ILD LOI Benchmark Simulation and Reconstruction

(Looking Back…)

Steve Aplin
DESY

2009 Linear Collider Workshop of the Americas
30th September 2009

Overview

Physics Performance Benchmarking

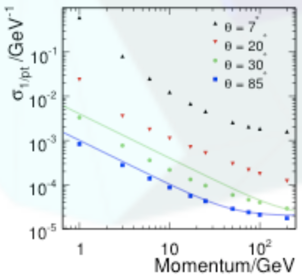
Detector Performance Benchmarking

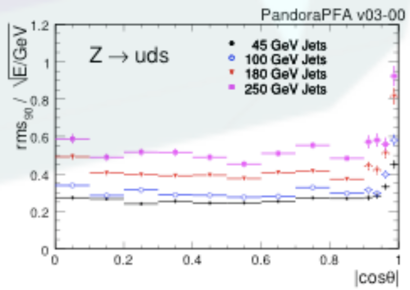
Algorithm Performance Benchmarking

Full Simulation and Reconstruction

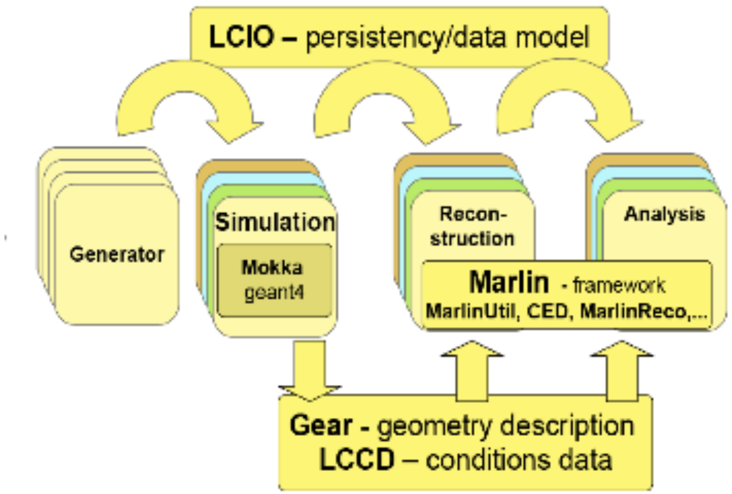
More than 50 million events produced

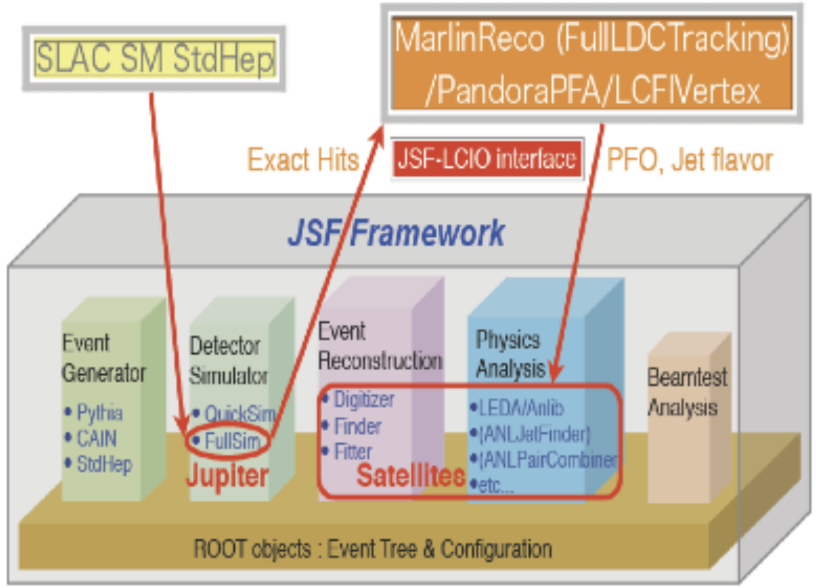
Analysis	$\sqrt{s}$	Observable	Precision	Comments
Higgs recoil mass	250 GeV	$\sigma(e^+e^- \to ZH)$	0.5 fb (5.1 %)	Model Independent
		m_H	74 MeV	Model Independent
		m_H	67 MeV	Model Dependent
Higgs Decay	250 GeV	$Br(H \to b\bar{b})$	$2 \oplus 5\,\%$	includes 5 %
		$Br(H \to c\bar{c})$	$14 \oplus 5\,\%$	from
		$Br(H \to gg)$	$29 \oplus 5\,\%$	$\sigma(e^+e^- \to ZH)$
$\tau^+\tau^-$	500 GeV	$\sigma(e^+e^- \to \tau^+\tau^-)$	0.3 %	$\theta_{\tau^+\tau^-} > 178°$
		A_{FB}	±0.003	$\theta_{\tau^+\tau^-} > 178°$
		P_τ	±0.015	$\tau \to \pi\nu$ only
Gaugino Production	500 GeV	$\sigma(e^+e^- \to \tilde{\chi}_1^+\tilde{\chi}_1^-)$	0.6 %	
		$\sigma(e^+e^- \to \tilde{\chi}_2^0\tilde{\chi}_2^0)$	2.1 %	
		$m(\tilde{\chi}_1^\pm)$	2.4 GeV	from kin. edges
		$m(\tilde{\chi}_2^0)$	0.9 GeV	from kin. edges
		$m(\tilde{\chi}_1^0)$	0.8 GeV	from kin. edges
$e^+e^- \to t\bar{t}$	500 GeV	$\sigma(e^+e^- \to t\bar{t})$	0.4 %	$(bq\bar{q})$ $(\bar{b}q\bar{q})$ only
		m_t	40 MeV	fully-hadronic only
		m_t	30 MeV	+ semi-leptonic
		Γ_t	27 MeV	fully-hadronic only
		Γ_t	22 MeV	+ semi-leptonic
Smuons in SPS1a'	500 GeV	$\sigma(e^+e^- \to \tilde{\mu}_L^+\tilde{\mu}_L^-)$	2.5 %	measurements
		$m(\tilde{\mu}_L)$	0.5 GeV	
Staus in SPS1a'	500 GeV	$m(\tilde{\tau}_1)$	$0.1\,\text{GeV} \oplus 1.3\sigma_{LSP}$	
WW Scattering	1 TeV	α_4	$-1.4 < \alpha_4 < 1.1$	
		α_5	$-0.9 < \alpha_5 < +0.8$	





