



HTTP Ecosystem & Federation Brainstorming

Oliver Keeble
Alexandre Beche, Fabrizio
Furano, Stefan Roiser
22/11/2013



Ecosystem components

- HTTP/DAV support in storage
 - dCache
 - DPM
 - LFC too
 - EOS
 - “will adapt xrd-http later” – different solution for now (NGINX proxy)
 - StoRM
 - <http://italiangrid.github.io/storm/documentation/webdav-guide/>
 - Xrootd
 - xrd-http currently being integrated into xrootd 4 (to be released)

Ecosystem components

- Access library
 - Gfal2 support
 - Davix – access library addressing shortcomings of existing clients
 - NB – existing clients will still work within their scope!
 - TDavixFile for ROOT
 - Performance numbers are promising...
- Transfer
 - FTS3 support
 - 3rd party copy implemented by DPM (dCache forthcoming)
- Federation

https://arioch.cern.ch/dpm/cern.ch/home/dteam/

/dpm/cern.ch/home/dteam/

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.
-rw-rw-r--	1	1	175.0K	Tue, 17 Apr 2012 07:52:14 GMT	a.devres
-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	f.furano
drwxrwxr-x	0	1	115	Fri, 01 Jun 2012 02:43:51 GMT	generate
-rwxrwxr-x	1	1	760.2M	Fri, 11 May 2012 11:20:29 GMT	group.te
-rwxrwxr-x	9	1	27	Fri, 04 May 2012 15:41:56 GMT	hadoop.c
-rwxrwxr-x	1	1	61.7K	Fri, 11 May 2012 07:37:46 GMT	higgs.jp
-rwxrwxr-x	9	1	27	Thu, 03 May 2012 11:58:47 GMT	http.std
-rwxrwxr-x	1	1	175.0K	Mon, 16 Apr 2012	image.jp

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

HTTP / DAV

- A browser click will download your file. Also...
- > aria2c <https://fed.cern.ch/mydata?metalink>
 - Parallel download from multiple replicas

DynaFed

- Dynamically federates HTTP endpoints
 - Including other catalogues
- Fast in-memory namespace cache
- Transparent redirection for clients
 - Closest replica chosen (geoip)
 - Other info sources could be integrated (eg perfsonar)
- Officially released and under evaluation
 - EUDAT, Victoria (CA)

