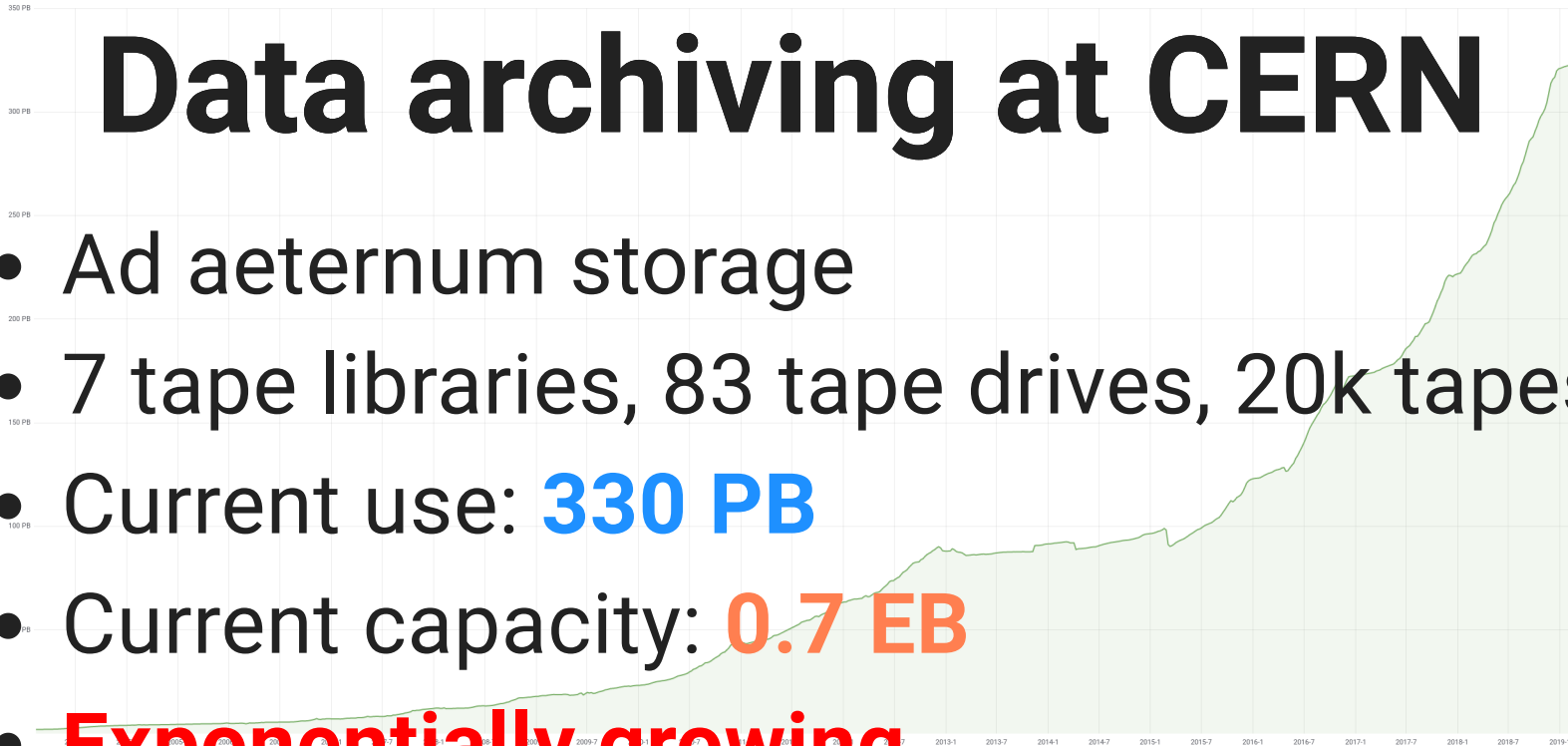


initial deployments

Julien Leduc

Data archiving at CERN

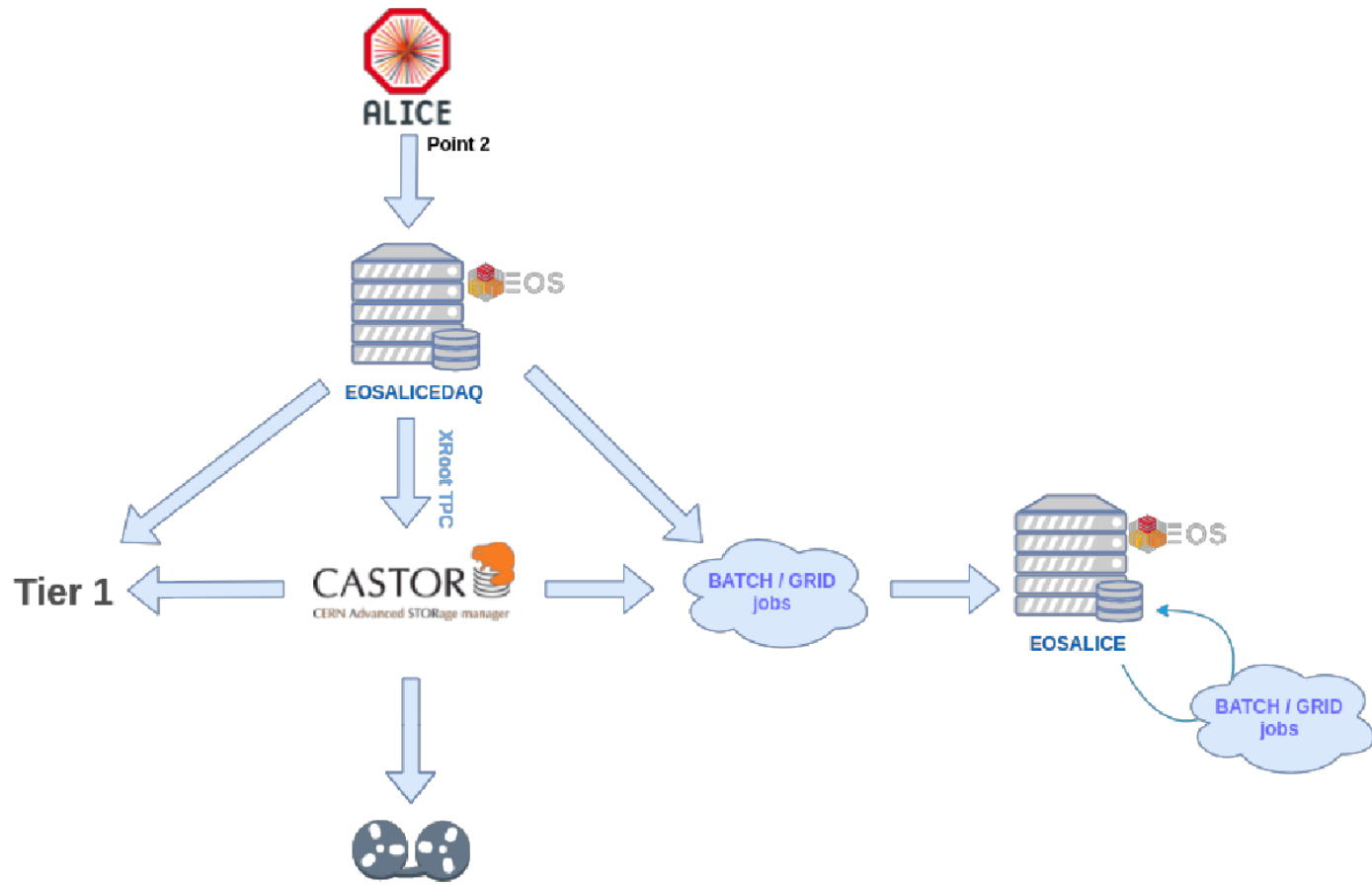
- Ad aeternum storage
- 7 tape libraries, 83 tape drives, 20k tapes
- Current use: **330 PB**
- Current capacity: **0.7 EB**
- **Exponentially growing**

