

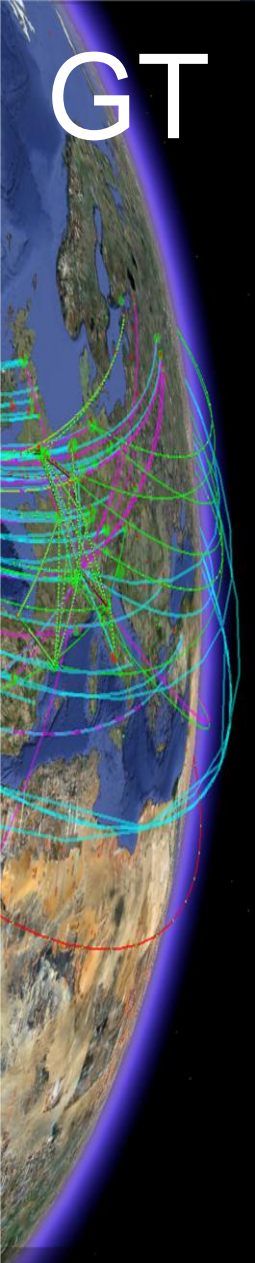
# DPM Collaboration Motivation and proposal

Oliver Keeble

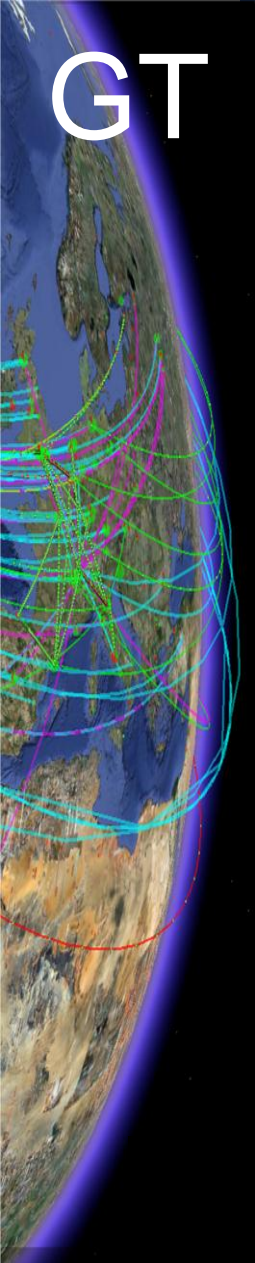
CERN

On behalf of DPM/LFC team

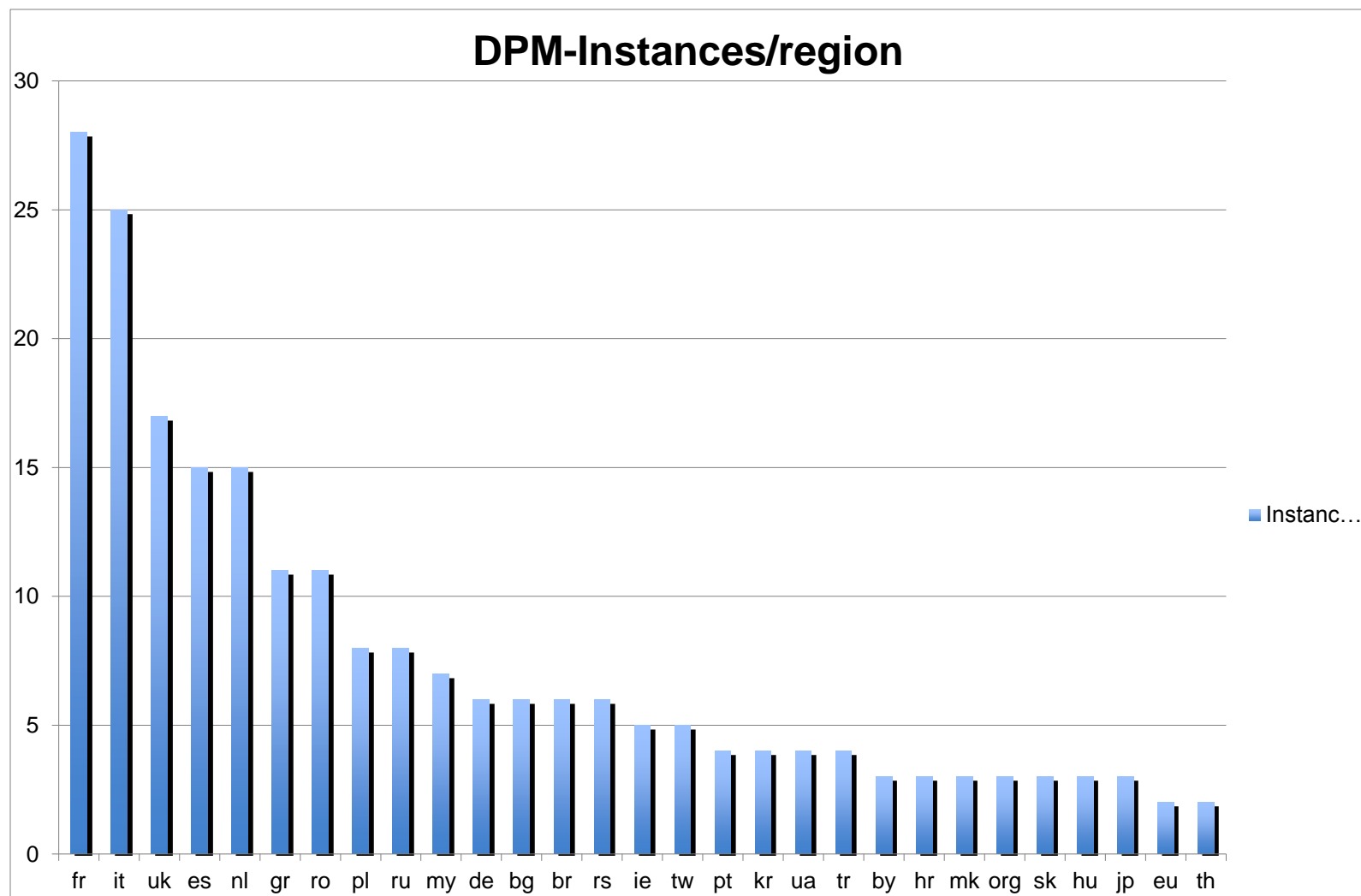
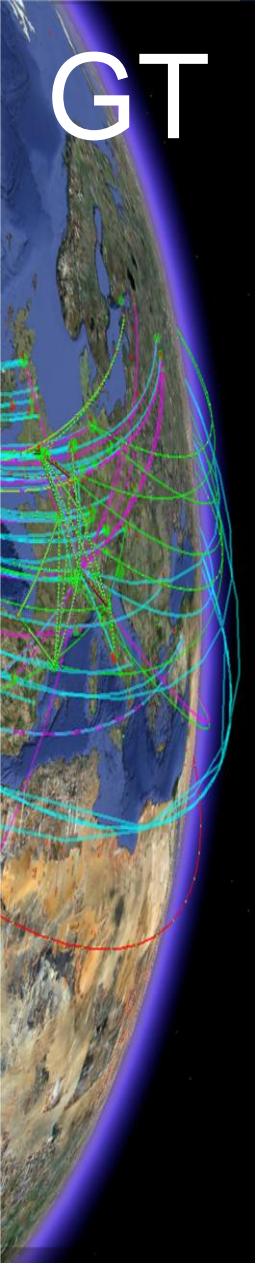
1. Context: middleware funding
2. Why DPM matters
3. To whom it matters
4. What DPM represents today
5. What to do about point 1 above...



