# 阿里云弹性计算稳定性建设和云上运维最佳实践

李刚 (东任)

阿里巴巴 高级技术专家



# 精彩继续! 更多一线大厂前沿技术案例

&广州站

QCon

全球软件开发大会

时间: 2023年5月26-27日 地点: 广州·粤海喜来登酒店

扫码查看大会详情>>



❷ 深圳站



时间: 2023年7月21-22日 地点: 深圳·博林天瑞喜来登酒店

扫码查看大会 详情>>



❷北京站



全球软件开发大会

时间: 2023年9月3-5日 地址: 北京·富力万丽酒店

扫码查看大会 详情>>



# 大纲

- 阿里云弹性计算概念
- 阿里云弹性计算稳定性建设实践
- 阿里云弹性计算云上运维最佳实践



# 什么是阿里云弹性计算?

- 又名云服务器ECS(Elastic Compute Service)
- 云计算最核心基础laaS服务之一
- 让大家像使用水、电、天然气等公共资源一样便捷、高效地使用服务器,实现计算资源的即开即用和弹性伸缩



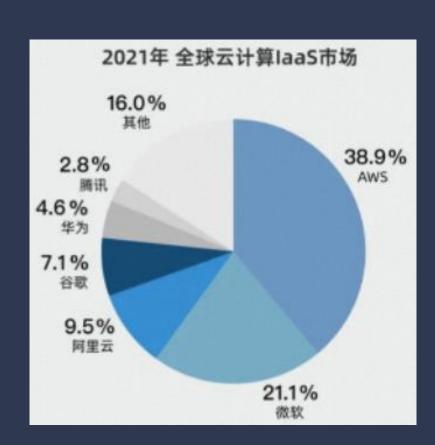


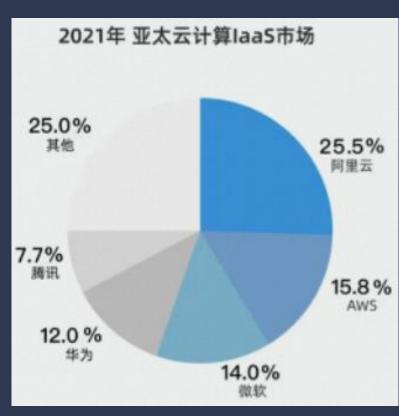
# 阿里云弹性计算稳定性建设实践



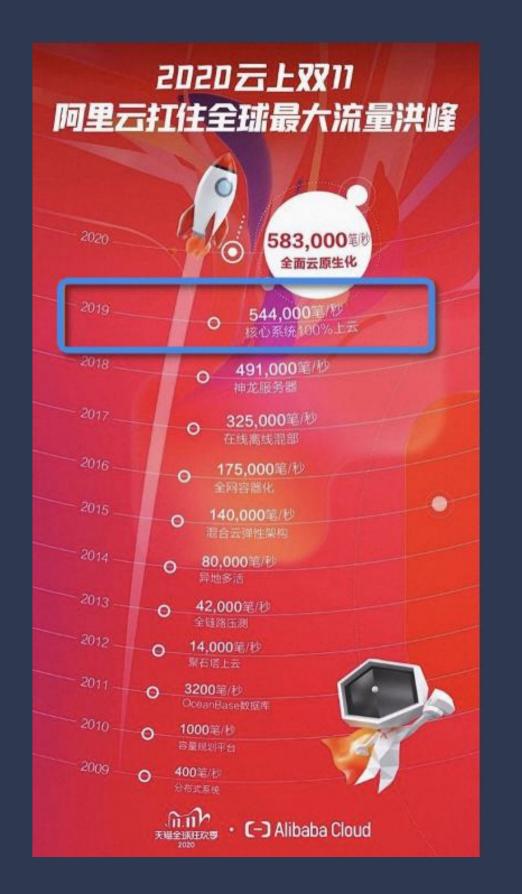
# 双轮驱动下的弹性计算极致稳定性诉求

### 百万级外部客户



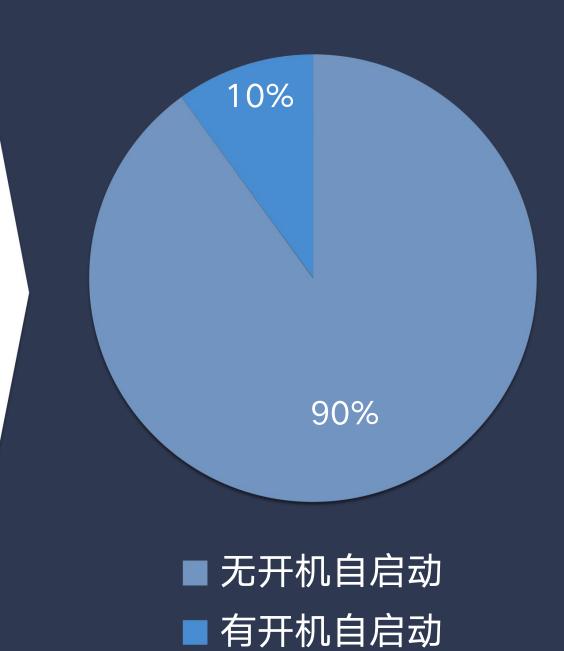


### 2019阿里全面上云



### 对弹性计算稳定性强依赖



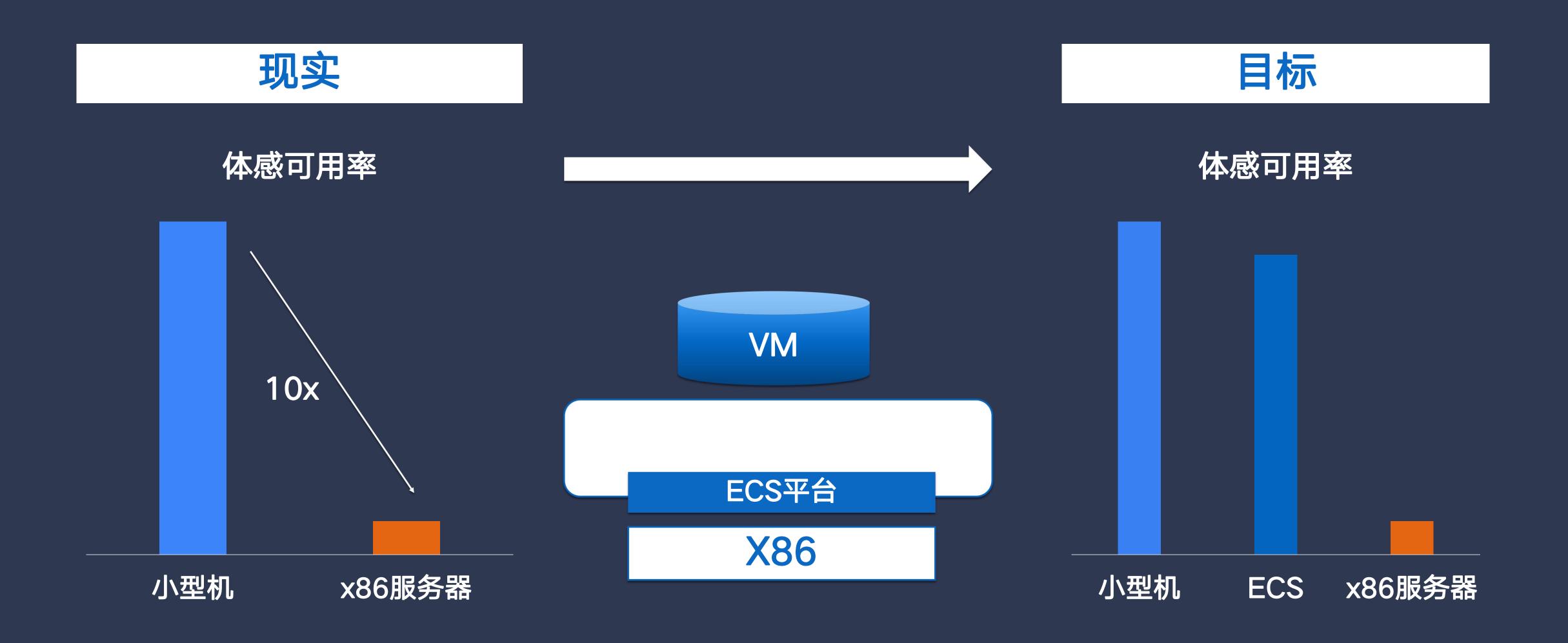


### 在线业务场景



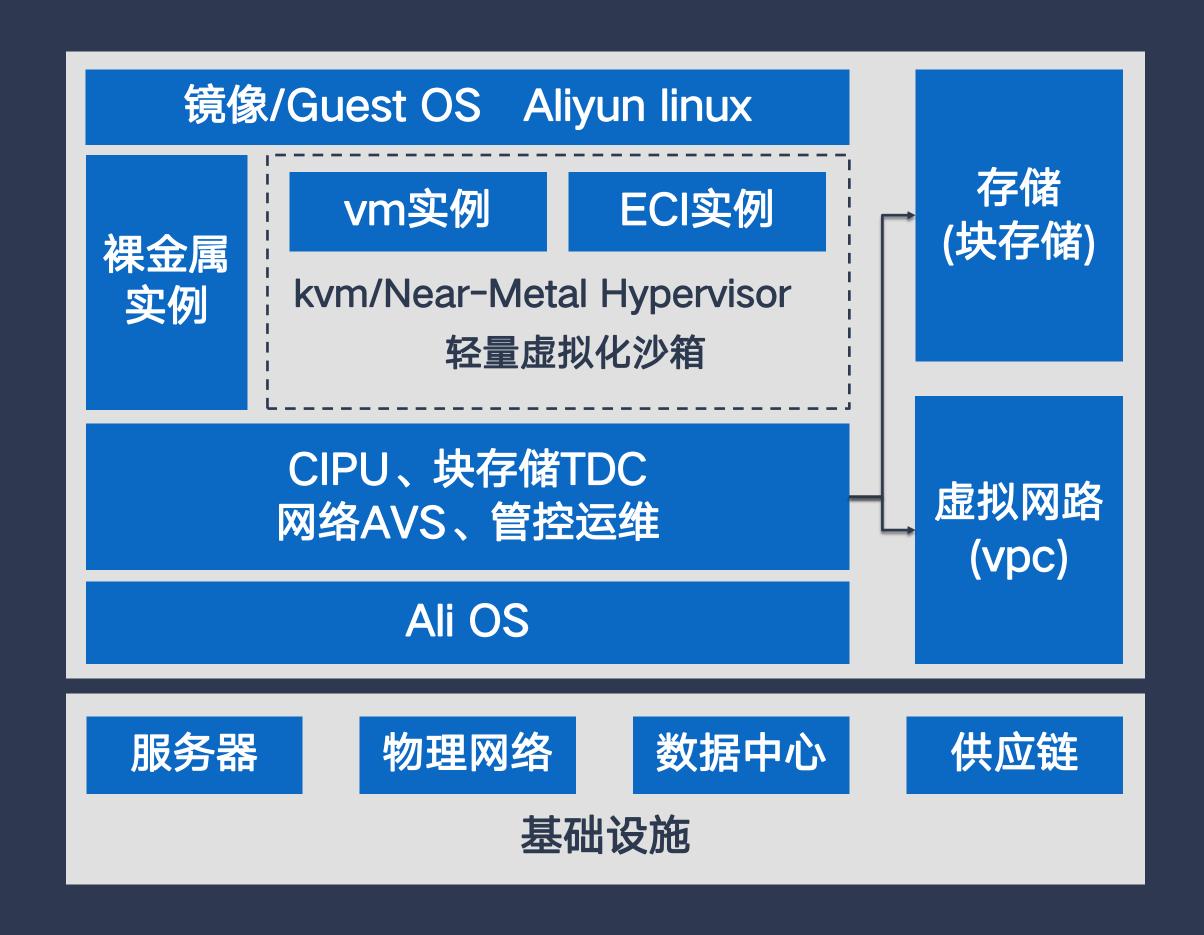


# 目标:用x86的硬件,提供小型机级别的稳定性





# 挑战: 复杂度和规模





**20+**数据中心区域(Region)

**80+** 可用区(AZ)

~5000+ 集群(Cluster) ~100000+

设备(Server)

