

Introduction à Xen

Mehdi Sebbar – MiNET

www.minet.net
intlab.minet.net

5 Novembre 2010



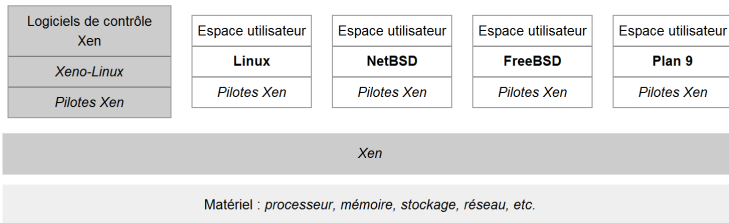
- 1 Principe de Xen
- 2 Principe de LVM
- 3 Installation de Xen
- 4 Création d'une VM
- 5 Configuration de la VM
- 6 Pour aller plus loin, commandes utiles
- 7 bibliographie

Qu'est ce que c'est que XEN ?

Université de Cambridge, Royaume-Uni

- 1 Xen permet d'exécuter plusieurs systèmes d'exploitation (et leurs applications) de manière isolée sur une même machine physique sur plate-forme x86, x86-64. . .
- 2 Les systèmes d'exploitation invités **domU** partagent ainsi les ressources de la machine hôte **dom0**.
- 3 Xen est « hyperviseur » de machines virtuelles. Les systèmes d'exploitation invités ont « conscience » du Xen sous-jacent, ils ont besoin d'être adaptés pour fonctionner sur Xen.
- 4 Xen 3 peut également exécuter des systèmes non modifiés comme Windows sur des processeurs supportant les technologies VT d'Intel ou AMD-V.

Architecture Xen



Différences avec les autres solutions :

QEMU, VirtualBox, VMWare :

- machines virtuelles complètes pour les systèmes d'exploitation invités
- BIOS logiciel (« firmware »).
- Le système d'exploitation invité « croit » tourner sur un matériel
- solutions peu satisfaisantes pour des besoins en performance.
- simples à mettre en œuvre.

VMWare ESX... :

- permet des machines virtuelles complètes pour les systèmes d'exploitation invités, incluant même un BIOS.
- empilage léger, la machine virtuelle se repose sur un noyau léger nommé « vmkernel ».
- similaire à Xen en ce qu'il n'y a pas de système hôte visible, en revanche les systèmes invités n'ont pas à être modifiés, et n'ont pas accès directement au matériel de la machine.

chroot, Linux-VServer, OpenVZ ou BSD Jail :

- isoler certains aspects ou ressources du système d'exploitation hôte comme les systèmes de fichiers ou les espaces mémoire.
- très performantes (peu de surcharge)
- peu ou pas complètement isolés.

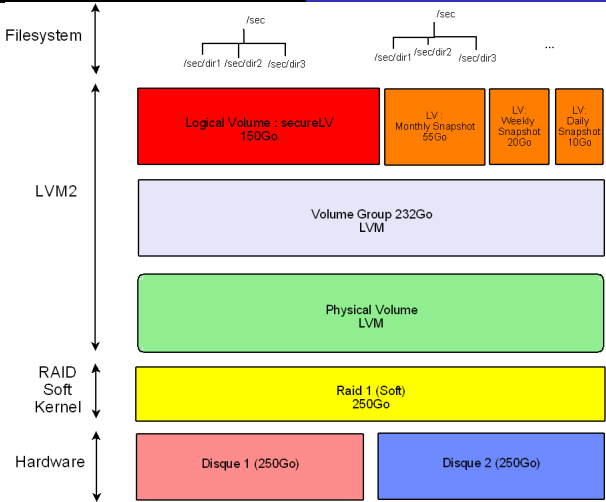


Du fait de cette paravirtualisation et de sa légèreté, Xen est un outil de virtualisation des plus performants. En effet les systèmes invités peuvent avoir un accès direct au matériel, ce qui offre un bien meilleur potentiel d'optimisation.

Qu'est ce que c'est que LVM ?

Logical Volume Manager

- 1 LVM permet la création et la gestion de volumes logiques sous Linux. L'utilisation de volumes logiques remplace en quelque sorte le partitionnement des disques. C'est un système beaucoup plus souple, qui permet par exemple de diminuer la taille d'un système de fichier pour pouvoir en agrandir un autre, sans se préoccuper de leur emplacement sur le disque.
- 2 Il n'y a pas de limitations « étranges » comme avec les partitions (primaire, étendue, etc.). On ne se préoccupe plus de l'emplacement exact des données. On peut conserver quelques giga-octets de libres pour pouvoir les ajouter n'importe où et n'importe quand. Les opérations de redimensionnement deviennent quasiment sans risques, contrairement au redimensionnement des partitions.
- 3 Toute l'administration de LVM se fait en ligne de commande.



