



DPM Status

Petr Vokáč

ATLAS Sites Jamboree

7th March 2019

DPM overview

- Popular grid storage system
 - multi-protocol
 - SRM, GridFTP, XRootD WebDAV
 - multi-vo
 - X.509 / VOMS and token based authentication, ACL
 - scalable to medium size sites ~ 10PB
 - headnode – maintains metadata, namespace and ACL + redirect clients
 - disknodes – data storage
 - built on top of common software
 - XRootD, Apache, Globus, MySQL
 - supported by CERN IT
 - relatively easy to deploy (puppet)
 - integrated with WLCG workflows



ATLAS Storage endpoints

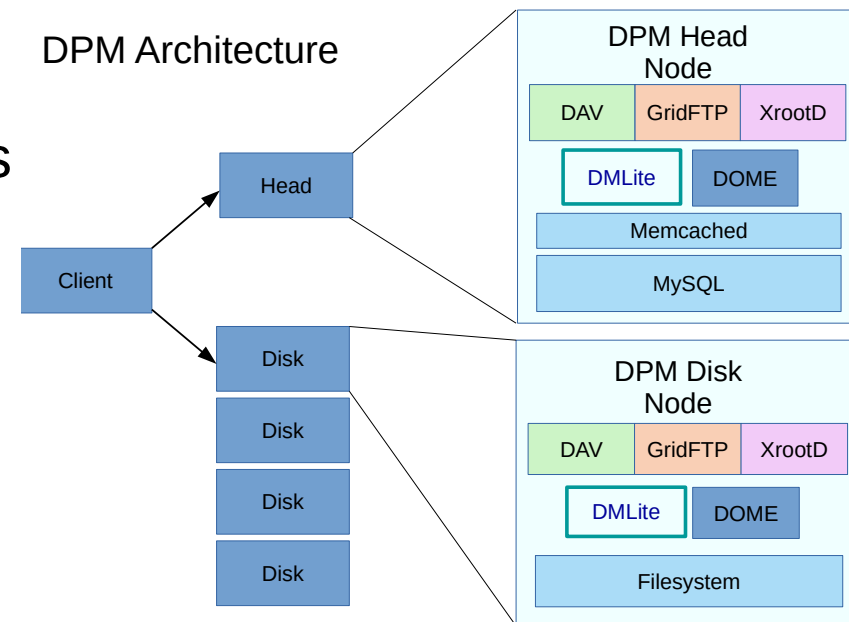
DPM	40
dCache	34
StoRM	18
Other	20
Σ	112

BDII registered

- 120 endpoints
- 100PB

DPM overview

- Popular grid storage system
 - multi-protocol
 - SRM, GridFTP, XRootD WebDAV
 - multi-vo
 - X.509 / VOMS and token based authentication, ACL
 - scalable to medium size sites ~ 10PB
 - headnode – maintains metadata, namespace and ACL + redirect clients
 - disknodes – data storage
 - built on top of common software
 - XRootD, Apache, Globus, MySQL
 - supported by CERN IT
 - relatively easy to deploy (puppet)
 - integrated with WLCG workflows



