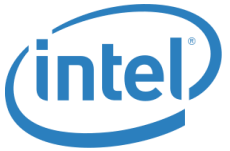
A Full GPU Virtualization Solution with Mediated Pass-Through

田坤 (Intel)
kevin.tian@intel.com

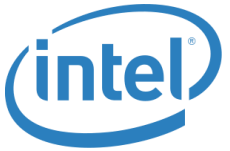
Legal Disclaimer

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 NEGLIGENT 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



Agenda

- Background of GPU Virtualization
- Full GPU Virtualization with Mediated Pass-Through
- Intel® Graphics Virtualization Technology Update

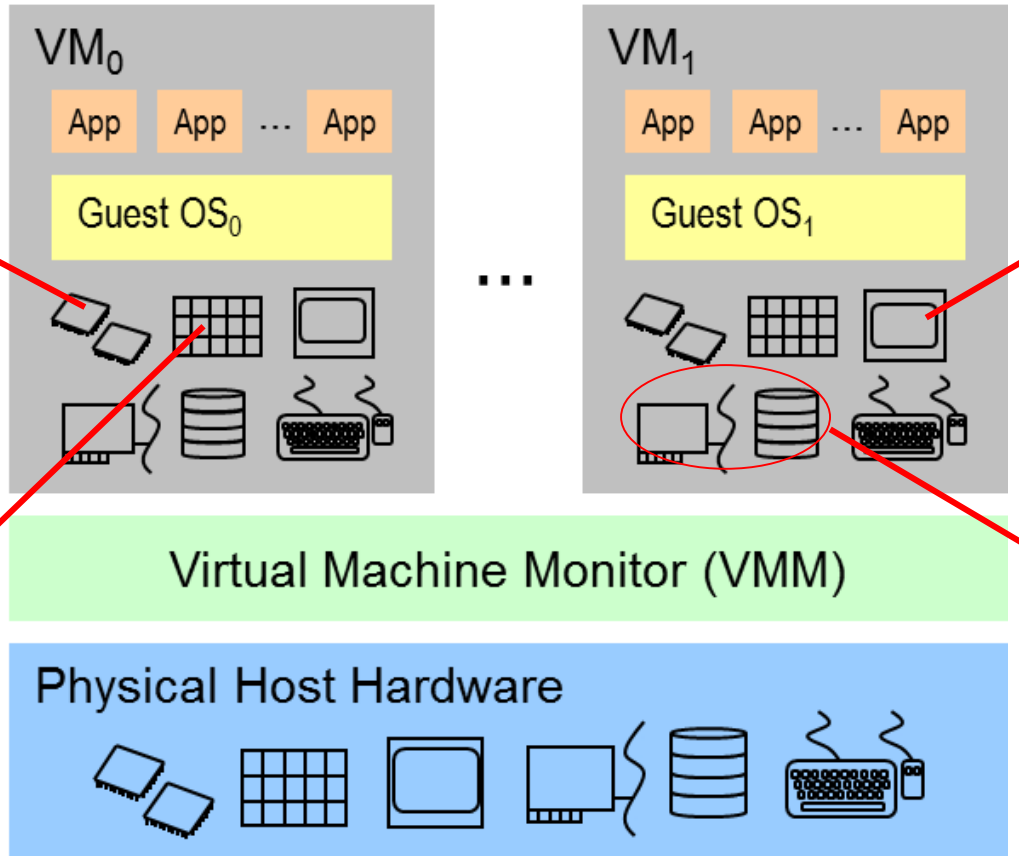


The Evolution of Virtualization Technologies

Binary Translation
Para-virtualization

2005-2006

HW Full Virtualization
(VT-x)



?

Shadow Page Table

2007-2008

HW 2-level Page Table
(EPT)

Emulated I/O
PV drivers

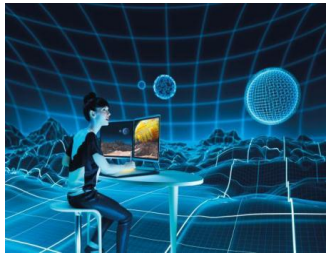
2008-NOW

Device Pass-through
(VT-d & SR-IOV)

Momentum of GPU Virtualization

Video Delivery

Store/Stream, Transcode



Cloud Graphics

Gaming, Remote Apps, Rendering

Visual Understanding

Search, Surveillance

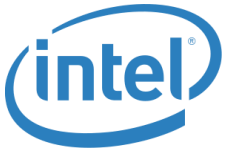
GPU-as-a-Service

Momentum of GPU Virtualization (Cont.)