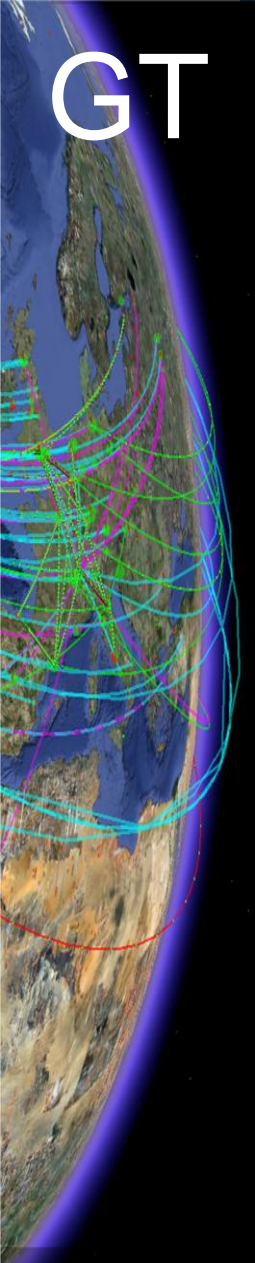
1. Context: middleware funding
2. Why DPM matters
3. To whom it matters
4. What DPM represents today
5. What to do about point 1 above...



- Middleware has enjoyed European funding for over 10 years
- This ends with the EMI project
- DPM/LFC is a project heavily reliant on this source of support
- At the same time, the project has never been in better shape...
- We propose a collaboration of stakeholders to drive the project forward



