



CERN IT-Storage Strategy Outlook

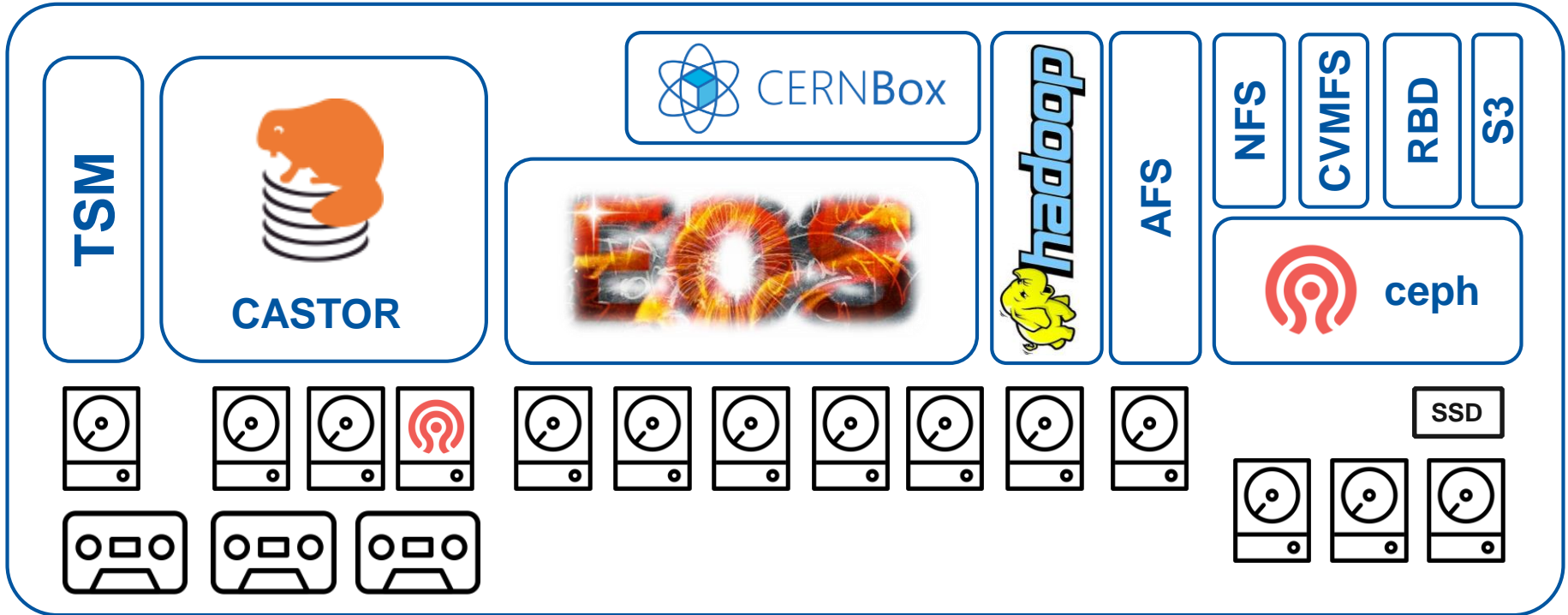
Alberto Pace, Luca Mascetti, Julien Leduc



CERN IT-Storage Group



Storage Services



Vision for CERN Storage

Build a flexible and uniform
infrastructure

Maintain and enhance (grid) data
management tools

Vision for CERN Storage

- **Unlimited storage for LHC experiments, at exabyte scale**
 - EOS Disk pools (on-demand reliability, on-demand performance)
 - currently 165 PB deployed
 - CERN Tape Archive (high reliability, low cost, data preservation)
 - currently 190 PB stored
- **A generic home directory service** – currently 3 PB
 - To cover all requirements of personal storage
 - Shared among all clients and services
 - Fuse mounts, NFS and SMB exports,
 - Desktop Sync and Web/HTTP/DAV Access

EOS Development Strategy



- Evolve **flexible** storage infrastructure as **scale-out solution** towards exabyte scale
 - one storage – many views on data – analysis/analytics service integration
 - suitable for LAN & WAN deployments
 - enable cost optimized deployments
- Evolve **file-system** interface /eos
 - multi-platform access
using standard protocol bridges CIFS, NFS, HTTP
- Evolve **namespace scalability**
 - scale-out implementation
- Evolve scale-out **storage back-end** support
 - object disks, object stores, public cloud, cold storage



