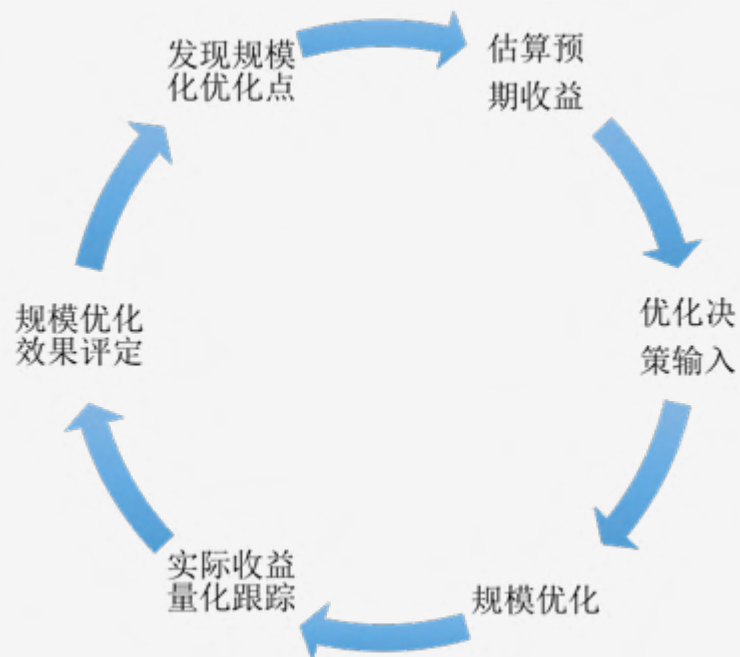
APMCon

-



数据驱动 - 全局优化

- 量化全网优化收益
- 建立全网性能成本模型
- 全网性能成本基线
- 发现规模优化点
-



Why

阿里数据库优化服务诉求

How

CloudDBA关键技术

Now

数据库智能优化探索

Q & A

Q & A

智能优化 – 自动化到智能化

- 异常检测/关联分析
- 主动预警
- 容量预估
- 自诊断，自优化
-



图片来源



