DPM Evolution: a Disk Operations Management Engine for DPM

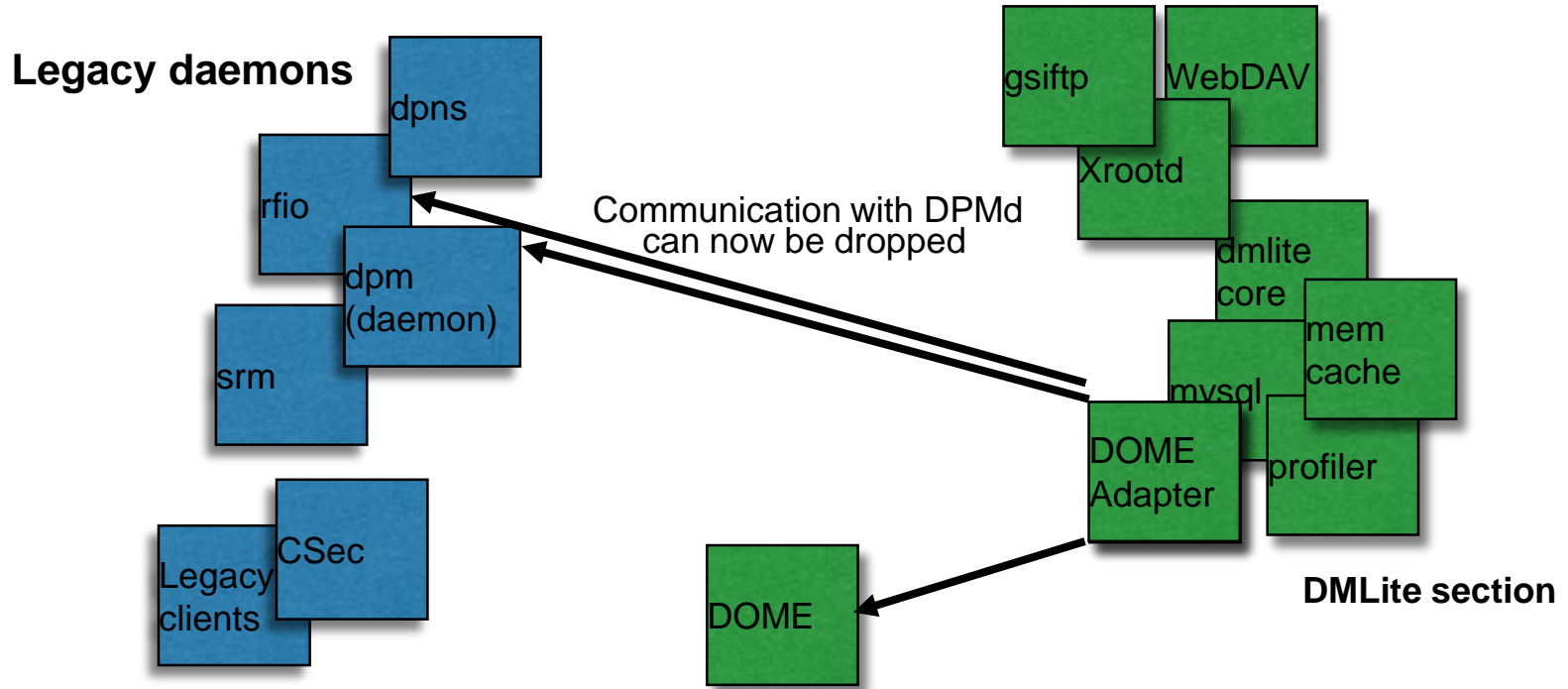Fabrizio Furano, Georgios Bitzes, Oliver Keeble (presenter), Andrea Manzi

Disk Pool Manager

- DPM is a system for managing disk storage at small and medium sites

- Over 70PB stored in the system at around 150 sites

- DPM has just finished a little revolution.
  - We now have our definitive platform for the future
  - This talk will describe that platform

The 3rd Generation DPM

- 1st generation – derived from Castor disk pool
- 2nd generation – introduced dmlite
  - Internal framework abstracting many functions, enabling multiple new frontends
- 3rd generation – introduces Dome
  - DPM 1.9 "Legacy flavour"
    - Legacy services still running
  - DPM 1.9 "Dome flavour"
    - Legacy services retired
      - No more dpmd, dpns, rfio, srm, csec

