

AFS Replacement project

- Motivation, process, alternatives
- Impact & opportunities

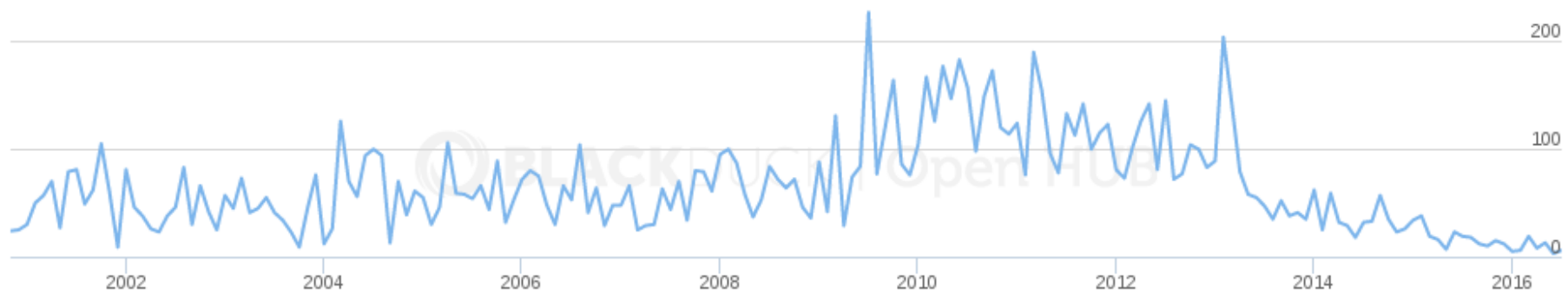
Why phaseout?

- OpenAFS project in **(slow) decline**
 - Various “soft” indicators: releases, traffic, people, conferences,..
 - Pent-up changes: IPv6, DES (backward compat.. ®)
 - Funding → ecosystem
 - Ongoing client upkeep (incl signed binaries on Win+Mac)
- Technical:
 - Widening gap
 - SPOF architecture vs ever-bigger machines
 - RX protocol vs high-latency
 - Odd limitations (32k files in directory)
- But
 - Project is still “functional” - new releases, slow changes

≡ need to move out, but in control

Commits per Month

Zoom 1yr 3yr 5yr 10yr **All**



Number of Contributors

Zoom 1yr 3yr 5yr 10yr **All**



How to phase out?

- Classify AFS subtrees (& AFS volumes)
 - Migrate
 - CERNBOX – human-generated content
 - CVMFS – software distribution
 - EOS – live data
 - CASTOR – dead data
 - Needs ACL review – ideally “public”
 - Delete ? – machine-generated junk & obsolete
 - Difficult decision, users need guidance.
 - [Escalate anything else ..]
- No full overlap use case / proposed solutions
 - But some use cases should change
 - Temp files : use local disk or memory
 - Interactive analysis?
 - Browsers, Mail: stay local

1998: "look at alternatives.."

2014: "rather bad state"

✓ 2015: NOISE & FUD & Discovery

- initial communications
- establish experiment contacts
- Use case discovery

➔ 2016: EASY

- Web → EOS
- Projects → EOS
- Software → CVMFS
- + Obvious cleanups
- + Test & Improve alternatives
- KPI: Cap AFS growth (rate)



2017: HARDER

- Dead experiments – DPHEP?
- More elaborate use cases
- Home directories

2018+19: HARD

- As in 'die-hard'..
- No more "LHC-stop-threat"

- Communication
 - Gently spread bad news to key people ✓
(No hard deadlines yet)
 - Formal communication: ITUM, IT-Exp., .. ✓
- Experiment/department phaseout contacts ✓
 - Awareness: Spread news in their exp/group ✓
 - Inventorize: Discover & report particular use cases ✓
 - (prevent new AFS-based solutions)
 - Internal tracking
- EOS & CERNBOX – getting ready + picking up use cases

- Communications, support & documentation ..
- EOS-FUSE is *key enabler*
 - Unlock PLUS / BATCH use cases
 - EOS project space (→ lock AFS project space)
 - EOSWeb backend (migration for AFSWEB)
- Remove AFS deadwood
 - Unused accounts (30%?), obsolete volumes
 - Old binaries: AIX, SGI, SUN4m, HP-UX → DPHEP decision
 - Archive 'obvious' subtrees
 - Flush out “lost tribes” and (unknown) external usage
- Improvements on EOS & CERNBOX
 - EOS namespace: boot time (2Q2016), scalability (4Q2016)
 - CERNBOX: SSO, integration, better sharing & scaling..

