Intel Graphics Virtualization Technology Update

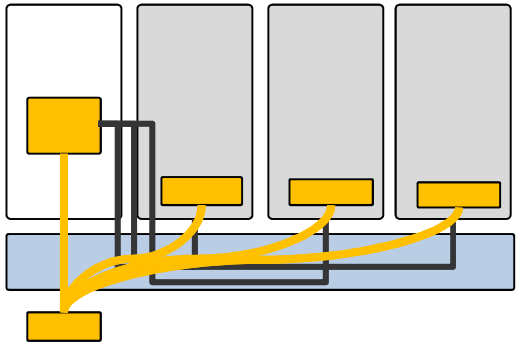
Zhi Wang, zhi.a.wang@intel.com

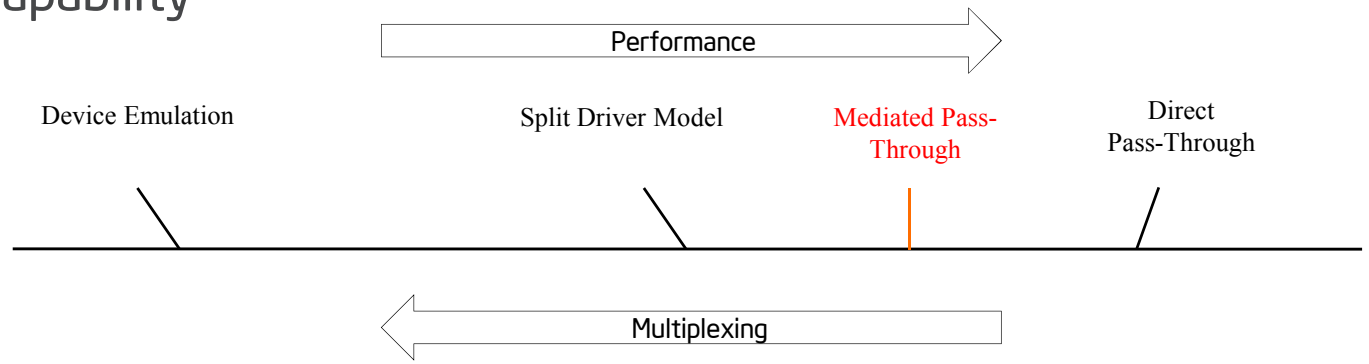
Agenda

- The History
- Intel Graphics Virtualization Technology Update
- New Usage Scenarios
- Upstream Status
- Summary

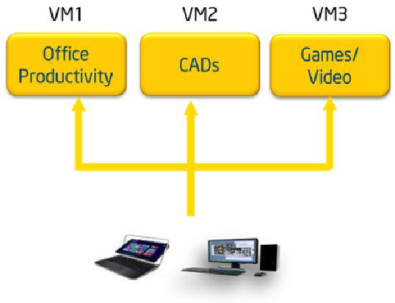
Intel GPU Virtualization Technology: GVT - g

