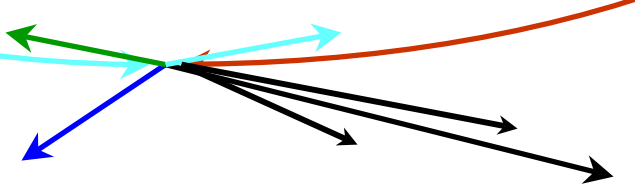
Monte Carlo at DØ
A Users Guide

Iain Bertram

DØ Collaboration Meeting

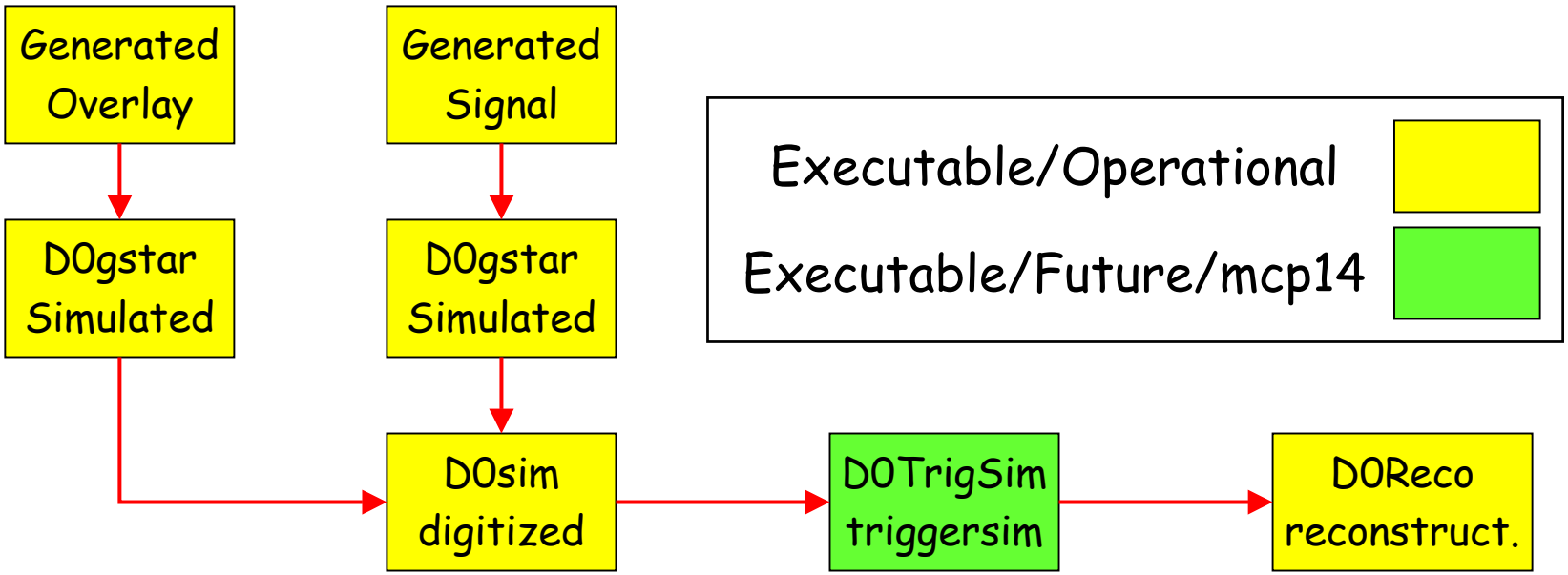
11 February 2003



Outline

- Monte Carlo Production
  - Outline of Processing
  - Runjob
  - Metadata
  - Creating a Request
    - How to get what you want and Need!!
    - How to find existing requests
    - How to make sure your request will run
  - What happened to my Request?
  - Finding your data once it has been produced
- Runjob
  - Producing your own data
  - Processing MC data sets previously produced

MC Farm Processing



- Standard MC Processing uses the runjob package
  - ➔ Five different executables
  - ➔ Store DST + Raw Chunk, and Thumbnail
  - ➔ Request System based on SAM
    - Still under development!

Runjob

- Runjob is the tool used to run MC Production
  - Macro driven tool to chain together many executables
  - Requests are processed using runjob keywords
  - Current web Page: http://www-clued0.fnal.gov/runjob/
  - Mailing List: d0_mc_runjob_users@fnal.gov

- All MC Requests are based on runjob
  - If you cannot run your own MC using runjob the production team cannot
  - Before making requests should test your request to ensure it works.
  - Group representatives need to know how to use runjob!

