DPM and Dynafed 2020

Fabrizio Furano
Nov2020

Intro - What's in this presentation

- Describe the path that led to DPM and Dynafed
- Understanding what constitutes the effort of maintaining these systems
- Hint at the milestones that the sites will have to face

- Promote once more the "open source" character of the modern DPM/Dynafed components
  - … which is the best chance for longer life and support with a healthy coordinated community

Modern DPM and Dynafed: DMLite

- in 2010-2012 the CERN “Software for Distributed Computing” group started developing the DMLite framework and libs. The activities were part of the EU EMI project
- DMLite is an abstract library to build systems for managing storage

- DMLite was a shy attempt to incrementally rewrite DPM, which was facing code obsolescence at the horizon (C code inherited from past aeons)
- This focused in 2017-2019 with the DOME daemon
  - Made possible to implement “modern” features, fix historical troubles and obsolescence and improve dramatically robustness and scalability of DPM

- DMLite spurred in 2010-2012 the writing of Dynafed to make federations of HTTP(s) storage
- The HTTP client Davix (used also by ROOT) came from Dynafed, and found an important place in the present and future of computing at LHC scale

DMLite —> DPM and Dynafed

- DPM: traditional “multi-protocol GRID storage element” system based on a DB catalog and the concept of “disk pools” and a full-featured redirector node (“head node”)

- Dynafed: minimalistic, very flexible HTTP(s) dynamic redirector
  - We can see it as a DPM headnode without DOME and not needing a static catalog
    - The catalog is built on-the-fly in memory
  - We used to call this “storage federations”, yet it became more than this when someone realised it works very well with S3 and caches in general

- Both systems support things like TPC, X509, VOMS, Macaroons, OpenID-Connect
- DPM comes with a puppet-based pre-made do-it-all config
- Dynafed comes with good docs, no pre-made deployments

DPM Status - November 2020

- Total disk space: 111PB (BDII says 96PB)
- Number of instances: 88 (BDII says 71 in 59 sites)

- The old components (dpnsdaemon, srm, dpm-daemon, rfio…) are deprecated since September 2019
  - No urgency, they will stay in the EPEL7 repos, simply not being ported to EPEL8 and not updated with newer version numbers

- Well up to date with the features needed in WLCG (e.g. Macaroons and OpenID-Connect, cross-protocol checksums, TPC, easier multi-site, pools as caches)

- The DPM upgrade TF has tracked/promoted the upgrade progress and the enabling of the new components (DOME) and of the WLCG Storage Reporting Record
  - https://twiki.cern.ch/twiki/bin/viewauth/LCG/DPMupgrade

- Roadmap of stability: the tech goals have been reached, hence no revolutions are foreseen in the DPM development

- Given the stability of the platform, the dev deltas that we see will mostly be little fixes and polishments, e.g. CLI commands, security fixes, etc.

Dynafed status - November 2020

- Relevant deployments:
  - BOINC CMS@home to store files in S3 at CERN
  - Canadian S3-based cloud (with redirectors at CERN)
  - ECHO at RAL
  - Belle-II in Italy
- Differently from DPM, we don’t track Dynafed installations

- Can be configured with TPC and OpenID-Connect
- Particularly flexible, well documented authorization subsystem

Maintenance main ingredients

- Facilities at CERN:
  - GitLab repository group: https://gitlab.cern.ch/lcgdm
  - Jenkins build system: https://jenkins-lcgdm.web.cern.ch/
  - Semi-automatic test procedures
  - Small testbed
    - Seriously testing the development releases of a plethora of other components as a side effect, e.g. xrootd, voms, davix, gfal, etc.
- Precise, simple, written down test/release workflow
  - Pretty quick turnaround in case of urgency
  - A few brave sysadmins also help checking things in their sites
  - EPEL is the last step of the release
- In the last years the development effort has been negligible with respect to running these necessary steps

Publish to Fedora/EPEL8

- The code does not need fixes, or maybe just minimal things
- Petr Vokac and Mattias Ellert already contributed several needed bits
- The specfiles might need juggling
- The transition to Python3 is not fully complete yet
- Full status: https://twiki.cern.ch/twiki/bin/view/DPM/DpmEpel

- The effort would include upgrading the build/test system and the testbeds dpmhead-trunk and dpmhead-rc
  - Xlation: “fight with specfile, Openstack, Jenkins and Puppet”
  - There can be no releases without builds/tests!
  - This is not development

