



DAMS

中国数据资产管理峰会

CHINA DATA ASSET MANAGEMENT SUMMIT

浙江移动 云运维体系建设与实践

演讲人：王晓征


王晓征

- ❑ Oracle 9I OCM (2003年)
- ❑ 中国移动通信集团浙江有限公司
信息技术部副总经理兼云计算中心主任
- ❑ 中国移动集团业务支撑高级技术专家
- ❑ 高级工程师
- ❑ 1997年中国足球乙级联赛注册球员
- ❑ 新浪微博：酒剑仙007
- ❑ 邮箱：wxz@zj.chinamobile.com





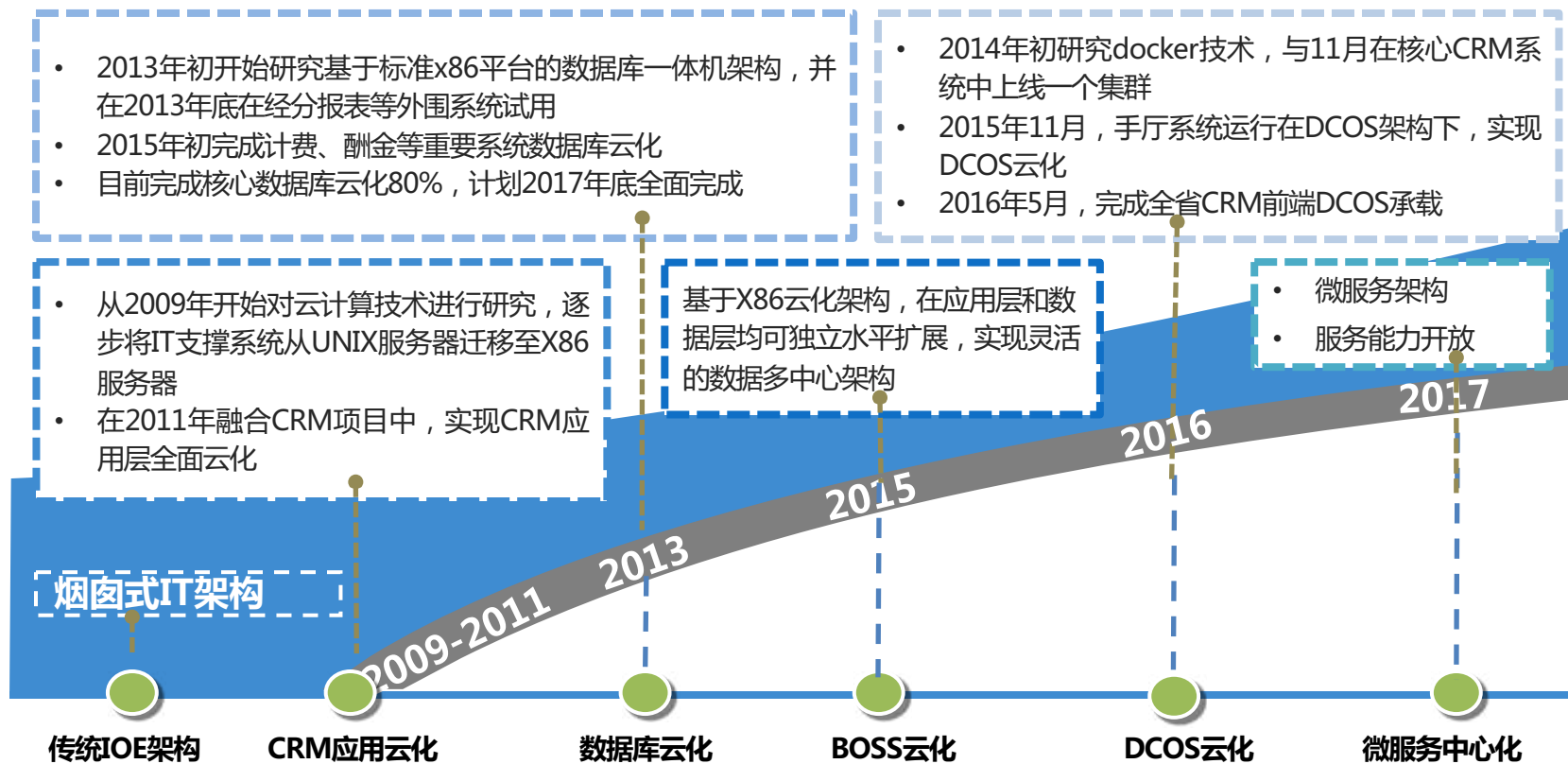
目录

- 
- 一 现状与挑战
 - 二 运维体系转型建设
 - 运维之道 道以明向
 - 运维之法 法以立策
 - 运维之器 器以成事
 - 三 展望未来



现状与挑战

浙江移动云计算建设历程



云计算中心的新环境

