



阿里云数据库 (ApsaraDB) 智能化支撑系统 天象

褚霸 2016/04/23

QCon

2016.10.20~22
上海·宝华万豪酒店

全球软件开发大会 2016

[上海站]



购票热线: 010-64738142
会务咨询: qcon@cn.infoq.com
赞助咨询: sponsor@cn.infoq.com
议题提交: speakers@cn.infoq.com
在线咨询 (QQ): 1173834688

团·购·享·受·更·多·优·惠

7折 优惠 (截至06月21日)
现在报名, 立省2040元/张

- 数据库云化带来的挑战
- 系统的数据化、智能化
- 天象数据产品
- 精细化服务

数据库云化带来的挑战

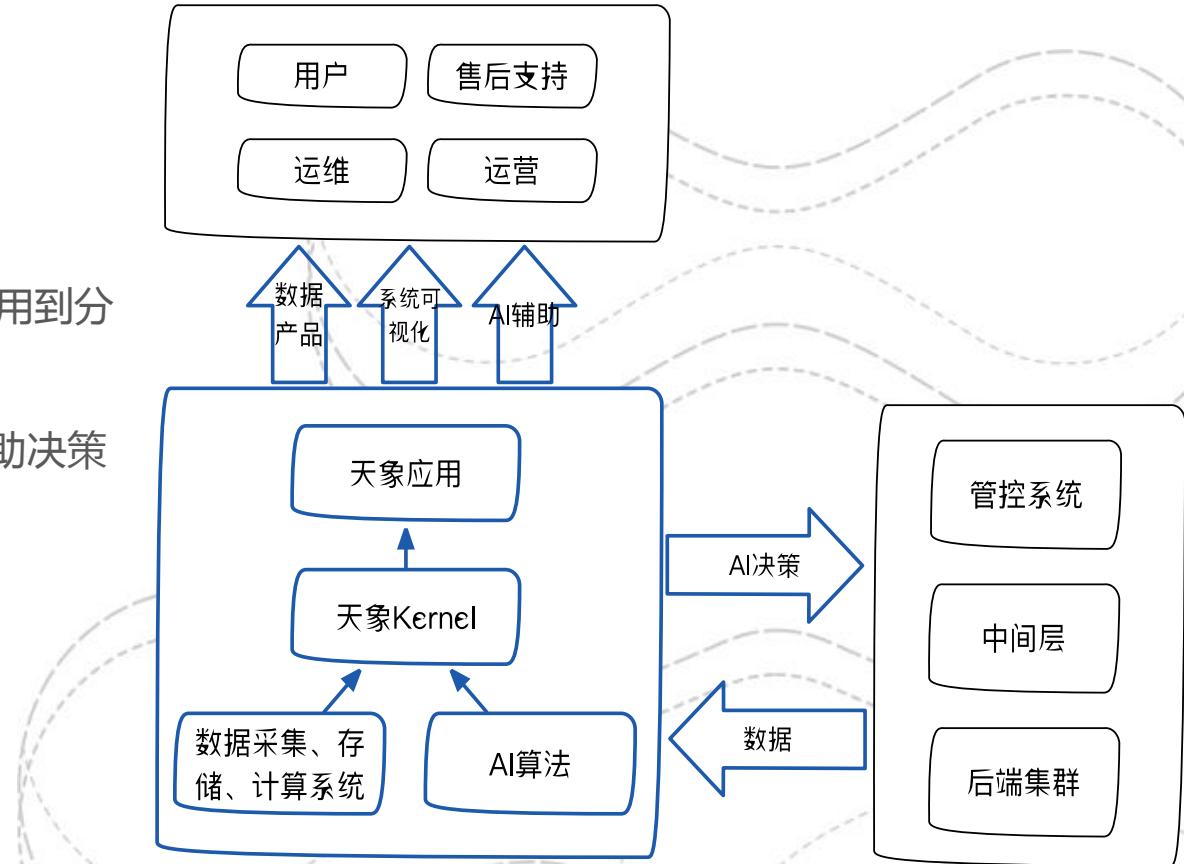
数据库云化带来的挑战

- 数据库服务本身服务质量要求苛刻
- 面对不同业务类型用户
- 用户上云后希望高附加值专业的服务
- 运维实体从传统几个到很多万级别
- 用户希望更省心，主动服务
- 高性能低成本



