

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>>



📍 深圳站

ArchSummit

全球架构师峰会

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>>



📍 北京站

QCon

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IaaS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



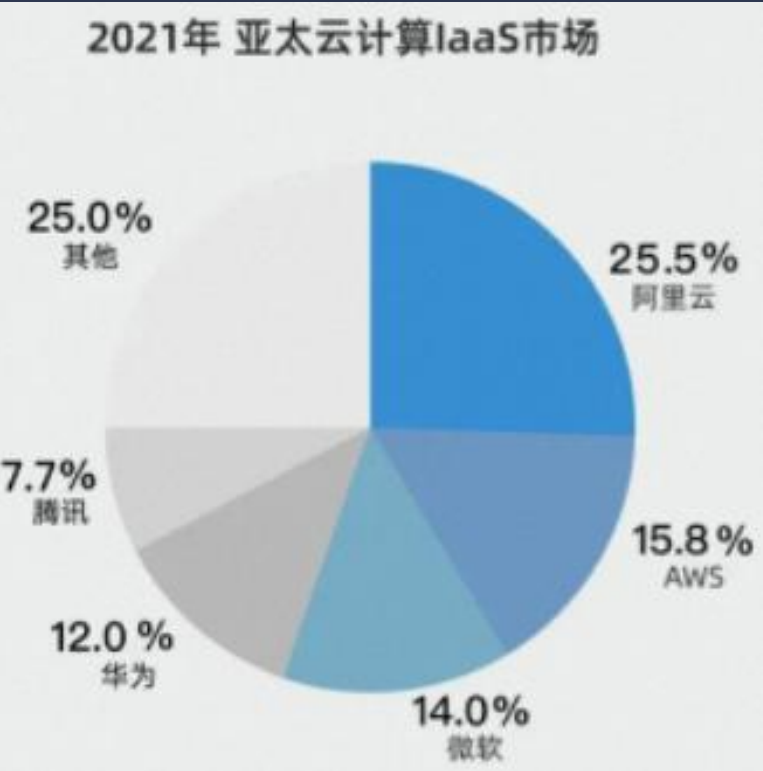
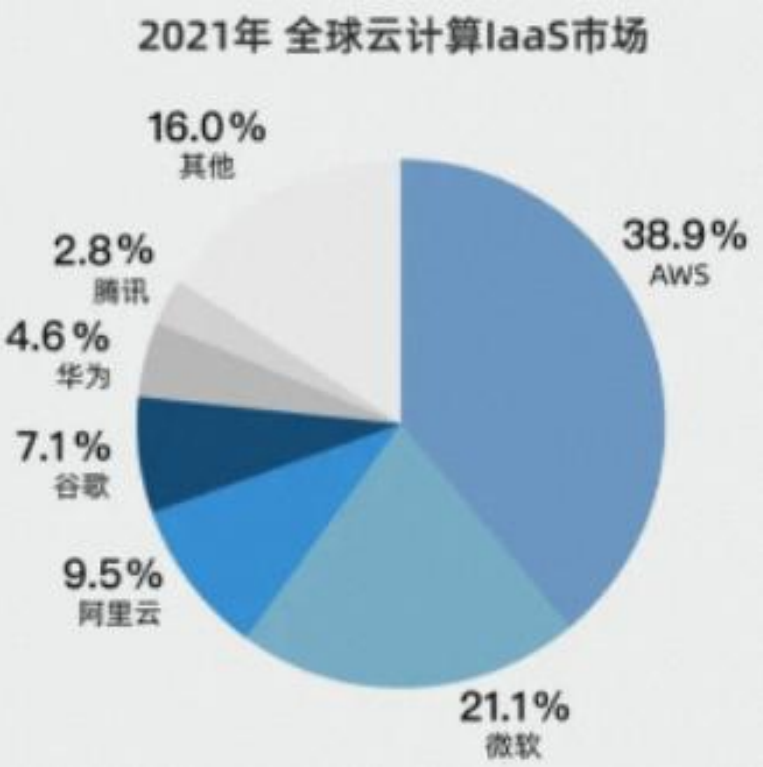
阿里云弹性计算稳定性建设实践

