

# 基于北海云计算试验平台的虚拟化技术探索

中国电信研究院  
李天正

1

北海云计算试验平台介绍

2

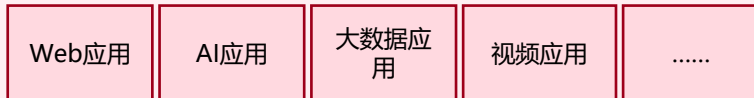
虚拟机管理技术探索与实践

针对问题：RISC-V云计算软硬件生态不成熟、缺乏规模化的应用验证。

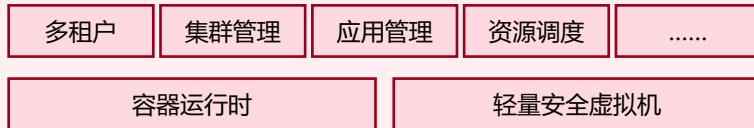
目标：以中国电信丰富应用场景牵引，推动RISC-V架构在云计算领域的成熟与规模应用。

## 北海RISC-V云计算试验平台架构

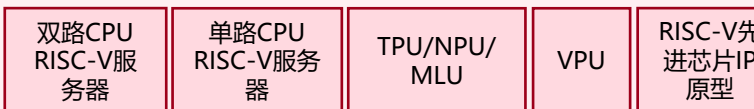
### ③ RISC-V云计算应用迁移适配与示范应用



### ② RISC-V云计算资源池化纳管试验



### ① RISC-V云计算基础设施研发试验



## 依托中国电信研究院RISC-V大科创装置



双路CPU RISC-V服务器



单路CPU RISC-V服务器



已构建运营商首个超千核RISC-V云计算集群



2024 RISC-V中国峰会发布

人民日报·正文

### 中国电信“北海”RISC-V云计算试验平台发布

发布时间：2024-08-26 08:32 来源：人民日报 作者：冀研

A A | 分享

8月21日，在第四届RISC-V中国峰会上，中国电信研究院、天翼云科技公司联合中国科学院软件所、达摩院等合作伙伴发布业界首个RISCV云计算试验平台“北海”，RVEI战略指导委员会主任、中国工程院院士倪光南共同参与发布。RISC-V指令集以其开源开放、精简高效、模块化可定制等技术优势，逐步从物联网、终端向高性能计算领域迈进。

中国电信研究院副院长傅志仁在发布会主题演讲中指出，依托中国电信云网融合大科创装置已完成“北海”平台部署，该平台以RISC-V CPU为核心，同时支持TPU、NPU、VPU

相关成果得到媒体报道

中国电信研究院研发首个RISC-V视频转码卡TeleVPU，内置高性能视频编解码加速器和AI加速器，具备“转码+AI”双能力。最大可降低视频大小80%以上，减少30%~50%的视频存储成本。

- 提供高效视频转码和AI分析能力，双宽全长全高PCIe卡形态，可集成到现有的2U、4U机架式服务器中，单服务器可插8卡。



VPU加速卡



VPU+服务器

- 软件形态：提供FFmpeg标准接口和自有接口，包含编码、解码、缩放和拼接、视频压缩、AI等不同视频处理软件接口。

## SDK软件

编码接口

解码接口

缩放接口

拼接接口

压缩接口

AI接口

中国电信研究院联合产业合作伙伴，依托北海平台开展RISC-V AI智算加速卡的大模型应用迁移适配验证试验及整体解决方案研发。目前已完成中国电信自研大语言模型TeleChat适配。

