Virtualization using Xen

References:

1. The definitive guide to the Xen hypervisor

2. XenWiki : http://wiki.xen.org

3. Xen and The art of virtualization
http://dl.acm.org/citation.cfm?id=945462

Contact: {deba,puru}@cse.iitb.ac.in

Xen Architecture

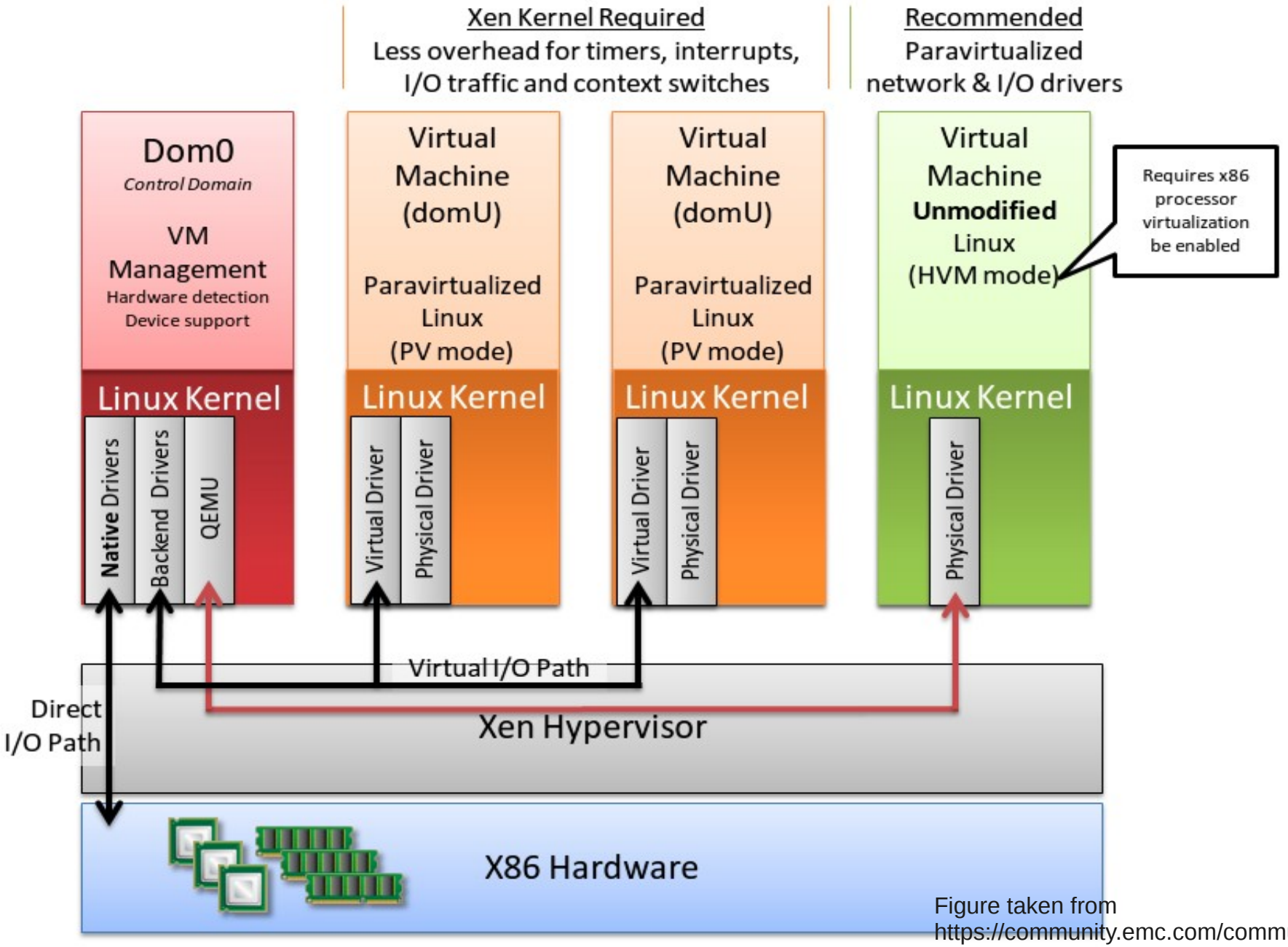


Figure taken from
https://community.emc.com/community/connec
verything_oracle/blog/2011/12/15/oracle-vm-
server-xen-architecture-knowledge-is-power

