ILD software

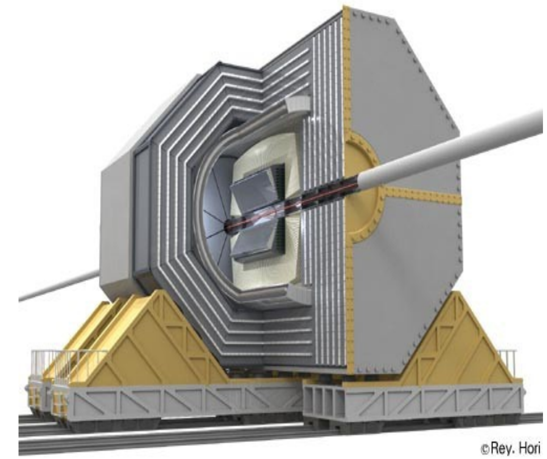
towards the LOI and beyond

Frank Gaede
DESY

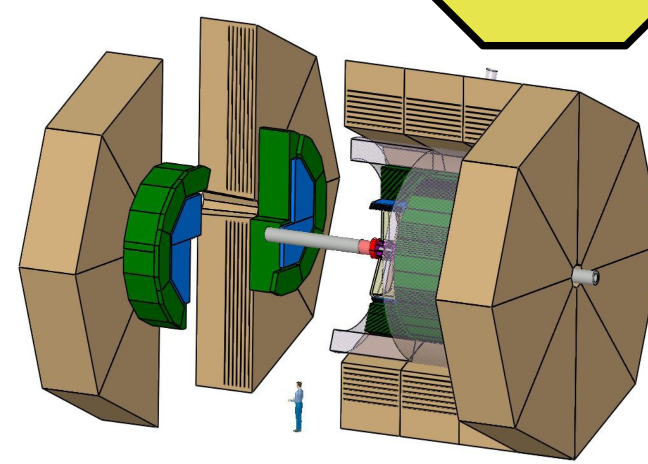
ILD Meeting Cambridge
September 11-13, 2008

Outline

- the past
  - ilc software frameworks
- the recent past
  - LDC & GLD-frameworks
  - detector models
  - (pre?)LOI Monte Carlo production
- the present
  - background studies
- the near future
  - getting ready for the LOI
- the future
  - beyond the LOI

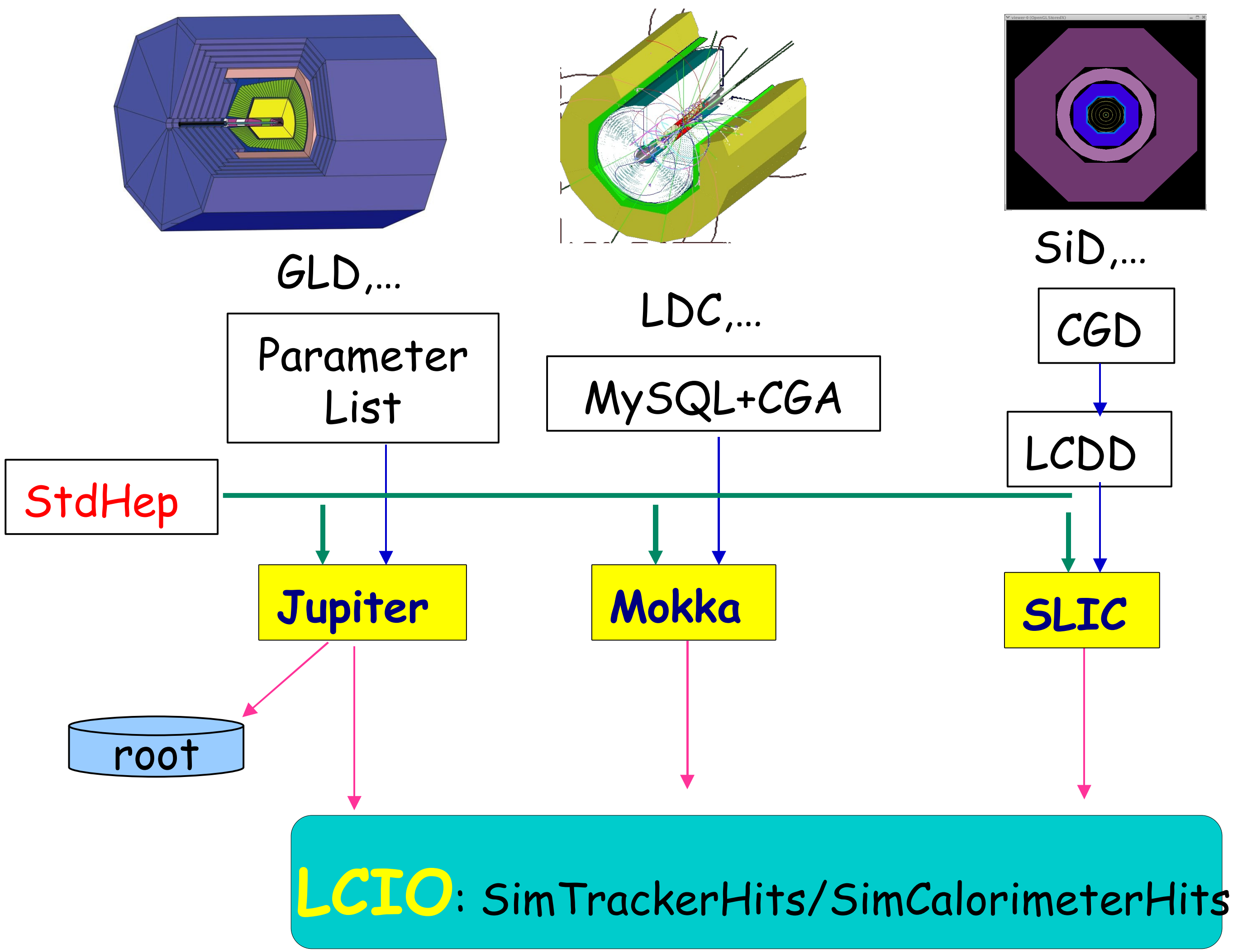


ILD



ILC Simulation Frameworks (Geant4)

- Geant4, StdHep and LCIO are common feature
- Each trying to be generic with different approach
  → different ways to define geometries



A.Miyamoto @LCWS 2005

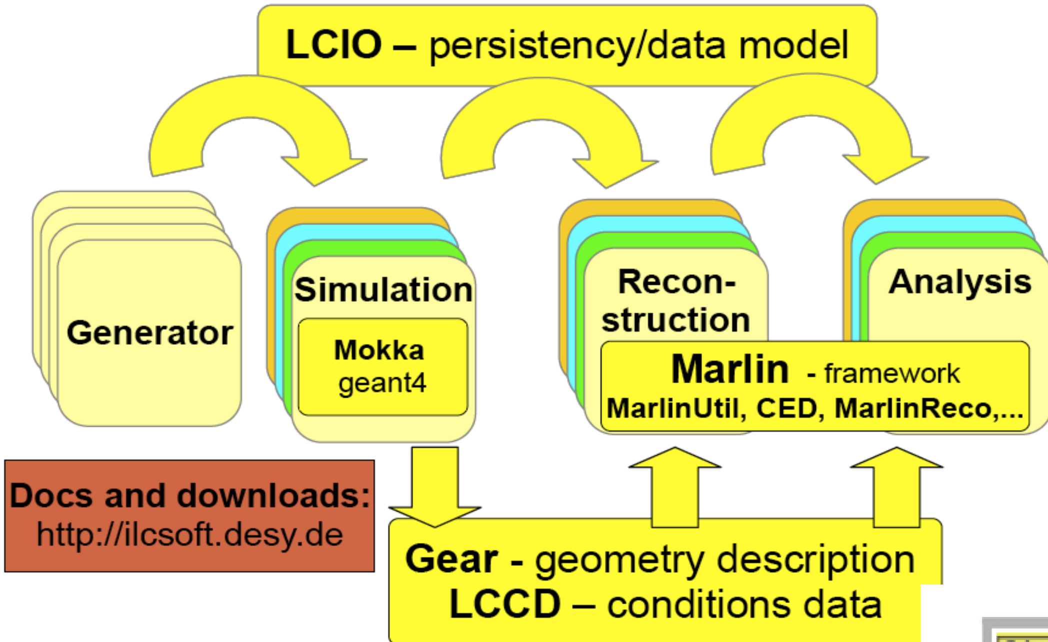
Note: 4th did not yet exist

ILC tools/frameworks (LCWS07)

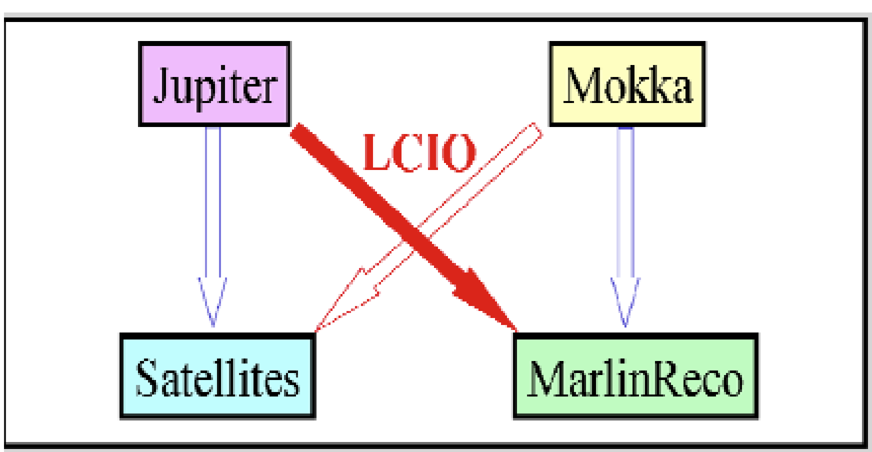
	Description	Detector	Language	IO-Format	users
Simdet	fast Monte Carlo	TeslaTDR	Fortran	**LCIO**	LDC,SID
SGV	fast Monte Carlo	simple Geometry, flexible	Fortran	None/LCIO	LDC,SiLC
LiCToy	fast Monte Carlo	simple trk. Geometry	C++	**LCIO**	LDC,SiLC
Lelaps	fast Monte Carlo	SiD, flexible	C++	SIO, **LCIO**	SID
Mokka	full simulation – Geant4	LDC, flexible	C++	**LCIO**	LDC
SLIC	full simulation – Geant4	SiD, flexible	C++	**LCIO**	SID
Jupiter	full simulation – Geant4	GLD	C++	Root/**LCIO**	GLD
ILCroot	full sim. – Geant4/Flukka/g3	4th	C++	Root	4th
Marlin	reconstruction and analysis application framework	Flexible	C++	**LCIO**	LDC
org.lcsim	reconstruction framework	SiD (flexible)	Java	**LCIO**	SID
Jupiter-Satelites	reconstruction and analysis	GLD	C++	Root	GLD
ILCroot	reconstruction and analysis	4th	C++	Root	4th
LCCD	Conditions Data Toolkit	All	C++	**LCIO**	LDC,Calice,...
GEAR	Geometry description	Flexible	C++	XML	LDC,Calice,...
LCIO	Persistency and datamodel	All	Java, C++, Fortran	**LCIO**	LDC,SID, GLD,Calice,...
JAS3/WIRED	Analysis Tool / Event Display	All	Java	xml,stdhep, eprep,LCIO,..	SID
root	Analysis Tool / Event Display	All	C++	Root	LDC,GLD,4th

