Unified System for Processing Real and Simulated Data in the ATLAS Experiment

Mikhail Borodin[1,2], Kaushik De[3], Jose Garcia Navarro[4], Dmitry Golubkov[1,5], Alexei Klimentov[1,6], Tadashi Maeno[6], David South[7] and Alexandre Vaniachine[8] on behalf of the ATLAS Collaboration

[1]*Big Data Laboratory, National Research Centre "Kurchatov Institute," Moscow, Russia*
[2]*National Research Nuclear University "MEPhI," Moscow, Russia*
[3]*University of Texas Arlington, TX, U.S.A.*
[4]*Instituto de Fisica Corpuscular, Universidad de Valencia, Spain*
[5]*Institute for High Energy Physics, Protvino, Russia*
[6]*Brookhaven National Laboratory, NY, U.S.A.*
[7]*Deutsches Elektronen-Synchrotron, Hamburg, Germany*
[8]*Argonne National Laboratory, IL, U.S.A.*
mikhail.borodin@cern.ch, kaushik@uta.edu, jose.enrique.garcia@cern.ch, dmitry.v.golubkov@cern.ch, aak@bnl.gov, maeno@bnl.gov, southd@mail.desy.de, vaniachine@anl.gov

Abstract
The physics goals of the next Large Hadron Collider run include high precision tests of the Standard Model and searches for new physics. These goals "at the Gates of Nature" require detailed comparison of data with computational models simulating the expected data behavior. To highlight the role which modeling and simulation plays in future scientific discovery, we report on use cases and experience with a unified system built to process both real and simulated data of growing volume and variety.

Keywords: Big Data, Grid-based Simulation and Computing, Parallel and Distributed Computing, Large Scale Scientific Instruments

1 Introduction

In 2015 the Large Hadron Collider will open "the Gates of Nature" by reaching instantaneous luminosities exceeding $2\cdot10^{34}$ cm^{-2}s^{-1} and center of mass energies above 13 TeV. The physics goals of the ATLAS experiment [1] include searches for physics beyond the Standard Model and high precision Higgs sector studies. These goals require detailed comparison of the expected physics and detector behavior with data. A rich set of computational models is employed to provide simulated data needed for this comparison with real data.

To address the corresponding Big Data processing challenge, the LHC experiments employ the computational infrastructure of the Worldwide LHC Computing Grid (WLCG) – the largest academic distributed computing environment. Thanks to the outstanding LHC performance, ATLAS manages over 160 petabytes of data on more than hundred computational sites. Following Big Data processing, more than 8000 scientists analyze LHC data in search of new discoveries. ATLAS leads the WLCG usage in the number of data processing jobs and the processed data volume.

Leveraging the underlying job management system PanDA developed by T. Maeno et al. [2], the production system orchestrates ATLAS data processing applications for efficient usage of more than hundred thousands of CPU-cores. In order to manage the diversity of LHC physics (exceeding 35K physics samples per year), the individual data processing tasks are organized into workflows. During data processing the system monitors site performance and supports dynamic sharing minimizing the workflow duration. In addition, the production system manages jobs and/or tasks failures enhancing the resilience.

In preparation for data taking, the ATLAS experiment is scaling up its Big Data capabilities by upgrading a multilevel production system that unifies processing of the real and simulated data. In this paper we describe representative data processing use cases handled by the production system highlighting the role which modeling and simulation plays in future scientific discovery.

2 Big Data Processing Use Cases

To process Big Data, the ATLAS experiment adopted the data transformation technique, where software applications transform the input datasets of one data type into the output datasets of another data types. In Big Data processing ATLAS deals with datasets, not individual files. Similarly a task (comprised of many jobs) has become a unit of the workflow in ATLAS Big Data processing. The successful validation of this technique was achieved through the exponential growth rate in the number of new data transformations and data types used for Big Data processing in the ATLAS experiment. Table 1 lists representative use cases described below.

