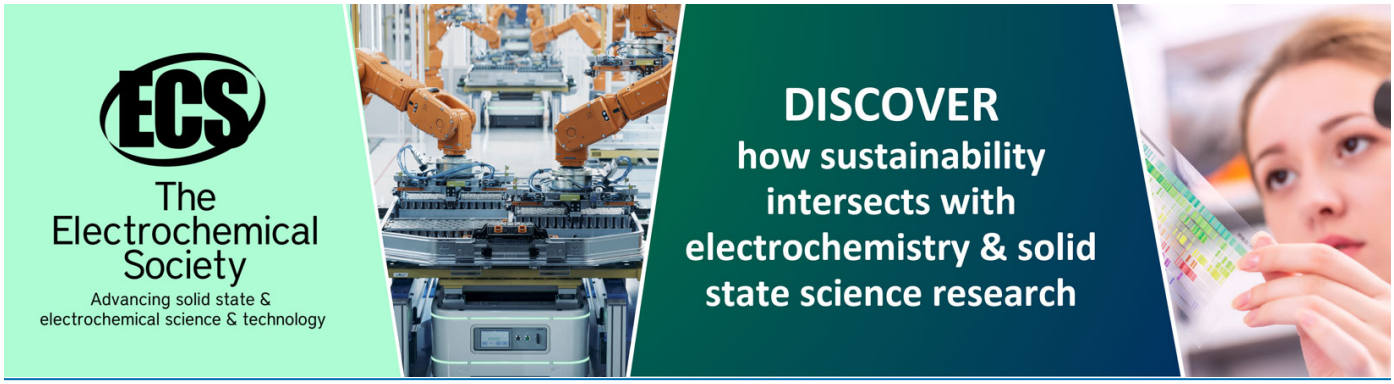
New developments in Generator Services project



Journal of Physics: Conference Series **331** (2011) 032025 doi:10.1088/1742-6596/331/3/032025

New developments in Generator Services project

A Karneyeu[2], M Kirsanov[2], D Konstantinov[3], W Pokorski[1], A Ribon[1], A Ryabov[3] and O Zenin[3]

[1] CERN, Geneva, Switzerland
[2] INR, Moscow, Russia
[3] IHEP, Protvino, Russia

E-mail: anton.karneyeu@cern.ch

Abstract.

The LCG Generator Services project provides validated, LCG compliant Monte Carlo generators code for both the theoretical and experimental communities at the LHC. In this paper we present the recent developments and the future plans of the project. We report on the current status of the generators repository, the new Autotools-based build system, as well as the new installation tools to create mirrors of the repository. We discuss new developments in testing and physics validation procedures in particular the use of HepMC Analysis Tool, as well as the Rivet validation tool. We also present a new activity, enlarging the scope of the Generator Services project, it is the involvement in the tuning of the Monte Carlo generators. This work, being essential for the understanding of the future LHC data, is now starting with the involvement of all the LHC experiments.

1. Introduction

The goal of the Generator Services project is to prepare validated LHC Computing Grid (LCG) compliant Monte Carlo (MC) generators code for both the theoretical and experimental communities at the LHC. It collaborates with the MC generators authors as well as with the LHC experiments and it is part of the LCG Simulation Project.

The project is composed of a number of work-packages, some of which are developed independently, and some are interconnected. The first work-package which we would like to mention here is the MC Generators Repository (GENSER) which was the initial motivation for establishing the project. Together with the repository comes the testing and the validation activity, as well as the user support activity which are, of course, strongly related to the installation of the generators. The two remaining work-packages are the HepMC event record and the Monte Carlo Event Database (MCDB).

In this paper we will review the status and the new developments in each of the work-packages. We start with reporting on the current status of the generators repository, the new Autotools-based build system, as the new installation tools to create mirrors of the repository. We discuss a number of new developments in the area of validation, in particular the use of HepMC Analysis Tool [2] and Rivet [1]. We then move to the other work-packages, namely HepMC [4] and MCDB [6], where we report on the current development. We conclude with a brief description of the new activity related to Monte Carlo generators tuning.

