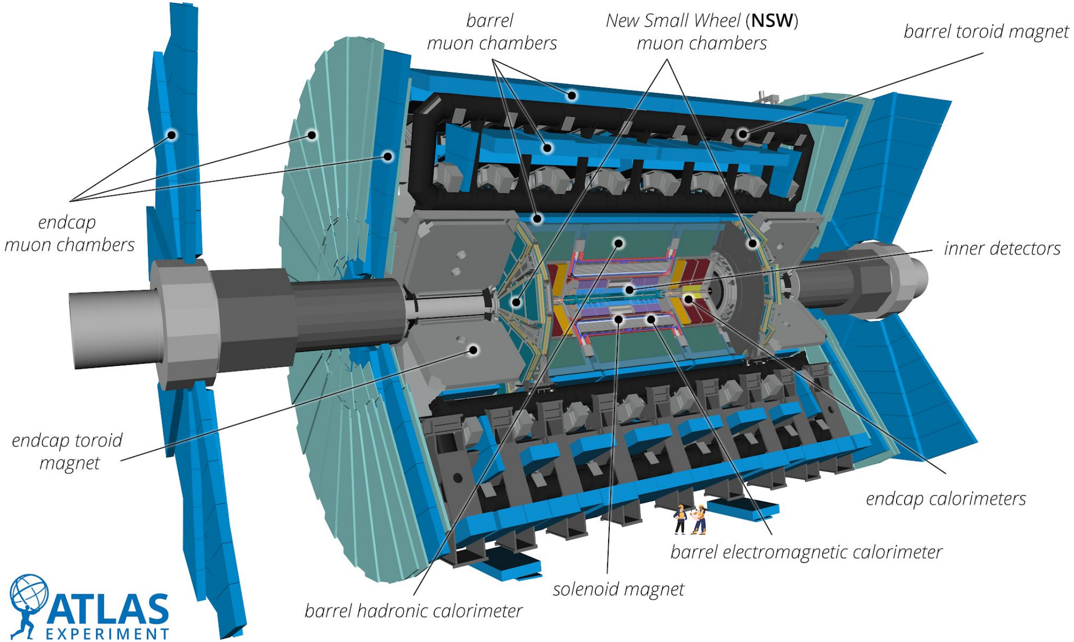
The ATLAS Monte Carlo detector simulation for Run 3 at the LHC

ICHEP, 20th July 2024

John Chapman (University of Cambridge) on behalf of the ATLAS Collaboration

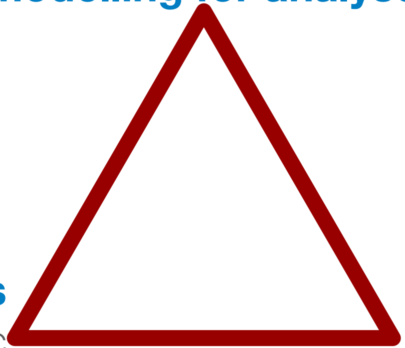
ATLAS Experiment



- Simulation of interactions within the ATLAS Experiment at the LHC is performed using Geant4 integrated into the ATLAS Offline Software framework (Athena).
- In Run 2 ATLAS used Geant4 10.1. In Run 3 ATLAS uses Geant4 10.6.

Competing factors for Simulation development

Stability of physics modelling for analysers



Improved physics modelling (data-MC agreement)

Improved technical performance (higher throughput, lower resource requirements)

- Analysers need a consistent set of Monte Carlo (MC) samples with matching physics modelling corresponding to all data periods in an LHC run.
- Physics modelling changes in MC require updated recommendations for physics objects.
  - This is an expensive process (personpower) which is only undertaken once or twice per LHC run.
- Optimizations which improve technical performance can be added to production releases between data-taking years in an LHC Run - *if they do not alter the physics modelling*.
- Changes (improvements) to physics modelling can only be included in production releases at points when new physics object recommendations are planned.

Methods of optimizing technical performance

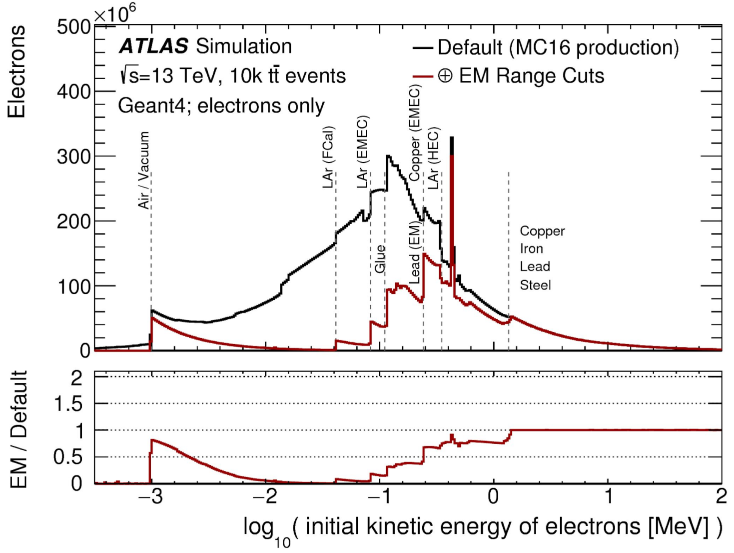
- There are two main ways to optimize the technical performance of simulation code:
  - Avoid simulating uninteresting particles (Simulation output changes.)
  - Speed up the simulation of interesting particles
    - Do the same thing, but faster. (Simulation output unchanged.)
    - Do something simpler. (Simulation output changes.)
- Output-changing optimizations require very careful validation, but can produce output which is compatible with being analysed together with previous MC production.
- This talk focuses on "Full Simulation", but ATLAS also maintains a "Fast Simulation" workflow.
- See "The Fast Simulation Program of ATLAS at the LHC" later in this session for a discussion of the other flavours of simulation used by ATLAS.

