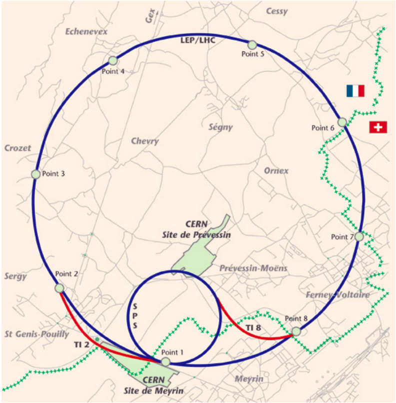
The ATLAS Simulation

Jochen Meyer
On Behalf Of The ATLAS Simulation Group

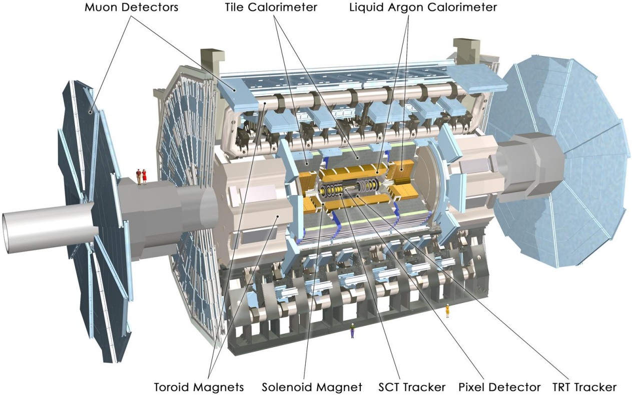
06 October 2016

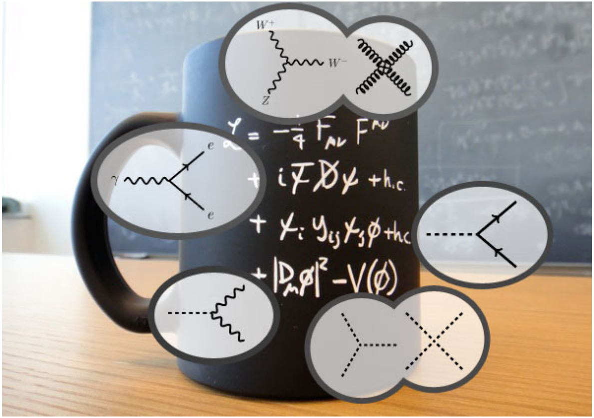
SCCTW'2016
South-Caucasus Computing and Technology Workshop



- **L**arde **H**adron **C**ollider
  - circular tunnel of 27 km diameter below Switzerland and France
  - machine to accelerate and collide particles (mainly protons, but also lead cores)

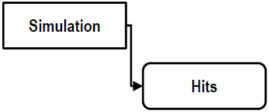
- **A T**oroidal **LHC A**pparatu**S**
  - one of two multi-purpose experiments at the LHC
  - detection of charged particles, photons and hadrons
  - measurement of particle trajectories and energy deposits





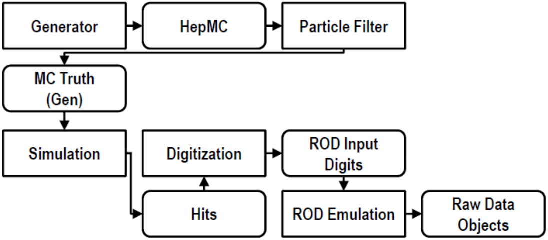
- Why The Experiment?
  - measurement of natural constants like the mass of the W boson
  - probing of the well-established Standard Model of particle physics
  - test for predictions made by extensions and other theories

- Why To Simulate The Experiment?
  - simulation helps to understand the detector and to separate experimental effects from actual new physics phenomena
  - some backgrounds cannot be extracted from data and an adequate Monte Carlo prediction is required to study physics models
  - compatibility of simulated new physics and data can be probed and eventually confirmed or rejected with certain degree of confidence

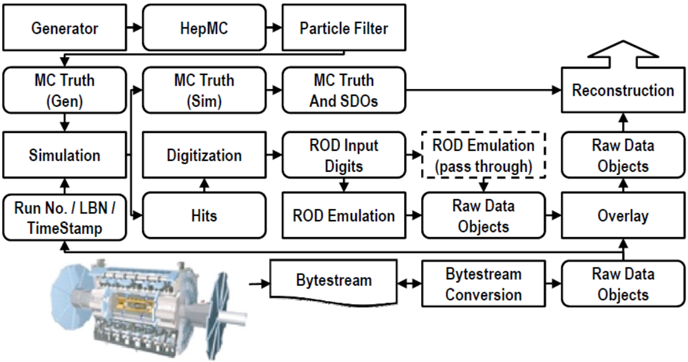


- Geant4 toolkit is used to ...

