Xen GPU Intro.

國網中心-格網技術組
Rock(郭文傑), Rider(涂哲源)
rock@nchc.org.tw, rider@nchc.org.tw

Why Xen ?

- Xen aims to be able to execute multiple operating systems on one physical x86 machine
  - Support popular OS (Linux, Windows XP, NetBSD)
  - Scalable up to around 100 VMs
  - Securely
  - With close-to-native performance
- Support heterogeneous applications using one physical machine (I.E. Win32 and Linux apps)
- Suspend / Resume & Live Migration

Why GPU ?

- Most of the Graphics Processing Units were designed to accelerate graphics tasks like image rendering.
- The power and flexibility of GPUs makes them an attractive platform for general-purpose computation.
- Potential computing power.

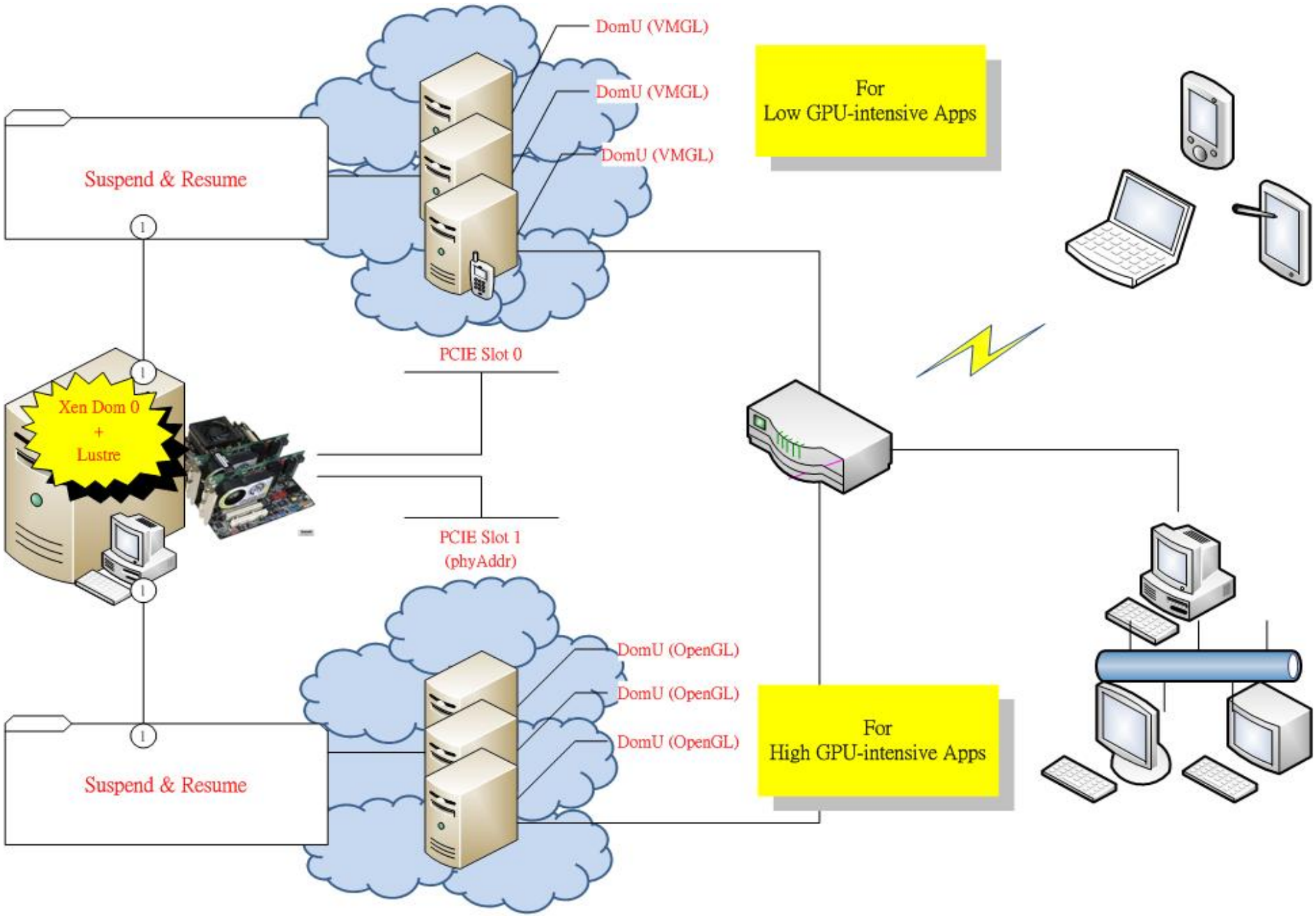
Target

- **GPGPU**
- **Running 3D Applications**
  - Different devices
  - Different purposes
- **Vision is the reality**

