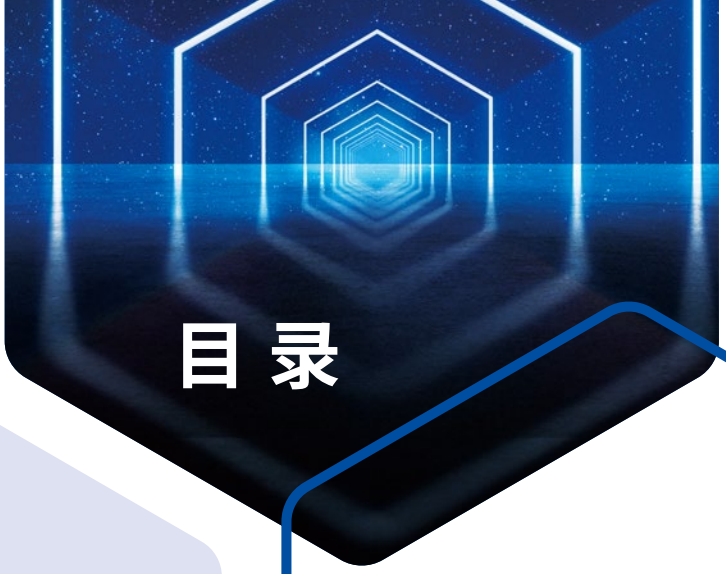
openEuler 22.03 LTS SP2

技术白皮书



目录

CONTENTS

概述 01

openEuler 社区，全称为 OpenAtom openEuler 社区，是一个面向数字基础设施操作系统的开源社区，简称 openEuler 或者 openEuler 社区。由开放原子开源基金会（以下简称“基金会”）孵化及运营。

openEuler 是一个面向数字基础设施的操作系统，支持服务器、云计算、边缘计算、嵌入式等应用场景，支持多样性计算，致力于提供安全、稳定、易用的操作系统。通过为应用提供确定性保障能力，支持 OT 领域应用及 OT 与 ICT 的融合。

openEuler 社区通过开放的社区形式与全球的开发者共同构建一个开放、多元和架构包容的软件生态体系，孵化支持多种处理器架构、覆盖数字基础设施全场景，推动企业数字基础设施软硬件、应用生态繁荣发展。

2019 年 12 月 31 日，面向多样性计算的操作系统开源社区 openEuler 正式成立。

2020 年 3 月 30 日，openEuler 20.03 LTS（Long Term Support，简写为 LTS，中文为长生命周期支持）版本正式发布，为 Linux 世界带来一个全新的具备独立技术演进能力的 Linux 发行版。

2020 年 9 月 30 日，首个 openEuler 20.09 创新版发布，该版本是 openEuler 社区中的多个企业、团队、独立开发者协同开发的成果，在 openEuler 社区的发展进程中具有里程碑式的意义，也是中国开源历史上的标志性事件。

2021 年 3 月 31 日，发布 openEuler 21.03 内核创新版，该版本将内核升级到 5.10, 还在内核方向实现内核热升级、内存分级扩展等多个创新特性，加速提升多核性能，构筑千核运算能力。

2021 年 9 月 30 日，全新 openEuler 21.09 创新版如期而至，这是 openEuler 全新发布后的第一个社区版本，实现了全场景支持。增强服务器和云计算的特性，发布面向云原生的业务混部 CPU 调度算法、容器化操作系统 KubeOS 等关键技术；同时发布边缘和嵌入式版本。

2022 年 3 月 30 日，基于统一的 5.10 内核，发布面向服务器、云计算、边缘计算、嵌入式的全场景 openEuler 22.03 LTS 版本，聚焦算力释放，持续提升资源利用率，打造全场景协同的数字基础设施操作系统。

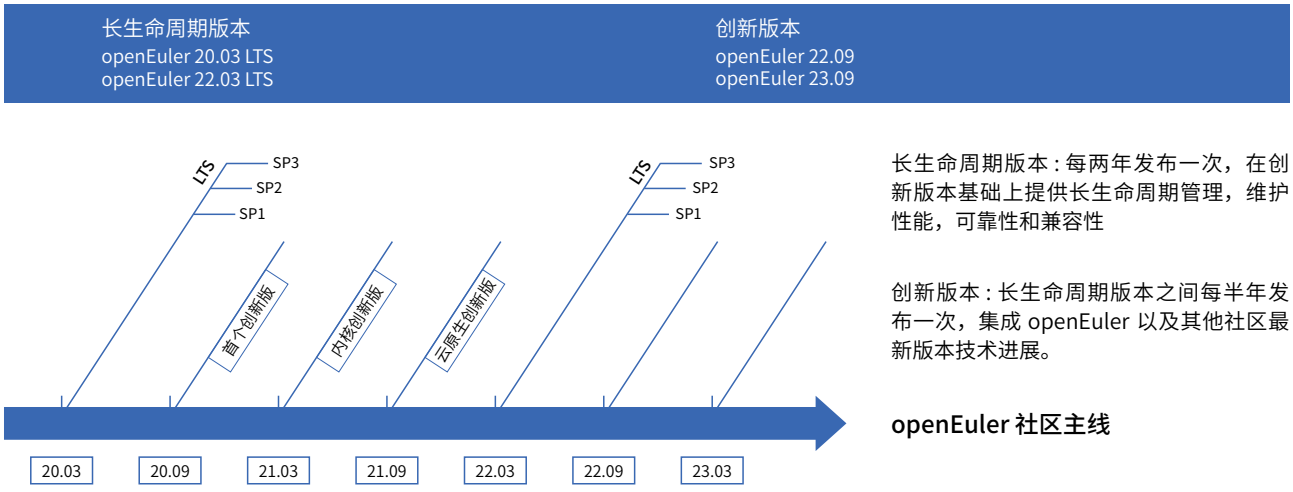
2022 年 9 月 30 日，发布 openEuler 22.09 创新版本，持续补齐全场景的支持。

2022 年 12 月 30 日，发布 openEuler 22.03 LTS SP1 版本，打造最佳迁移工具实现业务无感迁移，性能持续领先。

2023 年 3 月 30 日，发布 openEuler 23.03 内核创新版本，采用 Linux Kernel 6.1 内核，为未来 openEuler 长生命周期版本采用 6.x 内核提前进行技术探索，方便开发者进行硬件适配、基础技术创新绝上层应用创新。

2023 年 6 月 30 日，发布 openEuler 22.03 LTS SP2 版本，场景化竞争力特性增强，性能持续提升。

openEuler 版本管理



长生命周期版本：每两年发布一次，在创新版本基础上提供长生命周期管理，维护性能，可靠性和兼容性

创新版本：长生命周期版本之间每半年发布一次，集成 openEuler 以及其他社区最新版本技术进展。

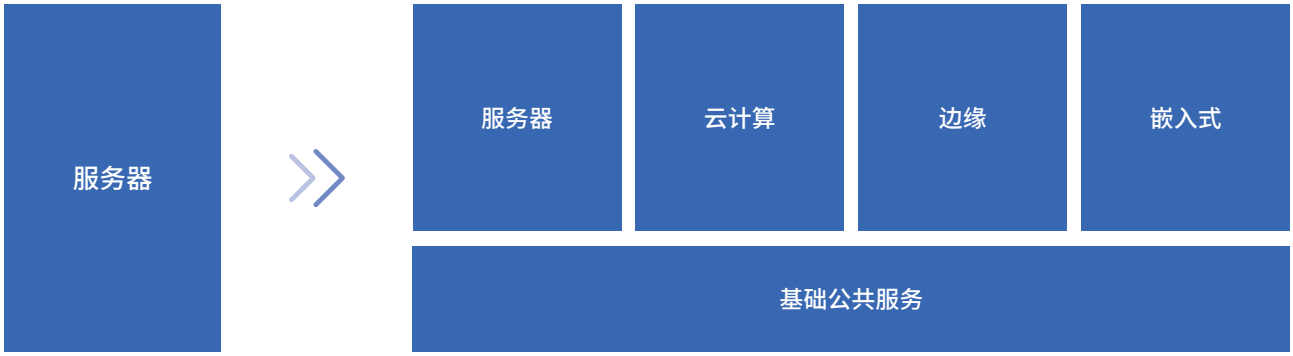
openEuler 社区主线

openEuler 作为一个操作系统发行版平台，每两年推出一个 LTS 版本。该版本为企业级用户提供一个安全稳定可靠的操作系统。

openEuler 也是一个技术孵化器。通过每半年发布一个创新版，快速集成 openEuler 以及其他社区的最新技术成果，将社区验证成熟的特性逐步回合到发行版中。这些新特性以单个开源项目的方式存在于社区，方便开发者获得源代码，也方便其他开源社区使用。

社区中的最新技术成果持续合入社区发行版，社区发行版通过用户反馈反哺技术，激发社区创新活力，从而不断孵化新技术。发行版平台和技术孵化器互相促进、互相推动、牵引版本持续演进。

openEuler 覆盖全场景的创新平台



openEuler 已支持 X86、ARM、SW64、RISC-V、LoongArch 多处理器架构，逐步扩展 PowerPC 等更多芯片架构支持，持续完善多样性算力生态体验。

openEuler 社区面向场景化的 SIG 不断组建，推动 openEuler 应用边界从最初的服务器场景，逐步拓展到云计算、边缘计算、嵌入式等更多场景。openEuler 正成为覆盖数字基础设施全场景的操作系统，新增发布面向边缘计算的版本 openEuler Edge、面向嵌入式的版本 openEuler Embedded。

openEuler 希望与广大生态伙伴、用户、开发者一起，通过联合创新、社区共建，不断增强场景化能力，最终实现统一操作系统支持多设备，应用一次开发覆盖全场景。

openEuler 开放透明的开源软件供应链管理

开源操作系统的构建过程，也是供应链聚合优化的过程。拥有可靠开源软件供应链，是大规模商用操作系统的基础。openEuler 从用户场景出发，回溯梳理相应的软件依赖关系，理清所有软件包的上游社区地址、源码和上游对应验证。完成构建验证、分发、实现生命周期管理。开源软件的构建、运行依赖关系、上游社区，三者之前形成闭环且完整透明的软件供应链管理。

02 平台架构

系统框架

openEuler 是覆盖全场景的创新平台，在引领内核创新，夯实云化基座的基础上，面向计算架构互联总线、存储介质发展新趋势，创新分布式、实时加速引擎和基础服务，结合边缘、嵌入式领域竞争力探索，打造全场景协同的面向数字基础设施的开源操作系统。

openEuler 22.03 LTS SP2 发布面向服务器、云原生、边缘和嵌入式场景的全场景操作系统版本，统一基于 Linux Kernel 5.10 构建，对外接口遵循 POSIX 标准，具备天然协同基础。同时 openEuler 22.03 LTS SP2 版本集成分布式软总线、KubeEdge+ 边云协同框架等能力，进一步提升数字基础设施协同能力，构建万物互联的基础。

面向未来，社区将持续创新、社区共建、繁荣生态，夯实数字基座。

夯实云化基座

- 容器操作系统 KubeOS：云原生场景，实现 OS 容器化部署、运维，提供与业务容器一致的基于 K8S 的管理体验。
- 安全容器方案：iSulad+shimv2+StratoVirt 安全容器方案，相比传统 Docker+Qemu 方案，底噪和启动时间优化 40%。
- 双平面部署工具 eggo：ARM/X86 双平面混合集群 OS 高效一键式安装，百节点部署时间 <15min。

新场景

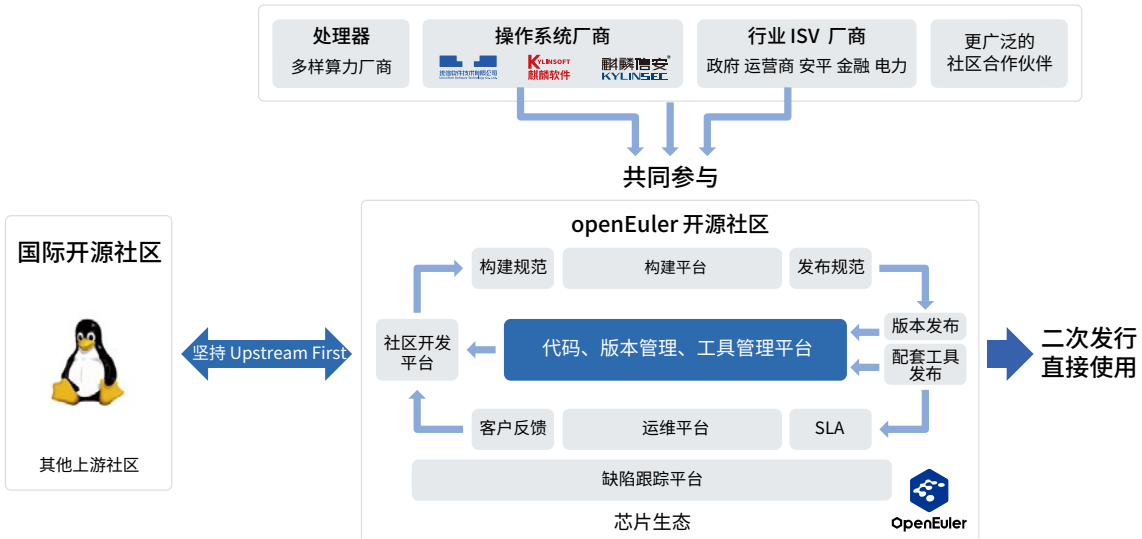
- 边缘计算：发布面向边缘计算场景的版本 openEuler 22.03 LTS SP2 Edge，支持 KubeEdge+ 边云协同框架，具备边云应用统一管理和发放等基础能力。
- 嵌入式： 发布面向嵌入式领域的版本 openEuler 22.03 LTS SP2 Embedded，镜像大小 <5M，启动时间 <5s。

