

Challenges of the ATLAS Monte Carlo production during Run 1 and beyond

J Chapman¹, W Ehrenfeld², J E Garcia Navarro³, C Gwenlan⁴,
S Mehlhase⁵, V Tsulaia⁶ and J Zhong⁴

¹Cavendish Laboratory, Cambridge, United Kingdom

²Physikalisches Institut, Universität Bonn, Bonn, Germany

³Universidad de Valencia, Valencia, Spain

⁴University of Oxford, Oxford, United Kingdom

⁵University of Copenhagen, Copenhagen, Denmark

⁶Lawrence Berkeley National Lab, Berkeley, USA

E-mail: wolfgang.ehrenfeld@cern.ch

Abstract. In this paper we will review the ATLAS Monte Carlo production setup including the different production steps involved in full and fast detector simulation. A report on the Monte Carlo production campaigns during Run 1 and Long Shutdown 1 will be presented, including details on various performance aspects. Important improvements in the work flow and software will be highlighted.

Besides standard Monte Carlo production for data analyses at 7 and 8 TeV, the production accommodates various specialised activities. These range from extended Monte Carlo validation, Geant4 validation, pile-up simulation using zero bias data and production for various upgrade studies. The challenges of these activities will be discussed.

1. Introduction

The luminosity of the Large Hadron Collider (LHC) increased significantly over the last years. The ATLAS experiment [1] collected 5 fb^{-1} of data at 7 TeV collision energy in the year 2011 and 20 fb^{-1} at 8 TeV in 2012. This large data set needs to be complemented with a sufficient number of Monte Carlo (MC) events for Standard Model measurements and new physics searches as every result relies on MC simulation. MC events are used in data analysis in many different ways: to estimate selection efficiencies, to study systematic effects from detector or pile-up mis-modeling, or to verify data-driven background estimation methods.

The challenges encountered in MC production have different origins: One is the grid, which provides large number of uniform processing slots with limited running time, memory, local and global disk space. The challenge is to structure and optimise the different MC production steps to run efficient on the provided resources. Another one are requirements from physics. On one side more MC events are always useful, which can be achieved by reducing running time or finding additional resources. On the other side the produced MC events need to reflect accurately the conditions from data taking, which is difficult to accomplish before data taking.

In the following, the different steps of the MC production are described including recent improvements. Details of the default MC production during Run 1 and Long Shutdown 1 (LS1) are given. Finally a selection of dedicated production activities are discussed.



2. Monte Carlo Production Steps

In the following the ATLAS MC production chain is described in some detail, including important performance numbers. Challenges and improvements are discussed.

2.1. Event Generation

Around 30 different MC generators are in use in ATLAS [2]. The standard multi-purpose MC generators, such as PYTHIA 6 and 8, HERWIG, HERWIG++ and Sherpa, are fully integrated into the ATLAS software framework. Other MC generators, such as POWHEG-BOX, Alpgen and MC@NLO, provide only 4-vectors of the hard process. The hadronisation is done in the ATLAS framework using PYTHIA 6, PYTHIA 8, HERWIG or HERWIG++. The 4-vectors are either pre-generated, stored on grid [3]¹ storage and downloaded by the job, or generated on-the-fly before the hadronisation process. The latter simplifies the event generation process significantly, and collects all relevant information in one place.

The common job option files for the generator configuration and the specific job option files for the different hard processes and possible filters are stored in Subversion [4] in order to provide versioning and ensure reproducibility. The distribution of the necessary configuration files is decoupled from the software release in order to accommodate frequent updates. These are fetched on demand by the job as an archive file via HTTP from a CERN web server. In the future the distribution will be done via CVMFS in order to reduce the number of files in the local working directory and to unify the software and configuration distribution.

