

DPM workshop 2019 - GDB report

Fabrizio Furano
Oliver Keeble

DPM 2019 - Bern - 13-14th of June

- Excuse was Bern being the first “multi-site DPM”, in production since > one year
- Perfect local organisation, attendance in line with the previous editions
- 15 registrations, 8-9 Vidyo participants
- Among the guests:
 - EGI rep (B.Grenier)
 - WLCG rep (S.Campana)
 - CERN mgmt rep (D.Duellmann)
- Big discussions about support, obsolescence of the “old codebase”, future plans
- <https://indico.cern.ch/event/776832>

Direction: stability and LTS

- No tech revolutions since quite some time, increased focus on quality
 - DPM stays one of the main storage techs used in Grid sites
 - BDII reports 102 instances, 96PB, largest is >10PB
- The architecture is definitive and stable, a TF is managing the sites' transition (supporting the WLCG SRR)
<https://twiki.cern.ch/twiki/bin/viewauth/LCG/DPMupgrade>
- All the past and present efforts are aimed at giving long term support and longevity
 - Longevity: can adapt to the future required features (e.g. scitokens, OIDC, caches, macaroons, ...)
 - Long term support through stability. We expect relatively little fixes and improvements
 - Most of them related to usability and UI

Direction: stability and LTS

- This is actually the goal sought several years ago with the idea of DMLite and of the “DPM collaboration”
 - A healthy open source project
 - Understandable by others willing to cooperate and contribute
- IMO this has been among the best accomplishments, many thanks to all who contributed!
- The explorative project took more than expected, the results have been good in our view
- And this is the reason why we did...

