Summary of the status and plans of FTS and LFC back-end database installations at Tier-1 Sites

Gordon D. Brown

Rutherford Appleton Laboratory

Overview

- Acronyms

- 11 Sites

- 26 Questions

- Developer overview

- Conclusions & Summary

FTS

Acronyms - FTS

- Full-Text Search (filename extension)

- Fourier Transform Spectroscopy

- Fault Tolerant System

- Federal Technology Service

- Flying Training Squadron (USAF)

- Flight Termination System

- Full-Time Support

Acronyms - FTS

- **Fourier Transform Spectroscopy**
  - is a measurement technique whereby spectra are collected based on measurements of the temporal coherence of a radiative source, using time-domain measurements of the electromagnetic radiation or other type of radiation.

Overview - FTS

- grid **F**ile **T**ransfer **S**ervice

- FTS Web Service
  - This component allows users to submit FTS jobs and query their status. It is the only component that users interact with.

- FTS Channel Agents
  - Each network channel, e.g CERN-RAL has a distinct daemon running transfers for it. The daemon is responsible for starting and controlling transfers on the associated network link.

- FTS VO Agents
  - This component is responsible for VO-specific parts of the transfer (e.g. updating the replica catalog for a given VO or applying VO-specific retry policies in case of failed transfers). Each VO has a distinct VO agent daemon running for it.

- FTS Monitor
  - This is currently a CERN only element.
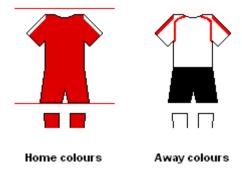
LFC

Acronyms - LFC

- Liverpool Football Club

- Lake Forest College (Lake Forest, IL)

- Level of Free Convection (meteorology)

- Los Fabulosos Cadillacs (Argentina band)

- Large Format Camera

- Land Forces Command (Canada)

- Load Frequency Control

Acronyms - LFC

- Full name: Liverpool Football Club
- Nickname(s): Pool, The Reds
- Founded: March 15, 1892
- Ground: Anfield, Liverpool, England
- Capacity: 45,362
- Manager: Rafael Benítez
- League: Premier League
- 2007–08: Premier League, 4th

Home colours   Away colours

Overview - LFC

- **L**CG **F**ile **C**atalog

- The LFC is a catalog containing logical to physical file mappings. Depending on the VO deployment model, the LFC is installed centrally or locally.

- The LFC is a secure catalog, supporting GSI security and VOMS.

- In the LFC, a given file is represented by a Grid Unique IDentifier (GUID). A given file replicated at different sites is then considered as the same file, thanks to this GUID, but (can) appear as a unique logical entry in the LFC catalog.

- The LFC allows us to see the logical file names in a hierarchical structure.

FTS

Do you have FTS running at your site?

CERN	Yes
CA-TRIUMF	Yes
DE-GridKa	Yes
ES-PIC	Yes
FR-IN2P3	Yes
IT-CNAF	Yes
NDGF	Yes
NL-SARA	Yes
TW-ASGC	Yes
UK-RAL	Yes
US-BNL	Yes

What back-end database does FTS run on?

CERN	Oracle
CA-TRIUMF	Oracle
DE-GridKa	Oracle
ES-PIC	Oracle
FR-IN2P3	Oracle
IT-CNAF	Oracle
NDGF	Oracle
NL-SARA	Oracle
TW-ASGC	Oracle
UK-RAL	Oracle
US-BNL	Oracle

Would you consider this database dev, test or production?

CERN	Prod
CA-TRIUMF	Prod
DE-GridKa	Prod
ES-PIC	Prod
FR-IN2P3	Prod
IT-CNAF	Prod
NDGF	Prod
NL-SARA	Prod
TW-ASGC	Prod
UK-RAL	Prod
US-BNL	Prod

If you have a prod copy, do you also have a dev or test copy?

CERN	Yes
CA-TRIUMF	No
DE-GridKa	No
ES-PIC	Pre-prod and Test
FR-IN2P3	No
IT-CNAF	Pre-prod
NDGF	No
NL-SARA	No
TW-ASGC	Test
UK-RAL	No
US-BNL	Prod copy and now implementing test copy

Is this database dedicated to FTS or does it share it with other schemas/applications?

