



# *RISC-V*数据中心软件生态 - 机遇与挑战

中兴通讯 RISC-V生态技术总监 贾云翔

2025-7 上海

X86服务器集采份额连续四年**蝉联第一**

中国移动服务器**第一大供应商**

中国移动2024年X86服务器集采份额**第一**

中国电信2024年X86服务器集采份额**第一**

电信

国有大行**6进4**

**12家**股份制银行全面进入

覆盖全国**50+**城商行/农信社

覆盖全国**20+**证券/保险机构

金融

规模部署重资产框架中标

**字节**定制化服务器超份额落地

**百度**文心一言大模型应用支持

京东/美团/滴滴/快手等**核心供应商**

互联网

市场规模国内前五，电信行业蝉联第一，互联网、金融、电力、交通、教育、大企业等全面应用



电力



交通



教育



政务



大企业

**ISA Extensions**  
Float, Bit, Vector

**non-ISA Extensions**  
Interrupt, IOMMU

**Software Extensions**  
BRS, ABI

**Custom Extensions**  
Vendor specific

**RVA**  
Application Processor

**RVB**  
Embedded Processor

**RVM**  
Microcontroller

Server

Mobile

Automotive

Embedded System

Edge IOT/AI

**Extensions**  
满足多样性

**Profiles**  
解决软件碎片化

**Platforms**  
OS/Firmware开箱即用



1

# 操作系统的支持

分类	关键需求	备注
固件接口	AIA、IOMMU、SMBIOS、LPI、SPCR、APEI、CPPC...	APEI/IOMMU上游尚未支持
内核	AIA、IOMMU、hwprobe、RVV、RVA23指令 (zbb优化位操作、zbc优化crc、Zawrs优化lock、Svtpct优化刷TLB...)	hwprobe上游还未完善
虚拟化	KVM (ONE_REG、PMU、irqbypass、AIA imsic、steal time...) qemu (匹配内核的扩展支持、vector支持、AIA...)	部分特性如中断透传上游尚未支持
编译器	RVA23指令支持、自动向量化相关优化...	
OS基础功能	kdump、livepatch...	
组件生态	RVA23支持优化 (openJDK、Go、openssl、ISA-L...)	

相关标准
RVA23 Profile: 2024 Ratified
Boot and Runtime:Frozen
RPMI: Frozen
RERI: Ratified
Server SoC:2025 Ratified
Server Platform: Under Development



6.16内核基本完善  
欠缺部分服务器特性  
性能优化与稳定性

当前社区版本主要支持RVA20  
OpenEuler24.03LTS 于 24 年 6 月正式支持  
OpenAnolis23 于 25 年 6 月发布初步支持  
OpenCloudOS9 于 25 年 5 月发布预览版本支持

当前商用主力仍然为4.19/5.10基线系列  
主要厂商24-25年发布新基线版本，内核6.6  
尚未有RISC-V通用服务器商业版本

RISC-V 通用服务器场景的落地，构建支持RVA23的**社区发行版**支撑是关键

数据中心业务的多样性，要求高性能的OS版本，x86/arm在主流OS发行版指令集基线均支持向量化，RISC-V发行版当前主要支持RVA20  
数据中心业务需要提升RISC-V基线到**RVA23 profile**

### 数据中心 workload类型

AI/ML

HPC

Bigdata

Video

Storage

Database

Crypto

.....

Profile	向量指令集特性	主流发行版
Intel V1	SSE2	Debian, Ubuntu
Intel V2	增加 SSE3、SSSE3、SSE4.1	OpenEuler, Fedora, openSUSE
Intel V3	增加 AVX、AVX2	Arch Linux, ChromeOS
Intel V4	增加 AVX-512	
ARMv8.0-A	NEON	Debian, Ubuntu, Fedora, Arch Linux
ARMv8.1-A	NEON + FP16 扩展	OpenEuler
ARMv9.0-A	增加SVE & SVE2	
RVA20	无	OpenEuler, Debian, Ubuntu
RVA23	增加RVV1.0	暂无