Table 1: Data processing use cases

Use Case	Frequency	Workflow Length	Number of Tasks	Tasks Duration	Data Loss
Trigger Data	Weekly	Short	Several	Day	no
Real Data	Yearly	Medium	Hundreds	Weeks	no
Simulated Data	Quarterly	Long	Thousands	Months	yes

2.1 Trigger Data Processing

Figure 1, provided by the ATLAS Collaboration [3], shows that new physics discoveries and high precision studies of rare events require rejection of “known” events by more than ten orders of magnitude. The multi-tier trigger system reduces Big Data volume to a manageable level. In 2015, the ATLAS experiment will have a two-tier trigger system:

1. The hardware-based Level 1 trigger.

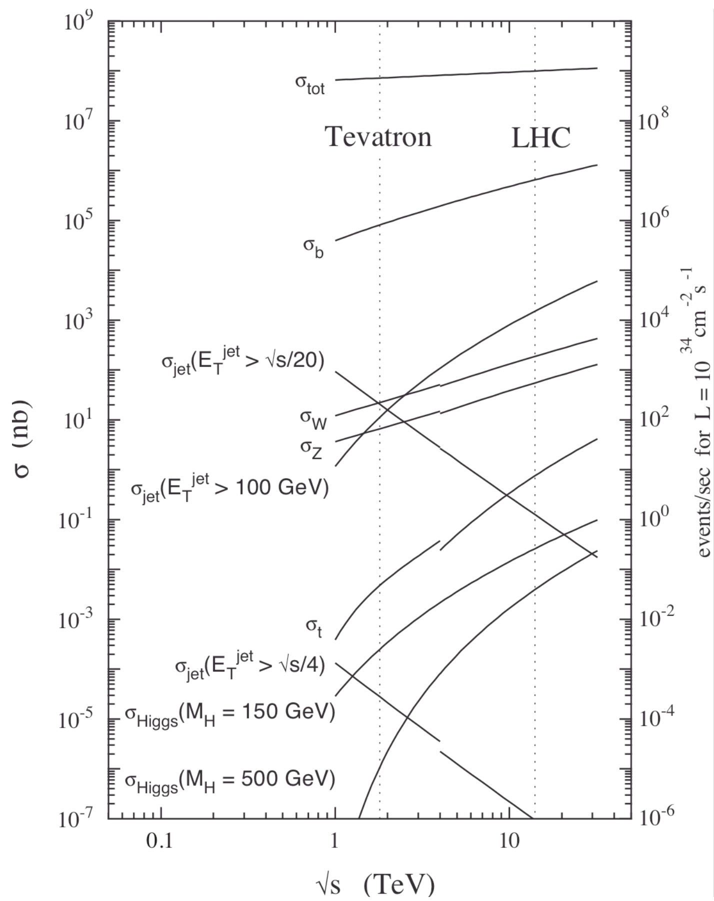

Figure 1: Comparison of cross-sections and rates for "known" and "rare" events in proton-(anti)proton

2. The software-based High-Level Trigger designed by Bartoldus et al. [4].

The Trigger Data Processing happens one step before the raw data recording. Thus, any inefficiencies or mistakes may lead to unrecoverable loss of real data. To eliminate such losses, the dedicated raw-to-raw data processing technique is employed to validate trigger software and other critical trigger changes during data taking. This technique is the main tool for commissioning the trigger for data taking.

2.2 Real Data Processing

The "raw" data from the ATLAS detector (Figure 2) are processed to produce the reconstructed data for physics analysis. During reconstruction ATLAS applications process raw detector data to identify and reconstruct physics objects such as leptons. Figure 3 shows data processing flow used in reconstruction. The distributed multi-tier data processing architecture handles petascale data flow analyzed by Vaniachine [5]. Since the detector data are comprised of independent events, massively parallel applications process one event at a time. Events taken during few minutes are collected in one file. Thousands of files with events that are close in time are collected in one dataset.