DPM flavours

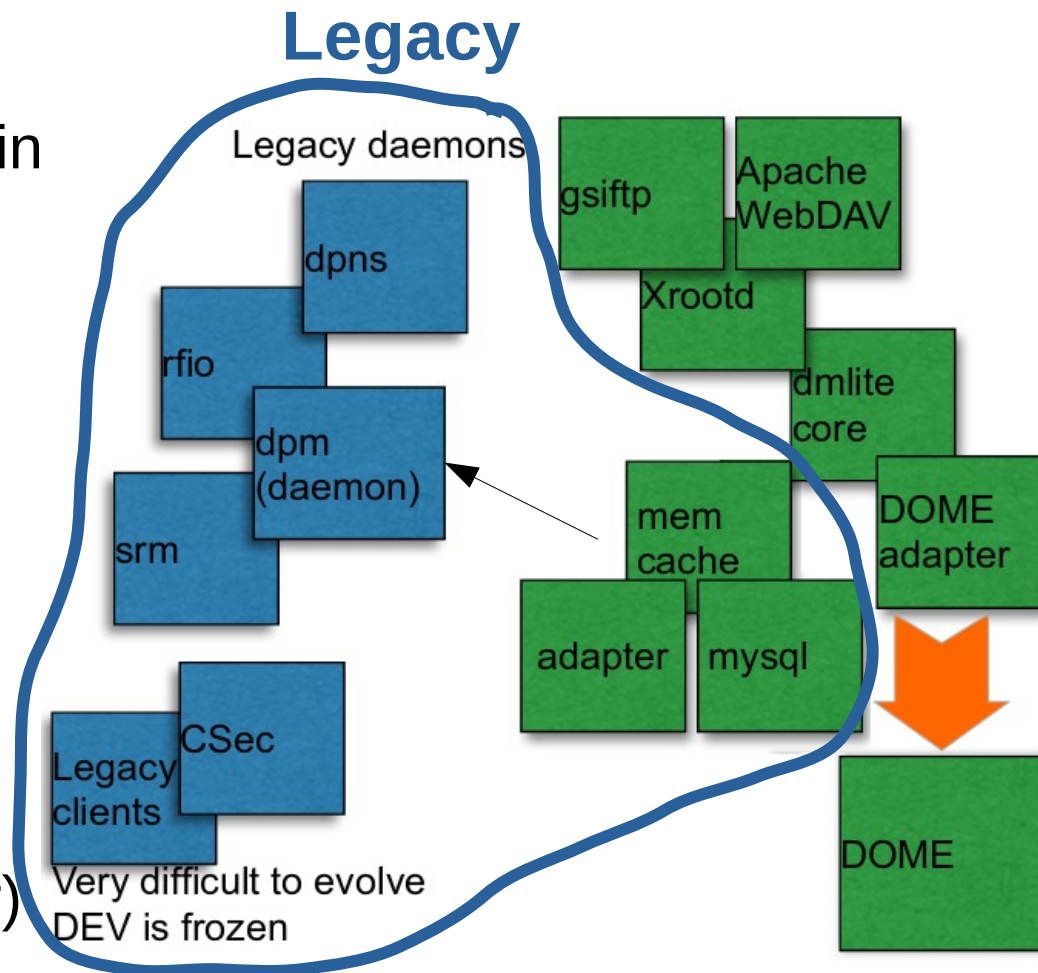
Legacy

VS.

DOME

DPM flavours

- Legacy DPM
 - old design and codebase – development frozen
 - dpm daemon does coordination
 - every process loads dmlite libraries
 - per process db connections
 - performance bottlenecks (caching)
 - sub-optimal resource usage
- DOME (Disk Operations Management Engine)
 - available since DPM 1.9
 - huge changes in DPM 1.10
 - apache fastcgi → XrdHTTP
 - DOME does coordination
 - dpm daemon for SRM (adapter)



Legacy DPM support
ends with 1st June 2019

DPM legacy EOL

- With a few exceptions all DPM sites still run legacy mode
 - less than 10 sites (including non-ATLAS sites) moved to DOME
 - you'll not get DOME without changing configuration
 - normal RPM upgrade just bring new DPM version
 - 7 ATLAS sites still use ancient DPM 1.8
 - 14 ATLAS sites use more than year old DPM 1.9
 - 19 ATLAS sites use DPM 1.10 (or newer)
 - only ½ bigger ATLAS DPM sites with 2PB+ use this version
 - DOME available since 1.9, “usable” since 1.10, good with 1.12(?)
- Legacy DPM vs. DOME + legacy vs. DOME only DPM
 - Legacy was all about SRM (GridFTP) + some XRootD & WebDAV
 - DOME is about GridFTP+XRootD+WebDAV
 - to support SRM have to run both DOME & legacy
 - WLCG experiments don't need SRM
 - getting rid of SRM is not effortless

WLCG DPM Upgrade TF

- Scope of [DPM upgrade task force](#) activities
 - coordinate upgrade of the DPM sites
 - DPM 1.12 should be released soon
 - upgrade to the DOME mode (legacy EOL in summer 2019)
 - enable WLCG [Storage Resource Reporting](#) (SSR)
 - provide guidance and support for upgrade and reconfiguration
 - validate SSR reporting and later integrate to CRIC and WLCG Storage Space Accounting
- Upgrade plan
 - phase 1 (end of 2018)
 - get experience DPM DOME experience with small number of sites
 - phase 2
 - by summer 2019 80% of DPM storage (in terms of capacity)
 - by the end of 2019 80% of DPM sites

