



DPM: Future Proof Storage

Ricardo Rocha

(on behalf of the DPM team)

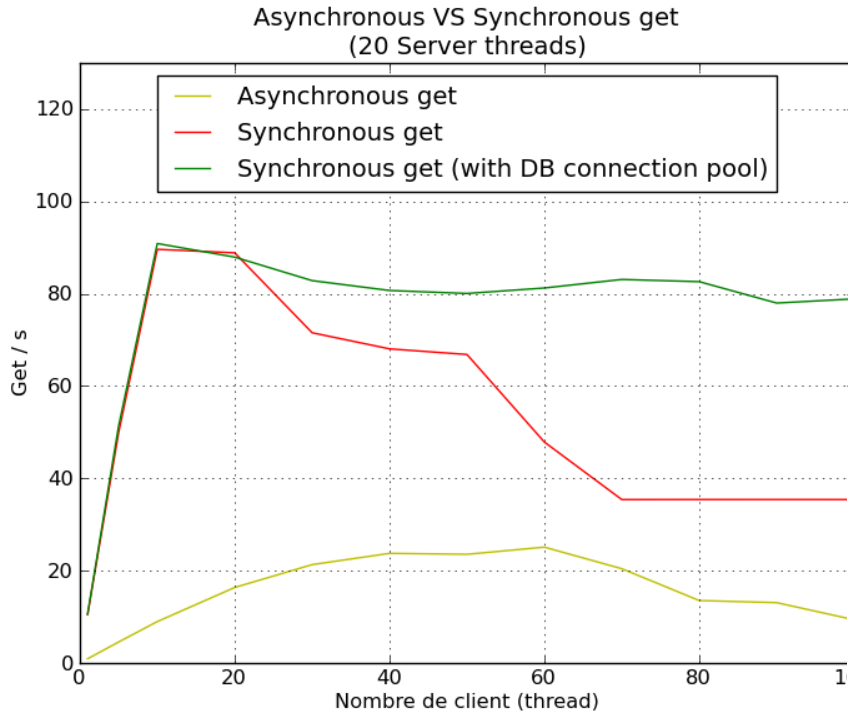
Goals

- Provide a lightweight, grid-aware storage solution
- Simplify life of users and administrators
- Improve the feature set and performance
- Use **standard protocols**
- Use **standard building blocks**
- Allow easy integration with new tools and systems

Quick News

- Current production version is 1.8.3
 - <https://svnweb.cern.ch/trac/lcgdm/blog/official-release-lcgdm-183>
 - In EPEL, EMI1 and EMI2
 - What's included
 - EPEL compliance
 - New HTTP/WebDAV frontend (old DPM httpd is gone)
 - Thread safe clients
 - Synchronous GET requests
 - ...
- What's coming next (1.8.4)
 - New DPM XROOT plugin
 - **Federation aware** (already deployed in several UK sites)
 - DMLITE with all its goodies (much more on this later)
 - Improved performance, easy integration with standard tools, ...