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



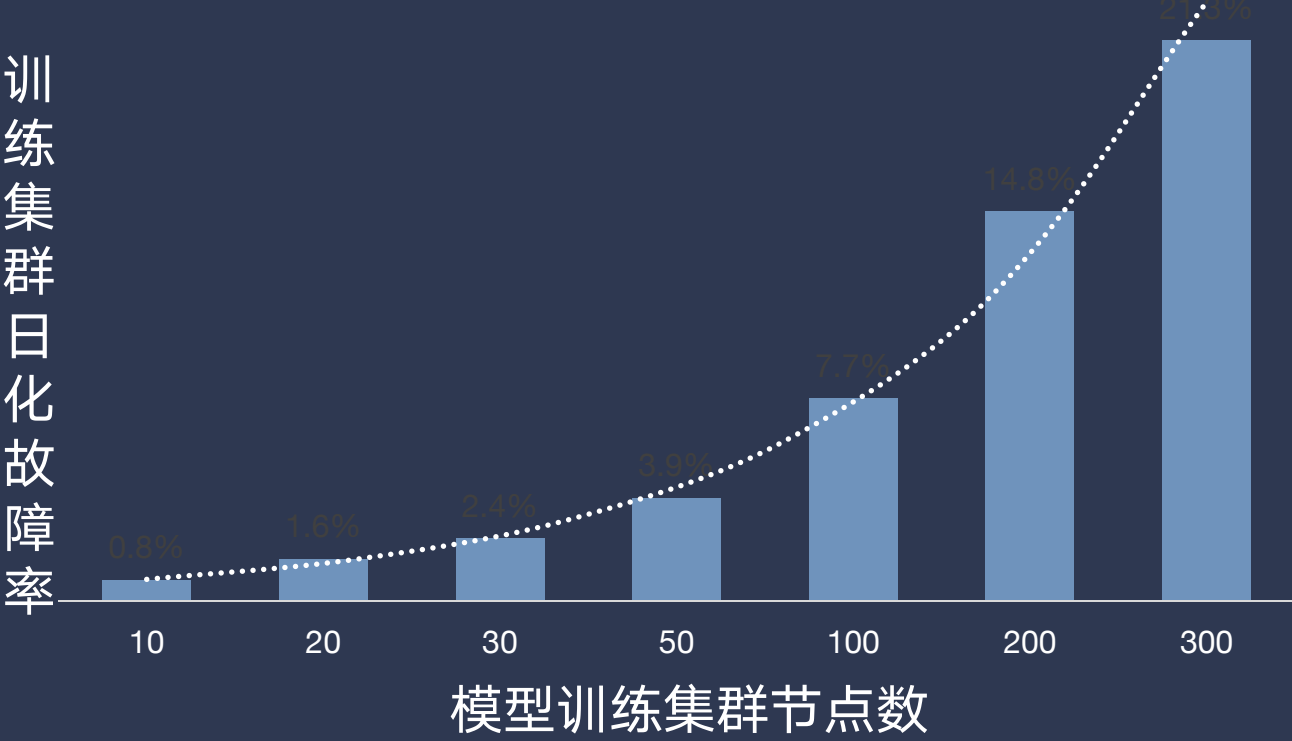
开机自启动客户占比



在线业务场景



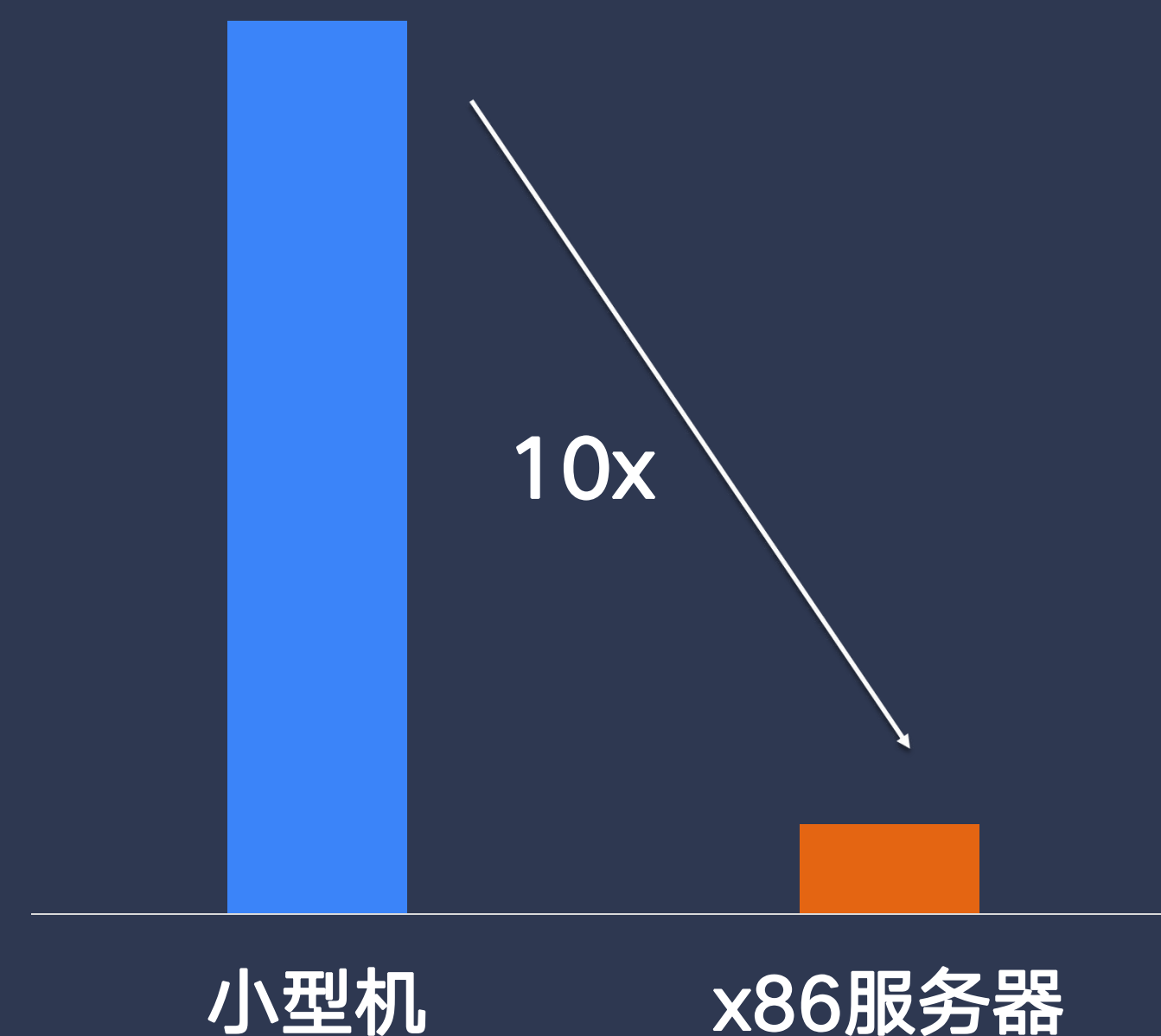
AI大模型离线场景



目标：用x86的硬件，提供小型机级别的稳定性

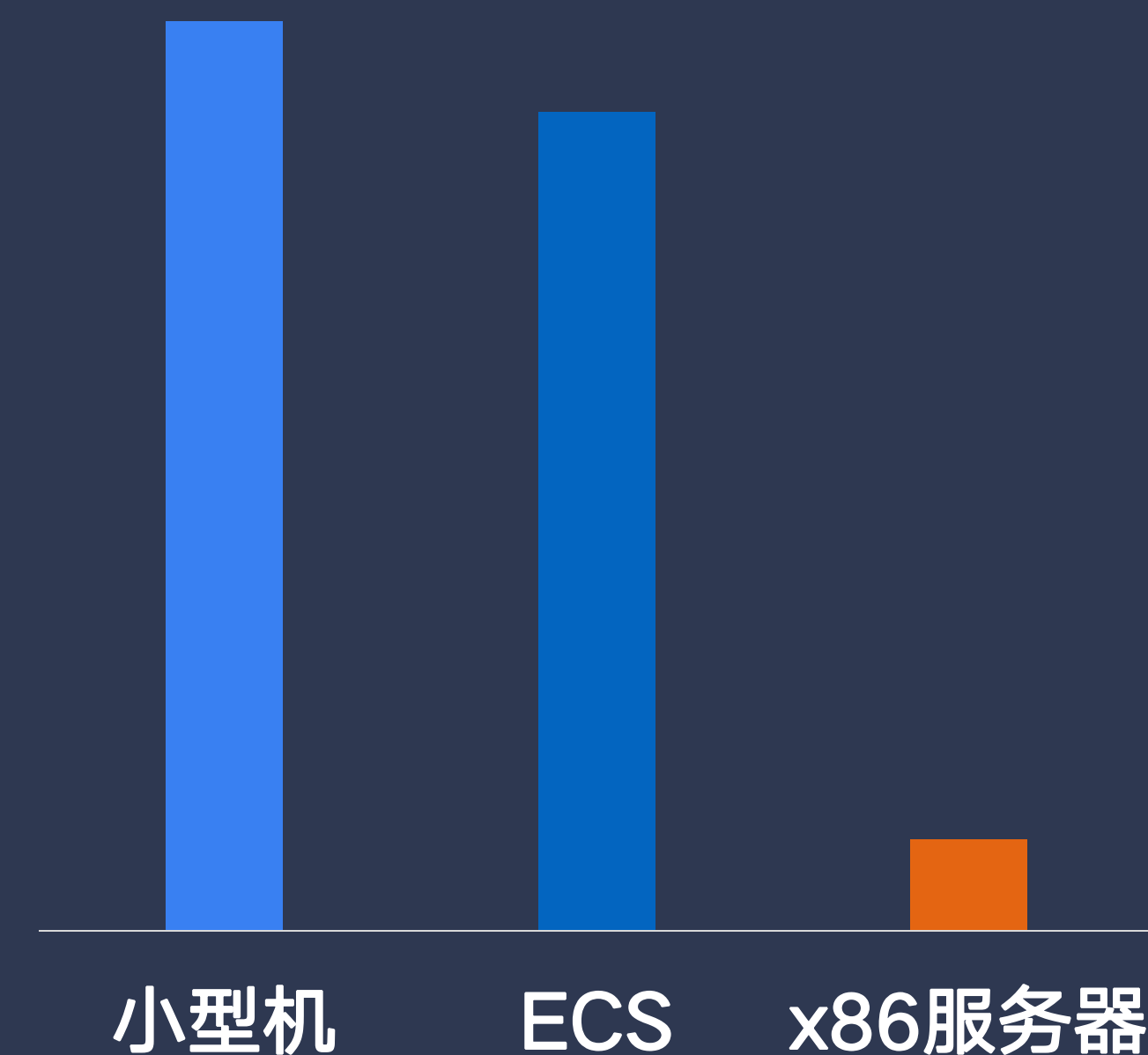
现实

体感可用率

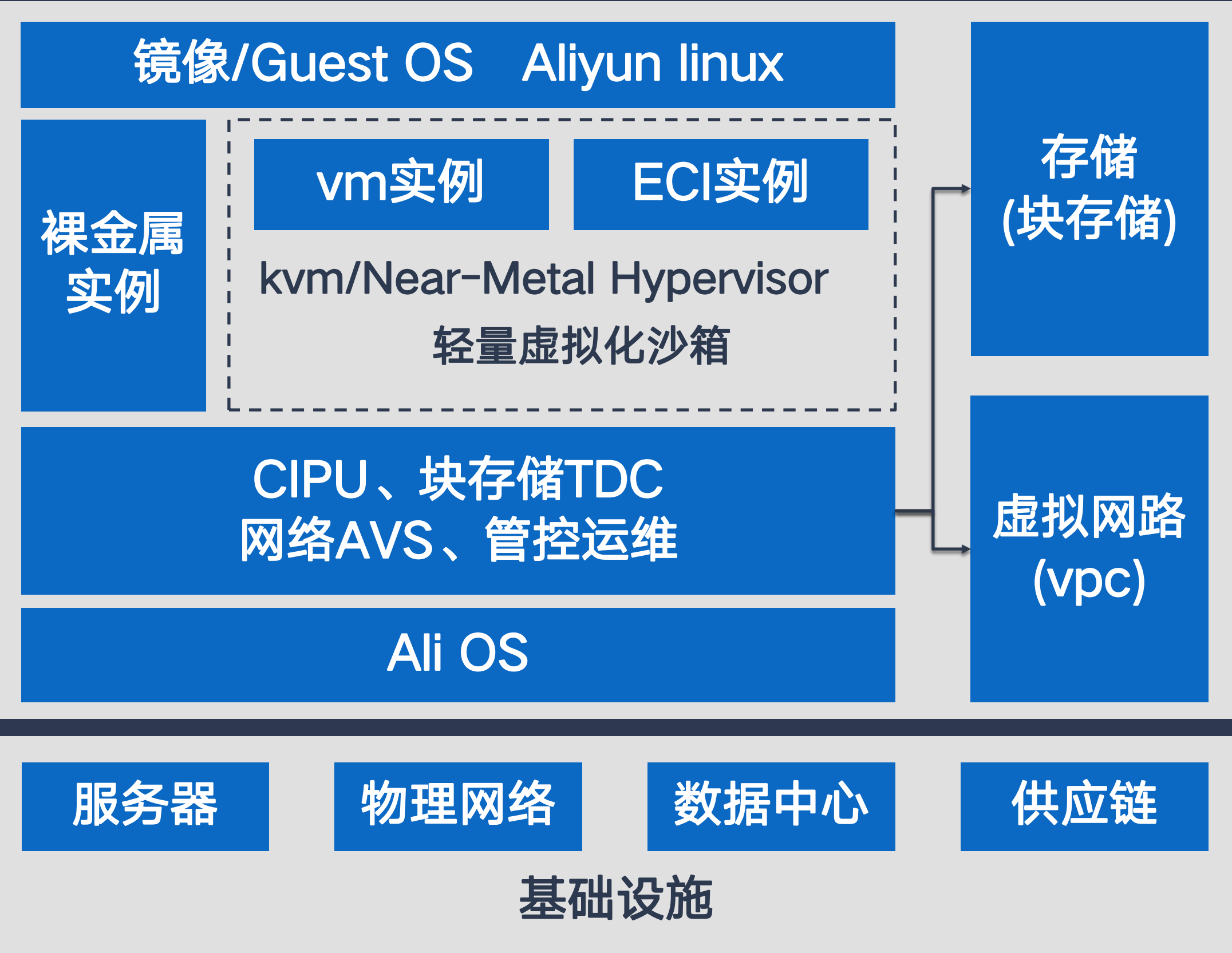


目标

体感可用率



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)

~10000000+
设备(Server)