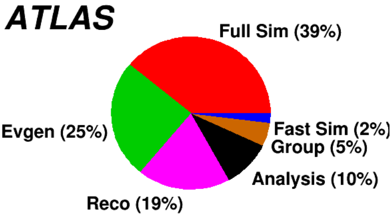
... track particles through the detector,
... simulate their multiple scattering,
... model their energy loss and
... steer their decay if needed



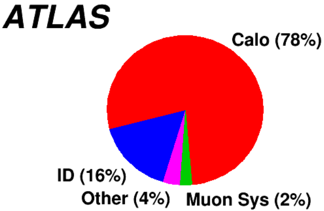
- further responsibilities of the ATLAS simulation group are ...
  - ... emulation of detector response during digitization step
  - ... managing some generator code which provides physics input



- further responsibilities of the ATLAS simulation group are ...
  - ... emulation of detector response during digitization step
  - ... managing some generator code which provides physics input
  - ... injection of backgrounds and underlying events closely to data
  - ... supplying events with “truth” information of all the above

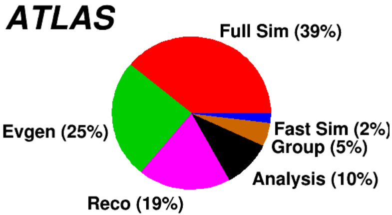


Wall clock time fraction for grid and HPC jobs
July 2015 - July 2016

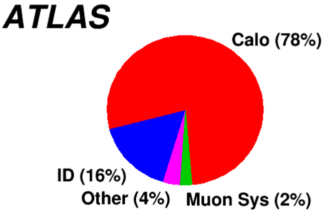


Subdetector CPU fraction for 50 ttbar events
MC16 Candidate Release

- real data rates (events per second):
  - LHC delivers about 600 million evts/sec
  - ATLAS hardware reduces rate to approximately 100.000 evts/sec
  - after pre-selection it is $\mathcal{O}(1000)$ evts/sec

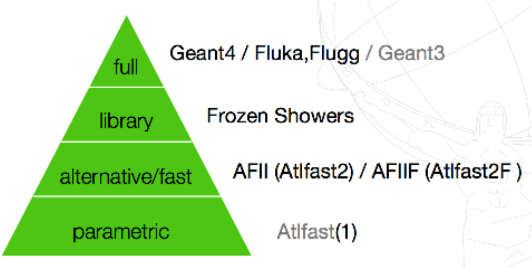


Wall clock time fraction for grid and HPC jobs
July 2015 - July 2016



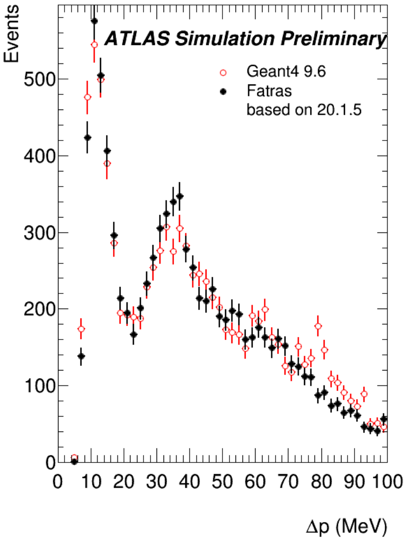
Subdetector CPU fraction for 50 ttbar events
MC16 Candidate Release

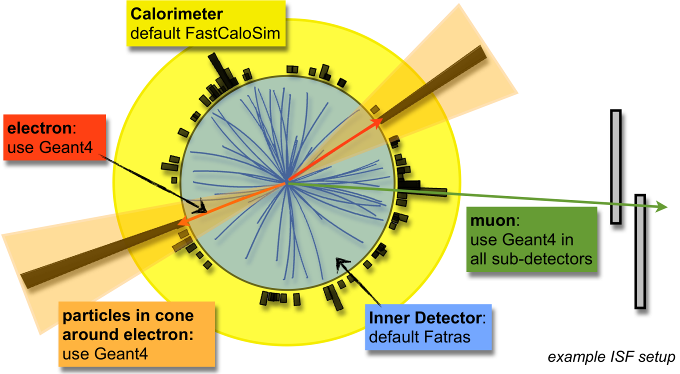
- billions of events are simulated as well, but impossible to provide equal or more Monte Carlo events for all physics processes
- simulation nevertheless consumes majority of available computing resources (particularly low energetic particles below 10 MeV in calorimeter)



