

# CERN's AFS Replacement Project

## Background



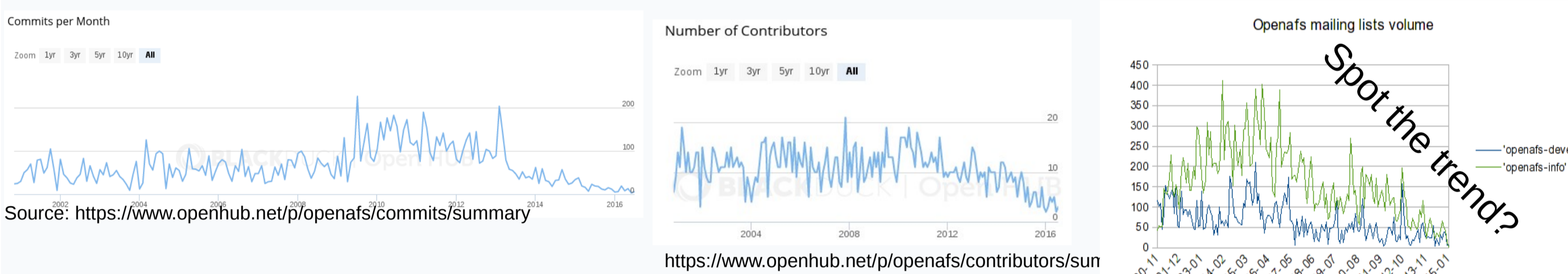
- AFS: distributed global shared filesystem
- CarnegieMellon (1983) → TransArc (1994) → IBM (1999) → OpenAFS (2000)

## AFS @ CERN

- Since 1990
- 35k users (5k active/day), 450TB data, **3.5B** files/dirs, **3.5B** accesses/day
- Small footprint: 50 (old=small) file servers, 5 DB servers, 1.2FTE / 3 people

## Why replace AFS?

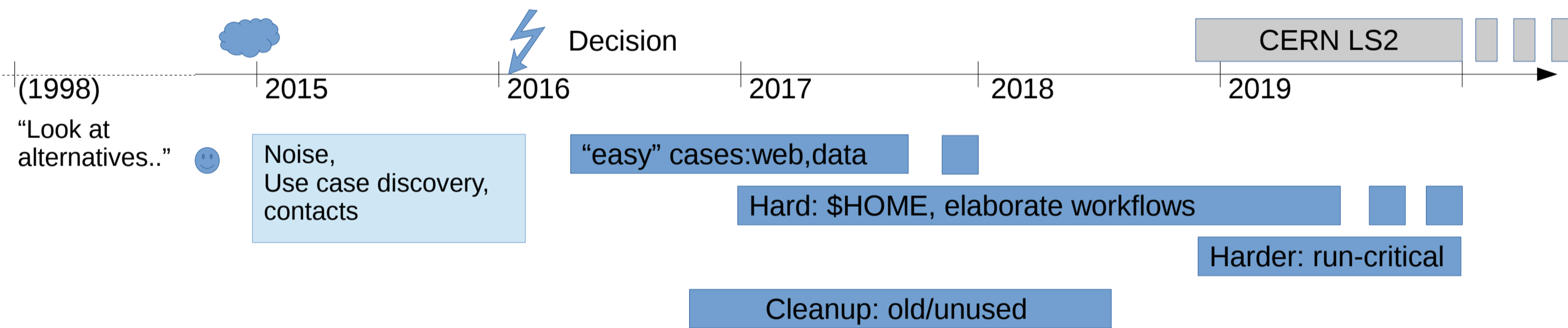
- Slow decline, lurking issues (single-DES, IPv4-only, RX protocol..)
- Architecture is prone to overloads (parallel access to single R/W volume)
- Difficulties with modern clients (MacOS X, Win10, Linux4)
- Turnaround + new lease of life is unlikely
- **“orderly phase-out”**. No need to rush... but cannot ignore.  
LHC Run3 should no longer depend on AFS



## Phaseout process

- nominated “phaseout coordinators”
  - Bottom-up (community) + top-down (hierarchy)
- **Discover** use cases, **classify** (common vs specific)
- **Migration is user-driven**
  - Assess, migrate to appropriate target, verify, delete from AFS
- **Communication**: Talks, meetings, mass-mailings, ..

Multi-year!



## AFS taxonomy @ CERN

- Personal \$HOME (2..10GB)
- Personal workspace (10..100GB)
- Shared project space (1GB..10TB)
- group shell environments

## Used for

- Per-user files content (analysis, scripts, documents, ..)
- Software development (engineering, experiments, ..)
- Batch input/programs/output area
- Software distribution (ready-to-run, tarballs)
- Websites (3.6k personal, 600 project)
- But also:
  - Workflow state machines, cheap “backup” service, remote home directory.. and lots of abandoned data.

