# Generative Models for Particle Shower Simulation in Calorimeters for Geant4

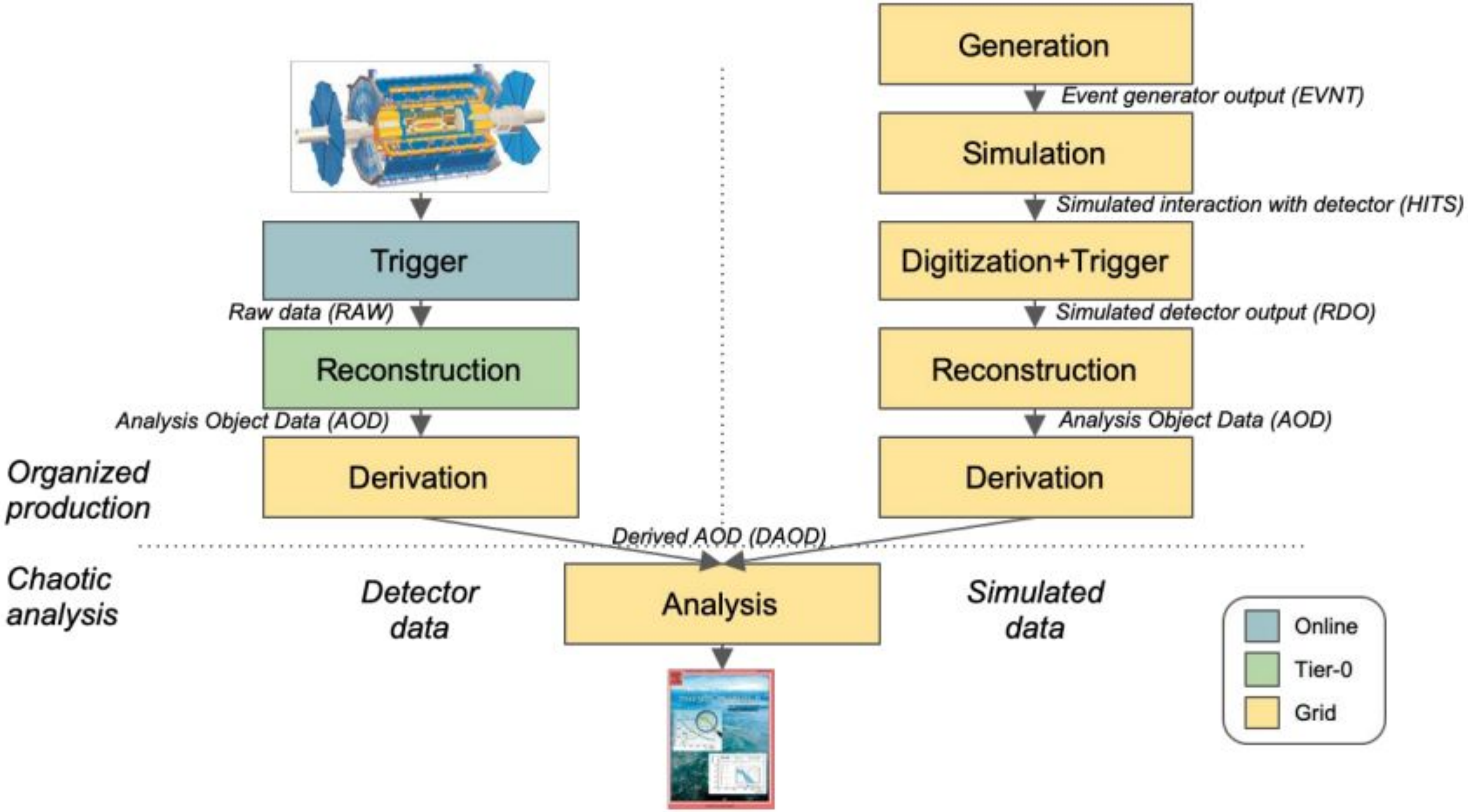
Mohammed Touami **(EP - SFT)**

**Supervisors:** Piyush Raikwar, Peter McKeown, Anna Zaborowska, Dalila Salamani

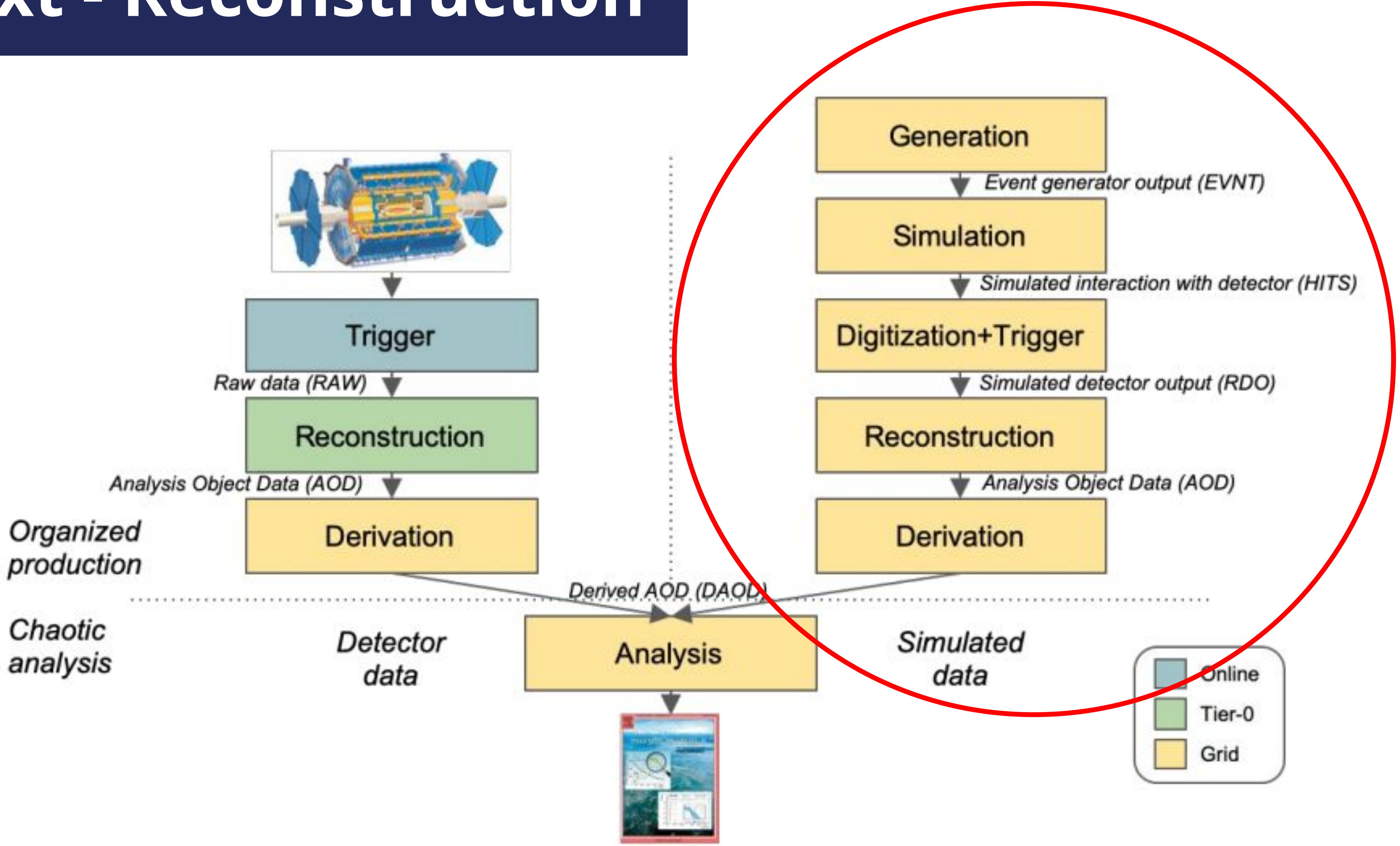
CERN openlab

# Context

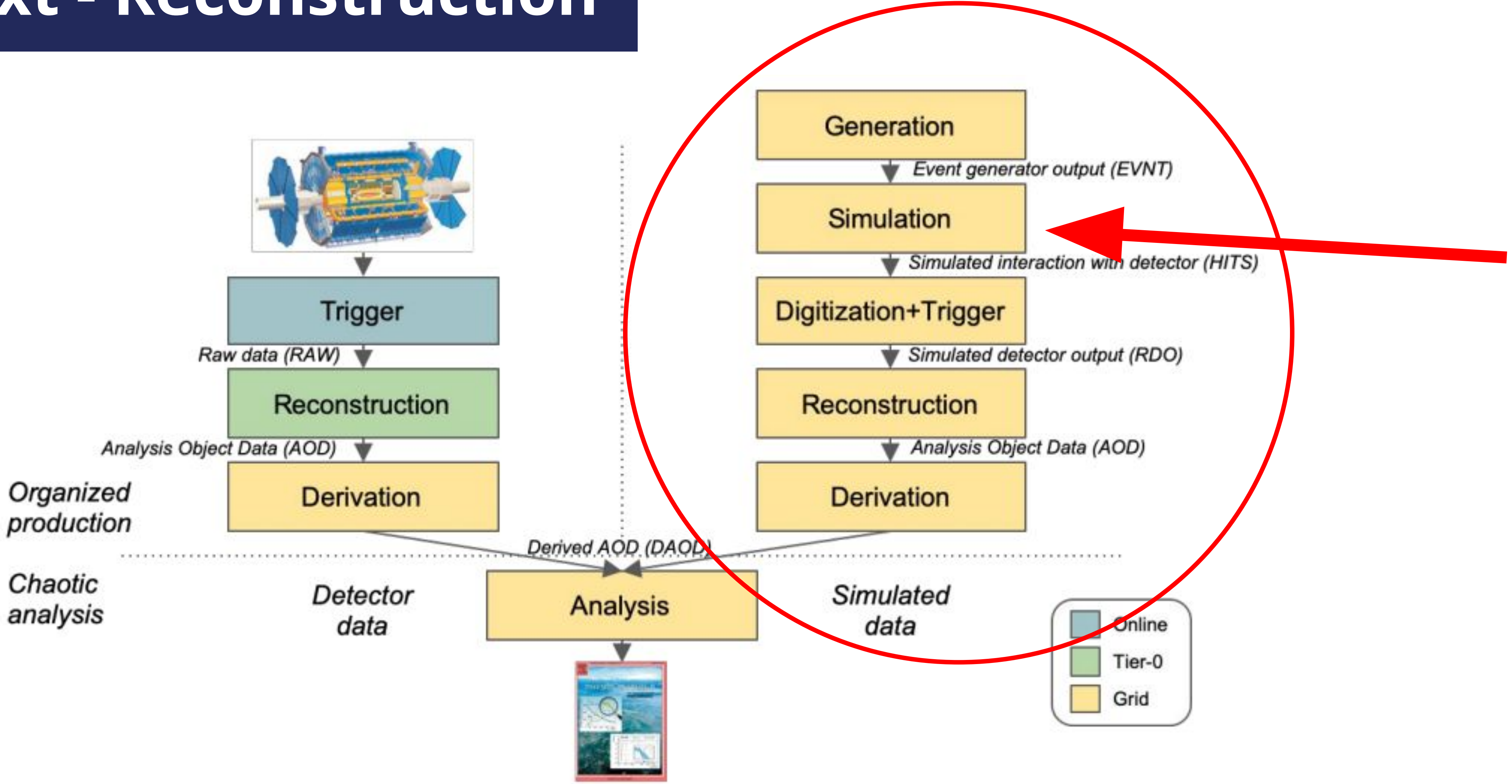
# Context - Reconstruction



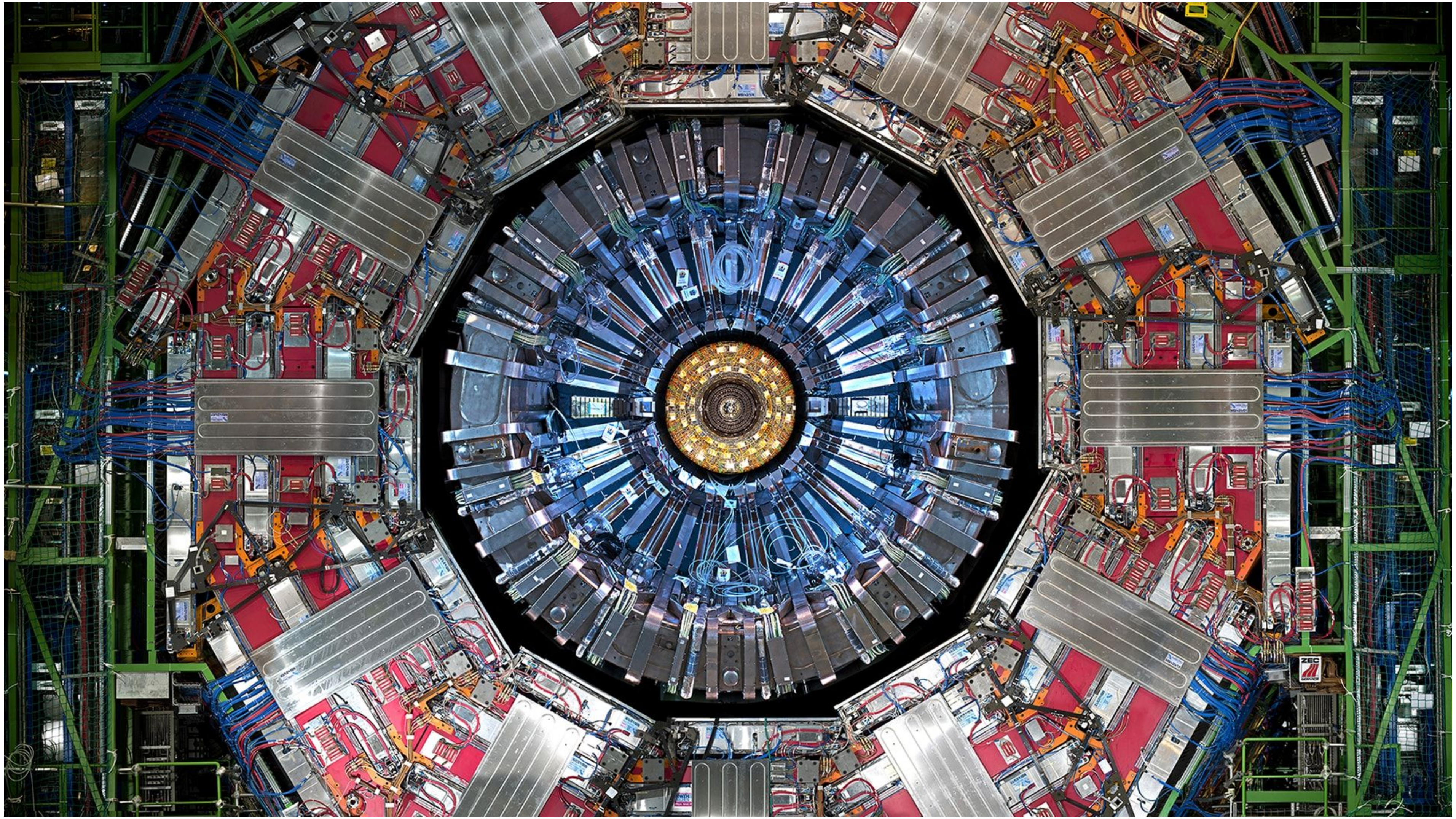