繁荣社区生态

- 友好桌面环境：UKUI、DDE 、Xfce、Kiran-desktop、GNOME 桌面环境，丰富社区桌面环境生态。
- 欧拉 DevKit：支持操作系统迁移、兼容性评估、简化安全配置 secPaver 等更多开发工具。

平台框架

openEuler 社区与上下游生态建立连接，构建多样性的社区合作伙伴和协作模式，共同推进版本演进。



硬件支持

openEuler 社区当前已与多个设备厂商建立丰富的南向生态，Intel、AMD 等主流芯片厂商的加入和参与，openEuler 全版本支持 X86、ARM、申威、龙芯、RISC-V 五种架构，并支持多款 CPU 芯片，包括龙芯 3 号 、兆芯 开先 / 开胜系列、Intel IceLake/ Sapphire Rapids、AMD EPYC Milan /Genoa 等芯片系列，支持多个硬件厂商发布的多款整机型号、板卡型号，支持网卡、RAID、FC、GPU&AI、DPU、SSD、安全卡七种类型的板卡，具备良好的兼容性。

支持的 CPU 架构如下：

硬件类型	X86	ARM
CPU	Intel、AMD、兆芯、海光	鲲鹏、飞腾

支持的整机如下：

硬件类型	X86	ARM
整机	Intel：超聚变、超微 AMD：超微 海光：曙光 / 中科可控 兆芯：兆芯	鲲鹏：泰山 飞腾：青松、宝德

支持的板卡类型如下：

硬件类型	X86	ARM
网卡	华为、Mellanox、Intel、星云智联、云芯智联	华为、Mellanox、Intel、星云智联、云芯智联
Raid	Avago、云芯智联	Avago、云芯智联
FC	Marvell、Qlogic、Emulex	Marvell、Qlogic、Emulex
GPU&AI	Nvidia	Nvidia
SSD	华为	华为

全版本支持的硬件型号可在兼容性网站查询：https://www.openeuler.org/zh/compatibility/

03 运行环境

服务器

若需要在物理机环境上安装 openEuler 操作系统，则物理机硬件需要满足以下兼容性和最小硬件要求。

硬件兼容支持请查看 openEuler 兼容性列表：https://openeuler.org/zh/compatibility/。

部件名称	最小硬件要求
架构	ARM64、x86_64
内存	为了获得更好的体验，建议不小于 4GB
硬盘	为了获得更好的体验，建议不小于 20GB

虚拟机

openEuler 安装时，应注意虚拟机的兼容性问题，当前已测试可以兼容的虚拟机及组件列表如下：

- centos-7.9 qemu 1.5.3-175.el7 libvirt 4.5.0-36.el7 virt-manager 1.5.0-7.el7
- centos-8 qemu 2.12.0-65.module_el8.0.0+189+f9babebb.5 libvirt 4.5.0-24.3.model_el8.0.0+189+f9babebb virt-manager 2.0.0-5.el8
- fedora 32 qemu 4.2.0-7.fc32 libvirt 6.1.0-2.fc32 virt-manager 2.2.1-3.fc32
- fedora 35 qemu 6.1.0-5.fc35 libvirt 7.6.0-3.fc35 virt-manager 3.2.0-4.fc35

部件名称	最小虚拟化空间要求
架构	ARM64、x86_64
CPU	2 个 CPU
内存	为了获得更好的体验，建议不小于 4GB
硬盘	为了获得更好的体验，建议不小于 20GB

边缘设备

若需要在边缘设备环境上安装 openEuler 操作系统，则边缘设备硬件需要满足以下兼容性和最小硬件要求。

部件名称	最小硬件要求
架构	ARM64、x86_64
内存	为了获得更好的体验，建议不小于 4GB
硬盘	为了获得更好的体验，建议不小于 20GB

嵌入式

若需要在嵌入式环境上安装 openEuler 操作系统，则嵌入式硬件需要满足以下兼容性和最小硬件要求。

部件名称	最小硬件要求
架构	ARM64、ARM32
内存	为了获得更好的体验，建议不小于 512MB
硬盘	为了获得更好的体验，建议不小于 256MB

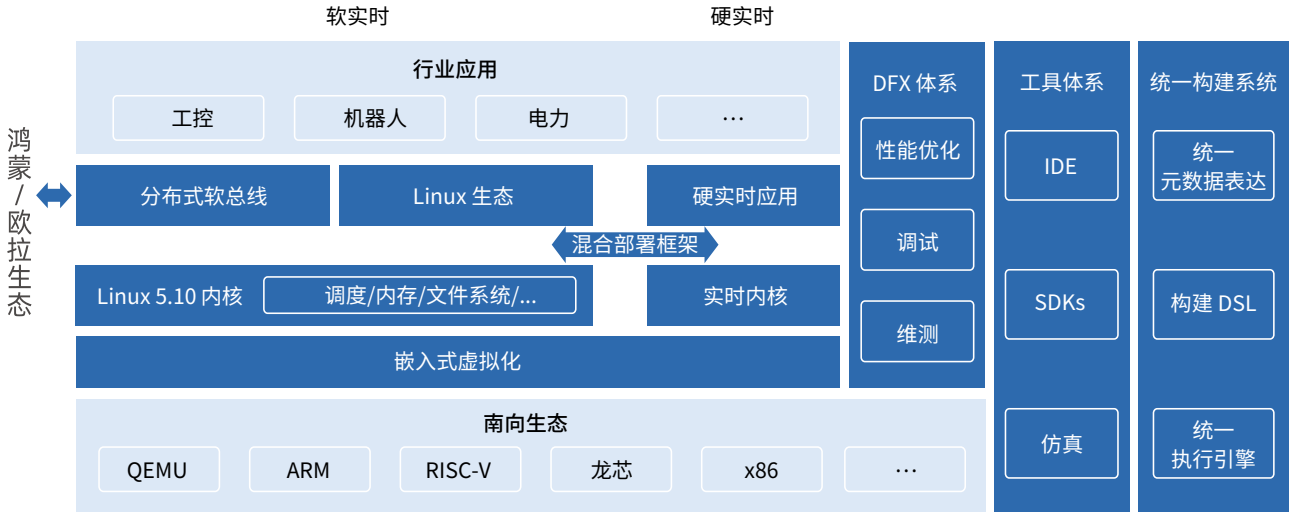
04 场景创新

嵌入式

openEuler发布面向嵌入式领域的版本openEuler 22.03 LTS SP2 Embedded，提供更加丰富的嵌入式软件包构建能力，支持实时 / 非实时平面混合关键部署，并集成分布式软总线。

openEuler Embedded 围绕工业和机器人领域持续深耕，通过行业项目垂直打通，不断完善和丰富嵌入式技术栈和生态。openEuler 22.03 LTS SP2 Embedded 支持嵌入式虚拟化弹性底座，提供 Jailhouse 虚拟化方案、openAMP 轻量化混合部署方案，用户可以根据自己的使用场景选择最优的部署方案。同时支持 ROS humble 版本，集成 ros-core、ros-base、SLAM 等核心软件包，满足 ROS2 运行时要求。未来 openEuler Embedded 将协同 openEuler 社区生态伙伴、用户、开发者，逐步扩展支持 RISC-V、龙芯等芯片架构，丰富工业中间件、ROS 中间件、仿真系统等能力，打造嵌入式领域操作系统解决方案。

功能描述



版本功能如下：

1. 轻量化能力：开放 yocto 小型化构建裁剪框架，支撑 OS 镜像轻量化定制，提供 OS 镜像 <5M，以及 <5s 快速启动等能力。
2. 多硬件支持：支持树莓派、X86、Hi3093、RK3568 作为嵌入式场景通用硬件。
3. 软实时内核：基于 Linux 5.10 内核提供软实时能力，软实时中断响应时延微秒级。
4. 嵌入式弹性虚拟化底座：提供多种虚拟化方案，满足用户不同硬件和业务场景需要：
   - baremetal：基于 openAMP 实现裸金属混合部署方案，支持外设分区管理，性能最好，但隔离性和灵活性较差。目前支持 UniProton/Zephyr/RT-Thread 和 openEuler 嵌入式 Linux 混合部署。

- 分区虚拟化：基于 Jailhouse 实现工业级硬件分区虚拟化方案，性能和隔离性较好，但灵活性较差。目前支持 FreeRTOS 和 openEuler 嵌入式 Linux 混合部署。
- 实时虚拟化：openEuler 社区自研虚拟化方案，兼顾性能、隔离性和灵活性，综合最优。目前支持 Zephyr 和 openEuler 嵌入式 Linux 混合部署。

5. 嵌入式软件包支持：350+ 嵌入式领域常用软件包的构建。
6. 硬实时内核（UniProton）：支持 POSIX 接口（103 个），上下文切换时延 3us、中断延迟 2us。

未来还将提供：

- 南向生态：RISC-V、龙芯支持。
- 混合关键性部署框架：围绕生命周期管理、跨 OS 通信、服务化框架、多 OS 协同构建 4 个方面持续打造标准混部框架，支持更多的软实时和硬实时 OS 接入 openEuler 生态中。
- 嵌入式弹性底座：持续完善 Jailhouse 和 ZVM 虚拟化能力，支持更广泛的南向生态，提供更好的时延优化。
- 硬实时（UniProton）中间件：提供丰富的 POSIX 接口支持和常用中间件，方便用户应 用开发和迁移。
- 泛工业泛嵌入式通用接口：围绕 RTOS 领域极致性能场景，定义高性能 API，为北向应用提供统一的接口。
- 行业安全认证：联合伙伴逐步支持面向行业安全认证，如面向 IEC61508、CC EAL 等。

应用场景

嵌入式系统可广泛应用于工业控制、机器人控制、电力控制、航空航天、汽车及医疗等领域。

05 内核创新

openEuler 内核中的新特性

openEuler 22.03 LTS SP2 基于 Linux Kernel 5.10 内核构建，在此基础上，同时吸收了社区高版本的有益特性及社区创新特性：

- SMT 驱离优先级反转特性：解决混部 SMT 驱离特性的优先级反转问题，减少离线任务对在线任务 QoS 的影响。
- CPU QoS 优先级负载均衡特性：在线、离线混部 CPU QoS 隔离增强 , 支持多核 CPU QoS 负载均衡，进一步降低离线业务 QoS 干扰。
- 潮汐 affinity 调度特性：感知业务负载动态调整业务 CPU 亲和性，当业务负载低时使用 prefered cpus 处理，增强资源的局部性；当业务负载高时，突破 preferred cpus 范围限制，通过增加 CPU 核的供给提高业务的 QoS。
- 支持进程、容器级别 KSM 使能：KSM 即 Kernel Same page Merge, 在引入本特性之前，KSM 的使用需要用户态程序显式调用 madvise 来指定参与去重的内存地址范围，而一些非 C 语言写的程序也无法调用 madvise 去做去重。本特性新增了 2 个功能方便程序使用 KSM 而无需显式调用 madvise：
  1. 进程粒度支持全范围去重：新增一个 prctl 系统调用接口，作为进程使能 KSM 的开关， 调用该接口可以使进程内所有地址范围的内存（私有匿名页）参与 KSM 去重，fork 后的进程也继承该去重方式。相比之下，进程只需要调用一次 prctl 接口，即可使能全范围 KSM 去重，而无需多次调用 madvise 分别指定去重地址范围。
  2. 容器粒度支持全范围去重：在 memory cgroup v1 目录下新增容器粒度的开关 memory.ksm，写 1 后该容器内所有进程都使能 KSM 全地址范围内存去重。
- Damon（Data Access MONitoring）特性增强：Damon 可在轻度内存压力下，实现主动、轻量级的线上内存访问监控及回收，用户根据监控结果定制策略对内存区域做相应操作。
- uswap 特性增强：增加用户态换出内存页面的机制，支持用户态灵活换出内存到后端存储，节省内存。
- Intel EMR（Emerald Rapids）新平台支持：EMR 是 Intel 下一代基于 Intel 7 制程的新一代 CPU 平台，性能更高，在增强已有硬件特性的同时，提供了诸如 TDX 等新的硬件特性。openEuler 提供 EMR 平台的支持，对关键应用以及客户计算平台的持续升级至关重要。
- 支持 ACPI for Arm64 MPAM 2.0：MPAM（Memory system component Partioning and Monitoring）是 Arm Architecture v8.4 的 Extension 特性。用于解决服务器系统中，混部不同类型业务时，由于共享资源的竞争（Cache，DMC，Interconnect），而带来的某些关键应用性能下降或者系统整体性能下降的问题。

SMT 驱离防止优先级反转特性

目前云场景中，在线业务与离线业务混合部署提升资源利用率的同时，如何保证在线业务的 QoS 是当前亟需解决的问题。在开启 SMT 场景中，同时运行在同一个物理核上的在离线业务与在线业务之间存在干扰。针对这一诉求，设计混部 SMT 驱离方案，用于隔离离线任务对在线任务的 IPC 干扰。对于 CFS 任务运行策略的改变可能会带来优先级的问题，该特性解决了由于被驱离离线任务占用临界资源无法释放的问题。

