

Toolstack

M1 RÉSEAUX & TÉLÉCOMS – RT0702

OLIVIER FLAUZAC



Libvirt - Généralités

Libvirt

Ensemble d'outils de gestion

- de machines virtuelles
- d'actions sur des machines virtuelles

types d'outils

- gestion du stockage
- gestion du réseau
- gestion de l'installation
- clonage
- sauvegarde

Libvirt : terminologie

Nœud

- Machine physique (hôte)

Hyperviseur

- Couche logicielle de virtualisation
- Propriétés génériques
- Propriétés spécifiques
- S'exécute sur un nœud

Domaine

- Instance d'un système d'exploitation
- Exécution d'un invité géré par l'hyperviseur

Libvirt : Hyperviseurs

XEN

Qemu /KVM

LXC

OpenVz

VirtualBox

VMware ESX - Workstation - Player

HyperV

IBM PowerVM

Virtuozzo

Bhyve

Libvirt : Objectifs

Assurer la gestion de domaines

Proposer une couche logicielle stable

- API
- binding dans de nombreux langages

Proposer une couche logicielle sécurisée

Utilisable à distance

- connexion socket
- connexion SSH

Libvirt : Opérations

Opérations sur les domaines

- Création
- Gestion du cycle de vie
- Modification
- Contrôle
- Sauvegarde
- Clonage
- Provisionning

Libvirt : Principe

Opérations sur les domaines

Connexion à un hyperviseur : URI

- hyperviseur
- adresse

Exécution de commandes

- en shell spécifique
- commande shell paramétrée
 - fichier XML

en API

- C, C++, C#, Java, Perl, PHP, Python, Ruby

Libvirt : Le réseau

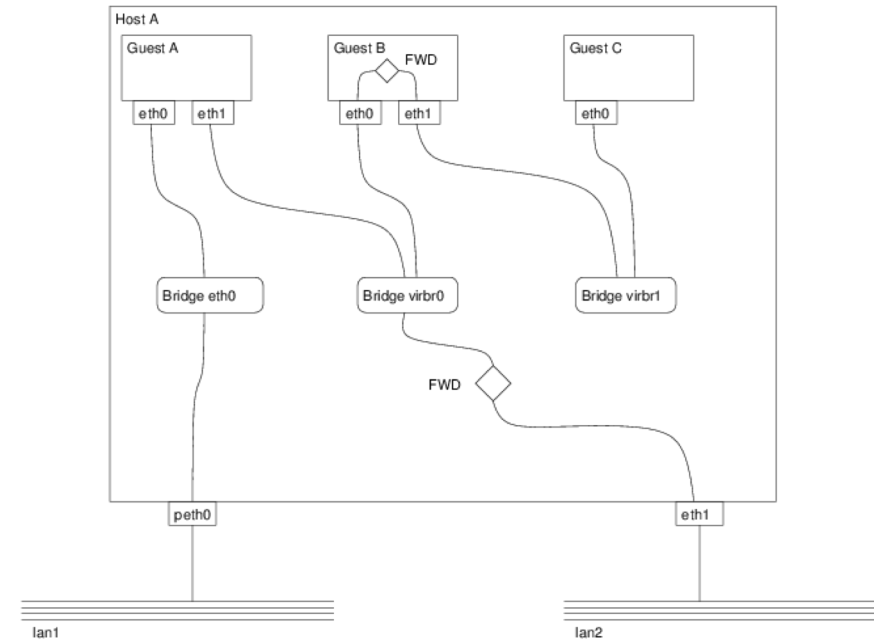
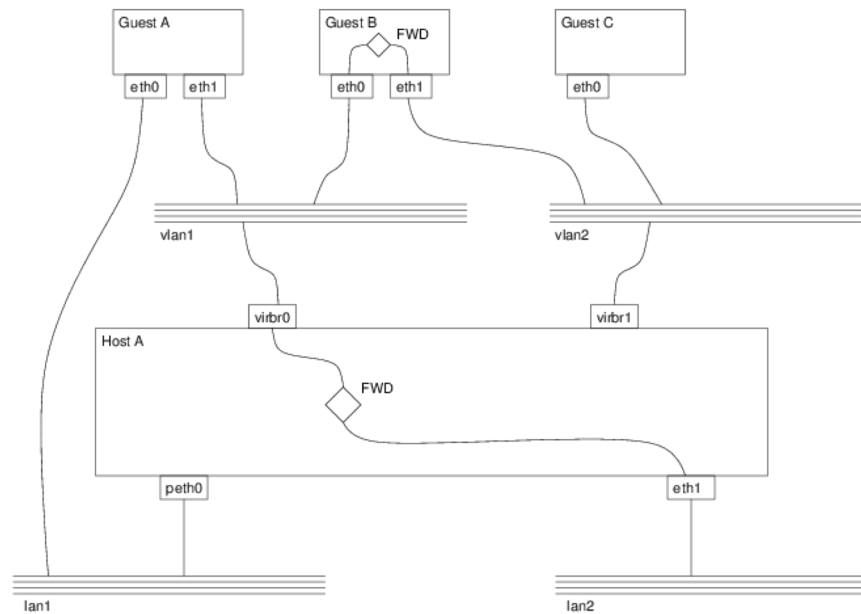
Connexion en NAT

