A Full GPU Virtualization Solution with Mediated Pass-Through

Kun Tian, Yaozu Dong, David Cowperthwaite
kevin.tian@intel.com,
eddie.dong@intel.com,
david.j.cowperthwaite@intel.com

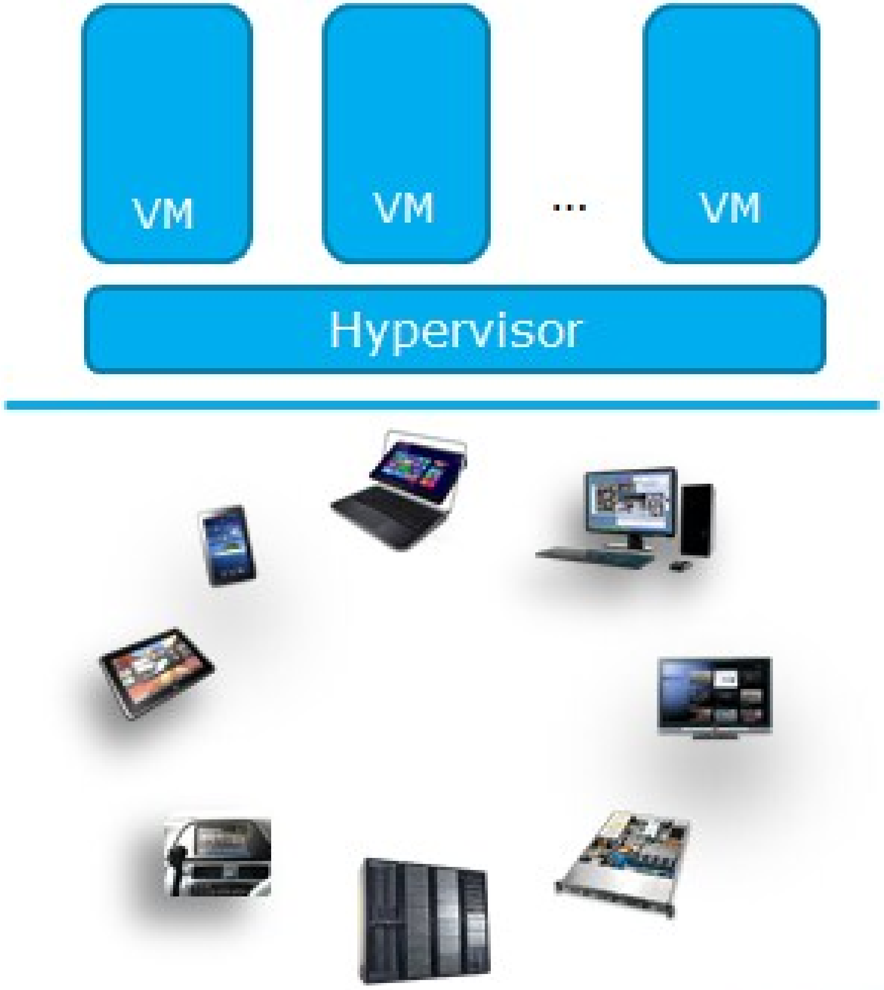
GPU Usages

3D Graphics

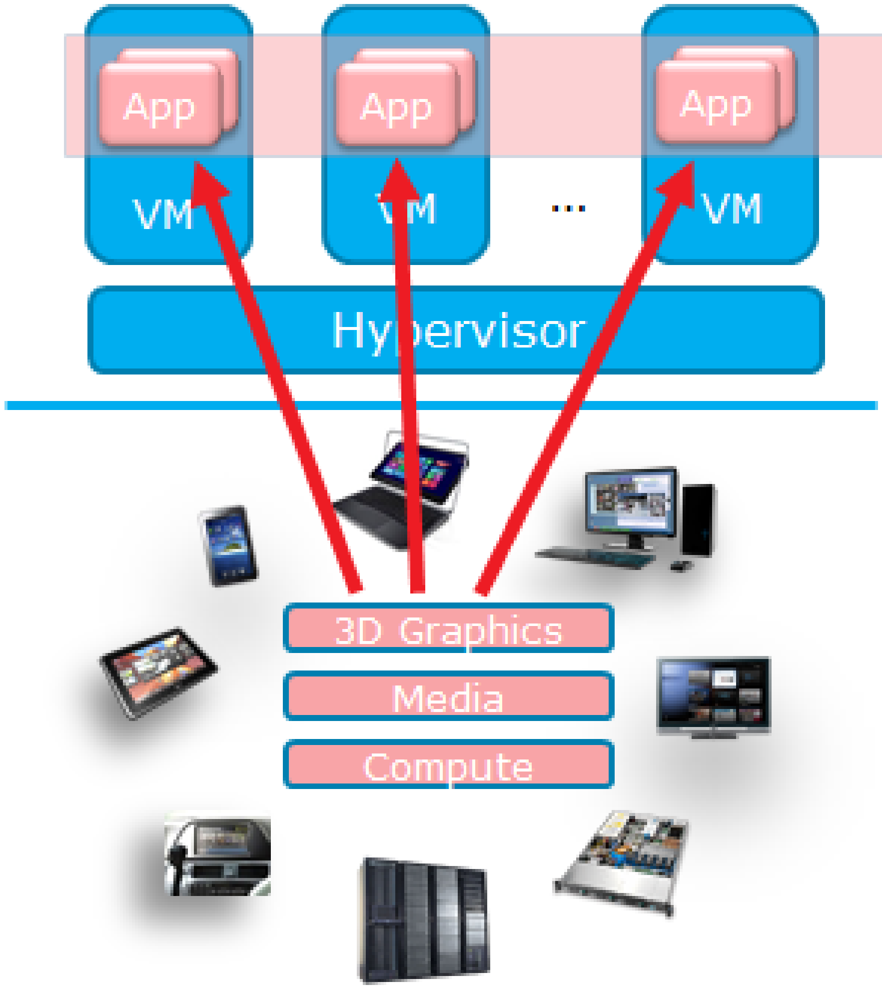
Media

Compute

Virtualization Usages



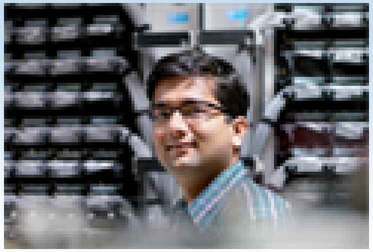
Motivations for GPU Virtualization



GPU Accelerated Tasks

Computer Aided Design
Video Playback
Media Transcoding
Web Experience
Office Productivity
User Interface
Games
Weather broadcast
…

GPU Virtualization Requirements



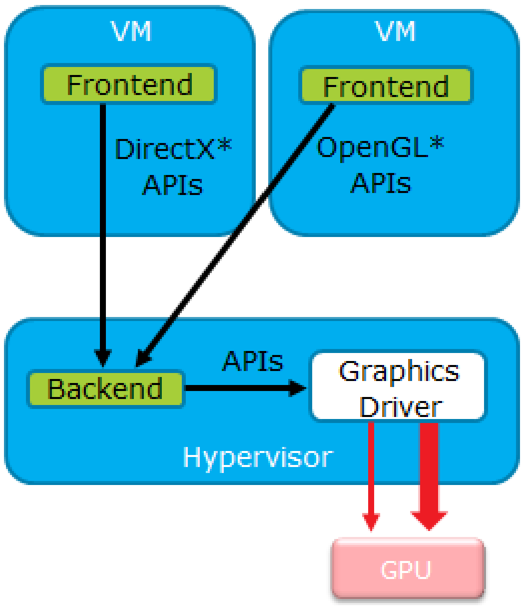
Performance → Direct GPU acceleration

Feature → Consistent visual experience

Sharing → Multiple Virtual Machines

GPU Virtualization Approaches

API Forwarding

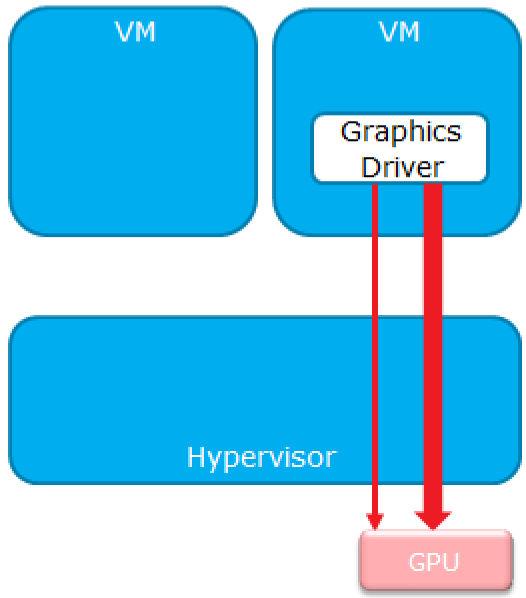


Performance

Feature

Sharing

Direct Pass-Through



Performance

Feature

Sharing

Performance

?