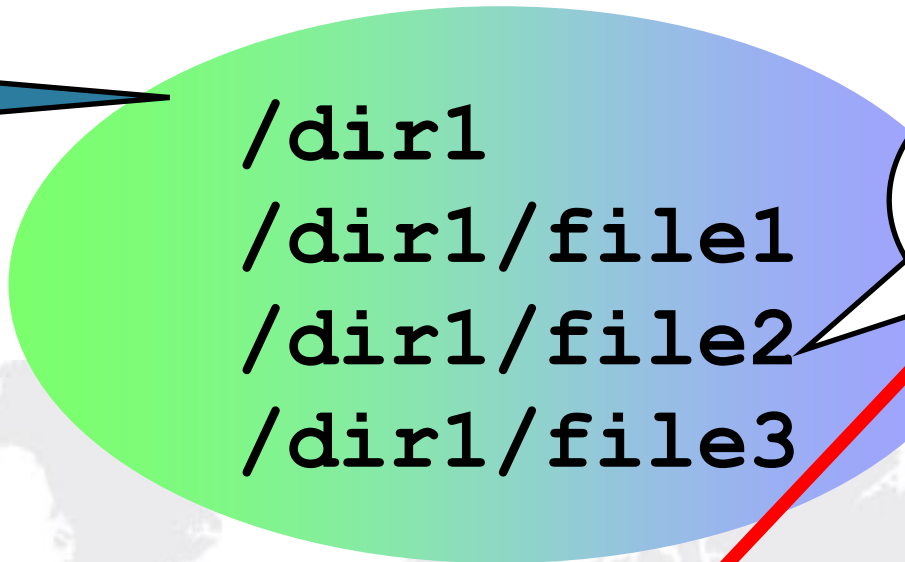
Aggregation

This is what we want to see as users

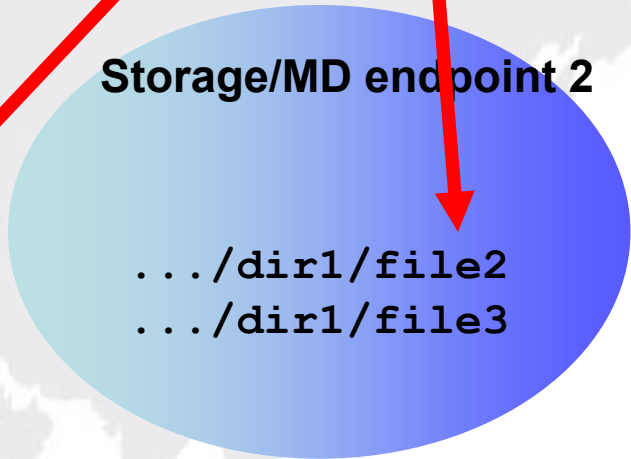
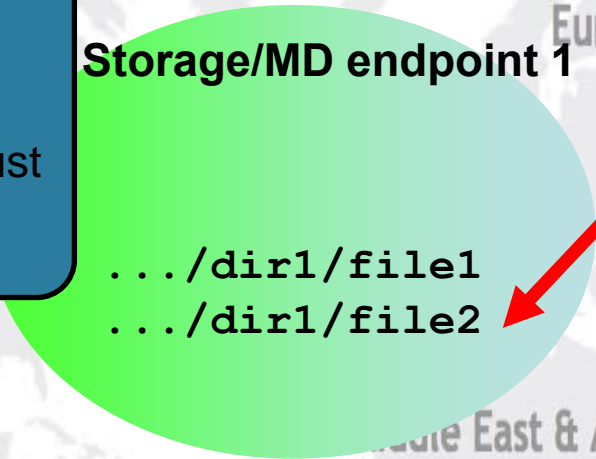
Sites remain independent and participate to a global view

All the metadata interactions are hidden and done on the fly

NO metadata persistency needed here, just efficiency and parallelism



With 2 replicas



http://federation.desy.de/fed

The screenshot shows a Mozilla Firefox browser window displaying a directory listing for the path `/fed/atlas/dq2/data13_8TeV/ESD/f511/data13_8TeV.00217556.physics_Background.recon.ESD.f511/`. The browser's address bar shows the URL `Federation.desy.de/fed/atlas/dq2/data13_8TeV/ESD/f511/data13_8TeV.00217556.physics_Background.recon.ESD.f511/`. The page content includes a table with columns for Mode, UID, GID, Size, Modified, and Name. Each entry in the table represents a file with a size ranging from 1.4M to 17.1M and a modification date of Sun, 26 Nov 4431684 23:46:32 GMT. The file names are `data13_8TeV.00217556.physics_Background.recon.ESD.f511_lb0002_SF0-ALL_0001.1` through `data13_8TeV.00217556.physics_Background.recon.ESD.f511_lb0019_SF0-ALL_0001.1`. At the bottom of the page, there is a footer that reads "Request by nobody (nobody) Powered by LCGDM-DAV 0.14.0".