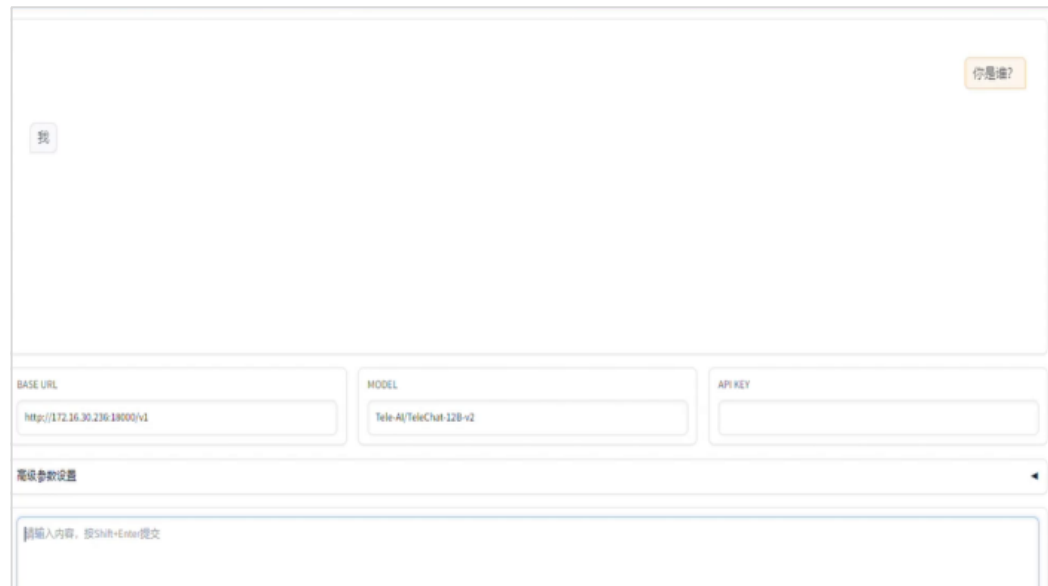
## RISC-V AI芯片国外进展

- **Tenstorrent**: 推出768核RISC-V AI芯片BlackHole，性能高达745 TOPS；
- **Meta**: 推出RISC-V AI芯片MTIAv2，稀疏算力达708TFLOPS。

## RISC-V AI智算加速卡应用进展

- **完成中国电信TeleChat大模型适配**:
  - AI卡配置：国产RISC-V AI卡
  - AI大模型配置：TeleChat 12B，FP格式（不进行量化）