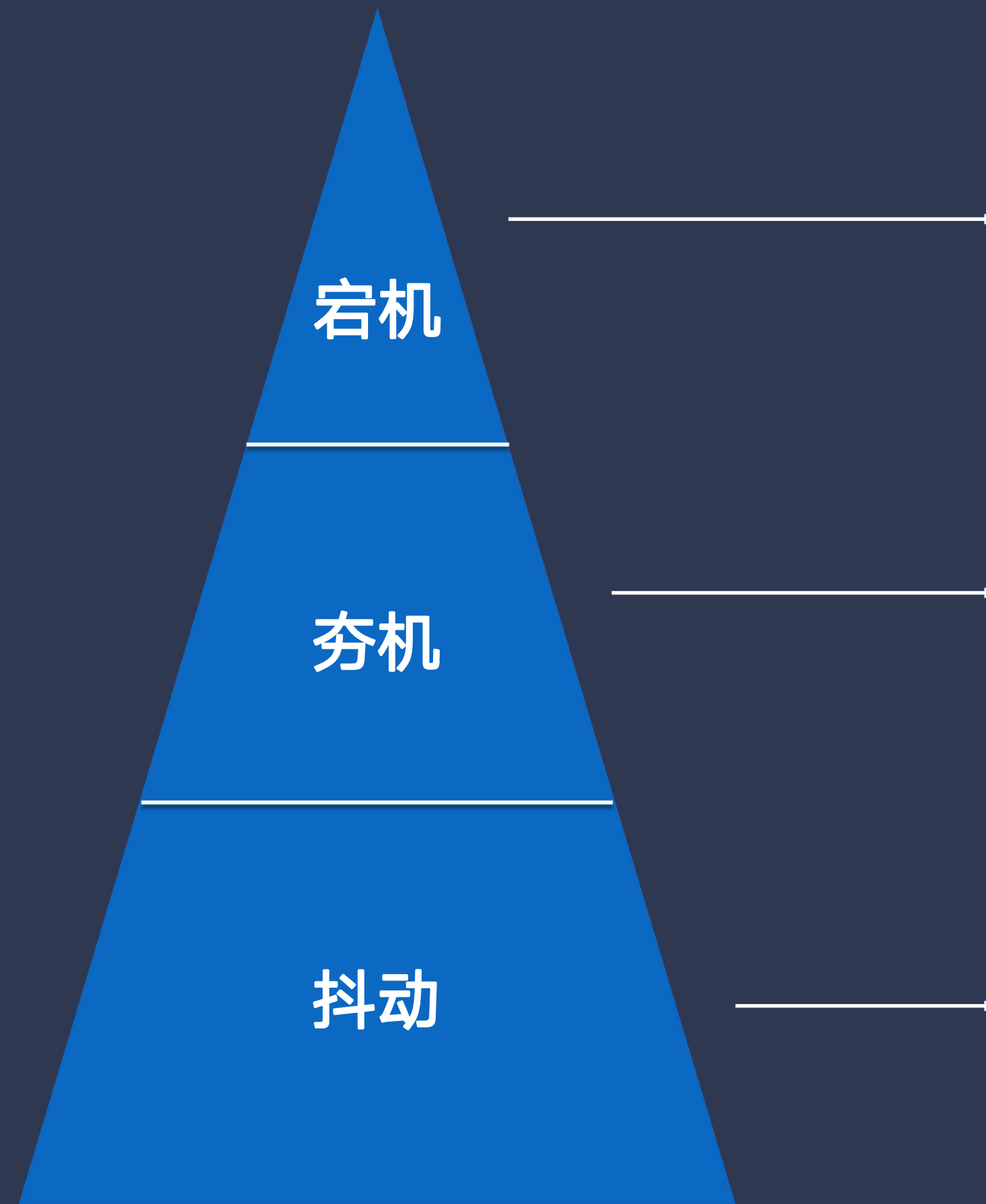
2023年4月

复杂度



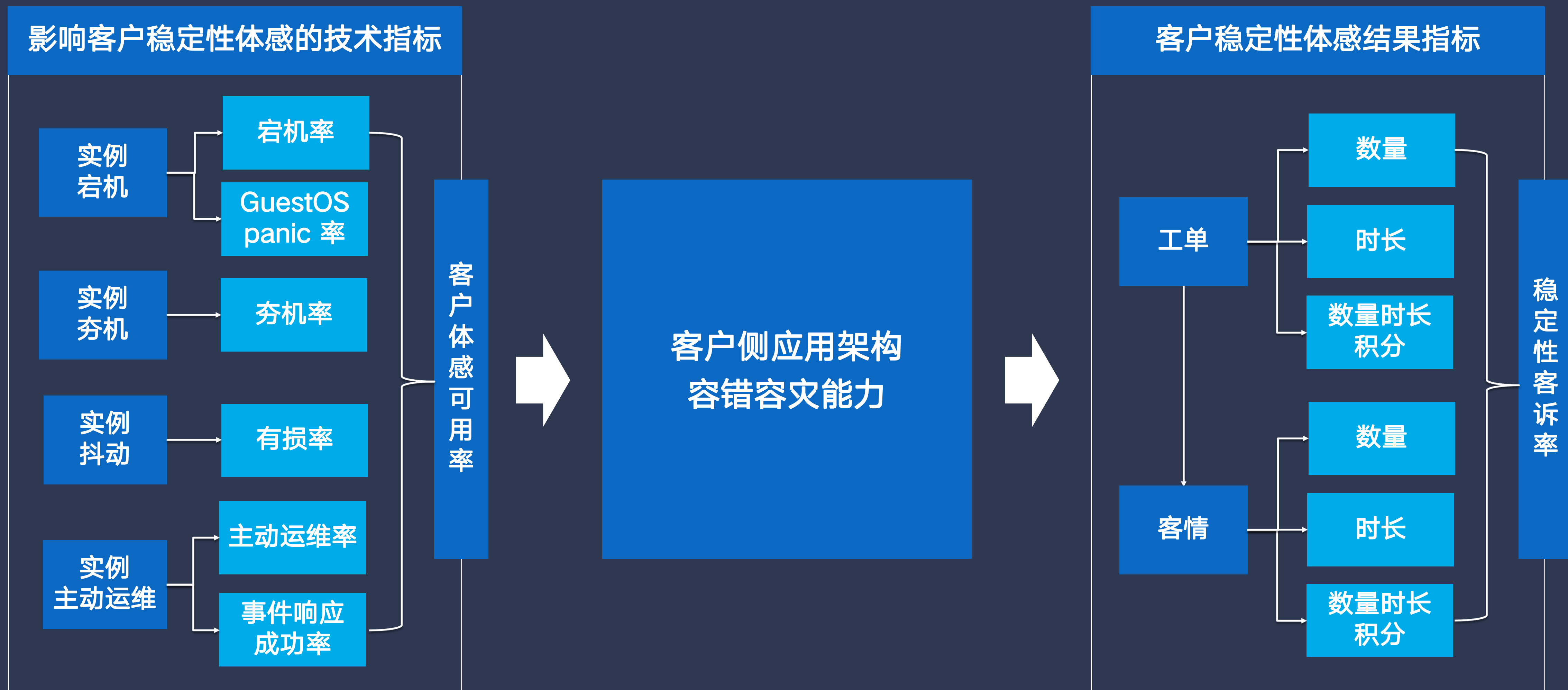
规模

识别问题：宕机、夯机、抖动

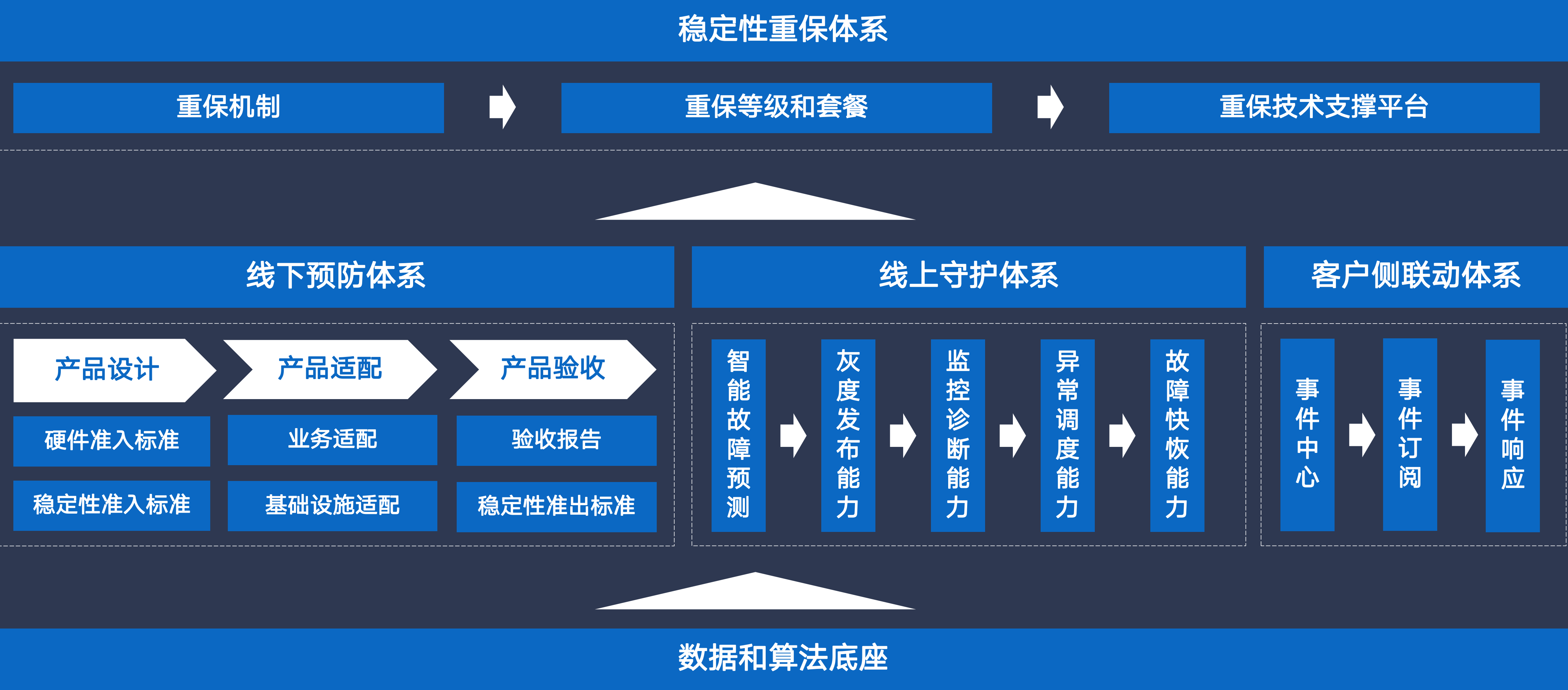


- **现象：**ECS资源100%不可用，多数因基础设施、服务器硬件或底层软件原因导致。
- **影响：**所有未持久化的数据和配置都将丢失，该ECS实例上的业务将完全中断。
- **现象：**ECS资源服务时断时续，甚至某些核心功能不可用或无法连接和操作。如：OS 夯，IO hang等。