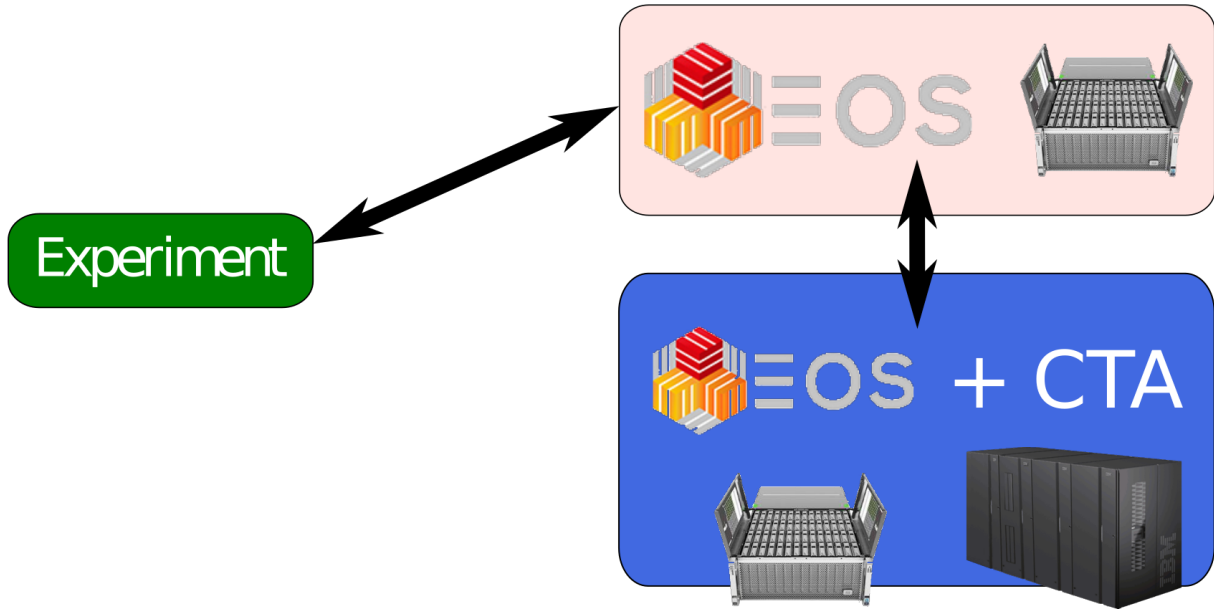


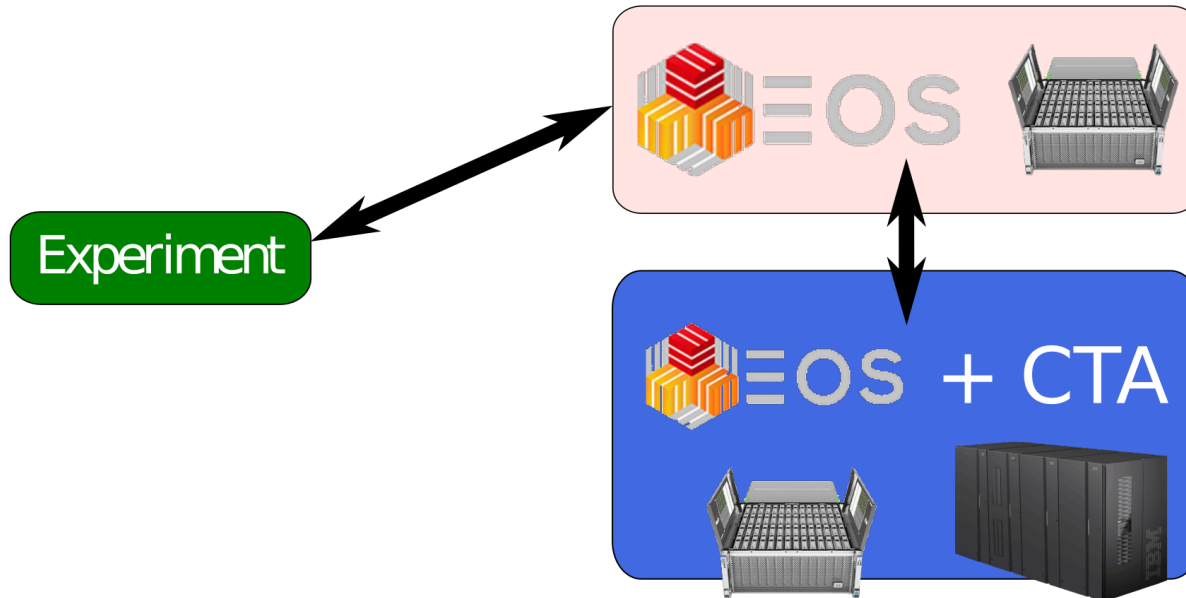
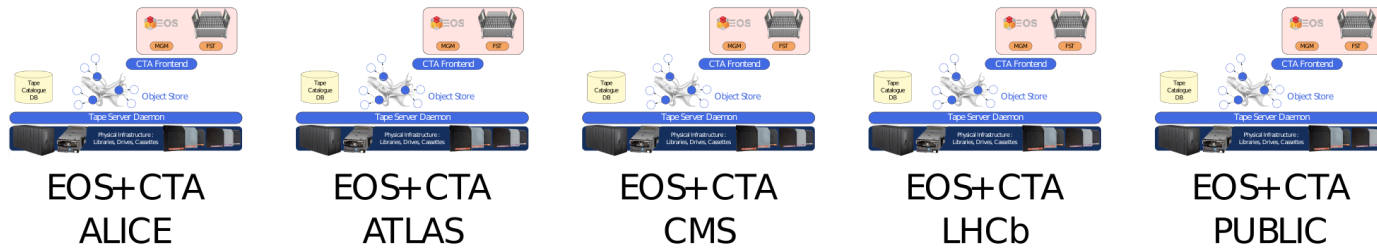
Data Archiving at CERN

