



Rencontres Mathrice

4 Octobre 2011 - Lyon

Un petit aperçu

Qu'est que DRBD ?

Définition Wikipedia (<http://fr.wikipedia.org>)

DRBD (Distributed Replicated Block Device en anglais, ou périphérique en mode bloc répliqué et distribué en français) est une architecture de stockage distribuée pour GNU Linux, permettant la replication de périphériques de bloc (disques, partitions, volume logiques etc...) entre des serveurs. DRBD est un logiciel libre, il est composé d'un module noyau, d'outils d'administration en espace utilisateur ainsi que de scripts shell.

La réplication des données se fait:

- En temps réel. En permanence, pendant que les applications modifient les données présentes sur le périphérique
- De façon transparente. Les applications qui stockent leurs données sur le périphérique répliqué n'ont pas conscience que ces données sont en fait stockées sur plusieurs ordinateurs
- De façon synchrone, ou asynchrone. En fonctionnement synchrone, une application qui déclenche une écriture de donnée est notifiée de la fin de l'opération seulement après que l'écriture a été effectuée sur tous les serveurs, alors qu'en fonctionnement asynchrone, la notification se fait après que la donnée a été écrite localement, mais avant la propagation de la donnée.

Drbd est intégré au noyau linux depuis la version 2.6.33 (février 2010)

Fonctionnalités de DRBD

- 2 nœuds au maximum (N dans la prochaine version)
- Synchronisation en temps réel (
- Utilisation transparente

Différents modes de synchronisation

asynchrone (mode A)

les écritures locales sont considérées terminées une fois que l'écriture a eu lieu sur le disque local et que l'information a été enregistrée dans le tampon TCP d'envoi

Semi-synchrone (mode B)

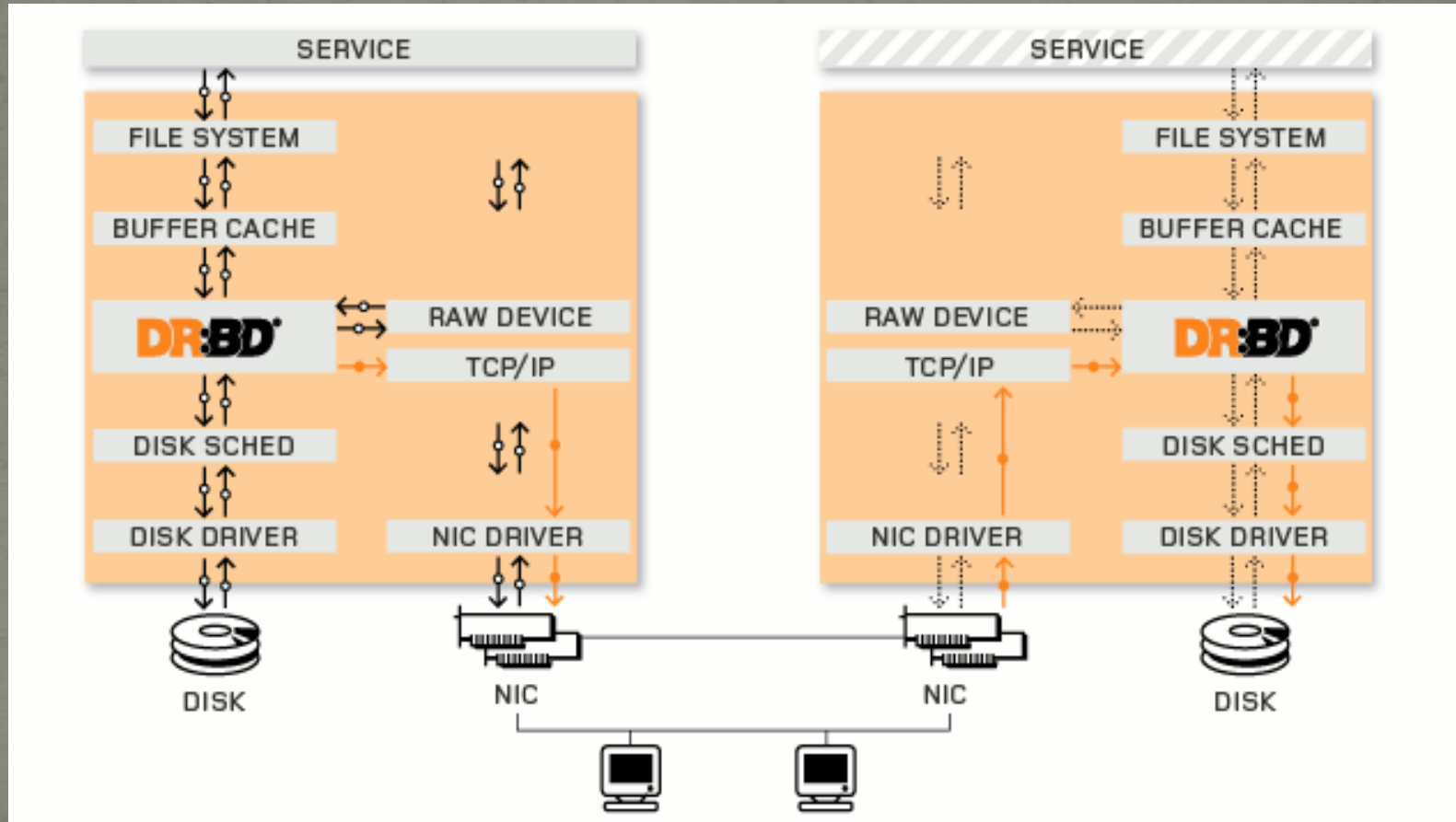
les écritures locales sont considérées terminées une fois que l'écriture a eu lieu sur le disque local et que l'information a été reçue par le deuxième serveur