2023年4月

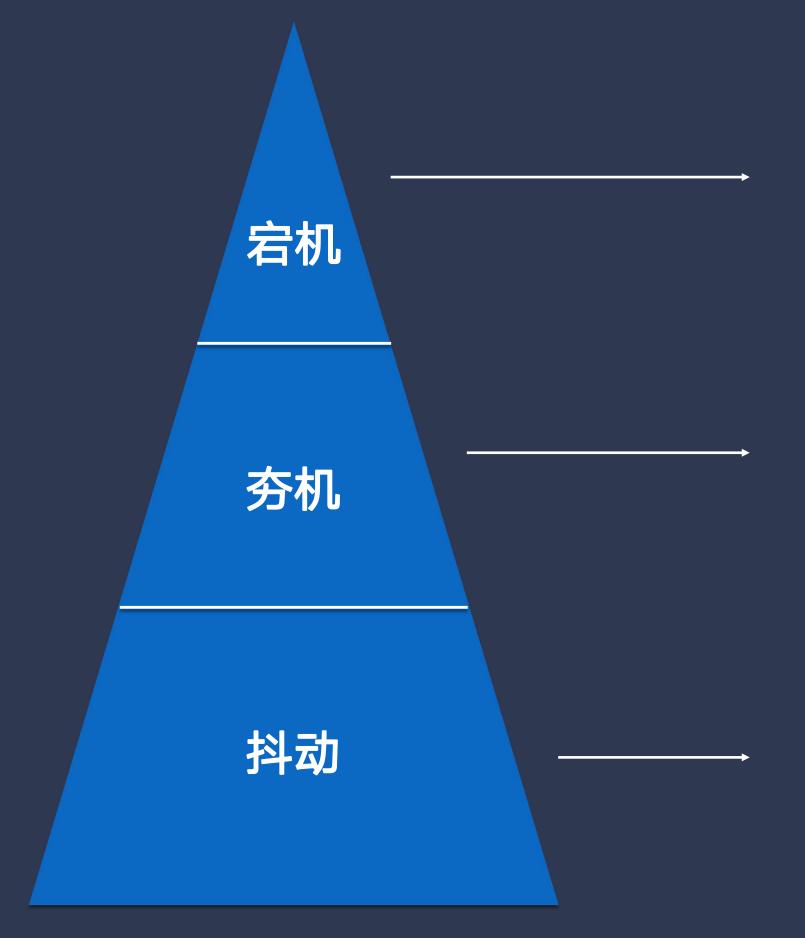
复杂度



规模



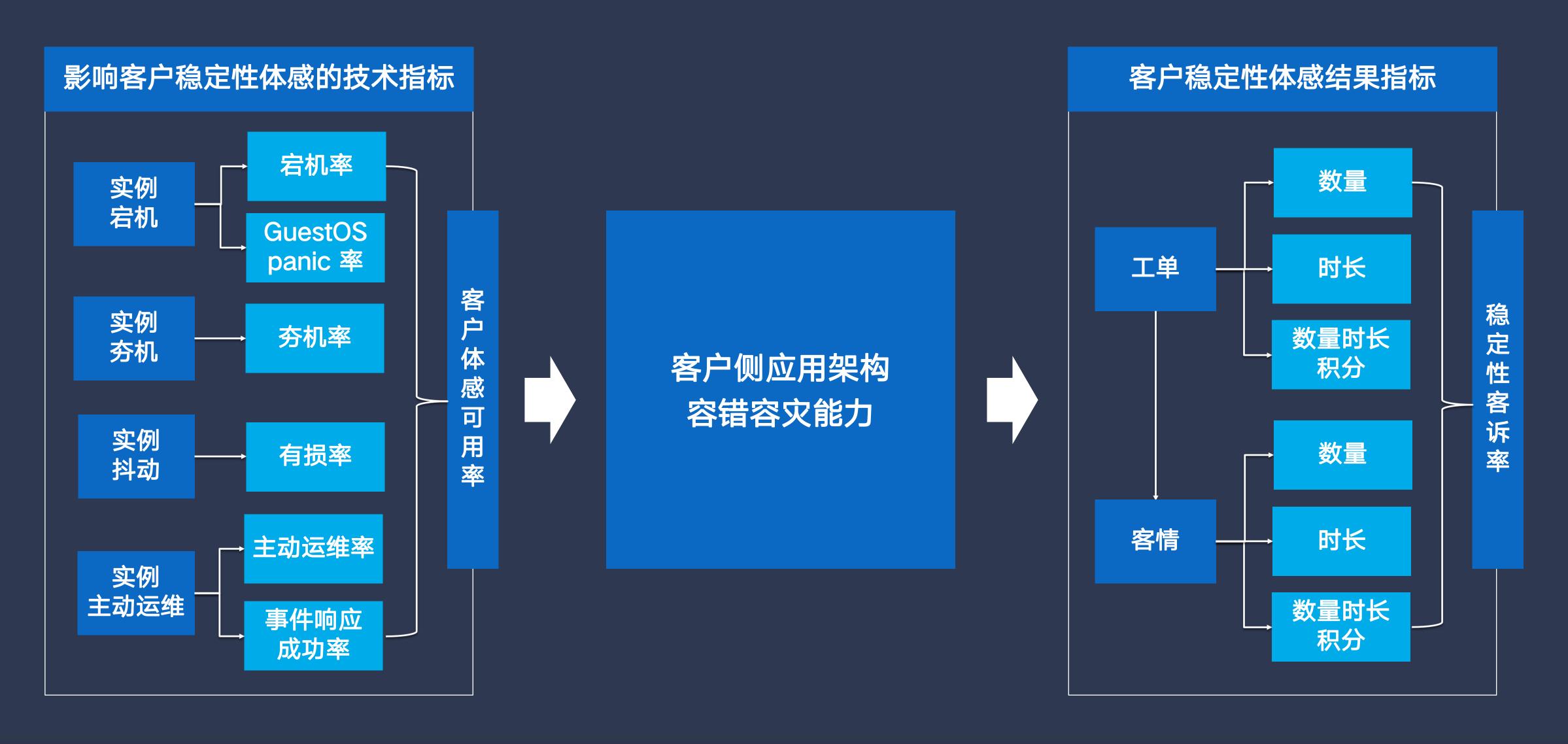
# 识别问题: 宕机、夯机、抖动



- **现象:** ECS资源100%不可用,多数因基础设施、服务器硬件或底层软件原因导致。
- **影响:** 所有未持久化的数据和配置都将丢失,该ECS实例上的业务将完全中断。
- **现象:** ECS资源服务时断时续,甚至某些核心功能不可用或无法连接和操作。如: OS 夯, IO hang等。
- **影响:** 未持久化数据尚未丢失,但整个ECS几乎无法使用,有时甚至无法恢复、没有备份的机会。
- **现象:** ECS资源核心服务可以正常使用,但在极端情况下会出现网络或性能抖动。
- **影响:**着重影响抖动敏感用户,性能抖动可能导致用户压测等容量规划付之东流,甚至可能因抖动引发用户应用系统雪崩效应,导致整体业务中断。

