

openEuler 22.03 LTS SP4 技术白皮书

1. 概述

OpenAtom openEuler（简称“openEuler”）社区是一个面向数字基础设施操作系统的开源社区。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，简称为 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，并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、基础技术创新及上层应用创新。

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。

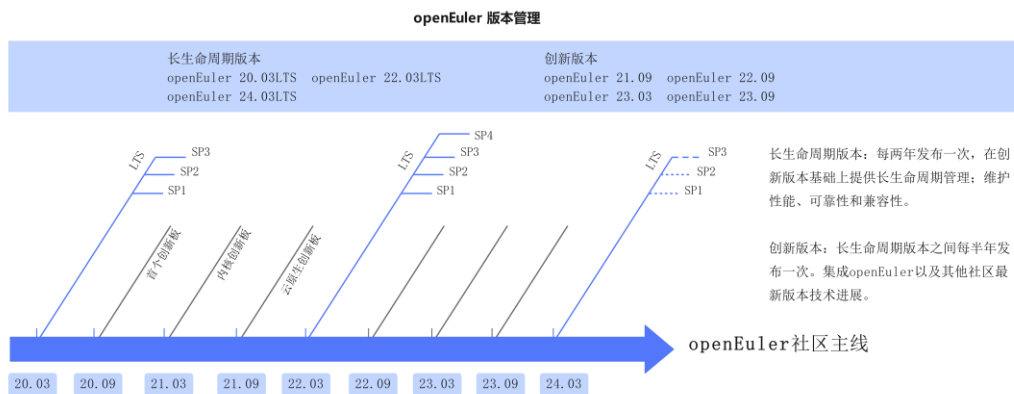
2023 年 9 月 30 日，发布 openEuler 23.09 创新版本，是基于 6.4 内核的创新版本（参见版本生命周期），提供更多新特性和功能，给开发者和用户带来全新的体验，服务更多的领域和更多的用户。

2023 年 11 月 30 日，发布 openEuler 20.03 LTS SP4 版本，其作为 20.03 LTS 版本的增强扩展版本，面向服务器、云原生、边缘计算场景，提供更多新特性和功能增强。

2023 年 12 月 30 日，发布 openEuler 22.03 LTS SP3 版本，是 22.03 LTS 版本增强扩展版本，面向服务器、云原生、边缘计算和嵌入式场景，持续提供更多新特性和功能扩展，给开发者和用户带来全新的体验，服务更多的领域和更多的用户。

2024 年 5 月 30 日，发布 openEuler 24.03 LTS，基于 6.6 内核的长周期 LTS 版本（参见版本生命周期），面向服务器、云、边缘计算、AI 和嵌入式场景，提供更多新特性和功能，给开发者和用户带来全新的体验，服务更多的领域和更多的用户。

2024 年 6 月 30 日，发布 openEuler 22.03 LTS SP4，是 22.03 LTS 版本增强扩展版本，面向服务器、云原生、边缘计算和嵌入式场景，持续提供更多新特性和功能扩展，给开发者和用户带来全新的体验，服务更多的领域和更多的用户。



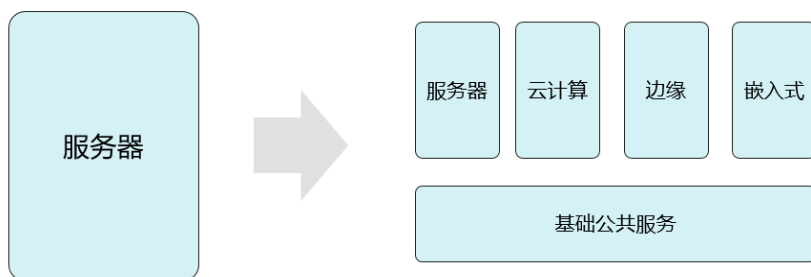
openEuler 作为一个操作系统发行版平台，每两年推出一个 LTS 版本。该版本为企业级

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。

2. 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 SP4 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 6.6 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 SP4 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 SP4 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 24.03 LTS 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。
- openEulerDevKit：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，比如 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 Sapphire Rapids/Emerald 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	LoongArch	SW64	RISC-V
CPU	Intel、AMD、兆芯、海光	鲲鹏、飞腾	龙芯	申威	Sophgo、T-Head 等

支持的整机如下：

硬件类型	x86	Arm
------	-----	-----

整机	<ul style="list-style-type: none"> •Intel: 超聚变、联想、浪潮 •AMD: 新华三、超微 •海光: 曙光/中科可控 	<ul style="list-style-type: none"> •鲲鹏: 泰山 •飞腾: 青松、宝德、新华三
	<ul style="list-style-type: none"> •兆芯: 兆芯 	

支持的板卡类型如下：

硬件类型	x86	Arm
网卡	华为、Mellanox、Intel、Broadcom、云芯智联、Netswift、云脉、沐创	华为、Mellanox、Intel、Broadcom、云芯智联、Netswift、云脉、沐创
Raid	Avago、云芯智联、PMC、华为	Avago、云芯智联、PMC、华为
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/>。

3. 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如下所示。

1. 以 openEuler 22.03 LTS SP4 为 HostOS，组件版本如下：

- libvirt-6.2.0-63.oe2203sp4
- libvirt-client-6.2.0-63.oe2203sp4
- libvirt-daemon-6.2.0-63.oe2203sp4
- qemu-6.2.0-91.oe2203sp4
- qemu-img-6.2.0-91.oe2203sp4

2.兼容的虚拟机列表如下：

HostOS	GuestOS(虚拟机)	架构
openEuler 22.03 LTS SP4	Centos 6	x86_64
openEuler 22.03 LTS SP4	Centos 7	aarch64
openEuler 22.03 LTS SP4	Centos 7	x86_64
openEuler 22.03 LTS SP4	Centos 8	aarch64
openEuler 22.03 LTS SP4	Centos 8	x86_64
openEuler 22.03 LTS SP4	Windows Server 2016	x86_64
openEuler 22.03 LTS SP4	Windows Server 2019	x86_64

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中安装 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中安装 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