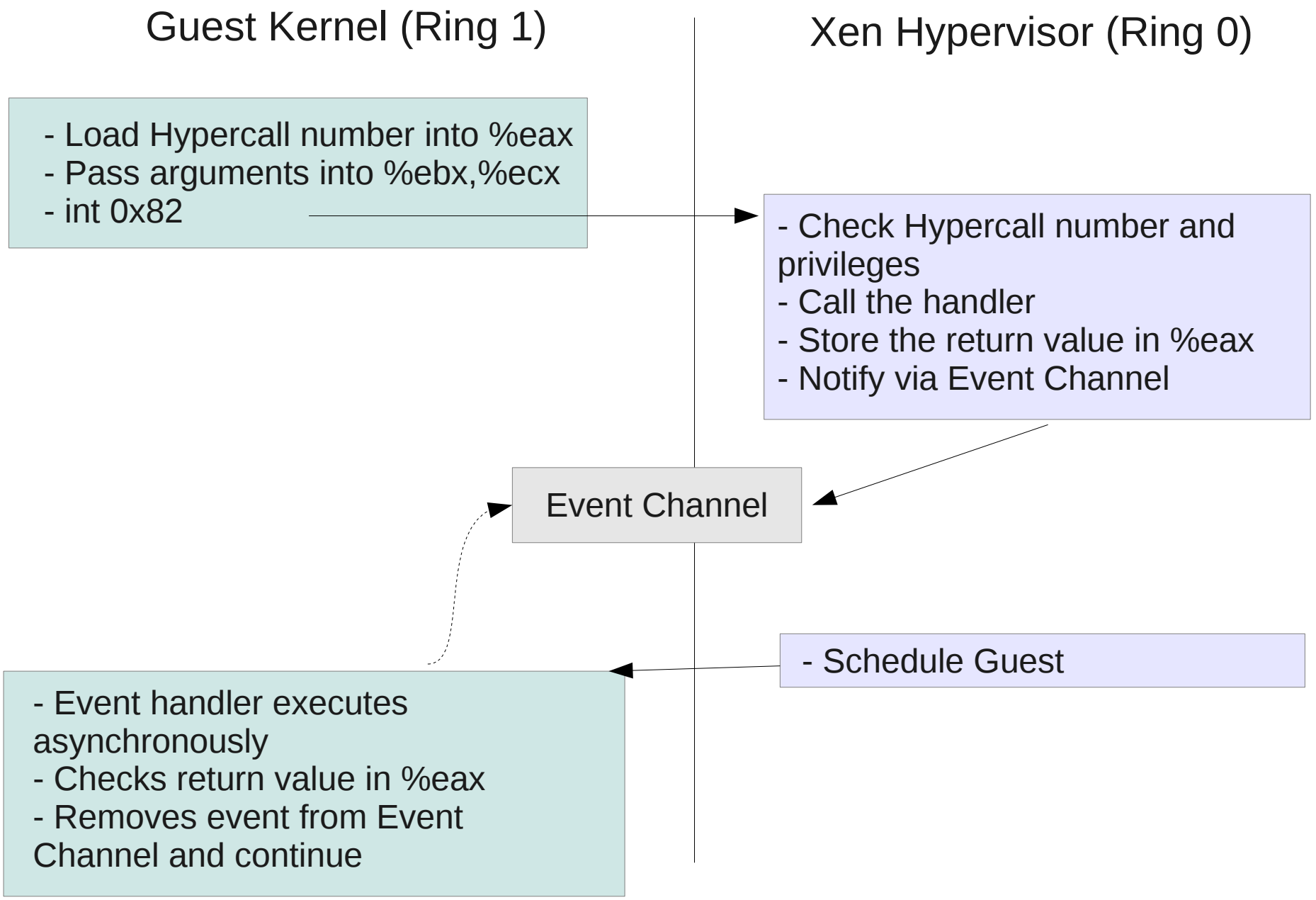
Virtualization techniques

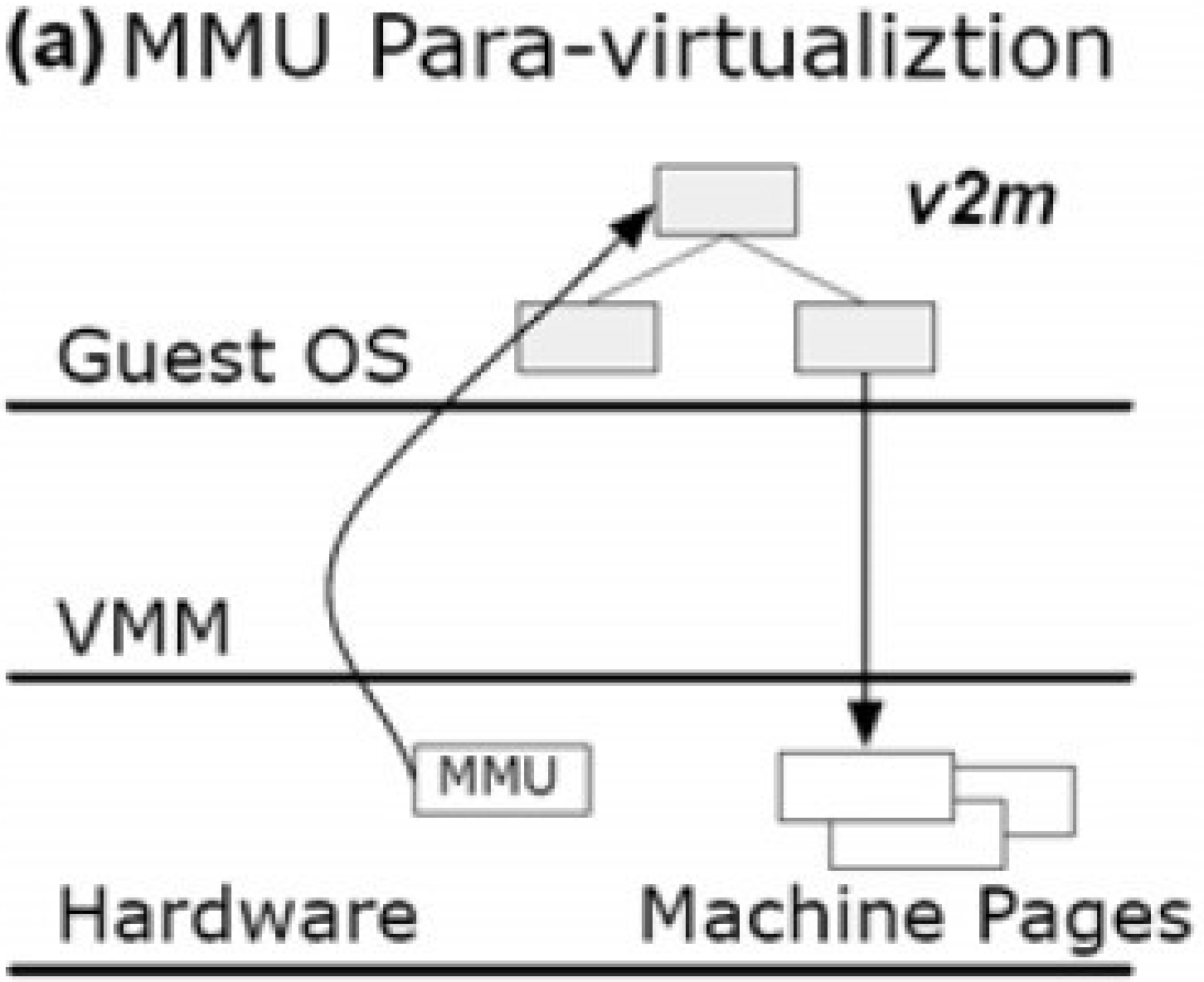
- Para-virtualization
  - VMM runs in ring-0 and guest kernel in ring-1
  - Guest modifications at compile (design) time
  - Hyper-calls to carry out sensitive tasks
  - Use hypercalls to optimize operations
  - Split device model
- Hardware assisted virtualization (HVM)
  - QEMU stub used for device emulation.
  - VMCALL to issue hypercalls
  - Guest kernel modules for optimized device operations.

Hypercalls in Xen

- Privileged Guest operation (Shut down)
  - Normal guest OS would execute HLT on CPU.
  - Hypervisor should be notified that guest is shutting down.
  - sched_op( ) call comes to rescue.
- Shared resource access (Global Desc. Table)
  - Normal method of loading a GDT : lgdt
  - GDT register is shared between guests.
  - set_gdt() is the Xen alternative.
- Administrative Operation (Create Domain)
  - Domain-0 has all the user APIs for domain mgt.
  - Xen provides hypercalls to be invoked from Domain-0.

Hypercall Implementation(x86)




(a) MMU Para-virtualiztion
v2m
Guest OS
VMM
MMU
Hardware
Machine Pages

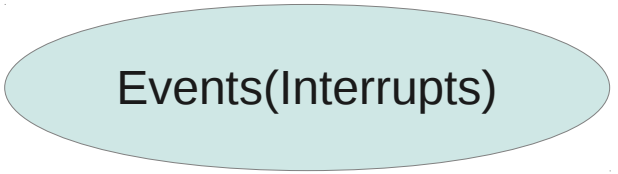
Direct mapping and Isolation

- Guest registration of new Page Table(s)
  - Pages containing page tables must be of special type (PAGE META type)
  - Contents of the page
    - Should point to frames that belong to the guest
    - If any entry is of type PAGE META, they must be read-only
- Guest modification of Page table entries
  - Has to be done through a hypercall.
  - The page fault handler is just the special case.