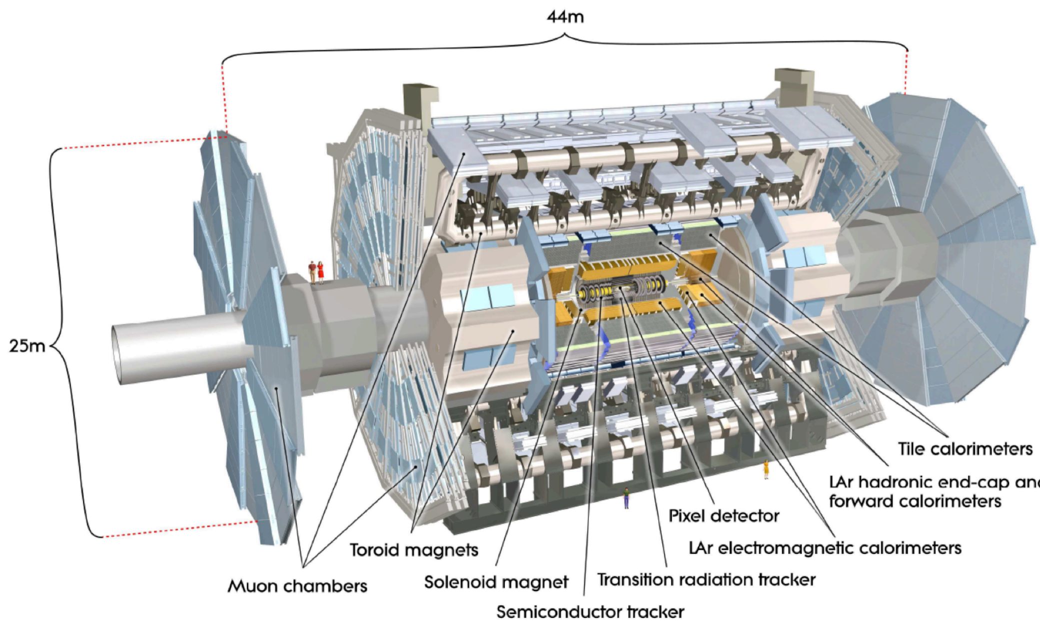

Figure 3: The real data for processing come from the ATLAS detector

Validating these techniques, the ATLAS collaboration completed four petascale data processing

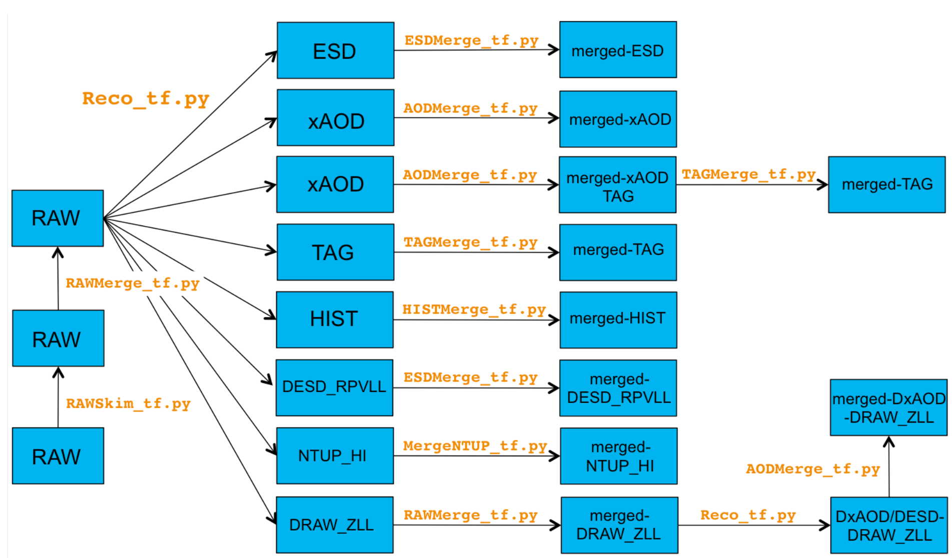

Figure 2: Data processing workflow for the real data. Arrows are labeled with data transformation applications; boxes are labeled with various data types produced.

