

OceanBase在蚂蚁金服 的智能运维实践之路

虞舜



OceanBase

About OceanBase

阿里巴巴、蚂蚁金服自主研发的金融级
分布式关系数据库



微信公众号
OceanBase

关注OceanBase微信公众号，回复关键词“[交流群](#)”

加入OceanBase技术交流群



00:05:22

技术能力再创新记录!!

支付峰值

同比增长超一倍

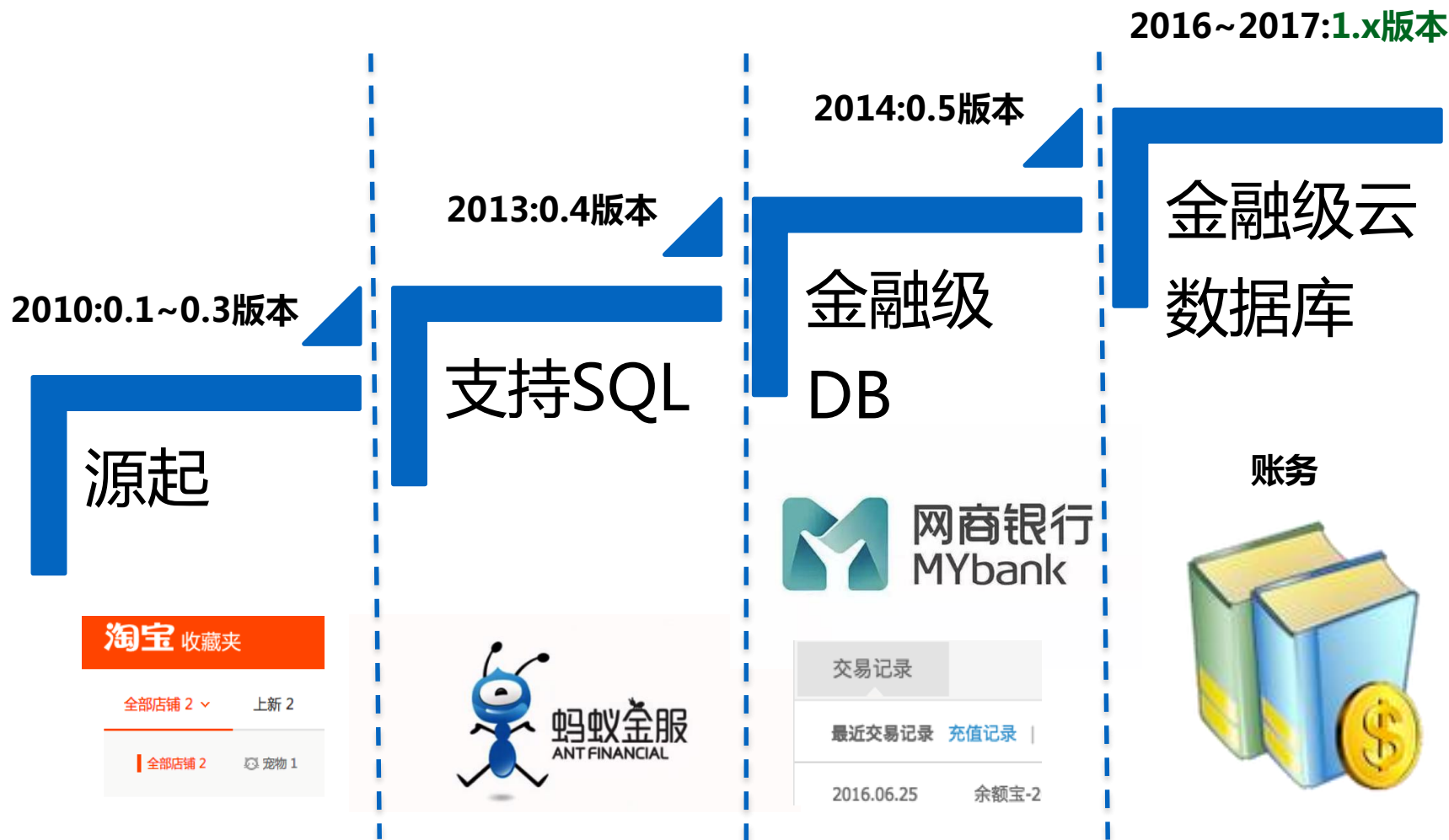
25.6 万笔/秒

中国自主研发数据库OceanBase处理峰值

4200 万次/秒



OceanBase: 回顾



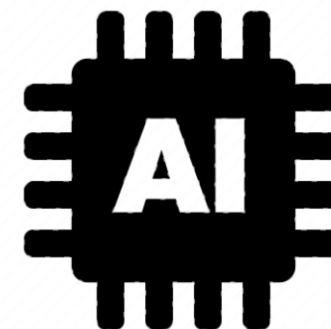
OceanBase : 现在



全局快照



混合负载



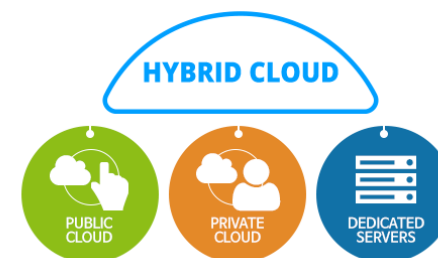
智能化的DB

ORACLE®

兼容性



性价比



容器&云

蚂蚁数据库的挑战



智能驱动的Self-Driving Database

- Self/Auto Tuning
 - SQL/配置/Schema
- Self/Auto Healing
 - 硬件故障/软件Bug/SQL异常
- Self/Auto Scaling
 - 容量/利用率/成本

