

Cloud1

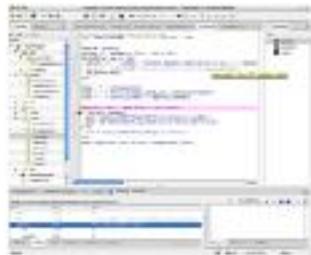
OpenNebula – KVM

Cloud Computing (VMware)



Application/Information (SaaS)

Sometimes referred to as Software-as-a-Service, a wide ranging services delivered via varied business models normally available as public offering.



Development (PaaS)

Sometimes referred to as Platform-as-a-Service, application development platforms enable application authoring and runtime environment.



Infrastructure (IaaS & ITaaS)

Sometimes referred to as elastic compute clouds or Infrastructure-as-a-Service, virtual hardware made available for varied uses.



Private

Behind a firewall for use by limited, pre-determined audience

Hybrid

Community Clouds or Managed Private Hosted Clouds

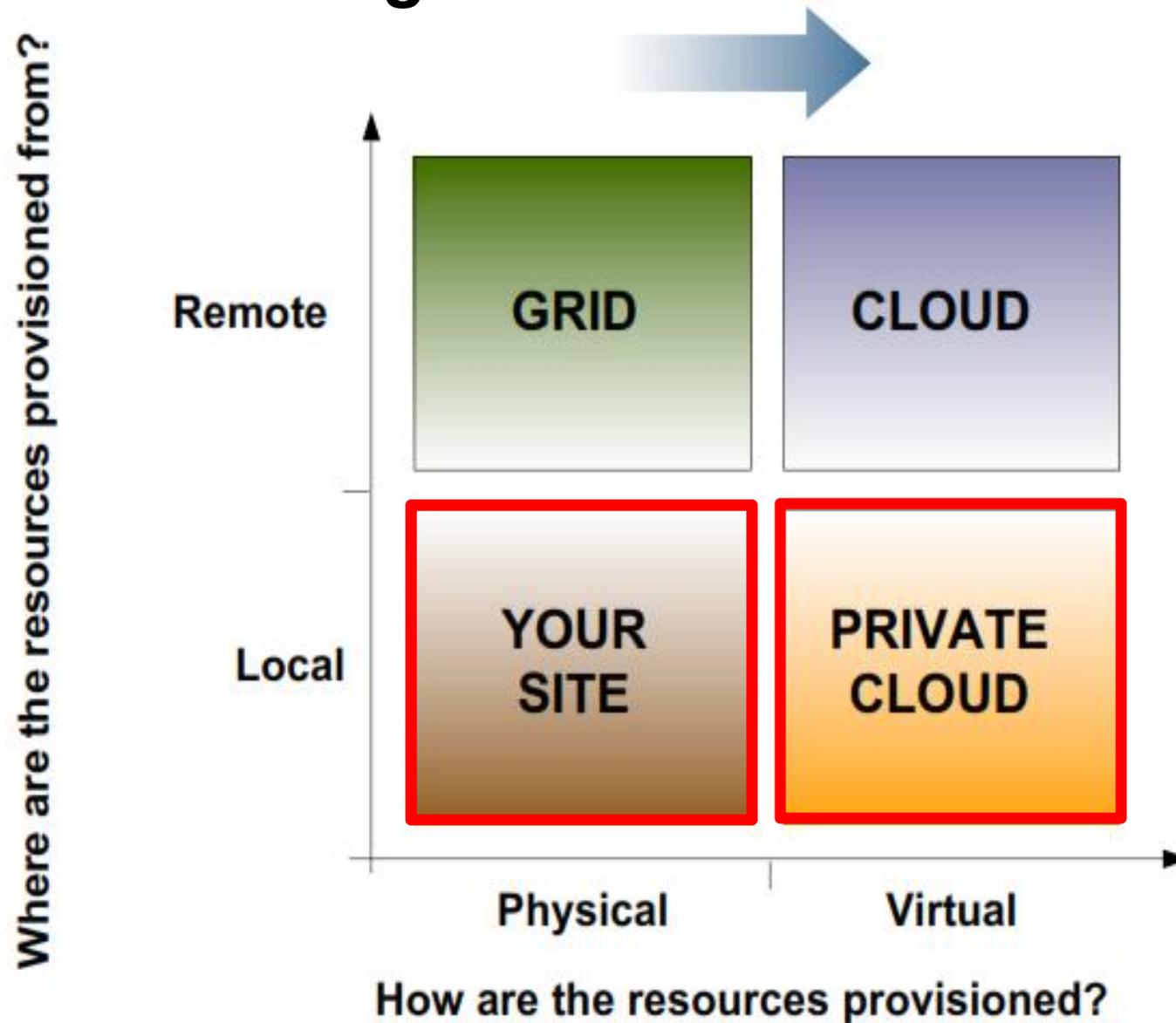
Public

Accessible over the Internet for general consumption

Capex / Opex

- Termes bien connus des comptables et des contrôleurs financiers
- **Capex** (*Capital Expenditure*) représente les dépenses **d'investissement** consacrées à l'achat de biens et d'équipements
- **Opex** (*Operating Expenses*) se réfère à l'ensemble des dépenses de **fonctionnement** destinées à exploiter un produit ou un système
- Imprimante
montant investi → Capex
coût annuel de papier et de toner consommés → Opex
- http://fr.wikipedia.org/wiki/CAPEX_-_OPEX

Resource Provisioning Models



- http://dsa-research.org/lib/exe/fetch.php?media=people:ruben:xteemeos_opennebula.pdf

Infrastructure matérielle SIACG utilisée par Korso

- Hyperviseurs → 4 x Cisco UCS B200 M3 Blade Server



80 pCPU – 2.8 = 223 GHz, 1 TB pRAM, 1 TB SSD en RAID5
200 VMs Win7

- SAN EMC-VNX = 2 systèmes RAID5

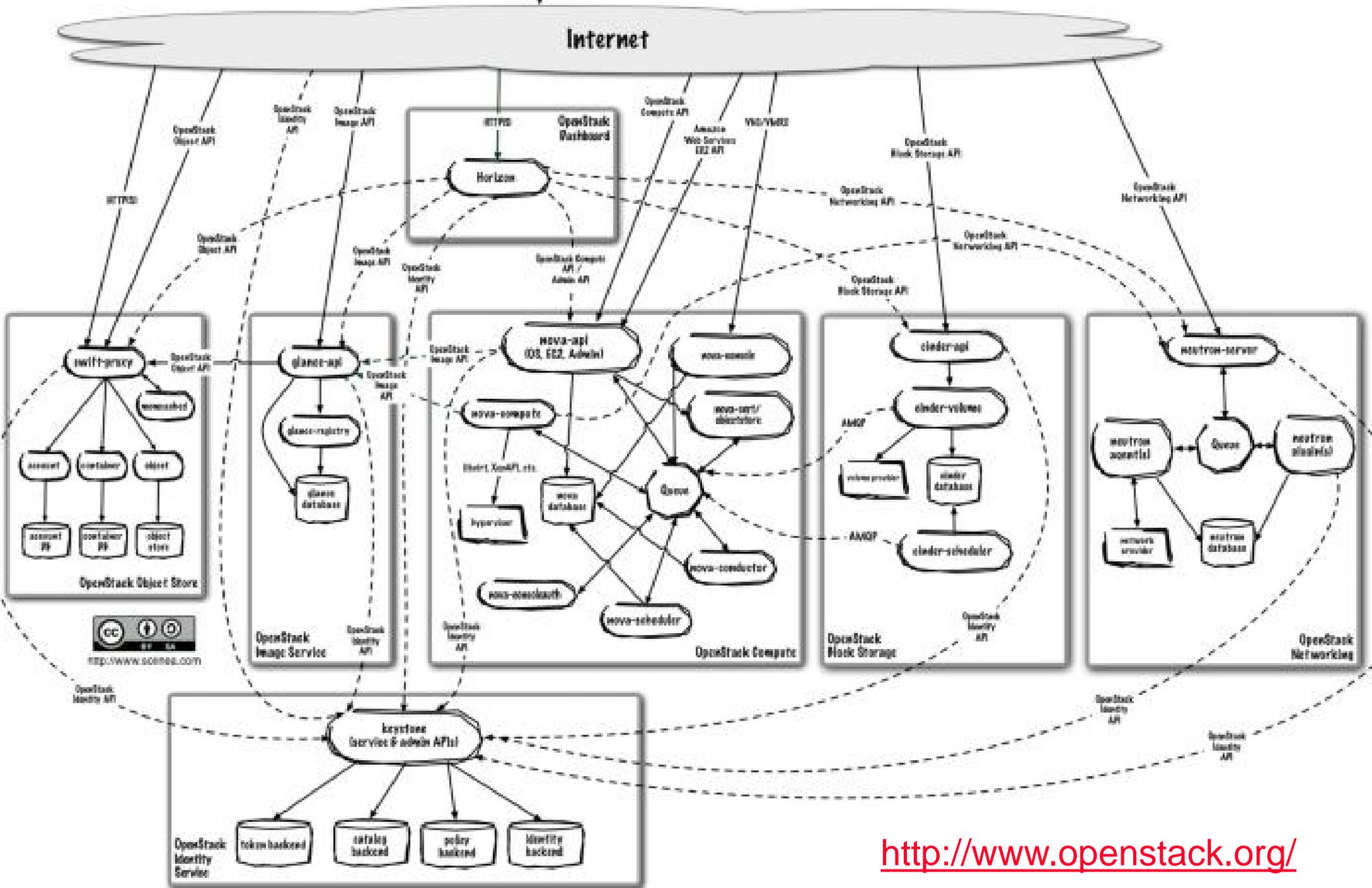
5 x SAS 600 GB = 3 TB = 2 x LUN (1.5 TB + 0.5 TB)

5 x SSD 200 GB = 1 TB

VMware Datastore

	vnx-lun160-Server	511.75 GB
	vnx-lun161-Lclone	1.53 TB
	vnx-lun170-Master	149.75 GB
	vnx-lun171-Profils	149.75 GB
	vnx-lun172-Replicas	299.75 GB

OpenStack Architecture



Project Name	Description	Layer	AWS Equivalent	Codename
Dashboard	Self-service, role-based web interface for users and administrators	UI	Console	Horizon
Compute	Provision and manage large pools of on-demand computing resources	Elastic Service	EC2	Nova
Block Storage	Volumes on commodity storage gear, and drivers for turn-key block storage solutions	Elastic Service	EBS	Cinder
Object Storage	Petabytes of reliable storage on standard gear	Elastic Service	S3	Swift
Networking	L2-focused on-demand networking with some L3 capabilities	Elastic Service	VPC	Quantum
Orchestration	Application orchestration layer that runs on top of and manages OpenStack Compute	Elastic Service	CloudFormation, CloudWatch	Heat
Metering	Centralized metering data for all services for integration to external billing	Shared Service	N/A	Ceilometer
Identity	Multi-tenant authentication system that ties to existing stores (e.g. LDAP) and Image Service	Shared Service	None	Keystone
Image Management	Upload, download, and manage VM images for the compute service	Shared Service	VM Import/Export	Glance

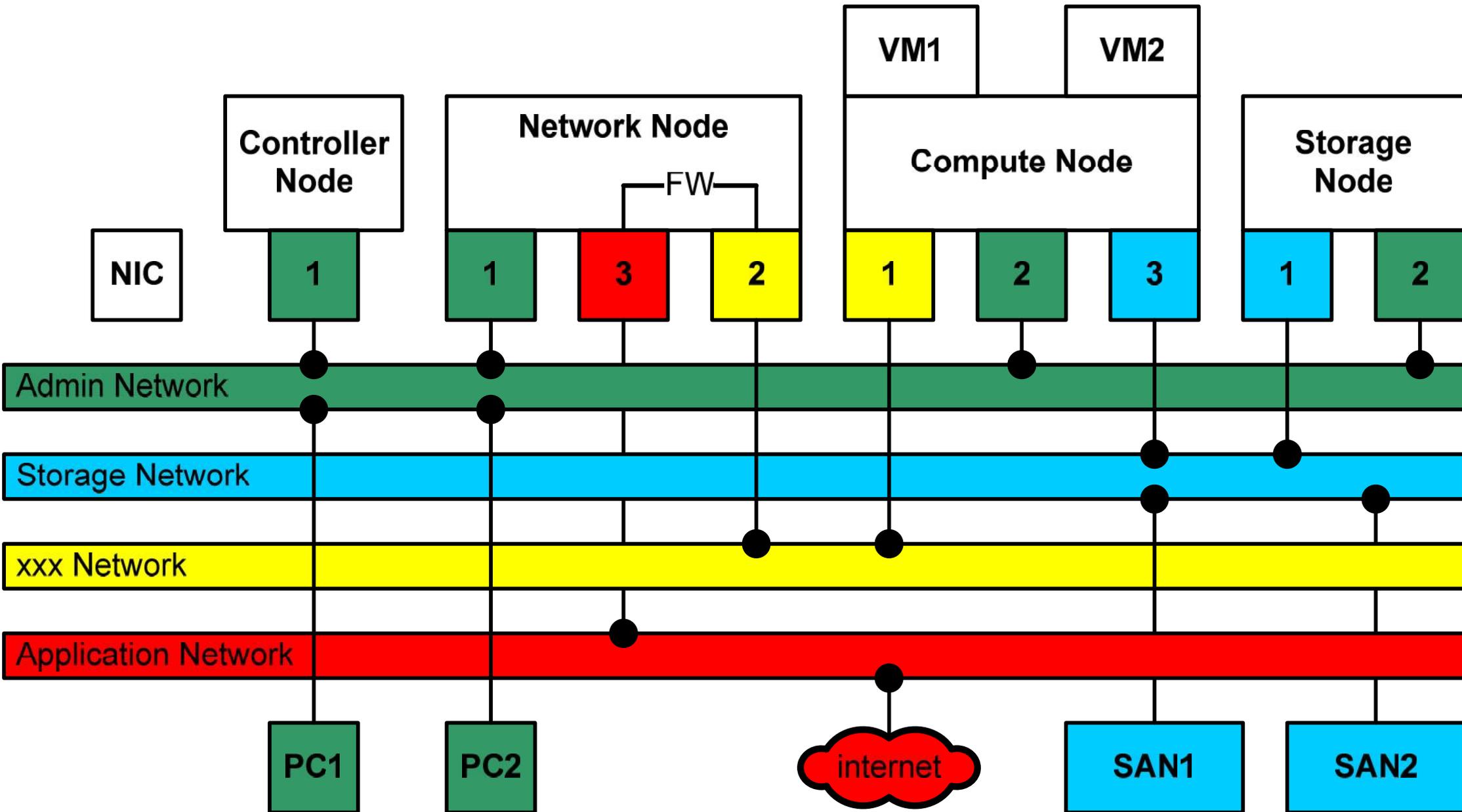
What's in a Complete Cloud ? (Randy Bias @ Cloudscale)

	OpenStack Relationship	Who?	OpenStack Score	Ecosystem Score	Vendor Target*
User Interface(s)	Horizon, CLI, ...	OpenStack, Vendor, Ecosystem	4	2	4
Elastic Resource Management	Nova, Swift, Quantum, Cinder, ...	OpenStack, Vendor, Ecosystem	4	1	4
Service Discovery	Scattered: Nova, Keystone, ...	OpenStack, Vendor	2	0	4
Authentication, Authorization, and Access Controls (AAA)	Keystone [authn/author], various projects [ACLs]	OpenStack	2	0	4
HW/SW Life Cycle Management	N/A	Vendor, Ecosystem	1	2	4
Service Management	N/A	Vendor	0	1	4
Health & Logging	N/A	Vendor, Ecosystem	0	0	4
Topology & Inventory	N/A	Vendor	0	0	4
Hardware Certifications	N/A	OpenStack, Vendor, Ecosystem	1	1	2**

* We're all trying to close this gap

** It's a hard problem no one will solve individually

Nodes & Networks (Physical View of TM Chalut 2014)



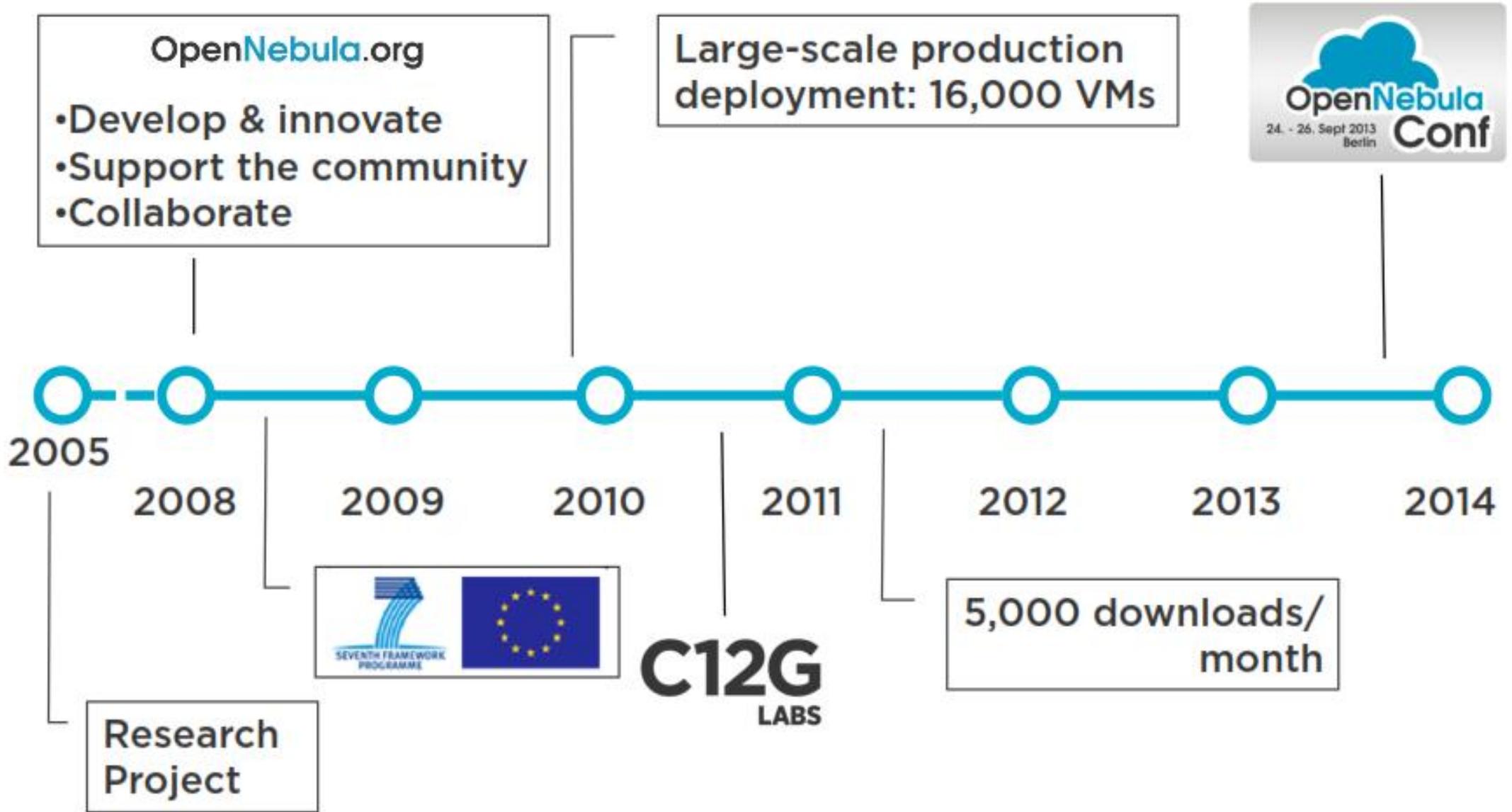
Caractéristiques du TM Chalut 2014

- Architecture distribuée (imposée par OpenStack)
- Cloisonnement (4 réseaux physiques, firewall, ...)
- Stockage (VM, données professionnelles, ...)
- **Haute disponibilité**
- http://www.tdeig.ch/linux/Chalut_RTM.pdf

Cloud : risques

- Aspects marketing : vos données vous suivent, vous ne payez que ce que vous consommez, ... nuage magique
- Principaux risques
 - Où mes données sont-elles stockées ?
 - Confidentialité de mes données ?
 - Intégrité ?
 - ROI (Return Of Investment) ?
- Le nuage magique repose sur des ressources matérielles partagées (infrastructure) qui doivent être parfaitement maîtrisées au niveau technique
 - http://www.supermicro.com/products/nfo/Ad_INTEL_MicroCloud.cfm

OpenNebula - <http://opennebula.org>



© OpenNebula Project

OpenStack (US) versus OpenNebula (EU)

- Source = Master Thesis Pashaie Oslo 2013
- <https://www.duo.uio.no/handle/10852/37451>

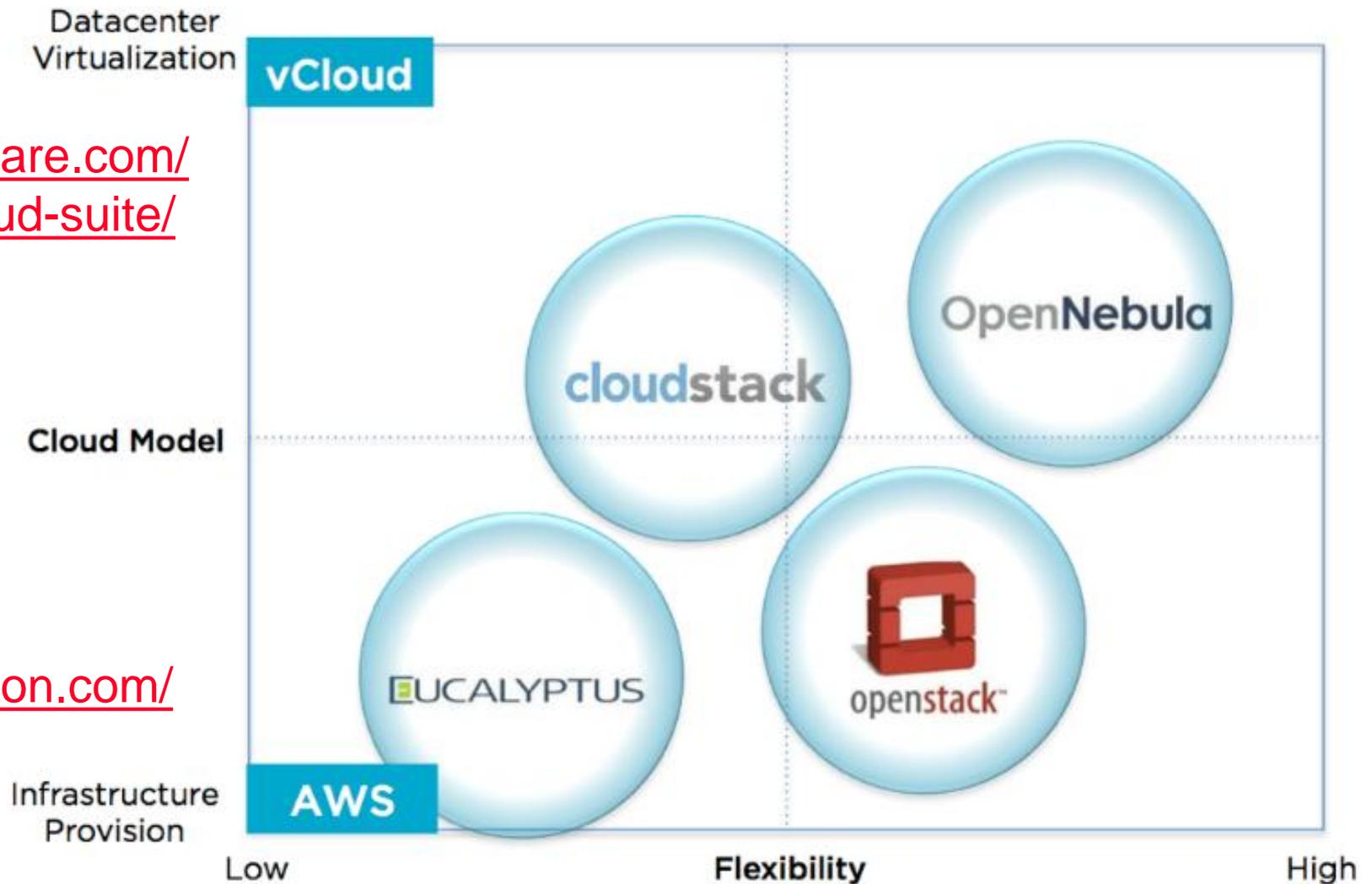
Project name	Cinder	Glance	Horizon	Keystone	Nova	Quantum	Swift	Open-Nebula
Age in days	391	4483	866	771	1095	878	1052	1799
Total amount of-Files	517	337	931	331	2558	783	284	1560
Total amount of Codelines	9959	82554	166424	43014	141838	229173	48920	293621
Total amount of Commits	1087	2597	2519	3261	20709	2658	2639	6683
Average commits per active day	4.0	3.9	4.4	5.6	20.2	4.9	3.8	6.6
Number of Authors	123	144	134	181	454	162	116	37

- *Deploying the OpenNebula was easier than deploying OpenStack cloud platform*
- *OpenStack user & developer communities were the most active*
- *OpenStack was found more complex than OpenNebula*

