

Исследования и разработка баз данных для эксперимента CMS/LHC в рамках коллаборации RDMS

А.Ш. Петросян, Д.А. Олейник, Р.Н. Семенов, И.А. Филозова,
В.В. Кореньков, Э.Г. Никонов, Е.А. Тихоненко
Лаборатория информационных технологий, ОИЯИ

Одним из экспериментов на Большом адронном коллайдере (LHC) является установка CMS (Compact Muon Solenoid). Главными задачами участников экспериментов на установке CMS на подготовительном этапе является проектирование, построение и калибровка детекторов CMS, разработка специального программного обеспечения. ОИЯИ и российские физические исследовательские центры активно участвуют в коллаборации CMS в рамках единой организационной структуры RDMS (Russia and Dubna Member States). RDMS CMS объединяет российские институты и институты стран-участниц ОИЯИ на основе общей ответственности за определенные направления в исследованиях и разработке детекторов экспериментальной установки CMS. Данная структура была создана для оптимизации затрат и повышения общей эффективности участия членов RDMS CMS в коллаборации CMS. Основной задачей RDMS является проектирование, создание, ввод в эксплуатацию и эксплуатация наиболее важных следующих детекторов установки CMS таких, как:

- передний (торцевой) калориметр HE (Endcap Hadron Calorimeter),
- передняя (торцевая) мюонная станция ME1/1 (Endcap muon station),
- передний калориметр HF (Very forward calorimeter).

Одной из основных задач RDMS CMS и участвующего в данных работах сектора распределенных информационных систем и баз данных ЛИТ [1] в рамках построения и дальнейшей эксплуатации перечисленных выше подсистем является разработка и сопровождение необходимых информационных систем и баз данных.

Задача по разработке баз данных для установки CMS и его подсистем является широкой по охвату и требует комплексного анализа в силу сложности и масштабности самих физических экспериментальных установок. Информация, которая должна быть сохранена в соответствующих базах данных, включает в себя данные, необходимые для сопровождения проектно-конструкторских работ всего детектора и его подсистем, калибровочную и другую информацию, необходимую для эффективной эксплуатации установки на действующей стадии эксперимента. В состав коллаборации CMS входит 160 институтов, более 2000 ученых и инженеров. Всем участникам коллаборации необходимо обеспечить быстрый и эффективный доступ к постоянно меняющейся информации по работам в рамках данного проекта. Поскольку работа над установкой проходит достаточно интенсивно, необходимо учитывать постоянно меняющиеся требования участников к структуре сохраняемой информации и формам доступа к ней. Кроме того, разработчикам баз данных необходимо прогнозировать возможные информационные структуры и формы доступа к информации для будущих участников эксперимента.

Структура и назначение баз данных для эксперимента CMS

Разработана структура баз данных для эксперимента CMS и созданы соответствующие базы данных для подсистем HE, HF, ME1/1 [2]. Все физические параметры и

данные напрямую связаны с конструкцией детекторов установки. Поэтому на первом уровне структуры баз данных располагается **Equipment database**, в которой должна храниться структурированная информация обо всех составных частях детекторов и отдельных элементах оборудования. На следующем уровне структуры находится **Construction Database**, в которой должна храниться информация о связях и способах взаимодействия между различными узлами и подсистемами экспериментальной установки. Элементы установки могут быть связаны также со специальной информацией, имеющей отношение к результатам тестирования, условиям и результатам калибровок. Следующими уровнями являются **Configuration database** и **Conditions database**. Иерархическая структура перечисленных баз данных имеет следующий вид.

- **Equipment database** — содержит структурированную техническую информацию обо всех подсистемах детектора и отдельных узлах оборудования.
- **Construction databases** — содержит информацию о взаимосвязях между различными подсистемами и узлами оборудования.
- **Configuration database** — содержит информацию, необходимую для конфигурирования детектора для различных режимов эксплуатации.
- **Conditions database** — содержит информацию, о состоянии детекторов в момент времени сбора информации.

Различные базы данных требуют различных интерфейсов в зависимости от вариантов использования хранящейся в них информации. Например, необходимым условием для использования Equipment Database является наличие Web-интерфейса, поскольку информация, хранящаяся в этой базе данных, предназначена для использования участниками коллаборации, которые могут находиться в любом месте земного шара и иметь различные (иногда минимальные) условия для работы в Интернет. В то же время, для **Calibration Database** (специальная база данных по калибровкам) необходимым является программный интерфейс API, поскольку эта база данных, в основном, используется программами обработки.

В базе данных **Equipment database** хранится техническая информация для различных подсистем CMS детектора. **Equipment database** должна предоставлять возможность хранить структурированную информацию об элементах детекторов, при этом для обеспечения универсальности такая структура не должна быть напрямую связана с конструкторскими особенностями детектора и его элементов, система должна предоставлять возможность параметризованного поиска, ведения журналов. Еще одной особенностью **Equipment database** является возможность для ввода неструктурированных данных в различных форматах таких, как текстовые документы, чертежи и рисунки, необходимые “сырые” данные и т.п. Объемы данных, предполагаемые для хранения, могут достигать десятков гигабайт. При этом необходимо обеспечить высокую скорость и простоту доступа к данным, взаимодействие с другими системами хранения информации, возможность автоматического резервного копирования. Разработка и реализация **Equipment database** состояла из нескольких этапов. На этапе формулировки требований были осуществлены: сбор и накопление информации от конкретных разработчиков и участников эксперимента, классификация различных элементов оборудования, распределение по группам и подгруппам, описание свойств, определение взаимосвязей между различными элементами оборудования. Анализ требований, объектов рассматриваемой предметной области и их свойств позволило разработчикам определить полный набор основных

сущностей для проектируемой базы данных. В результате была построена концептуальная модель базы данных. Данный подход дал возможность хранить в базе данных информацию о любом элементе оборудования, независимо от его свойств и системной принадлежности. Каждая из подсистем ME1/1, HE и HF обладает своими особенностями, которые были учтены при разработке соответствующих баз данных [3], [4].

