

vStation, le VDI sur ma station.

David DELAVENNAT

Cellule Informatique Mutualisée PHYMATH
CMLS (UMR 7640) / CMAP (UMR 7641) / CPHT (UMR 7644)
[Jean-Luc BELLON, Dan PHAM-KIM, Sylvain FERRAND]

Journée Mathrice de printemps 2017

Plan

- ✧ Motivation
- ✧ Constat
- ✧ Théorie
- ✧ Matériel

Motivation

Motivation

- ✦ rationaliser les méthodes de creation/déploiement/gestion :
 - ✦ instances de clusters virtuels sur infrastructure serveur CLOUD utilisant les cartes infiniband,
 - ✦ instances de stations de travail virtuelles sur infrastructure de station CLOUD utilisant de vraies cartes PCI (performances).

Constat

NVIDIA GRID vGPU

- ✦ ACCÉLÉRATION GPU DES BUREAUX VIRTUELS
 - ✦ NVIDIA GRID™ est une plateforme de virtualisation graphique qui s'appuie sur les technologies GPU les plus récentes de NVIDIA pour accélérer les systèmes d'applications et de bureaux virtuels afin d'offrir une expérience utilisateur aussi immersive que qualitative à tous vos collaborateurs - des concepteurs aux utilisateurs mobiles en passant par les employés de bureau
- ✦ prix public ~7k€HT (hors licences)

Théorie

VDI : wikipédia ?

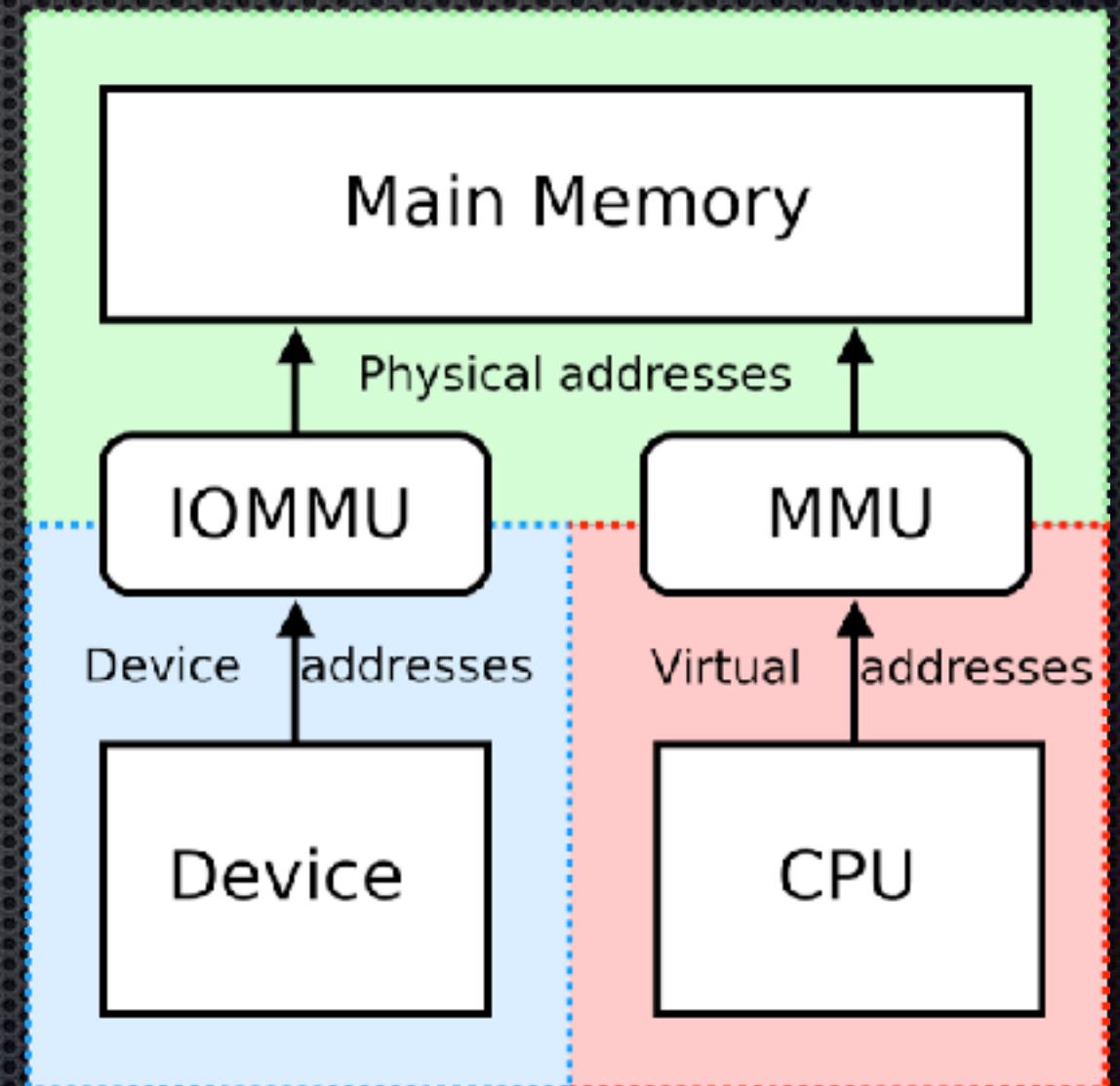
- ✦ La virtual desktop infrastructure (VDI, traduit par infrastructure de bureau virtuel ou machine virtuelle) est un système permettant la dissociation de la machine de l'utilisateur d'une part et de la machine physique d'autre part.

