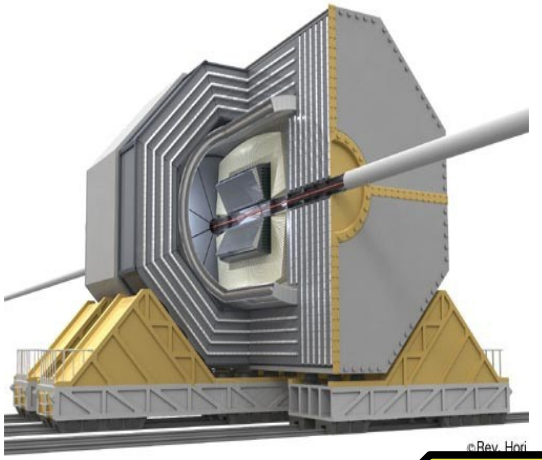
Marlin et al
Status of the ILD_LDC core software tools

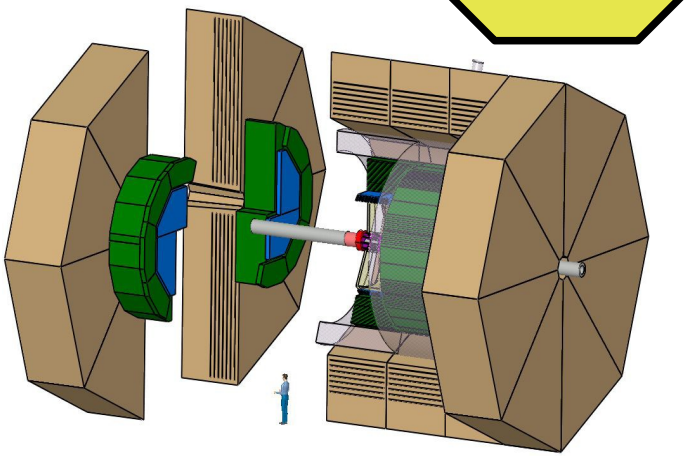
Frank Gaede
DESY
ECFA2008, Warsaw, Poland
June 9-12, 2008

Outline

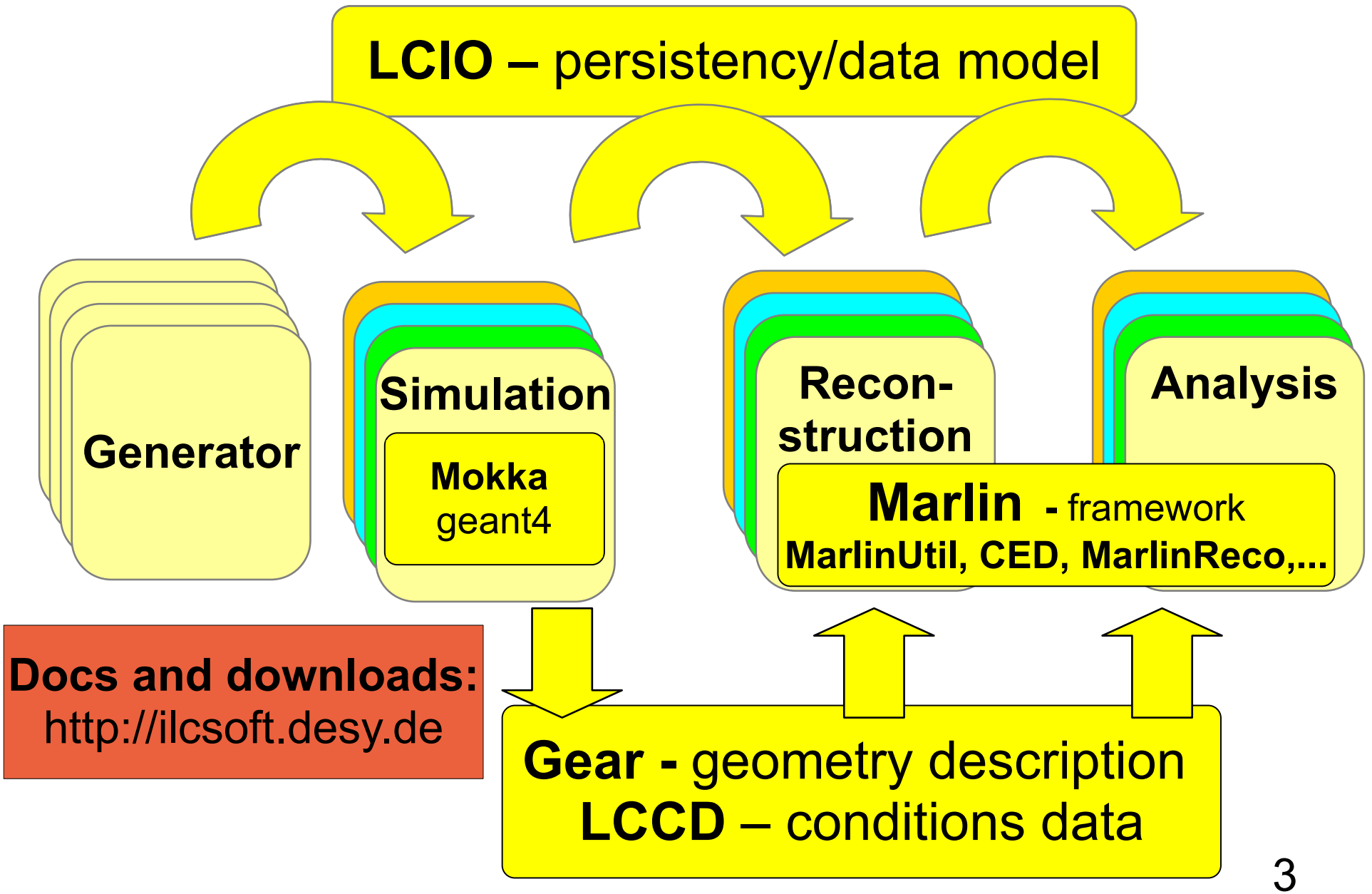
- introduction/overview
- core tools developments
  - LCIO, Marlin, Gear,...
- standard reconstruction
  - DST files
- Grid
  - sw-installation
  - Monte Carlo production
- summary



