

Linux Plumbers Conference

Dublin, Ireland September 12-14, 2022

A decorative graphic of a green pipe network with various fittings, valves, and elbows, framing the central text.

IOMMUFD Discussion

Eric Auger
Yi Liu
Baolu Lu
Jason Gunthorpe
Kevin Tian



Linux

Plumbers Conference | Dublin, Ireland **Sept. 12-14, 2022**

A decorative graphic of green pipes with valves and elbows, running vertically on the left side of the slide and curving at the top and bottom.

Session Goals

Overview of iommufd and its goals
Review progress so far
Open discussions from mailing list
Discussion merging threshold



Linux

Plumbers Conference | Dublin, Ireland Sept. 12-14, 2022

Overview

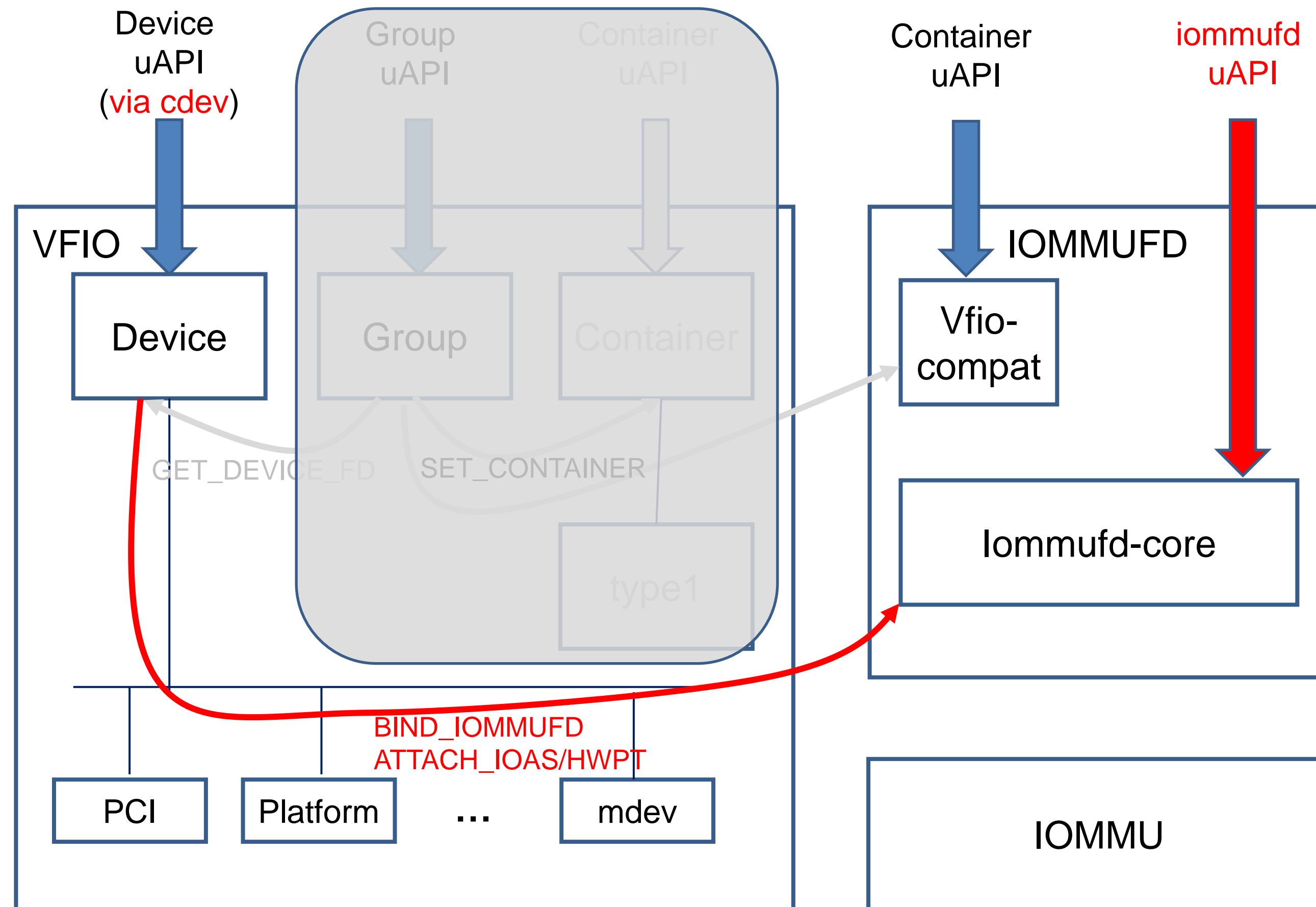
- Expose IOMMU HW to userspace control
- Provide advanced IOMMU features
- Provide high performance “kernel bypass” to support VMs
- Support all kernel subsystems equally: VDPA, VFIO, uacce, etc



