BUDKER INSTITUTE OF NUCLEAR PHYSICS

Alexander Zaytsev // A.S.Zaytsev@inp.nsk.su
On behalf of CMD-3 Collaboration
CHEP2009 (Prague, Mar 2009)

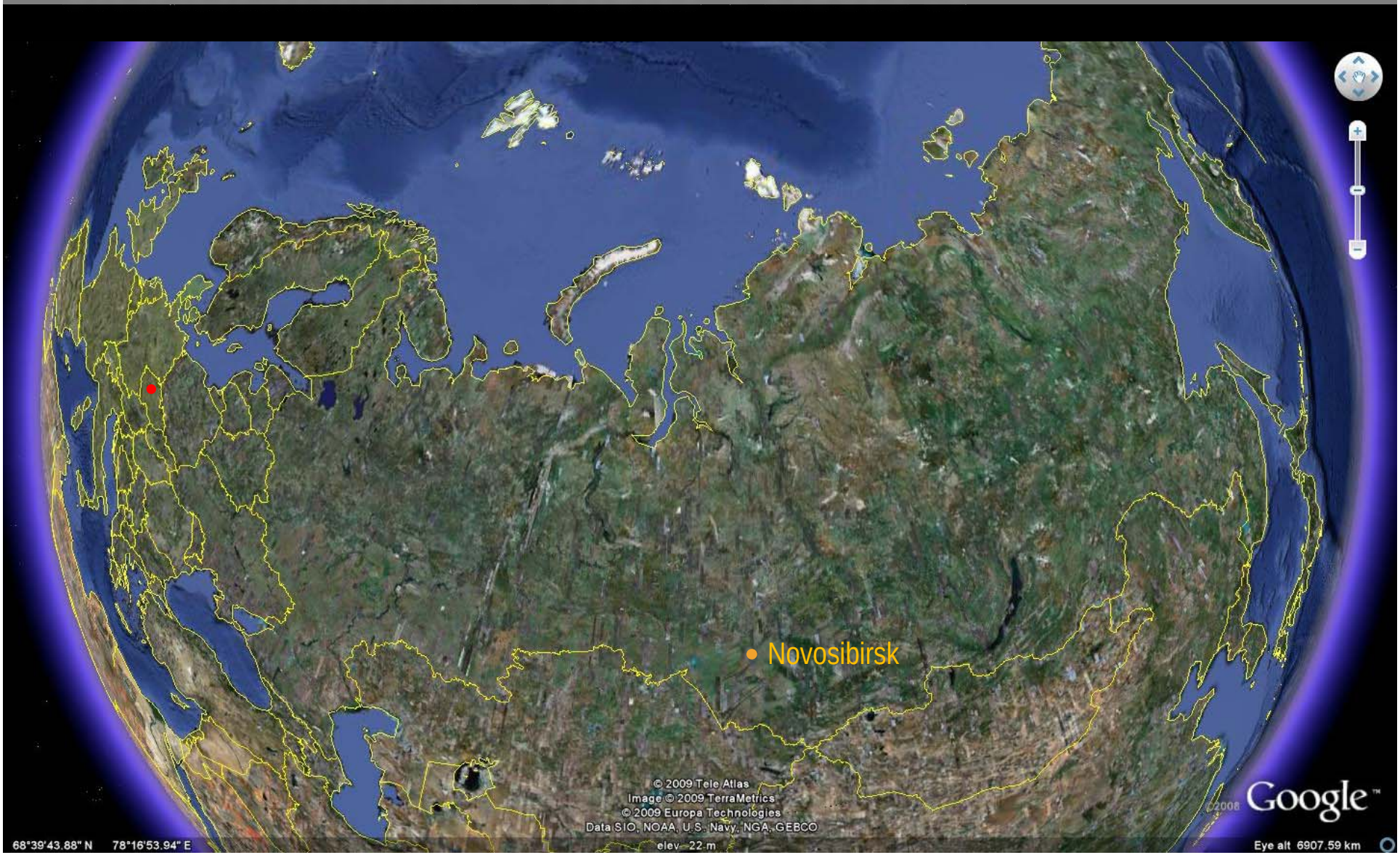
CMD-3 DETECTOR OFFLINE SOFTWARE DEVELOPMENT

OUTLINE

- CMD-3 Experiment Overview
- Basic Offline Software Requirements
- Generic Offline software Design
- Software Integration Solution
- Event Model and Data Persistency
- Status of the Implementation
- Summary

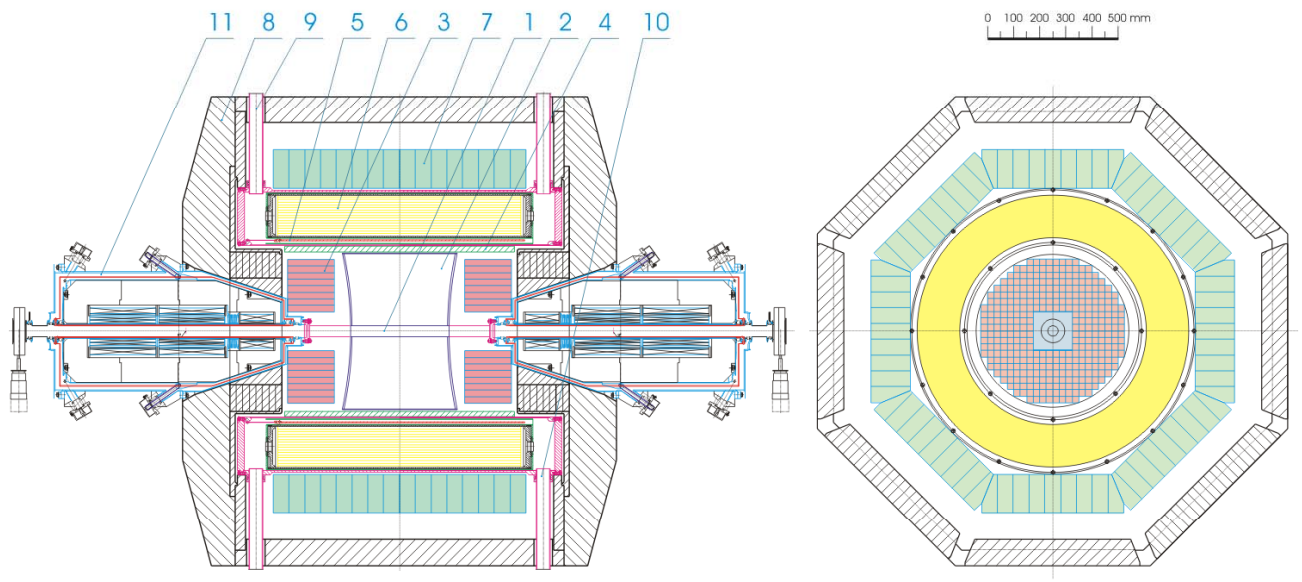
NOVOSIBIRSK / BINP / CMD-3

NOVOSIBIRSK / BINP / CMD-3



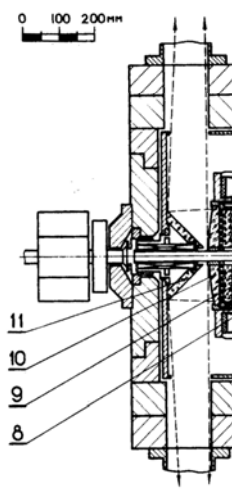
NOVOSIBIRSK / BINP / CMD-3

HISTORICAL PERSPECTIVE



1 - vacuum chamber
2 - drift chamber
3 - **BGO** endcap calorimeter
4 - **Z**-chamber
5 - superconducting solenoid
6 - liquid **Xe** calorimeter
7 - **CsI** barrel calorimeter
8 - iron yoke
9 - liquid **He** supply
10 - vacuum pumpdown
11 - **VEPP2000** superconducting magnetic lenses

CMD-3: many features are inherited from **CMD-2**



1 - vacuum chamber
2 - drift chamber
3 - **Z**-chamber
4 - main solenoid
5 - compensating magnet
6 - **BGO** endcap calorimeter
7 - **CsI** barrel calorimeter
8 - muon range system
9 - iron yoke
10 - storage ring lenses

CMD-2: 1992-2000
Drift chamber, crystal calorimeters (data archival on tapes, Linux cluster for offline processing)

The schematic cross section of the detector. 1 - high voltage feeding, 2 - joke, 3 - nitrogen shell, 4 - magnetic lens of the storage ring, 5 - compensating solenoid, 6 - spark chamber, 7 - main solenoid, 8 - outer MWPC, 9 - inner MWPC, 10 - optic lens, 11 - mirror.

