Utility of collecting metadata to manage a large scale conditions database in ATLAS

Elizabeth Gallas, Solveig Albrand,
Mikhail Borodin, and Andrea Formica

International Conference on Computing in High Energy and Nuclear Physics
October 14-18, 2013

Outline

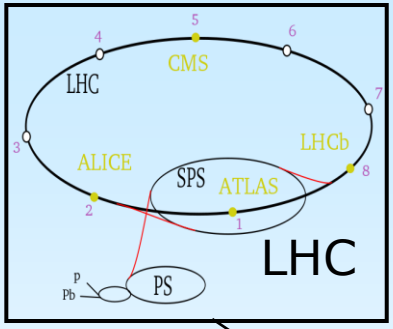
- Intro: “Conditions data”
- Motivation
- Goals
- Schema
- Data Sources & Cross Checks
- Amending content:
  - Connecting to other ATLAS systems
  - Adding useful metrics
- Interfaces
  - Browsing
  - Reporting
- Utility of the system during LS1 (current Long Shutdown)
  - In preparation for LHC Run 2
- Conclusions

This system is an extension of the ATLAS COMA system described at CHEP2012 (Run-level metadata)

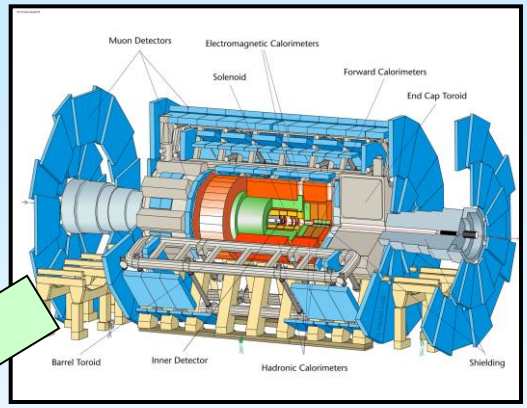
- with similar database and interface design principles

Ties between AMI (ATLAS Metadata Interface) with COMA have broadened into this new area of Conditions Data management

- As noted in the slides



“Conditions data”



Detector Control | Trigger DAQ | ... | Data Quality

ATLAS Conditions Database

- “Conditions”: general term for information which is not ‘event-wise’
  - reflecting the conditions or states of a system
  - valid for an interval ranging from very short to infinity
- ATLAS Conditions DB
  - Stores conditions data from a wide variety of subsystems which are needed at every stage of data taking, processing, analysis:
    - online calibrations, alignment, monitoring, to offline processing … more calibrations, further alignment … reprocessing … analysis …to final luminosity and data quality
  - Is based on LCG Conditions DB infrastructure using LCG ‘COOL’ API
    - Generic system which efficiently stores / delivers our data
  - Frontier makes that data is readily available for grid-wide access
- ATLAS exploits the wide variety of storage options available to optimize it for its content and for its use cases:
  - ‘inline’ payload (stored internally in the database tables): many data types
  - ‘reference’ payload (pointers to an external file or other table)

Motivation

The ATLAS Conditions database, by the end of LHC Run 1 is:

- **Large** (now many TB of data) & **Diverse** (65 active schemas)
  - **17 subsystems: 3 active instances** for LHC Run 1 in **2 domains**:
    (1) Simulation (2) Real Data replicated to Tier 1s (3) Real Data monitoring
    (1) Used Online (2) used for Offline processing (not used Online)
  - **>1400 Folders** (~database tables) in active schemas
    - Payload (columns): from 1 to 265; Many time larger volume variation;
  - **> 15000 Folder 'Tags'** (versions of conditions in IOV ranges)
  - **> 600 Global Tags** (collections of folder tags across schemas)
- Based on the LCG Conditions Database infrastructure: serves us well
  - Many methods for writing, reading the data (LCG COOL API)
    - Schema by Schema
  - Great for data taking, offline processing, monitoring
  - Very useful to have conditions from all systems in common infrastructure

But: Schema-specific access makes it difficult to

- Form an overview from a management/coordination perspective
- Find information without detailed subsystem-specific knowledge

And: The infrastructure does not easily allow us to

- Enhance content with ATLAS specific information and metrics
- Connect dynamically with other systems