Avoid Simulating uninteresting particles

- Beam-pipe killer
  - Kill particles entering certain forward beam-pipe volumes which mean they will be very unlikely to reach sensitive detector regions. (Statistical differences in output due to random number changes for remaining particles.)
- **EM Range Cuts**
  - Turn on range cuts for gamma processes (conv, phot, compt). See next slide.
- Neutron/Photon Russian Roulette
  - Kill low energy neutrons/photons with some probability when created. Upweight energy deposits of surviving particles accordingly. Speed-up: ~10%

EM Range Cuts

- Range cuts are a built-in way of optimizing Geant4 performance
- For each material-volume pair, range cuts can be specified in distance units (mm).
- Secondary particles that are expected to travel less than the range cut are not created and their energy is immediately deposited by the parent particle.



- By default Geant4 does not apply range cuts for the conversion, photo-electric or Compton-scattering gamma processes.
- Option provided by Geant4 to activate range cuts for these processes:
  - G4 command: `'/process/em/applyCuts true'`
- Range cut of 0.1 mm used (same as for electron processes).
- Speed-up: 6-10% - due to far fewer being particles created.

Simulate interesting particles more efficiently

ATLAS EXPERIMENT

- VecGeom
  - ATLAS uses VecGeom implementations for a subset of G4Solid implementations.
  - In testing it was found that switching only Cones, Tubes and Polycone implementations from the defaults to VecGeom gave the best performance. Speed-up: 2-7%
- EM Endcap Calorimeter (EMEC) Geometry Optimisation
  - Reduce the time needed for geometry navigation calls by dividing the EMEC inner (outer) wheel into 14 (21) thick slices along the z-axis. Speed-up: 5-6%
- Tailored magnetic field switch-off in LAr Calorimeters
  - Magnetic field switched off in central LAr calorimeter for all particles except muons. Speed-up: 3%.
- G4GammaGeneralProcess
  - Use a super-process that hides all six standard physics processes involving photons.
  - Only one mean free path needs to be calculated for a photon. Speed-up: 3%
- **Woodcock Tracking in the EMEC**
  - See following slides.
- **Big Library**
  - See following slides.
- **Link Time Optimization**
  - See following slides.

Woodcock Tracking in the EMEC

- Woodcock Tracking is a tracking optimisation technique for highly segmented detectors where the geometry boundaries rather than the physical interactions limit the simulation steps.
- In the EMEC region photon simulation is dominated by the transportation process.
- Photons don’t interact during transport (no continuous energy deposition), therefore safe to perform tracking of photons in a simplified EMEC geometry (i.e. without boundaries) made of the densest material from the standard EMEC geometry (Pb).
- Interaction then occurs with probability equal to ratio of cross-sections of the true material and Pb. Statistical changes in output only.
- Woodcock tracking is applied on top of the G4GammaGeneralProcess.
- This code will be part of G4HepEM, but was added as atlas patch on Geant4 10.6.patch03
    - Geant4 command: `'/process/em/useWoodcockTracking EMEC'`
- Speed-up: 17.5%
    - 50% reduction in number of steps for photons in EMEC.

Big Library

- Single shared object library for all Athena code with Geant4 dependencies - statically linked to Geant4 libraries.
- Implementation strategy:
  a. Identify all packages which depend on Geant4.
  b. Compile Geant4 external in static mode (e.g. create .a archives).
  c. For all the packages identified in (a) modify CMakeLists.txt to create a cmake library of OBJECT type (.o files).
  d. Create a meta-package that will create the “big library” using all OBJECT libraries and static link from G4.
- Speed-up: 5-7%

PkgA/CMakeLists.txt

```
atlas_subdir( PkgA )
atlas_add_library( PkgA
     src/*.cxx
     [...]
     OBJECT )
// ...
```

PkgB/CMakeLists.txt

```
atlas_subdir( PkgB )
atlas_add_library( PkgB
     src/*.cxx
     [...]
     OBJECT )
// ...
```