Vision

- **Cost Down – Save your money**
- **Green – Energy efficiency**
- **Access Anytime & Anywhere**

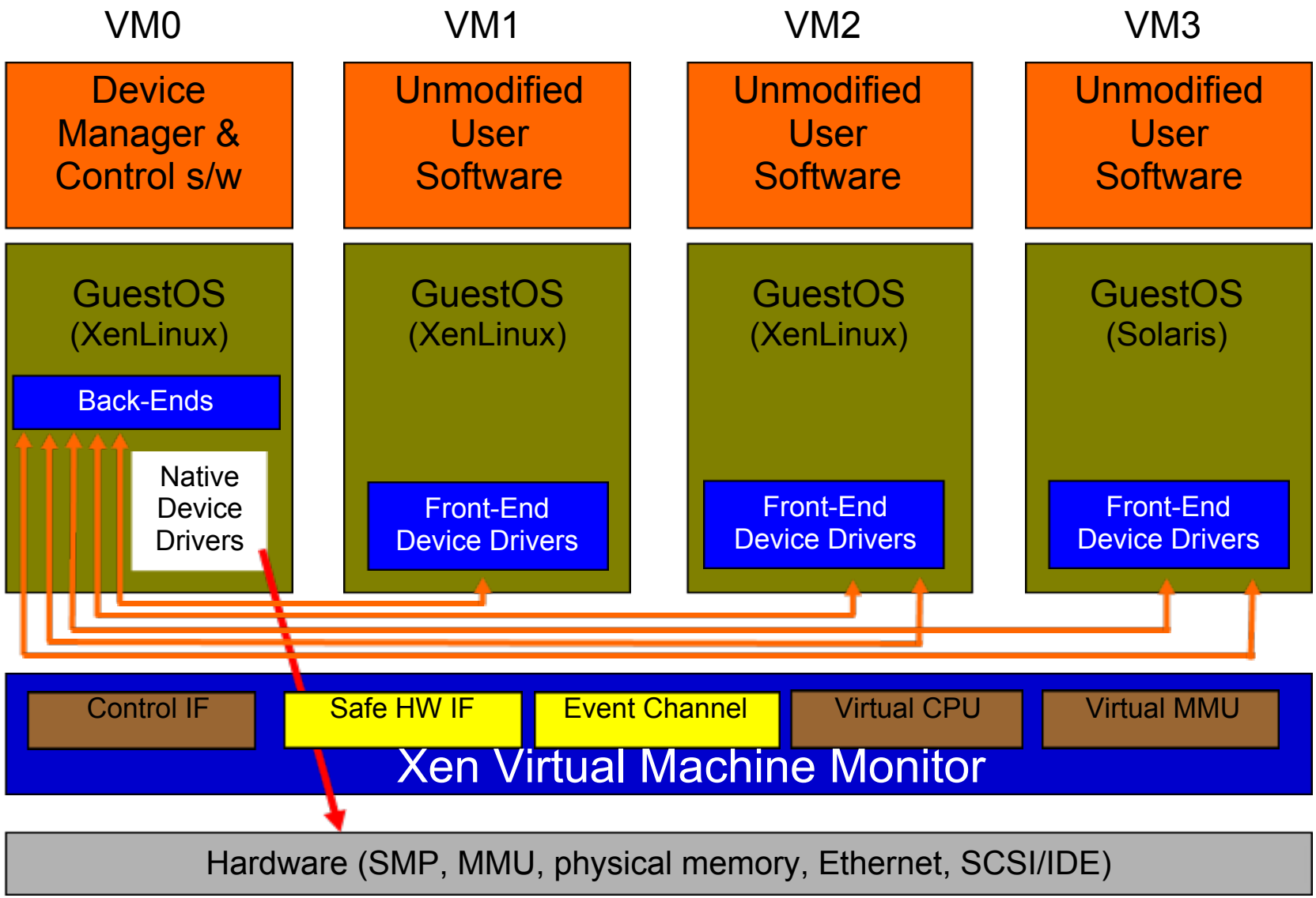
Xen GPU Architecture



Solution ?

