

Ecole informatique IN2P3 2014 : Maîtriser le Cloud
TP Dev : portage d'applications sur le Cloud
TP4 : Utilisation d'une base de données

Cécile Cavet

`cecile.cavet at apc.univ-paris7.fr`

Centre François Arago (FACe),

Laboratoire AstroParticule et Cosmologie (APC), LabEx UnivEarthS

APC, Univ. Paris Diderot, CNRS/IN2P3,

CEA/lrfu, Obs. de Paris, Sorbonne Paris Cité, France

4 Juillet 2014

Table des matières

0.1	Introduction	1
0.2	Configuration du disque persistant et montage des disques	2
0.2.1	Création d'un disque persistant	2
0.2.2	Lancement de la MV sur StratusLab	2
0.2.3	Lancement de la MV sur OpenStack	3
0.3	Utilisation de la base de données	5
0.3.1	Copie des données	5
0.3.2	Interface Web	5

0.1 Introduction

Vous allez créer dans ce TP un disque persistant afin de copier des données présentes sur un disque « read only ». Pour cela, vous devrez passer par une MV car le contenu des disques ne peut pas être accédé sans la MV. Le disque « read only » (DRO) est une image ISO qui est référencée sur les catalogues d'image de la même manière que les images disques. Les données du DRO ne sont accessibles qu'en lecture seule et il est donc nécessaire pour certaines applications de copier les données afin de pouvoir les manipuler. Dans ce TP, le DRO contient des données et une base de données associée. Le disque persistant (DP) permet de garder de manière pérenne des données même si la MV venait à défaillir. Dans ce TP, le DP doit être créée et formaté. Vous copierez ensuite les données du DRO et lancerez un utilitaire, Pipelet, afin de visualiser l'architecture des données via une interface Web.

Note : les disques persistants, comme les instantanées d'image disque, sont liés à une infrastructure de Cloud et donc non interopérable. L'utilisation d'un disque « read only » est un bon moyen d'utiliser un même jeu de données ou de transférer des données d'un Cloud vers un autre.

0.2 Configuration du disque persistant et montage des disques

Vous devez d'abord créer le DP sur les deux infrastructures. Vous utilisez ensuite une MV à laquelle vous connectez le DRO. Vous connectez dans un troisième temps le DP afin d'avoir les deux disques à disposition sur la MV. Comme dans les TP précédents, vous utiliserez une image disque customisée qui contient des modules Python et l'utilitaire Pipelet.

0.2.1 Création d'un disque persistant

Créez un DP de taille 5 GB à l'aide de la commande `stratus-create-volume` sur StratusLab et de la commande `cinder create` sur OpenStack. Le disque est référencé sur les listes de volumes par son UUID.

Solution StratusLab :

```
$ stratus-create-volume --size=5 --tag=data-disk-school --private
$ stratus-describe-volumes -f tag:data
```

Solution OpenStack :

```
$ cinder create --display-name data-disk-school 5
$ nova volume-list
$ cinder list
```

0.2.2 Lancement de la MV sur StratusLab

Lancez une MV qui va vous servir à monter les deux disques. Pour cela, utilisez la commande `stratus-run-instance` avec entre autres l'option `--readonly-disk=Bque07c1p11TVu_TV0EUCBUWE9i` (vous pouvez vérifier sur le Marketplace à quoi correspond cet identifiant) et l'image `Marketplace_ID = 00q5xec9-qQnLIUuUQ4cbJoAfKz`. Vous devez attacher le DP après l'instanciation de la MV via la commande :

```
$ stratus-attach-volume -i VM_ID UUID
```

Connectez-vous à la MV en tant que `root` et vérifiez les disques présents avec la commande `fdisk -l`.

```
root@vm $ fdisk -l
Disk /dev/sda: 10.7 GB, 10737418240 bytes
```

Device	Boot	Start	End	Blocks	Id	System
/dev/sda1	*	1	1306	10484736	83	Linux

```
Disk /dev/sdb: 1610 MB, 1610612736 bytes
```

```
Disk /dev/sdc: 2147 MB, 2147483648 bytes
```

```
Disk /dev/vda: 5368 MB, 5368709120 bytes
```