Feature

Sharing

gVirt

- Full GPU virtualization

  → Full-featured vGPU

  → Run native graphics driver in VM

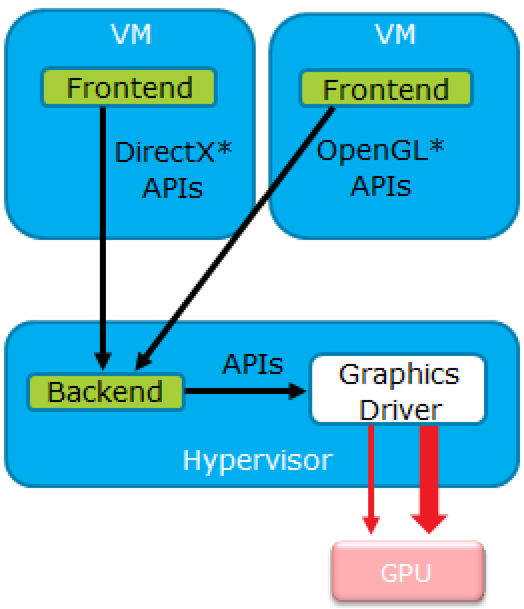
- Mediated Pass-through
  - Pass-through performance critical operations
  - Trap-and-emulate privileged operations

  → Up to 95% native performance

  → Scale up to 7 VMs

GPU Virtualization Approaches

API Forwarding

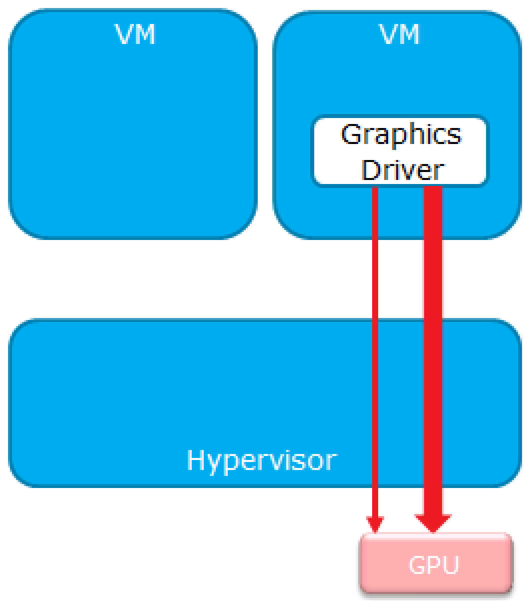


Performance ●●●

Feature ●●●

Sharing ●●●●

Direct Pass-Through

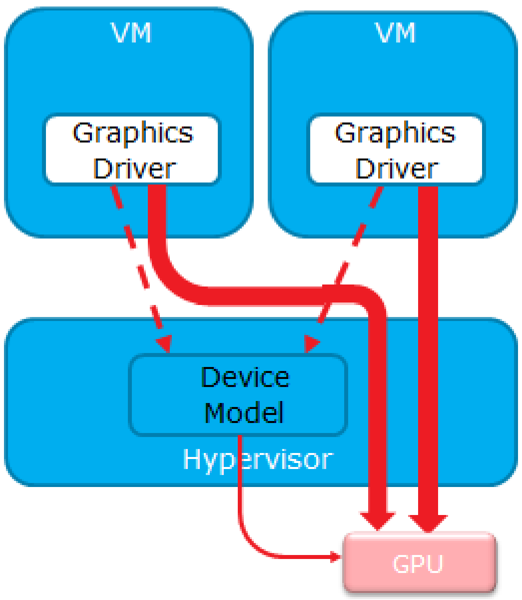


Performance ●●●●●

Feature ●●●●●

Sharing

Full GPU Virtualization



Performance ●●●●

Feature ●●●●●

Sharing ●●●

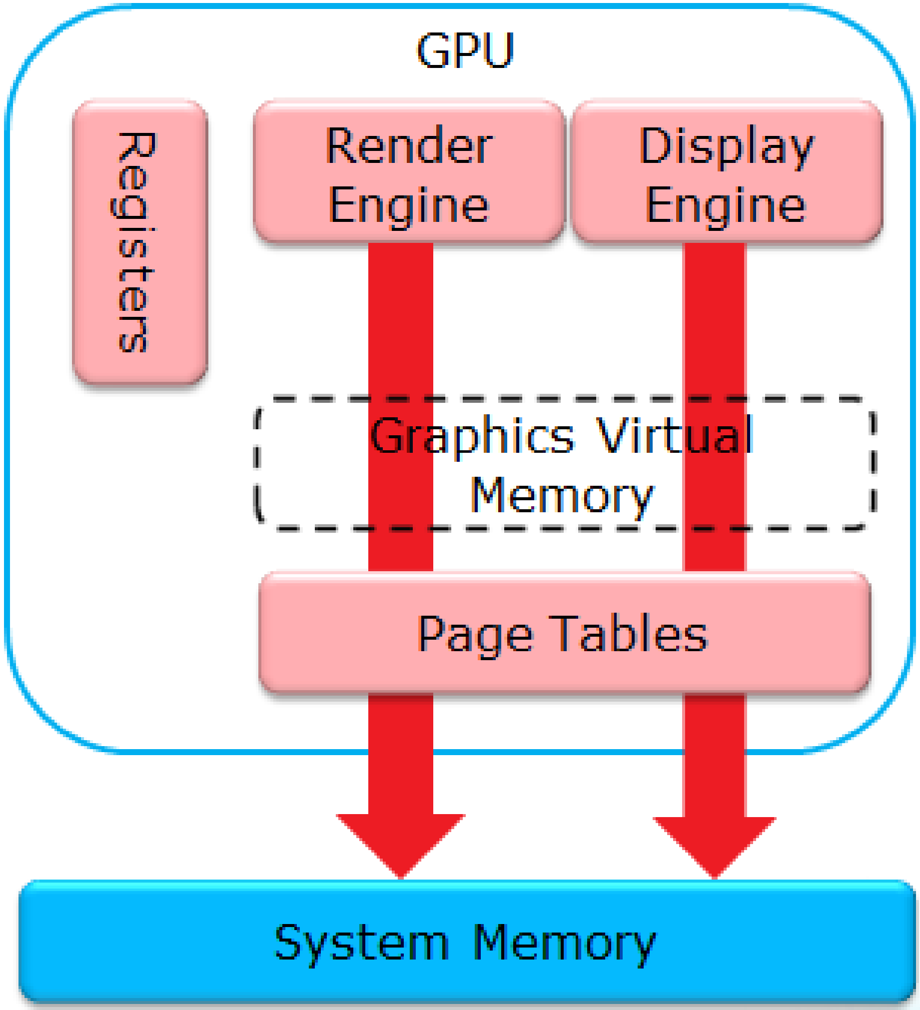
gVirt

- Open source implementation
  - GPL/BSD dual-license
  - Current based on Xen (codename as XenGT)
  - KVM support is coming
- Support Intel® Processor Graphics built into 4th generation Intel® Core™ processors
  - Principles apply to different GPUs
- Trademarked as Intel® GVT-g
  - Intel® Graphics Virtualization Technology for virtual GPU

Challenges

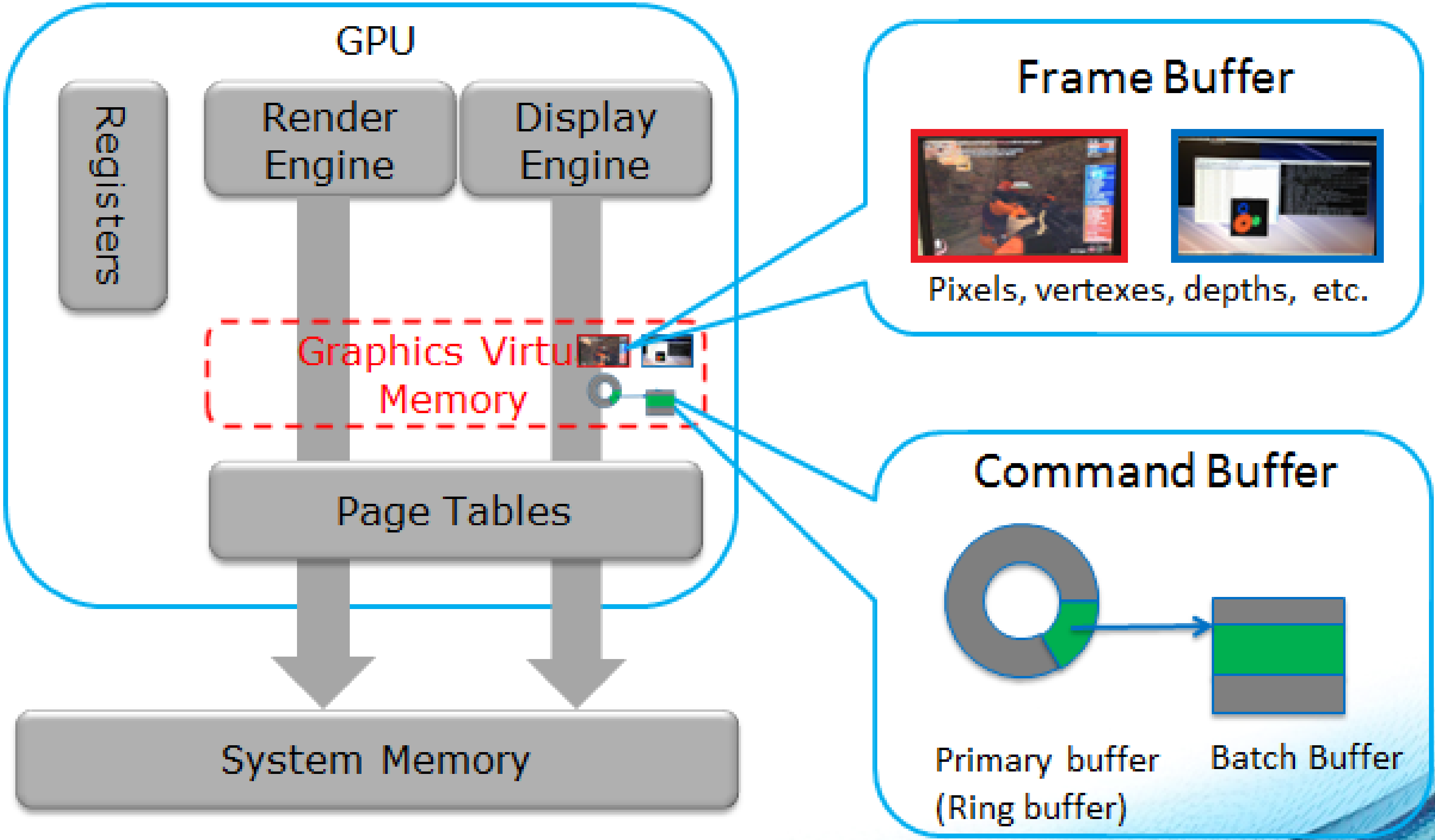
- Complexity in virtualizing a modern GPU
- Efficiency when sharing the GPU
- Secure isolation among the VMs

