Development of the CMS Databases and Interfaces for CMS Experiment

D.A. Oleinik, A.Sh. Petrosyan, R.N. Semenov, I.A. Filozova, V.V. Korenkov, E.G. Nikonov, E.A. Tikhonenko
Laboratory of Information Technologies, JINR

CMS DB activity was started many years ago by subdetectors groups. DBMS ORACLE was chosen as a preferred database technology. Data were stored in different databases. It was very difficult to support and update such heterogeneous system.

Our group started the work for CMS DB development in May 2004. A first stage of the work was collecting information about detector; investigation of data structure and data relations, analysis of end-user needs [1]. The considered subject domain is very extensive and complex. Informational requirements of the participants change, vary and extent permanently. It is difficult to consider all these factors in the varying conditions and create a static relational model for all the detector parts. The main task was to create the informational system allowing to describe the changeable and unspecified subject domain.

A second stage started at spring 2005. Storage system development of the information equipment began in spring too (equipment logs, tests results, unstructured information in the files). Usually, the typical decision for the databases creation is the relational approach application for the model description. But it is necessary to change the data structure because of the new objects input essentially in this case. Such an approach is impossible because of a great volume of different equipment. An abstract description method of the detector elements was applied for designing the databases structure [2]. Work has been carried out in cooperation with subdetectors developers. A new extensible database structure was developed. The first realization for Endcap Muon System database and initial filling of Endcap Hadronic Calorimeter database were done. Work has carried out in cooperation with subdetectors developers.

The developed informational system has such properties as flexibility, adaptability and efficiency. Abstract description method of the detector elements was applied for databases structure designing. Thus, the kernel of the informational system is Equipment database. It is located on the basic level of the databases structure. It is independent of the equipment unit physical model. DB supports tests results and accompanying documentation storing, new data input and informational search on many criterions. Information about equipment involves unstructured data such as document files, images, raw data and so on [3], [4].

The database structure considers the hierarchical relations (Fig.1). It gives the opportunity to store the information about communications between equipment elements.

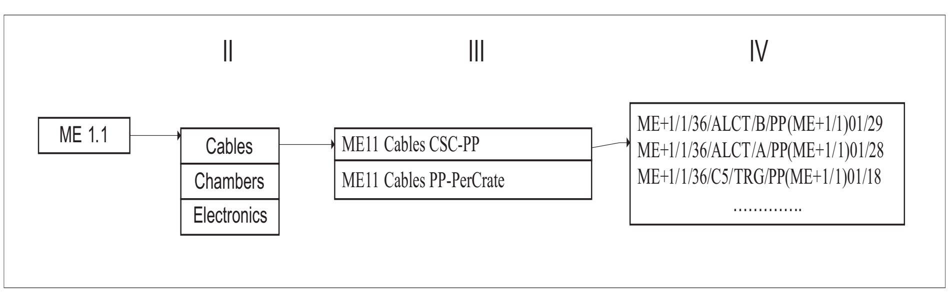

Fig. 1: Equipment Hierarchy Levels

At the group level: the general parameters set and values for all equipment of the given group are defined. For example:

Cabels: Chamber position [ME 1.1 position], Cable type [ALCTA, ALCTB...], Disk [YE+, YE-]

At the subgroup level: the characteristics set specific to the given subgroup is defined. For example:

Cabels: ME11 CABLES CSC-PP: Start Point (CSC), End Point (Patch panel), Length

At the equipment description level: communications with group attributes values are created, characteristics values are defined. For example:

Serial number: ME+1/1/14/LV/PP(ME+1/1)08/2
Cable type LV
Chamber ME+1/1/14
Disk YE+1
Start point (CSC) ME+1/1/14
End Point (Patch panel) PP(ME+1/1)08/2
Length (m) 2.25

The analysis of the requirements, objects of the considered subject domain has allowed one to define the full basic entities set of designed database. The database conceptual model was constructed as a result of the work. Such approach has provided an opportunity to store information on any equipment [6]. Elements have been defined by groups and subgroups. The general parameters with the fixed values set have been revealed for the each elements group. Specific attributes sets have been defined for each subgroup. Fig. 2 shows the general DB logical model.