Positionnement

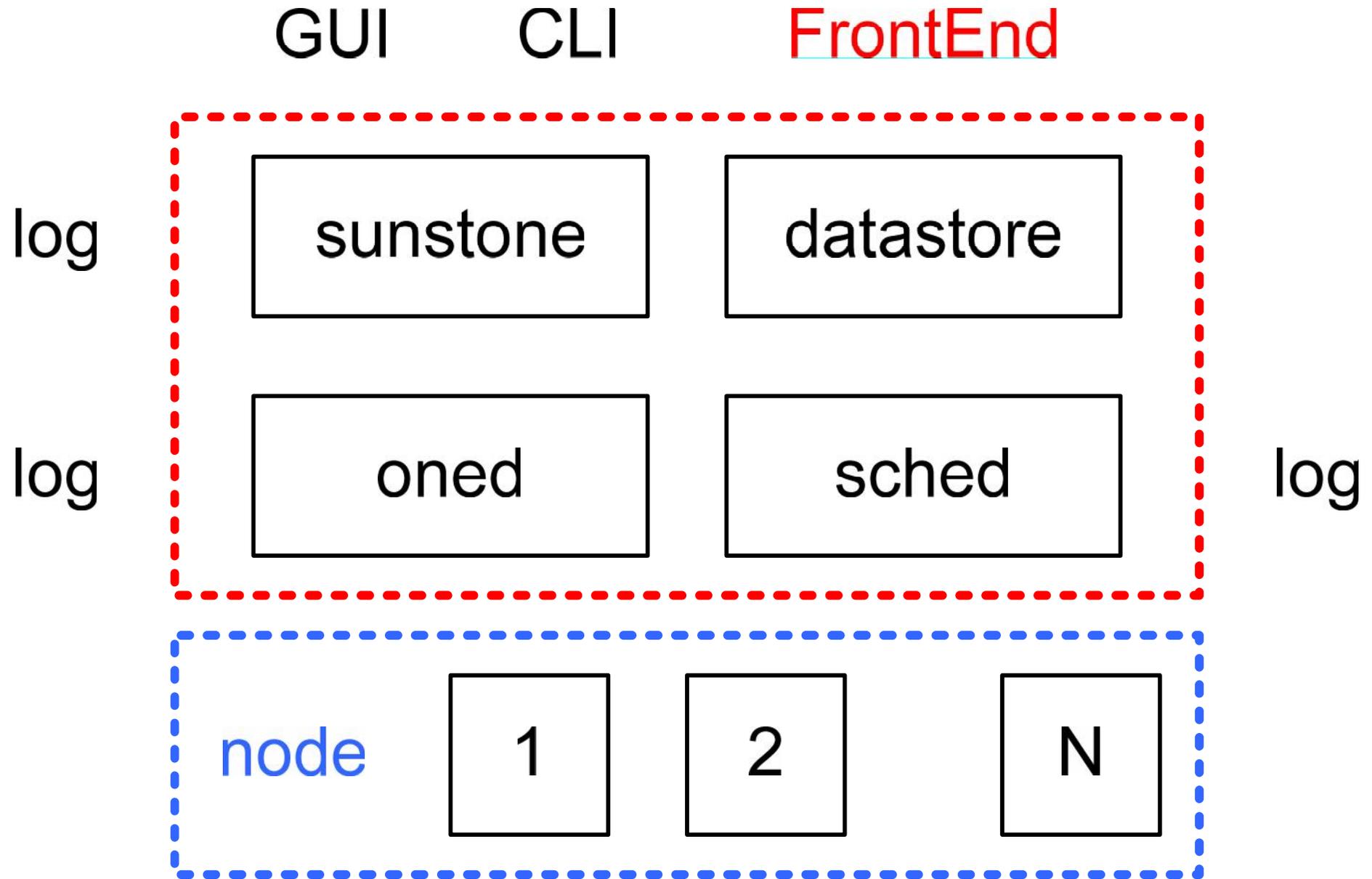
<http://www.vmware.com/fr/products/vcloud-suite/>

<http://aws.amazon.com/>



<http://opennebula.org/eucalyptus-cloudstack-openstack-and-opennebula-a-tale-of-two-cloud-models/>

OpenNebula Process View 1/2



OpenNebula Process View 2/2

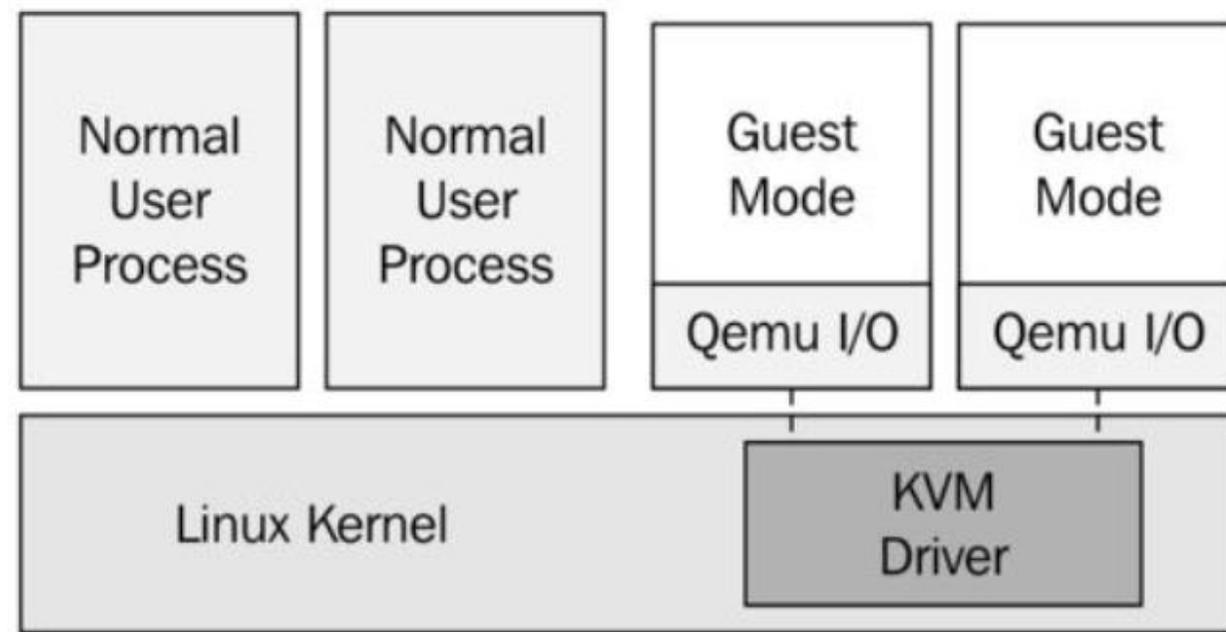
- oned Processus OpenNebula
- sunstone Processus pour administration web
→ <http://FrontEnd:9869>
- datastore Stockage des images, ...
sqlite (par défaut) ou mysql (stockage externe)
- node Hyperviseurs KVM ou ESXi ou ...
- sched Place la VM sur le nœud le plus approprié
→ Lab LiveMigration

Kernel-based Virtual Machine (KVM) history

- ~1950 The supervisor program of CTSS (MIT)
- ~1960 IBM Watson Research : M44/44X Project
- 1972 IBM : VM370 (Mainframe)
- 2001 VMware ESX
- Fév 2007 Linux 2.6.20 : Initial release by Qumranet
- Sept 2009 Red Hat Enterprise Linux 5.4 is released with KVM
- Mai 2011 BMC, HP, IBM, Intel, Red Hat, announce the Open Virtualization Alliance to promote open virtualization technologies including KVM

Hyperviseur KVM

- KVM = Linux kernel module
- Ce module transforme OS en **hyperviseur**
- Il permet à un *User-Space Program* (VM) d'accéder aux ressources virtualisées
- Il exige des CPUs possédant les extensions de virtualisation (Intel VT)
- Il crée un périphérique `/dev/kvm/` qui permet au noyau de fonctionner dans un **3^{ème} mode** → kernel – user – **guest**
- Chaque VM possède son propre espace mémoire via `/dev/kvm`
- QEMU-kvm (Quick Emulator) = User-Space Emulator



Sunstone Web GUI

<http://localhost:9869>

Storage

1 IMAGES 11MB USED

Users

3 USERS 2 GROUPS

Network

2 VNETS 1 USED IPs

Hosts

2 TOTAL 1 ON 0 OFF 1 ERROR

CPU



MEMORY



Allocated

Real

Total

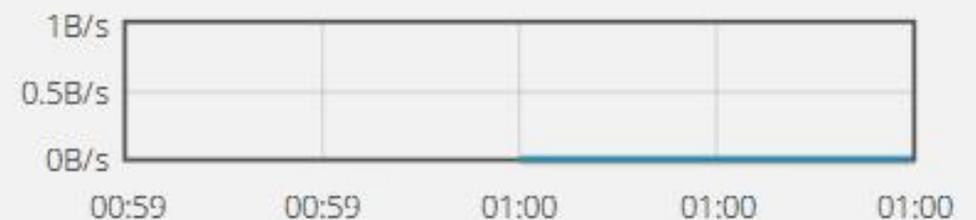
Virtual Machines

1 TOTAL 0 ACTIVE 0 PENDING 1 FAILED

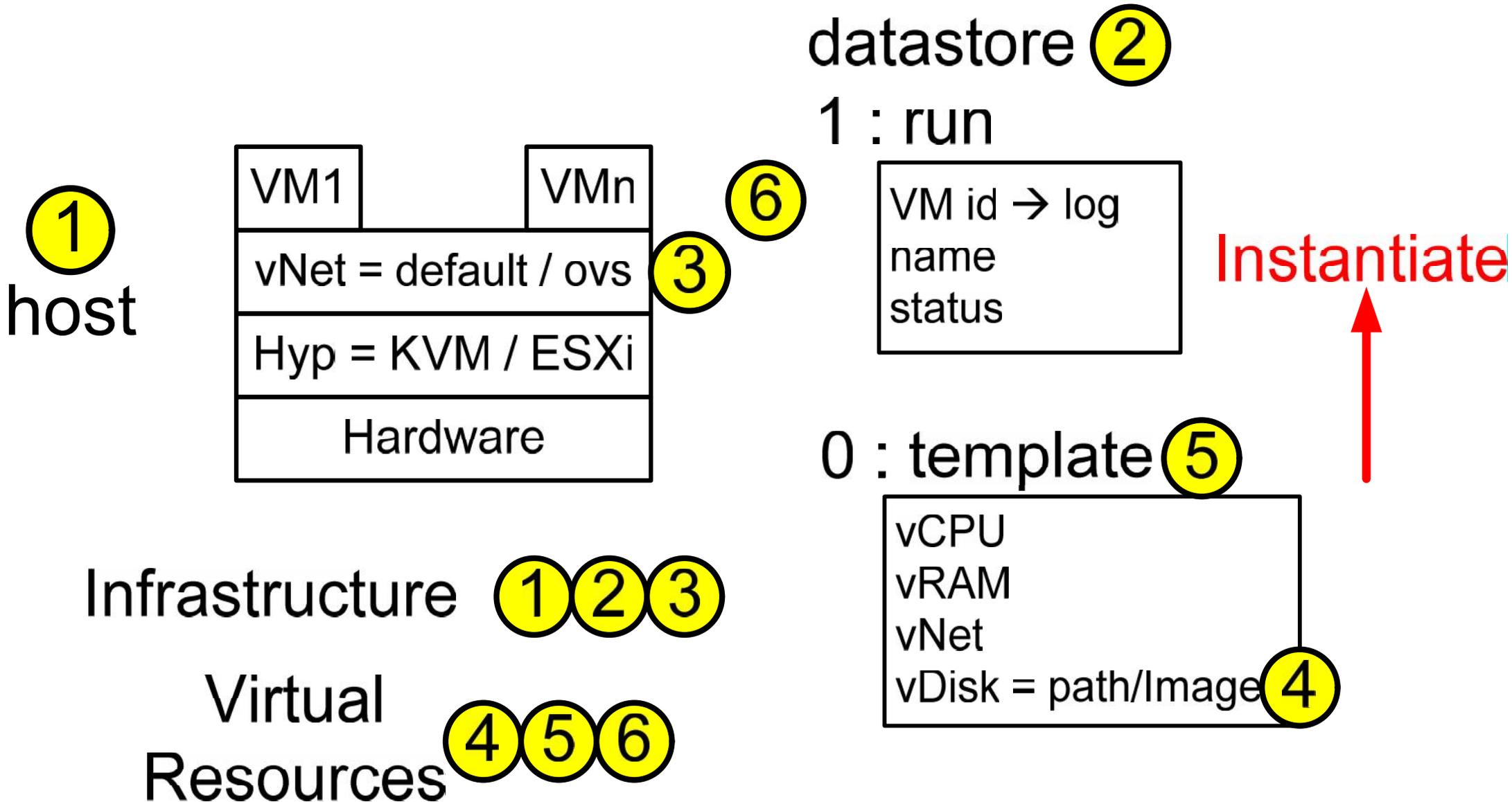
NET DOWNLOAD SPEED



NET UPLOAD SPEED



OpenNebula : 6 étapes de configuration (1/2)



OpenNebula : 6 étapes de configuration (2/2)

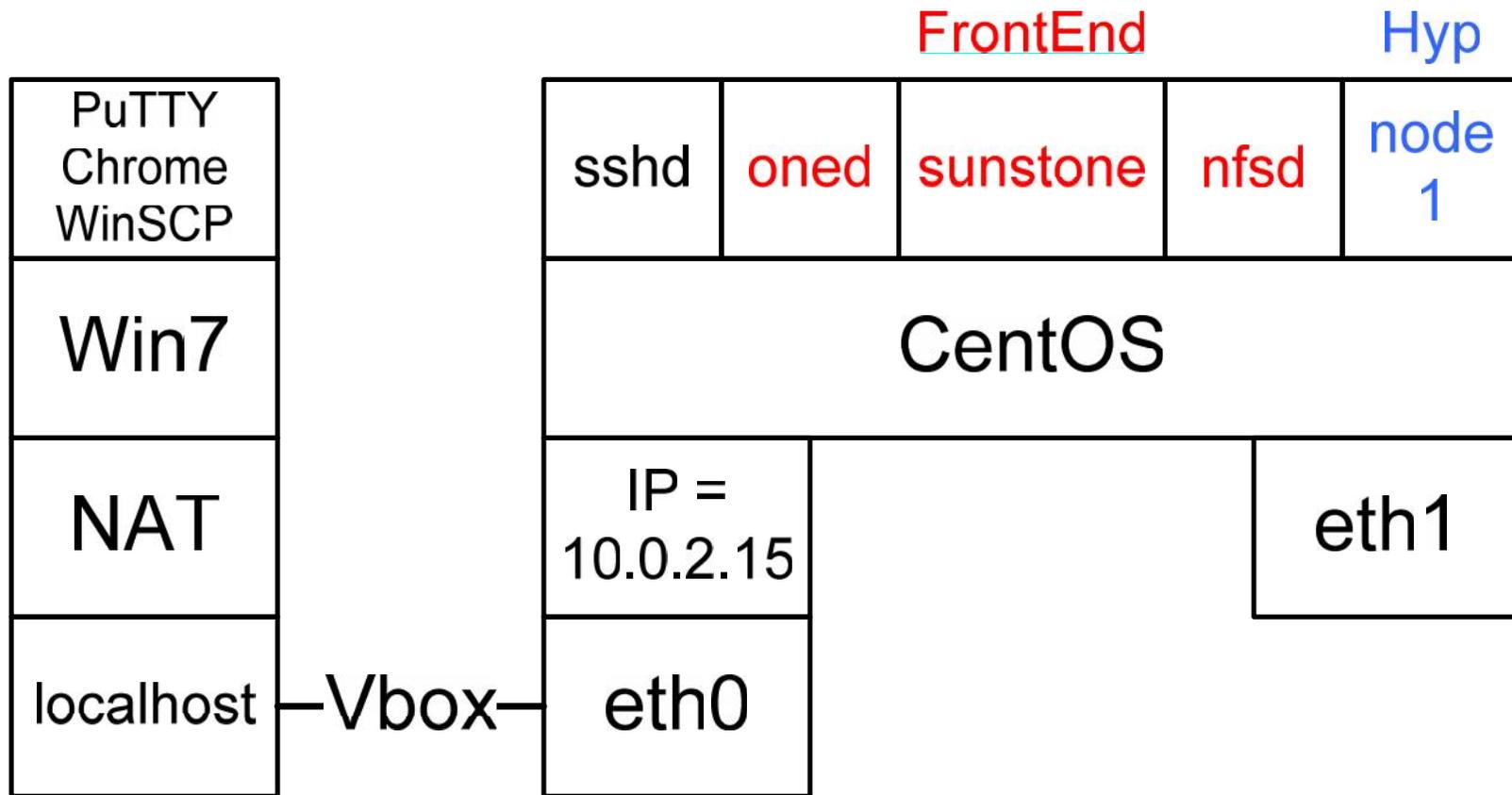
- ① N host = matériel (CPU, RAM, ...) & hyperviseur
- ② datastore = Dépôt (repository) des images
System Datastore (/var/lib/one/datastore/0) → run VM
Image Datastore (/var/lib/one/datastore/1) → create VM
Choix = shared datastore for **live migration**
- ③ vNet = choix de réseaux virtuels (**Open vSwitch**)
- ④ Image = disque virtuel stocké dans Datastore
- ⑤ **Template** = VM definition
- ⑥ **Instantiate** → VM active

Commandes utiles

- `onehost list / show id`
<http://archives.opennebula.org/doc/4.4/cli/onehost.1.html>
- `onedatastore list`
<http://archives.opennebula.org/doc/4.4/cli/onedatastore.1.html>
- `onevnet`
- `oneimage create --name --path --datastore --driver`
<http://archives.opennebula.org/doc/4.4/cli/oneimage.1.html>
- `onetemplate create --name --cpu --memory --disk
--nic --vnc --ssh`
- `onetemplate instantiate I --name`
<http://archives.opennebula.org/doc/4.4/cli/onetemplate.1.html>
- `onevm list / show id`
<http://archives.opennebula.org/doc/4.4/cli/onevm.1.html>

Lab OpenNebula

- Déterminer la configuration de l'appliance Vbox mise à disposition



- Administration GUI avec Sunstone
- Administration en lignes de commande (CLI)
- Appliance Vbox exige un Win7 – 64 bit ... attention à Nb_vCPU

Cloud2

Storage

Rappel : Slide 9

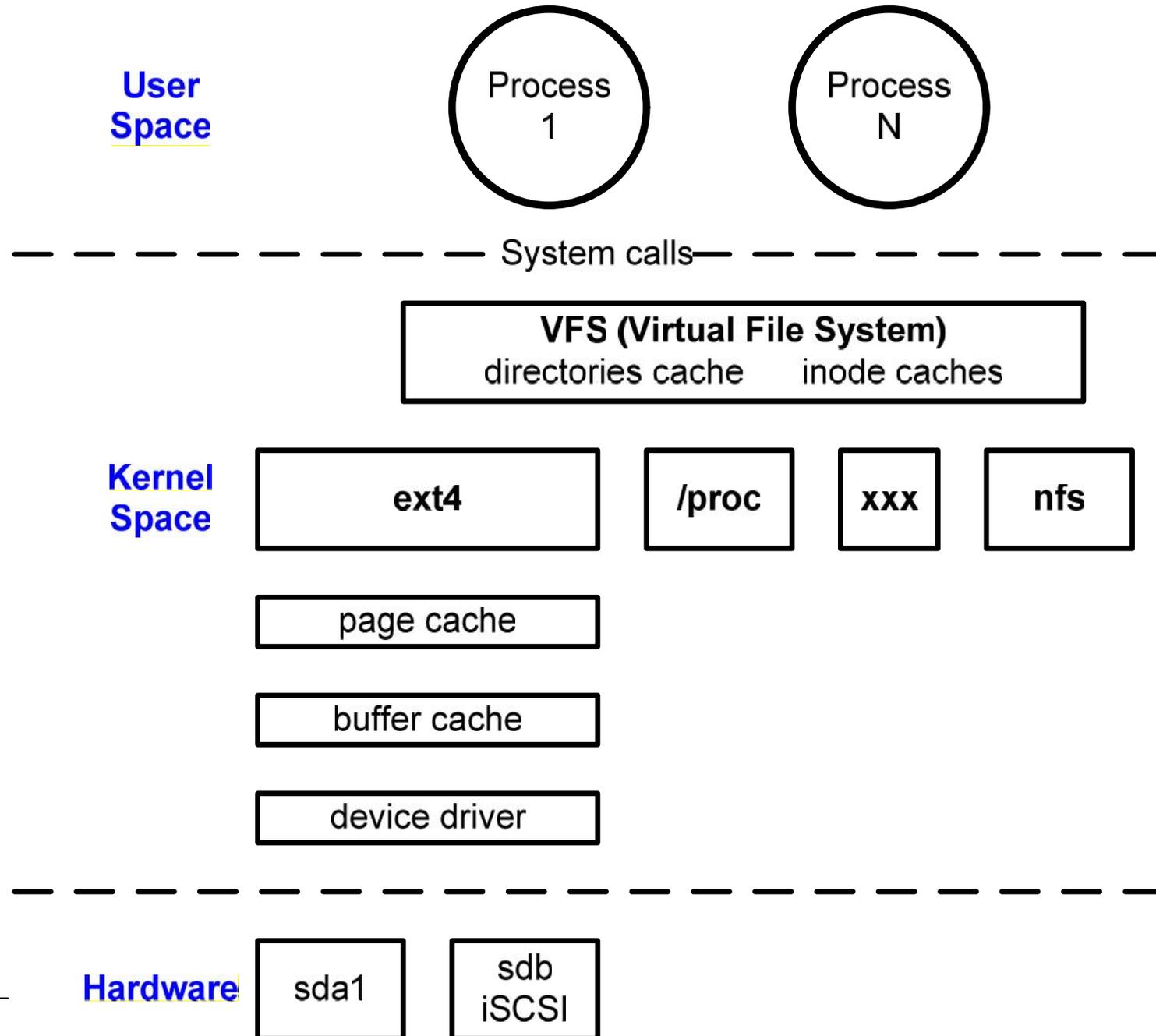
Systemes de stockage

- Essentiels dans le fonctionnement des SI (créer – **stocker** – ...)
- Utilisateur (processus) travaille au niveau **fichiers**
 - **Systemes de fichiers** ntfs (MS), ext4 (Linux), ...
- Les données sont stockées sur des supports physiques (disque magnétique, SSD, ...) → **Block Storage**
- Illustration : Disque de 500 GB (3 partitions) + CD-DVD

Volume	Layout	Type	File System	Status	Capacity
New Volume (E:)	Simple	Basic	NTFS	Healthy (Logical Drive)	153.32 GB
Windows7_OS (C:)	Simple	Basic	NTFS	Healthy (Boot, Page File, Crash Dump,...)	98.08 GB
New Volume (D:)	Simple	Basic	NTFS	Healthy (Logical Drive)	195.31 GB

Disk 0 Basic 465.76 GB Online	SYSTEM_DR 1.46 GB NTFS Healthy (Syst	Windows7_OS (C:) 98.08 GB NTFS Healthy (Boot, Page File	New Volume (D:) 195.31 GB NTFS Healthy (Logical Drive)	New Volume (E:) 153.32 GB NTFS Healthy (Logical Drive)
CD-ROM 0 DVD (F:)				

Linux Virtual File System (VFS)



Linux Virtual File System (VFS)

- L'accès aux systèmes de fichiers locaux (ext4) ou distants (nfs) est unifié par le couche VFS (Virtual File System)
- inode contient les **méta-données** = file type, access rights, owner & group, file length, nb of data blocks, timestamp, ... → ls command
→ équivalent à Windows file properties
- Divers mécanismes de **cache** (page, buffer)
- `/proc` = pseudo-filesystem which is used as an interface to kernel data structures → `procinfo` : statistiques RAM, CPU & IRQ (interruptions)

Disques et partitions Linux

```
lsblk
```

NAME	SIZE	TYPE	MOUNTPOINT
sda	298G	disk	
sda1	1M	part	
sda2	500M	part	/boot
sdb	40G	disk	

- **sda** = disque physique SATA de 300 GB
- **sda1, sda2** = partitions
- **sdb** = disque iSCSI (disque SCSI accessible par le réseau)
- Les données sont stockées par bloc sur le disque
- Liste chaînée de blocs (analogie avec les trames Ethernet)
- 2 systèmes de fichiers ne peuvent pas partager une partition disque

Stockage local (PC – serveur)

- **DAS** (*Direct Attached Storage*)

Disque(s) relié(s) physiquement au PC ou au *mainframe*

- Disques durs (*Hard Disk*)

Capacité (MByte), MTBF (*Mean Time Before Failure*)

Interfaces = ATA/IDE, SCSI (*Small Computer System Interface*),
SATA (*Serial ATA*), SAS (*Serial Attached SCSI*)

Performance = F (*blocksize, read/write, sequential/random, cache*)

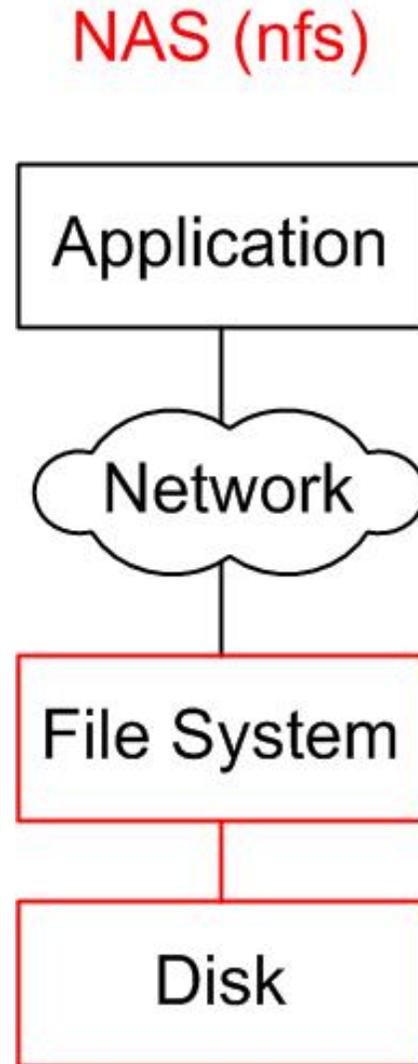
IOPS (*I/O per s*), *throughput* (Mbyte/s), *response time*

http://www.passmark.com/products/pt_advdisk.htm

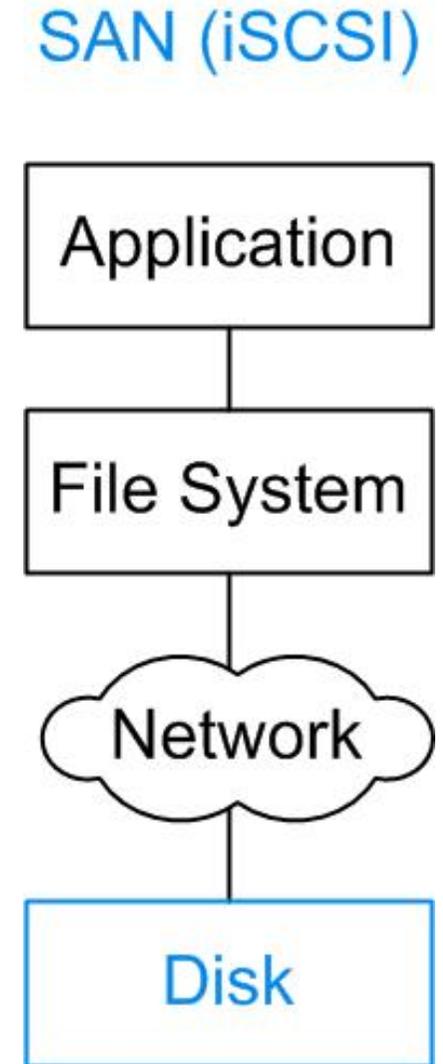
- Technologie = SSD (*Solid State Drive*)