EOS Challenges



1 *Remote access APIs* ⇒ ***filesystem API***



1st key project: EOS FUSE rewrite

2 *Meta data scale-up* ⇒ ***scale-out***

- *in-memory namespace* ⇒ *in-memory scale-up namespace cache + persistency in scale-out KV store*

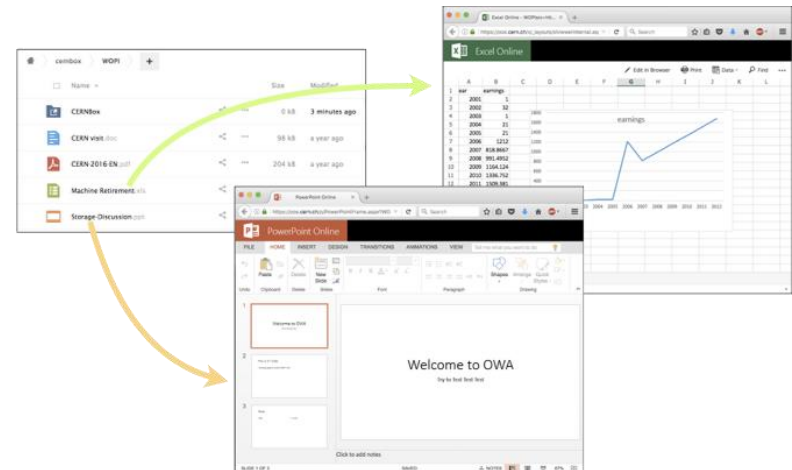


2nd key project: EOS Citrine & QuarkDB

CERNBox



- Strategic service to access EOS storage
 - Unprecedented performances
 - LHC experiments data workflows
 - Access to all CERN physics (and accelerator) bulk data
 - Disconnected / Offline Work
 - Easy Sharing model
 - Main Integration Access Point
 - SWAN, Office Online, ...
 - Checksum and consistency

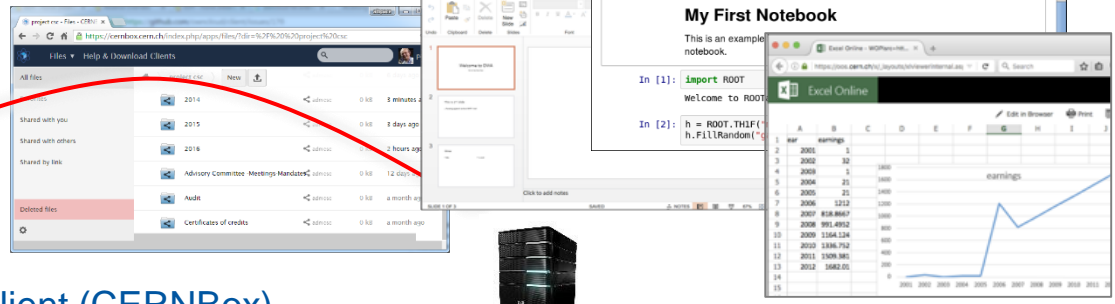


Architecture



Physicists
Engineers
Administrative staff

Web browser access

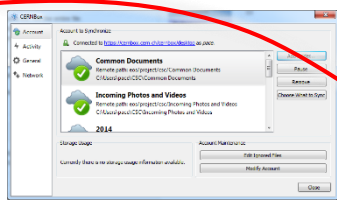


(web access)

Sync client (CERNBox)



(offline access)

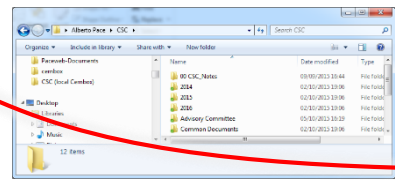


http sync protocol

CERNBOX web interface

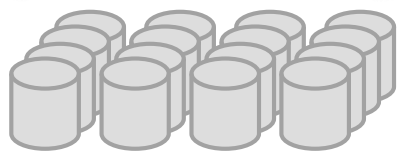


(online access)



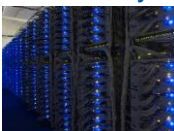
http, fuse
Webdav, NFS,
SMB, xroot

Tape
Archive



Disk Pools (EOS)

Clusters (Batch, Plus, Analytics, ...)