Macros: Basics

1. Attach processes (basic MC processes)
    1. samglobal – global job descriptions
    2. generators: pythia, herwig, isajet, single, cosmic
        1. Pythia: Comphep, Onetop, …
        2. d0mess MC event selection system.
        3. evtgen, QQ, for b-decays
    3. d0gstar
    4. d0sim
        1. requires overlap events
    5. d0trigsim
    6. d0reco
    7. TMBAnalyze
2. For each specify a series of control words.

Runjob Macros – Example I

MiniDB
StandardD0
SaveOnMake

Required Control Parameters

```
attach samglobal
cfg samglobal define string Phase mcp13
cfg samglobal define string Stream notstreamed
cfg samglobal define string Description higgs h->gammagamma 170 gev wz associated
cfg samglobal define string ProducedByName bertram
cfg samglobal define string OriginName lancs
cfg samglobal define string RunType monte carlo
cfg samglobal define string FacilityName lancs
cfg samglobal define string DestinationDir  ./dest
cfg samglobal define string ProducedForName qzli
cfg samglobal define string GroupName higgs
cfg samglobal define string JobName Request-4220-
cfg samglobal define string RequestID 4220
cfg samglobal define string CurrentDir ./worker
cfg samglobal define int UniquenessLevel 3
```

Global Configuration Parameters

Red Signifies required request parameters

Runjob Macros – Example II

```
attach pythia
cfg pythia define int NumRecords  500
cfg pythia define string D0Release  p13.05.00
cfg pythia define int UseMaxopt  1
cfg pythia define float CollisionEnergy  1960.0
cfg pythia define float HiggsMass  70.0
cfg pythia define string PDFLibFunc  CTEQ4L
cfg pythia define string CardfileDir  higgs
cfg pythia define string Production  h_wzassociated
cfg pythia define string Decay  gamma+gamma
cfg pythia define string CardfileVersion  v00-04-22
```

Pythia Parameters

Red Signifies required request parameters

```
attach d0gstar
cfg d0gstar define string KeepParticleCalEnergy off
cfg d0gstar define string Geometry plate
cfg d0gstar define string D0Release p13.05.00
cfg d0gstar define int UseMaxopt 0
```

DØgstar

Runjob Macros - Example

```
attach d0sim
cfg d0sim define string MergeMinBias on
cfg d0sim define string MinBiOpt Poisson
cfg d0sim define string CalorimeterNoise on
cfg d0sim define string MinBiDataset lancs_MB_mcp13_dataset
cfg d0sim define string MinBiDir  /prj_root/794/mc_20/minbias_events/p
cfg d0sim define string LinkToRundata on
cfg d0sim define float NumMinBi 0.5
cfg d0sim define int FarmBuildMBD 1
cfg d0sim define string D0Release p13.05.00
cfg d0sim define int UseMaxopt 1
```

DØsim

```
attach d0trigsim
cfg d0trigsim define string D0Release p13.08.00
cfg d0trigsim define string InPrefix d0sim
cfg d0trigsim define string WriteEvents on
cfg d0trigsim define string WriteTuple off
cfg d0trigsim define string RunOnData off
cfg d0trigsim define int UseMaxopt 1
```

DØtrigsim

Runjob Macros - Example

```
attach d0reco
cfg d0reco define string LinkToRundata on
cfg d0reco define string D0Release p13.05.00
cfg d0reco define int UseMaxopt 1

attach runjob
cfg runjob define string RunOption RunExec

repeat 1
reset chain
cfg pythia make seeds
cfg d0gstar make seeds
make job
save
print script
cfg runjob run
end
```

