Intel Graphics Virtualization Technology Update

Zhi Wang, zhi.a.wang@intel.com
Agenda

• The History
• Intel Graphics Virtualization Technology Update
• New Usage Scenarios
• Upstream Status
• Summary
Intel GPU Virtualization Technology: GVT - g

- A mediated pass-through solution for graphics virtualization
  - Pass-through performance critical resources
  - Trap-and-emulate privileged operations
    - Maintain a device model per VM
- Run native graphics driver in VM
- Achieve good performance and moderate multiplexing capability

---

Device Emulation  Split Driver Model  Mediated Pass-Through  Direct Pass-Through

Performance  Multiplexing
The History

- An open source project based on Xen.
- First published in 2013, codename XenGT.
- First implemented on Haswell - Intel 4th Generation Processor
  - Achieve 80% performance of native Linux in VM
  - Experimental formal support of Windows VM
  - Support up to 3 VMs
  - Target rich virtual client usage
Intel GPU Virtualization Technology Update
From 2013 - 2015
Brand Names

<table>
<thead>
<tr>
<th>API Forwarding (Intel(R) GVT-s)</th>
<th>Direct Pass-through (Intel(R) GVT-d)</th>
<th>Full GPU Virtualization (Intel(R) GVT-g)</th>
</tr>
</thead>
<tbody>
<tr>
<td>VM Frontend</td>
<td>VM Frontend</td>
<td>VM Frontend</td>
</tr>
<tr>
<td>DirectX® APIs</td>
<td>OpenGL® APIs</td>
<td>Graphics Driver</td>
</tr>
<tr>
<td>Backend</td>
<td>APIs</td>
<td>Graphics Driver</td>
</tr>
<tr>
<td>Graphics Driver</td>
<td>Hypervisor</td>
<td>Device Model</td>
</tr>
<tr>
<td>Hypervisor</td>
<td>Hypervisor</td>
<td>Hypervisor</td>
</tr>
</tbody>
</table>

**Pros:**
- Performance
- Sharing

**Cons:**
- No media/GPGPU
- Compatibility

**Pros:**
- Performance
- Capability

**Cons:**
- No sharing

**Pros:**
- Performance
- Capability
- Sharing
Project Information

• Official Website: https://01.org/igvt-g

• Quarterly release model
  • Starting from Q2'14
  • 6 releases till now

• Citrix releases XenClient 5.5 on Jan 2015
  • GVT-g support on HSW is one of the release highlights

• GVT-g guest support got merged into official Linux kernel 4.1 on Feb 2015

<table>
<thead>
<tr>
<th>Supported Guest OSes</th>
</tr>
</thead>
<tbody>
<tr>
<td>32/64bit Windows 7</td>
</tr>
<tr>
<td>32/64bit Windows 8</td>
</tr>
<tr>
<td>32/64bit Ubuntu</td>
</tr>
</tbody>
</table>
Papers and Presentation

Papers
USENIX Annual Technical Conference (USENIC ATC ’14), 2014
• A Full GPU Virtualization Solution with Mediated Pass-Through
USENIX Annual Technical Conference (USENIC ATC ’15), 2015
• Boosting GPU Virtualization Performance with Hybrid Shadow Page Tables

Presentations
Xen Summit 2013
• XenGT: a Software Based Intel Graphics Virtualization Solution

Linux Foundation Collaboration Summit 2013
• XenGT: A Full GPU Virtualization Solution with Mediated Pass-Through

KVM Forum 2014
• KvmGT: A Full GPU Virtualization Solution
New Platform Support - Broadwell

BDW is 5th Generation of Processor Graphics

No Major Changes to High Level Feature Support
- DX11.1+
- OpenCL 2.0*
- OpenGL 4.x