系统的数据化、智能化

- 实时数据处理、计算、可视化应用到分析系统自身
- 利用AI算法进行AI运维，及AI辅助决策



多样化的数据

用户真实的每个请求的访问延迟，网络探测数据，上百项操作系统指标，数据库内核核心指标，各组件自身日志

无损采集，不损耗正常服务

内核TCP / IP协议栈采集延迟，系统指标直接采集，CPU占用1%以内

压缩发送，本地初步聚合后传输，网络带宽100KB内

每日百TB级别新增数据，秒级时间序列数据

Kafka + JStorm实时数据处理架构

低成本数据存储，ApsaraDB For PetaData存储

多源异构数据使用

ETL转化为通用格式，建立数据仓库



面对的数据

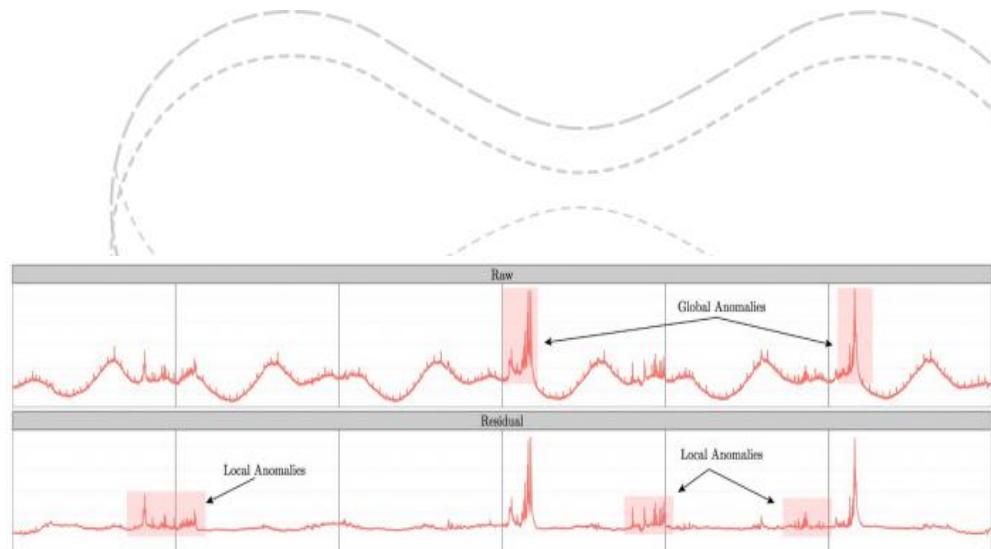
- 时间序列的指标
- 日志数据

传统的阈值和关键词只可应对简单场景

时间序列数据分析算法

人与机器，让机器懂专家经验

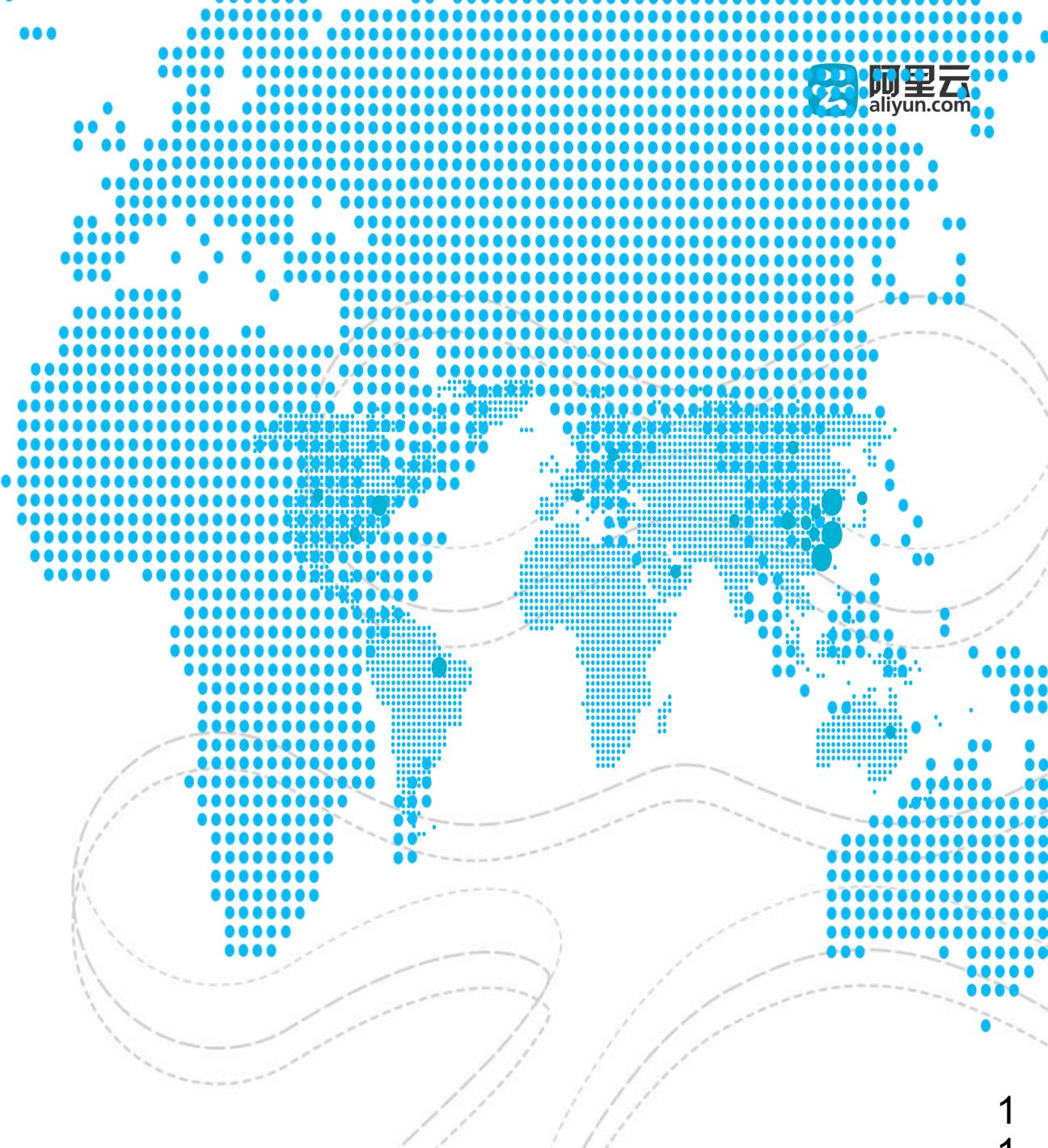
- 人逻辑分析能力强，但只能间歇性观察最多几十个指标
- 机器可以不停的观察上百万的指标
- 我们将人的分析逻辑模型化，针对问题建立判断模型，让机器根据此去判断