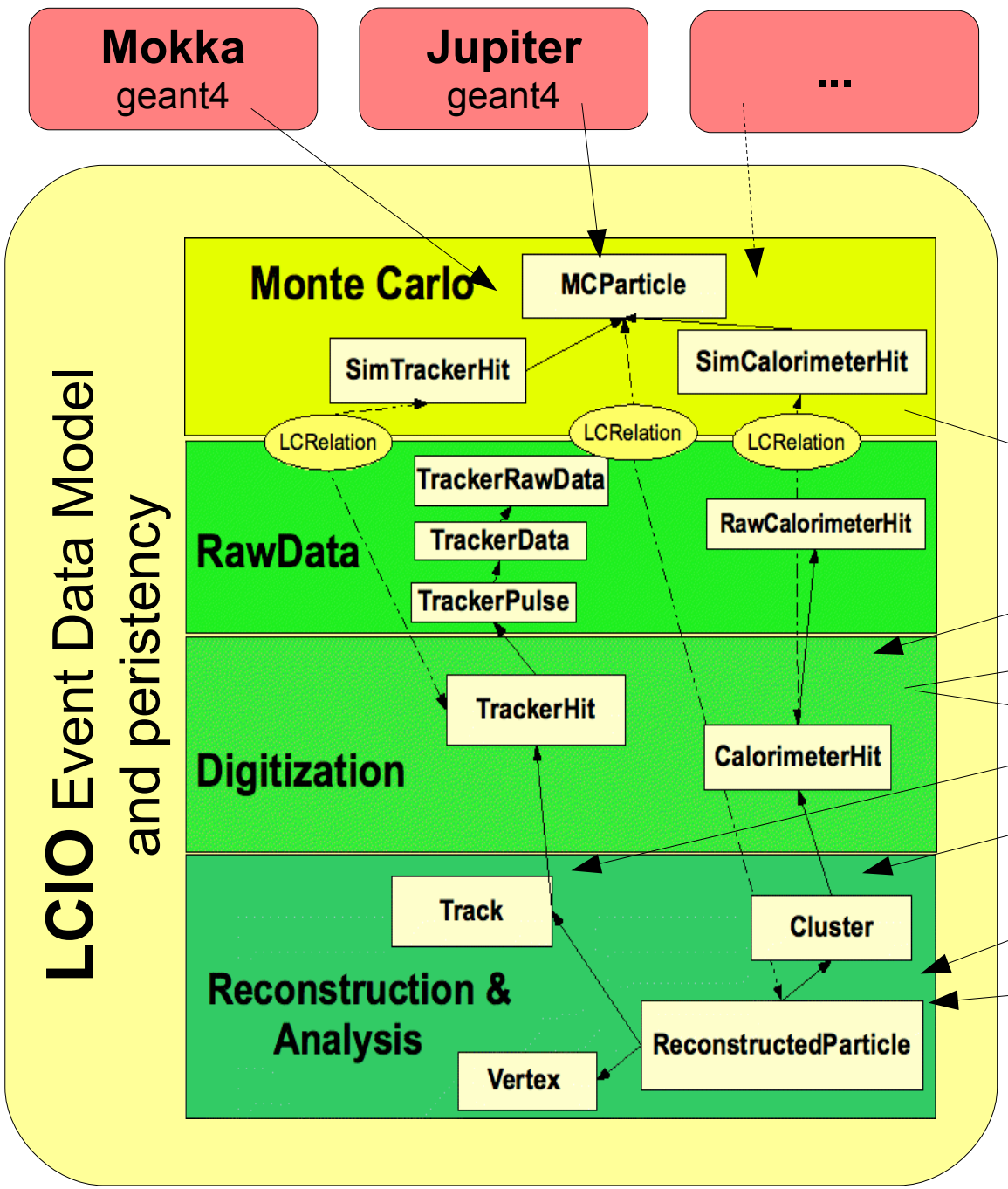
ILD



LDC sw-framework overview



overview LCIO & Marlin



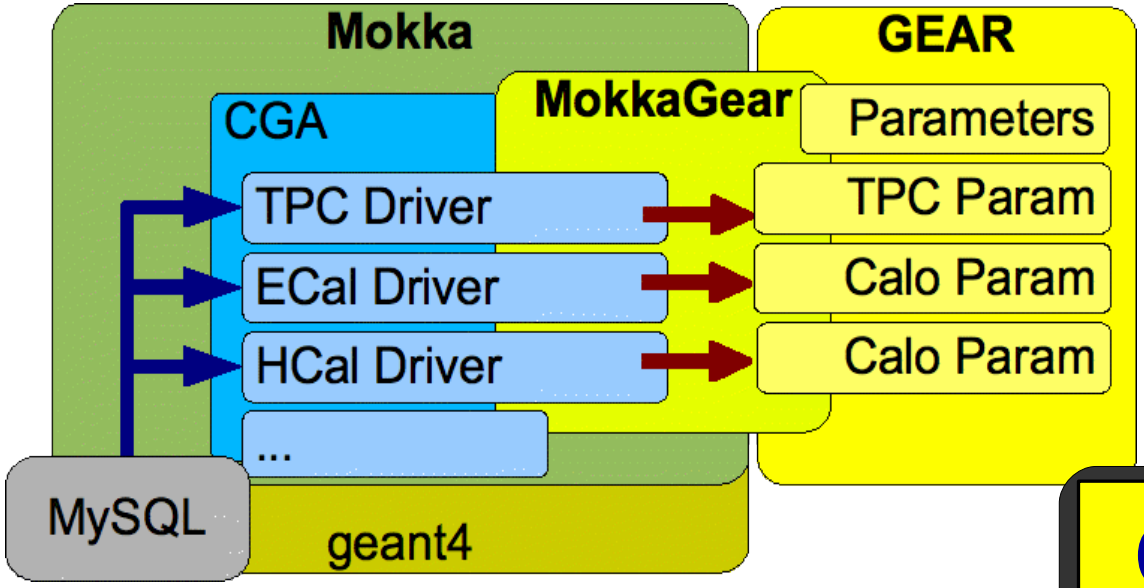
- **LCIO** provides a common data model for ILC studies (Java, C++, f77)
- **Marlin** is modular C++ application framework based on LCIO

-> all packages that speak **LCIO** can use Marlin modules

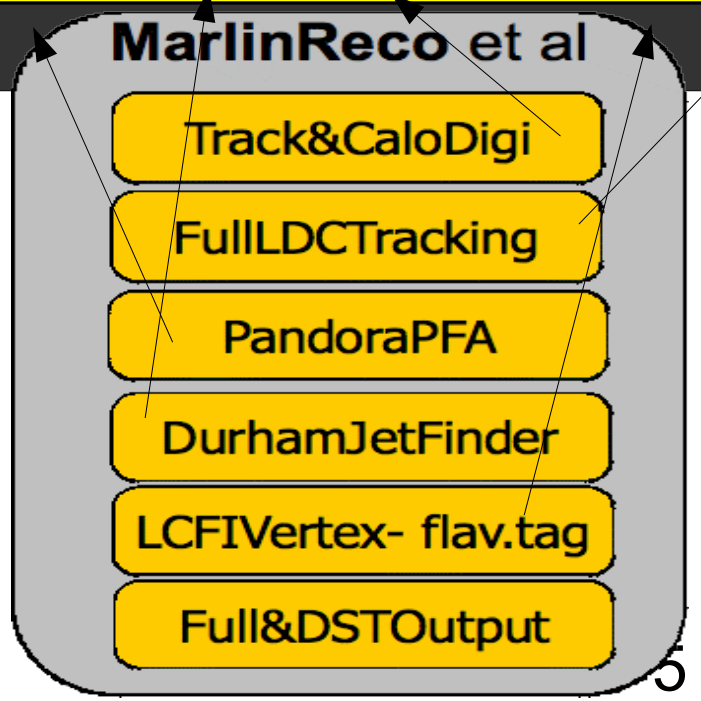
MarlinReco et al

- Track&CaloDigi
- FullLDCTracking
- PandoraPFA
- DurhamJetFinder
- LCFIVertex- flav.tag
- Full&DSTOutput