Software Tools

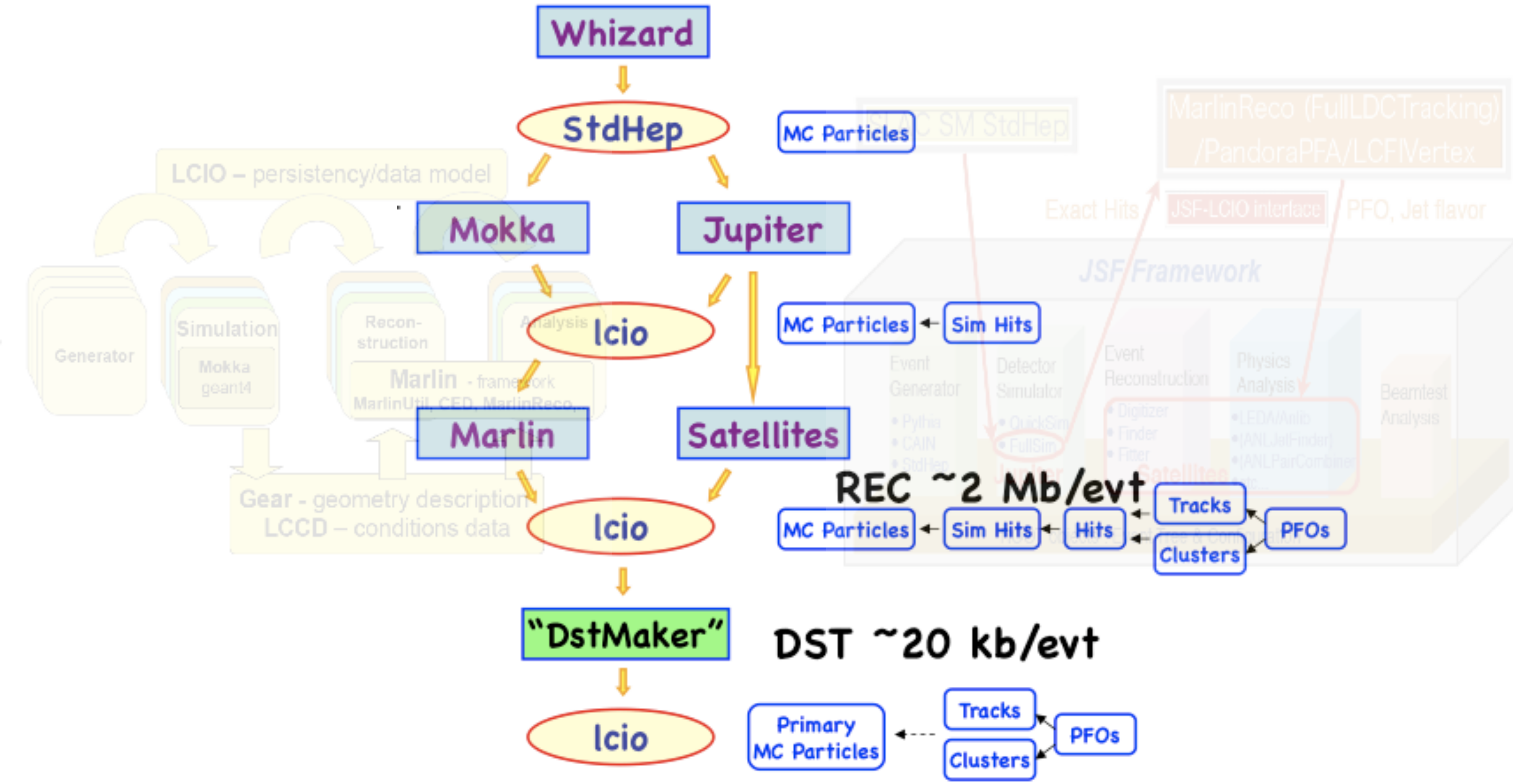




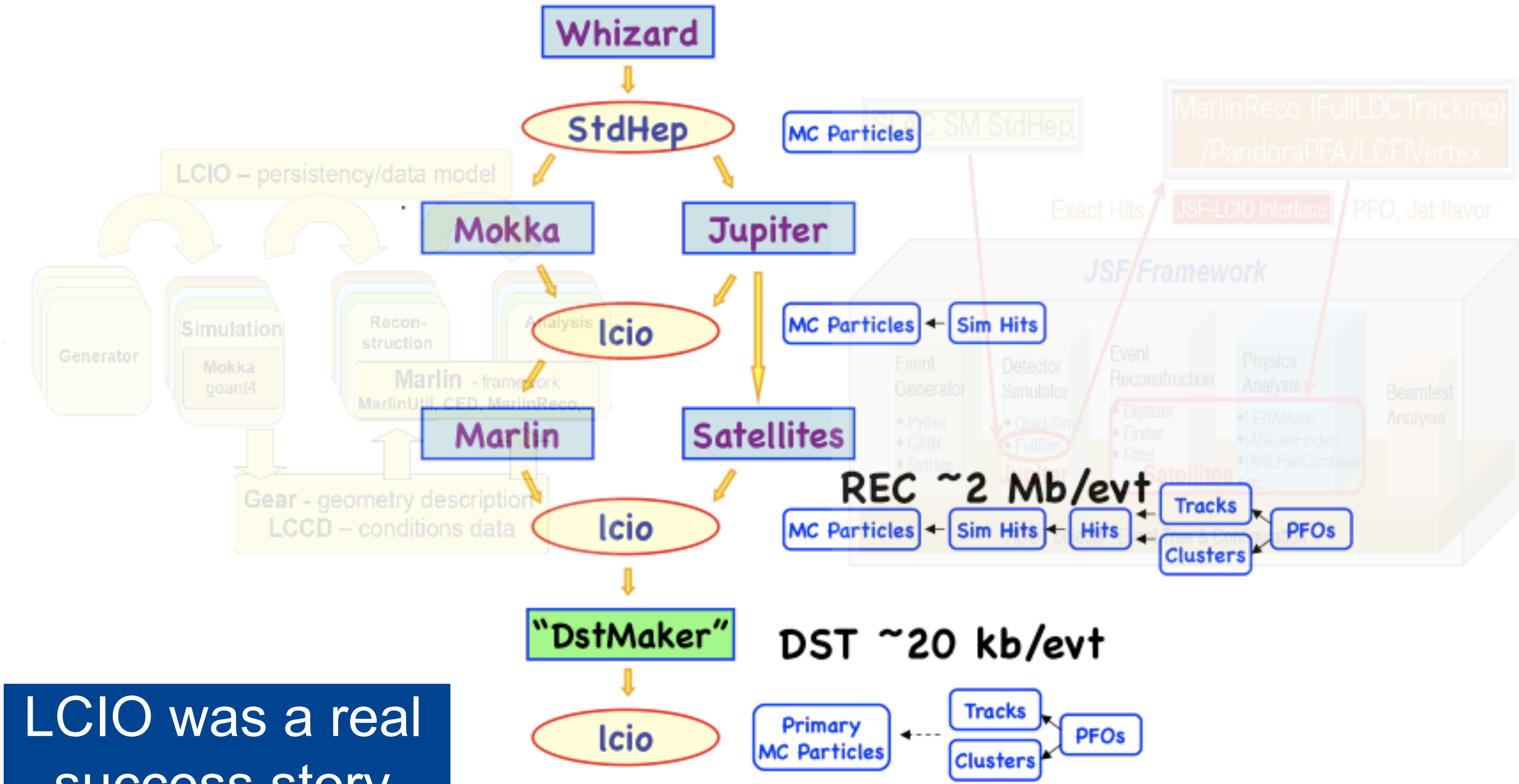
ILD members brought with them two fully developed software frameworks:

1) the JSF Framework used to develop the GLD concept
2) the Mokka – Marlin Framework used to develop LDC

Software Tools



Software Tools



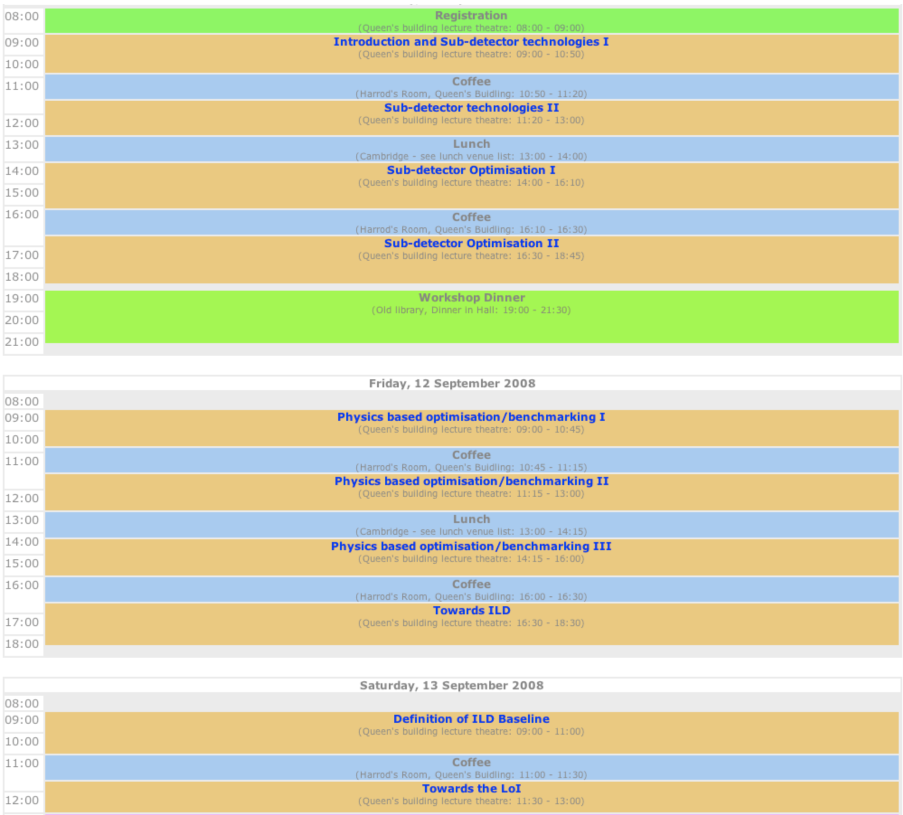
LCIO was a real success story

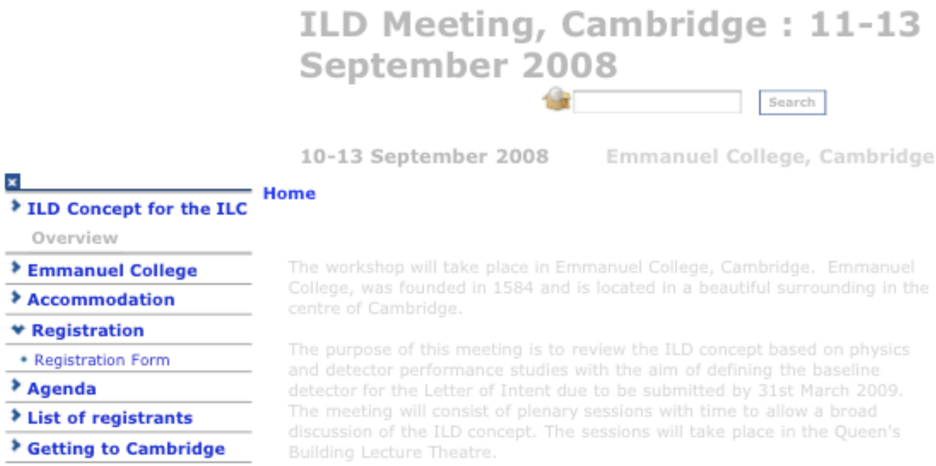
Input to Detector Optimisation

The maturity of the detector descriptions within both these frameworks allowed for studies to investigate the parameter space, e.g. B and R, needed for Detector Optimisation.

Model Name		GLD	GLD′	GLD4LDC	LDC4GLD	LDC′	LDC
Simulator		Jupiter			Mokka		
B field (T)		3.0	3.5	4.0	3.0	3.5	4.0
Beampipe R_{min}		15.0	14.0	13.0	15.5	14.0	13.0
Vertex	Geometry	cylindrical			ladders		
Detector	Layers	3 doublets			5		
	R_{min}	17.5	16.0	15.0	16.5	15.0	14.0
Barrel	Layers	4 cylinders			2 cylinders		
SIT	Radii	90, 160, 230, 300			161.4, 270.1		
TPC	R_{min}	437	435	371	371		
drift	R_{max}	1978	1740	1520	1931	1733	1511
region	z_{max}	2600	2350	2160	2498	2248	2186
TPC pad rows		256	217	196	260	227	190
ECAL	R_{min}	2100	1850	1600	2020	1825	1610
barrel	Layers	33			20(thin)+9(thick)		
	Total X_0	28.4			22.9		
ECAL endcap z_{min}		2800	2250	2100	2700	2300	2550
HCAL	Layers	46	42	37	48		
barrel	R_{max}	3617	3260	2857	3554	3359	3144
λ_I (ECAL+HCAL)		6.79	6.29	5.67	6.86		

