

EOS+CTA :

Adding tape storage capability to EOS

Michael Davis, Vladimir Bahyl, Cedric Caffy, Eric Cano, David Fernandez Alvarez, Oliver Keeble, Julien Leduc, Steven Murray

Presenting the CERN Tape Archive (CTA)



CERN
Tape Archive

Presenting the CERN Tape Archive (CTA)



CTA is the tape back-end to EOS

EOS+CTA in Production : End 1Q 2020



+

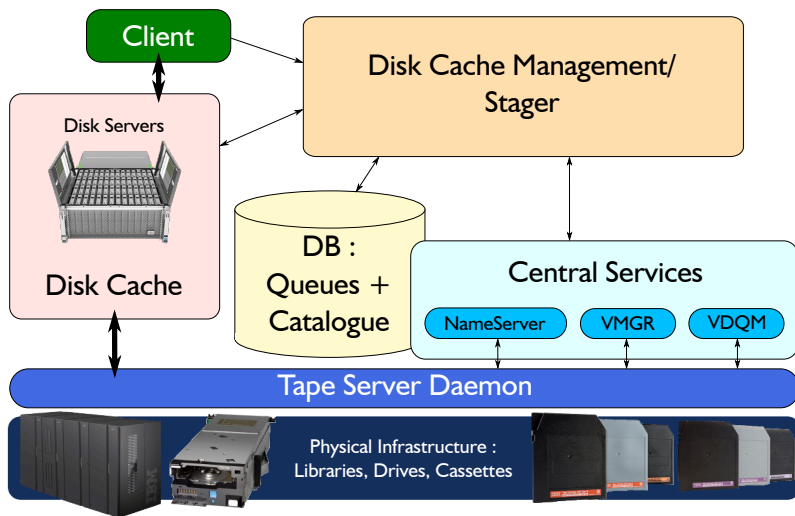


**CERN
Tape Archive**

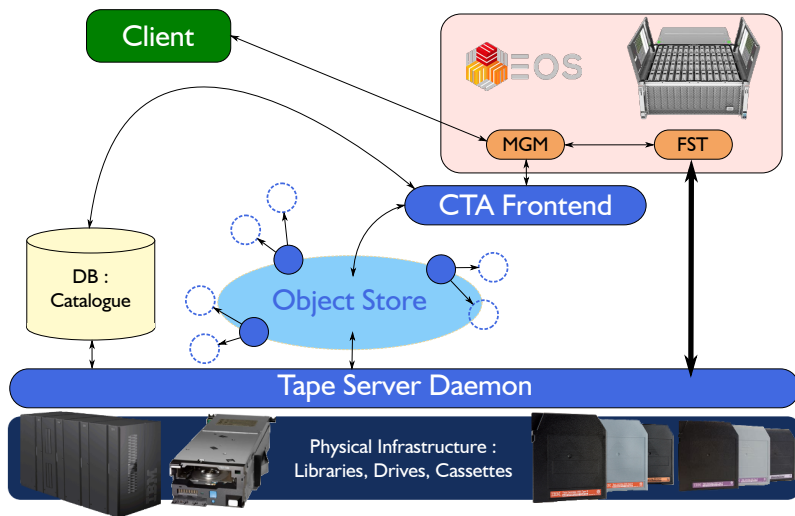
Data Archival at CERN/WLCG

- The custodial copy of all CERN physics data is stored on tape here at Tier-0.
- Additional copies of the data are distributed across the grid. Tier-1s also provide tape archival storage.
- The current Tier-0 production tape archival system ([CASTOR](#)) will be progressively replaced by EOS+CTA.
- Many different tape systems in use across the Tier-1s. Likely consolidation in the coming years.

CASTOR vs. EOS+CTA



CASTOR vs. EOS+CTA



EOS+CTA : Who Does What?

