

Nouveautés LVM en AIX 5.1

Alexandre BLANCKE / Jean-François LE MOAN

Introduction

L'AIX 5.1 comporte de nombreuses améliorations par rapport aux versions précédentes, y compris dans le domaine du LVM (*Logical Volume Manager*), améliorations permettant une administration à la fois plus simple et plus détaillée des volumes logiques.

Migration automatique de disques défaillants

En AIX 5.1, le LVM permet la mise en œuvre de la migration automatique de disques défaillants (*HOT SPARE*).

Cette facilité ne s'applique que sur les volumes logiques "*mirrored*".

Mise en œuvre

- Il faut d'abord définir l'un des disques du "*volume group*" comme disque "**HOT SPARE**". Ce disque doit être vide et avoir une taille au moins équivalente au plus petit des disques membres du groupe de volumes.

```
# chpv -hy hdisk4
```

Le disque "hdisk4" devient ainsi le disque "hotspare".

```
# lsvg -p hotsparevg  
  
hotsparevg:  
PV_NAME      PV STATE      TOTAL PPs   FREE PPs   FREE DISTRIBUTION  
hdisk4       active/hotspare 268         268        54..54..53..53..54  
hdisk2       active         268         266        54..52..53..53..54
```

Infos Produits

```
# lspv hdisk4

PHYSICAL VOLUME:      hdisk4          VOLUME GROUP:      hotsparevg
PV IDENTIFIER:       0000773709982e48  VG IDENTIFIER      0020923300005700000000f03918bf80
PV STATE:            active
STALE PARTITIONS:    0                ALLOCATABLE:       no
PP SIZE:             8 megabyte(s)    LOGICAL VOLUMES:   0
TOTAL PPs:          268 (2144 megabytes)  VG DESCRIPTORS:    1
FREE PPs:           268 (2144 megabytes)  HOT SPARE:         yes
USED PPs:           0 (0 megabytes)
FREE DISTRIBUTION:   54..54..53..53..54
USED DISTRIBUTION:   00..00..00..00..00
```

- Indiquer que le Groupe de Volumes doit fonctionner en mode “migration automatique” :

```
# chvg -hY -sy hotsparevg
```

Et le Groupe de Volumes devient également “hotspare” :

```
# lsvg hotsparevg

VOLUME GROUP:      hotsparevg    VG IDENTIFIER:      0020923300005700000000f03918bf80
VG STATE:          active        PP SIZE:            8 megabyte(s)
VG PERMISSION:     read/write    TOTAL PPs:          804 (6432 megabytes)
MAX LVs:           256          FREE PPs:           801 (6408 megabytes)
LVs:               2            USED PPs:           3 (24 megabytes)
OPEN LVs:          2            QUORUM:             2
TOTAL PVs:         3            VG DESCRIPTORS:     3
STALE PVs:         0            STALE PPs:          0
ACTIVE PVs:        3            AUTO ON:            yes
MAX PPs per PV:   1016          MAX PVs:            32
LTG size:          128 kilobyte(s)  AUTO SYNC:          yes
HOT SPARE:         yes (one to many)
```

Si l'un des disques membres du Groupe de Volumes devient défaillant, il sera remplacé par le disque “hotspare”.

Remarque :

Il est **IMPORTANT** d'installer le dernier niveau du *fileset* “**bos.rte.lvm**”.

Désactivation d'un espace de pagination

Depuis la version AIX 5.1, la commande “**swapoff**” permet de désactiver un espace de pagination.

Afin que cette commande puisse s'exécuter correctement, il faut :

- Qu'il reste au moins un espace de pagination actif sur le système.
- Que cet espace de pagination soit suffisant pour accueillir les données contenues dans l'espace de pagination à désactiver.

Réduction de la taille d'un espace de pagination

Avant la version AIX 5.1, la réduction de la place utilisée par l'espace de pagination nécessitait de nombreuses manipulations incluant des redémarrages du système.

