

RISC-V KVM的 调试与调优

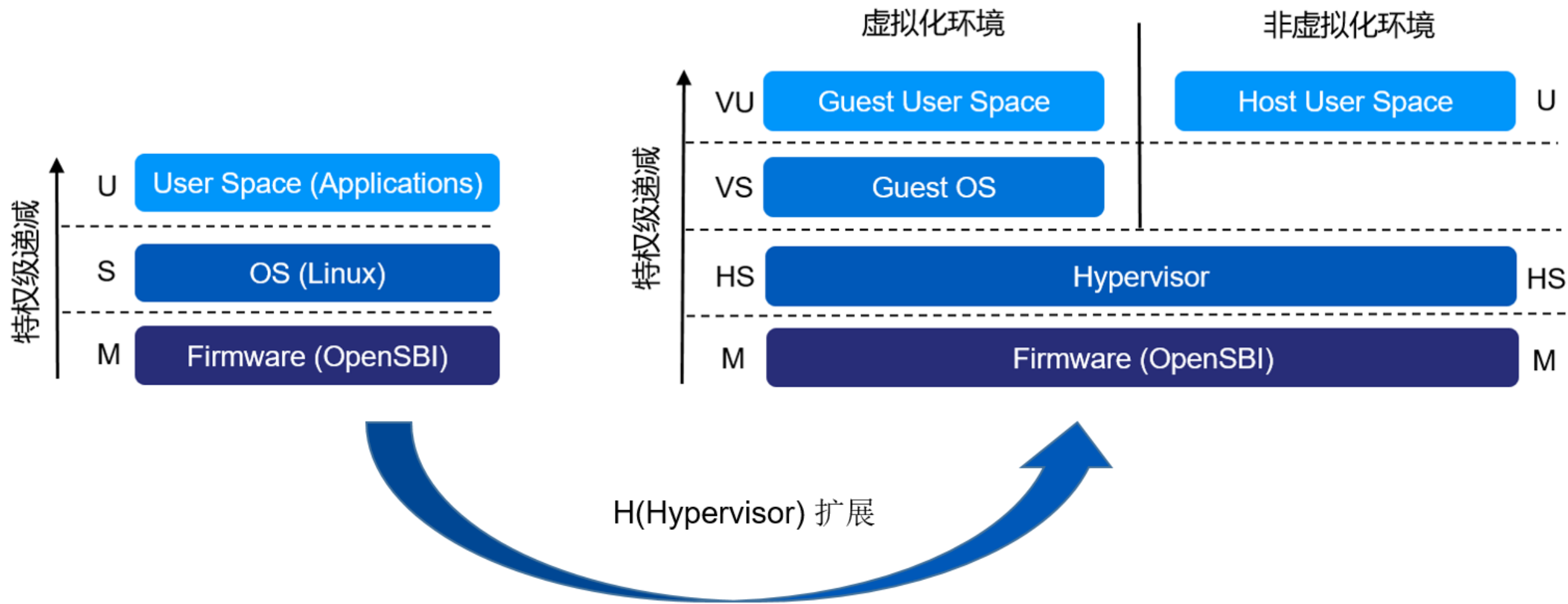
杜超

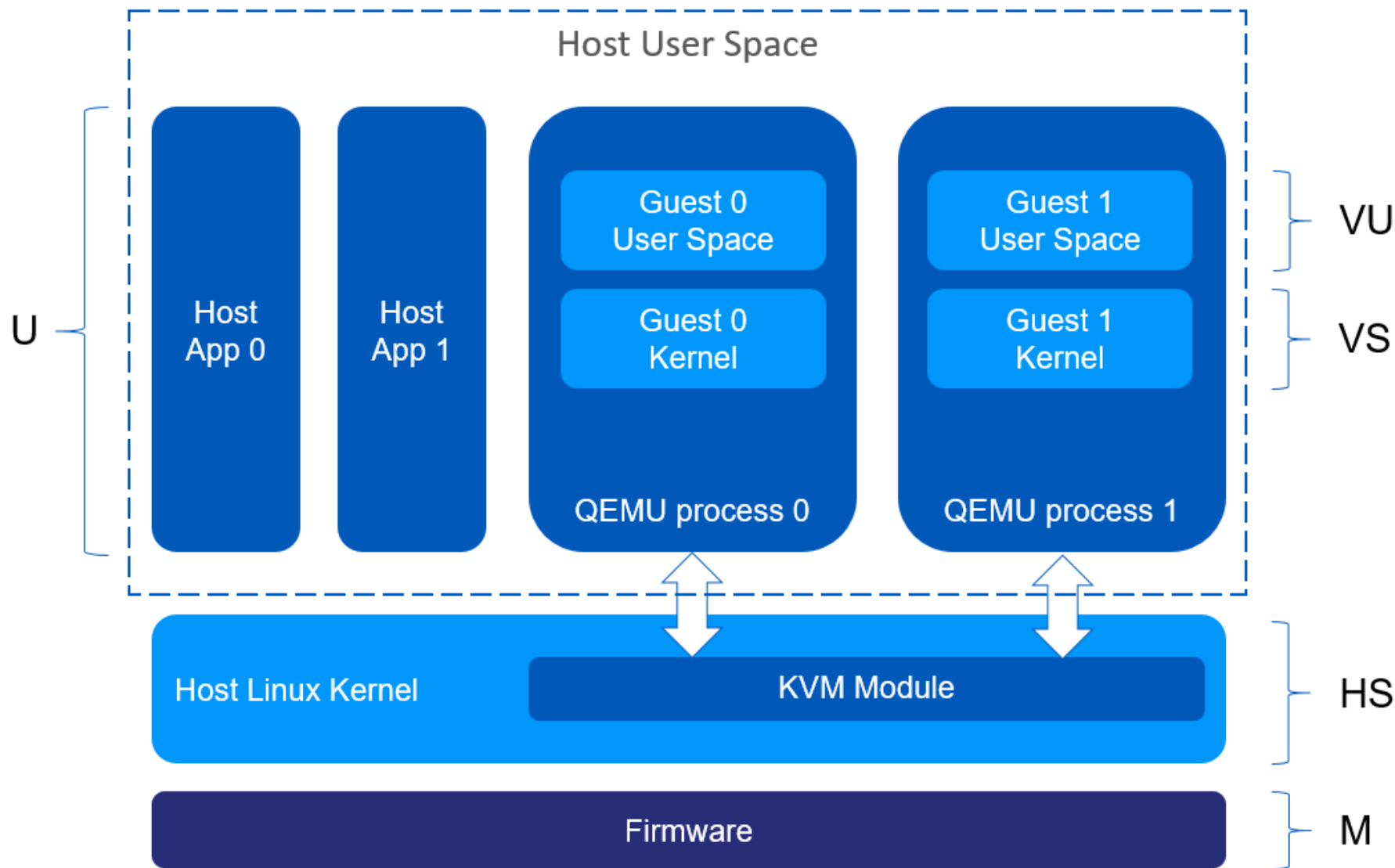
北京奕斯伟计算技术股份有限公司



目录

- I. RISC-V KVM-简介与现状
- II. RISC-V KVM调试-场景与方法
- III. RISC-V KVM调试-实践与展望
- IV. RISC-V KVM调优-实践与展望





软件生态支持情况:

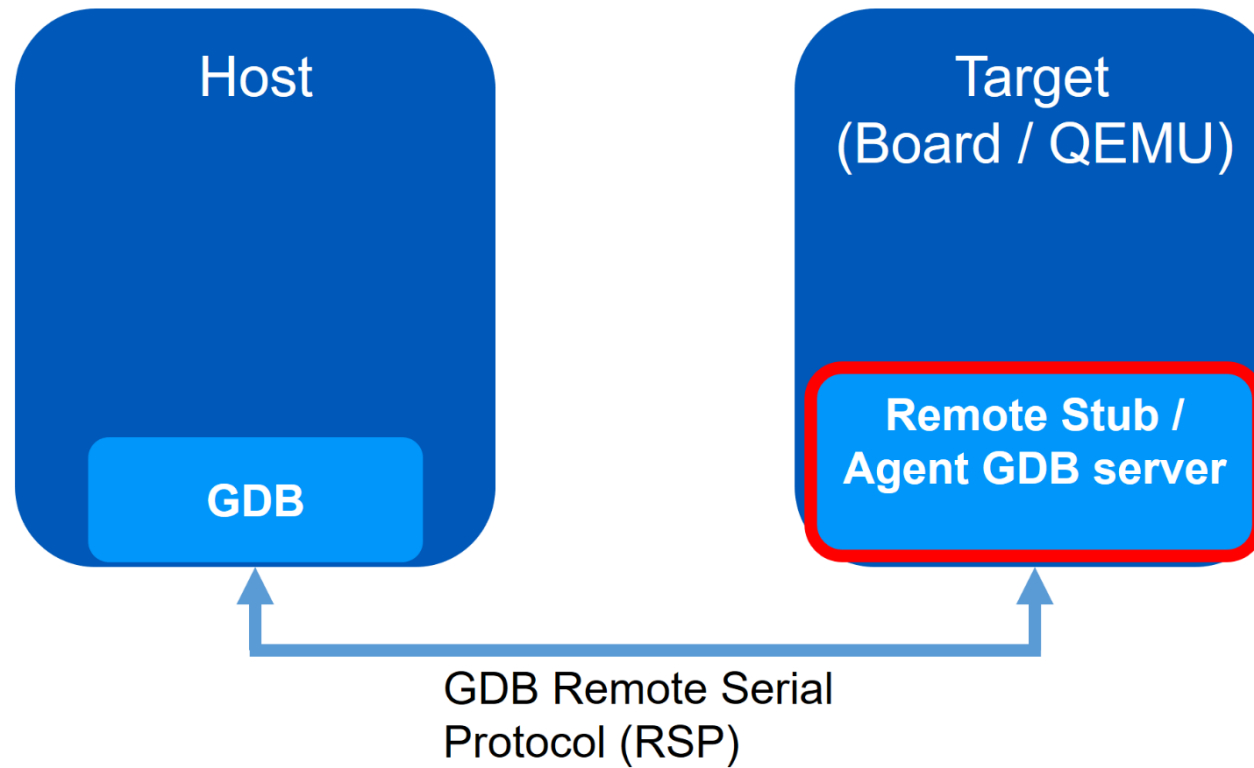
- Linux Kernel: from 5.16
- QEMU: from 7.0.0
- AIA/IOMMU: on-going

目录

- I. RISC-V KVM-简介与现状
- II. RISC-V KVM调试-场景与方法
- III. RISC-V KVM调试-实践与展望
- IV. RISC-V KVM调优-实践与展望

为什么需要调试客户机程序？

- 虚拟化相关驱动调试及Bug Fix
- Guest Kernel调试及Bug Fix
- BIOS / Bootloader调试及Bug Fix
- Hypervisor调试及优化
- 深入理解KVM VM的实际运行情况
-



Target	GDB Remote Stub
Board	OpenOCD
QEMU	TCG Stub
KVM	?

目录

- I. RISC-V KVM-简介与现状
- II. RISC-V KVM调试-场景及方法
- III. RISC-V KVM调试-实践与展望
- IV. RISC-V KVM调优-实践与展望

