VENTANA MICRO

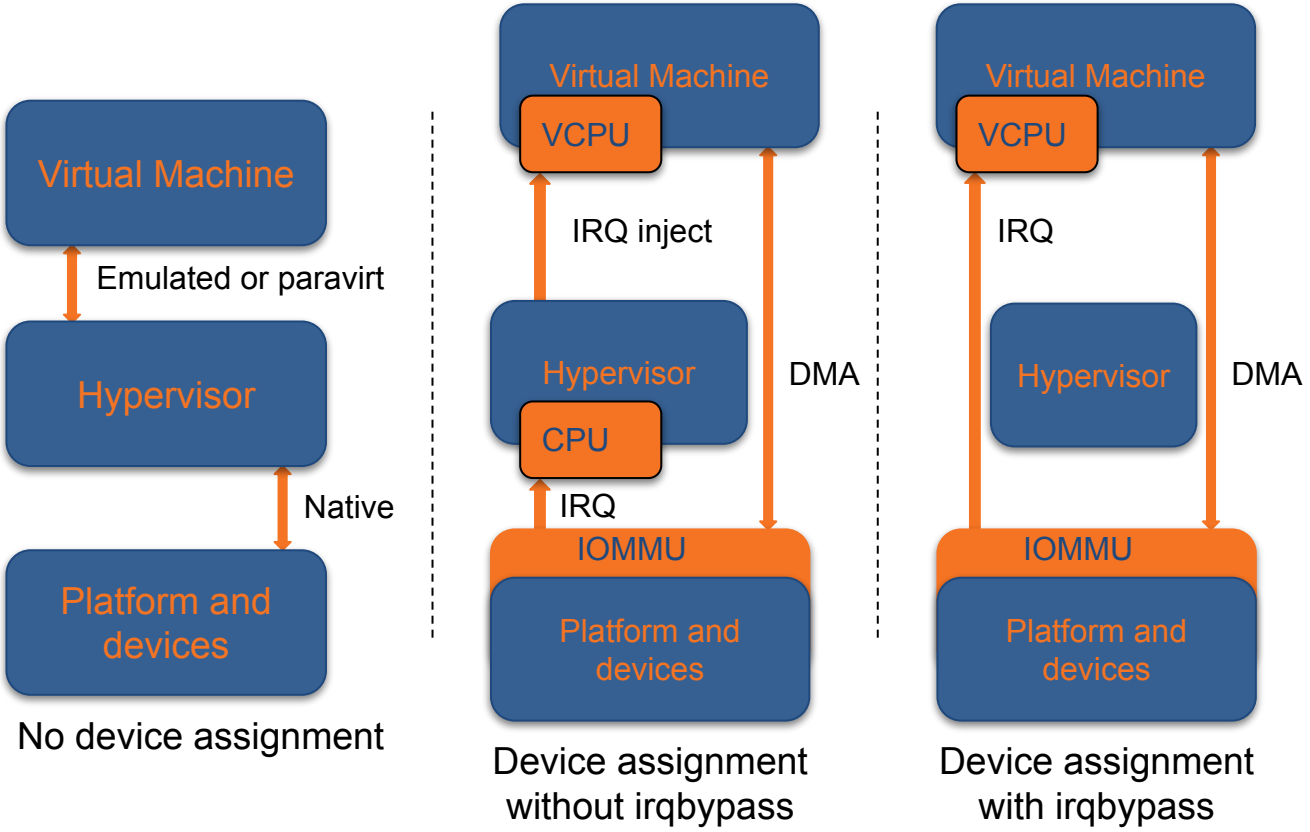
RISC-V irqbypass

Andrew Jones <ajones@ventanamicro.com>

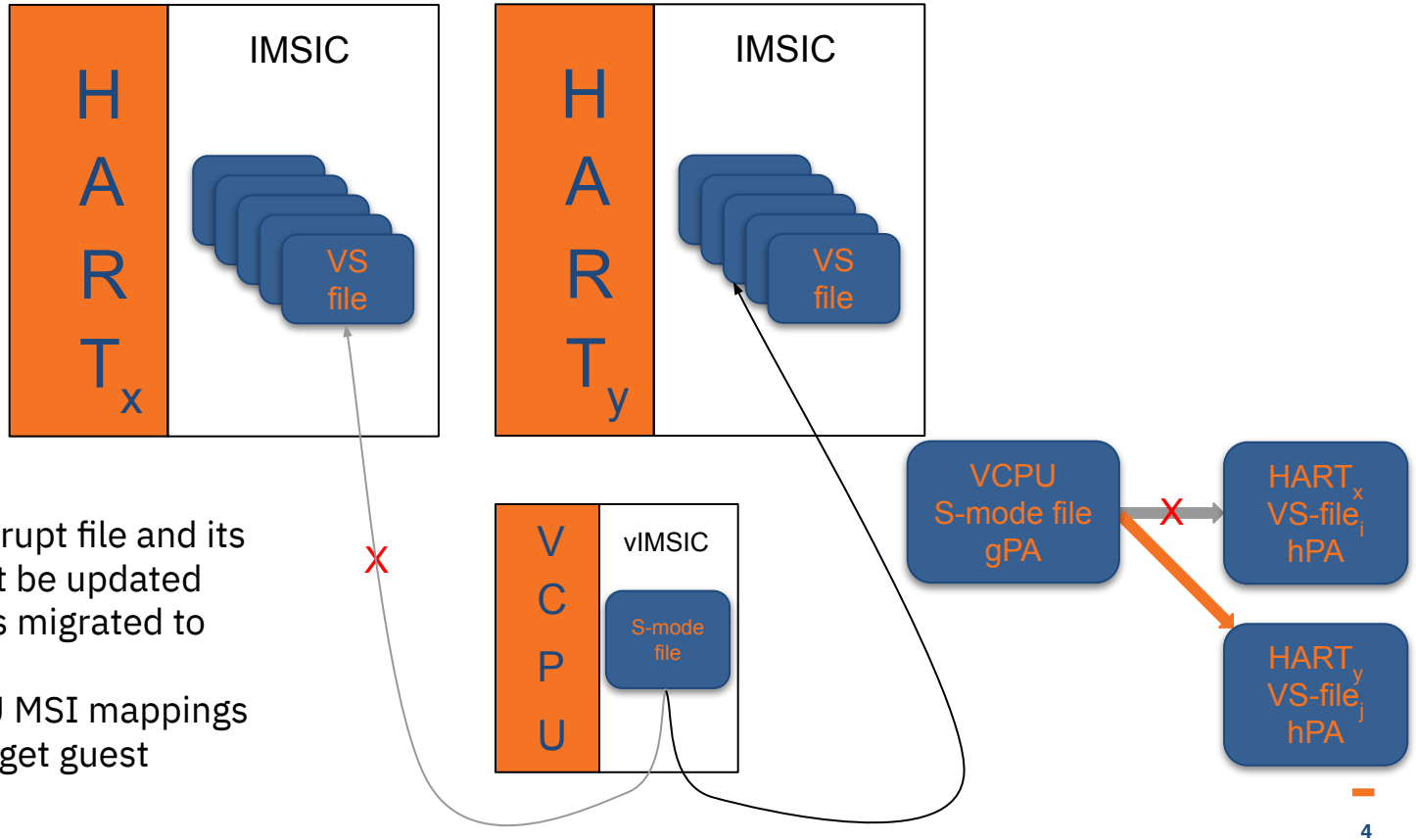
Outline

- Device assignment and irqbypass
- RISC-V architecture support
- PoC description and discussion
- References