随着IT架构朝着虚拟化、云化演进，运维环境带来了新的变化，呈现出网元海量化、组件多样化、调度动态化等特点，也给我们传统运维人员带来新的挑战。

云数据中心



集中资源池的共享

虚拟化、分时共享

动态调配、弹性伸缩

自动化、自服务

低成本、标准化硬件

如何把控动态调配、弹性伸缩带来的管理成本？

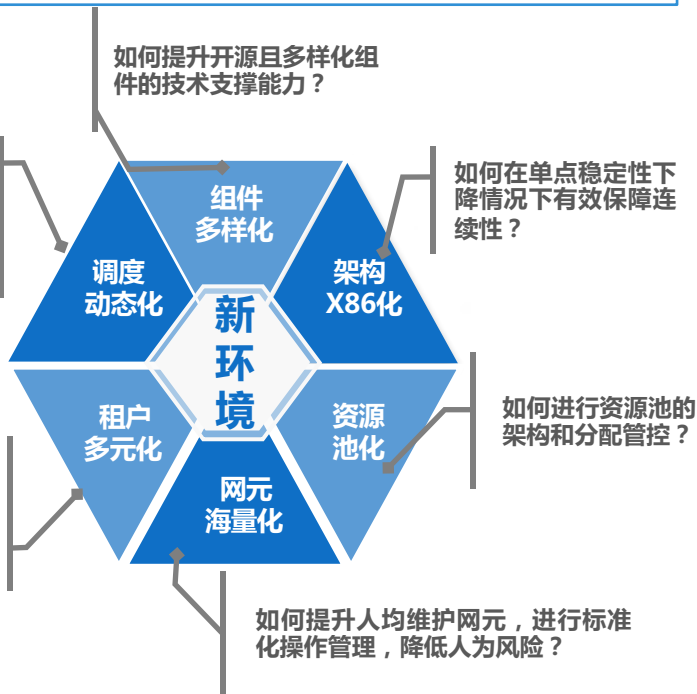
如何提升开源且多样化组件的技术支撑能力？

如何在单点稳定性下降情况下有效保障连续性？

如何进行资源池的架构和分配管控？

如何提升人均维护网元，进行标准化操作管理，降低人为风险？

如何有效保障租户的业务感知？

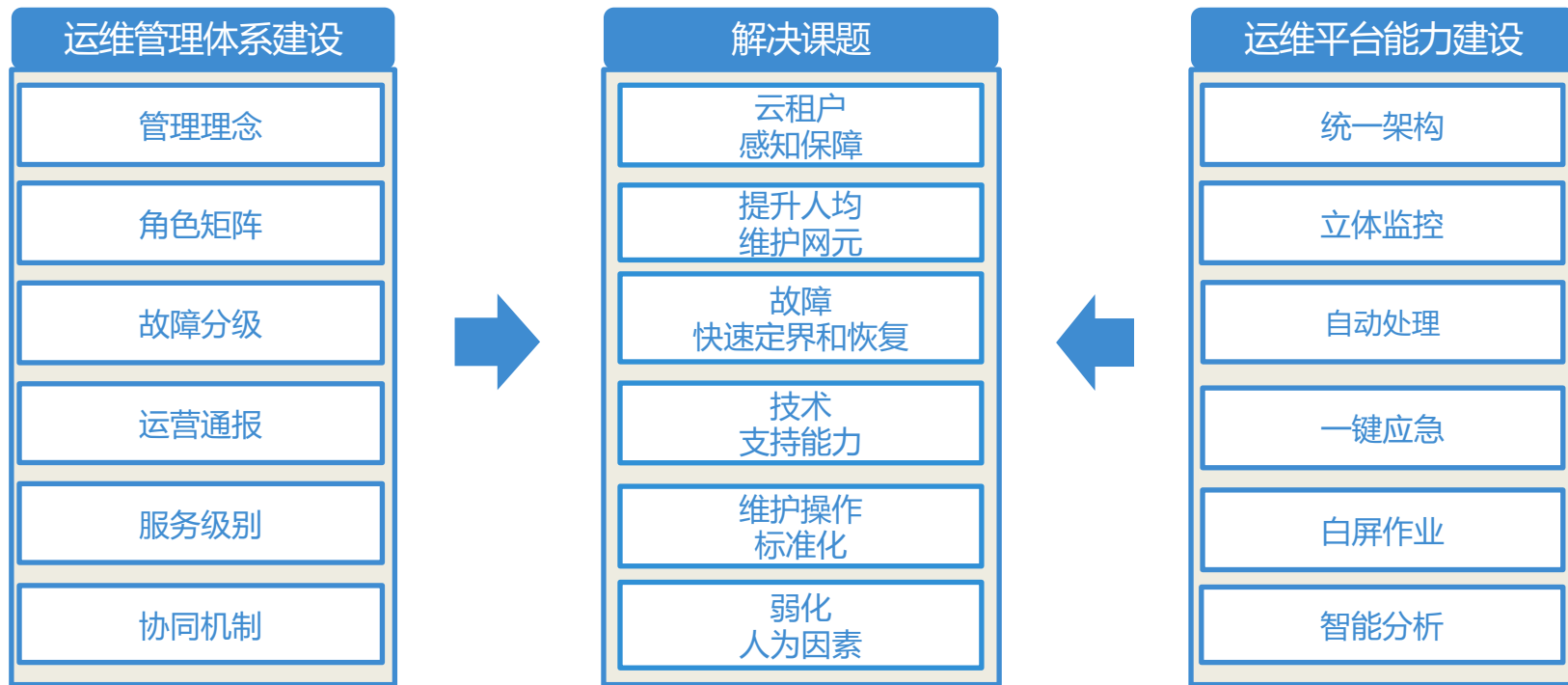


私有云运维的特定挑战



我们的应对

浙江移动云计算中心深入分析思考私有云运维下的新环境和新挑战，采取运维管理体系建设和运维平台能力建设的方式来应对，两手抓，两手都要硬。具体措施如下：





二

运维体系转型建设

云运维体系的“道、法、器”



