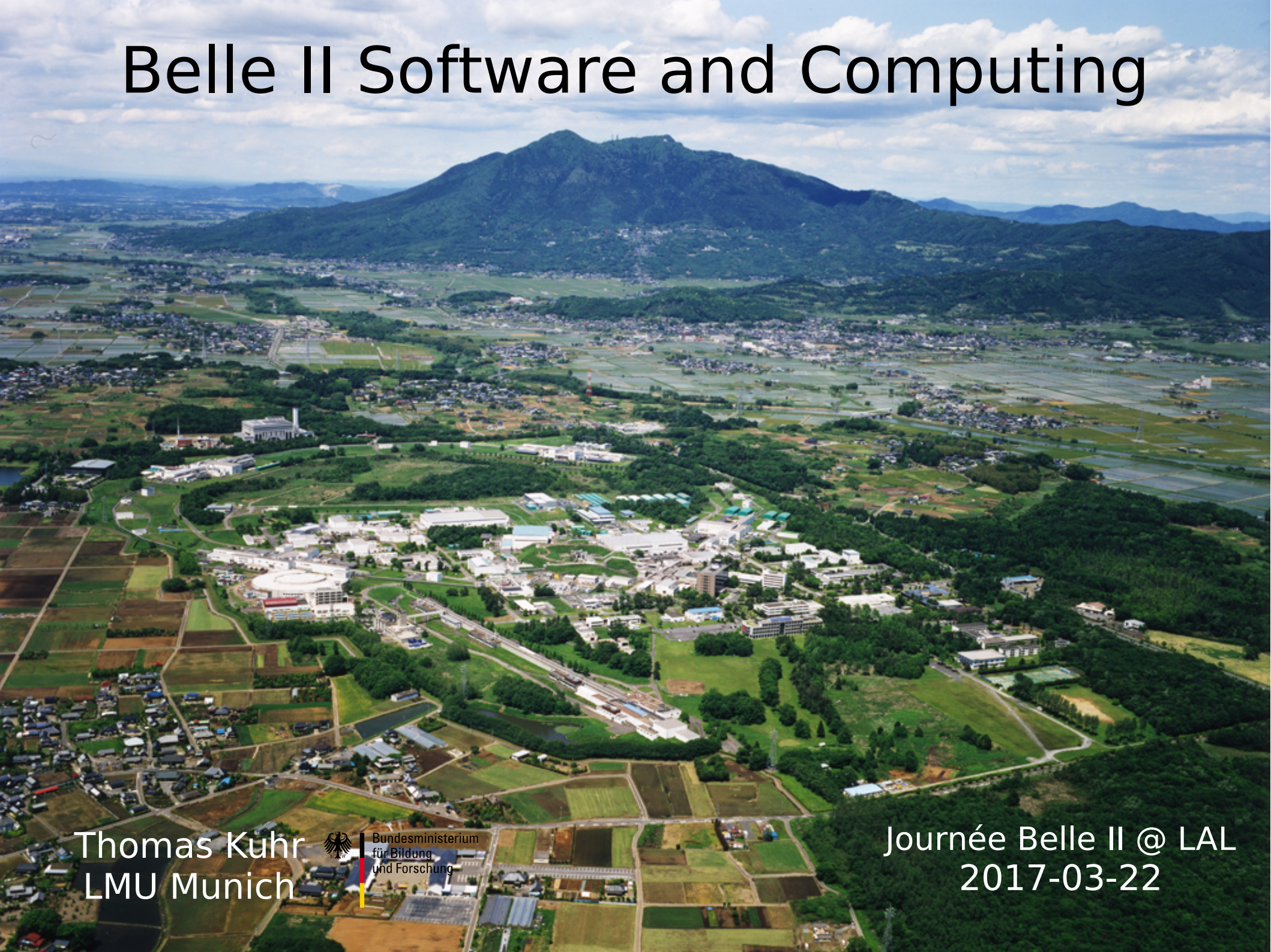
Belle II Software and Computing



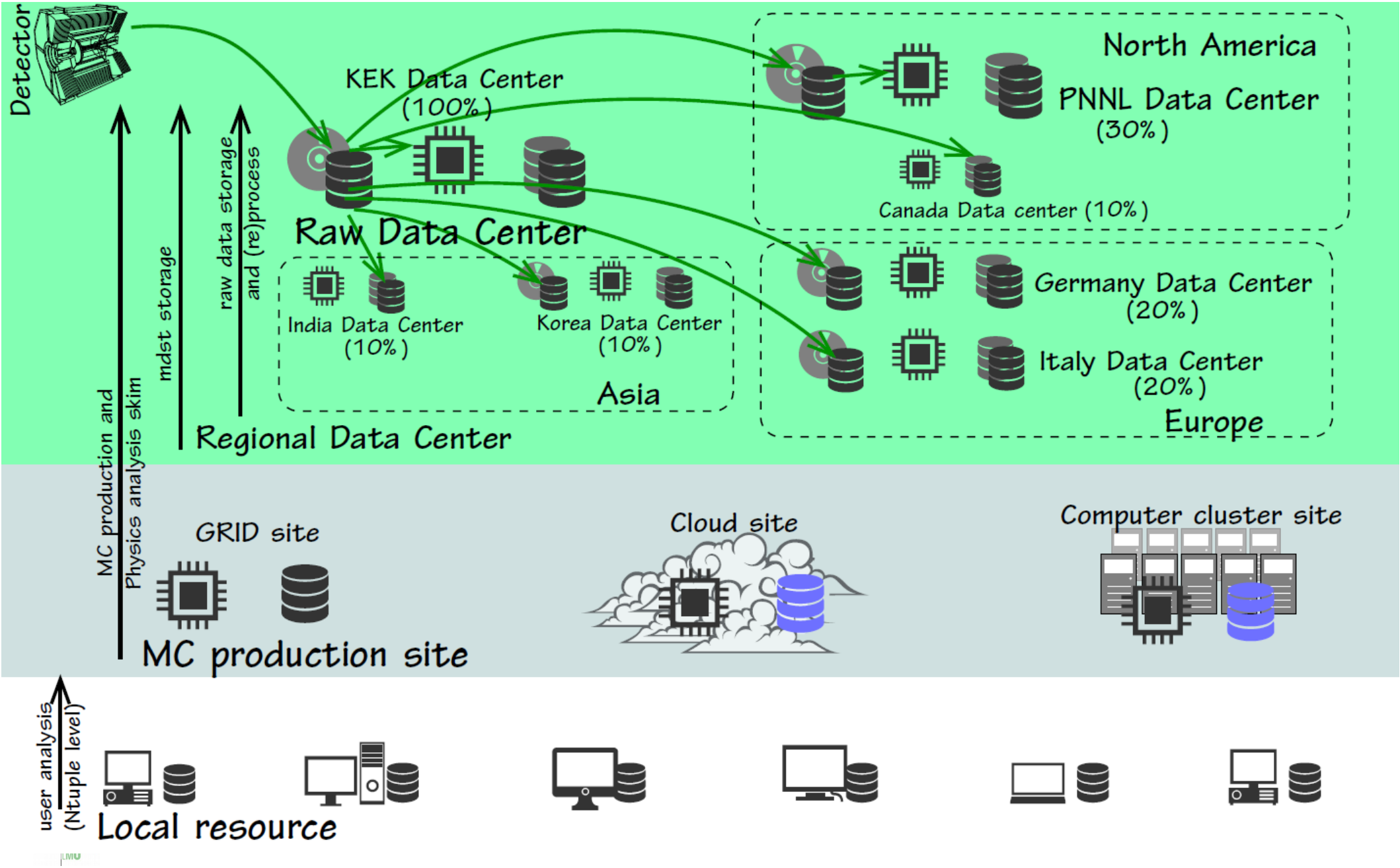
Thomas Kuhr
LMU Munich

Bundesministerium für Bildung und Forschung

Journée Belle II @ LAL
2017-03-22

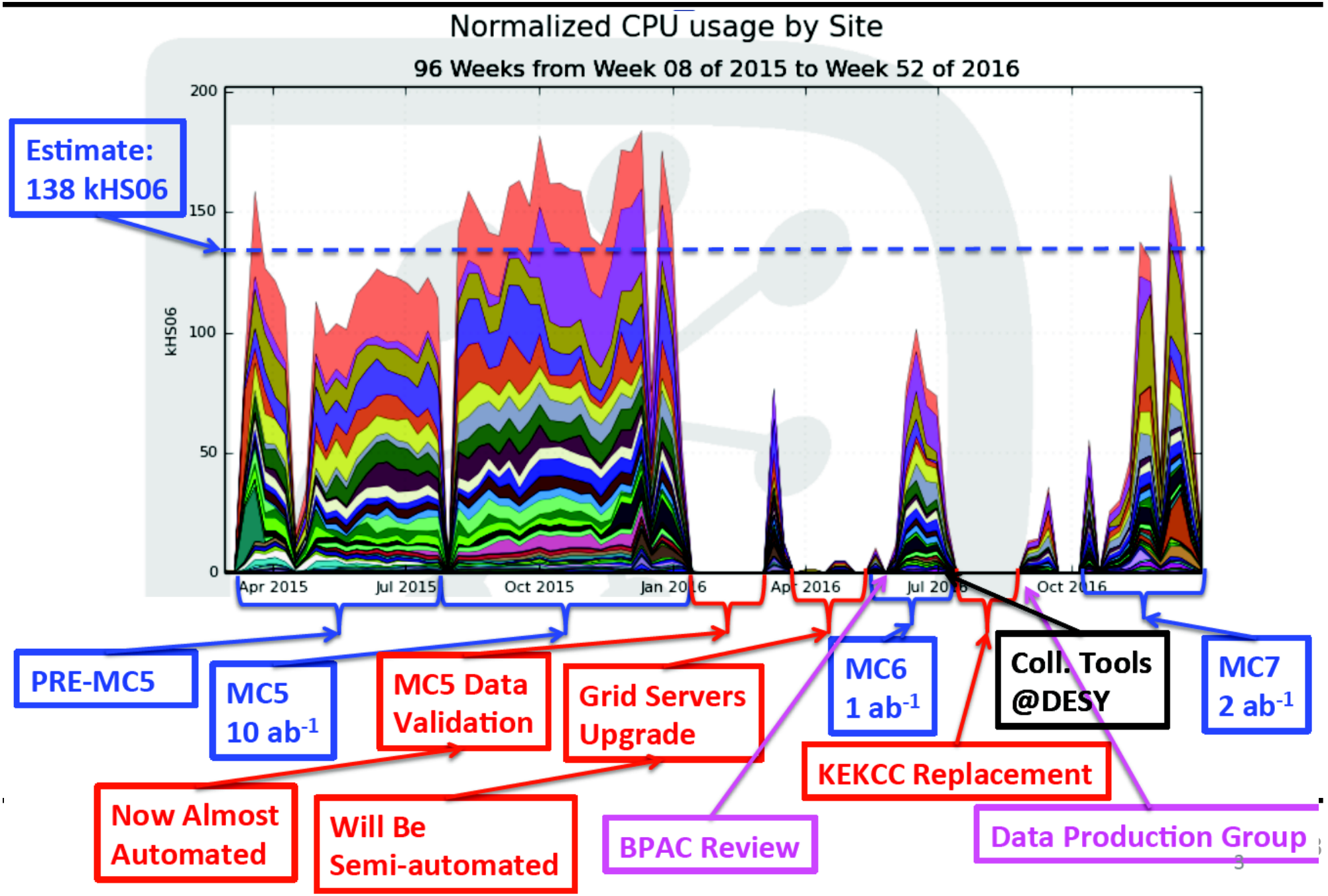
Computing Model

>10 PB of raw data per year
→ Distributed computing model
Share per fraction of PhDs



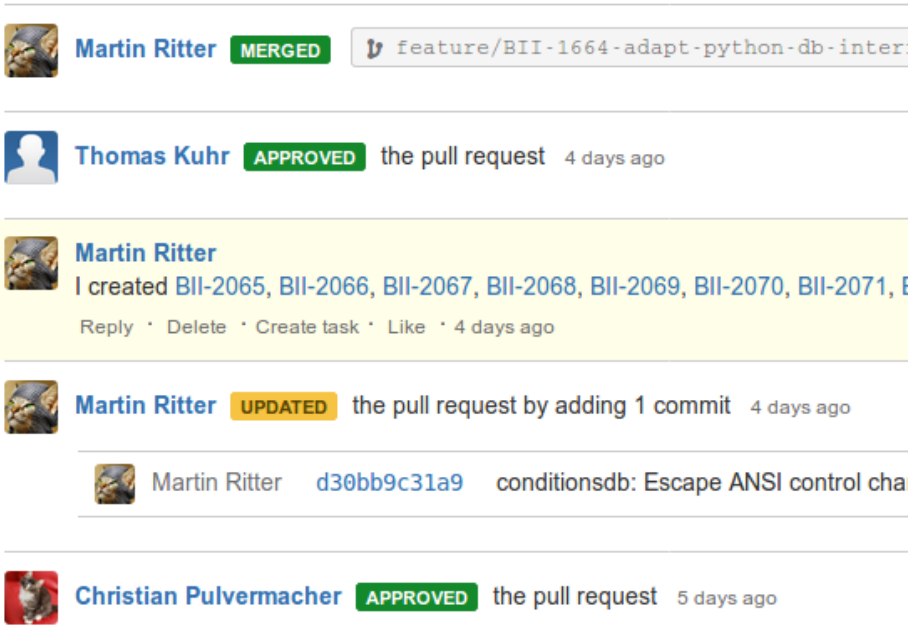