### TeleChat运行效果



## “北海”平台RISC-V AI智算卡方案试验

目前完成多种RISC-V推理卡、大显存训练卡试验验证，RISC-V在推理、微调性价比有优势。

大模型推理卡

大模型训练卡1

大模型训练卡2

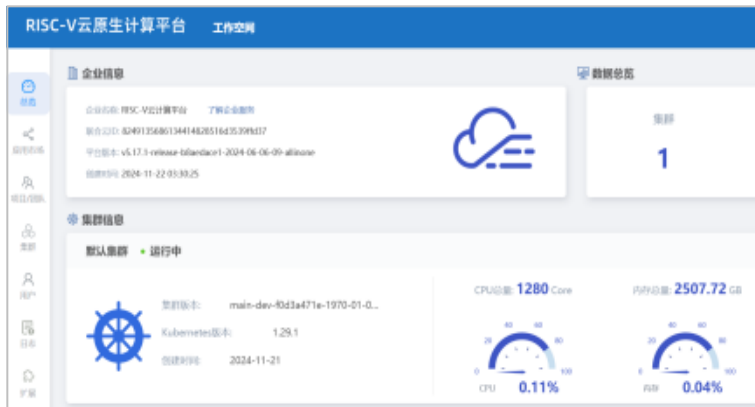
厂商A

厂商B

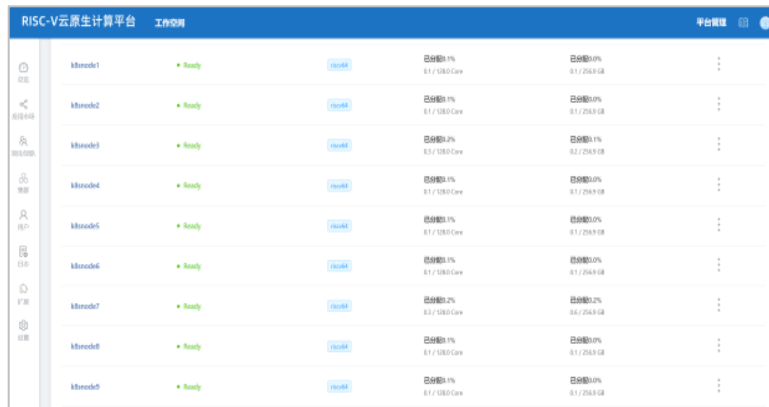
厂商C

# RISC-V云原生解决方案

**RISC-V云原生管理平台：基于国产RISC-V服务器集群构建RISC-V云原生管理平台，实现超千核服务器纳管，完成集群管理、应用管理、应用市场等主流商用级功能验证。**



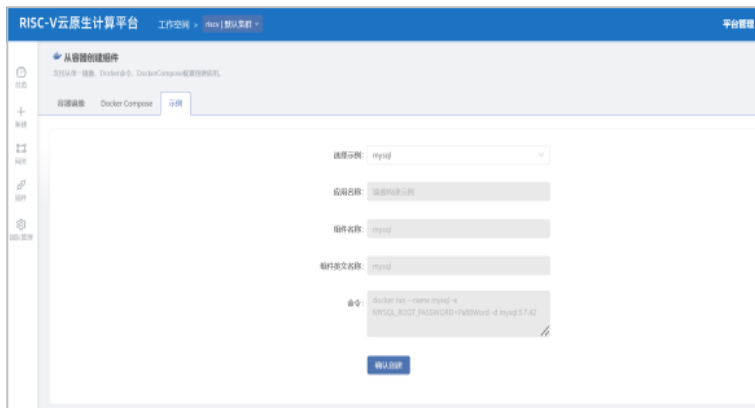
RISC-V云平台集群管理界面



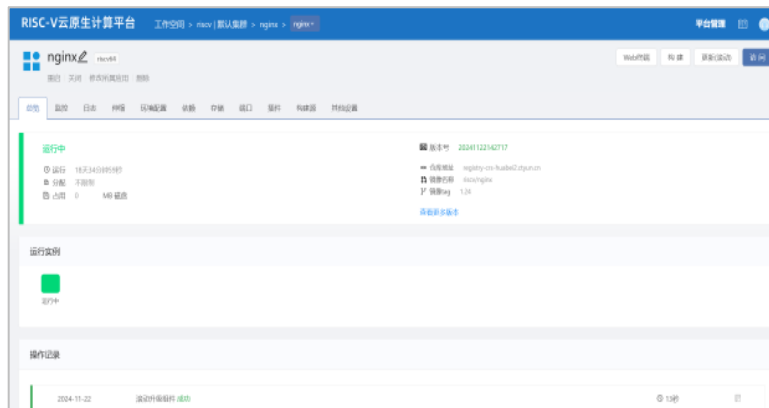
RISC-V云平台节点管理界面



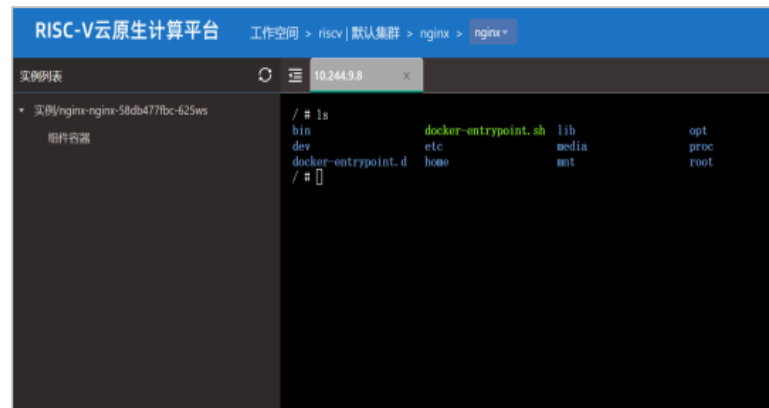
RISC-V云平台节点监控界面



RISC-V云平台应用创建界面



RISC-V云平台应用管理界面

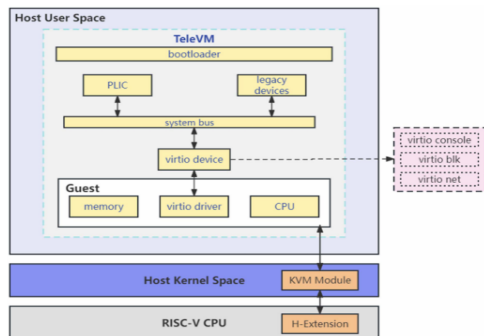


RISC-V云平台容器实例访问

**RISC-V轻量级虚拟机：创新性提出支持RISC-V虚拟化扩展的云原生轻量级虚拟机TeleVM，并完成与多种国产RISC-V IP核的联合验证。相比传统虚拟机内存消耗降低90%，启动时延降低80%。**

## RISC-V轻量级虚拟机

### ■ 首个支持RISC-V H扩展的云原生轻量安全虚拟机TeleVM



### ■ 完成了OpenEuler+TeleVM+香山开源芯片仿真验证

**OpenEuler+TeleVM+香山开源芯片仿真平台**

- TeleVM虚拟机二进制文件
- OpenEuler操作系统镜像
- 香山RISC-V芯片H扩展仿真平台

## RISC-V虚拟机获得业界认可

### ■ 成为EulixOS操作系统的虚拟化引擎



**Televm (Stratovirt) 支持:** Televm 作为基于 Stratovirt 的轻量虚拟化方案，是首个基于 RISC-V 架构实现 H 扩展虚拟化的软件。它显著提升了虚拟化性能和资源管理效率，为云计算环境中的高效虚拟化提供了坚实基础。