Memory operation hypercalls

mmu_update(req[ ],count,success_count)	Update the page table for a guest VM. The updates can be - normal MMU meta data update - Machine to physical mapping update - misc commands (cr3 access,tlb flush etc)
update_va_mapping(vaddr, newpte,flags)	Update the virtual address mapping of 'vaddr' in the MMU page table to point to the new physical address given in 'newpte'
memory_op(command, args)	Many operations can be done. Some example commands are - increase/decrease mem reservation - memory exchange - Get machine-physical mapping table address

Event Handling

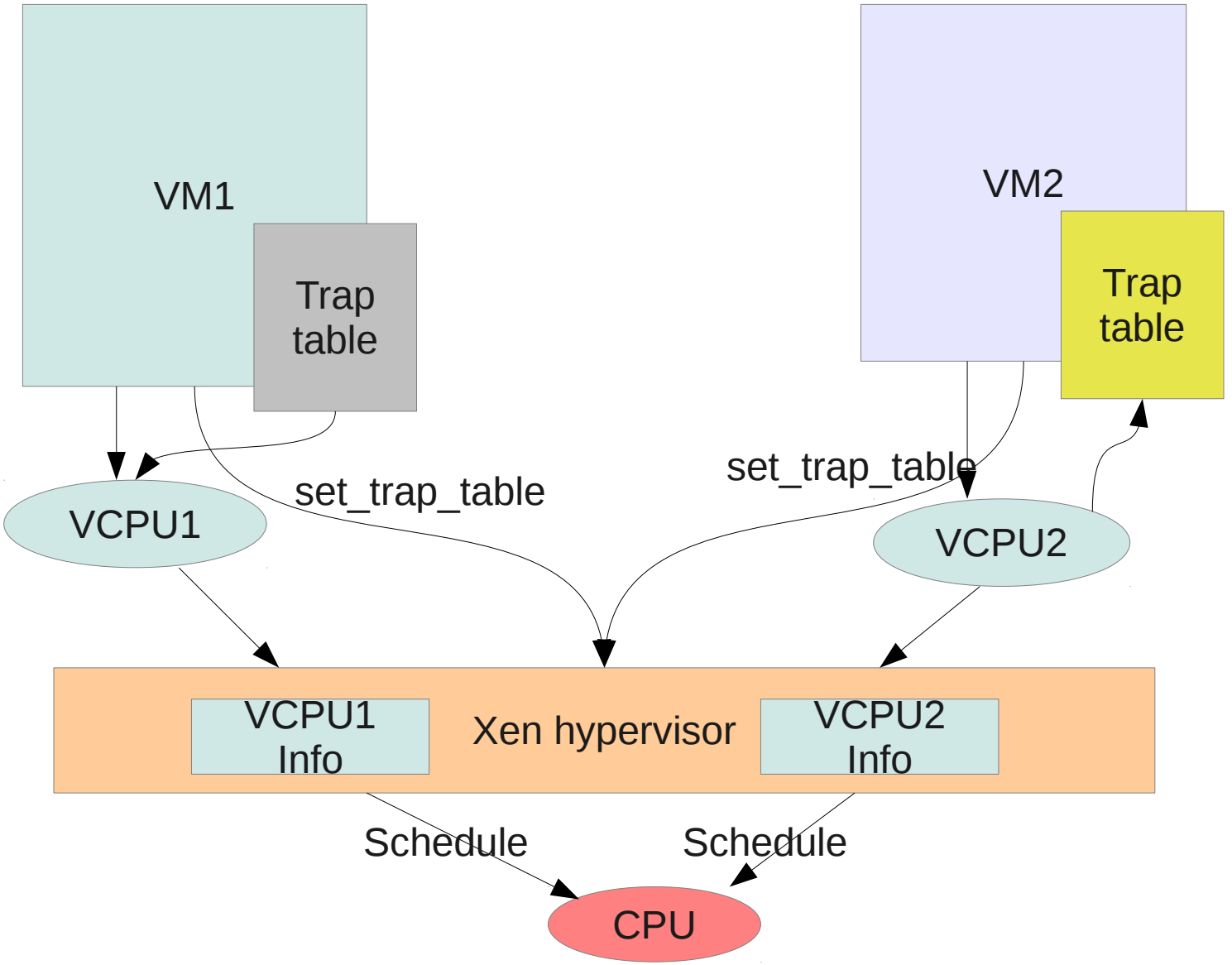


- Total 256 interrupts in x86.
- OS provides IDT for interrupt handling

-Occurs because of an instruction execution
- For virtualization traps are easier to handle.

- Occurs because of external event.
- Interrupts are difficult to implement.

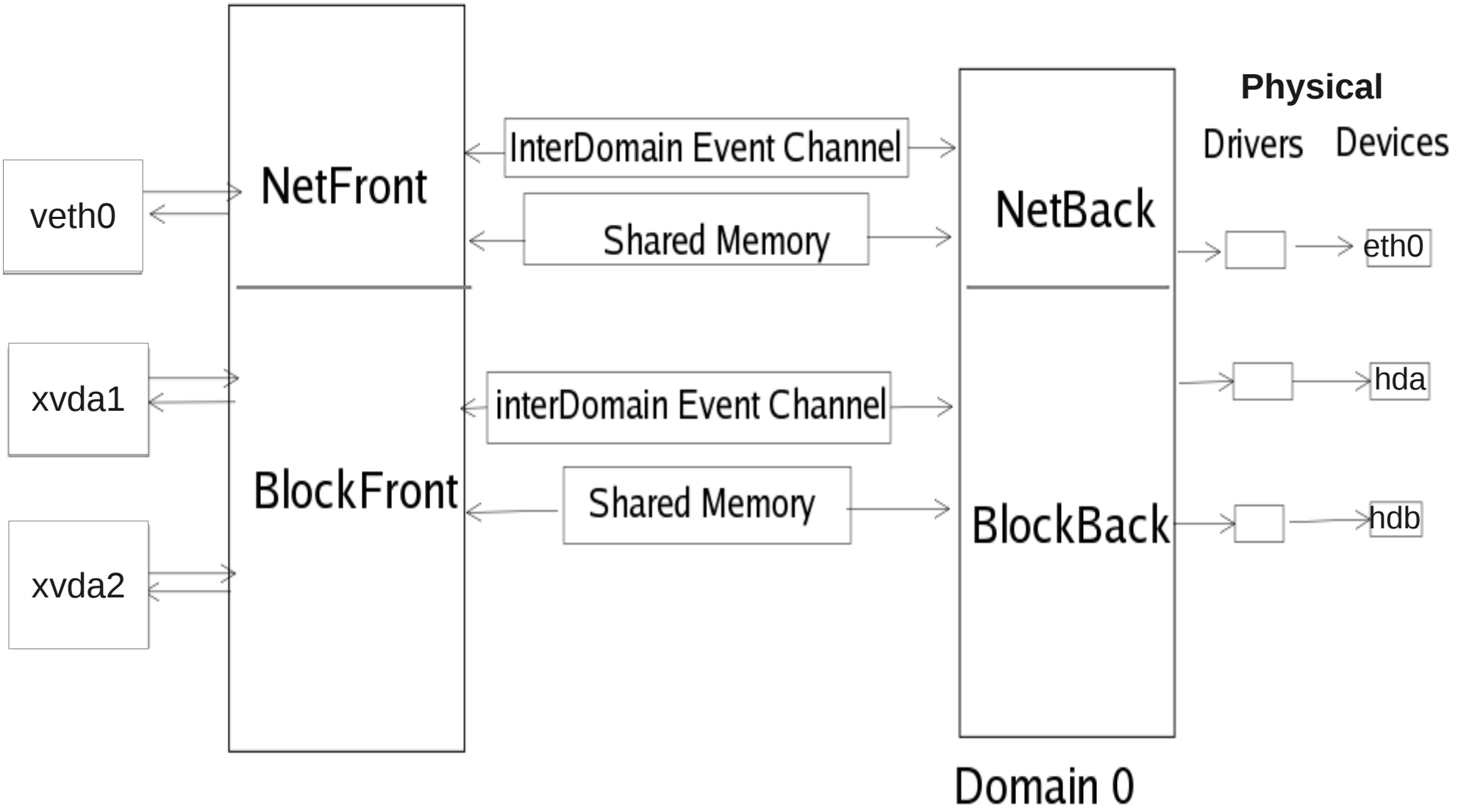
Trap handling in Xen



Device I/O (Bottom processing)