- 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





The History



XenClient

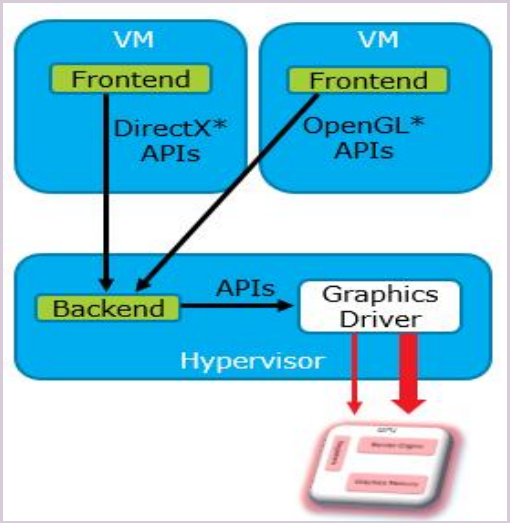
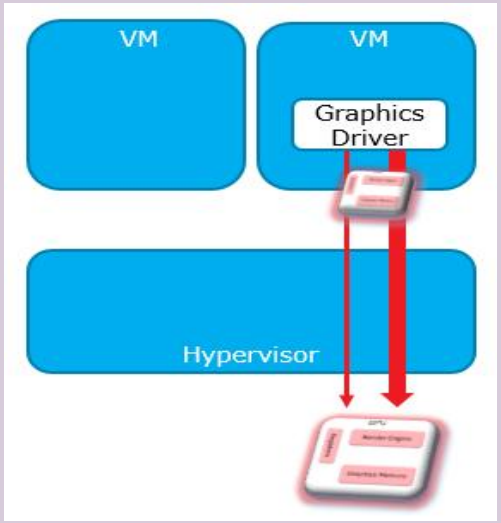
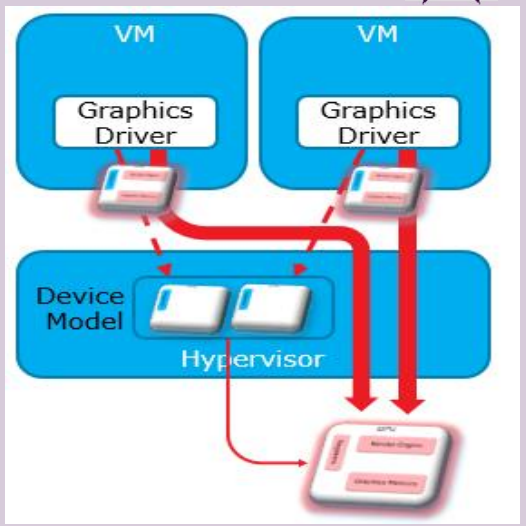
- An open source project based on Xen.
- First published in 2013, codename XenGT.
- First implemented on Haswell - Intel 4th Generation Processor

  - 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

API Forwarding (Intel(R) GVT-s)	Direct Pass-through (Intel(R) GVT-d)	Full GPU Virtualization (Intel(R) GVT-g)
		
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 offical linux kernel 4.1 on Feb 2015

Supported Guest OSes
32/64bit Windows 7
32/64bit Windows 8
32/64bit Ubuntu

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

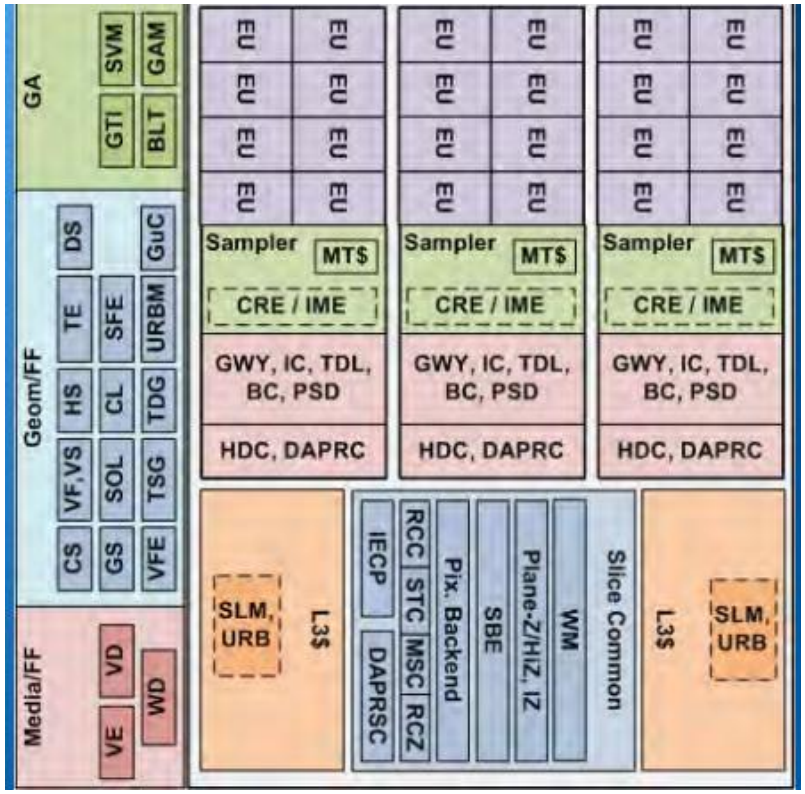
Linux Foundation Collaboration Summit 2013

