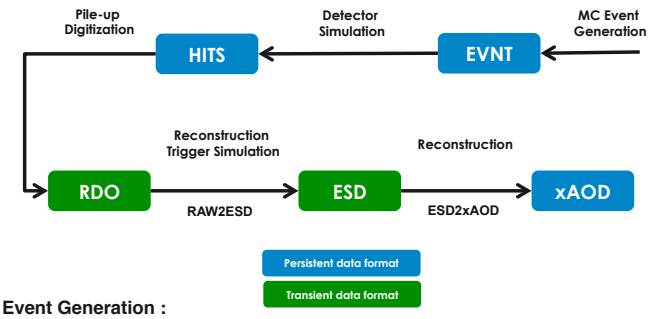
ATLAS Monte Carlo production Run-1 experience and readiness for Run-2 challenges

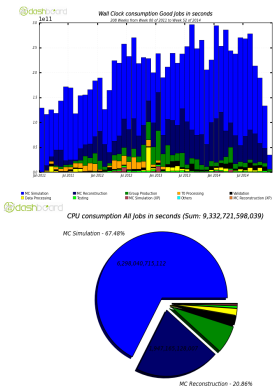
J. Chapman[1], J. Ferrando[2], J. Garcia Navarro[3], C. Gwenlan[4], S. Mehlhase[5], V. Tsulaia[6], A. Vaniachine[7], J. Zhong[4] and A. Pacheco Pages[8] for the ATLAS Collaboration

[1]University of Cambridge, [2]University of Glasgow, [3]Instituto de Fisica Corpuscular, [4]University of Oxford, [5]Ludwig-Maximilians-Univ. Muenchen, [6]Lawrence Berkeley National Lab. , [7]Argonne National Laboratory, [8]Institut de Física d'Altes Energies

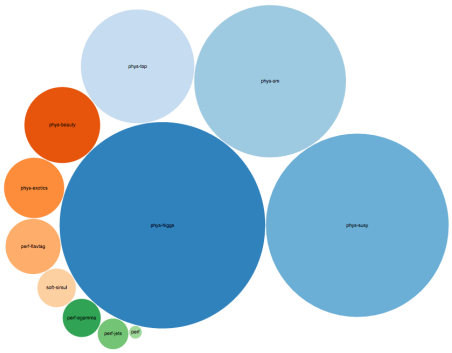
ATLAS Monte Carlo Simulation Flow



- **Event Generation :**
  - ~30 Monte-Carlo generators used to simulate physics processes:
    - Framework integrated and stand-alone generators.
  - Event generation workflows :
    - Single step generation: Pythia6/8, Herwig(++), Sherpa,…
    - Two-step generation: Parton level generator (Alpgen, MadGraph, …) coupled to framework generator for hadronisation.
- **Simulation :**
  - Geant4 (G4) full simulation:
    - All stable particles are tracked through the ATLAS geometry.
  - G4 full simulation with Frozen Showers in calorimeters: 25% Speed up.
    - Showers are tracked down to very low energy by G4 → Stop showering at a threshold and substitute by a pre-made list of energy deposits.
  - AtlFast-II (AF-II): Factor of 10 speed up in mc12.
    - Parameterize all particles except muons in the calorimeters.
  - Integrated Simulation Framework (ISF)
    - Balance accuracy with CPU consumption on a particle level, by selecting particles for accurate simulation only where needed and using fast simulation methods if possible.
- **Digitization :**
  - Simulate detector readout .
  - Simulate pile-up contributions (multiple pp interactions on top of hard scatter event).
  - Overlay a number of pre-simulated minimum bias events on each signal event.
  - Optimize pile-up event storage and access
    - Cache pile-up events in memory → Memory intensive.
    - Flush memory early and re-load from disk on demand → I/O and CPU intensive.
- **Reconstruction :**
  - Reconstruct simulated events in the same way as data.
  - Simulate Trigger.
  - Two step process:
    - RAWtoESD: main reconstruction → output is Event Summary Data (ESD).
    - ESDtoAOD: fast slimming process → output is Analysis Object Data (xAOD). xAOD is a new data format usable for both Athena jobs and lightweight ROOT (analysis) jobs introduced for the latest simulation campaign and is the default to be used in Run-2.



CPU consumption for the different steps of the production of simulated events.



Relative size of the event sample simulated of the more than 22.000 different datasets produced during Run-1 (mc12 campaign) by group.

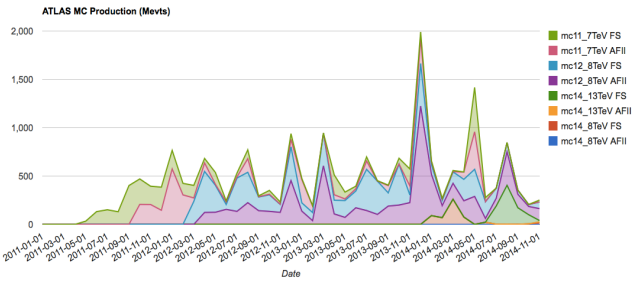
Campaign	Full Simulation (10^6 events)	Fast Simulation (10^6 of events)
mc11	3640	3270
mc12	6690	7150
mc14	1100	50

Run 1 MC Campaigns

- MC production campaigns correspond to data taking periods with same conditions: center-of-mass energy (COM), detector configuration, conditions, ...
  - **mc11**: Simulation configuration for 7 TeV in 2011. Four different sub-campaigns implementing pile-up condition, detector conditions and geometry closer to data.
  - **mc12**: Simulation configuration for 8 TeV in 2012 . Three sub-campaigns. Main part of the events were produced in sub-campaign mc12b. Late mc12c implemented an improved detector geometry description.

Run 2 preparations

- The LHC is expected to deliver increased COM energy to 13 TeV, twice luminosity and larger pileup. Different data and simulation conditions.
- **A new campaign was created to get ready for the new data taking in 2015.**
  - **mc14** : preparation for Run-2 production:
    - 8 TeV : Improved and updated simulation, digitization and reconstruction with same conditions as mc12 campaign.
      - ISF framework used as main simulation framework.
    - 13 TeV : Campaign with the COM for Run-2 and estimate of the pile-up and detector conditions.
      - Multicore processing becomes default for production (simulation, digitization and reconstruction).
  - **mc15** : Simulation configuration for 13 TeV. Improved detector description in simulation and new faster reconstruction.



Produced events (10^6) per month for the Run-1 period

Production Characteristics (mc14 example)

- **Simulation** :
  - Processing: 100 events per job → ~ 70 MB output file size → Merged up to 1000 events (0.7 GB file size) for better grid transfers and tape storage.
  - Low memory requirement: ~1 GB
  - Run time per event (averaged over grid CPUs) :
    - G4 full simulation without (335 s) and with Frozen Showers (250 s), and AtlFast-II (20 s).
- **Digitization and reconstruction :**
  - Processing 500 events per job → ~400 MB output file size → merged up to 5000 events (~2.0 GB file size) for better grid transfer and tape storage.
  - High memory usage: at least 3.6 – 3.8 GB (in 64 bit, serial processing).
- Both steps run software in multicore processing mode, using 8 cores simultaneously.