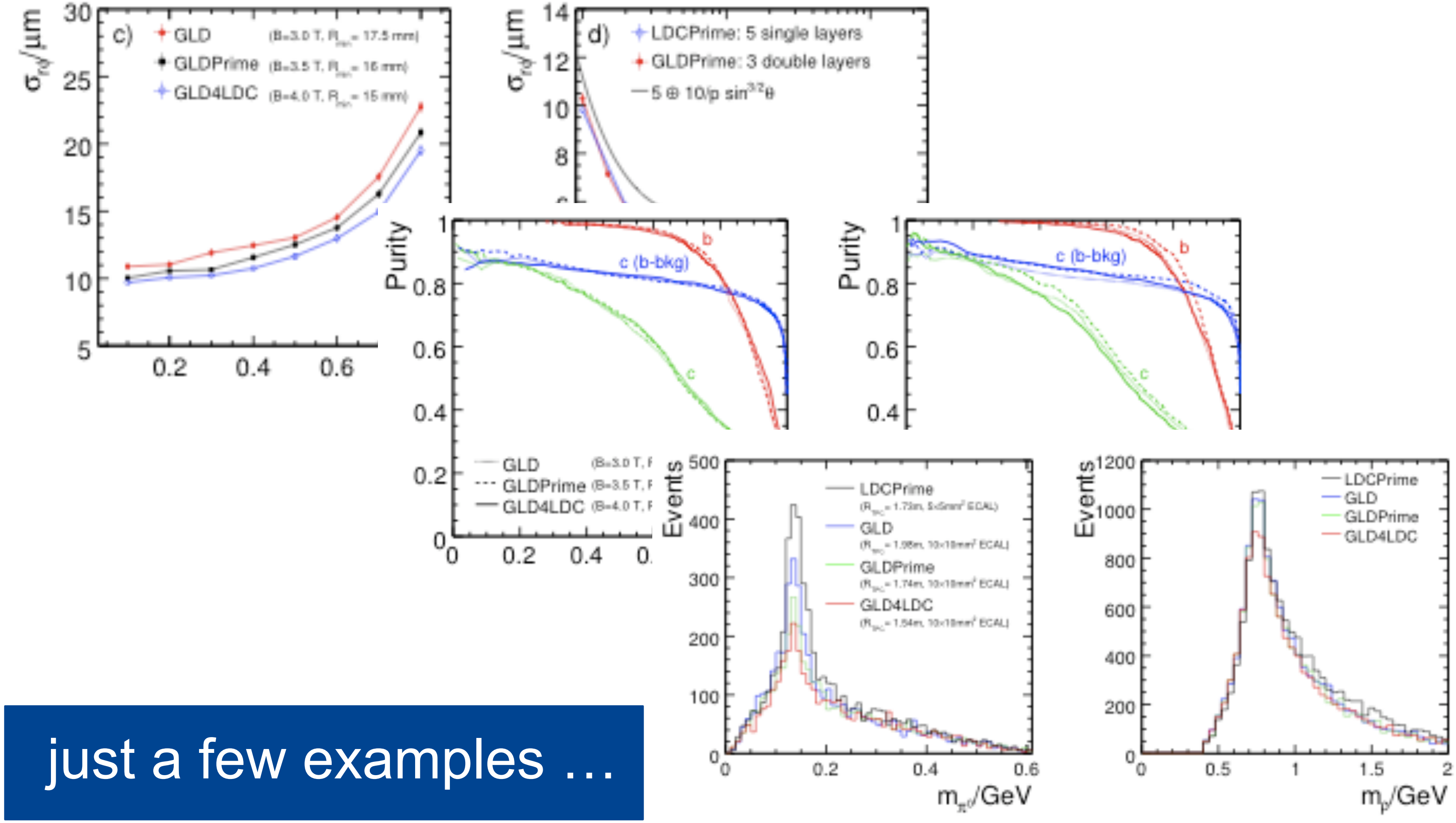
Input to Detector Optimisation





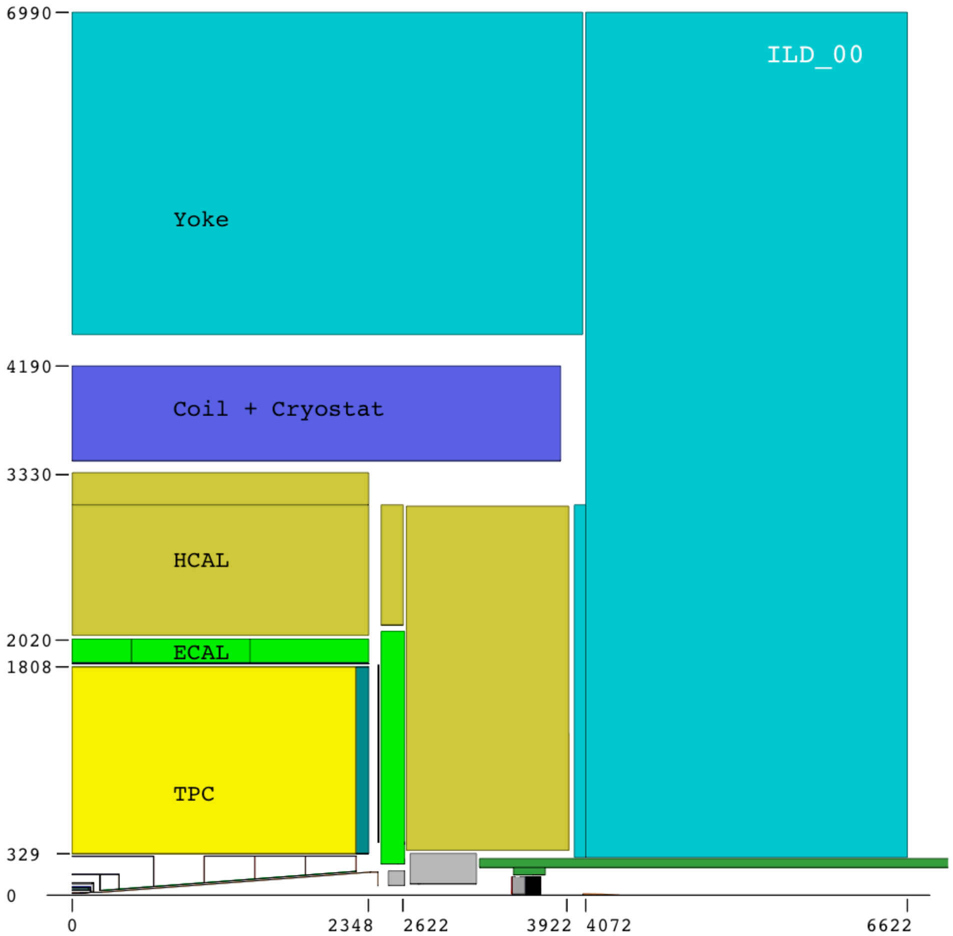
In preparation for the LOI such studies were presented in detail at the ILD workshop in Cambridge, September 2008

Input to Detector Optimisation



just a few examples …

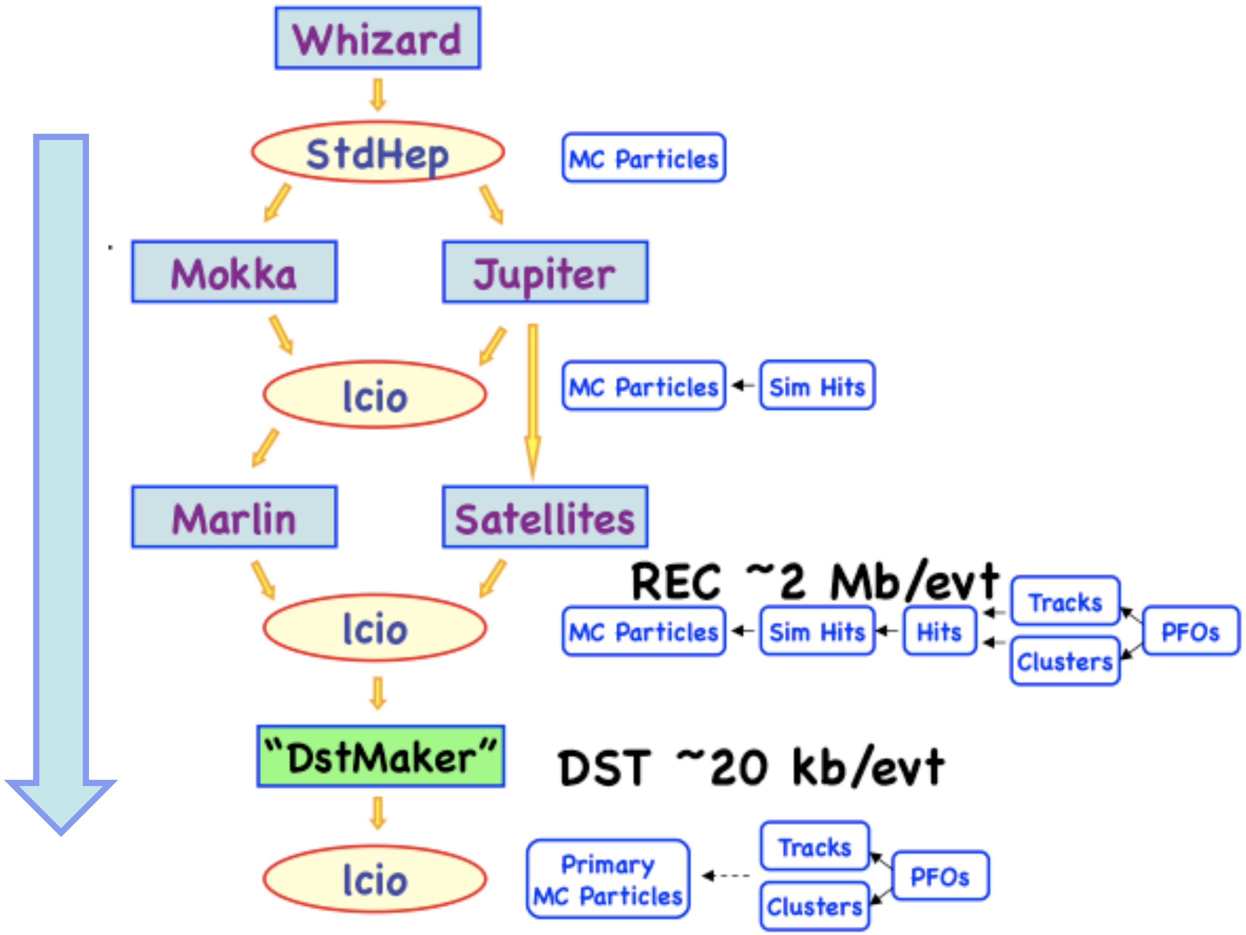
ILD Software Reference Model



Model Name		ILD
Simulator		Mokka
B field (T)		3.5
Beampipe R_{min}		14.5
Vertex	Geometry	ladders
Detector	Layers	3 doublets
	R_{min}	16.0
Barrel	Layers	2 cylinders
SIT	Radii	165, 309
TPC	R_{min}	395
drift	R_{max}	1739
region	z_{max}	2247.5
TPC pad rows		224
ECAL	R_{min}	1847.4
barrel	Layers	20+9
	Total X_0	23.6
ECAL endcap z_{min}		2450
HCAL	Layers	48
barrel	R_{max}	3330
λ_I (ECAL+HCAL)		6.86