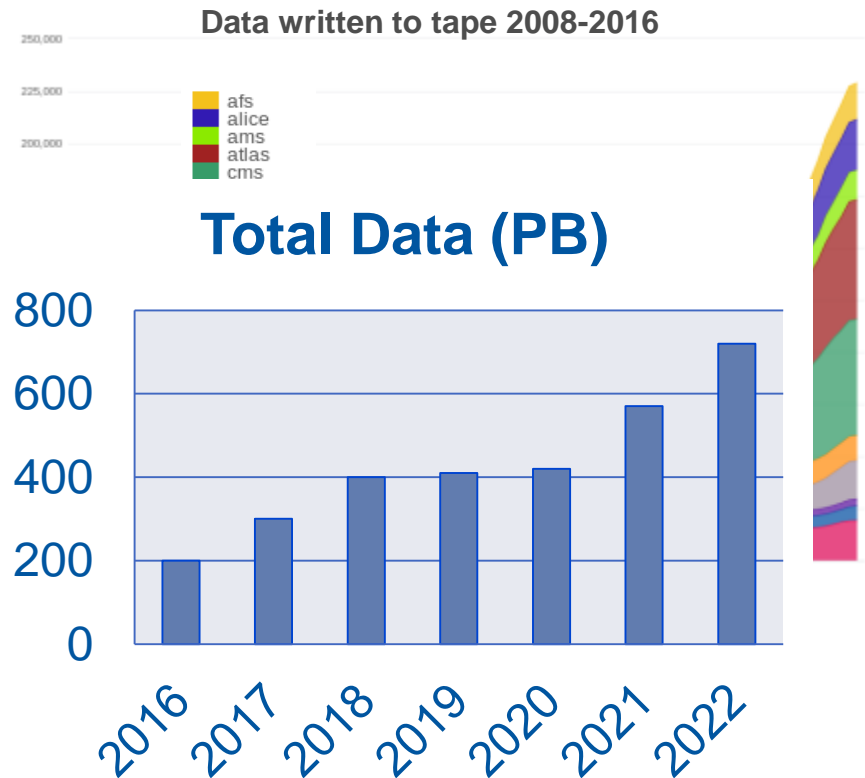
Applications
(Page 1, LHC Logging, ...)



