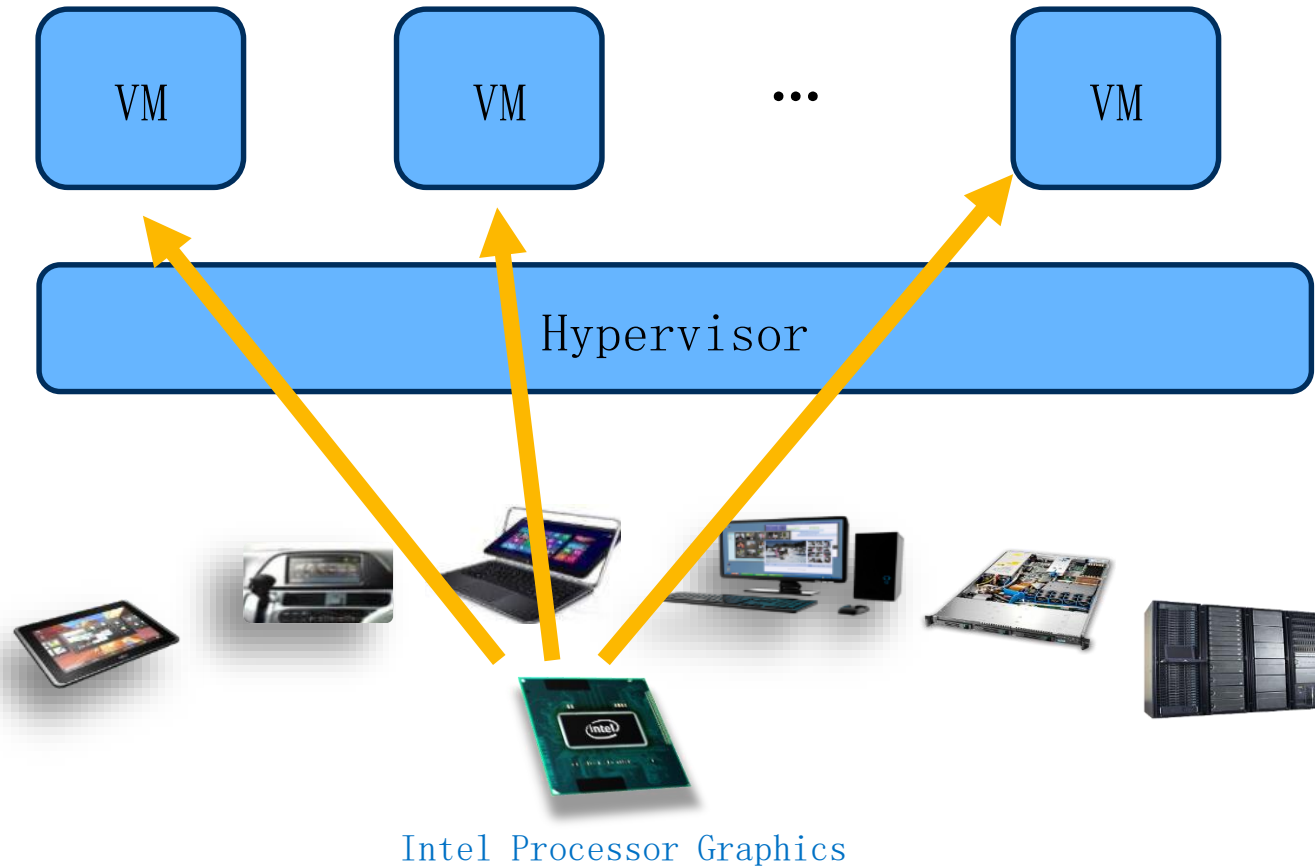


An Introduction to Intel GVT-g (with new architecture)

