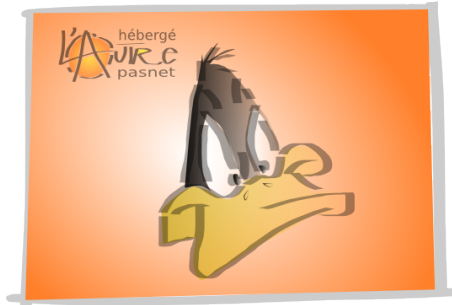


<http://daffyduke.lautre.net/spip/?Tips-Tricks-Sun-Clusters>



# Tips & Tricks Sun Clusters

- 1- Blog-Notes - Au boulot -

Date de mise en ligne : mardi 17 janvier 2006

---

Copyright © L'Imp'Rock Scénette (by @\_daffyduke\_) - Tous droits réservés

---

# Cas concret : un disque défaillant

Etat : Le disque défaillant (c2t0d0), après unplug/replug, reboot sans sds + drvconfig+disks+devlinks => a été revu OK par "iostat -En" après reboot.

Mais l'état des metadevices était pitoyable.

disque c2t0d0 malade /pci@6,4000/scsi@4,1/sd@0,0 (sd75) corrupt label - wrong magic number

On l'a dépluggé, puis repluggé + drvconfig + disks + devlinks, il faudrait maintenant rebooter.

```

cmapqlf01:root} metastat d50 d50: Mirror      Submirror 0: d51      State: Needs maintenance      Submirror
1: d52      State: Okay      Pass: 1      Read option: roundrobin (default)      Write option:
parallel (default)      Size: 308826 blocks d51: Submirror of d50      State: Needs maintenance      Invoke:
metareplace d50 c2t0d0s0 Size: 308826 blocks      Stripe 0:      Device      Start Block Dbase
State      Hot Spare      c2t0d0s0      0      No      Maintenance      d52: Submirror of d50
State: Okay      Size: 308826 blocks      Stripe 0:      Device      Start Block Dbase
State      Hot Spare      c2t1d0s0      0      No      Okay      cmapqlf01:root}
metadb -i      flags      first blk      block count      a m c lu o      16      1034
/dev/dsk/clt0d0s7      a c lu o      1050      1034      /dev/dsk/clt0d0s7      a
c lu o      2084      1034      /dev/dsk/clt0d0s7      a c lu o      16
1034      /dev/dsk/clt1d0s7      a c lu o      1050      1034      /dev/dsk/clt1d0s7
a c lu o      2084      1034      /dev/dsk/clt1d0s7      M c      unknown
unknown      /dev/dsk/c2t0d0s7      M c      unknown      unknown
/dev/dsk/c2t0d0s7      M c      unknown      unknown      /dev/dsk/c2t0d0s7 o - replica
active prior to last mddb configuration change u - replica is up to date l - locator for this replica was
read successfully c - replica's location was in /etc/lvm/mddb.cf p - replica's location was patched in
kernel m - replica is master, this is replica selected as input W - replica has device write errors a -
replica is active, commits are occurring to this replica M - replica had problem with master blocks D -
replica had problem with data blocks F - replica had format problems S - replica is too small to hold
current data base R - replica had device read errors cmapqlf01:root} metadb -d /dev/dsk/c2t0d0s7
cmapqlf01:root} metadb      flags      first blk      block count      a m c lu o      16
1034      /dev/dsk/clt0d0s7      a c lu o      1050      1034
/dev/dsk/clt0d0s7      a c lu o      2084      1034      /dev/dsk/clt0d0s7      a c lu o
16      1034      /dev/dsk/clt1d0s7      a c lu o      1050      1034
/dev/dsk/clt1d0s7      a c lu o      2084      1034      /dev/dsk/clt1d0s7
cmapqlf01:root} metastat d20      d20: Mirror      Submirror 0: d21      State: Needs maintenance
Submirror 1: d22      State: Needs maintenance      Pass: 1      Read option: roundrobin (default)      Write
option: parallel (default)      Size: 1231713 blocks d21: Submirror of d20      State: Needs maintenance
Invoke: metareplace d20 clt0d0s4 Size: 1231713 blocks      Stripe 0:      Device      Start Block
Dbase State      Hot Spare      clt0d0s4      0      No      Maintenance      d22: Submirror
of d20      State: Needs maintenance      Invoke: after replacing "Maintenance" components:
metareplace d20 clt1d0s4 Size: 1231713 blocks      Stripe 0:      Device      Start Block Dbase
State      Hot Spare      clt1d0s4      0      No      Last Erred      cmapqlf01:root}
metareplace -e d20 clt0d0s4 metareplace: cmapqlf01: clt0d0s4: is mounted on /var cmapqlf01:root}
metareplace -e d20 clt1d0s4 metareplace: cmapqlf01: d20: clt1d0s4: component in invalid state to replace -
Replace "Maintenance" components first cmapqlf01:root} metadetach -f d20 d22 metadetach: cmapqlf01: d20:
operation would result in no readable submirrors cmapqlf01:root} metadetach -f d20 d21 d20: submirror d21
is detached cmapqlf01:root} metaclear d20 metaclear: cmapqlf01: d20: attempted to clear mirror with