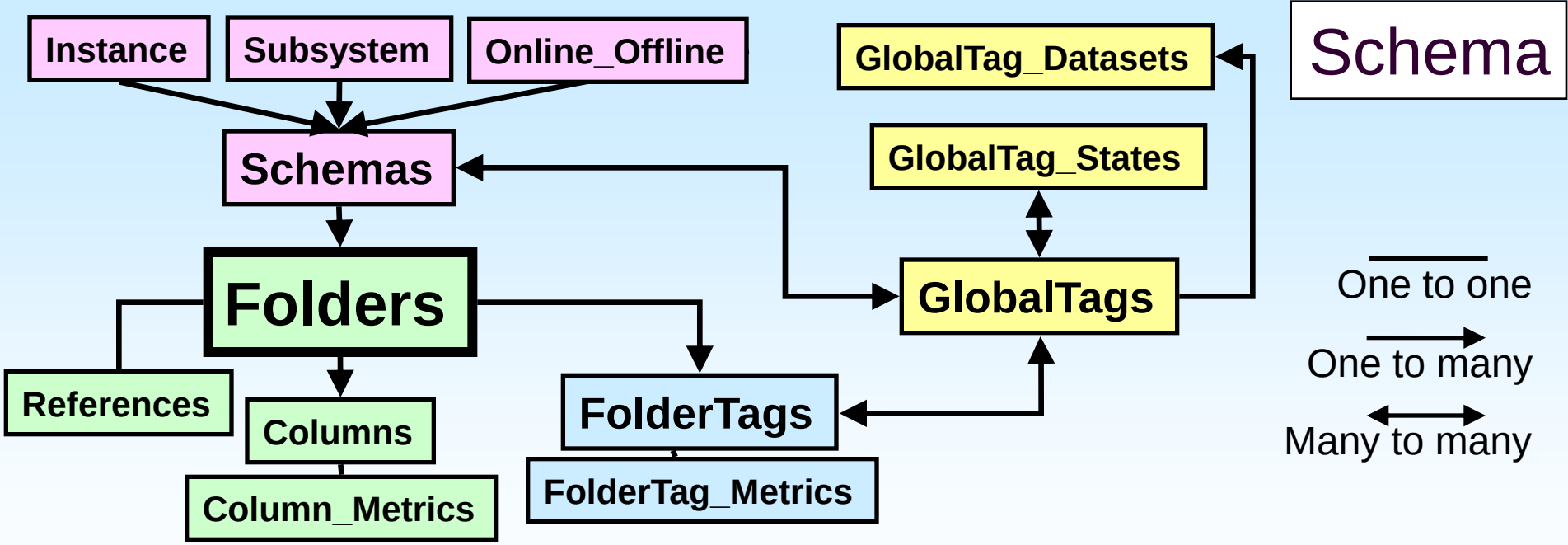
So: A dedicated repository has been developed to collect metadata on ATLAS Conditions Database structure to help fill the gap

Goals

- Enhance functionality: ATLAS COOL Tag Browser:
  - Please see the poster in this conference ! (#287)
    - "**A tool for Conditions Tag Management in ATLAS**"
      - A. Sharmazanashvili, G. Batiashvili, G. Gvaberidze
  - Opportunity for further extensions: Browsing the conditions data itself
- Collect structural metadata about content … Examples:
  - Channels, columns, rows, volume …which data changes most/least ?
  - Understand gaps in IOV coverage (gaps in conditions w/time)
  - Which folders use external references, their uniqueness
- Offer a global view of Conditions DB structure
  - Web-based Interfaces:
    - Browse: COOL structure using a variety of predicates
    - Report: Global Tag and Folder Reports
- Connect Conditions Data references to other ATLAS systems:
  - Which conditions are(/not) used in event-wise processing
    - Connect with AMI: ATLAS Metadata Interface (#260, this conference)
  - Which sets of conditions are "current", or in preparation "next"

→ Assist: general Conditions 'cleanup' during LS1 (current Long Shutdown)

→ In preparation for LHC Run 2 operations



- Database Design: Driven primarily by the Conditions DB structure
  - "Folder" centric: Folders represent Conditions DB tables
    - Each folder is owned by a specific Schema
      - Each has subsystem, instance, and if used offline or strictly online
    - Multi-version folders have one/more FolderTags
      - for conditions that allow different versions over time intervals
    - FolderTags may be included in one/more GlobalTags
      - When designated to be used in event-wise processing