DPM upgrade issues

- puppet (re)configuration is simple
 - still a lot of variations in configuration
 - issues caused by migration to CentOS7
 - database on dedicated machine
 - not all available configuration option optimal
 - necessary to follow [DPM tuning hints](#)
 - increase limit on number of open file handles
 - local DNS caching service (nscd or local DNS)
- GridFTP redirection (GridFTP still used exclusively for TPC)
 - necessary for pure GridFTP without SRM
 - in production on two sites (issues hopefully fixed in 1.12)
- Spacetokens vs. quotatoken
 - issues with (slowly) diverging counters – enforcing wrong quota
 - physical disk space issues (fix in 1.12)

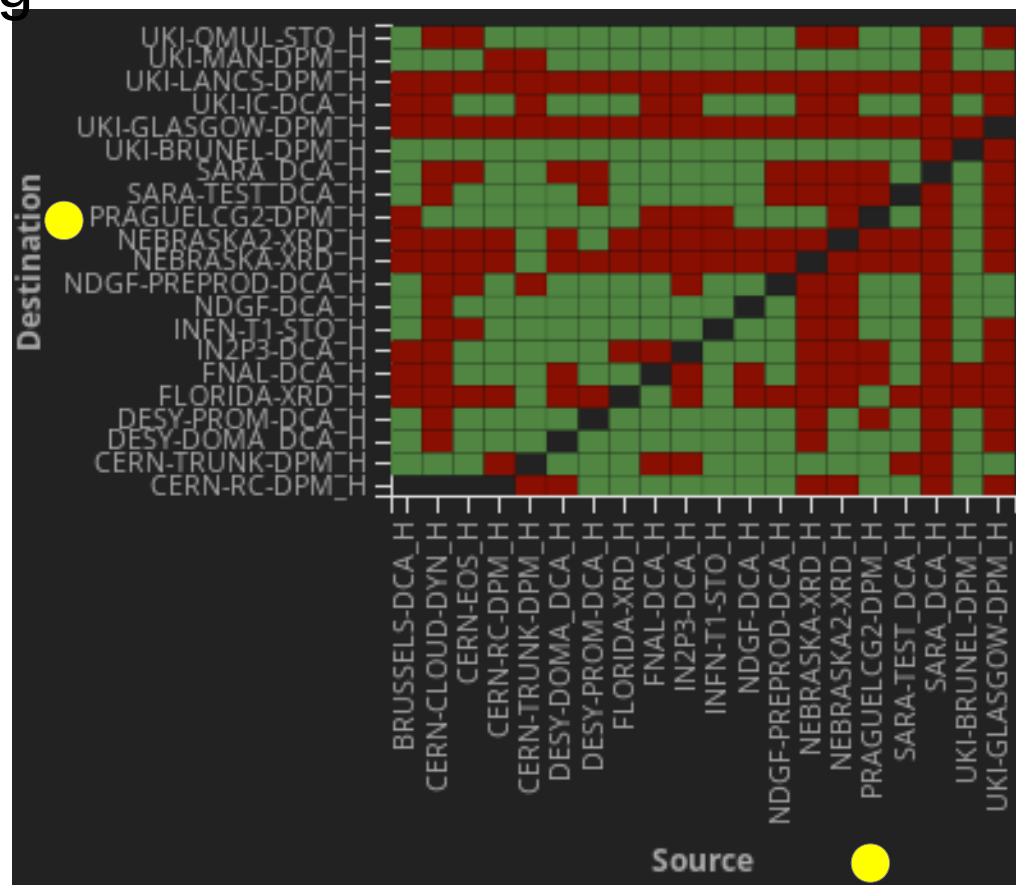
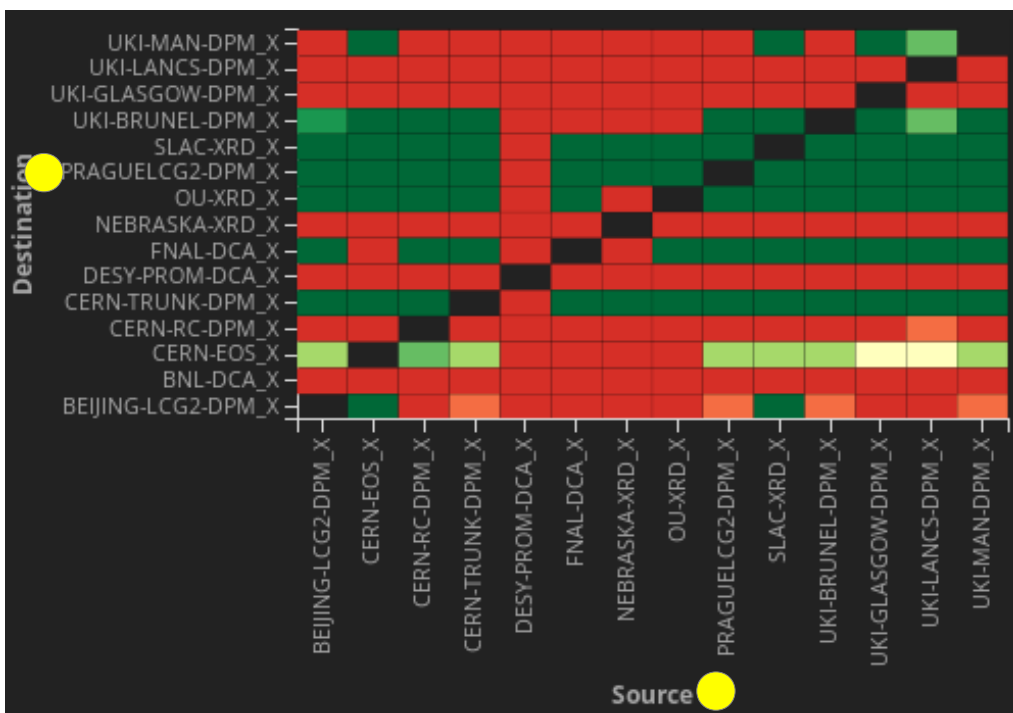
using legacy **SRM**
together **with DOME**
WebDAV, GridFTP,
XrootD within one
VO should be
avoided

