

## Database Architecture for the Calibration of ATLAS Monitored Drift Tube Chambers

**Elisabetta Vilucchi<sup>1</sup>**

*INFN Laboratori Nazionali di Frascati*

*Frascati 00044, Italy*

*E-mail: elisabetta.vilucchi@lnf.infn.it*

**Paolo Bagnaia**

*Università degli Studi di Roma "La Sapienza" & INFN Sezione di Roma*

*Roma 00185, Italy*

*E-mail: paolo.bagnaia@roma1.infn.it*

**Manuela Cirilli**

*University of Michigan*

*Ann Arbor, MI 48109-1040, USA*

*E-mail: Manuela.Cirilli@cern.ch*

**Paola Celio**

*Università di Roma Tre & INFN Sezione di Roma Tre*

*Roma 00164, Italy*

*E-mail: celio@roma3.infn.it*

**Gancho Dimitrov**

*Lawrence Berkeley National Laboratory*

*Berkeley CA 94720, USA*

*E-mail: Gancho.Dimitrov@cern.ch*

**Shawn McKee**

*University of Michigan*

*Ann Arbor, MI 48109-1040, USA*

*E-mail: smckee@umich.edu*

**Domizia Orestano**

*Università di Roma Tre & INFN Sezione di Roma Tre*

*Roma 00164, Italy*

*E-mail: Domizia.Orestano@cern.ch*

**Joe Rothberg**

*University of Washington*

*Seattle, WA 98195, USA*

*E-mail: rothberg@mail.cern.ch*

---

<sup>1</sup> Speaker

**Florbela Tique Aires Viegas***CERN**Geneva 1211, Switzerland**E-mail: florbela.tique.aires.viegas@cern.ch***Monica Verducci***CERN**Geneva 1211, Switzerland**E-mail: Monica.Verducci@cern.ch*

The size and complexity of LHC experiments raise unprecedented challenges not only in terms of detector design, construction and operation, but also in terms of software models and data persistency.

One of the more challenging tasks is the calibration of the 375000 Monitored Drift Tubes, that will be used as precision tracking detectors in the Muon Spectrometer of the ATLAS experiment. A high rate of muon tracks is needed to reach the design average resolution of 80 microns.

In this context, data suitable for MDT calibration will be extracted from the second level trigger and then streamed to three remote Tier-2 Calibration Centers. The Calibration sites will also need the ATLAS conditions data that are relevant for the calculation of MDT calibrations: part of the Conditions Database will thus be replicated at the remote sites via ORACLE streams.

At each centre, the computation of the actual calibration constants will be performed in several steps, including strict validation and data quality checks.

All information produced at every stage of the calibration procedure will be stored in local ORACLE Calibration databases that will be replicated to a central database located at CERN using ORACLE streams: this will allow each Calibration site to access the data produced by the others and to eventually provide back-up should one site become unavailable for any reason. The validated calibration constants will be extracted from the CERN Calibration DB and stored into the ATLAS Conditions database for subsequent use in reconstruction and data analysis.

This paper reviews the complex chain of databases envisaged to support the MDT Calibration and describes the actual status of the implementation and the tests that are being performed to ensure a smooth operation at the LHC start-up at the end of this year.

*XI International Workshop on Advanced Computing and Analysis Techniques in Physics Research  
Amsterdam, the Netherlands  
23-27 April, 2007*

## 1. Introduction

High pressure Monitored Drift Tube (MDT) [1] chambers will be used as precision tracking detectors in the Muon Spectrometer of the ATLAS experiment at the LHC at CERN [2]. An accurate knowledge of the space-time relation is needed to reach the design average resolution of 80 microns.

The model for the calibration of the MDT chambers defines the optimal number of muon tracks for a calibration, the procedures and the requirements to the calibration software. The model requires a substantial effort in data collection, software implementation and data management. The required statistics and calibration frequency cannot be met using the standard Atlas data-flow. An ad hoc mechanism allowing the extraction of a dedicated data stream [3] for the muon calibration has been proposed and is currently under test. Calibration data will be streamed to three remote Tier-2 Calibration Centres in Rome, Munich, Ann Arbor.

A dedicated software infrastructure [4], developed within the ATLAS offline framework (ATHENA), can host different algorithms devoted to the measurement of the time spectrum parameters, to the computation of the space-time relation, to the evaluation of the tube resolution. The output of these calibrations will have to be fed, after a validation process, to the conditions database accessed by the reconstruction. The calibration software is complemented by an ATHENA service providing the correct calibrations to the reconstruction.

