The ATLAS simulation infrastructure

Aad, G.; et al., [Unknown]; Bentvelsen, S.; Bobbink, G.J.; Bos, K.; Boterenbrood, H.; Colijn, A.P.; de Jong, P.; Doxiadis, A.; Ferrari, P.; Garitaonandia, H.; Gosselink, M.; Hartjes, F.; Hessey, N.P.; Igonkina, O.; Kayl, M.S.; Klous, S.; Kluit, P.; Koffeman, E.; Koutsman, A.; Lee, H.; Liebig, W.; Mechnich, J.; Mussche, I.; Ottersbach, J.P.; Rijpstra, M.; Ruckstuhl, N.; Salamanna, G.; Sandstroem, R.; Snuverink, J.; Tsiakiris, M.; Turlay, E.; van der Graaf, H.; van der Kraaij, E.; van der Poel, E.; van Kesteren, Z.; van Vulpen, I.; Verkerke, W.; Vermeulen, J.C.; Vreeswijk, M.

DOI
10.1140/epjc/s10052-010-1429-9

Publication date
2010

Document Version
Final published version

Published in
European Physical Journal C

Citation for published version (APA):
Aad, G., et al., U., Bentvelsen, S., Bobbink, G. J., Bos, K., Boterenbrood, H., Colijn, A. P., de Jong, P., Doxiadis, A., Ferrari, P., Garitaonandia, H., Gosselink, M., Hartjes, F., Hessey, N. P., Igonkina, O., Kayl, M. S., Klous, S., Kluit, P., Koffeman, E., ... Vreeswijk, M. (2010). The ATLAS simulation infrastructure. European Physical Journal C, 70(3), 823-874.
https://doi.org/10.1140/epjc/s10052-010-1429-9

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).
The ATLAS Simulation Infrastructure

The ATLAS Collaboration

G. Aad, B. Abbott, J. Abdallah, A. Abdelalim, A. Abdesselam, O. Abdinov, B. Abi, M. Abolins, H. Abramowicz, H. Abreu, B.S. Acharya, D.L. Adams, T.N. Addy, J. Adelman, C. Adorján, P. Adragna, T. Adye, S. Afsks, J.A. Aguilar-Saavedra, M. Aharrouche, S.P. Ahlen, F. Ahles, A. Ahmad, H. Ahmed, M. Ahsan, G. Aielli, T. Akdogan, T.P.A. Åkesson, G. Akimoto, A.V. Akimov, A. Aktas, M.S. Alam, M.A. Alam, S. Albrand, M. Aleksa, I.N. Aleksandrov, C. Alexa, G. Alexander, G. Alexandre, T. Alexopoulos, M. Alhroob, M. Aliev, G. Alimonti, J. Alison, M. Aliyev, P.P. Allport, S.E. Allwood-Spiers, J. Almond, A. Aloisio, T2a, T2b, R. Alonso, J. Alonso, M.G. Alviggi, T2a, T2b, S. Amako, C. Ameling, A. Amorim, N. Amram, C. Anastopoulos, T. Andeen, C.F. Anders, K.J. Anderson, A. Andreazza, V. Andrieu, X.S. Anduaga, A. Angerami, F. Angghinolli, N. Anjos, A. Annoij, A. Atonoki, M. Antongiuli, S. Antonelli, J. Antos, B. Antunovic, F. Anulli, S. Aoun, G. Arabidze, I. Arcacena, Y. Ara, T.H. Arce, J.P. Archambault, S. Arfaoui, J.-F. Arguin, T. Argyropoulos, M. Arlik, A.J. Armbruster, O. Arnaez, C. Arnauld, A. Artamonov, A. Arutinov, M. Asai, S. Asai, R. Asfandiyarov, S. Ask, B. Åsman, V. Ask, A. Aslaksen, X. Astară, A. Astvatsaturov, G. Atoian, V. Aubert, K. Auge, M. Aurousseau, N. Austin, G. Avolio, R. Avramidou, D. Axen, G. Ay, G. Azuelos, Y. Azuma, M.A. Baak, A.M. Bach, H. Bachus, M. Backes, E. Badescu, P. Bagnai, S. Baines, T. Bain, J.T. Baines, O.K. Baker, M.D. Baker, S. Baker, F. Baltasar, D. Santos, Pedroso, E. Banas, P. Banerjee, S. Banerjee, J. Banfi, S. Banfi, A. Bangerter, V. Bansal, S. Baranov, S. Baranov, A. Barashkou, T. Barber, E.L. Barberio, D. Barberis, D. Bard, D. Bard, D. Barillari, M. Barisonzi, T. Barklow, N. Barlow, B.M. Barnett, R.M. Barnett, A. Baronecchi, A.J. Barr, F. Barro, J. Barreiro Guimaraes, da Costa, P. Barrillon, S. Bartoldus, D. Bartsch, R.L. Bates, L. Batkova, J.R. Batley, A. Battaglia, M. Battistini, F. Bauer, H.S. Bawa, M. Bazarova, B. Beare, T. Beau, P.H. Beauchemin, N. Becerici, A. Bechtle, G.A. Beck, H.P. Beck, M. Beckingham, K.H. Becks, A.J. Beddall, A. Beddall, V.A. Bednyakov, C. Bee, M. Begg, S. Behar Harzpaz, P.K. Behera, M. Beinforde, C. Belanger-Champagne, P.J. Bell, W.H. Bell, G. Bella, L. Bellagamba, F. Bellina, M. Bellomo, A. Belloni, K. Belotskii, O. Beltrame, S. Ben Aim, O. Benary, D. Benckehroun, M. Bendel, B.H. Benedict, N. Benekos, Y. Benhamou, G.P. Benincasa, D.P. Benjamin, M. Benoît, J.R. Bensinger, K. Benslama, S. Bentvelsen, M. Beretta, D. Berge, E. Bergeaas Kuutmann, N. Berger, F. Berghaus, E. Berglund, J. Beringer, P. Bernat, R. Bernhardt, C. Bernius, T. Berry, A. Bertin, M.J. Besana, S. Besson, S. Bethke, R.M. Bianchi, M. Bianco, O. Biebel, J. Biesiada, M. Bighetti, H. Bilokon, M. Bindi, S. Binet, A. Bingui, C. Bini, C. Bisnar, C. Biscarot, U. Bitenc, K.M. Black, R.E. Blair, J.-B. Blanchard, G. Blanchot, C. Blocker, A. Blondel, W. Blum, U. Blumenschein, G.J. Bobbink, A. Bocci, M. Boehrer, J. Boek, N. Boelart, S. Böser, J.A. Bogaerts, A. Bogouch, C. Bohm, J. Bohn, V. Boisvert, T. Bold, V. Boldea, V.G. Bondarenko, M. Bondioli, M. Boonekamp, S. Bordoni, C. Borcher, A. Borisov, G. Borisov, I. Borjanovic, S. Borroni, K. Bos, D. Boscherini, M. Bosman, H. Boterenbrood, J. Bouchami, J. Boudreau, E.V. Boulova-Thacker, C. Boulahouache, C. Bourdarios, A. Boveia, J. Boyd, I.R. Boyko, I. Bozovic-Jelisavcevic, J. Bracini, A. Breem, P. Branchini, G.W. Brandenburg, A. Brandt, G. Brandl, O. Brandt, U. Bratzler, B. brau, J.E. Brau, H.M. Braun, B. Brelier, J. Bremer, J. Brener, S. Bressler, D. Britton, F. Brochu, I. Brock, R. Brock, E. Brodett, C. Crombberg, G. Brooijmans, W.K. Brooks, G. Brown, P.A. Bruckman de Renstrom, D. Bruenck, R. Bruneliere, S. Brunet, A. Brun, G. Brun, M. Brusch, F. Bucco, J. Buchanan, P. Buchholz, A.G. Buckley, I.A. Budagov, B. Budick, V. Bünser, L. Bugge, O. Bulekov, M. Bunse, T. Buran, H. Burckhart, S. Burdita, T. Burgess,
S. Persembe, P. Perus, V.D. Peshekhonov, B.A. Petersen, T.C. Petersen, E. Petit, C. Petridou, E. Petrol, F. Petrucci, D. Petschull, M. Petteni, R. Pezoa, A. Phan, A.W. Phillips, G. Piacquadio, M. Piccinini, R. Piegaia, J.E. Pilcher, A.D. Pilkington, J. Pina, M. Pinamonti, J.L. Pinfold, B. Pinto, C. Pizlo, R. Placakyte, M. Plamondon, M.A. Pleier, A. Plobaguev, S. Poddar, F. Podlisky, P. Poffenberger, L. Poggioli, M. Pohl, F. Polci, G. Polesello, A. Policicchio, A. Polini, J. Poll, V. Polychronakos, D. Pomeroy, K. Pommes, P. Ponsot, L. Pontecorvo, B.G. Pope, G.A. Popenecciu, D.S. Popovic, A. Poppleton, J. Popule, X. Portell Bueso, R. Porter, G.E. Pospelov, S. Pospisil, M. Poteckin, I.N. Potrap, C.J. Potter, C.T. Potter, K.P. Potter, G. Pouard, J. Poveda, R. Prabhu, P. Pralavorio, S. Prasad, R. Pravahan, L. Pribyl, D. Price, L.E. Price, P.M. Prichard, D. Prieur, M. Primavera, K. Prokofiev, F. Prokoshin, S. Protopescu, P. Proudfoot, X. Prudent, H. Przyseziak, S. Psoroulas, E. Ptacek, C. Puigdengoles, J. Purdue, M. Purohit, P. Puzo, Y. Pylypchenko, M. Qi, J. Qian, W. Qin, A. Quadri, D.R. Quarrar, W.B. Quayle, F. Quincez, M. Raas, V. Radeka, V. Radescu, B. Radios, T. Rador, F. Raguza, G. Rahal, J.A.M. Rahimi, S. Rajagopalan, M. Rammensee, M. Rammensee, F. Rauscher, E. Rauter, M. Raymond, A.L. Read, D.M. Rebuizi, A. Redelbach, G. Redlinger, R. Reece, K. Reeves, E. Reinerz-Aronis, A. Reinsch, I. Reisinger, D. Reljic, C. Rembser, Z.L. Ren, P. Renkel, S. Rescia, M. Rescigno, S. Resconi, B. Resende, P. Resznicki, R. Richards, R. Richter, E. Richter-Was, M. Ride, M. Rijpstra, A. Rimoldi, L. Rinaldi, R.R. Rios, I. Riu, F. Rizatdinova, E. Rizvi, D.A. Roa Romero, S.H. Robertson, E. Robichaud-Varneaud, D. Robinson, J.E.M. Robinson, M. Robinson, A. Robson, J.G. Rocha de Lima, C. Roda, R. Roda Dos Santos, D. Rodriguez, Y. Rodriguez, S. Roe, O. Rohne, V. Rojo, S. Rolli, A. Romanionk, V.M. Romanov, G. Romeo, D. Romero Maltrana, L. Roos, E. Ros, S. Rosati, G.A. Rosenbaum, L. Rossette, V. Rossetti, L.P. Rossi, M. Rotaru, J. Rothberg, D. Rousseau, C.I. Royon, A. Rozano, Y. Rozen, X. Ruan, B. Ruckert, N. Ruckstuhl, V.I. Rud, G. Rudolf, F. Rühr, F. Ruggieri, A. Ruiz-Martinez, L. Rumyantsev, Z. Rusirka, N.A. Rusakovich, J.P. Rutherford, C. Ruwiedel, P. Ruzicka, Y.F. Ryabov, P. Rybkin, S. Rzaeva, A.F. Saavedra, H.W. H. Sadrozinski, R. Sadykov, H. Sakamoto, G. Salamanna, A. Salamon, M.S. Saleem, D. Salihagic, A. Salnikov, J. Salt, J.F. Salman, D. Salvaterra, D. Salvatore, E. Salvucci, B. Salzburger, D. Sampsonid, B.H. Samset, H. Sandaker, H.G. Sander, M.P. Sanders, M. Sandhoff, P. Sandhu, R. Sandstrom, S. Sandvos, D.P. Sankey, B. Sanny, A. Sansonetti, C. Santamarina, R. Santonic, J.G. Saraiya, T. Sarang, E. Sarkisyan-Grinbaum, F. Sarra, N. Sasa, I. Satsounkevitch, G. Savauge, P. Savard, A.Y. Savine, V. Savinov, L. Sawyer, D.H. Saxton, L.P. Says, C. Sbarra, A. Sbrizzi, D.A. Scaccinio, J. Schaarschmidt, P. Schacht, U. Schäfer, S. Schaeetzle, A.C. Schafer, D. Schäfe, R.D. Schammer, A.G. Schanom, V.A. Schegelsky, D. Scheirich, M. Schernau, M.I. Scherzer, C. Schiavi, M. Schioppa, S. Schlenker, E. Schmidt, K. Schmied, M. Schmitz, M. Schott, D. Schouten, J. Schovancova, M. Schram, A. Schreiner, C. Schroeder, N. Schroer, M. Schroers, J. Schulze, H.C. Schultz-Coulon, J.W. Schumacher, A. Schumacher, B.A. Schumann, P. Schune, C. Schwanenberger, A. Schwartzman, J.P. Schwemling, R. Schwenhorst, R. Schwierz, J. Schwindling, W.G. Scott, J. Searcy, E. Sedykh, E. Segura, C.C. Seidel, A. Seiden, F. Seifert, J.M. Seixas, G. Sekhniaidze, D.M. Silverov, B. Sellden, N. Semprini-Cesari, L. Serin, R. Seuster, M.E. Sevior, A. Sfyrla, E. Shabalina, M. Shamim, L.Y. Shan, J.T. Shank, Q.T. Shao, P.B. Shatalov, K. Shao, D. Sherman, P. Sherwood, S. Shihata, M. Shimojima, T. Shim, A. Shmeleva, M.J. Shochet, M.A. Shupe, P. Sichov, A. Sidoti, F. Siegert, J. Siegrist, D.J. Sijacki, O. Silberth, J. Silva, Y. Silver, D. Silverstein, S.B. Silverstein, V. Simak, L.J. Simic, S. Simion, B. Simmons, M. Simonyan, P. Sinervo, N.B. Sinev, Y. Sipica, G. Siragusa, A.N. Sisakyan, S.Yu. Sivoklov, J. Sjoelin, T.B. Sjursen, K. Skovpen, T.P. Skubic, M. Slater, T. Slavicek, K. Silva, J. Sloper, T. Sluka, V. Smakhin, S.Yu. Smirnov, Y. Smirnov, L.N. Smirnova, B.C. Smith, D. Smith, K.M. Smith, M. Smizanska, K. Smolek, A.A. Snesarev, S.W. Snow, J. Snow, J. Snuverink, T. Snyder, A. Soares, R. Sobie, J. Sodomka, A. Soffer, C.A. Solans, M. Solar, J. Sol, E. Solfaroli Camillocci, A.A. Solodkov, O.V. Solovyov, R. Soluk, J. Sondericker,
Abstract The simulation software for the ATLAS Experiment at the Large Hadron Collider is being used for large-scale production of events on the LHC Computing Grid. This simulation requires many components, from the generators that simulate particle collisions, through packages simulating the response of the various detectors and triggers. All of these components come together under the ATLAS simulation infrastructure. In this paper, that infrastructure is discussed, including that supporting the detector description, interfacing the event generation, and combining the GEANT4 simulation of the response of the individual detectors. Also described are the tools allowing the software validation, performance testing, and the validation of the simulated output against known physics processes.

1 Introduction

ATLAS [1], one of the general-purpose detectors at the Large Hadron Collider [2], began operation in 2008. The detector will 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. During proton–proton collisions at the design luminosity of $10^{34}$ cm$^{-2}$ s$^{-1}$, beam bunches will cross every 25 ns (40 MHz) and provide on average 23 collisions per bunch crossing. ATLAS has been designed to record up to 200 bunch crossings per second, keeping only the most interesting interactions for physics analyses, including searches for new physics.

In order to study the detector response for a wide range of physics processes and scenarios, a detailed simulation has been implemented that carries events from the event generation through to output in a format which is identical to that of the true detector. The simulation program is integrated into the ATLAS software framework, Athena [3], and uses the GEANT4 simulation toolkit [4, 5]. The core software and large-scale production infrastructures are discussed further in Sect. 2.

The simulation software chain is generally divided into three steps, though they may be combined into a single job: generation of the event and immediate decays (see Sect. 3), simulation of the detector and physics interactions (see Sect. 5), and digitization of the energy deposited in the sensitive regions of the detector into voltages and currents for comparison to the readout of the ATLAS detector (see Sect. 6). The output of the simulation chain can be presented in either an object-based format or in a format identical to the output of the ATLAS data acquisition system (DAQ). Thus, both the simulated and real data from the detector can then be run through the same ATLAS trigger and reconstruction packages.

The ATLAS detector geometry used for simulation, digitization, and reconstruction is built from databases containing the information describing the physical construction and conditions data. The latter contains all the information needed to emulate a single data-taking run of the real detector (e.g. detector misalignments or temperatures). The same geometry and simulation infrastructure is able to reproduce the test stands and installation configurations of the ATLAS detector. The detector description is discussed in Sect. 4.

Large computing resources are required to accurately model the complex detector geometry and physics descriptions in the standard ATLAS detector simulation. This has led to the development of several varieties of fast simulation. Each is best suited to a particular use-case, and they are described in Sect. 7. Validation of the software, testing of the software performance, and validation of the physics performance and output of each piece of the simulation software chain is discussed in Sect. 8.

This paper reviews the status of the software and geometry used for large-scale production in 2008.

2 ATLAS offline software overview

The ATLAS software framework, Athena [3], uses PYTHON as an object-oriented scripting and interpreter language to configure and load C++ algorithms and objects. Rather than develop an entirely new high-energy physics data processing infrastructure, ATLAS adopted the Gaudi framework [6, 7], 

**e-mail: atlas.secretariat@cern.ch**
originally developed for LHCb and written in C++. Gaudi was created as a flexible framework to support a variety of applications through base classes and basic functionality. As much as possible, the infrastructure relies on the CLHEP common libraries [8], which include utility classes particularly designed for use in high-energy physics software (e.g., vectors and rotations).

Athena releases are divided into major projects by functionality [9], and all of the ATLAS simulation software (including event generation and digitization) resides in a single project. The dependencies of the “simulation” project are the “core” project, which includes the Athena framework, the “conditions” and “detector description” projects, which include all code necessary for the description of the ATLAS detector, and the “event” project, which includes descriptions of persistent objects. The number of lines of code by software language for the simulation project are summarized in Table 1, as calculated using cloc [10] in Athena release 14.4. Lines of code in the upstream Athena projects, excluding external dependencies like Gaudi and CLHEP, are summarized in Table 2.

### Table 1 Numbers of files, lines of code, and lines of comments in the ATLAS simulation project, by programming language for major contributors. External dependencies are not included

| Language        | Files | Comment | Code   |
|-----------------|-------|---------|--------|
| C++             | 930   | 24,000  | 120,000|
| FORTRAN         | 270   | 15,000  | 42,000 |
| C/C++ Header    | 1,100 | 13,000  | 34,000 |
| Python          | 430   | 16,000  | 27,000 |
| HTML            | 62    | 130     | 15,000 |
| Bourne Shell    | 390   | 1,000   | 7,300  |
| C Shell         | 380   | 210     | 3,800  |
| XML             | 52    | 1,200   | 3,400  |
| Sum             | 3,600 | 70,000  | 250,000|

### Table 2 Numbers of lines of code in each of the projects upstream of the ATLAS simulation project, versus the programming language. Most projects are dominated by C++ and PYTHON code. The most significant exception is the detector project, which contains 70,000 lines of XML and Java code.

| Project     | C/C++ code | C/C++ headers | PYTHON code | Total code |
|-------------|-------------|----------------|-------------|------------|
| Core        | 390,000     | 43,000         | 240,000     | 860,000    |
| Event       | 200,000     | 110,000        | 16,000      | 350,000    |
| Conditions  | 280,000     | 90,000         | 21,000      | 620,000    |
| Detector    | 38,000      | 6,100          | 8,400       | 140,000    |
| Sum         | 910,000     | 250,000        | 280,000     | 2,000,000  |

All Athena jobs consist of three distinct steps. First, in the initialization step, services and algorithms are loaded on demand using dynamic library loading. Generally, algorithms include methods to be called once per event, whereas services may be accessed many times during a single event. The configuration and initialization is controlled within a common PYTHON infrastructure which allows introspection, particularly useful in debugging and providing help for the user. Also, by using a scripting language for loading and configuring objects, there is no need to recompile C++ code or a script for each job. Small modifications can be made in the scripts (also called “fragments” or “job options”), or even in the midst of the job, without having to stop and recompile the libraries. This scripting method also lightens the load on the user, since there is, under normal circumstances, no need to compile anything prior to running a job.

Each algorithm and service can be configured differently for each step of the simulation software chain, allowing maximal sharing of infrastructure among the distinct steps of the chain. Algorithms can be added to a top list of methods to be run during the event loop.