Mise en place de VLAN

Gestion des Bridges

Gestion des réseaux privés

Configurations logiques et physiques



Libvirt : ligne de commande

La commande `virsh`

Commande de gestion de domaine

Shell de gestion de domaine

Utilisation en ligne de commande

- passage de tous les paramètres
- exploitation de fichier de définition XML

URI de connexion

Gestionnaire d'invité

Adresse de connexion

Type de connexion

- `session` : accès local utilisateur
- `system` : accès local root
- `tcp` : accès socket
- `ssh` : accès ssh

Exemples

- `test:///default`
- `qemu:///system`
- `qemu+ssh://root@192.168.1.40/system`
- `openvz+ssh:///system`
- `lxc:///`

Gestion de domaines

`define` enregistrement de domaine

`undefine` désenregistrement de domaine

`create` création à partir d'un fichier XML

`list` liste des domaines

`dominfo` information sur les domaines

`state` état d'un domaine

`dumpxml` exportation XML

Gestion de l'état des domaines

`start` démarrage de l'invité

`shutdown` arrêt propre

`destroy` arrêt brutal

`reboot` redémarrage

`save` sauvegarde d'un domaine

`restore` restauration d'un domaine

`suspend` suspension d'un domaine

`resume` réactivation d'un domaine

Gestion de l'état des ressources

`setmem` modification de la mémoire

`setvcpus` modification du nombre de CPU

`vcpuinfo` récupération de l'état CPU

Importation d'un invité dans un domaine

Solution de migration la plus simple

Obligatoire pour des invités non filesystem

Problème de gestion des propriétés / compatibilité

- entrées / sorties
- format des drivers ..

Exploitation de scripts

Unification des appels / paramètres

Gestion d'une chaîne de bout en bout

Pas toujours possible

Détachement des impératifs des hyperviseurs

Scripts

- `virt-install`
- `virt-clone`
- `virt-convert`

Script de création / paramétrage

`virt-install`

Script basé sur les commandes `virsh`

Création d'une image

Paramétrage de l'image

Gestion des driver

Gestion des domaines

- création
- enregistrement

Script de création / paramétrage

`virt-install`

De très nombreuses options !!!

Assure la compatibilité de la configuration

- depuis la création
- jusqu'à l'utilisation
- cas des drivers (entrées / sorties, disque ...)

