Peter Edwards

CSIRO Advanced Scientific Computing

CSIRO

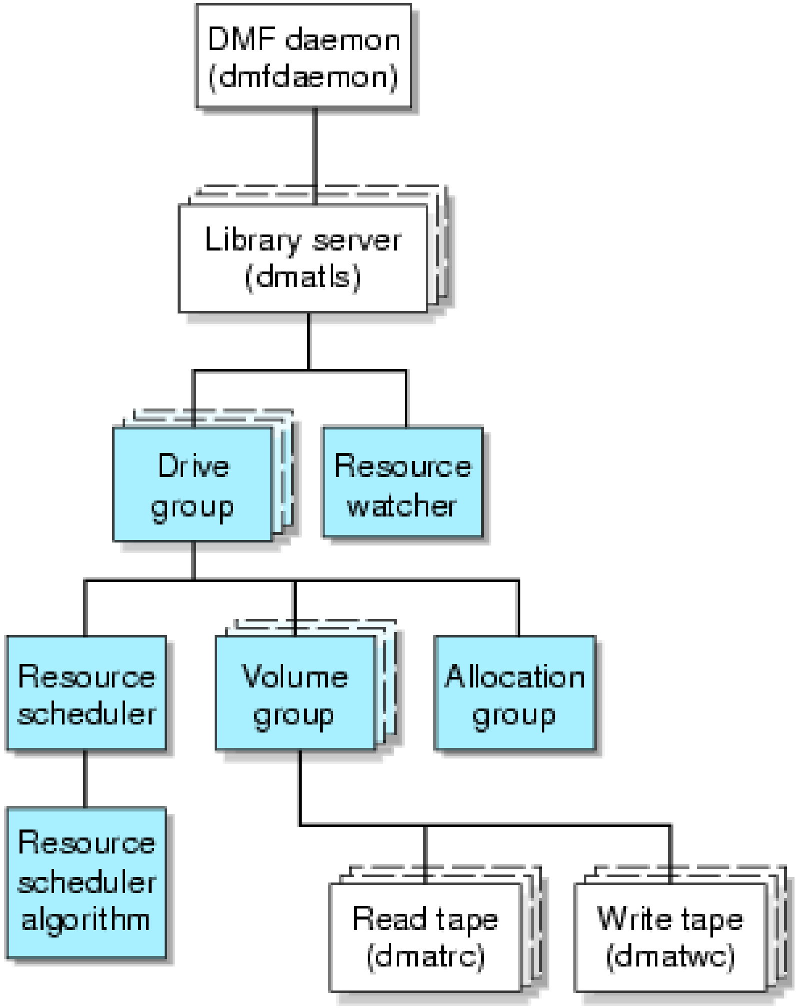
Introduction

- Avoidance of tape errors is one of the two major functions of the Library Server's Drive Group (DG) component.
  - This presentation is based on information from 2006 but is believed to be still current in 2010.
  - We thank SGI for their assistance. They bear no responsibility for any information which is incorrect or obsolete.
  - Apologies for the very heavyweight slides!

Background

- Tapes are unreliable
  - Drives are mechanical
  - There is physical contact between the fast-moving media and the heads and other drive components
  - Cartridges are moved around by humans and machines
- So make multiple copies of the data
- When (not if) a drive or volume fails, retry with a different drive or volume
- Repair the drive or replace the volume
- How hard can it be?

Volume Groups and Drive Groups



Error Categories

- At tape dismount, DG function “autopsy” records the volume serial number (VSN), drive name and error category (if any):

DRIVE_MS_DOWN	No connection to OV/TMF
DRIVE_NO_DRIVES	No drives available (eg: all configured down)
DRIVE_MEDIAUNKNOWN	No such tape in library
DRIVE_MEDIACONFIG	Media config problem, such as "ring-out"
DRIVE_MOUNT	Couldn't mount on a particular drive
DRIVE_LABEL	Problem reading or writing a tape label
DRIVE_POSITION4READ	Positioning prior to a read
DRIVE_POSITION4WRITE	Positioning to EOT prior to a write
DRIVE_READ	I/O error while reading
DRIVE_WRITE	I/O error while writing
DRIVE_DISMOUNT	Error during tape dismount

Special-case Processing

The first four categories of errors can be acted upon immediately, without regard to context or past history:

- DRIVE_MS_DOWN
  - Restart MS (ie: OV/TMF) if permitted by MAX_MS_RESTARTS
- DRIVE_NO_DRIVES
  - Restart MS if down and if permitted by MAX_MS_RESTARTS; otherwise just set the tape's HLOCK flag in the VOL database for a while to prevent repeated accesses
- DRIVE_MEDIAUNKNOWN
  - Set the tape's HLOCK flag for a while to prevent repeated accesses
- DRIVE_MEDIACONFIG
  - Set the tape's HLOCK flag for a while to prevent repeated accesses

No further action is taken for these categories, apart from email

Write Error Processing

After those errors have been detected and processed, the DG tests for errors relating to writing to tape, the DRIVE_POSITION4WRITE and DRIVE_WRITE categories.