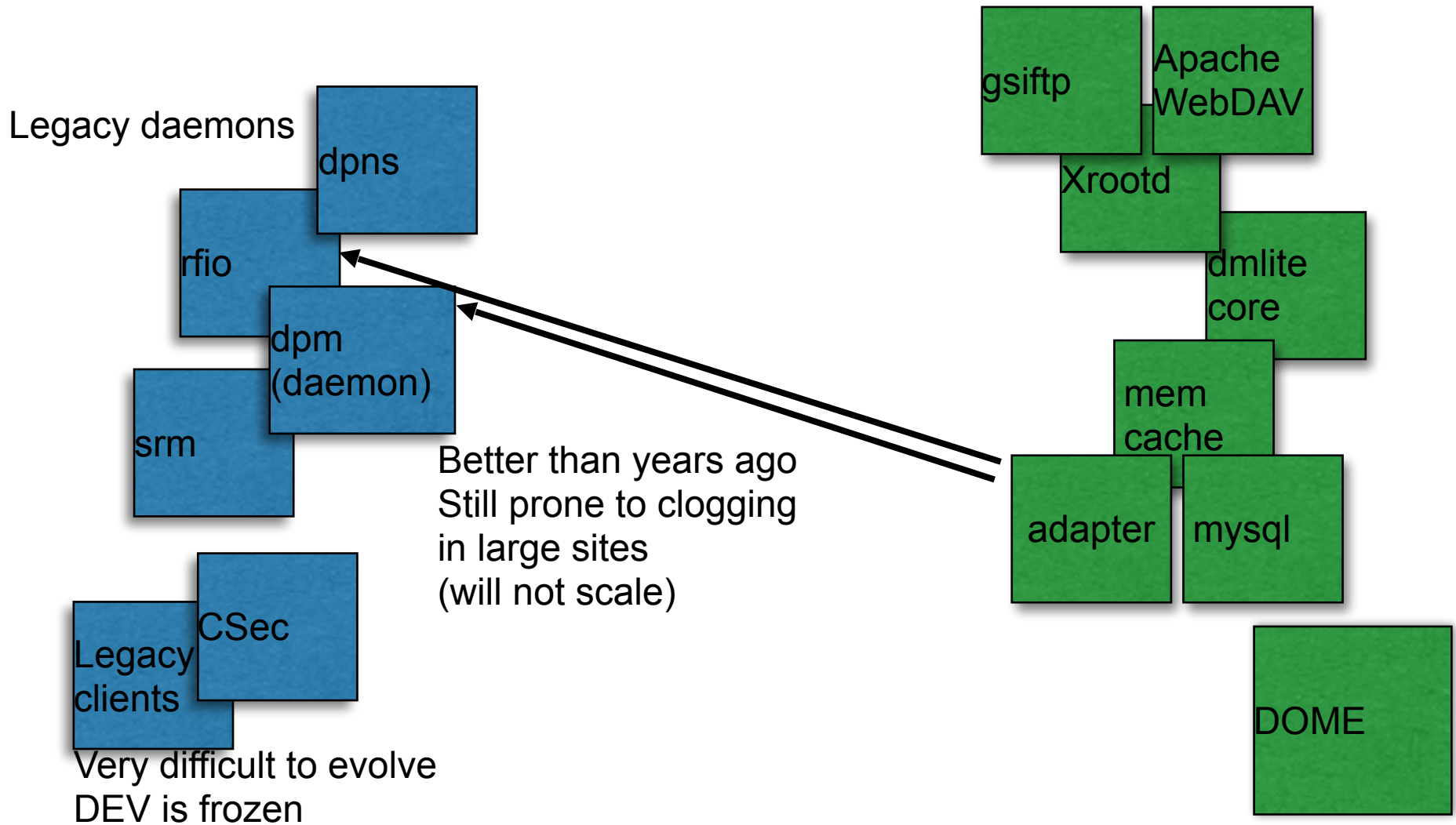
LCGDM support from 01/Jun/2019

- From **1st of June, 2019** our standard LCGDM support answer will be **“there is an alternative: upgrade to DOME flavour, please”**
 - That affects: dpns, dpmdaemon, rfio, CSec, dmlite::Adapter, SRM