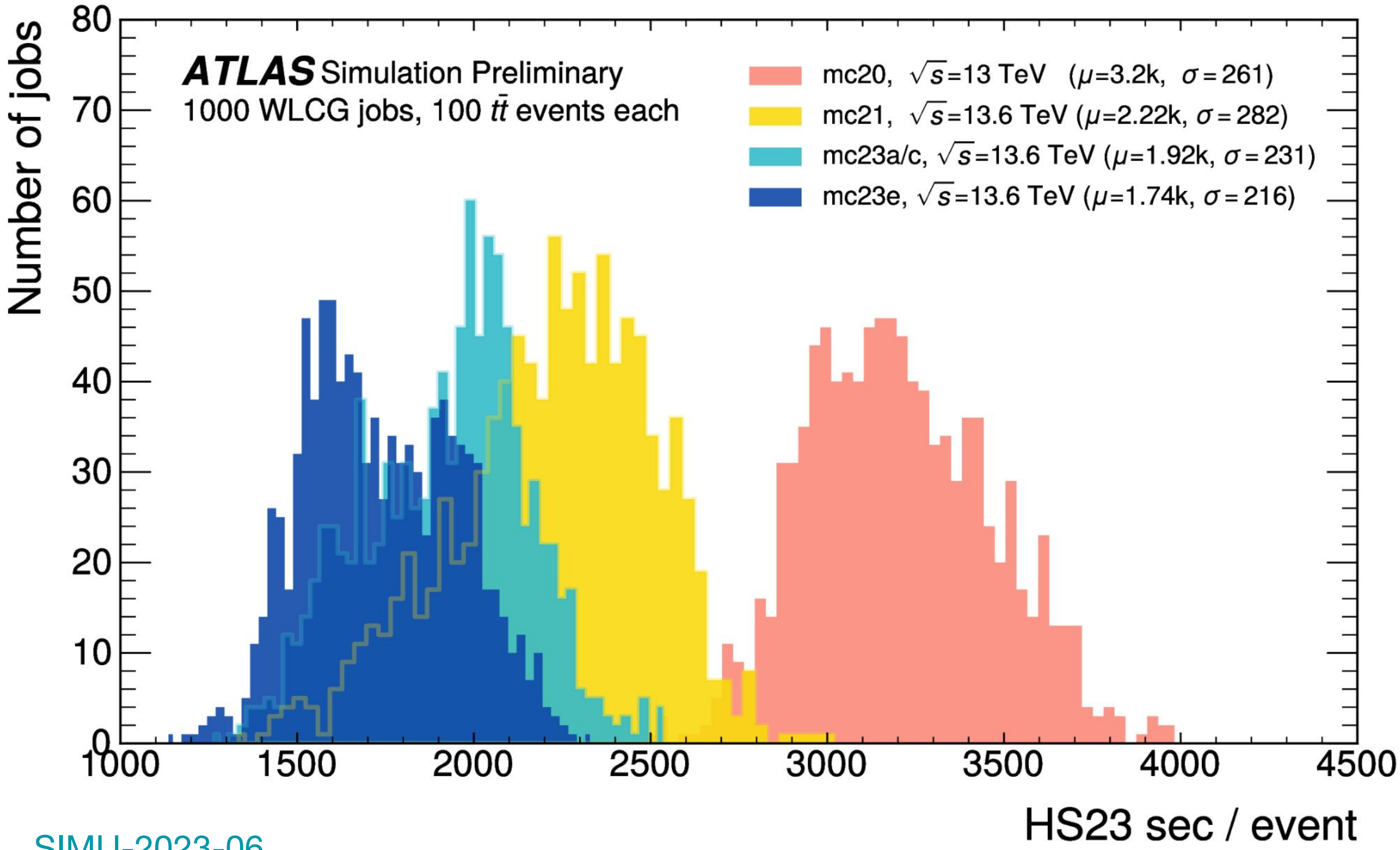
PkgC/CMakeLists.txt

```
atlas_subdir( PkgC )
atlas_add_library( PkgC
     [...]
     OBJECT_LIBRARIES
     PkgA
     PkgB )
// ...
```

Link-Time Optimization

- Once all files have been compiled separately into object files, traditionally, a compiler links (merges) the object files into a single file, the executable.
- However in LTO the compiler is able to dump its intermediate representation, so that all the different compilation units that will go to make up a single executable can be optimized as a single module when the link finally happens. [wikipedia]
- As all Athena code with Geant4 dependencies is statically linked together into a single shared-object library it is possible to use LTO on this shared-object library (instead of an executable).
- This only required changes in the CMake configuration of the Athena build.
- As expected, the simulation output is identical after this purely technical change.
- Speed-up: ~5%

CPU times



Throughput increased by x1.84 between mc20 and mc23e campaigns! (mc20 uses Geant4 10.1, mc21 & mc23x use Geant4 10.6)

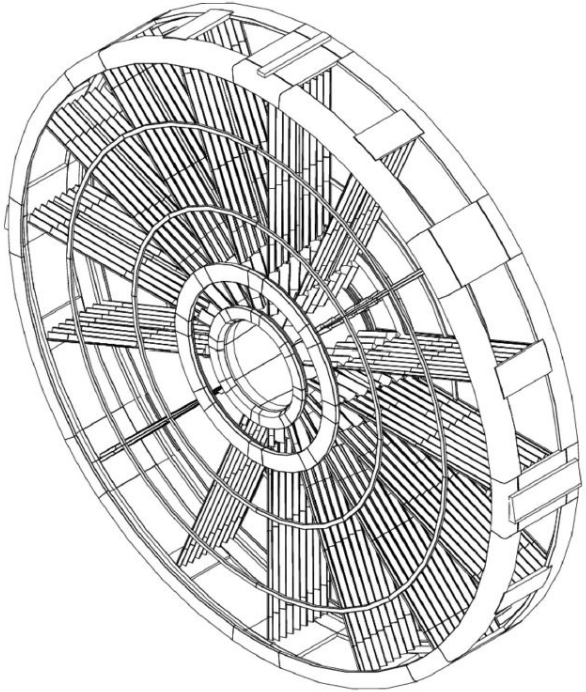
SIMU-2023-06

Possible future changes

- Switching off Energy Loss fluctuation
  - G4 Command "`/process/eLoss/fluct false`"
  - Physics output-changing. Postponed until next major MC campaign.
- Advanced Compiler Optimisations - PGO / AutoFDO
  - Next step after LTO - use profile driven feedback to further optimize the big library.
- **High-η particle rejection**
  - See following slides
- Parameter Tuning of In-Field Tracking
  - Customize G4 tracking parameters based on particle type, properties and location region to optimize CPU performance without compromising precision.
- Adoption of G4HepEM and the specialised transport
  - A new compact Geant4 EM library optimized to be used for HEP em showers development and transport. It provides significant speedup w.r.t the general Geant4 EM library.
- G4NeutronGeneralProcess
  - Super-process for neutron physics.
- **Re-implementation of EMEC geometry**
  - See following slides

