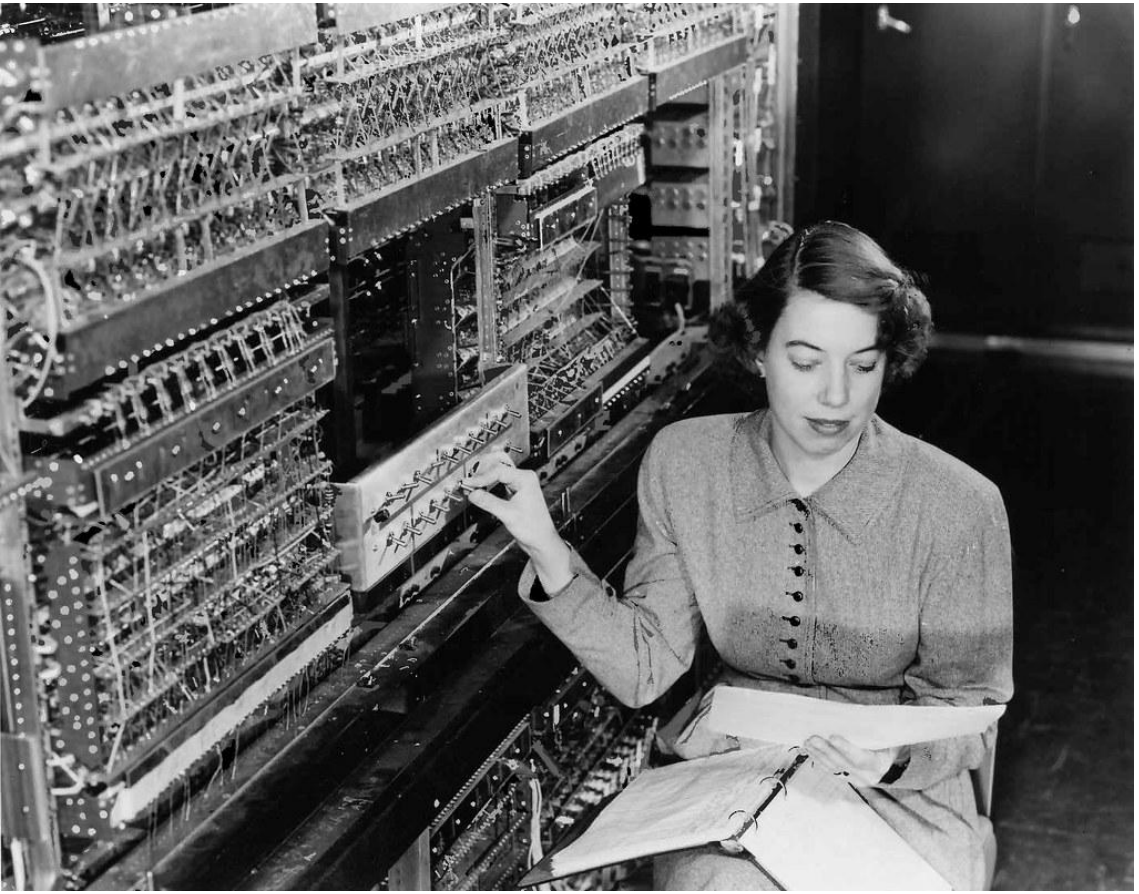


# ATHENA Software

EICUG Summer Meeting
2021-08-05

**The Software and Computing WG Conveners:**
Andrea Bressan (University of Trieste and INFN) ,
Dmitry Romanov (Jefferson lab) ,
Sylvester Joosten (Argonne National Laboratory) ,
Whitney Armstrong (Argonne National Laboratory) ,
Wouter Deconinck (The University of Manitoba)

# *Philosophy:* Let's prepare for our future at the EIC!

- **Build forward-looking team of software developers to ensure the long-term (decades!) success of the EIC scientific program.**
- Focus on modern scientific computing practices
  - Strong emphasis on modular, orthogonal tools
  - Integrate with HTC/HPC, CI workflows, and enable use of data-science toolkits
- Avoid "not-invented-here" syndrome, rather leverage cutting-edge CERN-supported software components where possible.
  - Build on top of mature, well-supported, and actively developed software stack.
- Actively work with the EICUG SWG to help develop and integrate new community tools.