Architecture of Intel Processor Graphics

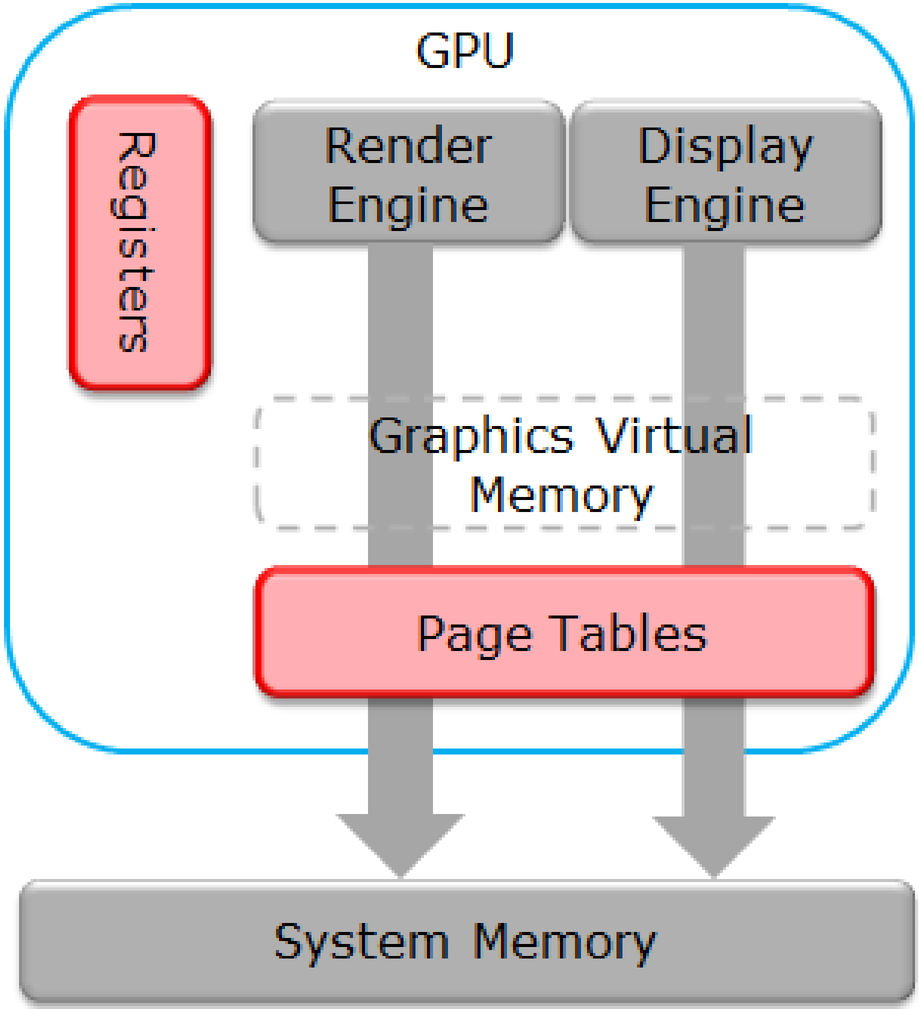


* Above abstraction applies to most modern GPUs,
with major difference in how graphics memory is implemented