Zhi Wang zhi.a.wang@intel.com

WW24'17

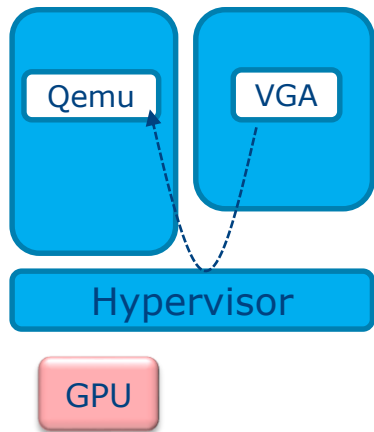
GPU Virtualization



Intel Processor Graphics

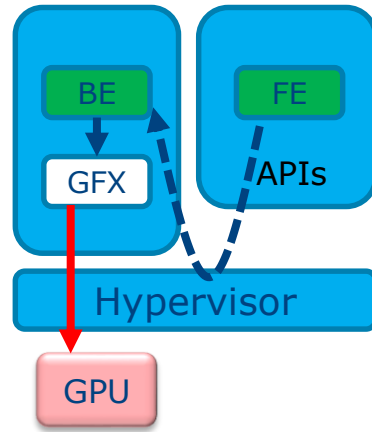
Existing Arts vs Intel GVT-g

Legacy VGA Emulation



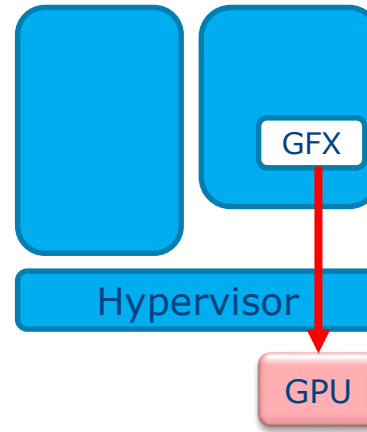
- 2D only

API Forwarding



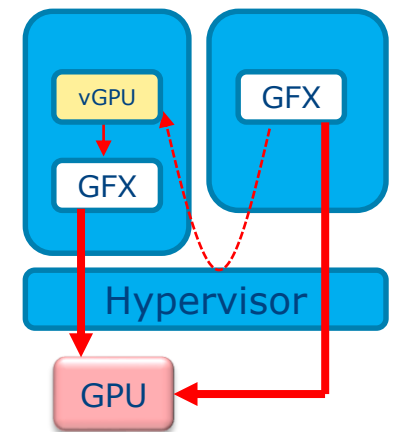
- +3D
- +Sharing
- Media/compute
- Compatibility