CMD-1: 1979-1984
Spark chamber, MWPCs (photo registration)

CMD-3 DETECTOR SUBSYSTEMS

- Trackers:
  - Drift Chamber (DC)
  - Z-Chamber (ZC)
- Calorimeters:
  - BGO Endcap Calorimeter
  - CsI Barrel Calorimeter
- Calorimeters with tracking capabilities:
  - Liquid Xe Calorimeter (LXe)

State-of-the-art sub-system most challenging from the reconstruction point of view

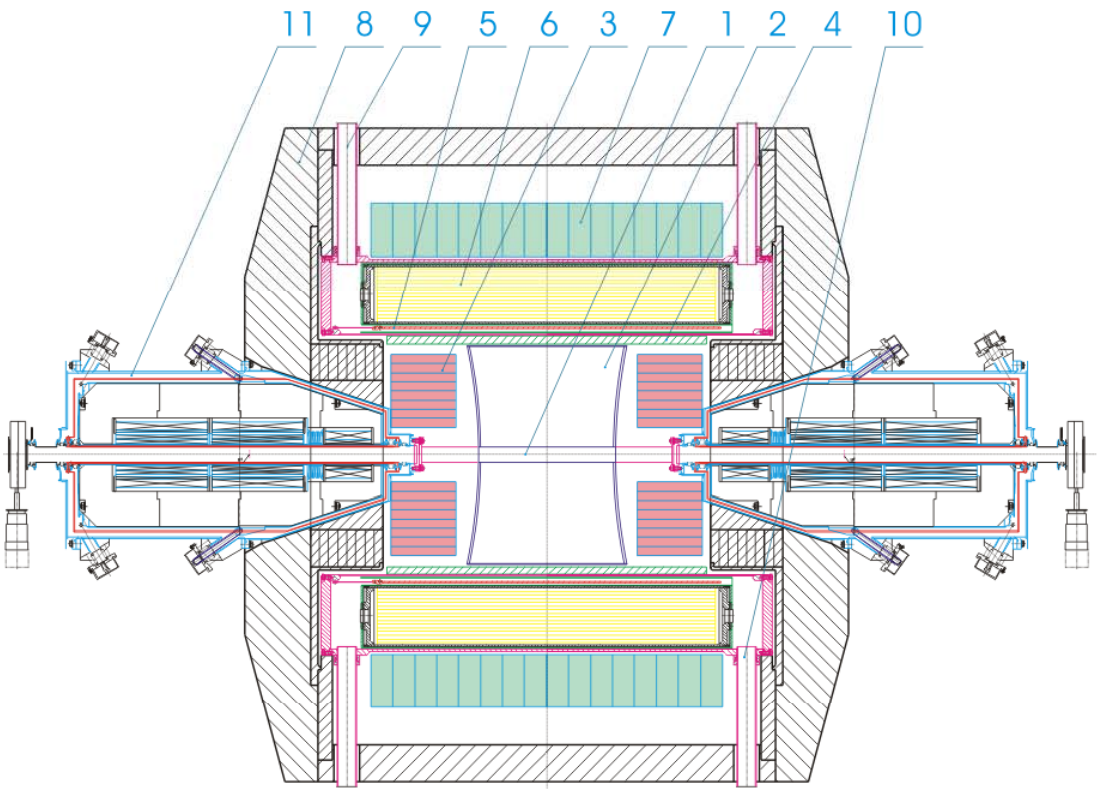
- Particle ID:
  - Muon
  - TOF
- Trigger:
  - Low Level
  - High Level

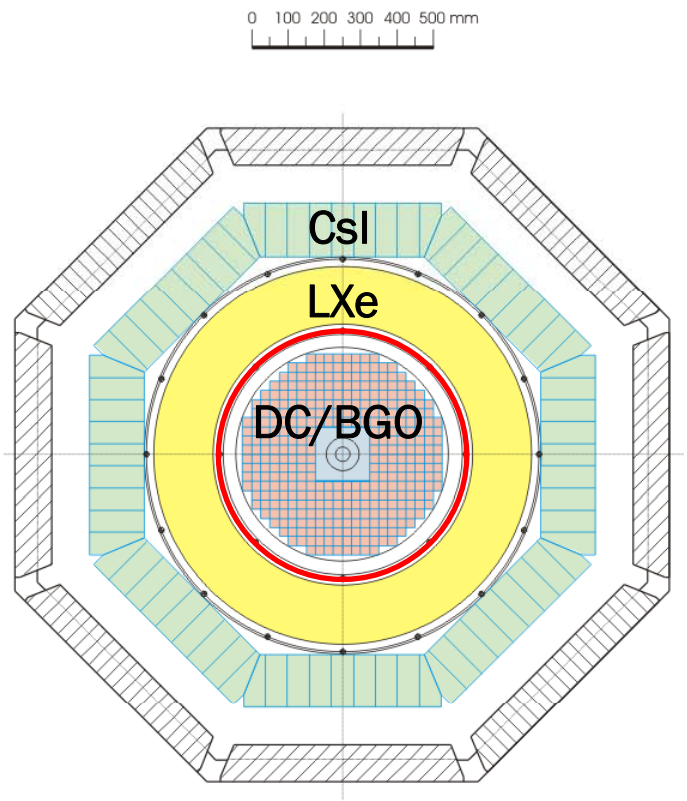
TDAQ rate: 1 kHz
RAW event size: 3 KB
RAW data rate: 3 MB/s

Total amount of RAW data stored for analysis: 10 TB

We are not limited by computing power: all the experimental data can be re-processed at once when needed, factor of 10 more of reconstruction/simulation data is to be produced

CMD-3 GENERAL LAYOUT





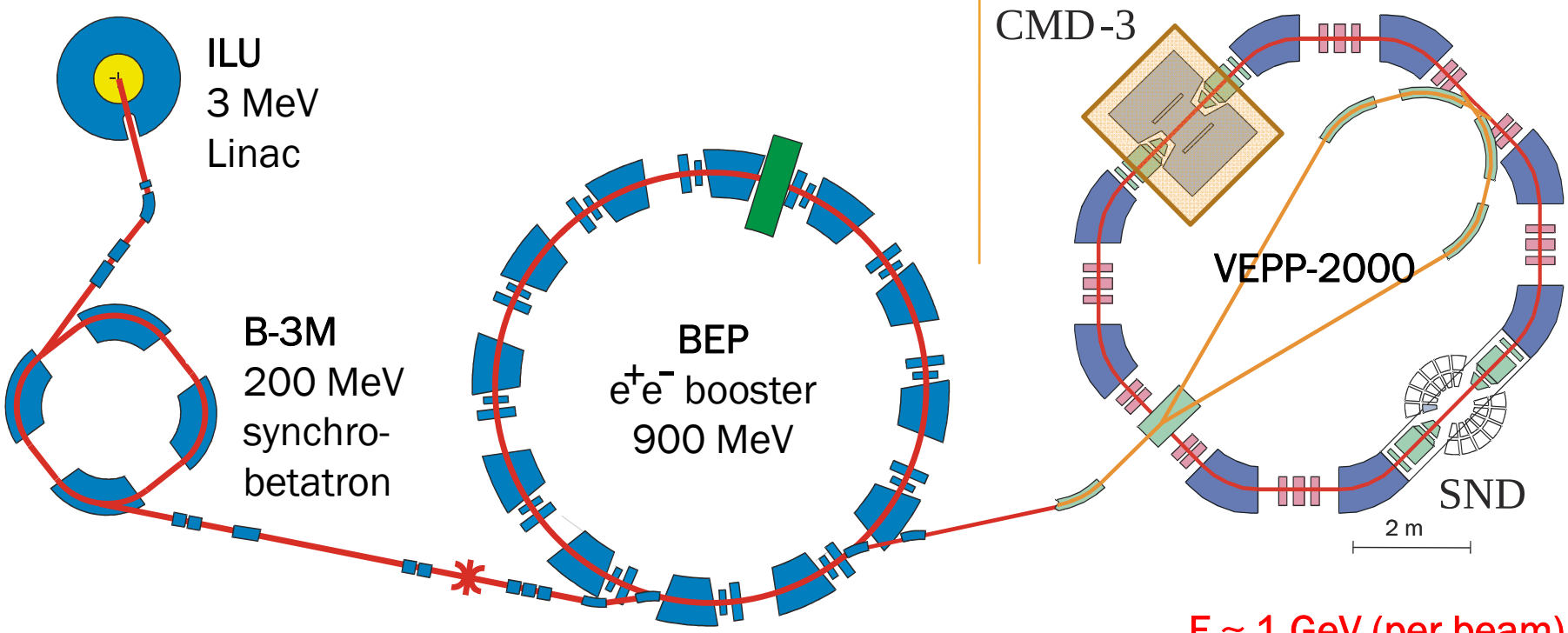
1 - vacuum chamber
2 - drift chamber
3 - **BGO** endcap calorimeter
4 - **Z**-chamber
5 - superconducting solenoid
6 - liquid **Xe** calorimeter
7 - **CsI** barrel calorimeter
8 - iron yoke
9 - liquid **He** supply
10 - vacuum pumpdown
11 - **VEPP2000** superconducting magnetic lenses