Recommended ATLAS DOME cfg

- DPM DOME dmlite version at least 1.12 + XRootD 4.9
 - Enable gridftp redirection (puppet gridftp_redirect)
 - Enable XRootD checksum (puppet configure_dpm_xrootd_checksum)
 - Optionally enable TPC (puppet configure_dpm_xrootd_delegation)
- **AGIS SE configuration**
 - GridFTP for wan_read, wan_write, third_party_copy
 - XRootD for lan_read, lan_write
 - can take up to ~ 2 hours till Rucio pick new configuration
 - rucio-admin rse info RSE_NAME
 - https://rucio-ui.cern.ch/rse?rse=RSE_NAME
- **AGIS Panda queues**
 - rucio mover
- SAM tests for non-SRM sites
 - update test templates to consider GridFTP protocol

DPM TPC

- Pretty good support for non-SRM TPC ([WLCG ThirdPartyCopy TF](#))
 - only supported with DPM DOME configuration
 - X.509 proxy delegation (XRootD 4.9, WebDAV with gridsite)
 - token based delegation with macaroons
- [Issues with Apache CRL reloading](#)
- Testing – [smoke-test](#), [matrix](#)



Summary

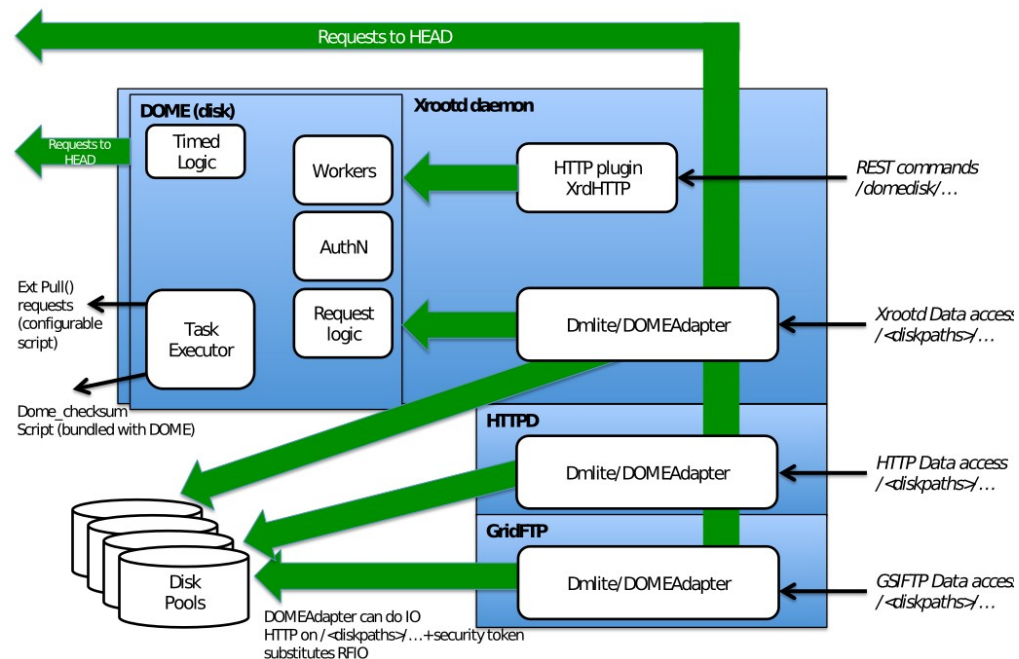
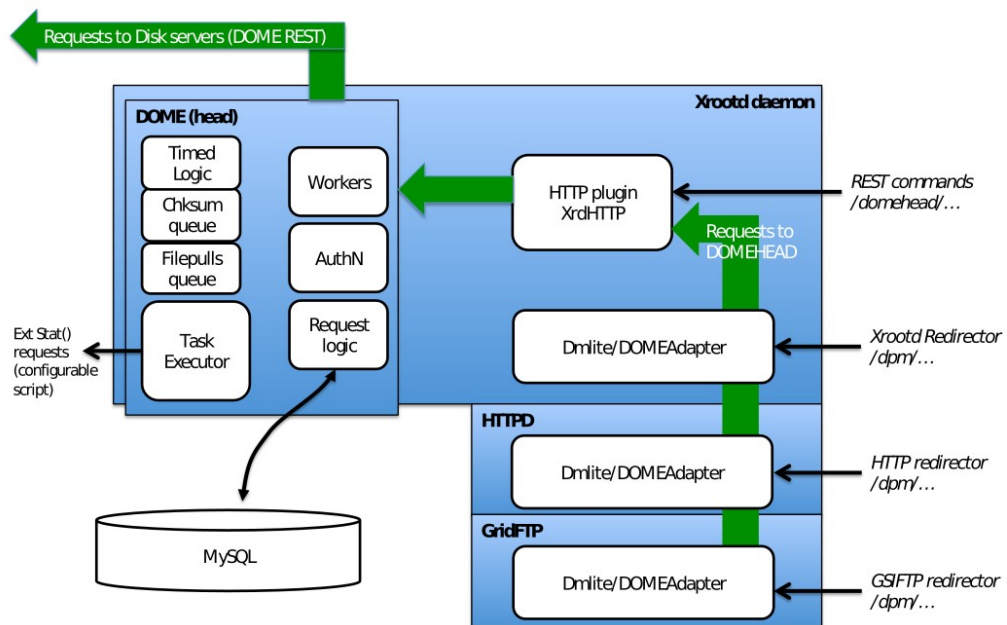
- Two flavors of DPM – legacy and DOME
- Active support of legacy DPM ends by June 1st 2019
 - upgrade to DOME DPM 1.12 with XRootD 4.9 and dpm-dsi 1.9.15
 - latest recommendation by [WLCG DPM Upgrade Task Force](#)
- New features only in DPM DOME
 - Third Party Copy for non-GridFTP protocols
 - WLCG Storage Resource Reporting
 - improved (metadata) performance compared to legacy DPM
 - at least 2x for pure GridFTP
 - order of magnitude for XRootD and WebDAV
- Missing support for SRM protocol in DOME DPM
 - WLCG experiments can be configured not to use SRM
 - SRM still available through DOME to legacy DPM adapter
 - SRM and DOME transfers in one VO/spacetime not recommended