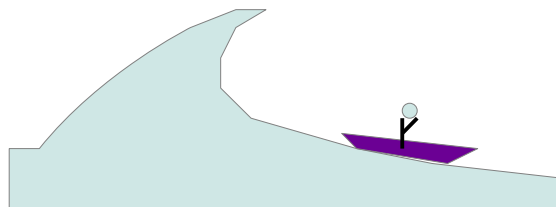
- EOS+FUSE: strategic direction = try to make it work
 - OK for data (big fast files)
 - bad for interactive/compile, \$HOME, small+short-lived
 - Missing: Namespace, file creation rate, missing POSIX (locks) etc
- Cluster filesystems? OK for unauth access from “few” machines
 - NFS (no auth): several services
 - CEPHFS: used on HPC cluster
 - Others (Lustre, GPFS, Gluster): not (recently) considered
- DFS
 - Secure, but small service; trying to replace: Samba+EOS



- Multi-role LXPLUS:
 - External SSH access gateway
 - LSF submission machine
 - “default” SLC6/CC7 validated environment
 - Analysis compile, debug, run
 - 'acrontab' recipient, mail reading, browsing..
 - disentangle from “AFS”
- BATCH: LSF → CONDOR migration
 - Opportunity for better efficiency
- Account: split “UNIX” account from “AFS” account
 - Home directory is optional.
- WEBAFS → WEBEOS: pilot stage
- AFS-the-free-backup: make explicit. We have tapes.

} Future
Computing
@CERN

- AFS Phaseout has to happen
 - Migration impact all over CERN
- Communication & perception
 - Preparation, guidance, common solutions
 - tools to ease migration
- CERN can control the migration (speed & direction)
 - Have started early enough ..
 - Goal: final cleanup in LS2
 - Strategic: physics computing in 2020 ?
 - Steer evolution via attractive services (e.g CERNBOX + EOS/FUSE)





Support material

Recap – AFS

- Global (*secure*) Multi-OS Filesystem

- Client caching

- Size

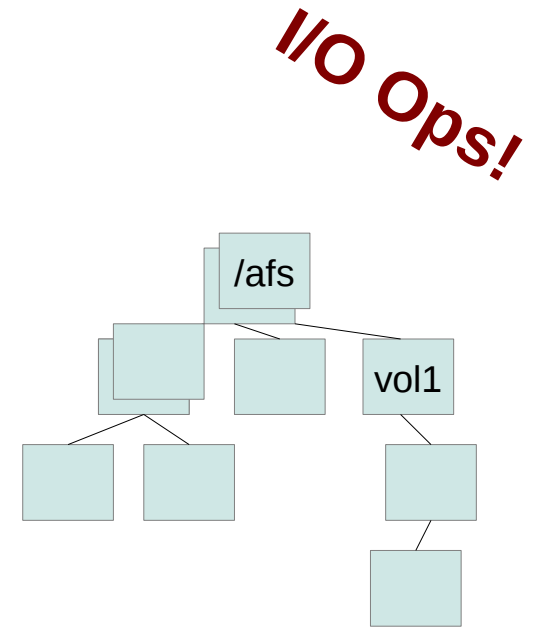
- 460TB used on ~50 servers

- **3.2G files** (Σ CASTOR+EOS: ~1G)

- 5G stats daily, 600M r/w ops daily

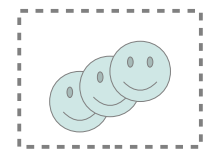
- Basic building block: “AFS volume”

- 100k of them, Size: 5k .. 300GB



- Own user database, **Own groups (!= E-Groups)**

- 33k users, ~6k active last week, 3k groups (ACLs)



- By usage:
 - “Work” user data (.doc, .ppt, code, analysis executables etc)
 - Home directories (config), group settings
 - Project space (=shared)
 - Experiment data
 - Software compiling & distribution
 - Web
- } building block for other services
- By creation:
 - Human-crafted (documents, code): **high-value**
 - Physics data: **high-value**
 - Physics derived data (MC, analysis): **medium-value**
 - Machine-generated for re-use (binaries): **medium-value**
 - temporary, write-once read-never (or -once): **junk**

Also in GIT, EDMS

EOS+T1 replicas

CVMFS

YUM

- Keep AFS
 - Self-support
 - Possible, need increasing manpower
 - Commercial support
 - Possible now, but unclear business perspective
 - Commercial alternative (AuriStor)
 - unclear business perspective
 - Non-HEP solution
 - Segregate + decay (firewall)
 - Risky but part of a solution
- = Does not address architectural issues. Delay until when?
- Use “somethingelse” - No single candidate. what to give up?
 - Secure authentication = untrusted clients, wide-area
 - Multi-OS support
 - Performance