天象数据产品

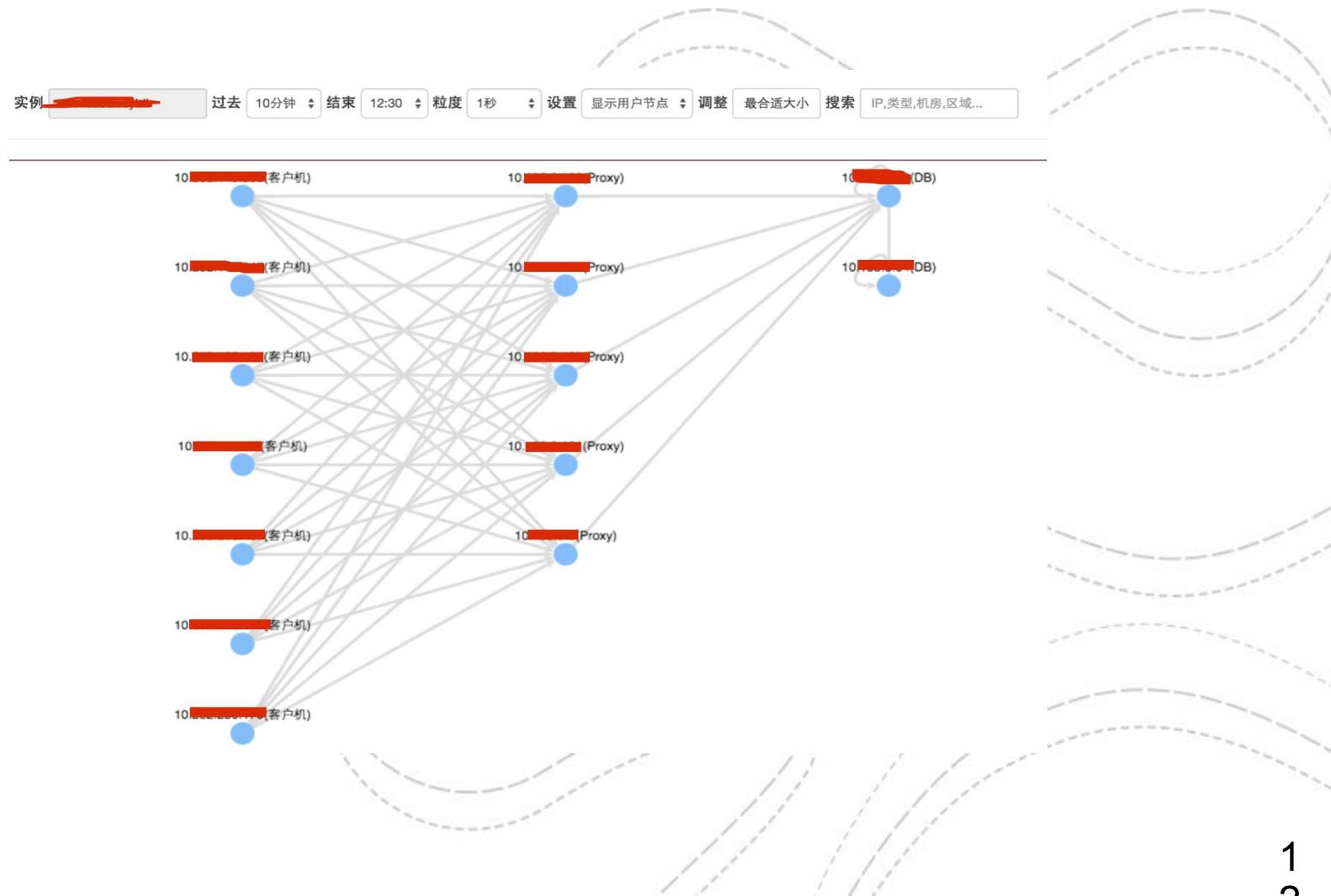
数据的魅力

- 服务质量量化
- 一键诊断
- 智能化运维
- 用户画像
- 移山资源调度



了解自身服务质量

用户真实的请求链路拓扑，每个请求都完整记录的服务质量数据



