

# 华为GPU虚拟化介绍

韩伟东

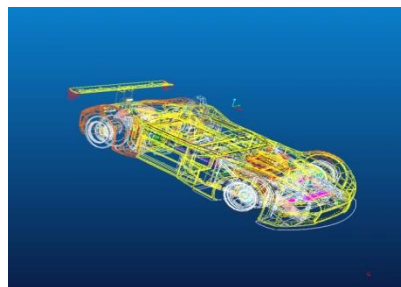
hanweidong@huawei.com

[www.huawei.com](http://www.huawei.com)

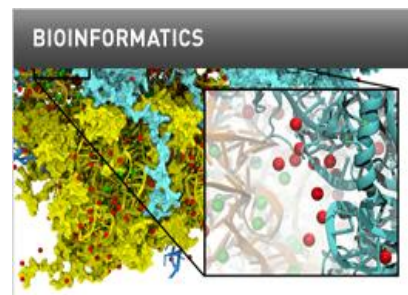
# GPU虚拟化应用场景



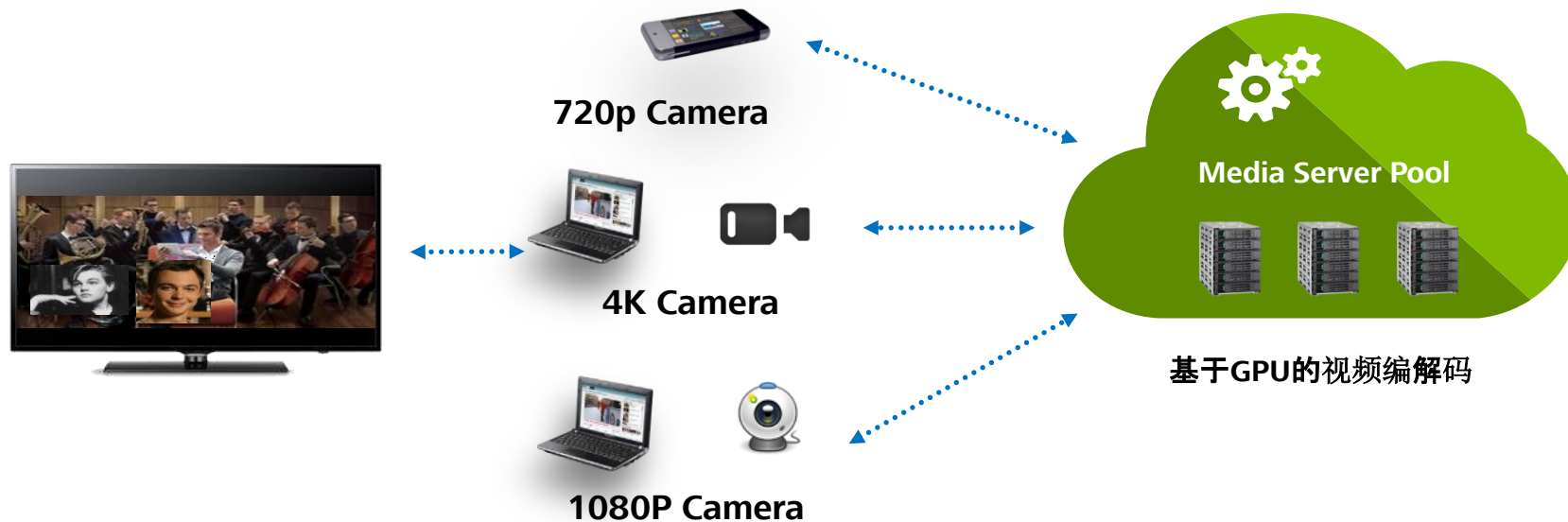
高清视频编辑