# Oversimplified software stack

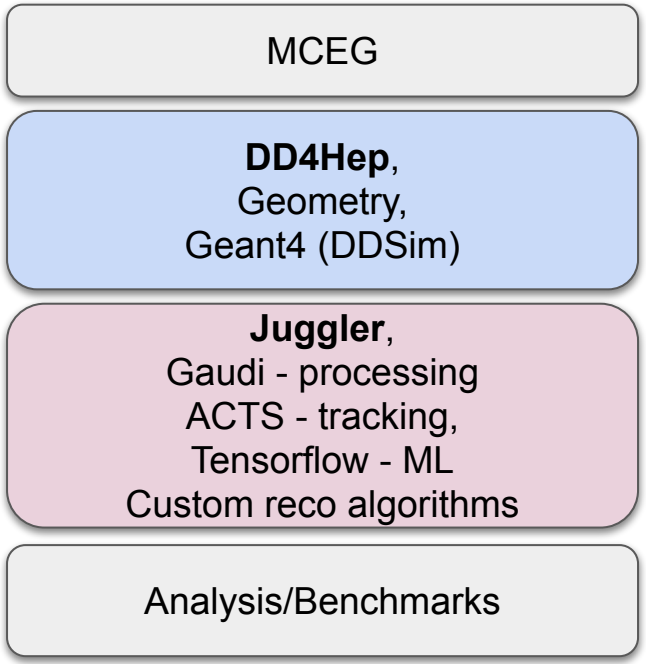
**DD4hep**: Geant4 geometry, detector plugin library, wrappers to run Geant4

**Juggler**: Digitization and reconstruction software (based on Gaudi with Podio-based data model and ACTS for tracking)

**Gaudi**: Generic open project for building event processing frameworks. Enables modern task-based concurrent execution in a heterogeneous computing environment. Used by ATLAS and LHCb.

**ACTS**: Experiment-independent tracking toolkit (ACTS' geometry constructed from DD4hep via plugin)

**Podio**: Robust data model definition to cross the boundaries between the tools



# ATHENA Software Ecosystem: emphasize modularity

**DD4hep** (CLiC, SLD): define active and passive geometry, readout segmentation

Simulation in npsim (DDG4) provides hits for Gaudi

**Gaudi** (originally LHCb): connect geometry and data to algorithms

**Flat output files** (podio) enable flexible workflows (ROOT, python, …) without additional dependencies, available worldwide (S3, xrootd)

DD4hep geometry service in Juggler

ACTS has built-in support for DD4hep geometries

G4 backend (can be swapped out in the future, e.g. celeritas, opticks, …)

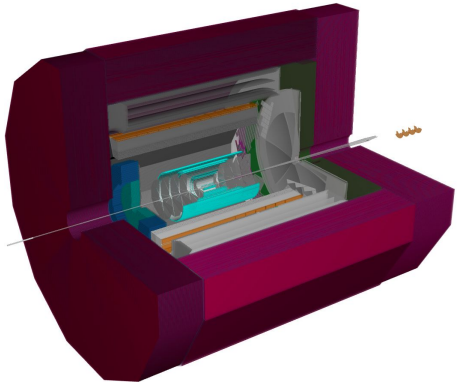
PyTorch TensorFlow

**Juggler**: I/O, digitization, reconstruction, clustering and tracking

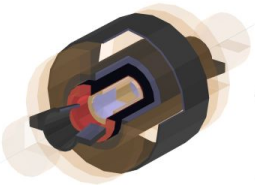
Call ACTS for tracking & vertexing, providing hits and detector geometries

**ACTS** (ATLAS): provide general track reconstruction algorithms

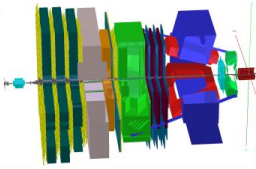
# The DD4hep community



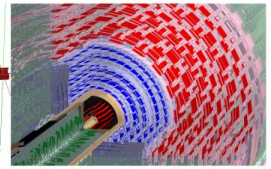
Production



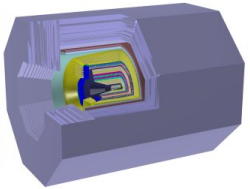
Production



Production
Run 3



Under
investigation
Run 3



Production

Super Charm
Tau Factories

"framework for providing a complete solution for full detector description (geometry, materials, visualization, readout, alignment, calibration, etc.)"

