Simulation strategies for the ATLAS Experiment at LHC

A.Rimoldi

University of Pavia & INFN

On behalf of the ATLAS Collaboration

adele.rimoldi@cern.ch

Abstract. A fully detailed simulation of the ATLAS experiment, in operation at the CERN Large Hadron Collider (LHC), has been implemented using the Geant4 toolkit. The simulation program, built within the ATLAS common framework (Athena) is being used for large-scale production of events on the LHC Computing Grid. Simulation software requires many components, from generators packages for simulation of particle collisions, to packages simulating the response of the various detectors and triggers. All of these components are steered by the ATLAS simulation infrastructure. The latest developments in this project addressed a better representation of the reality of the detector in all the possible details by providing increased functionality and robustness. The full simulation flow is constantly monitored and profiled. Performance improvements permit more effective use of the available resources without degradation in the quality and accuracy of the simulation. In this presentation, after a short introduction of the ATLAS detector itself, emphasis is put on validation efforts, performance tests, CPU and memory optimizations through the description of the entire process.

ATL-SOFT-PROC-2011-016
13 January 2011

1. Introduction

ATLAS [1] is the biggest general-purpose detector ever built in particle physics, operational at the Large Hadron Collider (LHC) [2] at the CERN Laboratories in Geneva. The detector has been built to collect data from proton-proton collisions with center-of-mass energies up to 14 TeV, as well as 5.5 TeV per nucleon pair in heavy ion (Pb-Pb) collisions. The ATLAS detector was built with a tremendous effort for two decades (1990- 2008) by a collaboration composed by ~3000 physicists, it weighs ~7000 tons, its height is ~25 m and length is ~46 m. The total number of equipped channels is $\sim 10^8$. The ATLAS detector is currently described in detail in Ref. [1]. A cut-away view of the detector is shown in Fig.1. The detector has been operational since 2008 and the experiment is currently performing in agreement with design specifications. The understanding of the detector response and Monte Carlo tools is extremely good and beyond expectations for this early phase. In this paper we will comment about the current status of simulation and its validation.