Synchrone (mode C)

les écritures locales sont considérées terminées une fois que l'écriture a eu lieu sur le disque local et sur le disque distant

Fonctionnalités de DRBD

<http://www.drbd.org/>



Installation de DRBD

Installer drbd

```
#aptitude install drbd8-utils  
#lsmod ( pour verifier si le module est chargé )  
#modeprobe drbd (pour le charger)
```

Plus loins un exemple de drbd.conf en Primary/Primary et Primary/Secondary

En **Primary/Primary** : Tous les neuds peuvent écrire en même temps, file system préconisé pour le stockage partagé : GFS/OCFS

En **Primary/Secondary**, pas de préconisation de File system

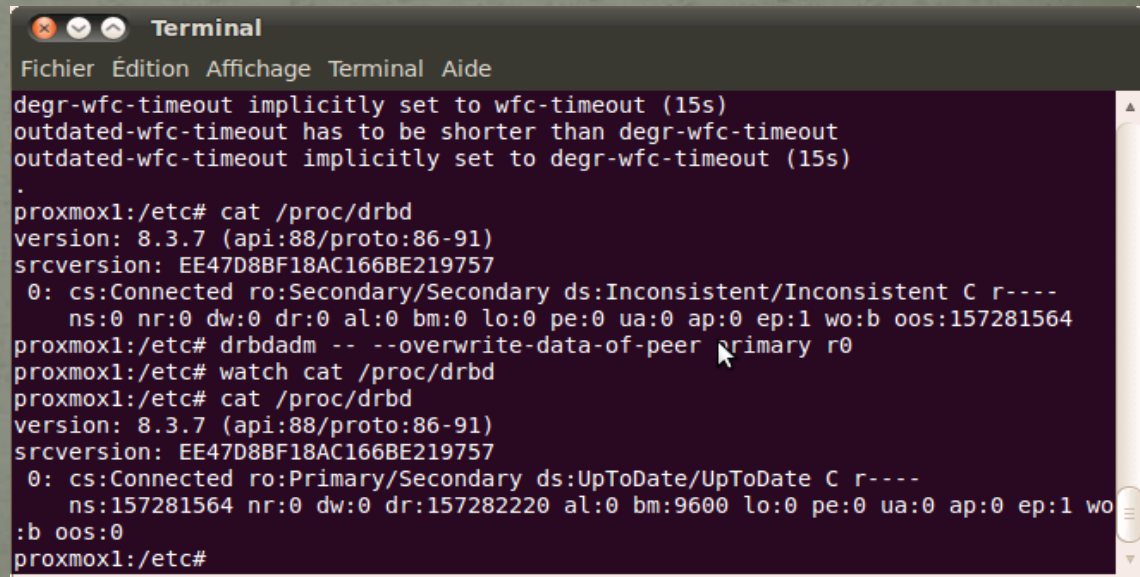
Démarrer drbd

```
sur les 2 noeuds  
#/etc/init.d/drbd start  
#drbdadm create-md ro  
#drbdadm up ro
```

