

CALICE Data Processing (From Raw to Physics Data)



Roman Pöschl
LAL Orsay



- Calice Testbeam Data Taking
- Data Management
- Event Building and Reconstruction Software
- Summary and Outlook

EUDET Annual Meeting Palaiseau/France October 2007

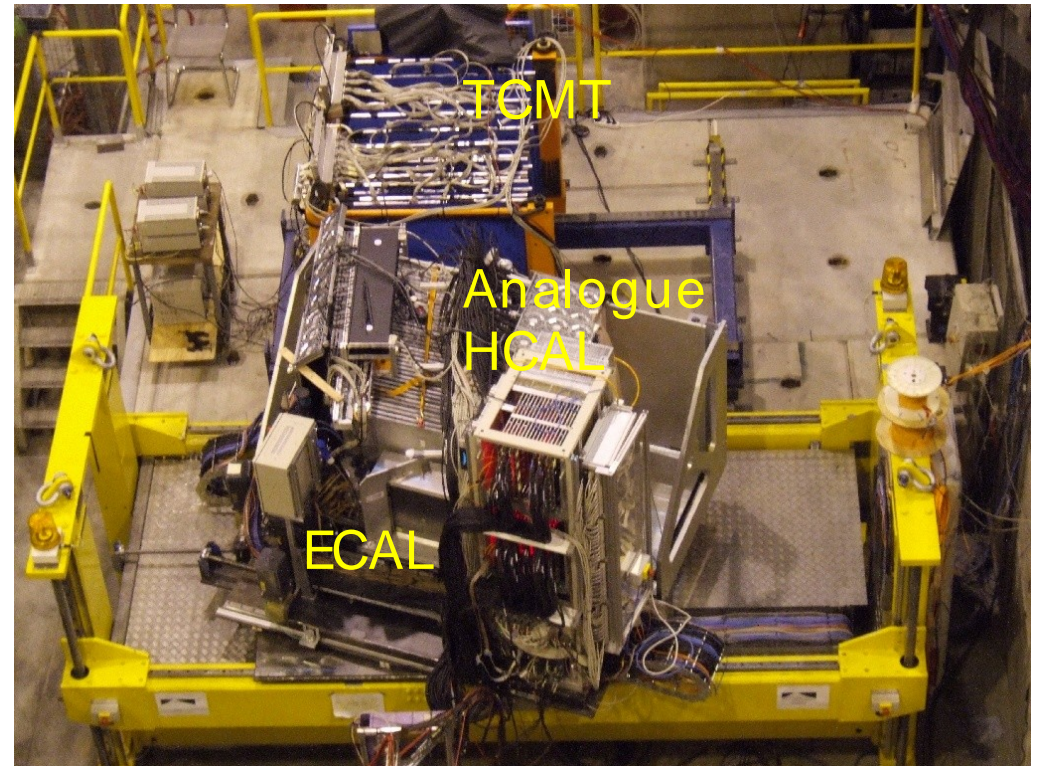
CALICE Testbeam Data Taking

CALICE collaboration is preparing/performing large scale testbeam
Data taking in Summer 2006/2007

Testbeam Setup at CERN 2007

Testbeam program poses
software/computing “
challenges”

- Data processing from
Raw Data to final
Clusters in a
coherent way
- Handling of Conditions Data
Detector Configuration
Calibration, Alignment etc.
- Comparison with simulated
data
'Physics' Output



O(15000) calorimeter cells
readout by Calice DAQ
No Zero Suppression

CALICE "TIER 0" – Infrastructure in the Control Room



Gigabit Uplink

- High Speed Connection to the outside world
- Serves all Calice Control Room Computers

caliceserv.cern.ch

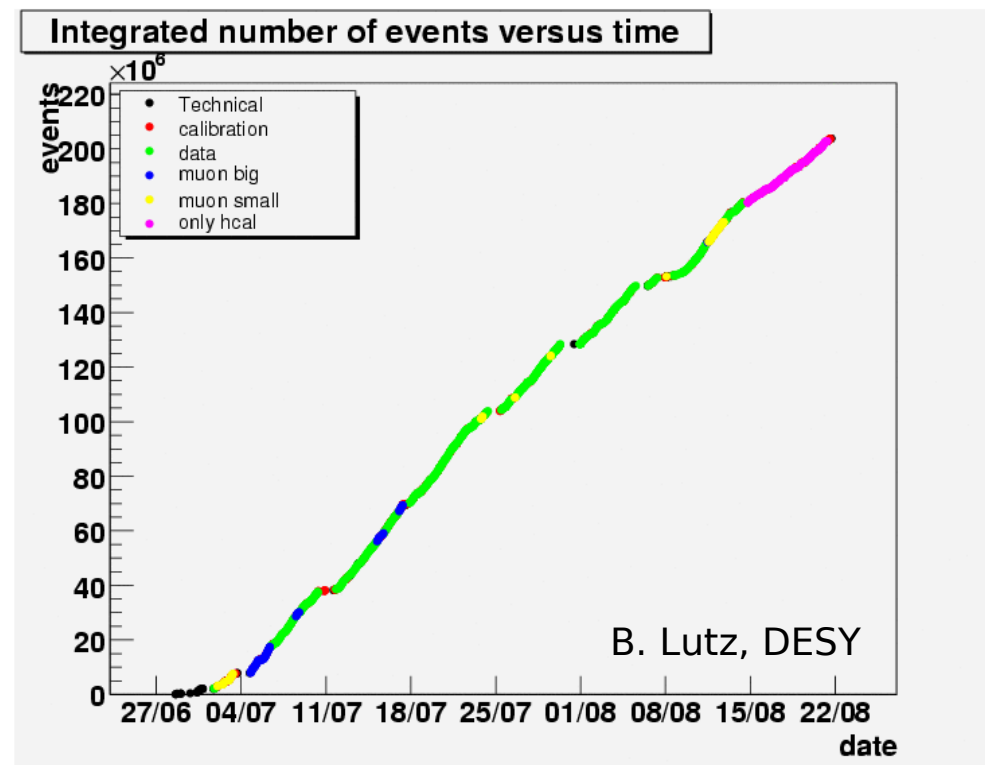
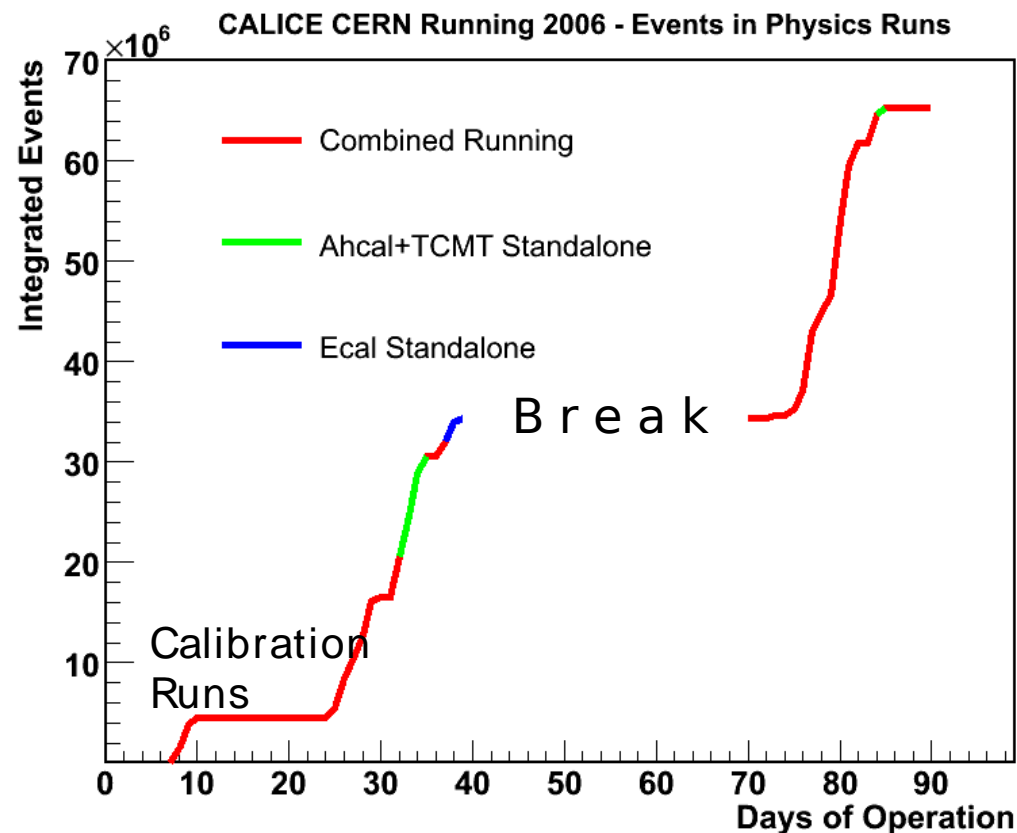
- Online Monitoring
- Grid Transfers

