



Contribution ID: 203

Type: **not specified**

Non-discoverable devices in PCI devices

Wednesday, 15 November 2023 10:30 (30 minutes)

Modern PCI devices can expose a whole slew of hardware behind a single PCI “device”. While the PCI device itself is discoverable, everything behind it (via BARs) is not. These devices aren’t fixed in what downstream devices are exposed nor their configuration. There’s already a solution for discovering devices and their configuration which is Devicetree. There’s also already a mechanism to dynamically add to a DT with DT overlays.

Specific use cases:

- AMD/Xilinx Alveo Accelerator cards - FPGA based PCIe card. Multiple downstream devices (hardware peripherals) exposed through PCI BARs. The downstream devices may be unrelated and already have a driver (platform bus) in the kernel.
- Microchip LAN9662 Ethernet controller - An SoC already supported upstream using Devicetree. This device also has a PCIe endpoint which can expose the entire SoC to Linux. Using DT to describe the SoC allows reusing all the drivers as-is.
- roadtest - A testing framework which exposes platform devices behind a virtual PCI device in UML. See <https://lore.kernel.org/all/20230120-simple-mfd-pci-v1-1-c46b3d6601ef@axis.com/>

What’s needed in the kernel:

- Generating PCI device DT nodes - PCI devices can already be described in DT, but usually are not. In order to have a base to apply DT overlays to, the PCI device needs a DT node. That can be solved by generating the DT node (and parent nodes) for the device. With this, the PCI driver for the device can load and apply DT overlays to its node. See <https://lore.kernel.org/all/1692120000-46900-1-git-send-email-lizhi.hou@amd.com/> (Should be upstream in 6.6)
- Enable Devicetree on Non-DT systems (x86_64) - A solution for ACPI based systems is desired as well. There’s SSDT overlays for ACPI which might work?, but many of the devices to support don’t use ACPI. With the generated PCI nodes, we just need the PCI host bridge and hooking up a few host resources (interrupts, MSI) to reuse the same DT overlay mechanism. A skeleton base DT is needed as well. That already exists as the DT unittest creates one for non-DT systems. There may be other issues lurking where we make DT vs. ACPI decisions in the kernel.

Questions:

- What about swnode and/or auxiliary bus?

Primary authors: HOU, Lizhi (AMD); HERRING, Rob (Arm)

Presenters: HOU, Lizhi (AMD); HERRING, Rob (Arm)

Session Classification: VFIO/IOMMU/PCI MC

Track Classification: LPC Microconference: VFIO/IOMMU/PCI MC