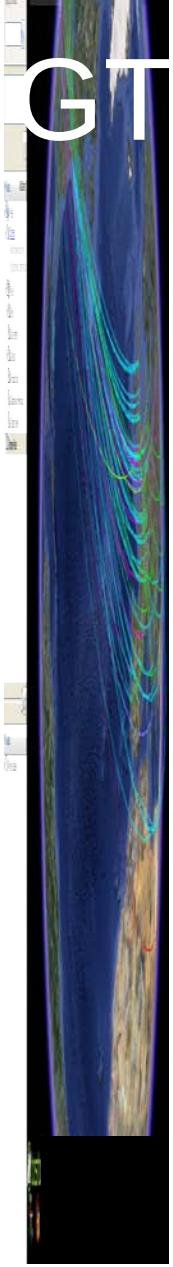
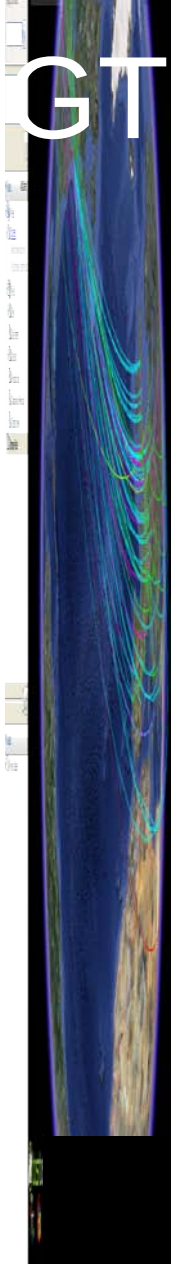


Services for Grid Data Management

Oliver Keeble & Jean-Philippe Baud

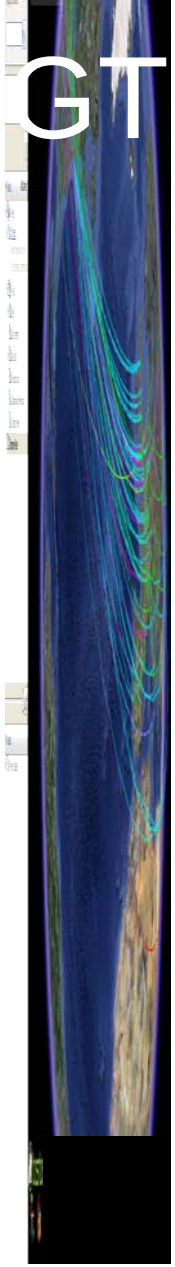
On behalf of IT-GT-DMS at CERN

CERN IT Department
CH-1211 Geneva 23
Switzerland
www.cern.ch/it



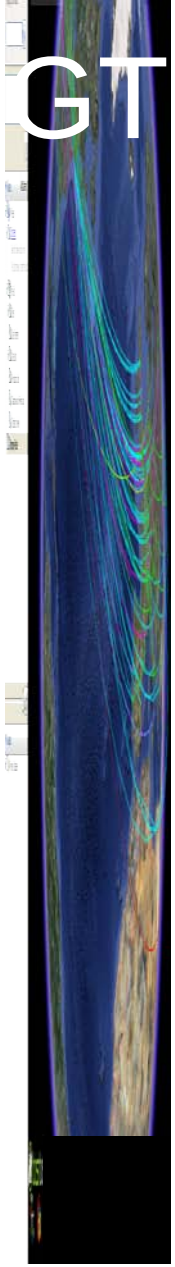
Overview

- Service overview
- Recent developments
- Status
- Short term outlook
- Longer term directions



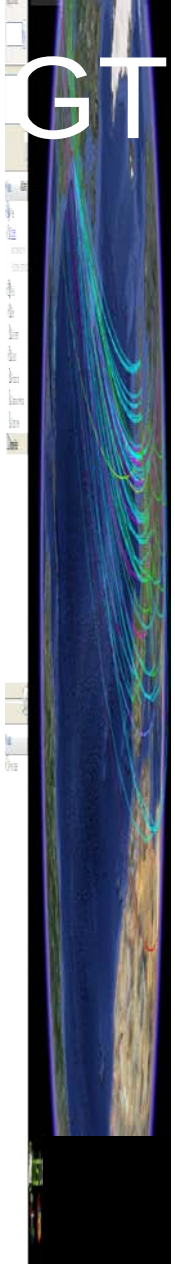
Data Management Services

- A coherent set of services for DM
- Storage : Disk Pool Manager (DPM)
- Catalogue: LCG File Catalogue (LFC)
- Transfer: File Transfer Service (FTS)
- Clients: gfal/lcg_util



