



# **Software Common Task Group Report**

**Akiya Miyamoto**

**ILC PAC**

**POSTECH**

**3 November 2009**



# Topics

- **Introduction**
- **Common data samples**
- **Common software tools**
- **The GRID**
- **Summary**



# Charge and Members

- **Charge**

- **Coordinate tools and data bases common to LOI groups and code compatibility for simulation studies.**
- **Work on any common software issues for ILC detector studies**

- **Members**

- **Re-organized after IDAG validation**
- **Members:**
  - **Akiya Miyamoto(Convener, KEK, ILD)**
  - **Norman Graf(Deputy Convener, SLAC, SiD)**
  - **Frank Gaede(DESY, ILD)**
  - **Tony Johnson (SLAC, SiD)**



## Software issues in RD's work plan until 2012

4. Develop a realistic simulation model of the baseline design, including faults and limitations.

7. Simulate and analyze benchmark reactions, which can be updated

8. Simulate and analyze reactions at 1 TeV, including realistic higher energy backgrounds demonstrating the detector performance.

8&9: Based on the work of the Physics and Software group.

The reaction will be chosen to show the strength of ILC compared to other facilities



# Brief summary of activities

- **LOI era:**
  - **Contacts to make MC data samples for benchmark processes common to all LOI groups.**
  
- **After LOI submission (TILC09)**
  - **Meetings: 2 WebEx , face-to-face meetings at ALCPG09, E-mail communications**
  - **Discussing common samples, tools, and plans until 2012.**
  - **Set up a web page for our activities**
    - [http://www.linearcollider.org/wiki/doku.php?id=swctwg:swctwg\\_home](http://www.linearcollider.org/wiki/doku.php?id=swctwg:swctwg_home)
  - **LC software workshop (May 28-29) @ CERN**



# Topics

- Introduction
- **Common data samples**
- Common software tools
- The GRID
- Summary



# Common Data Samples: LOI era

- **Sample productions**
  - **Mostly by SLAC:**
    - SM samples of 2 ~8 fermions +  $n_\gamma$  for collisions of  $ee/e\gamma/\gamma\gamma$
    - $O(3000)$  processes for 250GeV [ $250\text{fb}^{-1}$ ] / 500GeV [ $500\text{fb}^{-1}$ ] in general; Lower statistics if  $\sigma$  is large;  $O(3)$ TB in data size.
  - **Some signal samples were also produced at DESY/KEK ...**
- **Using common StdHep format and shared by ftp & GRID**
- **Useful to understand differences**
  - **Note that samples used by SiD and ILD were not completely exact.**
    - SiD sim/re. pre-mixed samples (80%/30% pol.)
    - ILD sim./rec. un-mixed samples (100% pol. ) and mix in anal.
    - ➔ No significant problem is recognized, but details will matter.

# CDS: Generator Issues

- **Whizard was used for all processes.**
- **Requests for improvements for the next round**
  - inclusions of all top decay modes
  - tau pol. in processes other than tau-pair process
  - Whizard may not be optimal for Bhabha, gamma-gamma, SUSY with long cascade, ...
  - Should we change the Pythia fragmentation from its default to the LEP-tuned values ?
  - Better process ID assignments and logging
- **Proposing to continue with StdHep, as no viable alternative exists ( HepMC is ascii only)**



- **Major production will be done after**
  1. **Determination of machine parameters ( for luminosity spectrum, energy spreads )**
  2. **Physics Common Task group's decision on energies and processes.**
  3. **Generator updates**

**For 2 & 3, meetings with Physics CTG will be necessary.**

- ***Load sharing of production works needs to be considered.***

- **SB2009 : Urgent**
  - **We need beam parameters for studies on**
    - Physics performance
      - Reweight LOI results using new luminosity spectrum or re-generate samples and re-analysis
    - Backgrounds
      - Pair background → vertex detector, BCAL/LCAL
  - **Preliminary studies of pair background at 500 GeV are in progress using GunieaPig and CAIN.**
  - **Both ILD and SiD are eager to know a concrete parameter sets**
    - Initial Energy spread at 500 GeV
    - Parameters at different energy ( 250 GeV, 350 GeV? ... )
  - **Common samples will be produced using GuniaPig as soon as the parameter set is available**

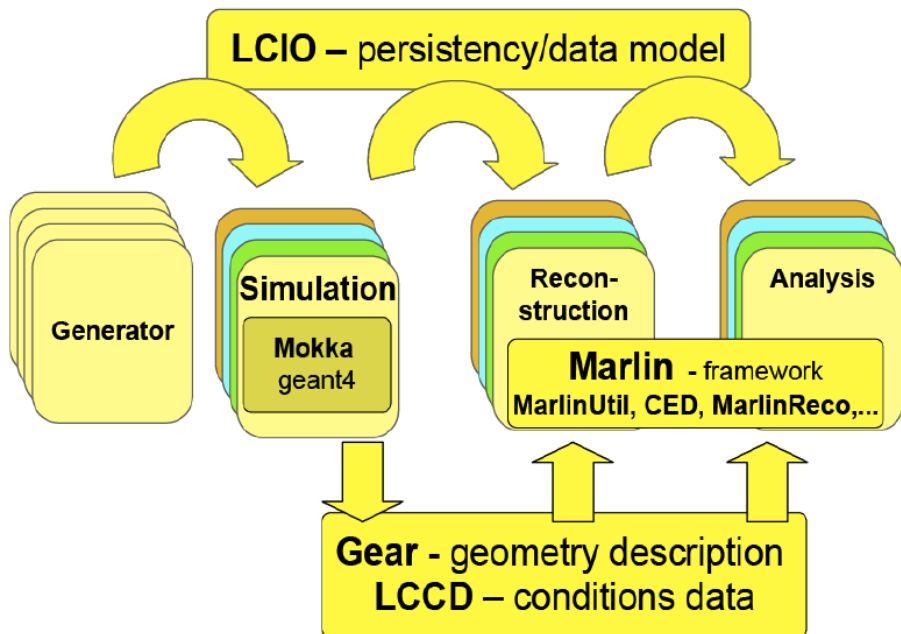
- **Samples under the 'SB2009' beam condition needs as soon as possible for timely feedback to GDE**
- **LCWS2010, Mar. 2010**
  - list of benchmarks ready ( from Physics CTG )
  - Machine parameter fixed
- **~ June 2010**
  - Updates of generator complete and start production of generator files.
  - Tools for generators should be ready by then
- **By the end of 2010:**
  - Generator files are available for simulations.
- **Simulations for DBD**
  - It will take ~ 1 year.
  - Expected to start by mid. 2011 or early 2012( Not solid yet )
  - Common tools for Sim/Rec should be updated by then