## 2. Database architecture for MDT calibrations

The MDT Calibration procedure computes all the information needed by the offline reconstruction to correctly compute drift radii of muon hits in the spectrometer. This information, as any other quantity relevant for data analysis, must be stored in the central ATLAS Conditions Database, in order to be accessible within the ATHENA framework.

MDT Calibration jobs will also produce a sizeable amount of information (~50 MB/day) that is essential to evaluate the quality of the calibrations but is not needed by the offline reconstruction. The quality and stability of the individual tube parameters, as well as of the space-time relation, must be continuously monitored. It is important to note that quality checks cannot be performed by the ATLAS online monitoring: not only a high statistics is needed to reach the desired accuracy, but only hits associated to good tracks will have to be used to avoid being overwhelmed by the noise. Validation is therefore a crucial part of the MDT Calibration procedure, and all data needed for it must be accessible from quality checks programs that will typically run outside ATHENA.

A “private” MDT database (Calibration Database) is thus being implemented to store the complete calibration information. Validation procedures will make use of the additional parameters to ensure that the calibration constants have been correctly computed. Also, the newly produced constants will be compared to those from the previous data taking to decide whether the Conditions DB must be updated. The full information produced at every stage will be stored in local ORACLE Calibration databases that will be replicated via ORACLE streams to a central database located at CERN: this will allow each Calibration site to access the data

produced by the others and to eventually provide back-up should one site become unavailable for any reason. ORACLE Streams enable the propagation and management of data, transactions and events in a data stream from one database to another. A capture process retrieves change data extracted from the redo log, either by hot mining the online redo log or by mining archived log files, the capture process formats it into a Logical Change Record (LCR) and places it in a staging area for further processing (can intelligently filter LCRs based upon defined rules). Then change records get transferred to receiving end and, if the receiving end is down, they are stored in a queue; eventually change records are applied at receiving and by the apply process.