# Context - Reconstruction



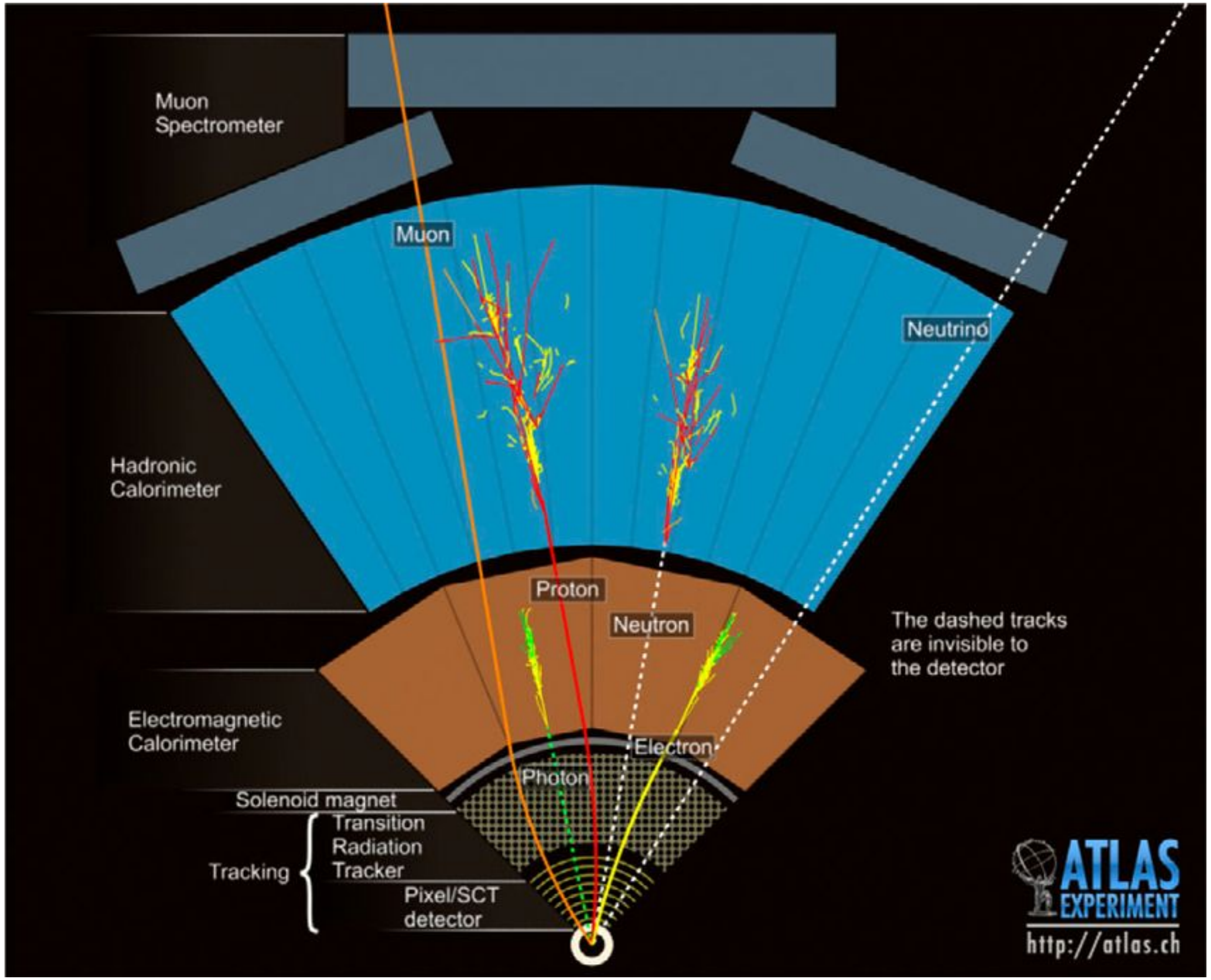
# Context - Reconstruction



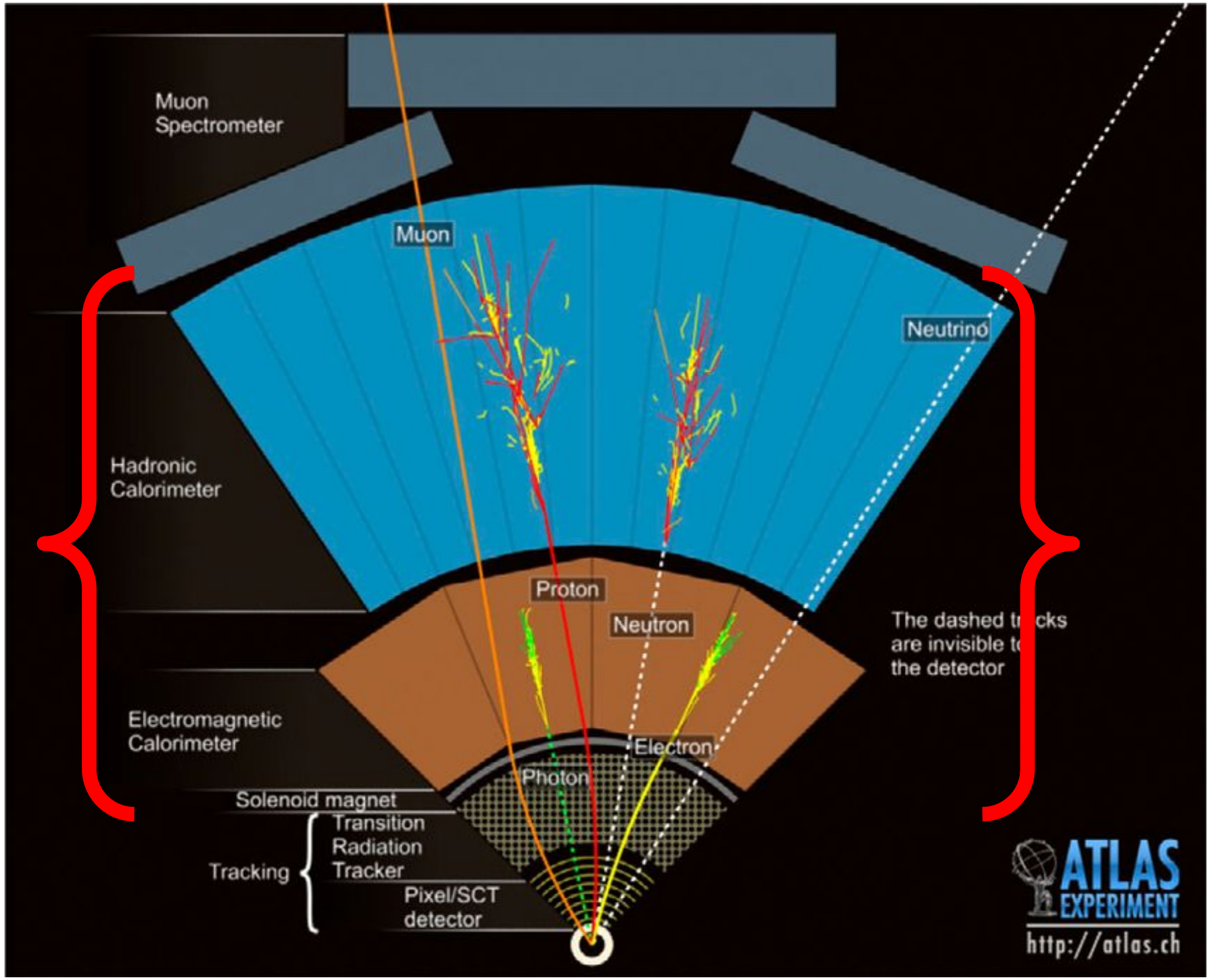
# Context - Simulation



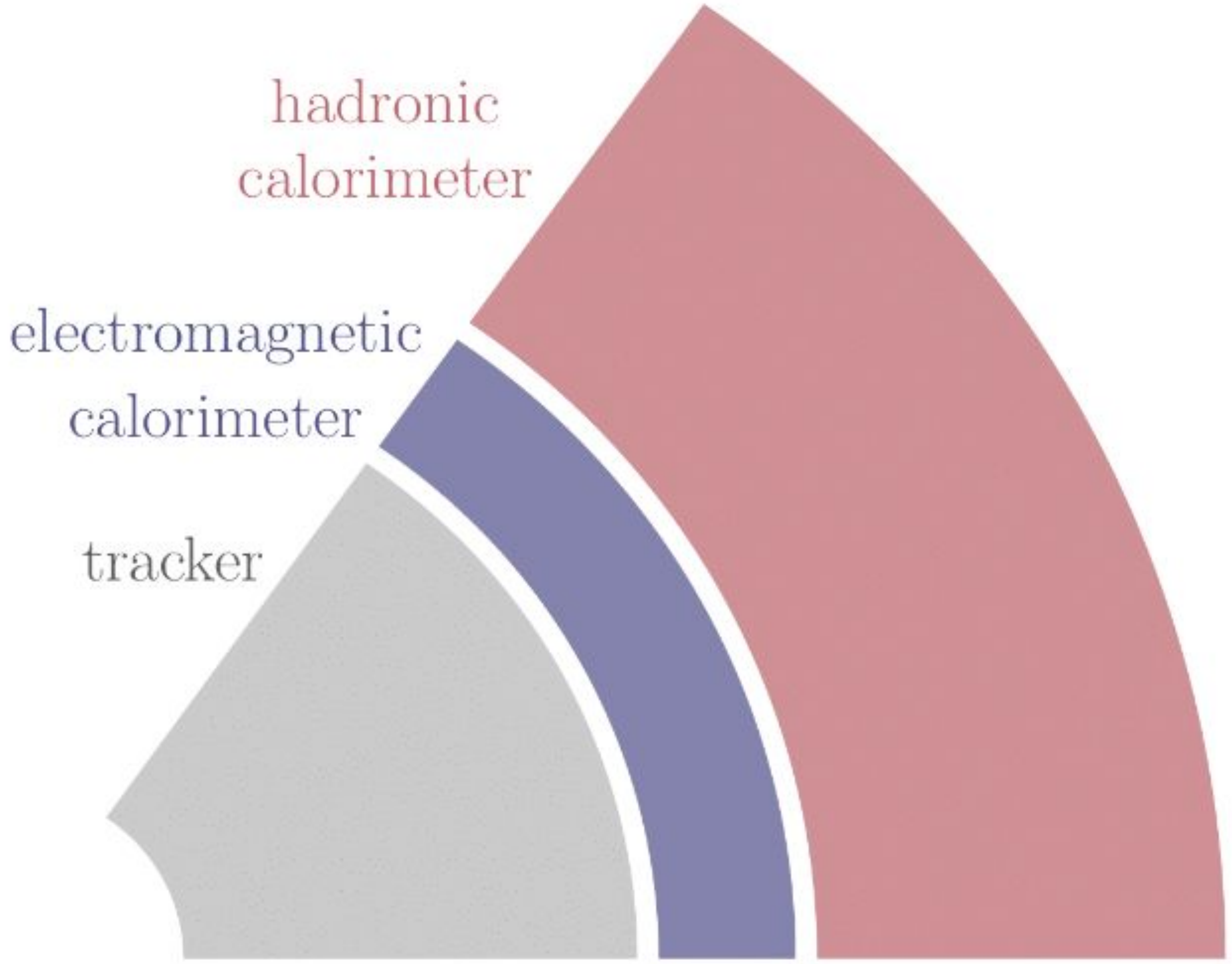
# Context - Simulation



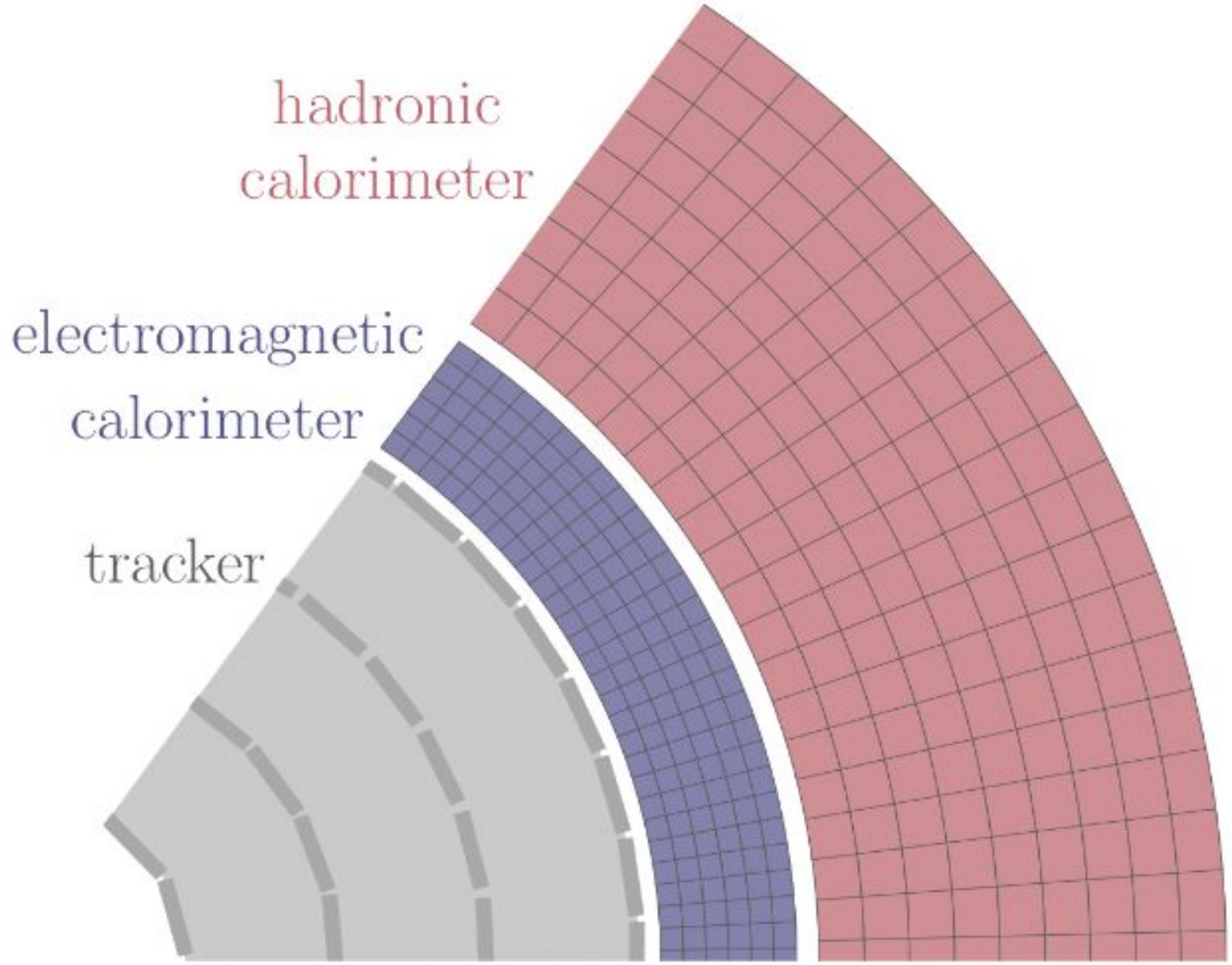
# Context - Simulation



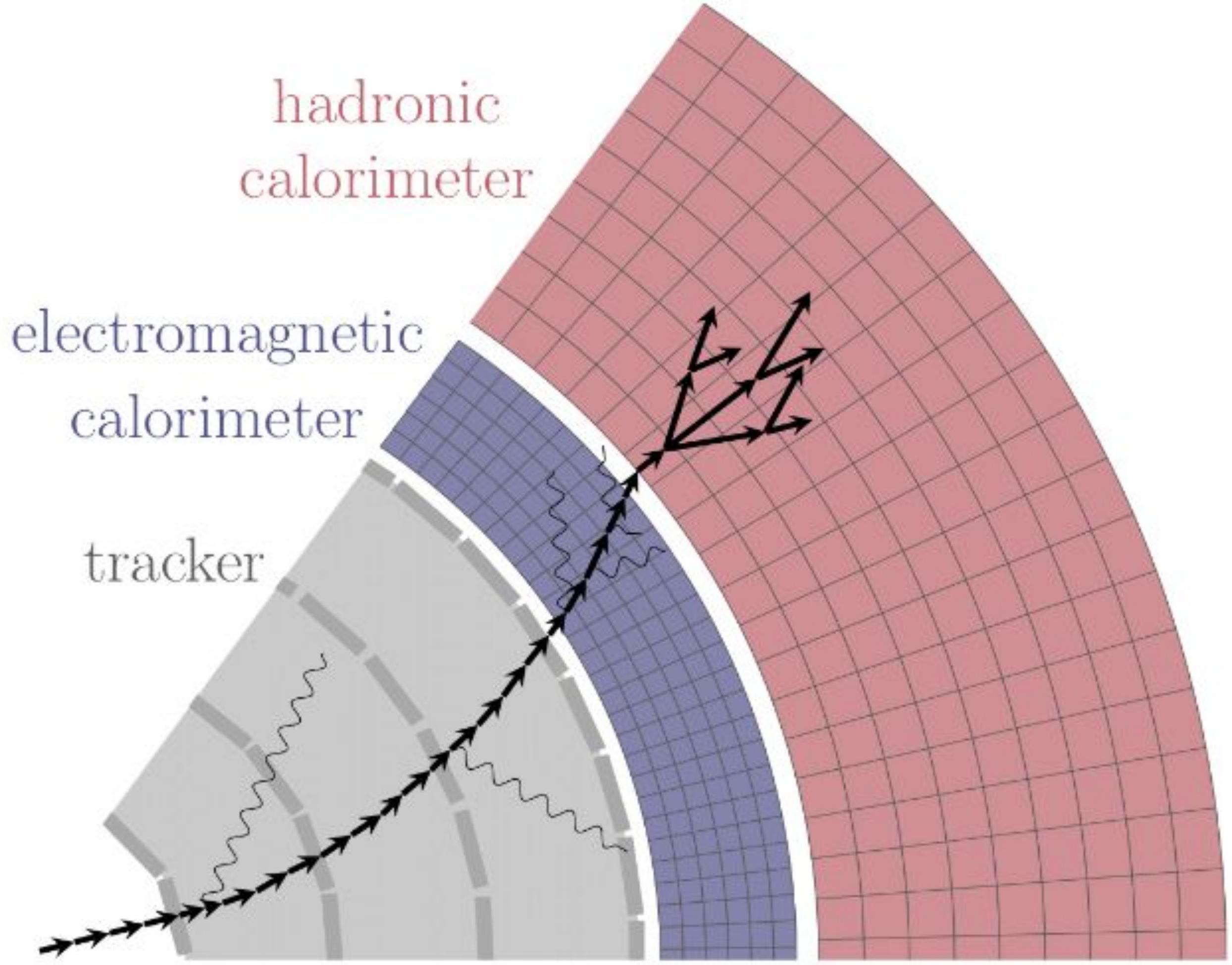
# Context - Simulation



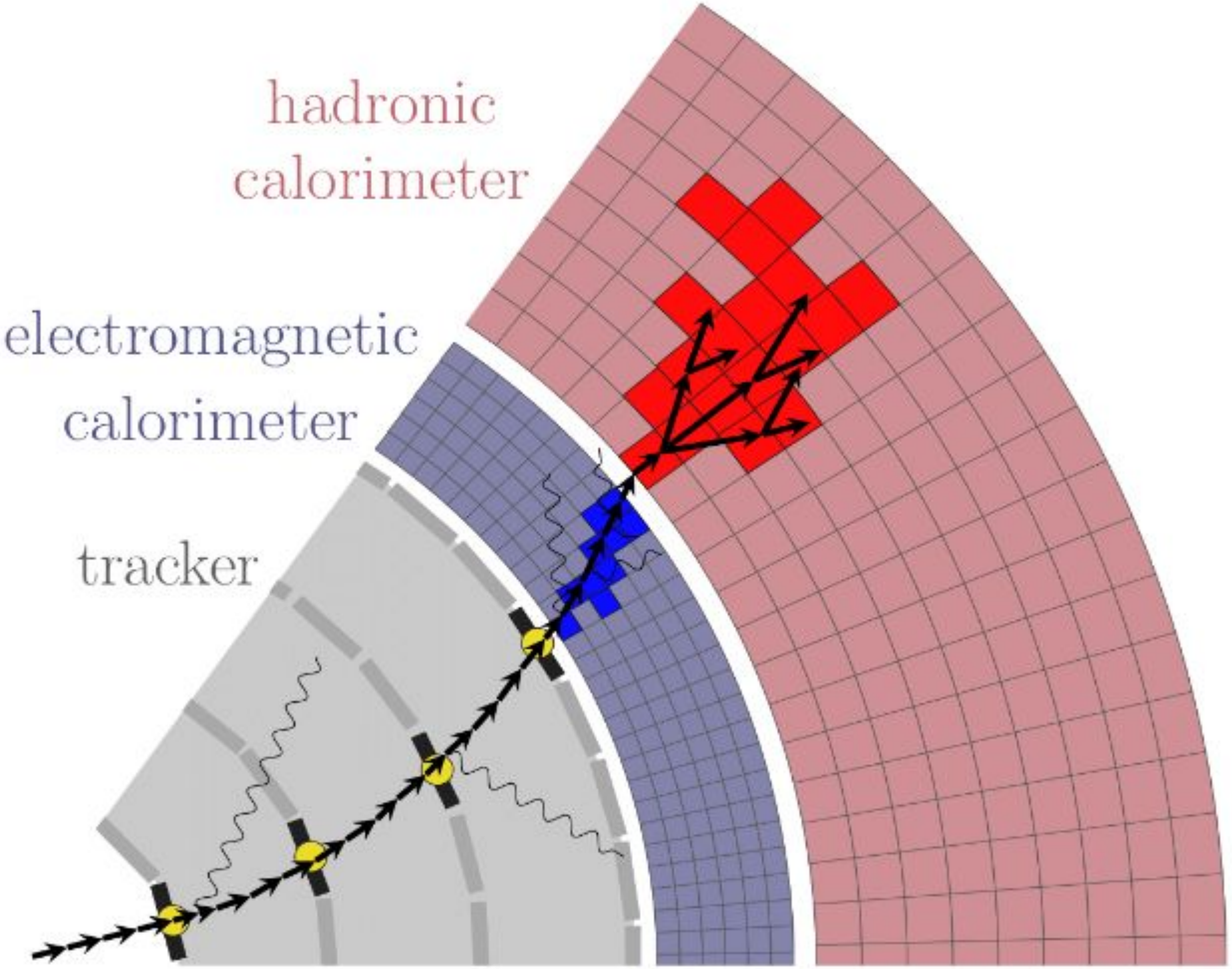