功能描述

开启混部 SMT 驱离特性后，假设 CPUA 和 CPUB 互为 SMT 核，将在线任务绑在 CPUA 上，离线任务绑在 CPUB 上。CPUA 上的在线任务长时间 100% 占用 CPU 资源，则 CPUB 上的离线任务因为被驱离无法运行，无法释放临界资源。此时如果有高优先级任务等待离线任务占有的临界资源，就会出现优先级翻转现象。该特性通过检测离线任务被压制的运行时间，来判断系统是否处于优先级反转的风险状态中，来决定是否需要将离线任务解除压制直到释放内核中的临界资源。

提供用户可配置的两个接口：

1) /proc/sys/kernel/qos_overload_detect_period_ms

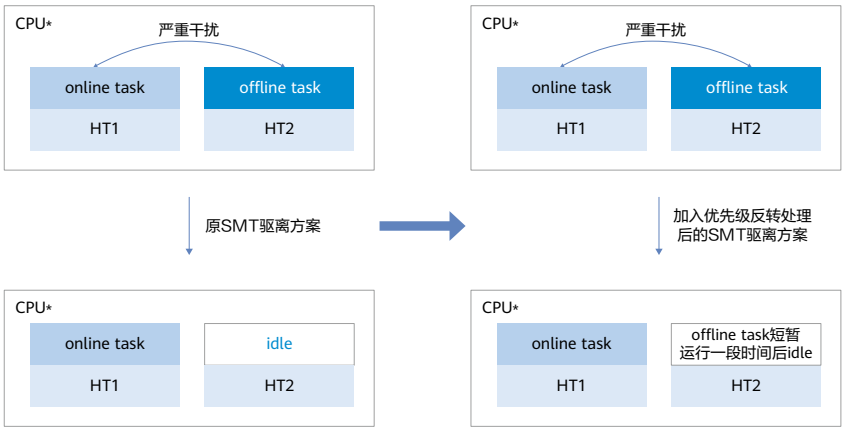
描述：非离线任务持续占有 CPU 超过此时间会触发优先级反转解决流程，单位 ms。

范围：100---100000

默认值：5000

配置建议：

- 过小，会频繁触发，对在线任务影响较大，容易存在误报。
- 过大，容易因为优先级反转问题导致系统卡住时间过长。



SMT 驱离优先级反转

2) /proc/sys/kernel/qos_offline_wait_interval_ms

描述：离线任务在超负载情况下返回用户态时，每次睡眠的时间，单位 ms。

范围：100—1000

默认值：100

配置建议：

- 过长，可能会导致在线任务停止运行后，离线任务处于 sleep 状态，CPU 一段时间内处于 idle 状态，降低 CPU 利用率。
- 过小，会导致离线任务频繁唤醒，干扰在线任务。

应用场景

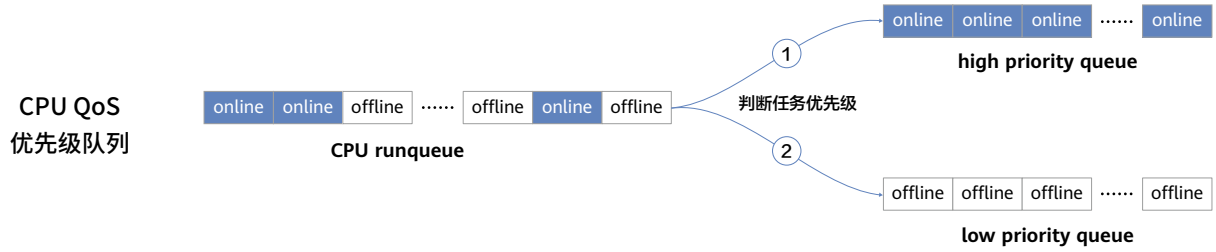
混部场景中，开启了 SMT 驱离离线任务特性，需要将 CONFIG_QOS_SCHED_SMT_EXPELLER 打开。

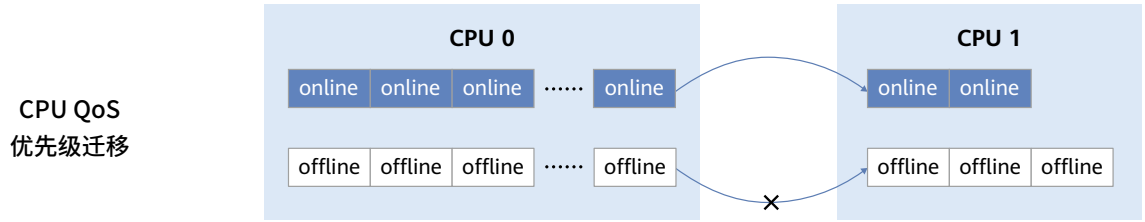
CPU QoS 优先级负载均衡特性

负载均衡 FIFO 任务迁移队列不区分优先级，无法解决跨核迁移抢占保障高优先级，特别是 CPU 敏感型任务的优先调度，针对在线、离线容器混部场景下，CFS 负载均衡需要提出一种优先级队列模型，支持高低优先级的 QoS 负载均衡，确保在线业务能更快得到调度和执行，最大化压制离线任务的 QoS 干扰，提高整机 CPU 资源利用率。

功能描述

实现一种 CFS 多优先级任务等待队列，在线任务和离线任务分别由不同优先级的 CFS 任务等待队列维护。多核 CPU 负载均衡时，优先从任务等待队列中选择高优先级任务，确保高优先级任务迁移优先得到调度；压制低优先级任务迁移，减少不必要的低优先级任务上下文切换、唤醒抢占等带来的 QoS 干扰及调度性能开销。





提供用户可配置的接口：/proc/sys/kernel/sched_prio_load_balance_enabled

描述 : 是否开启 CPU QoS 优先级负载均衡。

范围 : 0 和 1

默认值 : 0

应用场景

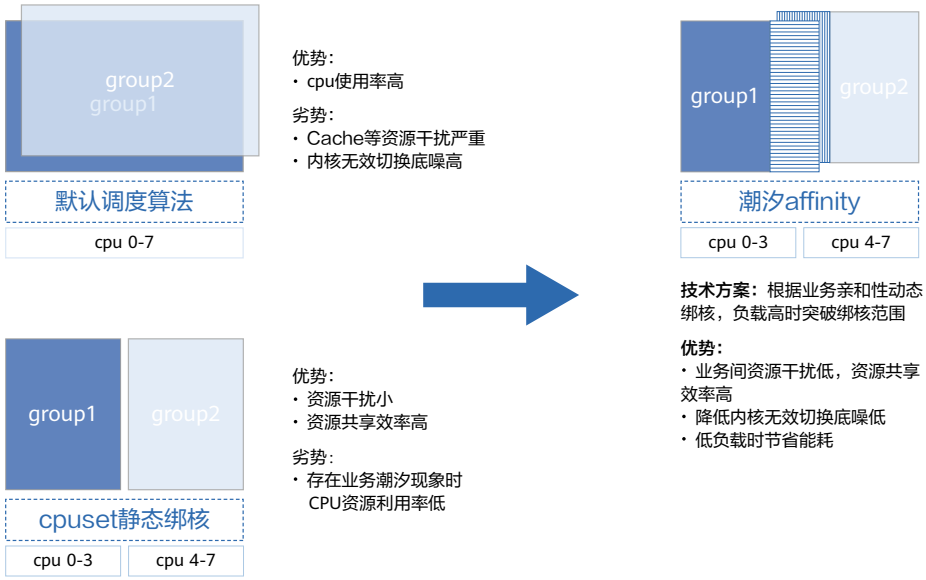
混部场景中，开启了 CPU QoS 优先级负载均衡特性，需要将 CONFIG_QOS_SCHED_PRIO_LB 打开。

潮汐 affinity 特性

当今服务器核数越来越多，为了充分利用 CPU 资源，很多业务混合部署在同一台机器上。一方面，不同业务混合部署，虽然能提高 CPU 利用率，但是也加剧了 cache 等资源的冲突。另一方面，对于相同业务，在满足 QoS 的前提下，可用 CPU 资源越多，由于迁移更加频繁，CPU 核 idle 切换变多，跨 NUMA 访存增加等问题，都会导致 CPU 使用效率下降。

功能描述

潮汐 affinity 技术通过感知业务负载变化，动态调整业务 CPU 资源的供给。具体来说，当业务负载低时，在满足 QoS 的基础上，使用更少的 CPU 资源，减少 CPU 迁移、idle 切换和 cache miss 等，提升 CPU 利用效率和能效比。当业务负载高时，通过分配更多 CPU 资源来提升 QoS 质量，提升 CPU 资源的利用率。



潮汐 affinity 技术动态感知业务的负载变化调节业务 CPU 范围，负载低时使用优先 CPU，增强资源局部性；负载高时突破优先 CPU 的限制，利用共享 CPU 来提升 QoS。

应用场景

在线 - 在线、在线 - 离线混部场景中，对绑核能带来性能提升；在由于负载动态变化无法准确绑核的业务场景中，利用潮汐 affinity 来提升性能和降低功耗是个不错的选择。

06 云化基座

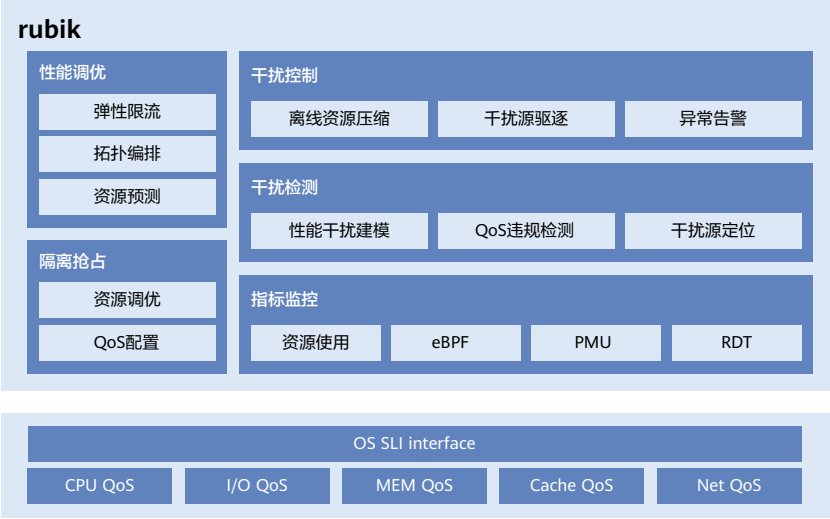
Rubik 混部支持精细化资源 QoS 感知和控制

云数据中心资源利用率低是行业普遍存在的问题，提升资源利用率已经成为了一个重要的技术课题。将业务区分优先级混合部署（简称混部）运行是典型有效的资源利用率提升手段。混部的核心技术是资源隔离控制。

功能描述

版本功能如下：

- 集群调度增强：增强 OpenStack Nova 能力，支持优先级语义调度。
- 功耗控制：通过对低优先级虚拟机的 CPU 带宽进行限制，以此达到降低整机功耗的同时保障高优先级虚拟机 QoS。
- Cache 及内存带宽控制：支持对低优先级虚拟机的 LLC 和内存带宽进行限制，当前仅支持静态分配。
- CPU 干扰控制：支持 CPU 时间片 us 级抢占及 SMT 干扰隔离，同时具有防优先级反转能力。
- memcg 异步内存回收：支持限制混部时离线应用使用的总内存，并在在线内存使用量增加时动态压缩离线业务内存使用。
- QuotaBurst 柔性限流：支持关键在线业务被 CPU 限流时允许短时间突破 limit 限制，保障在线业务运行的服务质量。
- PSI 指标观测增强：支持 cgroup v1 级别的压力信息统计，识别和量化资源竞争导致的业务中断风险，支撑用户实现硬件资源利用率提升。



应用场景

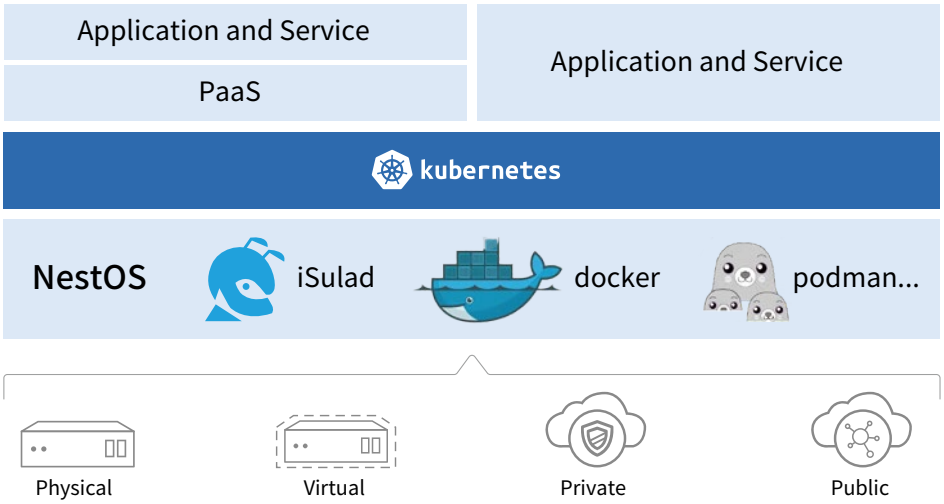
业务可根据时延敏感性分为高优先级业务和低优先级业务，将业务区分优先级混合部署以提高资源利用率。高优先级虚拟机业务推荐：时延敏感类业务，如 web 服务、高性能数据库、实时渲染、机器学习推理等。低优先级虚拟机业务推荐：非时延敏感类业务，如视频编码、大数据处理、离线渲染、机器学习训练等。