Universal
Cryogenic Magnetic Detector

VEPP-2000 COLLIDER COMPLEX

Inherited from the previous VEPP-2M machine

Built from the scratch



E ≈ 1 GeV (per beam)

L ≈ 1×10^{32} cm^{-2} sec^{-1} (1×1 bunch, round beams)

WHY ARE WE DOING IT?

- Fundamental constants of particle physics and the quantity $R = \frac{\sigma(e^+e^- \to hadrons)}{\sigma(e^+e^- \to \mu^+\mu^-)}$
- Exclusive channels of e^+e^- annihilation to hadrons $e^+e^- \to 2h,\ 3h,\ 4h$, with $h = \pi,\ K,\ \eta$
- Study of known and search for new vector mesons
- Search for exotic hadrons
- Study of $e^+e^- \to p\bar{p},\ n\bar{n}$
- CVC tests: comparison of $e^+e^- \to hadrons$ cross section with $\tau \to \nu_\tau + hadrons$ decay spectra
- Hadronic processes with hard photon emission $e^+e^- \to \gamma^*\gamma,\ \gamma^* \to hadrons$
- Two photon physics $e^+e^- \to e^+e^- + hadrons$: $\gamma\gamma \to \pi^0,\ \eta,\ \eta',\ a_0\,(980),\ f_0\,(980),\ 2\pi,\ 3\pi,\ n\pi,\ 2K,\ \eta\pi$, etc.
- Tests of high order QED

Mainstream task: extension of the work done with CMD-2 with broader energy range and much higher precision

The most challenging task from the point of view of building reconstruction algorithms (yet to be handled)

BASIC SOFTWARE REQUIREMENTS

- Modularity: the reconstruction application might consist of up to 100 components (modules) bonded in runtime or compilation time
- Flexibility: changing the layout of module interaction without recompiling the code
- Parallelization schema: splitting input data
- Number of threads: single thread mode
- Batch environment compatibility: extensive log management support, switching off all the interactive features
- Lifetime/Support: 10-15 years
- Persistent data lifetime: 15-20 years

The main challenge is to build the system which is suitable for both development and production environment

RECONSTRUCTION SOFTWARE DESIGN

- Software integration platform: custom
- Internal data exchange mechanism: custom
- Configuration representation: XML
- Generic containers/data representation: STL & ROOT containers wrapped by custom high level experimental event container
- Persistency platform: ROOT

- Standard build/runtime environment :
  - Platform: Scientific Linux (SL, both x86 and x86_64)
  - Main programming language: C++
  - Compilers involved: GCC
  - Build system: GNU/Make
  - API documentation: Doxygen, TWiki
  - Use of code generation: ROOT dictionaries
  - Source code management: CVS >> SVN
  - Coding convention exists from the early days of the project

SOFTWARE INTEGRATION PLATFORM

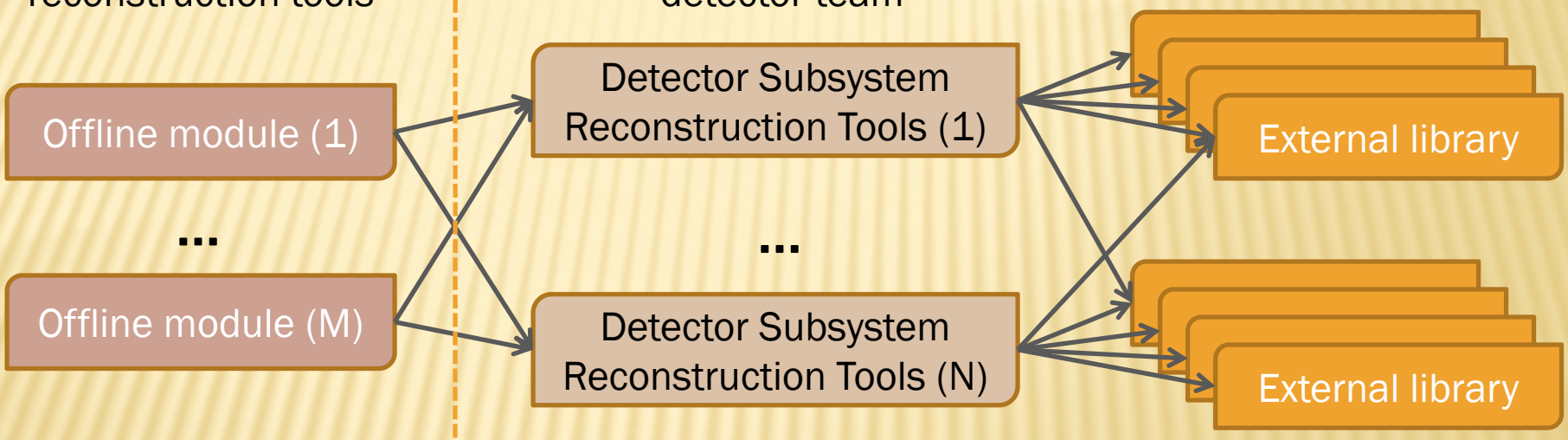
Cmd3Fwk Software Integration Framework

Development (DSO based) environment

Wrapper software encapsulating the reconstruction tools

Software developed by a particular sub-detector team

External software (ROOT, GSL, Boost, etc.)