MC Campaigns

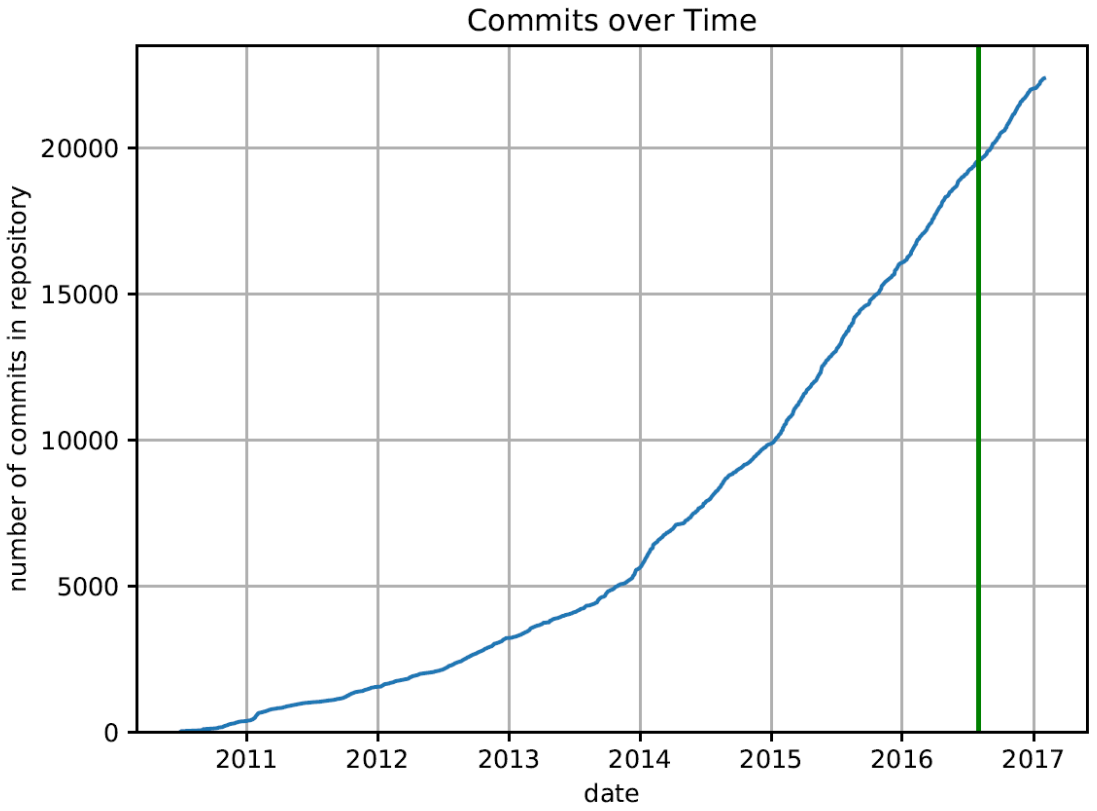
Computing system based on DIRAC
→ Many possibilities to contribute
Meeting with Takanori Hara tomorrow



Migration svn → git

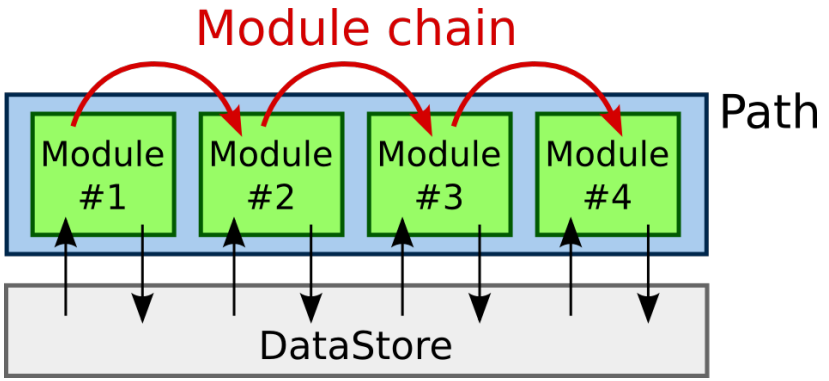
- Belle II decided last year to migrate collaborative services from KEK to DESY
- We used that opportunity to switch from svn to git
  - Adjustment of procedures and tools required