```

## Tips & Tricks Sun Clusters

```
submirror(s) in invalid state  cmapqlf01:root} metastat d20 d20: Mirror      Submirror 1: d22      State:
Needs maintenance      Pass: 1      Read option: roundrobin (default)      Write option: parallel (default)
Size: 1231713 blocks  d22: Submirror of d20      State: Needs maintenance      Invoke: after replacing
"Maintenance" components:      metareplace d20 clt1d0s4 Size: 1231713 blocks      Stripe 0:
Device      Start Block  Dbase State      Hot Spare      clt1d0s4      0      No
Last Erred      destruction  cmapqlf01:root} metaclear -f d20 d20: Mirror is cleared cmapqlf01:root}
metastat d20 metastat: cmapqlf01: d20: unit not set up      cmapqlf01:root} metastat d21 d21: Concat/Stripe
Size: 1231713 blocks      Stripe 0:      Device      Start Block  Dbase      clt0d0s4
0      No      cmapqlf01:root} metastat d22 d22: Concat/Stripe      Size: 1231713 blocks
Stripe 0:      Device      Start Block  Dbase      clt1d0s4      0      No
cmapqlf01:root} metainit d20 -m d21 d20: Mirror is setup cmapqlf01:root} metastat d20 d20: Mirror
Submirror 0: d21      State: Okay      Pass: 1      Read option: roundrobin (default)      Write
option: parallel (default)      Size: 1231713 blocks  d21: Submirror of d20      State: Okay
Size: 1231713 blocks      Stripe 0:      Device      Start Block  Dbase State      Hot Spare
clt0d0s4      0      No      Okay      cmapqlf01:root} df -k Filesystem
kbytes  used  avail capacity  Mounted on /dev/dsk/clt0d0s0      771110  352347  364786  50%  /
/dev/dsk/clt0d0s3      1018382  178193  779087  19%  /usr /proc      0      0      0
0%  /proc fd      0      0      0      0%  /dev/fd mnttab      0
0      0      0%  /etc/mnttab /dev/dsk/clt0d0s4      578351  365642  154874  71%  /var swap
4653176      24 4653152      1%  /var/run swap      4653160      8 4653152      1%  /tmp
/dev/dsk/clt0d0s5      4492386  1979687  2467776  45%  /sybase  cmapqlf01:root} metastat d50      d50:
Mirror      Submirror 0: d51      State: Needs maintenance      Submirror 1: d52      State: Okay
Pass: 1      Read option: roundrobin (default)      Write option: parallel (default)      Size: 308826
blocks  d51: Submirror of d50      State: Needs maintenance      Invoke: metareplace d50 c2t0d0s0 Size:
308826 blocks      Stripe 0:      Device      Start Block  Dbase State      Hot Spare
c2t0d0s0      0      No      Maintenance  d52: Submirror of d50      State: Okay
Size: 308826 blocks      Stripe 0:      Device      Start Block  Dbase State      Hot Spare
c2t1d0s0      0      No      Okay      cmapqlf01:root} metadetach -f d50 d51 d50:
submirror d51 is detached cmapqlf01:root} metaclear d50 d50: Mirror is cleared cmapqlf01:root} metastat d51
d51: Concat/Stripe      Size: 308826 blocks      Stripe 0:      Device      Start Block  Dbase
c2t0d0s0      0      No      cmapqlf01:root} metastat d52 d52: Concat/Stripe      Size:
308826 blocks      Stripe 0:      Device      Start Block  Dbase      c2t1d0s0
0      No
```