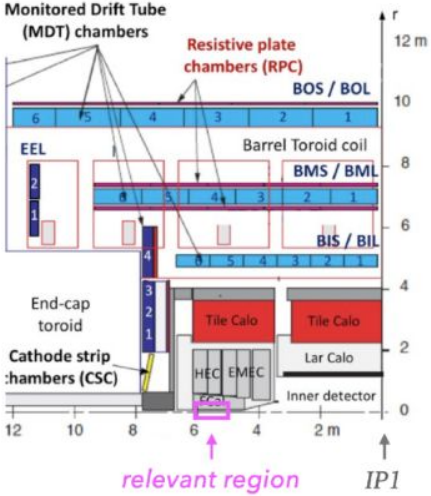
Re-implementation of EMEC Geometry

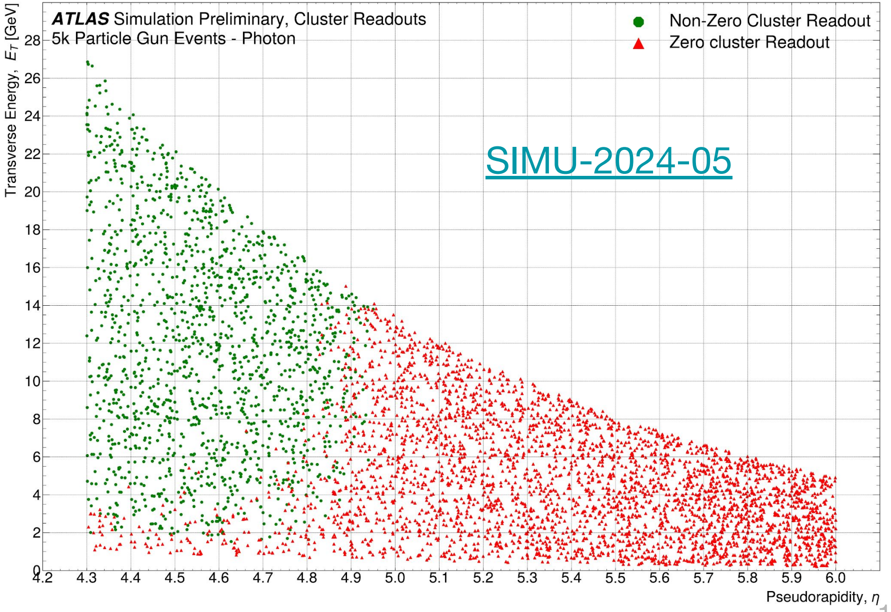
- Background:
  - The as-built EMEC has a complicated “Spanish Fan” geometry.
  - Efficient description using the G4Solids available in early versions of Geant4 was not possible.
  - Custom solids used to implement the geometry algebraically.
- New implementation based on data taken from the technical drawings
  - EMEC mother volume
  - Inner and Outer Wheel envelopes
  - Inner and Outer Wheel Absorbers and Electrodes defined as G4GenericTrap solids
- Includes option to subdivide the geometry into slices to aid voxelization.
- Initial tests in a stand-alone Geant4 example:
  - Tracking performance ~4 times faster even with extra debugging enabled.
- Code now integrated into Athena.
- Next step is to adapt the corresponding sensitive detectors to the new geometry implementation to allow assessment of physics modelling.



High-η particle rejection

- **Idea:** Kill primary particles generating secondaries close to the beam-pipe at ~ 5-6 m from the IP.
- Many particles in the collision are at high $|\eta|$ (no Inner Detector hits) with little energy compared to the calorimeter noise.
- Check if we can kill some particles early on which will have no or little effect on the simulated energy in the calorimeters to save CPU.
- Cutting $\eta > 5.0$ and $E_T < 0.5$ GeV looks promising.



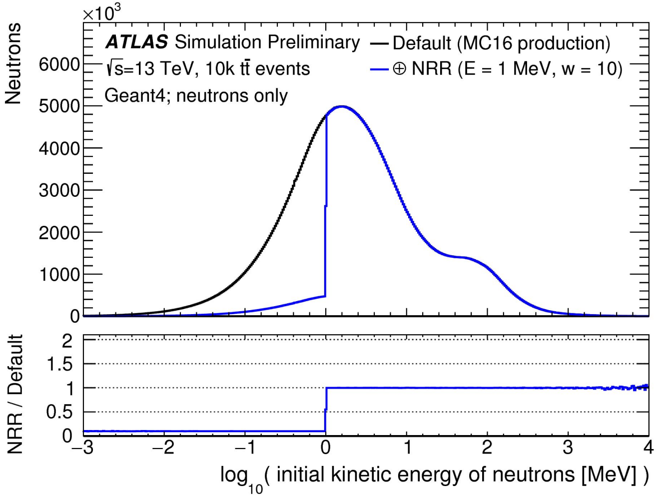


Summary

- The ATLAS Simulation group aims to:
  - provide Monte Carlo samples with consistent physics for entire LHC runs for analysers
  - include new optimizations without changing physics modelling between yearly sub-campaigns.
  - include physics modelling improvements between campaigns.
- Multiple optimizations were introduced between the Run 2 (mc20) campaign and the latest Run 3 campaign (mc23e - matching 2024 data), increasing throughput by a factor of 1.84!
- Close collaboration with the Geant4 Collaboration is key.
- Healthy programme of on-going development to include further optimizations in the future both from adopting new Geant4 features and improving code within Athena.
- Further performance improvements expected for the sub-campaign for 2025 data.

