

Research and application of GPU pass-through based on KVM in desktop cloud

Lin Zhou^{1,a,*}, Jianxin Song^{2,b}, Jinkun Yuan^{3,c} and Haifeng Han^{4,d}

¹ Student, Nanjing University of Posts and Telecommunications, nanjing 210000, China

² Prof., Nanjing University of Posts and Telecommunications, nanjing 210000, China

³ Engineer, NANJING ASTUTE SOFTWARE TECHNOLOGY CO.,LTD.nanjing 210000, China

⁴ Student, Nanjing University of Posts and Telecommunications, nanjing 210000, China

^az1443619637@163.com, ^biamsongjx@139.com, ^cyuan.jinkun@astute-tec.com, ^d1334574047@qq.com

ABSTRAT

At present,the desktop cloud has come into people's life.You'll see them in some companies, offices, and school labs.The problem I am trying to solve in this paper is that because the computing power of virtual desktop is provided by the virtual machine,and then transmit data to the virtual desktop to display through the network,so compared with the traditional PC in terms of performance, there are still gaps, it is difficult to cope with the application of high load ,such as 3D animation, high-definition video processing, etc.The approach I adopt to solve the problem is that we can use vfio-pci technology to direct GPU to a virtual machine,Enables the virtual machine to monopolize the graphics card,And be able to achieve more than 90% of the performance of graphics,This is a huge improvement for users, and many desktop cloud users are not limited to general enterprise applications,Such as Office, Web applications, Flash playback, video playback and so on.After pass-through,Users can carry out high load applications, 3D animation, high-definition video processing, etc.There are several companies doing virtualization, but KVM is an open source virtualization solution, its market share is not very high, KVM virtualization based on direct display card is not a mature scheme, this is another contribution to the open source.

Keywords: Openstack Cloud;KVM ;Vfio-pci;GPU Passthrough with qemu

1. INTRODUCTION

Desktop cloud refers to access a cross platform application through a thin client or any other network connected device.The user's desktop environment is centrally deployed in the enterprise's data center, and the local terminal is only a display device.Desktop virtualization has a lot of disadvantages, this paper studies and analyses the main performance problems of the virtual desktop, because virtual card cannot complete the high load task, this paper focuses on how to achieve the transmission of GPU based on QEMU-VFIO.

2. Related Technology (Pass-through)

2.1 Fixed Pass-throug

The fixed pass-through (fixed pass-through: A virtual machine using a GPU exclusively) is adopted in this article, enables virtual machines to use full performance of graphics cards.

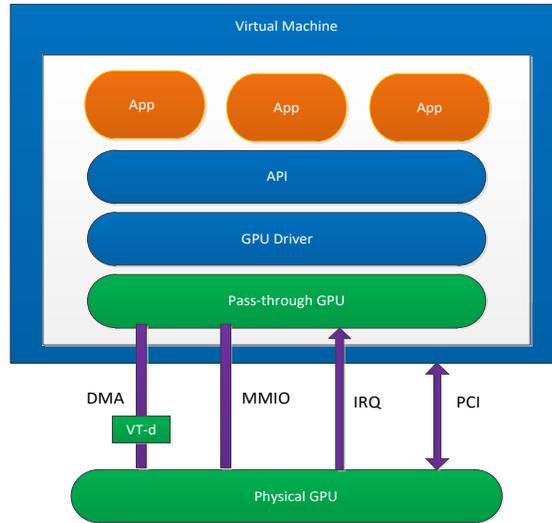


Fig.1 Fixed Pass-through

2.2 Using VFIO mode to make GPU pass through in KVM/QEMU

VFIO is supported after Linux kernel 3.6/qemu 1.4 and only supports PCI devices currently. VFIO is a user-space driver framework that provides two basic services: It provides device access interfaces to userspace, and provides configuration IOMMU interfaces to userspace. It just opens the IOMMU interface to userspace for the first time, configuring IOMMU entirely in user mode, mapping the DMA address space and limiting it within the process virtual address space. VFIO can be utilized to implement efficient user-space drivers. The virtualized scene can be used for PCI device pass-through. By configuring the IOMMU interface through the user space, the DMA address space mapping can be restricted to the process virtual space. This is especially important for a driver of high performance and device pass-through in virtualization scenarios. VFIO supports UEFI better than traditional methods, VFIO technology enables userspace to access devices directly. Users are not bound to have root permissions, more security, more functions.