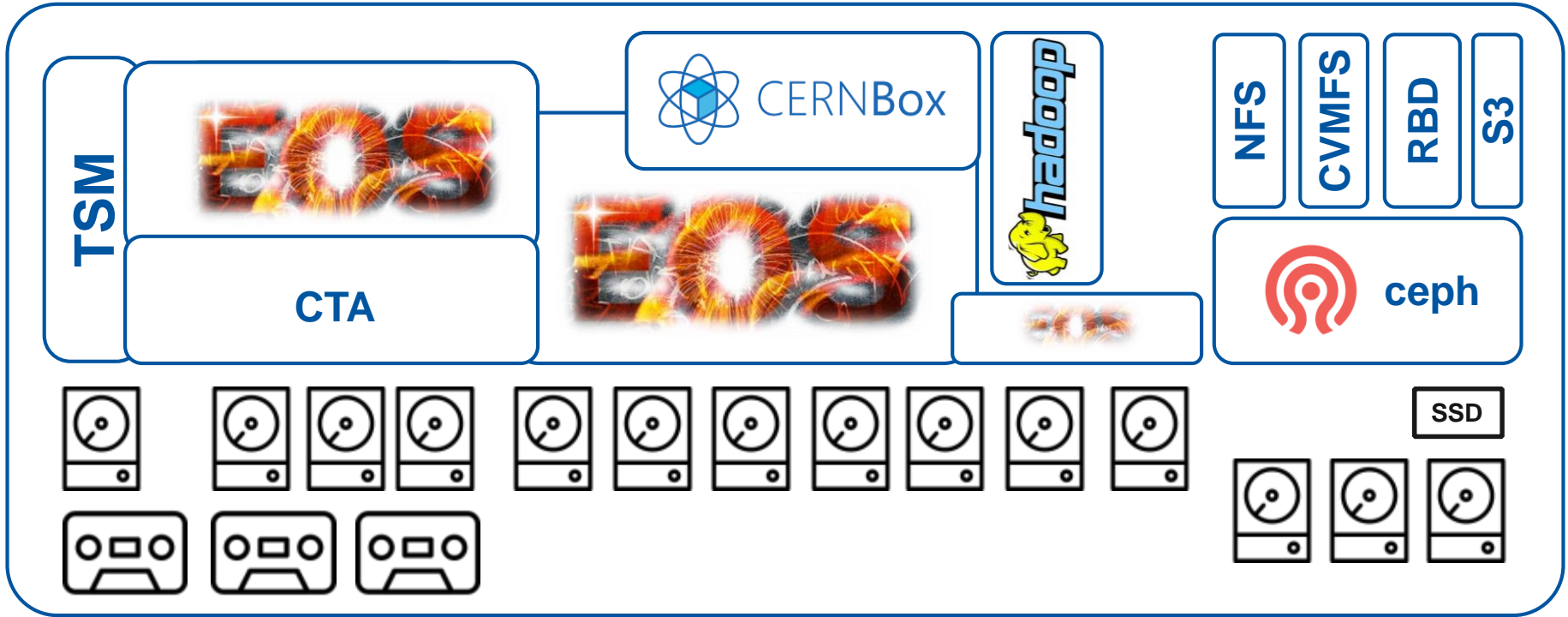
IT-Storage

Data Archiving evolution

- Preserving the past ...
 - Protect against bit rot (data corruption, bit flips, environmental elements, media wear out and breakage ...)
 - Migrate data across technology generations, avoiding obsolescence
 - Some of our data is 40 years old
- ... and anticipating the future
 - 2017-2018: 100PB/year
 - 2021++ : 150PB/year



