开放原子开源基金会
OPENATOM FOUNDATION

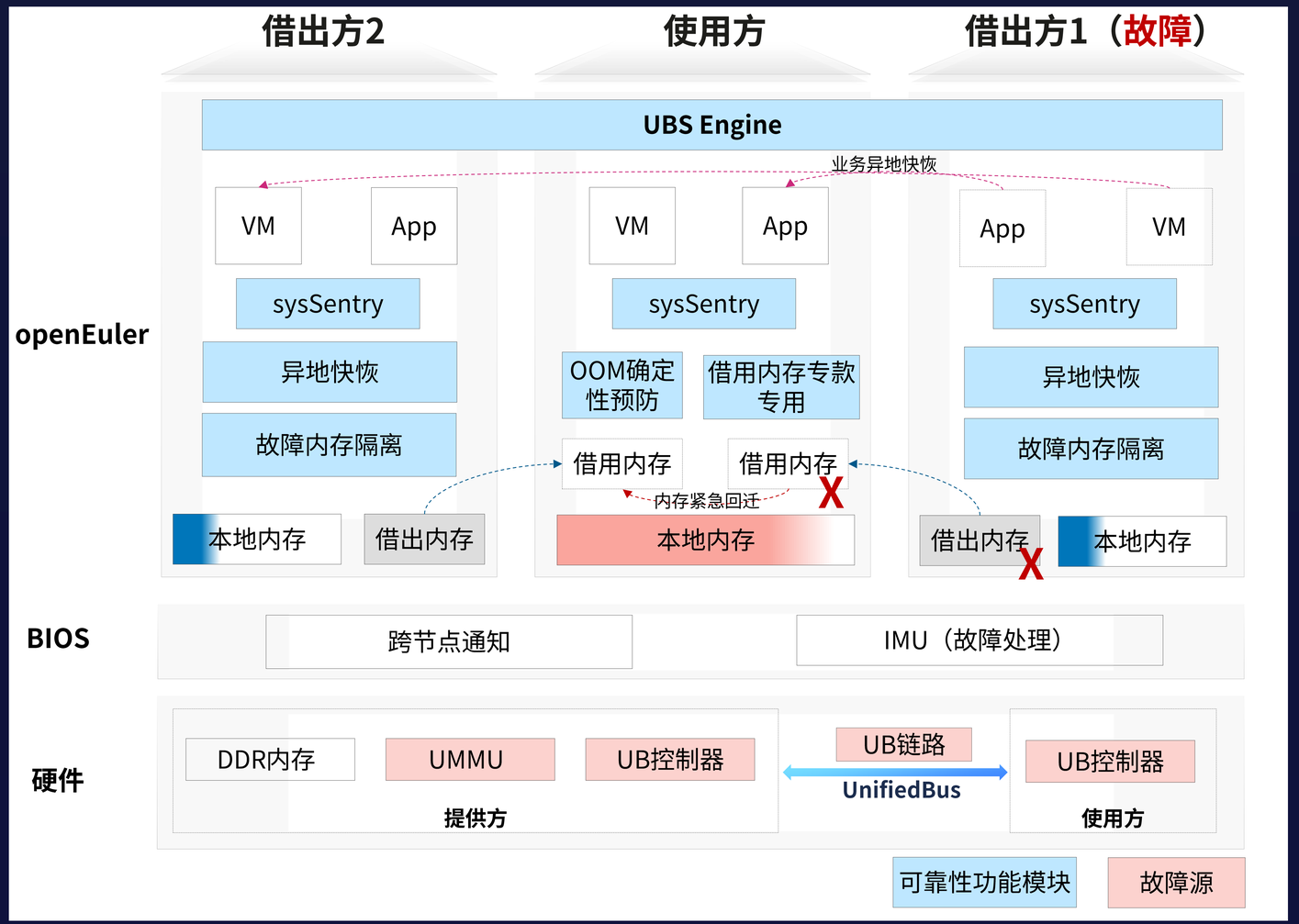
OpenEuler

超节点可靠性关键技术

演讲人
邓广兴

演讲人职位
openEuler社区Maintainer

灵衢可靠性：软硬配合消减故障扩散风险



方案说明

- **业务异地快恢**：节点故障时，通过灵衢高速互联总线进行业务快速迁移
- **借用内存专款专用**：防止故障影响内核造成系统panic
- **内存紧急回迁**：借出方panic触发内存紧急回迁，不影响业务正常运行
- **OOM确定性预防**：本地内存不足触发内存紧急借用，防止业务降级或被OOM kill