Linux

Plumbers Conference | Dublin, Ireland Sept. 12-14, 2022

iommufd-native



- iommufd replaces VFIO group and container FDs
- VFIO becomes a simple device driver
 - Device is opened via cdev
 - No use of group or container
- New VFIO APIs to manage iommufd connection
- New IOMMU features flow through iommufd



Basic uAPI

IOMMUFD_CMD_DESTROY
Destroy an object

IOMMUFD_CMD_IOAS_ALLOC
Allocate an IO Address Space (IOAS)

IOMMUFD_CMD_IOAS_ALLOW_IOVAs
Ensure a range of IOVA is available

IOMMUFD_CMD_IOAS_COPY
Copy the mapping from a range of IOVA to another

IOMMUFD_CMD_IOAS_IOVA_RANGES
Inquire valid ranges of IOVA

IOMMUFD_CMD_IOAS_MAP
Install user VA to IOVA

IOMMUFD_CMD_IOAS_UNMAP
Remove IOVA

IOMMUFD_CMD_VFIO_IOAS
Work with the VFIO compat IOAS



Linux

Plumbers Conference | Dublin, Ireland Sept. 12-14, 2022

iommufd features

Features	Ver
iommufd generic interface	RFC v2
VFIO integration with iommufd (compat. mode)	github
IOMMU dirty tracking	github
iommufd nested translation	github
VFIO integration with iommufd (via cdev)	github Qemu
I/O page fault	n/a
PASID support	n/a
Share KVM page table with IOMMU	n/a

Clean and in a reviewable state

In preparation for formal patch series

Prototype ready but not fully scrutinized

Prototype in progress



Linux

Plumbers Conference | Dublin, Ireland Sept. 12-14, 2022

Preparatory Series

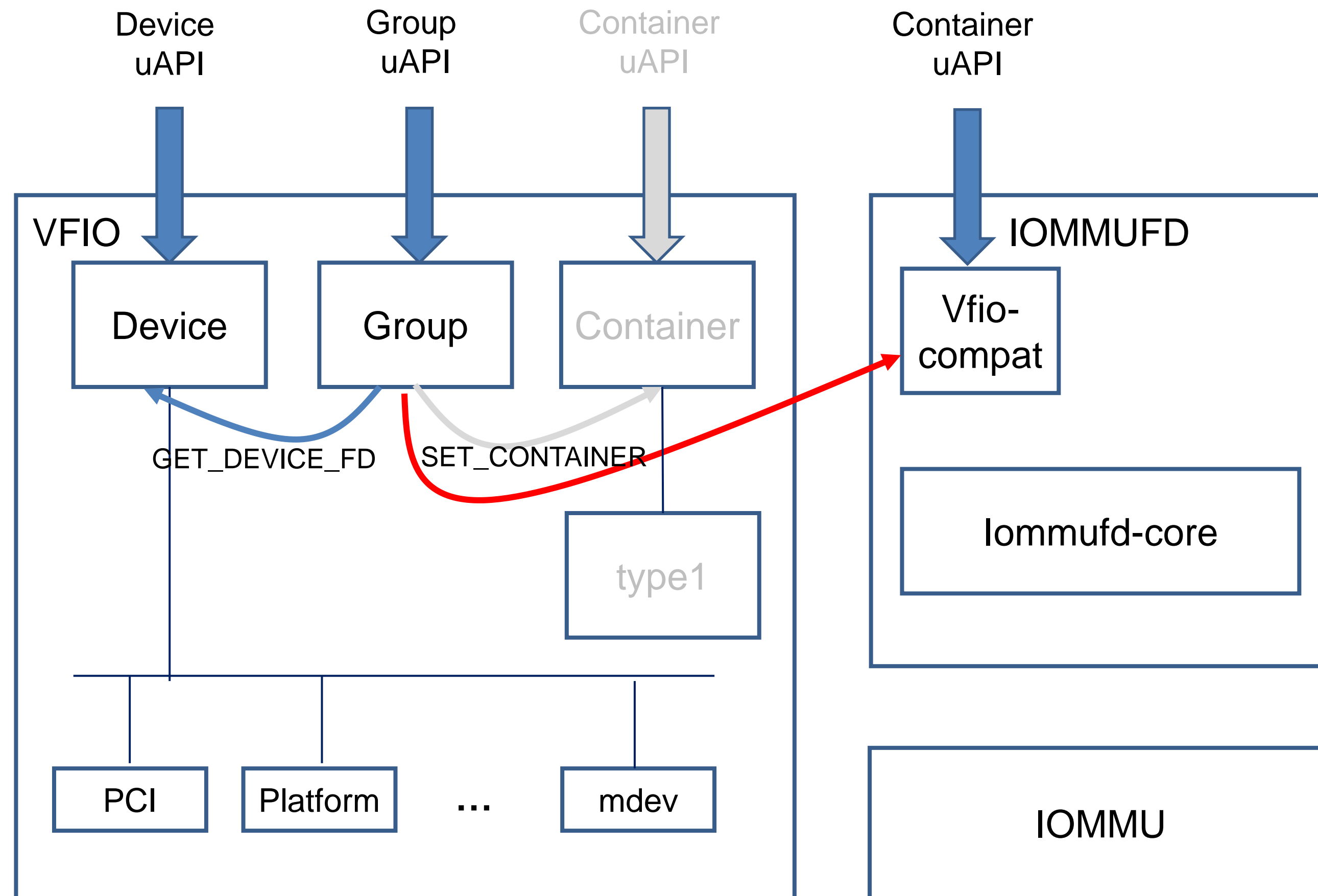
Series	Ver	Status
Simplify domain/device compat. checking	v6	Respin with different error codes
Redo SVA with domain-centric model	v13	Close to be merged
Introduce a device struct for vfiio_device	v2	Collecting reviewed-by
VFIO API for dirty tracking	v6	Merged
Complete restructuring the vfiio mdev model	v7	Still see breakage on ccw
VFIO dma_buf exporter for VFIO	v2	Requires major infrastructure work
Isolate vfiio container in preparation for iommufd	v1	Simple change with only a naming open.
Provide iommu_domain ops for power	v1	Not elegant but no owner to further improve if it's not picked by maintainer