- **影响：**未持久化数据尚未丢失，但整个ECS几乎无法使用，有时甚至无法恢复、没有备份的机会。
- **现象：**ECS资源核心服务可以正常使用，但在极端情况下会出现网络或性能抖动。
- **影响：**着重影响抖动敏感用户，性能抖动可能导致用户压测等容量规划付之东流，甚至可能因抖动引发用户应用系统雪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

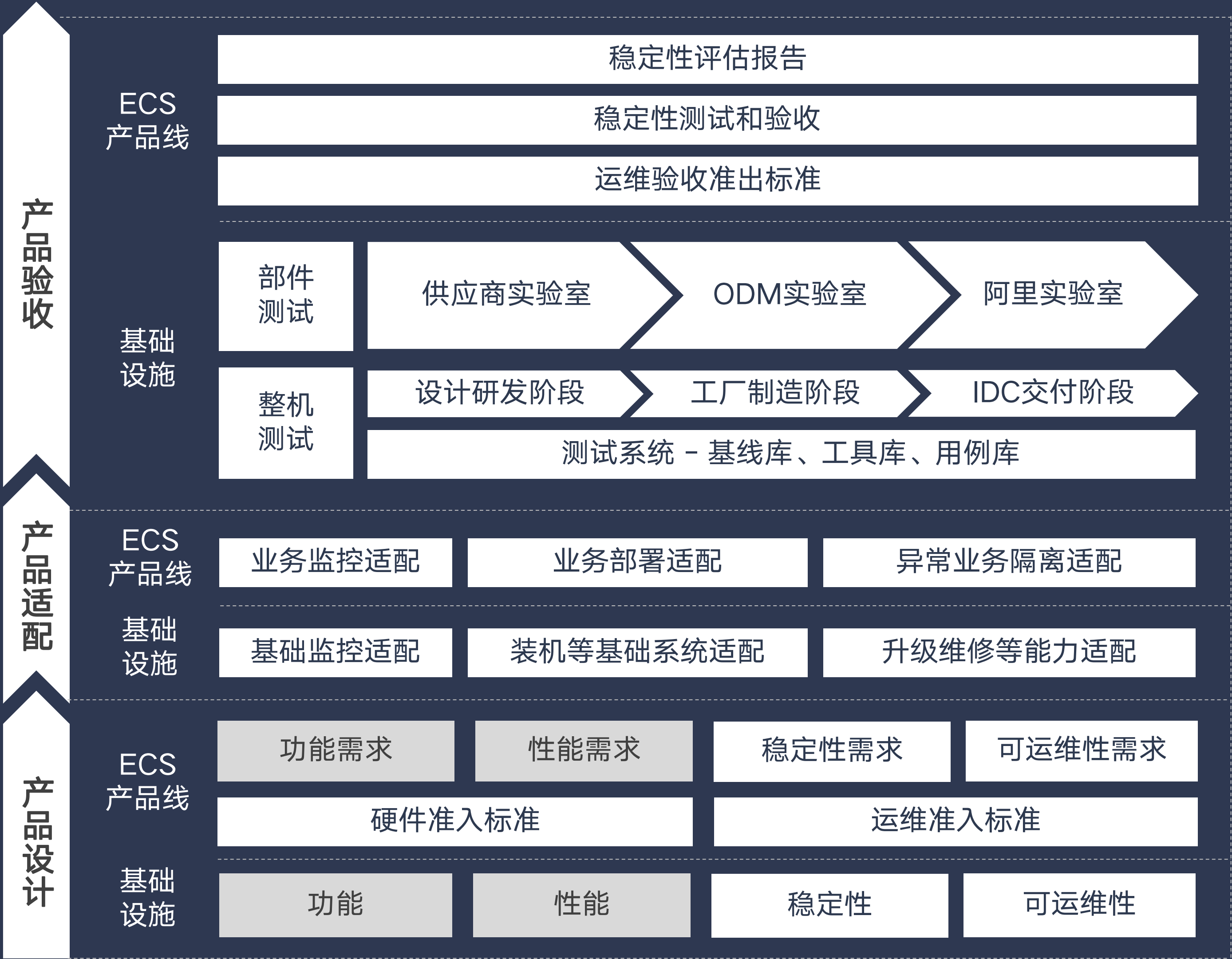
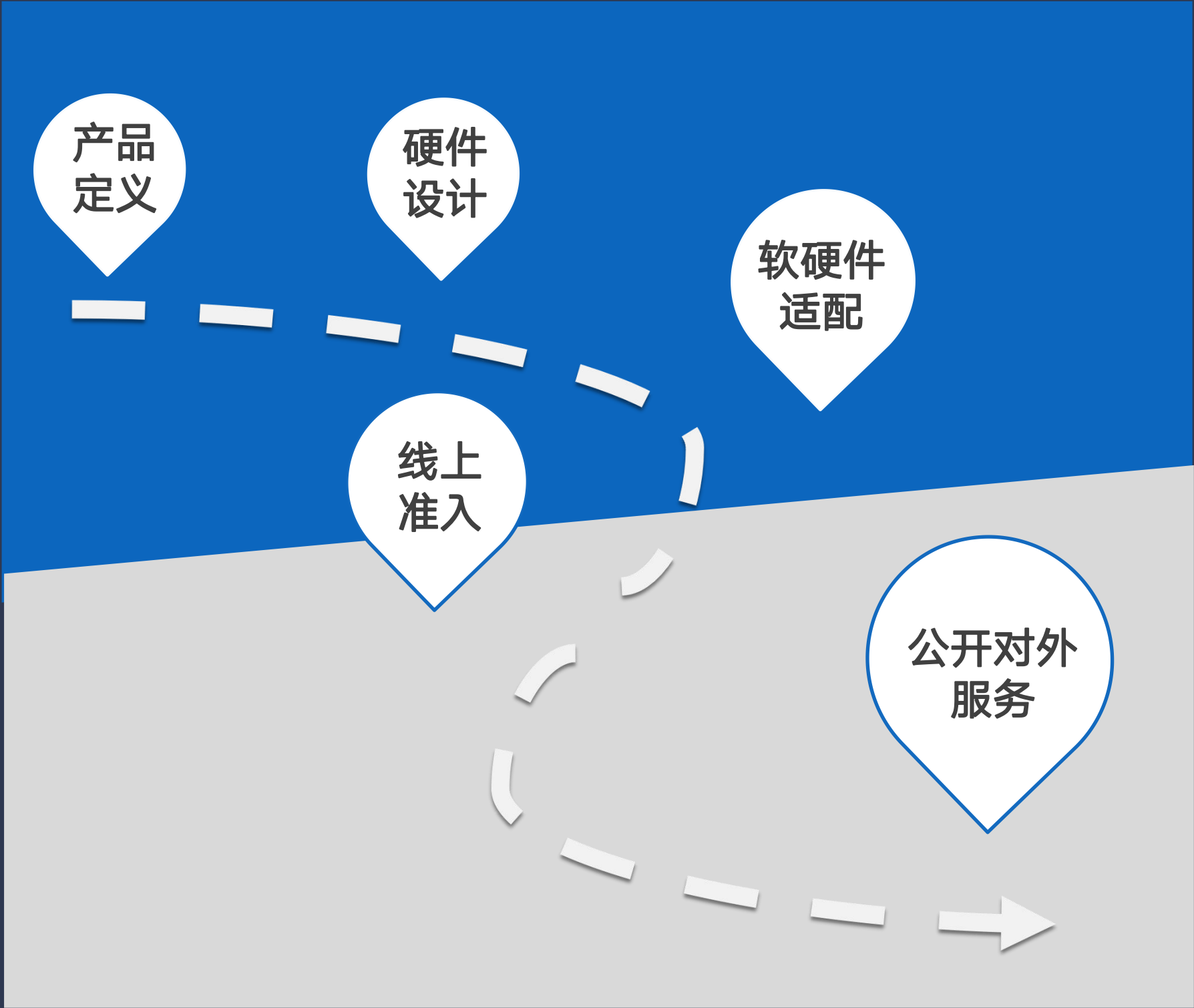
量化问题：从客户体感视角全面度量稳定性过程和结果