A l'ioostat -En, on voit que le disque c2t0d0 est OK

Mais au format, il demande à être labellé. => il faut y virer d'abord tous les metadevices

```
cmapqlf01:root} metaclear -f d51 d51: Concat/Stripe is cleared  cmapqlf01:root} metastat d60 d60: Mirror
Submirror 0: d61      State: Needs maintenance      Submirror 1: d62      State: Okay
Pass: 1      Read option: roundrobin (default)      Write option: parallel (default)      Size: 308826 blocks
d61: Submirror of d60      State: Needs maintenance      Invoke: metareplace d60 c2t0d0s1 Size: 308826 blocks
Stripe 0:      Device      Start Block  Dbase State      Hot Spare      c2t0d0s1
0      No      Maintenance  d62: Submirror of d60      State: Okay      Size: 308826 blocks
Stripe 0:      Device      Start Block  Dbase State      Hot Spare      c2t1d0s1
0      No      Okay      cmapqlf01:root} metadetach -f d60 d61 d60: submirror d61 is detached
cmapqlf01:root} metaclear d60 d60: Mirror is cleared cmapqlf01:root} metaclear -f d61 d61: Concat/Stripe is
cleared cmapqlf01:root} metastat d62 d62: Concat/Stripe      Size: 308826 blocks      Stripe 0:      Device
Start Block  Dbase      c2t1d0s1      0      No
```

Idem : d71 need maintenance

```
cmapqlf01:root} metadetach -f d70 d71 d70: submirror d71 is detached  cmapqlf01:root} metaclear d70 d70:
Mirror is cleared  cmapqlf01:root} metaclear -f d71 d71: Concat/Stripe is cleared  cmapqlf01:root} metastat
```

```
d72 d72: Concat/Stripe      Size: 2050461 blocks      Stripe 0:      Device      Start Block  Dbase
      c2t1d0s3                0      No      Idem: d81 need maint cmapqlf01:root} metadetach -f d80 d81
d80: submirror d81 is detached cmapqlf01:root} metaclear d80 d80: Mirror is cleared cmapqlf01:root}
metaclear -f d81 d81: Concat/Stripe is cleared      Idem d90 need maint cmapqlf01:root} metadetach -f d90 d91
d90: submirror d91 is detached cmapqlf01:root} metaclear d90 d90: Mirror is cleared cmapqlf01:root}
metaclear -f d91 d91: Concat/Stripe is cleared cmapqlf01:root} metastat d92 d92: Concat/Stripe
```

# Reformatage du disque c2t0d0

```
cmapqlf01:root} metainit d51 1 1 /dev/dsk/c2t0d0s0 d51: Concat/Stripe is setup cmapqlf01:root} metainit d61
1 1 /dev/dsk/c2t0d0s1 d61: Concat/Stripe is setup cmapqlf01:root} metainit d71 1 1 /dev/dsk/c2t0d0s3 d71:
Concat/Stripe is setup cmapqlf01:root} metainit d81 1 1 /dev/dsk/c2t0d0s4 d81: Concat/Stripe is setup
cmapqlf01:root} metainit d91 1 1 /dev/dsk/c2t0d0s5 d91: Concat/Stripe is setup      metinit d50 -m d51
metattach d50 d52
```

idem d60, d70, d80, d90 + attente synchro

refaire les newfs des d70 et d80, peut-être malmenés ?

Modif vfstab pour ré-intégrer tous les miroirs sauf pour / et swap

On prévoit de rebooter....