Backup

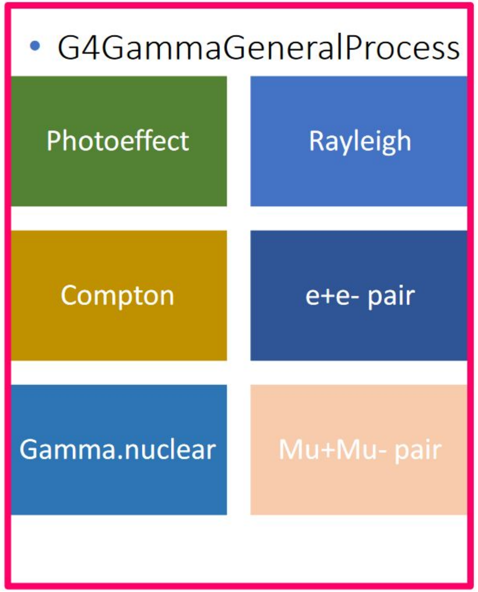
Neutron/Photon Russian Roulette

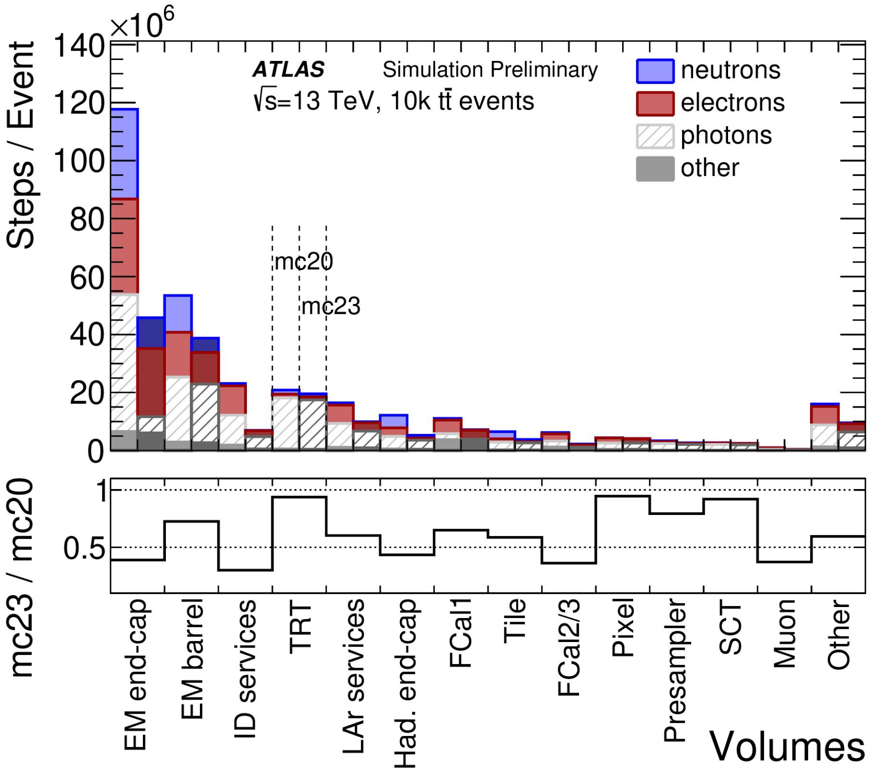


- Neutron Russian Roulette
  - Low energy neutrons take quite some CPU time in simulation, usually with many steps that are not really correlated with the point of their creation.
  - Randomly kill neutrons below some threshold energy with a probability `1/w` and apply a corresponding weight (`w`) to the remaining neutrons. The remaining neutrons would then deposit `w`-times the energy.
    - Based on an idea already used in CMS Simulations.
  - Parameters used in production: Threshold = 2.0 MeV, w = 10.
- Photon Russian Roulette:
  - Applied to photons produced in the LAr EM Calorimeters in ATLAS
    - Avoids unwanted effects in the Inner Detector.
  - Parameters used in production: Threshold = 0.5 MeV, w = 10.
- Both Implemented in Athena in a configurable G4UserStackingAction.
- Speed-up: ~10% overall