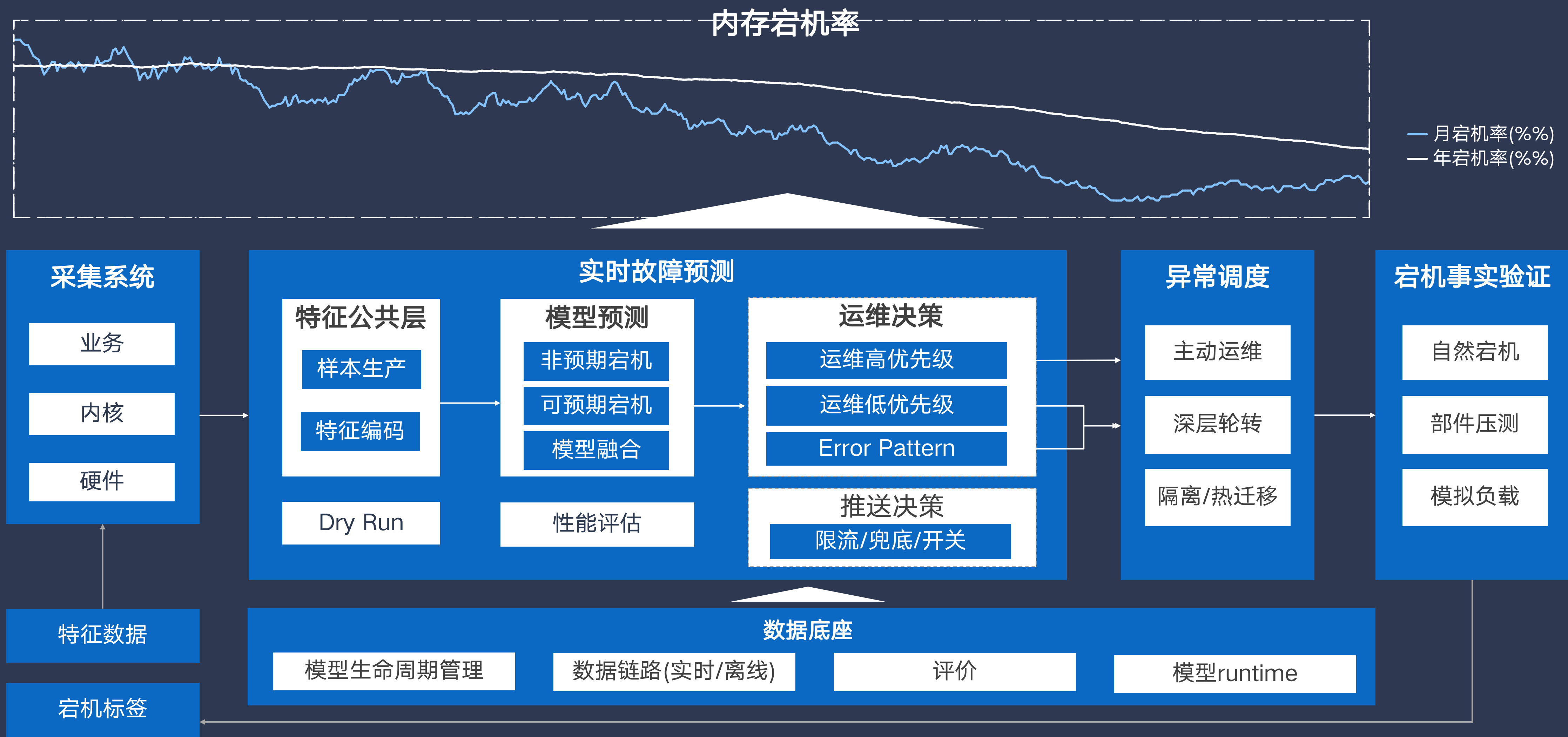
解决问题：弹性计算稳定性系统工程



线下预防体系



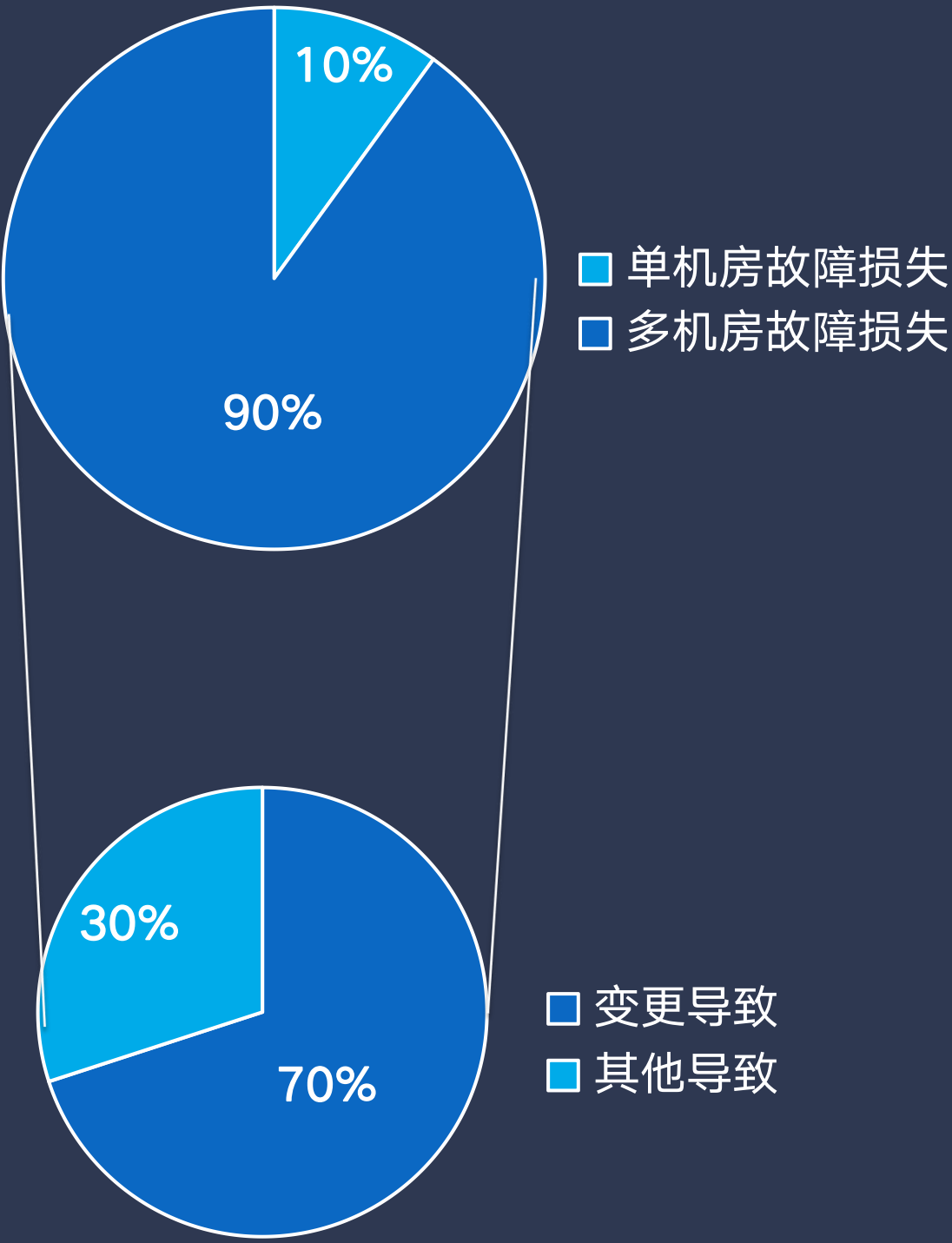
线上守护体系：智能故障预测



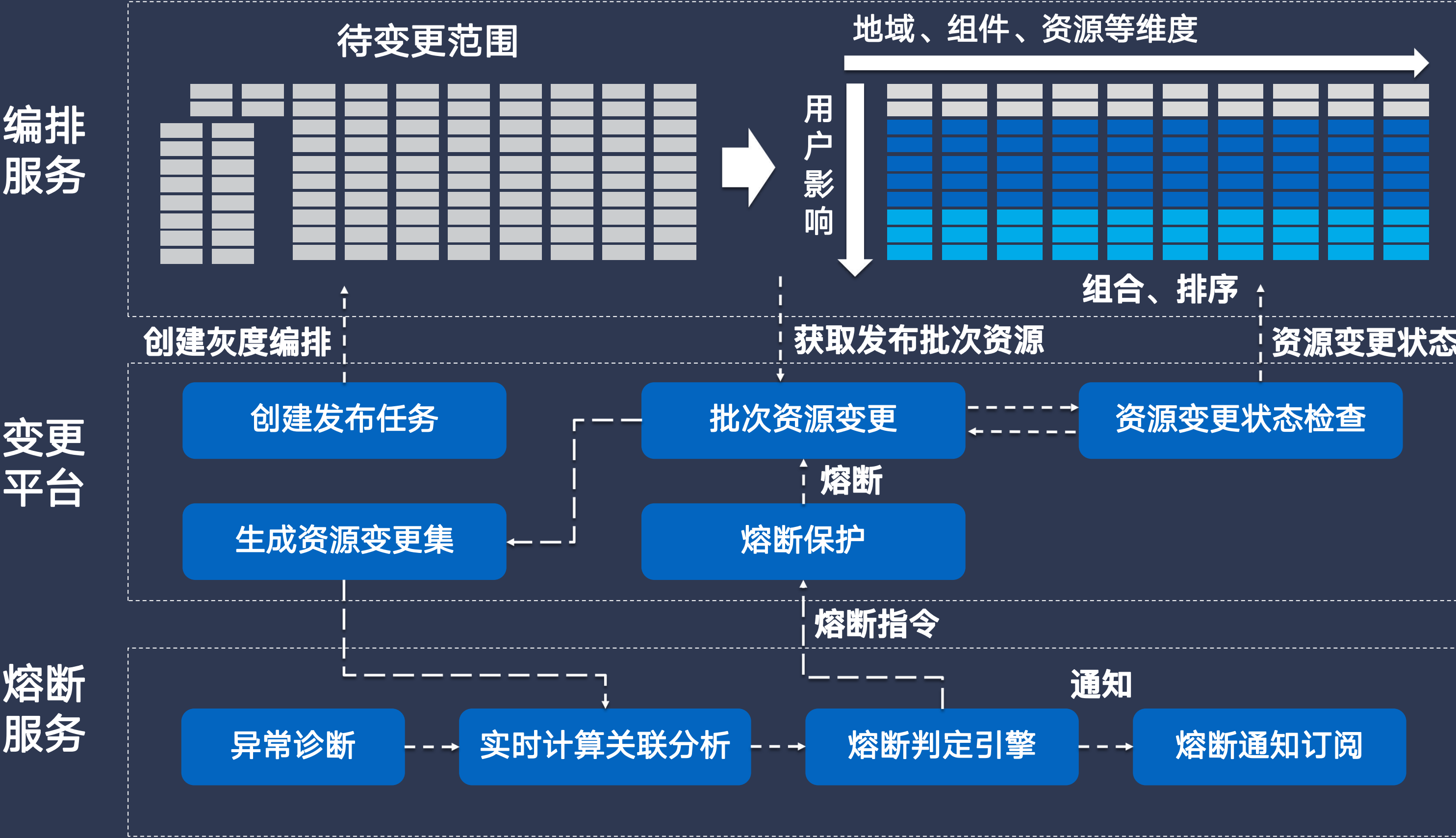
线上守护体系：灰度发布能力

问题

变更是导致故障和损失的头号杀手



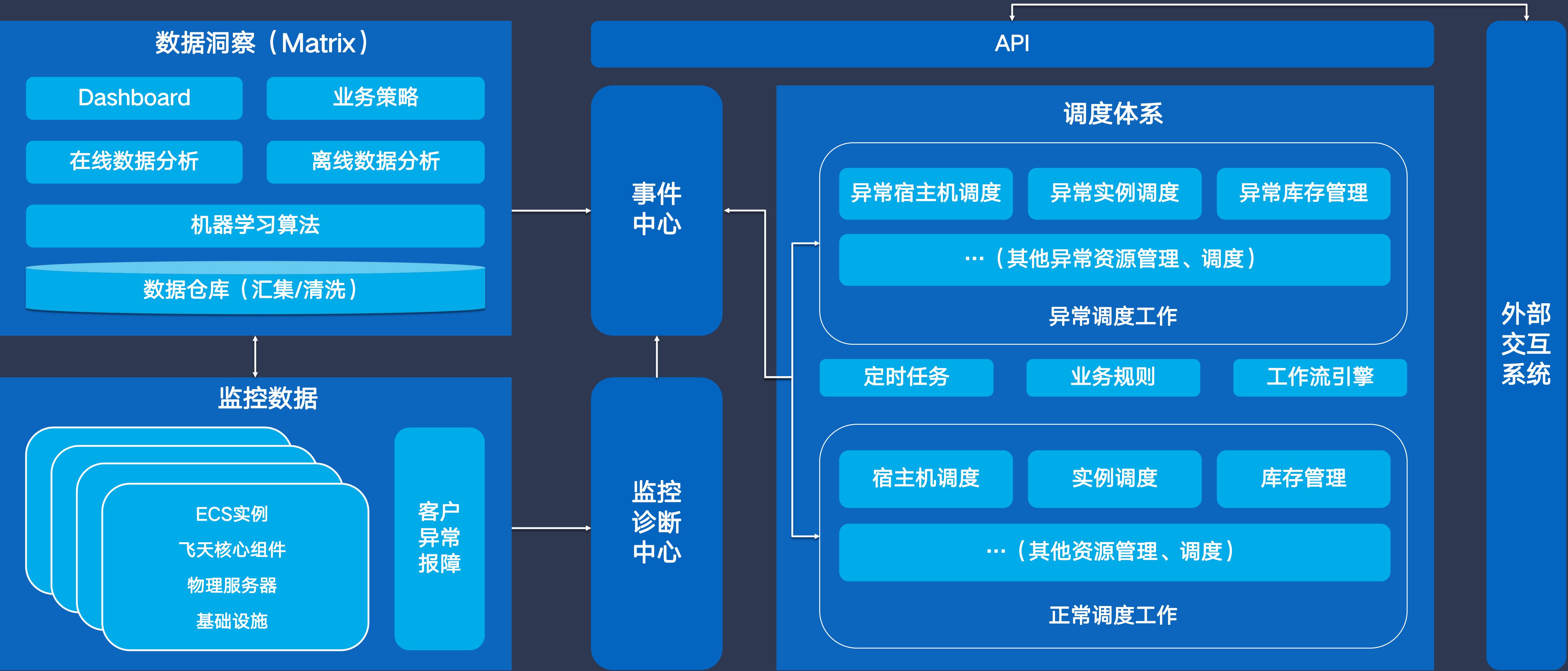
解决方案



线上守护体系：监控诊断能力

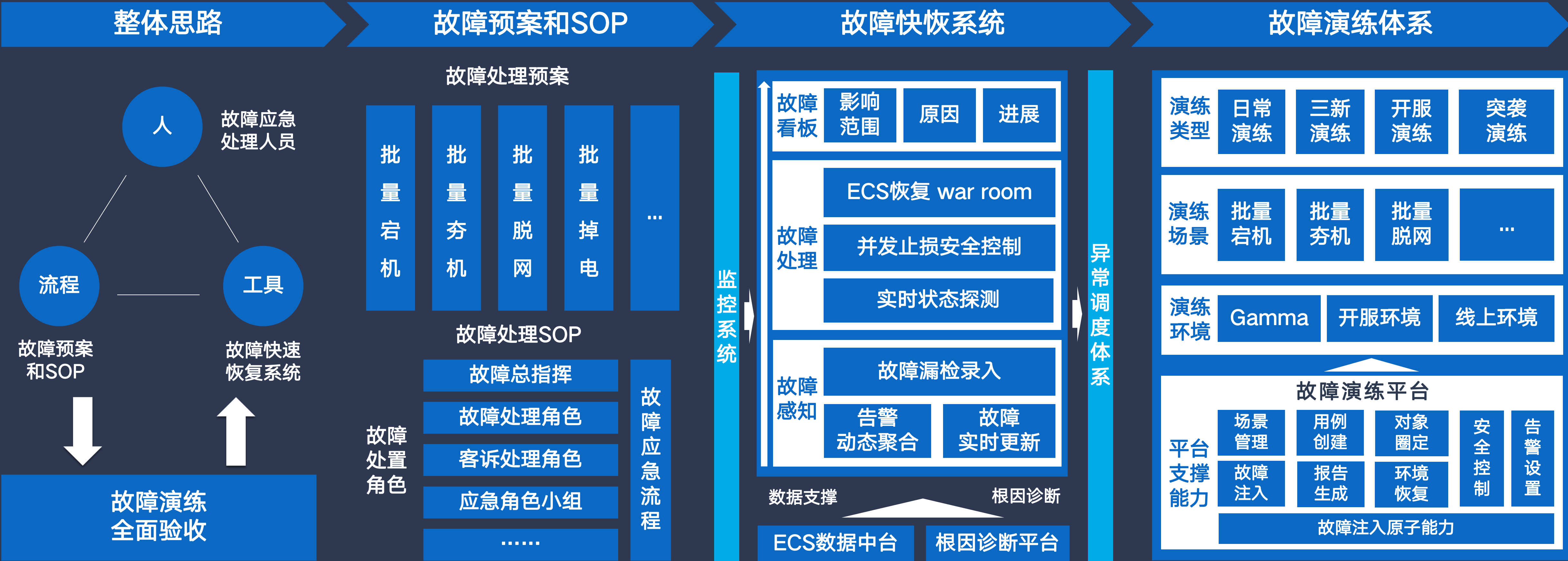


线上守护体系：异常调度能力

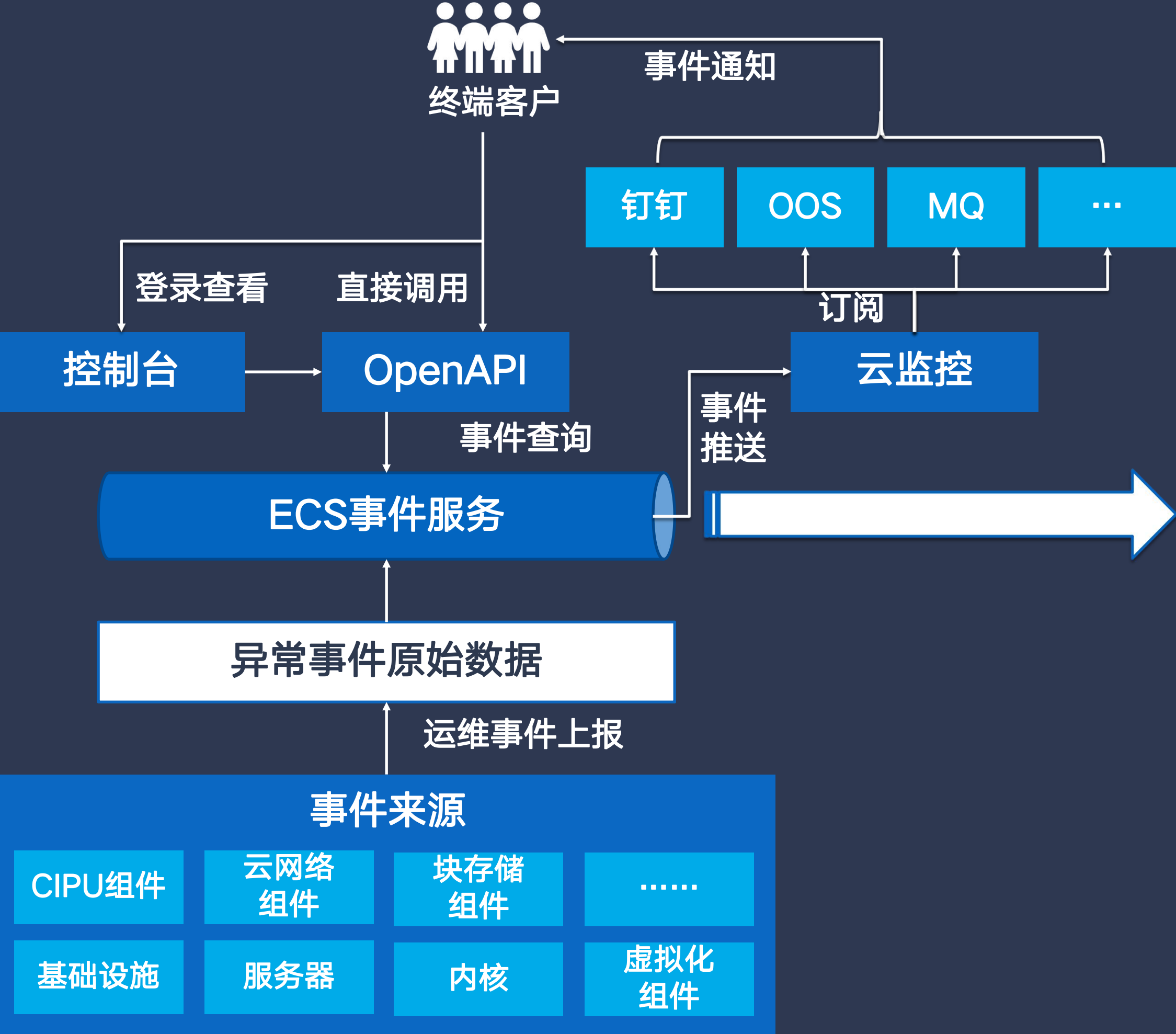


线上守护体系：故障快恢能力

目标：1 - 5 - 10



客户侧联动体系

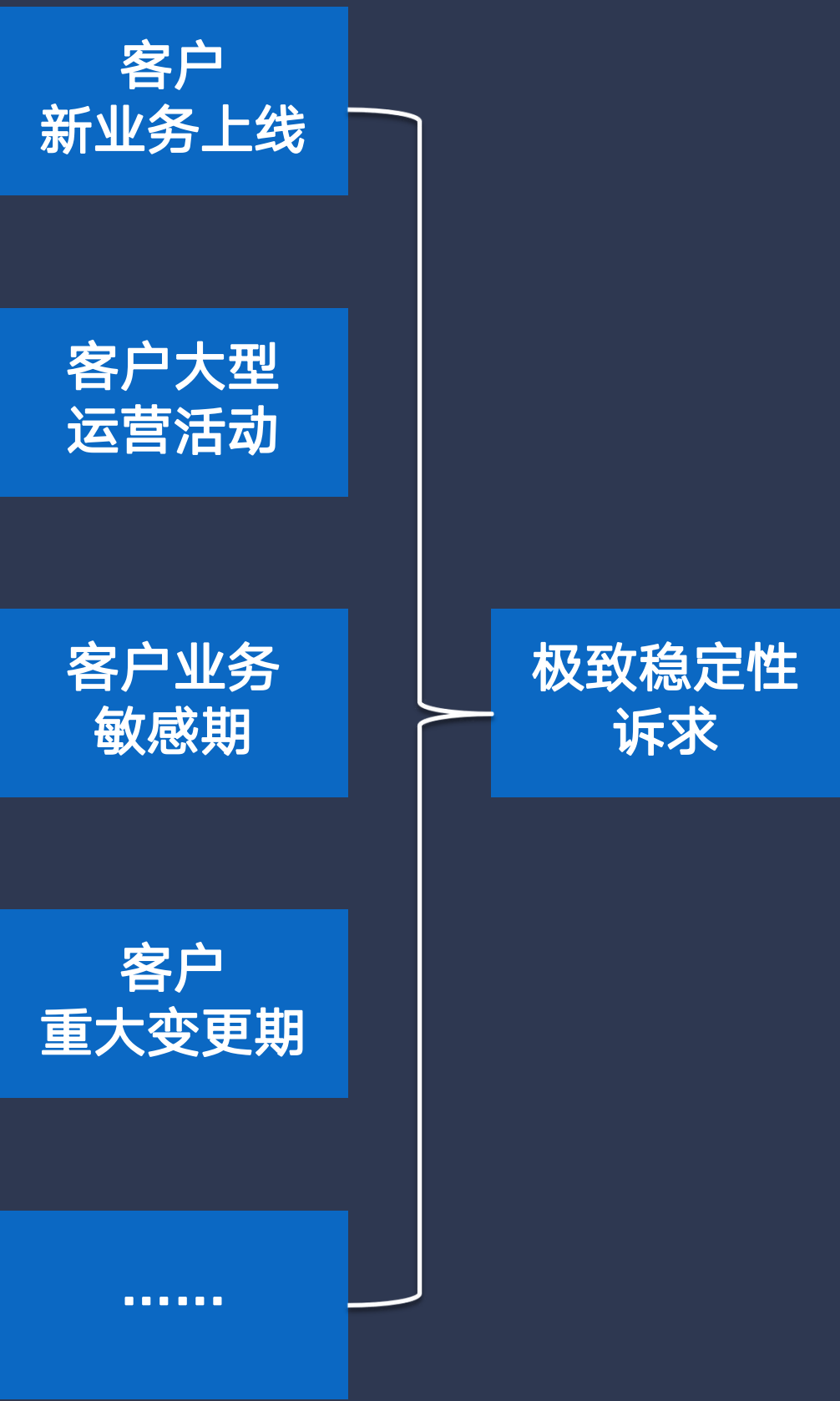


系统事件类型

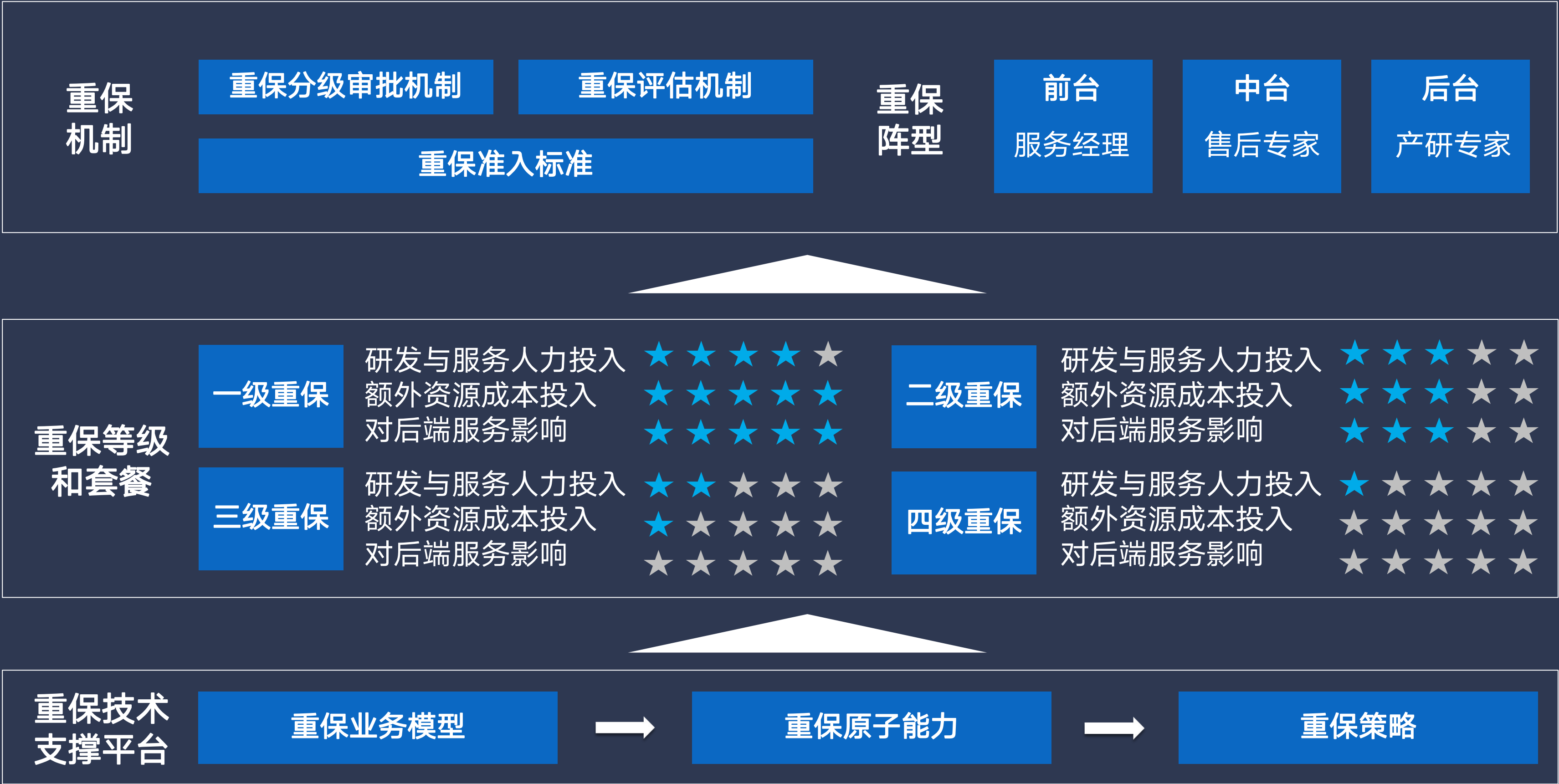
计划内运维事件	<ul style="list-style-type: none">为规避底层宿主机软硬件隐患发起的主动维护事件	非预期运维事件	<ul style="list-style-type: none">底层宿主机突发软硬件故障导致的实例重启或宕机发送的事件
本地盘实例事件	<ul style="list-style-type: none">因本地盘实例故障或单磁盘故障发送的事件	突发性能实例性能受限事件	<ul style="list-style-type: none">突发性能实例的CPU积分已耗尽，可能对实例性能产生影响的提醒事件
实例因底层升级需迁移事件	<ul style="list-style-type: none">升级或改造基础设施时，可能影响对应实例，提醒根据系统事件引导进行实例迁移	实例费用事件	<ul style="list-style-type: none">影响费用的事件，例如实例到期、账号欠费，导致实例即将被停止或释放
实例安全事件	<ul style="list-style-type: none">影响实例安全的事件，例如实例遭遇DDoS攻击或进入黑洞，导致实例安全受到威胁	状态变化事件	<ul style="list-style-type: none">实例的生命周期或其他状态发生变化时发送的事件

