

Lecture IV Server virtualization

Operating system concepts – Infrastructure concepts for applications

Chris Roets

2025-2026

(based on material from T. Deneut)

(Server) virtualization

Almost every single application is running within a virtual environment.

Why?

A normal server often has 24 or more processor cores and 128GB or more RAM.
Is there a single application that even knows how to address such amounts?



Virtualization

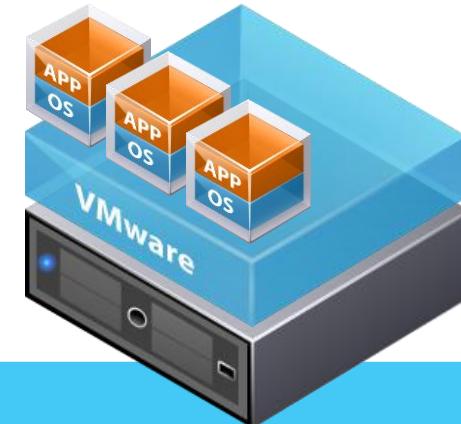
Physical machine

- Difficult to move or copy
 - Bound to specific hardware
- Often has a short lifecycle
- Requires personal contact to upgrade hardware



Virtual machine

- Easy to move and copy:
 - Encapsulated into files
 - Independent of physical hardware
- Easy to manage:
 - Isolated from other virtual machines
 - Insulated from hardware changes
- Provides the ability to support legacy applications
- Allows servers to be consolidated



Virtualization

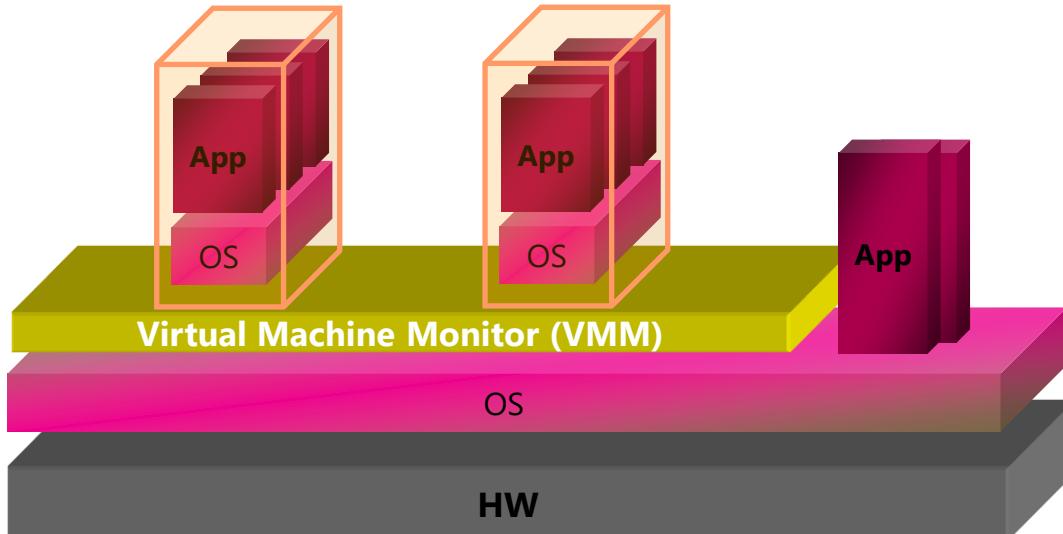
Different Kinds of virtualization

1. **Hypervisor-based** or ***bare metal*** virtualization (or **Type 1**)
 - Software based or hardware assisted (e.g. Intel VT-d, AMD-V)
2. **Hosted** virtualization (or **Type 2**)
 - Using a software inside an OS
3. Container-based virtualization
 - E.g. docker
4. LXC containers
5. ...Desktop, application, data, device ... virtualization
 - E.g. what a sandbox does (and most browsers)

Type 2 vs Type 1 virtualization

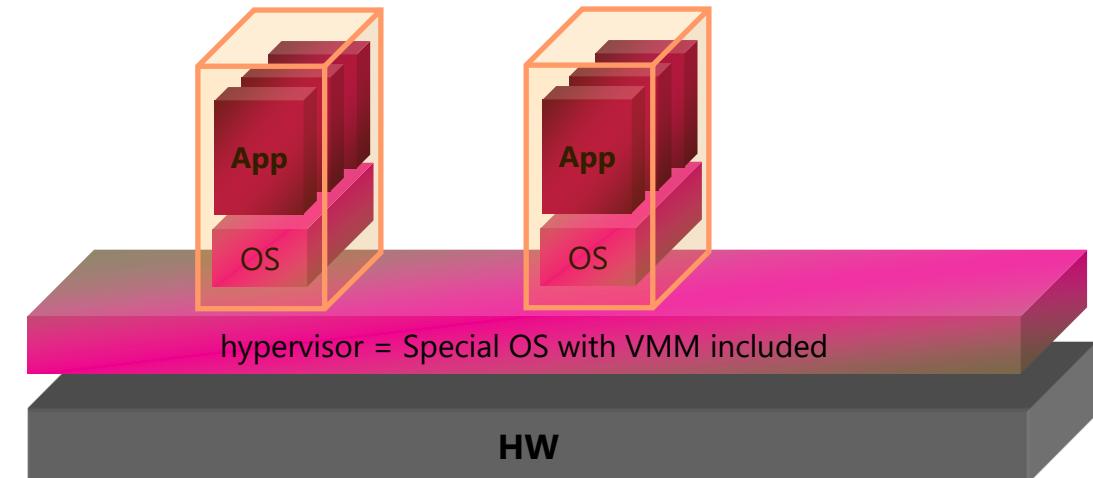
Hosted virtualization

- VMware Fusion and Workstation
- Oracle VM Virtualbox
- Parallels
- Microsoft Virtual Server and VirtualPC (discontinued)



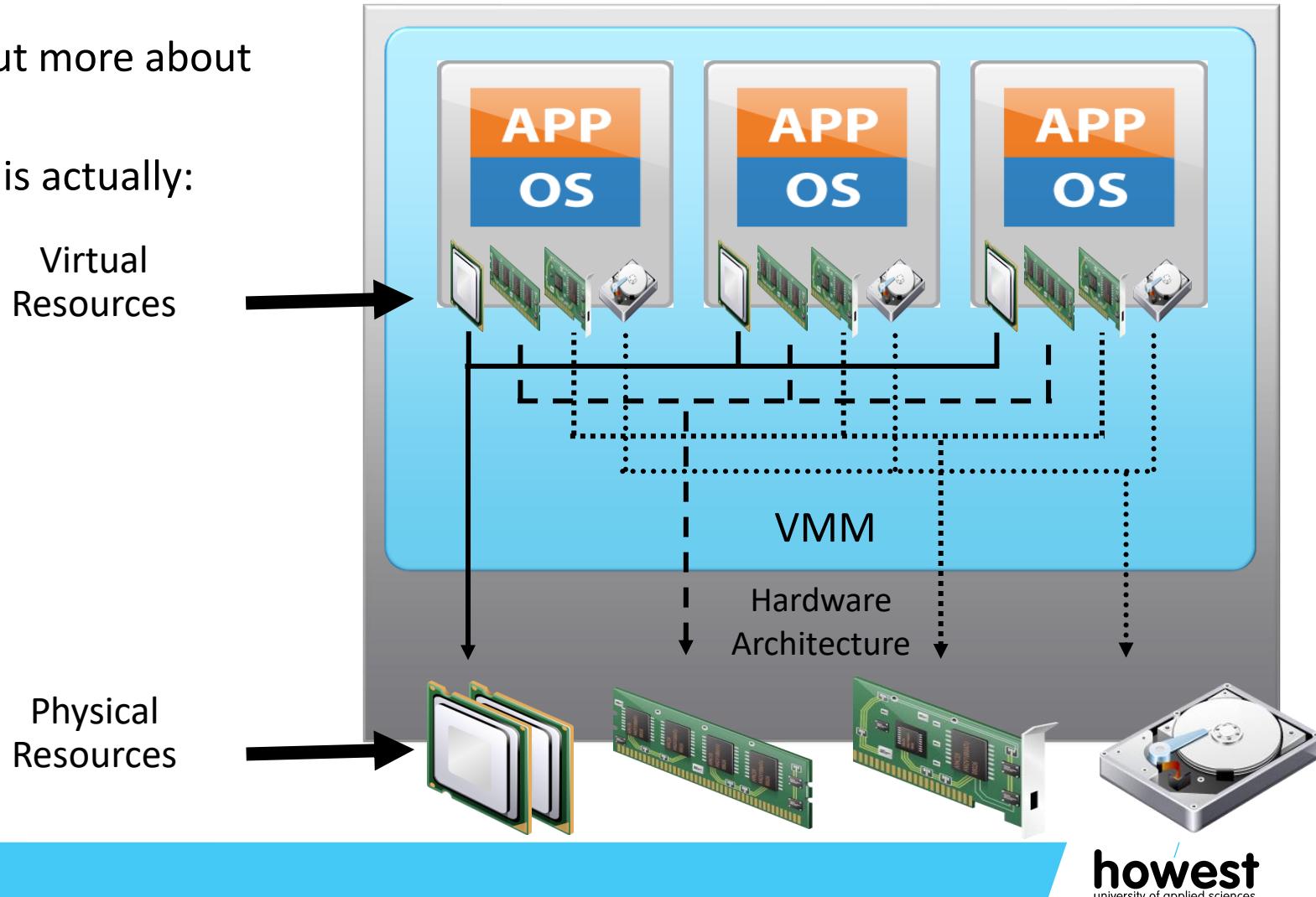
Bare-metal virtualization

- VMware ESXi
- Microsoft Hyper-V
- Citrix XenServer
- RedHat (& others) Kernel Virtual Machines (KVM)



First: General virtualization

- Virtualization is NOT emulation but more about **isolation**
- The concept of a Virtual Machine is actually:
a VM is a bunch of files



What is a VM?

a VM is a bunch of files:

VM firmware config

Virtual Hard Drive (like VHD)

snapshot overview (vmsn+vmdk)

VM config file (very important)

“Child” VMDKs (of the snapshots)

VM snapshot config

Detailed log files
(created upon every VM boot)

File Explorer navigation bar: ← → ⏪ ⏹ This PC > Documents > Virtual Machines > Debian 11.5

Name	Date modified	Type	Size
Debian 11.5.nvram	3/10/2022 21:32	VMware Virtual M...	9 KB
Debian 11.5.vmdk	19/09/2022 20:10	VMware virtual dis...	1,697,152 KB
Debian 11.5.vmsd	20/09/2022 15:10	VMware snapshot ...	1 KB
Debian 11.5.vmx	11/10/2022 12:32	VMware virtual m...	4 KB
Debian 11.5.vmxn	28/09/2022 9:19	VMware Team Me...	1 KB
Debian 11.5-000001.vmdk	20/09/2022 15:10	VMware virtual dis...	24,768 KB
Debian 11.5-000003.vmdk	11/10/2022 12:32	VMware virtual dis...	487,872 KB
Debian 11.5-Snapshot1.vmsn	19/09/2022 20:11	VMware virtual m...	28 KB
Debian 11.5-Snapshot2.vmsn	20/09/2022 15:10	VMware virtual m...	29 KB
vm.scoreboard	11/10/2022 9:30	SCOREBOARD File	8 KB
vm-10.scoreboard	3/10/2022 21:31	SCOREBOARD File	8 KB
vm-11.scoreboard	4/10/2022 9:33	SCOREBOARD File	8 KB
vm-12.scoreboard	10/10/2022 21:45	SCOREBOARD File	8 KB
vmware.log	11/10/2022 12:32	Text Document	198 KB
vmware-0.log	10/10/2022 22:23	Text Document	170 KB
vmware-1.log	4/10/2022 12:33	Text Document	179 KB
vmware-2.log	3/10/2022 23:14	Text Document	177 KB

This means interesting functionality

- A VM (being a bunch of files) can be moved very easily
 - Move to another machine and your VM runs on completely different hardware in minutes
- Not only moved but also copied (called a Clone), just copying the files of the VM
- No more “driver” issues; they do not require drivers for Nvidia, Intel, AMD, etc.
 - VMs allow for Operating Systems with small disk footprints
- We can create “snapshots”, **freezing** the state of a VM in a certain point in time.
 - Be aware that not only the VMDK is frozen but also the VM configuration.
 - Can be very powerful to have multiple states for one VM
 - Especially in combination with clones, called “linked clones”
- Entire network environments can be configured