FTS

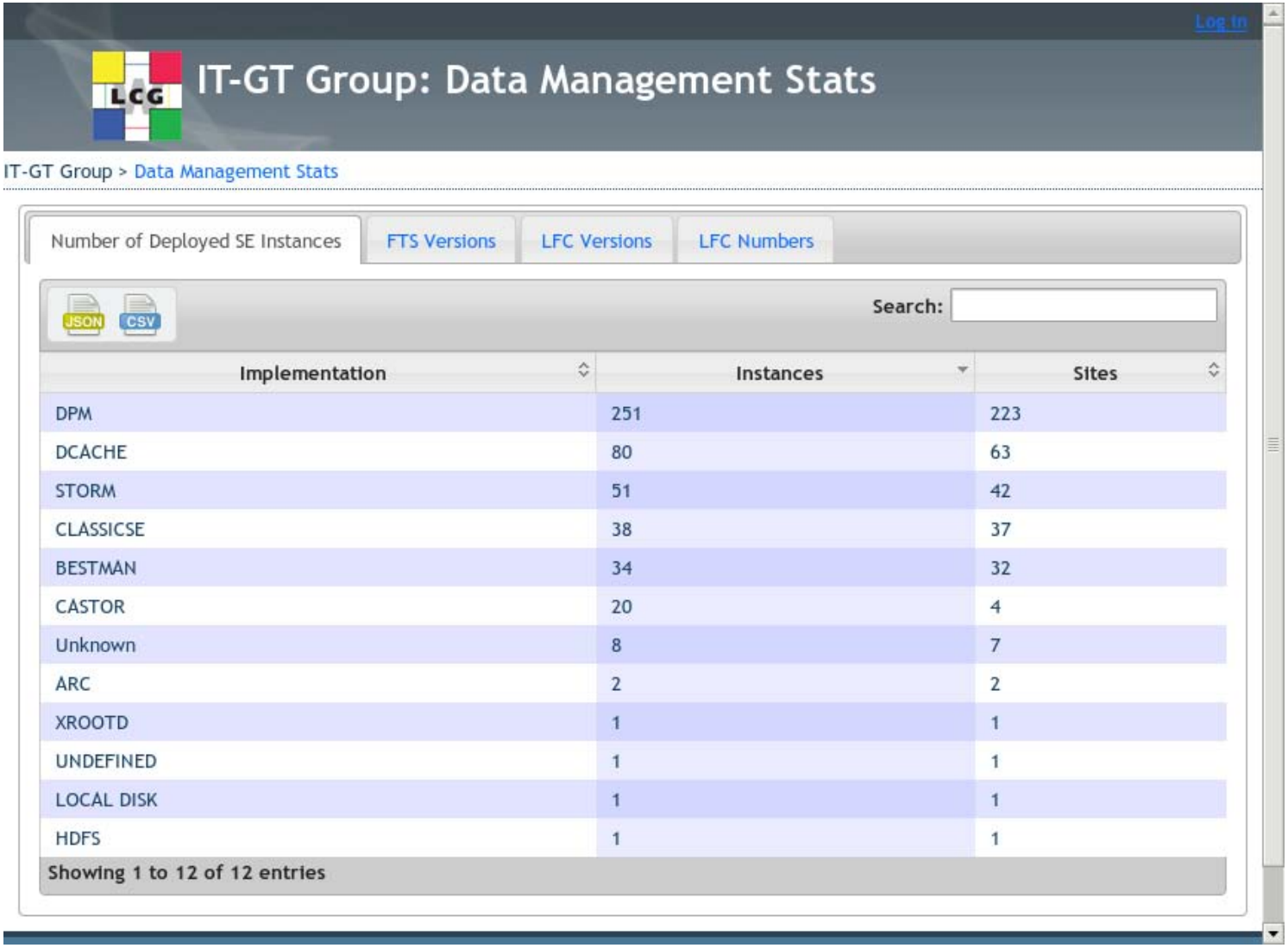
- Recent developments
- Available on SL5
  - SL4 security updates till Apr 2011
- 2.2.5 (in certification)
  - srmless endpoints
  - removal of voms server cert dependency
- 2.2.4
  - finalisation of checksum handling