Vérification de l'état de l'allocation
de la ressource ro par DRBD

```
#cat /proc/drbd
```

(pour voir Secondary/Secondary
Inconsistent/Inconsistent)



```
Terminal  
Fichier Édition Affichage Terminal Aide  
degr-wfc-timeout implicitly set to wfc-timeout (15s)  
outdated-wfc-timeout has to be shorter than degr-wfc-timeout  
outdated-wfc-timeout implicitly set to degr-wfc-timeout (15s)  
.  
proxmox1:/etc# cat /proc/drbd  
version: 8.3.7 (api:88/proto:86-91)  
srcversion: EE47D8BF18AC166BE219757  
 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:b oos:157281564  
proxmox1:/etc# drbdadm -- --overwrite-data-of-peer primary r0  
proxmox1:/etc# watch cat /proc/drbd  
proxmox1:/etc# cat /proc/drbd  
version: 8.3.7 (api:88/proto:86-91)  
srcversion: EE47D8BF18AC166BE219757  
 0: cs:Connected ro:Primary/Secondary ds:UpToDate/UpToDate C r----  
    ns:157281564 nr:0 dw:0 dr:157282220 al:0 bm:9600 lo:0 pe:0 ua:0 ap:0 ep:1 wo  
:b oos:0  
proxmox1:/etc#
```

Installation de DRBD

```
# /etc/init.d/drbd stop
```

```
#/etc/init.d/drbd start
```