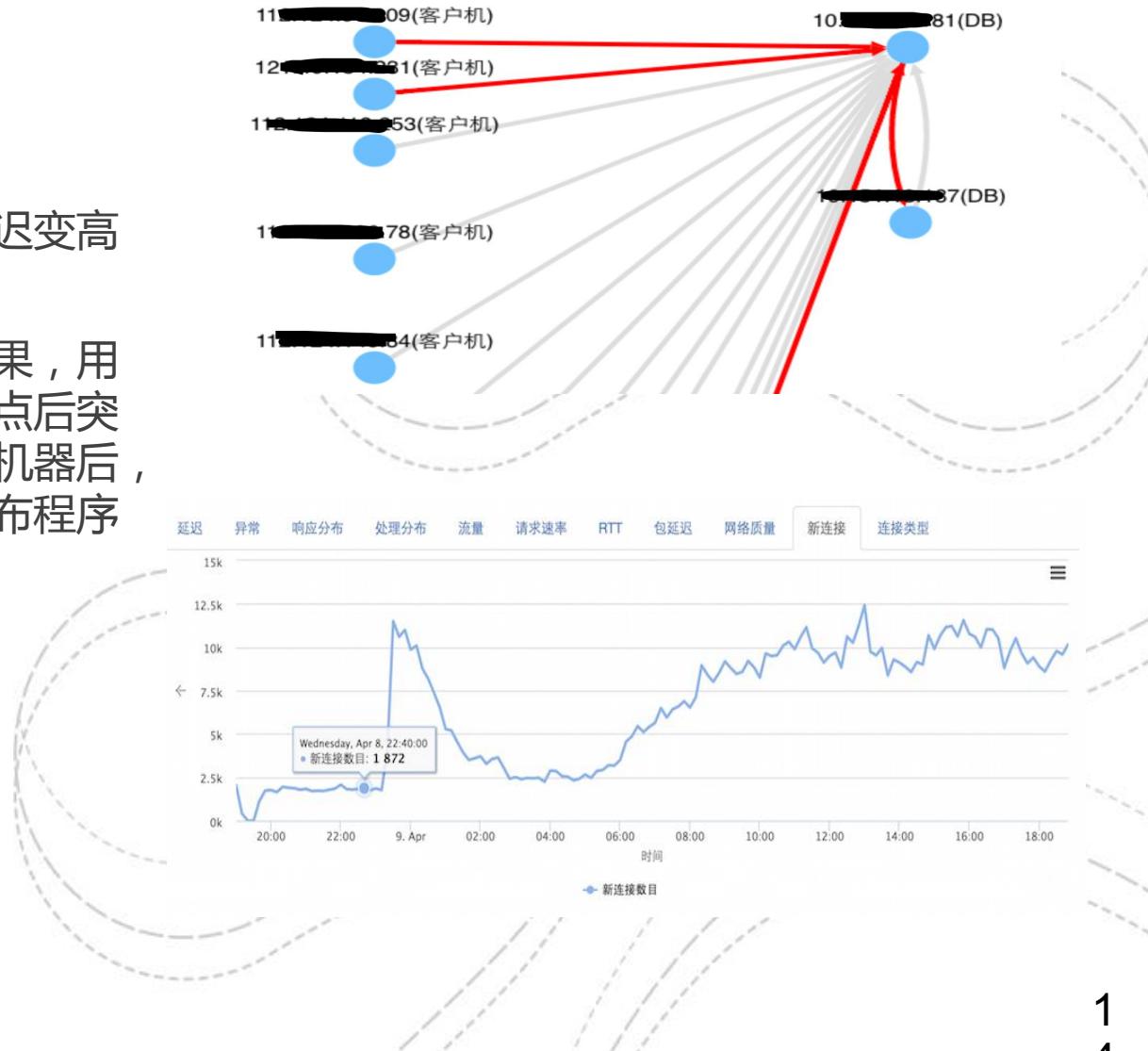
几十项指标量化服务

Proxy 10.████████



帮助用户发现程序问题

- 用户反馈数据性能降低，延迟变高
- 利用天象可视化数据分析结果，用户某两台机器，在某个时间点后突然长链接变短链接。定位到机器后，用户很快发现两台机器新发布程序有问题