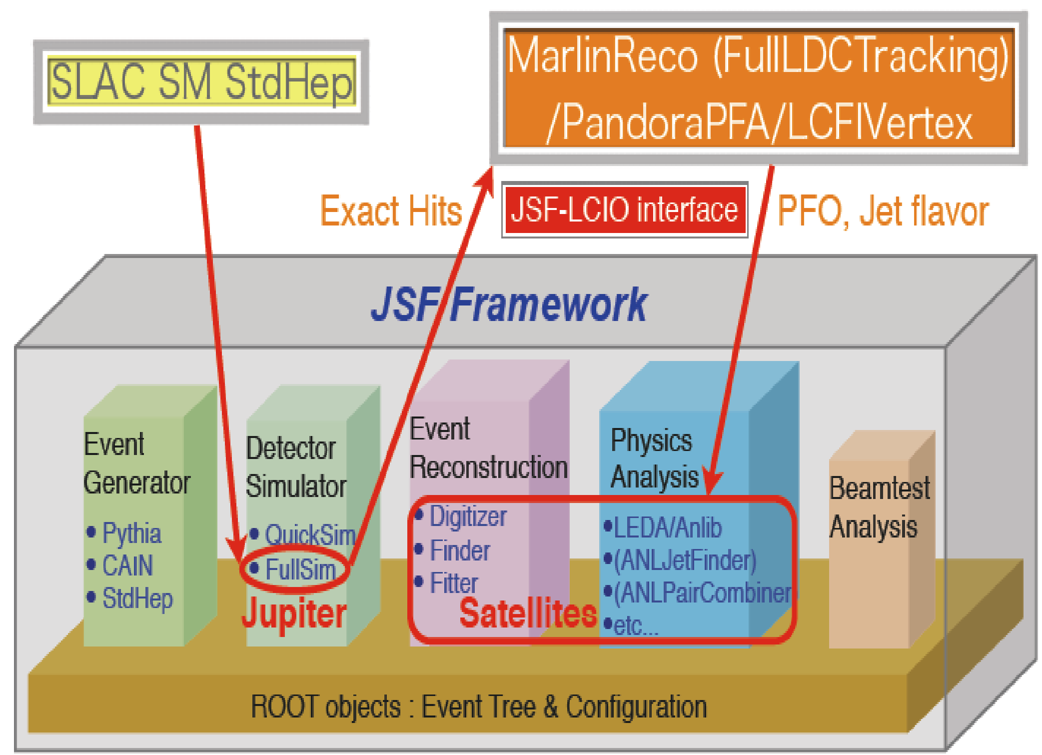
LCIO: basis for 'horizontal' collaboration

LDC & GLD sw frameworks

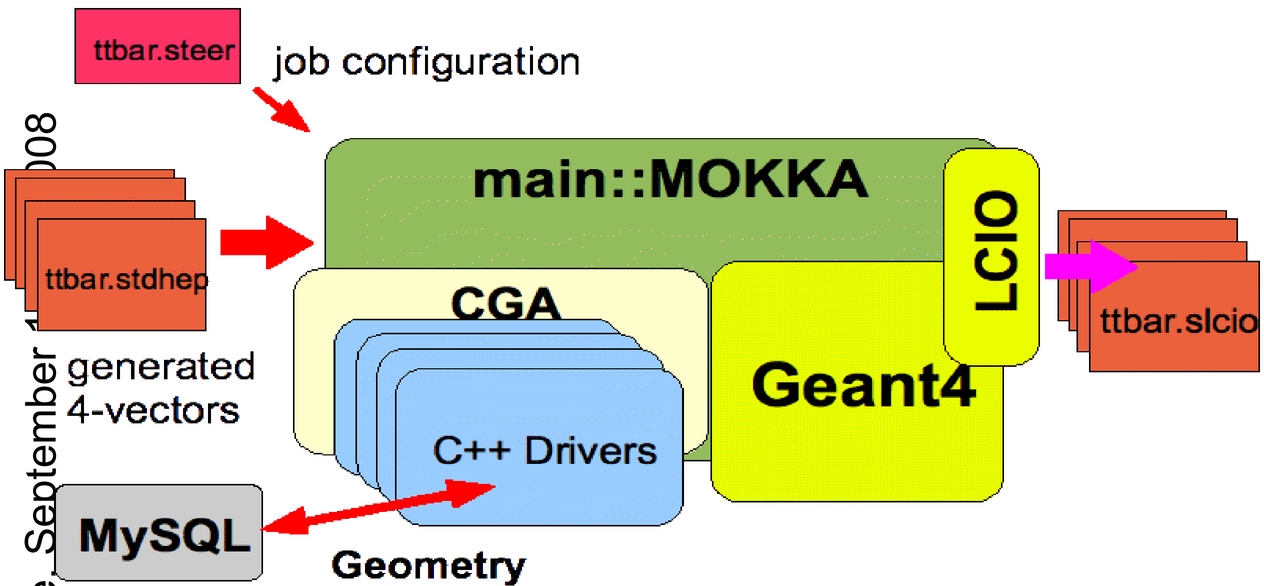


- two independent frameworks had been developed in the 2 regions
- -> LCIO & GEAR provide basis for interoperability

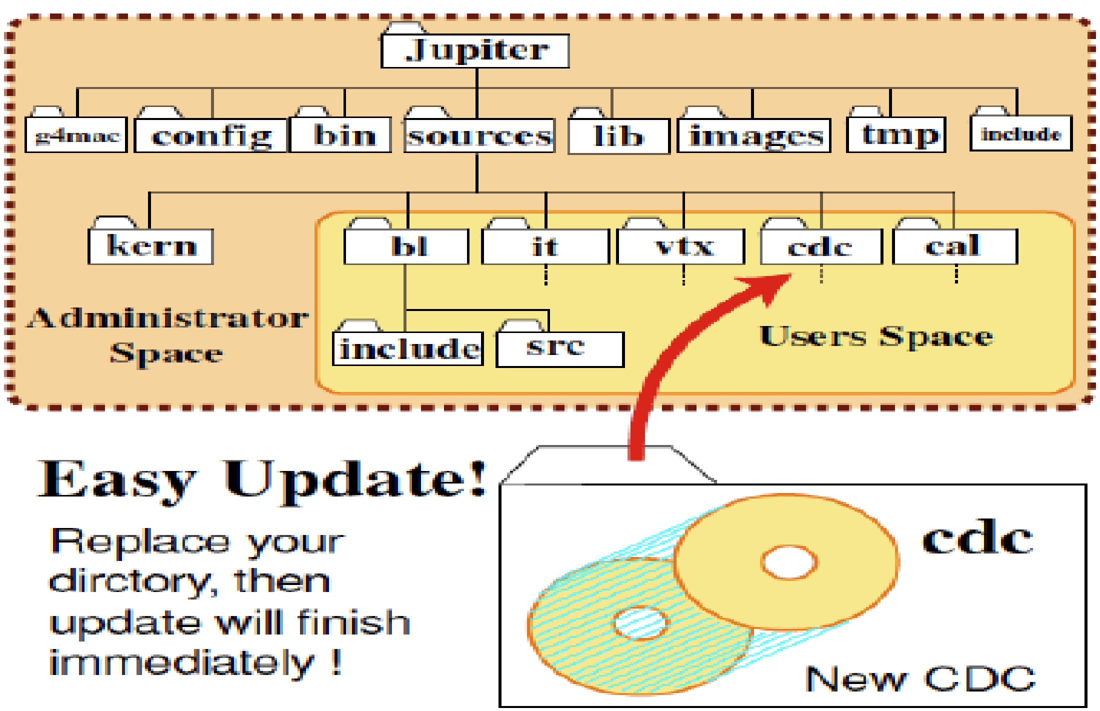




Mokka/Jupiter Geometry

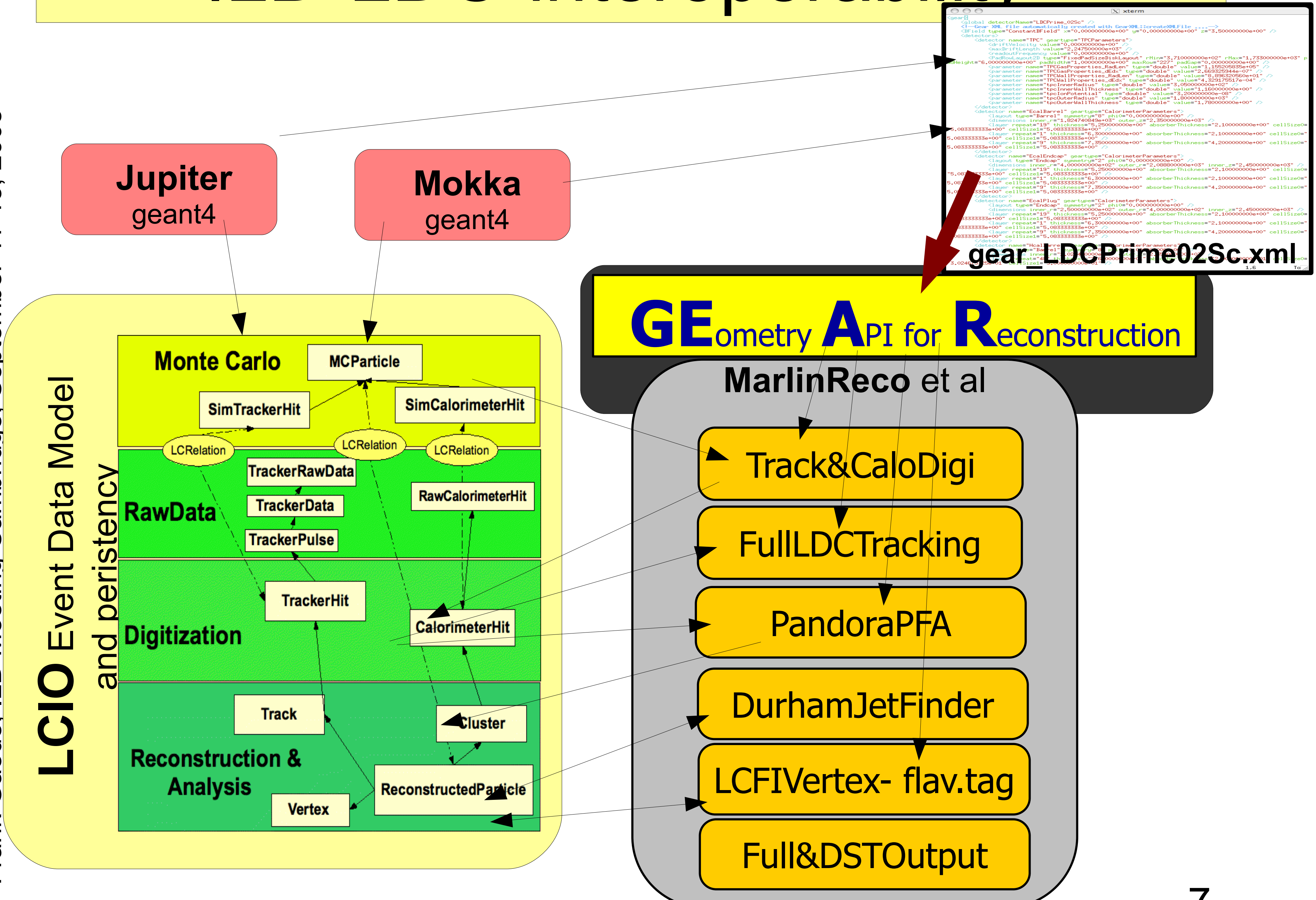


- Mokka geometry definition:
  - parameters in MySQL-db
  - actual placement in C++
  - flexible ,'perfect' bookkeeping
  - partly engineering level of details (Ecal,Hcal,VTX,..)
    - stored in GEAR for reco