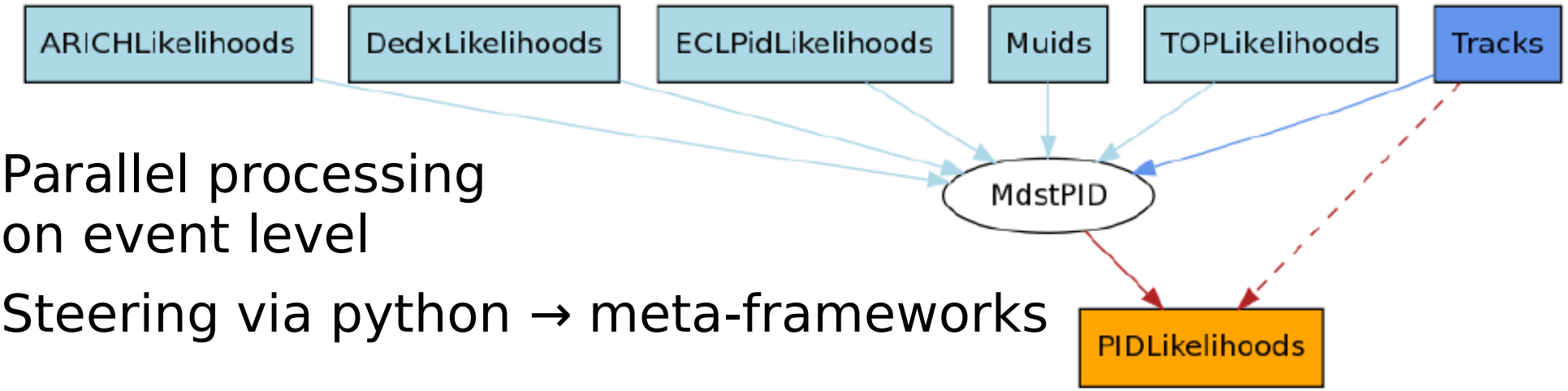


Software Framework

- Dynamic loading of modules
- Data exchange via DataStore
- Relations
- Conditions data interface
- Root I/O
- Belle data input (b2bii)

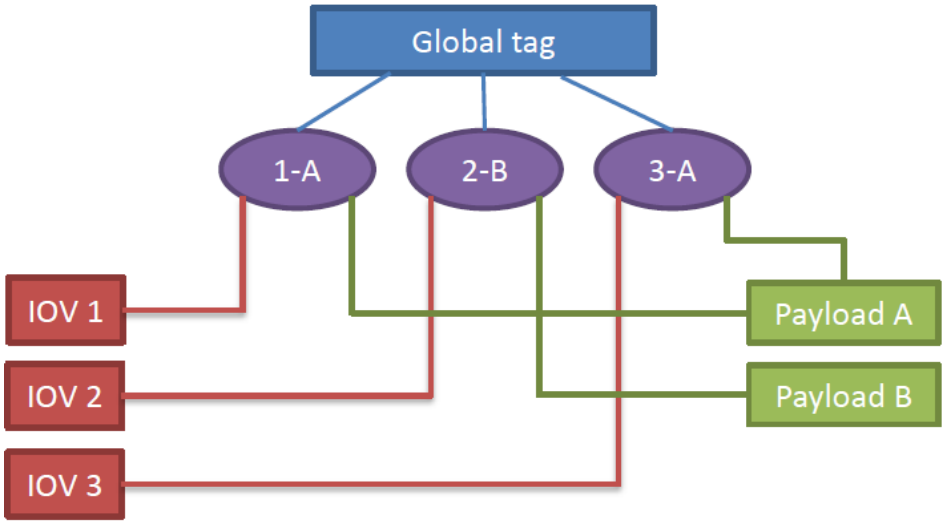


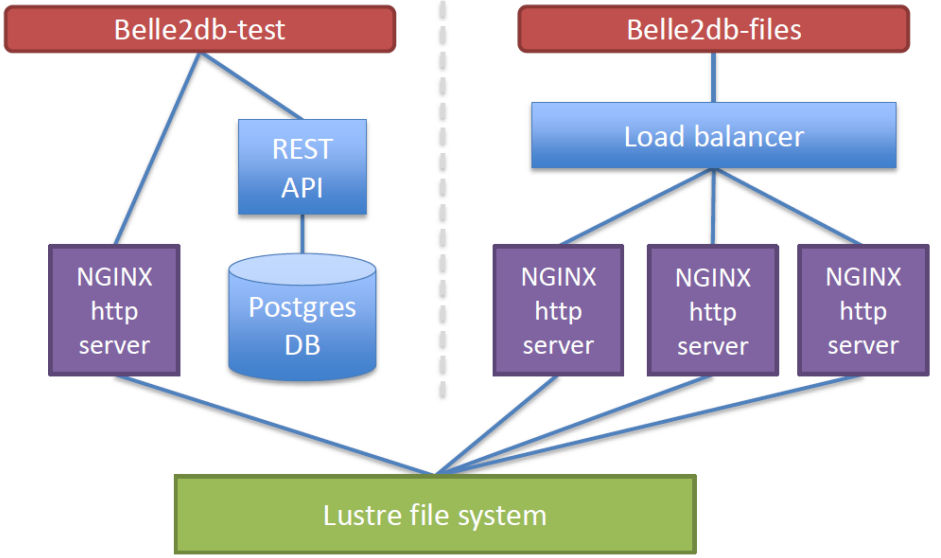

```
StoreArray<Track> tracks;
for (const Track& track: tracks) {
  const PIDLikelihood* pid =
    track->getRelated<PIDLikelihood>();
}
```