4. 场景创新

AI

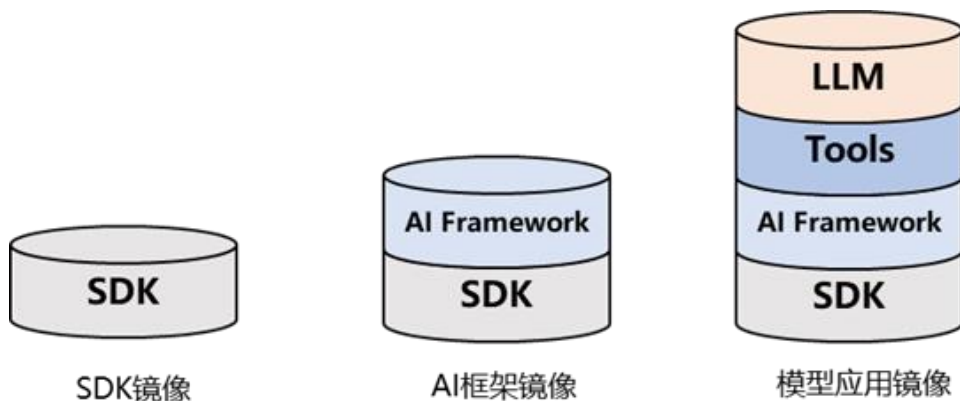
智能时代，操作系统需要面向 AI 不断演进。一方面，在操作系统开发、部署、运维全流程以 AI 加持，让操作系统更智能；另一方面，openEuler 已支持 Arm、x86、RISC-V 等全部主流通用计算架构，在智能时代，openEuler 也率先支持 NVIDIA、昇腾等主流 AI 处理器，成为使能多样性算力的首选。

OS for AI

openEuler 兼容 NVIDIA、Ascend 等主流算力平台的软件栈，为用户提供高效的开发运行环境。通过将不同 AI 算力平台的软件栈进行容器化封装，即可简化用户部署过程，提供开箱即用的体验。同时，openEuler 也提供丰富的 AI 框架，方便大家快速在 openEuler 上使用 AI 能力。

功能描述

1. openEuler 已兼容 CANN、CUDA 等硬件 SDK，以及 TensorFlow、PyTorch 等相应的 AI 框架软件，支持 AI 应用在 openEuler 上高效开发与运行。
2. openEuler AI 软件栈容器化封装优化环境部署过程，并面向不同场景提供以下三类容器镜像。



- SDK 镜像：以 openEuler 为基础镜像，安装相应硬件平台的 SDK，如 Ascend 平台的 CANN 或 NVIDIA 的 CUDA 软件。
- AI 框架镜像：以 SDK 镜像为基础，安装 AI 框架软件，如 PyTorch 或 TensorFlow。
- 模型应用镜像：在 AI 框架镜像的基础上，包含完整的工具链和模型应用。

相关使用方式请参考 [openEuler AI 容器镜像用户指南](#)。

应用场景

openEuler 使能 AI，向用户提供更多 OS 选择。基于 openEuler 的 AI 容器镜像可以解决开发运行环境部署门槛高的问题，用户根据自身需求选择对应的容器镜像即可一键部署，三类容器镜像的应用场景如下。

- SDK 镜像：提供对应硬件的计算加速工具包和开发环境，用户可进行 Ascend CANN 或 NVIDIA CUDA 等应用的开发和调试。同时，可在该类容器中运行高性能计算任务，例如大规模数据处理、并行计算等。
- AI 框架镜像：用户可直接在该类容器中进行 AI 模型开发、训练及推理等任务。
- 模型应用镜像：已预置完整的 AI 软件栈和特定的模型，用户可根据自身需求选择相应的模型应用镜像来开展模型推理或微调任务。

AI for OS

当前，openEuler 和 AI 深度融合，一方面使用基础大模型，基于大量 openEuler 的代码和数据，训练出 EulerCopilot，初步实现代码辅助生成、智能问题智能分析、系统辅助运维等功能，让 openEuler 更智能。

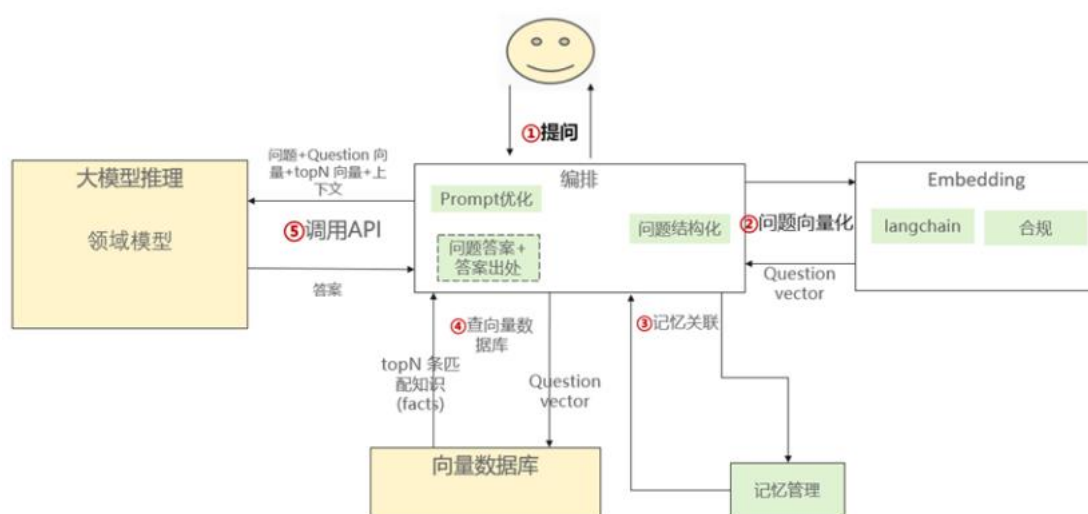
EulerCopilot-智能问答

功能描述

EulerCopilot 智能问答平台目前支持 web 和智能 shell 两个入口。

- Web 入口: 操作简单, 可咨询操作系统相关基础知识, openEuler 动态数据、openEuler 运维问题解决方案、openEuler 项目介绍与使用指导等。