By default 5000 events are requested per job, giving a file size of about 100 MB, which can be transferred easily on the grid. The memory consumption of MC generators is very low, usually below 0.5 – 1 GB. The running time for different processes and generators can vary significantly. Simple final states, or hadronisation-only, run within a few minutes. More complex final states or processes with applied generator-filters with low efficiency, run longer. The number of events per jobs can be used to adjust the running time. In order to provide a good turn around and to avoid cluttering job slots, it should be below 8 hours, but can be extended to one or two days. To reduce the running time even further, pre-made integration grids can be used for Sherpa, Alpgen and MadGraph.

2.2. Detector Simulation

The ATLAS detector simulation is based on Geant4 [5]. All stable particles from the event generation are tracked through the ATLAS geometry. The QGSP_BERT physics list [6] is used to simulate interactions in the material.

Different approaches have been developed to speed up the detector simulation:

Frozen Showers [7]: The tails of the electro-magnetic showers are not simulated by Geant4 but substituted with pre-made energy deposits. Since the 2011 MC campaign² frozen showers are used in the forward calorimeters. For the 2012 MC campaign the simulation time is reduced by 25%.

Fast Calorimeter Simulation [2]: The interaction of all particles except muons is parametrised in the calorimeters based on fully simulated showers for electrons and charged pions. The fast calorimeter simulation has been used since the end of the 2010 MC campaign. For the 2012 MC campaign the simulation time is reduced by a factor 10.

During LS1, the integrated simulation framework (ISF) [8] has been developed to better integrate the different simulation flavours. It allows the possibility to choose the simulation flavour not only by sub-detector but also by cones around a particle.

¹ See Section 4 for more details.

² See Section 3 for more details on MC campaigns.

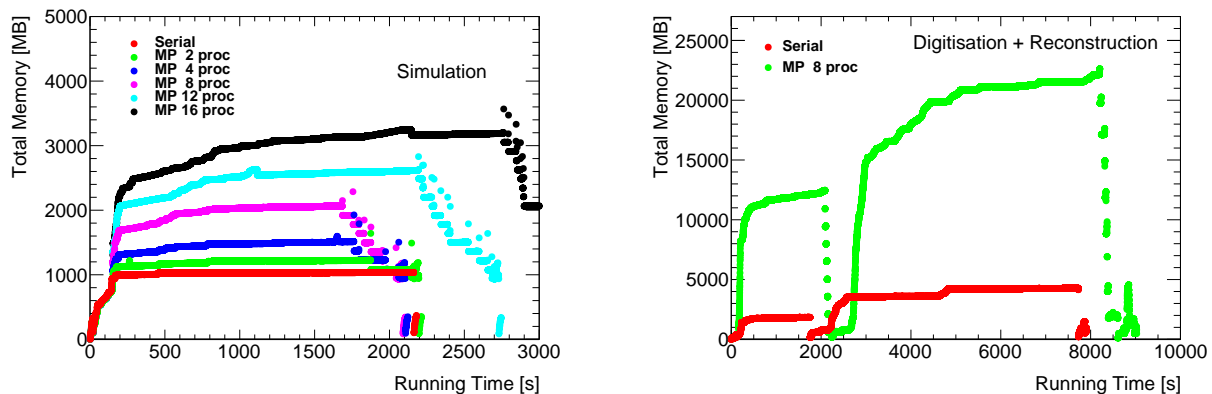


Figure 1. Total memory consumption as a function of the running time of (left) a simulation job and (right) of a digitisation+reconstruction job for different number of cores used.

By default, 100 events are simulated for a Geant4 + Frozen Shower job, which run for around 7 hours. The output files size is about 80 MB and files are merged and stored on tape, giving 1000 events per file, in order to give a better file size for grid transfer and tape storage. For Geant4 + Fast Calorimeter Simulation 1000 events are simulated per job, which run for around 6 hours. The file size is about 0.5 GB due to the reduced calorimeter information.

Simulation jobs are run in 64 bit, which give better performance compared to 32 bit, increasing the memory consumption only slightly. Further performance improvements have been achieved during the 2012 MC campaign by using the Intel math library and switching to the SIMD-oriented Fast Mersenne Twister random number generator.

