Dealing with Robot Failures

Peter Edwards
CSIRO Advanced Scientific Computing

CSIRO

Introduction

- DMF analyses errors to attempt to avoid future problems
  - http://hpsc.csiro.au/users/dmfug/Meeting_May2010/Presentations/tape_error_avoidance.pdf
- It normally assumes that drive and volume failures are unrelated
- When a robot in an Oracle STK SL8500 fails this is not so
- (Similarly for configurations using external pass-through ports)

Disclaimer

- We use TMF in a single STK SL8500 with four robots
- I have no recent experience with
  - OpenVault
  - Non-STK libraries
  - Multiple libraries interconnected with pass-through ports
  - An 8-robot SL8500

SL8500 Overview

- Four horizontal “rails”, containing:
  - One quarter of the library's tape slots
  - Either one or two single-handed robots
  - 0 - 16 tape drives, possibly of different types
  - Three of them have “Cartridge Access Ports” for injecting and ejecting tapes
- Two vertical “elevators” allowing tapes to be moved between rails
- If the original and final locations of a tape being moved are in the same rail, then each rail operates independently of the others
- If not, then the two rails coordinate via one of the elevators
- In ACSLS terminology, each rail is an LSM and the whole

DMF's actions when a robot fails

- Any attempt to mount a tape using a failed robot is rejected immediately, without any retries, as if it were missing from the library
  - The tape will be locked by setting the HLOCK flag
  - DMF tries to access the second copy, which had better not be in the same rail (see later)
  - After a site-selected time, or at DMF restart, the HLOCK is cleared
  - Repeat the above until the robot is fixed
- This is exactly the correct reaction, though with a misleading reason given
- Except for the fact that you will receive one “serious” email per failure (ie: hundreds or thousands of them!)

DMF's actions when a robot fails (cont'd)

- Any attempt to use a drive using a failed robot fails, **with retries**, as if it were a mount failure
  - Multiple attempts will be needed before the drive is downed, assuming the site allows it to be
    - http://hpsc.csiro.au/users/dmfug/Meeting_May2010/Presentations/tape_error_avoidance.pdf
  - After a site-selected time the drive is reinstated
  - Repeat the above until the robot is fixed
- This will only happen if the tape comes from a rail with a working robot, as the “cannot access tape” error masks the “cannot mount tape” one
- As ACSLS tries to minimise elevator use, this is infrequent
- If it does happen repeatedly, it could cause the tape to be wrongly locked
- A retry might attempt to use another drive on the same failed rail

Ideal world – a solution

- The Library Server learns of the rail failure somehow
  - Both ACSLS & SLConsole support SNMP, and SLConsole may support email; running an ACSLS or SLConsole query via ssh may be a last resort
- Given knowledge of the library configuration, it decides which drives (and tapes?) are affected
  - Manually maintained configuration file or something more dynamic?
- These are all downed/locked as appropriate
- A single email is sent to the administrator

- The process is reversed after the problem is fixed (best) or after a timeout

Real world – use a human

- Administrator notices a flood of emails apparently relating to unknown tapes and decides to act upon them
- The affected drives are manually configured down to prevent attempts to use them for tapes in one of the still-operative rails

- The process is reversed after the problem is fixed
- Optionally, all the HLOCK flags are cleared manually (this will happen automatically, given time)

Pre-emptive action

- Redistribute tapes amongst the rails to minimise the effects on DMF of a future robot failure
  - If using multiple drive types (hopefully installed in different rails), move tapes to the same rail(s) as their drives
  - Move tapes contain primary copies of files to a different rail from those containing secondary copies, so one rail failure does not deny access to both copies
  - Move some empty tapes, and maybe cleaning tapes, to each rail
- As tapes move around over time, normally being placed in slots near the last drive to use them, this process must be repeated regularly (eg: every night)

Example

- T9940B tapes are only for primary copies of files, T10000B's are for both primary and secondary
- Tape distribution at CSIRO 24 hours after previous redistribution:

LSM	Drive Type	Tape Type	Purpose	Empty slots
0	T9940B	mostly T9940B	mostly primary	0
1	T9940B	mostly T9940B	mostly primary	0
2	T10000B	both T9940B & T10000B	mostly primary	1
3	T10000B	mostly T10000B	mostly secondary	141

(We currently have too many T9940's to fit in just two rails)

Example (cont'd)

- Move tapes in four steps using the ACSLS cmd_proc utility:
  - T10k's from LSM 0 & 1 to LSM 3 *(7 tapes moved)*
  - T9940's from LSM 3 to LSM 0 & 1 as much as possible *(8)*
  - Primary T10k's to LSM 2, alternating with secondary T10k's to LSM 3 *(10 each)*
  - Try to place half the empty T10k's in each of LSM 2 & 3 *(0)*

- After redistribution:

LSM	Drive Type	Tape Type	Purpose	Empty slots
0	T9940B	T9940B	primary	0
1	T9940B	T9940B	primary	0
2	T10000B	T9940B & T10000B	primary primary	0
3	T10000B	T10000B	secondary	142

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