- 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

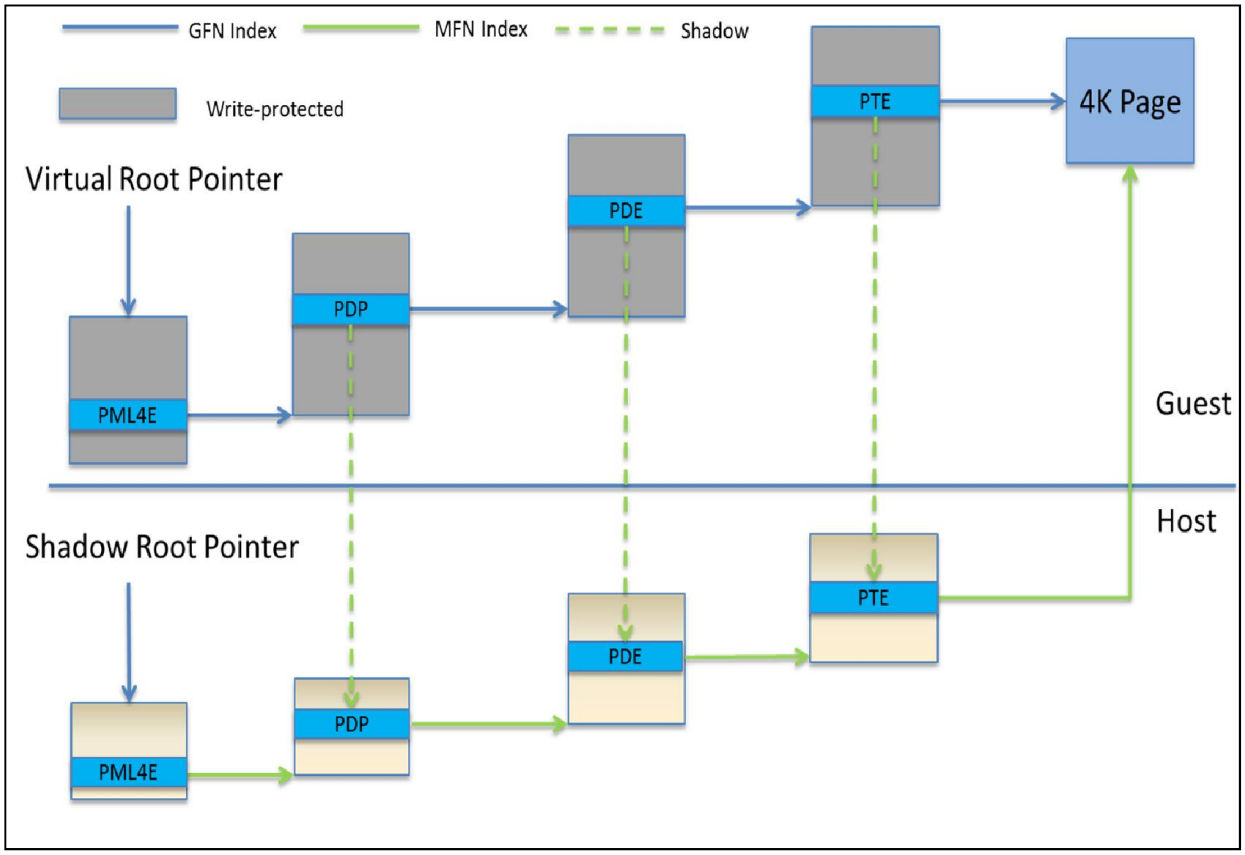
Broadwell vGPU

- More aperture resource allows better scalability (up to 7VMs)
- 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

