

# Virtualizace a Cloud computing

**Luboš Matějka**  
**KIV**  
**FAV ZČU Plzeň**

# Téma přednášky

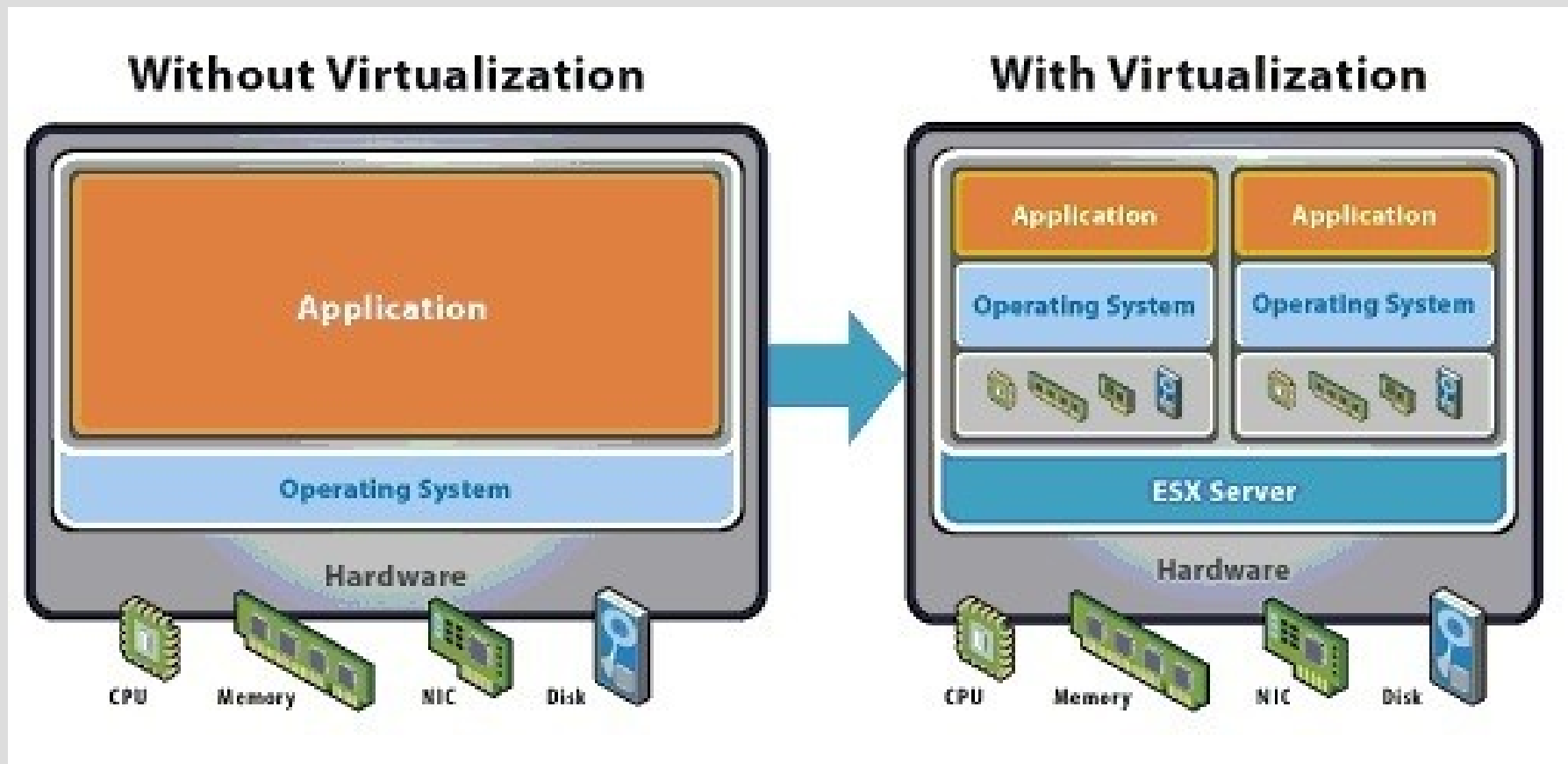
- Virtualizace
  - Historie
  - Typy virtualizací
  - Existující systémy
  - Virtualizace a cloudy
  - Podpůrné systémy
  - Praktické nasazení

# Virtualizace

- Následník emulátorů
  - Na PC můžu provozovat 8bity Atari, Commodore
- Více operačních systémů v jednom
  - Vývoj na různých platformách
  - Testování různých verzí, např IE
  - Runtime prostředí pro atypický software
  - Bezpečnost / Izolovanost
  - Škálovatelnost / Dostupnost služeb
  - Zálohování / snapshoty