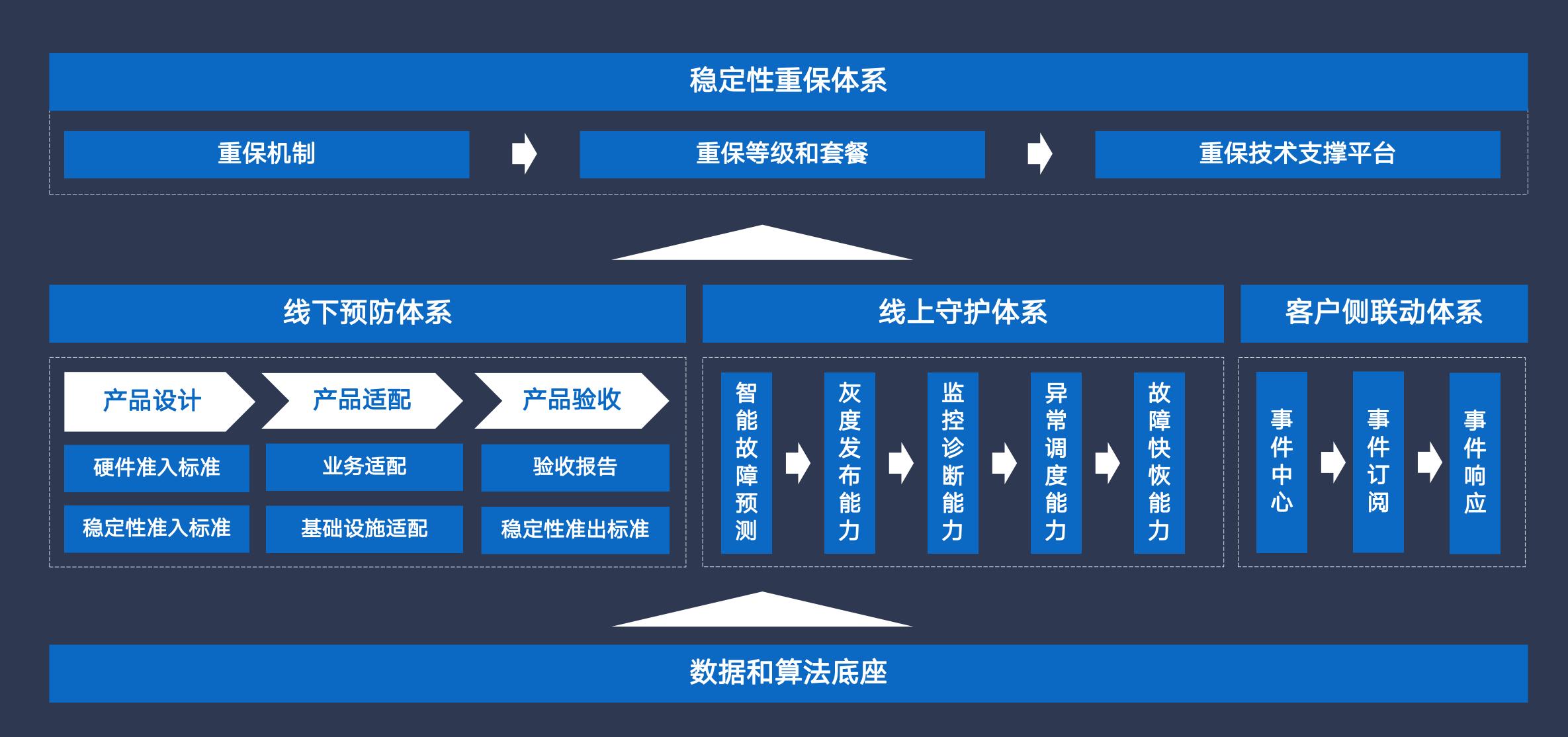


# 量化问题:从客户体感视角全面度量稳定性过程和结果





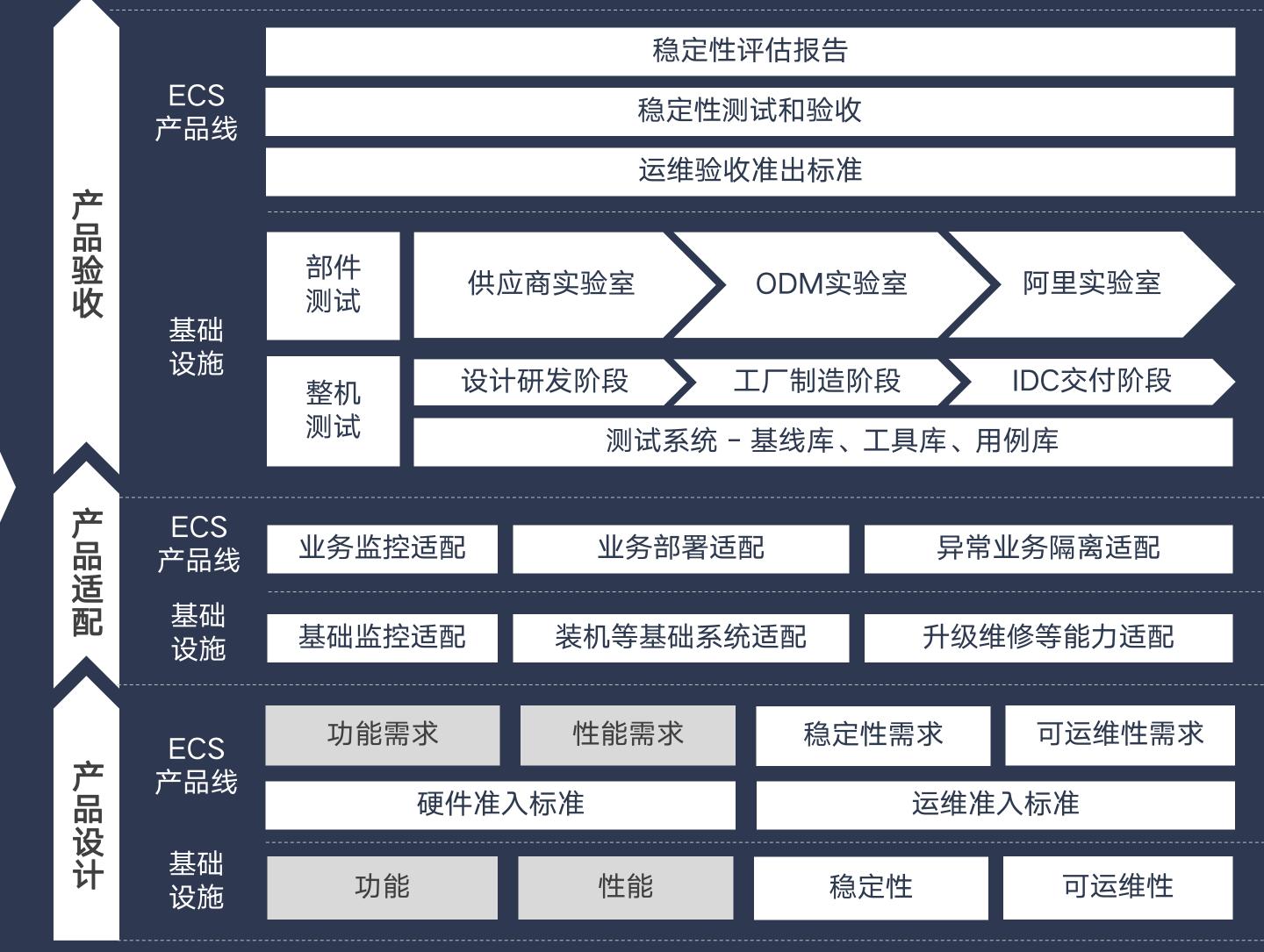
# 解决问题: 弹性计算稳定性系统工程





# 线下预防体系

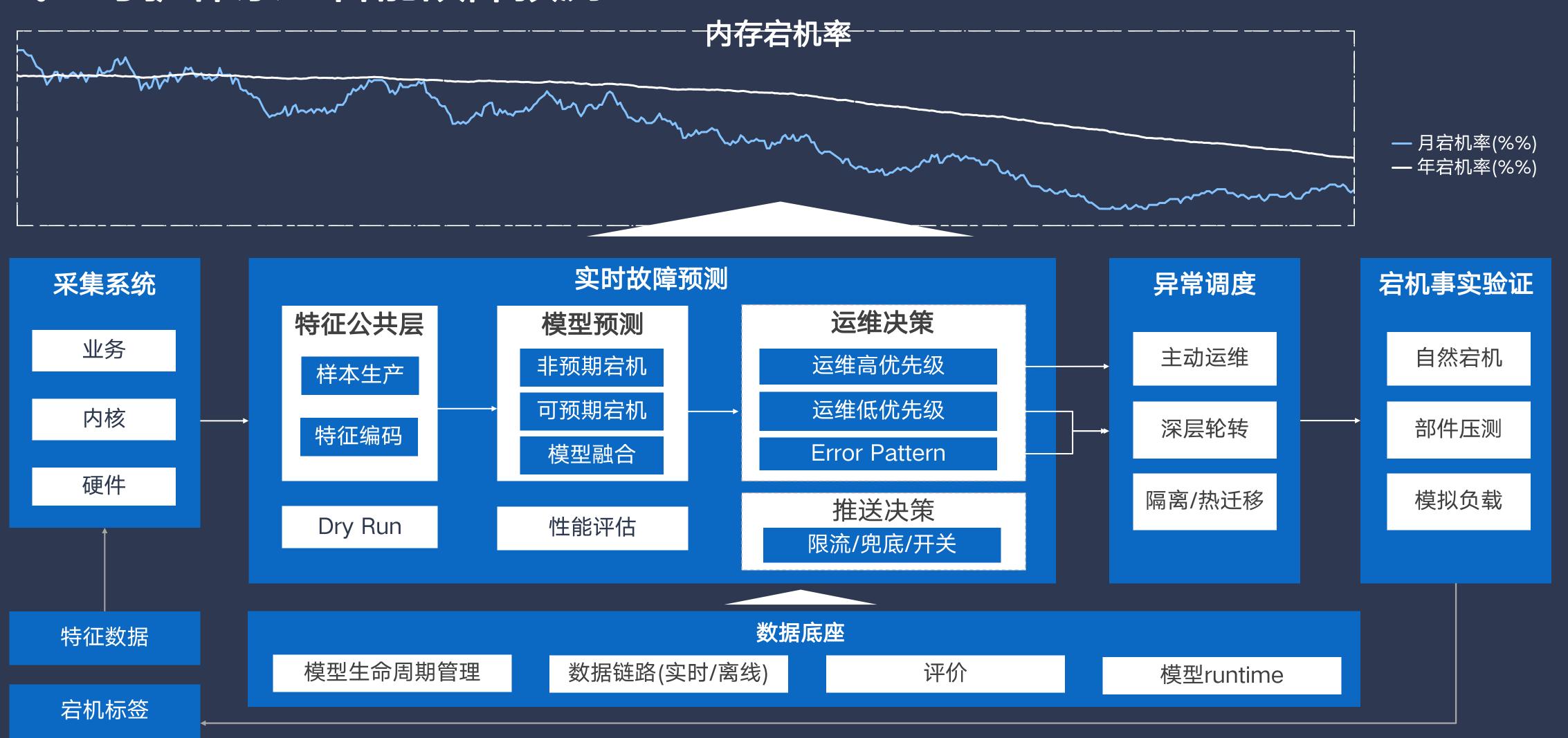








# 线上守护体系:智能故障预测



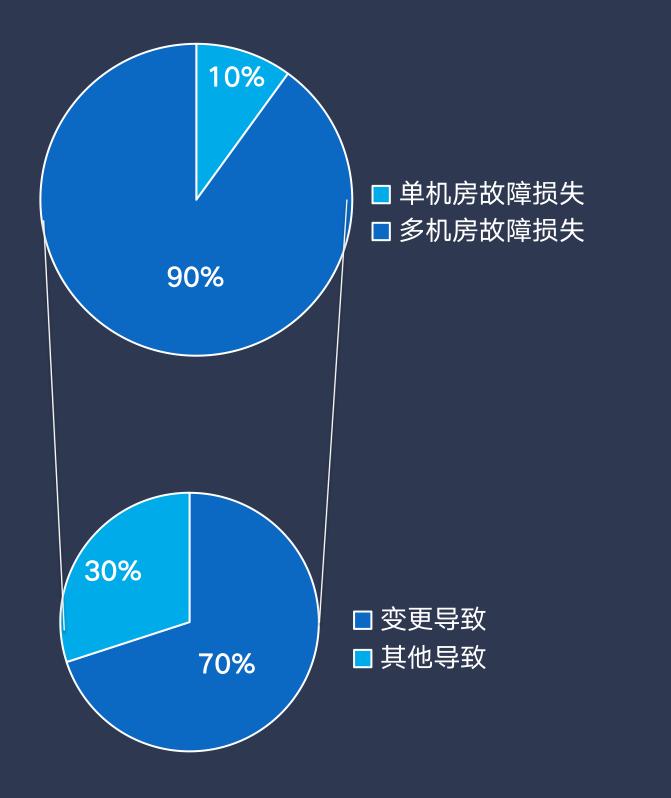




