Proxmox à UT1
retour d'expérience

UNIVERSITÉ TOULOUSE 1 CAPITOLE

Journée Capitoul du 15 décembre 2016

Philippe ORTH & Virginie GIROU
(Direction du Système d'Information)

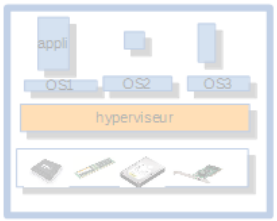
Du modèle physique au IAAS

Application
Data
Runtime
Middleware
OS
Servers
Storage
Network

Ludovic
KVM

Virtualization

Compute :
KVM
HyperV
Vmware ESXi

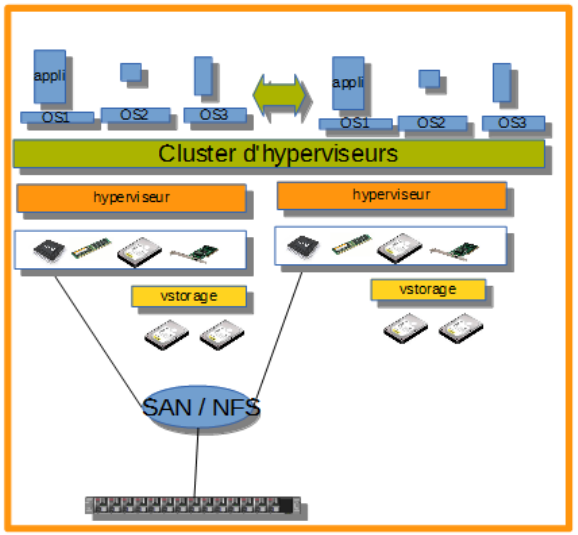


Virtualisation compute

Philippe O.
PROXMOX + CEPH

Clusters

Compute : KVM, ESXi
Storage : SAN, vSAN, CEPH
Network : vswitch (L2)

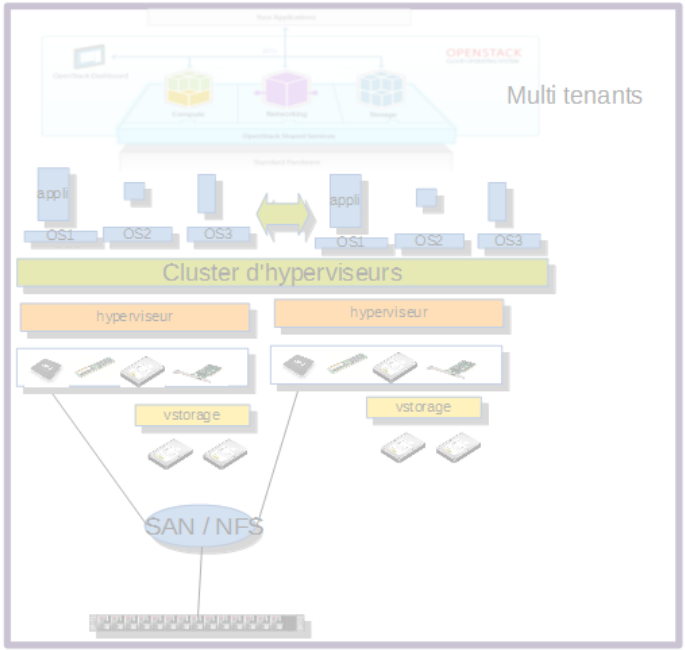


Virtualisation
compute (cluster)
± storage
± network L2
→ contexte LAN

Philippe S.
OPENSTACK + CEPH

IAAS

Compute : PROXMOX, Vmware VSPHERE
Storage : SAN, vSAN, CEPH
Network : Openvswitch (L2 + L3 + service)



Virtualisation
compute (cluster)
± storage
± network L2 + L3
→ contexte LAN, WAN

A. Pourquoi Proxmox ?
B. L’infra de « production » (Proxmox + SAN)
C. L’infra « système » (Proxmox + CEPH)

A - Pourquoi PROXMOX ?

- Un peu d'histoire à UT1:
  - ~2005 : 1ère infra de virtualisation : VMWare ESX (3 nœuds)
  - 2007 : arrivée de XEN
  - 2013 : remplacement de XEN par Proxmox VE à base de KVM
- CF présentation de F. Soulier (capitoul du 16/10/2012)
- Aujourd'hui :
  - VmWare toujours là (3 nœuds) plus d'évolution
  - Proxmox :
    - plus de 160 VM sur 2 infras
    - Quelques projets dans les cartons…

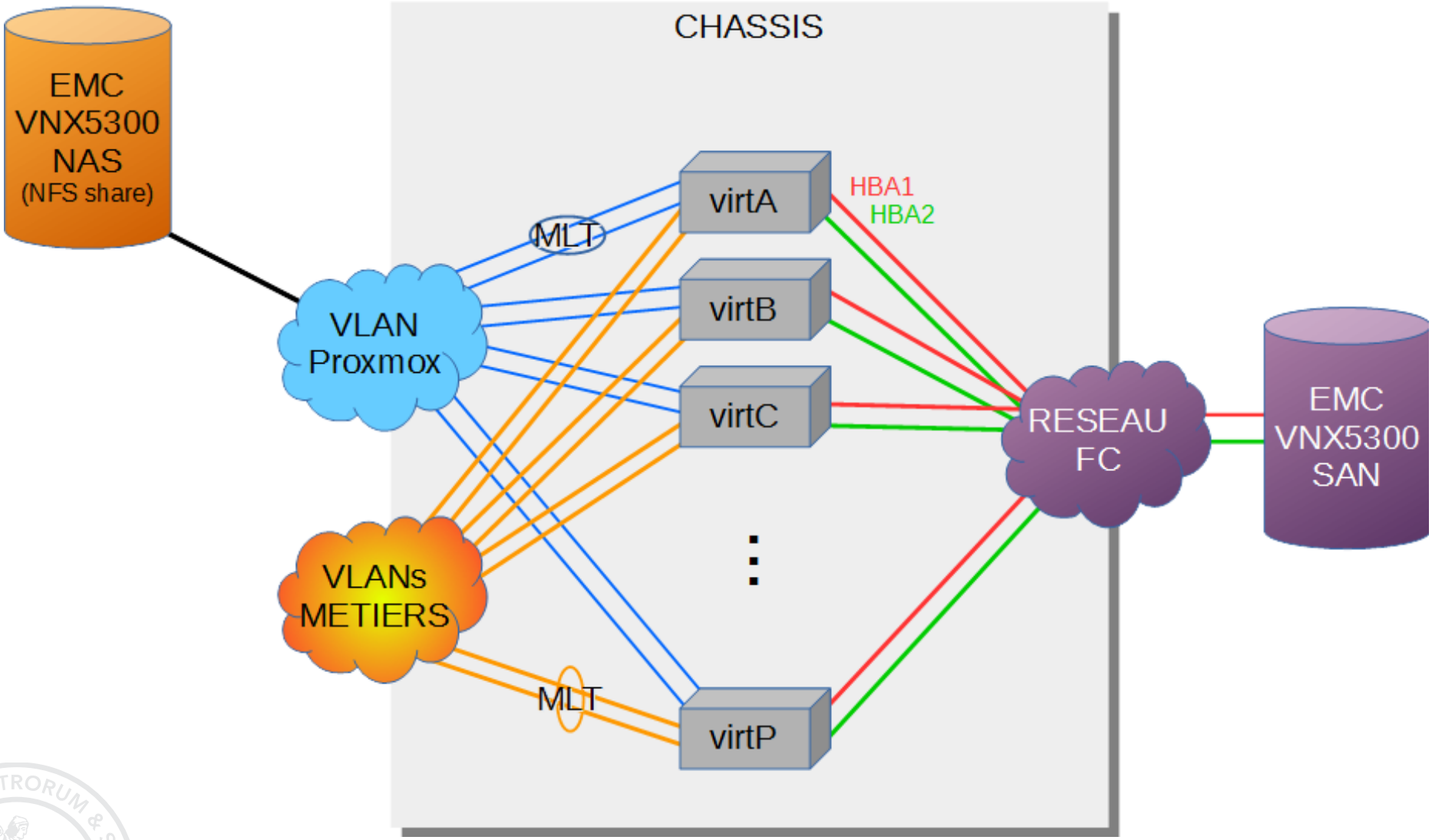
A - Pourquoi PROXMOX ?