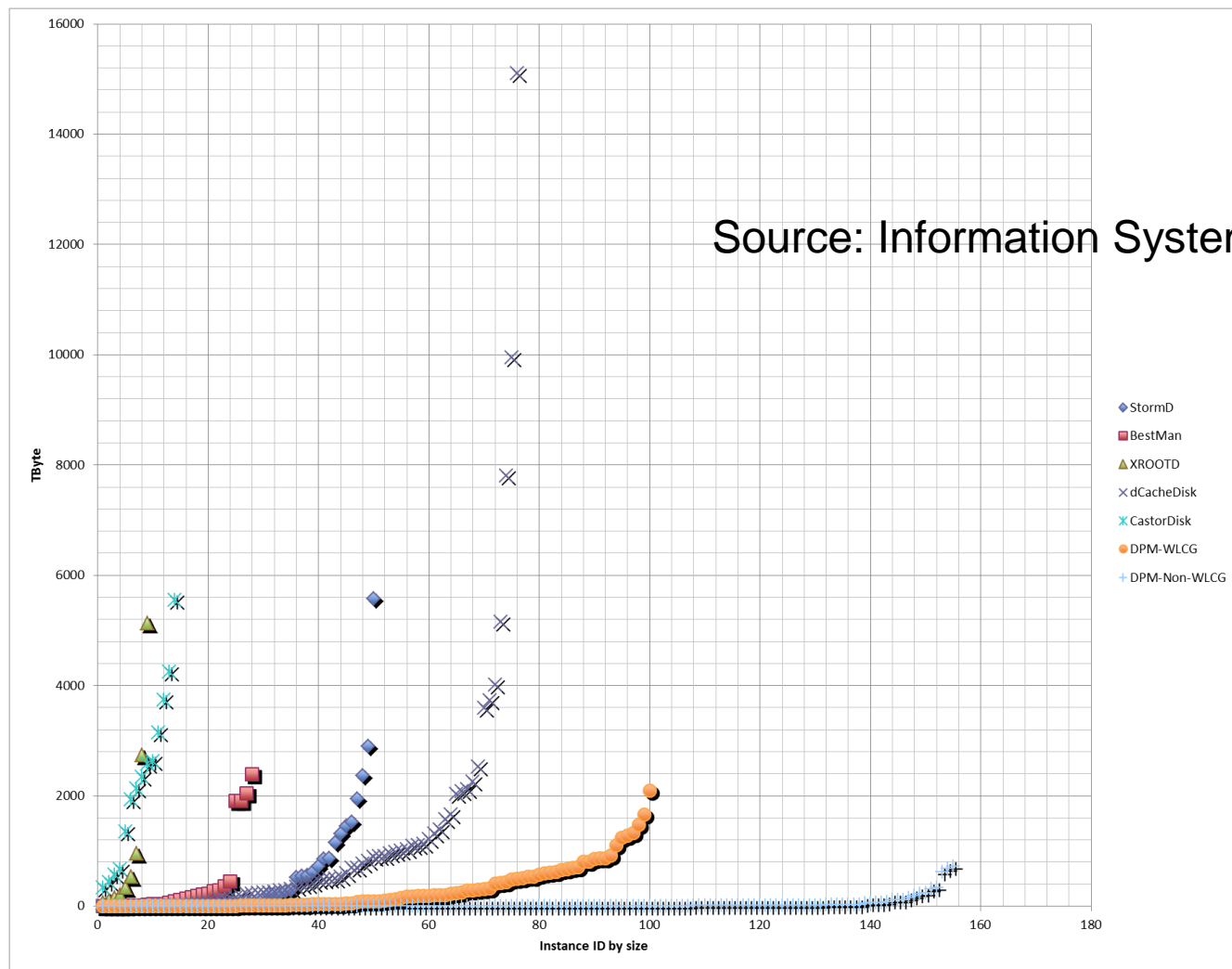
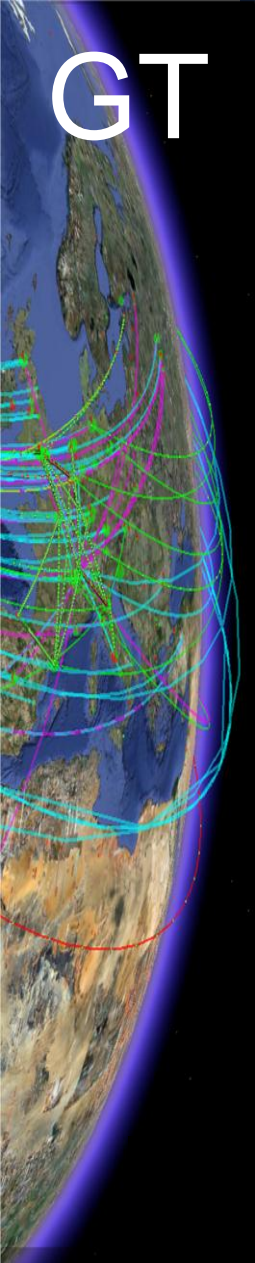
1. Context: middleware funding
2. Why DPM matters
3. To whom it matters
4. What DPM represents today
5. What to do about point 1 above...





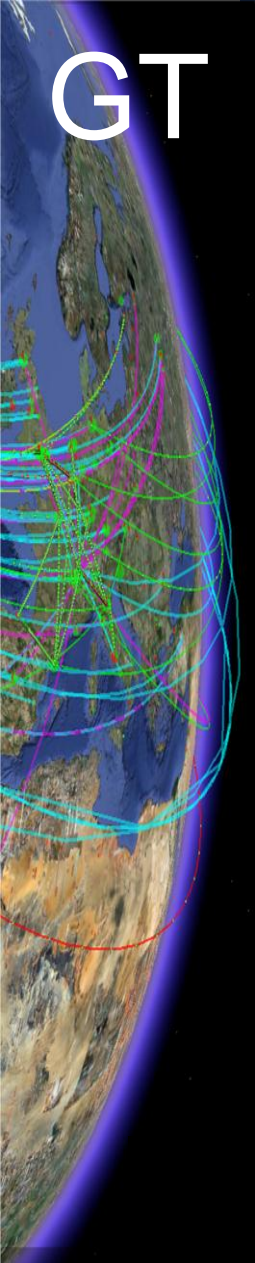
- 36PB
  - 10 sites with  $> 1$ PB
- Over 200 sites in 50 regions
  - ~90 WLCG sites
- Over 300 VOs have access to DPMs