3D专业制图设计

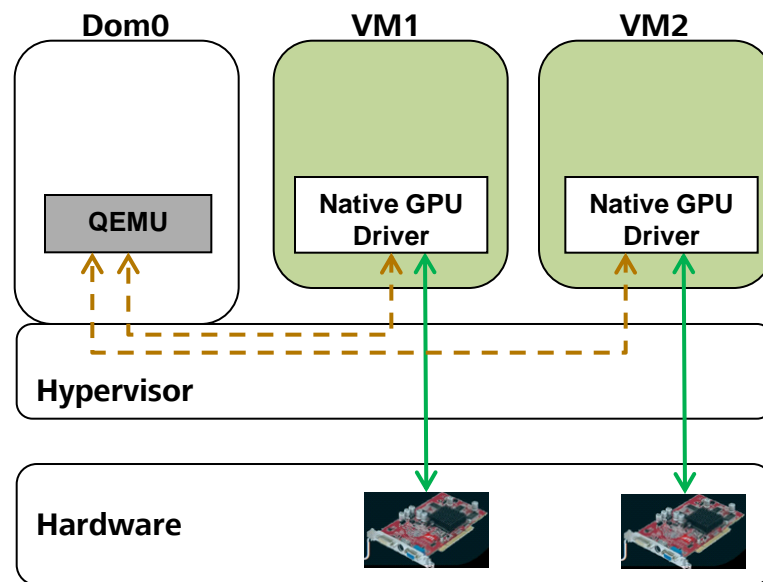


高性能计算

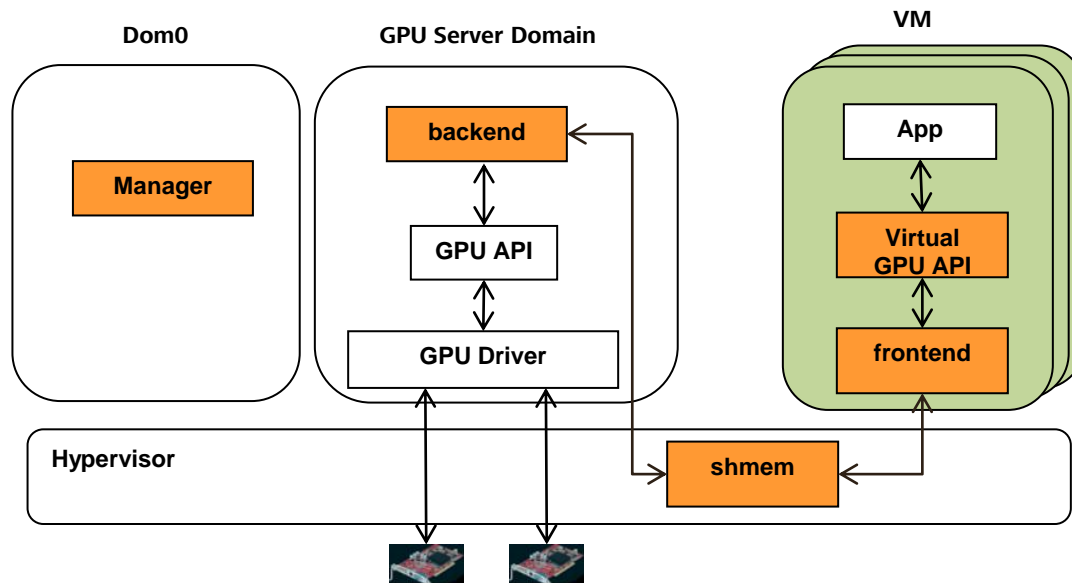


# GPU直通

- 通过IOMMU技术把物理GPU直通给虚拟机
  - DMA remapping
  - MMIO 映射
  - 中断通过hypervisor注入
  - PCI config访问通过QEMU模拟
- 性能接近物理GPU性能
- 不支持迁移