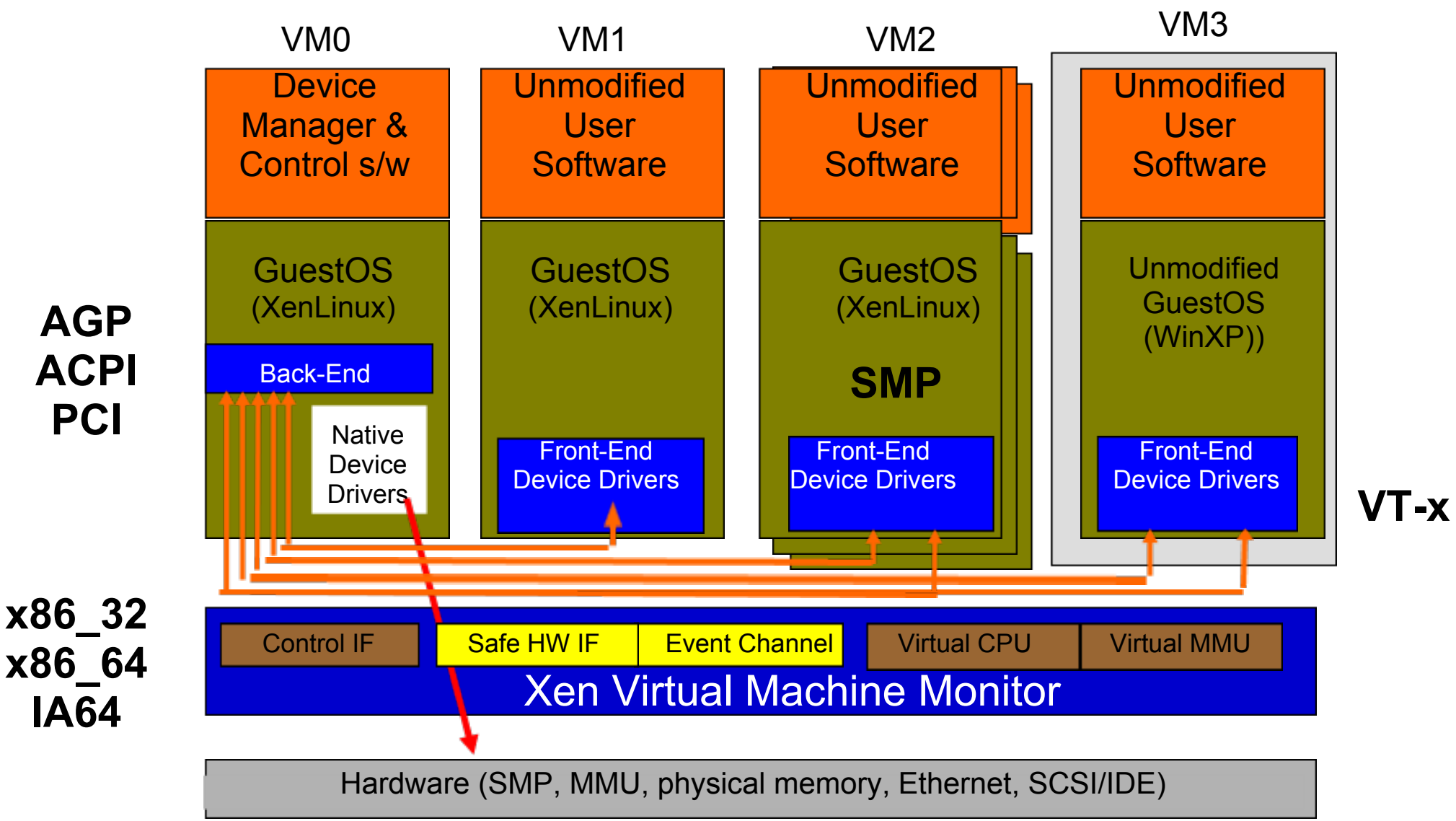
- **PCI Passthrough**
  - HW support ?
    - Intel,NVIDIA,ATI...
  - Bottleneck ?
- **OpenGL Virtualization - VMGL**
  - SW support ?