Installation :

On prendra un dom0 sur Debian, on le configurera pour utiliser LVM par la suite. Pour cela il faut créer une partition principale / (5G seront suffisants) et un swap. Le reste de l'espace disponible sera assigné pour LVM. Une fois la Debian démarrée, on installe les paquets pour LVM

Installation paquet LVM

```
apt-get install lvm2
```

Puis on crée une partition qui prend le reste de l'espace grâce à fdisk :

```
Command (m for help): p

Disk /dev/sda: 15.8 GB, 15869149184 bytes
255 heads, 63 sectors/track, 1929 cylinders
Units = cylinders of 16065 * 512 = 8225280 bytes
Disk identifier: 0x00071499

   Device Boot      Start         End      Blocks   Id  System
/dev/sda1            1           365     2931831   83   Linux
/dev/sda2            366          487     979965    82   Linux swap / Solaris
/dev/sda3            488         1929    11582865   83   Linux

Command (m for help): _
```


Les disques durs, partitions de disques durs, volumes RAID forment des « volumes physiques » (physical volumes ou PV), on doit les initialiser :

Initialisation des disques et partitions pour LVM :

```
#pvcreate /dev/sda3  
Physical volume "/dev/sda3" successfully created
```

Les commandes utiles :

pvremove, pvdisplay, et tous les pv*...

On concatène ces volumes physiques dans des « groupes de volumes » (volume groups ou VG). Ici on l'appelle lvmvolume, il est préférable de l'appeler xenvg.

Création du volume lvmvolume :

```
# vgcreate lvmvolume /dev/sda3  
Volume group "lvmvolume" successfully created
```

Les commandes utiles :

vgremove, vgdisplay, vgrename, vgscan et tous les vg*...

Des « volumes logiques » (logical volumes ou LV) sont alors découpés dans les groupes de volumes, puis formatés et montés dans des systèmes de fichiers ou utilisés en tant que raw devices.

Création du volume test

```
#lvcreate -name test -size 6G lvmvolume  
Logical volume "test" created
```

Les commandes utiles :

lvremove, lvdisplay, lvrename, lvscan, lvresize, lvextend, lvreduce, et tous les lv*...

Maintenant, on installe Xen :

Installation des paquets Xen :

```
apt-get install xen-hypervisor-3.2-1-amd64 xen-linux-system-2.6.26-2-xen-amd64  
xen-utils-3.2-1 xenstore-utils xenwatch xen-shell xen-tools
```

- xen-hypervisor** : The hypervisor is the "core" for XEN itself. It gets booted by the boot loader and controls cpu and memory, sharing them between your administrative domain (Domain 0) and the virtual guest systems.
- xen-utils** : The userspace tools to manage a system virtualized through the XEN virtual machine monitor.
- xenwatch** : This is a bunch of utilities for virtualization hosts...
- xen-shell** : Using the xen-shell a hosting company could easily allow their customers to fully manage individual Xen guest instances. The shell allows users to start, stop, reboot, or connect to the serial console of a Xen guest instance upon a Xen system.
- xen-tools** : This package contains tools to manage Debian based XEN virtual servers. Using the scripts you can easily create fully configured Xen guest domains (domU) which can be listed, updated, or copied easily.



Après un redemarrage on voit une nouvelle entrée dans le menu grub :

```
GNU GRUB  version 0.97  (639K lower / 261056K upper memory)

Xen 3.2-1-amd64 / Debian GNU/Linux, kernel 2.6.26-2-xen-amd64
Debian GNU/Linux, kernel 2.6.26-2-xen-amd64
Debian GNU/Linux, kernel 2.6.26-2-xen-amd64 (single-user mode)
Debian GNU/Linux, kernel 2.6.26-2-amd64
Debian GNU/Linux, kernel 2.6.26-2-amd64 (single-user mode)
```

`uname -a`

```
# uname -a
Linux pneg 2.6.26-2-xen-amd64 #1 SMP Thu Sep 16 16 :32 :15 UTC 2010 x86_64
GNU/Linux
```

Configuration de XEN :

Deux fichiers : `/etc/xen/xend-config.sxp` et `/etc/xen-tools/xen-tools.conf`.

On modifie le fichier `/etc/xen/xend-config.sxp` pour activer le mode bridge (le dom0 agit comme un hub, switch?)