Table 2: Processing campaigns for real data

Campaign year	Input Data Volume (PB)	CPU Time Used for Reconstruction ($10^6 h$)
2010	1	2.6
2011	1	3.1
2012	2	14.6
2013	2	4.4

campaigns on the Grid, with up to 2 PB of real data being processed every year. Table 2 lists parameters for the ATLAS yearly data processing campaigns. (In 2013 reprocessing, 2.2 PB of input data were used for selecting about 15% of all events for reconstruction, thus reducing CPU resources vs. the 2012 reprocessing.)

2.3 Simulated Data Processing

The data processing campaigns for the simulated data correspond to the data taking periods of the real data. As a result, the computational resources required to process the simulated data dominate the overall resource usage (Figure 4). The LHC data taking periods of the same conditions are characterized by the same center-of-mass energy, instantaneous luminosity, detector configuration,

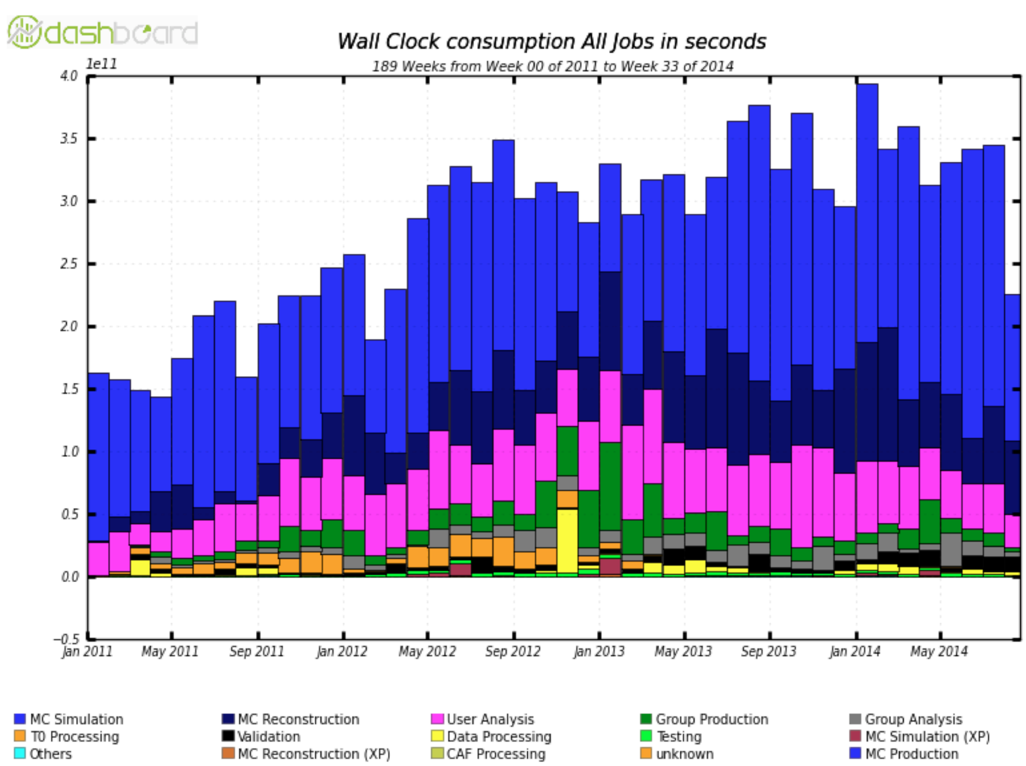

Figure 4: Processing of the simulated data (MC) dominates the consumption of computational resources

Table 3: Data processing campaigns for simulated data

Campaign Label	Data Taking Period for Real Data	Configuration	Full Simulation (10^9 events)	Fast Simulation (10^9 events)	Number of Sub-campaigns
mc11	2011	7 TeV	3.64	3.27	4
mc12	2012	8 TeV	6.37	6.43	3
mc14	2012 & 2015	8 & 13 TeV	0.85		2

etc. Table 3 lists the major data processing campaigns for the simulated data, while Figure 5 shows the rate of the simulated events produced.