- speed at cost of accuracy
  - full simulation is accurate but slow
  - trimmed pre-simulated showers injected for low energetic $e^{\pm}/\gamma$ after property matching
  - FastCaloSim/Atlfast2 loads parametrized response tuned to data into calorimeter cells
  - FATRAS uses simplified derivatives of detailed geometry to perform fast tracking of particles

- accuracy of simulation depends on know ledge of interaction between particles and material and therefore also on geometry
- ideally a cocktail of all flavors can be operated depending on requirements

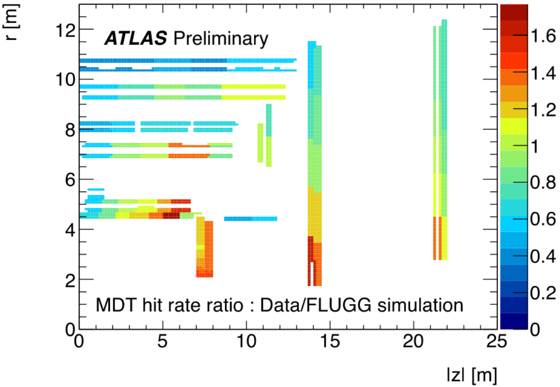




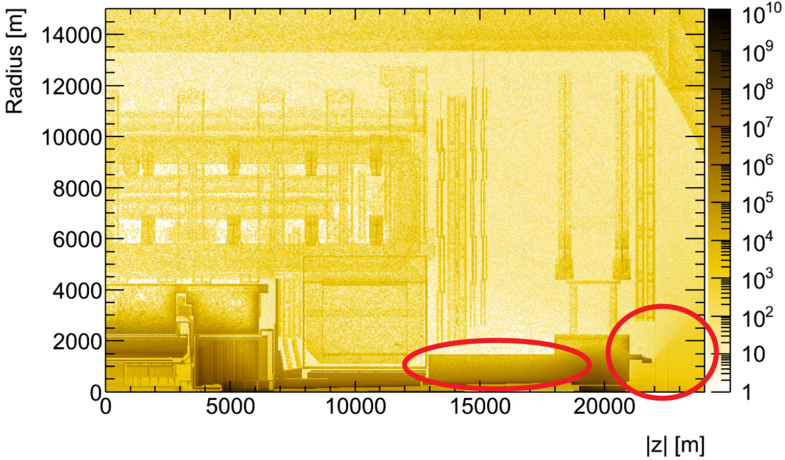
- **I**ntegrated **S**imulation **F**ramework allows to mix different flavors
- recent infrastructure developments give possibility for very specific configurations of Geant4 depending on volume and particle type

- accuracy of ATLAS Muon Spectrometer geometry can be tested by comparing ...
  - ... contribution of multi-scattering term in alignment studies
  - ... cavern background hit rates

- MDT hit rate (= flux $\otimes$ sensitivity) in data is compatible with FLUGG based simulation within a factor of 2



- early Geant4 geometry shows obvious flaws and differs from FLUGG geometry

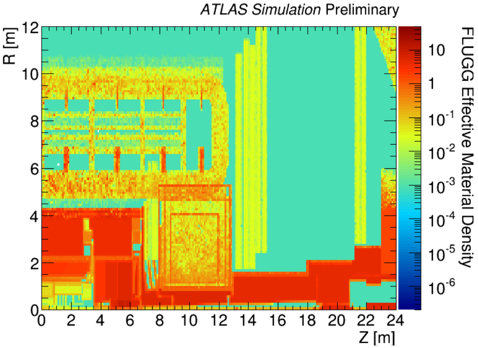


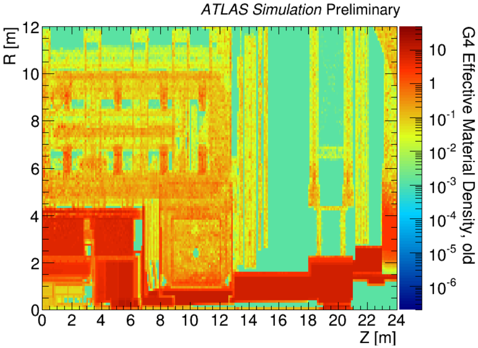
arbitrarily normalized total energy deposited in average minimum bias event

