



Using the Grid for the ILC

Mokka and Marlin on the Grid
ILC Software Meeting, Cambridge 2006

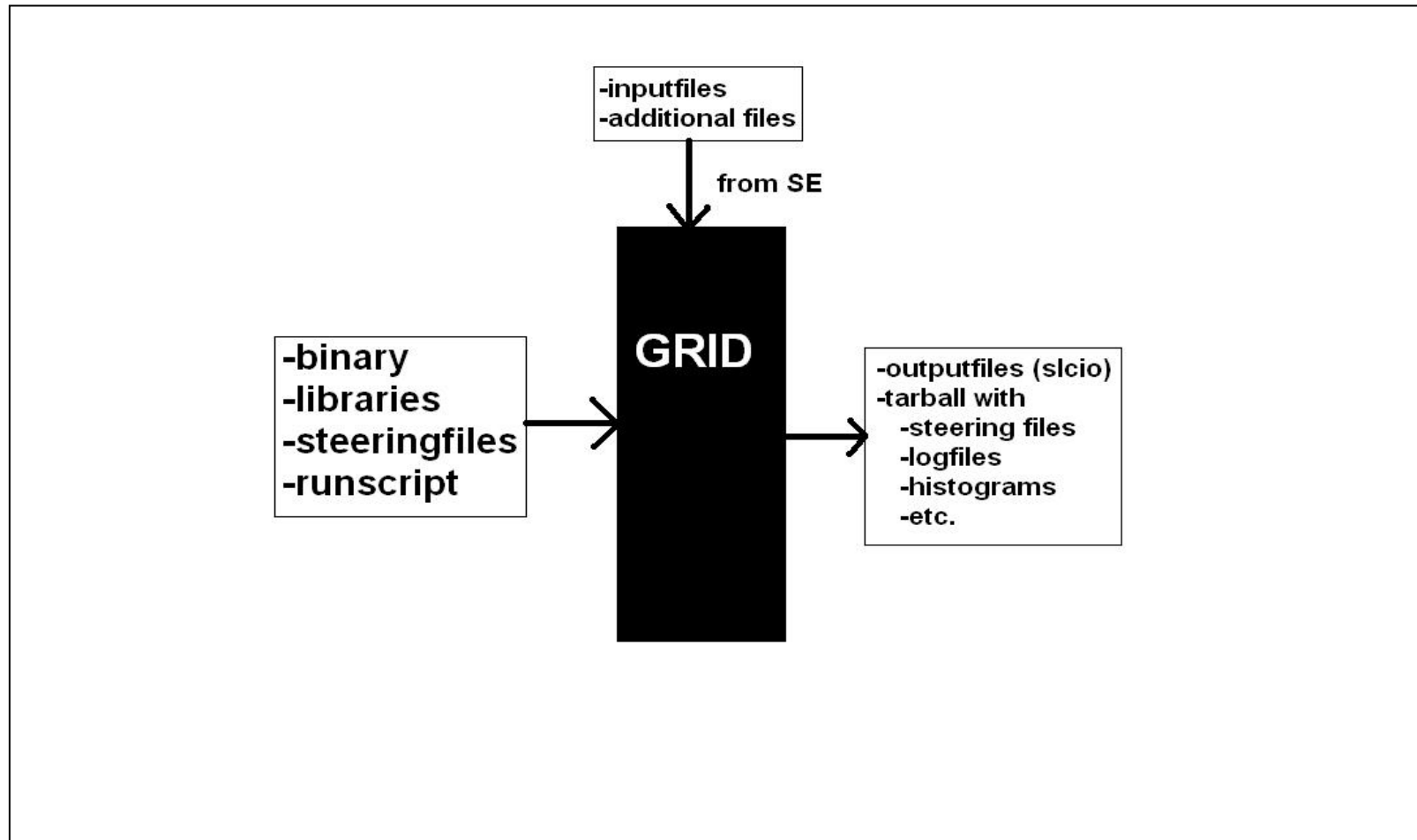


Motivation

- MC-Simulations are essential for analysis tool development and optimization studies
- “Divide and Conquer”-strategy speeds up computing
- Combining multiple computing environments into one array of computercenters is called a Grid
- In contrast to batch farms a Grid is NOT a regional computing facility
- All activities on a Grid are done via a “Grid Middleware”, therefore different systems can be combined with a common interface.
- The resulting computing capabilities are more powerful than any single batch system.