# Context - Simulation



# Context - Simulation



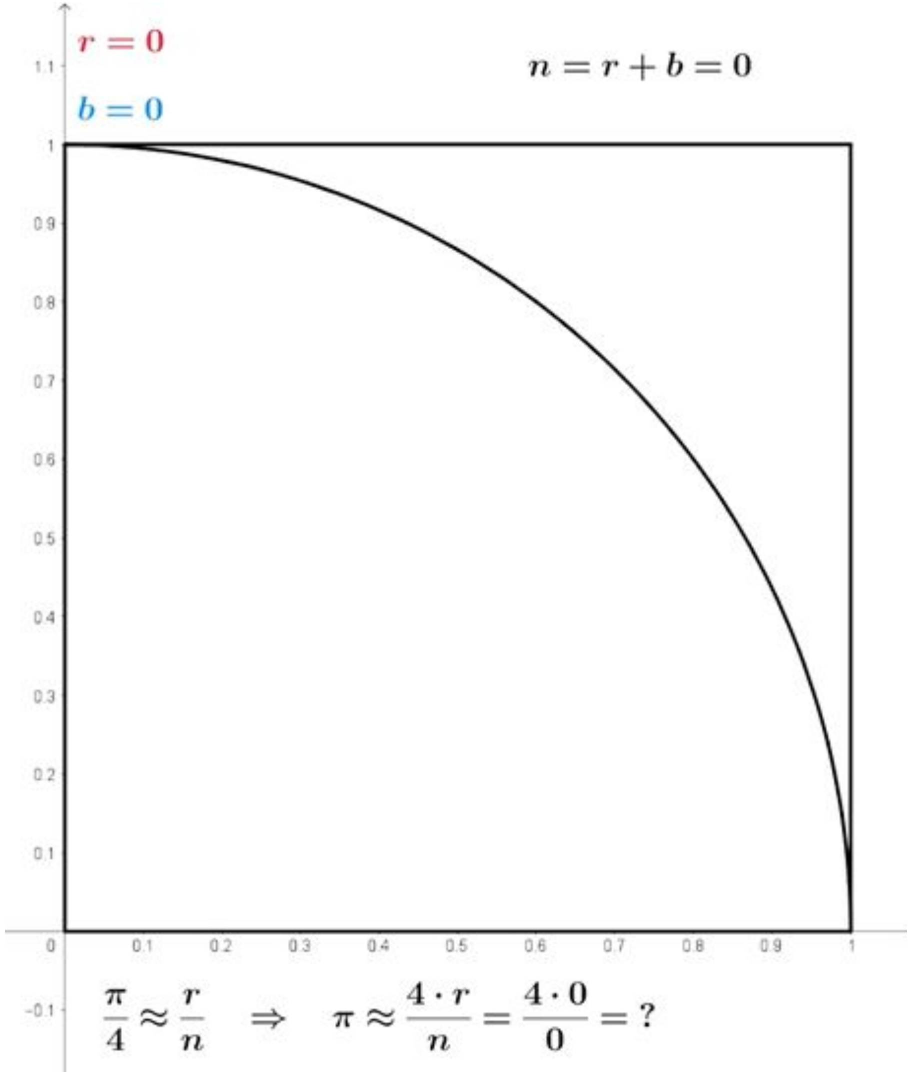
# Context - Simulation



# Context - Geant4

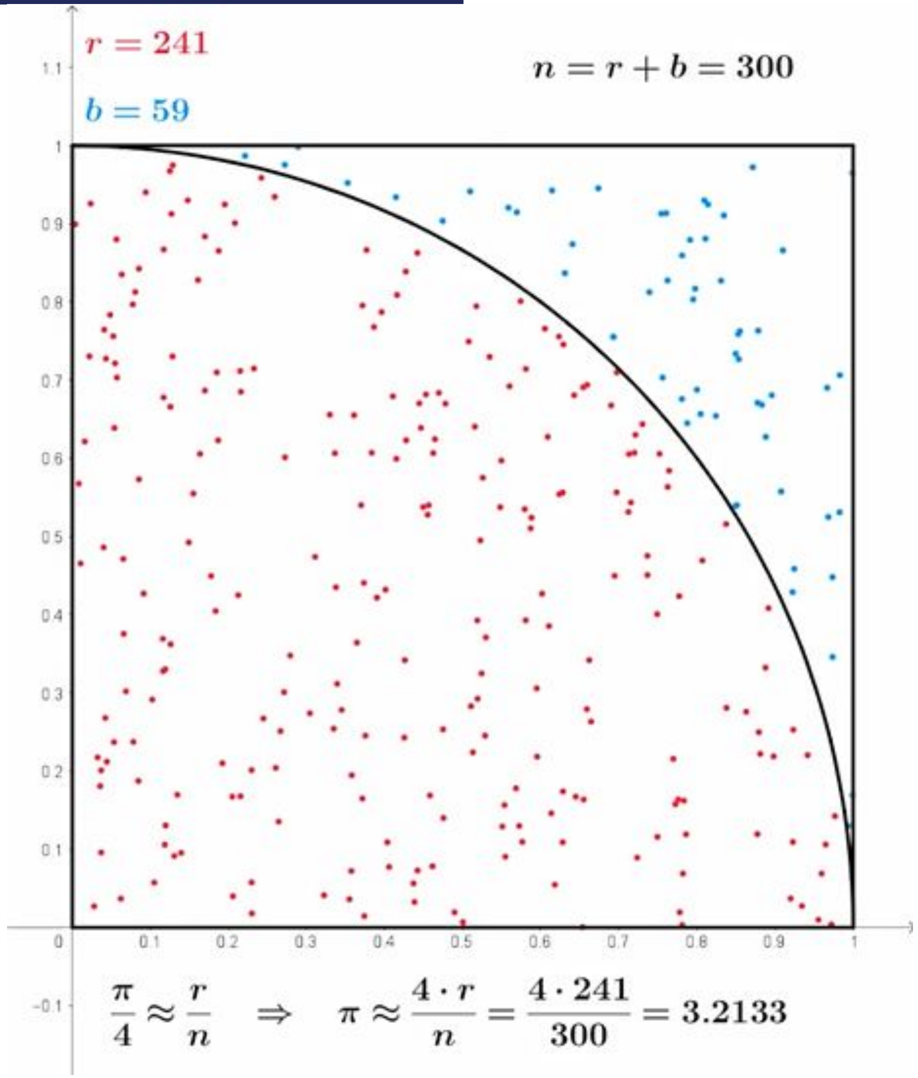
- Toolkit for the simulation of the passage of particles through matter (Open Source, C++).
- Allows creation of geometries, custom detectors, record particle hits & tracks…
