Intel® Graphics Virtualization Technology

Kevin Tian
Graphics Virtualization Architect
Legal Disclaimer

INFORMATION IN THIS DOCUMENT IS PROVIDED IN CONNECTION WITH INTEL PRODUCTS. NO LICENSE, EXPRESS OR IMPLIED, BY ESTOPPEL OR OTHERWISE, TO ANY INTELLECTUAL PROPERTY RIGHTS IS GRANTED BY THIS DOCUMENT. EXCEPT AS PROVIDED IN INTEL’S TERMS AND CONDITIONS OF SALE FOR SUCH PRODUCTS, INTEL ASSUMES NO LIABILITY WHATSOEVER AND INTEL DISCLAIMS ANY EXPRESS OR IMPLIED WARRANTY, RELATING TO SALE AND/OR USE OF INTEL PRODUCTS INCLUDING LIABILITY OR WARRANTIES RELATING TO FITNESS FOR A PARTICULAR PURPOSE, MERCHANTABILITY, OR INFRINGEMENT OF ANY PATENT, COPYRIGHT OR OTHER INTELLECTUAL PROPERTY RIGHT.

A "Mission Critical Application" is any application in which failure of the Intel Product could result, directly or indirectly, in personal injury or death. SHOULD YOU PURCHASE OR USE INTEL’S PRODUCTS FOR ANY SUCH MISSION CRITICAL APPLICATION, YOU SHALL INDEMNIFY AND HOLD INTEL AND ITS SUBSIDIARIES, SUBCONTRACTORS AND AFFILIATES, AND THE DIRECTORS, OFFICERS, AND EMPLOYEES OF EACH, HARMLESS AGAINST ALL CLAIMS COSTS, DAMAGES, AND EXPENSES AND REASONABLE ATTORNEYS' FEES ARISING OUT OF, DIRECTLY OR INDIRECTLY, ANY CLAIM OF PRODUCT LIABILITY, PERSONAL INJURY, OR DEATH ARISING IN ANY WAY OUT OF SUCH MISSION CRITICAL APPLICATION, WHETHER OR NOT INTEL OR ITS SUBCONTRACTOR WAS NEGLIGENCE IN THE DESIGN, MANUFACTURE, OR WARNING OF THE INTEL PRODUCT OR ANY OF ITS PARTS.

Intel may make changes to specifications and product descriptions at any time, without notice. Designers must not rely on the absence or characteristics of any features or instructions marked "reserved" or "undefined". Intel reserves these for future definition and shall have no responsibility whatsoever for conflicts or incompatibilities arising from future changes to them. The information here is subject to change without notice. Do not finalize a design with this information. The products described in this document may contain design defects or errors known as errata which may cause the product to deviate from published specifications. Current characterized errata are available on request. Contact your local Intel sales office or your distributor to obtain the latest specifications and before placing your product order.

Copies of documents which have an order number and are referenced in this document, or other Intel literature, may be obtained by calling 1-800-548-4725, or go to: http://www.intel.com/design/literature.htm

Copyright ©2014 Intel Corporation.
GPU Virtualization

VM

VM

...

VM

Hypervisor

3D

Media

Compute

Intel Processor Graphics
GPU Cloud

Video Delivery
Store/Stream, Transcode

Cloud Graphics
Gaming, Remote Apps, Rendering

Visual Understanding
Search, Surveillance

GPU-as-a-service

GPU Cloud
Internet video traffic is forecasted to grow at 29% CAGR and will represent 69% of consumer traffic by 2017.
Media Delivery Example

1. Video files store in cloud storage
2. Media transcoding before deliver to CDN
3. Media transcoding for client in Edge

Content Provider

Data Center

Content Deliver Network

Edge Cloud

Realtime, Video Conf, video awareness

Transport Edge (RAN, IP, Cable)

Service Provider

RAN

IP

Cable

Servers

Laptop

SDN

NFV

VM

Content Deliver Network

OTT, VOD, Offline Video ...

Servers

Video files store in cloud storage

Media transcoding before deliver to CDN

Media transcoding for client in Edge

Edge Cloud

Realtime, Video Conf, video awareness
Requirements of GPU Virtualization

- **Performance**: Direct GPU acceleration
- **Capability**: Consistent visual experience
- **Sharing**: Multiple Virtual Machines
# GPU Virtualization Approaches

<table>
<thead>
<tr>
<th>API Forwarding</th>
<th>Direct Pass-thru</th>
<th>Full GPU Virtualization</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image1.png" alt="Diagram" /></td>
<td><img src="image2.png" alt="Diagram" /></td>
<td><img src="image3.png" alt="Diagram" /></td>
</tr>
</tbody>
</table>

**Pros:**
- Performance
- Sharing

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

**Pros:**
- Performance
- Capability

**Cons:**
- No sharing

**Pros:**
- Performance
- Capability
- Sharing
Intel® Graphics Virtualization Technology

- Intel® GVT-s for API level GPU sharing
  - Make existing API forwarding protocols running best on Intel platform

- Intel® GVT-d for direct GPU passthrough
  - Xen GPU passthrough upstreaming in progress
  - KVM PoC patch in community

- Intel® GVT-g for vGPU based sharing
  - Achieve a good balance of performance, feature and sharing
  - Xen implementations (a.k.a XenGT) in production quality on HSW
    - BDW support is in alpha quality in 2015/Q1 release
  - KVM support (a.k.a KVMGT) in prototype quality on HSW
Intel GVT-g: Full GPU Virtualization

Performance
- 3DMark: 80%
- H.264 transcoding: 90% (of native performance)

Feature
- Native driver
  - DirectX 11.1, OpenGL 4.2, OpenCL 1.2, MediaSDK x.x

Sharing
- Simultaneously accelerate multiple VMs

- HSW
- BDW
- SKL
- ...
Building Blocks

- Host OS or VM
- OpenStack
  - Nova
  - Control point
- ISV Guest OS
  - Ubuntu* 14.04 LTS
  - Windows* 7 64bit
  - Windows 8 32bit
- Intel® Xeon® E3 Processor
- OpenStack
  - Network node
  - Storage/Image
- XeNT
  - Intel® HD Driver
- KVM Hypervisor
- Intel® Media SDK
- Intel HD Driver
- ISV Guest OS
  - Video Conference
  - Media Server
  - VoD streaming
- Real Time Transcoding
- Cloud Infrastructure
- Media Server
  - Real Time Transcoding
- Video Conference
- Media Server
  - VoD streaming
- Intel® Media SDK
- Intel® HD Driver
- 4th generation Intel® Core™ i7
- GPU instance
- ISV/Non Intel
- Intel Hardware
- Intel Own
- Intel Contributed
Video Conference Usage Case

4K Display

1080P Camera

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

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

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

High Density workload: 480p transcoding

High Density workload: 1080p transcoding

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

Avg S90% of native H.264 transcoding performance
CPU Transcoding vs. GPU Transcoding

Lab data. Config: I7 4770, Guest Windows 7 _x64, 4GB mem, 1.5G GraphicMem, MediaSDK

Performance boost with Intel GVT-g!
Q&A