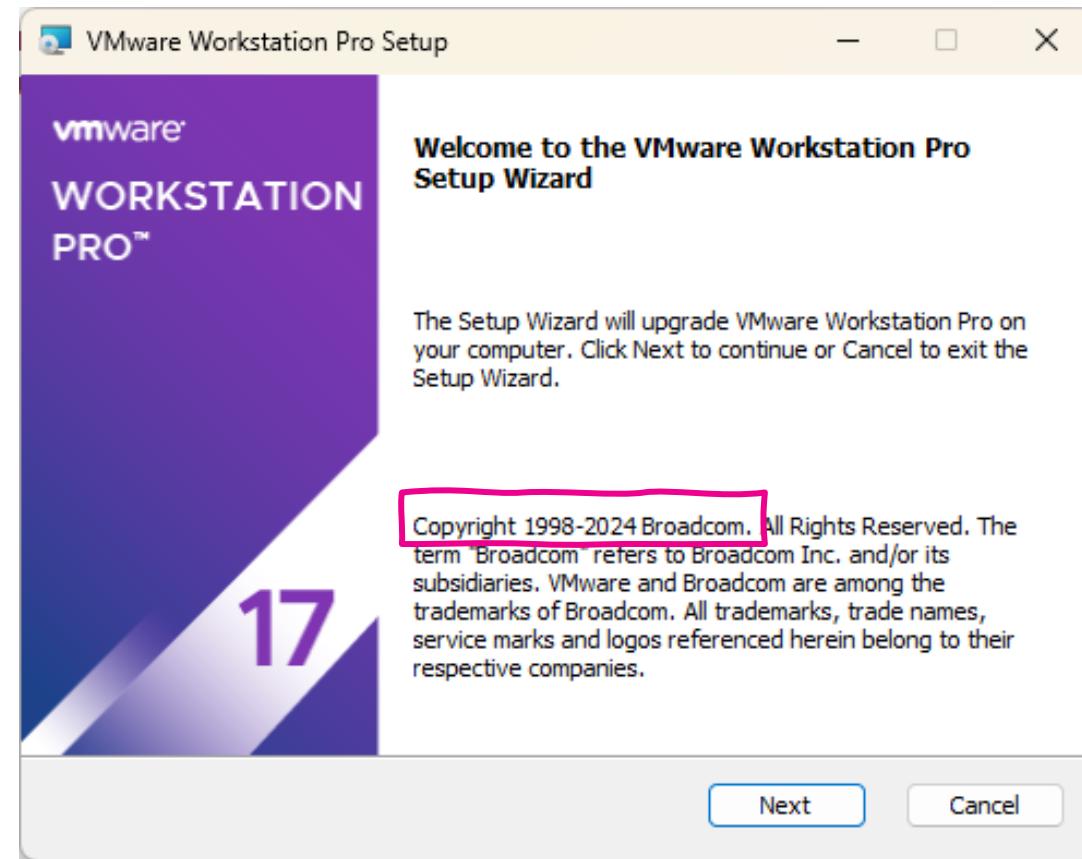
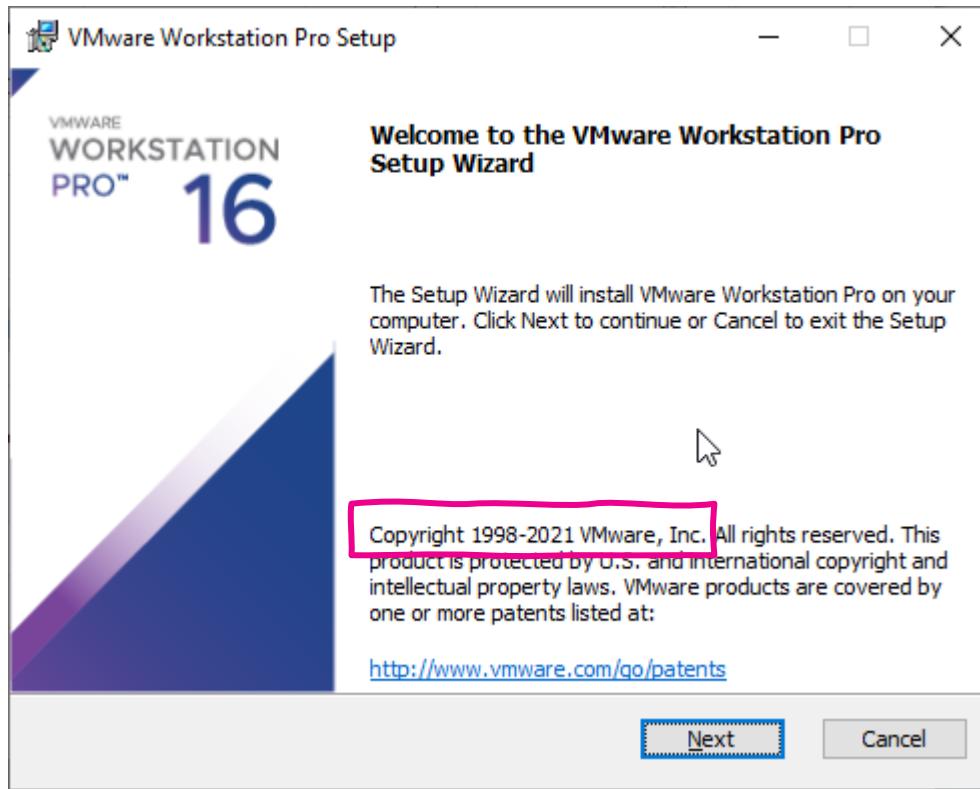
VMware server virtualization

VMware acquisition by Broadcom

- **Deal Value:** \$61 billion in cash and stock
- **Completion Date:** November 22, 2023
- **Criticism and Concerns**
 - Significant increase in licensing costs
 - Shift from perpetual licenses to subscriptions
 - Concerns over support quality and disrupted partnerships

VMware acquisition by Broadcom

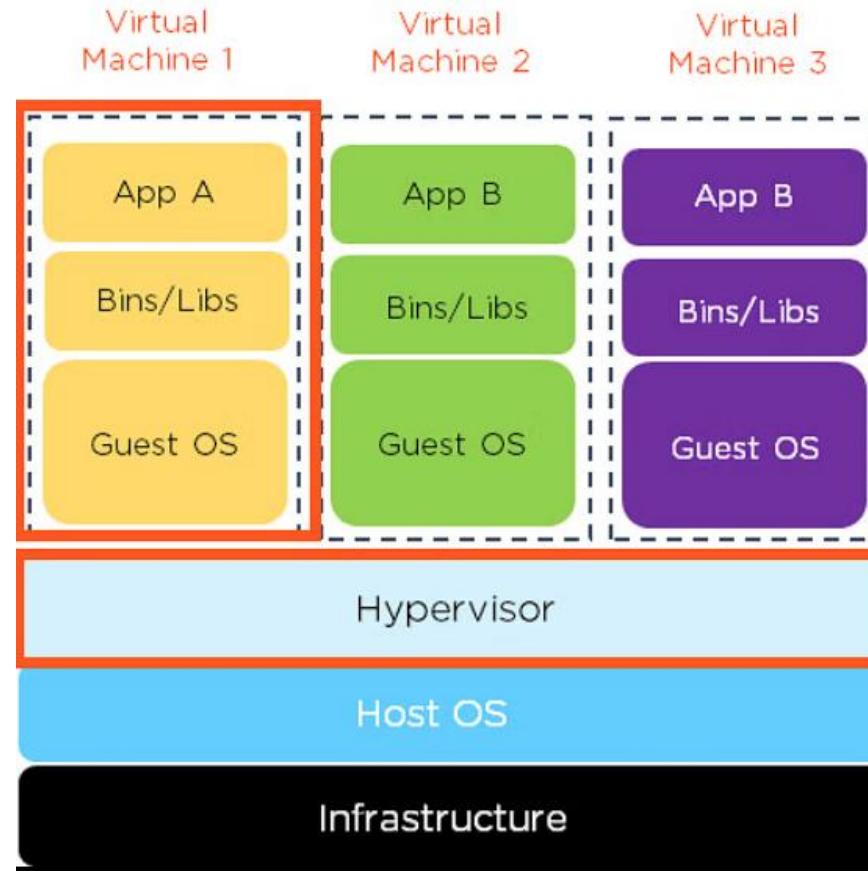
Since May 2024: VMware Workstation/Fusion Pro are free of charge for personal usage.



VMware workstation virtualization

VMware Workstation : hosted virtualization

- Your primary OS provides the virtualization



Workstation architecture

- **Type-2 Hosted Hypervisor**

- VMware Workstation operates as a hosted hypervisor, meaning it runs atop an existing host OS (Windows or Linux). It interfaces with the OS to manage guest VMs, rather than being installed directly on hardware.

- **Integration of a Virtual Machine Monitor (VMM)**

- It combines a hosted architecture with a VMM. This VMM includes a trap-and-emulate mechanism and a dynamic binary translator to virtualize x86 instructions effectively—even when hardware support for virtualization isn't available.

- **Full Virtualization via Binary Translation**

- VMware fully virtualizes guest OS environments by running non-sensitive instructions directly on hardware and using binary translation (or trapping) for sensitive ones. This allows the guest OS to remain unmodified while still maintaining high performance (typically achieving >80–95% of native speed).

- **Software-Emulated I/O Devices**

- The hypervisor emulates canonical I/O devices in software rather than exposing the full diversity of physical I/O hardware. This abstraction improves compatibility and portability across varied environments.

CPU

- Number of processors and number of core per processor
- Some features
 - Max or min Hz per CPU
 - Shares per VCPU
 - "low" → 500 shares per vCPU
 - "normal" → 1000 shares per vCPU (default)
 - "high" → 2000 shares per vCPU
 - or a numeric value (custom)
 - Example :
 - VM1: 2 vCPUs, sched.cpu.shares = "high" ($2000 \times 2 = 4000$ shares)
 - VM2: 2 vCPUs, sched.cpu.shares = "low" ($500 \times 2 = 1000$ shares)
 - Total shares = 5000.
 - VM1 gets $4000 / 5000 = 80\%$ of CPU time.
 - VM2 gets $1000 / 5000 = 20\%$ of CPU time.

Processors

Number of processors:

Number of cores per processor:

Total processor cores: 2

Virtualization engine

Virtualize Intel VT-x/EPT or AMD-V/RVI

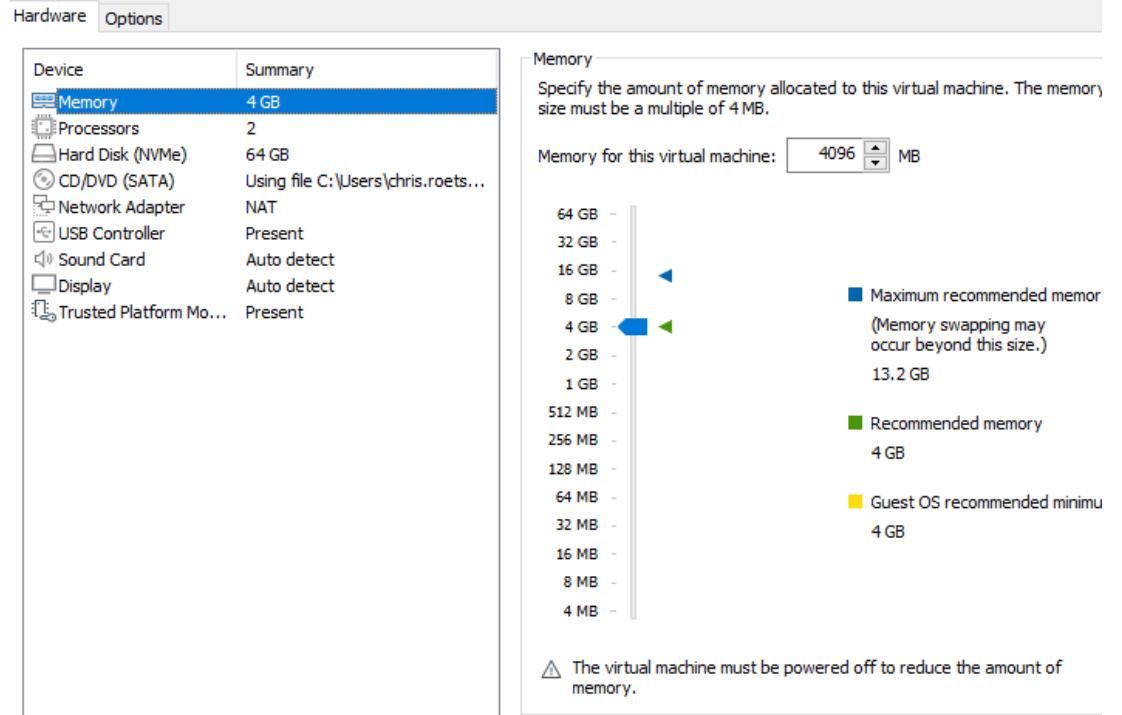
Virtualize CPU performance counters

Virtualize IOMMU (IO memory management unit)

```
#####
# CPU Settings
#####
numvcpus = "2"
sched.cpu.min = "1000"      # Reservation in MHz
sched.cpu.max = "2000"      # Limit in MHz
sched.cpu.shares = "normal" # Shares: low / normal / high / numeric
```

Memory

- dynamically allocating RAM to virtual machines (VMs) as they need it, with options to adjust settings like reserved memory, shares, limits, and reservations
- So the physical addresses are handed over to the address translation table of the guest OS
- Some features
 - **Memory Overcommitment**
 - **Memory Swapping**
 - **Memory Ballooning**
 - **Memory Reservation**
 - **Memory Limit**



```
#####
# Memory Settings
#####
memsize = "4096"          # Reservation in MB
sched.mem.min = "2048"      # Limit in MB
sched.mem.max = "4096"
mem.locked = "TRUE"         # Lock in host RAM
mem.force = "TRUE"
```

Disks

- Virtual disks
 - File on your host system (vmdk)
- Physical disks
 - Partition on your host system
- Features
 - Shares
 - through



```
#####
# Disk I/O Settings (per disk)
#####
scsi0:0.shares = "high"
scsi0:0.throughputCap = "20480"      # 20 MB/s limit
scsi0:0.throughputCap.enable = "TRUE"
```

Network

- Bridged :
- NAT : vmnet8
- Host only : vmnet0
- Custom
- Bandwidth
- Lan segment : compare with VLAN's

Device status

Connected
 Connect at power on

Network connection

Bridged: Connected directly to the physical network
 Replicate physical network connection state

NAT: Used to share the host's IP address
 Host-only: A private network shared with the host
 Custom: Specific virtual network
VMnet0

LAN segment:
LAN Segments... Advanced...

Incoming Transfer

Bandwidth: Unlimited
Kbps:
Packet Loss (%): 0.0
Latency (ms): 0

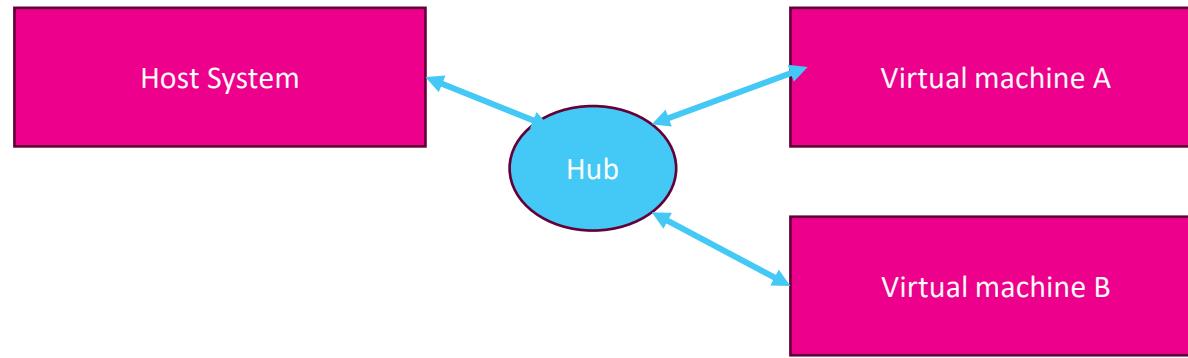
Outgoing Transfer

Bandwidth: Unlimited
Kbps:
Packet Loss (%): 0.0
Latency (ms): 0

MAC Address

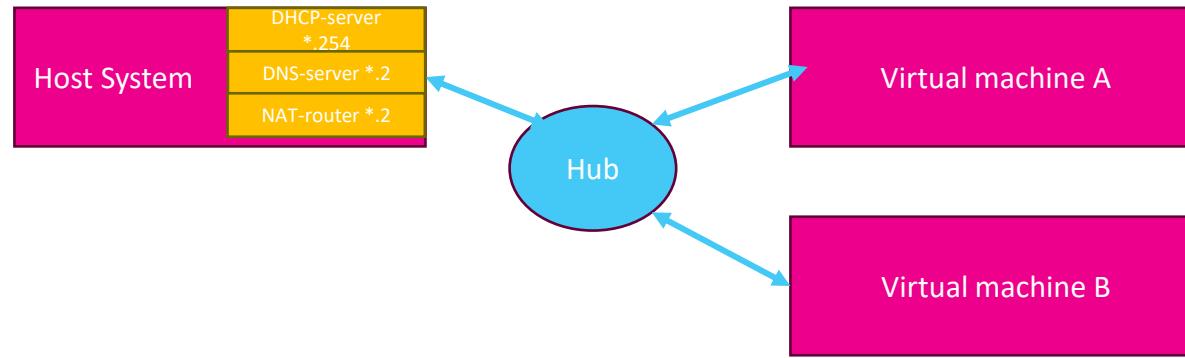
00:0C:29:56:D2:95
Generate

Virtual networking reexplained : Host only



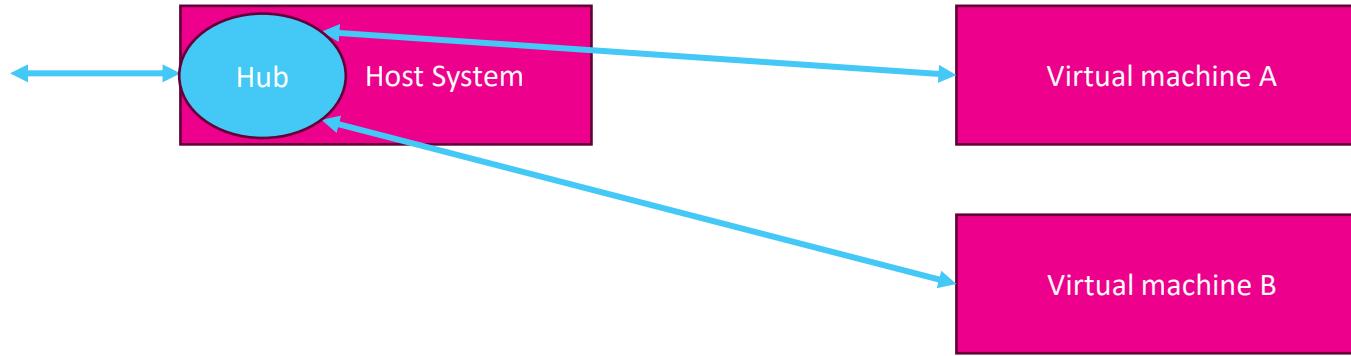
- IP address definitions is up to you

Virtual networking reexplained : Nat



- IP address are configured for you, using a network subnet defined in the virtual address editor