Folders

Schemas

FolderTags

GlobalTags

- Database derived/enhanced content:
  - *_Metrics tables: structural metadata about Columns and FolderTags
  - GlobalTag_* tables: information from and/or for other ATLAS systems

Data Sources and Cross Checks

- Sources of the metadata include
  - The ATLAS Conditions Database itself
    - COOL API; Underlying database tables; Oracle dictionaries;
  - Derived content from the AMI database
    - Specific to each Global Tag
  - Expert entry from experts via an AMI entry interface
- Cross checks on source content finds inconsistencies and typos in Conditions DB definitions, sending email to experts to correct these issues. Examples of issues found:
  - Global Tag Descriptions and Lock Status
    - Stored schema-wise, must always be consistent schema to schema … and are occasionally found to be out of sync
  - Folder definition parsing
    - Folder definitions contain xml: must conform to set standards if those folders need to be accessed by Athena

Amending content: Connect w/external systems

Global Tag amended content:

- Usage in event processing
  - AMI Team: populates this table
    - Collect usage by dataset project name
      - Adding information like time range of offline processing

GlobalTag_Datasets

- State designations: Time varying as experiment evolves
  - States:

    Current: The best knowledge Global Tag for usage (domain dependent)

    Next: A Global Tag in preparation
  - State flavors depend on domain of usage:
    - Online data taking (HLT)
    - Express Stream processing (ES)
      - Quasi-real time processing of the latest data
    - Offline processing (no suffix)
      - All offline bulk data processing
  - Putting States into a database makes them available to external systems needing this information

    (moved away from AFS file system used previously)

    Thanks to AMI team for collaboration in developing the entry interface !

GlobalTag_States

Amending content with metrics

Why add Metrics (structural metadata) ? During LS1, based on Run 1 experience, we are acting to considerably clean up the Conditions DB structure and content … the metadata has been useful in many respects.

An example: Folder payload can be a “reference” to external files:

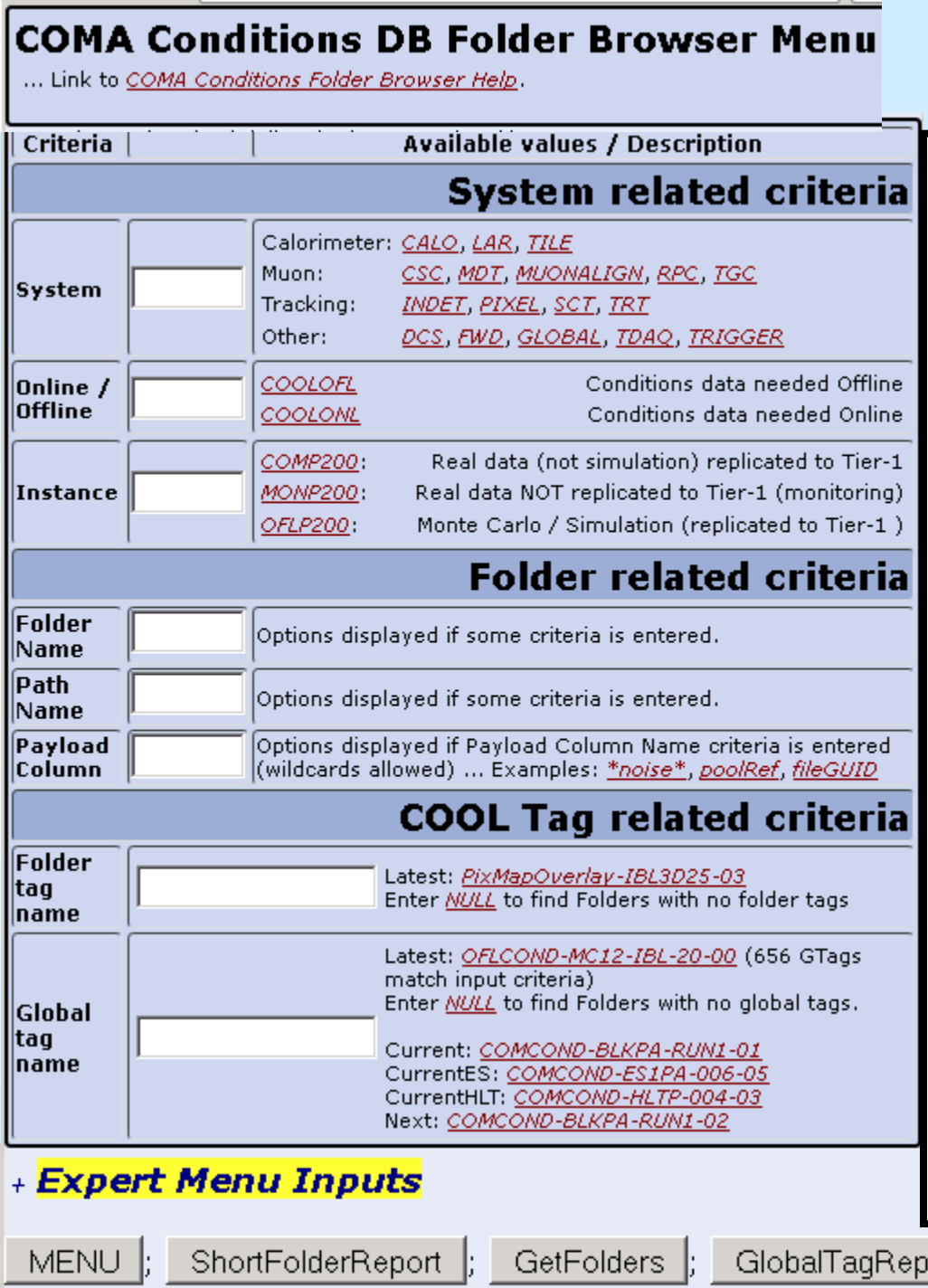
- But external files are problematic (Run 1 experience: ‘inline’ preferred):
  - Online: file movement around firewall is problematic
    - Requires special infrastructure, can cause delays
  - Offline: file movement on the grid
    - Files must be delivered to worker nodes for jobs on the grid
- LS1 directive: reduce/eliminate(?) external references
  - Using metadata: easy to identify at Coordination level:
    - folders using external references by subsystem (208 in 5 subsystems)
    - how many are used in current Global Tags (99 in the current GTag)
    - uniqueness of their content (some data did not change as anticipated)
  - Work with subsystems to evaluate/optimize storage:
    - Found: Sometimes good reasons for external files (volume/usage)
      - Decided: Keep these folders as they are for Run 2
    - Other times: Subsystems agree that ‘inline’ payload is better
      - Redefine these folders for Run 2: moving references to ‘inline’ content