- Its areas of application also include high energy, nuclear and accelerator physics, as well as studies in medical and space science.

# Monte Carlo



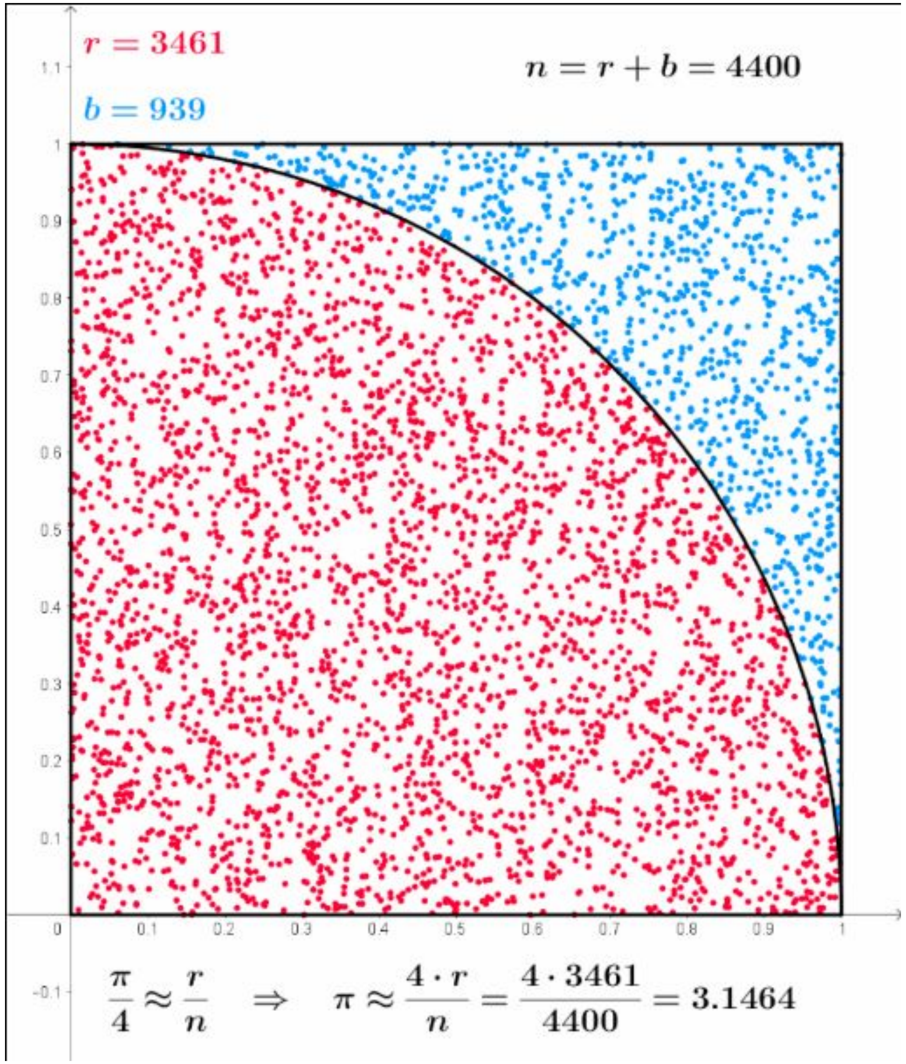
$r = 0$
$b = 0$
$n = r + b = 0$
1.1
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
-0.1
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
$\frac{\pi}{4} \approx \frac{r}{n} \quad \Rightarrow \quad \pi \approx \frac{4 \cdot r}{n} = \frac{4 \cdot 0}{0} = ?$

