



# CERN Storage Evolution

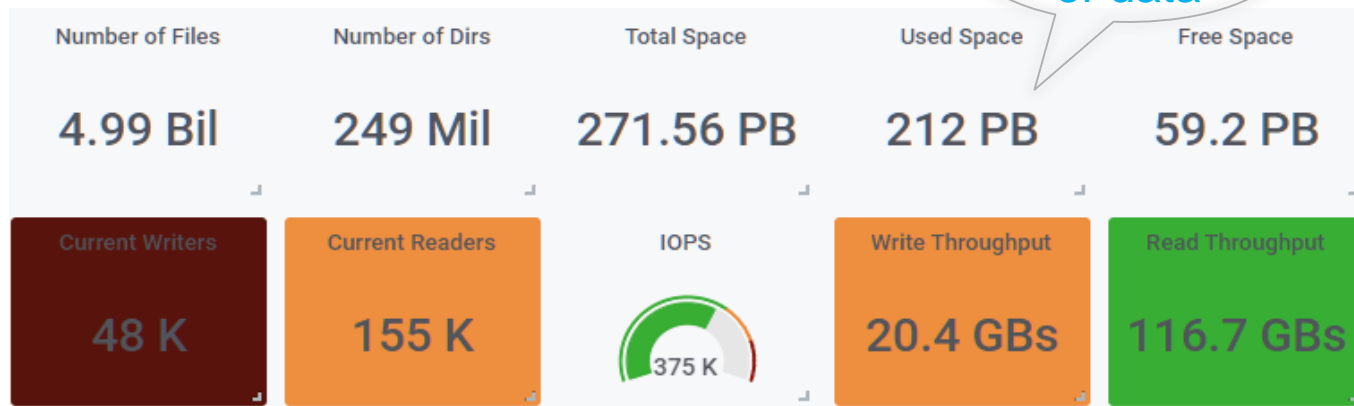
Giuseppe Lo Presti

*on behalf of the CERN IT Storage Group*



# Outline

- Disk Storage Evolution Overview
  - For Physics Data: EOS, CASTOR/CTA
  - For the IT infrastructure: Ceph, CephFS, S3
  - For the Scientific Community: AFS, CERNBox
- Deep dive into selected topics
  - Tape Archive metadata migration (CASTOR to CTA)
  - Storage for *MALT*
  - A new EU Project in the horizon



2019:

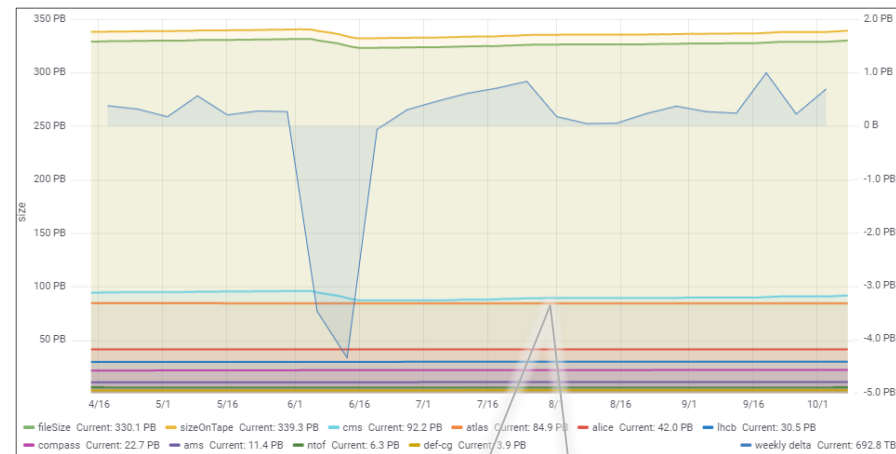
- Minimal space increase (no data-taking), yet large throughput increase
- **65 PB** replicated because of the Wigner Data Centre decommissioning
- **+70 PB** new and repatriated capacity coming soon

## EOS Instances

- 5 for LHC experiments
- EOSPUBLIC: non-LHC experiments
- 6 for CERNBox (including EOSBACKUP)
- 3 for Project Spaces
- EOSMEDIA: photo/video archival
- EOSUp2U: Pilot for Education and Outreach