Disk Array

DAQ Computer

Well organized setup of computing
Thanks to B. Lutz

CALICE - CERN Data taking 2006/2007



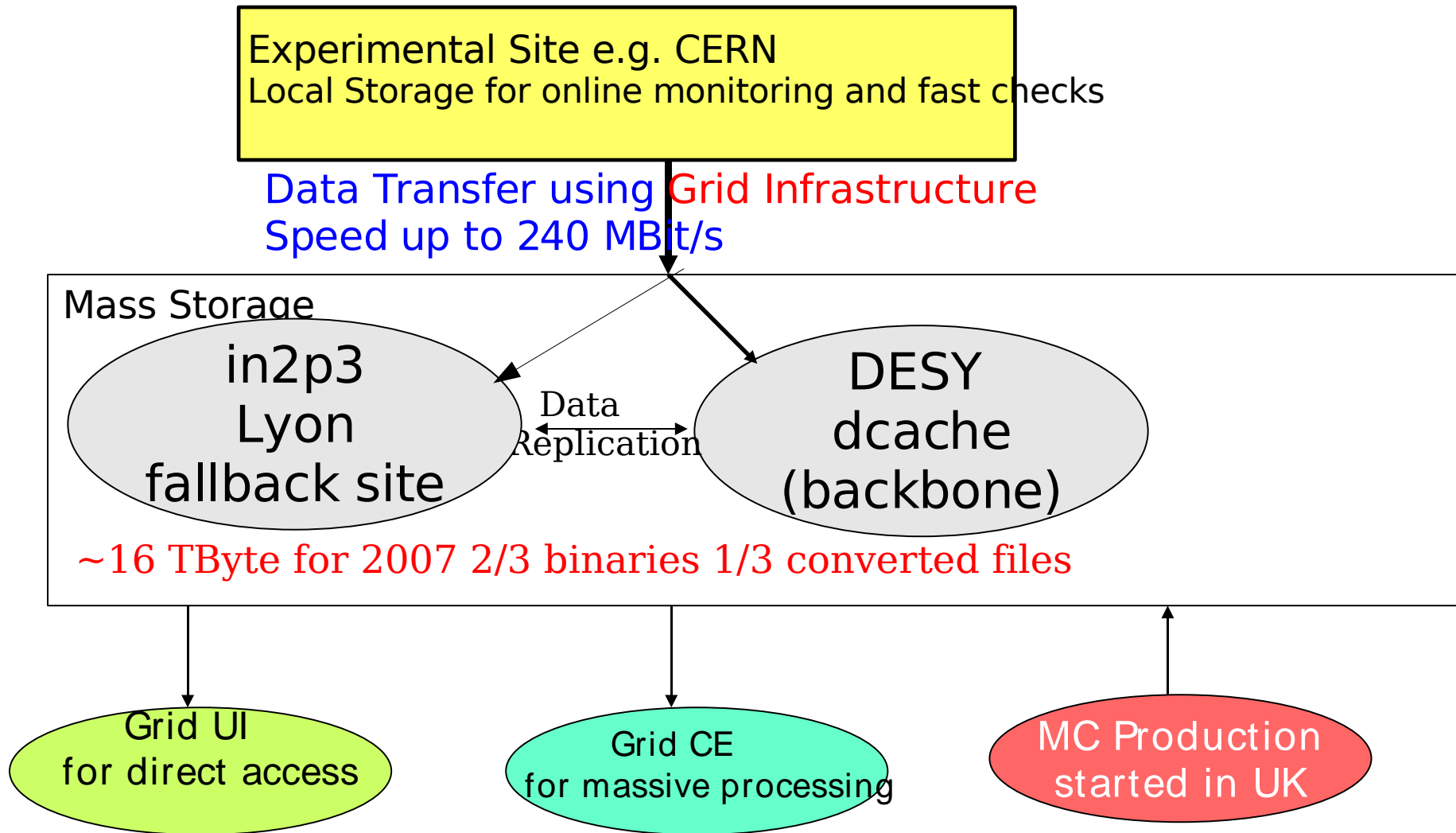
~200 Mio Events
in 'Physics' Runs

+

O(50 Mio). Muon
Calibration Events)

Efficient and fast
way of data distribution
and processing ?

Data Handling and Processing



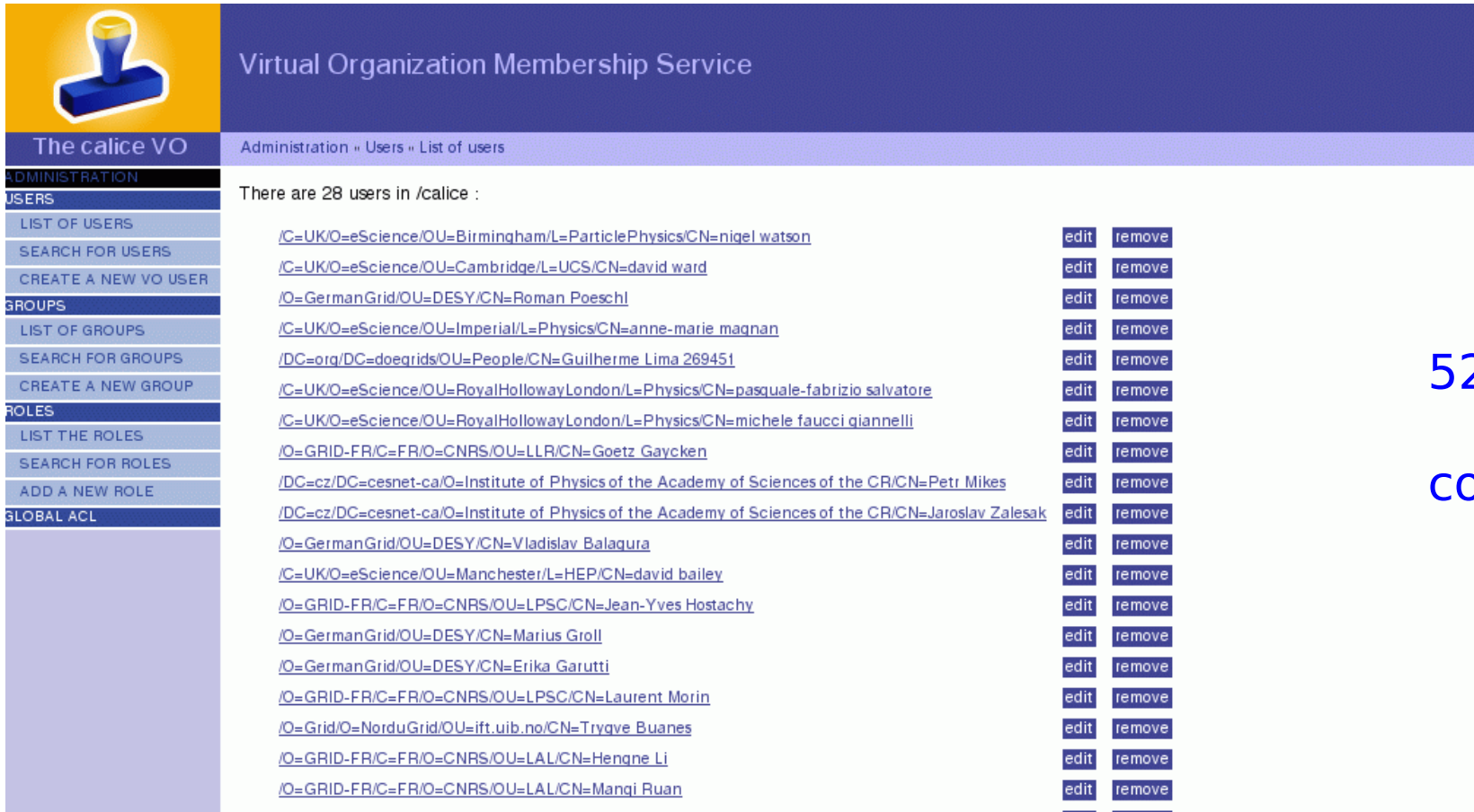
- Raw Data are (usually) available ~20 Min. after Run End
- Delay of Converted Files (usually) < 1 day

