What's coming in VIRTIO 1.2

New virtual I/O devices and features

Stefan Hajnoczi

stefanha@redhat.com

Red Hat

The VIRTIO standard

Open standard for virtual I/O devices

- ▸ Drivers available for Linux, BSD, macOS, Windows, and more
- ▸ Devices supported by multiple hypervisors and emulators
- ▸ Hardware implementations possible for many devices

New specification release every few years

- ▸ New features are often shipped in code once accepted into spec, you may already be using features from the upcoming spec release!

Red Hat

New features in 1.2

Core

- ▶ Per-virtqueue reset
- ▶ Shared memory resources
- ▶ Object resources for inter-device sharing
- ▶ Virtio-pci vendor-specific capabilities
- ▶ Virtio-pci queue_notify_data optimization

virtio-blk

- ▶ **Secure erase**
- ▶ Lifetime metrics for eMMC-like media

virtio-net

- ▶ **UDP Segmentation Offload**
- ▶ **Receive Side Scaling**
- ▶ **Per-packet hash reporting**
- ▶ Guest hdrlen optimization
- ▶ Link speed and duplex reporting

virtio-balloon

- ▶ Free page hinting & reporting
- ▶ Free page poisoning

Red Hat

New devices in 1.2

net
blk
serial
rng
balloon
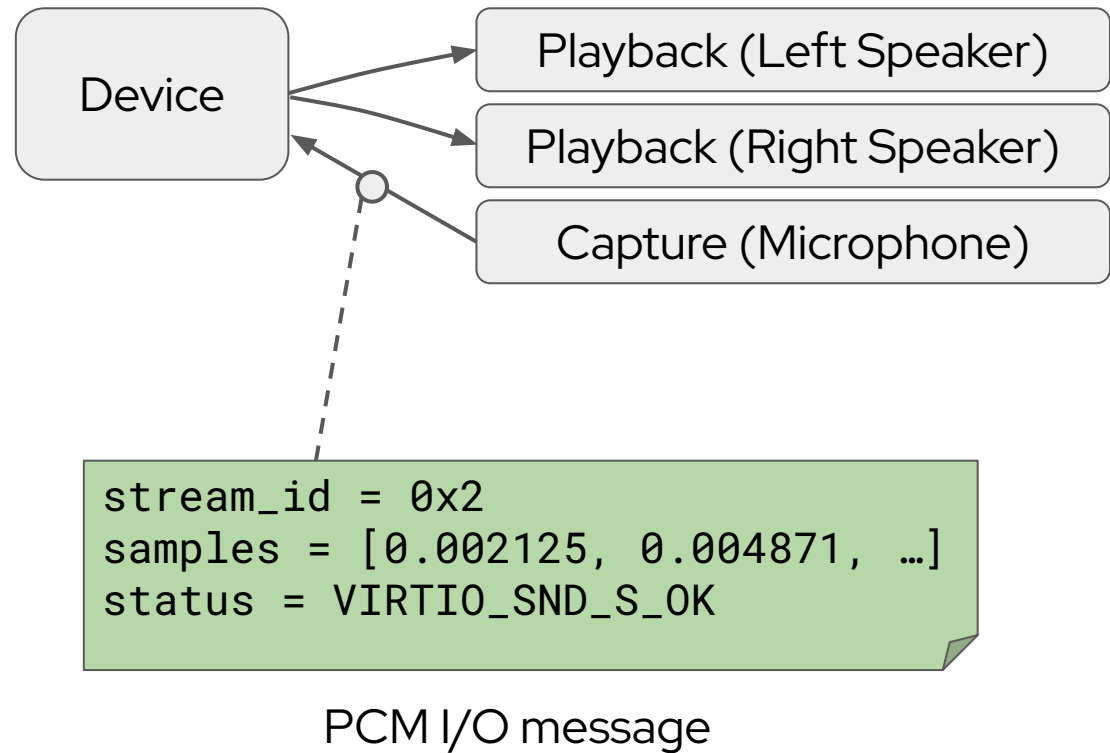scsi
gpu
input
crypto
vsock

+

iommu
sound
fs
pmem
mem
rpmb
scmi
i2c
gpio

Huge update with 9 new devices in the standard!

Even more device IDs reserved for future devices or non-standard applications.

We'll just look at the standard devices.