Device assignment



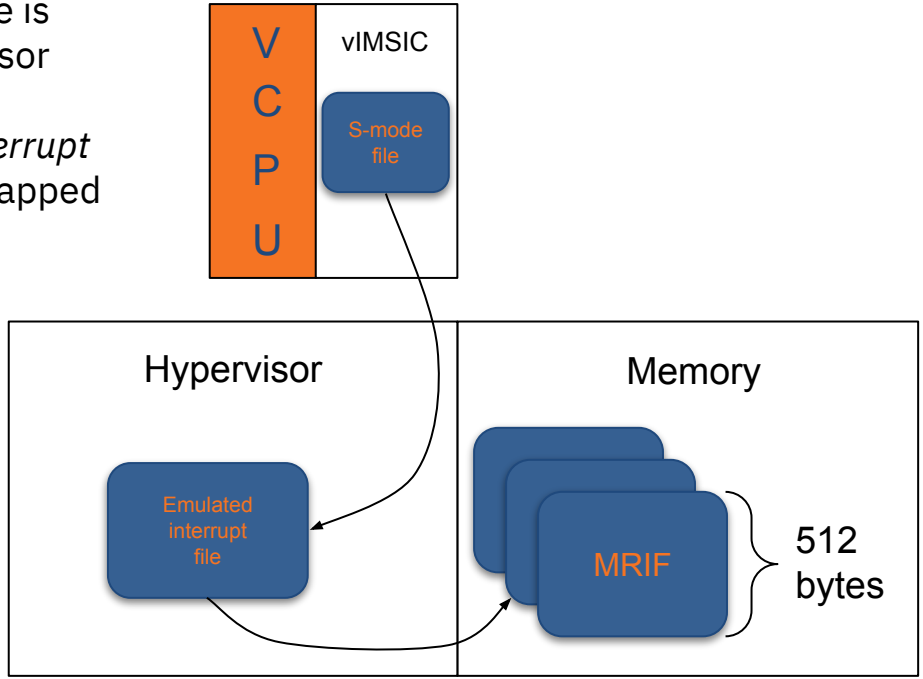
Guest MSI support



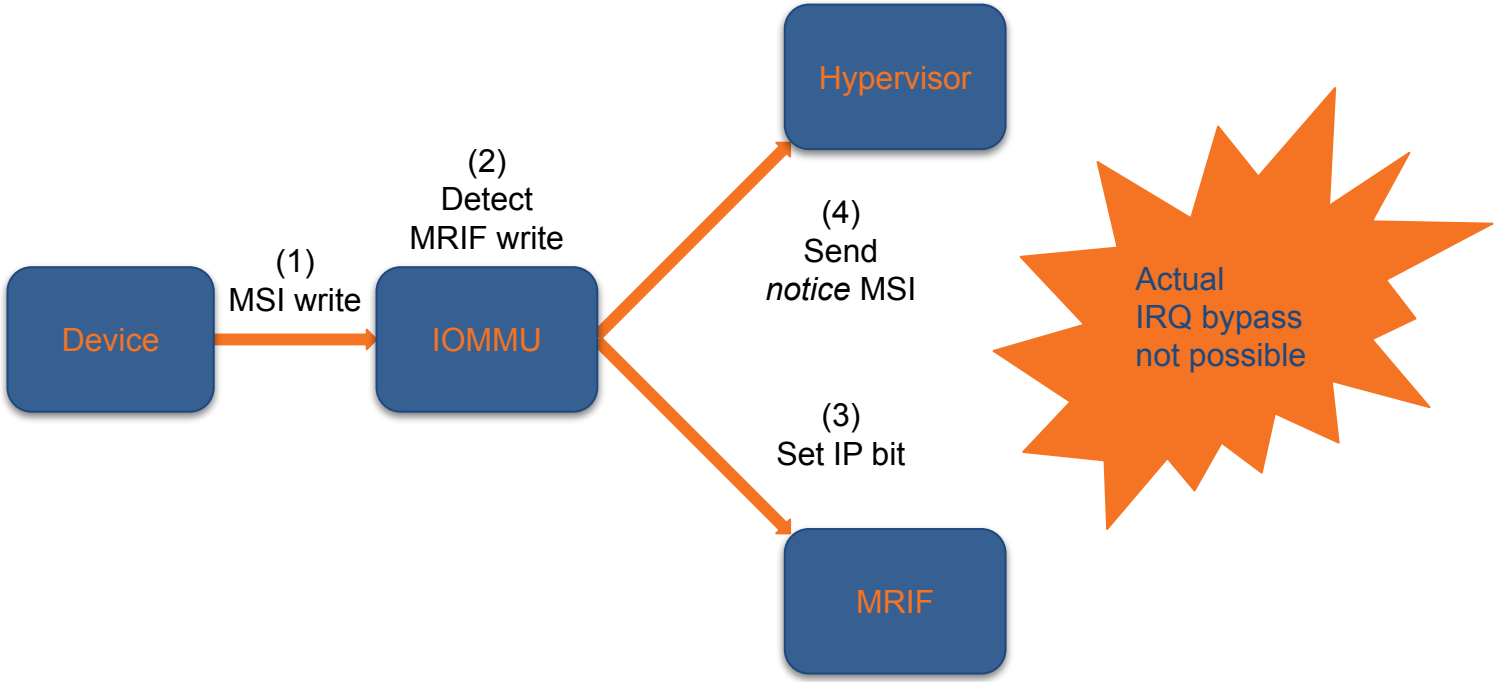
- The guest interrupt file and its mappings must be updated when a VCPU is migrated to another hart.
- RISC-V IOMMU MSI mappings can directly target guest interrupt files

