Storage and Analysis System for Data Intensive High Energy Physics Applications

Maciej Malawski[1], Marek Wieczorek[1], Marian Bubak[1,2], Elżbieta Richter-Wąs[3,4,5]

1. Institute of Computer Science AGH, Kraków, Poland
2. ACC Cyfronet-AGH, Kraków, Poland
3. Institute of Nuclear Physics, Kraków, Poland
4. Institute of Computer Science, Jagiellonian University, Kraków, Poland
5. CERN, IT Division, Geneva, Switzerland

{bubak,malawski}@agh.edu.pl, Elzbieta.Richter-Was@cern.ch

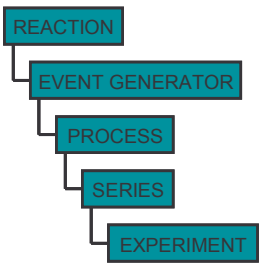
Why the Lhcmaster?

The High Energy Physics is one of the domains that deliver the highest number of great computational and database challenges in the world.

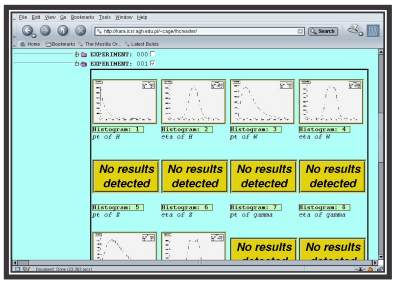
The simulators of detectors (e.g. ATLFAST) produce many hundreds of data files, what might mean many gigabytes of data that should be stored and reasonably managed. The experience gained in the INP (Institute of Nuclear Physics in Cracow) indicates, that not using of any specialized database system for storage of that data makes the work with the files very complicated, and leads to many unwanted mistakes. First of all, it is essential to represent every data file as a set of histograms, what might help a physicist recognize the file correctly. The Lhcmaster is first such system that fulfils all these requirements.

What is the Lhcmaster

The Lhcmaster is a system designed for storage and management of data being output of the ATLFAST detector simulator. The data stored in the system can be differentiated into two categories: the **data** (data files), and the **metadata** (database records and graphic files). The metadata is designed to support a logical order within the set of files.



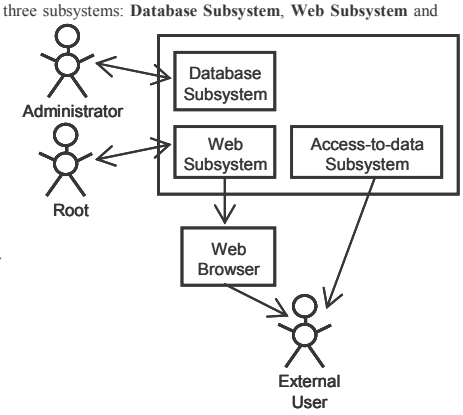
The data files are ordered in a hierarchical way.
There are 3 categories of data:
reaction - type of event represented by the file,
event generator - software simulator of the physical process (e.g., PYTHIA, HERWIG),
process - one of the ways out of which the simulator can produce the events corresponding to a given reaction.
Two additional categories: **experiment** (data file) and **series** (group of files).



Every data file is represented graphically by a set of histograms, so that a kind of graphic index of files is formed. **Graphic File Index (GFI)** is obtained during the Basic Data Analysis.

The system is composed of three subsystems: **Database Subsystem**, **Web Subsystem** and **Access-to-data Subsystem**.

3 types of system users:
Administrator - authorized user; creates, modifies and manages the data stored in the system; the Administrators are divided into groups
Root - superuser; create and removes the Administrators within the system, manage the groups of Administrators,
External User - unauthorized user; use the GFI to look through the files in system, gets the files from system.



Typical use case sequence in the system can be as follows:

1. Administrator stores some series of data files into the system with use of a command-line interface.
2. External User connects to the GFI by means of a web browser, and select the files he is interested in. As an output, text index of the files to be downloaded is generated.
3. External User uses the Access-to-data Subsystem to get the files he previously chose.

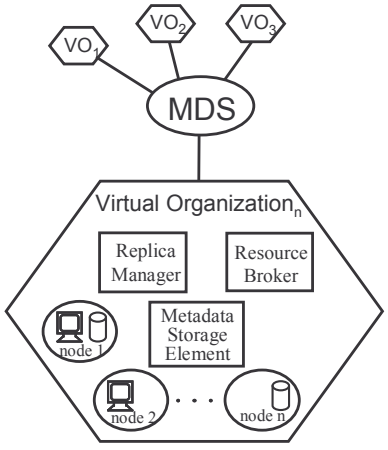
Towards the Grid: Lhcmaster-G

Even though the Lhcmaster fulfils all the requirements specified for it, there are several reasons that make us look for new solutions, to implement the system in a more flexible and scalable way:

- load on resources that is generated by the Lhcmaster is significant,
- the local file transfer should be replaced in the future by the network transport protocols (e.g., GridFTP), to make the system more scalable,
- at least two new functionalities might be added to the system: production of data within the system, and user own-designed analysis (apart from the Basic Data Analysis available now).

All these functionalities require a new distributed and powerful environment to be implemented in. We chose the Grid technology as the most proper for the system. We proposed also the new system - **Lhcmaster-G**, which will be a Grid-enabled system equivalent to the Lhcmaster, and moreover will solve the problems mentioned above.

The Lhcmaster-G grid will be composed of the elements of several types:
Storage Elements (SEs) and **Computing Elements** (CEs) will provide basic system functionalities,
Resource Broker (RB) and **Replica Management (RM)** will provide typical Grid functionalities (job management and data management, respectively),
Metadata Directory Service (MDS) - a central point of the Lhcmaster-G grid; will store both persistent information about the Grid structure, and the variable information gathered during the system operation.
Metadata Storage Element (MSE).



The MSE will be an equivalent of the Web Subsystem, but it will also store the database records within a singe Virtual Organization (generally speaking: it will store the whole metadata within the system). Aside from the user interface, the MSE will also provide a programming interface for user-design analysis. In order to provide a flexible interface possible to apply either for a user or for a program, the Web Services architecture will be applied for the MSE.

Information Society Technologies