**Istio 首次支持 RISC-V:** Istio 服务网格作为微服务架构中的重要组件，首次在 RISC-V 平台上获得了完整支持，进一步推动了 RISC-V 在微服务架构中的应用。

1

北海云计算试验平台介绍

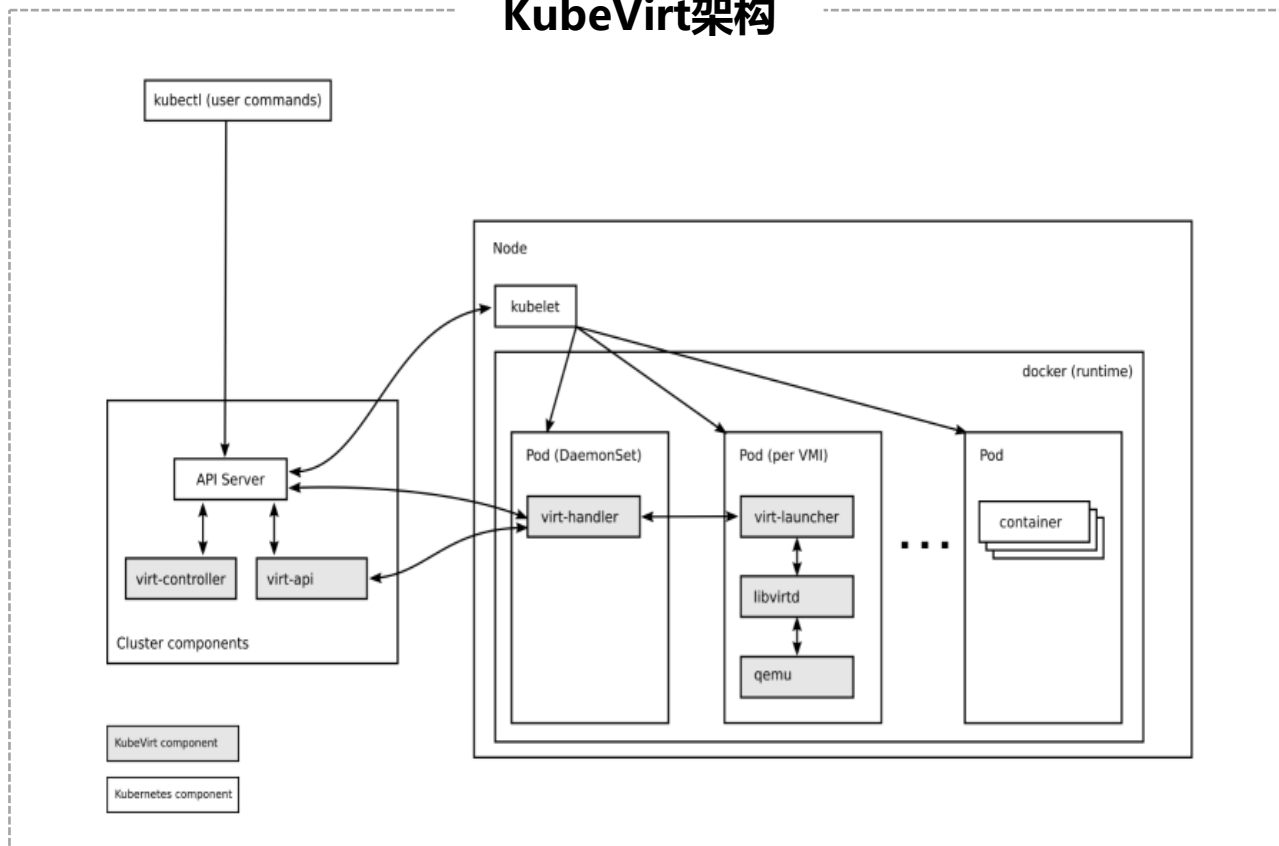
2

虚拟机管理技术探索与实践

# 目标：容器和虚拟机的统一管理

**KubeVirt虚拟机管理方案：通过扩展 Kubernetes 的功能，使虚拟机能够像容器一样被管理、调度和编排，从而实现容器和虚拟机的混合部署与管理，满足不同应用场景的需求。**

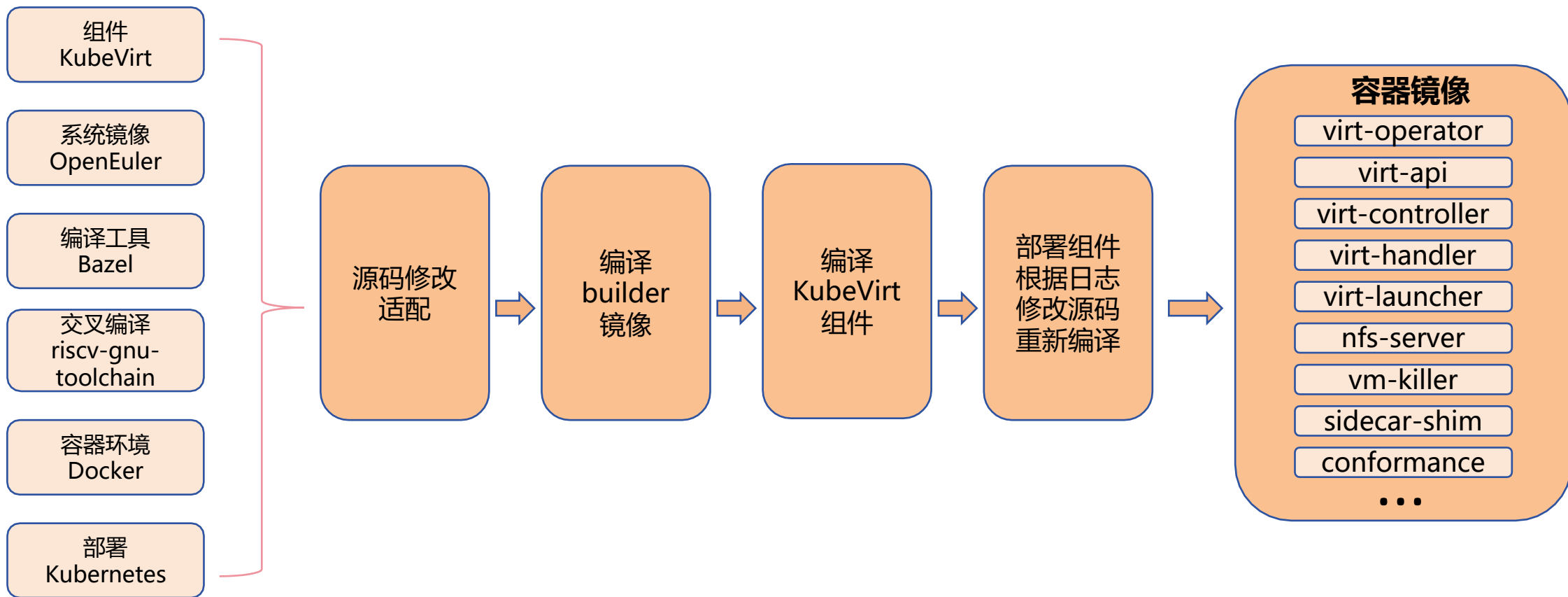
## KubeVirt架构



## 核心组件

- **virt-controller:** 作为控制器，负责监听和管理整个集群中虚拟机相关资源的状态。
- **virt-api:** 提供 RESTful API 接口，用于接收和处理用户对 KubeVirt 虚拟机资源的管理请求。
- **virt-handler:** 以 DaemonSet 形式运行在每个节点上，处理节点上的虚拟机生命周期操作，并与容器运行时交互。
- **virt-launcher:** 运行在每个虚拟机实例对应的 Pod 中，作为实际启动和管理虚拟机进程的引擎。