Virtual networking reexplained : Bridged

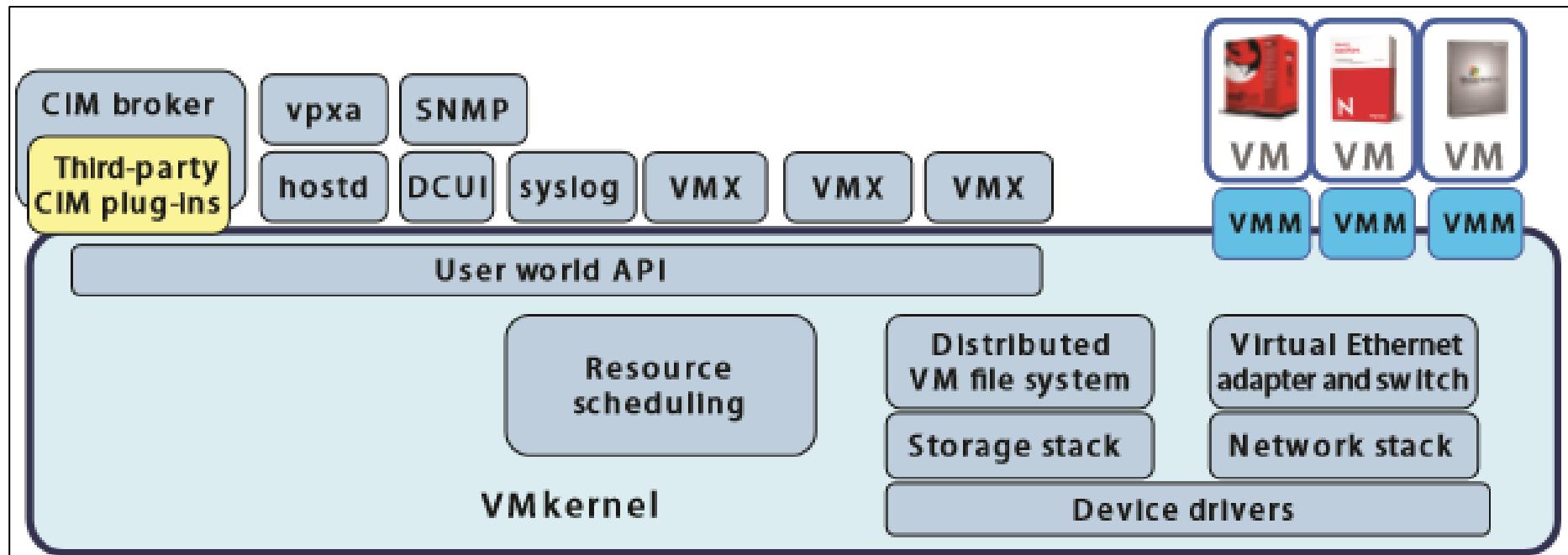


- IP addresses depend on external network of the host

VMware server virtualization

Server Virtualization: ESXi

- As we know: a Hypervisor is a special kind of **operating system** that contains a virtualization layer
- VMware's hypervisor product is the **vSphere Hypervisor**, better known as **ESXi** (or 'VMvisor')
- The "magic" is in the kernel. VMware calls this the **VMkernel**.



Hardware Compatibility List (HCL)

Compared to Windows and Linux only certain hardware is supported with ESXi.

There is an online list of supported hardware: the **ESXi Hardware Compatibility List (HCL)**:

- <https://www.vmware.com/resources/compatibility/>

Home / Resources / Compatibility Guides

VMware Compatibility Guide

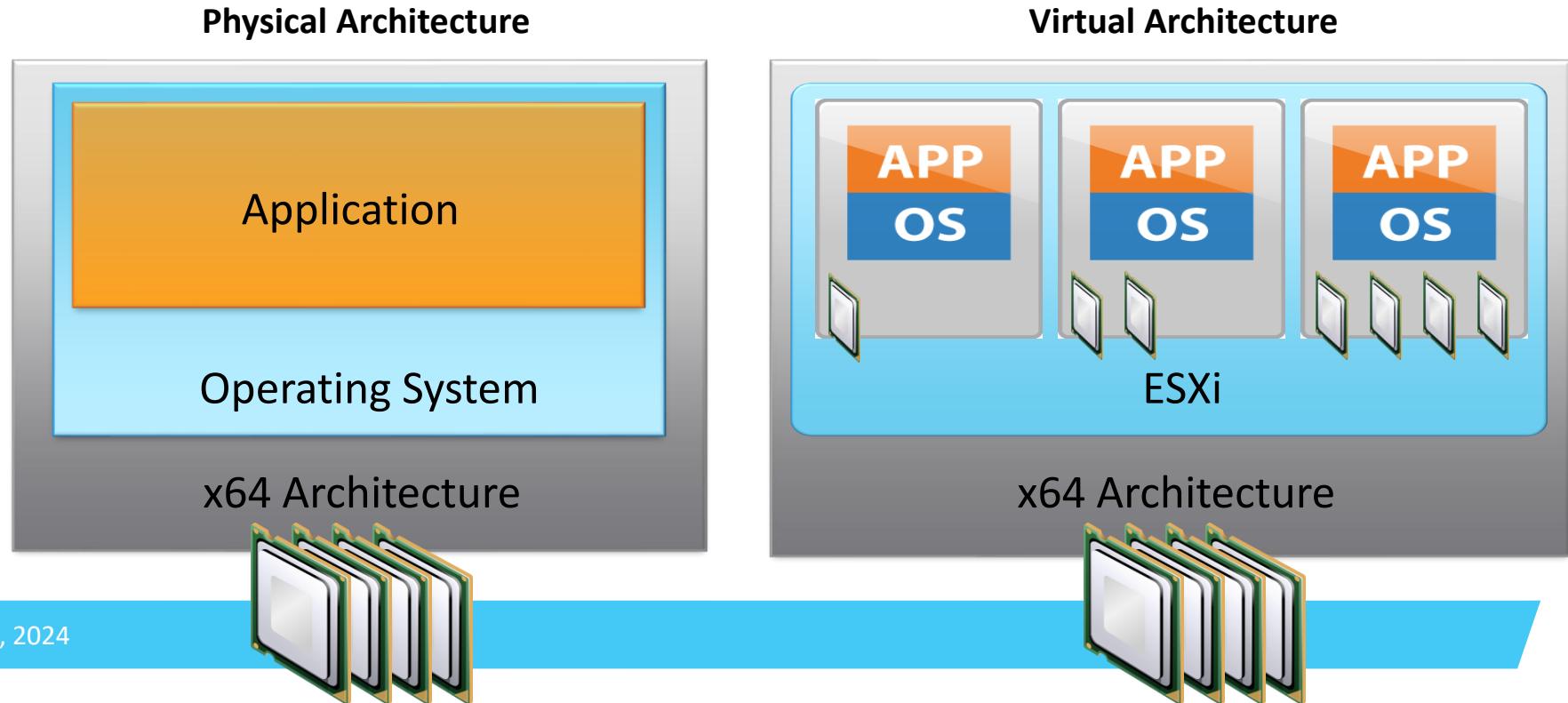
The screenshot shows the VMware Compatibility Guide search interface. At the top, there is a search bar with the placeholder "Search Compatibility Guide: (e.g. compatibility or esx or 3.0)" and a "Search" button. Below the search bar, the page title "VMware Compatibility Guide" is displayed. The main area contains several filter sections:

- What are you looking for:** A dropdown menu set to "Systems / Servers".
- Product Release Version:** A dropdown menu showing options like "All", "ESXi 8.0 U2", "ESXi 8.0 U1", "ESXi 8.0", "ESXi 7.0 U3", and "ESXi 7.0 U2".
- System Type:** A dropdown menu showing options like "All", "Blade", "Mother Board", "Rack or Tower", "Rackmount", and "Tower".
- Additional Criteria:** A link to "Collapse All".
- Min Certified Memory:** A dropdown menu set to "All".
- Max Certified Memory:** A dropdown menu set to "All".
- Sockets:** A dropdown menu set to "All".
- Enhanced vMotion Capability Modes:** A dropdown menu showing options like "All", "AMD Opteron™ Generation 1", "AMD Opteron™ Generation 2", "AMD Opteron™ Generation 3", "AMD Opteron™ Generation 3 without 3DNc", and "AMD Opteron™ Generation 4".
- Max Cores per Socket:** A dropdown menu set to "All".
- CPU Series:** A dropdown menu showing options like "All", "AMD EPYC 7001 Series", "AMD EPYC 7002/7Fx2/7Hx2 Series", "AMD EPYC 7003/7003X Series", "AMD EPYC 9004 Series", "AMD Opteron 6200 Series", and "AMD Opteron 6300 Series".
- Fault Tolerant Compatible Sets:** A dropdown menu showing options like "All", "AMD Bulldozer Generation", "AMD Opteron™ Generation 3", "AMD Piledriver Generation", "Intel® Haswell Generation", "Intel® Ivy-Bridge Generation", "Intel® Nehalem Generation", "Intel® Penny Generation", and "Intel® Sandy-Bridge Generation".
- Posted Date Range:** A dropdown menu set to "All".

At the bottom of the filter section, there are two buttons: "Update and View Results" and "Reset". To the right of the filters, the text "Help" and "Current Results: 2170" are displayed.

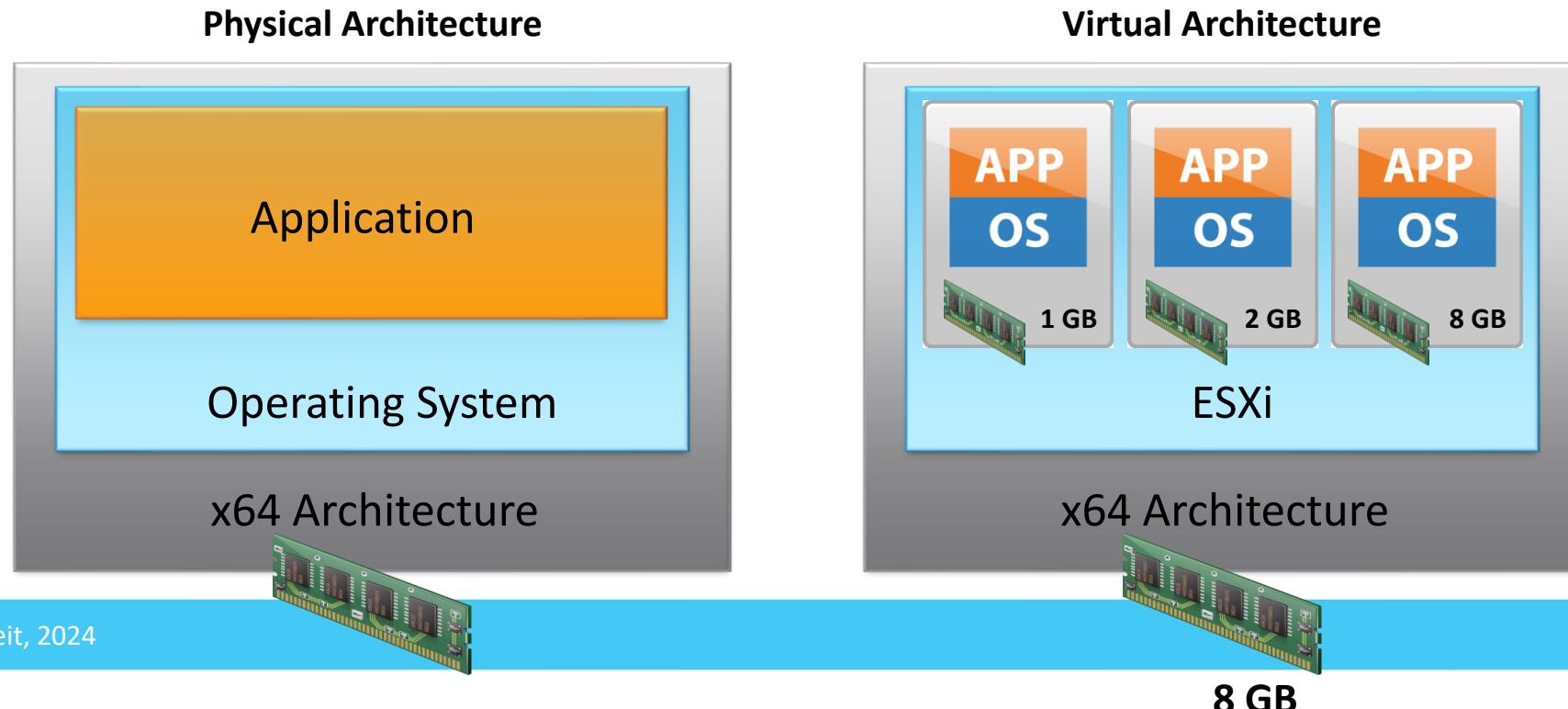
Some concepts: CPU

- It is perfectly possible to hand out more CPU's (or cores) in total for all VMs together than there are physically available in the device (“**overprovisioning**”)
- Only when heavily used, contention will exist and there is performance impact