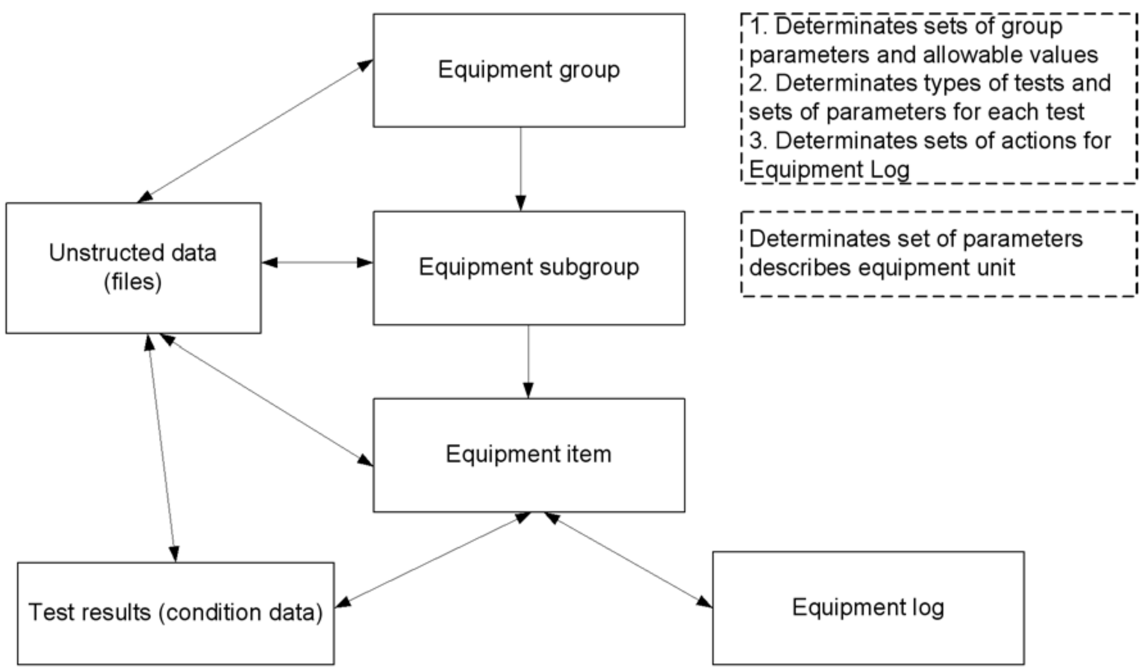

Fig. 2: DB Logical Model

In user view all information has a direct link to detector equipment: unstructured files; group parameters and values; equipment unit properties; equipment logs; tests results. The full Equipment DB structure holds all information sequences about detector equipment on the all hierarchy levels. The opportunity to support equipment logs (traveler lists) is realized. The events occurring with equipment elements are fixed in the traveler lists. Equipment log provides possibility of event registration like installation, maintenance, etc. List of events have been defined for each equipment group.

A lot of information is provided in an unstructured form:

- *Special raw data;*
- *Images;*
- *Files.*

These data set as files can be connected with each equipment element, group or subgroup. Files can be united in the folders. Tests results storage is realized. The tests techniques can be defined for the each equipment group. There are tested parameters sets for the each technique. Testing results are grouped in the revisions. The files group can be connected to the test in revision.

To manipulate this kind of data, a special subsystem connected with mass storage system was developed.

Different types of the interfaces to DB were realized:

- WEB interface,
- Application Program Interface (API) interface to conditions data,
- O2O interface.

Due to specificity of the DB scheme, API interfaces have been organized differently for different subdetectors. Now the API to HE calibration data is fully realized. API has based on OCCI (Oracle C++ Call Interface).

API can add, update, delete and select information. Selection can be done by different parameters:

- Tile code,
- Date,
- Date interval,
- Data processing operator name.

Web-interface provides the initial filling of database, different access levels for users, information search on different criterions, adding and updating data (Fig.3).

O2O (online to offline) — data transmission procedure from online (relational) DB to offline (POOL) database. Current realization of O2O code work has been done in CMSSW framework. OCCI interface is used for access to OnlineDB (Oracle), and access to OfflineDB is provided by PoolDBOutputService.

