

## Haute disponibilité d'un serveur FTP

### Activité 1 - Installation et configuration de DRBD

#### Le fonctionnement résumé de DRBD est le suivant :

- Une ou plusieurs partitions (déclarée(s) dans un fichier de configuration) contenant les données à répliquer sont identiques sur chaque serveur ;
- Dans la plupart des cas, un seul serveur (celui qui est actif) est habilité à recevoir des données (sauf à utiliser un système de fichier comme OCFS -Oracle Cluster File System- qui permet la synchronisation de deux partitions montées en lecture et écriture) ;
- La synchronisation est réalisée en temps réel car elle se fait à la volée, pendant que les données sont modifiées : lorsqu'une opération en écriture a lieu sur cette partition, DRBD l'intercepte et envoie les blocs du disque dur qui ont été modifiés vers l'autre serveur ;
- l'utilisation est transparente : les applications (comme dans notre cas, le service FTP), qui enregistrent leurs données sur le périphérique de stockage répliqué, le font sans même savoir qu'il s'agit d'une unité de stockage spéciale.

#### Préalable

##### Nous supposons ici que chaque nœud disposera à terme de trois partitions :

- */dev/sda1* : une partition primaire sur laquelle on trouve le système, les données des utilisateurs et les applications ;
- */dev/sda2* : une partition primaire (que nous créerons) destinée à accueillir les données des utilisateurs (le répertoire */home*) ;
- */dev/sda3* : la partition swap.



***La désignation des partitions est à adapter par rapport à votre propre configuration.***

Proftpd sera installé sur la première partition du disque dur. Chaque utilisateur du système (dont les développeurs) aura accès, après authentification, en lecture et écriture à son propre FTP dans son répertoire personnel. À terme, l'ensemble des répertoires personnels sera donc déporté sur */dev/sda5*.

Par mesure de simplification, l'authentification sera une authentification « classique » via les fichiers système (*/etc/passwd*, */etc/shadow*, */etc/group*). C'est le comportement par défaut de proftpd ; ainsi, après l'installation, aucune modification n'est à opérer sur le fichier de configuration.

Le client FTP utilisé sur les STA peut être en ligne de commande et/ou graphique (comme *filezilla* disponible sur Windows et Linux).

#### Vous disposez de la documentation suivante :

- documentation technique de GSB (organisation/outils/GSB-DocumentTechnique.doc) pour l'installation et la configuration du serveur *proftpd* (fiche 5) ;
- **document 1** : installation de DRBD ;
- **document 2** : configuration de DRBD ;
- **document 3** : test de de synchronisation.



Sur chacun des deux serveurs, la partition */dev/sda2* est celle que DRBD duplique (dans l'exemple). Comme nous faisons le choix (arbitraire) d'enregistrer les métadonnées sur cette partition, elle ne doit pas être formatée (ou si elle l'est, elle devra être remise à zéro et toutes les données seront perdues).

*Il est possible d'utiliser une autre partition (non formatée ou remise à zéro) pour les métadonnées.*

## Travail à faire

### 1. Préparation des serveurs

#### Sur chacun des serveurs

- Q1.** Installez *proftpd* (vous garderez les propositions par défaut) et testez le démarrage du service.

**apt update**  
**apt install proftpd**

Le service a normalement démarré : **root@intralabXX:~# service proftpd status**

• proftpd.service - LSB: Starts ProFTPD daemon  
Loaded: loaded (/etc/init.d/proftpd)  
Active: active (running) since sam. 2016-07-23 01:08:08 CEST; 12min ago  
...

La configuration par défaut n'est ni optimisée ni finement sécurisée mais reste valable dans la cadre de ce Côté Labo.

- Q2.** À partir d'un client FTP, testez le service : il est rappelé que chaque utilisateur du système à accès en FTP à son répertoire personnel (créez un utilisateur si vous n'en avez pas) ==> il suffit donc de créer un fichier dans ce répertoire en tant que "user" et de voir s'il est accessible en FTP.

