

PCI hotplug: movable BARs and bus numbers

Wednesday 26 August 2020 08:40 (25 minutes)

Hot-adding a PCI device requires gaps in the address space for new BARs, and extra bus numbers if this is a bridge. Usually these resources are reserved not by the kernel, but by BIOS, bootloader, firmware.

If a bridge have windows not big enough (or fragmented too much) for newly requested BARs, it is still may be possible to allocate a memory region for new BARs, if at least some working BARs can be moved after pausing the drivers supporting this feature.

This approach is also useful if a BIOS don't allocate all requested BARs, leaving some (for example, SR_IOV) unassigned, without gaps for bridge windows to extend. And it can help in allocating large (gigabytes in size) BARs.

Second (and optional) part is re-enumerating the buses allows hot-adding large switches in the middle of an existing PCIe tree, but it's problematic point is renaming entries in `/sys/bus/pci` and `/proc/bus/pci`.

I agree to abide by the anti-harassment policy

I agree

Primary author: MIROSHNICHENKO, Sergei (Yadro)

Presenter: MIROSHNICHENKO, Sergei (Yadro)

Session Classification: VFIO/IOMMU/PCI MC

Track Classification: VFIO/IOMMU/PCI MC