- analyzed simulated quantity (not directly observable in data)

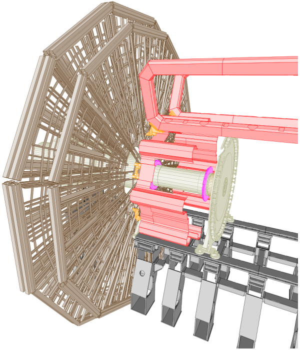
$$\text{effective material density [mass/volume]} = \frac{\text{energy desposition [energy/volume]}}{\text{dose [energy/mass]}}$$

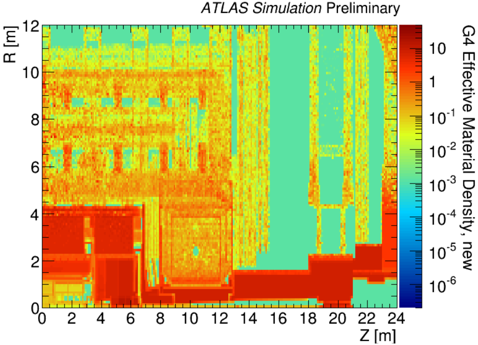
- comparison of simplified FLUGG geometry and first update of Geant4 geometry ("old") used for physics simulation
  - some structures not intended to be described in FLUGG



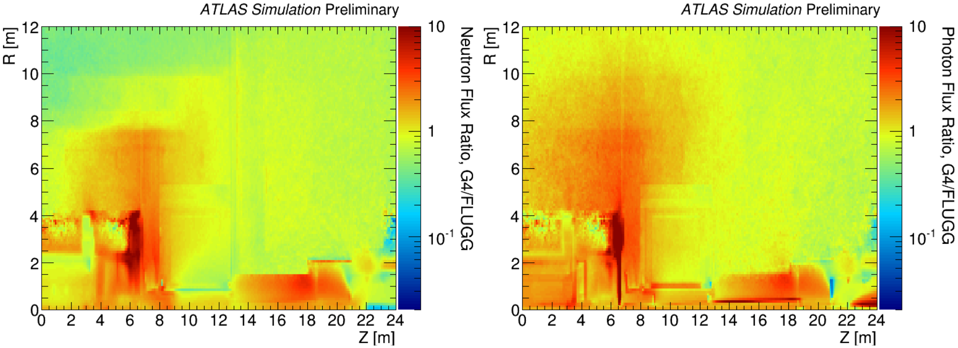


- major updates of Geant4 geometry ("new"):
  - thermal shielding for all toroid coils (red)
  - additional shielding inside end-cap toroid (magenta)
  - re-implementation of end-cap support (brown)
  - shielding installed during winter shutdown 2011/2012 (purple)
  - axial force return brackets (yellow) particularly important for scattering
- more updates to come also with help of **GTU**