QEMU VGAPassthrough

- ✦ Votre CPU doit supporté la virtualisation et la technologie IOMMU
- ✦ Vous devez avoir deux GPU distincts:
 - ✦ 1 pour le host
 - ✦ 1 pour la VM guest
- ✦ la VM guest va sortir son affichage directement sur la carte graphique qui lui est associée (non visible sur l'host)
 - ✦ 1 écran pour le host
 - ✦ 1 écran pour la VM guest
- ✦ Les périphériques d'interaction utilisateur doivent être dédiés à la VM guest

IOMMU : wikipédia ?

- In computing, an input-output memory management unit (IOMMU) is a memory management unit (MMU) that connects a direct-memory-access-capable (DMA-capable) I/O bus to the main memory. Like a traditional MMU, which translates CPU-visible virtual addresses to physical addresses, the IOMMU maps device-visible virtual addresses (also called device addresses or I/O addresses in this context) to physical addresses. Some units also provide memory protection from faulty or malicious devices.
- An example IOMMU is the graphics address remapping table (GART) used by AGP and PCI Express graphics cards on Intel Architecture and AMD computers.



Virtual Function I/O (VFIO)

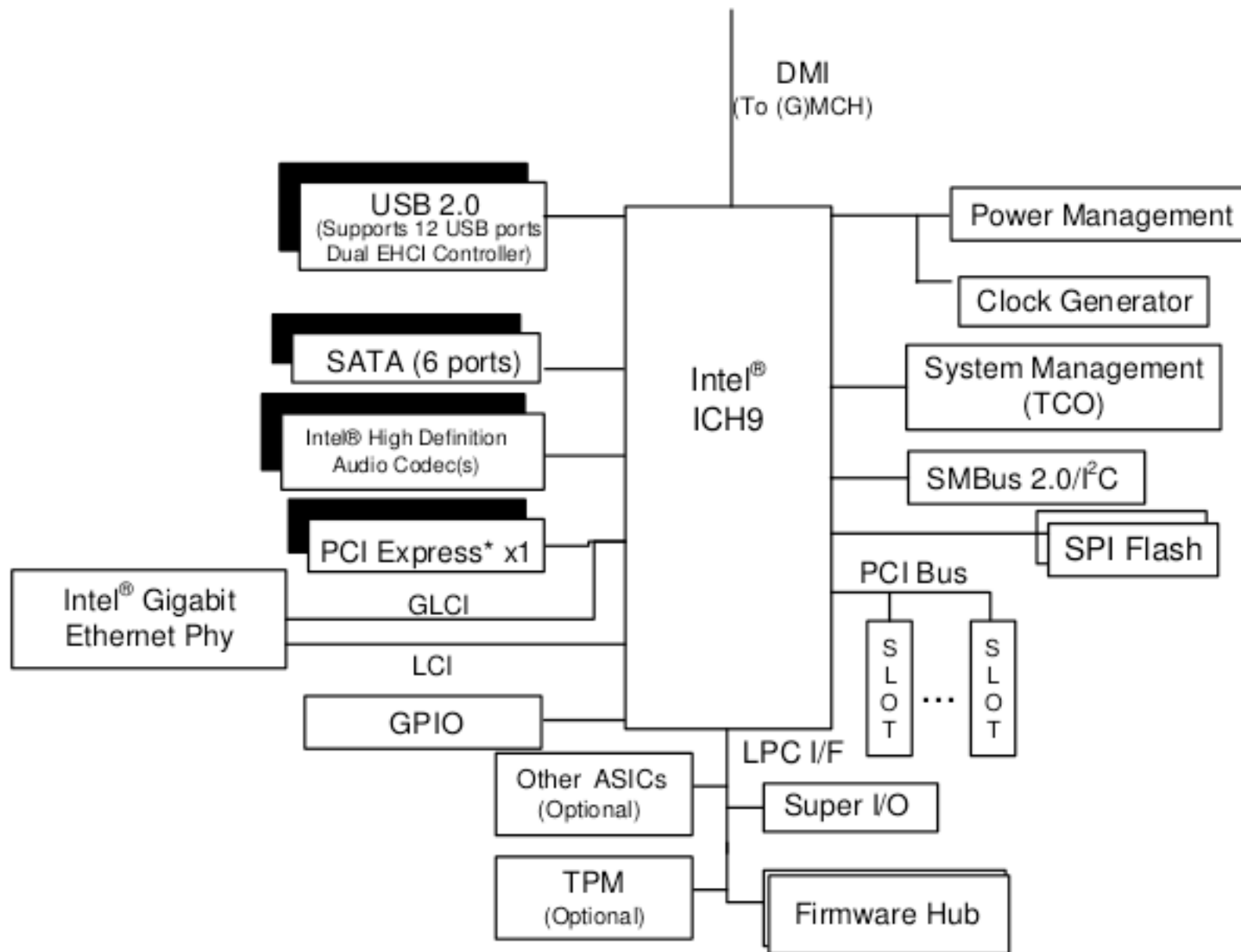
- Les systèmes modernes fournissent des mécanismes de redéfinition d'interruption et d'accès mémoire qui s'assurent que les périphériques fonctionnent dans le périmètre qui leur a été fixé
 - AMD-Vi
 - Intel VT-d,
 - POWER systems avec Partitionable Endpoints (PEs)
 - Freescale PAMU
- Le pilote VFIO est un framework IOMMU permettant l'accès direct aux périphériques
 - sûr
 - sans privilege
 - pilotes en espace utilisateur