# Monte Carlo



$r = 241$
$b = 59$
$n = r + b = 300$
$\frac{\pi}{4} \approx \frac{r}{n} \Rightarrow \pi \approx \frac{4 \cdot r}{n} = \frac{4 \cdot 241}{300} = 3.2133$

# Monte Carlo

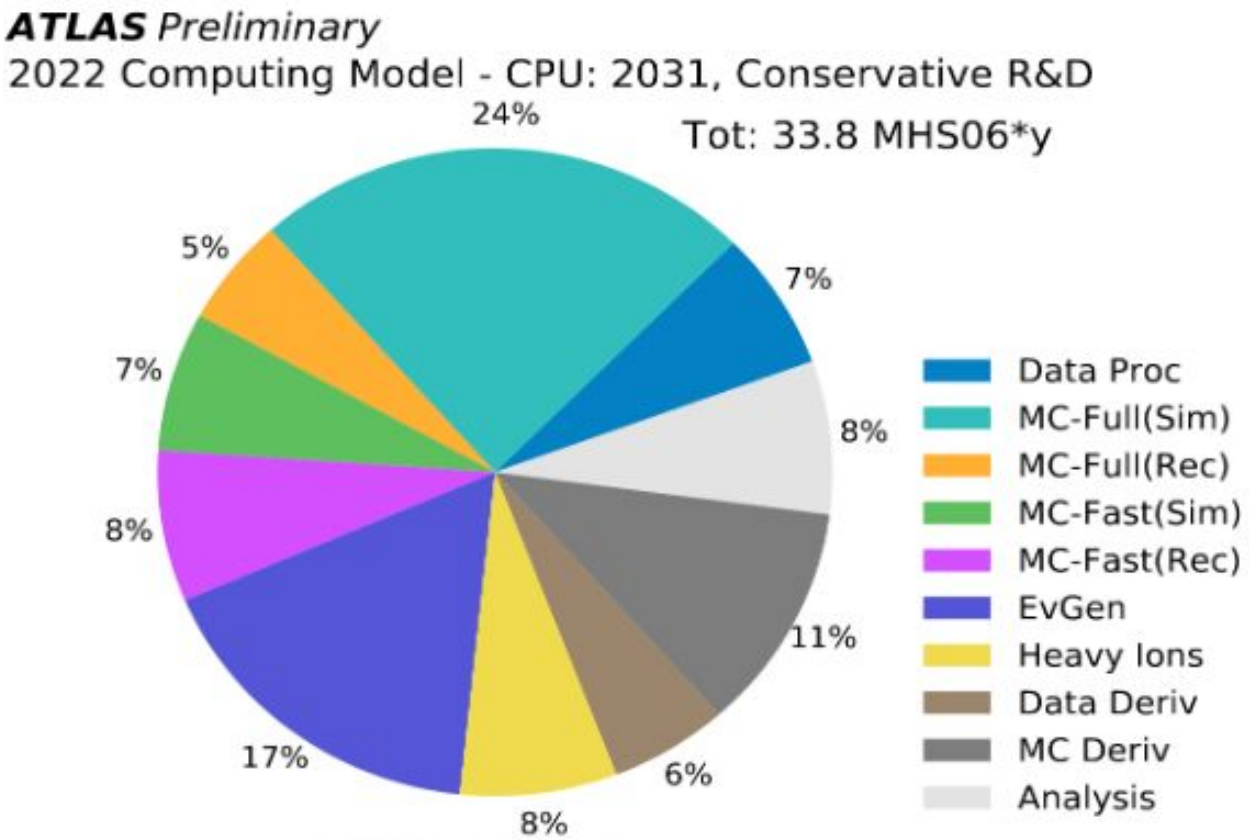


$r = 3461$
$n = r + b = 4400$
$b = 939$
$\frac{\pi}{4} \approx \frac{r}{n} \Rightarrow \pi \approx \frac{4 \cdot r}{n} = \frac{4 \cdot 3461}{4400} = 3.1464$

