

# CHEP 2018

Fabrizio's mini report

# Intro

- I am the third one today, maybe you have enough of intros
- This is somewhere between a summary and a ‘look at this’, sometimes with my personal opinion and interpretation
- Well organized CHEP conf, good content, very dense days.
- I also enjoyed Sofia
- I followed the “Data Track”, don’t know about other tracks
  - <https://indico.cern.ch/event/587955/sessions/266673/#all>
- Many people followed the DPM talks
- Several talks from people using Dynafed in very interesting ways
- Unexpected leap forward for the DOMA activities

# UK and Tier-2 caches

- Sam Skipsey and his controversial talk on caching studies in UK sites. Very detailed, a must read
- Main message: it's very difficult to have an advantage from caching:
  - “Files overwhelmingly accessed once”
- FF: In the talk (despite using WAN access) there is no mention of data access, TTreecache and latency. It seems they have not been understood in ATLAS. CMS is more efficient due to this
- IMO they should have characterised jobs and studied the various behaviours separately. Certain kind of jobs are obviously run once, and this is particularly the case for centrally managed campaigns.

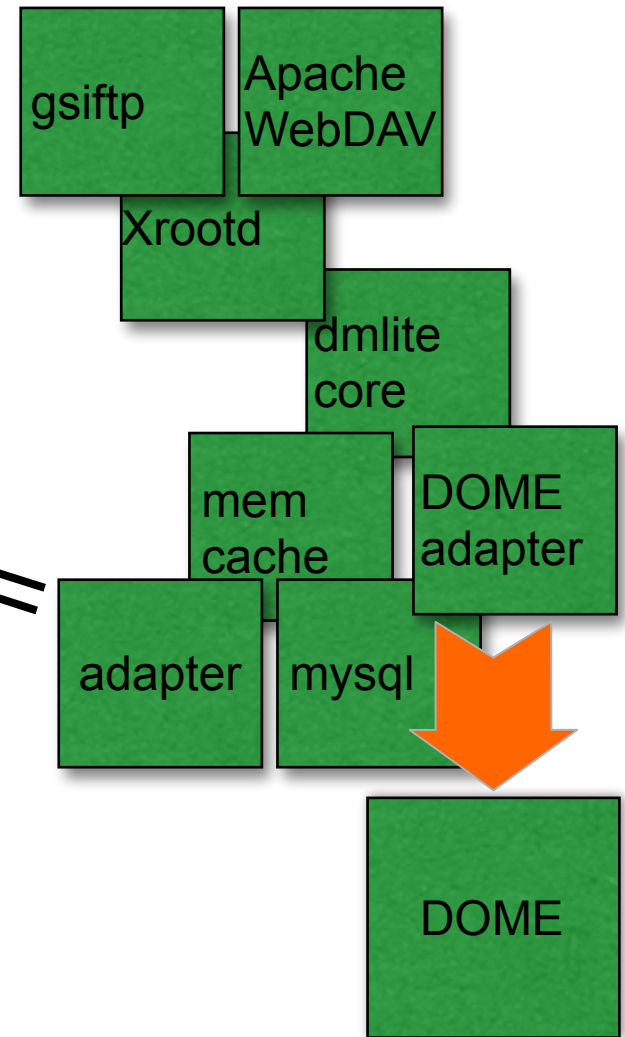
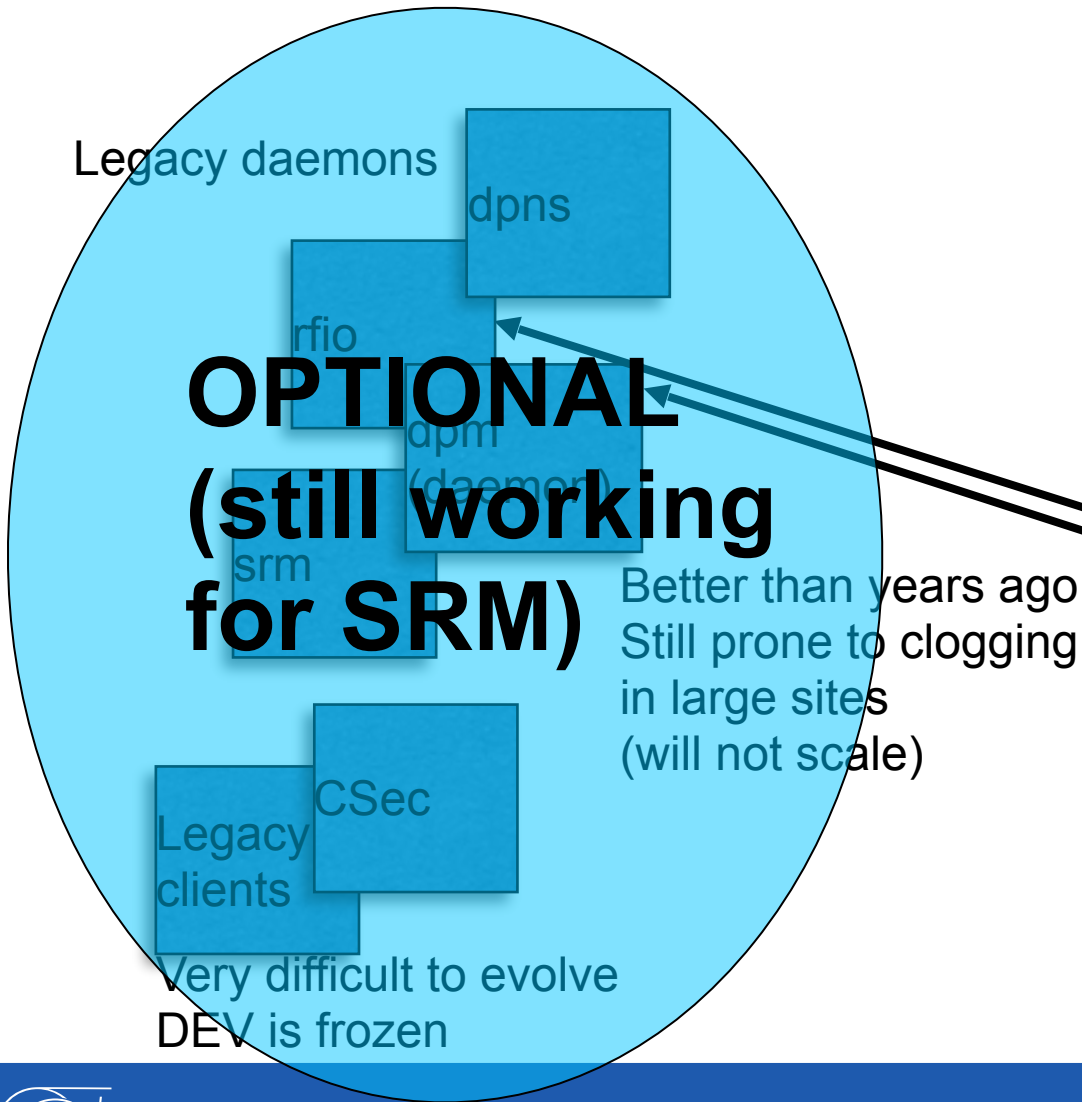
# WLCG space accounting in the SRM-less world