G4GammaGeneralProcess

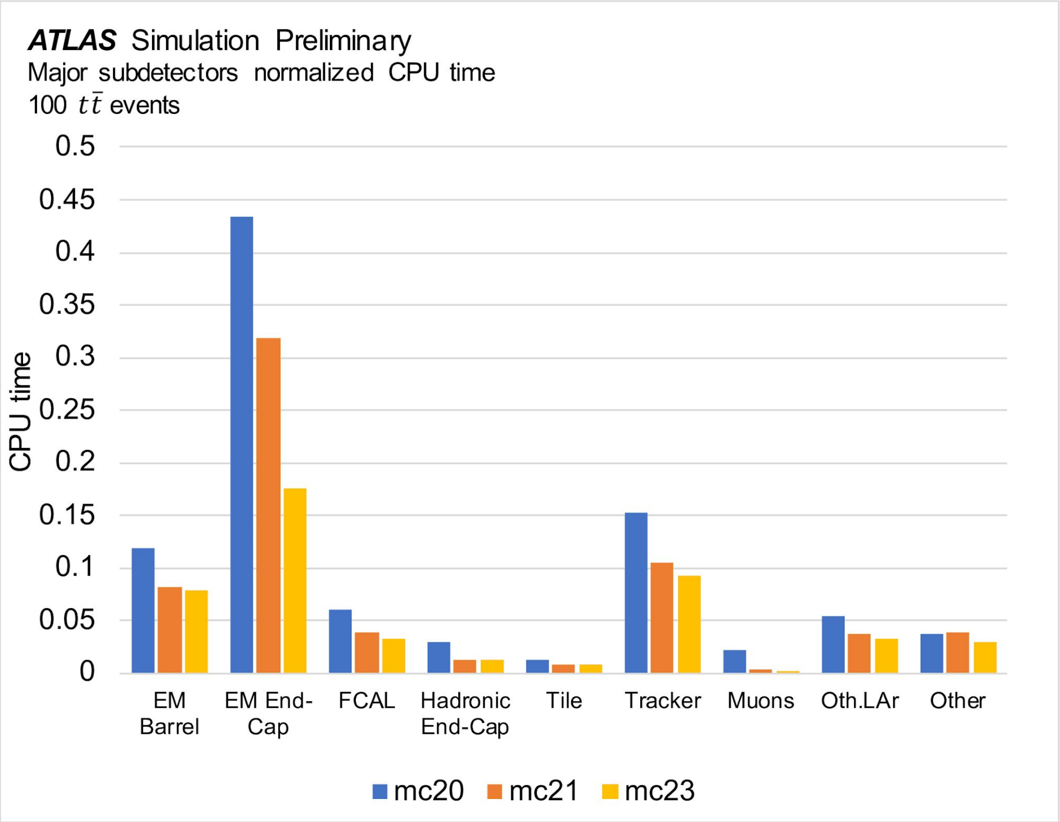
- The G4GammaGeneralProcess is a super-process that hides all six standard physics processes involving photons.
- The G4SteppingManager only sees one physics process rather than six, so only one mean free path needs to be calculated for a photon.
- The number of instructions is reduced at the cost of introducing extra physics tables shared between threads.
- Initial version included in Geant4 10.6 onwards. Additional patches backported to atlas patch on Geant4 10.6.patch03.
- Output unchanged.
- Geant4 command: `'/process/em/UseGeneralProcess true'`
- Speed-up: 3%





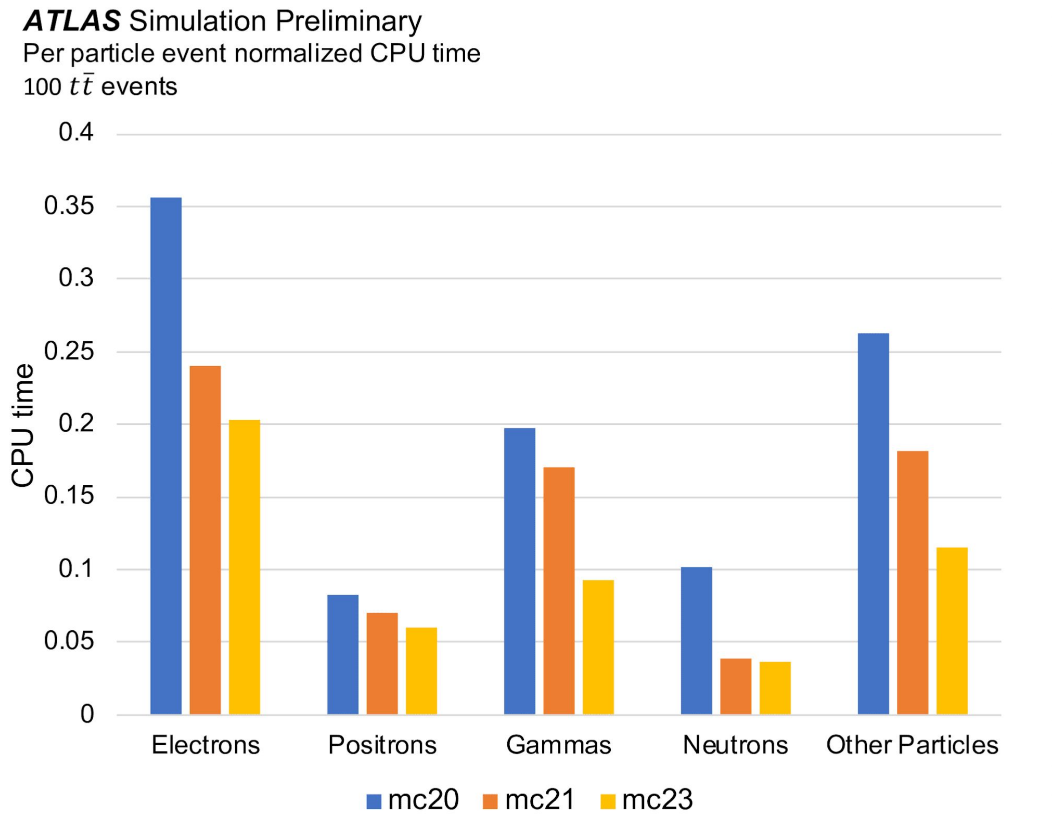
- Number of Geant4 steps per event for various ATLAS detector volumes. The left column in each section represents the Run 2 (mc20) setup and the right column represents the setup during Run 3 (mc23 = mc23c).
- 'FCal1' includes the first (electromagnetic) module of the forward calorimeter and 'FCal2/3' includes the subsequent two hadronic modules.
- 'ID services' includes ID services and the beam pipe.
- 'LAr services' includes LAr services and LAr cryostats.
- 'Other' includes all other particles and all other volumes that are simulated.