ACTS is rapidly becoming the standard across particle and nuclear physics

# The ACTS community

ATHENA

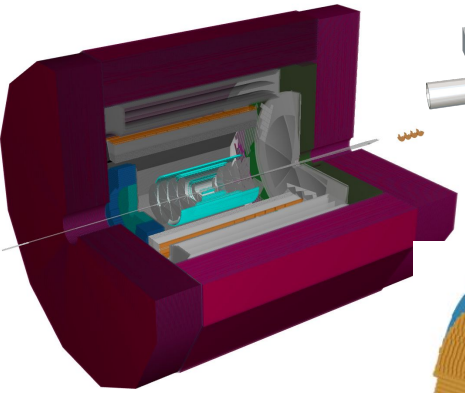
ATLAS EXPERIMENT

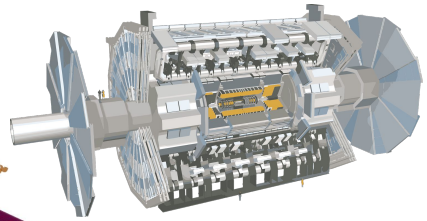
sPHENIX

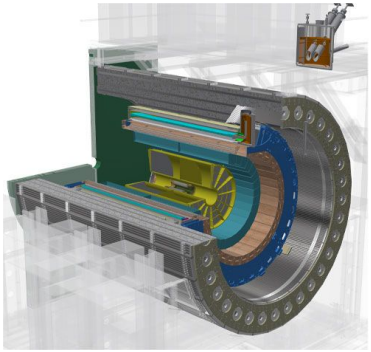
CEPC

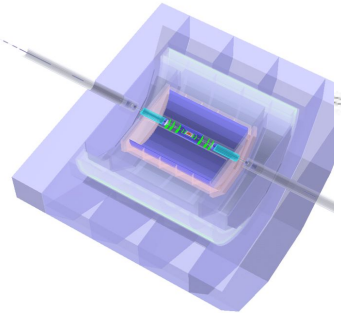
Belle II

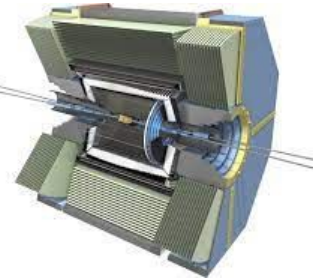
ATHENA

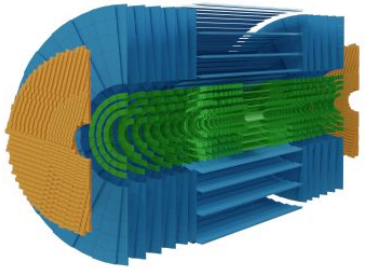


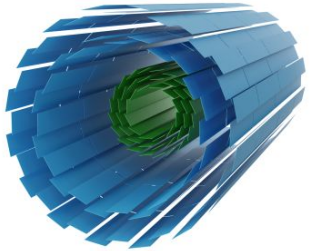


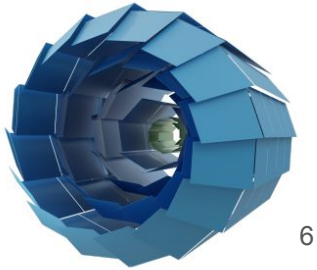












# Automated workflows at eicweb

**GitLab server** (eicweb.phy.anl.gov)

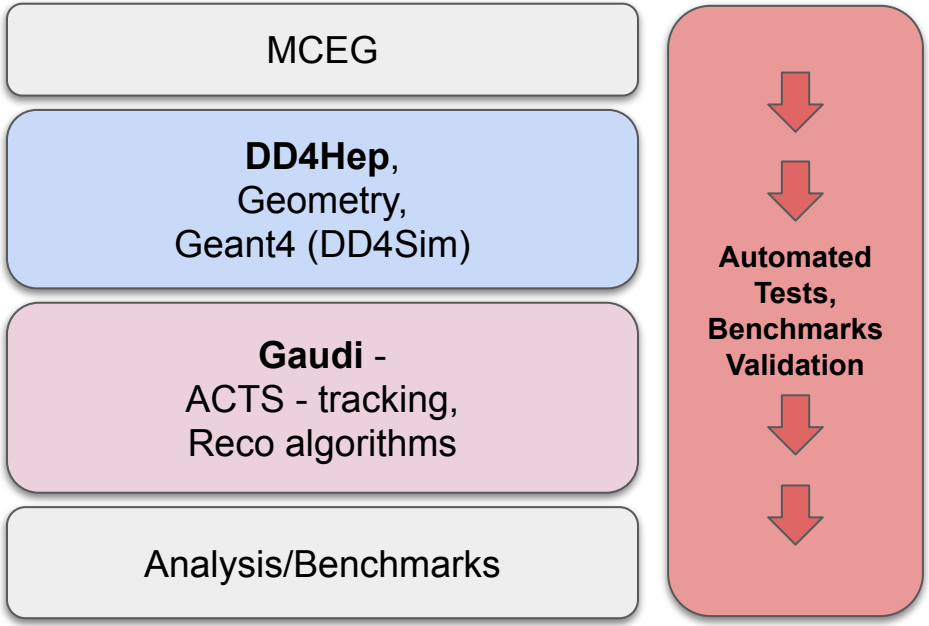
- continuous integration
- dedicated build cluster

Runs automatically on each user commit, executing workflows running multiple tests, benchmarks and analysis

## Automated containers

Both Docker and Singularity images are created nightly or on demand (commit) providing:

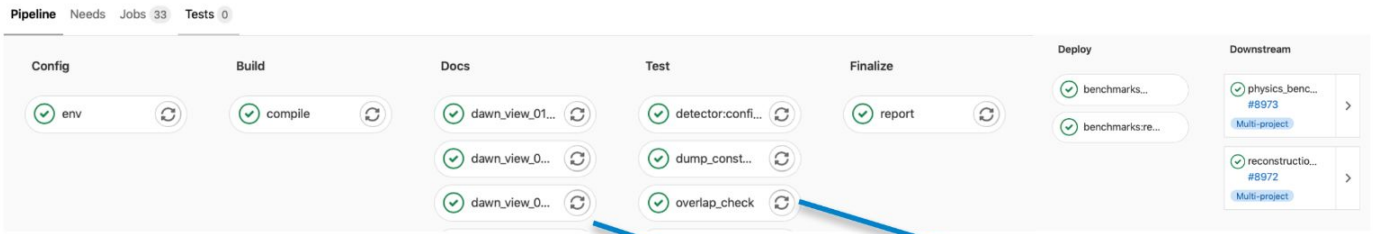
- reproducibility,
- production level images
- latest updates for those working locally