GEAR-Geometry description



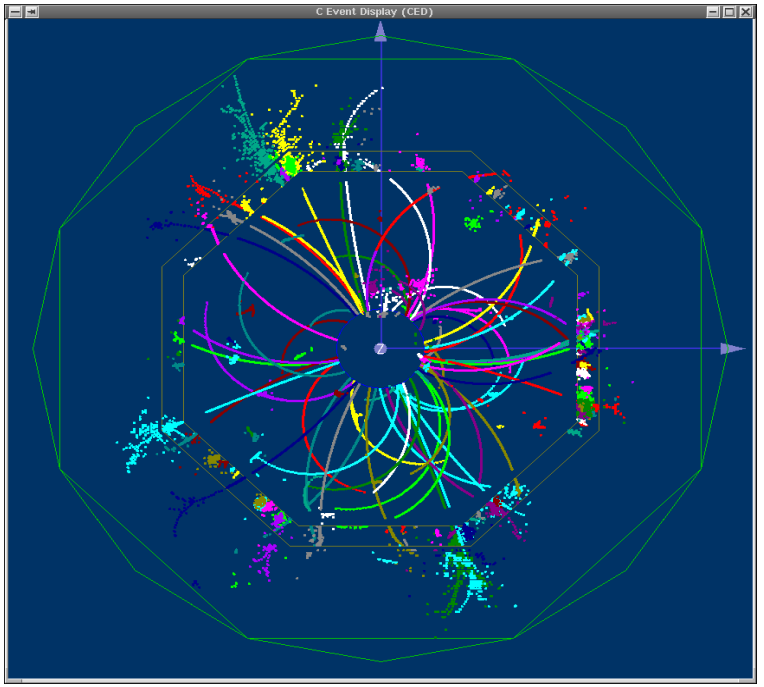
GEometry **A**PI for **R**econstruction

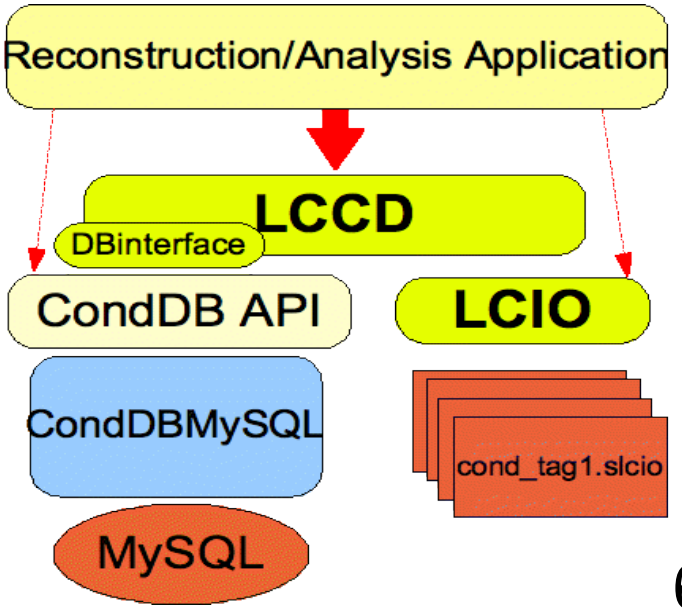


- Mokka writes GEAR xml files w/ **complete** geometry needed for reconstruction
  -> keep track of variations (optimization)
- GEAR provides **abstract API** for accessing geometry (+ some navigation) in Marlin
- xml-files provide **decoupling** from simulation engine -> all groups can use Marlin modules if appropriate GEAR xml file is created (even manually)

Marlin et al support packages

- MarlinUtil
  - Utility and Helper classes
    - helix fitter, cluster shapes,...
- RAIDA
  - AIDA histogramming w/ root
- CED
  - event display based on OpenGL
  - client server architecture
- CEDViewer
  - event display client processors
    - CEDViewer, GenericViewer
- LCCD
  - conditions data toolkit
- Overlay
  - overlay events (background)





Applications of Marlin et al

- LDC detector optimization (MonteCarlo)
  - MarlinReco – full reconstruction suite
    - Digitization Calo,TPC, Silicon, PatternRecognition/Tracking, clustering, ParticleFlow algorithms: **Wolf**, **TrackBased**
  - PandoraPFA
    - ParticleFlow algorithm
  - LCFIVertex
    - ZVTop/ZVKin vertex finding and fitting algorithms
  - various physics analyses ...
- testbeams (Data & MonteCarlo)
  - Calice - calorimeter
  - MarlinTPC – TPC tracking
  - EUTelescope – pixel telescope for silicon tracking

using the same core framework for MC/offline and testbeam/online facilitates exchange of knowledge !

recent developments

- LCIO (v01-10)
  - CalorimeterHit::getEnergyError()
  - added event weight
  - improved reading of stdhep
  - C++ version:
    - LCWriter::setCompressionLevel(I)
      - -> can write uncompressed files
    - simple direct access
      - -> use for background overlay
    - improved handling of ParticleIDs
    - started automation of tests
- CMake
  - new easy to configure cross platform build tool
  - build shared libraries (plugins)
  - all ilcsoft tools use CMake now
    - -> old makefiles will be phased out !
- Marlin (v00-10-01)
  - logging mechanism
  - plugins (shared libraries)
  - overlay mechanism
  - check of detector model in gear
- GEAR (v00-09)
  - GearMgr::getDetectorName()
    - -> use for consistency checks
  - added HcalRing and EcalPlug
  - bug fixes (VXDParameters ...)

•focus of development for core
•tools has been support for the
•planned LOI mass production
  • add features
  • fix bugs
as quickly as needed
•

ilcsoft installation

J.Engels

- **ilcinstall** (v01-03-05)
  - python script to install all of the LDC software in one go
  - **"start script – go to lunch – run application"**
  - fully configurable:
    - versions, dependencies/build options, links to existing packages/tools, e.g. root, CLHEP,...
  - used for reference installations in afs (SL3/SL4)
  - user can link their packages against these
  - even w/o installing any software on their computer

new manual for using ilcinstall, CMake and reference installations
http://ilcsoft.desy.de/portal/general_documentation

ilcsoft release v01-03-05

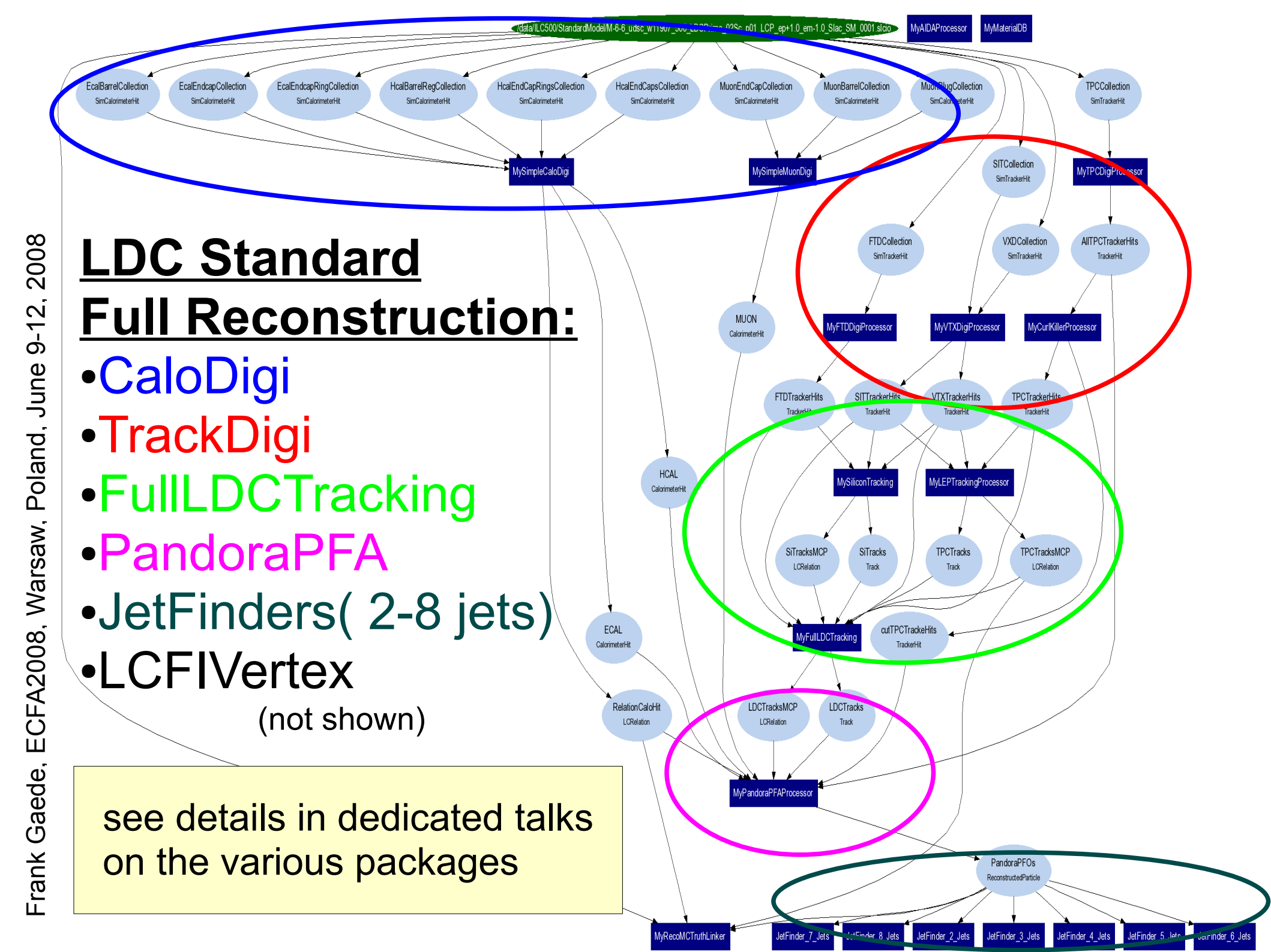
```
lcio/v01-10
gear/v00-09
Marlin/v00-10-01
lccd/v00-03-06

MarlinUtil/v00-08
CED/v00-03
CEDViewer/v00-05
Overlay/v00-02
SiliconDigi/v00-04-01
RAIDA/v01-04-02
StandardConfig/v00-01

MarlinReco/v00-07
PandoraPFA/v02-01-dev
LCFIVertex/v00-02-03-dev
LCFI_SGVbasedNets/v00-01
```

```
MarlinTPC/v00-02-06
Eutelescope/v00-00-06

Mokka/mokka-06-06-p03

CMakeModules/v01-07
CLHEP/2.0.3.2
QT/4.2.2
root/5.16.00
CondDBMySQL/CondDBMySQL_ILC-0-5-10
mysql/5.0.26
cernlib/2006
gsl/1.8
java/1.5.0
```