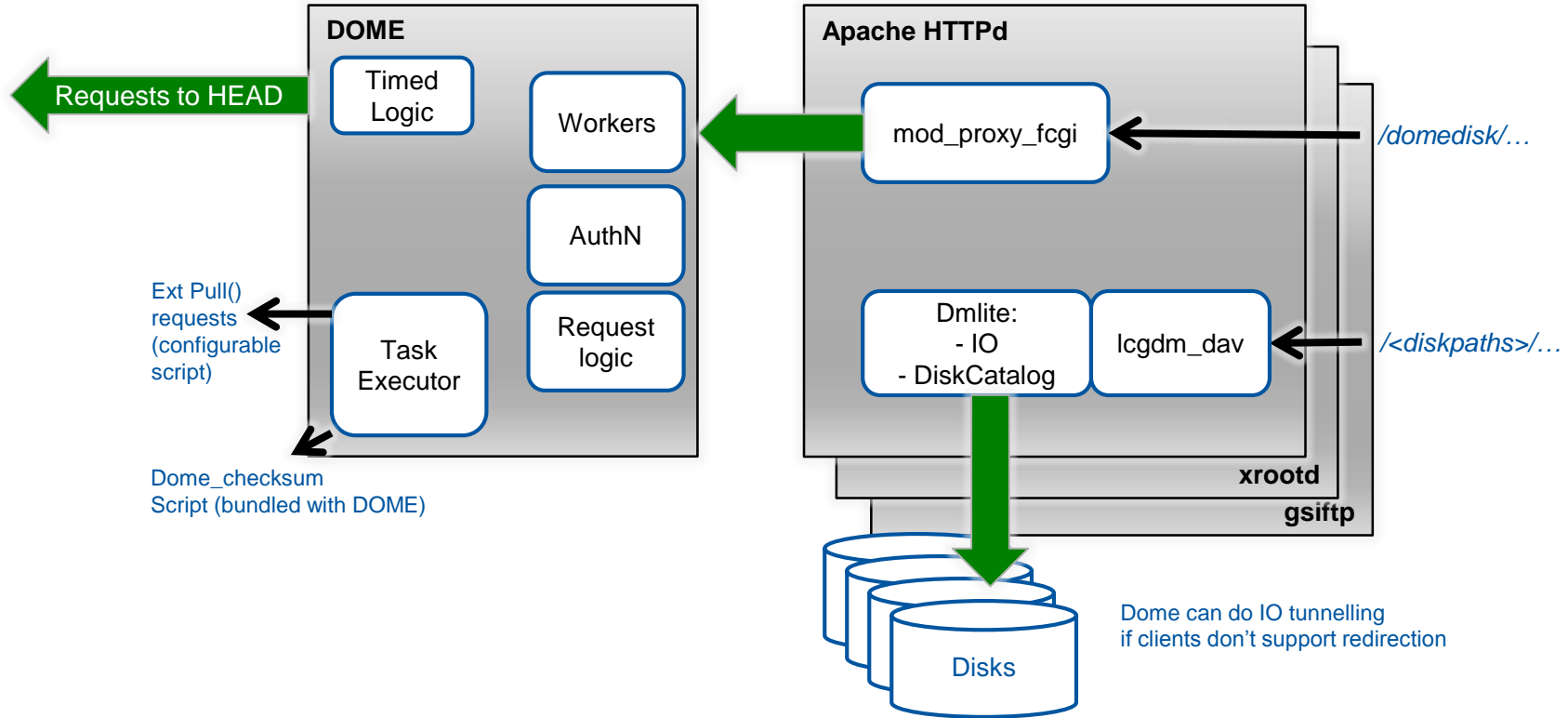# DPM headnode architecture

Legacy daemons

dpns

rfio

dpm
(daemon)

srm

Legacy
clients

CSec

Communication with DPMd
can now be dropped

gsiftp

WebDAV

Xrootd

dmlite
core

mem
cache

mysql

DOME
Adapter

profiler

DOME

DMLite section

DPM 1.9 "Dome flavour"