Direct Pass-Through



- +3D/media/compute
- +Performance
- No sharing

Full GPU Virtualization

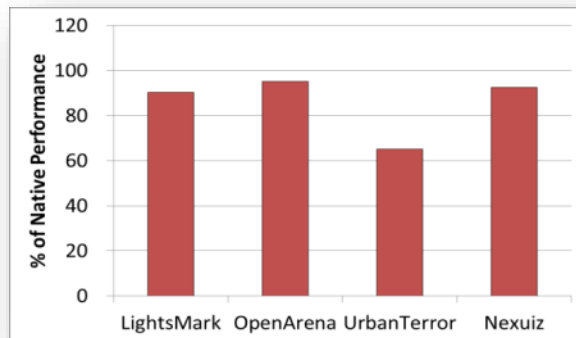


- +3D/media/compute
- +Performance
- +Sharing

Intel GVT-g Capabilities

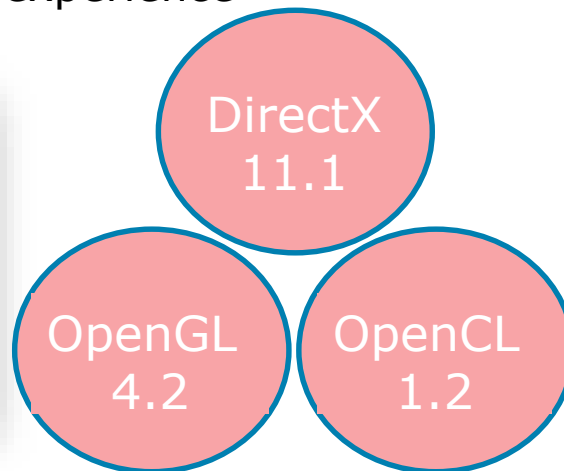
Performance

Near native performance with direct GPU execution



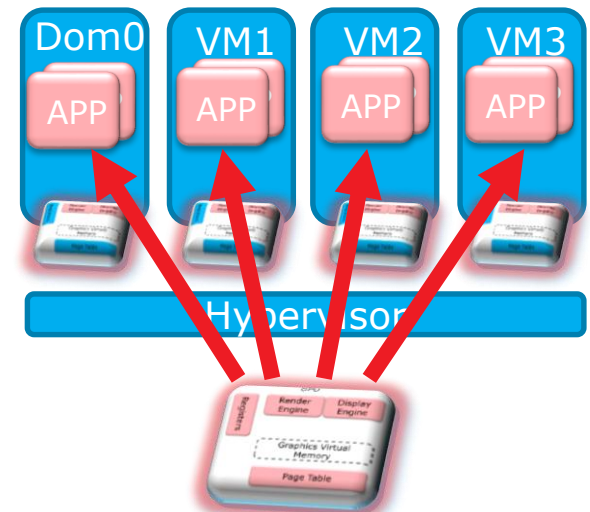
