

CERN Tape Archive (CTA) : status and plans

Julien Leduc for the CTA Team

Presenting the CERN Tape Archive (CTA)



CERN Tape Archive

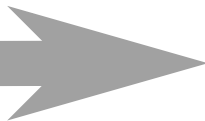
Presenting the CERN Tape Archive (CTA)



CTA is the tape back-end to EOS

EOS+CTA in Production by End Q1 2020

CASTOR
CERN Advanced STORage manager



+



CERN
Tape Archive

EOS+CTA : “Best of Both Worlds”

- EOS provides the interface, file operations and namespace
- CTA provides highly performant tape operations based on CASTOR Tape Server

CTA Design Principles

- Simplicity
- Scalability
- Performance

EOS+CTA Performance Enhancements

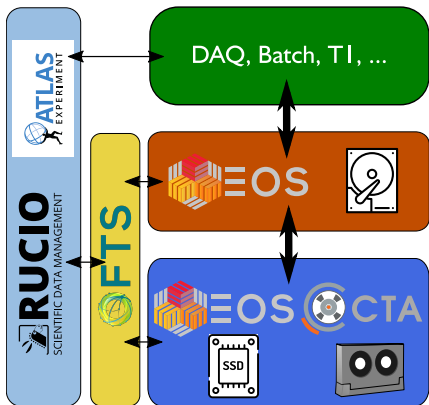
Reduced latency compared to CASTOR

- Advanced queue manager
- Just-in-time scheduling

New features in the **pipeline**

- *FTS check that file is safe on tape*
- *Recommended Access Order (RAO) for LTO media*
- *Preemptive scheduling*
- *Colocation hints*

Data Management Stack



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_cta_casette_backup.svg

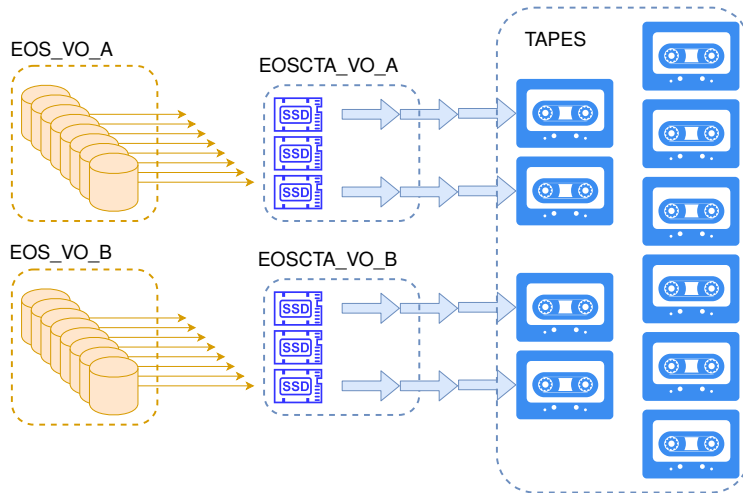
EOSATLAS “Big EOS”

- Tens of PB of storage for reconstruction, analysis and staging to Tier-1s
- Accounted as part of the pledge

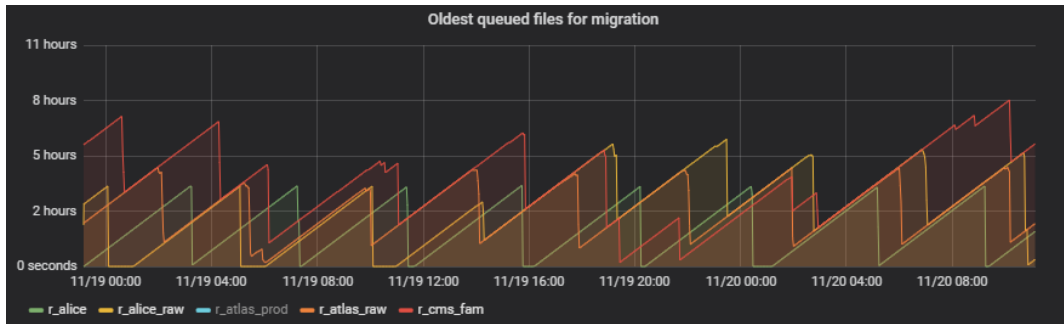
EOSCTAATLAS “Little EOS”

