

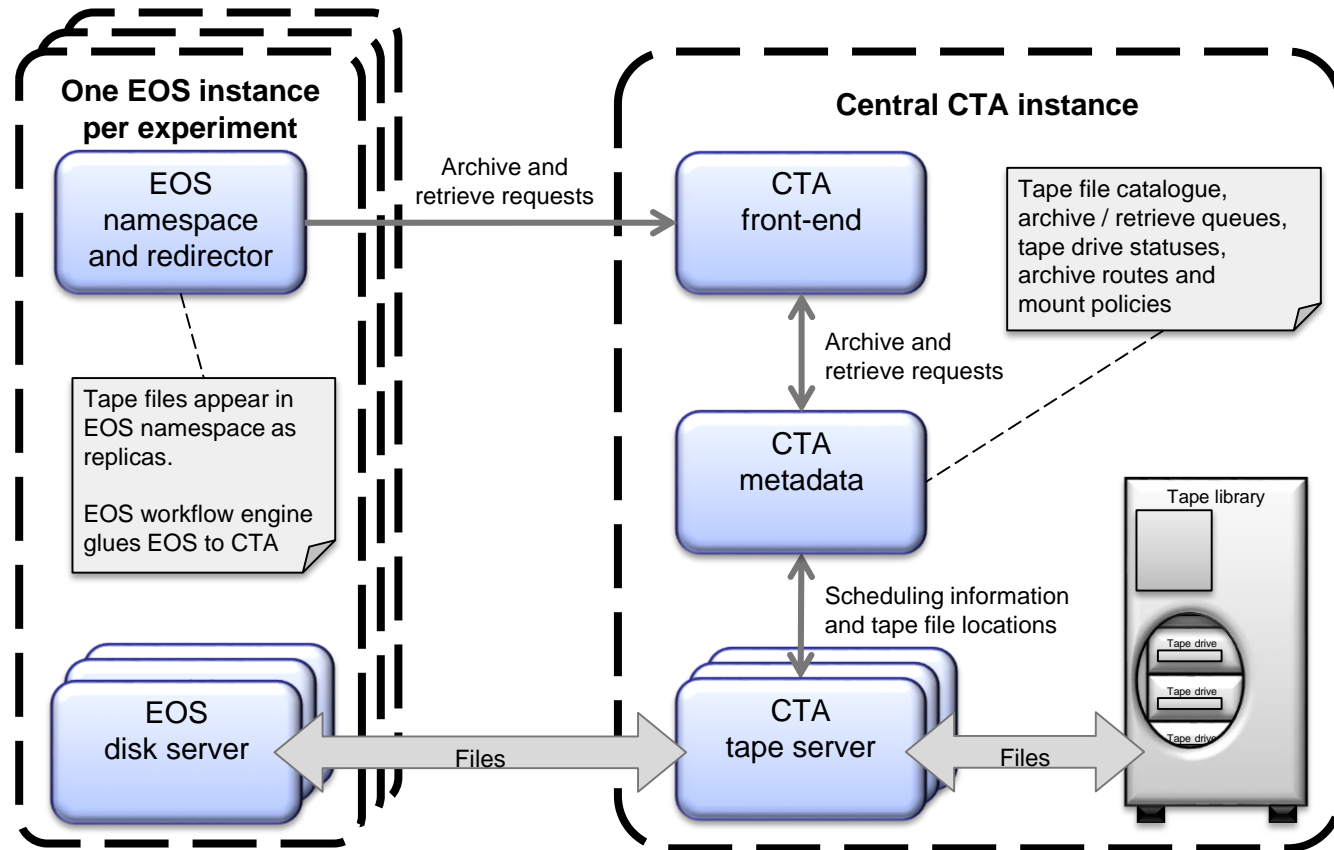
What, why and when

CTA and the Tier 1s

German Cancio, Eric Cano, Julien Leduc and Steven Murray

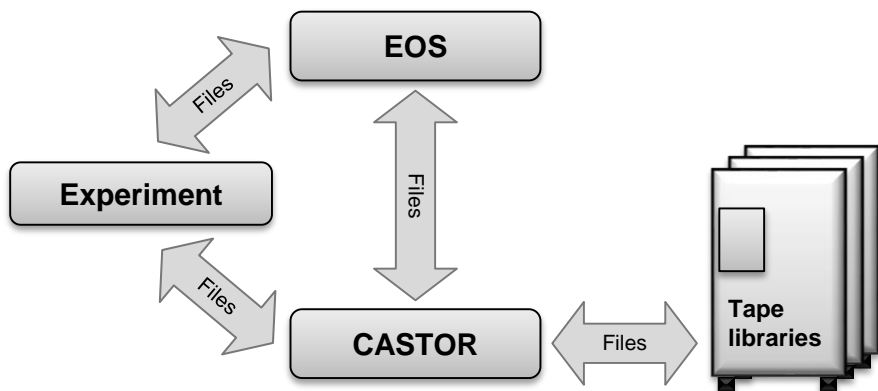
CTA is:

- Natural evolution of **CASTOR**
- A tape backend for EOS
- A preemptive tape drive scheduler
- A clean separation between disk and tape