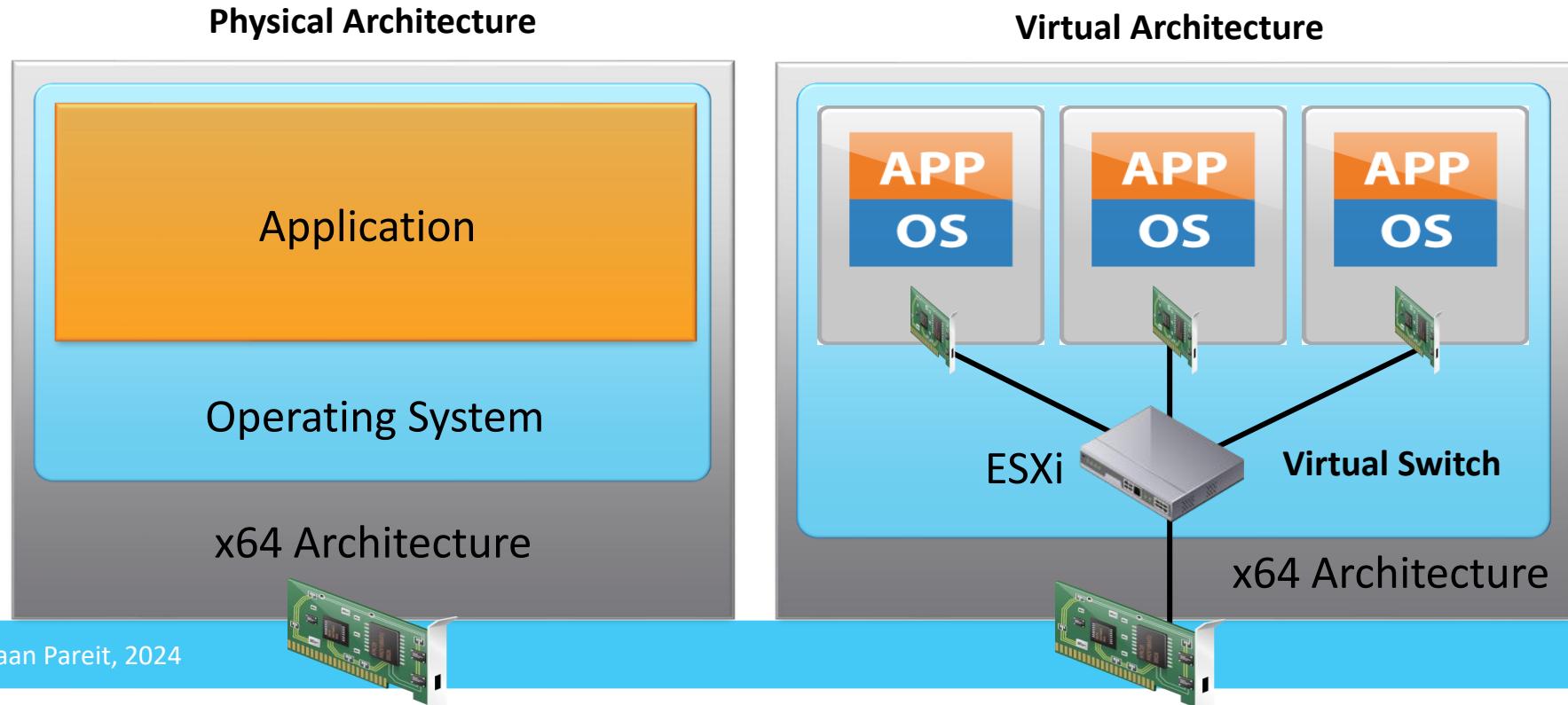
Some concepts: MEMORY

- It is also possible to hand out more memory than is physically present on the system
 - something which is only possible in VMware (thus not in e.g. Hyper-V)
 - ESXi uses several Memory Savings techniques
- vmem file is created when starting a VM: last resort for memory swapping, also for pausing a VM



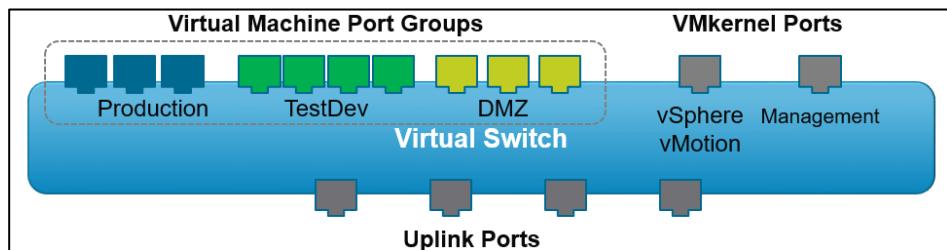
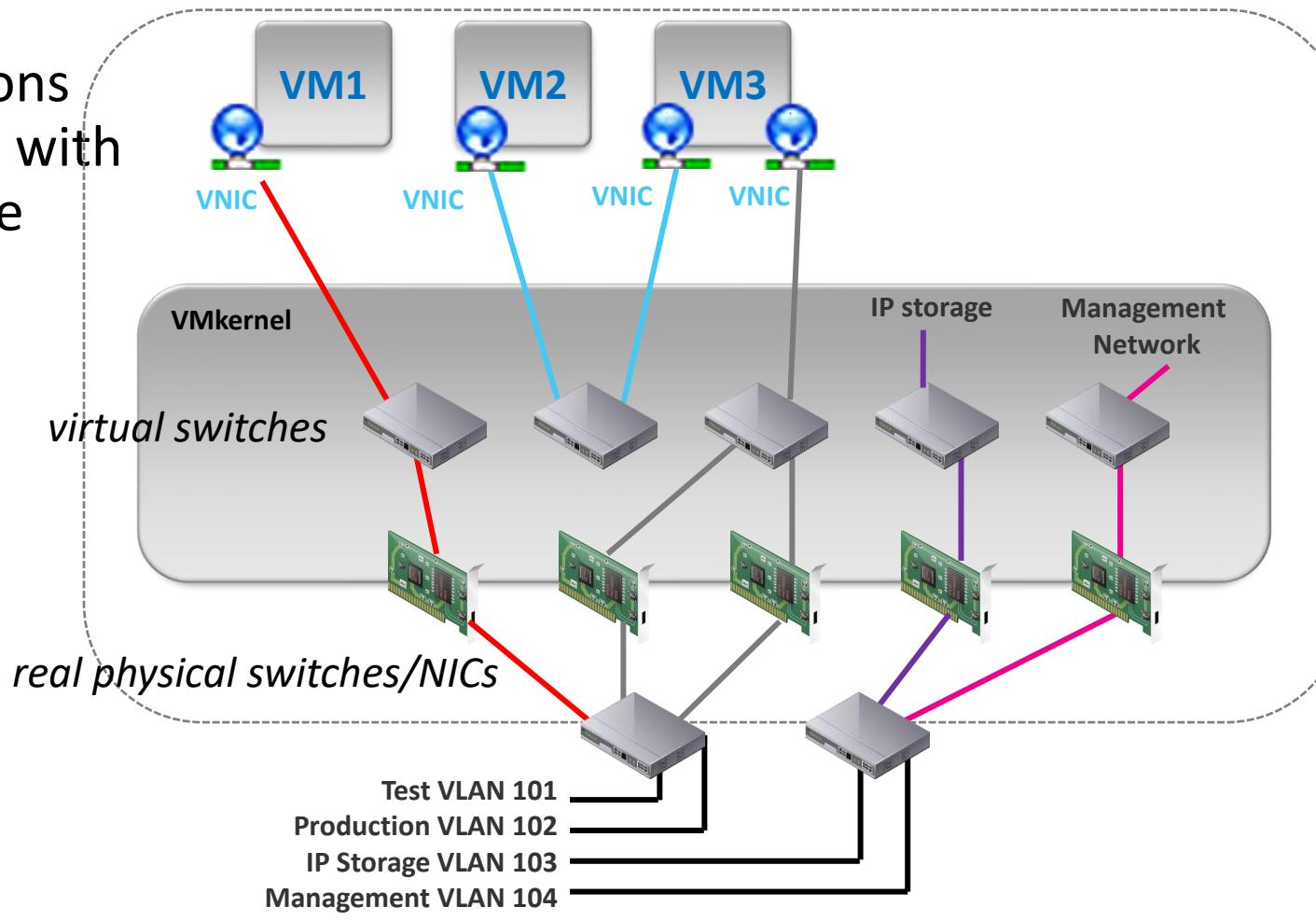
Some concepts: NETWORKING

- ESXi uses the concept of virtual switches,
 - it **does not have built-in features** like ‘Host-Only’ or ‘NAT’ (unlike Workstation/Fusion)
- vSwitches have two sides: ‘**virtual port groups**’ (towards VM) and ‘**uplink**’ (towards physical NIC).



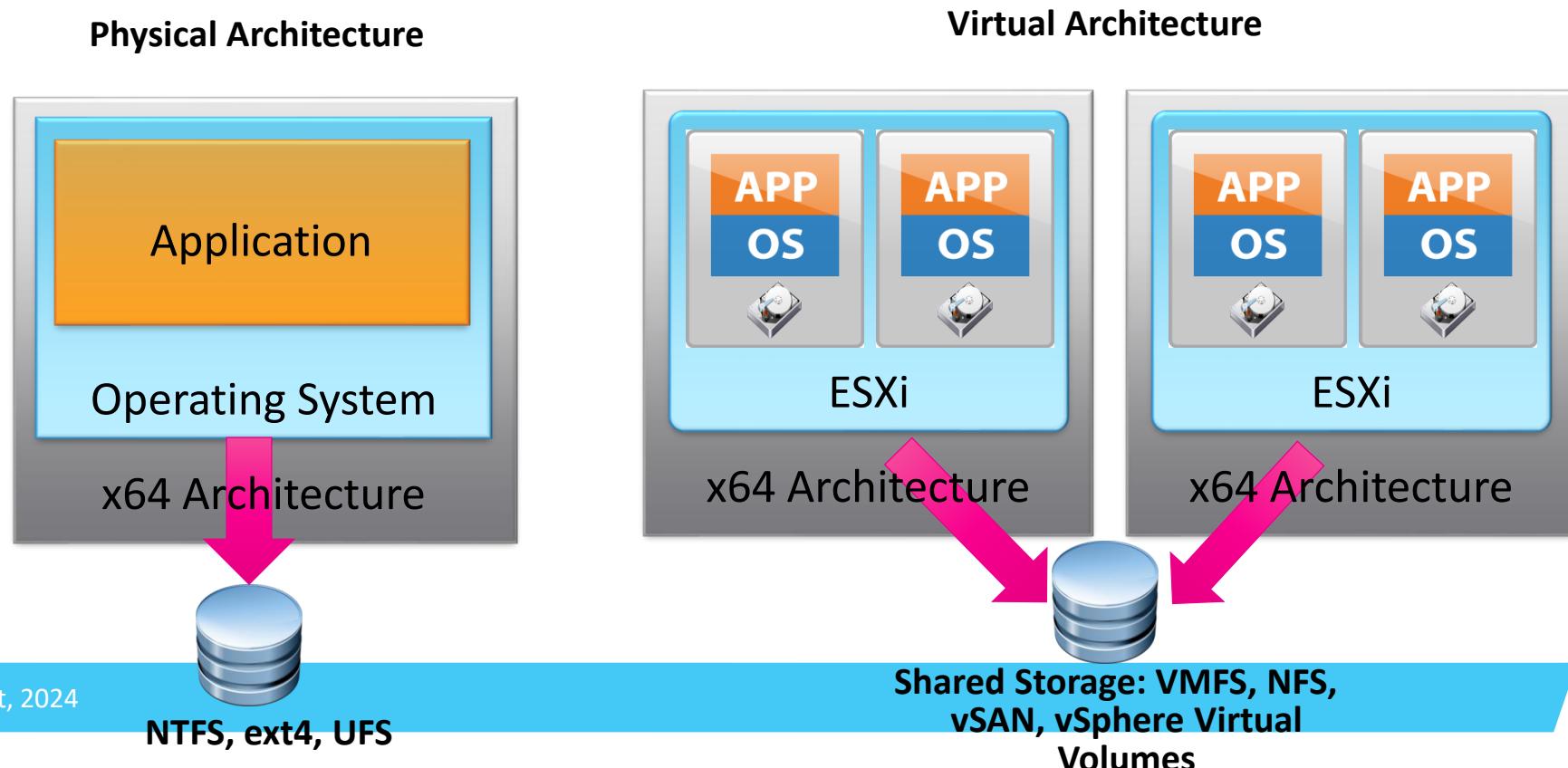
Components of Standard Switches

A standard switch provides connections for virtual machines to communicate with one another, whether they are on the same host or on different hosts.



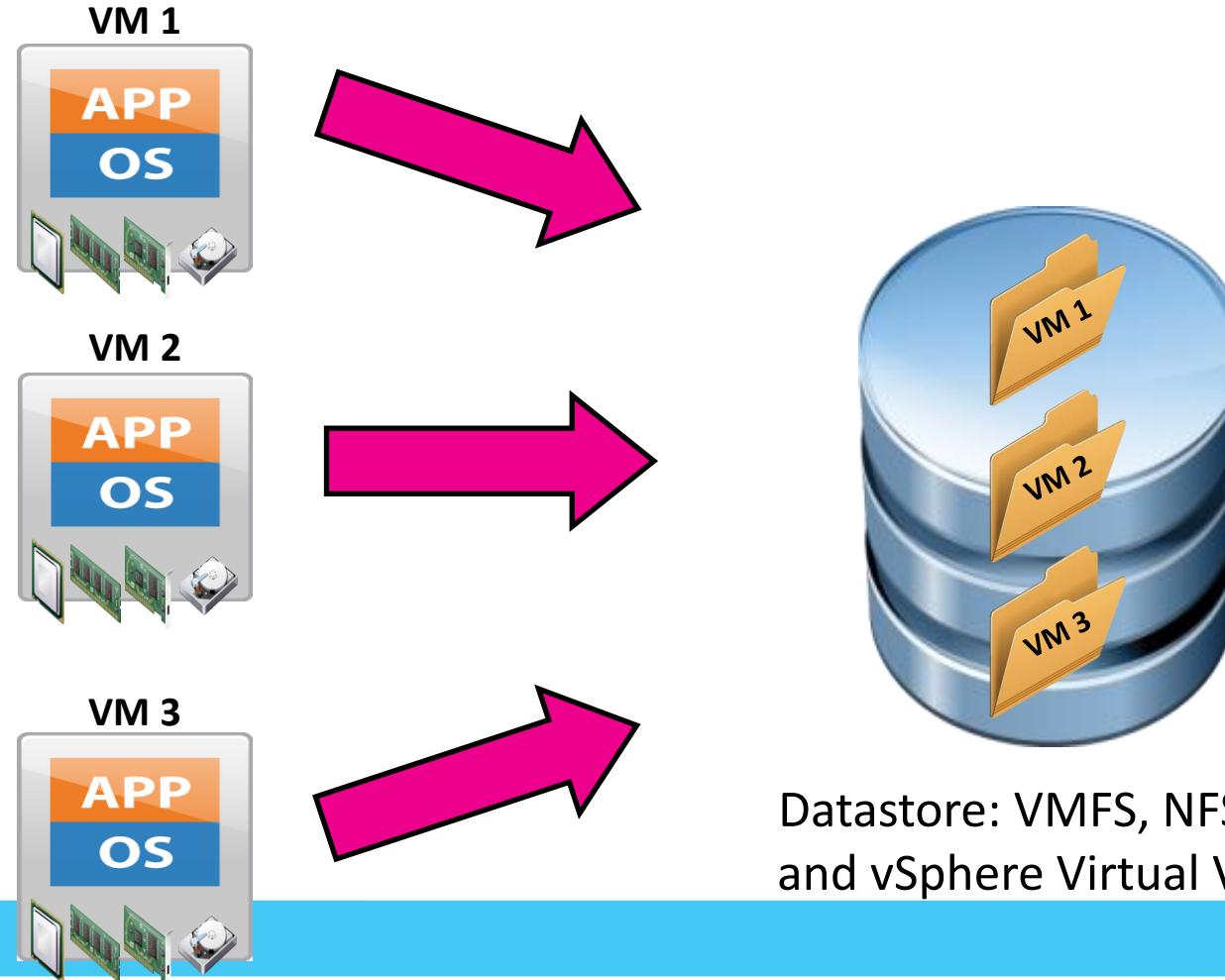
Some concepts: STORAGE

- VMware has created its own, special file system called **VMFS** or “Virtual Machine File System”
- It is optimized for Virtual Machine usage (large files),
- It supports simultaneous access (e.g. to the same SAN via iSCSI) by multiple ESXi servers (unique!)