The storage system has realized (Fig.4).

It allows to be connected to DB by API interfaces and Web-interfaces. The reserve copy system operates.

Table 1 shows some actual statistics about database capacity, available disk storage and data volumes [5].

Current works are unification Muon CSC equipment and construction DB. Database will content all information about chambers, electronics, cables and other equipment; relations between equipment (tracking), test results.

The future plans are following:

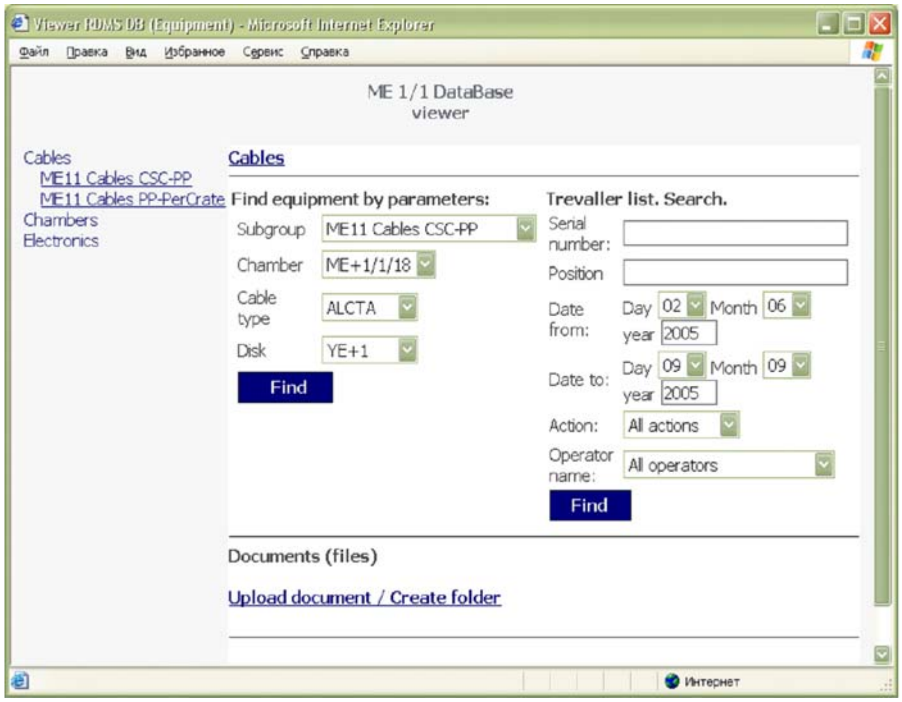

Fig. 3: Equipment Web-interface

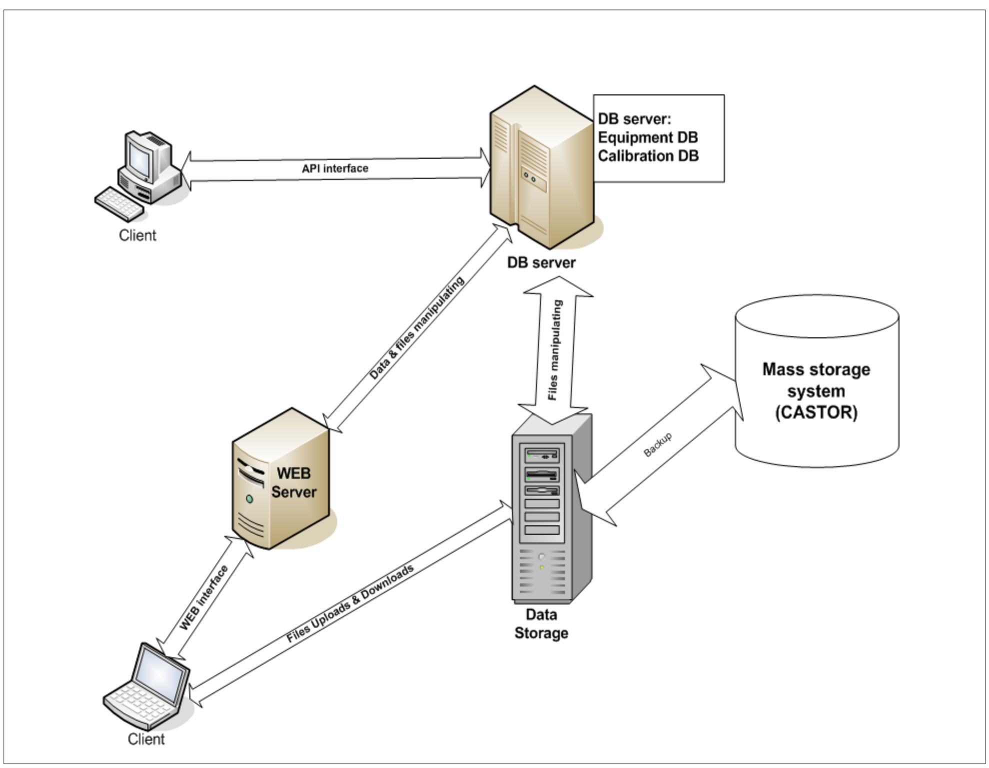

Fig. 4: Information System Scheme

Database	Current capacity: 1.2 GB
Disk storage	Capacity: 42 TB disk massive at JINR (dCache) 0.4 TB buffer storage at CERN + CASTOR resources
Data volumes	Database: 30,000 records ($\sim$ 500 MB) Raw files: 600 GB

Table 1: Statistics

- Positive feedback in pre-production operations prepared database for deployment in production,
- Software development for the updating chambers test results from files,
- Structure creation for the additional electronics information storing,
- API development for the access to the initial electronics calibration constants,
- Further DB filling,
- User interfaces improvement.

Expected growth of data volumes:

1. up to half million of the records (5 GB) in DB,
2. up to 6 TB in raw files.

References

[1] V.Ilyin, V.Gavrilov, I.Golutvin, O.Kodolova, V.Korenkov, S.Shmatov, E.Tikhonenko. RDMS Analysis Scenario & Data Bases & Data Flows. // 9th Annual RDMS CMS Collaboration Conference Minsk, 29 November 2004.