- Parallel processing on event level
- Steering via python → meta-frameworks

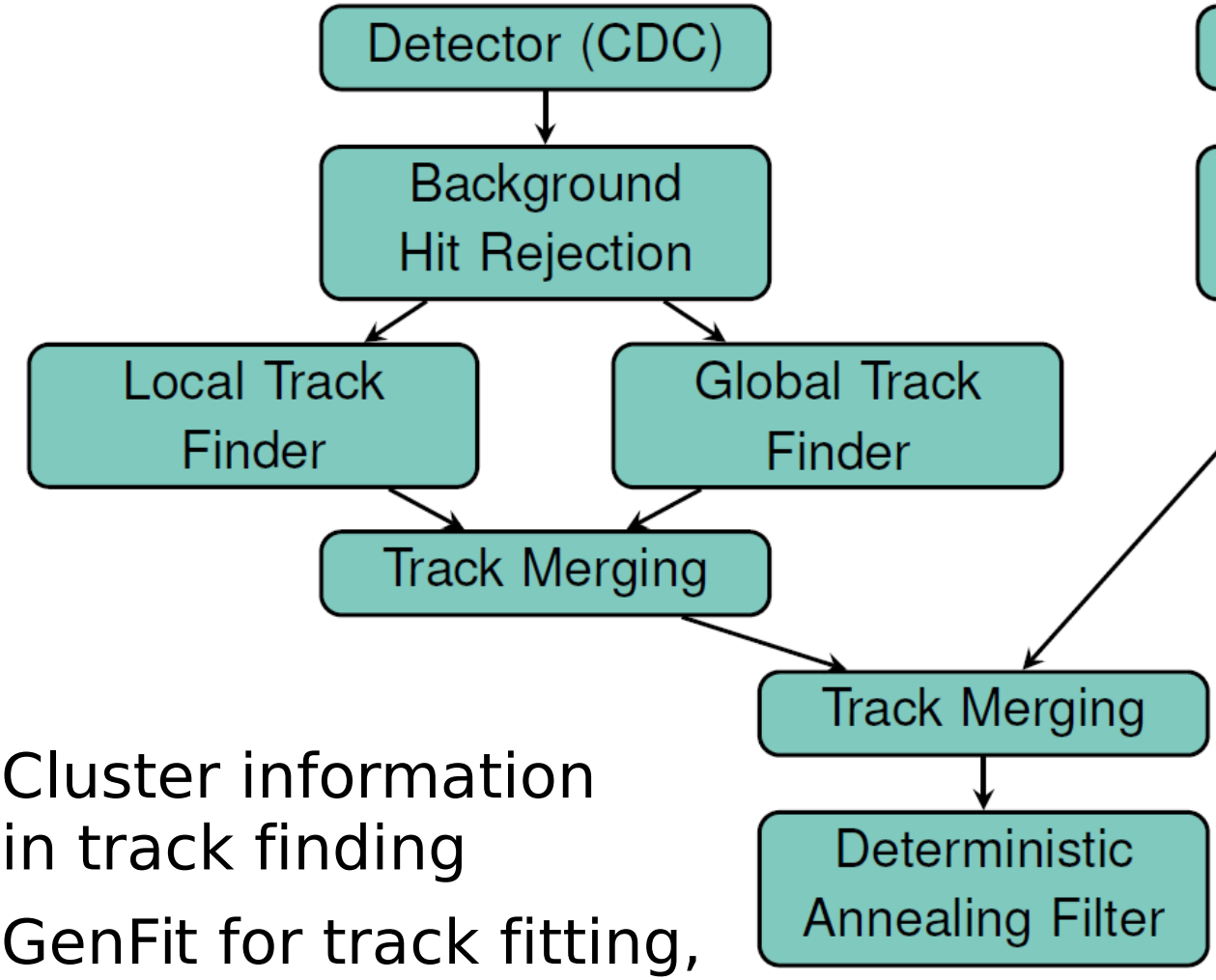
Conditions Database



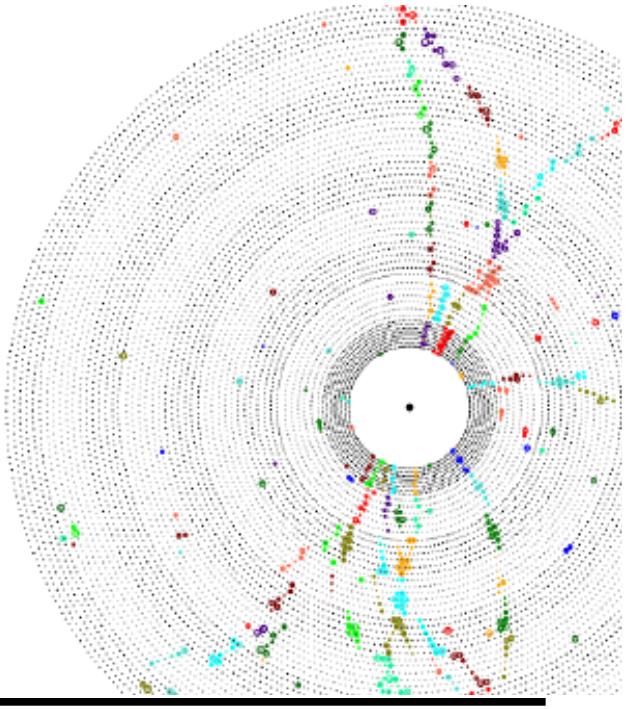


- Conditions data stored in objects in root files (payloads)
  → Provided via CVMFS or downloaded from server
- Global tag: Assignments of intervals of validity (IoV) to payloads → Database
- Distribution, redundancy
- Online integration

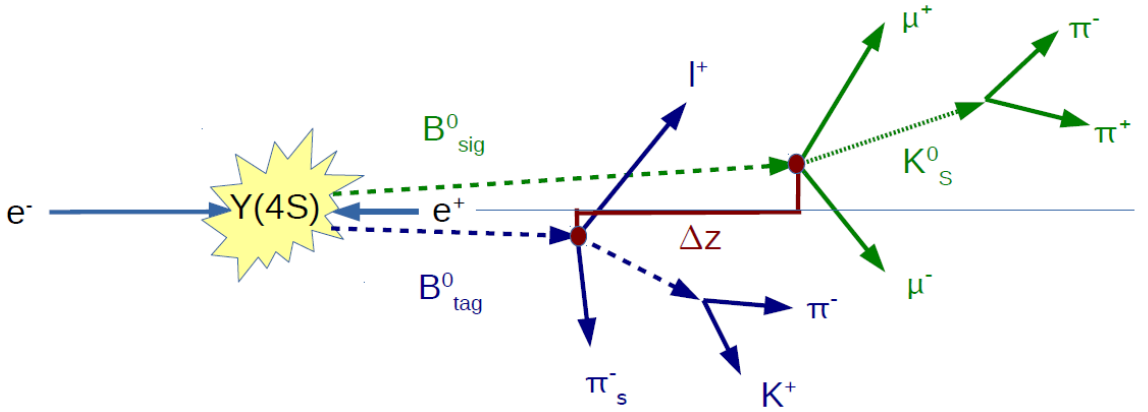
Tracking



- Cluster information in track finding
- GenFit for track fitting, ACTS as alternative?