NestOS 容器操作系统

NestOS 是在 openEuler 社区孵化的云底座操作系统，集成了 rpm-ostree 支持、ignition 配置等技术，采用双根文件系统、原子化更新的设计思路，使用 nestos-assembler 快速集成构建。并针对 K8S、OpenStack 等平台进行适配，优化容器运行底噪，使系统具备十分便捷的集群组件能力，可以更安全的运行大规模的容器化工作负载。

功能描述

- 开箱即用的容器平台：NestOS 集成适配了 iSulad、Docker、Podman 等主流容器引擎，为用户提供轻量级、定制化的云场景 OS。
- 简单易用的配置过程：NestOS 通过 ignition 技术，可以以相同的配置方便地完成大批量集群节点的安装配置工作。
- 安全可靠的包管理：NestOS 使用 rpm-ostree 进行软件包管理，搭配 openEuler 软件包源，确保原子化更新的安全稳定状态。
- 友好可控的更新机制：NestOS 使用 zincati 提供自动更新服务，可实现节点自动更新与重新引导，实现集群节点有序升级而服务不中断。
- 紧密配合的双根文件系统：NestOS 采用双根文件系统的设计实现主备切换，确保 NestOS 运行期间的完整性与安全性。



应用场景

NestOS 适合作为以容器化应用为主的云场景基础运行环境，解决了在使用容器技术与容器编排技术实现业务发布、运维时与底层环境高度解耦而带来的运维技术栈不统一，运维平台重复建设等问题，保证了业务与底座操作系统运维的一致性。

07 特性增强

SysCare 热补丁能力

在 Linux 世界，有一个困扰大家已久的难题：如何在不影响业务的情况下，快速可靠地修复漏洞、解决故障。

当前常见的方法是采用热补丁技术：在业务运行过程中，对问题组件直接进行代码级修复，业务无感知。然而，当前热补丁制作方式复杂，补丁需要代码级匹配，且管理困难，特别是用户态组件面临文件形式、编程语言、编译方式、运行方式的多样性问题，当前还没有简便统一的补丁机制。

为了解决热补丁制作和管理的问题，SysCare 应运而生。

SysCare 是一个系统级热修复软件，为操作系统提供安全补丁和系统错误热修复能力，主机无需重新启动即可修复该系统问题。 SysCare 将内核态热补丁技术与用户态热补丁技术进行融合统一，用户仅需聚焦在自己核心业务中，系统修复问题交予 SysCare 进行处理。后期计划根据修复组件的不同，提供系统热升级技术，进一步解放运维用户提升运维效率。

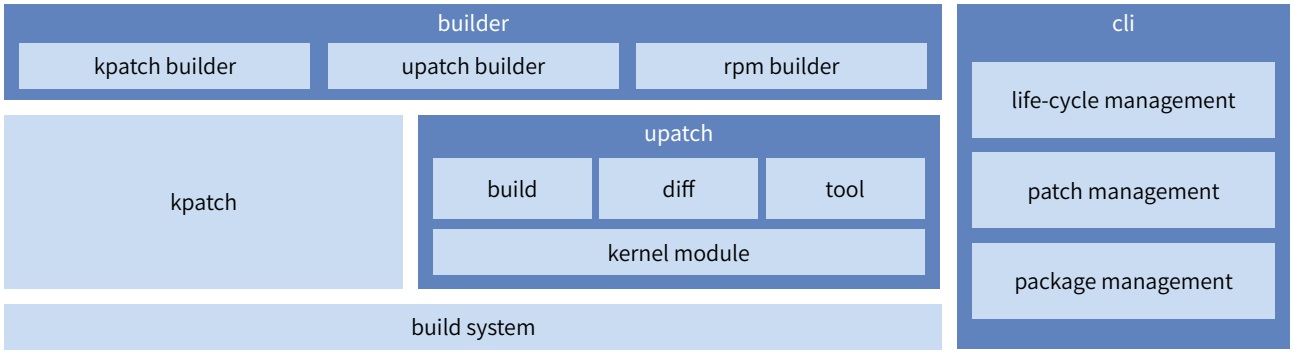
功能描述

1. 热补丁制作

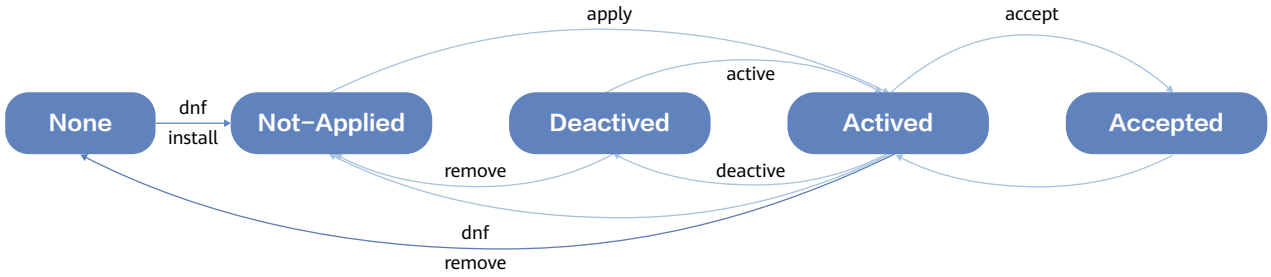
用户仅需输入目标软件的源码 RPM 包、调试信息 RPM 包与待打补丁的路径，无需对软件源码进行任何修改，即可生成对应的热补丁 RPM 包。

2. 热补丁生命周期管理

SysCare 提供一套完整的，傻瓜式补丁生命周期管理方式，旨在减少用户学习、使用成本，通过单条命令即可对热补丁进行管理。依托于 RPM 系统，SysCare 构建出的热补丁依赖关系完整，热补丁分发、安装、更新与卸载流程均无需进行特殊处理，可直接集成放入软件仓 repo。



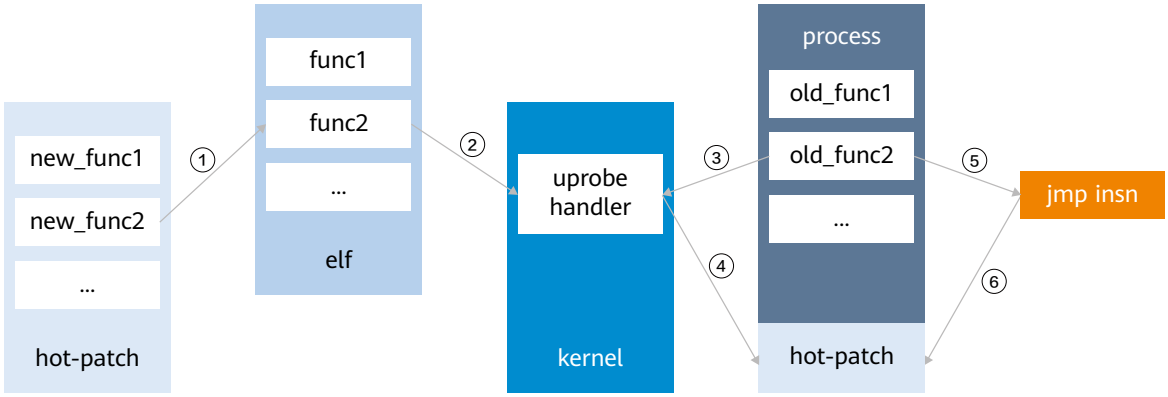
SysCare 整体架构



热补丁生命周期

3. 针对 ELF 文件（程序可执行文件）的用户态热补丁

使用 uprobe 技术，将热补丁与 ELF 文件绑定。在 ELF 文件运行时，通过 uprobe 触发补丁生效，这样无需监控进程。因此，无论用户进程是否已经运行都可以在打补丁后或新进程运行时使补丁生效。同时，该技术也可以给动态库打热补丁，解决了动态库热补丁的难题。补丁生效流程如下图所示。

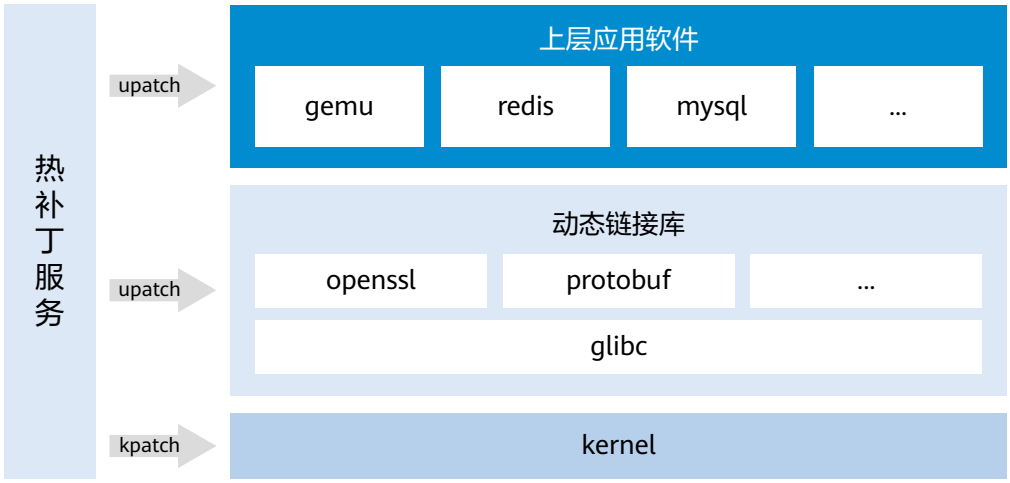


补丁生效流程

- 执行 uprobe 系统调用，在待修改函数 func 处增加 uprobe 断点。
- 注册 uprobe handler。
- 进程运行到 func 时调用 uprobe handler。
- 将 patch 映射到当前进程地址空间。
- 进行安全检查并将 func 的第一条指令修改为 jump 指令，指向 patch 地址。
- 跳转到 patch 地址执行。

4. 内核热补丁与用户态热补丁融合

SysCare 基于 upatch 和 kpatch 技术，覆盖应用、动态库、内核，自顶向下打通热补丁软件栈，提供用户无感知的全栈热修复能力。



热补丁应用范围

5. 新增特性

- 支持 aarch64 架构。
- 支持补丁制作参数自动推导。
- 支持补丁状态保存 / 恢复。
- 支持重启补丁状态恢复。
- 支持 syslog。

6. 约束限制

- 当前仅支持 64 位系统。
- 当前仅支持 ELF 格式的热修复，不支持解释型语言，不支持纯汇编修改。
- 当前仅支持 GCC / G++ 编译器，且不支持交叉编译。
- 暂不支持 LTO 优化。

应用场景

应用场景 1： CVE 补丁快速修复。

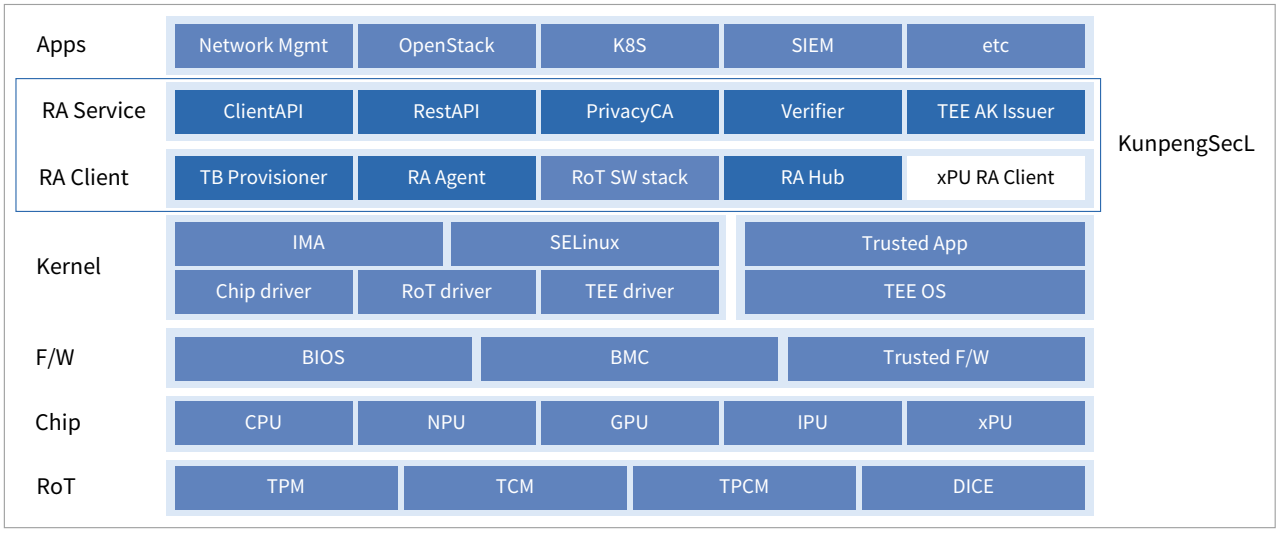
应用场景 2： 现网问题临时定位。

kunpengsecl 软件包支持平台和 TEE 远程证明

鲲鹏安全库（kunpengsecl）是开发运行在鲲鹏处理器上的基础安全软件组件，先期主要聚焦在远程证明等可信计算相关领域，使能社区安全开发者。

鲲鹏安全库的每个特性都可以由两大部分组成：组件和服务。组件部署在提供资源（计算、存储、网络）为用户运行工作负载的工作服务器节点上，将平台安全可信能力转化为软件接口，并将其提供给服务。服务则部署在专门的管理服务器节点上，汇聚来自所有工作服务器节点的安全可信能力，并将其提供给用户及其指定的管理工具以达成用户的对系统安全可信设计的具体要求。