Machine type

```
david_delavennat — -bash — 80x37
[station40:~ david_delavennat$ qemu-system-x86_64 -M ?
Supported machines are:
pc                Standard PC (i440FX + PIIX, 1996) (alias of pc-i440fx-2.5)
pc-i440fx-2.5     Standard PC (i440FX + PIIX, 1996) (default)
pc-i440fx-2.4     Standard PC (i440FX + PIIX, 1996)
pc-i440fx-2.3     Standard PC (i440FX + PIIX, 1996)
pc-i440fx-2.2     Standard PC (i440FX + PIIX, 1996)
pc-i440fx-2.1     Standard PC (i440FX + PIIX, 1996)
pc-i440fx-2.0     Standard PC (i440FX + PIIX, 1996)
pc-i440fx-1.7     Standard PC (i440FX + PIIX, 1996)
pc-i440fx-1.6     Standard PC (i440FX + PIIX, 1996)
pc-i440fx-1.5     Standard PC (i440FX + PIIX, 1996)
pc-i440fx-1.4     Standard PC (i440FX + PIIX, 1996)
pc-1.3           Standard PC (i440FX + PIIX, 1996)
pc-1.2           Standard PC (i440FX + PIIX, 1996)
pc-1.1           Standard PC (i440FX + PIIX, 1996)
pc-1.0           Standard PC (i440FX + PIIX, 1996)
pc-0.15          Standard PC (i440FX + PIIX, 1996)
pc-0.14          Standard PC (i440FX + PIIX, 1996)