Second, the event loop begins. All algorithms in the top list are run sequentially on each event. An external generator or algorithm controlling GEANT4 may be added to this list, for example. From these main methods, other services and algorithms can be called. A messaging service, called throughout the jobs, controls log file outputs with different levels of verbosity. The user may configure the total logging verbosity or configure the verbosity individually for a single algorithm, particularly useful for debugging.

During the finalization stage of the job, all algorithms are terminated and all objects are deleted. At this point, algorithms may output any statistics (e.g., memory or CPU usage) they track.

These three steps comprise each Athena job, but the infrastructure allows for the insertion of hooks at various places. Each step of the ATLAS simulation chain takes advantage of this infrastructure to provide maximal flexibility for the user. Only requested modules are loaded as plug-ins, keeping each step as light as possible in memory and as fast as possible during the event loop.

For storing data, ATLAS has adopted a scheme for separating transient from persistent objects. Most general C++ types, immediately prior to storage, are converted to a type that requires less space. Although, for example, energy is accumulated in the calorimeter by summing double-precision floating point numbers, at the end of each event and prior to storage, the total energy is converted into a single-precision floating point number (float). Summing with floats was found to alter the total energy because of truncation. For some types, more complicated storage schemes are implemented that rely on properties of the information to be stored (e.g., where it is possible to sacrifice some accuracy). Metadata, general property information for data collected in a file,
Fig. 1 The flow of the ATLAS simulation software, from event generators (top left) through reconstruction (top right). Algorithms are placed in square-cornered boxes and persistent data objects are placed in rounded boxes. The optional pile-up portion of the chain, used only when events are overlaid, is dashed. Generators are used to produce data in HepMC format. Monte Carlo truth is saved in addition to energy depositions in the detector (hits). This truth is merged into Simulated Data Objects (SDOs) during the digitization. Also, during the digitization stage, Read Out Driver (ROD) electronics are simulated.

2.1 ATLAS simulation overview

An overview of the ATLAS simulation data flow can be seen in Fig. 1. Algorithms and applications to be run are placed in square-cornered boxes, and persistent data objects are placed in round-cornered boxes. The optional steps required for pile-up or event overlay (see Sect. 6.2) are shown with a dashed outline.

A generator produces events in standard HepMC format [14]. These events can be filtered at generation time so that only events with a certain property (e.g. leptonic decay or missing energy above a certain value) are kept. The generator is responsible for any prompt decays (e.g. Z or W bosons) but stores any “stable” particle expected to propagate through a part of the detector (see Sect. 3). Because it only considers immediate decays, there is no need to consider detector geometry during the generation step, except in controlling what particles are considered stable. During this step, the run number for the simulated data set and event numbers for each event are established. Event numbers are generally ordered in a single job, though events may be omitted because of filtering at each step. Run numbers for simulated data sets derive from the job options used to generate the sample and mimic real run numbers used during data taking.

These generated events are then read into the simulation. A record of all particles produced by the generator is retained in the simulation output file (see Sect. 3.6), but cuts can be applied to select only certain particles to process in the simulation. Each particle is propagated through the full ATLAS detector by GEANT4. The configuration of the detector, including misalignments and distortions, can be set at run time by the user. The energies deposited in the sensitive portions of the detector are recorded as “hits,” containing the total energy deposition, position, and time, and are written to a simulation output file, called a hit file.

In both event generation and detector simulation, information called “truth” is recorded for each event. In the generation jobs, the truth is a history of the interactions from the generator, including incoming and outgoing particles. A record is kept for every particle, whether the particle is to be passed through the detector simulation or not. In the simulation jobs, truth tracks and decays for certain particles are stored. This truth contains, for example, the locations of the conversions of photons within the inner detector and the subsequent electron and positron tracks. In the digitization jobs, Simulated Data Objects (SDOs) are created from the truth. These SDOs are maps from the hits in the sensitive regions of the detector to the particles in the simulation truth record that deposited the hits’ energy. The truth information...
is further processed in the reconstruction jobs and can be used during the analysis of simulated data to quantify the success of the reconstruction software.

The digitization takes hit output from simulated events: hard scattering signal, minimum bias, beam halo, beam gas, and cavern background events. Each type of event can be overlaid at a user-specified rate before the detector signal (e.g. voltage or time) is generated. The overlay (called “pile-up”) is done during digitization to save the CPU time required by the simulation. At this stage, detector noise is added to the event. The first level trigger, implemented with hardware on the real detector, is also simulated in a “pass” mode. Here no events are discarded but each trigger hypothesis is evaluated. The digitization first constructs “digits,” inputs to the read out drivers (RODs) in the detector electronics. The ROD functionality is then emulated, and the output is a Raw Data Object (RDO) file. The output from the ATLAS detector itself is in “byte stream” format, which can be fairly easily converted to and from RDO file format. The two are similar, and in some subdetectors they are almost interchangeable. Truth information is the major exception. It is stripped in the conversion to bytestream.

The simulation software chain, divided in this way, uses resources more effectively than a single-step event simulation and simplifies software validation. Event generation jobs, typically quick and with small output files, can be run for several thousands of events at a time. By storing the output rather than regenerating it each time, it becomes possible to run identical events through different versions of the simulation software or with different detector configurations. The simulation step is particularly slow, and can take several minutes per event (see Sect. 8.2). Simulation jobs are therefore divided into groups of 50 or fewer events; only a few events may be completed in a single heavy ion simulation job. Digitization jobs are generally configured to run ~1000 events. This configuration eases file handling by producing a smaller number of RDO files. Each step is partially configured based on the input files. For example, the detector geometry used for a digitization job is selected based on the input hit file.

The ATLAS high level trigger\(^1\) (HLT) [15] and reconstruction [16] run on these RDO files. The reconstruction is identical for the simulation and the data, with the exception that truth information can be treated and is available only in simulated data. During data taking, the HLT is run on bytestream files, however all hypotheses and additional test hypotheses may be evaluated by translating the RDOs into bytestream format.

2.2 Large-scale production system

Because of the significant time consumption of the ATLAS simulation, only minimal jobs can be completed interactively on most computers. It is, therefore, desirable to distribute as much as possible production of the necessary simulated data for ATLAS. Complete software releases are built and distributed to production sites and users every few weeks, providing all Athena software and all external dependencies, including generators and GEANT4. These releases are patched several times with “production caches” before a new clean release is built and distributed. With each release or production cache, a small set of data files are packaged that include database replicas, any necessary external data files, and some sample output files. These sample files can be used to ensure that the locally installed release can be validated by processing events through the entire software chain, from generation through reconstruction.

Large-scale production is then done on the World-wide LHC Computing Grid (“WLCG” or “Grid”) [17]. A single task on the Grid (e.g. simulation of 500,000 \(\text{it} \) events) is separated into many jobs depending on the content and complexity of the task. A job can be completed by a single CPU within the maximum allowed time for a job on the Grid (typically 2–3 days). The output, including log files, of every Grid job is registered with the ATLAS Distributed Data Management system (DDM) [18]. The DDM uses DQ2 [19] for dataset bookkeeping, and allows users to search for datasets on the Grid, analyze them in place, and, if necessary, retrieve them. Separate Grid software controls the distribution of jobs to the various Grid sites. In a typical task, 10 jobs are queued and run as a test sample, and only once they finish successfully are the remainder of the jobs released to the Grid. In the case of a full chain of jobs being run (generation, simulation, and digitization), each subsequent step is automatically held in the queue until the required data are available from the previous step. Frequently, Grid jobs are configured to run two steps (e.g. simulation and digitization together, or digitization and reconstruction together). About one million events per day can be produced using GEANT4 on the Grid.

On the Grid, “job transforms” are run, which may only include well-defined, minor modifications to some standard job configuration after the input events have been specified. A task is given a random number seed, and each job increments the seed in sequence. The modifications to a generation job also may include a configuration file for the selected generator to be run. These configuration files are included with each release and may not be arbitrarily modified by the user during production. The modification to a simulation job may include detector geometry and conditions and specially designed job options fragments that are included with each release. These fragments are typically constructed for

\(^{1}\)The ATLAS high level trigger comprises two stages: level 2, and the event filter. Both are software triggers run with the reconstruction, and may be treated as a single unit for the purposes of this discussion.
a very specific purpose, for example a non-standard vertex smearing, simulation of cavern background, or propagation and late decay of long-lived exotic particles. Many of these modifications can be chained to provide maximal flexibility to the user, but if two fragments are sufficiently complex such chaining becomes impossible. The modifications to a digitization job may include geometry and conditions versions, calorimeter sampling fraction, trigger configuration, and noise control. These modifications are discussed further in the subsequent sections.

3 Event generation overview

Event generation consists of the production of a set of particles which is passed to either full or fast detector simulation. Event generation runs within the Athena framework, but most of the generators themselves are written and maintained by authors external to ATLAS. The ATLAS-specific implementation, therefore, consists mostly of a set of interface packages. These are designed to be as simple as practicable and wherever possible to be factorized from the external packages. This is essential to allow rapid feedback and bug reporting to the authors of the external packages. Most of the well-understood and thoroughly debugged generators are written in FORTRAN. Their interfaces transfer the event information, mostly contained in FORTRAN common blocks, into an object format that can be used by the ATLAS software. This ensures that any downstream algorithms are shielded from details specific to an individual generator. Events can either be stored as POOL files for later use or passed to simulation in the same Athena job.

Details of the framework and comments specific to each generator are listed below. Large-scale production has been run with PYTHIA [20] (including an ATLAS variant, PythiaB [21, 22], used for production of events with B-hadrons), HERWIG [23–25], Sherpa [26], Hijing [27], Alpgen [28], MC@NLO [29], and AcerMC [30]. Tauola [31] and Photos [32] are routinely used to handle tau decays and photon emission. EvtGen [33] is used for B-decays in cases where the physics is sensitive to details of the B hadron decays.2 ISAJET [34] is used for generating supersymmetric particles in conjunction with HERWIG. The newer C++ generators PYTHIA 8 [35] and HERWIG++ [36] are being tested. Both produce events in the HepMC format, so no translation is needed. They can be passed directly to simulation. As these new generators evolve and undergo extensive testing and validation, they are expected to enter the production shortly and eventually supersede their FORTRAN predecessors. Some production was also done with MadGraph [37] (vector boson scattering), CHARYBDIS [38] (black hole event generation), and CompHep [39, 40] (specific exotic physics models). Discussion of the generation of cavern background, beam halo, and beam gas events follow in Sects. 6.2.1 and 6.2.2. Single particle generators are also used to generate cosmic ray events and single particle events for performance studies and calibration of the detector.

Each generated event contains the particles from a single interaction with a vertex located at the geometric origin. Modifications to account for the beam properties are applied to the event before it is passed to GEANT4 (see Sect. 5.1). Particles with a proper lifetime \( c \tau > 10 \) mm are considered stable by the event generator. They can propagate far enough to interact with detector material before decaying. Their decays are handled by the simulation. Any particles with \( c \tau < 10 \) mm are decayed by the event generator, and their interactions with material or curving in the magnetic field of ATLAS are ignored.

3.1 Generator framework

Many external generator packages assume that the parameters for a particular job are set via a main program. This would require recompilation to change parameters. The Athena generator interfaces allow for the passing of all relevant parameters at run time, permitting a fixed software release to be used to produce different physics configurations. During initialization, the relevant parameters are passed via PYTHON fragments. The combination of the fragments, random number seeds, and the software release uniquely identifies the resulting data.3 The Athena event manager is run for each event, and a run number and an event number are created; then the event generator is asked to produce an event. This event is created in memory in the format specific to the generator itself. The event must then be mapped into a common format so that subsequent algorithms are independent of the generator used.

ATLAS uses the HepMC event record [14], initially developed by the ATLAS collaboration but now supported by WLCG [17]. This is a set of C++ classes which holds the full event as produced by the generator. Stable particles are used as input to simulation; unstable ones can be of use in physics studies and diagnostics. Each event generator produces a very large number of stable particles (e.g. muon, kaons, pions, electrons, photons), a much larger number of unstable particles (e.g. gluons, quarks, B mesons, heavy hyperons), and, possibly, other objects (e.g. “strings” or “color singlet clusters”) specific to an individual generator. The

---

2PYTHIA remains the default for current inclusive production, but EvtGen is likely to be used by default for the long-term production.

3Since pseudo-random number generators are chip architecture dependent, jobs are exactly reproducible only when run on the same type of processor (e.g. Intel or AMD).
HepMC record consists of a connected tree, navigation inside of which retains information on the event history including the parents of unstable particles. There is an important caveat here: the event generators are modeling quantum processes, and the event record has the structure of a classical decay chain. It is inevitable that compromises must be made and difficulties can arise from an over-literal interpretation of the tree structure. A very simple example is provided by events containing an $e^+e^−$ pair. The parent of the $e^+e^−$ pair cannot be uniquely specified, as the pair may arise from an intermediate Z boson, photon, or quantum interference. The HepMC event record is also used to contain the particle information from secondaries produced by interactions in the detector. This is discussed below in the section on Monte Carlo truth (see Sect. 3.6). Information about all interacting partons (e.g. momentum fractions $x_1$ and $x_2$) is saved, so that parton distribution function reweighting can be done without rerunning the event generation.

The FORTRAN generators usually use the HEPEVT common block [41] to store the information. Unfortunately, the different generators use slightly different structures. A separate translation into HepMC is needed for each one. The C++ generators such as HERWIG++ produce output in the HepMC format. No translation is required and the integrity of the HepMC event record is the responsibility of the generator authors.

3.2 General purpose generators

General purpose generators produce complete events starting from a proton–proton, proton–nucleus or nucleus–nucleus initial state. They are used standalone or with specialized generators that improve the description of certain final states. They have many parameters, some of which are related to fundamental parameters such as the QCD coupling constant and electroweak parameters, and some of which describe the models used to parametrize long distance QCD, soft QCD, and electroweak processes.

3.2.1 PYTHIA and PythiaB

PYTHIA [20] and HERWIG (see below) in their FORTRAN versions have been tested, used, and validated over many years in $e^+e^−$ and hadron colliders. They start with a hard scattering process calculated to lowest order in QCD. They then add additional QCD and QED radiation in a shower approximation which is most accurate when the radiation is emitted at small angle. The approximation is poorest in those cases with a large number of widely separated emissions of comparable energy. In addition, PYTHIA use a model for hard and soft scattering processes in a single event in order to simulate underlying activity. This model is used in the simulation of minimum bias events. While other generators may be used for specific final states, PYTHIA and HERWIG are the benchmarks.

ATLAS uses PYTHIA 6.4. There are two models of QCD radiation in PYTHIA. By default, ATLAS uses the showering model introduced in PYTHIA 6.3. This showering model is believed to better match the theoretical description of QCD showers. It produces somewhat more jet activity [42, 43], resulting in “busier” events than the older model which was used, for example, for detailed simulations at the Tevatron (see, for example, [44, 45]). In this model, the multiple scatterings which make up the underlying event are interleaved with the parton shower according to the hard scale of the scatter or the emission. At the end of the shower, a phenomenological model is used to combine the quarks and gluons into hadrons. This hadronization model, which has many parameters, has been tuned by comparison with data in $e^+e^−$, ep, and hadron colliders [46, 47]. The underlying event model was retuned within ATLAS [48] to recover an acceptable description of the Tevatron data [49, 50]. PYTHIA contains a very large number of built-in processes, and new ones can be added by modifying the code. Hard scattering events can also be generated in a separate program in a standard format and fed into PYTHIA for the addition of a parton shower and hadronization. PYTHIA is the default generator in ATLAS: many hundreds of millions of events have been generated using it. Its ease of use, speed, and robustness make it an ideal choice for the default. It is supplemented by other generators, either to obtain some estimate of the uncertainties, or when specialized generators are expected to give a better physical description in certain final states.

PythiaB [21, 22] is an ATLAS-specific modification of PYTHIA aimed at the efficient generation of events related to B-physics. In PYTHIA, most high $p_T$ bottom quarks are produced in the QCD shower of a high $p_T$ light quark or gluon from a hard scattering process. Most showers do not produce such a $b\bar{b}$ pair, so using PYTHIA to generate B-physics events is inefficient. PythiaB reuses those QCD showers that contain a $b$- or $c$-quark, hadronizing them several times to increase the probability of producing a $b$-hadron. Since the probability producing a $b$- or $c$-quark in a parton shower is low, this procedure results in more efficient procedure of making $b$-hadron events without introducing any bias in the distribution of $b$-hadrons within the event. If a specific decay mode is then required the $b$-hadron decay can be forced using a modified $b$-hadron decay table, either in PYTHIA itself or via EvtGen.

3.2.2 HERWIG

ATLAS uses HERWIG 6.5 [23–25], the last release of the FORTRAN HERWIG package which is now superseded by HERWIG++ (see below). It is a flexible generator with a
large number of built in processes and has been tuned to agree with the Tevatron data [49, 50]. In particular, most of the generation of supersymmetric processes is done with HERWIG using the ISAWIG package [23–25] with the particle spectra and decay modes generated by ISAJET. ATLAS uses HERWIG with the Jimmy [51] implementation of the underlying event.

3.2.3 Sherpa

Sherpa [26] is a generator written in C++ which implements the CKKW duplicate removal prescription [52] to match fixed-order QCD matrix elements to QCD showers. It uses an interface to PYTHIA’s hadronization model and produces complete events. It is expected to give better approximations for final states with large numbers of isolated jets than generators such as PYTHIA and HERWIG based on pure QCD showering. Sherpa generates underlying events using a simple multi-parton interaction model based on that of PYTHIA. For each new process to be generated, Sherpa must be recompiled to incorporate the specific libraries for the process of interest. On the Grid, this implies either re-compiling Sherpa at the production site or deploying updated libraries for new production jobs. Instead, Sherpa is run locally to produce event files in Sherpa’s native format. These files are then translated into the HepMC format with an additional Athena Grid job. It is also possible to run Sherpa entirely within an Athena job.

3.2.4 Hijing

Hijing [27] is a dedicated generator for the production of heavy ion events at all impact parameters. In a dense nuclear environment, such as appears in central collisions, a particle produced in a primary collision can re-interact several times as it propagates. Hijing models the propagation. It is also the only generator that can be used for proton–nucleus collisions occurring in beam–gas interactions. Hijing uses the PYTHIA hadronization model.

3.2.5 Single particle generators

A single particle event generator is frequently used for calibrating the detector, testing, and evaluating the reconstruction efficiencies. Although unphysical, these generators produce events with a single primary particle, for example a muon, electron, or charged pion, at a specified energy, position, and momentum direction. A range may also be specified for either the energy or direction. No underlying event, proton remnants, or other primary interactions are included when these events are generated.

A specialized single particle generator is used to produce cosmic ray events. Single muons are generated at the earth’s surface in a square region (typically 600 m by 600 m) above the ATLAS detector and with the standard cosmic ray $p_T$ spectrum [53, 54]. The upper and lower energy cutoffs of the spectrum are configurable. Those muons pointing to a sphere of configurable size (typically 20 m) centered at the geometric origin are propagated through the bedrock and the ATLAS cavern during simulation.

3.3 Specialized generators

Specialized generators do not produce complete events which can be passed directly to simulation. Rather, they are run in conjunction with one of the general purpose generators to improve the accuracy for specific decays or specific final states. Several of these specialized generators are “Les Houches” type generators. That is, they are run stand-alone using unmodified code from the generator author and produce an ASCII file containing partonic four-vectors in the “Les Houches” format [55, 56]. Athena uses a common interface that reads in these files and prepares them for processing in PYTHIA or HERWIG [55].

3.3.1 ISAJET

The FORTRAN generator ISAJET [34] is not used in large-scale production. However, it is used in conjunction with HERWIG for generation of supersymmetric events. Here, the ISASUGRA component of ISAJET is used to generate consistent sets of masses and decay modes for supersymmetric models. These are then loaded into HERWIG using the ISAWIG translation package, and HERWIG then generates complete final states.

3.3.2 Photos and Tauola

ATLAS uses the dedicated tau decay package Tauola to handle tau decays [31]. General purpose generators are set to treat tau leptons as stable. The events are passed to Tauola for decay. Because Tauola is a FORTRAN package, the events are extracted from the HEPEVT record. The Tauola interface is dependent on the generator that produced the tau, because helicities and helicity correlations are passed in generator-dependent formats. The original generator’s results must be replaced, so both the input and output formats of Tauola are in fact generator-dependent. Special attention is paid to the polarization of the tau. In certain cases, for example the decay $W^\pm \rightarrow \tau^\pm \nu_\tau$, the polarization is known for the tau. In others, such as $Z \rightarrow \tau^+ \tau^-$, there is a correlation between the polarization of the taus.

Photos handles electromagnetic radiation [32]. It is used by Tauola, and, therefore, Tauola cannot be used without Photos. Photos is also used to improve the description of electromagnetic radiation in, for example, the decay $W^\pm \rightarrow$
$e^\pm \nu_e$, where radiation distorts the electron energy distribution. In these cases the final state electromagnetic radiation is switched off in the general purpose generator, usually HERWIG or PYTHIA, to avoid double counting.