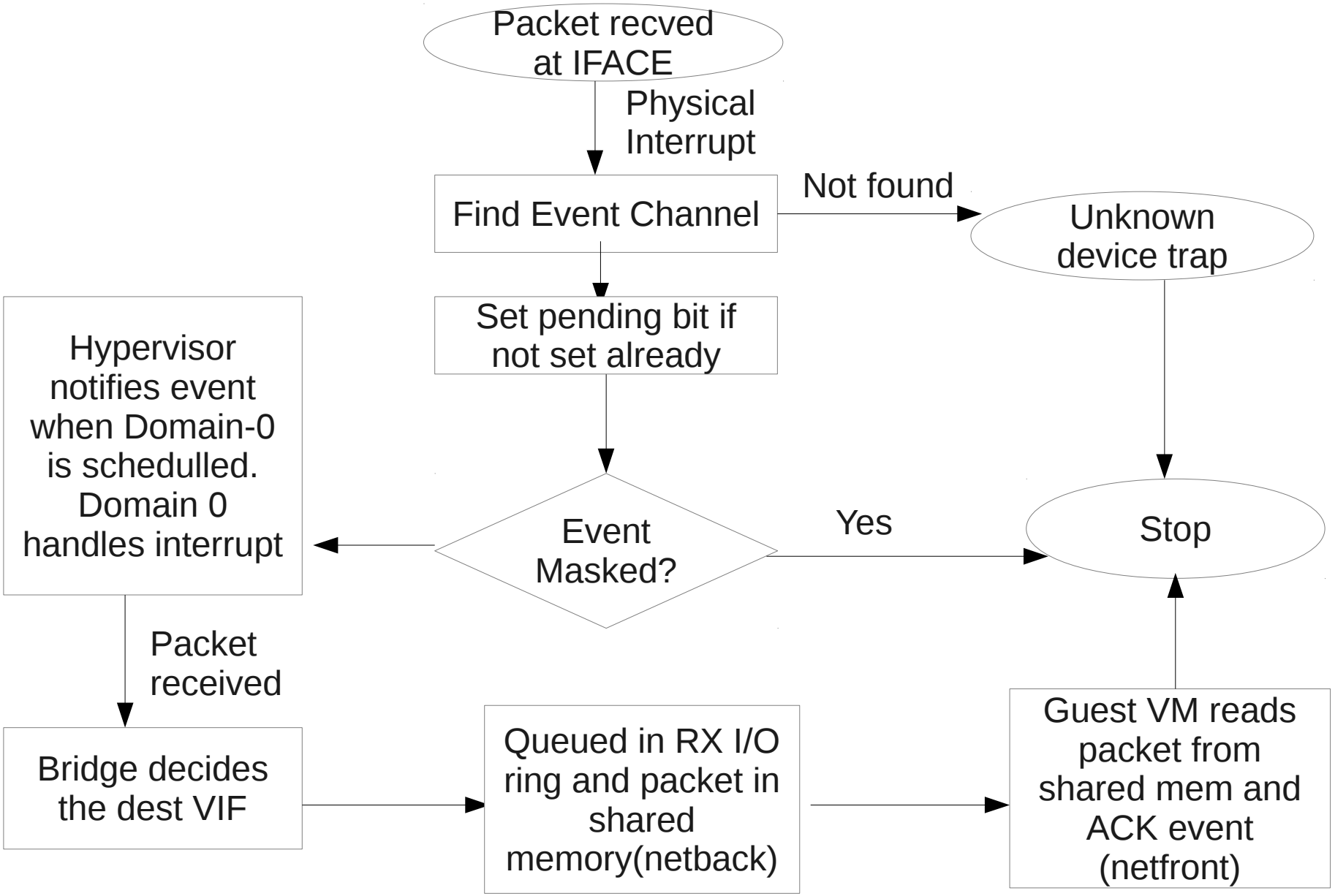
- Domain-0 registers *event channels* for each physical device.
- Hypervisor receives all the interrupt first.
- Except timer and serial interrupts all others are passed to Domain-0.
- Inserts a *pending event* in the corresponding event channel.
- Interdomain event handling managed by domain-0 (using I/O rings).
- Refer event_channel_op for more...