- Jupiter geometry definition:
  - parameters ascii files
  - actual placement in C++ drivers
  - flexible, easy to change
  - simple geometry for fast optimization
    - stored in root for reco

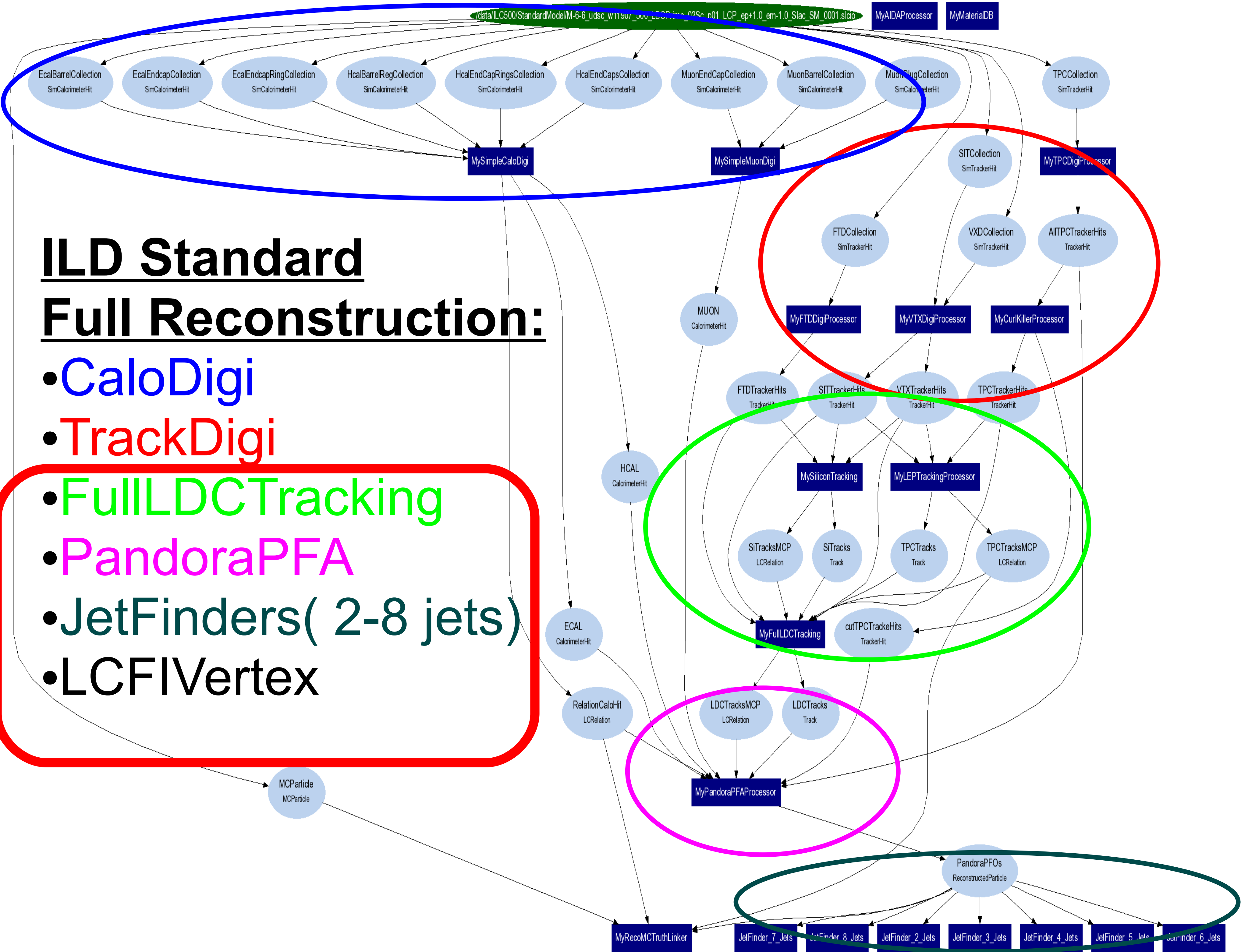
ILD-LDC-interoperability



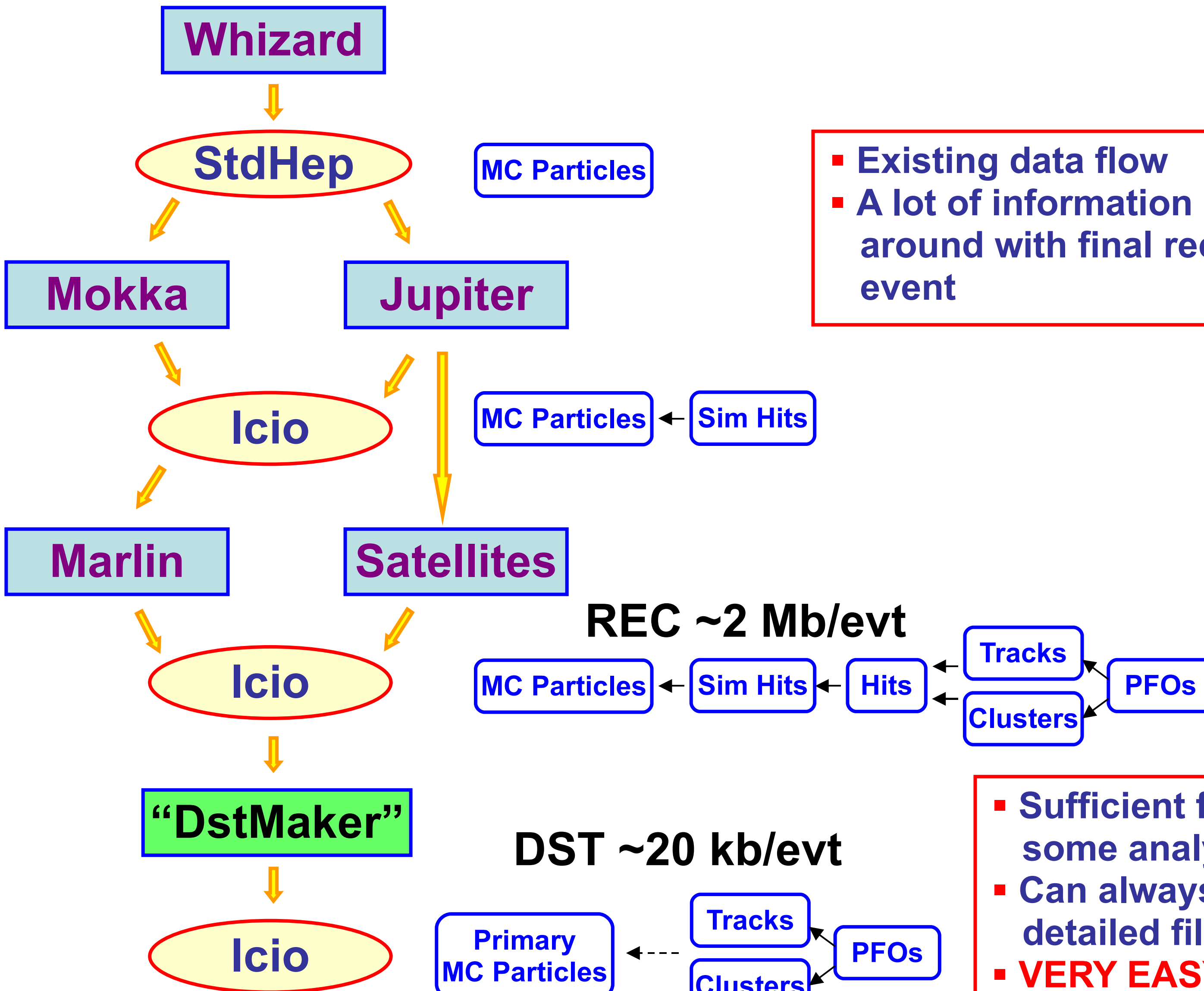
Frank Gaede, ILD Meeting Cambridge, September 11-13, 2008

ILD Standard Full Reconstruction:

- CaloDigi
- TrackDigi
- FullLDCTracking
- PandoraPFA
- JetFinders(2-8 jets)
- LCFIVertex



'combined' ILD software framework

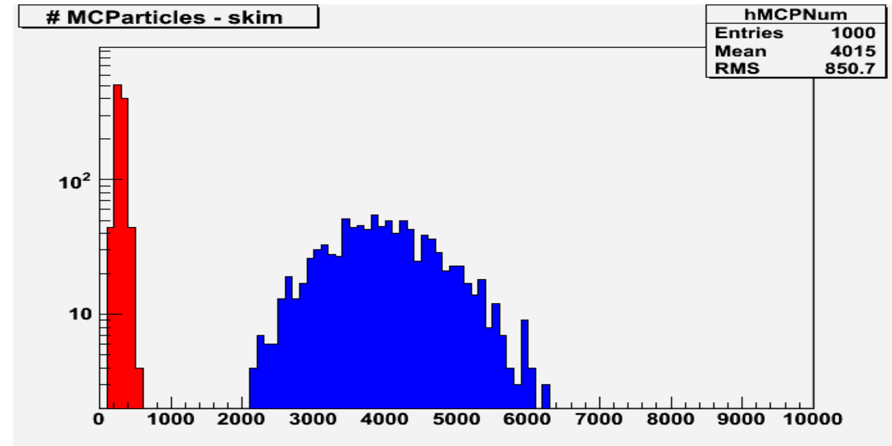


- Existing data flow
- A lot of information carried around with final reconstructed event

- Sufficient for some analyses
- Can always use detailed files
- **VERY EASY TO USE**

'DSTs' with LCIO

- DST format defined at Zeuthen:
  - Tracks, Clusters, ReconstructedParticles, Jets
  - MCParticle skim:
    - store full generated event
    - + reconstucted particles & parents
    - decays in flight & conversions
  - store LCFIVertex flavour tag in ParticleID objects
    - flavour tags b,c,b-bg
    - NN input quantities
    - true jet flavour & charge