# Princip Virtualizace

- Jeden HW více OS
- Zdroje se emulují nebo mapují



# Typy virtualizací 1.

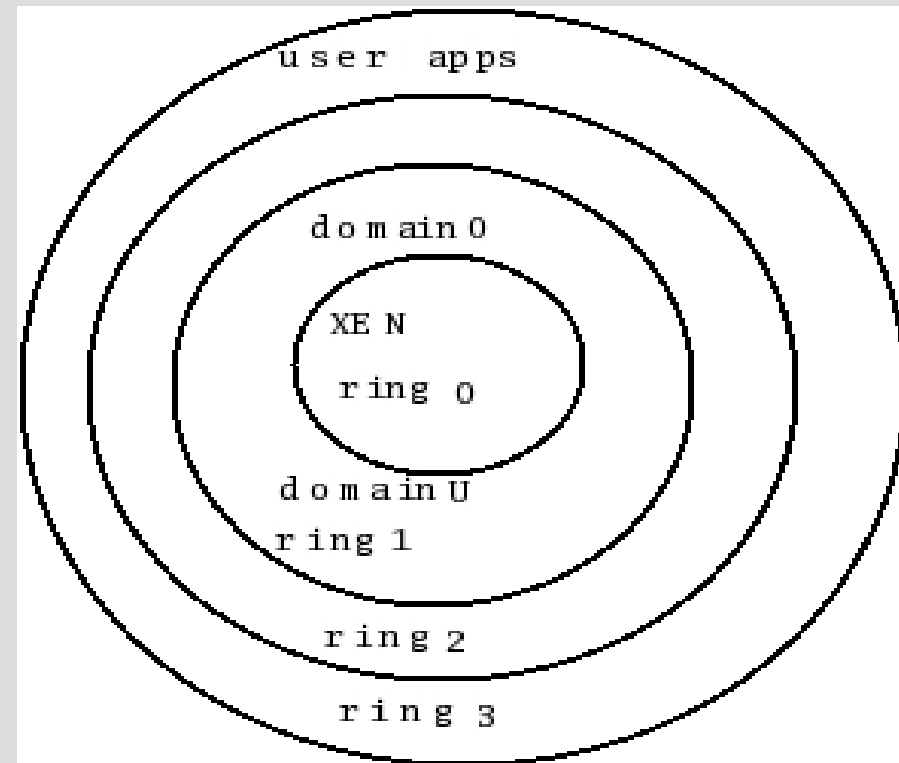
- **Emulace**
  - Veškeré požadavky se překládají
  - Můžeme emulovat téměř cokoliv
  - např. VirtualBox, Qemu
- **Aplikační**
  - Rozhraní/kontejner pro aplikace
  - Spouští speciální kód
  - JVM
- **Na úrovni OS**
  - Sdílíme jedno jádro
  - Musí přesně souhlasit verze OS
  - V zásadě se jedná o formu chrootu
  - Např vServer, BSD

# Typy virtualizací 2.

- **Plná virtualizace**
  - Vše se emulujeme, RAM, CPU, USB, ...
  - Vyžaduje podporu HW – vmx, svm
  - Nezávislost virtualizovaného OS
  - Nemodifikované OS
  - Možnost mapovat zdroje
  - Vyšší režie
  - VMware, KVM, XEN, VirtualBox
- **Paravirtualizace**
  - Vyžaduje modifikovaný kernel OS
  - Sdílí operace s jádrem hypervisoru
  - Nelze mezi platformami
  - Nejefektivnější
  - XEN

# Model virtualizace

- Ring – úroveň oprávnění / zabezpečení
- Hypervisor – ovladač HW
  - Ring -1
- Dom0 - správa virtualizací
  - Ring 0
- DomU – host
  - Ring 1
- Apps – Aplikace v DomU
  - Ring 3



# Existující systémy

- Není jen jeden virtualizační SW
- Různé úrovně podporovaných technik
- Někdy několik technologií v sobě
  - XEN + Qemu
- Různé licence a ceny
- Nejpoužívanější
  - VMware
  - XEN / KVM
  - Virtualbox
  - Hyper-V



# Virtualizace XEN

- Open Source převzatý firmou Citrix
- Dlouho nezačleněno do Kernel
  - DomU v 2.6.3X
  - Dom0 v 3.X
- Podporuje paravirtualizaci i plnou virtualizaci
- Konfigurace typicky v /etc/xen
- Příjemná obsluha pomocí xm / xl utils
  - xm list | create | destroy | pausa | migrate

selene:~# xm list

Name	ID	Mem	VCPUs	State	Time(s)
Domain-0	00	14565	16	r-----	934397.0
amos	16	1024	1	-b----	672709.1
fai	22	1024	1	-b----	53323.8
forge	11	2048	1	-b----	393803.8

# XEN - paravirtualizace

```
kernel      = '/boot/vmlinuz-2.6.26-2-xen-amd64'
ramdisk     = '/boot/initrd.img-2.6.26-2-xen-amd64'
memory      = '512'

root        = '/dev/sda2 ro'
Disk        = [ 'file:/home/xen/soubor1.img,sde,w',
                'phy:/dev/kiv.spos/spos-spos-swap,sda1,w',
                'phy:vgsp/spos-spos-disk,sda2,w',
                'drbd:drbd-disk3,sdb,w',
                'iscsi:iqn.2014-12.cz.zcu.kiv.storage1:webmath,sdb,w' ]
name        = 'spos'
vif         = [ 'bridge=brxen63,ip=147.228.63.12 ,mac=00:16:3E:FD:90:01' ]

on_poweroff = 'destroy'
on_reboot   = 'restart'
on_crash    = 'restart'
```

# XEN – plná virtualizace

```
builder='hvm'  
memory = 24576  
vcpus=58  
maxvcpus = 58;  
cpus="all"  
boot="c"  
disk = [ 'phy:/dev/mapper/mpath-herakles,ioemu:hda,w' ]  
name = 'herakles'  
vif = [ 'bridge=brxen63,ip=147.228.63.12 ,mac=00:16:3E:FD:FF:01' ]  
vnc=1  
vnclisten="127.0.0.1"  
vncconsole=1  
vncviewer=0  
vncpasswd='velice_Tajne-Hes.10'
```

# Virtualizace KVM

- Mladší alternativa pro XEN
- Jen plná virtualizace
- Nativně dostupné k kernelu
- Masivně podporované Redhatem
- Složitější ovládání z řádky
- Konfigurace v XML
- Využívá Qemu
- Většinou ovládané pomocí virsh