Feature

Run native graphics stack to sustain visual experience

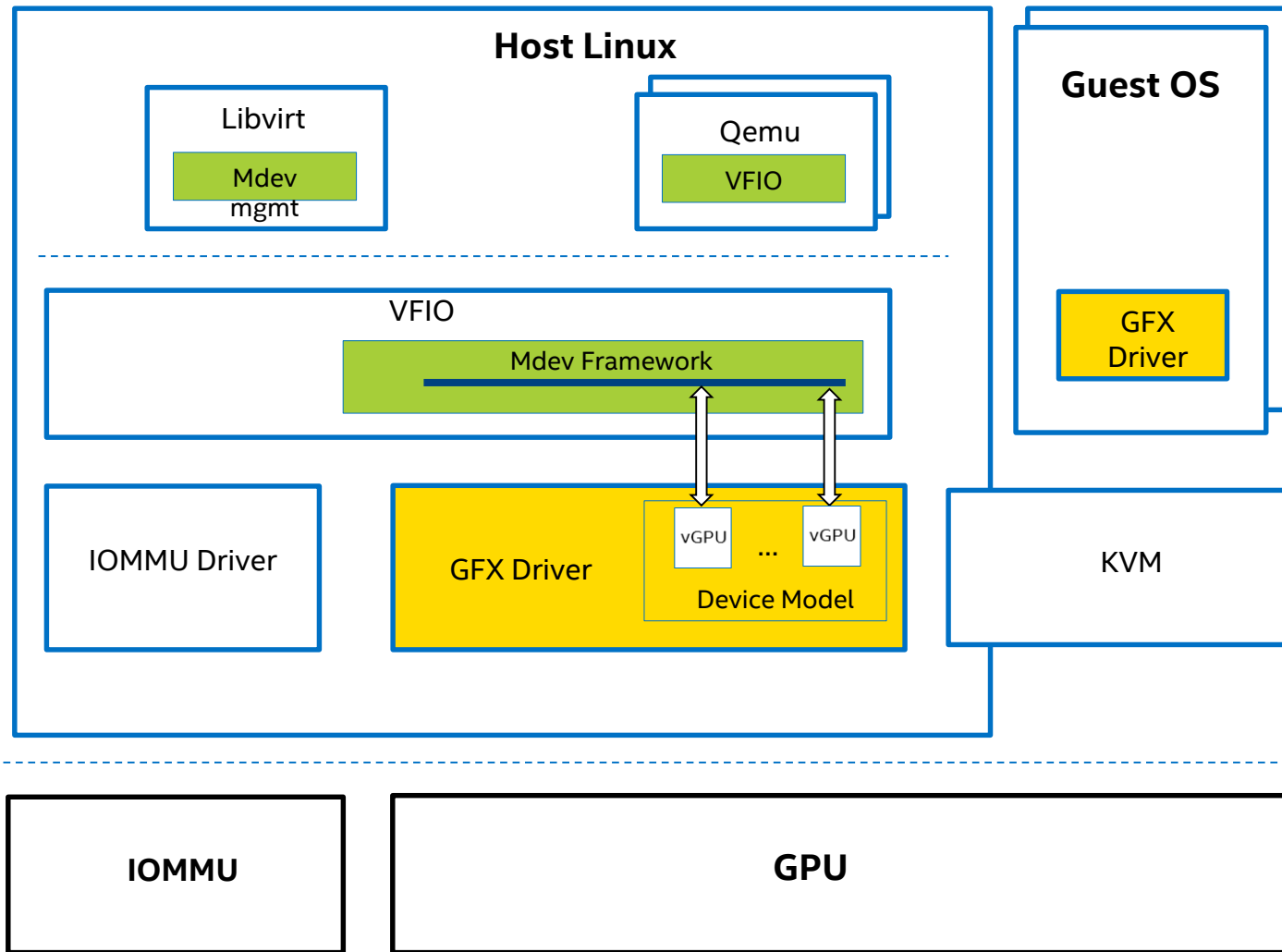


Sharing

Accelerate up to 8 VMs plus Dom0

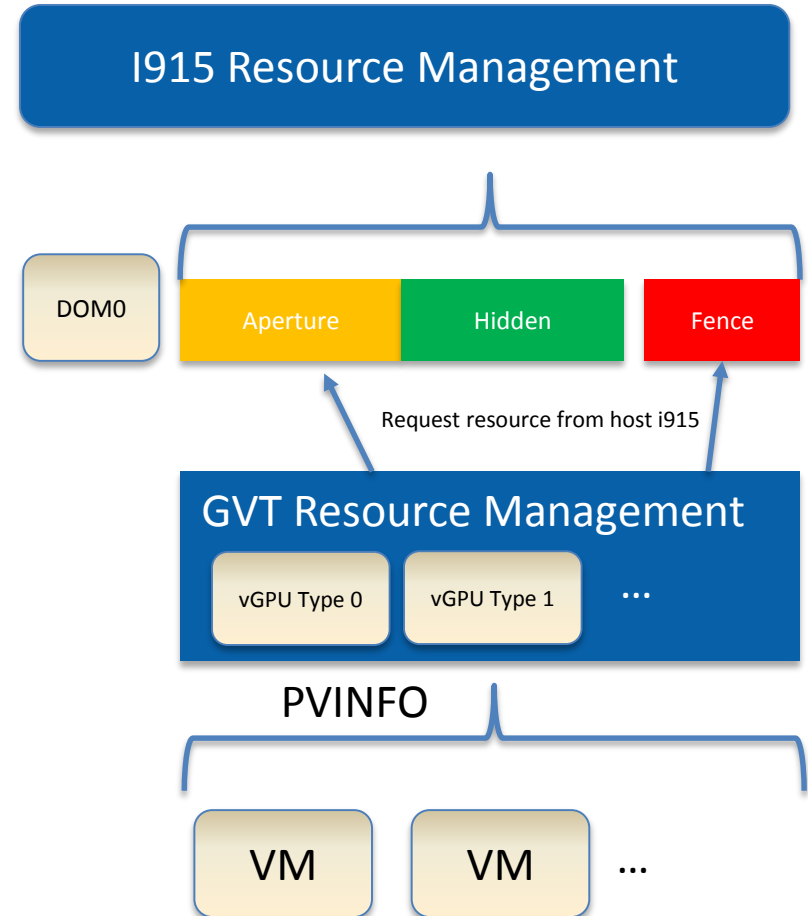


GVT_g (KVMGT) Architecture



Resource Management

- vGPU resources
 - Graphics memory
 - Fence registers
- Request GPU resource from host resource allocator