- **跨节点通知**：BIOS提供跨节点通知能力，故障事件快速通知
- **故障中断定制**：上送OS进行处理，避免过度panic

- 硬件层面新增故障源导致可靠性下降
- UB控制器超时代答，避免SEI故障触发系统panic

UB控制器&链路故障

UB控制器&链路故障类型

故障说明

- **新增故障模式：**新增UB控制器、链路、UMMU等期间故障点
- **故障域扩散：**超节点中资源提供方故障，可扩散至使用方，部分故障影响集群可用性

故障处理措施

序号	借用资源故障模式	故障分类	故障影响	目标	解决措施
⑤	•端口Down •链路损坏	UB链路故障	**集群跨节点无法访问**	故障不扩散	•硬件链路冗余 •**远端内存访问故障隔离**
⑤	UB链路降Lane	UB链路故障（亚健康）	影响集群访存性能	业务性能无影响	•支持链路质量检测和故障隔离 •业务处理亚健康场景
①	借出方内存颗粒不可纠正错误	借出方内存颗粒故障	**单点系统panic 业务终止**	远端内存故障主机不宕机	•**远端内存访问故障隔离** •内存故障预测，虚机热迁移 •**借用内存专款专用，防止系统挂死** •**内存故障精细化处理，特定场景免挂死**
②③⑥	•UMMU不可纠正错误 •借出方/使用方UB控制器不可纠正错误	UB控制器/UMMU故障			
④	•OS Panic •节点异常掉电/整机宕机(通过CPLD拉电重启) •节点计划内重启(升级重启/节点重启按钮重启)	整机故障	**借用节点业务终止**	业务无影响	•**Panic/Reboot场景借用内存数据回迁**

内存不足场景可靠性提升

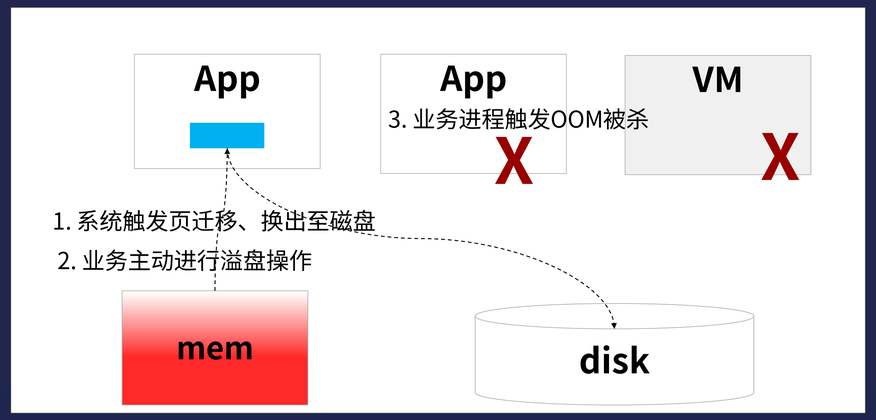
单机

关键挑战

- 传统计算节点存算固定配比，无法适应突发及峰值业务
- 故障内存隔离导致系统内存总量下降；内存替换平均需要24H，需保证期间业务可用性

影响

- 操作系统迁移回收内存，换出至磁盘 -> **业务性能下降**
- 操作系统阻塞业务内存申请，同步整理内存 -> **业务阻塞**
- 操作系统OOM killer杀死业务进程 -> **业务异常终止**