Synchronization & Cross Checks

Keeping the metadata in sync with COOL is a challenge

- Real time sync is not possible:
  - COOL schema, content changes: not reported to external systems (infrastructure is not set up to do so)
    - Nor is that desirable: We would only want sets of changes only after completion of a set of changes or records added, not incrementally
- Currently, metadata is synchronized once per day, and on demand
  - The program requires about an hour to execute
    - Uses pyCOOL methods and
    - direct underlying table access for information not available/efficient via pyCOOL
- Work is ongoing to speed up the synchronization process while adding additional useful metrics as the system expands
  - Splitting program: fast (critical) / slower (less critical) parts
    - To execute the critical components more often
  - Employing a new API: a RESTful service (Java) in a JBoss server, which obtains new metrics through dedicated direct PL/SQL
    - not available via pyCOOL
- Under discussion: expansion of schema to include bookkeeping details of changes made by subsystem experts using ATLAS specific tools
  - These tools, generally in python, are outside the LCG infrastructure
  - They can add metadata content directly as experts execute them

Folder Browser menu



This dynamic menu interface

- Shows the variety of selection criteria available to find Folders or Tags of interest

Buttons (bottom) generate reports

- Enter criteria into textbox at left:
  - → Type manually or
  - → Click on options at right

  <return> or MENU ;

  re-generates Menu applying selection
- Choose GlobalTagReport ;

  for the Global Report
- Choose ShortFolderReport ; GetFolders ;

  for the Folder Report(s)
- More Expert Criteria available under
  - + Expert Menu Inputs

Global Tag Report (1)

COMA Conditions DB Global Tag Report

COOL Instance (cbi) : COMP200

295 global tags found meeting the input criteria.
Choose Global tag name of interest to generate the full report.