Stockage distant (SAN – NAS)

NAS appears to the client OS as a **file server**



SAN, which provides **only block-based storage** and leaves file system concerns on the client side



SAN vs NAS

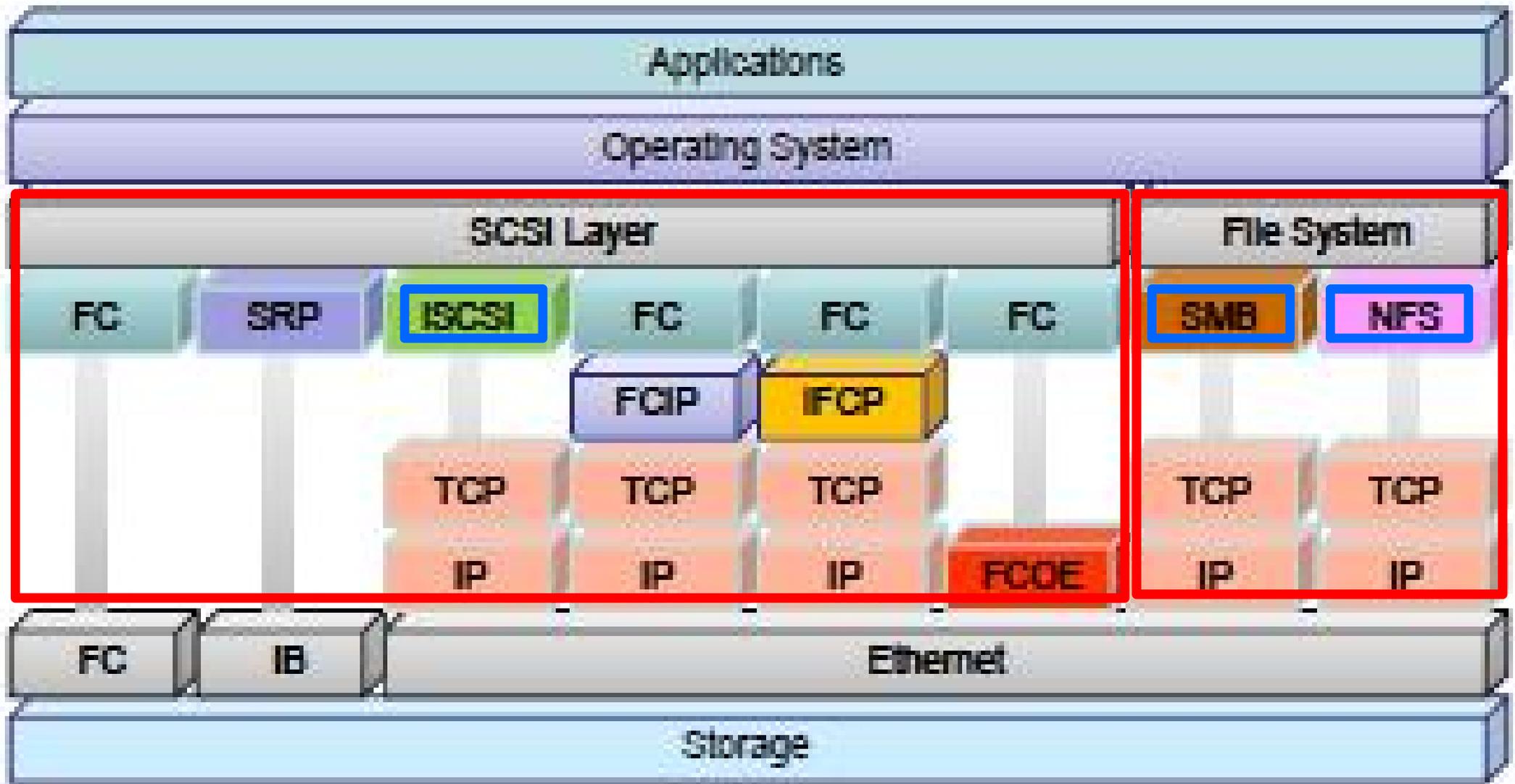
- **NAS** (Network-Attached Storage) provides both storage and a file system. This is often contrasted with **SAN** (Storage Area Network), which provides **only block-based storage and leaves file system concerns on the client side**. SAN protocols are SCSI, Fibre Channel, iSCSI, ...
- One way to loosely conceptualize the difference between a NAS and a SAN is that a **NAS appears to the client OS (operating system) as a file server** (the client can map network drives to shares on that server) whereas **a disk available through a SAN still appears to the client OS as a disk**, visible in disk and volume management utilities (along with client's local disks), and available to be formatted with a file system and mounted.
- Source = http://en.wikipedia.org/wiki/Network-attached_storage

QNAP TS-459 Pro II (SAN – NAS)

- Système de stockage 4 baies
 - Intel Atom 1.8 GHz Dual Core, 3 GB RAM
 - Noyau Linux 2.6, 2 NIC 1 Gbit/s
- Disque WD6000HLHX
 - Capacité = 600 GB
 - Interface SATA 6 Gbit/s
 - Vitesse de rotation = 10'000 rpm
 - Cache de 32 MB
 - Data Transfer Rate = 145 MB/s
 - MTBF = 1.4 mio heures
 - 600'000 cycles arrêt-démarrage
 - Consommation (W) = 4.3 (repos) = 6.4 (RW)
 - Température de fonctionnement = 5-55°C
 - Prix = 210.- en mai 2012



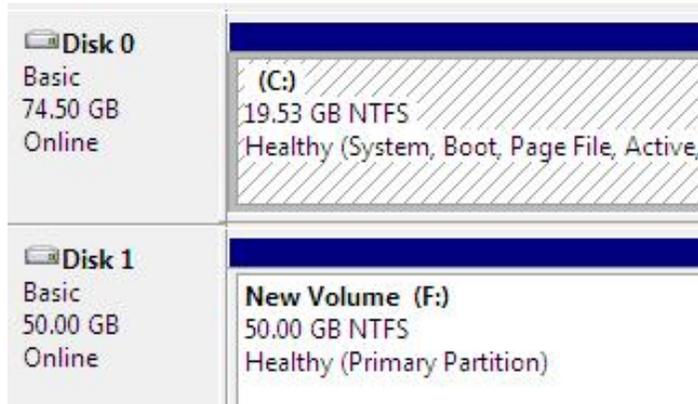
Storage Protocols



Variantes

- **FC** **Fibre Channel** Technologie éprouvée et chère
Gbps 1 2 4 8 16 32 64 ...
Année 1997 2001 2005 2008 2011 ? ? ...
- SRP/Infiniband protocole d'interconnexion de périphériques
- **iSCSI** 1 (10) Gbps Ethernet, solution économique
Charge CPU → TCP/IP Offload engine
- FCIP FC over IP for async. replication
- **FCoE** FC over Ethernet, 10-40-100 Gbps
- http://www.snia.org/images/tutorial_docs/Networking/MKJibbe-Wilson_Fibre_Channel_Technologies-rev.pdf L

Labo iSCSI 2012



Logical Unit Number (LUN)

- Subdiviser un disque iSCSI → N x (Client-Serveur)
- LUN (*Logical Unit Number*) désigne l'unité de stockage (adresse)

Device	Capacity	Available	SAN Identifier	LUN
vmhba32:3:0	48.84 GB	None	iqn.2006-01.com.openfiler:tsn.9ed4ca97e2...	0

- Besoin = LUN de boot, LUN des VMs, LUN de données, ...
- LUN *masking* permet d'autoriser le client à voir et à accéder des LUN spécifiques → segmentation (adressage) du système de stockage
→ <http://www.sansecurity.com/faq/lun-masking.shtml>
- iSCSI Naming & Discovery → <http://tools.ietf.org/html/rfc3721>

SCSI Protocol

- Client (*Initiator*) – Serveur (*Target*)
- LUN (*Logical Unit Number*)
- *Phase sequencing*

Bus Free

Arbitration/80 (7)

Select w/ATN/81 (7,0)

Message Out/80 (Identify)

Command/08 (Read) 00 34 22 01 00

Data In/00 01 02 03 04 05 06 07...

Status/00 (Good)

Message In/00 (Command Complete)

Bus Free

- http://en.wikipedia.org/wiki/SCSI_command

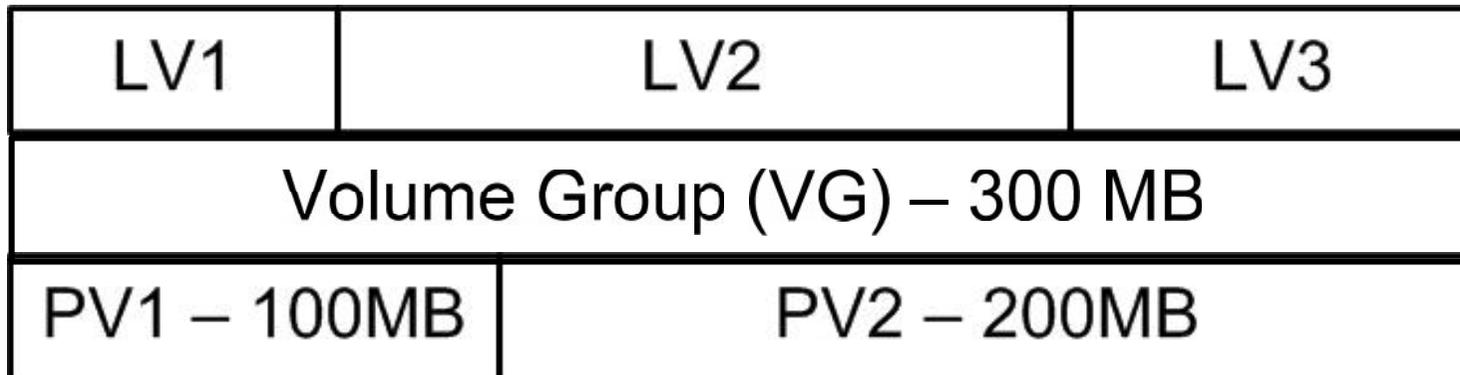
- Principaux paramètres

http://en.wikipedia.org/wiki/SCSI_CDB

- Exemple : **Read (LUN, Log. Block Adr, Len) : Status**

Logical Volume Manager (LVM) → Lab Storage §1

- Couche d'abstraction qui permet de créer des partitions indépendantes des tailles des disques physiques
- Imaginons un PC équipé de 2 disques : 100 MB + 200 MB
→ créer PV1 + PV2 (Physical Volume)
- Nous créons un VG (Volume Group)
- Nous créons 3 LV (Logical Volume) sans contrainte physique



- [http://en.wikipedia.org/wiki/Logical_Volume_Manager_\(Linux\)](http://en.wikipedia.org/wiki/Logical_Volume_Manager_(Linux))

nfs : Intro

- Network File System (NFS) is a **distributed filesystem protocol**
- Version 4 définie dans rfc 3530 (2003)
- Equivalent dans le monde MS = SMB (Server Message Block)
- **Principales différences** avec **file transfer protocols** (ftp, ...)
 - N clients se partagent le serveur
 - Illustration : utilisateur (PC) en entreprise et données sur serveur
 - Opérations = créer, modifier, supprimer, ... un fichier
 - Accès concurrents en lecture et écriture → mécanisme de verrou
 - Délégation par le serveur d'opérations gérées par le client (cache)

nfs : config Linux

- **Server**

```
mkdir /nfs
```

Ajouter dans `/etc/exports`

```
IP_CLIENT/32(rw,sync,no_root_squash,no_all_squash
```

```
systemctl start nfs-server.service
```

http://www.server-world.info/en/note?os=CentOS_6&p=nfs&f=1

- **Client**

```
mkdir /mnt/nfs
```

```
mount -t nfs4 IP_SERVEUR:/nfs /mnt/nfs
```

http://www.server-world.info/en/note?os=CentOS_6&p=nfs&f=2

```
cp f500MB > /mnt/nfs/f1
```

nfs : analyse Wireshark

- ACCESS FH: 0xcccf4a0a, [Check: RD LU MD XT DL]
- ← ACCESS, [Allowed: RD LU MD XT DL]
- LOOKUP DH: 0xcccf4a0a/nfs_file_100b
- ← LOOKUP Status: NFS4ERR_NOENT **file does not exist**
- OPEN DH: 0xcccf4a0a/nfs_file_100b
- ← OPEN **attributes = permission, owner, space, time**
- **WRITE** StateID: 0x16c1 **Offset: 0 Len: 100**
- CLOSE

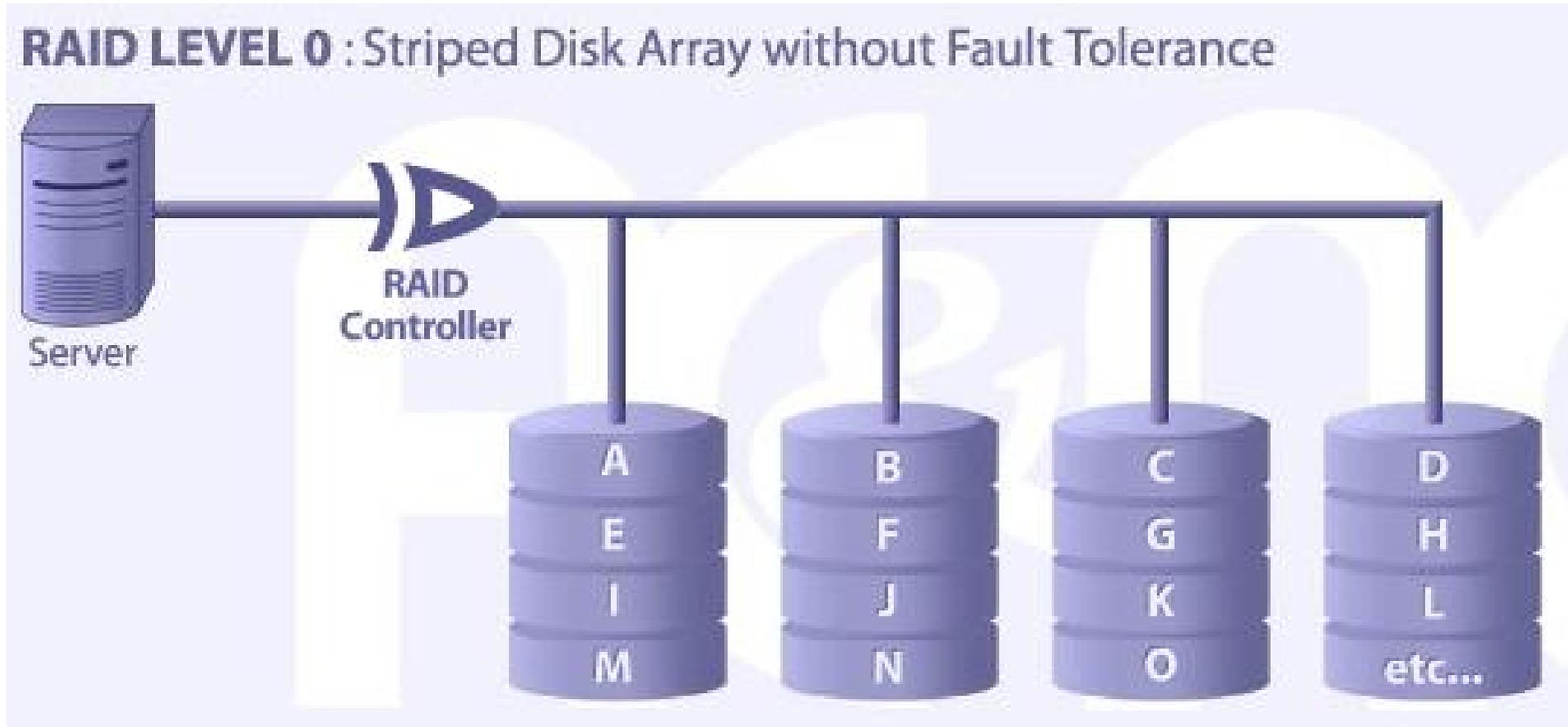
Labo Storage basé sur Fedora16GUI (Client + Serveur)

- §1 Logical Volume Manager (**LVM**) 5 min
Créer un LV de 40 GB qui servira de disque partagé iSCSI
- §2 Configurer le serveur et un client **iSCSI** 10 min
- §3 Mesures de **performance** (copie fichier sur le disque iSCSI) 15 min
Charge réseau avec System Monitor
Commande pv (copy + throughput)
Activité (LED) du disque
- §4 Configurer un partage **NFS** sur le serveur et le tester 15 min
- §5 Comparer les volumes échangés avec iSCSI et nfs 10 min

RAID (Redundant Array of Independent Disk)

- RAID est une technologie d'agrégation de disques visant des **performances élevées** et/ou une **haute disponibilité des données**
- ***Disk mirroring*** → plusieurs disques contiennent les mêmes données
- ***Block striping*** → les données sont réparties sur plusieurs disques
- ***Error correction*** → la redondance (parité, ...) permet de détecter l'erreur et parfois de corriger
- L'implémentation peut être logicielle intégrée au système d'exploitation : FreeBSD, Linux, Windows, ... → *software RAID*
- Elle peut aussi être intégrée dans une carte contrôleur RAID connectée au bus PCI → *hardware RAID*
- <http://en.wikipedia.org/wiki/RAID>

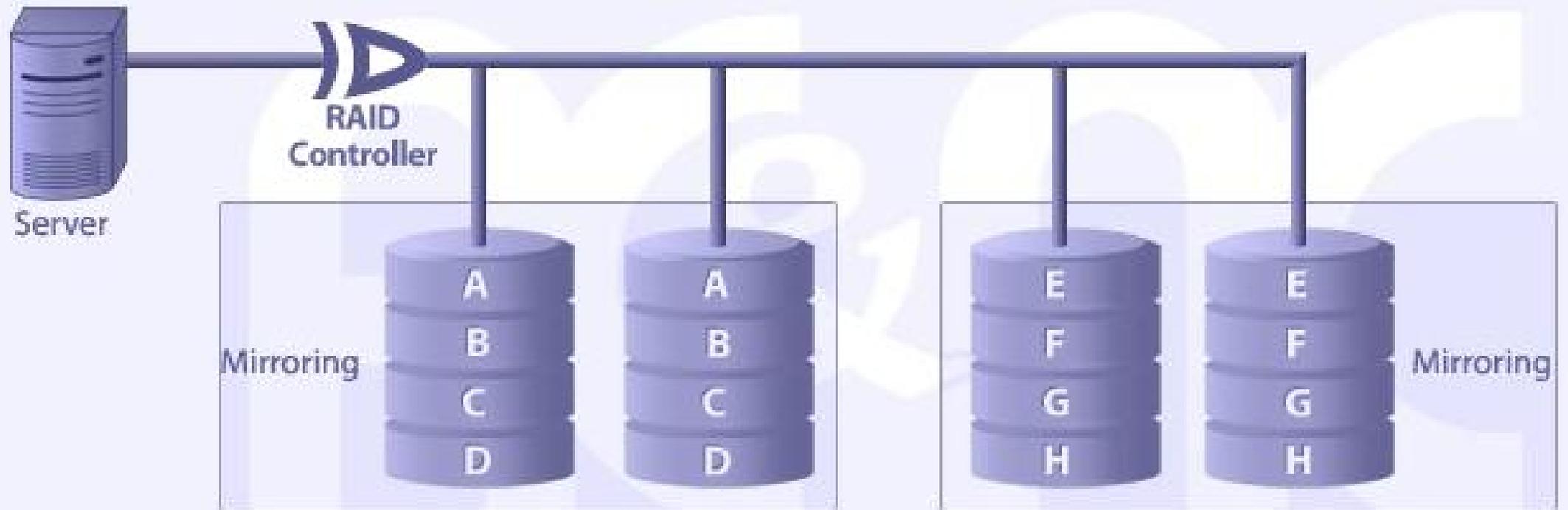
RAID 0



- Performance sans redondance
- Voir animation → http://www.acnc.com/04_01_00.html

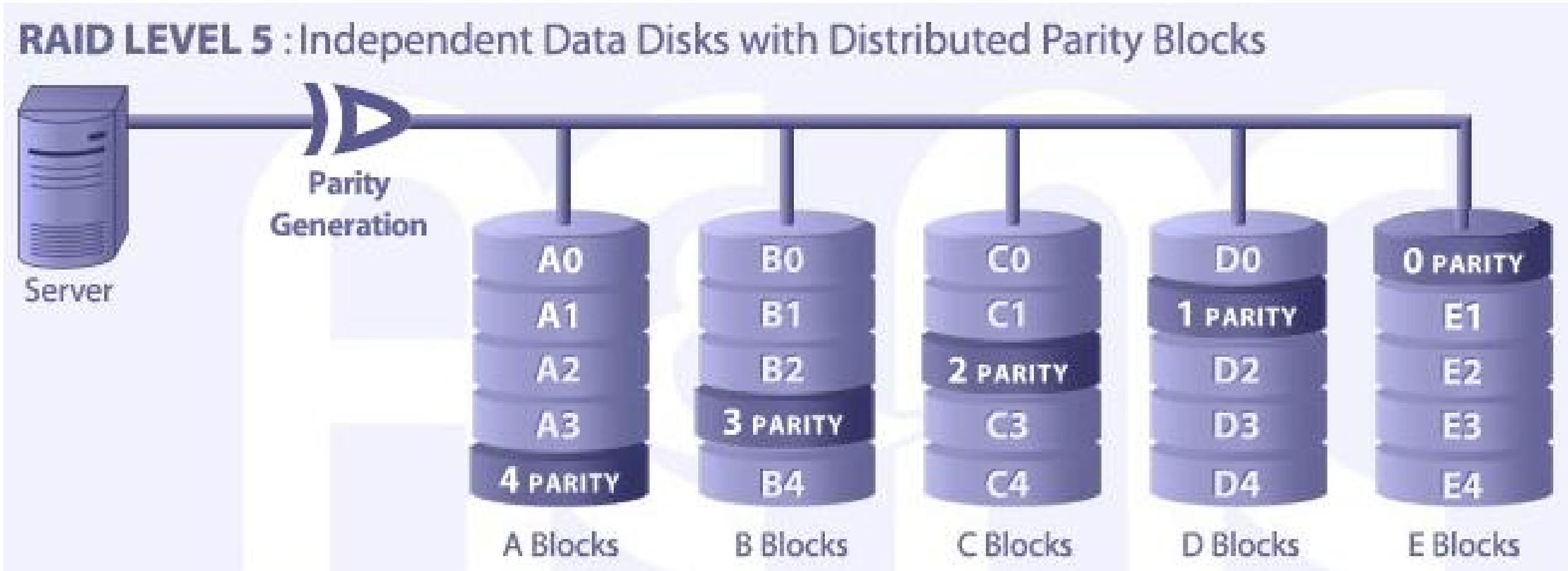
RAID 1

RAID LEVEL 1 : Mirroring & Duplexing



- Supporte la défaillance d'un disque
- Animation → http://www.acnc.com/04_01_01.html

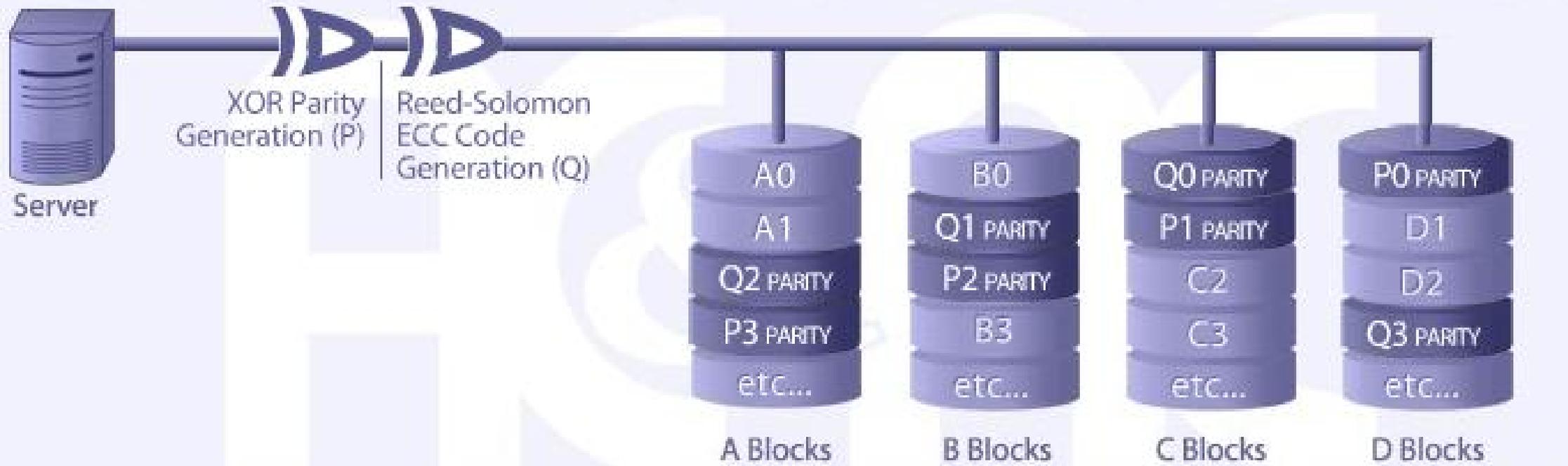
RAID 5



- Supporte la défaillance d'un seul disque qui doit être changé à chaud
- Performance dégradée en cas de reconstruction
- Animation → http://www.acnc.com/04_01_05.html