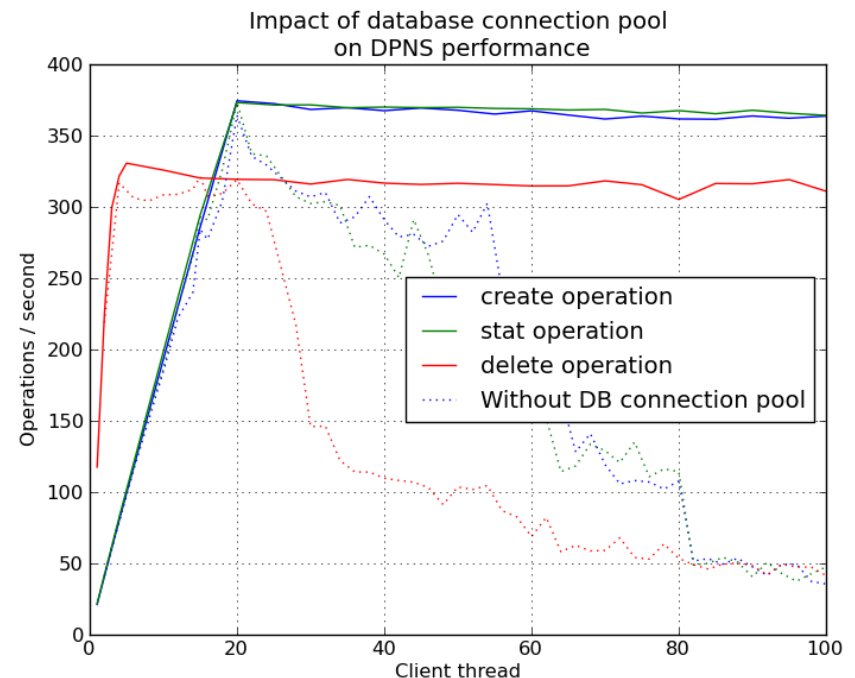
Quick News



- What's coming next (1.8.4)
 - New DPM XROOT plugin
 - DMLITE with all its goodies (m
 - Improved performance, easy int