MRIFs

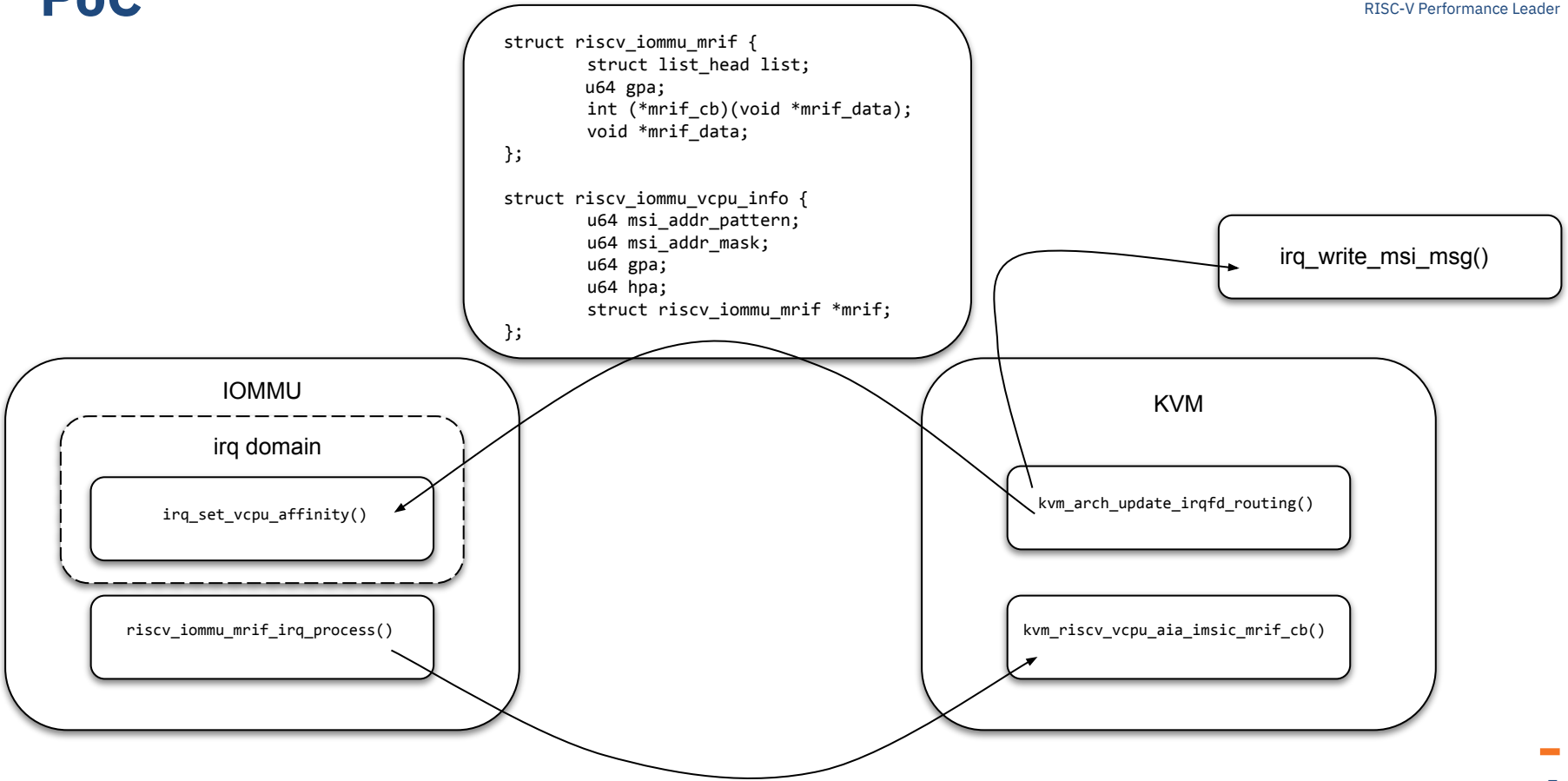
- When no VS-mode file is available, the hypervisor provides an MRIF (*memory-resident interrupt file*). Access to it is trapped and emulated.



MRIFs


Hypervisor
(2)
Detect
MRIF write
(1)
MSI write
Device
IOMMU
(4)
Send
notice MSI
Actual
IRQ bypass
not possible
(3)
Set IP bit
MRIF

PoC


struct riscv_iommu_mrif {
struct list_head list;
u64 gpa;
int (*mrif_cb)(void *mrif_data);
void *mrif_data;
};
struct riscv_iommu_vcpu_info {
u64 msi_addr_pattern;
u64 msi_addr_mask;
u64 gpa;
u64 hpa;
struct riscv_iommu_mrif *mrif;
};
irq_write_msi_msg()
IOMMU
irq domain
irq_set_vcpu_affinity()
riscv_iommu_mrif_irq_process()
KVM
kvm_arch_update_irqfd_routing()
kvm_riscv_vcpu_aia_imsic_mrif_cb()

References

- RISC-V IOMMU specification (ratified 1.0)
  https://github.com/riscv-non-isa/riscv-iommu
- RISC-V AIA specification (ratified 1.0)
  https://github.com/riscv/riscv-aia
- Linux RISC-V IOMMU support (under review)
  https://lore.kernel.org/all/cover.1689792825.git.tjeznach@rivosinc.com/
- Linux RISC-V AIA support (under review)
  https://lore.kernel.org/all/20231023172800.315343-1-apatel@ventanamicro.com/
- KVM AIA irqchip (including vIMSIC) support (merged for v6.5)
  https://lore.kernel.org/all/20230615073353.85435-1-apatel@ventanamicro.com/
- QEMU IOMMU model (under review)
  https://lore.kernel.org/all/cover.1689819031.git.tjeznach@rivosinc.com/
- QEMU AIA model (merged for v7.0.0)
  https://lore.kernel.org/all/20220220085526.808674-1-anup@brainfault.org/

Thank You