RAID 6

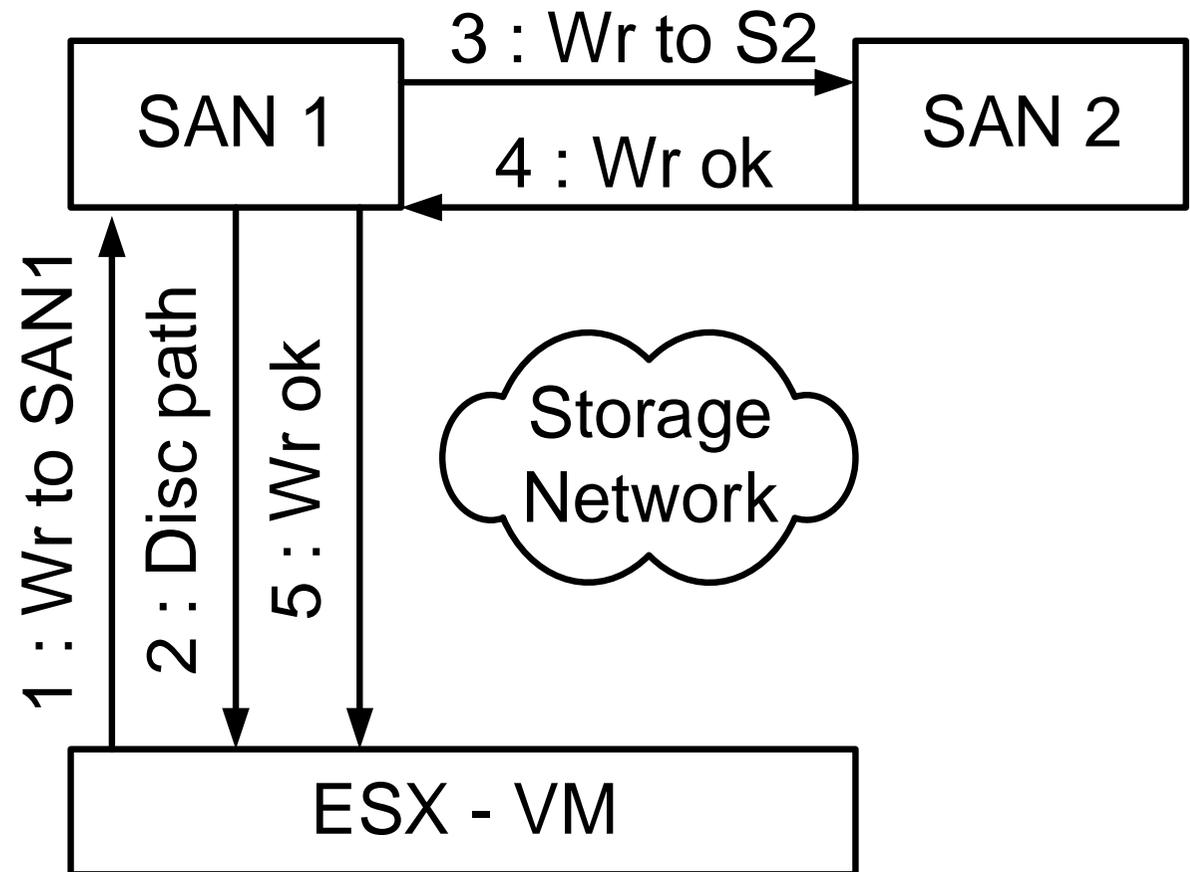
RAID LEVEL 6 : Independent Data Disks with Two Independent Distributed Parity Schemes



- Supporte la défaillance de 2 disques
- Plus on a de disque, meilleurs en sont les performances
- Animation → http://www.acnc.com/04_01_06.html

Replication

- La disponibilité du SAN étant vitale, on trouve généralement une architecture redondante (2 dans l'exemple)
- Un mécanisme de réplication est nécessaire entre les 2 SANs lors d'une écriture
- La réplication est dite **synchrone** avec accusé de réception qui garantit la transaction 1-2-3-4-5
- L'indisponibilité de S2 peut être critique !
- Les performances réseau sont critiques



Replication (suite)

- Comment fonctionne ce système ? *Hot swapping / Load balancing*

HDS : The main advantage of synchronous replication is that data can be recovered quickly. Operations at the remote, mirrored site can begin immediately at the point in time where the primary site stopped should operations at the primary site be disrupted. Only the few operations in process at the instant of disruption may be lost.

- La réplication **asynchrone** est utilisée lorsque les performances réseau ne permettent pas d'utiliser le mécanisme synchrone
- La copie de S1 vers S2 devient **transparente** et **non garantie** mais elle ne freine plus le producteur
- La sauvegarde traditionnelle sur bande est conseillée

Cloud3

OpenNebula Advanced

Problématique des systèmes de fichiers distribués

- Flexibilité (elasticity)

Pouvoir facilement adapter la capacité de stockage aux besoins
Ajouter (remplacer - supprimer) des disques

- Performance

Garantir un débit utile (lecture - écriture)

Solution de type RAID 0 (block striping) → [slide 44](#)

Le(s) serveur(s) de méta-données (table : filename - blocs)
deviennent des goulets d'étranglement avec l'augmentation du
nombre de fichiers

- Haute disponibilité (pas traité)

GlusterFS : linéarité (exemple = 100 Mbyte/s par serveur)

• Server	Disk=1TB	Eth (Gbit/s)	Capacity TB	Performance MB/s
2	12	1	24	200 100/server
2	16	1	32	200 idem
4	6	1	24	400 idem
8	12	1	96	800 idem
8	12	10	96	5000

- **Astuce = Client** GlusterFS calcule la position du fichier à l'aide d'une fonction de hachage → hash (path/filename)
- Si le fichier est renommé → pointeur est utilisé pour éviter de déplacer le fichier
- **Les performances sont proportionnelles au nombre de serveurs**
- http://moo.nac.uci.edu/~hjm/fs/An_Introduction_To_Gluster_ArchitectureV7_110708.pdf

GlusterFS : architecture

- Solution logicielle (rachetée 136 Mio \$ en 2011 par RedHat)
- Utilise un(des) FS natif(s) Linux = Ext4, ...
- Volume est composé de briques (exemple = 1 brique/serveur)
- Elles sont montées (mount) du côté client
- Traducteur (côté client) calcule la position du fichier
- **Traducteur de distribution capable de répartir un fichier sur N briques**
- Traducteur de réplication capable de gérer la redondance
- Augmentation à chaud de la capacité de stockage
- http://www.tdeig.ch/linux/Basbous_RPS.pdf p16
- <http://www.redhat.com/en/files/resources/en-rhst-introducion-red-hat-storage-architecture-8457567.pdf>

GlusterFS : configuration

- Distributed Setting

http://www.server-world.info/en/note?os=CentOS_6&p=glusterfs&f=1

- Replication Setting

http://www.server-world.info/en/note?os=CentOS_6&p=glusterfs&f=2

- Striping Setting

http://www.server-world.info/en/note?os=CentOS_6&p=glusterfs&f=3

- Distributed & Replication Setting

http://www.server-world.info/en/note?os=CentOS_6&p=glusterfs&f=4

- Stripping & Replication Setting

http://www.server-world.info/en/note?os=CentOS_6&p=glusterfs&f=5

OpenNebula (suite slides 19-23) : Template

TEMPLATE N INFORMATION

```
ID          : N
NAME        : T1
USER        : toto
GROUP       : oneadmin
REGISTER TIME :
PERMISSIONS
OWNER       : um-
GROUP       : ---
OTHER       : ---
```

TEMPLATE CONTENTS

```
CONTEXT=[
  NETWORK="YES",
  SSH_PUBLIC_KEY=
  CPU="1"
  DISK=[
    IMAGE_ID="2" ]
  GRAPHICS=[
    LISTEN="0.0.0.0",
    TYPE="VNC" ]
  MEMORY="64"
  NIC=[
    NETWORK_ID="1" ]
```

/var/log/one/id.log (WinSCP) – Lab §3.6

Date Time

[DiM][I]: New VM state is **ACTIVE**

[LCM][I]: New VM state is **PROLOG**

[LCM][I]: New VM state is **BOOT**

[VMM][I]: Generating deployment file:
/var/lib/one/vms/4/deployment.0

[VMM][I]: ExitCode: 0

[VMM][I]: Successfully execute network driver operation:pre

[VMM][I]: ExitCode: 0

[VMM][I]: Successfully execute virtualization driver
operation: deploy.

[VMM][I]: ExitCode: 0

[VMM][I]: Successfully execute network driver operation:post

[LCM][I]: New VM state is **RUNNING**

Etats d'une VM

- VM instantiate... suite des slides 20-21
- Etat = **Pending**
Scheduler choisit le host selon l'algorithme défini
- Etat = **Prolog**
Le disque virtuel est accessible avec nfs
→ Live Migration
- Etat = **Boot**
Contextualization
- Etat = **Running**
VM monitoring : **push** (node →) / **pull** (node ←) model

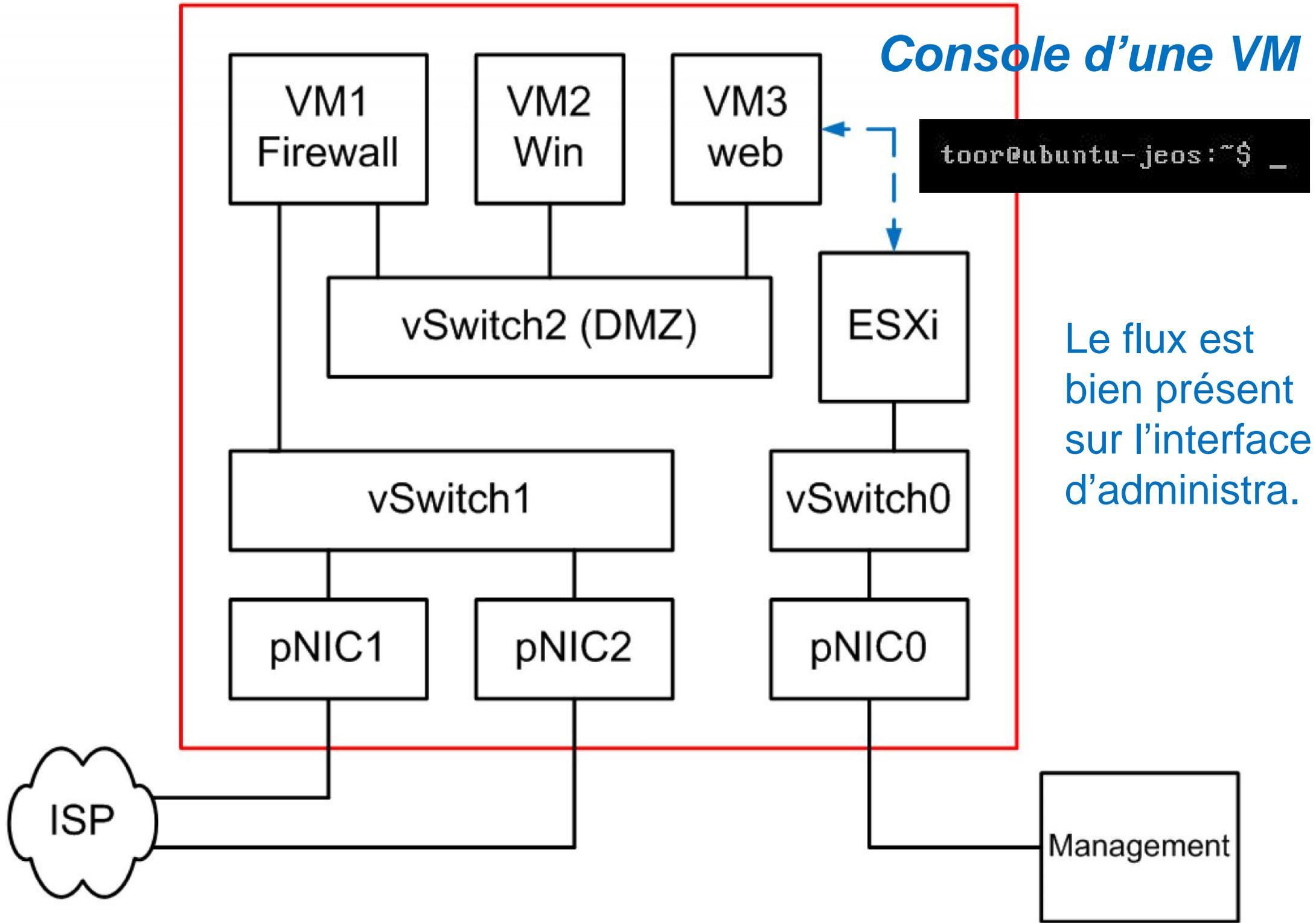
Scheduler

- Périodiquement (30 s), le processus sched prend en compte les VMs dans l'état Pending
- **Packing** → minimise le nb de hosts
cluster = N hosts
- **Striping** → utilise toutes les ressources du cluster
- **Load-aware** → choisit le host le moins chargé
- **Fixed** → choix du host selon une liste prédéfinie
- ... CPU over commitment, RAM over commitment, ...

Contextualization

- Donner aux VMs des configurations (réseau, ...) spécifiques
- Script exécuté au démarrage → [/etc/rc.d/init.d/vmcontexttty](#)
- OpenNebula génère une adresse Ethernet dont les 4 membres de poids faible sont identiques à l'adresse IP
HWaddr 02:00:C0:A8:38:0A
- Adresse IP est définie dans Virtual Network

Leases information	
IP Start	192.168.56.10 = C0:A8:38:0A
IP End	192.168.56.30



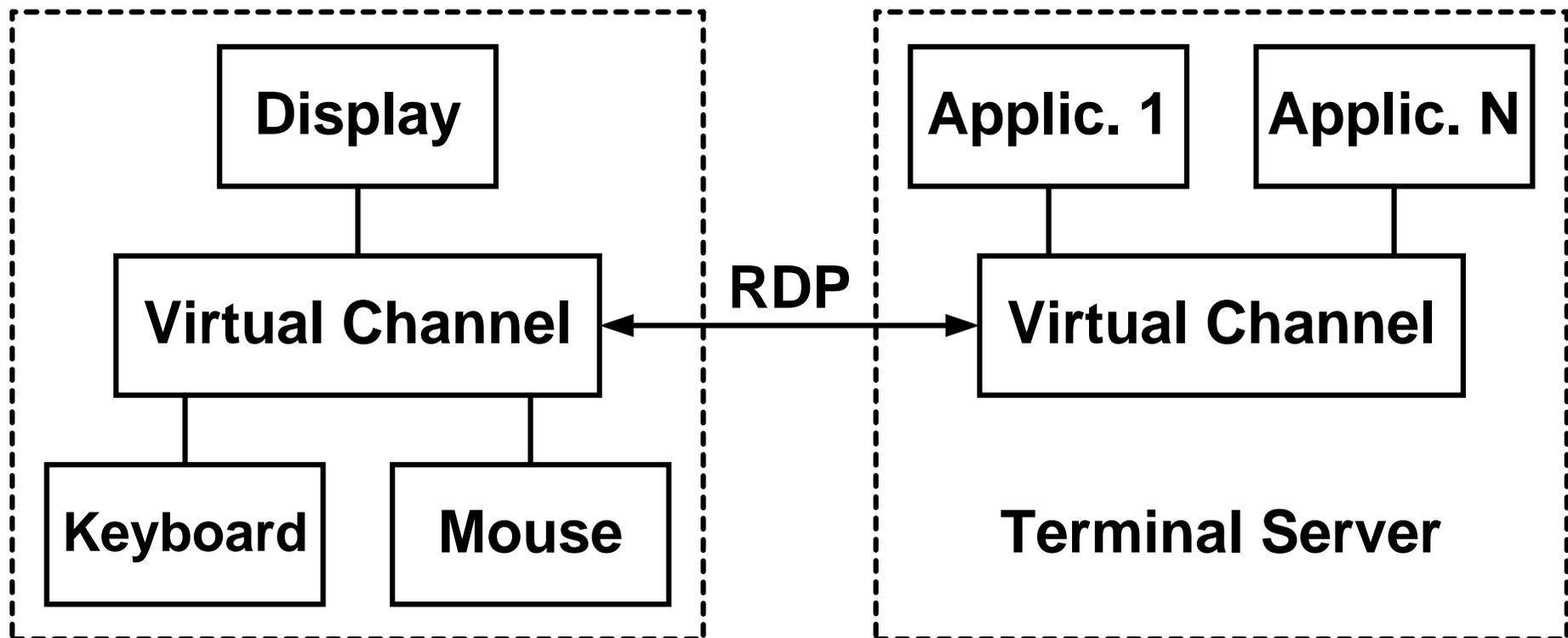
Console d'une VM

```
toor@ubuntu-jeos:~$ _
```

Le flux est bien présent sur l'interface d'administra.

Virtual Network Computing (VNC)

- VNC is a **graphical desktop sharing** system that uses the remote framebuffer protocol to **remotely control another computer**. It transmits the **keyboard and mouse events** from one computer to another, relaying the **graphical screen updates** back over a network.
- Architecture Remote Desktop Protocol de Microsoft



Page d'accueil du panneau de configuration

-  Gestionnaire de périphériques
-  Paramètres d'utilisation à distance
-  Protection du système
-  Paramètres système avancés

- Configuration Microsoft

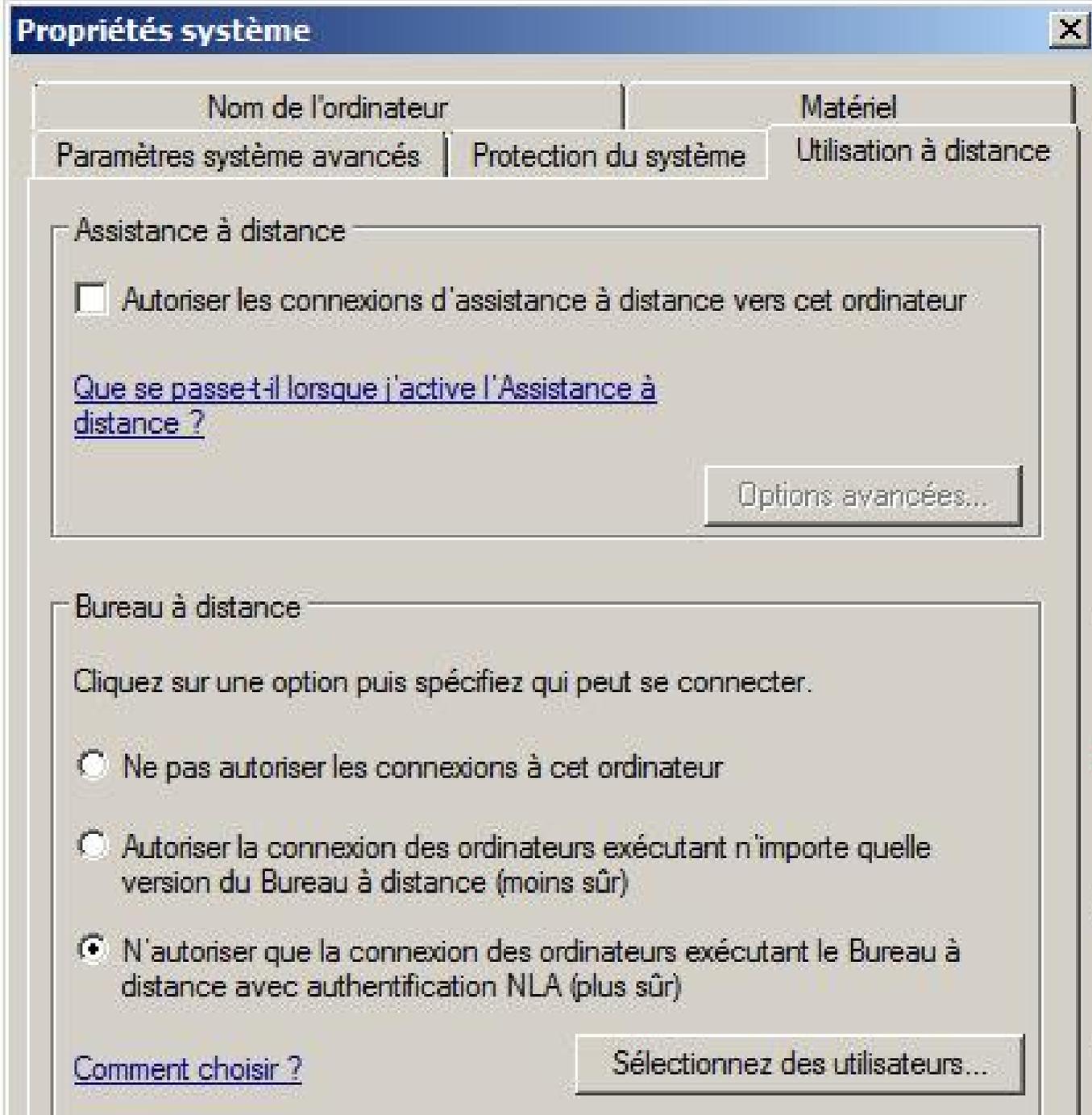
- Client

`\\system32\\mstsc.exe`

Voir aussi

[Centre de maintenance](#)

[Windows Update](#)



Propriétés système

Nom de l'ordinateur | Matériel

Paramètres système avancés | Protection du système | Utilisation à distance

Assistance à distance

Autoriser les connexions d'assistance à distance vers cet ordinateur

[Que se passe-t-il lorsque j'active l'Assistance à distance ?](#)

[Options avancées...](#)

Bureau à distance

Cliquez sur une option puis spécifiez qui peut se connecter.

Ne pas autoriser les connexions à cet ordinateur

Autoriser la connexion des ordinateurs exécutant n'importe quelle version du Bureau à distance (moins sûr)

N'autoriser que la connexion des ordinateurs exécutant le Bureau à distance avec authentification NLA (plus sûr)

[Comment choisir ?](#)

[Sélectionnez des utilisateurs...](#)

Exercices VNC

- 1 Que faut-il au minimum comme matériel du côté client ?
- 2 Où les applications (word, ...) doivent être installées ?
- 3 Quels types de données sont échangées entre client et serveur ?