Éventuellement "useradd user"

Puis :

- su user
- touch /home/user/essai\_fichier\_ftp

Pour le test, un navigateur ou un client FTP comme Filezilla est nécessaire.

- Q3.** Créez la partition destinée à être répliquée sur chacun des serveurs. **Ne pas la formater pour le moment.**

Dans une console, l'utilitaire *cfdisk* pour redimensionner, supprimer ou ajouter une partition est relativement convivial.

Ressource éventuelle pour le partitionnement : <https://debian-facile.org/doc:systeme:cfdisk>.

### 2. Installation et configuration de DRBD (aide sur le document 1 jusqu'à l'étape 4 comprise)

L'installation et la configuration de DRBD se fera dans cette activité en dehors des problématiques de Corosync et de Pacemaker.

- Q4.** Activez le module DRBD et installez les outils de gestion (drbd8-utils). Vous procéderez aux vérifications nécessaires

Voir document 1

- Q5.** Configurez DRBD sur chacun des serveurs. Vous procéderez, à chaque étape, aux tests nécessaires.



**Attention au nommage de la partition créée qui sera synchronisée via DRBD.**

Voir document 2 (étape 1 à 4)

### 3. Bascule des "home" et tests

**Q6.** Basculez les "home" sur la partition à répliquer (aide dans le document 2 – Étape 5)



**Cette manipulation ne doit être faite que sur le maître (intralabXX) car tout ce qui est écrit dans cette partition est répliquée sur l'esclave.**

```
mkdir /root/sauvhome  
mv /home/* /root/sauvhome/  
mount /dev/drbd0 /home/  
mv /root/sauvhome/* /home/  
rmdir /root/sauvhome
```

**Q7.** Testez la solution et complétez le tableau des tests permettant de vérifier que la solution est opérationnelle (aide dans le document 3).

Les tests doivent se faire à partir d'un client FTP sur l'adresse IP virtuelle pour vérifier l'écriture dans la partition dupliquée.

On part d'une situation où :

- le serveur actif du cluster est aussi le serveur maître pour drbd. Les ressources (VIP et Apache) sont lancées sur le serveur actif.
- 
- la commande "mount /dev/drbd0 /home" a été effectuée sur le serveur maître (la partition /home n'est pas montée sur le serveur esclave et n'est donc pas accessible directement) ;
- via l'IP virtuelle, le client FTP accède donc au service FTP lancé sur le serveur maître.

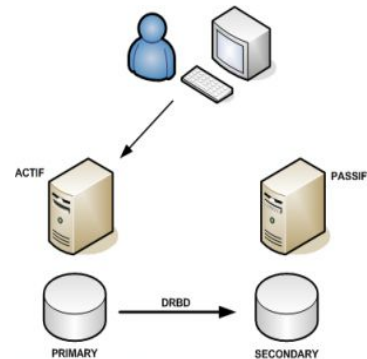
Actions à effectuer	Résultats attendus	Résultats obtenus	Statut
Sur le serveur maître, tester une connexion à partir d'un client FTP et créer un fichier dans un répertoire.	Le fichier doit être créé et téléchargeable.	Le fichier est créé et téléchargeable.	OK
On met le serveur maître en standby et on bascule manuellement le service drbd sur le serveur esclave (on démonte /home sur le serveur maître et on monte la partition sur le serveur esclave. On vérifie que le fichier a bien été dupliqué.	Les services (dont l'IP virtuelle) sont basculés sur le serveur esclave. Le fichier créé doit être présent sur le serveur esclave.	Le fichier créé est présent sur le serveur esclave.	OK
Sur le serveur esclave, tester une connexion à partir d'un client FTP et créer un autre fichier dans un répertoire.	Le fichier doit être créé et téléchargeable.	Le fichier est créé et téléchargeable.	OK
On met le serveur maître on line et on bascule manuellement le service drbd sur le serveur maître. On vérifie que le second fichier a bien été dupliqué.	Les services (dont l'IP virtuelle) sont basculés sur le serveur maître. Le fichier créé doit être présent sur le serveur maître.	Le fichier créé est présent sur le serveur maître.	OK

## Documents

### Document 1 : installation de DRBD

Le **module DRBD (Distributed Replicated Block Device)** est un autre composant du High-Availability Linux Project qui a pour rôle de synchroniser les données entre deux postes. DRBD offre une excellente garantie de réplication des données en fonctionnant en RAID-1 par IP : les données sont copiées simultanément sur deux postes via le réseau.