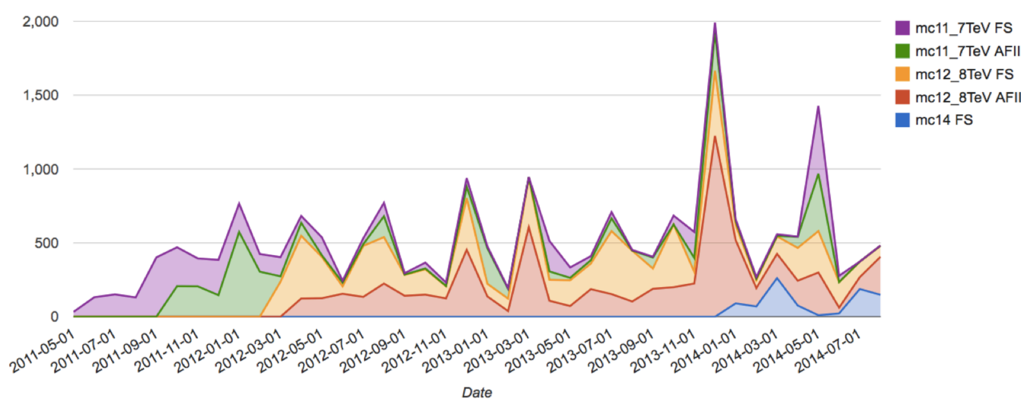

Figure 5: Monthly rate of the simulated events produced (10^6). The Full Simulations are labeled FS, the Fast Simulations are labeled AFII. Note, that the production system process several concurrent campaigns.

The LHC instantaneous luminosities result in the presence of a large number of simultaneous collisions in the same event, overlapping the hard scattering event of interest. The presence of the minimum bias events is usually called "pileup". To provide realistic simulation of these conditions, the data processing workflow for simulated data is composed of many steps (Figure 6): generate or configure hard-processes, hadronize signal and minimum-bias (pileup) events, simulate energy deposition in the ATLAS detector, digitize electronics response, simulate triggers, reconstruct data,

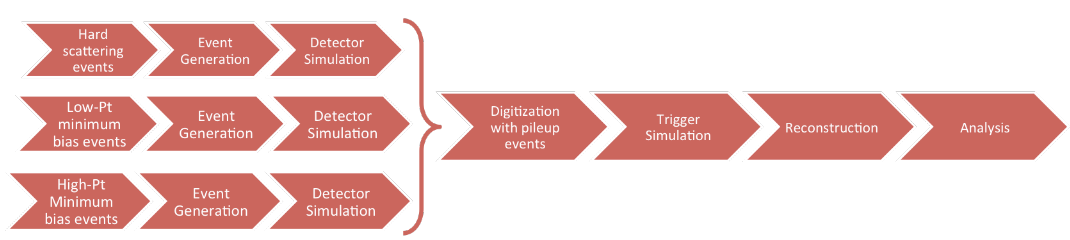

Figure 6: ATLAS simulations workflow is composed of many steps. Several initial steps are repeated for the hard-scattering events and the minimum bias events (simulated in two complementary samples).

