



DDM/DQ₂: Summary

Vincent.Garonne@cern.ch
CERN-PH-ADP

On behalf of the ATLAS \mathcal{DQ}_2 development team

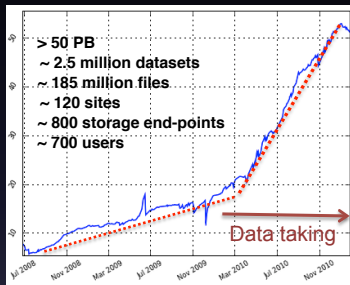
ATLAS Software&Computing Workshop, April, 2011

Outline

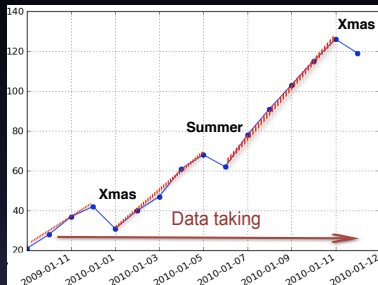
- Distributed Data Management session [Link]
 - Conveners: G.Steward, C.Serfon, V.Garonne
 - 7 talks
- Topics:
 - *DQ₂* software: Status, News and Future
 - *DDM* infrastructure: Deployment, Monitoring
 - *DDM* operations: Status, Current Issues and Challenges

Current scale

- Evolution of the total space (PB)



- Traces per month (M)

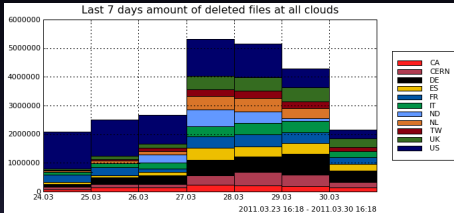


<http://bourricot.cern.ch/dq2/accounting/>

- Central services
 - 14M reads per day (162Hz)
 - 0.6M writes per day (7Hz)

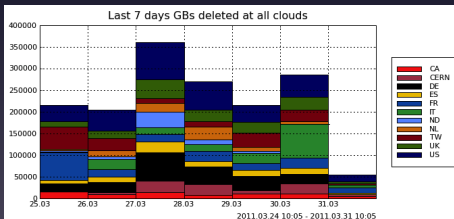
Common trend in the system

5M files/340 Tera deleted per day



Both experiment and user activity keep increasing

Development to ensure scalability and reliability, to protect and better manage the system



<http://bourricot.cern.ch/dq2/deletion/>

New accounting service

Storage space and usage information

- Possibility to break down volumes by metadata information
 - DQ_2 meta-data, e.g. location, custodality, owner ...
 - Few AMI meta-data, e.g. project, datatype, ...

Documentation&How-to [Link]

- Count all data10 datasets at CERN:
`dq2-list-accounting project=data10* location=cern*`
- Dataset name finder included for free
`dq2-find project=data10* location=cern*`

Timeline: *Mid-April in production*

- Pilot service - Contact: atlas-dq2-dev@cern.ch

Development & Deployment support

Mailing lists/Documentation under review

Packaging and distribution

- New package conventions (Python version independant)
- Drop python 2.3 and 2.4 support
- Support of python 2.5 and 2.6

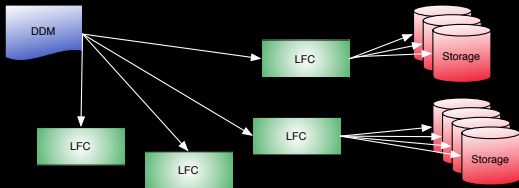
Infrastructure under continuous improvement

More info

- Cf. Angelos' talk - \mathcal{DQ}_2 Clients [link]
- Cf. Fernando's talk - \mathcal{DDM} Infrastructure [link]
- Cf. Graeme's talk - LFC Consolidation [link]

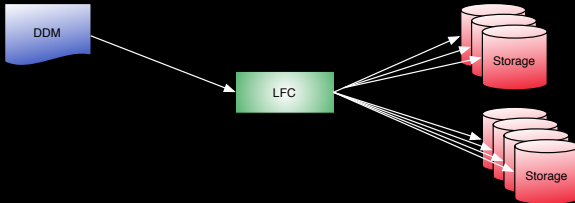
Current Architecture

- Each cloud has an LFC host at the Tier-1
 - Plus Tier-2 LFCs in US
- Distributes failures, so each failure is not so catastrophic
 - But more LFCs means there is a higher probability of a failure
 - More robust deployment scenarios require rather specialist knowledge
 - NL cloud caught by a weird Oracle DB restore problem, took cloud out for 2 weeks
- Many sources of lazily synced metadata that can drift apart



LFC Consolidation

- After discussion at the last S&C week and at Napoli
- Decision to consolidate to one LFC
- Addresses some of the drawbacks of the distributed LFC model