Xen 2.0 Architecture



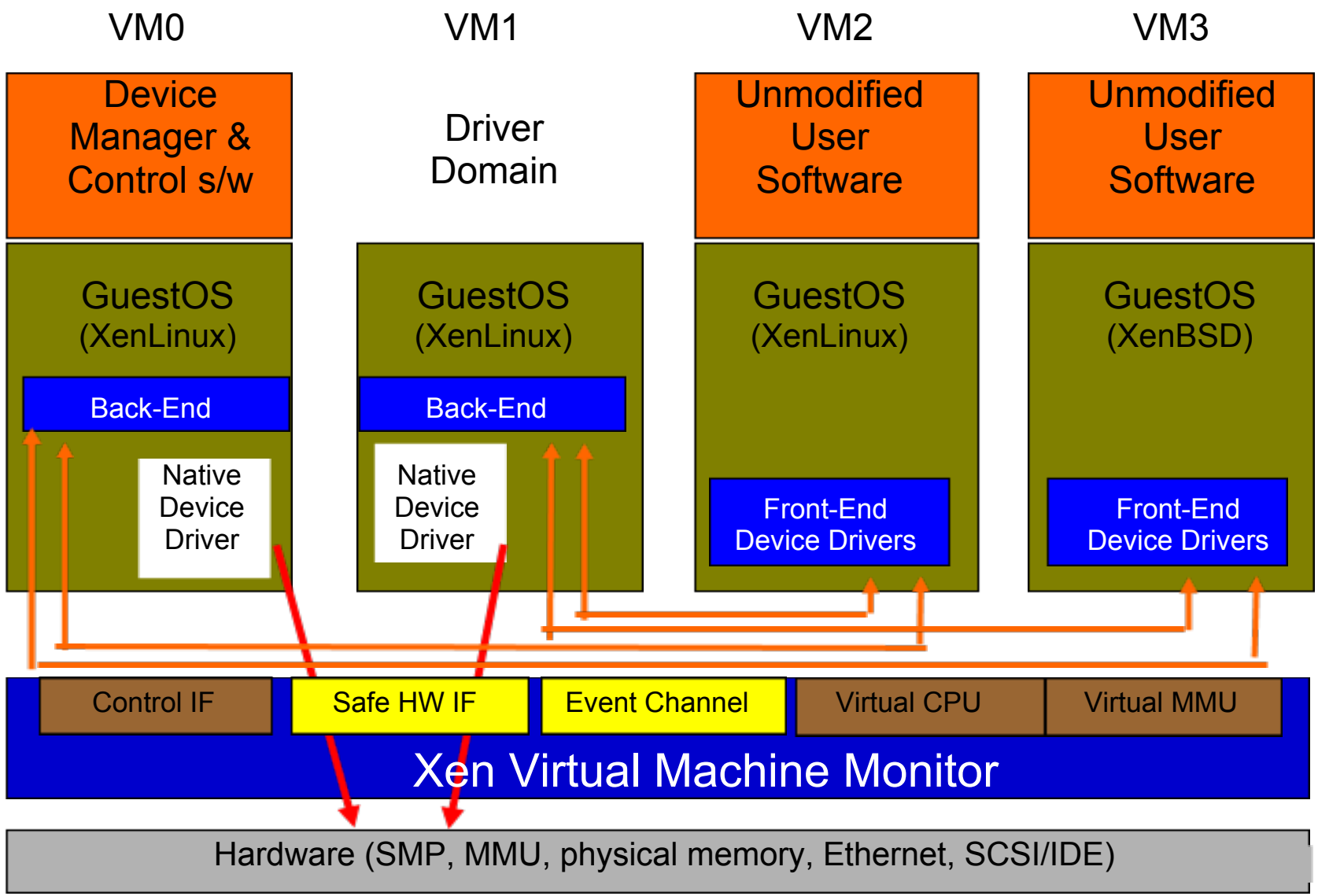
Reference: http://www.cl.cam.ac.uk/research/srg/netos/papers/2006-xen-fosdem.ppt

