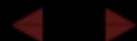


VFIO, OVMF, GPU, and You

The state of GPU assignment in QEMU/KVM

Alex Williamson / alex.williamson@redhat.com



The current state of VGA assignment



VGA assignment defined:

- Graphics card assigned as primary graphics for the VM
- Uses VGA BIOS for initialization and runtime services



Status:

- Works for discrete graphics cards
 - AMD Radeon (HD5xxx+)
 - NVIDIA GeForce (8-series+)
- Requires *experimental* vfio-pci option
 - x-vga=on
- Requires working VGA arbitration
- Still some device specific glitches
 - BSODs with AMD Catalyst + HDMI audio
 - Reset issues on some Radeon cards
 - Code 43 on GeForce



Issues:

- Requires *experimental* vfio-pci option, x-vga=on
 - Not supported by libvirt because experimental
 - Experimental because...
- Requires working VGA arbitration
 - Drivers don't participate (vgacon)
 - Drivers lie (i915)
 - Xorg DRI can't handle multiple arbitration participants



Can't we fix these?

- Drivers don't participate
 - fixable
- Drivers lie
 - Once upon a time...
 - i915 hardware allowed control of VGA resources
 - designers dropped this feature
 - nobody noticed?!
 - i915 continues to opt-out anyway
 - But if we fix that...
- Xorg DRI can't handle multiple arbitration participants



Start with fixing Xorg

DRI wants to mmap VGA MMIO space

Possible options...

- Provide a new VGA arbiter interface allowing mmmaps
- Remove the mmap requirement

But what about compatibility?

“WE DO NOT BREAK USERSPACE!”

-Linus Torvalds



VGA assignment status:

Stuck



What else could we do?



Do we really need VGA?



Secondary graphics

- Emulated VGA for VM primary display
- Assigned graphics card for secondary
- Supported by NVIDIA
 - Requires K-series Quadro/GRID/Tesla
 - Supported in RHEL7.0
- Some reports of success with Radeon

Depends on guest driver support



Legacy-free UEFI VM

- No VGA required

No VGA arbitration required!



Problem solved?

VGA Issues:

- Requires "experimental" vfio-pci option
 - Not supported by libvirt because experimental
 - Experimental because...
- Requires working VGA arbitration
 - Drivers don't participate (vgacon)
 - Drivers lie (i915)
 - Xorg DRI cannot handle multiple arbitration participants



No Blockers



How do UEFI?



Step 1: OVMF

Open Virtual Machine Firmware

An implementation of UEFI built on the EDK II code base

Replaces Seabios



OVMF Setup

via command line:

```
/usr/libexec/qemu-kvm ... \  
-drive if=pflash,format=raw,readonly,file=/path/to/OVMF_CODE.fd \  
-drive if=pflash,format=raw,file=/copy/of/OVMF_VARS.fd
```

via libvirt:

```
<domain type='kvm'>  
  ...  
  <os>  
    <loader readonly='yes' type='pflash'>/usr/share/edk2.git/ovmf-x64/OVMF_CODE-pure-efi.fd</loader>  
    <nvram template='/usr/share/edk2.git/ovmf-x64/OVMF_VARS-pure-efi.fd' />  
    ...  
  </os>  
</domain>
```



Step 2: EFI graphics card ROM

PCI ROMs can support multiple images

Most newer cards already support EFI



Step 3: UEFI compatible guest

Newer versions of Linux

Windows 8, Server 2012



That's it!



Other GPU Topics



440FX vs Q35

Q35 originally recommended for VFIO-VGA

Largely FUD

- Q35 *looks* more like the host platform
- Some Linux drivers *assume* an upstream PCIe port

In reality

- Windows doesn't care
- Linux is fixable

440FX is likely the *easier* path for Windows guests



NVIDIA: Code 43

Driver detects KVM hypervisor, fails to initialize^{*}
Nvidia - "Accidental" breakage, won't fix, unsupported

We can't solve it, but we can work around it

Hide the hypervisor

via command line:

```
-cpu [type],kvm=off
```

via libvirt:

```
<domain type='kvm'>  
  ...  
  <features>  
    <kvm>  
      <hidden state='on' />  
    </kvm>  
  ...  
</features>  
</domain>
```

*NVIDIA driver version 338.77+



NVIDIA: Code 43 Redux

Driver detects Hyper-V extensions, fails to initialize^{*}

Workaround: Disable use of Hyper-V extensions

Performance hit for Windows VMs

Have we entered an arms race?

^{*}NVIDIA driver version 344.11



Quirks

GeForce: One still required*

Radeon: Not needed!