- 智能 Shell 入口: 自然语言和 openEuler 交互, 启发式的运维。



应用场景

- 面向 openEuler 普通用户: 深入了解 openEuler 相关知识和动态数据, 比如咨询如何迁移到 openEuler。

- 面向 openEuler 开发者: 熟悉 openEuler 开发贡献流程、关键特性、相关项目的开发等知识。

- 面向 openEuler 运维人员: 熟悉 openEuler 常见或疑难问题的解决思路和方案、openEuler 系统管理知识和相关命令。

相关使用方式请参考 [EulerCopilot 智能问答服务使用指南](#)。

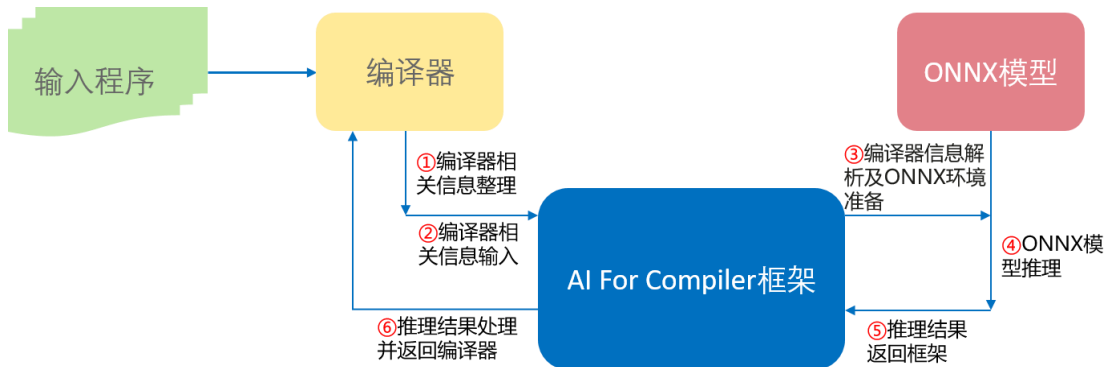
AI for Compiler

AI for Compiler 框架提供相关接口, 使得用户方便快速地在 openEuler 编译器上使用 AI

能力辅助编译选项调优和内存排布优化等优化方式。

功能描述

AI For Compiler 框架接受编译器相关信息作为输入，在框架内部进行编译器信息解析和 ONNX 环境准备之后，向 ONNX 模型发送推理请求。ONNX 模型将根据编译器信息进行推理，并将推理结果返回 AI For Compiler 框架。AI For Compiler 框架对推理结果进行解析、处理，将处理后的结果返回编译器，用于指导多种优化手段的实施。



应用场景

- 面向 openEuler 普通用户：通过开启特定选项，轻松使用 AI 模型指导调优，提升目标程序性能，提高使用体验。
- 面向 openEuler 开发者：编译器与模型解耦合，使得开发者可以轻松便捷地自主训练模型并使用 AI 模型结果辅助编译器调优。

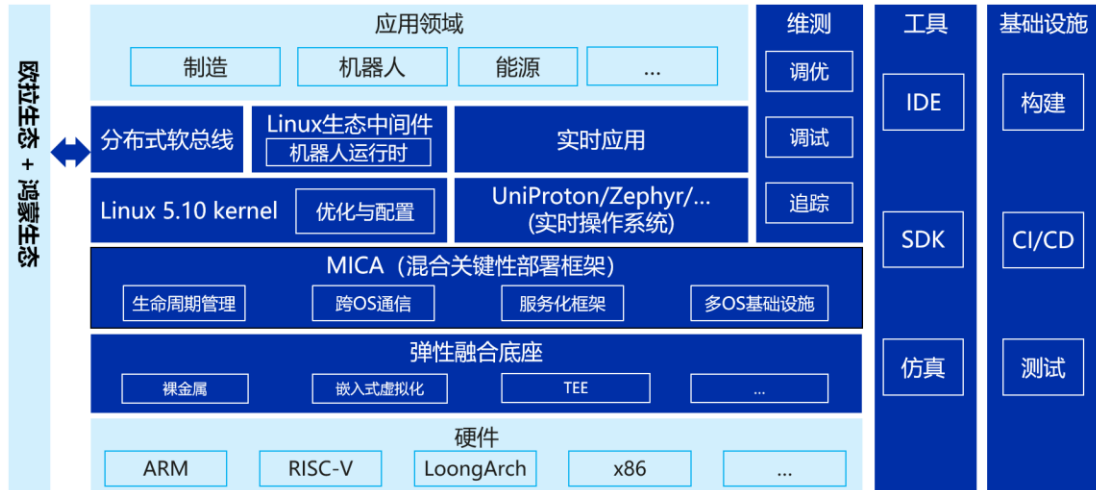
openEuler Embedded

openEuler 发布面向嵌入式领域的版本 openEuler 22.03 LTS SP4 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 SP4 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。

系统架构图



南向生态

openEuler Embedded 22.03 LTS SP4 当前主要支持 ARM64、x86-64 两种芯片架构，未来 openEuler Embedded Linux 还将支持龙芯、飞腾等芯片。

嵌入式弹性虚拟化底座

openEuler Embedded 的弹性虚拟化底座是为了在多核片上系统（SoC, System On Chip）上实现多个操作系统共同运行的一系列技术的集合，包含了裸金属、嵌入式虚拟化、轻量级容器、LibOS、可信执行环境（TEE）、异构部署等多种实现形态。不同的形态有各自的特点：

1. 裸金属：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 Embedded Linux 混合部署。

2. 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 Embedded Linux 混合部署。
3. 实时虚拟化：openEuler 社区自研虚拟化 ZVM，兼顾性能、隔离性和灵活性，综合最优。目前支持 Zephyr 和 Linux 混合部署。

混合关键性部署框架

openEuler Embedded 打造了构建在融合弹性底座之上混合关键性部署框架，并命名为 MICA (Mixed Criticality)，旨在通过一套统一的框架屏蔽下层弹性底座形态的不同，从而实现 Linux 和其他 OS 运行时便捷地混合部署。依托硬件上的多核能力使得通用的 Linux 和专用的实时操作系统有效互补，从而达到全系统兼具两者的特点，并能够灵活开发、灵活部署。

MICA 的组成主要有四大部分：生命周期管理、跨 OS 通信、服务化框架和多 OS 基础设施。生命周期管理主要负责从 OS (Client OS) 的加载、启动、暂停、结束等工作；跨 OS 通信为不同 OS 之间提供一套基于共享内存的高效通信机制；服务化框架是在跨 OS 通信基础之上便于不同 OS 提供各自擅长服务的框架，例如 Linux 提供通用的文件系统、网络服务，实时操作系统提供实时控制、实时计算等服务；多 OS 基础设施是从工程角度为把不同 OS 从工程上有机融合在一起的一系列机制，包括资源表达与分配，统一构建等功能。