3. System & environment

We need to do some environmental preparation to see if the hardware supports the pass-through of graphics cards. Not all cards support pass-through, FirePro cards, W2100, W3100, W9100 and so on are currently known in AMD graphics cards, and it needs Intel's vt-d technical support.

Hardware Information:

GPU: AMD Firepro W2100 CPU: Intel(R) Core(TM) i3-6100 CPU @ 3.70GHz MEM: 32G

SSD: 480G

OS: CentOS 7.3 + Ocata

4. Implementation

- ① Update kernel for 'CONFIG_VFIO_PCI_VGA' and 'i915 VGA arbiter patch'


```
[root@host016 gpu]# grep CONFIG_VFIO_PCI_VGA /boot/ -r
/boot/config-3.10.0-514.16.1.el7.centos.plus.x86_64:CONFIG_VFIO_PCI_VGA=y
```
- ② Modify grub parameter for intel_iommu and i915


```
[root@host016 gpu]# vi /etc/default/grub
```

```
GRUB_CMDLINE_LINUX="crashkernel=auto rhgb quiet idel=poll idel=poll intel_iommu=on
i915.enable_hd_vgaarb=1modprobe.blacklist=nouveau"
```

```
[root@host016 gpu]# grub2-mkconfig -o /boot/grub2/grub.cfg
```

- ③ Modify default load system module and paramter

```
[root@host016 home]# echo "blacklist radeon" >> /etc/modprobe.d/blacklist.conf
```

```
[root@host016 home]# echo "options vfio_iommu_type1 allow_unsafe_interrupts=1" >
```

```
/etc/modprobe.d/vfio_iommu_type1.conf
```

```
[root@host016 home]# echo "options kvm ignore_msrs=1" >> /etc/modprobe.d/kvm.conf
```

- ④ Check pci info

```
[root@host016 home]# lspci | grep AMD
```

```
01:00.0 VGA compatible controller: Advanced Micro Devices, Inc. [AMD/ATI] Oland GL [FireProW2100]
```

```
01:00.1 Audio device: Advanced Micro Devices, Inc. [AMD/ATI] Cape Verde/Pitcairn HDMI Audio
```

```
[Radeon HD 7700/7800 Series]
```

- ⑤ Detach gpu from host

```
[root@host016 home]# lspci -s 01:00.0 -v
```

```
01:00.0 VGA compatible controller: Advanced Micro Devices, Inc. [AMD/ATI] Oland GL [FirePro
W2100] (prog-if 00 [VGA controller])
```

```
Subsystem: Advanced Micro Devices, Inc. [AMD/ATI] Device 2120
```

```
Flags: bus master, fast devsel, latency 0, IRQ 16
```

```
Memory at e0000000 (64-bit, prefetchable) [size=256M]
```

```
Memory at f7200000 (64-bit, non-prefetchable) [size=256K]
```

```
I/O ports at e000 [size=256]
```

```
Expansion ROM at f7240000 [disabled] [size=128K]
```

```
Capabilities: [48] Vendor Specific Information: Len=08 <?>
```

```
Capabilities: [50] Power Management version 3
```

```
Capabilities: [58] Express Legacy Endpoint, MSI 00
```

```
Capabilities: [a0] MSI: Enable- Count=1/1 Maskable- 64bit+
```

```
Capabilities: [100] Vendor Specific Information: ID=0001 Rev=1 Len=010 <?>
```

```
Capabilities: [150] Advanced Error Reporting
```

```
Capabilities: [200] #15
```

```
Capabilities: [270] #19
```

```
Kernel driver in use: vfio-pci
```

```
Kernel modules: radeon
```

If the information is: *Kernel driver in use: vfio-pci*. That explains the successful separation of video cards.