Xen 3.0 Architecture



Reference: http://www.cl.cam.ac.uk/research/srg/netos/papers/2006-xen-fosdem.ppt

Driver domains



Reference: http://www.cl.cam.ac.uk/research/srg/netos/papers/2006-xen-fosdem.ppt

PCI-Passthrough

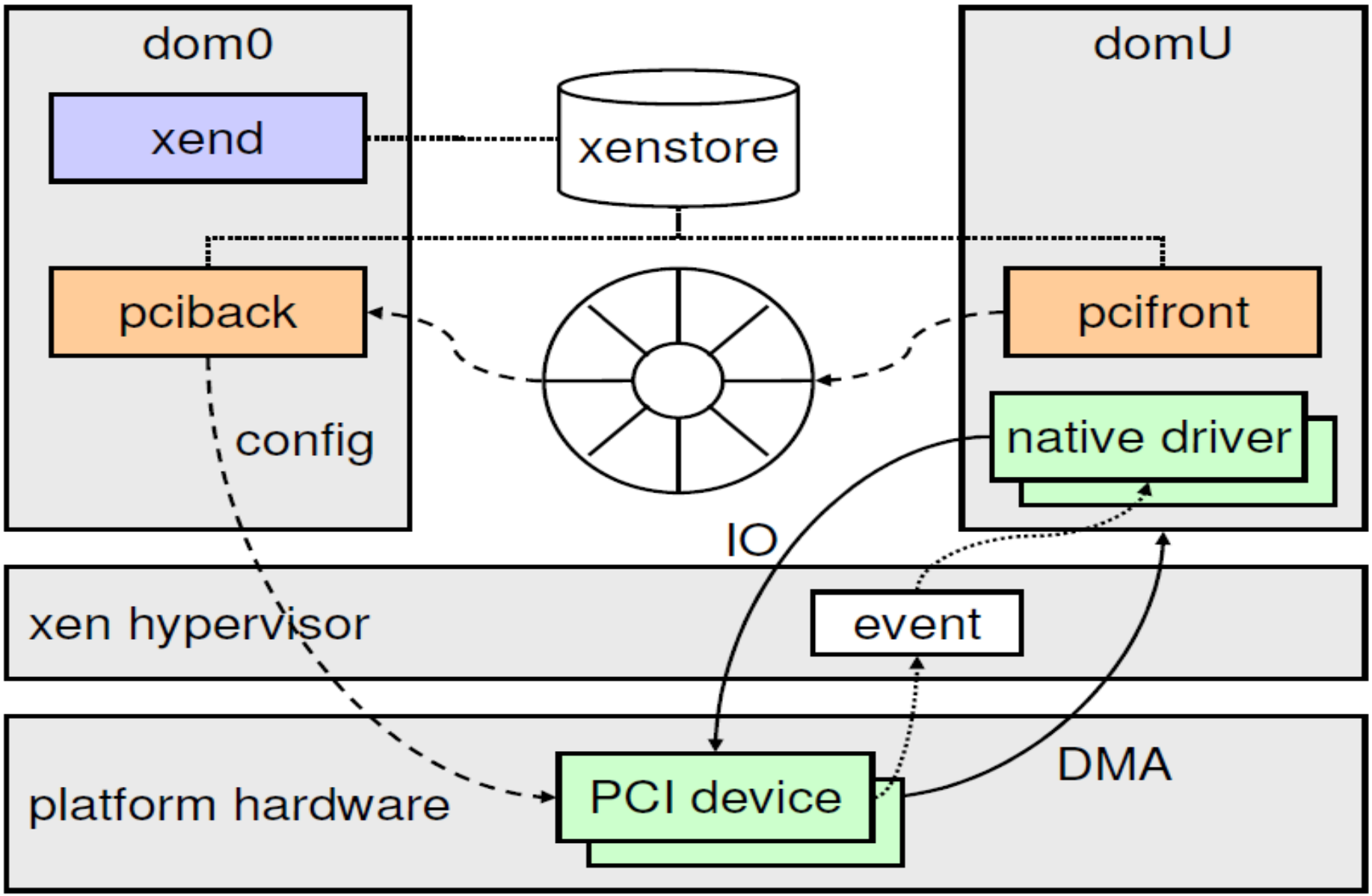
- **What is PCI Passthrough ?**
  - A guest VM directly accesses physical PCI devices
- **Xen supports PCI-passthrough**, which will give a virtual machine direct, exclusive access to a hardware device on the PCI-bus.
- This is useful for hardware testing on multiple distributions.

PCI-Passthrough (cont)

- **Advantages**
  - Improve I/O Performance
  - Improve VM scalability
- **Disadvantages**
  - A VM is bound to a device.
    - Save / restore state of Physical device
    - Problem - Live migrate with assigned device
    - Flexibility reduction
  - A device is bound to a VM.
    - Running VM preempt the device
    - Shortage of IO devices