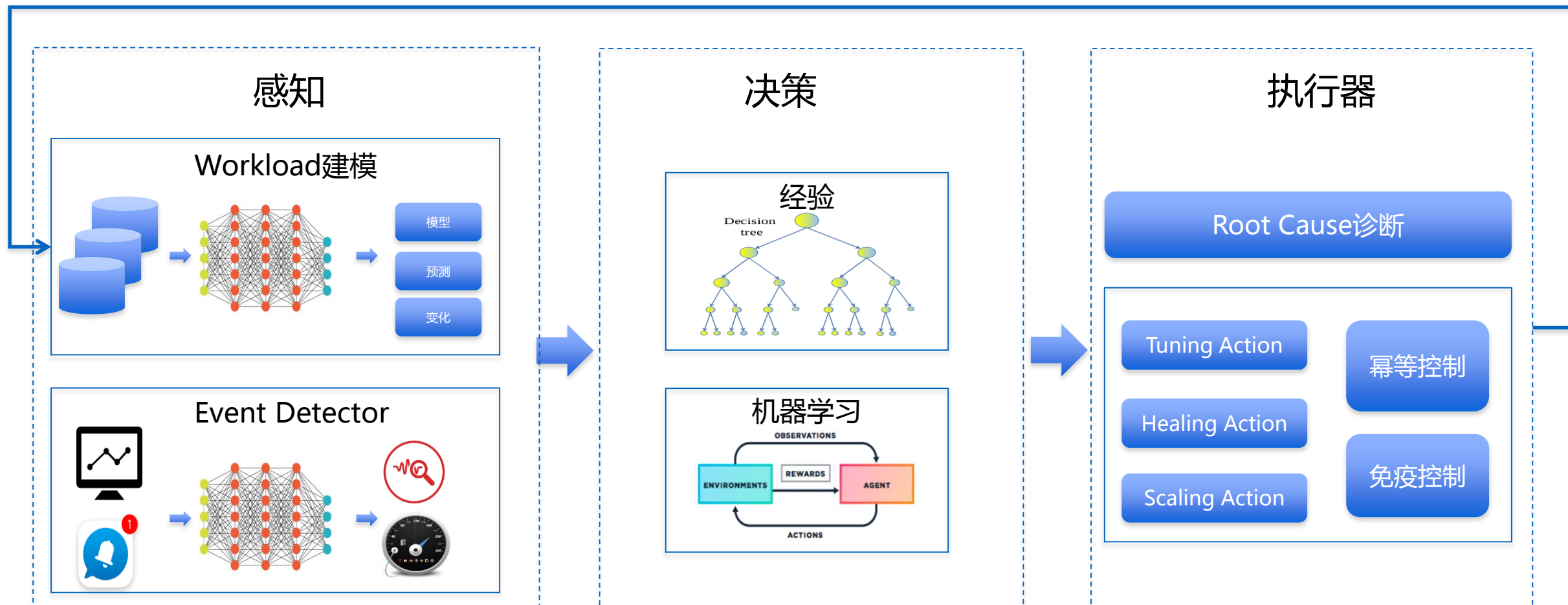
SELF-DRIVING CAR



SIGMOD以及业界趋势

- SIGMOD 2018 Paper
 - Query-based Workload Forecasting for Self-Driving Database Management Systems
 - P-Store: An Elastic Database System with Predictive Provisioning
- 工业界
 - Oracle 18c : Autonomous Database

智能化数据库系统的架构



智能化的基石

- 灵活可扩展、可定制的OceanBase
- 自动化、稳定、具备强大处理能力的平台
- 数据库专家经验

- Workload Modeling
 - 随机过程、回归、RNN
- 统一事件
 - Anomaly Detection
 - LSTM、ARIMA、HoltWinters、Ensemble
 - Engineered Events
 - 宕机、迁移、DB内部状态变化、异常巡检

决策

- 本质
 - Input : 事件、Matric Data
 - Output : Action Operator
 - 优化目标 : $\min(RT) \setminus \text{Max}(TPS) \setminus \text{Max}(\text{Success Rate})$
- 策略
 - 基于经验的决策
 - 基于学习的决策
 - 如何积累学习中需要的数据 ?

执行器

- 目标
 - 幂等、最小化系统影响
- Operator
 - 模型
 - 执行代价、系统影响、预计耗时
 - 可免疫、可回滚
 - 原子
 - Add Index/Change Index/Switch Leader/Add Replica...
 - 组合
 - 链路扩容/机房切换/Add Zone.....