`/etc/xen/xend-config.sxp (pneg)`

```
(xend-relocation-server yes)
(xend-relocation-port 8002)
(xend-relocation-address '')
(xend-relocation-hosts-allow '^192\\.168\\.1\\. [0-9]+$')
(network-script network-bridge)
(vif-script vif-bridge)
(dom0-min-mem 196)
(dom0-cpus 0)
```

Les `xend-relocation*` servent pour les migrations de vm.

On modifie maintenant `/etc/xen-tools/xen-tools.conf`

/etc/xen-tools/xen-tools.conf (pneq)

```
lvm = lvmvolume
install-method = debootstrap
size = 6Gb
memory = 256Mb
swap = 256Mb
fs = ext3
dist = lenny
image = sparse
gateway = 157.159.40.1
netmask = 255.255.255.128
broadcast = 157.159.40.127
bridge = eth0
passwd = 1
kernel = /boot/vmlinuz-`uname -r`
initrd = /boot/initrd.img-`uname -r`
arch = amd64
mirror = http://ftp.fr.debian.org/debian/
ext3_options = noatime,nodiratime,errors=remount-ro
ext2_options = noatime,nodiratime,errors=remount-ro
xfs_options = defaults
```

Suite /etc/xen-tools/xen-tools.conf (pneg)

```
reiserfs_options = defaults  
btrfs_options    = defaults  
serial_device   = hvc0  
disk_device     = xvda
```


Ce fichier servira de base pour xen-tools. Maintenant on crée une vm :

création d'un domU

```
#xen-create-image --hostname coursxen --role udev --ip 192.168.1.118
```

General Information

```
Hostname       : coursxen
Distribution    : lenny
Partitions     : swap           256Mb (swap)
                /               6Gb   (ext3)
Image type     : full
Memory size    : 256Mb
Kernel path    : /boot/vmlinuz-2.6.26-2-xen-amd64
Initrd path    : /boot/initrd.img-2.6.26-2-xen-amd64
```

Networking Information

```
IP Address 1   : 192.168.1.118 [MAC: 00:16:3E:B8:0F:0D]
Netmask       : 255.255.255.128
Broadcast     : 157.159.40.127
Gateway       : 157.159.40.1
```

Suite de la création du domU coursxen

```
Creating swap on /dev/lvmvolume/coursxen-swap
Done
Creating ext3 filesystem on /dev/lvmvolume/coursxen-disk
Done
Installation method: debootstrap
Done
Running hooks
Done
No role scripts were specified.  Skipping
Creating Xen configuration file
Done
Setting up root password
Enter new UNIX password:
Retype new UNIX password:
passwd: password updated successfully
All done
Logfile produced at:
/var/log/xen-tools/coursxen.log
```

Puis maintenant on peut lancer le domU :

Création du domU coursxen :

```
# xm create /etc/xen/coursxen.cfg  
Using config file "/etc/xen/coursxen.cfg".  
Started domain coursxen
```

```
# xm list
```

```
# xm list
```

Name	ID	Mem	VCPUs	State	Time(s)
Domain-0	0	466	2	r-----	933.9
coursxen	7	256	1	-b----	2.4
emi	1	256	1	-b----	419.0
newbies	3	256	1	-b----	82.6
raknorvm	4	256	1	r-----	284852.8
zabbix	5	512	1	-b----	65.3

Configuration du domU :

xm console 7

```
# xm console 7  
  
Debian GNU/Linux 5.0 coursxen hvc0  
  
coursxen login:
```

Première chose à faire, configurer le réseau du vlan 100, on modifie le fichier /etc/network/interfaces.

/etc/network/interfaces

```
auto eth0  
iface eth0 inet static  
    address 192.168.1.118  
    gateway 192.168.1.1  
    netmask 255.255.255.0  
    broadcast 192.168.1.255
```

On redémarre `/etc/init.d/networking` :

Redémarrage du service networking :

```
coursxen : # /etc/init.d/networking restart  
Reconfiguring network interfaces...done.
```

On vérifie enfin avec ping que l'on communique avec le dom0.

Pour avoir une ip dans le vlan2, il faut récupérer le paquet vlan sur le dom0 :

Installation paquet vlan sur le domU

```
root@pneg:~# scp vlan_1.9-3_amd64.deb 192.168.1.118:~/  
[...]  
vlan_1.9-3_amd64.deb      100%   39KB   39.2KB/s   00:00
```

Puis on l'installe sur le domU avec `dpkg`.

On lance le module `8021q` :

Chargement de 8021q

```
coursxen : # modprobe 8021q
```

On oublie pas d'ajouter `8021q` dans `/etc/modules` pour qu'il soit chargé au démarrage du système.