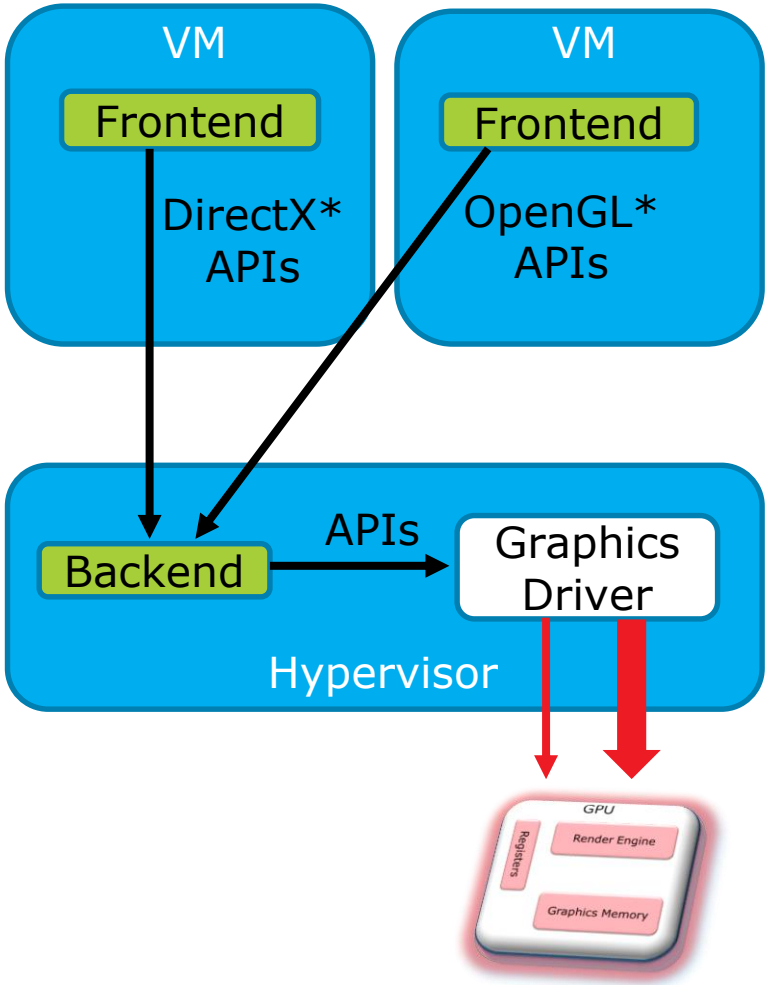
Rich Client Experience/ Multi-Personality

Head Consolidation/IVI

GPU HW acceleration
is becoming a basic feature in a VM !



