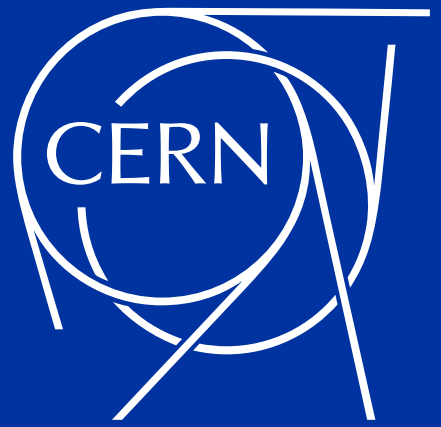


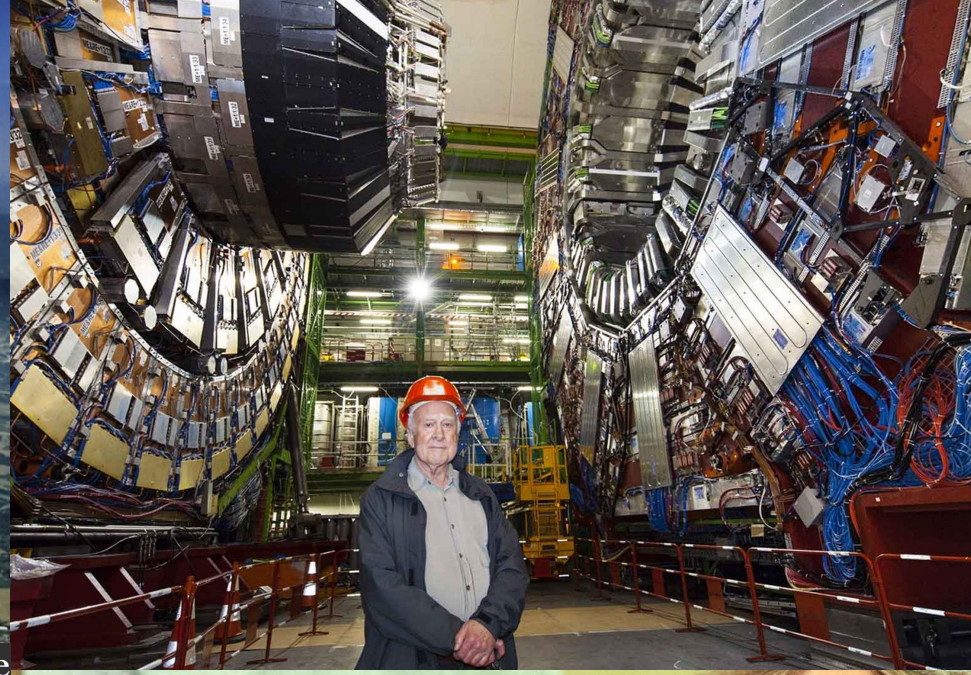
Site Status Report

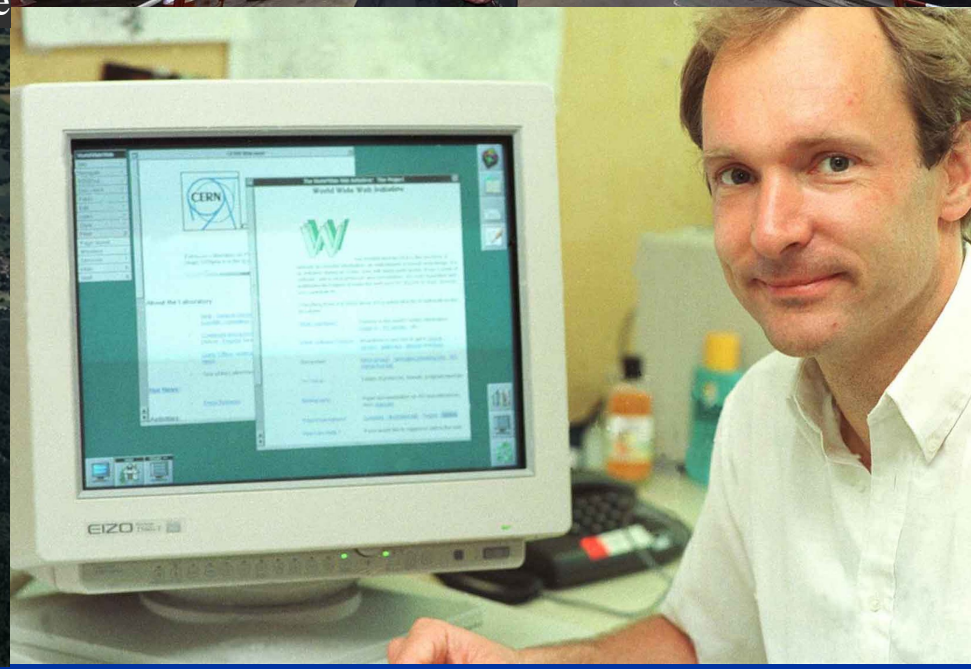
2022 AFS Technologies Workshop

Diogo Castro
diogo.castro@cern.ch

14/06/2022

Introduction

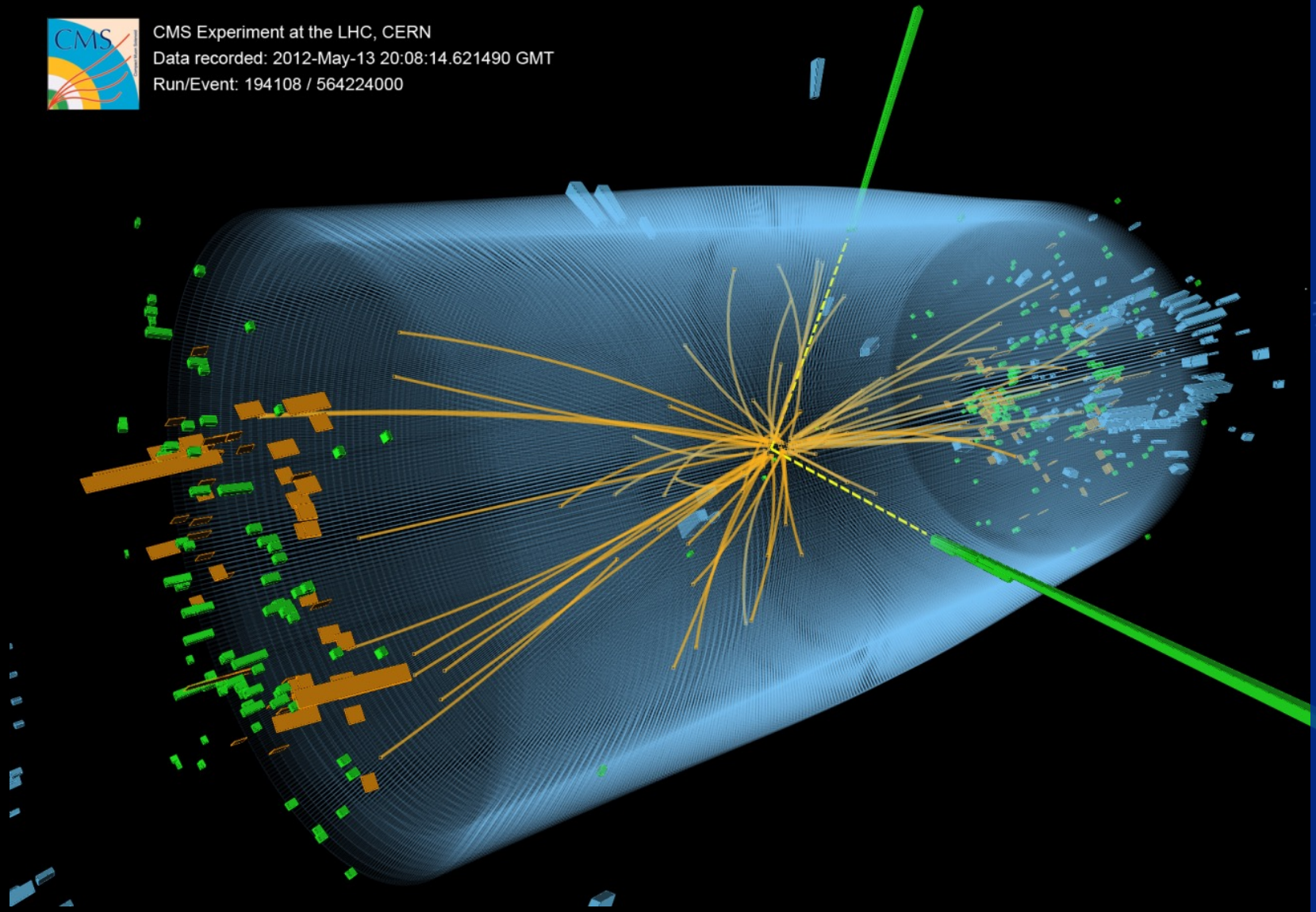




The Large Hadron Collider

Discovery of the Higgs boson

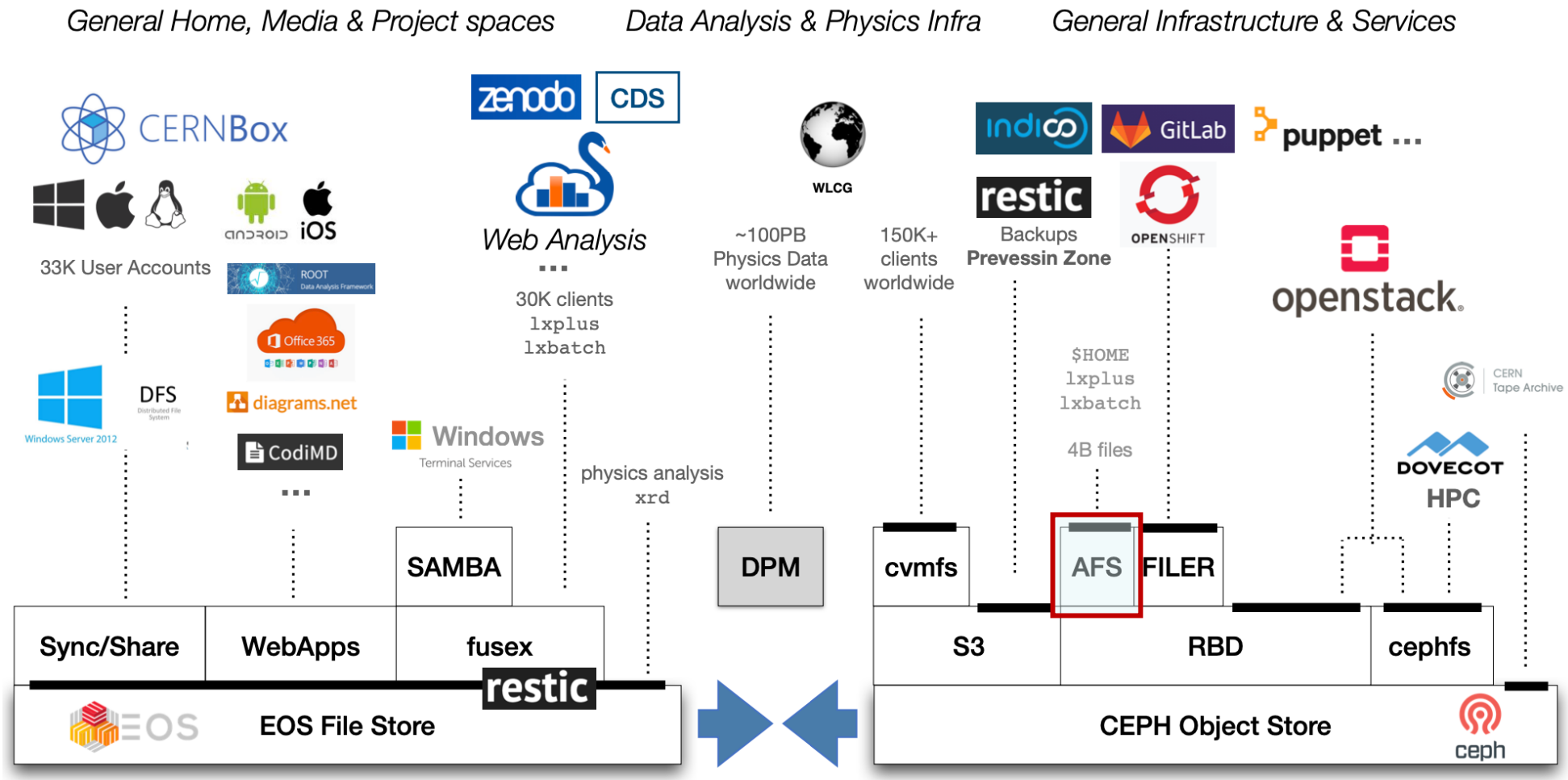
WWW invention



More than a billion particle collisions per second

Peak 60 GB/s

