



FTS3

WLCG robust, simplified and high-performance data movement service

CHEP 2013

Michail Salichos
IT/SDC

14/10/2013





Overview

- Background
- Features
- Status
- Road-map
- Summary



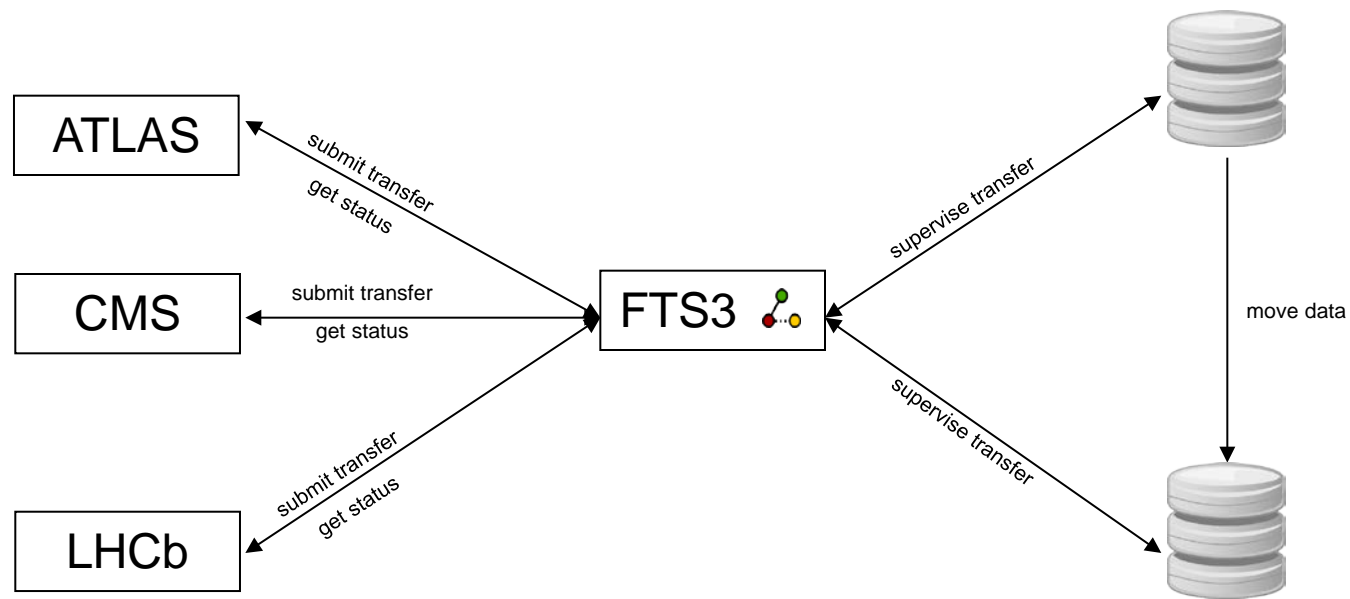
What is FTS?

- The service responsible for distributing the majority of LHC data across WLCG infrastructure
 - mature service - running for almost 10 years
- Low level data movement service, responsible for moving sets of files from one site to another while allowing participating sites to control the network resource usage



How it works

- Users interact with FTS by submitting transfer jobs, that simply say "copy <source URL> to <destination URL>"
 - FTS then queues, schedules and performs the transfer, retrying it if necessary