# Limiations

More Data = More Accuracy = More Resources

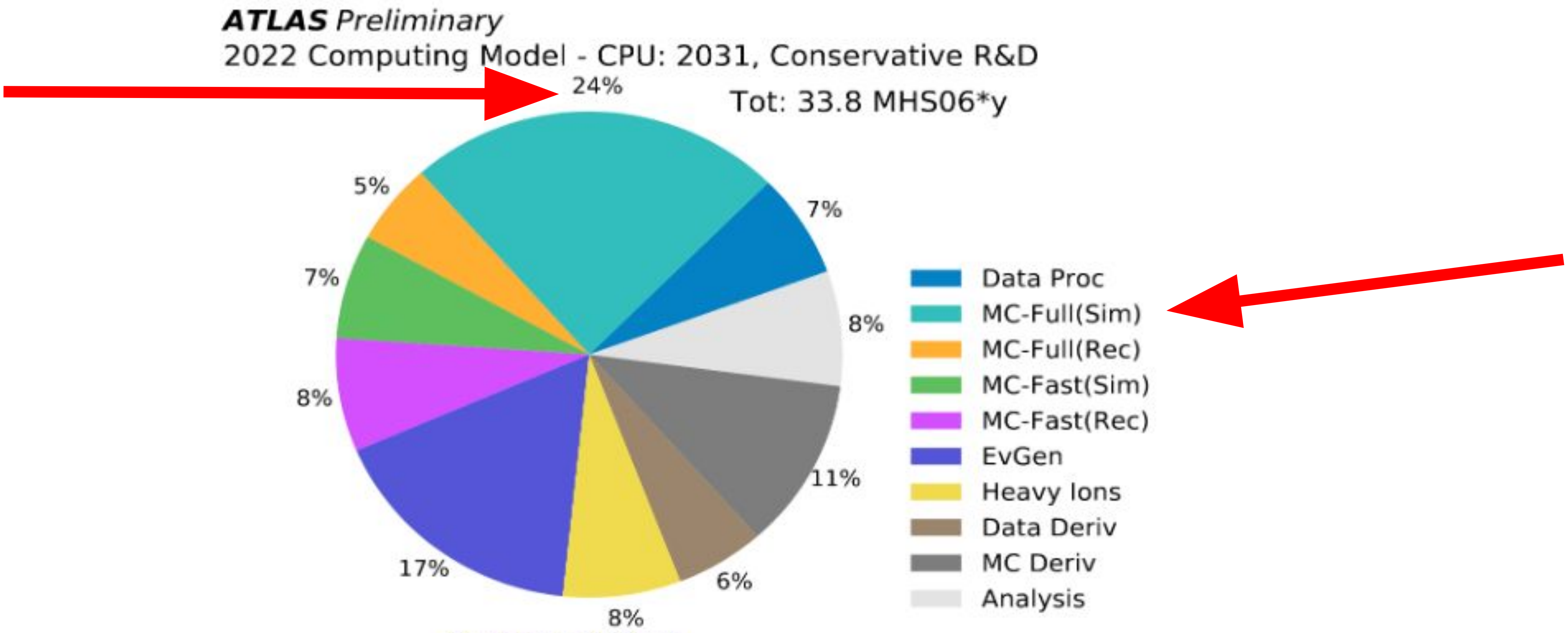
# Simulation Cost




CERN-LHCC-2022-005

# Simulation Cost

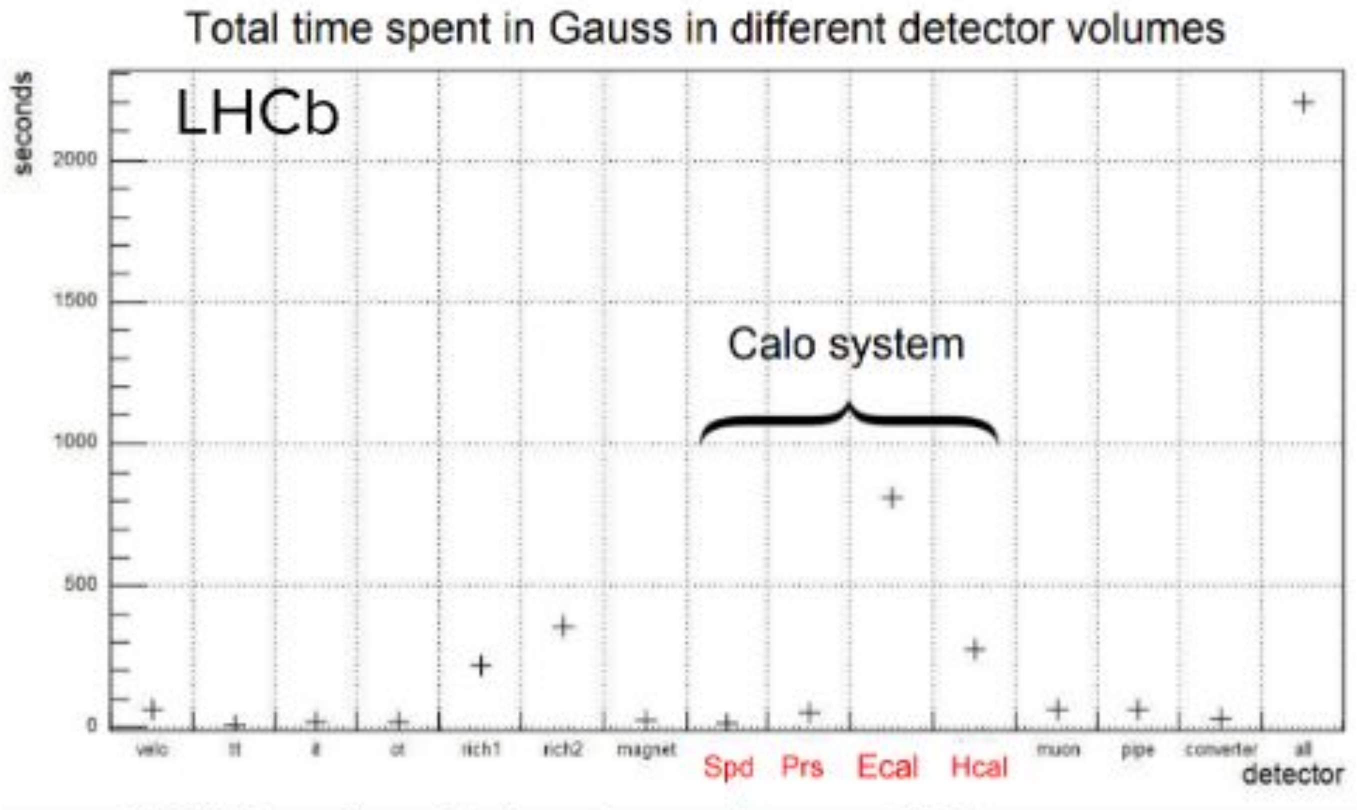
**One quarter of CPU usage!**




CERN-LHCC-2022-005

# Calorimeter Cost

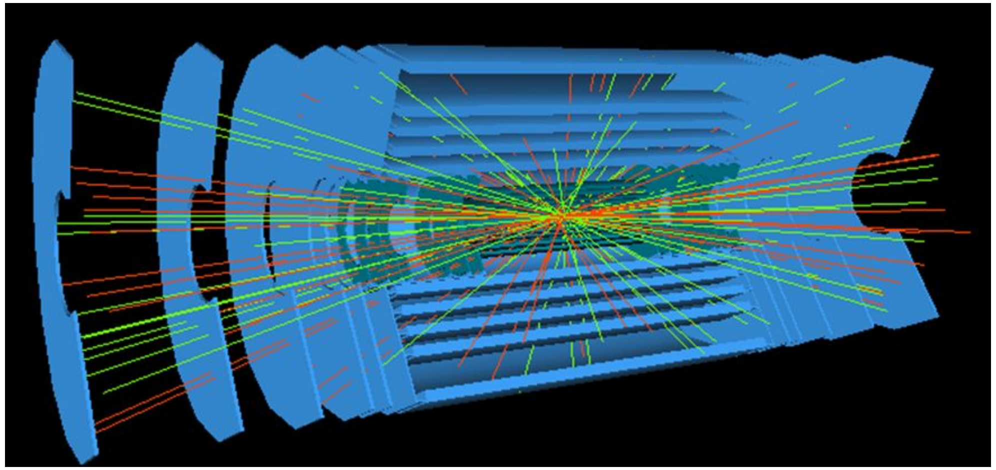
**One half of CPU usage!**




CPU time in calorimeter system: ~ **53%**

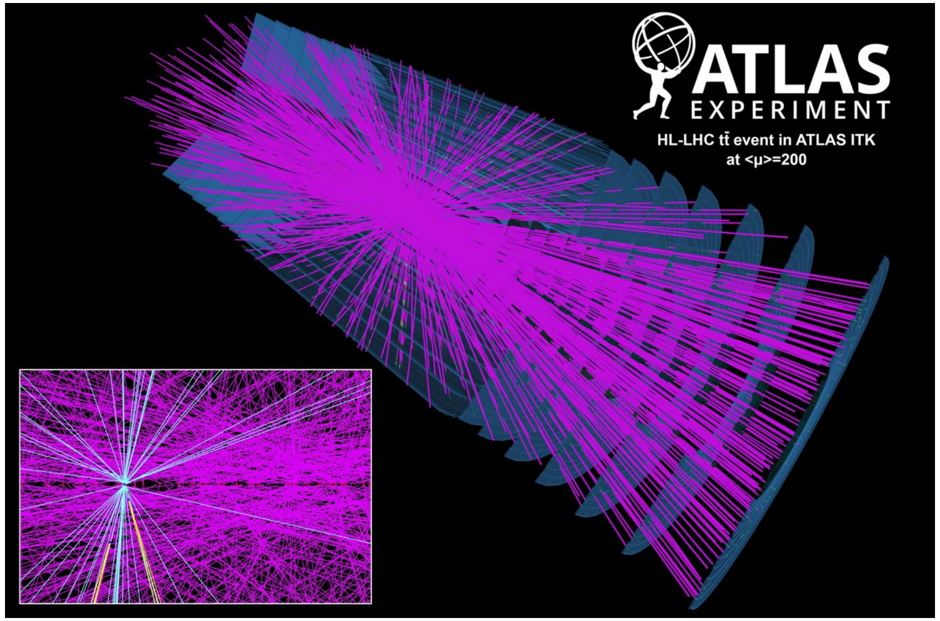
CHEP 2018, M. Rama

# Limitations

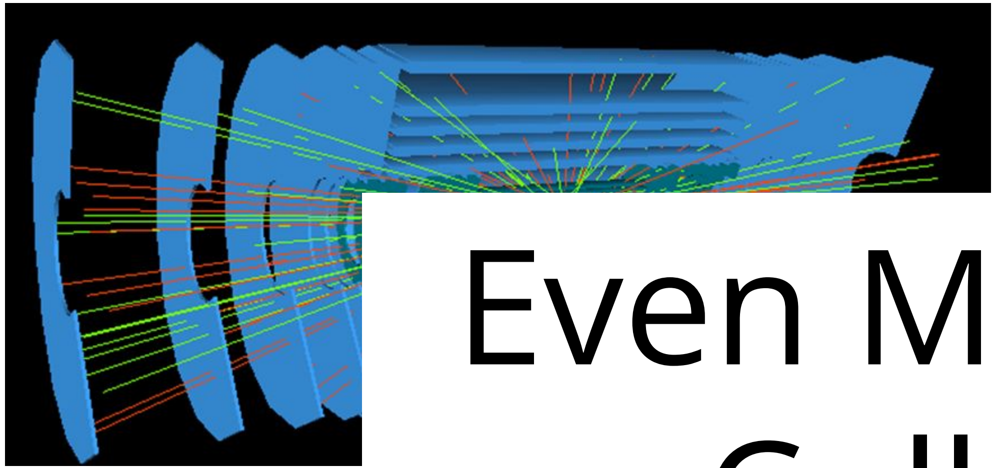


LHC - 2024

HL LHC - 2029



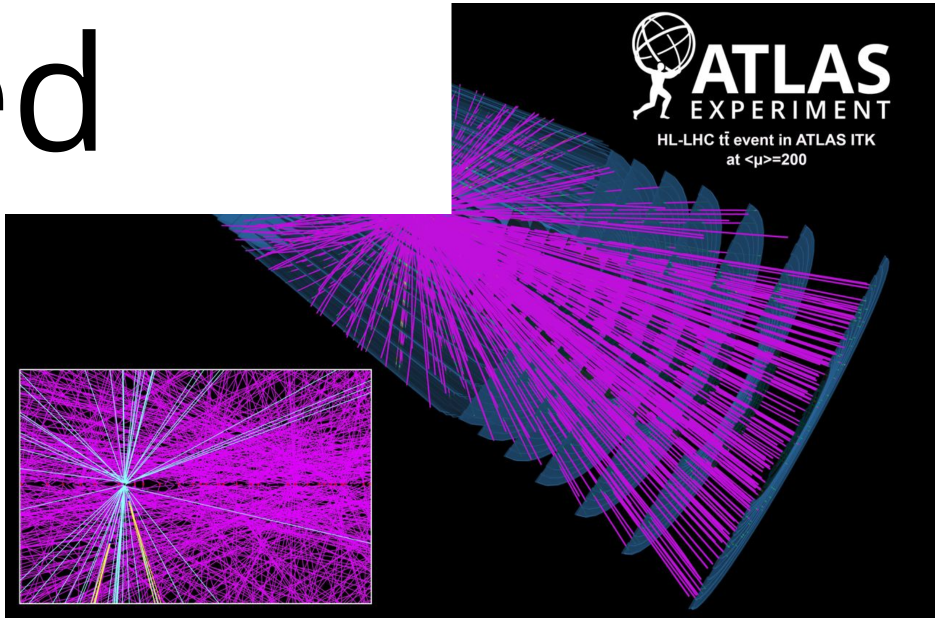
# Limitations



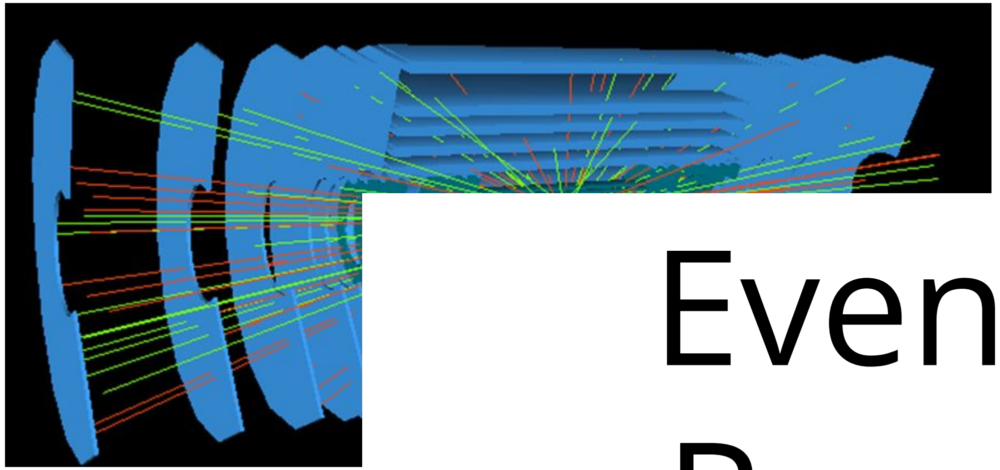
# Even More Data Collected

LH

HC - 2029



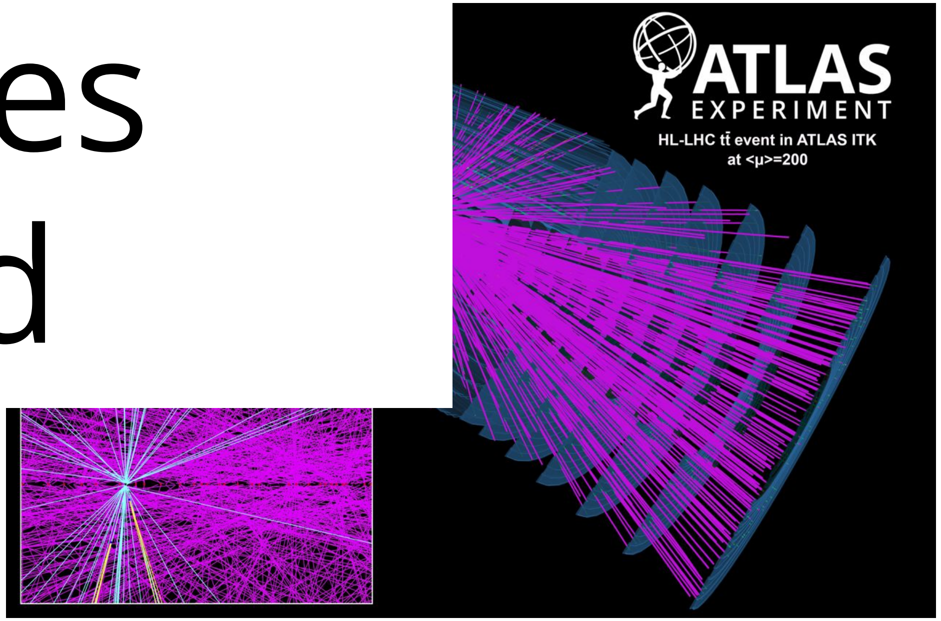
# Limitations



LH

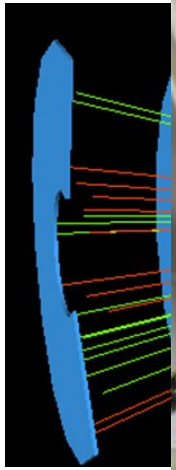
Even More Resources Needed

HC - 2029



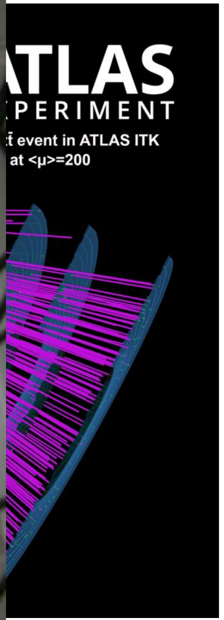
# Limitations

- 2029




Adobe Stock | #427228267

