

Linux Plumbers Conference 2024



Contribution ID: 226

Type: **not specified**

Live update: persisting IOMMU domains across kexec

Wednesday 18 September 2024 12:10 (20 minutes)

Live update is a mechanism to support updating a hypervisor in a way that has limited impact to running virtual machines. This is done by pausing/serialising running VMs, kexec-ing into a new kernel, starting new VMM processes and then deserialising/resuming the VMs so that they continue running from where they were. When the VMs have DMA devices assigned to them, the IOMMU state and page tables needs to be persisted so that DMA transactions can continue across kexec.

In this session we want to discuss a revised approach to solving this problem: introducing persistent iommufd IOAS and HW pagetable. The idea is to use the Kexec Hand Over (KHO) framework as a mechanism to pass the persisted data across kexec and to restore it after kexec: <https://lore.kernel.org/kexec/20231213000452.88295-1-graf@amazon.com/>

We'd like to have a discussion about what the correct abstraction is for marking IOMMU(FD) domains as persistent, setting up persistent mapping and discovering and restoring the domains after kexec. RFC patches will be posted before hand to make the problem clearer.

This session will be iterating on the live update concept which was discussed at last LPC, and it will be a revision to the idea of pkernfs which was floated as a potential solution: <https://lore.kernel.org/all/20240205174238.GC31743@ziepe.ca/>

Primary authors: GRAF, Alexander; GOWANS, James (Amazon EC2)

Presenter: GOWANS, James (Amazon EC2)

Session Classification: VFIO/IOMMU/PCI MC

Track Classification: VFIO/IOMMU/PCI MC