Evolution

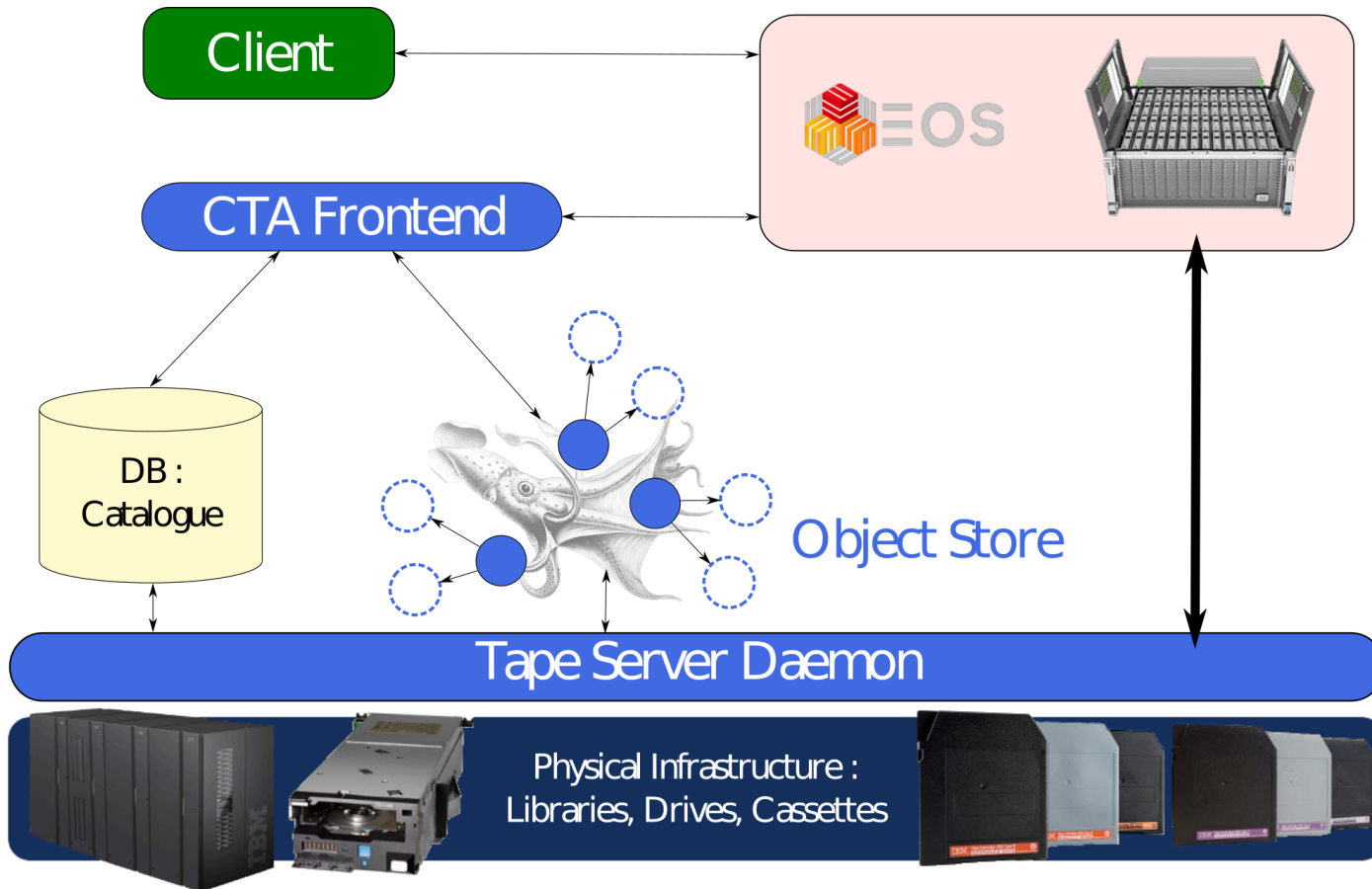
EOS+CTA *Deployment*







EOS+CTA *Architecture*



EOS+CTA *Timeline*

2018
Run 2

2019
LS2

2020
LS2

2021
Run 3

Migration during period
of reduced data taking.

Migration is a metadata-only
operation. No physical
movement of data on tape.

CTA FIELD
TESTING

DATA TAKING AND
RETRIEVAL : CASTOR

DATA TAKING AND
RETRIEVAL : CTA

EOS+CTA *Dev&oper*

Tightly coupled software ⇒ **tightly coupled developments**

Extensive and systematic testing is paramount to limit regressions

Extensive monitoring in place to ease debugging and target high performance from day 1