### 3.3.3 EvtGen

EvtGen [33], originally developed by the CLEO collaboration, provides a more complete description of B meson and hadron decays than that provided in PYTHIA or HERWIG. Recent modifications have been made to handle $B_S$ and $b$-baryon decays, incorporating measurements from the Tevatron, BaBar, and Belle. In particular, EvtGen incorporates the best measurements of branching ratios and has theoretical models for unmeasured decay modes. It includes angular correlations, which impact the acceptance for certain decay modes of $B$ mesons and baryons. It has been used for ATLAS studies involving the prospects for measurements of exclusive B decays.

### 3.3.4 Alpgen

Alpgen [28] is a “Les Houches” type generator enabling more sophisticated generation of certain final states. HERWIG or PYTHIA is then used to perform the hadronization and produce final (and initial) state QCD radiation. Alpgen is targeted at final states with several well-separated hadronic jets where the fixed order QCD matrix element is expected to give a better approximation than the shower approximation of PYTHIA or HERWIG. Alpgen is used, for example, to generate final states containing a $W$ or $Z$ and many jets. Alpgen also provides an algorithm to prevent double counting by event rejection. The Athena interface package includes the methods needed to pass events through HERWIG and veto those events that would contribute to double counting. This process can be very inefficient for final states with large numbers of jets, and generation time can be significant.

### 3.3.5 MC@NLO

MC@NLO [29], which is also a “Les Houches” type generator, runs standalone to produce ASCII files which are then processed by HERWIG running inside of Athena. MC@NLO uses fundamental (hard scattering) processes evaluated at next to leading order in QCD perturbation theory. It is used, for example, to generate top events as it gives a better representation of the transverse momentum ($p_T$) distribution of top quarks than PYTHIA or HERWIG. MC@NLO includes one loop corrections, with the consequence that events appear with negative and positive weight which must be taken into account when they are used. Any resulting distribution will contain entries from both types of event, and, given sufficient statistics, the result will by physical (i.e. positive).

MC@NLO has been used for large-scale production of top, $W$ and $Z$ events. Only the parts of MC@NLO needed to read these events and process them via HERWIG are included in Athena releases.

### 3.3.6 AcerMC

AcerMC [30] is a “Les Houches” type generator aimed primarily at the production of $W$ or $Z$ bosons with several jets, including jets originating from $b$-quarks. A partonic final state is obtained by running it standalone and making an external ASCII file. Only the parts needed to read these events and process them via PYTHIA are included in Athena releases.

### 3.4 New C++ generators

#### 3.4.1 PYTHIA 8

PYTHIA 8 [35] is a rewrite of the FORTRAN PYTHIA in C++ with new and expanded physics models. It provides a new user interface, transverse-momentum-ordered showers, and interleaving with multiple interactions. The program is under intensive tests and it will require some further tunings before it can replace the Pythia6 code as a leading generator. It is, however, interfaced to Athena and used for generator studies in ATLAS. It includes support for both “Les Houches” and HepMC event formats.

#### 3.4.2 HERWIG++

HERWIG++ [36] is the C++ based replacement for HERWIG. It contains only important processes from the Standard Model, the universal extra dimensions model, and supersymmetric models (whose details are specified via Supersymmetric Les Houches Accord model files [58, 59]). Additional hard scattering processes can be used via “Les Houches” input from specialized generators, and additional decay models can be added by users.

HERWIG++ will soon be used for generation of some Standard Model processes, notably $W$ and $Z$ production. It will also be used for supersymmetric processes, because it includes full spin correlations and QCD radiation in the supersymmetric decay chains. The current version of HERWIG++ also incorporates an underlying event model based on the extension of Jimmy [51] to include soft scatters [60] and can thus potentially generate minimum bias physics.

---

4 An alternative tool, POWHEG [57], implements essentially the same physics and produces events with only positive weight. Once it includes all the processes that MC@NLO does and has been validated, it is expected to take the place of MC@NLO.
3.5 Parton distribution functions

Parton distribution functions (PDFs) are used to describe the substructure of the proton and are used by all the event generators as external inputs. ATLAS uses the Les Houches Accord PDF Interface (LHAPDF [61]) library which is a replacement for PDFLIB [62] which provides a large repository of PDFs. CTEQ [63] PDFs are used by default (MC@NLO uses NLO PDFs, and all other generators use LO PDFs). There is a correlation between the PDFs and the tuning of parameters connected to initial state radiation [64, 65]; inconsistent results can be obtained by varying the PDFs in isolation. Therefore, when a new set of PDFs is used, the parameters of the event generator are retuned to produce consistent results [42].

3.6 Monte Carlo truth

The entire connected tree of the HepMC event record is stored as the Monte Carlo truth. Only the stable particles are propagated by the simulation. The various status codes and event history provided by the individual generators are retained within the HepMC event record. Unfortunately, much of this information is specific to a particular generator. Only status codes 1 (stable) and 2 (unstable) have a general meaning: the remaining values are used differently by the individual generators. As remarked in Sect. 3.1, there can be ambiguities resulting from the attempt to represent a quantum process by a classical tree. Some filters have been provided to select HepMC particles that, for example, are stable at the generator level or are non-interacting (e.g. neutrons).

When the simulation is run, the HepMC tree from the event generator is copied, and some particles resulting from decays within, or interactions simulated by, GEANT4 are added to the copy (see Sect. 5.3). In this way, a complete event including both the generator and simulation information is provided. In order to ensure consistency, a particle decayed by GEANT4 but considered stable by the generator (such as a $K_S$) has its status code changed when the copy is made. A particle that has status code 2 after simulation will be identified as stable at the generator level, if the decay took place in GEANT4. GEANT4 secondaries are distinguished from those from the generators by an offset applied to their numerical identifier. The resulting Monte Carlo truth record can be large and account for a significant fraction ($\sim 30\%$) of the disk space used by a simulated event after reconstruction.

3.7 Default parameters, tuning and bug fixing

The generator authors define default parameters. In some cases, however, these parameters are not tuned for use at the Large Hadron Collider and are superseded by parameters obtained by comparisons to data. The criteria for a particle to be considered stable are modified for use in ATLAS, for example. Once high-energy data appear, it is expected that retuning of the parameters will occur. These tunings can be made by varying parameters at run time. Once a new tuning is available, it can be loaded as a PYTHON fragment at run time or hard coding the values into the generator interfaces. In either case, the tuning becomes available as part of the next Athena software release and will be enabled by default. The settings can be overridden if needed or the previous defaults re-established. It is important to note that the parameters are often not independent and a complete set must be used. Arbitrary adjustments of a few parameters may result in inconsistent results. One of the most important sets of tunings is concerned with structure of minimum bias events and spectator processes in a hard scattering event: the underlying event. At present, these tunings are obtained for both PYTHIA and HERWIG [42] by first tuning to the Tevatron data and then extrapolating. The extrapolation from the Tevatron relies on the models used by PYTHIA and HERWIG. This extrapolation has had testing from comparisons of the Tevatron data at 630 and 1800 GeV [46, 66]. A high priority task for the ATLAS simulation as data accumulates is the testing of these tunings and changing of the parameters as needed.

4 ATLAS detector description

The ATLAS detector is described in detail in Ref. [1], but its main features will be summarized here. We discuss the geometry used in the simulation, which as much as possible matches the as-built detector. A cut-away view of the entire detector is shown in Fig. 2. ATLAS comprises several concentric components. The subdetectors are:

- A Beam Conditions Monitor (BCM) and Beam Loss Monitor used for detecting dangerous conditions and triggering an abort in the detector system. The BCM is located 1.84 m from the interaction point (IP) at $|\eta| \sim 4.2$.
- A tracking detector composed of a fine granularity pixel detector with three layers covering $|\eta| < 2.5$, a silicon strip tracker (SCT) with eight layers determining four space points covering $|\eta| < 2.5$, and a transition radiation tracker (TRT) which has 32 space points on a typical track, covering $|\eta| < 2.0$.\footnote{Pseudorapidity, $\eta \equiv -\ln\tan(\theta/2)$, where $\theta$ is the polar angle measured from the beam pipe. The other coordinate variables used are typically $r$, $z$ and $\phi$, where the $x$-axis points towards the center of the LHC ring, the $y$-axis points up, the $z$-axis defines a right-handed coordinate system, $r \equiv \sqrt{x^2 + y^2}$, and $\phi$ is the azimuthal angle defined such that $\phi = 0$ along the $x$-axis.}

 Springer
• Hermetic calorimetry composed of liquid argon (LAr) electromagnetic calorimetry covering $|\eta| < 3.2$, scintillating tile hadronic calorimetry in the barrel ($|\eta| < 1.7$), sampling LAr hadronic calorimetry in the endcap ($1.5 < |\eta| < 3.2$), and LAr electromagnetic and hadronic forward calorimetry covering $3.2 < |\eta| < 4.9$.

• Four different types of muon chambers, two of which are high precision (monitored drift tubes, and cathode strip chambers) and two of which have a rapid response for the muon trigger (thin gap chambers, and resistive plate chambers), covering $|\eta| < 2.7$.

• Luminosity detectors, including a zero-degree calorimeter that sits 140 m from the interaction point, a detector that performs a luminosity measurement using Cherenkov integration (LUCID), and an absolute luminosity detector for ATLAS.

The ATLAS magnetic field is formed by a solenoid, providing a 2.0 T uniform magnetic field in the tracking subdetectors, and a toroidal magnet system, composed of a barrel and two endcap toroid magnets. In the inner detector, the field has small $\phi$- and $z$-asymmetries due to the toroid field and perturbations from the iron nearby. The field in the toroidal system has approximate $z$- and eight-fold $\phi$-symmetry and provides on average 2.5 Tm of bending power in the barrel and 5 Tm in the endcap. During a simulation run, the field map required about 30 MB of memory.

In the standard production simulation (see Sect. 2.2) the luminosity detectors are not included. They can be simulated in dedicated jobs, but keeping particles in such a high pseudorapidity region increases simulation time by approximately 50% per unit of pseudorapidity ($|\eta|$) per event (see Sect. 5.1).

Several layouts of the complete detector are available, including those that were used for recording cosmic ray events while the detector was being completed. Test stands are also supported with the same infrastructure. All these layouts are described in Sect. 4.4. As much as possible, the details of the detector geometry are preserved in the simulation layout. Some approximations are necessary for describing dead materials, for example bundles of cables and cooling pipes in the service areas of the detector. In these cases, the description only aims to match the general distribution of the material, including inhomogeneities in $\phi$.

4.1 Simulated detector geometry

The geometry structure can be viewed in terms of solids, basic shapes without a position in the detector; logical volumes, solids with additional properties (e.g. name or material); and physical volumes, individual placements of logical volumes. Table 3 shows the number of materials, solids, logical volumes, physical volumes, and total volumes created when constructing various pieces of the ATLAS detector. Not all volumes are equivalent, however: in the case of repeating structures, as in the sampling portion of the LAr calorimetry in particular, it is possible to define a single logical volume that is repeated in hundreds of physical volumes (known as volume parameterization). Because of nesting, one can also define dependencies that create many total vol-
in each detector subsystem and the memory required to build efficiencies, and calorimeter response.

example, missing transverse energy, track reconstruction efficiency of the ATLAS detector, with hundreds of materials and a complex shape. One can see in this table the complexity of the ATLAS detector in GeoModel. Here “calorimeter” is simply the sum of the liquid argon and tile calorimetry volumes from the physical volumes used. In other cases, a single volume can correspond to a piece of shielding or support with a complex shape. One can see in this table the complexity of the ATLAS detector, with hundreds of materials and hundreds of thousands of physical volumes. Such a detailed detector description is crucial for accurately modeling, for example, missing transverse energy, track reconstruction efficiencies, and calorimeter response.

Table 4 shows the number of physical volumes contained in each detector subsystem and the memory required to build each using the GeoModel library (see Sect. 4.3) [67]. As expected, the two are correlated, although differences in volume complexity invalidate a direct correspondence. The entire geometry must be translated into a GEANT4 equivalent, so the total memory required for the geometry of the entire ATLAS detector is almost 300 MB (see Sect. 8).

In creating such a complex, dense geometry, removing volume overlaps and touching surfaces provides a particular challenge. Any overlap of more than 1 picometer and any place in which two volume faces touch can lead to stuck tracks during the simulation, a situation in which a track in GEANT4 may not know in which volume it belongs. These stuck tracks result in a loss of the event, but they can be overcome by introducing small gaps between volumes, at the cost of an extra step for each particle moving through the transition region.

Many layouts are available corresponding to the various revisions of material. The material budget is constantly updated, so that the geometry description is as realistic as possible. During any major updates of detector geometry, the subdetectors are generally required to make all changes backwards-compatible so that all older geometries can be configured and run as normal. This requirement allows for a fair comparisons between software releases with consistent geometries. During any job, the user may choose to enable or disable portions of the detector. Each subdetector is responsible for including any necessary materials and elements for its own construction. In this way, only required elements and materials are used during simulation, and memory loads are reduced when not using the entire ATLAS detector. The switches for disabling portions of the detector generally correspond to the highest level of the tree-structure in the detector geometry (i.e. entire subdetectors, not pieces).

It is possible to apply detector “conditions” modifications to each chosen geometry layout. The detector misalignment can be configured by selecting misaligned layouts either for each subdetector or for the full detector at once. Each ATLAS subdetector sits within a well defined envelope, allowing each to shift and distort without colliding with any other. One set of misalignments is provided specifically for the detector simulation, which includes appropriate levels of misalignment without any overlapping volumes. One may alternately use a misalignment directly derived from the data, but these misalignments may include overlapping volumes, which can cause missed hits. The misalignments are generally applied to more than simply the active volume. That is, if a silicon module in the inner detector is moved, any volumes associated with it (e.g. physically attached electronics) are moved as well.

In digitization and reconstruction jobs, conditions may include detector information beyond misalignments (e.g. dead channels). The infrastructure is in place to record detector conditions in a database and, at run time, allow the

| Subsystem       | Materials | Solids | Logical vol. | Physical vol. | Total vol. |
|-----------------|-----------|--------|--------------|---------------|------------|
| Beampipe        | 43        | 195    | 152          | 514           | 514        |
| BCM             | 40        | 131    | 91           | 453           | 453        |
| Pixel           | 121       | 7,290  | 8,133        | 8,825         | 16,158     |
| SCT             | 130       | 1,297  | 9,403        | 44,156        | 52,414     |
| TRT             | 68        | 300    | 357          | 4,034         | 1,756,219  |
| LAr calorimetry | 68        | 674    | 639          | 106,519       | 506,484    |
| Tile calorimetry| 8         | 51,694 | 35,227       | 75,745        | 1,050,977  |
| Inner detector  | 243       | 12,501 | 18,440       | 56,838        | 1,824,614  |
| Calorimetry     | 73        | 52,366 | 35,864       | 182,262       | 1,557,459  |
| Muon system     | 22        | 33,594 | 9,467        | 76,945        | 1,424,768  |
| ATLAS total     | 327       | 98,459 | 63,769       | 316,043       | 4,806,839  |

To overcome by introducing small gaps between volumes, at the cost of an extra step for each particle moving through the transition region.
user to select conditions from a specific data taking run. Conditions and geometry versions selected by the user can be transferred from the simulation jobs to the digitization and reconstruction jobs so that no additional user interaction is required. These default versions may at any time be overridden by job options.

In order to study the penalties of a poor material description on jet resolution and missing transverse energy bias, a special geometry layout with material distortions was created [68]. Material distortions correspond to additional material added to half of the detector ($y > 0$) to approximate a poor material description.

4.2 Databases and configuration

Two databases are used to construct the detector geometry chosen by the user: one to store basic constants (the ATLAS Geometry database), and one to store various conditions data (e.g. calibrations, dead channel, misalignments) for the specific run chosen (ATLAS Conditions database) [69]. At CERN, large (terabytes) Oracle data bases are used, primarily because they are well supported and straightforward to update. With any stable software release, a small subset of data needed for Athena jobs is replicated from Oracle into SQLite [70–72]—file based databases—and is distributed to the production centers. The large I/O requirements of production jobs can overwhelm a central Oracle server and are better handled by relatively small SQLite files. These files can also be replicated to individual production nodes for local and rapid access. The database replica version to be used can be chosen at run time for each Grid job.

Both the geometry and conditions databases support versioning of the data. The data are organized in a tree consisting of branch and leaf nodes. The nodes in this tree can be “tagged,” and one can create a hierarchy of the tags. Such tag hierarchies are uniquely identified by the tag of the root node, which is usually referred to as top level geometry or conditions tag.

A geometry database stores all fundamental constants for detector construction. Volume dimensions, rotations, and positions, as well as element and material properties including density, are all stored as database entries. New detector-specific tags may be created for inclusion in a global ATLAS geometry tag, where different tags generally correspond to different detector geometry revisions. At run time, the user can select a global geometry tag as well as detector-specific geometry tags to create the desired geometry. In addition to constants for detector construction, the geometry database contains links to external data files that may store, for example, magnetic field maps. These files are shipped with software distribution kits to production sites. By using links through the database, it is possible to select a magnetic field map based on the chosen geometry layout. The selection of field map based upon the name provided in the database, for example, can be overridden with job options.

A separate conditions database stores detector conditions data which are indexed by intervals of validity and tags. The entire detector may be optionally misaligned with a global misalignment tag, and the user may configure the job to use specific misalignment versions for each subdetector. The global misalignment is used frequently to study the performance of the entire ATLAS detector with misalignments of the expected as-built magnitude. The detector-specific misalignments allow studies of the effects of misalignment of a single subdetector assuming ideal alignment of the remainder of ATLAS. The inner detector, for example, has completed an alignment challenge, wherein simulated data was produced with misalignments, and the analysis group was challenged to align the detector as with data. The tile calorimeter, on the other hand, does not use any misalignments in its geometry. A variety of misalignments have been used in the lead-up to data taking in order to speed the process of global detector alignment and improve early physics studies.

During data collection, the alignment constants of the detector are recorded periodically in the central conditions database. The user is able to recreate the misalignment conditions for a specific run by selecting an alignment version, again by subdetector if desired, at run time.

4.3 GeoModel and translations

The ATLAS simulation, digitization, and reconstruction each run in distinct jobs, but they must be able to use the same detector geometry. Therefore, a complete geometry description is maintained that can be used by each step and is not specific to any. By using the geometry databases, it is already possible to read identical detector constants and run conditions.

For these reasons, ATLAS uses GeoModel [67], a library of basic geometrical shapes, to describe and construct the detector. GeoModel contains geometry features similar to those of GEANT4: basic volumes can be constructed, rotated, and shifted in space; subvolumes can be placed inside a volume; boolean volumes can be made by adding or subtracting primitives; volumes can be parameterized and repeated. For the digitization and reconstruction, this detector description is entirely sufficient to place hits, reconstruct tracks and objects, and complete all necessary calculations.

The GeoModel descriptions of most ATLAS subsystems are built using constants in the geometry database. However, a translator has been constructed that parses an XML description of a detector’s geometry and builds a transient representation from GeoModel primitives at run time. This generic package can translate any valid XML description of
detector geometry into GeoModel format. It has been used for describing the geometry of the muon system’s rather complicated dead material.

For the simulation, the geometry is translated entirely from the GeoModel to the GEANT4 format. All volumes and subvolumes are translated, constructed, and properly placed within the “world volume” (the volume allocated for the detector, at the edge of which particles cease to be simulated). All information tied to GeoModel, including position, rotation, and dimensions, are also translated into a GEANT4 equivalent. Once the geometry has been translated, all subsystems rely solely on their GEANT4 description. The GeoModel geometry is currently maintained in memory for the entirety of the job, though it may be released to ease memory pressure in the future. As shown in Table 4, this release can be expected to save 100 MB of memory. Sensitive detectors and particle range cuts (see Sect. 5.5), for example, are tied to the GEANT4 geometry by volume name and can be added at any time after the geometry has been constructed. Each change in detector description is particularly weighty in simulation, because any additional volumes must be built both in GeoModel and in the GEANT4 geometry.

4.4 Alternate layouts

In addition to the standard detector layouts, several commissioning layouts are available to the user for simulation of cosmic ray data taking. During detector assembly, cosmic ray data were taken for several runs using as many subdetectors as were available. Some of these subdetector configurations included calorimeter endcaps shifted out of position while the inner detector was being accessed and were missing large portions of the beam pipe that had not yet been installed. One such commissioning layout is shown in Fig. 3. For studies of cosmic rays and cavern background, it is possible to simulate the ATLAS cavern surrounding the detector as well as the bedrock surrounding the two shafts leading down from the surface.

Several different magnetic field configurations are also available for some of the full detector layouts. Fields with the toroidal magnets on and solenoid off or solenoid on and toroidal magnets off are provided. These magnet configurations have already been used for some cosmic ray data taking runs and may be used for brief periods during high-energy collisions as well. Field maps have also been constructed that reflect the as-built misalignments of the mag-

Fig. 3  Commissioning layout of the detector used for cosmic ray data taking during 2008. The endcap toroidal magnets and beampipe are not yet installed. The calorimeter endcaps (purple) are shifted by 3.1 m and the muon endcaps (green) are shifted to provide access to the inner detector during installation. The barrel toroid magnets are shown in yellow, and the inner detector is shown in blue
net system, for example a vertical shift of a 1.6 mm in the solenoid.