Mode	UID	GID	Size	Modified	Name
-rwxrwxr-x	0	0	1.4M	Sun, 26 Nov 4431684 23:46:32 GMT	data13_8TeV.00217556.physics_Background.recon.ESD.f511_lb0002_SF0-ALL_0001.1
-rwxrwxr-x	0	0	3.6M	Sun, 26 Nov 4431684 23:46:32 GMT	data13_8TeV.00217556.physics_Background.recon.ESD.f511_lb0003_SF0-ALL_0001.1
-rwxrwxr-x	0	0	17.1M	Sun, 26 Nov 4431684 23:46:32 GMT	data13_8TeV.00217556.physics_Background.recon.ESD.f511_lb0004_SF0-ALL_0001.1
-rwxrwxr-x	0	0	5.9M	Sun, 26 Nov 4431684 23:46:32 GMT	data13_8TeV.00217556.physics_Background.recon.ESD.f511_lb0005_SF0-ALL_0001.1
-rwxrwxr-x	0	0	111.6M	Sun, 26 Nov 4431684 23:46:32 GMT	data13_8TeV.00217556.physics_Background.recon.ESD.f511_lb0006_SF0-ALL_0001.1
-rwxrwxr-x	0	0	111.3M	Sun, 26 Nov 4431684 23:46:32 GMT	data13_8TeV.00217556.physics_Background.recon.ESD.f511_lb0007_SF0-ALL_0001.1
-rwxrwxr-x	0	0	111.3M	Sun, 26 Nov 4431684 23:46:32 GMT	data13_8TeV.00217556.physics_Background.recon.ESD.f511_lb0008_SF0-ALL_0001.1
-rwxrwxr-x	0	0	111.3M	Sun, 26 Nov 4431684 23:46:32 GMT	data13_8TeV.00217556.physics_Background.recon.ESD.f511_lb0009_SF0-ALL_0001.1
-rwxrwxr-x	0	0	111.3M	Sun, 26 Nov 4431684 23:46:32 GMT	data13_8TeV.00217556.physics_Background.recon.ESD.f511_lb0010_SF0-ALL_0001.1
-rwxrwxr-x	0	0	111.6M	Sun, 26 Nov 4431684 23:46:32 GMT	data13_8TeV.00217556.physics_Background.recon.ESD.f511_lb0011_SF0-ALL_0001.1
-rwxrwxr-x	0	0	110.9M	Sun, 26 Nov 4431684 23:46:32 GMT	data13_8TeV.00217556.physics_Background.recon.ESD.f511_lb0012_SF0-ALL_0001.1
-rwxrwxr-x	0	0	110.9M	Sun, 26 Nov 4431684 23:46:32 GMT	data13_8TeV.00217556.physics_Background.recon.ESD.f511_lb0013_SF0-ALL_0001.1
-rwxrwxr-x	0	0	110.9M	Sun, 26 Nov 4431684 23:46:32 GMT	data13_8TeV.00217556.physics_Background.recon.ESD.f511_lb0014_SF0-ALL_0001.1
-rwxrwxr-x	0	0	111.3M	Sun, 26 Nov 4431684 23:46:32 GMT	data13_8TeV.00217556.physics_Background.recon.ESD.f511_lb0015_SF0-ALL_0001.1
-rwxrwxr-x	0	0	111.2M	Sun, 26 Nov 4431684 23:46:32 GMT	data13_8TeV.00217556.physics_Background.recon.ESD.f511_lb0016_SF0-ALL_0001.1
-rwxrwxr-x	0	0	111.3M	Sun, 26 Nov 4431684 23:46:32 GMT	data13_8TeV.00217556.physics_Background.recon.ESD.f511_lb0017_SF0-ALL_0001.1
-rwxrwxr-x	0	0	111.2M	Sun, 26 Nov 4431684 23:46:32 GMT	data13_8TeV.00217556.physics_Background.recon.ESD.f511_lb0018_SF0-ALL_0001.1
-rwxrwxr-x	0	0	111.2M	Sun, 26 Nov 4431684 23:46:32 GMT	data13_8TeV.00217556.physics_Background.recon.ESD.f511_lb0019_SF0-ALL_0001.1

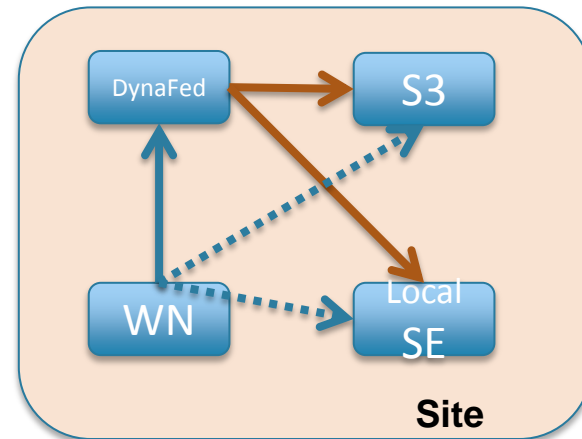
Request by nobody (nobody)
Powered by LCGDM-DAV 0.14.0

Use Case #1: Storage sharing

- Two well connected sites can share a dataset
 - Esp effective if each site concentrates on a subset but needs access to the whole
- Advantages
 - Transparent access via clients
 - Closest replica chosen in case of duplicates
 - Efficient use of storage
- Can be scaled up to national level or beyond
 - Additional monitoring work would be required (both for endpoints and federator).

Use Case #2: Integrating Cloud Resources

- A site wishes to elastically expand storage to support a short campaign
 - S3 storage is provisioned (privately or publically).
 - Data is imported
- Federation provides an integrated view of existing storage and elastic storage through namespace integration and translation



Use Case #3: Integration with caching

- eg Pure cache site
- Reuse existing cache tech
- Much easier if data is “group readable”
- Integration with CDN