9


TLAS
PERIMENT
event in ATLAS ITK
at <μ>=200

# Any Alternatives?

# Solution

Software Efficiency Enhancement

# Solution

Software Efficiency Enhancement

Machine Learning

# Solution

Software Efficiency Enhancement

Machine Learning

Generative Models

# Solution

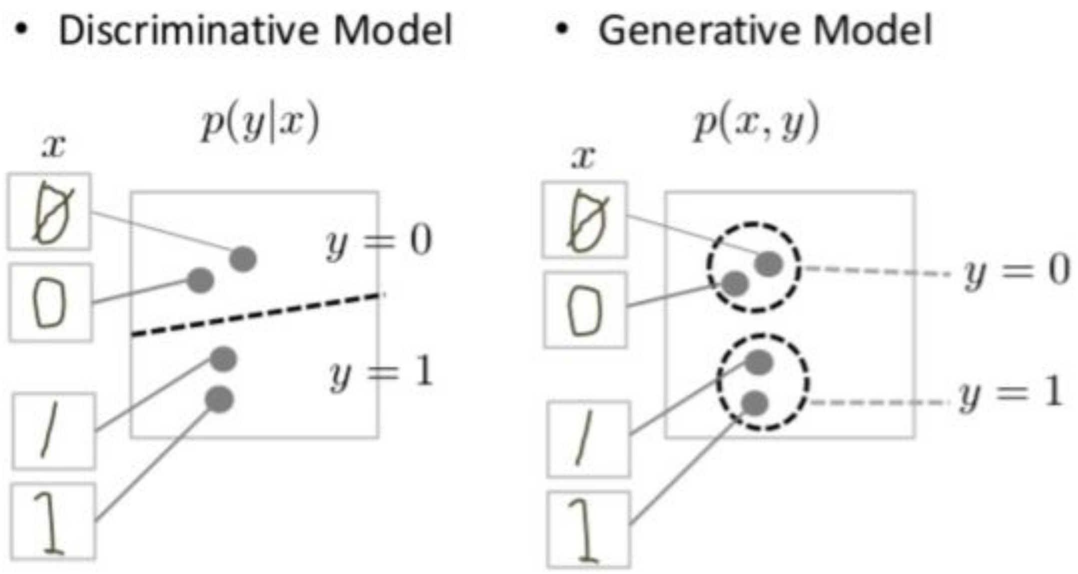
Software Efficiency Enhancement

Machine Learning

Generative Models

Particle Shower Inference

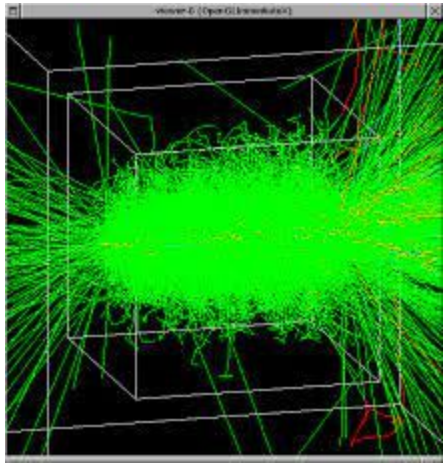
# Generative Models




Difference between standard ML classifiers:

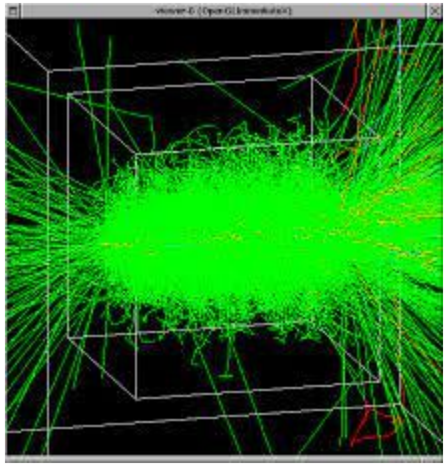
- Give the model a standard prior (usually Gaussian)
- Give the model the outputs (classes)
- Model tries to mimic an input that would yield such a class

# Gen AI in HEP

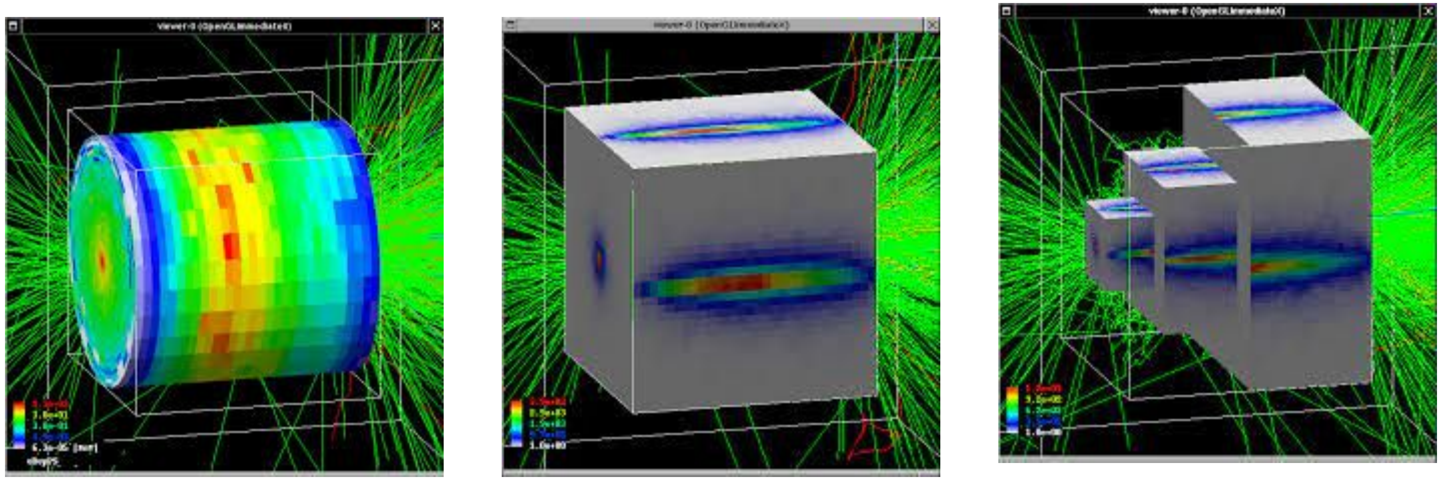


Simulated Collision in Detector

# Gen AI in HEP



Simulated Collision in Detector



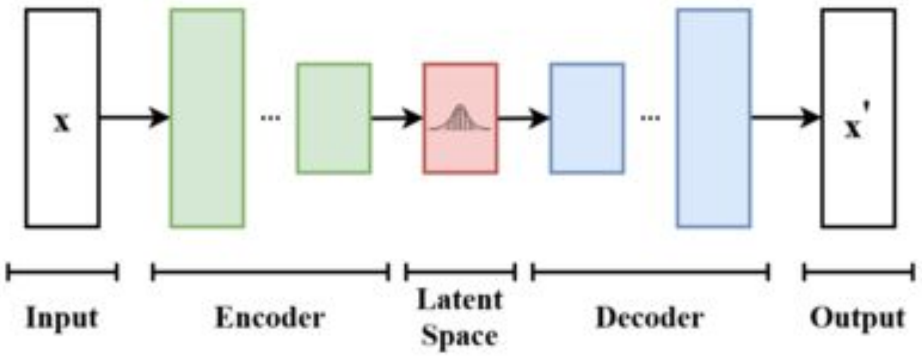
Recorded Scoring Mesh