- Small buffer for copying files to/from tape
- Not part of the pledge; not available for physics jobs
- Files are deleted as soon as they are safely archived on tape or copied to “Big EOS”
- SSDs: reduce contention and give the best price/performance ratio

Max. throughput, minimum contention



Max. throughput, minimum contention



Ion Run (2018): highest data rate/simultaneous recalls

- SLA target: no data stays on CASTOR disks for more than 8 hours
- ATLAS fill/write sawtooth: no file stayed on disk for more than 4 hours

EOS+CTA Production Status : Hardware

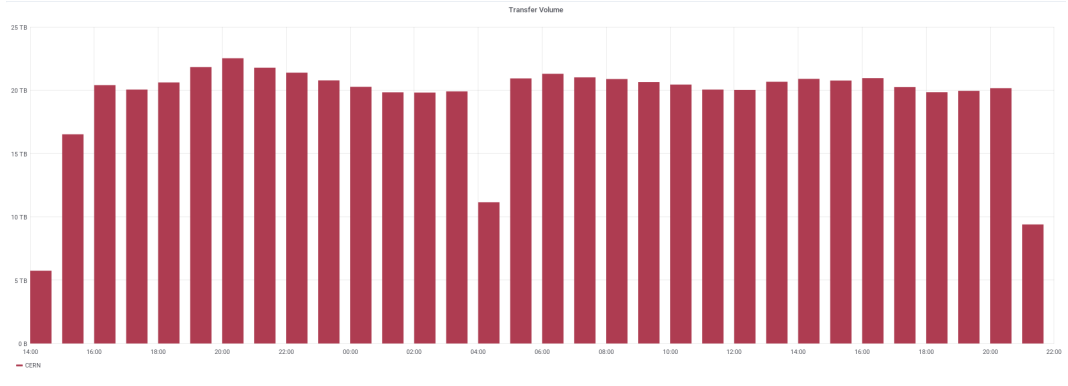
- 6× 100 Gb/s network links
- 32× hyper-converged servers
 - 16× 2 TB SSD
 - 25 Gb/s Ethernet
- 4 servers reserved for ATLAS, throughput up to 10 GB/s
- 20 to 30 Tape Drives

EOS+CTA Production Status : Software

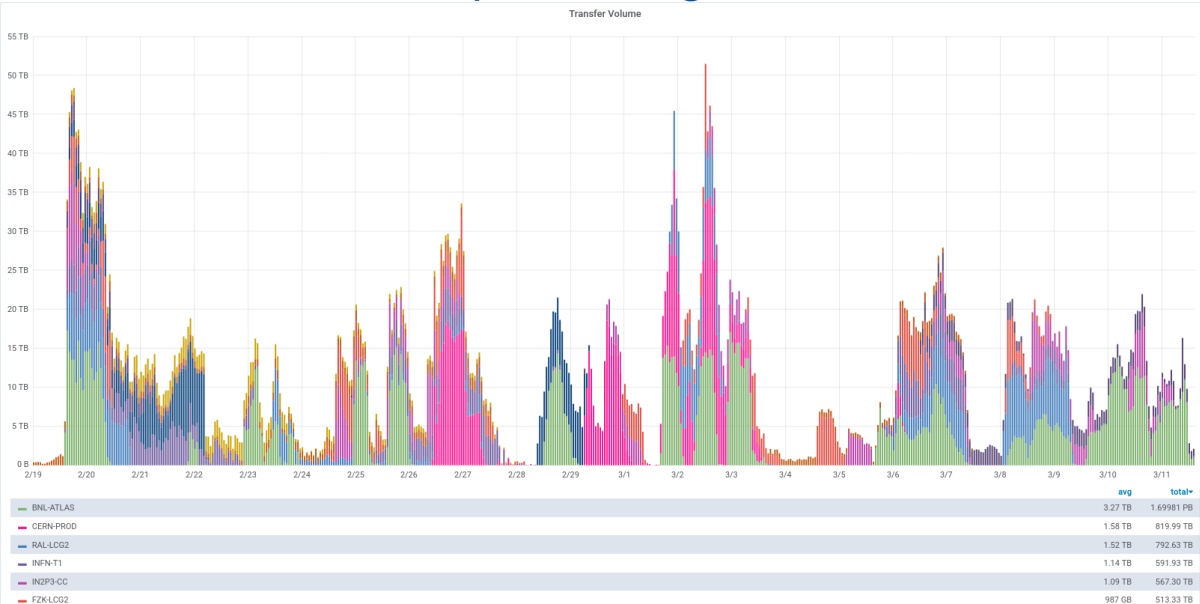
- CTA version 1.2 deployed
 - **Friday 13 2020: CTA version 2.0 will be deployed**
- Archival/retrieval workflows integrated with FTS and Rucio
- SFO handshake file is safely on CTA tape integrated