Première synchro sur un nœud

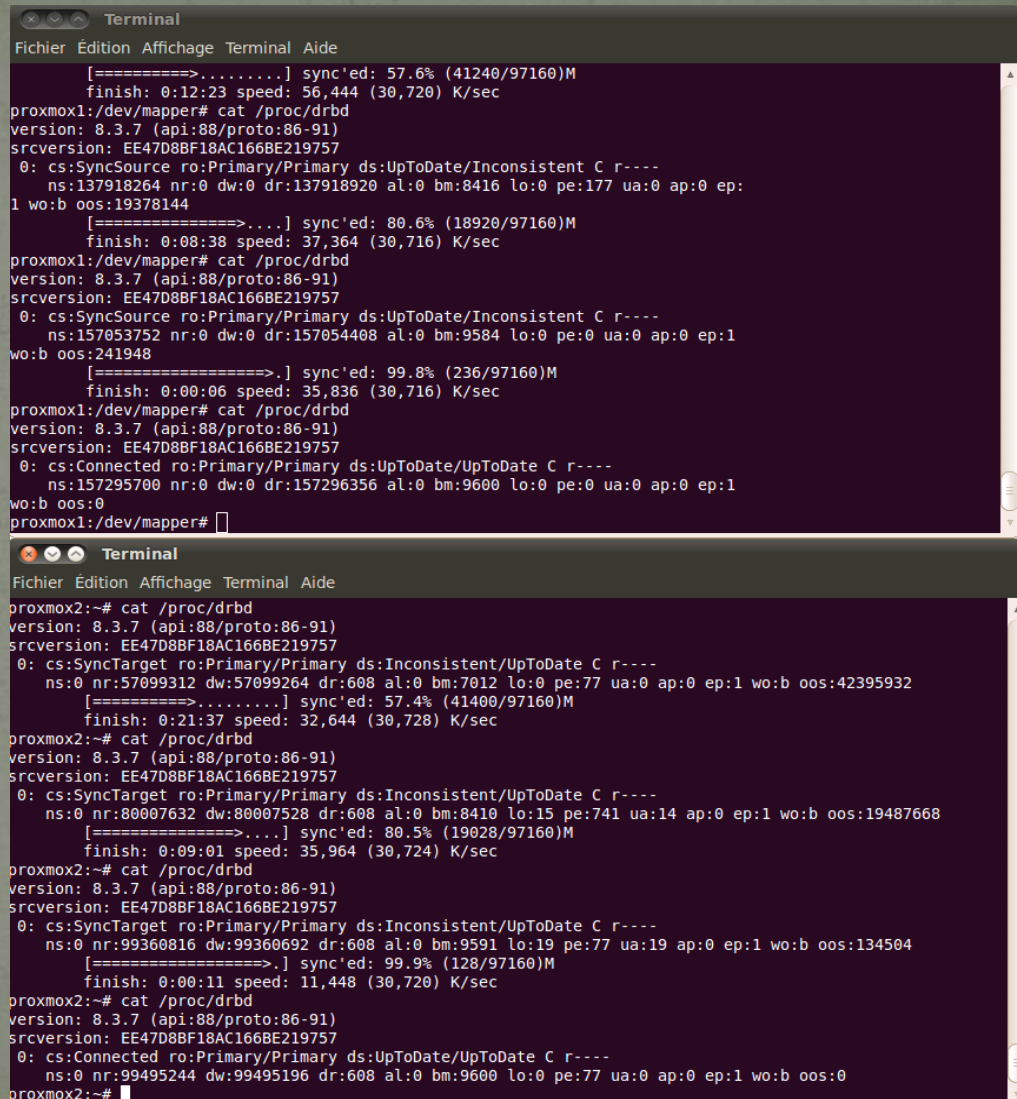
```
#drbdadm -- --overwrite-data-of-peer primary ro
```

```
# cat /proc/drbd
```