# KVM – spuštění virtuálu

- `qemu-system-x86_64 -enable-kvm -name windows8.1-adminit -S -machine pc-1.1,accel=kvm,usb=off -m 4096 -realtime mlock=off -smp 4,sockets=4,cores=1,threads=1 -uuid 5d28e7f6-88c9-9595-8d9b-a1ff33db4fec -no-user-config -nodefaults -chardev socket,id=charmonitor,path=/var/lib/libvirt/qemu/windows8.1-adminit.monitor,server,nowait -mon chardev=charmonitor,id=monitor,mode=control -rtc base=localtime -no-shutdown -boot strict=on -device ich9-usb-ehci1,id=usb,bus=pci.0,addr=0x5.0x7 -device ich9-usb-uhci1,masterbus=usb.0,firstport=0,bus=pci.0,multifunction=on,addr=0x5 -device ich9-usb-uhci2,masterbus=usb.0,firstport=2,bus=pci.0,addr=0x5.0x1 -device ich9-usb-uhci3,masterbus=usb.0,firstport=4,bus=pci.0,addr=0x5.0x2 -device ahci,id=ahci0,bus=pci.0,addr=0x6 -drive file=/dev/vgnero/kvm-windows8.1-adminit,if=none,id=drive-sata0-0-0,format=raw -device ide-hd,bus=ahci0.0,drive=drive-sata0-0-0,id=sata0-0-0,bootindex=1 -drive file=/mnt/DATA1/Download/ISO/Microsoft/Windows-8.1-CZ-64b-Prof.iso,if=none,id=drive-ide0-1-0,readonly=on,format=raw -device ide-cd,bus=ide.1,unit=0,drive=drive-ide0-1-0,id=ide0-1-0,bootindex=2 -netdev tap,fd=24,id=hostnet0 -device e1000,netdev=hostnet0,id=net0,mac=52:54:00:8c:cf:de,bus=pci.0,addr=0x3 -chardev pty,id=charserial0 -device isa-serial,chardev=charserial0,id=serial0 -device usb-tablet,id=input0 -vnc 127.0.0.1:0 -device VGA,id=video0,bus=pci.0,addr=0x2 -device virtio-balloon-pci,id=balloon0,bus=pci.0,addr=0x4 -msg timestamp=on`

# Virtualizace - LibVirt - VirSH

- Více virtualizačních platforem
  - V jedné společnosti ...
  - Na jednom stroji ...
- Jednotný nástroj na správu virtualizací
  - XEN, KVM, Qemu, OpenVZ, VirtualBox, ...
  - Filosofii i ovládáním vychází z xm utils

```
root@nero:~# virsh list
```

Id	Name	State
2	windows8.1-adminit	running

# Virtualizace Hyper-V

- Virtualizační platforma od Microsoftu
- Dostupné jako
  - Součást serverů od MS Server 2008R2
  - Samostatně běžící hypervisor
- Svázané s Microsoftem - Active Directory, Share,.....
- Plná virtualizace – možnost Linuxu, BSD, .....
- Grafické ovládání
- Podporuje formát image z Azure

# Hyper-V - ukázka


The screenshot displays the Hyper-V Manager interface. The main window shows a list of virtual machines with columns for Name, State, CPU Usage, Memory, and Uptime. The 'WIN7 - VS2008' VM is selected and highlighted. Below the list, the 'Snapshots' section indicates that the selected VM has no snapshots. The 'Details' section for 'WIN7 - VS2008' shows a thumbnail of the VM's desktop, along with its creation date (17.10.2009 16:50:51), heartbeat status (OK), and memory usage (1024 MB). The right-hand pane shows the 'Actions' menu for the selected VM, including options like 'New', 'Import Virtual Machine...', 'Hyper-V Settings...', 'Virtual Network Manager...', 'Edit Disk...', 'Inspect Disk...', 'Stop Service', 'Remove Server', 'Refresh', 'View', 'New Window from Here', 'Help', 'Connect...', 'Settings...', and 'Turn Off...'.

Name	State	CPU Usage	Memory	Uptime
UBUNTU9	Running	0 %	1024 MB	00:05:36
UBUNTU9 - BASE	Off			
WIN2008R2 - BASE	Off			
WIN2008R2 - BTS2009	Off			
WIN2008R2 - SP2007	Off			
WIN2008R2 - SP2010	Off			
WIN7 - BASE	Off			
WIN7 - DESKTOP	Running	0 %	1024 MB	00:38:36
WIN7 - VS2008	Running	4 %	1024 MB	00:10:36
WIN7 - VS2010	Off			

**Snapshots**

The selected virtual machine has no snapshots.

**WIN7 - VS2008**

 **Created:** 17.10.2009 16:50:51      **Heartbeat:** OK  
**Notes:** None      **Memory Usage:** 1024 MB

**Actions**

**GHOST**

- New
- Import Virtual Machine...
- Hyper-V Settings...
- Virtual Network Manager...
- Edit Disk...
- Inspect Disk...
- Stop Service
- Remove Server
- Refresh
- View
- New Window from Here
- Help

**WIN7 - VS2008**

- Connect...
- Settings...
- Turn Off...