稳定性重保体系

需求场景



重保体系



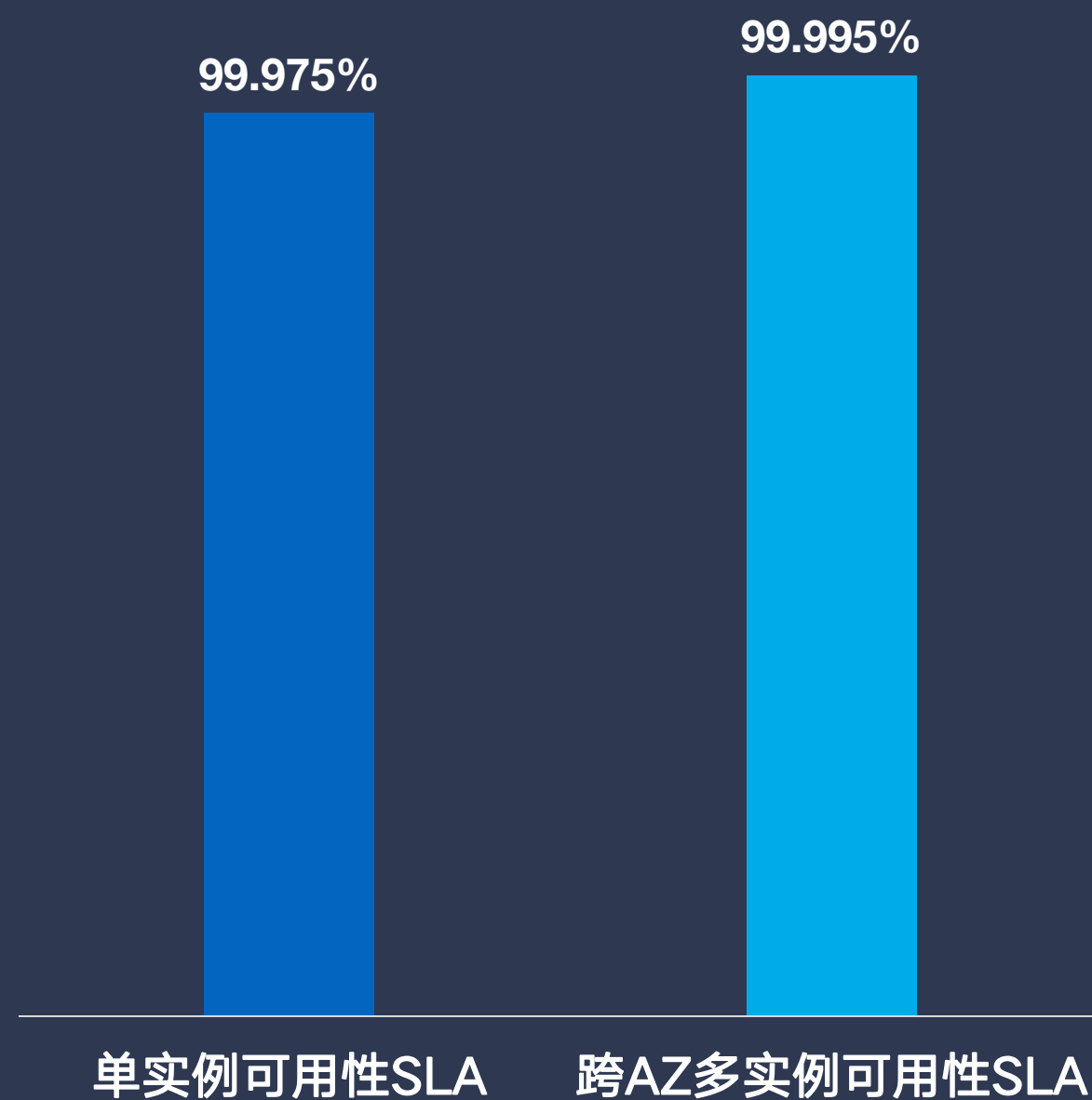
阿里云弹性计算云上运维最佳实践

ECS云上运维最佳实践概览



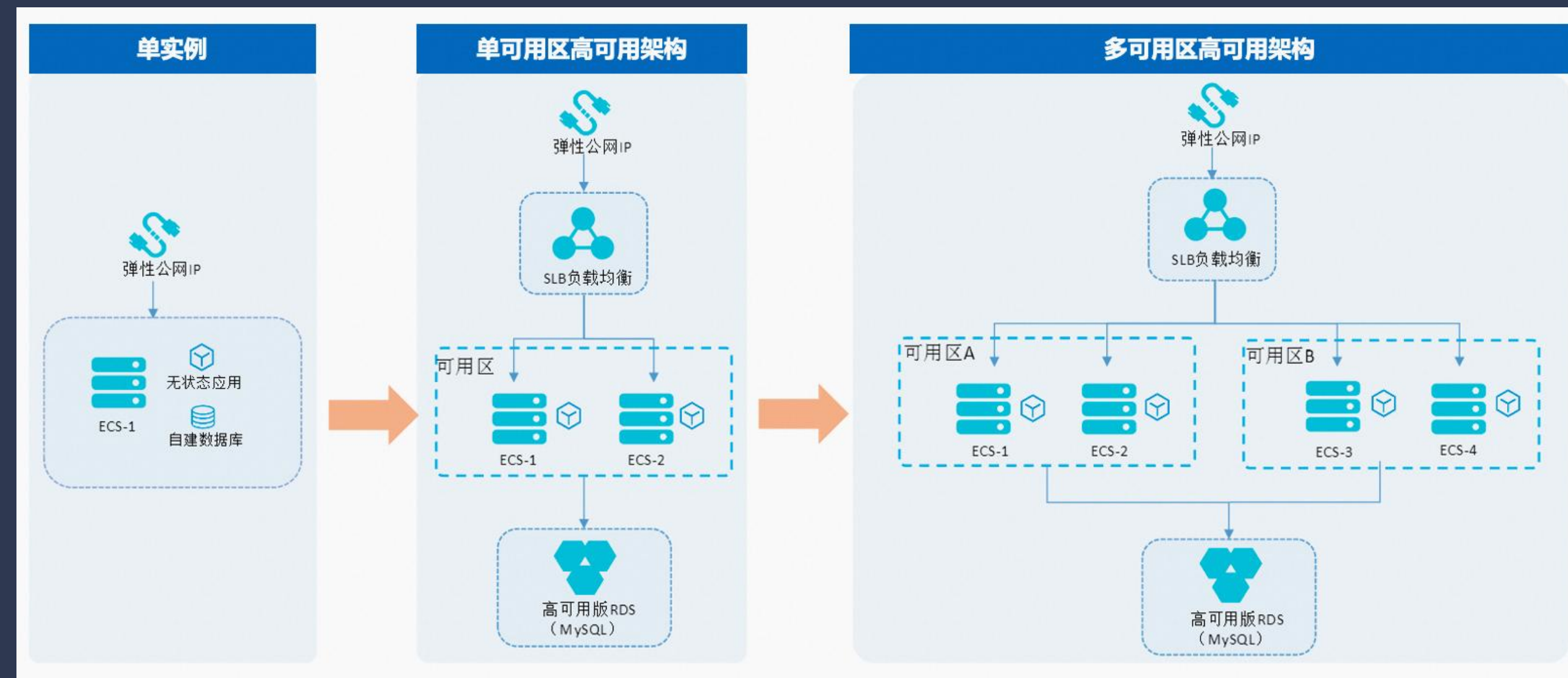
部署高可用架构，提升应用系统单实例宕机容错和极端场景容灾能力

阿里云ECS实例可用性SLA



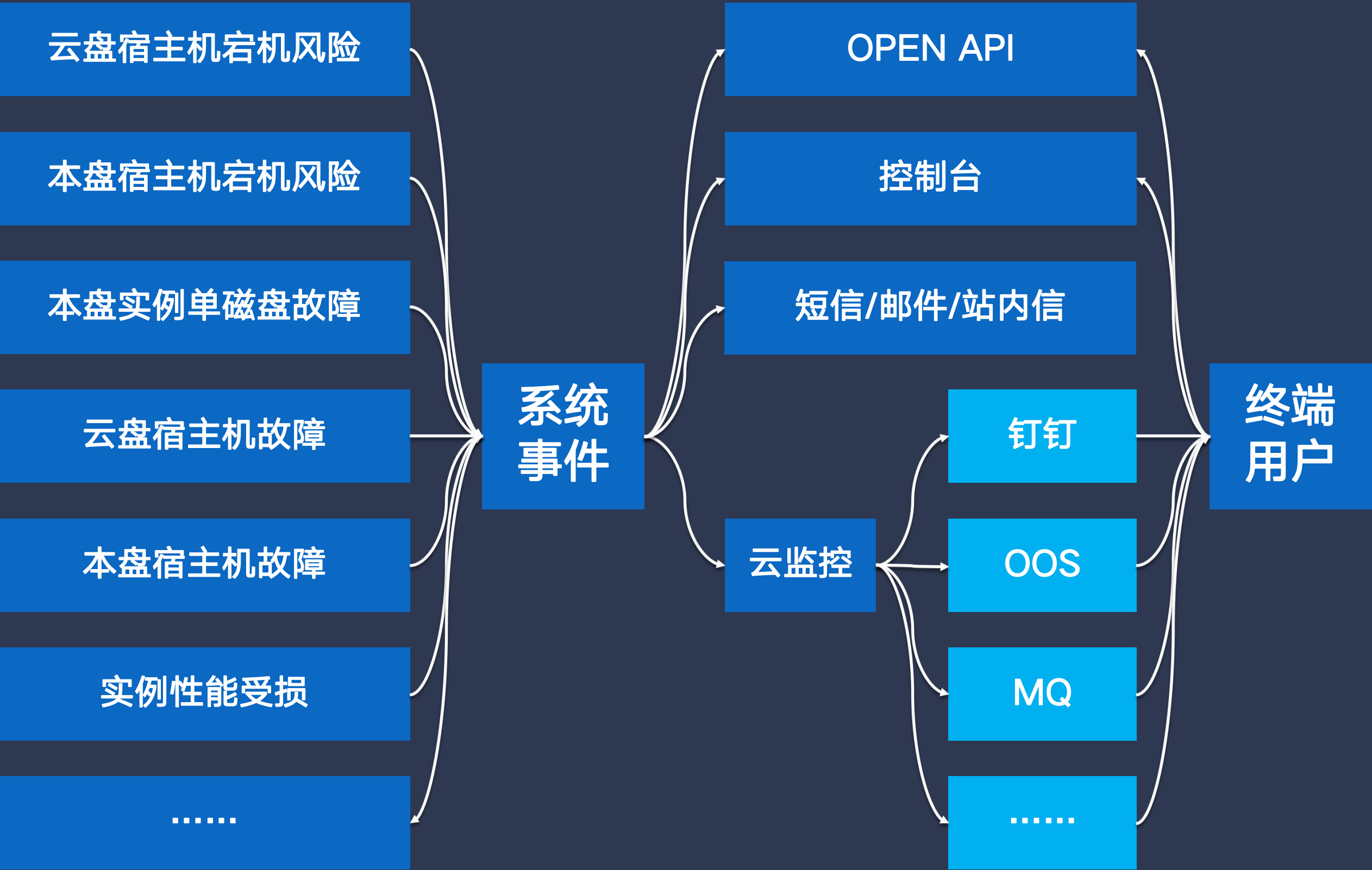
- 负载均衡SLB流量分发
- 容量规划N+1冗余

- 单元化部署，架构消除拧麻花
- 具备应用系统流量切换能力



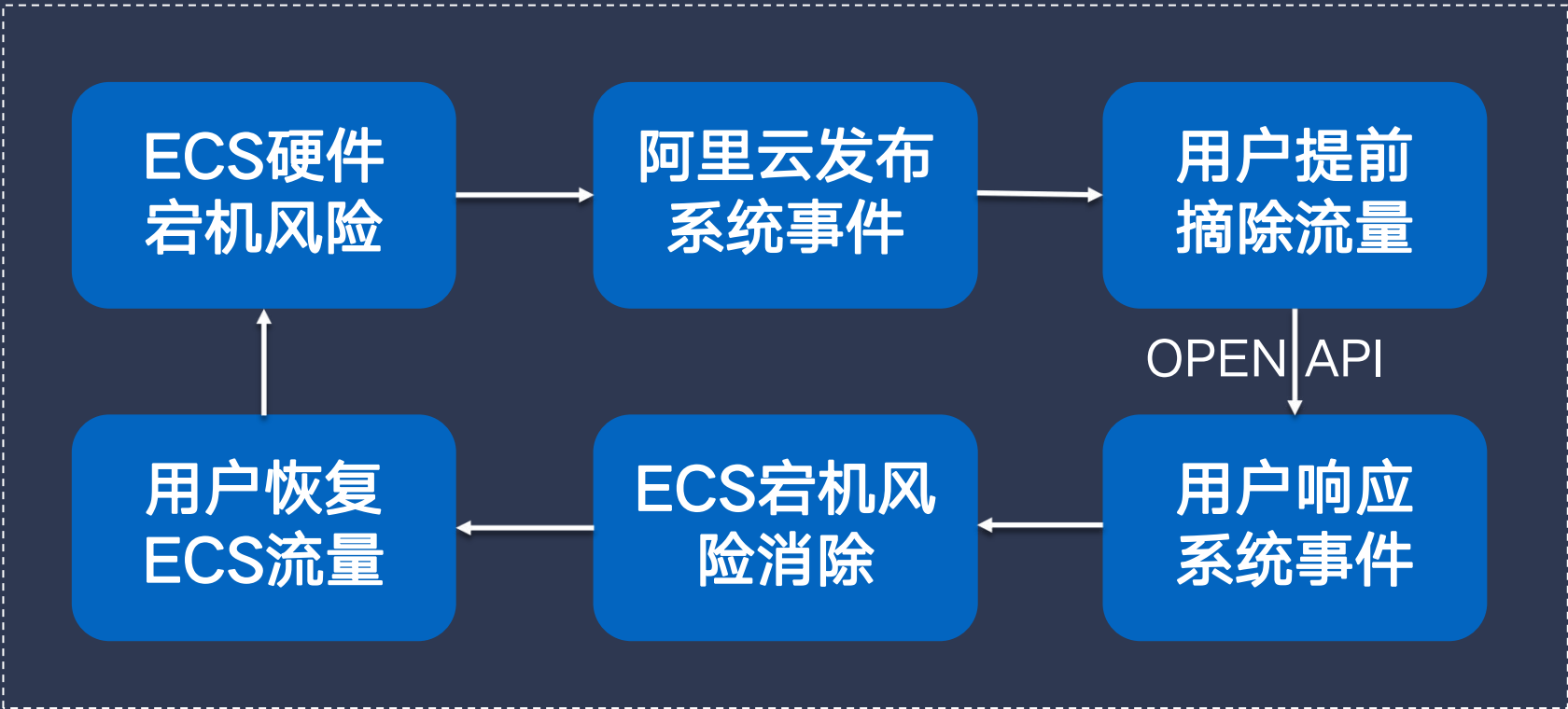
主动响应ECS系统事件，高效优雅规避实例运行风险

阿里云ECS事件体系

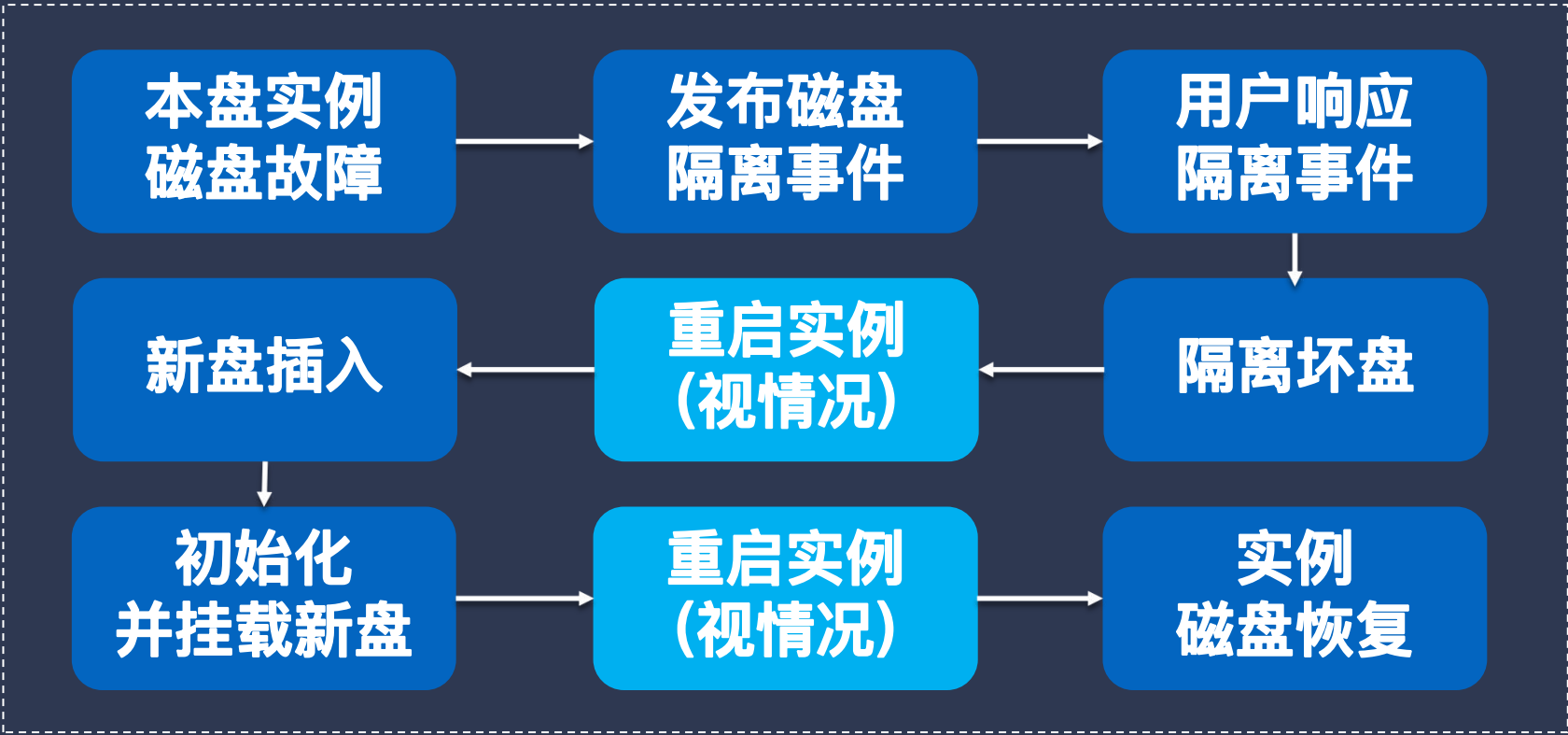


最佳实践典型示例

实例硬件
风险场景



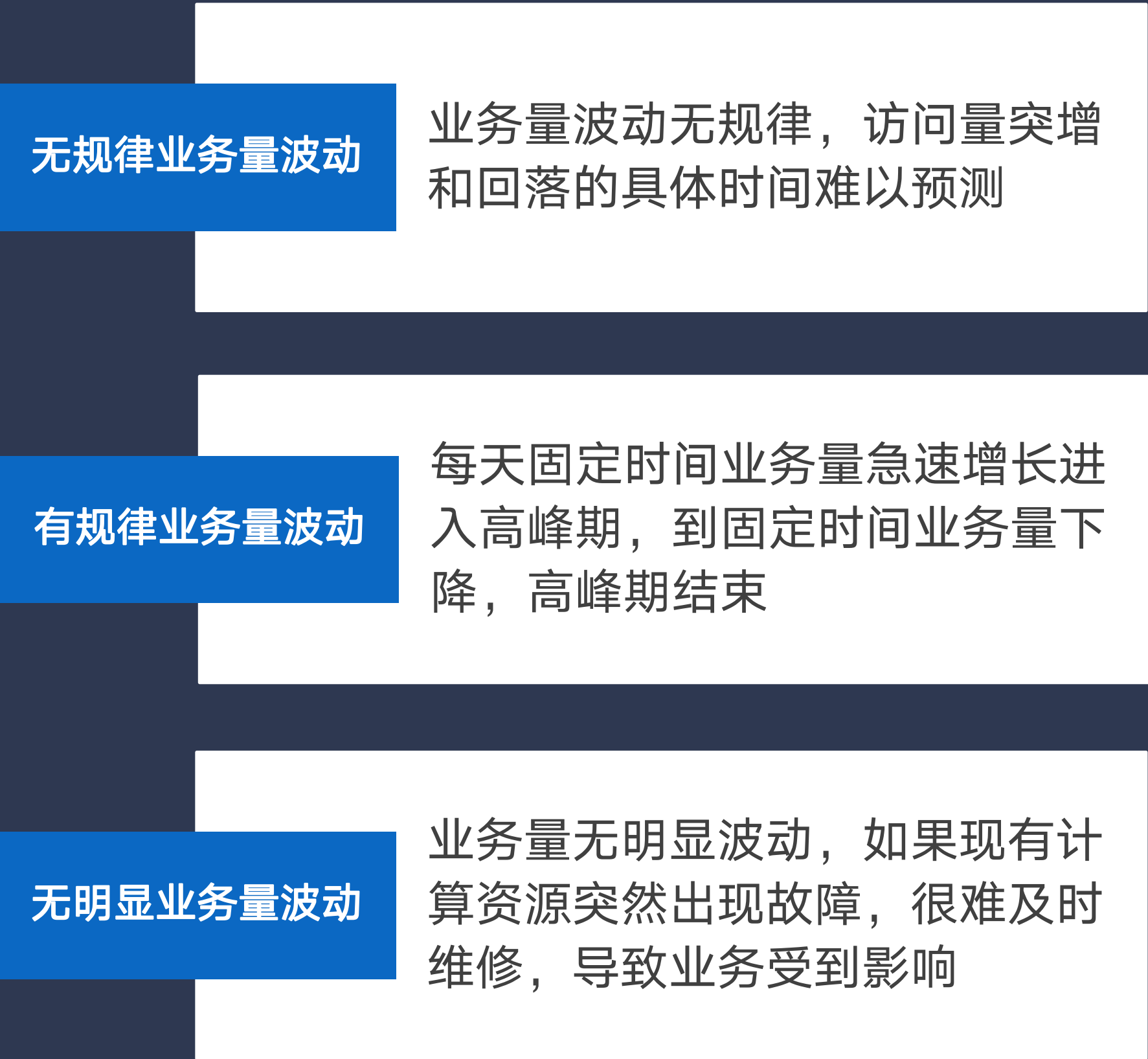
本地磁盘
故障场景



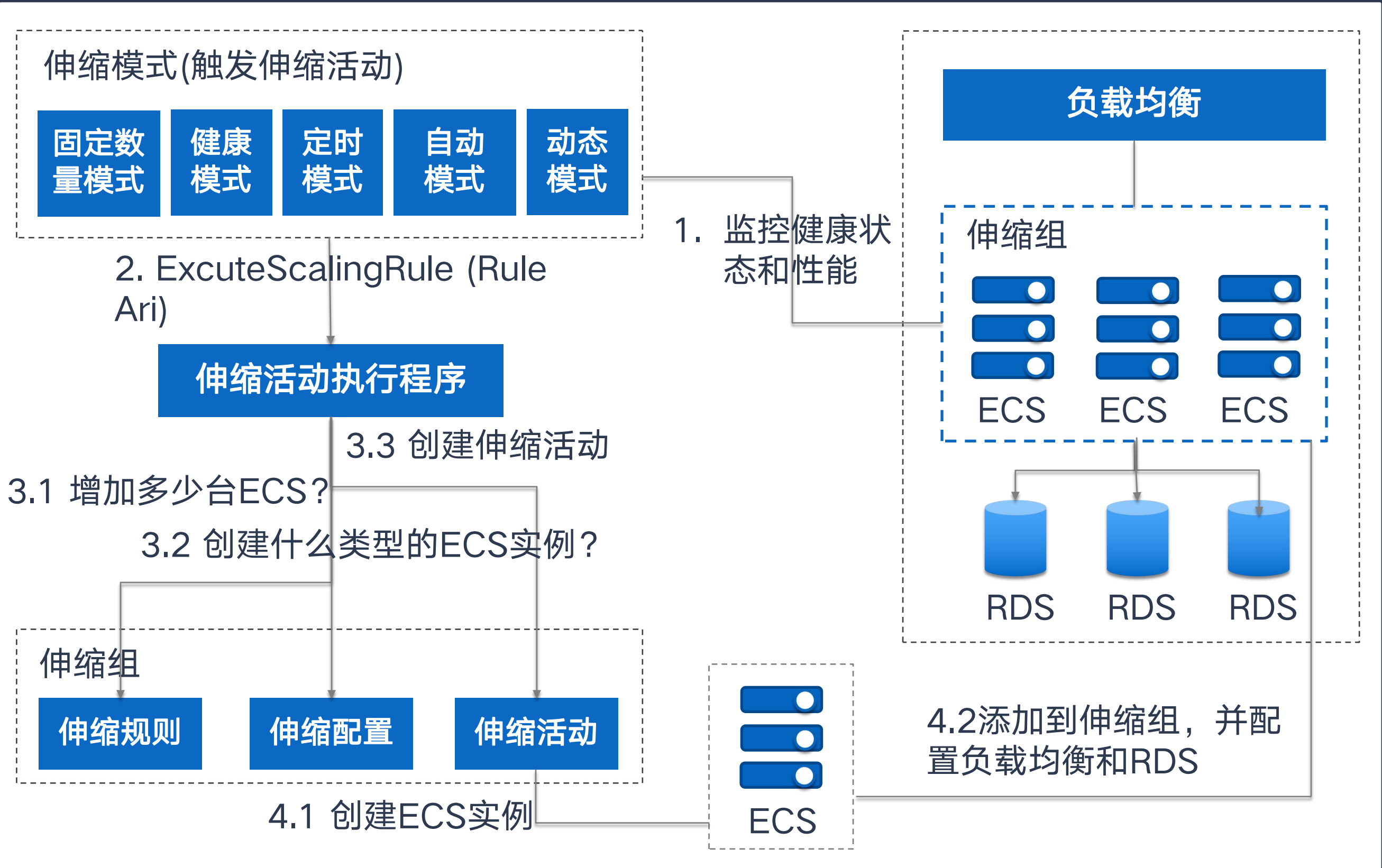
更多详见：[查询和响应ECS系统事件](#)

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



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

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？

实例健康诊断能力

<div>计算服务诊断</div> <div>诊断ECS实例底层的资源和虚拟化层</div>	<div>网络服务诊断</div> <div>检查实例内网络组件和外部网络环境异常</div>
<div>存储服务诊断</div> <div>检查实例磁盘和存储的运行状态</div>	<div>实例配置管理诊断</div> <div>检查实例启动或运行过程中，是否存在阻塞等操作</div>
<div>安全控制诊断</div> <div>检查实例关联的所有安全组相关状态</div>	<div>费用类诊断</div> <div>检查实例本身和实例关联组件的费用情况</div>
<div>资源配额诊断</div> <div>检查核心资源使用量是否即将达到配额</div>	<div>GuestOS内配置诊断</div> <div>检查实例OS内系统文件、关键进程等状态</div>

使用方式



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

Q&A

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