混合关键性部署框架当前能力：

- 支持裸金属模式下 openEuler Embedded Linux 和 RTOS (Zephyr/UniProton) 的生命周期管理、跨 OS 通信。
- 支持分区虚拟化模式下 openEuler Embedded Linux 和 RTOS (FreeRTOS) 的生命周期管理、跨 OS 通信。

北向生态

1. 向软件包支持：600+ 嵌入式领域常用软件包的构建。
2. ROS 运行时：支持 ROS2 humble 版本，集成 ros-core、ros-base、SLAM 等核心包，并提供 ROS SDK，简化嵌入式 ROS 开发。
3. 软实时内核：提供软实时能力，软实时中断响应时延微秒级。

UniProton 硬实时系统

UniProton 是一款实时操作系统，具备极致的低时延和灵活的混合关键性部署特性，可以适用于工业控制场景，既支持微控制器 MCU，也支持算力强的多核 CPU。目前关键能力如下：

- 支持 Cortex-M、Arm64、x86_64 架构，支持 M4、RK3568、RK3588、x86_64、Hi3093、树莓派 4B。
- 支持树莓派 4B、Hi3093、RK3588、x86_64 设备上通过裸金属模式和 openEuler Embedded Linux 混合部署。
- 支持通过 gdb 在 openEuler Embedded Linux 侧远程调试。
- 支持 890+ POSIX 接口，支持文件系统、设备管理、shell 控制台和网络。

应用场景

openEuler Embedded 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。

5. 内核创新

openEuler 内核中的新特性

openEuler 22.03 LTS SP4 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：

体系结构

- **支持 CPU 在线巡检**：静默数据损坏(SDC)可能导致数据丢失和数据被破坏。通过执行巡检指令，发现存在静默故障的核，提前对故障核进行隔离，避免出现更严重的故障，提高系统可靠性。
- **自适应 NUMA 特性**：随着摩尔定律放缓，硬件架构 Scale-up 的收益空间收窄，需要解决资源“locality”的问题，Scale-out 是硬件架构的主要发展方向。对于 NUMA 架构而

言，跨 NUMA 访存时延相比本地访存时延大大增加，而 linux 操作系统以吞吐为中心，强调负载均衡，打破全局跨 NUMA 访存最优的状态。本特性是以业务亲缘关系为中心的负载均衡机制：突破 Linux 调度模型中负载均衡的约束，以亲缘关系为中心，在资源未达到瓶颈时，将具有亲缘关系的任务 packing 在一起，减少跨 NUMA 访存。主要技术点包含如下：

- 1) 感知亲缘关系：通过软硬协同，感知线程间网络、内存的亲缘关系。
- 2) 感知资源瓶颈：基于 NUMA 维度的资源瓶颈感知，感知 NUMA 资源是否达到瓶颈。
- 3) 优化选核和负载均衡机制：基于可编程调度的框架，突破 linux 原有负载均衡约束，决策 NUMA 内还是 NUMA 间调度，在没有资源瓶颈前提下，将具有亲缘关系的任务调度到相同 NUMA 上。

调度

- **支持功耗感知调度**：在面向业务层面收集访存带宽，CPU 负载等数据，确保业务关键线程资源得到满足。引入物理拓扑，调压域感知新的维度，减少单 DIE 调频、单 DIE 调压带来的调频降功耗的局限，保障在低负载下能最小化功耗。

具体如下：

- 1) 根据物理拓扑构建逃逸通道，根据 CPU 负载和访存带宽瓶颈自动调节节能级别，低负载集中业务减少功耗，高负载扩散业务的策略保证 QoS。
- 2) 定时器收集负载，带宽等信息，感知访存带宽，避免跨 DIE 访问。
- 3) OS 新增调压域、idle 静态功耗感知、业务标签等机制，实现智能感知调频、CPU idle、DVFS，来优化业务性能。
- 4) 提供一套任务标签化的机制来为每个用户态进程来标注自己的 QoS 等级，并且支持将标记为不同优先等级的任务分别绑定到不同的分区上进行调度，以此来保证业务 QoS，每个 QoS 的功耗策略也可以通过 cpufreq 模块进行配置，以达到不影响业务 QoS 的前提下保障节能的需求。

内存

- **cgroup 粒度的 KSM 使能增强**：通过 memory.ksm 接口为 memcg 中的进程使能 KSM 时，仅配置已在 memcg 中的进程，后续新加入 memcg 的进程不会自动使能 KSM，需要重新调用该接口使能。该特性支持 memcg 的 KSM 状态位，向 memory.ksm 接口写入 1 后，新加入 memcg 的进程自动使能 KSM，简化了为进程使能 KSM 的操作。
- **支持地址随机化区域指定**：内核增加 CONFIG_NOKASLR_MEM_RANGE 选项，允许用户指定至多 4 个物理内存区域，系统在启动时避开这 4 个内存区域进行物理地址随机化。arm64 和 x86_64 架构的物理地址随机化支持从启动参数指定不随机化的范围，避免跟 kdump 预留内存冲突，解决内存多占用导致内存不足问题。
- **bufferio 限速自动绑定 memcg 和 blkcg**：该特性由美团贡献到 openEuler 社区。当前，cgroup v1 中的子系统独立，无法协同工作，如 blkio cgroup 在限制资源时无法感知 memory cgroup 中的资源使用，需手动创建映射关系。改进特性允许在内核态自动绑定同一进程的 blkio cgroup 和 memory cgroup 映射关系。用户可在编译阶段通过 CONFIG_CGROUP_V1_BIND_BLKCG_MEMCG 开关选择启用，也可在系统运行时通过 sysctl_bind_memcg_blkcg_enable 动态开关调整。

虚拟化

- **KVM TDP MMU**：该特性由 Intel 贡献到 openEuler 社区，是 Linux 5.10 之后，内存虚拟化领域中用于提升 KVM 可扩展性的一个改进方案。相对于传统的 KVM MMU，TDP MMU 提供了更高效的并发 Page Fault 处理机制，从而使得 KVM 对大型虚拟机（多 vCPU，大内存）有了更好的支持。同时，通过采用新的 EPT/NPT 遍历接口，KVM TDP MMU 摒弃了传统内存虚拟化中对 rmap 数据结构的依赖，使得主机内存有了更好的利用率。继 openEuler 22.03-SP3 中提供了 Linux 5.10 至 5.15 版本的 KVM TDP MMU 支持之后，openEuler 22.03-SP4 版本增加了 Linux 5.15 至 6.2 版本间对 KVM MMU 的后续优化和 bugfix，如提供优化大幅降低 KVM MMU unloading 的次数以及 unloading 的流程，优化 KVM MMU page fault 的处理流程，以及 memslot 更新时的相关优化等。
- **Dirty ring**：该特性由麒麟软件贡献到 openEuler 社区。原始内核中对 vcpu 访问内存