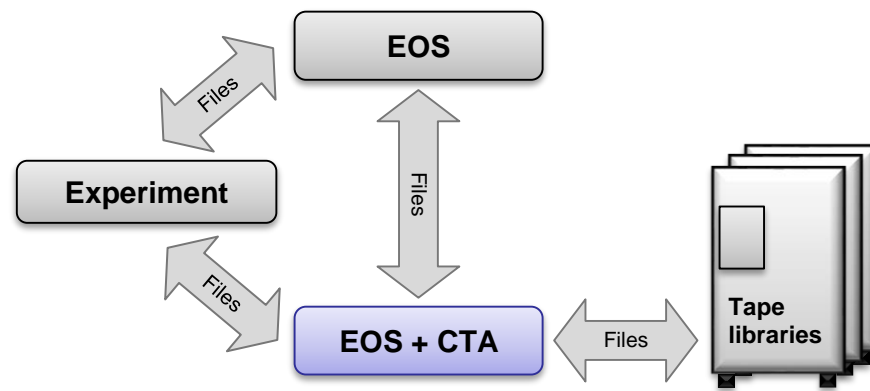


- EOS plus CTA is a “drop in” replacement for CASTOR

Current deployments with CASTOR



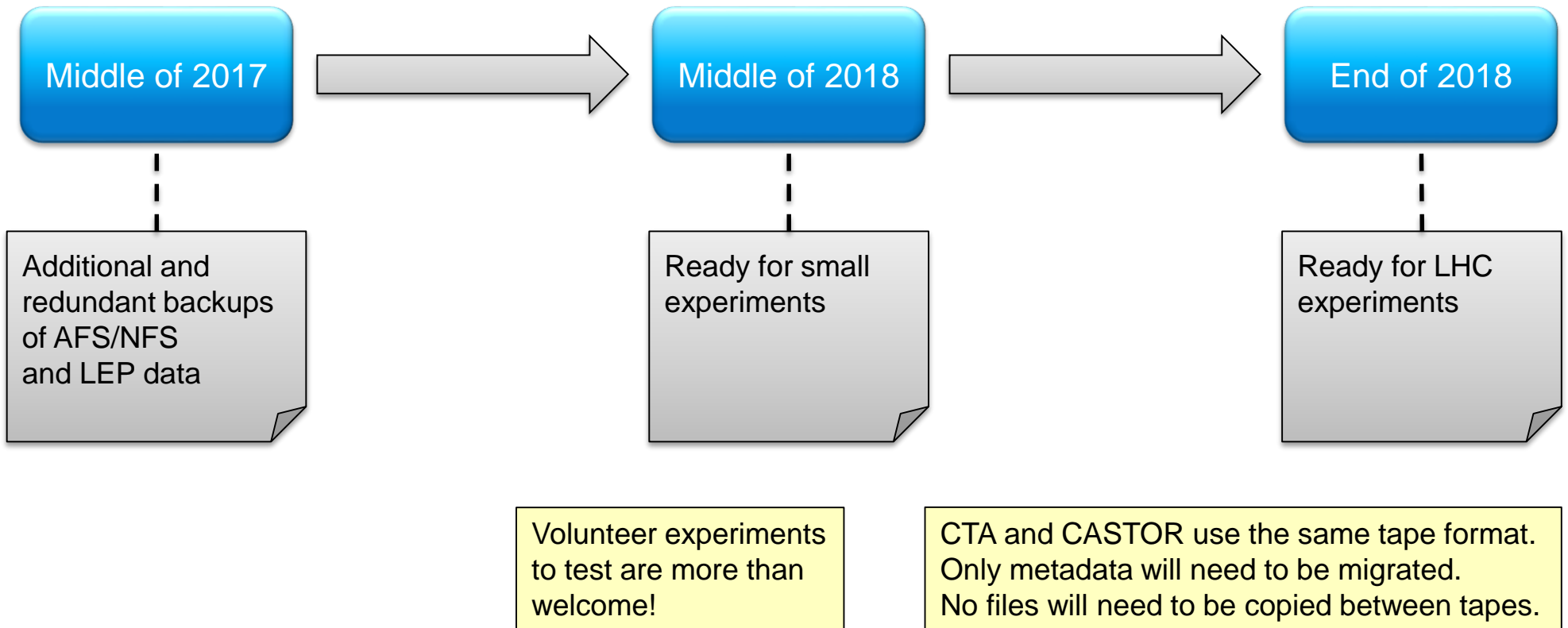
Future deployments with EOS plus CTA



- **EOS plus CTA is a “drop in” replacement for CASTOR**
 - **CASTOR like file lifecycle implemented by EOS workflow engine**
 - **Immutable tape files**
 - **Implicit archive to tape (directories that are tagged for tape)**
 - **Explicit retrieves from tape (`stager_get` replaced by `xrdfs prepare`)**
 - **Implicit retrieves from tape (open for read blocks until file is retrieved)**
 - **D1T1 – Garbage collected disk cache on top of permanent tape files**

- **EOS has become the de facto disk storage for LHC physics data**
- **Natural evolution from CASTOR**
 - **Remove duplication between CASTOR disk storage and EOS**
 - **Thin layer on top of existing CASTOR tape server**
 - **Stronger and more decoupled separation between disk and tape**

- **CTA preemptive scheduler**
 - **Use drives at full speed all of the time**
 - **Single step scheduling vs the partial step scheduling of CASTOR**
- **Same tape format as CASTOR – only need to migrate metadata**
- **Full flat catalogue of all tape files can be used for disaster recovery**
- **Less networked components than CASTOR (no CUPV, VDQM or VMGR)**



- **CTA will be usable anywhere EOS is used**
- **CTA could go behind another disk storage system if:**
 - **The disk storage system manages the disk and tape lifecycle of each file**
 - **The disk storage system can transfer files using one of the protocols supported by the CTA tape server (rfio, xrootd and Ceph object striper)**
 - **The CTA tape server can easily be modified to support other transport protocols**
- **CTA currently uses Oracle for the tape file catalogue**
 - **CTA has a thin RDBMS layer that isolates Oracle specifics from the rest of CTA**
 - **The RDMS layer means CTA could be modified to run with a different database technology**