智能化大促



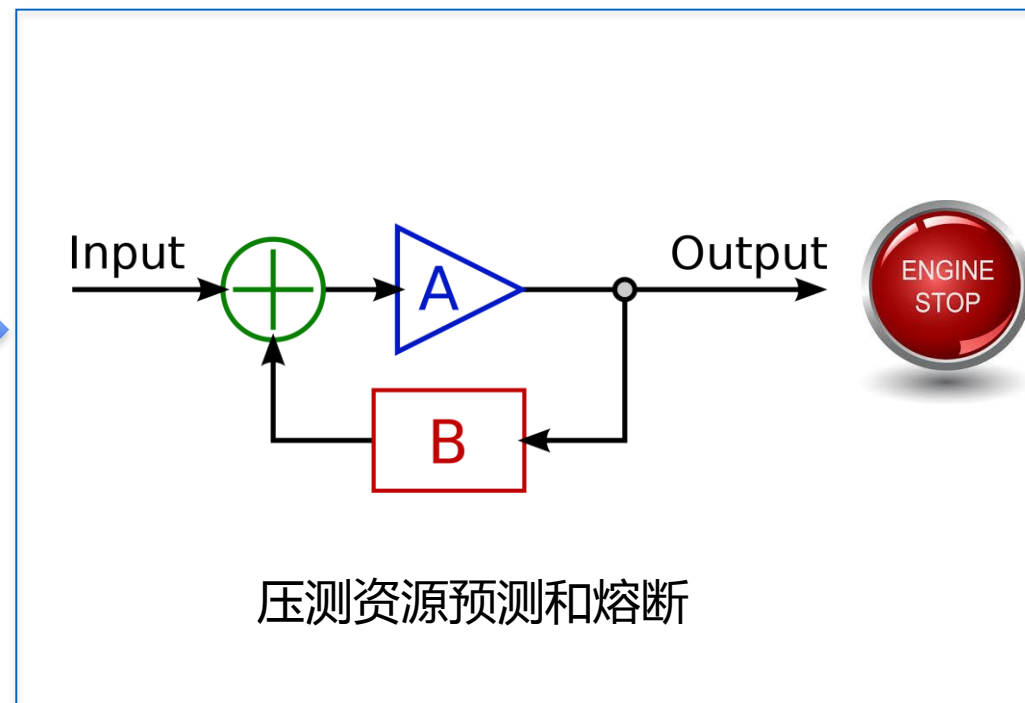
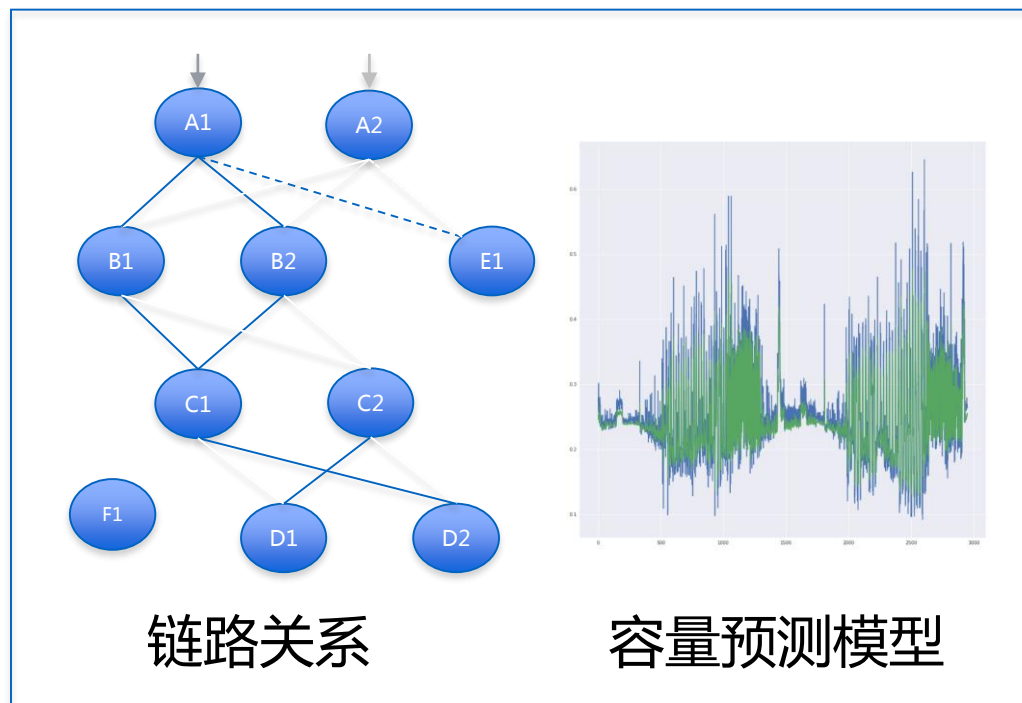
精细化的容量