Split Drivers Diagram



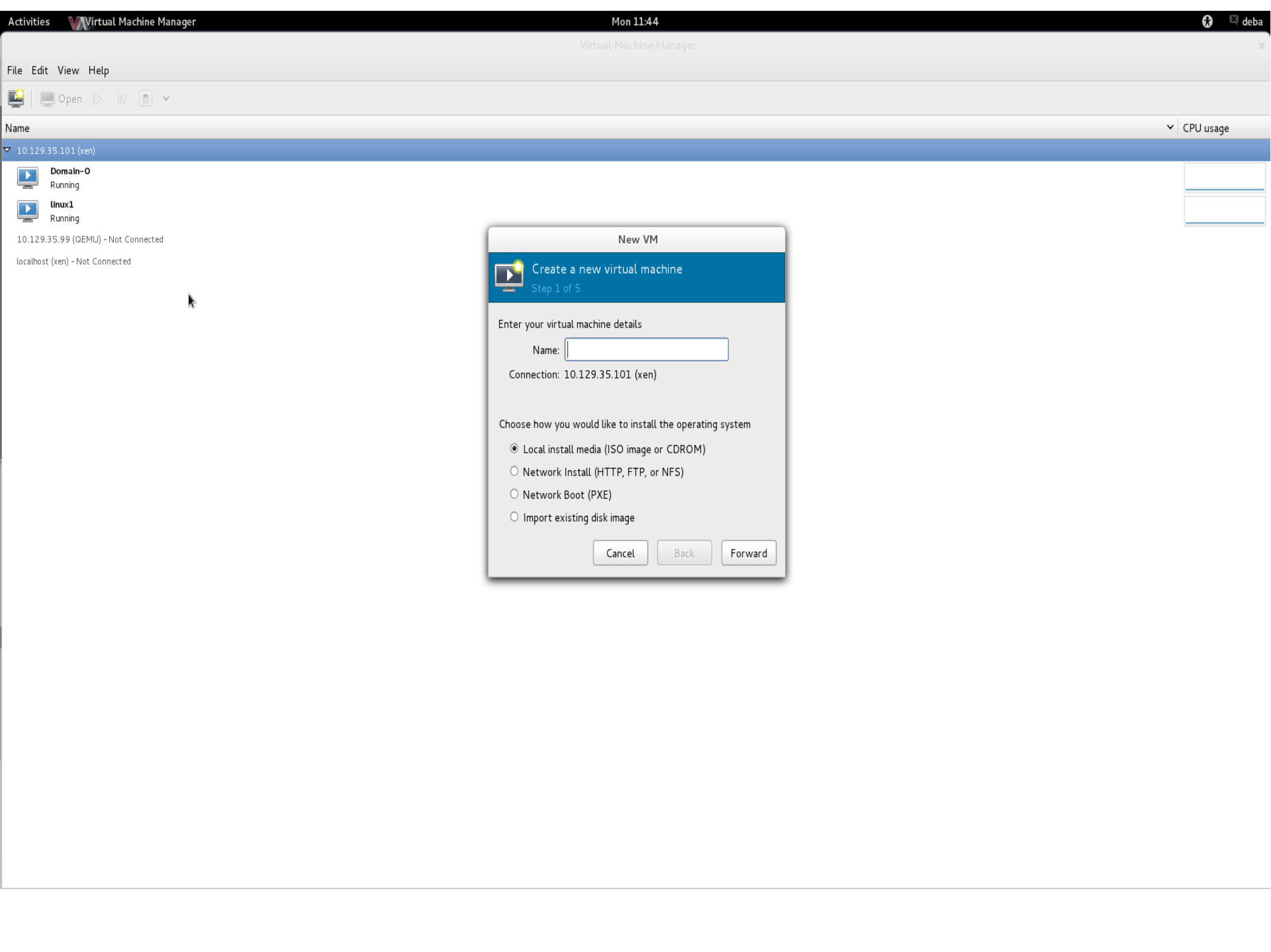
[2] XenWiki

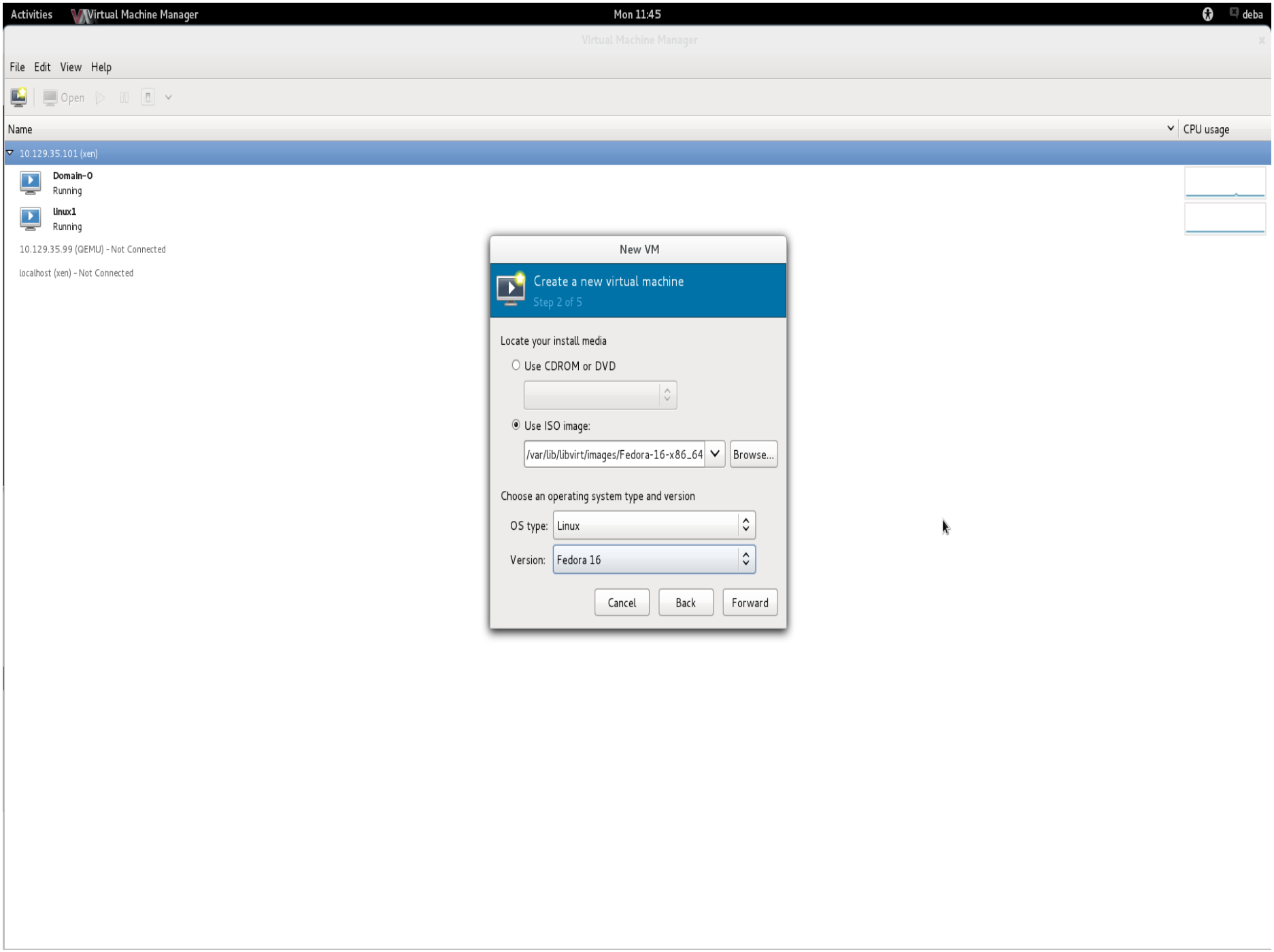
Network I/O (Guest receive a packet)