Linux

Plumbers Conference | Dublin, Ireland Sept. 12-14, 2022

vfio-compat



Differences

- TYPE1v2 only
- No `dma_entry_limit` (rely on memory cgroup)
- Pinned page accounting (user vs. mm)
- VFIO P2P is a followup
- TBD
VFIO_DMA_MAP_FLAG_VADDR
- errno differences
- SPAPR is left out (until it is moved to use `iommu_domain`)



A decorative graphic of a green pipe network with various fittings, valves, and elbows, framing the central text.

Threshold for merging

As-is VFIO integration and functional complete
dpdk, qemu functionally work, with known limitations

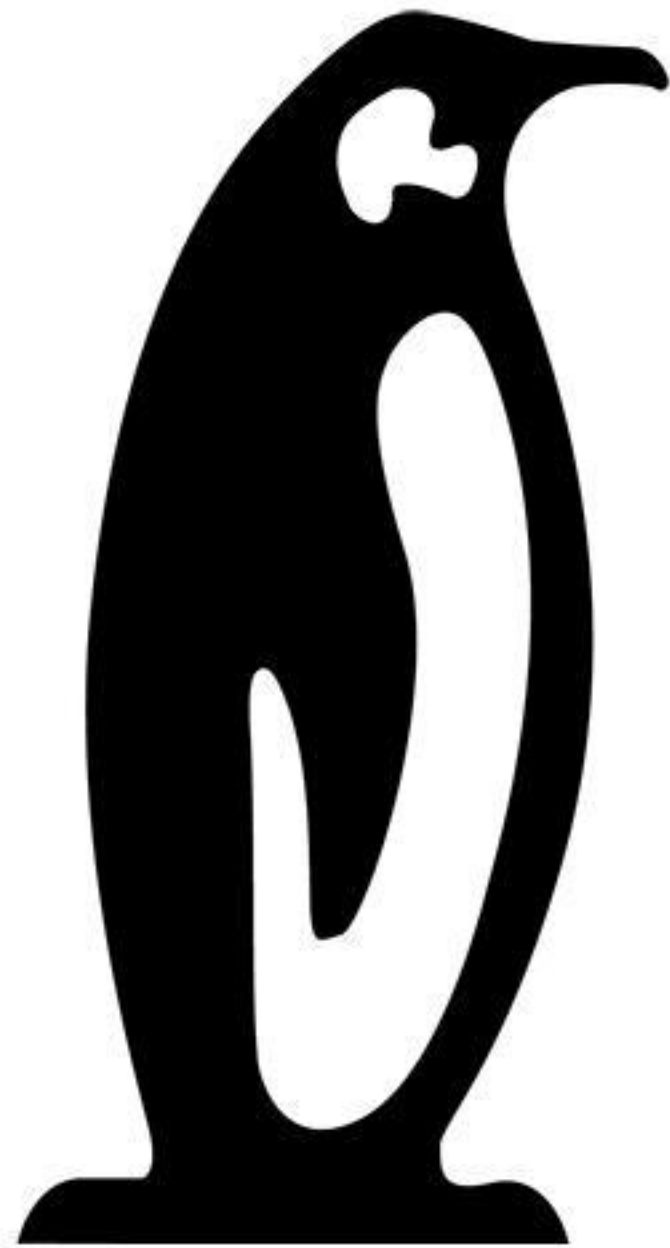
6.2 is a plausible target

Want to get parallel work into the community



Linux

Plumbers Conference | Dublin, Ireland Sept. 12-14, 2022



Linux Plumbers Conference

Dublin, Ireland September 12-14, 2022