持续优化

用户无感的压测

资源利用率提升

智能化的大促



ScaleUp & Rebalance
容量刻画 业务类型刻画

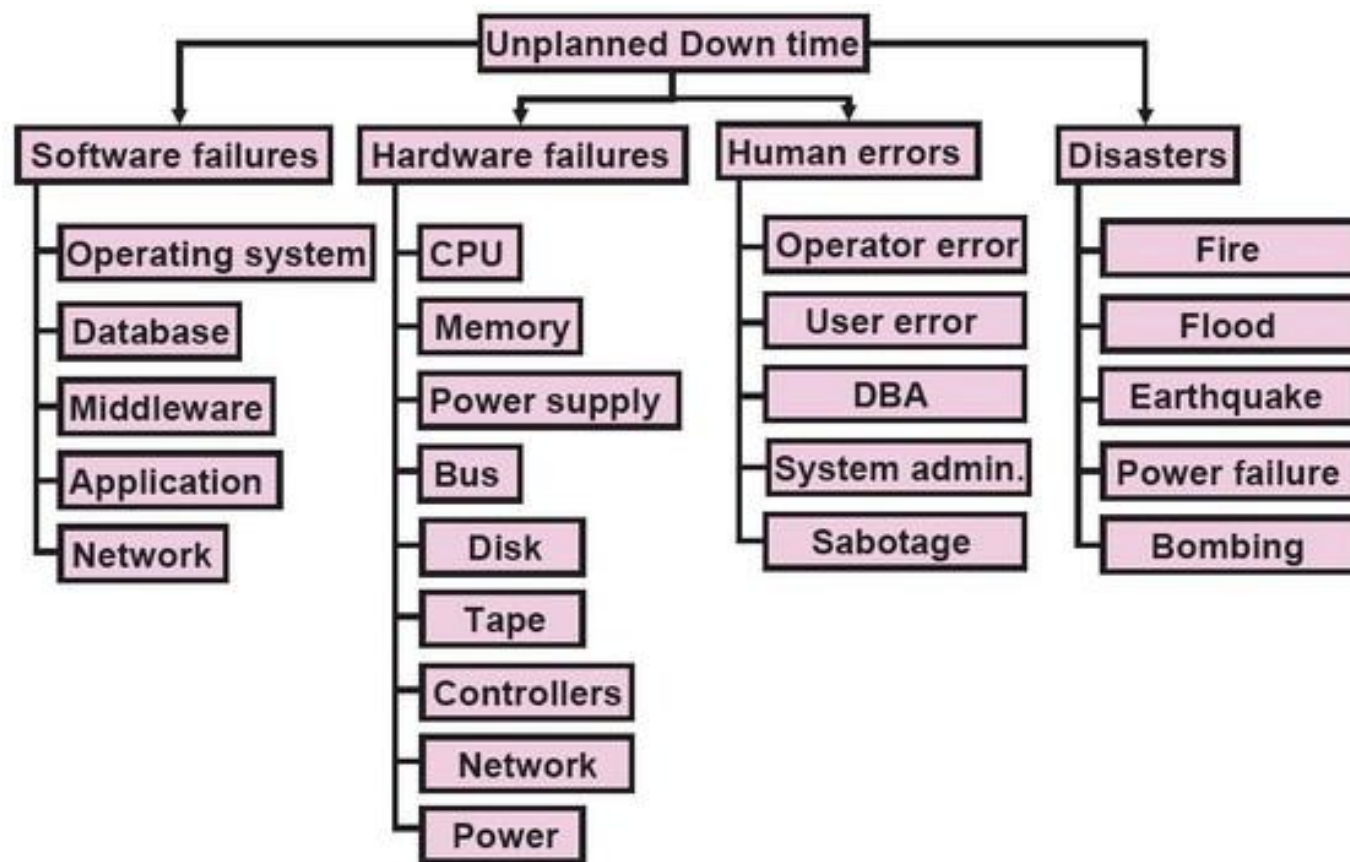
SQL Tuning
Plan Status/Resource Usage

稳定性



- 很少的异常会影响到很多用户的体验
- 从DC级容灾到精细化的异常恢复

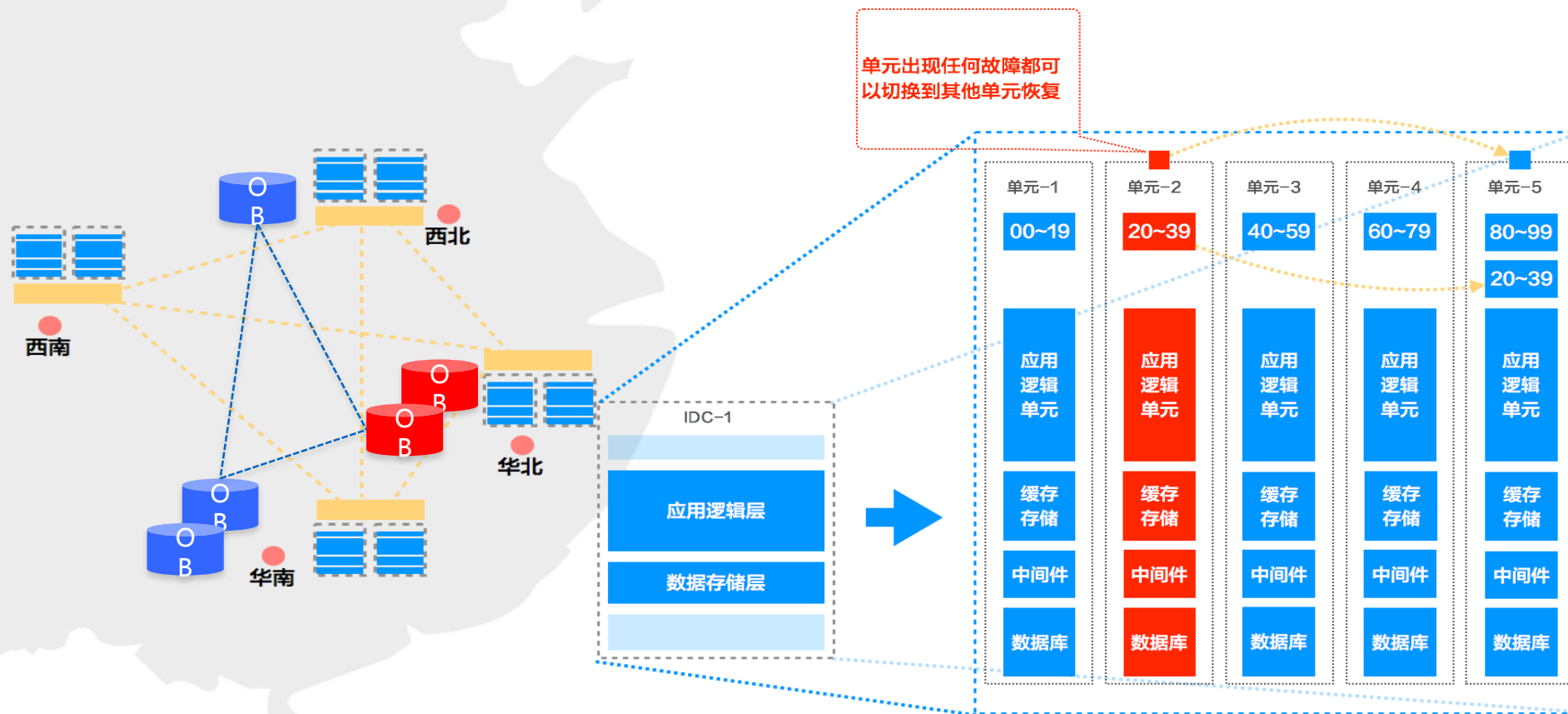
异常类别



