

CDF Software

- **Introduction**
- **Main Programs**
- **Code Organization**
- **Batch System**

Introduction

- HEP a computer-intensive field:
 - Tevatron: 3 million collisions/s
 - CDF: experiment with ~1 million channels
 - Write to tape ~25 MBytes/s
 - Have recorded ~ $\frac{1}{2}$ billion events
 - Over 100 Tbytes to tape and counting
 - Large software project (>1 million lines)
 - Large computing farms

Main Programs

- Generators
 - Calculation of particle collisions
 - Output a list of particles
- CDF Detector Simulation (cdfSim)
 - Simulates interaction of particles with CDF
 - Input a list of particles, output raw data banks
- Reconstruction
 - Turns raw data input into electron, muon, jet objects
- Analysis
 - Higher-level reconstruction
- ROOT

Generators

- Major packages:
 - PYTHIA
 - HERWIG
 - integrated in cdfSim
 - generate “events”
- Other packages:
 - CompHEP, MadGraph (Vegas)
 - Alpgen, VECBOS (W+jets)
 - Decay packages:
 - BBGEN, QQ, Tauola
- Output of Generators: HEPG Bank

Simulation and Reconstruction

- Simulation package is called “cdfSim”
 - Based on GEANT package from CERN
 - Some parts are parametrized
 - Output are raw data “D” banks e.g. “ISLD”
 - Takes ~ 10 s/event
- Reconstruction Package called: “ProductionEXE”
 - Output: tracks, electrons, muons, jets
 - Takes about ~ 2 secs/event

Analysis and ROOT

- Some groups have analysis packages
 - TopFind for top group
 - B group has a few
 - Input is ProductionExe output
- ROOT
 - Last software package we use in the chain
 - Used to make plots, histograms
 - Can also be used as analysis framework
 - Good documentation (tutorials, examples)

Code Organization

- All our software can be seen in “CodeBrowser”
- Code distributed in “packages” (>100 of them)
- In general have following structure
 - Geometry
 - Objects (containers)
 - Algorithms
 - Modules
- We use CVS to keep track of versions
- Releases (e.g. 4.10.4):
 - Associates a version to every file
 - Use stable releases (not “int” releases)

Code Organization (cont)

- Executables are assembled from modules
 - e.g. CalorimetryModule, JetCluModule
- Execution gives you “AC” prompt
 - AC stands for analysis control
 - Users can give instructions, parameters to modules (“talk-to”):
 - module talk JetCluModule
 - ConeRadius set 0.7
 - exit
- In general, instructions in .tcl files
- Ex: ProductionExe ProductionExe.tcl > test.log

Computer Resources

- CAF (Central Analysis Facility)
 - At Femilab
 - Has access to all CDF data
- Fish Cluster
 - ~30 dual cpus
 - Dedicated worker nodes
 - Desktops: only one cpu available
- Big Mac
 - 224 nodes, 30 Terabytes (almost available)
- Batch system is called FBSNG

Batch System

- Instructions by Reda on our Web page
 - setup fbsng
 - fbs : lists all commands
 - fbs status: gives details about submitted job
- To submit jobs:
 - fbs exec -q savard job.sh
 - fbs submit job.jdf
- We have 3 process types
 - short: 2 hours
 - meduim: 6 hours
 - Long: 48 hours