*QEMU now enables this independent of x-vga (2.2)



Unresolved Issues

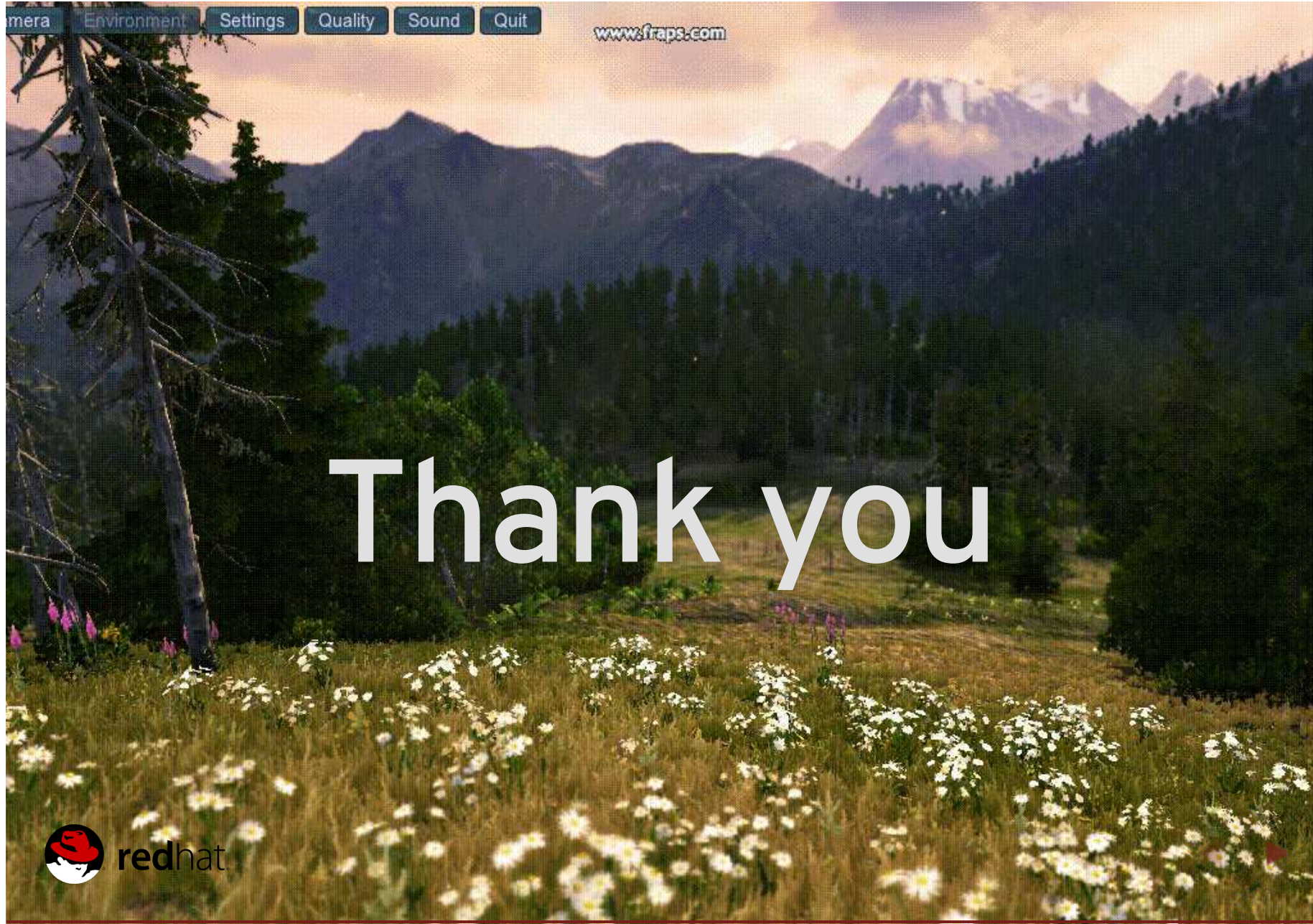
- Radeon device resets
 - Sea Island GPUs don't reset correctly
 - SMC firmware issue?
 - Southern Island GPUs indicate PM reset capability
 - doesn't work
- Radeon BSOD
 - Typically resolved by avoiding audio function
- Intel Graphics support
 - In progress? KvmGT?
- Host/Guest suspend
 - And ponies!



Resources

- VFIO tips and tricks blog:
<http://vfio.blogspot.com>
- This presentation:
<http://awilliam.github.io/presentations/KVM-Forum-2014>





Camera Environment Settings Quality Sound Quit

www.fraps.com

Thank you





Alex Williamson / alex.williamson@redhat.com