SIM-2024-003/



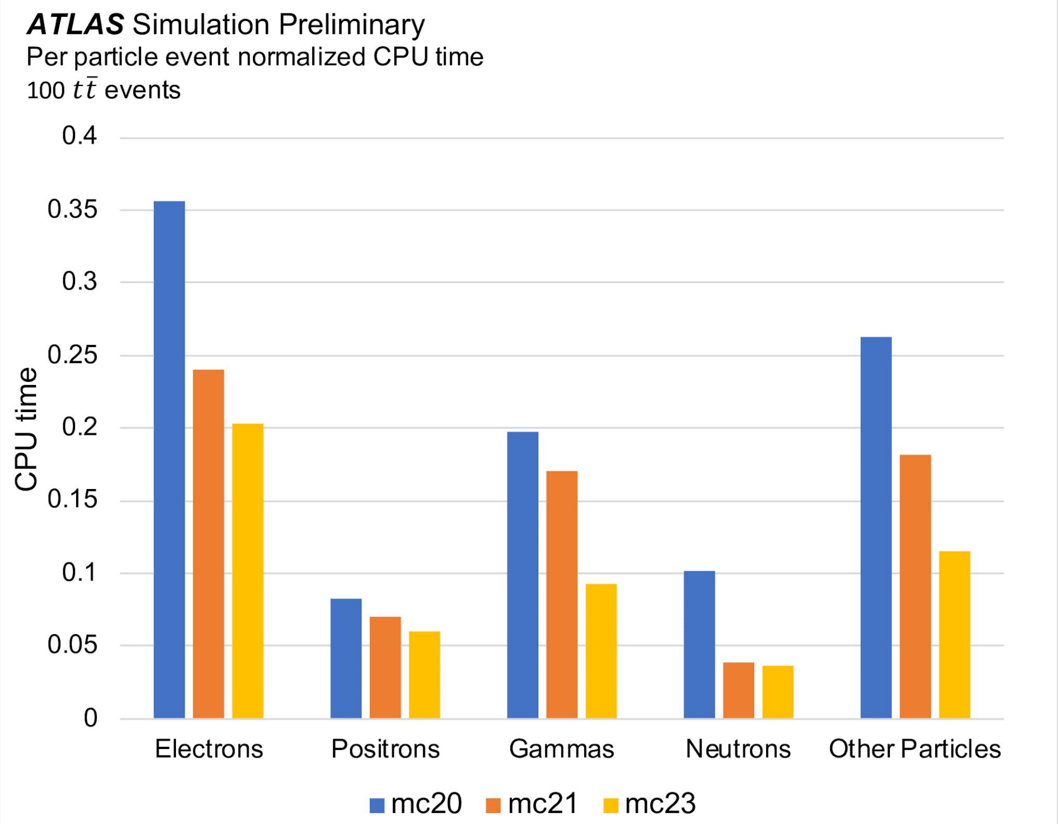
- Time spent per event simulating 100 $t\bar{t}$ events, normalized to the total time spent to simulate events in mc20, for each of the major subdetectors.
- The different colored bars represent different simulation configurations: the left most bar, for each subdetector, shows the simulation time for the Run 2 (mc20) configuration; the middle bar displays the time for an optimization used for the first Run 3 simulated samples (mc21); while the right bar shows the time spent on simulating events with the latest Run 3 configuration (mc23 = mc23c).

SIM-2024-003/



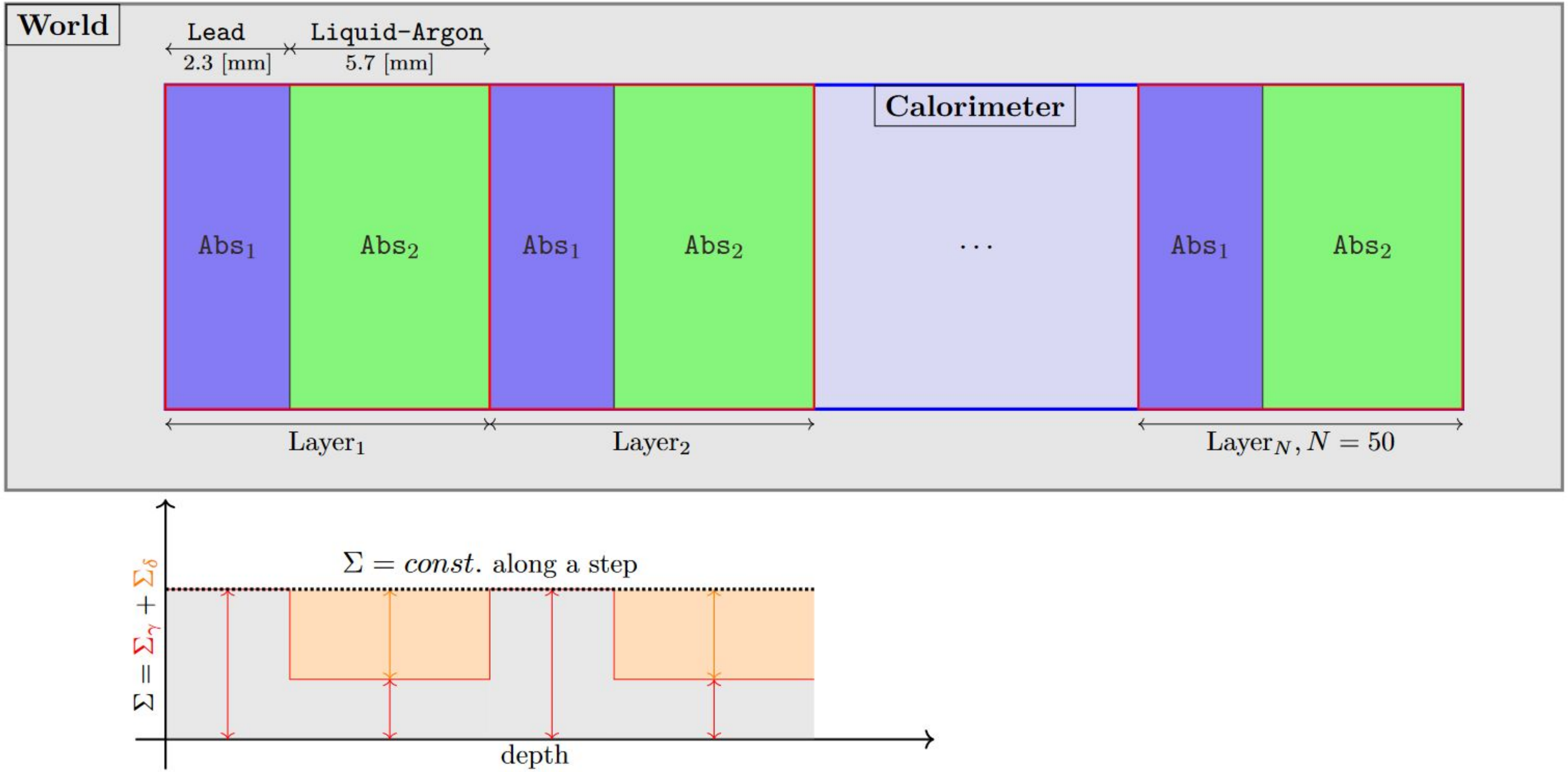
- Time spent per event simulating 100 $t\bar{t}$ events, normalized to the total time spent to simulate events in mc20, for each particle type.
- The different colored bars represent different simulation configurations: the left most bar, for each subdetector, shows the simulation time for the Run 2 (mc20) configuration; the middle bar displays the time for an optimization used for the first Run 3 simulated samples (mc21); while the right bar shows the time spent on simulating events with the latest Run 3 configuration (mc23 = mc23c).