CALICE is the first HEP Experiment which uses the grid systematically for real data!!!

The Virtual Organisation - vo calice

Hosted by DESY:

Page for registration is <https://grid-voms.desy.de:8443/voms/calice>



Virtual Organization Membership Service

The calice VO Administration » Users » List of users

There are 28 users in /calice :

/C=UK/O=eScience/OU=Birmingham/L=ParticlePhysics/CN=nigel watson	edit	remove
/C=UK/O=eScience/OU=Cambridge/L=UCS/CN=david ward	edit	remove
/O=GermanGrid/OU=DESY/CN=Roman Poeschl	edit	remove
/C=UK/O=eScience/OU=Imperial/L=Physics/CN=anne-marie magnan	edit	remove
/DC=org/DC=doegrids/OU=People/CN=Guilherme Lima 269451	edit	remove
/C=UK/O=eScience/OU=RoyalHollowayLondon/L=Physics/CN=pasquale-fabrizio salvatore	edit	remove
/C=UK/O=eScience/OU=RoyalHollowayLondon/L=Physics/CN=michele faucci qiannelli	edit	remove
/O=GRID-FR/C=FR/O=CNRS/OU=LLR/CN=Goetz Gaycken	edit	remove
/DC=cz/DC=cesnet-ca/O=Institute of Physics of the Academy of Sciences of the CR/CN=Petr Mikes	edit	remove
/DC=cz/DC=cesnet-ca/O=Institute of Physics of the Academy of Sciences of the CR/CN=Jaroslav Zalesak	edit	remove
/O=GermanGrid/OU=DESY/CN=Vladislav Balagura	edit	remove
/C=UK/O=eScience/OU=Manchester/L=HEP/CN=david bailey	edit	remove
/O=GRID-FR/C=FR/O=CNRS/OU=LPSC/CN=Jean-Yves Hostachy	edit	remove
/O=GermanGrid/OU=DESY/CN=Marius Groll	edit	remove
/O=GermanGrid/OU=DESY/CN=Erika Garutti	edit	remove
/O=GRID-FR/C=FR/O=CNRS/OU=LPSC/CN=Laurent Morin	edit	remove
/O=Grid/O=NorduGrid/OU=ift.uib.no/CN=Trygve Buanes	edit	remove
/O=GRID-FR/C=FR/O=CNRS/OU=LAL/CN=Hengne Li	edit	remove
/O=GRID-FR/C=FR/O=CNRS/OU=LAL/CN=Manqi Ruan	edit	remove

52 Members
and
counting ...

VO Manager: R.P./ LAL, Deputy: A. Gellrich/ DESY

Institutes which provide Grid support for Calice

Supported by: **DESY Hamburg**

LAL

LLR

DESY Zeuthen

Imperial College

Birmingham

cc in2p3 Lyon

Cambridge

Institute of Physics

Prague

University College

KEK

Manchester

CIEMAT Madrid

Fermilab

Univ. Liverpool

Univ. Regina

Hosting, Computing and Storage

Computing and Storage

Computing and Storage

Computing and Storage

Computing and Storage

Computing and Storage

Computing and Storage

Computing and Storage

Computing and Storage

(in preparation)

Computing and Storage

Computing and Storage

Computing and Storage

Computing and Storage

Computing and Storage

Exploit started between Fermilab and
NIU Colleagues

Resources Provided (not yet exploited)

Offer Received

- Most of the sites have been involved in recent data and MC processing
Smaller Problems at Manchester and KEK (about to be solved)

The Grid kept me/us busy – Problems encountered

- Transfers blocked due to server overload

Justification of having (at least) two major sites at hand
One site (desy or cc lyon) usually was well performing