运维之道

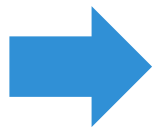
- 转变运维理念
- 连续性保障分层
- 运维自动化节奏

运维之法

- 组织架构 运维角色 调度机制
- 故障级别 通报体系 变更管理

运维之器

- 立体监控
- 自动处理



1

运维之道 道以明向

2

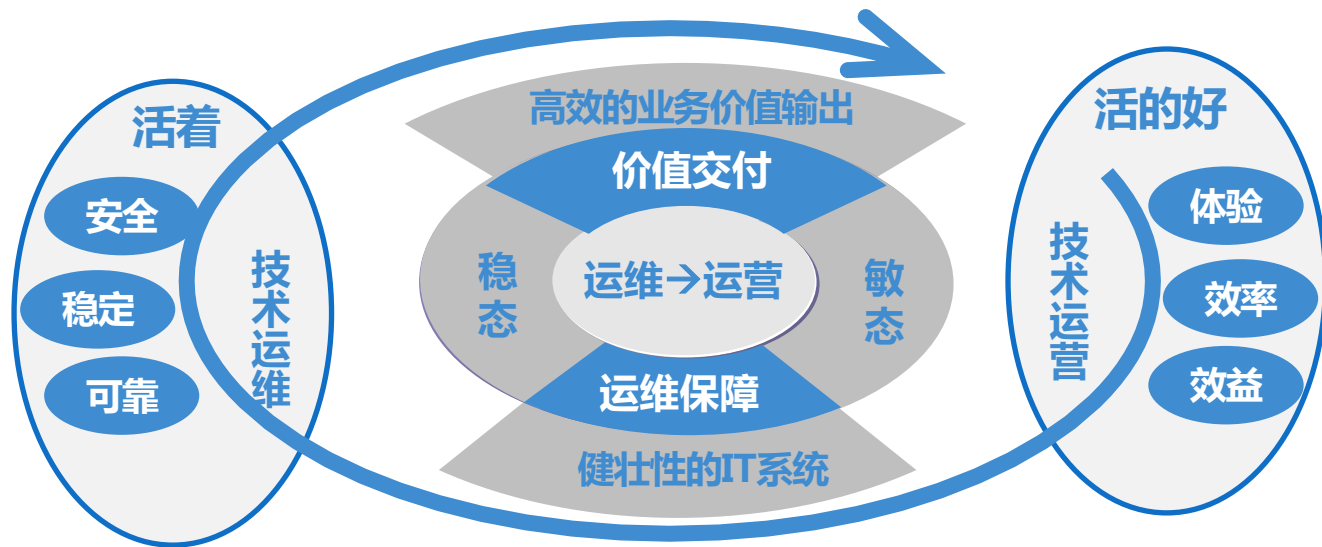
运维之法 法以立策

3

运维之器 器以成事

理念：从技术运维到技术运营，转变运维管理理念

- IT运维的宗旨是**保持系统高效的业务价值输出**。
- 云运维时代，IT运维除了关注系统安全稳定可靠以外，更加关注用户的体验、业务效率、企业效益。，运维管理应该从传统的**面向系统的技术运维模式向面向业务的技术运营模式**的理念上转变。



双态驱动 组织转型 流程优化 智慧运维

分层：三层连续性保障模型

浙江移动云计算中心引入SRE云运维的最佳实践，从运维的角度分解为三层保障，从上而下分为：业务连续性、应用连续性、平台连续性。

利用平台化的自动运维能力分面向平台、面向应用、面向业务的三层进行建设规划，从运维效率的角度解决对内对外的运维满意度问题。

面向业务

- 可视化业务流程
- 关联映射业务指标

业务连续性

充值业务

开户业务

停复机业务

系统登录

客户资料查询

实时费用查询

面向应用

- 以应用为单位整合资源管理
- 自动调度
- 统一监控

应用连续性

经营分析系统

管理信息系统

电子渠道系统

网管支撑系统

业务支撑系统

客户服务系统

面向平台

- 自动部署
- 弹性伸缩
- 资源管理

平台连续性

服务总线 (OSB/ESB/VSB)