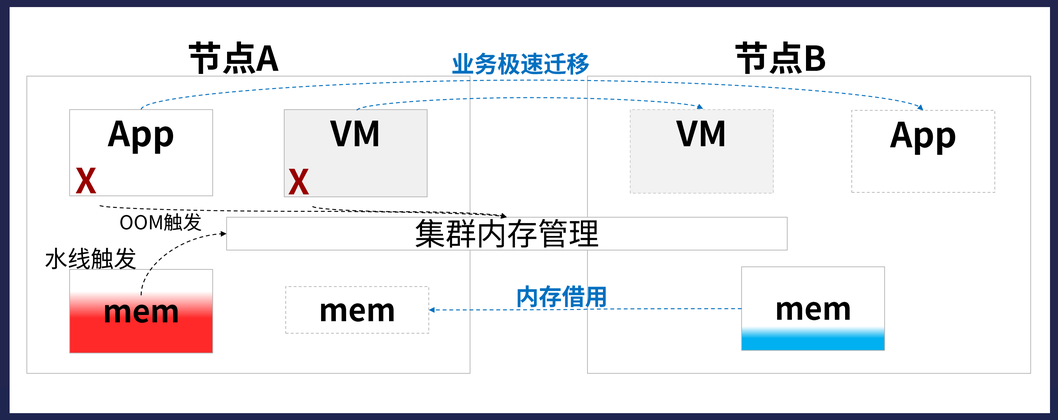
超节点

消减措施

- **内存域扩大至超节点**，消除单节点业务突发的内存影响

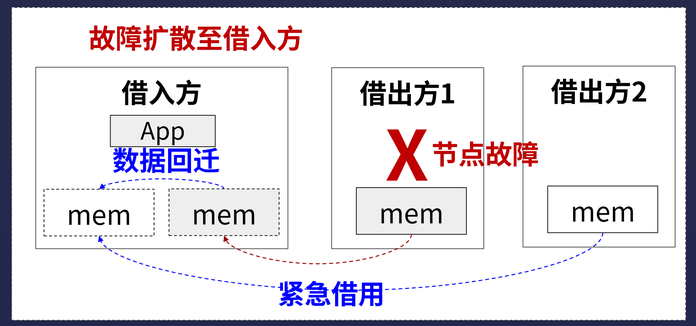
关键技术

- **OOM确定性预防：**进程及虚机OOM前触发紧急内存借用，避免OOM kill；OOM事件快速上报处理（亚秒级），借用内存快速上线（10ms/GB）
- **极速热迁移：**内存不足时触发虚机及业务迁移，避免影响性能；基于灵衢高速互联总线达成迁移中断时间<50ms
- **主动内存借用：**根据系统内存水线，主动触发内存借用，内存冷热搬移保障业务性能，业务性能下降小于10%



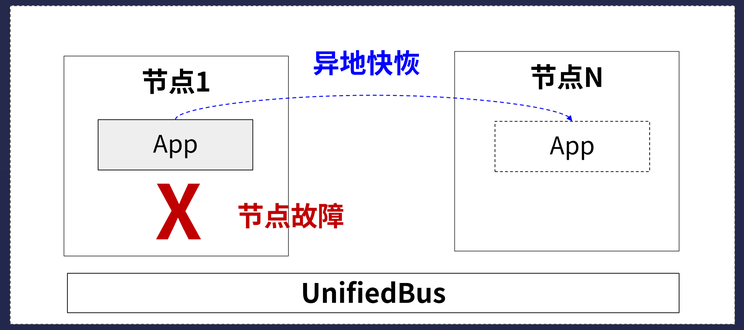
OS Panic/重启场景可靠性消减

超节点场景1：借出方OS Panic/Reboot



- **故障影响：**借出方panic/reboot影响借入方节点的业务由单点故障升级为节点故障。
- **可靠性消减：**借出节点故障后通过上报处理，借入方执行数据回迁将故障内存数据回迁至本地内存或紧急借用内存，消除故障扩散

超节点场景2：非借出方OS Panic/Reboot



- **故障影响：**与普通服务器场景一致，影响本节点的所有业务。
- **可靠性消减：**通过虚机&容器异地快恢将业务极速迁移到其它节点。

关键技术

- 故障快速通知：Panic及reboot信号捕获及快速通知，高性能内存极速决策调度（<3ms）,触发故障处理流程
- 借用内存快速上线：通过numa预上线及高位预留，加速内存上线时间 (<10ms/GB)
- 数据紧急回迁：panic/reboot触发借用内存高速搬移数据(8GB/s)，数据搬移过程应用无感
- 虚机异地快恢：通过池化内存及预创建达成panic场景虚机极速迁移，业务恢复时间小于1s
- 容器异地快恢：通过池化内存及迭代迁移方式，实现进程快速异地迁移恢复