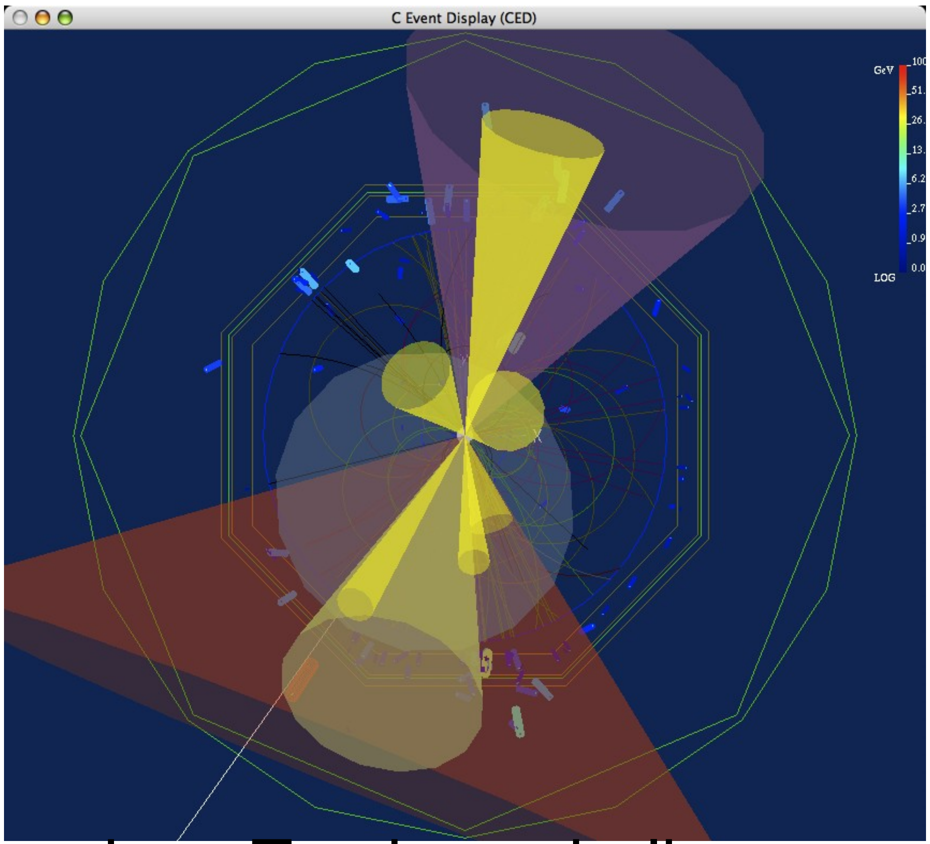


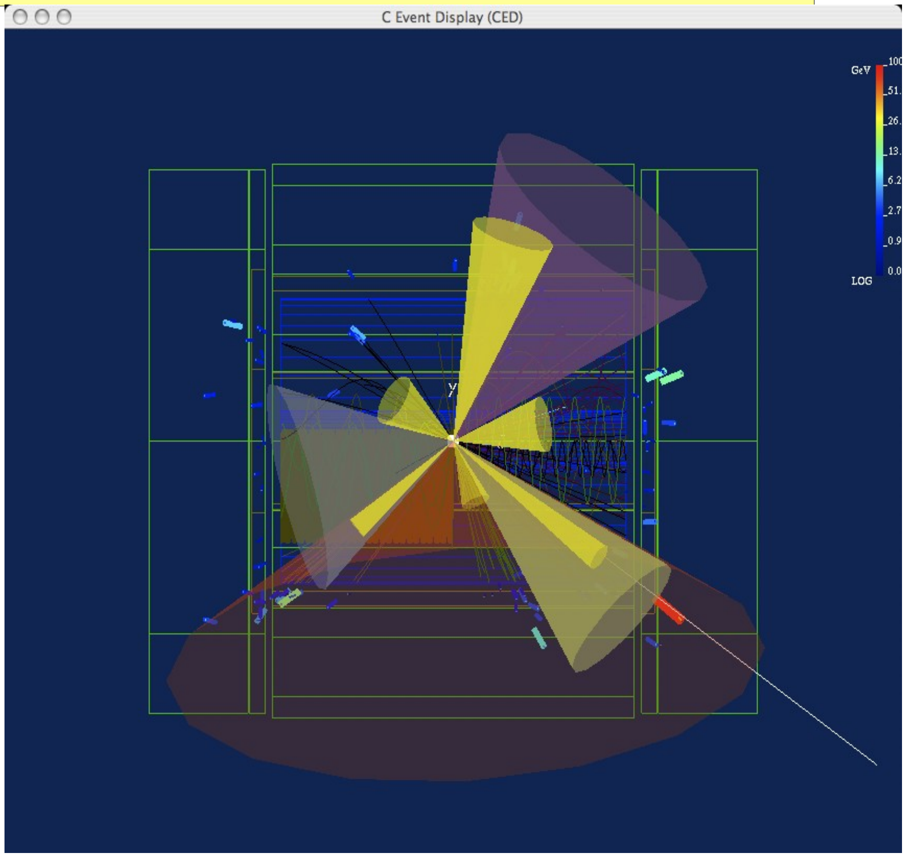

```
///////////////////////////////////
EVENT: 0
RUN: 2011907
DETECTOR: LDCPrime_02Sc
COLLECTIONS: (see below)
///////////////////////////////////

---------------------------------------------------------------------------
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

Visualizing DSTs





- show Tracks as helices
- Clusters as lines & Cylinders (scaled w/ E)
- jets as cones (E,p_t) + particles coloured
- 3 momentum of all particles in the event
- in cvs HEAD – to be released soon

S.Daraszevic

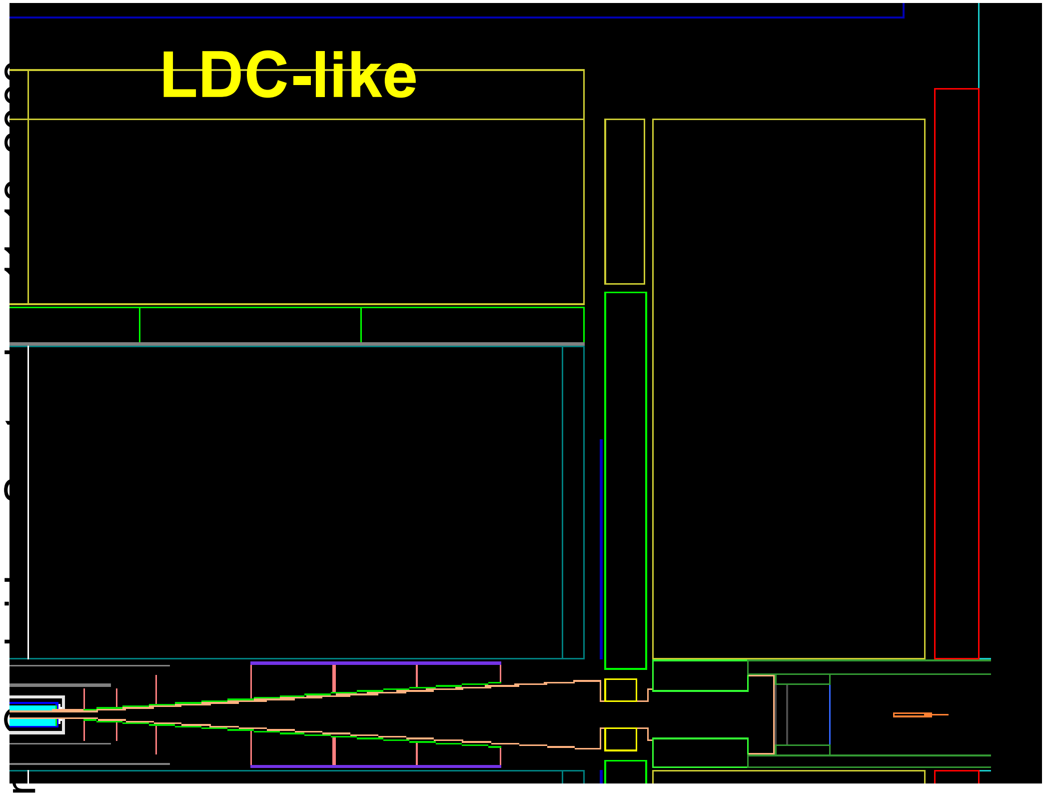
detector variants in Jupiter/Mokka

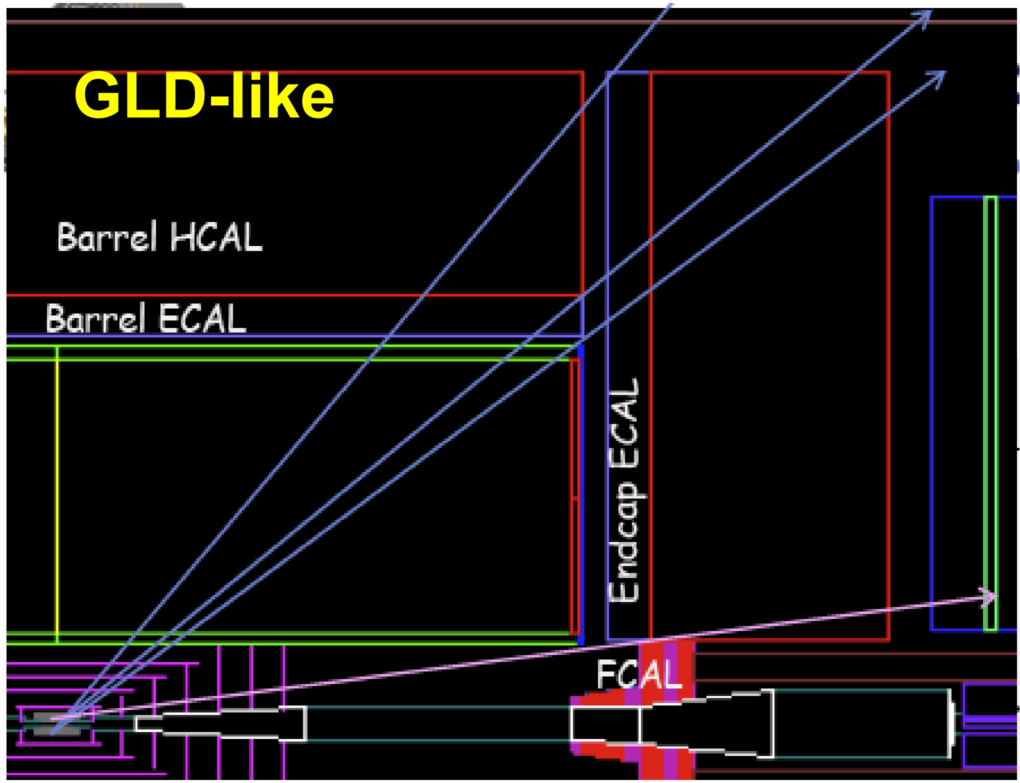
Sub-Detector	Parameter	GLD Jupiter	LDCGLD Mokka	J4LDC Jupiter	LDC Mokka	GLD' Jupiter	LDC' Mokka
TPC	R_{inner} (m)	0.44	0.37	0.34	0.37	0.45	0.37
	R_{outer} (m)	1.98	1.93	1.52	1.51	1.80	1.73
	Z_{max} (m)*	2.60	2.50	2.16	2.19	2.35	2.25
Barrel ECAL	R_{inner} (m)**	2.10	2.02	1.60	1.61	1.85	1.83
	Material	Sci/W	Sci/W	Sci/W	Si/W	Sci/W	Si/W
Barrel HCAL	Material	Sci/W	Sci/Fe	Sci/Fe	Sci/Fe	Sci/Fe	Sci/Fe
ECAL EndC	Z_{min} (m)***	2.80	2.70	2.10	2.30	2.25	2.55
Solenoid	B-field	3	3	4.0	4.0	3.5	3.5
VTX	Inner Layer (mm)	17.5	16.5	15	14	16	15

- at beginning of ILD optimization study (~09/2007) it was considered impossible to join the two different frameworks
- decided to define intermediate 'Prime' detectors in both frameworks and the other concept's detector
  - -> cross check of frameworks !

have seen lots of nice studies with these models !

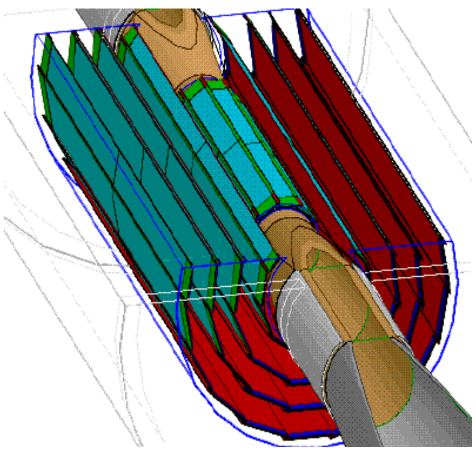
GLD and LDC like detectors

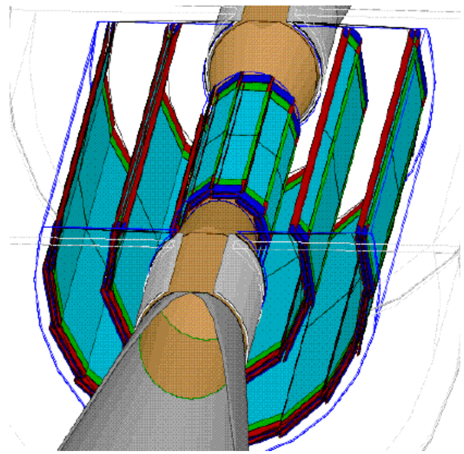




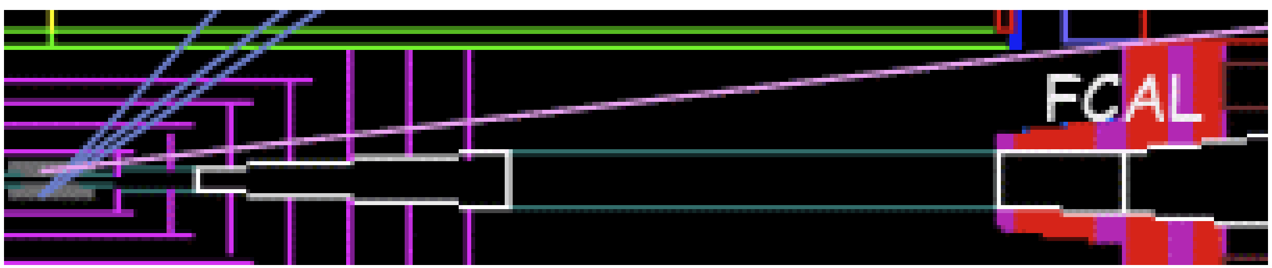
- LDClike and GLDlike detectors are reasonably similar
- however: still quite some differences/alternatives exist

silicon tracking detectors

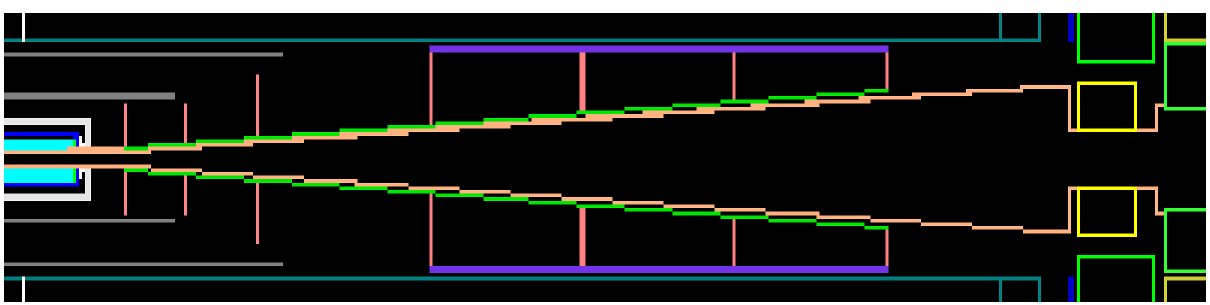




- several technologies for VTX
- LDC geometry: 5 single layers
- GLD geometry 3 double layers
  - (both implemented in Mokka)



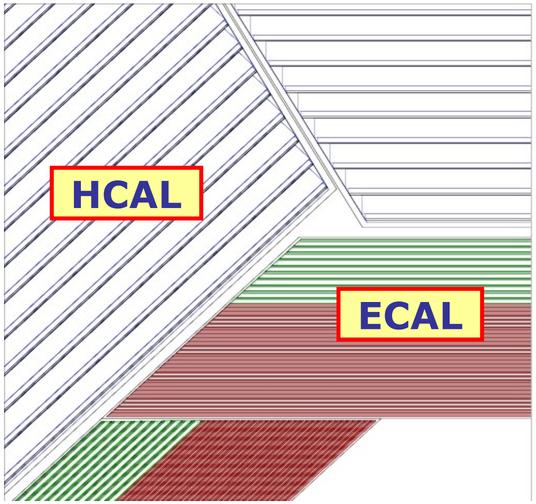
should the Si-Tracking geometry (#layers, positions, resolution) be defined in the ILD baseline ?

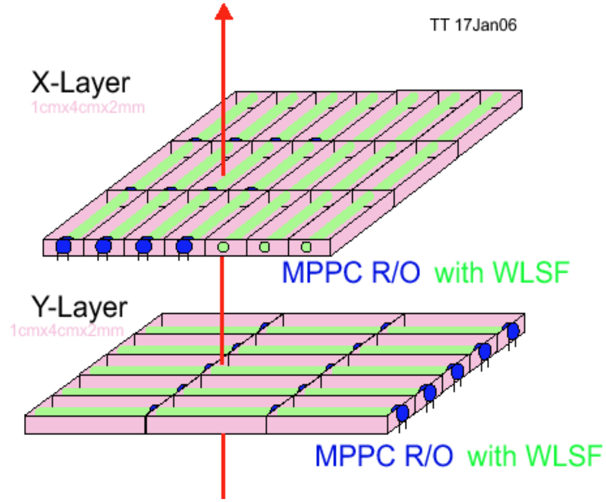


- intermediate Si-Tracker
  - LDC 2 layers
  - GLD 4 layers
- forward tracking
  - LDC/GLD: 7 discs
    - different z-position

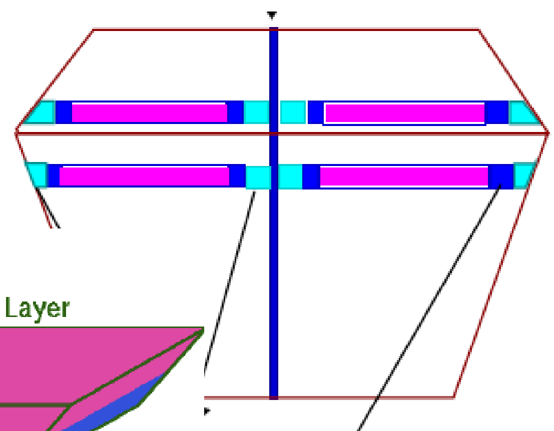
- LDC has (optional) outer Silicon Tracker (SET) used in MC production
- GLD ?

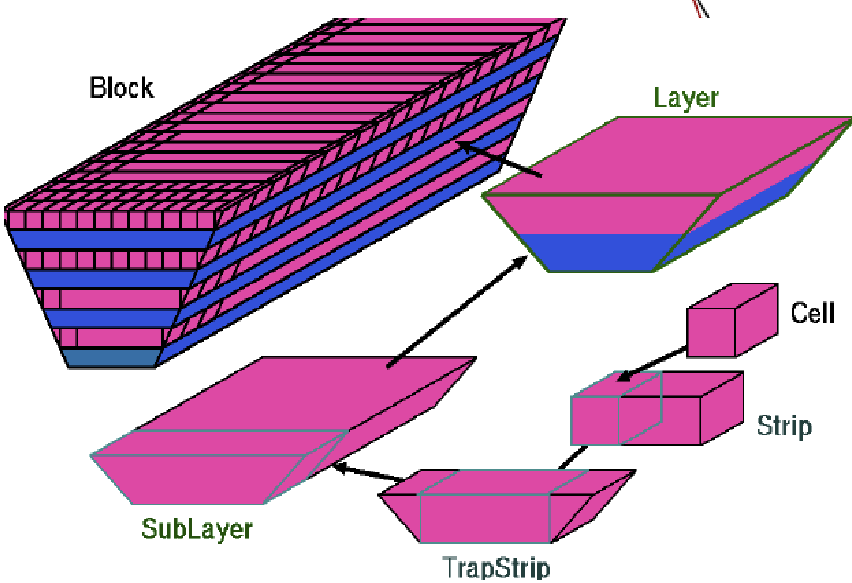
Calorimeters





- **Ecal** technologies
  - LDC: octagonal **SiW**
    - 0.5cm -1cm lateral cell size
  - GLD dodecagonal **Scint-W**
    - 1cmx4cm strips
    - (effective 1cmx1cm)





- **Hcal** technologies
  - LDC: octagonal **Scint-Fe**
    - 3 cm lateral cell size