- GVT manages vGPU resource according to vGPU types

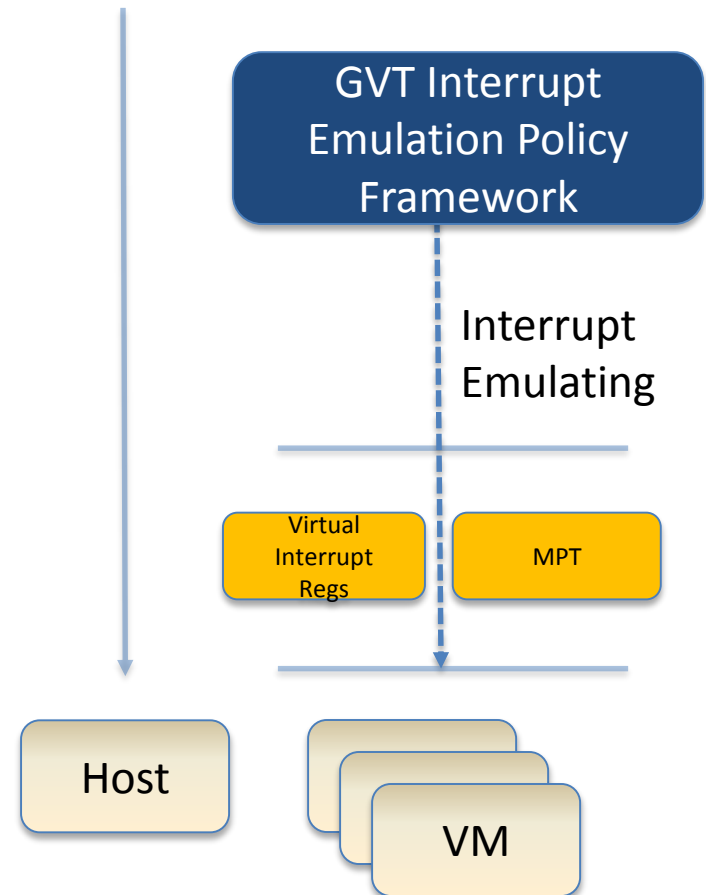


Interrupt

Full GPU interrupt virtualization

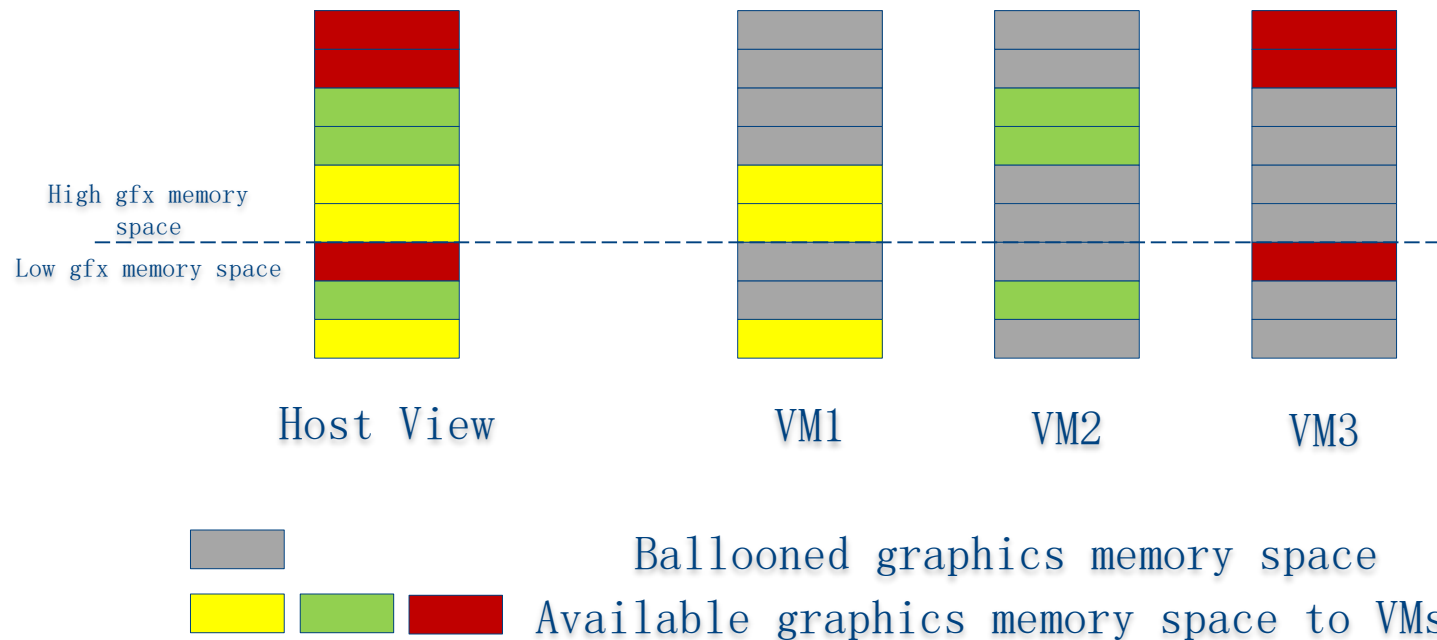
- Display Interrupts
 - VBlank
 - GMBUS
 - AUX Channel
- GPU command Interrupts
 - MI_USER_INTERRUPT
 - PIPE_CONTROL_NOTIFY
- Context Switch Interrupts