异地快恢(1/2)：资源池化下的虚机故障秒级恢复

问题场景

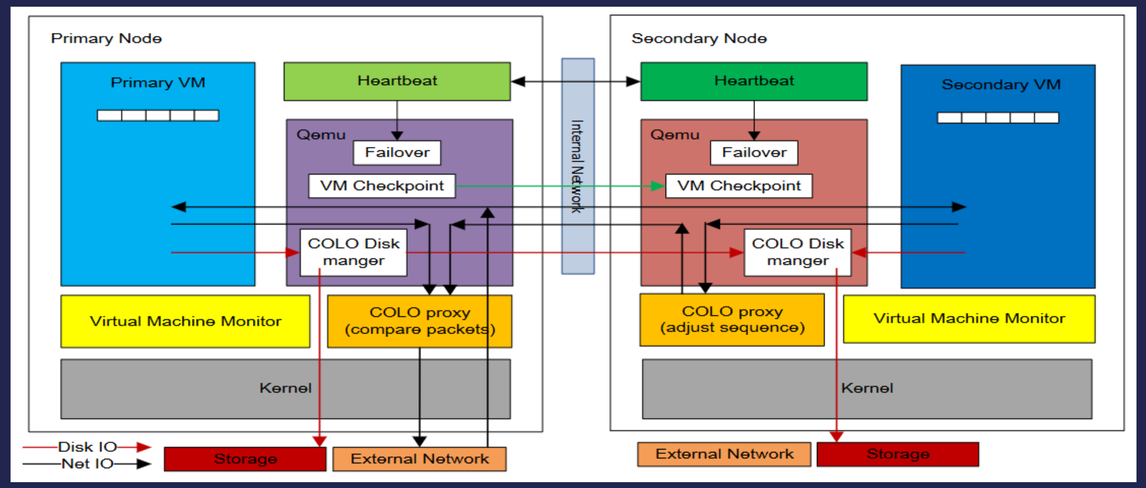
因软硬件故障（如服务器系统Panic）导致业务虚拟机意外宕机，传统HA方案恢复耗时较长（3~5min），严重影响客户业务，影响SLA指标。

故障检测 → 调度新主机 → 挂载共享卷 → 启动虚拟机 → 恢复业务

技术现状

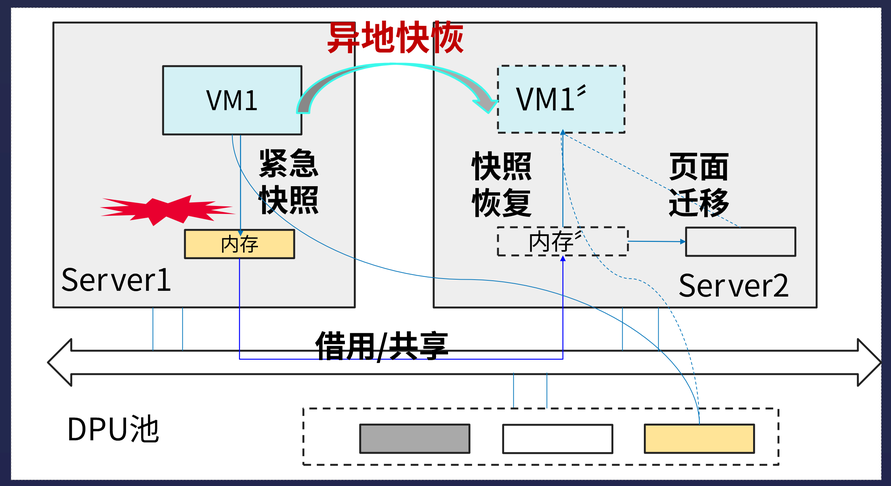
业界普遍采用主动容错，资源和业务性能限制大，以Qemu Colo为例：

- **主动同步**：虚拟机运行时时刻需要进行主备同步，性能受损严重；
- **资源冗余**：双主机同时运行主备虚拟机，内存和CPU资源需要双份；
- **设备限制**：VCPU数量限制（不超过8个），不支持直通设备等。