# CASTOR

- **330 PB** of data (+10 PB dual copy)
- ~0.6 EB capacity, *details in Steve's talk*
- Minimal space increase
- “Only” 16 PB of disk cache, 100% on CentOS 7



ATLAS and CMS have nearly 100 PB each

- Preparing the migration to **CTA**
- *More later*

CERN Ceph Clusters		Size	Version
OpenStack Cinder/Glance	<i>Production</i>	6.4 PiB	<i>mimic</i>
	<i>Wigner</i>	<del>4.6</del> 0 PiB	<i>nautilus</i>
	<i>Hyperconverged</i>	245 TiB	<i>mimic</i>
CephFS (HPC+Manila)	<i>Production</i>	1.09 PiB	<i>luminous</i>
	<i>Preproduction</i>	164 TiB	<i>mimic</i>
	<i>Hyperconverged</i>	356 TiB	<i>mimic</i>
<b>CASTOR</b>	<i>Tape Disk Buffer</i>	5.3 PiB	<i>nautilus</i>
<b>S3+SWIFT (CVMFS)</b>	<i>Production (4+2 EC)</i>	1.92 PiB	<i>luminous</i>

Decommissioned

# AFS (Phaseout)

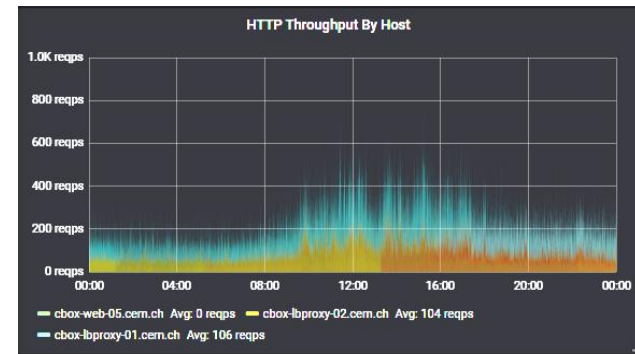


- **Reduced concerns about upstream OpenAFS longevity**
  - No specific urgency to replace current AFS service, it will be supported as necessary
  - **BUT: we still don't see long-term improvements for OpenAFS**
    - Therefore we will keep shrinking down the AFS service to the necessary minimum
- Several usages of AFS at CERN can profit from significant improvements when migrated to alternative storage services
  - For such cases we will continue to propose to data owners to opt-in for other solutions
- Requests for new AFS spaces (beyond home directories) will be examined whether the data should be hosted elsewhere
- CERN-IT will further review the long-term plan for remaining use-cases

# CERNBox



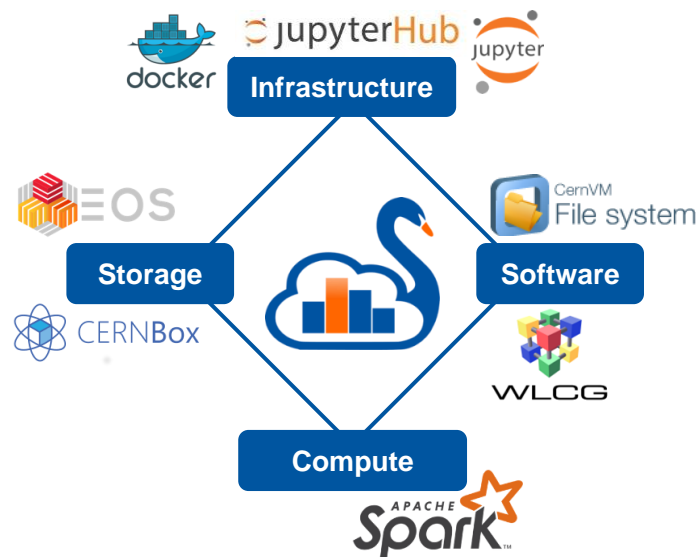
- 5 years of production, still growing service
  - 17K users, **~5K active daily**
  - 3.5+ PB, 1B+ files, **113K shares**
  - Migration to new EOS multi-instance infrastructure almost completed
- New core daemon (**Reva**) in production (collaboration with ownCloud)
  - *More later*