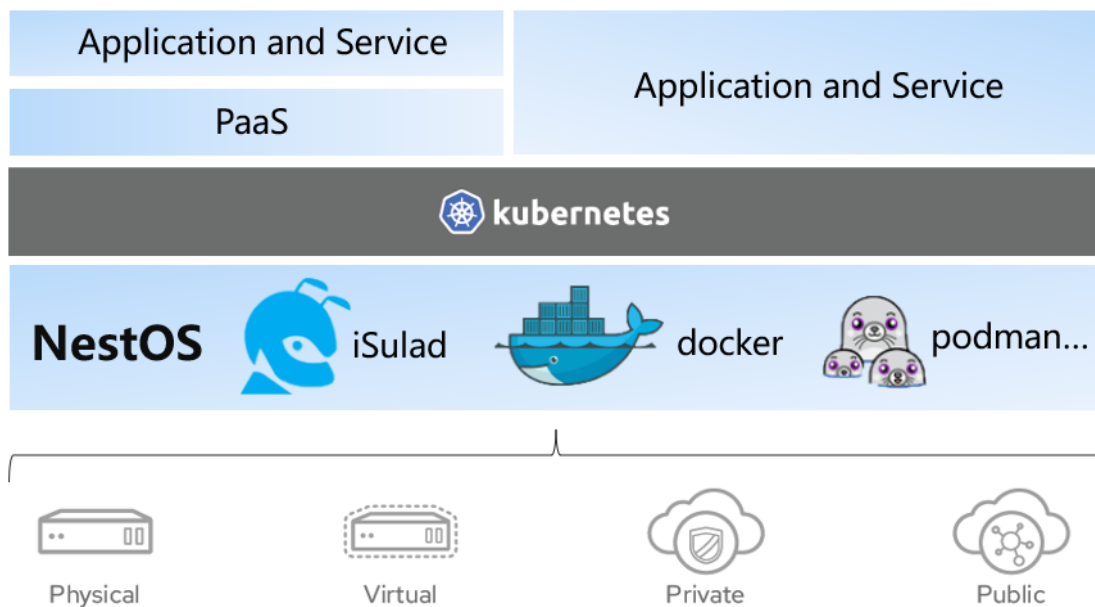
页的统计是通过 dirty bitmap 来实现的，dirty bitmap 的统计粒度是一个 kvm_mem_slot。而 dirty ring 的统计粒度更细，可以精确到一个 vcpu 甚至一个内存页，并且统计可以在用户空间进行，dirty ring 的扩展性较好，可以在大内存规格的虚拟机上开启此特性。

6. 云化基座

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建能力，可以更安全的运行大规模的容器化工作负载。

功能描述



1. 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。
2. 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。

3. 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。
4. 友好可控的更新机制：NestOS 使用 zncati 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。
5. 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

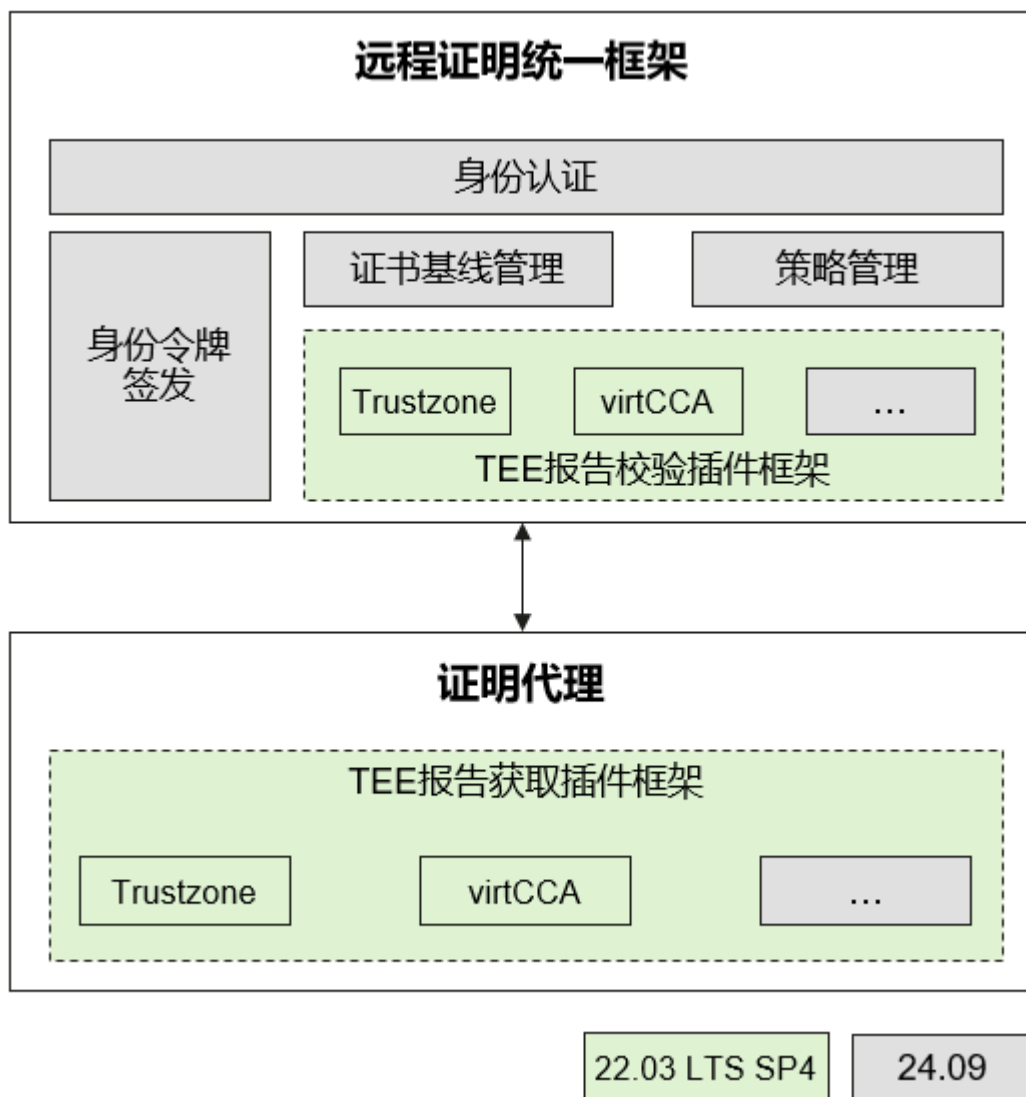
NestOS 适合作为以容器化应用为主的云场景基础运行环境，在本次 openEuler 24.03 LTS 版本中，引入社区孵化项目 NestOS-Kubernetes-Deployer，辅助 NestOS 解决在使用容器技术与容器编排技术实现业务发布、运维时与底层环境高度解耦而带来的运维技术栈不统一，运维平台重复建设等问题，保证了业务与底座操作系统运维的一致性。

