# Linux Plumbers Conference

Vienna, Austria | September 18-20, 2024

# Zoned Emulation Support for QEMU

Presenter: Sam Li

LINUX PLUMBERS CONFERENCE | Vienna, Austria
Sept. 18-20, 2024

# Collaborators

- Damien Le Moal
- Stefan Hajnoczi
- Dmitry Fomichev
- Hannes Reinecke
- The QEMU community

# 1 State of zoned storage in QEMU

- Virtio-scsi -> attach a zoned device (e.g. ZBC or ZAC HDD) to QEMU
- Virtio-blk emulation -> attach a zoned device or a qcow2 image file to QEMU
- PCI device passthrough -> attach an NVMe PCI device to QEMU
- NVMe device emulation -> ZNS emulation

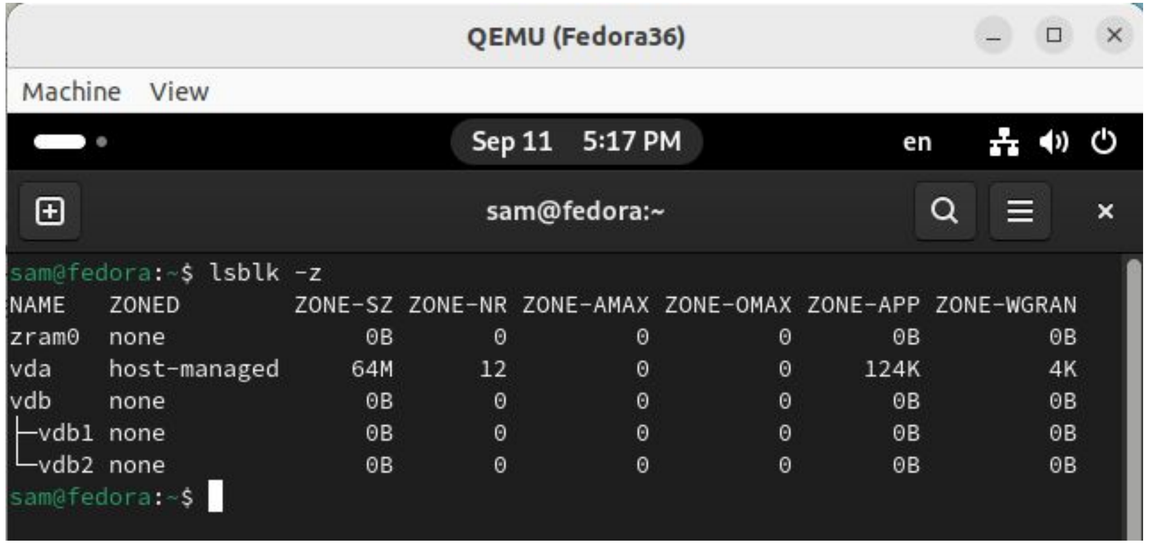
# How to play with the emulated zoned device?

Environment:

- Qemu: v8.1.0 supports zoned device via virtio-blk emulation
- Linux: suggested version > v6.3-rc1

Steps:

1. Create a zoned block device on the host
2. Boots a VM
3. Check if the zoned device is in the guest




# Store files to a zoned storage using btrfs on a QEMU VM

1. Open a QEMU VM
2. Operations on the VM:

   $ mkfs.btrfs -f /dev/vdb

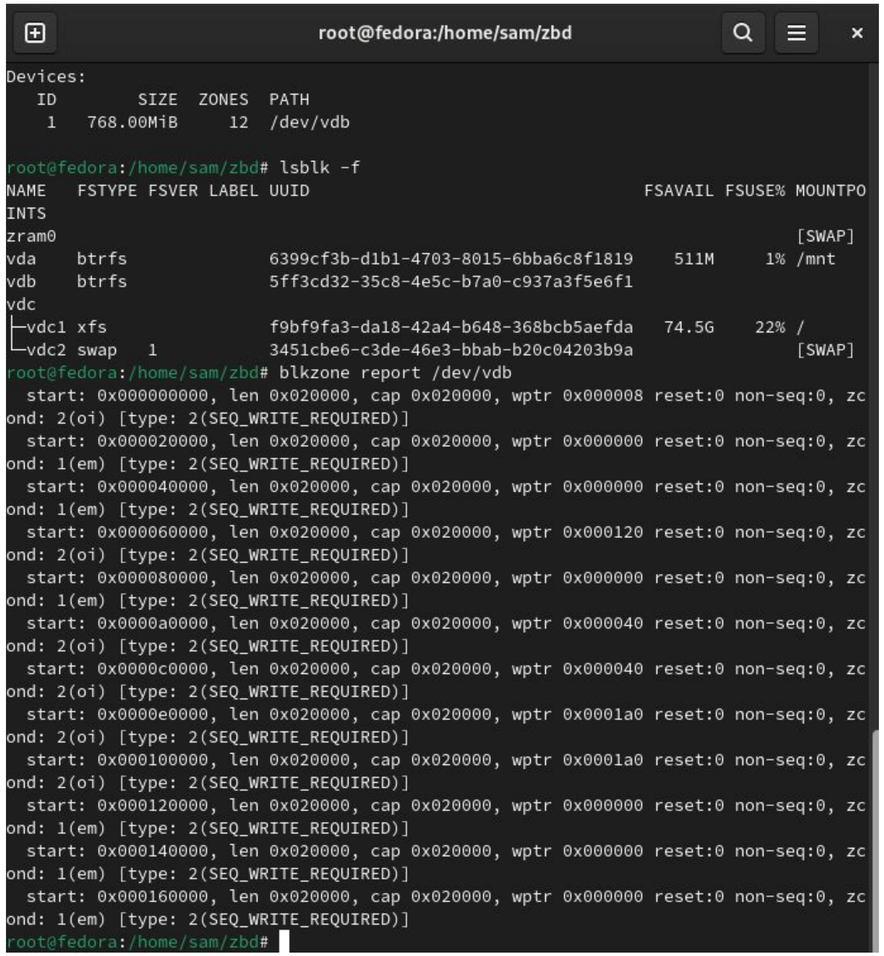
   $ mount -t btrfs /dev/vdb /mnt

   $ btrfs subvolume create /mnt/zoned

   $ echo “echo hello zbd” > hello.txt

   $ mv hello.txt /mnt/zoned

3. Shut down the vm and restart it
4. mount the btrfs again and check if the hello.txt file is there



# 2 Config: two block backends to pick from

1. Null_blk device: modprobe null_blk nr_devices=1 zoned=1
   https://zonedstorage.io/docs/getting-started/zbd-emulation#zoned-block-device-emulation-with-null_blk
2. A qcow2 file with zoned format

```
jli@groves:~/Desktop/infra/qemu$ ./build/qemu-img create -f qcow2 zbc.qcow2
 -o size=768M -o zone.size=64M -o zone.capacity=64M -o zone.conventional_zo
nes=0 -o zone.max_append_bytes=4096 -o zone.max_open_zones=6 -o zone.max_ac
tive_zones=8 -o zone.mode=host-managed
Formatting 'zbc.qcow2', fmt=qcow2 cluster_size=65536 extended_l2=off compre
ssion_type=zlib zone.mode=host-managed zone.size=67108864 zone.capacity=671
08864 zone.conventional_zones=0 zone.max_append_bytes=4096 zone.max_active_
zones=8 zone.max_open_zones=6 size=805306368 lazy_refcounts=off refcount_bi
ts=16
jli@groves:~/Desktop/infra/qemu$
```

# QEMU Command line

1. From the doc, to expose the host's zoned block device through virtio-blk, the command line can be:

   -blockdev node-name=drive0,driver=host_device,filename=/dev/nullb0,cache.direct=on \

   -device virtio-blk-pci,drive=drive0 \

2. To expose the qcow2 file with zoned format through virtio-blk, the command line can be:

   -blockdev node-name=drive1,driver=qcow2,file.driver=file,file.filename=test.qcow2 \