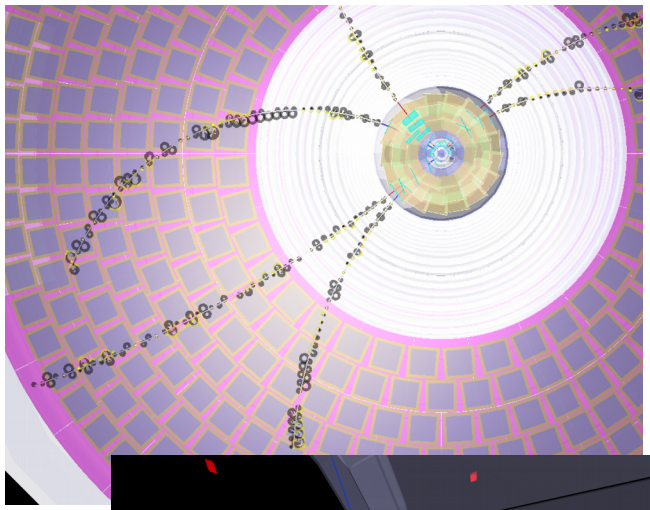
Modular Analysis

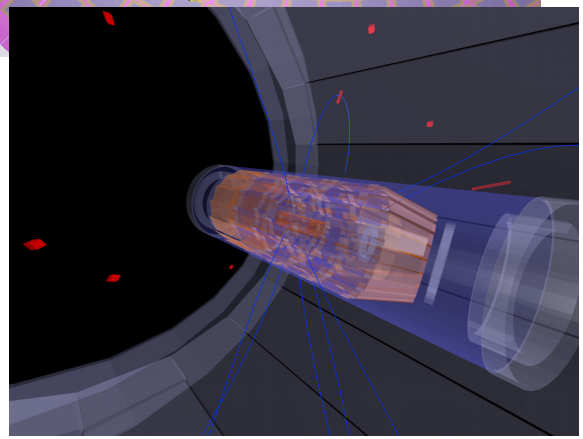


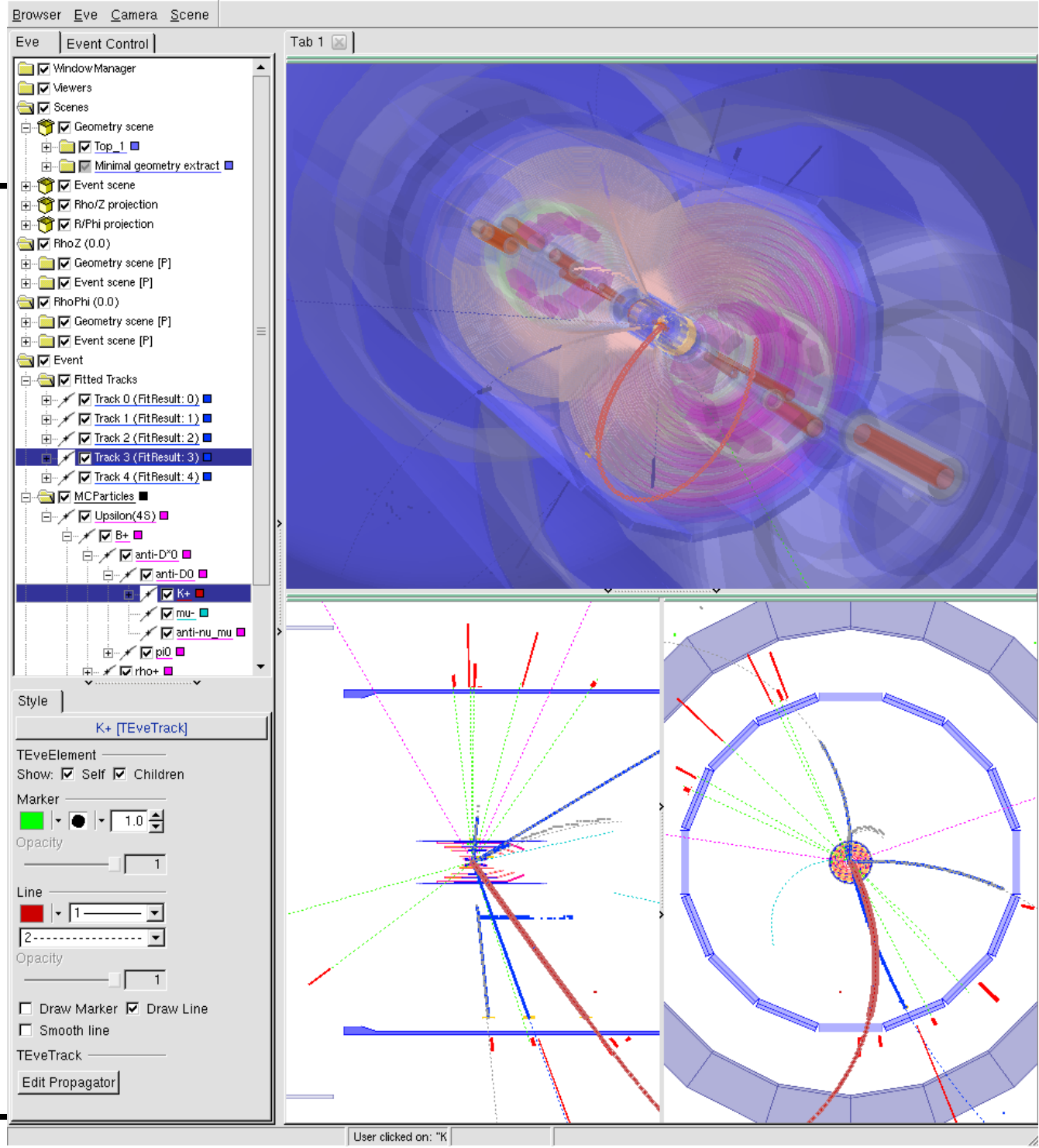

```
inputMdst(...)

# create "mu+:loose" ParticleList (and c.c.)
stdLooseMu()

# create Ks -> pi+ pi- list from V0
# keep only candidates with 0.4 < M(pipi) < 0.6 GeV
fillParticleList('K_S0:pipi', '0.4 < M < 0.6')

# reconstruct J/psi -> mu+ mu- decay
# keep only candidates with 3.0 < M(mumu) < 3.2 GeV
reconstructDecay('J/psi:mumu -> mu+:loose mu-:loose', '3.0 < M < 3.2')

# reconstruct B0 -> J/psi Ks decay
# keep only candidates with 5.2 < M(J/PsiKs) < 5.4 GeV
reconstructDecay('B0:jspiks -> J/psi:mumu K_S0:pipi', '5.2 < M < 5.4')

# perform B0 kinematic vertex fit using only the mu+ mu-
# keep candidates only passing C.L. value of the fit > 0.0 (no cut)
vertexRave('B0:jspiks', 0.0, 'B0 -> [J/psi -> ^mu+ ^mu-] K_S0')

# build the rest of the event associated to the B0
buildRestOfEvent('B0:jspiks')

# perform MC matching (MC truth association)
matchMCTruth('B0:jspiks')

# calculate the Tag Vertex and Delta t (in ps)
# breco: type of MC association.
TagV('B0:jspiks', 'breco')

# create and fill flat Ntuple with MCTruth, kinematic information and D0 FlightInfo
toolsDST = ['EventMetaData', '^B0']
toolsDST += ['MCTruth', '^B0 -> [^J/psi -> ^mu+ ^mu-] [^K_S0 -> ^pi+ ^pi-]']
toolsDST += ['Vertex', '^B0 -> [^J/psi -> mu+ mu-] [^K_S0 -> pi+ pi-]']
toolsDST += ['DeltaT', '^B0']
toolsDST += ['MCDeltaT', '^B0']

# write out the flat ntuples
ntupleFile('B2A410-TagVertex.root')
ntupleTree('B0tree', 'B0:jspiks', toolsDST)
```

- Analysis on steering file level using decay strings
- ✔ Particle reconstruction and selection
- ✔ MC matching
- ✔ Vertex fits
- ✔ Flavor tagging
- ✔ Continuum suppression

Event Display







- Virtual reality: https://vimeo.com/185549878

Roadmap

Version	Status	Progress	Start date	Release date	Description
release-00-09-00	UNRELEASED		01/Dec/16	01/May/17	Ready for Cosmic Data: raw data unpackers, cosmics simulation and reconstruction
release-01-00-00	UNRELEASED			01/Nov/17	Ready for BEAST Phase 2: alignment and calibration of outer detectors, generators for Upsilon(nS) verified, all software components ready for physics analysis
release-01-01-00	UNRELEASED			01/May/18	Ready for VXD cosmics
release-02-00-00	UNRELEASED			01/Oct/18	Ready for phase 3 data taking: full data processing chain working, proecdures for systematic error determinations defined
release-02-01-00	UNRELEASED	No issues		01/May/19	Tuning to real data
release-03-00-00	UNRELEASED	No issues		01/Jan/21	High luminosity phase: optimized simulation, reconstruction, and alignment/calibration procedures

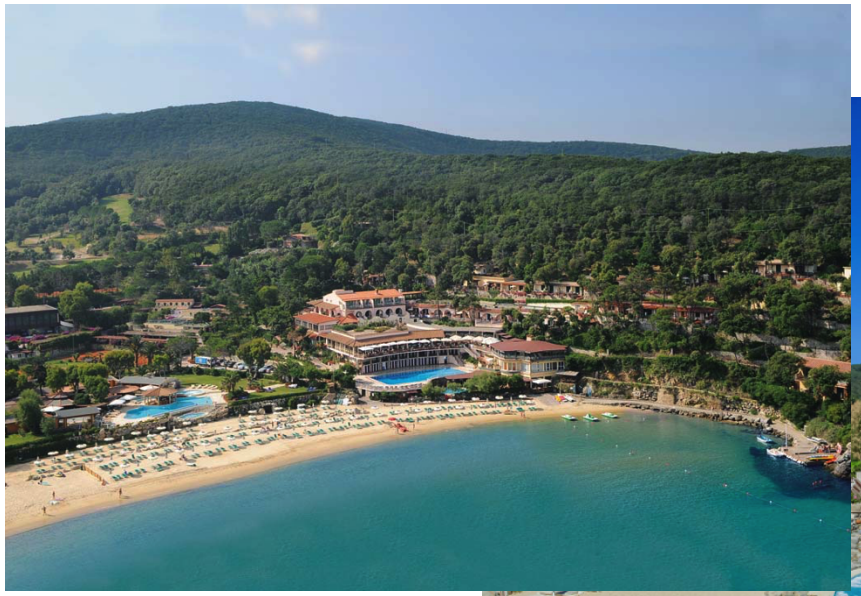
- Many exciting opportunities to contribute now!

Institutional Commitments

Institution	Commitment	Last Update	Timescale
Pisa	Development of the VXD Track Finder 2. In particular: • Development of the sector map, segment network and cellular automaton parts • Tuning, deployment and maintenance • Monitoring of the VXD TF2 performances during the first years of the data taking • Development of the algorithms to add PXD and SVD points to CDC-only tracks	2017-02	
CMU	• CDC dE/dx reconstruction and calibration • CDC dE/dx track-level simulation • charm related generators • reconstruction package librarian • data production: communication with physics, computing, detector, and daq/trigger groups • data production: generation of release validation samples	2016-11	data production tasks associated with coordinator role
KEK	• Background simulation (interface to machine) • Software related to the infrastructure of detector(TRG, DAQ)	2016-10	
Ljubljana	• ARICH reconstruction software and data base • TOP reconstruction software • database • background mixing • analysis package librarian • Belle to Belle II Conversion task force	2016-10	
MPP Munich	• vertexing • flavor tagging	2016-10	
Jülich	• genfit2	2016-10	3 years
Virginia Tech	• bklm package • ext • muid • KLM trigger simulation • KLM-related parts of b2bii	2016-10	
British Columbia	• ECL clustering	2016-09	2 years
LMU	• software tools • external software • build system • mdst package • geometry package • database interface in basf2 • index files • software installations on cvmfs • buildbot infrastructure, including automatic (monthly) builds, checks, tests, documentation generation • commit hooks • shift tools	2016-09	2 years
PNNL	• database back-end • TOP analysis software	2017-02	