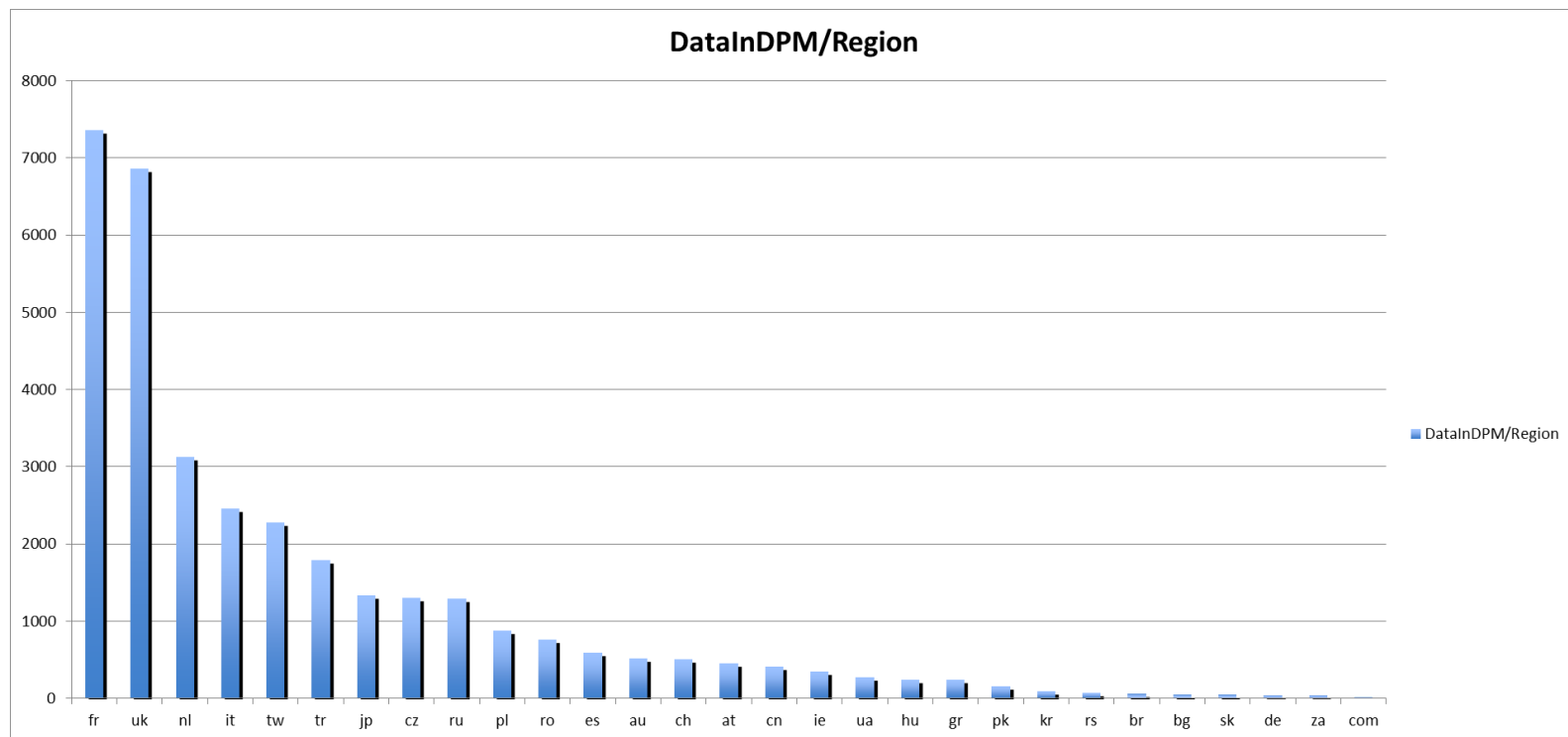




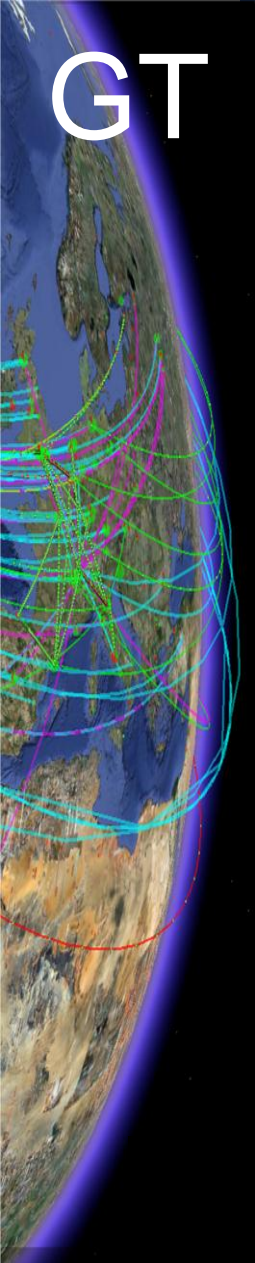


1. Context: middleware funding
2. Why DPM matters
3. To whom it matters
4. What DPM represents today
5. What to do about point 1 above...