There are also several test stand layouts that were constructed to model test beam and standalone cosmic ray runs. A sample of these various test stands, including subsystems, incident particles, and energies, are listed in Table 5. A combined test beam run was taken with a wedge of the full detector during 2004 [73], and standalone test beams were constructed for the muon detectors [74, 75], tile calorimeter [76], and liquid argon calorimeter subsystems [77, 78]. The combined test beam setup is shown in Fig. 4. Cosmic ray data were also collected with various pieces of the inner detector [79, 80] and with the muon chambers both prior to and after installation.

All test stand and commissioning layouts are available as a part of the same geometry infrastructure and can be selected at run time for simulation. By maintaining all detector configurations as a part of a common infrastructure, it is possible to ensure consistency between, for example, the test beam and full detector simulation. Conclusions drawn from analysis of the test beam simulated data are generally still valid for the full detector simulation. The extensive tuning of the detector simulation and digitization on test beam data can be applied directly to the full detector. As many common elements as possible are kept between the two, including GEANT4 version and physics list (see Sect. 5).

### Table 5 Examples of test stands for ATLAS simulated using GEANT4

| Subsystem                      | Incident particle | Energy       |
|--------------------------------|-------------------|--------------|
| Hadronic Endcap calorimeter   | $e^+/-$, $\pi^+/-$, $\mu^+/-$ | 6–245 GeV    |
| Electromagnetic Barrel calorimeter | $e^+$          | 10–245 GeV   |
| Electromagnetic Endcap calorimeter | $e^+$          | 10–200 GeV   |
| Combined Endcap calorimeter   | $e^+/-$, $\pi^+/-$, $\mu^+/-$ | 6–200 GeV    |
| Hadronic Barrel calorimeter   | $\pi^+/-$, $p$   | 5–350 GeV    |
| Entire detector endcap wedge  | $e^+/-$, $\pi^+/-$, $\mu^+/-$ | 1–350 GeV    |
| Muon Detectors                | $\mu^+/-$        | 20–350 GeV   |
| Silicon Pixel Tracker Endcap  | Cosmic Rays      | 0.5–200 GeV  |
| Silicon Strip Tracker Barrel  | Cosmic Rays      | 0.5–200 GeV  |

Each subsystem’s scoring is optimized and tailored to store only what is necessary for accurately reproducing the performance of that particular subsystem [81–88]. Athena code is necessary to add to the Monte Carlo truth record. Physics models are chosen and parameters optimized for the ATLAS detector. The results shown in this paper used GEANT4 version 8.3 with official patch #2 and two modifications: updates for boundary represented volumes and a patch to the G4Tubs code. The software is continuously evolving, and ATLAS has moved to newer Geant4 versions since the writing of this paper.

The GEANT4 Collaboration and ATLAS Simulation Group have benefitted from 15 years of close collaboration. Frequently, new GEANT4 features have allowed faster or more realistic simulation of the ATLAS detector. Feature requests from the ATLAS collaboration have helped drive the development of GEANT4. The ATLAS simulation has also provided one of the more complicated test-beds for the GEANT4 toolkit, and GEANT4 has been extensively evaluated and validated during large-scale simulation production.

In order to provide PYTHON flexibility to the GEANT4 simulation, an additional layer of infrastructure is necessary. “Standard” GEANT4 simulation typically runs from compiled C++, and in order to modify any of the parameters or the geometry used in the simulation it is necessary to recompile. The Framework for ATLAS Detector Simulation (FADS) [89] wraps several GEANT4 classes in order to allow selection and configuration without recompilation of any libraries. Since a PYTHON interface is used for configuration, all the usual introspection capabilities of PYTHON may be employed. FADS wraps GEANT4 base-classes for volumes, materials, and sensitive detectors for hit processing as well as GEANT4 physics process definitions. These wrappers serve a dual purpose: first, they ease translations between the GEANT4 and Athena standards of geometry, hits, and particle storage. Second, FADS can catalogue the options available to the user, loading only those that will be needed for the desired simulation configuration while still providing all possibilities without any recompilation. Through FADS, a user is free to select a physics list (see Sect. 5.4) for use during the simulation. The user may also modify the physics list by adding particles or processes not included in the GEANT4 toolkit but included in the FADS catalogues, for instance in the simulation of long-lived exotic particles. Similarly, the detector description is configured with PYTHON dictionaries and FADS catalogues before it is built in GEANT4 and may be modified by the user. For example, sensitive detectors may be assigned to any volume in the detector. Range cuts (see Sect. 5.5) may also be added in the PYTHON and FADS layer prior to their being applied to any constructed GEANT4 geometry. Once the PYTHON configuration is complete, FADS objects are translated into their GEANT4 equivalents and loaded. Even after
In order to fit into the Athena framework, a service for GEANT 4 and an algorithm that calls the service during the event loop have been implemented [90]. The service wraps the event loop of GEANT 4 and provides a few additional handles for user-configuration in the PYTHON layer. The service also takes care of initialization and finalization of each GEANT 4 event. The generated events are translated from HepMC format into the standard GEANT 4 event format prior to each event, and at the end of each event an analysis is done to ensure that the simulation finished without errors. Most of the functionality of the standard GEANT 4 run manager is included in this service, so that any Athena-specific modifications (e.g. event translation from HepMC to GEANT 4 format) to the usual GEANT 4 event sequence can be made. The service also provides for interaction with GEANT 4 through its standard user interface.

This section describes the possible user inputs, initialization, output, and various parameters of the simulation. Several useful features, including visualization, are also described.

5.1 Simulation input

The ATLAS simulation offers a choice for event generation. Events can be read from a file produced by any of the generators described in Sect. 3; one of the external generators can be configured and run concurrently; or commands can be provided for a single particle generator. The single particle generator can produce particles by the particle PDG identifier [91] with a configurable position and momentum. Neutral and charged gantinos (pseudo-particles without any interactions) are available for making material depth maps of the detectors and for debugging. The user may also choose to skip a certain number of events at the beginning of an input file, allowing 20 simulation jobs of 50 events each to a 1000 event input file without overlap or repetition.

Several cuts and transformations can then be made to the event. The vertex position is smeared to represent the luminous region within ATLAS. It can be shifted if the user desires, but both the shift and smear are given initial default values that represent ideal collisions within the ATLAS detector. The generated event can be rotated in any direction, though only rotation in \( \phi \) is physical. Primary particles are only passed through the detector simulation if they are within a specified range in \( \eta - \phi \). By default, primary particles with \( |\eta| \geq 6.0 \) are not simulated to save time. For example, typically beam-remnants are not simulated. This cut was chosen to ensure consistent response in the forward calorimeter: a cut at \( |\eta| = 6.0 \) allows a sufficient number of particles to scatter back into the forward detector from high-

\footnote{During early data taking, the beams will collide head-on. Therefore, no crossing angle is added to the simulated events for the time being.}
Fig. 5 CPU time per event as a function of the acceptance in $|\eta|$ of primary (generator) particles. If only central response is important, the CPU time required to simulate a single event is significantly different than if forward detectors must be included in the simulation. These two cases approximately represent the extremes of the plot. The time is normalized to CPU time for simulation of all primaries inside of $|\eta| < 3.0$ and increases between three-fold and eight-fold for simulation of all primaries in $|\eta| < 8.0$. The average of 200 simulated $t\bar{t}$ and minimum bias events was taken. Linear fits are overlaid.

$\eta$ without requiring an unacceptable amount of CPU time. Generally, an increase in one unit of pseudorapidity corresponds to an increase of 40–120% in CPU time, so that it is not possible to simulate the very forward detectors like LUCID during a standard simulation job. Figure 5 shows the $\eta$-dependence of the CPU time per event. The increase is approximately three-fold in $t\bar{t}$ events and eight-fold in minimum bias events from simulating particles in $|\eta| < 3.0$ to simulating particles in $|\eta| < 8.0$. The difference between the two types of events is primarily because the majority of activity in the minimum bias events is forward, and there is considerable central ($|\eta| < 3.0$) activity in the $t\bar{t}$ events.

At run time, either through job options or in the production system described in Sect. 2.2, seeds for pseudo-random number generators to be used by GEANT4, Athena, and any particle generators can be set. Different pseudo-random number generators may be configured for each. Since all random number seeds can be controlled, a single job is entirely reproducible. The seeds can also be written to a file or read from a file, providing an additional level of reproducibility.\footnote{Because of caching in GEANT4, it is not possible to reproduce an individual event or track without starting from the beginning of the job.}

The user must also select a layout for the detector. As described in Sect. 4, several layouts of the full detector and various test stands are available. Combined test beam simulation requires such a different configuration that an independent but similar core drives the PYTHON configuration and loading of user job options. The distinction is made at run time based on the detector or test stand configuration selected. The layout of the detector determines what other options are available to the user at run time. During simulation of the entire ATLAS detector, several additional options are available. For example, a neutron time cut (see Sect. 5.5) may be enabled. The magnetic field may be enabled and the field map may be selected.

The user may optionally select a set of run conditions for the simulation job, through which all options and flags are set and a pre-defined job is run. This option is particularly useful for testing and debugging.

5.2 Simulation initialization

Although the initialization in a standard Athena job occurs in a single step, for an ATLAS simulation job the initialization is broken into three steps to allow additional user intervention. Table 6 summarizes the processes that occur at each one of the three steps. The division of the initialization is such that most modifications to the simulation conditions can be accomplished in job options alone (i.e. without code modification). Normally, the user provides job options and allows the initialization to progress unhindered. Some parts of the job, for example the detector layout, are only loaded after the initialization has begun. In order to modify volumes after the layout is loaded, the user must intervene during the initialization. Only certain commands are effective at each stage of the initialization, since some parts of the GEANT4 simulation have been loaded and created while others have yet to be translated from dictionaries.

Stage one of the initialization occurs as soon as Athena is started. Several external PYTHON modules are loaded that provide basic functionality for any Athena job. The job properties provided by the user are read during this phase and
are locked. Once the job properties are locked, any significant modification to the running of a simulation job must be done by directly accessing the affected services and algorithms. This saves propagation of changes in the case of a late modification to a job property. Metadata that will be stored with the hit output file are gathered. External dependencies that require early initialization are loaded, providing a service for GeoModel, a service for database interaction, and a service for frozen showers (see Sect. 7.1). The event generation mode (reading external events, generating events from an external generator, or generating single particles on-the-fly) is determined, and any necessary configuration is included for the generator. A stream is opened for the output hit file, if necessary, and hit containers for each enabled subdetector are added to the new file. Finally, a service is created to interface with and control GEANT4, although at this point GEANT4 is not fully initialized.

Stage two of the initialization begins with the construction of the detector in PYTHON dictionaries. Dictionaries of physics regions, range cuts, and volumes in which to apply step limitation (see Sect. 5.5) are constructed, and all key properties are assigned to the detector facilities or built in dictionaries for later addition to the geometry. Each enabled piece of the ATLAS detector or test setup is recursively constructed in GeoModel according to the parameters specified in the geometry and conditions databases. After each subdetector has been constructed in GeoModel, it is translated recursively into an equivalent GEANT4 geometry. After this point in the initialization all volumes and regions are available to the user for modification. Next, the Monte Carlo truth strategies (see Sect. 5.3) are added to the simulation. The magnetic field is then loaded. Under normal circumstances, the field is a map loaded from an external data file, the name of which is specified in the geometry database according to the geometry layout selected. The user may optionally choose to load data from one of the magnetic field test configurations, rather than the standard ATLAS magnetic field, or to create a new, basic magnetic field. The physics list (see Sect. 5.4) to be used for simulation is also set at this point.

User actions are then initialized. GEANT4 allows a user to insert pieces of code in various places throughout the simulation event loop, including after each step, when each track is queued (“stacked”), before and after each event, before and after each track is simulated, and before and after each run. By default, ATLAS includes user actions that monitor simulation time, memory, and the number of tracks generated during each event. A neutrino cut (see Sect. 5.5) is also implemented as a track-stacking action. Whenever a new track is queued, its type is checked. The LAr calorimeters also use end-of-event actions for merging hits to save space prior to storage. All truth storage strategies (see Sect. 5.3) are implemented as stepping actions that store interesting interactions based on the type of process, detector region, and energies of the particles involved. Users may also configure their own actions and add them to the simulation in the same way. Examples have been constructed for integrating interaction lengths or radiation lengths through the detector when making geantino maps, for stopping or killing particles if certain conditions are met, and for turning on additional output only under specific conditions in order to study a bug or issue without having to sift through enormous log files.

Stage three of the initialization completes the job preparation. During this stage, the fast simulation models are built and added to the volumes to which they have been assigned. Any physics regions that will be used are constructed. Sensitive detectors are built and assigned to the regions of the detector that are to be made sensitive (i.e., in which hits will be stored). GEANT4’s run manager and physics models are initialized. Recording envelopes are added (see Sect. 5.3), and any visualization that has been enabled by the user is initialized (see Sect. 5.7).

Once the initialization is complete and all the necessary elements have been loaded into memory, the event loop begins.

### 5.3 Monte Carlo truth information

The GEANT4 simulation adds to the Monte Carlo truth record already defined during generation (see Sect. 3.6).

---

8Here, run is used in the GEANT4 sense to refer to a finite set of events within a simulation job. Several runs may comprise a job, and each run may include an arbitrary number of events.
Far too many secondary tracks are produced during detector simulation to store information for every interaction. Only those interactions which are of greatest relevance to physics analyses are saved, according to several saving rules ("strategies"). Most are applicable only to the inner detector. For each interaction that satisfies any of the storage criteria, the incoming particle, step information, vertex, and outgoing particles are included in the truth record. Later in the software chain, individual track segments are recombined so that, for example, a single electron that undergoes several bremsstrahlung events along its path is counted as only one "true" particle.

The strategies include (with all cuts on kinetic energy)\(^9\):

- In the inner detector, bremsstrahlung vertices are stored if the primary electron or muon has an energy above 500 MeV and the photon produced has an energy above 100 MeV.
- In the inner detector, ionization vertices are stored if the primary particle has an energy above 500 MeV and the electron generated has an energy above 100 MeV.
- In the inner detector, hadronic interaction vertices are stored if the primary particle has an energy above 500 MeV.
- In the inner detector, decay vertices are stored if the decaying particle has an energy above 500 MeV.
- In the inner detector, the conversions of photons above 500 MeV are stored.
- In the calorimeter, muon bremsstrahlung vertices are stored if the primary muon's energy was above 1 GeV and the photon generated is above 500 MeV.

All cuts and regions of applicability are made configurable, so that any energy cut-offs can be modified and a strategy can be assigned to any volume in the simulation. Additional rules could be constructed, for example, for tracking shower development within the calorimeter, but many would consume too much CPU time and disk space for use in standard simulation jobs.

Standard simulation jobs also define several volumes that are used to record all particles escaping part of the detector. All tracks above 1 GeV are typically recorded at the end of the inner detector, the end of the calorimeter, and the end of the muon system (and the end of the ATLAS world volume). It is possible for the user to configure the simulation at run time to add additional volumes to the list of these recording volumes. In each case, tracks are saved as they exit each volume.

### 5.4 Physics list

Physics lists include all numerical models that describe the particles' interactions in the GEANT4 simulation. Models are generally good for a single type of interaction and over a limited energy range. The GEANT4 collaboration provides several combinations of these models that have been tailored to various scenarios as standard physics lists that ship with each distribution. In order to enhance reproducibility and ensure that validated combinations of models are used, only those physics lists provided by the GEANT4 collaboration are used by the ATLAS simulation. One exception is allowed, namely transition radiation. Transition radiation is crucial for the tracking portion of the inner detector and is added to each physics list.

There are several physics lists that are used by ATLAS:

- **QGSP_BERT**—the physics list used for all simulation production after 2008. The list includes the Quark–Gluon String Precompound model (QGSP) and the Bertini intranuclear cascade model (BERT)\(^4\) as part of the hadronic physics package. The electromagnetic physics package includes step-limiting Multiple Coulomb Scattering (MSC).
- **QGSP_EMV**—the physics list used for simulation production before 2008. This list included the QGSP model, but without the Bertini cascade. The MSC of this list was not allowed to limit the step, so it is labeled an electromagnetic variant (EMV).
- **QGSP_BERT_HP**—the physics list used for neutron fluence studies and comparisons with the FLUKA simulation package\(^9\). This list includes the QGSP and Bertini models, step-limiting MSC, and additional "high-precision" low-energy neutron physics models.

A step limitation process that controls the maximum allowed step length of a charged particle was added in the inner detector when using the QGSP_EMV physics list. It helped the simulation to better reproduced test beam and cosmic ray data. The step-limiting MSC that is a part of QGSP_BERT was found to agree equally well with data, and therefore the step limitation was removed from simulation with QGSP_BERT.

These physics lists were studied in detail for each sub-detector\(^9\). Table 7 shows the number of steps, number of hits in sensitive detector regions, and number of secondary particles with kinetic energy above 50 MeV and 1 GeV within several regions of the detector and for the whole of ATLAS using both the QGSP_EMV and QGSP_BERT physics lists. Sensitive detectors to record calibration hits (described in Sect. 5.6) are included in the cryostat of the LAr calorimeter. The average was taken of 50 $t\bar{t}$ events, where there were on average 482 primary (generator-level) particles per event. The calorimeter clearly

\(^9\)For the most recent production, cuts are applied on transverse momentum, $p_T > 100$ MeV, rather than on kinetic energy. The lower cut allows for a study of tracking performance in minimum bias events where it may be possible to reconstruct tracks down to only a few hundred MeV.
Table 7  Number of steps, number of hits in sensitive detector (SD) regions, and number of secondary particles with kinetic energy above 50 MeV and 1 GeV within several regions of the detector and for the whole of ATLAS, using both the QGSP_EMV and QGSP_BERT physics lists. The average was taken of 50 $t\bar{t}$ events, where the average number of primary tracks per event was 482. Sensitive detectors to record calibration hits are included in the cryostat of the LAr calorimeter.

| Detector          | QGSP_EMV  | Hits in SD | Sec. above 50 MeV | Sec. above 1 GeV |
|-------------------|-----------|------------|-------------------|-----------------|
| Inner detector    | $1.80 \times 10^6$ | $3.10 \times 10^5$ | 1,570             | 260             |
| Calorimetry       | $1.87 \times 10^7$ | $6.87 \times 10^6$ | 39,900            | 2,040           |
| Muon system       | $1.90 \times 10^6$ | 1,030      | 7,820             | 332             |
| Total ATLAS       | $2.24 \times 10^7$ | $7.18 \times 10^6$ | 49,300            | 2,630           |

| Detector          | QGSP_BERT  | Hits in SD | Sec. above 50 MeV | Sec. above 1 GeV |
|-------------------|------------|------------|-------------------|-----------------|
| Inner detector    | $2.13 \times 10^6$ | $1.98 \times 10^5$ | 1,450             | 269             |
| Calorimetry       | $3.93 \times 10^7$ | $1.36 \times 10^7$ | 40,100            | 2,170           |
| Muon system       | $2.69 \times 10^6$ | 1,285      | 8,210             | 385             |
| Total ATLAS       | $4.41 \times 10^7$ | $1.38 \times 10^7$ | 49,700            | 2,820           |

Table 8  Number of steps for various processes and detector regions during simulation with the QGSP_EMV physics list. The average was taken of 50 $t\bar{t}$ events, where the average number of primary tracks per event was 482. The “other processes” in the inner detector are primarily step limitation processes.

| Process                  | Inner detector | Calorimeter | Muon system |
|--------------------------|----------------|-------------|-------------|
| Transportation           | $1.50 \times 10^6$ | $9.33 \times 10^6$ | $2.15 \times 10^5$ |
| MSC                      | 4,910          | 1.09 $\times 10^5$ | 5,200       |
| Photoelectric effect     | 6,060          | 1.32 $\times 10^6$ | 2.03 $\times 10^5$ |
| Compton scattering       | 12,800         | 1.43 $\times 10^6$ | 4.26 $\times 10^5$ |
| Ionization               | 1.08 $\times 10^5$ | 4.97 $\times 10^6$ | 8.10 $\times 10^5$ |
| Bremsstrahlung           | 6,310          | 1.28 $\times 10^6$ | 1.87 $\times 10^5$ |
| Conversion               | 434            | 82,400      | 17,300       |
| Annihilation             | 291            | 82,800      | 17,800       |
| Decay                    | 254            | 2,320       | 538          |
| Other hadronic interaction| 1,710        | 1.23 $\times 10^6$ | 21,600       |
| Other process            | 1.56 $\times 10^5$ | 4,800       | 831          |
| Total                    | $1.80 \times 10^6$ | $1.87 \times 10^7$ | $1.90 \times 10^6$ |

dominates the total number of steps and hits in sensitive detector for both physics lists. The muon system, though it has a comparable number of hits, consists mostly of shielding and therefore has far fewer hits in sensitive detector regions. The numbers of steps divided into different process types for QGSP_EMV and QGSP_BERT are listed in Tables 8 and 9. In both cases, transportation processes dominate the inner detector simulation, while electromagnetic physics and transportation dominate the calorimeter and the muon system.