Workshop @ Elba May 15-19 2017

- https://agenda.infn.it/conferenceDisplay.py?confId=1262



Backup

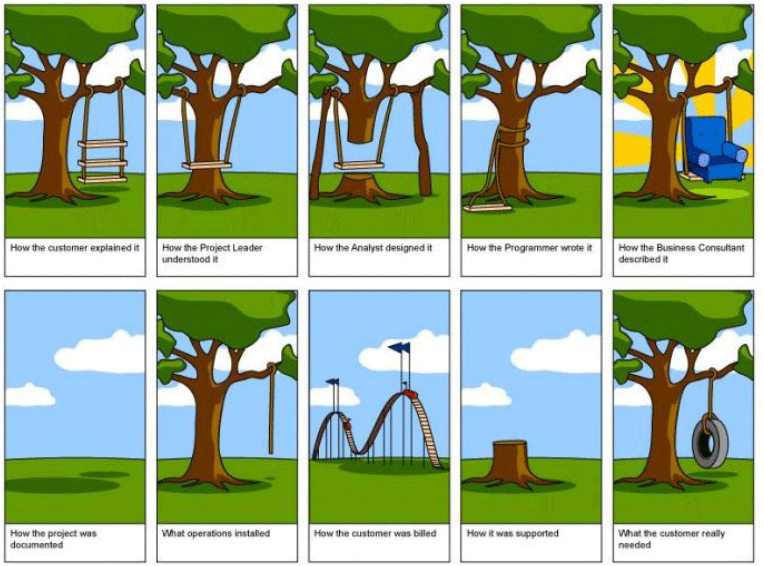
Software Development at Belle II

Aim:

- Reliable, sophisticated, and easy-to-use software for acquisition, simulation, reconstruction, and analysis of Belle II data

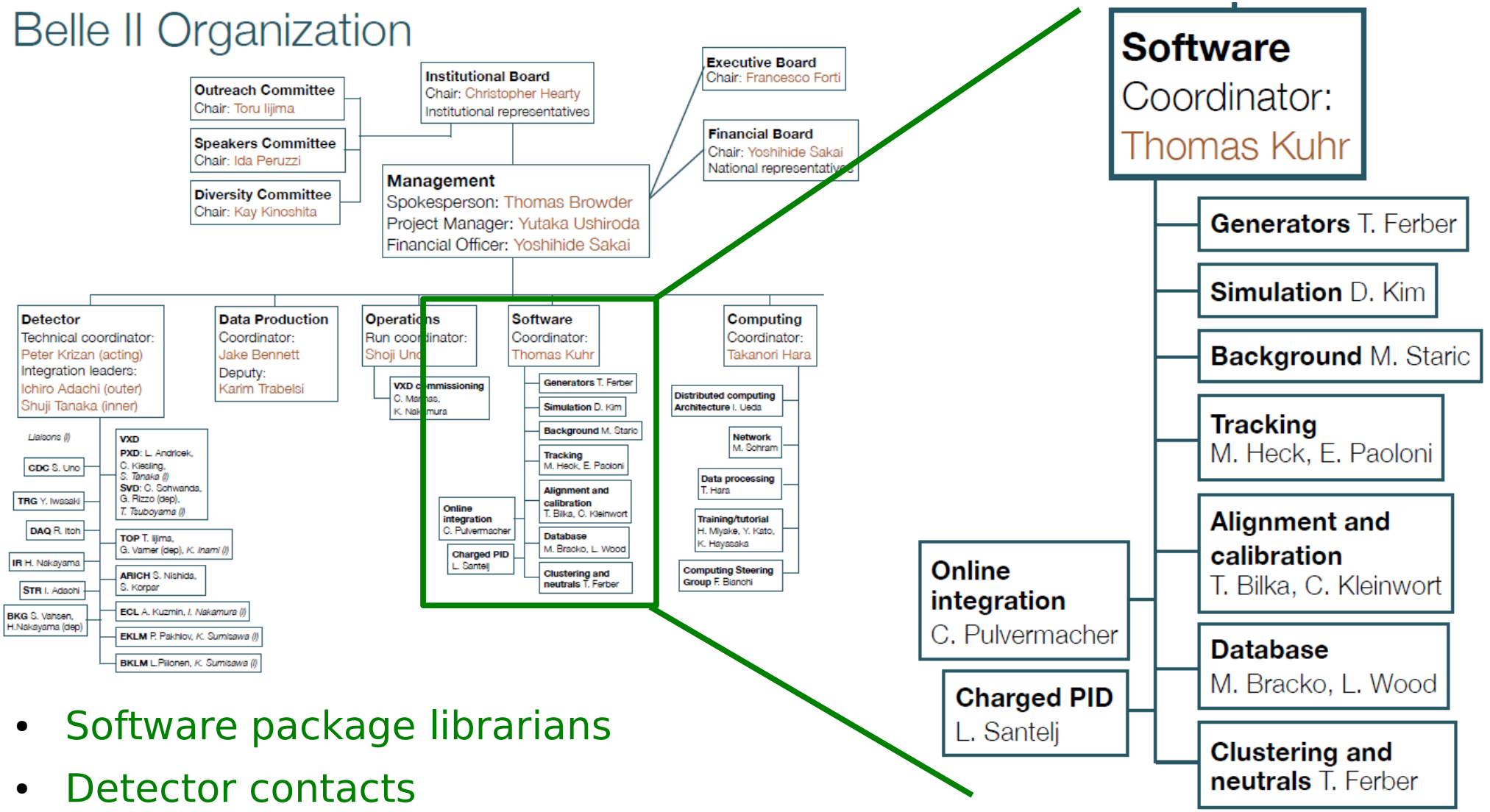
Challenge:

- Regional distribution, different (cultural) backgrounds and skills of developers



- ✔ State-of-the-art tools
- ✔ Commonly accepted rules and guidelines
- ✔ Well defined procedures
- ✔ Efficient communication channels

Software Group Structure



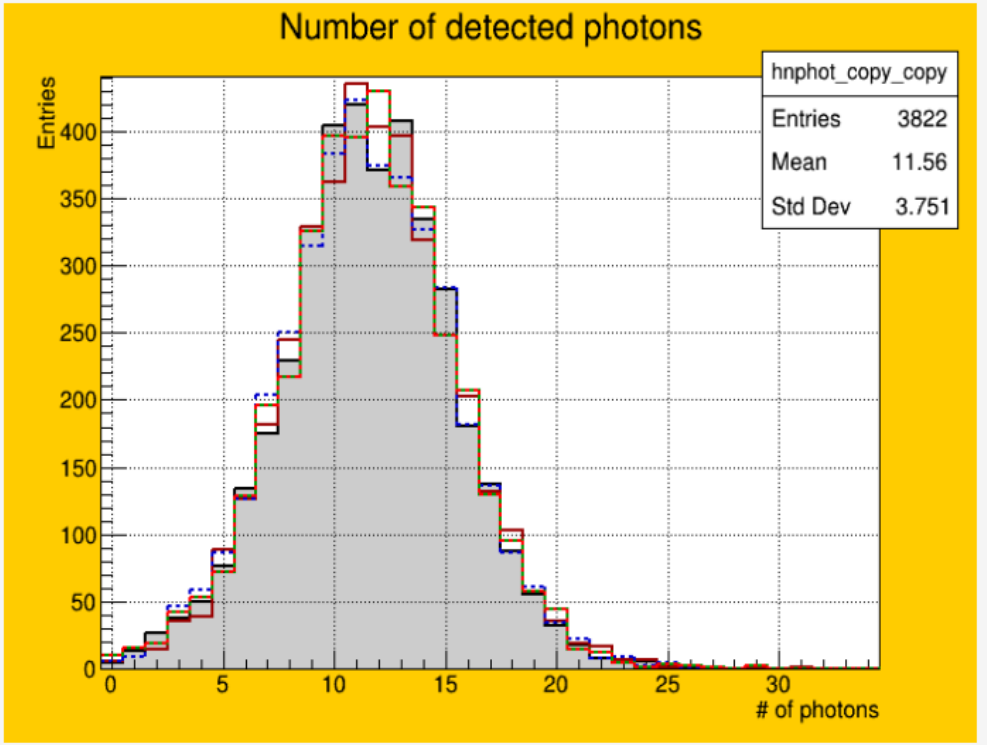
- Software package librarians
- Detector contacts

Software Quality Control

Automated checks:

- code style
- gcc/clang/icc
- cppcheck,
  clang static analyzer
- unit/execution tests
- Doxygen
- geometry overlaps
- valgrind memcheck
- execution time and
  output size monitoring
- high level validation plots
  using simulated samples

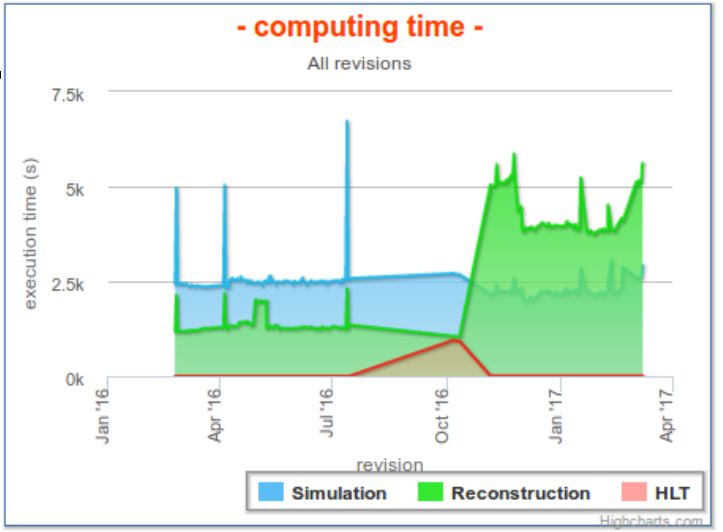
Plot: ARICHValidate_hnphot



Package: arich
Description: Number of detected photons in a 3#sigma band around the expected Cherenkov angle, for #pi with momenta 3.0-3.5 GeV.
ROOT file: ARICHValidate.root
Chi^2 Test: Performed Chi^2-Test between reference and 29456 (Chi^2 = 32.79074530336273 NDF = 28 Chi^2/NDF = 1.1710980465486691)
Comparison p-value: 0.2435353507685487 (Warning below 1 | Error below 0.01)
Contact: luka.santelj@kek.jp
Check for: Poissonian peak with average at ~11.5.