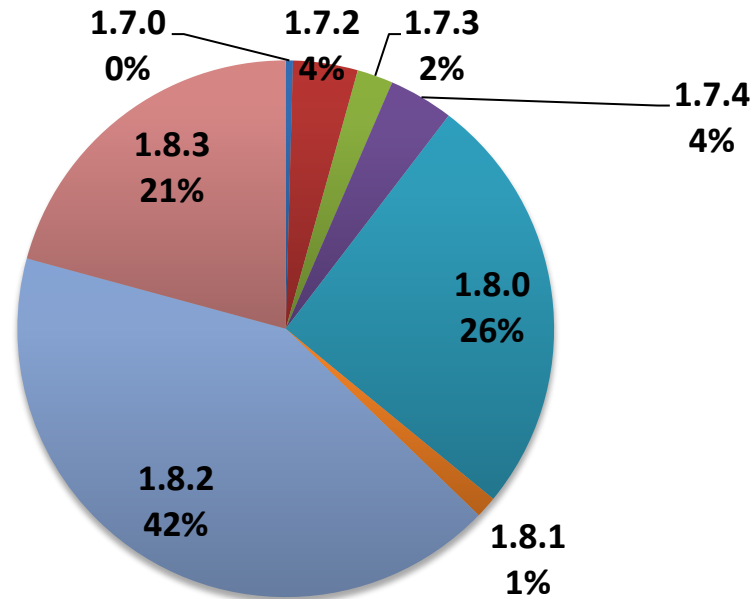
3.3

<http://blog/official-release-lcgdm-183>

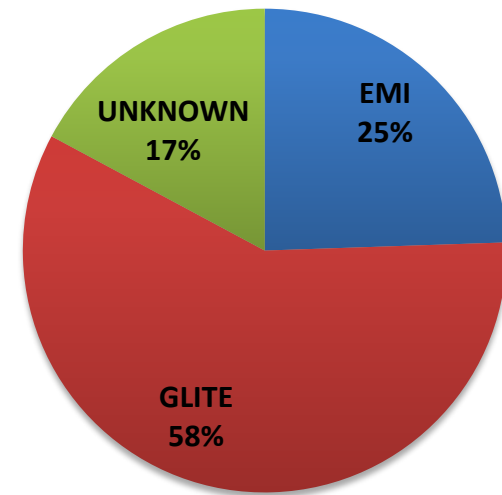


Deployment

Versions



Distributions



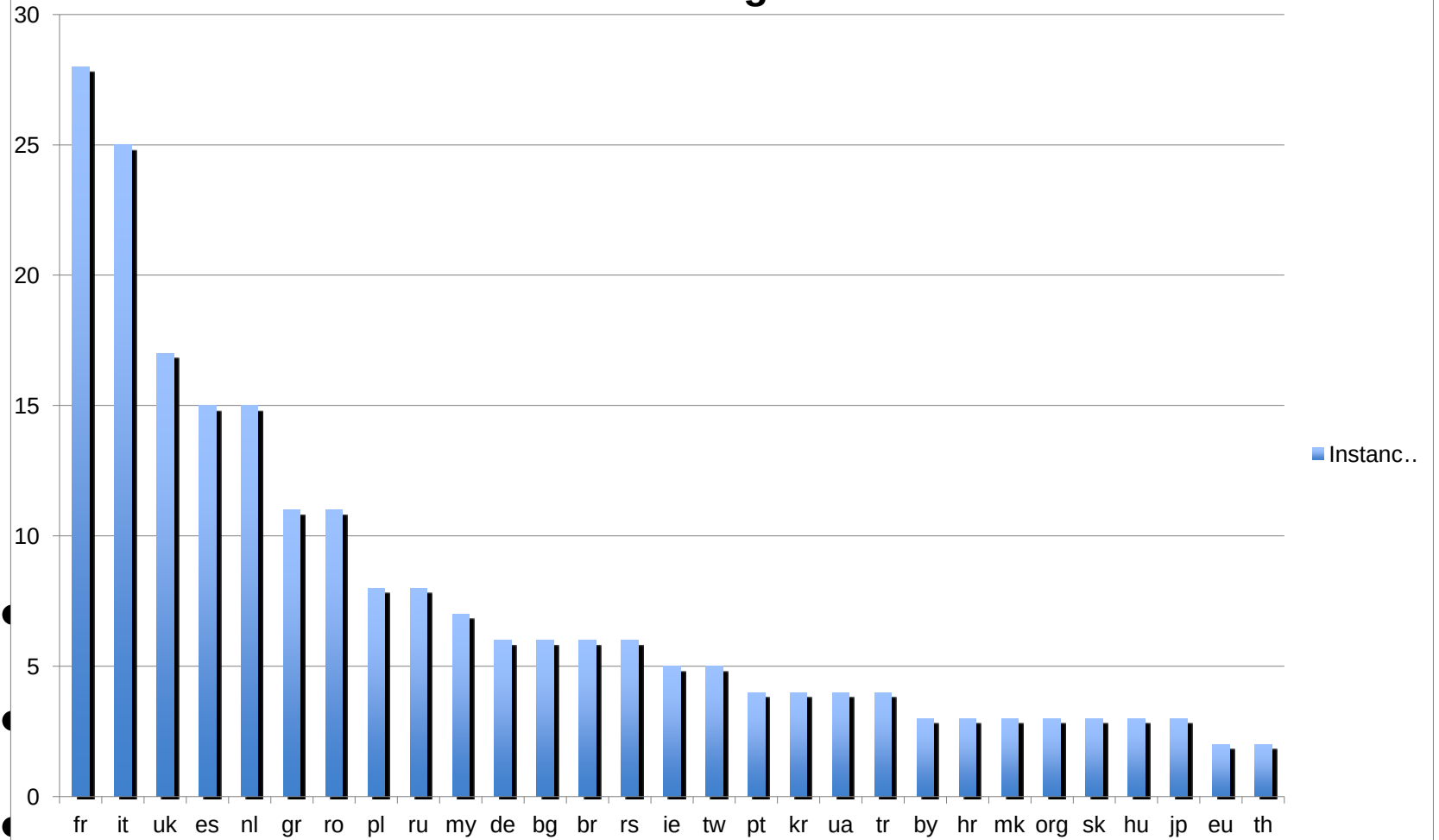
- 36 PB (10 sites with > 1PB)
- Over 200 sites in 50 regions
- Over 300 VOs