3. To expose the qcow2 file as an emulated zns drive, the command line can be:

   -drive file=${znsimg},id=nvmezns0,format=qcow2,if=none \

   -device nvme-ns,drive=nvmezns0,bus=nvme0,nsid=1,uuid=xxx \

# 3 Develop, test & debug

Test suits for zbd

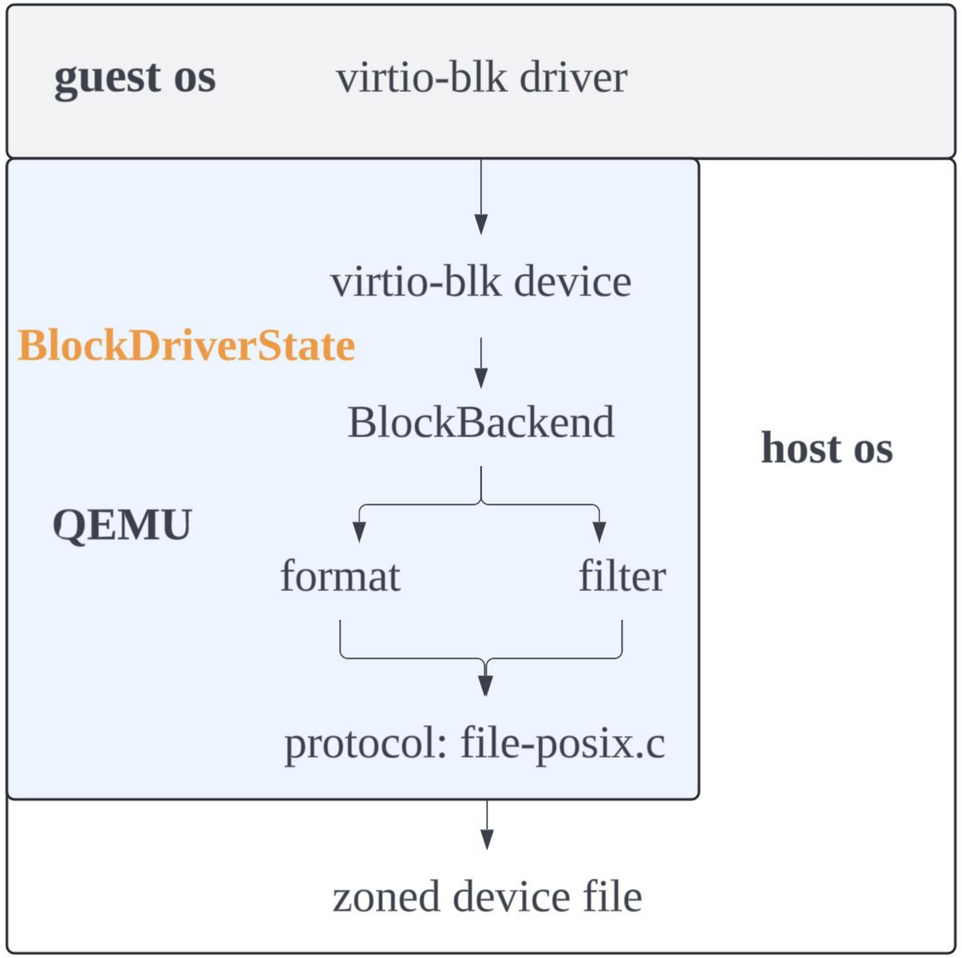
- Qemu-io or qemuio-tests (host)
  $ tests/qemu-iotests/check [<test-case>]
- zonefs-tools
  $ tests/zonefs-tests.sh /dev/vda
- fio/test-zbd, blktests
- dd (zone append), blkzone commands
  $ dd if=/dev/zero of=/mnt/seq/0 oflag=direct,append bs=4096 count=1 conv=notrunc