Software Quality Shifts

- Monitoring of quality indicators
- Interaction with developers and librarians
- Detailed instructions on confluence
- One week duty
- Reports in weekly developers meeting

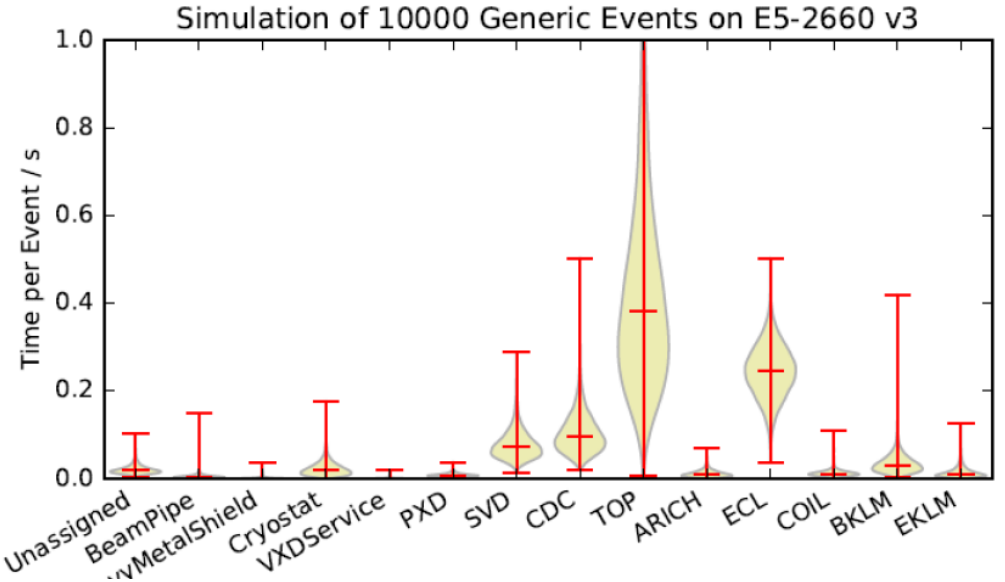


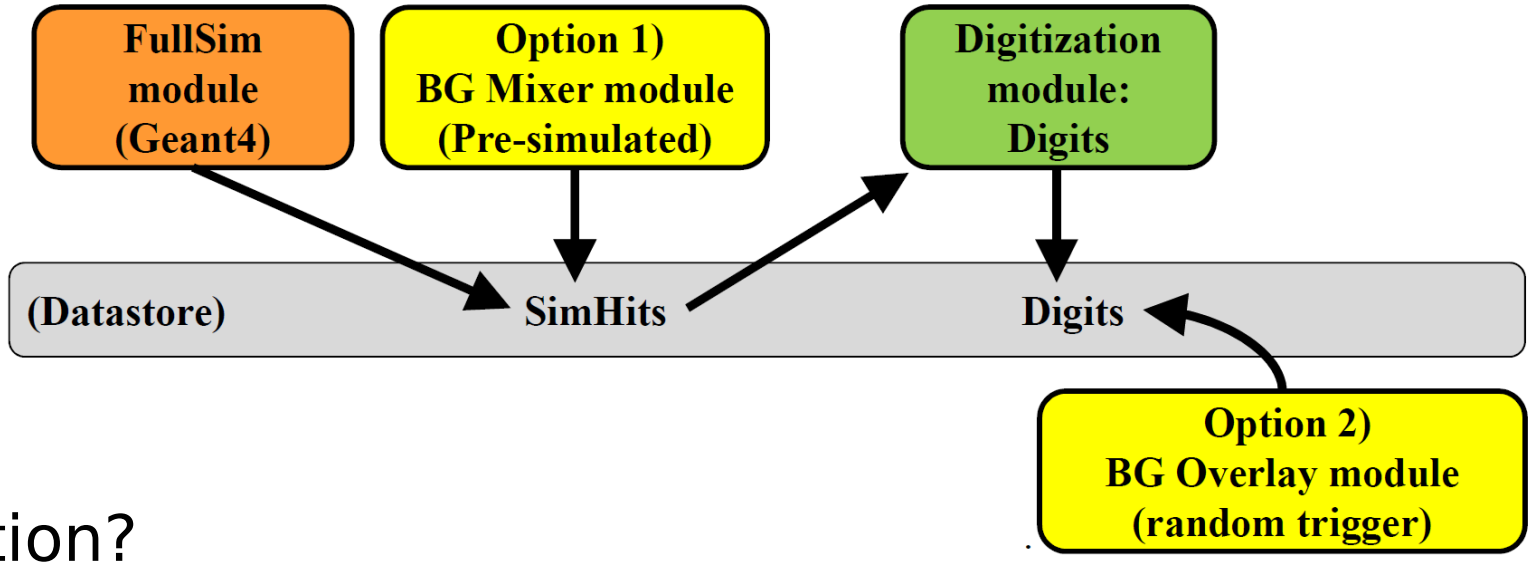
Package details

Package	Librarian	Build Result	Intel Build Result	Clang Build Result	Cppcheck	Test Result	Geometry	Code Documentation	Dependencies
alignment	Tadeas Bilka	OK	OK	OK	Remarks: 1	0/6, 0/0	OK	Missing: 1	Extra: 2 Missing: 2 OK
analysis	Anze Zupanc	OK	OK	OK	Remarks: 50	0/89, 0/11	OK	Missing: 1084	OK OK
arich	Luka Santelj	Warnings: 1	OK	OK	Remarks: 1	0/0, 0/1	Overlaps: 255	Missing: 55	OK OK
b2bii	Anze Zupanc	OK	OK	OK	OK	0/0, 0/1	OK	Missing: 306	OK OK
background	Marko Staric	OK	OK	OK	OK	None	OK	OK	OK OK
beast	Igal Jaegle	Warnings: 3	OK	Warnings: 7	Remarks: 9	None	OK	Missing: 404	OK OK
bklm	Leo Piilonen	OK	OK	OK	OK	0/0, 0/1	OK	OK	OK OK
calibration	Tadeas Bilka	OK	OK	OK	OK	0/0, 0/3	OK	Missing: 12	OK OK
cdc	Nakano Eiichi	OK	OK	OK	OK	0/0, 0/1	OK	OK	OK OK