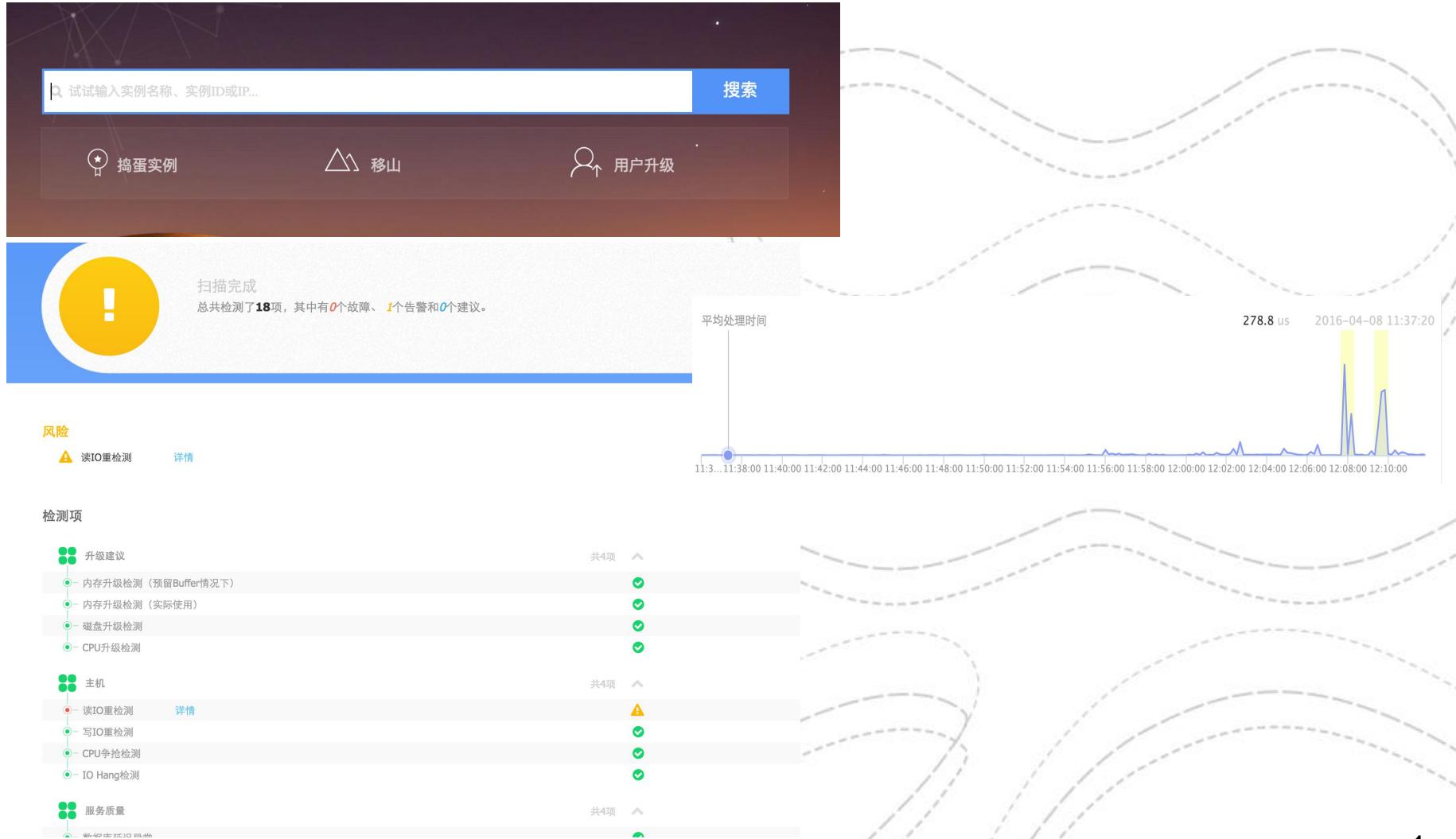
对比传统服务质量检测方式



	之前	现在
监控方式	探测	真实数据
粒度	分钟	1s
维度	单节点	链路拓扑
指标	单一	20+项
可视化	日志	拓扑图

一键诊断快速解决用户问题

天象辅助售后技术支持快速诊断用户工单，提升问题处理速度，从周缩短到秒

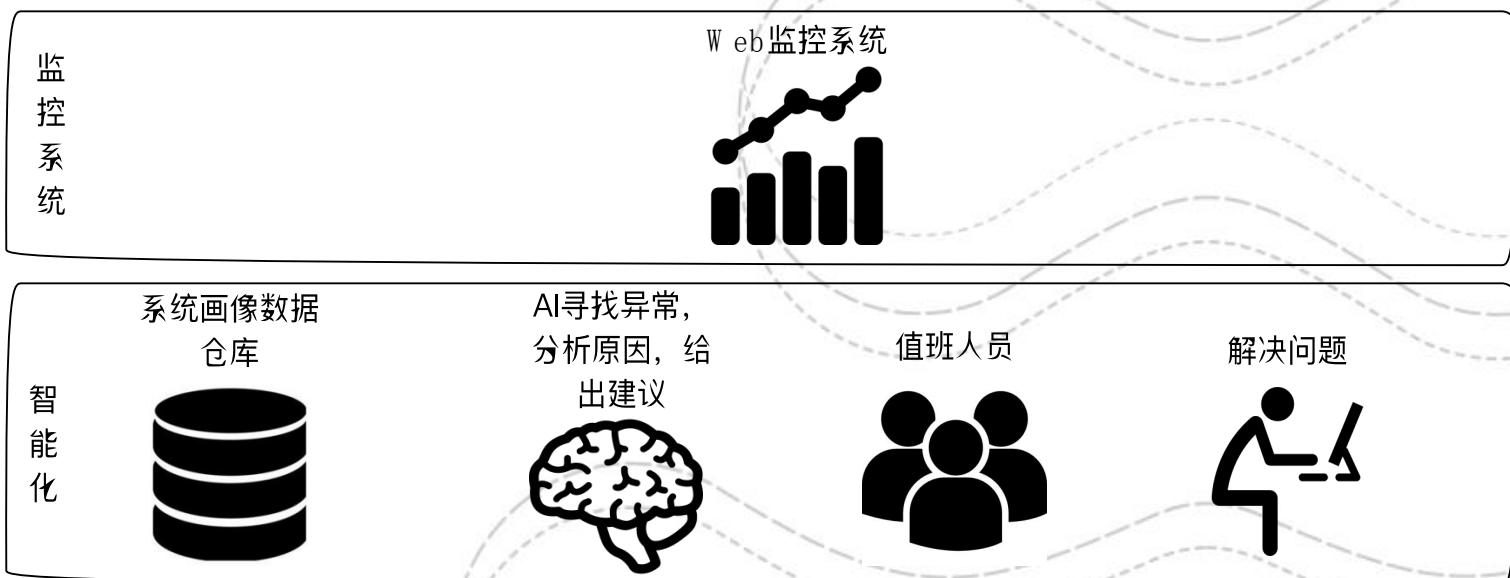


运维的演进

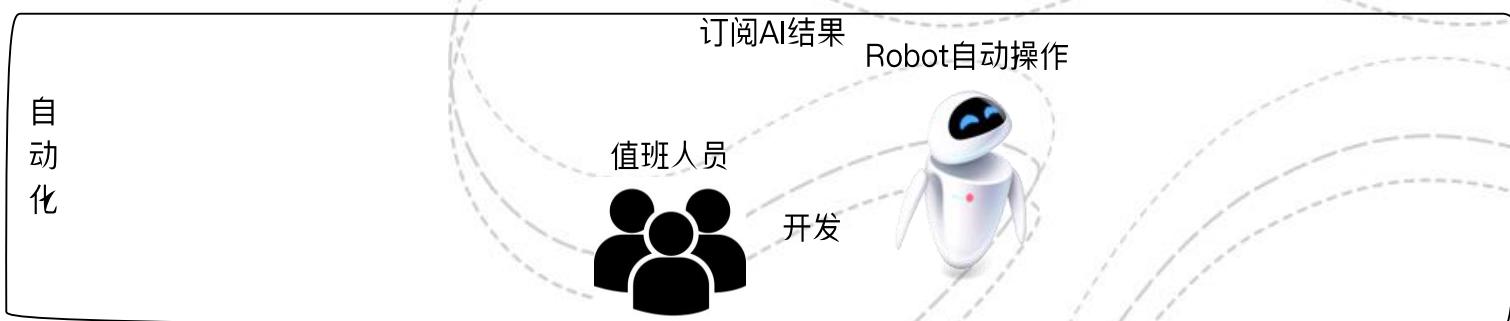
过去



目前



正在实现



异常发现和问题干预

实时发现系统异常，推送问题和影响面给相应负责人