- Activate Dome, what do you get?
  - SRM free operation
    - Quotas
      - "space token" concept generalised and mapped into namespace
    - Space reporting – used/free via HTTP/DAV
      - Reporting on "space tokens" and subdirectories
    - GridFTP redirection enables scalability
  - Caching hooks
  - Simplified system
    - All internal communication over HTTPS
      - Control and data tunnelling
  - Improved dmlite-shell

What does Dome do?

domesetquotatoken
domeaddpool
domegetspaceinfo
domegetreplicainfo
domedocheck sum
domemakespace
domedelreplica
domegetidmap
domepull
domegetstatin fo
domechksum
domegetquotatok en
domepfnrm
domeget
domeput
domestatpfn
domegetdir
domestatpool
domegetuser
domechksumstatus
domeupdatexattr
domeputdone
domermpool
domermfs
domegetdirspaces
domedelquotatok en
domemodif yfs

Disk server (simplified)

DOME

Timed Logic

Workers

AuthN

Request logic

Task Executor

Requests to HEAD

Ext Pull() requests (configurable script)

Dome_checksum Script (bundled with DOME)

Apache HTTPd

mod_proxy_fcgi

Dmlite:
- IO
- DiskCatalog

lcgdm_dav

xrootd

gsiftp

/domedisk/…

/<diskpaths>/…

Disks

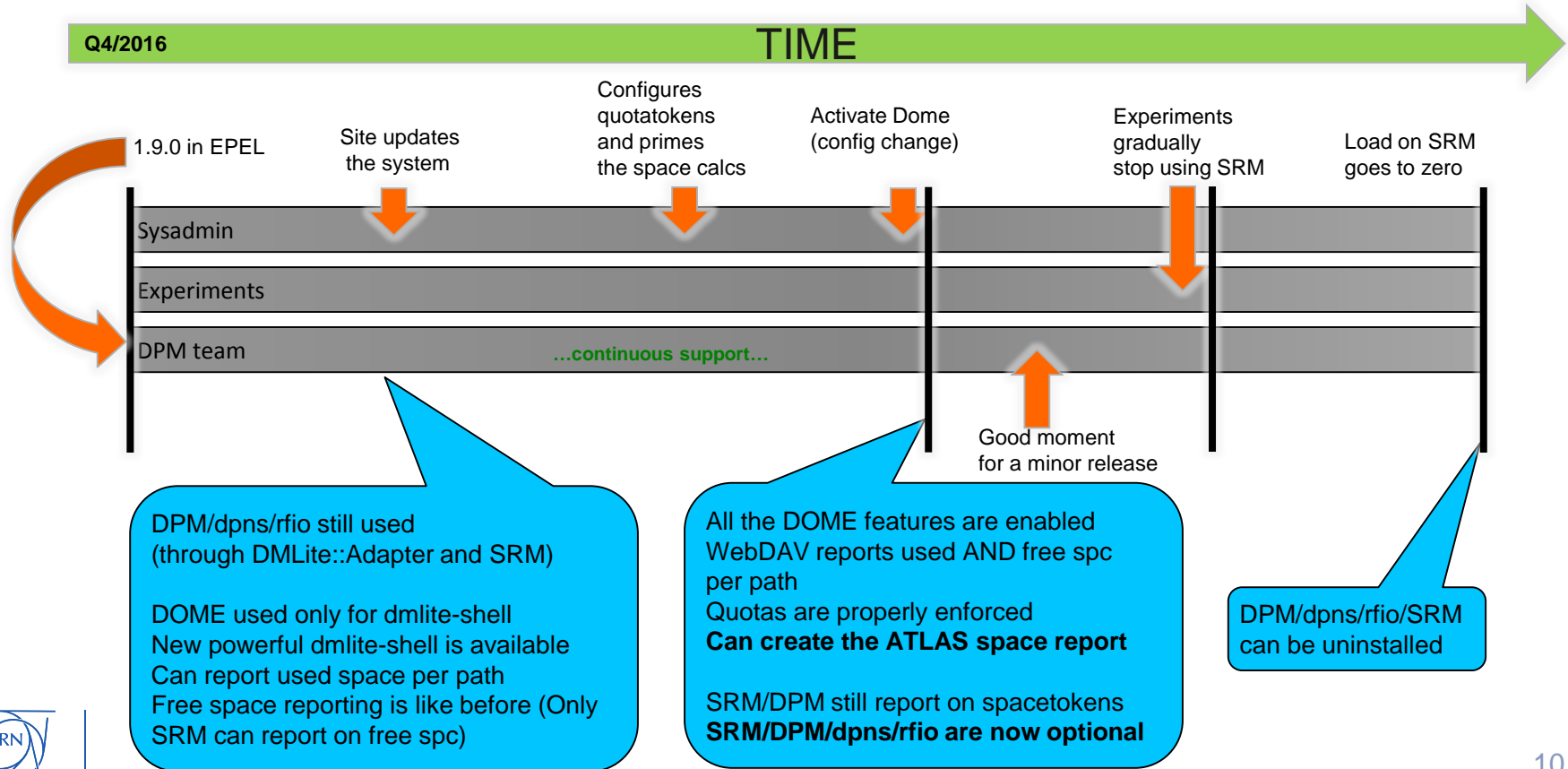Dome can do IO tunnelling if clients don't support redirection