API Forwarding



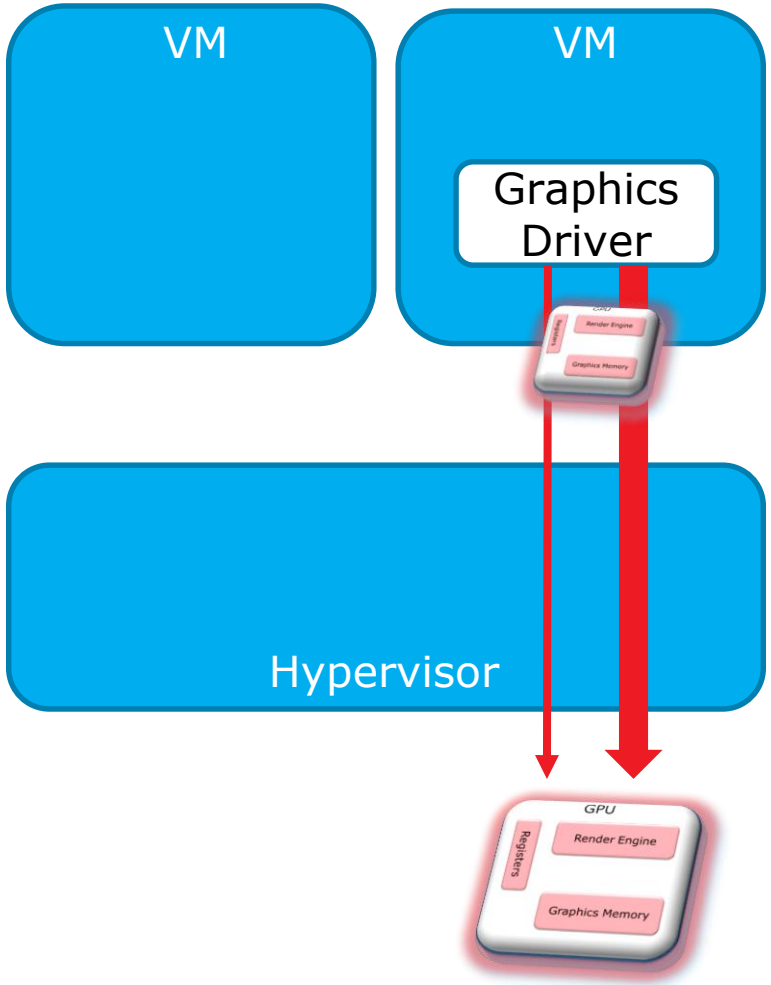
Pros

- Performance
- Scalability

Cons

- Lagging features
- Incompatible APIs
- Maintenance burden

Direct Pass-Through



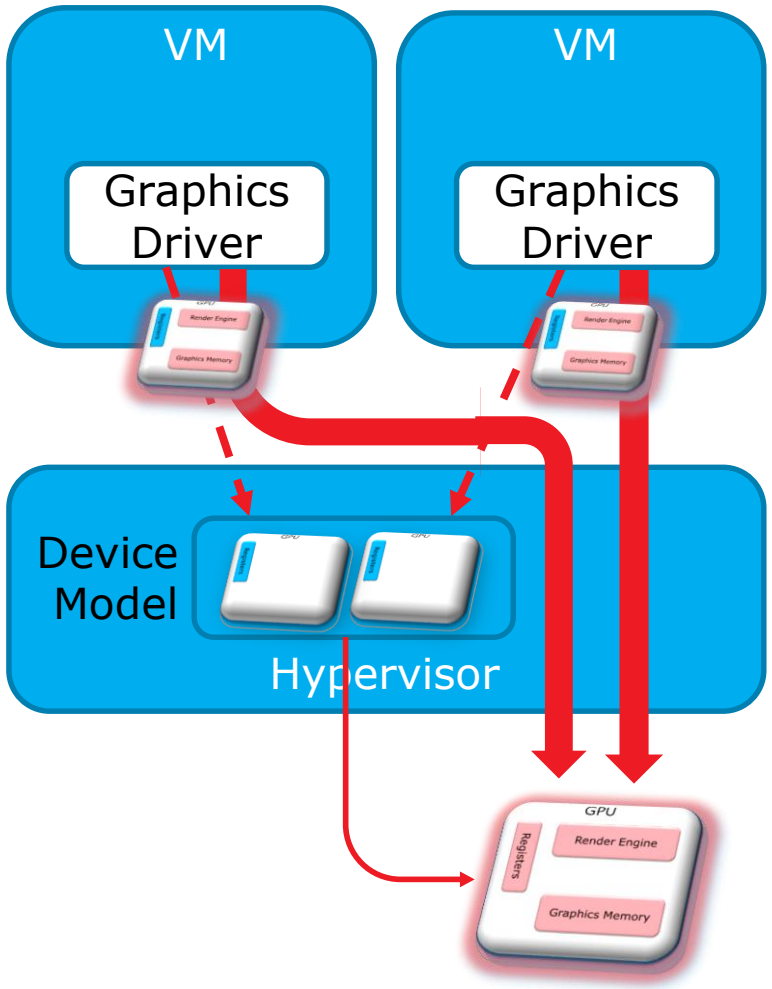
Pros

- Performance
- Full features

Cons

- No or limited sharing (w/ SRIOV)
- Break VM live migration

Mediated Pass-Through



Pros

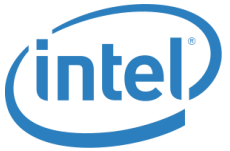
- Performance
- Full feature
- Scalability
- Allow VM live migration

Cons

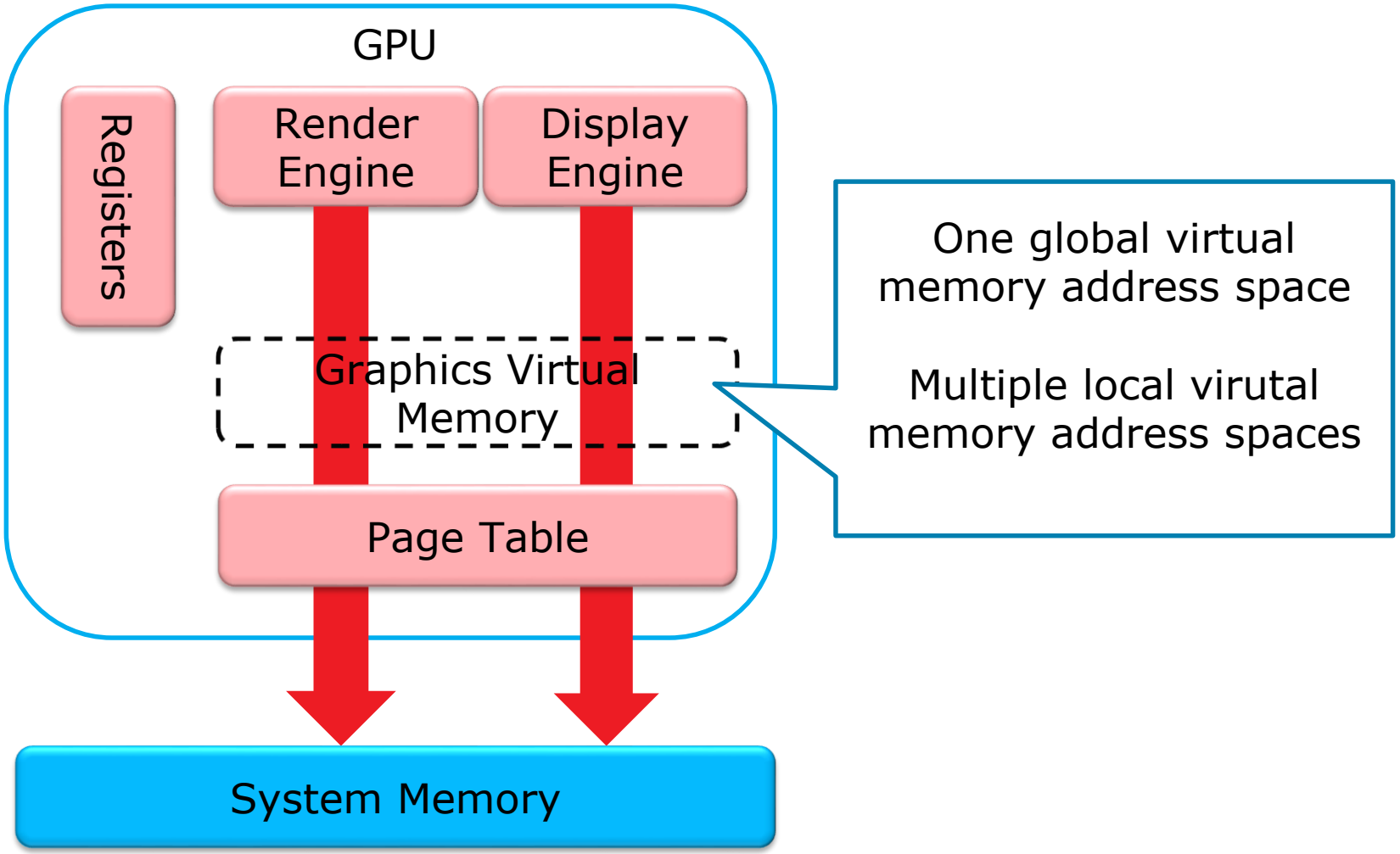
- Vendor specific

Agenda

- Background of GPU Virtualization
- Full GPU Virtualization with Mediated Pass-Through
- Intel® Graphics Virtualization Technology Update