- Proxmox c'est :
  - Une solution « clé en main » pour virtualiser (V4.4 13/12/2016)
    - Basé sur Linux debian
    - KVM (émulation matérielle) et LXC (isolation) (remplace OpenVz)
    - Interface d'administration graphique intuitive (pas de java ☺)
    - La console des VM utilise spice ou html5 (toujours pas de java ☺)
  - Un support applicatif réactif (en anglais ou en allemand)
  - Un système de licence simplissime
- Apports par rapport à KVM « seul »
  - Gestion native mode cluster, multi-maître
  - Gestion native des backups de VM
  - Interface graphique : « vmotion » (VM et storage)
  - Intégration du stockage CEPH

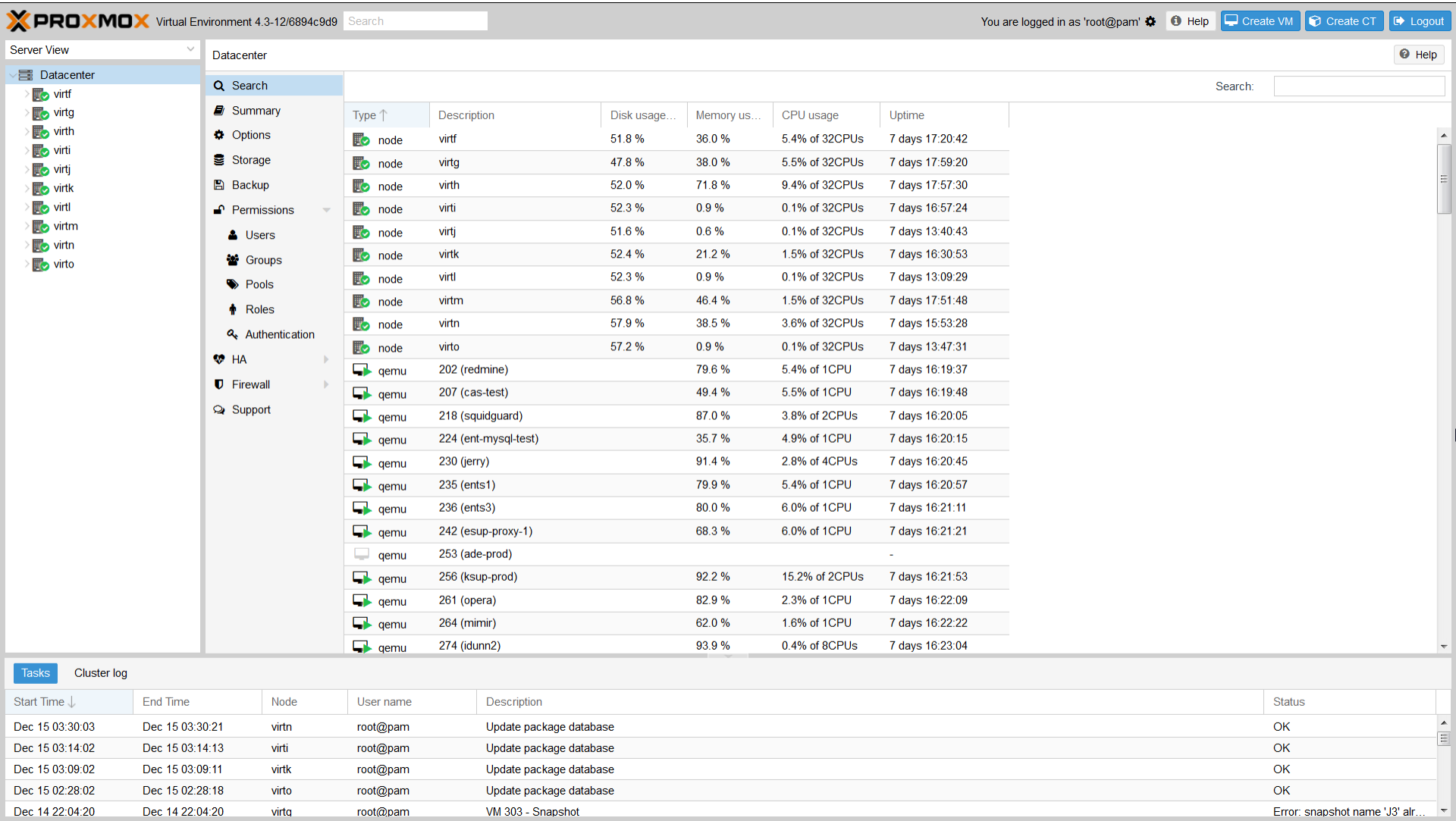
B - Le cluster de production

- Plus de 150 VM serveurs
- Un cluster de 16 nœuds
  - 1 Chassis Blade DELL M1000e
    - 16 lames DELL M630 (32 Cœurs et 256Go Ram par lame)
    - 2 commutateurs internes « 10G/40G ethernet » DELL MXL
    - 2 commutateurs internes « 16G Fiber Channel » BROCADE M6505
    - 32 cartes SD 32Go pour la partition système (!)
  - 16 licences proxmox
- 1 baie EMC VNX 5300 (SAN+NAS)
  - 1 volume en mode bloc (production)
  - 2 volumes NFS
    - Backup
    - Préproduction, tests, formations, etc (QCOW2)