Maintains Similar uArch Partitioning to IVB/HSW
Broadwell vGPU

- More aperture resource allows better scalability (up to 7 VMs)
- New command submission interface - Execution List
  - More self-contained with better programmability than ring buffer on HSW
- Enhanced Per-Process GTT (PPGTT)
  - Page table format changes - 3/4 levels and 64bit PTE
  - Driver usage change - true per-process PPGTT
Challenges to old vPPGTT implementation

- Limitations of old shadow PPGTT
  - only support 2 level page table
  - assume one PPGTT shared by all processes

- True PPGTT may incur more write-protection traps
Enhanced vPPGTT

Features
- 2/3/4 level page table
- True per-process PPGTT
- Page table cache
- Reference counting
- Out-of-Sync shadow
New Usage Scenarios
New Usage Scenario: IVI

Cluster/HeadUnit Consolidation

- Real Time
- TCB size
- Display Cluster
- Infotainment
- ByT POC completed
New Usage Scenario: Visual Cloud

Video Delivery
Store/Stream, Transcode

Cloud Graphics
Gaming, Remote Apps, Rendering

Visual Understanding
Search, Surveillance

GPU-as-a-service

Visual Cloud
GVT-g Case Study: Media Cloud
A Virtualized Media Server with GVT-g

Media Server Pool
(x86 platform with Intel® Processor graphics)

Intel Processor graphics-based Encoding, Decoding, and Composing.
OpenStack Media Cloud

4K Display

4K Camera

1080P Camera

4K Video Conferencing on Media Plane NFV Demo @ MWC’15
Performance challenges

- Memory Intensive
- Real Time Processing
- Mixed Media/3D

Increased GPU virtualization overhead!
Optimizations

15ch-720p_2Mbps_transcoding

Native bar: 757 fps

Optimized: 716.85 fps (95%)

Media workload optimized result

- Smart shadow GPU page table
- Enable cross-engine synchronization
- Increase system memory

Config: I7 4770, Guest Ubuntu* 14.04LTS, 4GB mem, 1.5G GraphicMem, MediaSDK
Architecture

Legacy VGA Emulation (w/o XenGT)

Dom0/Linux

VM1

VM2

Qemu

VG A

VGA Driver

Xen Virtual Platform

i915

Xen

GPU

Full GPU Virtualization (w/ XenGT)

Dom0/Linux

VM1

VM2

Qemu

vgpu

vPCI

VG A

GFX Driver

Balloo

Xen Virtual Platform

MPT Wrapper

MPT Services in Xen

Xen

GPU

Trap-and-emulation

Pass-through
Required MPT Services in Xen

**XEN SUPPORTED**
- Allow vGPU device model to register on given resource ranges
- Forward filtered I/O requests to vGPU device model
- Allow vGPU device model to inject virtual interrupt
- Selective I/O resource pass-through

**UPSTREAM IN PROGRESS**
- On-demand memory write-protection
  - RB-tree based range set

**UPSTREAM NOT START/UNDER DISCUSSION**
- Map/Unmap guest memory
- GPA->HPA translation service
MPT Wrapper Driver in DOM0 Kernel

- General MPT framework defined in vGPU device model
  - Hypervisor agnostic

- MPT wrapper driver as the glue layer
  - Between MPT framework and hypervisor specific services
  - Dynamic registration to MPT framework

- Implementation options
  - Could be a standalone module in host domain (e.g. Xen)
  - Or could be integrated in hypervisor (e.g. KVM)

UPSTREAM NOT START
Other Components

- Linux Guest Support in i915 driver
  - Haswell Support - **DONE**
  - Broadwell – **UPSTREAM IN PROGRESS**

- GVT-g Legacy PCI Device Emulation in QEMU – **UPSTREAM IN PROGRESS**

- GVT-g Trivals in Xen Toolstack – **UPSTREAM NOT START**
Summary

• Great Evolution from 2013
• More and More Usage Models
• Going Forward to Upstream!

Call for actions

• Try and feedback
• Help us to upstream
Thank You