$\mathcal{DDM}$: Data & Meta-Data

Vincent.Garonne@cern.ch

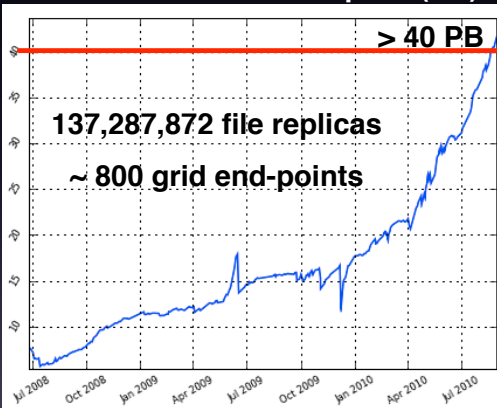CERN-PH-ADP

30 August 2010

ATLAS meta-data handling & AMI workshop, Grenoble, 2010

Data

Evolution of the total space (PB)

> 40 PB

137,287,872 file replicas

~ 800 grid end-points

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

$\mathcal{DDM}$ entries

- $1,623,356$ replicas
- $2,099,545$ datasets
- $820,265$ containers
- $334,017,070$ files
- Replica driven
- Archive
  - $19,379,050$ datasets
  - $587,891,321$ files

Meta-Data

- Only Meta-data relevant for data management activities
  - Placement, discovery, access, accounting, usage, deletion, history, consistency, etc.

- Information about data system objects
  - Dataset
    - Name, identifier, owner, timestamps,
    - Placement policy tags, location
  - File
    - Name, identifier, size, checksum, timestamps

Can we extend $\mathcal{DDM}$ with user meta-data ?

Limitations

Schema and APIs

- Conceived for a limited amount of arbitrary attributes
- Optimized for dataset/replica oriented queries
- Not flexible
- No generic attribute-oriented queries

Changes to add new meta-data

- APIs, tables modifications
- New indexes, queries
- Lot of Oracle tuning
- Random I/Os, large I/O, locks
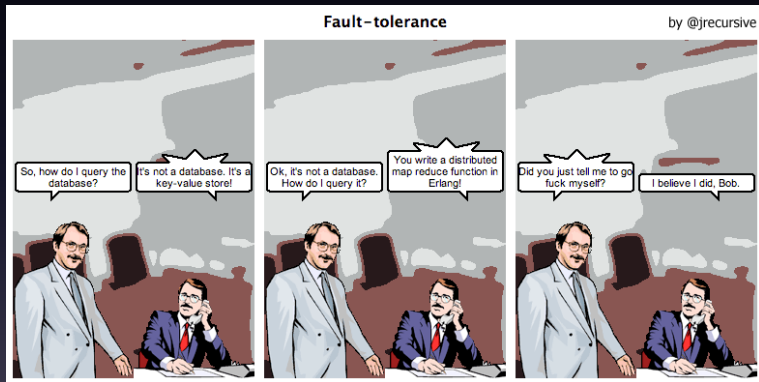
Example for new file meta-data

T_FILES

- GUID, LFN, size, checksum, timestamps
- Index Organized Table on GUID
- $314,441,787$ rows

Issues

- Extension might involve migration
- IOT → HEAP ?
  - Performance degradation
- New table ?
  - Consistency or performance issue
- Access by meta-data ?
  - Secondary indexes are expensive

Key-value store ?

Apache Cassandra [link]

http://howf**kedismydatabase.com/

$\Longrightarrow$ **Focus on a better integration with AMI**

Synchronization with AMI

AMI $\rightarrow \mathcal{DDM}$

- AMI metadata project & dataset naming convention
  - $\rightarrow$ $\mathcal{DDM}$ namespaces, permissions and ownerships

- Asynchronous collector

- Independent of external application performances

Synchronization with AMI

$\mathcal{DDM} \rightarrow$ AMI

- Events generated for each dataset registration and deletion

- Notifications based on HTTP

- Apache ActiveMQ [link] and STOMP protocol [link]

    - Pre-production

    - Robust and stable: transaction, fail-over, etc.

    - $277,406$ events sent in 123.360 seconds ($2200\,HZ$)

    - Support request to CERN/IT (LCG/MsgTutorial [link])