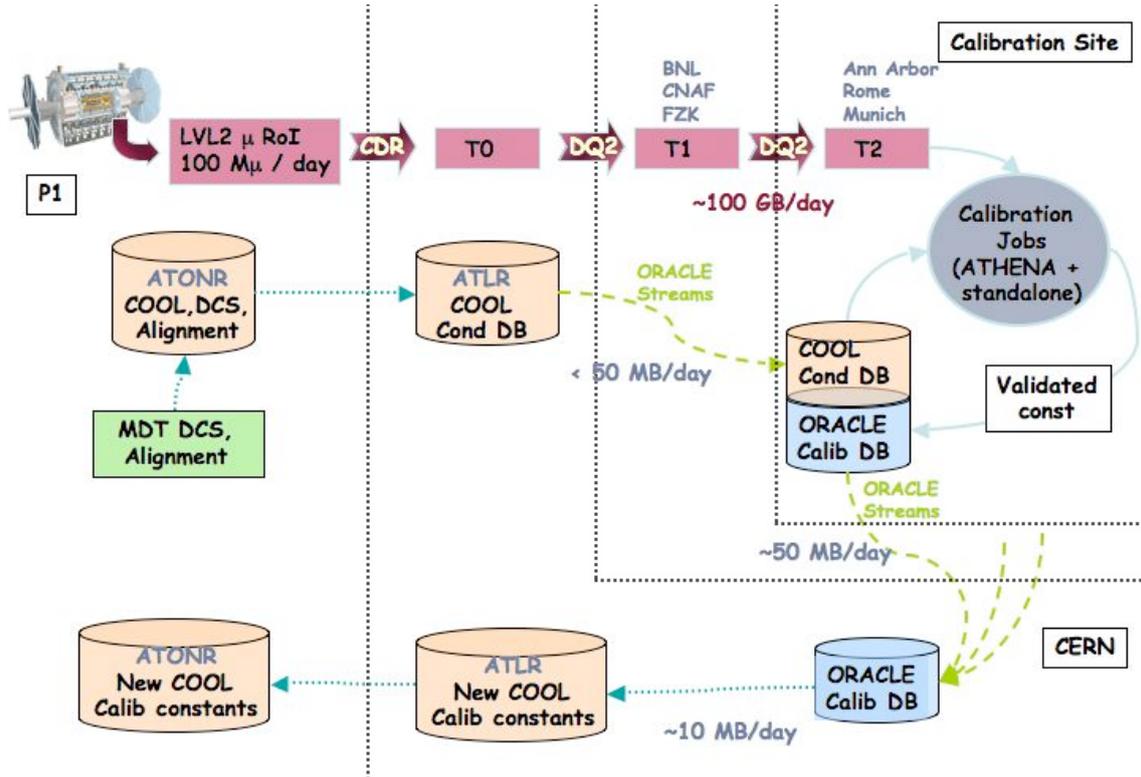


Figure 1: The calibration data flow. ATLR is the ATLAS Offline Database server sitting in the CERN Computing Centre. ATONR is the ATLAS Online Database Server, that sits behind a firewall at LHC Point1, where the detector is actually operating.

The validated calibration constants will be extracted from the CERN Calibration DB and stored into the ATLAS Conditions database for subsequent use in reconstruction and data analysis. This data management model has the major advantage that the Calibration DB is completely decoupled from the clients of the calibration and thus it can be modified without affecting the reconstruction; moreover, while the Conditions Database is optimized for reconstruction access, the Calibration Database is optimized for access by calibration and validation programs.

Finally, Calibration Centres will also need the ATLAS conditions data that are relevant for the calculation of MDT calibrations: alignment constants from the Muon Spectrometer optical alignment system, magnetic field and temperature maps, data quality information. Since Calibration Centres are Tier-2 Sites, the ATLAS Conditions DB is available through the standard distribution channels (SQLite files); however, through this standard distribution channels it is not possible to be notified in real time of changes in any of the ATLAS Conditions DB Table that are relevant for the calibration. For this reason, the Calibration Centres will profit of their local ORACLE servers to subscribe to the Conditions database tables that are needed by the calibration processes and get the data via ORACLE Streams, so that data will be automatically pushed from the Tier0 to the Tier2 whenever these tables are updated.

A general overview of the calibration data flow and database architecture can be seen in Figure 1, where also the data transfer between ATLAS Online and Offline Databases is highlighted. Calibration constants produced at the Centres will be written into the ATLAS Offline RAC Cluster (ATLR), which is installed in the CERN Computing Centre (Tier-0). The calibration constants will also have to be available for online usage (Online Monitoring, High Level Trigger), and so they will have to be replicated in the ATLAS Online RAC Cluster (ATONR), which is installed at LHC Point 1, where the detector is actually taking data.

Some replication tests between calibration sites and Cern have already been done; no issues of latency or bandwidth are expected because of the small amount of data that will be replicated and because in calibration sites there will be about 100 CPU working in asynchronous way, so constants will be produced, validated and replicated during all the 24 hours between the data taking and the reconstruction

### 3. MDT Calibration Conditions Database

The Conditions DB is being implemented using the COOL interface, developed in the LCG context.

COOL allows database applications to be written independently of the underlying database technology (ORACLE or MySQL). The use of COOL ensures long-term compatibility with the Calibration Software, since it is supported within ATHENA. The objects stored or referenced in COOL have an associated start and end time between which they are valid. COOL data are stored in folders, which are themselves arranged in a hierarchical structure of folder sets. Within each folder, several objects of the same type are stored, each with his own interval of validity range. COOL is optimized to store and retrieve objects associated to a given time.

To help with the storage of many objects of identical structure (e.g. chambers or tubes of the same chamber), objects in COOL folders can be optionally identified by a channel number (or channel ID) within the folder. Each channel has its own intervals of validity, but all channels can be dealt with together in bulk updates or retrieves. In the case of MDT Calibration, the channel ID can be a chamber or a sector of chambers of the Muon Spectrometer.

COOL implements each folder as a relational database table, with each stored object corresponding to a row in the table. COOL creates columns for the start and end times of each

object, and optionally the channel\_ID and tag if used. Several other columns are also created (e.g. insertion time and object ID), to be used internally by the COOL system, but these are generally of no concern to the user. The remaining payload columns are defined by the user when the table is created.

The MDT calibration data has been stored as an inline BLOB in the database, i.e. defining the payload to be a large character object (CLOB) which has internal structure invisible to the COOL database. COOL is then responsible only for storing and retrieving the CLOB, and its interpretation is up to the client (other possibilities to store data are: inline payload or reference payload). The CLOB appears as a sequence of characters, with a header and all the values separated by a comma; the header typically contains all the information needed by the reconstruction job.