Encapsulation

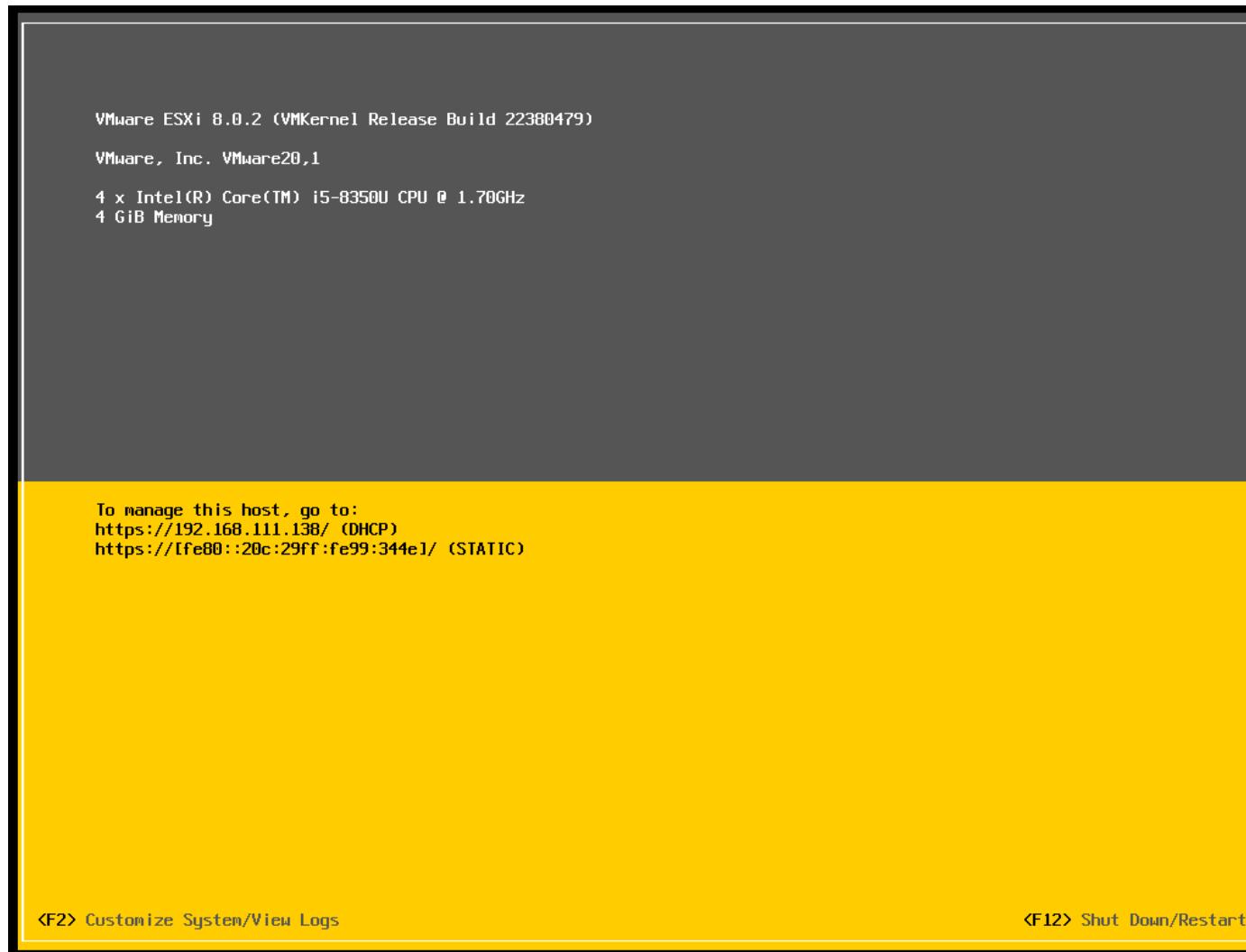
Virtual machine files are stored in directories on a VMFS, NFS, VMware vSAN™ or VMware vSphere® Virtual Volumes™ datastore.



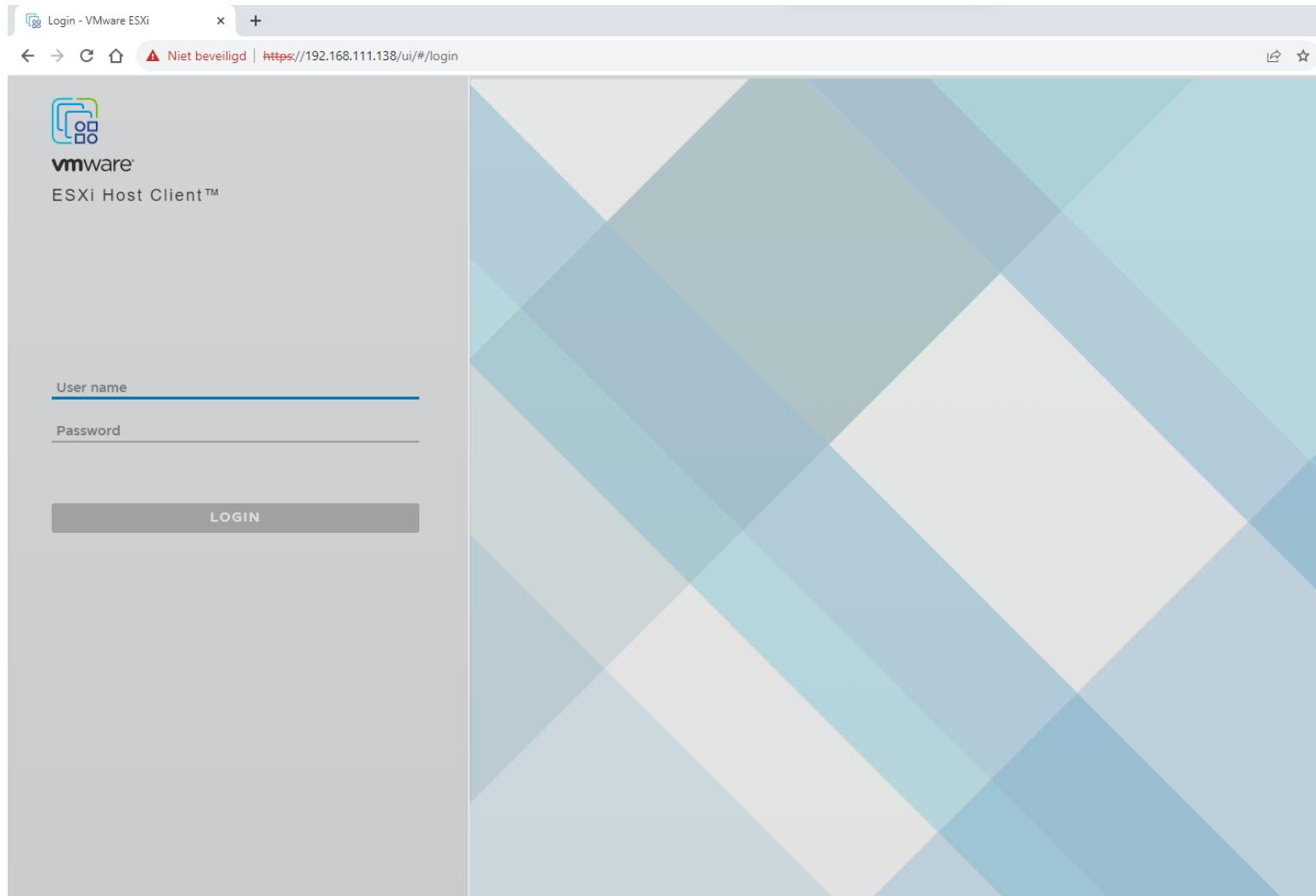
ESXi practicalities

- ESXi has a very simple installer:
 - Only a password and a hard drive selection is needed during installation
 - It also has a very **limited direct configuration**, it is **managed entirely via a Web interface**
- VMware used to recommend to install it to a USB thumb drive (or SD card, maybe in RAID1)
 - It loads into RAM and only uses the thumb drive to store ESXi image & configuration.
 - “ESXi configuration” is very limited: NIC settings, password, keyboard, service settings
- VMs are on a (VMFS) volume (not on thumb drives)
 - Resetting ESXi does not change the VMs ...

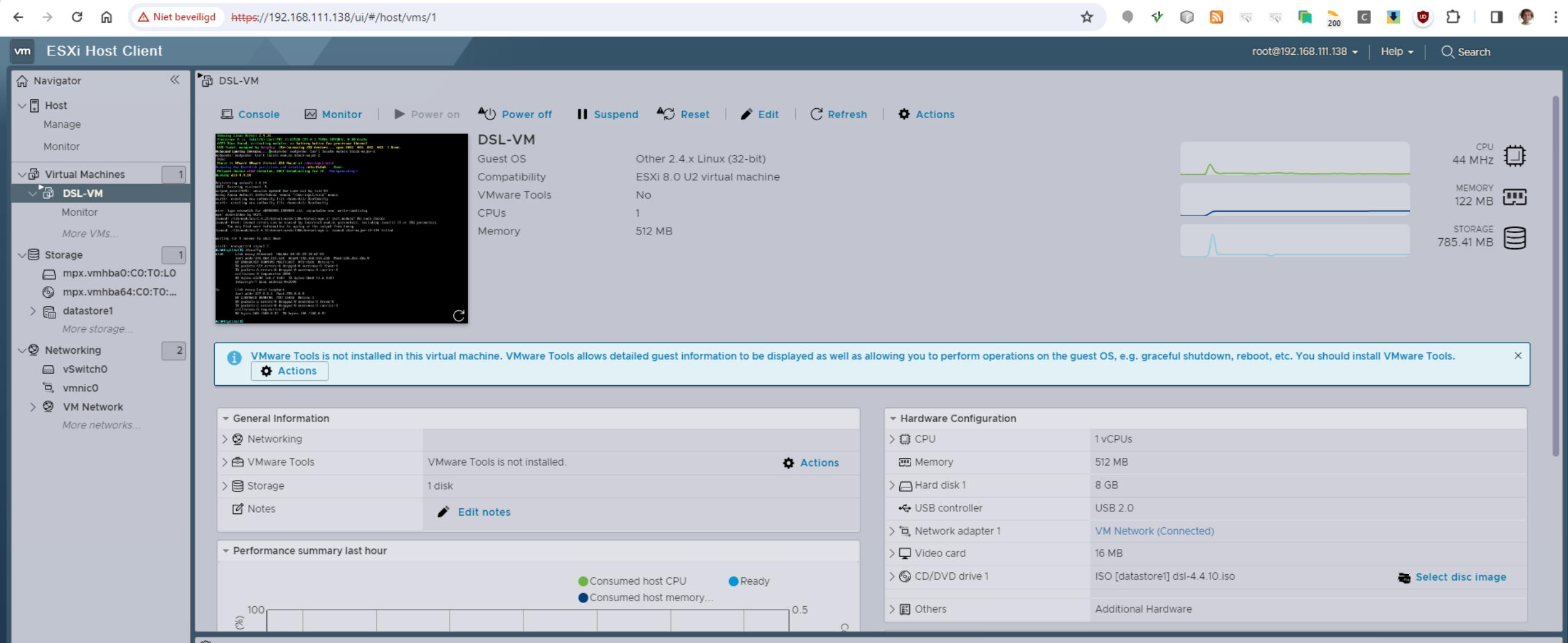
ESXi Direct Console User Interface (DCUI)



ESXi web interface



ESXi web interface



VMware Remote Console

DSL-VM - VMware Remote Console

VMRC | II | ⊞ | ⊞

```
Running Linux Kernel 2.4.31.  
Processor 0 is Intel(R) Core(TM) i5-8350U CPU @ 1.70GHz 1895MHz, 0 KB Cache  
ACPI Bios found, activating modules: ac battery button fan processor thermal  
USB found, managed by hotplug: (Re-)scanning USB devices... sync:[001 001 002 003] Done.  
Autoconfiguring devices... modprobe: modprobe: Can't locate module block-major-2  
modprobe: modprobe: Can't locate module block-major-2  
Done.  
Mouse is VMware VMware Virtual USB Mouse at /dev/input/mice  
Scanning for Harddisk partitions and creating /etc/fstab... Done.  
Network device eth0 detected, DHCP broadcasting for IP. (Backgrounding)  
Running dsl-4.4.10  
  
Registering unionfs 1.0.14  
INIT: Entering runlevel: 5  
su(pam_unix)[489]: session opened for user dsl by (uid=0)  
Using Xvesa default 1024x768x32 -mouse "/dev/input/mice" mouse  
xauth: creating new authority file /home/dsl/.Xauthority  
xauth: creating new authority file /home/dsl/.Xauthority  
  
mtrr: type mismatch for e8000000,1000000 old: uncachable new: write-combi  
apm: overridden by ACPI.  
insmod: /lib/modules/2.4.31/kernel/arch/i386/kernel/apm.o: init_module: No  
such file or directory
```

Instellingen voor virtual machine

Hardware Opties

Apparaat	Samenvatting
Geheugen	512 MB
Processors	1
Hard disk 1	8 GB (Thin provisioned)
CD/DVD drive 1	Het externe bestand [datast...
Network adapter 1	VM Network
USB controller	Aanwezig
Video card	Automatisch detecteren