1. Context: middleware funding
2. Why DPM matters
3. To whom it matters
4. What DPM represents today
5. What to do about point 1 above...



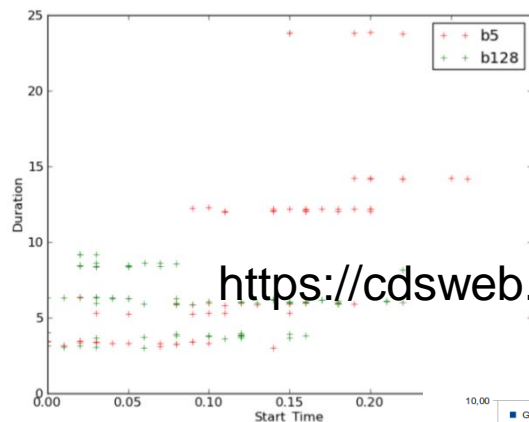
- Some DPM myths:
- DPM is only for small sites
  - No, there are several multiPB DPMs, incl. a T1
- DPM doesn't do xrootd
  - A new plugin has been created
- DPM doesn't represent much storage
  - DPM represents an respectable fraction of grid storage
- DPM hasn't seen investment
  - DPM has seen significant investment thanks to the EMI project
    - Not everything has been released yet

- HTTP/WebDAV: released with EMI2
- NFSv4.1: released r/o with EMI2
- Refactoring (dmlite) to make future additions much easier and allow performance improvements
- Perfsuite: internal performance testing suite, completed recently
- Nagios probes: released during EMI1
- EPEL compliance – you can find it there now
- + more!

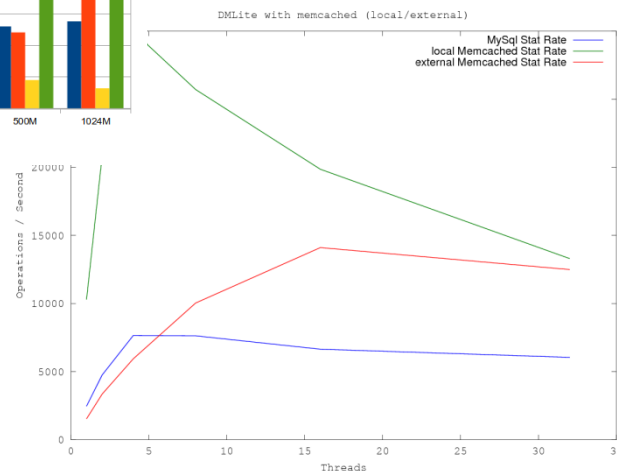
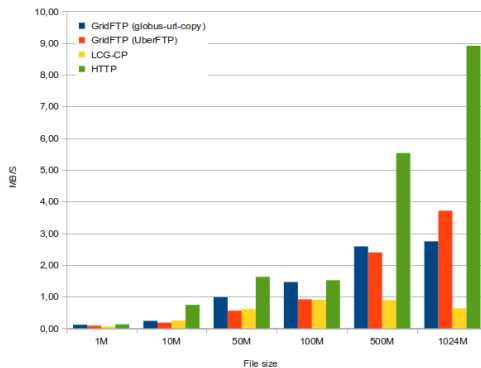
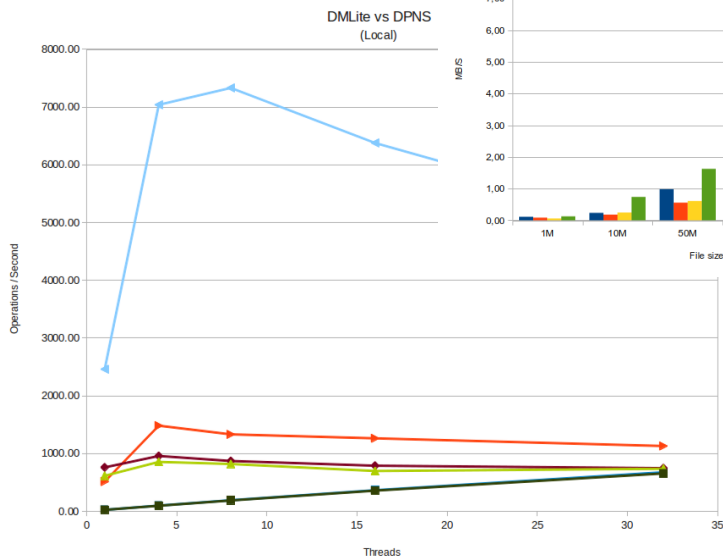
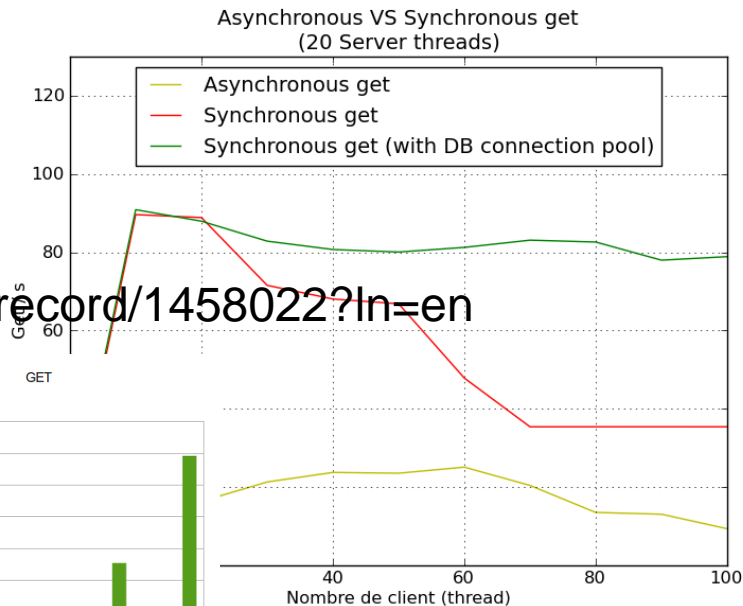