/afs/desy.de/group/it/ilcsoft/v01-03-05

reconstruction w/ Marlin et al

- ILD plans a massive Monte Carlo production for the LOI
- in recent month the focus of the ILD software groups has been on improving the Mokka simulation for that
  - -> Mokka-06-06-p03 & LDCPrime_02Sc model frozen
- no we need to focus on the reconstruction chain:
  - need to improve & fix the software
  - need standard reconstruction steering file:
    - best set of algorithms
    - best parameters for the algorithms
    - digitization constants
    - calibration constants
- released package StandardConfig v00-01 (ilcsoft v01-03-05)
  - provides Marlin steering file with best configuration from experts to date
  - will have to be tested & improved in coming weeks
  - use as starting point for your own reconstruction

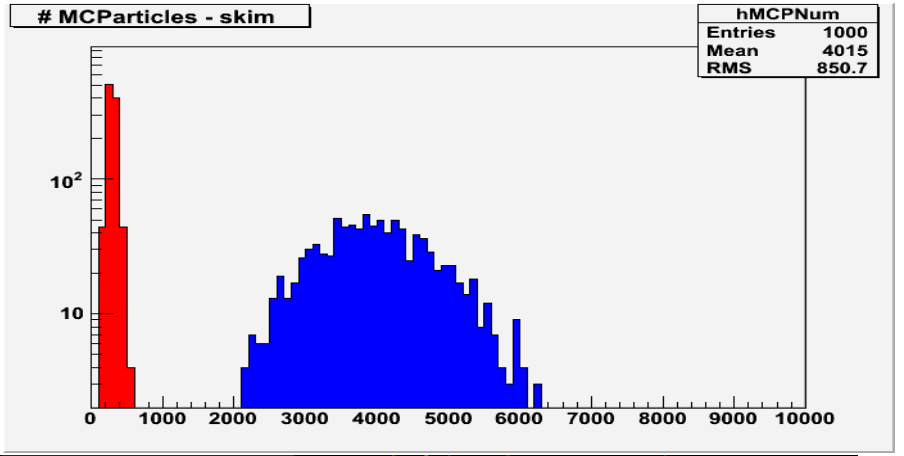

LDC Standard
Full Reconstruction:
•CaloDigi
•TrackDigi
•FullLDCTracking
•PandoraPFA
•JetFinders(2-8 jets)
•LCFIVertex
(not shown)
see details in dedicated talks on the various packages
Frank Gaede, ECFA2008, Warsaw, Poland, June 9-12, 2008
MyAIDAProcessor
MyMaterialDB
EcalBarrelCollection
EcalEndcapCollection
EcalEndcapRingCollection
HcalBarrelRegCollection
HcalEndCapRingsCollection
HcalEndCapsCollection
MuonEndCapCollection
MuonBarrelCollection
TPCCollection
SimCalorimeterHit
SimTrackerHit
MySimpleCaloDigi
MySimpleMuonDigi
SITCollection
FTDCollection
VXDCollection
AllTPCTrackerHits
TrackerHit
MyFTDDigiProcessor
MyVTXDigiProcessor
MyCurlKillerProcessor
MUON
CalorimeterHit
FTDTrackerHits
SITTrackerHits
VTXTrackerHits
TPCTrackerHits
HCAL
MySiliconTracking
MyLEPTrackingProcessor
SiTracksMCP
SiTracks
TPCTracks
TPCTracksMCP
LCRelation
Track
ECAL
MyFullLDCTracking
cutTPCTrackeHits
RelationCaloHit
LDCTracksMCP
LDCTracks
MyPandoraPFAProcessor
PandoraPFOs
ReconstructedParticle
MyRecoMCTruthLinker
JetFinder_7_Jets
JetFinder_8_Jets
JetFinder_2_Jets
JetFinder_3_Jets
JetFinder_4_Jets
JetFinder_5_Jets
JetFinder_6_Jets

'DSTs' with LCIO I

- at Zeuthen ILD meeting agreed that we want DST files with reconstructed data for the LOI mass production
- DST content:
  - Tracks
  - Clusters
  - ReconstructedParticles
    - PFOObjects (Pandora)
    - Jets: 2-8 (!) jets (DurhamNJet algorithm)
  - Vertices & flavour tag per jet
  - MCParticles (Monte Carlo truth)
  - LCRelation (Reco-MCTruth link)

'DSTs' with LCIO II

- DSTs produced with
  - StandardConfig/mc2008/stdreco.xml
- apply MCParticle skim:
  - store full generated event
  - + reconstucted particles & parents
  - decays in flight & conversions
- store LCFIVertex flavour tag in ParticleID objects (C.Lynch)
  - flavour tags b,c,b-bg
  - NN input quantities
  - true jet flavour & charge




```
////////////////////////////////
EVENT: 0
RUN: 2011907
DETECTOR: LDCPrime_02Sc
COLLECTIONS: (see below)
////////////////////////////////
```

COLLECTION NAME	COLLECTION TYPE	NUMBER OF ELEMENTS
FTFinal_2Jets	ReconstructedParticle	2
FTFinal_3Jets	ReconstructedParticle	3
FTFinal_4Jets	ReconstructedParticle	4
FTFinal_5Jets	ReconstructedParticle	5
FTFinal_6Jets	ReconstructedParticle	6
FTFinal_7Jets	ReconstructedParticle	7
FTFinal_8Jets	ReconstructedParticle	8
IPVertex	Vertex	1
LDCTracks	Track	70
MCParticlesSkimmed	MCParticle	202
PandoraClusters	Cluster	64
PandoraPFOs	ReconstructedParticle	65
RecoMCTruthLink	LCRelation	65
ZVRESVertices_2Jets	Vertex	1
ZVRESVertices_3Jets	Vertex	1
ZVRESVertices_4Jets	Vertex	1
ZVRESVertices_5Jets	Vertex	1
ZVRESVertices_6Jets	Vertex	1
ZVRESVertices_7Jets	Vertex	1
ZVRESVertices_8Jets	Vertex	1

some numbers udsc @ 500 GeV		
type	kB/evt	f_I/O /Hz
SIM	950	10
REC	1800	3
DST	23	250

Note: f_I/O numbers are examples only - simple Marlin job on my PC

Monte Carlo production on the Grid

Proposed (preliminary)		
Process	fb^{-1}	#events
ee->6f	500	1197236
ee->4f	50	3358252
ee->2f	20	1192784
ee->hX	500	299278
nn(n*g)	20	841726
ee->ee	0.1	6953510
eg->eg	0.1	344270
gg->X	0.1	554782
ee->gg(n*g)	10	306954
rest	1	517376
Total		**15566168**

Sub-Detector	Parameter	GLD	LDC	GLD'	LDC'
TPC	R_{inner} (m)	0.45	0.30	0.45	0.30
	R_{outer} (m)	2.00	1.58	1.80	1.80
	Z_{max} (m)*	2.50	2.16	2.35	2.35
Barrel ECAL	R_{inner} (m)**	2.10	1.60	1.85	1.82
	Material	Sci/W	Si/W	Sci/W	Sci/W
Barrel HCAL	Material	Sci/W	Sci/Fe	Sci/Fe	Sci/Fe
Endcap ECAL	Z_{min} (m)***	2.80	2.30	2.55	2.55
Solenoid	B-field	3.0	4.0	3.50	3.50
VTX	Inner Layer (mm)	20	16	18	18