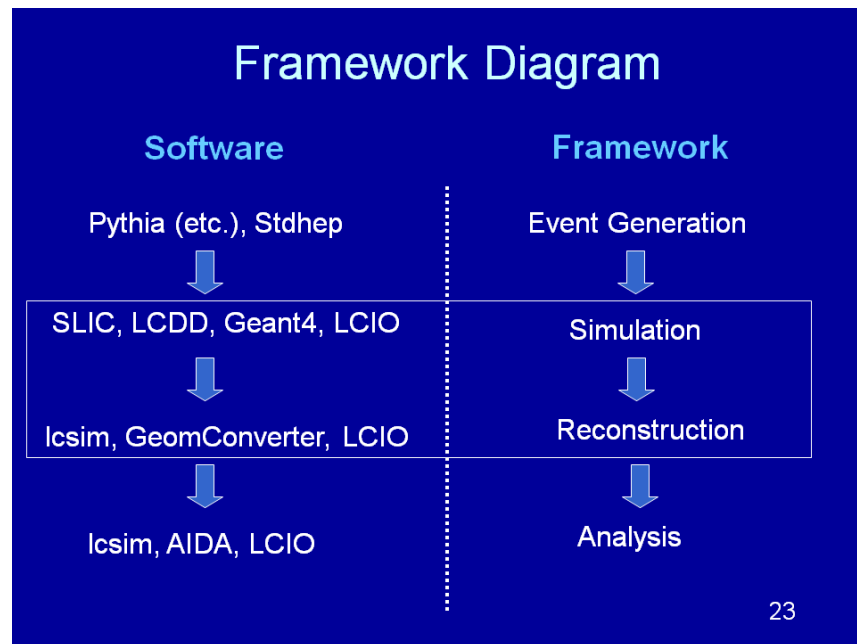
# Topics

- Introduction
- Common data samples
- **Common software tools**
- The GRID
- Summary

## ILD: C++ based



## SiD : Java based



### ■ Commonality in

- Detector Simulation (Geant4)
- Event reconstruction packages; Vertexing, PFA, ...

- **LCIO provides**
  - A common event data model and a common persistent format
- **LCIO has been a bases of**
  - Successful merger of GLD + LDC → ILD
  - Successful use of cross-concept software packages, eg. LCFIVertexing, PFA, ...
  - SiD Java-based reconstruction were further processed using ILD's C++ based MarlinReco
- **LCIO2.0**
  - User requests and LOI lessons initiated discussions. Works towards LCIO2.0 are in progress.
  - We will continue using LCIO



# Software update plan

- **ILD**

- Merge jsf goodies and Marlin
- Improve simulation, geometry system, reconstructions
- Develop a test system and new GRID production system

- **SiD**

- Improve detector models for simulation, better support for dual readout calorimeter, improve geometry tools
- Full reconstruction with new geometry components
- Use standard Geant4 physics list
- Improve for batch run system

*Foreseen many active field in software development*

*Software CTG encourages tool sharing when possible*



# LC Software Workshop @ CERN

- **Organization**
  - by CLIC software community to discuss software issues common to ILC
  - 27 participants from CLIC, ILD, SiD and 4<sup>th</sup>
- **Topics is selected CLIC community's point of view**
  - **Persistency: StdHep/HepMC, LCIO**
  - **data model: LCIOV2**
  - **Java  $\leftrightarrow$  C++ interfacing**
  - **Geometry tools: Gear, GDML, TGeo, Visualization**
  - **Common PFA**
  - **SW-Packaging, Virtualization, common Framework**

<http://indico.cern.ch/conferenceDisplay.py?confId=58717>





# Topics

- Introduction
- Common data samples
- Common software tools
- **The GRID**
- Summary

- In LOI era,
  - SID extensively used both the LCG and OSG grids (RAL, fermiGrid, DESY, in2p3, ...), in addition to SLAC Isf batch system
  - ILD used the LCG grids (DESY, in2p3, KEK, ... )
- **Number of issues with GRID job submission, monitoring and file transfers. Not a simple job, but it worked.**
- **Very LHC-centric, but GRID is responsible for successful processing of many tens of millions of events of a few hundreds TB data**
- **For the moment, we are not sure about the availability of computing resources in the next study phase.**  
**CPUs, network bandwidth, ... depends on LHC**

# Summary

- **Common data samples and code sharing backed up by LCIO were successful in LOI era.**
- **We wish to keep this direction in DBD era**
- **Major software works in coming years includes**
  - **New data samples for new benchmark processes**
  - **Updates of detector simulators and reconstruction tools**
- **Software Common Task Groups will continue to work**
  - **on common data samples**
  - **on common software tools and standards**