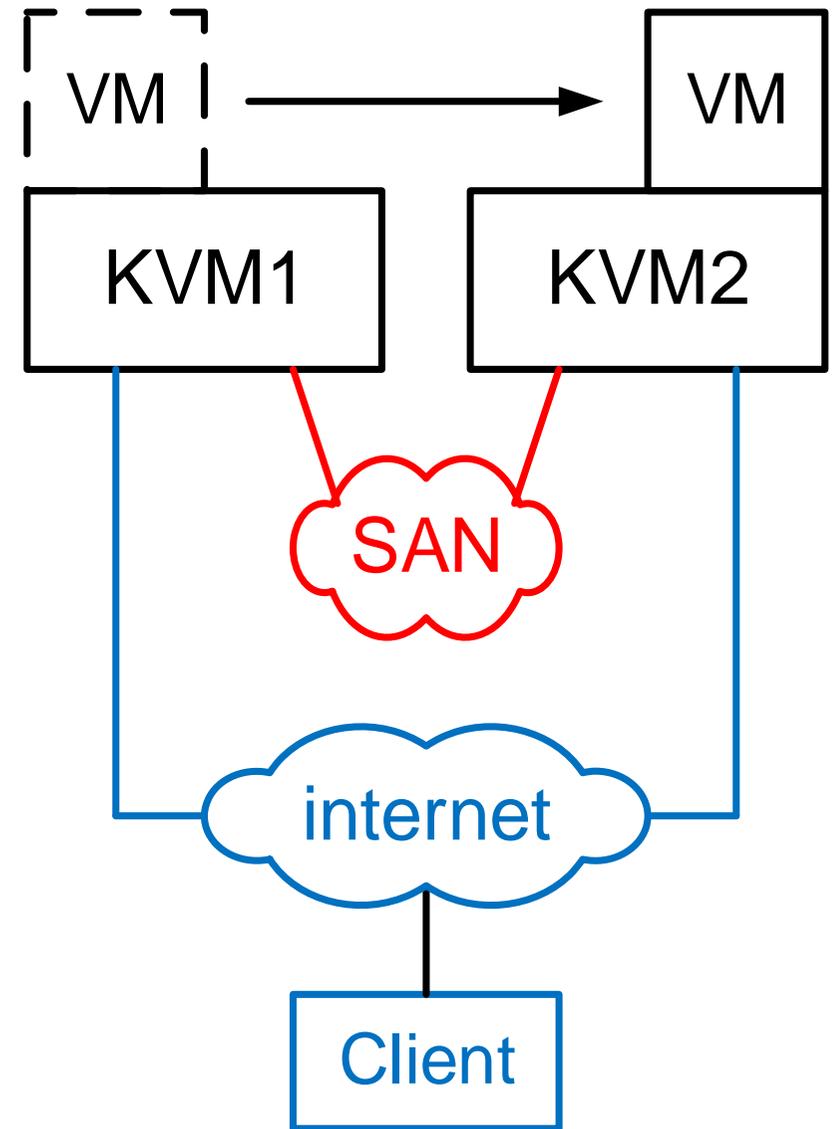
Exercices VNC (correction)

- 1 Que faut-il au minimum comme matériel du côté client ?
CPU + RAM + clavier + souris + contrôleur graphique + NIC
suite dans VDI → www.tdeig.ch/linux/VDI_miniLinux.pdf
- 2 Où les applications (word, ...) doivent être installées ?
Du côté serveur
- 3 Quels types de données sont échangées entre client et serveur ?
Caractère du clavier, position de la souris, bitmap de l'écran
→ flux très dissymétriques

Voir lab VDI – Spice optionnel + config Linux sur tdeig

Live Migration : utilité & exigences

- Déplacer une VM de KVM1 vers KVM2 sans perte de disponibilité
- Répartition de charge (déplacement auto)
- Maintenance (déplacement manuel)
- Exige une **capacité de stockage partagée** (nfs, iSCSI, ...)
- **Exige un réseau de stockage séparé**
- Exige une compatibilité des CPUs



Live Migration : fonctionnement

- Mécanisme transparent pour la VM

- Algorithme

KVM1 initie le transfert afin de détecter les changements en RAM

KVM1 copie toutes les pages de la mémoire RAM

KVM1 arrête la VM

KVM1 transfère VM states + les changements en RAM

KVM2 démarre la VM et envoie un paquet RARP

La MAC adresse est conservée

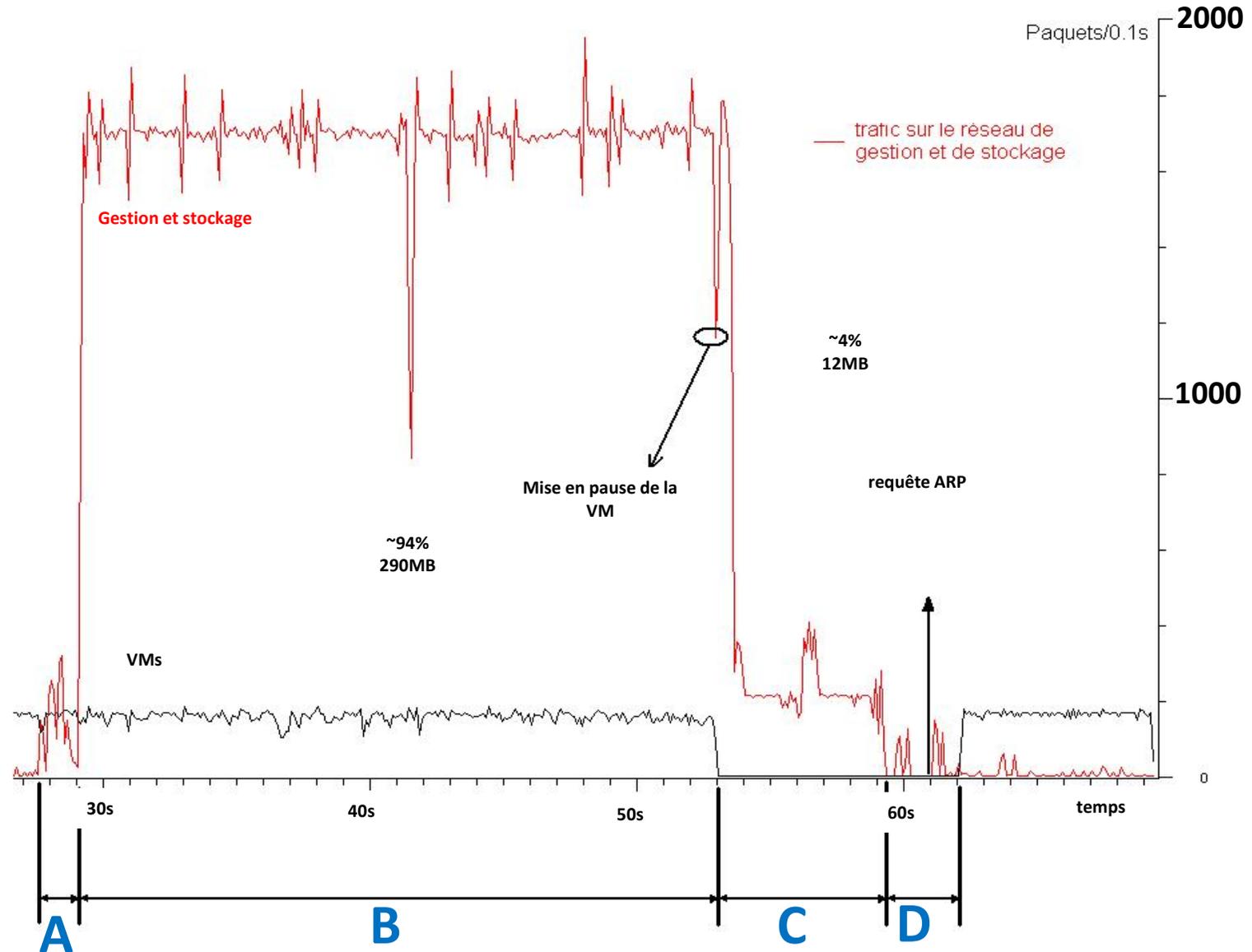
- Limitations

Perte du paquet ARP ?

A l'intérieur d'un réseau IP à haut débit

vMotion : mesures avec réseau lent (100 Mbit/s)

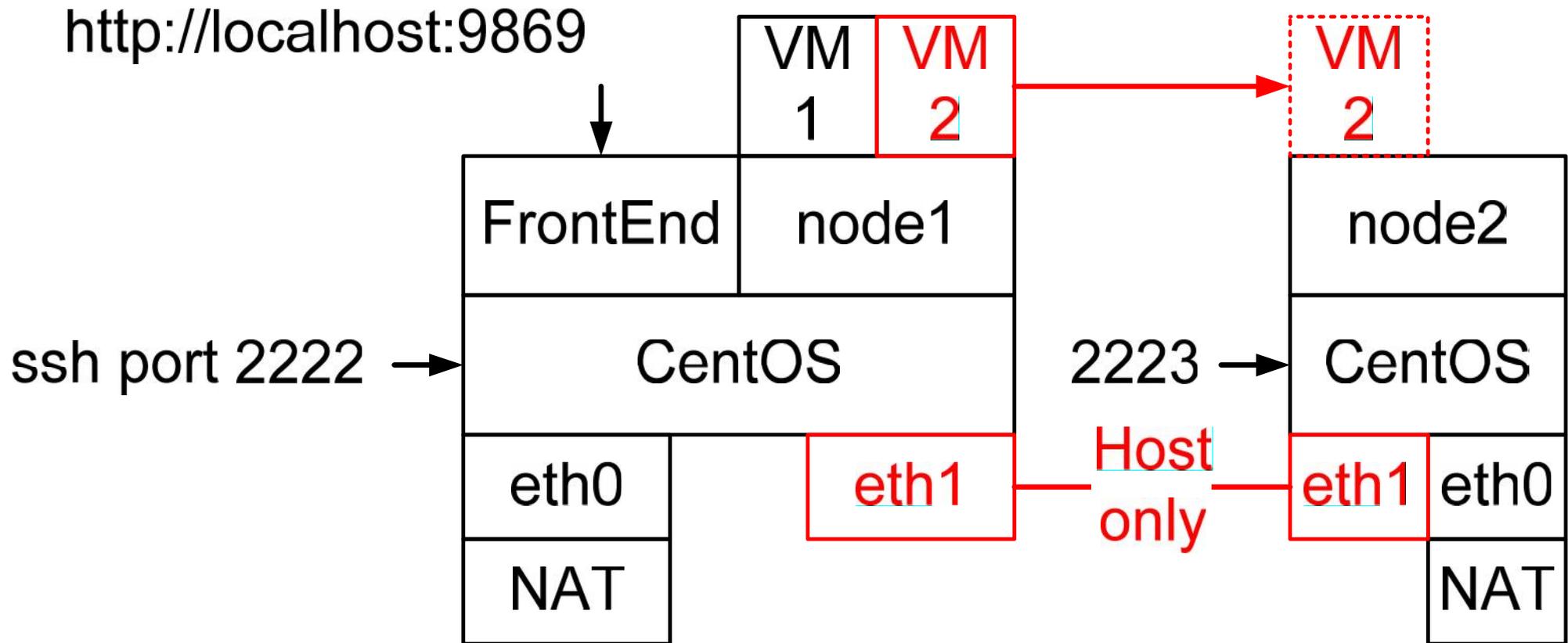
- Temps total = 70s
(A+B+C+D)
- Temps de transfert = 32s
(A+B+C+D)
- Temps d'indisp. = 8.8s (C+D)



- http://www.tdeig.ch/vmware/sandmeier_M.pdf

Lab LiveMigration

- A dispo : GL1.ova + GL2.ova (appliances Vbox)



- Configurer vCPU, scheduler Ajouter node2, tests DNS et nfs
- Live Migration → virtual network, logs, scheduling policies
- Contextualization

Cloud4

Network Virtualization

Virtualisation Réseau

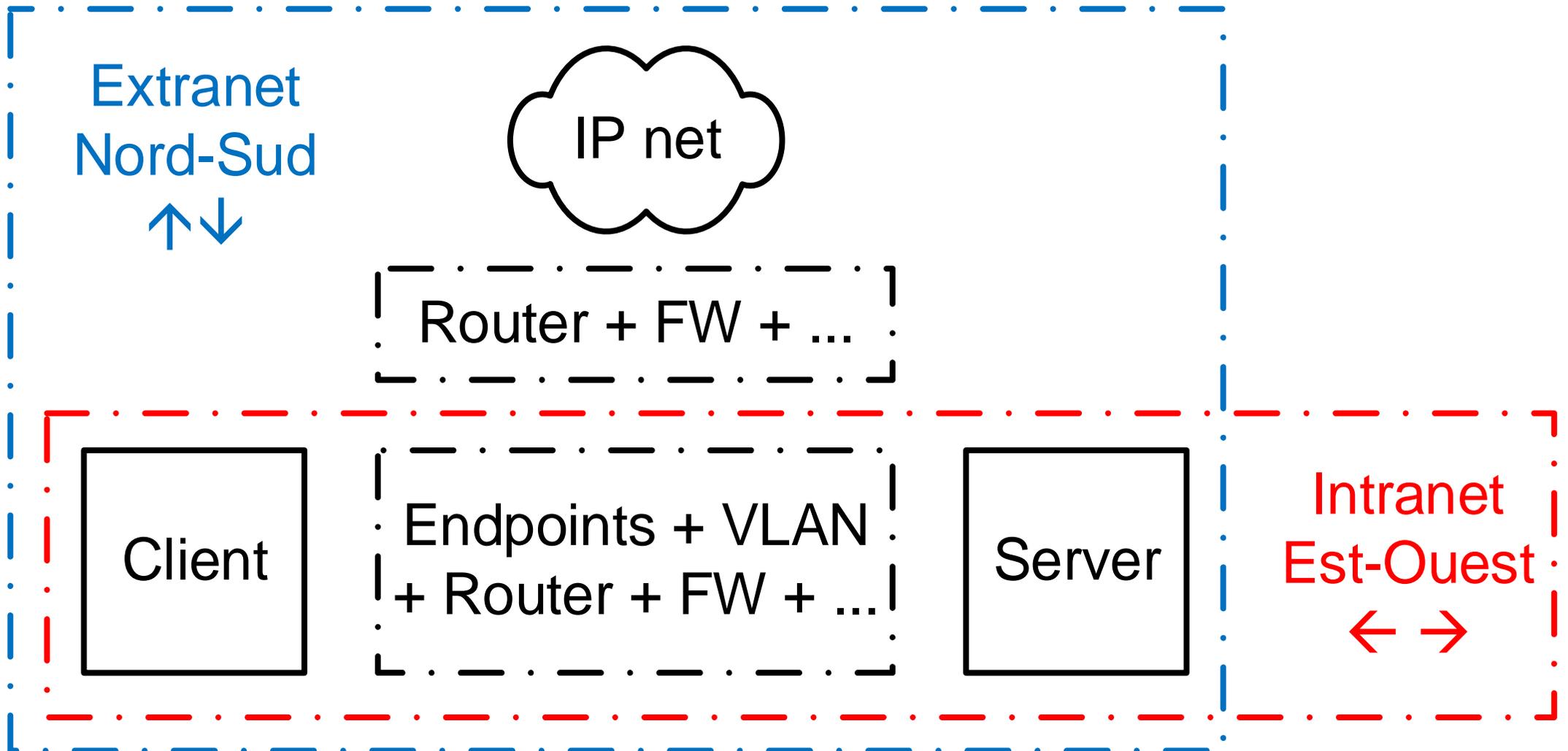
- Trafic Nord-Sud & Est-Ouest
- Spine-Leaf Architecture
- Administration à distance via module KVM (Keyboard-Video-Mouse)
Cisco Integrated Management Controller
- Architecture physique d'un hébergeur
VLAN, tunnel IPSec,
Distributed Switch, Logical Switch, VXLAN
- Scénario Multi-tenants
Edge Service Gateway
- Software Defined Networking
- Labo OpenvSwitch
- **Cours basé sur les résultats des travaux de diplôme 2016**

http://www.tdeig.ch/virtual/Udriot_RTb.pdf

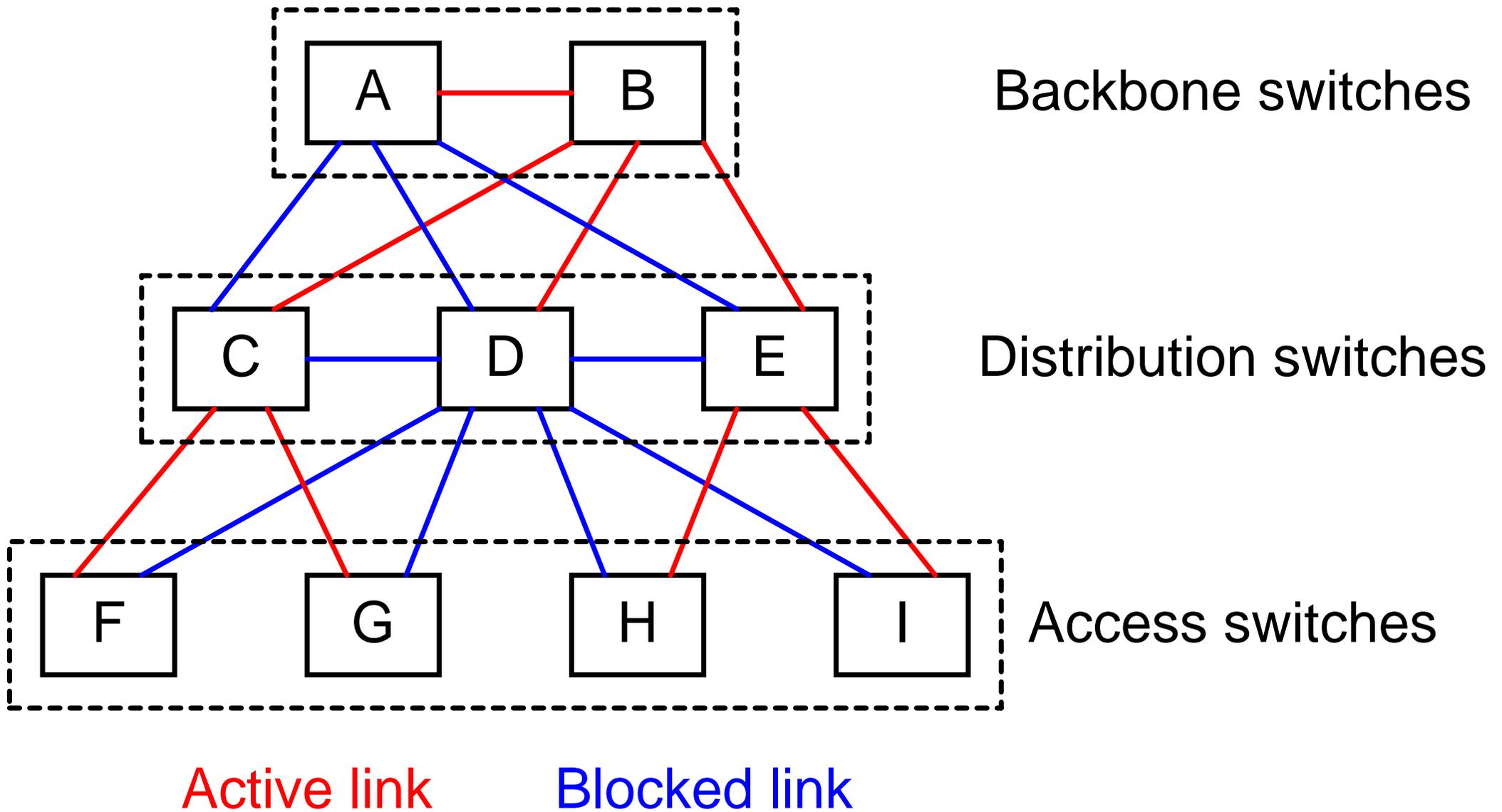
http://www.tdeig.ch/vmware/Ifakren_RTb.pdf

Trafic Nord-Sud & Est-Ouest

- Terminologie commune à plusieurs entreprises (Cisco, VMware, ...) pour distinguer les **flux internes & externes** à l'entreprise

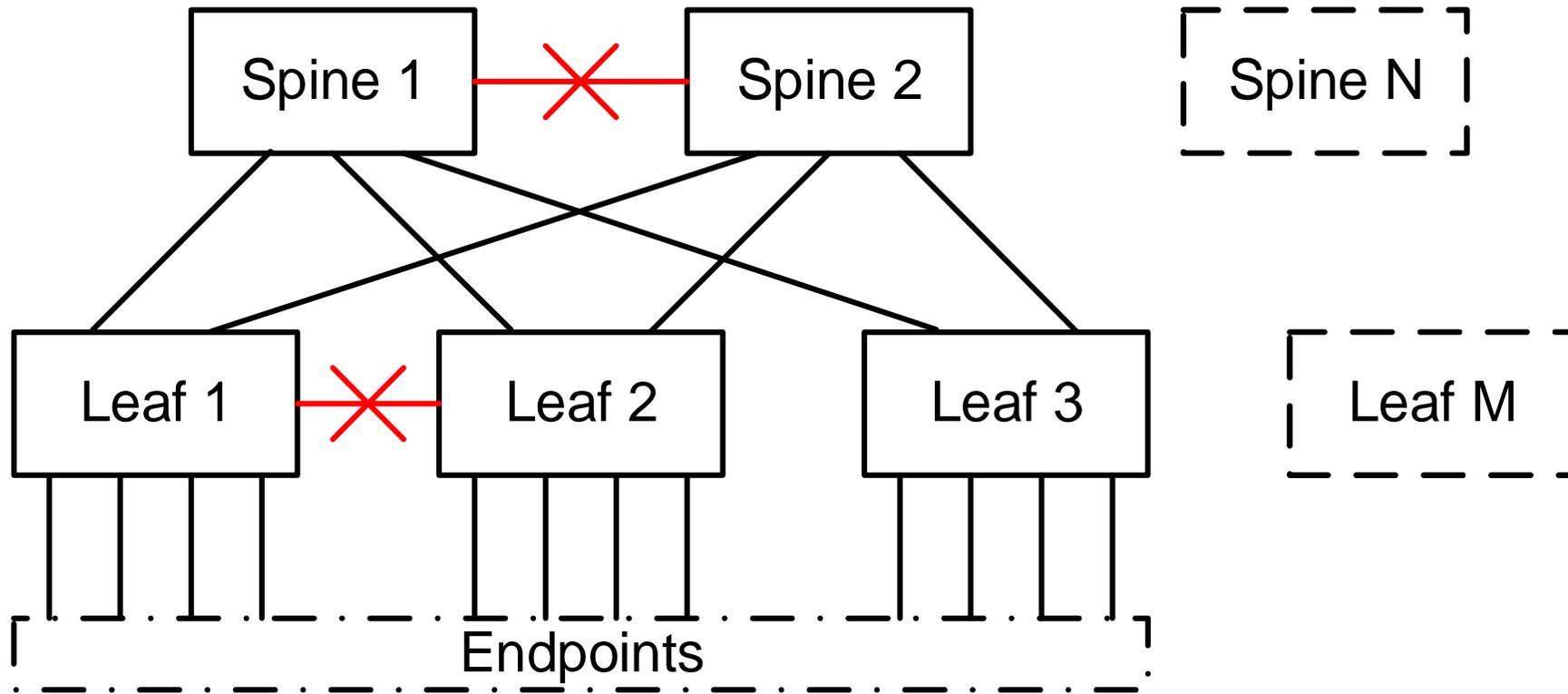


Switches in a hierarchical network



Architecture Spine – Leaf (colonne vertébrale – feuille)

- Modèle physique pour implémenter le **switch idéal (Layer 2)**



- Les **endpoints** sont reliés uniquement aux leaves
- Les **performances** (via 2 leaves + 1 spine) sont **déterministes**
- Facile d'ajouter des endpoints en ajoutant des leaves
- Facile d'augmenter la bande passante en ajoutant des spines

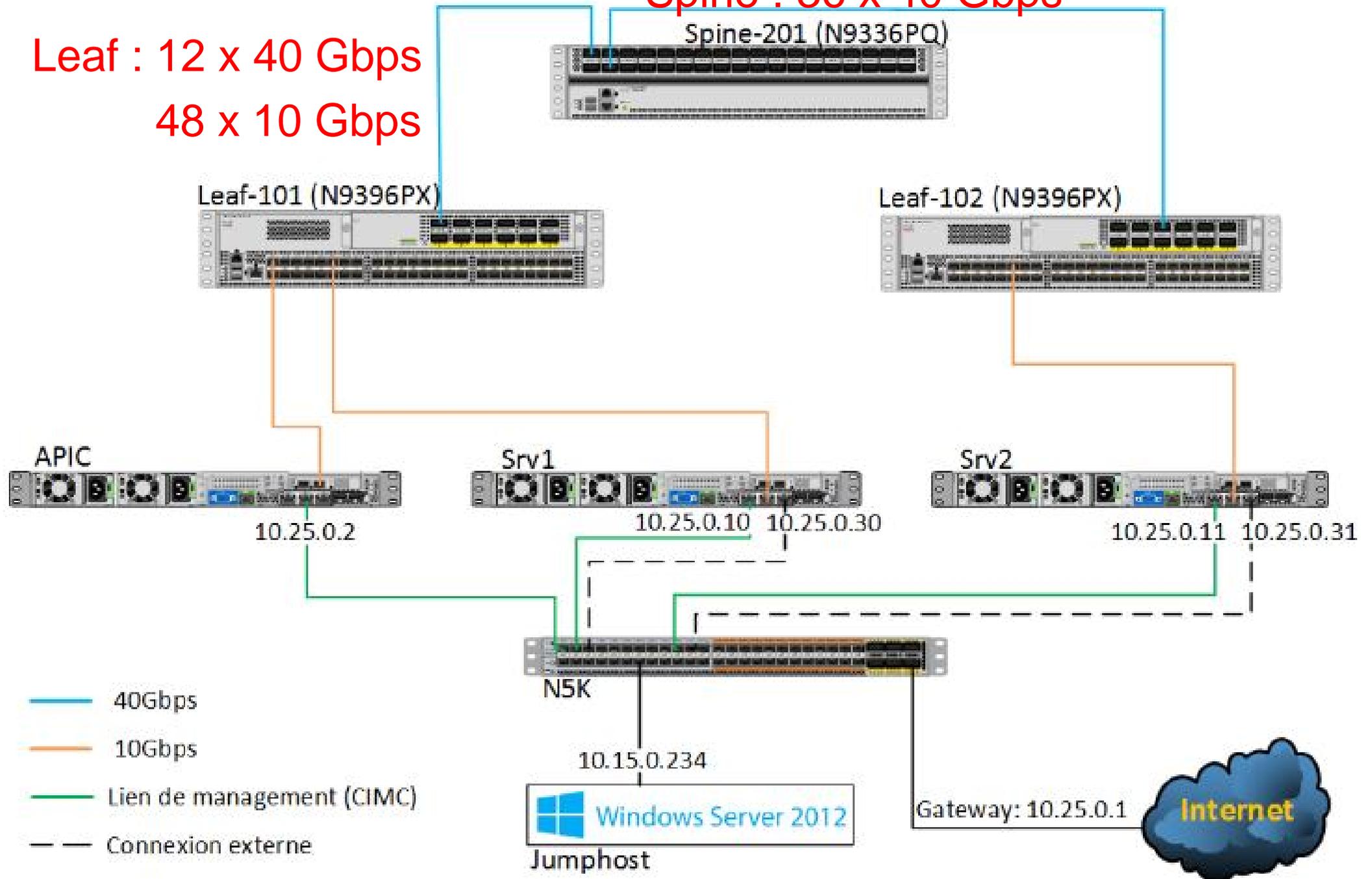
Caractéristiques

- Les équipements utilisateurs (endpoints) croient utiliser un commutateur Ethernet géant qui cache une certaine complexité
- Chaque spine connaît tous les endpoints (IP - Eth)
- Le commutateur leaf possède une vue locale
- Les paquets Leaf – Spine – Leaf (dans la fabric) utilisent le protocole VXLAN

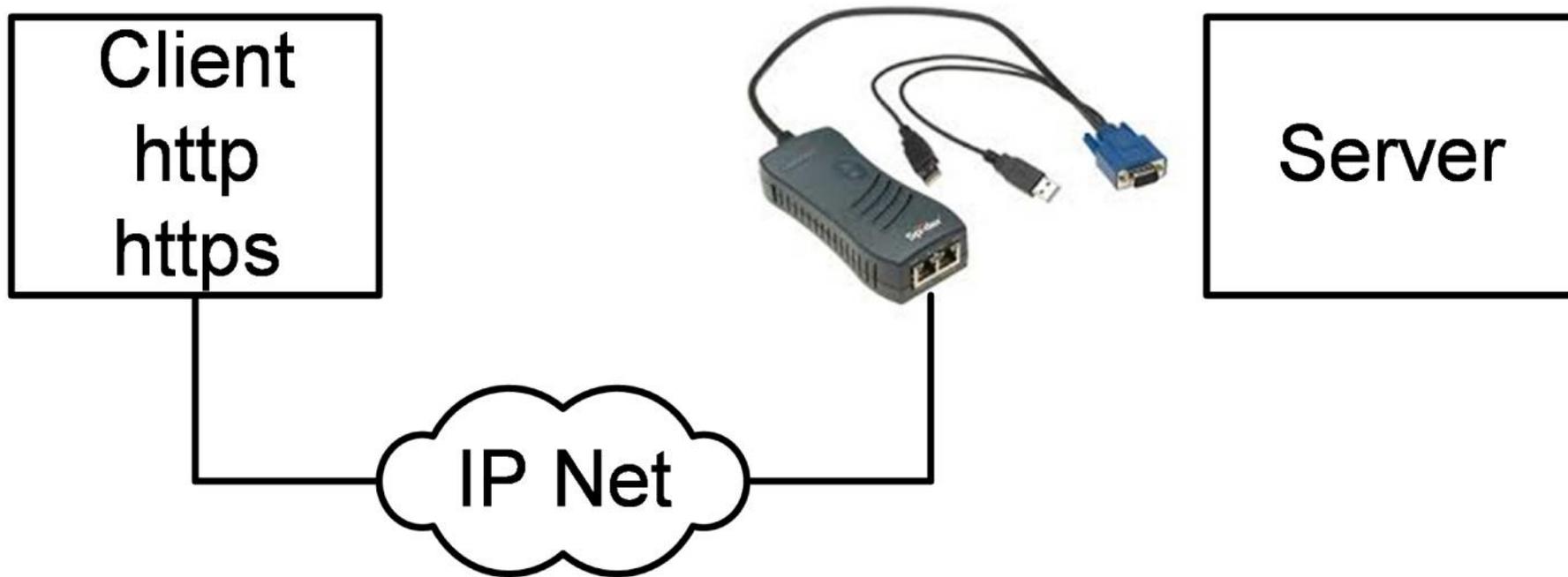
Cisco Nexus 9000 & ACI (dipl Udriot 2016)