异常发现的关键

- 准确、准确、准确

基础是有足够的为 系统画像的数据

- 系统数据仓库

专家知识

- 有能力教机器发现问题

闭环并且详细的实时指标

服务质量数据

操作系统指标

中间层指标

数据库指标

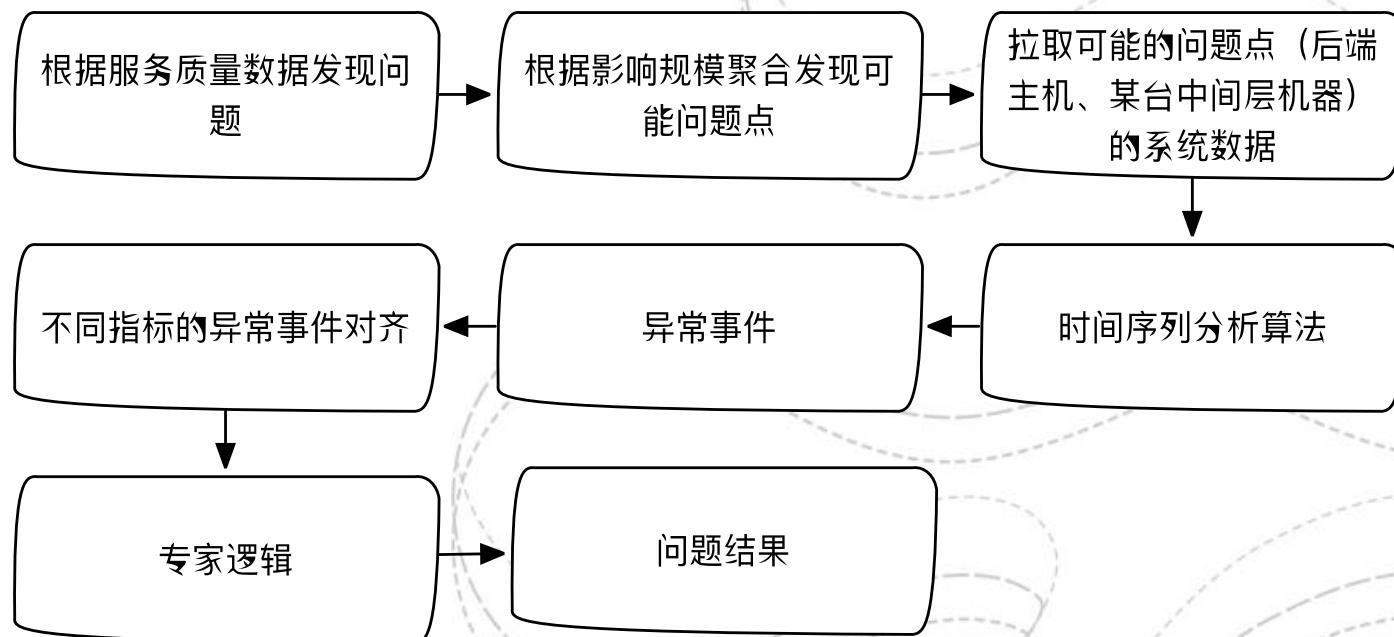
日志指标

如何实时准确挖掘异常（续）

○ 关键点

利用服务质量数据，从用户是否受到影响作为线索和判定依据

○ 分析流程



实时准确异常发现的意义



从被动运维到主动运维

- 故障后快速恢复能力

主动干预问题

天象异常发现

AI决策

管控系统

用户画像

哪一个才合适？

地域 :