2. Generators Repository and Testing

The GENSER generators repository is used by the LHC experiments in the MC productions. The structure has been stable for a number of years, and the new versions of the supported generators are regularly installed (see Figure 1). At present, the repository contains 31 generators out of which 28 are covered by regression tests. The repository also contains 6 auxiliary packages which serve the purpose of testing and validation. The currently supported platforms are SLC4, SLC5, Win32 and MacOSX 10.6.

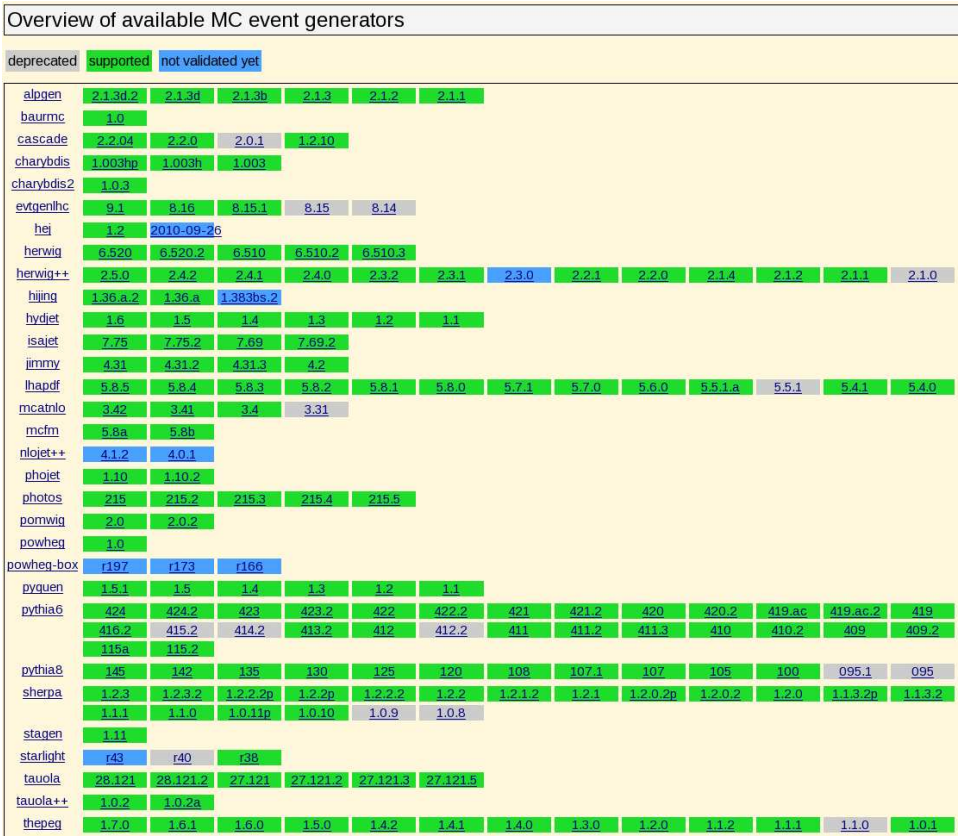


Figure 1. MC Generators repository.

The major development that has happened over the last year was to switch to Autotools as a default generator build system. We have provided Autotools-based build infrastructure for the generators which come without any specific build systems (like Pythia6 or Herwig). We have also provided feedback and we have collaborated in developing Autotools-based build systems for some new generators (like Tauola++). The goal of this development is to provide build systems which are on one hand generic in use (following the "configure and make" approach) and on the other hand have features specific for the Monte Carlo generators users, which are not always the same as for general software C++ packages (installation in users' directories, automatic definition of Fortran compilers, etc).

Another development that has largely facilitated the maintenance of the MC generators repository is the move to PKGSRC tool as the package management system. The combination of Autotools and PKGSRC simplifies the maintenance of generators on multiple platforms and allows to automatically build the entire set of all the supported generators and their dependencies at some other sites (to create 'mirror' installation of GENSER).