流程引擎

缓存服务

消息服务

容器

操作系统

DCOS

数据库

大数据平台

虚拟机

服务器

存储

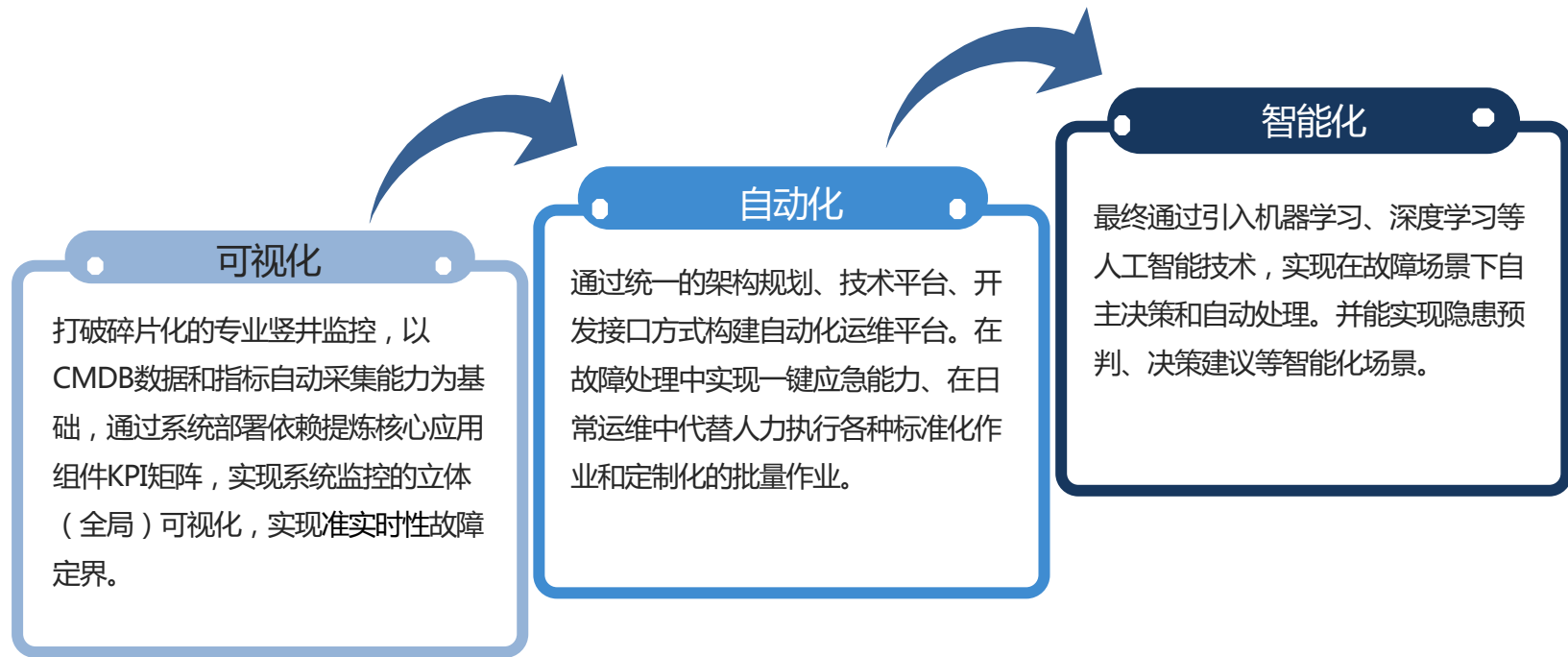
交换机

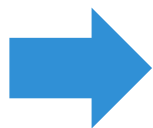
路由器

平台和应用都连续的情况下，超过99%的场景业务是连续的。云计算中心在连续性保障中重点关注平台，保障应用。

节奏：运维平台能力提升的节奏

自动化工具建设通过**可视化能力**解决故障快速定界，通过**自动化能力**提高效率，通过**智能化能力**提供智能分析/决策建议。根据互联网的发展经验，齐头并进，重点突破，现阶段优先实现可视化、再进行自动化、最终实现智能化。





1

运维之道 道以明向

2

运维之法 法以立策

3

运维之器 器以成事

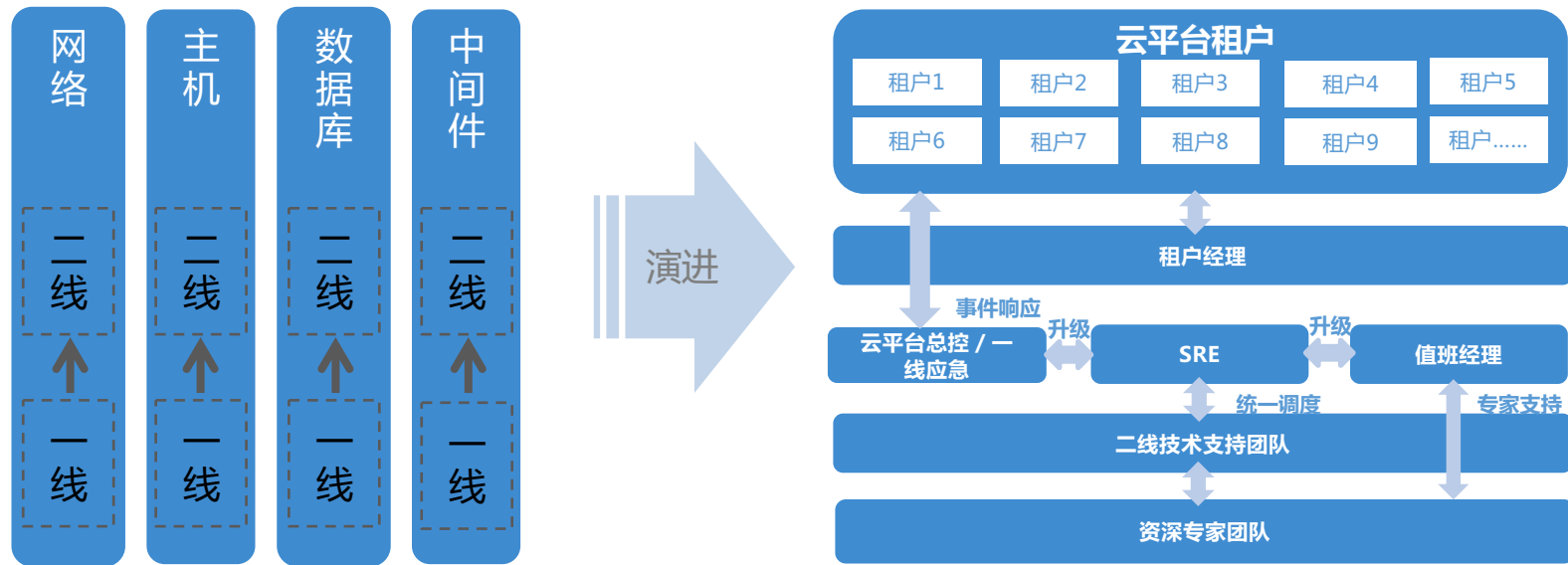
三层模型指导转型建设

通过将IT运维抽象出三层模型指导我们的运维管理体系转型建设。自下而上分为运维对象、运维活动、运维角色。



组织架构转型

“小前台，大中台”这种组织形式，已不再是新鲜事物，但仍不失为一种理想化的支撑模式。众多的互联网公司都在进行组织结构的优化调整，阿里集团在去年年底宣布组织结构全面升级。其实，所有人的最终目的只有一个：**简化流程冗余，降低沟通成本，有效整合资源。**



运维角色：云计算中心总控

浙江移动云计算中心总控是云计算中心运维的触点，技术运营的接口。依托于运维自动化平台在可视化告警和自动化处理的基础上实现全栈的基础运维集中化、运营服务标准化的一线团队。



运维角色：SRE

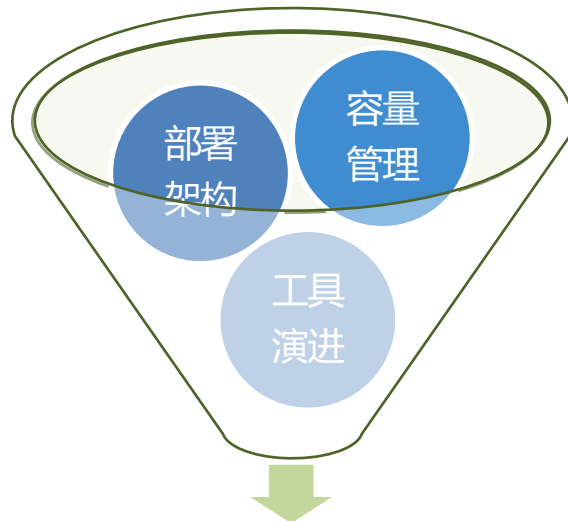
SRE是指Site Reliability Engineer(网站可靠性工程师)，起源于Google，具备较强的编程能力和系统全栈能力的维护工程师。云计算中心在总结出云运维的三层保障体系的基础上，并借鉴国内外优秀的互联网运维经验，我们对SRE进行了重新解读和定位：**SRE是以连续性保障为最终使命，对部署架构、容量管理、运维工具演进负责的全栈型运维人员**，分为两种角色：**平台SRE、应用SRE**。

