

Update on CEDAR Projects

Jonathan Butterworth

University College London

MC4LHC Meeting

17-26 July CERN

Update on CEDAR Projects

- What is CEDAR?
- HepForge
- HepData
- HZTool, HZSteer, Rivet
- JetWeb
- HepML

What is CEDAR?

- Main aim is to couple validation tools for MC programs and other physics calculational tools with data.
 - JetWeb, HZTool, Rivet for validation
 - HepData archive of high energy physics data.
- Also provides (since it needs them itself)
 - XML descriptions of HepData records and generator parameters (HepML)
 - Lightweight code development environment (HepForge).
- **Active core members:** Andy Buckley, JMB, James Monk, Emily Nurse, Ben Waugh, Mike Whalley, with lots of close collaborators on different subprojects.

- Handy, lightweight development platform for small reusable HEP software projects.
- Pick and mix of services available:
 - Version control, bug/issue tracking, download/release management, web space, wiki, mailing lists, shell account...
- Ask if you want to know more
 - <http://hepforge.cedar.ac.uk> (or <http://hepforge.org>)

- Data store for all HEP measurements
- Exists in legacy Berkeley DB and Beta MySQL DB

- Fortran library of generator-independent analysis routines.
 - Each routine reproduces the particle-level distributions of a single paper.
- HZTool status and plans:
 - HERA origins -> fairly large selection of HERA papers
 - Small selection of Tevatron measurements (adding more)
 - Badly need LEP hadronic measurements, but these may be better added to Rivet (see later).
 - The only planned major development (apart from including new generators and new data) is to export data header files directly from HepData for each release.
 - Current version 4.1 (5/4/2006).
 - See <http://hepforge.cedar.ac.uk/hztool>

- Fortran main program and I/O for HZTool
- Mainly intended for use by JetWeb, but also useful to others
- HZSteer status and plans:
 - Currently lots of development (beta releases).
 - Outputs proposed HepML parameter descriptions for HERWIG and PYTHIA.
 - Outputs histograms as AIDA (XML) or HBOOK RZ files.
 - See <http://hepforge.cedar.ac.uk/hzsteer>

Rivet and RivetGun

- Robust Independent Validation of Experiment & Theory
- Approximately equivalent to a C++ replacement of HZTool (Rivet) and HZSteer (RivetGun).
 - Will make greater use of existing external libraries (CLHEP, KtJet/FastJet etc)
 - Rivet is generator independent.
 - RivetGun must interface to ThePEG, Sherpa, Pythia8 and existing Fortran generators.
 - Configure generators using HepML (see later).

Rivet and RivetGun

- Design and development ongoing
 - see web pages, e.g.
<http://hepforge.cedar.ac.uk/rivet/code/pub/inherits>
- Discuss/try out a working demo this week.
 - Current developers; A. Buckley, L. Lonnblad, B. Waugh, JMB
 - Generator interface and relationship to HepMC needs discussion & work.
 - Hope to expand this a lot, as with HZTool, once the basic structure is working.
- See <http://hepforge.cedar.ac.uk/rivet> and <http://hepforge.cedar.ac.uk/rivetgun>

- Web and database server for archiving validated MC models.
- Uses HZSteer and HZTool running on LCG
 - Future releases will use Rivet/Gun as well
- Undergone major redevelopment after initial demonstrator version
 - New test version now available.
 - Uses HepML for describing validated models
 - Will soon use HepData as the single source for all measurements

- **Goals**
 - Build up database of validated models using wide range of existing data
 - Running HZSteer and RivetGun on LCG
 - CEDAR now a registered VO
 - would like to use GENSER distribution of generators
 - would like HZSteer (and eventually RivetGun) supported in GENSER.
 - Add new generators and data rapidly as they appear
 - Add more user front-end facilities for interactive tuning and analysis
- See <http://jetweb.cedar.ac.uk>