- LCGDM will stay in EPEL as long as it compiles untouched in Rawhide (EPEL rules will remove it the day it breaks)
- It's pure C, hence that can be even years, we don't give limits

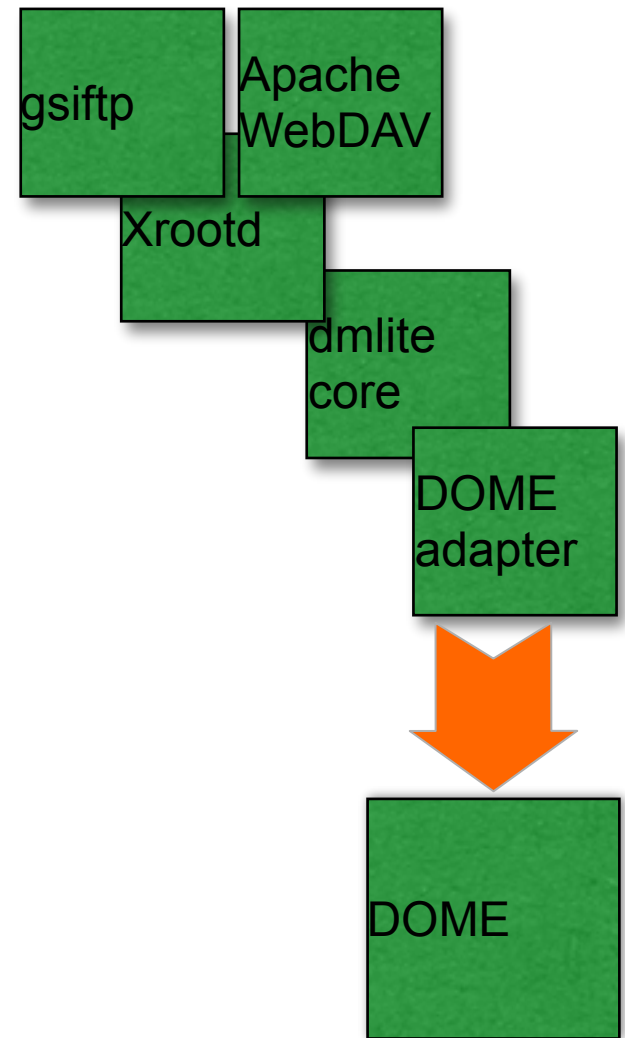
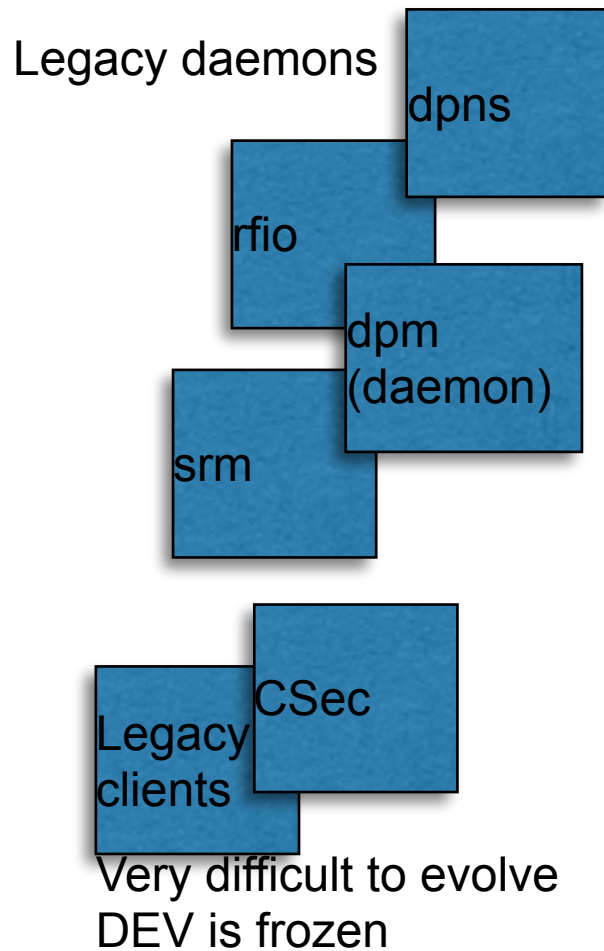
Longer legacy security support