7. 特性增强

机密计算远程证明服务：远程证明 TEE 插件框架

随着机密计算的逐渐成熟，机密计算硬件众多，信任根及远程证明差异，成为阻塞 TEE 之间建立信任关系的关键，构建远程证明统一框架，提供证明代理、TEE 插件框架等组件，实现不同 TEE 证明报告统一验证流程。

功能描述



- 如图所示当前版本证明代理支持鲲鹏 Trustzone、virtCCA 证明报告获取。
- TEE 报告校验插件框架支持运行时兼容鲲鹏 Trustzone、virtCCA 证明报告验证，并支持证明代理集成 TEE 报告校验插件框架，实现无证明服务时的点对点验证。

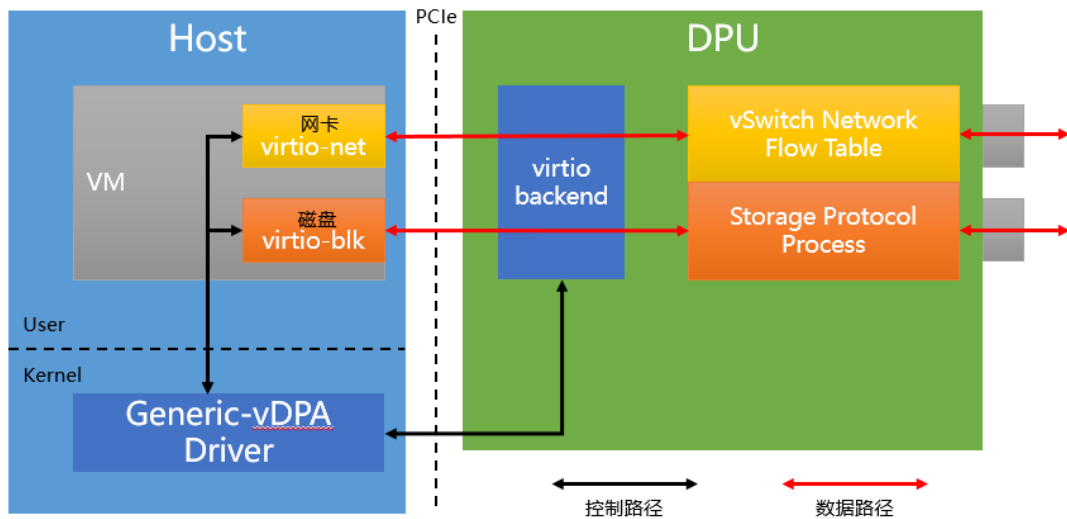
应用场景

挑战者对安全应用发起挑战，安全应用集成远程证明代理获取 TEE 远程证明报告并返回给挑战者，挑战者集成远程证明代理校验证明报告。

DPU 场景 vDPA 新增支持磁盘，并支持配置 vDPA 网卡和磁盘的虚拟机热迁移

内核态 vDPA 框架，为设备虚拟化提供了一种性能与直通持平，且支持跨硬件厂商热迁移的方案。新增对于 DPU 卡形态的硬件支持，支持 DPU 卡呈现给前端主机的网络/存储设备，接入通用 vDPA 框架，并支持热迁移。

功能描述



DPU 场景 vDPA 需求继承 openEuler 2203 LTS SP1/SP3 支持 vDPA 网卡设备以及支持 vDPA 网卡直通热迁移特性。新增支持对于 DPU 卡形态呈现存储设备的支持，并继承支持 vDPA 磁盘直通热迁移特性。

详细范围说明如下：

- 支持虚拟机配置 vDPA 磁盘作为数据盘。
- 支持虚拟机热插拔 vDPA 磁盘。
- 支持同时配置 vDPA 网卡和 vDPA 磁盘热迁移。

应用场景

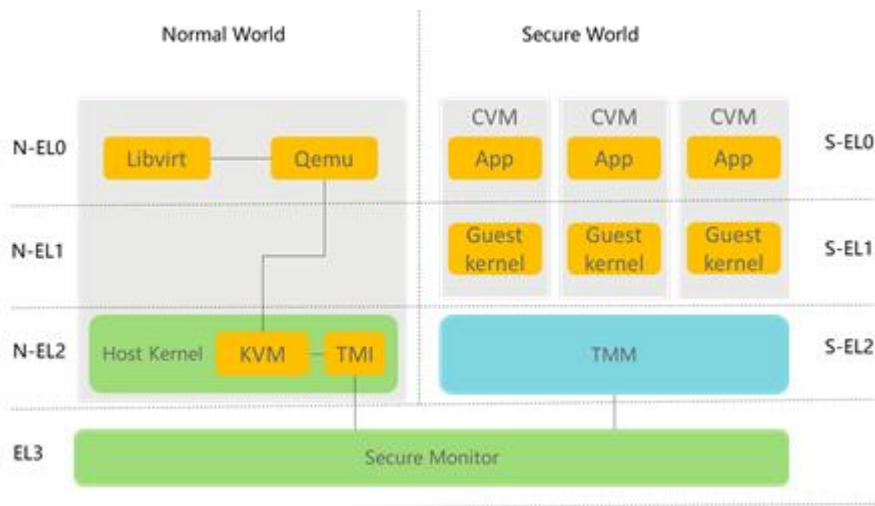
通用 vDPA 框架可以应用于云计算中 DPU 虚拟化全卸载场景。通过 generic-vDPA 框架