## RVA23在GCC不同版本支持的扩展清单

GCC12.3

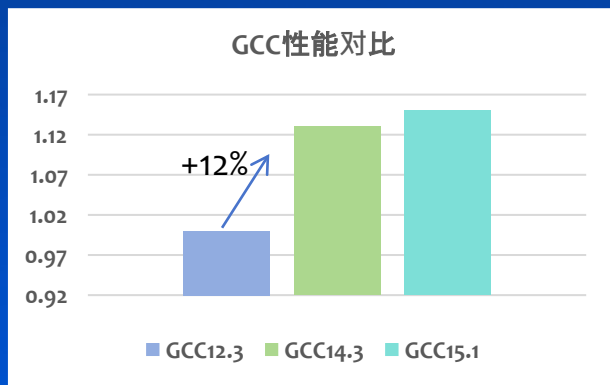
GCC14.3

GCC15.1

Profile		Extensions										
Mandatory	RVA23U64	I	M	A	F	D	C	Zicsr	B	Zicntr	Zihpm	Zkt
		Ziccif	Ziccrse	Ziccamoa	Zicclsm	Za64rs	ZihintPause	Zic64b	Zicbom	Zicbop	Zicboz	Zfhmin
		V	Zvfhmin	Zvbb	Zvkt	Zihintntl	Zicond	Zimop	Zcb	Zfa	Zawrs	Supm
Option	RVA23S64	Zifencei	Svbare	Sv39	Svade	Sscpptr	Sstvecd	Sstvala	Ssconterenw	Svpbmt	Svinval	Shgatpa
		Svnapot	Sstc	Sscopfpmf	Ssnpm	Ssu64xl	Ssstateen	Shcounterenw	Shtvala	Shvstvecd	Shvsatpa	
Option	RVA23U64	Zvkng	Zvksg	Zabha	Zacas	Ziccamoc	Zvbc	Zama16b	Zfh	Zbc	Zicfilp	Zicfiss
		Zvfh	Zfbfmin	Zvfbfmin	Zvfbfwma							
	RVA23S64	Sv48	Sv57	Zkr	Svadu	Sdtrig	Ssstrict	Svpptc	Sspm			

数据中心对于生态软件性能有较高的需要,  
**GCC14.3及更高版本**在特性支持以及性能上都有较大的提升

使用高版本工具链构建软件生态是性能的关键



类别	关键特性	状态
基础功能	AIA 中断虚拟化	完成
	IOMMU 设备透传	完成, RIMT规范已发布, IOMMU ACPI 驱动支持预期很快进入上游
场景拓展	嵌套虚拟化	进行中
	CoVE虚拟机	依赖于 smmmtt、CoVE 规范, 当前处于草稿阶段
性能优化	PMU snapshot	已支持, kernel 6.10
	irq bypass	patch review 中
	Supervisor Software Events	依赖 SBI3.0规范, 当前状态: 已冻结

虚拟化支持正从功能完善不断向性能优化深入, SBI3.0规范将带来更多性能优化特性

## 生态目标

构建完善的 RVA23 社区发行版生态，推进数据中心场景应用落地

## 实施方案

匹配操作系统发行版现状，减小下游生态适配成本

- 保持6.6内核基线，反合高版本特性
- 升级工具链，提升指令集基线，充分发挥硬件性能
- 基于RVA23指令集构建完整社区发行版
- 关键组件（基础库、编程语言及运行时）上游RVA23支持优化

## 当前进展

- 主导推进龙蜥社区 23.3 版本发布对 RISC-V 初步支持
- 联合社区生态伙伴（中科院软件所，阿里达摩院等）制订 RVA23 支持计划，预期社区年底发布版本
- 推进 6.6 内核 RVA23 功能反合，完成 AIA、hwporbe 等支持
- 基于 gcc14.3 完成RVA23U64 支持回合



2

# 数据中心开源生态



分类	开源软件	支持情况
虚拟化和容器化平台	Kubernetes	可运行, 官方未支持
	Docker	可运行
	QEMU	支持较好
	OpenStack	可运行, 官方未支持
存储管理软件	Ceph	支持
	SPDK	支持
网络管理软件	Open vSwitch	支持
	DPDK	支持rv64gc
数据库和大数据平台	PostgreSQL	支持
	Hadoop/Spark	可运行
监控与日志管理工具	Prometheus/Grafana	支持
	Elasticsearch/Logstash/Kibana	可运行
运行时	OpenJDK	jdk21支持rv64gc
	Go	实验性的支持rva22u64
	Rust	nightly版本支持rva23

