

IOASID API extensions for Intel Scalable IOV usages

Wednesday, August 26, 2020 7:40 AM (20 minutes)

As it currently stands in the mainline kernel, IOASID is a generic kernel service that provides PCIe PASID or ARM SMMU sub-stream ID allocations. On VT-d and Intel's Scalable IO Virtualization(SIOV) platforms, IOASID core serves a particularly important role as its usage spans the following dimensions:

- bare metal and guest SVM
- A slew of in-kernel users consists of VFIO, IOMMU, mm, VDCM*, KVM

To fulfill the requirements of SIOV, we are proposing adding the following functionalities:

1. Extend IOASID set to support permission checking, token sharing, quota management
2. Add reference counting for life cycle management
3. Add per IOASID set the private ID for non-identity guest-host PASID mappings
4. Add notifiers to keep IOASID users synchronized on state change events, e.g. FREE, BIND/UNBIND, etc.

At LPC 2020. We are trying to get a consensus on the principles of these API extensions. If time permits, we would like to walk through the life-cycle of an IOASID on Intel's SIOV enabled platforms. Kernel documentation will be included in the patchset submission.

I agree to abide by the anti-harassment policy

I agree

Primary authors: LIU, Yi; PAN, Jacob

Presenters: LIU, Yi; PAN, Jacob

Session Classification: VFIO/IOMMU/PCI MC

Track Classification: VFIO/IOMMU/PCI MC