平台SRE

运维思路：**以云平台
角度向应用延伸**

职权：

- 平台应急与决策
- 平台部署架构管理
- 平台高可用管理
- 平台容量管理
- 平台性能优化
- 平台工具演进



连续性保障为最终使命

应用SRE

运维思路：**以应用架
构角度向业务延伸**

职权：

- 应用应急与决策
- 应用上线发布管理
- 应用高可用管理**
- 应用容量管理
- 应用性能优化
- 应用工具演进

运维角色：租户经理

租户经理是云租户和云计算中心之间更为通畅的沟通渠道，可以更快调度更为充分的资源响应客户需求，建立良好的客户关系。



调度机制：快速响应、扁平调度

浙江移动云计算中心总控统一接口事件管理快速响应，SRE则完成扁平调度，拥有应急重启、高可用接管、容灾切换等常规应急处理的直接决策权。故障处理上遵循**先恢复，后修复**的原则。

原则：先恢复，后修复

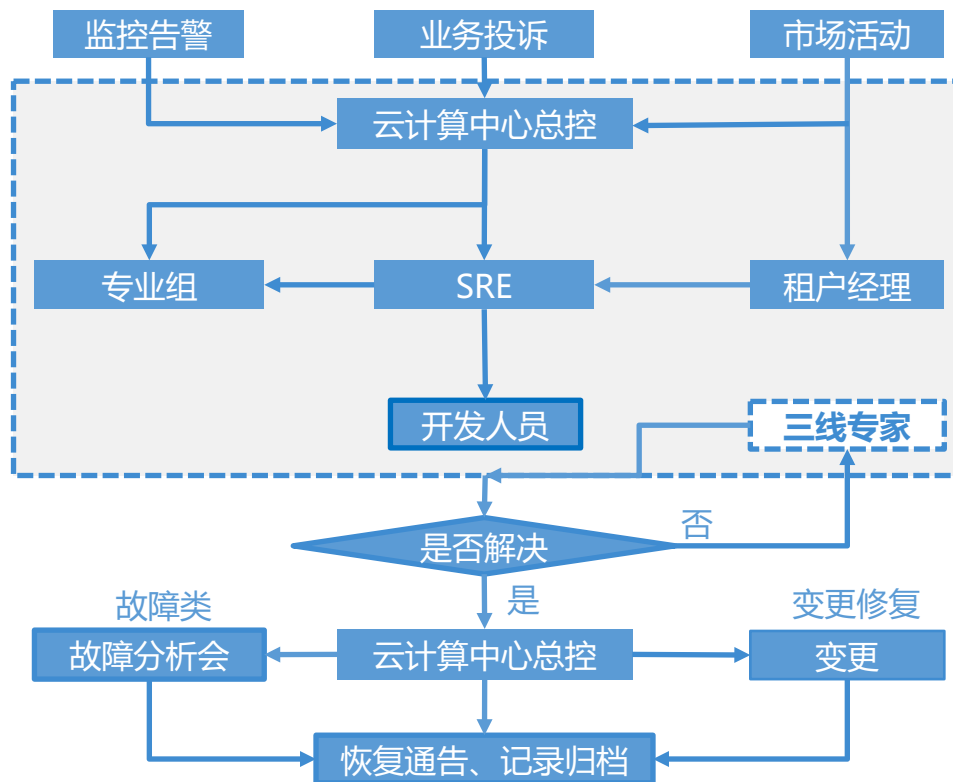
云总控作为事件统一入口，根据自动化工具和标准化手册进行处理。

业务投诉：云总控>SRE>专业组

监控告警（无业务影响）：云总控>专业组

恢复之后：

故障类需要经过故障分析会进行总结后归档；
需要后续修复的则经过变更修复后归档。



故障分级：明确标准，张弛有度

故障分级理念

- 坚持云平台为根、云服务为本的理念
- 从云平台服务可用性、云租户业务感知、影响租户数量三个方面综合考虑定制故障标准。
- 从业务的全中断和高可用性出发，制定故障的等级。

故障五大分级

- 五级事件【灰-关注事件】
- 四级事件【蓝-一般故障】
- 三级事件【黄-主要故障】
- 二级事件【橙-严重故障】
- 一级事件【红-重大故障】

云服务业务全中断时间	一般系统		重要系统		核心系统
	<=3 系统	>=4 系统	<=3 系统	>=4 系统	>=1系统
<=	五级事件	五级事件	五级事件	五级事件	五级事件
> <=	五级事件	五级事件	五级事件	四级事件	四级事件
> <=	五级事件	五级事件	四级事件	三级事件	三级事件
> <=	五级事件	四级事件	三级事件	二级事件	二级事件
> <=	四级事件	四级事件	二级事件	二级事件	二级事件
> <=	四级事件	三级事件	二级事件	一级事件	一级事件
> <=	二级事件	二级事件	一级事件	一级事件	一级事件

高可用中断时间	异地灾备		本地高可用	
	<=2 服务组件	>=3 服务组件	<=2 服务组件	>=3 服务组件
<	五级事件	五级事件	五级事件	五级事件
> <=	五级事件	五级事件	五级事件	四级事件
> <=	五级事件	四级事件	四级事件	三级事件
> <=	四级事件	四级事件	四级事件	三级事件
> <=	三级事件	三级事件	三级事件	三级事件