DØreco

Job Control Parameters

Cardfiles

- Cardfiles Package
  - ls /d0dist/dist/packages/cardfiles/vxx-xx-xx/

bid/ calib/ ckm/ ctf/ higgs/ hit/ mc/ np/ qcd/ reco/ tauid/ top/ wz/

  - cardfiles maintained for each group
  - If you want production MC must have cardfiles in here!
  - SAM rule: must be in lower case
  - File naming: generator_production_decay.cards
- QQ, d0_mess, EVTgen

ls /d0dist/dist/packages/cardfiles/v00-04-22/ckm/

EvtGen/ QQ/ d0_mess/

ls /d0dist/dist/packages/cardfiles/v00-04-22/ckm/EvtGen

user.dec_BDStarLNu user.dec_BJPsiMuMu user.dec_BaBarJPsiKs user.dec_BsJPsiPhi

ls /d0dist/dist/packages/cardfiles/v00-04-22/ckm/QQ/

user.dec_bjpsimumu112702 user.dec_bs-dsphi-mu user.dec_bsjpsikstar-mu …

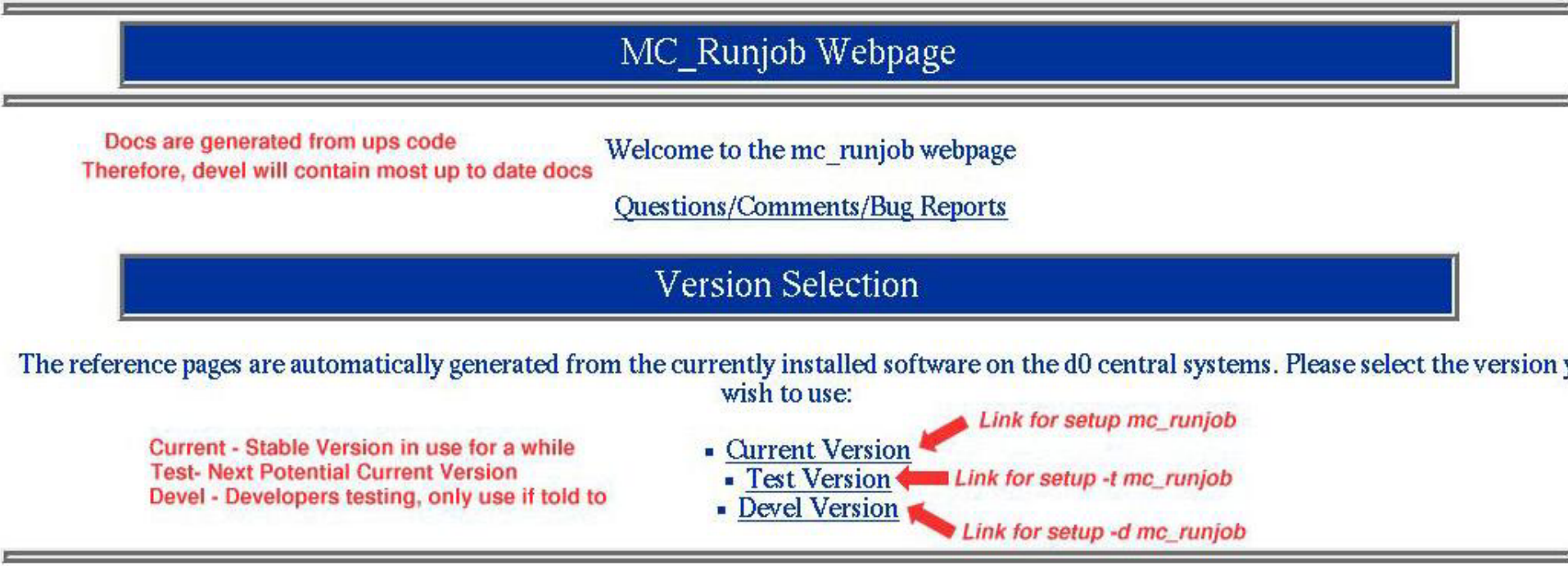
/d0dist/dist/packages/cardfiles/v00-04-22/ckm/d0_mess/