BACKUP

DPM sites > 2PB

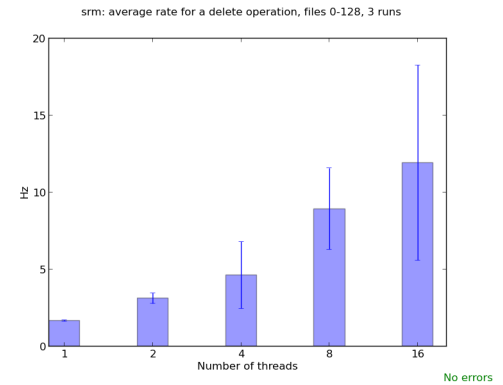
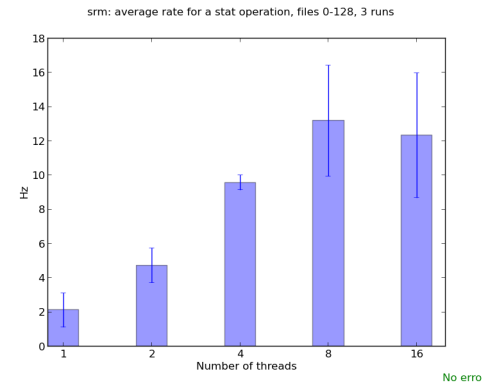
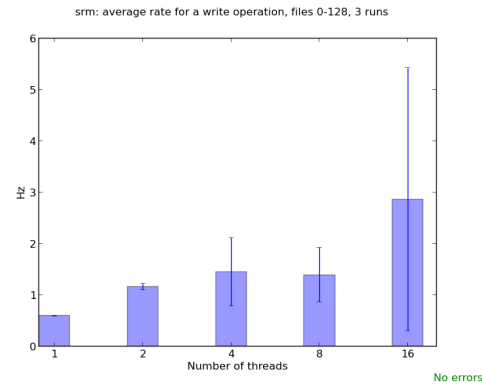
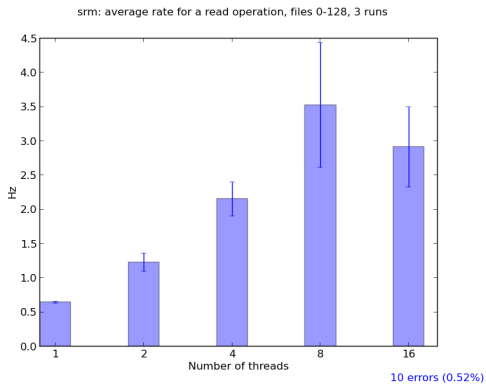
Site	Headnode	Size	DPM Version	XRootD protocol	GridFTP
TOKYO-LCG2	lcg-se01.icepp.jp	10559536	DPM/1.9.0-1	0x10030000	GridFTP Server 12.5
ICN-UNAM	tlapiacalli.nucleares.unam.mx	5248671	N/A	N/A	GridFTP Server 11.1
GRIF	node12.datagrid.cea.fr	4554739	DPM/1.10.0-1	0x10030000	GridFTP Server 13.9
UKI-NORTHGRID-MA N-HEP	bohr3226.tier2.hep.manchester.ac.uk	4537796	N/A	0x30000	GridFTP Server 9.1
praguelcg2	goliath100.farm.particle.cz	4456037	DPM/1.10.0-1	0x40000	GridFTP Server 13.9
UKI-SCOTGRID-GLA SGOW	svr018.gla.scotgrid.ac.uk	3816622	DPM/1.8.10-1	0x10030000	GridFTP Server 12.4
Taiwan-LCG2	f-dpm001.grid.sinica.edu.tw	3269468	DPM/1.8.11-1	N/A	N/A
TR-10-ULAKBIM	torik1.ulakbim.gov.tr	3161019	DPM/1.10.0-1	0x10030000	GridFTP Server 11.3
INDIACMS-TIFR	se01.indiacms.res.in	3107872	DPM/1.9.0-1	0x10030000	N/A
UKI-NORTHGRID-LA NCS-HEP	fal-pygrid-30.lancs.ac.uk	3074570	DPM/1.10.0-1	0x10030000	GridFTP Server 13.9
IN2P3-CPPM	marsedpm.in2p3.fr	2642392	DPM/1.9.0-1	0x10030000	GridFTP Server 12.4
INFN-NAPOLI-ATLAS	t2-dpm-01.na.infn.it	2399002	DPM/1.10.0-1	0x10030000	GridFTP Server 13.9
RO-07-NIPNE	tbit00.nipne.ro	2367783	N/A	0x30000	GridFTP Server 7.26
NIKHEF-ELPROD	tbn18.nikhef.nl	2353756	DPM/1.9.0-1	0x10030000	N/A
IN2P3-LAPP	lapp-se01.in2p3.fr	2187150	DPM/1.9.0-1	0x10030000	GridFTP Server 12.5
GRIF	lpnse1.in2p3.fr	2113058	DPM/1.10.0-1	0x10030000	N/A
INFN-FRASCATI	atlasse.lnf.infn.it	2111236	N/A	0x10030000	GridFTP Server 11.8
IN2P3-IRES	sbgse1.in2p3.fr	2032105	DPM/1.10.0-1	0x10030000	GridFTP Server 13.9