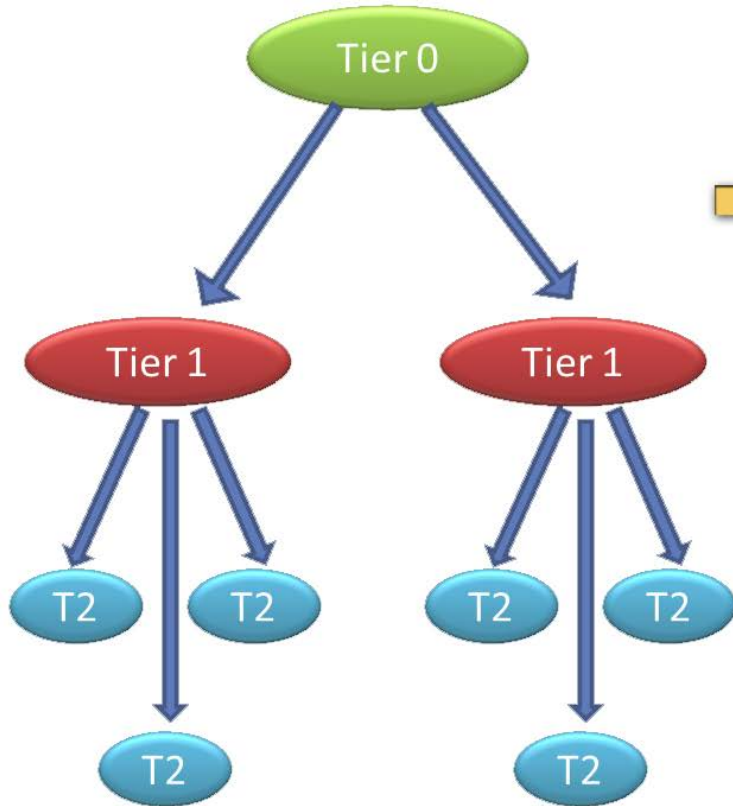


Motivation behind FTS3

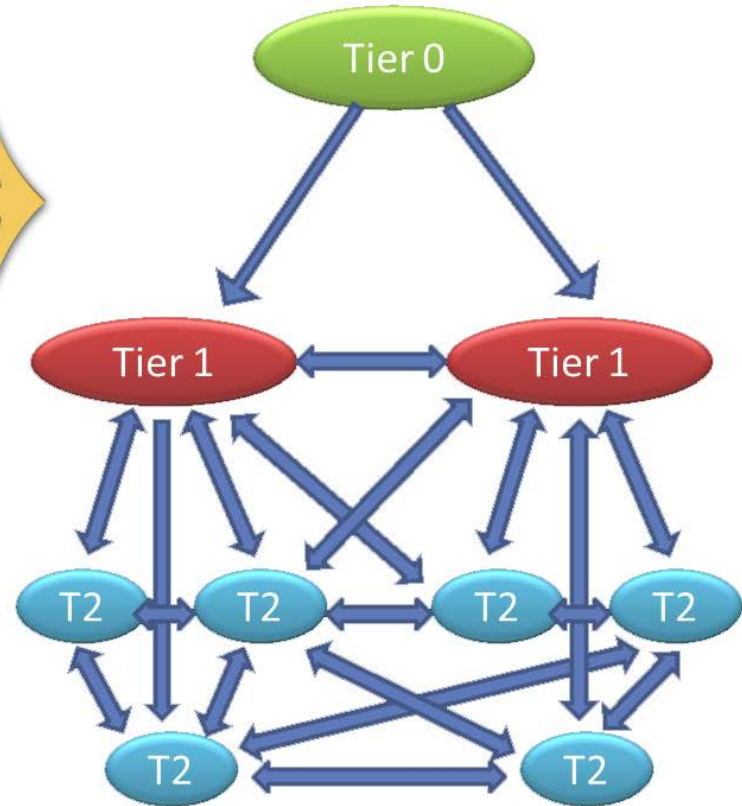
- address a particular set of FTS2 shortcomings, e.g.
 - relax the requirement to configure channels
 - protocols support
 - code maintenance issues
- simple to install and configure
- easy to maintain and support
- light-weight and not “resource hungry”
- support transferring large volume of data
- scale well horizontally
- control and efficiently use resources (network, SEs)

FTS3 goal

from hierarchical topology...



to mesh...



... and move data at a very large scale



How to access the service

- Clients and interfaces
 - FTS2 clients compatibility
 - FTS3 clients with many new features
 - RESTful API for standard clients using JSON
 - Python bindings for custom clients



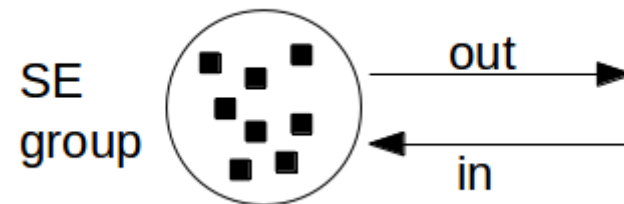
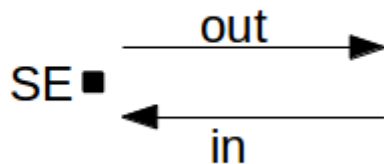
Resource optimization

- Adaptive optimization – let FTS3 decide
- Session reuse
 - GridFTP channel caching
 - SRM KeepAlive
 - HTTP SSL context reuse
- Multiple replicas support
- Smart transfer retry mechanism



Resource management

- Protocols support
 - GridFTP, SRM, HTTP, xroot
 - On top of GFAL2 – provides protocol plug-ins
- Blacklisting users (DN) and SEs
- Endpoint-centric configuration