DRBD est constitué d'un module du noyau (le cœur du système), qui se comporte comme un driver de système de fichiers, et d'outils de gestion regroupés dans le paquetage **drbd8-utils**.



### Vérifications et installation des outils de gestion

Le noyau installé avec *debian Jessie* (version stable actuel) dispose du module DRBD dans sa version 8.4.3.

Sur chaque serveur, Vous pouvez le constater en chargeant le module (***modprobe drbd***). À noter qu'il ne sera pas nécessaire de gérer le module soi-même à chaque redémarrage car, après installation des outils de gestion **drbd8-utils**, DRBD va le charger automatiquement via le service (script `/etc/init.d/drbd`).

Pour vérifier qu'il soit bien chargé : ***lsmod | grep drbd***

**Pour obtenir des informations sur le module installé : *modinfo drbd***

```
filename:    /lib/modules/3.16.0-4-amd64/kernel/drivers/block/drbd/drbd.ko
alias:       block-major-147-*
license:     GPL
version:     8.4.3
description: drbd - Distributed Replicated Block Device v8.4.3
...
```

Et la commande ***cat /proc/drbd*** doit renvoyer (avant configuration) :

```
version: 8.4.3 (api:1/proto:86-101)
srcversion: 1A9F77B1CA5FF92235C2213
```

Il est nécessaire ensuite d'installer les outils de gestion **drbd8-utils**.

## Document 2 : configuration de DRBD

### Étape 1 : les fichiers de configuration et démarrage du service

La configuration s'articule autour des fichiers suivants qui doivent être identiques sur les deux serveurs.

- `/etc/drbd.conf` : fichier général qui inclut les autres (nous ne le modifierons pas).
- `/etc/drbd.d/global_common.conf` : pour les paramètres généraux et communs à toutes les ressources.
- `/etc/drbd.d/nom_ressource.res` : un fichier pour chaque ressource (une ressource correspond à UNE partition (block device) sur laquelle DRBD va agir).

**`/etc/drbd.d/global_common.conf` (à modifier sur chaque serveur)** : ce fichier contient les directives communes à toutes les ressources. Ces directives peuvent être surchargées lors de la déclaration des ressources. Dans ce contexte d'expérimentation, on n'ajoutera qu'une seule ligne :

Directives et valeurs	Explications
<code>global {</code> <code>  usage-count no;</code> <code>}</code>	Statistiques sur l'usage des versions collectées par le projet DRBD. Si la valeur est à "yes", un serveur web est contacté à chaque fois qu'une nouvelle version de DRBD est installée sur un serveur.
<code>common {</code> <code>  <b>protocol C;</b></code> <code>...</code> <code>}</code>	C'est le protocole le plus sécurisé : les opérations d'écriture en local sur le nœud principal sont considérées comme achevées lorsque que les écritures sur les 2 partitions ont été confirmées.

**La déclaration des ressources : fichier `/etc/drbd.d/ftp.res` (à créer sur chaque serveur) :**

Directives et valeurs	Explications
<code>resource resftp {</code>	Définition de la ressource « resftp »
<code>  device /dev/drbd0;</code>	DRBD va créer un disque virtuel...
<code>  disk /dev/sda2;</code>	... Qui sera monté sur le disque physique <code>/dev/sda2</code>
<code>  meta-disk internal;</code>	Gestion des méta-données (ici, enregistrés sur le même disque)
<code>  on intralabXX {</code> <code>    address 10.22.100.212:7789;</code> <code>  }</code>	Déclaration du premier nœud (adresse IP et port)
<code>  on hdintralabXX {</code> <code>    address 10.22.100.215:7789;</code> <code>  }</code> <code>}</code>	Déclaration du deuxième nœud (adresse IP et port)

**`device /dev/drbd0`** : le block device est conventionnellement nommée `/dev/drbdX`, ou X est le numéro de périphérique mineur. Si une autre ressource (correspondante à une autre partition) est créée, le « device » sera `/dev/drbd1`.