B - Schéma architecture production



B - Interface administration Proxmox

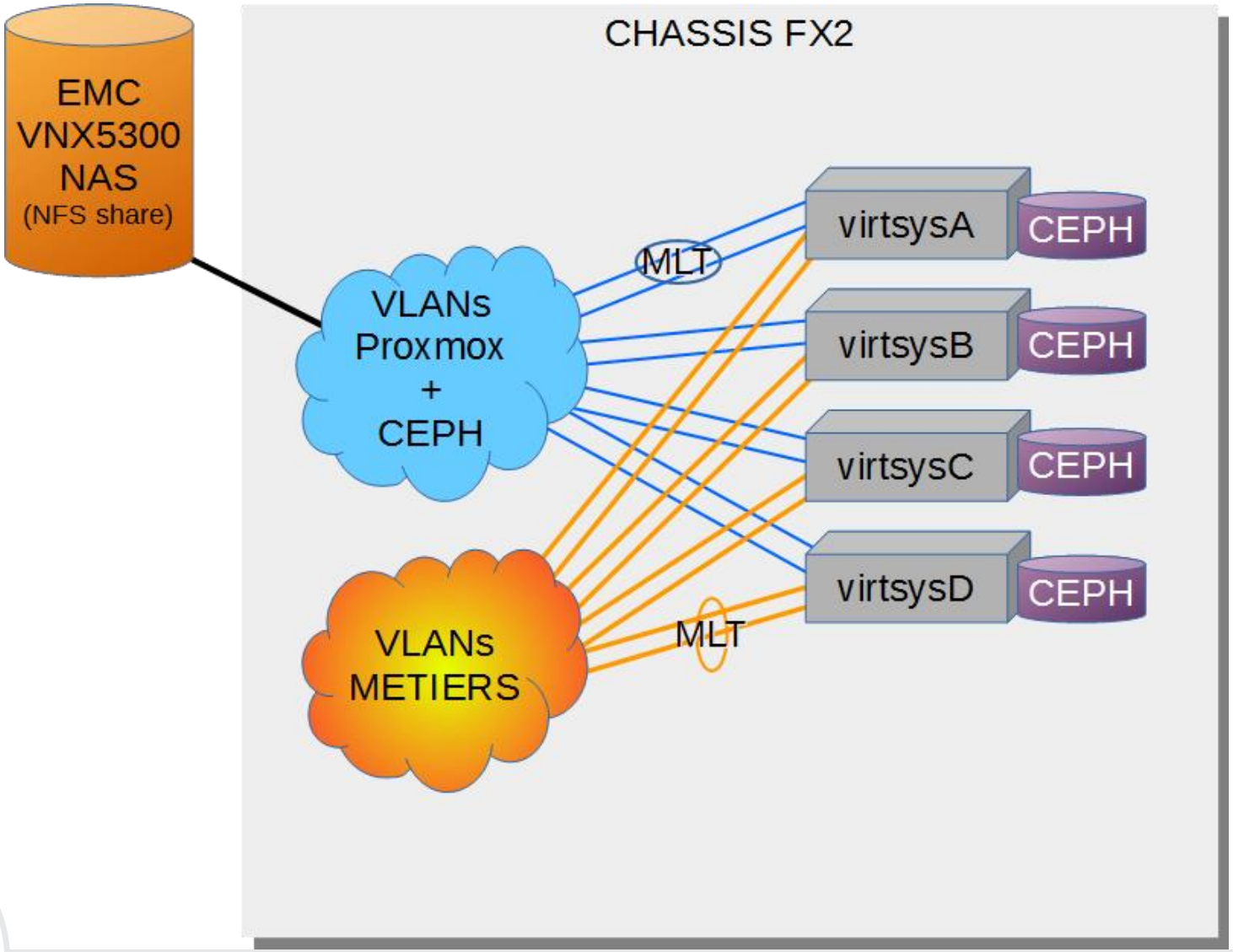


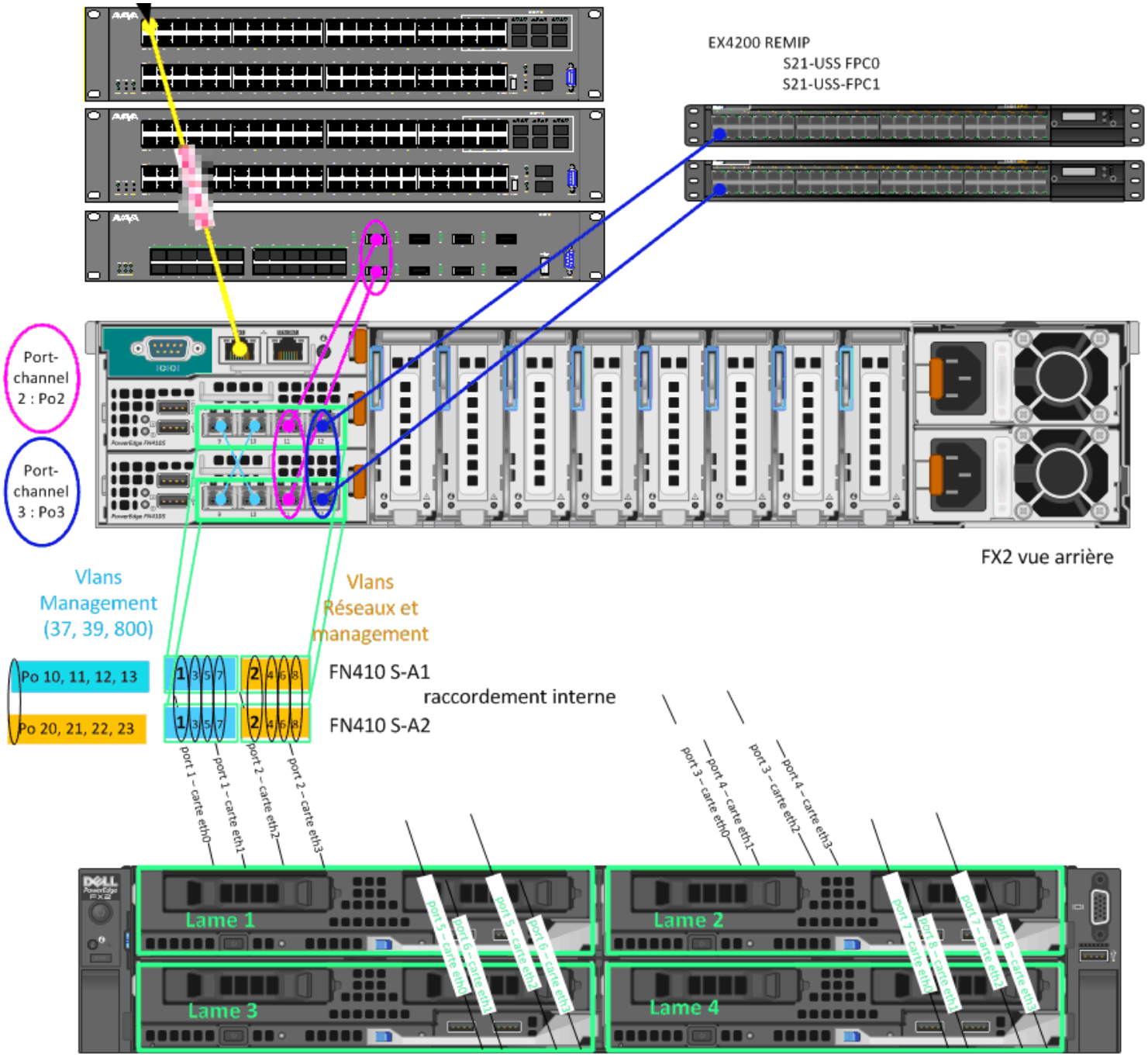
C - L'infra « système »