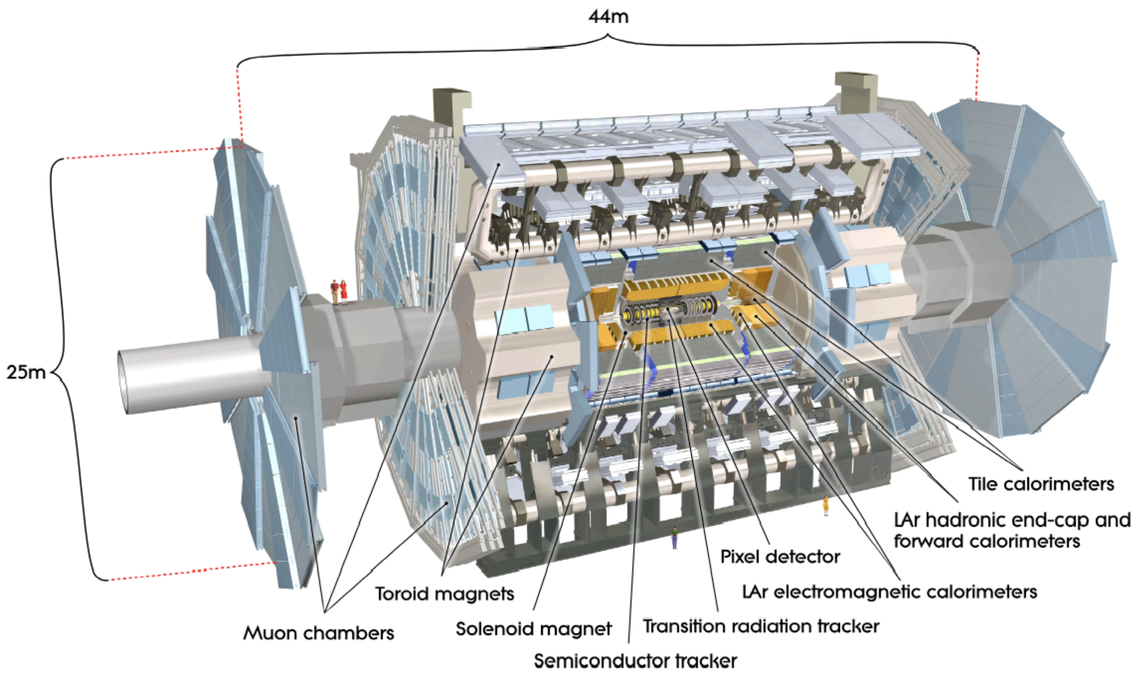

Fig.1 Atlas Detector view

2. Simulation in ATLAS

Simulation is a component of the ATLAS Software Project and it uses the Geant4 simulation toolkit [3]. The simulation software developed in ATLAS has been used for large scale productions at the Computing Grid through production of about 10^9 full physics events [4]. This process is generally divided into three steps, though they may be combined into a single job: 1) event generation 2) physics and detector response 3) digitization of physics quantities with production of final output. Fig.2 shows the flow of ATLAS Simulation software from event generation (top left) to reconstruction (top right).