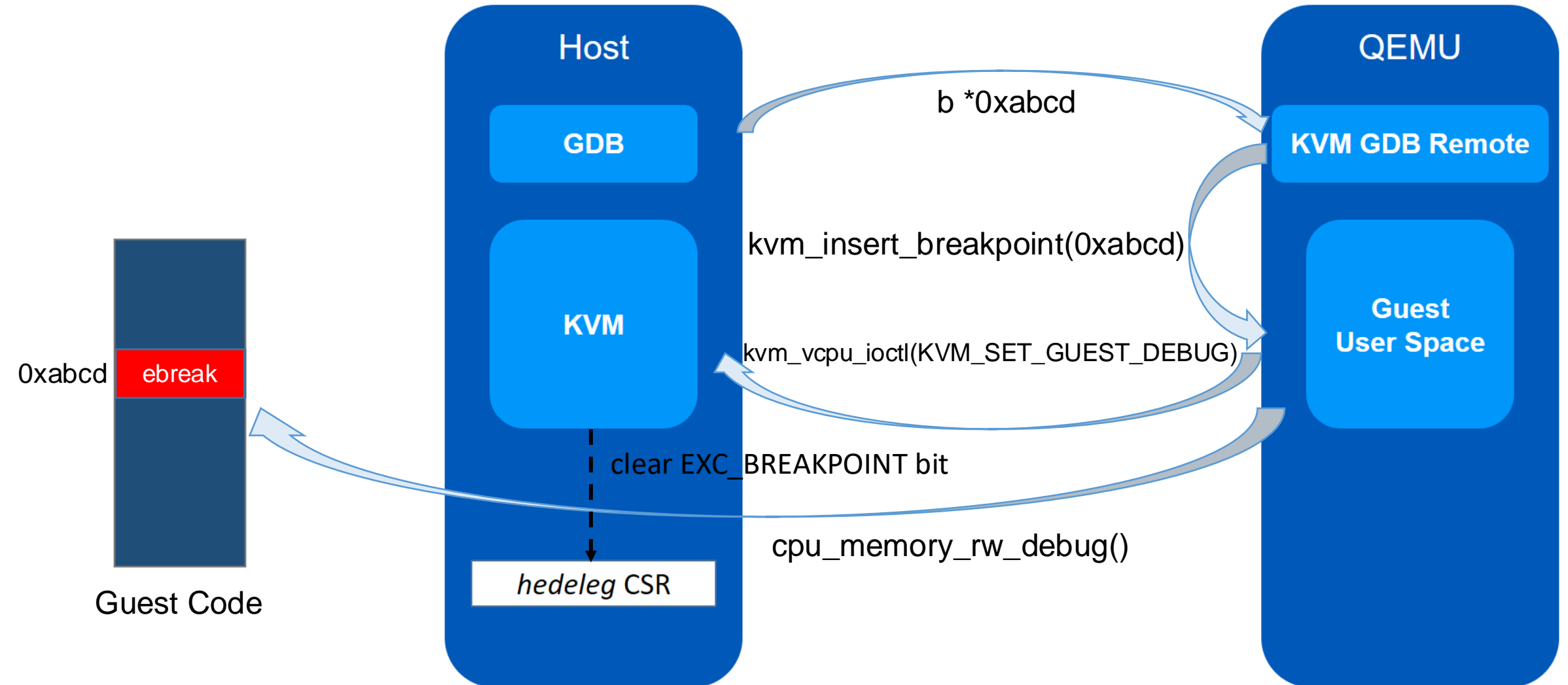
支持RISC-V KVM GDB Remote Stub（也被称为**KVM Guest Debug**）：

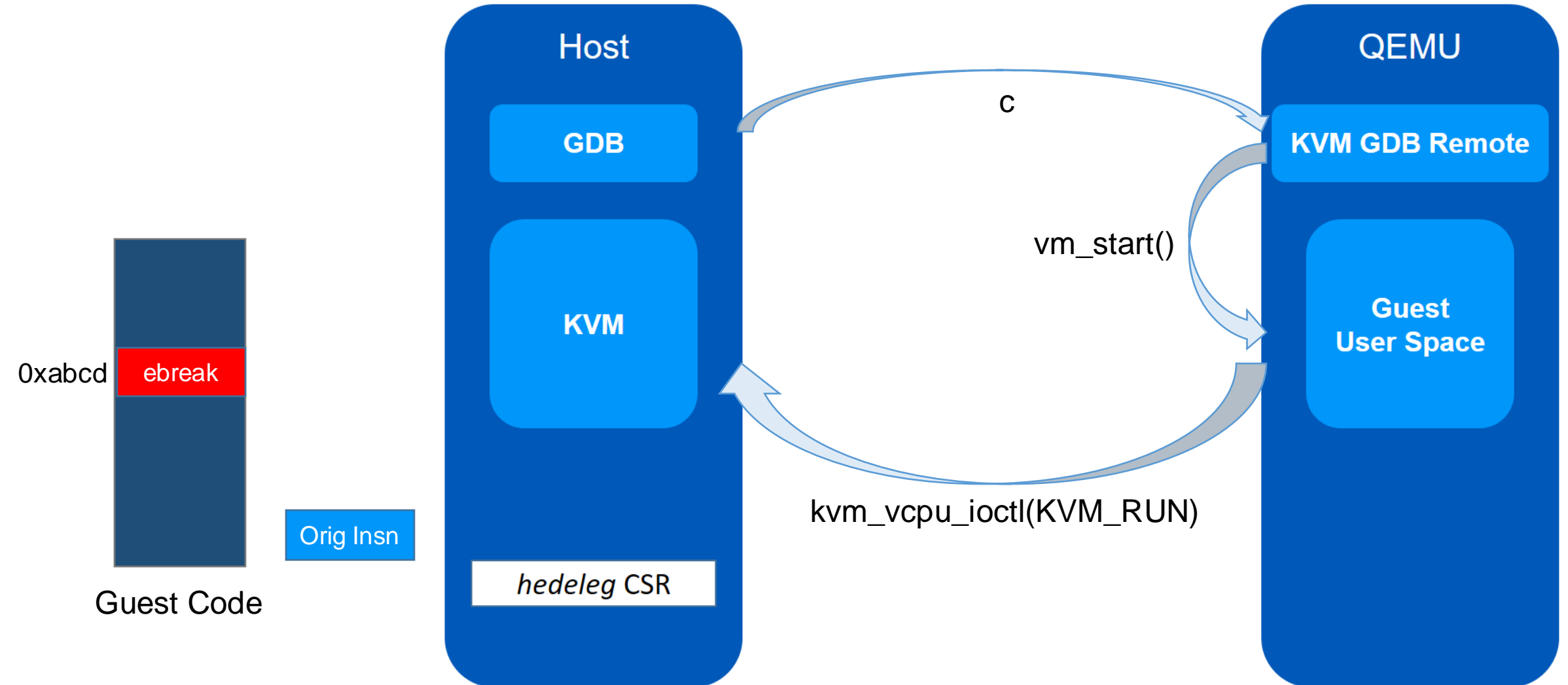
KVM:

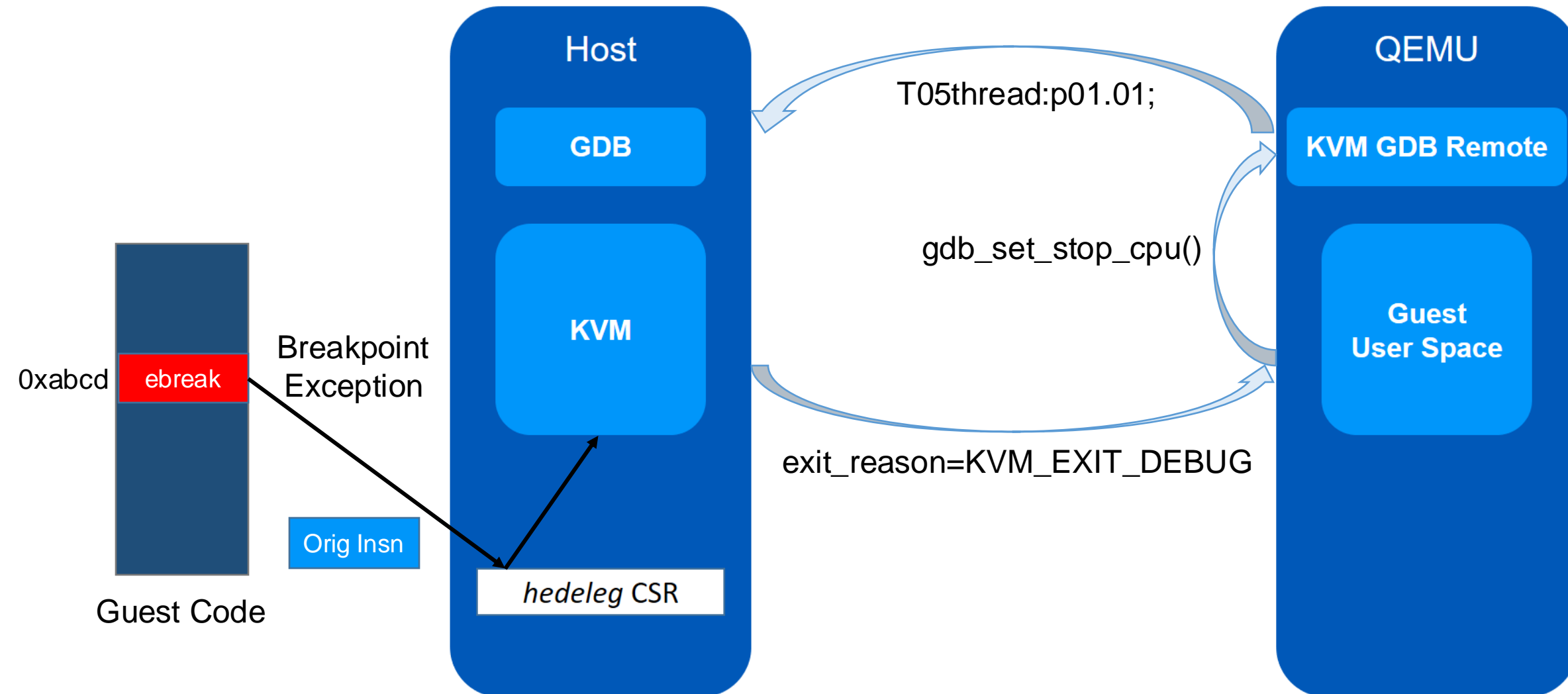
- KVM_CAP_SET_GUEST_DEBUG > 0
- 动态配置 *hedeleg* 寄存器
- 增加 KVM_EXIT_DEBUG 类型的exit