Au final Primary/Primary

```
# /etc/init.d/drbd stop
```

```
#/etc/init.d/drbd start
```



```
Terminal
Fichier Edition Affichage Terminal Aide
[=====>.....] sync'ed: 57.6% (41240/97160)M
finish: 0:12:23 speed: 56,444 (30,720) K/sec
proxmox1:/dev/mapper# cat /proc/drbd
version: 8.3.7 (api:88/proto:86-91)
srcversion: EE47D8BF18AC166BE219757
0: cs:SyncSource ro:Primary/Primary ds:UpToDate/Inconsistent C r----
ns:137918264 nr:0 dw:0 dr:137918920 al:0 bm:8416 lo:0 pe:177 ua:0 ap:0 ep:
1 wo:b oos:19378144
[=====>...] sync'ed: 80.6% (18920/97160)M
finish: 0:08:38 speed: 37,364 (30,716) K/sec
proxmox1:/dev/mapper# cat /proc/drbd
version: 8.3.7 (api:88/proto:86-91)
srcversion: EE47D8BF18AC166BE219757
0: cs:SyncSource ro:Primary/Primary ds:UpToDate/Inconsistent C r----
ns:157053752 nr:0 dw:0 dr:157054408 al:0 bm:9584 lo:0 pe:0 ua:0 ap:0 ep:1
wo:b oos:241948
[=====>.] sync'ed: 99.8% (236/97160)M
finish: 0:00:06 speed: 35,836 (30,716) K/sec
proxmox1:/dev/mapper# cat /proc/drbd
version: 8.3.7 (api:88/proto:86-91)
srcversion: EE47D8BF18AC166BE219757
0: cs:Connected ro:Primary/Primary ds:UpToDate/UpToDate C r----
ns:157295700 nr:0 dw:0 dr:157296356 al:0 bm:9600 lo:0 pe:0 ua:0 ap:0 ep:1
wo:b oos:0
proxmox1:/dev/mapper#

Terminal
Fichier Edition Affichage Terminal Aide
proxmox2:~# cat /proc/drbd
version: 8.3.7 (api:88/proto:86-91)
srcversion: EE47D8BF18AC166BE219757
0: cs:SyncTarget ro:Primary/Primary ds:Inconsistent/UpToDate C r----
ns:0 nr:57099312 dw:57099264 dr:608 al:0 bm:7012 lo:0 pe:77 ua:0 ap:0 ep:1 wo:b oos:42395932
[=====>.....] sync'ed: 57.4% (41400/97160)M
finish: 0:21:37 speed: 32,644 (30,728) K/sec
proxmox2:~# cat /proc/drbd
version: 8.3.7 (api:88/proto:86-91)
srcversion: EE47D8BF18AC166BE219757
0: cs:SyncTarget ro:Primary/Primary ds:Inconsistent/UpToDate C r----
ns:0 nr:80007632 dw:99360692 dr:608 al:0 bm:8410 lo:15 pe:741 ua:14 ap:0 ep:1 wo:b oos:19487668
[=====>...] sync'ed: 80.5% (19028/97160)M
finish: 0:09:01 speed: 35,964 (30,724) K/sec
proxmox2:~# cat /proc/drbd
version: 8.3.7 (api:88/proto:86-91)
srcversion: EE47D8BF18AC166BE219757
0: cs:SyncTarget ro:Primary/Primary ds:Inconsistent/UpToDate C r----
ns:0 nr:99360816 dw:99360692 dr:608 al:0 bm:9591 lo:19 pe:77 ua:19 ap:0 ep:1 wo:b oos:134504
[=====>.] sync'ed: 99.9% (128/97160)M
finish: 0:00:11 speed: 11,448 (30,720) K/sec
proxmox2:~# cat /proc/drbd
version: 8.3.7 (api:88/proto:86-91)
srcversion: EE47D8BF18AC166BE219757
0: cs:Connected ro:Primary/Primary ds:UpToDate/UpToDate C r----
ns:0 nr:99495244 dw:99495196 dr:608 al:0 bm:9600 lo:0 pe:77 ua:0 ap:0 ep:1 wo:b oos:0
proxmox2:~#
```

Installation de DRBD

Primary/Primary

```
/etc/drbd.conf (identique sur tous les nodes)
global { usage-count no; }
common { syncer { rate 30M; } }
resource ro { protocol C;
  startup { wfc-timeout 15;
            degr-wfc-timeout 60;
            become-primary-on both; }
  net {
    cram-hmac-alg sha1;
    shared-secret "my-secret";
    allow-two-primaries;
    after-sb-opri discard-zero-changes;
    after-sb-1pri discard-secondary;
    after-sb-2pri disconnect; }
  on proxmox1 {
    device /dev/drbd0;
    disk /dev/pve/hotshare;
    address 192.168.1.101:7788;
    meta-disk internal; }
  on proxmox2 {
    device /dev/drbd0;
    disk /dev/pve/hotshare;
    address 192.168.1.102:7788;
    meta-disk internal; }
}
```

Primary/Secondary

```
/etc/drbd.conf (identique sur tous les nodes)
global { usage-count no; }
common { syncer { rate 30M; } }
resource ro { protocol C;

  net {
    cram-hmac-alg sha1;
    shared-secret "my-secret";

    }
  on proxmox1 {
    device /dev/drbd0;
    disk /dev/pve/hotshare;
    address 192.168.1.101:7788;
    meta-disk internal; }
  on proxmox2 {
    device /dev/drbd0;
    disk /dev/pve/hotshare;
    address 192.168.1.102:7788;
    meta-disk internal; }
}
```

Installation de DRBD

Modification du filtre lvm (/etc/lvm/lvm.conf)