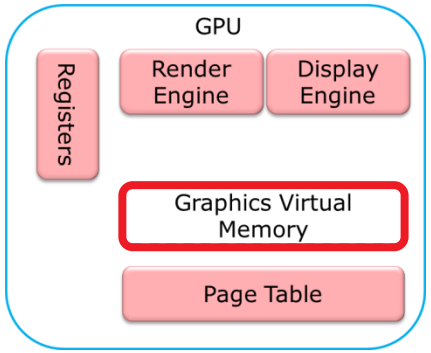


Processor Graphics Overview



Mediated Pass-Through Policies

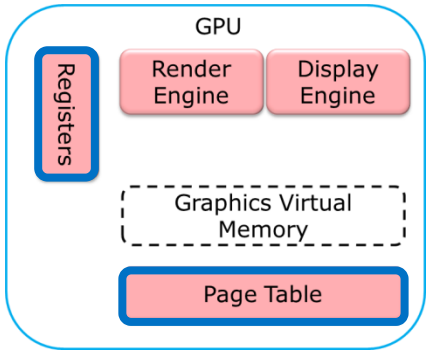
Pass-Through



- Partitioning global address space
- Switching local address spaces

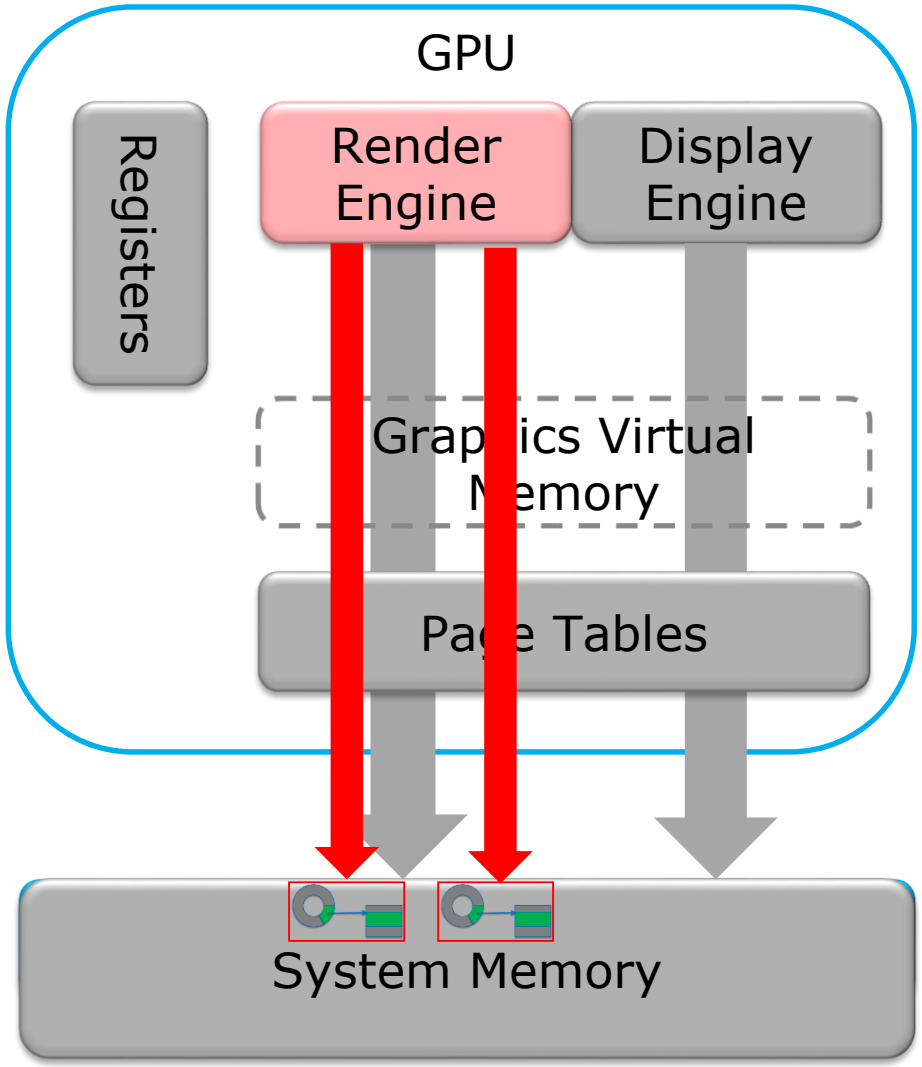
10x-100x more accesses

Trap-and-Emulation



- Emulate vGPU device model
- Shadow GPU page tables

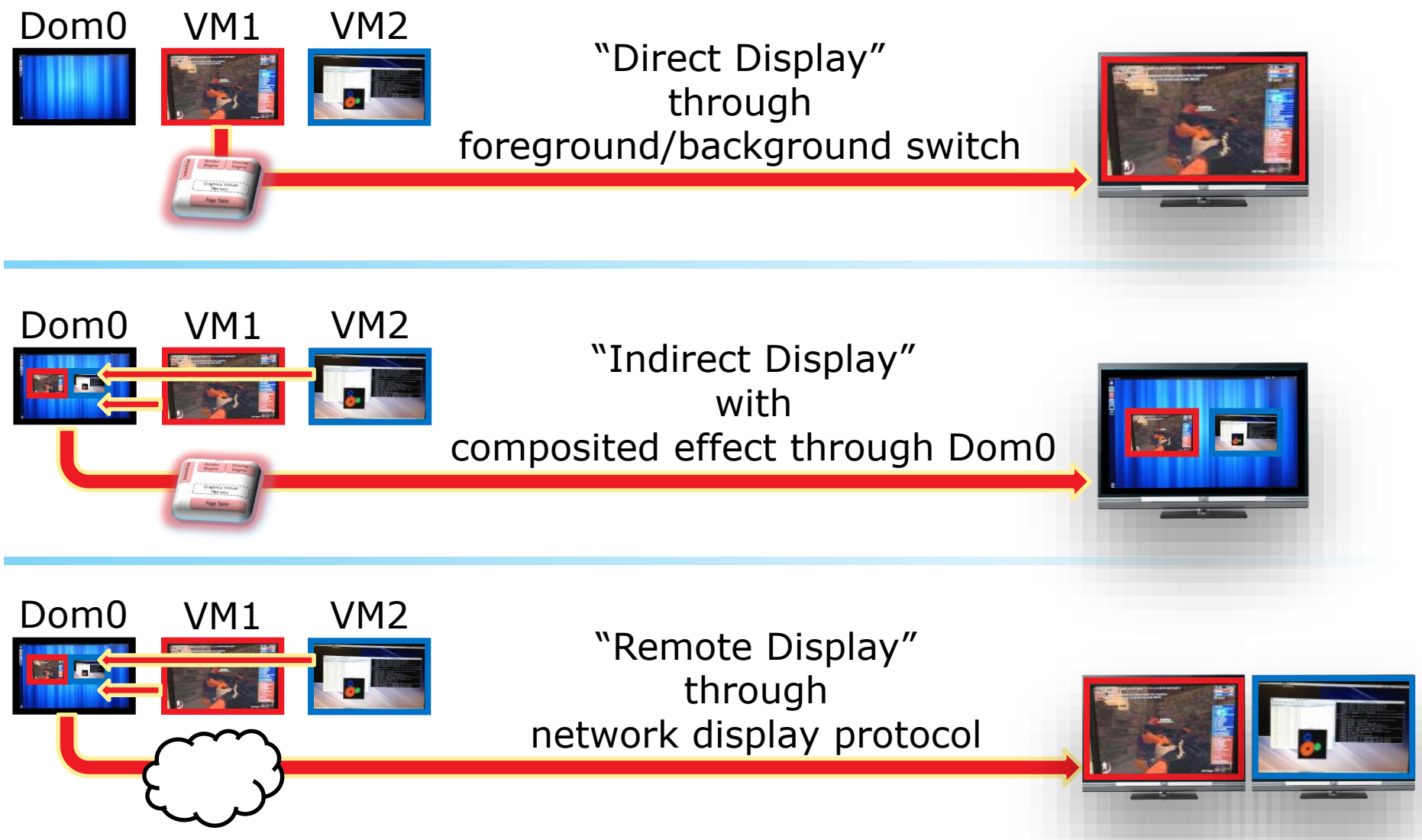