All **1,000+** Pending **0** Running **0** Finished **1,000+** Branches Tags

Run Pipeline

Clear Runner Caches

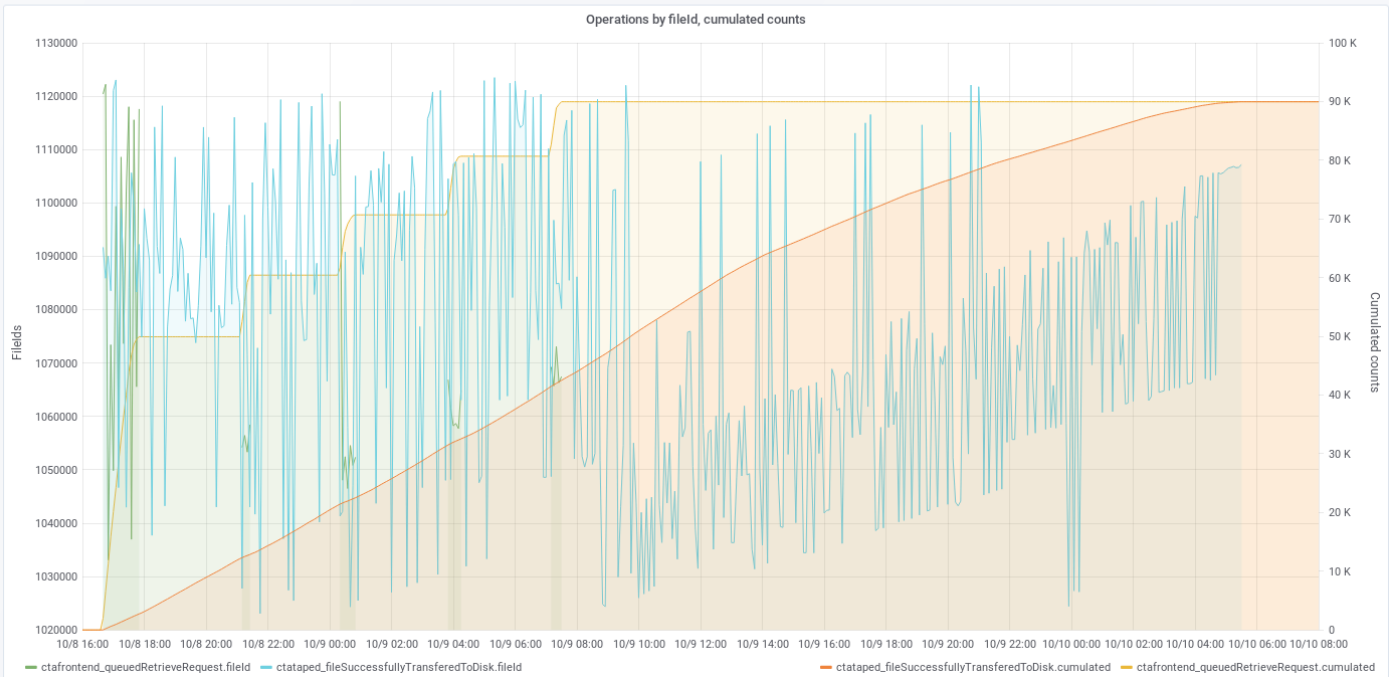
CI Lint

Status	Pipeline	Commit	Stages		
passed	#689856 by latest	Ψ postgres → bb099e43 adjust PostgresCatalogue after ma...		⌚ 00:35:44 📅 5 hours ago	▾
passed	#689743 by latest	Ψ generic05to... → 4060188b [os-failedrequests] Generalise...		⌚ 00:25:44 📅 6 hours ago	▾
passed	#689688 by latest	Ψ repack-dev → c74a10db Changed the return type of...		⌚ 01:06:22 📅 5 hours ago	▾
passed	#689583 by latest	Ψ repack-dev → e3e0cfc4 The getNextSucceededRetrieveRe...		⌚ 00:55:35 📅 6 hours ago	▾
passed	#688245 by latest	Ψ master → 622ec286 Removed unused maxTriesTo...		⌚ 01:20:50 📅 21 hours ago	▾
passed	#687380 by latest	Ψ master → 622ec286 Removed unused maxTriesTo...		⌚ 00:58:14 📅 1 day ago	▾
passed	#687037 by latest	Ψ master → 622ec286 Removed unused maxTriesTo...		⌚ 01:03:57 📅 2 days ago	▾
passed	#687021 by latest	Ψ master → 7a43dc08 Removed unused maxTriesTo...		⌚ 00:32:12 📅 2 days ago	▾

CTA timings (in house)

Oct 8, 2018 16:00:00 to Oct 10, 2018 08:00:00

resolution (seconds) 300 Puppet commits cta-taped started cta-frontend started test runs

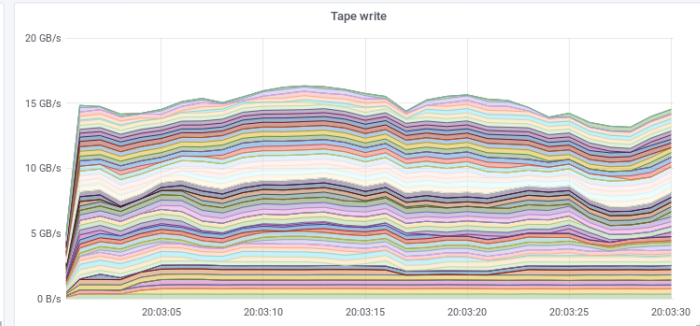
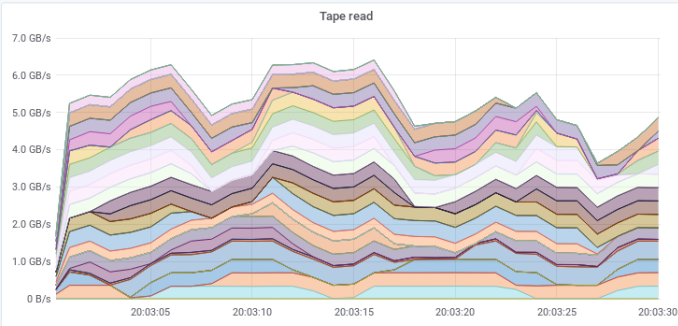


perf3 -

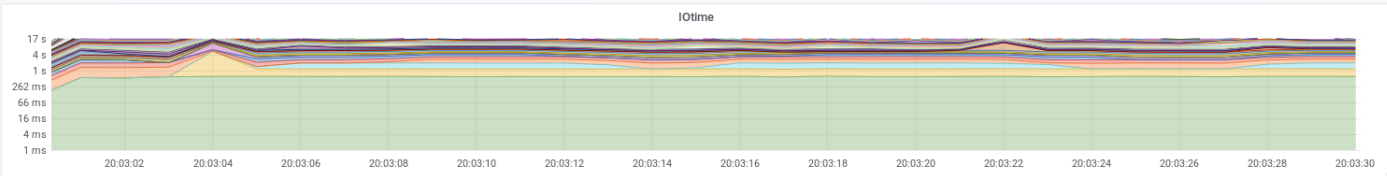
📊 ⭐ 📄 ⚙️ ⏪ 🔍 ⏩ Sep 14, 2018 20:03:00 to Sep 14, 2018 20:03:30 ↻

tapeserver All ▾ resolution (seconds) 1 ▾ instance All ▾ CTA tests

Drives



Drives IO time



CTA metadata stress tests

Small EOS instance VTL backed

- Few retired 1Gb/s disk servers and a decent
MGM for **EOS**
- A few VMs for the **CTA tape part**