• 通断类

• 非通断

Zone/Region级别容灾

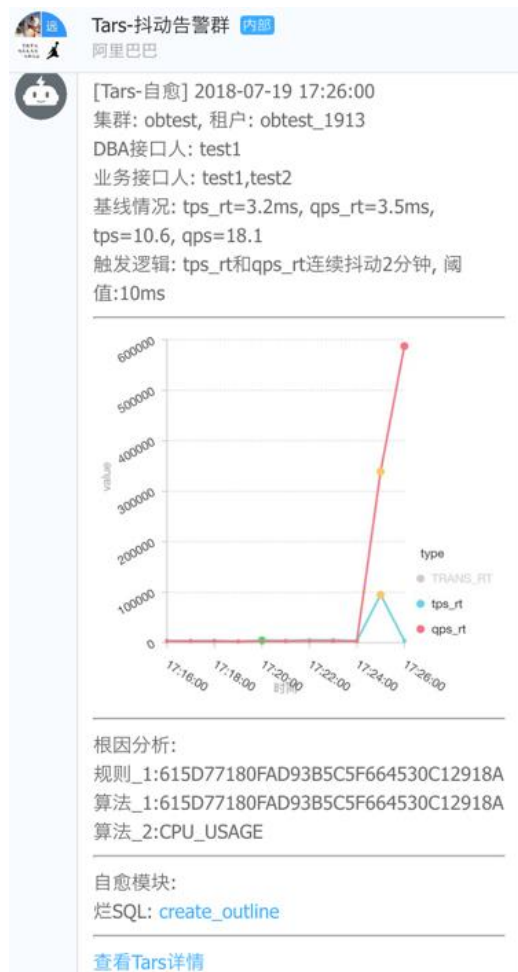


- 通常是有损的，有比较大的业务代价

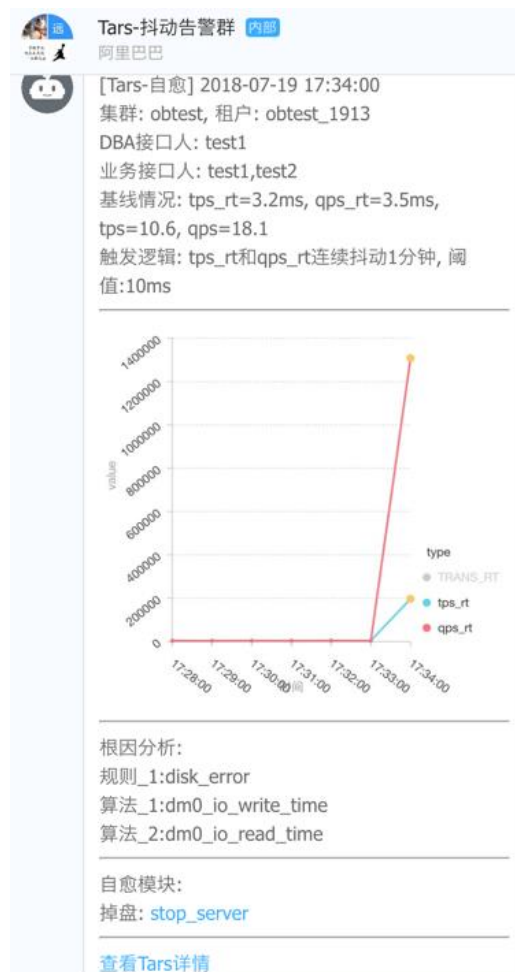
Self-Healing - 精细化异常恢复

- 目标异常：非通断异常
 - Bad SQL、IO Hung、Software Bug
- 5分钟快速恢复