Introduces the subdetector tools to the software integration framework

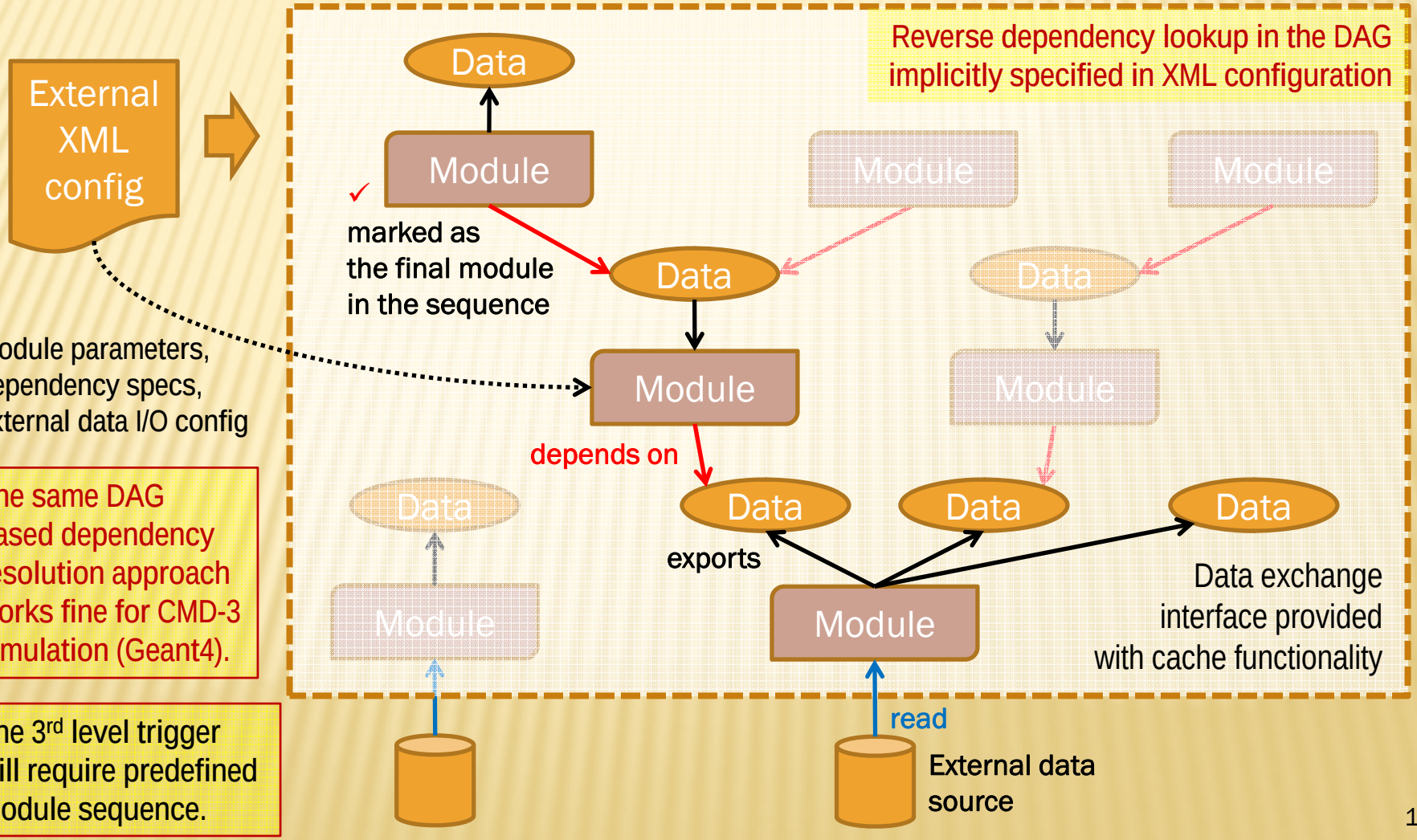
Each block represent a part of subsystem specific reconstruction algorithm

No data exchange between the consequent parts of reconstruction occur here!

SOFTWARE INTEGRATION PLATFORM

Cmd3Fwk Software Integration Framework

Development (DSO based) environment



The same DAG based dependency resolution approach works fine for CMD-3 simulation (Geant4).

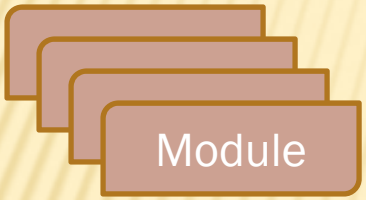
The 3rd level trigger will require predefined module sequence.

SWITCHING TO PRODUCTION MODE

Cmd3Fwk Software Integration Framework

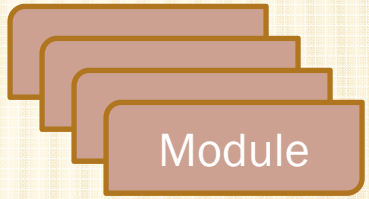
Not yet available in current Cmd3Fwk version branch 1.x

Development mode: short term OS environment compatibility



Cmd3Fwk CLI Tools
(DSO based modules)

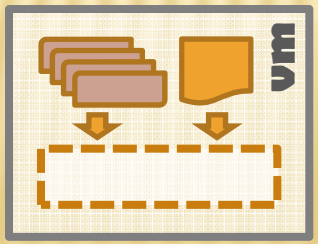
Production mode:
1-3 years of OS environment compatibility



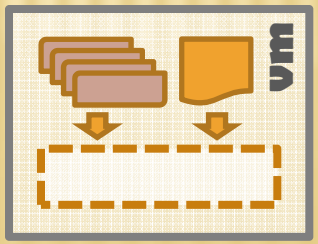
Cmd3Fwk Static Snapshot (no DSO modules)

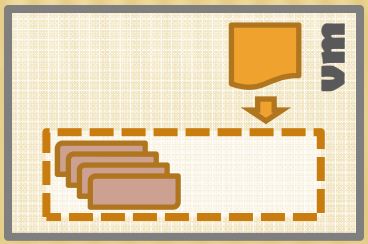
Building VM images of the production environment

Production mode, long term OS environment compatibility, high reproducibility



...