- ⑥ Check USB keyboard and mouse info

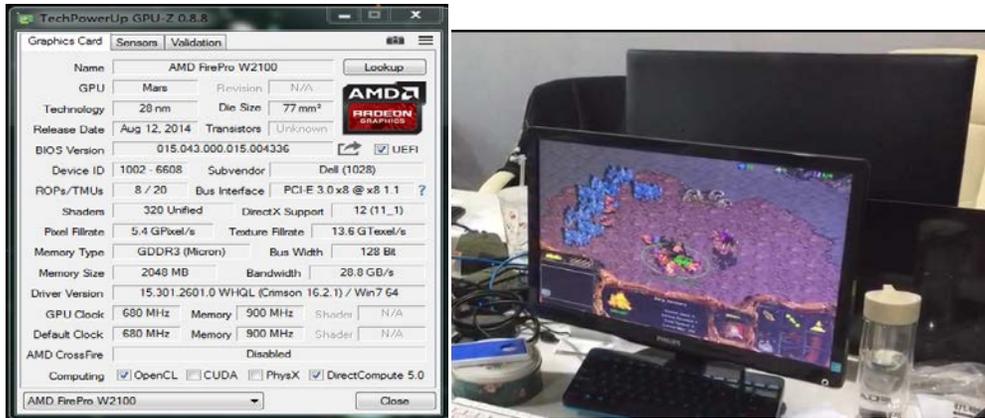
- ⑦ Create vm network(bridge)

- ⑧ Build start vm script

- ⑨ Boot vm and start

5. Results and discussion

Download the 3DMark software and GPU-Z software in the virtual machine, tested the performance of the AMD card, the results are as follows, we also used to play star games, The effect is very good and I can't feel the game is in the virtual machine.



As you can be clear from the above picture, the performance is very strong after passing through. The pass through of video cards has a certain complexity. There are different schemes for passing through. The hardware condition is very harsh. After a few tests, it's not very stable yet.

6. Conclusion

OpenStack and KVM have attracted much attention in the IT community, and KVM is an open source virtualization scheme, it has a strong development capacity, although the current market share is not high, maturity is not enough, But KVM has a very important place in the OpenStack computing virtualization layer, Therefore, in the KVM program, the video card pass through is still very long-term strategic significance.

Reference

- [1] Virtualization with KVM, Irfan Habib , February 2008 Linux Journal: Volume 2008 Issue 166, February 2008
- [2] Managing KVM deployments with Virt-Manager Michael J. Hammel January 2011 Linux Journal: Volume 2011 Issue 201, January 2011
- [3] How to assign devices with VT-d in KVM. http://www.linux-kvm.org/page/How_to_assign_devices_with_VT-d_in_KVM .
- [4] Adding a PCI device with virsh. https://access.redhat.com/documentation/en-US/Red_Hat_Enterprise_Linux/5/html/Virtualization/chap-Virtualization-PCI_passthrough.html#sect-Virtualization-PCI_passthrough-AI_device_to_a_host .
- [5] Guest virtual machine device configuration. https://access.redhat.com/documentation/en-US/Red_Hat_Enterprise_Linux/6/html/Virtualization_Administration_Guide/chap-Guest_virtual_machine_device_configuration.html .
- [6] Intel?Virtualization Technology for Directed I/O (VT-d). <http://ark.intel.com/search/advanced/?s=t&VTD=true> .
- [7] Lahabar S, Agrawal, Narayanan P J, High Pformance patern recognition on GPU [C] //Proc of National conference on Computer Vision pattern Recognition Image Processing and Graphics, 2008: 154-159.
- [8] Xen VGAPassthrough [EB/OL]. [2012-12-06]. [wiki.Xen-source.com/Xenwiki/Xen VGAPassthrough](http://wiki.Xen-source.com/Xenwiki/Xen_VGAPassthrough).
- [9] Shi L, Chen H , Sun J. vCUDA: GPU accelerated high performance computing in virtual machines [C] //Proc of International Parallel & Distributed Processing Symposium , 2009: 1-11