For these, the DG disqualifies that tape for writing for a little while. This is to force DMF to retry its writing on another volume.

This is done even if subsequent analysis suggests that there was nothing wrong with the volume; reacting promptly to a filling filesystem is more important than waiting until you're sure of the likely cause of the problem.

General Processing

All errors, including those in the DRIVE_POSITION4WRITE and DRIVE_WRITE categories, are then analysed by function “autopsy_history”. This is done by comparing the current problem with records of previous drive and volume errors.

- For example, if a given tape volume has been unreadable several times in a row, even though different drives were used, then the DG concludes that the problem is most likely due to the volume rather than the drive. So it suspends use of that tape for a while, forcing DMF to recall the file from another tape held by a different VG.
- But if instead, a several volumes fail on a specific drive, but are usable on other drives, then a drive problem is likely, and the drive may be automatically configured down if permitted by the administrator.

This requires knowledge of the drive name, but TMF (unlike OV) only provides this after a successful mount. If the failure happened before then, the DG tries to guess which drive (if any) was involved. See a slide at the end for details of these guesses.

Control Parameters

The following parameters control the disabling of faulty drives or volumes, and their names are used in the description below.

These are hard-coded:

DRIVE_THRESHOLD	1.0
CONSECUTIVE_DRIVE_ERRORS	4
SOLE_VOLUME_MULT	2
MIN_DRIVE_USES	5
VOLUME_THRESHOLD	1.5
CONSECUTIVE_VOLUME_ERRORS	3
SOLE_DRIVE_MULT	2

These are settable through the *dmf.conf* file, with defaults shown:

DRIVES_TO_DOWN	0
MAX_MS_RESTARTS	1 (TMF) or 0 (OV)
MOUNT_TIMEOUT	0 (ie: never)
REINSTATE_DRIVE_DELAY	1440 (1 day)
REINSTATE_VOLUME_DELAY	1440 (1 day)

Definitions

- Use

  An attempt to mount, position and read/write a volume on a drive and then dismount it, which might or might not fail at any stage in the process.

- Condition Labels

  A code such as D3 or V1 used in the description below, and visible in the LS logs if MESSAGE_LEVEL is set to 3 or above.

- Ignored

  Drives are "ignored" if they are down or have been used less than MIN_DRIVE_USES times since last coming up. This means they are excluded from all statistical calculations and comparisons, either as a suspect or as a non-suspect. It doesn't mean "won't be used".

- HLOCK, HVFY, HSPARSE

  Flags in the database entries describing a tape volume, which can be manipulated by *dmvoladm.*

- BOT, EOT

  Beginning of Tape, End of Tape.

Drives

- A drive is configured down only if all of the following are true:
  - D1. It is not being ignored
  - D2. At least one other drive in the same drive group is not being ignored
  - D3. Its number of consecutive errors of the same type since coming up (or in the last month if it's been up for longer) exceeds the mean of the same errors for all non-ignored drives plus DRIVE_THRESHOLD times their standard deviation
  - D4. The last CONSECUTIVE_DRIVE_ERRORS uses all failed with the same category of error, and involved more than one volume. If only a single volume, then CONSECUTIVE_DRIVE_ERRORS multiplied by SOLE_VOLUME_MULT failures are required instead
  - D5. There are fewer than DRIVES_TO_DOWN drives already down for any reason (DRIVES_TO_DOWN comes from *dmf.conf* and defaults to zero, which prevents drive downing)

Drives (cont'd)

If a drive is downed by DMF, a "critical" email is sent to the administrator.

Such a drive may be manually configured up again at any time by the administrator using *tmconfig* or *ov_drive.*

This happens automatically after a delay of REINSTATE_DRIVE_DELAY minutes (default of one day; zero is interpreted as infinite). This option is intended to make mistakes in the diagnosis self-healing, and for sites with unattended systems.

Volumes

- A volume is considered faulty if:

  - V1. Its number of consecutive errors of the same type in the last month exceeds the mean of the same errors for all other volumes by VOLUME_THRESHOLD times.

  and either

  - V2. The last CONSECUTIVE_VOLUME_ERRORS uses all failed with the same category of I/O error, and involved more than one drive. (If only a single drive, then CONSECUTIVE_VOLUME_ERRORS multiplied by SOLE_DRIVE_MULT failures are required instead.)

  or

  - V3. The last CONSECUTIVE_VOLUME_ERRORS uses all failed with a mount error.