 - ML : Anomaly Detection → RCA → Action Operator
 - Rule : 异常场景→DBA经验→Action
- 如何防止问题恶化？
 - 5Redo、切主不杀事务、幂等控制、柔性强切

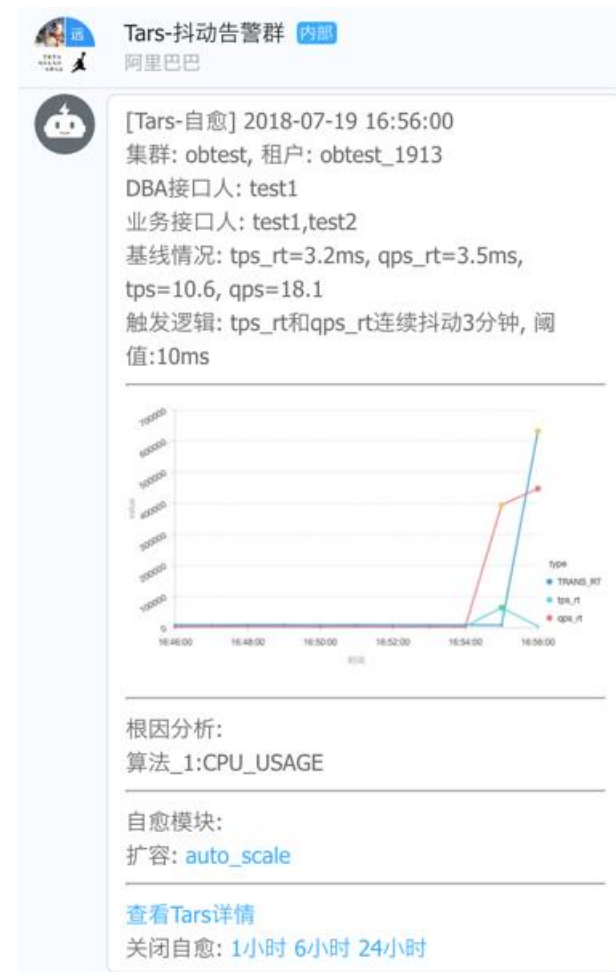
Self-Healing



SQL异常



IO异常



容量不足

下一步计划



**Autonomous
OceanBase**

谢谢

THANK YOU

虞舜



OceanBase