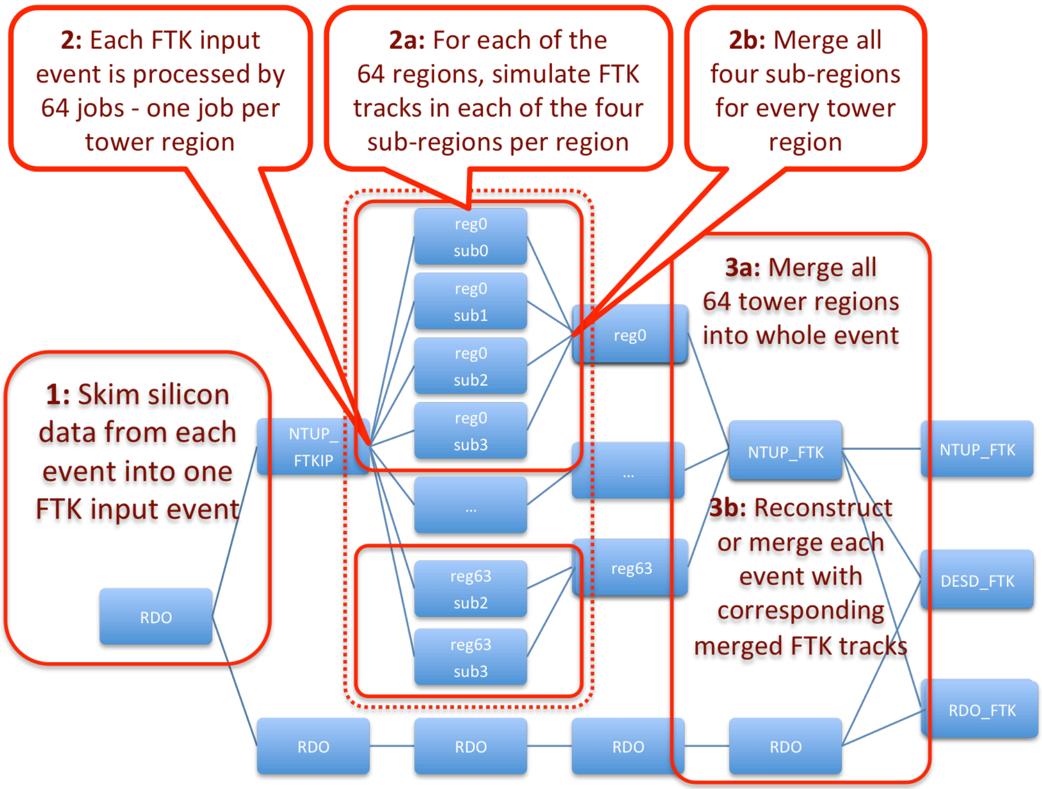

Figure 7: The simulation of the FTK hardware splits every event into 256 sub-events

transform the reconstructed data into data types for physics analysis, etc. The intermediate outputs are merged and/or filtered as necessary to optimize the chain.

An example of a more complex workflow used to simulate the ATLAS trigger using dedicated hardware for fast tracking (FTK) designed by Annovi et al. [5] is shown on Figure 7, where to keep the computational resources for the FTK simulation below practical limits, Adelman et al. [6] split every event into 256 η-ϕ sub-regions of each event. In the three-step workflow, each event is processed by 64 jobs; each job simulates tracks in four FTK sub-regions one after another. The sub-region merging is done in two steps: producing whole regions, then whole events in the NTUP_FTK files. The final step uses FTK tracks in trigger simulations producing the reconstructed data in DESD_FTK files or adds FTK tracks to the simulated events in the RDO_FTK files.

Validating our Big Data processing techniques, four different sub-campaigns of the mc11 campaign implemented the pileup conditions, detector conditions and geometry increasingly closer to those in real data. During the mc12 campaign, the most of the events was simulated in the sub-campaign mc12b. Later, the mc12c sub-campaign implemented an improved detector geometry description. Figure 8 shows the variety of the simulated of event samples sizes for more than 22000 different datasets produced during mc12 campaign. The goal of the mc14 campaign was to prepare for the 2015 data taking. The 8 TeV events were processed with improved and updated simulation, digitization and reconstruction software while using the same conditions as in the mc12 campaign. The 13 TeV campaign with the center of mass energy expected for the 2015 data taking and estimated pileup and detector conditions. In the mc14 campaign the new ATLAS Integrated Simulations