- [4th of June] CERN and EGI agreed on postponing the deadline for security support for the legacy codebase to end of September
- As DPM team we agree it's important, and we don't expect troubles from the core legacy components in CC7 or EL6
- Also the external older components (e.g. globus, gSOAP) are pretty stable

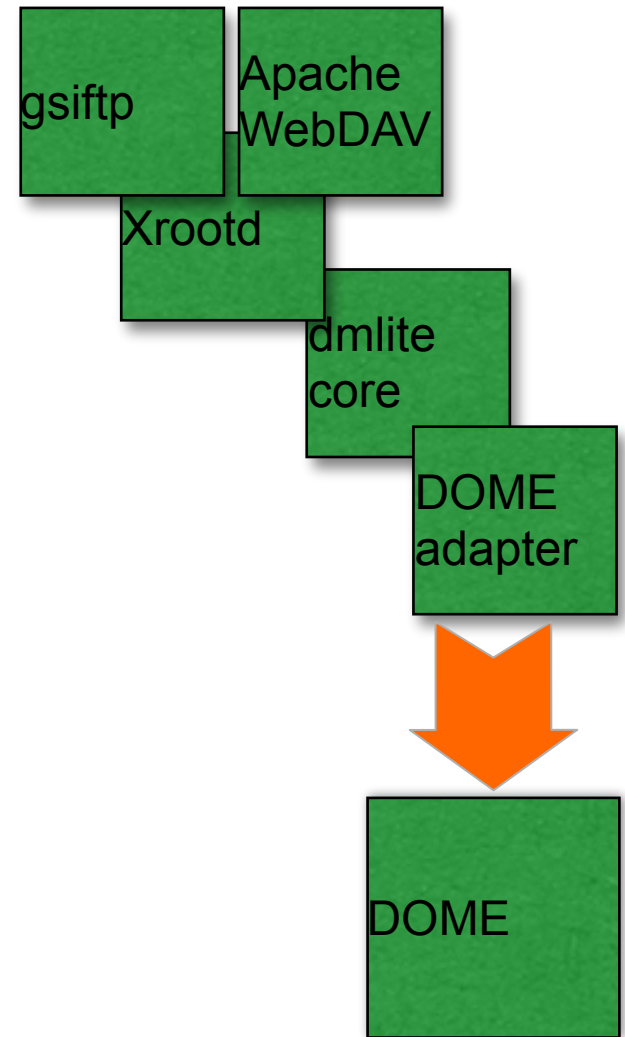
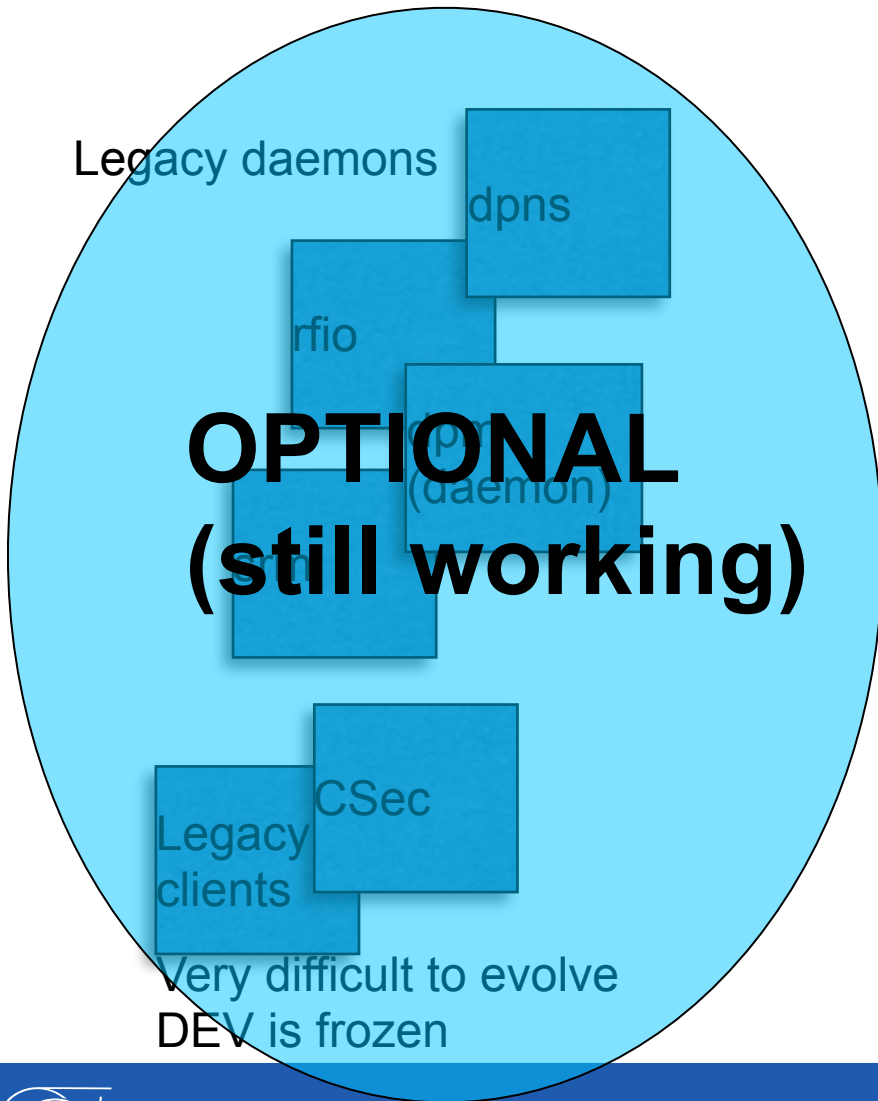
DPM components and plugins (2017/2018)



DPM components and plugins (2017/2018)



DPM components and plugins (2017/2018)



What about SRM ?

- The legacy lcg-dm stack also includes SRM
- A serious attempt to pull the community (at least for disk-only sites)
 - out of SRM
 - towards more modern protocols
- DPM is positioned very well to interact with the stakeholders (sites, EGI, VOs, WLCG)
- During the workshop it has become clear that the message is being understood

Summary: 3 DPM setup flavours

- **Full legacy**: dpmdaemon+dpnsdaemon+rfio+srm
 - DMLite loads the Adapter plugin to give http+gridftp+xrootd
 - The DOME daemon runs in the head node and does basically nothing (dormant)
 - gridftp can be used only through SRM
 - **This is the status of many sites...This legacy setup cannot scale up and is based on legacy components**
- **DOME plus legacy**: legacy dpmdaemon+dpnsdaemon+rfio+srm plus DOME supporting http+gridftp+xrootd
 - DMLite uses DOMEAdapter, legacy part stays legacy and does only SRM+rfio
 - gridftp can be used through SRM and directly with a gridftp client towards the headnode (gridftp redirection)
 - **This setup can scale up the number of servers, the max transaction rate (http/xrootd) is more than an order of magnitude higher than the legacy flavour (which is still being used as an optional part)**
- **DOME without legacy**: take the previous option, stop and/or uninstall srm,dpmdaemon,dpnsdaemon and rfio.

