

Virtualisation avec Xen

Stéphane Bortzmeyer

<stephane+blog@bortzmeyer.org>

Première rédaction de cet article le 1 Février 2006. Dernière mise à jour le 14 Janvier 2007

<http://www.bortzmeyer.org/xen.html>

La virtualisation, qui consiste à créer des machines virtuelles (typiquement plusieurs sur une machine physique), est très à la mode. Plusieurs exposés à JRES 2005 (<http://www.bortzmeyer.org/jres-2005.html>) en ont parlé. Il existe plusieurs façons différentes de virtualiser et nous allons commencer par Xen.

Selon le cahier des charges qu'on a, on choisira une ou l'autre solution de virtualisation. Il existe :

- Les émulateurs de processeur, comme Qemu (<http://www.bortzmeyer.org/qemu.html>) ou Bochs (<http://bochs.sourceforge.net/>), qui permettent de donner une illusion complète d'une machine et donc de faire tourner des systèmes d'exploitation non modifiés (mais à un coût élevé, question performances, puisqu'il faut traverser l'émulateur à chaque instruction machine),
- Les machines compartimentées, comme les "*jails*" (<http://docs.freebsd.org/44doc/papers/jail/jail.html>) de FreeBSD ou comme les Vservers (<http://linux-vserver.org/>) de Linux, qui sont limités à un seul noyau, ne permettent pas, par exemple, de modifier la table de routage mais qui suffisent à, par exemple, un hébergeur qui veut laisser ses clients être root, sans pour autant leur confier une machine physique entière,
- Les hyperviseurs comme l'historique VM d'IBM ou comme Xen, qui est décrit ici. L'hyperviseur n'émule pas un processeur, le système d'exploitation doit être porté sur Xen mais ils offrent les meilleures performances, et permettent de faire tourner plusieurs systèmes d'exploitation sur la même machine physique.

Notons que l'apparition récente de processeurs offrant des fonction de virtualisation comme le Pacifica d'AMD brouille un peu les frontières. Ces technologies permettent par exemple de faire tourner sur Xen un système d'exploitation non modifié.

Un système Xen se compose d'un système d'exploitation **hôte** ("*host*"), nommé dom0 et de plusieurs systèmes d'exploitation **invités** ("*guests*"), les domU. L'hôte, aujourd'hui, avec Xen version 3, doit être un Linux ou, depuis peu, un NetBSD. Les invités que j'ai testé sont Linux, NetBSD ou FreeBSD.

Lorsque l'hôte (dom0) consomme 100 % du CPU, les invités (domU) sont complètement à genoux. Il est donc conseillé d'avoir un dom0 minimum, très sécurisé, et qui ne fait rien d'autre que de lancer les domU.

L'hôte, dom0 nécessite un noyau Xen. Actuellement, le noyau Linux standard n'a pas de support de Xen et il faut récupérer les patches dans la distribution de Xen (<http://www.cl.cam.ac.uk/Research/SRG/netos/xen/downloads/>) avant de compiler, ou bien utiliser un noyau déjà compilé par les gens de Xen. Attention : le noyau utilisé dépend étroitement de l'hyperviseur qui tourne, même de la version 3.0.0 à 3.0.2 de Xen, les changements obligent à changer tous les composants ensemble.

Le logiciel de lancement, grub dans mon cas, doit lancer l'hyperviseur qui lancera ensuite le noyau xenisé :

```
title          XEN with Debian GNU/Linux, kernel 2.6-custom-brassens-xen0
root          (hd0,0)
kernel       /boot/xen-3.0-i386.gz dom0_mem=32M
module       /boot/vmlinuz-2.6.16-xen-brassens root=/dev/sda1 ro console=tty0
savedefault
boot
```

`/boot/xen-3.0-i386.gz` est l'hyperviseur, disponible dans le `xen-hypervisor-3.0-i386`. `/boot/vmlinuz-2.6.16-xen-brassens` est le noyau, que j'ai compilé.

Si on utilise Debian comme système hôte, on dispose désormais dans la version "etch" de Debian d'un jeu complet de paquetages pour Xen.

Pour chaque système invité, il faut créer au moins un disque dur virtuel. Cela peut être une partition ou bien un fichier (créé avec `dd`, par exemple `ddif=/dev/zero of=/xen-machines/images/peggy-root bs=1024k count=1000`, puis `mkfs`; ce disque virtuel peut être monté, par exemple pour examiner et modifier les fichiers, avec `mount-text2-oloop/xen-machines/images/polly-root/mnt`). Ce disque dur virtuel doit être rempli avec les fichiers du système souhaité. Si c'est le même que le système hôte, ou bien si une machine du système souhaité est disponible sur le réseau, une copie avec `cp` ou `scp` suffit. Si le système invité est un Debian ou un Ubuntu, la commande `debootstrap` est très pratique (par exemple `debootstrap --arch=i386 breezy/mnt http://archive.ubuntu.com/ubuntu`). Sinon, il faut trouver une image du système convoité, à copier sur le disque dur virtuel. (Pour NetBSD, ce n'est pas nécessaire, l'installeur peut tourner sur Xen.) Deux excellents sites, <http://jailtime.org/> et <http://www.xen-get.org/> en fournissent pour beaucoup de systèmes.

