

Reemplazo de discos en caso de degradación de RAID ZFS

En primer lugar vamos a ver el estado del pool

```
# zpool status
pool: rpool
state: DEGRADED
status: One or more devices could not be used because the label is missing or
invalid. Sufficient replicas exist for the pool to continue
functioning in a degraded state.
action: Replace the device using 'zpool replace'.
see: http://zfsonlinux.org/msg/ZFS-8000-4J
scan: scrub repaired 384K in 1070h3m with 0 errors on Tue Feb 23 14:27:56 2021
config:

NAME                STATE      READ WRITE CKSUM
rpool                DEGRADED   0     0     0
  mirror-0           ONLINE     0     0     0
    wwn-0x50014ee264214812-part1 ONLINE     0     0     2
    wwn-0x50014ee20e5acbbe-part1 ONLINE     0     0     4
  mirror-1           DEGRADED   0     0     0
    wwn-0x50014ee261dfc7ff-part1 ONLINE     0     0     0
    wwn-0x50014ee2b7358a0a-part1 UNAVAIL    0     0     0

errors: No known data errors
```

Como podemos ver en este caso la partición `wwn-0x50014ee2b7358a0a-part1` está dando problemas. Si hacemos un `ls -lah /dev/disk/by-id` podemos ver que el identificador `wwn` se corresponde con el disco.

Preparar el disco para sacarlo

Una vez identificado el disco, nos tenemos que asegurar de que las particiones no están actualmente montadas o forman parte de un RAID.

Para ver si montadas puedes usar `df` o `mount`.

Para ver si forma parte de un raid por software podemos usar "`fdisk -l /dev/(id del disco)`" para ver las particiones. En caso de que nos encontremos un tipo de partición "Linux RAID" o "Solaris /usr & Apple ZFS" tendremos que tenerlo en cuenta a la hora de reemplazar el disco ya que tendremos que marcar esta partición como fallida o "offline".

En el caso de zfs, podría darse el caso de tener 2 o mas pools como en este caso:

```
# zpool status
  pool: bpool
  state: ONLINE
status: Some supported features are not enabled on the pool. The pool can
still be used, but some features are unavailable.
action: Enable all features using 'zpool upgrade'. Once this is done,
the pool may no longer be accessible by software that does not support
the features. See zpool-features(5) for details.
  scan: scrub repaired 0B in 0 days 00:00:01 with 0 errors on Sun Mar 14 00:24:03 2021
config:

NAME                STATE      READ WRITE CKSUM
bpool                ONLINE      0    0    0
  mirror-0
  wwn-0x50014ee20cac5e8c-part2 ONLINE      0    0    0
  wwn-0x50014ee20cac9ddc-part2 ONLINE      0    0    0
  mirror-1
  wwn-0x50014ee2b9051b25-part2 ONLINE      0    0    0
  wwn-0x50014ee262017b82-part2 ONLINE      0    0    0

errors: No known data errors

  pool: rpool
  state: ONLINE
  scan: scrub repaired 0B in 0 days 11:07:22 with 0 errors on Sun Mar 14 11:31:25 2021
config:

NAME                STATE      READ WRITE CKSUM
rpool                ONLINE      0    0    0
```

```
□ mirror-0                ONLINE      0      0      0
□ wwn-0x50014ee20cac5e8c-part3 ONLINE      0      0      0
□ wwn-0x50014ee20cac9ddc-part3 ONLINE      0      0      0
□ mirror-1                ONLINE      0      0      0
□ wwn-0x50014ee2b9051b25-part3 ONLINE      0      0      0
□ wwn-0x50014ee262017b82-part3 ONLINE      0      0      0
```

```
errors: No known data errors
```

En este caso si por ejemplo tuviéramos un problema con `wwn-0x50014ee262017b82-part3` tendríamos que reemplazar también `wwn-0x50014ee262017b82-part2` ya que son particiones del mismo disco. Para marcar como offline esa partición ejecutamos lo siguiente:

```
zpool offline bpool wwn-0x50014ee262017b82-part2
```

Una vez hemos verificado que no tenemos nada montado y que no hay ninguna partición que esté activa y formando parte de un RAID, ya podemos sacar el disco.

Identificar físicamente el Disco

Una vez sepamos qué disco está fallando, lo podemos hacer iluminar para identificarlo ejecutando el siguiente comando:

```
dd if=/dev/disk/by-id/(id del disco) of=/dev/null
```

Reemplazar el disco

Una vez sabemos que disco es y cuando la luz del disco esté apagada, podemos sacarlo y reemplazarlo por otro. Esto conviene coordinarlo con alguien que esté atento del syslog.

Copiar tabla de particiones

Teniendo en cuenta que `/dev/sdd` en el disco "mirror" del que está fallando y que el reemplazo es `/dev/sdc`, ejecutamos el siguiente comando para copiar la tabla de particiones:

```
sgdisk /dev/sdd -R /dev/sdc
```

Después tenemos que volver a generar los GUIDs de las particiones del disco nuevo

```
sgdisk -G /dev/sdc
```

Reemplazar particiones en ZFS

En este caso para reemplazar la partición "wwn-0x50014ee26201e74e-part3" por "wwn-0x50014ee2b9051b25-part3" en "rpool" ejecutamos el siguiente comando:

```
zpool replace rpool wwn-0x50014ee26201e74e-part3 wwn-0x50014ee2b9051b25-part3
```

Instalar grub en el nuevo disco

Para instalar grub (legacy) en el nuevo disco (/dev/sdc)

```
grub-install /dev/sdc
```

Reemplazar con discos mas grandes

1. Identificar un disco a la vez por número de serie
 - `zpool offline <zpool_name> <disk ID>`
2. Retirar la unidad y reemplázala con la unidad más grande
 - `zpool replace <zpool_name> <old disk ID> <new disk ID>`
3. Espera a que resilver se complete
4. Repite desde el punto 1. hasta que se reemplacen todos los discos
 - `zpool online -e <zpool_name> <new disk ID>` para cada disco

El grupo se expande automáticamente

Revision #2

Created 5 August 2023 06:04:04 by etaboada

Updated 8 October 2023 06:29:49 by etaboada