M.Thomson

- use the Grid to produce a significant Monte Carlo data set
- use Standard Model generator files produced at SLAC
- -> produce 15M events for LDCPrime configuration
- + signal samples, detector variations,....
- -> computing infrastructure set up by DESY group
  - job submissions scripts, databases, monitoring tools,...
  - grid software installations

grid sites with ilcsoft v01-03-05

J.Engels

CE	SW-VER	SW-OS	DATE	TIME	SAM	JOB	TAGGED	HIST-LOGS
cclcgceli01.in2p3.fr	v01-03-05	sl4	2008-06-09	12-18-27	OK	OK	VO-ilc-ilcsoft-v01-03-03-sl4 VO-ilc-ilcsoft-v01-03-05-sl4	History
cclcgceli02.in2p3.fr	v01-03-05	sl4	2008-06-09	12-18-27	OK	OK	VO-ilc-ilcsoft-v01-03-03-sl4 VO-ilc-ilcsoft-v01-03-05-sl4	History
ce01.dur.scotgrid.ac.uk	v01-03-05	sl4	2008-06-09	12-18-27	OK	OK	VO-ilc-ilcsoft-v01-03-05-sl4	History
ce02.tier2.hep.manchester.ac.uk	v01-03-05	sl3	2008-06-09	14-17-04	OK	OK	VO-ilc-ilcsoft-v01-03-03-sl3 VO-ilc-ilcsoft-v01-03-05-sl3	History
ce1.pp.rhul.ac.uk	v01-03-05	sl3	2008-06-09	14-17-04	OK	OK		History
ce2.ppgrid1.rhul.ac.uk	v01-03-05	sl4	2008-06-09	12-18-27	OK	OK	VO-ilc-ilcsoft-v01-03-05-sl4	History
dg10.cc.kek.jp	v01-03-05	sl3	2008-06-09	14-17-04	OK	OK	VO-ilc-ilcsoft-v01-03-03-sl3 VO-ilc-ilcsoft-v01-03-04-sl3 VO-ilc-ilcsoft-v01-03-05-sl3	History
fal-pygrid-18.lancs.ac.uk	v01-03-05	sl4	2008-06-09	12-18-27	OK	OK	VO-ilc-ilcsoft-v01-03-03-sl4 VO-ilc-ilcsoft-v01-03-05-sl4	History
grid-ce3.desy.de	v01-03-05	sl4	2008-06-09	12-18-27	OK	OK	VO-ilc-ilcsoft-v01-03-03-sl4 VO-ilc-ilcsoft-v01-03-05-sl4	History
grid10.lal.in2p3.fr	v01-03-05	sl4	2008-06-09	12-18-27	OK	OK	VO-ilc-ilcsoft-v01-03-03-sl4 VO-ilc-ilcsoft-v01-03-05-sl4	History
lcg-ce0.ifh.de	v01-03-05	sl4	2008-06-09	12-18-27	OK	OK	VO-ilc-ilcsoft-v01-03-03-sl4 VO-ilc-ilcsoft-v01-03-05-sl4	History
lcg-ce1.ifh.de	v01-03-05	sl4	2008-06-09	12-18-27	OK	OK	VO-ilc-ilcsoft-v01-03-03-sl4 VO-ilc-ilcsoft-v01-03-05-sl4	History
lcgce02.gridpp.rl.ac.uk	v01-03-05	sl4	2008-06-09	12-18-27	OK	OK	VO-ilc-ilcsoft-v01-03-03-sl4 VO-ilc-ilcsoft-v01-03-05-sl4	History
polgrid1.in2p3.fr	v01-03-05	sl4	2008-06-09	12-18-27	OK	OK	VO-ilc-ilcsoft-v01-03-03-sl4 VO-ilc-ilcsoft-v01-03-05-sl4	History

x

- http://ilcsoft.desy.de/grid/results/User.html

First round of the mass production

- first round of production started last week (~4 days)
- Detector Model **LDCPrime_02Sc** and **Mokka-06-06-p03**
- currently : ~10-15% of the final sample
- expect ~2 months in total

Proposed (preliminary)			Produced		
Process	fb^{-1}	#events	Process	fb^{-1}	#events
ee->6f	500	1197236	ee->6f	108	258462
ee->4f	50	3358252	ee->4f	7	452332
ee->2f	20	1192784	ee->2f	1.5	86016
ee->hX	500	299278	ee->hX	27	16277
nn(n*g)	20	841726	nn(n*g)	4	173508
ee->ee	0.1	6953510	ee->ee	0	748
eg->eg	0.1	344270	eg->eg	0	0
gg->X	0.1	554782	gg->X	0	2005
ee->gg(n*g)	10	306954	ee->gg(n*g)	0	0
rest	1	517376	rest	0	0
			calibration		7400
Total		**15566168**			**996748**

event processing times

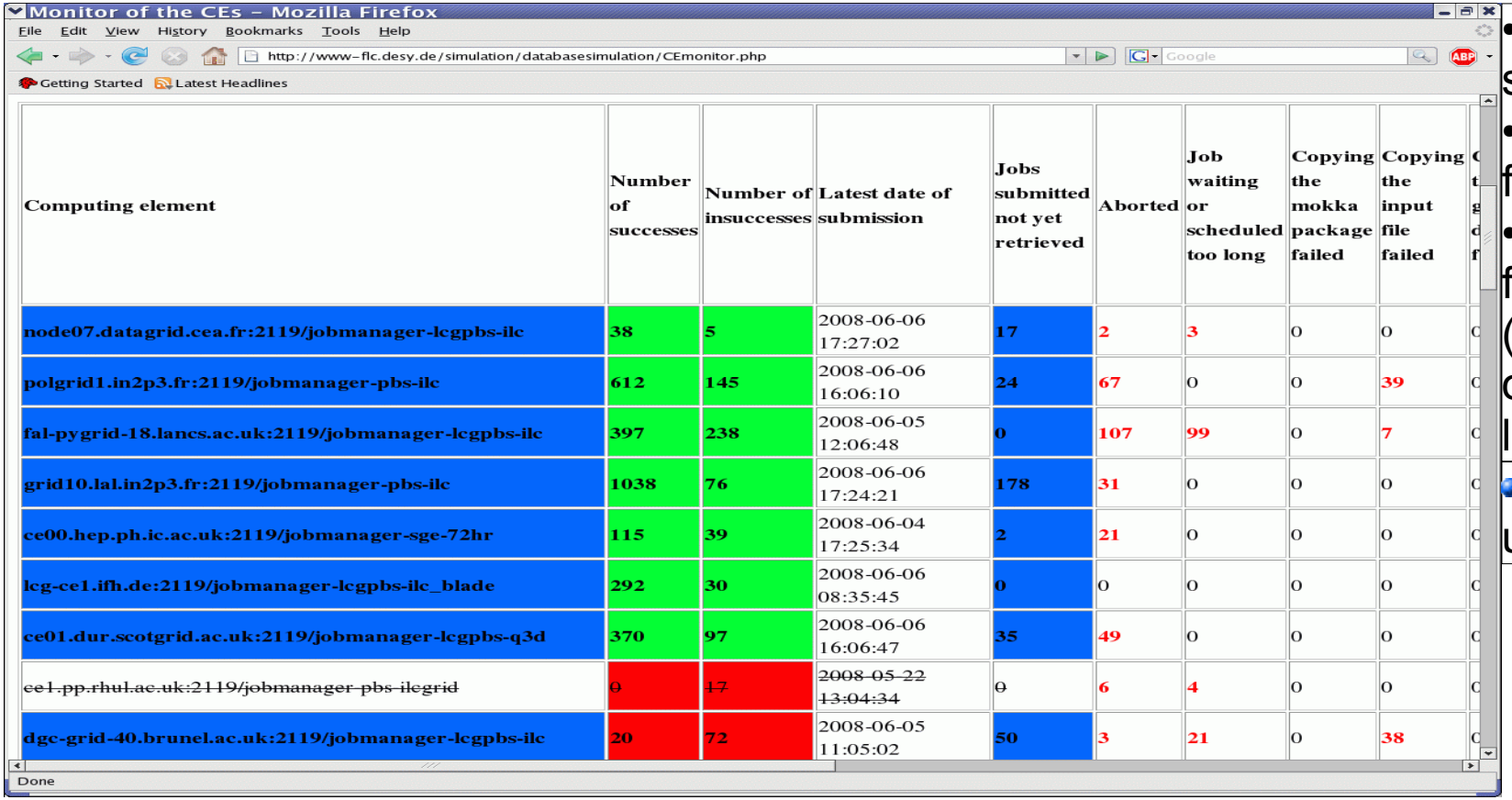
Average simulation time	
Final state	**average time/event (minutes)**
ee->6f	**3.69**
ee->4f	**1.99**
ee->2f	**1.9**
ee->hX	**0.84**
nn(n*g)	**0.16**
bbuddu	3.74
cccsdu	4.07
bbn3e3du	2.46
n1e1sc	2.39
uddu	2.69
e1e1	1.23
e2e2	0.42
e3e3	0.97
uu	1.82
cc	2.15
dd	1.95
ss	2.07
bb	2.05
n1n1h	0.69