pc-0.13          Standard PC (i440FX + PIIX, 1996)
pc-0.12          Standard PC (i440FX + PIIX, 1996)
pc-0.11          Standard PC (i440FX + PIIX, 1996)
pc-0.10          Standard PC (i440FX + PIIX, 1996)
q35              Standard PC (Q35 + ICH9, 2009) (alias of pc-q35-2.5)
pc-q35-2.5       Standard PC (Q35 + ICH9, 2009)
pc-q35-2.4       Standard PC (Q35 + ICH9, 2009)
pc-q35-2.3       Standard PC (Q35 + ICH9, 2009)
pc-q35-2.2       Standard PC (Q35 + ICH9, 2009)
pc-q35-2.1       Standard PC (Q35 + ICH9, 2009)
pc-q35-2.0       Standard PC (Q35 + ICH9, 2009)
pc-q35-1.7       Standard PC (Q35 + ICH9, 2009)
pc-q35-1.6       Standard PC (Q35 + ICH9, 2009)
pc-q35-1.5       Standard PC (Q35 + ICH9, 2009)
pc-q35-1.4       Standard PC (Q35 + ICH9, 2009)
isapc            ISA-only PC
none            empty machine
station40:~ david_delavennat$
```

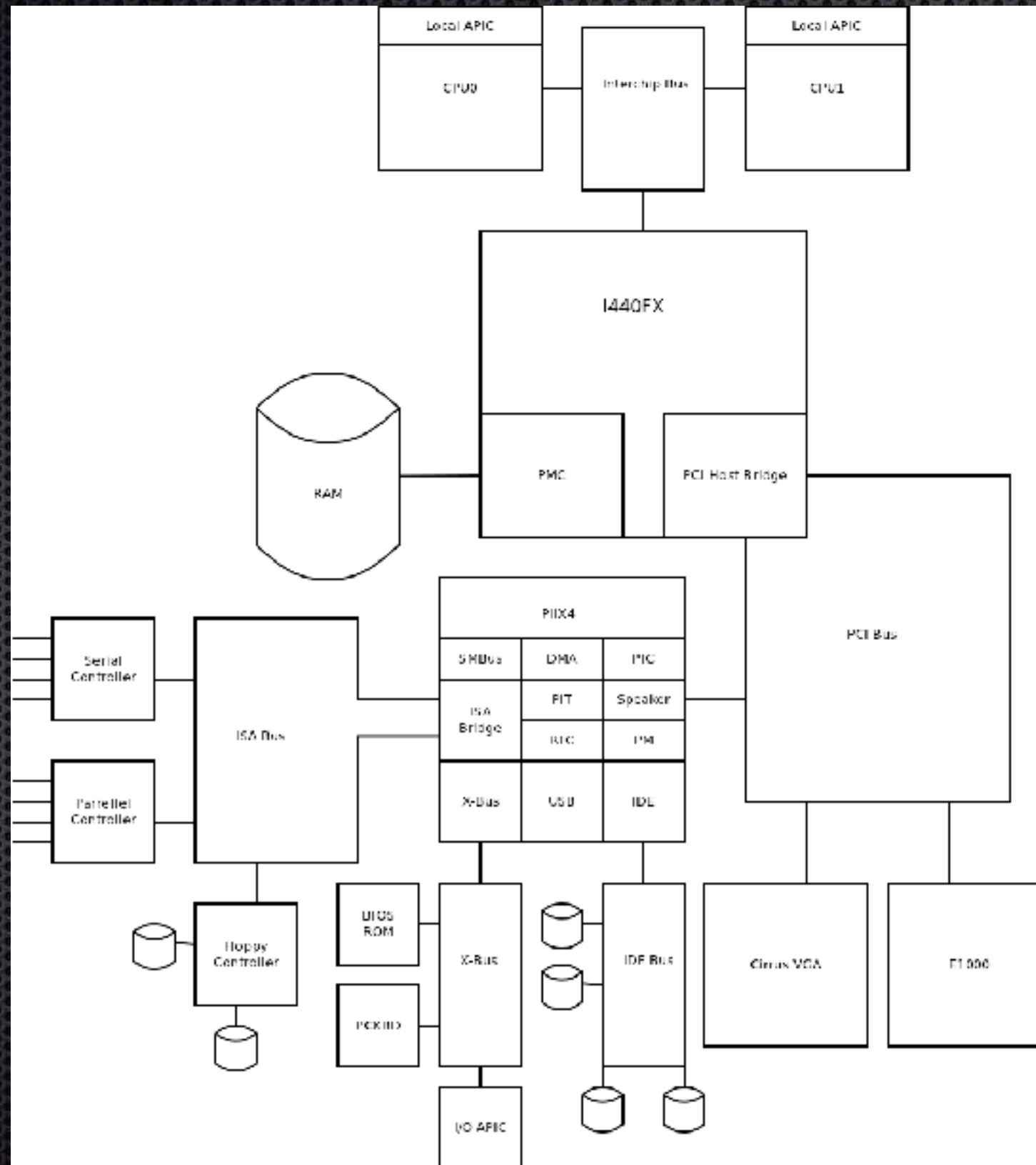
better support of
PCI-E passthrough
since ICH9 uses a
PCI-E bus whereas
the I440FX only
supports a PCI
bus.