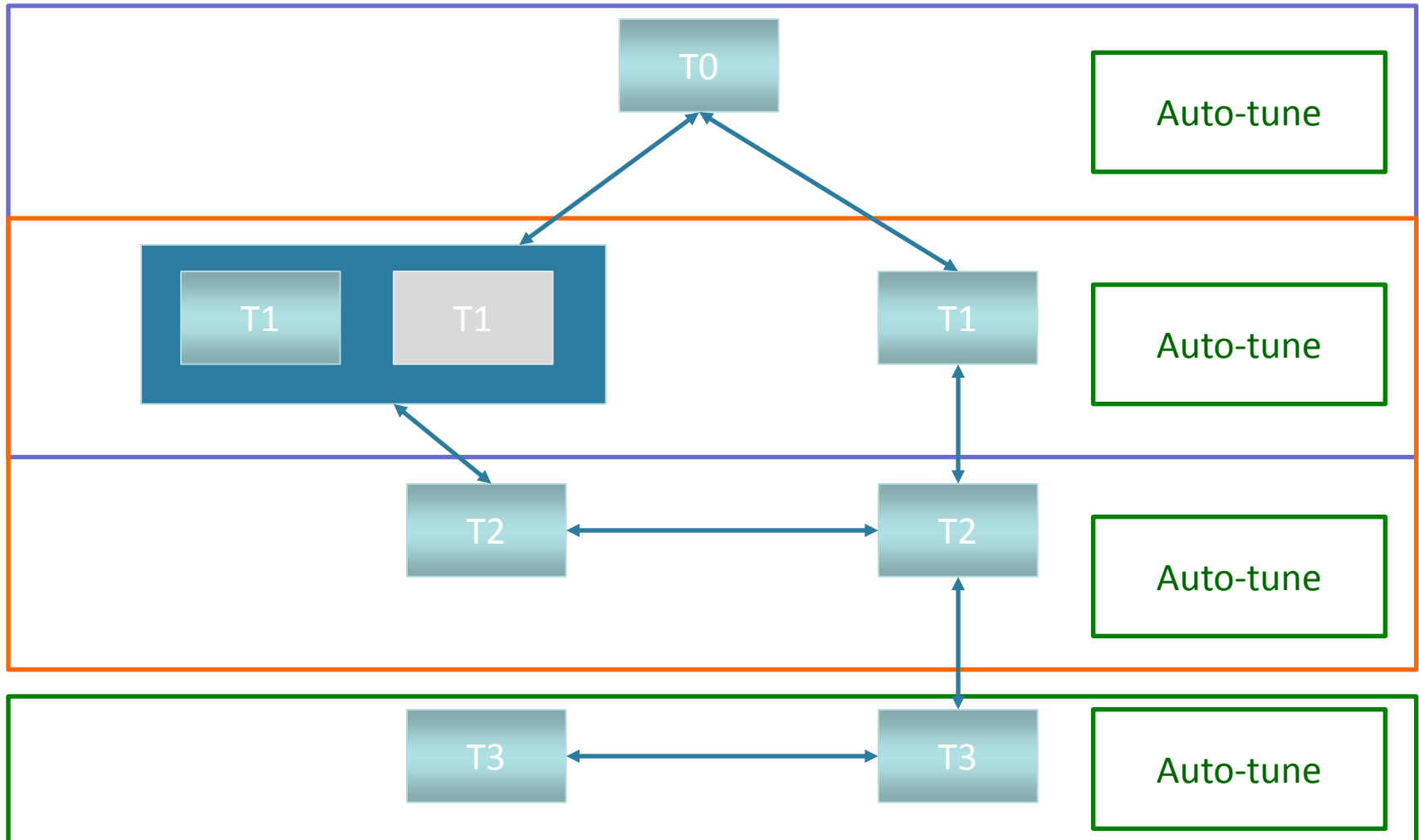


Deployment

- Horizontal scalability
- Minimal initial configuration
 - Mostly stored into the database
- DB backend support
 - MySql, Oracle
 - SQLite and/or PostgreSQL if requested