- Some examples of Mokka simulation time for this first statistic.

grid production - efficiency I

- Bug in the grid scripts had to be corrected.
- Nevertheless efficiency (in #jobs) ~87%.
- CE monitor in number of jobs and reasons of failure at:

http://www-flc.desy.de/simulation/databasesimulation/CEmonitor.php



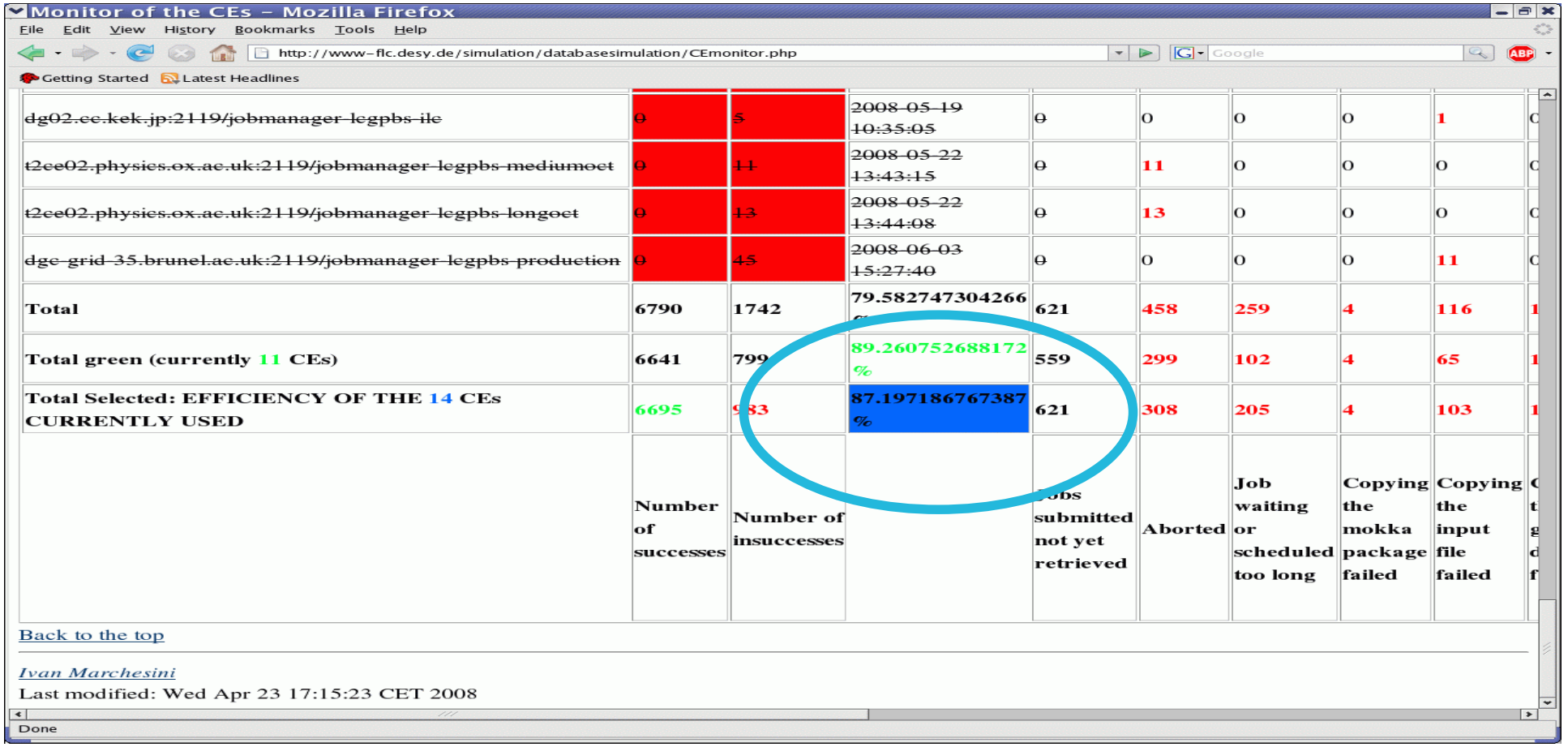
Computing element	Number of successes	Number of insuccesses	Latest date of submission	Jobs submitted not yet retrieved	Aborted	Job waiting or scheduled too long	Copying the mokka package failed	Copying the input file failed	
node07.datagrid.cea.fr:2119/jobmanager-lcgpbs-ilc	38	5	2008-06-06 17:27:02	17	2	3	0	0	0
polgrid1.in2p3.fr:2119/jobmanager-pbs-ilc	612	145	2008-06-06 16:06:10	24	67	0	0	39	0
fal-pygrid-18.lancs.ac.uk:2119/jobmanager-lcgpbs-ilc	397	238	2008-06-05 12:06:48	0	107	99	0	7	0
grid10.lal.in2p3.fr:2119/jobmanager-pbs-ilc	1038	76	2008-06-06 17:24:21	178	31	0	0	0	0
ce00.hep.ph.ic.ac.uk:2119/jobmanager-sge-72hr	115	39	2008-06-04 17:25:34	2	21	0	0	0	0
lcg-ce1.ifh.de:2119/jobmanager-lcgpbs-ilc_blade	292	30	2008-06-06 08:35:45	0	0	0	0	0	0
ce01.dur.scotgrid.ac.uk:2119/jobmanager-lcgpbs-q3d	370	97	2008-06-06 16:06:47	35	49	0	0	0	0
~~ce1.pp.rhul.ac.uk:2119/jobmanager-pbs-ilcgrid~~	~~0~~	~~17~~	~~2008-05-22 13:04:34~~	~~0~~	6	4	0	0	0
dgc-grid-40.brunel.ac.uk:2119/jobmanager-lcgpbs-ilc	20	72	2008-06-05 11:05:02	50	3	21	0	38	0

- Number of successes;
- Number of failures;
- Reasons of failure (introduced only for the latest jobs).