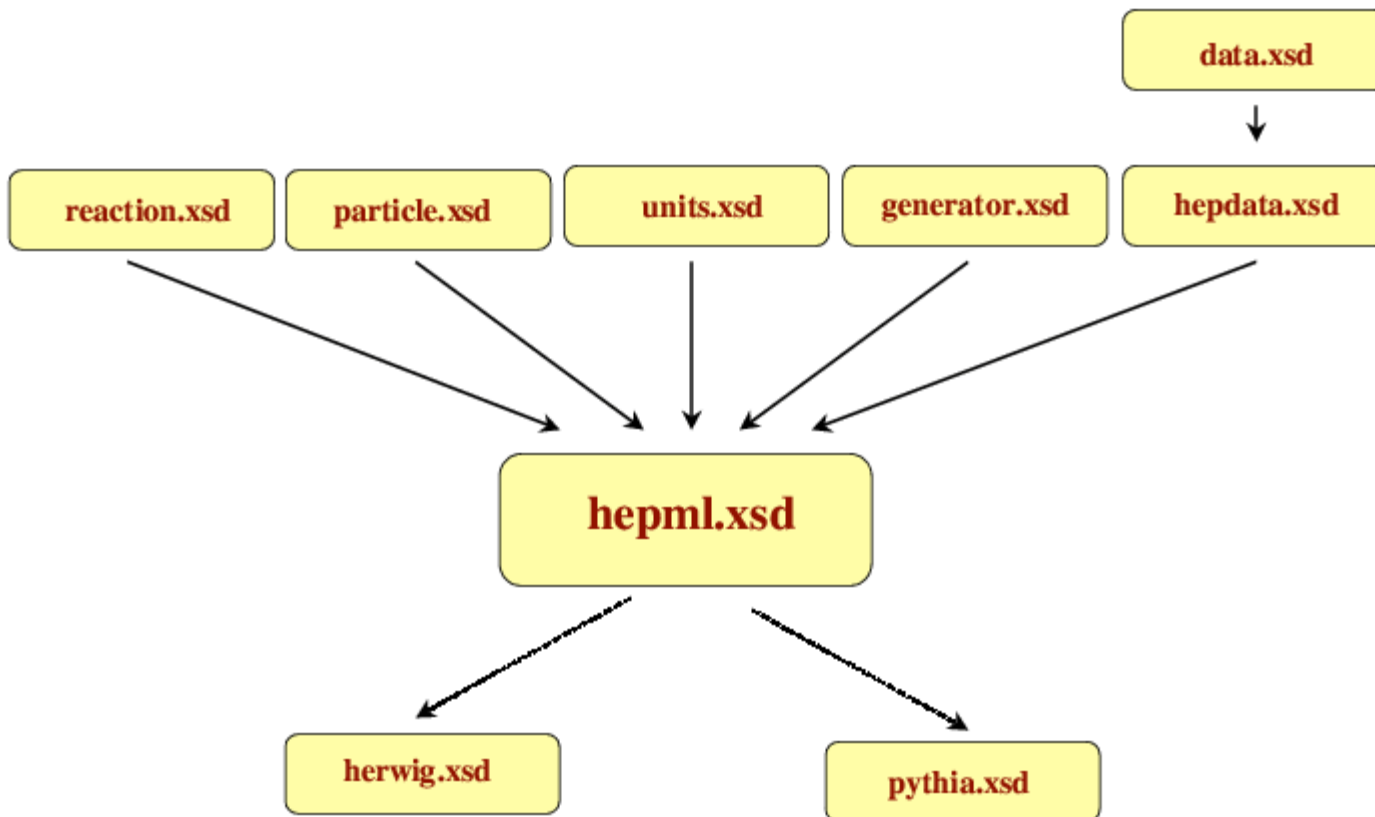
- Why HepML?
 - Need to be able to reliably reproduce a particular generator run.
 - Need to be able to match existing MC datasets with known, validated models
 - Need to be able to import/export configuration files from a generator, JetWeb or experiment MC store.

- Could be satisfied by any well defined file format.

Benefits of XML:

- “Self documenting” and can be validated against the schema by third-party software.
- “Future proof”: ASCII (not binary) and versioned.
- Well defined hierarchical syntax with standard parsing tools
- Can reflect the object-model or fortran-list structure of old or new generators
- Can use “subclassing”, meaning that a program which can read one generator's parameters can read them all (though it will not necessarily know what to do with them, of course!).

- General XML schemas for describing HEP objects.
 - Name also used by MCDB project for planned MC event-record I/O
 - Ideally would like to merge/reuse common elements with MCDB subschemas when available.
- CEDAR HepML schemas now available for MC parameters and HepData records
 - I/O for JetWeb and HepData
 - Each generator author can effectively “subclass” the general parameter description, thus restricting the names and types of parameters to be only those relevant for their generator.



- Some elements may be common to all generators
 - Beam energy and particle type
 - PDFs
 - Luminosity/number of events/cross section
 - ...?
- In general should be no constraint on parameters which may be declared valid for a given generator.
 - lists, arrays, maps, tables.... double & integer... others?
 - types specified in superclass so they can be read by anyone, selection of these types (and their names) completely controlled by generator authors.

- Proposed schemas for Fortran Herwig and Pythia currently under discussion with authors
 - Test versions of Fortran output available in HZSteer
 - Used by JetWeb when uploading new MC data.
- See <http://hepforge.cedar.ac.uk/hepml>
- Planned/possible use cases:
 - JetWeb upload & download
 - RivetGun (and maybe HZSteer) configuration
 - Generator configuration GUI?
- Will also be used for HepData I/O.
 - Some common elements, not MC dependent.

- **Relation to LCG**
 - LCG MCDB group (Lev Dudko et al) propose HepML schemas with use cases including I/O for intermediate states of generators, matrix elements and final states.
 - Quite a bit of common overlap (units, particles, reactions, generator parameters...)
 - First schemas, as well as an attempt at merging CEDAR/LCG schema, were released yesterday by MCDB, for discussion
 - <https://twiki.cern.ch/twiki/bin/view/Main/HepML>

Wish-list for this workshop

- Agree design principles of HepML with authors and MCDB, and work on concrete generator parameter schema.
- Agree priority datasets for validation, and start getting the missing ones into JetWeb (HZTool or Rivet).
- Determine some key versions and parameter sets from the experiments to start including in JetWeb.
- Understand better how we should use GENSER, and maybe get a release of HZTool and HZSteer incorporated.
- Demonstrate and discuss Rivet and RivetGun and a common generator interface.