

Adventures in Virtualisation for Gaming

Adventures in Virtualisation for Gaming

Or, let's learn systemd, bcache, btrfs, EFI, KVM, libvirt, and PCIe ACS in one week...

Adventures in Virtualisation for Gaming

Or, “why I don’t have screenshots”

Why I use computers

```
vilab@vilab-A560-SH6AK: ~/catkin_ws
roscore http://localhost:11311/
scores.
####
#### Running command: "make cmake_check_build_system" in "/home/vilab/catkin_ws/build"
####
#### Running command: "make -j8 -l8" in "/home/vilab/catkin_ws/build"
####
[ 0%] [ 0%] [ 4%] [ 4%] [ 4%] [ 4%] Built target geometry_msgs_generate_messages_cpp
Built target std_msgs_generate_messages_py
Built target geometry_msgs_generate_messages_py
Built target std_msgs_generate_messages_cpp
Built target std_msgs_generate_messages_lisp
Building CXX object view_controller_msgs/CMakeFiles/testsrv.dir/src/testsrv.cpp.o
[ 4%] Built target geometry_msgs_generate_messages_lisp
[ 16%] Built target rviz_animated_view_controller
[ 16%] Built target _view_controller_msgs_generate_messages_check_deps_CameraPlacement
[ 29%] [ 29%] [ 29%] Built target view_controller_msgs_generate_messages_py
Built target view_controller_msgs_generate_messages_lisp
Built target _oroca_ros_tutorials_generate_messages_check_deps_srvTutorial
[ 29%] [ 33%] [ 33%] Built target _beginner_tutorials_generate_messages_check_deps_Num
Built target view_controller_msgs_generate_messages_cpp
Built target _oroca_ros_tutorials_generate_messages_check_deps_msgTutorial
[ 41%] [ 41%] Built target beginner_tutorials_generate_messages_cpp
Built target beginner_tutorials_generate_messages_lisp
[ 58%] [ 66%] Built target oroca_ros_tutorials_generate_messages_py
[ 70%] Built target oroca_ros_tutorials_generate_messages_cpp
[ 75%] Built target beginner_tutorials_generate_messages_py
Built target view_controller_msgs_generate_messages
[ 83%] Built target oroca_ros_tutorials_generate_messages_lisp
[ 83%] Built target beginner_tutorials_generate_messages
[ 87%] [ 87%] [ 91%] Built target ros_tutorial_srv_client
Built target oroca_ros_tutorials_generate_messages
Built target ros_tutorial_msg_publisher
[ 95%] Built target ros_tutorial_msg_subscriber
[100%] Built target ros_tutorial_srv_server
In file included from /home/vilab/catkin_ws/src/view_controller_msgs/src/rviz_animated_view_controller.h:35:0,
                 from /home/vilab/catkin_ws/src/view_controller_msgs/src/testsrv.cpp:2:
/opt/ros/indigo/include/rviz/view_controller.h:35:19: fatal error: QCursor: No such file or directory
#include <QCursor>
                 ^
compilation terminated.
make[2]: *** [view_controller_msgs/CMakeFiles/testsrv.dir/src/testsrv.cpp.o] Error 1
make[1]: *** [view_controller_msgs/CMakeFiles/testsrv.dir/all] Error 2
make: *** [all] Error 2
Invoking "make -j8 -l8" failed
vilab@vilab-A560-SH6AK:~/catkin_ws$
```

Why I use computers



And So....



Virtualisation is popular, how hard could this be?

Gaming Priorities

1. GPU
2. IO
3. CPU

Filesystem

1x NVMe SSD

bcache; (install media support is lacking...)

4x 1TB HDD

BTRFS RAID1; (may complicate bootloader config...)

Both complicated my initramfs

But: bcache *vastly* improved latency

CPU

VM through libvirt (KVM)

BIOS 'helpfully' disables virtualisation, iommu, etc, etc

Nothing particularly special here

'virt-manager' GUI suffices

GPU – The ‘fun’ part

1x nVidia 1070: for gaming

1x AMD <whatever’s cheap>: for admin

PCIe has an ‘Access Control System’

IOMMU can isolate arbitrary PCIe devices for VMs

But...

GPU – The ‘fun’ part

Problem:

IOMMU disabled by default in the kernel:

grub.cfg:

```
linux /path/to/kernel iommu=on amd_iommu=on
```

GPU – The ‘fun’ part

Problem:

IOMMU groupings are... inflexible...

Solution:

/usr/src/linux: pcie_acs_override.patch

grub.cfg: pcie_acs_override=downstream

Let's pretend everything 'just works'

GPU – The ‘fun’ part

Problem:

Host GPU drivers initialise the hardware too early

Solution:

```
/etc/modprobe.d/vfio.conf
```

```
options vfio_pci ids=1002:1b81,10de:10f0
```

```
options vfio_iommu_type1 allow_unsafe_interrupts=1
```

GPU – The ‘fun’ part

Problem:

UEFI/BIOS GPU boot priority initialises one GPU *really* early

Solution:

Open your case and swap your GPUs around.

Don't use this GPU for your guests

GPU – The ‘fun’ part

Problem:

nVidia have ‘convenient bugs’ that prevent PCIe passthrough

Solution?:

```
virsh edit <vmname>
```

```
<vendor_id state='on' value='totally_not_kvm' />
```


Ongoing Tasks

nVidia GPU forwarding

USB controller forwarding

CPU pinning

But, overall, it seems to work.