EOS clone/back-up

EOS-specific support for consistent back-ups

Overall Logic

- Clone directory to be backed up (aka "snapshot")
    - marks files as belonging to "this" clone
        - all files ("full back up") as belonging to "this" clone
        - files modified after specified date ("incremental")
        - produces a catalogue (all files, "json" mark-up)
        - supports nested clones
- Back up files based on catalogue to archive storage
- Erase the clone
    - unmark all files belonging to "this" clone
- Time-critical operations performed directly in MGM

Clone Logic

- Cloning is a "short" operation - e.g. 10s per 50000 files
    - But is not "atomic" (which would be possible but very heavy)
    - All files on the same directory level are "strictly" consistent, the whole tree "mostly"
- Changes to files belonging to a clone are handled via copy-on-write
- Clones are cheap, but not free
        - Meant to last only while date are being backed up
        - Copy-on-write overhead
            - for updates other than new files, deletes or complete re-writes
            - supports "reflink" copies where kernel and filesystem do
            - otherwise a real copy on the FST(s)

Back-up Logic

- Relatively feature-rich example program in Python
- You can write your own, easily! Or modify the example to suit your needs
- Back-up logic: create a clone, back up data, remove the clone, produce catalogue
- Handles full/incremental back-ups through time-stamps
- A list of archives can be fed into "restore" which restores a tree in a specified (local) file system
- Backs data up in GB-sized "blobs", suitable for e.g. CTA
- Uses XrootD for backing-up files and writing "blobs", e.g. into CTA
- Blobs for one back-up are written in parallel (but from a single client); BIG files are written into stand-alone blobs using 3rd-party-copy
- The "archive" is a [CTA-] directory containing the catalogue (!) and all blobs
- Backs up symbolic links and, within limits, hard links
- Backs up ACLs

Restore Logic

- Work in progress… currently:
  - scans a list of full-incremental back-up catalogues
  - reconstructs FULL tree in mounted file system
  - restores symbolic links
  - produces a list of blobs needed to restore all files matching a certain pattern
- Does not yet
  - restore ACLs
  - restore the original mtime/ctime/mode bits
  - restore hard-link relations
  - trigger CTA "recalls from tape"

Status

- Currently in testing, available since 4.7.0
- To be fixed, revealed in stress tests
  - recycle-bin interferes
  - versioning probably as well
  - no clones created for updates to erasure-encoding files - live files backed up instead ("replica" layout ok)
  - "git" repositories are a challenging testbed!