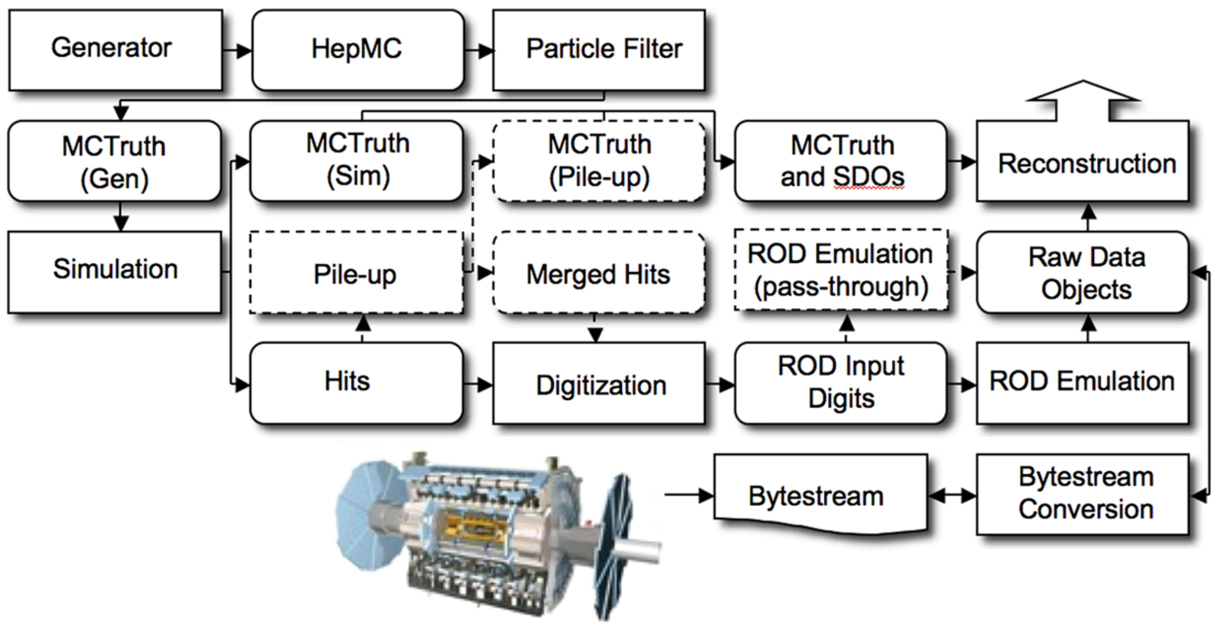

Fig.2 The ATLAS Simulation software flow

The event generation is provided by about 40 different generators interfaced to the Athena framework, all of which are continually validated for physics performance. These generators cover a wide range of LHC physics processes, as well as configurable "particle guns". Events from these generators can either be written to a file for further processing, or passed directly into the further simulation stages. Each event contains particles from a single interaction with vertex at (0,0,0). All the beam properties are applied at a later stage before being passed through Geant4. The physical construction parameters and conditions data are contained in databases to allow an identical description for simulation digitization and reconstruction processes and they are connected at runtime. Test stand layouts as well as some installation configurations (as alternate layouts in the commissioning phase) and the data-taking setup are all available for running. A description of the full ATLAS detector with inclusion of forward and very forward detectors has been made available to simulation. Misalignments in the realization of the detector, material distortions and extra materials are allowed, described and made as well available in the full simulation.
The current version of Geant4 used for the ongoing productions is 9.2.p2.a4: fixes to the main version (9.2) are provided by both the Geant4 team (p2) and the ATLAS team (a4). The simulation chain works under the ATLAS framework (Athena) [5] and the current version is based on Atlas Software release 15.6.12.5, ATLAS-GEO-16-00-00 geometry release and condition data as in OFLCOND-SRD-BS7T-02 database.

3. ATLAS Core Simulation

In order to provide Python flexibility to the Geant4 simulation, an additional layer of infrastructure is necessary. Standard Geant4 applications normally run from compiled C++ and any modification of parameters or geometry in the code requires a recompilation. The framework for ATLAS Geant4 simulation wraps several Geant4 classes (for volumes, sensitive detectors, physics process definitions, etc.) in order to allow selections and configuration without recompilation of any libraries. This additional framework for simulation helps in ease transitions between Geant4 and Athena standards for geometry as well as in cataloguing the options available to the user. The detector description is configured with Reflex dictionaries and simulation framework catalogues, before it is built in Geant4. Once the configuration is set in the Python layer, the objects created by the infrastructure are translated into their Geant4 equivalents and loaded. The event flow is then monitored in Athena. Simulation parameters as range cuts can be tuned in the simulation via the Python configuration.

4. ATLAS Fast Simulations

Much effort has been dedicated to development of fast simulations, to optimize the necessary limited computing resources needed to model the complexity of both detector and physics. For some studies increased statistics or faster turn-around are needed. To make such fast simulation possible, several avenues have been pursued to reduce the computational requirements of ATLAS simulation jobs.
ATLAS developed three approaches to fast simulations named ATLFAST I [6,7], ATLFAST II, ATLFASTIIF [8] and FAST-G4 (frozen showers approach) [9]. In ATLFAST I, formerly developed for physics parameter space scans, objects are smeared by detector resolution; in ATLFAST II parameterized particle showers are used (FastCaloSim) and simplified detector description tracking (Fatras). In FAST-G4 the standard full simulation uses pre-simulated shower libraries for energies from 1 GeV to 10 MeV, the rest being dealt as standard full simulation with Geant4. The relative gain in processing time for the same sample of events (ttbar) is shown in Fig. 3. Many comparisons have been made for the different “faster” approaches to simulation on key variables, as shown in the following Fig. 4, all indicating a substantial good agreement among the different flavours of fast simulation with respect to full simulation. Although still in development, the present tests show that for FAST-G4/ full simulation comparisons, in the forward region, an agreement of ~1% is achieved.

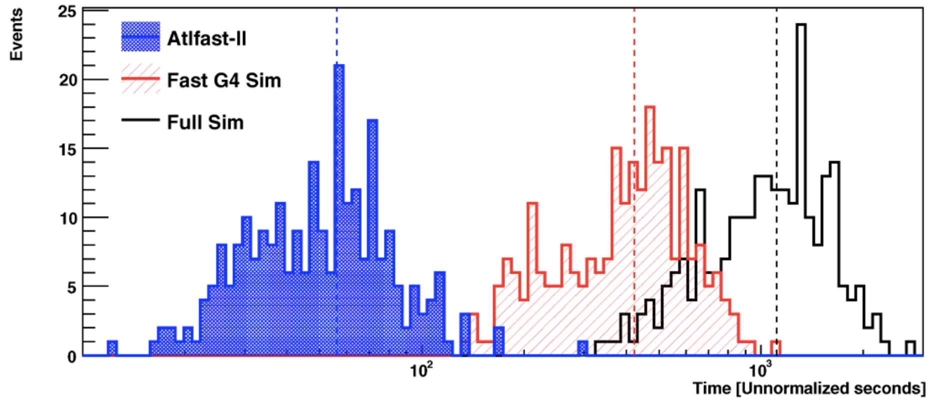

Fig.3 Distributions of CPU time for 250 ttbar events in full, FASTG4 and ATLFAST II simulations. Vertical dotted lines denote the averages of the distributions