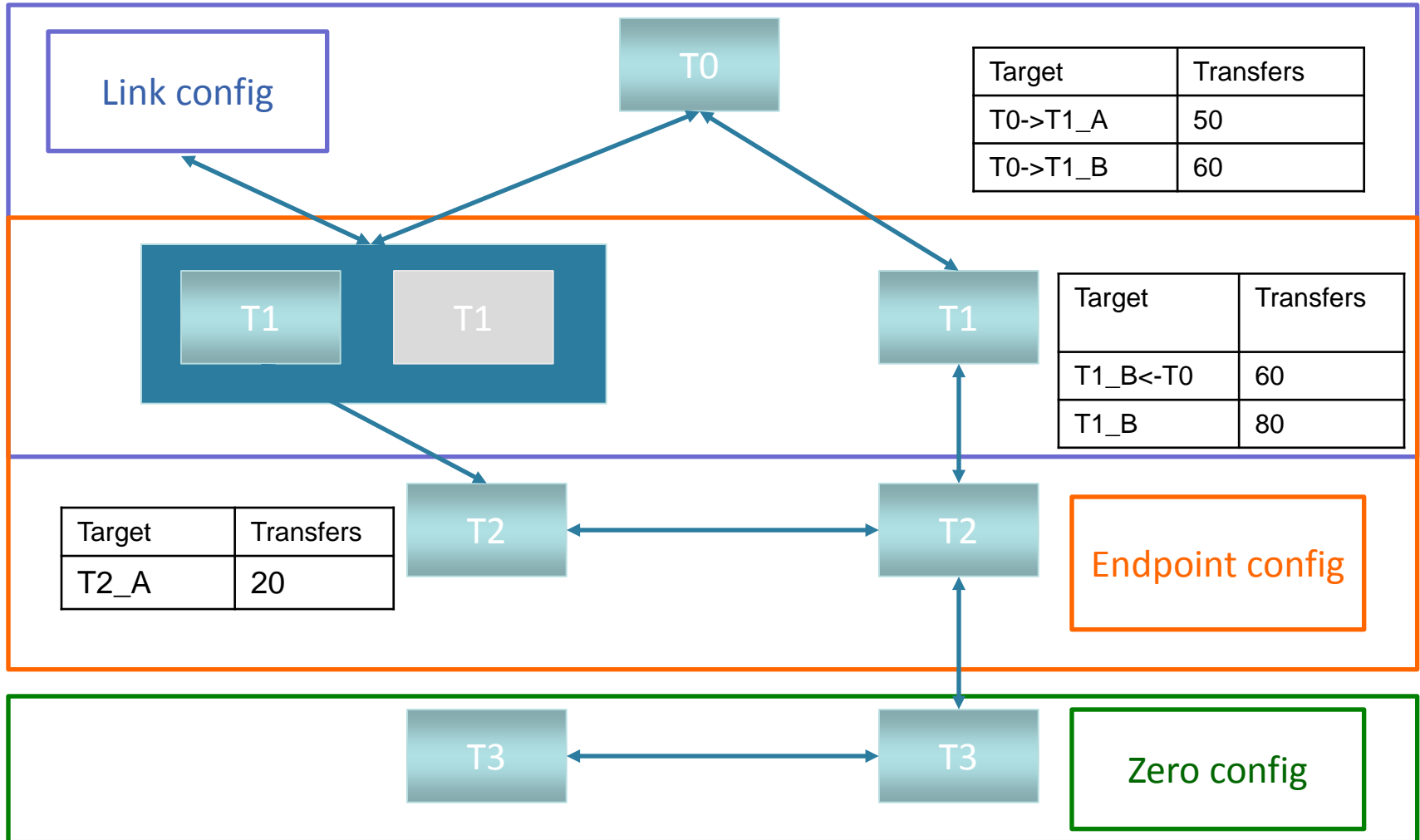


Configuration model





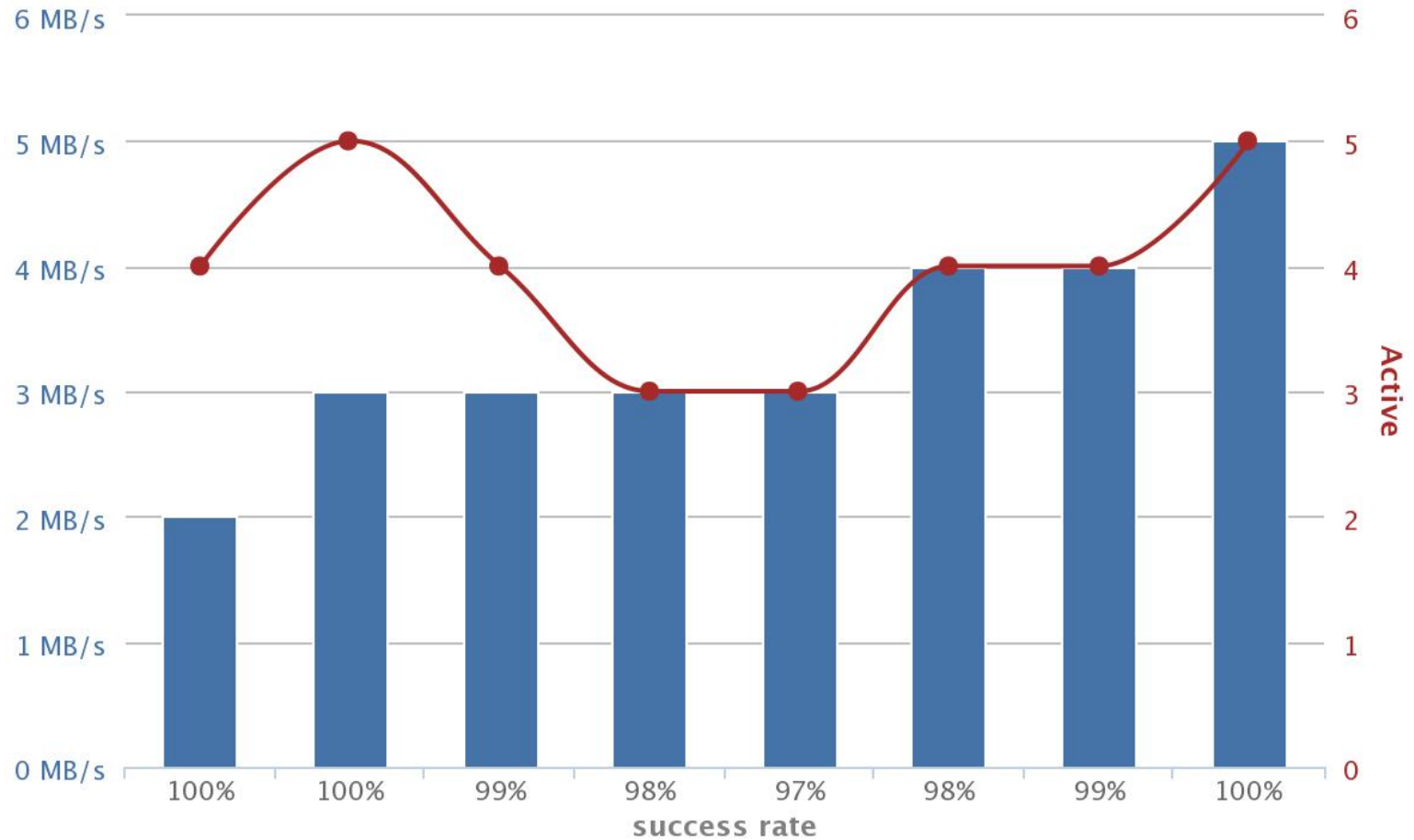
Configuration model (2)





Core features

Adaptive optimization



Highcharts.com



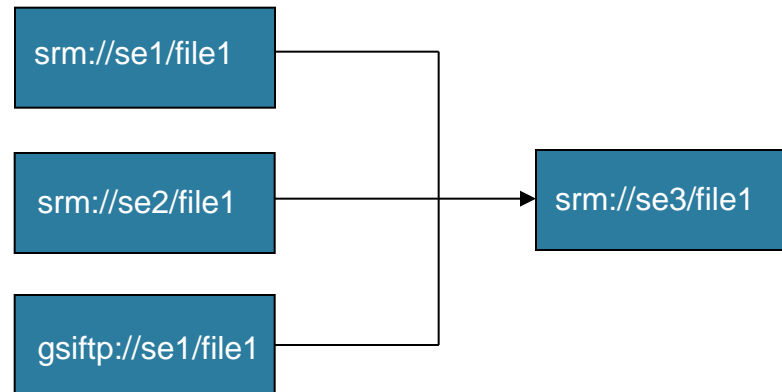
Core features (2)

- RESTful API
 - clients installation not needed
 - standard clients and/or libraries can be used
 - [lib]Curl, Python's urllib2...
 - well defined JSON schema



Core features (3)

- Multiple replicas support
 - Modes
 - Automatic – let FTS3 decide the order of replicas based on historical information
 - Manual - respect order set






Core features (4)

- Retry failed transfers
 - per individual job or globally set
 - failures classified as recoverable or not
 - Non-recoverable
 - No such file or directory
 - No space left on device
 - Permission denied
 - Read-only file system
 - etc
 - Recoverable – to be retried
 - All the rest
- More information in FTS3 wiki page



Monitoring

- WLCG Dashboard transfers UI 
 - Developed by CERN Dashboard team
 - A single entry point to the monitoring data collected from the distributed systems of the LHC
 - Monitor multiple FTS3 instances
 - each FTS3 server publishes messages to a message bus to report transfer status and state transitions
- Web interface for individual FTS3 server monitoring
 - In-depth details about job information, queued jobs, audit-trails, etc
- Nagios probes