华北 2 (北京) 华东 1 (杭州) 华北 1 (青岛) 香港 华南 1 (深圳) 美西 (硅谷)
亚太 (新加坡)

不同地域之间的产品内网不互通；订购后不支持更换地域，请谨慎选择 教我选择>> 查看我的产品地域

可用区 :

单可用区 请选择可用区

数据库类型 :

MySQL

版本 :

5.6 5.5

存储空间 :

250GB 10 GB 2000GB 745 GB

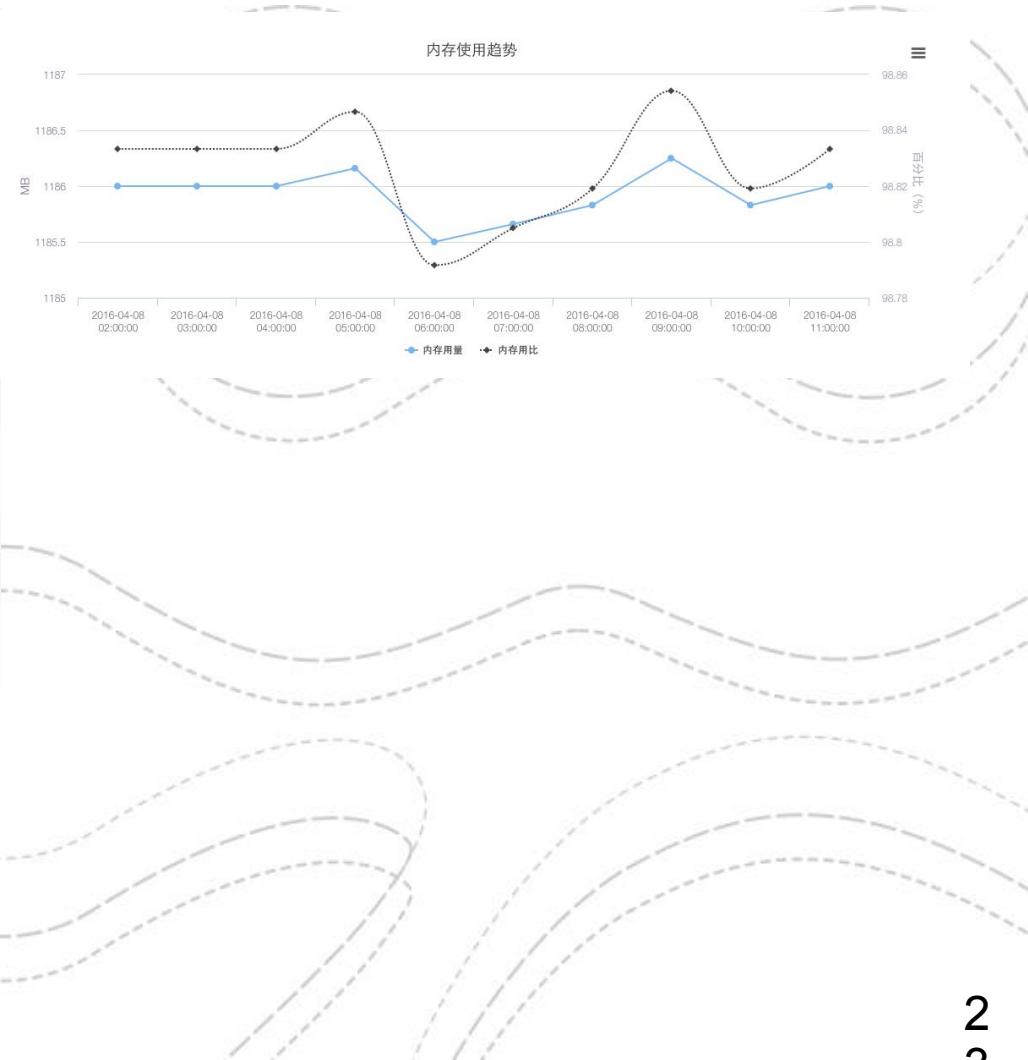
步长为5GB

内存 :

240MB 2400MB 240MB 600MB

用户画像

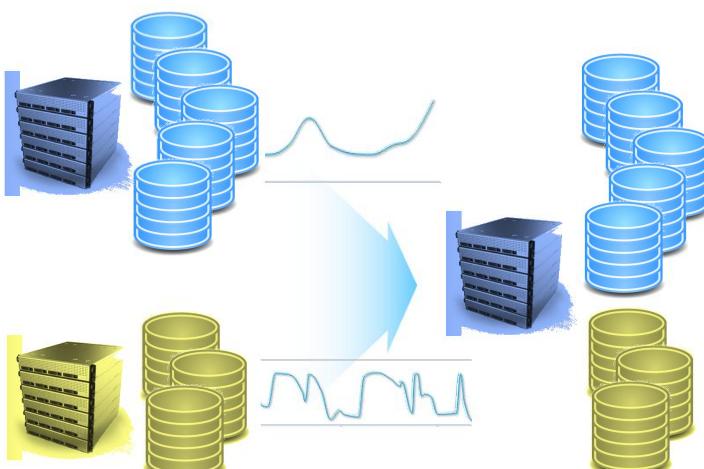
实例名	当前状况	建议升级项	升级评分
[REDACTED]	在过去30天里 ...	磁盘	21474836.47
[REDACTED]	在过去10天里 ...	内存	230540.41
[REDACTED]	在过去10天里 内存消耗日均使用超过1.8达到9天	内存	7204.4
[REDACTED]	在过去10天里 ...	内存	4878.3
[REDACTED]	在过去30天里 ...	磁盘	4358.68
[REDACTED]	在过去10天里 ...	内存	1632.64
[REDACTED]	在过去10天里 ...	内存	1492.23