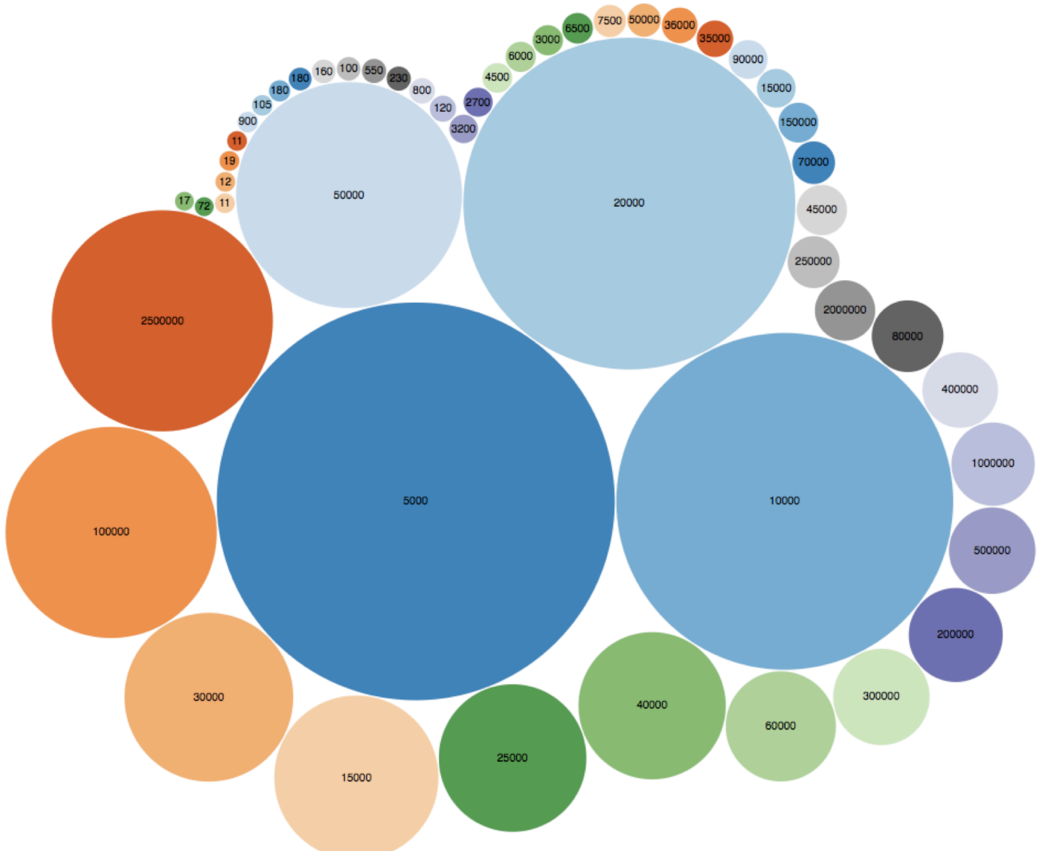

Figure 8: Variety of event samples sizes for more than 22000 different datasets produced during mc12 campaign

Framework was used, with multicore processing becomes the default for major simulated data processing steps: simulation, digitization and reconstruction.

3 Multilayer Data Processing System

The LHC shutdown provided an opportunity for upgrading the production system, making implementations more scalable, whilst retaining most valued core capabilities. To assure scalability, the production system was upgraded with extra layers. Avoiding inherent fragility of the monolithic systems, we separated the core concerns: the system logic layer is separated from the presentation layer, providing a familiar but improved interface for task requests.