Global_Tag_Name	Lock Stat	Description	Create Date	Folder Tag Count	AMI Dataset Count	Processing Date Range	AMI Project (s)
COMCOND-BLKPA-RUN1-02 **Next**	0	New run-1 conditions BLK tag, based on COMCOND-BLKPA-RUN1-01 TWiki:ConditionsTagComCondBlkpaRun102	2013-Jun-26 16:07	266			
COMCOND-BLKPA-RUN1-01 **Current**	1	First run-1 conditions BLK tag, based on last BLK tag COMCOND-BLKPA-006-11 TWiki:ConditionsTagComCondBlkpaRun101	2013-Jun-11 09:22	266			
COMCOND-BLKPA-006-11	1	Based on COMCOND-BLKPA-006-10, new BULK tag with mu=0 conditions for run 200805 TWiki:ConditionsTagComCondBlkpa00611	2013-May-22 14:14	266	3352	2013-May-27: 2013-Sep-16	data12_8TeV
COMCOND-BLKPA-006-10	1	Based on COMCOND-BLKPA-006-09, new tag for 2011 2.76 TeV BLK repro TWiki:ConditionsTagComCondBlkpa00610	2013-Mar-11 15:52	266	336	2013-Mar-26: 2013-Apr-10	data11_2p76TeV
COMCOND-BLKPA-006-09	1	Based on COMCOND-BLKPA-006-08, new tag for 2011 2.76 TeV ES1 repro TWiki:ConditionsTagComCondBlkpa00609	2013-Mar-04 16:25	266	20	2013-Mar-19	data11_2p76TeV
COMCOND-BLKPA-006-08	1	Based on COMCOND-BLKPA-006-07, new BULK tag for 2013 partial reprocessings TWiki:ConditionsTagComCondBlkpa00608	2013-Feb-16 11:01	266	88	2013-Feb-23: 2013-Mar-07	data11_2p76TeV data13_hip
COMCOND-BLKPA-006-07		Based on COMCOND-BLKPA-006-06, new					
COMCOND-ES1PA-006-05 **CurrentES**							

- Multi-tag report: Generated when >1 Tag matches input criteria
  - Shows Tag States (Current, Next), Lock status, descriptions, link to TWiki, create date, folder tag counts, and which were used when in data processing (from AMI)

Global Tag Report (2)

COMA Conditions DB Global Tag Report

COOL global tag name (cbgt) : COMCOND-BLKPA-006-11

Global_Tag_Name	Lock Stat	Description	Create Date	Folder Tag Count	AMI Dataset Count	Processing Date Range	AMI Project (s)
COMCOND-BLKPA-006-11	1	Based on COMCOND-BLKPA-006-10, new BULK tag with mu=0 conditions for run 200805 TWiki:ConditionsTagComCondBlkpa00611	2013-May-22 14:14	266	3352	2013-May-27: 2013-Sep-16	data12_8TeV

Evolution of Global Tag State designations for this Global Tag:

GTag_State	Starting	Ending	Description	Created by (Modified)
Current	2013May22 18:31	2013Jun26 18:48	COMA initialize	mborodin
Next	2013May22 18:30	2013May22 18:31	COMA initialize	mborodin

Global Tag **COMCOND-BLKPA-006-11** includes 266 Folder Tags
A summary of folder tag count per subsystem is shown here.
Use links here to jump down this page to the folder tag detail.

System	SubSystem	Folder Tag Count	COOLOFL Count	COOLONL Count
Calorimeter	CALO	79	4	75
"	LAR	28	17	11
"	TILE	53	27	26
Muon	CSC	20	8	12
"	MDT	9	5	4
"	MUONALIGN	5	5	0
"	RPC	5	1	4
"	TGC	1	0	1
Other	FWD	1	1	0
"	GLOBAL	17	14	3
"	TRIGGER	1	1	0
Tracking	INDET	4	4	0
"	PIXEL	10	10	0
"	SCT	3	1	2
"	TRT	30	15	15

Single-Tag Report:

Summary section

- For this Global Tag
  Description, status, usage, …

Subsections show details:

1. Evolution of States
2. Processing details
   When >1 project uses it
   (in this case: only one)
3. Count Summary Table
   (266 Folders, Tags in this GTag)
   Showing counts per subsystem
4. Details of all Folders, Tags in this Global Tag
   Too much to show here.
   It appears below the counts.