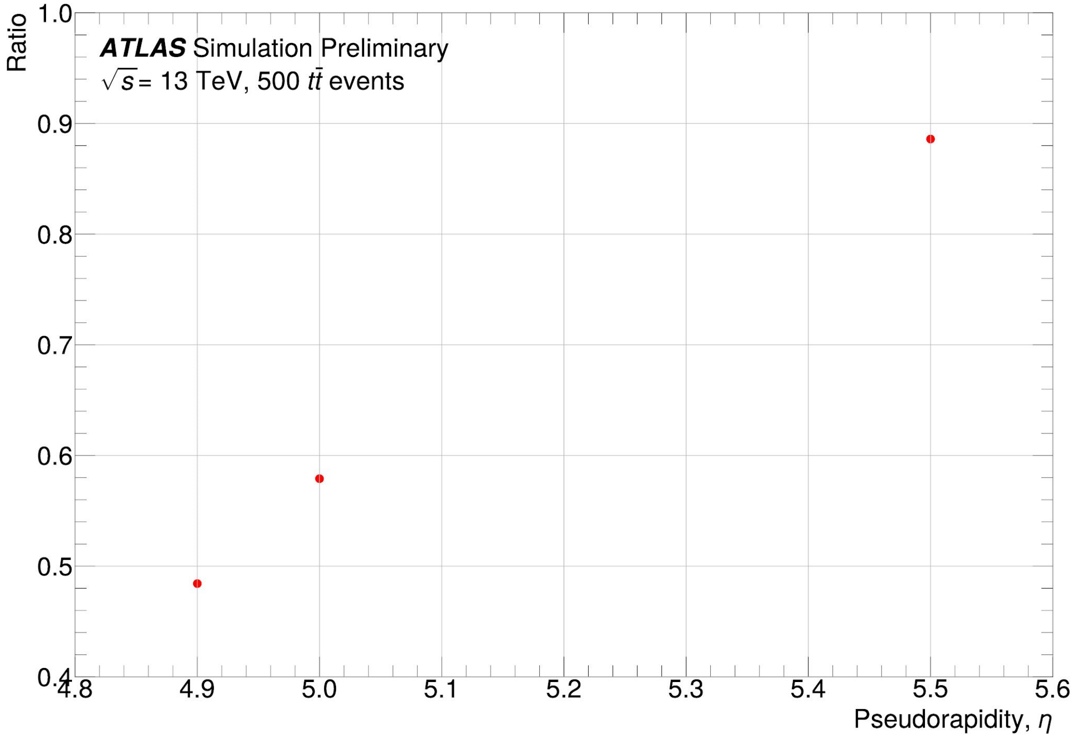
SIM-2024-003/



- Time spent per event simulating 100 $t\bar{t}$ events, normalized to the total time spent to simulate events in mc20, for each particle type.
- The different colored bars represent different simulation configurations: the left most bar, for each subdetector, shows the simulation time for the Run 2 (mc20) configuration; the middle bar displays the time for an optimization used for the first Run 3 simulated samples (mc21); while the right bar shows the time spent on simulating events with the latest Run 3 configuration (mc23 = mc23c).

SIM-2024-003/


World
Lead
2.3 [mm]
Liquid-Argon
5.7 [mm]
Calorimeter
Abs$_1$
Abs$_2$
Abs$_1$
Abs$_2$
...
Abs$_1$
Abs$_2$
Layer$_1$
Layer$_2$
Layer$_N$, $N = 50$
$\Sigma = const.$ along a step
$\Sigma = \Sigma_\gamma + \Sigma_\delta$
depth



- Benchmark study for the relative CPU speed for simulating the particles in 500 ttbar events with pseudorapidities larger than 4.3 and smaller than the pseudorapidity value indicated in the figure. The CPU speed is given relative to the simulation of particles in ttbar events in the pseudorapidity range of = 4.3 ~ 6.0.

SIM-2024-003/