Depuis la version AIX 5.1, et notamment grâce à la possibilité de désactiver un espace de pagination sans qu'un redémarrage du système ne soit nécessaire, il est possible de réduire dynamiquement un espace de pagination aux conditions suivantes :

- Le Groupe de Volumes "ROOTVG" doit disposer de suffisamment d'espace libre pour créer (en plus de l'espace de pagination existant) un espace temporaire d'une taille équivalente à la taille ciblée.
Par exemple, si l'on veut réduire "hd6" de 512 Mb à 496 Mb, il faut qu'il y ait 496 Mb libres dans "ROOTVG".
- Dans le cas du "HD6", il doit être possible d'exécuter la commande "bosboot".
En effet, pendant l'opération, la commande crée une pagination temporaire et remplace le "hd6" par cet espace temporaire dans l'image de "boot" afin qu'un redémarrage soit toujours possible.

Exemple

Avant l'opération, le "hd6" fait 512 Mb :

```
# lspvs -a
Page Space Physical Volume Volume Group Size %Used Active Auto Type
hd6 hdisk0 rootvg 512MB 1 yes yes lv
```

On réduit le *paging* de 2 partitions (le "PPSIZE" est de 16 Mb) :

```
# shrinkps hd6 2
shrinkps: Temporary paging space paging00 created.
shrinkps: New boot image created with temporary paging space.
shrinkps: Paging space hd6 removed.
shrinkps: Paging space hd6 recreated with new size.
shrinkps: New boot image created with resized paging space.
shrinkps: Resized and original paging space characteristics differ,
check the lslv command output
```

Une fois l'opération terminée, la taille du "hd6" est bien réduite de 2 partitions :

```
# lspvs -a
Page Space Physical Volume Volume Group Size %Used Active Auto Type
hd6 hdisk0 rootvg 480MB 1 yes yes lv
```

Analyse des performances interactives

Dans les versions d'AIX précédentes, l'analyse détaillée de l'activité Entrée/Sortie (I/O) sur les Volumes Logiques nécessitait l'utilisation d'outils de traçage non-interactifs, comme "filemon".

L'analyse interactive s'arrêtant au niveau du disque physique, c'est-à-dire au niveau du "device driver", interrogé par la commande "iostat" par exemple.

La commande "lvmstat"

Depuis la version AIX 5.1, la commande "lvmstat" permet d'interroger le LVM sur les statistiques, au niveau du "Groupe de Volumes" ou du "Volume Logique".

Ceci permet, dans une analyse, d'approfondir les résultats fournis par la commande "iostat", afin de déterminer rapidement quels sont les Volumes Logiques les plus utilisés et, dans ces Volumes Logiques, les partitions logiques les plus utilisées.

Par défaut, ces statistiques ne sont pas collectées. Avant toute analyse, il faut donc :

- Commencer par mettre en œuvre ces statistiques (option "-e") sur le Groupe de Volumes (option "-v") ou le Volume Logique (option "-l").

Exemple :

```
# lvmstat -v rootvg -e ; lvmstat -l hd9var -e
```

- Ensuite, on peut utiliser "lvmstat" pour obtenir l'activité en direct.
Par exemple : toutes les secondes pendant dix secondes.

```
# lvmstat -s -v rootvg 1 10

Logical Volume      iocnt      Kb_read    Kb_wrtn    Kbps
hd2                  44          40         156        0.00
hd6                  39          0          204        0.00
hd9var               36          0          204        0.00
hd8                  36          0          144        0.00
hd4                  34          40         156        0.00

..... Période d'inactivité

hd9var               177         708         0          475.17
hd6                   22          68          20          59.06
hd2                    7           28           0          18.79

hd9var               662        2648         0          1765.33
hd9var               660        2640         0          1748.34
hd9var               645        2580         0          1686.27
hd9var               658        2632         0          1462.22
hd9var               661        2644         0          1774.50
hd9var               663        2652         0          1768.00
```

La commande “lvmstat” commence toujours par un résumé de l’activité, depuis la mise en œuvre (“-e”) ou la remise à zéro (“-C”) des statistiques LVM. Puis l’affichage des statistiques ne se fait qu’en cas d’activité constatée, comme ici sur “hd9var”. En cas d’inactivité, “lvmstat” affiche des “.” à chaque itération.