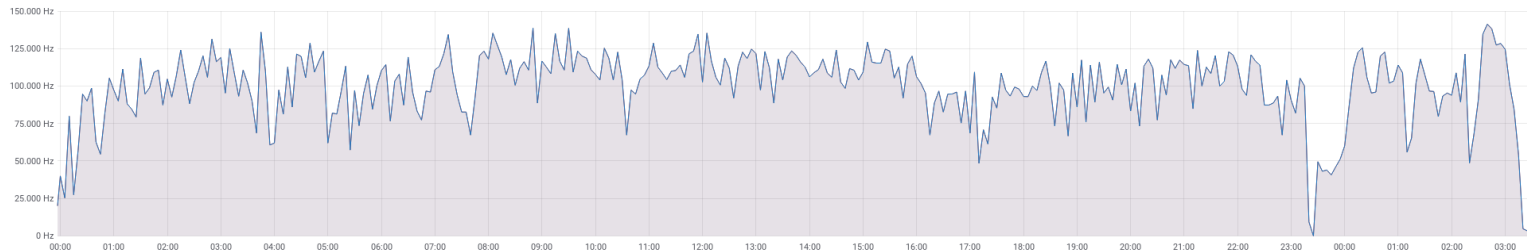
**Each VTL tape server exposes 20 virtual tape
drives and can sustain >1kHz 1kB files**