Debugging tools

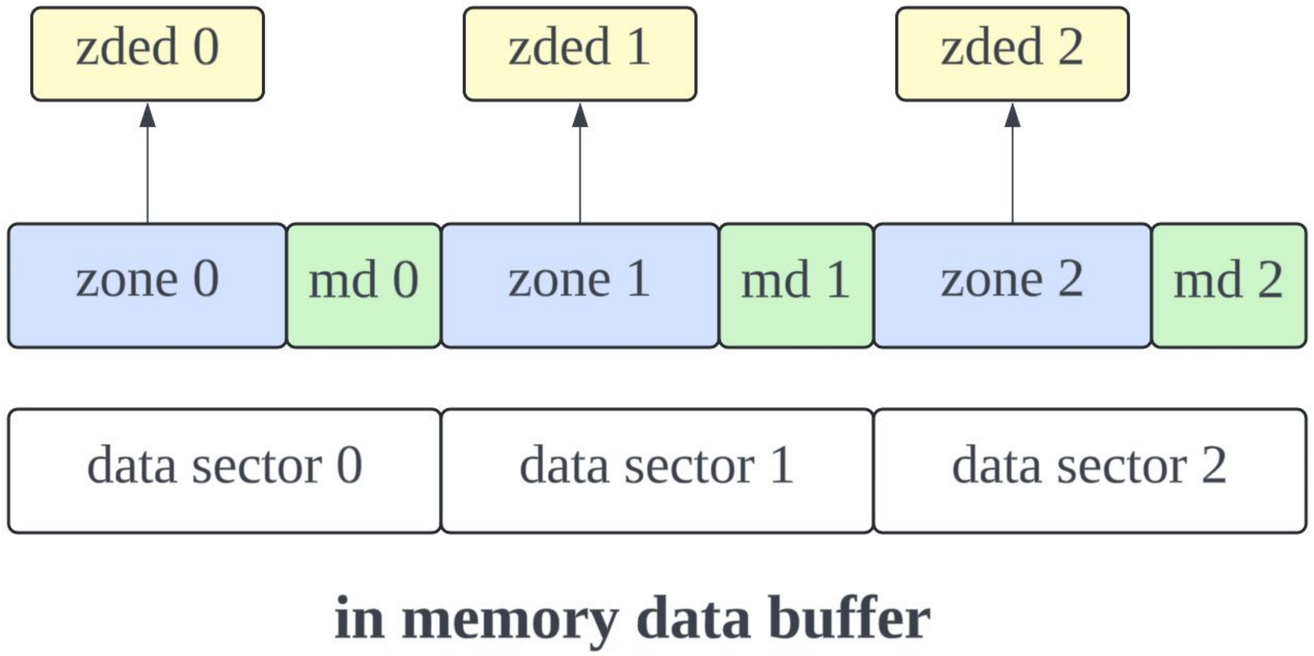
- address sanitizer (config with --enable-sanitizers) or valgrind (host)
- Gdb (bt) + coredump debug control (host)
  $ coredumpctl debug
- ftrace/blktrace
- strace

# 4 Virtio-blk: zoned emulation

- Zoned models: conventional, host-managed, host-aware
- Zone management command: report, open, close, finish
- Zone append: uses write pointer emulation |zone type (1)| write pointer (63)|