Avant : mise aux droits sybase des devices correspondants aux miroirs + aux sub-miroirs des  
raw qu'utilisera sybase

```
cmapqlf01:root} ls -l /dev/ASE* total 12 lrwxrwxrwx  1 sybase  sybase      16 Sep 30 16:19 log_rdev01 ->
/dev/md/rdisk/d90 lrwxrwxrwx  1 root      other      18 Oct  7 09:18 master -> /dev/rdisk/c2t0d0s0
lrwxrwxrwx  1 sybase  sybase      16 Sep 30 16:08 master_1 -> /dev/md/rdisk/d50 lrwxrwxrwx  1 root
other      18 Oct  7 09:19 subsystem -> /dev/rdisk/c2t0d0s1 lrwxrwxrwx  1 sybase  sybase      16 Sep
30 16:17 subsystem_1 -> /dev/md/rdisk/d60 lrwxrwxrwx  1 sybase  sybase      17 Sep 30 16:29 user_rdev01
-> /dev/md/rdisk/d110  cmapqlf01:root} metastat -p d50 d50 -m d51 d52 1 d51 1 1 c2t0d0s0 d52 1 1 c2t1d0s0
cmapqlf01:root} ls -l /dev/rdisk/c2t0d0s0 lrwxrwxrwx  1 root      root      46 Aug 20 18:56
/dev/rdisk/c2t0d0s0 -> ../../devices/pci@6,4000/scsi@4,1/sd@0,0:a,raw  cmapqlf01:root} ls -l
/devices/pci@6,4000/scsi@4,1/sd@0,0:a,raw crw-r-----  1 sybase  sybase      32,600 Oct  7 09:21
/devices/pci@6,4000/scsi@4,1/sd@0,0:a,raw  cmapqlf01:root} ls -l /dev/rdisk/c2t1d0s0 lrwxrwxrwx  1 root
root      46 Aug 20 18:56 /dev/rdisk/c2t1d0s0 -> ../../devices/pci@6,4000/scsi@4,1/sd@1,0:a,raw
cmapqlf01:root} ls -l /devices/pci@6,4000/scsi@4,1/sd@1,0:a,raw crw-r-----  1 root      sys      32,608 Aug
20 18:56 /devices/pci@6,4000/scsi@4,1/sd@1,0:a,raw  cmapqlf01:root} chown sybase:sybase
/devices/pci@6,4000/scsi@4,1/sd@1,0:a,raw  cmapqlf01:root} ls -l /devices/pci@6,4000/scsi@4,1/sd@1,0:a,raw
crw-r-----  1 sybase  sybase      32,608 Aug 20 18:56 /devices/pci@6,4000/scsi@4,1/sd@1,0:a,raw
cmapqlf01:root} metastat -p d60 d60 -m d61 d62 1 d61 1 1 c2t0d0s1 d62 1 1 c2t1d0s1  cmapqlf01:root} ls -l
/dev/rdisk/c2t0d0s1 /dev/rdisk/c2t1d0s1 lrwxrwxrwx  1 root      root      46 Aug 20 18:56
/dev/rdisk/c2t0d0s1 -> ../../devices/pci@6,4000/scsi@4,1/sd@0,0:b,raw lrwxrwxrwx  1 root      root
46 Aug 20 18:56 /dev/rdisk/c2t1d0s1 -> ../../devices/pci@6,4000/scsi@4,1/sd@1,0:b,raw  cmapqlf01:root} ls -l
/devices/pci@6,4000/scsi@4,1/sd@0,0:b,raw /devices/pci@6,4000/scsi@4,1/sd@1,0:b,raw  crw-r-----  1
sybase  sybase      32,601 Aug 20 18:56 /devices/pci@6,4000/scsi@4,1/sd@0,0:b,raw crw-r-----  1 root      sys
32,609 Aug 20 18:56 /devices/pci@6,4000/scsi@4,1/sd@1,0:b,raw  cmapqlf01:root} chown sybase:sybase
/devices/pci@6,4000/scsi@4,1/sd@1,0:b,raw  cmapqlf01:root} metastat -p d90 d90 -m d91 d92 1 d91 1 1 c2t0d0s5
d92 1 1 c2t1d0s5  cmapqlf01:root} ls -l /dev/rdisk/c2t0d0s5 /dev/rdisk/c2t1d0s5 lrwxrwxrwx  1 root      root
46 Aug 20 18:56 /dev/rdisk/c2t0d0s5 -> ../../devices/pci@6,4000/scsi@4,1/sd@0,0:f,raw lrwxrwxrwx  1
root      root      46 Aug 20 18:56 /dev/rdisk/c2t1d0s5 -> ../../devices/pci@6,4000/scsi@4,1/sd@1,0:f,raw