# GT

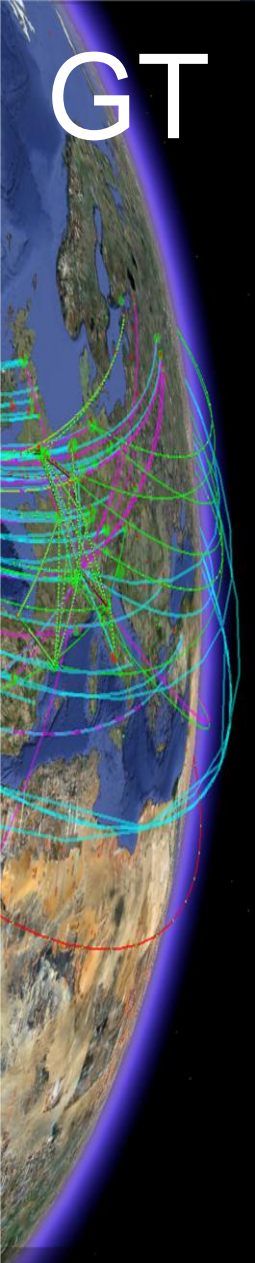
# Performance, performance, performance



<https://cdsweb.cern.ch/record/1458022?ln=en>



- LFC is a file catalogue derived from DPM's nameserver
- It is currently used by Atlas and LHCb
  - Both have declared their intention to move away from LFC
  - Timescales are not clear though
- It is also used by other VOs...



- HTTP Federation: ongoing work with dCache
- Memcache: scheduled for Q3 release
  - Fast in-memory namespace cache
- Improved xrootd support
  - In production testing
- Different backends have been prototyped:
  - HDFS
  - S3
- Disk/tape decoupling offers new deployment scenarios for DPM as a radically simpler solution for larger sites