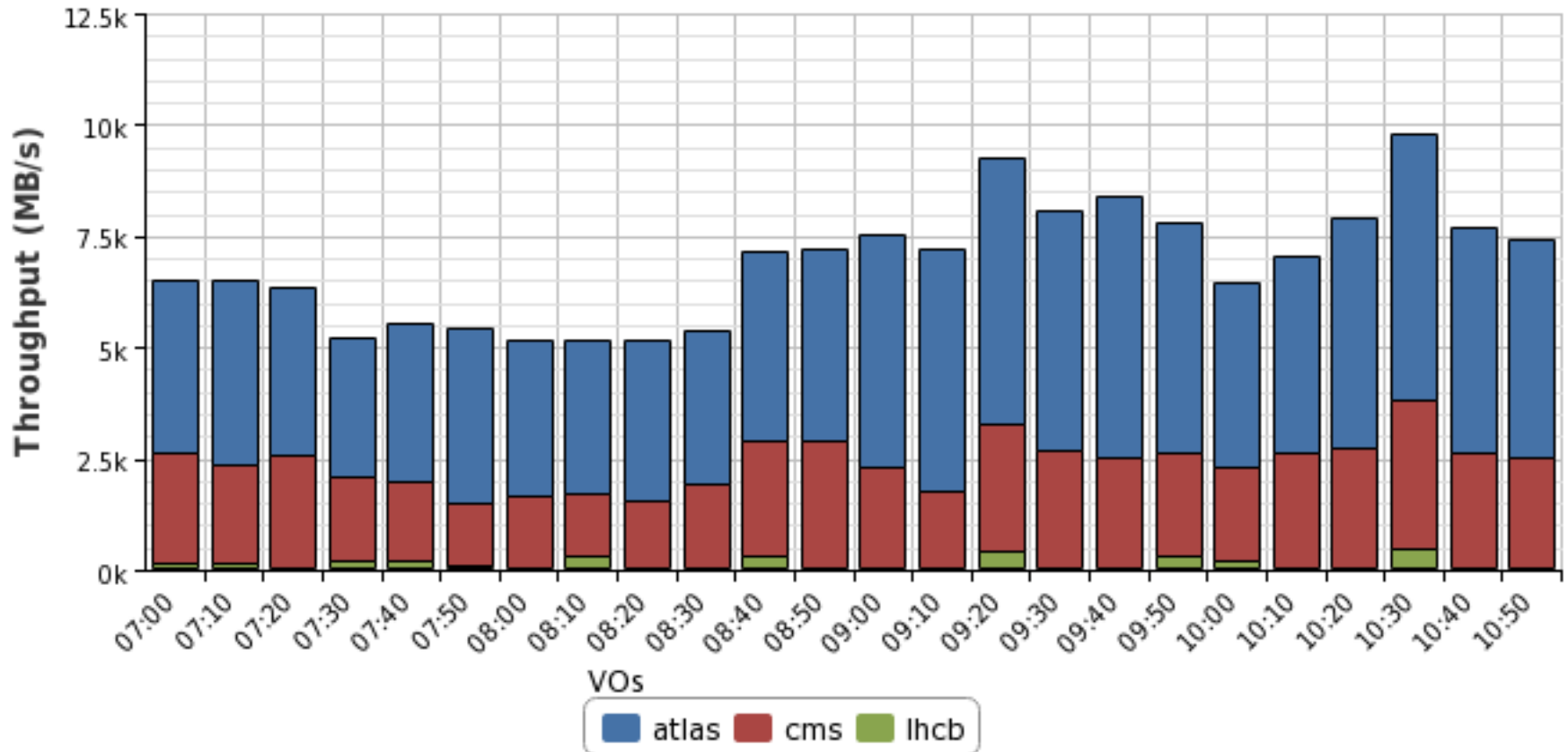


Global monitoring



Throughput

2013-10-10 07:00 to 2013-10-10 11:00 UTC



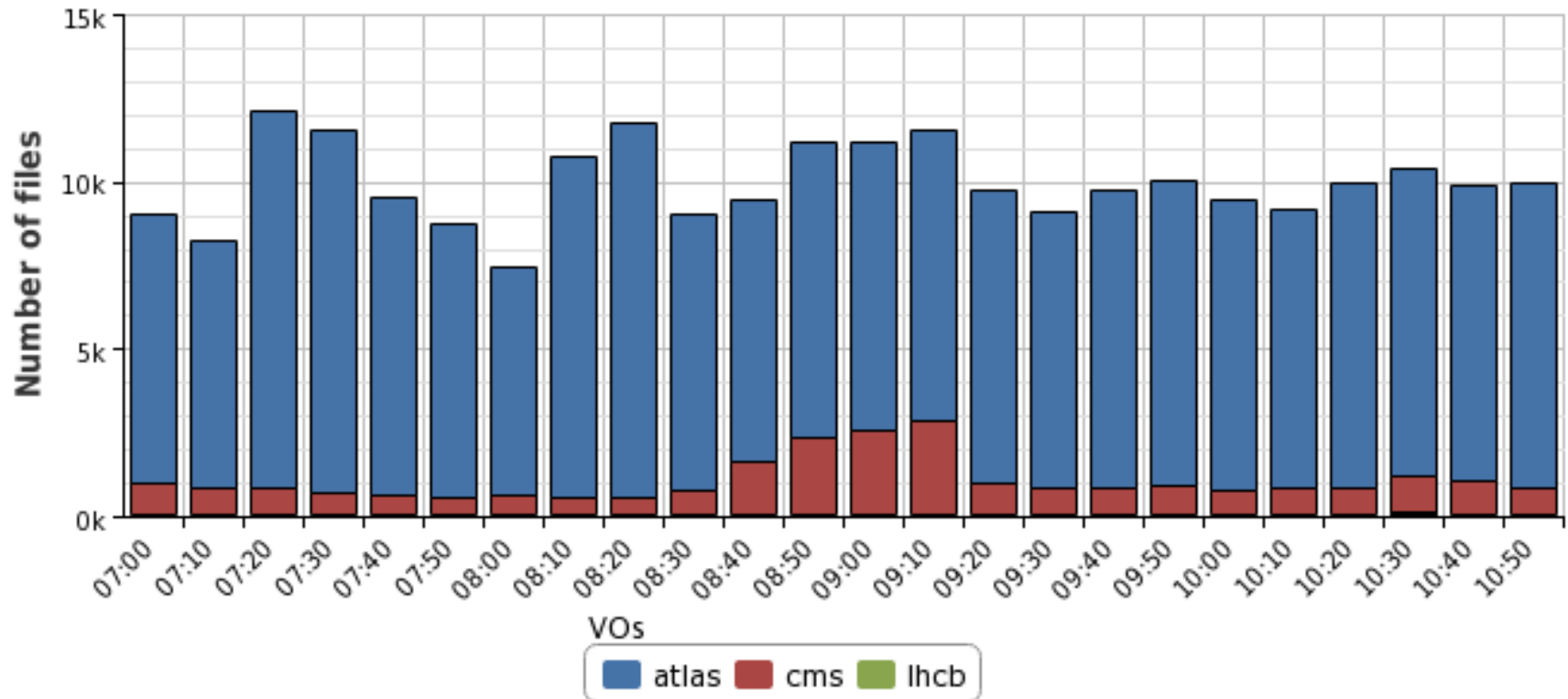


Global monitoring (2)