CTA metadata stress tests

Stage in

Tests until CTA routinely archived **1M files**

- **10M file test:**



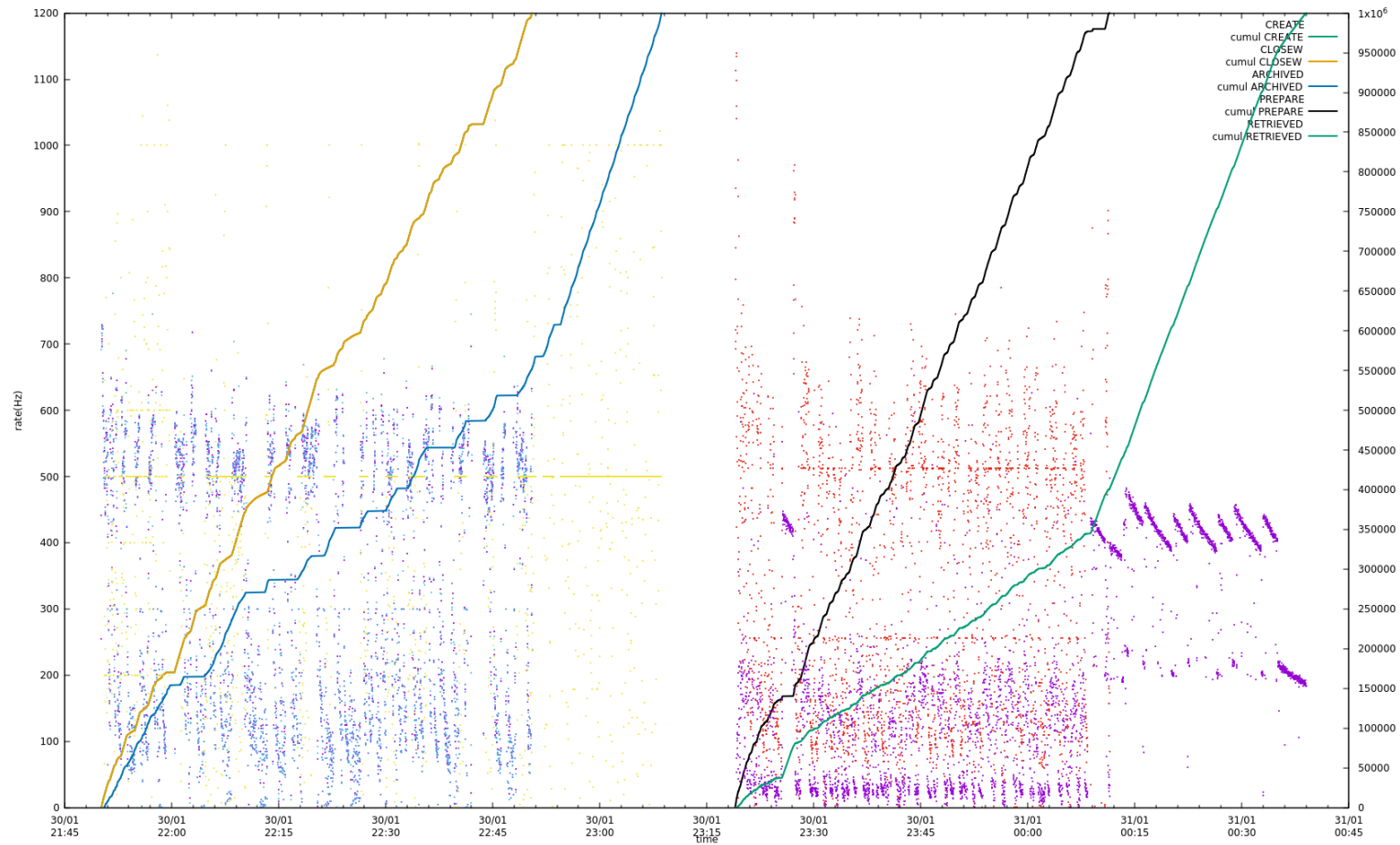
CTA metadata stress tests

Stage out

Tests until CTA routinely retrieved **1M files**
- **4M file test:**



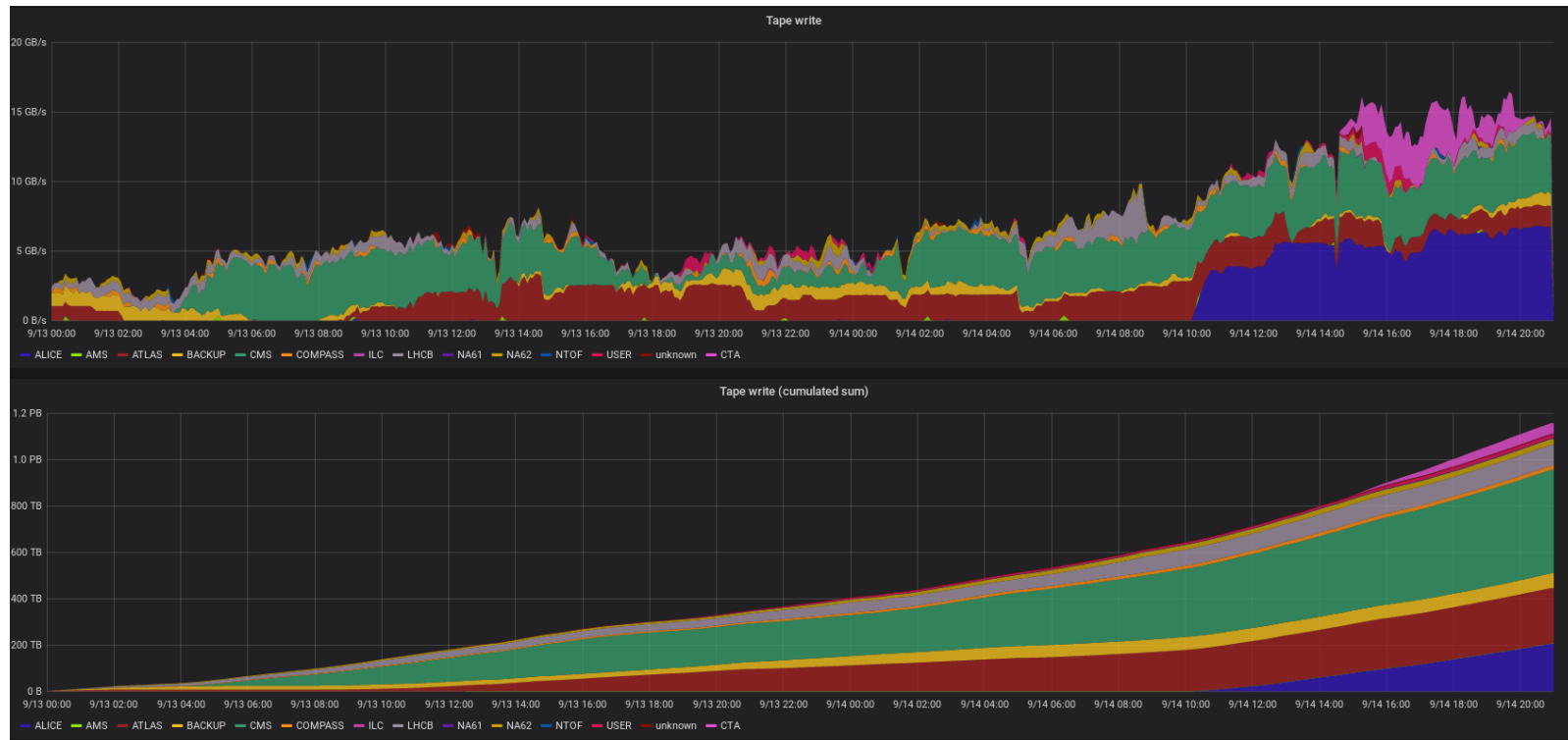
CTA metadata stress tests *CI*



CTA HI Data challenge

Bigger EOS instance

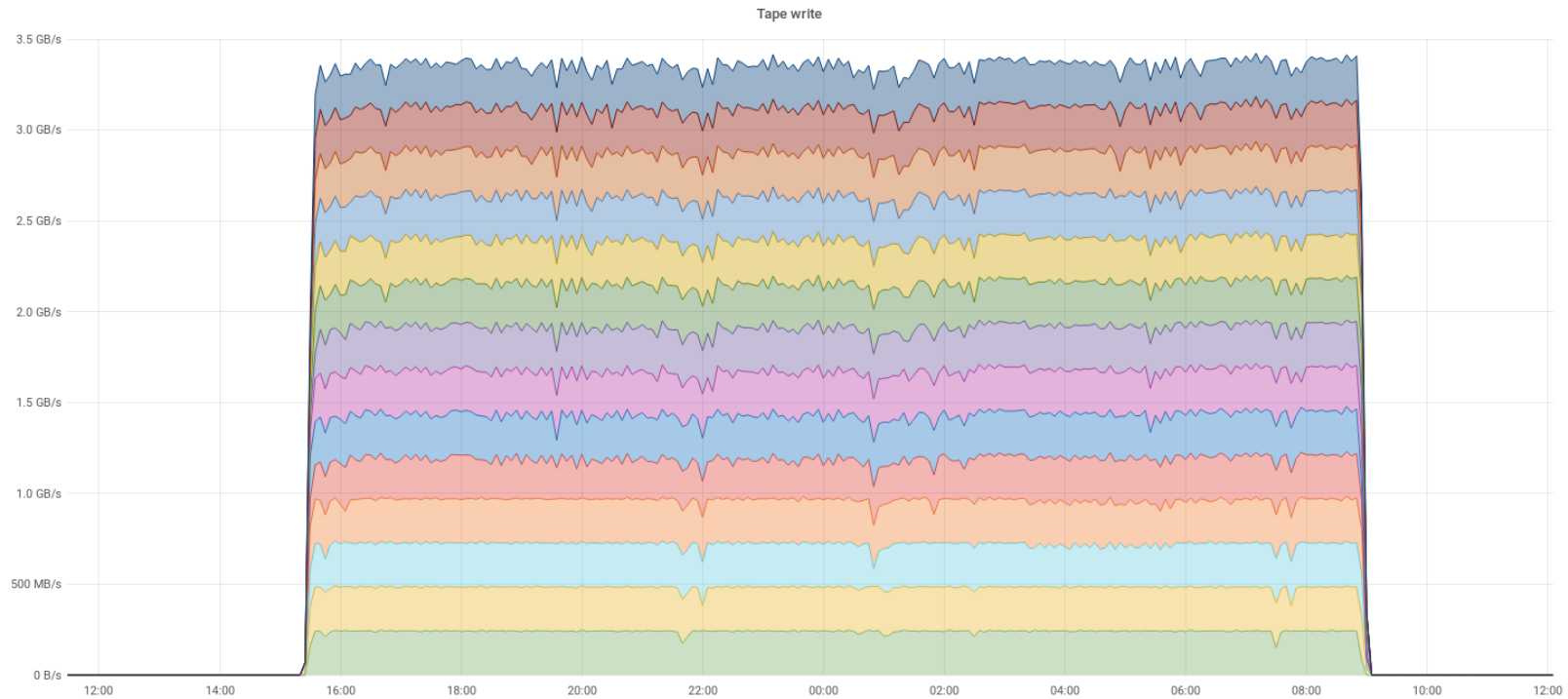
- 10x10Gb/s disk servers and a decent MGM for EOS (~2.4PB)
- 14xT10KD tape drives for the CTA tape part (3.5GB/s)