QEMU:

- 实现kvm_arch_insert_sw_breakpoint()和kvm_arch_remove_sw_breakpoint()
- 实现kvm_arch_update_guest_debug()
- 实现对KVM_EXIT_DEBUG退出原因的处理







QEMU启动命令在-accel kvm基础上加上-S -s

```
For help, type "help".
Type "apropos word" to search for commands related to "word".
(gdb) target remote:1234
Remote debugging using :1234
warning: No executable has been specified and target does not support
determining executable automatically. Try using the "file" command.
0x0000000080200000 in ?? ()
(gdb) add-symbol-file u-boot 0x80200000
add symbol table from file "u-boot" at
        .text_addr = 0x80200000
(y or n) y
Reading symbols from u-boot...
(gdb) b call_harts_early_init
Breakpoint 1 at 0x80200032: file arch/riscv/cpu/start.S, line 145.
(gdb) c
Continuing.

Breakpoint 1, call_harts_early_init () at arch/riscv/cpu/start.S:145
145         jal     harts_early_init
(gdb) l
140         mv      s0, a0
141
142
143         /* Configure proprietary settings and customized CSRs of harts */
144     call_harts_early_init:
145         jal     harts_early_init
146
147     #if !CONFIG_IS_ENABLED(XIP)
148         /*
149         * Pick hart to initialize global data and run U-Boot. The other harts
(gdb) █
```

回馈社区 by 奕斯伟计算：

- Linux Kernel - 6.10
- QEMU - 9.1.0

下一步计划与展望：

- 添加硬件断点支持
- Guest / KVM / User Space tool 同步调试

目录

- I. RISC-V KVM-简介与现状
- II. RISC-V KVM调试-场景及方法
- III. RISC-V KVM调试-实践与展望
- IV. RISC-V KVM调优-实践与展望**



KVM性能**初步**对比测试 (RISC-V 与 ARM)

- QEMU仿真平台
- 尽量相同配置
- benchmark: SPEC2006

关注Guest与Host结果的**比值**, **RISC-V约为0.6**, **ARM约为0.8**。且不同case间浮动较大。

影响KVM性能的主要因素: KVM Exit次数

因此, 提升RISC-V KVM性能的主要方法: **分析、减少KVM的Exit次数**

perf kvm for RISC-V:

```
A simple test go below:
# ./perf kvm stat record -a
[ perf record: Woken up 3 times to write data ]
[ perf record: Captured and wrote 8.502 MB perf.data.guest (99338 samples) ]

# ./perf kvm stat report --event=vmexit
```

Event name	Samples	Sample%	Time (ns)	Time%	Max Time (ns)
STORE_GUEST_PAGE_FAULT	26968	54.00%	2003031800	40.00%	3361400
LOAD_GUEST_PAGE_FAULT	17645	35.00%	1153338100	23.00%	2513400
VIRTUAL_INST_FAULT	1247	2.00%	340820800	6.00%	1190800
INST_GUEST_PAGE_FAULT	1128	2.00%	340645800	6.00%	2123200
SUPERVISOR_SYSCALL	1019	2.00%	245989900	4.00%	1851500
LOAD_ACCESS	986	1.00%	671556200	13.00%	4180200
INST_ACCESS	655	1.00%	170054800	3.00%	1808300
HYPERVISOR_SYSCALL	21	0.00%	4276400	0.00%	716500

回馈社区 by 奕斯伟计算:

- Linux Kernel - 6.11

下一步计划与展望:

- 完善perf kvm工具
- 分析RISC-V KVM的性能瓶颈

Thanks

北京奕斯伟计算技术股份有限公司

BEIJING ESWIN COMPUTING TECHNOLOGY CO., LTD.

www.eswincomputing.com



奕斯伟计算官方公众号

