Steps in the Production of a Monte-Carlo Sample

The CMS software framework is organized such that the production of a dataset for analysis has to be done in successive steps, with I/O from disk. Arrangement of the steps is dictated by : software specifics, storage and software requirements, flexibility for further reprocessing, etc. In this document, a request is the logical container of the information required to perform one of such step. McM allows for the formation of any combination of steps.

Building the Request Configuration

A request document holds a sequence of arguments to a configuration utility, bound to the software release used in the production. It can be modified from the values defined in the campaign in specific cases.

Evolution in Between Requests

Requests are prepared by the contacts from various Physics groups. McM can produce validation histograms to check the physics content prior to full production, or at each relevant stage in the production chain. The request is then processed through the desired chain of requests.

Actions to Be Taken on Samples

Production operators set the required action on the root requests or preparation of the next request. The output is fed back to McM for completion of the production of samples for analysis.

Validation of The Physics Content

The configuration of all varieties of event generator is complex. The full production of a samples can take a long time. In order to not waste resource ans delay delivery, a validation job on a selected number of events, is driven by McM and upload a set of plots to the Data Quality Monitoring (DQM) server.

Cycles of Production

Requests prepared in McM are sent for production to the computing system, monitored during processing, and the output is fed back to McM for completion of the production or preparation of the next request.

Documentation and Glossary

A set of documentation twiki [http://twiki.org/] is written with instructions and glossary for users, contacts and operators. The documentation is a collaborative effort.

Overview and Outlooks

McM uses evolutive contemporary technology. It incorporates the experience gained from past years of operation of production tools at the interface of Analysis and Computing. Novel concepts (flow, chained requests, chained campaigns, actions) were introduced to simplify and automatize the production of a sample through the logical and technical steps of processing. McM enforces testing and validation of the single production. McM enforces testing and validation of the single requests prior to production to improve manpower and computing efficiencies. Special and vanilla requests can be handled within targeted campaigns of production. McM serves as bookkeeping of sample production. It permits the aggregation of information on production from within or from other CMS services. Extensive documentation and tutorials are provided to the users and contacts. McM is the service for central production of samples for analysis in CMS.

Chains of Campaigns

Campaigns that are connected with “flows” form chains of campaigns. A “chained campaign” defines what are the necessary operations toward the finalization of the production of a sample for analysis.

Chains of Requests

The completion of the production of a given Monte-Carlo sample.

<table>
<thead>
<tr>
<th>Campaign #</th>
<th>Detector campaign</th>
<th>Flow</th>
<th>Action</th>
<th>Preprod</th>
<th>Prepared</th>
<th>Chained request</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Service Infrastructure and Technology

McM has a very simple and common infrastructure, with schema oriented documents stored in a database, object manipulation run on server side in a python application, visualization and operation in JavaScript on the client side.

* CouchDB [http://couchdb.apache.org/]
* CherryPy [http://cherrypy.org/]
* JavaScript, Angularjs [http://angularjs.org/]

Jean-Roch Vlimant, CERN

McM Home page

Welcome to the McM developers' home page

Jean-Roch Vlimant, CERN

Physics Performance and Dataset Project

Physics Data & MC Validation Group

Jean-Roch Vlimant for the CMS Collaboration

Steps in the Production of a Monte-Carlo Sample

The CMS software framework is organized such that the production of a dataset for analysis has to be done in successive steps, with I/O from disk. Arrangement of the steps is dictated by : software specifics, storage and software requirements, flexibility for further reprocessing, etc. In this document, a request is the logical container of the information required to perform one of such step. McM allows for the formation of any combination of steps.

Building the Request Configuration

A request document holds a sequence of arguments to a configuration utility, bound to the software release used in the production. It can be modified from the values defined in the campaign in specific cases.

Evolution in Between Requests

Requests are prepared by the contacts from various Physics groups. McM can produce validation histograms to check the physics content prior to full production, or at each relevant stage in the production chain. The request is then processed through the desired chain of requests.

Actions to Be Taken on Samples

Production operators set the required action on the root requests or preparation of the next request. The output is fed back to McM for completion of the production of samples for analysis.

Validation of The Physics Content

The configuration of all varieties of event generator is complex. The full production of a samples can take a long time. In order to not waste resource ans delay delivery, a validation job on a selected number of events, is driven by McM and upload a set of plots to the Data Quality Monitoring (DQM) server.

Cycles of Production

Requests prepared in McM are sent for production to the computing system, monitored during processing, and the output is fed back to McM for completion of the production or preparation of the next request.

Documentation and Glossary

A set of documentation twiki [http://twiki.org/] is written with instructions and glossary for users, contacts and operators. The documentation is a collaborative effort.

Overview and Outlooks

McM uses evolutive contemporary technology. It incorporates the experience gained from past years of operation of production tools at the interface of Analysis and Computing. Novel concepts (flow, chained requests, chained campaigns, actions) were introduced to simplify and automatize the production of a sample through the logical and technical steps of processing. McM enforces testing and validation of the single requests prior to production to improve manpower and computing efficiencies. Special and vanilla requests can be handled within targeted campaigns of production. McM serves as bookkeeping of sample production. It permits the aggregation of information on production from within or from other CMS services. Extensive documentation and tutorials are provided to the users and contacts. McM is the service for central production of samples for analysis in CMS.

<table>
<thead>
<tr>
<th>Campaign #</th>
<th>Detector campaign</th>
<th>Flow</th>
<th>Action</th>
<th>Preprod</th>
<th>Prepared</th>
<th>Chained request</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Service Infrastructure and Technology

McM has a very simple and common infrastructure, with schema oriented documents stored in a database, object manipulation run on server side in a python application, visualization and operation in JavaScript on the client side.

* CouchDB [http://couchdb.apache.org/]
* CherryPy [http://cherrypy.org/]
* JavaScript, Angularjs [http://angularjs.org/]

Jean-Roch Vlimant, CERN

McM Home page

Welcome to the McM developers' home page

Jean-Roch Vlimant, CERN

Physics Performance and Dataset Project

Physics Data & MC Validation Group

Jean-Roch Vlimant for the CMS Collaboration