Missing PVH bits

Roger Pau Monné
roger.pau@citrix.com

June 26th, 2023

What is PVH?

- HVM domain without a (mandatory) QEMU device model.

Brief history

- 2013: PVHv1 domU support merged
- 2014: PVHv1 dom0 support merged
- 2016: PVHv2 domU introduction
- 2017: remove PVHv1
- 2018: PVHv2 dom0 introduction

PVHv2 design

- Do not use a different domain type internally in the hypervisor: PVHv2 is an HVM domain without any ioreq server.
- Do not blindly propagate PV interfaces into PVHv2.
- Consider what is required in order to make use of the hardware provided assistances: vAPIC or posted interrupts.
- Avoid introducing (Xen) PV specific interfaces when possible.
- New PVHv2 specific entry point ABI.

PVHv2 dom0 challenges

- dom0 requires access to physical devices, and for HVM guests that involves a device model (QEMU).
- On PV some interactions with devices involve using hypercalls.
- In order to allow a mostly transparent interaction with devices PCI config space emulation is required.
- Legacy PCI interrupts also require an IO-APIC.
- Identifying device MMIO regions is hard.

domU shortcomings

- UEFI (OVMF) firmware.
- PCI device passthrough.

dom0 shortcomings

- Wider testing.
- Some PCI capabilities won't work: Resizable BARs, SR-IOV (?).
- NMI handling: nmi=dom0 command line option not implemented.
- MCA support.
- Lack of PCI passthrough support.
- Linux: support for C and P state reporting.
- Physical CPU hotplug.
- Slowness of hypercalls on HVM.
- Security support.

Summary

- PVH domU is overall in a better position.
- PVH dom0 needs more work, and parties interested in using it.

Q&A

Thanks

Questions?