The testing and the validation of the generators is an essential part of our project. This

area sees continuously new developments and improvements. The range of the tests goes from repository consistency tests (checking the presence of the libraries, etc) up to physics validation using tools like Rivet [1], HepMCAnalysis [2] or MCTester [3]. One can basically distinguish three levels of the tests:

- Level 0 - automatic check that all the libraries and tarballs are built
- Level 1 - regression tests, check for deviations of different observables (total cross sections, etc)
- Level 2 - check of physics distributions using Rivet, HepMCAnalysis tool and MCTester

On Figure 2 we see an example of the physics validation using Rivet. The experimental data is plotted against simulation done using two versions of Herwig++. Such a validation allows not only the comparison of the generators output to the data, but also shows any changes between different versions of the same generator.

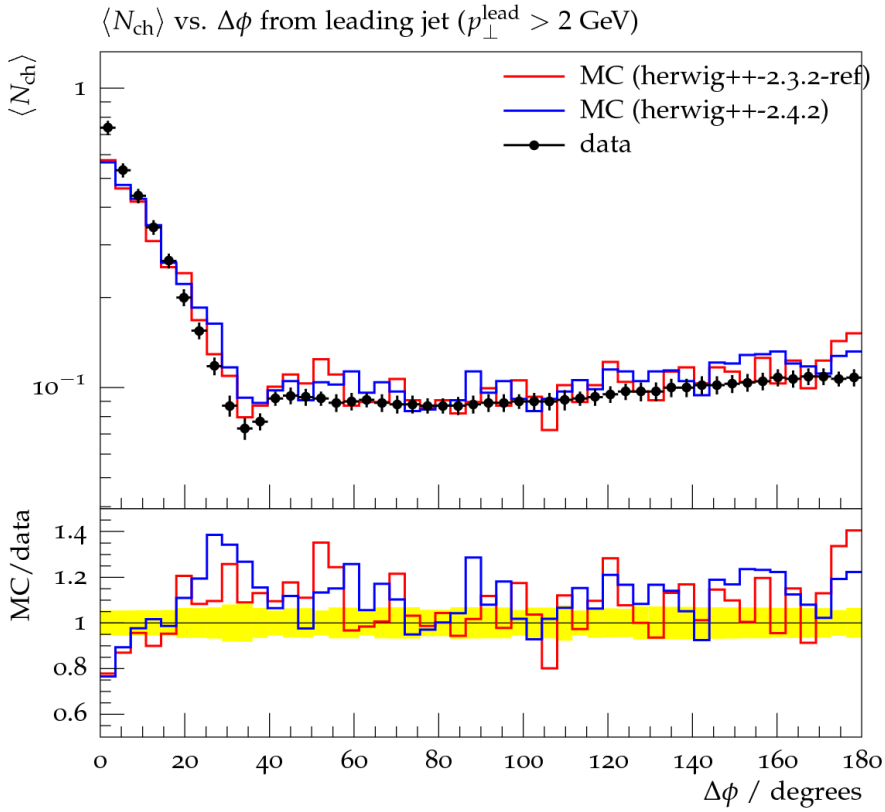


Figure 2. Rivet validation.

Figure 3 shows a snapshot of our web page containing some validation results using the HepMC Analysis tool. This tool is used to compare different versions of a generator for a specific physics process.