**`disk /dev/sda2`** : attention à bien utiliser les noms de *devices* indiqués par *fdisk* ou *lvdisplay* (pas ceux du *fstab* qui peuvent être des labels, ni ceux du "mount" qui peuvent être différents si LVM est utilisé).

**`meta-disk internal`** : une autre configuration possible est d'externaliser les méta-données sur une autre partition (recommandée par DRBD mais décriée par certains spécialistes à cause du délai d'accès disque plus long que lorsque les méta-données sont sur le même disque) ; il faut préciser dans ce cas la partition et l'emplacement des données dans la partition, par exemple *meta-disk /dev/sda6[0]*. Dans ce cas, la partition qui contient les données peut être formatée normalement ou laissée telle quelle si elle contient déjà les données utiles.

**Les noms de machines après les "on..."** doivent correspondre à ce que retourne "uname -n".



Ne pas encore démarrer le service ***drbd***

## Étape 2 : initialisation des métadonnées sur la partition qui va contenir les données

Initialisation des méta-données sur **chaque nœud** :

```
drbdadm create-md resftp
```

Plusieurs cas peuvent se présenter :

### Cas N°1

La partition vient d'être créée et n'est pas formatée. La commande devrait fonctionner sans problème et renvoyer le résultat suivant :

```
initializing activity log
NOT initializing bitmap
Writing meta data...
New drbd meta data block successfully created.
```

**Vous pouvez démarrer le service *drbd* sur les 2 serveurs : `service drbd start`**



Faire éventuellement un service drbd restart si vous avez le message "Warning: Journal has been rotated since unit was started. Log output is incomplete or unavailable."

**root@intralabXX:~# /etc/init.d/drbd status**

```
• drbd.service - Distributed Replicated Block Device
  Loaded: loaded (/lib/systemd/system/drbd.service; enabled)
  Active: active (exited) since sam. 2016-07-23 13:26:49 CEST; 9min ago
  Process: 10740 ExecStart=/sbin/drbdadm sh-b-pri all (code=exited, status=0/SUCCESS)
...
juil. 23 13:26:49 intralabXX drbdadm[10688]: [
juil. 23 13:26:49 intralabXX drbdadm[10688]: prepare net: resftp
juil. 23 13:26:49 intralabXX drbdadm[10688]: adjust net: resftp
juil. 23 13:26:49 intralabXX drbdadm[10688]: ]
```

### Cas N°2

La partition a déjà été formatée (éventuellement des données sont écrites). Dans ce cas, vous devriez avoir le message d'erreur ci-dessous :

```
you are the 1012th user to install this version
md_offset 2600464384
al_offset 2600431616
bm_offset 2600349696

Found ext4 filesystem
  2539520 kB data area apparently used
  2539404 kB left usable by current configuration

Device size would be truncated, which
would corrupt data and result in
'access beyond end of device' errors.
You need to either
  * use external meta data (recommended)
  * shrink that filesystem first
  * zero out the device (destroy the filesystem)
Operation refused.

Command 'drbdmeta 0 v08 /dev/sda2 internal create-md' terminated with exit code 40
drbdadm create-md r0: exited with code 40
```

Il est nécessaire de détruire le système de fichier. Pour cela :

**Sur le nœud maître :**

- sauvegarder (dans /root par exemple), droits sur les objets compris, le contenu de la partition s'il s'agit de la partition qui contient les données (et que des données utiles sont déjà dessus).
- démonter la partition : ***umount /dev/sda2*** ;
- détruire le système de fichier : ***shred -zvf -n 1 /dev/sda2*** ;
- exécuter la commande ***drbdadm create-md resftp*** qui doit maintenant fonctionner.