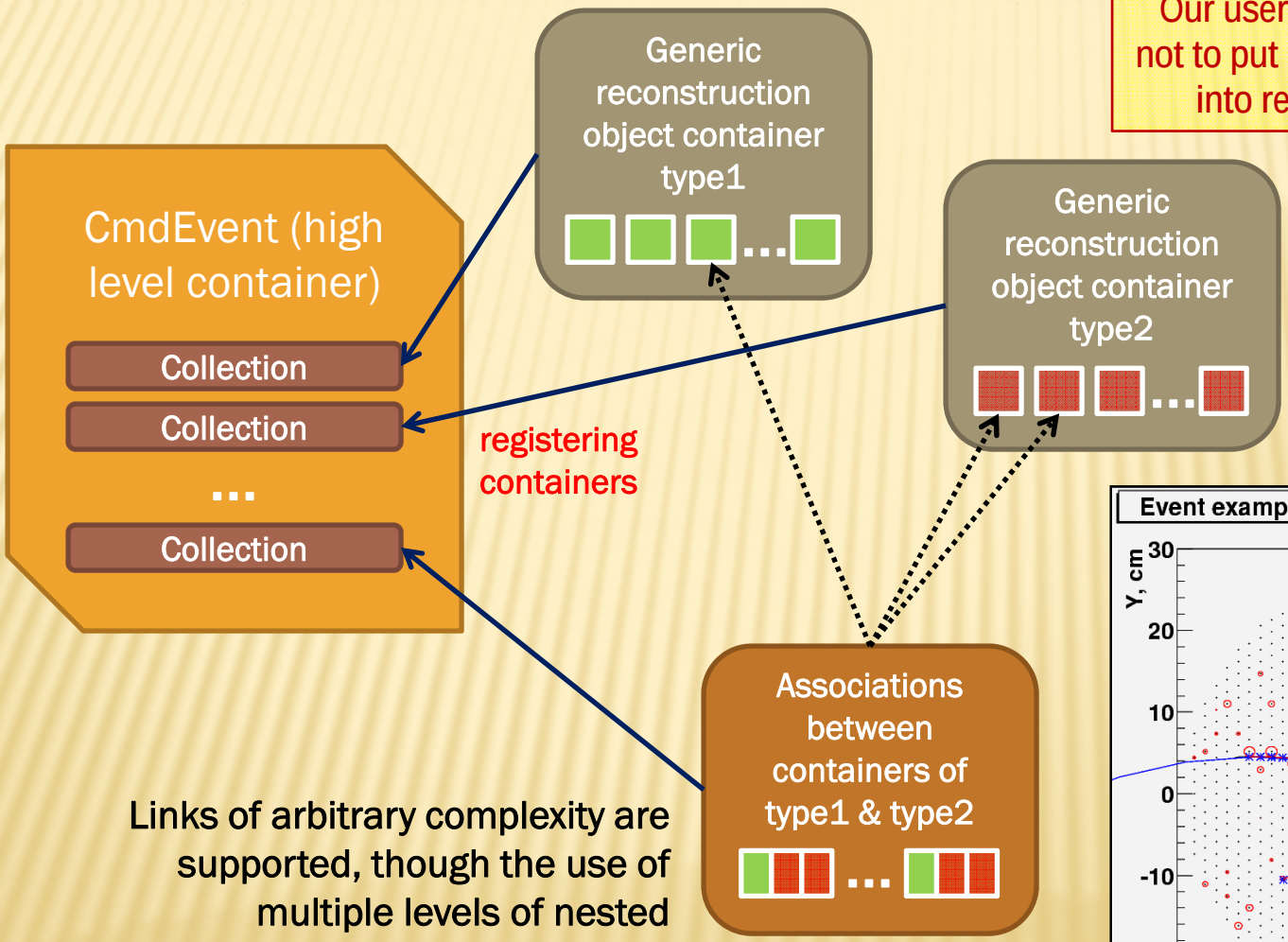


INTEGRATION PLATFORM IMPROVEMENTS

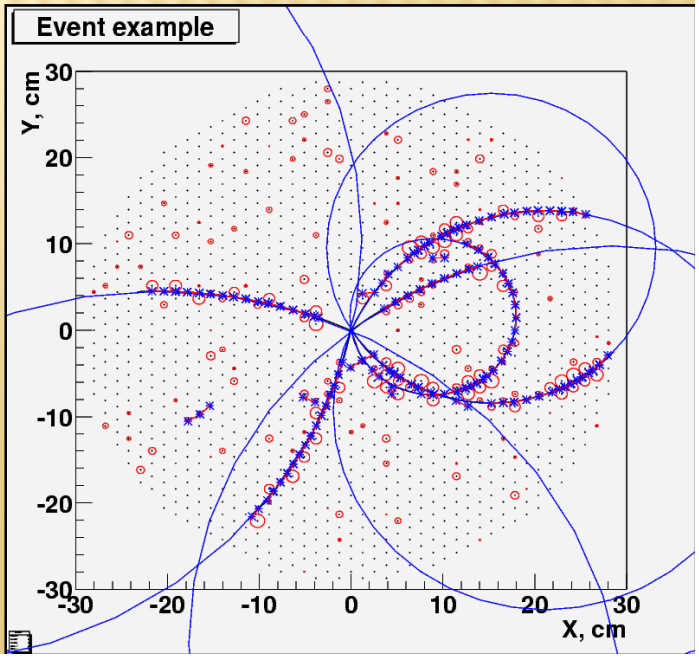
- A new version branch of Cmd3Fwk software integration product (v2.0) which should replace current production version v1.3.x in the near future is being prepared for the release
- The new version should provide the features missing which are required both for production and development environment, e.g.:
  - Centralized log messages management
  - Internal debugging mechanisms
  - Extensible XML configuration support
  - Configuration time validation mechanisms
  - Multiple dependency resolution modes
  - Revised module registration interface
  - Much more advanced data access mechanism
  - … and more
- Migration to the new version will require changes only in a small part of the code written by users

EXPERIMENTAL EVENT MODEL

Our users are strongly encouraged to not to put interdependency information into reconstruction objects directly

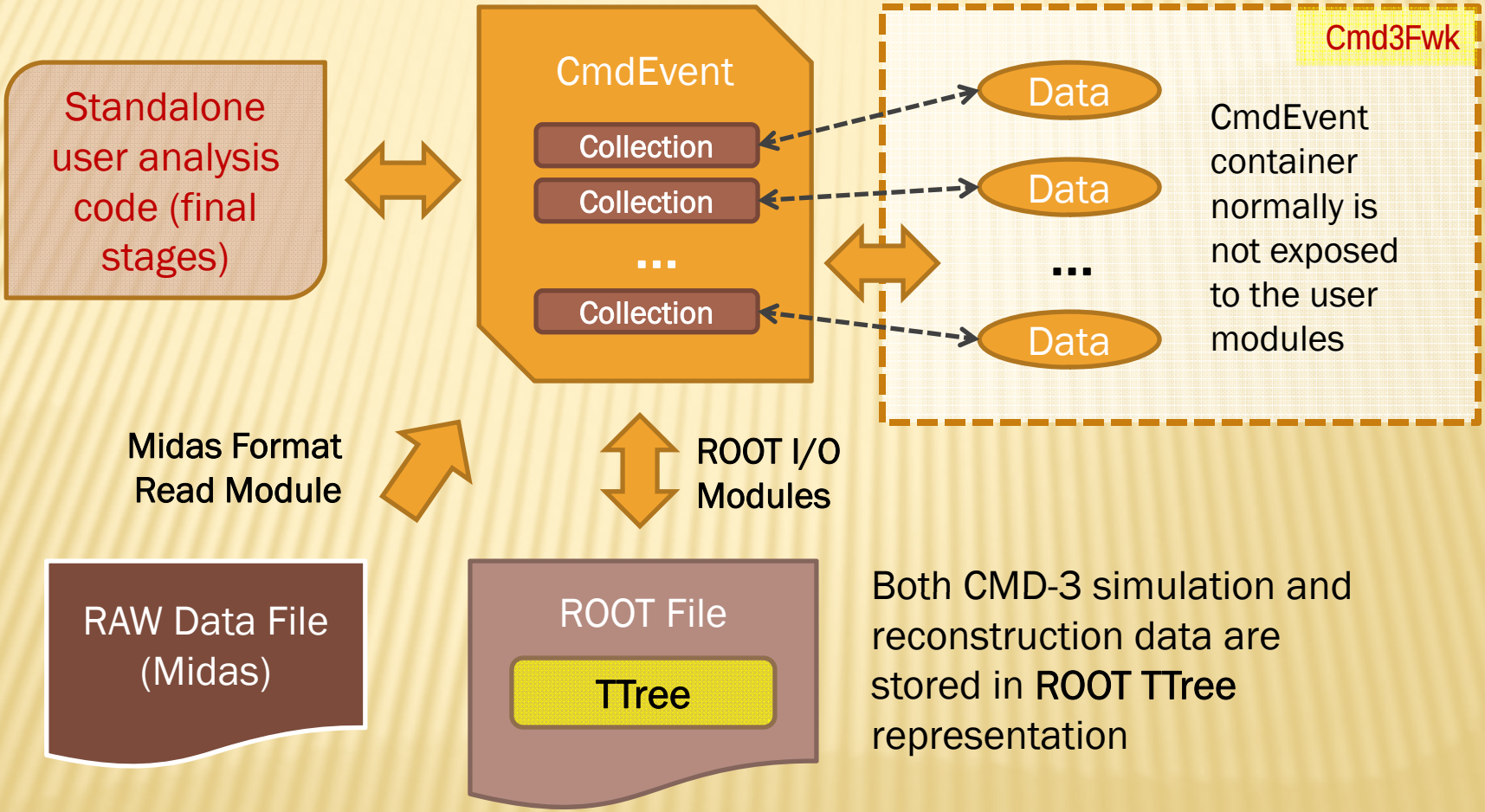


Example:
type1: list<Track>
type2: list<ClbrHit>
link(t1-t2): list< Track, list<ClbrHit> >



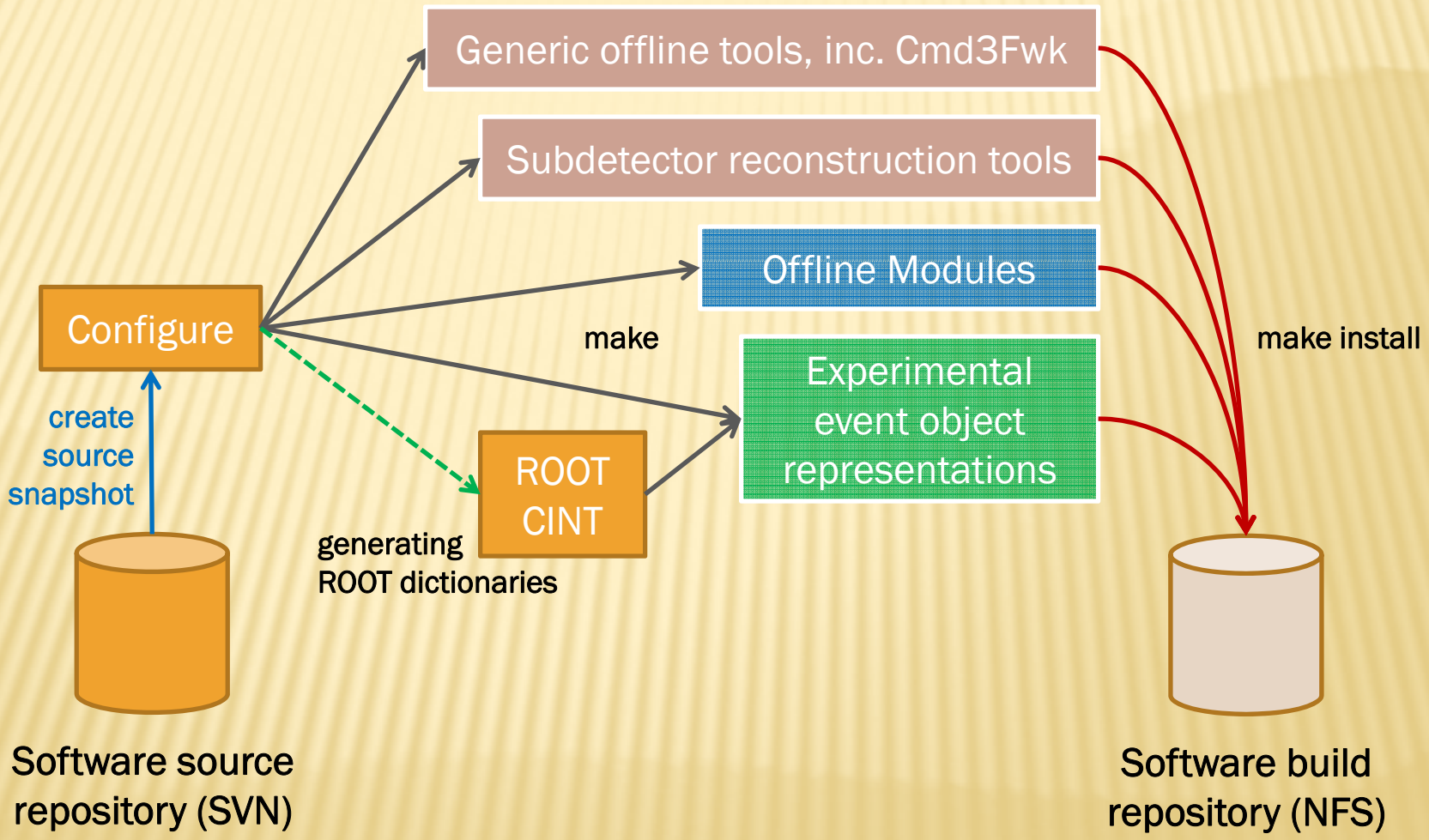
Links of arbitrary complexity are supported, though the use of multiple levels of nested containers must be justified