DPM/LFC

- Recent developments
- 1.8.0 (in certification)
  - user banning with Argus support
  - internal 3rd party rfcp
- 1.7.4
  - configurable RFIO readahead buffer size on client
  - db cleanup
  - SURL bulk lookup
  - dual architecture installations 32/64
- 1.7.3
  - xrootd plugin update

DPM around the grid



IT-GT Group: Data Management Stats

Log in

IT-GT Group > Data Management Stats

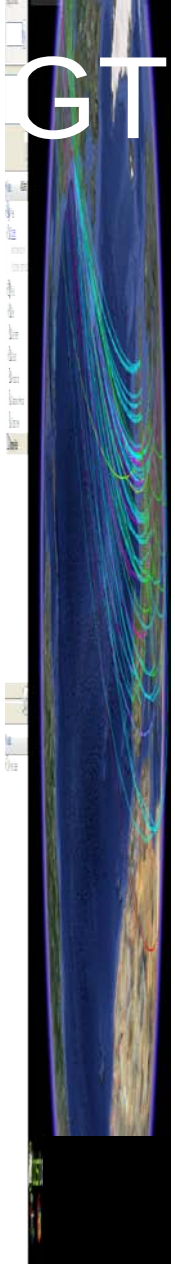
Number of Deployed SE Instances | FTS Versions | LFC Versions | LFC Numbers

JSON CSV

Search:

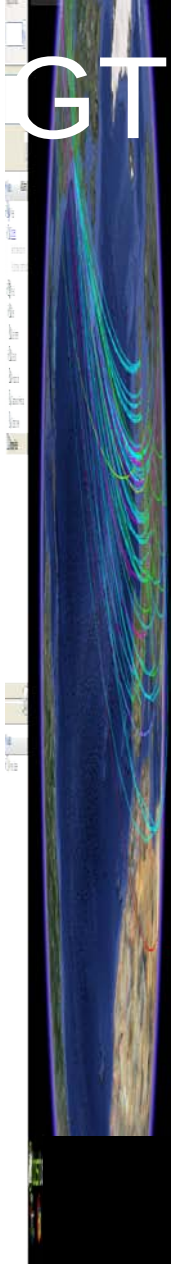
Implementation	Instances	Sites
DPM	251	223
DCACHE	80	63
STORM	51	42
CLASSICSE	38	37
BESTMAN	34	32
CASTOR	20	4
Unknown	8	7
ARC	2	2
XROOTD	1	1
UNDEFINED	1	1
LOCAL DISK	1	1
HDFS	1	1