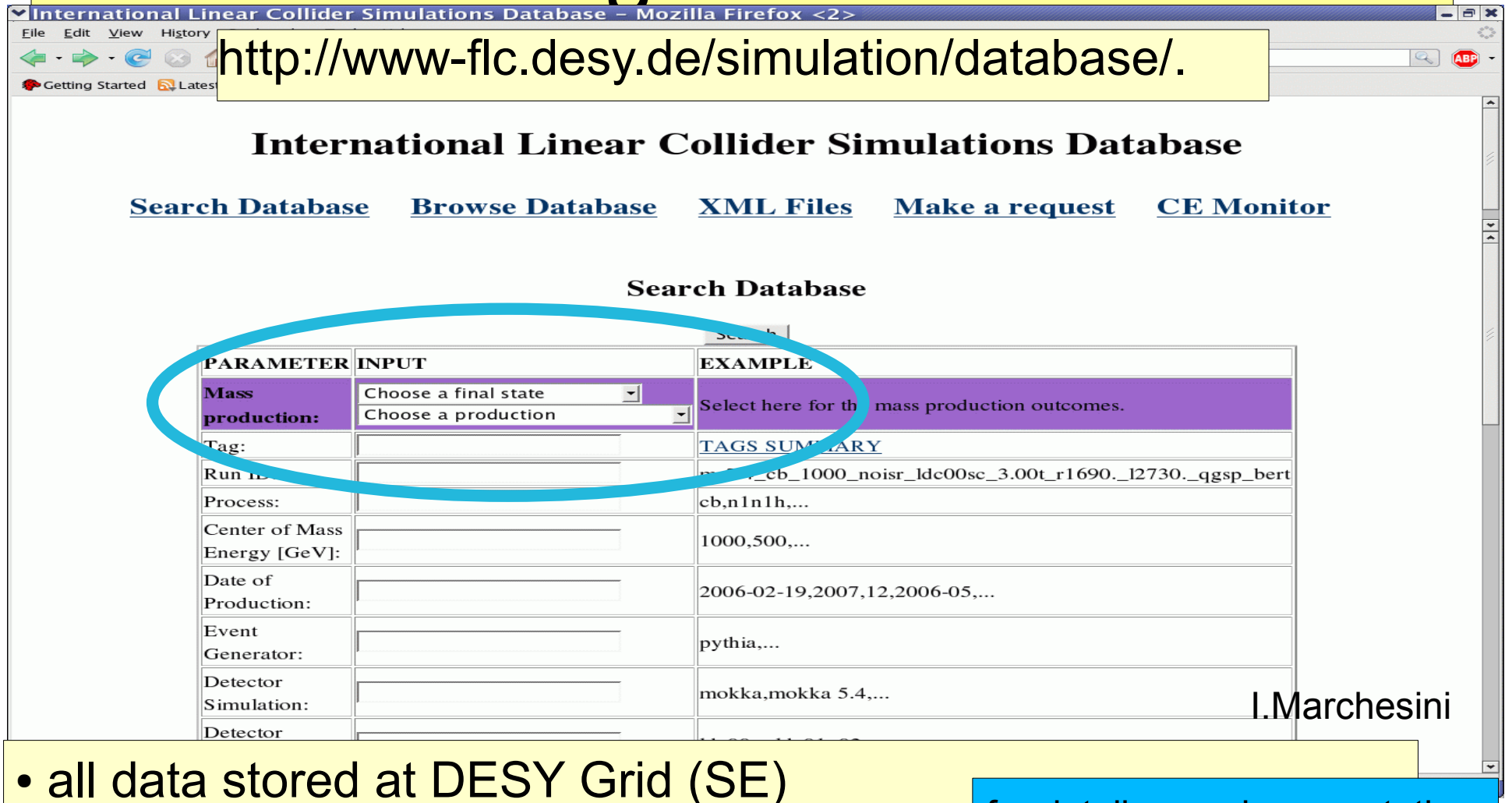
- CE currently used.

grid production - efficiency II

- Summary in the end of the page: currently 14 CEs used.



accessing the data files



I.Marchesini

- all data stored at DESY Grid (SE)
- browse the data catalogue on the web
- -> retrieve logical grid file name
  - copy the data to your computer using Grid tools
  - or analyze the data on the grid

for details see documentation, eg. @ http://grid.desy.de

Summary

- LDC (ILD) has a mature and powerful software framework based on LCIO, Marlin, GEAR,...
  - recent focus is improving the core tools for LOI mass production
  - full Marlin based reconstruction is available w/ MarlinReco, FullLDCTracking, PandoraPFA and LCFIVertex (see other talks)
  - now putting together standard reconstruction (incl. DST files for physics analysis)
  - software reference - and Grid installations (v01-03-05)

Outlook

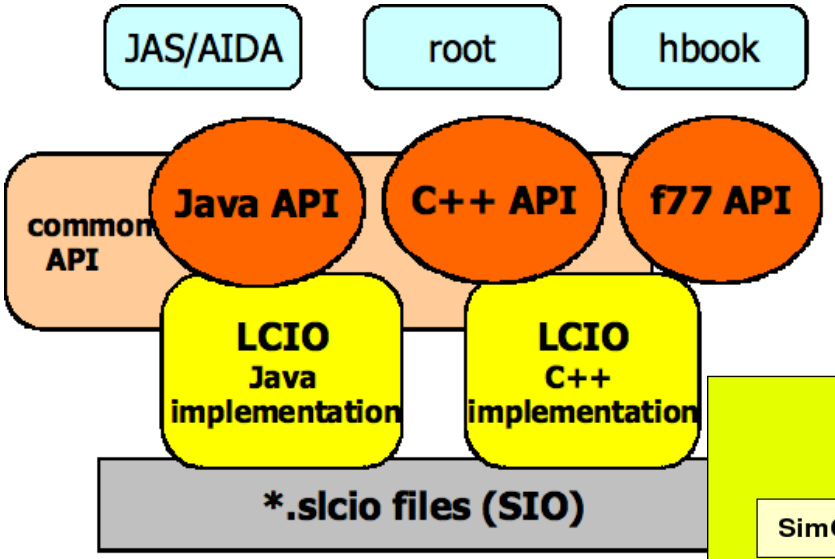
- finish simulation production on the grid
- include signal samples (which?)
- improveand run reconstruction to provide DSTs
- ... optimize detector and write LOI ....

RecoMCTruthLink

- processor in MarlinReco/Analysis/RecoMCTruthLink
- create collection with 1-1 relation from ReconstructedParticle to MCParticle:
  - charged: particle w/ most hits used in track fit
  - neutral: particle / largest energy contribution to cluster
  - weight: relative contribution of MCParticle to Reco.Part.
- can use this for checking the reconstruction algorithms even on DSTs:
  - momentum resolution
  - jet energy resolution:
    - if true jet energy defined as true energy of MCParticles w/ largest contribution to reco particle used
- -> do we need to run perfect PFA for the DST ?