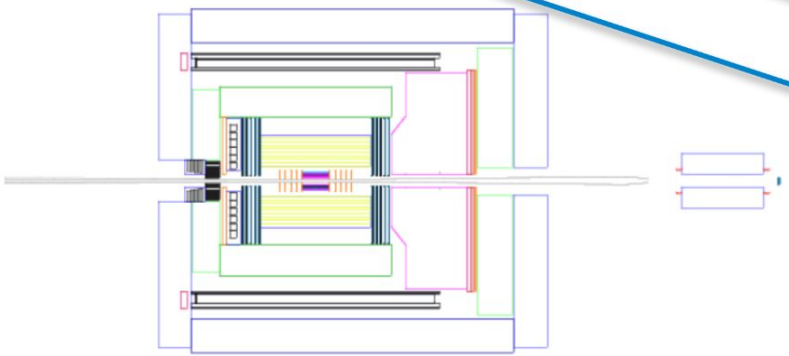


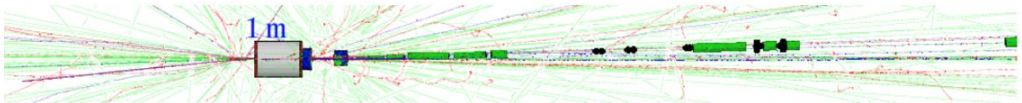
# Easy to get started … takes 1 line for local deploy!

```bash
curl https://eicweb.phy.anl.gov/containers/eic_container/-/raw/master/install.sh | bash
```

# Benchmarks, documentation, conterization



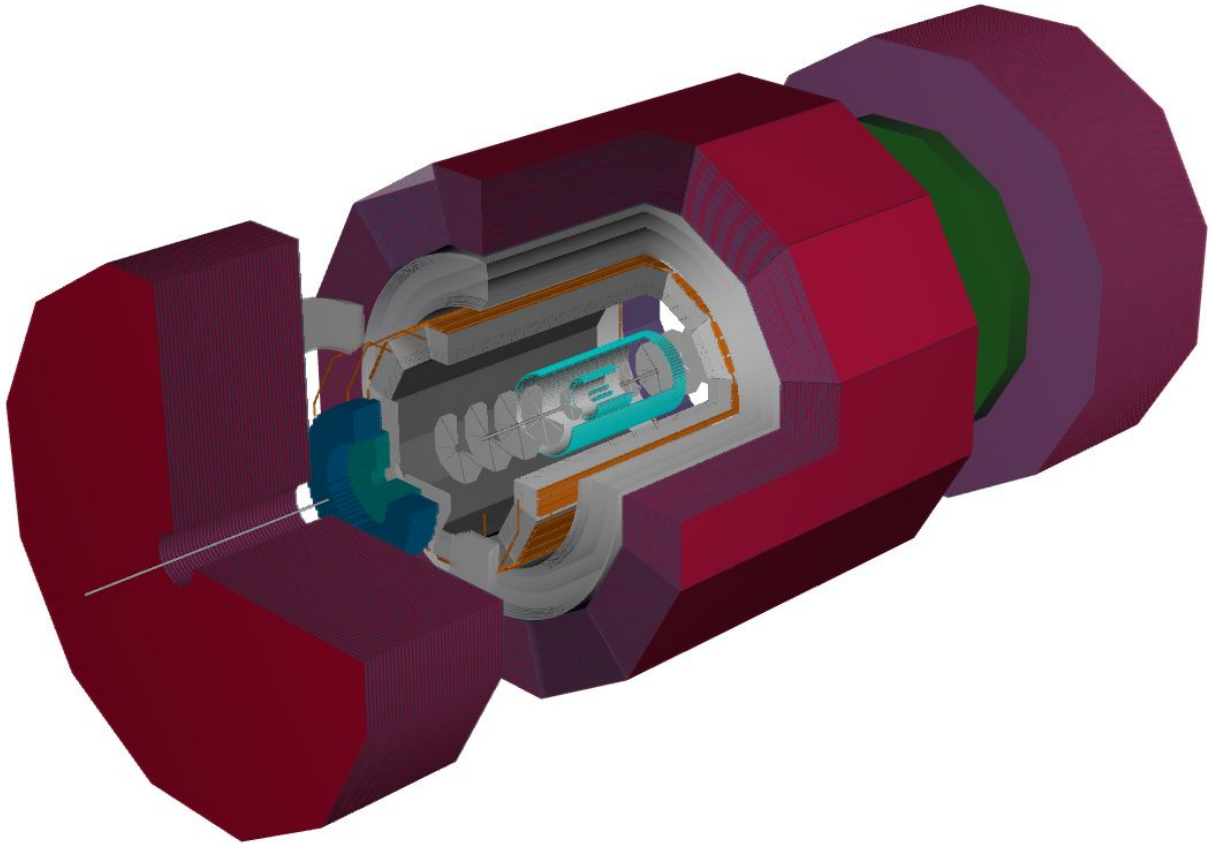






Geometry overlap checks running as part of every merge request

Automatic visualizations for detector geometries, saved as job artifacts (browsable!)