鲲鹏安全库的首个安全特性就是远程证明，目的就是帮助用户获取工作服务器节点的软硬件可信状态，支持端到端的可信计算远程证明解决方案，让各种资源管理工具可以根据可信报告制定策略，对各种服务器资源进行差异化的调度和使用。



鲲鹏安全库的远程证明特性目前支持：

1. 基于 TPM 的通用平台远程证明。
2. 对鲲鹏服务器 TEE 的远程证明。

详情请参见项目 readme：https://gitee.com/openeuler/kunpengsecl/blob/master/README.md

功能描述

针对平台远程证明

- 远程证明服务 RAS 通过 PCA 来为工作服务器上的 TPM 提供远程证明密钥证书，通过 TrustMgr 管理工作服务器可信相关数据信息，通过 ClientAPI 来接收工作服务器的可信报告，通过 Verifier 验证目标的可信状态，通过 Cache 提供可信状态缓存服务，通过 Config 管理策略等配置信息，最终通过 RestAPI 向用户提供远程证明服务。
- 远程证明客户端 RAC 通过 TBProvisioner 解决部署阶段平台可信启动能力的检测和使能，通过 RAC Tools 来获取远程证明所需的各种数据信息，最后由 RA Agent 负责与 RAS 通信完成注册和可信报告发送。

- RAC Tools 负责屏蔽与可信模块及系统的交互细节，承担未来向不同可信模块扩展的责任。
- RA Hub 负责在需要时对本地域 RAC 进行通信汇聚和代理的功能，同时也将在未来负责提供 RAC 与 RAS 间通信通道适配的能力。

针对 TEE 远程证明

- TEE 证书服务 TAS（TEE AK Service）提供 TEE AK issuer 来使能用户部署自定义的证书服务为远程证明身份密钥颁发自定义证书。
- TEE 验证库（TEE verifier lib）以及验证示例 TEE Attester 为用户提供了获取 TEE 远程证明报告并进行验证的接口和工具。
- RAC 和 RAS 也同时将对服务器平台上的 TEE 环境的远程证明集成到了整体的远程证明架构中，通过 RAS 即可获取特定 TA 的可信状态。

应用场景

应用场景 1：可信云主机

通过云物理服务器的可信启动与平台远程证明的结合，对虚机运行的主机环境进行可信验证，为云主机用户提供安全可信的底层支持，同时借助虚机 vtpm 的特性完成对虚机可信启动和虚机远程证明的支持，进而使能可信云主机用户对可信云主机自身的安全可信状态进行直接感知，增强用户对云主机安全可信的信心。

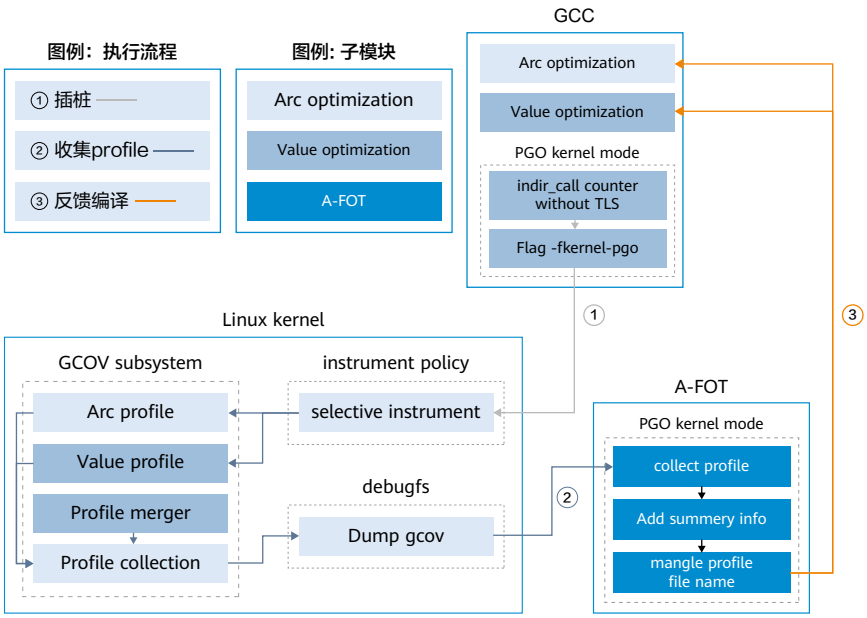
应用场景 2：密钥缓存管理

利用平台远程证明，TEE 远程证明以及 TEE 本地证明对可信应用（TA）需要从企业或云基础设施的密钥管理服务（KMS）获取并缓存密钥的场景进行安全加固，让密钥的传输、存贮和使用中的安全性得到更好的保障。

GCC for openEuler

GCC 编译器基于开源 GCC 10.3 版本开发，支持自动反馈优化、软硬件协同、内存优化、SVE 向量化、矢量化数学库等特性。

- 充分使能ARM架构的硬件特性，运行效率更高，在 SPEC CPU 2017 等基准测试中性能大幅优于上游社区的 GCC 10.3 版本。
- 支持 mcmodel=medium、fp-model、四精度浮点、矢量化数学库等功能。
- 支持自动反馈优化特性，实现应用层 MySQL 数据库等场景性能大幅提升。
- 多版本 GCC 共存支持：提供以 GCC 12.2.0 为基线的 gcc-toolset-12 系列软件包，支持 Intel SPR 相关特性。
- 本次新增支持内核反馈优化特性。通过增强内核与 GCC，实现内核支持编译器反馈优化；用户可通过自动反馈优化工具 A-FOT，一键构建针对特定场景优化的内核。



功能描述

内核反馈优化

内核：支持完整编译器 PGO 能力，包括 arc 与 value profile。

GCC：支持内核反馈优化模式，新增选项 -fkernel-pgo。

A-FOT：支持一键启动内核反馈优化。

具体性能提升效果与目标应用热点在内核中的占比有关。

应用场景

通用计算领域，运行 SPECCPU 2017 测试，相比于上游社区的 GCC 10.3 版本可获得 20% 左右的性能收益。

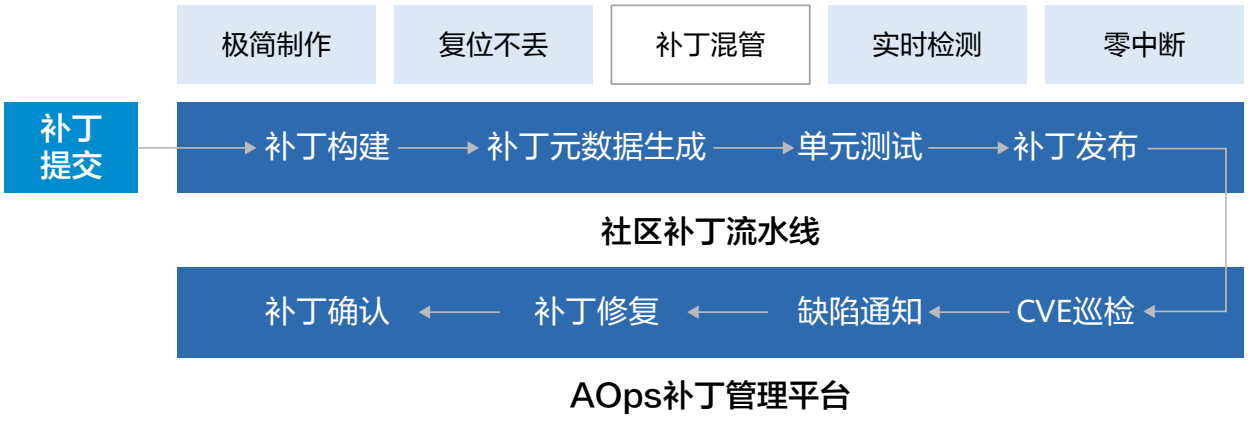
高性能计算领域，运行 WRF、NEMO 应用，相比于上游社区的 GCC 10.3 版本能够获得 10% 左右的性能收益。

其他场景领域，使能自动反馈优化后，MySQL 性能提升 15% 以上；使能内核反馈优化后，实现 Unixbench 性能提升 3% 以上。

A-Ops 智能运维

A-Ops 是一款基于操作系统维度的故障运维平台，提供从数据采集，健康巡检，故障诊断，故障修复的到智能运维解决方案。A-Ops 项目包括了若干子项目：覆盖故障发现（Gala），故障定位支撑（X-diagnosis），缺陷修复（Apollo）等。

本次发布的 apollo 项目是智能补丁管理框架，提供 CVE/Bug 实时巡检，冷热补丁修复，实现自动发现和零中断修复。



功能描述

新增社区热补丁流水线

- 热补丁制作：支持在冷补丁 PR 内指定软件包的目标版本和 patch 文件，构建热补丁。
- 热补丁发布： 支持通过热补丁 issue 自动收集待发布热补丁，复用冷补丁 udapte 版本发布逻辑进行发布。

• 漏洞管理能力增强

- 智能补丁巡检：支持单机 / 集群的 CVE 巡检和通知能力，一键式修复和回退。
- 热修复：支持部分 CVE 通过热补丁修复，做到业务零中断修复。
- 补丁服务：支持冷热补丁订阅，提供补丁在线获取能力。

应用场景

应用场景 1：社区热补丁流水线

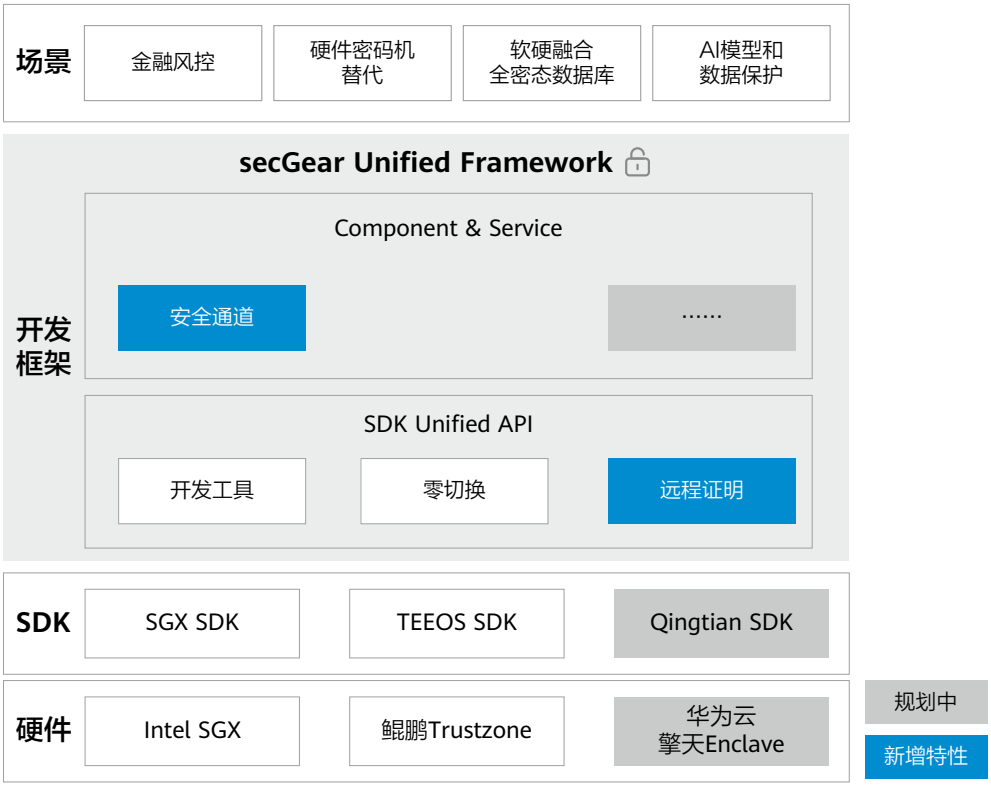
针对现网环境往往不支持重启服务或机器进行修复的问题，从个人制作和版本角度，提供便捷的热补丁制作工具和流水线，根据补丁文件快速完成规范化的热补丁制作、下载、验证和发布。

应用场景 2：漏洞管理

针对有运维诉求的管理员或个人开发者，可以通过漏洞管理服务快速实现漏洞感知和修复，极大减少补丁管理成本，保障集群安全，提升漏洞修复效率。

secGear 特性增强

secGear 是面向计算产业的机密计算安全应用开发套件，屏蔽不同的 TEE（Trusted Execution Environment）SDK 差异提供统一的开发框架，同时提供开发工具、通用安全组件等，帮助安全应用开发者聚焦业务，提升开发效率。



secGear 的整体架构如图所示，主要提供三大能力：

- 架构兼容：屏蔽不同 SDK 接口差异，提供统一开发接口，实现不同架构共源码。
- 易开发：提供开发工具、通用安全组件等，帮助用户聚焦业务，开发效率显著提升。
- 高性能：提供零切换特性，在 REE-TEE 频繁交互、大数据交互等典型场景下提升 REE-TEE 交互性能 10 倍 +。