Ensuite on ajoute la nouvelle interface dans `/etc/network/interfaces` :

`/etc/network/interfaces` pour vlan2

```
auto lo
iface lo inet loopback

auto eth0.2
iface eth0.2 inet static
    address 157.159.40.121
    gateway 157.159.40.1
    netmask 255.255.255.128

auto eth0
iface eth0 inet static
    address 192.168.1.118
    netmask 255.255.255.0

dns-domain minet.net
dns-nameservers 192.168.1.55 192.168.1.54
```

On redemarre le service `/etc/init.d/networking` et on vérifie la config à l'aide de `ifconfig` :



ifconfig

```
coursxen:~# ifconfig
eth0  Link encap:Ethernet  HWaddr 00:16:3e:00:be:d9
      inet addr:192.168.1.118  Bcast:192.168.1.255  Mask:255.255.255.
      inet6 addr: fe80::216:3eff:fe00:bed9/64 Scope:Link
[...]
```

```
eth0.2 Link encap:Ethernet  HWaddr 00:16:3e:00:be:d9
      inet addr:157.159.40.121  Bcast:157.159.40.127  Mask:255.255.255.128
      inet6 addr: fe80::216:3eff:fe00:bed9/64 Scope:Link
[...]
```

```
lo    Link encap:Local Loopback
      inet addr:127.0.0.1  Mask:255.0.0.0
      inet6 addr: ::1/128 Scope:Host
[...]
```

On vérifie ensuite la configuration à l'aide de ping.

Ne pas oublier de modifier /etc/resolv.conf si l'on décide de ne garder que l'interface dans le vlan2.



Pour aller plus loin, commandes utiles :

Démarrage automatique du domU :

```
mkdir /etc/xen/auto  
ln -s /etc/xen/mon-domU /etc/xen/auto/
```

Les commandes xm :

console	Attach to <Domain>'s console.
create	Create a domain based on <ConfigFile>.
destroy	Terminate a domain immediately.
help	Display help.
list	List information about all/some domains.
mem-set	Set the current memory usage for a domain.
migrate	Migrate a domain to another machine.
pause	Pause execution of a domain.
reboot	Reboot a domain.
restore	Restore a domain from a saved state.
resume	Resume a Xend managed domain
save	Save a domain state to restore later.
shell	Launch an interactive shell.
shutdown	Shutdown a domain.
start	Start a Xend managed domain
suspend	Suspend a Xend managed domain

top	Monitor a host and the domains in real time.
unpause	Unpause a paused domain.
uptime	Print uptime for all/some domains.
vcpu-set	Set the number of active VCPUs for allowed for the domain.

Problèmes :

Impossible de se connecter en ssh :

```
mehdi@coursxen's password:  
stdin: is not a tty
```

Solution : Installer udev sur le domU.

Le problème du *Time went backwards* :

```
[528025.354359] clocksource/0 : Time went backwards : ret=104c7d2d8df64  
delta=-1586800815081502 shadow=104c7a82db02b offset=2aab8d90
```

Problème d'horloge.

/etc/sysctl.conf sur le dom0

```
[...]  
xen.independent_wallclock=1
```

/etc/xen/coursxen.conf sur le dom0

```
[...]  
extra="clocksource=jiffies"
```

Puis pour appliquer sans redemarrer le domU

On applique

```
# sysctl -p  
# echo "jiffies" >  
/sys/devices/system/clocksource/clocksource0/current_clocksource
```

3 solutions sur <http://wiki.debian.org/Xen>.

Pour aller plus loin :

Ganeti : <http://code.google.com/p/ganeti/>

Ganeti is a cluster virtual server management software tool built on top of existing virtualization technologies such as Xen or KVM and other Open Source software.

Live migration support, Plain LVM volumes...

Chef : Chef is a systems integration framework, built to bring the benefits of configuration management to your entire infrastructure.

With Chef, you can :

- Manage your servers by writing code, not by running commands. (via Cookbooks)
- Integrate tightly with your applications, databases, LDAP directories, and more. (via Libraries)
- Easily configure applications that require knowledge about your entire infrastructure ("What systems are running my application?" "What is the current master database server?") (via Search)

Bibliographie :

<http://fr.wikipedia.org/wiki/Xen>

http://fr.wikipedia.org/wiki/Gestion_par_volumes_logiques

<http://wiki.debian.org/Xen>

<http://www.pouf.org/uploads/img/lmg-diverse/LVM-RAID-2.png>

<http://www.howtoforge.com/virtualization-with-xen-on-debian-lenny-amd64>

<http://docs.ganeti.org/ganeti/current/html/>