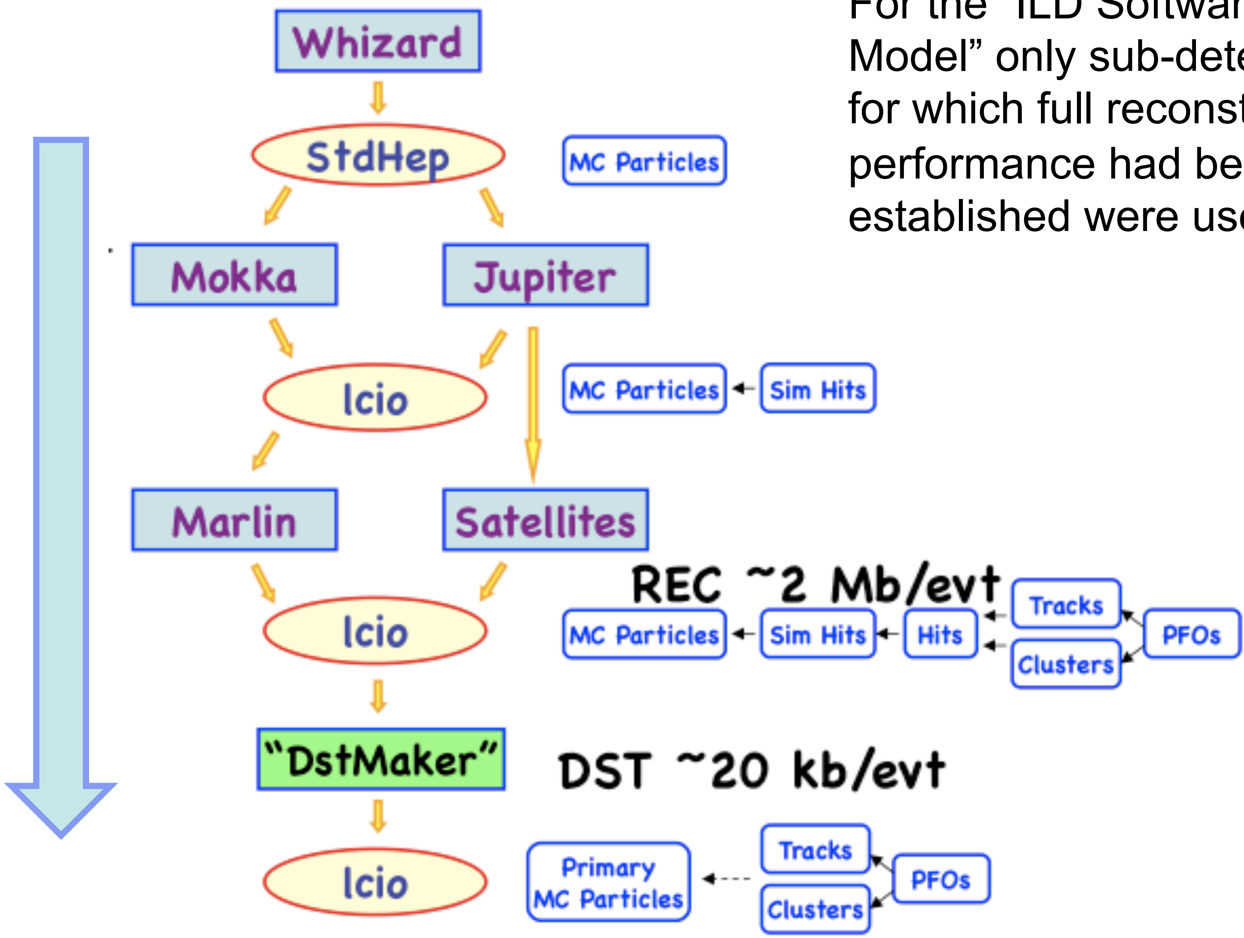
Software Choices

Given the limited amount of time and the number of events required the Mokka – Marlin chain was used for the central simulation and reconstruction of the LOI physics benchmark samples



Software Choices

Given the limited amount of time and the number of events required the Mokka – Marlin chain was used for the central simulation and reconstruction of the LOI physics benchmark samples



For the "ILD Software Reference Model" only sub-detector models for which full reconstruction performance had been established were used.

Input Data – Generators

ILD used the Standard Model background sample which was generated by SLAC and provided to all concepts.

This included all 2 → 2, 4, 6 and some 8 processes in the e+e-, eγ, γγ channels generated via WHIZARD/OMEGA employing full matrix elements.

PYTHIA was used for final state QED and QCD parton showering, fragmentation, and decay.

Backgrounds arising from interactions between virtual and beamstrahlung photons were included via Guinea-Pig.

Event samples were weighted to reflect the expected ILC baseline beam polarization configuration of Pe- =80% and Pe+=30%. 50 fb-1 was generated at 500 GeV and weighted by a factor of 10, and a somewhat smaller sample was generated and appropriately weighted for a collision energy of 250 GeV

Input Data – Generators

ILD used the Standard Model background sample which was generated by SLAC and provided to all concepts.

This included all 2 → 2, 4, 6 and some 8 processes in the e+e-, eγ, γγ channels generated via WHIZARD/OMEGA employing full matrix elements.

PYTHIA was used for final state QED and QCD parton showering, fragmentation, and decay.

Backgrounds arising from interactions between virtual and beamstrahlung photons were included via Guinea-Pig.

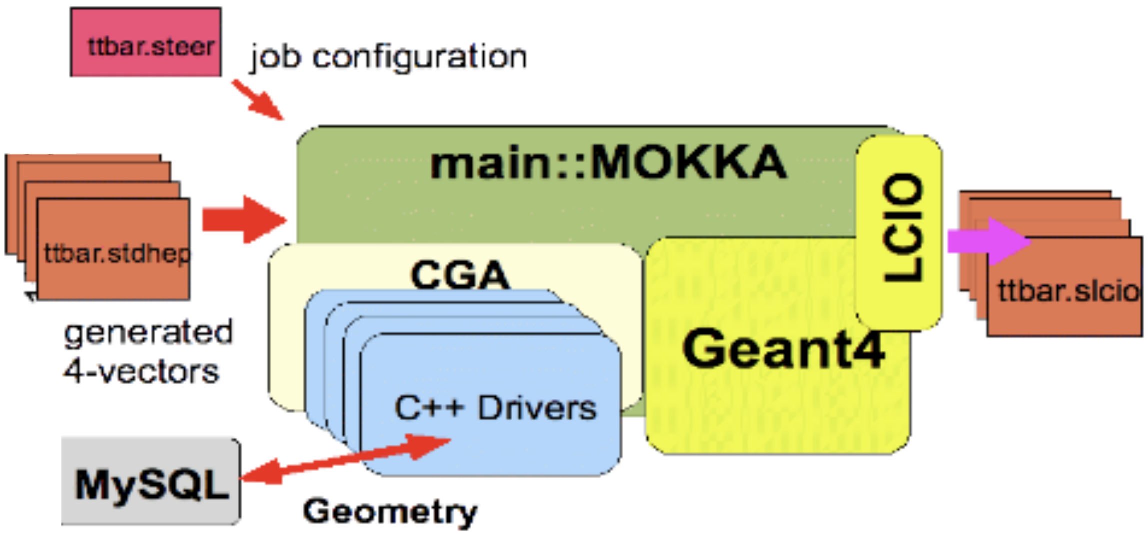
Event samples were weighted to reflect the expected ILC baseline beam polarization configuration of Pe- =80% and Pe+=30%. 50 fb-1 was generated at 500 GeV and weighted by a factor of 10, and a somewhat smaller sample was generated and appropriately weighted for a collision energy of 250 GeV

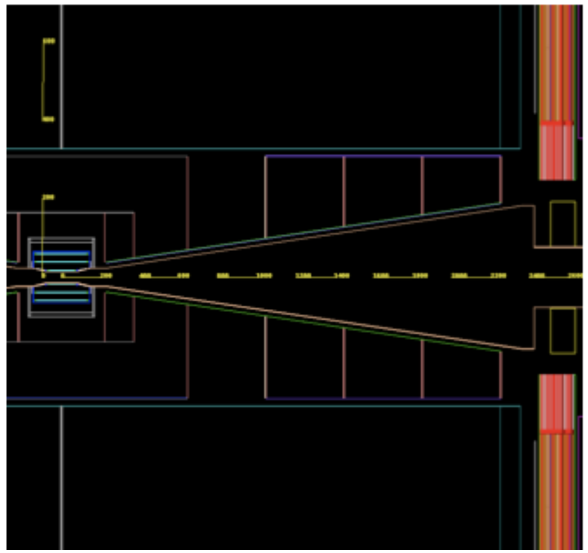
many thanks to the guys @ SLAC

GEANT4 Simulation

Mokka v06-07-patch01

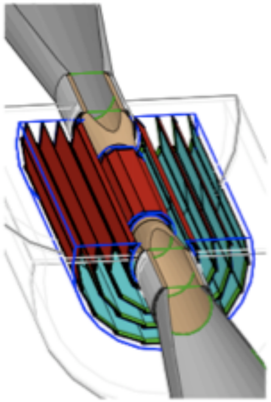
Geant4 – version 9.1 patch 01
LCPhysics physics list
Geometry updated for IDAG Background Studies