Tentative Timetable

- As of March 7th:
 - May 15th
 - Instance of ATLAS central LFC in production. It merges entries from local LFCs at CERN, NL, TW
 - Panda central registration for those clouds
 - Jul 15th
 - 50% of ATLAS resources are served by the central LFC
 - Read-only replica at BNL in place
 - Oct 15th
 - All non US sites are served by central LFC
 - Massive scalability test
 - Dec 15th
 - All sites served by central LFC

Current status

We have now a more dynamic ATLAS computing model

- Based on caching, data importance and popularity
 - Key DQ_2 services: Tracer, Popularity, Location, Victor, Deletion, Accounting
 - Strong interests from CMS/LHCb
- Break of cloud boundaries
 - T2Ds - Simone's talk [\[link\]](#)
 - Cf. UK experience, Alessandra's talk [\[link\]](#)
 - Multi-hop/Direct transfer - Kaushik's talk [\[link\]](#)
 - PD2P - Fernando's talk [\[link\]](#)

Future plans

New high level use cases and client work-flows

- Group support
- Data placement policy management

Evolution of middleware

- \mathcal{DQ}_2 on the cloud
- More features provided, i.d. faster `dq2-get` or even better ☺

New technologies / paradigms

- High availability service
- Meta-data / Messaging

⇒ Beyond the original \mathcal{DQ}_2 vision

⇒ Difficult to extend the existing system with new paradigms

DQ_2 – Version 3 (Code name *Rucio*)

Rucio



⇒ This is not an April Fool's Joke 😊

⇒ "YES !" from the audience

Rucio – Overview and plans I

Extensive internal improvements

- Gain for scalability/deployment/operation/development

More high level features

- E.g. group support, group/user quota

Heterogeneous resources

- Grid/Cloud Storages
- Transfer/file access protocols
- Data discovery services

Rucio – Overview and plans II

Support of the current version

- Bug fixes/Small changes/Protections
- New features in Rucio
- Migration / Synchronization foreseen

Timescale

- Roadmap will come
 - Extensive systems analysis is required
- Organizing requirements
 - Side note: DA user survey useful

Everyone is invited to suggest and discuss !

DQ₂ – Version 2

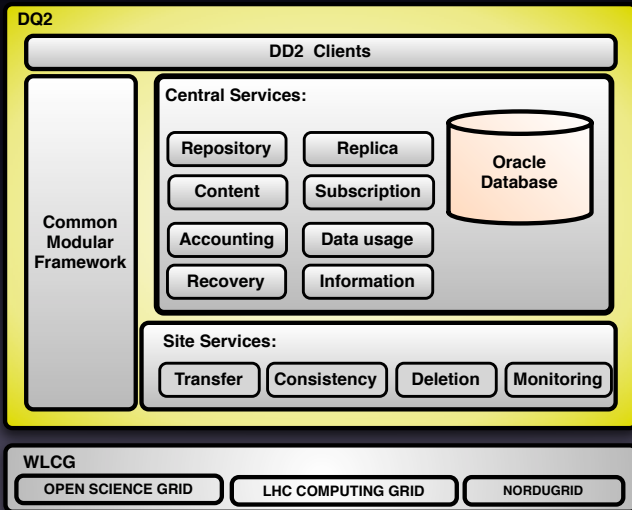
Production system

Analysis system

Data export

End-user

Physics meta-data system



DQ₂ – Version 3 (Code name Rucio)

Production system

Analysis system

Data export

End-user

Physics meta-data system

DQ2
V3
(Rucio)

High level services

Group support

Group/User quota

Data placement policy management

Testing framework

...

Core services

API / CLI Clients

Common Modular Framework

Central Services:

Repository

Replica

Content

Subscription

Accounting

Data usage

Recovery

Information

Technologies

Oracle Database

NoSQL

Messaging

Site Services:

Transfer

Consistency

Deletion

Monitoring

Grid + Cloud computing

OPEN SCIENCE GRID

LHC COMPUTING GRID

NORDUGRID

EC2

AMAZON (S3)

EBS

+ *Storages:* T3s, Proof, Xrootd, NFS4.1, HDFS, etc.

Functionalities: File discovery, Global namespace, Fail-over, caching

See you soon !



Example – High level use cases

New DQ_2 concepts for activity group support

- Logical location and sub-location
 - E.g. `T1s_groupdisks = {BNL_groupdisks, FZK_groupdisks, ...}`
 - E.g. `BNL_groupdisks/phys-higgs`
- Global/Local group quota
- **I.d. Symlinks, hierarchical namespace**
- Replication factor
 - E.g. 2 copies at `BNL/phys-higgs`, `FZK/phys-higgs`
- **I.d. Data placement policy engine**
- Transparent and automatic
 - E.g. `dq2-register-subscription --fact=2 <dataset>`
`T1s_groupdisks/phys-higgs`