DATA I/O MECHANISMS

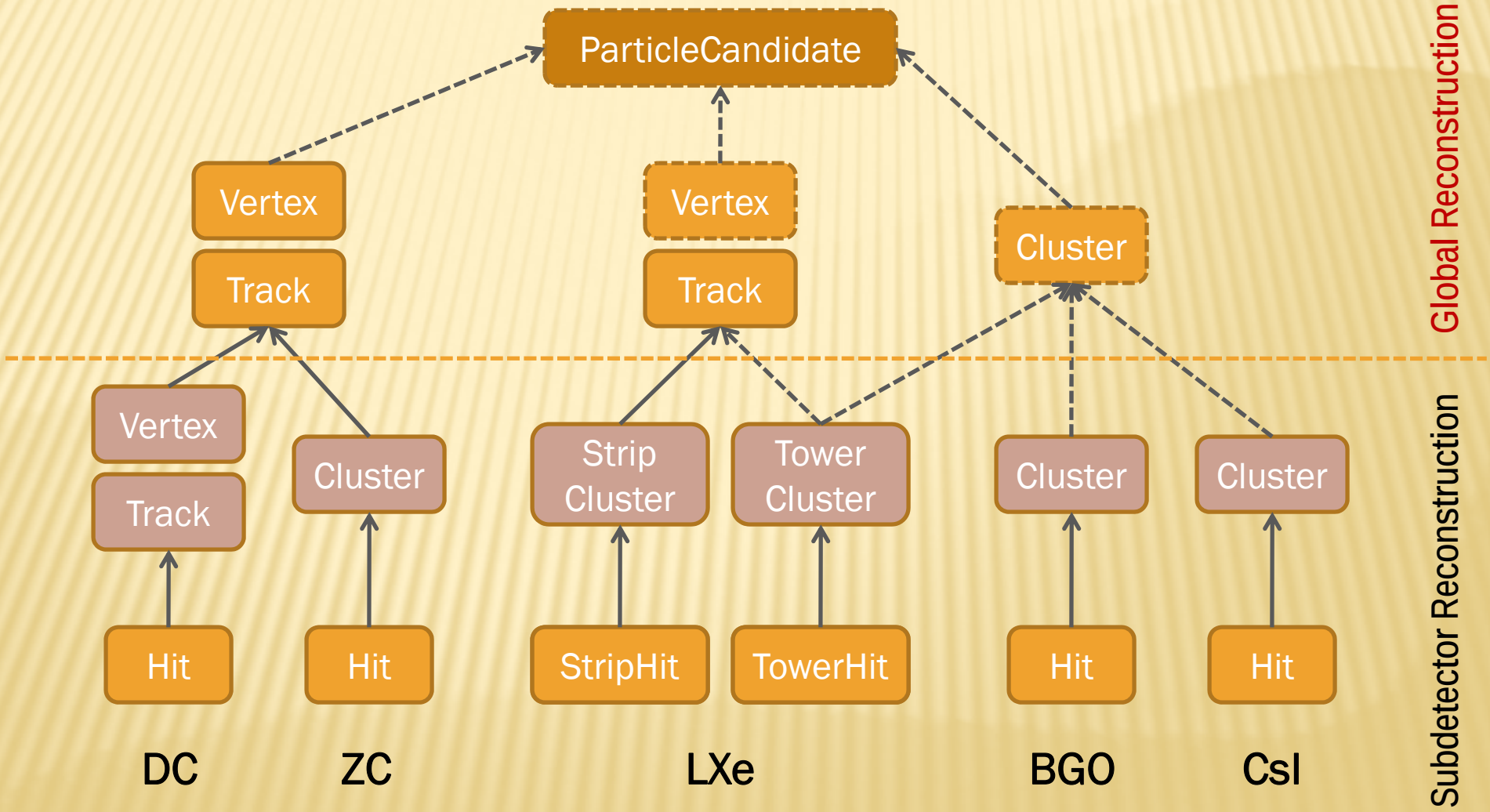


Both CMD-3 simulation and reconstruction data are stored in **ROOT TTree** representation