Machine : Q35



- controller AHCI
- controller USB
- réseau ethernet
- carte son
- bus PCI
- bus PCI-E
- bus LPC
- pas de bus ISA

Machine : i440fx



BIOS type

- ✧ Legacy (SEABIOS)
 - ✧ pas réussi à le faire fonctionner
- ✧ UEFI (TIANOCORE)
 - ✧ Projet OVMF
 - ✧ Pas possible pour l'instant de faire du PXE
 - ✧ oui mais... en fait si! ;-) (se compiler une rom iPXE)

Matériel

DELL PowerEdge T130



- Marché MATINFO3 ~560€ TTC
 - Intel® Xeon® E3-1220 v5 à 3.0GHz, 8Mo cache, 4C /8T, turbo (80W)
 - Passage à 16Gb en 4 x 4Go 2133 Mhz UDIMM - Faible Voltage
 - 1 disque dur 500 Go SATA II
 - On-Board LOM 1 GbE Dual Port (BCM5720 GbE LOM)
 - Passage à carte iDRAC 8 Enterprise
 - Lecteur interne SATA DVD+/-RW

Deport de la console

DELL PowerEdge T130



- ❖ `racadm -r <drac_ip> -u root -p calvin get system.thermalsettings`
 - ❖ `ThirdPartyPCIFanResponse=Enabled`
- ❖ `racadm set System.ThermalSettings.ThirdPartyPCIFanResponse 0`
 - ❖ `ThirdPartyPCIFanResponse=Disabled`

Passe en mode silence

Carte vidéo

- Host

- Carte video Matrox G200 intégrée

- Guest

- Radeon HD 5450 || Radeon HD 6450 : 30€HT

- Carte vidéo NVIDIA ?

- le pilote Windows ne s'installe pas en environnement virtualisé (kvm=off) -> NVIDIA veut vous faire acheter des cartes GRID !!!!

Carte son

- ✦ Host

- ✦ Pas de carte son intégrée au T130

- ✦ Guest

- ✦ Carte PCIe Creative Sound Blaster Audigy FX : 30€HT
 - ✦ Carte USB : « glitch » (brusque augmentation de tension provenant du dysfonctionnement d'un dispositif électrique ou électronique)
 - ✦ Son via le HDMI : son parfait mais HDMI à déconseiller

Clavier et souris

- Clavier

- idrac (host)
- USB (guest)

- Souris

- idrac (host)
- USB (guest)

QEMU : `—usb-device`

ou

USB PASSTHRU :
Chip NEC RENESAS
Carte silverstone EC04-E

Attention au connecteur interne
19pin dual port !!!