Deployment

Versions

Distributions

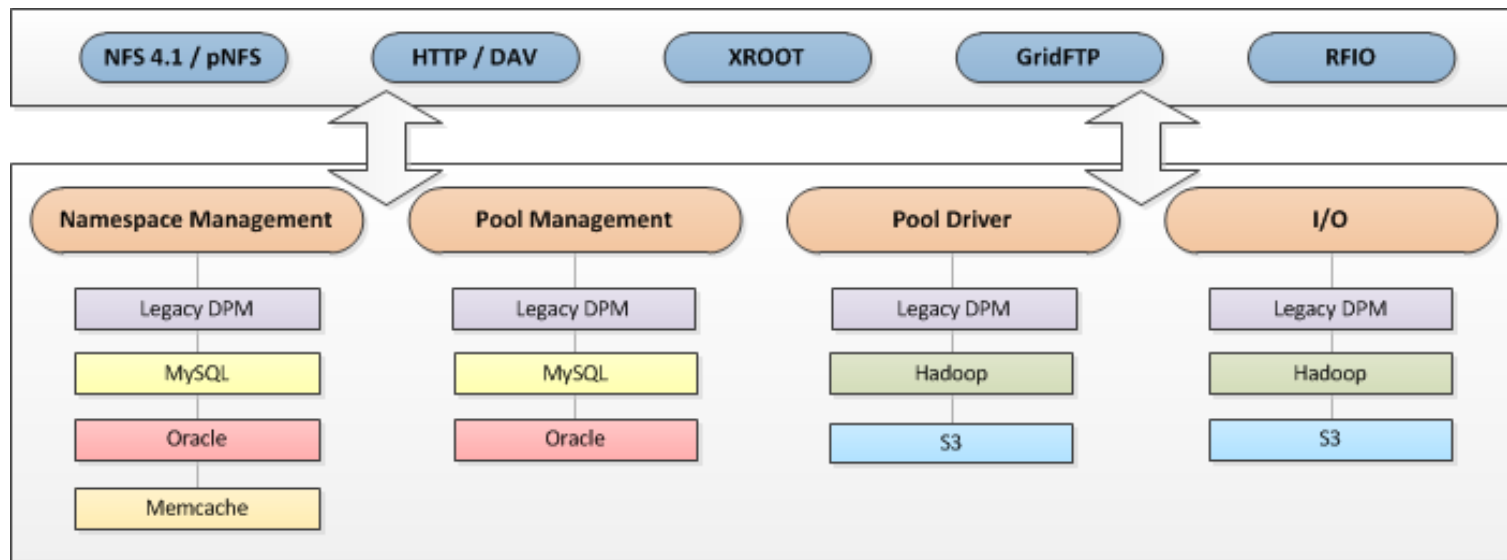
Per Region



gLite to EMI migration

- 25% of the sites have already done it
 - At least for the Head Node
- A multi-flavor setup works fine
 - You can upgrade Head/Disk nodes independently
- Why do it?
 - New features go to EMI only
 - Support ... see Markus's talk
- A procedure? Yes!
 - <https://svnweb.cern.ch/trac/lcgdm/wiki/Dpm/Admin/Glite2EMI>

Refactoring & DMLITE



- DMLITE is the result of a significant refactoring effort
- Better separation between frontends and backend
 - Cleaner, more open, much improved performance
- Improved integration with standard building blocks
 - Hadoop, Memcache, S3, Lustre, ...

Improved Frontends

- Standard protocols, standard clients
 - HTTP/DAV, NFS 4.1/pNFS
 - Ubiquitous access to grid storage
- But we ensure other protocols also benefit from our recent improvements
 - GridFTP, XROOT will also rely on DMLITE

Improved Frontends

Browser window showing the DPM frontend interface for the path `/dpm/cern.ch/home/dteam/`. The interface displays a table of files and folders with columns for Mode, UID, GID, Size, Modified, and Name.