Render Engine Sharing



Direct execution of guest command buffer

Time-based sharing

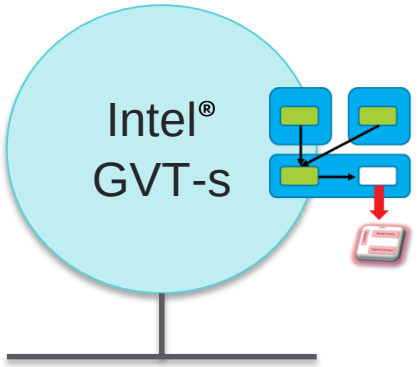
Display Virtualization



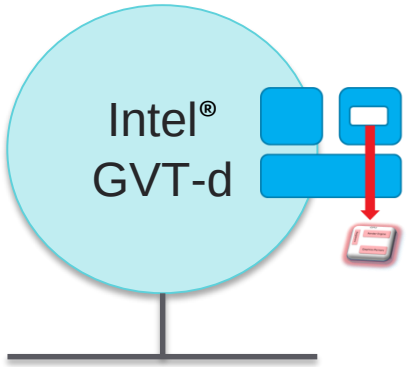
Agenda

- Background of GPU Virtualization
- Full GPU Virtualization with Mediated Pass-Through
- Intel® Graphics Virtualization Technology Update

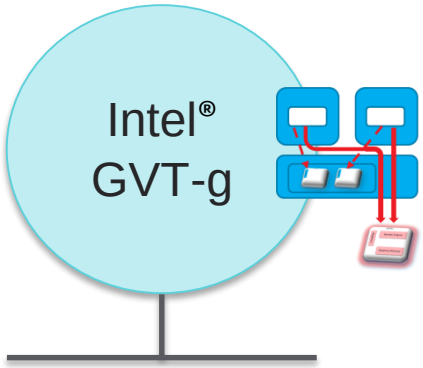
Intel® Graphics Virtualization Technologies