Spine : 36 x 40 Gbps

Leaf : 12 x 40 Gbps
48 x 10 Gbps



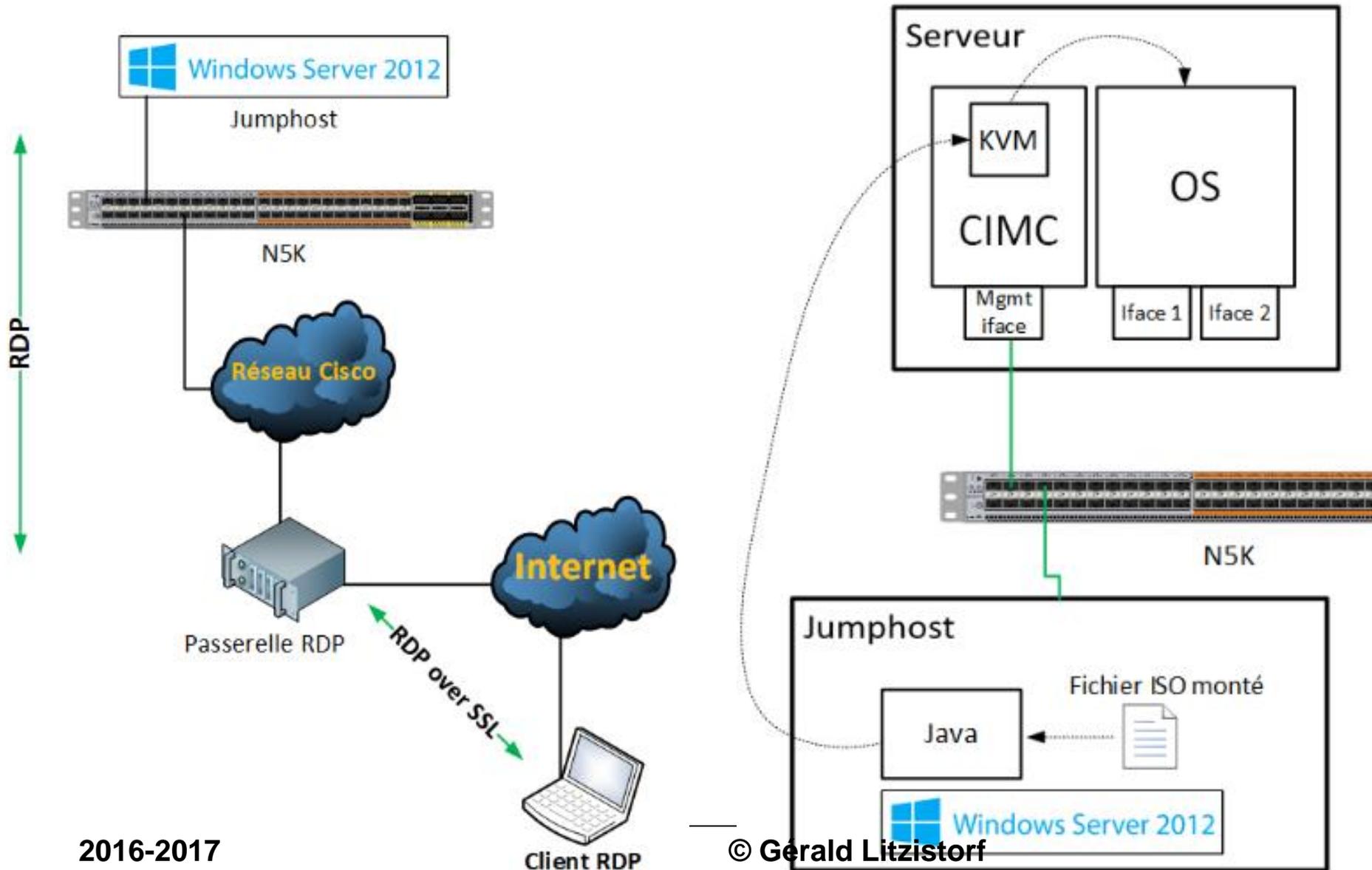
Remote access via **KVM module** to Server



- La console (clavier-souris-écran) des serveurs physiques est accessible à distance via un réseau IP
- 1 module **KVM (Keyboard-Video-Mouse)** par serveur
- Navigateur côté client
- Illustration avec module Spider de Lantronix
- <https://www.lantronix.com/products/lantronix-spider/>

CIMC (Cisco Integrated Management Controller)

- Les serveurs professionnels (Cisco, Dell, Lenovo, HP, ...) possèdent une interface d'administration propre au constructeur



Overall Server Status

 Good

Server

Admin

Storage

Summary

Inventory

Sensors

Remote Presence

BIOS

Power Policies

Faults and Logs

Troubleshooting

Server Summary

Actions

 Power On Server

 Power Off Server

 Shut Down Server

 Power Cycle Server

 Hard Reset Server

 Launch KVM Console

 Turn On Locator LED

Server Properties

Product Name: **UCS C220 M3S**

Serial Number: **FCH1945V1BA**

PID: **UCSC-C220-M3S**

UUID: **031C9653-704F-482F-BA53-28A18530B548**

BIOS Version: **C220M3.2.0.4b.0 (Build Date: 04/28/2015)**

Description:

Server Status

Power State:  On

Overall Server Status:  Good

Temperature:  Good

Overall DIMM Status:  Good

Power Supplies:  Good

Fans:  Good

Locator LED:  Off

Overall Storage Status:  Good

Cisco Integrated Management Controller (Cisco IMC) Information

Hostname: **C220-FCH1945V1BA**

IP Address: **10.25.0.10**

MAC Address: **58:AC:78:F2:60:B2**

Firmware Version: **2.0(6d)**

Current Time (UTC): **Thu Jun 23 07:26:52 2016**

Local Time: **Thu Jun 23 07:26:52 2016 UTC +0000**

Timezone: **UTC** ([Select Timezone](#))

Overall Server Status
 Good



LSI MegaRAID SAS 9271-8i (SLOT-1)

Server Admin **Storage**

Controller Info Physical Drive Info Virtual Drive Info Battery Backup Unit Storage Log

- LSI MegaRAID SAS 9271-8i (SLOT-1)
- Cisco FlexFlash

- Actions**
- Create Virtual Drive from Unused Physical Drives**
 - Create Virtual Drive from an Existing Virtual Drive Group
 - Import Foreign Config
 - Clear Foreign Config
 - Clear Boot Drive
 - Get TTY Log

Health/Status

Composite Health:  Good

Controller Status: **Optimal**

RAID Chip Temperature: **68 °C**

TTY Log Status: **Not Downloaded**

Firmware Versions

Product Name: **LSI MegaRAID SAS 927...**

Serial Number: **SV51822622**

Firmware Package Build: **23.33.0-0023**

PCI Info

PCI Slot: **SLOT-1**

Vendor ID: **1000**

Device ID: **5b**

SubVendor ID: **1000**

SubDevice ID: **9271**

Manufacturing Data

Manufactured Date: **2015-04-28**

Revision No: **33B**

Settings

Predictive Fail Poll Interval: **300 sec**

Rebuild Rate: **30 %**

Patrol Read Rate: **30 %**

Consistency Check Rate: **30 %**

Reconstruction Rate: **30 %**

Cache Flush Interval: **4 sec**

Max Drives To Spin Up At Once: **2**

Delay Among Spinup Groups: **12 sec**

Physical Drive Coercion Mode: **1 GB**

Cluster Mode: **false**

Battery Warning: **true**

ECC Bucket Leak Rate: **1440 min**

Expose Enclosure Devices: **true**

Maintain PD Fail History: **false**

Enable Copyback on SMART: **false**

Enable Copyback to SSD on SMART Error: **true**

Native Command Queuing: **enabled**

JBOD: **false**

Enable Spin Down of Unconfigured Drives: **true**

Enable SSD Patrol Read: **false**

AutoEnhancedImport: **true**

Capabilities

RAID Levels Supported: **Raid 0
Raid 1
Raid 5
Raid 6
Raid 10
Raid 50**

LSI MegaRAID SAS 9271-8i (SLOT-1)

Controller Info

Physical Drive Info

Virtual Drive Info

Battery Backup Unit

Storage Log

Physical Drives

Controller	Physical Drive Number	Status	Health	Boot Drive	Drive Firmware	Coerced Size	Manufacturer Model	Type
SLOT-1	1	Online	Good	false	5701	285148 MB	TOSHIBA	HDD
SLOT-1	2	Online	Good	false	5701	285148 MB	TOSHIBA	HDD
SLOT-1	3	Online	Good	false	5701	285148 MB	TOSHIBA	HDD
SLOT-1	4	Online	Good	false	5701	285148 MB	TOSHIBA	HDD

General

Enclosure Device ID: **252**
Physical Drive Number: **1**
LSI Drive Number: **0**
Power State: **active**
Device ID: **11**
Sequence Number: **2**
Media Error Count: **0**
Other Error Count: **0**
Predictive Failure Count: **0**
Raw Size: **286102 MB**
Non Coerced Size: **285590 MB**
Coerced Size: **285148 MB**
Block Size: **512**
SAS Address: **0: 5000039678112c02**
1: 0x0
Link Speed: **6.0 Gb/s**
Media Type: **HDD**
Interface Type: **SAS**

Actions

- Make Global Hot Spare
- Make Dedicated Hot Spare
- Prepare For Removal

Status

Locator LED: Turn On
 Turn Off
State: **online**
Fault: **false**
Online: **true**

Operation Status

Operation: **No operation in progress**
Progress in %: **0**
Elapsed Time (secs): **0**
[Refresh](#)

Inquiry Data

Product ID: **AL135XB300N**

LSI MegaRAID SAS 9271-8i (SLOT-1)

- Controller Info
- Physical Drive Info
- Virtual Drive Info**
- Battery Backup Unit
- Storage Log

Virtual Drives

Virtual Drive Number	Name	Status	Health	Size	RAID Level	Boot Drive
0	RAID1_1234	Optimal	Good	570296 MB	RAID 1	true

Actions

- Initialize
- Set as Boot Drive
- Delete Virtual Drive
- Edit Virtual Drive

Operation Status

Operation: **No operation in progress**
 Progress in %: **0**
 Elapsed Time (secs): **0**
[Refresh](#)

General

Name: **RAID1_1234**
 Strip Size: **64 KB**
 Drives Per Span: **4**
 Span Depth: **1**
 Access Policy: **Read-Write**
 Cache Policy: **Direct**
 Read Ahead Policy: **Always**
 Requested Write Cache Policy: **Write Back Good BBU**
 Current Write Cache Policy: **Write Back**
 Disk Cache Policy: **Unchanged**
 Allow Background Init: **true**

Physical Drives

Virtual Drive Number	Physical Drive Number	Span	Starting Block	Number Of Blocks	State
0	1	0	0	583983104	online
0	2	0	0	583983104	online
0	3	0	0	583983104	online
0	4	0	0	583983104	online

Activate W

Server Summary

Actions

- Power On Server
- Power Off Server
- Shut Down Server
- Power Cycle Server
- Hard Reset Server
- Launch KVM Console**
- Turn On Locator LED

10.25.0.10 - KVM Console

Open a new tab (Ctrl+T)

File View Macros Tools Power Virtual Media Help



Press <F2> Setup, <F6> Boot Menu, <F7> Diagnostics, <F12> Network Boot

Bios Version : C220M3.2.0.4b.0.042820150826
Platform ID : C220M3

Cisco IMC IPv4 Address : 10.25.0.10
Cisco IMC MAC Address : 58:AC:78:F2:60:B2

/ Loading LSI EFI SAS Driver
Total Memory = 64 GB Effective Memory = 64 GB
Memory Operating Speed 1600 Mhz

Entering boot selection menu...



Remote Presence

Virtual KVM Virtual Media Serial over LAN

Actions

- Launch KVM Console

vKVM Properties

Enabled:

Max Sessions: 4

Active Sessions: 0

Remote Port: 2068

Enable Video Encryption:

Enable Local Server Video:

10.25.0.10 admin 6.8 fps 28.247 KB/s

Inventory

CPU2

CPU1

Memory

Power Supplies

PCI Adapters

Cisco VIC Adapters

Network Adapters

CPU1

Socket Name: **CPU1**

Vendor: **Intel(R) Corporation**

Status: **Enabled**

Family: **Xeon**

Speed(Mhz): **3300**

Version: **Intel(R) Xeon(R) CPU E5-2643 0 @ 3.30GHz**

Number of Cores: **4**

Signature: **Type 0, Family 6, Model 45, Stepping 7**

Number of Threads: **8**

Num

Num

CPU1

Memory

Power Supplies

PCI Adapters

Cisco VIC Adapters

Network Adapters

Storage Ad

Summary

Memory Speed: **1600 MHz**

Failed Memory: **0 MB**

Total Memory: **65536 MB**

Ignored Memory: **0 MB**

Effective Memory: **65536 MB**

Number of Ignored DIMMs: **0**

Redundant Memory: **0 MB**

Number of Failed DIMMs: **0**

Memory RAS Possible: **Independent Mirroring Lockstep**

Memory Configuration: **Independent**



[DIMM location diagram](#)

Network Adapter Inventory

CPU's

Memory

Power Supplies

PCI Adapters

Cisco VIC Adapters

Network Adapters

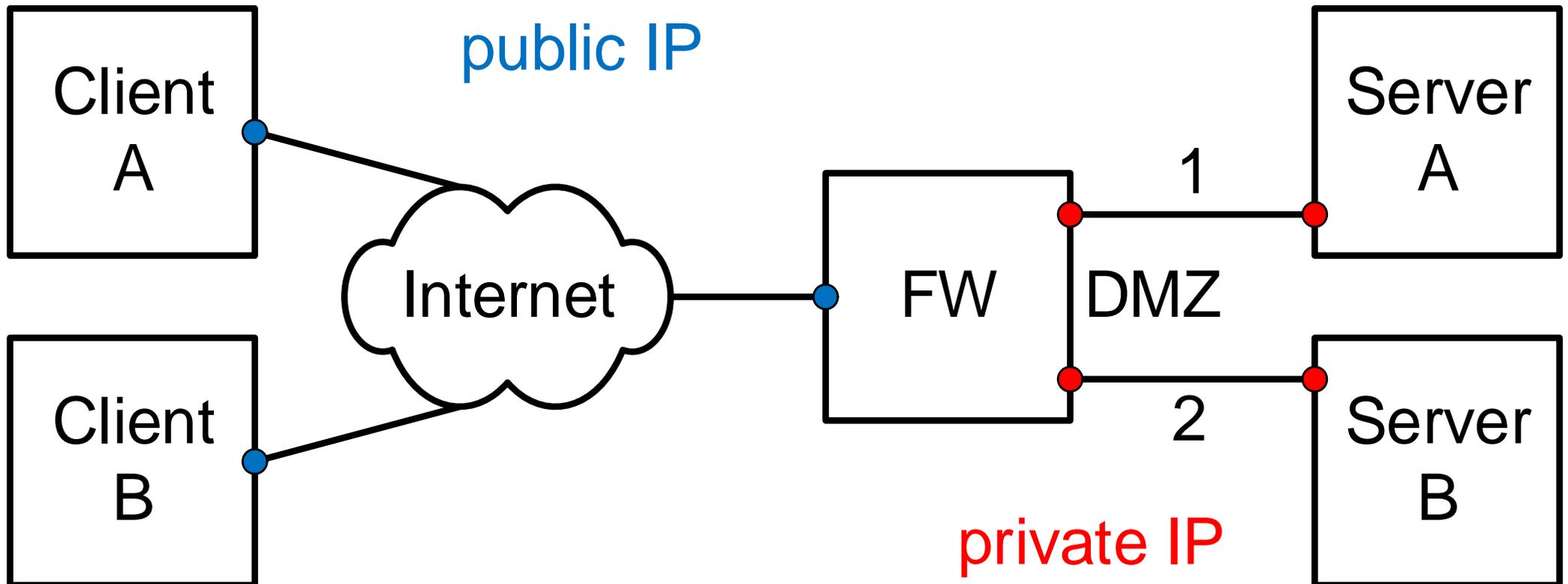
Network Adapters

Slot ID	Product Name
2	Intel Ethernet Server Adapter I350-T4
L	Intel Onboard 1Gbps Ethernet Adapter

External Ethernet Interfaces

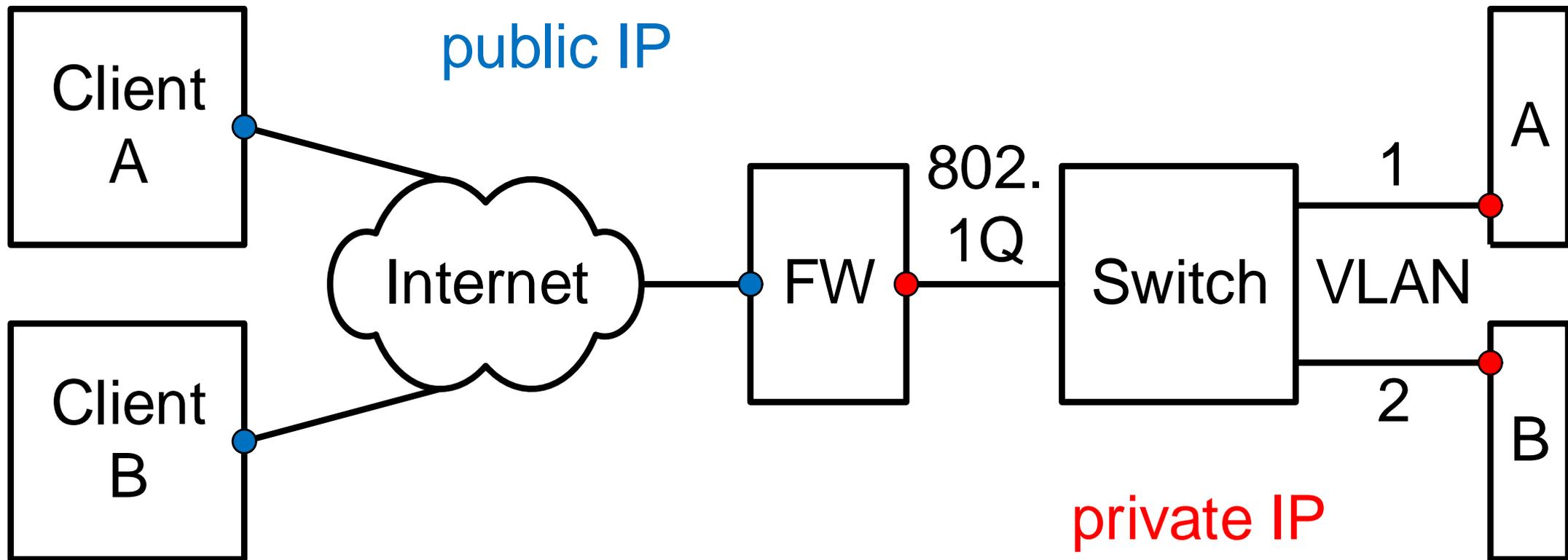
ID	MAC Address
1	a0:36:9f:88:65:28
2	a0:36:9f:88:65:29
3	a0:36:9f:88:65:2a
4	a0:36:9f:88:65:2b

Architecture physique d'un hébergeur web



- Firewall isole Server A de Server B → **1 interface DMZ / Server**
- 1 adr. publique par Server
- FW = pfSense : proxy ARP + IP forwarding + rule
- http://www.tdeig.ch/Schema_Reseau.pdf

Architecture physique d'un hébergeur web avec VLAN



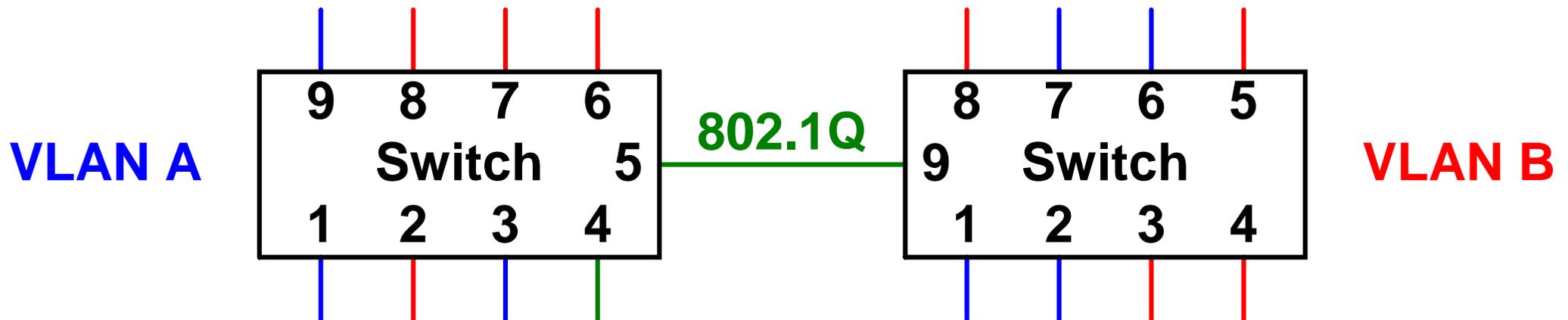
- Switch isole Server A de Server B → **1 VLAN / Server**
- Firewall doit permettre des règles contenant l'identifiant de VLAN
- **Plus de limitation du nombre de servers**

Avantages & inconvénients des VLAN

- Avantages

Chaque VLAN est isolé des autres

→ les paquets du **VLAN A** ne sont pas visibles (wireshark) sur le **VLAN B**



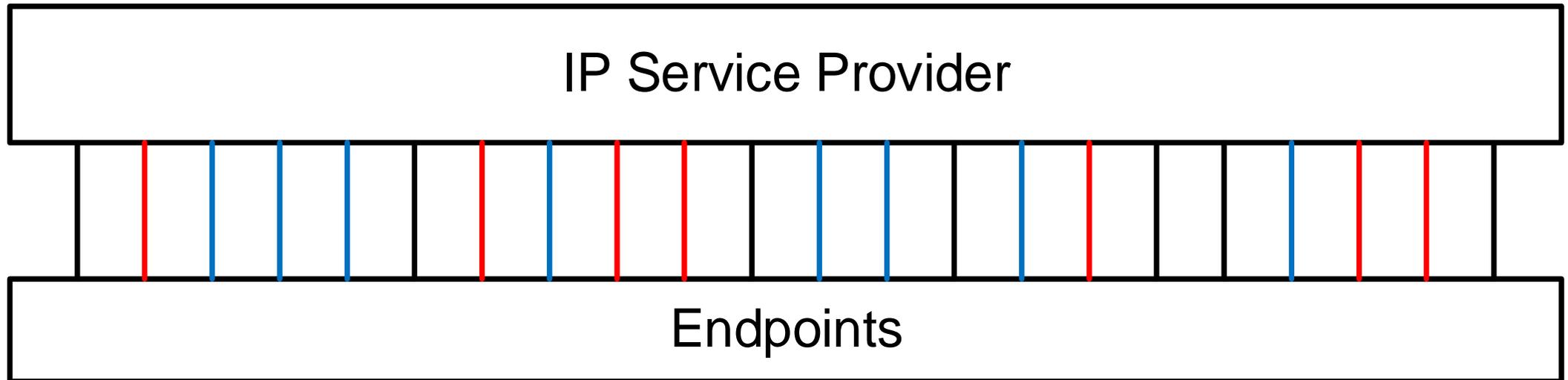
- Inconvénients

Leur portée est locale (LAN) car ils ne sont pas routables

→ ils utilisent une couche 802.1Q entre les en-têtes IP et Ethernet

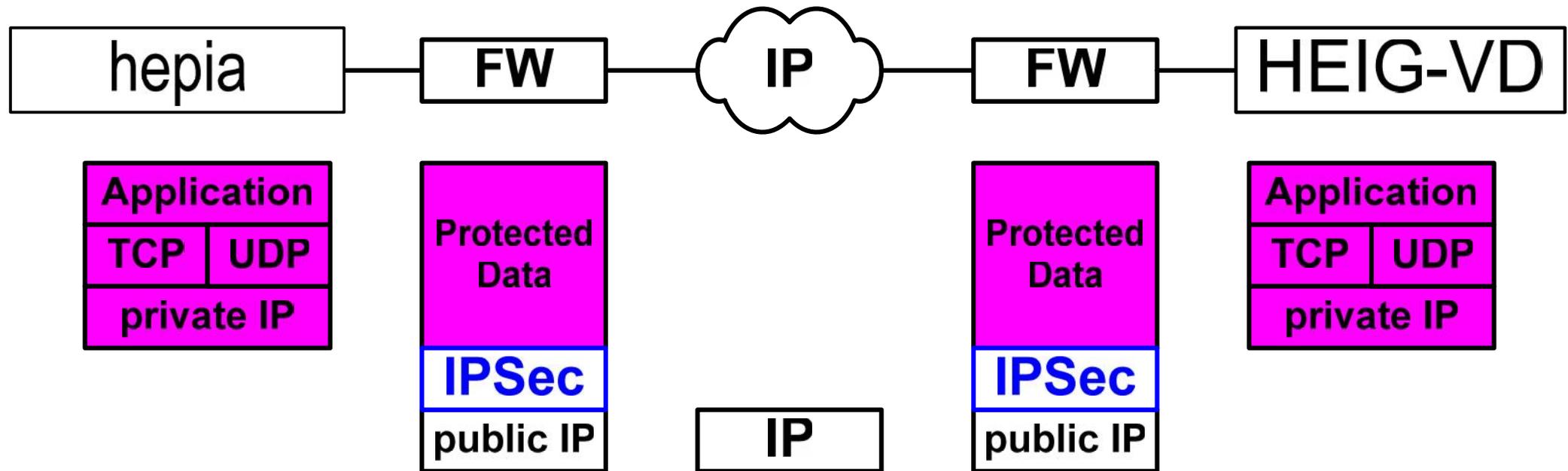
Réseau privé virtuel (Virtual Private Network)

- Comment un fournisseur de services IP peut garantir l'isolation des flux entre ces endpoints : **Tenant 1** **Tenant 2** Tenant 3 ?
- On parle d'architecture multi tenancy, ...