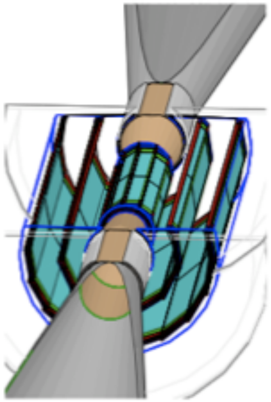


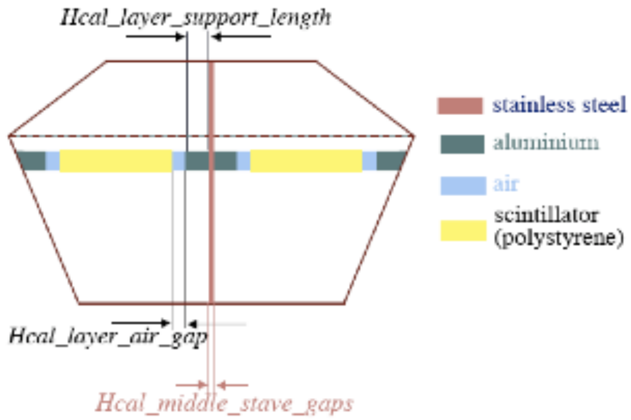


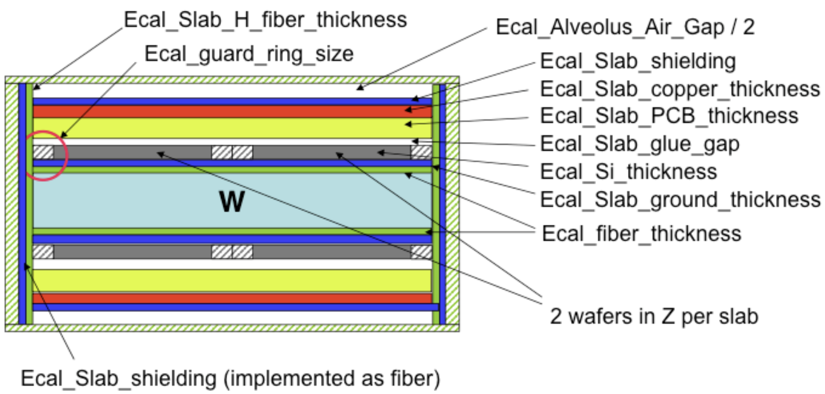
Mokka

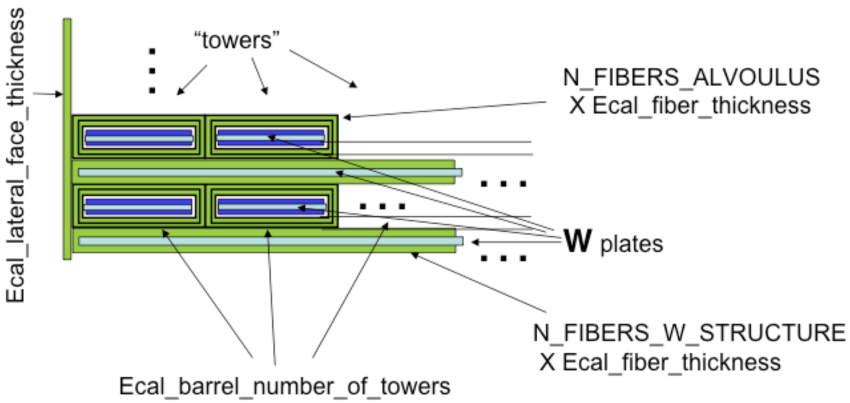
A lot of effort made to include as much engineering detail as possible



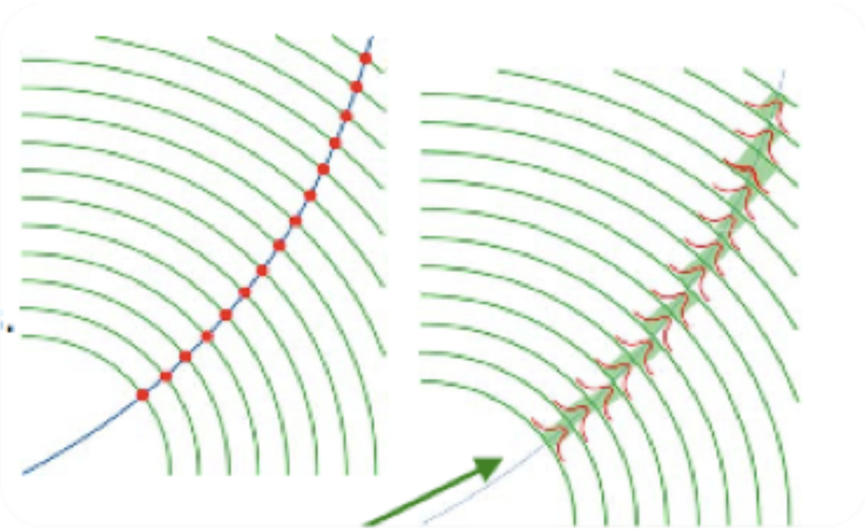








Digitisation – Trackers



	$\sigma_{r-\phi}/\mu\mathrm{m}$	$\sigma_z/\mu\mathrm{m}$		$\sigma_{r-\phi}/\mu\mathrm{m}$	$\sigma_z/\mu\mathrm{m}$
VTX	2.8	2.8	FTD	5.8	5.8
SIT/SET	7.0	50.0	ETD	7.0	7.0
TPC	$\sigma^2_{r\phi} = 50^2 + 900^2 \sin^2\phi + ((25^2/22) \times (4/B)^2 \sin\theta)\, z\, \mu\mathrm{m}^2$ $\sigma^2_z = 40^2 + 8^2 \times z\, \mu\mathrm{m}^2$				

Digitisation – Calorimeters

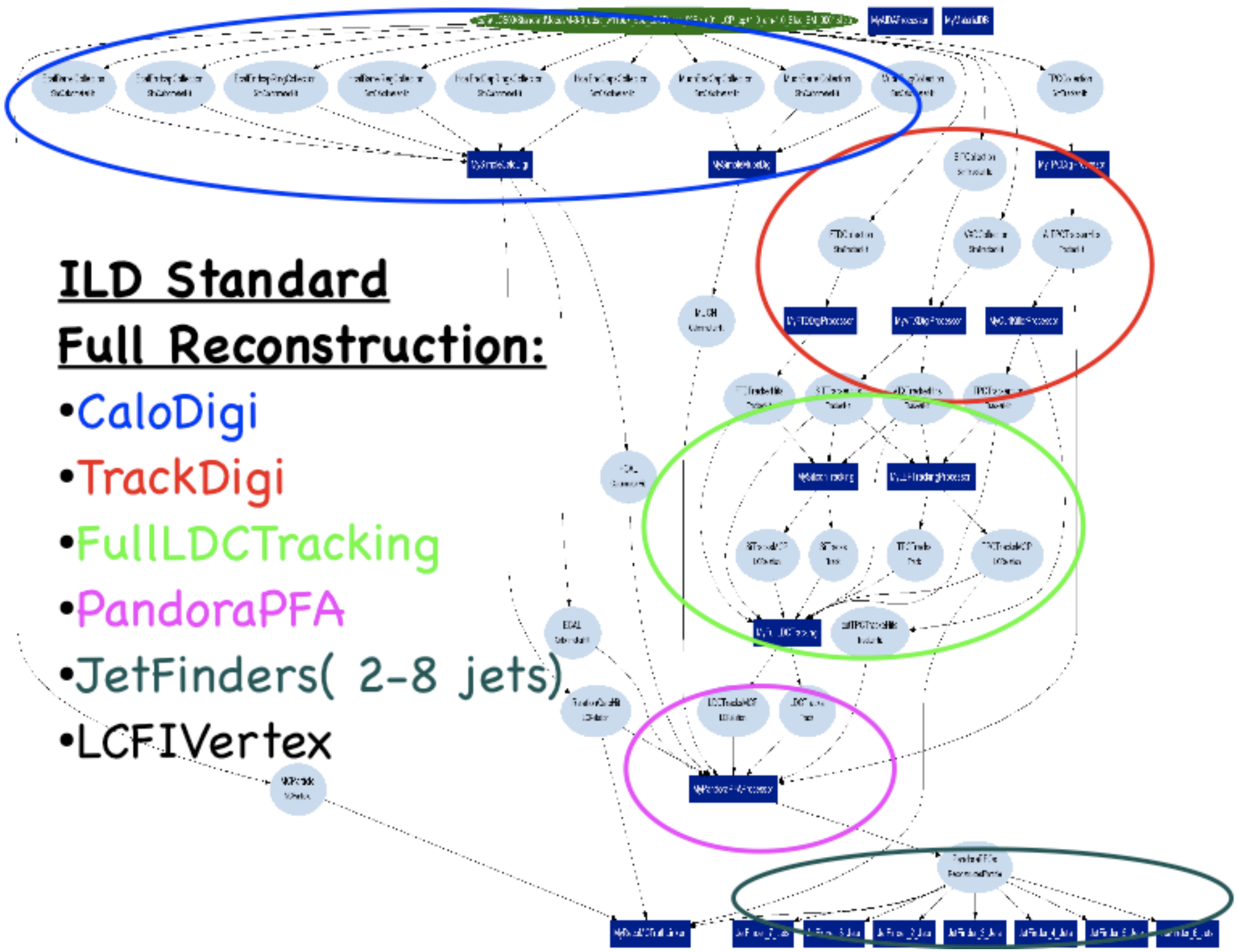
Takes SimCalorimeterHits converts to CalorimeterHits

Converts simulated energy deposits in active layers into physical energy taking into account sampling fractions.

Conversion factors calibrated using single particles scans.

Different regions may use different conversion factors.

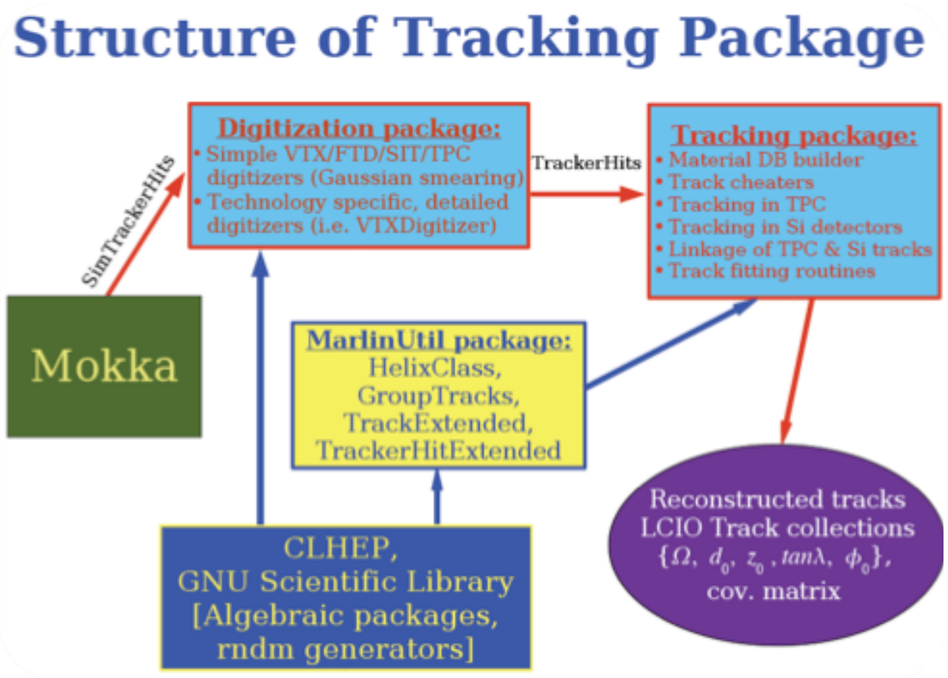
Reconstruction Software



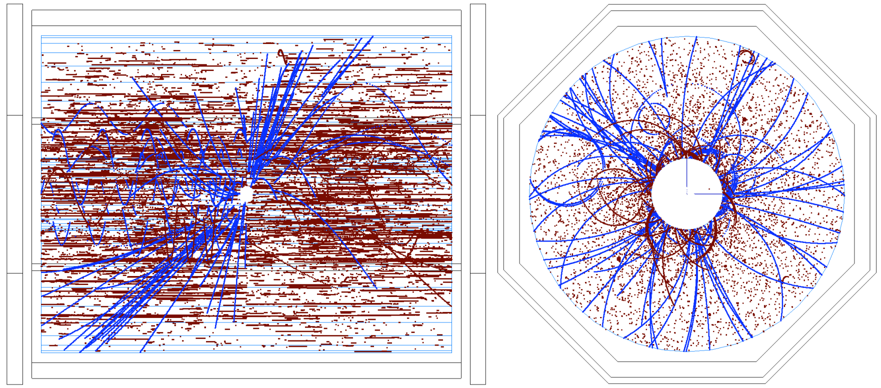
ILD Standard
Full Reconstruction:

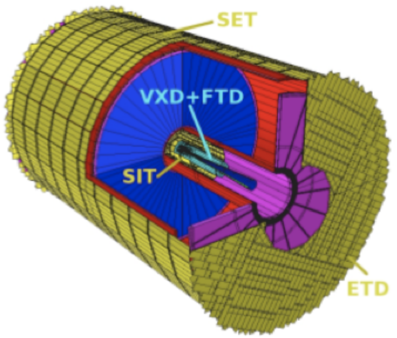
- CaloDigi
- TrackDigi
- FullLDCTracking
- PandoraPFA
- JetFinders(2-8 jets)
- LCFIVertex

Tracking

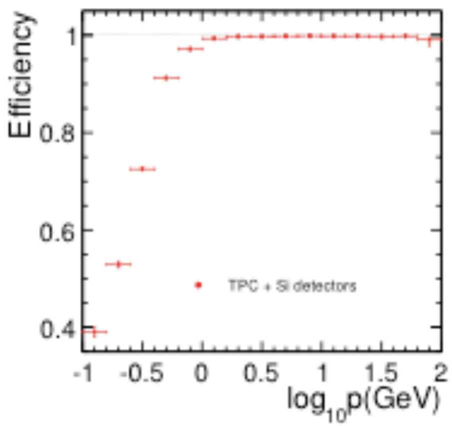


Structure of Tracking Package
Digitization package:
• Simple VTX/FTD/SIT/TPC digitizers (Gaussian smearing)
• Technology specific, detailed digitizers (i.e. VTXDigitizer)
TrackerHits
Tracking package:
• Material DB builder
• Track cheaters
• Tracking in TPC
• Tracking in Si detectors
• Linkage of TPC & Si tracks
• Track fitting routines
SimTrackerHits
Mokka
MarlinUtil package:
HelixClass,
GroupTracks,
TrackExtended,
TrackerHitExtended
Reconstructed tracks
LCIO Track collections
{Ω, d0, z0, tanλ, φ0},
cov. matrix
CLHEP,
GNU Scientific Library
[Algebraic packages,
rndm generators]

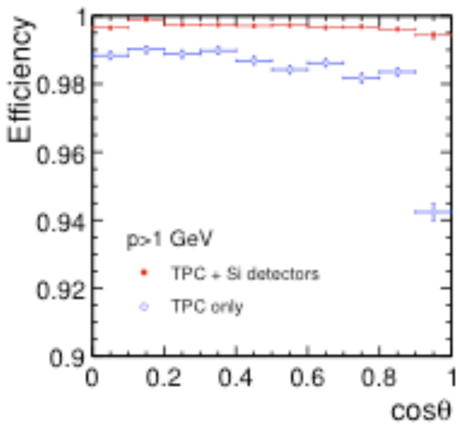




SET
VXD+FTD
SIT
ETD

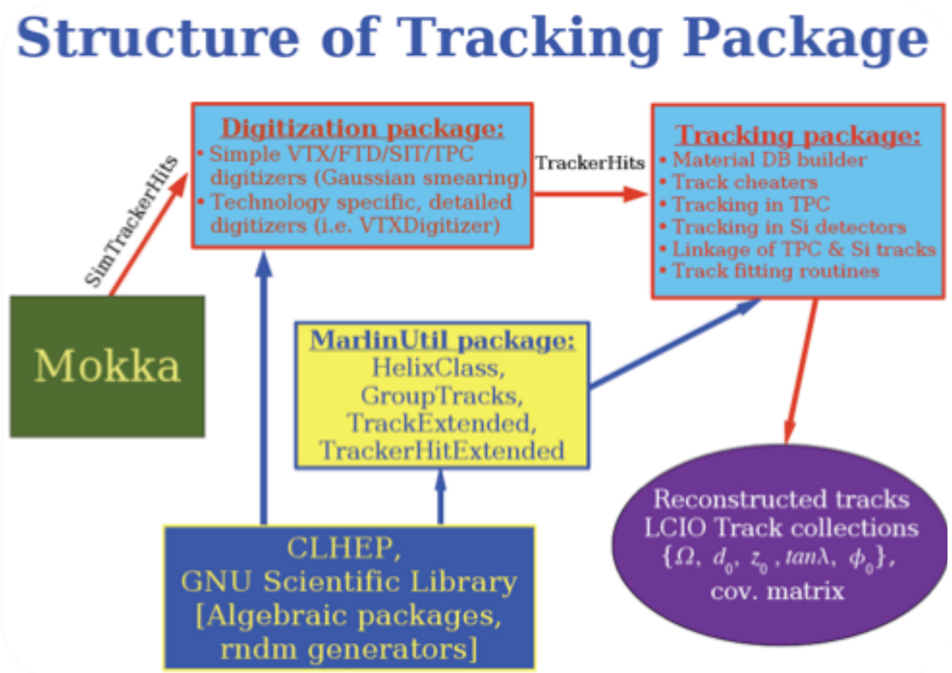


Efficiency
1
0.8
0.6
0.4
-1 -0.5 0 0.5 1 1.5 2
log10p(GeV)
TPC + Si detectors

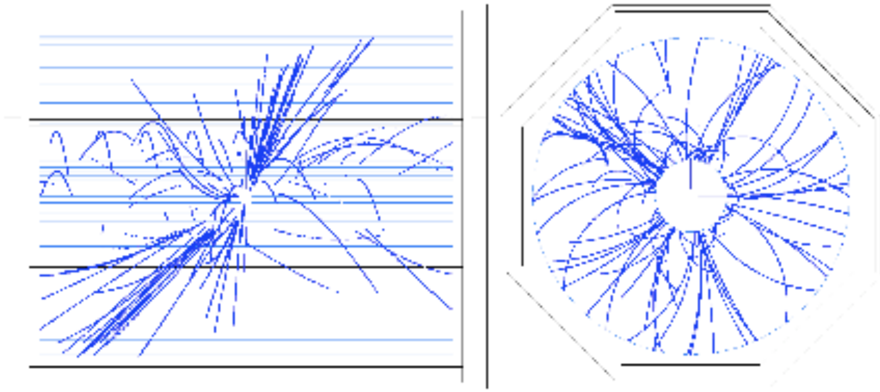


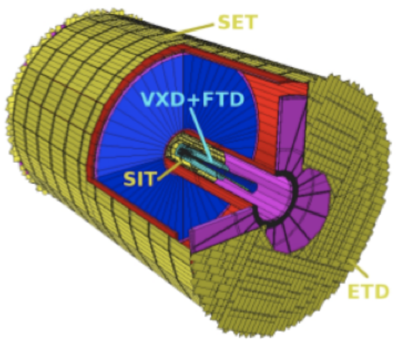
Efficiency
1
0.98
0.96
0.94
0.92
0.9
0 0.2 0.4 0.6 0.8 1
cosθ
p>1 GeV
TPC + Si detectors
TPC only

Tracking

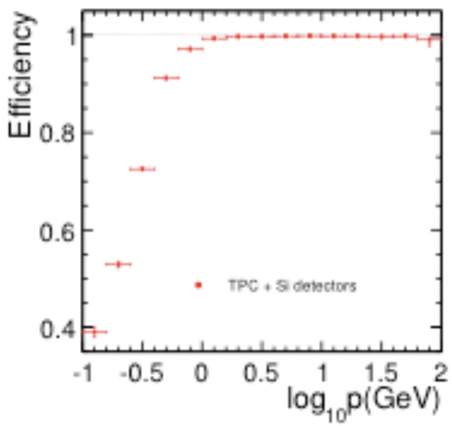


Structure of Tracking Package
Digitization package:
• Simple VTX/FTD/SIT/TPC digitizers (Gaussian smearing)
• Technology specific, detailed digitizers (i.e. VTXDigitizer)
TrackerHits
Tracking package:
• Material DB builder
• Track cheaters
• Tracking in TPC
• Tracking in Si detectors
• Linkage of TPC & Si tracks
• Track fitting routines
SimTrackerHits
Mokka
MarlinUtil package:
HelixClass,
GroupTracks,
TrackExtended,
TrackerHitExtended
Reconstructed tracks
LCIO Track collections
{Ω, d0, z0, tanλ, φ0},
cov. matrix
CLHEP,
GNU Scientific Library
[Algebraic packages,
rndm generators]