Pass Through



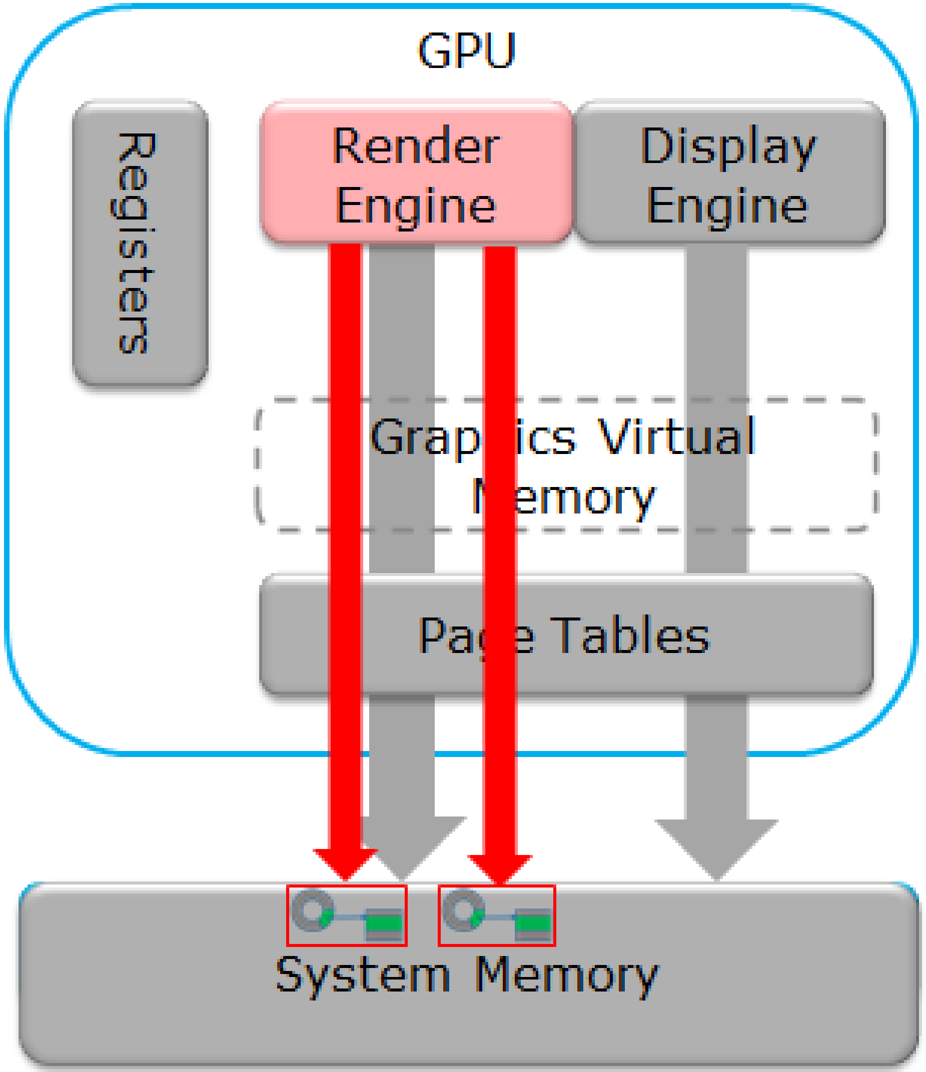
Trap and Emulation



Full-featured vGPU device model

Shadow GPU page table

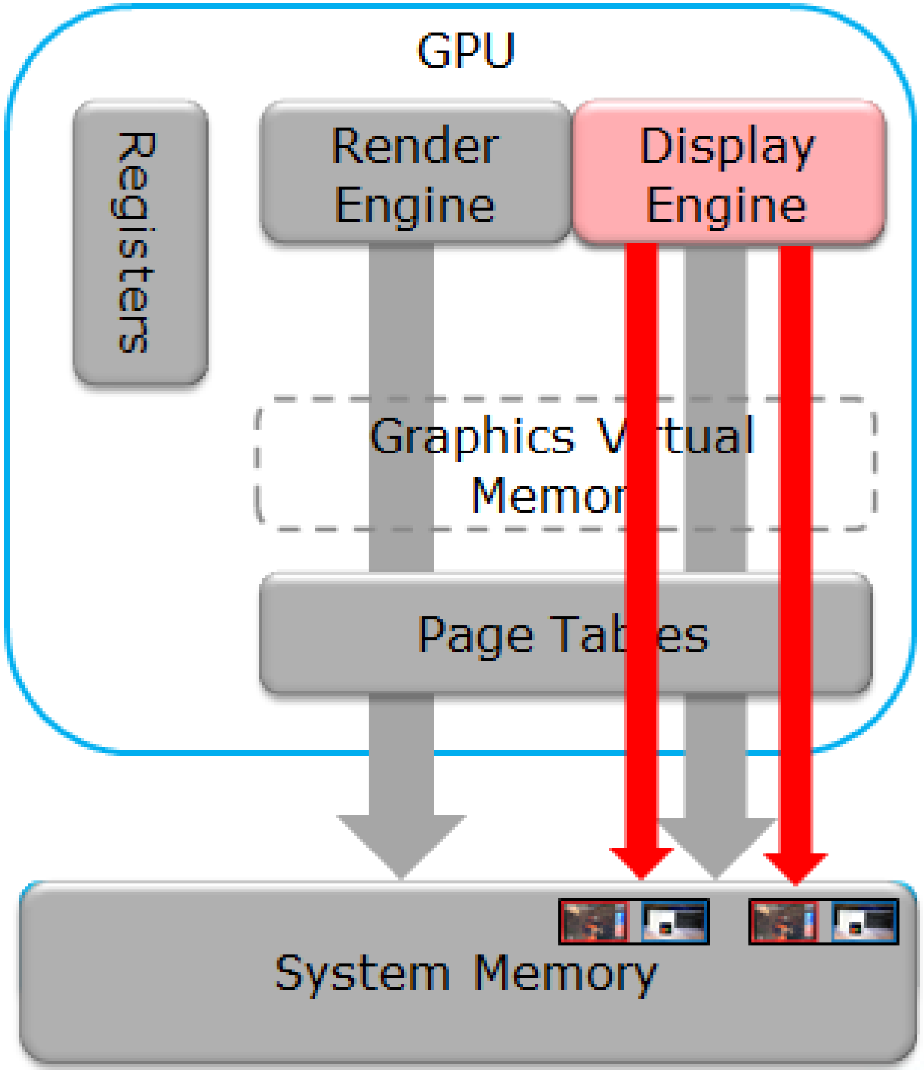
Render Engine Sharing



Direct execution of guest command buffer

Time-based sharing

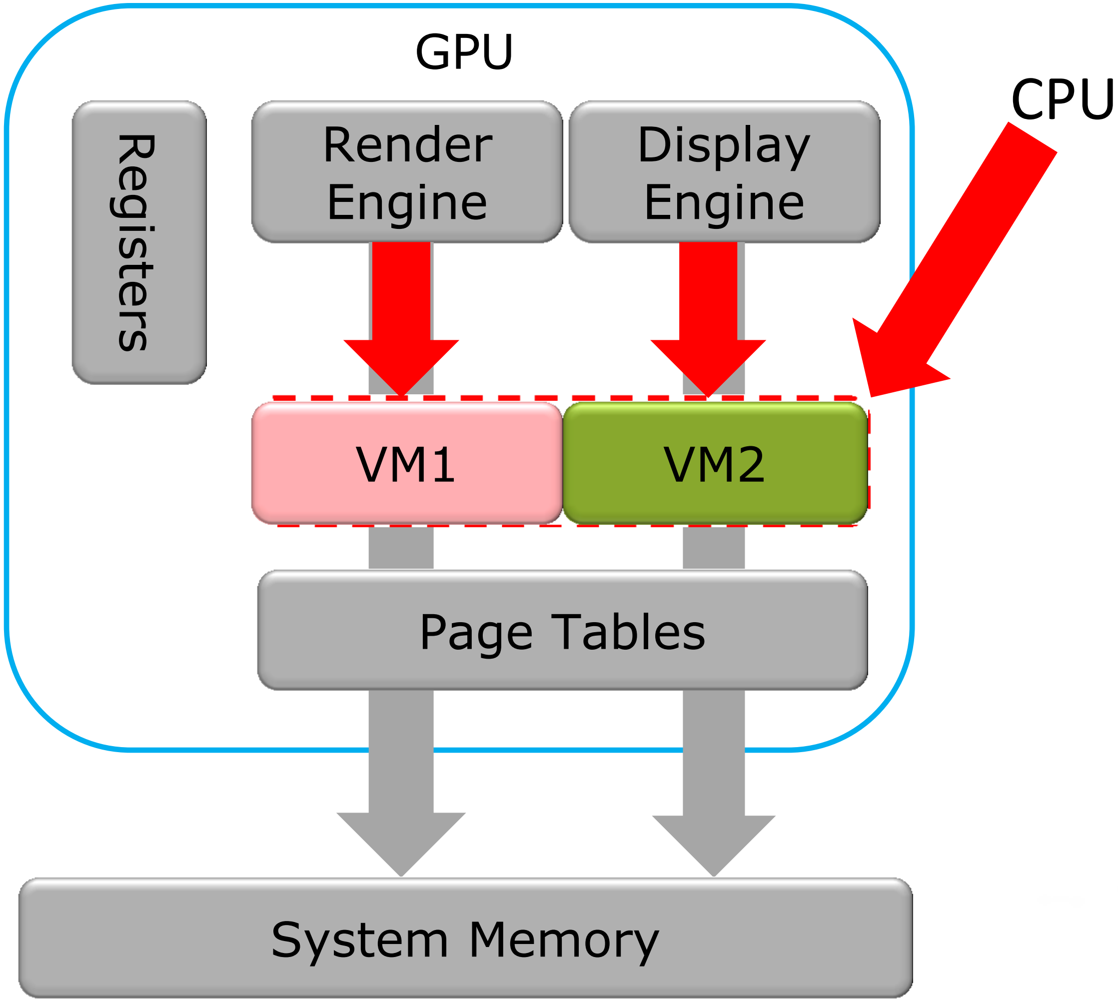
Display Management



"Foreground VM"
vs.
"Background VM"

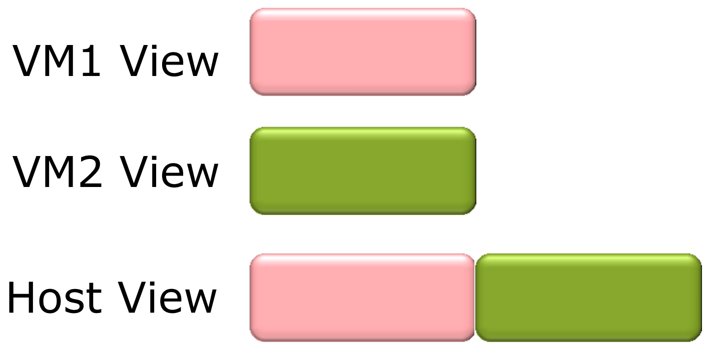
User-initiated switch

Graphics Memory Resource Partitioning



Address Space Ballooning

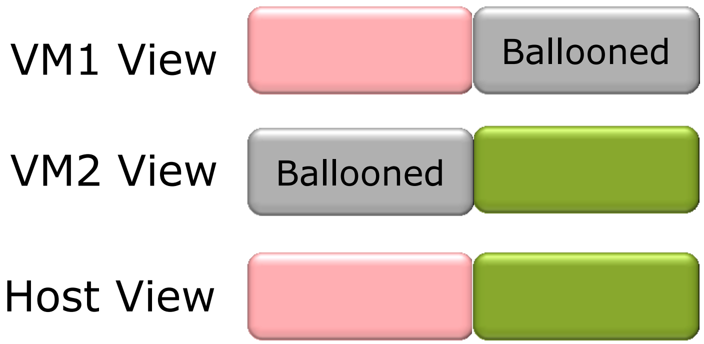
Inconsistent view due to graphics memory partition



(VM's view)
- Reduced size
- Start from address ZERO

Avoid address translation!

Consistent view with address space ballooning



(VM's view)
- Full size
- Start from random address
- Other VM's resource are reserved

Secure Isolation in gVirt

- A VM may map u… …phics memory pages

Mediated
Pass-through

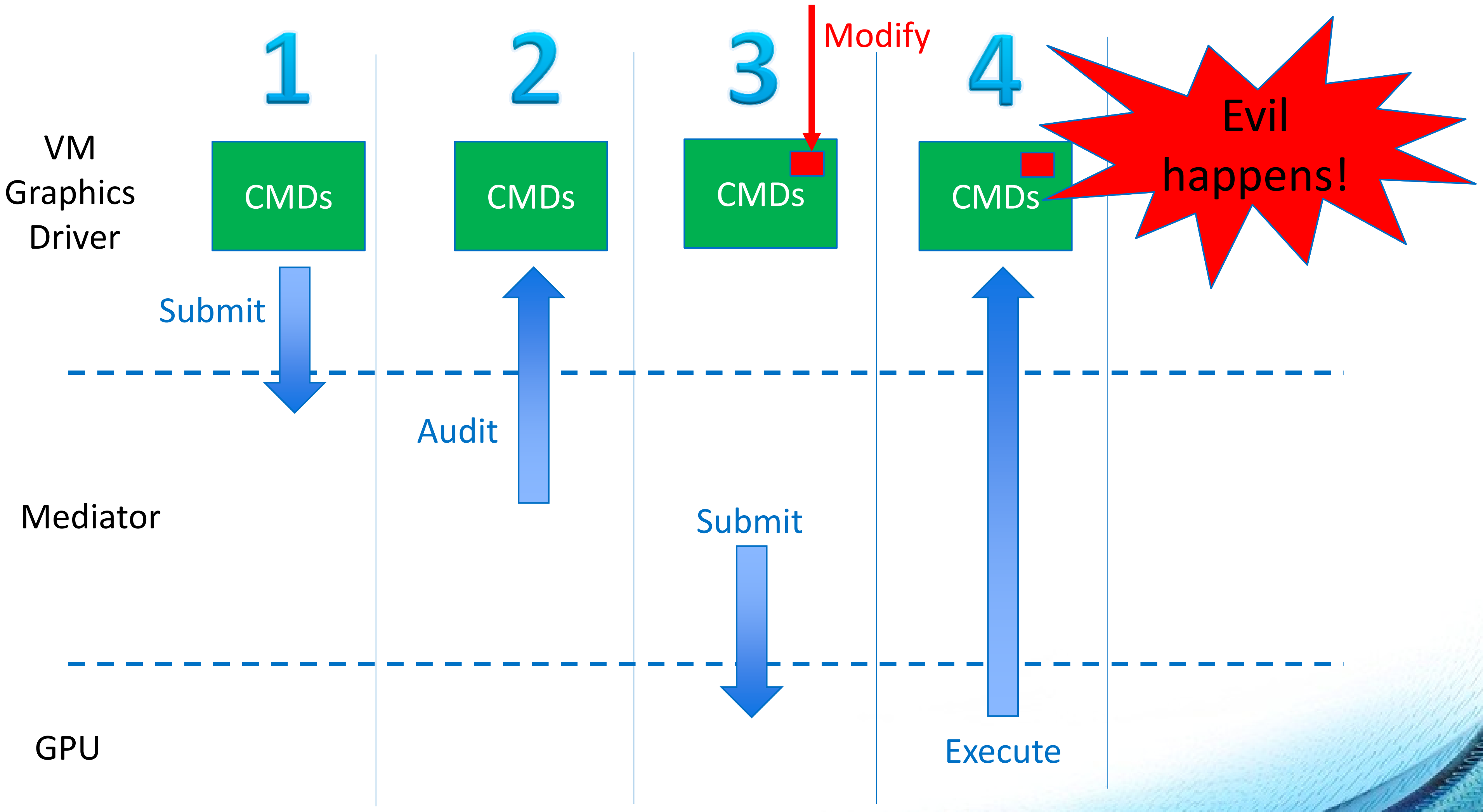
- A VM may progr… …l graphics memory addresses … …nd commands