本次版本新增支持远程证明，安全通道两个特性。

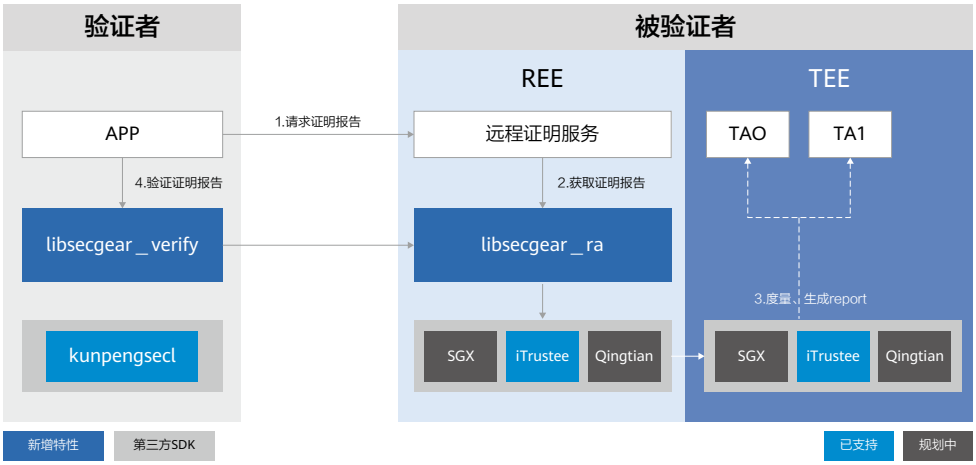
远程证明

机密计算厂商纷纷推出远程证明技术，可以让租户随时检测云上可信执行环境及应用的可信状态，彻底打消租户的顾虑。

远程证明是一种动态度量技术，可以对可信执行环境和运行在环境里的应用进行实时度量，生成证明报告，并使用预置根密钥签名，防止证明报告被篡改或伪造。

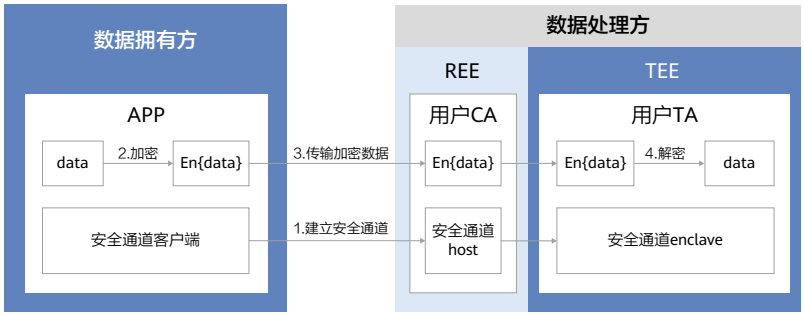
secGear 远程证明基于各厂商 SDK 远程证明能力，封装统一远程证明接口，当前仅支持鲲鹏平台，其中 secGear 依赖 kunpengsecl 的 TEE 验证库来验证证明报告。

secGear 远程证明与 kunpengselc TEE 远程证明区别在于：secGear 是统一开发框架，提供统一远程证明 SDK 接口（获取远程证明报告、校验报告接口），供客户基于 secGear 开发统一远程证明服务和验证程序。kunpengsecl TEE 远程证明提供鲲鹏平台的远程证明服务、证明报告校验，可供用户直接部署使用。



安全通道

数据拥有者在请求云上机密计算服务时，需要把待处理数据上传到云上 TEE 环境中处理，由于 TEE 没有网络，用户数据需要经过网络先传输到 REE，REE 接收到数据的明文后，再传入 TEE 中。用户数据的明文暴露在 REE 内存中，存在安全风险。



secGear 安全通道基于远程证明和密钥协商，实现数据拥有方与 TEE 之间完成密钥协商，并使用协商出来的密钥加密数据并传输，在 REE 侧转发的数据是密文，在 TEE 中收到数据的密文后解密，达到保护数据安全的传入 TEE 中的效果。

功能描述

远程证明的功能：

- 支持远程证明报告获取。
- 支持远程证明报告校验。
- 支持同物理机 TA 对其他 TA 发起本地证明。

安全通道：

安全通道 SDK 分为客户端、host、enclave 三部分。

- 客户端：支持建立安全通道、加解密，供数据拥有方应用调用。
- host：支持初始化安全通道服务，供用户 CA 调用。
- enclave：支持加解密，供用户 TA 调用。

应用场景

应用场景：密态数据库

密态数据库在 TEE 中提供 SQL 查询和计算能力，当数据库客户端请求查询时会将客户端密钥通过安全通道传入 TEE 中，在安全通道协商前需要先通过远程证明验证服务端 TA 的合法性，校验通过后再将建立安全通道并传输密钥，否则终止。

sysmonitor 功能

sysmonitor 是一款系统运维监控软件，支持监控系统磁盘、CPU、内存、进程 / 线程数量、句柄数量等系统资源使用情况，支持监控关键进程并在其异常时将其恢复，支持监控文件系统异常，支持监控记录系统网卡和文件被操作日志，支持自定义监控，执行用户监控动作。

功能描述

文件系统监控	关键进程监控	文件监控	磁盘分区监控
网卡状态监控	CPU 监控	内存监控	进程数 / 线程数监控
系统句柄数监控	磁盘 inode 监控	磁盘 io 延时监控	僵尸进程监控
自定义监控			

- 支持文件系统、磁盘分区、网卡状态、CPU、内存、进程数、系统句柄数监控。
- 支持关键进程监控，进程异常时快速恢复业务进程。
- 支持关键文件监控，记录文件操作信息，方便定位文件异常问题。
- 支持自定义监控框架，支持用户拓展监控功能。

应用场景

适用于操作系统运维监控场景，方便运维人员观测操作系统运行情况、监控系统运行异常。 系统异常发生后辅助运维人员快速定界定位。

网卡监控场景，系统运行过程中可能出现人为原因或者异常导致网卡状态或 IP 发生改变，导致服务或系统异常。sysmonitor 记录网卡的启停、增加和删除 ip 事件，并记录对应操作者。运维人员可以通过日志快速定位网卡变化的时间点、操作者。

系统资源监控场景，sysmonitor 支持监控系统磁盘、CPU、内存、进程 / 线程数量、句柄数量等系统资源使用情况，系统资源使用率超过配置阈值时，记录日志方便运维人员定位系统资源异常类问题。

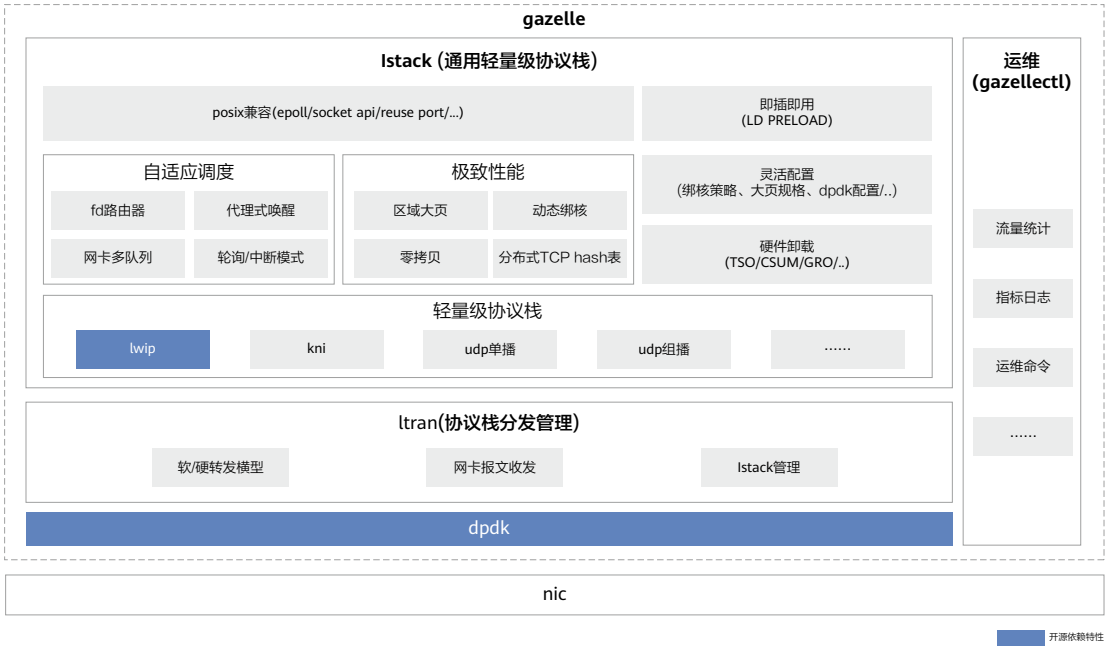
关键进程监控场景，业务中关键进程的正常运行至关重要，进程难免因为系统或人为的原因导致异常奔溃，此时 sysmonitor 监控到进程异常快速恢复关键进程，保证进程正常运行。

Gazelle 特性增强

Gazelle 是一款高性能用户态协议栈。它基于 DPDK 在用户态直接读写网卡报文，共享大页内存传递报文，使用轻量级 LwIP 协议栈。能够大幅提高应用的网络 I/O 吞吐能力。专注于数据库网络性能加速，兼顾高性能与通用性。本次版本新增 UDP 协议及相关接口支持，丰富用户态协议栈。

功能描述

gazelle 功能架构图：



- 高性能（超轻量）：基于 dpdk、lwip 实现高性能轻量协议栈能力。
- 极致性能：基于区域大页划分、动态绑核、全路径零拷贝等技术，实现高线性度并发协议栈。
- 硬件加速：支持 TSO/CSUM/GRO 等硬件卸载，打通软硬件垂直加速路径。
- 通用性（posix 兼容）：接口完全兼容 posix api，应用零修改，支持 udp 的 recvfrom 和 sendto 接口。
- 通用网络模型：基于 fd 路由器、代理式唤醒等机制实现自适应网络模型调度，udp 多节点的组播模型，满足任意网络应用场景。
- 易用性（即插即用）：基于 LD_PRELOAD 实现业务免配套，真正实现零成本部署。
- 易运维（运维工具）：具备流量统计、指标日志、命令行等完整运维手段。

应用场景

适用于提升网络协议栈成为性能瓶颈点的应用提升业务处理性能。

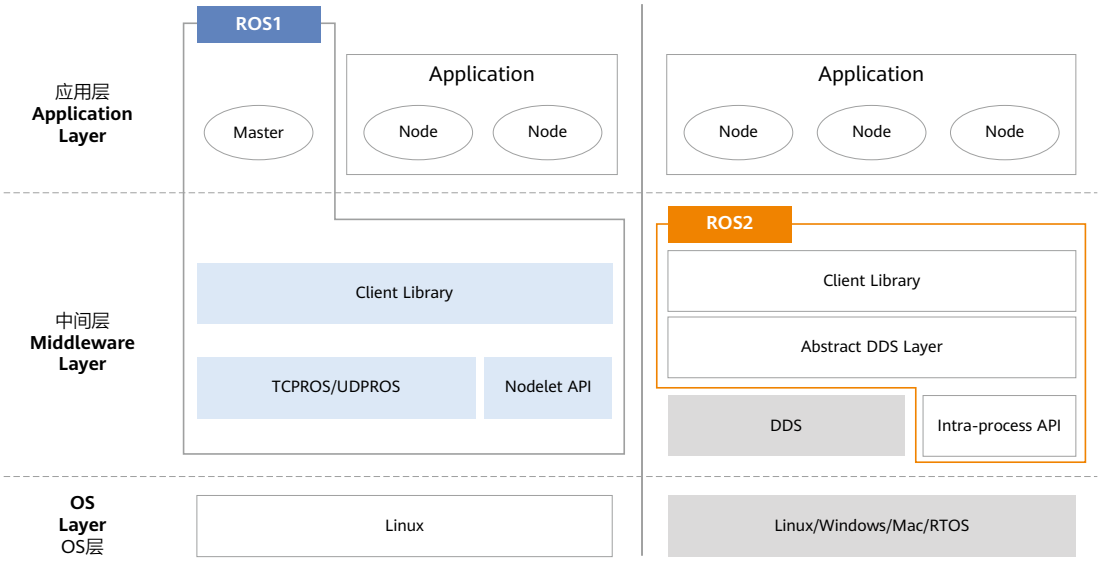
ROS 组件支持

ROS（Robot Operating System）是一组软件库和工具，可帮助开发者构建机器人应用程序。从驱动程序到算法，开发工具，ROS 把原本松散的零部件耦合在了一起，为开发者们提供了通信架构。ROS 虽然叫做操作系统，但并非传统意义的操作系统，它只是连接了操作系统和开发者的 ROS 应用程序，所以它是一个中间件，基于 ROS 的应用程序之间建立起了沟通的桥梁，所以也是运行在 Linux 上的运行时环境，在这个环境上，机器人的感知、决策、控制算法可以更好的组织和运行。

本特性提供了基于 ROS2 的基础核心 ROS 组件，包括通信架构、编译工具，填补了在 openEuler 上使用 ROS 的空白。

功能描述

ROS 功能架构图：



22.03 LTS SP2 支持 ROS humble 版本，特性主要提供：

- 在 openEuler 服务器 / 边缘版本提供 ros-core 和 ros-base 全部软件包。
- 在 openEuler 嵌入式版本支持 SLAM 建图和导航。
- 支持基于 openEuler 开发、构建、调试 ROS 应用（支持 rqt 系列工具，暂不支持 rviz 和 gazebo）。

应用场景

适用于各种机器人设备开发，例如人形及机器人、服务机器人、工业机器人等。用户可根据自己的需求安装对应的 ROS 功能软件包，然后开发自己的机器人应用。

openEuler WSL 支持

WSL（Windows Subsystem for Linux）是支持在 Windows 上运行 Linux 用户态软件的适配层，通过在微软应用商店，用户可以快速在 Windows 中获得原生的 openEuler 体验。