Gestion complète pour Qemu / KVM, XEN

Gestion Partielle pour OpenVz, LXC

Propriétés générales

`connect` connexion à l'hyperviseur

`name` nom du domaine

`arch` architecture processeur si différente de l'hôte

`os-type` type système d'exploitation

`os-variant` distribution (si disponible)

Paramètres de l'instance

`memory` Mémoire allouée au domaine

`name` nom du domaine

`vcpus` nombre de processeurs alloués

`accelerate` utilisation de KVM

Entrées / sorties et disques

`cdrom` CD physique ou image ISO

`location` source de l'installation

- répertoire, disque partagé, HTTP, FTP

`disk` Gestion des disques

- `path`
- `device` (disk,file)
- `bus` (ide, scsi, usb, virtio, xen)
- `size`

Réseau

`network` connexion réseau

- `bridge` connexion à un bridge
- `network` connexion à un réseau virtuel
- `user` mode SLIRP
- `default` création juste d'une carte, ajouté à un bridge contenant une carte physique, sinon création du réseau default

`mac` si pas précisé, allocation aléatoire

Accès et console

graphics

- keymap
- vnc, spice
- port
- listen
- password

noautoconsole

video

- cirrus
- vga
- qxl
- vmvga

Commande

```
virt-install
--connect qemu:///system
--os-type linux -n debianOli2
--memory=512
--vcpus=1
--accelerate
--cdrom='/var/lib/libvirt/images/debian-8.2.0-amd64-CD-1.iso'
--disk path=/var/lib/libvirt/images/debianOli2.qcow2,device=disk,bus=virtio,format=qcow2,size=10
--network=user,model=virtio
--noautoconsole
--graphics vnc,listen=0.0.0.0,keymap=fr
```

```
root@salon:~# virsh list --all
```

ID	Nom	État
-	debian0li	fermé
-	debian0li2	fermé

```
root@salon:~# virsh dominfo debian0li2
```

```
ID : -
Nom : debian0li2
UUID : c685bfbf-8f44-4d36-86bf-4ab21fe00efe
Type de SE : hvm
État : fermé
CPU : 1
Mémoire Max : 524288 KiB
Mémoire utilisée : 0 KiB
Persistent: yes
Démarrage automatique : disable
Managed save: no
Security model: apparmor
Security DOI: 0
```

```
root@salon:~# virsh domstate debian0li2
```

```
fermé
```

```
root@salon:~# virsh start debianOli2
```

```
Domaine debianOli2 démarré
```

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

ID	Nom	État

24	debianOli2	en cours d'exécution

```
root@salon:~# virsh domstate debianOli2
```

```
en cours d'exécution
```

```
root@salon:~# virsh shutdown debianOli2
```

```
Le domaine debianOli2 est en cours d'arrêt
```

```
root@salon:~# virt-clone --original=debianOli2 --name=debianOli3 \
                        --file=/var/lib/libvirt/images/debianOli3.qcow2
Allocating 'debianOli3.qcow2' | 10 GB 00:39
```

Clone 'debianOli3' created successfully.

```
root@salon:~# virsh list --all
```

ID	Nom	État

-	debianOli	fermé
-	debianOli2	fermé
-	debianOli3	fermé

```
root@salon:~# virsh start debianOli2
```

Domaine debianOli2 démarré

```
root@salon:~# virsh start debianOli3
```

Domaine debianOli3 démarré

Retour sur le clonage

Permet la mise en place de banques d'images

Duplication du disque

- logiciels / service / configuration
- configuration réseau ?
- configuration hostname

Changement d'adresse MAC

libvirt : format XML

Définition d'un format

Description de tous les éléments d'un domaine

- généraux
- matériel
- configuration réseau
- ...