  - GLD dodecagonal **Scint-Fe**
    - strips ? size ?

calorimeter technologies and geometrical layout fairly different -> issue for common simulation tool

ILD simulation reference design

- by now we should have agreed upon the options and parameters that serve as a simulation-reference design for ILD...

- ... if not – we need to to so really soon (1-2 weeks) – if we want a new production for the LOI

LDC Monte Carlo production on the Grid

	√s (GeV)	Observable	Comments
ZH → eeX	250	σ, m_H	m_H=120GeV, test materials and γ_{ID}
ZH → μμX	250	σ, m_H	m_H=120GeV, test ΔP/P
ZH, H→cc, Z→νν	250	Br(H→cc)	Test heavy flavor tagging and anti-tagging of light quarks and gluon
ZH, H→cc, Z→qq	250	Br(H→cc)	Same as in multi-jet event
Z* → ττ	500	σ, A_{FB}, Pol(τ)	Test π^0 rec. and τ rec. aspects of PFA
tt, t→bW, W→qq'	500	σ, A_{FB}, m_{top}	Test b-tag. and PFA in multi-jet events. m_{top}=175GeV
$\chi^+ \chi^-$, $\chi_2^0 \chi_2^0$	500	σ, m_χ	Pint 5 of Table 1 of BP report. W/Z separation by PFA.

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

- use the Grid to produce a significant Monte Carlo data set
  - (as proposed by WWS software panel)
- 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

LOI MonteCarlo production LDC

LDCPrime_02Sc			
EventType	NEvent	L[1/fb]	delta_L
2f	1092584	19.89	0.15
4f	2666806	48.88	2.69
6f	446028	471.36	66.21
aa_X	181408	0.28	0.02
ee	6931010	0.1	0
eaea	344270	0.19	0
nnNa	806700	17.45	3.93
aaNa	261954	9.42	0.82
hX	276728	465.86	75.74
other	120000	0	0
Zh_ee_mur	20000	1332.98	0
Zh_qqnn	10000	223.26	0
Zh_qqqq	25000	158.73	0
tautau	100000	22.03	2.8
6f_bbqqq	450217	486.37	5.42
sp5_ch_ne	82305	464.01	112.54
sp5_x	78570	692	248.26
sps1ap	1617133	891.37	195.94
ZZ	50000	74.59	39.59
total	15560713		

LDC01_06Sc			
EventType	NEvent	L[1/fb]	delta_L
4f	4000	46.02	76.38
Zh_ee_mumu	19000	1266.33	66.65
Zh_qqnn	10000	223.26	0
Zh_qqqq	25000	158.73	0
sps1ap	1621157	879.67	190.59
ZZ	55000	82.21	44.76
total	1734157		

ilcsoft v01-04

LDC_GLD_01Sc			
EventType	NEvent	L[1/fb]	delta_L
Zh_ee_mumu	20000	1332.98	0
Zh_qqnn	9000	200.94	0
Zh_qqqq	23000	146.03	0
sps1ap	1528657	873.32	197.43
ZZ	55000	82.21	44.76
total	1635657		

SM model background sample (SLAC)

canonical signal samples LOI (WWS-SW)

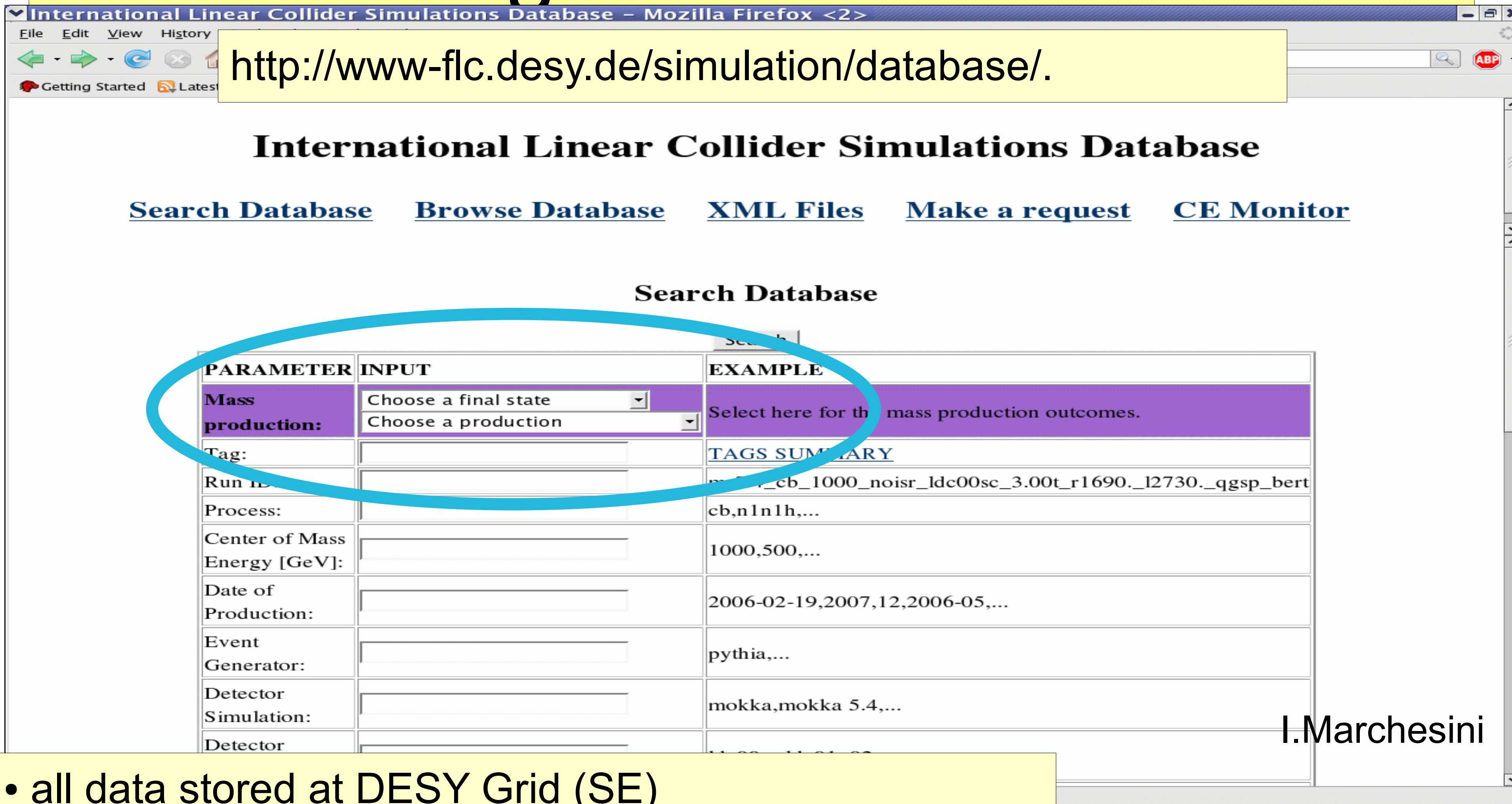
other physics signal samples

massive MC production for 3 LDC like detectors – details at:
http://www-flc.desy.de/simulation/database
http://ilcsoft.desy.de/loi/reco (summary)

total effort: simulation ~2 month
reconstruction ~1 month

accessing the LDC data files

http://www-flc.desy.de/simulation/database/.



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

should try and make also GLD DSTs available !

summary table LOI reconstruction

Reconstructed Monte Carlo Files ILD – Mozilla Firefox