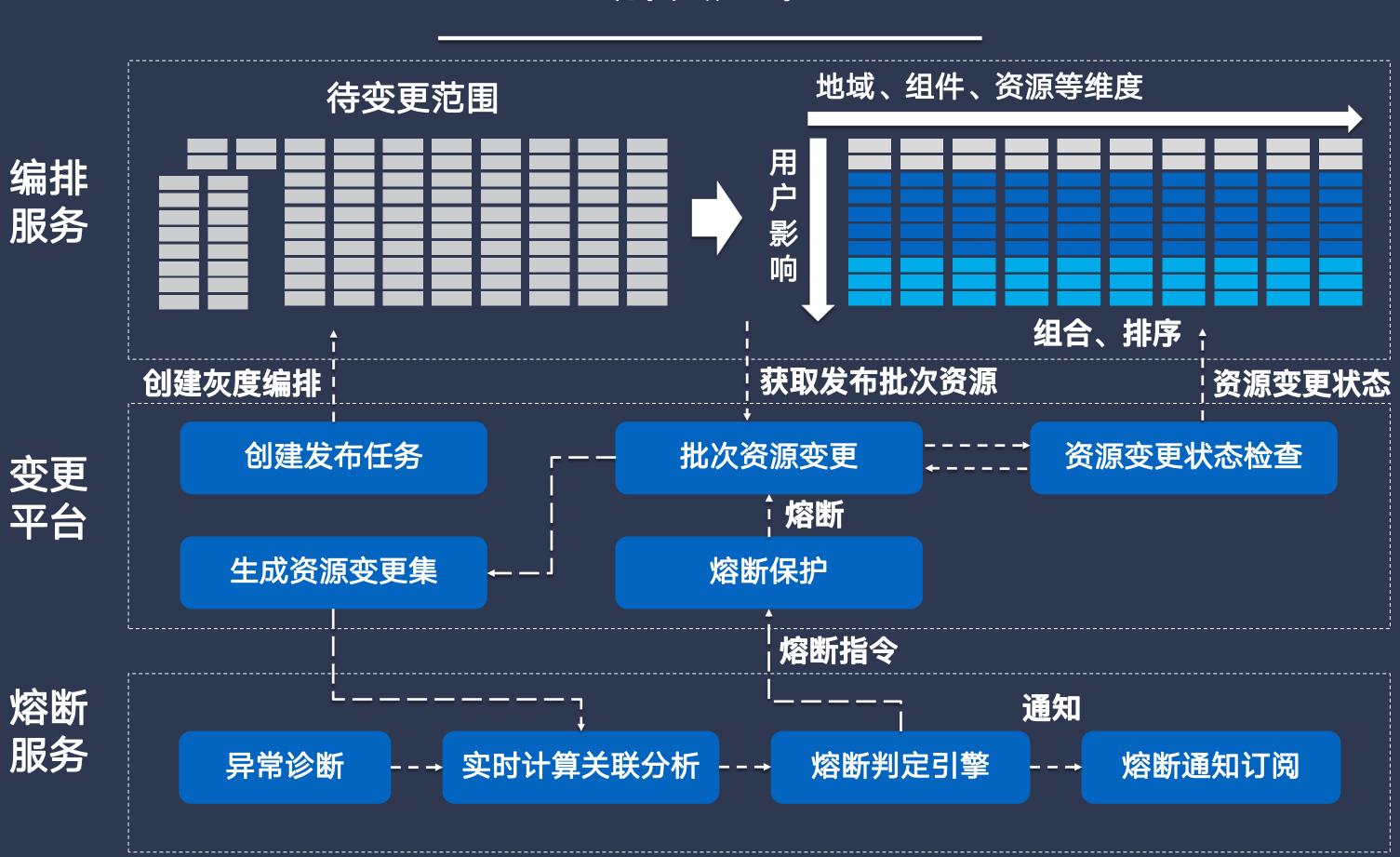
# 线上守护体系: 灰度发布能力



### 变更是导致故障和损失的头号杀手



### 解决方案



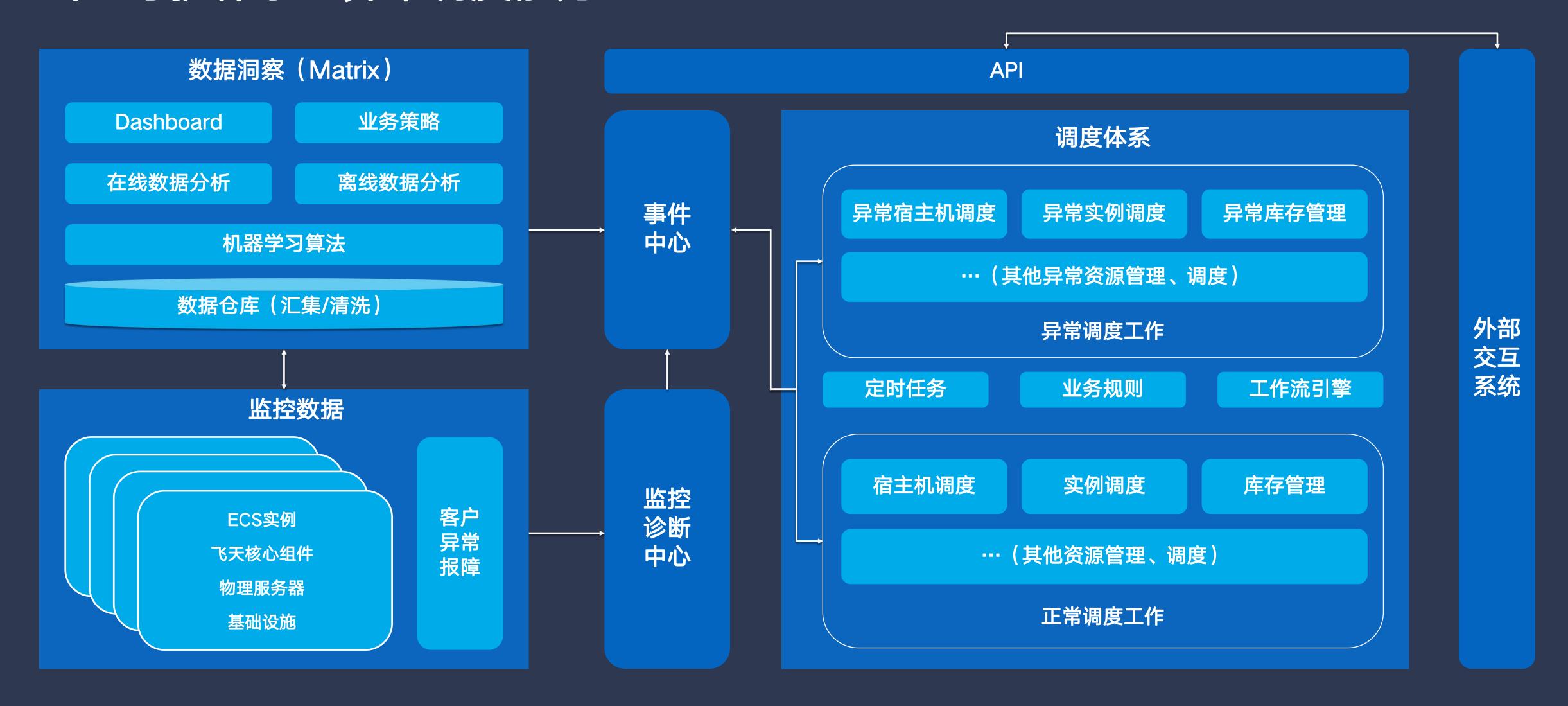


# 线上守护体系: 监控诊断能力



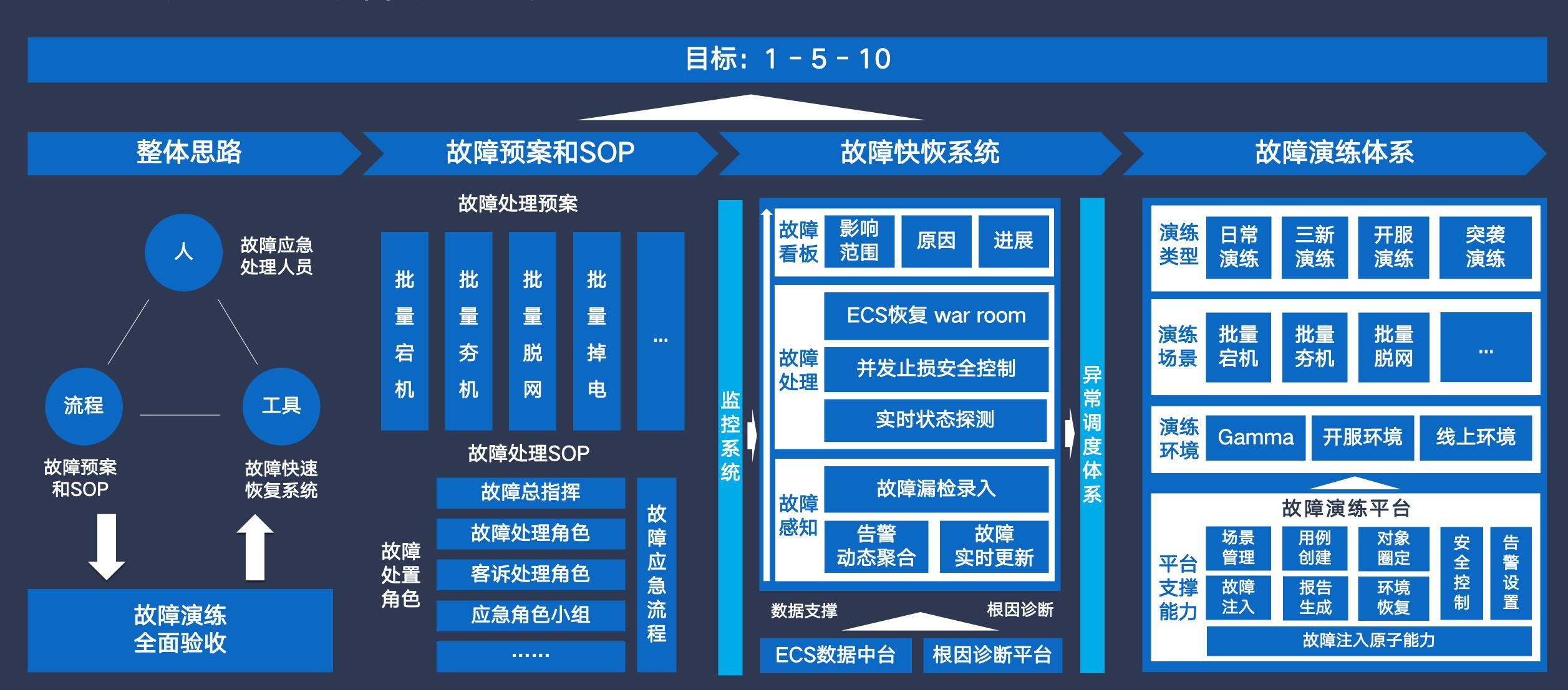


# 线上守护体系: 异常调度能力



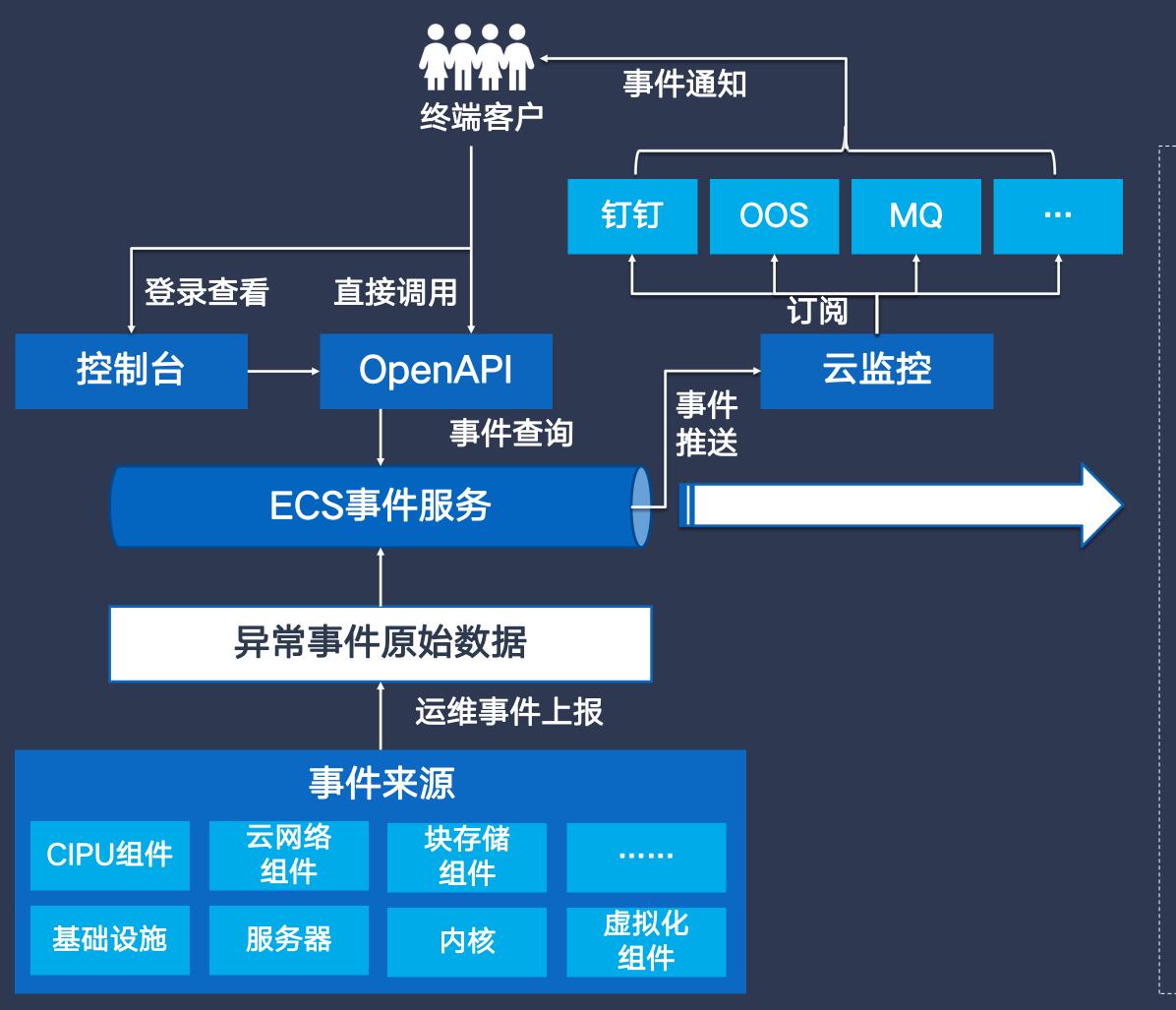


# 线上守护体系:故障快恢能力





# 客户侧联动体系



### 系统事件类型

• 为规避底层宿主机软 • 底层宿主机突发软硬 非预期 计划内 硬件隐患发起的主动 件故障导致的实例重 运维事件 运维事件 维护事件 启或宕机发送的事件 突发性能实例的CPU 突发性能实 因本地盘实例故障或 本地盘 积分已耗尽,可能对 例性能受限 单磁盘故障发送的事 实例事件 实例性能产生影响的 件 事件 提醒事件 • 影响费用的事件,例 • 升级或改造基础设施 实例因底层 实例 如实例到期、账号欠 时,可能影响对应实 升级需迁移 费用事件 费,导致实例即将被 例,提醒根据系统事 事件 件引导进行实例迁移 停止或释放 • 影响实例安全的事件, 实例的生命周期或其 例如实例遭遇DDoS攻 状态 实例 他状态发生变化时发 变化事件 安全事件 击或进入黑洞,导致 送的事件 实例安全受到威胁



# 稳定性重保体系

需求场景

极致稳定性

诉求

重保体系

客户 新业务上线

客户大型 运营活动

客户业务 敏感期

客户 重大变更期

.....