- Slow dcache doors at desy

- Hacker attack on desy

- Human Errors

e.g. Corrupt mapping file at desy

Very fast and efficient response by experts

Thanks to their support

- Processing sometimes failed due to several problems

Connectivity problems

Full disks at the various sites

- Processing slowed down in general by 'poor' connectivity between sites

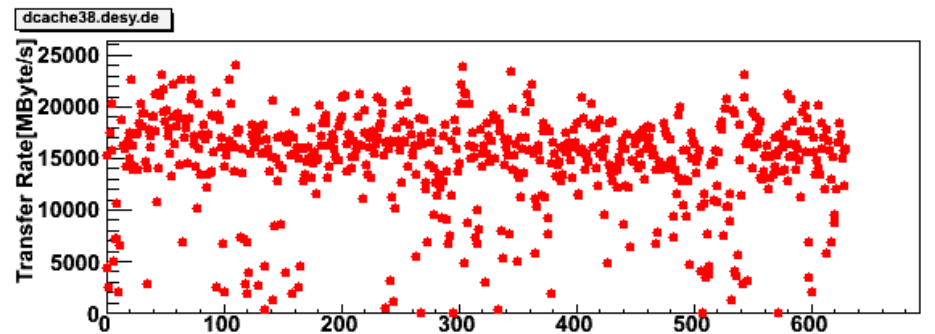
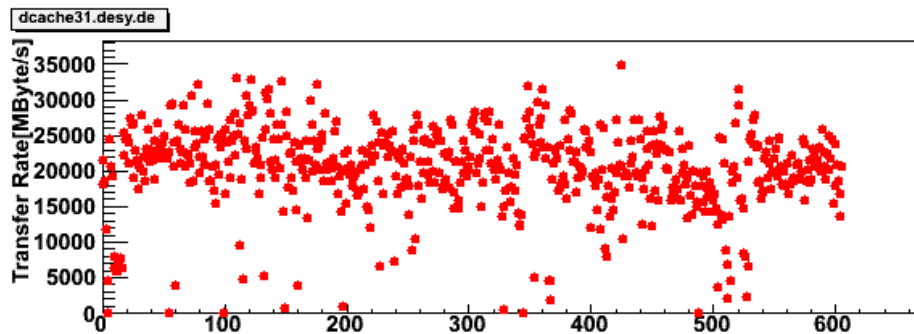
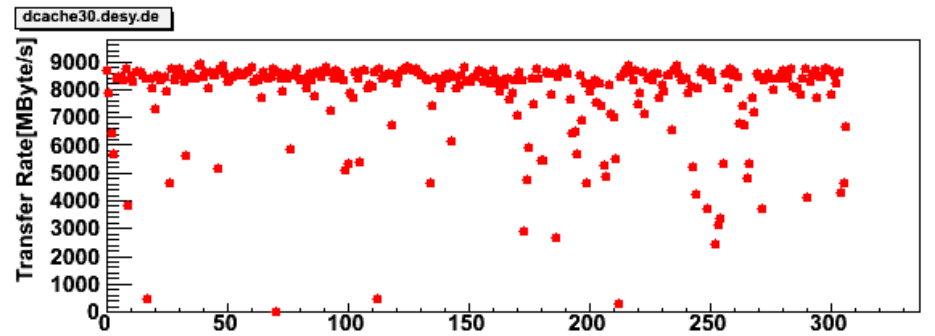
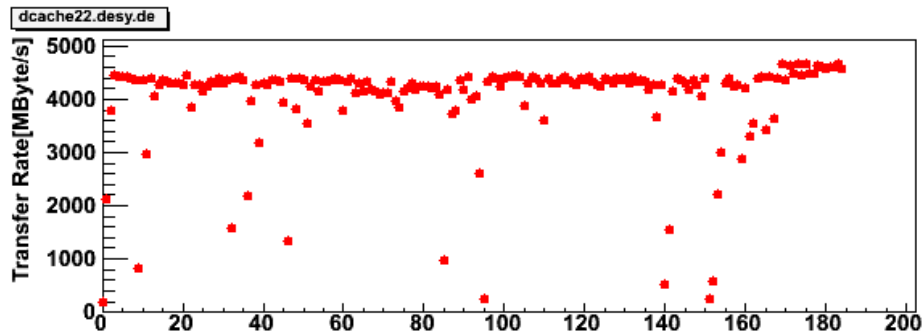
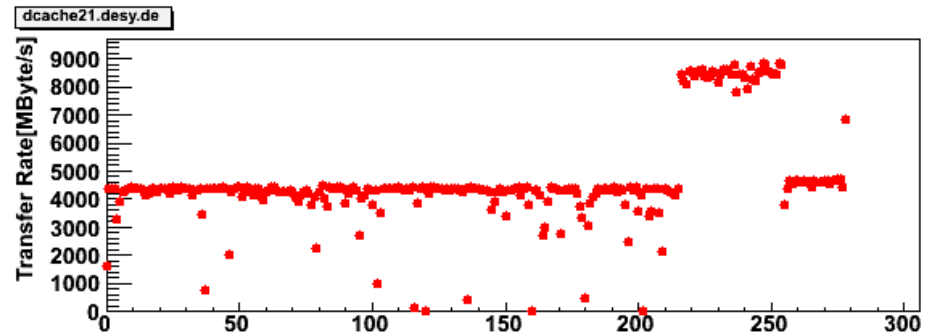
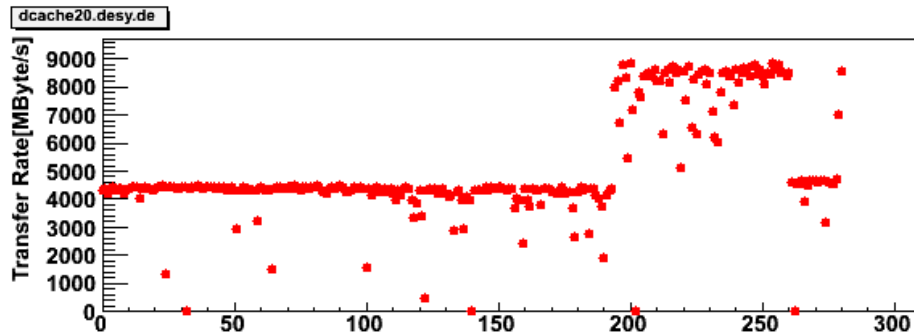
It looks to me as if all sites are well connected to cern

but badly connected among each other (e.g. ccin2p3 <-> desy 2-3 Mbyte/s, ridiculous!!!).

A major infrastructural issue!!!!

Details of Data Transfers – DESY I

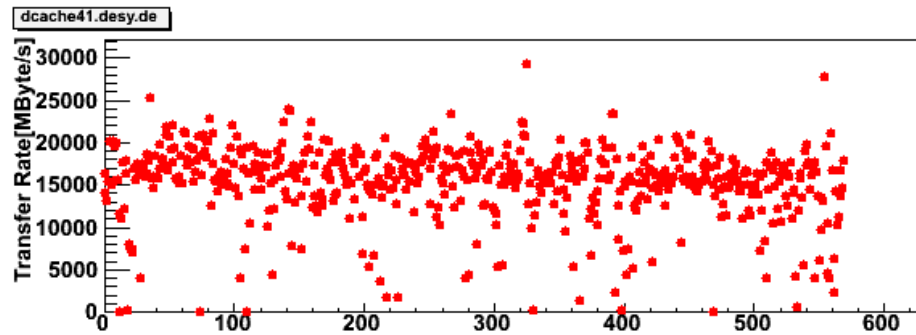
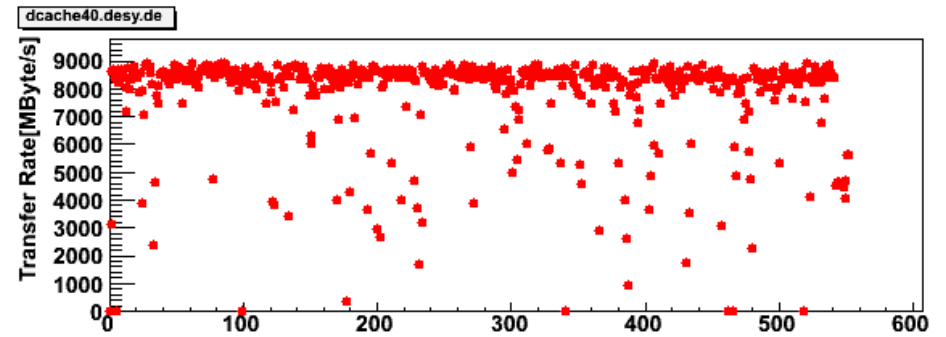
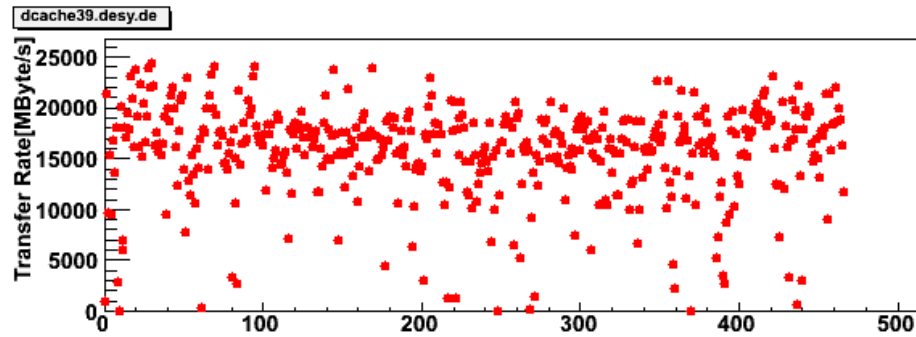
Transfer Rates Control Room -> Mass Storage obtained with different dcache doors
Transfers realized using lcg software and with 4 parallel streams



Average Rates between 4 and 22 MByte/s <- Proof of quality of our equipment
Different performance of different dcache doors not yet understood