## 总结

- 大部分已适配RISC-V
  - 缺少官方RISC-V Release版本
  - 缺乏稳定性和性能优化的版本
  - 社区缺少RISC-V的CI/CD基础设施
  - 编程语言运行时还不成熟
- 
- 随着RVA23的芯片上市, 上面的问题会逐步解决
  - 开源生态的性能和稳定性仍然有较大的工作量, 需要大家共同建设

## 如何兼顾兼容性 vs 差异化

Step1: 基于RVA23指令集优化相关基础软件

Step2: 针对性能场景, 通过多实现等技术优化生态软件

Step3: 持续演进

一个二进制可以既可以兼容多种硬件又可以发挥性能优势

Software Bin

一次编译  
多处运行

RVA23  
mandatory

RVA23  
mand+crypto

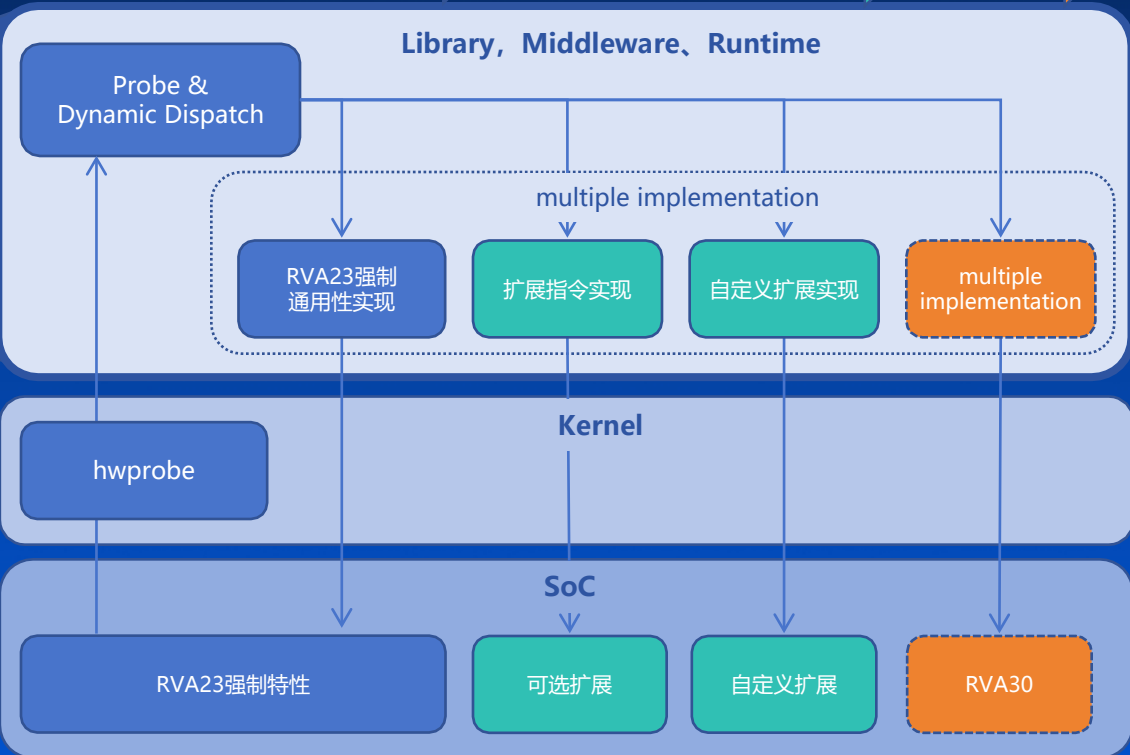
RVA23  
mand+matrix

RVA23.1

RVA30

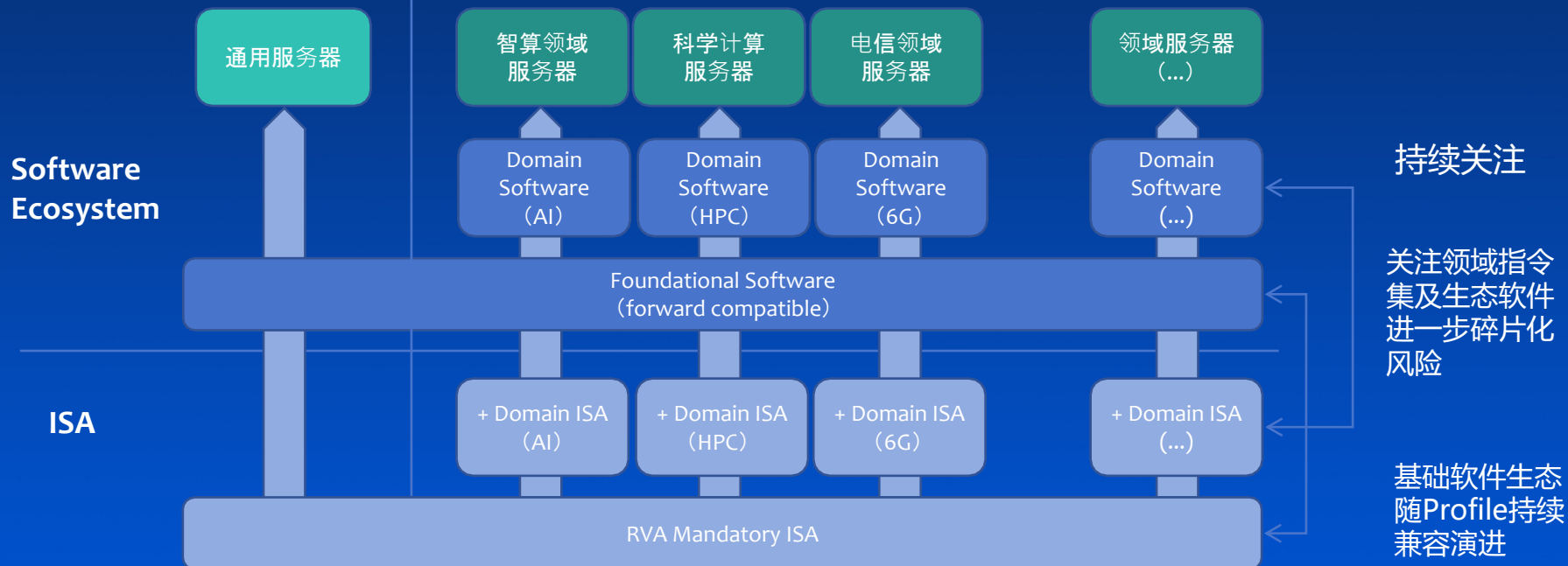
.....

Hardware(多种可能性)



① 围绕RVA23强制特性快速构建基础软件生态

② 基础软件持续兼容性演进，充分发挥RISC-V指令集**多样性**以及**快速演进**的特点，在数据中心细分领域形成标杆解决方案，带动RISC-V在数据中心领域不断发展。



中兴通讯积极参与RISC-V相关开源建设，相关成果也会反馈给开源社区