# SWAN



- Turn-key data analysis platform
  - Accessible from everywhere via a web browser
  - Support for ROOT/C++, Python, R, Octave
- Fully integrated in CERN ecosystem
  - Storage on EOS, Sharing with CERNBox
  - Software provided by CVMFS
  - Massive computations on Spark
- Generic docker-based deployment
  - Adopted by e.g. AARNET (Australia)



*SWAN Users Workshop last Friday:*  
<https://indico.cern.ch/event/834069/>



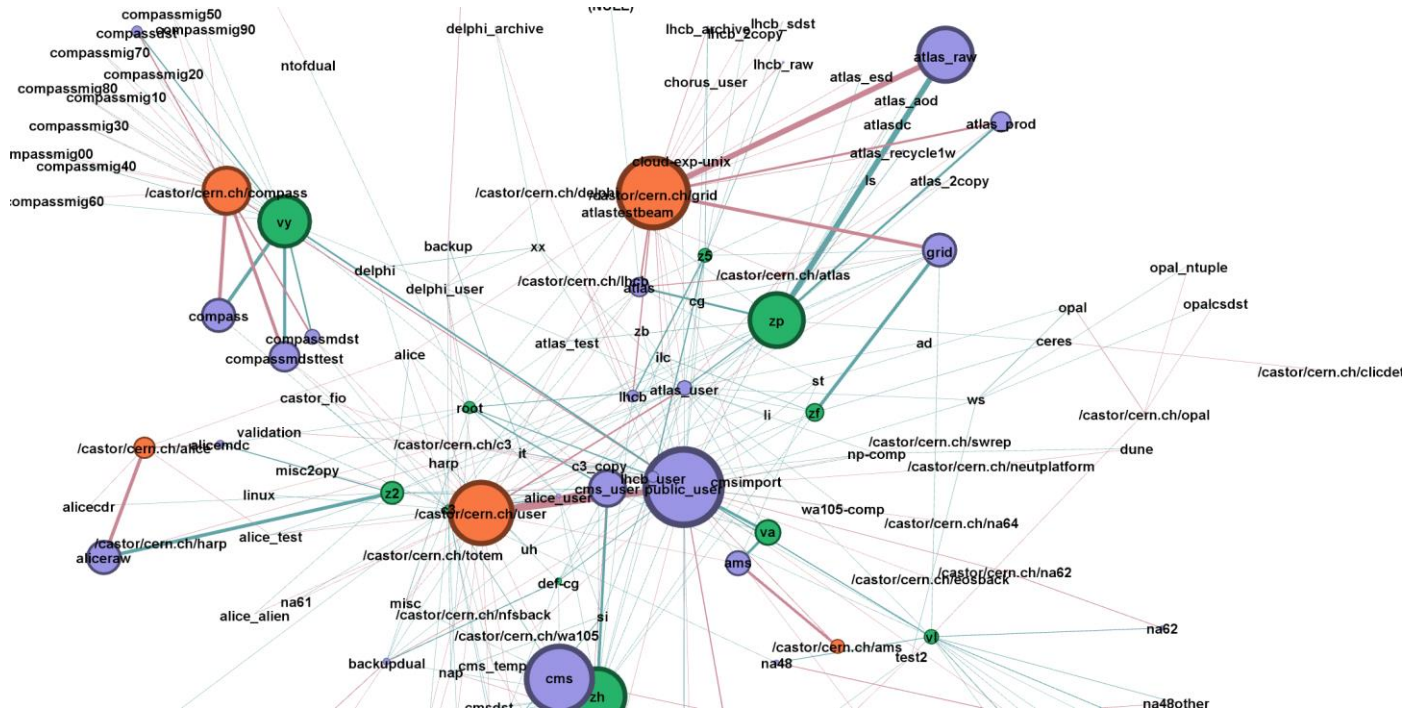
# Deep dive into selected topics

- Tape Archive metadata migration
- Storage for *MALT*
- New EU Project in the context of EOSC