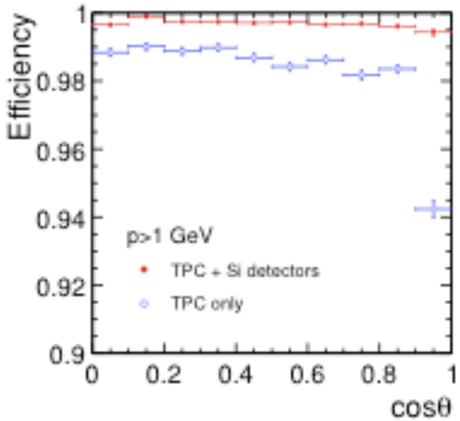




SET
VXD+FTD
SIT
ETD



Efficiency
1
0.8
0.6
0.4
-1 -0.5 0 0.5 1 1.5 2
log10p(GeV)
TPC + Si detectors



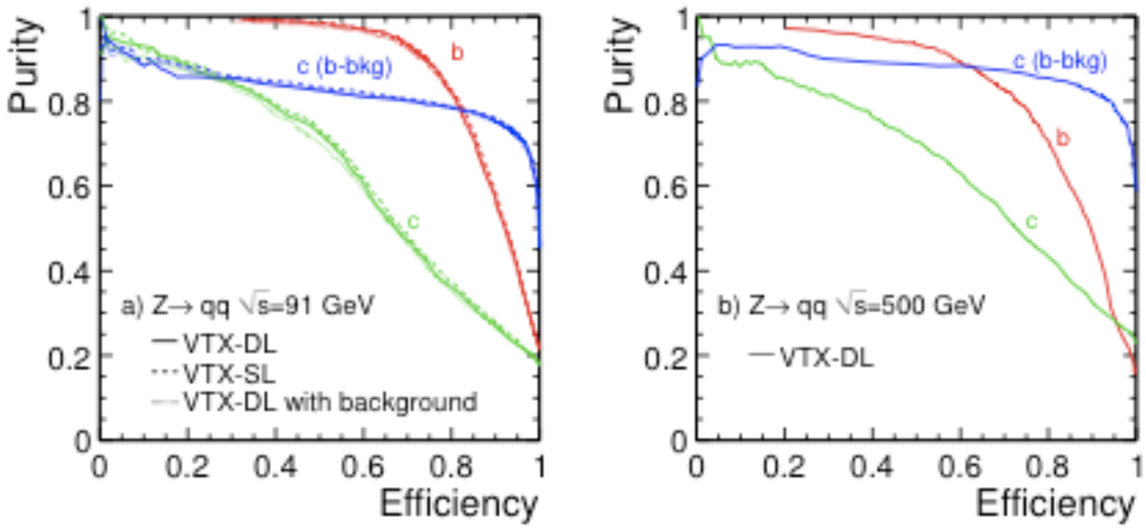
Efficiency
1
0.98
0.96
0.94
0.92
0.9
0 0.2 0.4 0.6 0.8 1
cosθ
p>1 GeV
TPC + Si detectors
TPC only

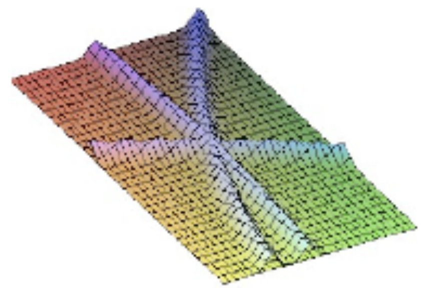
Flavour Tagging

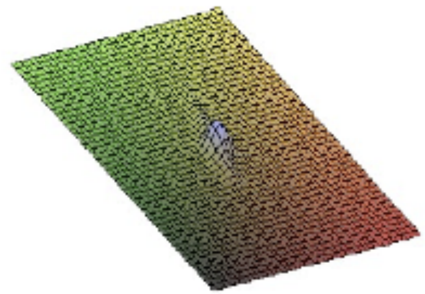
LCFIVertex flavour tagging uses Artificial Neural Networks to discriminate between jets of different quark flavour.

The ANN's were trained using samples consisting of 150k events Z → qq at the Z pole equally distributed among the three decay modes q=b,c and light quarks

Test samples of 10k events Z → qq at √s = 91 GeV and s=√500 were then used to evaluate the ILD flavour tagging performance



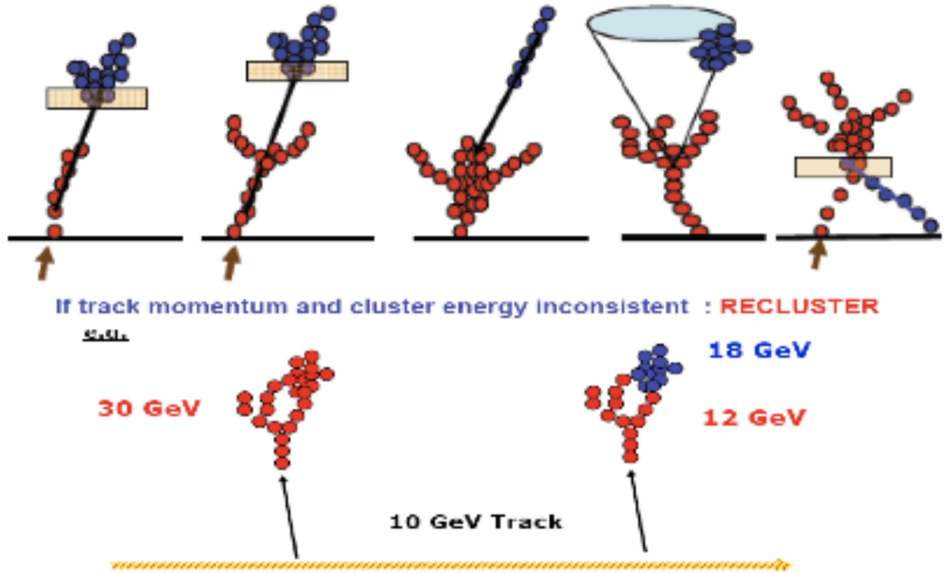


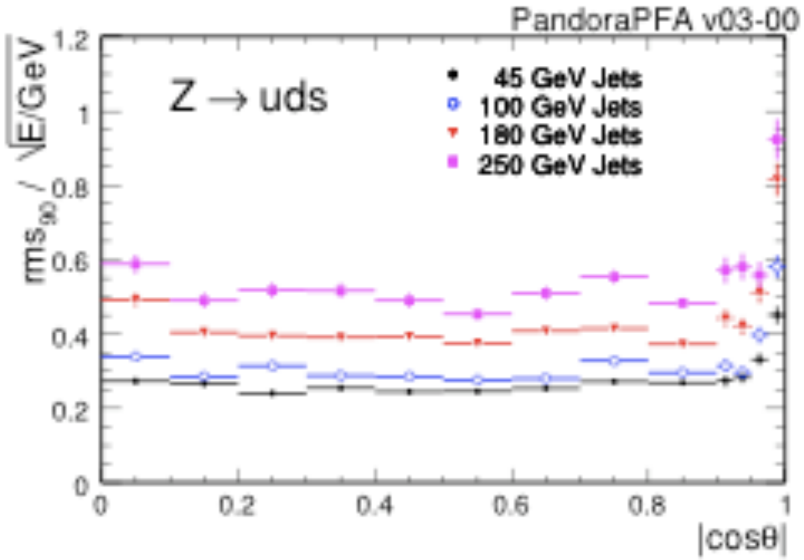


Particle Flow

The **PandoraPFA** package is currently the most sophisticated and highly performing particle flow algorithm available.

One of the key outcomes of which has been its demonstration of the performance of PFA at s = √1TeV





Jet Energy	raw rms	rms_{90}	$rms_{90}/\sqrt{E_{jj}/GeV}$	σ_{E_j}/E_j
45 GeV	3.3 GeV	2.4 GeV	25.0 %	(3.71 ± 0.05) %
100 GeV	5.8 GeV	4.1 GeV	29.5 %	(2.95 ± 0.04) %
180 GeV	11.2 GeV	7.5 GeV	40.1 %	(2.99 ± 0.04) %
250 GeV	16.9 GeV	11.1 GeV	50.1 %	(3.17 ± 0.05) %

Production – Grid

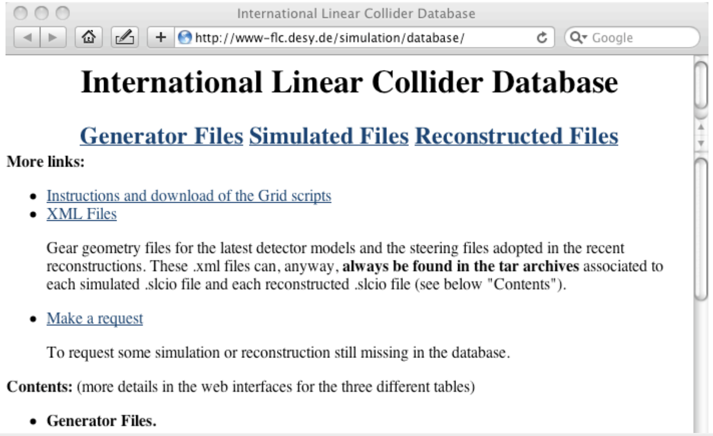
Over 50 Million events fully simulated and reconstructed