# GPU共享

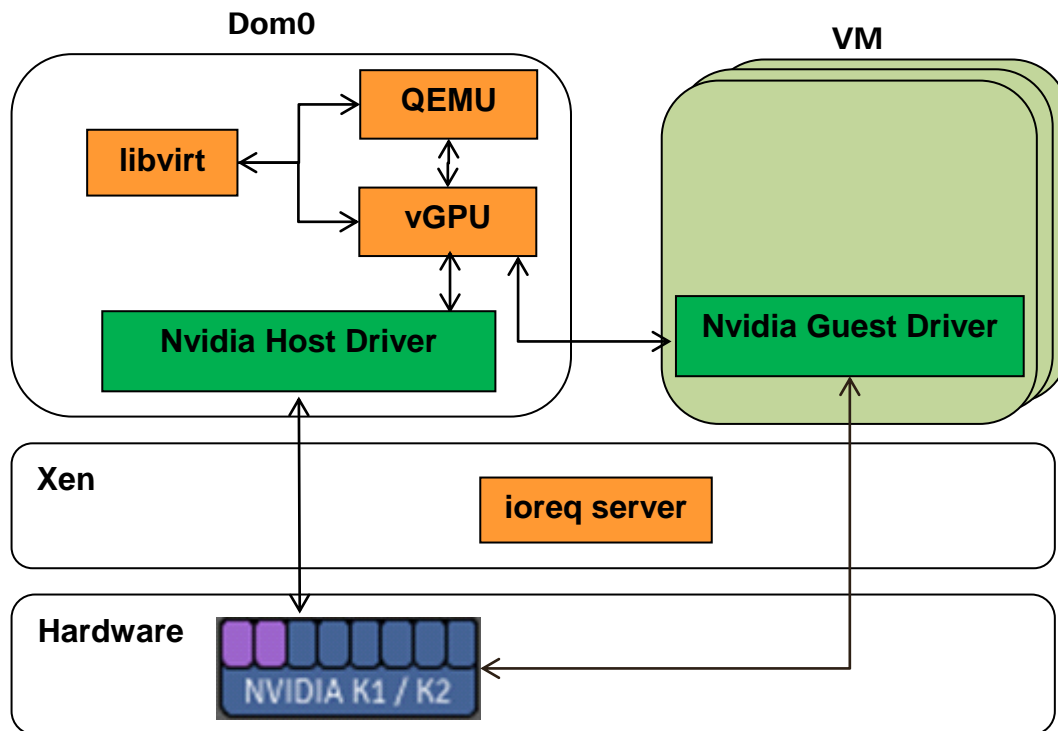


- 一个**GPU**可加速多个**VM**
- 引入**GPU Server Domain**，为其他虚拟机提供**GPU**加速服务
  - **GPU**直通给**GPU Server Domain**
- **API remoting**
  - 在**VM**中实现虚拟驱动截获和转发**GPU API**调用，实现**GPU**加速

# GPU共享

- 虚拟图形驱动：虚拟OpenGL ICD和虚拟WDDM驱动
  - 按照驱动规范实现虚拟驱动，对上层应用透明
  - 对截获的指令和数据进行聚合后再发送，提升性能
- 虚拟机间共享内存通信，提升GPU数据传输性能
  - 实现了Xen上的ivshmem
  - 实现一套共享内存接口供上层应用调用

# GPU硬件辅助虚拟化



- **GPU**硬件上专门为虚拟化设计，一个物理**GPU**可以虚拟出多个**vGPU**给不同的虚拟机使用
  - **vGPU**设备模拟
  - 引入**ioreq server**支持两个设备模拟进程
  - **Nvidia**驱动实现了硬件辅助虚拟化的能力

# 总结与展望

- 不掌握GPU架构和驱动，很难实现出高效的GPU虚拟化方案
- GPU硬件辅助虚拟化是大势所趋
  - 各GPU厂商已经在发力
  - 最好能把硬件虚拟化的能力开放出来，让更多的人来开发（不仅仅是GPU厂商），加速技术发展
- GPU硬件辅助虚拟化目前看都还不支持热迁移
- 容器场景下，GPU虚拟化怎么做？

**Thank you**

[www.huawei.com](http://www.huawei.com)