# CASTOR to CTA Migration

- A.k.a. how to migrate 300+ PB without physically moving a byte (...almost)
  - Developed tools to export/import all CASTOR metadata
    - To 4+1 EOS instances (Oracle to QuarkDB) for the namespace part
    - To a single CTA catalogue (Oracle to Oracle) for the tape-related metadata
  - Namespace analysis being done with the experiments
    - 20+ years of accumulated legacy often requires their agreement as for what / how to migrate
- Proven to work at the required scale
  - The entire ATLAS metadata (~**80M** files) was exported at **2 kHz** in about half day
- Ready to migrate all experiments

# CASTOR to CTA Migration



Credits:  
Michael Davis

# Storage for MALT



- MALT: CERN/IT globally moving to Open Source
  - Replacing proprietary software wherever possible
- Mail service: pilot based on **Kopano**, approaching production (*cf. Thomas' talk*)
  - **CephFS** for the mail attachments
    - Foreseen integration also with **CERNBox**
  - OpenStack + **Ceph** hyperconverged infrastructure



# Storage for MALT

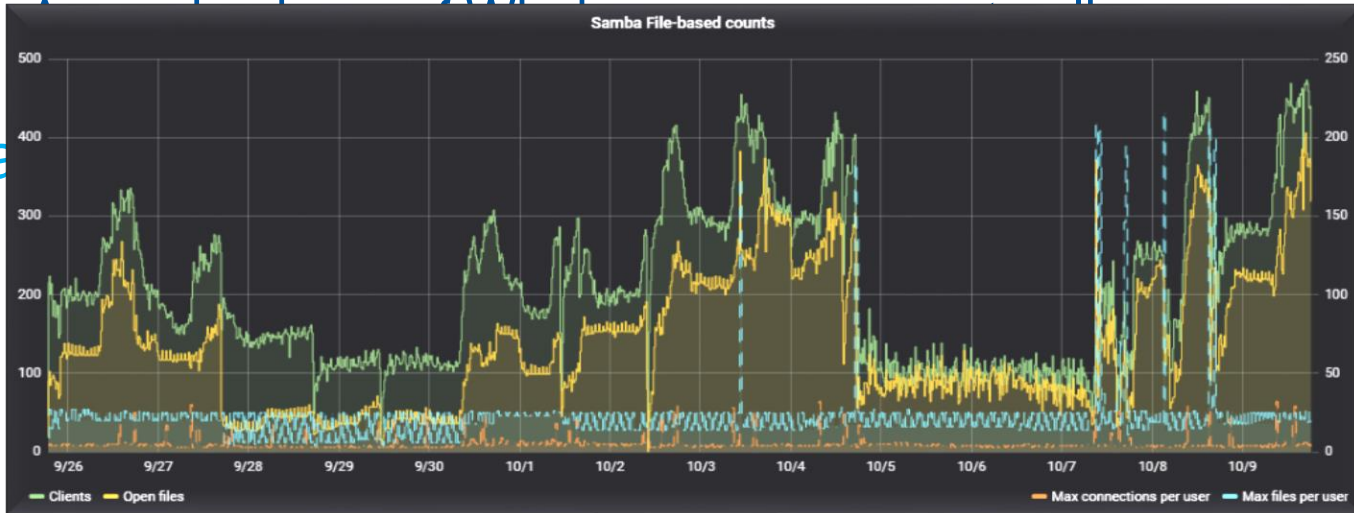


- Windows DFS going to be decommissioned
  - Users being migrated to **CERNBox**, *cf. Vincent's talk*
  - A growing base of Windows users expect online (Samba-based) access to EOS
- Samba gateways from prototype to production
  - High Available cluster in place since Sep 2019
  - Monitoring and alarming + performance analysis to identify shortcomings of the SMB/CIFS protocol
  - CERN Terminal Service permanently connected to it as of Oct 2019

