Preparation of pilot cases for AFS phaseout

G. Iadarola

Many thanks to:
L. Deniau, G. Rumolo, R. De Maria, J. Iven, Jakub Moscicki

- As you probably already heard, **the IT department is planning to decommission the AFS filesystem**
  - For many of us AFS is a **vital part of our workflow** (used in lxplus, lxbatch and mounted on several local machines)
- The **solution proposed by IT is to use EOS** (mounted via FuseX)
  - This is the same space used to provide the CernBox service
  - User EOS space (1 TB) available for all of you at (in /eos/user/)

The timeline foreseen by IT **aims at completing the phaseout by the end of LS2**
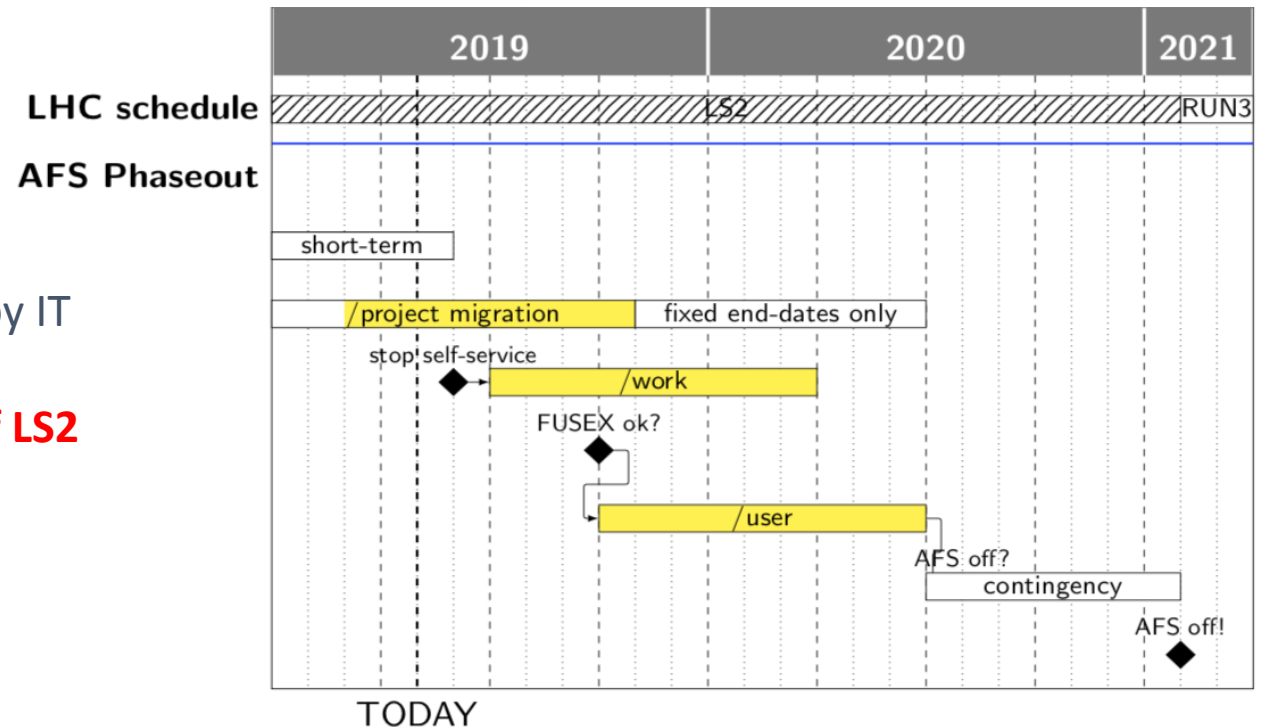
- Many of us have already **tried to test their workflow on EOS**

- **Several issues** were encountered:

  o Random I/O errors, files randomly disappearing, non-reproducible errors when making basic installations, issues with compilation and testing, slow reaction time to commands, not working with HTCondor, high failure rate on receiving results, limitation on number of files per workspace

- The general feeling in ABP is that **EOS is not yet mature** to function reliably as a replacement for AFS → **impact on ABP work could be very significant**

- From the discussions with IT it became clear that **they need input from our side** to identify main issues and solve them

  → we cannot just "do nothing" and wait to EOS to become better

A strategy to address the problem while preserving the continuity of our work has been defined at the ABP group level (CWG, SLs, GL) and agreed with IT:

1. **Starting now:** ABP will create a series of **pilot use cases** (in the form of shell scripts) representative of ABP workflows relying on AFS

2. **Mid-May 2019:** pilot cases will be **passed to IT**

3. **May-October 2019: IT will use these pilot cases to test the functionality and reliability of EOS** and **make the necessary improvements**

   - Use cases should work without modifications on EOS fuse-mounted paths, with reliability and performance comparable with AFS

   - If modifications are required these should be agreed with ABP (after assessing the impact on our work)

4. **October 2019: performance** of pilot cases on EOS will be **evaluated by ABP and IT** → **if successful ABP workload migration starts**

   - If unsatisfactory a different solution and/or timescale should be proposed by IT

5. **Oct 2019 - June 2020:**(assuming positive decision): **migration of ABP workload** from AFS to EOS → goal is to be out of AFS in June 2020

Contact persons defined in each section (G. Iadarola for HSC) to follow up the process

We need to **identify the activities in the section that are presently relying on AFS**:

- For each of them we should prepare a pilot case to be passed to IT

We need to move quite fast:

- If you feel you are concerned, **please prepare your pilot cases and send to me by Tue 6 May**

- Feel free to contact me in case you encounter issues or need advice

We should try to avoid duplications → coordinate within working groups

A **pilot case** should be a **shell script** which mimics your workflow starting form an arbitrary empty folder (AFS, then EOS):

- The script should run on **lxplus**

- It **can access the internet** to retrieve software, data, etc.

- **Installations**, **compilations** etc. should be included (this is where EOS has troubles)

- You can **launch jobs on HTCondor** from the script
  - You can **include a check script** that verifies that the the jobs were successful and that output files are readable

You should **make sure that your pilot case works reliably on AFS**, no need for you to perform tests on EOS

An working example can be found here:

https://github.com/giadarol/afs_phaseout_ecloud_test/
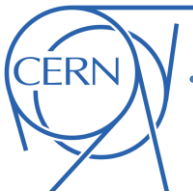
Please send me your pilot cases by Tue 6 May

Please contact me in case you need advice