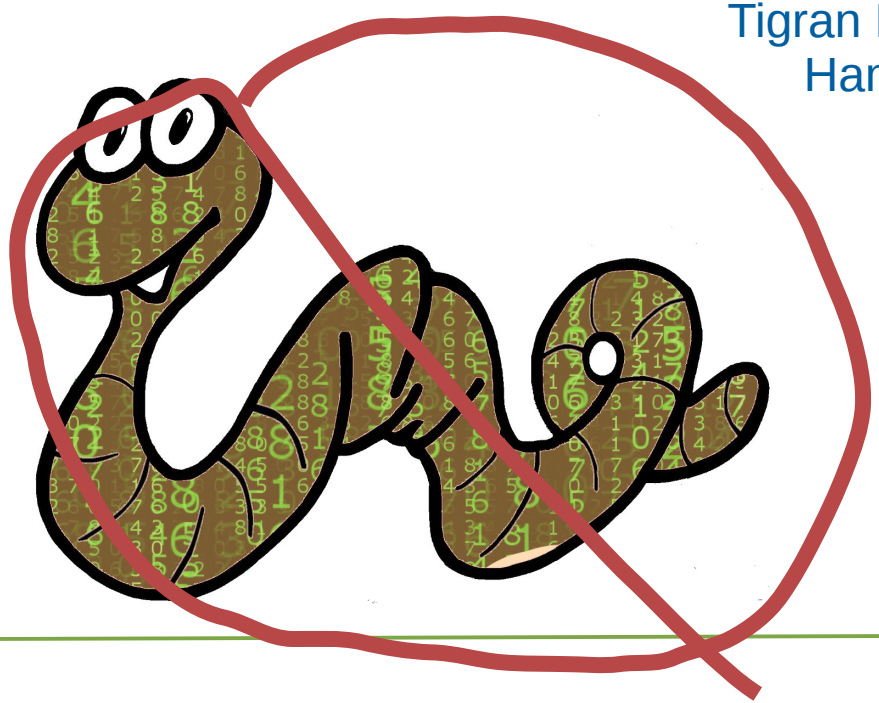
dCache.org

noWORM

Tigran Mkrtchyan for dCache

Hamburg, 29.05.2018

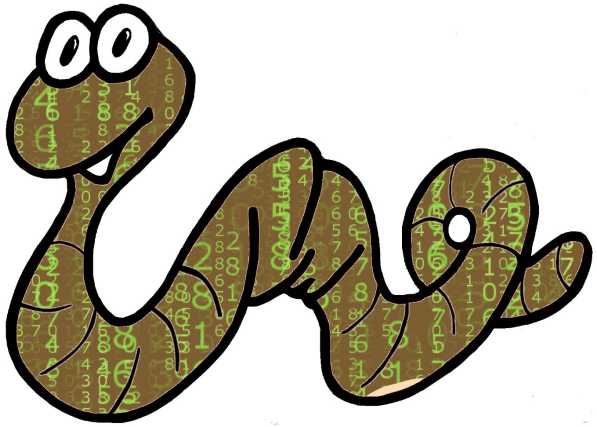


DESY NDGF Nordic DataGrid Facility Fermilab

HELMHOLTZ RESEARCH FOR GRAND CHALLENGES

WORM (SEC Rule 17a-4)

- Write Once,Read Many
- Immutable data
  - PNFSID always points to the same data
  - PNFSID used by dCache to address data
  - Multiple copies are possible
- Provided by High-End storages
  - NetApp
  - EMC
  - IBM



Who needs it?

- Legal organizations
- Video hosting providers
  - Youtube
- Streaming Audio
  - Spotify
  - iTunes
- Photo hosting services
  - Flickr
  - Google Photo

So, why WORM in not enough?

- Too many random-IO NFS workloads
  - xCloud
  - Samba
  - Scratch space
- Show stopper in some workloads
  - HDF5 files (random update)
  - xCloud applications

But wait…..

- PNFSID points to different data!
  - What about HSM?
  - Replicas?
  - Checksums?!