GPU Interrupts



Shadow GPU Page Table - GGTT

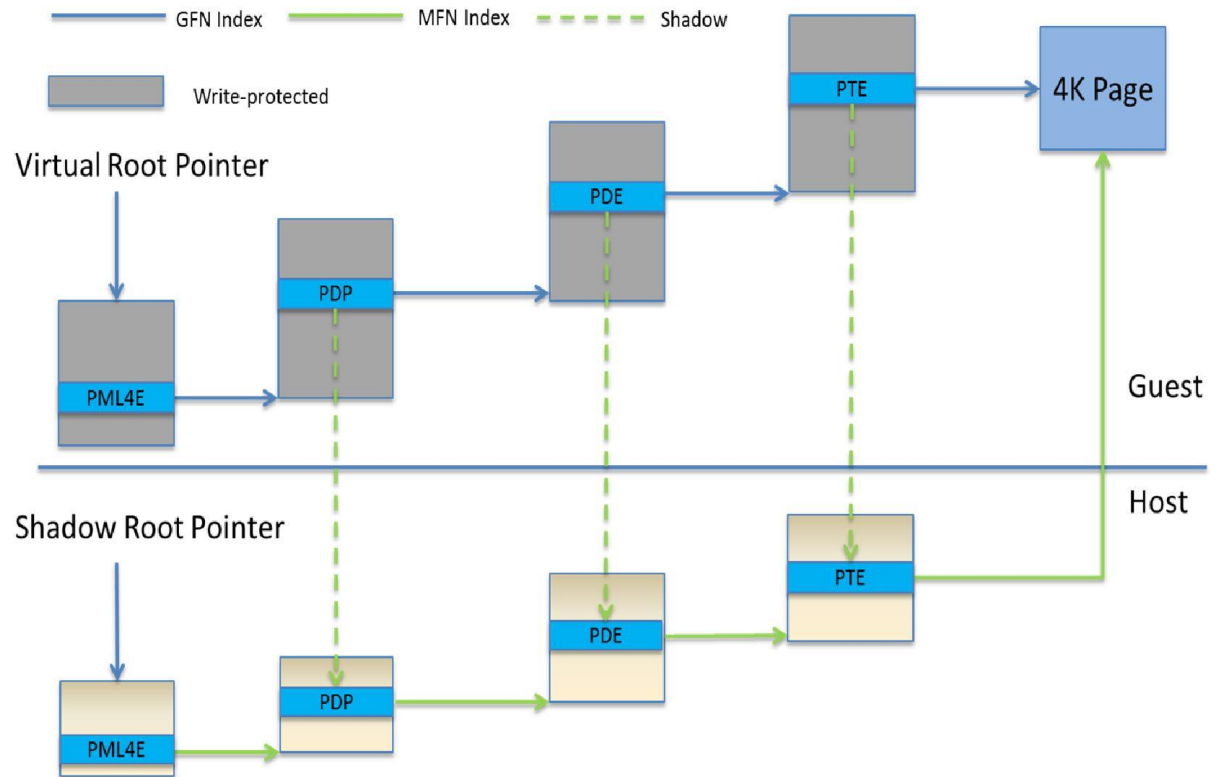
- Global graphics memory space (GGTT) is partitioned
 - Dedicated resource for each VM