CERN	There are other schemas
CA-TRIUMF	Dedicated
DE-GridKa	Will be shared with LFC after migration
ES-PIC	Dedicated
FR-IN2P3	Shared with LFC
IT-CNAF	Shared on RAC with LFC but dedicated nodes
NDGF	Dedicated
NL-SARA	Dedicated
TW-ASGC	Shared with LFC
UK-RAL	Dedicated, will be shared with LFC in future
US-BNL	Dedicated

Is this database a cluster? If so, how many nodes?

CERN	4 node cluster
CA-TRIUMF	Single instance
DE-GridKa	3 node cluster
ES-PIC	2 node cluster
FR-IN2P3	4 node cluster
IT-CNAF	3 node cluster
NDGF	Single instance
NL-SARA	Single instance
TW-ASGC	3 node cluster
UK-RAL	Single instance (going to 2/3 node cluster)
US-BNL	2 node cluster

What is the backup policy on this database?

CERN	Ondisk backup + tape backup every hour (arch logs)
CA-TRIUMF	- RMAN Sun & Wed weekly L0 backup - M,T,TH,FRI,SAT daily L1 backup - Every 30 minutes archivelog backup
DE-GridKa	2 full backup copies kept (2 weeks)
ES-PIC	- Full a week, - Differential Mon/Tue/Wed - Cumulative on Thurs - Differential on Wed and Sun - Archive logs backup up every day
FR-IN2P3	1 full each week, incremental on other days

What is the backup policy on this database?

IT-CNAF	
NDGF	Daily backups
NL-SARA	- Daily dumps to the filesystem - Tivoli Storage Manager takes care of the filesystem backup
TW-ASGC	Daily incremental. L0 on Mon, L1 on Tue-Sun.
UK-RAL	Full backup every week, incremental every day
US-BNL	- Full image copy updated every day - Archivelogs every 6 hours - Recovery window 14 days

How is this database monitored?

CERN	OEM + home-brewed monitoring
CA-TRIUMF	OEM
DE-GridKa	OEM, nagios and ganglia
ES-PIC	OEM, nagios and ganglia
FR-IN2P3	OEM
IT-CNAF	OEM
NDGF	Host: nagios and Ganglia. No DB monitoring.
NL-SARA	Scripts
TW-ASGC	OEM and nagios
UK-RAL	OEM and nagios
US-BNL	Ganglia, nagios

Are you replicating this database?

CERN	No
CA-TRIUMF	No
DE-GridKa	No
ES-PIC	No
FR-IN2P3	No
IT-CNAF	No
NDGF	No
NL-SARA	No
TW-ASGC	No
UK-RAL	No
US-BNL	No

Do you have any other redundancy built in to this database?

CERN	No
CA-TRIUMF	Oracle Data Guard
DE-GridKa	Mirrored FTS DB storage in SAN (i.e. 140 GB for data + 140 GB for recovery + 2 x 140 GB for mirroring).
ES-PIC	No
FR-IN2P3	SAN devices with RAID redundancy
NDGF	No
NL-SARA	No
TW-ASGC	Disk storage for RAID 6
UK-RAL	SAN RAID 1
US-BNL	Storage has hardware RAID controller 2 hots spares disks.

Do you have any other redundancy built in to this database?

IT-CNAF	- EMC storage device (CX3-80) which is highly redundant in all his parts (dual controller, double fibre channel connections and so on) - We use RAID 1 mirroring and ASM as storage management software (even if is configured with redundancy=external because we prefer to exploit hardware RAID1 redundancy, as recommended by EMC best practices). - We have a 3-nodes RAC, each node is connected to a different network switch and is equipped with dual power supply, dual port HBAs, RAID 1 local discs.

Do you plan to move FTS to another database? If so, which?

CERN	No
CA-TRIUMF	No
DE-GridKa	No
ES-PIC	No
FR-IN2P3	No
IT-CNAF	No
NDGF	No
NL-SARA	No
TW-ASGC	No
UK-RAL	No
US-BNL	No

What are your plans for the FTS database?