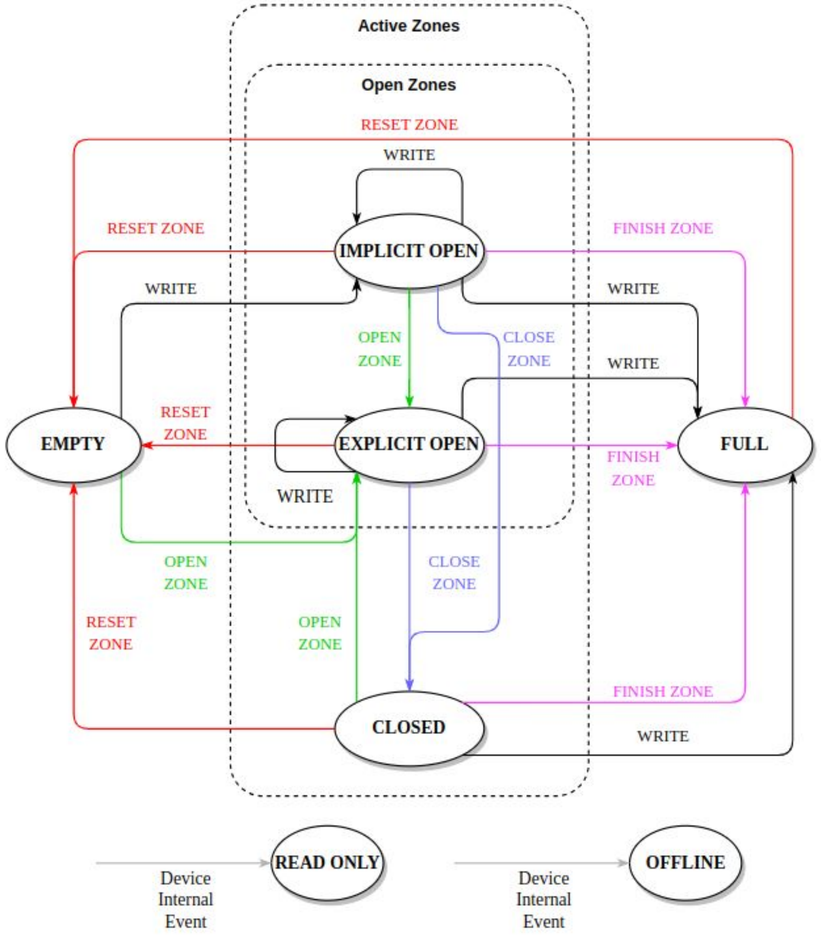


# 5 QCow2: full emulation



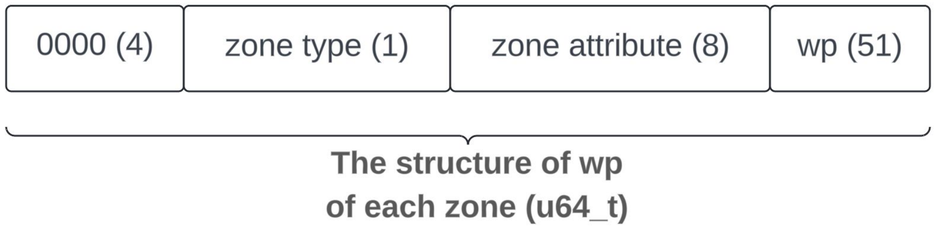
# Metadata persistence


**the layout of disk img**




*Zone State Transitions Overview*

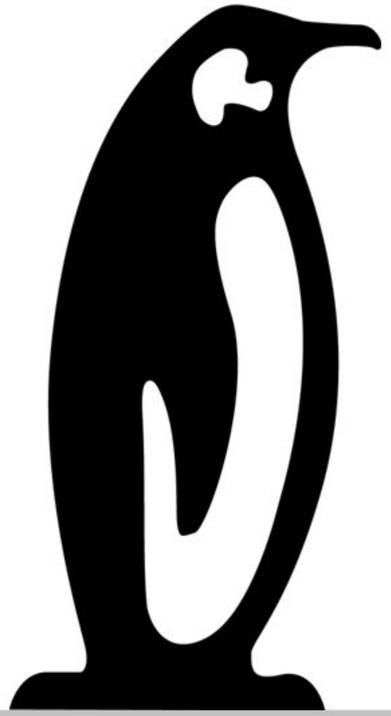
# 6 Persistent states for ZNS emulation




- Address translation
  Nvme_blk_zone_append
  -> dma_blk_zone_append (DMA ops)
  -> blk_aio_zone_append (block layer API)
- Zone attributes
  - ZRWA (zone random write area)
  - ZDED (zone descriptor extension data)

# Contributions

- Added zoned storage APIs to the block layer
- Implemented zoned storage support in virtio-blk emulation
- Add full zoned storage emulation to qcow2 driver (ongoing)
- Add persistence to NVMe ZNS emulation (ongoing)



# Linux Plumbers Conference

Vienna, Austria | September 18-20, 2024