EOS+CTA Status : Commissioning

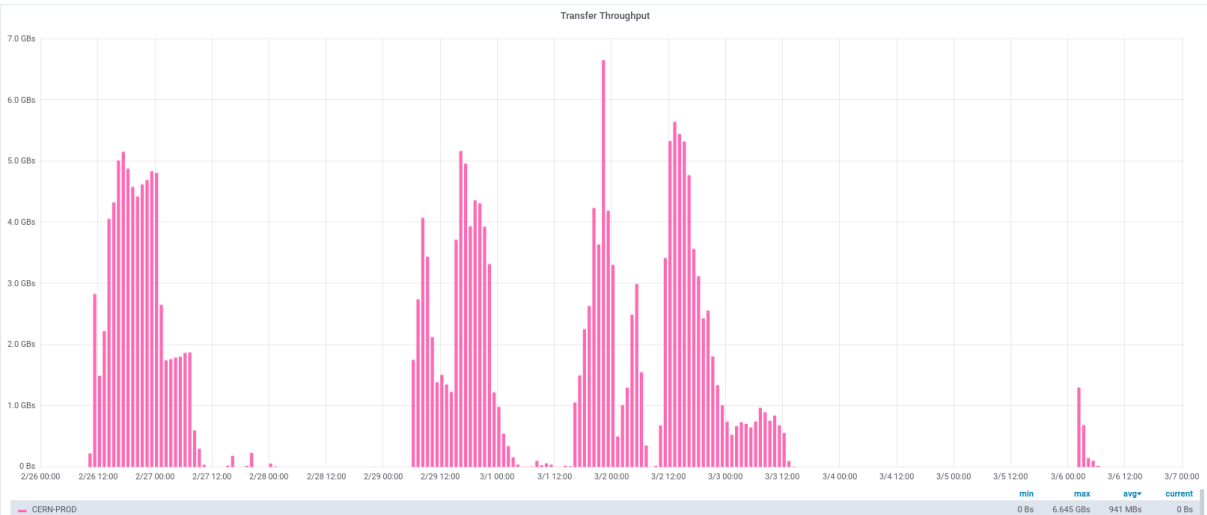
Write Stress Test : 0.6 PB at sustained 5.6 GB/s
(17 tape drives in parallel)



ATLAS data17 Reprocessing: Volume



ATLAS data17 Reprocessing: Bandwidth



EOS+CTA Status : Commissioning

- Limiting factor on throughput was tape drive availability
- CTA Service scaled smoothly as more hardware was added

Upcoming Acceptance Tests

- **16 March 2020: SFO full chain integration test**
from TDAQ buffer to file is safely on tape
- Multi-hop; Metadata stress test?; Mixed Read/Write test;

Migration from CASTOR to CTA

Migration is a Metadata-only Operation

- Tape file format between CASTOR and CTA is identical
- No physical rewriting of data
- Migration process has been tested many times (including ATLAS recall exercises)
 - ≈ 90 million files. Import to CTA ≈ 15 hours

Risk Mitigation

- Tapes imported from CASTOR are read-only in CTA
- To return a tape to CASTOR : disable the tape in the CTA catalogue and re-enable the tape in CASTOR

Conclusion

- Production hardware is installed; software is deployed
- Integration with SFO, Rucio and FTS is done
- **Current phase: Commissioning tests**
- \approx End March: ATLAS migration CASTOR \rightarrow EOS+CTA
- Other LHC experiments to be migrated before end of 2020



www.cern.ch