功能描述

- 一键安装，开箱即用：在支持 WSL 的 Windows 设备上，通过 Windows Store 即可一键下载体验最新的 openEuler LTS 版本。
- 全生命周期支持：22.03 LTS 版本的 WSL 应用将会更新至 22.03-LTS-SP2 版本 。
- 自定义友好：用户既可以通过微软商店下载已经打包的 openEuler WSL 应用，也可以基于 openEuler WSL 仓库的开源代码构建自己的 WSL 应用。
- 新增 metalink 支持：openEuler 的用户在使用 dnf 安装软件包时，metalink 服务会引导 dnf 到当前 IP 附近的镜像站点下载 / 更新软件包，可以较大提升软件包安装的体验。

应用场景

- 在 Windows 中快速部署和体验 openEuler LTS 版本。
- 利用 vs code 和 openEuler WSL 打造流畅跨平台开发体验。
- 在 openEuler WSL 中搭建 K8S 集群。
- 使用你喜爱的 openEuler command-line 程序或脚本处理 Windows 或 WSL 中的文件和程序。

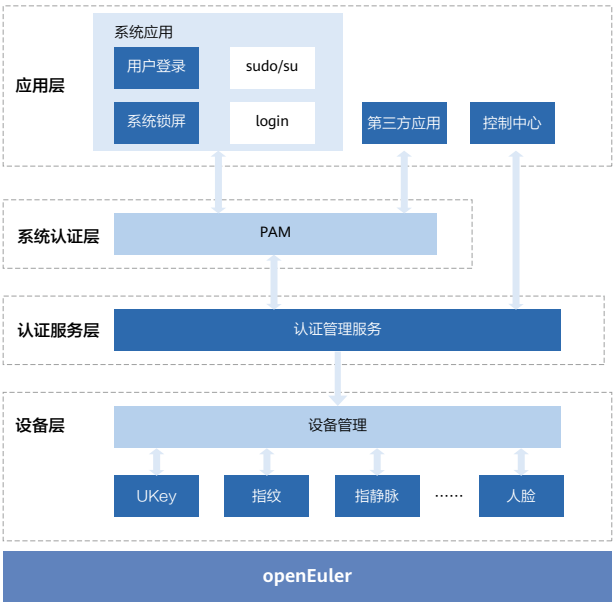
发布 kiran-desktop 2.5 版本

Kiran-desktop 2.5 版本新增多因子认证特性，Kiran 桌面的多因子认证方案是将多个认证方式进行组合对用户进行认证的方案。组合方式分为或模式和与模式。在或模式下，只要其中一种认证方式通过即可；在与模式下，需要所有认证方式通过才能认证成功。

功能描述

控制中心提供认证管理功能，包括：

- 认证方式开关等功能。
- 支持特征的录入和删除。
- 支持指定默认认证设备。
- 支持设备驱动的开关控制。
- 支持图形和非图形两种认证场景。
- 图形场景包括登录、锁屏和提权等。
- 非图形场景包括 sudo、su 等。
- 支持 Ukey，虹膜，人脸，指纹，指静脉等认证方式。
- 提供认证管理服务，解决设备抢占问题。在开启多个认证会话时，认证服务层可以通过调度算法来让多个会话共享一个认证设备。
- 提供设备管理服务。屏蔽不同设备接口之间的差异，为上层服务提供统一的设备操作接口。

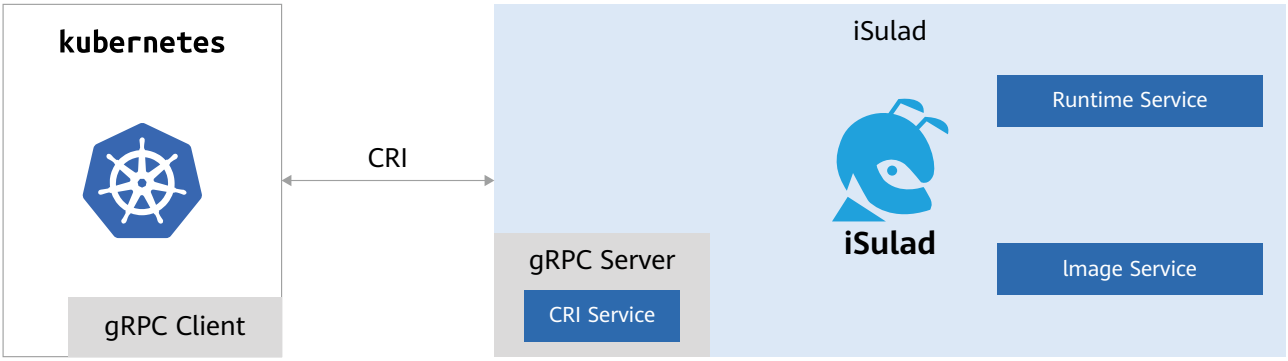


iSulad 支持 K8S 1.24/1.25 版本

作为云原生技术的基石，使用 Kubernetes + 容器引擎构建云原生应用已经在业界被广泛推广。iSulad 作为一款轻量级容器引擎，北向支持 CRI 标准，可以作为容器底座对接 Kubernetes。

功能描述

由于 Kubernetes 社区不断的发展，CRI 标准也在不断的更新迭代，iSulad 支持的 Kubernetes 版本已经远落后于当前最新的版本。为了更好的完善生态，方便用户使用，iSulad 对支持的 CRI 版本进行了升级。从之前 v1aplha2 接口的 1.1X 版本，升级到 v1alpha2 接口的 1.24/1.25 版本。



根据 1.25 版本的 CRI 接口规范进行差异化排查，升级涉及到的主要接口包括：

- 完善 CreateContainer、UpdateContainerResources、ContainerStatus、ContainerStats、ListContainerStats 等接口。
- 新增 PodSandboxStats、ListPodSandboxStats 接口。

应用场景

可以应用于 Kubernetes 1.24/1.25 版本 + 容器引擎的场景中。由于 CRI 接口向前兼容，因此对于一些较老版本的 Kubernetes，iSulad 同样支持。

sysMaster 系统管理大师

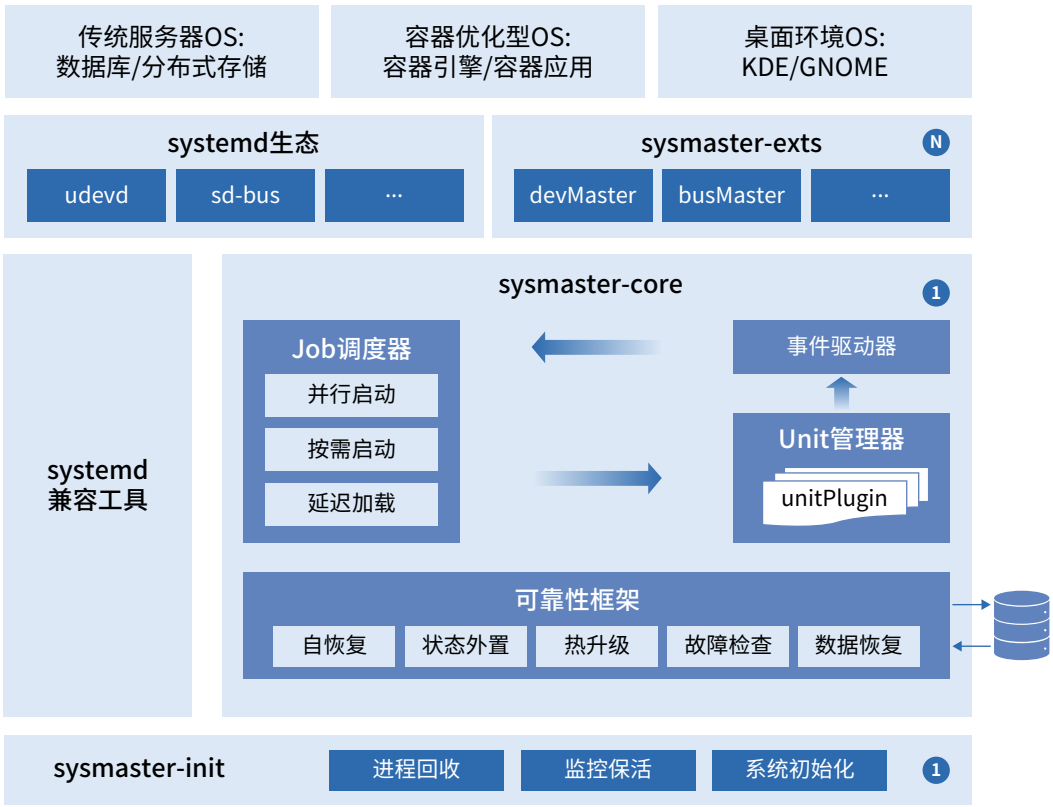
sysMaster 是一套超轻量、高可靠的服务管理程序集合，是对 1 号进程的全新实现，旨在改进传统的 init 守护进程。它使用 Rust 编写，具有故障监测、秒级自愈和快速启动等能力，从而提升操作系统可靠性和业务可用度。

sysMaster 支持进程、容器和虚拟机的统一管理，并引入了故障监测和自愈技术，从而解决 Linux 系统初始化和服务管理问题，其适用于服务器、云计算和嵌入式等多个场景。

功能描述

sysMaster 实现思路是将传统 1 号进程的功能解耦分层，结合使用场景，拆分出 1+1+N 的架构。如下面 sysMaster 系统架构图所示，主要包含三个方面：

- sysmaster-init：新的 1 号进程，功能极简，代码千行，极致可靠，提供系统初始化 / 僵尸进程回收 / 监控保活等功能，可单独应用于嵌入式场景。
- sysmaster-core：承担原有服务管理的核心功能，引入可靠性框架，使其具备崩溃快速自愈、热升级等能力，保障业务全天在线。
- sysmaster-extends：使原本耦合的各组件功能独立，提供系统关键功能的组件集合（如设备管理 devMaster，总线通信 busMaster 等），各组件可单独使用，可根据不同场景灵活选用。



sysMaster 组件架构简单，提升了系统整体架构的扩展性和适应性，从而降低开发和维护成本。其主要特点如下：

- 具有自身故障秒级自愈和版本热升级能力。
- 具备快速启动的能力，更快的启动速度和更低的运行底噪。
- 采用插件化机制，支持按需动态加载各种服务类型。
- 提供迁移工具，支持从 Systemd 快速无缝迁移到 sysMaster。
- 结合容器引擎 (iSulad) 和 Qemu，提供统一的容器实例和虚拟化实例的管理接口。

本次随 SP2 发布的 0.2.4 版本，仅支持在容器场景下，以 sysMaster 的方式管理系统中的服务。

新增特性：

- 支持 aarch64、x86_64 架构。
- 支持 service、target 两种 unit 服务类型。
- 支持 10+service 配置。
- 支持 sctl 命令行管理服务的生命周期。
- 支持日志输出到文件。

约束限制：

- 当前仅支持 64 位系统。
- 当前仅支持在系统容器中运行。
- 当前仅支持 sysMaster 使用的 toml 配置格式。

未来，sysMaster 将继续探索在多场景下的应用，并持续优化架构和性能以提高可扩展性和适应性。同时，我们还将开发新的功能和组件以满足容器化、虚拟化、边缘计算等场景的需求。让 sysMaster 成为一个强大的系统管理框架，为用户提供更好的使用体验和更高的效率。

应用场景

sysMaster 致力于替代容器、虚机、服务器及边缘设备上现有 1 号进程。

支持 OpenStack Train、Wallaby 多版本

OpenStack 是目前全球部署最广泛的、经过大规模生产环境验证的开源云基础设施平台，其中包括一系列软件组件，为云基础架构提供通用服务。

功能描述

OpenStack 包含云服务众多，结构复杂，部署难度大，为了方便用户在 openEuler 上方便使用 OpenStack，openEuler 社区提供了基于 RPM 的 OpenStack 部署方式，用户只需要通过一些简单命令就能快速部署一套符合要求的 OpenStack 集群。同时由于 OpenStack 版本众多，在每个 LTS 版本中，openEuler 会根据用户的反馈，发布最热门的几个 OpenStack 版本，并以多版本支持的方式向更多用户提供服务。

在 22.03 LTS SP2 中，openEuler 提供了 OpenStack Train 和 Wallaby 的支持。其中：

- Train、Wallaby 版都支持的服务有：Keystone、Glance、Nova、Cinder、Neutron、Tempest、Horizon、Ironic、Placement、Trove、Kolla、Rally、Swift、Ceilometer、Heat、Aodh、Cyborg、Gnocchi。
- Wallaby 版在 T 的基础上新增一些服务，包括：Designate、Manila、Barbican、Octavia、Cloudkitty、Masakari、Mistral、Senlin 和 Zaqar。

应用场景

当用户希望部署基于 openEuler 的云计算平台时，可以使用 openEuler 提供的 OpenStack 组件，通过 RPM 软件包的方式快速部署。OpenStack 可以用作构建私有云和混合云、云计算服务提供商、学术研究和开发、测试和评估云计算技术和应用程序等众多方面。

Lustre Client 软件包支持

Lustre 是一个开源、分布式并行文件系统软件平台，具有高可扩展、高性能、高可用等特点。Lustre 运行于 Linux 系统之上，提供符合 POSIX 标准的 UNIX 文件系统接口。

功能描述

openEuler 22.03 LTS SP2 引入了 Lustre v2.15.2 client 端相关组件。如需 server 端组件可以从社区源码编译，目前社区最新代码已支持 openEuler 22.03 LTS。Client 端的引入之后，用户可以在计算节点在 openEuler 的操作系统上可以使用 Lustre 并行文件系统，体验 Lustre 文件系统的功能。