CERN IT Storage



AFS Deployment

Timeline

2014

- CERN uses AFS since ca 1993
  (see CERN AFS Site Report 2014)

2016

- CERN decides to phase out AFS

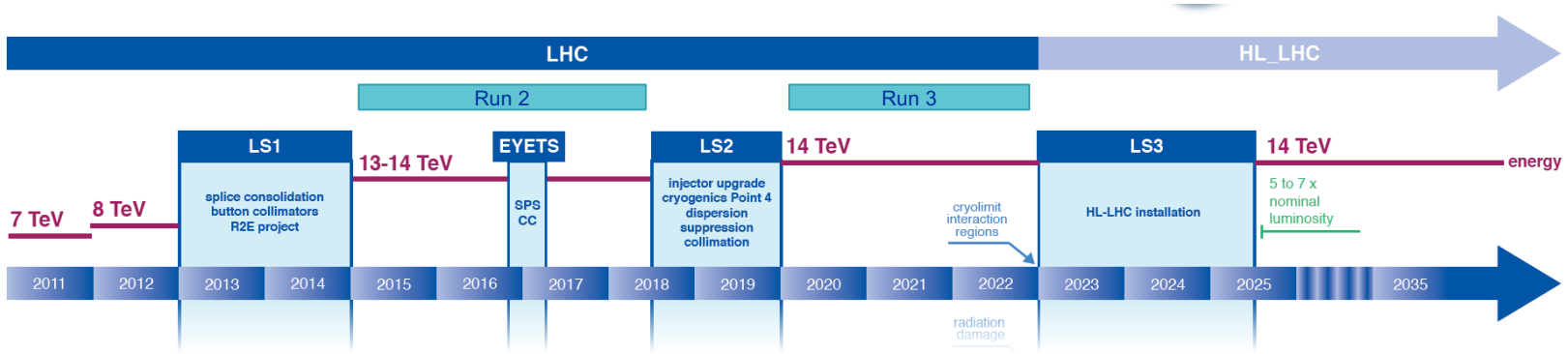
2020

- The AFS phaseout decision is reassessed

2022

- AFS is still supported but enters maintenance mode
  - Dependency should still be removed

LHC plan



Phaseout why? How?

- **Potential risks identified**
  - 2012 community split, dev effort decline, longstanding issues that affect US funding i.e IPv6/crypto
  - Risk of “sudden death”?
- **No changes should be made in the middle of an LHC Run period**
- **No drop-in replacement identified**
  - Migrate use case-by-use case
  - Hoped for shift towards web tools
  - Attempt to move most to our internal storage solution (EOS)

More information: CERN AFS phaseout: status & plans, J. Iven, A. Pace, CHEP 2021

Use cases

- **Not migrated**
  - $Home (interactive logon service) → user volumes (/afs/cern.ch/user)
  - Filesystem interface for local BATCH access → preferably workspace volumes (/afs/cern.ch/work)
  - Small Experiments data, projects → /afs/cern.ch/project, …
  - Software compilation, websites... → all (mainly workspaces)
- **Migrated**
  - Large Experiments data
  - Most of Project spaces
  - Archival of old data/projects (first website in the world!)
  - Software distribution → now on a dedicated, RO system (CVMFS)