8

Used/free space: WebDAV

```
$ davix-http -P grid -X PROPFIND --header 'Depth: 0' --header 'Content-Type: text/xml;
charset=UTF-8' "https://domehead-trunk.cern.ch/dpm/cern.ch/home/dteam" --data '<?xml
version="1.0" ?><D:propfind xmlns:D="DAV:"><D:prop><D:quota-used-bytes/><D:quota-
available-bytes/></D:prop></D:propfind>'


<?xml version="1.0" encoding="utf-8"?>
<D:multistatus xmlns:D="DAV:" xmlns:ns0="DAV:">
<D:response xmlns:lp1="DAV:" xmlns:lp2="http://apache.org/dav/props/"
xmlns:lp3="LCGDM:">
<D:href>/dpm/cern.ch/home/dteam/</D:href>
<D:propstat>
<D:prop>
<lp1:quota-used-bytes>24677181319</lp1:quota-used-bytes>
<lp1:quota-available-bytes>75322818681</lp1:quota-available-bytes>
</D:prop>
<D:status>HTTP/1.1 200 OK</D:status>
</D:propstat>
</D:response>
</D:multistatus>
```

Going SRM-less with your DPM

TIME

Q4/2016

1.9.0 in EPEL

Site updates the system

Configures quotatokens and primes the space calcs

Activate Dome (config change)

Experiments gradually stop using SRM

Load on SRM goes to zero

Sysadmin

Experiments

DPM team

…continuous support…

Good moment for a minor release

DPM/dpns/rfio still used
(through DMLite::Adapter and SRM)

DOME used only for dmlite-shell
New powerful dmlite-shell is available
Can report used space per path
Free space reporting is like before (Only
SRM can report on free spc)

All the DOME features are enabled
WebDAV reports used AND free spc
per path
Quotas are properly enforced
Can create the ATLAS space report

SRM/DPM still report on spacetokens
SRM/DPM/dpns/rfio are now optional

DPM/dpns/rfio/SRM
can be uninstalled

Caching laboratory

- DPM 1.9 with Dome will allow investigation of operating WLCG storage as a cache
- Scenarios
  - Data origin a regional federation of associated sites
  - Data origin the global federation
- A volatile pool can be defined which calls out to a stager on a miss
  - Caching logic implemented in a pluggable way
  - Hybrid cache/conventional setup
- Questions to investigate
  - Cache management logic
  - Different client strategies on miss
    - blocking read, async read, redirection to origin
  - Authentication solutions
  - Workflow adaptation for locality

Summary

- DPM 1.9 "Dome flavour" represents DPM's platform for the future.
- Full modernisation of the stack complete
  - SRM-less operation supported
  - Improved system management
  - Caching operation possible
  - Significant reduction in maintenance and operation costs

References

- http://lcgdm.web.cern.ch/
- Contact us via GGUS or hep-service-dpm@cern.ch
- Join dpm-users-forum@cern.ch for further discussion
  - https://e-groups.cern.ch/e-groups/EgroupsSubscription.do?egroupName=dpm-users-forum
- DPM Workshop 2016, 23-24 Nov, Paris
  - https://indico.cern.ch/event/559673/