解决方案

基于灵衢资源池化的优势，探索新型被动容错机制，破除主动容错限制的同时，实现秒级虚拟机故障异地快恢。

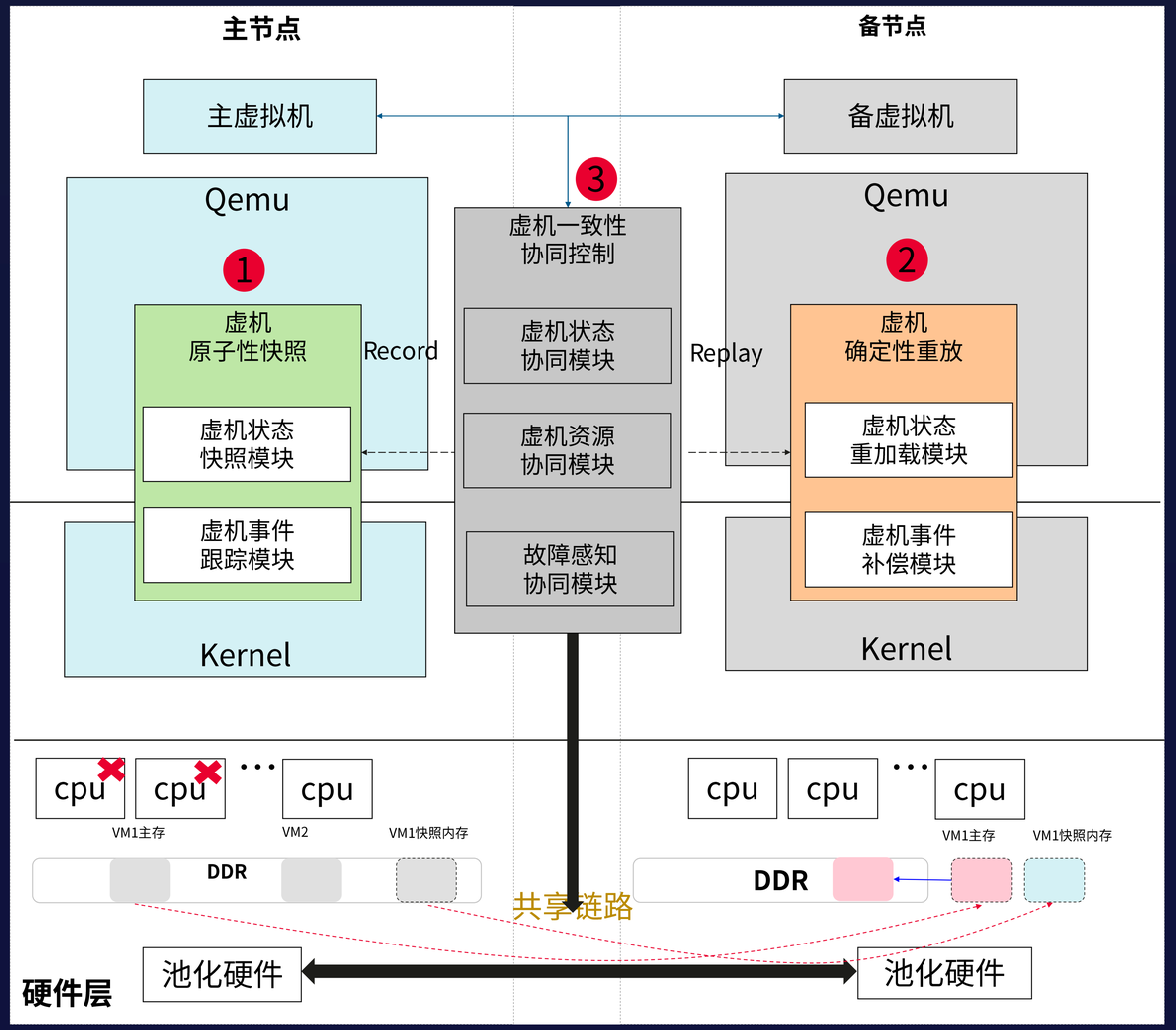


技术指标	主动容错技术	被动容错技术
资源约束	冗余备份	无
虚拟机规格约束	<= 8 vcpu	无限制
性能受损程度	>20%	<5%
设备约束	仅支持模拟设备	无限制

- **挑战一**：随机故障瞬间如何保证关键状态数据的完整性？
- **挑战二**：故障恢复瞬间如何保证海量状态数据的准确性？

异地快恢(2/2)：资源池化下的虚机故障秒级恢复

架构设计



关键技术

① 虚机原子性快照

设置CheckPoint，对虚机运行关键状态（VCPU陷出事件、IO、网络包等）进行实时原子性快照，利用双缓存机制进行状态数据完整性保护，应对瞬时随机故障。利用池化内存进行状态数据同步。

② 虚机确定性重放

从池化内存恢复快照数据，将虚机运行状态回溯到上一个检查点，对可能丢失的In-Flight事件进行确定性重放，保证虚拟机状态一致性。

③ 虚机一致性协同控制

故障前在备节点预创建备虚拟机，主虚拟机状态发生变化时进行备虚拟机的状态同步。发生故障后构建基于UB链路的快速故障通知能力（ sysSentry ），实现故障通知时间<100ms。

验证效果&下一步计划

□ 验证结果

虚拟机业务恢复时间：0.788 秒

异地快恢执行时长：3.4 秒

□ 下一步计划

1、基于灵衢池化设备能力，支持更多类型设备（SPDK、直通设备等）的异地快恢能力；

2、进一步优化业务恢复时间；

开放原子开源基金会 OPENATOM FOUNDATION | OpenEuler

Thank You

开源计划

调度层 UBS Engine ③

功能层 sysSentry ② SMAP 异地快恢

Kernel ① UB-RAS APEI驱动 UBUS驱动

2025年11月首批开源

26年H2增强特性逐步开源

③ https://gitee.com/openeuler/ubs-engine

② https://gitee.com/openeuler/sysSentry

① https://gitee.com/openeuler/kernel