File Edit View Go Bookmarks Tools Help

http://ilcsoft.desy.de/loi/reco

Go

LCIO ilcsoft simulation/geant4 Google Linear Collider Forum: ... DESY IT Group MyHome »

complete file lists per sample REC and DST !

Reconstructed Monte Carlo Files ILD

The following tables provide an overview of the physics samples that have been fully simulated and recons
LDCPrime_02Sc, LDC01_06Sc and LDC_GLD_01Sc. The tables show both Standard Model and Signal eve

The events "Sig 1-6:" are the "Benchmark Reactions for the LOI" as defined by the WWS software pan

Click on the corresponding links to get a complete list of logical file names for the given type

LDCPrime_02Sc					
EventType	NEvent	L[1/fb]	delta_L	REC_files	DST_files
2f (w/o ee,tautau)	1092584	19.8890	0.1509	2f_LDCPrime_02Sc_REC.list	2f_LDCPrime_02Sc_DST.list
4f	2666806	48.8792	2.6930	4f_LDCPrime_02Sc_REC.list	4f_LDCPrime_02Sc_DST.list
6f (w/o bbqqqq)	446028	471.3608	66.2062	6f_LDCPrime_02Sc_REC.list	6f_LDCPrime_02Sc_DST.list
gg->X	181408	0.2770	0.0212	aa_X_LDCPrime_02Sc_REC.list	aa_X_LDCPrime_02Sc_DST.list
ee	6931010	0.0997	0.0002	ee_LDCPrime_02Sc_REC.list	ee_LDCPrime_02Sc_DST.list
eg->eg	344270	0.1924	0.0005	eaea_LDCPrime_02Sc_REC.list	eaea_LDCPrime_02Sc_DST.list
nn(Ng)	806700	17.4457	3.9255	nnNa_LDCPrime_02Sc_REC.list	nnNa_LDCPrime_02Sc_DST.list
gg(Ng)	261954	9.4225	0.8234	aaNa_LDCPrime_02Sc_REC.list	aaNa_LDCPrime_02Sc_DST.list
hX	276728	465.8559	75.7372	hX_LDCPrime_02Sc_REC.list	hX_LDCPrime_02Sc_DST.list
other	120000	0.0000	0.0000	other_LDCPrime_02Sc_REC.list	other_LDCPrime_02Sc_DST.list
Signal Events:					
Sig1: ZH,H->ee/mumu	20000	1332.9779	0.0000	Zh_ee_mumu_LDCPrime_02Sc_REC.list	Zh_ee_mumu_LDCPrime_02Sc_DST.list
Sig2: ZH,H->qq,Z->nn	10000	223.2641	0.0000	Zh_qqnn_LDCPrime_02Sc_REC.list	Zh_qqnn_LDCPrime_02Sc_DST.list
Sig3: ZH,H->qq,Z->qq	25000	158.7302	0.0000	Zh_qqqq_LDCPrime_02Sc_REC.list	Zh_qqqq_LDCPrime_02Sc_DST.list
Sig4: ee->tautau	100000	22.0349	2.8003	tautau_LDCPrime_02Sc_REC.list	tautau_LDCPrime_02Sc_DST.list
Sig5: ee->tt->bbqqqq	450217	486.3730	5.4159	6f_bbqqqq_LDCPrime_02Sc_REC.list	6f_bbqqqq_LDCPrime_02Sc_DST.list
Sig6: ee->chch,nene[SP5]	82305	464.0133	112.5413	sp5_ch_ne_LDCPrime_02Sc_REC.list	sp5_ch_ne_LDCPrime_02Sc_DST.list

Done

GLD physics sample production

				Jupiter samples		
Calibration samples			# Events	gldapr08	gldprim_v04	j4ldc_v04
	Thomson's uds91		10000	done	done	done
	Thomson's uds200		10000	done	done	done
	PythiaZPole uds		10000	done	done	done
	PythiaZPole ccbar		10000	done	done	done
	PythiaZPole bbbar		10000	done	done	done
	jsf's uds 91		2500	done	done	done
	jsf's uds 200		2500	done	done	done
	jsf's uds 500		20000	done	done	done
				Jupiter Production		
	Process	Int. Lum(1/fb)	# Events	gldapr08	gldprim_v04	j4ldc_v04
250 GeV	zh->eeH	250	5000	done	done	done