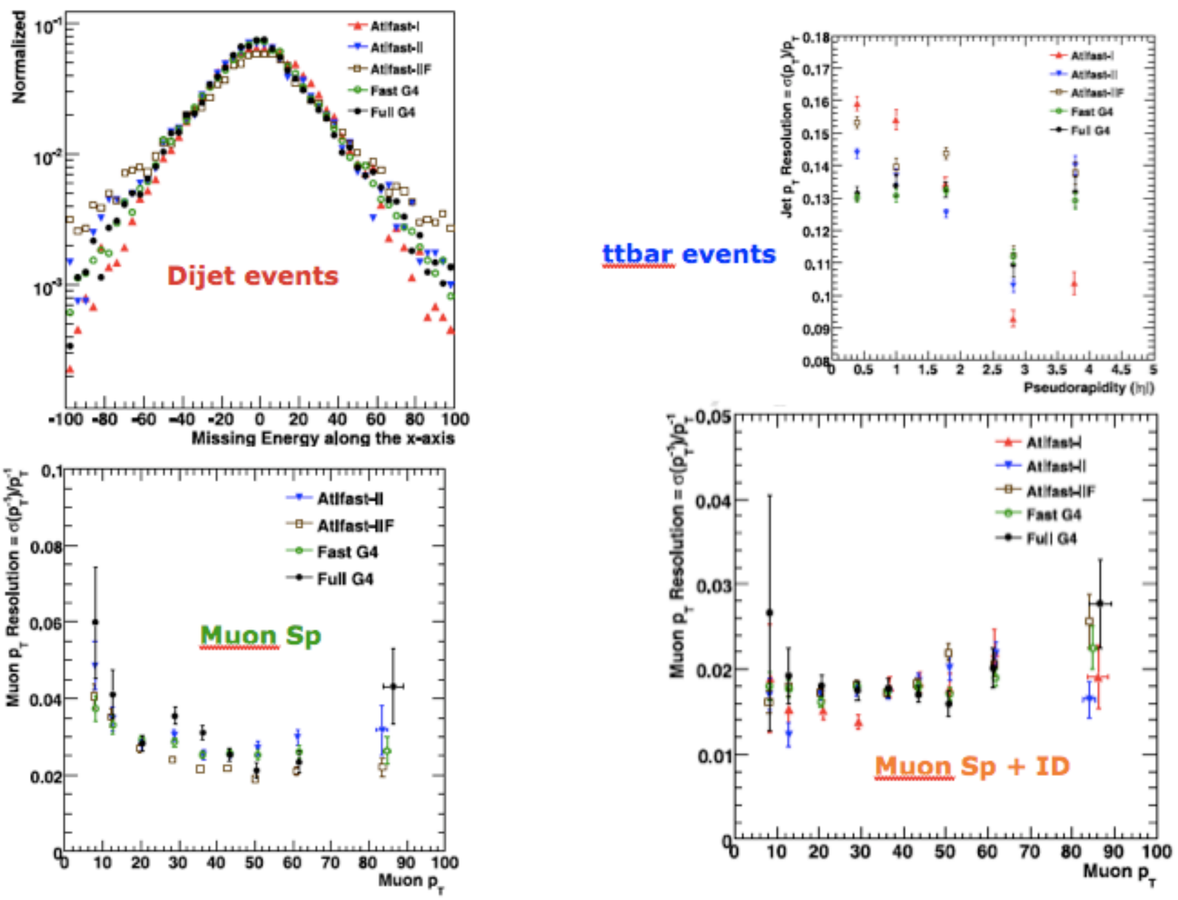

Fig. 4 Fast simulations and full simulation comparisons for missing transverse energy in dijet events, pseudorapidity in ttbar events, as well as reconstructed muon pt resolutions using muon spectrometer only (Muon Sp) or inner detector (ID) and muon spectrometer combined