Dans cette liste, `/dev/sdc` correspond au DRO et `/dev/vda` au DP. Formatez le DP avec la commande `mkfs.ext4`. Vérifiez à présent le type de disques présents avec la commande `blkid` :

```
root@vm $ blkid
/dev/sda1: UUID="4e6ea1d0-c280-42a8-9709-e28315ea7d3d" TYPE="ext4"
/dev/sr0: LABEL="_STRATUSLAB" TYPE="iso9660"
/dev/sdb: UUID="62bb0372-6d22-4efd-95b2-fc6e41cdd733" TYPE="swap"
/dev/sdc: LABEL="CDROM" TYPE="iso9660"
/dev/vda: UUID="5c4e6527-a0a3-4f1b-9e7e-e47f6149110c" TYPE="ext4"
```

Vous devez retrouver le DRO sous le type ISO et le DP sous le type `ext4`.

Créez un sous-répertoire `/tmp/Data` (ce chemin spécifique est nécessaire au bon fonctionnement de la base de données) et montez le DP sur le point de montage `/tmp/Data` avec l'aide de la commande `mount`. Montez le DRO sur le point de montage `/mnt`. Regardez le contenu des deux disques.

Solution :

```
$ stratus-run-instance --vm-name=school_cloud_tp4
--readonly-disk=Bque07c1p1lTVu_TV0EUCBUWE9i Marketplace_ID
$ stratus-attach-volume -i VM_ID UUID
$ ssh -X -i ~/.ssh/id_rsa root@vm
root@vm $ fdisk -l
root@vm $ mkfs.ext4 /dev/vda
root@vm $ blkid
root@vm $ mkdir /tmp/Data
root@vm $ mount /dev/vda /tmp/Data
root@vm $ mount /dev/sdc /mnt

root@vm $ ls /mnt/
log          simu_4AVUT5I  simu_GERDNY  simu_VXZAVVY
simu_2LIPFSI simu_AWHYGYA simu.tgz

root@vm $ ls /tmp/Data/
lost+found
```

0.2.3 Lancement de la MV sur OpenStack

Dans le cas d'OpenStack, il semble pas (encore) possible de démarrer une instance avec un volume attaché de type image. Les données sont donc stockées directement sur une image disque customisée. Lancez une MV qui va vous servir à monter le DP. Pour cela, utilisez la commande `nova boot` et l'image `Glance_ID = SL6.5_cloud_school_TP4`. Vous devez attacher le DP après l'instanciation de la MV via la commande :

```
$ nova volume-attach school_cloud_tp4 UUID
```

Connectez-vous à la MV en tant que `root` et vérifiez les disques présents avec la commande `fdisk -l`.

```
root@vm $ fdisk -l
Disk /dev/vda: 10.7 GB, 10737418240 bytes
```

Device	Boot	Start	End	Blocks	Id	System
/dev/vda1	*	1	1306	10484736	83	Linux

```
Disk /dev/vdb: 42.9 GB, 42949672960 bytes
```

```
Disk /dev/vdc: 5368 MB, 5368709120 bytes
```

Dans cette liste, `/dev/vdc` correspond au DP. Formatez le DP avec la commande `mkfs.ext4`. Vérifiez à présent le type de disques présents avec la commande `blkid` :

```
root@vm $ blkid
/dev/vda1: UUID="4e6ea1d0-c280-42a8-9709-e28315ea7d3d" TYPE="ext4"
/dev/vdb: LABEL="ephemeral10" UUID="a6926655-ccc6-464e-8e21-38a1126ab42f" SEC_TYPE="ext2" TYPE="ext2"
/dev/vdc: UUID="f986f622-e7d4-4b09-ad5f-d05f990956aa" TYPE="ext4"
```

Vous devez retrouver le DP sous le type `ext4`.

Créez un sous-répertoire `/tmp/Data` (ce chemin spécifique est nécessaire au bon fonctionnement de la base de données) et montez le DP sur le point de montage `/tmp/Data` avec l'aide de la commande `mount`. Regardez le contenu du DP.