Simulation time was also examined for each physics list. Simulation using the QGSP_BERT physics list consumes ~2.5 times more CPU time than does simulation with the QGSP_EMV physics list. However, applying a neutron time cut (see Sect. 5.5) with the QGSP_BERT list reduces simulation time by more than 30%. Simulation with QGSP_BERT_HP requires approximately five times more CPU time than QGSP_EMV. Therefore, the QGSP_BERT_HP physics list cannot be used for standard simulation.

5.5 Simulation optimizations

In order to optimize use of both disk space and CPU time, several other modifications are made to the standard GEANT4 simulation [94, 95]. Comparing the QGSP_BERT physics list to the QGSP_EMV physics list, approximately three times as many neutrons are generated in typical hard scattering events, and they travel approximately three times further. These neutrons cause an increase in the output hit file size of approximately 75% as well as an increase in CPU time per event for hard scattering events. A GEANT4 neutron time cut is, therefore, applied which removes all neutrons...
150 ns after the primary interaction. This cut was found to be sufficient time for the hadronic shower development and did not degrade the energy scale or energy resolution of the calorimeters. The total energy deposition of the shower is changed by $\sim 1\%$. Output files are the same size when using the QGSP_BERT physics list with this cut enabled as they are when using the QGSP_EMV physics list without a neutron time cut. The simulation time required for QGSP_BERT is reduced by 10–15% when the neutron cut is enabled.

Neutrinos are also removed as soon as they are created in the simulation. No particle is allowed by GEANT 4 to step through more than one volume at a time. Therefore, neutrinos may require several thousand steps to exit the entire ATLAS detector. They may therefore consume a noticeable fraction of simulation time, even though their interaction probability is practically null. The removal is done when the particles are stacked.

Range cuts are GEANT4 parameters that control the creation of secondary electrons or photons during bremsstrahlung and ionization processes. If the expected range of the secondary is less than some minimum value, the energy of that secondary particle is deposited at the end of the primary particle's step and no separate secondary is produced. Effectively, this parameter defines an energy scale at which particle propagation may be ignored. By increasing the range cuts throughout the detector one can decrease the CPU time required per event. Particularly near boundaries and thin materials, the detector’s sampling fraction may be affected if the range cuts are too large. Range cuts can be specified separately for electrons, positrons, and photons, but in ATLAS the same distance is used for all three. Range cuts are specified as a distance, and for each material the distance is translated into an energy based on the average energy loss of a particle in that material. For the majority of the ATLAS detector, range cuts take a default value of 1 mm. Exceptions are listed in Table 10. Deviations usually occur in sensitive volumes that are very thin, where it is important to correctly calculate the sampling fraction of the detector or model the energy deposition. Reduced range cuts are also applied to very thin volumes that are adjacent to sensitive volumes for the same reason. In the monitored drift tube muon chambers, for example, range cuts are only reduced in the thin aluminum tubes surrounding the sensitive detector (gas)—the gas itself takes the standard 1 mm cuts. In some shielding volumes it may be possible to relax range cuts considerably without degrading physics performance.

GEANT4 uses a set of parameters to control errors and accumulated biases on charged particles transportation through a magnetic field. Because the equation of motion is solved numerically, the user must select the numerical integration method to be used, including the order of integration, and the tolerances on the errors of the step. ATLAS has chosen to use the GEANT4 standard fourth-order Runge–Kutta method with the default error parameters for the majority of the detector. These parameters are generally satisfactory and result in errors and biases that are less than the position resolution of the detector. In the inner detector, however, tracks were found to be shifted sufficiently that detector residuals were affected. Here, stepping parameters were tightened by an order of magnitude. Further optimization of the stepping algorithms of GEANT4 has been undertaken, including the configuration of the choice of stepper and stepping parameters as a function of the initial particle type, energy, and position within the detector. Such a configuration can allow more careful stepping of muons in the calorimetry without degrading the total performance of the simulation significantly. Muons in particular can accumulate a significant bias after passing through all the sampling layers of the calorimeter, making more accurate tracking necessary. As a fourth-order stepper requires four values of the magnetic

| Process                          | Inner detector | Calorimeter | Muon system |
|----------------------------------|----------------|-------------|-------------|
| Transportation                   | $1.76 \times 10^6$ | $1.46 \times 10^7$ | $2.31 \times 10^5$ |
| MSC                              | $2.31 \times 10^5$ | $1.48 \times 10^7$ | $5.200$ |
| Photoelectric effect             | $6.760$        | $1.37 \times 10^6$ | $2.32 \times 10^5$ |
| Compton scattering               | $14.800$       | $1.66 \times 10^6$ | $5.03 \times 10^5$ |
| Ionization                       | $1.03 \times 10^5$ | $4.81 \times 10^6$ | $9.71 \times 10^5$ |
| Bremsstrahlung                   | $6.060$        | $1.22 \times 10^6$ | $1.92 \times 10^5$ |
| Conversion                       | $416$          | $86.800$    | $18.100$ |
| Annihilation                     | $212$          | $1.670$     | $402$ |
| Decay                            | $2,190$        | $6.66 \times 10^5$ | $1.23 \times 10^5$ |
| Other hadronic interaction       | $426$          | $25.400$    | $5.720$ |
| Other process                    | $2,190$        | $6.66 \times 10^5$ | $1.23 \times 10^5$ |
| Total                            | $2.13 \times 10^6$ | $3.93 \times 10^7$ | $2.69 \times 10^6$ |
field to be calculated, optimization of magnetic field map access will also be key to improving the performance of the simulation’s tracking.

5.6 Hit storage format

The output from the simulation is a hit file, containing some metadata describing the configuration of the simulation during the run, all requested truth information, and a collection of hits for each subdetector. The hits are records of energy deposition, with position and time, during the simulation. Each subdetector is responsible for implementing their own sensitive detector for the selection, processing, and recording of these hits. In most subsystems, including the inner detector and muon system, this consists simply of recording all hits that occur in sensitive regions of the detector for subsequent storage. Some additional manipulation is done at the end of each event to compress the output as much as possible; still, the files are typically 2 MB per event for hard scattering events (e.g. $t \bar{t}$ production).

The file size is large, mostly due to the inner detector, for which the majority of hits are independently stored. Merging hits there is difficult, since they tend to be isolated and cannot normally be merged across readout channels. These consume typically 60% of the disk space in a hit file (e.g. 65% of the hit file for $t \bar{t}$ events). In the calorimetry, there are far too many hits created by electromagnetic and hadronic showers for the individual storage of a four-vector for each (see Sect. 5.4). Instead, hit merging occurs at the end of each event. By optimizing time binning, hits can be compressed to a large extent. About 10% of the hit file is consumed by optional “calibration hits” for the calorimeters, hits in dead material, stored to improve the detector calibration and missing energy calculation and to study simulation-based calorimeter calibration schemes. Under normal circumstances, the muon systems contribute a negligible portion of the hit file. The contributions by subdetector can be found in Table 11 for the average of 50 simulated $t \bar{t}$ events. Here and elsewhere in this paper, file sizes are without compression and are taken from ROOT [96]. In practice, compression reduces the actual disk space required for the files, but file-level metadata adds several hundred kilobytes.

By comparing Tables 7 and 11, one can understand these numbers in terms of hits in the sensitive detector region. Although the muon system is large, the majority of it is shielding. Therefore, it collects far fewer hits than the other subsystems and requires less disk space for the hit records. The calorimetry produces 95% of the hits in sensitive regions during simulation. Because of the compression applied prior to storage, the calorimetry comprises only 25% of the hit file.
5.7 Visualization

Visualization is used to understand anomalies or features in odd events, occasionally to debug errors due to geometry, and to check for overlaps and touching volumes in the geometry that can be spotted by eye. Although GEANT4 contains viewing software of its own, because the geometry of ATLAS must be translated from GeoModel into GEANT4 format it is useful to use a viewer that can construct geometry directly from its GeoModel description. A general purpose three-dimensional event display program, VP1 [97], has been developed specifically for ATLAS. It is optimized specifically for the visualization of the ATLAS geometry and is arguably the most useful tool for understanding and debugging of detector description across all ATLAS subsystems. Two examples of VP1 event displays are in Figs. 6 and 7. Ray Tracer, the GEANT4 visualization utility, has also been used to visualize portions of the detector containing some exotic shapes.

VP1, as well as other event display programs used in ATLAS (e.g. Atlantis [98] and Persint [99]), are mostly used for visualizing real and simulated events after they have run through the common reconstruction software. The VP1 viewer can be injected directly into the simulation job in order to visualize events immediately after the simulation step.

6 Digitization

The ATLAS digitization software converts the hits produced by the core simulation into detector responses: “digits.” Typically, a digit is produced when the voltage or current on a particular readout channel rises above a pre-configured threshold within a particular time window. Some subdetector’s digit formats include the signal shape in detail over this time, while others simply record that the threshold has been exceeded within the relevant time window.

The peculiarities of each subdetector’s charge collection, including cross-talk, electronic noise and channel-dependent variations in detector response are modelled in subdetector-specific digitization software [79, 82, 100–102]. The various subdetector digitization packages are steered by a top-level PYTHON digitization package which ensures uniform and consistent configuration across the subdetectors. The properties of the digitization algorithms were tuned to reproduce the detector response seen in lab tests, test beam data, and cosmic ray running. Dead channels and noise rates are read from database tables to reproduce conditions seen in a particular run. In some cases, dead channels are removed during the reconstruction step.

The digits of each subdetector are written out as Raw Data Objects (RDOs). For some subdetectors this requires the digits produced to be converted to RDOs by a second algorithm during the digitization process. For others there is no intermediate digit object and RDOs are produced di-

---

\[\text{VP1 is short for Virtual Point 1.}\]
rectly from the hits. In addition to RDOs, the digitization algorithms can also produce Simulated Data Objects (SDOs). These SDOs contain information about all the particles and noise that contributed to the signal produced in the given sensor and the amount of energy contributed to the signal by each. The relationship between RDOs and SDOs depends on the particular subdetector. For example, in the SCT each RDO represents a group of consecutive strips which recorded a hit, whereas one SDO is produced for each strip where energy was deposited by a particle in the Monte Carlo truth tree. No SDOs are created in the calorimeter. SDOs are mainly used for determining tracking efficiency and fake track rates.

Simulating the detector readout in response to a single interesting hard scattering interaction is unrealistic. In reality, for any given bunch crossing there may be multiple proton–proton interactions. In addition to the hard scattering which triggers the detector readout, many inelastic, non-diffractive proton–proton interactions may appear. These interactions must be included in a realistic model of detector response. The effects of beam gas and beam halo interactions, as well as detector response to long-lived particles, must be incorporated. These interactions are treated separately at the event generation and simulation stages. Within a digitization job, hits from the hard scattering are overlaid with those from the requested number of these additional interactions before the detector response is calculated. Because of long signal integration times, most subdetector responses are affected by interactions from neighboring bunch crossings as well. Therefore, additional interactions offset in time are overlaid as necessary. The overlaying off these various types of events, known collectively as “pile-up,” is described in Sect. 6.2.

Before reconstruction can be run, bytestream data from the real detector must be converted into RDO format. As mentioned above, the digitization usually avoids this step by writing out RDOs directly. However, in order to do simulation studies with the High Level Trigger it is necessary to translate the RDO files into bytestream format. There is some loss due to truncation in the first conversion from RDO to bytestream, but the inverse operation is basically lossless. Having the ability to convert output in both directions also allows evaluation of the conversions themselves.

6.1 Digitization configuration

The ATLAS digitization takes as input hit files produced by the ATLAS simulation. For pile-up simulation, there are also input hit files for each type of background interaction to be overlaid. In such cases it is the main hard scattering event which sets the run number and event number. Run and event numbers from overlaid events are ignored.

The digitization steering package exists entirely in the PYTHON layer and configures how the digitization will be performed before the event loop starts. This configuration is highly flexible, but also ensures that sensible default values are given for each configurable property of the job. In the configuration of digitization jobs, the user may specify the number of events to digitize, the number of leading events to skip in the input file, the input hit file(s), and the output file. Digitization and writing out of RDOs may be enabled or disabled by subdetector. In order to ensure consistency,
Digitization options also include the following:

Detector Noise Simulation: Detector noise simulation can be turned off in the inner detector, calorimeter or muon spectrometer or any combination thereof. This is useful for data overlay jobs where noise is taken from real data events and for studies using a noise-free detector.

Random Number Services: The type of random number engine to be used in all digitization algorithms can be specified (Ranlux64, the default, or Ranecu [103]). Each algorithm has one or more random number streams. Random number seeds can be initialized from a text file or set in job options. The user may alternately specify an offset from the default values of the seeds, to be used in all streams.

Metadata: In the default configuration, metadata from the simulation stage are used to configure the physics list (for setting the sampling fraction of the calorimeters) and the detector layout. The metadata can be overridden.

Pile-up Background Events: The overlay of minimum bias, cavern background, beam gas and beam halo events can all be configured separately. In each case the mean number of events (if any) per bunch crossing to be overlaid and a collection of files containing the events to be overlaid onto the signal events can be specified.

Beam Properties: The LHC beam bunch spacing can be configured, as can the number of bunch crossings to overlap before and after the hard scattering event.

Detector Conditions: Default detector conditions (including, e.g., dead electronics and noisy channels) are associated with each detector layout. Non-default conditions may be specified globally or by subdetector for use in digitization.

After a check to make sure that at least one subdetector has been left switched on, the input and output streams are initialized. GeoModel is initialized using the detector layout and conditions versions read from the hits file metadata or specified by the user. Setting the detector layout version to be different from that used in the simulation is possible, but considered to be an expert action. The magnetic field service is initialized at this point. It is necessary because the magnetic field affects charge propagation from the active regions of the detector to the readout surfaces.

At this point, caches for pile-up events are created and configured with the appropriate collection of hits files as well as the number of events to be overlaid per bunch crossing. These caches are controlled by an overall pile-up manager service. A second pile-up service is created to hold information about the time window within which interactions can affect the response recorded by each subdetector. During the initialization stage, this information can be combined with the bunch spacing to calculate the number of bunch crossings which should be simulated for each subdetector for each event.

Subsequently, the subdetector digitization algorithms are configured and added to the sequence of algorithms to be run in the job. The collections of RDOs, hits, and truth information which are to be recorded are added to the output stream. Digitization algorithms exist for the following subdetectors:

Inner Detector: BCM, silicon pixel tracker, SCT, and TRT.

Calorimeter: LAr and tile calorimeters. Separate algorithms also exist to simulate the formation of trigger towers in the calorimeters, which serve as inputs to the level one trigger.

Muon Spectrometer: Cathode Strip Chambers, Monitored Drift Tubes, Resistive Plate Chambers and Thin Gap Chambers.

If requested, the level one trigger simulators are added to the algorithm sequence, provided that the digitization of the relevant parts of ATLAS have been turned on. The default mode of simulation production is to run the level one trigger simulation during the reconstruction step rather than as part of the digitization step.

As the digitization algorithm for each subdetector is configured, the names and seeds for the random number streams it requires are added to a list. In the case where seeds are to be read in from a file, the default list of stream names and their seeds are replaced by the file contents. Once all algorithms have been configured, the list is used to configure the random number service. Separate random number streams are used for each subdetector digitization algorithm and give the same result independent of what is used for the other subdetectors.11

Much of the job configuration information, along with the detector layout version, is written to the output file as digitization metadata. The run number provided in the simulation metadata is used to establish a validity range for the digitization metadata corresponding to the current run only. At this point the digitization job is fully configured and the event loop begins.

6.2 Pile-up

To simulate pile-up, various types of events are read in, and hits from each are overlaid. The different types considered can be configured at run time, and normally comprise signal, minimum bias, cavern background, beam gas, and beam halo events. The number of events to overlay of each type can affect the response recorded by each subdetector. During the initialization stage, this information can be combined with the bunch spacing to calculate the number of bunch crossings which should be simulated for each subdetector for each event.

Subsequently, the subdetector digitization algorithms are configured and added to the sequence of algorithms to be run in the job. The collections of RDOs, hits, and truth information which are to be recorded are added to the output stream. Digitization algorithms exist for the following subdetectors:

Inner Detector: BCM, silicon pixel tracker, SCT, and TRT.

Calorimeter: LAr and tile calorimeters. Separate algorithms also exist to simulate the formation of trigger towers in the calorimeters, which serve as inputs to the level one trigger.

Muon Spectrometer: Cathode Strip Chambers, Monitored Drift Tubes, Resistive Plate Chambers and Thin Gap Chambers.

If requested, the level one trigger simulators are added to the algorithm sequence, provided that the digitization of the relevant parts of ATLAS have been turned on. The default mode of simulation production is to run the level one trigger simulation during the reconstruction step rather than as part of the digitization step.

As the digitization algorithm for each subdetector is configured, the names and seeds for the random number streams it requires are added to a list. In the case where seeds are to be read in from a file, the default list of stream names and their seeds are replaced by the file contents. Once all algorithms have been configured, the list is used to configure the random number service. Separate random number streams are used for each subdetector digitization algorithm and give the same result independent of what is used for the other subdetectors.11

Much of the job configuration information, along with the detector layout version, is written to the output file as digitization metadata. The run number provided in the simulation metadata is used to establish a validity range for the digitization metadata corresponding to the current run only. At this point the digitization job is fully configured and the event loop begins.

6.2 Pile-up

To simulate pile-up, various types of events are read in, and hits from each are overlaid. The different types considered can be configured at run time, and normally comprise signal, minimum bias, cavern background, beam gas, and beam halo events. The number of events to overlay of each type

11Here “digitization algorithm” does not include the calorimeter trigger tower simulation algorithms, which require the corresponding calorimeter digitization to be performed. Similarly, the level one trigger simulation requires the simulation and digitization of the expected trigger inputs to give meaningful results.
per bunch crossing may also be set at run time and is a function of the luminosity to be simulated. The mean number of interactions per bunch crossing (BC), for example 23 at the design luminosity of $10^{34}$ cm$^{-2}$ s$^{-1}$ with 25 ns bunch spacing, depends linearly on luminosity and bunch spacing. However, this number is Poisson-distributed, with a long tail beyond the most probable value. Thus, a substantial fraction of the bunch crossings will have more than the average number of interactions. In addition, the ATLAS subdetectors are sensitive to hits several bunch crossings before and after the BC that contains the hard scattering event (which triggers the readout). Table 12 shows the simulation window for each detector along with the corresponding number of bunch crossings for 25 ns and 75 ns bunch spacing. All of these detector and electronic effects are taken into account during the pile-up event merging.

6.2.1 Cavern background

Neutrons may propagate through the ATLAS cavern for a few seconds before they are thermalized, thus producing a neutron–photon gas. This gas produces a constant background, called “cavern background,” of low-energy electrons and protons from spallation. The cavern background consists mainly of thermalized slow neutrons, long-lived neutral kaons and low-energy photons escaping the calorimeter and the forward shielding elements. Muon detectors are most affected by high cavern-background rates. The radiation levels to be expected in the ATLAS cavern scale with luminosity, and they have been simulated during pile-up jobs, along with the corresponding numbers of bunch crossing simulated in the case of 25 ns bunch spacing and 75 ns bunch spacing.

Table 12  The time window (relative to the current bunch crossing) during which interactions in each subdetector are simulated during pile-up events. The mean number of interactions per bunch crossing may also be set at run time and is a function of the luminosity to be simulated. The mean number of interactions per bunch crossing (BC), for example 23 at the design luminosity of $10^{34}$ cm$^{-2}$ s$^{-1}$ with 25 ns bunch spacing, depends linearly on luminosity and bunch spacing. However, this number is Poisson-distributed, with a long tail beyond the most probable value. Thus, a substantial fraction of the bunch crossings will have more than the average number of interactions. In addition, the ATLAS subdetectors are sensitive to hits several bunch crossings before and after the BC that contains the hard scattering event (which triggers the readout). Table 12 shows the simulation window for each detector along with the corresponding number of bunch crossings for 25 ns and 75 ns bunch spacing. All of these detector and electronic effects are taken into account during the pile-up event merging.

| Subdetector      | Simulation window [ns] | No. Bunch crossings (25 ns bunch spacing) | No. Bunch crossings (75 ns bunch spacing) |
|------------------|------------------------|------------------------------------------|------------------------------------------|
| BCM              | $-50, +25$             | 4                                        | 1                                        |
| Pixel trackers   | $-50, +25$             | 4                                        | 1                                        |
| SCT              | $-50, +25$             | 4                                        | 1                                        |
| TRT              | $-50, +50$             | 5                                        | 1                                        |
| LAr calorimeter  | $-801, +126$           | 38                                       | 12                                       |
| Tile calorimeter | $-200, +200$           | 17                                       | 5                                        |
| Muon chambers    | $-1000, +700$          | 69                                       | 23                                       |

The simulation is then carried out using the full detector geometry and GEANT4, and hits are stored. The cavern events are mixed, with a safety factor of up to 10, at the digitization level with the minimum bias and signal events.