API forwarding



Direct Pass-through



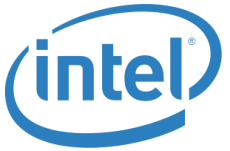
Full GPU Virtualization

Driver extension to support 3rd party remoting protocol

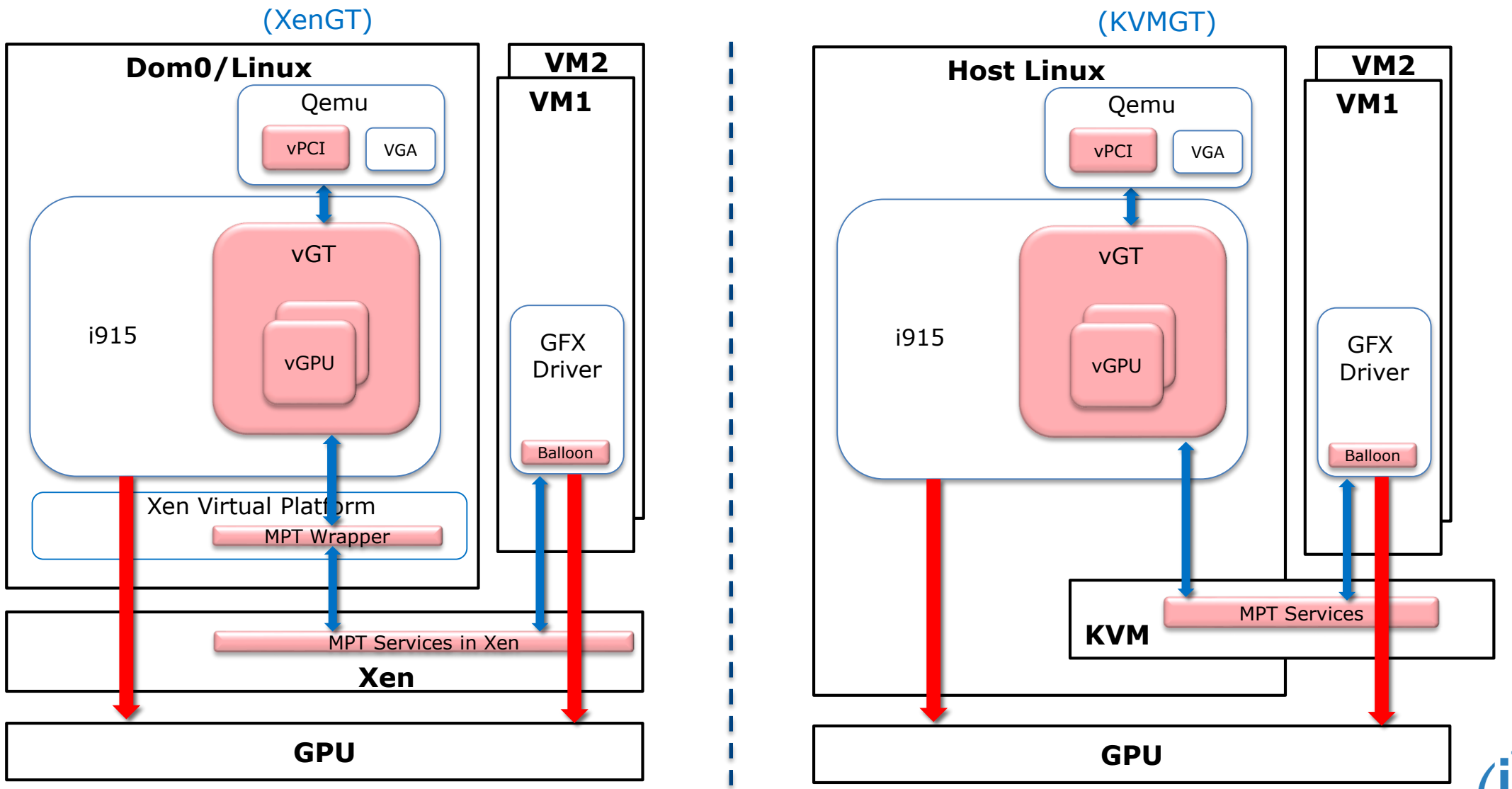
Intel core hypervisor technologies

Intel® GVT-g

- Project started in 2012, for client virtualization usage
  - Formally released on HSW
- Now mainly focus on server/cloud usages
  - Target BDW+ server platform
  - An SoC/BYT porting is also available for IVI virtualization
- Dual-licensed (MIT/GPL) open source project
  - Quarterly releases (latest 2015-Q3)
- Support both Xen (XenGT) and KVM (KVMGT)
  - Share vGPU device model (~90% LLC)