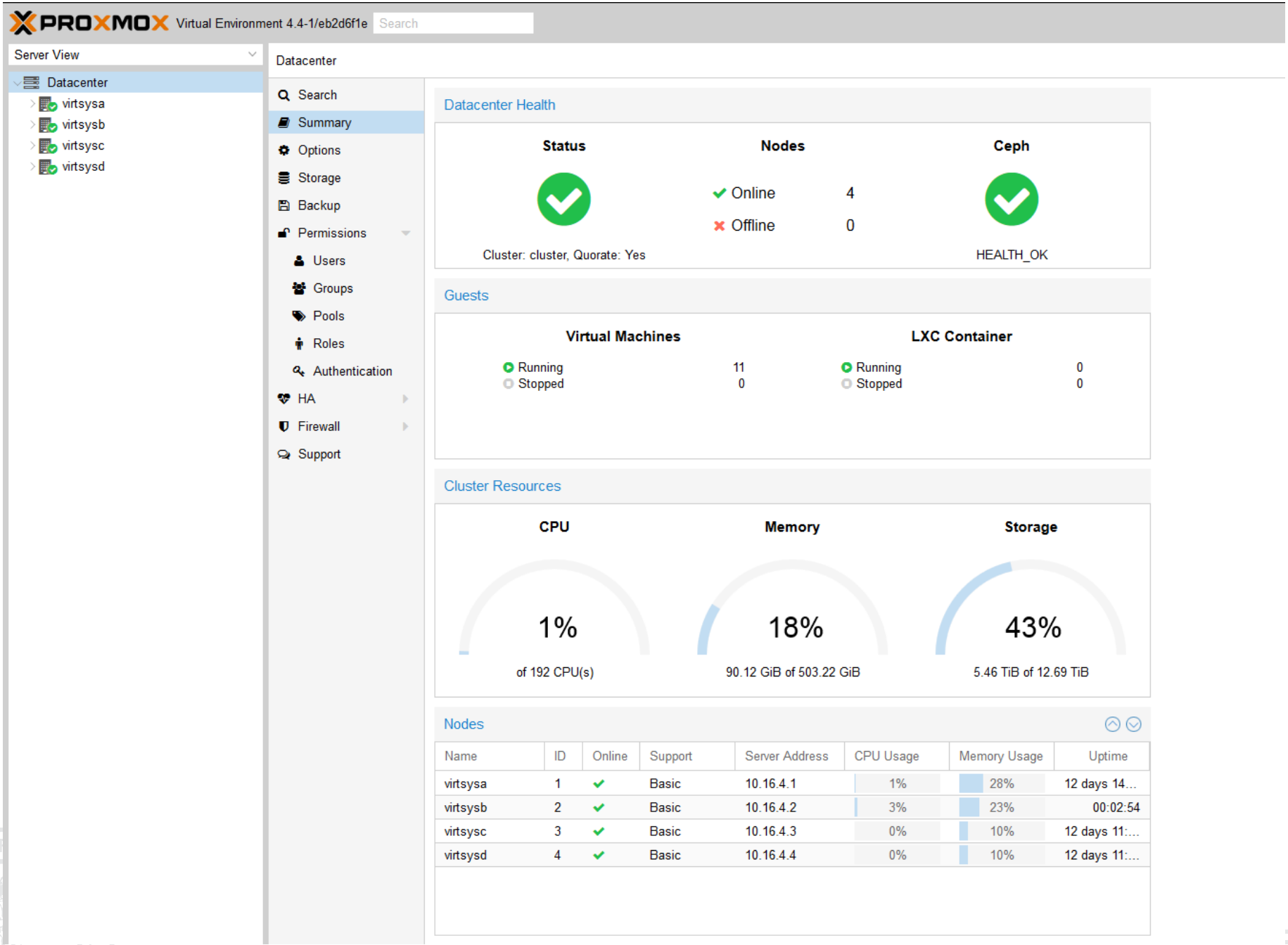
- But
  - Hébergement des VMs primaires (dns, dhcp, parefeu, etc)
  - Etre indépendant de tout autre système (démarrage en 1er)
- Chassis DELL FX2
  - 4 lames FC630 avec pour chacune :
    - CPU : 48 cœurs (Xeon E5-2650 v4)
    - RAM : 128Go
    - Disques internes :
      - 2x 480Go ssd
      - 2x 800Go ssd