3. HepMC Event Record

The HepMC event record [4] has become a 'de facto' standard as far as the high-energy physics event records are concerned. It is primarily used as the interface between the event generators and the analysis and simulation frameworks. The current version of HepMC is 2.06. Some of the improvements with respect to the previous version are:

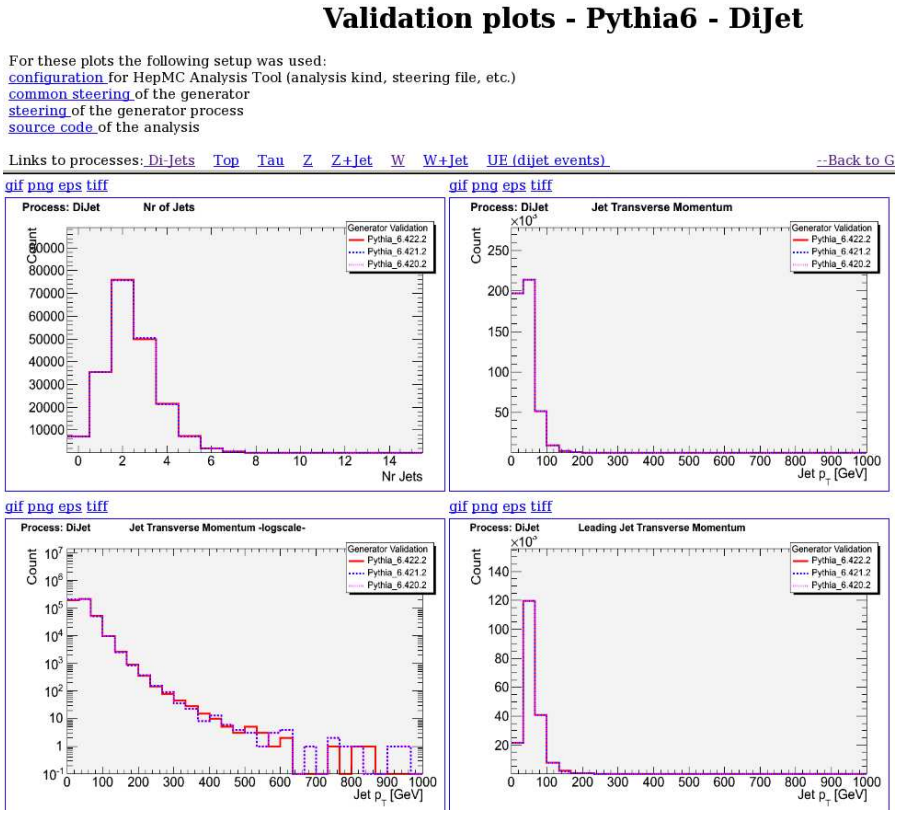


Figure 3. HepMC Analysis tool validation.

- removed deprecated ParticleData classes
- condensed Pythia and Herwig Fortran wrappers
- resolved output issues
- resolved iterator issues
- implemented default install directory structure
- allowed named weights
- provided a check to see if polarization has been defined
- improved "make check" not to complain when units are MeV

The detailed description of all the features of the latest release can be found on the HepMC web page [5].

4. Monte Carlo Event Database

MCDB is a database of Monte Carlo simulated events (for the list of MCDB authors see Reference [6]) used by the experiments in different simulation studies.

Currently it contains over 15000 parton level event samples. The database can be accessed through the web interface (see Figure 4) used both by the authors as well as the users. There are also mechanisms for automatic uploading and documenting the new LHEF files, as well as automatic access to the content of MCDB using dedicated C++ libraries. The interface to MCDB has been implemented in CMS software framework and it is regularly used in production. This database could potentially be also used by the other LHC experiments.

Recent developments in this area consist of the migration to the new hardware platform (maintained by the CERN IT services) and the implementation of the new features in MCDB

