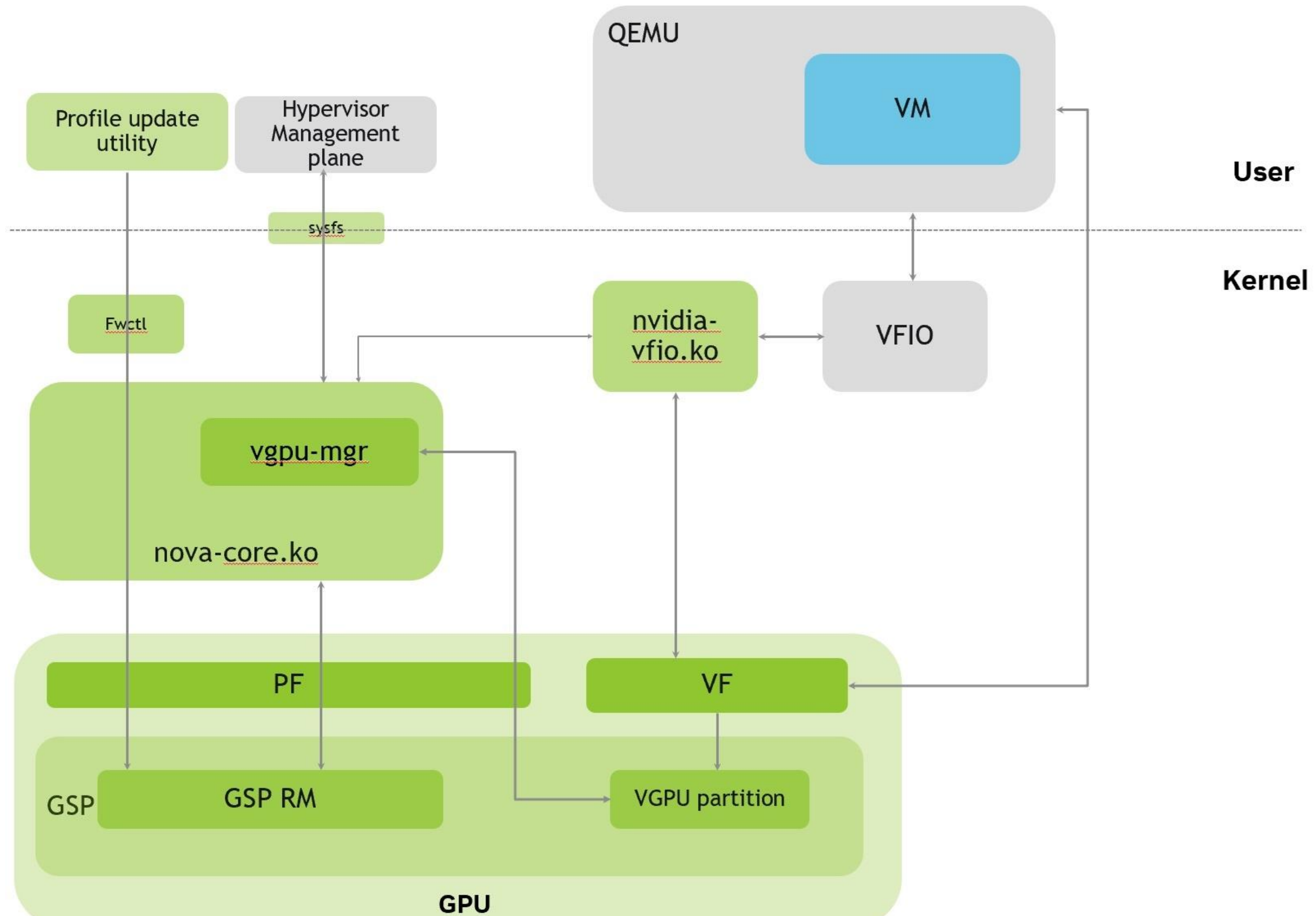




vGPU Upstream Plan

Zhi Wang and NVIDIA vGPU team | Linux Plumber Conference 2025 – Nova GPU
workshop / Dec, 10th, 2025

Architecture



<https://pretalx.com/kvm-forum-2025/speaker/DXULWY/>

vGPU Upstream Plan

Stage 1: Boot GSP RM with vGPU enabled

RFC: <https://lore.kernel.org/all/20251206124208.305963-1-zhiw@nvidia.com/>

Deps on rust-for-linux

- PCI configuration space access abstraction
- PCI SR-IOV helpers

Stage2: Bootstrap vGPU

Deps on rust-for-Linux

- Fwctl abstraction

Stage 3: Boot up VMs

Deps on rust-for-linux

VF enabling: <https://lore.kernel.org/all/20251119-rust-pci-sriov-v1-0-883a94599a97@redhat.com/>



Thanks!

vGPU Upstream Plan

Nova-core

- GSP command/message queue refinement
- GSP command support: RM_ALLOC/RM_FREE/RM_CONTROL
- GPU interrupt support (VFN)
- GPU memory management
- BAR1 mapping support
- Channel allocation and reservation
- Fwctl driver

Nova-core

- Engine enumeration
- Channel Initialization (RAMFC/GPU context....)
- CE workload submission

vGPU Upstream Plan

Stage 3: Enable VFs and boot up VMs

Deps

Rust-for-linux

VF enabling

<https://lore.kernel.org/all/20251119-rust-pci-sriov-v1-0-883a94599a97@redhat.com/>

Nova-core

- Engine enumeration
- Channel Initialization (RAMFC/GPU context....)
- CE workload submission