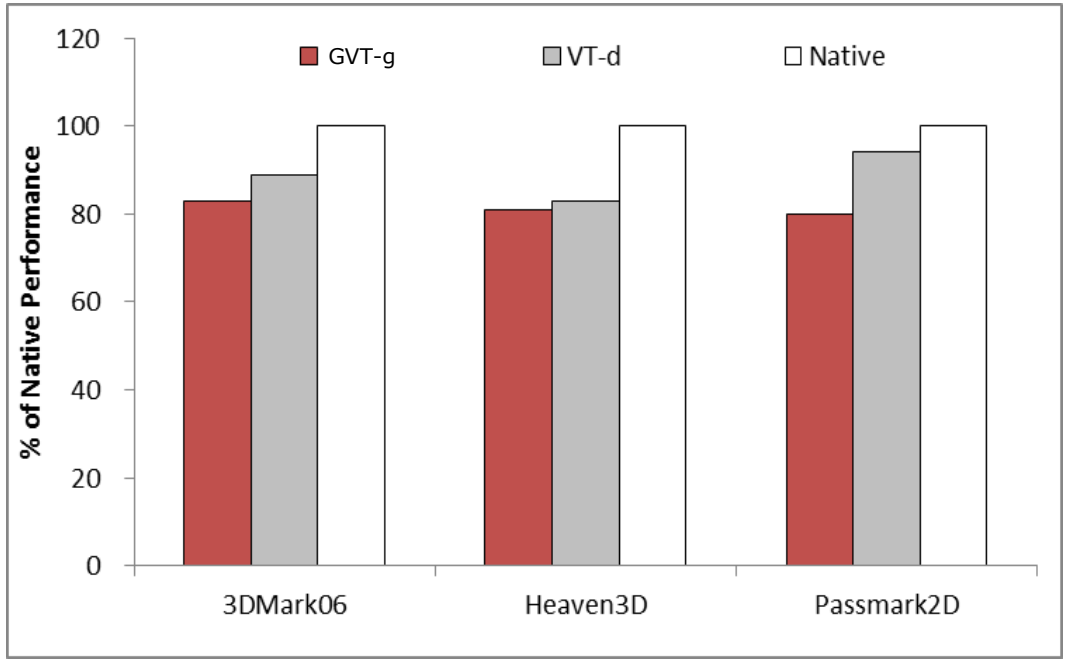


XenGT vs. KVMGT

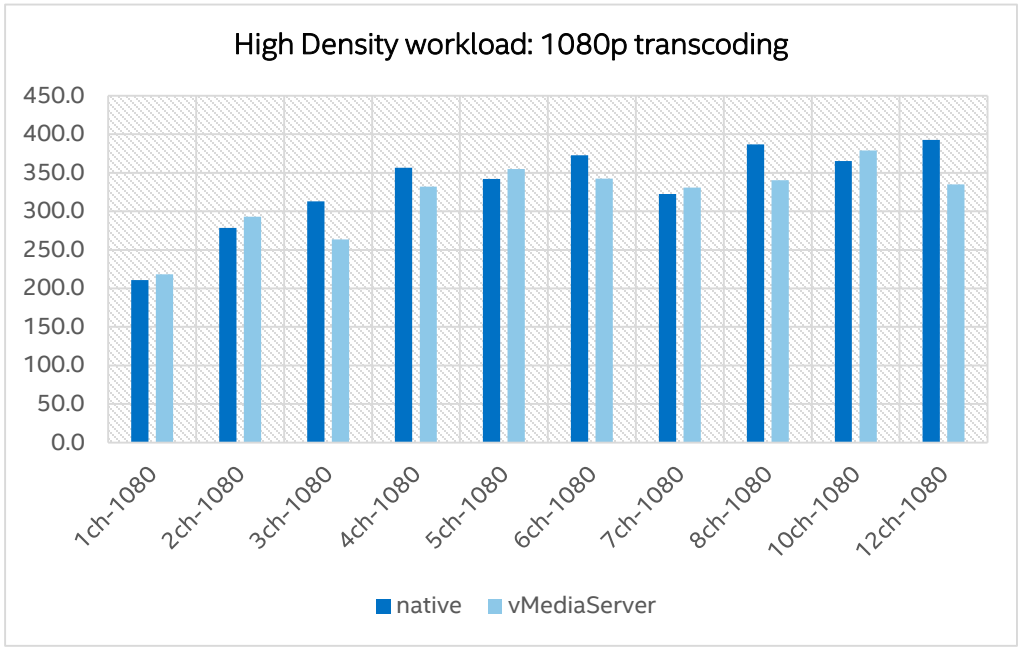


Performance

Average 80% of native 3D performance

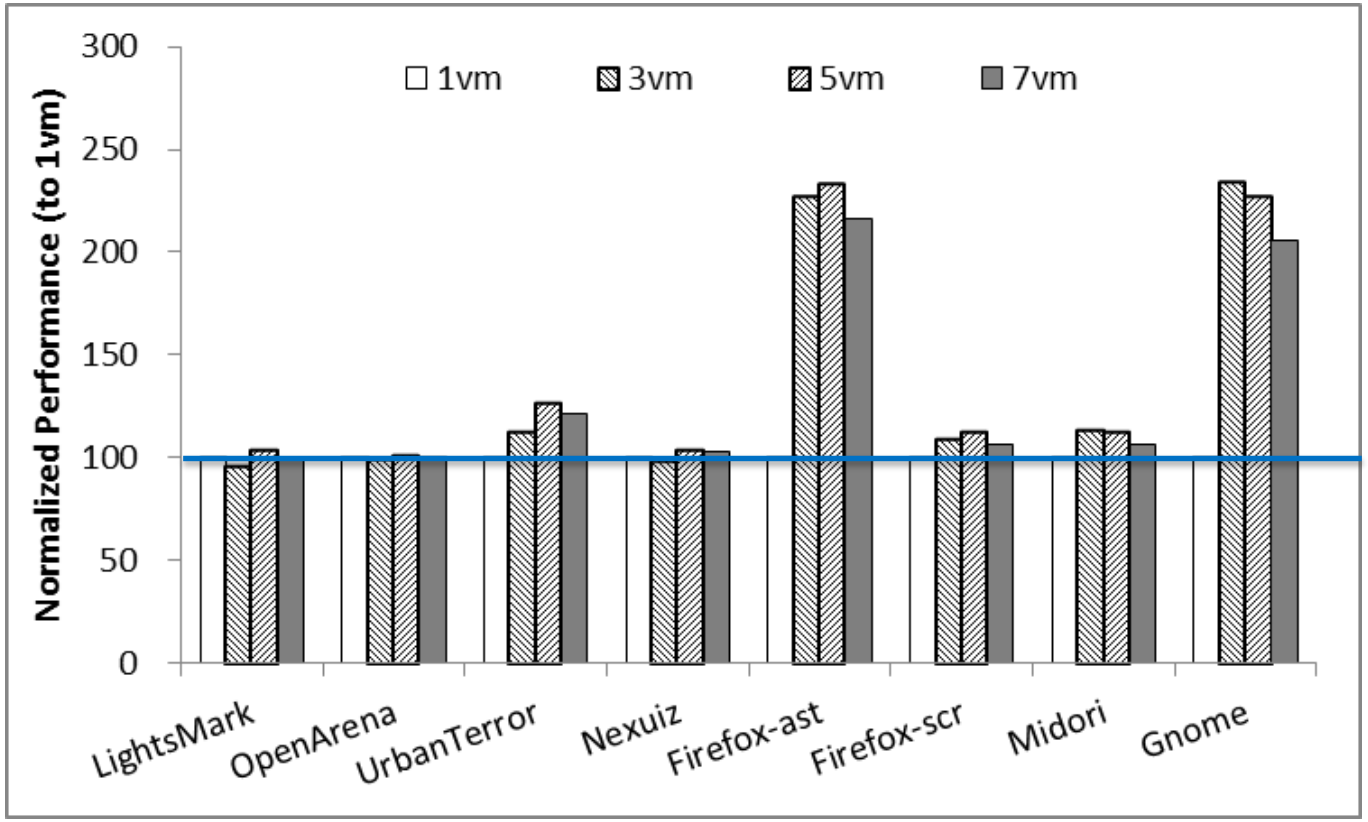


Average 90% of native media transcoding performance



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

Scalability



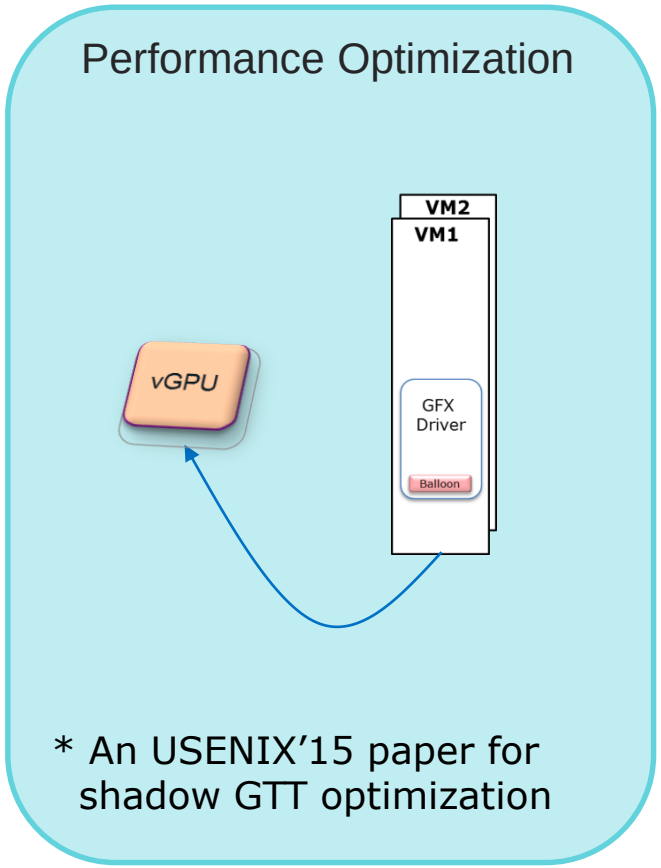
Project Status

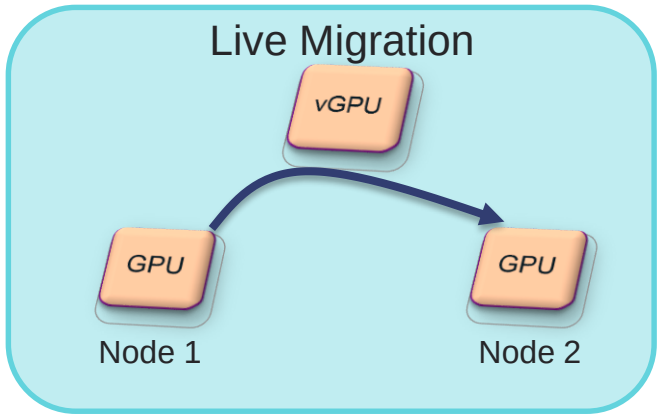
- XenGT
  - HSW stable release in 2014-Q2
  - BDW stable release scheduled in 2015-Q4
  - SKL preliminary release scheduled in 2015-Q4
  - An USENIX'14 paper based on XenGT

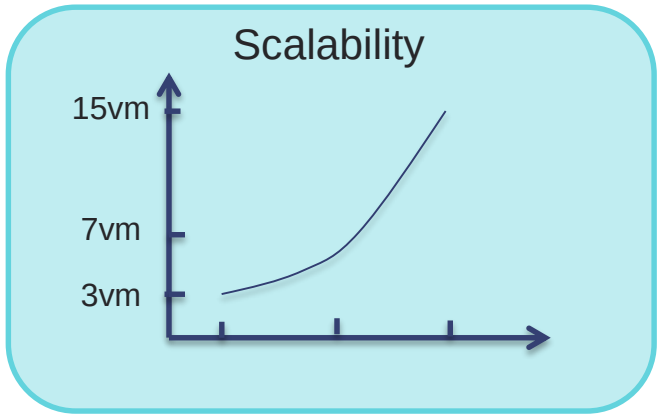