# Virtualizace VMware

- Lídr v oblasti virtualizace
- Dostupné jako free i placená verze
- Support a certifikace
- Součást jiného OS – VMwarePlayer
- „Samostatný“ hypervisor – VMware ESX
- Jednotlivé verze podle licence mají omezené možnosti
- Postavené na – Redhat Linux
- Umožňuje sdílet paměť!!!
- Velké množství modulů a rozšíření – velmi drahé
  - HA availability
  - Cluster / vCenter
  - vSan
- Elegantní zálohování Veem

# VMware - vCenter

VC2.RWVDEV.INTRA - vSphere Client

File Edit View Inventory Administration Plug-ins Help

Home Inventory Hosts and Clusters Search Inventory

VC1.RWVDEV.INTRA VMware vCenter Server, 4.0.0, 208111

Getting Started Datacenters Virtual Machines Hosts Tasks & Events Alarms Permissions Maps

**What is the Hosts & Clusters view?**

This view displays the set of computing resources that run on a particular host, cluster, or resource pool. Using the Hosts & Clusters view, you can manage and organize your inventory of computing resources.

**Basic Tasks**

Create a datacenter

Hosts & Clusters view

Explore Further

Learn more about inventory views

**Recent Tasks**

Name	Target	Status	Details	Initiated by	vCenter Server	Requested Start Time	Start Time	Completed Time
Rename resource pool	RP-Tier5	Completed		RWVDEV\Ad...	VC1.RWVDEV...	12/18/2009 2:00:58 ...	12/18/2009 2:00:58 ...	12/18/2009 2:00:58 ...
Power On virtual mach...	VM-SERVER6	Completed		System	VC1.RWVDEV...	12/18/2009 2:00:14 ...	12/18/2009 2:00:14 ...	12/18/2009 2:00:14 ...
Initialize powering On	RWVDEV1	Completed		RWVDEV\Ad...	VC1.RWVDEV...	12/18/2009 2:00:14 ...	12/18/2009 2:00:14 ...	12/18/2009 2:00:14 ...

Tasks Alarms 2 vCenter Server licenses expire in 60 days RWVDEV\Administrator

Start 2:04 AM 12/18/2009

# VMware – zálohování Veem

The screenshot displays the Veeam Backup & Replication console interface. The top navigation bar includes 'Home' and 'Backup' tabs. Below the navigation bar is a toolbar with icons for 'Instant VM Recovery', 'Entire VM', 'VM Files', 'Guest Files', 'Application Items', and 'Remove from Disk'. The main area is divided into a left-hand navigation pane and a central job list.

**Backup & Replication**

- Jobs
  - Backup
  - SureBackup
  - Replication
  - VM Copy
- Backups
  - Disk
  - Imported
  - SAN
- Replicas
- Last 24 hours

**Virtual Machines**

**Files**

**Backup Infrastructure**

**SAN Infrastructure**

**Job List:**

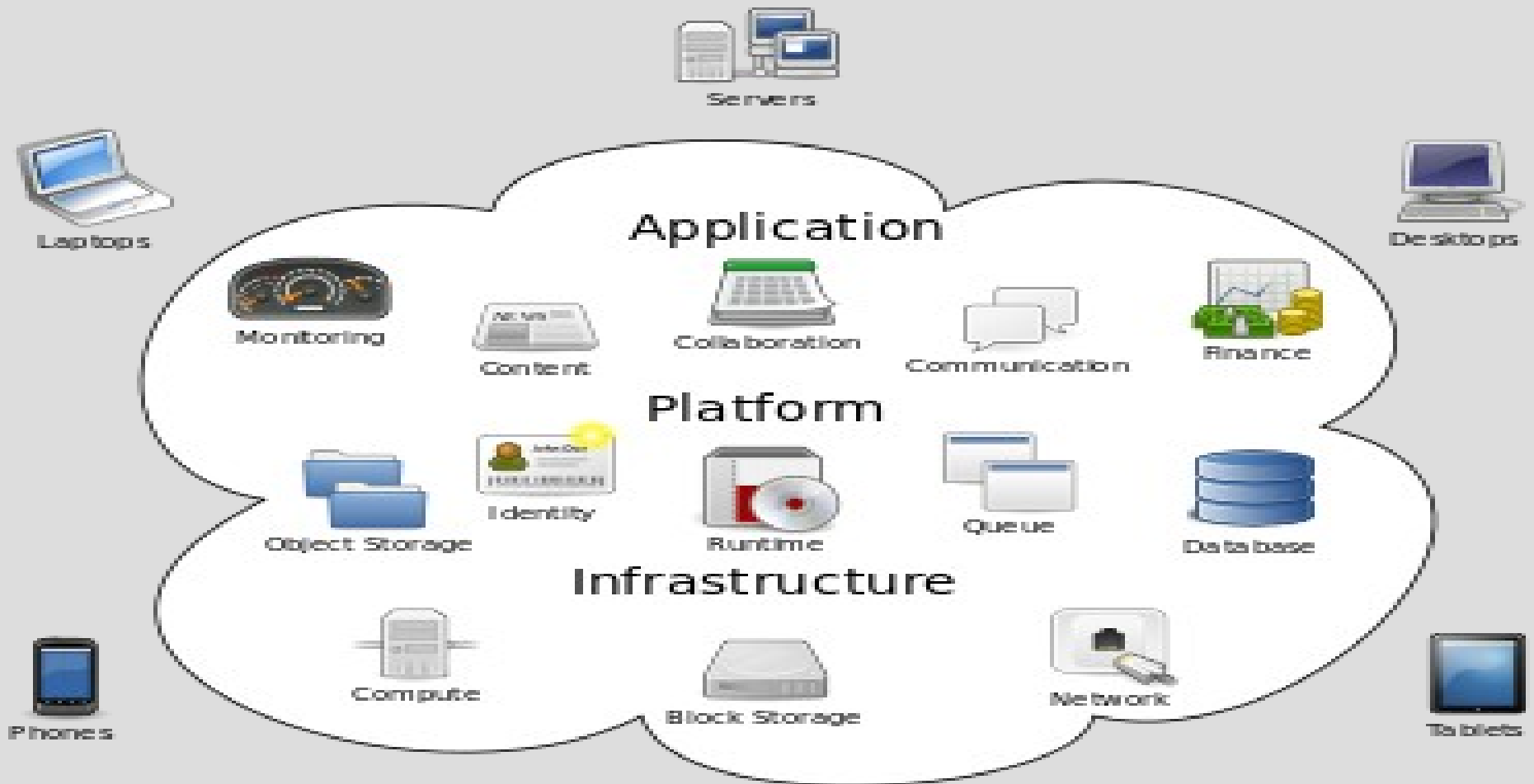
Job name	Creation time	Restore points	Repository
Exchange Servers	9/8/2012 2:00 AM		Backup Share
exch01	9/8/2012 3:00 AM	12	
exch02			
File server backup			Backup Share
Intranet Servers			Backup Share

