Monte-Carlo Events Data Base

P. Bartalini, L. Dudko, A. Kryukov, A. Sherstnev,
I. Seluzhenkov, A. Vologdin

Outline

- Introduction
- CMS MCDB
- Main features of LCG MCDB
- General structure of LCG MCDB
- Present status of LCG MCDB
- Plans and milestones

Introduction

- Correct MC simulation of complicated processes requires sophisticated expertise and expert knowledge
- Therefore, a physics group in an experimental collaboration requests experts and/or authors of MC generators to create MC samples for particular process.
- It is common, that the same physics process is investigated by various physics groups in need of the same MC event samples.

The main motivation behind the Monte-Carlo Data Base (MCDB) project is to make MC event samples, as prepared by experts, available for various physics groups.

CMS MCDB

http://cmsdoc.cern.ch/cms/generators/mcdb

- Based on web technologies
- stores parton level events with a standard interface to the next level of simulation (based on Les Houches Accord I)
- Stores detailed documentation for each set of event samples
- Provide rapid communication between users and experts possible via its web pages
- Divided in two zones:
  - a public area for all users where they can find all necessary information about available event samples and download the corresponding event files.
  - a restricted are for authors, where experts can change the content of MCDB dynamically, e.g. upload and document event files for new processes, change the exist files and reply to user's comments.
- The MCDB requires users to reference the authors of event sample in case the events are used in a physics analysis.

CMS MCDB, problems

- CMS MCDB is designed to store parton level events. This implies that the size of event files should not be too large (typically smaller then 100 Mb).
- The expected number of physics processes in the database is several hundred.
- CMS MCDB is not a SQL database, therefore its search engine can not serve complex queries, only keyword phonetic searches.

We can expect that in a few years users will request a more powerful MCDB without the restrictions above.

LCG MCDB

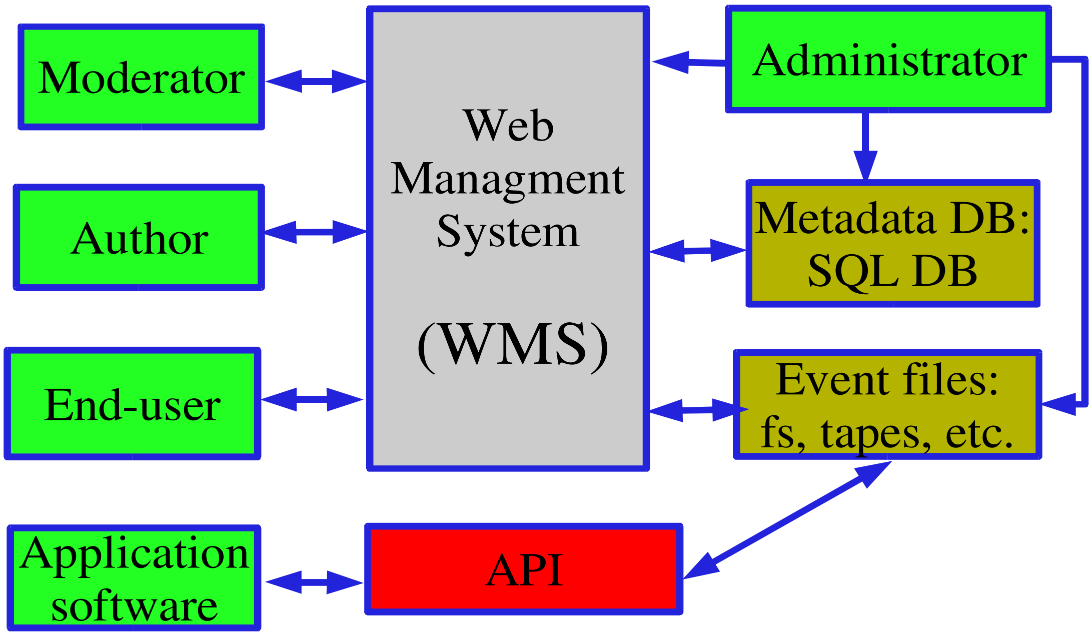
LCG MCDB will use ideas and experience of CMS MCDB
Main new features in LCG MCDB are the following:

- it is based on SQL database (MySQL)
  therefore it is possible to keep information at high level of structurization and serve the sophisticated search queries
- it will be possible to keep events at the particle level as well as at the parton level, also one can keep some other types of events
- new interfaces for the event files uploading
- direct programing interface from the collaboration software to LCG MCDB