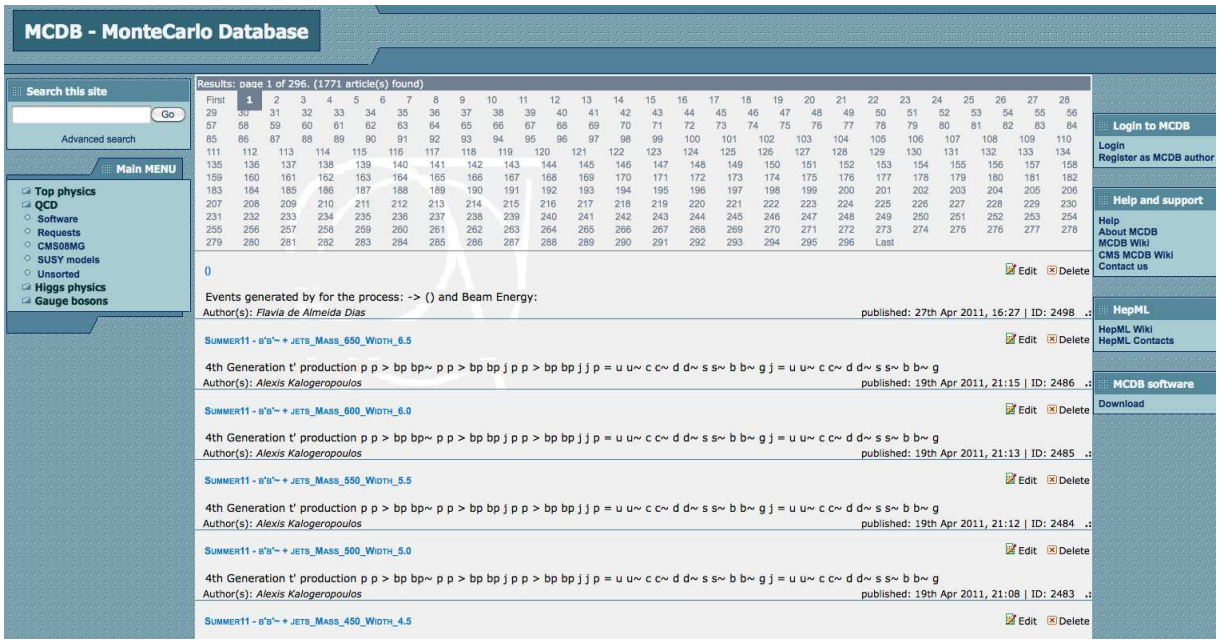


Figure 4. Monte Carlo Database (MCDB).

allowing to attach images (plots) to better describe the samples. This activity also covers the development of the unified XML-based description of parton level Monte Carlo events called HepML [7]. During the past year, a new version of HepML has been developed. It consists of a new XML schema and C++ libraries to write, read and modify HepML blocks. Interfaces to it have already been implemented in MCDB, CMSSW and CompHep.

The future plans for the MCDB work-package are to improve the GRID storage allowing easier and faster access to the samples, extend the event description with the new HepML features making it more detailed and improve the interfaces to allow more flexible access of the database.

5. Monte Carlo event generators tuning and validation web page

The Monte Carlo event generators tuning and validation web page (MCPLOTS) [8] development is a new activity that has been started in the first half of 2010. The purpose of it is to systematize the distributions of observables available from different generators, tunes and data sets. Its aim is also to provide a graphical front-end allowing users to browse easily through the relevant plots, and to organize the plots for easy reference (for an example see Figure 5).

The page is under continuous development and the content of it will, of course, be growing.

6. Conclusion

The LCG Generator Services project is mature and is used in production by the LHC experiments. The structure of the generators repository is stable and new versions are being continuously added. The recent developments include many improvements allowing easier maintenance of the repository, as well as extended validation of the generators.

The HepMC event record has become a 'de facto' standard for the high-energy physics Monte Carlo events. Its structure is stable and the new releases include mainly technical improvements.

The Monte Carlo database (MCDB) is used by CMS in large productions. The recent developments have improved its stability and reliability