Systeme d'exploitation

- Host

- Centos 7
- Fedora 25

- Guest

- Centos 7
- Fedora 25
- Ubuntu 16.04, 16.10
- Windows 10 (ok avec carte ATI)

Fabrication des images

- ✧ packer, systemd-nspawn
 - ✧ construction offline
 - ✧ possibilité de repartir d'une *couche* précédente
 - ✧ intégration continue

Liens

- ✦ https://fr.wikipedia.org/wiki/Virtual_desktop_infrastructure
- ✦ <http://www.nvidia.fr/grid/>
- ✦ VFIO
 - ✦ <https://www.kernel.org/doc/Documentation/vfio.txt>
 - ✦ <https://github.com/andre-richter/rVFIO>
- ✦ Passthrough
 - ✦ <https://wiki.debian.org/VGAPassthrough>
 - ✦ https://wiki.archlinux.org/index.php/QEMU#Starting_QEMU_virtual_machines_on_boot
 - ✦ https://wiki.archlinux.org/index.php/PCI_passthrough_via_OVMF#Using_vfio-pci
 - ✦ <https://gist.github.com/cuibonobo/d354440fecdd37c35ecd> (VGA Passthrough on virtual machines in CentOS 7)
- ✦ <https://www.packer.io/docs/>

- ✧ Architecture matérielle

- ✧ <http://wiki.qemu-project.org/Features/Q35>

- ✧ UEFI / PXE

- ✧ <https://github.com/openSUSE/kiwi/wiki/Setup-PXE-boot-with-EFI-Using-GRUB2>
 - ✧ <https://doc.rogerwhittaker.org.uk/ipxe-installation-and-EFI/>

- ✧ CLOUD / VFIO

- ✧ http://docs.opennebula.org/5.0/deployment/open_cloud_host_setup/pci_passthrough.html

- ✧ Linux / USB

- ✧ <http://www.tyan.com/manuals/LinuxUSB3DriverPortingManual.pdf>

- ✧ QEMU audio

- ✧ http://nairobi-embedded.org/qemu_audio.html

- ✦ <https://github.com/tannevaled/qemu-contrib>
 - ✦ ROM iPXE pour QEMU/UEFI
 - ✦ Script sprinkle de configuration d'un HOST VDI
 - ✦ Script sprinkle de configuration d'un HOST OpenNebula

LOONEY TUNES

“Go home, Folks!”

A WARNER BROS. CARTOON

QEMU Tips

```
qemu-system-x86_64 -machine ?
```

```
Supported machines are:
```

```
pc                Standard PC (i440FX + PIIX, 1996) (alias of pc-i440fx-2.0)
```

```
pc-i440fx-2.0      Standard PC (i440FX + PIIX, 1996) (default)
```

```
...
```

```
q35                Standard PC (Q35 + ICH9, 2009) (alias of pc-q35-2.0)
```

```
pc-q35-2.0         Standard PC (Q35 + ICH9, 2009)
```

```
...
```

```
# qemu-system-x86_64 -net nic,model=? -nographic /dev/null
```

```
qemu: Supported NIC models: ne2k_pci,i82551,i82557b,i82559er,rtl8139,e1000,pcnet,virtio
```

```
# qemu-system-x86_64 -device ?
```

```
# qemu-system-x86_64 -device nec-usb-xhci,?
```

```
nec-usb-xhci.msi=on/off
```

```
nec-usb-xhci.msix=on/off
```

```
nec-usb-xhci.intrs=uint32
```

```
nec-usb-xhci.slots=uint32
```

```
nec-usb-xhci.p2=uint32
```

```
nec-usb-xhci.p3=uint32
```

```
nec-usb-xhci.addr=pci-devfn
```

```
nec-usb-xhci.romfile=str
```

```
nec-usb-xhci.rombar=uint32
```

```
nec-usb-xhci.multifunction=on/off
```

```
nec-usb-xhci.command_serr_enable=on/off
```