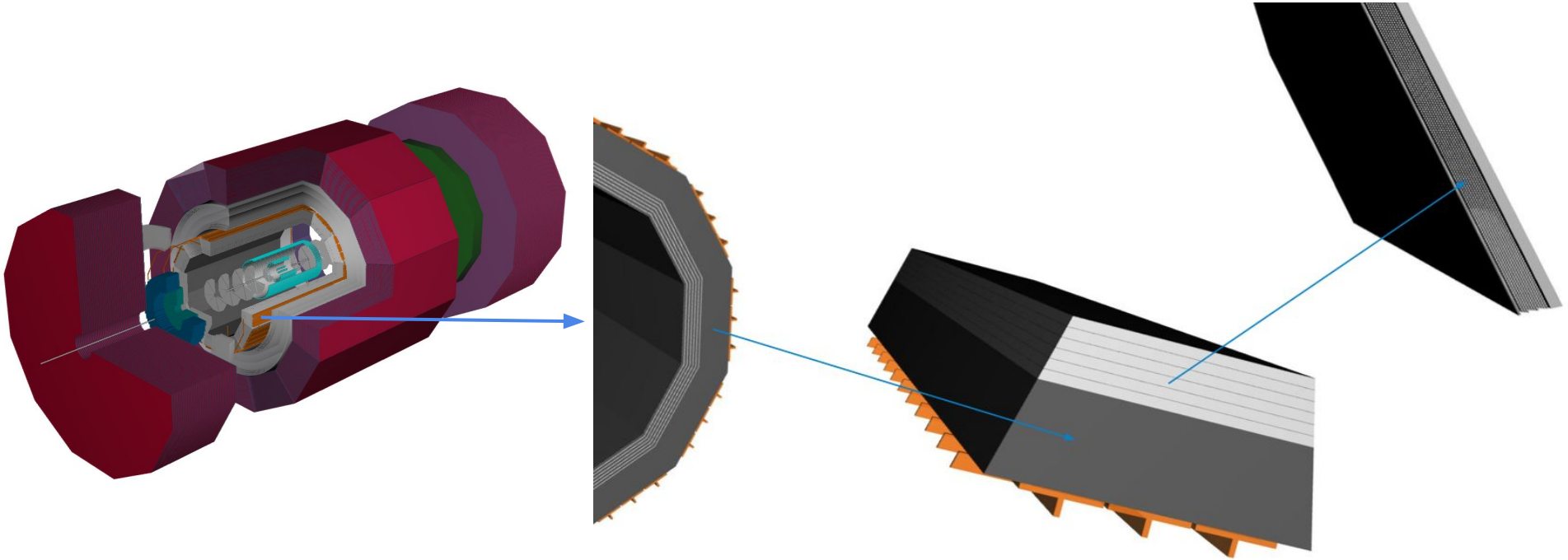
# Web Integration: GeoViewer, XRootD



Live link

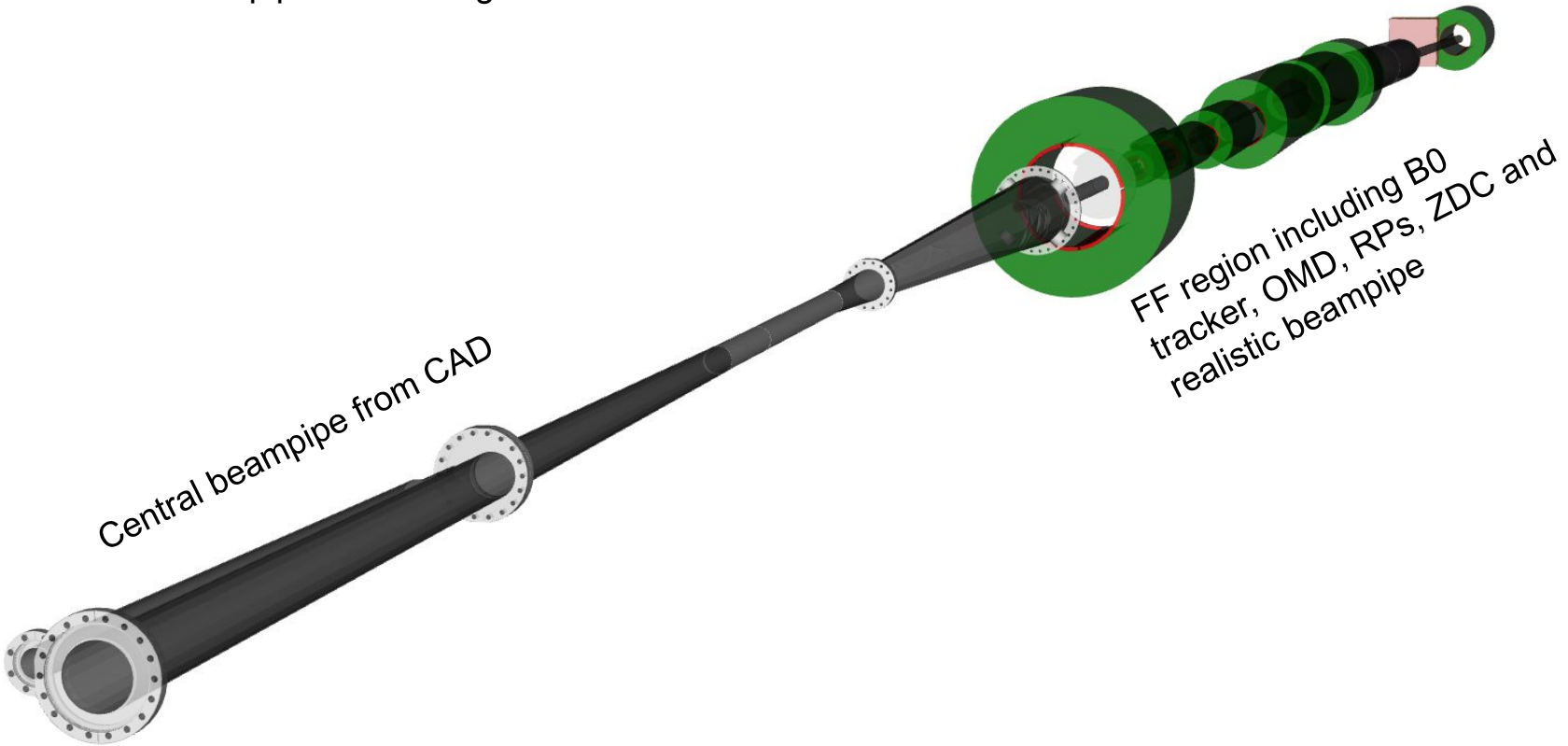
# Detailed geometry implementation

EXAMPLE: Hybrid WSi/WSciFi BECAL with support



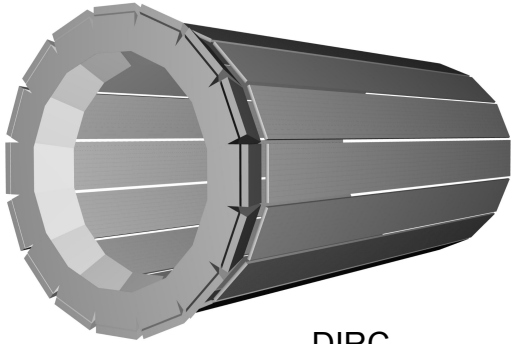
# Detailed geometry implementation

EXAMPLE: Beampipe and FF region

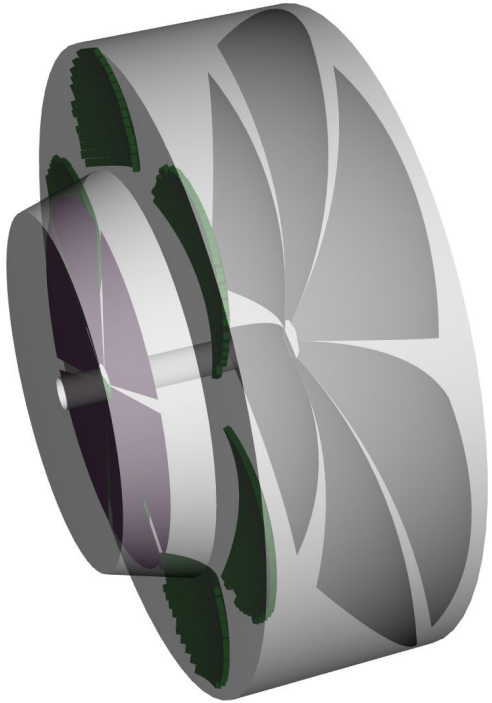


# Detailed geometry implementation

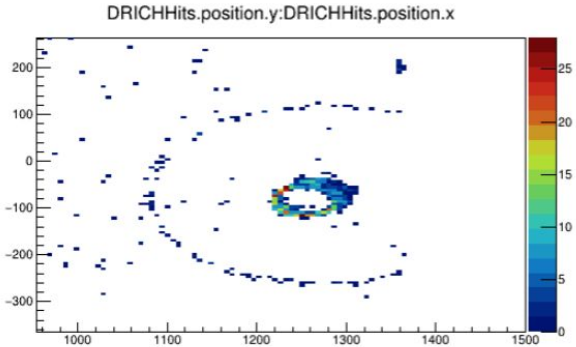
EXAMPLE: PID Systems



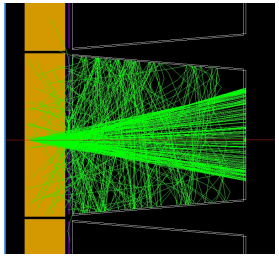
DIRC

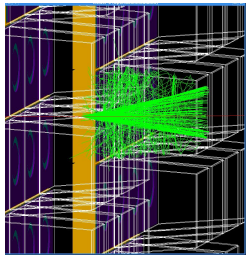


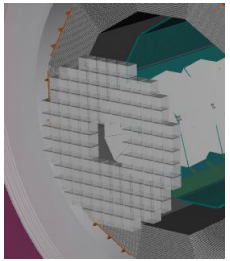
dRICH




mRICH

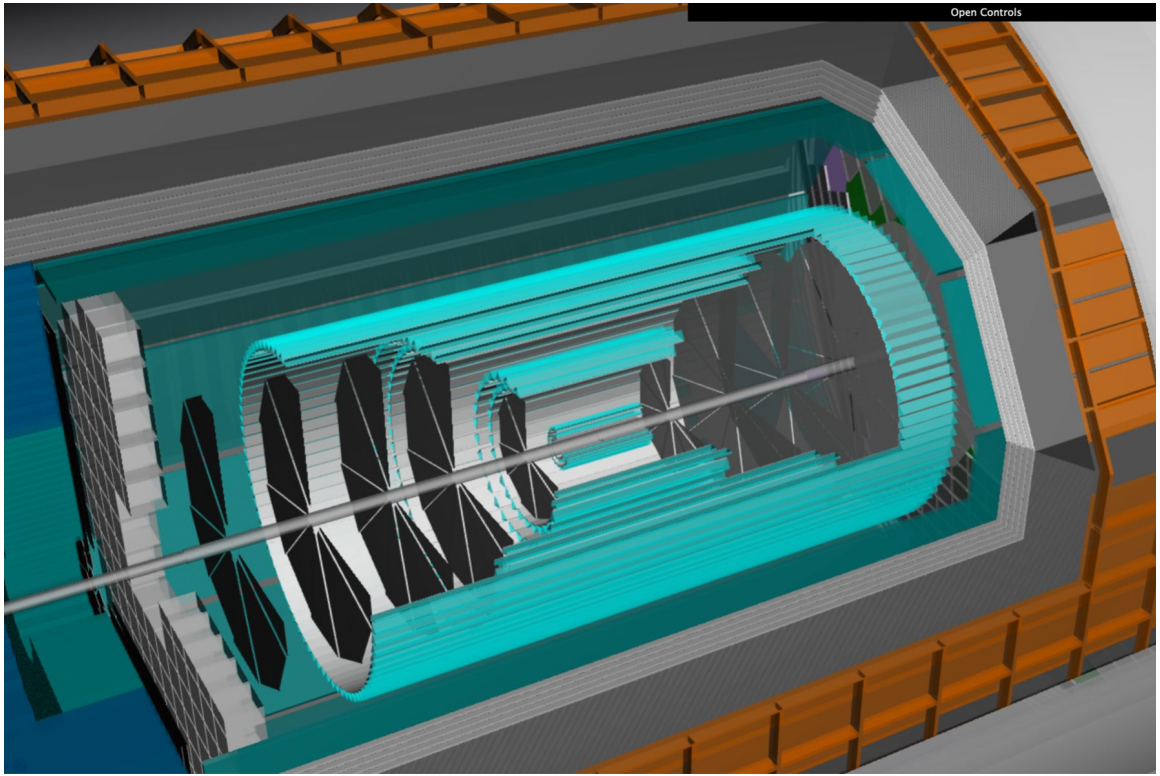






# Detailed geometry implementation

EXAMPLE: Tracking Systems