Reference: http://xen.org/files/xensummit_tokyo/27_Yosukelwamatsu_en.pdf

Overview of PCI Passthrough

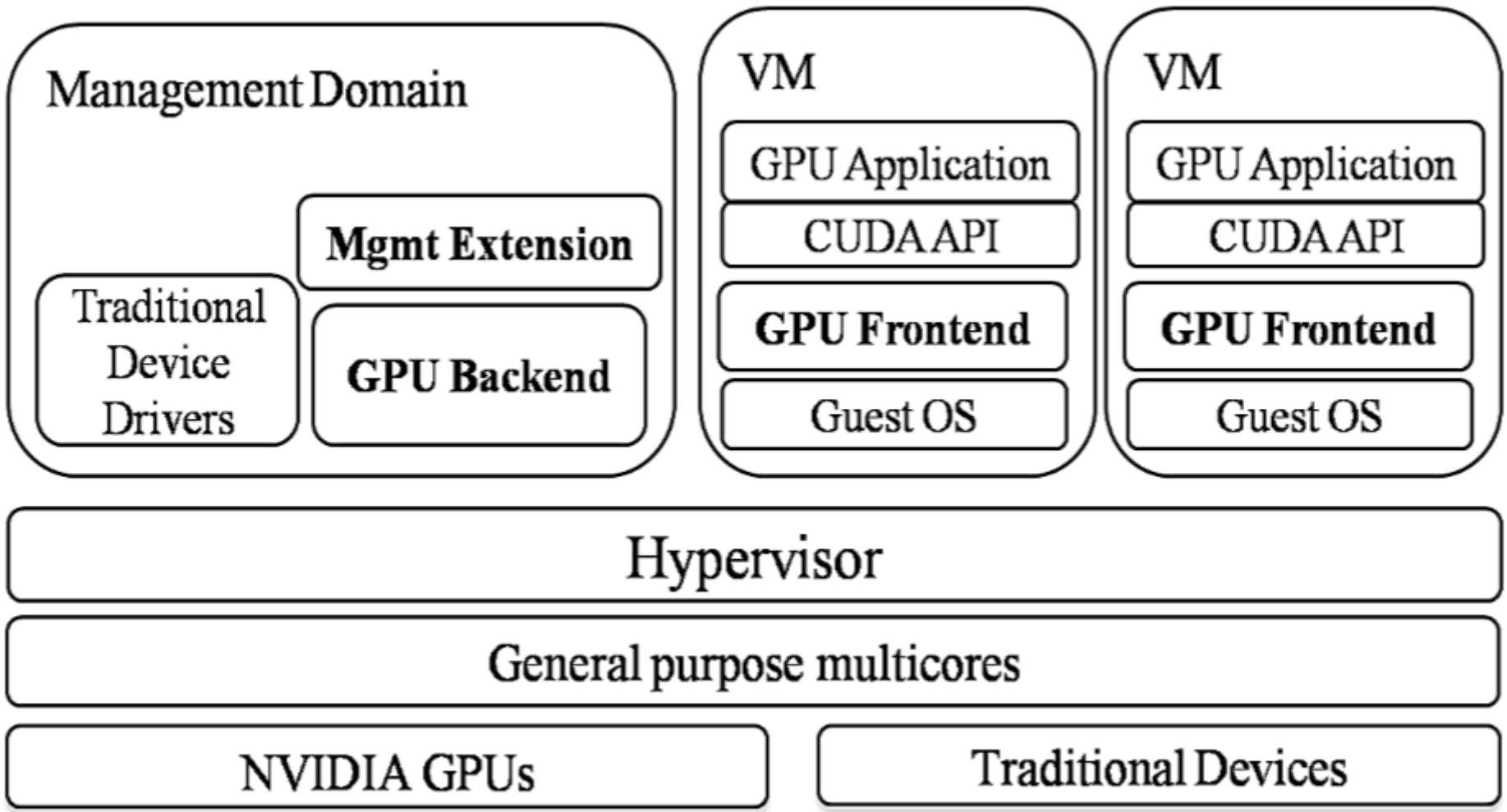


Reference: http://xen.org/files/xensummit_tokyo/27_Yosukelwamatsu_en.pdf

GPGPU On VMs

- **GViM**
  - Split-driver model
  - VMM-bypass mechanism
  - Lower-level memory management mechanism
- **Xen privileged domain(Dom0)**
  - Memory/Module management
  - Communication/Scheduling method