Per-VM Display

Display Cluster

Infotainment

TCB size

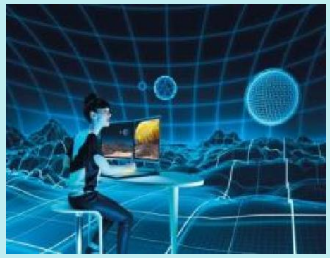
Consolidated SoC

* BYT POC completed

New Usage Scenario: Visual Cloud

Video Delivery

Store/Stream, Transcode



Cloud Graphics

Gaming, Remote Apps, Rendering

Visual Understanding

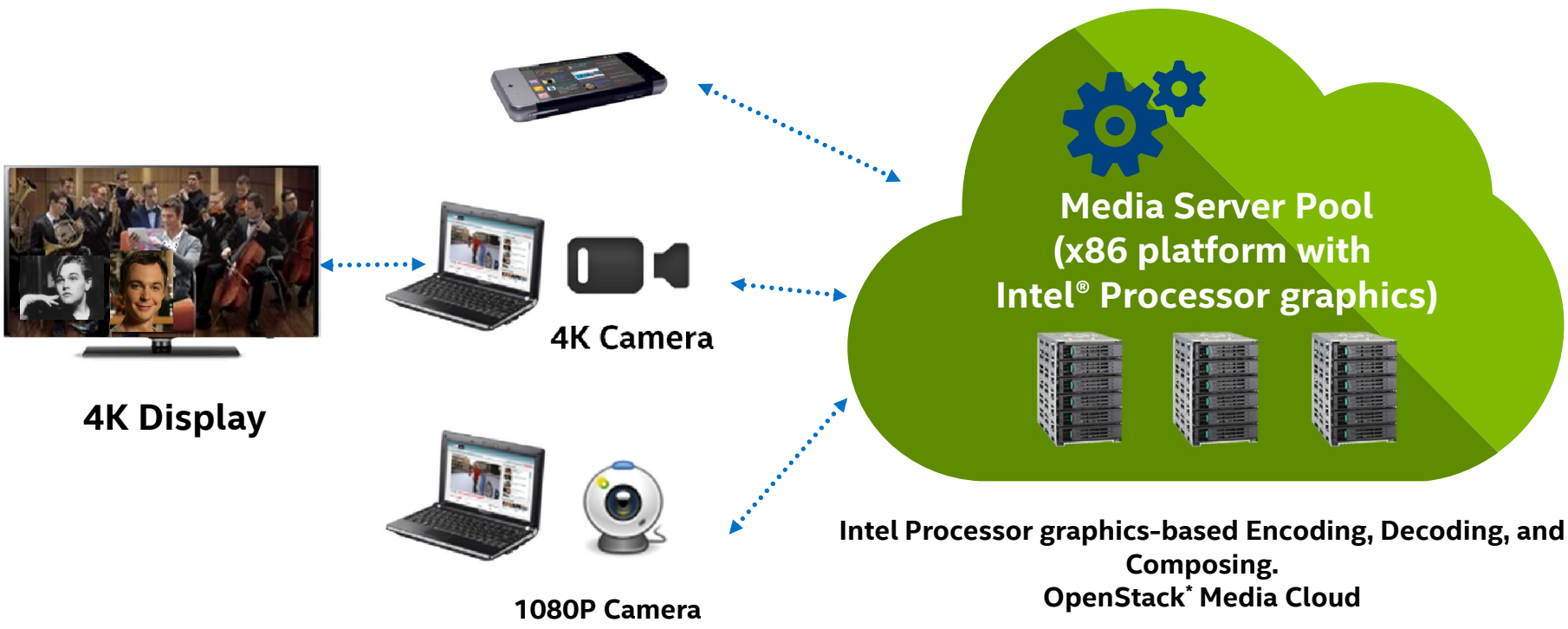
Search, Surveillance

GPU-as-a-service

Visual Cloud

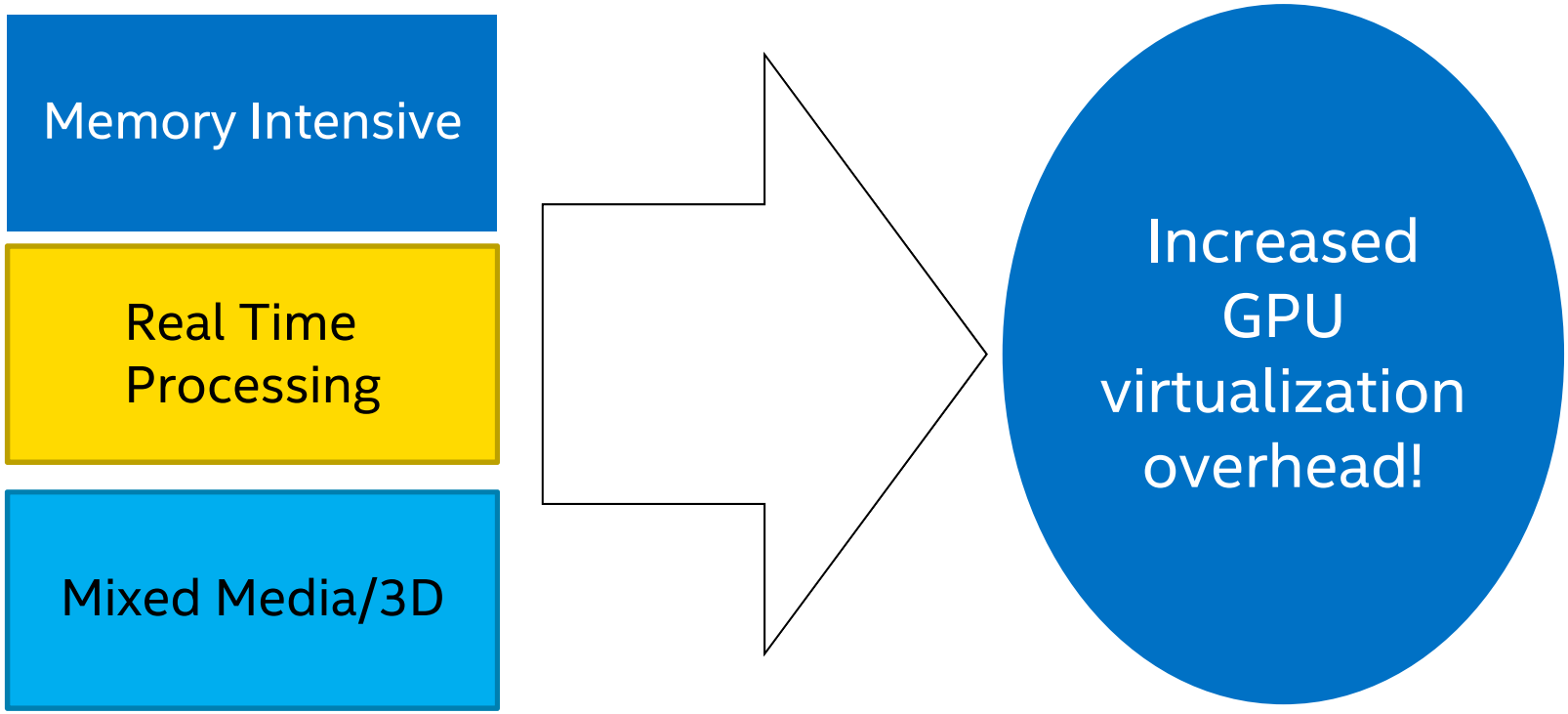
GVT-g Case Study: Media Cloud

A Virtualized Media Server with GVT-g

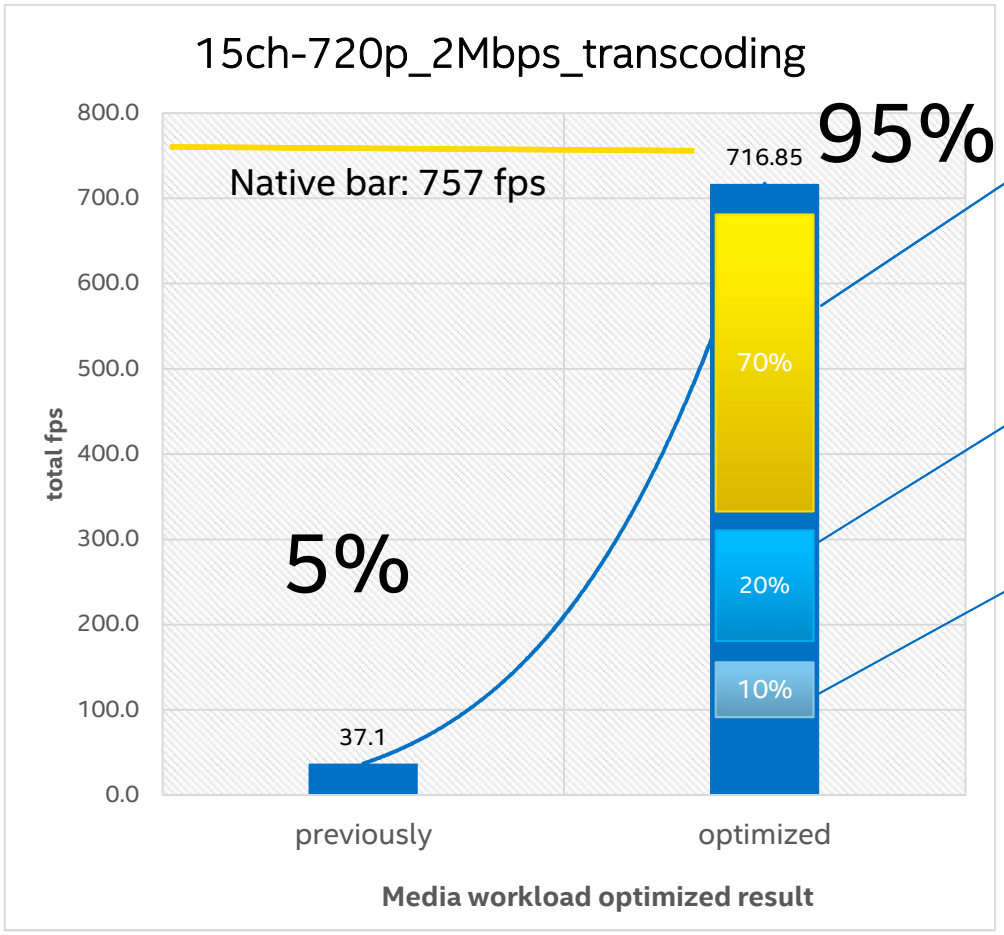