智能优化探索 - 空间预测

Goal: 基于历史增长趋势预测空间可用天数

Value: 自动扩容决策提供输入

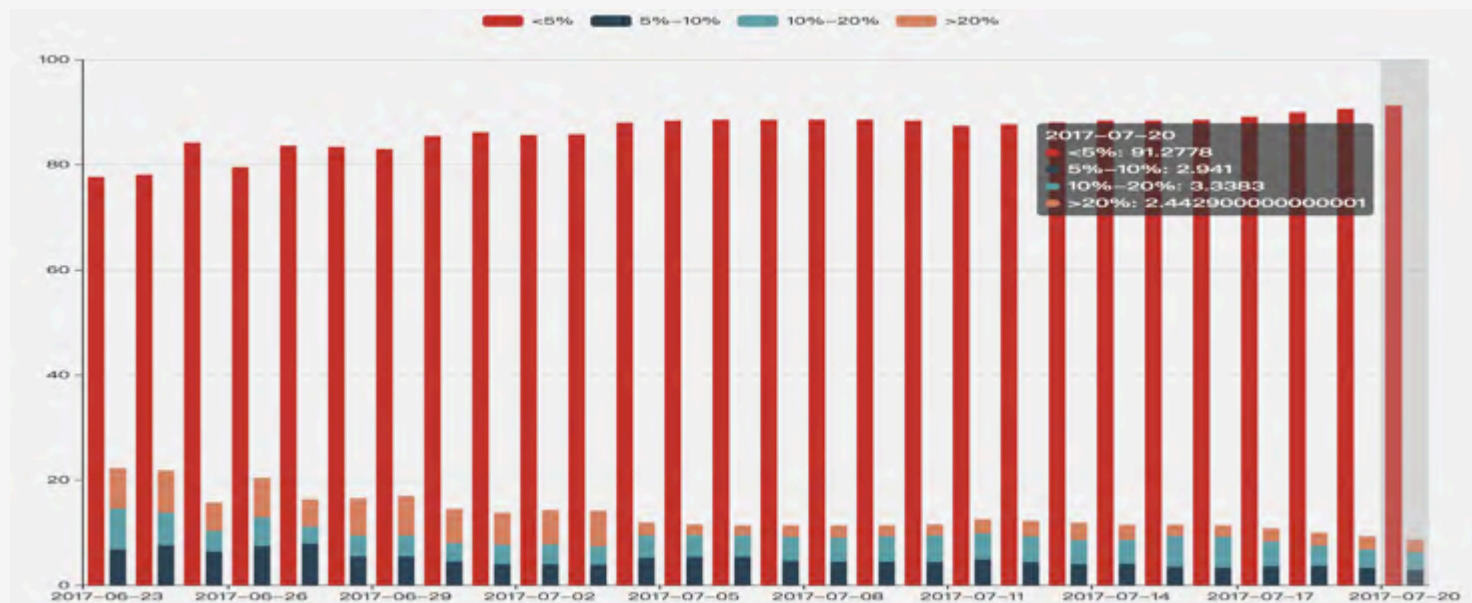
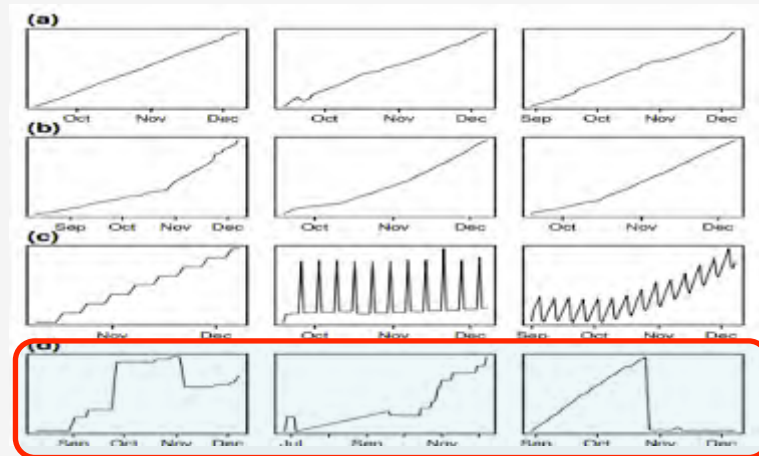
全局实例调度策略输入

Challenges: 无规律导入/删除数据

Algorithm: [Prophet](#)(Facebook)

[RDP](#)(Ramer Douglas Peucker)

Result:



➤ 被动诊断 -> 主动诊断

- ✓ 未报警DB已经异常？
- ✓ 历史特征学习？
- ✓ 快速发现异常？



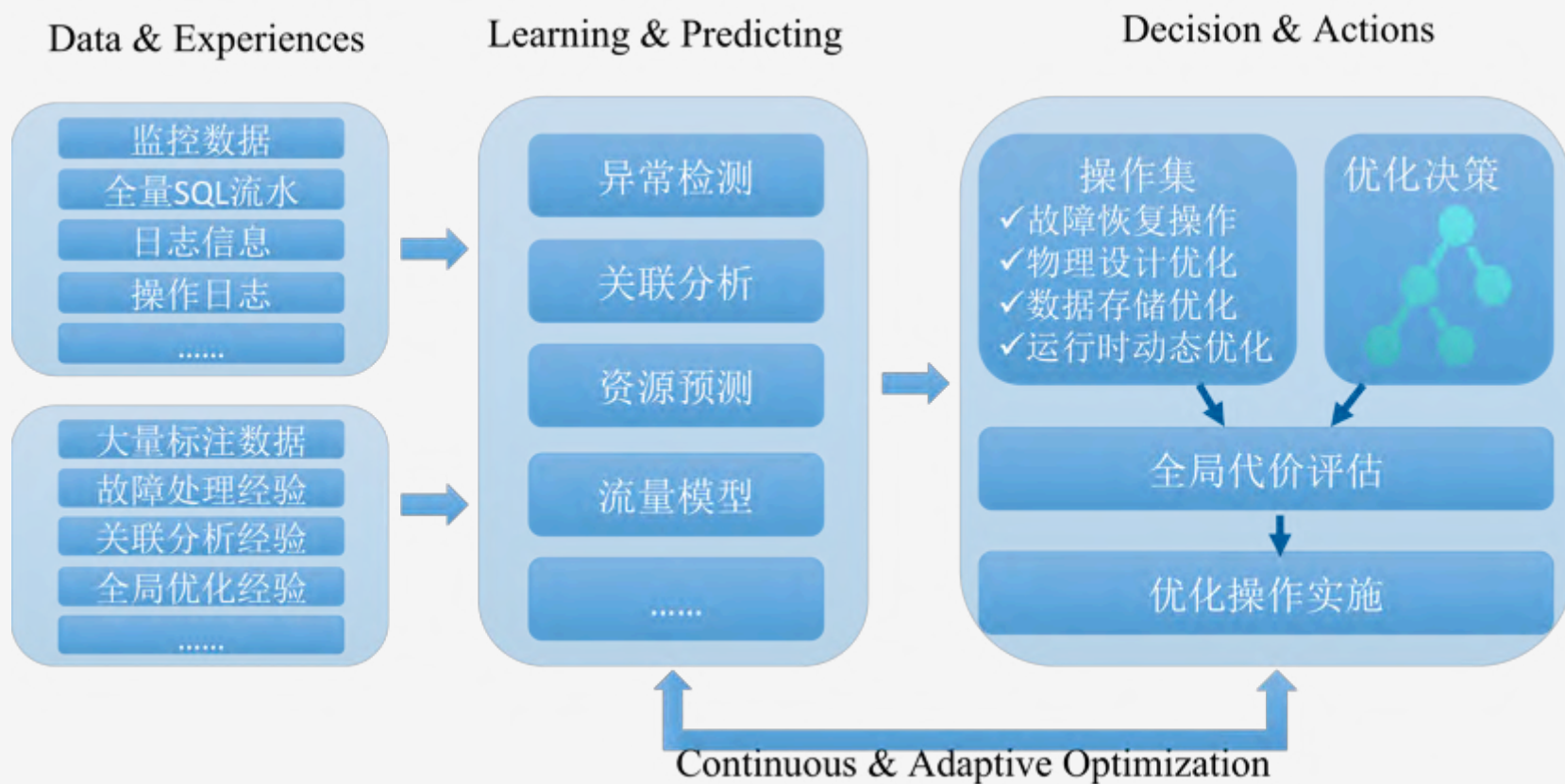
- 训练数据筛选
- 人工标注工具开发
- 异常形式化定义

- 基于RNN(LSTM)
- 基于VAE/VRNN
- F-Score

挑战：

1. 数据预处理(数据缺失, 主备切换等)
2. 特征提取: 秒级KPI局部抖动特征形式化描述(趋势/幅度/邻近窗口相关性/残差等)
3. 秒级数据深度学习模型层数较多, 训练成本高
4. 大规模场景模型的泛化能力

智能优化探索 - 自诊断，自优化



Why

阿里数据库优化服务诉求

How

CloudDBA关键技术

Now

数据库智能优化探索

We are hiring!!!

Java技术专家/高级专家

- ✓ 数据库诊断优化
- ✓ Java后端开发
- ✓ 数据处理/数据分析(Storm/Spark等)

算法专家/高级专家

- ✓ 数据挖掘
- ✓ 机器学习

Mail: kenny.ghl@alibaba-inc.com
qiaohonglin@gmail.com



THANK YOU