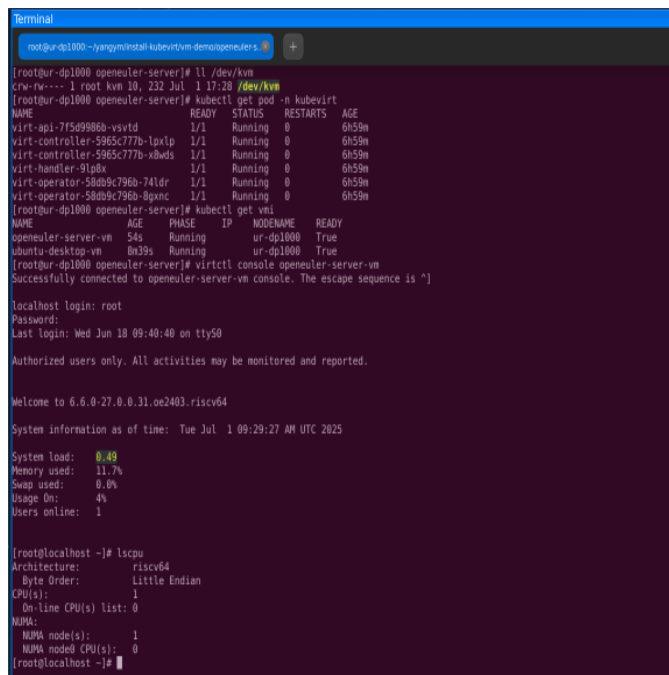
通过在x86环境使用交叉编译方式编译完成KubeVirt所有组件，成功实现KubeVirt对RISC-V架构的适配。



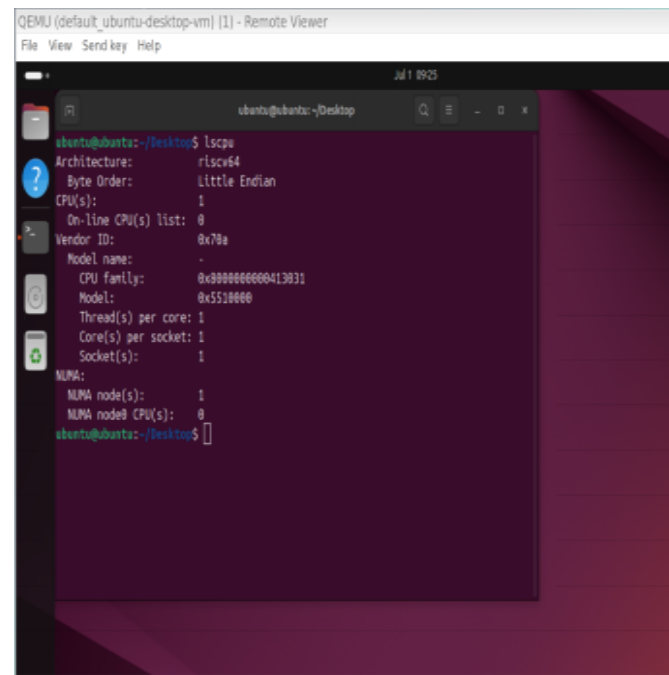
北海平台新增支持H扩展的RISC-V板卡，基于现有Kubernetes集群部署KubeVirt，成功启动OpenEuler和Ubuntu虚拟机，实现业界首个RISC-V容器和虚拟机资源的统一部署与管理。