**Context Menu for exch02:**

- Instant VM recovery...
- Restore entire VM...
- Restore VM files...
- Restore VM hard disks...
- Restore guest files (Windows)...
- Restore guest files (Other OS)...
- Restore Exchange items...
- Remove from Disk

1 backup selected | License: Enterprise, Period: 275 days remaining | Veeam

# Cloud Computing



Cloud Computing

# Cloud Computing - Virtualizace

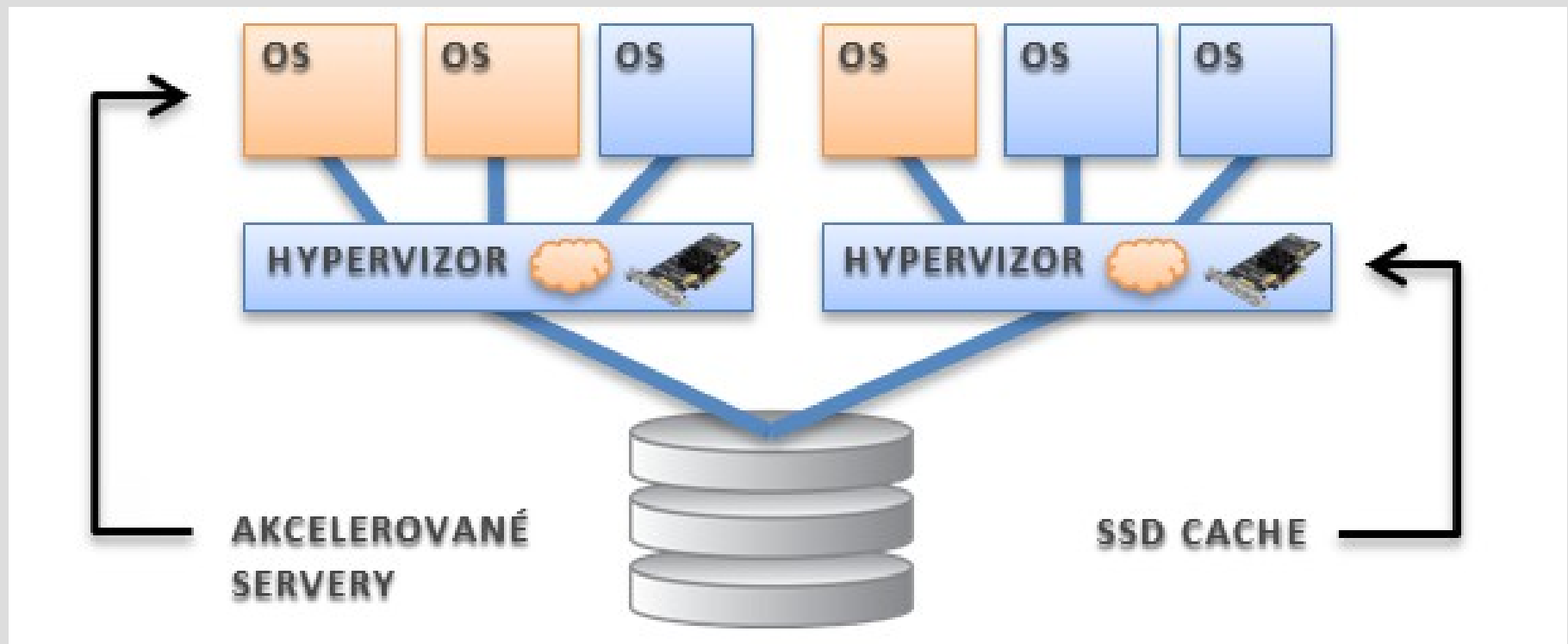
- Základní myšlenky
  - Úspora nákladů
  - Efektivní využití zdrojů
  - Levnější a rychlejší modifikace zdrojů
    - SW i HW
  - Využívat jen to, co opravdu NYNÍ potřebuji
  - Platit jen to, co jsem opravdu použil
  - Ideově vychází a využívá virtualizaci
    - Škálovatelnost
    - Dostupnost – stabilita
      - FailOver / Balancer

