Handling device identity mappings in the IOMMU API

AKA: Please stop abusing RMRRs

Alex Williamson / alex.williamson@redhat.com

What are identity mappings?

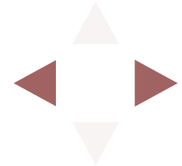
RMRR - Reserved Memory Region Reporting Structure

- Defined in the Intel VT-d spec
- Specified by the BIOS (platform) via ACPI tables
- Identifies a memory region and set of devices
- Requires persistent access between memory & device(s)
- 1:1 address mapping

The spec says...

> *"The RMRR regions are expected to be used for legacy usages (such as USB, UMA Graphics, etc.) requiring reserved memory.* ***Platform designers should avoid or limit use of reserved memory regions since these require system software to create holes in the DMA virtual address range available to system software and its drivers.***"

VT-d spec, rev 2.2, section 8.4



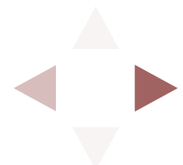
Platform designers...

```
IOMMU: Setting RMRR:
IOMMU: Setting identity map for device 0000:02:00.0 [0xbdf7f000 - 0xbdf8efff] (Smart Array)
IOMMU: Setting identity map for device 0000:01:00.0 [0xbdf7f000 - 0xbdf8efff] (iLO)
IOMMU: Setting identity map for device 0000:01:00.2 [0xbdf7f000 - 0xbdf8efff] (iLO)
IOMMU: Setting identity map for device 0000:03:00.0 [0xbdf7f000 - 0xbdf8efff] (BCM5719)
IOMMU: Setting identity map for device 0000:03:00.1 [0xbdf7f000 - 0xbdf8efff] (BCM5719)
IOMMU: Setting identity map for device 0000:03:00.2 [0xbdf7f000 - 0xbdf8efff] (BCM5719)
IOMMU: Setting identity map for device 0000:03:00.3 [0xbdf7f000 - 0xbdf8efff] (BCM5719)
IOMMU: Setting identity map for device 0000:02:00.0 [0xbdf8f000 - 0xbdf92fff] (Smart Array)
IOMMU: Setting identity map for device 0000:01:00.0 [0xbdf8f000 - 0xbdf92fff] (iLO)
IOMMU: Setting identity map for device 0000:01:00.2 [0xbdf8f000 - 0xbdf92fff] (iLO)
IOMMU: Setting identity map for device 0000:03:00.0 [0xbdf8f000 - 0xbdf92fff] (BCM5719)
IOMMU: Setting identity map for device 0000:03:00.1 [0xbdf8f000 - 0xbdf92fff] (BCM5719)
IOMMU: Setting identity map for device 0000:03:00.2 [0xbdf8f000 - 0xbdf92fff] (BCM5719)
IOMMU: Setting identity map for device 0000:03:00.3 [0xbdf8f000 - 0xbdf92fff] (BCM5719)
IOMMU: Setting identity map for device 0000:02:00.0 [0xbdf93000 - 0xbdf94fff] (Smart Array)
IOMMU: Setting identity map for device 0000:01:00.0 [0xbdf93000 - 0xbdf94fff] (iLO)
```

HP ProLiant DL360p Gen8

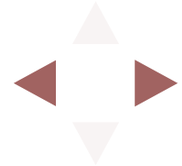
Why is this a problem?

- dma_ops
  - IOMMU driver defines device IOVA space
  - RMRRs are handled properly (mostly)
- IOMMU API
  - API user defines device IOVA space
  - RMRRs are ignored
    - device may continue to write to RMRR area
      - may result in IOMMU faults
    - RMRR IOVA may be remapped
      - potentially worse failure modes



Current solution:

Exclude devices with RMRRs from the IOMMU API



Support requirements

- IOMMU API
  - Retain RMRR mappings
  - Expose RMRR mappings
    - How to handle existing domain mappings?
- IOMMU Groups
  - Group devices that share an RMRR area?!
    - All linked to management controller?
- IOMMU API users - VMs
  - Reserve RMRR region in VM address space
    - No hotplug for RMRR devices
    - What can a guest exploit through an RMRR?

Should we try to support it?

Discuss...