On doit ensuite créer un fichier de configuration par machine virtuelle. Un des miens dit :

```
kernel = "/boot/vmlinuz-2.6-xenU"
memory = 64
name = "polly"
vif = ['']
disk = ['file:/xen-machines/images/polly-root,sda1,w', 'file:/xen-machines/images/polly-swap,sda2,w']
root = "/dev/sda1 ro"
```

Il spécifie le noyau à lancer (`/boot/vmlinuz-2.6-xenU`), la quantité de mémoire allouée, les disques virtuels (ici, deux, de type fichier) et la racine du futur système. On peut aussi passer des arguments au noyau du système invité, par exemple `extra="gentoo=nodevfs"` pour arriver à démarrer une Gentoo.

Si on fait tourner Xen sur une machine ayant peu de RAM, il peut être utile d'abaisser la quantité de mémoire dédiée à Xen lui-même, ce qui se fait dans le fichier `xend-config.sxp`, par exemple :

<http://www.bortzmeyer.org/xen.html>

```
# Setting it too low has strange effects (freezes of dom0 or domU)
(dom0-min-mem 16)
```

Ensuite, il ne reste plus qu'à lancer la machine virtuelle, avec `xmcreatefichier-configuration-c`. On a alors la joie de voir les messages de démarrage du noyau s'afficher dans le terminal et, lorsque le démarrage est fini, on a une nouvelle machine à sa disposition.

Xen permet d'afficher la liste des systèmes qui tournent :

```
aslan:~ % sudo xm list
Name                               ID Mem(MiB) VCPUs State  Time(s)
Domain-0                            0    46      1 r----- 268.7
diggory                              2    64      1 -b----- 214.3
polly                                6    64      1 ----- 785.6
```

Ici, `dom0` a 46 mégas de mémoire, les invités, `polly` et `diggory` en ont 64 chacun. Avec Xen, les partitions mémoire sont statiques (il faut arrêter une machine virtuelle pour la changer ; si une machine virtuelle ne fait rien, elle bloque quand même la quantité de mémoire allouée). Il existe aussi un `xmtop` pour suivre en temps réel l'activité des machines virtuelles.

Xen crée aussi une carte réseau virtuelle, donc les machines virtuelles ont un accès complet à Internet :

```
polly:~ % ifconfig eth0
eth0      Link encap:Ethernet  HWaddr 00:16:3E:05:19:FF
          inet addr:172.19.1.32  Bcast:172.19.255.255  Mask:255.255.255.0
```

Le préfixe de l'adresse Ethernet, `00:16:3E` est celui alloué à Xen mais on peut le changer.

Des manipulations plus intéressantes peuvent être tentées avec le réseau (configurer plusieurs ponts, router, etc) mais je n'ai pas encore eu beaucoup de succès sur ce terrain.

Lorsqu'on veut arrêter la machine virtuelle, c'est un `xmshutdown` et on a la surprise de voir ladite machine virtuelle afficher "*Power down*" à la fin.

Ce texte a été écrit sur un système Ubuntu invité pendant qu'un autre système invité, un Gentoo compilait X.org (<http://www.x.org>), le système hôte étant une Debian. J'ai aussi démarré un FreeBSD sur la même machine physique pendant ce temps.

Pour en savoir plus, on peut consulter :

- Sur l'installation d'Ubuntu, <http://wiki.xensource.com/xenwiki/UbuntuHoaryHowTo> ou bien http://www.howtoforge.com/perfect_xen_setup_debian_ubuntu. **Attention à ne pas supprimer tous les getty, il faut laisser le premier.**
- Pour FreeBSD, <http://www.fsmware.com/xenofreebsd/> (attention, c'est assez rude).
- Pour NetBSD, voir <http://wiki.xensource.com/xenwiki/NetBSDdomU> (attention, Xen 2 et Xen 3 sont très différents et, si NetBSD tourne sur Xen 3, c'est depuis peu : j'ai testé la version 3.1 de NetBSD sans problème. NetBSD semble le seul système où Xen est disponible en standard).

À noter qu'il existe aussi des hébergeurs qui vous louent une domU Xen pour pas cher comme SliceHost (<http://www.bortzmeyer.org/slicehost-debut.html>).