База данных подсистемы ME1/1

База данных ME1/1 является составной частью базы данных передних (торцевых) детекторов (Endcap Data Base). Для построения структуры базы данных ME1/1 применялось абстрактное описание элементов детектора, что позволит упростить создание аналогичных баз данных для других подсистем детектора. Элементы были разбиты на группы и подгруппы — для каждой группы элементов были выявлены общие параметры с фиксированным количеством значений, для каждой подгруппы были определены наборы специфичных признаков — значения этих признаков определялись для каждого элемента отдельно. Отдельным блоком выступают журналы оборудования. В журналах фиксируются события, происходящие с каждым элементом оборудования, для каждой группы оборудования определяется свой набор допустимых событий.

С каждым элементом оборудования, подгруппой или группой оборудования может быть связан набор неструктурированных данных в виде файлов, файлы могут объединяться в группы (папки).

В качестве отдельного блока реализована система хранения результатов тестов. Для каждой группы оборудования могут быть определены свои методики тестов, для каждой методики определяются наборы тестируемых параметров. Результаты тестирования группируются в ревизии. К каждому тесту в ревизии для каждого элемента оборудования может быть привязана группа файлов.

Общий подход к структуре системы хранения и баз данных для Endcap Data Base представлен на рисунке Рис. 1. Данная система хранения информации для ME1/1

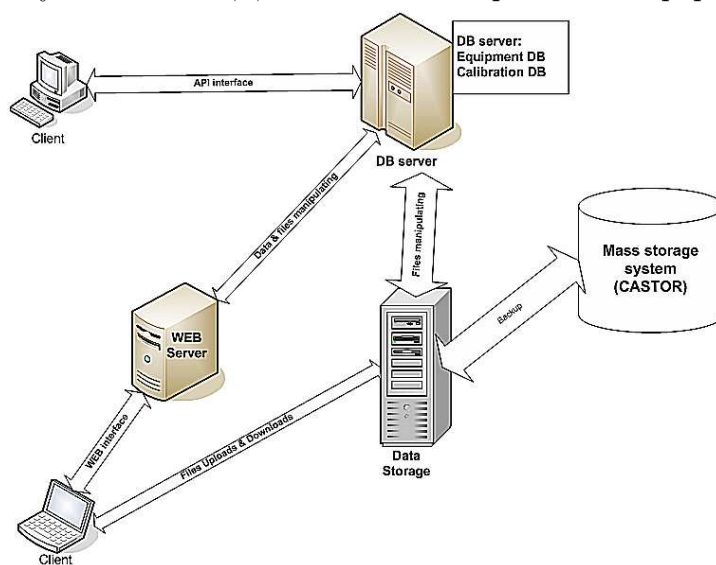


Рис. 1: Структура системы хранения и баз данных для передних (торцевых) детекторов (**Endcap**)

реализована в ОИЯИ и CERN одинаковым образом, дает возможность значительную часть работы по моделированию, тестированию и калибровке модуля ME1/1

проводить в ОИЯИ. Информационная система для модуля ME1/1 реализована на идентичном базовом программном обеспечении:

- Сервер баз данных — ORACLE.
- WEB сервер - Internet Information Server.

Web-интерфейс пользователя для базы данных ME1/1 обеспечивает начальное заполнение базы данных, отдельный доступ к данным, поиск информации по различным критериям, добавление и редактирование данных.

База данных подсистемы HE

База данных HE предназначена для хранения технических параметров различных элементов детектора: трубок, индекс-баров, ридаут-боксов, мегатайлов и др. и результатов калибровок радиоактивным, проволочным источниками, лазером, калибровок LED и RBX. Для каждого элемента хранится справочная информация, например для трубок: длина пластиковой трубки, длина металлической трубки, расчетная длина трубки, измеренная длина трубки, метки проблемных трубок.

База данных подсистемы HF

База данных HF содержит данные для каждого сегмента детектора HF: результаты различных калибровок и результатов сканирования поверхности сегментов.

План дальнейших работ

В ближайшие планы сотрудничества в рамках RDMS CMS входит продолжение работ по развитию и дальнейшей разработке баз данных подсистем CMS детектора, в частности:

- развитие Construction базы данных;
- дальнейшая интеграция Calibration и Equipment баз данных;
- дальнейшее развитие Web-интерфейсов для унификации доступа к другим специализированным базам данных;
- создание распределенной системы хранения данных.

Список литературы

- [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, <http://agenda.cern.ch/fullAgenda.php?ida=a044186>.
- [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. XX International Symposium on Nuclear Electronics & Computing (NEC'2005) Book of Abstracts, JINR; E10,11–2005–122, p.32–33.
- [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, <http://agenda.cern.ch/askArchive.php?base=agenda&categ=a052044&id=a052044s7t2/transparencies>.
- [4] Design and realization of CMS sub detectors Database. / *D.Oleynik, A.Petrosyan, V.Korenkov, P.Moissenz, V.Karjavine, S.Movchan, N.Gorbynov, V.Perelygin* // 10th RDMS CMS Conference, PNPI, September 13, 2005, <http://agenda.cern.ch/askArchive.php?base=agenda&categ=a052044&id=a052044s7t3/transparencies>.