Red Hat

sound


```
stream_id = 0x2
samples = [0.002125, 0.004871, …]
status = VIRTIO_SND_S_OK
```
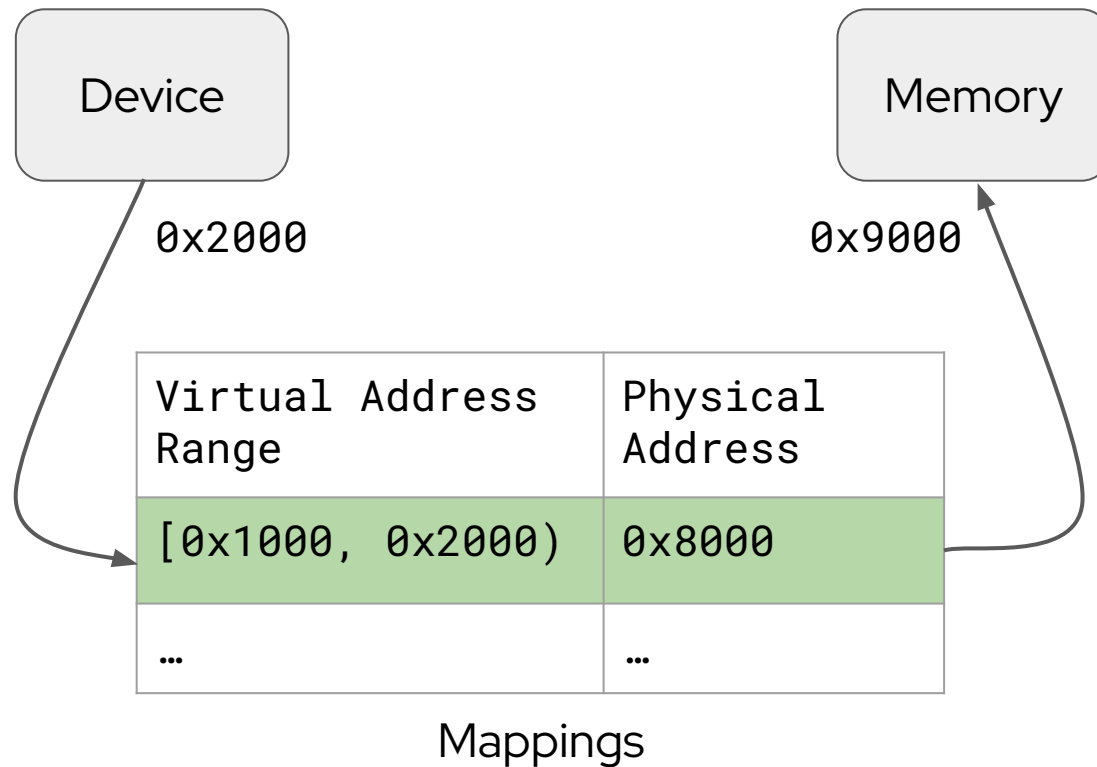
PCM I/O message

Sound card for multichannel PCM audio playback and capture with up to 384 kHz sample rate and 64-bit sample depth.

Reuses parts of High Definition Audio (HDA) spec.

Use cases:

▸ **Listening to music and movie/game audio.**

▸ **Voice calls.**

▸ **Pro audio/recording.**

Red Hat

Linux 5.3   QEMU 5.0   Libvirt TBD

iommu

Device

Memory

0x2000

0x9000

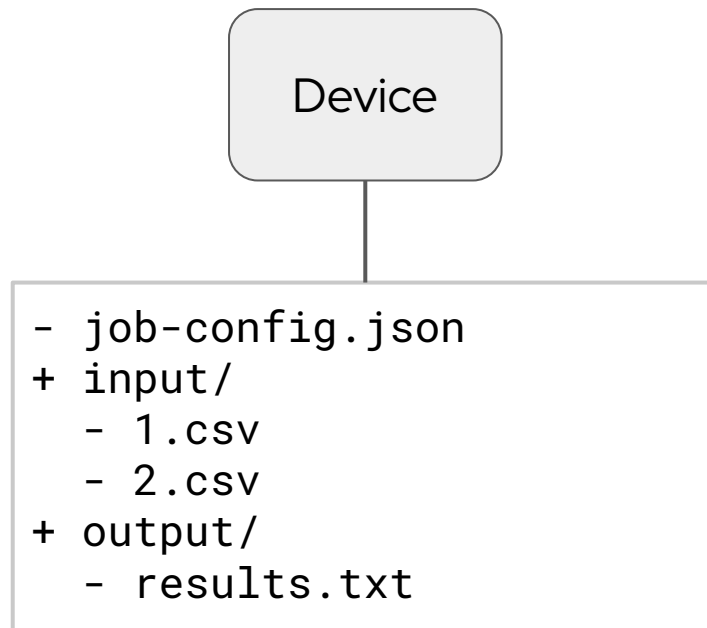| Virtual Address Range	Physical Address
[0x1000, 0x2000)	0x8000
...	...

