Table des matières

Renommer des disques /dev/emcpowerX 3
Problème de cohérence chemins ⇔ devices PowerPath 4
Found Duplicate PV 5
Rescanner les devices 5
Cartes Emulex 5
Cartes QLogic 5
Trouver les WWN 5
Erreur lors du renommage des disques 5
Export / import de conf PowerPath 5
Déterminer le modèle d'une carte SAN 6
Supprimer un disque SAN 6
Montage des FS sur SAN 6
Faire clignoter les LEDs d'une carte Emulex 6
forcerescan-scsi-bus.sh / read-only 6

Renommer des disques /dev/emcpowerX

Soit 2 noeuds d'un cluster : *noeud1* et *noeud2*. Chaque noeud dispose de plusieurs disques SAN */dev/emcpowerX* visibles par chacun. Pour monter un cluster on doit avoir les mêmes noms de disques correspondants au même ID logique de la même baie. En effet chaque noeud doit travailler sur les mêmes disques. Parfois sous Linux PowerPath perd les pédales au boot et on peut avoir la la conf suivante :

- **noeud1**

```
root@noeud1:~#  powermt display dev=emcpowerh
Pseudo name=emcpowerh
Symmetrix ID=000287750899
Logical device ID=0040
state=alive; policy=SymmOpt; priority=0; queued-IOs=0
==============================================================================
---------------- Host ---------------   - Stor -   -- I/O Path -  -- Stats ---
### HW Path                I/O Paths    Interf.   Mode    State  Q-IOs Errors
==============================================================================
   1 lpfc                     sdah      FA  7bA   active  alive      0      0
   0 lpfc                     sdk       FA 10bA   active  alive      0      0
```

- **noeud2**

```
root@noeud2:~# powermt display dev=emcpowerh
Pseudo name=emcpowerh
Symmetrix ID=000287750899
Logical device ID=0544
state=alive; policy=SymmOpt; priority=0; queued-IOs=0
==============================================================================
---------------- Host ---------------   - Stor -   -- I/O Path -  -- Stats ---
### HW Path                I/O Paths    Interf.   Mode    State  Q-IOs Errors
==============================================================================
   1 lpfc                     sdam      FA  7bA   active  alive      0      0
   0 lpfc                     sdp       FA 10bA   active  alive      0      0
```

Ce qui nous intéresse ici est le *Logical device ID*. On prend une machine de référence, ici c'est le *noeud1*. Soit pour le disque **emcpowerh** :

	noeud1	**noeud2**
Pseudo name	emcpowerh	emcpowerh
Symmetrix ID	000287750899	0040
Logical device	000287750899	0544

On voit bien que le disque **emcpowerh** n'est pas le même pour les 2 noeuds. On pourrait avoir une baie différente (c'est-à-dire un Symmetrix ID différent). Dans ce cas le *Logical device ID* suffirait quand même (à moins d'avoir le même ID sur 2 baies différentes). Voici la marche à suivre :

- On renomme d'abord l'*emcpowerh* sur le *noeud2* avec un nom libre (sinon on ne peut pas bosser dessus) :

```
root@noeud2:~# emcpadm getfree -n 5 -b emcpowerh

Next free pseudo device name(s) from emcpowerh are:

Pseudo Device Name      Major# Minor#
        emcpoweru         120    320
        emcpowerv         120    336
        emcpowerw         120    352
        emcpowery         120    384
        emcpowerz         120    400
```

```
root@noeud2:~# emcpadm rename -s emcpowerh -t emcpoweru
```

- Quel est le disque ayant l'ID 0040 (cf. noeud1)

```
root@noeud2:~# powermt display dev=all|grep -B 2 0040
Pseudo name=emcpowerm
Symmetrix ID=000287750899
Logical device ID=0040
```

- On le renomme :

```
root@noeud2:~# emcpadm rename -s emcpowerm -t emcpowerh
```

Et voila résultat :

```
root@noeud2:~# powermt display dev=emcpowerh
Pseudo name=emcpowerh
Symmetrix ID=000287750899
Logical device ID=0040
```

```
state=alive; policy=SymmOpt; priority=0; queued-IOs=0
==============================================================================
---------------- Host ---------------   - Stor -   -- I/O Path -  -- Stats ---
### HW Path                I/O Paths    Interf.   Mode    State  Q-IOs Errors
==============================================================================
   1 lpfc                     sdah      FA  7bA   active  alive      0      0
   0 lpfc                     sdk       FA 10bA   active  alive      0      0
```


Quand tout est OK ne pas oublier de faire un **powermt save** sur le noeud modifié.