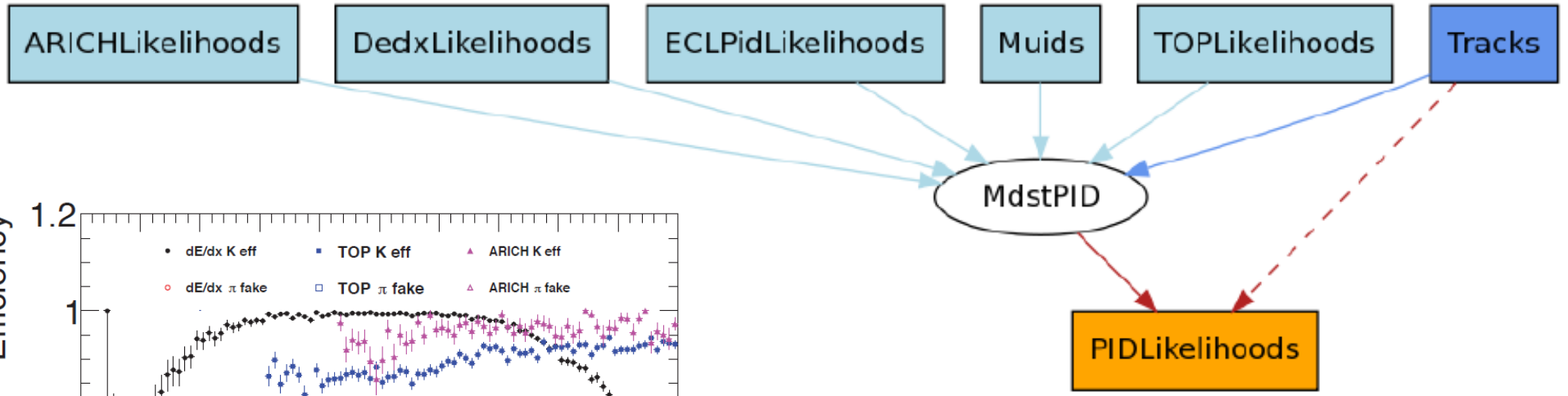
Simulation

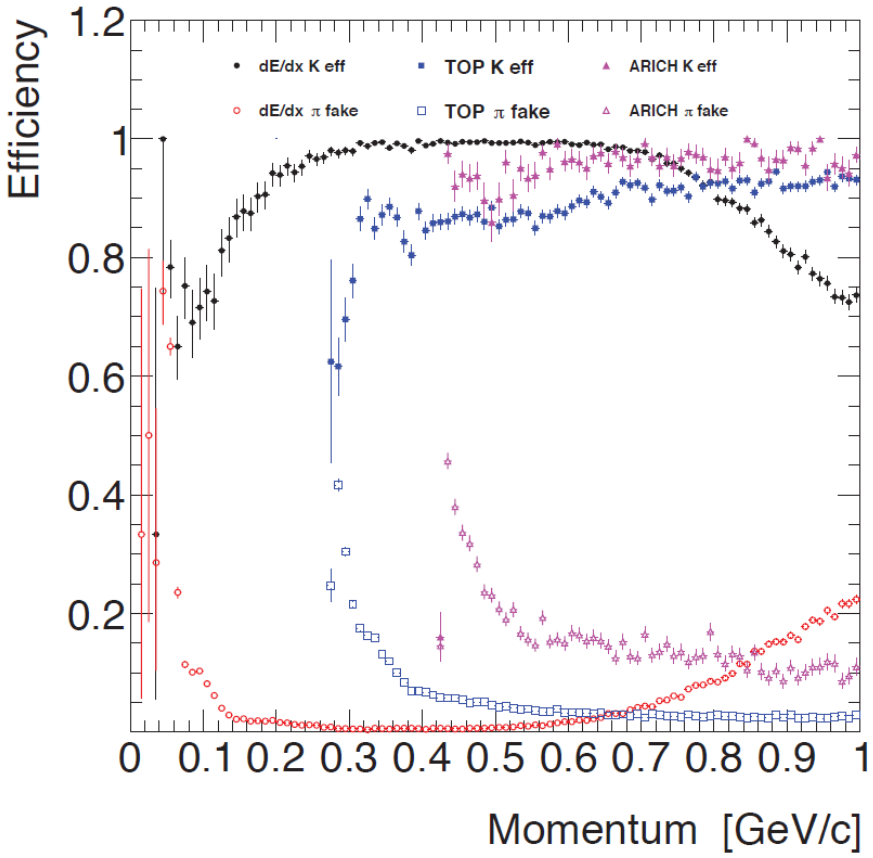
- Detector geometry implemented in Geant4
- Parameters obtained from xml file/database
- Energy deposits stored as SimHits
- Digitization in modules
- Background mixing
- Back-ground overlay
- Fast simulation?





Charged Particle Identification

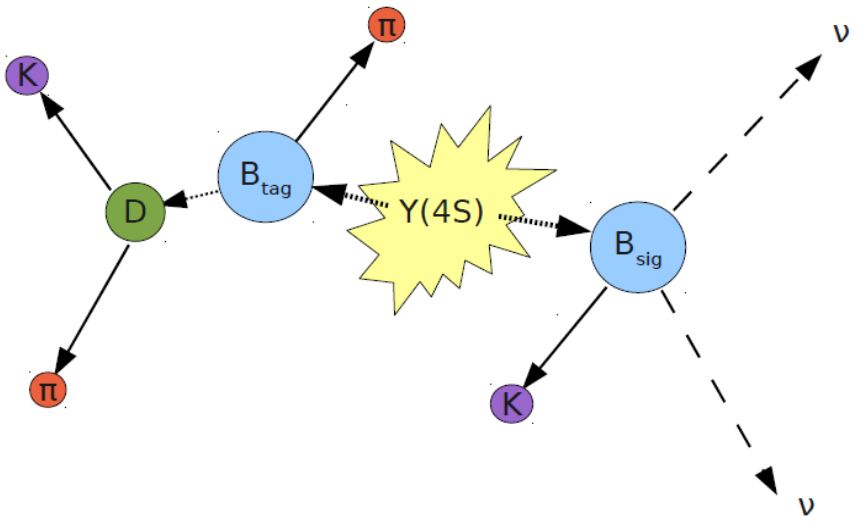


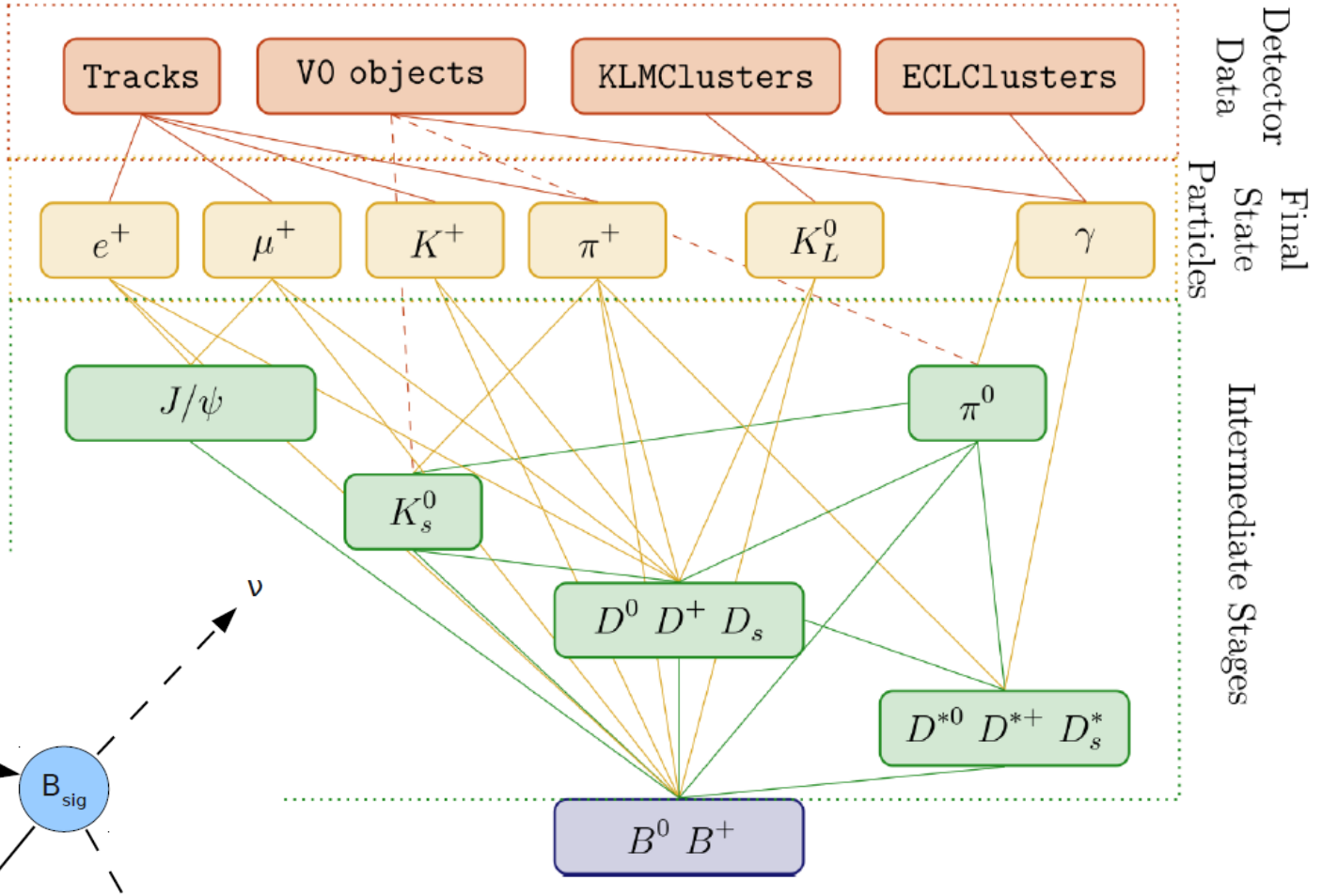


- Neyman Pearson lemma
- Likelihood for each detector:
  L(detector response|part. type)
- Combination:
  product of likelihoods
- Probability can be calculated with analysis dependent priors

Full Event Interpretation

- Huge number of B meson decay modes
- ➔ Hierarchical reconstruction
- ➔ Multivariate classifiers





➔ Tools for analysis specific training of classifiers