通报体系：自下而上、逐层升级

为了提升租户满意度，并对租户负责的态度，云计算中心对四级及以上事件会向租户提供故障分析报告。云计算中心内部会上故障分析会，进行分析总结归档。

逐层升级 ↑	故障信息通报	通报对象	通报方式	备注
	一级事件【红-重大故障】	公司和租户高层领导	短信	通报对象和通报方式向下兼容。
	二级事件【橙-严重故障】		短信	
	三级事件【黄-主要故障】		短信	
	四级事件【蓝-一般故障】		短信	
	五级事件【灰-关注事件】	租户工作群	微信	

变更管理：统一收口、降低风险

根据云计算中心的运维经验，70%+的故障都是由变更引起的，云计算中心租户众多，不同租户应用运维体系建设水平参差不齐，云计算中心建设统一变更流程纳管云平台所有变更尤为重要。



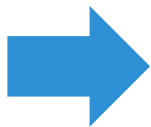


1

运维之道 道以明向

2

运维之法 法以立策



3

运维之器 器以成事

智慧运维平台立体化监控的必要性

系统的云化部署之后，应用和平台的关系呈现动态调整的趋势，且在多网元、多组件的情况下，通过传统的竖井化工具发现单点问题来故障定界捉襟见肘，通过优化管理体系，设立SRE来加强全局分析能力，但是人的因素也成了瓶颈，故障快速定界成了运营商行业面临的难题。



业务和应用依赖关系复杂
应用和平台动态调整
平台组件之间的依赖
多网元 多组件

1. 传统工具和告警：

传统工具与告警虽全面，但在复杂云化架构下的全局定界能力捉襟见肘，且阈值式设定造成告警风暴，使运维人员困扰。

告警很多，但是不知道问题在哪里？！



2. SRE：

全栈型维护人员凭着过硬的知识与经验一定程度上增强了快速定界能力，但人为因素大、不稳定、成本高，最终成为瓶颈。

问题很怪异！这次我也没辙！



3. 立体化监控：

梳理应用层与平台层拓扑，形成依赖关系网，并对所有网元和组件进行健康度模型评判，联动下钻分析定位，可视化呈现。

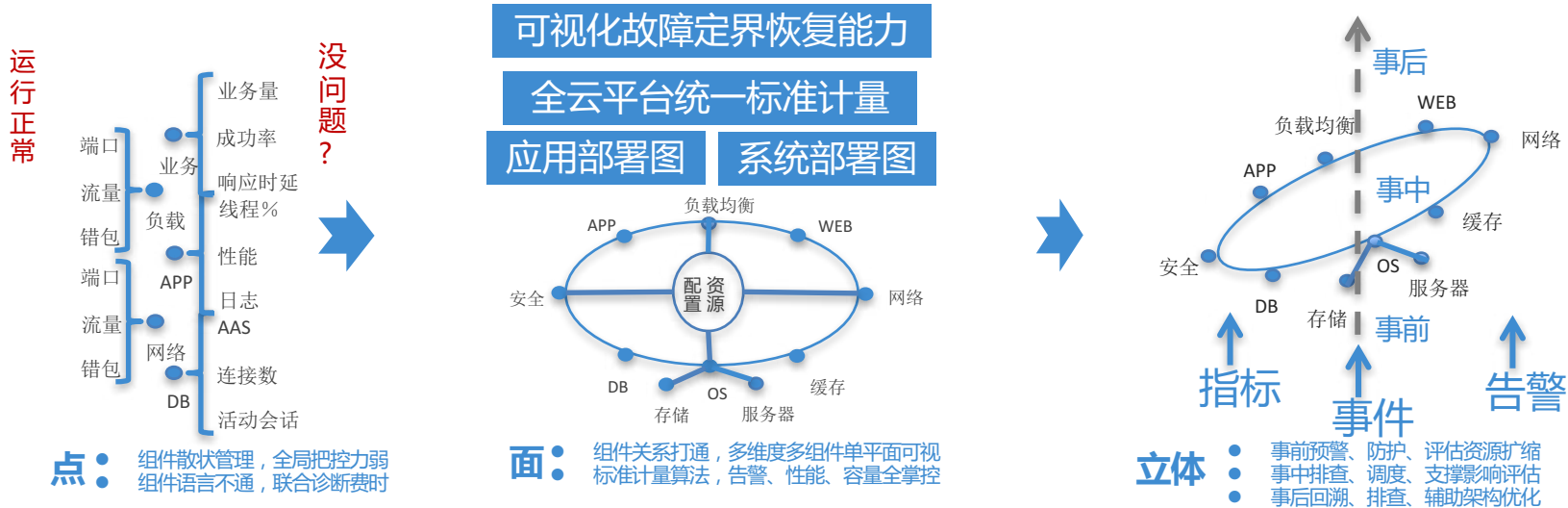
看！问题就在这里！



故障原因

智慧运维平台建设思路

在基础采集能力、配置能力支撑下，布局云平台场景化运维建设，逐步实现可视化、自动化、智慧化的云平台场景化运维能力，大大降低了运维门槛，实现租户自助式、无障碍故障定位与根因排查、恢复，同时实现运维人员价值转型。