实现对于磁盘/网卡等多种 virtio 设备的支持，同时消减了用户态 vDPA 设备对于 dpdk 的依赖，减少全卸载场景下前端主机的资源占用以及管控面的架构复杂度。

virtCCA 机密虚拟机特性合入

virtCCA 机密虚拟机特性基于鲲鹏 920 高性能版 S-EL2 能力, 在 TEE 侧实现机密虚拟机能力, 实现现有普通虚拟机中的软件栈无缝迁移到机密环境中。

功能描述



基于 Arm CCA 标准接口，在 Trustzone 固件基础上构建 TEE 虚拟化管理模块，实现机密虚拟机间的内存隔离、上下文管理、生命周期管理和页表管理等机制，支持客户应用无缝迁移到机密计算环境中。

技术约束:

1. 机密虚拟机每次启动需要对 Guest Kernel 等内存信息进行度量, 因此不支持 reboot 能力。
2. 基于安全性考虑, 宿主机 BIOS 中 CPU 超线程需要关闭。

应用场景

可用于通用应用场景、云化场景以及数据隐私保护等各种场景。

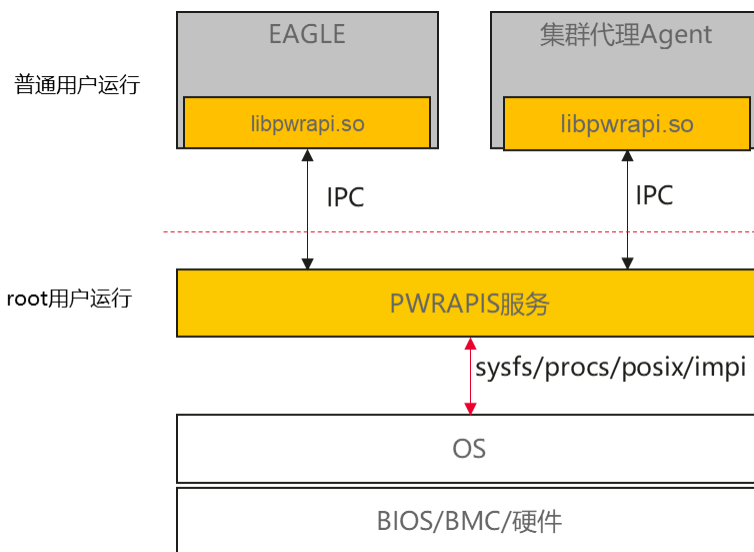
PowerAPI 功耗管理统一 API

PowerAPI 是一组用于支持系统功率管理的 API，它们提供了一种标准化方法来管理系统的功率使用情况，包括监控，调整和优化系统的功率消耗。这些 API 可以帮助系统管理员更好地管理系统的能源消耗，从而提升系统的效率和可靠性，并减少能源成本。

功能描述

如下图所示，PowerAPI 含 pwrapis 服务程序和 libpwrapi.so 动态库。PowerAPI 实际上是对 OS 提供的各类接口的统一封装，方便上层管理关键对功耗进行统一管理，其目标为：

- 屏蔽底层 sci/sysfs/procs/impi 等接口差异，屏蔽平台差异。
- 屏蔽功耗管理场景下，root 特权需求(如集群代理 Agent 不要求采用 root 用户运行)。



PowerAPI 接口已提供功能：

- 数据采集：系统功耗、CPU 利用率、LLC cache Miss 等。
- CPU 调频、CPU 休眠管理。
- 瓦特调度管理、分域调度管理 。

应用场景

希望降本增效，对服务器能效有要求的客户和业务场景。比如集群调度场景下，集群的

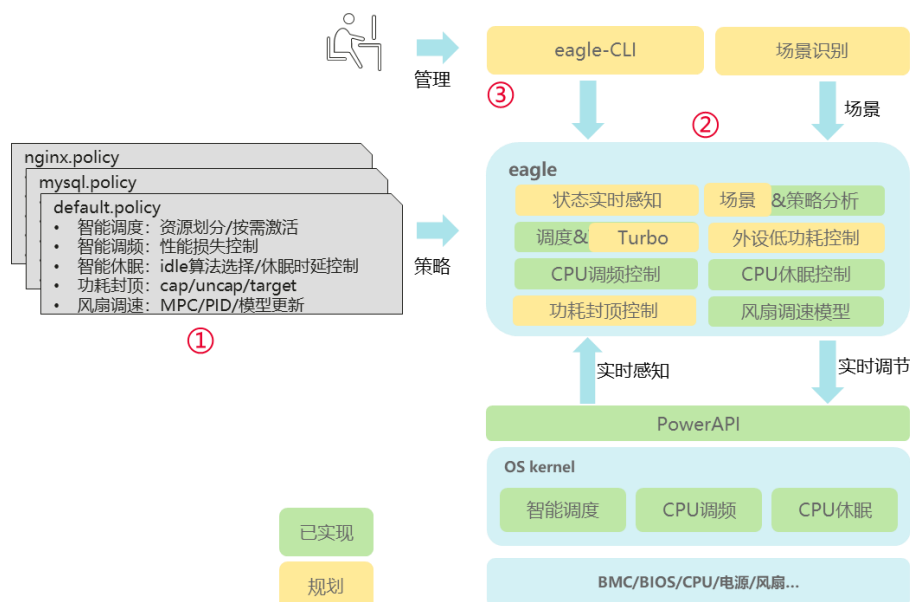
agent 可集成 PowerAPI，根据任务的繁忙程度对当前节点的功耗进行控制。

eagle 实现整机能耗管理

Eagle（Energy Aware intelliGent scheduler）智能功耗感知调节器是一个以提升能效为目标的调优服务。用户可通过配置策略的方式，让 eagle 实现整机功耗管理，进而提升能效，实现降本增效。

功能描述

Eagle 基于 PowerAPI 构建，实现基于策略的能效调优。



①调优策略：用户可根据业务情况，定制调优策略。

②智能场景感知功耗调节：eagle 通过 powerapi 实时感知并结合场景特点(CPU 密集/访存密集/时延敏感)，应用优先级，通过智能调度、调频、Turbo、休眠等调节手段，保证业务体验的前提下，自适应地实现能效最优。