cmapqlf01:root} ls -l /devices/pci@6,4000/scsi@4,1/sd@0,0:f,raw /devices/pci@6,4000/scsi@4,1/sd@1,0:f,raw
crw-r-----  1 root      sys      32,605 Aug 20 18:56 /devices/pci@6,4000/scsi@4,1/sd@0,0:f,raw crw-r-----
```

```

1 root      sys      32,613 Aug 20 18:56 /devices/pci@6,4000/scsi@4,1/sd@1,0:f,raw  cmapqlf01:root} chown
sybase:sybase /devices/pci@6,4000/scsi@4,1/sd@0,0:f,raw /devices/pci@6,4000/scsi@4,1/sd@1,0:f,raw
cmapqlf01:root} ls -l /devices/pci@6,4000/scsi@4,1/sd@0,0:f,raw /devices/pci@6,4000/scsi@4,1/sd@1,0:f,raw
crw-r-----  1 sybase  sybase    32,605 Aug 20 18:56 /devices/pci@6,4000/scsi@4,1/sd@0,0:f,raw crw-r-----
1 sybase  sybase    32,613 Aug 20 18:56 /devices/pci@6,4000/scsi@4,1/sd@1,0:f,raw  cmapqlf01:root} metastat
-p d110 d110 -m d111 d112 1 d111 1 1 c3t0d0s1 d112 1 1 c3t1d0s1  cmapqlf01:root} ls -l /dev/rdisk/c3t0d0s1
/dev/rdisk/c3t1d0s1 lrwxrwxrwx  1 root      other          45 Sep 29 17:59 /dev/rdisk/c3t0d0s1 ->
../../../../devices/pci@1f,4000/scsi@3/sd@0,0:b,raw lrwxrwxrwx  1 root      other          45 Sep 29 17:59
/dev/rdisk/c3t1d0s1 -> ../../../../../../devices/pci@1f,4000/scsi@3/sd@1,0:b,raw  cmapqlf01:root} ls -l
/devices/pci@1f,4000/scsi@3/sd@0,0:b,raw /devices/pci@1f,4000/scsi@3/sd@1,0:b,raw crw-r-----  1 root
sys      32,  1 Sep 29 17:59 /devices/pci@1f,4000/scsi@3/sd@0,0:b,raw crw-r-----  1 root  sys      32,
 9 Sep 29 17:59 /devices/pci@1f,4000/scsi@3/sd@1,0:b,raw  cmapqlf01:root} chown sybase:sybase
/devices/pci@1f,4000/scsi@3/sd@0,0:b,raw /devices/pci@1f,4000/scsi@3/sd@1,0:b,raw  cmapqlf01:root}
cmapqlf01:root} ls -l /dev/md/rdisk/d50 lrwxrwxrwx  1 root      other          37 Sep 30 11:01
/dev/md/rdisk/d50 -> ../../../../../../devices/pseudo/md@0:0,50,raw  cmapqlf01:root} ls -l
/devices/pseudo/md@0:0,50,raw crw-r-----  1 sybase  sybase    85, 50 Oct  7 09:16
/devices/pseudo/md@0:0,50,raw  cmapqlf01:root} ls -l /dev/md/rdisk/d60 lrwxrwxrwx  1 root      other
37 Sep 30 11:01 /dev/md/rdisk/d60 -> ../../../../../../devices/pseudo/md@0:0,60,raw  cmapqlf01:root} ls -l
/devices/pseudo/md@0:0,60,raw crw-r-----  1 sybase  sybase    85, 60 Sep 30 11:01
/devices/pseudo/md@0:0,60,raw  cmapqlf01:root} ls -l /dev/md/rdisk/d90 lrwxrwxrwx  1 root      other
37 Sep 30 11:01 /dev/md/rdisk/d90 -> ../../../../../../devices/pseudo/md@0:0,90,raw  cmapqlf01:root} ls -l
/devices/pseudo/md@0:0,90,raw crw-r-----  1 sybase  sybase    85, 90 Sep 30 11:01
/devices/pseudo/md@0:0,90,raw  cmapqlf01:root} ls -l /dev/md/rdisk/d110 lrwxrwxrwx  1 root      other
38 Sep 30 11:01 /dev/md/rdisk/d110 -> ../../../../../../devices/pseudo/md@0:0,110,raw  cmapqlf01:root} ls -l
/devices/pseudo/md@0:0,110,raw crw-r-----  1 sybase  sybase    85,110 Sep 30 11:01
/devices/pseudo/md@0:0,110,raw