Function	Provided by
File Metadata Operations	EOS (MGM/XRootD)
Namespace	EOS (QuarkDB)
Disk Buffer for Staging	EOS (FST)
Tape File Metadata Ops	CTA (Frontend)
Archive/Recall Requests	CTA (Objectstore)
Tape File Catalogue	CTA (Catalogue DB)
Tape Operations (libraries, drives, cassettes)	CTA (Tape Server)

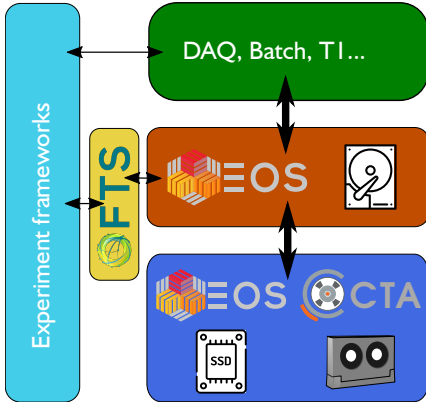
EOS+CTA : “Best of Both Worlds”

- Interface and file operations provided by EOS
- Tape operations provided by CTA

CTA design principles

- Simplicity
- Scalability
- Performance

“Big EOS” and “Little EOS”



HDD icon: <https://commons.wikimedia.org/wiki/File:Hard-drive.svg>
SSD icon: <https://commons.wikimedia.org/wiki/File:Ssd.svg>
Tape icon: https://commons.wikimedia.org/wiki/File:Tape_cista_cassette_backup.svg

“Big EOS”

- Tens of PB of storage for physics jobs and staging to Tier-1s.
- File replicas have a long lifetime.
- Spinning disks.

“Little EOS”

- Small buffer for copying files to/from tape.
- File replicas have a very short lifetime. Deleted as soon as tape copy exists (archival) or copied to “Big EOS” (retrieval).
- SSDs: reduce contention and give the best price/performance ratio.

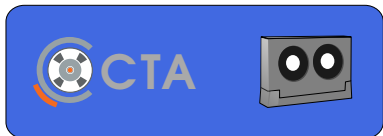
Tape Workflows



- “Big EOS” is not directly connected to tape
- To archive/recall files, transfer them to/from “Little EOS” (e.g. using `xrdcp`)



- “Little EOS” has tape-backed directories
- Copying a file into one of these directories triggers an archival request



- CTA manages the tape queues:
- Archival requests streamed to mounted tape
- Retrieval requests grouped by tape

“Little EOS” Tape Configuration

- Tape-backed directories are configured by setting EOS extended attributes

```
sys.archive.storage_class="single"
```

- The **storage class** of a file is used to specify how many copies of the file should be archived to tape, and which **tape pool** each copy will be allocated to
- A **tape pool** is a logical collection of tapes used to manage :
 - file ownership, e.g. files belonging to different VOs should not be mixed on the same tape
 - where the file should be physically stored, e.g. two copies of the same file should be stored in different tape libraries

Stages in Archiving a File

```
$ xrdcp my-important-file root://little-eos//...  
$ eos ls -y /eos/tape/backed/directory/my-important-file
```

File staged: disk replica but no tape copies (yet)

```
d1::t0  -rw-r-----  myuser mygroup  1TB  my-important-file
```

File archived: disk replica and a tape copy

```
d1::t1  -rw-r-----  myuser mygroup  1TB  my-important-file
```

Buffer cleared: no disk replicas, just the tape copy

```
d0::t1  -rw-r-----  myuser mygroup  1TB  my-important-file
```

Stages in Retrieving a File

- Recall the file from tape

```
xrdfs prepare ...
```

- Wait for the file to be recalled to “Little EOS”

```
xrdfs query prepare ...
```

- Copy the file from “Little EOS” to “Big EOS”

```
xrdcp ...
```

- Remove the replica from the cache

EOS+CTA : Summary



- EOS+CTA is EOS disk with CTA tape archival
- Interface and disk operations using familiar protocols and commands (EOS shell, `xrdcp`, ...)
- Recommended set-up :
 - “Big EOS” instance for user workflows (physics analysis and reconstruction, etc.)
 - “Little EOS” instance optimised for tape workflows

DevOps Hands-on Session

Hands-on with EOS and the CTA tape backend
Julien Leduc

A practical overview of operating EOS with a tape infrastructure

Wednesday 5 February at 11:25
513/1-024



home.cern