  - 2 commutateurs 10G FN410s
    - 8 ports internes
    - 4 ports externes

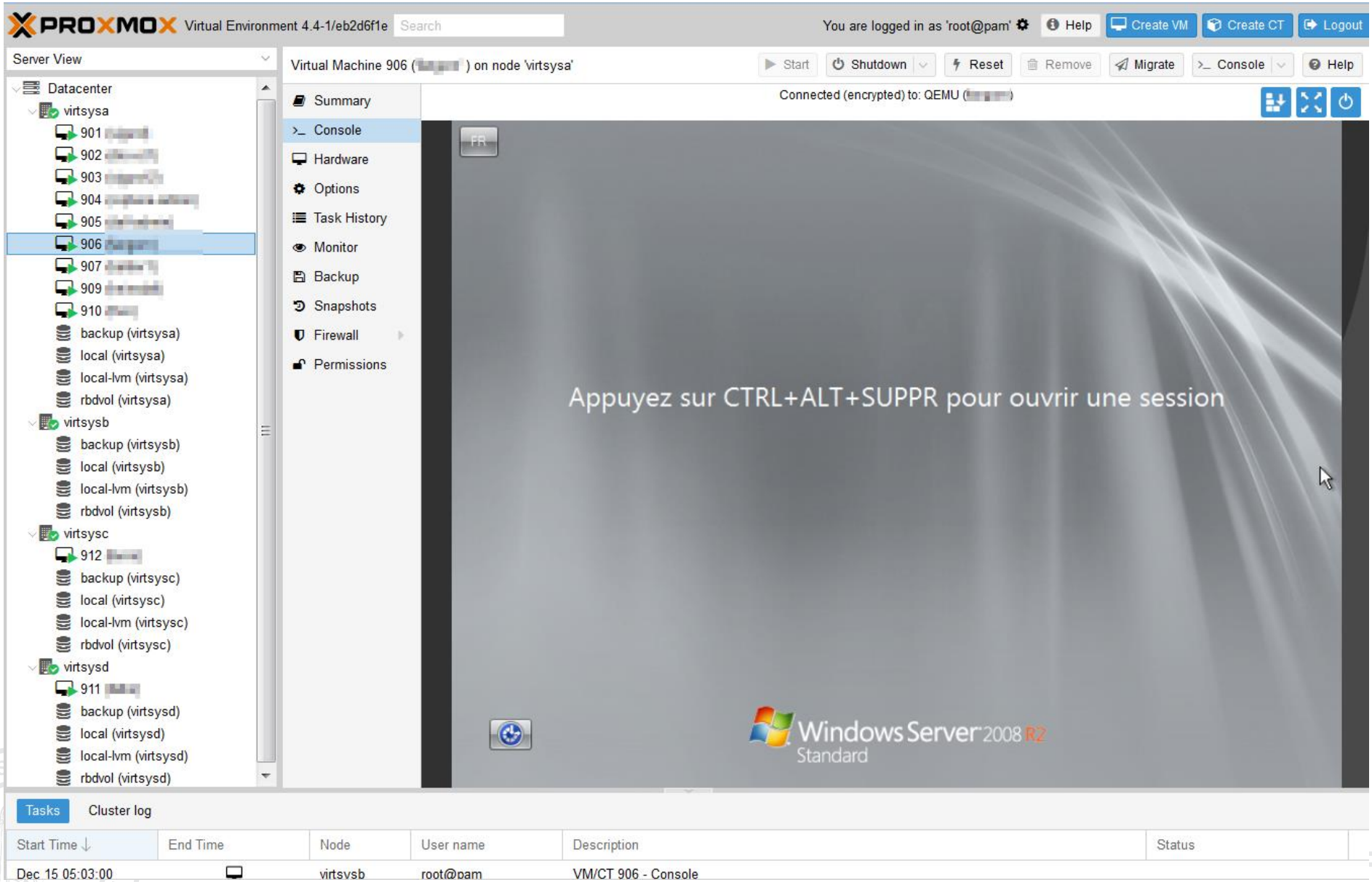
C - L'infra « système » - schéma



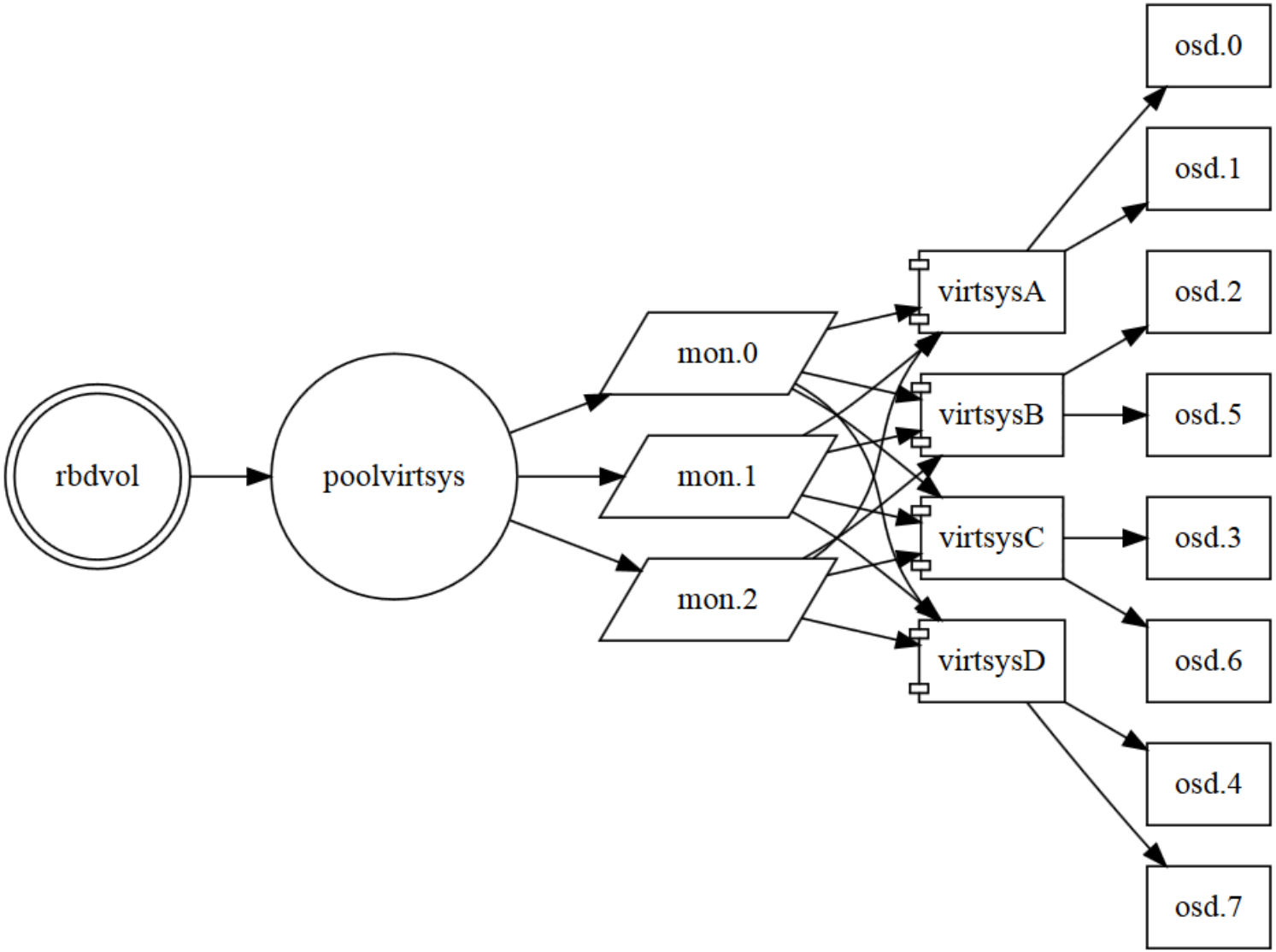

EX4200 REMIP
S21-USS FPC0
S21-USS-FPC1
Port-channel 2 : Po2
Port-channel 3 : Po3
FX2 vue arrière
Vlans Management (37, 39, 800)
Vlans Réseaux et management
Po 10, 11, 12, 13
Po 20, 21, 22, 23
FN410 S-A1
FN410 S-A2
raccordement interne
port 1 – carte eth0
port 1 – carte eth1
port 2 – carte eth2
port 2 – carte eth3
port 3 – carte eth0
port 4 – carte eth1
port 3 – carte eth2
port 4 – carte eth3
Lame 1
Lame 2
Lame 3
Lame 4
FX2 vue avant