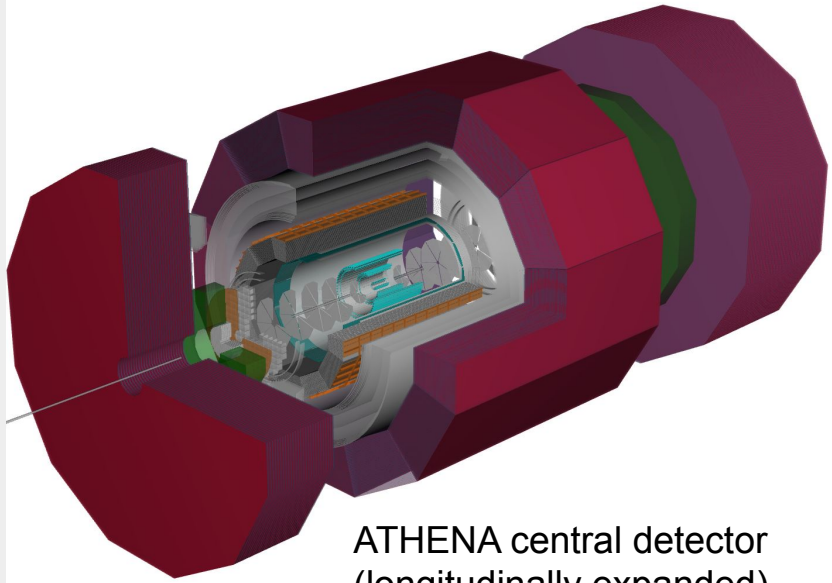
# Large scale ATHENA data productions


- Input: HepMC files preferred (`mcconv` developed for other formats)
- Full simulation with current detector model, all bells and whistles:
  - Typical: 0.25 to 3.0 s/event, <500MB RAM RSS, 30 kB to 750 kB output size/event
  - Full ROOT files on S3 under `ATHENA/FULL/` (but likely only need reco files)
- Full reconstruction (reconstruction_benchmarks/benchmarks/full):
  - Calorimetry clustering (Ecal, ScFi, Hcal), tracking (up to inner tracker), RICH hits/digi
  - Reco ROOT files on S3 under `ATHENA/RECO/` and `sci-xrootd.jlab.org`
    - Working on jsroot and file browser support on `sci-xrootd`
- Full simulation: ~weekly repetition; reconstruction: every few days