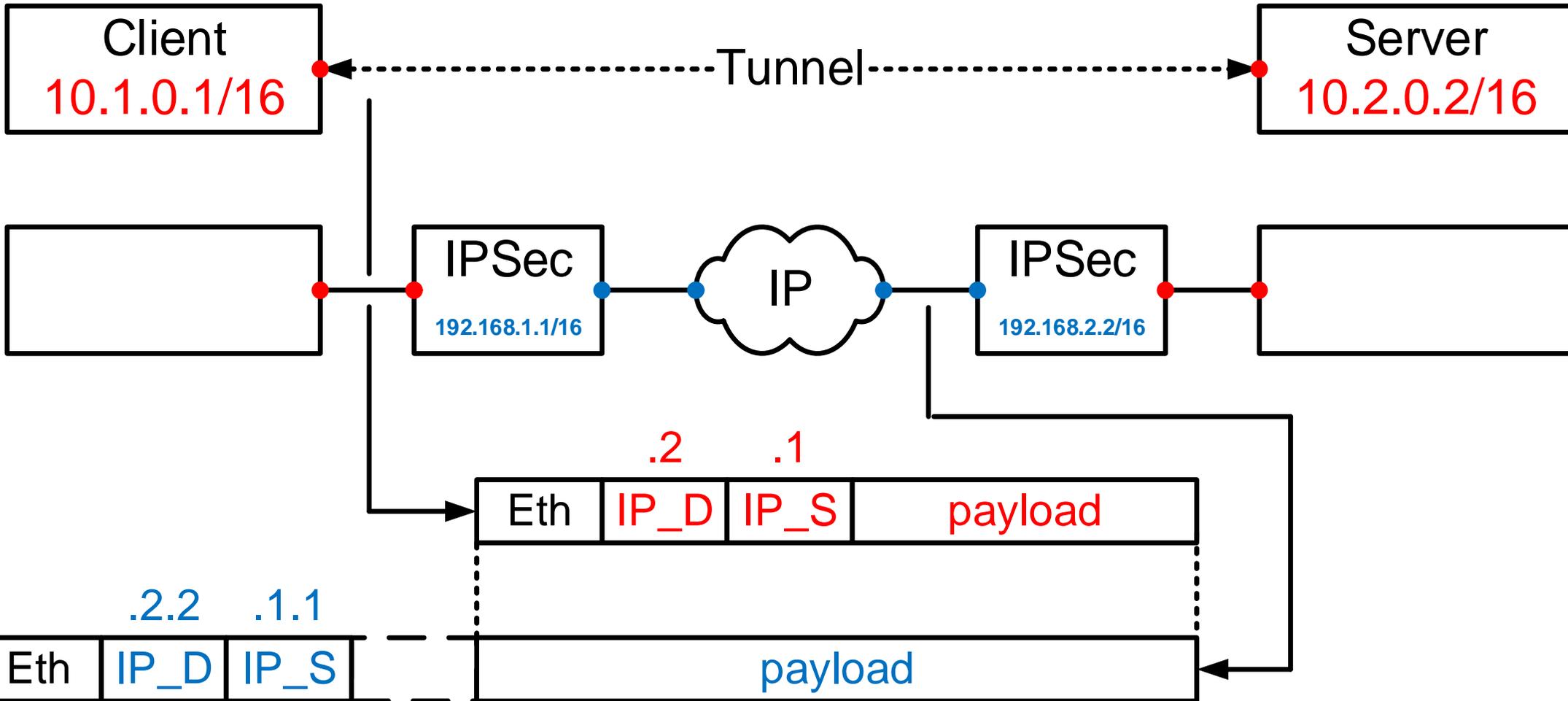
- 1975 → X.25 1985 → Frame Relay 1995 → MPLS
- 2005 → **IPSec**
- 2015 → **VXLAN**

IPSec-VPN (Virtual Private Network) hepia – HEIG-VD



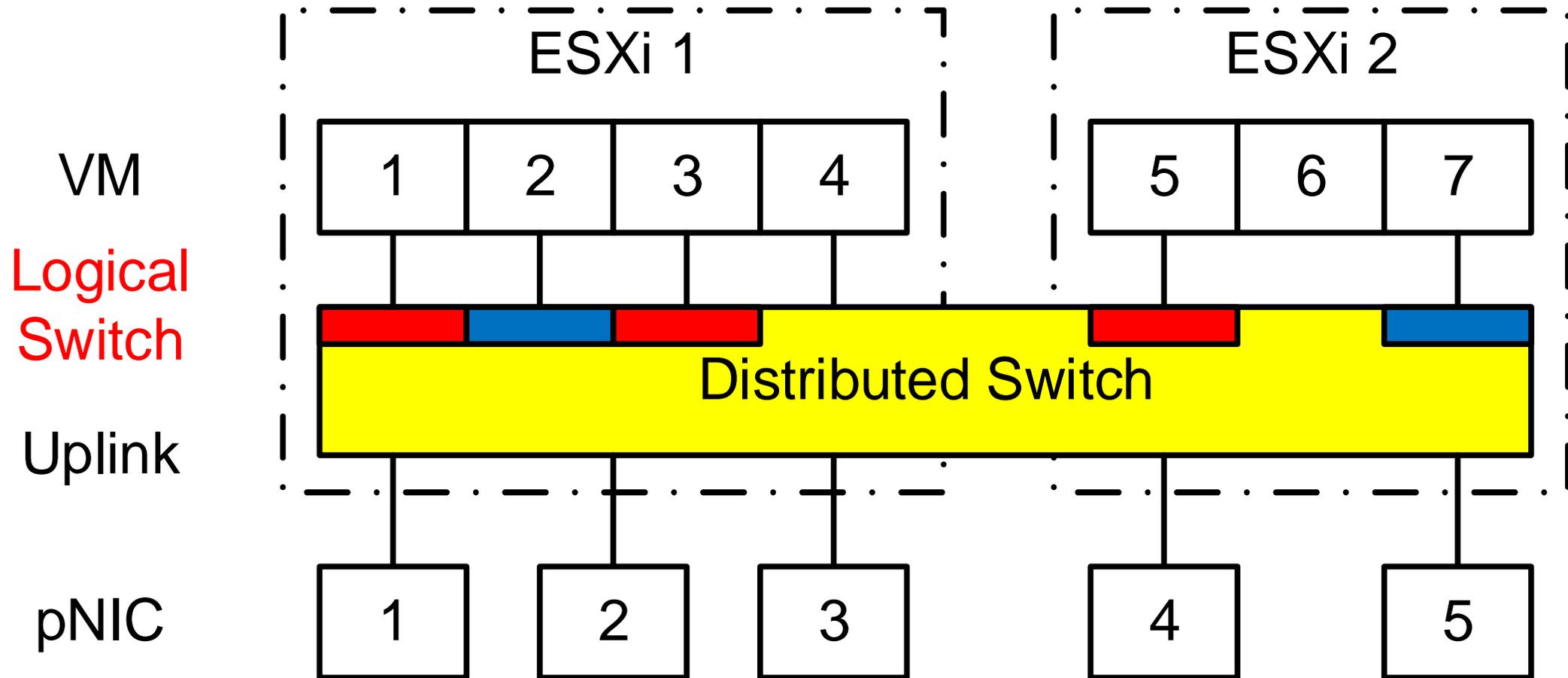
- 2 firewalls (Checkpoint - Linux) gèrent IPSec (analogue à SSL)
- **Adresses IP privées et publiques** → **mode tunnel**
- Tunnel créé entre les firewalls
- **Host ne voit pas IPSec**
- Services d'authentification mutuelle, de confidentialité et d'intégrité

Comment le tunnel IPsec est-il construit ?



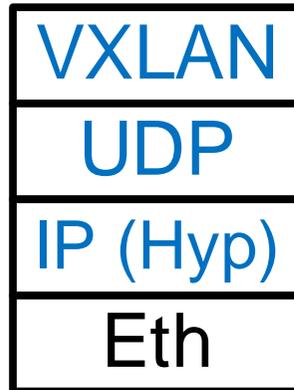
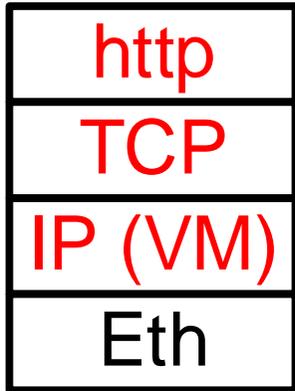
- Module IPsec ajoute puis retire l'en-tête gérée par le réseau IP
- Associations fixes $IP_D : IP_D$ dans les règles du module IPsec
- Overlay networks using tunnels → connectivity between isolated networks

VMware Distributed Switches



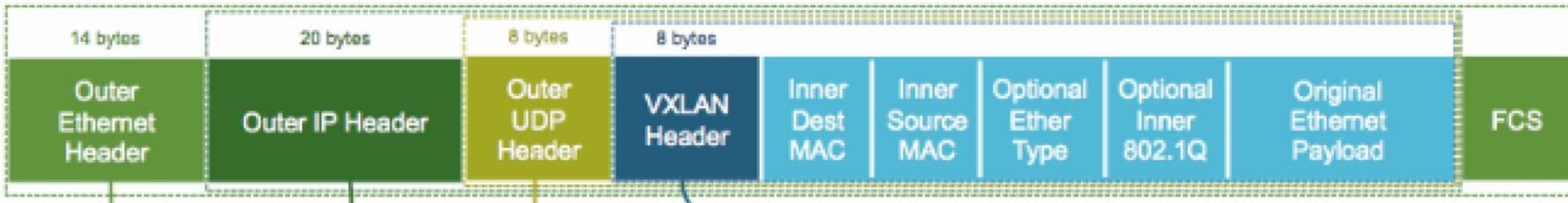
- **Un commutateur transversal** supporte les flux **Est-Ouest (VM to VM)**
- Chaque VM (via son **Logical Switch**) peut faire partie d'un ou plusieurs VLANs
- **Uplink** = lien logique avec un pNIC
- **Le flux (paquets) VM1 \leftrightarrow VM5** doit passer par le monde physique
- **Il va être encapsulé dans un VXLAN**

VXLAN = VLAN routable : encapsulation (tunnel)

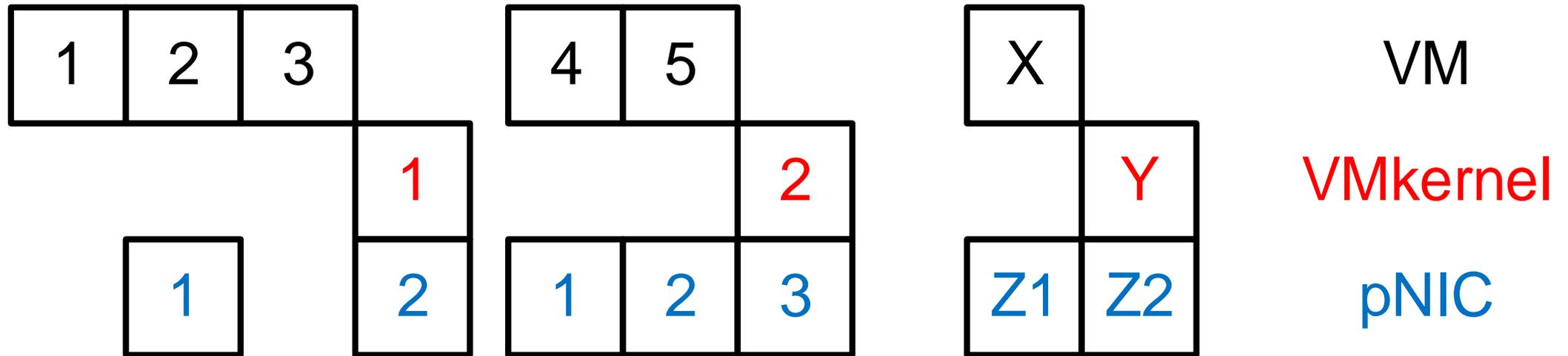


- Le paquet issu de **IP1** (destiné à **IP5**) est **encapsulé** dans un paquet VXLAN
- On parle d'un **tunnel** terminé par 2 endpoints = **VTEP** = **VXLAN tunnel endpoint**
- Default MTU = 1600 byte
- Rfc 7348 Supported by VMware, Cisco, Huawei, ...

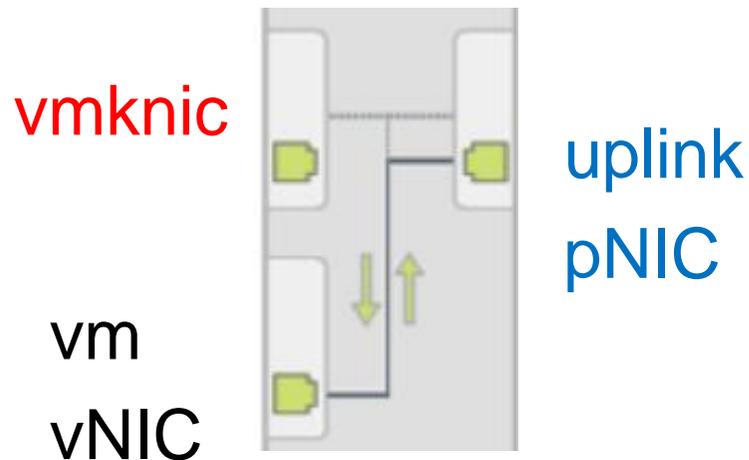
← Original Ethernet Frame →



Virtual Distributed Switch → vNIC – vmknic – pNIC



vCenter



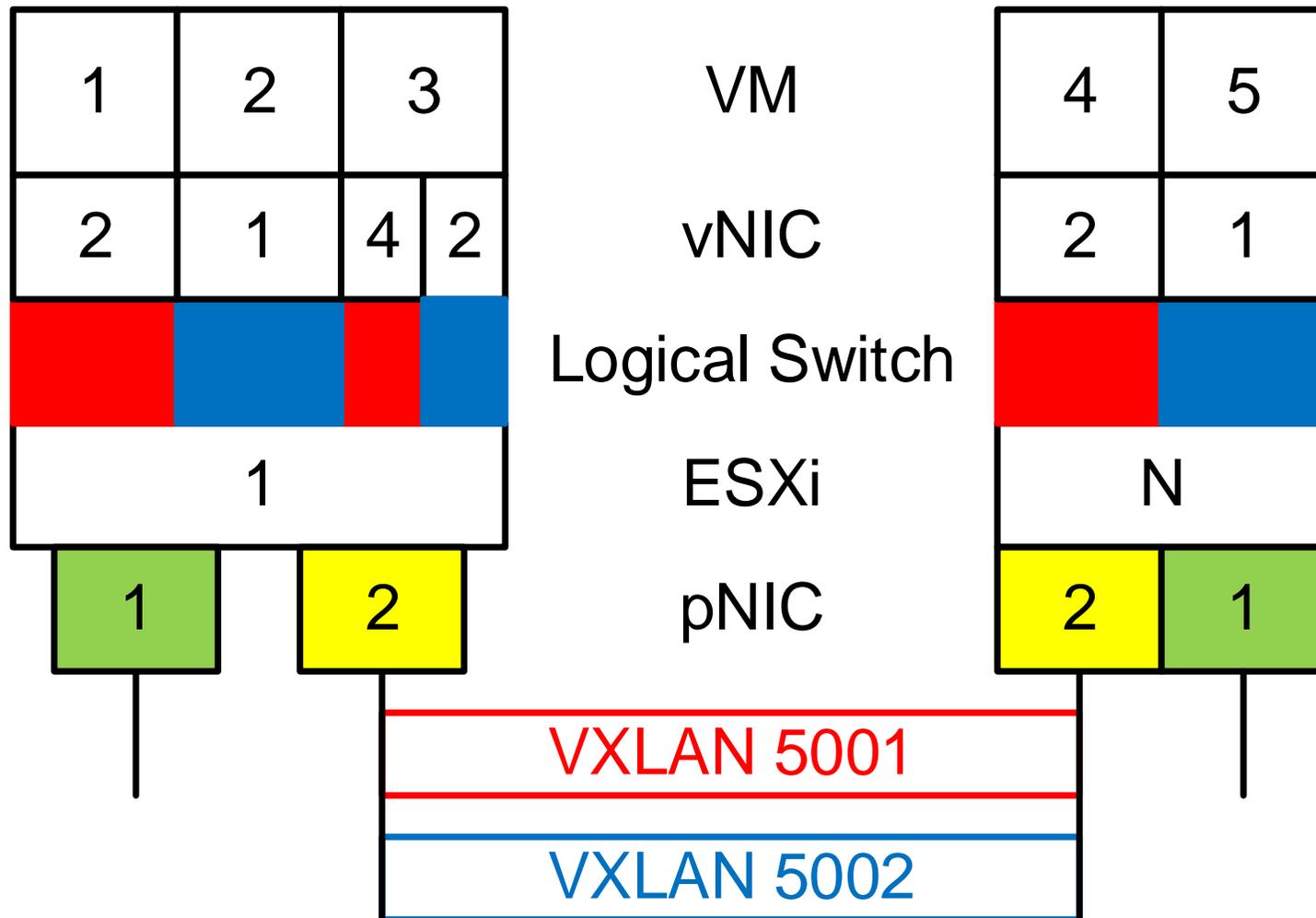
Réseau comprenant :

- X VMs (machine virtuelle-vNIC)
- Y hyperviseurs (VMkernel-vmknic)
- Z physical NIC (uplink) par hyperviseur

Logical Switch & VXLAN

Réseau comprenant :

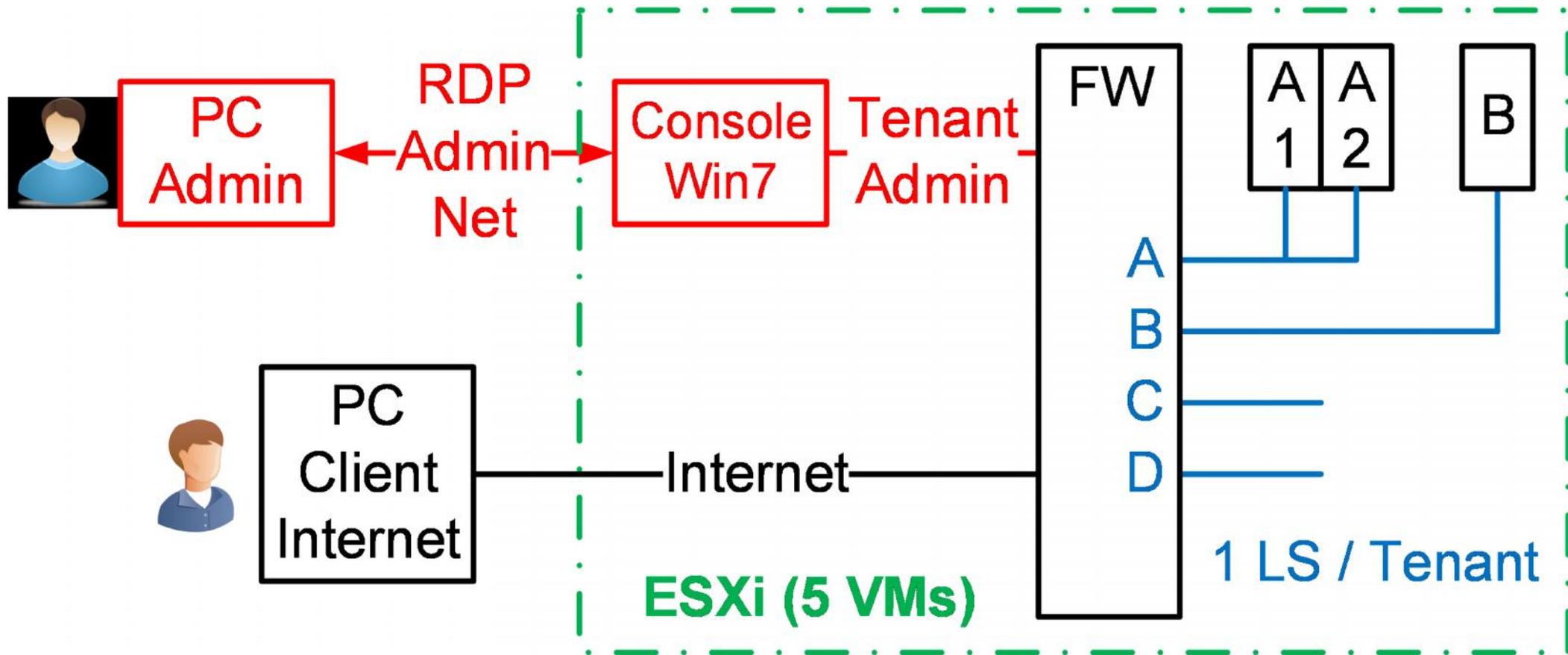
- 5 VMs
- **VLAN rouge**
VM 1-3-4 utilisent
VXLAN 5001



- **VLAN bleu** → VM 2,3,5 utilisent VXLAN 5002
- **Les flux VXLAN utilisent un réseau physique propre (pNIC2)**
1 adr IP (pool) au minimum par ESXi

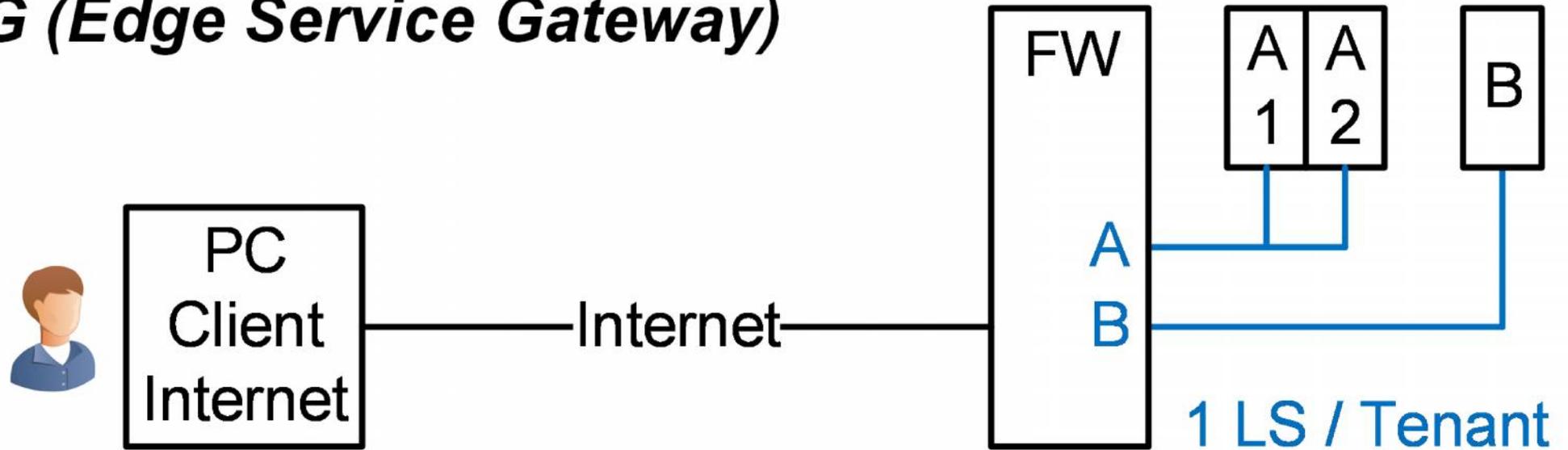
- Flux d'administration utilisent pNIC 1

Dipl Ifakren : Scénario multi-tenants (Tenants A – B)



- **Client internet** accède aux services hébergés dans les VMs A1, ... (scalability)
- **1 LS (VLAN) / Tenant** → isolation entre Tenant (pas d'isolation dans Tenant_A)
- **Hébergeur** administre, via RDP, toutes les VMs depuis la VM Console
- **3 réseaux physiques** (Admin / Internet / VXLAN)

ESG (Edge Service Gateway)



- Prévu pour autoriser les flux **Nord-Sud** (en liaison avec le **monde physique**)
- Implémenté dans une VM
- Fonctionnement identique à un firewall physique stateful

Source	Destination	Service	Action
any IP	129.194.184.81	http	allow
LS internet	LS Tenant 1		

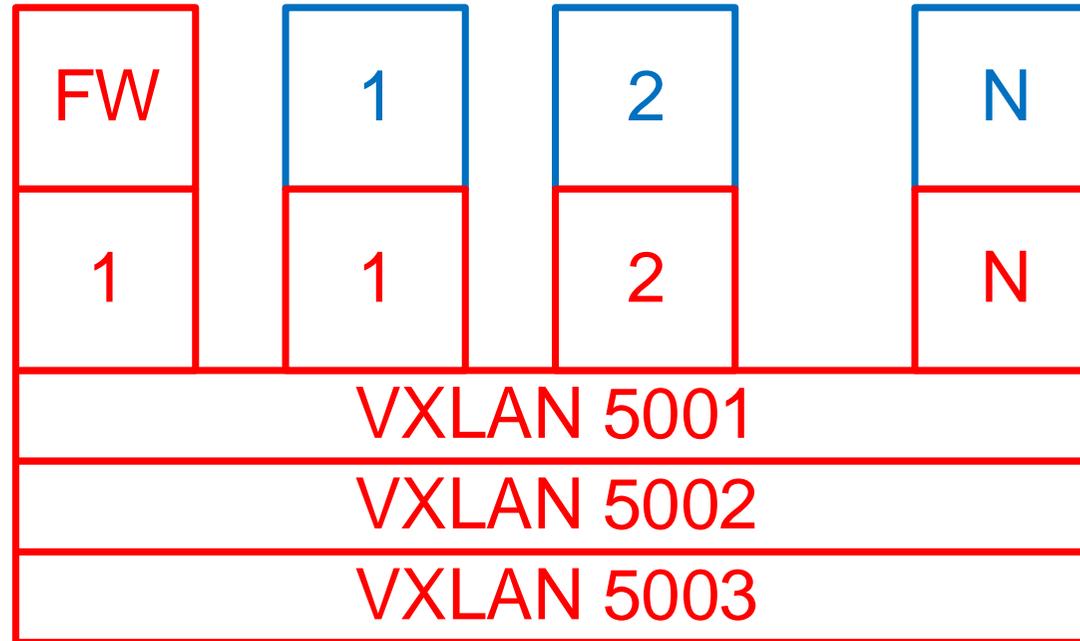
- Les flux dans les réseaux Tenant sont de type VXLAN