PROXMOX Virtual Environment 4.4-1/eb2d6f1e
Search
Server View
Datacenter
virtsysa
virtsysb
virtsysc
virtsysd
Datacenter
Search
Summary
Options
Storage
Backup
Permissions
Users
Groups
Pools
Roles
Authentication
HA
Firewall
Support
Datacenter Health
Status
Nodes
Ceph
Online 4
Offline 0
Cluster: cluster, Quorate: Yes
HEALTH_OK
Guests
Virtual Machines
LXC Container
Running 11
Stopped 0
Running 0
Stopped 0
Cluster Resources
CPU
Memory
Storage
1%
18%
43%
of 192 CPU(s)
90.12 GiB of 503.22 GiB
5.46 TiB of 12.69 TiB
Nodes
Name ID Online Support Server Address CPU Usage Memory Usage Uptime
virtsysa 1 Basic 10.16.4.1 1% 28% 12 days 14...
virtsysb 2 Basic 10.16.4.2 3% 23% 00:02:54
virtsysc 3 Basic 10.16.4.3 0% 10% 12 days 11:...
virtsysd 4 Basic 10.16.4.4 0% 10% 12 days 11:...

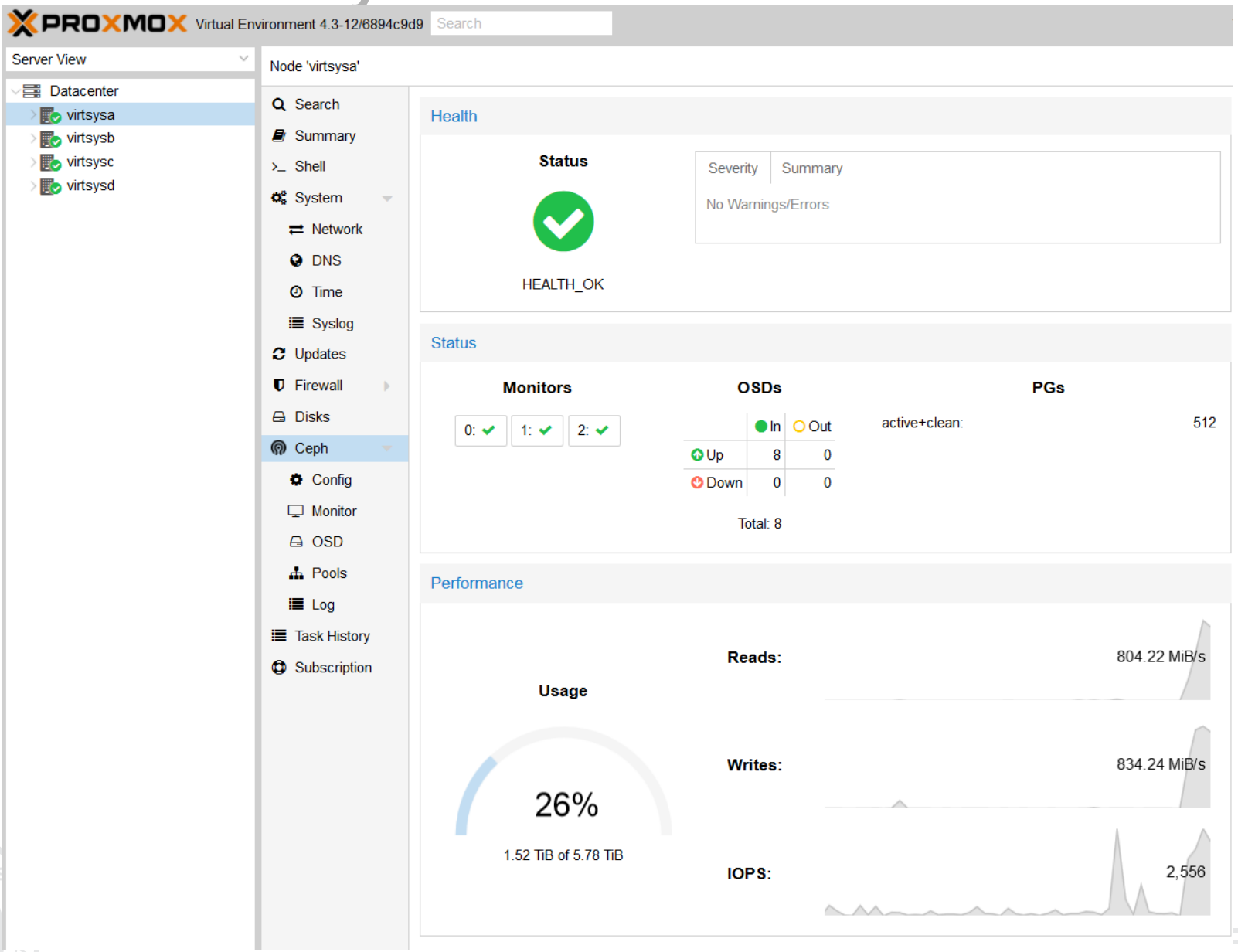
C - L’infra « système »