CTA HI Data challenge

Extended

**Can we saturate the tape drive
speed longer?**



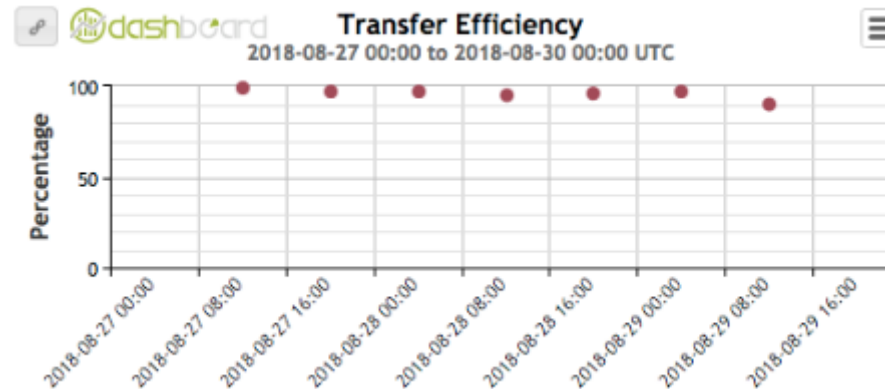
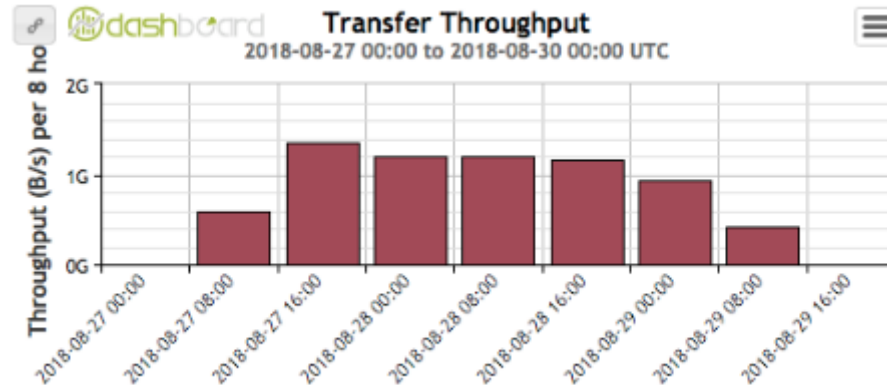
CTA VS experiment data transfers

ATLAS stage in

Several tests conducted with Atlas DDM team using Rucio and FTS.

- 2 stage in tests of 200TB each
 - ~90k files of 2.6GB archived to tape
 - sub-optimal EOS instance (2 slow disk servers)