```

=> tout est bien à sybase !

Et rétablissement des bons liens vers les miroirs :

```

cmapqlf01:root} ls -l /dev/ASE01 total 12 lrwxrwxrwx  1 sybase  sybase    16 Sep 30 16:19 log_rdev01
-> /dev/md/rdisk/d90 lrwxrwxrwx  1 sybase  sybase    16 Sep 30 16:08 master -> /dev/md/rdisk/d50
lrwxrwxrwx  1 root      other          18 Oct  7 09:18 old_master -> /dev/rdisk/c2t0d0s0 lrwxrwxrwx  1 root
  other          18 Oct  7 09:19 old_sybsystem -> /dev/rdisk/c2t0d0s1 lrwxrwxrwx  1 sybase  sybase    16
Sep 30 16:17 sybsystem -> /dev/md/rdisk/d60 lrwxrwxrwx  1 sybase  sybase    17 Sep 30 16:29 user_rdev01
-> /dev/md/rdisk/d110

```

## Ajout d'un replica

```

cmapqlf01:root} metadb      flags      first blk      block count      a m pc luo      16
 1034      /dev/dsk/clt0d0s7      a pc luo      1050      1034
/dev/dsk/clt0d0s7      a pc luo      2084      1034      /dev/dsk/clt0d0s7      a pc luo
 16      1034      /dev/dsk/clt1d0s7      a pc luo      1050      1034
  /dev/dsk/clt1d0s7      a pc luo      2084      1034      /dev/dsk/clt1d0s7
cmapqlf01:root} metadb -a /dev/dsk/c2t1d0s7  cmapqlf01:root} metadb      flags      first blk
block count      a m pc luo      16      1034      /dev/dsk/clt0d0s7      a pc luo
 1050      1034      /dev/dsk/clt0d0s7      a pc luo      2084      1034
/dev/dsk/clt0d0s7      a pc luo      16      1034      /dev/dsk/clt1d0s7      a pc luo
 1050      1034      /dev/dsk/clt1d0s7      a pc luo      2084      1034
  /dev/dsk/clt1d0s7      a u      16      1034      /dev/dsk/c2t1d0s7

```

reboot, donc....

Avec tous les miroirs sauf / (root), c'est OK.

Si on fait juste l'échange des # dans vfstab pour / (d40 au lieu de c1t0d0s0 ) => alors ça marche pas au boot (et il faut à nouveau booter cdrom pour inverser la vapeur dans vfstab).

Après un nouveau reboot sur c1t0d0s0, on corrige :

```
metadetach d40 d42 metaclear d40
```

*(d40 disparaît, reste les deux sous-miroirs)*

```
metainit d40 -m d41 metaroot d40 lockfs -fa reboot
```

Après reboot (ok) : `metattach d40 d42` (*synchro*)

Il faut refaire un essai de boot sur le miroir....