Problème de cohérence chemins ⇔ devices PowerPath

Lorsqu'on utilise des commandes LVM on peut obtenir ces erreurs :

```
root@server1101582:/# vgchange -an vg_vrk1 vg_vrk2 vg_vrk2_raw
  0 logical volume(s) in volume group "vg_vrk1" now active
  Can't deactivate volume group "vg_vrk2" with 5 open logical volume(s)

Found duplicate PV ZjelLvmkIdZxspuWDvgbHtWe1Mt6dWy0: using /dev/sdq not /dev/emcpowere
Found duplicate PV iboistKzZhnuI7vamjgSCbjNx98BNLfa: using /dev/sdag not /dev/emcpowerf
Found duplicate PV 3ePFHbke6WBeWbWRjHmk0DipIbzo2IxH: using /dev/sdb not /dev/emcpowerg
Found duplicate PV SakxwlEATnYsYk6GpUhq52QCGQTzFaDW: using /dev/sdr not /dev/emcpowerb
```

De la même façon lors d'un *vgdisplay* on ne voit que les chemins et non pas les *emcpowerX*. Ce qui est incorrect car on **doit** utiliser les disques PowerPath, dans le cas contraire on risque de sérieux soucis si un chemin lâche.

```
  --- Physical volumes ---
  PV Name               /dev/sdj
  PV UUID               iboist-KzZh-nuI7-vamj-gSCb-jNx9-8BNLfa
  PV Status             allocatable
  Total PE / Free PE    871 / 0

  PV Name               /dev/sdy
  PV UUID               3ePFHb-ke6W-BeWb-WRjH-mk0D-ipIb-zo2IxH
  PV Status             allocatable
  Total PE / Free PE    871 / 0
```

En il peut s'agir soit d'une erreur de manipulation (un *pvcreate* sur un /dev/sdX) ou d'un mauvais filtre utilisé dans */etc/lvm/lvm.conf*. Par défaut le filtre ci-dessous est utilisé :

filter = [“a/.*/”]

Comme on utilise des *emcpowerX* il faut modifier le filtre. En effet la couche LVM trouve des PV identiques si elle scanne l'emcpowerX et ses *n* chemins. Et donc elle utilise le premier qu'elle trouve. On peut modifier le filtre pour prendre tout ce qui contient *emcpower*, *c0d0* (ou autre), etc :

filter = [“a/emcpower/”,“a/c0d0.*/”,“r/.*/”]

ou

filter = [“a/emcpower/”, “a/sda[1-9]$/”, “r/sd.*/”, “r/disk.*/”, “r|/dev/cdrom|”, “a/.*/”]

A adapter selon les besoins.

Ensuite un **vgscan** pour remettre d'équerre :

```
  --- Physical volumes ---
  PV Name               /dev/emcpowerf
  PV UUID               iboist-KzZh-nuI7-vamj-gSCb-jNx9-8BNLfa
  PV Status             allocatable
  Total PE / Free PE    871 / 0

  PV Name               /dev/emcpowerg
  PV UUID               3ePFHb-ke6W-BeWb-WRjH-mk0D-ipIb-zo2IxH
  PV Status             allocatable
  Total PE / Free PE    871 / 0
```


Parfois en RHEL3 le package *lvm* est trop ancien et ne permet pas de travailler avec le LVM sur des devices PowerPath ... Il faut donc mettre à jour le package.

Found Duplicate PV

Le filtre dans le fichier */etc/lvm/lvm.conf* n'est pas correct. Le système voit plusieurs PVs identiques à cause de PowerPath : un device PowerPath = *2n* devices. Pour cela on peut utiliser ce filtre (à adapter selon ses besoins) :

```
filter = [ "a/emcpower/","a/sda/","r/.*/" ]
```

On accepte tous les devices *emcpower*, *sda* et on rejette les autres. Penser à lancer un *vgscan* par la suite.