Permet l'instanciation / configuration des machines

Création d'un fichier XML

- à la main
 - fastidieux ou incomplet
- depuis un dumpxml d'un domaine existant
 - trop complet


```
<domain type='kvm' id='25'>
  <name>debian0li2</name>
  <uuid>c685bfbf-8f44-4d36-86bf-4ab21fe00efe</uuid>
  <memory unit='KiB'>524288</memory>
  <currentMemory unit='KiB'>524288</currentMemory>
  <vcpu placement='static'>1</vcpu>
  <resource>
    <partition>/machine</partition>
  </resource>
  <os>
    <type arch='x86_64' machine='pc-i440fx-utopic'>hvm</type>
    <boot dev='hd' />
  </os>
  <features>
    <acpi/>
    <apic/>
    <pae/>
  </features>
  <cpu mode='custom' match='exact'>
    <model fallback='allow'>Nehalem</model>
  </cpu>
  <clock offset='utc'>
    <timer name='rtc' tickpolicy='catchup' />
    <timer name='pit' tickpolicy='delay' />
    <timer name='hpet' present='no' />
  </clock>
  <on_poweroff>destroy</on_poweroff>
  <on_reboot>restart</on_reboot>
  <on_crash>restart</on_crash>
```

Définition des disques

```
<devices>
  <emulator>/usr/bin/kvm-spice</emulator>
  <disk type='file' device='disk'>
    <driver name='qemu' type='qcow2' />
    <source file='/var/lib/libvirt/images/debian01i2.qcow2' />
    <backingStore/>
    <target dev='vda' bus='virtio' />
    <alias name='virtio-disk0' />
    <address type='pci' domain='0x0000' bus='0x00' slot='0x04' function='0x0' />
  </disk>
  <disk type='block' device='cdrom'>
    <driver name='qemu' type='raw' />
    <backingStore/>
    <target dev='hda' bus='ide' />
    <readonly/>
    <alias name='ide0-0-0' />
    <address type='drive' controller='0' bus='0' target='0' unit='0' />
  </disk>
```

```
<controller type='usb' index='0' model='ich9-ehci1'>
  <alias name='usb0' />
  <address type='pci' domain='0x0000' bus='0x00' slot='0x05' function='0x7' />
</controller>
<controller type='usb' index='0' model='ich9-uhci1'>
  <alias name='usb0' />
  <master startport='0' />
  <address type='pci' domain='0x0000' bus='0x00' slot='0x05' function='0x0'
multifunction='on' />
</controller>
<controller type='usb' index='0' model='ich9-uhci2'>
  <alias name='usb0' />
  <master startport='2' />
  <address type='pci' domain='0x0000' bus='0x00' slot='0x05' function='0x1' />
</controller>
<controller type='usb' index='0' model='ich9-uhci3'>
  <alias name='usb0' />
  <master startport='4' />
  <address type='pci' domain='0x0000' bus='0x00' slot='0x05' function='0x2' />
</controller>
```

Contrôleurs

```
<controller type='pci' index='0' model='pci-root'>
  <alias name='pci.0' />
</controller>
<controller type='ide' index='0'>
  <alias name='ide0' />
  <address type='pci' domain='0x0000' bus='0x00' slot='0x01' function='0x1' />
</controller>
```

```
<interface type='user'>
  <mac address='52:54:00:50:21:dc' />
  <model type='virtio' />
  <alias name='net0' />
  <address type='pci' domain='0x0000' bus='0x00' slot='0x03' function='0x0' />
</interface>
<serial type='pty'>
  <source path='/dev/pts/1' />
  <target port='0' />
  <alias name='serial0' />
</serial>
<console type='pty' tty='/dev/pts/1'>
  <source path='/dev/pts/1' />
  <target type='serial' port='0' />
  <alias name='serial0' />
</console>
<input type='mouse' bus='ps2' />
<input type='keyboard' bus='ps2' />
<graphics type='vnc' port='5900' autoport='yes' listen='0.0.0.0' keymap='fr'>
  <listen type='address' address='0.0.0.0' />
</graphics>
<video>
  <model type='cirrus' vram='16384' heads='1' />
  <alias name='video0' />
  <address type='pci' domain='0x0000' bus='0x00' slot='0x02' function='0x0' />
</video>
```

```
<memballoon model='virtio'>
  <alias name='balloon0' />
  <address type='pci' domain='0x0000' bus='0x00' slot='0x06' function='0x0' />
</memballoon>
</devices>
<seclabel type='dynamic' model='apparmor' relabel='yes'>
  <label>libvirt-c685bfbf-8f44-4d36-86bf-4ab21fe00efe</label>
  <imagelabel>libvirt-c685bfbf-8f44-4d36-86bf-4ab21fe00efe</imagelabel>
</seclabel>
</domain>
```