	zh->μμH	250	5000	done	done	done
	zh->ννH	250	12500	done	done	done
	zh->qqH	250	40000	done	done	done
	zz->eeqq	250	20000	done	done	done
	zz->μμqq	250	20000	done	done	done
	zz->ννqq	250	77500	done	done	done
	zz->qqqq	250	168000	9300	done	93000
	zz->ττqq	250	20000	0	done	0
	ww->enuenu	250				
	ww->munumunu	250				
				Jupiter Production		
	Process	Int. Lum(1/fb)	# Events	gldapr08	gldprim_v04	j4ldc_v04
500 GeV	smuon(e-L)	500	14750	done	done	done
	smuon(e-R)	500	61000	done	done	done
	xcxc(e-L)	500	79000	done	done	done
	xcxc(e-R)	500	500	done	done	done
	xn2xn2(e-L)	500	14750	done	done	done
	bblnqq(e-L)	100	54000	0	done	0
	bblnqq(e-R)	100	24000	0	done	0
	bbqqqq(e-L)	150	126000	0	done	0
	bbqqqq(e-R)	150	51000	0	done	0
	tau-pair	12.4	57500	done	done	done
	tau-pair	100		0	in progress	0

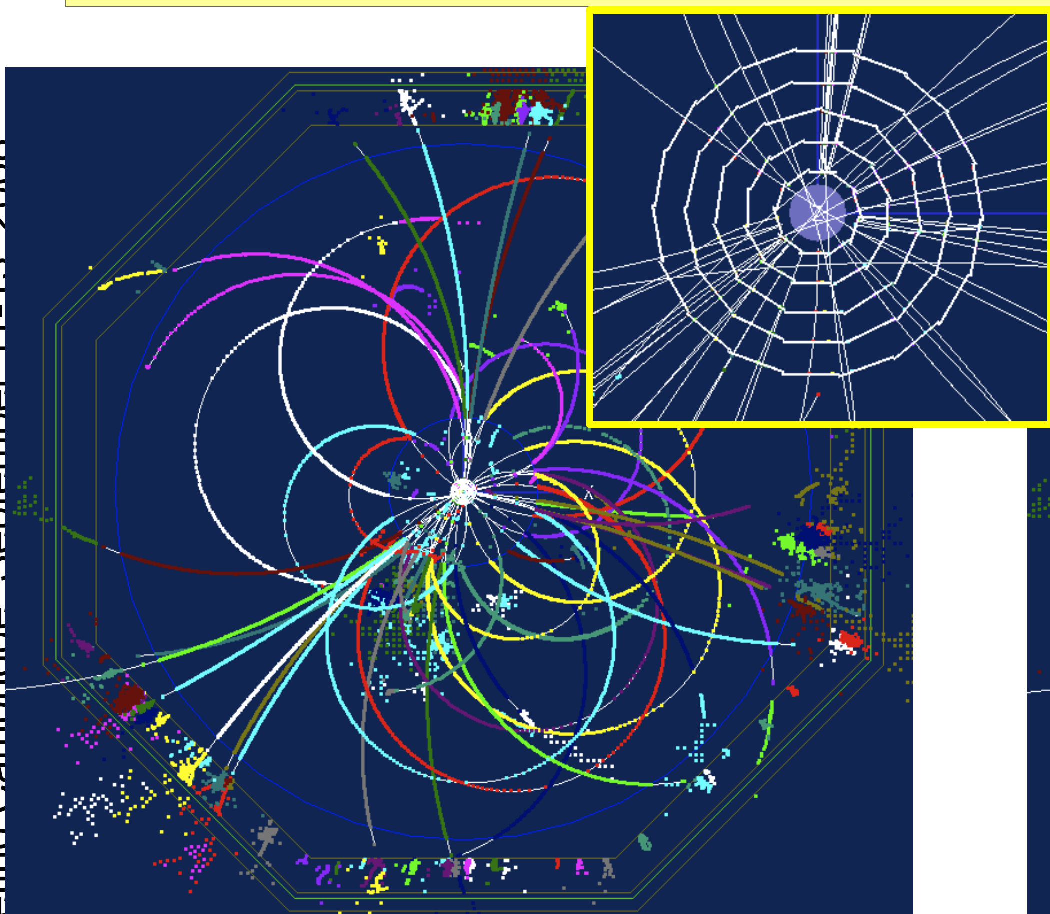
Job summary: http://ilcphys.kek.jp/soft/samples/apr08/

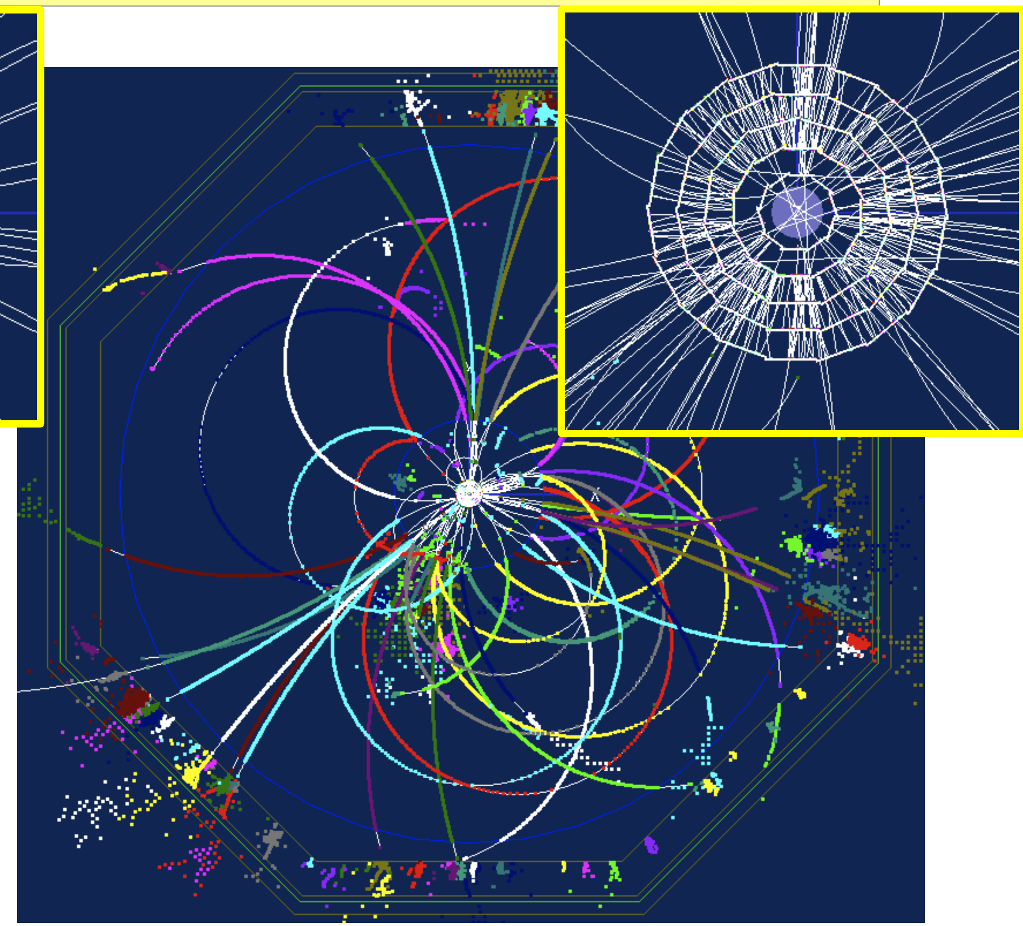
- GLD-ILD focused on producing dedicated physics signal samples
- three detector models for optimization: GLD, GLDPrime, j4LDC
-

Studying machine background

- IDAG has made studying the machine background a requirement for the LOI
- so far studies have investigated the hit rates to be expected from pair production (based on current parameters from MDI) (A.Vogel)
- effect of background on PFA yet be addressed !
- overlay of fully simulated pairs is realistically not feasible
- -> start w/ 'salt&pepper' background w/ expected rates integrated over N BXs (depending on subdetector readout)

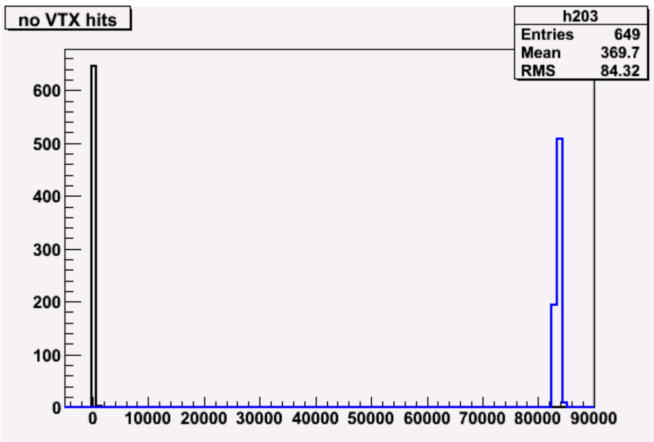
Random Noise in VTX detector

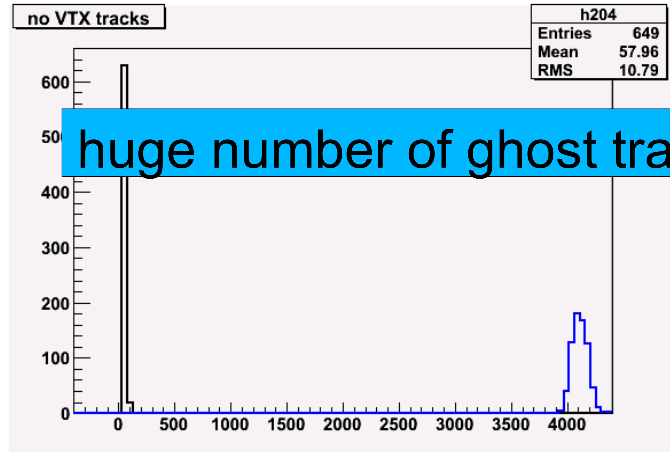




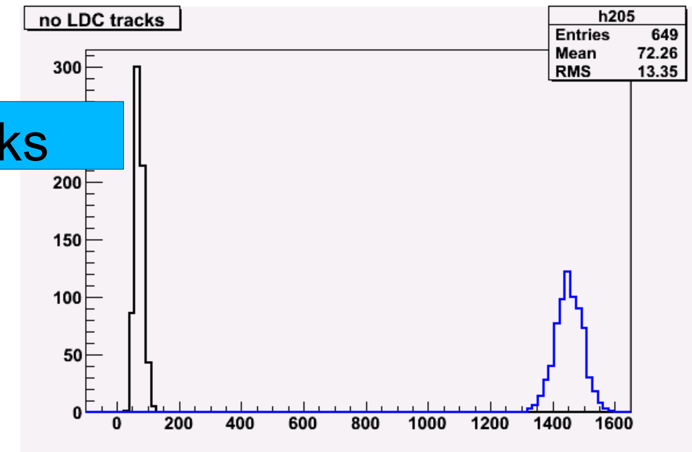
- example: VTXNoiseHits – processor that adds 'Salt'n Pepper' hits to the VTX,e.g:
- 400,40,15,10,3 hits/cm^2 in layers 1-5
  - SiTracks: 43 -> 1647
  - LDCTracks 54 -> 1658
  - **PandoraPFO 122 -> 131**

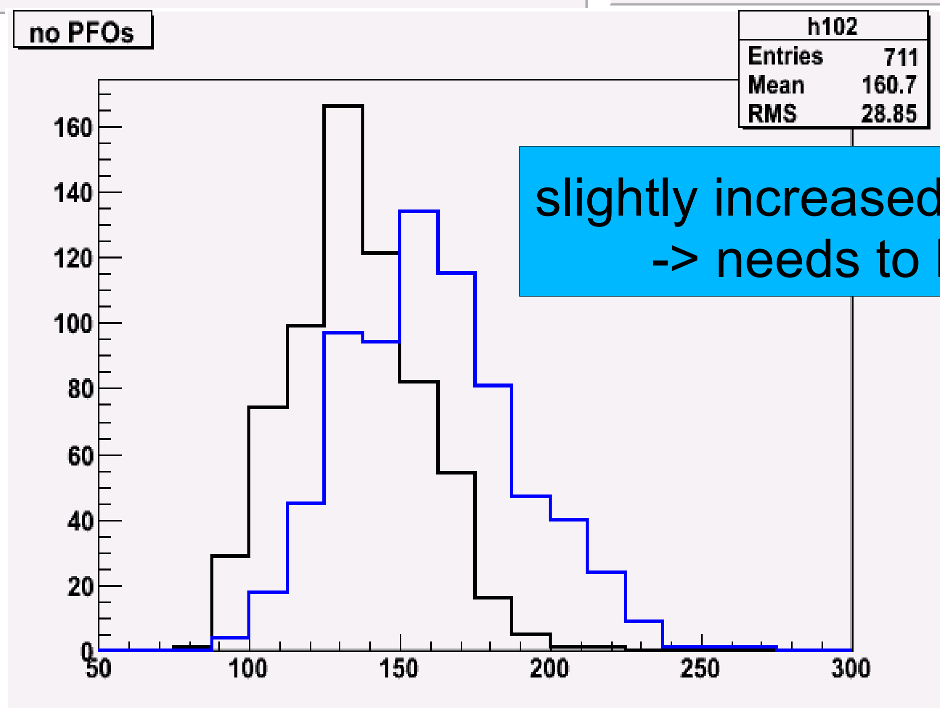
Random Noise in VTX detector II

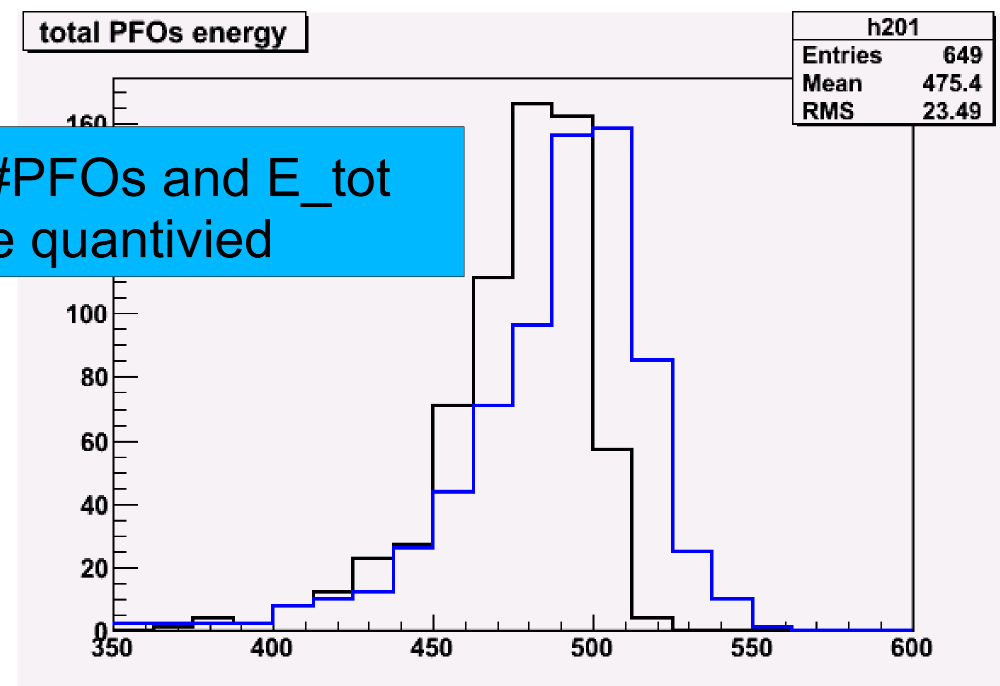




huge number of ghost tracks





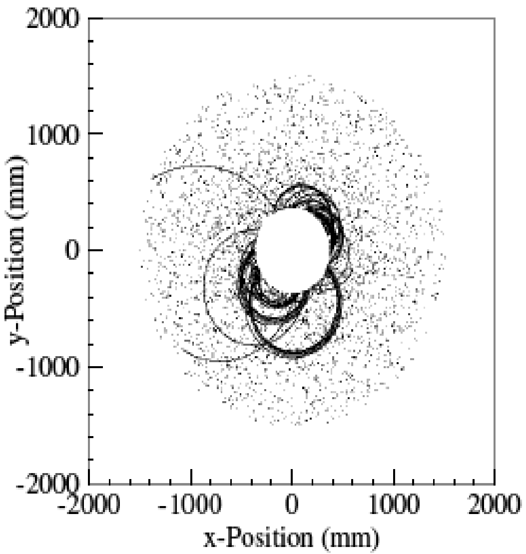


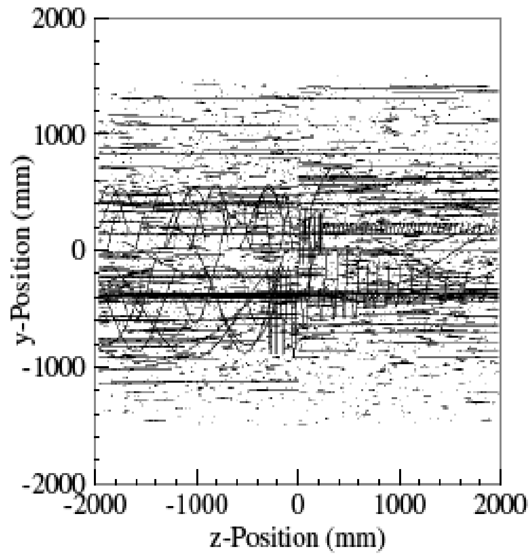
slightly increased #PFOs and E_tot
-> needs to be quantivied

- some first plots (ttbar events)
- **MarlinReco/PandoraPFA: out of the box !**

K.Klimek

Studying machine background II





- include other silicon trackers (SIT, FTD)
- should try and parameterize the 'shape' of the background to get more realistic results !?
  - **need to adopt PatRec & PFA in order to cope with bg !**
    - note: only positive result is to show that we can cope – if not successful we might not have tried hard enough
    - -> this might be difficult on the time scale of the LOI !

towards the LOI - I

- ideally we should not have two frameworks but rather on common ILDsoft framework