CERN	None
CA-TRIUMF	Move to Oracle RAC
DE-GridKa	None
ES-PIC	Upgrade with the last schema given by the FTS team on 18 March, in order to use the history and admin packages
FR-IN2P3	Plans are : monitor and take appropriate actions to increase QOS !! This may lead to change hosts, OS , etc
IT-CNAF	None

What are your plans for the FTS database?

NDGF	Possibly move it to the university's central database group. Also consolidating it to use the 3D installation is considered.
NL-SARA	To change the backup policy so that RMAN is used instead of dumps on the filesystem.
TW-ASGC	None
UK-RAL	Move to Oracle RAC
US-BNL	Optimize if needed

Does the same DBA looking after the FTS and 3D databases?

CERN	Yes – same team
CA-TRIUMF	Yes
DE-GridKa	Yes
ES-PIC	Yes
FR-IN2P3	Yes
IT-CNAF	Yes – same team
NDGF	No
NL-SARA	No
TW-ASGC	Yes
UK-RAL	Yes – same team
US-BNL	Yes

LFC

Do you have LFC running at your site?

CERN	Yes
CA-TRIUMF	Yes
DE-GridKa	Yes
ES-PIC	Yes
FR-IN2P3	Yes

Do you have LFC running at your site?

IT-CNAF	- LHCb: we have a replica of the central catalog hosted at CERN - ATLAS: we have a local catalog, this is not replicated. The catalogs are installed on 2 different oracle RACs, they are both in production and we have a pre-production installation. The same considerations made for pre-production FTS apply here for LFC.
NDGF	We will soon
NL-SARA	Yes
TW-ASGC	Yes
UK-RAL	Yes
US-BNL	No

What back-end database does LFC run on?

CERN	Oracle
CA-TRIUMF	mySQL
DE-GridKa	mySQL
ES-PIC	mySQL
FR-IN2P3	Oracle
IT-CNAF	Oracle
NDGF	mySQL, but would prefer PostGreSQL support
NL-SARA	mySQL and Oracle
TW-ASGC	Oracle
UK-RAL	Oracle
US-BNL	n/a

Would you consider this database dev, test or production?

CERN	Prod
CA-TRIUMF	Prod
DE-GridKa	Prod
ES-PIC	Prod
FR-IN2P3	Prod
IT-CNAF	Prod
NDGF	Test – prod soon
NL-SARA	Prod
TW-ASGC	Prod
UK-RAL	Prod
US-BNL	n/a

If you have a prod copy, do you also have a dev or test copy?

CERN	Yes
CA-TRIUMF	No
DE-GridKa	No
ES-PIC	Test
FR-IN2P3	No
IT-CNAF	No
NDGF	No
NL-SARA	No
TW-ASGC	No
UK-RAL	No
US-BNL	n/a

Science & Technology Facilities Council
e-Science

Is this database dedicated to LFC or does it share it with other schemas/applications?

CERN	There are other schemas
CA-TRIUMF	Dedicated
DE-GridKa	Dedicated
ES-PIC	Dedicated
FR-IN2P3	Shared with FTS
IT-CNAF	Shared with FTS
NDGF	Dedicated
NL-SARA	MySQL: Dedicated. Oracle: Shared with the 3D database.
TW-ASGC	Shared with FTS
UK-RAL	Dedicated – will share with FTS in future
US-BNL	n/a

Is this database a cluster? If so, how many nodes?

CERN	4 node cluster
CA-TRIUMF	Single instance
DE-GridKa	Single instance
ES-PIC	Single instance
FR-IN2P3	4 node cluster
IT-CNAF	- LHCb: we have a 2 node RAC which hosts LFC and LHCb Conditions DB - ATLAS: we have a 3 node RAC which hosts LFC and FTS
NDGF	Single instance

Is this database a cluster? If so, how many nodes?

NL-SARA	MySQL: No. Oracle: Yes, 2 nodes.
TW-ASGC	3 node cluster
UK-RAL	Atlas: Single instance (going to 2/3 node cluster) LHCb: Part of 3D 2 node cluster
US-BNL	n/a

What is the backup policy on this database?