Mappings

Simpler cross-arch VIRTIO alternative to emulated IOMMUs (Intel, AMD, ARM, etc).

Translates device memory accesses and enforces memory protection.

Use cases:

▶ **Linux VFIO apps (e.g. DPDK) inside guest.**

▶ **Isolating passthrough device DMA.**

▶ **Nested virtualization.**

Linux Plumbers Conference 2017 presentation (Jean-Philippe Brucker): https://blog.linuxplumbersconf.org/2017/ocw/sessions/4811.html
KVM Forum 2017 presentation (Eric Auger): https://www.linux-kvm.org/images/8/8e/Viommu_arm.pdf

fs

Device

```
- job-config.json
+ input/
  - 1.csv
  - 2.csv
+ output/
  - results.txt
```
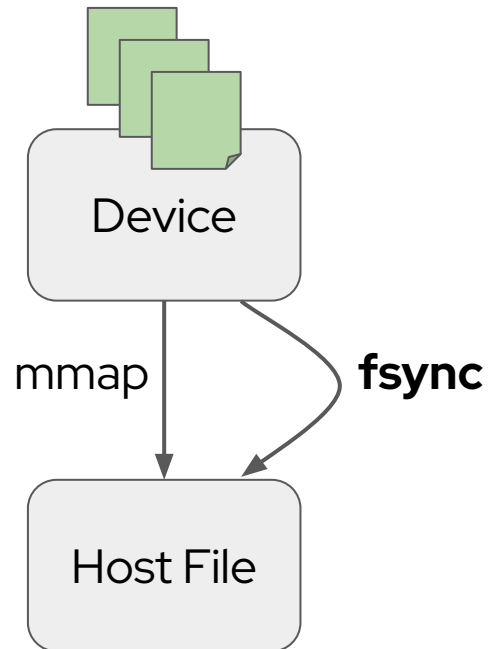
Shared host <-> guest file system with optional direct access to host page cache.

Based on Linux FUSE interface.

Use cases:

▸ **Secure container VMs like Kata Containers.**
▸ **Sharing source and build directories for development/testing in a VM.**
▸ **Providing files to VMs at install time.**

Website: https://virtio-fs.gitlab.io/

pmem

Device

mmap | **fsync**

Host File
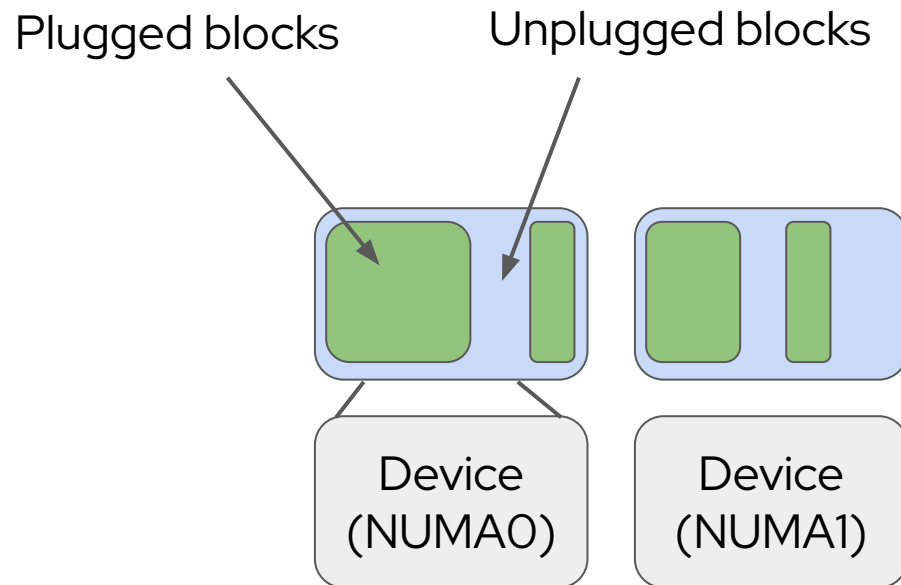
Persistent memory (NVDIMM) device with flush mechanism.

Real NVDIMM devices may lack hypervisor-trappable mechanisms to flush written data to persistent storage.