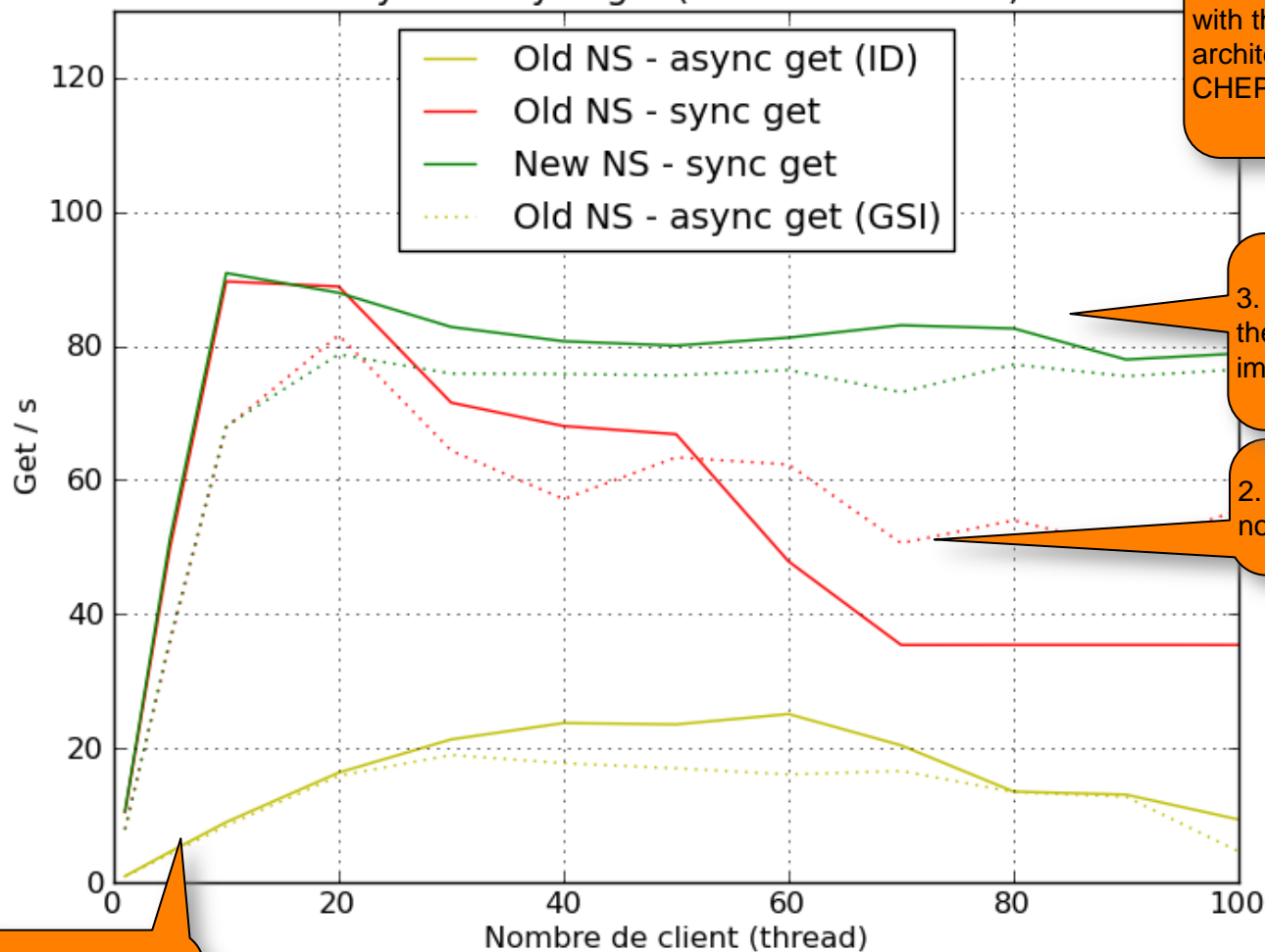
GOAL: functional and performance improvements.

- Performance is of primary importance
  - Profit from changes in DPM
- Allow xrootd clients to access and write files to DPM
  - Use X.509 authentication, with VOMS extensions
    - DPM does the usual authorization
  - Optionally allow the ALICE “tkauthz” mechanism
- Mixed multi-VO setups
  - A single DPM can serve more than one VO
- Support for per-VO xrootd federations
  - Following the evolution of the ATLAS LFC-based name translations
- Currently under beta testing by Atlas





Async VS Sync get (20 Server threads)



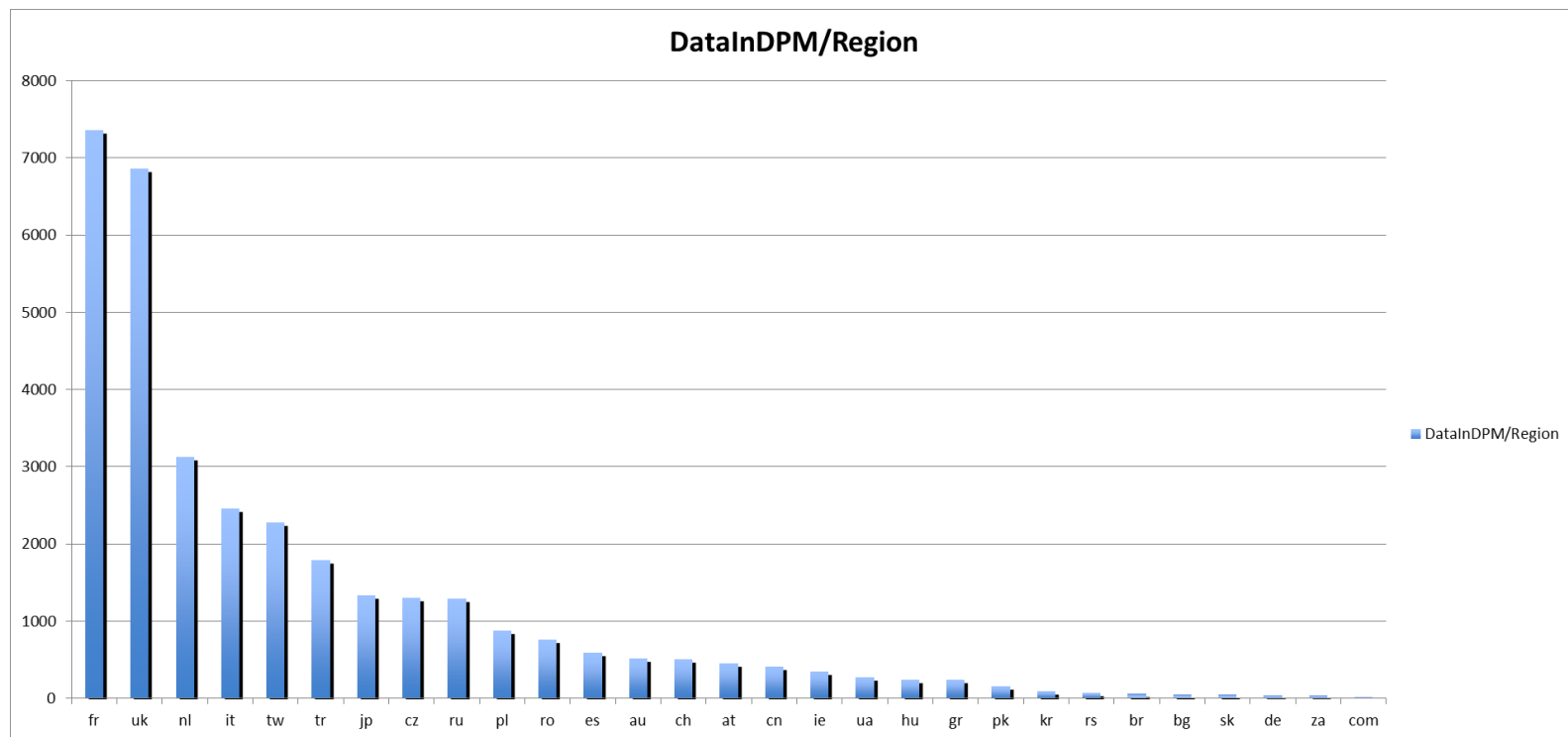
4. Will be higher with the newer architecture (see CHEP contribs)