Cache mode - Volatile pools

- There since Q1/2018, works interchangeably with all the protocols (modulo SRM/rfio)
- INFN-NA has an advanced testbed that was described at the workshop
 - <https://indico.cern.ch/event/776832/contributions/3378598/>
- Functionally it's quite solid and well integrated in the idea of the DPM pools
- It's a full-file buffer, AFAIK more than sufficient to give the 'cache experience' (and transferring the file at the first access is way quicker than tunnelling all its chunks with some latency)
- Supports pre-populating by construction, can also be written into normally
- If/when there is any content that is worth caching we will be able to understand if its cache purging algorithm is good enough
 - I would be in favour of improving it, IF it's useful to some clear use case and IF we can document that a different algorithm would make a meaningful difference
 - Until then, the current purging algorithm is fine and the feature can be used normally

Remote pools - disk-only sites

- Theoretically it has always been possible, yet quite tricky with libshift (among the oldest components from CERN IT!) and rfiio (not much younger)
 - Without forgetting firewall rules and reliability
- This workshop was triggered by Gianfranco Sciacca, who one day popped out and said “do you know I did that in production ? it works”
 - Hence, more details in Gianfranco’s talk
 - <https://indico.cern.ch/event/776832/contributions/3378586/>
- In pure DOME mode the setup of a disk-only site becomes simpler, and more robust, because the intercluster communication is more solid

Security - Macaroons

- DPM has pioneered macaroons together with dCache, a few years ago.
- They work fine, being used in the DOMA-TPC exercise, and they are quite easy to configure

A view on security

- A.Ceccanti contributed a very high quality talk on the next steps for WLCG security
 - <https://indico.cern.ch/event/776832/contributions/3378561/>
- In practice, standard security tokens, including WLCG fields
- We'll soon see the first beta services distributing preliminary WLCG tokens
- OpenID-Connect works fine in the DPM HTTP frontend, we'll need just to know the new rules for authorization
- Things will be clearer when the WLCG authorisation WG publishes its conclusions (or even better, when there's a working prototype we can use)
- FF: I don't see big problems with this business, maybe some small improvement on the way

Belle-II

- Belle-II is for the moment relying on older workflows (LFC+SRM)
- DPM is well represented among the Belle-II sites
- The advanced R&D presented at the DPM workshop shows results obtained with:
 - HTTP protocol
 - DPM volatile pools
 - Federations of regular and volatile pools
- <https://indico.cern.ch/event/776832/contributions/3378530/>

“The future of DPM”

- Unsurprisingly we start being posed this question more often
- The small DPM team secured the project from the technical point of view, and made it able to compete technically with the upcoming known technical challenges (e.g. higher load, scaling up, TPC, WebDAV, macarons, xrootd, multi-site, caches, etc.). Sites can work well.
- We know that there will be the necessity of adapting “deltas” in the future, e.g. enabling OpenID-Connect in Apache. Easy things and fixes are not a big problem
- **We (DPM team) don't know what to expect for the technical challenges that we don't know about yet.**
 - A random invented example? Interfacing the replica scheduling of a multi-site DPM with some georeferenced information from Google Earth
 - Our very low manpower may decrease at the end of 2019. Who will support the setup in 2020?
- **At the same time, the funding and the support of any scientific open-source project depends 90% on its users**
 - **Many users, well organised —> long and prosperous project life, good support**
- Bottom line: if you are concerned by this, you should talk to your WLCG contacts and together raise the questions to the appropriate place.

“The future of DPM”

- The discussion on this subject was pretty long and articulated
- It converged on preparing a document to be circulated and escalated, likely by the collaborating sites
- Ideas about what can be done to contribute to the future of the project in a wider community, that may include CERN and other NGIs as well