- CERN will not contribute directly the EPEL8 porting, as per statement
  - https://indico.cern.ch/event/813745/contributions/3766117/note/

Roles involved

- Which roles are necessary for a project in maintenance mode?
  1. Community (utilises, tests, discusses and contributes)
  2. Partial time dev (reviews contribs and contributes himself sometimes)
  3. QA manager (looks at the test clusters dpmhead-*, e.g. for DOMA tests or nightly internal ones, reviews the tests results)
  4. Build/test system maintenance (Jenkins+OpenStack+Puppet take some time to maintain)
  5. Epel master (for the final release steps)

- For DPM, points 2, 3, 4 have been inherited by Fabrizio
- PlZ consider this as a stopgap, the ownership should better be shared with the community

“New” EPEL packager

- Oliver Keeble moved to other things, and can’t anymore guarantee this role
- Many thanks to Andrea Manzi from EGI who kindly accepted to take this role

- The last version v1.14.2 had been pushed in the last weeks

- No burning fixes in the pipeline
- Emergencies (if any) are covered

Status

- Very stable services and codebases
- The code is in Gitlab, totally open
- For DPM, the scalability range is pacifically over the “tens of PBs”, the performance is more than adequate for this. Dynafed is one level up, depending on the usage, being a more lightweight-featured system
- Thanks to all the contributors !
- Will stay in EPEL7 until EPEL7 exists
- We see only reasons for DPM/Dynafed to work well in the next few years, depending on the interest of the community

- Oliver Keeble has been the previous EPEL packager, now it’s Andrea Manzi
- No plans so far to push to EPEL8, a non-CERN packager could do it

“Maintenance mode”

- On the CERN side the projects are in maintenance mode
  - Security fixes/updates if necessary
  - Following the reqs of DOMA TPC (e.g. OIDC, macaroons)
  - Slowly accepting external contributions
  - No new features planned
  - Build/test system stays alive, low maintenance
  - Build/test system being simplified (will be more similar to real sites)

Conclusions

- Modern DPMs profit from a healthy technical platform. This includes Dynafed as a successful spin-off
- The system accommodates all the current requirements (including TPC and bearer tokens) and is technically well-placed to accommodate future ones that may come

- A lot of effort has been put by CERN in order to enable the community to incrementally take ownership of these open source components. One of the key points has been the stability, and we are there
- On the CERN side the project is in “Maintenance mode”

- We suggest that sites in the medium term either get involved and promote a shared know-how or consider alternative scenarios
- We all put hope into the ongoing cooperation of the DPM/Dynafed community, WLCG, CERN and EGI so that such an important asset benefits sites for the years to come

Ref… The March statement

Position from the ST group and the CERN development perspective:

CERN, through the IT department, has been one of the main driving forces behind the Disk Pool Manager (DPM) design, development, evolution and support in the last 15 years. This made the DPM software a pillar for storage solution in a distributed scientific computing research environment in the LHC Worldwide Computing Grid as well as in other sciences.

Given the large scale and the increasing requirements for storage that the LHC experiments need in the coming years, CERN has developed the next generation of open source storage software (EOS) that is currently in Tier-0 production for all the LHC experiments and that is believed to sustain the increased performance demanded for the next LHC run.

CERN is very proud of the existence of a large community worldwide using DPM since several decades. As the community has expressed concerns on the future support that can be expected from CERN on DPM, this documents should summarize the position of the Storage group of the IT department.

The IT department will continue to support DPM on the existing deployments with the present set of features. CERN will continue to ensure to have a minimum set of skills to be able to ensure bug fixes on the current releases and the necessary coordination of the external contributions.

The Storage group is also committed to continue the implementation of the new authentication technology that will become necessary in the next generation of the WLCG software. This effort will be consolidated and released in an upcoming new version.

Finally, there is no plan from CERN to add additional features to DPM, and we encourage the community to continue to invest resources to develop further this package, for implementing additional features and to port it to newer operating systems as necessary.

#####

From the WLCG perspective:

* We should assume that operations with the existing DPM product and support effort at CERN and in the DPM community is sustainable for the duration of RUN-3
* We should initiate a discussion in the next couple of months (Lund was an opportunity, need now to think an alternative) to evaluate storage technologies with a production timescale by HL-LHC. The discussion should consider current storage- and caching-products.
* The evaluation should consider target functionalities for WLCG (e.g. the DOMA outcomes) as well as long term support and sustainability.

#####