Mode	UID	GID	Size	Modified	Name
-rwxrwxr-x	9	1	20	Wed, 25 Apr 2012 12:34:01 GMT	a.alvarez.auth
-rwxrwxr-x	9	1	5	Wed, 25 Apr 2012 12:30:57 GMT	a.beche.auth
-rw-rw-r--	1	1	175.0K	Tue, 17 Apr 2012 07:52:14 GMT	a.devresse.thx
-rwxrwxr-x	1	1	463.2K	Thu, 14 Jun 2012 13:04:51 GMT	cave.png
-rwxrwxr-x	1	1	252	Mon, 07 May 2012 14:23:53 GMT	
drwxrwxr-x	0	1	115	Fri, 01 Jun 2012 02:43:51 GMT	
-rwxrwxr-x	1	1	760.2M	Fri, 11 May 2012 11:20:29 GMT	
-rwxrwxr-x	9	1	27	Fri, 04 May 2012 15:41:56 GMT	
-rwxrwxr-x	1	1	61.7K	Fri, 11 May 2012 07:37:46 GMT	
-rwxrwxr-x	9	1	27	Thu, 03 May 2012 11:58:47 GMT	
-rwxrwxr-x	9	1	27	Mon, 16 Apr 2012	

Request by /DC=ch/DC=cern/OU=Organic Units/OU=Users/CN=aalvarez/CN=678984/CN=A
Powered by LCGDM-DAV (Arioch)

and clients

range

File manager window titled "dteam" showing a directory view of the `dteam` directory. The window includes a menu bar (File, Edit, View, Go, Bookmarks, Help) and a sidebar showing the file system hierarchy. The main pane displays a grid of files and folders, including:

- generated
- perfsuite
- private
- public
- a.alvarez.auth
- a.beche.auth
- a.devresse.thx
- f.furano.auth
- group.test.hc
- hadoop.cpp
- higgs.jpg
- http.std
- image.jpg
- m.hellmich.auth
- o.keeble.auth
- r.rocha.auth
- s3.cpp
- shu.ting.thx
- webdav.std

The status bar at the bottom indicates: "higgs.jpg" selected (63.2 kB)