WORM as QoS

- DISK-ONLY + noWORM
  - No HSM
- Default policy to create WORM files
- Can be changed with QoS transition
- Directory tag **worm** (true/false) controls policy

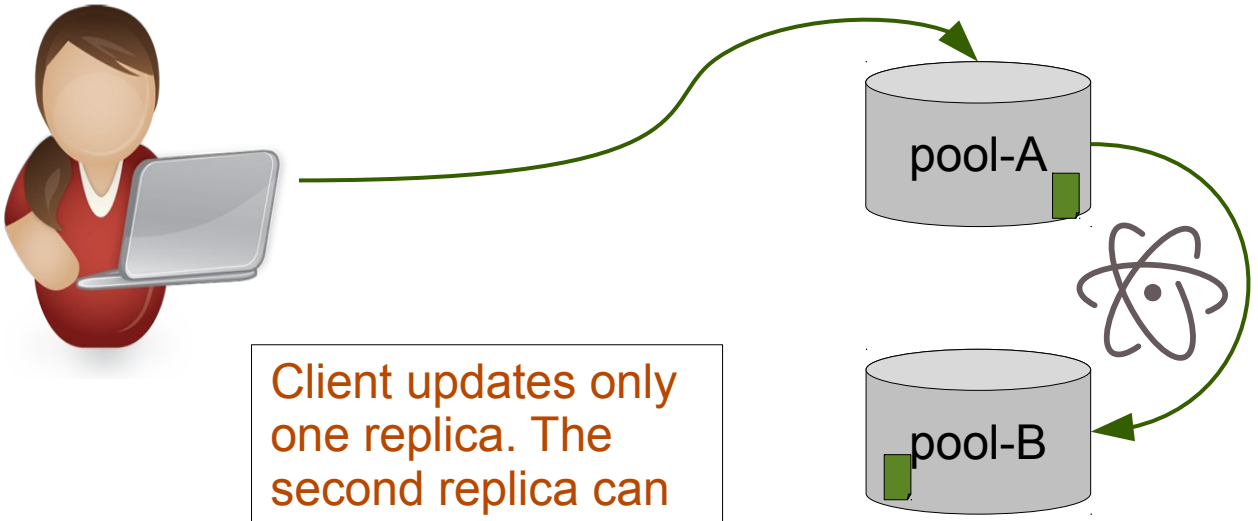
Worse – POSIX compliance!

- How much compliance is enough?
- Multi writer support
- Multi-protocol support

DEMO

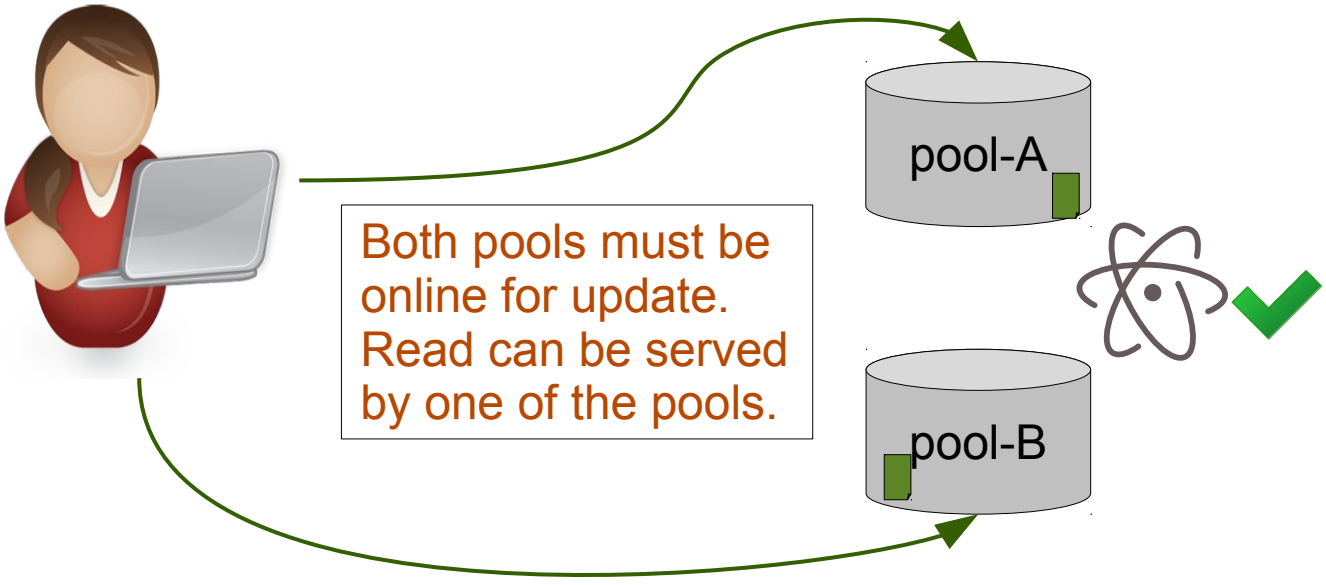
Cool, But what about replicas?!