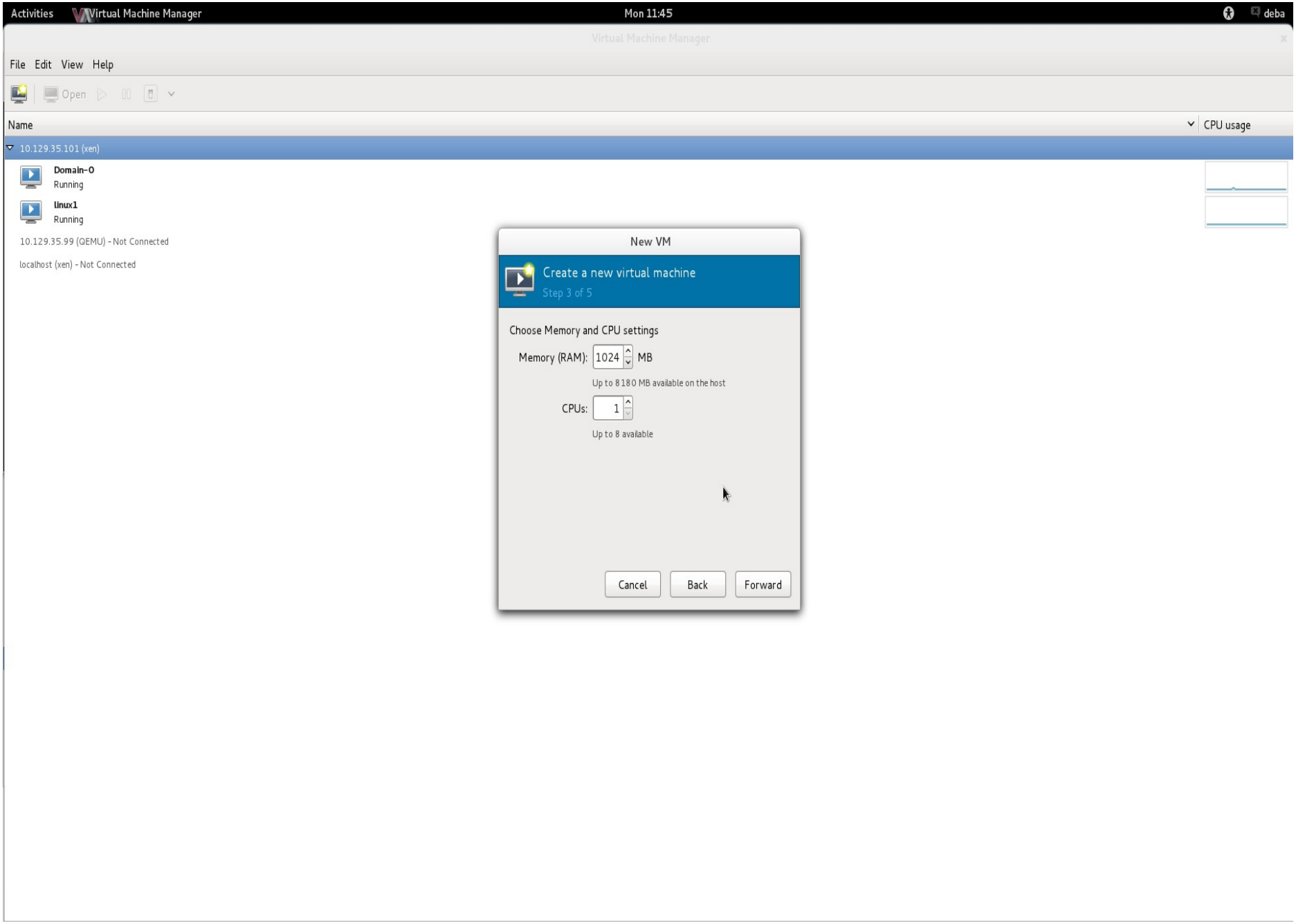


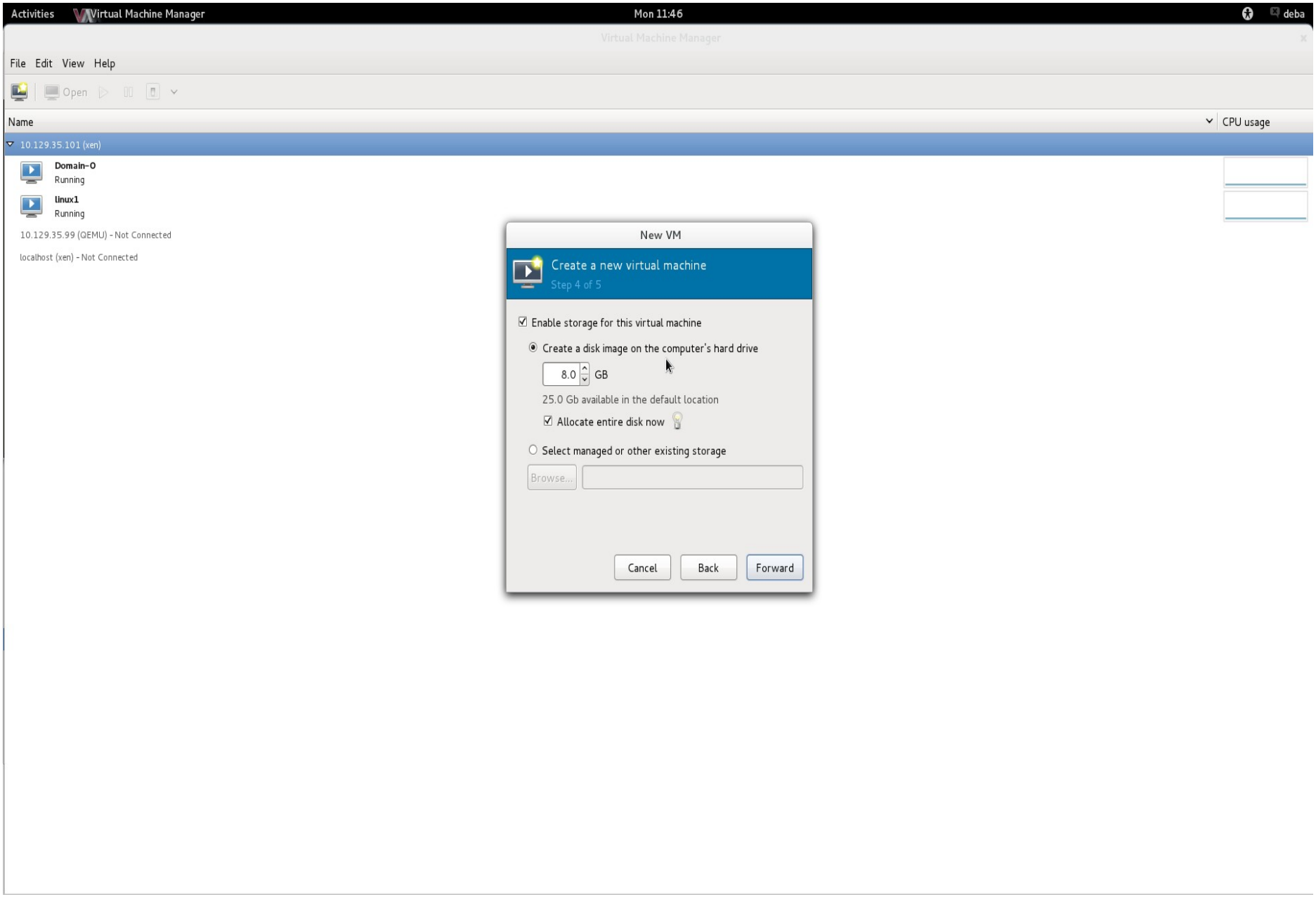
Xen Installation

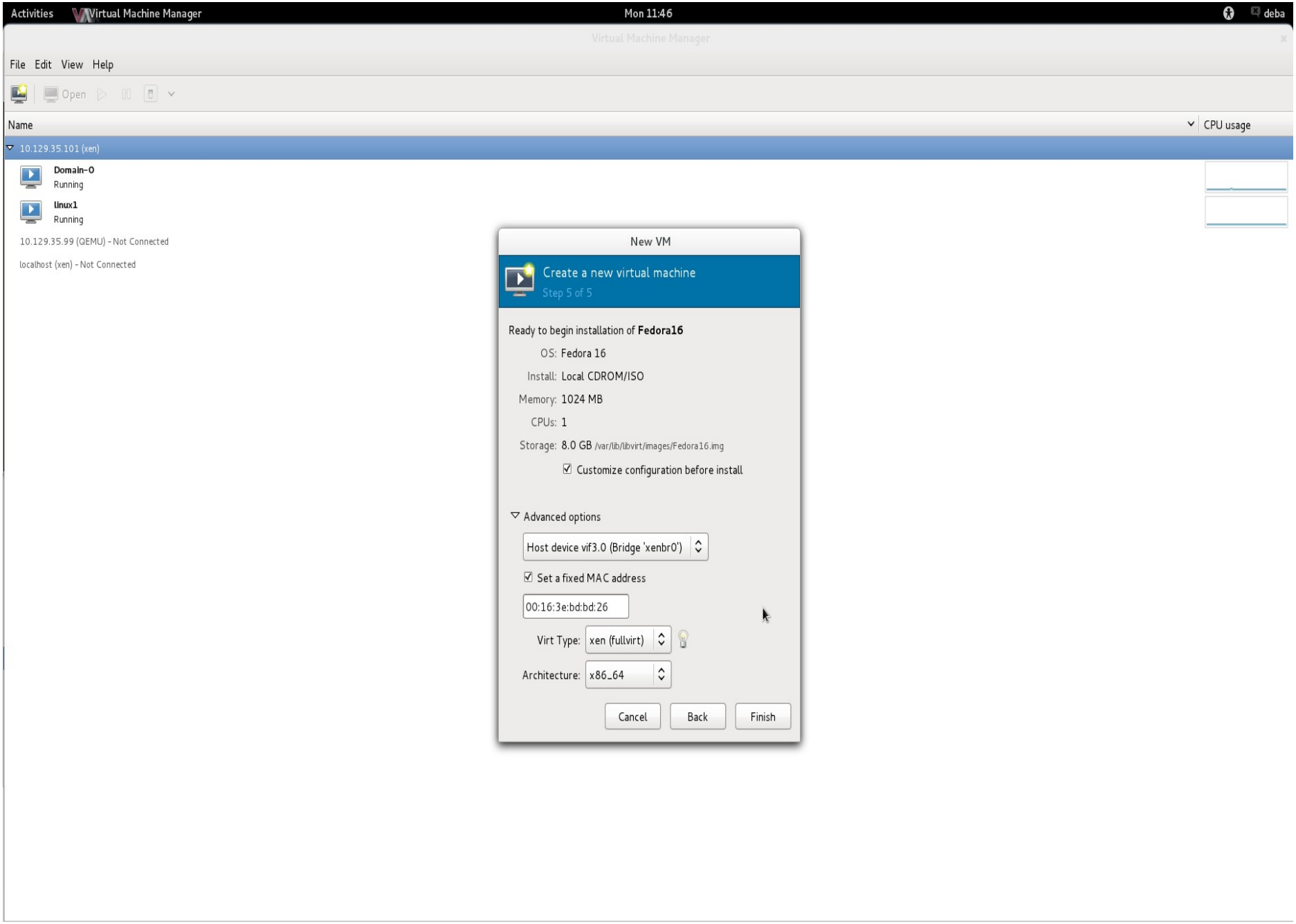
- Install Fedora 16, Disable SELinux
- Install Xen using yum repos
- systemctl disable NetworkManager.service || systemctl restart network.service
- chkconfig network on
- Create a bridge that has your eth0 as a member.
- Restart the VM and check xl info
- Install libvirt and virt-manager
- Ref [http://wiki.xen.org/wiki/Fedora_Host_Installation]

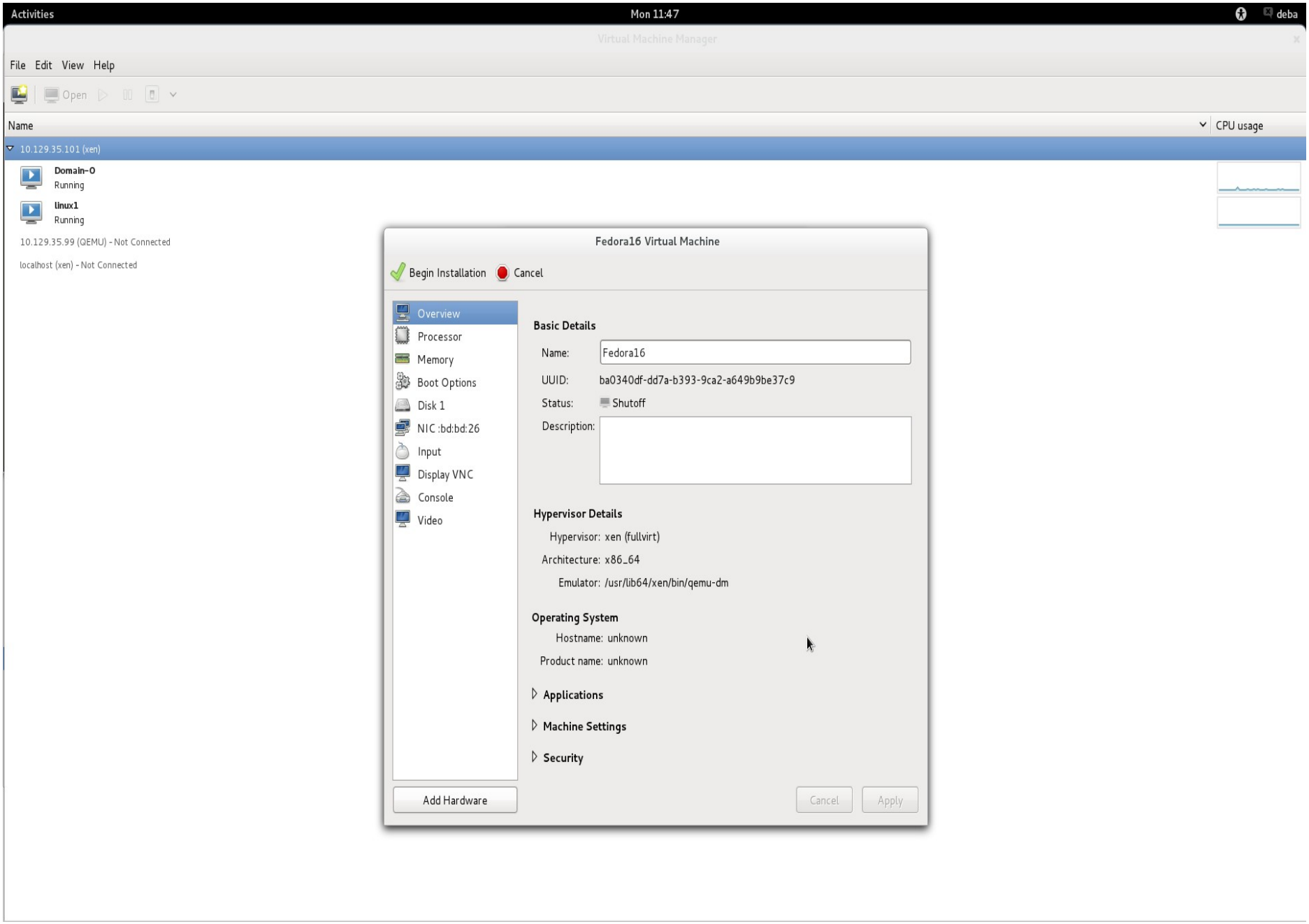

Activities
Virtual Machine Manager
Mon 11:44
deba
Virtual Machine Manager
File Edit View Help
Open
Name
CPU usage
10.129.35.101 (xen)
Domain-0
Running
linux1
Running
10.129.35.99 (QEMU) - Not Connected
localhost (xen) - Not Connected
New VM
Create a new virtual machine
Step 1 of 5
Enter your virtual machine details
Name:
Connection: 10.129.35.101 (xen)
Choose how you would like to install the operating system
Local install media (ISO image or CDROM)
Network Install (HTTP, FTP, or NFS)
Network Boot (PXE)
Import existing disk image
Cancel
Back
Forward