Showing 1 to 12 of 12 entries



DPM & LFC deployment

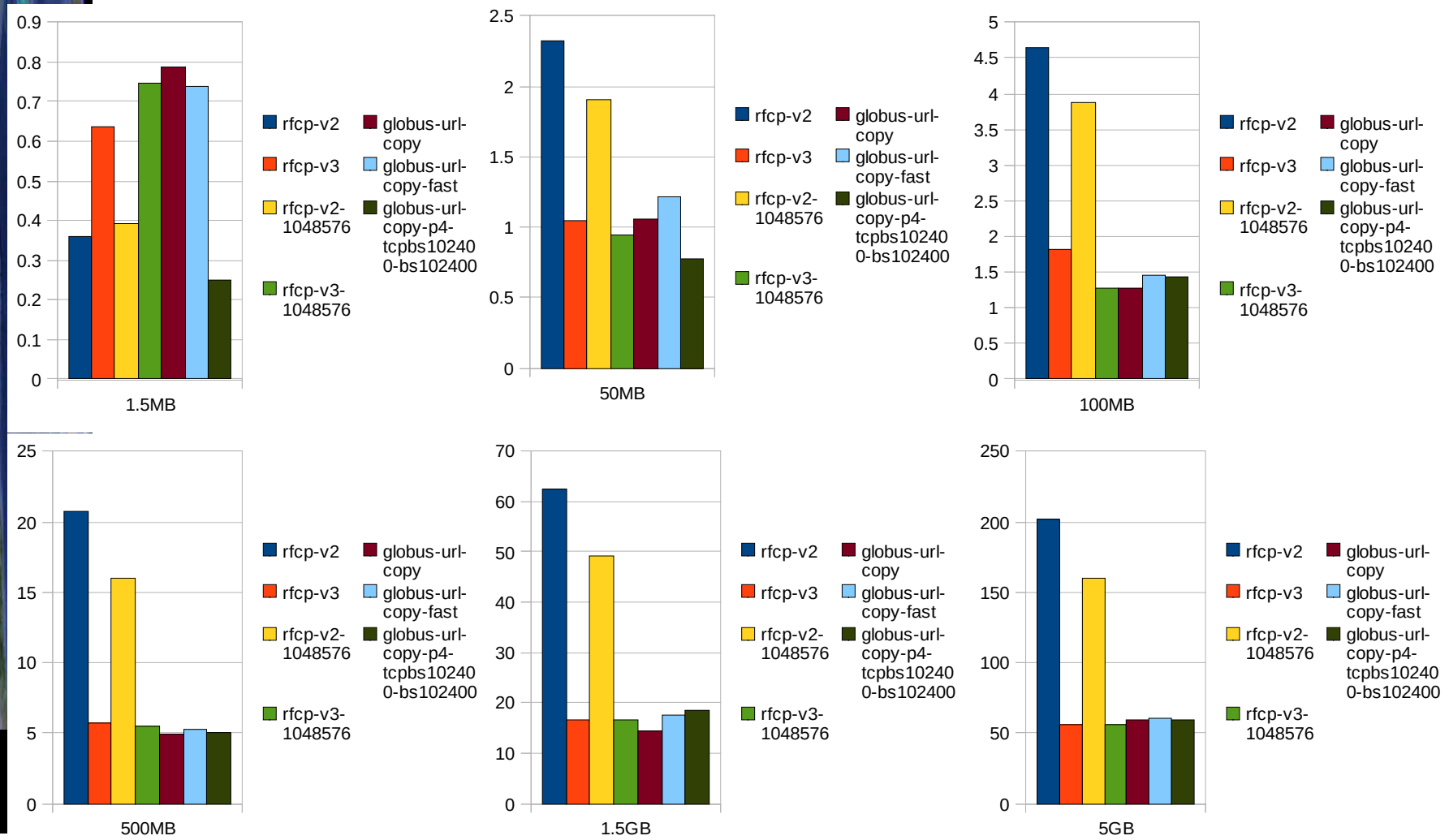
- DPM
  - 251 instances at 223 sites
  - Largest installations are over 1PB
  - In total manages over 15PB of storage
    - ~15% of online storage
- LFC
  - running at 56 sites
  - de-facto standard catalogue

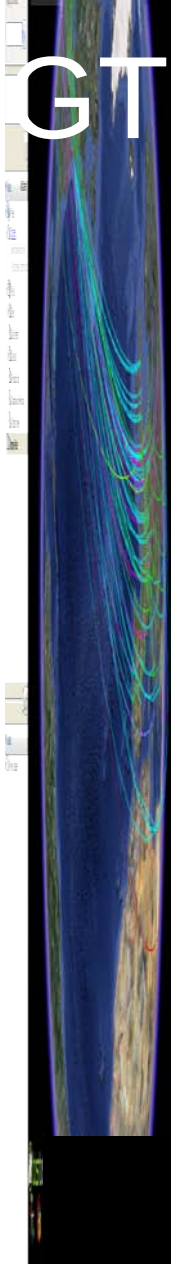


DPM/LFC future

- Adapt to new requirements
  - Performance & manageability
  - DPM: Instrumentation of disk servers
    - monitoring
    - filesystem selection
    - traffic management
  - DPM: Usage information
    - accounting
    - quotas
  - DPM: Replication
    - drain
    - hot files
  - DPM/LFC: access to logs
  - Catalogue Consistency

