Backup and Restore of TruCluster System Disks

27. DECUS Symposium 2004 in Bonn

Reinhard Stadler
HP Services

Agenda

- TruCluster overview
- Backing up TruCluster system disks
- Recover from failures:
  - quorum disk
  - member boot disk
  - cluster_root
- How to create bootable copies of TruCluster system disks
- Steps to restore a cluster from a backup to new H/W

Disks required to create a Cluster

- Common Cluster Root disk(s) (/, /usr, /var)
  - Can reside on different disks
  - H/W mirror or LSM volume
  - Root can be a multi volume domain
- Create a H/W mirror set for the cluster root
  - Use a small Partition to hold the quorum disk
  - Keep in mind that you need at least 50% free disk space to run clu_upgrade

Disks required to create a Cluster

- One disk for each member to boot from
  - Use H/W mirroring to protect against failures
  - Holds a Copy of Connection Manager Data in it's h – Partition (cnx)

```
# disklabel -r dskxx
```

- Create mirror sets for member boot disks
  - Mirror set can hold all member's boot disks
  - LSM is not supported for member boot disks

Disks required to create a Cluster

- A quorum disk (for an even number of cluster members)
- The disk used for installation of the Tru64 UNIX Operating System
  - Local or shared disk
  - Keep this disk for recovery !!!
- Configure a spare disk that can be used for disaster recovery
- Set Identifiers to locate the disks !

26 April 2004 5

Hardware Management

- Device Special Files are unique in a Cluster
- Hardware Database to maintain persistent device information
- major/minor device numbers required to reference the device
- HW Database files are located in cluster_root and member boot partitions
- Consistent copy of all files required

26 April 2004 6

Hardware Database

- Hardware Component Databases
  - `/etc/dec_hwc_ldb` Local (CDSL)
  - `/etc/dec_hwc_cdb` Cluster
  - `/etc/dec_scsi_db` Local (CDSL)
- Hardware Persistence Database
  - `/etc/dec_hw_db` Local (CDSL)
- Device Special File Data Files
  - `/etc/dfsl.dat` Local (CDSL)
  - `/etc/dfsc.dat` Cluster
- Unique ID Database
  - `/etc/dec_unid_db` Cluster

Backing Up System Disks

- H/W database files
  - distributed on cluster_root and member boot disks
  - Take care to save a consistent copy
- Make sure, that the backup can be accessed after booting the OS install disk
  - Keep backup on disk
  - Consider keeping bootable copies of system disks
- A restore of the cluster to new H/W also requires copies of the CNX partitions
  - dd to the cluster_root file system

Connection Manager and Quorum

- Voting Mechanism
  - A Cluster is operational only if the majority of votes are present (the Cluster has Quorum)
- Cluster members can have either 1 or 0 node votes
- A quorum disk can have either 1 or 0 votes
- Expected votes: the number of votes configured
- Current votes are the actual number of votes

26 April 2004 9

Booting after the Cluster lost Quorum

- Use `clu_quorum` to adjust node votes, quorum disk votes and expected votes as long as the cluster is alive
- If the Cluster loses quorum all members hang until they get enough votes to regain quorum
- Halt and reboot members to adjust expected votes

```
>>>boot -fl ia
Enter kernel_name [option_1 ... option_n]
...
clubase:cluster_expected_votes= ...
clubase:cluster_qdisk_votes= ...
clubase:cluster_node_votes= ...
clubase:adjust_expected_votes=0
```

26 April 2004 11

Replace a failed Quorum Disk

- As long as the Cluster does not lose quorum, you can replace the failed quorum disk by using the clu_quorum command

```
# clu_quorum -f -d remove
# hwmgr -scan scsi
# hwmgr -view device
# clu_quorum -f -d add
```

26 April 2004 12

clubase subsystem attributes

```
# sysconfig -q clubase
...
cluster_node_votes = 1
cluster_expected_votes = 3
cluster_qdisk_major = 19
cluster_qdisk_minor = 159

cluster_qdisk_votes = 1
cluster_seqdisk_major = 19
cluster_seqdisk_minor = 175
```

quorum disk CNX Partition

CNX Partition of member‘s boot disk

- Cluster root is stored in CNX Partitions

26 April 2004 13

Repairing a Member's Boot Disk

- Use `clu_bdmgr` to
  - Configure a member‘s boot disk
  - Back up and repair h - partition
- Steps to repair a member‘s boot disk
  - Select a new disk
  - Use `clu_bdmgr -c` to configure it
  - Mount the domain and restore it from backup
  - Edit sysconfigtab
  - Restore the h - partition using `clu_bdmgr -h`
  - Unmount the domain
- You can now boot the member into the Cluster

26 April 2004 14

Restore Cluster Root Disk

- Requires a disk, that is already known to the cluster (major / minor device number)
- OS installation disk to boot one member and perform the restore
- Steps
  - Boot one member from the OS installation disk or CD
  - Find the device to restore to (Identifier)
  - Label the disk, create file domains and filesets
  - Mount the disk and restore it‘s content
  - Modify /etc/fdmns directories
  - Shutdown the system and boot with the restored cluster disk

Specifying cluster_root at boot time

```
>>> boot -fl ia
(boot dkb200.2.0.7.0 -flags ia)
...
Enter kernel_name [option_1 ... option_n]
Press Return to boot default kernel
'vmunix': vmunix                        \
cfs:cluster_root_dev1_maj=19            \
cfs:cluster_root_dev1_min=221
```

- The System will remember the new cluster_root on subsequent boots

If LSM is in use

- As of V5.1a LSM can be used to mirror cluster root
  - Not supported to mirror member boot disks
  - Of course not supported for the quorum disk
- rootdg configuration is required at startup

How to duplicate cluster disks

- cluster_root
  - vdump | vrestore to new disk
  - /etc/fdmns directories need modification
- cluster_usr, cluster_var
  - vdump | vrestore without modifications
- Quorum disk
  - h-partition holds connection manager data (location of cluster_root and LSM rootdg configuration)

How to duplicate cluster disks

- Member boot disks
  - h-partition is used by the connection manager
  - /etc/sysconfigtab points to
    - swap devices
    - major / minor device number of the h-partition
    - major / minor device number of the quorum disk

How to restore a cluster to new H/W

- Problem
  - H/W database doesn‘t match the new H/W
  - Don‘t know the device names of the new disks
  - CNX partitions

How to restore a cluster to new H/W

- Solution
  - Install a standalone OS first
  - Restore the cluster to new disks
  - Copy the H/W database files from the standalone OS to the restored cluster disks
  - Restore the CNX Partition of the boot disk
  - Modify system configuration
  - Boot in interactive mode to single user and build new kernel
  - The new kernel boots to multiuser mode
  - The cluster is now up and running with one member

Conclusions

- As long as the common cluster root isn‘t affected everything can be repaired online
- Restoring cluster_root to a disk that is already in the H/W database is easy
  - Consider keeping a spare disk for recovery
  - Keep documentation of your device configuration
- There are tools available to duplicate all system disks so that you can boot straight of it
- Recovering everything to new H/W requires deep knowledge of TruCluster functionality


Fragen ?


hp
invent