The retrieving and storing of the data inside the Athena framework is possible using the IoVService, an interface with the COOL tables via IoV range, and a packing and unpacking algorithm of the data (currently read from ascii file). The IoVService permits to access to the right CLOB data in time with the event, and to retrieve the payload, the unpacking algorithm changes the format of the CLOB structure, obtaining the calibration constants and finally stores them in the Transient Data Store TDS to be used by the reconstruction algorithms via the Calibration Service.

The usage of COOL within the MDT Calibration framework has been tested and the code has been finalized. Calibration data have been stored as an inline large character object (CLOB), which has internal structure invisible to the COOL database. COOL is then responsible only for storing and retrieving the CLOB, while its interpretation is up to the client.

#### 4. MDT Calibration Database

Since the validation procedures will run outside the ATHENA framework, there is no obligation to use COOL. The main reason for choosing a different database implementation is that COOL, being a high-level API, does not possess the complete relational features of an ORACLE or MySQL database.

The MDT Calibration schema is the same in the databases at the calibration centers and at Cern where they merge in. At the moment it is made of ten tables, nine triggers, seven sequences, two procedures and some materialized views.

Calibration constants are distributed in the tables as follows.

The SITES table has one single field containing the initials of calibration sites (Roma, Ann Arbor, Munich) and it is used to merge the three databases in one DB at Cern.

The MDT\_HEAD table is the table of the calibrations. It contains the calibration date, the site name, the total amount of data used, the link to the histogram (and ntuple) files, etc.... It is used as a header, to identify a single "calibration job", i.e. about a day of normal data-taking; it references the SITES table, so, any calibration is uniquely associated to the calibration centre where has been done.

The MDT\_TUBE table holds the calibration constants relative to a single tube, like the " $t_0$ " (measured time for a muon which hits the tube wire), the parameters of the fit to the single tube time spectrum, useful for DQ estimates and efficiencies, the parameters of the ADC (charge

collected by the wire) spectrum, necessary to equalize the ADC of different tubes and the full correlation matrix on these quantities, as well as the primary key of the MDT\_HEAD, so it references the header table.

The MDT\_RT table holds information on the "r-t relation" (from measured drift time to position in space) of one calibration region, roughly equivalent to one (or half) single chamber. Then it contains the amount of events used, some parameters of the calibration region (temperature, mag. field, high voltage, etc...) and the data quality of the calibration region. Moreover it also references the header table. The rt function itself, together with its ancillary functions (its inverse  $t \rightarrow r$ , i.e. the resolution as a function of  $r$ ), may be stored in different formats and resides in other tables. Therefore, for each calibration region, at least one of the following tables must be present: the MDT\_RT\_CHEBY table that contains the  $rt$ ,  $tr$  and resolution, parameterized as a sum of Chebychev polynomials, and the MDT\_RT\_MAP\_R, MDT\_RT\_MAP\_S, MDT\_RT\_MAP\_T tables that contain the  $rt$ ,  $tr$  and resolution, parameterized as a map of points; and any of these four table must reference the MDT\_RT\_TABLE (see Figure 2).