d0_mess_Bs-etag.rcp d0_mess_bbbar-incl.rcp

Example Cardfile Names

onetop_2to2.cards
onetop_Wbbbar.cards
onetop_Wgfusion.cards
onetop_Wt.cards
onetop_mcrunjob.cards
onetop_schannel.cards
onetop_ttbar.cards
pythia_bbbar_bmunu.cards
pythia_bbbar_incl.cards
pythia_ccbar_incl.cards
pythia_foronetop.cards
pythia_gam+jets.cards
pythia_onetop_2to2_wmunu.cards
pythia_onetop_Wgfusion_wenu.cards
pythia_onetop_Wgfusion_wmunu.cards
pythia_onetop_schannel_wmunu.cards
pythia_qcd.cards
pythia_ttbar_wjj+wjj_loq.cards
pythia_ttbar_wjj+wjj_mrst.cards
pythia_ttbar_wjj+wlnu.cards
pythia_ttbar_wjj+wmunu.cards
pythia_ttbar_wlnu+wjj.cards
pythia_ttbar_wlnu+wlnu.cards
pythia_ttbar_wmunu+wenu.cards
pythia_ttbar_wmunu+wjj.cards
pythia_ttbar_wmunu+wjj_hiq.cards
pythia_ttbar_wmunu+wjj_loq.cards
pythia_ttbar_wmunu+wjj_mrst.cards
pythia_onetop_2to2_wenu.cards
pythia_ttbar_wmunu+wmunu_hiq.cards
pythia_ttbar_wmunu+wmunu_loq.cards
pythia_onetop_schannel_wenu.cards
pythia_ttbar_wtaunu+wenu.cards
pythia_ttbar_wtaunu+wmunu.cards

A word on documentation



MC_Runjob Webpage

Docs are generated from ups code
Therefore, devel will contain most up to date docs

Welcome to the mc_runjob webpage

Questions/Comments/Bug Reports

Version Selection

The reference pages are automatically generated from the currently installed software on the d0 central systems. Please select the version y
wish to use:

Current - Stable Version in use for a while
Test- Next Potential Current Version
Devel - Developers testing, only use if told to

- Current Version — *Link for setup mc_runjob*
- Test Version — *Link for setup -t mc_runjob*
- Devel Version — *Link for setup -d mc_runjob*

MC_Runjob Configurator Docs Index

Main Index

Complain Here

This page contains links to the mc_runjob documentation generated from the currently installed UPS chains. Version information is available below, along with an array of tutorials and reference sections. The best starting place is in the tutorial section, many of which will refer you to the appropriate reference pages indexed below. If you have any questions, please send them to the mailing list.

Version Information

Page Listing Version numbers for c,t,d versions

Information on the package versions of mc_runjob that are installed as UPS chains can be found on the Version Page. If a version is declared as "None" it means that there is no version installed with that particular UPS chain.

mc_runjob Tutorials

Tutorial Webpage index

A variety of tutorials for using mc_runjob in various use cases can be found on the Tutorial index page.

Reference Documents

Each mc_runjob tool generates its own reference documentation which can be viewed by clicking on the list of links below:

Information for individual tools

- Linker
- D0TrigsimConfigurator
- D0ReadJobMixin
- mc_jobscript — Automatic script generator for CAB and Clued0 jobs
- CABJobscript
- SamInputMixin
- D0RecoConfigurator
- D0MessMixin
- TupleMakerMixin
- Clued0Jobscript
- SamStreamConfigurator
- PythiaConfigurator
- CopyD0omConfigurator
- FileStreamConfigurator
- D0Configurator

Example Macro Generator to be added soon...

Coming Features

- define string FrameworkRCPName
  - ➔ Will allow you to pick up any official framework RCP in the release area of the package
  - ➔ Will be extended to arbitrary packages so groups can have their own set of framework rcp files
- Released packages can be included in a straightforward manner…
  - ➔ release your packages properly
- TMBStream_x thumbnail streaming.
- PMCS – coming soon
- SAM Input Files
  - ➔ Runjob being linked to SAM
  - ➔ will be able to run on sam files data/MC for chained processing
  - ➔ eg. Trigsim -> d0reco -> thbmanalyze
  - ➔ will use d0tools in future as running method.

Basic Session

- Setups

setup D0RunII p13.08.00

setup sam

satup mc_runjob

- To run
  - ➔ create macro e.g. Example.macro
  - ➔ mc_runjob –macro=Example.macro
- To run on CAB
  - ➔ (use medium queue for non-sam jobs)
  - ➔ Instructions: http://www.nuhep.nwu.edu/~schellma/cab/cab_doc_v2.html
  - ➔ mc_jobscript -cab -h
- To run on clued0: mc_jobscript -clued0 -h

Creating Requests

- Sam Command Based on python script
  1. Create a python script describing job
  2. Based on runjob macro
  3. Basic Keyword Types: Named Via DataTier
     Global (SamGlobal)
     Generated (Generators, pythia, herwig,….)
     Simulated (Døgstar)
     Digitized (Døsim)
     Triggersimulated (Døtrigsim)
     Reconstructed (Døreco)
  4. Specify the minimum number of keyword to get what you need
  5. submit your request
     sam submission instructions: sam create request
     http://d0db.fnal.gov/sam_user/samCreateRequest.html

sam create request

Usage:

sam create request <--requiredOptions> [--options] [-flags]