Storage Services Evolution



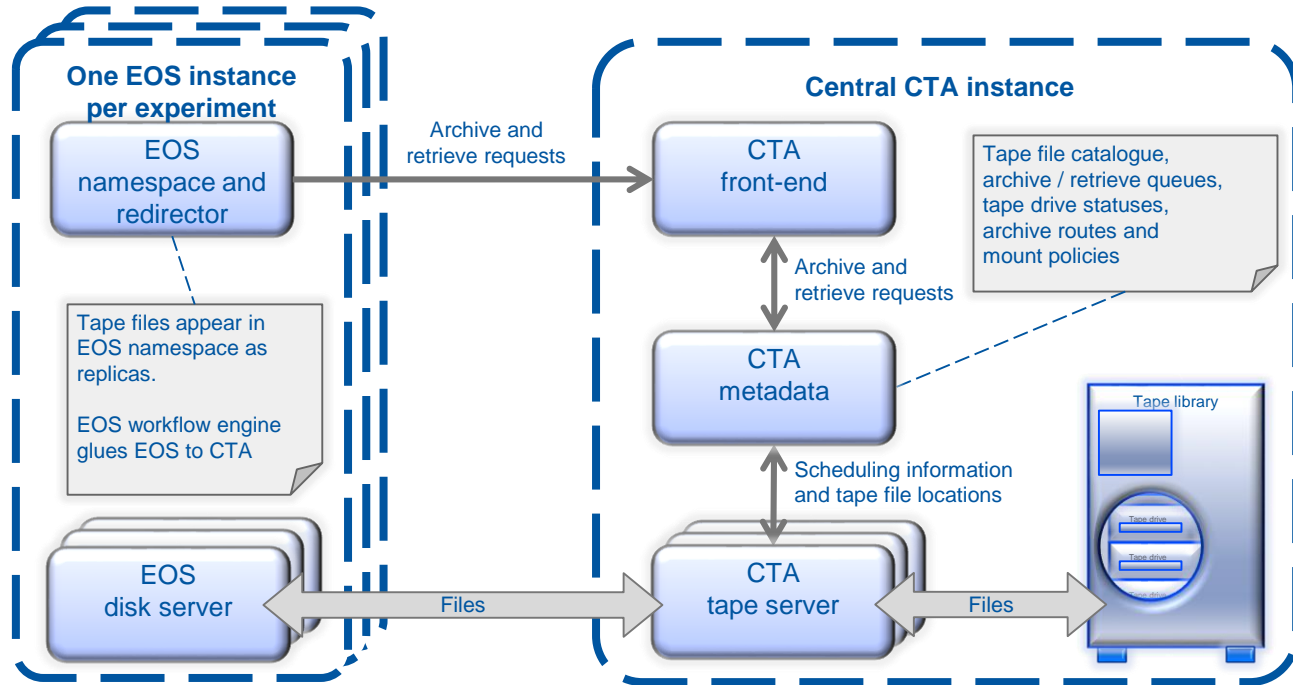
Data Archiving: Evolution

- EOS + tape ...
 - EOS is the strategic storage platform
 - Tape is the strategic long term archive medium
- EOS + tape = 
 - Meet CTA : the CERN Tape Archive
 - Streamline data paths, software and infrastructure

What is CTA?

CTA is

- A tape backend for EOS
- A preemptive tape drive scheduler
- A clean separation between disk and tape

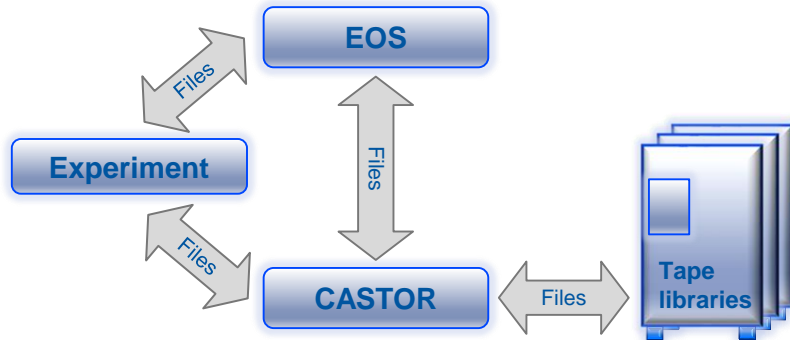


What is CTA?

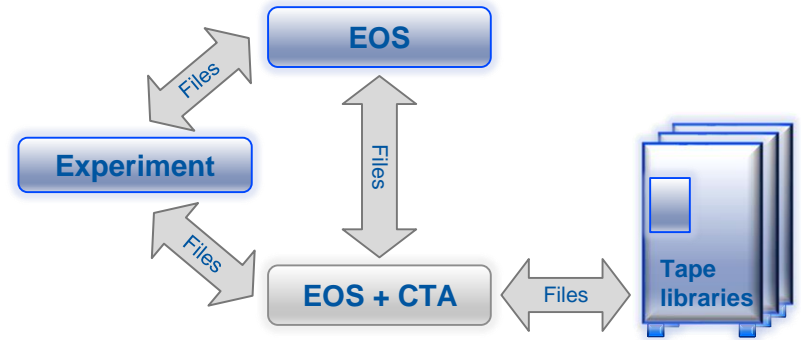
EOS plus CTA is a “drop in” replacement for CASTOR