C - L'infra « système » - CEPH



C - L'infra « système » - tableau de bord CEPH



C - L'infra « système » - tableau de bord CEPH

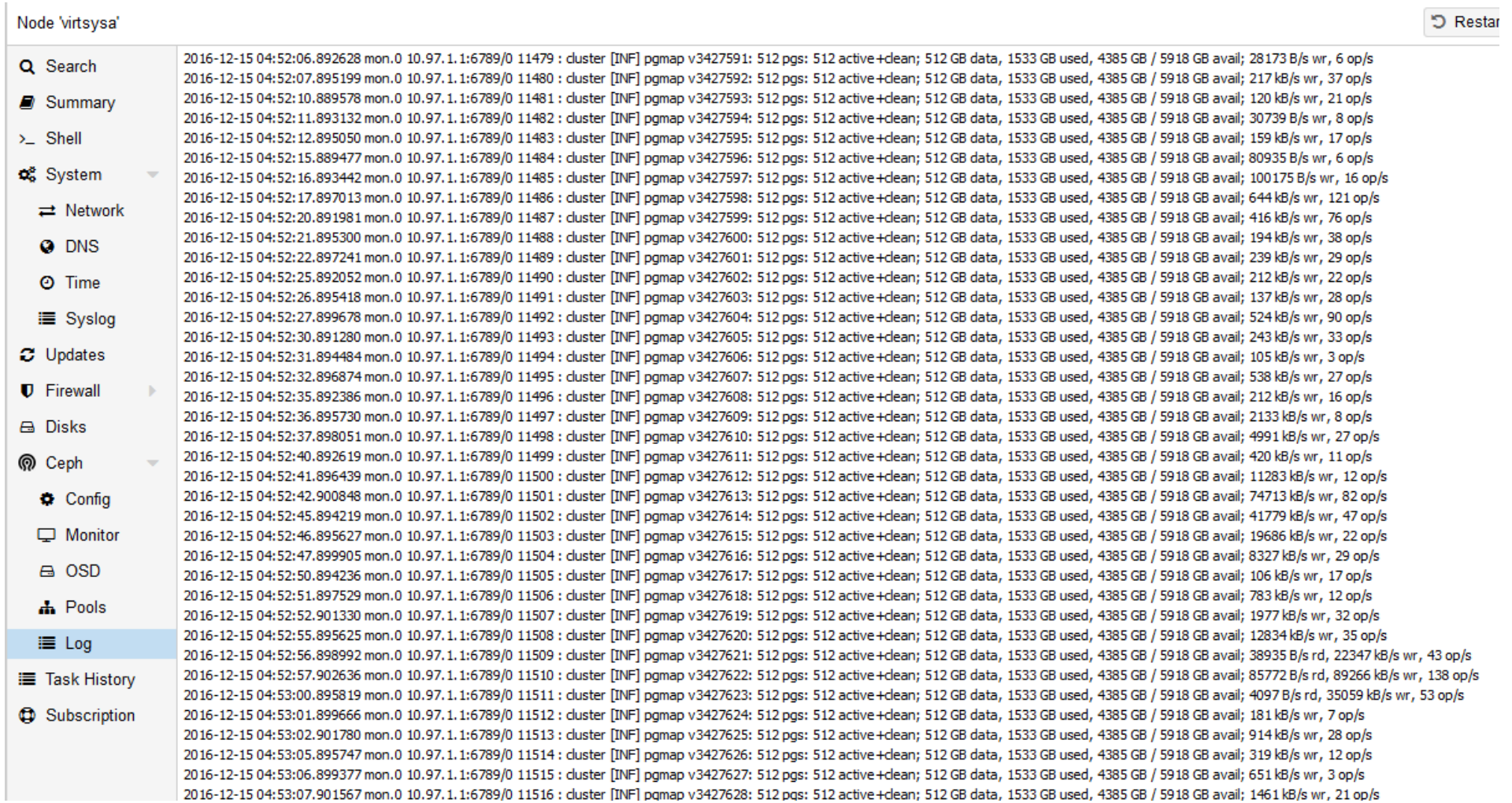
Node 'virtsysa'

- Search
- Summary
- Shell
- System
  - Network
  - DNS
  - Time
  - Syslog
- Updates
- Firewall
- Disks
- Ceph
  - Config
  - Monitor
  - OSD
  - Pools
  - Log
- Task History
- Subscription

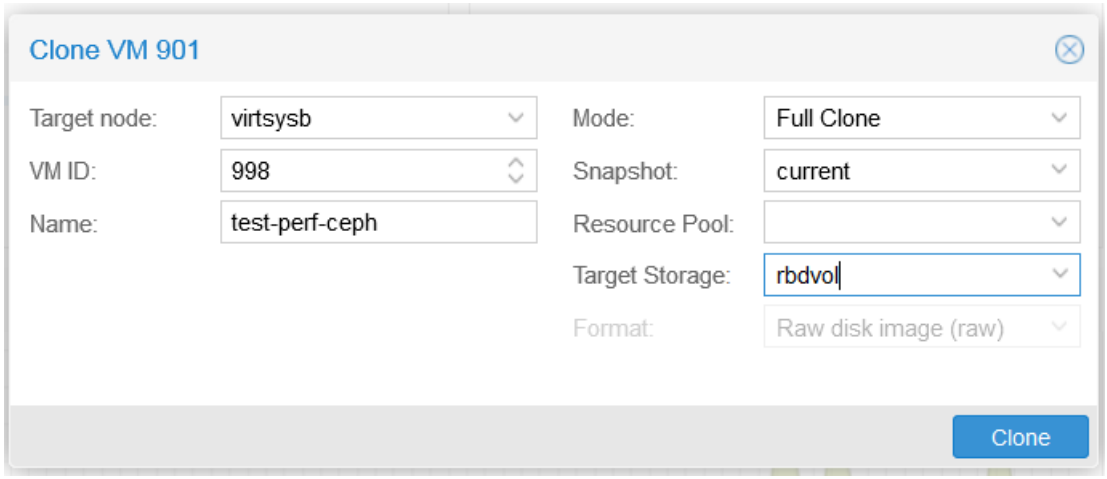
Create: OSD | Reload | Set noout

Name	Type	Status	weight	reweight	Used		Latency (ms)	
					%	Total	Apply	Commit
default	root							
virtsysa	host							
osd.0	osd	up / in	0.719986	1	26.04	739.85 GiB	4	3
osd.1	osd	up / in	0.719986	1	27.16	739.85 GiB	10	8
virtsysb	host							
osd.2	osd	up / in	0.719986	1	26.40	739.85 GiB	6	4
osd.5	osd	up / in	0.719986	1	25.46	739.85 GiB	1	0
virtsysc	host							
osd.3	osd	up / in	0.719986	1	25.39	739.85 GiB	2	1
osd.6	osd	up / in	0.719986	1	24.83	739.85 GiB	7	5
virtsysd	host							
osd.4	osd	up / in	0.719986	1	27.01	739.85 GiB	10	9
osd.7	osd	up / in	0.719986	1	24.97	739.85 GiB	2	1