Rescanner les devices

Cartes Emulex

```
/usr/sbin/lpfc/lun_scan all
```

Ce qui revient à faire :

```
echo '- - -' > /sys/class/scsi_host/host[1,2,3,...n]/scan
```

Cartes QLogic

```
echo "scsi-qlascan" > /proc/scsi/qla2xxx/1
echo "scsi-qlascan" > /proc/scsi/qla2xxx/1
echo "scsi-qlascan" > /proc/scsi/qla2xxx/2
echo "scsi-qlascan" > /proc/scsi/qla2xxx/3
echo "scsi-qlascan" > /proc/scsi/qla2xxx/4
echo "- - -" > /sys/class/scsi_host/host1/scan
echo "- - -" > /sys/class/scsi_host/host2/scan
echo "- - -" > /sys/class/scsi_host/host3/scan
echo "- - -" > /sys/class/scsi_host/host4/scan
```

Trouver les WWN

- RHEL4

```
root@server7410521:/sys/class/scsi_host# cat host3/port_name
0x10000000c9840edb
root@server7410521:/sys/class/scsi_host# cat host4/port_name
0x10000000c9840ca5
```

- RHEL5

```
root@server9011354:/sys/class/fc_host/host1# cat port_name
0x10000000c97565ca
```

Erreur lors du renommage des disques

Un disque à priori non utilisé ne peut pas être renommé et la commande *emcpadm renamepseudo* renvoit une erreur du style *cannot renamed pseudo device currently in use*. Il suffit d'utiliser la commande **emcppurge** (sans paramètre).

Export / import de conf PowerPath

```
emcpadm export -f /tmp/emcpower.xml
emcpadm import -f /tmp/emcpower.xml
```

ou

```
emcpadm export_mappings -f /tmp/emcpower.xml
emcpadm import_mappings -f /tmp/emcpower.xml
```

⇒ évite se de retaper tous les *emcpadm rename* sur d'autres noeuds d'un cluster par ex.

Déterminer le modèle d'une carte SAN

- Si aucun outil n'est installé (hbanyware par ex.)

```
lpsci
grep -i emulex -A 4 /etc/sysconfig/hwconf|grep deviceId|sort -u
```

Puis checker sur un site comme http://www.pcidatabase.com/vendor_details.php?id=607

- Avec hbanyware :

```
/usr/sbin/hbanyware/hbacmd listhbas
/usr/sbin/hbanyware/hbacmd hbaAttributes "Port WWN" de la carte
```

Supprimer un disque SAN

Pour supprimer proprement un device SAN sur un linux :

- supprimer le disque du VG (*vgreduce*) ou du DG (si VxFs → v*xdg -g DG rmdisk DISK*)
- si PowerPath : *powermt remove* du emcpower
- si multipath : *multipath -ll* pour récupérer l'ID, puis *dmsetup remove ID*

Enfin on supprime les /dev/sdX avec :

```
echo 1 > /sys/block/sdX/device/delete
```

Montage des FS sur SAN

Rajouter *_netdev* pour les FS sur du SAN si PowerPath ne démarre pas comme prévu (si filtré dans */etc/lvm.conf*)

Faire clignoter les LEDs d'une carte Emulex

```
# hbacmd ListHBAs

Manageable HBA List

Port WWN   : 10:00:00:00:c9:4f:d2:05
Node WWN   : 20:00:00:00:c9:4f:d2:05
Fabric Name: 10:00:08:00:88:a0:92:2f
Flags      : 8000f980
Host Name  : parsl3004325
Mfg        : Emulex Corporation

Port WWN   : 10:00:00:00:c9:34:24:d9
Node WWN   : 20:00:00:00:c9:34:24:d9
Fabric Name: 10:00:08:00:88:04:2a:2f
Flags      : 8000f980
Host Name  : parsl3004325
Mfg        : Emulex Corporation
```

```
# hbacmd GetBeacon 10:00:00:00:c9:4f:d2:05
Beacon state = 'Off'

# hbacmd SetBeacon 10:00:00:00:c9:4f:d2:05 1
Beacon state successfully set to ON.

# hbacmd GetBeacon 10:00:00:00:c9:4f:d2:05
Beacon state = 'On'
```

forcerescan-scsi-bus.sh / read-only

- Issue : When I execute rescan-scsi-bus.sh –forcerescan on a system with SAN LUNs, all filesystems become read-only.
- Resolution :

Use –hosts paramater to ignore local adapters.

```
host0: local device
host1: local device
host2: SAN device

rescan-scsi-bus.sh --forcerescan --hosts=2
```

From:
https://unix.ndlp.info/ - **Where there is a shell, there is a way**

Permanent link:
https://unix.ndlp.info/doku.php/informatique:nix:linux:linux_san

Last update: **2018/01/11 07:43**