There are a number of issues with the current simulation of the cavern background. The original primary cavern events were generated in an older version of PYTHIA where the generated particle density is a factor of two lower than in the newer versions of PYTHIA [42, 43]. The statistics for the available cavern events are limited: 40,000 events are available with a safety factor of 1; 10,000 events are available with a safety factor of 2 or 5; and 5,000 events are available with a safety factor of 10. Because of the limited statistics, a number of monitored drift tubes in the muon detector fire more often than expected (i.e. there are spikes in the hit response of the detector). Additionally, neutral particles are tracked through the entire detector during simulation, thus producing additional hits from particles that should have been removed at the edges of the muon chamber envelopes (multiple counting).

In the short term, the problem of limited statistics of the cavern events has been alleviated by taking advantage of the $\phi$-symmetry of the muon spectrometer: the cavern events are rotated and re-simulated eight times or more (in multiples of eight). Further improvement in the available cavern statistics can be achieved by repeating the simulation of the cavern events many times with different random number seeds, since the probability of a neutral interaction is very low, of the order 1%.
Table 13  Container size on disk in RDO files. Columns two and three show the average of 50 \( t\bar{t} \) events digitized in the absence of pile-up. These events were chosen because they produce large energy deposits throughout the detector. Columns four and five show the average of 50 \( t\bar{t} \) events digitized in the presence of pile-up with \( 10^{13} \text{ cm}^{-2} \text{s}^{-1} \) luminosity and 25 ns bunch spacing.

| Category                  | No Pile-up | No Pile-up | Pile-up | Pile-up |
|---------------------------|------------|------------|---------|---------|
|                           | space on disk [kB/event] | percentage of file | space on disk [kB/event] | Percentage of file |
| Inner detector RDOs       | 187        | 7.6        | 322     | 11.5    |
| Inner detector SDOs       | 247        | 10.0       | 333     | 11.9    |
| Calorimeter raw channels  | 995        | 40.3       | 1006    | 35.9    |
| Calorimeter calibration hits | 601      | 24.3       | 601     | 21.4    |
| Muon spectrometer RDOs    | 1          | 0.04       | 27      | 1.0     |
| Muon spectrometer SDOs    | 1          | 0.05       | 59      | 2.1     |
| Level one trigger         | 289        | 11.7       | 300     | 10.7    |
| Truth                     | 147        | 6.0        | 151     | 5.4     |
| Headers                   | >1         | 0.01       | 2       | 0.1     |
| Total                     | 2469       | 100.00     | 2801    | 100.00  |

6.2.2 Beam halo and beam gas

Beam halo is the background resulting from interactions between the beam and upstream accelerator elements. The flux from upstream (in the tunnel and collimators) is provided by the LHC Machine Division [106, 107]. Beam halo events are generated as discrete particle losses against the upstream collimators. The LHC machine division has estimated the proton loss rate in design conditions as being on the order of 1 MHz. FLUKA simulation of the last 150 m of the beamline indicates that daughter particles from these proton losses will reach the cavern wall (23 m from the interaction point) at a rate of \( \sim 400 \text{ kHz} \). This flux is input to the normal GEANT4 simulation to produce hit files.

Beam gas includes the residual hydrogen, oxygen, and carbon gasses in the ATLAS beam pipe. Beam gas interaction events are generated with Hijing (see Sect. 3.2.4) with appropriate time offsets. The interactions are allowed to take place anywhere in the beam pipe of ATLAS, 23 m in either direction from the interaction point.

6.2.3 Pile-up with real data

The pile-up mechanism described above will not work with real data, because it begins at the hit level. One must instead overlay events beginning from detector electronics output (RDOs). One may collect minimum bias, cavern background, beam halo, and beam gas backgrounds from the same “zero bias” trigger used to understand detector electronic noise. Then, one would overlay hits from simulated hard scattering events onto the zero bias trigger data to simulate the pile-up. The zero bias trigger data needed for this type of event overlay can be selected at random from the unfilled bunch crossings. The subdetectors should be read out with as little zero-suppression as is possible and with the HLT in “pass-through” mode (i.e. without further filtering). One can use bunch-by-bunch luminosity information to correctly weight the event sample for pile-up studies.

In principle one needs as many zero bias events as generated events, but in practice zero bias events can be reused with independent simulated data sets without introducing any bias. During data taking, zero bias events are sampled at all times, because detector and cavern conditions are likely to vary with time. The zero-bias events could, for example, be collected exactly one orbit after a high-\( p_T \) trigger has fired. Appropriately pre-scaled to the output rate needed for simulations (on the order of 1–2 Hz, or about 1% of the recorded events), this means that the rate follows the luminosity and that the bunch structure is guaranteed to be right.

6.3 RDO storage format

The ATLAS detector electronics produce data in bytestream format. The RDO format can be thought of as a POOL-compatible version of the byte stream.\(^{13}\) The file size on disk is typically around 2.5 MB per event for hard scattering events (e.g. \( t\bar{t} \) production) and increases in the presence of pile-up. Table 13 shows an example of the disk consumption by container for 50 \( t\bar{t} \) events without pile-up and with pile-up at \( 10^{13} \text{ cm}^{-2} \text{s}^{-1} \). In the absence of pile-up, one of the main consumers of disk space are the calibration hit collections, as described in Sect. 5.6, which are copied directly from the hit file to the RDO. As pile-up luminosity increases however the inner detector containers become increasingly significant.

\(^{12}\)The zero bias trigger is not a minimum bias trigger.

\(^{13}\)POOL compatibility requires separate transient and persistent object representation.
7 Fast simulations

Because of the complicated detector geometry and detailed physics description used by the ATLAS GEANT4 simulation, it is impossible to achieve the required simulated statistics for many physics studies without faster simulation. To that end, several varieties of fast simulation programs have been developed to complement the GEANT4 simulation. In this section, the standard GEANT4 simulation will be referred to as "full simulation."

Almost 80% of the full simulation time is spent simulating particles traversing the calorimeter, and about 75% of the full simulation time is spent simulating electromagnetic particles. The Fast G4 Simulation aims to speed up this slowest part of the full simulation [108, 109]. The approach taken, therefore, is to remove low energy electromagnetic particles from the calorimeter and replace them with pre-simulated showers stored in memory. Using this approach, CPU time is reduced by a factor of three in hadronic events (e.g. $t\bar{t}$ production) with little physics penalty. This simulation may eventually become the default simulation for all processes that do not require extremely accurate modeling of calorimeter response or electromagnetic physics.

ATLFAST-I has been developed for physics parameter space scans and studies that require very large statistics but do not require the level of detail contained in the full simulation [110, 111]. Truth objects are smeared by detector resolutions to provide physics objects similar to those of the reconstruction. Object four-vectors are output, without any detailed simulation of efficiencies and fakes. A factor of 1000 speed increase over full simulation is achieved with sufficient detail for many general studies.

ATLFAST-II is a fast simulation meant to provide large statistics to supplement full simulation studies. The aim is to try to simulate events as fast as possible while still being able to run the standard ATLAS reconstruction. ATLFAST-II is made up from two components: the Fast ATLAS Tracking Simulation (Fatras) for the inner detector and muon system simulation [112] and the Fast calorimeter Simulation (FastCaloSim) for the calorimeter simulation. Optionally, any subdetector can be simulated with GEANT4 to provide the higher level of accuracy without the same CPU time consumption as full simulation of the entire detector. An improvement over full simulation time of a factor of 10 is achieved with full GEANT4 inner detector and muon simulation and FastCaloSim, and a factor of 100 is achieved with Fatras and FastCaloSim.

7.1 Fast G4 simulation

The Fast G4 Simulation reduces CPU time consumption without sacrificing accuracy by speeding up the slowest parts of the full simulation. By treating (as described below) electromagnetic showers in the sampling portions of the calorimeters, a reduction in CPU time of a factor of three can be achieved even in hadronic events. Although the calorimetry dominates simulation time for the full simulation, after treatment of the electromagnetic showers the simulation time is evenly distributed throughout all subdetectors. One particular advantage of this fast simulation over the other varieties is that its output file matches identically the format of the output of the full GEANT4 simulation. The data can therefore be run through the identical tests and digitization software following simulation, and the standard ATLAS trigger and reconstruction can be run.

There are three treatments applied to electromagnetic showers. For very high energy ($>10$ GeV) electrons and positrons, a tuned shower parameterization is available. For medium energy (10 MeV to 1 GeV) electrons, positrons, and photons, libraries of pre-simulated showers can be applied during the event. For very low energy ($<10$ MeV) electrons and positrons, a single hit can be deposited to recreate detector response. Each one of these treatments can be turned on by the user in each compartment of the electromagnetic calorimeter and the forward hadronic calorimeter. The energy ranges can be set for each method, compartment, and particle.

For electrons and positrons above a sufficiently high energy, around 10 GeV in the central calorimeters, the sampling calorimeter is sufficiently homogeneous to apply a shower parameterization. Small steps are taken in the direction of the original particle, depositing energy according to several tuned functions as it traverses the detector. The longitudinal profiles of showers are parameterized and normalized with an energy scale to approximate the sampling fraction in each subdetector. The radial profile changes as a function of depth in the shower and is normalized by the longitudinal profile. Energy is deposited in hits in order to mimic the full simulation. Fluctuations are introduced in three separate places, representing the random characteristics of shower length and shape, the sampling resolution of the calorimeter, and the geometric fluctuations in the energy collected.

Particles captured by the fast simulation in the appropriate energy range (typically $<1$ GeV) are replaced by a shower from a pre-simulated library, rotated and scaled to match the primary particle. Shower libraries are generated in bins of pseudorapidity and energy for electrons and photons. Only hits in the sensitive detectors are stored in order to save space on disk and in memory. The binning reproduces the fine structure in the calorimeters. The libraries are read into memory the first time they are requested by the simulation, ensuring minimal memory overhead. They consume about 200 MB when all are in use. Showers are randomly selected with linear weighting from the energy bin above or below the primary particle and from the pseudorapidity bin above or below the particle. The shower is then rotated to
match the primary particle’s original direction, and the energy is scaled to match the primary particle’s energy. For example, an electron at pseudorapidity of 2.37 and with an energy of 12 MeV might use a shower from the electron and positron library’s 2.4 bin in pseudorapidity and 10 MeV bin in energy. The shower would then be moved and rotated to match the position and direction of the original particle, and its energy would be scaled up by 20%.

Electrons and positrons with energies below about 10 MeV typically deposit only one hit in the sensitive region of the calorimeter. These particles are removed when inside the regular sampling region of any of the calorimeters (“killed”), and a single hit is placed in the calorimeter. The position of the hit is determined by a random exponential number times the radiation length in the detector in order to approximate the particle’s range. The energy of the hit is scaled to the response of the detector and smeared by its resolution.

The standard combination of strategies is shown in Table 14. The strategies used in a particular subdetector is optimized for maximum CPU time improvement with minimal complexity. The upper energy bound for shower libraries, 1 GeV, balances memory use with speed. Libraries at higher energies also may not correctly reproduce the tails of electromagnetic shower shape distributions as well as low-energy libraries do. The minimum energy for application of the parameterization model is based purely on CPU time. In most subdetectors, it is faster to allow GEANT4 to produce 1 GeV secondaries and apply shower libraries to those secondaries than it is to apply the shower parameterization. The same argument applies to high energy photons. They pair produce sufficiently quickly that treating them separately only adds complexity to the models. The speed of the parameterization is limited by random number generation and locating hits within the detector geometry.

### 7.2 ATLFAST-I

ATLFAST-I performs a fast simulation of the ATLAS detector, including object reconstruction, in order to produce high statistics samples of signal and background events. The lowest possible CPU time per event is achieved by replacing detailed detector simulation with parameterizations of the desired detector and reconstruction effects. The high speed of simulation in ATLFAST-I makes it possible to study channels where the statistics involved would otherwise be prohibitive. For example, the background to a $Z \rightarrow \ell^+\ell^-$ study from fake taus in di-jet events is expected to require $O(10^9)$ events in 100 fb$^{-1}$ of data. Some searches also require many datasets to be simulated in order to scan across parameter space for the model being tested, such as SUSY.

ATLFAST-I is the least detailed simulation method. There is no realistic detector description, so studies of detector-based quantities, such as calorimeter sampling energies and track hit positions, are not possible. There is no simulation of reconstruction efficiency or misidentification rates, discussed later on, which means the presence of genuine physics objects are overestimated while fake objects are not modeled, with two exceptions. Because jet-flavor tagging efficiencies are applied, fake b-jets and taus are simulated. However, ATLFAST-I provides a useful method of making quick estimates of systematic uncertainties in early data analyses due to the simple process of re-parameterizing the detector and modeling reconstruction effects. The speed of operation enables datasets to be reproduced with different generator configurations, allowing quick estimates of systematic uncertainties arising from generators.

Common to the reconstruction of all objects in ATLFAST-I is that by default no reconstruction efficiencies are applied. These efficiencies can be taken from full simulation and accounted for by the user in the analysis. This applies to electrons, photons and jets as well as to ATLFAST-I tracks. It should be noted that tagging efficiency factors are implicitly taken into account in the tau- and b-tagging procedures. A system to apply a common set of efficiencies and misidentification rates at the analysis stage is in development. The misidentification rates will allow the modeling of fake objects as well.

ATLFAST-I takes input in HepMC format, enabling it to read the output of all ATLAS generators. Generator input is filtered to choose only particles that are useful in the current step. For example, only charged particles are considered in the tracking stage, and all particles are required to be a part of the final state.
The following sections describe steps taken in ATLFAST-I.

7.2.1 Tracks

Charged particle tracks from the generator with $p_T > 500$ MeV and with $|\eta| < 2.5$ are considered as reconstructed ATLFAST-I tracks, and five track parameters are associated to them. These parameters are calculated from the true particle properties by applying parametrized resolution functions which account for the measurement precision, energy loss, and multiple scattering as well as for hadronic interactions in the inner detector material. The resolution functions are taken from fully simulated events. The non-Gaussian tails resulting from hadronic interactions are taken into account by applying a double-Gaussian correlated smearing to the track parameters of hadrons. No vertex smearing is applied. In ATLFAST-I, three types of charged particles are distinguished: hadrons, electrons and muons. Due to the relatively large energy loss from bremsstrahlung, high-$p_T$ electrons are treated separately, and an additional energy loss correction is applied. It should be noted that while these tracks are used for specific studies in B physics, they are not used for lepton identification or b-tagging.

7.2.2 Track-based tau identification

Track-based tau identification is split into two distinct parts, namely reconstruction and identification of tau candidates. The reconstruction part applies a parameterized efficiency to the tracks to calculate the charged component in a tau candidate, while the neutral component is calculated directly from neutral particles in the generated event.

Once a sample of tau candidates has been reconstructed, the identification part is carried out by separating the sample into true and fake taus. True taus are defined as those matched to a hadronic decay in the truth record with $\Delta R \equiv \sqrt{\Delta \eta^2 + \Delta \phi^2} \leq 0.2$, whereas the remainder are considered fakes. Subsequently, a parametrization of the identification efficiency is applied based on the number of tracks.

7.2.3 Calorimetry

Stable charged particles from the event generators are propagated through the magnetic field along a simple helix. The primary vertex is assumed to be at the geometric origin. Using a helix model and assuming a perfectly homogeneous magnetic field inside the central tracking volume, the impact point on the calorimeter surface is calculated. To calculate this point, no interactions of the particle with the detector material (i.e. no multiple scattering, energy loss, or nuclear interactions) are taken into account. In particular, this implies that no energy loss due to bremsstrahlung for electrons and no pair production from photons in the inner detector media are simulated for the energy depositions in the calorimeters. The effects of these interactions are, however, implicitly taken into account by the application of appropriate resolution functions. For the calculation of track parameters, the four-momenta and the starting point of the particles (e.g. for stable decay products of long-lived particles) are taken from the generator information.

The energies of the electrons, photons, and hadrons are deposited in a calorimeter cell map. The response of the calorimeter is assumed to be unity and uniform over the full detector. No smearing (i.e. no resolution function) is applied. The energy of the particle is entirely deposited in the hit calorimeter cell, assuming a granularity of the calorimeter cell map of $\eta \times \phi = 0.1 \times 0.1$ up to $|\eta| < 3.2$ and $\eta \times \phi = 0.2 \times 0.2$ for $3.2 < |\eta| < 5.0$. Neither lateral nor longitudinal shower development is simulated. Therefore, the longitudinal fine structure of the calorimeters is not taken into account. There is also no separation between the electromagnetic and the hadronic calorimeter compartments.

Based on the map of deposited cell energies, cluster reconstruction is carried out using either SISCone or FastKt algorithms via the FastJet libraries. The default clustering routine is SISCone with a cone size of 0.4. The cluster transverse energy must pass a threshold, typically 5 GeV. The clusters may get re-classified as electrons, photons, taus or jets in one of the following steps. If they are associated to one of these objects, they are removed from the list of clusters.

7.2.4 Electrons and photons

For each true electron or photon, the reconstructed energy is obtained by smearing the true energy according to a resolution calculated by interpolating between resolutions measured in fully simulated events at precise values of $\eta$ and energy. If, after smearing, the candidate electron or photon has transverse energy exceeding a threshold value, typically 5 GeV, and has $|\eta| < 2.5$, then it is recorded with the $\eta$ and $\phi$ directions of the true particle.

Electrons and photons are matched to calorimeter clusters in $(\eta, \phi)$ space, with a maximum allowed separation of $\Delta K = 0.15$. If there is a matching cluster then it is removed from the list of clusters to be considered as jet candidates later on.
7.2.5 Muons

For each true muon with $p_T > 0.5$ GeV, the reconstructed momentum is calculated from the true muon momentum. A Gaussian resolution function which depends on $p_T$, $\eta$, and $\phi$ is applied. After smearing, muons with $p_T > 5$ GeV and with $|\eta| < 2.5$ are kept.

7.2.6 Isolation

In order to define isolated electrons, photons, and muons, the following criteria are applied: the difference in the unsmeared energy in a cone of $\Delta R = 0.2$ around the object direction and its smeared energy needs to be below 10 GeV. In addition, there should be no further clusters reconstructed with $\Delta R < 0.4$ around the object direction.

7.2.7 Jets

All clusters that have not been assigned to a true electron or photon are considered jets if their transverse energy exceeds 10 GeV. The jet energy is taken to be the cluster energy, after adding non-isolated muons within $\Delta R = 0.4$, and is smeared according to the jet energy resolution.

These functions do not account for pile-up, although there is a “high luminosity” mode available which adds a pile-up term to the resolution. The pile-up correction is constant with respect to jet transverse energy and is dependent on the size of the jet.

The jet direction is taken to be the cluster direction. Since the response function of the calorimeter is set to one, no jet calibration is needed to correct for the non-compensation of the calorimeter. However, an out-of-cone energy correction is needed. This correction is applied in a separate jet calibration step [110, 111].

7.2.8 Tagging

For each jet found, a label is attached to indicate whether the true jet originated from a light quark, $b$-quark, $c$-quark, or tau. This label is based on matching $b$ or $c$ partons or the visible decay products of hadronically-decaying taus at truth-level with $\Delta R < 0.3$ to a reconstructed jet. In the case of hadronically-decaying taus, the ratio between the true visible energy and the jet energy is also required to be larger than unity minus $2\sigma$, where $\sigma$ is the jet energy resolution as above.

The results of $b$- and tau-tagging are then simulated by applying identification efficiencies and fake rates to the labels. These efficiencies are determined from full simulation studies and are parameterized as a function of $p_T$ and $\eta$.

7.2.9 Missing $E_T$

The missing transverse energy is calculated from all reconstructed objects: isolated electrons, photons, muons, taus, jets and non-isolated muons, and remaining calorimeter clusters not associated to jets. In addition, cells not associated to clusters are included in the missing $E_T$ calculation. The cell energies are smeared by applying the jet resolution functions.

7.3 ATLFAST-II

ATLFAST-II directly simulates the input to the standard Athena reconstruction algorithms to mimic the full simulation. Unlike ATLFAST-I, which provides only momenta for the reconstructed objects, reconstructed ATLFAST-II output includes all the properties associated with a reconstructed object. In the case of Fatras these include the hits in the inner detector and muon system, and for FastCaloSim these include the energies in the calorimeter cells. Because the standard reconstruction is run, it is possible to work with a combination of full and ATLFAST-II simulated events without modifying any analysis code. Both Fatras and FastCaloSim run together with the event reconstruction. The simulation time is reduced by making use of the simplified detector description used for reconstruction [115]. By default, ATLFAST-II uses full simulation for the inner detector and muon system and FastCaloSim in the calorimeter. ATLFAST-IIIF uses FastCaloSim in the calorimetry and Fatras in the inner detector and muon system.

As input, Fatras uses input in HepMC format, performs a smearing of the primary vertex position to represent the luminous region within ATLAS, and records truth information in a way similar to the full simulation. FastCaloSim uses the truth information of all interacting particles at the end of the inner detector volume as input to the calorimeter simulation. In order to simulate pile-up, generated events must be overlaid prior to detector simulation.