- KVMGT
  - Feature on-par with XenGT, in 2015-Q3
  - 1st stable release on BDW scheduled in 2016-Q1

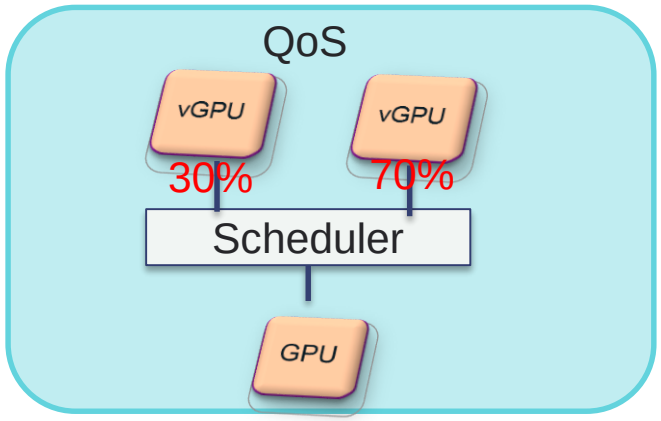
Future Work

- Lots of funny things can be done thru mediated pass-through









High Availability

GPU utilization monitoring

Multi-GPU combination

Mediated pass-through for single VM

Mediated pass-through for other devices

...

Visit US!

- Project home page
  - https://01.org/igvt-g

- Mailing list
  - https://lists.01.org/mailman/listinfo/igvt-g

- Contact us for research/production cooperation!
  - kevin.tian@intel.com, software architect
  - hongbo.wang@intel.com, project manager