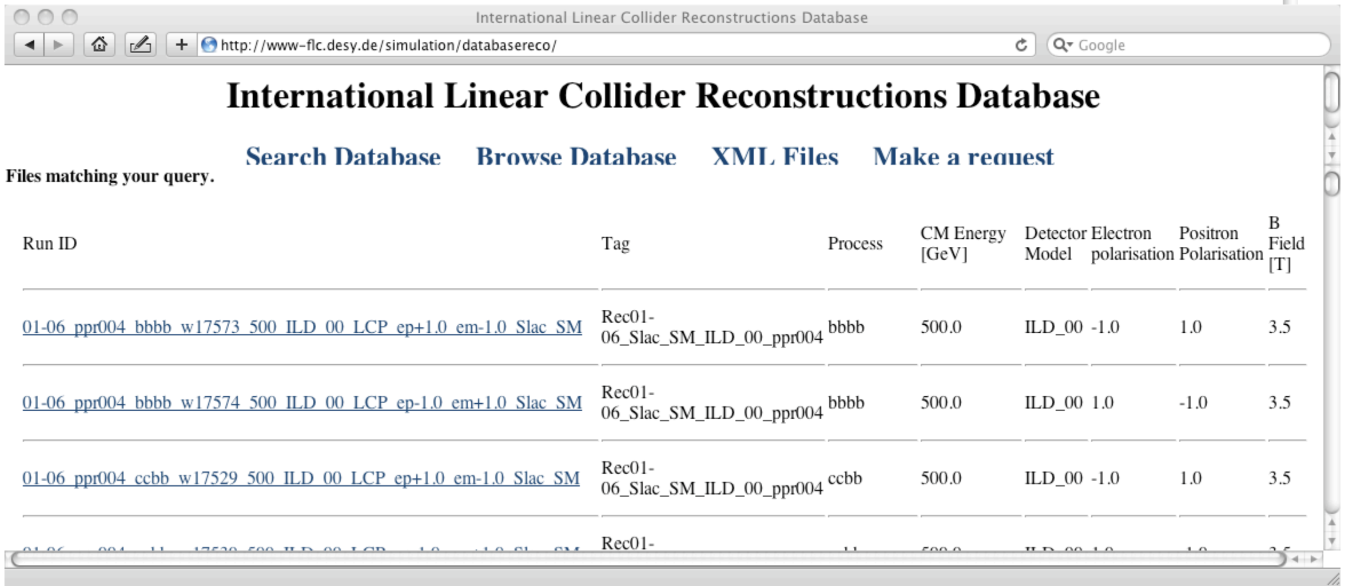
Producing > TB of Data

Over 350k Grid jobs

Production done at only a handful of sites DESY, Lyon, LAL …

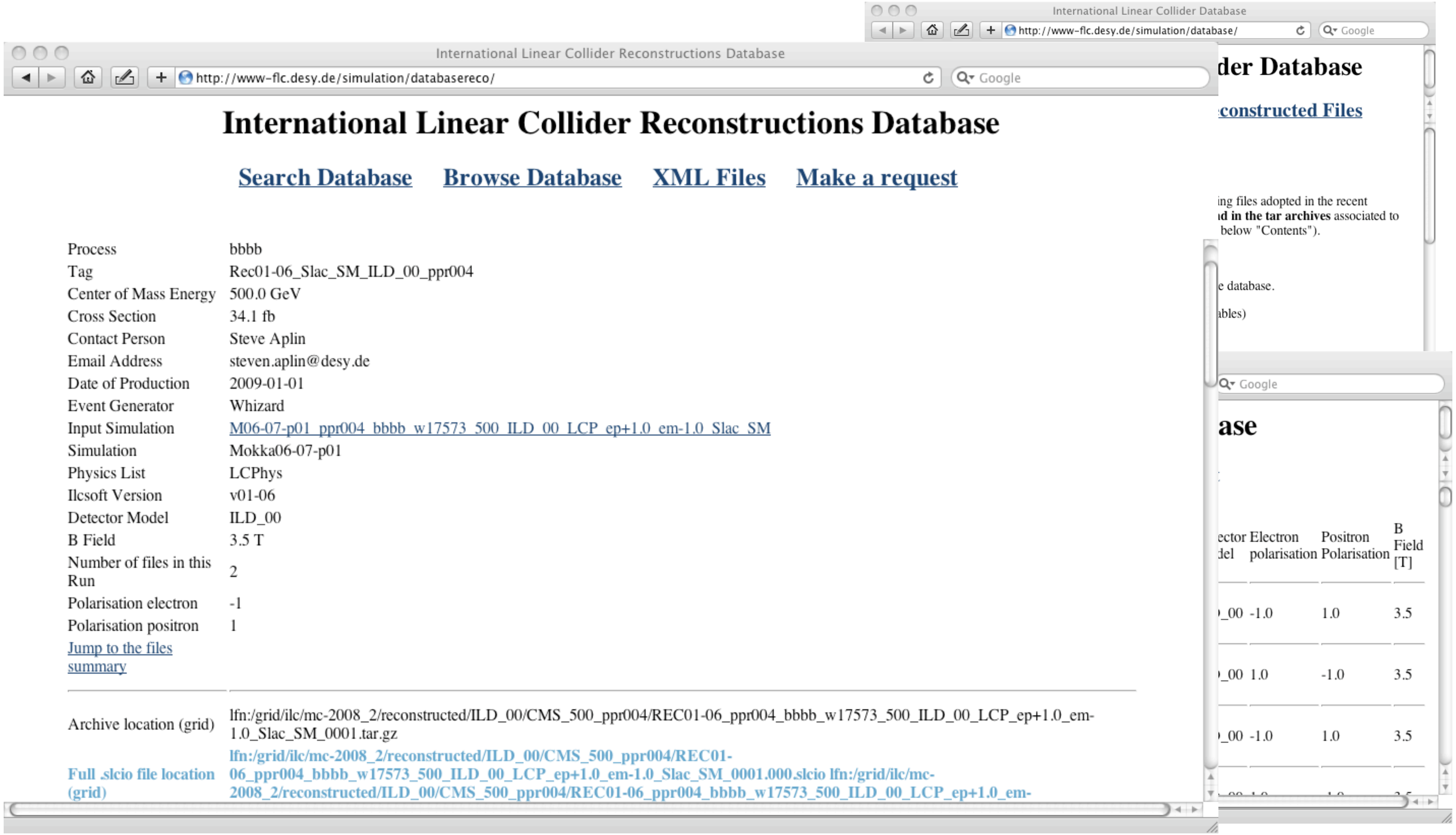
Data sets published to users (also in combined DST's) via DB as it became available





Run ID	Tag	Process	CM Energy [GeV]	Detector Model	Electron polarisation	Positron Polarisation	B Field [T]
01-06_ppr004_bbbb_w17573_500_ILD_00_LCP_ep+1.0_em-1.0_Slac_SM	Rec01-06_Slac_SM_ILD_00_ppr004	bbbb	500.0	ILD_00	-1.0	1.0	3.5
01-06_ppr004_bbbb_w17574_500_ILD_00_LCP_ep-1.0_em+1.0_Slac_SM	Rec01-06_Slac_SM_ILD_00_ppr004	bbbb	500.0	ILD_00	1.0	-1.0	3.5
01-06_ppr004_ccbb_w17529_500_ILD_00_LCP_ep+1.0_em-1.0_Slac_SM	Rec01-06_Slac_SM_ILD_00_ppr004	ccbb	500.0	ILD_00	-1.0	1.0	3.5

Production – Grid



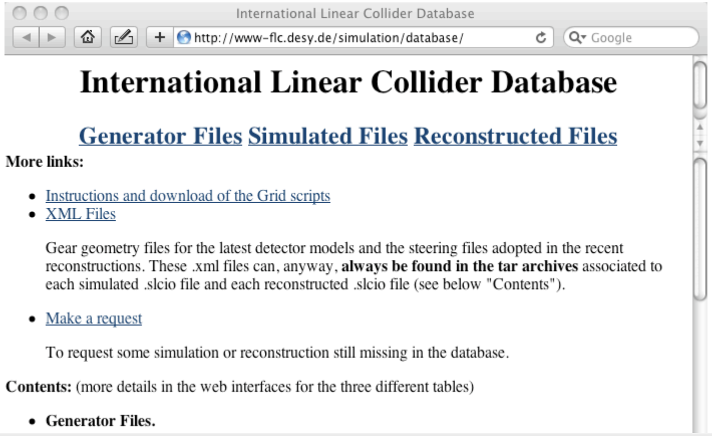
Production – Grid

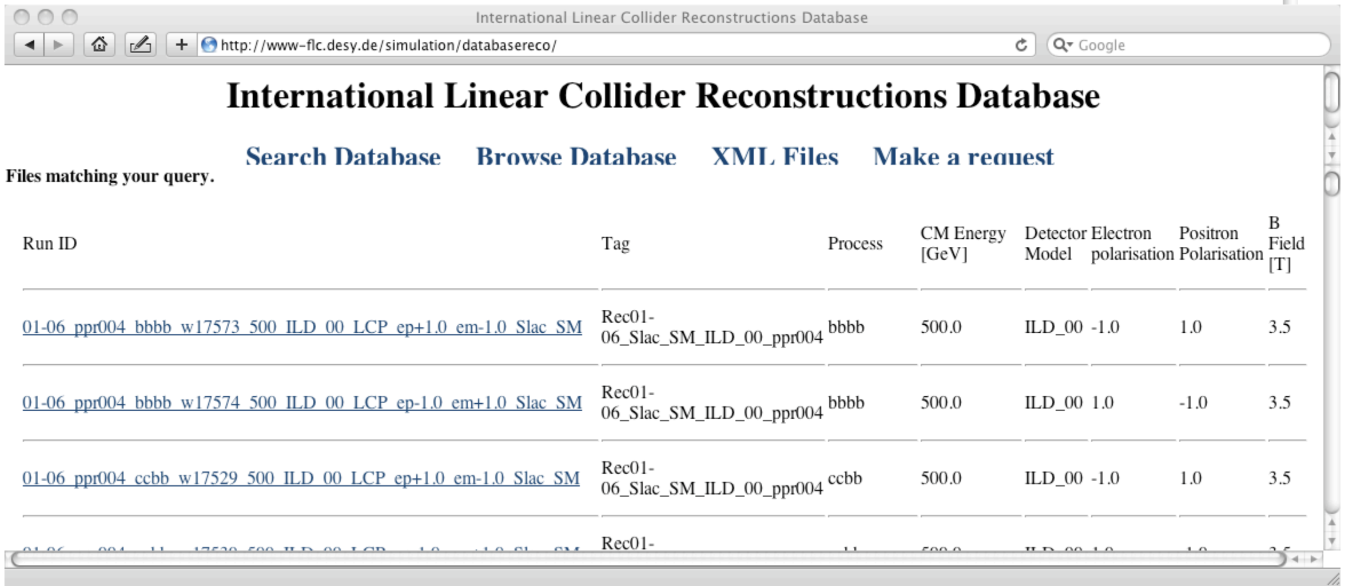
Production suite grown organically from from prototyping …

Based on bash scripting and mysql bookkeeping

Just in time development

Hard lessons learnt are being invested into new production system





Summary

Strategic decisions concerning software, taken long before the LOI process appeared on the horizon, have allowed a very considerable amount of work to be done in a short time under a fair amount of pressure.

Creating modular frameworks to work in, coupled with the provision of common software tools such as LCIO, allowed groups, as well as single authors, to contribute effectively to a complete, detailed and realistic software chain.

Centralised production of Monte Carlo data still remains necessary to effectively utilise the computing resources offered by the GRID. Whilst solutions are at hand which may alleviate this, it remains a man power intensive procedure.

Some form of testing and validation suite would have been invaluable.

"*The ILD efforts on simulating the physics benchmark processes have been impressive.*"

IDAG Report on the Validation of Letters of Intent for ILC detectors