  - So far simulated 285M events (excluding the CI system!)
- Written to work on any slurm batch system; performed at Compute Canada
  - Trial runs on OSG at the ~2k job scale for single particle events
  - Trial runs on HPC (THETA@ALCF)
- **Will start larger-scale campaign to study detector variations in 3 days!**

# EIC AI/ML in ATHENA

- Current use of AI/ML
  - ✅ e/π PID with 3D shower profiles from imaging calorimeter in center barrel region.
- Near-term anticipated use:
  - 🚧 ACTS: Track finding
  - 🚧 PID: Pattern recognition in RICH, DIRC
  - ✅ Calorimetry clustering (2D, 2+1D and 3D clustering)
  - 🚧 DNN-based fast simulation
  - 🚧 DNN-based detector optimization (Bayesian optimization)
  - 🚧 DNN-based reconstruction
- Implications on computing infrastructure:
  - 🚧 Many exascale GPU accelerators, but lack of support in current software tools limited by IO/memory bandwidth

✅ = working, 🚧 = in progress, ❌ = planned



ATHENA central detector (longitudinally expanded)

# Software & Computing WG

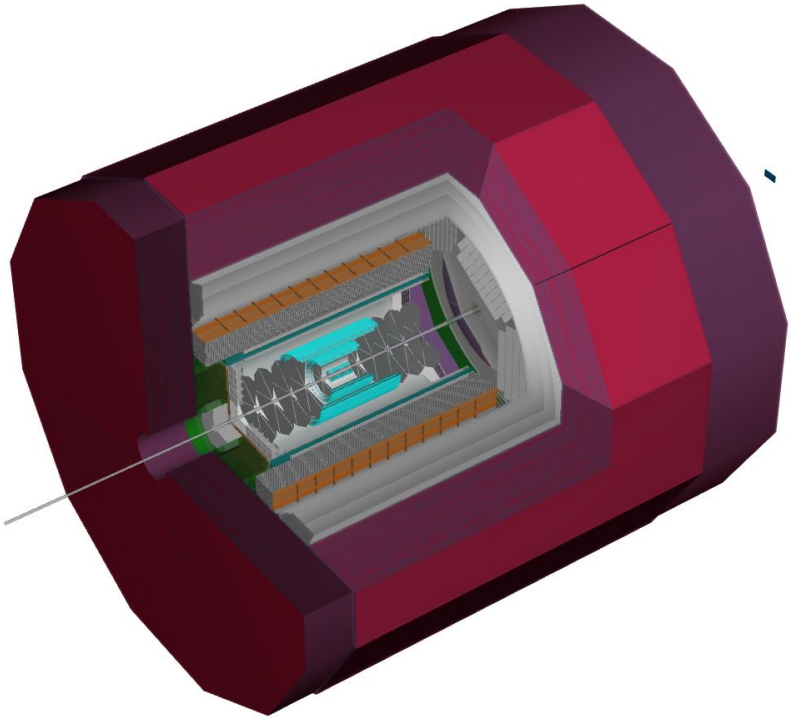
**Software & Computing Conveners**:
Whitney Armstrong, Andrea Bressan(*), Wouter Deconinck, Sylvester Joosten, Dmitry Romanov

(*)- liaison to EICUG software group

Documentation portal:
doc.athena-eic.org

Full simulation tutorials

eic-ip6-software-l@lists.bnl.gov
**#software-helpdesk** on Slack



ATHENA central detector

*Bi-weekly software meeting: Thursday 12:00pm EDT*