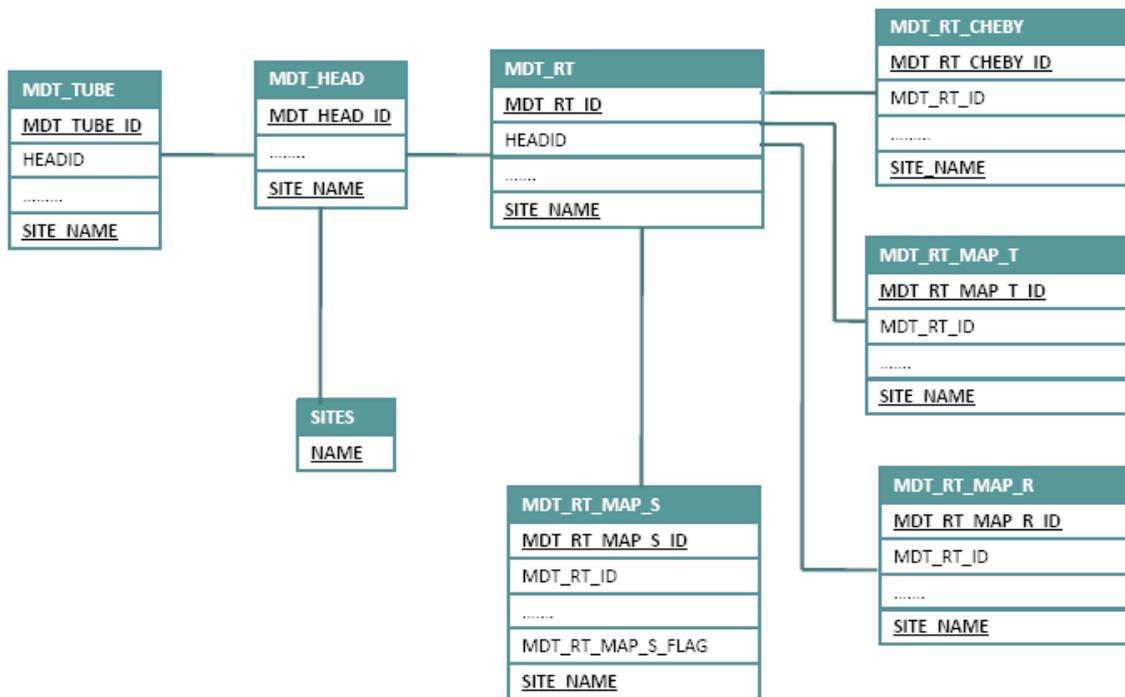


Figure 2. Calibration database tables

For any table, except SITES, the primary key is given by the couple of a unique ID, generated from a suitable sequence of sequential numbers, and the SITE\_NAME. In fact, for any table there is a sequence and a before-insert trigger such that at any insert time the trigger fills the ID field of the table with a term of the sequence. Since the tables are the same in the three databases, the ID values will be repeated, but since the primary key is given by the couple (ID, SITE\_NAME), then it remains unique even when databases merge in one database at Cern. Some check constraints have been created on any table to assure the correctness of inserted data.

To improve selects performances, some materialized views have been created with the queries made from the validation program. This program will even use calibrations of the last weeks to make some comparisons between channels. Then at least some weeks of calibrations will be maintained at each calibration center, while all the history will be stored at Cern database, that will even undertake a further backup functionality.

Any table has a validation flag field to signal if constants have even been submitted to validation procedure. In fact at the end of calibration computed constants are inserted into tables leaving the validation flag with the default value: 0, then materialized views are refreshed by two procedures (using also two further tables and two triggers), one run at the end of tubes calibration and another one run at the end of rt calibration. At this point validation programs can start reading constants from materialized views, then they perform all necessary checks and update, with a different from 0 value, the validation flag.

Once validation flags are updated, ORACLE Streams, with suitably configured rules, send new records, with different from 0 validation flag, to Cern Calibration database. The use of filtering rule for Streams avoid to send record to Cern twice: before and after validation; in this way, in fact, constants are send to Cern only once: after the running of validation process. In the end, using Coral, constants are written into Conditions database and, since they are only used by validation programs running at calibration sites, materialized views are not replicated at Cern database.

## 5. Conclusions

The MDT Calibration Software is fully integrated in the ATLAS offline framework. The parameters that must be stored in the Calibration DB have been defined, and validation procedures are being developed and tested on MDT test beam data.

Some tests have already been done on inserting and retrieving a massive amount of data, and on replicating data between calibration sites and Cern. These tests will be repeated in the next months within the Atlas Computing Data Challenge progressively including all elements of data flow to ensure a smooth operation at the LHC start-up.

## References

- [1] M.Livan, "Monitored drift tubes in ATLAS.", Nucl. Inst. And Meth. in Phys. Res A384 (1996) pp. 214.
- [2] ATLAS Muon Collaboration, "The ATLAS Muon Spectrometer Technical Design Report", CERN-LHCC/97-22, May 31 1997. Available online at <http://atlas.web.cern.ch/Atlas/GROUPS/MUON/TDR/Web/TDR.html>
- [3] E.Pasqualucci, S.Falciano, A. De Salvo, A.Di Mattia, F.Marzano, L.Luminari, D.Orestano, B.Martin, C.Meirosu, "Muon Calibration Data Extraction and Distribution for the ATLAS Experiment" presented at NuclearScience Symposium, Nuclear Science Symposium Conference Record, 2005 IEEE, Volume 1, Page(s): 830-833 (October 23th-29th, 2005 - Porto Rico)
- [4] M.Cirilli et al., "Conditions database and calibration software framework for ATLAS Monitored Drift Tube chambers", Prepared for CHEP 06. Electronic proceedings available at <http://indico.cern.ch>

POS (ACCAT) 031