5. Simulation as an evolving Project

Validation of the ATLAS simulation chain is done starting from technical tests on software performance and then moved to physics performance, this one being tested with comparisons to data through a wide set of key physical variables. Test results in the first phase are evaluated to check for stability and performance. A complex testing infrastructure has been set up in ATLAS to early detect possible failures [10]. Since the inception of this project emphasis was put in optimizing use of disk space and CPU versus time. At each software release a careful validation of some performance indicators was put in place. Fig. 5 shows two of these indicators, the mean event size and CPU measurements as a function of release number for different categories of physics events (SUSY, jets, Higgs into four leptons, Zeta into two leptons, minimum bias).

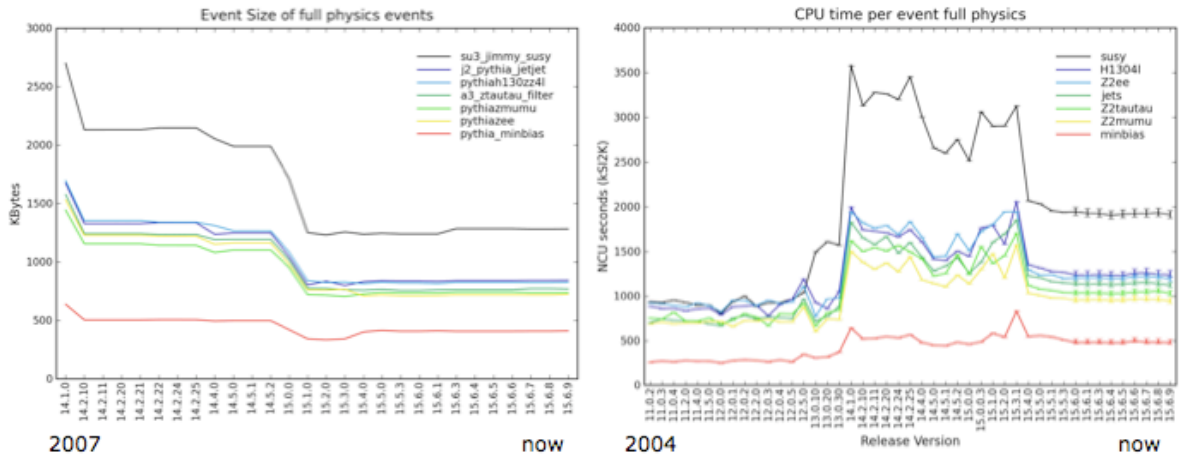

Fig. 5 Monitoring of performance indicators for the simulation framework as a function of the release number

The variable mean event size for the different physics samples has been collected since 2007 (ATLAS Release 14.1) showing, after a consistent optimization done in the containers of all subdetectors, a constant behavior in the last year, decreasing its mean value on the different physics samples at its half. The CPU time per event, monitored since 2004, shows a large increase for all physics samples due to adoption of the QGSP_BERT Geant4 physics list. Due to optimization efforts, this increase in CPU has been largely reduced since initial adoption, and the past year shows stability and even a slight decrease in CPU requirements. Other quantities that were under strict observation were the Geant4 memory usage (250 MB) as a function of time (or release number) and the memory usage in a job (700MB).

6. Simulation/Data matching

After the successful start of operation of the LHC, comparisons data/MC were made available at the colliding energy of 7 TeV. The following example shows the calorimeter E_t for the cluster of electron candidates after preselection for data and Monte Carlo candidates, broken down into the various signal and background components. The total number of Monte Carlo candidates is normalised to the number of observed data candidates. In Fig. 6, the total QCD background is broken down into its constituents: hadrons misidentified as electrons, electrons from conversions, and electrons from semi-leptonic decays of heavy.