Audit before
GPU access

- A VM may delibe… …GPU as a denial-of-servic… …at…

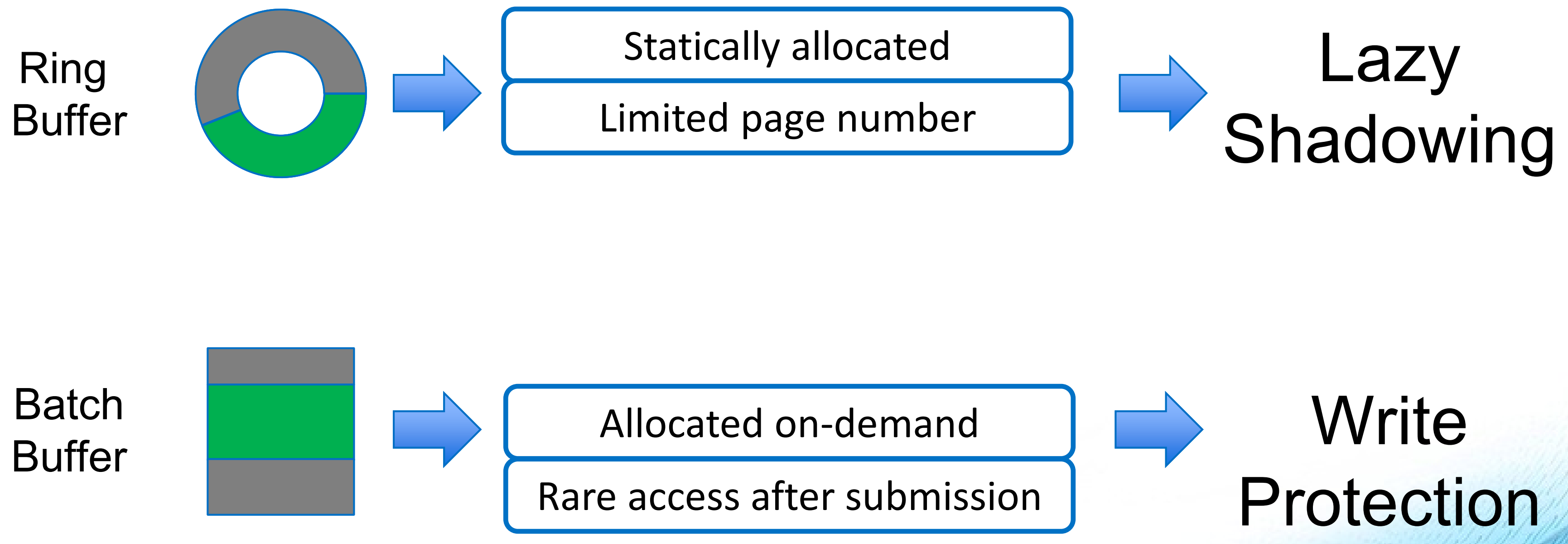
Detect and
recover

Vulnerability from Direct Execution

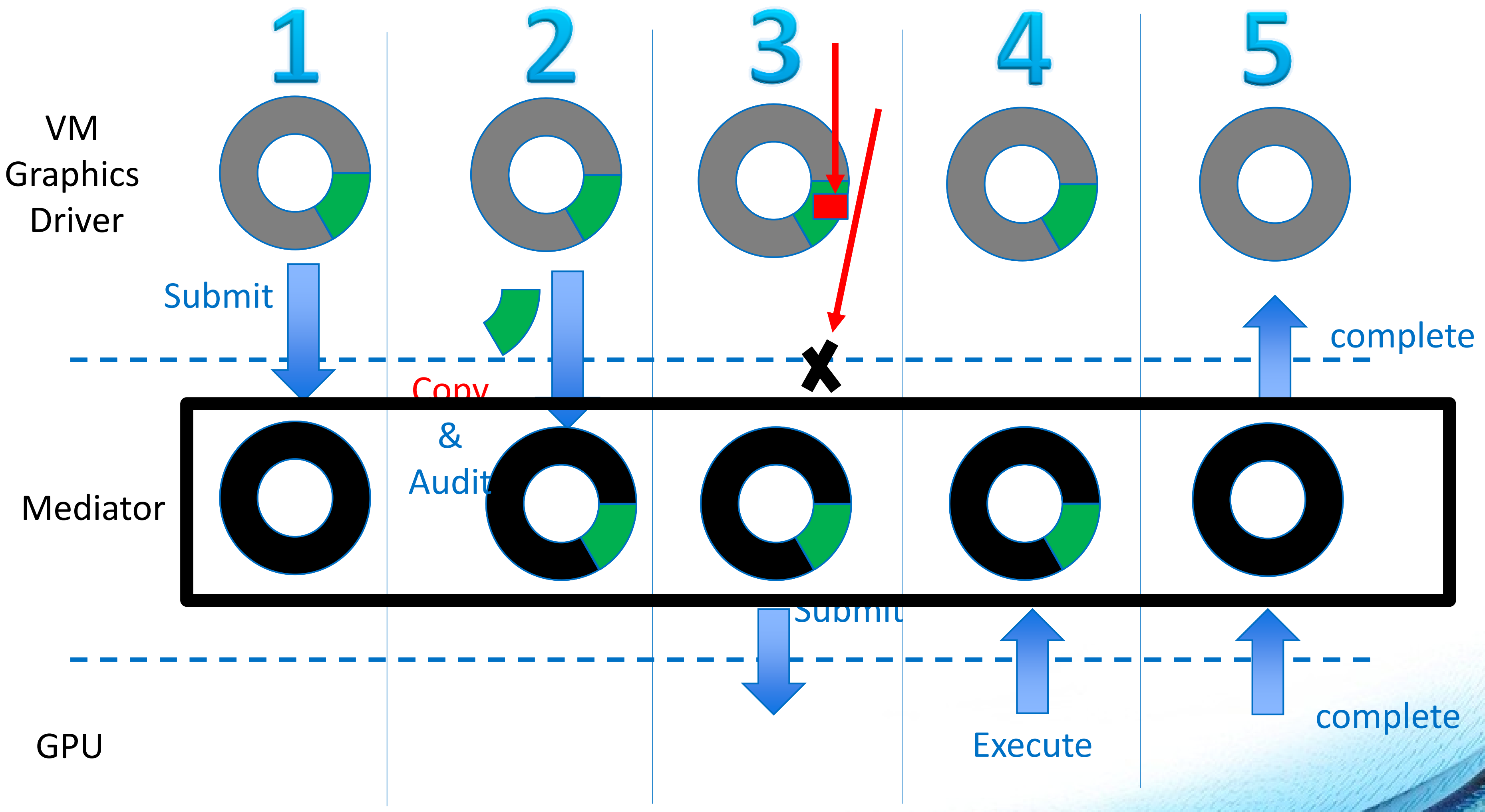


Smart Shadowing

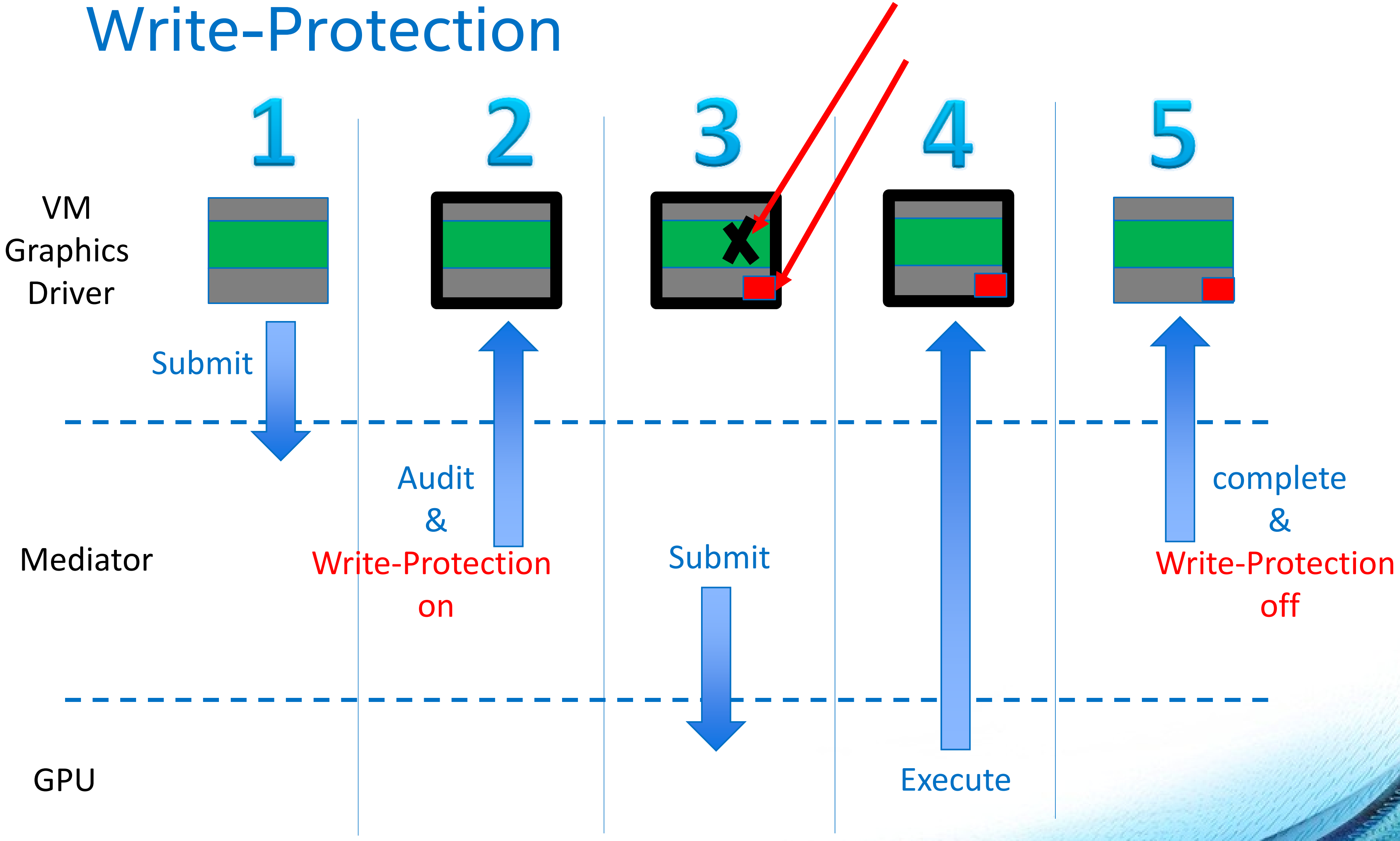
- Utilize specific programming model



Lazy Shadowing



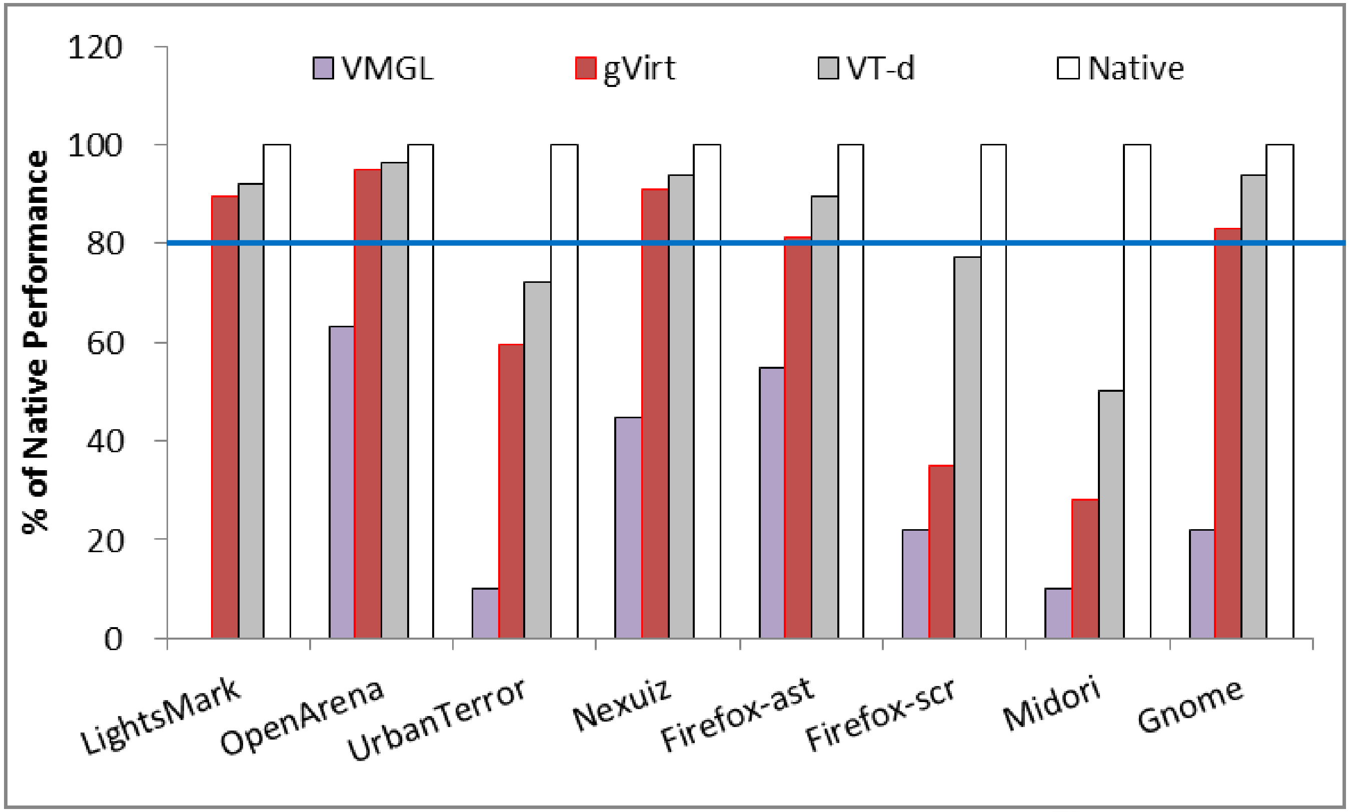
Write-Protection



Configurations

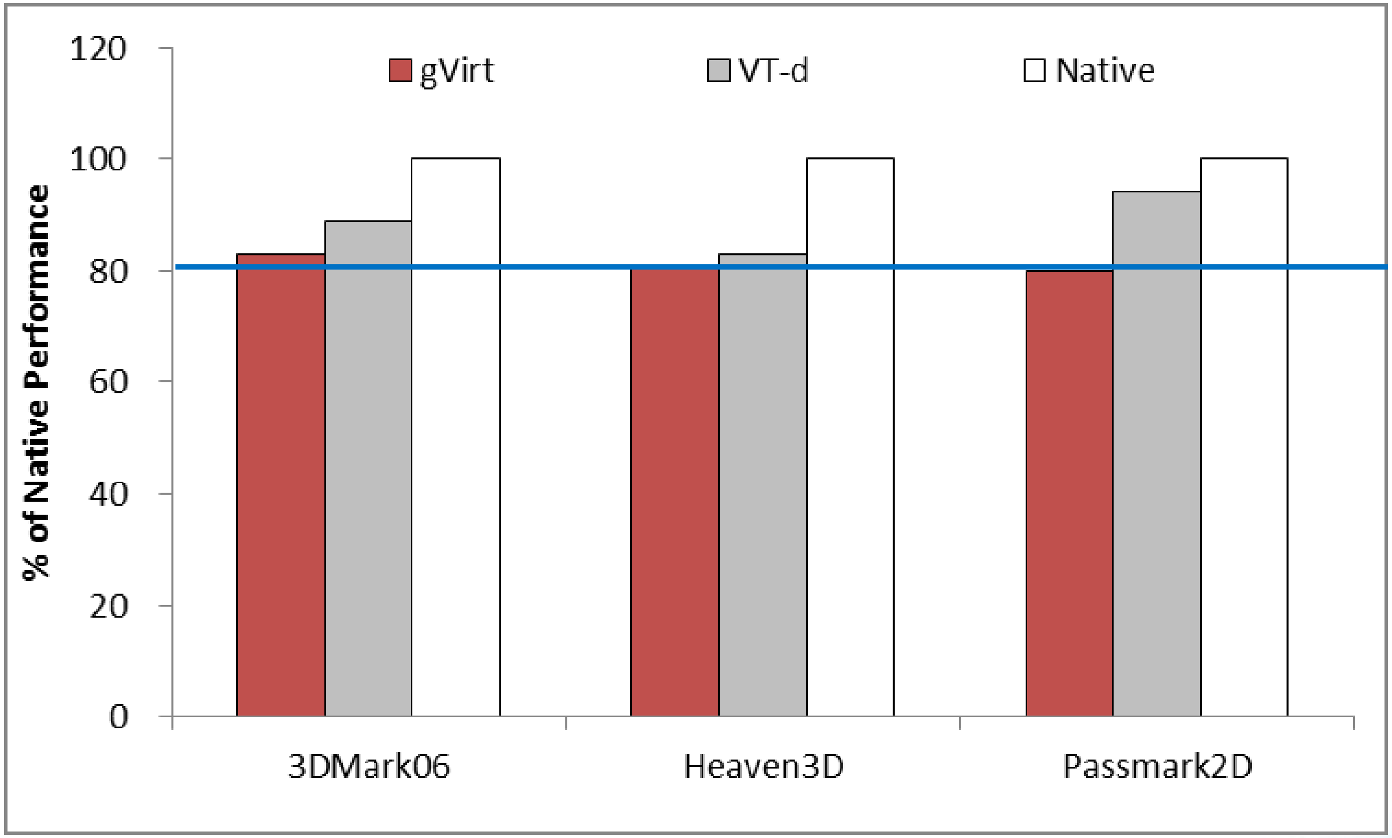
- Hardware with the 4th Intel® Core™ Processor
  - 4 CPU cores (2.4Ghz)
  - 8GB system memory
  - 256GB Intel® 520 series SSD
  - Intel® Processor Graphics
    - A 2GB global graphics memory
    - Multiple 2GB local graphics memory
- Software
  - Dom0/Linux VM: 64bit Ubuntu 12.04 (3.8 kernel)
  - Windows VM: 64bit Win7
  - Xen: 4.3
  - VM configuration
    - 4 VCPUs and 2GB system memory
    - Evenly partitioned global graphics memory (e.g. 512MB per VM in a 3-VM configuration)