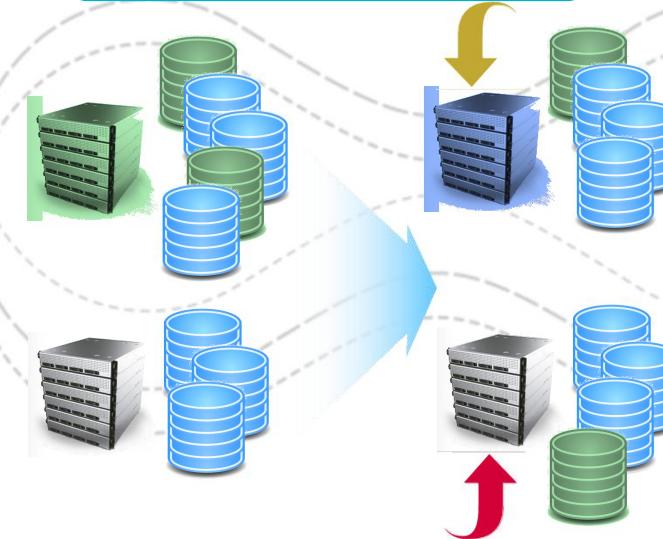
- 利用性能历史数据建立基线，评估CPU、内存、网络、磁盘等资源的使用量
- 主动预测用户未来使用情况，在出现瓶颈前，主动运维



成本压缩

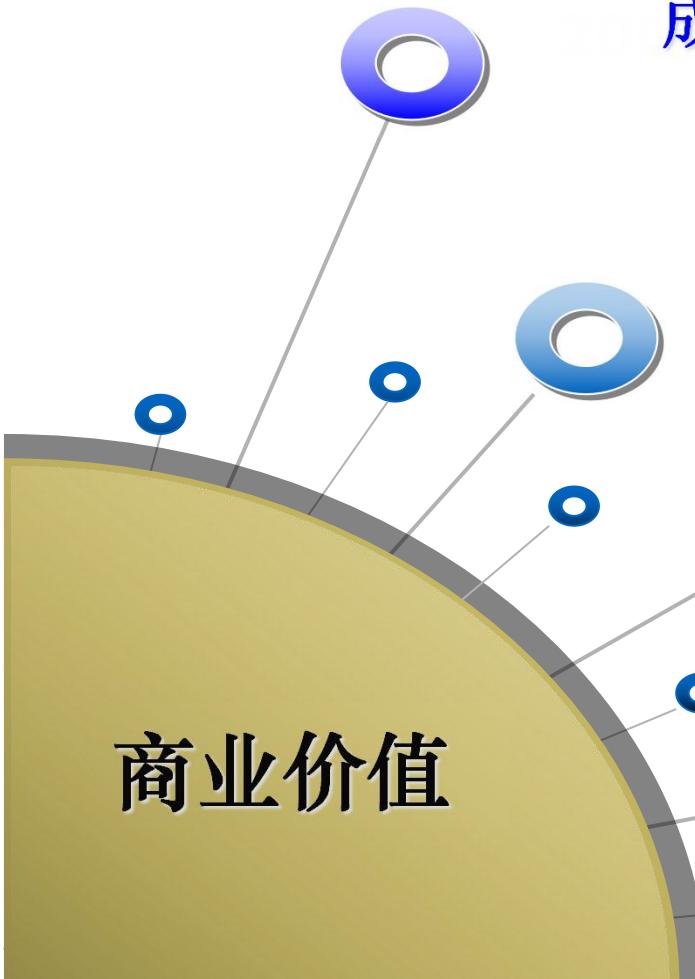


资源打散



- 平日阶段，进行成本压缩，最大化降低硬件成本，以及运维成本
- 重大活动期间，均衡负载，以应对高峰期承载





成本压缩

- ◆ 成本压缩
通过调度算法合理分配资源，提升主机利用率，节约成本
- ◆ 资源打散
均衡资源，以同样的基础设施，应对大促活动时的线上高峰

用户服务体验保障

- ◆ 用户打散
以用户为纬度，均衡打散其拥有实例在集群中的分布，最小化故障发生时，用户受到的影响

多种云产品混合部署

- ◆ 根据各个云产品业务的使用特性，混合部署，充分利用物理资源

自动弹性扩缩容

- ◆ 根据用户的使用行为及趋势，实现云产品的自动扩、缩容，以提供高弹性的增值云服务

精细化服务

SLA量化服务质量

- 精细化的服务质量评价
- 延迟波动、网络丢包等因素被全盘考虑
- 不再是传统的单一评价“是否可连接”
- 将用户的使用感受量化成指标

用户护航

- 整个云服务的视图，而不是单个实例是否有问题
- 帮助用户自己建立天象，完成系统数据化、智能化
- 服务化



Thanks!

ApsaraDB欢迎您