The memory requirement per job is very moderate and is around 1 GB. Recently the detector simulation was successfully validated for athenaMP [9], the implementation of multi-core usage in the ATLAS framework. Tests have shown, as expected, that a large amount of the memory (0.8 GB) is shared among the different cores (see figure 1). Only 0.16 GB is needed per core for non-shareable data.

2.3. Digitisation

At the digitisation step the response of the readout electronics is simulated. In addition, the effect of multiple pp interactions per bunch crossing, named pile-up, is simulated [10].

Pre-simulated minimum bias events are overlayed on each signal event. The number of pile-up events per signal event ($\langle\mu\rangle$) is Poisson distributed. The average is either fixed to a single value or follows a predefined $\langle\mu\rangle$ profile, similar to the one observed in data. In the latter case the full $\langle\mu\rangle$ profile is sampled over 5000 signal events. This implies that smaller samples should be a multiple of 5000 events.

The minimum bias pile-up sample is separated into a low-Q and a high-Q part ($Q=35$ GeV) in order to reuse low-Q events in one job and to increase the available statistics of the high-Q events. Each of the low/high-Q samples consists of 10×10^6 events. For the 2012 MC campaign the $\langle\mu\rangle$ profile spans from $\langle\mu\rangle = 0$ to $\langle\mu\rangle = 40$, with an average $\langle\mu\rangle$ of 20. The total size of the minimum bias pile-up samples is 6.3 TB (1.5 TB for the low-Q sample and 4.8 TB for the high-Q sample). The number of events per file is 5000 and 500 for the low-Q and high-Q sample, respectively. A standard digitisation job processes 500 signal events and uses ten pile-up files, five from the low Q and five from the high Q sample. Per job, 4.8 GB of input files are needed, dominated by the pile-up files.

In order to avoid massive file transfers from Tier-1 to Tier-2 sites, the minimum bias pile-up

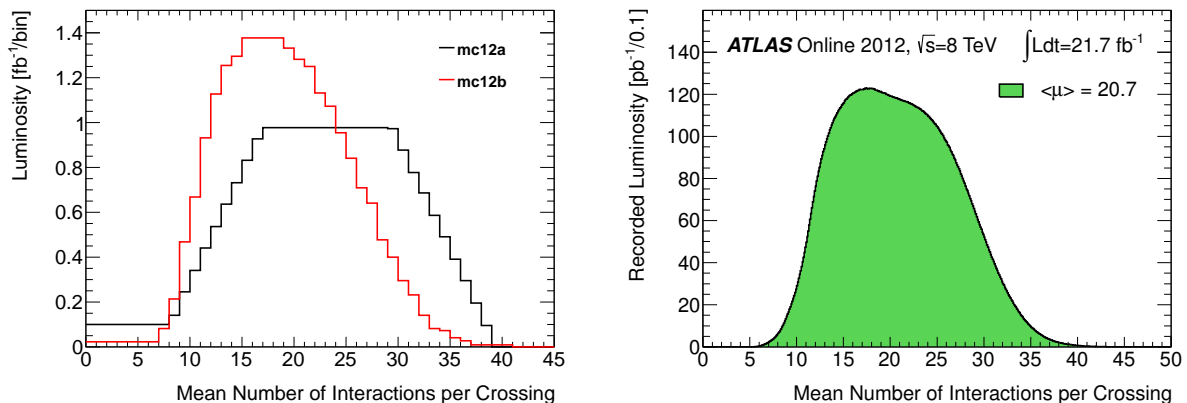


Figure 2. Pile-up profile as used in 2012 MC campaign for the mc12a and mc12b setups (left) and observed in data (right).

samples are not only stored at Tier-1 sites but also at larger Tier-2 sites. The total occupied disk space is 0.3 – 0.4 PB. For significant changes in detector geometry, conditions or centre-of-mass energy, the full minimum bias pile-up sample needs to be re-simulated and distributed to the grid. For smaller campaigns, eg upgrade production, either the number of events or the number of replicas on the grid is reduced in order to minimise the occupied disk space.