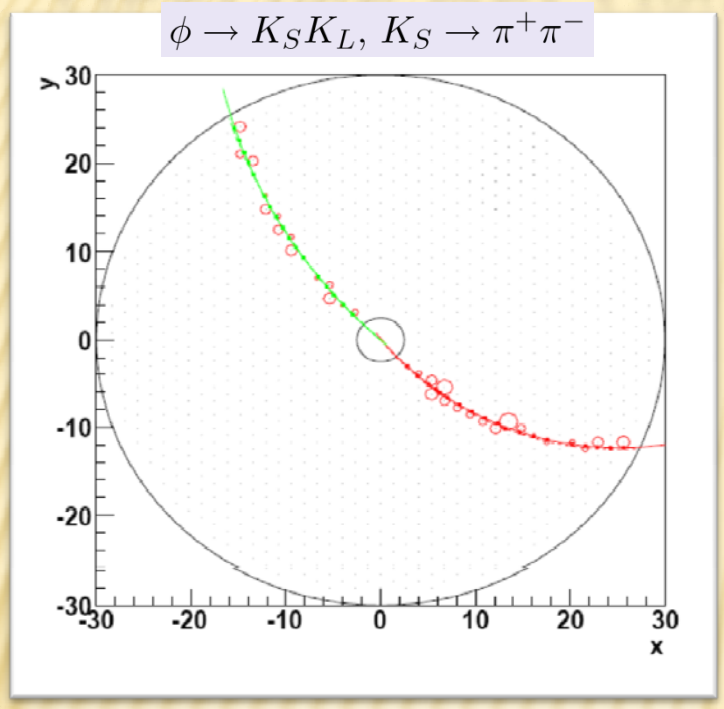
SOFTWARE BUILD SYSTEM

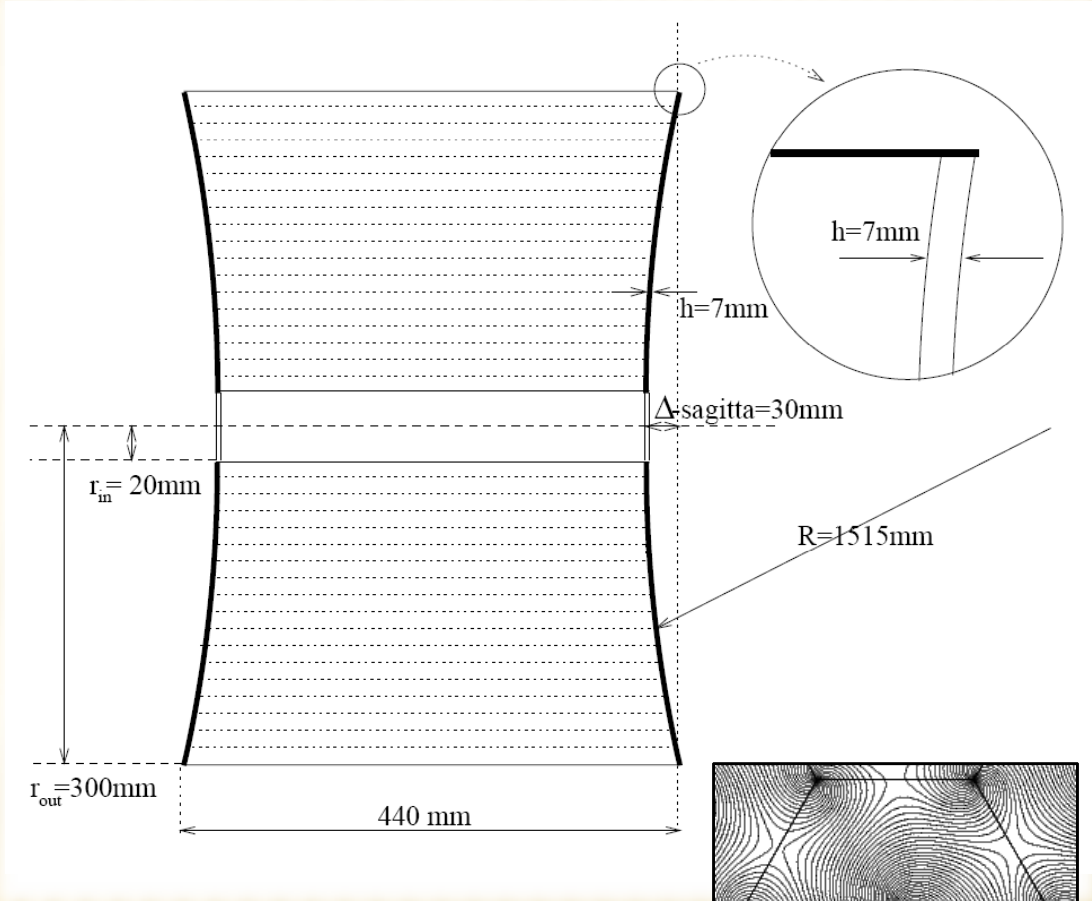


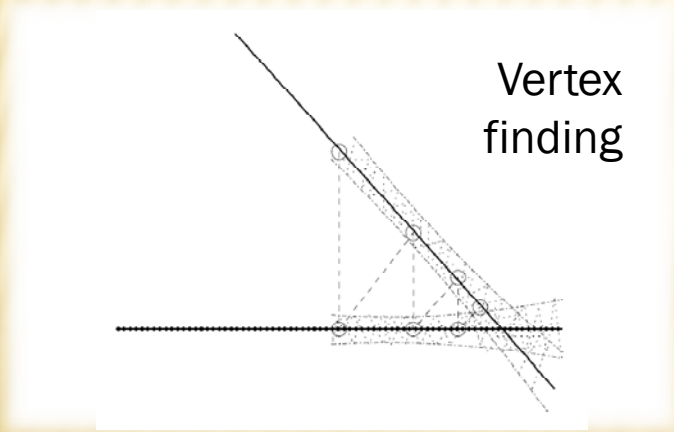
GENERIC RECONSTRUCTION LAYOUT



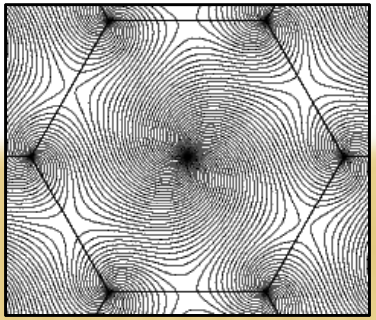
OFFLINE RECONSTRUCTION: DC







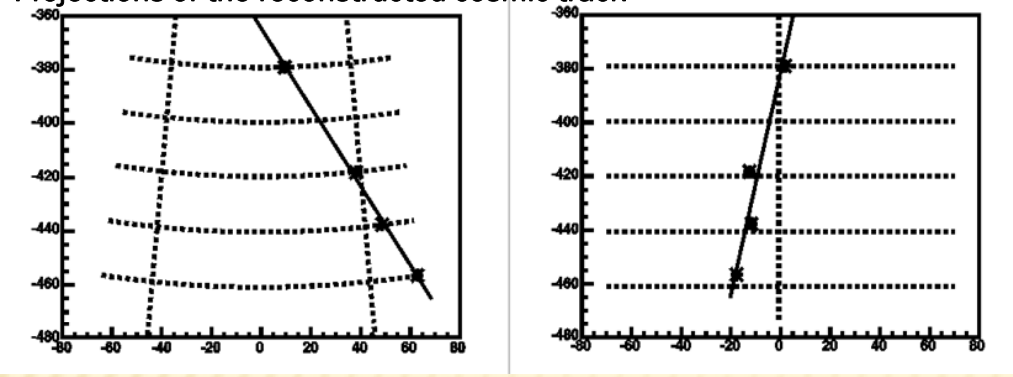
Vertex finding

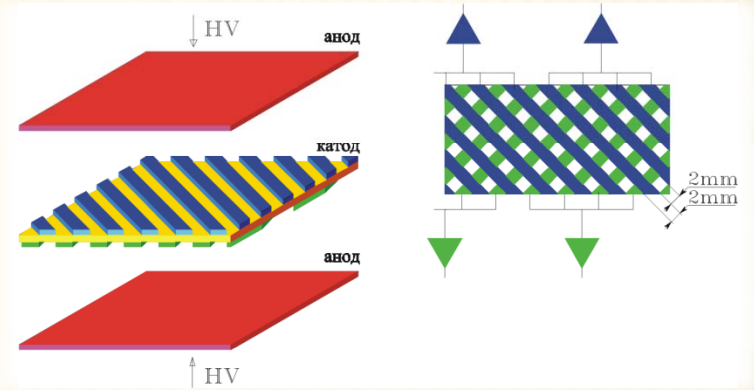


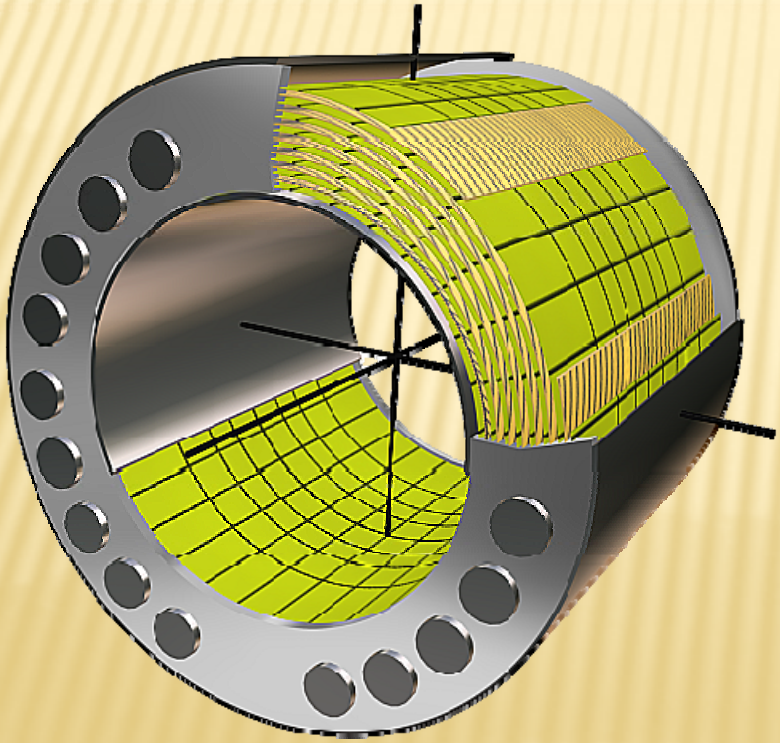
Garfield simulation of the drift lines in 1.5T magnetic field