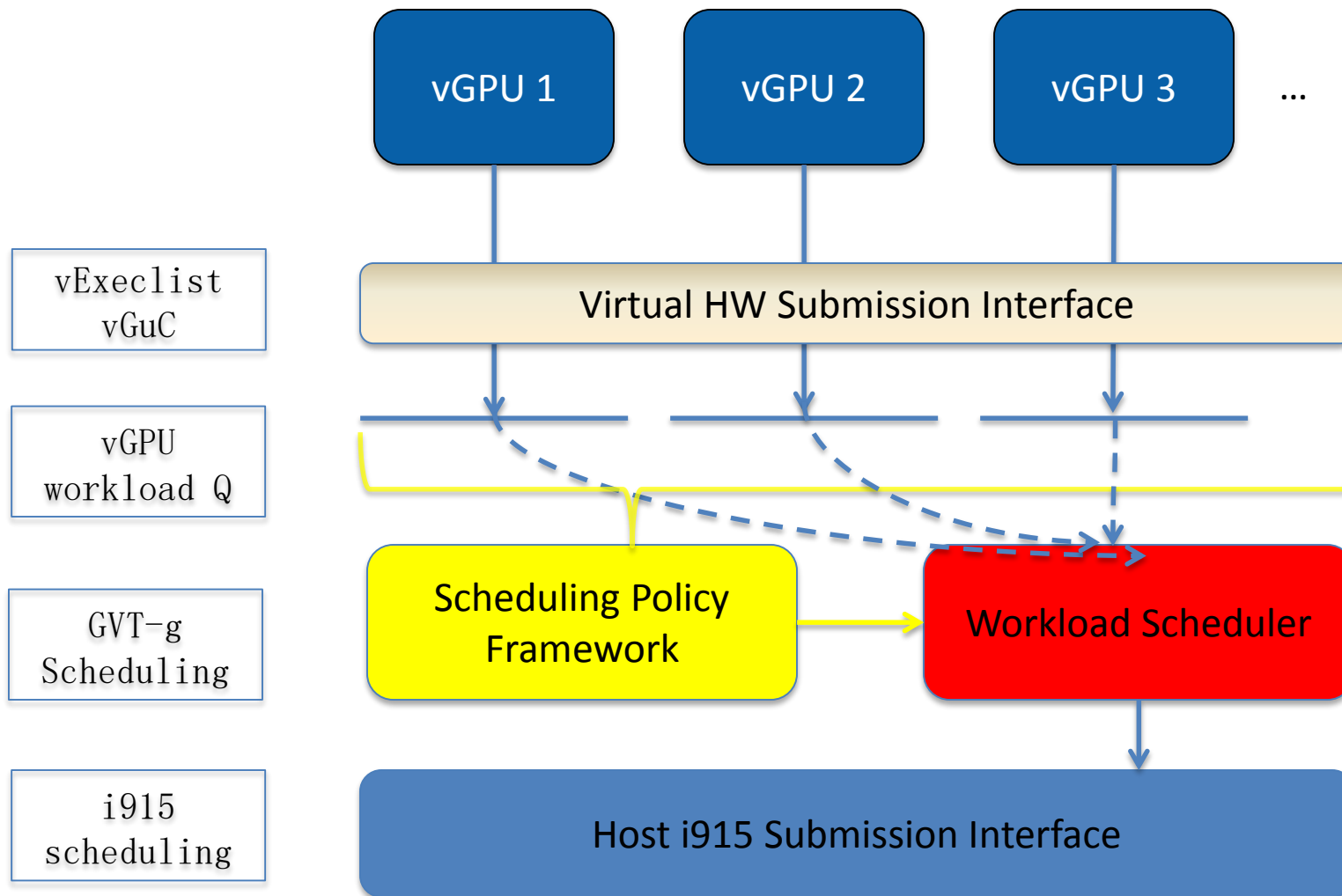
Shadow GPU Page Table - PPGTT

Features

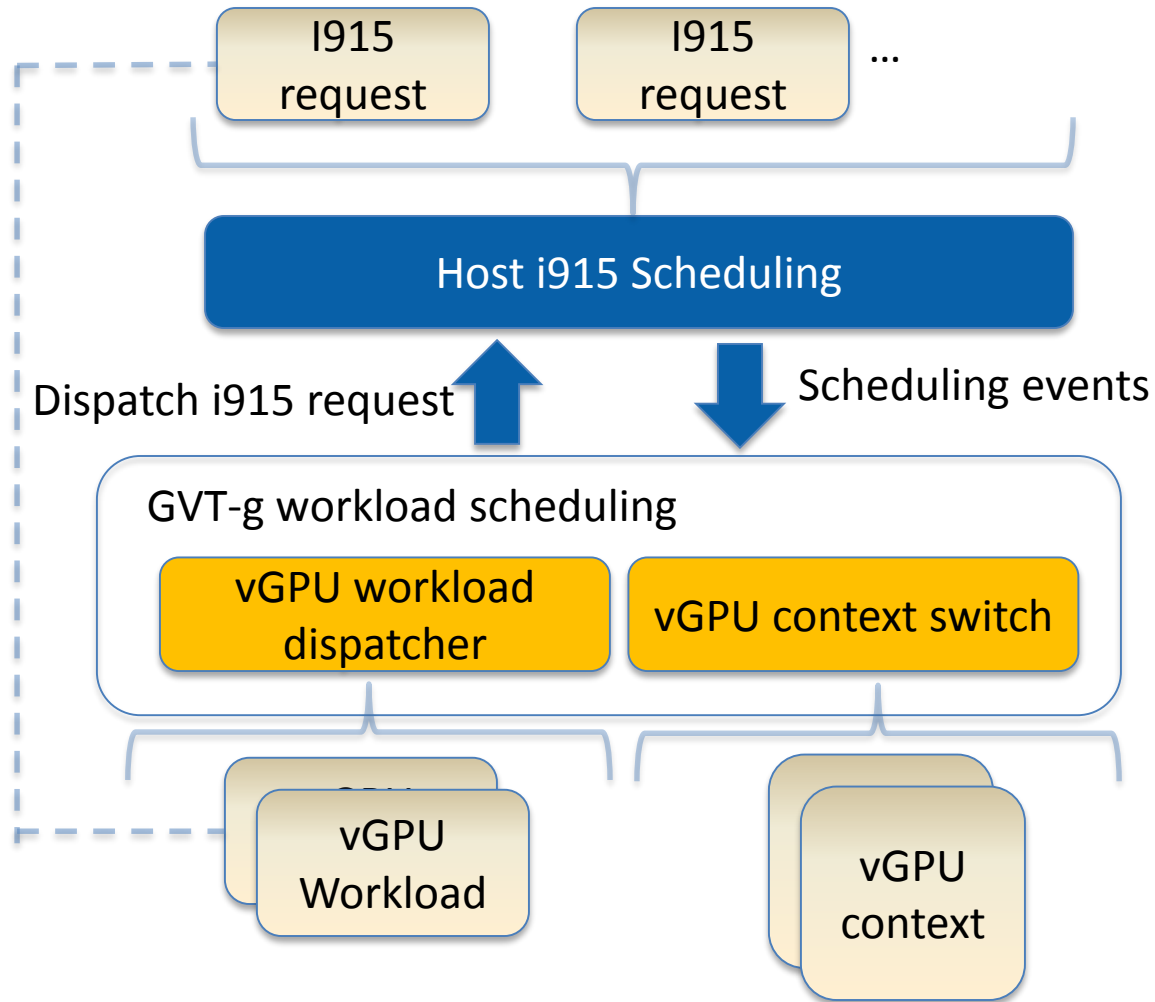
- 2/3/4 level page table
- True per-process PPGTT
- Page table cache
- Reference counting
- Out-of-Sync shadow