重保等级 和套餐

一级重保

三级重保

额外资源成本投入 对后端服务影响

研发与服务人力投入

额外资源成本投入

对后端服务影响

 $\star\star\star\star\star$  $\star\star\star\star\star$ 

 $\star\star\star\star\star$ 

 $\star\star\star\star\star$ 

二级重保

额外资源成本投入 对后端服务影响

 $\star\star\star\star\star$ 

四级重保

研发与服务人力投入 ★ ★ ★ ★ 额外资源成本投入 对后端服务影响

 $\star\star\star\star\star$ 

重保技术 支撑平台

重保业务模型



重保原子能力



重保策略



# 阿里云弹性计算云上运维最佳实践





## ECS云上运维最佳实践概览

### 响应ECS系统事件

通过OPENAPI响应ECS系统 事件,优雅规避实例运行风险

### 弹性能力规划容量

使用ECS弹性能力,自动规 划管理应用系统容量水位

### 选择适合的实例规格

结合应用实际场景合理选择 实例规格,兼顾成本和稳定

### 保持GuestOS更新

不选择已经EOL的OS,定期 更新系统补丁和软件驱动

### 问题规避

### 建设高可用架构

通过HA能力提升应用单实例 宕机容错和极端场景容灾能力

### 建立应用防抖动能力

设置合理的访问超时重试机制,规避实例短暂抖动对业 务应用的影响

### 选择适合部署方式

结合应用场景,使用专用宿 主机或部署集,实现亲和性 或反亲和性部署

### 实例健康自诊断

对实例本身的系统、网络、磁盘等状态进行全方位诊断

### 网络连通性自诊断

对不同ECS实例之间网络连通 性进行全面诊断,定位网络不 通的原因

### 问题报障

可通过OPEN API等方式将 ECS实例问题报给阿里云

### 应用侧止损

可在应用层通过摘除流量或 流量切换等方式先行止损

问题止损

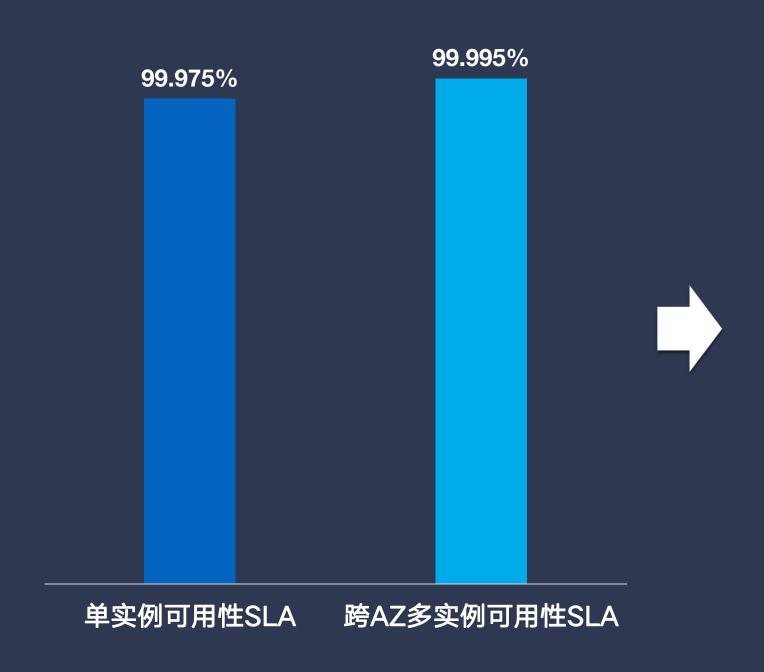
问题诊断

问题容错



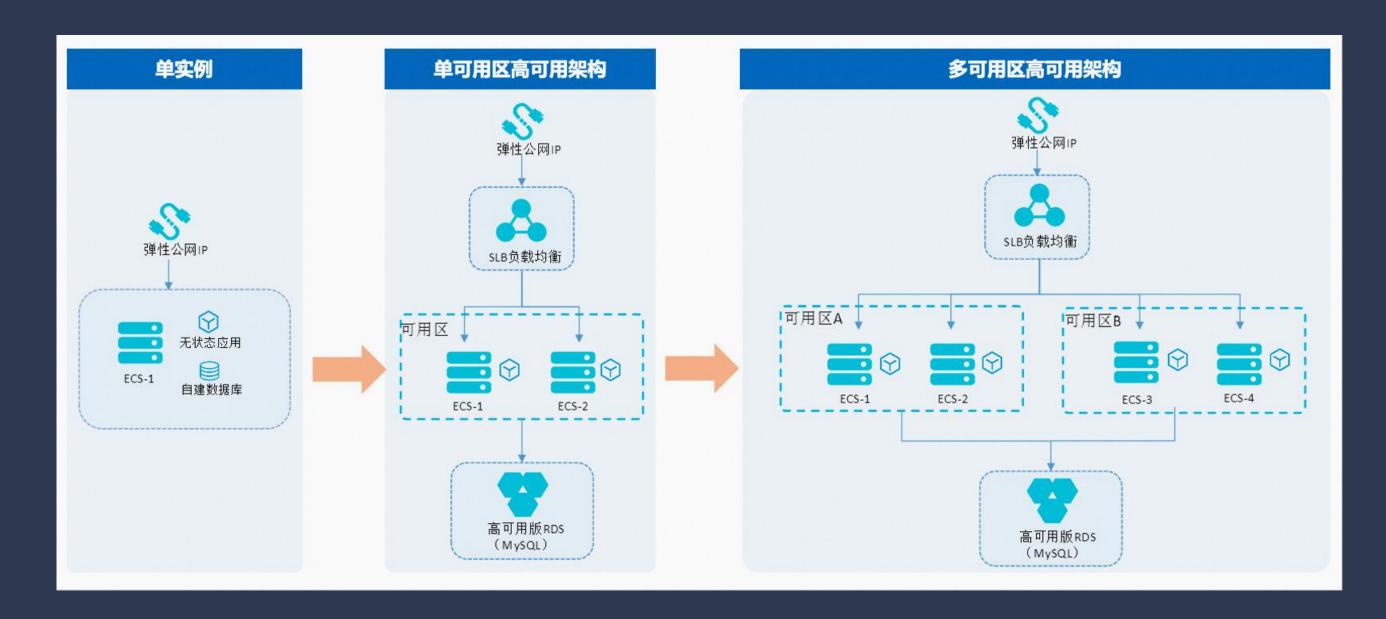
# 部署高可用架构,提升应用系统单实例宕机容错和极端场景容灾能力

### 阿里云ECS实例可用性SLA



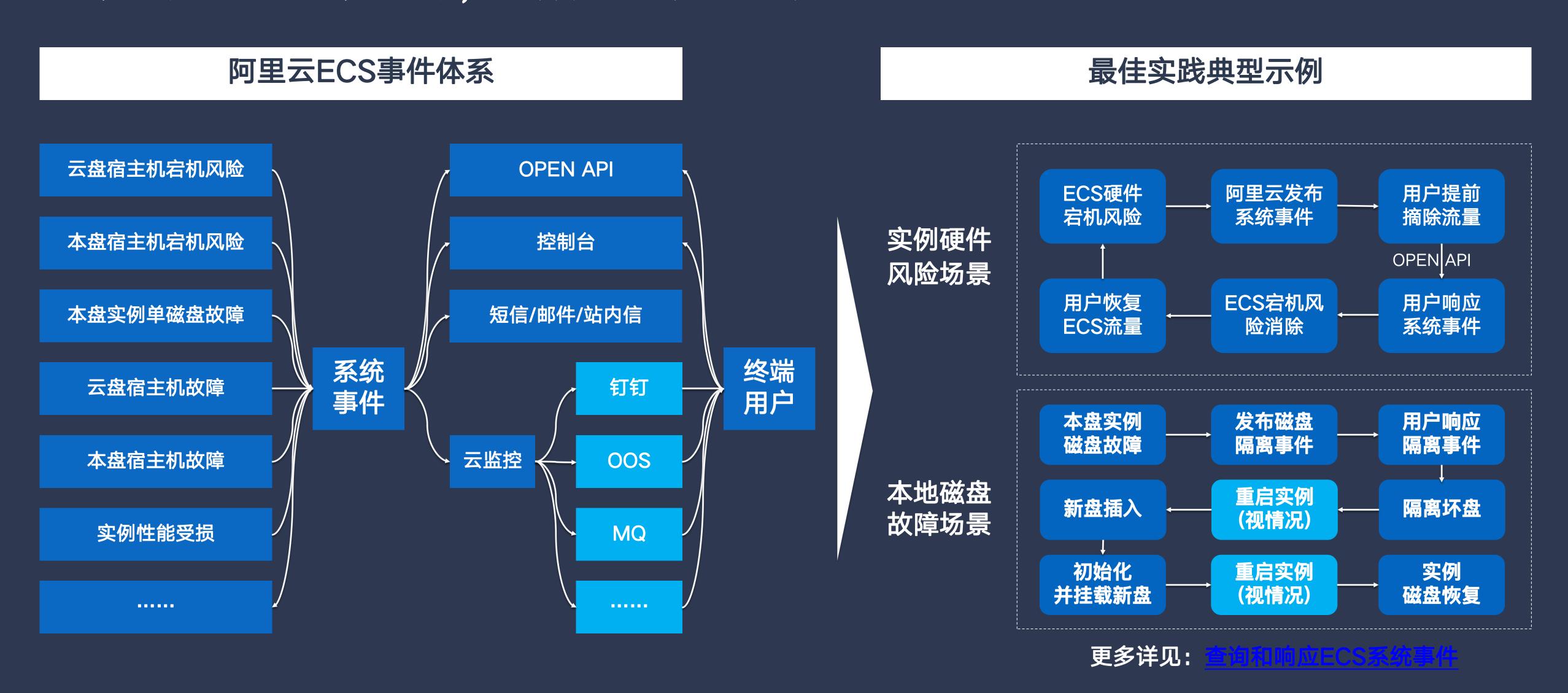