点： 组件散状管理，全局把控力弱
组件语言不通，联合诊断费时

面： 组件关系打通，多维度多组件单平面可视
标准计量算法，告警、性能、容量全掌控

双模驱动

租户、管理员、大一线

- 数据开放，实时数据化展示平台运行全貌
- 秒级、自动故障根源定位能力
- 可视化评估故障影响面
- 统一标准计量算法，客观评判云平台状态

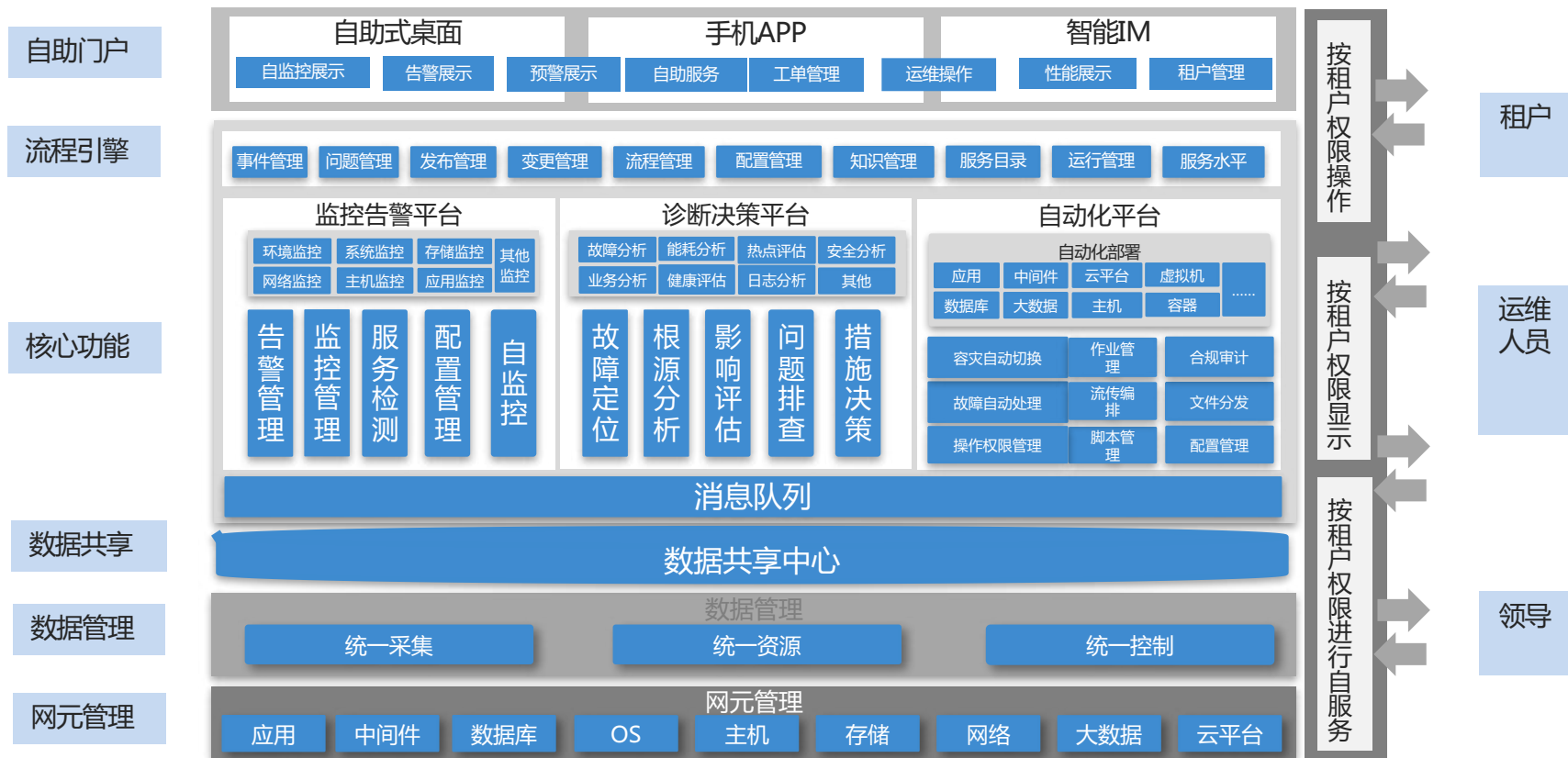
支撑租户
聚焦业务

维护二线、三线

- 实时问题发现、快速介入避免问题扩大
- 完善数据支撑，优化、完善平台架构
- 可视化故障全景，快速修复故障根源
- 开放平台支撑，逐步推进人工到自动落地

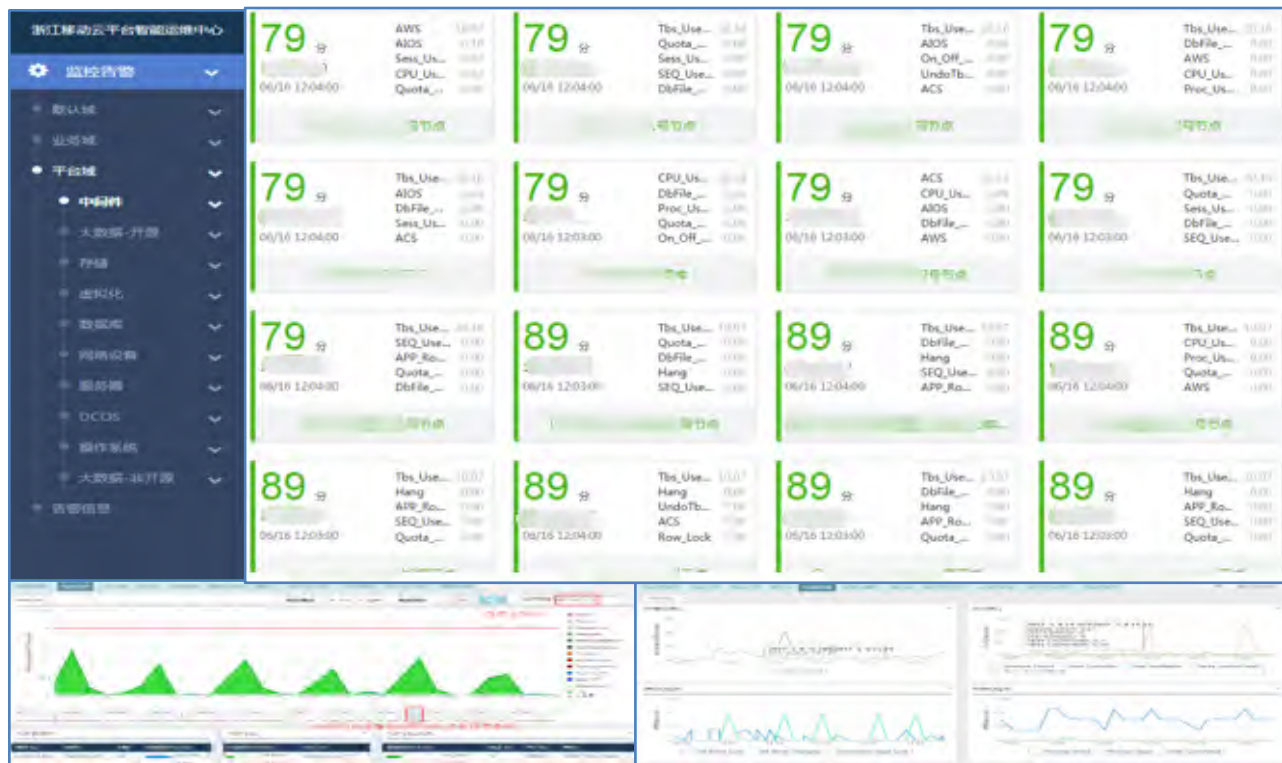
支撑运维
价值转型

智慧运维平台整体功能规划



已实现的可视化场景