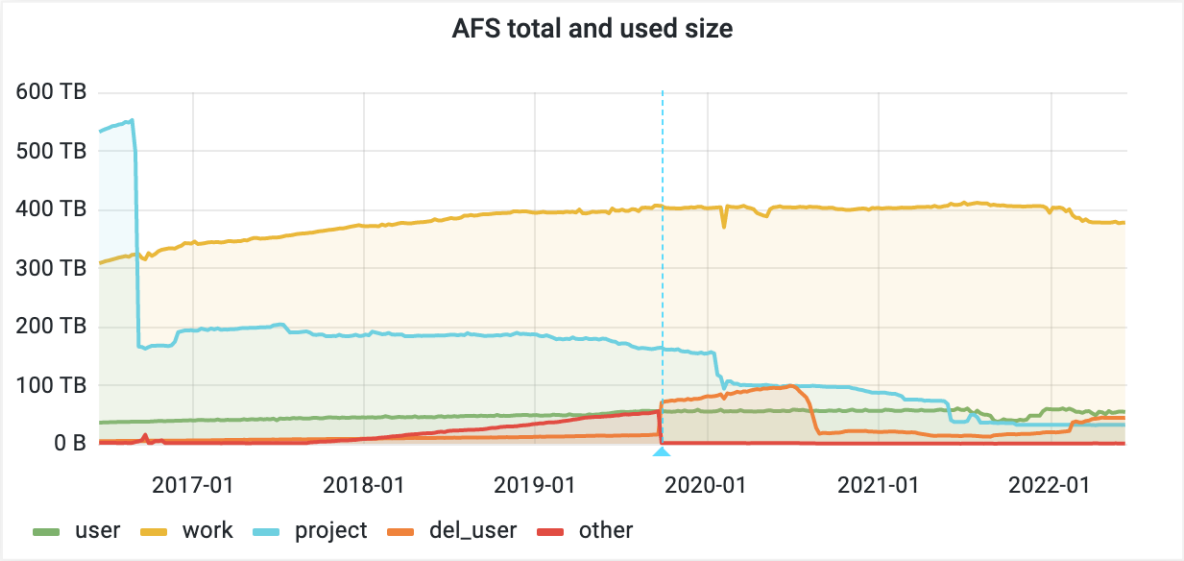
Did we phaseout?