Use cases:

- **Sharing host page cache to avoid copying data into all VMs.**
- **Persistent memory workloads where backing storage is not a real NVDIMM.**

https://qemu.readthedocs.io/en/latest/system/devices/virtio-pmem.html

mem

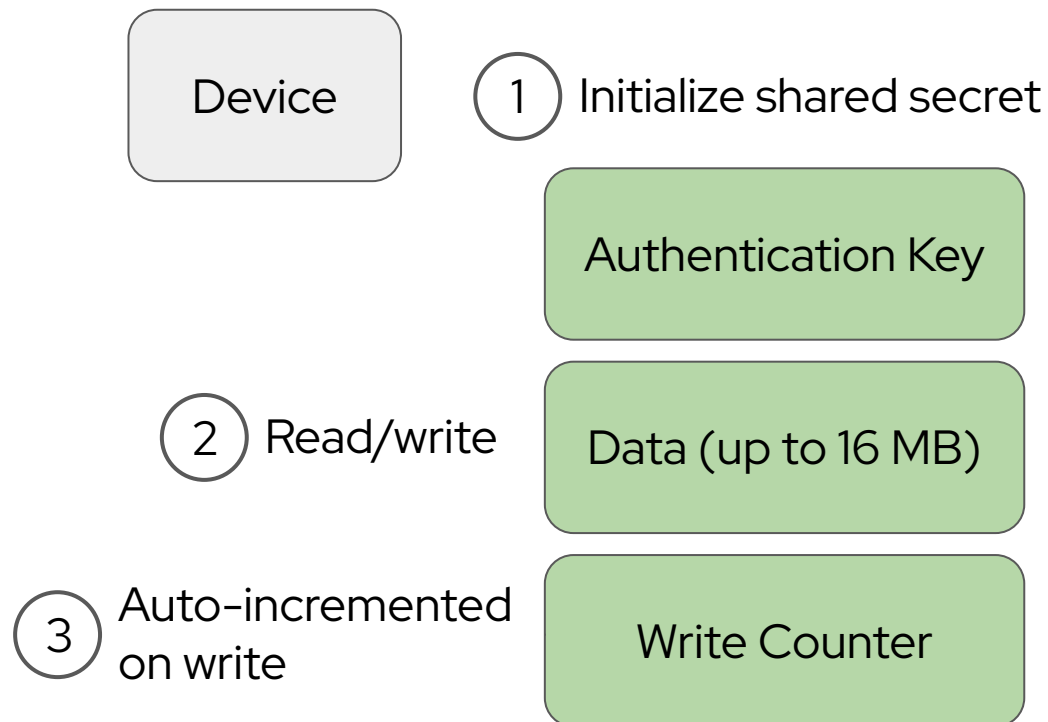Plugged blocks   Unplugged blocks

Device (NUMA0)   Device (NUMA1)

Cross-arch and NUMA-aware memory hotplug.

Supports fine-grained hotplug, not limited to arch-specific DIMM granularity.

Use cases:

▸ Adding/removing memory from guest at runtime.

KVM Forum 2018 (David Hildenbrand): :
https://events19.linuxfoundation.org/wp-content/uploads/2017/12/virtio-mem-Paravirtualized-Memory-David-Hildenbrand-Red-Hat-1.pdf
VEE 2021 paper (David Hildenbrand, Martin Schulz): https://dl.acm.org/doi/10.1145/3453933.3454010