CERN	Ondisk backup + tape backup every hour (archive logs)
CA-TRIUMF	Nightly database backup
DE-GridKa	bin-logs; replication (hot standby) on master: daily diff-backup in TSM.
ES-PIC	Full backup every day
FR-IN2P3	1 full each week, incremental each other days
IT-CNAF	
NDGF	Daily offsite backups are planned

What is the backup policy on this database?

NL-SARA	- mySQL: Daily dumps to the filesystem. - TSM for the filesystem. - Oracle: RMAN talking to a TSM plugin.
TW-ASGC	Incremental – daily. level 0 on Monday, level 1 on Tuesday to Sunday
UK-RAL	Full backup every week, incremental every day
US-BNL	n/a

How is this database monitored?

CERN	OEM + home-brewed monitoring
CA-TRIUMF	Scripts
DE-GridKa	DB: MySQL Query Browser Host: nagios, ganglia).
ES-PIC	Nagios and ganglia
FR-IN2P3	OEM
IT-CNAF	OEM
NDGF	Host is monitored through Nagios and Ganglia. Plans for testing the database also with Nagios.
NL-SARA	Nagios
TW-ASGC	OEM and nagios
UK-RAL	OEM and nagios
US-BNL	n/a

Are you replicating this database?

CERN	LFC for LHCB is being replicated
CA-TRIUMF	No
DE-GridKa	Yes
ES-PIC	No
FR-IN2P3	The LHBc LFC replica
IT-CNAF	No
NDGF	No
NL-SARA	No
TW-ASGC	No
UK-RAL	Possibly
US-BNL	n/a

Do you have any other redundancy built in to this database?

CERN	RAC and replication for LHCB. Only RAC for rest.
CA-TRIUMF	No, but currently testing replication
DE-GridKa	No
ES-PIC	No
FR-IN2P3	- Databases datafiles are hosted on SAN devices with RAID redundancy and are backuped each day. - Our Tivoli Storage Manager system give use the ability to have at the same time a copy of those backups on disk and on tape.

Do you have any other redundancy built in to this database?

IT-CNAF	- EMC storage device (CX3-80) which is highly redundant in all his parts (dual controller, double fibre channel connections and so on) We use RAID 1 mirroring and ASM as storage management software (even if is configured with redundancy=external because we prefer to exploit hardware RAID 1 redundancy, as recommended by EMC best practices). - We have a 3 node RAC, each node is connected to a different network switch and is equipped with dual power supply, dual port HBAs, RAID 1 local discs.

Do you have any other redundancy built in to this database?

NDGF	No
NL-SARA	MySQL: No. Oracle: Well, the fact that it is a RAC cluster should provide some redundancy...
TW-ASGC	We have RAC DB with disk storage for RAID 6.
UK-RAL	SAN RAID 1
US-BNL	n/a

Do you plan to move LFC to another database? If so, which?

CERN	No
CA-TRIUMF	Oracle RAC
DE-GridKa	MySQL DB to be migrated to Oracle to the FTS 3-node RAC (see above). Then: 2 preferred LFC nodes (+ 1 standby LFC node) and LFC DB storage in SAN of 4 x 140 GB, i.e. 140 GB for data + 140 GB for recovery + 2 x 140 GB for mirroring.
ES-PIC	Move to 2 node Oracle RAC in April

Do you plan to move LFC to another database? If so, which?

FR-IN2P3	No
IT-CNAF	No
NDGF	PostGreSQL, if the LFC supports it.
NL-SARA	No
TW-ASGC	No
UK-RAL	No
US-BNL	n/a

What are your plans for the LFC database?

CERN	None
CA-TRIUMF	Migrate to Oracle RAC
DE-GridKa	MySQL DB to be migrated to Oracle to the FTS 3-node RAC (see above). Then: 2 preferred LFC nodes (+ 1 standby LFC node) and LFC DB storage in SAN of 4 x 140 GB, i.e. 140 GB for data + 140 GB for recovery + 2 x 140 GB for mirroring.
ES-PIC	The LFC database will have the same backup policy than FTS, and will be set up in the same storage device than 3D and FTS databases.