The $\langle\mu\rangle$ profiles, as observed in the 2012 data and used in the two 2012 MC campaigns mc12a and mc12b, are shown in figure 2. The mc12a $\langle\mu\rangle$ profile was a best estimate of the expected data $\langle\mu\rangle$ profile prepared before the start of data taking. After the 2012 data taking the $\langle\mu\rangle$ profile was updated for the mc12b campaign to better match the observed data $\langle\mu\rangle$ profile and hence increase the statistical significance of the simulated events.

2.4. Reconstruction

The simulated MC events are reconstructed in the same way as data recorded by the ATLAS detector. In addition the trigger system is simulated. The main reconstruction process produces the Event Summary Data (ESD), which contains low level quantities as tracks and clusters and high level quantities as jet, electron, photon, muon and tau candidates. In a second, fast step the Analysis Object Data (AOD) is produced, which contains mainly high level information.

For the 2012 MC campaigns at 8 TeV, 500 events are reconstructed per job, resulting in an average running time of about 8 hours. The memory requirements are stringent. Reconstruction is run on 32 bit and needs 3.6 – 3.8 GB. Running in 64 bit would raise this above 4 GB, which is provided only by a few grid sites for a limited number of nodes.

Recently the digitisation and reconstruction were successfully validated for athenaMP. Tests have shown that the amount of shareable memory is much smaller compared to simulation, but still promising. In 64 bit the memory consumption is 4.3 GB for a single-core job and 22.6 GB for an eight-core job, respectively (see figure 1). 2.8 GB per core is needed, which is already significant less than the 4.3 GB of a single-core job. The aim is 2 GB per core.

By default, ESD files are stored temporary and only AOD files are stored on the grid. On request ESD files can be permanently stored, but must be moved to group disk space. Ntuples and derived formats are produced by group production [11] either from ESD or AOD files. Exceptions are made for certain work flows. For example the heavy ion group works with ESD and ntuple files, which are both produced in the reconstruction step.

AOD files are merged into files of 5000 events in order to increase the file size for better grid

transfers and tape storage. The average file size is about 2.2 GB.

2.5. Digitisation and Reconstruction

Every MC production step can be run as an independent grid job storing its output files on the grid. There are sometimes benefits to run more than one step in one job.

One example is the fast detector simulation, where the complete simulation chain can be run in one grid job. This avoids storing intermediate output files, as the large digitisation output, only keeping the final AOD and therefore simplifies the data management.

Another example is the combination of the digitisation and reconstruction step introduced for the 2011 MC campaign. The pile-up profile changed significantly between the start and end of the data taking period. At the same time the trigger menu was adjusted to the higher collision rates. This was implemented into MC production by splitting the pile-up simulation in the digitisation job into subsets identified by a run number. This number was used in reconstruction to pick the corresponding trigger menu.

3. Production Campaigns

MC production is divided into campaigns, where the centre-of-mass energy, geometry and conditions used in production correspond to a running period of the LHC. Major campaigns correspond to the calendar year. Minor campaign versions usually reflect improvements in reconstruction software, trigger menu or pile-up simulation. In contrast to data taking and reprocessing campaigns, MC campaigns are run as long as MC events are needed for data analysis.

In the following, the last two major MC campaigns are described in more detail including the produced number of events to date:

mc11: Simulation configuration for 7 TeV in 2011. Produced 2.4×10^9 full simulation and 2.1×10^9 fast simulation events.

mc11a: Digitisation+reconstruction configuration for 7 TeV in 2011, before start of data taking: minimum bias pile-up samples generated with PYTHIA 8; pile-up profile estimated from LHC run plans; beam spot information estimated from 2010 data. Produced 0.8×10^9 events.

mc11b: Improved digitisation+reconstruction configuration based on mc11a: updated trigger menu and improved pile-up profile based on four run periods. Produced 1.0×10^9 events.