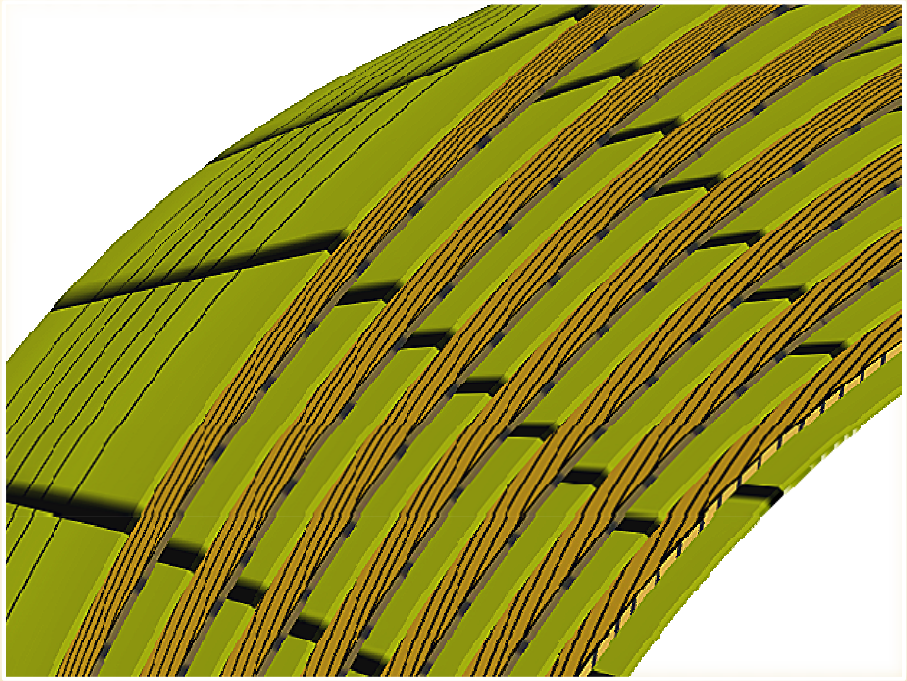
OFFLINE RECONSTRUCTION: LXE

Projections of the reconstructed cosmic track



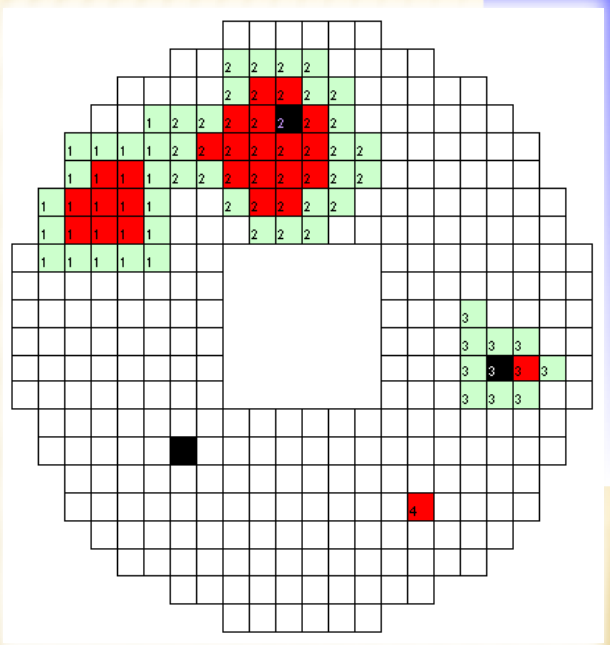


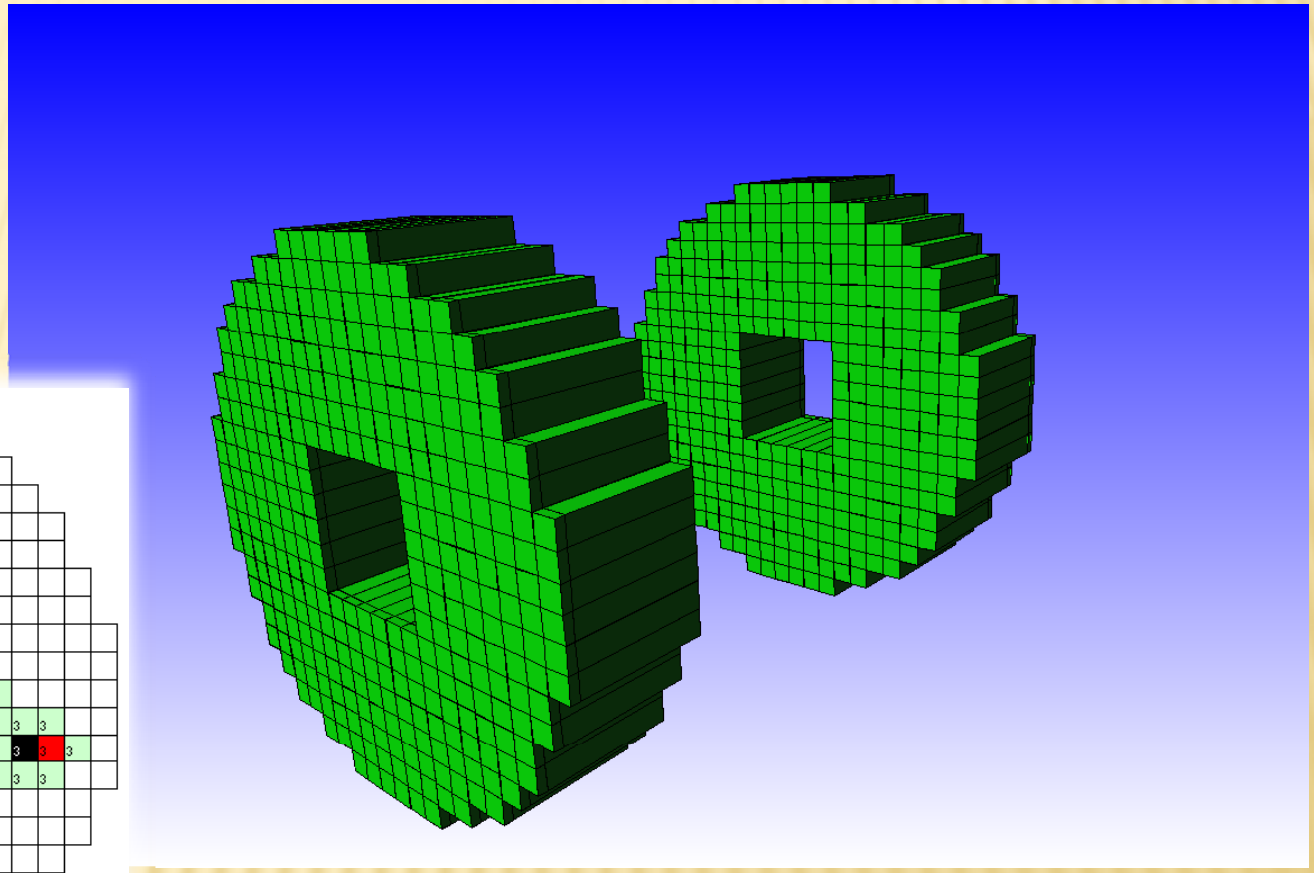




OFFLINE RECONSTRUCTION: BGO

Finding clusters
in BGO calorimeter





OTHER OFFLINE SOFTWARE COMPONENTS

- Detector geometry and material description (used also in Geant4 simulation)
- Detector and experimental event visualization (custom solution)
- Trigger simulation (required also for the offline processing)
- Offline calibration DBs and calibration procedures for detector subsystems
- Reconstruction XML configuration DB and a web interface on top of it
- Test production environment (SL5 x86_64):
  - host systems (Linux cluster)
  - test XEN virtual machines
- Tools for running production simulation/reconstruction jobs

CONCLUSION

- CMD-3 detector commissioning is now being finalized
- Main reconstruction algorithms required by detector subsystems individually are implemented, the work on their improvement and validation is in progress
- Many auxiliary components are in production state and extensively used in test runs (calibration tools, event display, etc.)
- The prototype of the offline data processing environment is implemented and being exploited on the data produced by Geant4 detector simulation and the first experimental data obtained during the test runs
- Next step is to put the whole system into production with a new version of the offline software integration framework
- Global reconstruction algorithms are yet to be implemented