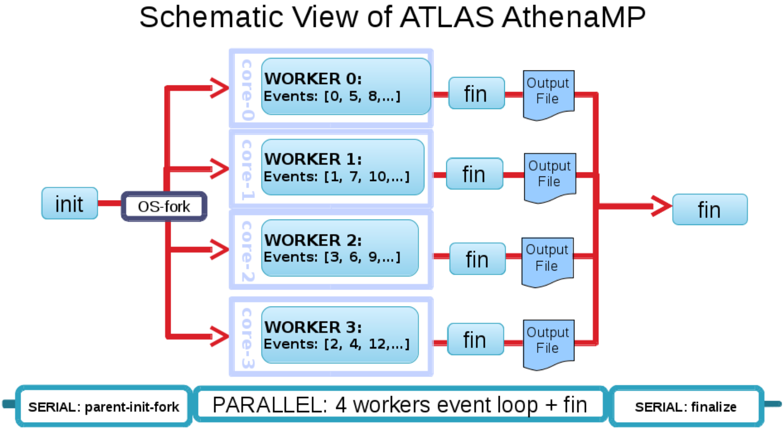


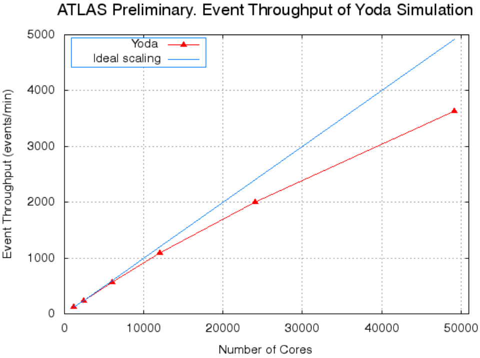


- ratio of neutron flux (left) and photon flux (right) simulated with Geant4 divided by corresponding flux simulated with FLUGG
- initial significant overshoot by Geant4 is gone and both simulations are in fairly good agreement
- comparison of hit rate from data will show if reduced flux of Geant4 where FLUGG is higher indeed holds expectations it promises



- encapsulation of event loop to sub-process(es)
  - master appoints events to workers and provides general configuration
  - workers run in parallel and return results when done

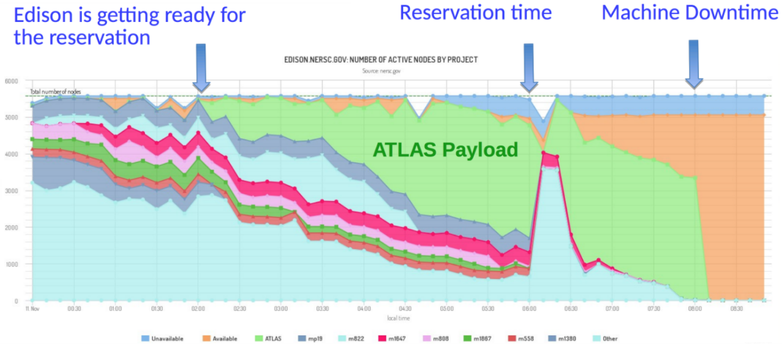




- approach is compatible with high-performance computing (HPC) model based on worker nodes
- new resources besides common grid become available
- first test with Yoda scheduler show good scaling behavior even though there is still room for improvement

- one step further is to parallelize simulation of events themselves
- dedicated functionality is available in Geant4 already
- implementation of basic requirements in ATLAS simulation software almost completed with re-written infrastructure components:
  - sensitive detectors
  - user actions
  - magnetic field and geometry
- to have a dedicated simulation release is of great advantage in this context

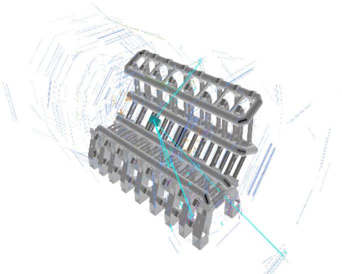
- more exceptional HPC resources (e.g. supercomputers like Cori and Edison [*] from the **N**ational **E**nergy **R**esearch **S**cientific **C**omputing Center) are now available



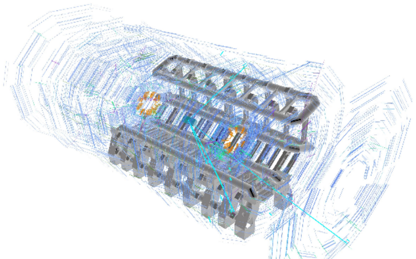
[*] Cray XC30, peak performance of 2.57 petaflops/sec, 133,824 cores, 357 terabytes of memory, 7.56 petabytes of disk

- embedding of real data
  - simulation of process does not describe reality with required accuracy
  - replacement of portions of actual events similar to desired process with simulated output
- overlay with real data
  - required statistics for desired process cannot or only badly be simulated
  - simulated physics process is overlaid with detector output of real data
- similar techniques are common to both approaches, but attitude and underlying problem are different

- decay of Higgs boson and different amount of background



$H \rightarrow ZZ \rightarrow 4\mu$ @ 2.6e33



$H \rightarrow ZZ \rightarrow 4\mu$ @ 2.6e34

Summary

- simulation is important to understand all details of real data
- Monte Carlo event generation is a technically challenging task taking a lot of computing power
- a balance between speed and accuracy has to be found
- parallelization speeds up and opens the way to new resources
- combinations of real data and simulated data achieve goals simulation alone would not be able to