Management – Control – Data planes

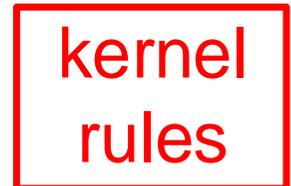
Gestion centralisée depuis vSphere
(Admin Console)



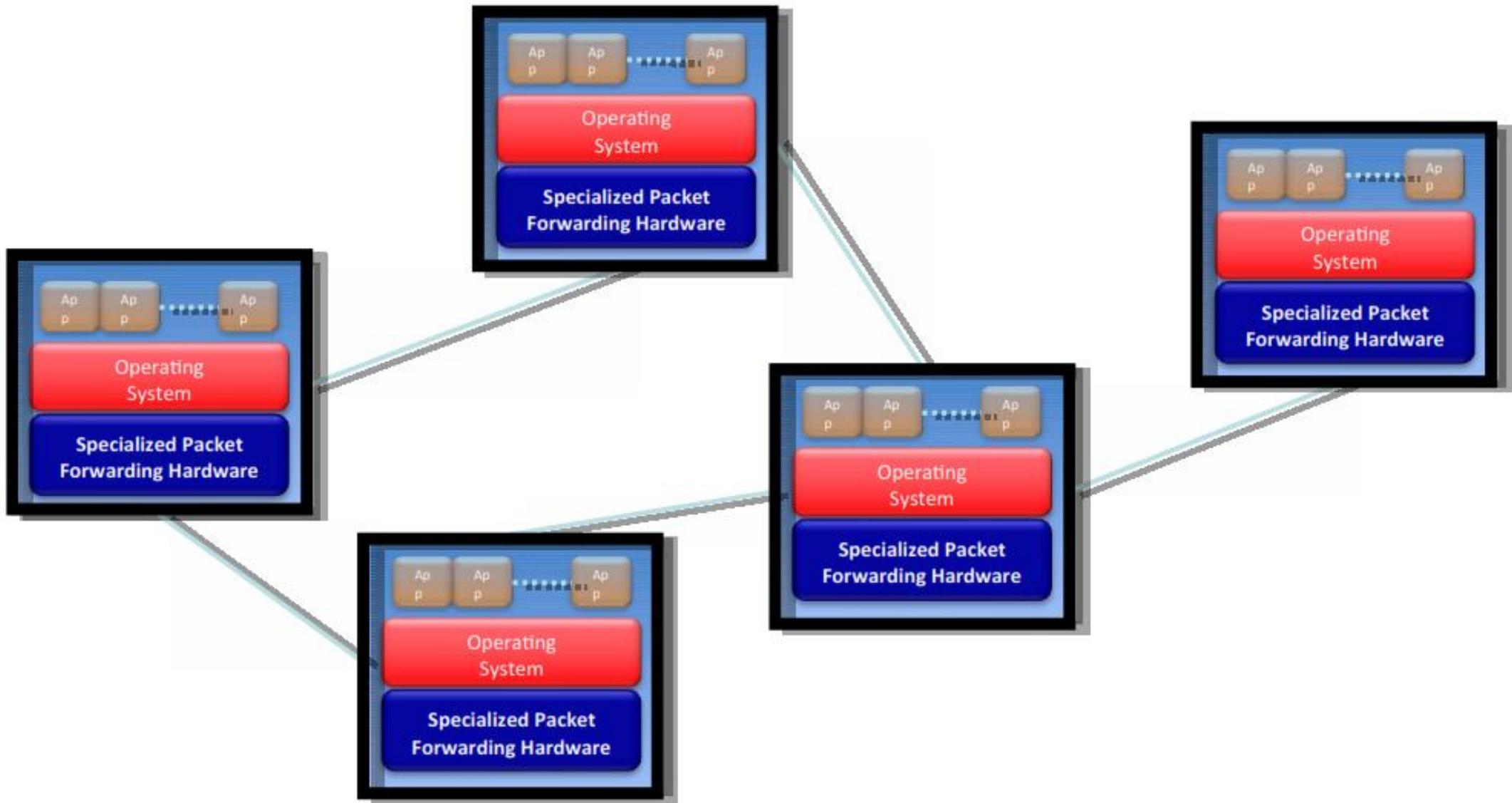
Gestion de NSX (management
plane) depuis webGUI



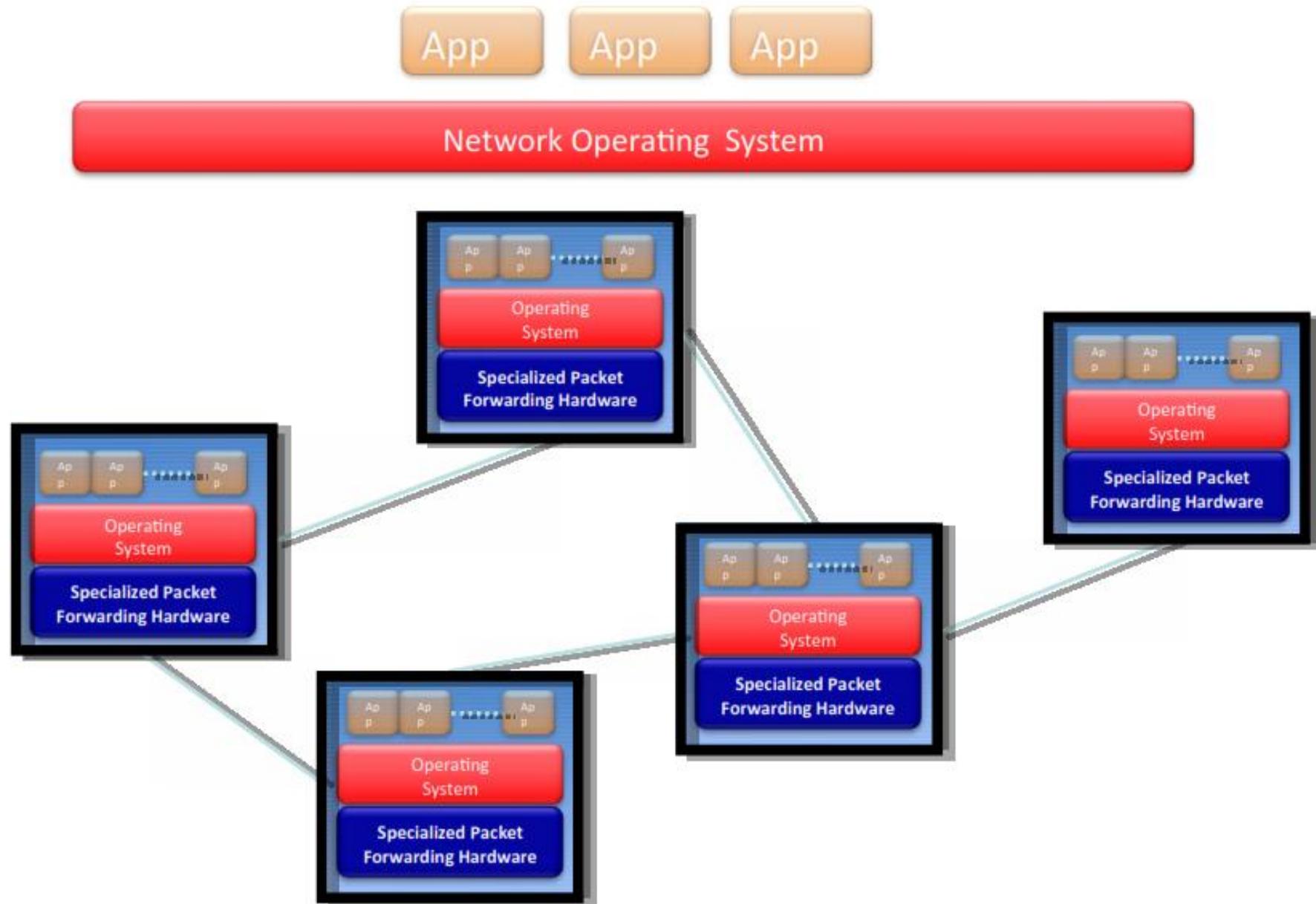
VM



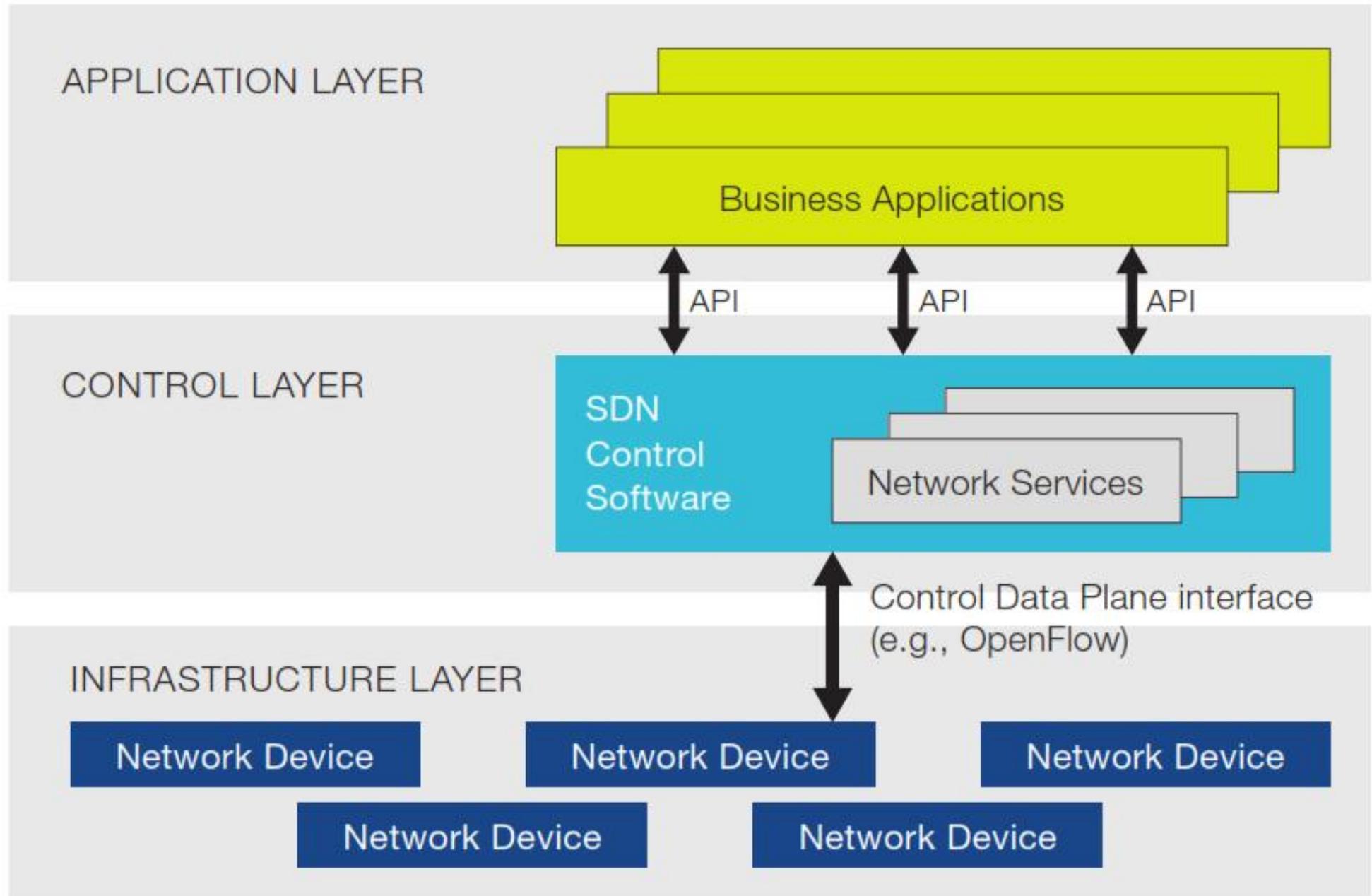
Internet – Intranet : closed boxes, distributed control



Software Defined Networking : centralized control



Software Defined Networking (SDN)

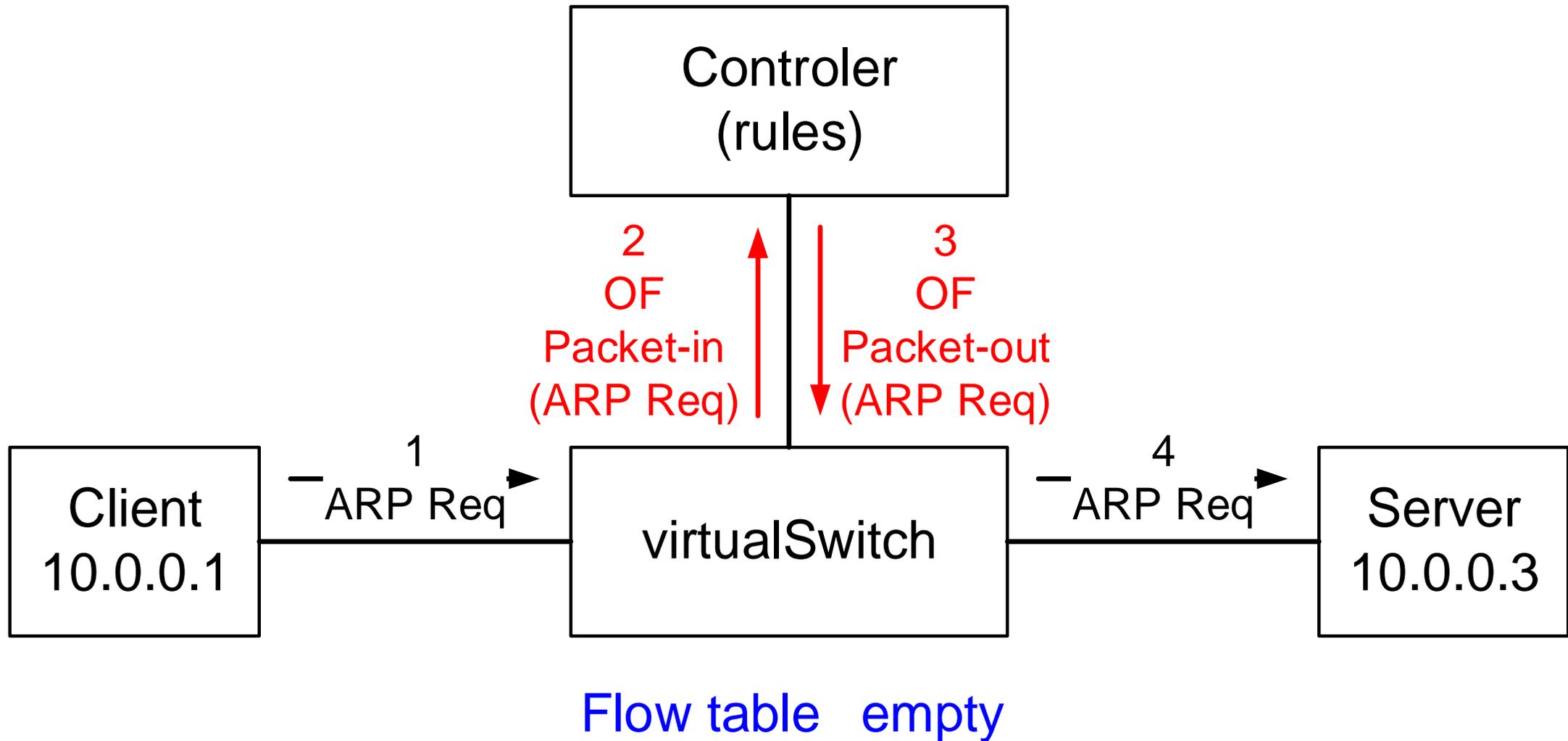


SDN

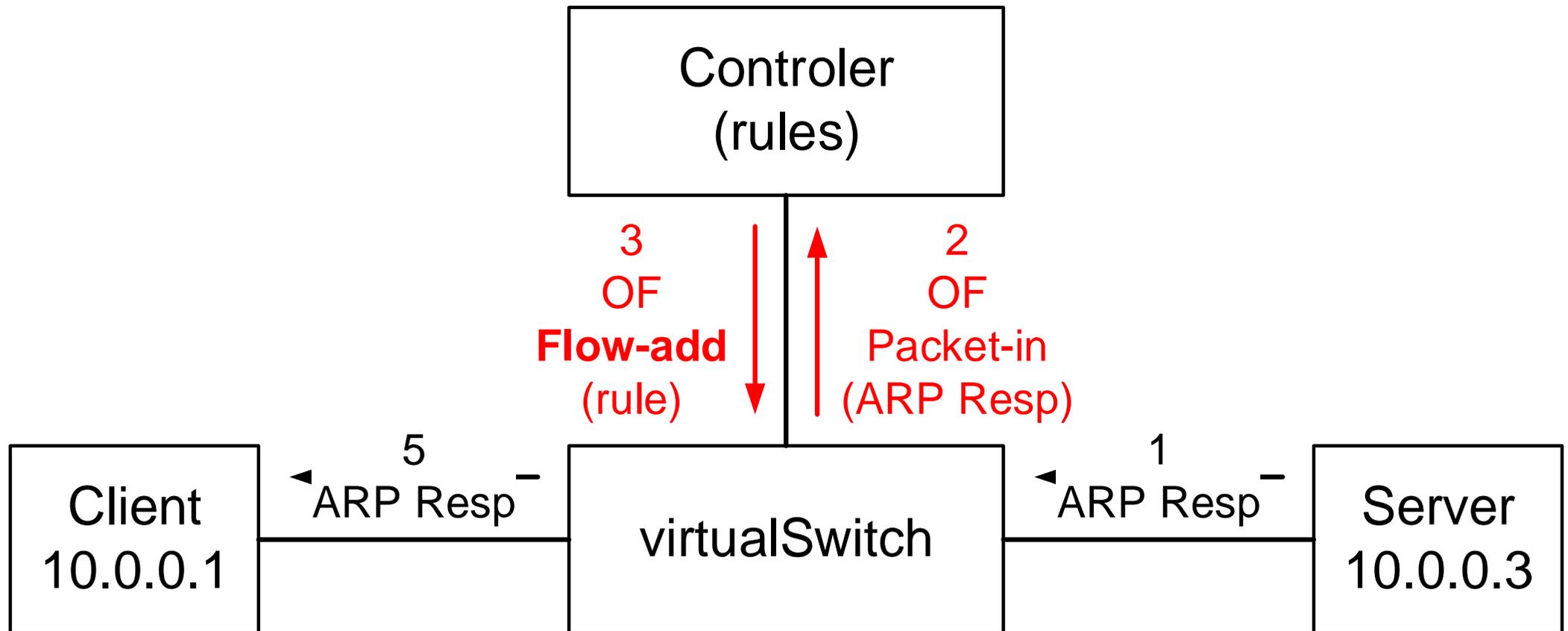
- *In the SDN architecture, the **control and data planes are decoupled**, network intelligence and state are logically **centralized**, and the underlying network infrastructure is **abstracted** from the applications.*
- *As a result, enterprises and carriers gain unprecedented programmability, automation and network control, enabling them to build highly scalable, flexible networks that readily adapt to changing business needs.*
- Source → <https://www.opennetworking.org/images/stories/downloads/sdn-resources/white-papers/wp-sdn-newnorm.pdf>
- Cisco → <http://blogs.cisco.com/news/three-truths-about-networking-the-next-chapter/>

Projet de semestre SDN par Mme Ifakren – 2015 (1/2)

- OpenFlow : packet-in & packet-out



Projet de semestre SDN par Mme Ifakren – 2015 (2/2)



Flow table

eth_src, eth_dst, ipv4_src, ipv4_dst, timeout

http://www.tdeig.ch/linux/lfakren_RPS.pdf page 27

OpenFlow Basics

< Match | Action >



Packet + byte counters

1. Forward packet to ports
2. Encapsulate and forward to controller
3. Drop packet
4. Send to normal processing pipeline
5. Any extensions you add!

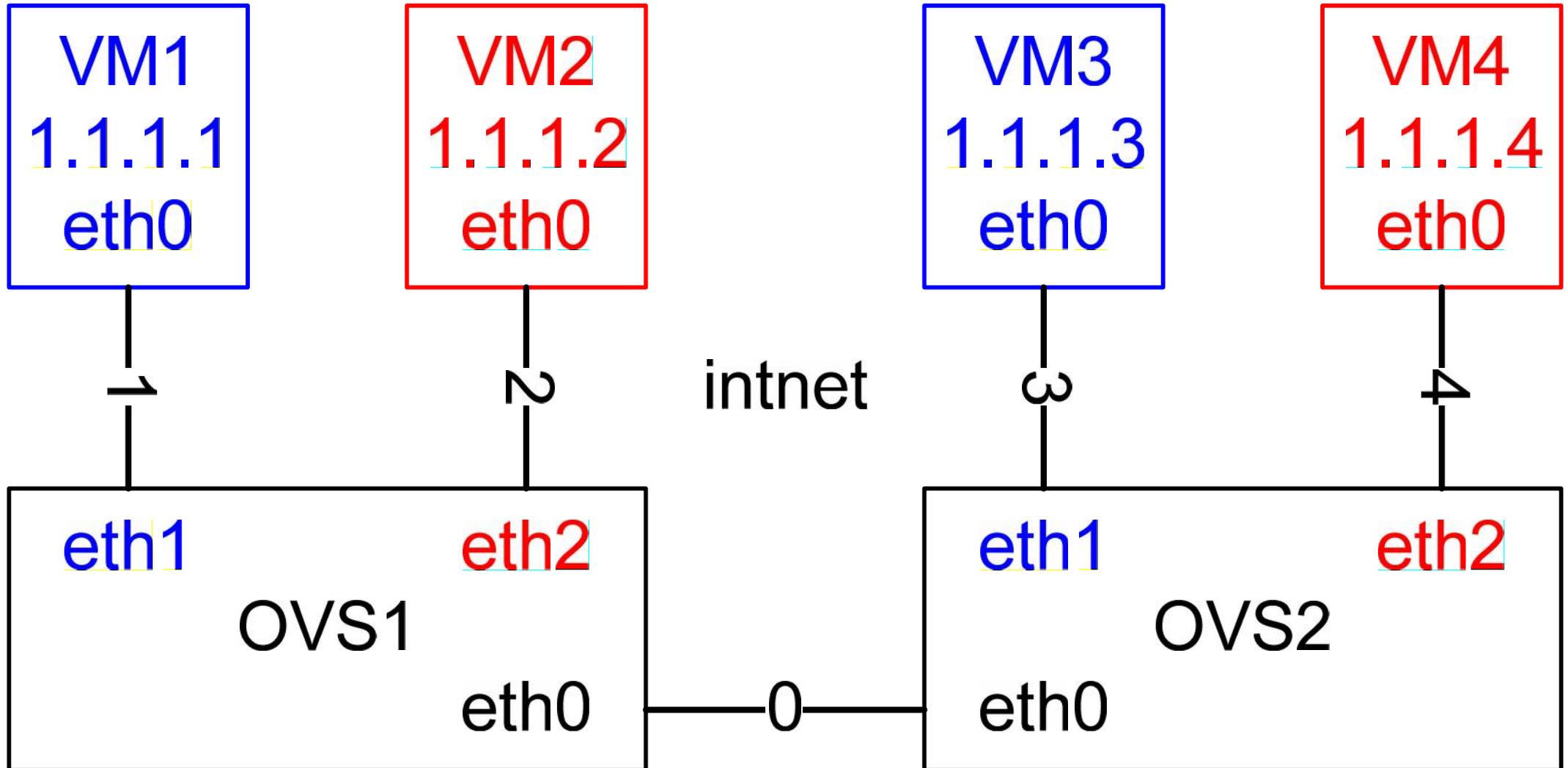
Switch Port	VLAN ID	VLAN pcp	MAC src	MAC dst	Eth type	IP Src	IP Dst	IP ToS	IP Prot	L4 sport	L4 dport
-------------	---------	----------	---------	---------	----------	--------	--------	--------	---------	----------	----------

Avantage du SDN

- **Contrôle centralisé du réseau** (vue d'ensemble, isoler des sous-réseaux pour maintenance, modifier les routes, déplacer des VMs avec LiveMigration, ...)
- Le **contrôleur** peut gérer des modules **physique** et **virtuel**
- Un **commutateur Ethernet compatible OpenFlow** peut devenir un **routeur**, un **firewall** ou un **proxy** grâce à l'intelligence présente dans le contrôleur
- Essentiel pour les grandes infrastructures (datacenter, cloud, ...) pour offrir **l'élasticité** et la **capacité à montée en charge** avec évent. la **qualité de service**

Lab OpenvSwitch : VLANs gérés par OpenvSwitch

- Configurer et tester ce réseau virtuel avec PC-Win7 + Vbox



Open vSwitch

Contrôle initial (aucune règle)

```
ovs-vsctl show
```

Créer bridge=switch a

```
ovs-vsctl add-br a
```

Connecter le port ethx au bridge

```
ovs-vsctl add-port a ethx
```

Port en mode access → le switch gère 802.1Q

```
ovs-vsctl add-port a ethx vlan_mode=access tag=10
```

Port en mode trunk → le switch transmet le(s) tag(s)

```
ovs-vsctl add-port a ethx vlan_mode=trunk trunks=10
```