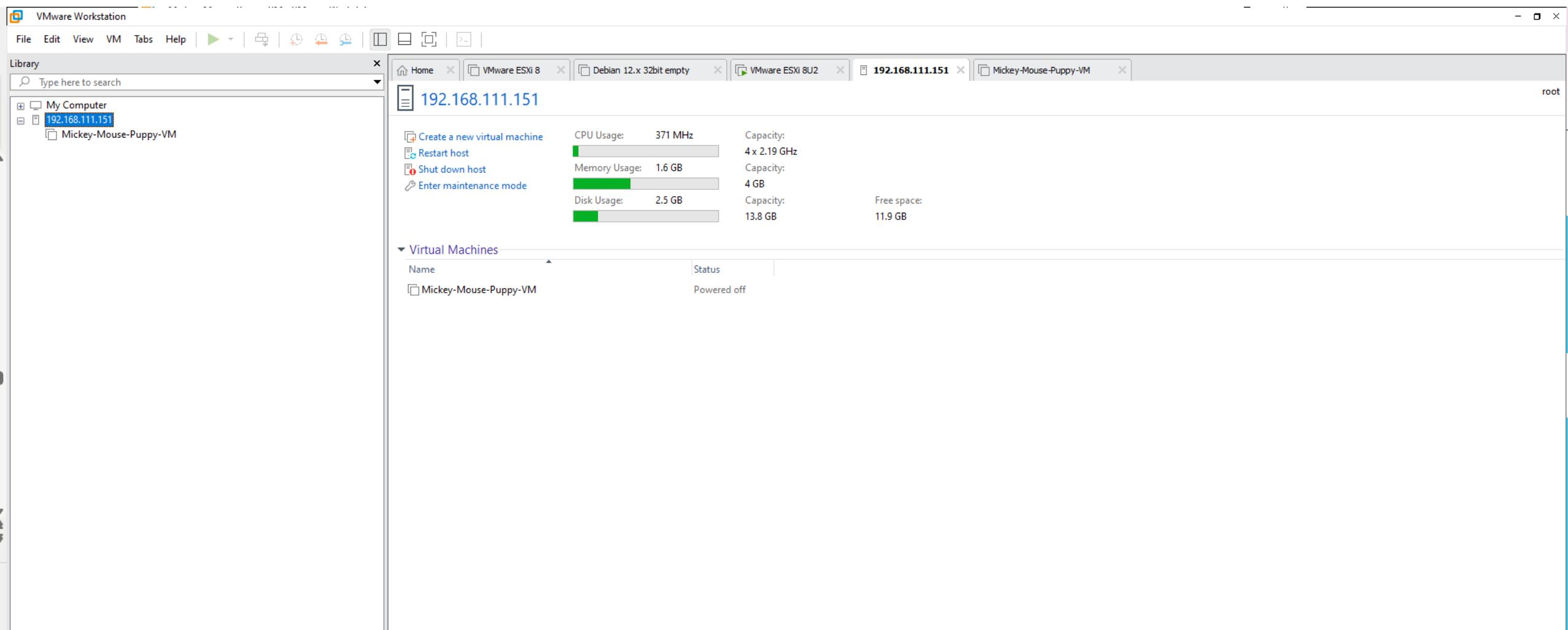
Geheugen

Geef de hoeveelheid geheugen op die aan moet een veelvoud zijn van 4 MB.

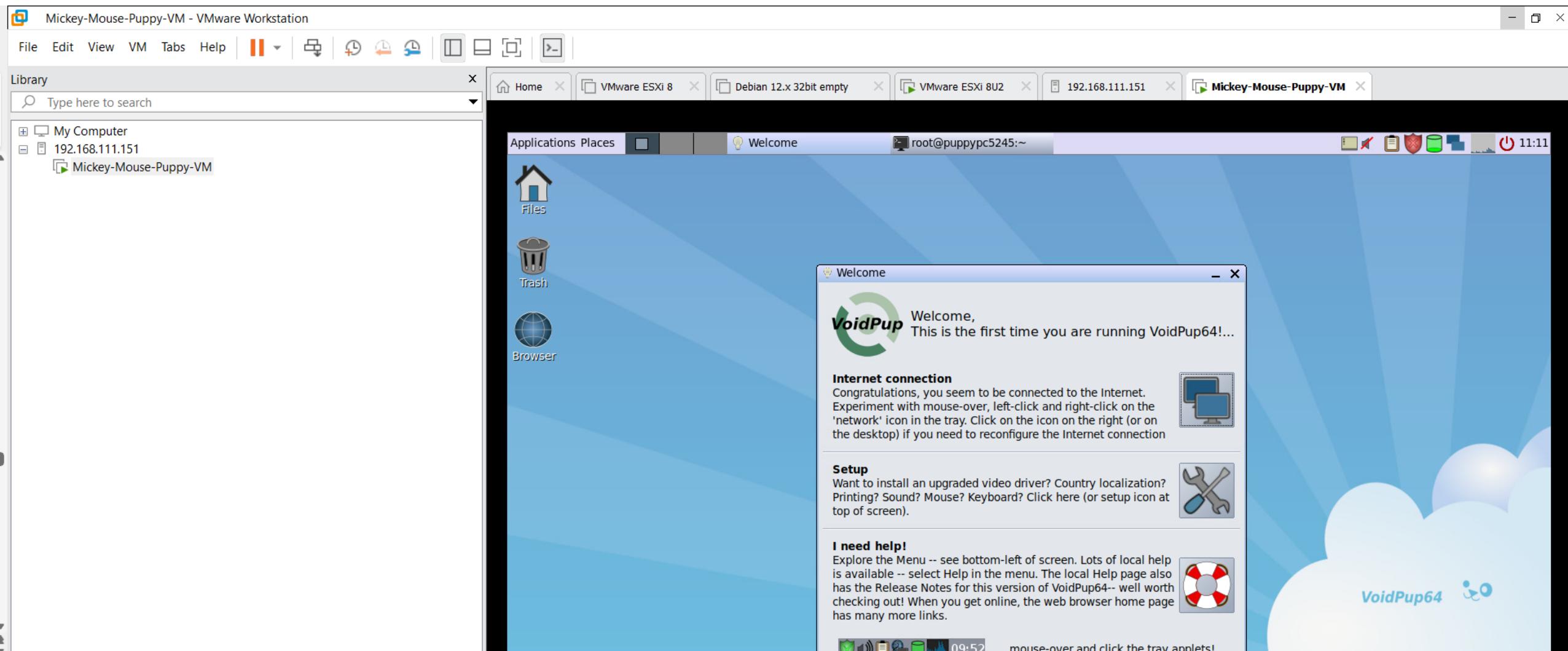
Geheugen voor deze virtual machine:

24.0 TB
16 TB
8 TB
4 TB
2 TB
1 TB
512 GB
256 GB
128 GB
64 GB
32 GB
16 GB
8 GB
4 GB
2 GB
1 GB
512 MB
256 MB
128 MB
64 MB
32 MB
16 MB

VMware workstation – “Connect to server”



VMware workstation – “Connect to server”



End Of General Availability of the free vSphere Hypervisor (ESXi 7.x and 8.x)



Article ID: 345098



Updated On: 07-09-2024

Products

VMware vSphere ESXi

VMware vSphere ESXi 6.0

VMware vSphere ESXi 7.0

VMware vSphere ESXi 8.0

Issue/Introduction

VMware vSphere Hypervisor (free edition) is no longer available on the VMware website.

Resolution

As part of the transition of perpetual licensing to new subscription offerings, the VMware vSphere Hypervisor (Free Edition) has been marked as EOGA (End of General Availability). At this time, there is not an equivalent replacement product available.

For further details regarding the affected products and this change, we encourage you to review the following blog post: [VMware End Of Availability of Perpetual Licensing and SaaS Services](#).

Central management in VMware

VMware vSphere Suite

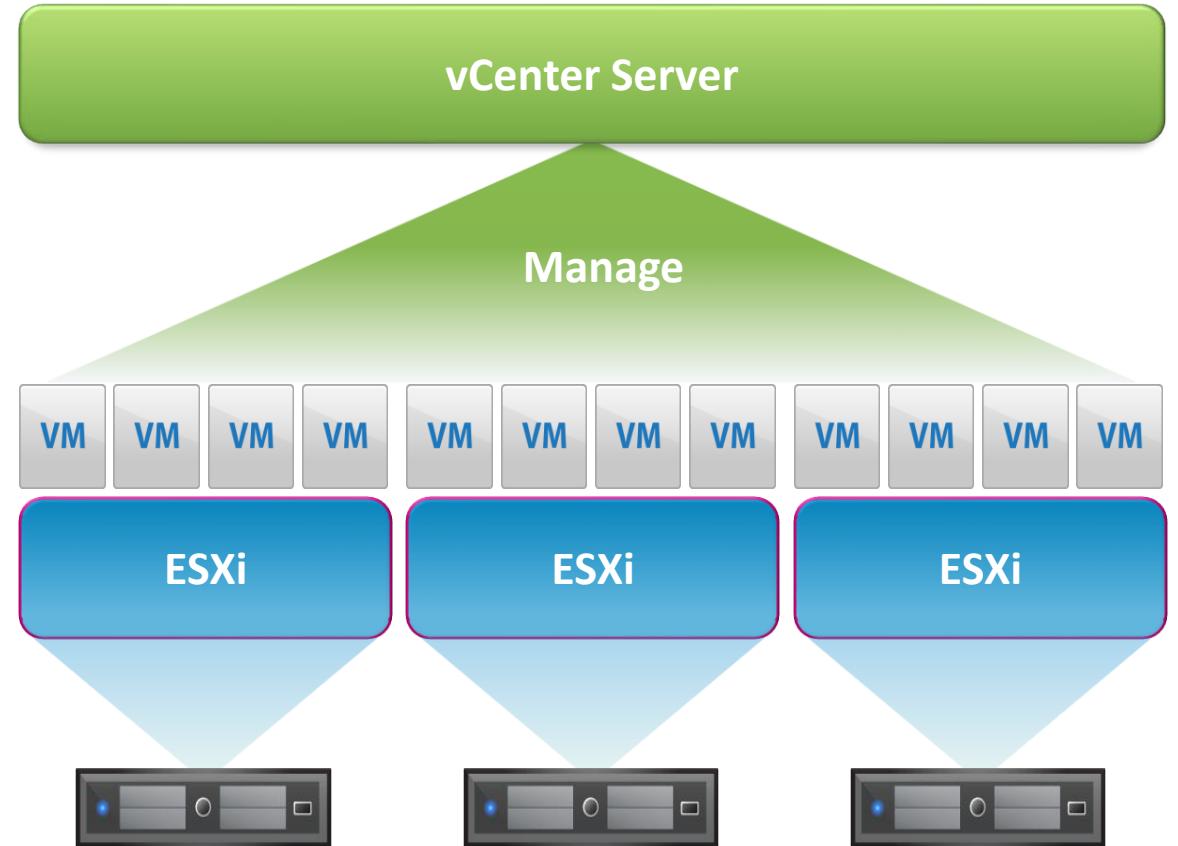
- When managing multiple ESXi hosts, a central management solution is necessary
 - This is called the “**VMware vSphere Suite**”
 - Multiple editions: ‘Essentials (Plus)’, ‘Standard’, ‘Enterprise Plus’, ‘Desktop’, etc.
- This suite contains multiple products
 - The minimal is “**vSphere Hypervisor (ESXi)**” and “**vCenter Server**”
- **vCenter Server is really a VM or Appliance**
 - **Installing this is as simple as deploying a ready-made virtual machine**

Essentials	
VMware vSphere Hypervisor (ESXi) 8.0U2	2023-09-21
VMware vCenter Server 8.0U2	2023-09-21
VMware NSX 4.1.1.0 For vShield Endpoint	2023-08-15
VMware Tools 12.3.0	2023-08-31
Essentials Plus	
VMware NSX 4.1.1.0 For vShield Endpoint	2023-08-15
VMware vSphere Hypervisor (ESXi) 8.0U2	2023-09-21
VMware vCenter Server 8.0U2	2023-09-21
VMware vSphere Replication 8.8.0	2023-09-21
VMware Tools 12.3.0	2023-08-31
Standard	
VMware vSphere Hypervisor (ESXi) 8.0U2	2023-09-21
VMware vCenter Server 8.0U2	2023-09-21
VMware vSphere Replication 8.8.0	2023-09-21
VMware Aria Automation Orchestrator 8.13.1	2023-09-07
VMware NSX 4.1.1.0 For vShield Endpoint	2023-08-15
VMware Tools 12.3.0	2023-08-31

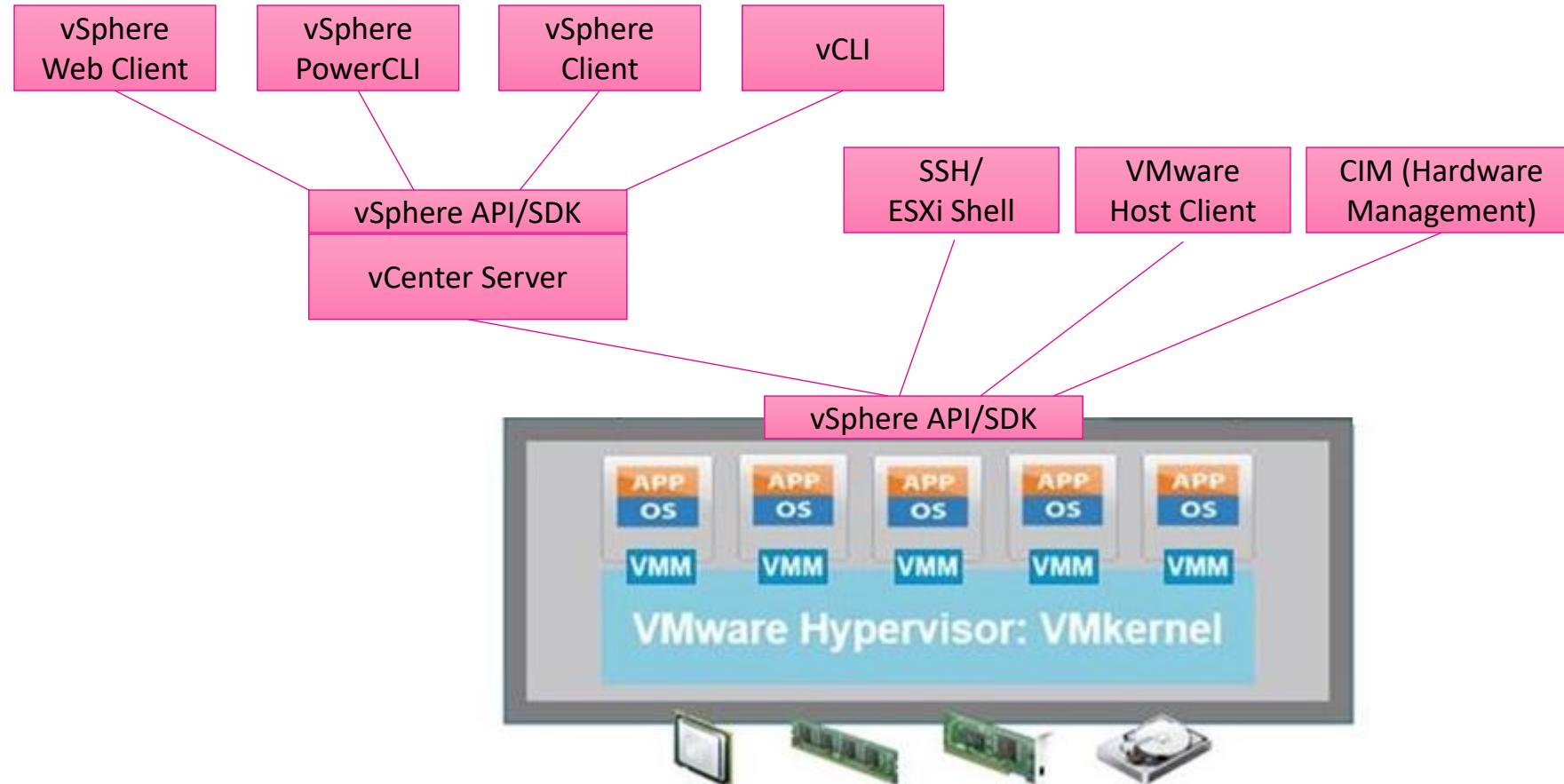
vCenter Server

vCenter Server adds features that are not available in a ~~free~~ standalone ESXi