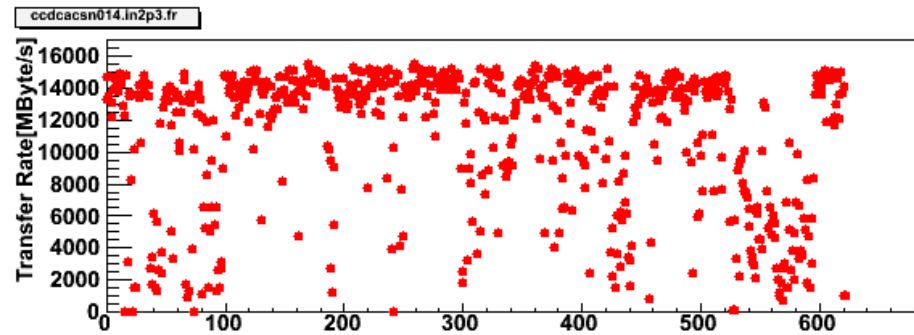
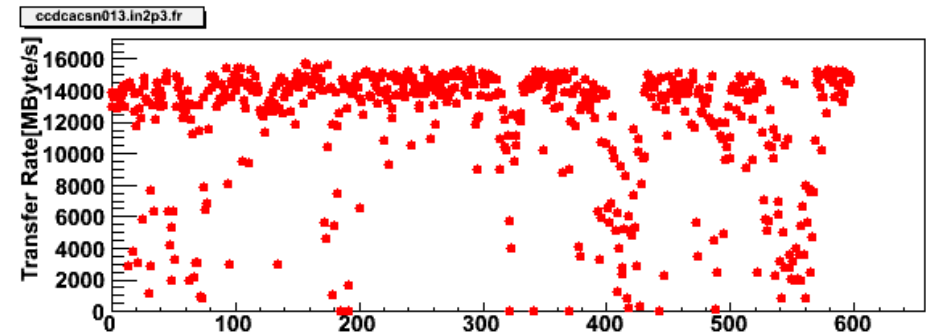
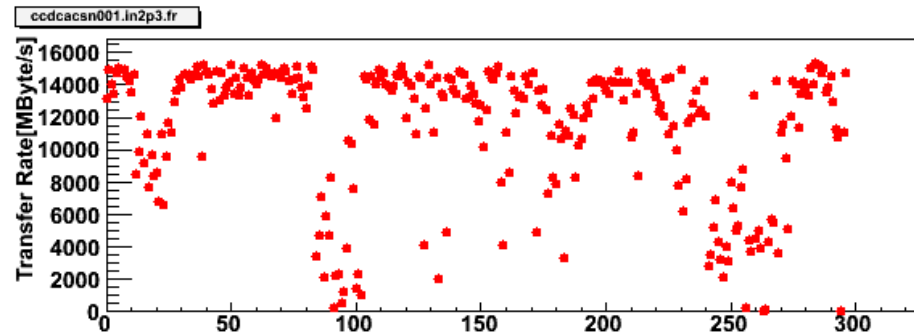
Details of Data Transfers – DESY II

Transfer Rates Control Room -> Mass Storage obtained with different dcache doors
Transfers realized using lcg software and with 4 parallel streams



Details of Data Transfers – CC in2p3 Lyon

Transfer Rates Control Room -> Mass Storage obtained with different dcache doors
Transfers realized using lcg software and with 4 parallel streams

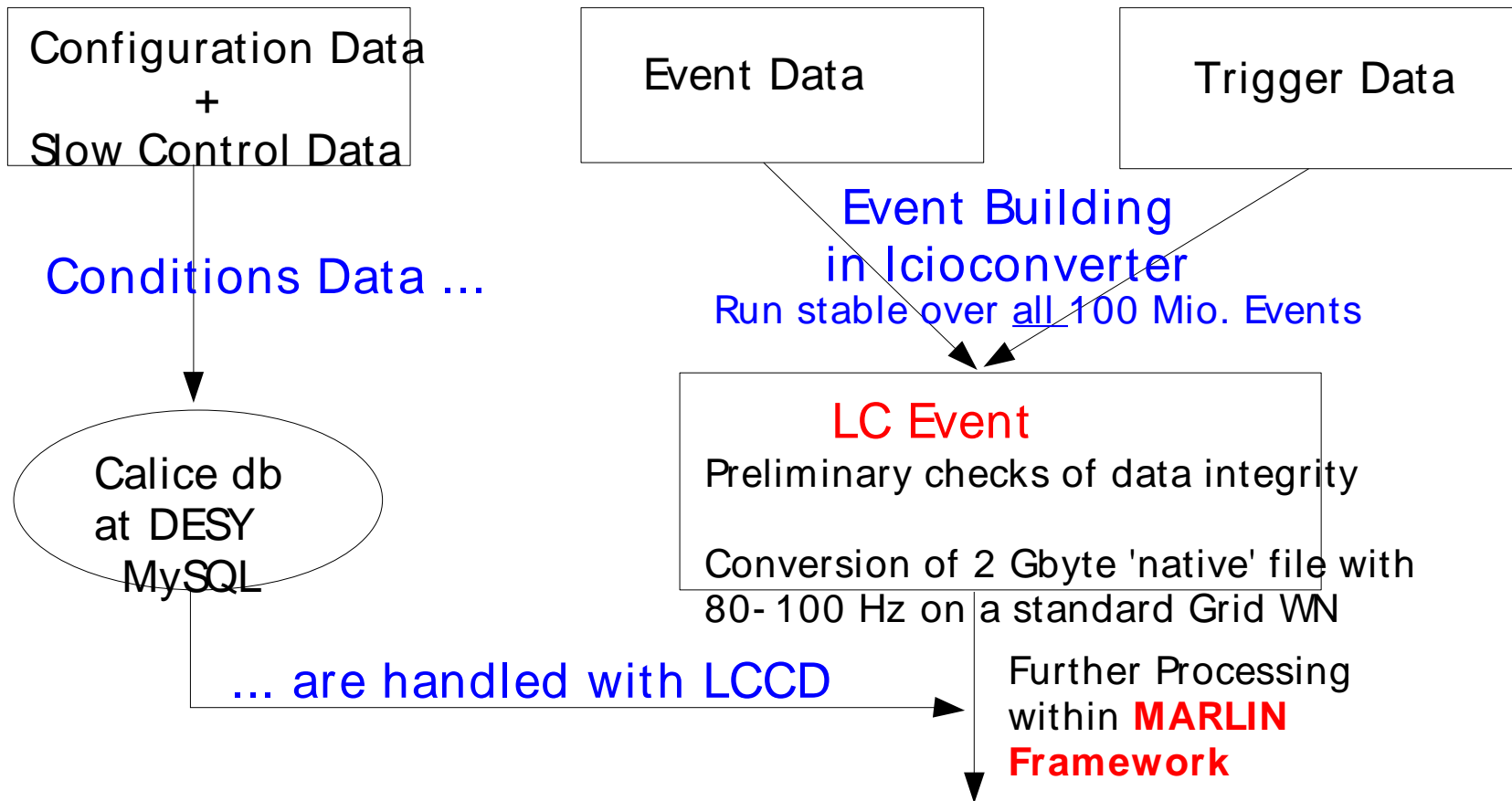


Average Rates ~14 MByte/s - Independant of dcache door

Conversion to LCIO

DAQ data types are converted/ wrapped into LCIO on the basis of [LCGenericObjects](#)

DAQ Data Files/ Types



Remark: LCIO and ILC software framework is not needed to analyze calice data but using it delivers important input for future ILC s/w development
-> General ILC Concept for low level data handling

Intermezzo – Conditions Data Handling

- LCCD – Linear Collider Conditions Data Framework:
 - Software package providing an Interface to conditions data
 - database
 - LCIO files

Author Frank Gaede, DESY

LCCD works and is heavily used within calice !!!