On peut également avoir un affichage sous forme de colonnes (“-F”), afin de pouvoir traiter les résultats ultérieurement.

Exemple :

```
# lvmstat -s -l hd9var -F 1 10
1:1:5118:20380:116:0.04
2:1:4193:16384:611:0.03

1:1:662:2648:0:1777.18

1:1:649:2596:0:1719.21

1:1:662:2648:0:1777.18

1:1:663:2652:0:1768.00

1:1:462:1848:0:1223.84
2:1:197:788:0:521.85

2:1:661:2644:0:1762.67

2:1:375:1472:52:996.08
...
```

La migration de partitions

Grace à l’utilisation de la commande “lvmstat” vue précédemment, il est aisé d’identifier les partitions logiques les plus utilisées.

Si vous constatez que plusieurs des partitions fortement utilisées se trouvent sur le même disque physique, il peut être intéressant d’en déplacer une ou plusieurs vers des disques moins utilisés. Ceci peut se faire grâce à la commande “migratelp”.

Exemple

Avant la migration de partition :

```
# lspv -p hdisk0
PP RANGE   STATE   REGION   LV NAME   TYPE      MOUNT POINT
1-1        used   outer edge   hd5       boot      N/A
2-2        used   outer edge   lv00      sysdump   N/A
3-24       free   outer edge
25-25      used   outer edge   hd10opt   jfs       /opt
26-27      used   outer edge   hd3       jfs       /tmp
28-29      free   outer middle
30-54      used   outer middle   hd6       paging    N/A
55-59      used   center      hd6       paging    N/A
60-60      used   center      hd8       jfslog    N/A
61-61      used   center      hd4       jfs       /
62-80      used   center      hd2       jfs       /usr
81-81      used   inner middle   hd2       jfs       /usr
82-83      used   inner middle   hd9var    jfs       /var
84-84      used   inner middle   hd1       jfs       /home
85-92      used   inner middle   hd2       jfs       /usr
93-107     free   inner middle
108-134    free   inner edge
```

Infos Produits

Nous constatons que la partition 3 est libre sur le disque 0

```
# migratelp hd9var/1 hdisk0/3
```

Après la migration :

```
# lspv -p hdisk0
hdisk0:
PP RANGE    STATE    REGION    LV NAME    TYPE    MOUNT POINT
1-1         used    outer edge    hd5        boot    N/A
2-2         used    outer edge    lv00       sysdump N/A
3-3         used    outer edge    hd9var     jfs     /var
4-24        free    outer edge
25-25       used    outer edge    hd10opt    jfs     /opt
26-27       used    outer edge    hd3        jfs     /tmp
28-29       free    outer middle
30-54       used    outer middle    hd6        paging  N/A
55-59       used    center        hd6        paging  N/A
60-60       used    center        hd8        jfslog  N/A
61-61       used    center        hd4        jfs     /
62-80       used    center        hd2        jfs     /usr
81-81       used    inner middle    hd2        jfs     /usr
82-82       free    inner middle
83-83       used    inner middle    hd9var     jfs     /var
84-84       used    inner middle    hd1        jfs     /home
85-92       used    inner middle    hd2        jfs     /usr
93-107      free    inner middle
108-134     free    inner edge
```

En cas de **mirroring**, vous pourrez préciser quelle partition doit être migrée. Par exemple, si vous voulez migrer la partition 2 (voir la commande “`lslv -m hd9var`”) :

```
# migratelp hd9var/1/2 hdisk0/3
```

La commande “recreatevg”

La commande “**recreatevg**” (qui existait déjà en AIX 433) permet, lorsque l’on fait de la copie disque à disque, de régénérer une structure valide de Groupe de Volumes sur le disque cible avec des noms de volumes logiques et de système de fichiers différents.

```
# recreatevg -y testvg -Y newlv hdisk1
```

Cette commande créera un Groupe de Volumes “**testvg**” dont tous les volumes logiques auront comme préfixe “**newlv**” à partir du disque “**hdisk1**” qui pourrait être la copie d’un autre disque contenant un Groupe de Volumes existant.