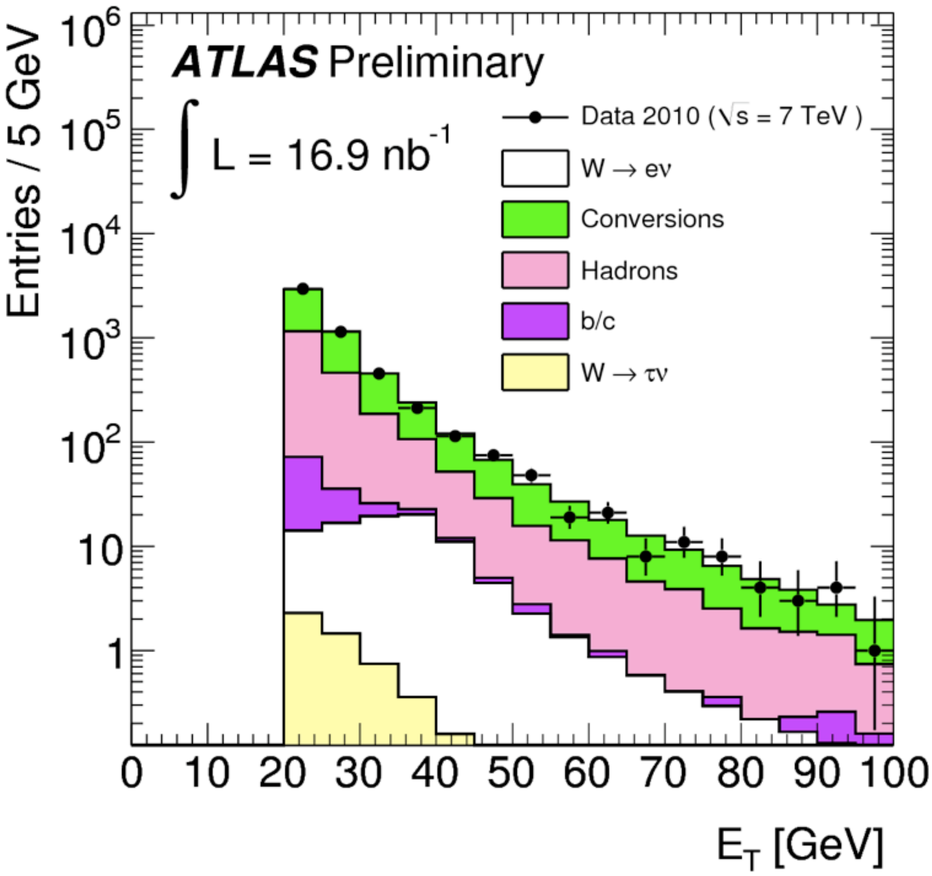

Fig. 6 Calorimeter E_t for the cluster of electron candidates after pre-selection for data and Monte Carlo candidates, broken down into the various signal and background components at $\sqrt{s}$ = 7 TeV and $\int L$ = 16.9 nb^{-1}

Data to simulation comparison is impressive for the shown distribution over more than 6 orders of magnitude. The same matching quality is observed for track variables in minimum bias samples where, as in Fig.7, data and MC agree very well for average number of Pixel and SCT hits of the inner detector in the whole eta interval (positive and negative) [11].