支持虚拟化扩展的RISC-V板卡



启动无图形界面的OpenEuler虚拟机



启动图形界面的Ubuntu虚拟机

# RISC-V硬件虚拟化可实现性能大幅提升

Coremark性能测试：经过测试，使用Qemu纯软件模拟方案性能损失约85%，使用Qemu+KVM硬件加速方案性能损失约6%。

## Coremark测试

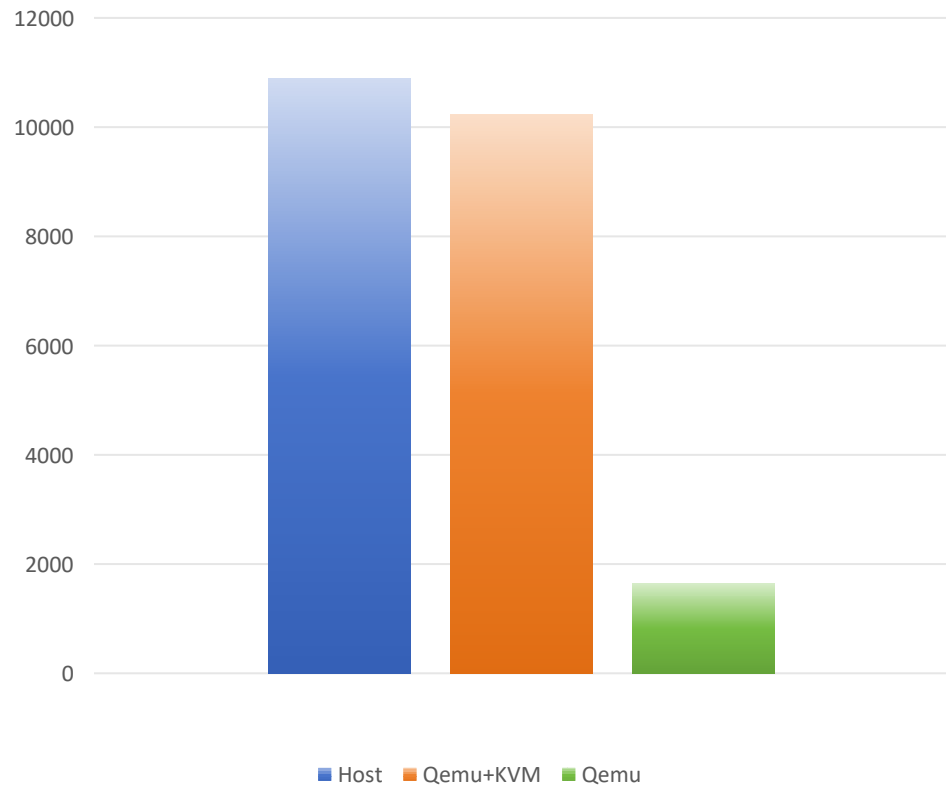
```
[root@localhost ~]# ./coremark
OK performance run parameters for coremark.
CoreMark Size : 666
Total ticks : 19548
Total time (secs): 19.548099
Iterations/Sec : 18235.414534
Iterations : 288660
Compiler version : GCC12.3.1 (openEuler 12.3.1-30.oe2483)
Compiler flags : -O2 -DPERFORMANCE_RUN=1 -DMULTITHREAD=1 -DUSE_PTHREAD -DPERFORMANCE_RUN=1 -lrt
Memory location : Please put data memory location here
(e.g. code in flash, data on heap etc)
seedcrc : 0xe9f5
0)crclist : 0xe714
1)crcmatrix : 0x1fd7
2)crystate : 0x0e3a
3)crfinal : 0x4963
Correct operation validated. See README.md for run and reporting rules.
CoreMark 1.0 : 18235.414534 / GCC12.3.1 (openEuler 12.3.1-30.oe2483) -O2 -DPERFORMANCE_RUN=1 -DMULTITHREAD=1 -DUSE_PTHREAD -DPERFORMANCE_RUN=1 -lrt / He
ap
```

RISC-V硬件虚拟化单核CPU跑分

```
[root@localhost ~]# ./coremark
OK performance run parameters for coremark.
CoreMark Size : 666
Total ticks : 18333
Total time (secs): 18.333060
Iterations/Sec : 1636.393389
Iterations : 30099
Compiler version : GCC12.3.1 (openEuler 12.3.1-30.oe2483)
Compiler flags : -O2 -DPERFORMANCE_RUN=1 -DMULTITHREAD=1 -DUSE_PTHREAD -DPERFORMANCE_RUN=1 -lrt
Memory location : Please put data memory location here
(e.g. code in flash, data on heap etc)
seedcrc : 0xe9f5
0)crclist : 0xe714
1)crcmatrix : 0x1fd7
2)crystate : 0x0e3a
3)crfinal : 0x5275
Correct operation validated. See README.md for run and reporting rules.
CoreMark 1.0 : 1636.393389 / GCC12.3.1 (openEuler 12.3.1-30.oe2483) -O2 -DPERFORMANCE_RUN=1 -DMULTITHREAD=1 -DUSE_PTHREAD -DPERFORMANCE_RUN=1 -lrt / He
ap
```

RISC-V软件虚拟化单核CPU跑分

## 性能对比



## 持续更新优化

基于最新版本KubeVirt持续更新优化，跟踪RISC-V标准增加新特性。提高集群稳定性，整体达到可商用级别。

## 多场景RISC-V云化方案

实现GPU、VPU等设备透传虚拟机，在虚拟机搭建相关服务，针对AI、视频处理等场景推出完整的RISC-V云化解决方案。

## 共建RISC-V云计算生态

将北海平台向业界开放，联合合作伙伴共建RISC-V云计算生态，加速RISC-V在云计算场景的落地应用。

感谢聆听!