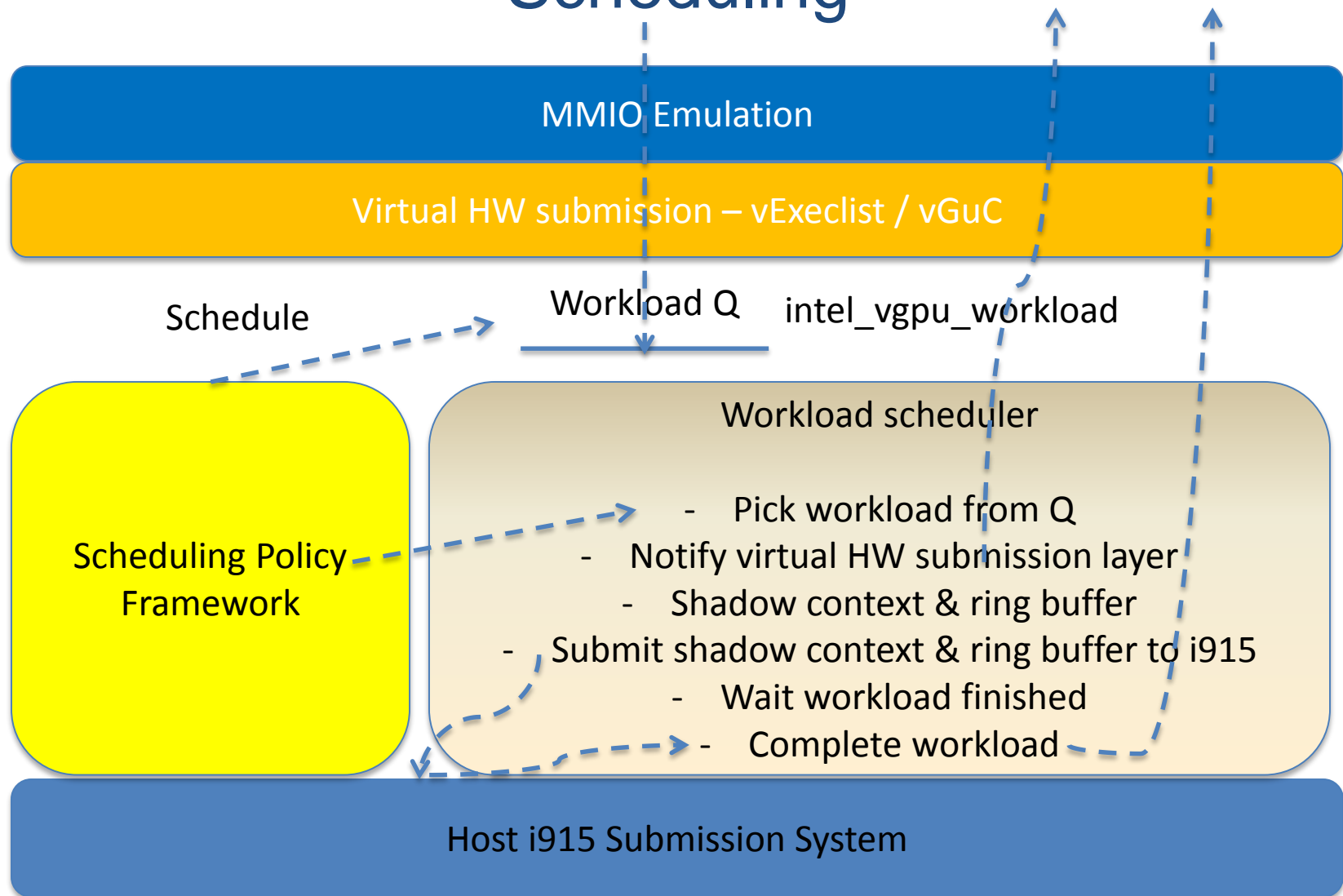
Scheduling



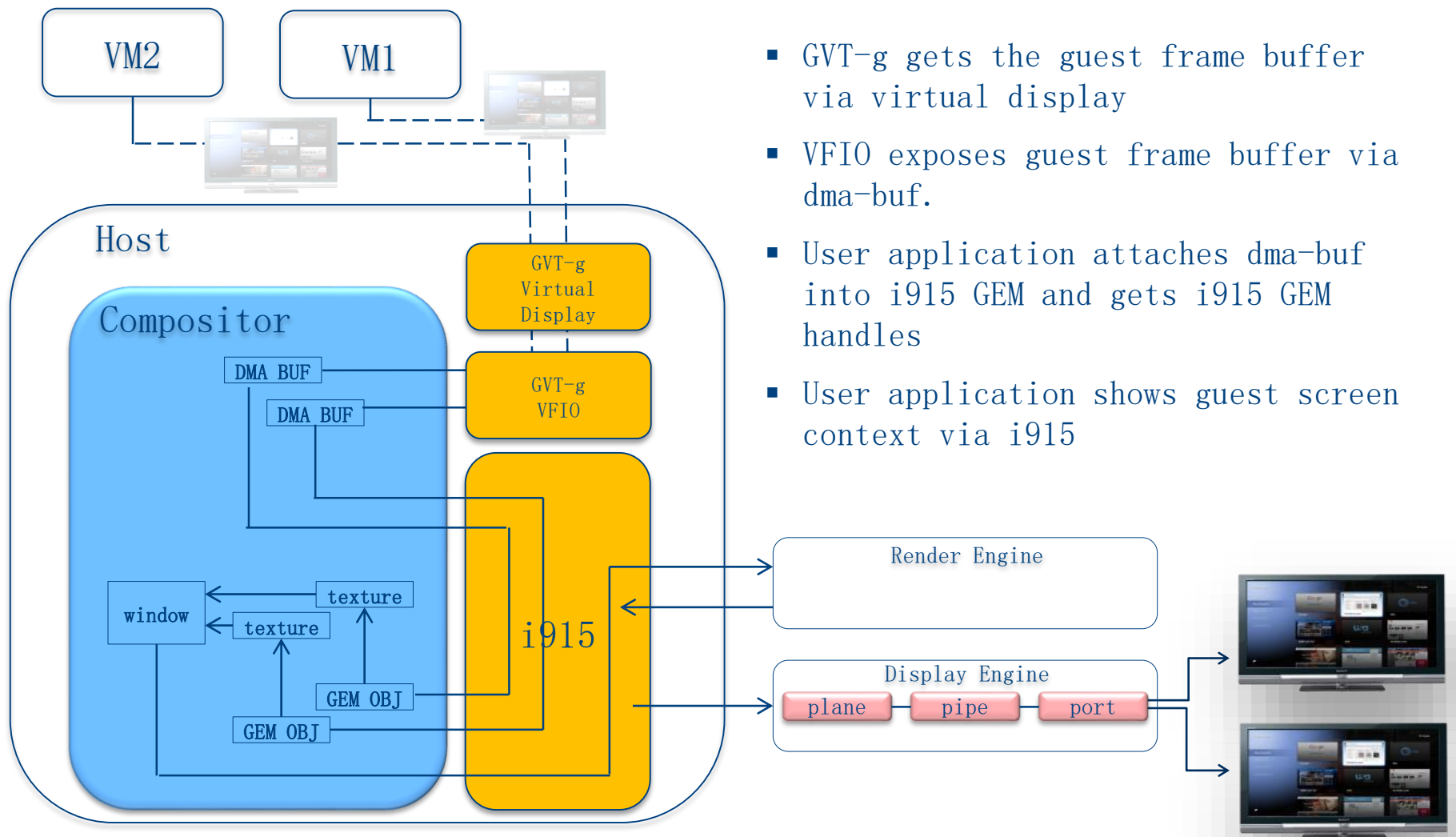
Scheduling



Scheduling



Display



Thanks