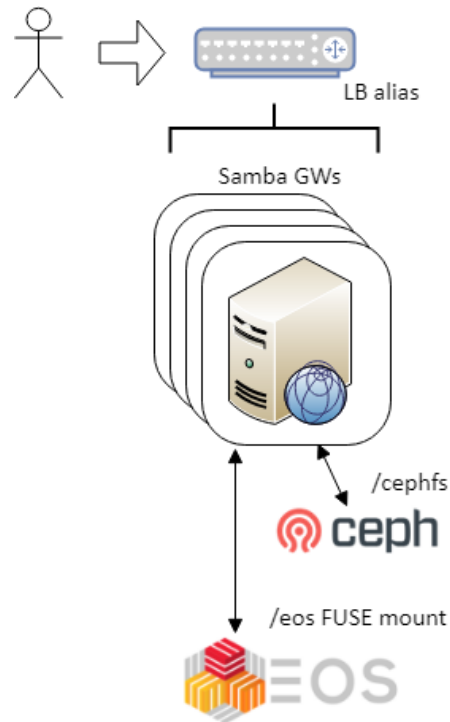
# Storage for MALT



- Windows DFS going to be decommissioned
  - Users being migrated to **CERNBox**, cf. Vincent's talk



# HA Samba Architecture



- The clustered Samba software requires
  - *Floating IPs*, one per host
  - A small `/cephfs` mount to share the state
    - Yes, to provide a shared filesystem we need a(nother) shared filesystem...
- Windows Domain (AD) joined in **dedicated keytab** mode, using the **credentials** of the load-balanced alias
- Forbidden to join with hosts credentials, because they are resolved with their own IPs, not the floating ones!
- EOS FUSE mounted
- Fully puppetized configuration

# Storage for MALT: Apps

- **CERNBox** is the Application Hub for Office productivity and more
  - New apps being proposed in collaboration with IT/CDA





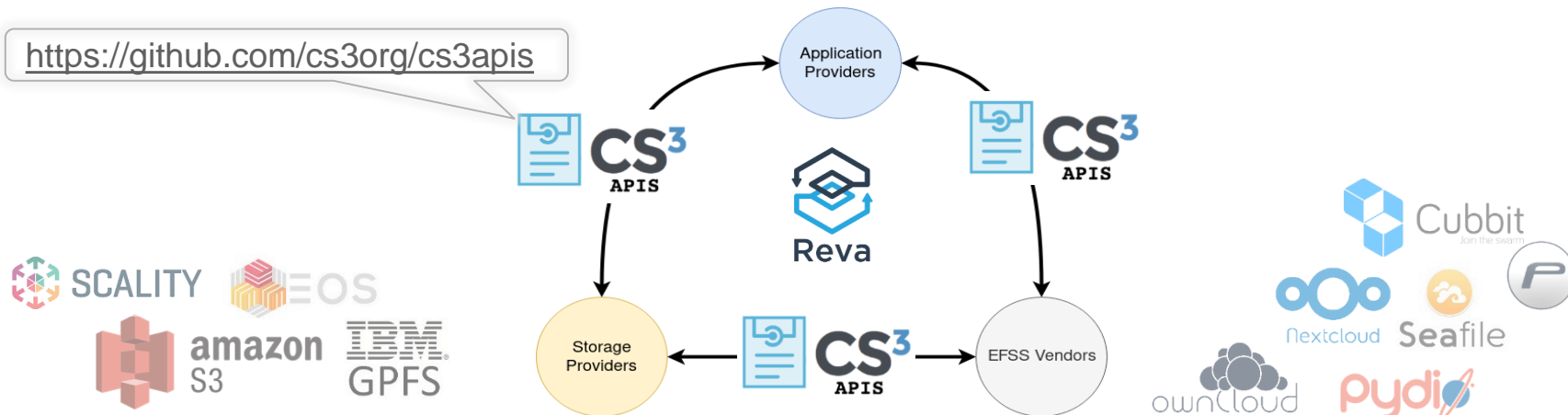
# CERNBox: a multi-purpose platform



- Multi-apps is only one dimension ...
- Expanding CERNBox towards a multi-storage, multi-vendors, multi-apps **mesh**
  - Interoperability ensured by the **CS<sup>3</sup> APIs**, currently being defined