4K Video Conferencing on Media Plane NFV Demo @ MWC'15

Performance challenges



Optimizations



Smart shadow GPU page table

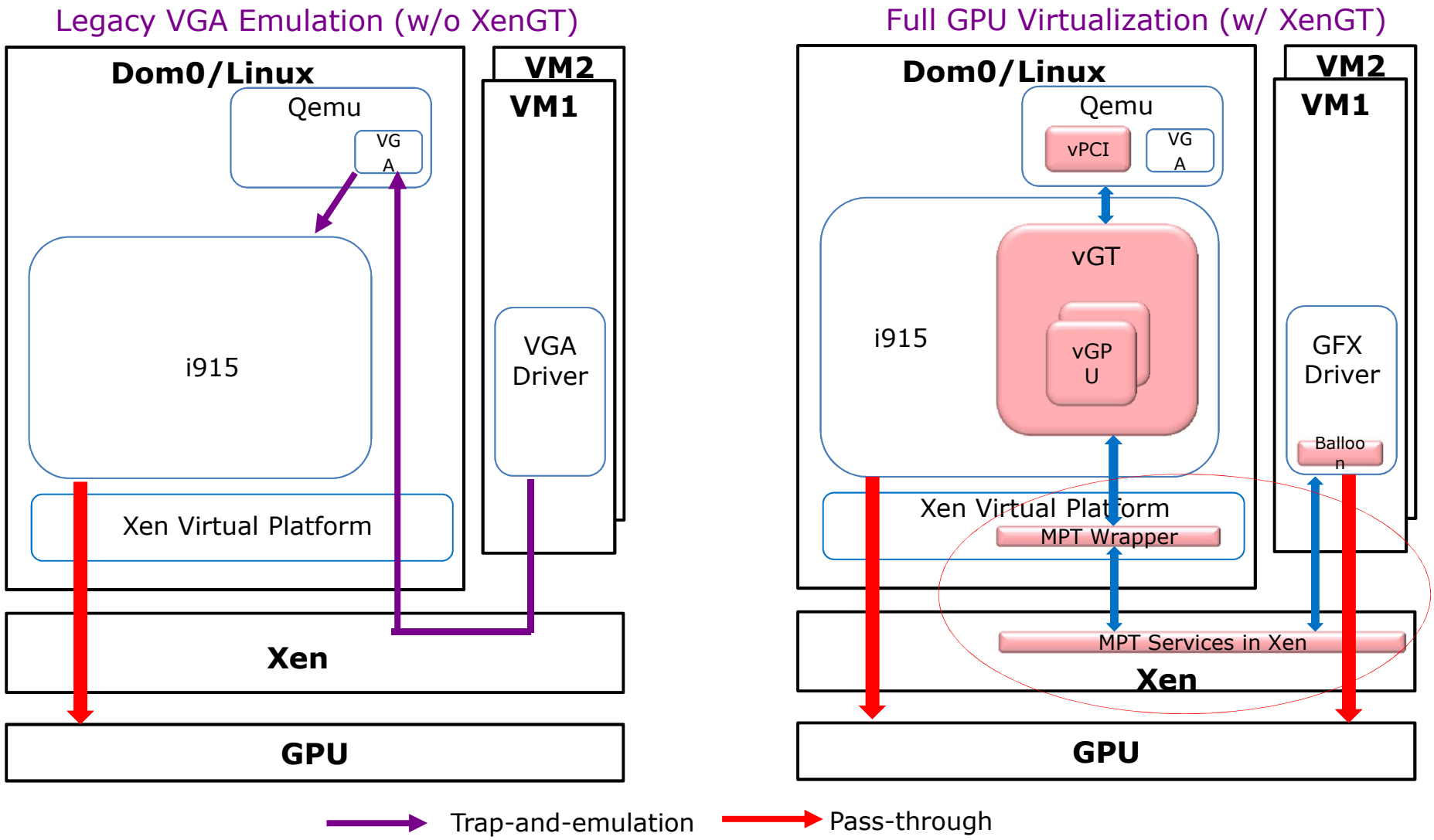
Enable cross-engine synchronization

Increase system memory

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

GVT-g Upstream

Architecture



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


TURE INTEL LINUX WIRELESS
OP
CS YOCTO CONNMAN XEN GUPNP KVM POKY
SYNCEVOLUTION SIMPLE FIRMWARE INTERFACE (SFI) OFONO LINUX KERNEL
ENTERPRISE SECURITY INFRASTRUCTURE
INTEL OPEN SOURCE
TECHNOLOGY CENTER