The upgraded data processing system generates actual workflow tasks and their jobs are executed across more than a hundred distributed computing sites via PanDA – the ATLAS workload management system. Figure 9 shows that on top, the Task Request interface encapsulates the presentation layer for users, while the lower Task Definition layer implements the core data processing logic that empowers production managers with templated workflow definitions through the Database

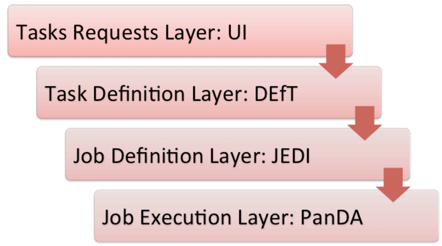

Figure 9: Multilayer architecture of the ATLAS production system for processing real and simulated data

Engine for Tasks (DEfT). At the layer below, the Job Execution and Definition Interface (JEDI) is integrated with PanDA to provide dynamic job definition tailored to the sites capabilities.

In the WLCG distributed computing environment, PanDA provides transparency of data and processing. As a result, the production system sees a unified computing facility that is used to run all data processing for the experiment, even though the sites are physically located all over the world. The production system supports a diverse range of workflows handling centrally ATLAS petascale data processing of the real and simulated data, including the mixture of both.

4 Conclusion

During LHC data taking, the ATLAS production system unified a diverse range of workflows and special use cases including processing of the real and the simulated data at large scales. The ATLAS production system fully satisfies the Big Data processing requirements of the ATLAS experiment through the unified approach for real data processing and simulations as well as the mixture of both. Golubkov et al. [8] reported that this technique enabled to address a much wider range of physics analyses, with a higher level of precision, surpassing the most optimistic expectations. In addition, detailed physics studies established that the simulated data are of unprecedented quality compared to previous generations of experiments, describing the real data behavior quite well in most analyses. The unified capabilities for real and simulated data processing significantly enhanced ATLAS physics output, and they motivated production of higher than foreseen simulated data volumes.

In preparations for the LHC data taking, De et al. [9] designed the production system upgrade to further improve the performance and accommodate a growing number of new requirements and use cases. Currently, the upgraded production system is undergoing commissioning, addressing a growing number of tasks for processing simulated and real data for future scientific discoveries.

Acknowledgements

We wish to thank all our colleagues who contributed to ATLAS Big Data processing activities. This work was funded in part by the U. S. Department of Energy, Office of Science, High Energy Physics and ASCR Contract No. DE-AC02-98CH10886 and Contract No. DE-AC02-06CH11357. NRC KI team work was funded by the Russian Ministry of Science and Education under Contract №14.Z50.31.0024.

References

[1] The ATLAS Collaboration. The ATLAS experiment at the CERN Large Hadron Collider. *J. Inst.* 2008; 3:S08003.

[2] Maeno T et al. Evolution of the ATLAS PanDA production and distributed analysis system *J. Phys.: Conf. Ser.* 2012; 396:032071.

[3] The ATLAS Collaboration. *ATLAS high-level trigger, data-acquisition and controls: Technical Design Report*. LHCC Report CERN-LHCC-2003-022, Geneva: CERN; 2003.

[4] Bartoldus R et al. *Technical Design Report for the Phase-I Upgrade of the ATLAS TDAQ System,* LHCC Report CERN-LHCC-2013-018, ATLAS-TDR-023, Geneva: CERN; 2013.

[5] Vaniachine AV for the ATLAS Collaboration. ATLAS detector data processing on the Grid. In: *IEEE Nuclear Science Symposium and Medical Imaging Conference;* 2011, p. 104-107.

[6] Annovi A et al. *ATLAS Fast TracKer (FTK) Technical Design Report,* LHCC Report CERN-LHCC-2013-007, ATLAS-TDR-021, Geneva: CERN; 2013.

[7] Adelman J et al. *ATLAS FTK challenge: simulation of a billion-fold hardware parallelism,* ATLAS Note ATL-DAQ-PROC-2014-030, Geneva: CERN; 2014.

[8] Golubkov D et al. ATLAS Grid Data Processing: system evolution and scalability. *J. Phys.: Conf. Ser.* 2012; 396:032049.

[9] De K et al. Task management in the new ATLAS production system. *J. Phys.: Conf. Ser.* 2014; 513:032078.