Volumes (actions)

- If Appending

  If this happened while attempting to append additional data to a non-empty tape, further attempts to do so are prevented by setting the HVFY database flag, and notifying the administrator by email. The existing data on the tape is still accessible.

  HVFY remains set until cleared manually. It is expected that the tape will be allowed to become empty by merging or hard-deletes and then tested and/or replaced prior to being returned to production by clearing this flag. An explanation of how to do this is included in the email sent to the administrator at the time.

- Otherwise

  The tape is prevented from being used either for writing from BOT or for reading by setting the HLOCK flag, and an email is sent to the administrator. HPARSE is cleared. The volume may be unlocked again at any time by the administrator with *dmvoladm*.

  This happens automatically after a delay of REINSTATE_VOLUME_DELAY minutes (default of one day; zero is interpreted as infinite) and at DMF restart. This option is intended to make mistakes in the diagnosis self-healing, and for sites with unattended systems.

Volumes (end-case)

- A pathological end-case:
  - V9. Only one drive is available, and only one particular volume is attempting to use it, and failing every time.

    The user's request is rejected after the greater of CONSECUTIVE_DRIVE_ERRORS and CONSECUTIVE_VOLUME_ERRORS attempts, but the volume is not HLOCKed.

    This should prevent a retry loop for now, but allow another attempt later on, when there might be more drives available.

Miscellaneous Notes (1)

- In addition, as noted earlier, writing to a volume may be prevented before these conditions are satisfied. This prevents attempts to write to a possible damaged (physically or logically) volume, without preventing attempts to read from it.
- Records of scheduled drive reinstatements may be lost over a DMF restart; a message to this effect is sent to the logfile during shutdown. The sysadmin will have to do them manually.

  All HLOCKed volumes are reinstated at DMF startup.

Miscellaneous Notes (2)

- If a positioning error is encountered prior to appending to a tape, it is possible that the error actually happened when the tape was last written to. So the drive in use at that time is checked out as well as the current one.

- It is possible, though unlikely, for just one additional error to disable both the current drive, the previous one and the volume; they are all processed independently of each other.

- Note that these techniques perform poorly if there are multiple concurrent errors, such as a faulty drive at the same time as a corrupted volume is accessed. This is another reason for ensuring that the decisions made by the DG are repealed after a period of time, as they may be wrong.

Miscellaneous Notes (3)

- If MOUNT_TIMEOUT is set to a non-zero value, any tape mount which takes longer than this number of minutes to complete will trigger a MS restart.

  Don't make this value too restrictive, as any non-LS tape activity (including *xfsdump*) can legitimately delay a VG's tape mount, which could result in this timeout being exceeded.

Guessing TMF Drive Names

The techniques used to guess at the name of a TMF drive which failed to mount include (in order):

1. Asking TMF for details of all mounted volumes, in the hope that the one we want is mounted on a drive, despite the failure of the mount request.

2. Failing that, look for a drive with the correct Resource Scheduler reservation number. (This is the number you see after the '#' on a "*tmstat -l*" command and in the library server log file.)

3. For TMF with the TS tape driver, look for certain TMF messages in the *daemon.stderr* file belonging to the TMF daemon. To make sure that whatever is found is appropriate to the current error, the details are matched against the contents of the per-request *tmf_msg.<pid>* file.

4. Look for drives which are dismounting unknown volumes. If there is exactly one such drive, then we assume it is the one we want.

Guessing TMF Drive Names (cont'd)

5. If there were more than one unloading drive and the mounting service is TMF, a quite different procedure is followed. All such drives are configured down. Then after delay of UNLOAD_TIME seconds, (typically 60) to allow the unloads to complete, it attempts to configure those same drives up again. If this succeeds, TMF may have reset the drive(s) and cleared the problem.

   If it fails, then the faulty drive(s) will remain down, the RS will eventually find out that this has happened, and it then schedules requests based on the reduced number of drives. The administrator will have to investigate the downed drives and reinstate them by hand.

   (An equivalent procedure is not followed for OV because it doesn't have this ability to reset the drives.)

6. Assume that the problem was really a mount failure, and use the fictitious _*none*_ drive to track them.

CSIRO Advanced Scientific Computing

Peter Edwards

Phone: 03 8601 3812
Email: peter.edwards@csiro.au
Web:http://hpsc.csiro.au/users/dmfug/

www.csiro.au

Thank you

Contact Us
Phone: 1300 363 400 or +61 3 9545 2176
Email: enquiries@csiro.au Web: www.csiro.au

CSIRO