What are your plans for the LFC database?

FR-IN2P3	Plans are : monitor and take appropriate actions to increase QOS !! This may lead to change hosts, OS , etc
IT-CNAF	None
NDGF	
NL-SARA	To add nodes to the cluster
TW-ASGC	None
UK-RAL	Move to 2/3 node Oracle cluster
US-BNL	n/a

Does the same DBA looking after the LFC and 3D databases?

CERN	Yes – same team
CA-TRIUMF	No
DE-GridKa	Yes
ES-PIC	Yes
FR-IN2P3	Yes
IT-CNAF	Yes – same team
NDGF	No
NL-SARA	MySQL: No; Oracle: Yes
TW-ASGC	Yes
UK-RAL	Yes – same team
US-BNL	n/a

Conclusions

FTS Summary

Site	DB	Test/ Dev	Dedi- cated	Nodes	Backups	Monitoring	3D DBA
CERN	ORACLE	D/T	Other	4	Hourly dsk/tpe	OEM + own	Yes
CA-TRIUMF	ORACLE	No	Yes	1	RMAN daily	OEM	Yes
DE-GridKa	ORACLE	No	LFC	3	RMAN daily	OEM, ng, gg	Yes
ES-PIC	ORACLE	PP/T	Yes	2	RMAN daily	OEM, ng, gg	Yes
FR-IN2P3	ORACLE	No	LFC	4	RMAN daily	OEM	Yes
IT-CNAF	ORACLE	PP	LFC	3		OEM	Yes
NDGF	ORACLE	No	Yes	1	RMAN daily	ng, gg	No
NL-SARA	ORACLE	No	Yes	1	Dumps daily	scripts	No
TW-ASGC	ORACLE	T	LFC	3	RMAN daily	OEM, ng	Yes
UK-RAL	ORACLE	No	Yes	1	RMAN daily	OEM, ng	Yes
US-BNL	ORACLE	PP/T	Yes	2	Image daily	ng, gg	Yes

LFC Summary

Site	DB	Test/Dev	Dedi-cated	Nodes	Backups	Monitoring	3D DBA
CERN	ORACLE	D/T	Other	4	Hourly dsk/tpe	OEM + own	Yes
CA-TRIUMF	MySQL	No	Yes	1	RMAN daily	scripts	Yes
DE-GridKa	MySQL	No	Yes?	1	Hot standby	OEM, ng, gg	Yes
ES-PIC	MySQL	T	Yes	1	RMAN daily	ng, gg	Yes
FR-IN2P3	ORACLE	No	FTS	4	RMAN daily	OEM	Yes
IT-CNAF	ORACLE	No	FTS	2/3		OEM	Yes
NDGF	MySQL	No	Yes	1	Offsite plan	ng, gg	No
NL-SARA	MySQL ORACLE	No	Yes	1/2	RMAN dy/dmp	nagios	N/Y
TW-ASGC	ORACLE	No	FTS	3	RMAN daily	OEM, ng	Yes
UK-RAL	ORACLE	No	Yes	1/2	RMAN daily	OEM, ng	Yes

FTS Developer View on Database Plans

- **We pretty much leave it to the DBA really…**

- **In terms of plans, we have the ongoing plan to move more monitoring into the database – which means more summary and raw data stored. We'll also do the analytic summarization in PL/SQL, so you should expect and increase in CPU use as we start to do this. It's kinda hard to quantify though…**

Personal View

- DBAs should work together to tackle issues

- Not just DB setup but application issues

- 3D good infrastructure for databases

    - community
    - experience in many areas
    - plenty of people to solve problems

- FTS & LFC (and CASTOR) tie in with 3D and other Oracle databases

Summary

- Database are important to WLCG

- Work more with developers where needed

- Availability and tuning is key

- 3D community/experience is helping with other database deployment areas

- Use list, use meetings, help each other

Questions and (hopefully) Answers

g.d.brown@rl.ac.uk