3. Will be here with the NS improvements

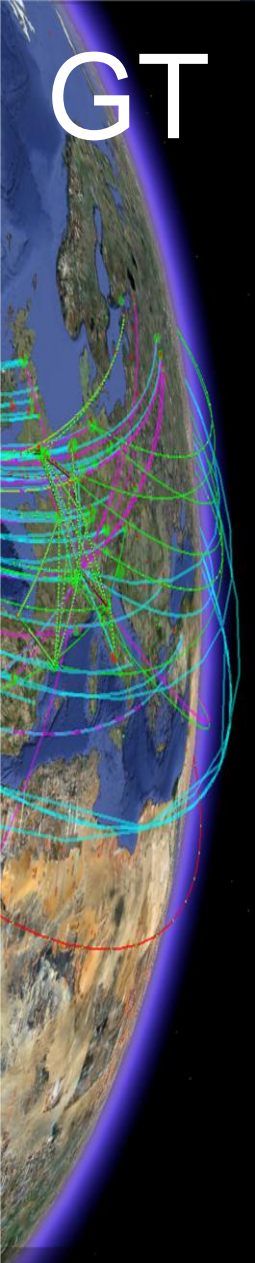
2. Aiming here now (red line)

1. We were here (~1-20 req/sec)





1. Context: middleware funding
2. Why DPM matters
3. To whom it matters
4. What DPM represents today
5. What to do about point 1 above...



- **Those to whom DPM matters are invited to express their interest in a collaboration**
- **Who**
  - All stakeholders, sites, projects, NGIs...
- **Why (ie the benefits)**
  - A sustainable vehicle for evolving the software on which they depend
    - Ensure continued support, updates and development
  - Protection of investment. Ensure that existing DPM installations remain well maintained and able to evolve with the hardware and OS.
  - Influence according to contribution
    - Credit system?
    - Nature of collaboration to be decided by involved parties
    - Potentially very important for larger sites considering investment in DPM
  - Opportunity to seek funding in the context of participation in an international project

- Development and research
  - thanks to recent improvements this has become easy
    - examples: memcached modules
    - interesting for experienced sites and universities
- Staged Rollout/stress testing
  - as important as development
- Operational tools
- Support
  - including improved documentation for sites
  - configuration management

- Current reality:
  - direct influence: WLCG, CERN, EMI
  - indirect influence: sites and volunteers
  - a lot depends on CERN IT
  - classical “producer – consumer” relation
    - not helpful to establish a collaborative approach
- We need something where effort is translated into influence



- The shape of the collaboration will be defined by the members in the end
- **straw-man:**
  - influence based on contribution ( by effort )
    - all forms have equal weight
  - main contributors (3-5) agree on a MoU
    - CERN will be one of these
    - without a MoU local needs “win” too often
      - see experience with staged rollout
      - volunteers are welcome and encouraged
        - low threshold for contributing
  - major directional changes agreed once or twice a year by a collaboration board meeting
    - Combined with community workshop

- DPM will be in the best state possible by the end of EMI
- If we can't find the resources:
  - Try to maintain DPM during the transition phase after the end of EMI
    - To allow for an orderly migration of the data
    - not clear how long this will take ( 1-2 years??)
    - decreasing effort needed ( see LCG-CE )
  - CERN's contribution to a collaboration could be used to maintain a static DPM for an extended period
    - Until natural obsolescence

- Collaboration
  - Markus.schulz@cern.ch
  - Oliver.keeble@cern.ch
- DPM technical
  - Ricardo.rocha@cern.ch
- CHEP 2012 papers:
  - Overview
    - <https://cdsweb.cern.ch/record/1458022?ln=en>
  - HTTP/dav
    - <https://cdsweb.cern.ch/record/1457962?ln=en>
  - Dynamic federations
    - <https://cdsweb.cern.ch/record/1460525?ln=en>