Solution :

```
$ nova boot --key-name cloudkey --image SL6.5_cloud_school_TP4 --flavor m1.medium school_cloud
$ nova volume-attach school_cloud_tp4 UUID
```

```
$ nova add-floating-ip school_cloud_tp4 134.158.246.XX
$ ssh root@vm_cloud
root@vm $ fdisk -l
root@vm $ mkfs.ext4 /dev/vdc
root@vm $ blkid
root@vm $ mkdir /tmp/Data
root@vm $ mount /dev/vdc /tmp/Data

root@vm $ ls /tmp/Data/
lost+found
```

0.3 Utilisation de la base de données

Vous allez utiliser une base de données qui référence l'architecture des produits issues d'un pipeline de traitement de données. Vous devez d'abord copier les données sur le DP afin de pouvoir écrire sur la base de données, démarrer l'outil Pipelet et enfin visualiser le contenu de la base de données via une interface Web.

0.3.1 Copie des données

Toujours en tant que `root`, copier du DRO (ou de l'espace disque) les fichiers `simu.tgz` et `.sqlstatus` vers `/tmp/Data` du DP. Dé-tarez `simu.tgz` et changez le propriétaire de `/tmp/Data` avec la commande `chown` (de `root` à `cloud-user`, voir le script du TP2). Changez les permissions de la base de données avec la commande `chmod` pour que l'utilisateur puisse écrire et exécuter celle-ci.

Solution :

```
root@vm $ cp /mnt/simu.tgz /tmp/Data/
root@vm $ cp /mnt/.sqlstatus /tmp/Data/
root@vm $ cd /tmp/Data/
root@vm $ tar xvzf simu.tgz
root@vm $ chown -R cloud-user:cloud-user /tmp/Data/
root@vm $ chmod u+wx /tmp/Data/.sqlstatus
```

0.3.2 Interface Web

Quittez l'utilisateur `root` pour devenir `cloud-user`. La base de données `.sqlstatus` est utilisable via l'outil Pipelet qu'il faut démarrer par les commandes suivantes (la deuxième commande ajoute un nom, la dernière commande ajoute un utilisateur et demande de choisir un mot de passe associé) :

```
cloud-user@vm $ pipeweb start
cloud-user@vm $ pipeweb track data_base /tmp/Data/.sqlstatus
cloud-user@vm $ pipeutils -a user -l 2 /tmp/Data/.sqlstatus
```

Sur la MH (votre poste local), il reste à pouvoir accéder à une interface Web en utilisant le tunnel ssh suivant :

```
$ ssh -N -f -i ~/.ssh/id_rsa cloud-user@onevm-XX.lal.in2p3.fr
-L8081:onevm-XX.lal.in2p3.fr:8080 sleep 60
```

À partir de votre navigateur préféré en local (testé pour Chrome et Firefox), affichez l'URL :

```
localhost:8081/photoz/
```

Vous devez accéder à l'interface Web visible sur la Figure 1. Les données de la base de données sont protégées par le compte que vous avez fourni. À l'aide de l'interface Web, naviguez dans les répertoires afin de retrouver la Figure 2. Téléchargez-là sur votre poste local depuis son emplacement sur le DP.

Solution StratusLab :

```
$ scp -i ~/.ssh/id_rsa cloud-user@onevm-XX.lal.in2p3.fr:/tmp/Data/simu_AWHYGYA/
catalogue_7B4IGAQ/reco_TU2HJQJ/visu_JCFXBGY/data/visu_zz_MAG_REF_1.png .
```

Solution OpenStack :

```
$ scp cloud-user@vm_cloud:/tmp/Data/simu_AWHYGYA/
catalogue_7B4IGAQ/reco_TU2HJQJ/visu_JCFXBGY/data/visu_zz_MAG_REF_1.png .
```