ATLAS stage in

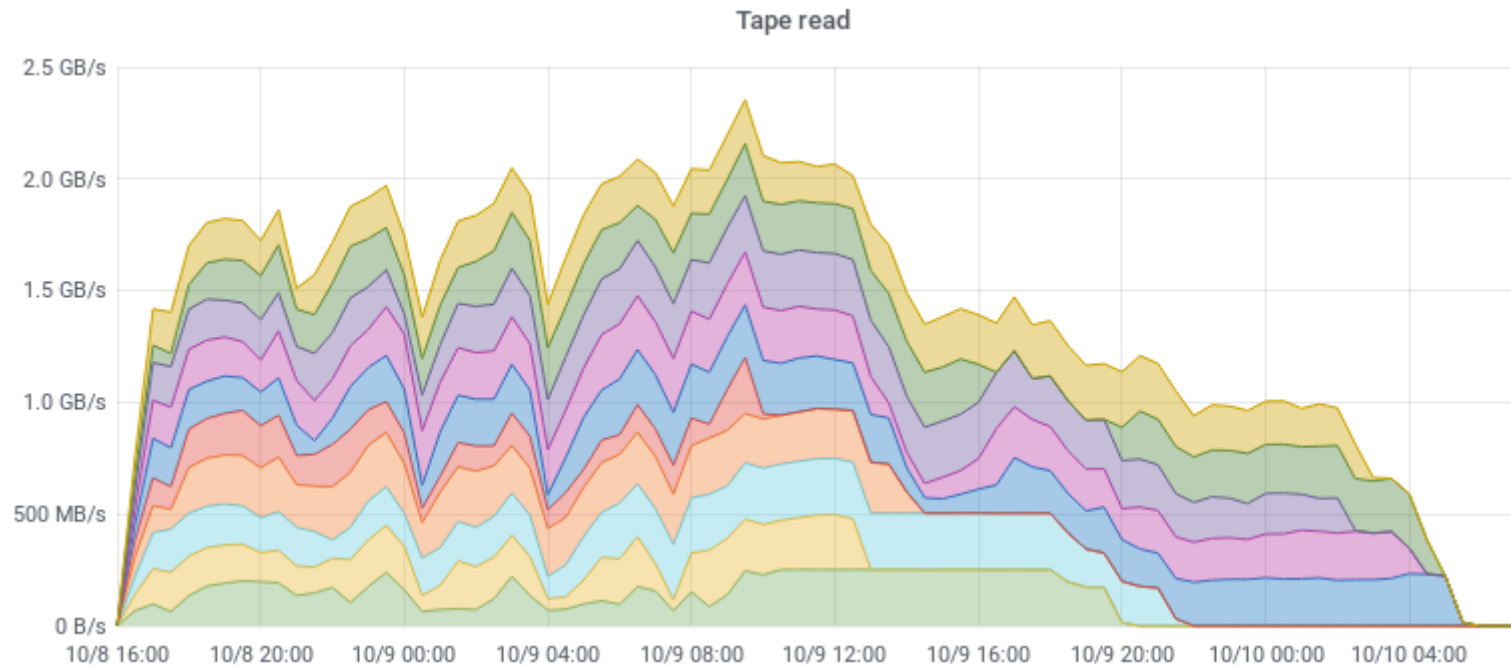


ATLAS stage out

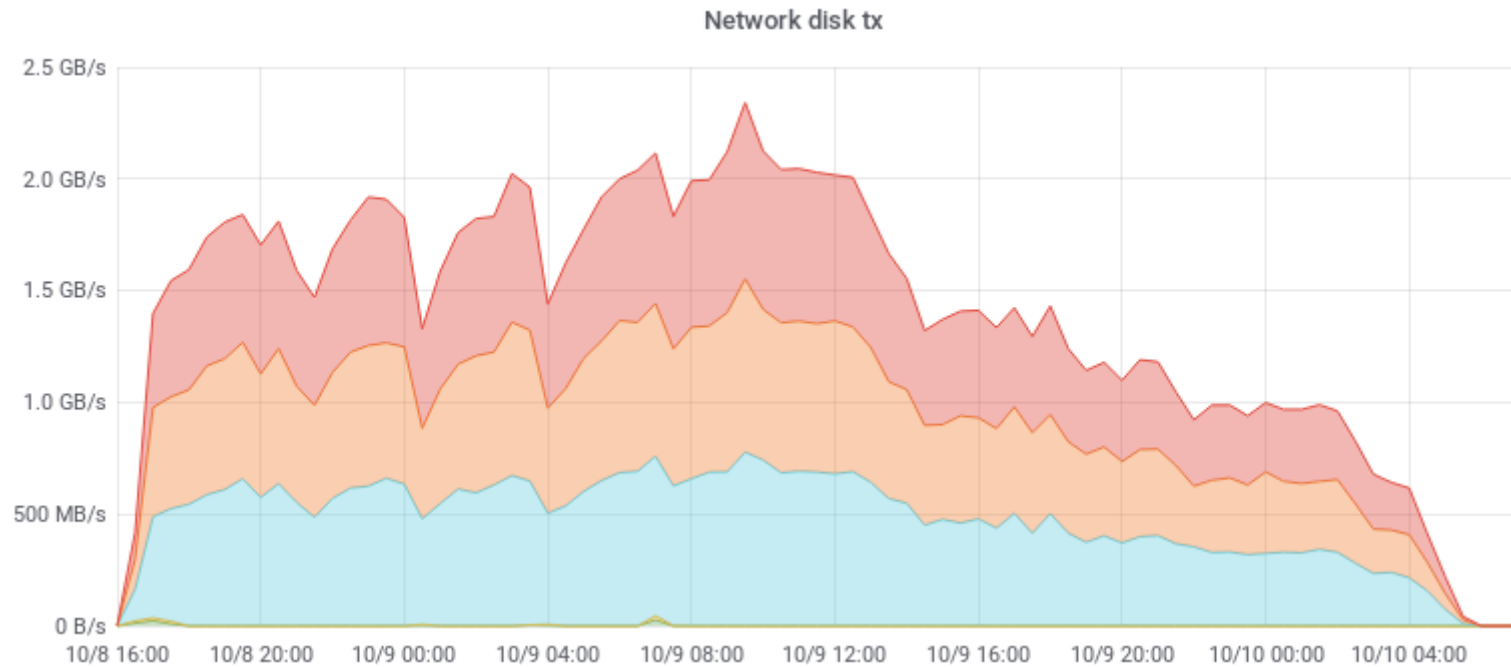
aka *Tape carousel* test took place during Hepix:

- 3 x EOS disk servers (~3x260TB of raw JBOD space)
- 6-10 x T10KD tape drives
- 90k files retrieved from EOSCTAATLASPPS (tape) to EOSATLAS by rucio through FTS

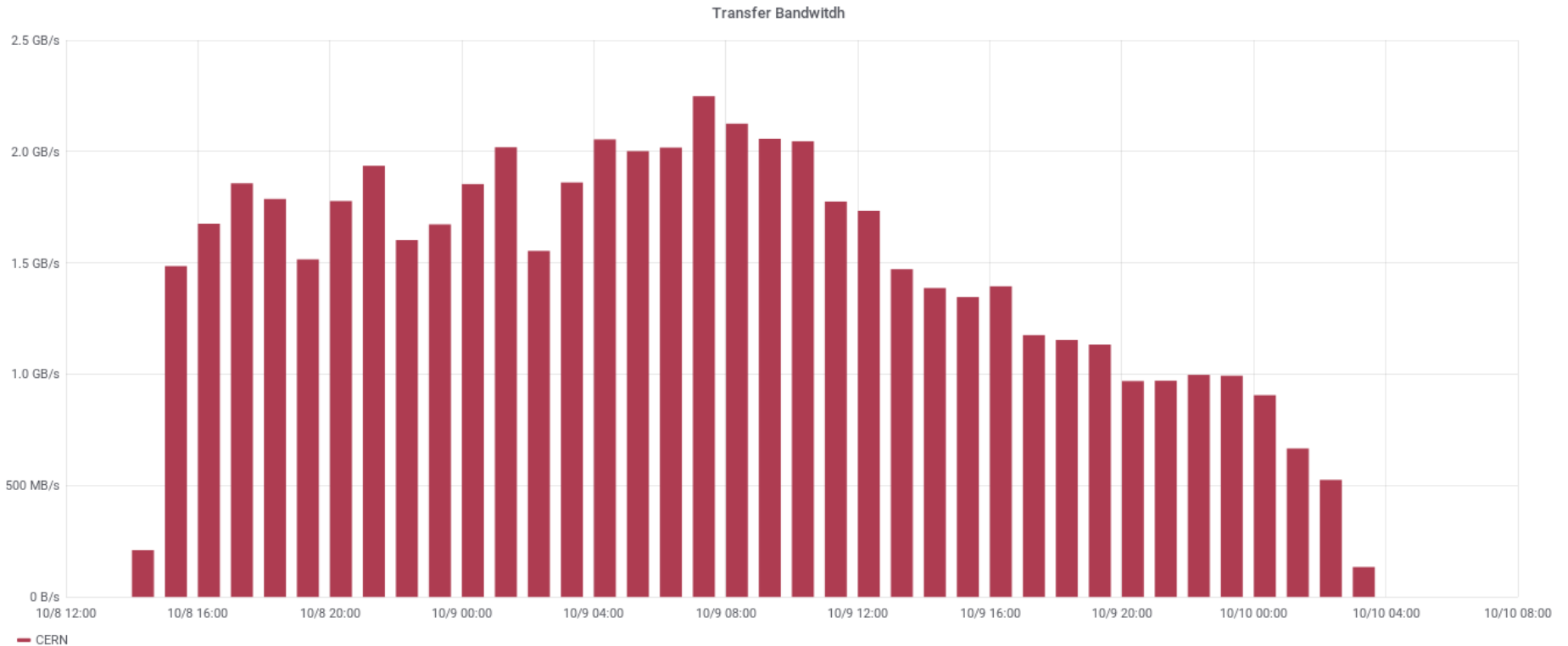
ATLAS stage out



ATLAS stage out



ATLAS stage out DDM



CTA & LHC experiments

ATLAS

- PPS instance is **up and running**
- Will push a redundant 5% share of production into CTA

CMS

- PPS instance **is up and running**
- Starting work with the Rucio team
 - no PhEDEx

ALICE

- ALICE and CERN IT already moved CASTOR behind EOS
 - During HI production DAQ went to EOSALICE first
- Agreed to provide a CTA PPS instance

LHCb

- relies heavily on SRM
 - CTA does not provide SRM
 - SRMless transfers hidden behind GFAL2
- Agreed to write DAQ data to EOSLHCb instance first