DOME head / disk node architecture

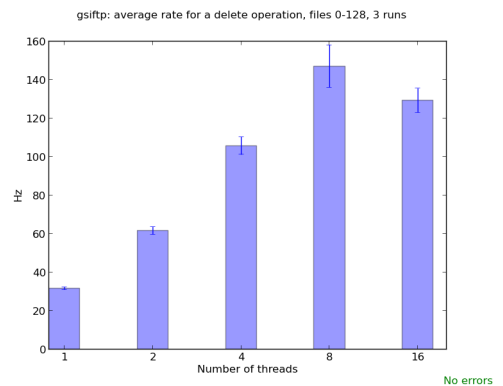
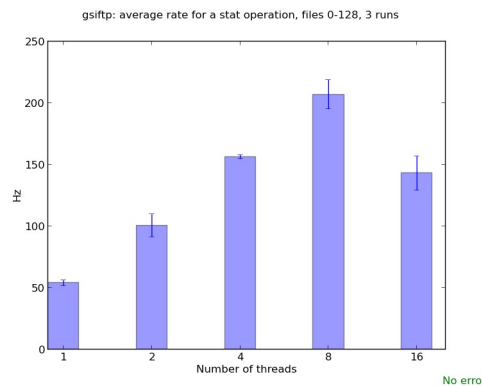
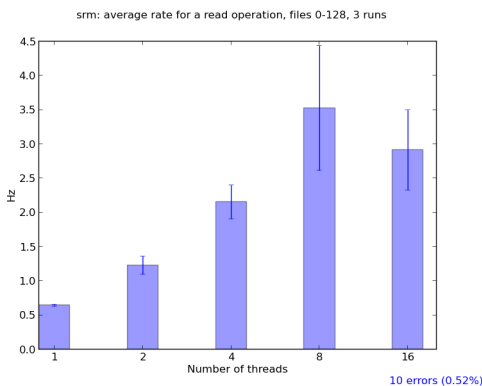
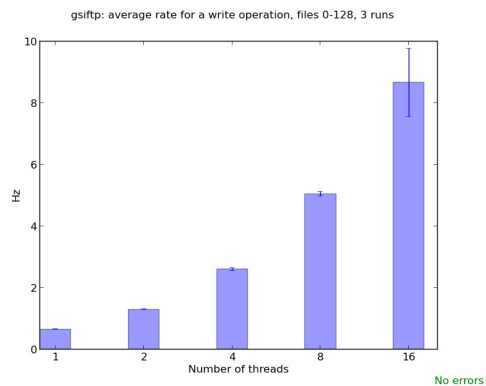


SRM vs. GridFTP performance

- DOME DPM 1.11.1, XRootD 4.9, dpm-dsi 1.9.15 with redirection
- SRM performance – write, read, stat, delete



- GridFTP performance – write, read, stat, delete



Issues – failing SAM SE tests

- 2018

- January 1st – UPS / power failure
- January 26-30 – performance issues (non-local DNS queries)
- May 11, 21 – troubles with one diskserver
- June 28 – DPM 1.10.3 upgrade
- July 23 – DPM reconfigured to DOME mode
- July 29-31 – enabled/disabled GridFTP redirection
- August 11 – spacetoken synchronization issues
- December – troubleshooting core switch / network problems

- 2019

- January 30 – new core switch
- February 10 – DPM DOME upgrade to 1.11.1
- February 17 – enabled GridFTP redirection

Test history golias100.farm.particle.cz using ATLAS_CRITICAL

10392 hours from 2018-01-01 01:00 to 2019-03-10 01:00