- Cloning VM's
- Templates
- Updates via GUI
- Replication of VM's
- High Availability (more later)
- DRS / DPM (more later)



vCenter Server



Managing vCenter Server

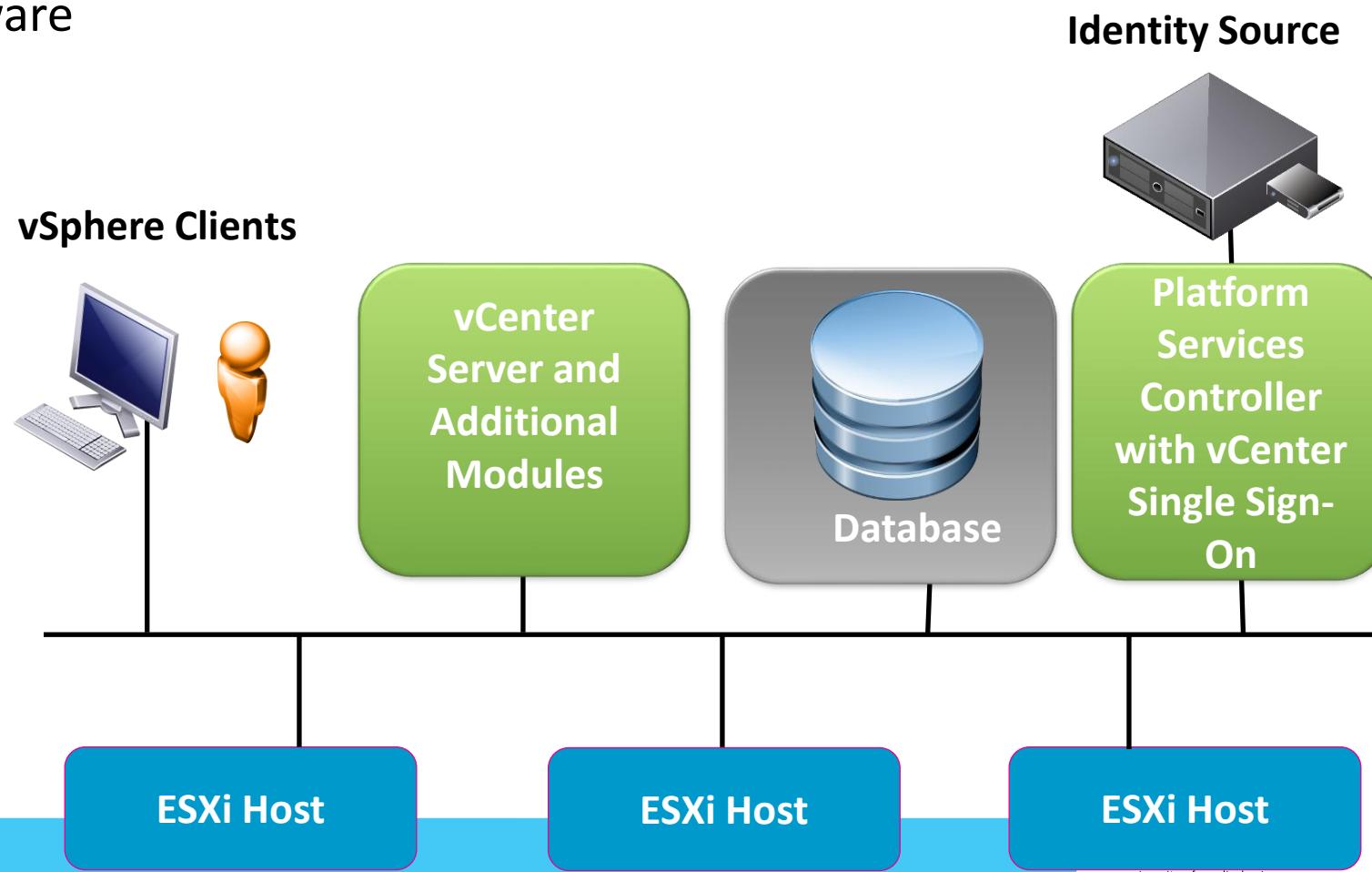
vCenter Server is modular, many third-party vendors have supporting software

- E.g. Dell EMC or HP EVA modules

vCenter can be managed via

- Web interface (HTML5)
- PowerShell
(Install-Module
-Name VMware.PowerCLI)

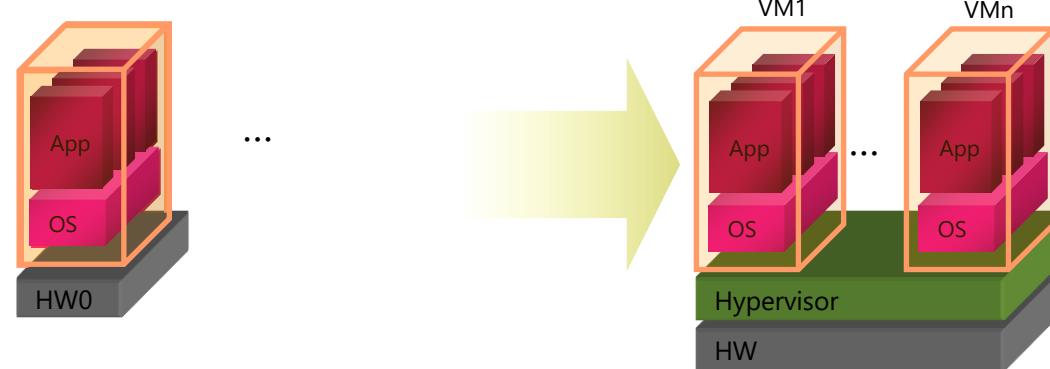
Uses “Single Sign-On” and is usually linked to AD for easy login



Shared Storage in datacenters

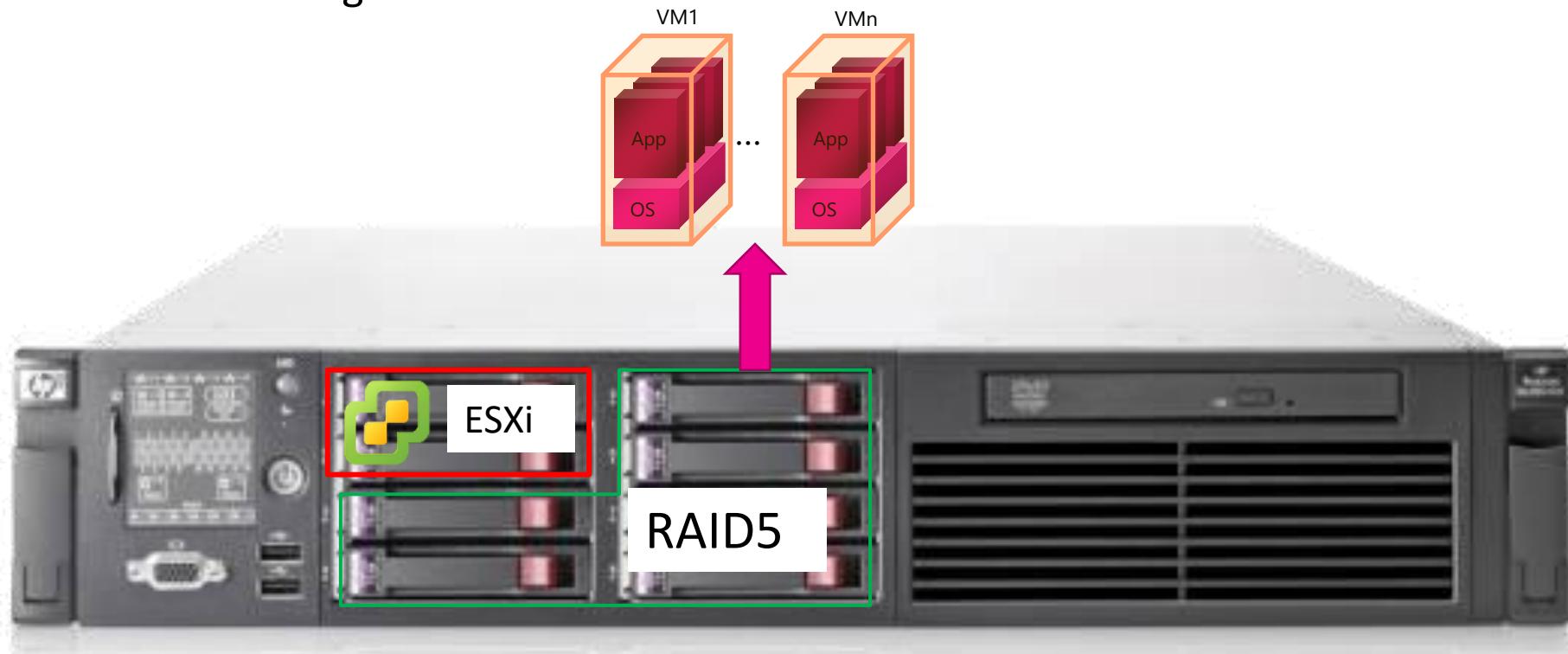
Advanced Features (Bare-Metal only)

- First phase of Virtualization: **Consolidation**
- Running multiple Operating Systems on a single system



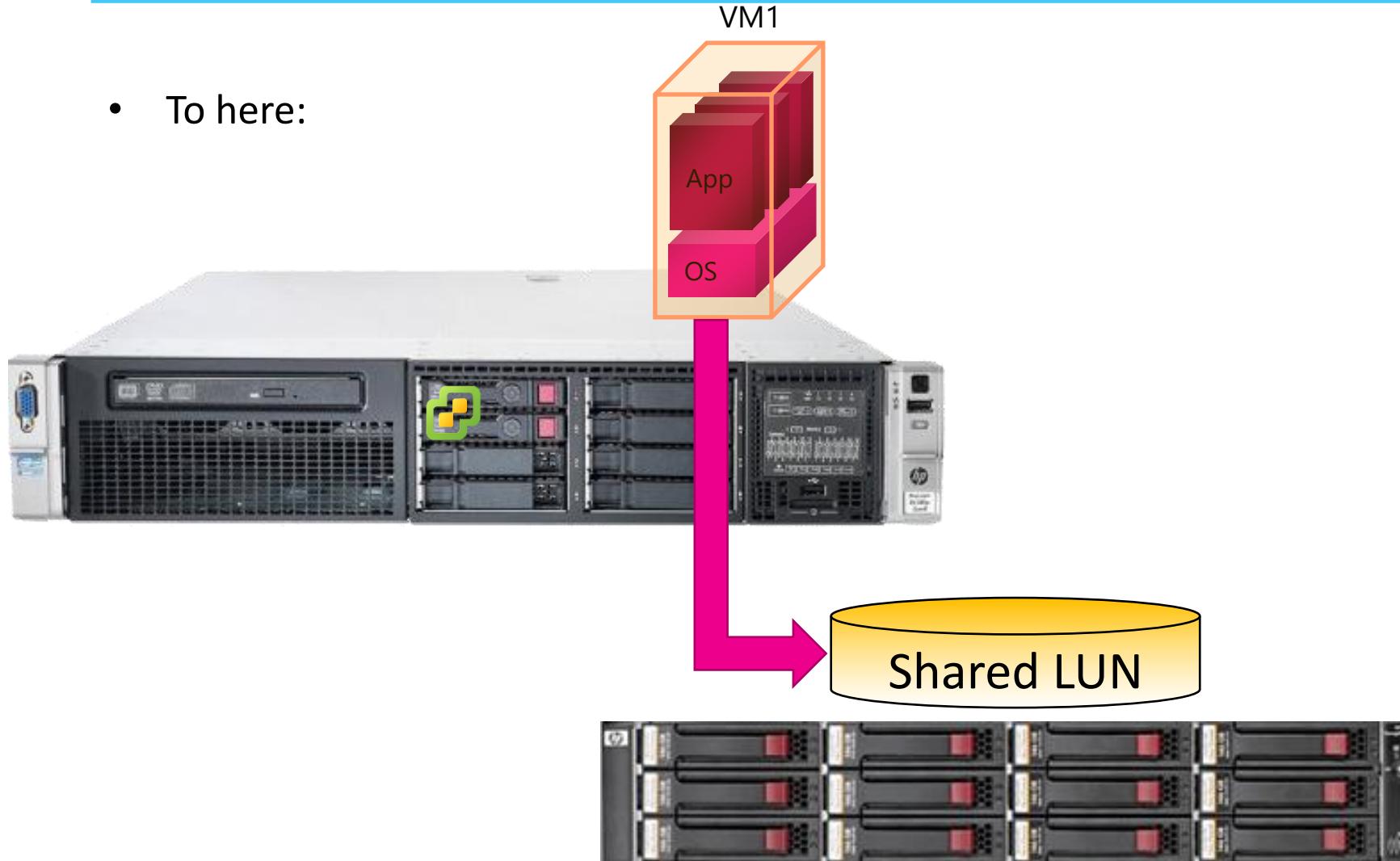
Datacenters “think” differently

- By using **Shared Storage**
- Datacenters are moving from here:

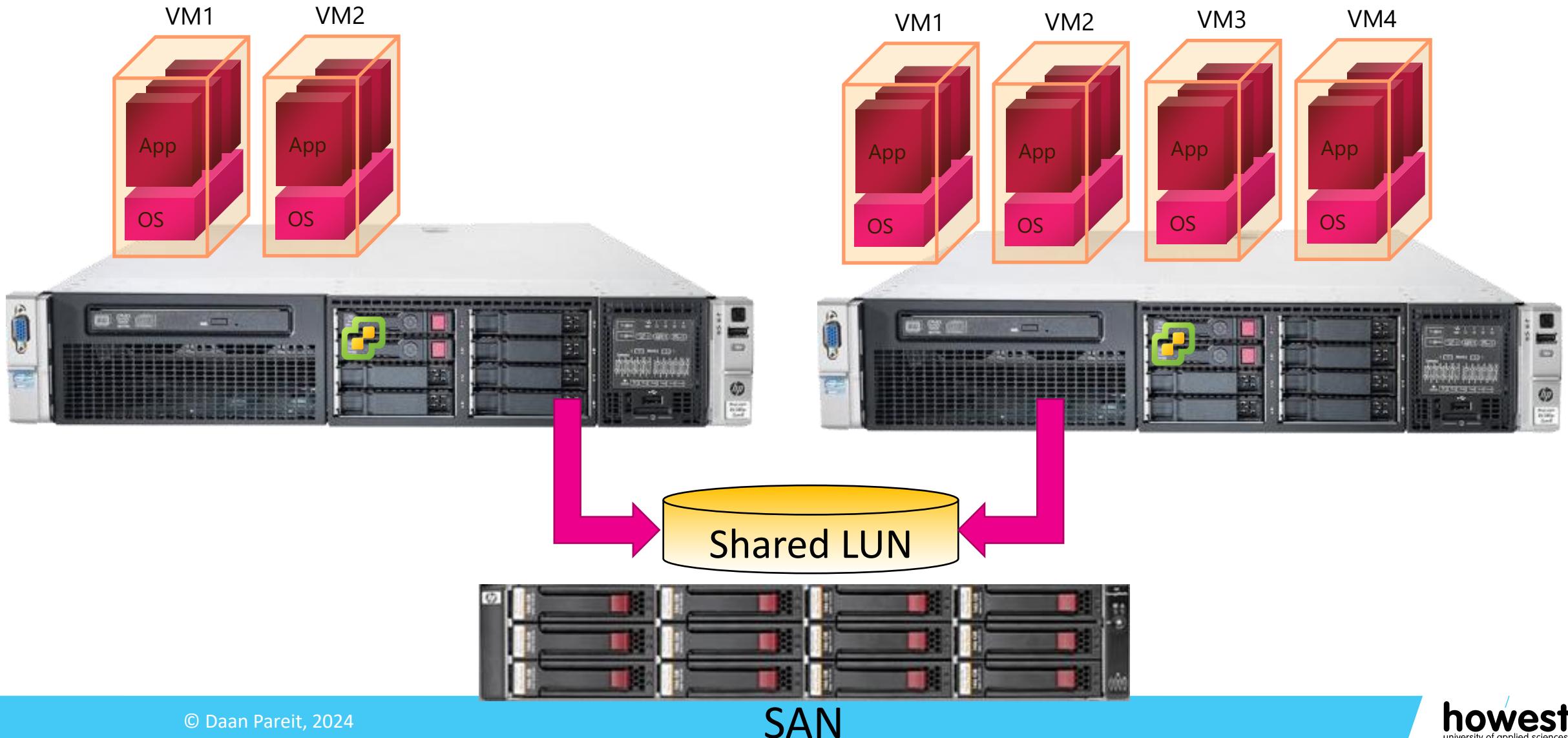


Shared Storage

- To here:

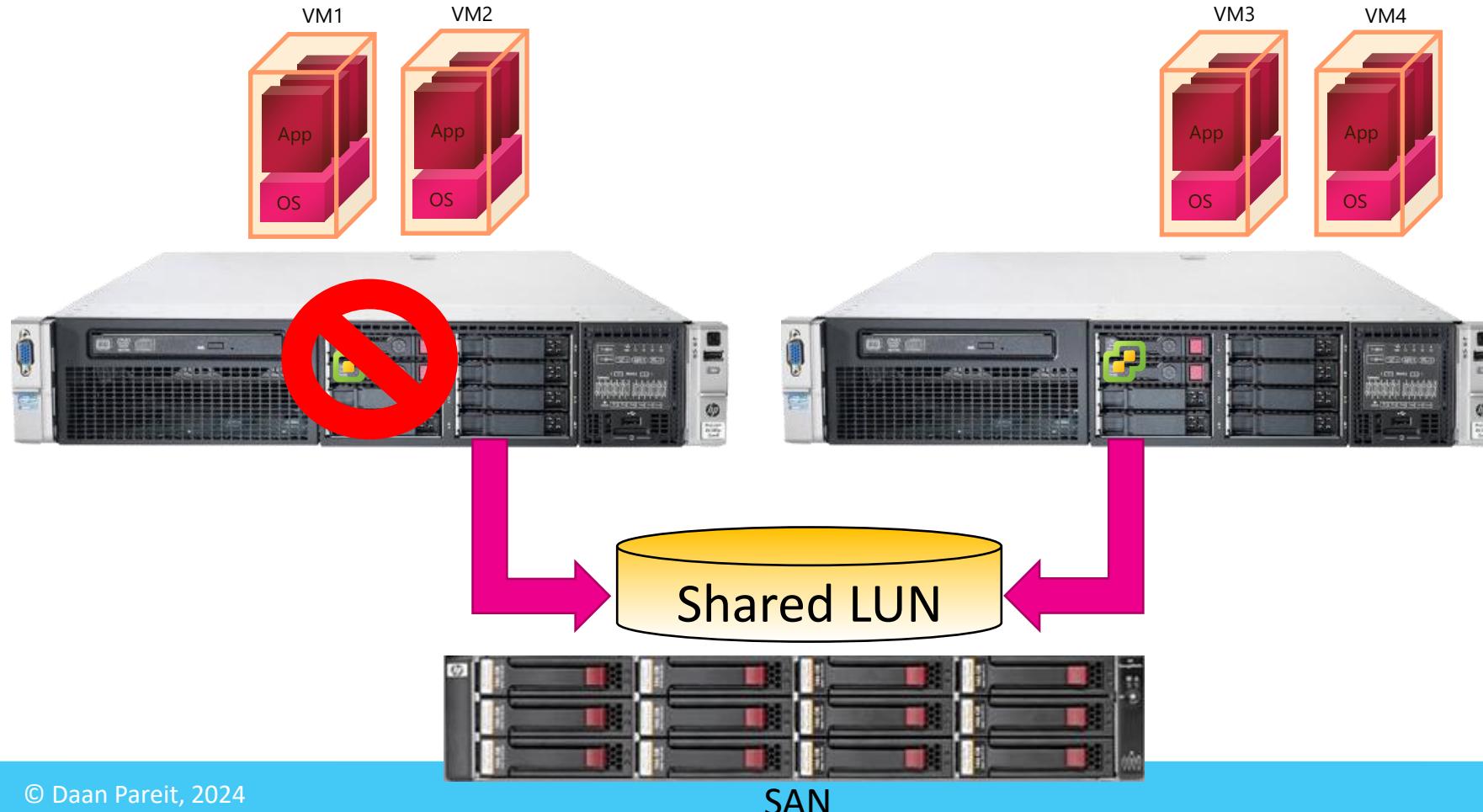


Movement 1: High Availability

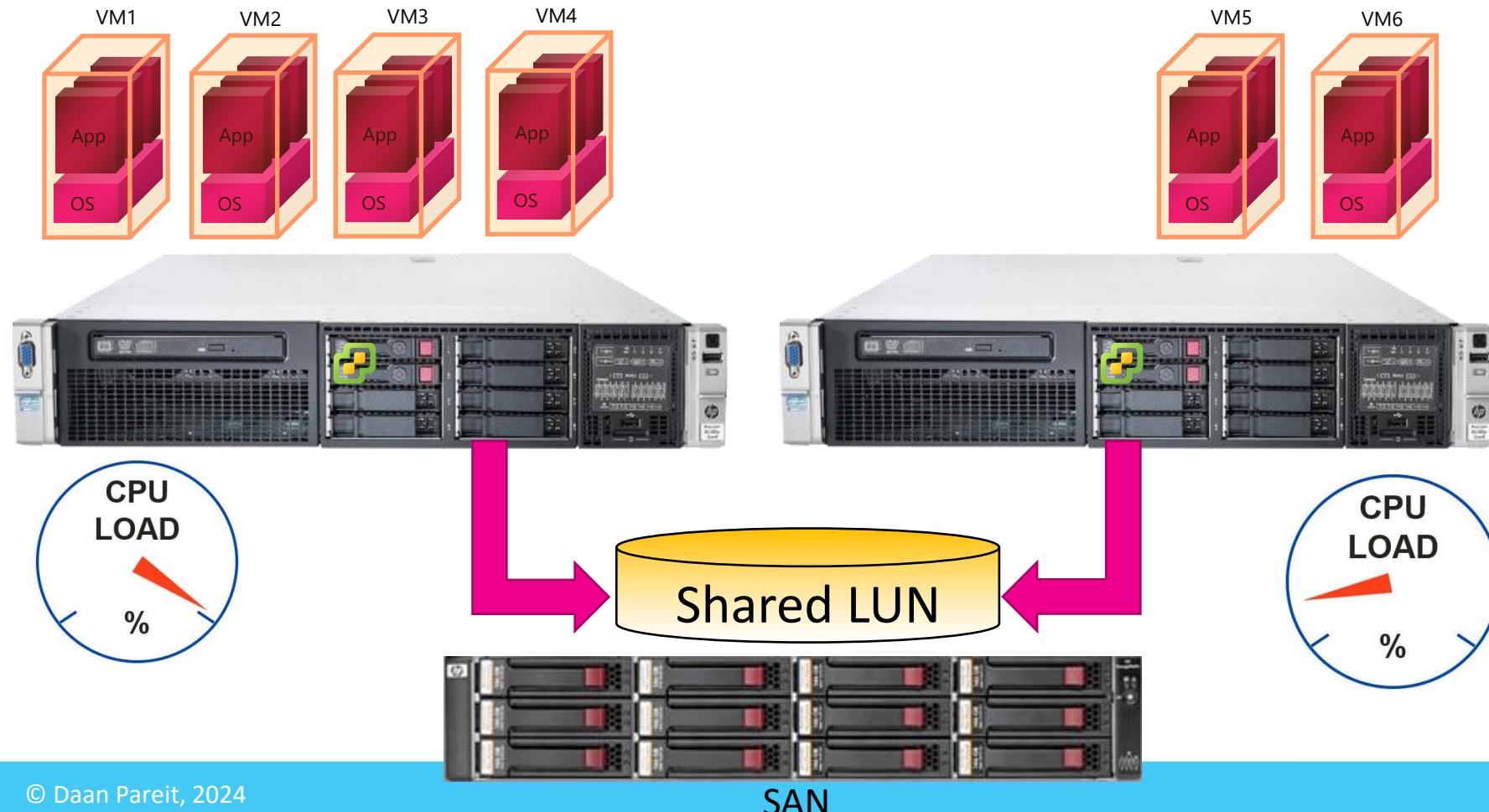


Movement 2: vMotion (~2010)

Or: *Migration without downtime*

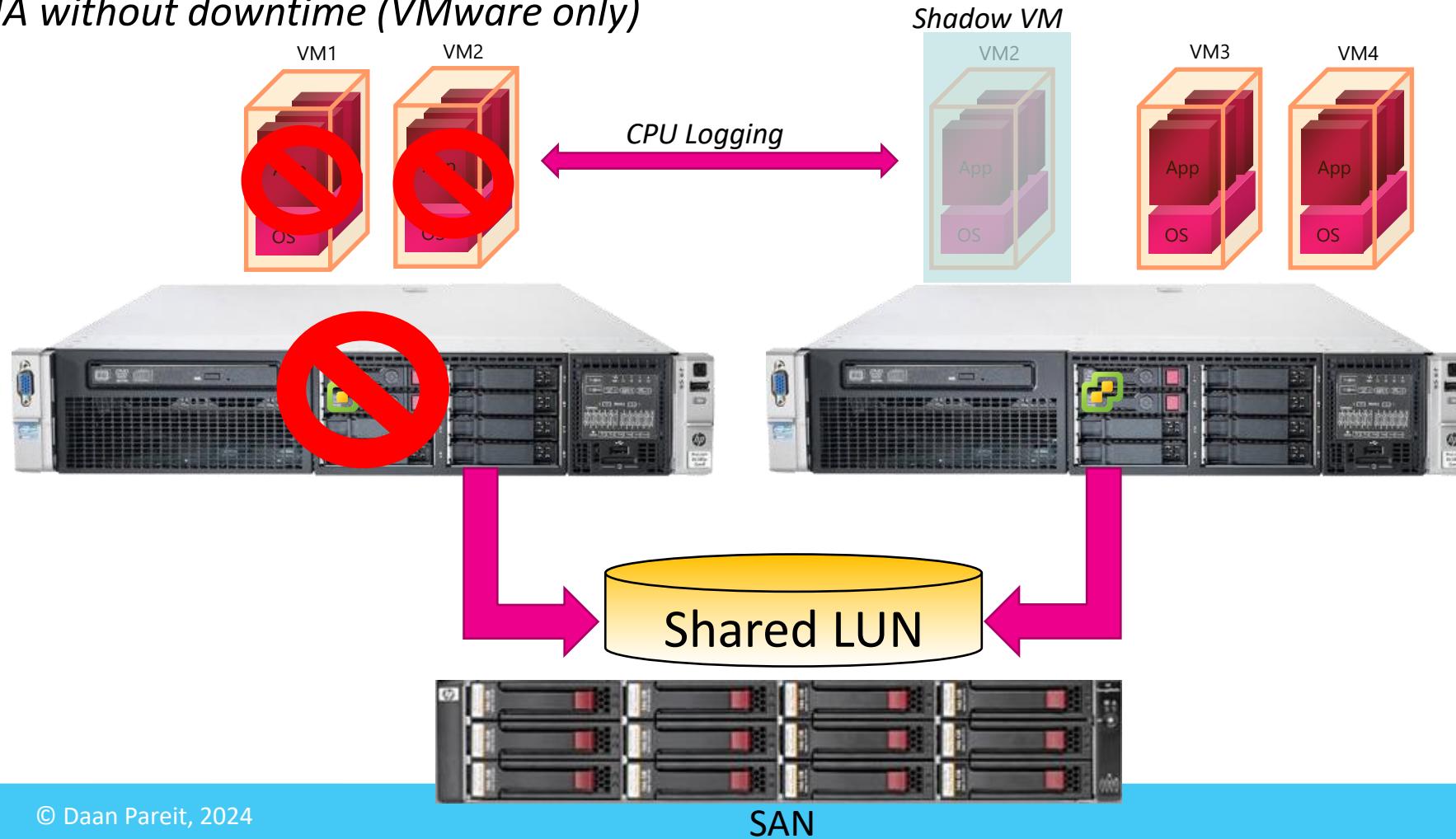


Dynamic Resource Scheduling



Movement 3: Fault Tolerance (~2012)

Or: *HA without downtime (VMware only)*



Vmware alternatives

Mostly used alternatives

- VMware workstation
 - Oracle virtual box (vm not directly accessible though NAT)
- ESXI
 - Linux KVM (also used by Oracle Linux)
 - Hyper-V (Microsoft)
 - Openshift (RedHat)
- Vsphere
 - Proxmox -> used in our Labo's
 - Openstack -> more Cloud based
- But the future : Containers

The Lab: VMware Virtualization

Hosted virtualization (Workstation / Fusion)

+

Bare-metal virtualization (ESXi)

Lab

- Remember: in datacenters, ESXi is installed on powerful servers (which meet the supported hardware compatibility list, HCL) as its dedicated operating system
- Big surprise: we can't donate a server to every student 🙀. You'll thus have to install ESXi within your Workstation/Fusion
 - Realization 1: you're installing a type 1 hypervisor within your type 2 host virtualization 🤯
 - Realization 2: you will be running VMs in ESXi, which itself is a VM in your Workstation/Fusion
→ nested virtualization
- Note: Hypervisor VMs on Workstation are only possible if Hyper-V extensions are disabled on the host
 - more details in the lab and in 'hyperv-vs-vmware.txt' on Leho