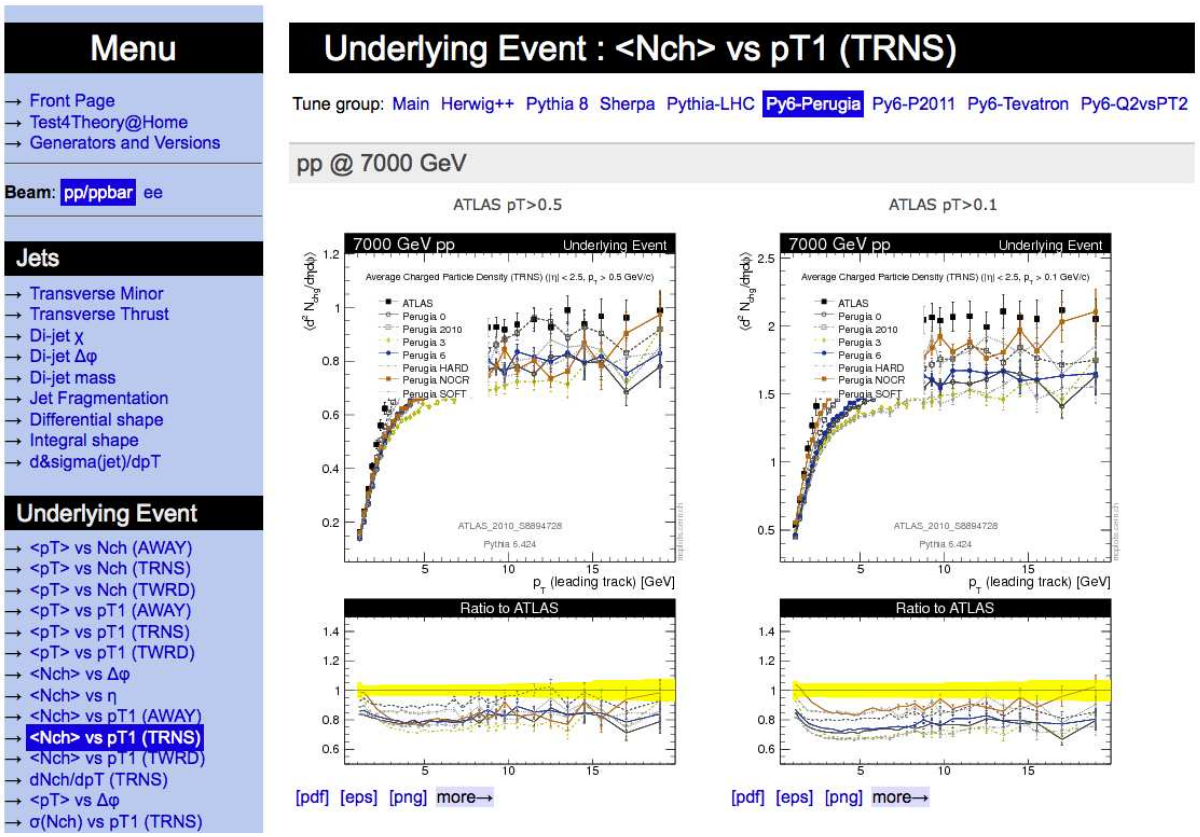


Figure 5. MCPlots tuning and validation page.

A new activity has been started in the scope of the project, namely the development of the Monte Carlo event generators tuning and validation web page. The importance of such a web page has particularly become evident with the appearance of the new LHC data.

References

[1] Buckley A, Butterworth J, Lonnblad L, Hoeth H, Monk J, Schulz H, von Seggern J E, Siegert F and Sonnenschein L 2010 Rivet user manual (MCnet/10/03) arXiv:1003.0694v5 [hep-ph]
[2] Ay C, Johnert S, Katzy J and Qin Z 2010 *J. Phys.: Conf. Ser.* **219** 032029 doi
[3] Golonka P, Pierzchala T, and Was Z 2004 *Comput. Phys. Commun.* **157** 39
[4] Dobbs M and Hansen J B 2004 *Comput. Phys. Commun.* **134** 41
[5] http://lcgapp.cern.ch/project/simu/HepMC/
[6] Belov S, Dudko L, Galkin E, Gusev A, Pokorski W and Sherstnev A 2008 *Comput. Phys. Commun.* **178** 222
[7] Belov S, Dudko L, Kekelidze D and Sherstnev A 2010 *Comput. Phys. Commun.* **181** 1758
[8] mcplots.cern.ch