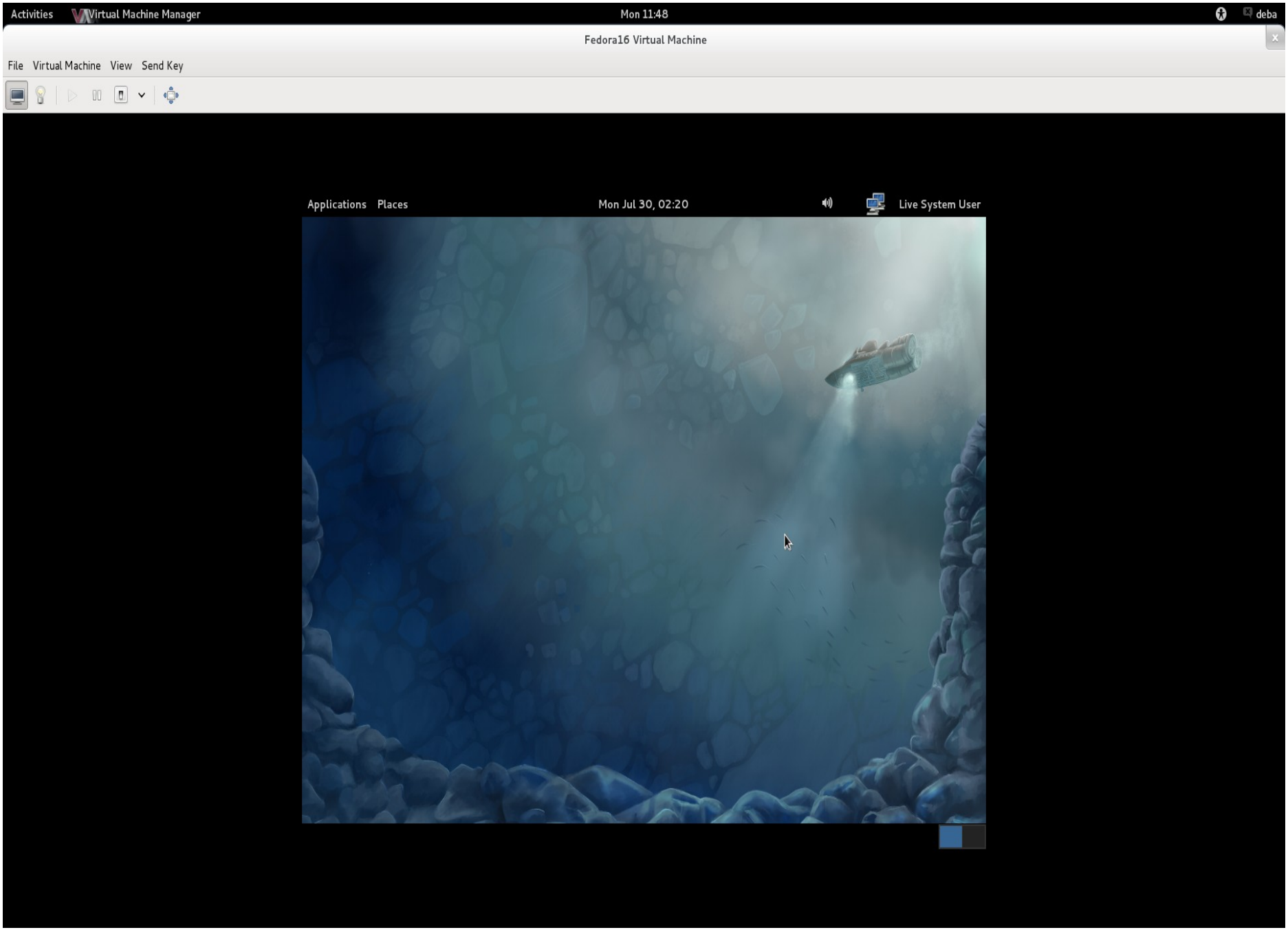

Activities
Virtual Machine Manager
Mon 11:45
deba
Virtual Machine Manager
File Edit View Help
Open
Name
CPU usage
10.129.35.101 (xen)
Domain-0
Running
linux1
Running
10.129.35.99 (QEMU) - Not Connected
localhost (xen) - Not Connected
New VM
Create a new virtual machine
Step 2 of 5
Locate your install media
Use CDROM or DVD
Use ISO image:
/var/lib/libvirt/images/Fedora-16-x86_64
Browse...
Choose an operating system type and version
OS type:
Linux
Version:
Fedora 16
Cancel
Back
Forward


Activities
Virtual Machine Manager
Mon 11:45
deba
Virtual Machine Manager
File Edit View Help
Open
Name
CPU usage
10.129.35.101 (xen)
Domain-0
Running
linux1
Running
10.129.35.99 (QEMU) - Not Connected
localhost (xen) - Not Connected
New VM
Create a new virtual machine
Step 3 of 5
Choose Memory and CPU settings
Memory (RAM): 1024 MB
Up to 8180 MB available on the host
CPUs: 1
Up to 8 available
Cancel
Back
Forward


Activities
Virtual Machine Manager
Mon 11:46
deba
Virtual Machine Manager
File Edit View Help
Open
Name
CPU usage
10.129.35.101 (xen)
Domain-0
Running
linux1
Running
10.129.35.99 (QEMU) - Not Connected
localhost (xen) - Not Connected
New VM
Create a new virtual machine
Step 4 of 5
Enable storage for this virtual machine
Create a disk image on the computer's hard drive
8.0
GB
25.0 Gb available in the default location
Allocate entire disk now
Select managed or other existing storage
Browse...
Cancel
Back
Forward


Activities
Virtual Machine Manager
Mon 11:46
deba
Virtual Machine Manager
File Edit View Help
Open
Name
CPU usage
10.129.35.101 (xen)
Domain-0
Running
linux1
Running
10.129.35.99 (QEMU) - Not Connected
localhost (xen) - Not Connected
New VM
Create a new virtual machine
Step 5 of 5
Ready to begin installation of Fedora16
OS: Fedora 16
Install: Local CDROM/ISO
Memory: 1024 MB
CPUs: 1
Storage: 8.0 GB /var/lib/libvirt/images/Fedora16.img
Customize configuration before install
Advanced options
Host device vif3.0 (Bridge 'xenbr0')
Set a fixed MAC address
00:16:3e:bd:bd:26
Virt Type: xen (fullvirt)
Architecture: x86_64
Cancel
Back
Finish


Activities
Mon 11:47
deba
Virtual Machine Manager
File Edit View Help
Open
Name
CPU usage
10.129.35.101 (xen)
Domain-0
Running
linux1
Running
10.129.35.99 (QEMU) - Not Connected
localhost (xen) - Not Connected
Fedora16 Virtual Machine
Begin Installation
Cancel
Overview
Processor
Memory
Boot Options
Disk 1
NIC :bd:bd:26
Input
Display VNC
Console
Video
Basic Details
Name: Fedora16
UUID: ba0340df-dd7a-b393-9ca2-a649b9be37c9
Status: Shutoff
Description:
Hypervisor Details
Hypervisor: xen (fullvirt)
Architecture: x86_64
Emulator: /usr/lib64/xen/bin/qemu-dm
Operating System
Hostname: unknown
Product name: unknown
Applications
Machine Settings
Security
Add Hardware
Cancel
Apply


Activities
Virtual Machine Manager
Mon 11:48
deba
Fedora16 Virtual Machine
File Virtual Machine View Send Key
Applications Places
Mon Jul 30, 02:20
Live System User

Some tools

- Xenstore is the common interface between all user level configuration and management and hypervisor.
- Xm (example xm list,xm migrate etc)
- Virsh (Interactive Shell built above libvirt)
- Xl (similar to xm)
- Xentop (performance monitoring)

Misc

- Xen latest stable version is 4.1.2.
- Downloadable from http://xen.org/products/downloads.html
- Compilation is similar to Linux kernel compile.
- Questions?