Linux Plumbers Conference 2025

Contribution ID: **129** Type: **not specified**

A Hyper-V Based pvIOMMU Built on Generic I/O Page Table

We present a Hyper-V based pvIOMMU implementation for Linux guest, built upon the community-driven Generic I/O Page Table framework. Our approach leverages stage-1 page tables in the guest(w/ nested translation) to drive DMA remapping(including vSVA). This also eliminates the need for complex device-specific emulation and map/unmap overhead, meanwhile staying scalable across architectures.

This session will cover:

1. The current development status and our planned upstream contributions prior to LPC.
2. Integration strategy with the Generic I/O Page Table infrastructure.
3. Demonstration that the Generic I/O Page Table can support not only hardware IOMMUs but also paravirtualized transports, with internal proof-of-concepts validating both virtio-iommu and Hyper-V pvIOMMU atop stage-1 DMA remapping.
4. Expected performance improvements compared to traditional emulation.
5. Open challenges, including modeling IOTLB invalidation in a paravirtualized context and exploring the use of hardware-virtualized command queues.
6. A brief discussion on differing pvIOMMU requirements between the Linux Dom0 on Azure hosts and Linux guest VMs, and our proposed code structure to support coexistence across both environments.

We seek feedback on integration boundaries and the minimal common infrastructure required to upstream the Hyper-V pvIOMMU.

Primary author: ZHANG, Yu

Co-authors: Mr HARIHARAN, Easwar (Microsoft); Mr PAN, Jacob (Microsoft)

Presenter: ZHANG, Yu

Session Classification: VFIO/IOMMU/PCI MC

Track Classification: VFIO/IOMMU/PCI MC