# Cloud Computing - Škálovatelnost

- „Nafukování služeb“
  - virtualizace
- Postupné spouštění služeb
  - Začnu s nejmenší možnou jednotkou
  - Při vzrůstající zátěži „startuji“ další nody
  - Nody mohou být různě výkonné
  - Když odpadne zátěž, stroje se zastaví
  - Nutnost optimalizace aplikací
  - Decentralizované prostředí
    - Úložiště - dostupnost všech dat na všech nodech
    - Session - synchronizace envirementu webových app
    - Databáze – nutnost škálovat do šířky, NON-SQL DB

# Virtualizace - Live Migrace

- Možnost za chodu přesunout běžící OS
- Krátký výpadek pro externí pozorovatele
- Nulový výpadek pro vnitřní systém



# Stavební kameny Cloudů

- „Výkonný HW“
- Virtualizace OS / Aplikací
- Balancing / FailOver
  - Síťových cest
  - Datových uložišť



# FailOver / Balancer

- FailOver
  - Horká / studená záloha
  - Problém identické konfigurace
  - Problém rozpoznání výpadku a přepnutí
- Balancer
  - Násobně spuštěný server / služba / aplikace
  - Balancer – aplikace/HW pro rozdělování provozu
  - Může vyžadovat úpravu balancovaného zdroje
  - Umožňuje škálovat do šířky

# Sítě v cloudu

- NAT
  - Nejjednodušší, překlad adres
- Bridge
  - Virtuální interface, přímý přístup k síti
  - Podpora VLAN – dnes nutnost

```
auto brxen67
iface brxen67 inet static
    address 0.0.0.0
    bridge_ports eth5.67
    bridge_stp off
    vlan-raw-device eth5
```

```
root@xen2:~# brctl show
```

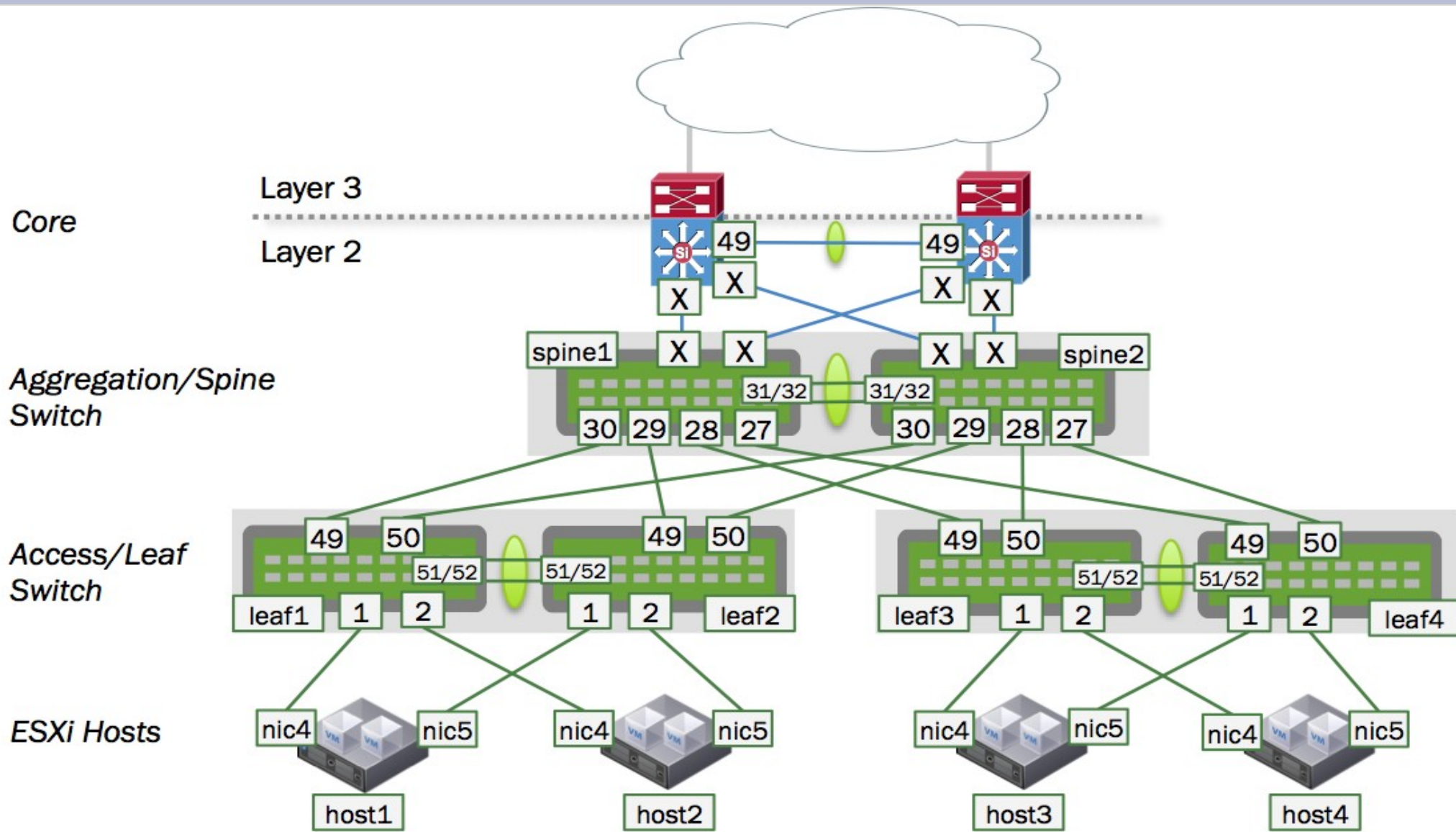
bridge name	bridge id	STP enabled	interfaces
brxen63	8000.000e1e72bd42	no	eth5 vif10.0 vif9.0
brxen67	8000.000e1e72bd42	no	eth5.67