How things work





Mokka & Marlin

- Binaries and inputfiles are on SE
- Steeringfiles are generated locally and send at job-submission

- Outbound connectivity required (DB)
- G4 Datasets on SE
- Output:
 - tarball with slcio, steering- and logfile

- Needed libraries in tarball on SE
- Output:
 - reconstructed slcio-file
 - tarball with log- and steeringfile

- Easy to get other software working on the GRID



Runscript

- Display some info about WN
- Copy all necessary files from SE onto WN
- Set environment
- Run binary (either Mokka or Marlin)
- Put outputfiles on SE



Status (Mokka)

- Mokka was successfully tested
 - 10k Z0pole events took ~8hours (10 jobs á 1000evt)
 - 1k ttbar events took ~9hours (10 jobs á 100evts)
- Events for 8 different detector configurations are on Grid Storage.
- Files are available by our Online Database at ilcsoft.desy.de



Status (Marlin)

- Marlin was also successfully tested
 - 10k Zpole events took ~8hours (1 job)
- Reconstructed events for all simulated events are on GRID Storage.
- There are files with all different reconstruction chains
- Not yet available via our database



Available data

- Simulated data for:
 - $e^+e^- \rightarrow t\bar{t}$ (3000evts)
 - $e^+e^- \rightarrow W^+W^-$ (3000evts)
 - $e^+e^- \rightarrow cb$ (3000evts)
 - $e^+e^- \rightarrow dus$ (3000evts)
 - $e^+e^- \rightarrow Zh(120)$ (3000evts)
- Available at 500GeV and 1TeV
- Additionally there are files at ZPole (10k evts).
- 8 different detector configurations



How to access the data

- Requirements:
 - access to a machine with LCG-software package installed or SL-installation and access to afs (see grid.desy.de)
 - valid grid-certificate for the VO „ilc“
 - LFN from MC-Database



How to access the data (cont.)

- Browse our database at <http://ilcsoft.desy.de>

The screenshot shows the 'International Linear Collider Monte Carlo Database' website. The search results page displays a table of MC data files matching a query. The table has columns for Run ID, Process, CM Energy [GeV], and other details. The search criteria on the left include Run ID, Process, Center of Mass Energy [GeV], Date of Production, Event Generator, Detector Simulation, Detector Model, Physics List, and B Field [T].

Run ID	Process	CM Energy [GeV]	Other
M-5-4_zpole_dus_noir_LDC00Sc_2.0T_r1690_I2730_QGSP_BERT	Z0 -> uds	91.2	
M-5-4_zpole_dus_noir_LDC00Sc_3.00T_r1690_I2730_QGSP_BERT	Z0 -> uds	91.2	
M-5-4_zpole_dus_noir_LDC00Sc_3.00T_r1890_I2930_QGSP_BERT	Z0 -> uds	91.2	2006-03-23 3.0
M-5-4_zpole_dus_noir_LDC00Sc_4.0T_r1690_I2730_QGSP_BERT	Z0 -> uds	91.2	2006-03-23 4.0
M-5-4_zpole_dus_noir_LDC00Sc_4.0T_r1890_I2930_QGSP_BERT	Z0 -> uds	91.2	2006-03-23 4.0

The details page shows the following information:

- Simulation: Madka 5.4
- Physics List: QGSP_BERT
- Detector Model: LDC00Sc
- B Field: 2.0 T
- Number of files in this Run: 10
- Logical Filename on the Grid: M-5-4_zpole_dus_noir_LDC00Sc_2.0T_r1690_I2730_QGSP_BERT_1.tar.gz
- Steering File: M-5-4_zpole_dus_noir_LDC00Sc_2.0T_r1690_I2730_QGSP_BERT_1.steer
- Macro File: M-5-4_zpole_dus_noir_LDC00Sc_2.0T_r1690_I2730_QGSP_BERT_1.g4
- Log File: M-5-4_zpole_dus_noir_LDC00Sc_2.0T_r1690_I2730_QGSP_BERT_1.log
- Comments:
- Logical Filename on the Grid: M-5-4_zpole_dus_noir_LDC00Sc_2.0T_r1690_I2730_QGSP_BERT_10.tar.gz
- Steering File: M-5-4_zpole_dus_noir_LDC00Sc_2.0T_r1690_I2730_QGSP_BERT_10.steer
- Macro File: M-5-4_zpole_dus_noir_LDC00Sc_2.0T_r1690_I2730_QGSP_BERT_10.g4
- Log File: M-5-4_zpole_dus_noir_LDC00Sc_2.0T_r1690_I2730_QGSP_BERT_10.log

- Get the LFNs for the files you're interested in



How to access the data (cont.)

- Use LCG-Commands to retrieve files from storage
 - `lcg-cp --vo ilc lfn:{LFNfromDB} file:{absolute path}`



Summary & Outlook

- Using the GRID is a promising way for simulation and reconstruction (fast, reliable & easy to use)
- Mokka & Marlin are working fine on the grid
- ~500GB of data were simulated beginning of 2006
- Reconstruction of this data in 2 different chains was done quite fast
- Visit <http://grid.desy.de> for more information