Virtualization of GPUs



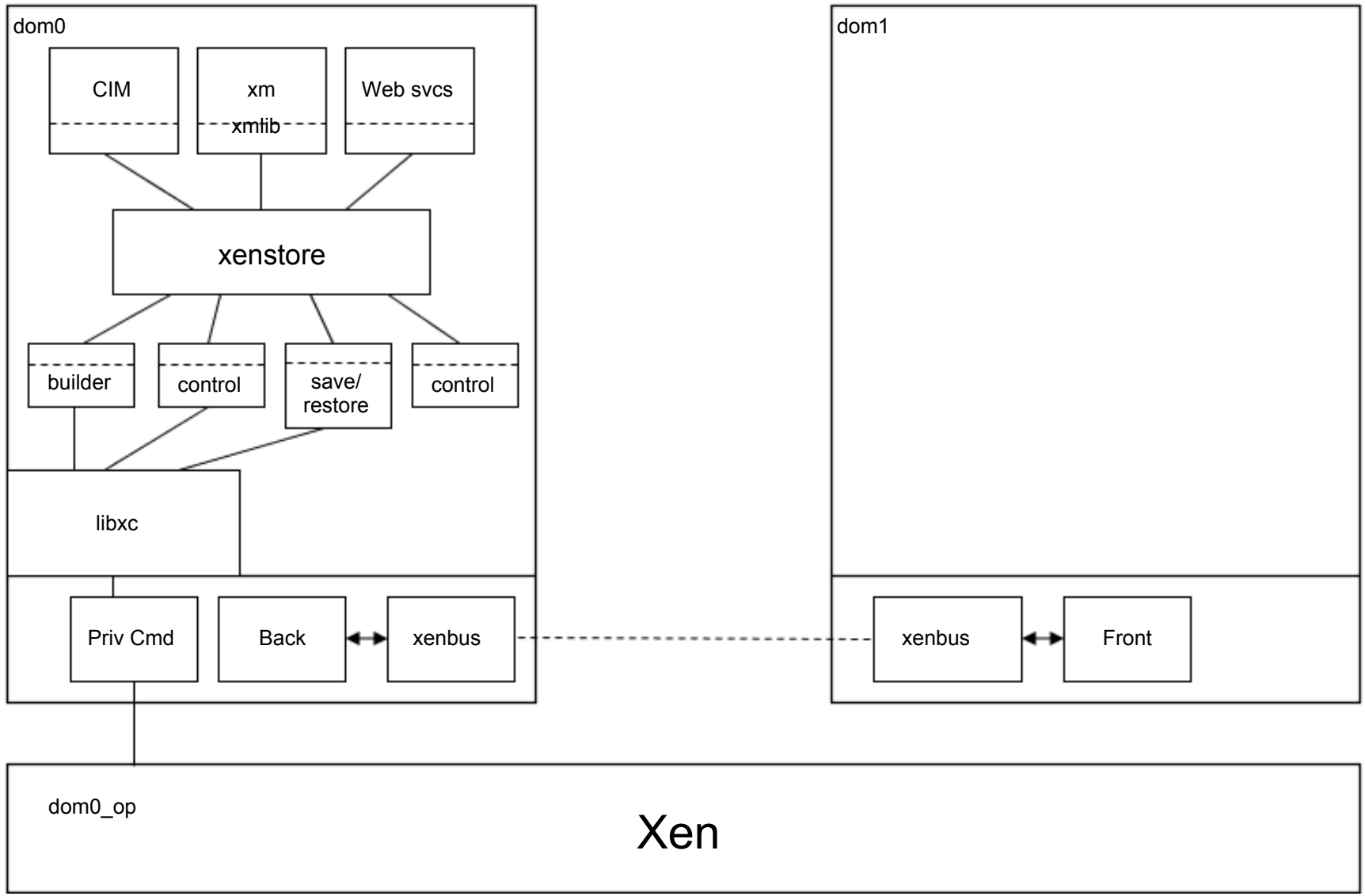
NVIDIA's CUDA – Compute Unified Device Architecture for managing GPUs