Where:

--requiredOptions:

--dict=<value> # python dictionary describing keywords, datatier and application family

--group=<value> # sam group user is a member of

--num-events=<value> # number of events to be processed for request

--options:

--comment=<value> # description of request

--email=<value> # email address of contact for request

--job-name=<value> # name associated with application or process creating the request

--priority=<value> # initial priority of request

--user=<value> # alternate user name to use when creating request

-flags:

-d

-v

Example Request

- To be carried out under group supervision for official production
- Should create a request for all MC.
  - ➔ even MC generated by-group for the group
  - ➔ allows ability to search for MC efficiently

- Commands:
  - ➔ Create python file: http://www-d0.fnal.gov/computing/mcprod/Tutorial/demo.py
  - ➔ submit request:

    sam create request --dict=demo.py --group=bphysics --user=bertram --num-events=10000 --comment='tutorial example' --email=bertram@fnal.gov --job-name='MC Request' --priority=5
  - ➔ Check request on request web page.
  - ➔ Ask group representative to approve request

Metadata

- Requests are Based on Metadata
  - ➔ Extensible system
  - ➔ Can add new keyword value pairs if required
  - ➔ Can add additional generators
  - ➔ Can add additional data tiers

- Warning
  - ➔ Aimed at storing useful information about each file
  - ➔ random number seeds, processing locations, timestamp, etc
  - ➔ Not always good for searching

Current Requests

- Several ways of finding them:
  - SAM Web page → Browse the SAM Meta-data → Request Query
  - Being Updated to allow more sophisticated searches (temp web page)
  - MC Production Web Page (DØ at work) → Current Requests
- Structure is very uniform
  - Important to test your requests before running
  - Recall SAM only knows lower case
  - Production system still undergoing debugging

Accessing MC Data

- Best Method
  - ➔ Create Data set definitions based on request ID

sam translate constraints --type=mcrun --dim='global.requestid 4594‘

tsim-p13.08.00_IsData-off_Iain-Bertram_algo_recocert_Iancs_4594_03040185737
sim-p13.08.00_Noise-on_NMB-0.0_MB-Fixed_Iain-Bertram_algo_recocert_Iancs_4594_03040185708
tmb-p13.08.00_Iain-Bertram_algo_recocert_Iancs_4594_03040185737
tsim-p13.08.00_IsData-off_Iain-Bertram_algo_recocert_Iancs_4594_03040185708
reco-p13.08.00_Iain-Bertram_algo_recocert_Iancs_4594_03040185737

  - ➔ To specify a particular data tier

--dim='global.requestid 4594 and data_tier thumbnail‘

tmb-p13.08.00_Iain-Bertram_algo_recocert_Iancs_4594_03040185737
tmb-p13.08.00_Iain-Bertram_algo_recocert_Iancs_4594_03040185708 …

  - ➔ Or use the the data set definition editor: http://d0db.fnal.gov/sam_project_editor/DatasetEditor.html

How to find requests of Interest

- Not quite online (sam query problems getting in way)
  - ➔ temp web page: will be released soon once testing complete
  - ➔ SAM Web page → Browse the SAM Meta-data → Request Query
- Base search metadata keywords used to form requests
  - ➔ for example to find all requests that are top production processes

Metadata for Files

- SAM Web page → Browse the SAM Meta-data → Data Files Parameter Query
  - ➔ use to check all parameters with any file
  - ➔ Will be updated soon with drill trough capabilities allowing access to all parent files
  - ➔ Useful for checking exactly what is in a file after processing
  - ➔ All request metadata should be the same for stored file and request.

My Generator isn't There!

- Volunteer some help to get it into runjob
- Storing own MC for reprocessing
  - ➔ Needs metadata…
  - ➔ should be created under a request all MC should have an associated request…
- Production Team doesn't care if output is rubbish
  - ➔ users and developers responsibility to test
  - ➔ you need to test your stuff

We need you!

Reprocessing

- Reprocessing
  - ➔ Can use a sam data set definition as an input method
  - ➔ Probably CAB is best location at the moment
    - should be ready for mcp14
    - runjob will take sam input
    - File naming needs checking to ensure no duplicates

Questions

- Over to you