mc11c: Improved digitisation+reconstruction configuration based on mc11b: minimum bias pile-up samples generated with PYTHIA 6. Produced 4.8×10^9 events.

mc12: Simulation configuration for 8 TeV in 2012. Produced 3.8×10^9 full simulation and 3.0×10^9 fast simulation events.

mc12a: Digitisation+reconstruction configuration for 8 TeV in 2012, before start of data taking: minimum bias pile-up samples generated with PYTHIA 8; pile-up profile estimated from LHC run plans; beam spot information estimated from 2011 data. Produced 5.9×10^9 events.

mc12b: Improved digitisation+reconstruction configuration based on mc12a after stop of data taking: updated, realistic beam spot information and pile-up profile. Produced 0.5×10^9 events.

mc12c: Improved geometry description for precision measurements. Simulation configuration based on mc12 and digitisation+reconstruction configuration based on mc12b. Produced 0.2×10^9 events.

4. Production on the Grid

For ATLAS the grid is organised in a tier structure. The Tier 0 is located at CERN. ATLAS has currently ten Tier 1 sites and around 70 Tier 2 sites. Twenty of the ATLAS Tier 3 sites are

enabled for official MC production. These sites provide on average 90 000 single-core job slots to ATLAS for MC production.

ATLAS distributed computing is based on grid technology, but other technologies, such as cloud computing, have been evaluated. MC production has been successfully run on different cloud implementations, as Amazon's EC2 cloud, Google's computing engine and various open alternatives [12]. Also, ATLAS investigates the use of opportunistic computing resources. During LS1 the ATLAS online farm is used for MC production using cloud technologies when not used otherwise [13]. Another area of research and development is the usage of high performance computing resources.

5. Upgrade Production

MC simulation at 14 TeV with the ATLAS detector and foreseen detector upgrades is needed for Run 2 preparations or to study performance and physics reach for Phase 2 and beyond. During LS1 two major upgrade campaigns are being done:

ATLAS+IBL: For Run 2 the pixel tracker will be extended by the Insertable b-Layer (IBL).

ATLAS+ITK: For Phase 2 the inner detector will be replaced by a silicon only tracker.

For Run 2 the LHC bunch spacing will most likely be 25 ns, while 50 ns is studied as an alternative. The anticipated pile-up level will range between 40 and 60. For ATLAS+IBL the following $\langle\mu\rangle$ values are studied: 20, 40, 60 and 80.

At the same pile-up level, digitisation and reconstruction will run longer and need more memory for the 25 ns bunch-spacing configuration since the level of out-of-time pile-up will be higher. When scaling the 8 TeV reconstruction setup to 14 TeV, the memory consumption for reconstruction for $\langle\mu\rangle=60$ starts to exceed the 4 GB limitation. The trigger menu was reduced to bring back the memory consumption below 4 GB. This is not sufficient for the 25 ns bunch-spacing $\langle\mu\rangle=80$ configuration, which need to run on dedicated high memory queues providing up to 6 GB of memory.

For Phase 2 the LHC bunch spacing will be 25 ns and the anticipated pile-up level will range between 80 and 200. For ATLAS+ITK simulation the following average $\langle\mu\rangle$ values are studies: 80, 140 and 200. Trigger simulation is not supported at the moment reducing the memory required for reconstruction significantly and keeping it below 4 GB in 32 bit mode.

6. Validation

Thorough validation of the software is mandatory before using a software release in production. Physics validation is one of the final steps for validating new software releases and configurations. Two million events split over different physics and performance samples are processed for each validation. These samples are checked and compared to previous validations by physics and performance contacts, concentrating on relevant distributions using realistic event selections.

The aim of this validation is manifolded and only a few aims are discussed here: The statistics are sufficiently large to show rare run time problems, eg within Geant4. Potential problems originating from the grid environment can be detected, as the local environment differs from the grid one in a few aspects. Finally, since mainly physics quantities are checked, the effect of expected changes on physics results can be estimated.