- 基于能效域的智能调度：通过划分能效域，按需激活资源、均衡调度、智能 Turbo 等实现能效最优。
- 性能损失可控的调频：通过实时感知业务性能变化，严格控制 CPU 调频导致的性能下降。
- 更精确的智能休眠：实时感知业务特点，制定更适合业务的 CPU、外设休眠策略。

- **多维度功耗封顶控制：**在功耗限额下，通过多维度、多设备的控制，保证高优先级应用的前提下，实现能效最优。
 - **基于预测的风扇调速：**通过建模，智能预测整机热变化，风扇提前响应，提升整机热可靠性，节省能耗。
- ③**人机交互：**管理员可通过 eagle-CLI 实时监控系统状态，或对调优策略等进行配置。

应用场景

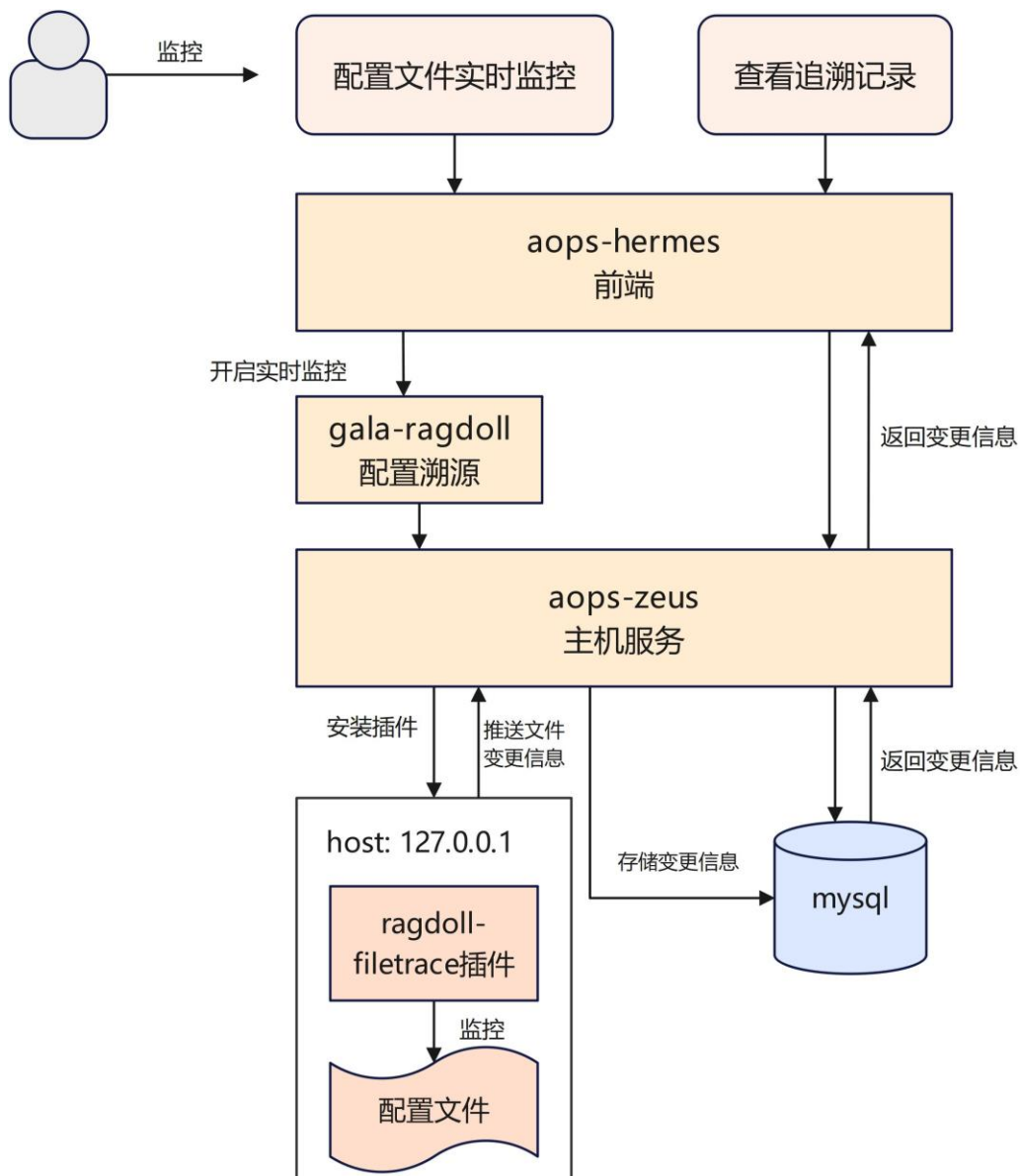
业务负载较低（一般 < 30%），希望降本增效，对服务器能效有要求的客户和业务场景。

gala-ragdoll 支持配置变更溯源

监控和记录节点上配置文件被修改的信息，可以帮助运维人员快速追踪变更配置文件的信息。

功能描述

利用 ragdoll 内部插件 ragdoll-filetrace 可以实现对配置文件的实时监控，当配置文件被修改的时候，会抓取操作文件的系统调用的信息保存到 mysql 数据库表中，ragdoll 通过接口形式将数据库的数据提供给 aops-hermes，用户可以在页面端查看配置文件的变更信息，比如：修改时间，修改方式，进程名称。



目前可以支持监控系统中常用文本编辑工具，比如 vi/vim、sed、echo、mv 等。

应用场景

配置文件在被修改时，无法监控到是谁修改的。为了保证可追溯性，用户在需要对配置文件修改有记录追踪时候，可以对配置域打开监控开关。当主机中配置文件被修改的时候，会追溯到相关的进程以及子进程是哪个，这些数据会存入数据库中，用户可以在“业务域详情 > 当前配置”页面中追踪到这些进程信息。

7. 著作权说明

openEuler 白皮书所载的所有材料或内容受版权法的保护，所有版权由 openEuler 社区拥有，但注明引用其他方的内容除外。未经 openEuler 社区或其他方事先书面许可，任何人不得将 openEuler 白皮书上的任何内容以任何方式进行复制、经销、翻印、传播、以超级链路连接或传送、以镜像法载入其他服务器上、存储于信息检索系统或者其他任何商业目的的使用，但对于非商业目的的、用户使用的下载或打印（条件是不得修改，且须保留该材料中的版权说明或其他所有权的说明）除外。

8. 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。

9. 附录

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