General structure of LCG MCDB:

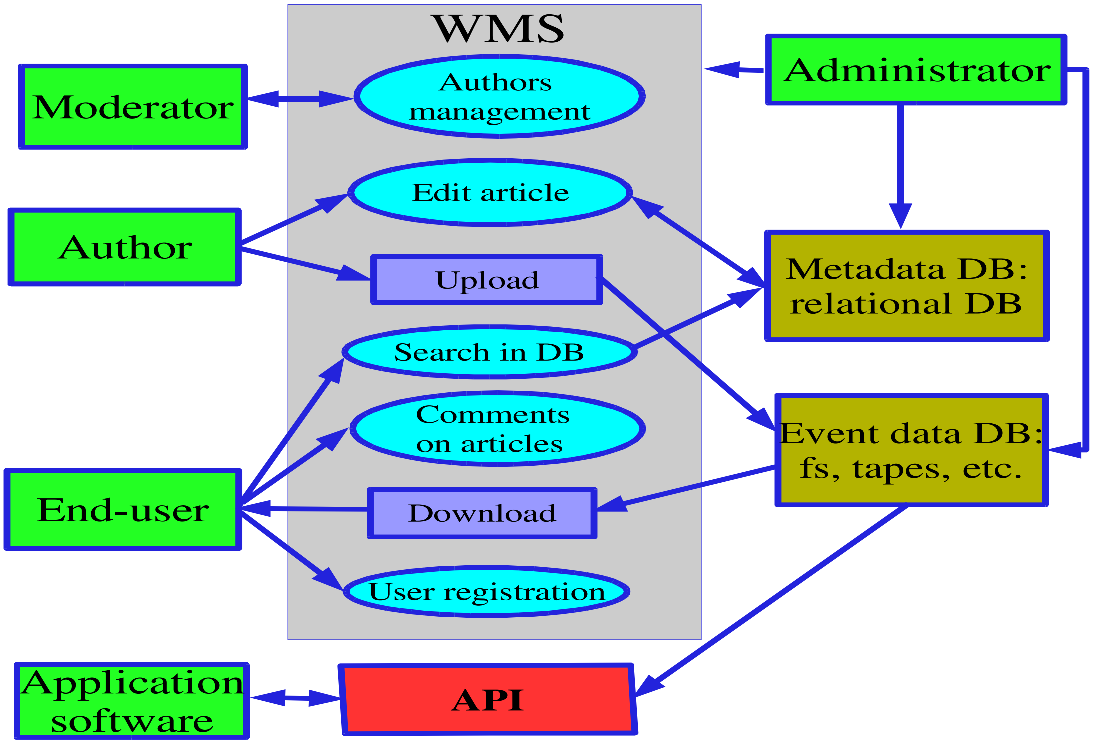
WMS; Data Base; API.

Authors, End Users, Administrator



- MC experts (Authors) provide files and its description;
- End Users can find necessary file, download and discuss it;
- API library provide direct way for collaboration software to download and process particular MC sample.

Development of LCG MCDB:



- Split each part of MCDB to independent sub-blocks
- Document input and output for each subblock
- Publish LCG Note with detailed description of MCDB
- Share design of subblocks between people who involved in the project

Present status of the project

- CMS MCDB is described in the proceedings of Les Houches MC workshop 2003
- LCG MCDB is described in LCG Note (CERN-LCGAPP-2004-01) the current version is attached to my talk on the agenda server
- During the development process we plan to write the technical LCG Note with detailed description of the software.
- The working prototype is designed at the MSU server at present. We plan to move this project to CERN server and CERN CVS to involve more people to the developing process.

Plans and Milestones (time scale can be decreased)

- 2004 year. Evaluation of SQL DB, WMS software
  - development of scripts for meta-data management and authorization blocks
  - design of structure and connections of SQL tables in MCDB
  - design of MCDB web site
  - development of blocks to manage author information, comments and block to post articles
  - testing the first release of LCG MCDB
- 2005 year.
  - Development of API for application (collaboration) software
  - Development of log system and templates management block
  - Development of new uploading methods
  - Adaptation of LCG MCDB to POOL and distributed storage systems
  - Adaptation of LCG MCDB to standard event format (HEPML) and developing of parsing software.