Linux VM Performance



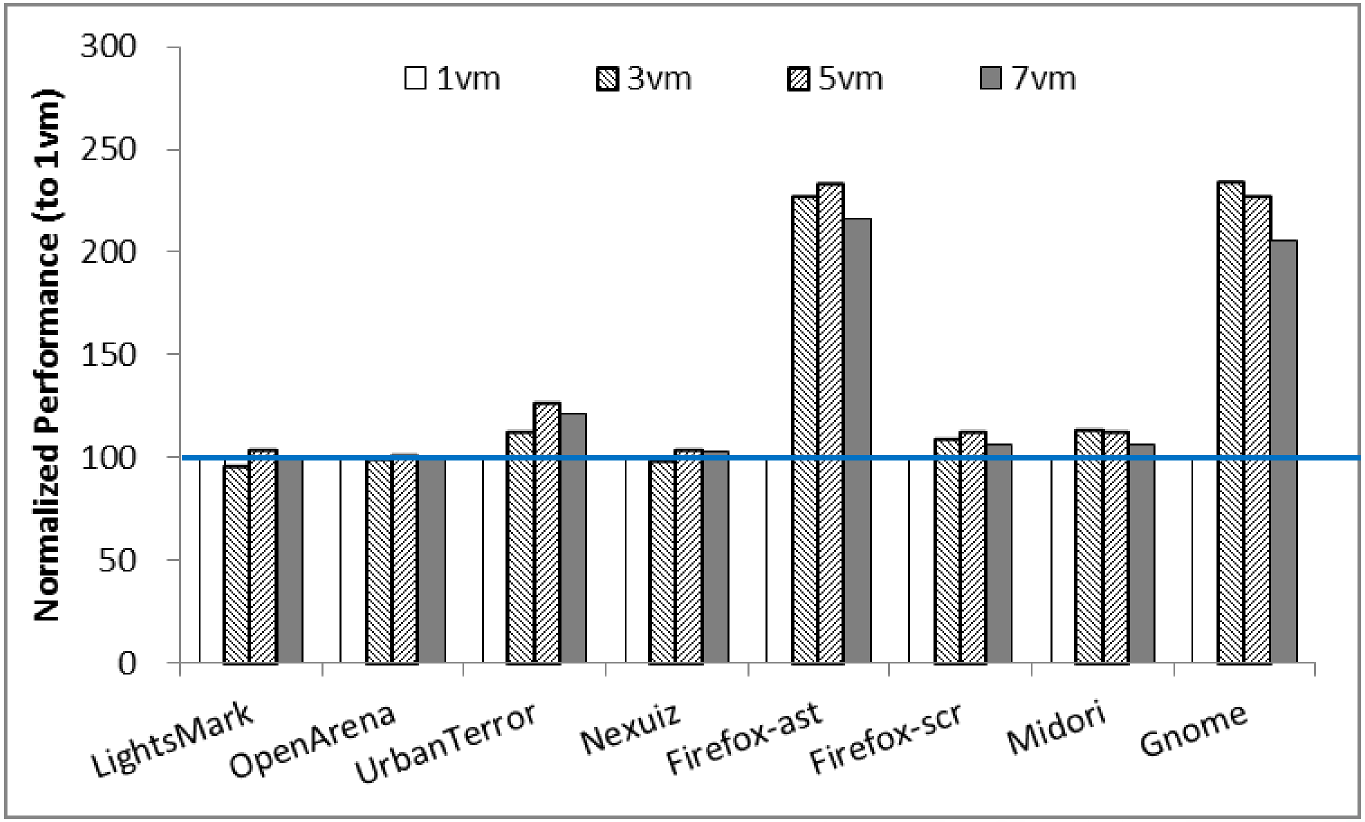
- 3D Benchmark: Phoronix Test Suite
  - LightsMark, OpenArena, UrbanTerror, Nexuiz
- 2D Benchmark: Cairo-perf-trace
  - Firefox-asteroids, firefox-scrolling, midori-zommed, gnome-system-monitor

Windows VM Performance



- 3D Benchmark: 3DMark06, Heaven3D
- 2D Benchmark: Passmark2D

Scalability



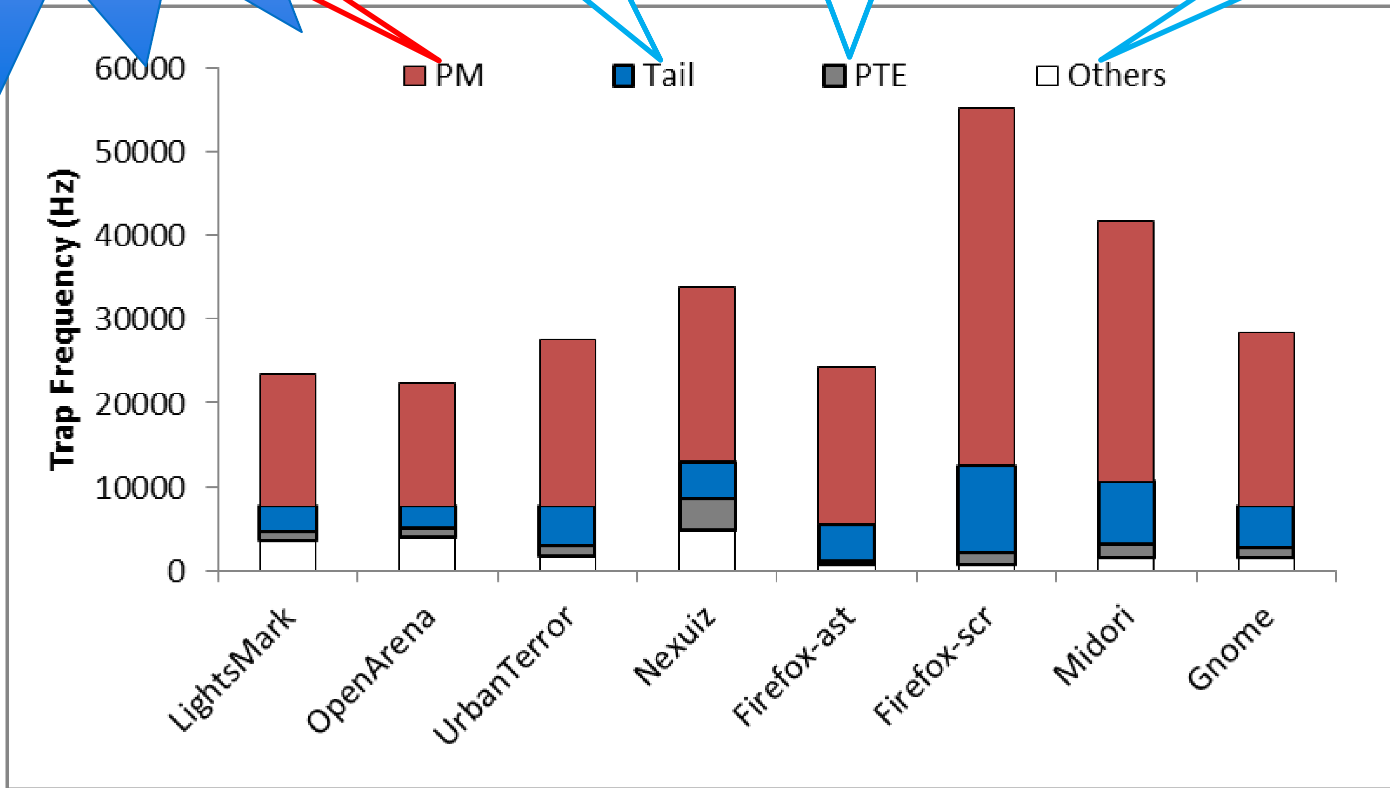
Overhead Analysis

Unnecessary in VM!