New software releases for simulation and digitisation+reconstruction are regularly validated. Besides these standard validation tasks, performance improvements, such as the usage of the Intel math library, improved random number generators or improved database access are also validated. Another area is the validation of simulation-related changes, eg updated geometry descriptions and different Geant4 physics lists, in order to improved data-to-MC agreement. The validation of Geant4 patches is done using dedicated project releases to avoid polluting production releases.

7. Zero-Bias Data Overlay

The accuracy of the pile-up simulation can be improved by using data events recorded in the zero-bias trigger stream [10]. This approach has the advantage that the background is determined in a truly data-driven way, and the dependence on MC event generator modelling and MC simulation is eliminated. These recorded zero-bias events are overlaid one-by-one on top of simulated signal events. The signal must be simulated with the same conditions, including the same vertex position of the primary vertex, as the background event, in order to ensure detector consistency. To ensure this coupling of simulated and zero-bias event, the simulation and overlay steps run in the same job.

The selection of the zero-bias overlay events needs to be done carefully. The zero-bias events are grouped into sets of 50 000 events each. Within each set the full pile-up distribution and detector conditions are sampled, hence the signal sample size needs to be a multiple of 50 000 events. Within one set, the events are randomly selected weighted by the instantaneous luminosity, sampling over the full data taking period.

Each zero-bias file contains 100 events, matching the number of events per simulation job. The file size is 0.3 GB. The input data for each simulation+overlay job is 0.4 GB, dominated by the size of the zero bias data, but much smaller compared to the standard digitisation job.

The full 2012 zero-bias overlay data set consists of 50×10^6 events. The total size is 160 TB, which is much bigger than the minimum bias pile-up sample. The storage of the full sample at every Tier-1 site is difficult and a more efficient distribution scheme needs to be defined.

This method is extensively used in ATLAS for heavy ion event simulation. The enormous particle background in a heavy ion collision can be taken directly from the data, and an additional collision can be overlaid on top of it. Performance and physics groups started to look into zero-bias pile-up simulation for high-energy pp data analysis.

8. Summary

The ATLAS MC production has been running, and continues to run, very successfully during Run 1 and Long Shutdown 1, and has produced more than 6.2×10^9 full simulation and 5.1×10^9 fast simulation events. Software and setups have been improved and adjusted where necessary in order to meet new requirements, or improve performance. Special activities, such as upgrade production or zero-bias data overlay, have been successfully run within the MC production.

Preparations for a large scale data challenge campaign planned for 2014 are underway. The aim is to test and evaluate new software features and LHC data taking conditions of Run 2, to be prepared for the next data taking period.

References

- [1] ATLAS Collaboration 2008 *JINST* **3** S08003
- [2] ATLAS Collaboration 2010 *Eur. Phys. J.* **C70** 823 (*Preprint* 1005.4568)
- [3] The LCG TDR Editorial Board 2005 LHC Computing Grid, Technical Design Report URL <http://cdsweb.cern.ch/record/840543>
- [4] Apache Subversion: <http://subversion.apache.org/>
- [5] Agostinelli S *et al.* (GEANT4) 2003 *Nucl. Instrum. Meth.* **A506** 250
- [6] Ribon A *et al.* 2010 Status of Geant4 hadronic physics for the simulation of LHC experiments at the LHC physics program CERN-LCGAPP-2010-02
- [7] Barberio E *et al.* 2009 *J.Phys.Conf.Ser.* **160** 012082
- [8] Debenedetti C 2013 *This proceedings*
- [9] Binet S *et al.* 2012 *J.Phys.Conf.Ser.* **368** 012018
- [10] Marshal Z 2013 *This proceedings*
- [11] Laycock P *et al.* 2013 *This proceedings*
- [12] Panitkin S *et al.* 2013 *This proceedings*
- [13] Ballestrero S *et al.* 2013 *This proceedings*