Transfer Successes

2013-10-10 07:00 to 2013-10-10 11:00 UTC





Global monitoring (3)



WLCG TRANSFERS DASHBOARD

Latest statistics update: 2013-06-26T13:30:00.176536

TRANSFER (2013-06-26 09:30 to 2013-06-26 13:30 UTC SLIDING)

MAP: TIERS ▾ LINK ▾ STYLE ▾

Summary

Interval

Last 4 hours

VOs

atlas

cms

lhcb

Technologies:

fts

xrootd

access type:

remote access

access mode:

Remote IO

Sources

Countries:

Sites:

Host:

Grouping: COUNTRY

Destinations

Countries:

Sites:

Host:

Grouping: COUNTRY

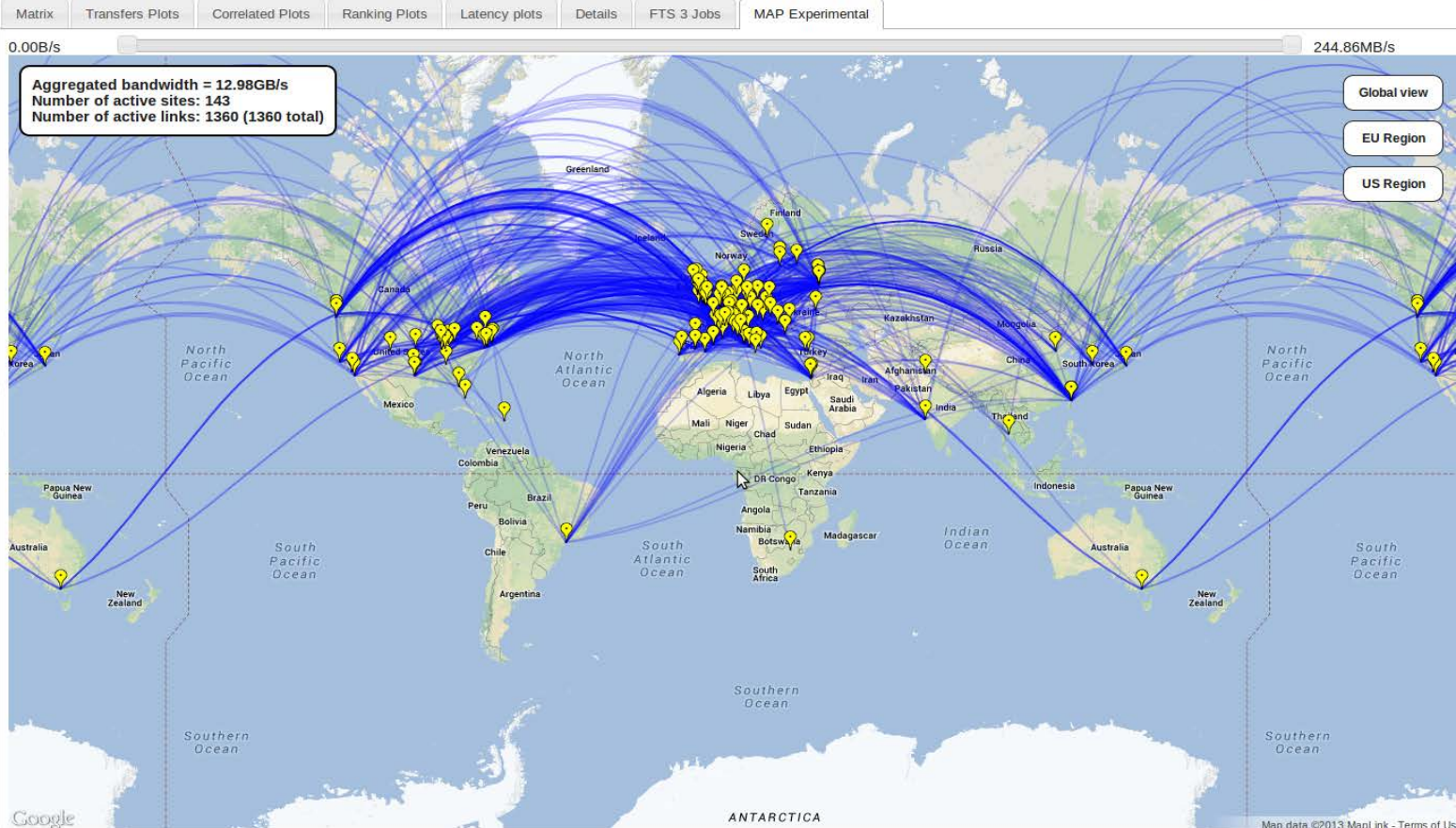
Interval

VOs

Technologies

Sources

Destinations



IT-SDC

FTS3 – WLCG data movement service



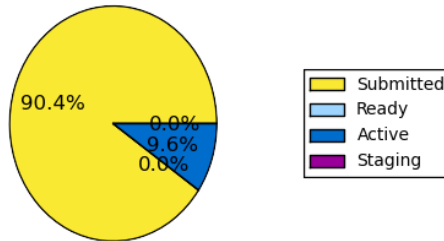
Standalone monitoring

Statistics / Overview

Queue

Submitted	Ready	Active	Staging
4602	0	487	0

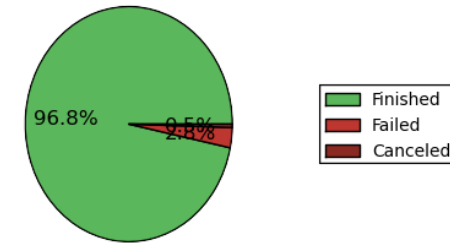
Queue



For the last hour

Succeeded	Failed	Canceled
6707	192	32

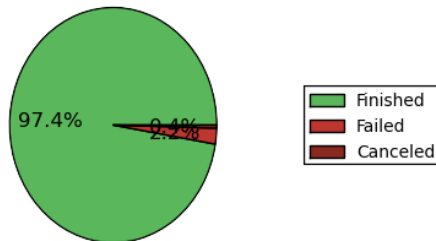
1h



For the last 24 hours

Succeeded	Failed	Canceled
188929	4216	821

24h





Releases

- Available in
 - EPEL 6 (fts-*)
 - our continuous integration repository (stable)
- Platform supported
 - SL6 / 64 bit



Testing and evaluation

- Installed at CERN, RAL, PIC, KIT, ASGC, BNL, IN2P3 and PNL
 - production and testing
 - > 1 year as a Pilot service
- Heavily used by ATLAS for prod jobs
 - avg weekly transfer volume from RAL ~1.5PB
- Tested by
 - LHC experiments
 - EGI/EUDAT against globus GridFTP, dCache GridFTP and GridFTP interface for iRODS (Griffin)
 - many other VOs already tested it successfully:
snoplus.sno.ubc.ca, ams02.cern.ch, vo.paus.pic.es, magic, T2K, NA62, etc
- Planned to run a “service challenge”
- Entering production!