- By Julia Andreeva. I suggest people to read it
- Mildly technical, it gives an idea of how the Grid is evolving about storage, monitoring, SRM, etc..
- Julia asked for the collaboration of the DPM team because DPM has everything ready and involves a majority of sites
- The large campaign for the DPM upgrades will be triggered (also) by this space accounting transition

# DPM talk

- Pretty good shape for DPM (Disk Pool manager)
- We've been able to make it independent from the legacy code
  - ... and set a date for the end of support for the old stuff (rfio, csec, shift, srm, ...)
- Metadata performance/scalability are OK too, no problem for sites that plan to grow (if they update and reconfigure...)
- People are doing interesting experiments with caches and volatile pools and federations

# DPM components and plugins (2018)



# One plugin to rule them all

- **In non-legacy mode** DPM now loads only DMLite::DOMEAdapter
- dmlite-adapter, dmlite-memcache, dmlite-mysql are no longer necessary
- Resource consumption (FDs, mysql, etc.) is reduced by an order of magnitude, and so complexity and cost for us all



# LCGDM support from 01/Jun/2019

- From **1st of June, 2019** our standard LCGDM support answer will be **“there is an alternative: upgrade DPM to DOME flavour, please”**
  - A large part of DPM support requests is about dpm+rfio clogging, these problems will instantly be reduced and simply disappear as the SRM usage decreases
- LCGDM will stay in EPEL as long as it compiles untouched in Rawhide (EPEL rules will remove it if it breaks)
  - It's pure C code, hence that can be years, we don't give limits
- LHC Tier-2s in general can work without SRM, at least there are recipes to do it
  - e.g. for ATLAS: <https://indico.cern.ch/event/699602/contributions/2944281/>



# BOF on macaroons

- Macaroons are an implementation of ‘bearer tokens’ (HTTP jargon)
  - “The bearer of this secure token is authorised to do XYZ”
- Supported by dCache, DPM, XrdHTTP
- They work, still pretty experimental
- Likely to have a role in the post-X509 scenarios, together with SciTokens



# DOMA BOF

- A leap forward for the DOMA activities, beyond my expectations
- Sometimes the discussion has been pretty “intense” (Tigran, Andy, Wei, Brian on protocols). That IMO has been the moment of highest value
- For TPC with HTTP, it seems that XrdHTTP has become a very important component, also thanks to the contributions made by Brian Bockelman
- I said am happy to review all the present and future pull reqs to XrdHTTP
  - And also very interested, as the DPM core is based on it. Also EOS has foreseen activities with it
  - So far this cooperation has been extremely fruitful

# Dynafed

- Quite some third party talks about Dynafed on Thu afternoon:
  - Building a global file system for data access using Large Scale CVMFS and DynaFed (RAL, A.Dewhurst)
  - An http data-federation eco-system with caching functionality using DPM and Dynafed (INFN, S.Pardi)
  - Advancements in data management services for distributed e-infrastructures: the eXtreme-DataCloud project (XDC, D.Cesini)
  - Using a dynamic data federation for running Belle-II simulation applications in a distributed cloud environment (UVic, M.Ebert)
- The only one with names from our team (Oliver) was the XDC overview. I did not like it, I expect much more from XDC. It hope it improves in the next iterations, as the project is ongoing
- Dynafed also had a poster (Furano-Keeble)

# DynaFed poster

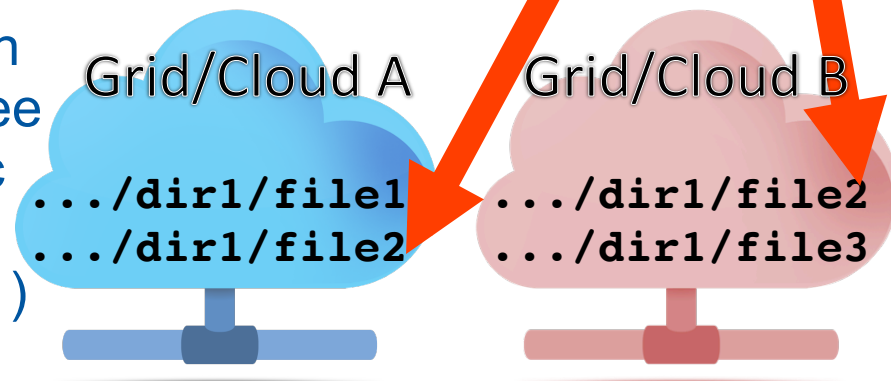
- Dynafed allows the creation of flexible and seamless storage federations out of sites that expose WebDAV, HTTP, S3 or Azure interfaces. Works great with CEPH and volatile storages (caches)
- Read/write support, FTS-friendly, matches Cloud and Grid on the fly
- Flexible, open, high performance authorization subsystem
- Part of the XDC project (see talk 537), in production at CERN for LHC@home (see talk 88), preproduction for ATLAS@Uvic (see poster 161, talk 105), Belle-II (see talk 479), deployed at RAL (see talk 421)

On the fly friendly  
visualization  
Full **WebDAV**  
access  
Redirection-based  
Fully scalable



With 2  
replicas

`/dir1`  
`/dir1/file1`  
`/dir1/file2`  
`/dir1/file3`



# Data track conclusions

- I would have liked to see a more energetic plenary track conclusion talk, more summarizing than random citing
- 4 Dynafed talks + poster summarized into one strange line
  - “DIRAC, Rucio, CVMFS, Dynafed and many others are adopted by an expanding community outside CERN: SoLid, Belle II, LIGO”
- Also ‘the line’© about DPM is difficult to understand:
  - “DPM streamlines DB operations by deprecating legacy & SRM-related services” ... very obscure to me