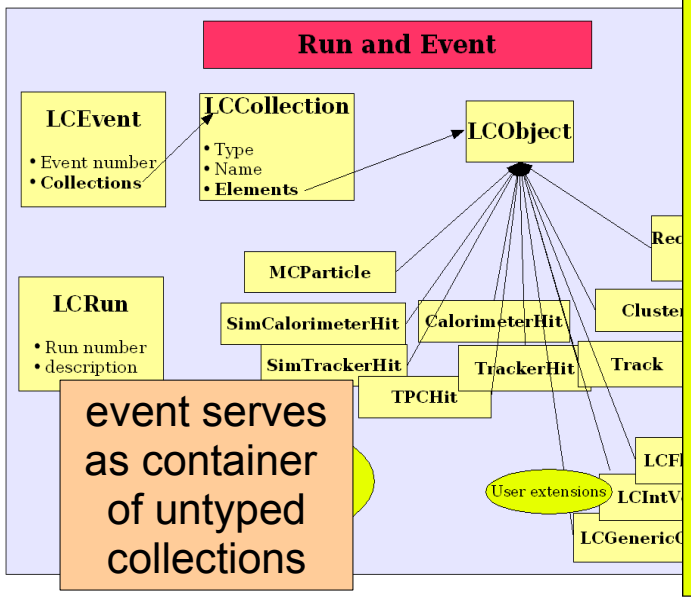
LCIO: persistency & event data model

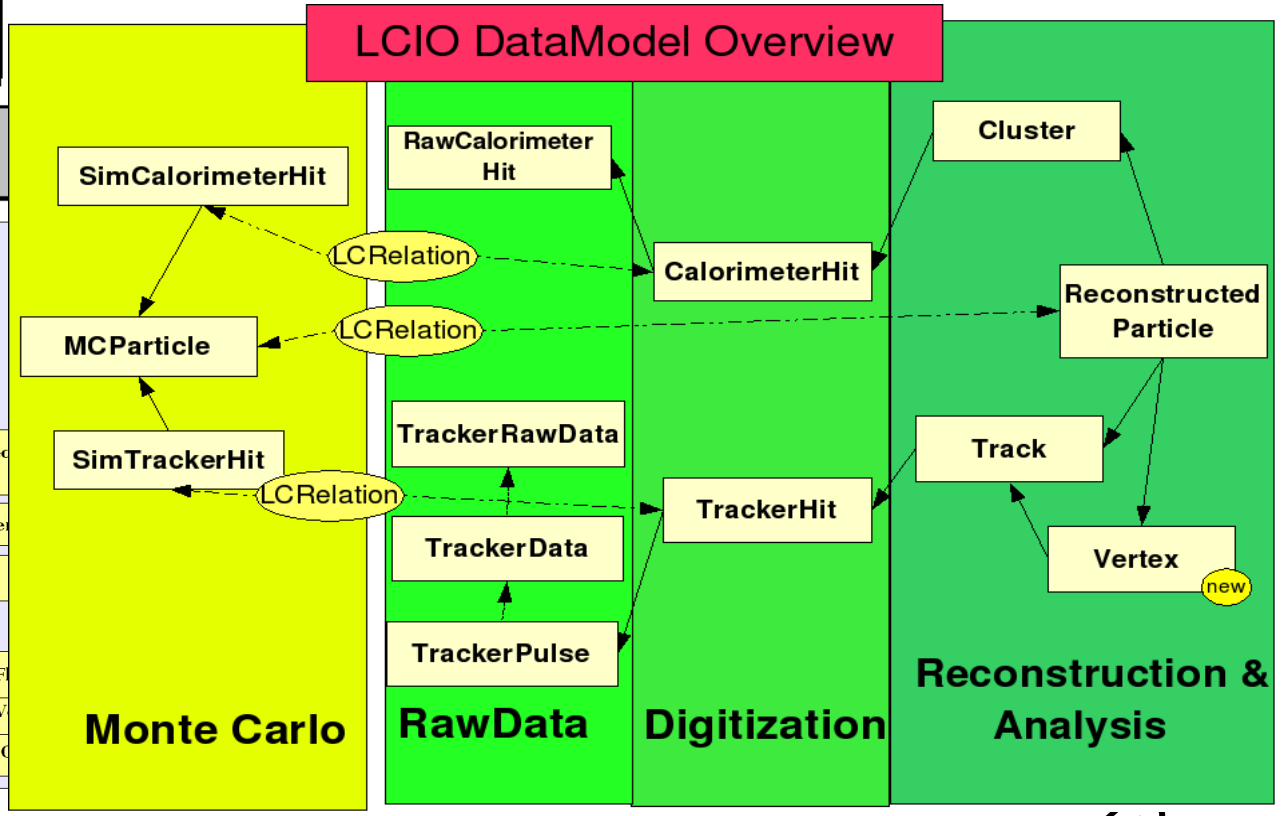
LCIO is the standard output format for LDC/GLD and SID



LCIO recent developments (v01-09):

- preparation for LOI MC-mass production
- improved handling of stdhep generator files
  - bug fixes, stdhep reader
- added event weight to LCEvent
- added experimental support for direct access (C++)
- added runtime extensions (C++)





example – GEAR API VXD

Gear: gear::VXDParameters class Reference - Mozilla Firefox

File Edit View Go Bookmarks Tools Help

http://ilcsoft.desy.de/gear/v00-03/doc/html/classgear_1_1VXDParameters.html

simulation/geant4 LCIO Linux Conferences DESY IT Group LEO English/Ger... Google MyHome Ctime

virtual const **VXDLayerLayout** & **getVXDLayerLayout** () const=0
The layer layout in the Vertex.

virtual int **getVXDType** () const=0
The type of Vertex detector: VXDParameters.CCD, VXDParameters.CMOS or VXDPara

virtual double **getShellHalfLength** () const=0
The half length (z) of the support shell in mm (w/o gap).

virtual double **getShellGap** () const=0
The length of the gap in mm (gap position at z=0).

virtual double **getShellInnerRadius** () const=0
The inner radius of the support shell in mm.

virtual double **getShellOuterRadius** () const=0
The outer radius of the support shell in mm.

virtual double **getShellRadLength** () const=0
The radiation length in the support shell.

virtual bool **isPointInLadder** (**Point3D** p) const=0
returns whether a point is inside a ladder

virtual bool **isPointInSensitive** (**Point3D** p) const=0
returns wheter a point is inside a sensitive volume

virtual **Vector3D** **distanceToNearestLadder** (**Point3D** p) const=0
returns vector from point to nearest ladder

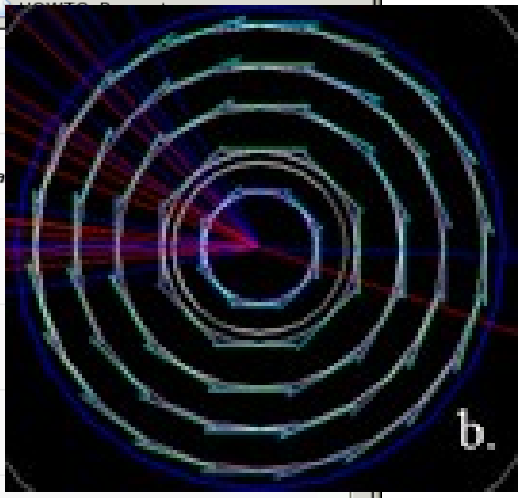
virtual **Vector3D** **distanceToNearestSensitive** (**Point3D** p) const=0
returns vector from point to nearest sensitive volume

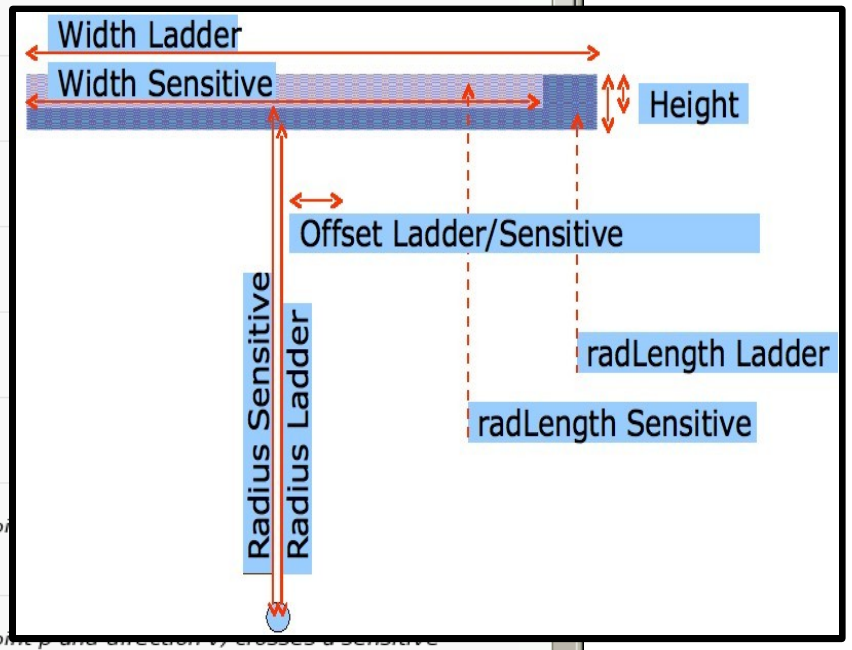
virtual **Vector3D** **intersectionLadder** (**Point3D** p, **Vector3D** v) const=0
returns the first point where a given strainght line (parameters po
is returned if no intersection can be found.

virtual **Vector3D** **intersectionSensitive** (**Point3D** p, **Vector3D** v) const=0
returns the first point where a given strainght line (parameters po
volume (0,0,0) is returned if no intersection can be found.

Find: VXD Find Next Find Previous Highlight Match case

Done





Storage of the files

- Using only the DESY storage element:
  - Generator files 1.7 T;
  - This run of the simulation 603G.

- Files can be accessed via the grid catalogue.

- The information about files can be accessed via the web interface of the database:
  - **http://www-flc.desy.de/simulation/databasesimulation/.**