- 负载均衡SLB流量分发
- · 容量规划N+1冗余

- 单元化部署,架构消除拧麻花
- 具备应用系统流量切换能力





# 主动响应ECS系统事件,高效优雅规避实例运行风险



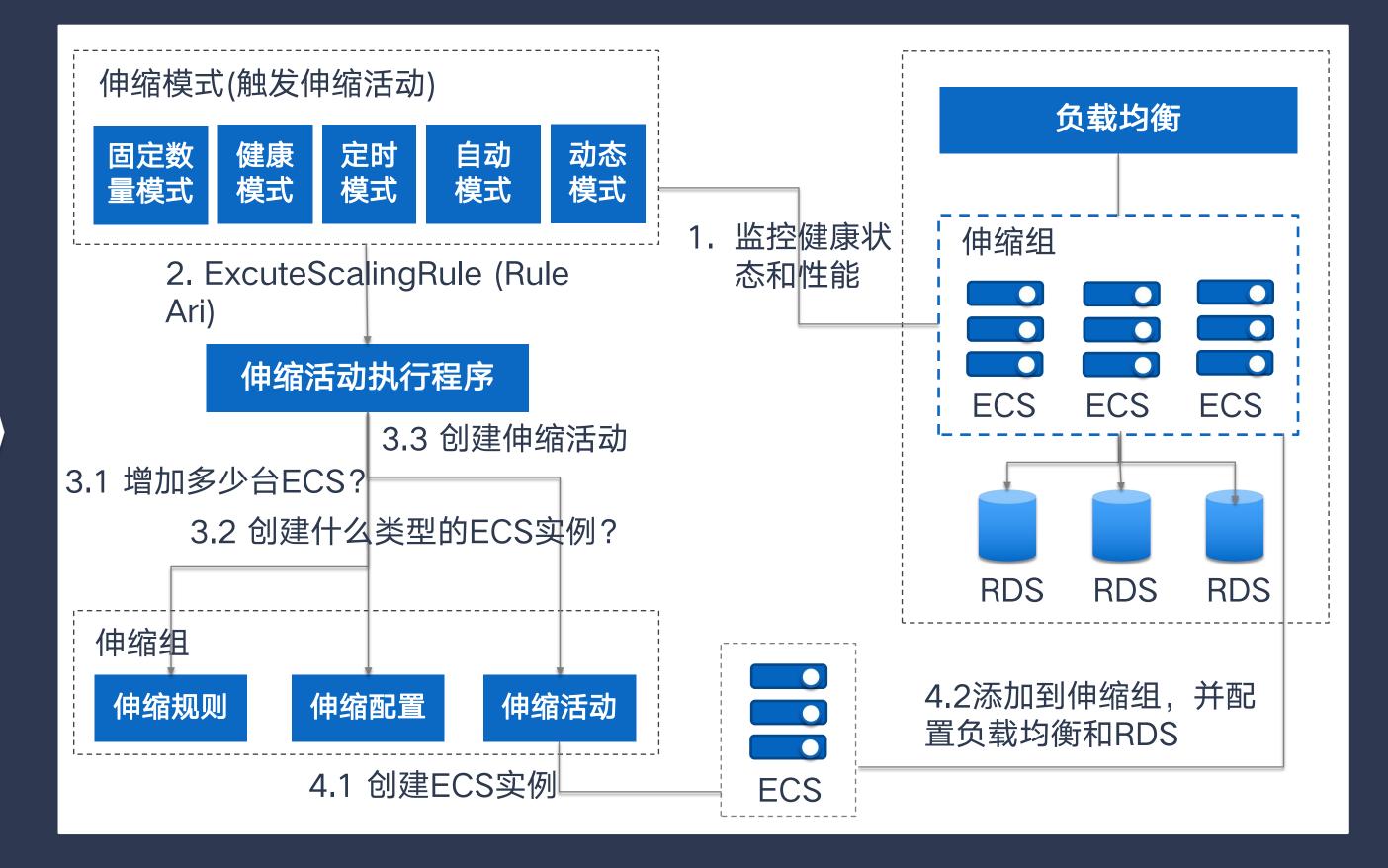


# 通过ECS弹性能力,低成本自动规划管理应用系统容量水位

### 业务场景

ESS弹性伸缩服务

业务量波动无规律,访问量突增 无规律业务量波动 和回落的具体时间难以预测 每天固定时间业务量急速增长进 有规律业务量波动 入高峰期, 到固定时间业务量下 降,高峰期结束 业务量无明显波动,如果现有计 无明显业务量波动 算资源突然出现故障,很难及时 维修,导致业务受到影响





# 使用实例健康诊断能力,快速定位问题根因,高效完成ECS实例深度巡检

痛点场景

实例健康诊断能力

使用方式

1



应用系统触发监控告警,根因难以定位,到底是云的问题还是自身系统问题?

2



业务大促前,需要对云资 源做一次全面深度巡检, 如何高效实施?



诊断ECS实例底层的资 源和虚拟化层

### 存储服务诊断

检查实例磁盘和存储的运 行状态

### 安全控制诊断

检查实例关联的所有安全 组相关状态

### 资源配额诊断

检查核心资源使用量是否 即将达到配额

### 网络服务诊断

检查实例内网络组件和外 部网络环境异常

### 实例配置管理诊断

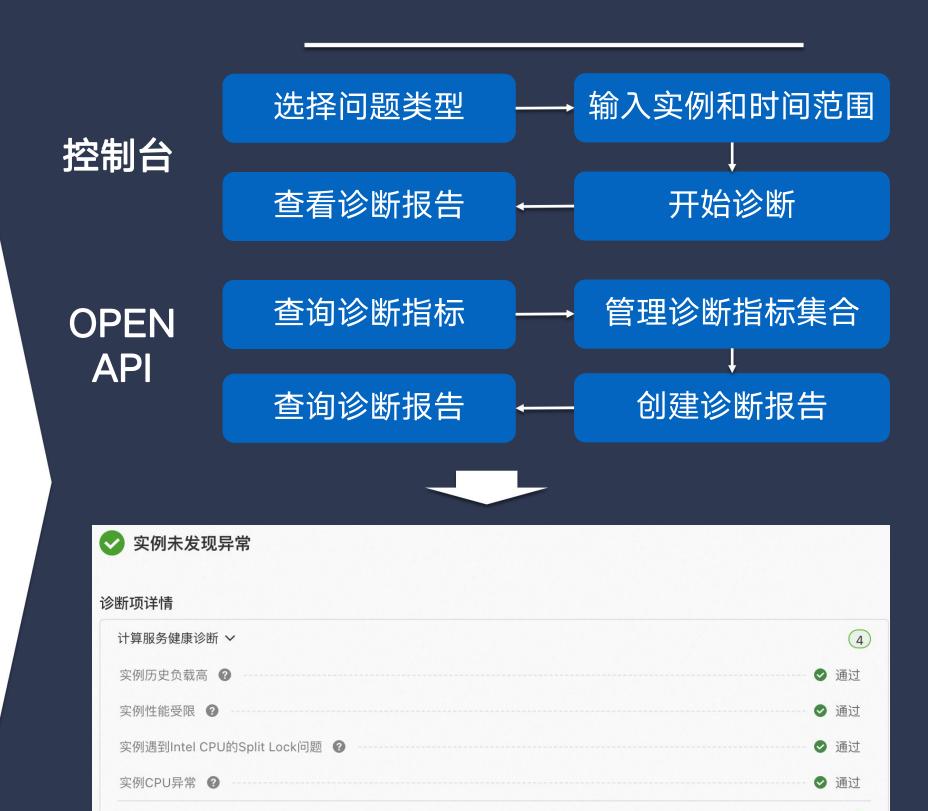
检查实例启动或运行过程 中,是否存在阻塞等操作

### 费用类诊断

检查实例本身和实例关联 组件的费用情况

### GuestOS内配置诊断

检查实例OS内系统文件、 关键进程等状态



存储服务健康诊断 ^

安全控制健康诊断 ^

网络服务健康诊断 へ

实例操作系统内相关配置诊断 へ



## 回顾总结

1) ------ • 什么是弹性计算?

2) ------- • 阿里云弹性计算稳定性建设实践

- 稳定性度量体系
- 线下预防体系
- 线上守护体系
- 客户侧联动体系
- 稳定性重保体系

3) ——— • 阿里云弹性计算云上运维最佳实践

- 最佳实践概览
- 部署高可用架构
- 响应主动运维事件
- 弹性能力管理容量
- 自诊断定位根因





# 谢谢

为了永不停机的计算服务!



# 想一想,我该如何把这些技术应用在工作实践中?

THANKS