Libvirt et le réseau

Le mode carte

```
<interface type='user'>  
  <mac address='52:54:00:50:21:dc' />  
  <model type='virtio' />  
  <alias name='net0' />  
  <address type='pci' domain='0x0000' bus='0x00' slot='0x03' function='0x0' />  
</interface>
```


Mode réseau virtuel

Définition des propriétés

Définition des éléments d'infrastructure (DHCP ...)

Modes proposés

- NAT
- Routed Network : gestion du réseau par le serveur libvirt
- private
- bridge

Mode NAT

```
<network>
  <name>default</name>
  <bridge name='virbr0' />
  <forward mode='nat' />
  <ip address='192.168.122.1' netmask='255.255.255.0'>
    <dhcp>
      <range start='192.168.122.2' end='192.168.122.254' />
    </dhcp>
  </ip>
</network>
```

Routed network

```
<network>
  <name>local</name>
  <bridge name='virbr1' />
  <forward mode='route' dev='eth1' />
  <ip address='192.168.122.1' netmask='255.255.255.0'>
    <dhcp>
      <range start='192.168.122.2' end='192.168.122.254' />
    </dhcp>
  </ip>
  <ip family='ipv6' address='2001:db8:ca2:2::1' prefix='64' />
</network>
```

Réseau isolé

```
<network>
  <name>private</name>
  <bridge name='virbr2' />
  <ip address='192.168.152.1' netmask='255.255.255.0'>
    <dhcp>
      <range start='192.168.152.2' end='192.168.152.254' />
    </dhcp>
  </ip>
  <ip family='ipv6' address='2001:db8:ca2:3::1' prefix='64' />
</network>
```

Bridge hôte

```
<network>  
  <name>host-bridge</name>  
  <forward mode='bridge' />  
  <bridge name='br0' />  
</network>
```

Libvirt et LXC

Problématique des conteneurs

Problématique

- Pas de création de filesystem
- Extension d'un système existant
- Script libvirt pas toujours utilisables

Mode opératoire

- Création / clonage avec les outils des conteneurs
- Intégration du domaine
- Cycle de vie avec libvirt

```
<domain type='lxc'>
  <name>olilxc</name>
  <memory>327680</memory>
  <os>
    <type>exe</type>
    <init>/sbin/init</init>
  </os>
  <vcpu>1</vcpu>
  <clock offset='utc' />
  <on_poweroff>destroy</on_poweroff>
  <on_reboot>restart</on_reboot>
  <on_crash>destroy</on_crash>
  <devices>
    <emulator>/usr/lib/libvirt/libvirt_lxc</emulator>
    <filesystem type='mount'>
      <source dir='/var/lib/lxc/olilxc/rootfs' />
      <target dir='/' />
    </filesystem>
    <interface type='network'>
      <source network='default' />
    </interface>
    <console type='pty' />
  </devices>
</domain>
```



```
root@salon:~# lxc-create -t debian -n olilxc
```

```
root@salon:~# virsh -c lxc:/// define olilxc.xml
```

Domaine olilxc défini depuis olilxc.xml

```
root@salon:~# virsh net-list --all
```

Nom	État	Démarrage automatique	Persistent

default	actif	yes	yes

```
root@salon:~# virsh -c lxc:/// start olilxc
```

Domaine olilxc démarré