## ISA-L

CRC算法增强  
- 性能提升19倍

EC算法增强  
- 性能提升16倍

已提交PR到Upstream

## OpenSSL

SM2算法增强  
- 加密：提升19倍  
- 解密：提升2.3倍

已提交PR到Upstream

## DPDK

预取：开发验证中  
向量化：开发验证中

## GCC

SPECCPU：  
- 性能提升：2%

SPEC子项525：  
- 性能提升：15%

## AI相关

vLLM：已适配RISC-V  
Ray：已适配RISC-V  
Triton：已适配RISC-V

开源贡献中

3

机遇与挑战

## 机遇

### 芯片及服务器逐步成熟

- 标准规范逐步完善
- RVA23芯片逐步上市

### 基础软件生态已逐步成熟

- RVA23的OS发行版逐步发布
- GCC/LLVM工具链成熟
- 生态软件适配情况较好

### 软件性能还需提升

- Rust/Java/Go等运行时性能还有提升空间
- 基础库，中间件还需要高性能实现，需要大家共同建设

### 领域相关指令集还在发展

- 数据中心场景还缺少商用案例
- AI Matrix指令集标准还在讨论中
- 关注进一步碎片化的风险

### 认证相关还处于起步阶段

- RVA23 Verification和Certification仍然还有较大gap
- 积极参与CSC相关工作

## 挑战

ZTE中兴

合作共赢 共创RISC-V未来