- meeting of software experts from both frameworks yesterday:
- general agreement to move towards a common framework !
- however only some first steps can reasonably be done before the LOI
  - try to see if calorimeter (scint-W Ecal) options from GLD can be implemented in Mokka to allow comparison within one framework
  - work closer together on Grid-Mass production
  - have common Grid-data catalogue with DSTs
  - review the existing DST format: is it complete and well suited for the existing analyses in both 'regions' ?

towards the LOI - II

- create the new models for the ILD simulation reference detector in Mokka:
  - B = 3.5 Tesla
  - RI_ecal = 1.85 m TPC: r= ~1 , z/2 = 230 cm
  - VTX detector layout
  - Silicon layout (SIT, FTD, SET ,...)
  - Calorimeters (Ecal 22 X0, .5x.5, Hcal
- redo the full Monte Carlo production :
  - Physics samples & SM background (~15MEvents)
  - -> 're-run' the physics analyses on these new data
  - keep mixed framework approach for LOI studies

have bi-weekly software meetings w/ experts to allow fast feedback and communication on all software issues

beyond the LOI

- take a deep breath and then start to work towards a truly common ILD software framework
- need to evaluate the exisiting software tools ('ILD software review')
- try and identify strengths and weaknesses in both frameworks and see how things can be merged allowing a 'smooth' transition for the users, e.g.
  - create **ILDReco** merging best tools from MarlinReco and Satelies
  - improve geometry description in simulation and reconstruction
  - think about LCIOv2 with a revised data model (learn from Jupiter&Satelites)
  - ...

Summary

- LDC&GLD have mature and powerful software frameworks that have been used as a 'combined' framework for the first round of LOI MC production
- have seen many nice and interesting analyses from this mixed framework
- need to keep momentum for the next round of simulations and work towards the LOI

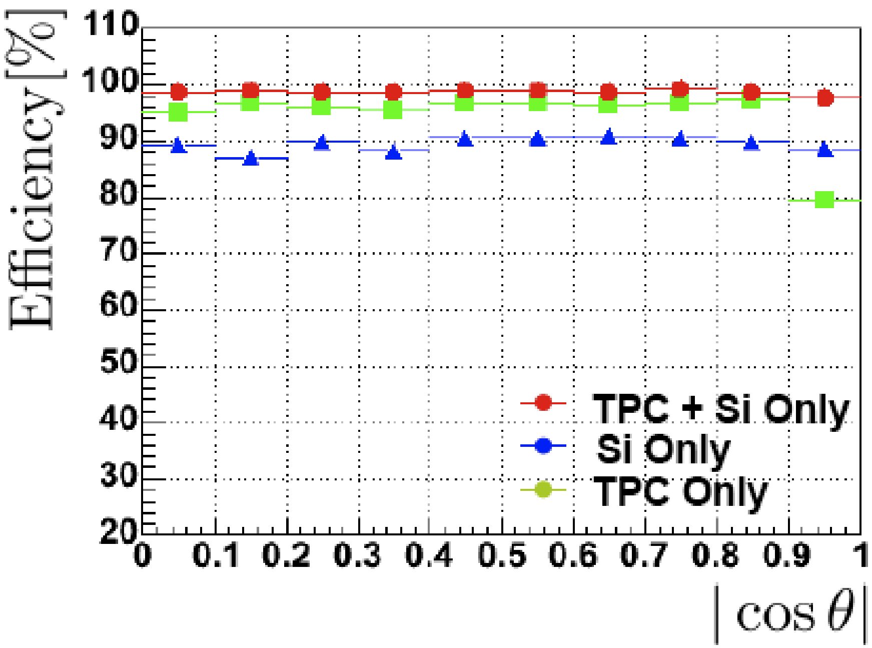
after LOI we will have the time to take on the next big project:

work together on a truly **common ILDsoft framework**

that combines the best of the two worlds and allows us to further optimize the ILD detector by comparing the different technology (and geometry) options on a 'fair' basis additional material

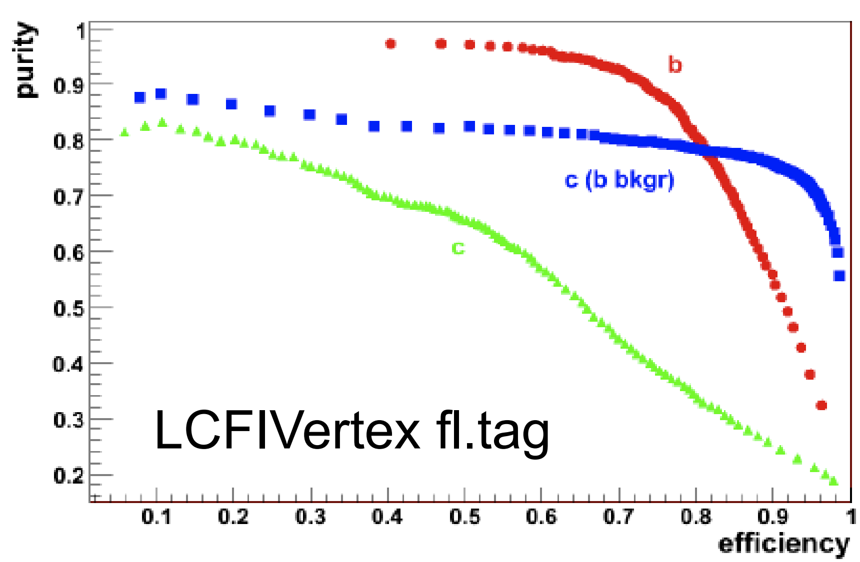
reconstruction tools

Tracking efficiency



ilcsoft v01-04: PandoraPFA - LDCPrime

E_{JET}	$\sigma_E/E = \alpha/\sqrt{E_{jj}}$ $\lvert\cos\theta\rvert<0.7$	σ_E/E_j
45 GeV	24.9 %	3.7 %
100 GeV	30.7 %	3.1 %
180 GeV	43.0 %	3.2 %
250 GeV	52.2 %	3.3 %
500 GeV	78.8 %	3.5 %



- reconstruction tools have been improved
- and modified to be compatible with
- the LDCPRime and GLDPrime model
-