# Gen AI in HEP

$$\texttt{def ML4FastSim}(\texttt{particle}_{(\texttt{energy,angle})}, \texttt{detector}):$$

$$\texttt{return } f(\texttt{shower} \mid \texttt{particle}_{(\texttt{energy,angle})}, \texttt{detector})$$

# Gen AI in HEP



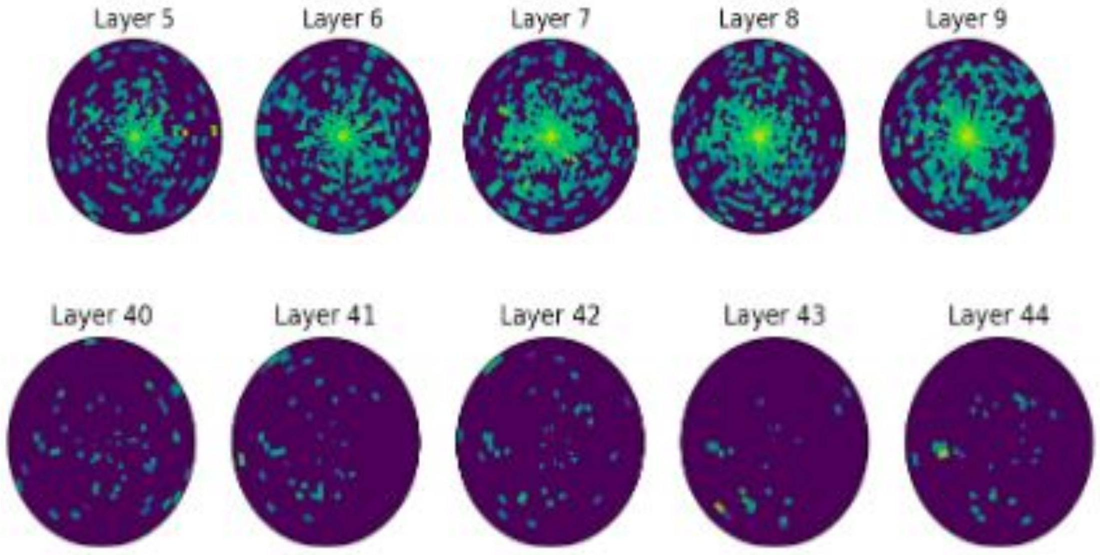

def ML4FastSim(particle$_{(energy,angle)}$, detector):

return f(shower|particle$_{(energy,angle)}$, detector)

The returned shower (3D Scoring Mesh) is used as final simulation output

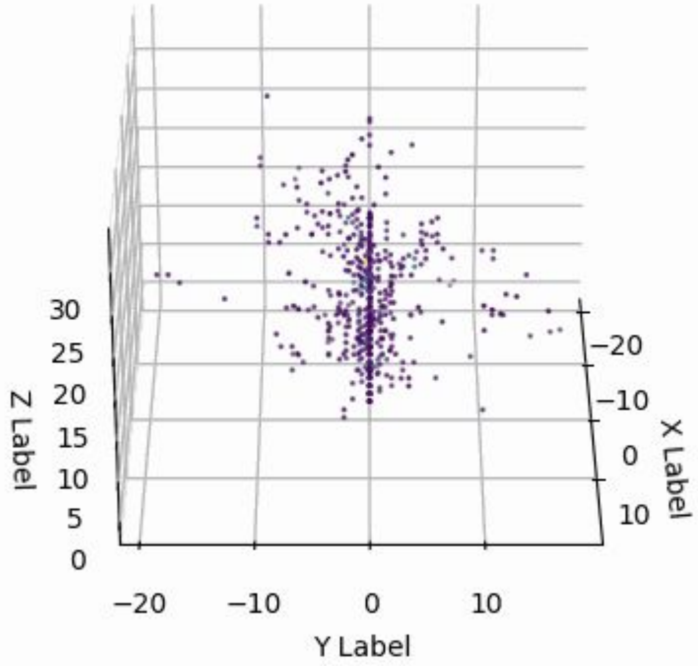
# Model Limitations

Sparse Data!



# Model Limitations

Alternate Data Representation (Point Clouds)



# Gains

- A substantial increase in shower simulation time in calorimeters.
- More efficient simulations (taking only the necessary information at different abstraction levels).
- Flexible level of physics and experimental bias for each experiment simulation

# Thank you!

Any Questions?

CERN openlab