- Same tape format as CASTOR – only need to migrate metadata

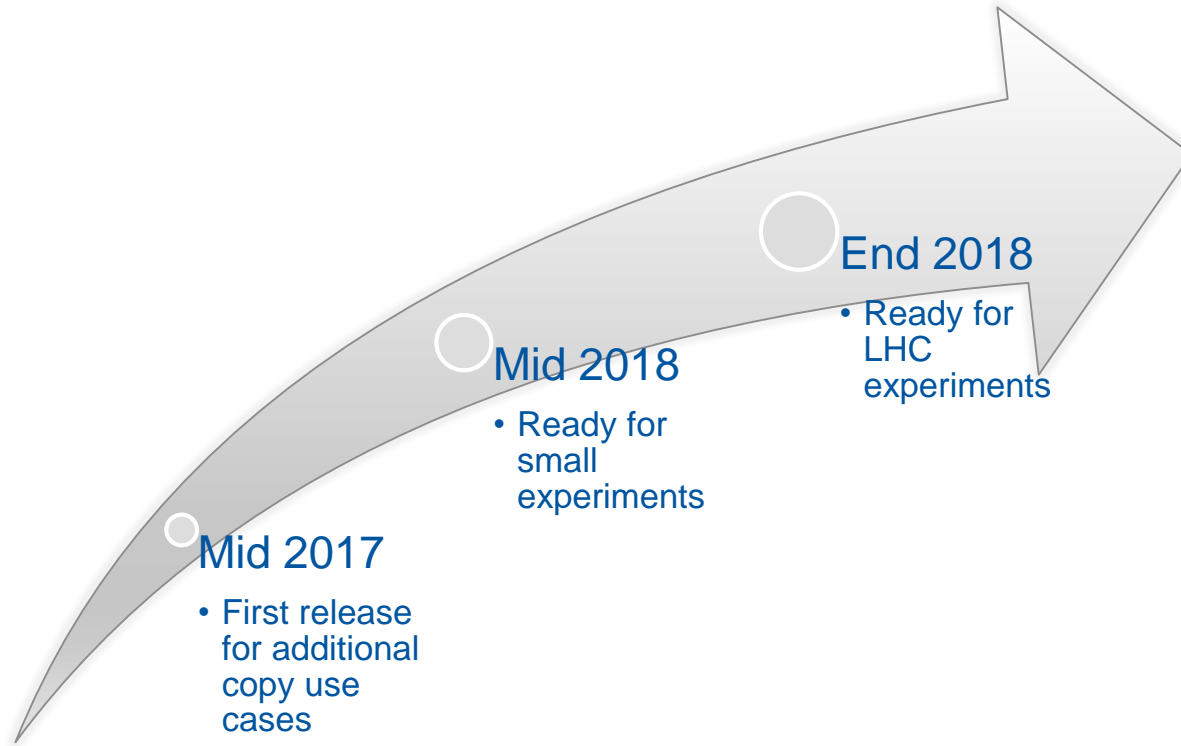
Current deployments with CASTOR



Future deployments with EOS plus CTA



When is CTA?

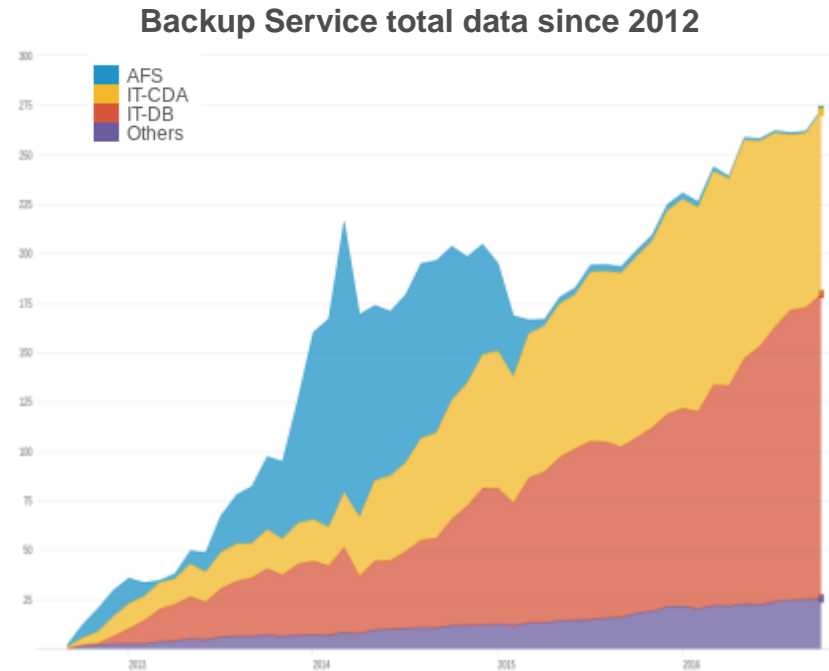


CTA outside CERN

- CTA will be **usable anywhere EOS is used**
- CTA could go behind another disk storage system but would require contributing some development:
 - The disk storage system must manage the disk and tape lifecycle of each file
 - The disk storage system needs to be able to transfer files from/to CTA (currently xroot - CTA can easily be extended to support other 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 extended to run with a different database technology

Backup Service

- TSM cost is proportional to backed up volume
 - Minimize backup volume in view of upcoming license retendering
- Ongoing efforts to limit annual growth rate: 24% (50% in the past):
 - Moved AFS backups to CASTOR
 - Working with the two other large backup customers



Question?



www.cern.ch