Improved Frontends

- Sta

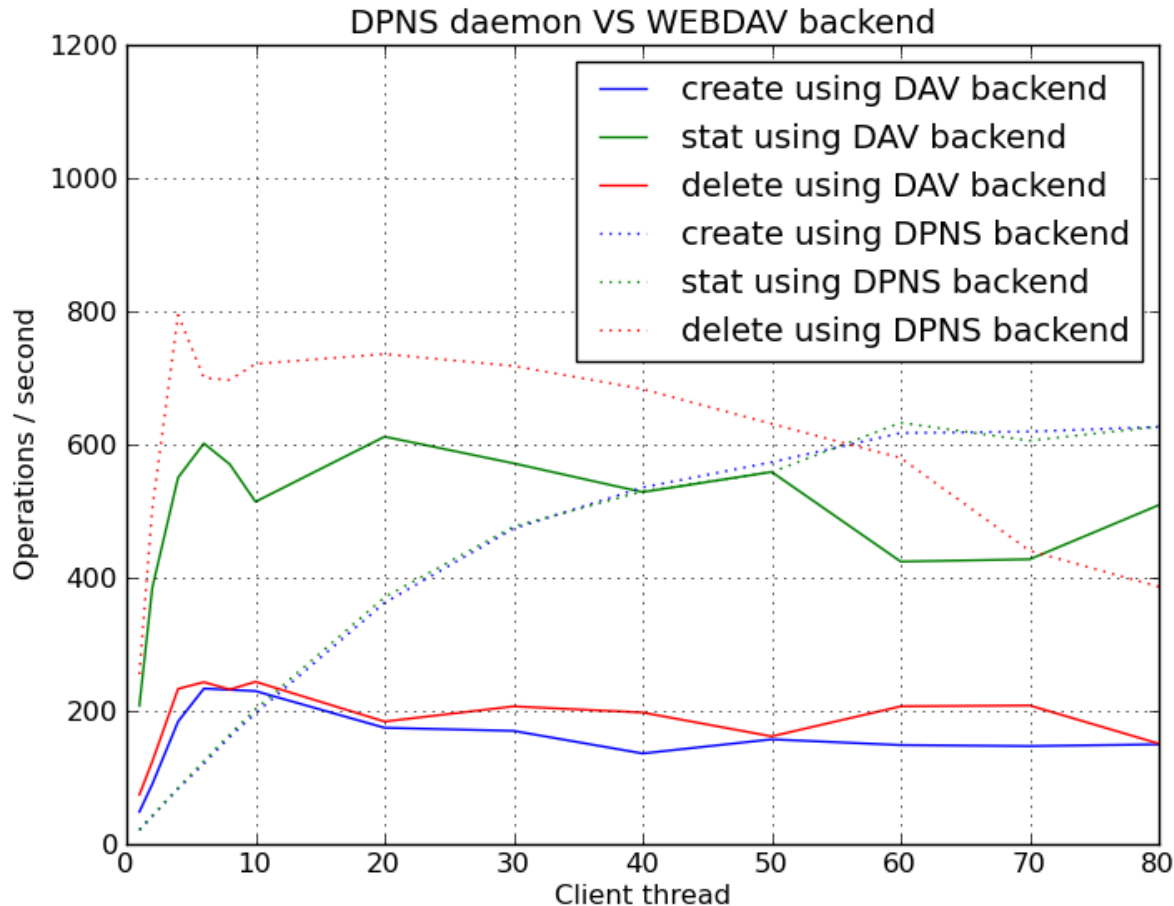
— t

— l

- But

fro

— c

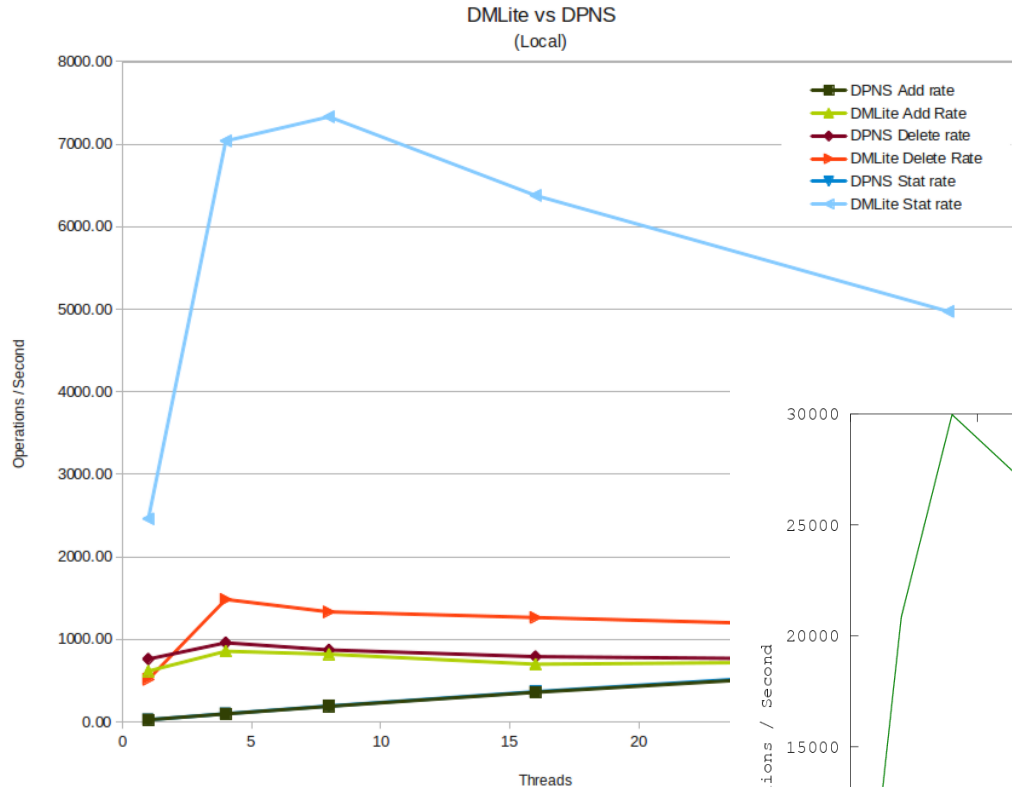


fit

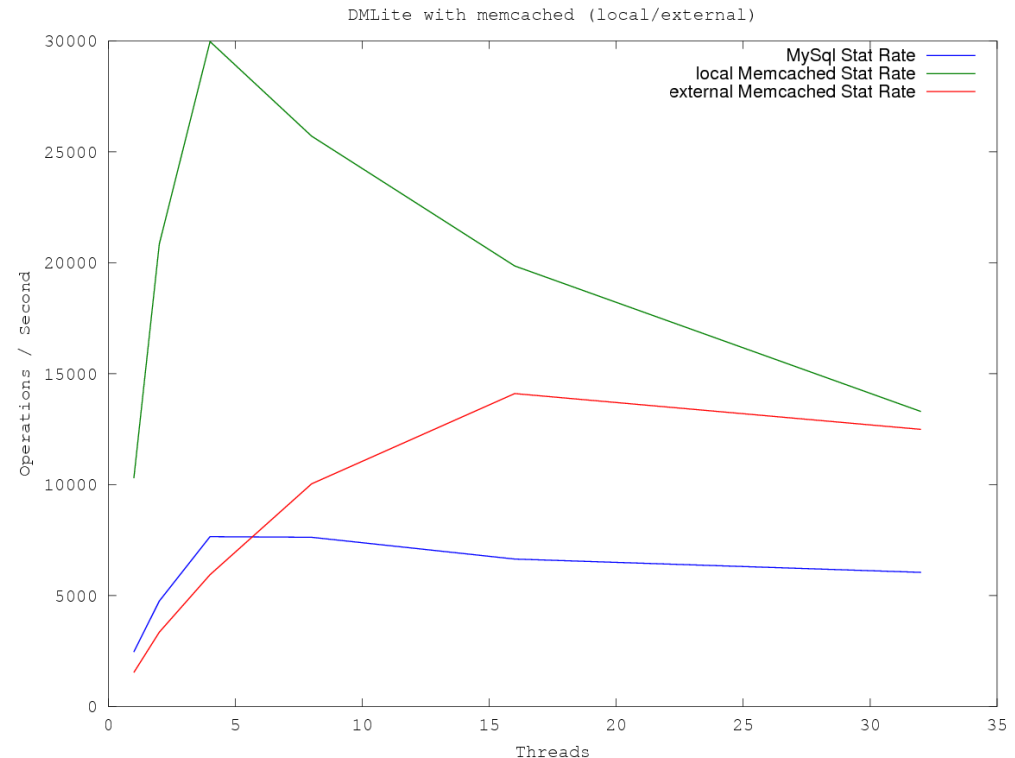
Improved Backend

- This is where DMLITE shines
 - Plugin based, open for constant evolution
 - Single API for all frontends and tools
- Improved nameserver performance
 - Connection pooling, improved SQL, memcache layer, ...
- Support for multiple pool types
 - Legacy DPM, Hadoop/HDFS, S3, ...
 - Sharing a single namespace if desired
 - Possibility for *opportunistic pools*
- And this is the beginning, much more coming
 - Python bindings, Lustre, VFS, ...