C - L'infra « système » - tableau de bord CEPH

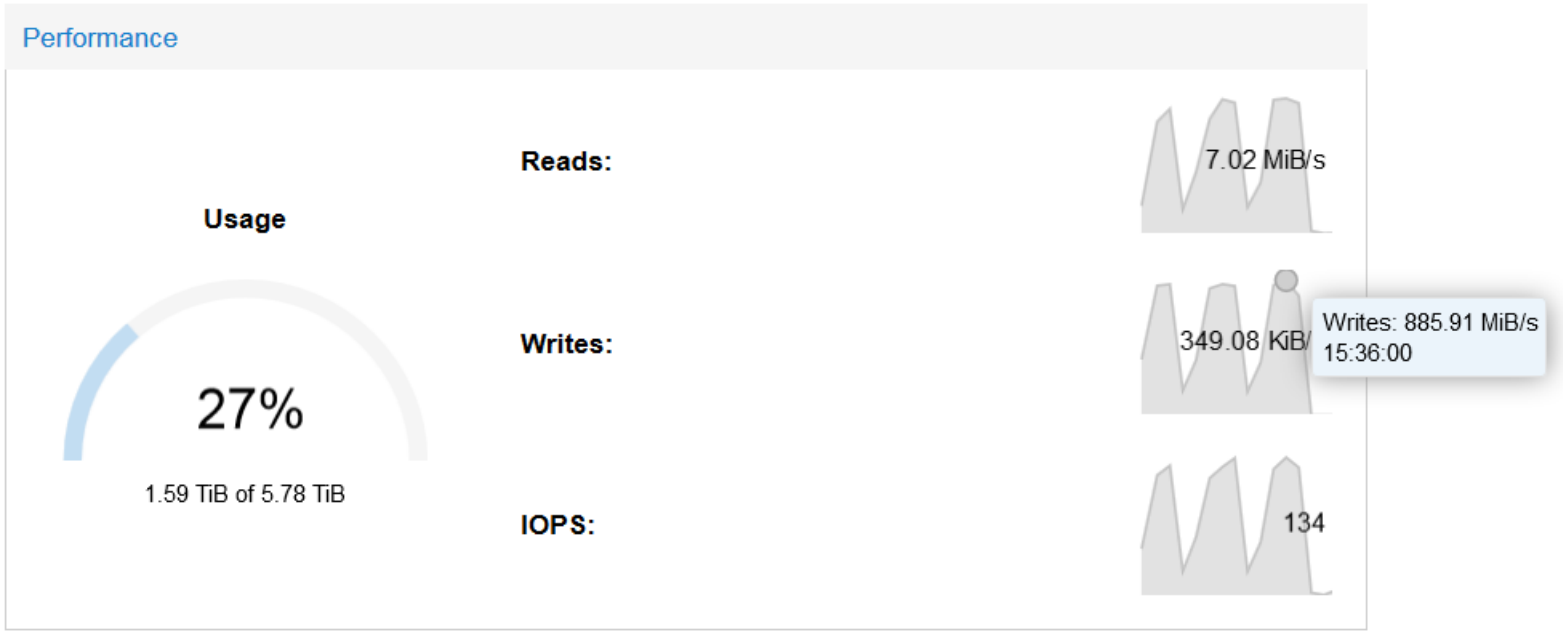


C - L'infra « système » - test Perf CEPH

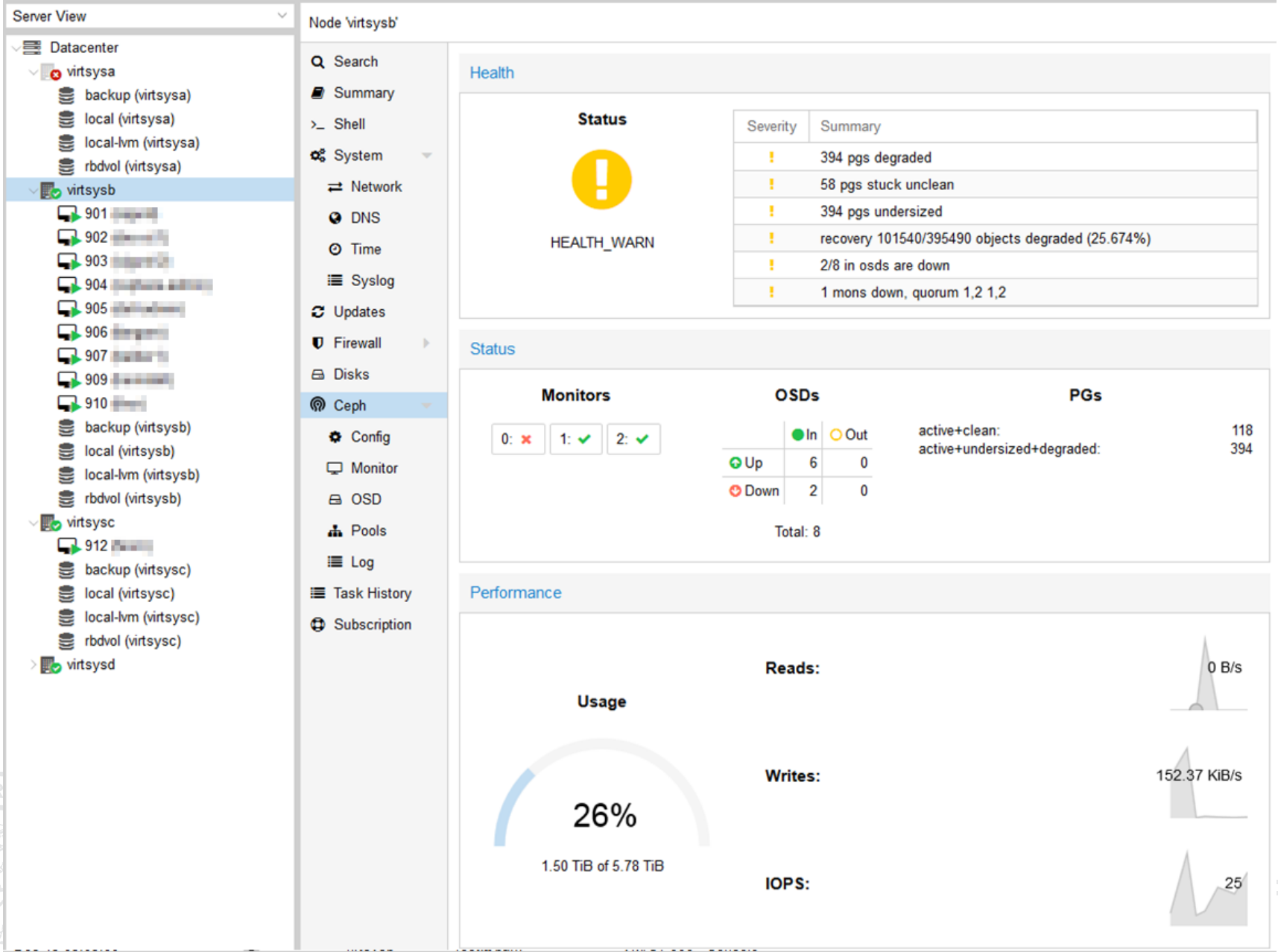


Test de clone d'une VM de 30Go

Durée : 1mn 10s (430Mo/s moyenne)



C – L'Infra « système » - CEPH – reboot



Et ensuite ?

- Aujourd'hui :
  - VmWare toujours là (3 nœuds) plus d'évolution
    - Migration des VMs vers Proxmox et/ou Cloud UFTMIP
  - Proxmox
    - plus de 160 VM sur 2 infras
- Demain : quelques projets dans les cartons…
  - Infra de virtualisation dédiés aux enseignants (UFR info)
  - Mise à dispo de VM pour les étudiants
  - Architecture OPEN STACK locale et/ou répartie