Folder Report

COMA Conditions DB Folder Report

Tracking: COOLOFL_SCT/COMP200 /SCT/Derived/Monitoring

Self Link: /SCT/Derived/Monitoring

Node Description:

```
<timeStamp>run-lumi</timeStamp>
<addrHeader>
<address_header service_type="71" clid="1238547719" />
</addrHeader>
<typeName>CondAttrListCollection</typeName>
```

Documentation: twiki, LXR

Created: 2009-07-13_09:26

Last Modified: 2009-07-13_09:26

Node/CBF ID: 3 / 525

Channel Count: 4070

Payload Count: 8

IOV Base: run-lumi

Attribute Type: CondAttrListCollection

Versioning: 1 (Multi version folder)

Payload Columns:

Column	Type
SampleSize	Int32
BarrelEndcap	Int32
Layer	Int32
Eta	Int32
Phi	Int32
DefectType	String4k
Threshold	Float
DefectList	String4k

Folder Tags (12)

Folder Tag	GTag Count	Lock Stat	Created / Last Insert	Rows
SctDerivedMonitoring-NOMINAL	10	1	09Jul13_09:27:	0
SctDerivedMonitoring-UPD4-003	7	1	12Oct26_22:41: 13Feb14_23:10	281189
SctDerivedMonitoring-UPD4-002	4	1	12Apr18_15:12: 12Nov02_12:10	164132
SctDerivedMonitoring-UPD4-001	115	1	10Jun07_09:04: 11Oct31_01:11	60679
SctDerivedMonitoring-UPD1-001	49	1	10Apr16_16:42: 10Aug29_19:10	401
SctDerivedMonitoring-UPD4-000	8	1	10Feb11_12:19: 10Jun26_12:21	492
SctDerivedMonitoring-003-00	25	1	10Apr17_12:25: 10Jun06_09:10	28462
SctDerivedMonitoring-UPD1-000	0	0	10Feb05_12:37	129
SctDerivedMonitoring-002-00	16	1	09Oct23_09:56: 09Dec17_08:37	22945
SctDerivedMonitoring-001-00	0	1	09Jul13_09:28	5847
SctDerivedMonitoring-000-01	1	1	09Jul13_09:27: 09Jul13_09:28	18809
SctDerivedMonitoring-000-00	0	1	09Jul13_09:27	445

- Folder details include:
  - Folder type
  - Links to TWiki and code repository
  - channels
  - IOV basis
  - Payload column details
  - …
- Its Folder Tags
  - Lock status
  - Association to Global Tags
  - Dates: creation, last data insertion
  - Associated rows of data in this Tag
  - …

LS1 evolution

- LS1: an excellent time to assess where we are and envision how to best refine Conditions for Run 2, while retaining Run 1 processing capacity.
  - A major cleanup is underway based on extensive Run 1 experience:
    - Refining folder definitions
    - Consolidating Global Tags
- As of LS1:
  - 3 current 'active' instances of the ATLAS Conditions DB contain all data conditions utilized over the last ~5 years (including all of Run 1)
    - Considerable development/evolution over Run 1
      - Many Folders, Folder Tags, Global Tags are now obsolete !
    - ATLAS Global Tagging procedures have reached maturity
      - We now believe that a single Global Tag for data and MC, respectively, can be consolidated for any future Run 1 analysis (called "Best Knowledge" Tags)
- Going forward: We are preparing new instances for use in Run 2:
  - Highest volume tables which are active can start freshly in Run 2
    - Important for the performance of the underlying infrastructure
    - Leaving behind the obsolete Folders and content
  - Carry forward only the multi-version folders, tags needed for future processing of Run 1 (under the Best Knowledge Tags)
    - Leaving behind the obsolete Global and associated Folder Tags

→ The metadata system has been very useful in this consolidation process

Conclusions

- Metadata about the ATLAS Conditions DB structure has been aggregated into a dedicated system
  - It is part of a broader integrated ATLAS Metadata program sharing information/infrastructure:
    - AMI: Dataset-level metadata
    - COMA: Run-level metadata
      - → now extended into Conditions DB management (described here)
      - → Supplemental information: from AMI content and infrastructure
    - TAGs – Event-level metadata
      - COMA : TAG relationship well established: See CHEP 2012
- This system delivers unique data and services to experts and users
  - Fulfils all goals of slide 5
  - Ongoing work: refine and expand content and utility
    - Improve Conditions DB management and coherence generally
    - Further enhance functionality of the CTB (Cool Tag Browser): poster #287
- Every moderate/large scale experiment needs to efficiently store, access and manage Conditions-type data
  - when it grows in size and diversity:

    collecting metadata about its structure: useful in many respects