[2] D.A Oleinik, A.Sh. Petrosyan, R.N.Semenov, I.A. Filozova, V.V Korenkov, P.V. Moissenz, A. Vishnevskii, V. Karjavin. RDMS-CMS Data Bases: Current Status, Development and Plans // Nuclear Electronics & Computing (NEC'2005). XX International Symposium, Varna, Bulgaria, Sept.12-18, 2005: JINR, E10,11-2006-45. pp.216-221.

[3] A.Petrosyan, D.Oleynik, V.Smirnov, V.Korenkov, P.Moissenz, V.Karjavine, S.Movchan, N.Gorbynov, V.Perelygin, R.Semenov, I.Filozova, A.Oulianov, A.Vishnevsky. Database system for CMS RDMS sub detector information support // 10th RDMS CMS Conference, PNPI, September 13, 2005.

[4] D.Oleynik, A.Petrosyan, V.Korenkov, P.Moissenz, V.Karjavine, S.Movchan, N.Gorbynov, V.Perelygin. Design and realization of CMS sub detectors Database // 10th RDMS CMS Conference, PNPI, September 13, 2005.

[5] D.A Oleinik, A.Sh. Petrosyan, R.N.Semenov, I.A. Filozova, V.V. Korenkov, E.G. Nikonov, E.A. Tikhonenko. Development of the CMS Databases and Interfaces for CMS Experiment: Current Status and Future Plans // XXI International Symposium on Nuclear Electronics & Computing (NEC'2007), Varna, Bulgaria, Sept.10-17, Book of Abstracts, 2007 - Dubna: JINR , 2007. (JINR; E10,11-2007-119).

[6] Д. Олейник, А. Петросян, Р. Семенов, И. Филозова, В. Кореньков, Э. Никонов. Исследования и разработка баз данных для эксперимента CMS/LHC в рамках коллаборации RDMS // Scientific Report 2004-2005, Joint Institute for Nuclear Research Lab. of Information Technologies.