Reference: http://www.cc.gatech.edu/~vishakha/files/GViM.pdf

Virtualization Components

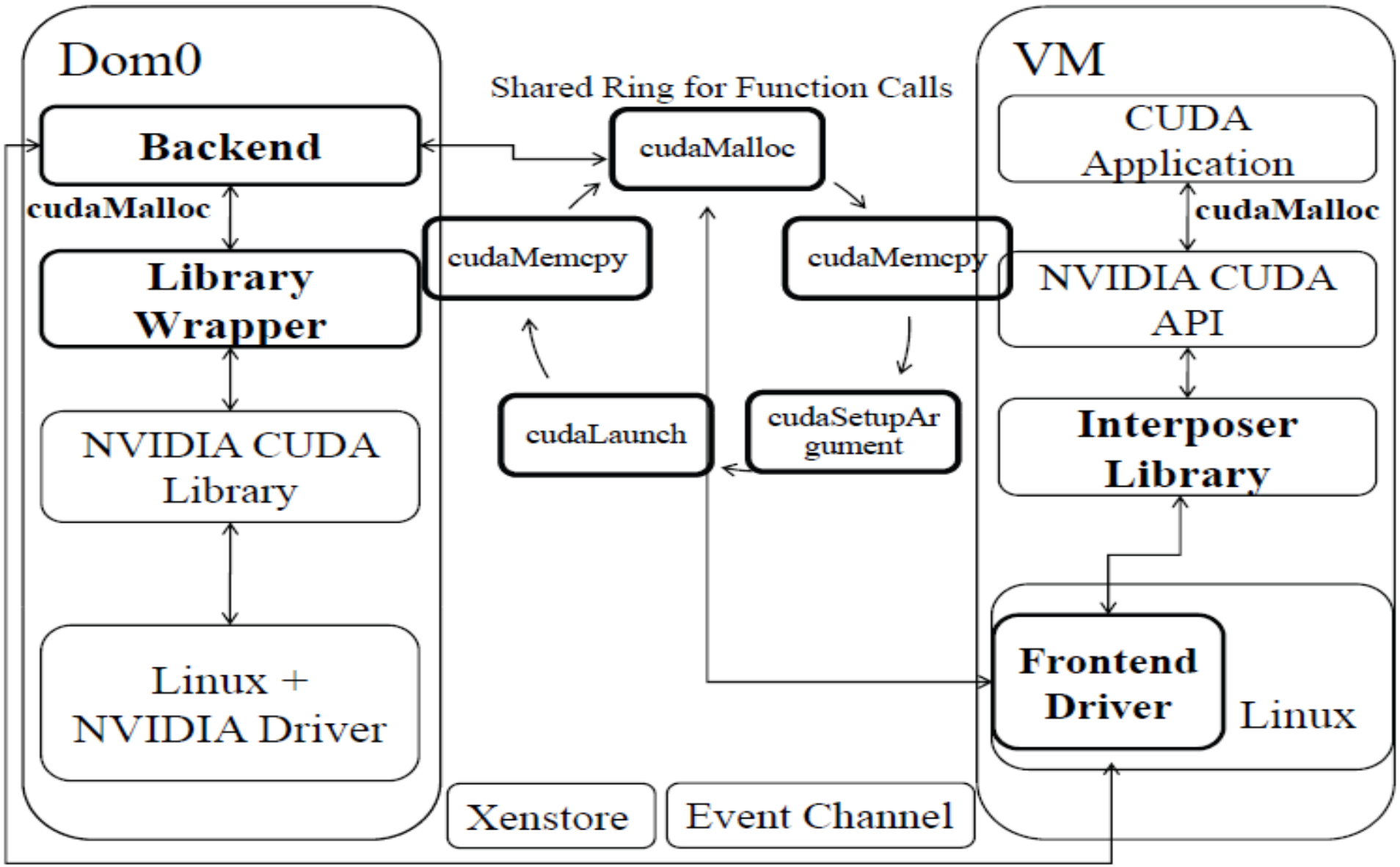
- **Using OpenGL/CUDA API**
- **The CUDA interposer library.**
- **Fronted driver**
  - Xen-bus
  - Xenstore
  - Shared call buffer

Xen Tools



Reference: http://www.cl.cam.ac.uk/research/srg/netos/papers/2006-xen-fosdem.ppt

Exp: Virtualization components



Reference: http://www.cc.gatech.edu/~vishakha/files/GViM.pdf

Conclusion

- Strengths
  - Cross-Platform(VMs), Cross-Architecture(HW)
- Weaknesses
  - Interrupt / Preemption enforcement under Xen 3.3
- Applications
  - GPGPU
- New trends
  - Intel VT-d (HW), NVIDIA SLI Multi-OS

Reference

- NVIDIA
  - http://www.nvidia.com/object/cuda_home.html
  - http://forums.nvidia.com/index.php?
  - David Kirk/NVIDIA and Wen-mei W. Hwu, 2007 ECE 498AL, University of Illinois, Urbana-Champaign
- GViM
  - http://www.cc.gatech.edu/~vishakha/files/GViM.pdf
- VMGL
  - http://www.cs.toronto.edu/~andreslc/xen-gl/
  - http://www.cs.toronto.edu/~andreslc/publications/LagarCavillaVEE07.pdf
  - http://www.cs.toronto.edu/~andreslc/publications/slides/Xen-Summit-2007/vmgl.pdf
- Xen
  - http://trac.nchc.org.tw/grid/wiki/Reading/XenP
  - http://www.virtuatopia.com/index.php/Xen_Virtualization_Essentials

Discussion