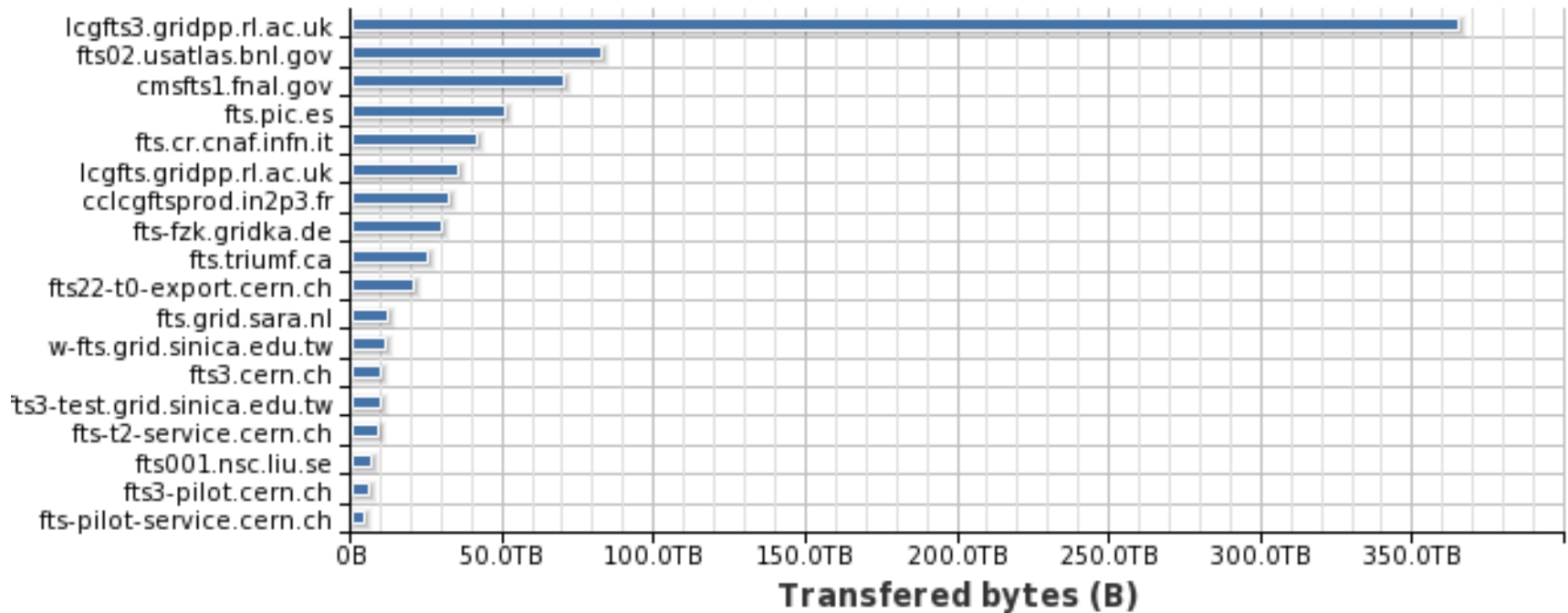


Sample volume



Total number of bytes transferred group by server

2013-09-09 08:50 to 2013-09-10 08:50 UTC

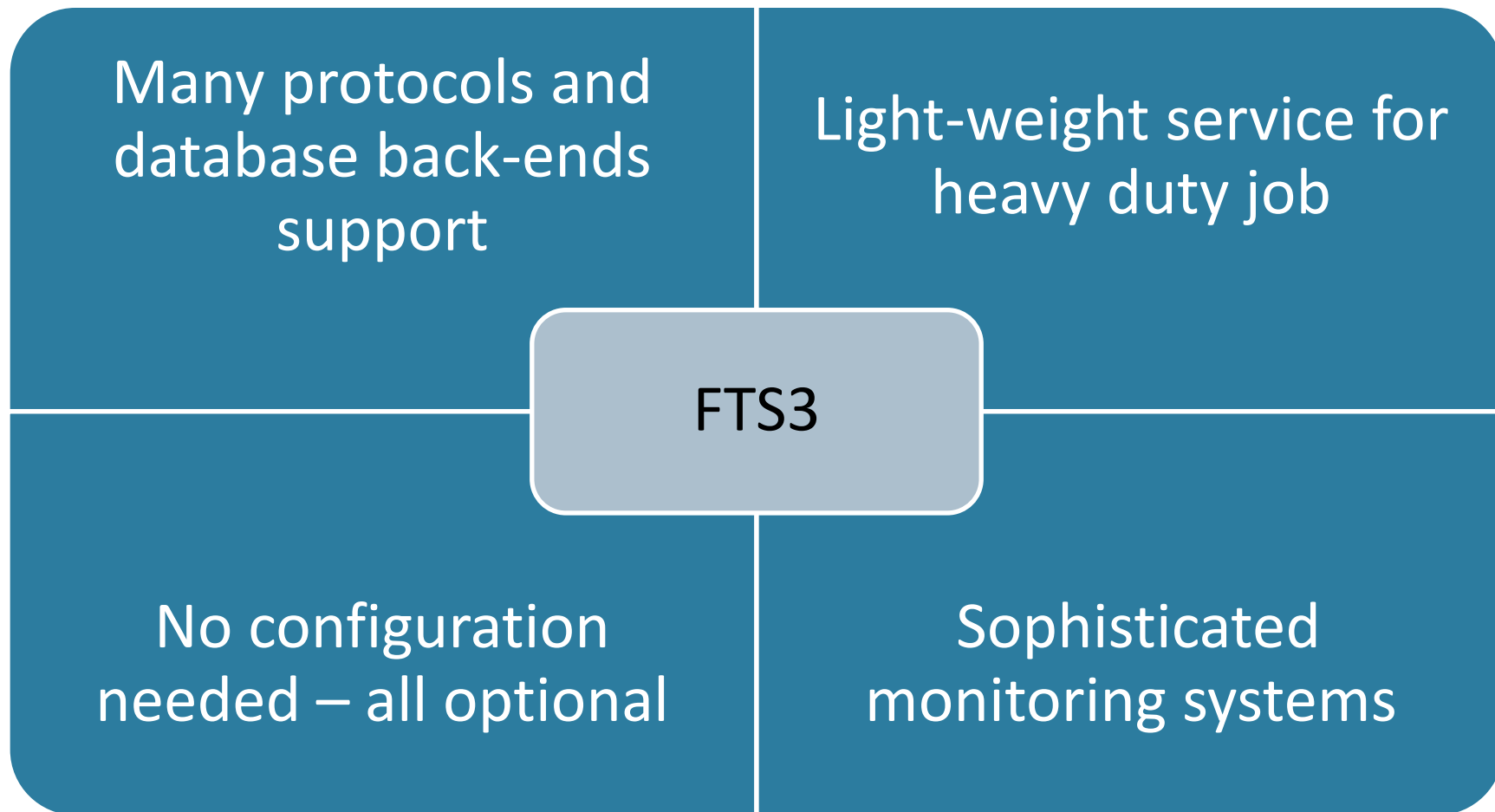




Roadmap

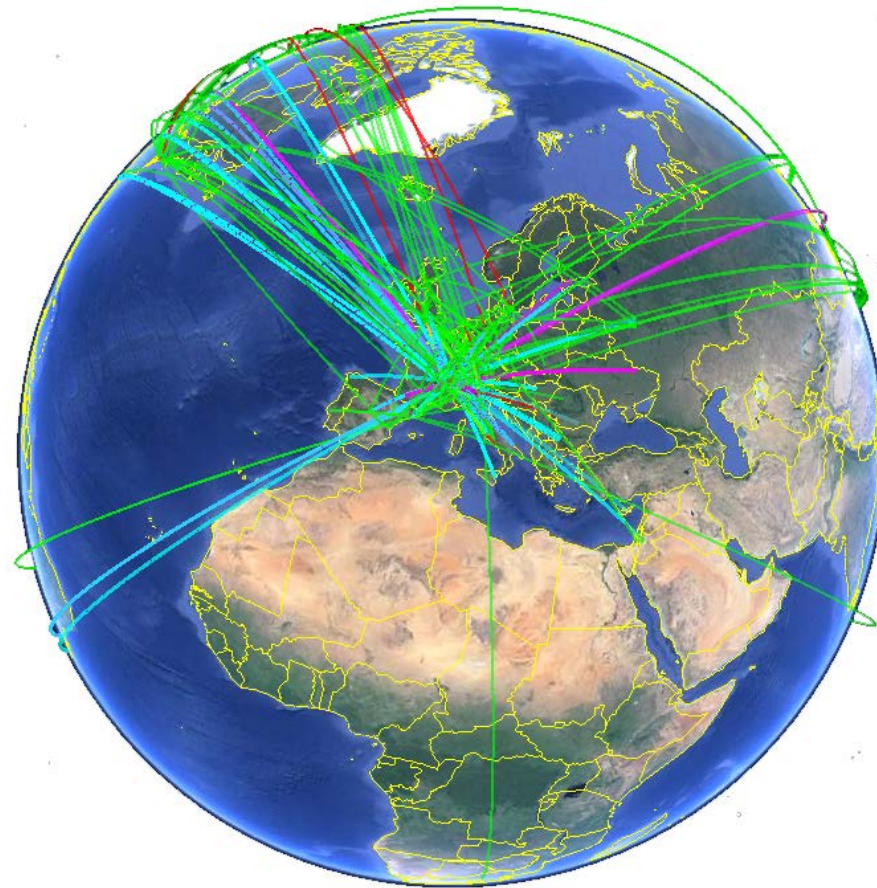
- Entirely determined by experiment requirements and prioritization
- What's next
 - Global scheduling and shared VO configuration across distributed FTS3 servers
 - Multi-hop transfers
 - VO shares per activity (primary, production, secondary, tier0, tier1, etc)
 - Integration and testing of perfSonar information (bandwidth & ping tests) for transfer optimization
 - deeper integration with archival storage and include high performance file management capabilities (deletes, renames...)
 - Keeping an eye on bandwidth reservation evolution

FTS3 – WLCG new data movement service





FTS3



Thank you!