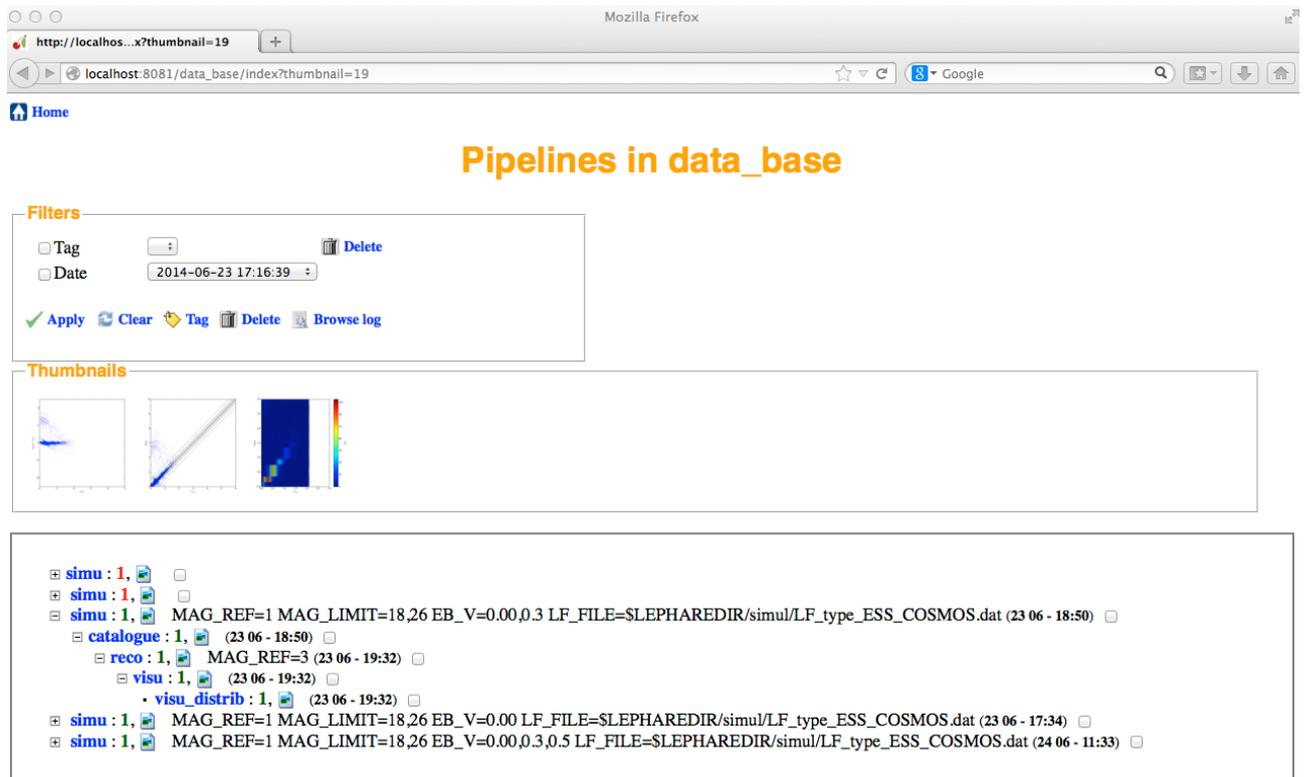


FIGURE 1 – Interface Web générée par l'outil Pipelet : l'architecture des répertoires présents sur le disque persistant est visible via l'interface.

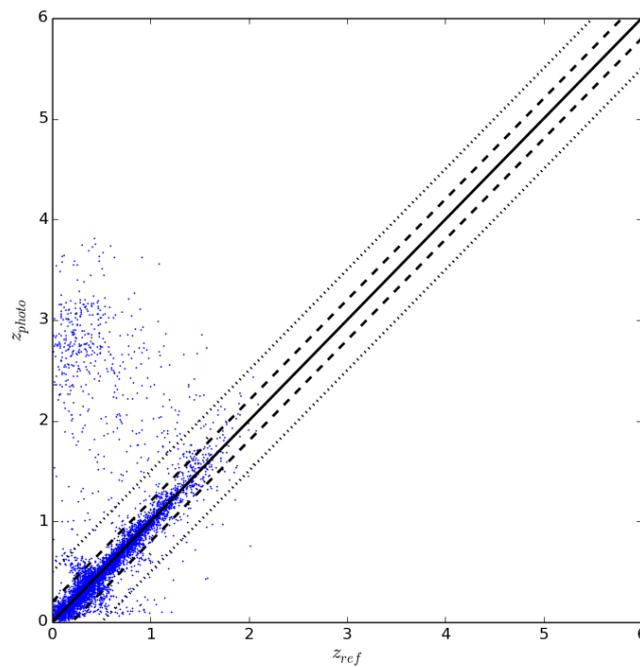


FIGURE 2 – Un des résultats du pipeline.