rpmb

Device

① Initialize shared secret

Authentication Key

② Read/write   Data (up to 16 MB)
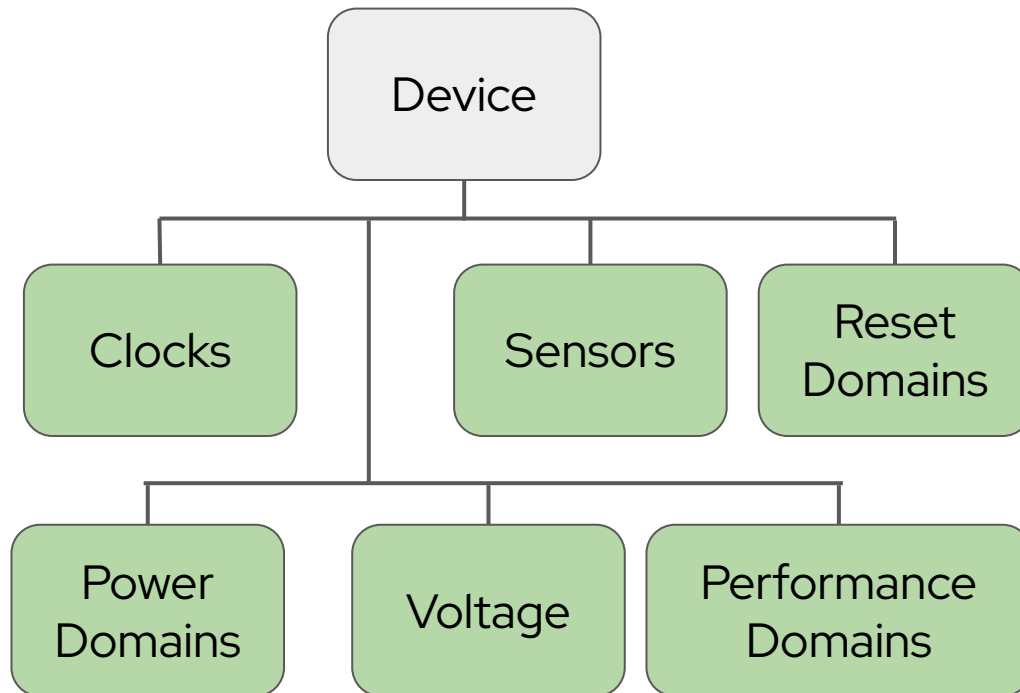
③ Auto-incremented on write   Write Counter

Replay Protected Memory Block device for tamper-resistant storage.

Based on feature from e.MMC and UFS flash devices.

Use cases:

▸ Protect system data (firmware or configuration) from modification.
▸ Storage for applications with sensitive data.

Red Hat

scmi

```
                    Device
         ┌────────────┼────────────┐
       Clocks      Sensors       Reset
                                 Domains
    ┌──────┼──────────┐
   Power   Voltage  Performance
  Domains           Domains
```
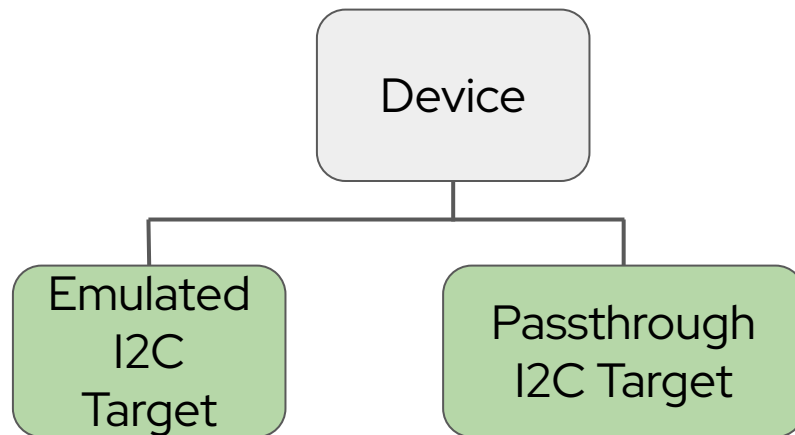
Arm System Control and Management Interface device for power and performance management.

Use cases:

▸ Reusing bare metal Arm power and performance management code inside VMs.

i2c

I2C adapter that gives access to emulated or passthrough I2C targets (EEPROMs, clocks, sensors, displays, etc).

Use cases:

▸ Emulation for applications that rely on I2C devices.

▸ Passthrough of physical I2C devices into VMs.

```
                    Device
                   /      \
          Emulated        Passthrough
          I2C             I2C Target
          Target
```

Red Hat

gpio

Device

Line Name	Direction	Value
Power LED	Out	1
Reset Button	In	0
…	…	…

General Purpose Input/Output device with support for edge- and level-triggered interrupt lines.

Lines can have human-readable names.

Use cases:

▸ Interfacing external devices and logic signals with VMs.

Project ACRN: https://projectacrn.github.io/latest/developer-guides/hld/virtio-gpio.html

More about VIRTIO

Want to write a driver? Want to implement a device?

Discussion and questions:

virtio-comment@lists.oasis-open.org

Contribute to the specification:

https://github.com/oasis-tcs/virtio-spec/blob/master/README.md

Thank you

Red Hat is the world's leading provider of enterprise open source software solutions. Award-winning support, training, and consulting services make Red Hat a trusted adviser to the Fortune 500.

linkedin.com/company/red-hat

facebook.com/redhatinc

youtube.com/user/RedHatVideos

twitter.com/RedHat

Red Hat