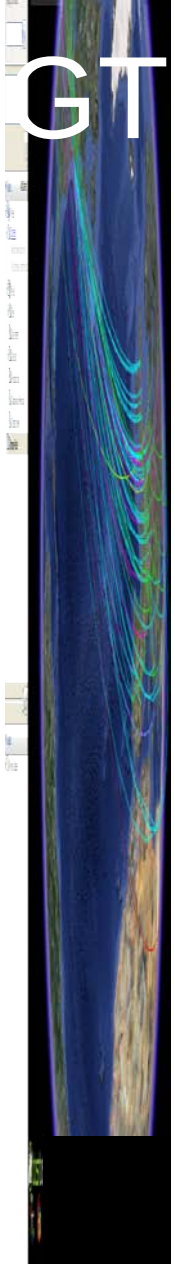
"3rd party" gridftp/rfio comparison





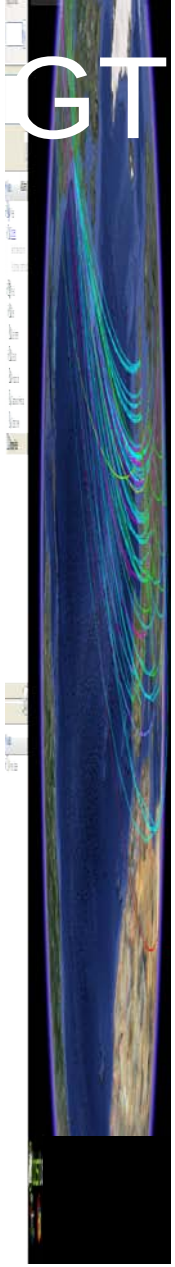
DPM future

- Standards
  - NFS4.1
    - access via file protocol with standard clients
    - performance to be established
      - reasons for optimism
  - SSL
    - httpg → https for SRM
    - sessions
  - http file access
    - allowing standard clients, eg curl
  - http access to LFC



Catalogue consistency

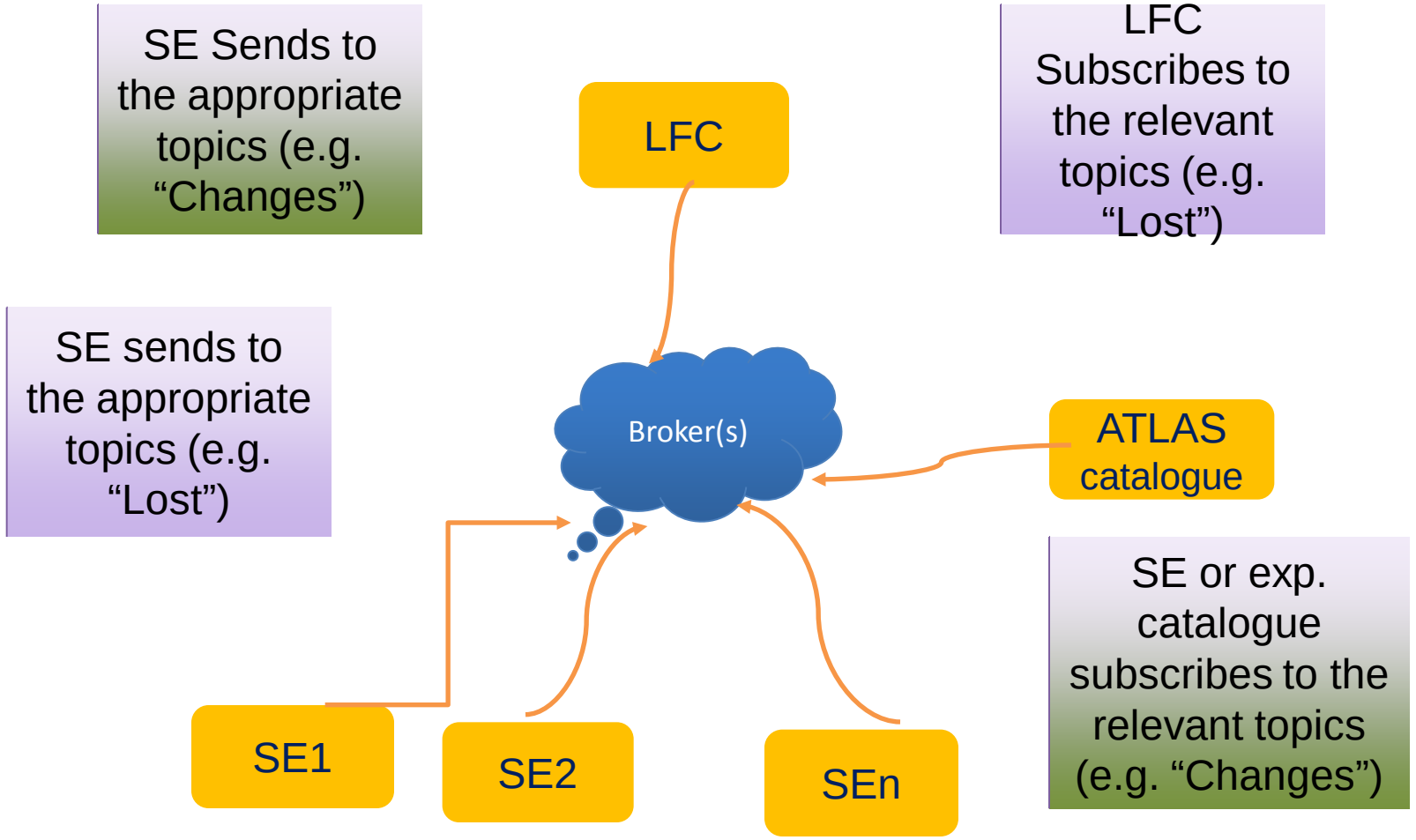
- The problem
  - If an SE looses a file, the LFC does not know
  - Absent files are expensive errors right now
  - A change in the permissions of a file in LFC is not automatically reflected by the peripheral catalogues
- Synchronisation now
  - Triggered by application software (copy/reg)
    - not atomic
  - Maintained by slow, periodic interventions