应用场景

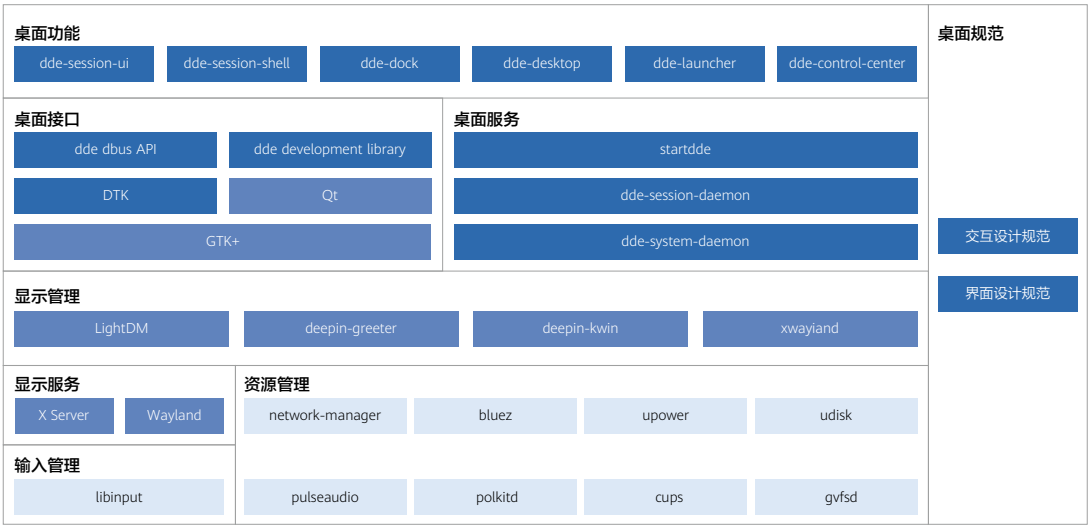
Lustre 并行文件系统适用于需要海量文件存储和高性能文件存储的应用场景。根据社区调查报告 Lustre 广泛应用于研究机构、科学计算、媒体、工业制造、金融、教育等领域行业的 HPC 和 AI 计算场景。

DDE 组件更新

统信桌面环境（DDE）是统信软件为统信操作系统（UniontechOS）开发的一款桌面环境，统信桌面操作系统、统信操作系统服务器版和统信操作系统专用设备版均在使用统信桌面环境。

功能描述

统信桌面环境专注打磨产品交互、视觉设计，拥有桌面环境的核心技术，主要功能包含：登录锁屏、桌面及文件管理器、启动器、任务栏（DOCK）、窗口管理器、控制中心等。由于界面美观、交互优雅、安全可靠、尊重隐私，一直是用户首选桌面环境之一，用户可以使用它进行办公与娱乐，在工作中发挥创意和提高效率，和亲朋好友保持联系，轻松浏览网页、享受影音播放。



显示服务、输入管理、资源管理较为底层，一般是基于 golang 开发的后端服务，为上层 GUI 程序提供桌面环境中所需功能接口，如创建用户、设置屏幕亮度、设置设备音量、管理网络连接等功能。

显示管理、桌面接口、桌面服务属于 shell 层，一般是基于 DBus 接口协议与后端服务进行通信，为定义用户界面、交互操作提供支撑，如登录界面、窗口外观、GUI 应用程序控件等。

应用场景

桌面功能属于应用层，一般是面向用户可操作的功能界面，比如启动器、任务栏（DOCK）等。

KAE 模块支持 nosva 特性

鲲鹏加速引擎（KAE）是基于鲲鹏硬件能力的加速解决方案，包含了 KAE 加解密、KAEZlib 解压缩模块，分别用于加速 SSL/TLS 应用和数据压缩，可以显著降低处理器消耗，提高处理器效率。此外，加速引擎对应用层屏蔽了其内部实现细节，用户通过 OpenSSL、zlib 标准接口即可以实现快速迁移现有业务。

功能描述

KAE 加解密

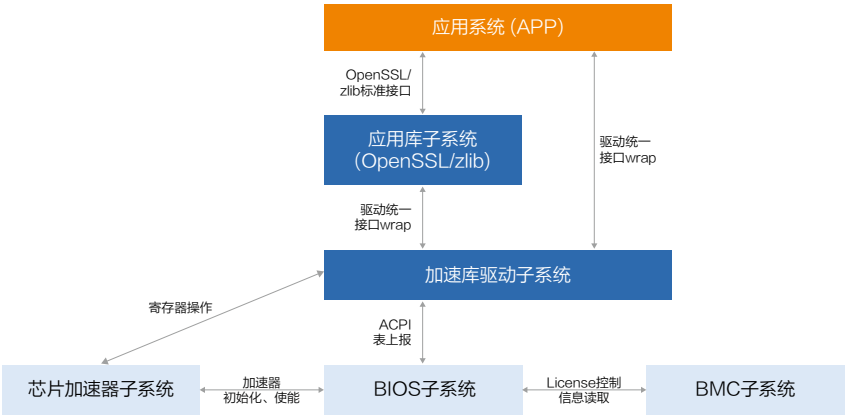
KAE 加解密是鲲鹏加速引擎的加解密模块，使用鲲鹏硬加速引擎实现 RSA/SM3/SM4/DH/MD5/AES 算法，结合无损用户态驱动框架，提供高性能对称加解密、非对称加解密算法能力，兼容 OpenSSL 1.1.1a 及其之后版本，支持同步 & 异步机制。

- 摘要算法 SM3/MD5，支持异步模型。
- 对称加密算法 SM4，支持异步模型，支持 CTR/XTS/CBC/ECB/OFB 模式。
- 对称加密算法 AES，支持异步模型，支持 ECB/CTR/XTS/CBC 模式。
- 非对称算法 RSA，支持异步模型，支持 Key Sizes 1024/2048/3072/4096。
- 密钥协商算法 DH，支持异步模型，支持 Key Sizes 768/1024/1536/2048/3072/4096。

KAEzip

KAEzip 是鲲鹏加速引擎的压缩模块，使用鲲鹏硬加速模块实现 deflate 算法，结合无损用户态驱动框架，提供高性能 Gzip/zlib 格式压缩接口。

- 支持 zlib/Gzip 数据格式，符合 RFC1950/RFC1952 标准规范。
- 支持 deflate 算法。
- 支持同步模式。
- 单处理器（鲲鹏 920 处理器）最大压缩带宽 7GB/s，最大解压带宽 8GB/s。
- 支持的压缩比 ≈ 2，与 zlib 1.2.11 接口保持一致。



应用场景

提供高性能对称加解密、非对称加解密算法能力，兼容 OpenSSL 1.1.1a 及其之后版本，支持同步 & 异步机制。

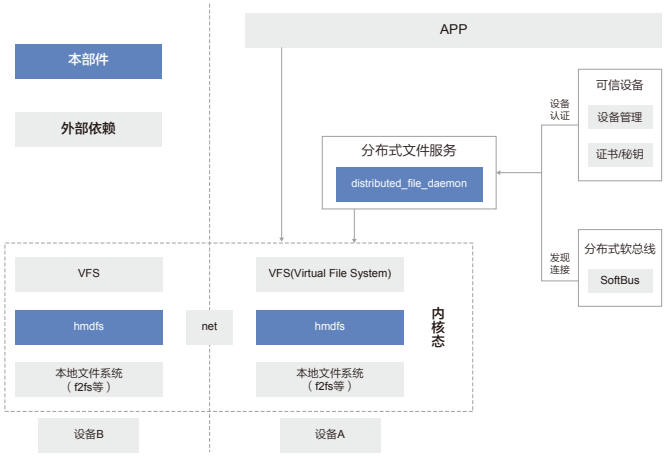
通过加速引擎可以实现不同场景下应用性能的提升，例如在分布式存储场景下，通过 zlib 加速库加速数据压缩和解压。

基于软总线的分布式文件系统（HMDFS）

HMDFS 是从 OH 社区迁移而来的在软总线生态之上的一个分布式文件系统，其在分布式软总线动态组网的基础上，为网络上各个设备结点提供一个全局一致的访问视图，支持开发者通过基础文件系统接口进行读写访问，具有高性能、低延时等优点。

功能描述

- distributed_file_daemon：分布式文件管理常驻用户态服务，负责接入设备组网接口控制，并负责挂载 hmdfs 及权限管理。
- 可信设备：统一管理不同业务建立的本设备与其他设备间的互信关系。
- 分布式软总线：网络链路层的发现连接。
- VFS：内核在物理存储介质上的文件系统和用户之间的一个虚拟文件系统的软件抽象层。
- hmdfs：分布式文件系统核心模块，是一种面向移动分布式场景的、高性能的、基于内核实现的、堆叠式文件系统。



文件查看视图

- 统一融合视图：HMDFS 将不同设备的本地文件系统目录融合成统一视图，从而实现跨设备、统一的用户访问视图。
- 分设备视图：分设备视图可以让用户和应用清晰地看到每个设备的目录，将本地和远端操作解耦，同时也能方便的支持远端创建操作。

数据同步方式

采用订阅发布设计，详细如下：

- Cache 订阅发布，按需同步，默认仅同步目录树，基于目录树访问管理订阅关系。
- 基于事件触发和超时的一致性核查机制，保证最终一致性。
- 懒加载方式，仅发失效通知给订阅端，元数据获取依赖下次访问重新获取。

应用场景

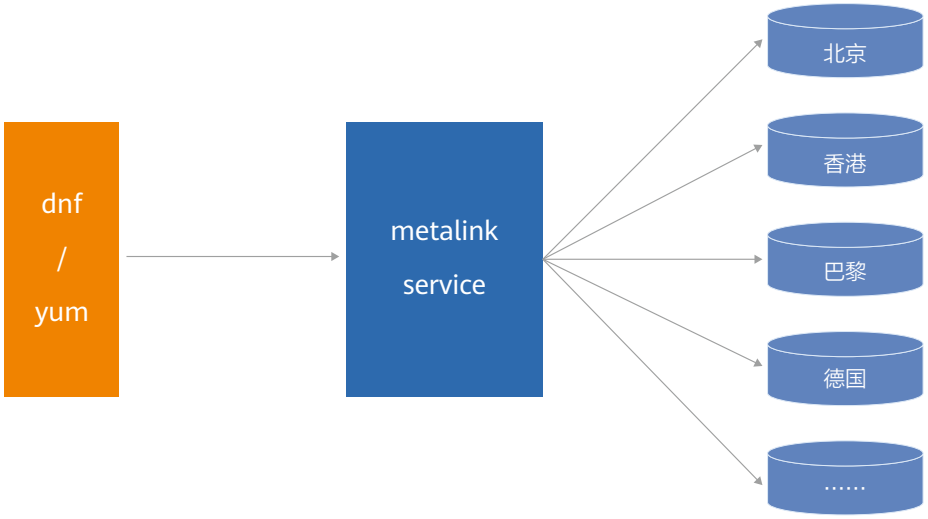
基于 openEuler 的分布式软总线技术，集成至工业互联总线 SDK 中，发挥极简、高性能的文件同步优势将工业产线嵌入式设备数据（图片采集、配置信息）实时同步至边缘服务器处理。

根据用户情况自动优化下载软件包

当前有 30 个站点为 openEuler 提供了镜像服务，这些站点分布在亚洲、欧洲和北美。支持用户从最近的多个站点下载软件包，进而提高其下载软件包的速度。

功能描述

在 dnf 或 yum 的配置文件中配置 metalink（随版本发布），其值是 metalink service 提供的 API 的 URL 地址。当用户下载软件包时，dnf 或 yum 客户端会请求 metalink 指向的 URL，该 URL 对应的服务会返回 xml 格式的数据，该数据包含离用户最近的多个镜像站点的地址。dnf 或 yum 客户端会从这些地址中选择最优的站点进行下载，进而提高其下载软件包的速度。



应用场景

用户使用 dnf、yum 下载软件包。

08 著作权说明

09 商标

openEuler 白皮书上使用和显示的所有商标、标志皆属 openEuler 社区所有，但注明属于其他方拥有的商标、标志、商号除外。未经 openEuler 社区或其他方书面许可，openEuler 白皮书所载的任何内容不应被视作以暗示、不反对或其他形式授予使用前述任何商标、标志的许可或权利。未经事先书面许可，任何人不得以任何方式使用 openEuler 社区的名称及 openEuler 社区的商标、标记。

附录 10

附录 1：搭建开发环境

环境准备	地址
下载安装 openEuler	https://openeuler.org/zh/download/
开发环境准备	https://gitee.com/openeuler/community/blob/master/zh/contributors/prepare-environment.md
构建软件包	https://gitee.com/openeuler/community/blob/master/zh/contributors/package-install.md

附录 2：安全处理流程和安全批露信息

社区安全问题披露	地址
安全处理流程	https://gitee.com/openeuler/security-committee/blob/master/security-process.md
安全披露信息	https://gitee.com/openeuler/security-committee/blob/master/security-disclosure.md

商标声明

在本手册中以及本手册描述的产品中，出现的商标，产品名称，服务名称以及公司名称，由其各自的所有人拥有。

免责声明

本文档可能含有预测信息，包括但不限于有关未来的财务、运营、产品系列、新技术等信息。由于实践中存在很多不确定因素，可能导致实际结果与预测信息有很大的差别。因此，本文档信息仅供参考，不构成任何要约或承诺，不对您在本文档基础上做出的任何行为承担责任。可能不经通知修改上述信息，恕不另行通知。