**Sur le nœud esclave**, même s'il s'agit d'une partition qui contient les données, il n'est pas besoin d'une quelconque sauvegarde puisque les données vont être répliquées à partir du maître :

- Démonter la partition : ***umount /dev/sda2***.
- Détruire le système de fichier : ***shred -zvf -n 1 /dev/sda2***
- Exécuter la commande ***drbdadm create-md resftp***

**Vous pouvez démarrer le service *drbd* sur les 2 serveurs : *service drbd start***

**root@intralabXX:~# /etc/init.d/drbd status**

```
● drbd.service - Distributed Replicated Block Device
   Loaded: loaded (/lib/systemd/system/drbd.service; enabled)
   Active: active (exited) since sam. 2016-07-23 13:26:49 CEST; 9min ago
   Process: 10740 ExecStart=/sbin/drbdadm sh-b-pri all (code=exited, status=0/SUCCESS)
...
juil. 23 13:26:49 intralabXX drbdadm[10688]: [
juil. 23 13:26:49 intralabXX drbdadm[10688]: prepare net: resftp
juil. 23 13:26:49 intralabXX drbdadm[10688]: adjust net: resftp
juil. 23 13:26:49 intralabXX drbdadm[10688]: ]
```



Faire éventuellement un service drbd restart si vous avez le message "Warning: Journal has been rotated since unit was started. Log output is incomplete or unavailable."

**État des 2 nœuds avec la commande "cat /proc/drbd" :**

IntralabXX	hdintralabXX
0: cs:Connected ro:Secondary/Secondary ds:Inconsistent/Inconsistent C r----- ns:0 nr:0 dw:0 dr:0 al:0 bm:0 lo:0 pe:0 ua:0 ap:0 ep:1 wo:f oos:3144572	0: cs:Connected ro:Secondary/Secondary ds:Inconsistent/Inconsistent C r----- ns:0 nr:0 dw:0 dr:0 al:0 bm:0 lo:0 pe:0 ua:0 ap:0 ep:1 wo:f oos:3144572

La ressource, « 0 » pour drbd0, est connectée (*cs:Connected*) et les 2 nœuds ont été trouvés. Aucun des nœuds n'est maître (*ro:Secondary/Secondary* : le premier *Secondary* représente le nœud sur lequel la commande est exécutée). *ds:Inconsistent/Inconsistent* indique le statut des disques et est l'état normal à ce stade (aucune synchronisation n'a encore été réalisée).

**Il faut maintenant procéder à la première synchronisation.**



## Étape 3 : première synchronisation

Sur le nœud maître (intralabXX), pour lancer la première synchronisation :

```
root@intralabXX:~# drbdadm -- --overwrite-data-of-peer primary resftp
```

Le statut a changé :

```
srcversion: 1A9F77B1CA5FF92235C2213
0: cs:SyncSource ro:Primary/Secondary ds:UpToDate/Inconsistent C r-----
   ns:245760 nr:0 dw:0 dr:246672 al:0 bm:15 lo:0 pe:0 ua:0 ap:0 ep:1 wo:f oos:2898812
   [>.....] sync'ed: 8.0% (2898812/3144572)K
   finish: 0:01:10 speed: 40,960 (40,960) K/sec
```

La ressource est en train de se synchroniser (cs:SyncSource) à partir du nœud sur lequel a été lancée la commande qui est en « Primary ». Le statut du disque du nœud secondaire « Inconsistent » est normal tant que la synchronisation n'est pas terminée.

**Remarque** : les chiffres indiqués ne sont pas très représentatifs de ce que l'on peut obtenir avec des disques réels (tous ces tests sont faits à partir d'une machine virtuelle).