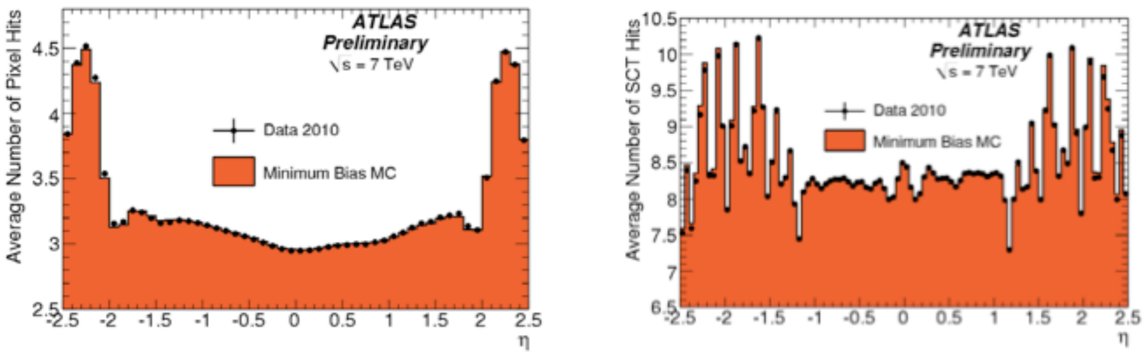

Fig. 7 Simulation data comparisons at 7 TeV for minimum bias events for average number of pixel (left) and SCT (right) hits

7. Digitization Pile-up and Overlay

The ATLAS digitization software transforms the hits into detector responses. The peculiarities of each subdetector charge collection including cross-talk, electronic noise and channel-dependent variations are modeled in subdetector–specific digitization software. The various subdetector digitization packages are steered by a top-level Python digitization package. Dead channels and noise rate are read from a database to reproduce run-dependent conditions. The optimization and tuning phase of each algorithm is ongoing at subdetector level.

To simulate pile-up, various types of events are read by mixing their hits. The different types could be specified at runtime so that beam-gas, beam-halo, cavern background (neutron haze) and additional interactions off-time are all overlaid to hard scattering events. Two approaches are presently adopted in ATLAS:

1) all above mentioned processes are simulated in Monte Carlo and then mixed together in proper ratios with realistic timing in the so called pile-up [12]

2) After simulation of signal event in Monte Carlo an overlay with it is done with a "random" data event to include all background (overlay).

While the first approach is standard, the second is what is needed is to merge data and Monte Carlo and produce a data-like event to be passed to reconstruction. Pile-up in ATLAS is deployed and in validation phase, overlay is developed and it will be soon deployed, after the necessary validation.

8. Conclusions

In this early phase of ATLAS operation we observed that simulation is extremely close to data. Event production is progressing at a rate of 7-8 M events/day with substantially no failures or recursive errors. The optimization phase for this project is not over, refinements in algorithms for each detector are still in development. To further match simulation with data, all the new versions of Geant4 are constantly under test and are evaluated through technical and physical validation for possible future deployment. Pileup and overlay are ready or in the deployment phase, while fast simulations are in the validation phase and ready to be used for Grid production.

References

[1] The ATLAS Collaboration, Performance of the ATLAS Detector using First Collision Data JHEP 1009:056, 2010

[2] Evans, L. The Large Hadron Collider. New J. Phys. 9, 335 (2007)

[3] Agostinelli et al., GEANT4: a Simulation Toolkit, Nucl. Instr. & Meth. A506 (2003) 250-303

[4] The ATLAS Collaboration, The ATLAS Simulation Infrastructure, EPJC, vol 70, n3, December1, 2010

[5] Obreshkov E. et al., Organization and management of ATLAS Offline software releases, Nucl. Instr. & Meth. A584 (2008) 244-251.

[6] Richter-Was, E. et al ATLFAST 2.0 a fast simulation package for ATLAS; CERN-ATL-PHYS-98-131 (1998),

[7] Dean, S. et al Athena-ATLFAST http://www.hep.ucl.ac.uk/atlas/atlfast, 2008

[8] The ATLAS Track Extrapolation Package, ATL-SOFT-PUB-2007-05 (2007)

[9] Barberio et al., Fast simulation of electromagnetic showers in the ATLAS calorimeter: Frozen showers. J.Phys.Conf.Ser.160:012082,2009.

[10] Nevski P., Large scale data movement on the GRID (2006)

[11] The ATLAS Collaboration, Charged-particle multiplicities in pp interactions at sqrt(s) = 900 GeV measured with the ATLAS detector at the LHC Phys. Lett. B688: 21-42,2010

[12] Chapman,J. et al., The ATLAS detector digitization project for 2009 data taking, J.Phys.Conf.Ser.219:032031,2010.