Standard (async) Replication



Client updates only one replica. The second replica can be out of date!

Client-side Replication



Client side replication (mirroring)

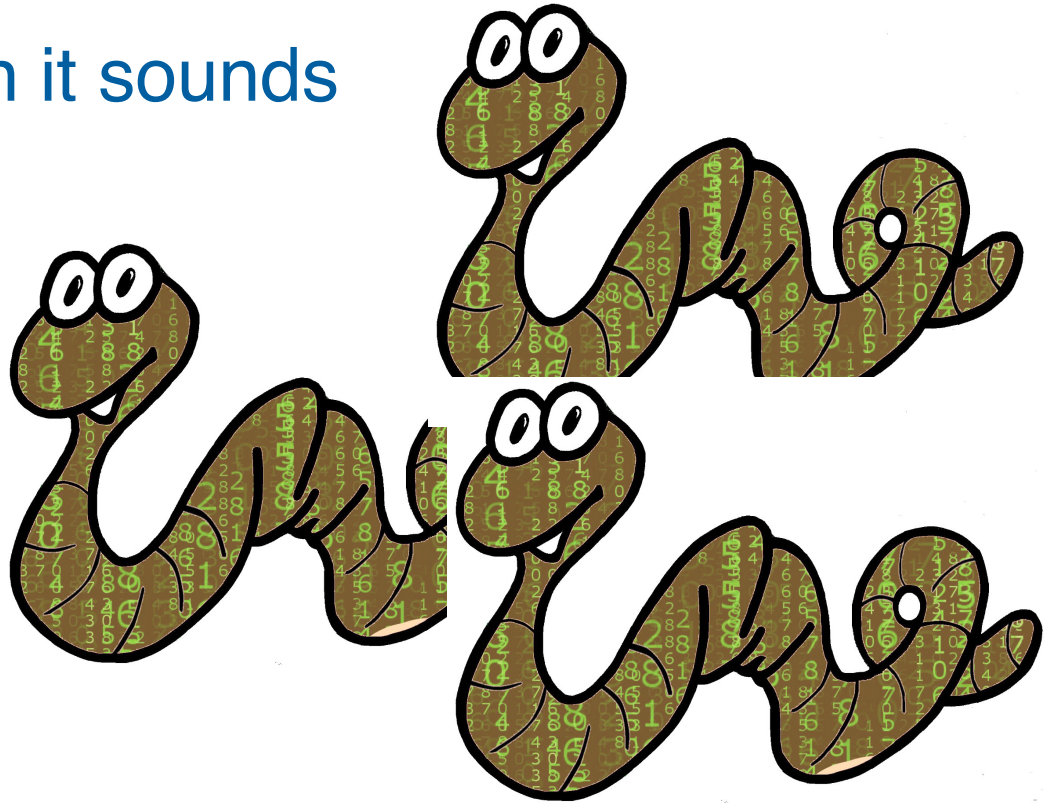
- Supported by pNFS-client with flexfile
  - RHEL 7.5 (3.10.0-862.el7) or kernel 4.15
- dCache team is working on getting it ready
  - Incremental changes as too many components have to be changed.

Internals

- No HSM, Replication, p2p for noWORM files
- Pool selection selects defined number of pools
  - First upload with non-NFS client will kick replication
  - Following updates with non-NFS client - rejected
- On first upload locations are ‘locked’
- Client instructed to update all mirrors
  - client reports IO errors to the door

Summary

- We have demand to provide noWORM solution
- It’s more complicated than it sounds
- But we are working on it!



Thank You!