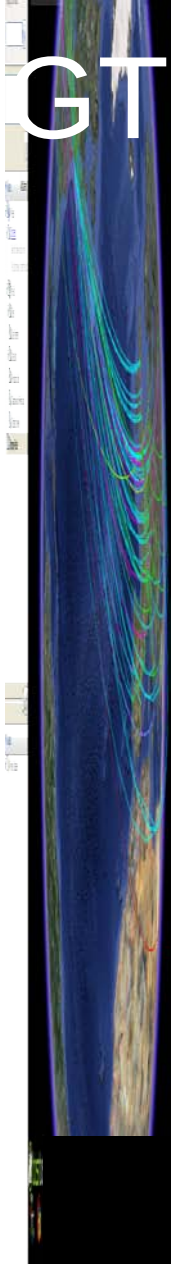


Catalogue consistency

- The solution
  - Status changes communicated over messaging
  - Eventual consistency
- Standardise message format (EMI)
  - Any SE or catalogue can become a participant in the consistency framework
- Demonstrator
  - using DPM/LFC
  - permissions changes and lost files
  - goal: demo December 2010

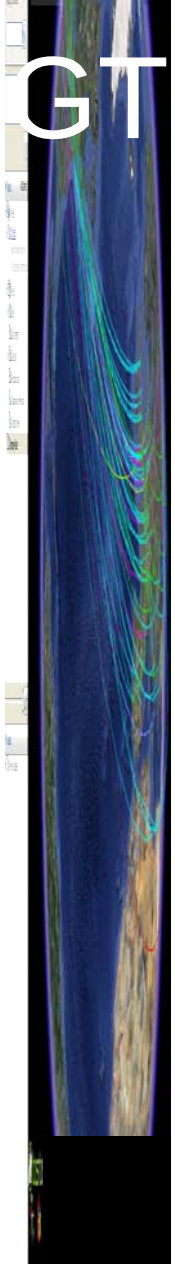
Catalogue consistency





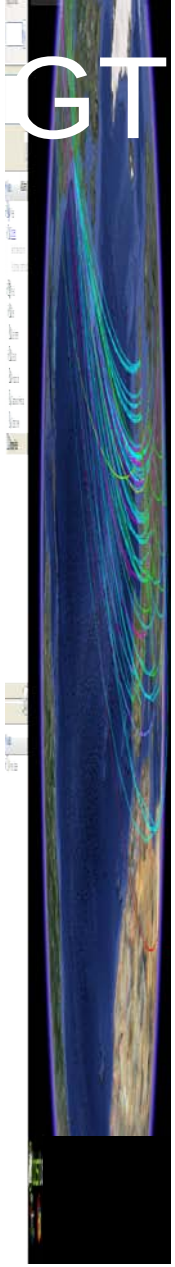
FTS - limitations

- While FTS operates as a workhorse for certain distribution scenarios, it has reached limitations in adapting to the evolution of infrastructure and workflows
  - Grid architecture
    - designed for the monarc model
    - now require something much less structured
      - T1/T1, T1/T2, T2/T2...
    - current FTS architecture will not scale, N squared problem
      - star channels are difficult to manage
      - move from channel paradigm to endpoint-centric
  - Configuration Model
    - Currently requires lots of configuration, with very little discovery
      - need opposite emphasis, where FTS decides number of concurrent transfers
      - permitted concurrent transfers become property of and endpoint not a channel
    - There is currently limited information on network or SE state
      - will require more information to be available to start with
      - inc feedback from recent transfers



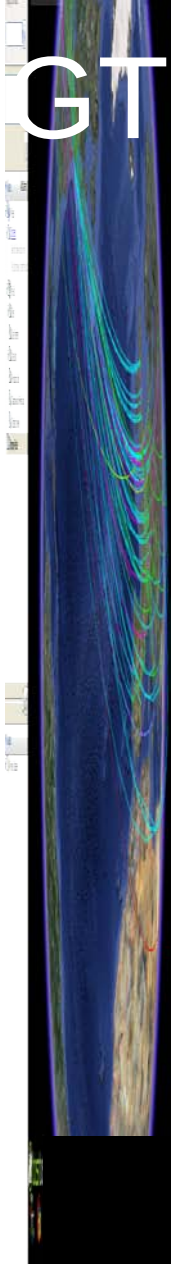
FTS - opportunities

- Create a more generic file transfer scheduler serving other use cases, such as 'chaotic' user initiated transfers
  - migration of output files
    - needs site reconfiguration as output files are typically removed at end of job
      - intermediate copy? stage area?
    - lcg-cr could hand over to asynchronous scheduling
  - cache population
    - remote files can be used via
      - remote byte access