7.3.1 Fatras

ATLFAST-II with the fast track simulation engine Fatras (ATLFAST-IIIF) reduces simulation time in the inner detector and muon system. Fatras is an ATLAS specific development and establishes a complete simulation within the track reconstruction framework. The reconstruction geometry is a simplified description of the full detector geometry, which keeps the same descriptive accuracy for sensitive detector parts, while approximating all other detector components as simplified layers that carry a high-granularity density map. This detector material description can be sufficient. A factor of 100 reduction in CPU time is obtained with only small physics performance degradation. The propagation of the
particles through the tracking detectors is carried out by the extrapolation engine [116] used in the offline track reconstruction applications.

The interactions of the particles with the simplified detector layers are simulated using several methods. Multiple Coulomb scattering is implemented as a Gaussian mixture model to account for tail effects from single large-angle scattering processes; ionization and radiative energy loss are simulated according to the Bethe-Bloch and Bethe-Heitler models; conversion of a photon into an electron and positron is carried out depending on the thickness of the traversed material; hadronic interactions of particles with the detector layers are simulated from a parametric model that has been obtained from GEANT4 simulation results. The decay of unstable particles is enhanced by a dedicated wrapper of the associated GEANT4 modules [112]. The calorimeter simulation of ATLFAST-IIIF is typically FastCaloSim, and Fatras provides the input particle collection. Energy deposition for muons in the calorimeter layers is also recorded according to the material description of the reconstruction geometry and is further used for cluster simulation in the FastCaloSim application.

Fatras was first established as a validation tool for the newly deployed inner detector reconstruction sequence. It has already been used for noise studies in the Transition Radiation Tracker and first simulations for a potential future upgrade of the ATLAS inner detector. The validation of Fatras against the full simulation results to be used for first collisions data from LHC is ongoing. An extension of the fast track simulation within the reconstruction geometry has taken place that also allows the use of Fatras in the muon spectrometer. The particles being simulated at the end of the inner detector volume are filtered. Muons are transported through the calorimeter, and their deposited energy is stored as an input to the FastCaloSim module. The trajectories of the muons are then simulated in the muon spectrometer, and the hits within sensitive detector elements are recorded. Standard digitization is applied on top of the simulated hits to account for the detailed calibration that must be included for a comparison to data.

### 7.3.2 FastCaloSim

Instead of simulating the particle interactions with the detector material, the energy of single particle showers is deposited by FastCaloSim directly using parametrizations of the longitudinal and lateral energy profile. The distribution of active and inactive material in the calorimeter needs to be respected by the parametrization, so a fine binning of the parametrization in the particle energy and pseudorapidity is needed. Furthermore, the energy deposition depends strongly on the origin of the shower in the calorimeter, so all parametrizations are also binned versus the longitudinal depth of the shower center.

The parametrizations are based on a 30 million event sample of fully-simulated (i.e. simulated with GEANT4) single photons and charged pions in an energy range between 200 MeV and 500 GeV, evenly distributed in $|\eta| < 5.0$ and $-\pi < \phi < \pi$. All electron and photon showers are approximated by the photon parametrization and all hadronic showers are approximated by the charged pion parametrization. The simplified reconstruction geometry of the calorimeter is used with details at the level of the readout cells.

The parameterization of the longitudinal energy distribution is constructed from histograms of the total energy in all calorimeter layers, the longitudinal depth of the shower center, and the energy fraction in each layer for the fully-simulated single-particle events. The dominant correlations between fractional energy deposits in each calorimeter layer (i.e. those related to the longitudinal depth of the shower’s origin) are accounted for in the parameterization binning. Gaussian correlations between fractional energy deposits in each calorimeter layer (i.e. those describing shower development) are stored in a correlation matrix and are applied to improve the parameterized energy distribution. During fast simulation, the parametrization closest in energy and pseudorapidity to the particle is taken, and then the total shower energy and the shower depth are chosen randomly from the stored histograms and rescaled to match the true particle energy. It was found that after rescaling no interpolation between parametrizations is necessary. Afterwards, the energy fractions in all calorimeter layers are generated randomly, taking into account the correlation matrix. The lateral energy distribution inside each calorimeter layer is simulated using a symmetric average radial shape function. The shape functions are extracted from fits to fully-simulated single-particle events and are constructed for bins of particle type, primary particle energy, position in $\eta$, and shower depth in the calorimeter. The asymmetry of shower shapes for particles entering the calorimeter at large incident angles is absorbed in a shape function describing a pseudorapidity-dependent asymmetry term. During simulation, the energy of a calorimeter cell is determined by the integral of the shape function over the cell surface area. Fluctuations derived from the intrinsic resolutions of each calorimeter are applied to the cell energy. The total energy of all cells in one calorimeter layer is normalized to the total energy in the layer making use of the longitudinal shower shape.

The histograms and shape functions needed as input for the parametrizations use about 200 MB of memory. Since no simulation of particle interactions is done, the dominant part of the simulation time is spent on the numerical integration of the lateral shape functions. Overall, the calorimeter simulation time for a single particle is a few microseconds, and a typical (e.g. $t\bar{t}$) event needs a few seconds.

The parameterization of FastCaloSim differs in several important ways from that of the Fast G4 Simulation.
For the same system, the peak SpeCmark floating point rating was 1794. The normalization was done using the peak SpeCmark integer rating, 1794.

15Measurements were performed on Sun Fire X2200 M2 units with dual dual-core 2.6 GHz AMD Opteron 2218 processors. All times assume the jobs run on a single core, with the other cores loaded. Normalization was done using the peak SpeCmark int 2000 rating, 1794. For the same system, the peak SpeCmark floating point 2000 rating was 3338. The normalization follows the published results, rather than the

FastCaloSim fills the readout geometry of ATLAS and applies a parameterization from the edge of the inner detector, whereas the Fast G4 Simulation places hits like those of GEANT4 into the full ATLAS detector geometry and is only applied in the sampling portion of the calorimeter (e.g. excluding the cryostats surrounding the calorimeter). As a result, the Fast G4 Simulation output can be run through the standard digitization software, whereas the FastCaloSim output is fed directly into the reconstruction.

### 7.4 Computing performance

Examples of simulation times in kSi2K seconds [117] for various types of events in the full and fast simulations are provided in Tables 15 and 16. In single central (|η| < 3) electron events the simulation time is decreased by a factor of ten or more by the fast G4 simulation, and in hard scattering events the simulation time is decreased by a factor of 2–5. ATLFAST-II without Fatras decreases simulation time by a factor of 20–40, and ATLFAST-II decreases simulation time by a factor of 100. FastCaloSim accounts for about 10% of the total simulation time in ATLFAST-II and 60–70% of the total simulation time in ATLFAST-IIF. ATLFAST-I requires a relatively negligible amount of CPU time even for hard scattering events. ATLFAST-I, FastCaloSim, and Fatras run during the reconstruction step, the amount of time taken by FastCaloSim could not be measured in that sample. It was estimated as 10% of the full inner detector simulation time, consistent with the other hard scattering events.

---

#### Table 15 Simulation times per event, in kSI2K seconds, for single particles generated with $|\eta| < 3.0$ and with the same transverse momentum. All times are averaged over 500 events. ATLFAST-II uses full simulation for the inner detector and muon system and FastCaloSim in the calorimetry. ATLFAST-IIF uses FastCaloSim in the calorimetry and Fatras in the inner detector and muon system

| Sample               | Full Sim | Fast G4 Sim | ATLFAST-II | ATLFAST-IIF | ATLFAST-I |
|----------------------|----------|-------------|------------|-------------|-----------|
| 5 GeV $\mu$±        | 0.879    | 0.899       | 1.28       | 0.633       | 0.011     |
| 50 GeV $\mu$±       | 1.63     | 1.15        | 2.71       | 0.606       | 0.011     |
| 500 GeV $\mu$±      | 12.0     | 10.4        | 11.8       | 0.615       | 0.011     |
| 1 GeV $e$±          | 3.62     | 0.734       | 0.825      | 0.513       | 0.011     |
| 5 GeV $e$±          | 17.8     | 1.64        | 1.00       | 0.542       | 0.011     |
| 50 GeV $e$±         | 179      | 4.86        | 1.25       | 0.588       | 0.013     |
| 1 GeV $\pi$±        | 2.40     | 1.48        | 0.701      | 0.515       | 0.011     |
| 5 GeV $\pi$±        | 10.4     | 4.27        | 0.811      | 0.540       | 0.011     |
| 50 GeV $\pi$±       | 94.7     | 30.3        | 1.04       | 0.569       | 0.011     |

#### Table 16 Simulation times per event, in kSI2K seconds, for the full simulation, Fast G4 simulation, ATLFAST-II, ATLFAST-IIF, and ATLFAST-I. ATLFAST-II uses full simulation for the inner detector and muon system and FastCaloSim in the calorimetry. ATLFAST-IIF uses FastCaloSim in the calorimetry and Fatras in the inner detector and muon system. All times are averaged over 250 events, except heavy ion times which were averaged over only 50 events. Because the memory required to reconstruct heavy ion events exceeds 3 GB and because FastCaloSim runs during the reconstruction step, the amount of time taken by FastCaloSim could not be measured in that sample. It was estimated as 10% of the full inner detector simulation time, consistent with the other hard scattering events.

| Sample               | Full Sim | Fast G4 Sim | ATLFAST-II | ATLFAST-IIF | ATLFAST-I |
|----------------------|----------|-------------|------------|-------------|-----------|
| Minimum bias         | 551      | 246         | 31.2       | 2.13        | 0.029     |
| $t\bar{t}$           | 1990     | 757         | 101        | 7.41        | 0.097     |
| Jets                 | 2640     | 832         | 93.6       | 7.68        | 0.084     |
| Photon and jets      | 2850     | 639         | 71.4       | 5.67        | 0.063     |
| $W^\pm \rightarrow e^\pm \nu_e$ | 1150 | 447        | 57.0       | 4.09        | 0.050     |
| $W^\pm \rightarrow \mu^\pm \nu_\mu$ | 1030 | 438        | 55.1       | 4.13        | 0.047     |
| Heavy ion            | 56,000   | 21,700      | ~3050      | 203         | 5.56      |
the data must be digitized and reconstructed. For ATLFAST-II, the inner detector and muon system must be digitized and reconstructed, but the calorimeter requires only reconstruction. For ATLFAST-IIF, only the muon system must be digitized before reconstruction is performed. The output of ATLFAST-I is in a format similar to that of the reconstruction and needs no further processing. The CPU time required for these additional steps is discussed in Sect. 8.2.

7.5 Physics performance

The fast simulations have been compared to full simulation in both low-level analyses with single particles entering the calorimeter and high-level analyses of detector observables with jets and active hard scattering events. The Fast G4 Simulation agrees to about 1–2% in jet energy scale after the standard calibration procedure and agrees to within 5% percent in electron identification efficiencies. Due to the simplifications in the calorimeter simulation, FastCaloSim differs at the 5% level from full simulation after reconstruction, especially in properties that are sensitive to the shape of hadronic showers. The jet energy scale differs by 1–2% after recalibration, and electron identification efficiency differs by about 5%. Since all particles are simulated using an average lateral shape function, visible effects like electromagnetic subshowers in charged pion showers are not described. These differences can be reduced by applying additional object-dependent correction functions after reconstruction. Fakes and calorimeter punch-through are not well modeled in ATLFAST-II and ATLFAST-IIF.

Figure 9 shows missing transverse energy along the x-axis for the full and fast simulations in di-jet events with a leading parton $p_T$ between 560 and 1120 GeV, as well as jet $p_T$ resolution as a function of $\eta$ in $t\bar{t}$ events for jets with $20 < p_T^{\text{true}} < 40$ GeV. ATLFAST-II and the Fast G4 Simulation agree well with full simulation in missing transverse energy spectrum, even in the tails of the distribution. ATLFAST-I does not sufficiently populate the tails of the missing transverse energy distribution, and ATLFAST-IIF has too wide a distribution. ATLFAST-I, ATLFAST-IIF, and ATLFAST-II show 10–20% deviations from full simulation in jet transverse momentum resolution. These fast simulations mostly deal with the response of individual hadrons, which is not always sufficient to model the threshold effects common in jet and missing transverse energy measurements. Fast G4 simulation is consistent with full simulation through the entire range in pseudorapidity.

Figure 10 shows reconstructed muon $p_T$ resolution as a function of muon $p_T$ in $Z \rightarrow \mu^+\mu^-$ events. Muons reconstructed using the muon spectrometer alone and those reconstructed using both the muon spectrometer (“standalone”) and inner detector (“combined”) are shown. Only one type of muon is provided by ATLFAST-I, so it is only included in the combined reconstruction plot. The difference between full simulation and ATLFAST-I is due to an older parameterization. In the cases of ATLFAST-II and the Fast G4 simulation, muon spectrometer simulation is done by GEANT4 and should, therefore, be identical to full simulation. The fast simulations show generally good agreement over the entire range of $p_T$. ATLFAST-IIF has standalone muon resolution that is 10% better than full simulation in some bins of $p_T$, but since the muon system simulation of ATLFAST-IIF is still under development, the agreement is expected to improve. It is generally left to the physics groups to evaluate the fast simulations with their analyses and determine which is acceptable.

8 Validation

Validation of the ATLAS simulation chain is done in two distinct phases. First, the software performance must be assessed. Then, the physics performance must be tested and
Fig. 9 *Left*, fast simulations (color) and full simulation (black) comparison of missing transverse energy along the x-axis in di-jet events with a leading parton $p_T$ between 560 and 1120 GeV. *Right*, a comparison of jet $p_T$ resolution as a function of pseudo rapidity in $t\bar{t}$ events for jets with $20 < p_T^{\text{True}} < 40$ GeV.

Fig. 10 Fast simulations (color) and full simulation (black) comparison of reconstructed muon $p_T$ resolution as a function of muon $p_T$ for central ($|\eta| < 1.2$) muons in $Z \rightarrow \mu^+ \mu^-$ events for muons reconstructed using only the muon spectrometer (*left*) and using both the inner detector and spectrometer (*right*). ATLFAST-I only provides one type of muon, which is included in the right plot.

Compared to available data. The first step includes testing robustness, testing software performance, and testing basic functionality. The second step includes comparison to test beam, cosmic data, and physics results obtained from previous simulation productions. In this section the infrastructure for each stage of validation is described.

ATLAS has a fresh software build every night for coordinating software development. Each nightly build is run through a rigorous test cycle, and as a release deadline approaches the test results are increasingly scrutinized to evaluate stability and performance. Thanks to the evaluation prior to release, generally only rare bugs appear in pro-
duction for the first time. The automatic testing infrastructure also allows evaluation of many different versions of the Athena software. Separate bug-fixing and development branches are employed, for example, and significant interface changes or low-level code migrations take place in separate branches until they are sufficiently stable to be merged into the main branch. Each version of the software comes in several flavors for different system architectures, operating systems, compilers, and so on. The simulation production in 2008 used 32-bit builds with gcc 3.4.6 [120] on CERN’s Scientific Linux 4 [121]. External dependencies include CLHEP 1.9.3.1 and WLCG 54G.

A web portal has been constructed, using the Savannah bug tracking software [122], for monitoring problems with the various facets of the simulation software. Bugs can be reported and tracked as they are diagnosed and solutions are found, and new features can be requested.

8.1 Automated testing

The software performance of the simulation is monitored in three types of automated tests: ATLAS Testing Nightly (ATN) tests, Run Time Tester (RTT) tests, and Full Chain Tests (FCT). ATN tests are run every night on every software build and are basic functionality tests. RTT tests are run on a subset of builds and include 50 simulation tests to ensure functionality and, in some cases, consistent results. FCT tests are run on only a few builds each night and test the entire software chain in a production-like environment. Releases are required to pass a minimal number of milestones before being declared ready for production. For details about the ATN and RTT, refer to [9].

FCT tests are run daily on a small set of jobs. The aim of the FCT is to verify the readiness of a production cache release candidate for Grid production. The FCT runs jobs that test the functionalities of generation, the different flavors of fast and full simulation, digitization, bytestream conversion, and reconstruction of Standard Model processes, black hole production, and heavy ion collisions. In the case of Standard Model processes, a full chain of jobs are run per release, with hard scattering events that stress the software. If successful, the output from each day’s run is saved for use in the next step of the test the following day (e.g. Monday’s generation provides input for Tuesday’s simulation, which provides input for Wednesday’s digitization). The typical number of events processed (50) is limited by the CPU requirements for the full simulation.

As part of the FCT, 1000 events that were simulated with an old validated release are reconstructed. This long test allows better evaluation of the reconstruction’s stability. Moreover, the relatively large sample is used to make a preliminary check on the quality of the reconstruction for final state objects (jets, electrons, muons, etc.). All the other tests only check for the success or failure of the job, the number of events in the output file, and unknown error messages in the log file. If any of these checks fail, the release candidate is rejected, and an additional iteration of bug fixing is undertaken. Only once a release succeeds in all FCT tests is it distributed to the Grid.

8.2 Computing performance benchmarking

Event generation jobs are typically fast enough that not a large effort is made to test their software performance. In general, the jobs take a tens of milliseconds per event. Generation with PYTHIA or HERWIG requires about 450 MB of memory, and generation with HiJing requires about 170 MB of memory. The files produced in generation jobs are tens of kB per event (e.g. 40 kB/event for tðt events).

CPU time, memory consumption, and output file size for the simulation are tested in each stable release using a variety of physics processes. Single muons, electrons, and charged pions are used, as well as di-jets in bins of leading parton pt, a supersymmetric benchmark point,15 minimum bias, Higgs boson decaying to four leptons, Z → e+e−, Z → μ+μ−, and Z → τ+τ− events. The same input events are used every time to ensure fair comparison of simulation results independent of changes to the event generation.

Simulation of the full detector requires typically ~750 MB of memory (i.e. VSIZE) and includes loading of almost 400 libraries into memory. Memory consumed during simulation is also broken down into its three key components by taking snapshots of memory use during initialization: GeoModel, the ATLAS-side detector geometry, typically about ~100 MB; G4Atlas, the purely GEANT4 component of the memory, typically about ~300 MB; and load modules, the remaining algorithms and services loaded during the job, typically about ~300 MB. Significant changes in any of these can indicate the proper source of a change in memory. The memory requirement is independent of the number of events in the job and varies by only a few percent for different physics processes. Although up to 2 GB of memory may be reserved for a Grid job, by keeping the memory requirements of a typical simulation job under 1 GB, more machines can be used. The memory required to build each piece of the detector can be found in Table 4. Of major concern is any increase in memory (“leaks”) during the event loop once all libraries have been loaded and setup is complete. Some increase due to caching is expected

15 A single chain of jobs runs all steps from event generation through reconstruction, sequentially, using the output of one step as the input of the next.

18 ATLAS mSUGRA benchmark point SU3: m0 = 100 GeV, m1/2 = 300 GeV, A0 = −300, tan β = 6, and μ > 0.
during the processing of the first few events. However, if the memory required by the application continues to grow beyond the system limits, memory corruption and memory pressure can result in serious problems. The memory required by the ATLAS simulation has been found to increase by less than 0.25 MB per event under normal circumstances. The increases are not steady, but come in large (∼10 MB) and sporadic jumps. The source of these increases is not fully understood, but a 50 event simulation job still consumes well under 1 GB of memory.

The CPU time consumed by generation, full simulation, digitization, and reconstruction for various types of events is shown in Tables 15 and 17 and is typically several minutes per hard scattering event. All times are normalized to kSi2K seconds [117]. For the purposes of testing, logfile output was suppressed and no output files (e.g. hit or RDO files) were created. CPU time is also measured as a function of other simulation input parameters prior to significant changes, for example using different physics lists. For these runs, output files were disabled; in simulating $t\bar{t}$ events the time per event is increased by ∼0.5% when file writing is enabled. The hard scattering events shown in Table 17 were generated with a 14 TeV center of mass energy; for 10 TeV center of mass energy the simulation time is reduced by 17% for $t\bar{t}$ events. The distribution of CPU time for simulation of 250 $t\bar{t}$ events is shown in Fig. 8.

For the samples in Tables 16 and 17, event generation of W production, minimum bias interactions, di-jet events, and photon and jet events was done using PYTHIA. $t\bar{t}$ production was done using MC@NLO for the hard scattering and Herwig for hadronization and showering. In this case, the generation time includes only the time required by Herwig. Heavy ion event production was done using Hijing.

Digitization jobs are generally fast, but memory consumption can be a serious concern during jobs with many overlaid events. Table 18 shows how resource consumption can be a serious concern during jobs with many overlaid events. Table 18 shows how resource consumption scales with pile-up luminosity. The memory required for digitizing with a luminosity of $10^{34}$ cm$^{-2}$ s$^{-1}$ is sufficiently large that the memory limit of the testing machine was reached, and, therefore, swapping resulted in a significant increase in CPU time. The allocated memory per event is provided for some benchmark of the change in memory over the course of a single event.