# Bond

- Více síťových karet v serveru
- Více switchů
- Bond – slučování síťových karet
- Balancer i FailOver

```
auto bond0
iface bond0 inet static
    address 192.168.2.27
    netmask 255.255.255.0
    network 192.168.2.0
    slaves eth0 eth1
    bond_miimon 100
    bond_mode 0
```

# Bond - zapojení



# Datová uložště

- Nejzásadnější prvek
- Pokud mám data už se vždy dá něco dělat

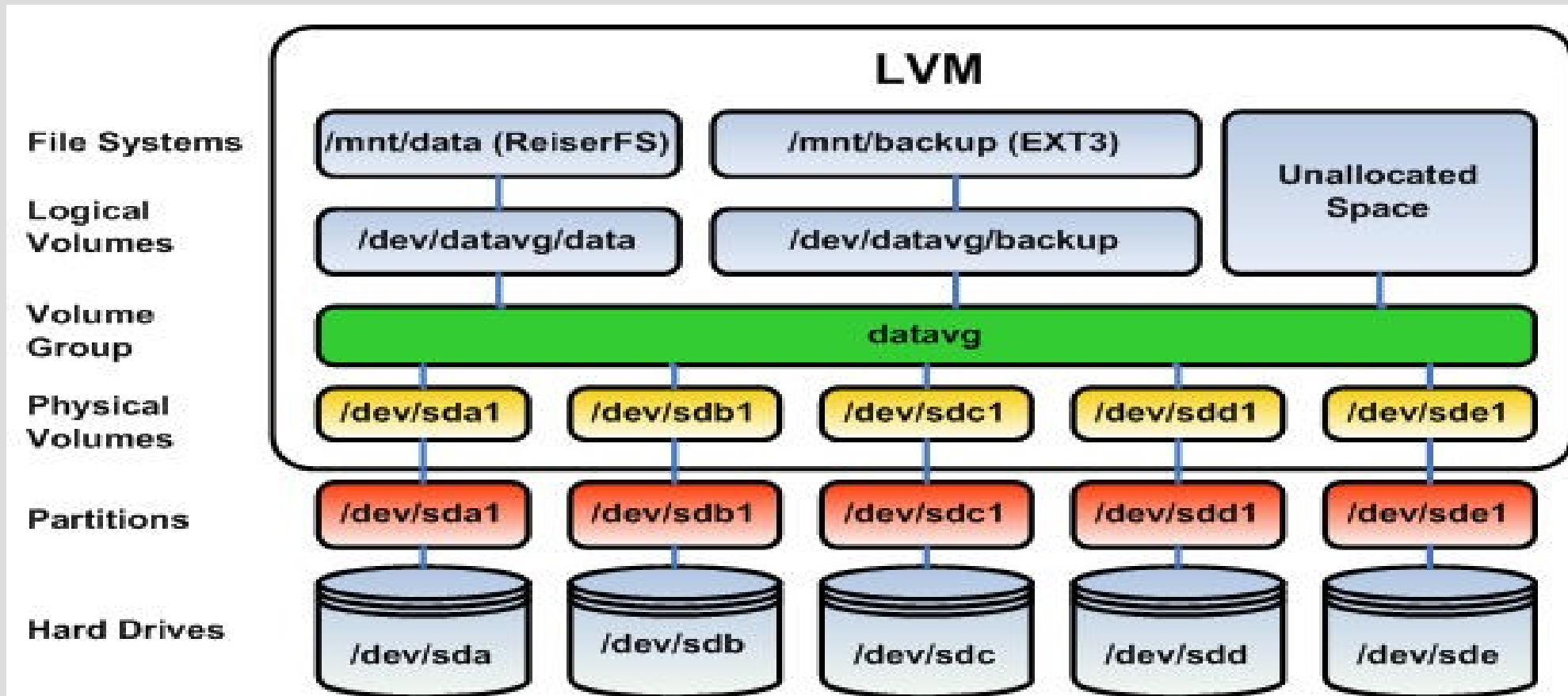
```
Disk = [ 'file:/home/xen/soubor1.img,sde,w'  
        'phy:/dev/kiv.spos/spos-spos-swap,sda1,w',  
        'phy:vgsp/spos-spos-disk,sda2,w',  
        'drbd:drbd-disk3,sdb,w',  
        'iscsi:iqn.2014-12.cz.zcu.kiv.storage1:webmath,sdb,w' ]
```

# Lokální datové uložení

- Soubor
  - file://
- Disk – IDE / SATA / SAS / SSD / RAID
  - phy://
  - drbd://
- Datové disky vždy v raidu 1/5/6/10
- LVM – Logical volume management