Submit commands

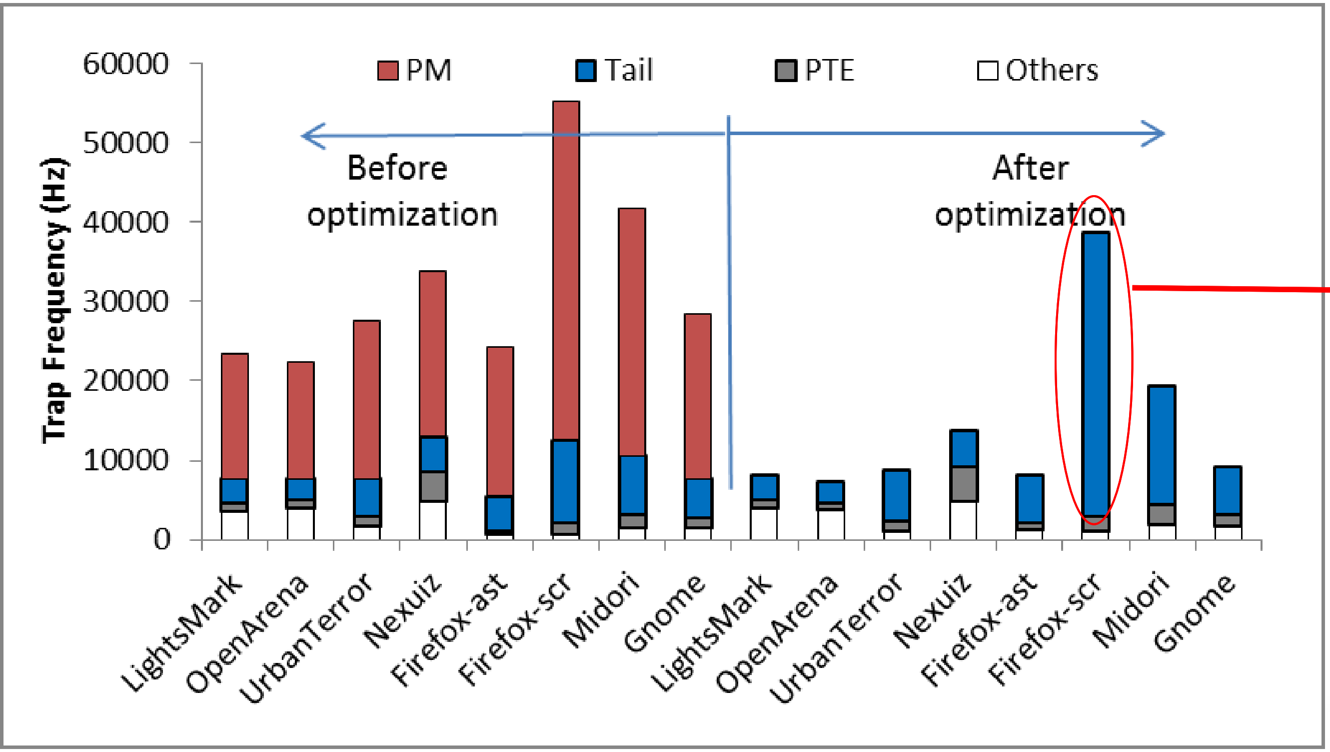
Allocate/free graphics memory

The rest



Software and workloads used in performance tests may have been optimized for performance only on Intel microprocessors. Performance tests, such as SYSmark* and MobileMark*, are measured using specific computer systems, components, software, operations and functions. Any change to any of those factors may cause the results to vary. You should consult other information and performance tests to assist you in fully evaluating your contemplated purchases, including the performance of that product when combined with other products. For more information go to http://www.intel.com/performance.

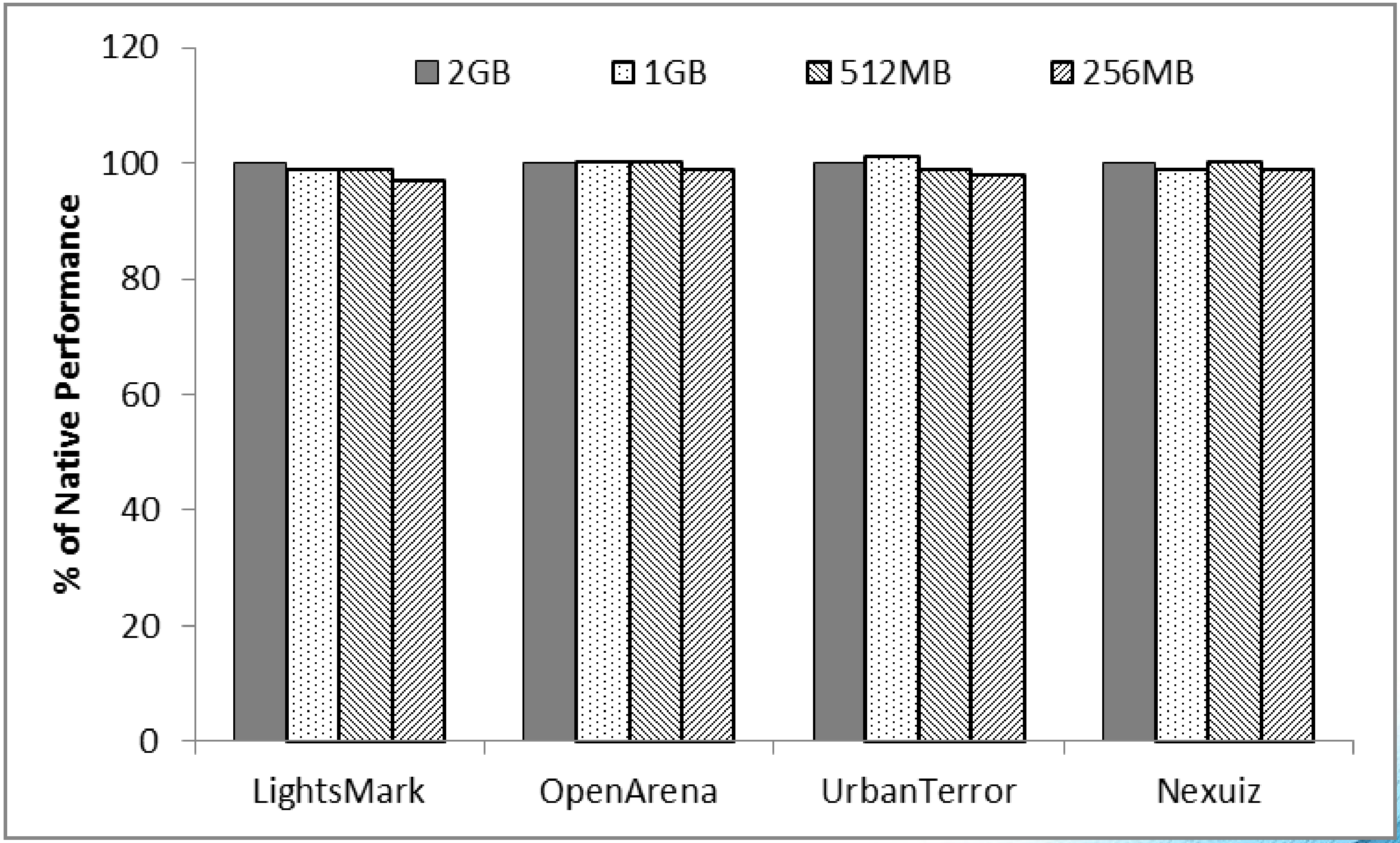
Optimization: Removal of PM Access



Average 60% reduction of trap rate!

2x more commands submitted with 2x higher performance!

Graphics Memory Resource Partitioning



Summary

- Full GPU virtualization + mediated pass-through
- Run native graphics driver in VM
- Good balance for performance, feature and sharing capability
- Publicly available patches
  - https://github.com/01org/XenGT-Preview-xen
  - https://github.com/01org/XenGT-Preview-kernel
  - https://github.com/01org/XenGT-Preview-qemu

Q&A

Legal Disclaimer

Risk Factors

The above statements and any others in this document that refer to plans and expectations for the first quarter, the year and the future are forward-looking statements that involve a number of risks and uncertainties. Words such as “anticipates,” “expects,” “intends,” “plans,” “believes,” “seeks,” “estimates,” “may,” “will,” “should” and their variations identify forward-looking statements. Statements that refer to or are based on projections, uncertain events or assumptions also identify forward-looking statements. Many factors could affect Intel’s actual results, and variances from Intel’s current expectations regarding such factors could cause actual results to differ materially from those expressed in these forward-looking statements. Intel presently considers the following to be the important factors that could cause actual results to differ materially from the company’s expectations. Demand could be different from Intel's expectations due to factors including changes in business and economic conditions; customer acceptance of Intel’s and competitors’ products; supply constraints and other disruptions affecting customers; changes in customer order patterns including order cancellations; and changes in the level of inventory at customers. Uncertainty in global economic and financial conditions poses a risk that consumers and businesses may defer purchases in response to negative financial events, which could negatively affect product demand and other related matters. Intel operates in intensely competitive industries that are characterized by a high percentage of costs that are fixed or difficult to reduce in the short term and product demand that is highly variable and difficult to forecast. Revenue and the gross margin percentage are affected by the timing of Intel product introductions and the demand for and market acceptance of Intel's products; actions taken by Intel's competitors, including product offerings and introductions, marketing programs and pricing pressures and Intel’s response to such actions; and Intel’s ability to respond quickly to technological developments and to incorporate new features into its products. The gross margin percentage could vary significantly from expectations based on capacity utilization; variations in inventory valuation, including variations related to the timing of qualifying products for sale; changes in revenue levels; segment product mix; the timing and execution of the manufacturing ramp and associated costs; start-up costs; excess or obsolete inventory; changes in unit costs; defects or disruptions in the supply of materials or resources; product manufacturing quality/yields; and impairments of long-lived assets, including manufacturing, assembly/test and intangible assets. Intel's results could be affected by adverse economic, social, political and physical/infrastructure conditions in countries where Intel, its customers or its suppliers operate, including military conflict and other security risks, natural disasters, infrastructure disruptions, health concerns and fluctuations in currency exchange rates. Expenses, particularly certain marketing and compensation expenses, as well as restructuring and asset impairment charges, vary depending on the level of demand for Intel's products and the level of revenue and profits. Intel’s results could be affected by the timing of closing of acquisitions and divestitures. Intel's results could be affected by adverse effects associated with product defects and errata (deviations from published specifications), and by litigation or regulatory matters involving intellectual property, stockholder, consumer, antitrust, disclosure and other issues, such as the litigation and regulatory matters described in Intel's SEC reports. An unfavorable ruling could include monetary damages or an injunction prohibiting Intel from manufacturing or selling one or more products, precluding particular business practices, impacting Intel’s ability to design its products, or requiring other remedies such as compulsory licensing of intellectual property. A detailed discussion of these and other factors that could affect Intel’s results is included in Intel’s SEC filings, including the company’s most recent reports on Form 10-Q, Form 10-K and earnings release.

Rev. 1/16/14

intel