      - transfer for local byte access
    - with vanilla local disk caching there is no scheduling
      - hand over instead to a scheduling service
    - namespaces for access
      - global logical namespace
        - » remote site not mounted directly (eg by NFS)
        - » Implementation could trigger transfer transparently



FTS – design considerations

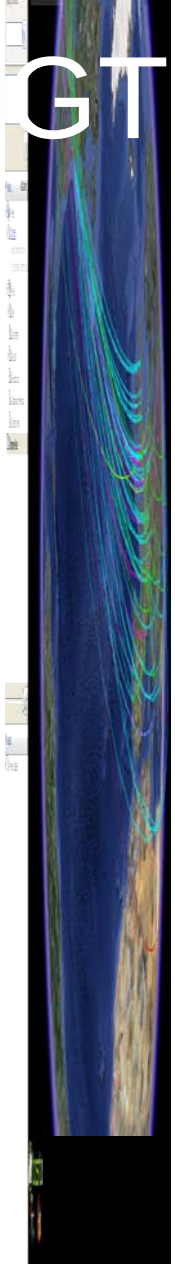
- Zero configuration
  - Through messaging and new information in the Infosys.
  - Access to more information
    - Info on SEs
      - free space
      - load
    - Info on networking
      - different for dedicated / shared links
      - what other info required?
- Endpoint-centric information model
- Dynamic selection of FTS server for a particular transfer
  - this is useful even without new use cases
  - can be implemented via messaging
    - promote interoperability between providers
- Catalogue Interaction
  - Should copy-register style operations be directly supported?
- Allow other db backends, even non relational



FTS - information sources

CERN IT Department

- How can site policy, network info and SE state be encapsulated?
  - Instrument SEs
    - query directly by FTS
    - populate Infosys
  - Messaging
    - FTS could consume 'not available' messages from SEs
    - messaging good for incidents
    - catalogue synchronisation work
  - Direct FTS config (last resort)



Summary

- DPM/LFC/FTS are alive and well
  - and have just recruited 2 new developers
- DPM/LFC evolutionary development
  - standards
    - Dedicated talk by Ricardo Brito Da Rocha
  - catalogue consistency
- FTS has more radical plans under discussion
  - generic transfer scheduler
  - existing service will be maintained
- A coherent set of services is required
  - using standards promotes participation
- Program of Work:
  - https://twiki.cern.ch/twiki/bin/view/LCG/DMProgramOfWork