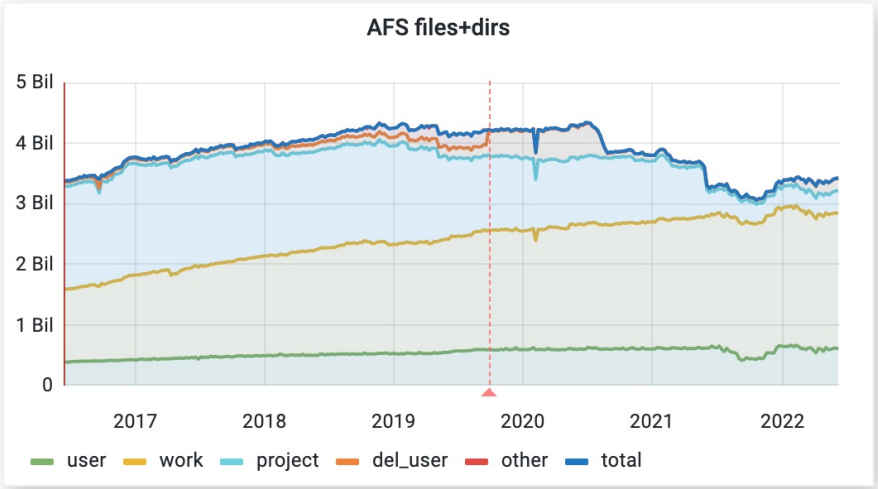
No

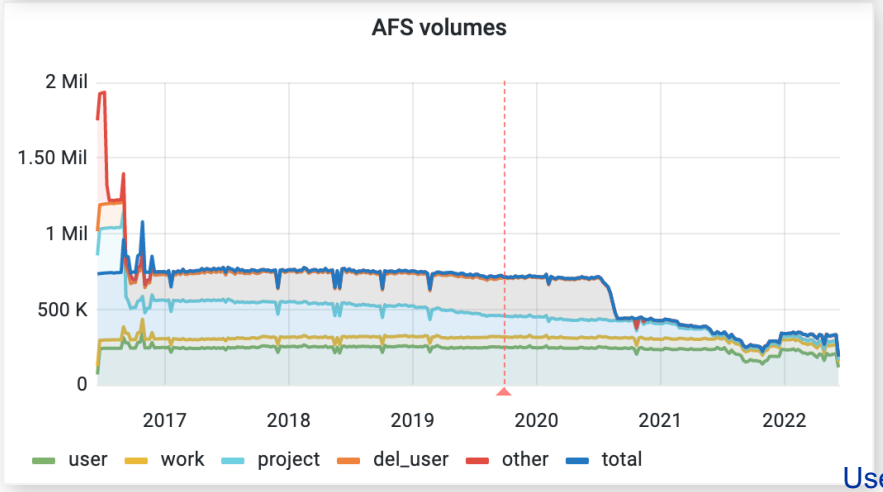
What happened?

- **The risks did not materialize**
  - OpenAFS did not die: community reduced but alive
  - Kernel support for latest OSes was provided, new releases, more development momentum
  - Service is stable
- **Not enough time to finish it all**
  - Reached the start of Run 3 without all use cases migrated (and no longer urgent to do it)
  - Some use cases still don’t have a clear replacement (maybe deprecate the use cases?)
- **We will continue to decrease the dependency on it**
  - The service entered “maintenance mode”
  - Still supported in Run 3 (2025)
  - Usage has stabilized (no new usecases), archival & cleanup ongoing

/afs/cern.ch in numbers







~ 510 TB (+96% than 2014)
~ 3.5 billion files (+66% than 2014)

- Increased after 2014, but has already stabilized
- AFS is global, but its usage at CERN is mostly local
  (external disconnection test 2021)

User quota: 10GB
Workspace quota: 100GB
(vs CERNBox quota: 1TB)

/afs/cern.ch in maintenance mode

- **Running on OpenStack VMs**
  - 59 Filesystems (shared, 4cores/7.5G/2partitions or 8cores/15G/4partitions)
  - 3 DBs*
  - CEPH block storage
  - Some FS+DB under critical power ($Home + critical projects)
- **Software**
  - OpenAFS 1.8.8 (clients are a mix of 1.8.6 and latest)
  - OS: CERN Centos (CC) 7
- **Backup to Tape**

(* + 2 Physical DBs to be decommissioned → moved from CellServDB to DNS to allow further changes)

Challenges

Batch jobs scalability

- **Service reaching its limits**
  - Local Batch capacity growing faster than what AFS can cope → DoS ”attack”
  - Users’ "bad" practices (enormous-files [Cache invalidation], "conda" instead of CVMFS, etc)
  - Often see huge (multi-second) latency on FS
- **Mitigations in place**
  - Volumes segregated by type (& sometimes experiment)
  - Increased number of (smaller) machines for higher user distribution
  - Automatic shuffling to reduce colocation of active users
  - Slowdown and (manual) blockage of heavy load users
  - Soon: automatic throttle of batch jobs

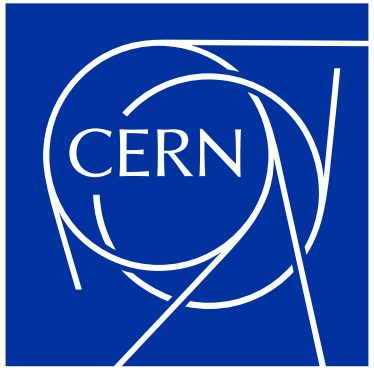
Batch jobs scalability

- **In the past we had a throttle mechanism**
  - “R.Többicke-patch”
  - Changes to Rx and clients to throttle jobs on the client
  - Lost with the migration from 1.6.7
    - Deliberate decision to go upstream

Conclusion

Conclusion

- **The current usage is stable**
- **AFS will be supported (at least) until the end of 2025**
  - Efforts to reduce dependencies will continue
- **Ways to improve the batch integration are under discussion**
  - New ways to interface with it (without the need of AFS) are being considered



home.cern