By default we accept every block device:

```
filter = filter = ["a|sd.*|", "a|drbd.*|", "r|..*|"]
```

```
#pvcreate /dev/drbd0
```

```
#vgcreate drbdvg /dev/drbd0
```

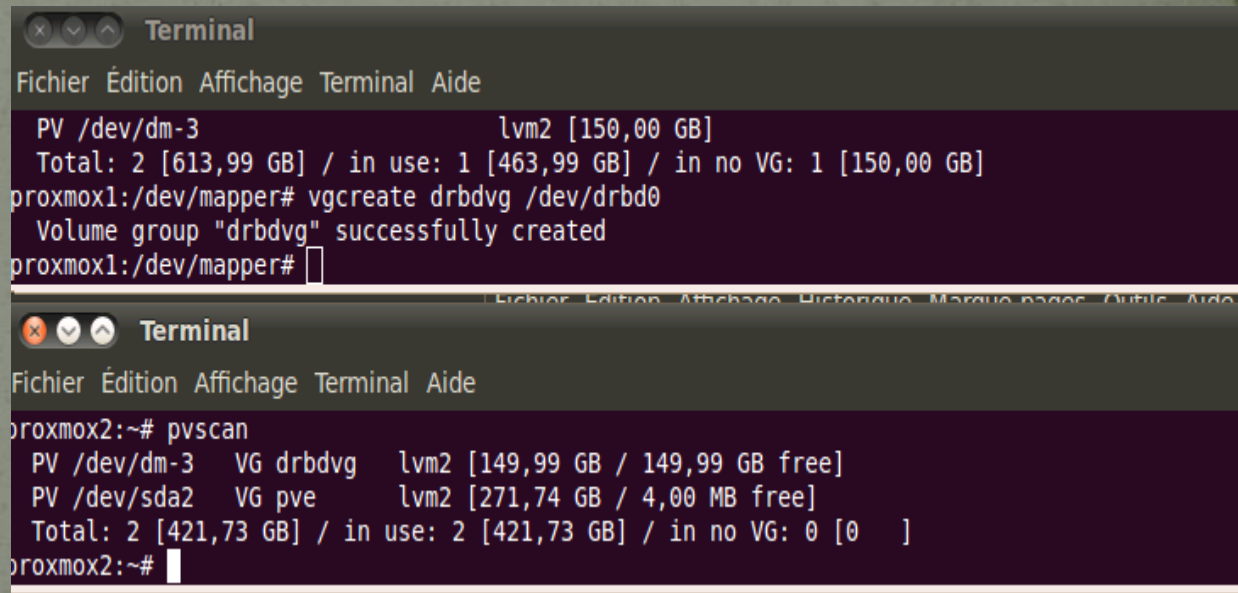
```
# mkfs.ext3 /dev/drbd0 (P/S)
```

```
# mkfs.gfs /dev/drbd0 (P/P)
```

```
# mkdir /hotshare
```

```
#mount /dev/drbd0 /hotshare
```

```
#/etc/init.d/drbd start
```



```
Terminal
Fichier Édition Affichage Terminal Aide
PV /dev/dm-3          lvm2 [150,00 GB]
Total: 2 [613,99 GB] / in use: 1 [463,99 GB] / in no VG: 1 [150,00 GB]
proxmox1:/dev/mapper# vgcreate drbdvg /dev/drbd0
Volume group "drbdvg" successfully created
proxmox1:/dev/mapper#

Terminal
Fichier Édition Affichage Historique Marque pages Outils Aide
proxmox2:~# pvscan
PV /dev/dm-3   VG drbdvg   lvm2 [149,99 GB / 149,99 GB free]
PV /dev/sda2  VG pve     lvm2 [271,74 GB / 4,00 MB free]
Total: 2 [421,73 GB] / in use: 2 [421,73 GB] / in no VG: 0 [0 ]
proxmox2:~#
```

Ajout du Storage LVM par l'interface http.