Improved Backend



evolution
tools



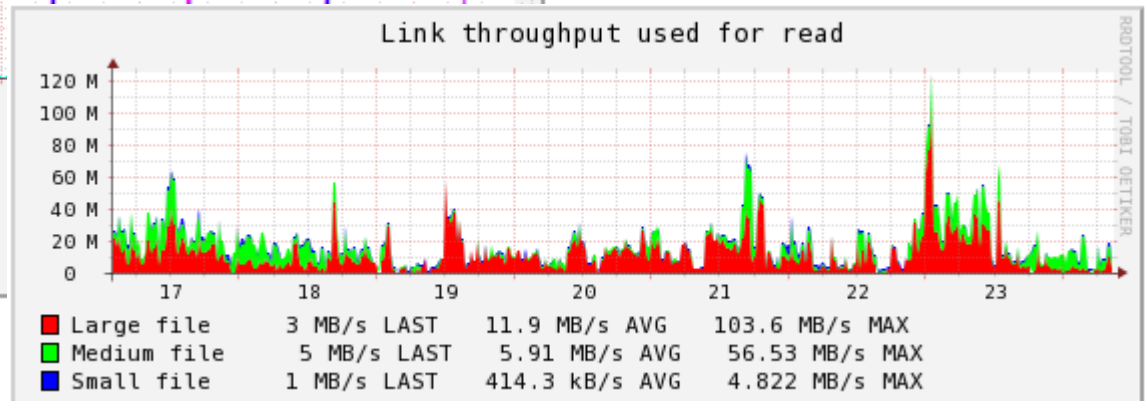
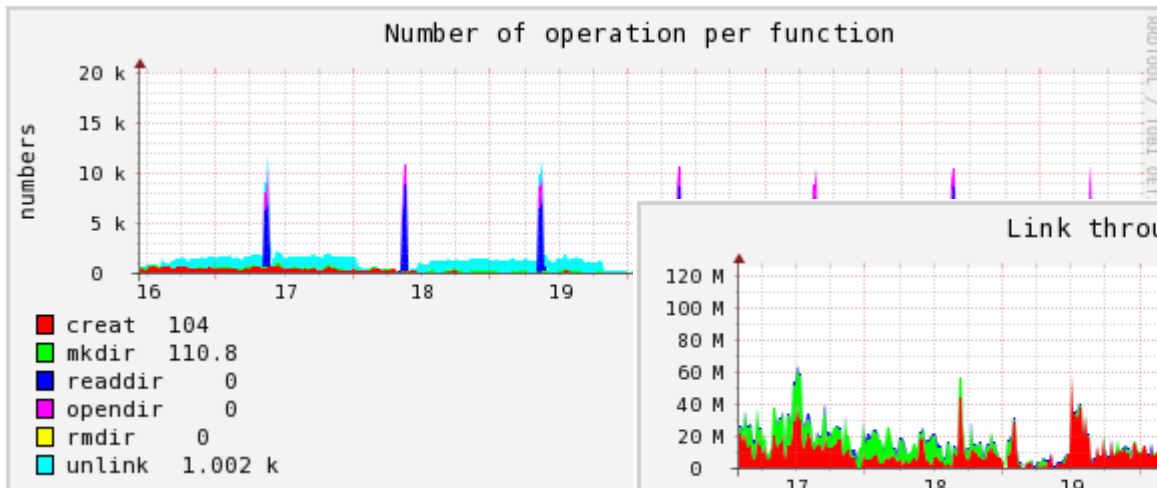
- sharing a single name
- Possibility for *opport*
- And this is the begin
 - Python bindings, Lus

Easy administration

- Puppet for configuration
 - Popular among large data centers
 - Lots of modules for popular tools (which we now rely on – apache, memcache, nagios, ...)
- Nagios for monitoring
 - We reuse as much as possible
 - And there's a lot already available
 - Added specific plugins for detailed status and performance monitoring

Easy administration

```
node 'dpm01.cern.ch' inherits dpm-service {  
  include dpm::headnode  
  include dpm::nfsserver  
  include dpm::nagios::headnode  
  
  # setup supported domain/vo(s)  
  dpm::headnode::domain { 'cern.ch': require => Service['dpmns'], }  
  dpm::headnode::vo { 'dteam': domain => 'cern.ch', require => Dpm::Headnode::Domain['cern.ch'], }  
  
  # setup pools  
  dpm::headnode::pool { 'pool1': require => Service['dpm'] }  
}
```



Summary

- DPM provides standards based data access
 - HTTP/DAV, NFS 4.1/pNFS
 - While remaining *grid-aware*
 - X509, GridFTP, XROOTD, ...
- Future proof with DMLITE
 - Easily extensible, plugin based
 - Open to external contributions
 - Reusing standard building blocks where possible
 - Apache, Memcache
 - Integration with Hadoop/HDFS, S3, ...

<http://cdsweb.cern.ch/record/1458022/>