News from the ENDIT world

Efficient Nordic Dcache Interface to TSM

norden

NordForsk

Nordic e-Infrastructure Collaboration

NordForsk

Nordic e-Infrastructure Collaboration

Overview

- ENDIT is based on doing “dsmc archive” and “dsmc retrieve” to store files in TSM as dCache HSM backend
- Four parts:
  - HSM provider (previous: script, now: plugin)
    - Creates files/links that the deamons work on
  - tsmdeleter.pl daemon, does removals
  - tsmarchiver.pl daemon, does archiving
  - tsmretriever.pl daemon, does retrievals
- Batching due to TSM's nature of only reordering reads within a session

NordForsk

Nordic e-Infrastructure Collaboration

Major news

- Native HSM provider plugin instead of script
  - Great performance increase, possible to have >20k outstanding restores
- Support for parallel retrieves from a single pool
  - Depends on getting a tapes.hint file of “<pnfsid> <tape>” that covers most files in order to get parallelism
  - Then launches workers, one per tape, up to a configurable max
- Parallelism on write is trivial, just run more write pools
  - But for reads cleverness needed, otherwise they'll just want the same tape all the time

State of the art

- Restored 1M small files in 8.5h from file-backed TSM storage
  - 8 workers, peaking at 200k-300k outstanding requests
  - Typical restore of about 50-200 files per “tape mount”
  - Steady incoming stream of “rh restore” during the first few hours
  - Larger tests limited by RAM on my test tape pools (only 12G)
- Speedrun of 10k small files
  - 8 workers, all restores preloaded before starting retriever daemon
  - On adjacent “tapes”, so about 2000 files per “tape mount”
  - 70 seconds, 150Hz, if I remove 10s overhead of starting the daemon

Future plans

- Finish testing and roll out on production at NDGF-T1
  - Also some issues still to be fixed
    - Like restoring some files twice in some corner cases and leaving them in a temporary directory
- Multiple pools per host without multiple endit installs
- Make a proper release of both plugin and daemons
  - Does anyone need more than a tar for /opt/endit and a jar?
  - Should happen at the latest this summer
- Until then you can find the bleeding edge on github :)

NordForsk

Nordic e-Infrastructure Collaboration

Questions?