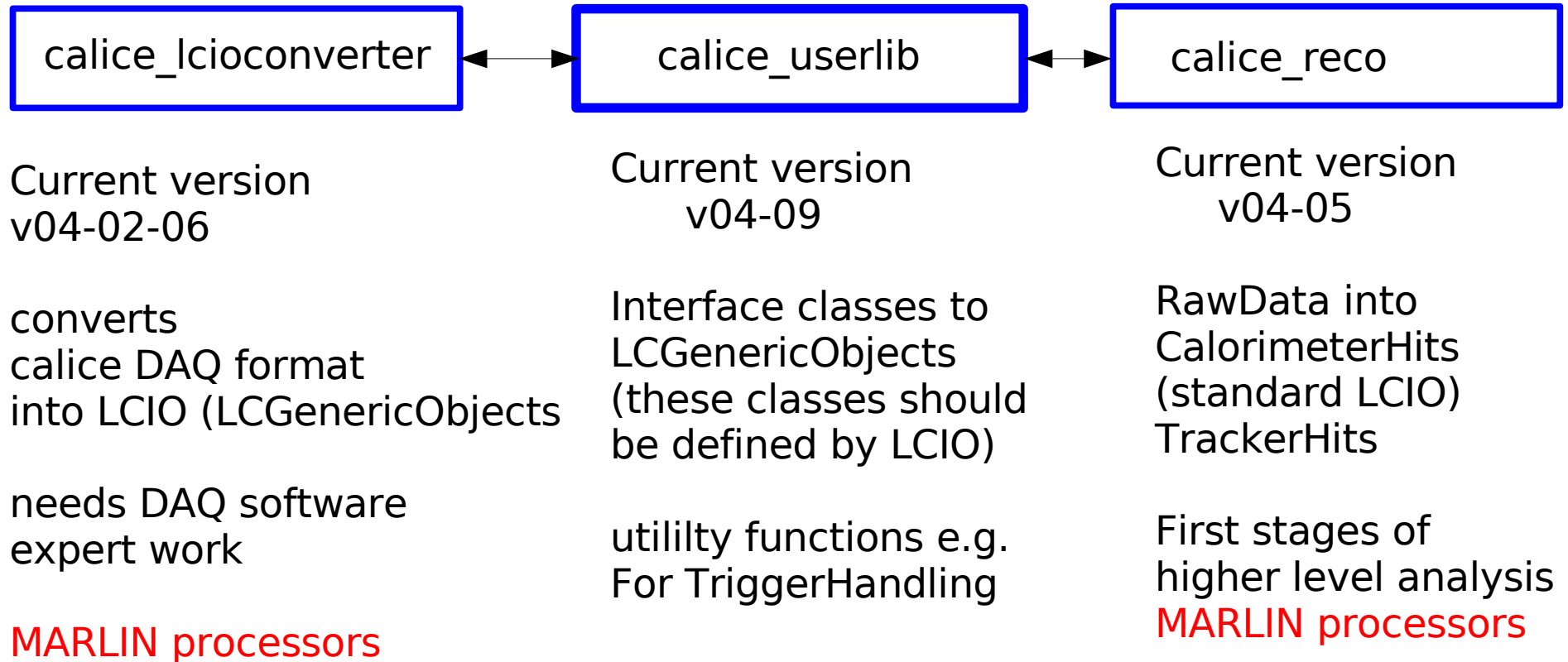
The importance of conditions data (not only) for 'real' data renders the development of a fully functional cd data toolkit to be a fundamental !!! piece of the ILC Software

- Efficient storage and access to conditions data
Browsing, convenient interfaces
- How to 'distribute' conditions data (e.g w.r.t to grid) ?
BTW.: LHC does have some headache with that!

Calice Software

Three main packages

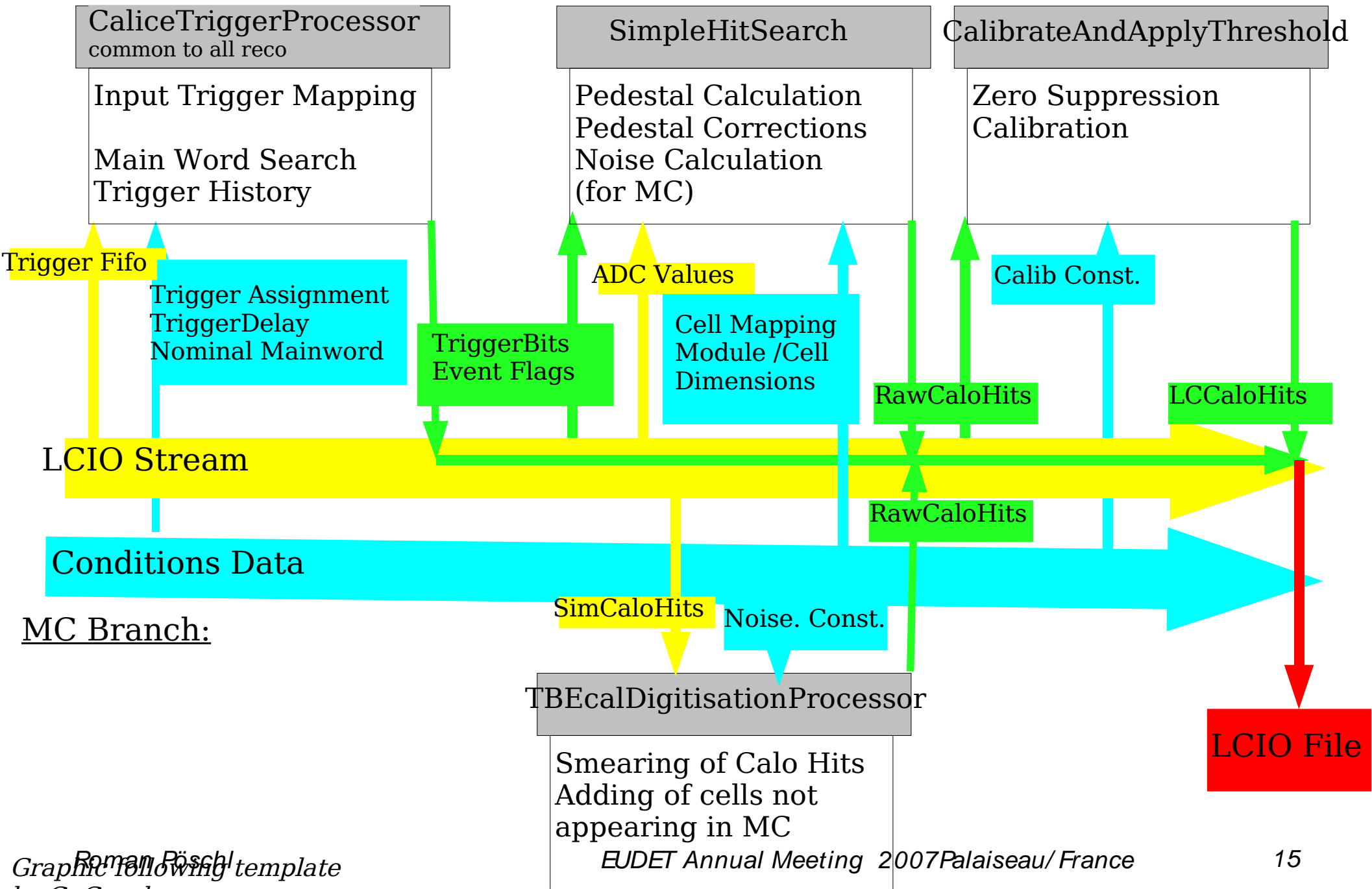
Contributions by groups from
DESY, Imperial, LAL, LLR, NIU, RHUL



225 classes or functions

**Data of four different Calorimeter Prototypes are
available in LCIO format**

Example for Data Processing - SiW Ecal



Real Data Branch:

CaliceTriggerProcessor
common to all reco

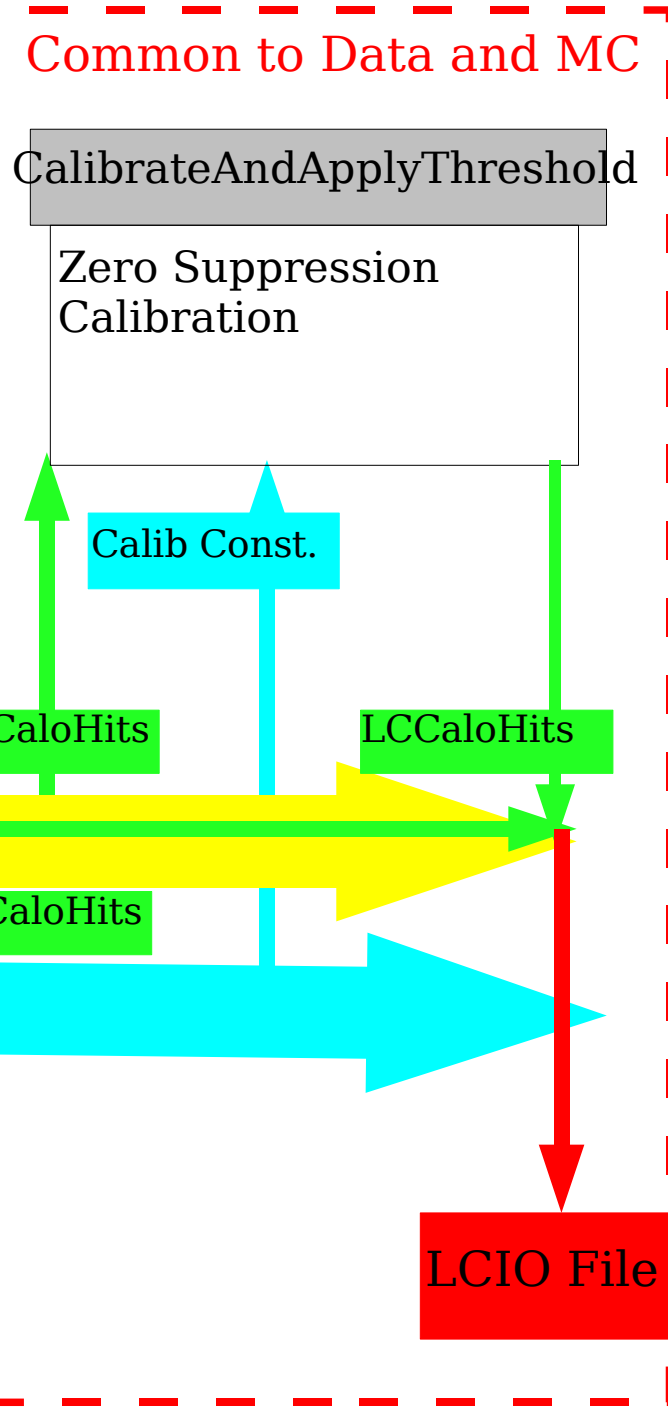
Input Trigger Mapping
Main Word Search
Trigger History

SimpleHitSearch

Pedestal Calculation
Pedestal Corrections
Noise Calculation
(for MC)

CalibrateAndApplyThreshold

Zero Suppression
Calibration



Trigger Fifo

Trigger Assignment
TriggerDelay
Nominal Mainword

ADC Values

Cell Mapping
Module /Cell
Dimensions

Calib Const.

TriggerBits
Event Flags

RawCaloHits

LCCaloHits

LCIO Stream

Conditions Data

RawCaloHits

MC Branch:

SimCaloHits

Noise. Const.

TBEcalDigitisationProcessor

Smearing of Calo Hits
Adding of cells not
appearing in MC

LCIO File

Reconstructed LCIO files are entry point for newcomers

... and starting point of high level analysis

Contain 'familiar' CalorimeterHits

Benefits:

Application of Standard ILC Software allows e.g. For analysis of Ecal data 7 different institutes without major startup problems

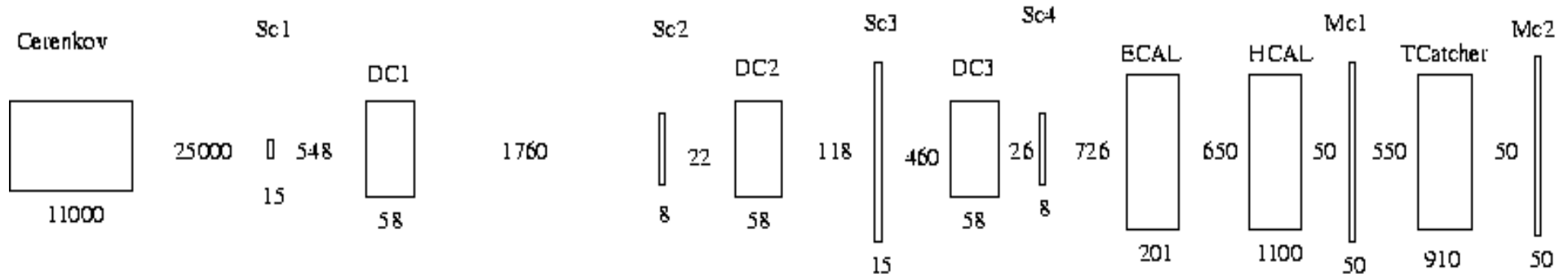
Transfer of knowledge into full detector simulations (and vice versa)
First attempts are ongoing

Though not the whole story – Still have to understand fundamentals of detectors -> required digging in raw data (less standard by definition)

A view to the Monte Carlo Branch

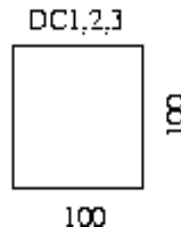
- Model for the simulation of the CERN (and DESY) test beam is available (in release 06-04-p03 of Mokka)

TOP



FRONT

Sc1 is 30x30
 Sc2 and Sc4 are 100x100
 Sc3 is 200x200



Mc1 and Mc2 are 1000x1000 All distances are in mm

DESY, RAL, LLR

Common effort of groups at RHUL, DESY, LLR, NIU

Will use grid for MC production

Estimation ~ 10000 kSI2kd for simulation of CERN 2006 data

Summary and Outlook

- Calice uses ILC Software for processing of Testbeam Data

ILC Datataking in a (big) nutshell

Very important input for current and future developments of ILC Software
Allows for a revision of the ILC Software concepts on a 'living' beast

- Calice uses systematically Grid tools

First (and only?) R&D project within ILC effort
24h/24h 7h/7h during CERN testbeam

A big thank you to all experts of IT divisions, particularly DESY, CERN and LAL, LLR, who support our effort!!!!

- Experience with testbeam data clearly reveals the needs for a coherent concept to handle 'low level' data within ILC Software