云计算中心与合作伙伴联合研发的智能运维平台实现了业界首创的IT组件健康度模型，直观的展现出组件的健康状态，目前已经实现了数据库、操作系统、网络设备等多个组件的健康度评分。



已实现的自动化场景

智能运维平台网页和手机APP“太空堡垒”的自动化模块让云平台运维人员实现了白屏化的运维操作，不仅提高效率而且规避了人为风险，目前实现了容灾切换、X86主机一键重启、数据库堵塞会话自动查杀及数据库夯自动重启等自动化运维功能。



月均自动介入
100+次

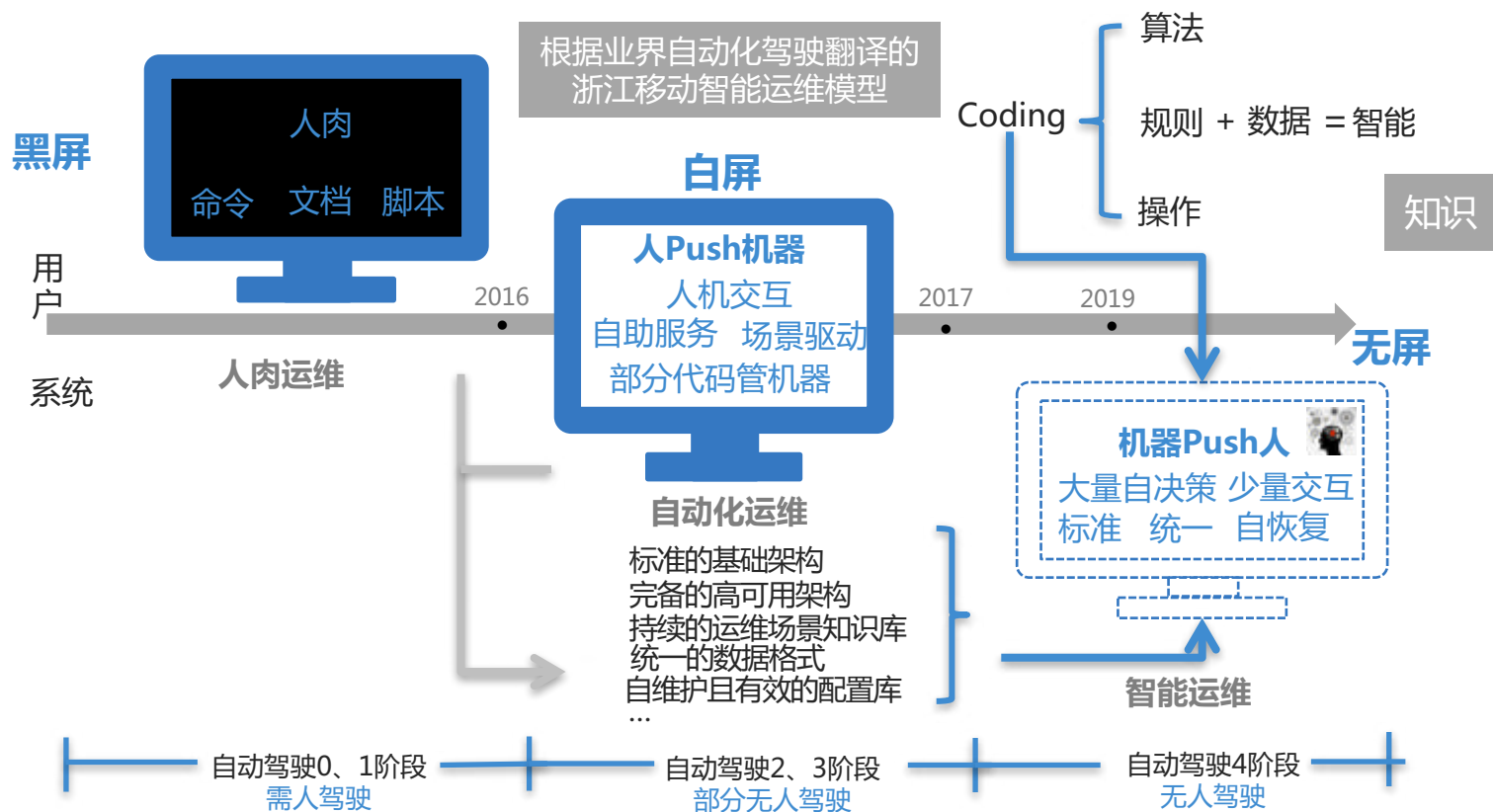
一键容灾切换
5分钟



三

展望未来

智慧运维三阶段愿景



欢迎关注三墩IT人





DAMS

中国数据资产管理峰会

CHINA DATA ASSET MANAGEMENT SUMMIT



THANK YOU!