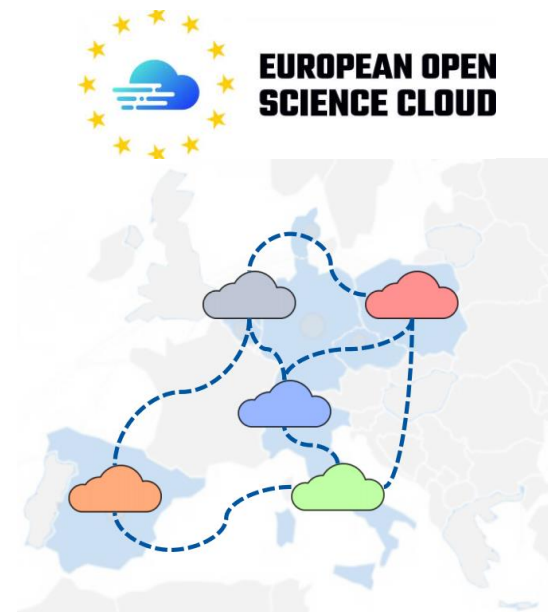
<https://github.com/cs3org/cs3apis>



# CS3Mesh: a new EU Project



- Federation of existing CS3 sites
  - **30+ sites** (e.g. CERNBox, DesyBox, GarrBox, universities...)
  - 300K+ users
- Global collaborative environment for research
  - Within EOSC
  - Share documents, files, projects, data, ...
  - Connected application hubs
  - Data/metadata-aware workflows
    - FAIR: Find, Access, Interoperate, Reuse
- EU-funded project
  - **6MEUR, 12 partners, 2020-2022**



# CS3Mesh in HEP



- CS3Mesh puts users and collaborative workflows/concurrent editing at the center
  - Users in the HEP community will be able to more easily collaborate on common projects
    - Shared documents, Jupyter notebooks-based analyses, access to data, etc.
  - Functionality available to all users with CERN accounts AND all users of participating sites (e.g. SURFsara T1)



# CS3 Workshops



Contact  

## Cloud Storage Services for Synchronization and Sharing (CS3)

This is a community of providers, developers and users of innovative storage and sync&share systems. The CS3 services are integrated with user environments and higher-level application services. CS3 reports on the progress in data science at all levels: local laboratories, regional collaborations and global science. CS3 applications range from innovative big-data analysis to science outreach and education.

### Conferences organized



**2019**

Rome, IT  
Conference Organized by  
INFN

DOI [10.5281/zenodo.2545482](https://doi.org/10.5281/zenodo.2545482)

[Website](#)

[Programme](#)



**2018**

Krakow, PL  
Conference Organized by  
Cyfronet

DOI [10.5281/zenodo.1157141](https://doi.org/10.5281/zenodo.1157141)

[Website](#)

[Programme](#)



**2017**

Amsterdam, NL  
Conference Organized by  
SURFSara

DOI [10.5281/zenodo.254064](https://doi.org/10.5281/zenodo.254064)

[Website](#)

[Programme](#)



**2016**

Zurich, CH  
Conference Organized by  
ETH Zurich

DOI [10.5281/zenodo.44783](https://doi.org/10.5281/zenodo.44783)

[Website](#)

[Programme](#)



**2014**

Geneva, CH  
Conference Organized by  
CERN

DOI [10.5281/zenodo.2546420](https://doi.org/10.5281/zenodo.2546420)

[Website](#)

[Programme](#)

5 editions since 2014

### Last edition – Rome:

<http://cs3.infn.it/>

- 55 contributions
- 147 participants
- 70 institutions
- 25 countries

### Industry participation:

- Start-Ups: Cubbit, pydio, ...
- SMEs: OnlyOffice, ownCloud
- Big: AWS, Dropbox, ...

### Community website:

<http://www.cs3community.org/>



# CS3 Workshop, 6<sup>th</sup> edition

CS3 COMMUNITY



<http://cs3.deic.dk/>

Workshop on **Cloud Services**  
for **Synchronisation** and **Sharing**

27-29 January 2020  
Copenhagen, Denmark

Abstracts submission  
open, register to the  
conference!





Thank you! Questions?