# LVM

- Logické členění pevných disků
- Možnost za chodu dynamicky měnit partition

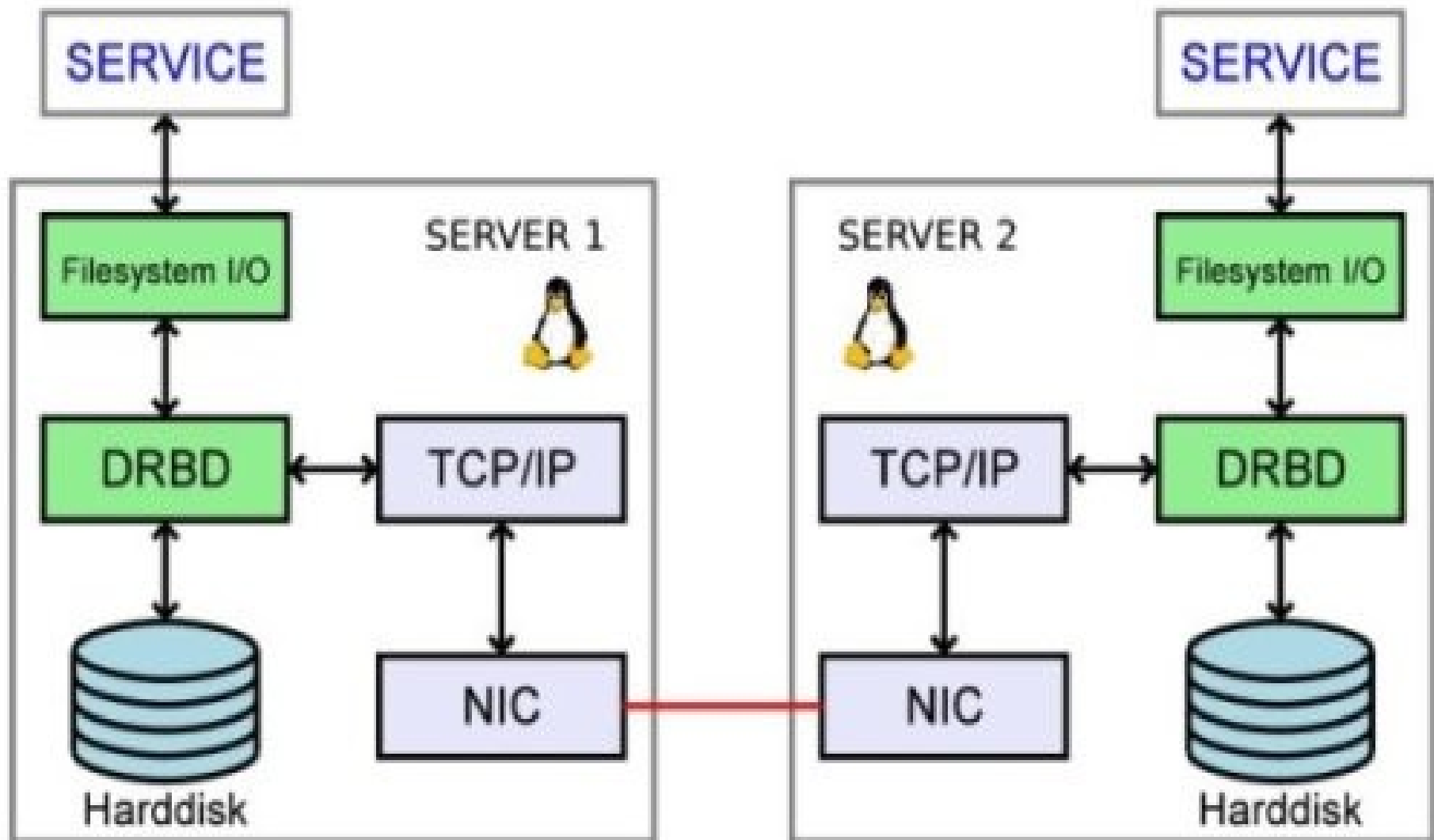


# Datové uložště DRBD

- V zásadě Raid1 po síti
- Primary / Slave
- Primary / Primary
- Kontrola kolize operací operací
- Problém zpoždění a routování v síti
- Vyhrazená síť
- Jumbo Frame MTU – 9000
- Umožňuje live migraci



# DRBD zapojení



# Sít'ová datová uložiště

- NFS
  - Jednoduché a rychlé
  - Sdílení souborů
- Fibre Channel
  - Speciální disky, case, switch HBA
  - Optické spoje
- iSCSI
  - Alternativa k drahému Fibre Chanelu
  - Postavitelné nad libovolnou sítí
  - Je součástí nových serverů
  - Postavitelné nad Linuxem
- Nutnost počítat s nestabilitou sítě

# Multipath

- Rozšíření síťového datového uložště
- Využívá více fyzických cest v síti
- Slučuje více STEJNÝCH blokových zařízení
- Možnost FailOver i Balancer
- Problém zpoždění v síti

```
mpath-webmath (3300000005851bdcf) dm-13 FreeBSD,iSCSI DISK
size=34G features='0' hwhandler='0' wp=rw
|-+- policy='service-time 0' prio=1 status=active
|  `- 32:0:0:0 sdaa 65:160 active ready running
`-+- policy='service-time 0' prio=1 status=enabled
   `- 33:0:0:0 sdab 65:176 active ready running
```

# Typické zapojení

- Univerzální zapojení neexistuje
- Podle potřeb a možností se kombinuje
- Vysoký stack => pevná pravidla operací
- Jednotlivé entity se aktivně brání chybě
- Raději nedostupnost než přijít o data

# Cloud Computing – Příklady dostupných řešení

- AWS
- Google
- Microsoft Azure
- Virt.io
- Virtual Master
- .....

# Virtualizace – Pro a Proti

- Pro +++
  - Lepší využití zdrojů
  - Lepší škálovatelnost
  - Bezpečné oddělení služeb / zákazníků
  - Dostupnost více platforem na jednom HW
- Proti ---
  - Další nutný SW – vrstvení požadavků
  - Snížení výkonu v jednotkách procent
  - Vysoké nároky na I/O
  - „Vzájemná interakce virtuálů“
  - CENA při dlouhodobém provozu

**Děkuji Vám za pozornost**

Vaše Dotazy ???