Une fois la synchronisation terminée, le statut de DRBD doit être le suivant :

IntralabXX	hdintralabXX
0: cs:Connected ro:Primary/Secondary ds:UpToDate/UpToDate C r----- ns:3144572 nr:0 dw:0 dr:3145484 al:0 bm:192 lo:0 pe:0 ua:0 ap:0 ep:1 wo:f oos:0	0: cs:Connected ro:Secondary/Primary ds:UpToDate/UpToDate C r----- ns:0 nr:3144572 dw:3144572 dr:0 al:0 bm:192 lo:0 pe:0 ua:0 ap:0 ep:1 wo:f oos:0

## Étape 4 : formatage de la partition sur le nœud primaire

Sur le nœud primaire, il est nécessaire de formater le disque virtuel.



Le disque à formater n'est pas /dev/sda2 mais /dev/drbd0 :

```
root@intralabXX:~# mkfs.ext4 /dev/drbd0
```

```
mke2fs 1.42.12 (29-Aug-2014)
```

```
Rejet des blocs de périphérique : complété
```

```
En train de créer un système de fichiers avec 786143 4k blocs et 196608 i-noeuds.
```

```
UUID de système de fichiers=374d9e91-d39a-4f3c-8159-f3ce150033e6
```

```
...
```

```
Création du journal (16384 blocs) : complété
```

```
Écriture des superblocs et de l'information de comptabilité du système de fichiers : complété
```



## Étape 5 : bascule de /home et montage de la partition sur le nœud primaire

Il est nécessaire ensuite de monter la partition (le but ici est de synchroniser les « /home ») :



### **Le montage se fera donc sous /home et doit utiliser /dev/drbd0**

/home est déjà utilisé (il y a notamment les données de l'utilisateur *user*) et les données ont été écrites sur la partition correspondante. Pour basculer /home sur la partition répliquée via *drbd*, une série de manipulations doit être opérée.

### **Sur le nœud primaire**

1. Créer dans /root un répertoire sauvhome et déplacer le contenu de /home dans ce répertoire (il n'y a théoriquement que le répertoire *user* correspondant au compte *user* créé).
2. Monter la partition (le contenu de la nouvelle partition doit être lisible à partir du répertoire /home).  
**root@intralabXX:~# mount /dev/drbd0 /home/**
3. Remettre le répertoire /user dans /home.
4. Tester l'accès au service FTP qui devrait être de nouveau opérationnel.

### **Sur le nœud secondaire**

La synchronisation de la partition est effective ==> les données ont aussi été écrites sur la partition du serveur esclave ==> Il suffit de supprimer tout le contenu du répertoire /home.



À noter que si /home était monté sur une partition spécifique (ce qui n'est théoriquement pas notre cas), il faut le démonter et supprimer, sur chaque serveur, la ligne correspondante dans */etc/fstab*.

## Document 3 : test de de synchronisation

### Tester la synchronisation

Par défaut, DRBD fonctionne en mode Primary/Secondary. Ceci signifie que l'un des disques supporte les lectures et écritures (c'est le nœud primaire), mais que le second n'accepte que les opérations en lecture.



La partition ne doit jamais être montée sur les deux nœuds en même temps. Le nœud sur lequel la partition est montée doit être déclaré en "primary" et le nœud sur lequel la partition n'est pas montée est déclarée en "secondary"

Pour pouvoir tester la synchronisation, il est nécessaire sur le nœud primaire de :

- démonter la partition : **`umount /dev/drdb0`** ;
- changer le statut de la ressource pour « secondaire » : **`drbdadm secondary resftp`**.

Et sur le nœud secondaire :

- changer le statut de la ressource pour « primaire » : **`drbdadm primary resftp`** ;
- monter la partition : **`mount /dev/drdb0 /home`**

Vous devriez retrouver les données restaurées sur le nœud primaire dans /home... et vous pouvez y créer un fichier pour vérifier qu'après l'opération inverse, celui-ci se retrouve bien sur le nœud primaire.

Le service FTP doit être aussi opérationnel sur le nœud secondaire (et ne plus l'être sur le nœud primaire puisque l'accès aux données n'est plus effectif).



Il est possible d'obtenir des disques répliqués en mode primary/primary. Cela signifie que les 2 disques supportent les écritures en simultanée. Mais il faut pour cela un système de fichiers qui le supporte comme OCFS2 ou GFS.