## Challenges

- CERN (local) computing is **built around AFS**
  - Desktop – LXPLUS – LXBATCH affected
- Many unknown use cases / workflows, **abandoned** areas
- External sites depend on it (Remember, “global” FS..)
- AFS (despite its age) is **pretty good!**
  - Scalable namespace
  - Fast: Client-cache + loose consistency
  - “enough” POSIXy to keep many clients happy
  - “Secure enough” for untrusted clients

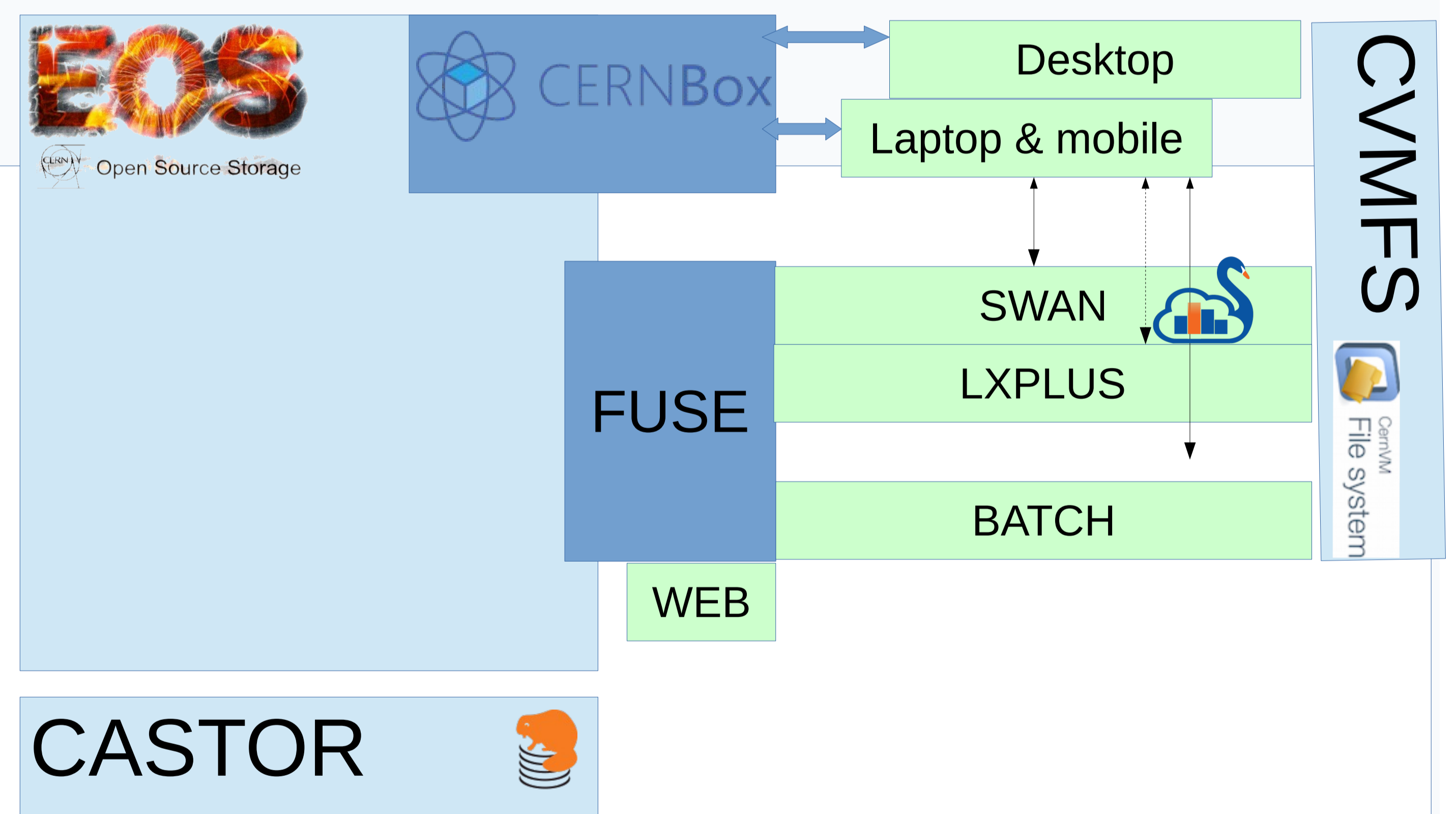
External users: Please get in touch..

→ **No single drop-in replacement yet**

## Alternatives & migration targets

- Primary target: **EOS (-FUSE)**
  - Main CERN disk system: heavily used for LHC data, 100PB
  - Already FUSE-mounted by users – now deploying by default
  - Weak areas: slow file creation, limited namespace scaling  
Are being addressed.
- Desktops + laptops: **CERNBOX** sync client (into EOS)
- Archive for historic read-never-but-don't-delete stuff: **CASTOR**
- **CVMFS** for software distribution
- NFS, CEPH(FS), .. for special cases

See EOS developments talk!



## Risks

- Long process + soft deadlines = never-ending project = AFS stays.
  - Isolate – mitigate network threats. Will eventually die out with clients
- Keep old use cases on new (but inferior) system? Some use cases might not even “fit”
- Talk about AFS decline: self-fulfilling prophecy

VS

## Opportunity: review (non-Grid) computing @ CERN

- Compute:
  - Sandboxing means much better scaling / **efficiency**
  - break-out to (opportunic cloud) capacity
- **Disconnected** operation
- **User experience** – workbooks instead of “ssh+bsub”

What should computing look like for the next 15 years?