### Table 17

| Sample          | Generation | Simulation | Digitization | Reconstruction |
|-----------------|------------|------------|--------------|----------------|
| Minimum bias    | 0.0267     | 551        | 19.6         | 8.06           |
| $t\bar{t}$ Production | 0.226      | 1990       | 29.1         | 47.4           |
| Jets            | 0.0457     | 2640       | 29.2         | 78.4           |
| Photon and jets | 0.0431     | 2850       | 25.3         | 44.7           |
| $W^\pm \rightarrow e^\pm \nu_e$ | 0.0788     | 1150       | 23.5         | 8.07           |
| $W^\pm \rightarrow \mu^\pm \nu_\mu$ | 0.0768     | 1030       | 23.1         | 13.6           |
| Heavy ion       | 2.08       | 56,000     | 267          | –              |

### Table 18

Digitization computing resources for 50 $t\bar{t}$ events as they scale with luminosity. CPU times are normalized to the time required by a no pile-up job. Cavern background was overlaid during these jobs with a safety factor of one. Beam gas and beam halo were ignored.

| Resource                        | No Pile-up | $10^{33}$ cm$^{-2}$ s$^{-1}$ | $3.5 \times 10^{33}$ cm$^{-2}$ s$^{-1}$ | $10^{34}$ cm$^{-2}$ s$^{-1}$ |
|---------------------------------|------------|------------------------------|----------------------------------------|------------------------------|
| CPU time factor                 | 1.0        | 2.3                          | 5.8                                    | 160                          |
| Memory leak [kB/event]          | 10         | 270                          | 800                                    | 2,100                        |
| Virtual memory [MB]             | 770        | 1,000                        | 1,300                                  | 2,000                        |
| Allocated memory [MB/event]     | 12         | 21                           | 40                                     | 985                          |

8.3 Physics validation

Once a new release is distributed to the Grid sites, a set of several physics samples is produced. Typically, a “validation”...
tion sample” includes 10,000 events for each process, a total of 110,000 single particle events and 250,000 hard scattering events. This standard validation sample includes single muons, pions, and electrons, Standard Model processes ($t\bar{t}$ production, vector boson production, B-physics), and exotic processes (e.g. supersymmetric events and black hole production). The composition of the validation sample has been chosen to test all aspects of the event reconstruction.

The running of the validation sample on the Grid usually exposes rare software problems in the release. It is unlikely that software bugs that appear with a frequency much lower than 1/1000 events are caught by the automatic validation procedure (1000 events is the size of the “long” jobs of the FCT, described in Sect. 8.1). This first round of production provides a feedback mechanism for the developers, who produce bug fixes before the next production cycle.

The last step before using a release for production is physics validation. A dedicated group of experts, including representatives from every detector performance (e.g. tracking, b-tagging, and jet reconstruction groups) and physics group (e.g. Standard Model, supersymmetry, and exotics search groups) in ATLAS, runs physics analyses on the validation samples. Their task is to verify the quality of the single object reconstruction (e.g. jets, electrons, and muons) and the results of more complex physics analyses (e.g. mass reconstruction in $Z \rightarrow \mu^+\mu^-$, $Z \rightarrow e^+e^-$, and $t\bar{t}$ events). The relatively large validation samples may expose minor problems that could not be found with lower statistics, for example a shift of a few percent in the reconstructed energy. In order to properly validate each version of the software, the results from each release are typically compared to those of previous validated releases. The software must, therefore, maintain backwards-compatibility in order to allow fair comparisons. Shifts in file format are carefully coordinated, and maintenance of the old format is continued for as long as necessary to ensure result consistency. The physics validation procedure is also used for checking major changes in the fast and full simulation (detector description, change in the simulation parameters, etc.).

The GEANT4 simulation has also been validated in a physics sense with all available detector data. Combined test beam studies have proven invaluable in understanding the performance of each of the subsystems, and the standalone test beam analyses have provided crucial input towards the optimization of the simulation and choice of parameters [123–126]. In 2008, a significant sample of cosmic ray data was collected with multiple subdetectors. The data have provided an important test of the simulation [127].

Although the detector simulation relies heavily on GEANT4, a significant effort was put into comparing tile calorimeter test stand response with the FLUKA simulation package [128]. For this comparison, the test stand geometry was translated into the FLUKA geometry format, and the output from the FLUKA simulation was translated back into a format comparable to that of GEANT4. It was eventually concluded that little would be gained by attempting a transition to FLUKA that could not already by gained by modifications to parameters and a different choice of physics models within GEANT4. FLUKA has also been used to study neutron flux and radiation levels throughout the detector [104], but many of these studies are being updated in GEANT4 with the high-precision neutron physics list (see Sect. 5.4).

Extensive efforts are underway to compare simulated data to real data and validate the output from each detector. Thanks to the multiple detector descriptions, several analyses have already been prepared and tested to find discrepancies between the detector description of the simulation and that of the as-built detector. For example, subdetectors can be “weighed” in the simulation to ensure that the amount of material is within a few percent of the constructed detector. Although the agreement with first high-energy collision data is not expected to be perfect, a great deal of experience has been gained. The effects of modifications to GEANT4 parameters have also been studied in some detail, so that differences between data and simulation might be remedied rapidly.

Digitization algorithms have been tuned against laboratory test results, test beam data, and, where possible, cosmic ray data taken during the detector commissioning. The studies continue with the data.

9 Summary and conclusions

We have presented the status of the ATLAS simulation project, including all steps from event generation to digitization. A robust and flexible framework is required to cope with the demands of complex detector descriptions and physics models. The software project has been prepared for data since late 2008 and is ready for data.

A variety of event generators are available to provide the user with a complete set of tools for testing new physics models. The simulation is highly configurable to ensure maximal flexibility in the face of the uncertain challenges approaching. The detector description itself, conditions of the detector, and many parameters used in the simulation can be modified at run time. The digitization is also made configurable to cope with uncertainty in machine performance, detector conditions, and cavern conditions. Three varieties of fast simulation have been made available to ease the difficulties caused by the time consumption of the full detector simulation. They each complement the full simulation.

Generation, simulation, and digitization tasks are running continually on the Grid. The validation program has produced a high quality simulation sample for the ATLAS experiment data.
Acknowledgements We are greatly indebted to all CERN’s departments and to the LHC project for their immense efforts not only in building the LHC, but also for their direct contributions to the construction and installation of the ATLAS detector and its infrastructure. We acknowledge equally warmly all our technical colleagues in the collaborating Institutions without whom the ATLAS detector could not have been built. Furthermore we are grateful to all the funding agencies which supported generously the construction and the commissioning of the ATLAS detector and also provided the computing infrastructure.

The ATLAS detector design and construction has taken about fifteen years, and our thoughts are with all our colleagues who sadly could not see its final realisation.

We acknowledge the support of ANPCyT, Argentina; Yerevan Physics Institute, Armenia; ARC and DEST, Australia; Bundesministerium für Wissenschaft und Forschung, Austria; National Academy of Sciences of Azerbaijan; State Committee on Science & Technologies of the Republic of Belarus; CNPq and FINEP, Brazil; NSERC, NRC, and CFI, Canada; CERN; CONICYT, Chile; NSFC, China; COLCIENCIAS, Colombia; Ministry of Education, Youth and Sports of the Czech Republic, Ministry of Industry and Trade of the Czech Republic, and Committee for Collaboration of the Czech Republic with CERN; Danish Natural Science Research Council and the Lundbeck Foundation; European Commission, through the ARTEMIS Research Training Network; IN2P3-CNRS and Dapnia-CEA, France; Georgian Academy of Sciences; BMDF, DFG, HGF and MPG, Germany; Ministry of Education and Religion, through the EPEAEK program PYTHAGORAS II and GSRT, Greece; IFI, MINEVA, GIF, DIP, and Benoyzio Center, Israel; INFN, Italy; MEXT, Japan; CNRST, Morocco; FOM and NWO, Netherlands; The Research Council of Norway; Ministry of Science and Higher Education, Poland; GRICES and FCT, Portugal; Ministry of Education and Research, Romania; Ministry of Education and Science of the Russian Federation, Russian Federal Agency of Science and Innovations, and Russian Federal Agency of Atomic Energy; JINR; Ministry of Science, Serbia; Department of International Science and Technology Cooperation, Ministry of Education of the Slovak Republic; Slovenian Research Agency, Ministry of Higher Education, Science and Technology, Slovenia; Ministerio de Educación y Ciencia, Spain; The Swedish Research Council, The Knut and Alice Wallenberg Foundation, Sweden; State Secretariat for Education and Science, Swiss National Science Foundation, and Cantons of Bern and Geneva, Switzerland; National Science Council, Taiwan; TAEK, Turkey; The Science and Technology Facilities Council and The Leverhulme Trust, UK; DOE and NSF, USA; Gordon and Betty Moore Foundation, USA.

The authors would like to thank the GEANT4 developers for their work and support.

Open Access This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

References

1. G. Aad et al., The ATLAS experiment at the CERN Large Hadron Collider. J. Instrum. 3, S08003 (2008)
2. L. Evans, The Large Hadron Collider. New J. Phys. 9, 335 (2007)
3. ATLAS Collaboration, ATLAS computing technical design report. ATLAS-TDR-017, CERN-LHCC-2005-022 (2005). See also http://atlas-computing.web.cern.ch/atlas-computing/packages/athenaCore/athaCore.php
4. S. Agostinelli et al., Geant4—a simulation toolkit. Nucl. Instrum. Methods Phys. Res. A 506, 250–303 (2003)
5. J. Allison et al., Geant4 developments and applications. IEEE Trans. Nucl. Sci. 53, 270–278 (2006)
6. M. Cattaneo et al., Status of the GAUDI event-processing framework, in Computing in High Energy and Nuclear Physics 2001 Conference (CHEP2001), IHEP, Beijing, China, September 3–7, 2001, ed. by H.S. Chen. (e-article, 2001)
7. G. Barrand et al., GAUDI—a software architecture and framework for building LHcb data processing applications, in Computing in High Energy and Nuclear Physics 2000 Conference (CHEP2000), ed. by M. Mazzucato, Pavia, Italy, February 7–11, 2000 (Padua, INFN, 2000)
8. L. Lonnblad, CLHEP: a project for designing a C++ class library for high-energy physics. Comput. Phys. Commun. 84, 307–316 (1994)
9. E. Obreshkov et al., Organization and management of ATLAS offline software releases. Nucl. Instrum. Methods A 584, 244–251 (2008)
10. CLOC—Count Lines of Code. http://cloc.sourceforge.net, 2006
11. D. Dullmann et al., POOL development status and plans, in Computing in High Energy and Nuclear Physics 2004 (CHEP2004), ed. by A. Aimer, J. Harvey, N. Knoors, Interlaken, Switzerland, September 27–October 1, 2004 (CERN, Geneva, 2004)
12. R. Chytracek et al., POOL development status and production experience. IEEE Trans. Nucl. Sci. 52, 2827–2831 (2005)
13. D. Dellmann, The LCG POOL project, general overview and project structure, in Computing in High Energy and Nuclear Physics 2003 (CHEP2003), La Jolla, CA, March 24–28, 2003. (CERN, Geneva, 2003). physics/0306129
14. M. Dobbs, J.B. Hansen, The HepMC C++ Monte Carlo event record for high energy physics. Comput. Phys. Commun. 134, 41–46 (2001)
15. M. Grothe et al., Architecture of the ATLAS high level trigger event selection software. Nucl. Instrum. Methods Phys. Res. A 518, 537–541 (2003). In Proceedings of the 2003 Conference for Computing in High Energy and Nuclear Physics, La Jolla, CA, USA, 24–28 March 2003
16. V. Boisvert et al., Final report of the ATLAS reconstruction task force. ATL-SOFT-2003-010 (2003)
17. I. Bird et al. (eds.), LHC computing Grid. Technical design report, CERN-LHCC-2005-024 (2005)
18. P. Nevski, Large scale data movement on the GRID, (2006)
19. M. Branco et al., Managing ATLAS data on a petabyte-scale with DQ2. J. Phys. Conf. Ser. 119, 062017 (2007). In Proceedings of Computing in High Energy and Nuclear Physics 2007 Conference (CHEP2007), Victoria, BC, Canada, September 2–7, 2007, ed. R. Sobie, R. Tafirjou, J. Thomson
20. T. Sjöstrand, S. Mrenna, P. Skands, PYTHIA 6.4 physics and manual. J. High Energy Phys. 05, 026 (2006). hep-ph/0603175
21. M. Smizanska, S.P. Baranov, J. Hirviance, E. Kneringer, Overview of Monte Carlo simulations for ATLAS B-physics in the period 1996–1999. ATL-PHYS-2000-025 (2000)
22. C. Anastopoulos et al., Physics analysis tools for beauty physics in ATLAS. J. Phys. Conf. Ser. 119, 032003 (2007). In Proceedings of Computing in High Energy and Nuclear Physics 2007 Conference (CHEP2007), Victoria, BC, Canada, September 2–7, 2007, ed. R. Sobie, R. Tafirjou, J. Thomson
23. G. Marchesini, B.R. Webber, G. Abbiendi, I.G. Knowles, M.H. Seymour, L. Stanco, HERWIG: a Monte Carlo event generator for simulating hadron emission reactions with interfering gluons. Version 5.1, April 1991. Comput. Phys. Commun. 67, 465 (1992)
24. G. Corcella et al., HERWIG 6: an event generator for hadron emission reactions with interfering gluons (including supersymmetric processes). J. High Energy Phys. 01, 010 (2001). hep-ph/0011363
25. G. Corcella et al., HERWIG 6.5 release note. CERN-TH/2002-270 (2005). hep-ph/0210213v2
26. T. Gleisberg et al., SHERPA 1. alpha, a proof-of-concept version. J. High Energy Phys. 02, 056 (2004). hep-ph/0311263
70. SQLite,  http://www.sqlite.org, 2008
71. R.F. van der Lans, The SQL Guide to SQLite, 1st edn. (Sams, 2004)
72. B. Di Girolamo et al., Beamline instrumentation in the 2004 combined ATLAS testbeam. ATL-TECH-PUB-2005-001 (2005)
73. C. Adorisio et al., Nucl. Instrum. Methods Phys. Res. A 593, 232–254 (2008)
74. C. Adorisio et al., System test of the ATLAS muon spectrometer in the H8 beam at the CERN SPS. Nucl. Instrum. Methods A 593, 232–254 (2008)
75. M. Hurwitz. Performance of ATLAS tile calorimeter production modules in calibration testbeams. Nucl. Instrum. Methods Phys. Res. A 572, 80–81 (2007)
76. P. Strizenez, A. Minaenko (ATLAS LAr Endcap group), J. Phys. Conf. Ser. 160, 012078 (2009)
77. H. Bartko, Performance of the combined ATLAS liquid argon end-cap calorimeters in beam tests at the CERN SPS. Diploma Thesis ATL-LARG-2004-007, 2004
78. ATLAS Pixel Collaboration, Pixel offline analysis for endcap A cosmic data. ATL-INDET-PUB-2008-003 (2008)
79. B. Demirkoz, Construction and performance of the ATLAS SCT barrels and cosmic tests. CERN-THESIS-2008-001 (2008)
80. Z. Broklova, Z. Dolezal, Simulations of ATLAS silicon strip detector modules in ATHENA framework. CERN-THESIS-2005-007 (2004)
81. T.H. Kittelmann, Slepton spin determination and simulation of the transition radiation tracker at the ATLAS experiment. PhD in physics, Niels Bohr Institute, University of Copenhagen, 2007
82. D. Salihagic, Comparison of beam test results of the combined ATLAS liquid argon endcap calorimeters with Geant3 and Geant4 simulations, in 11th International Conference on Calorimetry in High Energy Physics, ed. by C. Cecci, P. Cenci, P. Lubrano, M. Pepe-Altarelli (World Scientific, Singapore, 2004), p. 314
83. A. Dotti, A. Lupi, C. Roda, Results from ATLAS tile calorimeter: a comparison between data and Geant4 simulation. Nucl. Phys. B, Proc. Suppl. 150, 106–109 (2006)
84. N.C. Benekos, D. Rebuzzi, GEANT4 simulation of the ATLAS muon spectrometer, in 9th ICATPP Conference on Astroparticle, Particle, Space Physics, Detectors and Medical Physics Applications, ed. by M. Barone, E. Borch, A. Gaddi, C. Leroy, L. Price, Villa Erba, Como, Italy, 17–21 October, 2005 (World Scientific, Singapore, 2005)
85. K. Kordas, G. Parroux, S. Simion, GEANT4 for the ATLAS electromagnetic calorimeter, in 9th Conference on Calorimetry in High Energy Physics (CALOR 2000), ed. by B. Aubert, J. Colas, P. Nédélec, L. Poggioili, Annecy, France, 9–14 October 2000 (Lab. Ancyev Phys. Part., Ancyev-le-Vieux, 2000)
86. A. Kuryunin, D. Salihagic, P. Strizenez (ATLAS/LAr-HEC Collaboration), GEANT4 simulations of the ATLAS hadronic end cap calorimeter, in 9th Conference on Calorimetry in High Energy Physics (CALOR 2000), ed. by B. Aubert, J. Colas, P. Nédélec, L. Poggioili, Annecy, France, 9–14 October 2000 (Lab. Ancyev Phys. Part., Ancyev-le-Vieux, 2000)
87. P. Loch, R. Mazini, Comparison of experimental electron signals with GEANT3 and GEANT4 simulations for the ATLAS forward calorimeter prototype, in 9th Conference on Calorimetry in High Energy Physics (CALOR 2000), ed. by B. Aubert, J. Colas, P. Nédélec, L. Poggioili, Annecy, France, 9–14 October 2000 (Lab. Ancyev Phys. Part., Ancyev-le-Vieux, 2000)
88. A. Dell’Acqua et al., Development of the ATLAS simulation framework, in Computing in High Energy and Nuclear Physics 2001 Conference (CHEP2001), IHEP, Beijing, China, September 3–7, 2001, ed. by H.S. Chen, (e-article, 2001)
89. A. Dell’Acqua et al., ATLAS simulation readiness for first data at LHC. J. Phys. Conf. Ser. 119, 001001 (2007). In Proceedings of Computing in High Energy and Nuclear Physics 2007 Conference (CHEP2007), Victoria, BC, Canada, September 2–7, 2007, ed. R. Sobie, R. Tafrout, J. Thomson
90. C. Amsler et.al.(Particle Data Group), Phys. Lett. B 667, 1 (2008)
115. A. Salzburger, S. Todorova, M. Wolter, The ATLAS tracking geometry description. ATL-SOFT-PUB-2007-004 (2007)
116. A. Salzburger, The ATLAS track extrapolation package. ATL-SOFT-PUB-2007-05 (2007)
117. J.L. Henning, S.P.E.C. CPU2000, Measuring CPU performance in the new millennium. IEEE Comput. 28–35 (2000)
118. C.P.U. Normalization Criterion, http://www.egee.cesga.es/EGEE-SA1-SWE/accounting/guides/NormalizationCriterion, 2009
119. A. De Salvo, F. Brasolin, Benchmarking the ATLAS software through the kit validation engine, in Proceedings of Computing in High Energy and Nuclear Physics 2009 Conference (CHEP2009), Prague, Czech Republic, March 22–28, 2009
120. GCC, the GNU Compiler Collection. http://gcc.gnu.org, 2008
121. Scientific Linux CERN, 4 (SLC4). http://linux.web.cern.ch/linux/scientific4, 2008
122. Y. Perrin, F. Orellana, M. Roy, D. Feichtinger, The LCG Savannah software development portal. CERN Yellow Report CERN 2005-002 (2005), pp. 609–612
123. T. Carli, H. Hakobyan, A. Henriques-Correia, M. Simonyan, Measurement of pion and proton longitudinal shower profiles up to 20 nuclear interaction lengths with the ATLAS time calorimeter. ATL-TILECAL-PUB-2007-008 (2007)
124. P. Adragna et al., Testbeam studies of production modules of the ATLAS tile calorimeter. ATL-TILECAL-PUB-2009-002 (2009)
125. P. Speckmayer, Comparison of data with Monte Carlo simulations at the ATLAS barrel combined testbeam 2004. J. Phys. Conf. Ser. 160 (2009). In Proceedings of the XIII International Conference on Calorimetry in High Energy Physics (CALOR 08), Pavia, Italy, May 26–30, 2008, ed. M. Fraternali, G. Gaudio, M. Livan
126. A.E. Kiryunin et al., GEANT4 physics evaluation with testbeam data of the ATLAS hadronic end-cap calorimeter. J. Phys., Conf. Ser. 160 (2009). In Proceedings of the XIII International Conference on Calorimetry in High Energy Physics (CALOR 08), Pavia, Italy, May 26–30, 2008, ed. M. Fraternali, G. Gaudio, M. Livan
127. M. Cooke et al., In situ commissioning of the ATLAS electromagnetic calorimeter with cosmic muons. ATL-LARG-PUB-2007-013 (2007)
128. M. Campanella, A. Ferrari, P.R. Sala, S. Vanini, First calorimeter simulation with the FLUGG prototype. ATL-SOFT-99-004 (1999)