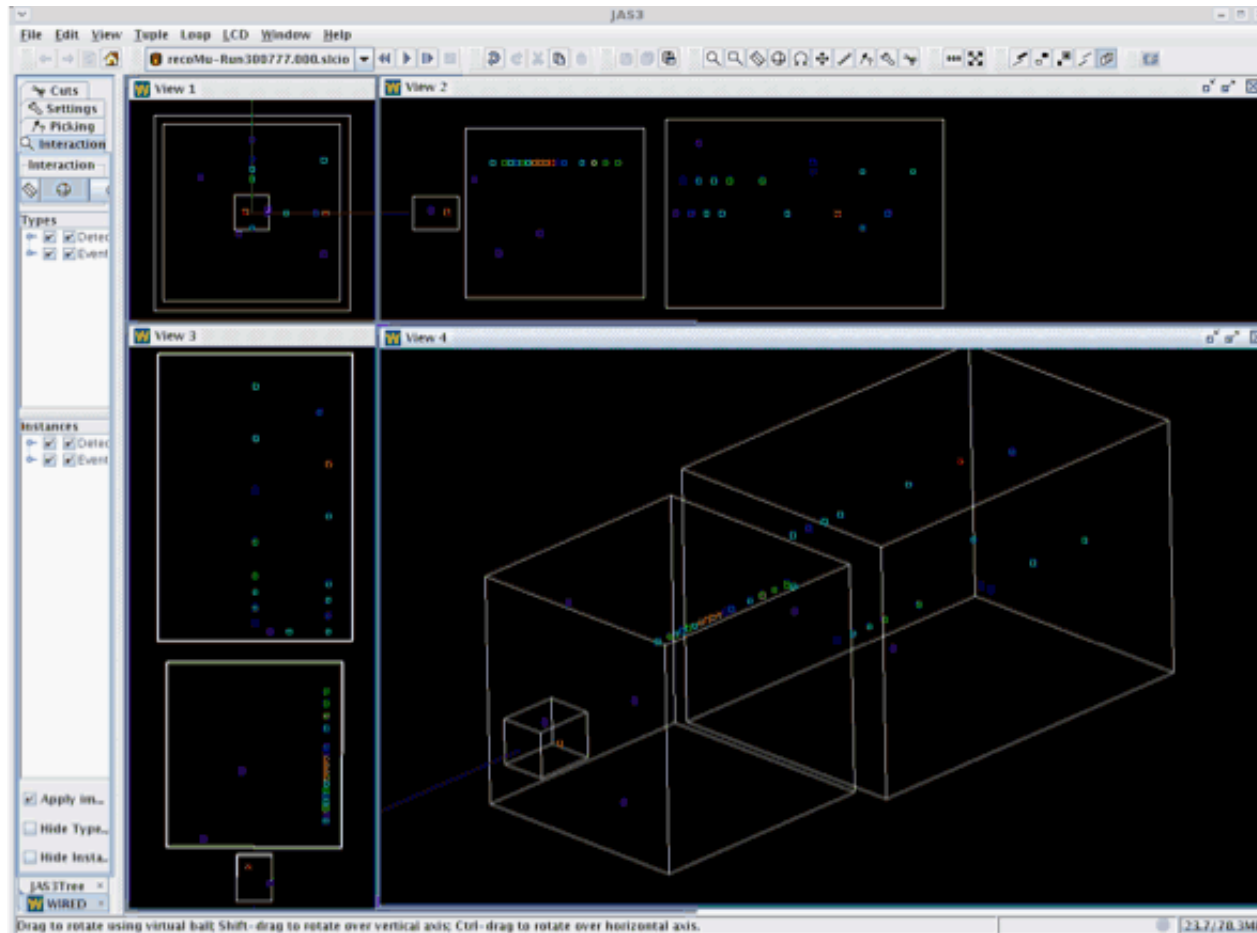
- (Latest) Next generation R&D projects should be used to develop a complete data processing/handling strategy for the ILC.
Avoid 'island' solutions and work on an integrated effort

CALICE does not only hardware-prototyping but also 'computing prototyping'

Computing benefits from collaborative effort and application of ILC software tools

On Event Displays

jas3/wired for based on HEPREP file from simulation

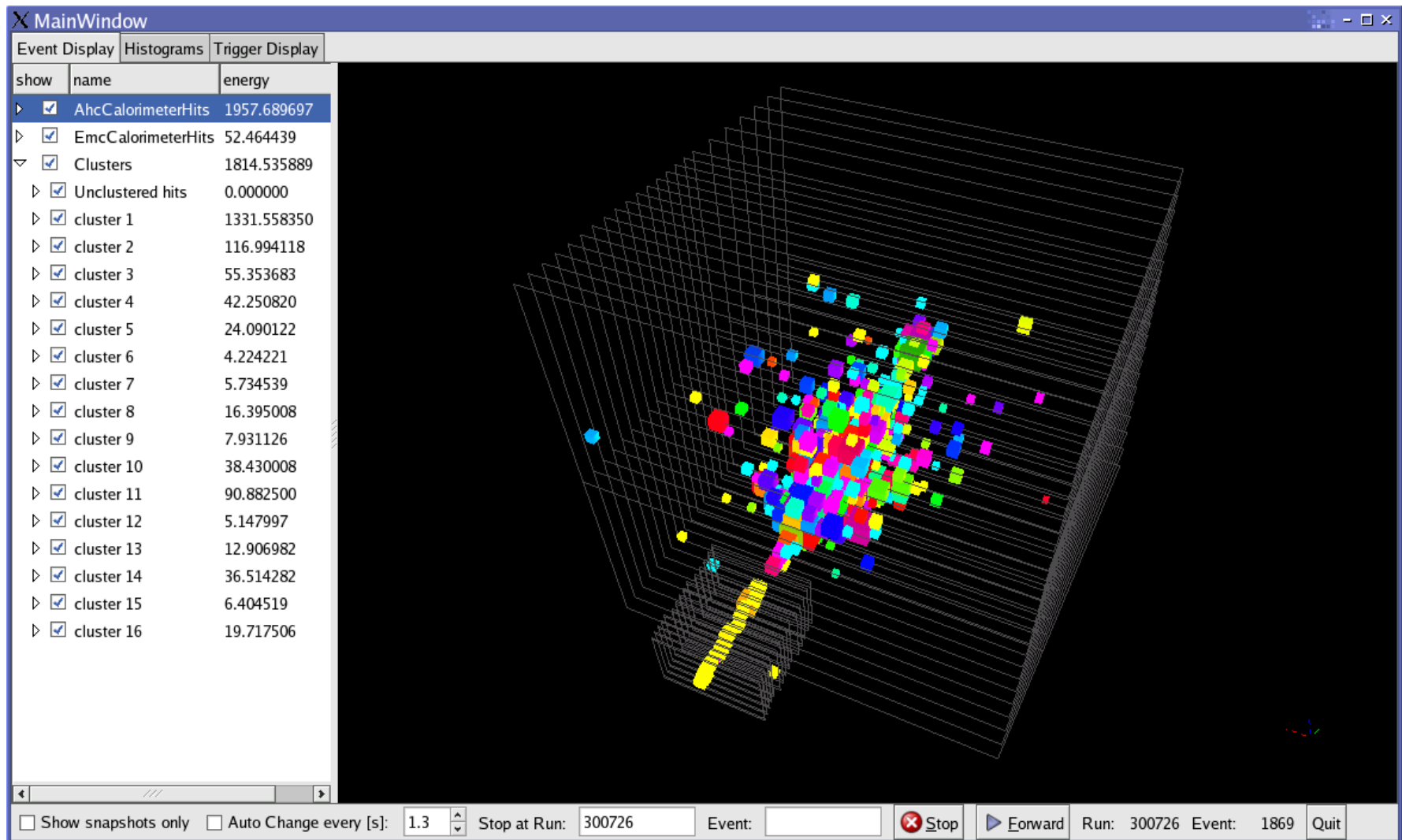


Pro: Easy to install and use, LCIO integrated

(Small) Con: Better would be to take geometry from calice db

Event Display based on OpenGL

G. Gaycken, S. Schmidt



Pro: Geometry based on info from calice db
Con: Heavy graphics package

AHCAL Reconstruction Framework – S.Schmidt/ DESY

- MappingI
 - ADCBlocks → CaliceHits1
- PedestalCalibration
 - CaliceHits1 → CaliceHits2
- GainCalibration
 - CaliceHits2 → CaliceHits3
- InterCalibration
 - CaliceHits3 → CaliceHits4
- SaturationCorrection
 - CaliceHits2, CaliceHits4 → CaliceHits5
- MIPCalibration
 - CaliceHits5 → CaliceHits6
- MappingII
 - CaliceHits6 → CalorimeterHits

Calibration steps modularized
in MARLIN processors

(more details → Calo Session)

